

**UNIVERSIDADE DE LISBOA**

**INSTITUTO DE EDUCAÇÃO**



**A APRENDIZAGEM DA COVARIÇÃO ESTATÍSTICA COM  
RECURSO AO *TINKERPLOTS*: UM ESTUDO COM ALUNOS  
DO 10.º ANO**

**Patrícia Alexandra Robalo Antunes**

**MESTRADO EM EDUCAÇÃO**

**Área de Especialidade em Didática da Matemática**

**Trabalho de Projeto orientado pela  
Prof.<sup>a</sup> Doutora Ana Cláudia Correia Batalha Henriques**

**2015**



**UNIVERSIDADE DE LISBOA**  
**INSTITUTO DE EDUCAÇÃO**



**A APRENDIZAGEM DA COVARIÇÃO ESTATÍSTICA COM  
RECURSO AO *TINKERPLOTS*: UM ESTUDO COM ALUNOS  
DO 10.º ANO**

**Patrícia Alexandra Robalo Antunes**

**MESTRADO EM EDUCAÇÃO**

**Área de Especialidade em Didática da Matemática**

**Trabalho de Projeto orientado pela  
Prof.<sup>a</sup> Doutora Ana Cláudia Correia Batalha Henriques**

**2015**

Trabalho realizado no âmbito do Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* financiado pela FCT – Fundação para a Ciência e Tecnologia (Contrato PTDC/CPE-CED/117933/2010).

## RESUMO

Este trabalho de projeto incide sobre a Unidade de Ensino - Distribuições Bidimensionais, no âmbito da qual realizei um estudo visando compreender as aprendizagens e dificuldades dos alunos de uma turma profissional do 10.º ano, dos cursos profissionais de Técnico de Auxiliar de Saúde e de Técnico de Apoio à Gestão Desportiva, no que respeita à covariação estatística, analisando o seu raciocínio estatístico covariacional quando utilizam o *software TinkerPlots* na resolução de tarefas envolvendo este conceito. O estudo, de natureza qualitativa e interpretativa, teve por base dados recolhidos através de observação participante, com registo áudio e vídeo das aulas e recolha documental (notas de campo, resoluções escritas dos alunos das tarefas propostas e gravações do trabalho realizado no *TinkerPlots*).

Os resultados do estudo evidenciam que, de modo geral, os alunos usam diferentes representações gráficas, construídas no *TinkerPlots* para explorarem a existência de relações entre variáveis e suportar as suas conjecturas e revelam uma evolução significativa na utilização do diagrama de dispersão, reconhecendo a sua adequabilidade para estabelecer o grau dessa associação. O trabalho desenvolvido nas tarefas, com o *software*, permitiu que os alunos evoluíssem na sua conceção local dos dados, passando a focar-se em tendências globais para avaliar e suportar as suas inferências. A utilização de contextos reais ou familiares aos alunos, nas tarefas, mostrou-se favorável para fazer emergir interesse e motivação por parte dos alunos e a procura de relações nos dados, proporcionando a sua significação e a busca de possíveis causas de variação nos padrões identificados. Algumas dificuldades ainda persistem no final, no que respeita à justificação da associação entre variáveis recorrendo a dados isolados e não admitindo exceções, revelando uma conceção determinista. Além disso, fazem inferências causais baseadas nos dados, interpretando-os com base no seu conhecimento do contexto.

*Palavras-chave:* Raciocínio estatístico; Representações gráficas; Covariação; Ensino básico; *Tinkerplots*.



## **ABSTRACT**

This project work focuses the content “Bidimensional Distributions”, according to which I developed a study to understand the learning process and the difficulties, concerning the statistical covariation, of the students of a class of the year 10, of the professional courses of health technician and sports management technician, by analyzing their covariational statistical reasoning, when they use the *software TinkerPlots* to carry out tasks involving this concept. This study, with an interpretative and qualitative nature, was based on data collected through a participating observation, the recording of the lessons and the gathering of documents (notes, the written answers of the students to the proposed tasks and the recordings of the work elaborated with the *TinkerPlots*).

The results of this study show that the students generally use different graph representations, build with the *TinkerPlots*, to explore the existence of relationships among variables and to support their conjectures, and they also show a significative evolution when using the dispersion diagram, recognizing its suitability to establish the degree of that association the work developed by the students, with the software *TinkerPlots*, enabled them to evolve their local conception of the data, making them focus on global tendencies to evaluate and support their conclusions. The use of real or familiar contexts by the students when carrying out the tasks, helped to stimulate the students’ interest and motivation and the search for relationships among the data, allowing, therefore, their meaning and the discovery of possible causes for the variation in the identified patterns. Some difficulties still persist at the end, concerning the justification of the association among variables by resorting to single data and by not admitting exceptions, which reveals a deterministic conception. Besides that, the students come to causal conclusions based on the data and interpret them taking into account their knowledge of the context.

*Keywords:* Statistical reasoning; Graph representations; Covariation; Basic literacy, *Tinkerplots*.



## **AGRADECIMENTOS**

A concretização deste trabalho só foi possível devido à colaboração, direta e indireta, de algumas pessoas. A todas deixo aqui expresso o meu reconhecimento.

À minha orientadora, Professora Doutora Ana Henriques, com quem muito aprendi, e cuja orientação contribuiu uma mais-valia para a conclusão deste trabalho, por quem nutro um enorme respeito. Pela forma interessada e disponível com que sempre acompanhou, pela compreensão e pertinência dos seus comentários e sugestões, bem como pelos estímulos constantes.

Ao Jorge, que me acompanhou ao longo de todo o percurso e com quem pude partilhar todas as dúvidas e ansiedades. Obrigada por cada minuto que dedicaste a ouvir-me, a apoiar-me e por não me deixar desistir.

Às minhas colegas deste mestrado, Irina e Raquel, pelos momentos inesquecíveis e inigualáveis de trabalho, convívio, apoio e amizade.

À Direção e aos meus colegas de departamento, da escola onde se realizou este trabalho, pelo encorajamento e apoio que me deram na concretização deste trabalho. À Rita, professora da turma do 10.º ano com quem tive o privilégio de trabalhar, por ter aceite o meu desafio.

A todos os alunos que participaram neste trabalho pela sua disponibilidade, entusiasmo e colaboração, sem os quais não teria sido possível a concretização deste trabalho.

A todos os amigos, por saberem ouvir e pela força que me deram ao longo deste período, que contribuíram para a concretização deste trabalho. A todos os colegas que, em seminários ou conversas informais, ajudaram-me a refletir sobre este estudo.

Ao meu avó Domingos, que apesar de já não estar entre os vivos, é das pessoas mais importantes da minha vida e que fez parte do caminho que percorri até aqui.

Por último, um agradecimento à minha família, mãe, avó Isabel, Luciana e João Pedro, por todo o incondicional apoio ao longo desta jornada. Pelas incansáveis e preciosas ajudas, incentivo e paciência que foram fundamentais na concretização de mais uma etapa importante da minha vida.



# Índice

## CAPÍTULO 1

INTRODUÇÃO.....	1
1.1. Motivação e pertinência do estudo .....	1
1.2. Objetivo e questões do estudo.....	5
1.3. Organização do estudo.....	5

## CAPÍTULO 2

FUNDAMENTAÇÃO E ENQUADRAMENTO CURRICULAR.....	9
2.1. O ensino e a aprendizagem da Estatística .....	9
2.2. A Tecnologia no ensino da Estatística .....	16
2.3. O raciocínio estatístico covariacional.....	21

## CAPÍTULO 3

METODOLOGIA DO ESTUDO .....	33
3.1. Opções metodológicas.....	33
3.2. Estudo exploratório .....	35
3.3. Caracterização do contexto escolar e dos participantes do estudo.....	41
3.4. Recolha e Análise de dados .....	43

## CAPÍTULO 4

A UNIDADE DE ENSINO.....	47
4.1. Aspectos gerais e planificação da unidade de ensino .....	47
4.2. Tarefas e recursos.....	53
4.3. A avaliação das aprendizagens .....	57

## CAPÍTULO 5

RACIOCÍNIO COVARIACIONAL DOS ALUNOS .....	59
5.1. Tarefa – “Desenvolvimento da criança” .....	59
5.2. Tarefa – “Indicadores Demográficos” .....	73
5.3. Tarefa – “Bateria de Testes de Fitnessgram” .....	87
5.4. Prova de Avaliação Final .....	100

## CAPÍTULO 6

CONCLUSÕES E REFLEXÃO FINAL .....	105
6.1. Síntese do estudo .....	105
6.2. Principais conclusões.....	106
6.3. Reflexão final.....	110
REFERÊNCIAS BIBLIOGRÁFICAS.....	113
ANEXOS .....	127

## ÍNDICE DE FIGURAS

Figura 1. Modelo de literacia estatística de Gal (2002).....	10
Figura 2. Formas de representar a covariação estatística e capacidades para as traduzir (Adaptado de Moritz, 2004, p.230). .....	27
Figura 3. Modelo de raciocínio estatístico dos alunos em ambientes gráficos com tecnologia (Adaptado de Fitzallen, 2006, p. 206). .....	31
Figura 4. Momentos de recolha de dados ao longo da unidade de ensino.....	45
Figura 5.1.1. Representações gráficas de vários grupos (G2 – A e G3 – B, C e D).....	61
Figura 5.1.2. Gráfico de covariação para o peso e perna [TG6]. .....	62
Figura 5.1.3. Gráficos de covariação entre: A- “altura aos 2 anos” e “altura aos 18 anos” [TG8] e B – “peso” e “força aos 9 anos” [TG10].....	63
Figura 5.1.4. Gráfico de dispersão representando o peso e largura da perna [TG3].....	64
Figura 5.1.5. Gráfico de dispersão para o peso, largura da perna e estrutura para crianças com 18 anos [TG2]. .....	64
Figura 5.1.6. Representação gráfica para a largura da perna e o peso com 18 anos [TG10]. .....	65
Figura 5.1.7. Gráfico de covariação para o peso e altura [TG2]. .....	65
Figura 5.1.8. Gráfico de covariação entre peso e altura aos 9 anos [TG10]. .....	67
Figura 5.2.1. Comparação entre o número de mortes e o número de nascimento, através do boxplot [G3] (A e C). Diagrama de extremos e quartis, nos diferentes dados de distribuição na variável “Taxa bruta de mortalidade” [G8] (B e D). .....	74
Figura 5.2.2. Diagrama de dispersão com as variáveis “Taxa bruta de natalidade” com a “Taxa bruta de mortalidade” [G2]. .....	77
Figura 5.2.3. Gráfico de pontos com duas escalas (contínua e intervalar) [G8].	77
Figura 5.2.4. Resposta à questão 2 [G8].....	79
Figura 5.2.5. Diagrama de dispersão representando as variáveis: “Idade média da mulher ao 1.º casamento” e “Idade média do Homem ao 1.º casamento” [G3]. .....	79
Figura 5.2.6. Diagrama de dispersão representando as variáveis: A - Idades médias ao 1.º casamento, relativamente ao homem e mulher [G10] e B – Relação de masculinidade e taxa de nupcialidade [G8]. .....	81

Figura 5.2.7.	Resposta à questão 3.3. [G6].	82
Figura 5.2.8.	Diagrama de dispersão representando as variáveis “índice sintético de fecundidade” com: “Taxa bruta de mortalidade” (A); “Taxa de mortalidade infantil” (B) e “Taxa bruta de natalidade” (G8).	83
Figura 5.2.9.	Resposta da parte III (G6).	85
Figura 5.3.1.	Representação gráfica de “Altura” e “Velocidade” [G2].	88
Figura 5.3.2.	Gráficos de dispersão representando relações entre “Peso” com: “IMC” (A) e “Velocidade” (B) [G3].	89
Figura 5.3.3.	Representação gráfica com duas ou três variáveis (G1 – A e B; G3 – C e D).	90
Figura 5.3.4.	Gráficos de dispersão representando relações entre diversas variáveis [G10].	91
Figura 5.3.5.	Representação gráfica representando relações entre diversas variáveis [G10].	92
Figura 5.3.6.	Representação gráfica representando relações entre variáveis (A – G8; B – G10 e C – G3).	93
Figura 5.3.7.	Diagramas de dispersão representando relações do “IMC” com a “Altura” (A- G10) e “Peso” (B – G8).	95
Figura 5.3.8.	Representação gráfica representando relações entre diversas variáveis [DC].	96
Figura 5.4.1.	Resposta à questão 3.2 da prova de avaliação [aluna do G5].	101
Figura 5.4.2.	Resposta à questão 3.3. da prova de avaliação [aluna do G10].	102
Figura 5.4.3.	Resposta à questão 3.3. da prova de avaliação [aluno do G2].	103

## ÍNDICE DE QUADROS

Quadro 1. Síntese do modelo de raciocínio estatístico desenvolvido por Garfield (2002):.....	23
Quadro 2. Formas de representar a fonte de dados.....	45
Quadro 3. Objetivos da unidade de ensino – distribuições bidimensionais (ME, 2001a; Me, 200b; ME, 2005).....	47
Quadro 4. Planificação da unidade de ensino.....	48

## ÍNDICE DE ANEXOS

Anexo 1. Pedido de Autorização aos alunos participantes no Estudo-Piloto.....	128
Anexo 2. Pedido de autorização à Direção da Escola para Estudo-Piloto .....	129
Anexo 3. Guião da entrevista/tarefa.....	130
Anexo 4. A tarefa da entrevista (suporte em ficheiro <i>TinkerPlots</i> ) .....	131
Anexo 5. Pedido de autorização ao Diretor da Escola.....	132
Anexo 6. Pedido de Autorização aos Encarregados de Educação .....	133
Anexo 7. Tarefa: “Nenana Ice Classic” .....	134
Anexo 8. Tarefa: “Uma experiência com peixes” .....	136
Anexo 9. Tarefa: “Desenvolvimento da criança” .....	138
Anexo 10. Tarefa: “Indicadores demográficos” .....	140
Anexo 11. Tarefa: “Bateria de testes Fitnessgram” .....	142
Anexo 12. Prova de Avaliação Final.....	145

# CAPÍTULO 1

## INTRODUÇÃO

Neste capítulo apresento o problema e as questões que me proponho estudar, bem como as razões que me motivaram à realização deste estudo sobre a aprendizagem do conceito de covariação estatística com recurso ao *Tinkerplots* por alunos do 10.º ano, no quadro de uma experiência de ensino.

### 1.1. Motivação e pertinência do estudo

A estatística, hoje em dia, desempenha um papel de muito importante no nosso mundo. Todos os dias, nos jornais, nas revistas, na rádio e na televisão, somos confrontados com tabelas, gráficos, sondagens que contêm informação estatística sobre várias atividades e fenómenos. A literacia estatística (Branco & Martins, 2002; Shaughnessy, 2007; Watson, 2006), sendo reconhecida como fundamental para a compreensão dessa informação e para a tomada de decisões com base nela, deverá ser desenvolvida ao longo do percurso escolar dos alunos.

A Estatística é, por isso, uma área de estudo onde tem ocorrido um desenvolvimento significativo, não só no que respeita às suas aplicações na sociedade, fortemente apoiadas pela tecnologia, mas também ao nível dos conteúdos e abordagens curriculares (GAISE, 2007; NCTM, 2008). Por exemplo, a análise exploratória de dados e a realização de investigações estatísticas são aspetos salientes nas orientações curriculares. De acordo com o NCTM (2008), os alunos devem trabalhar com os dados “para compreender as ideias fundamentais da estatística” (p. 52), envolvendo-os em novas ideias e procedimentos e a partir

daí “fazer comparações estatisticamente válidas” (p. 54) e adquirir a “capacidade de avaliar a validade de argumentos baseados em dados” (p. 55).

Abrantes, Serrazina e Oliveira (1999) defendem que o ensino da Estatística deve ser desenvolvido experimentalmente, apelando sempre às capacidades de raciocínio e comunicação. Assumem que a recolha e a análise de dados são o ponto fulcral do pensamento estatístico, pois com a recolha de dados promove a aprendizagem pela experiência e relaciona o processo de aprendizagem com a realidade e a análise dos dados permite responder às questões ou hipóteses conjecturadas.

Apesar disso, existem evidências na literatura de que os alunos, frequentemente, têm dificuldades com o raciocínio estatístico e o seu uso na tomada de decisões sobre questões do dia-a-dia (Ben-zvi & Garfield, 2004). Ao longo do meu percurso profissional também tenho verificado que, para além das dificuldades de aprendizagem específicas de cada aluno, existem temáticas onde a maioria dos alunos revelam grandes dificuldades. Esta ideia surge, não só da minha experiência de lecionação mas, também, do contacto e do trabalho colaborativo que tenho mantido com outros colegas.

Está ainda por compreender a melhor forma de ajudar os alunos a ultrapassar os seus erros e as dificuldades conceptuais e procedimentais, em diferentes níveis etários, particularmente no que respeita ao raciocínio sobre a covariação, que é um tópico que se tem revelado muito difícil para os alunos (Fitzallen, 2012). Por exemplo, ao analisarem as distribuições bivariadas representadas em gráficos de dispersão, os alunos revelam dificuldades na sua interpretação e no estabelecimento de covariação uma vez que se focam em dados individuais em vez de observar a tendência global dos dados ou no conhecimento prévio que têm sobre o contexto (Estepa & Batanero, 1996; Fitzallen, 2012).

De acordo com Carvalho e César (2000), a forma como os alunos trabalham as tarefas e as instruções na aula de Estatística, com os colegas e os professores, influenciam na aprendizagem e no seu desempenho. É importante, por isso, que os professores criem situações didáticas com ambientes de “interações sociais ricas e frutuosas” (Carvalho & César, 2000, p.207), onde os alunos terão possibilidade de explorar questões e ideias que envolvem o raciocínio estatístico, bem como ultrapassar possíveis dificuldades, como as identificadas na literatura. O processo

de aprendizagem dos alunos poderá ser influenciado pela introdução de propostas de trabalho que dão ênfase à utilização das tecnologias. A tecnologia pode ser incorporada no processo de ensino e aprendizagem para aumentar a experiência de aprendizagem e melhorar as suas compreensões de importantes conceitos.

Nos Princípios e Normas para a Matemática Escolar (NCTM, 2008) está bem patente a importância atribuída à tecnologia. Salienta-se nesse documento que as relações “podem ser eficazmente demonstradas, através da utilização de programas informáticos, com os quais os alunos poderão controlar um valor e observar os efeitos dessa alteração” (NCTM, 2008, p. 295) e “refletir sobre a forma como diferentes gráficos representam as características mais importantes dos conjuntos de dados” (NCTM, 2008, p. 296). Além disso, e também segundo Vieira (2008), a aplicação de ferramentas tecnológicas em sala de aula, proporciona aos alunos efetuar interpretações dos conceitos envolvidos, sem ter que se preocupar com os cálculos ou construções com pouco significado para eles. Assim, a aplicação de *software* e ferramentas tecnológicas permitirá que os alunos atribuam significado à Estatística facilitando a introdução de novas representações. Deste modo, permitem mudar a forma de trabalhar a Estatística, em sala de aula, que ainda é focada no cálculo e nos aspetos procedimentais em vez das investigações e da análise de dados e sua interpretação (Russell & Friel, 1989; Shaughnessy, 2007).

Durante a minha experiência profissional também tenho verificado ser vantajoso o uso das tecnologias disponíveis na escola, pois permite envolver os alunos em atividades que proporcionam uma melhor aprendizagem, visível através de um domínio mais aprofundado de conceitos e métodos adequados de análise de dados estatísticos. Estes aspetos concordam, com o que se configura no Programa de Matemática A do 10.º ano (ME, 2001a) e no Programa de Componente de Formação Científica da disciplina de Matemática, de cursos profissionais de nível Secundário (Martins, et al., 2005).

Na prática de sala de aula, a utilização da tecnologia para a realização de tarefas que poderiam ser difíceis de implementar de outro modo (por exemplo, tarefas envolvendo grande quantidade de dados reais e situações da vida real), permite que o aluno seja liberto de atividades rotineiras e repetitivas, dando-lhe a hipótese de explorar uma diversidade de conjeturas por ele formuladas. A utilização do

*Tinkerplots*, em particular, facilita o ensino e aprendizagem de conceitos estatísticos (Fitzallen, 2012; Rubin, Hammerman & Konold, 2006), bem como na compreensão de relações entre variáveis.

Desde sempre tento desempenhar a minha função docente da forma que acredito ser melhor para os alunos, procurando também refletir sobre a minha prática. Foi este contexto que suscitou o meu interesse em compreender como é que os alunos utilizam as tecnologias, dando especial atenção às suas potencialidades para o ensino e aprendizagem da Estatística. Além disso, atendendo ao exposto, torna-se pertinente analisar como os alunos tomam decisões acerca de associação estatística com base na análise de dados e no conhecimento do contexto do problema, nomeadamente em relação à covariação. Neste trabalho irei procurar compreender, também, as dificuldades que se colocam aos alunos durante a interação com o *software* e refletir sobre as suas potencialidades quando é usado no ensino e na aprendizagem do tópico das distribuições bidimensionais (abordagem gráfica e intuitiva) do Programa de Componente de Formação Científica da disciplina de Matemática, de cursos profissionais de nível Secundário (Martins, et al., 2005), em particular do conceito da covariação estatística. Espero que os resultados deste estudo possam contribuir para a melhoria do ensino da Estatística, nomeadamente informando outros professores que se deparem com as mesmas dificuldades e preocupações que tenho vindo a sentir na minha prática, no que respeita à criação de ambientes de sala de aula promotores do raciocínio estatístico e da aprendizagem da covariação.

Através da análise de dados, o desenvolvimento do raciocínio estatístico na exploração da relação entre duas variáveis, através da covariação, com os gráficos de dispersão. Cada ponto de dados de um conjunto disperso corresponde a uma unidade de análise entre as duas variáveis e os seus valores podem envolver a relação, associação, dependência ou até mesmo correspondência (Fitzallen, 2012). Com os scatterplots do *software Tinkerplots* permite de uma forma rápida, organizar grandes quantidades de informações e descreve a covariação de dois conjuntos que variam numa escala numérica bidimensional. Com este tipo de organização, viabilizam-se informações aos alunos, em que as duas variáveis podem não ser necessariamente dependentes uma da outra, mas exhibe a correspondência na ordenação de cada uma. Assim, o scatterplots auxilia o aluno a identificar a

relação entre duas variáveis e a variação da própria relação. Acredito que a sua interatividade facilitará a apreensão e compreensão dos conhecimentos na sala de aula, ajudando os alunos a organizar, representar e interpretar os dados nos diversos problemas e a estabelecer relações entre as variáveis.

## **1.2. Objetivo e questões do estudo**

Este trabalho de projeto incide sobre a Unidade de Ensino - Distribuições bidimensionais - no 10.º ano dos cursos profissionais de Técnico de Auxiliar de Saúde (TAS) e de Técnico de Apoio à Gestão Desportiva (TAGD), no âmbito da qual foi realizado um estudo com o objetivo de compreender as aprendizagens e dificuldades dos alunos no que respeita à covariação estatística, analisando o seu raciocínio covariacional quando utilizam o *software Tinkerplots* na resolução de tarefas envolvendo este conceito.

Tendo em conta estes objetivos, formulei um conjunto de questões mais específicas que servem de suporte ao trabalho de investigação:

- 1) Como é que os alunos exploram relações entre variáveis e estabelecem a existência de covariação, quando utilizam o *TinkerPlots* na resolução de tarefas envolvendo este conceito? Em particular, quais as representações que utilizam e qual a sua adequabilidade no estabelecimento dessa relação?
- 2) Quais as aprendizagens realizadas pelos alunos no final da unidade de ensino, no que respeita à covariação? Que dificuldades manifestam na compreensão do conceito?

## **1.3. Organização do estudo**

O presente estudo encontra-se organizado em seis capítulos. No primeiro capítulo – *Introdução* – procuro identificar a problemática em que o estudo se insere, apresentando a motivação para a sua realização bem como, a sua pertinência e os seus objetivos, concretizados num conjunto de questões às quais pretendo dar resposta.

O segundo capítulo – *Fundamentação e Enquadramento curricular* – diz respeito ao quadro teórico do estudo, que integra uma revisão de literatura sobre os resultados de investigações realizadas a nível nacional e internacional, relacionados com os temas principais deste trabalho de projeto e onde é feito também o enquadramento curricular do estudo. Este capítulo encontra-se dividido em três subcapítulos. No primeiro subcapítulo, *O ensino e a aprendizagem da Estatística*, mencionam-se algumas das razões que justificam o ensino da Estatística, dando-se uma ênfase particular à importância do ensino e aprendizagem da análise exploratória de dados e do desenvolvimento do raciocínio estatístico, bem como as orientações presentes no programa do ensino secundário e do ensino profissional. No segundo subcapítulo – *A Tecnologia no ensino da Estatística* – foco o papel da tecnologia no ensino da Estatística, articulada com as orientações preconizadas e, em particular, as potencialidades e as dificuldades a ter em conta na utilização do software *TinkerPlots*. No terceiro subcapítulo – *O Raciocínio covariacional estatístico* - discuto diversos aspetos (orientações curriculares, representações estatísticas, tecnologia e dificuldades dos alunos) relacionadas com o raciocínio covariacional.

O terceiro capítulo do estudo – *Metodologia* – apresenta as principais opções metodológicas e os resultados de um estudo exploratório que contribuiu para formar a proposta de intervenção que está na base deste estudo. Nele incluí, também, a caracterização dos participantes e do contexto escolar onde o mesmo se realizou. Por último, menciono como foi realizada a recolha dos dados e os métodos de análise dos dados recolhidos.

No quarto capítulo – *A unidade de ensino* – descrevo os princípios gerais da unidade de ensino que serve de base ao estudo. Começo por apresentar e fundamentar a planificação da unidade de ensino, com destaque, para as tarefas propostas aos alunos e os recursos utilizados. Refiro, de igual modo, a avaliação das aprendizagens dos alunos.

No quinto capítulo – *Aprendizagens dos alunos na unidade de ensino* – apresento os resultados do estudo, analisando o *raciocínio covariacional dos alunos* na exploração das tarefas com recurso ao *TinkerPlots* e as aprendizagens que evidenciam na *Prova de Avaliação Final*.

Finalmente, no sexto capítulo – *Conclusões e reflexões finais* – começo por fazer uma síntese do estudo, apresento as principais conclusões organizadas através das questões de investigação e faço uma reflexão final sobre o trabalho realizado e eventuais implicações e recomendações para a prática letiva futura.



## **CAPÍTULO 2**

### **FUNDAMENTAÇÃO E ENQUADRAMENTO CURRICULAR**

Neste capítulo apresento as orientações curriculares para o ensino e a aprendizagem da Estatística e uma revisão de literatura considerada relevante para enquadrar a problemática do estudo. Neste sentido, abordo aspetos gerais relacionados com a literacia e o raciocínio estatístico, enquanto finalidades do ensino da Estatística, refiro a importância do uso de tecnologia neste processo e foco-me depois, mais especificamente, sobre a covariação e o raciocínio covariacional dos alunos, aspeto fundamental do raciocínio estatístico.

#### **2.1. O ENSINO E A APRENDIZAGEM DA ESTATÍSTICA**

Nas últimas décadas, a Estatística têm tido uma importância crescente na educação devido ao seu papel na cidadania (NCTM, 2008). Daí, tenha começado a fazer parte das orientações curriculares em Matemática, um pouco por todo o mundo (GAISE, 2005; NCTM, 2008). Desde cedo, os alunos deverão desenvolver a capacidade de interpretar, avaliar criticamente e comunicar acerca da informação estatística, tendo em conta a incerteza (Batanero & Diaz, 2010; Gal, 2002).

No nosso quotidiano, contactamos com informação estatística diversa, que influencia as decisões que tomamos. No entanto, podemos não a conseguir compreender, caso não tenhamos conhecimentos de Estatística. Desta forma, é importante que os alunos compreendam o processo de geração das informações estatísticas, para que possam ser capazes de argumentar e tomar decisões

conscientes, para prosseguirem os seus estudos com sucesso e se inserirem no mundo do trabalho (NCTM, 1991). Como refere Branco e Martins (2002, p.9), relativamente à literacia, “o seu significado é fundamentalmente o seguinte: capacidade do indivíduo para ler, escrever e falar na sua língua materna, efetuar cálculos e resolver problemas do dia-a-dia, de forma a cumprir as tarefas que lhe são exigidas tanto no emprego como na sociedade”. Gal (2002) acrescenta que literacia não é só de conhecimentos factuais e de destrezas formais e informais, mas também de crenças, hábitos e atitudes, face a uma perspetiva crítica desse conhecimento.

É, por isso, fulcral para um aluno desenvolver a literacia estatística, porque só deste modo ele será capaz de compreender um conteúdo publicado num jornal ou na televisão e ter uma atitude crítica e reflexiva na nossa sociedade. Desta forma, Gal propõe um modelo (Figura 1) que requer o envolvimento de uma componente de conhecimento, em que estão incluídos elementos cognitivos (competências de literacia, conhecimento de estatística, conhecimento matemático, conhecimento do contexto e questionamentos críticos) e uma componente afetiva, que envolve atitude crítica, crenças e atitudes.

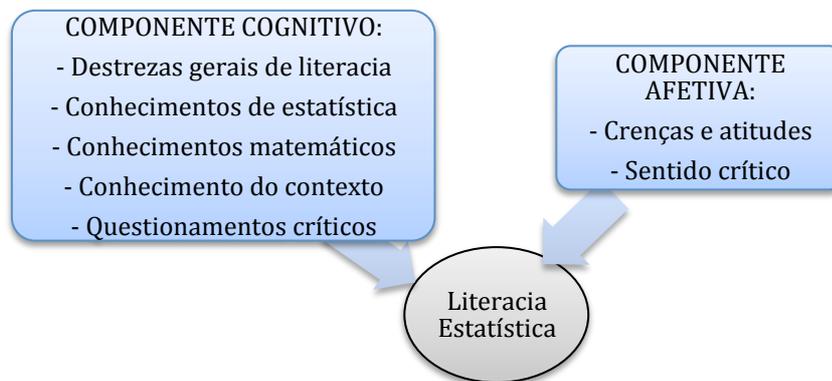


Figura 1 – Modelo de literacia estatística de Gal (2002).

Para Ben-Zvi e Garfield (2004), a literacia estatística envolve a capacidade de interpretar gráficos, ler e dar sentido à informação estatística apresentada nos meios de comunicação social e a compreensão da linguagem estatística, tais como palavras, símbolos e termos. Envolve, também, a capacidade de exprimir opiniões

e argumentar baseados em dados e em vários contextos, nomeadamente na vida profissional ou pessoal.

Assim, o desenvolvimento da literacia estatística nos alunos é essencial para fomentar o espírito crítico em relação à informação divulgada, entender e comunicar com base nessa informação e potenciar a tomada de decisões. No entanto, não se pretende criar especialistas na área da Estatística, mas simplesmente desenvolver nos alunos a capacidade de compreender os processos mais simples de recolha de dados e proporcioná-los a entender o que está subentendido no raciocínio estatístico, preparando-os assim, para uma melhor inserção na vida social (Branco & Martins, 2002; Watson, 2006).

A literacia estatística representa um dos objetivos curriculares, independentemente do nível de escolaridade e da área científica associada, como tem vindo a ser apontado em algumas orientações curriculares, nomeadamente o programa de Matemática do Ensino Básico (ME, 2007) e no NCTM (2008). Para o programa de Matemática para o Ensino Básico (ME, 2013), a ênfase deixou de ser para o desenvolvimento da literacia estatística e passou a ser para o conhecimento de procedimentos e técnicas.

Atualmente, a literacia estatística é reconhecida como uma aptidão essencial a todos os indivíduos, que deverá resultar da educação. Deste modo, justifica-se a inclusão da Estatística nas orientações curriculares (Martins & Ponte 2010).

Como sugerido no relatório *Guidelines for Assessment and Instruction in Statistics Education* (GAISE, 2005), há seis princípios a que se deve atender para se desenvolver a literacia estatística nos alunos: (1) evidenciar a literacia estatística e desenvolver o pensamento estatístico; (2) utilizar dados reais; (3) realçar a compreensão dos conceitos, em vez de ser só teoria e procedimentos; (4) fomentar uma aprendizagem ativa na sala de aula; (5) utilizar tecnologia para desenvolver a compreensão dos conceitos e a análise dos dados; e (6) utilizar a avaliação para conhecer e melhorar a aprendizagem dos alunos (pp.16-17).

Os autores deste relatório concebem a literacia estatística como o objetivo principal no ensino da Estatística, em que como nos outros campos de conhecimento científico, a Estatística possui linguagens e conceitos próprios. Neste documento, considera-se que o desenvolvimento da literacia estatística associa-se à

capacidade de compreender as ideias fundamentais da Estatística e a linguagem utilizada durante a sua comunicação, nomeadamente em termos estatísticos, símbolos de escrita e gráficos, que integra o uso de tecnologias.

Este documento defende, ainda, a realização de investigações estatísticas pelos alunos a partir de um tema do seu interesse, favorecendo a compreensão da importância dos dados para o dia-a-dia, e o porquê e o como uma investigação deve ser realizada. Numa investigação estatística, a formulação de uma questão representa o seu ponto de partida: “Os estudantes devem ser capazes de formular questões e determinar como os dados podem ser recolhidos e analisados para fornecer uma resposta “ (GAISE, 2005, p. 61). Isto sugere a realização de experiências nas quais os alunos façam escolhas sobre o que irão investigar e como.

As investigações de Campos (2007), Gal (2000), Martins e Ponte, 2010, Rumsey (2002) e Watson (1997) definem a literacia estatística dos alunos, como a capacidade para discutir opiniões e que os alunos precisam de aprender a usar estatísticas para evidenciar, argumentar e justificar situações quotidianas. Segundo Konold e Higgins (2003), na formulação de questões, os alunos devem elaborar questões estatísticas, a partir do problema geral, para que possam ser respondidas através da análise de dados. Desta forma, os alunos poderão descobrir não só uma variedade de respostas possíveis, como também poder antecipar as várias interpretações de resposta. Relativamente à formulação de conclusões, Henriques & Oliveira (2012), refere que é necessário refletir sobre a investigação realizada, de modo a poder responder às questões iniciais ou até mesmo, formular novas questões.

Como refere Ponte e Fonseca (2001) e Ponte e Sousa (2010), o ensino da Estatística, apesar de estar inserido no currículo de Matemática, não deve ter uma orientação semelhante à que é utilizada no ensino da Matemática. Pretende-se que os alunos desenvolvam a sua capacidade de ler, interpretar e usar os dados criticamente, para além de aprender e privilegiar um conjunto de habilidades para determinação e interpretação de medidas estatísticas bem como para a construção e interpretação de tabelas e gráficos.

Frequentemente, o ensino da Estatística é caracterizado por aulas expositivas de conteúdos teóricos, com a resolução de exercícios de aplicação onde se valoriza a

aprendizagem da representação de dados em tabelas e gráficos e a aplicação de fórmulas e procedimentos de cálculo de medidas estatísticas. Atualmente, o ensino da estatística assume uma perspectiva diferente, no sentido de desenvolver nos alunos a capacidade de leitura e interpretação de dados, permitindo-lhes responder de modo fundamentado às várias questões que lhes são colocadas em situações do seu quotidiano e ao mesmo tempo, desenvolver a capacidade de planear e executar investigações e o seu sentido crítico face ao modo como a informação é apresentada (Carvalho, 2009; ME, 2001a; Ben-Zvi & Garfield, 2004; Martins & Ponte, 2010).

O foco do ensino deve ser alterado, substituindo as tarefas estruturadas com base em conjuntos de dados já organizados, de dimensão reduzida, onde se pede a construção de gráficos específicos, por uma abordagem orientada para os dados, considerados números com um contexto. Espera-se que os alunos sejam capazes de exprimir descrições, juízos, inferências e opiniões fundamentadas em dados.

Este objetivo sugere, assim, a importância de desenvolver o raciocínio estatístico nos alunos, através de um trabalho exploratório e investigativo, orientado para os dados (GAISE, 2005; NCTM, 2008).

No ensino da Estatística, o uso de dados reais poderá ajudar os alunos, a desenvolver a compreensão estatística e a proporcionar o seu envolvimento no tema, pois favorecem o interesse e a relevância que estes atribuem à atividade escolar (Hall, 2011). Por isso, quando se recorre a situações da vida real do aluno ou com ele relacionadas, tende-se a verificar um aumento do significado, interesse, motivação por parte dos alunos, uma vez que não está centrado no professor e em atividade de rotina, onde só seja a aplicação de fórmulas e procedimentos, sem nunca desenvolver a interpretação (Batanero, 2000).

Os Princípios e Normas para a Matemática Escolar (NCTM, 2008) vieram reforçar a importância da Estatística nas orientações curriculares, bem como salientar aspetos que deverão ser valorizados no seu ensino. Neste documento (NCTM, 2008, pp. 53-55), no tema da Análise de Dados e Probabilidades e para todos os níveis de ensino (do pré-escolar até ao 12.º ano), sugere-se que todos os alunos deverão: (i) formular questões que possam ser abordadas por meio de dados e recolher, organizar e apresentar dados relevantes que permitam responder a essas questões, pois a natureza de dados pode ser influenciada pelo

modo como é feita a sua escolha; (ii) selecionar e usar métodos estatísticos adequados à análise de dados, em que devem analisar e compreender os dados como um todo, caracterizando-os e comparando-os, de modo a verificar como é que eles estão distribuído e daí transmitir a informação desejada de uma forma mais clara; (iii) desenvolver e avaliar inferências e previsões baseadas em dados, onde devem conseguir chegar a conclusões baseadas neles e prever como podem variar, podendo realizar investigações mais exaustivas sobre o comportamento dos seu dados; (iv) compreender e aplicar conceitos básicos de probabilidade, descrevendo os conhecimentos como prováveis ou improváveis e discuti-los.

A exploração das medidas estatísticas, tais como localização e dispersão, são também referidas como fundamentais à aprendizagem, uma vez que estas medidas permitem a caracterização da distribuição estatística e a sua compreensão é uma componente importante da literacia estatística (Curcio, 1987; Gal, 2002; Groth, 2006). No que se refere ao ensino secundário, os alunos também deverão “compreender que existem maneiras de quantificar o grau de certeza de alguns resultados estatísticos” (NCTM, 2008, p.55) e “a descrição do centro , da dispersão e da forma é essencial para a análise tanto de dados univariados como de bivariados” (NCTM, 2008, p. 385).

Tanto as orientações dos programas de Matemática A e B (ME, 2011a, 2011b), programa de componente de Formação Científica da disciplina de Matemática (Martins, et al., 2005), como o NCTM (2008) apontam para o desenvolvimento da literacia estatística, pois ambos desenvolvem a análise e interpretação de dados estatísticos, proporcionando aos alunos o desenvolvimento da avaliação crítica da informação a que têm acesso do seu quotidiano e estimulando-os a colocar questões e encontrar as respetivas respostas, partindo da análise dos dados estatísticos.

Atualmente, a Estatística faz parte de todos os programas de matemática do ensino secundário – Matemática A (ME, 2001a), Matemática B (ME, 2001b), Matemática Aplicada às Ciências Sociais (ME, 2001) e Programa do Ensino Básico (2013) e bem como, em cursos profissionais de nível Secundário (Martins, et a.l. 2005). No ensino secundário, os programas de Matemática recomendam um reinvestimento e aprofundamento de noções estatísticas, já desenvolvidas no 3.º ciclo, referindo que os alunos deverão “ficar a saber organizar, representar e

tratar dados recolhidos em bruto (ou tabelados) para daí tirar conclusões numa análise sempre crítica e sempre consciente dos limites do processo de matematização da situação” (ME, 2001, p. 29).

A interpretação de tabelas e gráficos, o planeamento e recolha de dados, questões éticas relacionadas com a experimentação, vantagens, desvantagens e limitações das medidas estatísticas, introdução gráfica à análise de dados bivariados quantitativos, modelos de regressão linear, dependência estatística e a relação entre variáveis qualitativas, são temas no programa do ensino profissional (Martins, et al., 2005). Estes temas são importantes para os alunos, pois através da “fase de organização dos dados recolhidos em que se procura reduzir” (Salomé & Silva, 2014, p. 2) e a “procura de conclusões e tomadas de decisões para um conjunto, onde se recolhem os dados”(Salomé & Silva, 2014, p. 2), os alunos conseguem tirar conclusões e inferir situações de correlação ou não correlação, “de modo a responderem a questões ou tomarem decisões fundamentadas no local de trabalho em situações do dia a dia” (NCTM, 2008, p. 383).

A partir da representação gráfica em forma de diagrama de dispersão ou diagrama de pontos, os professores devem proporcionar oportunidades para a exploração da existência de associação linear entre duas variáveis, e transmitir aos seus alunos uma “ideia intuitiva de correlação” (Martins, et al., 2005,p. 23), recorrendo a exemplos gráficos de correlação positiva, negativa ou nula.

Os autores do programa de matemática do ensino secundário – Matemática A (ME, 2001a), Matemática B (ME, 2001b), Matemática Aplicada às Ciências Sociais (ME, 2001) e Programa de Componente de Formação Científica da disciplina de Matemática, de cursos profissionais de nível Secundário (Martins, et al., 2005), bem como, Garfield e Ben-Zvi (2008) e Rubin (2007) referem a possibilidade e a importância do uso de tecnologia, para trabalho regular na sala de aula ou para demonstrações com todos os alunos, pois não se consegue atingir os objetivos e as competências do conteúdo sem recorrer à parte gráfica. A tecnologia permite tornar mais acessíveis os conceitos complexos e as ideias estatísticas, devido às suas potencialidades, nomeadamente nos domínios do tratamento de dados e da representação gráfica e da simulação, permitindo assim atividades não só de exploração, como também de pesquisa. Para além disso,

aborda as incompreensões dos alunos e desenvolve o seu raciocínio estatístico (Alias, 2009; Ben-zvi, 2004; Fitzallen & Watson 2010; Watson & Donne, 2009).

No 10.º ano, no ensino profissional (ME, 2005), a Estatística é lecionada como a ciência que trata os “dados”, enfatizando duas fases do trabalho estatístico. Numa primeira fase, a recolha e organização de dados, reduzindo-os e representando-os de uma forma adequada que permitam, numa segunda fase, retirar conclusões para um conjunto mais vasto, realizando inferência estatística. Segundo o NCTM (2008) é importante que os alunos desenvolvam a capacidade de propor e justificar previsões e conclusões a partir de dados, por forma, a investigá-los. No estudo dos dados numérico bivariados, dá-se relevância à representação gráfica sob a forma de diagrama de dispersão, a partir do qual se identificam padrões, tendências ou aglomerações nos dados, a existência de associação, a avaliação da intensidade da correlação entre duas variáveis quantitativas e do estudo da qualidade de ajuste de uma linha de regressão ao diagrama de dispersão. A representação gráfica permitirá estabelecer e compreender relações bivariadas, mais concretamente, estabelecer conexões entre os valores de coeficiente de correlação e diagramas de dispersão, em que os alunos desenvolvem melhor o sentido dos diferentes níveis de covariação e entendimento relativamente a fatores que podem influenciar o valor do coeficiente de correlação, sendo esse maior ou menor (Estepa & Batanero, 1996, 1999).

Com efeito, o programa da estatística no 10.º ano, no ensino profissional, trabalha-se o desenvolvimento do raciocínio sobre dados bivariados – raciocínio covariacional (Garfield & Ben-Zvi, 2008; Quintas, Oliveira & Ferreira, 2013), partindo de diagramas de dispersão, identificando tendências de associação entre as variáveis, quantificando o grau dessa associação através da correlação e expressando-a usando a reta de regressão.

## **2.2. A TECNOLOGIA NO ENSINO DA ESTATÍSTICA**

O rápido desenvolvimento da tecnologia na nossa sociedade tem proporcionado à escola, a utilização de tecnologias educacionais, em sala de aula. O papel central

que a tecnologia assume no desempenho das nossas funções faz com que ocorram mudanças na forma de estar e agir em sociedade.

A procura de novas formas de chegar aos alunos e de os auxiliar nas suas aprendizagens, tem exigido às escolas o recurso a ferramentas cuja manipulação nem sempre é simples, mas que podem proporcionar melhores condições de ensino e uma inovação curricular em Matemática. Assim, a introdução da tecnologia no ensino da Estatística não só é reforçada no sentido de alargar o que é ensinado, como também de modo a influenciar os conceitos estatísticos que são aprendidos (Batanero, Estepa & Godino, 1997; NCTM, 2008). Por forma a atingir os objetivos definidos para o ensino da Estatística, os diferentes documentos curriculares (ME, 2001a, 2001b, 2001; Martins, et al., 2005; NCTM, 2008) referem a possibilidade e a importância do uso de calculadoras e de *software* específico, para trabalho regular na sala de aula, utilizando situações reais que permitam a aplicação dos conceitos estatísticos de modo significativo (GAISE, 2005).

O uso da tecnologia, para além de estimular a criatividade e a dinâmica da aprendizagem na sala de aula, possibilita a exploração de situações reais de uma forma que antes não era praticada, pois grande parte do tempo que era dedicado ao cálculo passa agora ser possível utilizá-lo na interpretação e discussão dos resultados, estando centrado em aspetos interpretativos (Jolliffe, 2007; Pocinho & Gaspar, 2012). Ou seja, o aluno passa a ter um papel mais ativo no processo educativo, no qual poderá dedicar-se à pesquisa, descoberta, colaboração e não só, à realização de exercícios rotineiros.

Carvalho (2006) refere que as tecnologias são consideradas como um recurso importante para trabalhar os conteúdos de Estatística, possibilitando ao aluno a exploração de uma situação estatística, onde tem possibilidade de aplicar conceitos, formular inferências e descobrir padrões ou regularidades, adquirindo, assim, o poder estatístico.

No estudo das medidas de tendência central (a média, a mediana e a moda de uma distribuição), a tecnologia é apontada como um bom auxílio, uma vez que permite aos alunos estruturarem, investigarem e registarem a informação retida, de maneira mais dinâmica.

Segundo os autores, Engel e Sedlmeier (2011) e Garfield e Ben-Zvi (2008), a tecnologia permite a visualização e estabelecimento de conexões de diversificadas representações gráficas, bem como a própria exploração e manipulação de dados nelas representados.

Mas o uso de tecnologia não é a solução para todos os problemas que enfrentamos, quando pretendemos ajudar os alunos na compreensão de um conceito estatístico e melhorar a experiência de aprendizagem (Alias, 2009). Por isso, o professor deve refletir sobre de que forma a tecnologia pode ajudar nas aprendizagens dos alunos e o seu contributo, antes de a utilizar na sala de aula.

É destacada por vários autores (Fitzallen, 2012; Goos & Bennison, 2008; Konold, 2005; Pocinho & Gaspar, 2012), a importância do computador e do *software* educativo na aprendizagem da Estatística, pelas suas potencialidades. As características da tecnologia computacional torna-a uma ferramenta adequada para processos de descoberta e construção do conhecimento, permitindo, ainda, diferentes ritmos de aprendizagem dos alunos e o desenvolvimento da sua autonomia. Esta ferramenta permite aos alunos experimentar e explorar todas as fases que envolve os estudos estatísticos, nomeadamente nos domínios do tratamento dos dados, da representação gráfica, da simulação bem como na interpretação e compreensão de conceitos abstratos e das inter-relações entre os conceitos e a comunicação dos resultados (Batanero, 2001; Carvalho & Silva, 2001; Pocinho & Gaspar, 2012).

A utilização de tecnologia na sala de aula, proporciona uma nova dinâmica, onde os alunos desempenham um papel ativo na construção dos seus conhecimentos, sendo estes conduzidos pela orientação do professor. Além disso, facilita a representação e manipulação de dados, possibilita a compreensão de conceitos em vez de focar só a atenção nos procedimentos de cálculo, a realização de tarefas de carácter investigativo e exploratório e promove espaços de discussão, contribuindo para a consciencialização e valorização do papel da Estatística na sociedade relacionado com o desenvolvimento do pensamento estatístico. Desta forma, o uso da tecnologia permite desenvolver a capacidade de raciocinar dos alunos, de modo a que consigam valorizar, compreender e avaliar a evidência estatística que influencia o seu dia-a-dia, trabalhando grandes conjuntos de dados que conduzem à formulação de hipóteses e conclusões (Ridgway, Nicholson & McCusker, 2011).

O uso da tecnologia no ensino da Estatística possibilita, também, trabalhar muitos dados ao mesmo tempo, que seria difícil de usar de outra maneira (Chaput, Girard & Henry, 2011; Martins & Ponte, 2010).

De acordo com Goos e Bennison (2008), a integração do computador no ensino da Estatística proporciona a adoção de um conjunto de propostas de trabalho, como sejam os exercícios, a resolução de problemas, as tarefas de carácter exploratório ou investigativo, a modelação e própria comunicação e discussão coletiva de ideias. Além disso, desenvolve o espírito crítico, uma vez que o *software* estatístico torna-se indispensável para avaliar a veracidade e a razoabilidade dos resultados fornecidos por estas ferramentas, formulando e testando conjecturas próprias e criadas pelos alunos, bem como a exploração de várias ideias (Canavarro & Ponte, 1997). Contudo, pode exigir, por vezes, a tomada de decisões por parte do professor, quer na sua planificação quer na sua implementação, pois cria “um ambiente de aula com mais movimento, mais ruído, mais sobressaltos e receios para o professor” (Amado & Carreira, 2008, p. 287).

Ao longo de vários estudos (Ben-Zvi & Arcavi, 2001; Fitzallen, 2012; Fitzallen & Watson, 2010; Konold, Harradine & Kazak, 2007; Watson, 2012; Watson & Donne, 2009) tem-se evidenciado as potencialidades do uso de *software dinâmico* gráfico, o *TinkerPlots<sup>TM</sup>* (Konold & Miller, 2001; Konold & Miller, 2005), no desenvolvimento da compreensão de conceitos estatísticos, visto ser uma ferramenta que permite a representação de dados e a sua análise, quer seja com alunos do ensino básico quer com alunos do secundário.

Segundo esses autores, este *software* motiva o aluno a realizar atividades dentro da sala de aula e fora dela (Konold, 2006). É considerado um *software* interativo, de fácil acesso e de compreensão pelos alunos, pois permite-lhes inserir os seus próprios dados, importar dados de outras aplicações (tais como, bases de dados em Excel), criar representações gráficas a partir de pontos, organizando-os progressivamente (usando as operações básicas da barra de ferramentas, tais como, “Stack”, “Order” e “Separate”).

A utilização deste *software* gráfico, *Tinkerplots<sup>TM</sup>*, permite aos alunos observar uma grande multiplicidade de gráficos, onde comparam e contrastam diferentes dados, através do uso de um gradiente de cores sobrepostas para detetar variações. Esta

utilidade permite auxiliar na leitura dos dados, para além de se poder analisar um dado de uma forma variável, no sentido de buscar mais uma terceira informação acerca daquele próprio dado.

O uso deste *software* como instrumento de investigação, de demonstração ou de exploração, permite que os alunos construam diversos gráficos com facilidade e rapidez, o que lhes proporcionará conjecturar as suas próprias descobertas, que por seu turno promove a discussão entre os alunos, favorecendo a argumentação na procura da justificação para as afirmações proferidas ou na indagação de uma visão global que integre a multiplicidade de visões apresentadas (Ben-Zvi, 2006; Fitzallen, 2012; Watson, 2008; White, 2009).

Isto possibilita aos alunos o desenvolvimento de um sentimento de posse individual sobre a interpretação, sem se limitarem a aceitar o que o professor propõe, ocorrendo, assim, grandes diferenças na dinâmica da sala de aula (Dick, 1996). Desta forma, os alunos passam a ter a oportunidade de formular as suas próprias hipóteses, testando-as e ao mesmo tempo, enquadrando os resultados na teoria que estão a tentar formular (Grant & Searl, 1996).

A utilização deste *software* permite aos alunos, a qualquer momento, alterar facilmente as variáveis introduzidas e verem de imediato as mudanças nas propriedades estatísticas e nas representações gráficas, através das cores dos *Plots*. Desta forma, incentiva os alunos a um estudo mais profundo de um fenómeno, a testagem de conjecturas ou o estudo da influência da variação de dados de uma determinada amostra (Konold, 2005).

As funcionalidades deste *software* proporciona um ambiente de aprendizagem, onde os alunos experimentam e exploram diversos elementos e conceitos estatísticos, desenvolvendo as suas próprias compreensões da Estatística. Por isso, o seu uso no ensino da Estatística é inquestionável, pois com esta ferramenta, desenvolve-se nos alunos a compreensão crítica, onde o valor dos resultados depende do valor dos dados e da forma escolhida para o seu tratamento (Ponte & Canavarro, 1997).

O *TinkerPlots*<sup>TM</sup> pode ser visto como uma tecnologia que permite reorganizar, através das características dinâmicas de arrastar, articular múltiplas representações, sobrepor medidas estatísticas (por exemplo, média, mediana e quartis) sobre os

gráficos, estendendo a representação gráfica. Esta visualização não é possível sem o recurso à tecnologia, a qual proporciona aos alunos uma forma de reorganizar as suas concepções sobre diversas distribuições (Lee & Hollebrands, 2008; Lee, Kersaint, Harper, Driskell & Leatham, 2012).

Fitzallen e Watson (2010) no seu estudo com alunos do 2.º ciclo, demonstrou que o uso dos *Bins*, para além de serem úteis a separar os dados por completo, ilustra a flexibilidade do *TinkerPlots*, pois permite aos alunos escolher, quais os atrai visualmente. Destacou, ainda, que os alunos desenvolveram a capacidade de utilizar o *TinkerPlots* de forma criativa, discriminando as suas necessidades de forma apropriada, sem estarem constantemente, a seguir instruções do professor. Considerada como uma forte potencialidade deste *software*, este estudo demonstrou que os alunos criaram gráficos com significado para eles, raciocinando sobre os dados neles representados, e recuaram e avançaram entre a formulação de hipóteses e a criação de gráficos para darem sentido aos dados.

Alguns estudos (Bakker, *et al*, 2006; Ben-Zvi, 2006; Fitzallen, 2012; Rubim & Hammerman, 2006; Watson, 2008) referem que a utilização do software educativo reforça o raciocínio estatístico dos alunos. Por exemplo, para Ben-Zvi (2006) o uso do *Tinkerplots* proporcionou aos alunos desenvolver aspetos da inferência informal e o raciocínio argumentativo, como também, discutir e explicar os seus argumentos a respeito dos dados e sugerir a possibilidade dos mesmos, conseguirem representar os dados em escalas diferentes que contribuiu para o teste de hipótese. Igualmente, a interação com as informações e a possibilidade de utilizar uma ferramenta do *Tinkerplots*, onde é possível desenhar uma linha ou fazer demarcações nos *cases*, em gráficos de dispersão, que favorecem a análise da variação e da média dos dados.

### **2.3. O RACIOCÍNIO ESTATÍSTICO COVARIACIONAL**

Não existe uma definição unanimemente aceite para raciocínio estatístico. Na literatura, os conceitos de literacia, pensamento e raciocínio estatísticos são usados, algumas vezes, de modo indiferenciado atendendo à sua inter-relação (Ben-Zvi & Garfield, 2004; Martins & Ponte, 2010).

A definição mais comum de raciocínio estatístico dá especial atenção ao modo como as pessoas raciocinam com as ideias estatísticas, por forma a dar sentido às informações estatísticas, englobando a compreensão e a capacidade de explicar os processos estatísticos (Andrade, 2008; Garfield & Ben-Zvi, 2007). Desta forma, envolve fazer interpretações baseadas em representações, conjuntos ou resumos de dados, ou seja a capacidade de “interpretar completamente resultados estatísticos” (Ben-Zvi & Garfield, 2004, p. 7). Segundo Garfield e Ben-Zvi (2007), o raciocínio estatístico envolve a realização de interpretações e inferências estatísticas, estabelecendo articulações entre conceitos estatísticos (tais como, entre medidas de localização e medidas de dispersão) e entre diferentes representações estatísticas ou mesmo ideias estatísticas envolvendo o acaso.

Segundo Garfield e Gal (1999) para o aluno desenvolver o seu raciocínio estatístico, o ensino deve proporcionar-lhe condições para: (1) compreender a lógica das investigações estatísticas; (2) compreender os processos presentes numa investigação estatística; (3) dominar procedimentos estatísticos de modo que os mesmos desenvolvam uma ideia clara da natureza e dos processos que estão envolvidos numa investigação estatística; (4) estabelecer ligações nos procedimentos estatísticos, envolvendo o que se faz com a Matemática e quais as ideias matemáticas presentes; (5) terem a noção de probabilidade e de incerteza, através de atividade, onde possam ser simuladas e depois discutidas estas ideias; (6) desenvolver a capacidade de comunicar estatisticamente incentivando-os, à utilização de terminologia estatística de uma forma crítica; (7) desenvolver atitudes estatísticas seguindo uma metodologia de investigação.

Garfield (2002) afirma que os professores devem ensinar os conceitos e procedimentos estatísticos, bem como a utilização dos dados, mas refere que o raciocínio deve ser desenvolvido ao longo desse processo, como uma consequência imediata. Não havendo assim, consenso entre os investigadores sobre uma forma única de ajudar os alunos a desenvolverem o raciocínio estatístico ou de determinar o correto nível de raciocínio. Assim, a autora propõe um modelo que descreve e identifica cinco níveis de raciocínio estatístico, que devem ser desenvolvidos nos alunos, sintetizado no seguinte quadro:

Quadro 1 – Síntese do modelo de raciocínio estatístico desenvolvido por Garfield (2002):

<b>RACIOCÍNIO</b>	<b>DESCRIÇÃO</b>
<b>Idiossincrático</b> <i>(Idiosyncratic reasoning)</i>	Conhece algumas palavras e símbolos estatísticos relacionados com a distribuição de amostragem, que os usa sem compreendê-los. Frequentemente mistura-os com informações não relacionadas.
<b>Verbal</b> <i>(Verbal reasoning)</i>	O aluno seleciona e comunica corretamente uma definição correta, mas sem compreender o seu significado. Isto é, compreende verbalmente alguns conceitos, mas não o consegue aplicar num procedimento real.
<b>Transitório</b> <i>(Transitional reasoning)</i>	Identifica corretamente uma ou duas dimensões de um conceito estatístico, mas sem integrá-los na totalidade.
<b>Processual</b> <i>(Procedural reasoning)</i>	Identifica corretamente as dimensões de um conceito ou processo estatístico, mas não integra totalmente estas dimensões. Pode prever corretamente que a amostragem de distribuição corresponde aos parâmetros dados, mas pode não explicar o processo.
<b>Processual integrado</b> <i>(Integrated Process Reasoning)</i>	Entende todo o processo estatístico: regras e conceitos, usando suas próprias palavras para explicar um conceito e/ou fazer previsões corretas com confiança.

Os modelos de raciocínio estatístico são essencialmente considerados como modelos descritivos onde a análise e a interpretação contínua de dados se podem misturar procurando correlações, que na falta destas podem construir uma interpretação subjetiva, mas que permite ajudar a esclarecer a forma como os alunos pensam sobre estatística: o que conhecem, compreendem e as dificuldades que apresentam (Shaughnessy, 2007).

Tanto Mugabe, Fernandes e Correia (2012) como Zieffler e Garfield (2009) sugerem que se deve despende tempo nas aulas, para poder desenvolver o raciocínio dos alunos sobre as distribuições univariadas, para melhorar o raciocínio dos dados quantitativos bivariados.

Para Selmer, Bolyard e Rye (2012), as atividades que recorrem a estudos estatísticos, em contexto próximo à realidade dos alunos, são importantes, pois os alunos envolvem-se no raciocínio estatístico através da formulação de questões significativas, da recolha de dados e da representação e análise de dados e da utilização dos resultados para formular novas questões. Ainda como refere Lopes (2012, p.167), o raciocínio estatístico tem a “tem a variabilidade como o centro do

processo de fazer relações sobre o problema investigado, de elaborar a construção e a análise dos dados”.

Em sala de aula deve-se desenvolver nos alunos uma compreensão profunda e significativa do que é a Estatística, para ajudá-los a desenvolver a sua capacidade de pensar e raciocinar estatisticamente, como por exemplo, incentivar os alunos a descreverem verbalmente o processo estatístico que estão a realizar.

É a utilização combinada de estratégias e recursos como textos, tarefas, discussão e tecnologia, que permite criar um ambiente de aprendizagem para os alunos desenvolverem o seu raciocínio (Garfield & Ben-Zvi, 2010). Cobb, McClain e Gravemeijer (2003) descrevem seis princípios a serem considerados no ensino da Estatística, nomeadamente: I) desenvolver as ideias centrais da estatística, em vez de apresentar um conjunto de procedimentos e de ferramentas; II) usar conjuntos de dados reais e motivadores, por forma a envolver os alunos na elaboração de conjecturas; III) usar atividades para apoiar o desenvolvimento do raciocínio dos alunos, focando-se na importância das tarefas; IV) integrar o uso de tecnologia adequadas que permita aos alunos testar as suas conjecturas, analisar dados e desenvolver o seu raciocínio estatístico; V) promover o discurso e o debate, que inclui argumentos estatísticos com ideias estatísticas significativas e VI) usar a avaliação para saber o que os alunos sabem e o desenvolvimento da aprendizagem estatística, por forma a criar situações no sentido de incrementar e desenvolver o pensamento estatístico.

A noção de associação estatística (covariação) é necessária para compreender outros conceitos e procedimentos estatísticos, por isso o seu ensino deve ser preparado para tal, pois não se adquire de forma espontânea. Esta capacidade em lidar com dados bivariados e compreender associações entre duas variáveis são importantes para a literacia estatística dos alunos dos níveis secundário e superior, que está complementada nas atuais orientações curriculares nacionais e internacionais para o ensino da Estatística (GAISE, 2005; NCTM, 2007).

No ensino da Estatística, sobretudo para o secundário, a compreensão do raciocínio sobre as relações entre dois atributos, que variam ao longo de escalas numéricas e à forma de como eles variam em relação um ao outro, conhecido por raciocínio covariacional, deve ser considerado mais do que somente raciocinar sobre diagramas de dispersão, correlação e regressão (Carlson, Jacob, Coe, Larsen

& Hsu, 2002; Garfield & Ben-Zvi, 2008; Zeiffler & Garfield, 2009). Desta forma, o raciocínio covariacional é sobre o raciocínio sobre a covariação estatística, quando se observa à forma como as variáveis variam entre si, e que envolve saber como avaliar e interpretar essas relações (Zeiffler & Garfield, 2009). No secundário, nas distribuições bivariadas, dá-se atenção à representação gráfica conhecida por diagrama de dispersão, à interpretação do coeficiente de correlação como medida do grau de associação linear entre duas variáveis e à ideia de reta de regressão que permite modelar os dados e obter um valor estimado de uma variável, conhecido o valor de outra (ME, 2001a, 2001b).

Num sentido geral, a covariação é uma correspondência entre variáveis, que podem variar de forma previsível e em simultâneo, que por vezes, essa relação poderá ser modelada através de uma associação (Batanero, et al., 1997; Grafel & Ben-Zvi, 2010; Moritz, 2004; Zeiffler & Garfield, 2009). Contudo, essa relação não tem que implicar uma causalidade, embora possa existir uma relação causal, que poderá ser determinada por um teste específico de causa e efeito.

Esta forma de abordagem, que será o foco deste estudo, permite aos alunos saber distinguir a “associação” entre duas variáveis e uma relação de “causa-efeito” que, não estando explícita nos programas portugueses, é salientado nos objetivos e aprendizagem das orientações curriculares internacionais (GAISE, 2005). É importante não confundir associação entre variáveis, estabelecida pelo coeficiente de correlação com a relação *causa-efeito*. Pois, um diagrama de dispersão e uma covariação estatística não permitem provar a existência de uma relação causa-efeito, ou seja, que uma das variáveis tenha causada a ocorrência da outra variável. Um aluno deve ser capaz de julgar e interpretar as relações entre variáveis através de diagramas de dispersão e entender que uma forte correlação entre duas variáveis não significa que uma seja a causa de outra (Garfield & Gal, 1999).

A natureza da covariação pode ser classificada de acordo com a variação possível na medida de cada variável envolvida. O raciocínio sobre a covariação envolve processos entre os dados estatísticos, representações gráficas e justificações verbais sobre a covariação e a associação (Ben-Zvi & Garfield, 2004; Moritz, 2004). É importante salientar que no ensino da relação bivariada, deve-se trabalhar com os alunos situações onde possam comparar conceitos e avaliar a maneira mais

apropriada de analisar um conjunto de variáveis, onde consideram as ideias de forma, centro e dispersão, bem como a compreensão do papel da relação bivariada em modelos e na previsão de eventos. Para desta forma, permita aos alunos promover o desenvolvimento do raciocínio estatístico (Garfield & Ben-Zvi, 2008; Fitzallen, 2012; Moritz, 2004).

Ben-Zvi e Garfield (2004) destacam como conceitos-chave para o desenvolvimento do raciocínio sobre a covariação dos alunos, os seguintes: correspondência, classificação e seriação. Para esses autores, os alunos devem ser capazes de compreender e interpretar estatísticas sumárias de dados, estabelecer ligações entre conceitos estáticos e entre diferentes estatísticas ou até mesmo, relacionar ideias estatísticas com o acaso. Os mesmos autores (2008) consideram, ainda, que ao estabelecerem conexões entre os valores de correlação e os diagramas de dispersão, permite aos alunos compreender diferentes níveis de covariação e entender acerca dos fatores que influenciam o maior ou o menor valor do coeficiente de correlação, focando em termos de sinal e força/intensidade.

Para Moritz (2004), o raciocínio covariacional envolve processos de tradução entre dados numéricos, representações gráficas e afirmações verbais sobre covariação estatística e associação causal, que será desenvolvida, quando os alunos são solicitados a conjecturar e a formular hipóteses sobre as relações entre as variáveis e daí, representá-los ou analisá-los numericamente e tirar conclusões sobre a relação. Para isso, Moritz (2004) apresenta um modelo (Figura 2) de raciocínio para a covariação, que relaciona as várias formas de a representar e os seus processos: i) geração especulativa de dados, demonstrados por um gráfico, por forma a representar uma informação verbal de covariação, requerendo a compreensão da covariação numérica e a compreensão contextual dos dados; ii) interpretação verbal de um gráfico, a partir da descrição de um gráfico de dispersão, através de uma explicação; iii) interpretação verbal dos dados, a partir da demonstração e interpretação de um valor. Estes dois processos de interpretação estão associados, devido à capacidade de ambos descreverem verbalmente as características das variáveis em estudo. Embora a leitura dos dados apresentados em gráficos de dispersão seja uma componente importante para a

leitura de gráficos, mas é através da visualização do gráfico, que o aluno consegue interpretar e generalizar a forma de como os dados são apresentados.

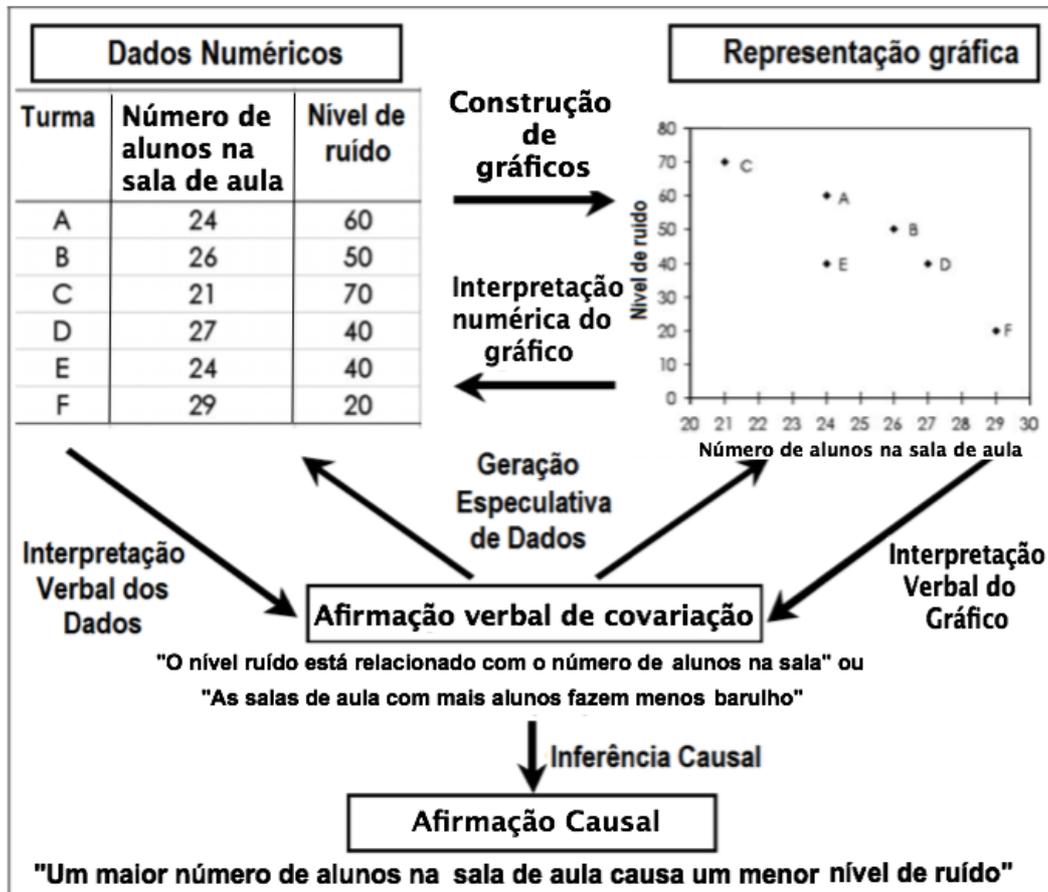


Figura 2 – Formas de representar a covariação estatística e capacidades para as traduzir (Adaptado de Moritz, 2004, p.230).

A covariação estatística, frequentemente designada também por correlação, refere-se à variação entre a correspondência de duas variáveis que variam ao longo de escalas numéricas. Essa covariação pode ser representada em diagramas de dispersão, utilizando um sistema de coordenadas que mostra a correspondência da ordenação de cada variável. Os gráficos de dispersão, descrevem covariação entre duas variáveis que variam ao longo de escalas (escalas contínuas e escalas segmentadas em intervalos) num espaço bidimensional. A produção de gráficos estatísticos de dados numéricos e a sua interpretação são os processos mais solicitados aos alunos durante a resolução de tarefas escolares e os mais investigados e discutidos relativamente à compreensão da covariação, por serem os mais fáceis para eles, permitindo-lhes organizar grandes quantidades de informação e visualizar a existência de relação entre dois conjuntos de medições

que variam ao longo de escalas numéricas. Deste modo, os alunos conseguem captar todas as características da relação entre duas variáveis, fundamentais para a identificação da covariação (Estepa, 2008).

O raciocínio covariacional pode ser desenvolvido quando os alunos são solicitados a formular hipóteses sobre a relação entre duas variáveis, a representá-las graficamente ou a analisá-las numericamente e daí, retirar conclusões dessa relação. Para isso, é necessário que os alunos dominem processos de tradução entre dados numéricos entre diferentes representações, tais como, as representações gráficas, as descrições verbais sobre a covariação e o valor do coeficiente (Moritz, 2004).

Têm sido bastante diversas as investigações sobre a compreensão da covariação e do raciocínio covariacional, permitido perceber como os alunos adquirem a noção de correlação e, ao mesmo tempo, destacando as dificuldades e as incompreensões sentidas por eles no que respeita a este conceito. Num desses estudos, Batanero, et al. (1997) salientam como uma dificuldade dos alunos a interpretação de relações em que o sinal do coeficiente de correlação é negativo. Nestes casos, os alunos consideram como independência a associação inversa. Os alunos tendem a compreender a existência de relação só quando a mesma é positiva, designada pelos autores por compreensão unidirecional de associação. Têm igualmente tendência para formularem relações causais, a partir da análise de correlação (Estepa, 2008). Um outro aspeto, salientado por Estepa e Batanero (1996), é relativamente às diversas conceções erradas que as autoras identificaram nas estratégias utilizadas pelos alunos do secundário para avaliar a associação estatística em diagramas de dispersão, nomeadamente: (i) conceção determinística, em que os alunos esperam que haja uma correspondência perfeita entre os valores da variável resposta e da variável explicativa, não admitindo exceções na relação entre as variáveis; (ii) conceção local, em que os alunos formam os seus julgamentos utilizando apenas uma parte dos dados fornecidos e acreditam que é o tipo de associação de todo o conjunto de dados; (iii) conceção causal, em que os alunos acreditam que a covariação pode ser atribuída a relações causais entre elas. Cobo, Estepa e Batanero (2000) afirmam que os erros mais frequentes ocorrem na representação gráfica a partir de uma descrição verbal ou durante a realização de tarefas que não estão familiarizados. Em alunos do 7.º ano, Chick (2004)

constatou que os alunos podem ser criativos nas suas próprias representações, quando tentam lidar com os dados por forma a representar a associação. Segundo Rubick e Yoon (2002), é necessário abordar no ensino a importância de uma representação para evidenciar afirmações sobre os dados. Além disso, devem agrupar e ordenar dados, enquanto estratégias de compreensão e representação de dados.

No estudo de Moritz (2000, 2004), os alunos foram capazes de traduzir afirmações verbais em gráficos e diagramas de pontos em afirmações verbais, quando analisam a covariação. Contudo, continuam a revelar algumas dificuldades, destacadas também por outros autores (por exemplo, Ben-Zvi & Arcavi, 2001), quando se focam em pontos individuais, em vez de procurar uma tendência global nos dados (conceção local) ou quando consideram uma das variáveis isoladamente. Outros problemas destacados por estes autores são referentes à forma como os alunos lidam com a covariação negativa, sobretudo quando as mesmas são contraditórias às suas crenças prévias sobre a relação entre duas variáveis. Konold (2002), no seu estudo, verificou que os alunos do 3.º ciclo não têm dificuldades em fazer julgamentos corretos sobre a covariação, mas apresentam problemas em descodificar o modo como essas relações são apresentadas.

A integração da tecnologia no ensino da Estatística pode mudar gradualmente o que se ensina e como se ensina (Garfield & Ben-Zvi, 2008), disponibilizando aos alunos ferramentas poderosas para as múltiplas representações que lhes permite desenvolver o significado dos conceitos (Batanero & Godino, 1998). Além disso, a tecnologia tem potencial para tornar mais acessíveis os conceitos complexos e as ideias estatísticas e desta forma, é inserido para abordar as incompreensões dos alunos e desenvolver o seu raciocínio estatístico (Alias, 2009; Hammerman & Rubin, 2004; Fitzallen & Watson, 2010; Watson & Donne, 2009).

No que concerne ao ensino e aprendizagem da covariação, o uso de *software* educativo na sala de aula permite apoiar o avanço dos alunos, em vários aspetos, tais como: (i) leitura de dados específicos para um olhar mais abrangente das diferenças entre dados, tentando compreender os dados como um todo; (ii) descrições qualitativas de tendências de dados para a estimação de retas de regressão através do método dos mínimos quadrados; (iii) fazer previsões

resultantes em dados particulares para interpolar; (iv) ter uma visão determinística para a tendência dos dados, quando reconhece a variação ao acaso; e (v) focar na construção de gráficos para interpretar, para que os alunos comecem a ter uma visão mais holística do gráfico (Ben-Zvi & Arcavi, 2000; Foster, 2007).

As investigações no desenvolvimento da compreensão de conceitos estatísticos, em particular o de covariação, recorrem às potencialidades do *software* gráfico, como o *TinkerPlots* (Konold & Miller, 2005). Este *software* permite passo a passo, controladamente, avaliar as alterações que os alunos realizam nas representações motivadas para um objetivo, bem como adicionar comentários escritos com facilidade e rapidez, por forma a criar oportunidades para conjeturarem as suas próprias descobertas, favorecendo, assim, a argumentação durante a procura de justificação para as afirmações ditas (Ben-Zvi, 2006; Fitzallen, 2012; Watson, 2008). Permitirá ainda, aos alunos, fazer julgamentos quando usam o *TinkerPlots*, possivelmente, porque este *software* oferece uma variedade de formas gráficas alternativas ao diagrama de dispersão para representar e interpretar a covariação.

Para analisar o modo como os alunos raciocinam em ambientes gráficos proporcionados pelo *TinkerPlots*, Fitzallen (2006) desenvolve um modelo teórico (Figura 3), que integra elementos considerados chave, para analisar a construção e interpretação gráfica de outros modelos de raciocínio estatístico.

O modelo apresenta quatro categorias de análise mas que estão relacionadas entre si: a) ser criativo com os dados; b) compreender os dados; c) pensar sobre os dados e d) conhecimento genérico. Este último porque se reconhece que quando se consideram ambientes com tecnologia devemos ter atenção conhecimentos/compreensões genéricas, que podem ser inerentes a todos os níveis de construção e interpretação gráfica.

CATEGORIAS		ELEMENTOS CHAVE
<i>Conhecimento genérico</i>	<i>Criatividade com dados</i>	Sumariar dados. Construir diferentes formas de gráficos. Descrever dados a partir dos gráficos.
Falar a linguagem dos dados e dos gráficos. Reconhecer as características dos dados e dos gráficos. Compreender como usar as características do <i>software</i> e do ambiente tecnológico.	<i>Compreensão dos dados</i>	Dar significado a dados e a gráficos. Compreender a relação entre tabelas, gráficos e dados. Identificar mensagens a partir dos dados. Responder a questões sobre dados. Reconhecer o uso apropriado de diferentes formas de gráficos.
	<i>Pensar sobre os dados</i>	Formular questões sobre os dados. Reconhecer as limitações dos dados. Interpretar dados. Fazer inferências causais baseadas nos dados. Procurar possíveis causas de variação. Procurar relações entre variáveis nos dados.

Figura 3 – Modelo de raciocínio estatístico dos alunos em ambientes gráficos com tecnologia (Adaptado de Fitzallen, 2006, p. 206).

Este modelo conceptual de Fitzallen (2006) permite caracterizar o desempenho dos alunos, no que diz respeito ao conhecimento genérico, à criatividade com os dados, à compreensão dos dados e ao pensamento sobre os dados, relativamente à covariação em ambientes gráficos.



## **CAPÍTULO 3**

### **METODOLOGIA DO ESTUDO**

Neste capítulo, dou a conhecer e fundamento os aspetos relacionados com a metodologia adotada no estudo, os participantes e os instrumentos e métodos utilizados na recolha e análise dos dados. Apresento, ainda, o estudo exploratório realizado com o objetivo de informar a unidade de ensino que está na sua base.

#### **3.1. OPÇÕES METODOLÓGICAS**

A escolha da metodologia a ser seguida numa investigação está estreitamente relacionada com a natureza do problema em análise e com as questões de investigação (Cohen & Manion, 1990), bem como com as condições em que ele ocorre (Abrantes, 1994; Ludke & André, 1986).

O objetivo deste estudo prende-se com a análise do raciocínio covariacional dos alunos do 10.º ano de um curso profissional quando utilizam o *software Tinkerplots* na resolução de tarefas envolvendo o conceito de covariação estatística e as aprendizagens e dificuldades decorrentes desse trabalho.

Atendendo a que o foco é nos aspetos qualitativos da construção de conhecimento matemático dos alunos, antevi uma metodologia que privilegiasse o contato direto e prolongado entre a investigadora e os participantes em ambiente natural - a sala de aula. Deste modo, o contexto pode ser observado ‘em ação’, permitindo uma interpretação do fenómeno essencialmente descritiva e fundamentada em dados empíricos. Além disso, a minha preocupação centra-se sobretudo na interpretação, compreensão e explicação dos acontecimentos do ponto de vista dos intervenientes, tendo em conta a sua singularidade e os contextos de interação social.

Pelo exposto, o estudo congregou particularidades de uma investigação qualitativa, de natureza interpretativa, onde os dados são recolhidos no ambiente natural dos alunos - em sede de sala de aula, permitindo uma observação participante por parte da investigadora (Bogdan & Biklen, 1994).

Ao procurar compreender o contributo da unidade de ensino que suporta este estudo, investigo também a minha própria prática enquanto docente. Existindo uma identidade entre a investigadora e a professora que orienta o trabalho dos alunos, pode considerar-se que se trata, também, de um estudo sobre a minha prática profissional, com as potencialidades e dificuldades específicas que isso comporta (Bogdan & Biklen, 1994).

Ao trabalhar conjuntamente com a docente da disciplina, permitiu “controlar” a atenção, concentração e comportamento dos alunos, tornando-se mais fácil para os alunos superar as dúvidas que fossem surgindo, proporcionar mais apoio, quer personalizado quer individualizado, aos alunos e colmatar a dificuldade que o professor, muitas vezes tem em conseguir chegar a todos alunos. Este trabalho colaborativo possibilitou trocar e criar materiais juntos e planificar e partilhar experiências com o objetivo comum de obter a efetiva participação dos alunos. A dificuldade deste trabalho remete-se à participação dos dois professores ao mesmo tempo no ensino dos alunos, pois os estilos de ensino eram diferentes entre os dois professores e para trabalhar facilmente em equipa, teve-se que escolher um modelo que permitisse partilhar o ensino dos alunos. Para tal, a abordagem escolhida, foi uma das professoras (investigadora) ensinar os alunos enquanto que a outra professora circulava pela sala e ajudava os alunos à medida que estes iam demonstrando dificuldades e controlava o comportamento dos alunos. A razão pela escolha desta abordagem deve-se ao fato da professora de disciplina usar um método de aula mais expositiva. A desvantagem para os professores, também, é o tempo que terão de dispendir mesmo antes das aulas começarem, na preparação e discussão das estratégias a adotar. Para os alunos, ter mais um professor na sala de aula, possibilita mais oportunidade de aprendizagens, relativamente no esclarecimento de dúvidas (Goetz, 2000).

## **3.2. ESTUDO EXPLORATÓRIO**

### **3.2.1. OBJETIVOS**

O estudo exploratório, que agora apresento, foi realizado com o objetivo de contribuir para a formação das principais características do estudo principal, sobretudo para a preparação da unidade de ensino que esteve na sua base. A partir dos seus resultados, tive oportunidade de reformular e refinar as questões iniciais do estudo, delimitando e focando a problemática para permitir um adequado aprofundamento das questões. Além disso, uma vez que a realização de tarefas foi a forma escolhida para ter acesso ao raciocínio dos alunos, o estudo exploratório permitiu ganhar compreensão sobre a estrutura e natureza das tarefas a construir e a propor aos alunos durante a unidade de ensino, de modo a constituírem uma fonte rica de evidência para o estudo do seu raciocínio covariacional e as suas aprendizagens e dificuldades.

### **3.2.2. ASPETOS METODOLÓGICOS**

Neste estudo exploratório procurei analisar como é que os alunos do 10.º ano exploram relações entre variáveis e estabelecem a existência de covariação, quando utilizam o *TinkerPlots* na resolução de tarefas envolvendo este conceito e de que modo o uso deste *software* pode contribuir para o desenvolvimento do seu raciocínio covariacional (Henriques & Antunes, 2014).

Os participantes, neste estudo, foram quatro alunos (2 rapazes e 2 raparigas) dos vinte e quatro alunos de uma turma do 10.º ano, que frequentaram a disciplina de Matemática A, no ano letivo de 2012/13, num estabelecimento de ensino particular e cooperativo de ensino básico e secundário, da região interior centro do país. O estudo, realizado no final do 3.º período do referido ano letivo, foi autorizado pela direção pedagógica da escola e pelos encarregados de educação,

tendo sido garantido o anonimato dos alunos intervenientes (de nomes fictícios Afonso, Cris, Leo e Margarida). O estudo não seguiu uma lógica de amostragem, envolveu antes uma escolha criteriosa de modo a permitir maximizar aquilo que eu poderia aprender acerca do fenómeno que estava a investigar (Stake, 1994). Assim, foram considerados três critérios fundamentais na escolha dos alunos: empenho, uma razoável capacidade de expressão escrita e oral e também a predisposição para participar no estudo. Atendendo às características do presente estudo, procurei definir um número de casos que permitisse ter uma variedade de aspetos interessantes do raciocínio covariacional dos alunos e, simultaneamente, constituísse uma dimensão de trabalho a que eu pudesse dar resposta.

Os dados utilizados, neste estudo exploratório, são oriundos de entrevistas clínicas individuais aos quatro alunos, seguindo os princípios e técnicas propostas por Goldin (2000). Na base do propósito de realizar entrevistas está a intenção de compreender o conhecimento do aluno em relação ao tópico em estudo e o significado que este atribui a esse conhecimento (Seidman, 2006). Ao aplicar este tipo de entrevista, como método de recolha de dados, baseei-me em três pressupostos: I) as respostas dadas pelo aluno refletem as suas ideias, com sinceridade e imparcialidade, procurando sempre, fazer o seu melhor; II) durante a entrevista, o aluno elabora interpretações para as situações que lhe são apresentadas em termos de conhecimento conceptual que ele considera relevante; III) a validade das respostas referentes ao momento a que foram obtidas, pelo que não se pressupõe que outro aluno na mesma situação, responda a mesma forma, usando as mesmas palavras (Koichu & Harel, 2007).

O protocolo usado na entrevista (em anexo 3) tem por base uma tarefa que pretende desafiar os alunos entrevistados a explorarem relações entre variáveis, enquanto interagem com o *TinkerPlots*. Enquanto trabalham no computador, os alunos são solicitados a descrever um conjunto de dados, a construir representações gráficas a partir deles, a tirar conclusões e a fazer inferências sobre os dados observados nos gráficos, permitindo-lhes demonstrar a compreensão sobre a covariação que puderam desenvolver ao longo da unidade de ensino e que fornecem, ao investigador, um conjunto rico de dados da unidade de ensino e que fornecem, ao investigador, um conjunto rico de dados descritivos sobre o raciocínio dos alunos (Seidman, 2006). Os dados da tarefa, que já se encontravam

introduzidos em cartões do *TinkerPlots*, referem-se aos resultados dos testes de condição inicial de uma turma de 9.º ano de escolaridade que incluem a observação das seguintes características dos alunos: peso, altura, resistência, diversos tipos de forças (superior, abdominal e inferior), flexibilidade, velocidade, impulsão vertical, IMC e sexo.

As entrevistas, com a duração aproximada de 30 a 40 minutos, foram realizadas na escola, de acordo com a disponibilidade apresentada pelos alunos, sem prejudicar o funcionamento das aulas e registadas em áudio e vídeo e transcritas posteriormente. Além disso, usei o *software AutoScreenRecorder 3.1 Pro* para gravar em vídeo as ações que os alunos realizaram com o *TinkerPlots* na exploração da tarefa do protocolo. Os ficheiros contêm a gravação de todos os movimentos do cursor enquanto os alunos trabalharam no computador, fornecendo um registo cronológico de como acederam às ferramentas do *TinkerPlots* e permitindo a sua articulação com os registos áudio das entrevistas.

Os alunos não tinham experiência com este *software*, apenas um contacto resultante de duas aulas de Estatística, onde foram dinamizados, por mim, momentos de apresentação das ferramentas do *software* na resolução de tarefas (do manual ou dos exames nacionais) que recorriam à construção de diversas representações gráficas e ao cálculo de medidas de localização.

Tal como sugerido por Bogdan e Biklen (1994), para que os alunos entrevistados se sentissem à vontade para expressar livremente os seus pontos de vista, abstive-me de fazer apreciações que pudessem ser encaradas como juízos de valor ou até mesmo inibir e influenciar a sua interpretação. Também procurei auxiliá-los na utilização das ferramentas do *software* e assumi uma postura de questionamento, por forma a clarificar as suas perspetivas e a incentivá-los a serem mais explícitos e específicos no seu discurso.

As transcrições e os ficheiros gravados das entrevistas foram analisados com base no quadro conceptual de Fitzallen (2006) tendo em mente dois objetivos, documentar o raciocínio covariacional dos alunos e o suporte fornecido ao mesmo pelo *TinkerPlots*.

### 3.2.3. RESULTADOS

A análise da interação dos alunos com este *software* teve por base o modelo conceptual de Fitzallen (2006) que se mostrou pertinente para caracterizar o seu desempenho, durante a entrevista, no que respeita ao conhecimento genérico, à criatividade com os dados, à compreensão dos dados e ao pensamento sobre dados, relativamente à covariação. Além disso, as questões da tarefa proposta requeriam a representação e a análise exploratória de dados, bem como a realização de inferências baseadas nessas representações (e na variação evidenciada). Esta complexidade da situação apresentada permitiu incorporar todas as categorias do modelo de modo relacionado, tal como a autora sugere.

*Conhecimento genérico.* Todos os alunos utilizaram, de modo criativo mas limitado, as ferramentas básicas do *TinkerPlots* para construírem várias representações gráficas para evidenciar a existência de relações entre duas variáveis e suportar as suas conjeturas formuladas verbalmente. De modo geral, todos os alunos evidenciaram compreender o significado das cores associadas aos dados, uma característica específica e facilitadora dos gráficos deste *software*, como evidenciado pela resposta: “os pontos são de cor diferente porque tem a ver com o valor [do atributo] que cada aluno tem”(Margarida). Quase todos os alunos necessitaram de apoio em relação à utilização de algumas ferramentas do *TinkerPlots*, contudo esta situação não surpreende, uma vez que foi a primeira vez que interagiram com este *software*. Todos compreenderam a linguagem usada na tarefa e perceberam o que lhes é pedido para fazer, não requerendo a clarificação de termos ou instruções sobre a aplicação dos recursos base do *TinkerPlots* na criação de gráficos. Contudo, ocasionalmente, recorreram a uma linguagem informal ou expressões inapropriadas para descrever os gráficos ou a informação neles representada, de que é exemplo a explicação: “há seis bolas que estão a crescer e ao puxa-las posso colocar num intervalo” (Cris).

*Criatividade com dados.* Os alunos criaram gráficos de covariação com escalas divididas em intervalos que permitiu resumir e organizar dados bivariados mas nem sempre os tentaram alterar recorrendo às ferramentas do *software* para os ajudar na identificação de relações ou confirmação/justificação das suas conjeturas.

A utilização do diagrama de dispersão, uma representação ajustada ao estudo da covariação, pois permite estabelecer e simultaneamente caracterizar a natureza das relações, apenas surge quando foram incentivados para tal, à exceção de um dos alunos - Afonso. Este aluno utilizou o diagrama de dispersão, de forma autónoma e intencional com o objetivo de facilitar a obtenção de evidências para as suas conjeturas sobre as relações identificadas, induzido pela sua experiência escolar e facilidade de interação com o *software*. Contudo, a falta de familiaridade que os alunos evidenciaram com o *TinkerPlots* limitou a sua criatividade. Esta situação remete para a importância dos alunos compreenderem as potencialidades do *software*, em particular como os diversos gráficos são construídos e a relação entre diferentes elementos do gráfico para serem capazes de avaliar e descrever a covariação (Fitzallen, 2012).

*Compreensão dos dados.* Os dados foram disponibilizados nos cartões do *TinkerPlots*, permitindo a todos os alunos aceder à informação com facilidade. Todos foram capazes de os articular com as representações gráficas que construíram, aspeto essencial para o estabelecimento de eventuais relações: “Conseguimos ver que há uma aproximação entre a altura e a velocidade. Os alunos com mais altura tendem a percorrer esse espaço rapidamente, que é o caso dos alunos Jéssica, Camila, Bruno e Mirian” (Leo). O desempenho dos alunos sugere que o desenvolvimento da compreensão de covariação e a avaliação de relações entre variáveis pode ocorrer de forma intuitiva antes da formalização de procedimentos associados a este conceito (Fitzallen, 2012; Moritz, 2004). Na sua maioria, foram capazes de reconhecer o uso apropriado das diferentes representações gráficas construídas, bem como, de descrever e interpretar os diagramas de dispersão construídos, focando-se tanto em características individuais, “Esta pessoa tem aqui muita resistência mas não tem muita altura” (Margarida), como em tendências

mais globais dos dados, “se o peso aumenta, também a altura aumenta... não contabilizando as exceções que vimos” (Afonso).

*Pensar sobre os dados.* Os alunos tiveram facilidade em verbalizar as suas conjecturas sobre a relação entre duas variáveis mas as conjecturas que formulam, inicialmente, tiveram por base o seu conhecimento pessoal e a sua experiência do contexto: “Hum... não sei. Talvez seja por causa da prática... esse aluno deve fazer mais corridas e aí desenvolver a velocidade” (Leo). Na sua maioria, os alunos também reconheceram o uso apropriado das diferentes representações gráficas construídas e foram capazes de descrever e interpretar os diagramas de dispersão que construíram, focando-se tanto em características individuais como em tendências mais globais dos dados, como descrito por Ben-Zvi e Arcavi (2001). Embora tenham começado por interpretar os gráficos focando-se em alguns dados individuais para avaliar e suportar as suas inferências, numa perspetiva local, quando incentivados, estes alunos conseguiram desenvolver uma perspetiva mais global para descrever a relação entre duas variáveis, evidenciada nos diagramas de dispersão, incluindo a consideração de dispersão e a identificação de tendências nos dados. Contudo, nenhum dos alunos chegou a mencionar a variação dos dados de modo a indicar adequadamente covariação.

Outros aspetos que emergem deste estudo, cujos resultados mais alargados podem ser encontrados em Henriques e Antunes (2014), permitiram compreender como os alunos usam o *TinkerPlots*, informando os professores sobre o modo como devem intervir no processo de ensino e aprendizagem da covariação.

Um dos aspetos a salientar é que os alunos tiveram facilidade em verbalizar as suas conjecturas sobre a relação entre duas variáveis mas basearam-nas maioritariamente no seu conhecimento e experiência com o contexto dos dados. Apesar do conhecimento do contexto os ter auxiliado a articular o seu raciocínio sobre covariação e a suportar o seu pensamento sobre a relação apresentada nos diagramas de dispersão, quando ela não se ajustava às suas expectativas os alunos valorizaram a informação que obtiveram dos gráficos e usaram-na para validar as suas afirmações e não tanto para explorarem relações entre variáveis. As possíveis causas de variação que procuraram e as inferências que fizeram baseadas nos dados também estão fortemente relacionadas com o seu conhecimento do

contexto dos dados. Deste modo, os alunos foram capazes de fazer conexões com o contexto dos dados, foram influenciados por ele e usaram-no, algumas vezes, de modo significativo (Fitzallen & Watson, 2011). No entanto, é preciso que os alunos se distanciem do contexto para procurarem relações entre as variáveis e, nesse sentido, Fitzallen (2012) sugere a necessidade de selecionar dados a partir de contextos menos familiares aos alunos, desafiando-os a ‘interrogar’ os dados e a tirar conclusões a partir das representações gráficas criadas.

Outro aspeto, remete para a interação com o *software*, uma vez que uma aprendizagem efetiva não decorre só da sua utilização, mas sim da forma como é usado continuamente em sala de aula e trabalhadas as suas potencialidades na promoção do raciocínio covariacional dos alunos. Assim, o professor tem que ter um papel ativo no estabelecimento e suporte do desenvolvimento do raciocínio dos alunos (Cobb, et al., 2003).

Estes resultados permitiram, ainda, informar e reformular algumas questões, no enunciado da tarefa, de modo a evitar respostas simples e imediatas e a promover a construção de diversas representações gráficas e a assegurar a sua exploração e de outras medidas estatísticas para apoiar a formulação e justificação de conjecturas sobre a covariação (Fitzallen, 2012).

Os resultados deste estudo exploratório permitiram, assim, perceber quais as aprendizagens e dificuldades mais importantes dos alunos que serão aprofundadas no estudo principal.

### **3.3. CARATERIZAÇÃO DO CONTEXTO ESCOLAR E DOS PARTICIPANTES DO ESTUDO**

A investigação desenvolvida decorreu no ano letivo 2013/14, num estabelecimento de ensino particular e cooperativo da região do interior centro do país, com 2.º e 3.º ciclo e Secundário. A escola conta com aproximadamente 400 alunos distribuídos por 21 turmas que abrangem desde o 5.º ao 12.º ano, Cursos Profissionais e Cursos de Educação e Formação para Jovens (CEF).

Os participantes neste estudo são todos os alunos de uma turma dos cursos profissionais de Técnico de Auxiliar de Saúde (TAS), que conta com 11 alunos e de Técnico de Apoio à Gestão Desportiva (TAGD), igualmente com 11 alunos. Dos 22 alunos que constituem a turma, 8 são do sexo feminino e 14 são do sexo masculino e têm idades compreendidas entre 15 e 19 anos. De referir, ainda, que 11 alunos (50%) já tiveram alguma retenção ao longo do seu percurso escolar, sendo que três desses alunos, reprovaram duas vezes durante o seu percurso escolar. Além disso 5 dos alunos são provenientes do CEF (curso de educação e Formação) e outros três são referenciados com necessidades educativas especiais.

As razões frequentemente apontadas, pela docente da disciplina e Conselho de Turma, para as dificuldades evidenciadas são: a falta de atenção e motivação, a falta de estudo, o facto de não gostar do tópico abordado e da própria Matemática bem como a complexidade da mesma. Nenhum dos alunos utiliza, habitualmente, o computador ou a máquina de calcular no estudo da Matemática, a não ser para realizar cálculos.

De salientar, ainda, que alguns alunos da turma têm um comportamento desadequado na sala de aula, são conflituosos, agitados e conversadores, o que se tem refletido no aproveitamento geral da turma. Estes alunos também manifestam falta de assiduidade, problemas de atenção e concentração, o que conduz a uma desmotivação fácil e pouco empenhamento nas atividades escolares. No entanto, os restantes elementos da turma mostram-se interessados e responsáveis, com bom aproveitamento em todas as áreas curriculares.

A maioria dos alunos sente dificuldades em argumentar e comunicar as suas ideias aos colegas e ao professor, embora esse grau de dificuldade dependa dos tópicos em estudo.

De uma forma geral, os pais não se deslocam regularmente à escola e só o fazem quando convocados. A origem socioeconómica dos alunos é diversa, sendo que, na sua maioria, são provenientes de agregados familiares com dificuldades económicas. As profissões dos encarregados de educação são, também, bastante variadas e muitos deles desempregados, revelando uma grande assimetria nas suas habilitações literárias e contexto cultural.

### 3.4. RECOLHA E ANÁLISE DE DADOS

Antes de iniciar a unidade de ensino e o respetivo processo de recolha de dados procedi às diligências necessárias para obter as autorizações para a realização do estudo. Todos os encarregados de educação autorizaram a participação dos seus educandos no estudo e a captação de som e imagens nas aulas.

A recolha de dados seguiu as orientações de Yin (2003) que remete para a importância da recolha de dados não ter uma única fonte de evidência, pois deve-se dispor de um leque alargado de fontes de informação. Assim, para assegurar a validade deste estudo foi previsto o recurso a vários métodos de recolha de dados, tais como a observação participante (através de gravações áudio e vídeo das aulas, registo das ações dos alunos no ecrã do computador pelo *software AutoScreenRecorder 3.1 Pro* e elaboração de notas de campo) e recolha documental (com base nas resoluções escritas dos alunos das tarefas e de uma prova final do módulo A3 – Estatística e nos ficheiros do *TinkerPlots*). A utilização destas fontes de recolha de dados permite articular os dados obtidos, contribuindo para garantir a credibilidade do estudo, minimizando a influência do observador (Stake, 2007).

A observação é um método de recolha de dados fundamental em qualquer estudo de carácter interpretativo, pois tem a finalidade de compreender as aprendizagens, bem como as dificuldades que os alunos manifestam durante a realização das tarefas. É através da observação que o investigador acede às perspetivas dos participantes e entende o que motivou as reações observadas e o seu significado (Foster, 1996; Gay, 1986). Como defende Merriam (1988, p. 89):

A observação é melhor técnica para utilizar quando uma atividade, acontecimento ou situação podem ser observados em primeira mão, quando uma perspetiva nova é desejável, ou quando os participantes não são capazes ou não estão dispostos a discutir o tópico em estudo.

A escolha do tipo de observação a efetuar, deve estar de acordo com o tipo de investigação, segundo Tuckman (2000), o qual refere que na investigação qualitativa, a observação visa examinar o ambiente de sala de aula, através de um esquema geral para nos orientar e o seu produto é registado em notas de campo.

Em sintonia com a afirmação anterior Vale (2000, p. 233), refere que “a observação é a melhor técnica de recolha de dados do indivíduo em atividade, em primeira mão, pois permite comparar aquilo que diz, ou que não diz, com aquilo que faz”. A observação participante é melhor técnica de recolha de dados neste tipo de estudo, de acordo com Bogdan e Biklen (1994), pois caracteriza-se pela integração do investigador no contexto, interagindo continuamente com os alunos, dialogando e apoiando-os no desenvolvimento do trabalho.

O duplo papel de professora/investigadora contribuiu para a compreensão das ações dos alunos, aquando da realização das tarefas, sem influenciar o decorrer normal dos acontecimentos (Tuckman, 2000).

As observações, realizadas em contexto sala de aula, possibilitaram analisar e compreender, com mais detalhe, as ideias e o raciocínio dos alunos participantes. Os dados recolhidos, neste caso, são essencialmente descritivos e o foco está centrado na forma como os alunos interpretam as situações problemáticas, nas estratégias e conhecimentos usados para estabelecer a covariação. Ao participar na experiência em sala de aula, procuro compreender, igualmente, a interação dos alunos com o *software* na resolução das tarefas propostas.

Quando a investigação incide sobre a prática dos alunos, há interesse em analisar os registos do seu trabalho na realização de tarefas com o *Tinkerplots*, através da videografia dos trabalhos no computador com o *software AutoScreenRecorder 3.1 Pro*. As aulas em que decorreu a realização das tarefas, relativas ao tema em estudo, foram gravadas em áudio e vídeo, a fim de captar aspetos que possam passar despercebidos ou até mesmo, para o investigador não ter de se preocupar em simultâneo em apreender todos os detalhes expressos pelo aluno (Bassey, 1999). Este recurso permitiu rever as situações, num número ilimitado de vezes, mantendo intacta a informação recolhida. Para além desse recurso, recorreu-se ao registo de notas de campo.

A observação foi, ainda, complementada com outra técnica de recolha de dados, a recolha de documentos, uma vez que é importante a corroboração e confirmação de evidências sugeridas por outras fontes de dados (Bell, 2004). Com isto, obtém-se mais elementos para compreender os procedimentos realizados, as generalizações e justificações formuladas e as dificuldades que revelam. A recolha de documentos incluiu, por isso, as produções escritas dos alunos e os ficheiros do

*TinkerPlots* com as suas resoluções, nas diversas tarefas ao longo do estudo e na prova final da unidade, por forma a uma melhor compreensão dos seus raciocínios e das suas aprendizagens e dificuldades, no tópicó da covariação.

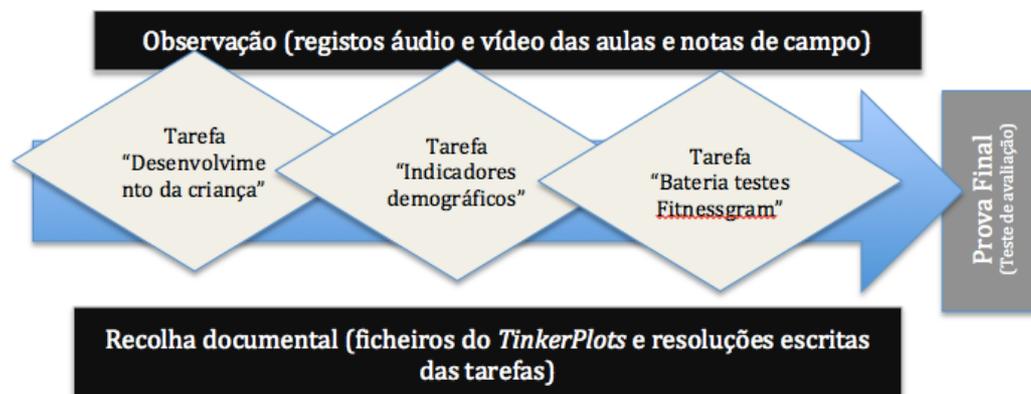


Figura 4 – Momentos de recolha de dados ao longo da unidade de ensino.

A análise de dados é um processo de compreensão e sistematização da informação recolhida, e como afirma Ludke e André (1986) para compreender e sistematizar a informação, devemos trabalhar todo o material obtido durante a pesquisa, organizando, dividindo, relacionando e procurando tendências e padrões relevantes. Neste estudo, a análise dos dados assumiu um carácter interpretativo, com base na análise de conteúdo realizada aos documentos recolhidos e às transcrições parciais dos registos áudio e vídeo do trabalho dos alunos. Na análise do raciocínio covariacional estatístico dos alunos usei as categorias referidas por Fitzallen (2006). A descrição das aprendizagens e dificuldades dos alunos, no que respeita à covariação, reporta-se aos objetivos definidos para o tópicó nas orientações curriculares.

Na descrição, há a necessidade de criar siglas para facilitar a identificação das fontes de dados, durante a sua análise geral. Neste estudo utilizei as siglas representadas no quadro 2.

Quadro 2 . Formas de representar a fonte de dados.

Sigla	Fonte de dados
TG	Trabalho de Grupo Ex: TG1- trabalho de Grupo 1
DC	Discussão de grupo Ex: DC 1 – discussão de Grupo 1

PA	Prova de Avaliação Ex: PA1 – prova de avaliação de um elemento do grupo 1
GTG	Gravação do trabalho de grupo Ex: GTG1 – gravação do trabalho de grupo 1
NC	Notas de campo Ex: NC1 - notas de campo relativas do comentário do grupo 1

## CAPÍTULO 4

### A UNIDADE DE ENSINO

Neste capítulo apresento a unidade de ensino, incluindo os seus objetivos de aprendizagem, a sua planificação e a descrição das tarefas. Incluo, também uma referência à avaliação das aprendizagens dos alunos.

#### 4.1. ASPETOS GERAIS E PLANIFICAÇÃO DA UNIDADE DE ENSINO

A unidade de ensino que serve de base à realização do estudo foi lecionada durante o 3.º período do ano letivo 2013/14, numa turma do 10.º ano dos cursos profissionais de Técnico de Auxiliar de Saúde (TAS) e de Técnico de Apoio à Gestão Desportiva (TAGD), e visa promover a aprendizagem de conceitos e procedimentos no tópico das distribuições bivariadas (Quadro 3), através de uma abordagem exploratória com recurso ao *software TinkerPlots*.

Quadro 3. Objetivos da unidade de ensino – distribuições bidimensionais (ME, 2001a; Me, 200b; ME, 2005)

<b>Tópico</b>	<b>Subtópico</b>	<b>Objetivos</b>
<b>Distribuições bidimensionais</b>	<b>Diagrama de dispersão</b>	<ul style="list-style-type: none"><li>- Formular conjecturas sobre dados representados em tabelas.</li><li>- Construir diagramas de dispersão a partir de dados representados em tabelas.</li><li>- Ler e interpretar diagramas de dispersão.</li><li>- Identificar diferentes tipos de tendência existentes nos dados.</li><li>- Avaliar, de forma intuitiva a partir da análise do diagrama de dispersão, a existência de uma associação linear entre as duas variáveis.</li></ul>

	<b>Correlação</b>	<ul style="list-style-type: none"> <li>- Observar a distribuição dos pontos nos quadrantes do diagrama de dispersão para classificar a correlação entre duas variáveis .</li> <li>- Definição de centro de gravidade de um conjunto finito de pontos e a sua interpretação física.</li> <li>- Calcular e interpretar a intensidade e a direção (positiva ou negativa) da correlação entre duas variáveis, relacionando-a com a linearidade/ dispersão e com a orientação da nuvem de pontos obtida no diagrama de dispersão.</li> </ul>
--	-------------------	---

Tendo por base o programa de Componente de Formação Científica da disciplina de Matemática (ME, 2005), o módulo A3 – Estatística, inclui três tópicos: Estatística-generalidades; Organização e interpretação de caracteres estatísticos (qualitativos e quantitativos); e Referência a distribuições bidimensionais (abordagem gráfica e intuitiva). Este estudo tem por base uma unidade de ensino respeitante ao tópico “Referência a distribuições bidimensionais” e foi planeada (Quadro 4) tendo em conta o plano anual de atividades escolares definido pela escola e o horário atribuído ao módulo A3 – Estatística.

Quadro 4. Planificação da unidade de ensino

Aulas	Subtópico	Tarefa	Metodologia de trabalho
17.04.2014 (60 minutos)	Estatística: generalidades	-----	-Discussão sobre a utilidade na vida moderna
22.04.2014 (60 minutos)	Organização e representação de caracteres estatísticos	“Nenana Ice Classic” (sala de informática)	-Leitura da tarefa, em voz alta -Realização da tarefa em grupo no <i>TinkerPlots</i>
24.04.2014 (120 minutos)			
29.04.2014 (60 minutos)			-Discussão da tarefa em grupo
05.05.2014 (60 minutos)	Organização e representação de dados	Exercícios de aplicação	- Trabalho individual
06.05.2014 (60 minutos)	Amostragem	“Uma experiência com peixes”	- Apresentação dos slides sobre Amostragem - Trabalho de grupo
08.05.2014 (120 minutos)	Medidas de dispersão	“Uma experiência com peixes” (sala de informática)	- Trabalho de grupo - Discussão em grupo
13.05.2014 (60 minutos)	Medidas de	“Desenvolvimento	- Leitura do problema - Discussão do tema

	localização central	da criança” (sala de informática)	- Trabalho de grupo
15.03.2014 (120 minutos)	Medidas de dispersão		- Trabalho de grupo
19.05.2014 (60 minutos)	Diagrama de dispersão		- Discussão coletiva
20.05.2014 (60 minutos)	Medidas de localização central Medidas de dispersão	Exercícios de aplicação	-Resolução individual -Correção participativa e explicativa pela turma
22.05.2014 (120 minutos)	Diagrama de dispersão	“Indicadores Demográficos” (sala de informática)	-Leitura do problema -Discussão do tema -Trabalho de grupo
27.05.2014 (60 minutos)	Centro de gravidade Covariação		-Trabalho de grupo -Discussão coletiva
29.05.2014 (120 minutos)	Coeficiente de correlação e sua variação no intervalo Reta de regressão	Exercícios de aplicação	-Resolução individual -Correção participativa e explicativa pela turma
02.06.2014 (60 minutos)			
03.06.2014 (60 minutos)	Diagrama de dispersão Dependência estatística e correlação positiva e negativa	“Bateria de testes de Fitnessgram” (sala de informática)	-Leitura do problema -Discussão de grupo sobre o tema -Trabalho de grupo
05.06.2014 (120 minutos)	Definição de centro de gravidade de um finito de pontos e a sua interpretação física		-Trabalho de grupo Elaboração de relatório sobre os dados analisados no <i>software</i>
09.06.2014 (60 minutos)	Interpretação e limitações da reta de regressão		-Discussão de grupo
10.06.2014 (60 minutos)	Distribuições bidimensionais Organização e interpretação de caracteres estatísticos	Exercícios de consolidação	Resolução individual -Correção participativa e explicativa pela turma (revisões para a prova)
12.06.2014 (120 minutos)		Prova de Avaliação (Escrita)	-Realização individual da prova -Correção da prova em grupo

Esta abordagem exigiu algumas alterações na dinâmica das aulas e no papel do professor, bem como a introdução de tarefas de natureza diversa dos exercícios rotineiros habitualmente usados. Procurei proporcionar um ambiente de aprendizagem rico, em sede de sala de aula, estimulando os alunos a

empenharem-se na realização de uma sequência de tarefas envolvendo situações reais e com contextos significativos para os alunos e que promovem, simultaneamente, o desenvolvimento de conceitos e ideias estatísticos (Lesh & Kelly, 2000; Steffe & Thompson, 2000).

Uma parte das aulas foi utilizada para a realização de 3 tarefas de exploração relacionadas com o tópico das distribuições bidimensionais (abordagem gráfica e intuitiva).

A realização de cada uma destas tarefas envolveu três fases em sala de aula: (i) a introdução da tarefa (apropriação da tarefa, para sala realização), (ii) a exploração da tarefa (trabalho autónomo dos grupos), (iii) a discussão de resultados e sistematização das aprendizagens (Stein, Engle, Smith & Hughes, 2008; Tudella, et al., 1999).

Na primeira fase, a professora apresentou a tarefa à turma, lendo-a, após ter distribuídos os exemplares escritos, com o objetivo assegurar a compreensão do seu propósito, bem como estimular a participação e o empenho dos alunos na realização da mesma.

Na exploração das tarefas predominou o trabalho de grupo. Os grupos foram constituídos pelas professoras (investigadora e a professora da disciplina) para garantir a heterogeneidade a nível das aprendizagens, integrando 2 a 3 elementos e mantiveram-se inalterados ao longo de todo o módulo. Em situações originadas pela ausência de alguns alunos, os elementos do grupo trabalharam sozinhos. Pelo facto dos alunos revelarem poucos hábitos de trabalho, usa-se o trabalho de grupo a fim de proporcionar aos alunos um ambiente estimulante de partilha e discussão. Desta forma, na resolução das tarefas, os alunos têm a oportunidade de expor as suas ideias, ouvir as dos seus colegas, colocar questões uns aos outros, discutir e comparar as suas estratégias e resoluções e ainda, desenvolver os seus próprios argumentos. Além disso, neste ambiente de partilha e apoio, os alunos podem superar mais facilmente as suas dificuldades. Segundo Ponte, Costa, Lopes, Moreirinha & Salvado (1997) os “alunos podem assim participar em dois níveis do discurso da aula – o coletivo e o que desenvolvem com o seu parceiro de aprendizagem” (p. 94). Para não originar perda de tempo com a organização da sala, os alunos estavam sentados de acordo com a formação dos grupos, quer em sala de aula, quer na sala de informática.

O professor deve também organizar o desenvolvimento do trabalho pela turma, determinando tempo a que se deve dedicar às diferentes fases e orientar os recursos a usar e definindo os modos de trabalho dos alunos (Anghileri, 2006). Nesta fase, o professor deve ter em consideração dois pontos importantes: a promoção da aprendizagem da covariação estatística e a gestão da aula. Para garantir o desenvolvimento da tarefa pelos alunos, deve colocar questões, dar pistas, sugerir representações, pedir clarificações e justificações e manter o desafio cognitivo e a autonomia dos alunos. Quanto à gestão da aula, o professor necessita regular as interações entre alunos, pedir os registos escritos, dar tempo aos alunos para preparar a apresentação e para organizar a discussão a fazer, deve identificar e selecionar resoluções variadas (com erro a explorar, mais ou menos completas, mas com representações relevantes) e sequenciar as resoluções selecionadas. Esta segunda fase exigiu o apoio aos alunos, no respetivo trabalho autónomo sobre a tarefa, de forma a garantir a participação produtiva de todos os elementos do grupo. É uma das fases que se revela extremamente desafiante para o trabalho do professor, uma vez que envolve o esclarecimento de dúvidas por parte dos alunos, sem que esse esclarecimento seja demasiado revelador da estratégia mais adequada ou que reduza o nível de exigência cognitiva da tarefa ou até mesmo, que os possa direcionar para uma uniformização de resoluções, prejudicando a posterior discussão coletiva dos diferentes grupos e as suas aprendizagens. Requer, também, a partir da rápida observação e apreciação das produções dos alunos, conseguir selecionar as resoluções que avalia como contribuições positivas para a discussão coletiva e estabelecer a sequência da apresentação pelos alunos (Stein, et al., 2008).

A terceira fase tem como propósito dar a conhecer à turma o trabalho realizado pelos alunos e a sua discussão, tendo como foco as representações gráficas construídas pelos alunos, as conjeturas formuladas sobre os dados representados e a sua avaliação, no sentido de estabelecer a existência ou não de uma associação linear entre as variáveis estatísticas em estudo. Os alunos apresentam o seu trabalho aos colegas, oralmente e recorrendo ao *software* (projetado no quadro interativo). Segundo Bishop e Goffree (1986), é uma ocasião apropriada para serem expostas conexões e significados, já que é um bom momento de reflexão para os alunos realizarem conexões com outras ideias e tópicos. Esta fase revelou-

se uma das grandes mais-valias da metodologia de trabalho adotada, por ser um caminho propício à troca de ideias tendo sido conseguido, na maioria das aulas, um clima de aprendizagem que proporcionou aprendizagens significativas sobre a covariação, através da comunicação estabelecida entre os alunos e entre alunos e professora. Nesta discussão, o professor não só tem que gerir as intervenções e interações dos diferentes grupos, mas também promover a qualidade nas verbalizações das justificações das suas conjeturas sobre a relação entre os dados (Ruthven, Hoffman & Mercer, 2011). Para além disso, nessas discussões deve-se manter um clima positivo e de genuíno interesse na discussão, garantindo a participação de todos os alunos. Pois é importante que a discussão tenha mais do que o objetivo de comparação e o confronto das resoluções dos alunos mas contribua para estes realizem novas aprendizagens relevantes. Por fim, a sistematização das aprendizagens mais relevantes e dos conceitos envolvidos nas tarefas propostas, também ela um momento de grande importância, favorável à apreensão e consolidação de novos conhecimentos.

Contudo, esta metodologia trouxe também algumas dificuldades, no que respeita à gestão do tempo durante o trabalho autónomo dos alunos com o *software* e na dinamização das apresentações coletivas. Para além disso, houve grande resistência em iniciar de forma produtiva o trabalho autónomo, por parte de alguns grupos, sendo necessário ajudá-los a ganhar ritmo e contagiá-los com entusiasmo, focando-os na tarefa que tinha que realizar. Outros grupos também tendiam a demorar-se no trabalho autónomo, dispersando-se em questões alheias à tarefa.

Atendendo à importância de diversificar tarefas na aprendizagem (Ponte, 2005) e visando a consolidação de conhecimentos adquiridos, as tarefas de natureza mais exploratória foram realizadas em alternância com outras aulas que contemplaram momentos de exposição teórica dos conteúdos estatísticos e tarefas de estrutura fechada (Ponte, 2005), nomeadamente exercícios e problemas do manual adotado.

A calendarização das tarefas propostas aos alunos teve em conta a articulação com os outros temas estatísticos, a serem lecionados, neste módulo, a realização das tarefas ocorreu no período letivo entre o dia 22 de abril e 09 de junho de dois mil e catorze, num total de 17 blocos de sessenta minutos, que contemplaram

igualmente a realização de atividades de consolidação, como referido, e por último a realização de uma prova de avaliação de conhecimentos (quadro 4).

## **4.2. Tarefas e recursos**

A seleção das tarefas a propor aos alunos é um dos aspetos essenciais do trabalho de qualquer professor. Devem estar articuladas entre si, em sequências coerentes, por forma a promover as aprendizagens (Ponte & Serrazina, 2009). Tendo em conta os objetivos gerais do Programa do Ensino secundário (ME, 2001a, 2001b), o programa de componente de Formação Científica da disciplina de Matemática (Martins, et al., 2005) e as orientações curriculares internacionais preconizadas para o ensino da Estatística, concebi uma sequência de tarefas (quadro 4) centradas no tópico das *distribuições bidimensionais*.

As tarefas aplicadas durante a unidade de ensino foram pensadas para serem realizadas em grupo e tiveram como objetivo geral promover nos alunos a aquisição de conhecimentos e procedimentos estatísticos com significado e desenvolver o seu raciocínio estatístico e a compreensão da covariação estatística. Bem, como de acordo com as possibilidades de organização e representação dos dados em gráficos, que o *software TinkerPlots* potencializa (Konold & Miller, 2005).

Procurei que as tarefas fossem diversificadas, na sua natureza, com o intuito de “no seu conjunto, proporcionem um percurso de aprendizagem coerente” (ME, 2007, p.11). Segundo a ideia de Carvalho (2009), a escolha das tarefas por parte do professor e o modo como os alunos se envolvem na sua resolução, é bastante importante para a qualidade dos seus desempenhos. Por isso, as tarefas elaboradas foram de encontro aos interesses dos alunos, retratando temas próprios da área de formação, ou seja, na área de saúde e na área do desporto. Como afirmam Martins e Ponte (2010), as tarefas estatísticas ao utilizar dados reais, motiva os alunos, envolvendo-os na aprendizagem. Para além disso, os contextos selecionados para as tarefas foram o contexto da realidade, com o intuito de proporcionar aos alunos uma visão acerca da aplicabilidade do raciocínio covariacional no dia-a-dia, alertando para os problemas da atualidade:

aquecimento global, vantagens de uma aquacultura, a evolução de um país, crescimento físico de crianças e avaliação da aptidão física.

Konold e Higgins (2003) defendem que os alunos aprendem mais se utilizarem dados relevantes, nas questões estatísticas que construíram e na análise de dados em estudo, sem trivializarem a questão inicial e assumir os dados produzidos como abstrações do contexto que os gerou sem esquecer que se trata de “números num contexto”.

As questões formuladas nas tarefas para a análise de dados devem: apoiar-se na vertente visual da representação e em técnicas estatísticas, assumindo uma atitude de “dar sentido” ao conjunto de dados numa perspectiva global; saber a razão pela qual um tipo específico de tabela, gráfico ou medida estatística se adequa a um certo tipo de dados; compreender porque as medidas centrais e de dispersão fornecem informação sobre um conjunto de dados e que as medidas estatísticas centrais e de dispersão são úteis para comparar conjuntos de dados (Garfield, 2002). Assim, os alunos poderão recorrer a diferentes representações, procedendo a comparações e procurar extrair o máximo de informação possível para gerar modelos a partir do conjunto, identificando assim importantes regularidades, desvios e variações dos dados. Além disso, procurei que as questões de interpretação de gráficos não incidissem exclusivamente a níveis cognitivos baixos, como sugere Batanero (2001).

Ao longo desta unidade de ensino, as tarefas foram sequenciadas pelo seu grau de dificuldade crescente e pelos vários tipos de representações gráficas requeridas, para que os alunos ficassem familiarizadas com elas, procurando assim, contribuir para o desenvolvimento do raciocínio covariacional.

A sequência de tarefas elaboradas tinham como objetivo contribuir para o desenvolvimento progressivo do raciocínio estatístico covariacional dos alunos.

Cada tarefa tinha como objetivo: (1) explorar, analisar, interpretar e utilizar informação de natureza estatística; (2) construir gráficos a partir de dados representados em tabelas; (3) interpretar os resultados de estudos associados à área de formação do curso e formular conjeturas a partir deles, usando a linguagem estatística para demonstrar as possíveis dependências.

Neste estudo foram propostas três tarefas de natureza exploratória: “Desenvolvimento da criança”, “Indicadores demográficos” e “Bateria de testes Fitnessgram”. Contudo, foram aplicadas mais duas – “Nenana Ice Classic” e “Uma experiência com peixes” – que foram usadas, inicialmente, para a familiarização dos alunos com as ferramentas do software e com o tipo de tarefas a realizar.

Todas as tarefas começaram pela sua apresentação e só era esclarecido o enunciado, quando a maioria dos alunos não entendia o que era pedido.

As três tarefas seguintes serão o foco da análise de estudo.

A tarefa 3, “Desenvolvimento da criança”, tem como objetivo estudar a população não só numa variável isolada, mas sim, em duas ou mais variáveis, podendo estar ou não relacionadas. Pretende-se comparar as dispersões dos pontos a partir do mesmo referencial ou em referenciais diferentes, considerando as diferentes medidas de localização e de dispersão, para estabelecer relações entre as variáveis. Nas representações gráficas construídas, os alunos têm que interpretar, comparar e explicar as eventuais relações que possam existir nas distribuições, recorrendo à dispersão dos dados entre as variáveis. O tema da tarefa foca o crescimento físico de rapazes e raparigas da Califórnia, desde o seu nascimento até atingir os 18 anos. Com base nos dados inseridos no ficheiro do software, os alunos tem que caracterizar essas crianças relativamente ao seu crescimento físico, identificando a possibilidade ou não, de existir associação estatística entre os valores das variáveis. Esta comparação dos dados possibilita aos alunos, a reflexão de quais os melhores meios que devem utilizar para tomar decisões e ponderar o grau de incerteza associado a previsões sobre o crescimento físico das crianças. Esta tarefa é constituída por duas partes. A parte I, os alunos exploram os dados no ficheiro, partindo das questões formuladas por eles. A parte II, os alunos recorrem aos dados para responderem às questões.

A tarefa 4, “Indicadores demográficos”, tem como objetivo construir diagramas de dispersão a partir de dados representados em tabelas, para verificar se existe alguma relação entre as variáveis em estudo, observar a distribuição dos pontos nos quadrantes do diagrama de dispersão para classificar a associação estatística entre duas variáveis e interpretar o centro de gravidade de um conjunto finito de pontos.

Nas questões da tarefa, destaca-se a aplicação das medidas de localização, que medem a variabilidade dos dados. Pressupõe-se que os alunos afirmem que os dados das variáveis possam estar ligados através de uma relação funcional ou de uma relação estatística, ou não estão ligados através de qualquer relação. Aqui, supõe-se que os alunos entendam que a relação estatística exprime o meio termo no grau de conexão entre variáveis, a relação funcional (termina com precisão desejada que um dado parte de outro dado) e a ausência de relação. Os alunos têm que avaliar a relação das variáveis quanto ao sentido, intensidade e linearidade nas representações gráficas construídas.

O tema da tarefa foca para as possíveis relações ou associações entre os vários indicadores demográficos utilizados (taxa de natalidade, taxa de mortalidade, índice de fecundidade, idade média do homem no 1.º casamento, a idade média da mulher no 1.º casamento, taxa bruta de nupcialidade, taxa de mortalidade infantil), para influenciam o crescimento do país.

Esta tarefa é constituída por duas partes. A parte I, os alunos recorrem aos dados para responderem às questões. A parte II, os alunos devem elaborar um pequeno texto com as conclusões que permita responder ao problema do estudo e explicar as evidências em que basearam.

A tarefa 5, “Bateria de testes Fitnessgram”, tem como objetivo construir diagramas de dispersão, como modo eficaz de visualizar e compreender a associação entre duas variáveis e formular conjecturas sobre os dados representados em tabelas e avaliar, de forma intuitiva a existência de uma associação linear entre as variáveis. Nas questões das tarefas, os alunos devem associar a linearidade/dispersão e orientação dos *plots* para interpretar a intensidade e a direção (positiva ou negativa) da associação estatística entre duas variáveis. Os alunos devem descrever os dados, retirar conclusões e fazer inferências sobre as relações identificadas nos diagramas de dispersão. O tema da tarefa refere-se aos resultados de testes de condição física inicial dos alunos de duas turmas do 9.º ano de escolaridade: peso, altura, resistência, diversos tipos de forças (superior, abdominal e inferior), flexibilidade, impulsão vertical, IMC e sexo.

Esta tarefa é constituída por três partes. A parte I, os alunos exploram os dados no ficheiro, partindo das questões formuladas por eles. A parte II, os alunos recorrem aos dados para responderem às questões. A parte III, os alunos devem elaborar

um pequeno texto com as conclusões que permita responder ao problema do estudo e explicar as evidências em que basearam.

### **4.3. A AVALIAÇÃO DAS APRENDIZAGENS**

No NCTM (2008) pode ler-se que “a avaliação deve apoiar a aprendizagem de uma Matemática relevante e fornecer informações úteis quer para os professores quer para os alunos”(p. 23). Neste sentido, a avaliação deverá ser um processo sistemático e contínuo, com o intuito de melhorar a formação dos alunos (NCTM, 2008).

Durante a realização das tarefas nas aulas pude observar o envolvimento e empenho de cada aluno, bem como os seus progressos na aprendizagem, através da análise dos seus registos escritos e dos ficheiros com a gravação das suas ações realizadas no computador. Esta observação permitiu-me também identificar as dificuldades dos alunos e reavaliar a estrutura das restantes tarefas. Nas discussões coletivas em turma, houve a oportunidade de avaliar a capacidade de cada aluno na expressão oral, nomeadamente na argumentação e na explicação dos seus raciocínios e estratégias usadas na resolução da tarefa.

No final da unidade de ensino, foi realizada uma prova escrita individual, que permitiu avaliar as aprendizagens dos alunos no que respeita à compreensão do conceito de covariação.

Para concluir o módulo, fiz uma prova escrita, com o objetivo de avaliar as aprendizagens realizadas ao longo da unidade de ensino. A prova final é constituída por 6 questões, e estas, por sua vez, com vários itens. Na maioria das questões, são inquiridas sobre a associação estatística entre as variáveis, por forma a verificar a tendência e interpretar a existência de uma associação linear entre as variáveis e relacionar duas variáveis através de uma reta.

No final do módulo, a prestação dos alunos foi satisfatória, perante as características da turma a nível de resultados à disciplina de Matemática. Existe ainda alguns alunos (22,7% dos alunos) que ainda não conseguem perceber o

conceito de correlação e associar o grau de correlação com a distribuição dos pontos nos diagramas de dispersão.

Além dos problemas na interpretação das tabelas e gráficos (9% dos alunos), há alunos que demonstram lacunas na comunicação escrita como é visível em algumas das questões dadas e são, muitas vezes, pouco claras ou incertas, ou que não têm sentido.

No global, todos os alunos conseguiram realizar o módulo, tendo-se registado 9% de alunos com 10 valores, 14% de alunos com 11 valores, 36% dos alunos com 12 valores, 23% dos alunos com 14 valores e 18% dos alunos com 16 valores.

## **CAPÍTULO 5**

### **RACIOCÍNIO COVARIACIONAL DOS ALUNOS**

Neste capítulo, foco-me no raciocínio covariacional dos alunos participantes neste estudo, apresentando uma análise do seu trabalho nas três tarefas propostas na unidade de ensino, orientadas para a aprendizagem da covariação estatística e no teste final de avaliação (Anexo 12).

Para cada uma das tarefas, começo por fazer um breve enquadramento e síntese do trabalho realizado pelos alunos em sala de aula e depois apresento a análise do raciocínio dos alunos, estruturada nas quatro dimensões do quadro conceptual de Fitzallen (2006) utilizado: (i) conhecimento genérico, (ii) à criatividade com os dados, (iii) à compreensão dos dados e (iv) ao pensamento sobre dados, relativamente à covariação.

A análise é suportada com excertos do trabalho dos alunos que permitem evidenciar uma variedade de aspetos interessantes do seu raciocínio covariacional bem como as aprendizagens e dificuldades que evidenciaram neste tópico.

#### **5.1. TAREFA – “DESENVOLVIMENTO DA CRIANÇA”**

##### **ENQUADRAMENTO**

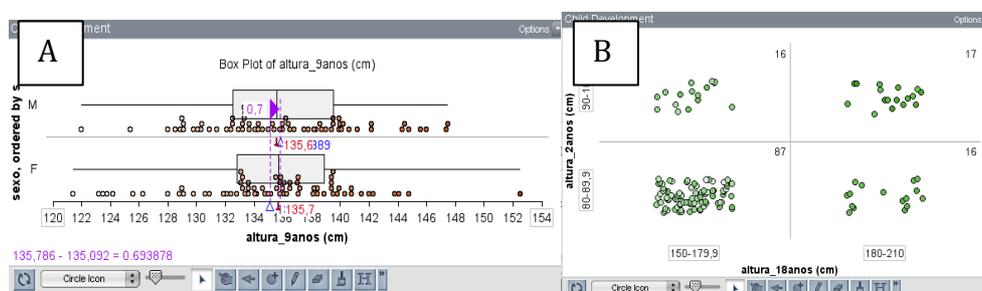
Esta tarefa tinha como objetivo comparar as dispersões dos pontos a partir do mesmo referencial ou em referenciais diferentes, considerando as diferentes medidas de localização e de dispersão, para estabelecer relações entre as variáveis. Desta forma, pretende-se introduzir a covariação, de modo informal, usando uma

abordagem gráfica. Espera-se que os alunos interpretem a informação apresentada de diversas formas, compreendam qual a representação gráfica que melhor evidencia a relação entre duas variáveis e interpretem diagramas de dispersão, através da análise da forma e da direção da associação entre duas variáveis.

A tarefa está estruturada em duas partes e foi realizada em três aulas. Na primeira aula (60 minutos), introduzi o tema e realizei uma discussão em grande grupo focada na questão geral da tarefa; o trabalho autónomo, em grupo, na exploração da tarefa, com base em questões focadas no conceito de covariação, foi realizado na segunda aula (120 minutos); a discussão em grande grupo, sobre o trabalho realizado na tarefa ocupou a terceira aula (60 minutos). Os alunos demoraram mais tempo do que o previsto na sua exploração e optei por deixar que a finalizassem, fazendo a discussão geral na aula seguinte. As resoluções dos alunos da tarefa foram recolhidas no final da segunda aula, digitalizada e devolvida no início da aula seguinte, para que os alunos pudessem acompanhar a discussão em grande grupo e efetuar a sua correção.

## RACIOCÍNIO COVARIACIONAL

*Conhecimento genérico.* Nesta tarefa, os alunos acederam, de modo autónomo, às ferramentas do *TinkerPlots*. Uma vez que esta é a terceira tarefa que realizam usando este *software*, os alunos construíram com facilidade diversas representações gráficas.



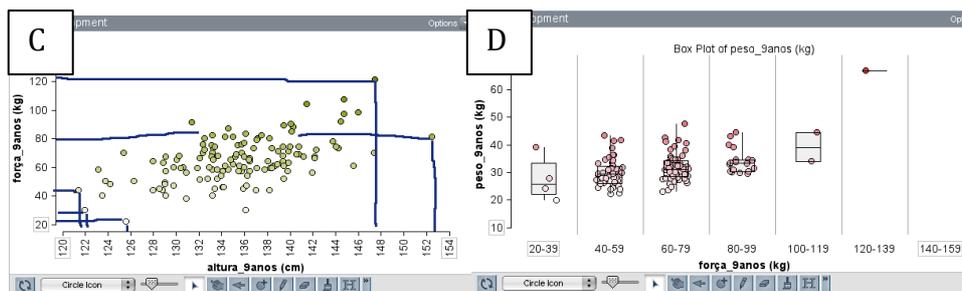


Figura 5.1. 1 – Representações gráficas de vários grupos (G2 – A e G3 – B, C e D).

O grupo 3 mostrou dificuldade em reconhecer as características dos gráficos, tendo começado por construir vários gráficos mas sem uma intenção clara de obter evidência a partir dos dados. Quando questionados para justificar a existência ou não de uma relação entre o peso e a força, os alunos deste grupo, apesar de não saberem que tipo de representação tinham que fazer, iniciaram por representar uma variável, numa escala segmentada em intervalos e depois estenderam-na para contínua, inserindo outra variável, do mesmo modo, até serem capazes de identificar o melhor gráfico que permitia demonstrar a associação estatística. Esta exploração foi facilitada pela dinâmica do software, que lhes permitiu criarem diversas representações com rapidez e pelo conhecimento que os alunos tinham das suas ferramentas. O diálogo seguinte ilustra a interação dos alunos do grupo 3, durante o trabalho com o *software*, na tentativa de obterem evidência para responder às questões. No diálogo seguinte, um dos elementos do grupo chamou a atenção para a característica do gráfico, reconhecendo que precisa de uma outra variável, entre a força e a altura:

Aluno 2: Espera... [pede ao colega que ia fazendo um gráfico com uma só variável] assim está mal. Não os podes fazer em gráficos diferentes. Assim, não dá para comparar!

Aluno 1: Eu conseguia. Vê. Se puxar a bola e colocar em intervalos, dá para saber quantas crianças estão nesses valores. Vê. A criança 130 é a única que mantém a lógica de ser mais alta e ter a mesma força. [Olhava para o gráfico com a altura aos 9 anos e para o outro gráfico, com o peso aos 9 anos e clicava no *plot* para ver a posição dos dados]

Aluno 2: (...) Nada disso, assim não vamos fazer. Primeiro temos que utilizar uma linha com valores contínuos e não em intervalos. Estás a ver [arrasta um *plot* para a direita e transforma num eixo contínuo]! E vamos só fazer dois gráficos com duas variáveis com os valores em linha e cada gráfico de acordo com as duas idades que temos [figura 5.1.1, gráfico C].

Nestes diálogos verifica-se também o uso de linguagem incorreta, como por exemplo, as expressões “puxar a bola”, “valores em linha”, “transforma num eixo”. Usam linguagem informal ou expressões inapropriadas para descrever os componentes dos gráficos ou a informação representada: “ao pegar nesta bola para cima, as outras da outra cor ficam em baixo” [TG1]. Pouco foram os grupos em que não se verificou essa dificuldade em utilizar a linguagem dos dados e dos gráficos.

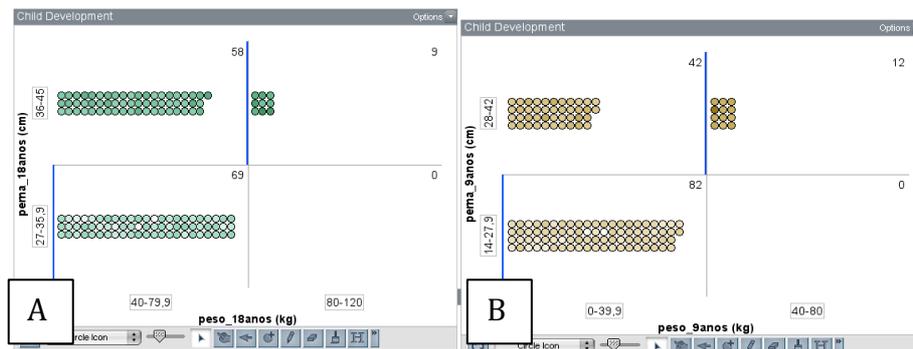


Figura 5.1. 2 – Gráfico de covariação para o peso e perna [TG6].

Os alunos do grupo 1, 5 e 6, ao criar gráficos com escala segmentada em intervalos (figura 5.1.2) para comentar a afirmação: “as crianças com maior largura na perna tendem a ser mais pesadas”, reconheceram que a afirmação é verdadeira, mas expressaram-se através de uma linguagem informal (e frequentemente inapropriada) quando explicaram como pensaram: “há um quadrado [quadrante 2], o de baixo do lado direito que não tem nenhuma bola e o outro quadrado de baixo, ao crescer [o peso] diminuiu o número de bolas” [TG 1] e “ao crescer, há sete bolas que continuam a ser [corresponder a] as mais pesadas e com perna mais larga do grupo” [TG6].

Os grupos 8 e 10 não mostram dificuldade em expressarem-se sobre termos e gráficos estatísticos, dando respostas corretas a questões envolvendo ideias chave, como variável, tendência, associação, relação e linearidade. Estes grupos compreenderam a linguagem dos dados e dos gráficos, pois não precisaram de esclarecimentos sobre os termos usados no enunciado ou sobre o uso de *TinkerPlots*.

Nas representações gráficas construídas, os grupos recorreram a algumas ferramentas do *TinkerPlots* (botão de contagens e frequências, linhas de referência, ferramenta de desenho e cores), para justificarem as relações entre variáveis. Os grupos 3, 6, 8 e 10 compreenderam como deviam usar algumas características do software: “com estas duas variáveis não podemos recorrer ao *icon* Box Plot, porque estão em linhas contínuas e ele não aparece. Temos que colocar um em intervalos e a outra em linha contínua”[TG3].

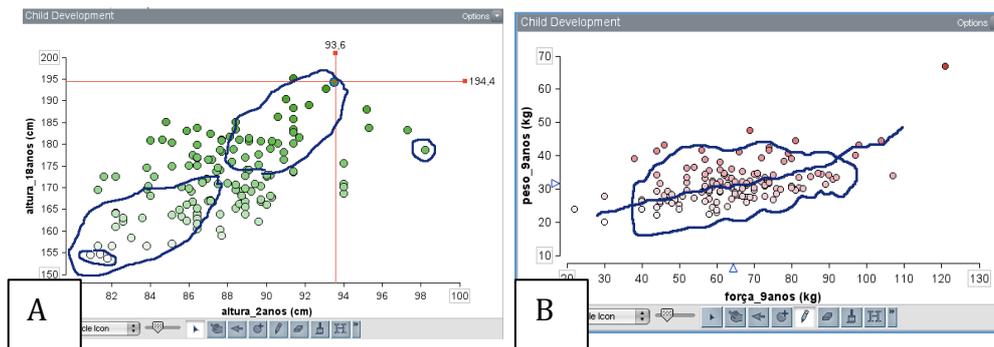


Figura 5.1. 3 – Gráficos de covariação entre: A- “altura aos 2 anos” e “altura aos 18 anos” [TG8] e B – “peso” e “força aos 9 anos” [TG10].

Na figura 5.1.3, o grupo 8 usou a ferramenta de desenho para destacar zonas na representação criada, facilitando a sua explicação para a associação que assume existir entre as variáveis.

Outros grupos perceberam que a melhor forma para estudar a relação entre duas variáveis é começar por construir a nuvem de pontos ou diagrama de dispersão. Usaram intencionalmente esta representação para comentar a afirmação: “as crianças de 9 anos com maior largura na perna tendem a ser as mais pesadas” [NC], reconhecendo as características deste tipo de gráfico. Os grupos reconheceram que as características do diagrama de dispersão permitia afirmar a existência de covariação, de acordo com a distribuição dos dados: “ao olhar para as bolas, elas estão alinhadas que parecem seguir um padrão que está a crescer” [TG10] e “só que aos 9 anos estão juntas enquanto nos 18 anos começaram a afastar” [TG3] (figura 5.1.4).

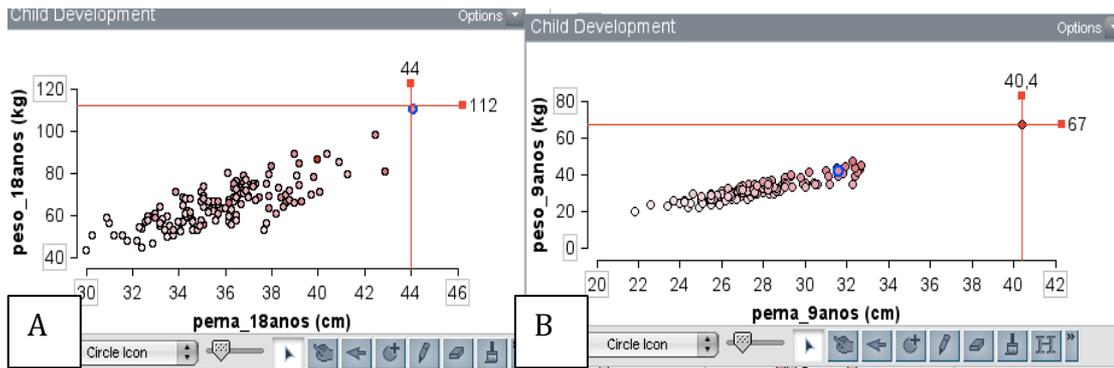


Figura 5.1. 4 – Gráfico de dispersão representando o peso e largura da perna [TG3]

O grupo 2, durante a exploração da conjectura “Será que as crianças aos 18 anos com maior peso e maior largura da perna, são [definidos como] os mais gordos?”, recorreu ao diagrama de dispersão para confirmar a associação entre as variáveis. No entanto, como a conjectura envolvia as variáveis: o peso, a largura da perna e a estatura, os alunos, em vez de construir mais do que uma representação gráfica, recorreram às características específicas deste *software* e usaram as cores dos *plots* como a terceira variável (figura 5.1.5): “a bola branca tem o menor valor [o mais delgado] e a bola azul escuro, o mais gordo”[TG2]. Quando solicitados a explicar, o grupo 2 afirmou que: “as bolas mais escuras são as crianças [consideradas] gordas e a bola branca é a criança mais magra [delgada], que neste caso as mais magras não são as que tem menor peso e largura de perna menor”. Este grupo evidencia compreender o significado das cores associadas aos dados.

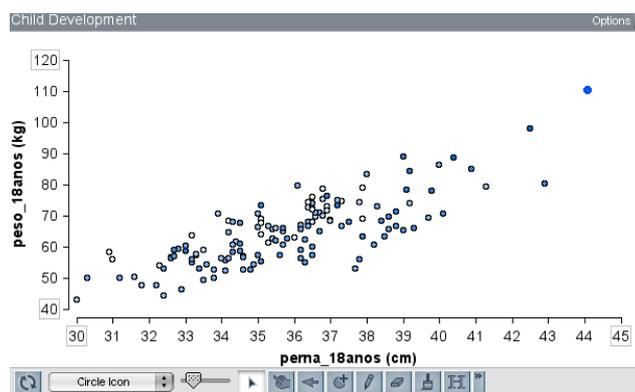


Figura 5.1. 5 – Gráfico de dispersão para o peso, largura da perna e estatura para crianças com 18 anos [TG2].

*Criatividade com os dados.* Na globalidade, os alunos procuraram evidências para confirmar as suas conjecturas, mas a criatividade no uso de diferentes representações gráficas não foi muito evidenciada. A maioria dos grupos recorreram a representações gráficas com escalas segmentadas em intervalos e diagramas de extremos e quartis (figura 5.1.6), sendo que só três grupos, recorreram a gráficos de dispersão, como da figura 5.1.3.

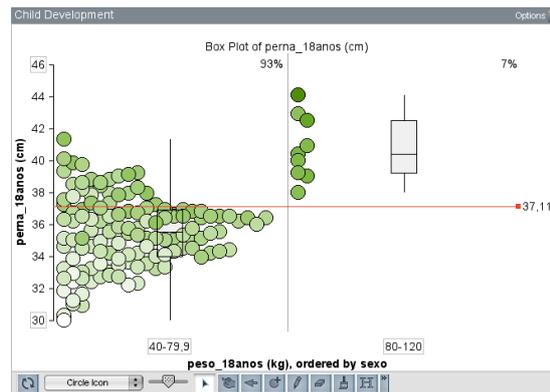


Figura 5.1. 6 – Representação gráfica para a largura da perna e o peso com 18 anos [TG10].

Durante o trabalho autónomo dos alunos na realização da tarefa, os grupos criaram diferentes gráficos para obterem as evidências sobre a existência de uma relação entre duas variáveis. O grupo 2 , tal como a maioria dos alunos, construiu representações gráficas com escalas segmentadas em intervalos em vez de escalas contínuas, para verificar a relação entre as variáveis (Figura 5.1.7). Não sendo o gráfico mais intuitivo, permitiu aos alunos identificar a existência de relação.

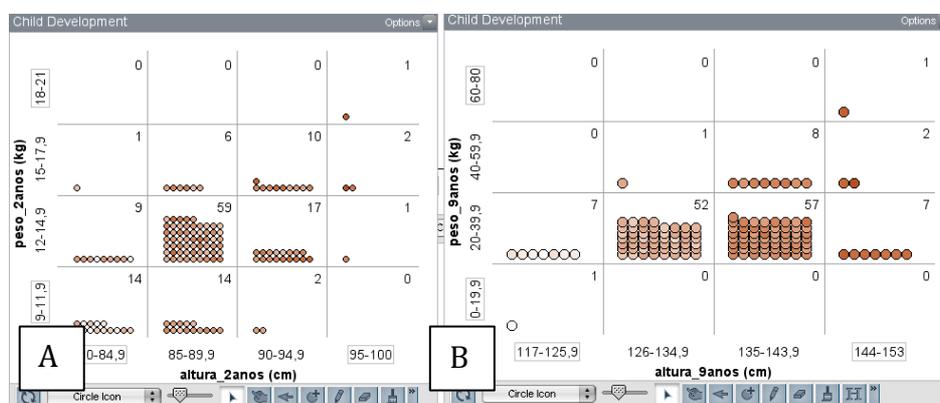


Figura 5.1. 7 – Gráfico de covariação para o peso e altura [TG2].

Relativamente aos gráficos da figura 5.1.7, o grupo 2 afirma que “no primeiro gráfico [A] as bolas estão mais distribuídas, enquanto o outro [B] não” [DC2] e “o valor mais alto de altura e de peso é sempre o mesmo nos dois gráficos e o menor é também igual” [DC2]. Este grupo descreve os dados a partir de particularidades da representação gráfica, focando na quantidade de dados que está distribuído pelos intervalos: “no intervalo com menor peso [0-19,9], no de 9 anos, vê-se que menor peso tem menor altura e apesar da maior altura não englobar os dados no intervalo maior do peso, mantém a mesma lógica “[TG2] e “no segundo intervalo do peso [20-39,9], o número aumentou de 86 crianças [de 2 anos] para 123 [crianças de 9 anos] e no terceiro intervalo, verificou-se uma descida de 8 crianças, mas com intervalos diferentes de altura” [DC2].

Ao longo do trabalho de grupo, verificaram-se vários exemplos em que os grupos conseguiram descrever dados a partir das representações gráficas construídas. Nomeadamente, na figura 5.1.2, o grupo 10 afirma que “aos 9 anos, os menos pesados, que estão entre 0 e 39,9, estão mais concentrados entre 26 e 28 cm na largura da perna e os mais pesados tendem a ter 31 e 32 cm na largura da perna” [TG10] e “quando atinge os 18 anos, esses valores estão mais dispersos” [TG10]. Outro exemplo de descrição dos dados, com base na figura 5.1.6, é quando o grupo 1 recorre à diferença entre o número de *plots* com maior peso e maior largura de perna, para descrever a relação entre as variáveis, durante o crescimento: “aqui [maior peso e maior largura] há menos 3 bolas, quando atingem os 18 anos” [TG1] e o “quando o peso é menor e a largura de perna maior, o número de bolas aumenta quando atinge os 18 anos” [TG1].

Apenas três grupos verificaram a existência de relação entre as variáveis altura e idade, a partir de diagramas de dispersão (figura 5.1.3). Aí, descrevem como utilizaram esta representação gráfica para averiguar a associação estatística: “as bolas que estão mais ou menos dentro do círculo que desenhei estão dentro da regra, se aumenta a outra também aumenta, apesar da criança 98 ter a maior altura aos dois, mas aos 18 anos, está muito abaixo dessa lógica” [TC8].

*Compreensão dos dados.* As respostas dadas pelos grupos são, na generalidade, baseadas na observação de pontos individuais. Revelam uma conceção local de associação, pois a compreensão dos dados é feita pela posição em que alguns

pontos individuais estão e, apresentam dificuldades em descrever a tendência geral da distribuição, à exceção dos grupos 3, 8 e 10. Os cartões do *TinkerPlots* são usados para identificar as crianças representadas por esses pontos. Os comentários destes outros grupos mostram que recorreram à distribuição dos dados, representada graficamente: “baseado no gráfico podemos concluir que quanto maior é a altura, maior é a força e podemos concluir também que das oito crianças com maior força e maior altura são todos os rapazes” [TG3]. O grupo 3 recorre à representação gráfica e aos dados da tabela para compreender a relação. Questionando sobre a conjectura “os mais altos aos 2 anos também são os mais altos aos 18 anos” [TC8], o grupo 8 baseia-se no gráfico 5.1.3 para afirmar a sua veracidade, mas reconhecendo exceções:

Há tendência para que quem seja mais pequeno aos dois anos, também seja mais pequeno aos 18 anos. Mas nem sempre é assim, como o gráfico mostra o mais alto aos 2 anos, tem uma altura de 98,2 cm e aos 18 anos tem 178,7 cm e o mais alto aos 18 anos (195,1) tinha aos 2 anos, 91,4 cm. [DC8]

Quando o grupo se foca em valores mais altos da altura nas duas idades, reconhece a tendência global dos dados na representação gráfica (figura 5.1.3) e a sua variabilidade. Ao referir a palavra “há tendência” mostra que o grupo compreende que existem exceções na tendência global dos dados.

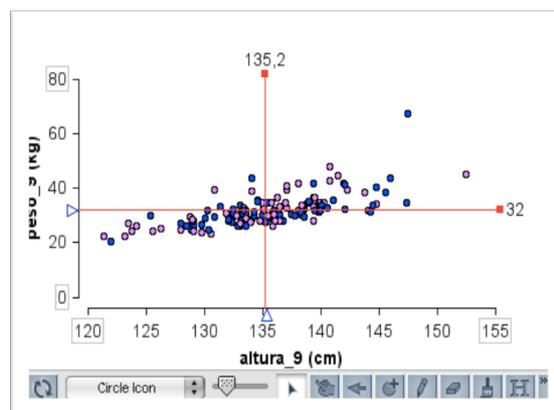


Figura 5.1. 8 – Gráfico de covariação entre peso e altura aos 9 anos [TG10].

Há grupos que conseguem identificar mensagens a partir dos dados:

Olhando para os pontos do gráfico, parece que eles organizam-se e parece que estão a crescer, nas duas variáveis. Mas nem todos crescem da mesma forma e

não é proporcional. A criança 55 é sempre a mais baixa, aos 2 anos e aos 18 anos. As crianças 72 e 119, que tem peso acima da média aos dois anos, mantém o peso acima da média aos 18 anos, mas a criança 119 que aos dois anos era mais alta do que a outra, aos 18 anos é mais baixa que a criança 72 [TG3].

Aqui, este grupo avalia de forma intuitiva a existência de uma associação linear entre as variáveis peso e altura, a partir da explicação que faz da figura 5.1.8. Desta análise, o grupo 3 foca-se na globalidade dos dados, onde se a altura aumenta, o peso também aumenta. Mas, entende que pode haver exceções, onde nem sempre se verifica essa tendência.

O grupo 10 e 1 focam-se na representação gráfica com escala em intervalos (Figura 5.1.2.B) para mostrar a veracidade da relação “as crianças de 9 anos com maior largura na perna tendem a ser as mais pesadas”. Ambos os grupos, para avaliar a existência de relação entre elas, foca a sua análise em valores individuais e não na tendência global dos dados. O diálogo seguinte mostra a forma como os grupos usam a informação da representação gráfica:

Professora: As crianças mais pesadas têm maior largura na perna, aos 9 anos?

Alunos (G1): Não. No *plot* verificamos que só a partir dos 80-120 do peso, só há 8 crianças que tem a perna mais grossa. Mas há mais crianças... 46... em que tem peso mais baixo... 40 a 79,9 quilos, com perna mais larga... entre 28 e 42.

Professora: então?

Alunos (G1): então... Não se verifica essa ideia.

Os grupos conseguiram identificar mensagens a partir dos dados, mas nem sempre recorreram à mesma análise. Por exemplo, o grupo 8 usa os dados individuais para determinar a existência de relação entre o peso e a largura da perna aos dois anos e assume que “a criança 104 é a que tem menor peso e menor largura” [TG8] e “a criança 130 tem maior peso e maior largura” [TG8]. Enquanto que o grupo 3 recorreu às características da representação gráfica (figura 5.1.4) em que mostra o aspeto relacional entre as duas variáveis, na sua globalidade: “através do gráfico podemos concluir que à medida que o peso aumenta, aumenta também a largura da perna. Pois quando mais peso, maior é a massa muscular” [TG3].

Para a questão: “As crianças de 9 anos com maior largura na perna tendem a ser mais pesadas”, o grupo 8 compreende que a largura da perna está associada ao peso dessa criança, no sentido que quanto mais pesado for a criança, mais larga a sua perna será. Quando afirma que “os pontos do gráfico estão a crescer” [figura 5.1.5] e a “sua disposição é a mesma, parecendo um padrão”. Ao serem questionados, quando referem “um padrão”, o grupo 8 explicou que “os pontos estão a crescer no eixo de baixo [eixo horizontal] e este ao crescer, está a crescer da mesma forma no eixo da lateral [eixo vertical]” [TG8] a partir da análise da distribuição dos dados, na representação gráfica. Mas ao mencionar que “apenas com uma diferença, aos 18 anos os pontos estão mais dispersos” [DC3], o grupo 3 compreende que este crescimento é mais linear para quando tem 9 anos e não linear para quando tem 18 anos.

Estas últimas afirmações dos alunos, a partir de diagramas de dispersão [figura 5.1.5], mostra que os grupos identificaram relações entre variáveis e estabeleceram a sua intensidade com base na dispersão dos dados: “verifica-se que quanto maior peso maior é a largura da perna e todos os pontos estão muito próximos um dos outros e estão a crescer no mesmo sentido” [DC8], que permitiu aos grupos responder a questões sobre os dados.

*Pensar sobre os dados.* Nesta tarefa, os alunos evidenciam compreender o contexto dos dados e usam-no nas suas afirmações, como mostram os comentários “nós sabemos que se nós aumentamos de peso, a largura da perna será grande e daí ser considerado gordo” [TG2] e “uma pessoa alta não quer dizer que seja mais pesada, pois tem a ver com a massa muscular que a criança têm” [DC8].

As respostas dadas pelos grupos são, na generalidade, baseadas no senso comum, relativamente aos fatores que estão subjacentes ao desenvolvimento de uma criança.

Os alunos revelam familiaridade com o contexto dos dados e usam-no nas suas interpretações. No entanto, houve a necessidade de clarificar o significado de algumas variáveis para que os alunos fossem capazes de as interpretar corretamente, como mostra o seguinte diálogo:

Aluno (G10): Professora... as variáveis não são as mesmas para as crianças de 2 anos.

Professora: E porquê?

Aluno (G2): Se calhar por serem pequenas.

Professora: talvez. E mais?

(...) Vejamos, porque é que só está aí [estatura do corpo] a partir dos 18 [anos]? Vamos aqui analisar (...) 2, 9 e 18 [anos], referem a que fases do crescimento?

Aluno (G2): A criança.

Professora: A infância. E dos 9... é o quê?

(...) Vocês já deram isto! Ocorre o quê?

Aluno (G2): É adolescente.

Professora: Adolescência (...) ocorre a puberdade, que é a alteração do corpo, a partir dos 9 (...) e dos 18 [anos]?

Aluna (G8): É adulto.

Professora: Têm aqui as 3 fases de desenvolvimento.

As conjeturas formuladas, inicialmente, sobre a relação entre duas variáveis, foram formuladas com base no seu conhecimento das variáveis em estudo. Todos os grupos recorreram à comparação entre os géneros, no que respeita a todas as outras variáveis: “Será que as raparigas na adolescência são maiores do que os rapazes?” [TG1] ou “Será que os rapazes são sempre mais altos do que as raparigas ou as raparigas mais altas do que os rapazes?” [TG2].

As outras conjeturas formuladas foram elaboradas, tendo em conta a relação entre duas variáveis, com base no seu conhecimento sobre o desenvolvimento das três fases em estudo (a infância, adolescências e adulta): “Poderá a altura influenciar o peso das crianças?” [TG10] ou “Será que a idade e o peso influenciam a estatura das crianças?” [TG1] ou “Será que a estatura interfere no peso das crianças com 18 anos?” [TG8] ou “Será que a estatura da criança é adequada à idade?” [TG9]. Houve, ainda, grupos que consideram outros aspetos como revelantes de serem estudados: “Será que as crianças mais pesadas tem maior largura da perna?” [TG10] ou “será que o peso tem influência na força?” [TG3] ou “Será que a força aumenta com a idade da criança?” [TG3] ou “O aumento do tamanho da perna é diretamente proporcional à idade da criança?” [TG4]. Na generalidade, os alunos não apresentaram argumentos para as suas respostas a essas questões, apenas “sim” ou “não”.

Na discussão, cada grupo selecionou uma das questões formuladas para ser discutida e usaram a natureza relacional dos diagramas de dispersão (figuras 5.1.3, 5.1.5, 5.1.6 e 5.1.7) para identificar as relações entre os dois atributos.

O grupo 10 estabelece a relação entre dois atributos afirmando “se o peso aumenta, a largura da perna também aumenta”[DC10]. Quando questionados por outros grupos, para a eventualidade de a largura da perna não ser devido ao peso, mas sim ao exercício físico, estes alunos recorrem a certos dados individuais com base no seu conhecimento sobre o contexto em causa: “o exercício físico dá massa muscular, que acaba por alargar...a massa muscular dá peso, isto influencia no peso da pessoa” [DC3].

Ao conjecturarem o crescimento de uma criança com uma altura de 85cm, os alunos afirmam que “essas crianças aos 18 anos irão ter entre 155cm e 180 cm” [DC2] porque “as crianças que têm 85 cm aos 2 anos vão ter alturas diferentes aos 18 anos, pois os estilos de vida e hábitos alimentares mudam de pessoa para pessoa” [DC8]. O grupo 6, ao analisar a figura 5.1.3, para esta questão afirma que não podemos generalizar e que depende de muitos fatores, constatando a partir dos dados que “visualizamos que dois alunos (consideramos os mais baixos) estão acima da média e os outros estão todos abaixo da média” [TG6]. Neste conjunto de comentários dos alunos, consegue-se observar como os alunos ao descrever os dados a partir dos gráficos, justificam as exceções da tendência global dos dados com o conhecimento do seu dia-a-dia.

Alguns alunos focam que a herança genética é o fator mais previsível e importante do crescimento, pois “a influência da altura dos pais é muito forte” [DC3], ou seja, “se existe apenas pessoas altas na família, a criança tem mais probabilidade de crescer mais [de crescer com o mesmo padrão] ou se são todos baixos, a criança provavelmente não será diferente” [DC10]. Para além disso, defendem que a altura de uma pessoa deve-se a outros fatores: “a saúde da mãe durante a gestação”[DC3] e “alimentação da criança depois do nascimento, porque se não receber os nutrientes na quantidade que precisa, irá ter problemas no seu crescimento” [DC8].

Alguns grupos fazem inferências causais recorrendo aos dados: “a perna pode ser mais larga e o peso ser menor, devido à altura, hábitos alimentares e o exercício

físico” [TG1] e “se os pais são baixos, o seu filho será também baixo, devido à hereditariedade” [DC10].

Os alunos procuram argumentos para justificar as relações entre as variáveis dos dados ou as variações entre os dados das variáveis, partindo do conhecimento do dia-a-dia. O grupo 2 remete para o “senso comum”, quando recorre à informação que obtém dos gráficos, na relação entre a altura e a estatura (gorda ou delgada): “se fosse muito gorda, era extremamente baixa” [TC2] e “há pessoas muito gordas, que são muito altas” [TC2]. Nesta situação, tanto o grupo 2 como o grupo 3, refere que esta situação ocorre devido aos “cuidados alimentares que as pessoas têm e a prática desportiva que essas pessoas fazem” [DC3]. Especulam, ainda, que ao caminhar e aumentar a atividade física, as pernas se estabilizam, enquanto que os músculos ficam mais desenvolvidos: “fazerem desporto, o corpo da criança irá sempre desenvolver” [DC8] e que “se os músculos estiverem mais desenvolvidos, terão mais força” [DC10].

Para justificar as contradições dos dados, os grupos acabam por dizer que uma rapariga é alta, em todas as idades: “como vimos na aluna com o número 10” [DC8], mas existe algumas meninas que podem ser altas durante a infância, mas terminam com uma estatura adulta mais pequena (figura 5.1.7) e tentam explicar que pode ser devido a “ter deixado de praticar exercício físico” [DC3].

## **SÍNTESE**

Os alunos compreenderam como usar as ferramentas do *software* e o ambiente dinâmico que ele proporciona. Construíram representações gráficas que mostram a relação entre dois atributos e foram capazes de modificá-las, de forma autónoma e intencionalmente, utilizando diversas ferramentas do *TinkerPlots* (por exemplo: escalas contínuas, linhas de referência e medidas estatísticas).

Os resultados da análise desta tarefa salientam que alguns alunos têm facilidade em verbalizar as suas hipóteses, com base no seu conhecimento, na análise de relações entre duas variáveis. O uso das ferramentas do *TinkerPlots*, ajudou a focar as suas ideias e explicá-las, com base em diferentes representações, compreendendo assim, os dados em estudo.

Durante a resolução da tarefa, os grupos focaram-se em casos individuais para identificar mensagens nos dados. Alguns grupos não evidenciaram dificuldades em identificar mensagens nos dados e nos gráficos sobre a existência de covariação entre as variáveis. Os argumentos dos alunos foram muitas das vezes, baseados no senso comum e na sua experiência, sobre os fatores que estão subjacentes ao desenvolvimento da criança. Quanto à variedade de gráficos para demonstrar a associação entre os atributos, reconheceram que o diagrama de dispersão, é uma representação gráfica adequada para identificar as relações entre dois atributos. Na generalidade, os grupos valorizaram a informação presente nos gráficos para justificar as conjeturas formuladas a partir dos atributos.

## **5.2. TAREFA – “INDICADORES DEMOGRÁFICOS”**

### **ENQUADRAMENTO**

Esta tarefa tinha como objetivo explorar a existência de relações entre as características da população portuguesa em 2011. Espera-se que os alunos intuitivamente verifiquem a existência de associação entre duas variáveis, quando a nuvem de pontos se condensa à volta de uma reta e a sua relação com o centro de gravidade de distribuição.

A tarefa realizou-se em duas aulas. A primeira, de 120 minutos, integrou três partes. Na primeira parte, cada grupo leu a parte introdutória do tema em estudo. Depois, antes do trabalho autónomo, discutiu-se em grande grupo, o problema colocado na tarefa, com a finalidade de assegurar a correta interpretação do enunciado. Na segunda parte realizaram trabalho autónomo. Na última parte dessa aula, ocorreu a discussão final em grande grupo. Na segunda aula, fez-se uma revisão do que foi discutido em grupo na aula anterior, para os ajudar a realizarem a terceira parte da tarefa, correspondente a um relatório onde apresentavam as conclusões.

## RACIOCÍNIO COVARIACIONAL

*Conhecimento genérico.* Como já verificado na tarefa anterior, os alunos utilizaram as funcionalidades disponíveis do software, mais nuns grupos do que noutros. Constroem com facilidade e rapidamente os gráficos, sem necessidade de ajuda no acesso aos recursos do *TinkerPlots*. Salienta-se que os alunos, para responderem às questões da tarefa, utilizaram essencialmente diagramas de dispersão e de extremos e quartis (figura 5.2.1), reconhecendo as características destes gráficos.

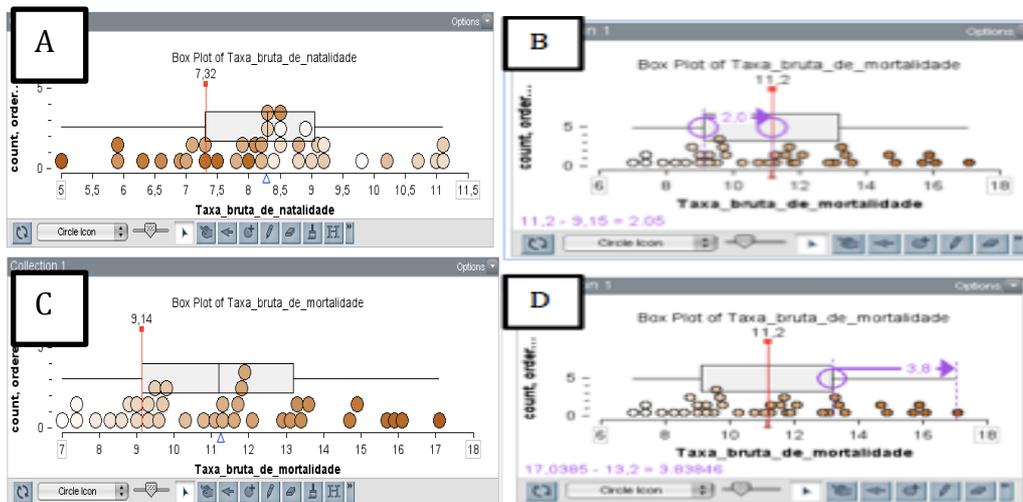


Figura 5.2. 1 – Comparação entre o número de mortes e o número de nascimento, através do boxplot [G3] (A e C). Diagrama de extremos e quartis, nos diferentes dados de distribuição na variável “Taxa bruta de mortalidade” [G8] (B e D).

Ao comparar variáveis, os grupos recorreram à utilização das medidas de localização: “devemos utilizar as médias nas duas taxas” [NC] e “ordenar os pontos para colocar os quartis” [NC]. Conclui-se, que identificam corretamente os intervalos, a moda, a média e descrevem corretamente, o que “hat” representa. Mas, certos grupos, usam linguagem informal para descrever as representações gráficas: “esta barra é maior do que esta, que significa que os valores da mortalidade estão mais afastados do que na natalidade” [TG2].

A partir da construção do gráfico da figura 5.2.1, os alunos representaram os valores dos dados e visualizaram a sua distribuição, ponderando a sua resposta, mediante a distribuição da amostra por quartis, quer do número de mortes quer do número de nascimentos:

Na natalidade, o 1º quartil é 7,31, logo 8 regiões do país estão pelo menos a 25% da amostra e a diferença entre o 3º quartil e o 1º quartil, 50% da taxa de natalidade tem entre 7,31 e 9 [DC10].

O grupo 10 compreende o significado determinado pela amplitude interquartílica e é perceptível que têm a noção de que entre o 3º e o 1º quartil, está a 50% da distribuição. Contudo, revela alguma falta de rigor na linguagem, pois é “pelo menos 50%” e não “50%”, como foi referido.

Os grupos revelaram ter compreendido a forma como os dados estão distribuídos (figura 5.2.1), nomeadamente que do extremo inferior até ao terceiro quartil (inclusive) estão pelo menos 75% dos dados e que a dispersão dos mesmos varia, podendo observar-se no diagrama, onde é maior ou menor.

Na comparação dos dois diagramas, referiram as semelhanças e as diferenças entre a forma e o centro, como os dados estão distribuídos e a maior ou menor dispersão dos dados. As figuras 5.2.1 A e C, confirmam a ideia de simetria ou de enviesamento dos dados, pois o grupo 8 e 10, com as funcionalidades do *rule*, concluíram que a distribuição da mortalidade é simétrica, enquanto na distribuição da natalidade há um enviesamento do lado esquerdo, partindo da distância entre a linha indicada na mediana e os lados do retângulo e o comprimento das linhas dos lados do retângulo.

Na interpretação da distribuição dos dados só 4 grupos (G2, G3, G6 e G8) é que reconheceram globalmente este enviesamento à esquerda e a simetria. O diálogo seguinte é indicativo do modo como foi usada a informação dos quartis, dos gráficos da figura 5.2.1:

Aluno (G10): Com esta coisa que está na barra [ícone *rule*], consigo saber o valor da diferença entre os extremos inferior e superior e dos quartis... Q1... Q2... Q3.

Professora: Que conclusões retiras daí?

Aluno (G10): A natalidade é assimétrica e a mortalidade é simétrica.

Professora: Como?

Aluno (G10): Na natalidade, entre o Q1 e o Q2 é 1,00 e entre o Q2 e o Q3 é 0,75. O primeiro é maior que o segundo.

Professora: E...

Aluno (G10): No segundo estão mais concentrados e no primeiro estão mais dispersos. No outro, no retângulo, as diferenças entre o Q1 e Q2 e entre Q2 e Q3, são iguais... é como se fosse um eixo de

simetria... um espelho no Q2. Ao dobrar o Q1 e o Q3 iriam coincidir.

Nesta tarefa, os grupos usaram as características do TinkerPlots estabelecer a existência de associação entre variáveis. Durante a discussão dos gráficos (figura 5.2.1), há grupos que recorrerem à cor dos plots:

Aluno (G2): oh, stora... estava aqui a brincar com os gráficos e vi que podemos comparar as cores.

Professora: Como assim? Explica-te.

Aluno (G2): Coloco o rato no atributo da natalidade, as cores mudam... o escuro é o maior e o mais claro é o menor.

Professora: Sim e?

Aluno (G2): No gráfico onde está a natalidade, a cor escuro está para os valores mais altos, mas ao olhar para o gráfico de baixo [da mortalidade], a cor está no início da reta... E se fizer o contrário [clicar no atributo da mortalidade], o escuro na natalidade está no início e na mortalidade no fim.

Professora: Que conclusão tiras dessa afirmação que fizeste?

Aluno (G2): oh, stora... está claro! A relação entre a natalidade e mortalidade. Se há mais mortalidade, há pouca natalidade.

Este aluno [G2] foca-se no gradiente da cor representado nos dois gráficos, para verificar a existência da relação entre os dados das duas variáveis, comentando ainda “Pinhal Interior Sul é um exemplo disso... é o que tem mais mortalidade e o que tem menor natalidade” [DC2]. Ao usar as linhas de referência, o grupo 2 procurou encontrar os valores das variáveis para salientar que dois pontos do país com a mesma natalidade, a percentagem da mortalidade não é igual: “os Açores e a Grande Lisboa que tem 11,1 % de nados vivos... na mortalidade, a Grande Lisboa tem 0,8% que os Açores” [DC2] (figura 5.2.2). As linhas de referência são uma característica que a maior parte dos grupos utilizou. O grupo 1 utilizou para explicar a posição dos dados ou um valor dos dados, mediante um valor, a média de cada uma das variáveis, onde refere que “existe quatro zonas em Portugal que há mais casamentos, a Tâmega, o Dão-Lafões, os Açores e Minho-Lima” [TG1].

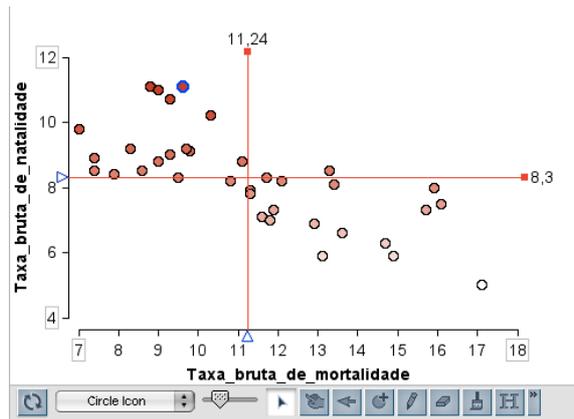


Figura 5.2. 2 – Diagrama de dispersão com as variáveis “Taxa bruta de natalidade” com a “Taxa bruta de mortalidade” [G2].

Este grupo passou de duas representações gráficas com *Box Plot*, para um diagrama de dispersão, pois reconheceu ser a melhor representação para justificar que “quando o valor de nascimento aumenta, o valor de mortes é menor” [TG2]. Na sua análise, este grupo recorre constantemente à representação gráfica e aos dados da tabela, quando descreve a relação estatística entre os atributos, mas foca-se em valores individuais.

O grupo 8, ao representar graficamente a idade média ao primeiro casamento entre o homem e a mulher, criou duas escalas, uma contínua para as mulheres e em intervalos para o homem (figura 5.2.3).

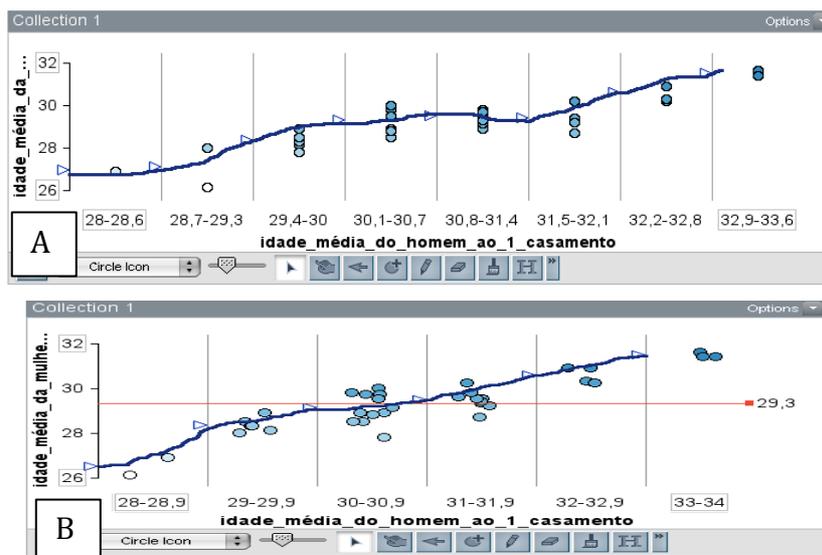


Figura 5.2. 3 – Gráfico de pontos com duas escalas (contínua e intervalar) [G8].

Quando questionados para este facto, referem que “ao colocar esta variável dentro de um intervalo e procurar a média, verificamos ao unir numa linha as médias, estas estão a subir” [TG8]. E pela análise dos dados no gráfico, mencionam a relação entre as idades do homem e da mulher, pois “tanto os homens como as mulheres tem mais idade quando é no primeiro casamento” [TG3] e recorrendo aos dados nos cartões, afirmam que em todo o país, “Lisboa, a Grande Lisboa e o Alentejo litoral são as zonas que casam pela primeira vez mais tarde” [TG8]. Esta ideia é mantida, mesmo quando mudaram a escala do intervalo (figura 5.2.3).

Nesta tarefa, alguns grupos de trabalho (grupo 2, 3 e 8) usam, na generalidade, gráficos de dispersão para reconhecer a possível relação entre as variáveis, focando nos aspetos da sua distribuição e direção. Relativamente à representação gráfica (figura 5.2.3), modificaram o gráfico de dispersão, transformando um dos eixos em segmentos. Ao determinar a média de um dos eixos e usar o recurso de desenho para unir as médias de cada intervalo, permitiu ao grupo dizer que “os pontos da média estão a crescer e ao unir forma um traço crescente” [TG8] e daí justificar que “à medida que a idade média do homem aumenta a idade média da mulher também” [TG8]. Desta forma, verifica-se que os grupos conseguem usar as características do software para demonstrar a associação entre as variáveis.

*Criatividade com os dados.* Todos os grupos criaram as mesmas representações gráficas, mas as mais usuais para responderem às questões propostas, foram na base de diagramas de dispersão (figura 5.2.2 e 5.2.3) e por vezes, diagramas de extremos e quartis (figura 5.2.1). Desta forma, verifica-se que não houve uma grande diversidade de gráficos construídos.

Partindo da figura 5.2.1, os grupos sumariam os dados, quando afirmam que “o número de mortes é muito superior ao nascimentos” [DC3], pois “a média da mortalidade (11,25) é superior ao da natalidade (8,28)” [DC10].

Todos os grupos recorreram à representação gráfica (figura 5.2.1) e descreveram os dados a partir dele (figura 5.2.4).

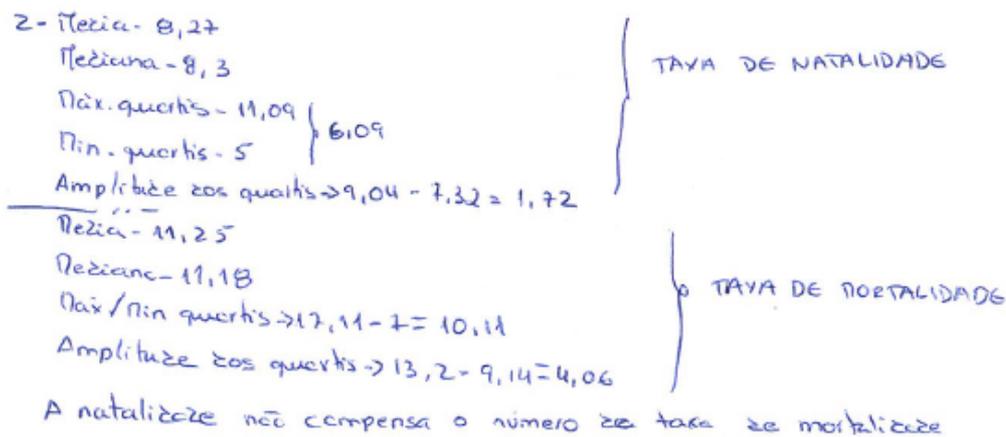


Figura 5.2. 4 – Resposta à questão 2 [G8].

Da figura 5.2.4 através dos valores retirados do diagrama de extremos e quartis, descrevem os dados que observam das representações gráficas construídas. Daí, alguns grupos, afirmaram que “na natalidade, os dados estão mais próximos, enquanto que na mortalidade, os dados estão mais dispersos” [TG8], pois a “diferença entre o máximo e o mínimo é menor [6,09] em relação à natalidade [10,11]” [TG8], e noutros atestam a mesma coisa, mas recorrem ao comprimento do desenho do quartil para afirmar a dispersão dos dados, “o retângulo do quartil 1 até ao quartil 3 é mais comprido na mortalidade que na natalidade” [DC10], que associaram esse comprimento do desenho ao valor da diferença de interquartis, “menos o Q3 e o Q1 da natalidade é 6,09 (...) e na mortalidade é 4,06” [DC6].

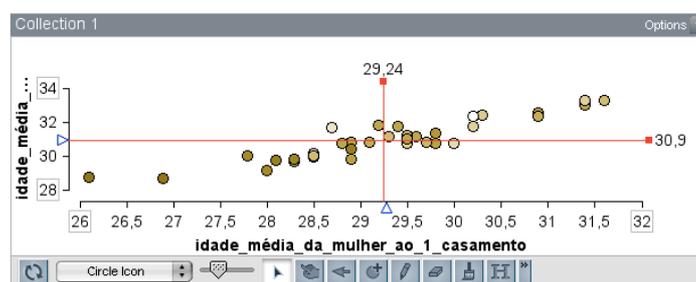


Figura 5.2. 5 – Diagrama de dispersão representando as variáveis: “Idade média da mulher ao 1.º casamento” e “Idade média do Homem ao 1.º casamento” [G3].

Na construção do gráfico, da figura 5.2.5, os alunos apresentaram dificuldades em encontrar no diagrama, o centro de gravidade, apesar de identificarem corretamente as variáveis e o seu significado, sendo necessário o apoio da

professora. Na análise dessa representação gráfica, o grupo 3 afirma que “a média da mulher ao 1.º casamento é de 29,32 anos e nos homens a média do 1.º casamento é de 30,92 anos” [TG3], sintetiza que “tanto as mulheres e os homens casam entre 29 e os 30” [TG3].

O grupo 2 conclui, ainda, “há mais homens a casar com idade superior ou igual a 30 do que mulheres” [TG2], como mostra no seguinte excerto:

Na linha de baixo e de cima, coloco esses nomes [variáveis], mas em cada linha, coloco-a em intervalos... nas mulheres faço quatro intervalos [26;27,9], [28;29,9], [30;31,9] e [32;34]. No homem, coloco três intervalos [28;29,9], [30;31,9] e [32, 34]... daí vejo, que acima dos 30 anos, nas mulheres há 9 e nos homens há 28. (...) e ainda mais... com o *Counts*, consigo ver que dos sete homens com [32-34] e das sete mulheres [30;31,9] estão em zonas desenvolvidas do país... Alentejo litoral, baixo Alentejo, Algarve, Grande Lisboa, Alentejo Central, Lisboa e a Península de Setúbal. (G2)

Este aluno mostra um erro na descrição dos dados, pois associa aos *plots* o número de pessoas por sexo, em vez de afirmar que há sete regiões, onde a média das idades no primeiro casamento nos homens é entre 32 e 34 e nas mulheres é entre 30 e 31,9.

Relativamente ao gráfico B da figura 5.2.4, o grupo 8, depois de determinar o centro de gravidade, a partir da interseção das médias de cada variável, afirma que os *plots* estão “espalhados pelos 4 quartis” e “não se consegue ver nesses 4 espaços [quadrantes], qual é que tem mais” [TG8].

*Compreender os dados.* As respostas dadas pelos grupos são, na generalidade, focadas no senso comum, requerendo por vezes algum questionamento da minha parte aos grupos, durante a resolução da tarefa, quer a nível da compreensão dos dados quer a nível da representação gráfica.

Além disso, da análise dos diagramas de extremos e quartis (figura 5.2.1), mais concretamente a partir da média e da mediana, os grupos afirmam que a taxa de natalidade é tão baixa que não compensa a taxa de mortalidade, logo não evidencia “nenhuma relação entre elas” [DC10]. Nesta análise, os alunos procuraram evidências para confirmar as suas ideias, reconhecendo que não existe relação. Contudo, quando o grupo 3 representou as duas variáveis num diagrama de dispersão, a sua conclusão foi diferente: “há relação, apesar de não ser igual,

uns aumenta e outros desce!” [TG3]. Aqui, os alunos referem-se à distribuição dos valores ao longo do eixo horizontal, pois analisam a forma e a direção da nuvem de pontos que representa as duas variáveis. Partindo deste diálogo com os dados, a partir da representação gráfica (figura 5.2.7), estes grupos identificam mensagens a partir dos dados, recorrendo aos cartões do TinkerPlots para lhes atribuir significado.

Quando questionados sobre a relação, atendendo ao número de dados que estão presentes no gráfico, cujas variáveis correspondem às idades médias ao 1.º casamento entre o homem e a mulher, afirmam que “no 1.º e no 3.º quadrante há maior quantidade de dados em relação ao 2.º e 4.º quadrante, [que] há menor quantidade de dados” [GTG8], ou seja, os pontos distribuem-se maioritariamente nos quadrante 1 e 2 (figura 5.2.6A).

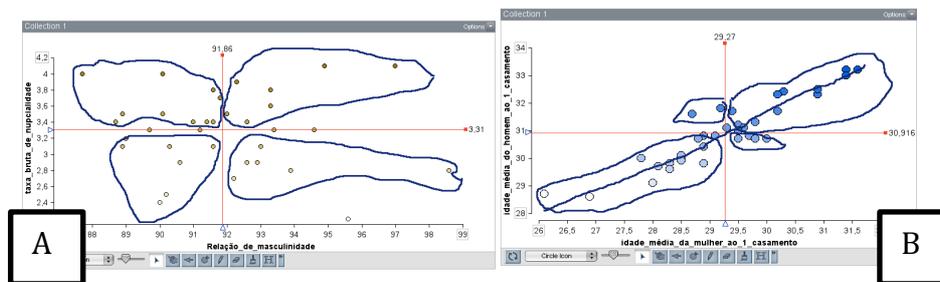


Figura 5.2. 6 – Diagrama de dispersão representando as variáveis: A- Idades médias ao 1.º casamento, relativamente ao homem e mulher [G10] e B – Relação de masculinidade e taxa de nupcialidade [G8].

Confrontados com o aspeto relacional, os alunos afirmaram que “os pontos não estão dispersos... aproximam-se uns dos outros” [DC10] e que “os valores da idade média da mulher aumentam, ao mesmo tempo que os valores da idade média dos homens” [DC3]. Devido a esse facto, este gráfico “dá a entender que [os dados] forma[m] uma reta que vai crescendo, de forma ascendente” [TG6]. Ao usar a expressão “dá a entender” mostra que este grupo compreende a existência de exceções na tendência global dos dados.

Relativamente ao gráfico B da figura 5.2.6, os alunos referem que “os dados estão dispersos e, como não há declive, não há relação” [CDC10], pois as variáveis não se relacionam entre si, sendo então considerados como independentes, como é indicativo na justificação do grupo 6 (figura 5.2.7):

3.3 → Tem tendência a crescer a idade média, ou seja, acontece tanto nos homens como nas mulheres, ou seja quando os homens aumentam nas mulheres também vai aumentar. Enquanto que no outro gráfico os dados estão dispersos, ou seja, nem aumenta nem diminui.

Figura 5.2. 7 – Resposta à questão 3.3. [G6].

No que se refere ao aspeto relacional de diagramas envolvendo outras variáveis (taxa de mortalidade, taxa de natalidade e taxa de mortalidade infantil), quando o índice sintético de fecundidade tende a aumentar (figura 5.2.6), os alunos conseguem analisá-lo em termos de dispersão de dados, dizendo que “no gráfico da taxa de mortalidade infantil e na taxa de mortalidade não existe relação” [DC7], pois o gráfico A tem “os pontos espalhados pelos quatro quadrantes, mais ou menos da mesma forma” [DC6] e desta forma, “tanto sobe o ponto como desce” [DC2]. No caso do gráfico B, os alunos ao analisarem a distribuição dos dados constatarem que “os pontos estão distribuídos da mesma forma no 1.º e 2.º quadrante” [DC3] e no “3.º quadrante tem 4 pontos e no 4.º quadrante tem 2 pontos” [TG3]. Desta forma, os grupos identificam que não há uma relação entre estas duas variáveis, e a ocorrência da taxa de mortalidade infantil é independente do índice sintético de fecundidade. No que se refere ao diagrama de dispersão (gráfico C), os grupos interpretam a inclinação da reta formada pelos pontos, pois “os dados estão mais agrupados” [TG8] e “quando a taxa aumenta, o índice também vai aumentar” [DC8].

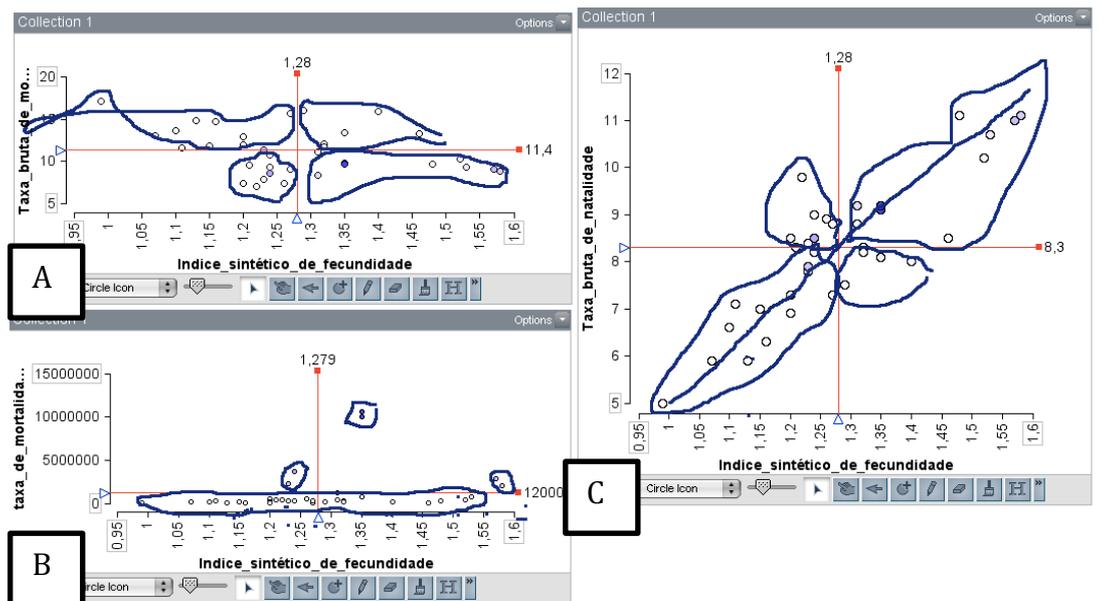


Figura 5.2. 8 – Diagrama de dispersão representando as variáveis “índice sintético de fecundidade” com: “Taxa bruta de mortalidade” (A); “Taxa de mortalidade infantil” (B) e “Taxa bruta de natalidade” (G8).

Nas representações gráficas, onde identificam a existência de relações, o grupo 8 e 10 interpretaram o sentido e a intensidade da associação entre as variáveis, com base na sua distribuição:

[Figura 5.2.2] os pontos distribuem-se no sentido descendente e a sua inclinação é decrescente (...) é óbvio se aumenta a mortalidade a natalidade tem que diminuir... nem que seja por Portugal ser um país desenvolvido (...) A sua intensidade parece ser forte e como é decrescente, o valor do coeficiente é negativo um bocadinho acima dos 0,5. [TC8]

Neste gráfico [C- figura 5.2.9], os pontos distribuídos tem declive positivo (...) parecem que vão aumentando nos quadrantes 1 e 3 (...) os pontos aproximam-se, criando a ilusão de uma reta, mas os pontos não formam uma linha... mas aproximam-se dessa linha ou reta, se nós a traçarmos. A ligação entre os pontos desse gráfico aproximam de uma intensidade forte, onde a dependência dessa relação é entre 0,5 e talvez 0,8. Nos outros [A e B, da figura 18], como os pontos estão dispersos pelos quadrantes, sem inclinação, o tipo e a intensidade é nula. [TG10]

A explicação desses alunos permite concluir que conseguem retirar a mensagem a partir da distribuição dos dados, interpretando-a para avaliar a existência de uma associação, quer positiva ou negativa, entre duas variáveis em estudo. As relações tornam-se mais fortes quando os pontos estão alinhados, criando a noção de que a razão entre os correspondentes valores das duas variáveis aproxima-se de uma constante. Aqui, estes grupos interpretam a intensidade e o sinal da covariação

através do coeficiente de correlação, relacionando com a dispersão ou linearidade da distribuição da nuvem de pontos.

Contudo, houve alguns grupos que deram algumas respostas incorretas: “[gráfico com as variáveis do 1º casamento] os dados estão a crescer e estão mais juntos e que faz uma reta, embora não seja perfeita, então a linearidade é muito forte” [G3]. O grupo 3 baseou-se na distribuição dos dados para justificar a variação da nuvem de pontos, mas como a janela de visualização utilizada era pequena, induzi-os a que fosse “muito forte”, por estarem muito próximos.

Ao analisar todo o trabalho dos grupos, verifica-se que os mesmos compreenderam a utilidade do diagrama de dispersão, permitindo observar o aspeto relacional entre as variáveis, no qual falaram sobre a dispersão de dados. Partindo das figuras 5.2.6, 5.2.7 e 5.2.8, reconheceram quando havia associação ou não.

Alguns grupos (como os grupos 3, 6, 8 e 10) consideraram a existência de associação entre as duas variáveis através da tendência da variação da nuvem de pontos, no mesmo sentido: “depende, porque, de uma forma geral, a região com maior natalidade tende a ser aquela com índice de fecundidade maior” [DC8]. Esta resposta baseia-se numa conceção correta de covariação entre duas variáveis. O grupo 10, na sua explicação, partindo do gráfico B da figura 5.2.8, verificou que a nuvem de pontos se encontra bastante dispersa, o que fez prever que não existe uma associação clara entre as duas variáveis, e desta forma, compreenderam que não havia ausência de variação. Salienta-se, ainda, que através das características das representações gráficas, os grupos procuraram avaliar as relações entre variáveis, baseando as suas respostas na nuvem de pontos que se condensa à volta de uma reta ou em pontos discrepantes, que acabam por influenciar o valor da correlação.

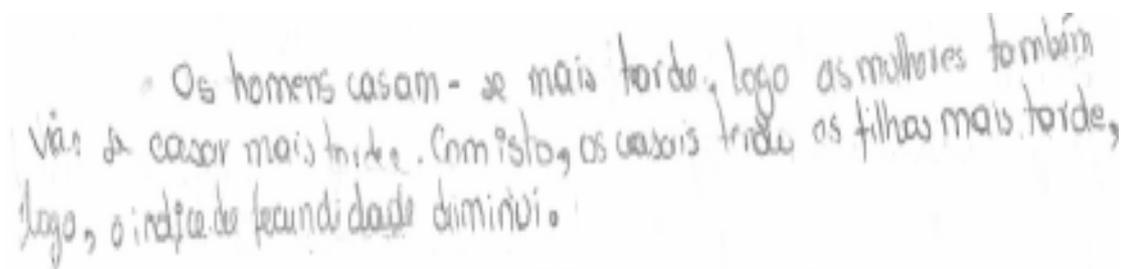
Nos gráficos onde existe relação, para justificar a associação entre essas variáveis, os alunos recorreram ao ajustamento de uma reta nos diagramas de dispersão: “sim, porque a maneira como os dados estão distribuídos no gráfico, mostram no fundo a ideia de uma possível linha crescente, neste caso e decrescente, no outro” [DC3]. O grupo 3 mostrou que as variáveis podem ser negativamente associadas (figura 5.2.2), quando os maiores valores de uma correspondem, de uma maneira geral, os menores valores da outra.

Relativamente à intensidade da associação dessas relações, os grupos partiram do pressuposto de que a maior ou a menor dispersão dos pontos implica uma fraca ou forte associação entre as variáveis em estudo:

“Existe relação, porque os pontos apresentam uma menor dispersão e por isso aproximam-se” [TG2].

*Pensar sobre os dados.* No gráfico F da figura 5.2.4, o grupo 10 concluiu que existem limitações nos dados que não permite interpretar corretamente a associação, quando recorreram aos pontos individuais para interpretar os dados; “na taxa de mortalidade infantil, há dois pontos bastante afastados dos outros, o Continente e Portugal, porque é o somatório dos outros pontos das regiões” [DC10] fazendo com que a “distribuição seja só em dois quadrantes, o primeiro e o segundo, sem que se consiga ver alguma associação entre os dados” [DC10].

Os grupos procuraram encontrar argumentos para justificar porque é que só em alguns casos ocorre a relação ou o porquê de contradizerem as ideias generalizadas sobre os indicadores geográficos (figura 5.2.9).



Os homens casam-se mais tarde, logo as mulheres também vão se casar mais tarde. Com isto, os casais têm as filhas mais tarde, logo, o índice de fecundidade diminui.

Figura 5.2. 9 – Resposta da parte III (G6).

A fim de justificar a relação causal entre a taxa de mortalidade e a taxa de natalidade [gráfico H, da figura 5.2.5], o grupo 6 afirmou que “há decréscimo na natalidade, pois há diminuição da população rural e aumento da população urbana” [DC6] e ainda, “com a integração da mulher na vida profissional, [elas] têm filhos mais tarde” [DC6]. Relativamente à última ideia mencionada pelo grupo 6, o mesmo diz que “a taxa de natalidade e de fecundidade estão inter-relacionadas” [DC6], uma vez que a quebra da taxa de natalidade resulta,

fundamentalmente da redução do índice de fecundidade, apontando para a distribuição dos dados no gráfico C da figura 5.2.7.

Outra situação, de inferências causais baseadas nos dados, é quando o grupo 3 refere que “as mulheres como a partir de uma certa idade entram na menopausa e que faz com que não possam ter filhos”[TG3] e que “isto reflete-se através da relação entre os gráficos da idade da mulher no 1.º casamento e a taxa de natalidade e o índice de fecundidade” [TG3]. Ainda, em relação à taxa de mortalidade e de natalidade, o grupo 3 reconhece algumas limitações nos dados da taxa de natalidade, pois:

a taxa de natalidade é uma análise deficiente, porque para achar esse valor, os investigadores relacionam os nascimentos com o total da população daquela região, sem considerar o seu sexo e a sua idade (...) só as mulheres tem filhos e algumas delas, são muito jovens ou já são idosas, com incapacidade de procriar. [TG3]

O grupo 2, durante o trabalho de grupo e em discussão coletiva, procurou encontrar causas para a variação, quando refere que “em Portugal, o envelhecimento da população e a morte natural da velhice são controlados pela qualidade dos cuidados dados pelos hospitais e dos níveis alimentares” [TG2], bem como, “haverá poucos nascimentos, pois há mais velhos vivos e as velhas já não engravidam”[DC2].

O grupo 10, procurou argumentos que permitam justificar as relações encontradas e para caracterizar a evolução da população, tendo em conta o comportamento dos vários indicadores demográficos utilizados. Isto é, evidente, quando afirma que:

a idade do primeiro casamento, este cada vez aumenta mais, ou seja, a idade do primeiro casamento entre ambos os sexos, proporciona-se cada vez mais tarde a procriação (...) a falta de tempo, a ocupação a tempo inteiro pela vida profissional, leva a que os nascimentos sejam mais tarde (...)Esta situação, também acontece devido à redução das taxas de nupcialidade e aumento da taxa de divórcios. Nas zonas onde há maior natalidade, normalmente são as do interior do país, para terem mais mão de obra no campo. Contudo, a partir do gráfico da natalidade, isso não verifica, se calhar é devido à falta de trabalho no campo e pouco financiamentos, que levam à emigração da população jovem, acabando por diminuir os nascimentos. Pois são há idosos nas aldeias do interior. [TG10]

## **SÍNTESE**

É assim, ainda evidente, ao longo da tarefa, que alguns alunos não usam corretamente a linguagem dos gráficos e dos dados. Contudo, quanto à capacidade de análise de dados a partir das representações, os alunos melhoraram o seu desempenho em todos os aspetos apreciados. Embora, alguns grupos, ainda, cometam erros de interpretação.

Na interpretação dos dados, a resposta dada pelos alunos parte diretamente da informação explícita no gráfico de dispersão, através da distribuição dos pontos. Ao longo da tarefa, não são visíveis situações onde os alunos não estabelecem relações entre tabelas, gráficos e dados. Mas nas representações gráficas construídas, os grupos compreendem que para mostrar a relação entre duas variáveis devem recorrer ao aspeto da dispersão, forma e direção do diagrama de dispersão. Relativamente, às conjecturas sobre as relações entre as variáveis, recorrem ao seu conhecimento pessoal. Ou seja, a interpretação dos dados é parcialmente baseada no conhecimento que os alunos têm do contexto e parcialmente, baseadas da interpretação direta das representações gráficas.

### **5.3. TAREFA – “BATERIA DE TESTES DE FITNESSGRAM”**

#### **ENQUADRAMENTO**

A tarefa, “Bateria de Testes do Fitnessgram”, tinha como objetivo identificar, descrever e fazer inferências sobre os dados, resultantes de testes de condição física de todos os alunos que frequentam o 9.º ano de escolaridade, identificando a relação de dependência entre as variáveis e avaliar, de forma intuitiva, a partir da análise de diagramas de dispersão, a associação entre os dados, ou seja, a covariação entre duas variáveis. A tarefa está estruturada em três partes e foi realizada em três aulas. A parte I, que correspondeu a uma aula de 60 minutos,

iniciou-se com a leitura da parte introdutória relativamente ao tema, com discussão sobre o problema em grupo, passando depois para o trabalho em grupo. A parte II e III, em trabalho de grupo, foi realizada noutra dia, numa aula de 120 minutos. A discussão em grande grupo sobre o trabalho realizado na tarefa ocupou a terceira aula de 60 minutos.

A resolução desta tarefa, seguiu a dinâmica já aplicada nas outras tarefas. Nesta tarefa, pretende-se que os alunos identifiquem e informalmente descrevem relações entre as variáveis, a partir da observação, comparação e interpretação da informação veiculada em diagramas de dispersão. Identificar e interpretar o sentido e a intensidade da correlação entre duas variáveis, associando a covariação com a análise crítica do valor do coeficiente de correlação associado, avaliar criticamente a informação e concluir sobre a sua veracidade. Bem como, reconhecer que pode não fazer sentido interpretar a associação das variáveis como de “causa-efeito”.

## RACIOCÍNIO COVARIACIONAL

*Conhecimento genérico.* Constatou-se que, na generalidade, os alunos foram capazes de aceder facilmente às ferramentas do *TinkerPlots*. Não mostraram dificuldades na construção de gráficos e, quando solicitados, foram capazes de descrever uma representação: “os que estão pintados de verde escuro são os que demoram mais tempo a percorrer os 60 metros e estão com nos intervalos com menor altura” (CTG2) (figura 5.3.1).

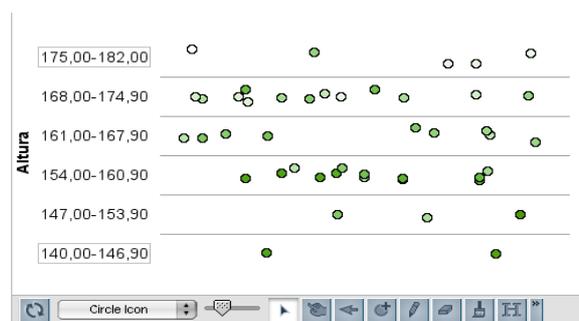


Figura 5.3. 1- Representação gráfica de “Altura” e “Velocidade” [G2].

O grupo 2 inicia a sua exploração criando uma representação gráfica para representar os dados relativos à altura usando uma escala segmentada (figura 5.3.1). Nesta representação gráfica, o grupo estabelece uma associação entre as duas variáveis, altura e velocidade, sem introduzir a última, numa escala. Este grupo recorre a uma característica específica das representações gráficas do *TinkerPlots*, a cor, para explicar a existência dessa relação. Na figura 5.3.1, o grupo compreende o significado das cores associadas aos dados, recorrendo ao cartão dos dados. Este processo repete-se para a maioria das representações gráficas construídas, antes de passar para diagramas de dispersão. Desta forma, revela familiaridade com o ambiente tecnológico e compreensão das características do software.

Nesta altura, já não necessitaram de assistência para aceder às principais características do *TinkerPlots* ou na construção de representações gráficas. A maior parte das representações gráficas que criam são diagramas de dispersão e para justificar ou dar respostas às questões colocadas, utilizam frequentemente termos corretos como tendência, associação, relação, força, inclinação, reta, entre outras. Os diagramas de dispersão foram construídos intencionalmente pelos grupos, pois compreendem que esta representação permite estudar duas características conjuntamente e que o dados observados aparecem sob a forma de pares de valores, como um sistema de eixos coordenados: “cada altura corresponde um valor de peso, como um par “(TG10) e “entre a altura dos alunos e a resistência desse aluno há uma relação estatística acentuada entre as duas variáveis” (TG8).

O grupo 3 recorre a diagramas de dispersão para observar o tipo de associação entre as variáveis em estudo (figura 5.3.2). Para além de analisar a representação gráfica pela forma ou a direção dos plots. Este grupo recorre a médias, linhas de referência e o lápis para salientar a distribuição dos dados pelos quadrantes.

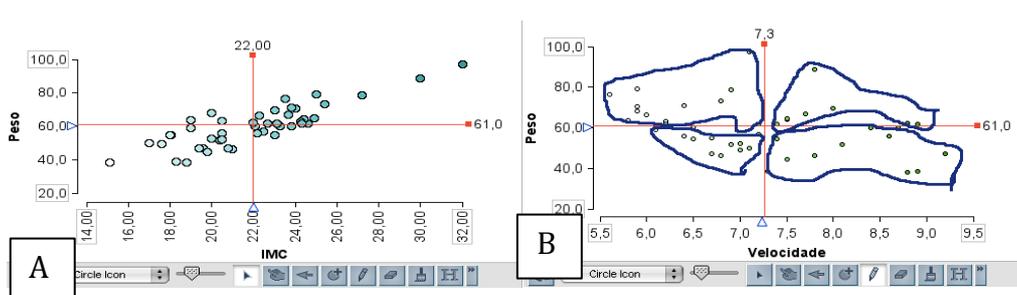


Figura 5.3. 2 – Gráficos de dispersão representando relações entre “Peso” com: “IMC” (A) e “Velocidade” (B) [G3].

Contudo, salienta-se que a linguagem utilizada pelos grupos 1, 5, 7 e 9, durante a resolução da tarefa, ao descrever os gráficos e as suas características, foi uma linguagem informal e incorreta, associada ao quotidiano: “esta bola é mais alta, porque tem os maiores valores” [G5] ou “as bolas estão a fugir umas das outras, não são todas iguais, podem ter maior velocidade e ter menor peso ou o contrário”[G9].

*Criatividade com os dados.* Os grupos foram capazes de criar gráficos para representar duas ou três variáveis, recorrendo ao gradiente de cor para quando fosse mais de duas variáveis, por forma a traduzir ou obter evidências para as suas conjeturas.

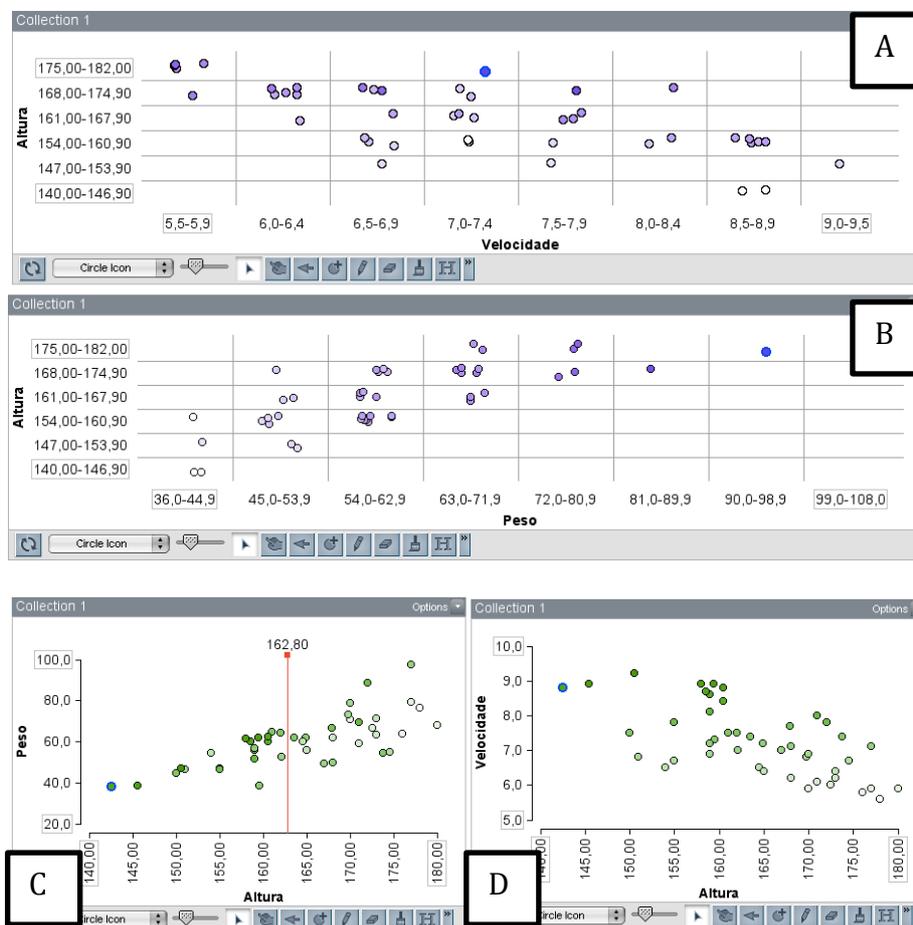


Figura 5.3. 3 – Representação gráfica com duas ou três variáveis (G1 – A e B; G3 – C e D).

Para provar a existência de relação entre a possibilidade dos mais altos serem os mais pesados e os mais rápidos, o grupo 1 limitou-se a construir gráficos de

covariação, em que ambos os eixos estão organizados em intervalos (figura 5.3.3).

A partir da descrição de dados desses gráficos construídos, afirmam que:

alunos com menos tempo [a percorrer os 60 minutos], entre 5,5 e 5,9, nem são os mais pesados nem os mais leves, quando olhamos para o peso... estão entre 63,0 e 80,9. Mas aqui valores que não batem certo com a lógica... o António que é um dos mais altos e dos mais pesados de todos, encontra-se entre 7,0 e 7,4 de velocidade. Mas devia estar errado, pois devia ser o mais lento, devido ao peso. [TG1]

No caso do grupo 3, descrevem os dados a partir da representação gráfica, focados numa margem identificada por uma linha de referência, que é evidente na frase:

podemos considerar que um grupo de pessoas com altura compreendidas entre 140 cm e 162 cm, são os que demoram mais tempo a percorrer os 60 metros. E o grupo compreendido entre 162 cm e 180 cm são os mais altos e os que demoram menos tempo a percorrer os 60 metros [TG3].

Para analisar esta possível relação, alguns grupos (G2, G3, G4, G6, G8, G9 e G10) criaram somente diagramas de dispersão, com diferentes atributos (figura 5.3.4).

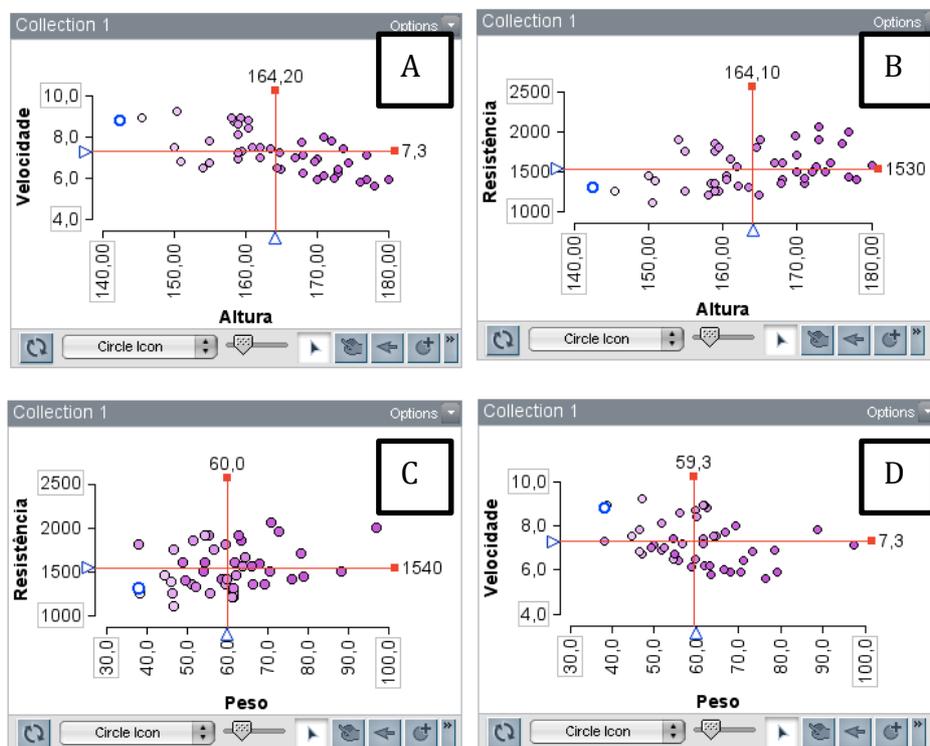


Figura 5.3. 4 – Gráficos de dispersão representando relações entre diversas variáveis [G10].

A criatividade do grupo é limitada a um tipo de gráfico, pois recorre só ao diagrama de dispersão, e quando questionados para o facto, dizem que “consequimos ver melhor a distribuição dos pontos... na inclinação e pela cor que têm” [TG10]. É de salientar que este grupo acrescentou a todos os gráficos linhas de referência para ler valores e determinar o centro de gravidade a partir das médias dos atributos.

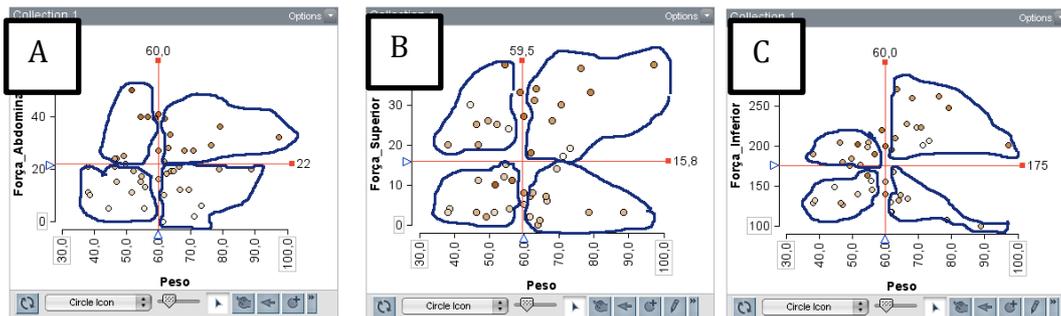


Figura 5.3. 5 – Representação gráfica representando relações entre diversas variáveis [G10].

Na generalidade, os alunos construíram diagramas de dispersão (figura 5.3.5) para obter evidências na identificação de relações. Para tal, o grupo 10 utilizou a função de desenho para destacar o número de pontos que estão distribuídos pelos quadrantes, ao fim de marcar o centro de gravidade através dos valores médios dos atributos. Ao usar esta ferramenta, agrupou os plots e afirmou que:

neste gráfico [figura 5.3.5A], nos quatro quadrantes, o número de grupo de crianças está distribuído da mesma maneira... a... no debaixo à esquerda [1.º quadrante] tem mais ou menos 13 pontos... no lado direito tem 12 pontos... no de cima à esquerda há 8 e no outro, há 10 pontos... bem... aqui os números tanto sobem como descem. Não são iguais. [TG10]

O uso da ferramenta de desenho ajudou este grupo a expor o seu raciocínio, descrevendo e resumindo, assim, os dados. Através de cálculos básicos a partir dos valores observados, o grupo 10 descreve os dados, resumindo-os, para mais tarde passar à sua interpretação.

*Compreender os dados.* A estrutura das representações gráficas que os grupos criaram influenciou na interpretação dos dados, quando foi pedido para descrever a possível existência de relações entre dois ou três variáveis.

Quando questionados “Será que o peso interfere, tanto no sexo masculino como feminino, na resistência que é a distância percorrida por cada aluno, durante 10 minutos?”, o grupo 8 constrói o gráfico da figura 5.3.6B e afirma que não se confirma porque “temos 3 rapazes com o peso mais baixo, nesses intervalos, só um que corre mais, o Marco. Ah... e mais uma rapariga” e “o mais alto, de todos nem é o mais pesado e nem o mais leve, ele está entre 63 e 71,9”. Em relação a esse gráfico e quando transformado para representar ambas as variáveis num eixo contínuo, o grupo 10 também rejeita a conjectura, justificando que os pontos vão oscilando e que “o baixo tanto percorre menos distância ou maior, o mesmo acontece para o mais alto e aqueles que estão a meio [gráfico 28B, discussão coletiva]”. Já para comparar rapazes e raparigas, o mesmo grupo afirma que “não é concreto” [TG10], porque os valores dos rapazes e das raparigas, nem sempre estão relacionados entre si, devido a outros fatores, como a prática desportiva que realiza diariamente.

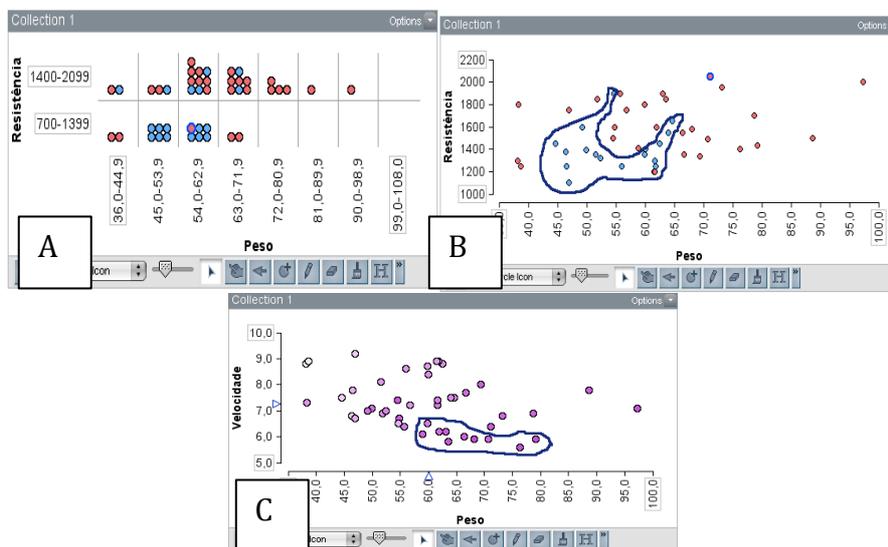


Figura 5.3. 6 – Representação gráfica representando relações entre variáveis (A – G8; B – G10 e C – G3).

Relativamente à questão “Os alunos mais altos tendem a ser os mais pesados e os mais rápidos”, houve uma diversidade de respostas pelos grupos. Aqui, os grupos

escolheram entre duas variáveis “resistência” ou “velocidade”, para comparar a rapidez dos alunos em estudo, com as outras duas variáveis “altura” e “peso”.

No trabalho com os gráficos, à exceção do grupo 7, os alunos consideram que a afirmação é verdadeira:

Segundo a representação, os alunos mais altos tendem a ser mais pesados e verificou-se que de facto os mais altos, apresentam uma menor velocidade, logo são os mais rápidos. [TC10]

Nesta análise, conclui-se que os alunos revelam compreensão das relações entre os dados e gráficos: “pelo gráfico que construímos conseguimos ver uma linha imaginária ascendente, logo quanto mais alto, mais pesado se encontra” [DC8] e “os mais rápidos são os mais altos, pois aqui o gráfico é descendente e conforme a altura aumenta, a velocidade vai diminuindo, que indica que percorrem os 60 metros em pouco tempo” [DC6]. Nesta discussão, os grupos descrevem o padrão geral de um diagrama de dispersão pela forma, direção e a intensidade da relação, ou seja, se a distribuição dos pontos observa-se numa associação linear que evidencia “uma reta”.

No entanto, o grupo 7 afirma que afirmação não é verdadeira, porque “o Gonçalo é o mais alto, mas não é o mais pesado e nem o mais rápido”. Aqui o grupo recorre a dados individuais do gráfico, para refutar esta afirmação. Para eles, o fato de haver uma exceção, já não permite a generalização.

Na procura de evidências numa possível relação entre o peso dos alunos e a sua força, os grupos afirmam que não existe tal associação (figura 5.3.5).

No gráfico do peso dos alunos com a força abdominal não existe nenhuma relação, os dados estão muito dispersos, mas os dados também se encontram mais no 1.º e 3.º quadrante, logo à tendência para que quando são mais pesados têm mais força. Na força inferior com o peso conseguimos ver alguma relação. Mas há exceções, conseguimos ver que temos uma linha imaginária ascendente, logo quanto mais pesado é, mais força inferior se tem. No gráfico do peso e a força superior, não encontramos nenhuma relação, os dados estão todos dispersos, que terá influência sim, pelo exercício que as pessoas fazem e os hábitos de vida. [TG8]

Pela descrição deste grupo, compreende-se que utilizaram o conjunto de dados de cada gráfico para tomar uma decisão quanto à possível relação entre o peso e a

força. E a estrutura do gráfico 5.3.5B, ao estar perto de uma ligeira inclinação, permite-lhes dar a ideia de que “os dados estão subindo juntos”.

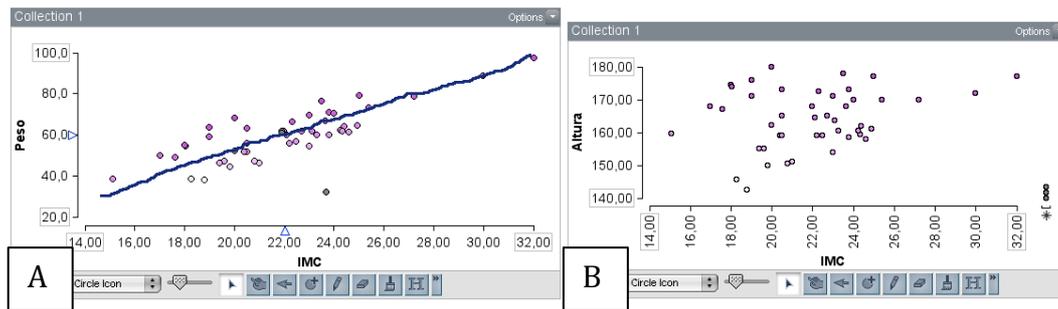


Figura 5.3. 7 – Diagramas de dispersão representando relações do “IMC” com a “Altura” (A- G10) e “Peso” (B – G8).

A partir das representações gráficas (gráfico 5.3.7), os grupos 8 e 10 afirmam que “o IMC ao aumentar, o peso dá a ideia de que vai aumentando” [TG8], pois “os dados estão mais agrupados no 1.º e 3.º quadrante e podemos traçar uma linha imaginária ascendente [TG10]”. Nesta análise, o grupo 10 foca-se não em dados individuais, mas em tendências globais do gráfico, quando constroem o gráfico 5.3.7B. Verificam que os dados estão mais dispersos, não “criando a ilusão de estarem agrupados sob uma linha imaginária” [DC10] e mostram que não há uma relação clara entre estas duas características e por isso, concluírem que estão fracamente relacionados. Aqui, o grupo 10 não faz a interpretação relativamente à intensidade da tal associação. Quando comparam as três variáveis, o grupo 3 afirma que “quanto mais tem IMC, são mais pesados, mas não são mais altos” e que “o peso e o IMC têm uma relação direta” [TG3], refutando assim, a ideia da existência de uma relação entre o IMC com a altura.

A partir do gráfico 5.3.5B, em discussão coletiva, os grupos referem que à medida que a altura vai aumentando, a nuvem de pontos “não forma uma reta perfeita, pois há mais dados no 1.º e 3.º quadrante, no entanto são dispersos e daí a relação é fraca” [TG 3]. Além disso, os alunos tendem a percorrer mais distância durante o tempo estipulado, contudo “não se verifica que para todos os pontos” [TG 3]. Aqui, os grupos recorrem ao ponto médio ou centro de gravidade da nuvem de pontos para, de forma análoga, associarem a sua resposta ao padrão da nuvem de pontos que pode ser aproximadamente modelada por uma reta. O facto dos valores das variáveis na nuvem de pontos estarem dispersos, sem nenhum padrão

definido, não permite afirmar que quanto maior a altura, maior será a resistência. Assim, o grupo 10 compreende que não existirá correlação linear entre essas variáveis, como acontece com a variável altura e velocidade.

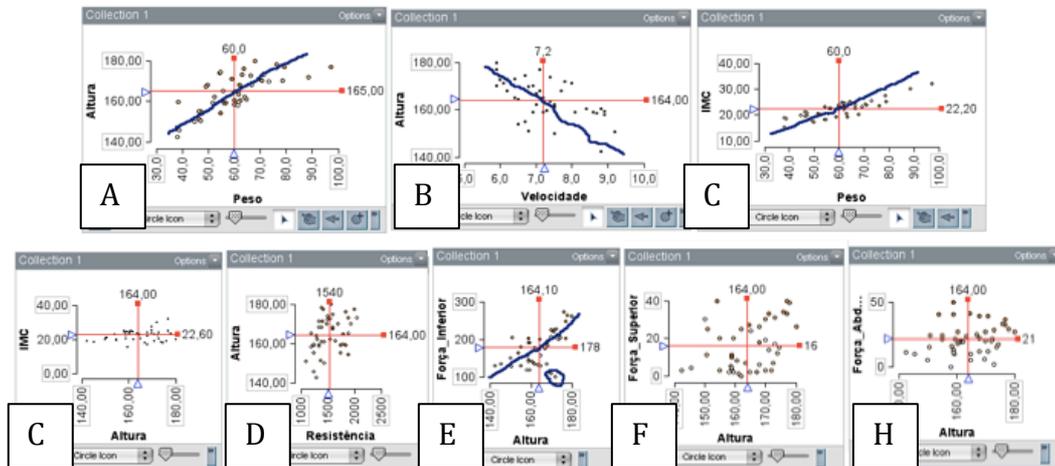


Figura 5.3. 8 – Representação gráfica representando relações entre diversas variáveis [DC].

A partir das representações gráficas construídas (figura 5.3.8), os grupos explicaram as características dos gráficos que permite estabelecer a existência de relação entre as variáveis: “a disposição dos pontos revela uma certa tendência, a altura aumenta à medida que a força inferior aumenta” [TG 10] e “verifica-se a tendência para a altura descer à medida que a velocidade aumenta” [TG 10]. Nestas respostas, o grupo 10 ao observar o diagrama de dispersão, intuitivamente, verificou a não existência de associação entre as duas variáveis em estudo. Aí, para quantificar o grau da associação linear entre as duas variáveis utilizaram uma estatística e interpretaram a intensidade e o sinal da covariação, com base no possível valor do coeficiente de correlação, relacionado com a linearidade/dispersão e com a própria orientação dos pontos que forma a nuvem de pontos:

No gráfico do peso com a altura, existe uma relação forte, que é uma relação positiva e aproxima-se de 1, pois os dados estão muito juntos uns aos outros. No gráfico do peso e resistência, não tem uma relação, logo é nula e o seu valor será 0. [TC 8]

O grupo 10 analisa a distribuição dos pontos e a sua aproximação/afastamento, estabelecendo depois uma tendência para a relação e sua intensidade, como se

pode apurar durante o momento da discussão coletiva das representações gráficas [figura 5.3.8]:

Professora: Daqui o que podemos concluir com todos estes gráficos?

G10: sabemos que o tipo de correlação vimos a partir da inclinação que os pontos estão distribuídos.

Professora: sim... e?

G8: só um é negativo...o segundo de cima [velocidade e altura]... eles estão a descer. E os outros são positivos.

Professora: Todos?

G10: stora, ele está errado... os de positiva, tem que estar os pontos próximos. Na fila de baixo, o gráfico 1, 4 e 5 os dados estão distribuídos aleatoriamente, sem orientação nenhuma... Logo, não tem inclinação positiva e nem negativa.

Professora: isso, em relação ao valor do coeficiente de correlação o que querera dizer?

G10: só pode ser zero. Nestes dois [altura vs peso e peso vs IMC], a inclinação é positiva, mas a ligação é mais forte no peso e IMC, porque os pontos estão próximos e alinhados formando uma linha imaginária.

Nesta discussão, verifica-se que os grupos compreendem o aspeto relacional dos diagramas de dispersão e a partir das características das representações gráficas identificam mensagens nos dados. Identificam que para estabelecer uma relação entre as variáveis, deve-se ter em conta a forma e a direção que os pontos apresentam num diagrama de dispersão, quer seja para uma associação positiva, negativa ou até, quando não há uma associação clara entre as variáveis. Da análise das diversas respostas, constata-se que a maioria dos grupos conclui que para avaliar a relação das variáveis, recorrem à nuvem de pontos e fazem referência aos pontos, quando estes se condensam à volta de uma reta.

*Pensar sobre os dados.* Os alunos revelaram familiaridade com este contexto e foram capazes de fazer conexões com a natureza dos dados, caracterizados pelos seus comentários: “a condição física dos alunos do 9.º ano é normal para a idade deles, em que existe uma boa relação entre o peso e a altura” [TC3]. As conjeturas formuladas pelos grupos têm por base a sua experiência no contexto. O grupo 10, ao afirmar inicialmente que “a altura só poderá ser influenciada pela velocidade”,

indicia que não recorre à natureza relacional dos diagramas de dispersão para justificar. Pois, para justificar o seu pensamento, afirma que:

uma pessoa mais baixa terá uma velocidade maior que uma pessoa mais alta, porque uma pessoa mais baixa [para] ganhar velocidade dá mais passadas, enquanto que uma pessoa mais alta dá menos passadas. [TG 10]

Por outro lado, este grupo para justificar a sua conjectura “será que os rapazes têm mais resistência que as raparigas?”, evoca a constituição física dos alunos: “Os rapazes podem ser mais resistentes que as raparigas, porque apresentam mais massa muscular que as raparigas” [TG 10].

Nalgumas representações gráficas, quando ocorrem discrepâncias em relação ao esperado a nível da relação entre as variáveis em estudo, os grupos tendem a justificar e procurar causas com base no seu conhecimento do contexto: “o aluno apesar de ser alto, pode correr menos, porque a altura das pernas é pequena em relação à altura do tronco” [TG 6] e “o aluno ao ser pesado, ele vai cansar-se mais rápido e demora mais tempo” [TG 2].

As explicações e as descrições das representações gráficas construídas, mostram que os grupos conseguiram avaliar a existência de covariação a partir da distribuição dos dados, interpretando as representações gráficas e articulando-as com o conhecimento que têm do contexto:

Os valores obtidos dizem que dependem da condição física, uns podem ter uma força mais desenvolvida que a outra, mesmo que seja mais altos podem ser mais pesados e não irão ter resultados positivos. [TG8]

Quando foi proposto descreverem a forma como o gráfico mostrava a relação, o grupo 2 respondeu: “ele está a ir para cima, como se fosse uma diagonal [desenhando uma linha crescente]” [TG2], evidenciando serem capazes de interpretar tendências globais nos dados.

A realização das inferências causais baseadas nos dados é pouco evidenciada. Só os grupos 3 e 10, é que descrevem uma regularidade nos dados e especulam a ocorrência das mesmas relações para alunos mais velhos, do 11.º ano:

Alunos do 11.º ano encontram-se mais desenvolvidos, a nível da massa muscular e pressupõe-se que sejam mais altos e segundo as representações gráficas, os alunos mais rápidos, são os mais altos, por isso, este facto é evidente. Ou seja, os

alunos do 11.º ano serão os mais altos, mas poderá ser influenciado por vários factores, sendo um deles, a prática desportiva regular. Mas pode ser também genético, e manter sempre as mesmas características relacionais. [TG 10]

## **SÍNTESE**

Nesta tarefa, os grupos criaram, praticamente, só um tipo de representação gráfica, uma vez que reconhecem a importância da construção do diagrama de dispersão para compreender o aspeto relacional das variáveis em estudo. Por isso, a sua construção foi autónoma e intencional para a compreensão da covariação e a avaliação de relações entre as variáveis. A maioria dos grupos conseguiu articular as representações gráficas criadas com os dados disponibilizados pelos cartões, que foram solicitados constantemente para a identificação de mensagens a partir dos dados. A nível dos recursos utilizados nos gráficos, estes grupos recorreram ao ponto médio resultante da interseção das médias das duas variáveis e daí, retirar as suas conclusões relativamente à forma, intensidade e distribuição, para inferir a associação entre as duas variáveis. Alguns grupos identificaram a existência da covariação, tanto quando duas variáveis se movem na mesma direção ou em direções opostas. Foram, ainda capazes de descrever e interpretar os diagramas de dispersão, focando-se tanto em características individuais como globais dos dados, identificando a tendência dos dados. Outro aspeto a salientar é a capacidade de alguns grupos quantificarem o grau da associação linear entre duas variáveis representadas em diagramas de dispersão.

## **5.4. PROVA DE AVALIAÇÃO FINAL**

A Prova final (em anexo) foi realizada individualmente com o objetivo de verificar quais as aprendizagens evidenciadas pelos alunos e as dificuldades que revelam, depois da realização da unidade de ensino, no que respeita à compreensão de covariação.

A Prova é constituída por 6 questões, sendo quatro dessas questões (questão 3, 4, 5 e 6) focadas no tópico das distribuições bidimensionais. Na resolução da prova os alunos poderiam utilizar a calculadora gráfica, também usada durante a unidade de ensino, facilitando a representação os dados num diagrama de dispersão e a determinação de alguns dos resultados pedidos.

A prova foi realizada numa folha de resposta e teve duração 60 minutos. As respostas dos alunos a estas questões foram analisadas relativamente ao objetivos de aprendizagens propostos para o tópico das distribuições bidimensionais. Para tal, foca-se em como os alunos compreendem a relação entre os dados e os gráficos e as mensagens a ser retiradas deles, bem como, interpretar e fazer inferências causais baseadas nos dados.

Na questão 3 averiguou-se a capacidade dos alunos lerem o gráfico de dispersão, estimar o valor de uma das variáveis a partir do valor dado de outra variável e avaliar o grau de associação entre as duas variáveis apresentadas no gráfico de dispersão. Na primeira alínea da questão, 80% dos alunos conseguiram identificar a jogadora que marcou mais pontos e que “apesar de ter mais pontos, não foi a que teve mais tempo em jogo” [PA8].

Na segunda alínea, observa-se uma variedade de respostas dadas pelos alunos, quando identificam e justificam a possível existência de uma associação entre o tempo jogado por cada atleta e os pontos marcados. Cerca de 50% dos alunos avaliaram, de forma intuitiva a partir da análise do diagrama de dispersão, a existência de uma associação linear entre as duas variáveis, minutos jogados e total de pontos marcados. Aí, para saber se existia alguma relação entre as duas variáveis, recorreram à distribuição dos pontos no diagrama de dispersão: “É possível existir, porque à partida o jogador com mais pontos marcados foi o que teve mais tempo no jogo” [PA10]. Nesta questão, houve 6 alunos que para justificar essa possível associação, recorreram a uma reta que se ajusta à nuvem de pontos, ou seja, como justificação focaram-se na ideia de que os dados concentram-se ao longo de uma “linha” com uma forma alongada, significando que ao aumentarem os valores de uma das variáveis, os da outra aumentam também: “os pontos estão a crescer aproximados a uma linha imaginária crescente, ou seja, à medida que aumenta os minutos jogados, o total de pontos também aumenta” [PA3]. Dois alunos, utilizaram como estratégia de resolução, a

observação da distribuição dos pontos nos quadrantes do diagrama de dispersão, partindo do centro de gravidade, obtido pela média calculada em cada variável e aí, analisaram a forma como os pontos estavam distribuídos nos quadrantes que resultaram da interseção das médias das duas variáveis: “há mais pontos nos quadrantes I e III, logo ocorre uma relação positiva entre o tempo de jogo e o total de pontos marcados” [PA8].

Cinco alunos, confirmaram que havia relação entre as duas variáveis, mas recorreram simplesmente à dispersão dos pontos na representação gráfica, sem referir o que acontece a uma das variáveis quando a outra aumenta ou diminui. Isto é, justificaram que maior ou menor dispersão entre os pontos, implica uma forte ou fraca associação entre as variáveis, sem estabelecer a ligação entre os valores das variáveis, como é exemplificado na figura 5.4.1.

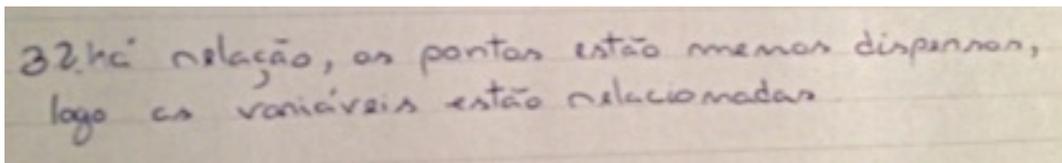


Figura 5.4. 1 – Resposta à questão 3.2 da prova de avaliação [aluna do G5].

27% dos restantes 50% dos alunos deram respostas incorretas, pois basearam-se em pontos isolados para justificar a associação: “não é possível existir relação, porque não é igual. O jogador com 34 minutos jogados, marca 18 pontos como o jogador de 44 minutos” [PA2] e “o [jogador] de 60 minutos marca menos 4 pontos que o jogador de 53 minutos” [PA1]. Dos alunos que deram resposta incorreta, 14% deles recorreram às suas crenças prévias relativamente à relação entre o número de pontos marcados e o tempo de jogo, em vez de recorrer às evidências dos dados que estavam disponíveis no diagrama de dispersão: “há relação, porque um jogador que jogar mais tempo, irá marcar mais pontos do que um jogador que estiver menos tempo” [PA2]. Outros, ainda, referem que a associação entre as variáveis poderá resultar da influência de outras variáveis na variável dependente: “Não é possível dizer que há relação, porque os pontos marcados podem influenciar na flexibilidade do jogador em causa. O jogador pode jogar muito tempo, e como é bastante flexível consegue marcar mais poucos ou o contrário” [PA7].

Nas questões 3.3., 4.2, 4.3 e 5, os alunos deveriam prever o valor de uma variável a partir do conhecimento do valor de outra variável. Verifica-se que para os alunos foi fácil encontrar esse valor na questão 5, uma vez que nesse gráfico os alunos tinham representado a reta de regressão. Nessa questão, os alunos recorreram às retas paralelas aos eixos coordenados para poder estimar os pontos de abscissa 25. Relativamente à questão 3.3, só três alunos (que pertenciam aos grupos 3, 8 e 10) responderam corretamente à questão. Estes alunos, através da observação dos dados no diagrama de dispersão, estimaram os valores das variáveis e obtiveram a equação da reta de regressão, onde estimaram o número de pontos que equivaleria aos 50 minutos. Como mostra a figura 5.4.2, três alunos (que pertenciam aos grupos 3, 8 e 10) identificaram os valores das variáveis em tabela e calcularam analiticamente as médias das duas variáveis, para determinar o centro de gravidade da distribuição. Aí, recorreram às funcionalidades da máquina de calcular para determinar a expressão da reta de regressão e substituíram o  $x$  por 50 (Figura 5.4.2).

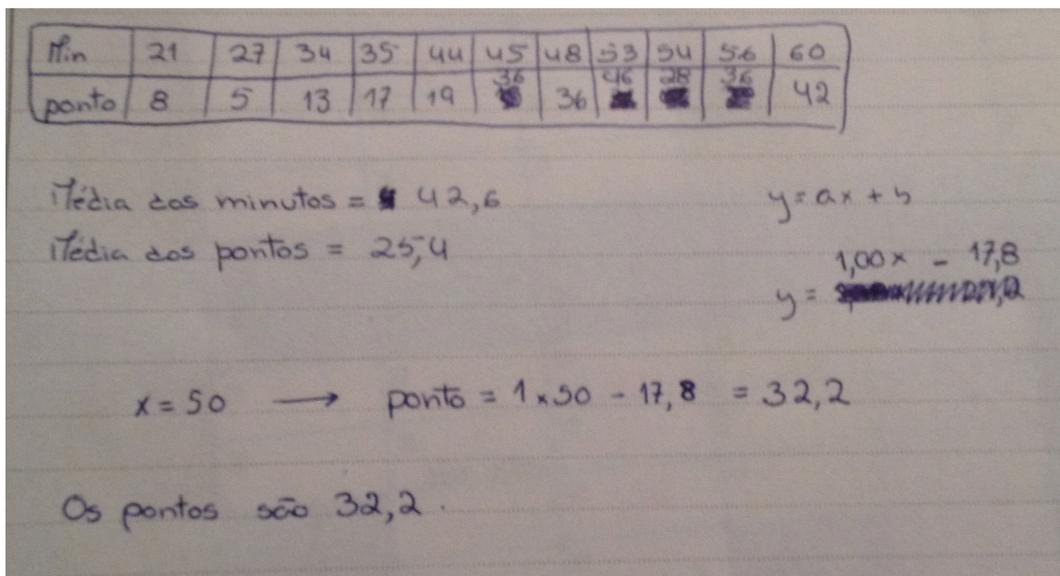


Figura 5.4. 2 – Resposta à questão 3.3. da prova de avaliação [aluna do G10].

Constata-se, ainda, que alguns alunos dos grupos (G10, G6, G2 e G9) estimaram para o número de pontos do jogador que tinha jogado durante 50 minutos, não um valor específico, mas num conjunto de valores, associado a um intervalo (Figura 5.43).

3.3.  
~~///~~  $(\bar{x}, \bar{y}) = (42,6; 25,4)$  Se o ponto médio é 25,4 e ao traçar no gráfico o centro e fazer uma reta que passe no quadrante 1º e 3º, os valores possíveis estão num intervalo. o mínimo é acima dos 30 e o máximo 35, por causa da distribuição dos pontos no gráfico.

Figura 5.4. 3 – Resposta à questão 3.3. da prova de avaliação [aluno do G2].

Cerca de 27% dos alunos não respondeu a esta questão ou então afirmou que “é impossível estimar o valor do número de pontos, porque não há relação linear entre o tempo do jogo e o número de pontos obtidos” [PA7].

Relativamente à questão 4.1, pretendia que os alunos analisassem a relação entre duas variáveis em estudo, a idade e altura de um rapaz. Todos os alunos construíram corretamente um diagrama de dispersão a partir de dados representados em tabelas e, a partir da sua análise, respondem que se trata de uma relação muito forte entre duas variáveis ao atenderem à distribuição dos dados:

Acho que na idade e altura deste rapaz há uma relação. À medida que a idade avança, a sua altura começa a aumentar, em que os pontos estão distribuídos no gráfico de uma forma alongada, próxima de uma linha reta crescente, se os unir. A correlação entre as variáveis analisadas é forte, pois parece quase uma função afim e quando está nesta forma, o valor do coeficiente é próximo de 1. Assim, a idade elevada exerce naturalmente influência sobre a sua altura. [PA10]

Alguns alunos (G8, G6, G3 e G2) estabelecem que existe relação afirmando que “as variáveis estão fortemente correlacionadas no sentido positivo, pois os pontos do gráfico estão muito próximos dos outros, que até dá para formar uma reta (reta de regressão) que está a crescer” [PA2], logo “as duas variáveis (idade e altura) estão dependentes uma da outra” [PA8].

Quanto à questão 6, os alunos associaram corretamente o gráfico de dispersão ao tipo de correlação linear correspondente. Contudo, um aspeto que ainda é ainda não está totalmente superado, é quando se pede para justificar a razão associar dessa associação. Alguns alunos, não justificam as suas respostas. Na análise e interpretação do coeficiente de correlação, os alunos descreveram o coeficiente de correlação em termos de sinal e força/intensidade das relações entre as variáveis

representadas. Desta forma, as suas respostas tiveram por base a dispersão/concentração dos dados na representação gráfica, recorrendo à interpretação geométrica dos dados:

O gráfico 3 e 4, ambos estão a crescer, pois os valores de  $x$  estão a aumentar e os valores de  $y$  também. Contudo, a nível da dispersão há diferenças. O gráfico 3, os pontos estão mais dispersos que no gráfico 4, enquanto nesse, os pontos estão quase alinhados em torno de uma reta. Assim, 4 é positiva forte e o outro é positiva fraca. O gráfico 6 é negativa fraca, porque os pontos estão um pouco mais afastados que no gráfico 1. E ela é negativa, porque uma variável aumenta e a outra [tem] a tendência a diminuir. [PA10]

Com esta resposta, compreende-se que o aluno consegue interpretar a intensidade e a direção, quer seja positiva ou negativa, entre duas variáveis, relacionando-a com a linearidade/ dispersão e com a orientação dos pontos obtidos no diagrama de dispersão.

Na avaliação da associação estatística, a maior parte dos alunos recorreu à tendência de crescimento dos dados representados no diagrama de dispersão para justificar a dependência das variáveis ou à forma desses dados estarem em torno de uma linha reta. Contudo, destaca-se que alguns alunos (27%) ainda baseiam os seus julgamentos em dados individuais ou no seu conhecimento do contexto.

## **CAPÍTULO 6**

### **CONCLUSÕES E REFLEXÃO FINAL**

Neste capítulo começo por apresentar uma síntese do estudo e as suas principais conclusões no que diz respeito ao raciocínio covariacional dos alunos.

Para concluir, faço uma reflexão pessoal sobre os seus contributos para o meu desenvolvimento pessoal e profissional e apresento algumas sugestões para estudos que se venham a realizar sobre esta temática.

#### **6.1. SÍNTESE DO ESTUDO**

Tendo em conta a importância de desenvolver a literacia estatística dos alunos, defendida nas orientações curriculares, o meu gosto por este tema da Matemática e a vontade de melhorar a minha prática letiva, ajudando os alunos a ultrapassarem as reconhecidas dificuldades na aprendizagem da covariação estatística, realizei uma experiência de ensino no tópico das distribuições bidimensionais. O estudo foi desenvolvido numa turma do 10.º ano dos cursos profissionais de Técnico de Auxiliar de Saúde (TAS) e de Técnico de Apoio à Gestão Desportiva (TAGD), em contexto de sala de aula, suportado na realização de tarefas de natureza exploratória e com recurso ao *software TinkerPlots*.

Com este objetivo, procurei responder às seguintes questões:

- 1) Como é que os alunos exploram relações entre variáveis e estabelecem a existência de covariação, quando utilizam o *TinkerPlots* na resolução de tarefas envolvendo este conceito? Em particular, quais as representações

estatísticas que utilizam e qual a sua adequabilidade no estabelecimento dessa relação?

- 2) Quais as aprendizagens realizadas pelos alunos no final da unidade de ensino, no que respeita à covariação? Que dificuldades manifestam na compreensão do conceito?

Para recolher os dados recorri à observação participante, gravações áudio e vídeo das aulas, registo das ações dos alunos no ecrã do computador pelo *software AutoScreenRecorder 3.1 Pro* e elaboração de notas de campo, bem como a documentos escritos, nomeadamente as resoluções escritas dos alunos das tarefas, nos ficheiros do *TinkerPlots* e da prova final do módulo A3 – Estatística.

## 6.2. PRINCIPAIS CONCLUSÕES

Neste ponto apresento as principais conclusões deste estudo, organizadas segundo as questões de partida.

*Como é que os alunos exploram relações entre variáveis e estabelecem a existência de covariação, quando utilizam o TinkerPlots na resolução de tarefas envolvendo este conceito? Em particular, quais as representações estatísticas que utilizam e qual a sua adequabilidade no estabelecimento dessa relação?*

Os alunos, de modo geral, usaram apropriadamente as diferentes representações gráficas que construíram no *TinkerPlots* para evidenciar a existência de relações entre duas variáveis e suportar as conjeturas formuladas.

Os alunos começaram por gerar, de modo criativo mas limitado, várias representações como gráficos de pontos, gráficos de barras e diagramas de extremos e quartis. No entanto, estes gráficos limitam a descrição dos dados com foco no conceito de covariação estatística. Alguns alunos, antes de utilizarem o diagrama de dispersão para caracterizar a natureza da relação entre as variáveis, partiam de uma representação gráfica de pontos, com escalas divididas em intervalos, onde resumiam e organizavam os dados bivariados. Ao longo do trabalho realizado nas tarefas, os alunos reconheceram intuitivamente que, para verificar o aspeto relacional das variáveis, seria necessário atender ao aspeto da

dispersão dos dados, nomeadamente a sua forma e direção, cuja visualização é facilitada pelo diagrama de dispersão. Desta forma reconheceram, intuitivamente, que o diagrama de dispersão é a representação gráfica adequada para explorar a associação entre dois atributos, tal como tinha ocorrido nos estudos de Fitzallen (2012) e Mortiz (2004).

A partir dos diagramas de dispersão, os alunos puderam verificar o grau de associação entre duas variáveis aleatórias de acordo com a forma da distribuição dos dados. É de salientar, que o uso das ferramentas do *TinkerPlots* (incluindo o gradiente de cor que é uma característica particular deste software), ajudou na observação de uma grande multiplicidade de variáveis representadas em diagramas de dispersão permitindo estabelecer associações entre as variáveis, como referiu Konold (2006).

Na interpretação da associação estatística, com base nos gráficos de dispersão, houve alunos que revelaram dificuldades em associar ao padrão global de variação da nuvem de pontos, a existência de covariação positiva, negativa ou ausência de covariação, bem como quantificar o grau dessa associação estatística, tal como verificado em outros estudos que não tiveram o apoio do *software* (Batanero, et al., 1997). Apesar disso, a maioria dos alunos compreendeu que podem identificar se dois atributos variam de forma conjunta, recorrendo à distribuição dos *plots* no diagrama de dispersão, focando-os na distribuição dos dados pelos quadrantes. Assim, quando a maioria dos pontos estavam nos quadrantes I e III, formados pelos valores médios das variáveis, os alunos estabeleceram a existência de associação estatística, classificando-a como covariação positiva, recorrendo à ideia de que quando uma cresce, a outra também cresce. O mesmo aconteceu quando as duas variáveis se relacionavam de forma negativa. Ainda com base no diagrama de dispersão, os alunos compreenderam que não havia relação linear entre as variáveis quando os dados estavam todos dispersos. Este resultado mostra que é possível superar a dificuldade mencionada por Batanero, et al., (1997), que os alunos só compreendiam a associação positiva.

Todos os grupos conseguiram articular as representações gráficas criadas com os dados disponibilizados nos cartões. A apresentação simultânea, no écran do *TinkerPlots*, dos cartões e das representações gráficas criadas durante a resolução

das tarefas, facilitou a identificação de mensagens a partir dos dados, aspecto já referido por Fitzallen (2012).

Inicialmente, os grupos focaram-se em dados individuais (conceção local) das representações gráficas para avaliar e suportar as suas inferências, quanto à dispersão e identificação de tendências nos dados, dificuldade identificada também em Ben-Zvi e Arcavi (2001). No entanto, o aumento da familiaridade dos alunos com o *software* e o trabalho desenvolvido nas tarefas permitiu que os alunos passassem a focar-se também na tendência global dos dados, identificando mensagens nos dados a partir também das representações gráficas.

Nesta altura, quando os dados individuais evidenciavam exceções à tendência global dos dados, os alunos procuravam justificações no seu conhecimento do contexto de dados e nas suas vivências do dia-a-dia.

Neste estudo, a seleção de contextos pouco familiar aos alunos, não funcionou como desafio para “interrogar” os dados e tirar conclusões a partir das representações gráficas criadas, como sugeria Fitzallen (2012). Na verdade, a tarefa “indicadores demográficos” foi a que trouxe maiores dificuldades aos alunos, no que respeita à formulação de inferências ou questões, uma vez que os alunos não estavam familiarizados com o contexto e ao procurar relações entre os dados, focaram-se unicamente no significado de cada variável. As outras duas tarefas, contextualizadas em situações da vida real dos alunos, devido às áreas curriculares do curso, fez emergir interesse e motivação por parte dos grupos e a procura de relações entre os dados em estudo, dando-lhes significado na procura de possíveis causas de variação nos padrões identificados, como verificado noutros estudos (Batanero, 2000; Fitzallen & Watson, 2011; Hall, 2011).

*Quais as aprendizagens realizadas pelos alunos no final da unidade de ensino, no que respeita à covariação? Que dificuldades manifestam na compreensão do conceito?*

No final desta unidade de ensino, é evidente que a maioria dos alunos identifica e compreende a covariação entre duas variáveis. Para além de conseguirem construir diagramas de dispersão a partir de dados também representados em tabelas, os alunos conseguem interpretá-los. A maioria dos alunos consegue avaliar, de forma intuitiva, a existência de uma associação linear entre duas variáveis, recorrendo à distribuição dos pontos no diagrama de dispersão. A partir

a distribuição dos *plots*, e quando os dados se concentram ao longo de uma “linha”, os alunos conseguem interpretar a intensidade e a direção, quer positiva ou negativa, da associação estatística entre as duas variáveis, relacionando a linearidade ou dispersão.

Os alunos revelam uma evolução significativa na utilização do diagrama de dispersão para retirar mensagens dos dados e compreender relações entre eles. No entanto, identifiquei ainda alguns erros que se mantêm na interpretação deste tipo de representação e dificuldades em verbalizar as suas conjeturas sobre a relação entre variáveis. Os alunos apresentam dificuldades a justificar o efeito sobre o valor central decorrente das alterações de alguns dados no conjunto, bem como distinguir de que a correlação estatística pode ser uma relação de causa-efeito e saber aplica-la corretamente. Ou seja, os alunos identificam que as variáveis estão a crescer, mas não entendem que podem ser num certo intervalo de tempo e não, porque uma é a causa da outra, como se verificou na tarefa – “Indicadores demográficos”, as variáveis apresentavam uma relação de crescimento, mas por estar toda a economia do país em crescimento. Para alguns alunos, apresentam dificuldades a nível da representação e a interpretação gráfica, no entendimento de escalas e a própria construção.

No final da unidade de ensino, a maioria dos alunos é capaz de identificar relações entre variáveis com base na observação da distribuição dos pontos nos quadrantes do diagrama de dispersão, partindo do centro de gravidade. Para efetuarem julgamentos de associação estatística em situações representadas em diagramas de dispersão, a maioria dos alunos usa o crescimento, decrescimento ou a forma constante da nuvem de pontos do diagrama de dispersão para justificar o tipo de dependência, e a sua dispersão para verificarem se é forte ou não.

Algumas dificuldades que persistem no final da unidade de ensino, como se pode constatar na prova final. Uma delas é a justificação da existência de associação entre variáveis. Apesar de alguns alunos conseguirem associar corretamente o gráfico de dispersão ao tipo de correlação linear correspondente, para justificarem as suas respostas recorrem a dados isolados e não a uma tendência global. Outra dificuldade verificada, é a realização de inferências causais baseadas nos dados. Os alunos, frequentemente, interpretam os dados, pelo menos de uma forma parcial, no conhecimento que têm sobre o contexto em causa, além da interpretação

direta da distribuição do diagrama de dispersão, focando a linearidade/dispersão, intensidade e direção. Outra situação, que alguns alunos mostram como dificuldade é de não admitir exceções na relação entre as variáveis, mantendo uma conceção determinista (Ben-Zvi & Arcavi, 2001; Chapman & Chapman, 1983; Estepa & Batanero, 1996).

### **6.3. REFLEXÃO FINAL**

A realização deste trabalho constituiu um importante momento de aprendizagem para mim, dando-me a oportunidade de refletir e aprofundar conhecimentos sobre o ensino e a aprendizagem da Estatística. Permitiu-me, por exemplo, contactar com o que as orientações curriculares e a investigação mais recente recomenda no que respeita ao processo de ensino-aprendizagem da Estatística e a compreender melhor como os alunos aprendem e as dificuldades que habitualmente apresentam, e vivenciando-o na realização de uma experiência de ensino.

Em termos globais, considero que os alunos realizaram muitas das aprendizagens pretendidas, embora não revelem sempre consistência na aprendizagem, registando-se avanços e recuos ao longo da realização das tarefas. O trabalho de grupo, em sala de aula, foi importante porque permitiu aos alunos tomarem consciência da sua responsabilidade na resolução da tarefa e para compreenderem a importância da partilha dos conhecimentos para a conclusão das tarefas com sucesso. As discussões, quer nos grupos, quer em grupo-turma, permitiram aos alunos expor as suas estratégias e aprender com as estratégias dos colegas, bem como desenvolver a comunicação, ao justificarem as suas estratégias e o seu raciocínio quando formularam conjecturas e tentaram generalizá-las. Os resultados evidenciam, também, que a construção e interpretação das representações gráficas, facilitadas pelo uso do *software TinkerPlots*, promoveram o desenvolvimento de ideias e conceitos estatísticos.

Além disso, o facto de a turma não estar familiarizada com as estratégias adotadas na experiência de ensino teve implicações no seu desempenho. Por exemplo, não

estavam habituados a justificar as suas respostas e nem a refletir sobre os significados dos conceitos. Mas esta situação é compreensível, uma vez que os alunos não estavam habituados a discutir sobre as formas de resolver as tarefas, nem a explicar as mesmas, que fez com que as discussões geradas na sala de aula não fossem tão ricas como pretendia. Com salienta Carvalho (2009), é importante que os alunos compreendam que a elaboração de uma resposta envolve a apresentação e explicação da origem dos passos dados. Na minha perspetiva, considero importante continuar a trabalhar com os alunos a apresentação e explicação de argumentos válidos para sustentar as suas respostas, bem como a utilizar representações para que as relacionem e interpretem os dados a partir delas.

Ao longo dos diversos momentos da minha intervenção pedagógica, senti algumas dificuldades, quer à turma pelas suas características e à minha prática. O fato da turma não ser acessível, não permitia desenvolver a estratégia delineada da melhor forma, uma vez que os alunos mostram alguma resistência em participar, expor as suas ideias e até mesmo, justificar os seus raciocínios efetuados na resolução das tarefas. Relativamente à minha prática, as minhas dificuldades estiveram em parte relacionadas com a escolha das tarefas que pretendia utilizar no estudo, para que pudessem corresponder aos objetivos definidos, onde os alunos pudessem utilizar os seus conhecimentos prévios para aprendizagem de novos conhecimentos e que nesse processo, a informação estivesse veiculada através das representações gráficas. Em determinadas alturas, não foi fácil, conseguir conjugar essas intenções, porque as tarefas tiveram ser modificados/alteradas para conseguir atingir o propósito deste estudo. Apesar destes constrangimentos, permitiu compreender, que é fundamental conjugar diversas fontes de informação e adaptá-las, conforme os objetivos de cada aula.

A minha maior dificuldade na realização deste estudo, diz respeito ao meu papel de investigadora e ao desafio que foi o processo de análise dos dados, pois nem sempre foi fácil gerir a quantidade de dados e analisá-los de acordo com a categorização, mediante as respostas dadas pelos alunos.

Este estudo constituiu, sem dúvida, uma grande valia para o meu desempenho enquanto professora, por me ter permitido compreender e refletir mais profundamente sobre a sequência de aprendizagem e sobre as dificuldades dos

alunos na aprendizagem dos conceitos das distribuições bidimensionais, no ensino da estatística. Além disso, provocou a curiosidade de saber mais sobre os processos de aprendizagem dos alunos noutras áreas da Matemática.

Ao desempenhar o papel de professora-investigadora, revelou-se uma tarefa um pouco difícil, uma vez que alguns pormenores importantes passaram despercebidos durante a minha observação. Foi difícil dar explicações aos alunos e, ao mesmo tempo, tirar notas sistematicamente sobre o que observava e ouvia, enquanto os alunos realizavam as tarefas e as discutiam.

É importante a realização de mais estudos sobre esta temática para fornecer informações que ajudem os professores a proporcionar ambientes de aprendizagem que permitam ultrapassar as dificuldades sentidas pelos alunos na aprendizagem da Estatística e, em particular, nas distribuições bidimensionais.

Como recomendações para estudos futuros aconselho que se valorize mais o trabalho dos alunos, na exploração dos conceitos estatísticos através das representações gráficas, recorrendo a tarefas de contexto real, que permite fazer inferências baseadas nos dados por estarem fortemente relacionadas com o seu conhecimento do contexto dos dados, por forma a motivar os alunos na realização das atividades e para que reconheçam a aplicabilidade da Estatística no quotidiano.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Abrantes, P. (1994). O trabalho de projeto e a relação dos alunos com a Matemática: a experiência do Projecto Mat<sub>789</sub>. (Tese de Doutoramento, Universidade de Lisboa). Lisboa: APM
- Abrantes, P., Serrazina, L. & Oliveira, I. (1999). *A Matemática na educação básica*. Lisboa: Ministério da Educação.
- Ainley, J., Pratt, D., & Nardi, E. (2001). Normalizing: children's activity to construct meanings for trend. *Educational Studies in Mathematics*, 45, 131–146.
- Aires, Luísa (2011). *Paradigma Qualitativo e Práticas de Investigação Educacional*, Lisboa: Universidade Aberta.
- Alias, M. (2009). Integrating technology into classroom instructions for reduced misconceptions in statistic. *International Electronic Journal of Mathematics Education*, 4 (2), 77-91.
- Amado, N., & Carreira, S. (2008). Utilização pedagógica do computador por professores estagiários de Matemática – diferenças na prática da sala de aula. In A. P. Canavarró, D. Moreira & M. I. Rocha (Orgs), *Tecnologias e educação matemática* (pp. 286-299). Secção de Educação Matemática da Sociedade Portuguesa de Ciências da Educação.
- Amado, N. (2007). *O professor estagiário de Matemática e a integração das tecnologias na sala de aula – Relações de mentoring numa constelação de prática*. (Tese de Doutoramento). Lisboa: APM.
- American Statistical Association (ASA) (2005). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: a Pre-K-12 Curriculum Framework*. Alexandria: American Statistical Association.
- Andrade, M. M. (2008). *Ensino e aprendizagem de estatística por meio da modelagem matemática: uma investigação com ensino médio* (Dissertação de Mestrado em Educação Matemática – Universidade Estadual Paulista, Rio Claro).
- Anghileri, J. (2006). Scaffolding practices that enhance mathematics learning. *Journal of Mathematics Teacher Education*, 9, 33-52.
- Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht, the Netherlands: CD Beta Press.
- Bakker, A., Derry, J., & Konold, C. (2006). Using technology to support diagrammatic reasoning about center and variation. In Proceedings of the 7<sup>th</sup> International Conference on Teaching Statistics – ICOTS, Salvador. Disponível em: <http://iase-web.org/documents/papers/icots7/2D4:BAKK.pdf>

- Balachowski, M. M. (1998). Trends in the statistics Classroom since NCTM standards. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee & W. K. Wong (Eds), *Proceedings of the Fifth International Conference on Teaching of Statistics* (vol. 1, pp. 75-76). Vooburg: International Statistical Institute.
- Batanero, C. (2000). Dificultades de los estudiantes en los conceptos estadísticos elementales: el caso de las medidas de posición central. In C. Loureiro, F. Oliveira & L. Brunheira (Orgs.), *Ensino e aprendizagem da estatística* (pp. 31-48). Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamentos de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.
- Batanero, C. (2001). Dificultades de los estudiantes en los conceptos estadísticos elementales: el caso de las medidas de posición central. In C. Loureiro, F. Oliveira & L. Brunheira (Orgs), *Ensino e aprendizagem da estatística* (pp. 31-48). Lisboa: SPE,APM, FCUL.
- Batanero, C., & Díaz, C. (2010). Training Teachers to teach statistics: What can we learn from research? *Statistique et enseignement*, 1(1), 5-20. Online: <http://statistique-et-enseignement.fr/ojs/>
- Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics: Proceedings of the 1996 IASE Round Table Conference* (pp. 191-205). Voorburg, The Netherlands: International Statistical Institute.
- Batanero, C., & Godino, J. D. (1998). Understanding graphical and numerical representations of statistical association in a computer environment. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee & W. Wong (Eds.). *Proceedings of the Fifth Conference on Teaching Statistics* (vol. 2, pp. 1017-1024). Voorburg, The Netherlands: International Statistical Institute.
- Bell, J. (2004). *Como Realizar um Projecto de Investigação*. Lisboa: Gradiva.
- Ben-Zvi, D. (2004). Reasoning about data analysis. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 121-145). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Ben-Zvi, D. (2006). *Scaffolding students informal inference and argumentation*. (Disponível em [http://www.researchgate.net/publication/228637727\\_Scaffolding\\_students\\_informal\\_inference\\_and\\_argumentation](http://www.researchgate.net/publication/228637727_Scaffolding_students_informal_inference_and_argumentation), accedido em 10 de junho de 2013).
- Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, 45, 35-65.
- Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning and thinking: goals, definitions and challenges. In D. Ben-Zvi & Garfield (Eds.),

- The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 3-15). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Bishop, A., & Goffree, F. (1986). Classroom organization and dynamics. In B. Christiansen, A. G. Howson & M. Otte (Eds), *Perspectives on mathematics education*. Dordrecht: D. Reidel, 309-365.
- Bogdan, R., & Biklen, S. K. (1997). *Investigação qualitativa em educação: Uma introdução à teoria e aos métodos*. Porto: Porto Editora.
- Branco, J., & Martins, M. E. (2002). Literacia estatística. *Educação e Matemática*, 62, 9-13.
- Campos, C. R. (2007). A Educação Estatística: uma investigação acerca dos aspetos relevantes à didática da estatística em cursos de graduação [online]. Tese de Doutoramento, Universidade Estadual Paulista (UNESP), Brasil. Disponível em: [http://www.pucrs.br/famat/viali/tic\\_literatura/teses/Campos\\_Celso\\_Ribeiro.pdf](http://www.pucrs.br/famat/viali/tic_literatura/teses/Campos_Celso_Ribeiro.pdf)
- Campelos, S. M. S. (2014). *Desenvolvimento da Literacia Estatística: uma abordagem no 3.º ciclo*. Tese de Doutoramento. Universidade Aberta, Portugal.
- Canavarro, A. P., & Ponte, J. P. (1997). *Matemática e novas tecnologias*. Lisboa: Universidade Aberta.
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamics events: A framework and a study. *Journal for Research in Mathematics Education*, 33, 352–378.
- Carmo, H., & Ferreira, M. M. (1998). *Metodologia da Investigação – Guia para Auto-aprendizagem*. Lisboa: UA.
- Carvalho, C. (2001). *Interações entre pares: contributos para a promoção do desenvolvimento lógico e do desempenho estatístico no 7.º ano de escolaridade*. Lisboa: APM
- Carvalho, C. (2006). Olhares sobre a educação estatística em Portugal. In *Anais do SIPEMAT*. Recife: Programa de Pós-Graduação em Educação, centro de Educação, Universidade Federal de Pernambuco.
- Carvalho, C. (2009). Reflexões em torno do ensino e da aprendizagem da Estatística. O exemplo dos gráficos. In J. A. Fernandes, F. Viseu, M. H. Martinho & P. F. Correia (Orgs). *Atas do II Encontro de Probabilidades e Estatística na Escola* (pp. 22-36). Braga: Centro de Investigação em Educação da Universidade do Minho.
- Carvalho, C., & César, M. (2000) As aparências iludem: reflexões em torno do ensino da estatística no ensino básico. Em C. Loureiro, O. Oliveira. e L. Brunheira. (Orgs.), *Ensino e aprendizagem da estatística*, (pp.212 - 225). Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamento de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.

- Chapman, L. J., & Chapman, J. P. (1982). Test results are what you think they are. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: heuristics and biases* (pp. 239-248). Cambridge: Cambridge University Press.
- Chaput, B., Girard, J. C., & Henry, M. (2011). Frequentist approach: Modelling and simulation in statistics and probability teaching. In C. Batanero, G. Burrill & C. Reading (Eds), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 85-95). New York, NY: Springer.
- Chick, H. (2004). *Representing association: Children manipulating data sets*. Presented at ICME-10, Denmark, TSG11: Research and development in the teaching and learning of probability and statistics. Disponible em <http://www.icme-10.dk>.
- Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*, 1-78.
- Cobo, F., Estepa, A., & Batanero, C. (2000). Un estudio experimental de la estimación de la correlación a partir de diferentes representaciones. *Enseñanza de las Ciencias, 18*(2), 297-310.
- Cohen, L., & Manion, L. (1990). *Métodos de Investigación Educativa*. Madrid: La Muralla.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin, 90*, 2, 272-292.
- Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382-393.
- Curcio, F. R. (1989). *Developing graph comprehension: elementary and middle school activities*. Reston, VA: National Council of Teachers of Mathematics.
- Dick, T. (1996). Much more than a toy. Graphing calculators in secondary school calculus. In Gómez, P. & Waits, B. (Eds.), *Roles of Calculators in the Classroom*. Bogotá (pp. 31-46).
- Engel, J., & Sedlmeier, P. (2011). Correlation and regression in the training of teacher. In C. Batanero, G. Burrill & C. Reading (Eds), *Teaching statistics in school mathematics-challenges for teaching and teacher education*. New York: Springer. pp. 247-258.
- Erickson, F. (1989). Qualitative methods in research on teaching. In M. C. Wittrock (Ed), *Handbook of research on teaching*, (pp. 119-161). New York, NY: Macmillan.
- Estepa, A. (2008). Interpretación de los diagramas de dispersión por estudiantes de bachillerato. *Enseñanza de las Ciencias, 13*(2), 155-170.
- Estepa, A., & Batanero, C. (1995). Judgments of association in scatterplots. In J. Garfield (Ed.). *Research Papers from the 4<sup>th</sup> International Conference on Teaching Statistics* (pp. 117-124). Minnesota, MN: The International Study Group for Research on Learning Probability and Statistics.

- Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatterplots: An empirical study of students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, 4, 25-41.
- Estepa, A., & Sanchez-Cobo, F. T. (2003). Evaluación de la comprensión de la correlación y regresión a partir de la resolución de problemas. *Statistics Education Research Journal*, 2(1), 54-68.
- Fernandes, J. A. (2009). Ensino e aprendizagem da Estatística: Realidades e desafios. *Actas do XIX EIEM: Números e Estatística*. Vila Real.
- Fitzallen, N. (2006). A model of students' statistical thinking and reasoning about graphs in an ICT environment. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Proceedings of 29<sup>th</sup> Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 203-210). Sydney: MERGA.
- Fitzallen, N. (2012). *Reasoning about covariation with TinkerPlots*. (Tese de doutoramento, University of Tasmania, Tasmania).
- Fitzallen, N., & Watson, J. (2010). Developing statistical reasoning facilitated by TinkPlots. In C. Reading (Ed), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010)*. Ljubljana, Slovenia. Voorburg, The Netherlands: ISL.
- Fitzallen, N., & Watson, J. (2011). Graph creation and interpretation: Putting skills and context together. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thorton (Eds.), *Proceedings of the Joint Conference of the Australian Association of Mathematics Teachers and Mathematics Education Research Group of Australasia* (pp. 253-260). Sydney: MERGA.
- Foster, P. (1996). *Obseerving Schools. A methodological guide*. London: Paul Chapman Publishing.
- Franklin, C., Kader, G., Mewborn, D. S., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *A curriculum framework for K-12 statistics education. GAISE report*. American Statistical Association. Online: [www.amstat.org/education/gaise/](http://www.amstat.org/education/gaise/).
- Freitas, C. (2011). *Literacia Estatística no 5.º ano: uma experiência de ensino*. Tese de mestrado, Universidade de Lisboa, Portugal.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Garfield, J. (2002). The Challenge of Developing Statistical Reasoning" *Journal of Statistics Education*, 10(3). (Disponível em: [www.amstat.org/publications/jse/v10n3/garfield.html](http://www.amstat.org/publications/jse/v10n3/garfield.html))
- Garfield, J., & Ben-Zvi, D. (2007). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*. Emeryville, CA: Key College Publishing.

- Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning. *Research and teaching practice*. Dordrecht, The Netherlands: Springer.
- Garfield, J., & Ben-Zvi, D. (2010). Developing students' statistical reasoning. *Connecting research and teaching practice*. Dordrecht, The Netherlands: Springer.
- Garfield, J., & Gal, I. (1999). Teaching and assessing Statistical reasoning. In L. Stiff & F. Curcio (Eds), *Developing mathematical reasoning in grades K-12*. 1999 Yearbook (pp. 207-219). Reston: NCTM.
- Garfield, J., & Franklin, C. (2011). Assessment of learning, and a learning in Statistics Education. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching Statistics in School Mathematics – Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study* (pp. 133-145). New York: Springer.
- Goetz, K. (2000). Perspectives on team. *Egallery*, 1 (4). Disponível em: <http://people.ucalgary.ca/~egallery/goetz.html>. Acedido em 10-05-2014.
- Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 517–545). Mahwah, NJ: Laurence Erlbaum Associates.
- Gonçalves, C. V., Fernandes, J. A., & Correia, P. F. (2013). Aprendizagem de estatística com tecnologia no 7.º ano de escolaridade. In a J. A., Fernandes, F., Viseu, M. H, Martinho, & P. F., Correia (Orgs.). *Atas do III Encontro de Probabilidades e Estatística na Escola*. Braga: Centro de Investigação em Educação na Universidade do Minho.
- Goos, M., & Bennison, A. (2008). Surveying the technology landscape: teacher' use of technology in secondary mathematics classrooms. *Mathematics Education Research Journal*, 20 (3), 102-130.
- Grant, F., & Searl, J. (1996). Mathematical modeling with a graphics calculator. In Gómez, P. & Waits, B. (Eds). *Roles of Calculators I nthe Classroom*. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.480&rep=rep1&type=pdf>
- Groth, R. E. (2006). An exploration of students' statistical thinking. *Teaching Statistics*, 28(1), pp. 17-21.
- Hall, J. (2011). Engaging teachers and students with real data: Benefits and challenges. In *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 335-346). Dordrecht, The Netherlands: Springer.
- Hammerman, J., & Rubin, A. (2004). Strategies for manging statistical complexity with new software tools. *Statistics Education Research Journal*, 3(2), pp. 17-41.
- Henriques, A., & Nascimento, M. M. (2013). Ensino e Aprendizagem de probabilidades e estatística. J. A., Fernandes, M. H., Martinho, J., Tinoco, & F., Viseu (Orgs). *Atas do XXIV Seminário de Investigação em Educação Matemática*.

Braga: APM & CIEd da Universidade do Minho. Disponível em: [http://www.apm.pt/files/\\_S2-OTI\\_529d279e26d76.pdf](http://www.apm.pt/files/_S2-OTI_529d279e26d76.pdf)

- Henriques, A. & Oliveira, H. (2012). Investigações estatísticas: um caminho a seguir? *Educação e Matemática*, 120, 3-8.
- Henriques, A & Antunes, P (2014). A exploração de covariação estatística por alunos do 10.º ano. *Quadrante*, 23(2), pp. 95-122.
- Jennings, D. L., Amabile, T. M., & Ross, L. (1982), Informal covariation assessment: Data-base versus theory-based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgments under uncertainty: Heuristic and biases* (pp. 211-230). New York, NY: Cambridge University Press
- Jolliffe, F. (2007). The changing brave new world of statistics assessment. In B., Phillips and L., Weldon (Eds.), *The Proceedings of the ISI/LASE Satellite on Assessing Student Learning in Statistics*, Voorburg: International Statistical Institute, The Netherlands, CD-Rom.
- Koichu, B., & Harel, G. (2007). Triadic interaction in clinical task-based interview with mathematics teachers, Department of Mathematics, Springer Science, San Diego: University of California, pp.349 -365
- Konold, C. (2002). Alternatives to scatterplots. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa. [CDROM] Voorburg, The Netherlands: International Statistical Institute.
- Konold, C. (2006). Designing a Data Analysis Tool for Learners. In M. Lovett & P. Shah (Eds.), *Thinking with data: The 33<sup>rd</sup> Annual Carnegie Symposium on Cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Konold, C., Harradine, A., & Kazak, S. (2007). Understanding distributions by modeling them. *International Journal of Computers for Mathematical Learning*, 12, pp. 217-230.
- Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin & D. Shifter (Eds.), *A research companion to Principles and standards for school mathematics* (pp. 193-215). Reston, VA: NCTM.
- Konold, C., & Miller, C. (2001). *Tinkerplots, Data analysis software for the middle school*. Amherst, MA: University of Massachusetts.
- Konold, C., & Miller, C., D. (2005). *TinkerPlots: Dynamic data exploration*. Emeryville, CA: Key Curriculum Press.
- Lakatos, E. M., & Marconi, M. de A. (2008). *Técnicas de pesquisa: Planejamento e execução de pesquisas, Amonstragens e técnicas de pesquisa. Elaboração, análise e interpretação de dados*. São Paulo: Editora Atlas.
- Lee, H. S., Kersaint, G., Harper, S., Driskell, S. O., & Leatham, K (2012). Teachers' statistical problem solving with dynamic technology: Research results across multiple institutions. *Contemporary Issues in Technology and*

Teacher Education, 12 (3). Disponível em:  
<http://www.citejournal.org/vol12/iss3/mathematics/article1.cfm>

- Lee, H. S., Kersaint, G, Harper, S., Driskell, S., Jones, D., & Leatham, K. (2014). Teachers' use of transnumeration in solving statistical tasks with dynamic statistical Software. In *Statistics Education REsearch Journal*, 13 (1), pp. 25-52. Disponível em: [http://iase-web.org/documents/SERJ/SERJ13\(1\)\\_Lee.pdf](http://iase-web.org/documents/SERJ/SERJ13(1)_Lee.pdf)
- Lee, H. S., & Hollebrands, K. F. (2008). Preparing to teach data analysis and probability with technology. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI study 18 and 2008 IASE Round Table Conference*.
- Lesh, R., & Kelly, A. A. E. (2000). Multitiered teaching experiments. In A. E. Kelly & R. Lesh (Eds). *Handbook of research design in mathematics and science education* (pp. 197-230). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lopes, C. (1998). *A Probabilidade e a Estatística no ensino fundamental: Uma análise curricular*. Tese de mestrado apresentado à Faculdade de Educação da Universidade Estadual de Campinas.
- Lopes, C. E. (2012). A educação Estocástica na Infância. *Revista electrónica de Educação*, 6, 160-174.
- Ludke, M. & André, M. (1996). *Pesquisa em Educação: Abordagens Qualitativas*. São Paulo: Editora Pedagógica e Educativa.
- Martins, A. A., Fonseca, C. M. C., Lopes, I. M. C., Martins, M. P. G., Fonseca, M. C., & Silva, J. M. M. C (Coord.) (2005). *Programa – Componente de Formação Científica. Disciplina de Matemática. Cursos Profissionais de Nível Secundário*. Lisboa: D. G. F. V., Ministério da Educação.
- Martins, M. E. G., & Ponte, J. P. (2010). *Organização e tratamento de dados*. Lisboa: DGIDC.
- Merriam, S. B. (1988) *Case study research in education: a qualitative approach*. San Francisco: Jossey-Bass.
- Ministério da Educação (2001a). *Matemática A – 10º ano*. Lisboa: ME, DES.
- Ministério da Educação (2001b). *Matemática B – 10º ano*. Lisboa: ME, DES.
- Ministério da Educação (2001). *Programa de Matemática aplicada às Ciências Sociais, 10.º, 11.º e 12.º anos*. Lisboa: ME,
- Ministério da Educação (2007). *Programa de Matemática do Ensino Básico*. Lisboa: Direcção Geral de Inovação e Desenvolvimento Curricular, Ministério da Educação.

- Ministério da Educação (2013). *Programa e Metas Curriculares em Matemática para o Ensino Básico*. Lisboa: ME.
- Monteiro, C. (2005). *Investigating critical sense in the interpretation of media graphs*. (Tese de doutoramento, University of Warwick).
- Morais, P. (2010). *Construção, leitura e interpretação de gráficos estatísticos por alunos do 9.º ano de escolaridade*. (Tese de Mestrado, Universidade do Minho).
- Moritz, J. B. (2000). Graphical representations of statistical associations by upper primary students. In J. Bana & A. Chapman (Eds.), *Proceedings of the 23<sup>rd</sup> Annual Conference of the Mathematics Education Research Group of Australasia* (v. 2, pp. 440–447). Perth: MERGA.
- Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–255). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Mugabe, D., Fernandes, J. A., & Correia, P. (2012). Avaliação da associação estatística num diagrama de dispersão por estudantes universitários. In H. Pinto, H. Jacinto, A. Henriques, A. Silvestre, C. Nunes, *Atas do XXIII Seminário de Investigação em Educação Matemática* (pp.403- 414). Lisboa: APM.
- NCTM (1991). *Normas para o currículo e a avaliação em matemática escolar* (tradução de E. Veloso, F. Nunes, H. Guimarães, J.F. Matos, J. M. Duarte, L. Leal, L. Moreira, L. Serrazina & R. Carvalho]. Lisboa: APM/IEE.
- NCTM (2008). *Princípios e normas profissionais para a Matemática Escolar*. Lisboa: APM e IIE.
- Pfannkuch, M., Rubick, A., & Yoon, C. (2002). Statistical thinking and transnumeration. In B. Barton, K. Irwin, M. Pfannkuch & M. Thomas (Eds.), *Proceedings of the 25<sup>th</sup> annual conference of the Mathematics Education Research Group of Australasia* (pp. 567-574). Sydney: MERGA.
- Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 17-46). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Pocinho, R. & Gaspar, J. /2012). O uso das TIC e as alterações no espaço Educativo, *Exedra* [Online],n.º. disponível em: <http://www.exedrajournal.com/docs/N6/09-Edu.pdf>
- Ponte, J. P. (2002). Literacia Matemática. In M. N. Trindade (Org.), *Actas do Encontro Internacional Literacia e cidadania: Convergências e interfaces* (em CD-Rom). Universidade de Évora: Centro de Investigação em Educação Paulo Freire.
- Ponte, J. P. (2006). Estudos de caso em educação matemática. *Bolema*, 25, 105-132
- Ponte, J. P., Brocado, J., & Oliveira, H. (2006). *Investigações Matemáticas na Sala de Aula*. Belo Horizonte: Autêntica.

- Ponte, J. P., & Canavarro, A. P. (1997). *Matemática e Novas Tecnologias*. Lisboa: Universidade Aberta.
- Ponte, J. P., Costa, F., Lopes, H., Moreirinha, O., & Salvado, D. (1997). Histórias da aula de Matemática. Lisboa: APM.
- Ponte, J. P., & Fonseca, H. (2001). Orientações curriculares para o ensino da estatística: Análise comparativa de três países. *Quadrante*, 10(1), 93-115. (disponível em [repositorio.ul.pt/bitstream/10451/3035/1/01-Ponte-Fonseca%20Quadrante.pdf](http://repositorio.ul.pt/bitstream/10451/3035/1/01-Ponte-Fonseca%20Quadrante.pdf))
- Ponte, J. P., Matos, J. M., & Abrantes, P. (1998). *Investigações em Educação Matemática*. Lisboa: IIE
- Ponte, J. P., & Sousa, H. (2010). Uma oportunidade de mudança na matemática no ensino básico. In GTI (Org), *O professor e o programa de matemática do ensino básico* (pp. 11-41) Lisboa: APM.
- Quintas, S., Oliveira, H., & Ferreira, R. (2009). Estudo exploratório sobre perspetivas e práticas de professores de matemática, do ensino secundário no domínio do ensino da estatística. In J. A. Fernandes, M. H. Marinho, F. Viseu & P. F. Correia (Orgs.), *Atas do II Encontro de Probabilidades e Estatística na Escola* (pp. 1-13). Braga: Centro de Investigação em Educação da Universidade do Minho.
- Quintas, S.; Oliveira, H., & Ferreira, R. T. (2013). O conhecimento didático de uma professor no ensino da relação bivariada na Estatística. In J. A. Fernandes, M. H. Martinho, J. Tinoco, & F. Viseu (Orgs). *Atas do XXIV Seminário de Investigação em Educação Matemática*. Braga: APM & CIEd da Universidade do Minho. Disponível em: [http://apm.pt/files/S2-C7-Quintas\\_529d2823b1e09.pdf](http://apm.pt/files/S2-C7-Quintas_529d2823b1e09.pdf).
- Ridgway, J., Nicholson, J., & McCusker, S. (2011). Developing statistical literacy in students and teachers. In C. Batanero, G. Burrill & C. Reading (Eds). *Teaching Statistics in school mathematics: Challenges for teaching and teacher education* (pp. 311-322). New York, NY: Springer.
- Ridgway, J., Nicholson, J., & McCusker, S. (2011). Reconceptualising ‘statistics’ and ‘education’. In *Statistics Education in School Mathematics: Challenges for Teaching and Teacher Education*, Ed. C. Batanero. Heidelberg: Springer.
- Rubin, A. (2007). Much has changed; little has changed; revisiting the role of technology in statistics education. *Technology Innovations in Statistics Education*, 1(1). Online: <http://escholarships.org/uc/item/833239sw>.
- Rubin, A., & Hammerman, J.K. (2006). Understanding data through new software representations. In G. F. Burrill (Ed), *Thinking and reasoning with data and chance: Sixty-eight NCTM yearbook*. Reston, Va: National Council of Teachers of Mathematics. pp. 241-256.
- Rubin, A., Hammerman, J. K. & Konold, C. (2006). Explore informal inference with interactive visualization software. In A. Rossman & B. Chance (Eds),

- Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador, Bahia, Brazil: International Association for Statistical Education. Disponível em : <http://iase-web.prg>.
- Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, 10(3). Disponível em: <http://www.amstat.org/publications/jse/v10n3/rumsey2.html>
- Russell, S., & Friel, S. (1989). Collecting and analyzing real data in the elementary school Classroom. In P.R. Trafton & A.P. Shulte (Eds), *New directions for elementary school mathematics* (pp. 134-148). Reston, VA: National Council of Teachers of Mathematics.
- Ruthven, K., Hoffman, R. & Mercer, N. (2011). A dialogic approach to plenary problem synthesis. In B. Ubuz (Ed.), *Proceedings of the 35<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, 4 (pp. 81-88). Ankara, Turkey: PME.
- Salomé, H.; Silva, L.P (2014). *Módulo A3 – Estatística*. Matemática A3 – Cursos profissionais de nível secundário. Lisboa Editora: Lisboa.
- Savenye, W., & Robinson, R. (2001). Qualitative research issues and methods: an introduction for educational technologists. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp.1171-1195). New Jersey: Lawrence Erlbaum Associates.
- Schoenfeld, A. (1999). Looking toward the 21<sup>st</sup> century: challenges of educational theory and practice. *Educational Researcher*, 28 (7), 4-14.
- Seidman, I. (2006). *Interviewing as qualitative research – a guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.
- Selmer, S. J. Bolyard, J. J., & Rye, J. A. (2012). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, 95(6), pp. 323-332.
- Shaughnessy, J. M. (2007). Research on Statistics Learning and Reasoning. In F. Lester (Eds.) *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957-1009). Greenwich, CT: Information Age Publishing and NCTM.
- Silva, A. A. (2006). *Gráficos e mapas: Representação de informação estatística*. Lisboa: LIDEL.
- Stake, R. (1994). *The art of case study research*. London: Sage.
- Stake, R. E. (2007). *A Arte da Investigação com Estudos de Caso*. Lisboa: Fundação Calouste Gulbenkian.
- Steffe, L., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. *Research design in mathematics and science education*. Hillsdale, NJ.

- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Helping teachers learn to better incorporate student thinking. *Mathematical Thinking and Learning*, 10(4), 313-340.
- Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School*, 3, 268-275.
- Thompson (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grows (Ed), *Handbook of research on mathematics teaching and learning* (pp. 127-146). Nova York, NY: MacMilan.
- Tuckman, B. (1999). *Conducting educational research*. Orlando: Harcourt Brace College Publishers.
- Tuckman, B. (2000). *Manual de Investigação em Educação*.(4), Fundação Calouste Lisboa: Gulbenkian.
- Tudella, A. Ferreira, C., Bernardo, C., Pires, F., Fonseca, H., Segurado, I., & Varandas, J. (1999). Dinâmica de uma aula com investigações. In P. Abrantes, J. P. Ponte, H. Fonseca & L. Brunheira (Orgs). *Investigações matemáticas na aula e no currículo*. Lisboa: Projeto MPT e APM.,pp. 87-96.
- Vale, I. (2000). *Didáctica da Matemática e Formação inicial de professores num contexto de Resolução de Problemas e de Materiais Manipuláveis*. Universidade de Aveiro.
- Vieira, M. (2008). *Análise exploratória de dados: uma abordagem com alunos do ensino médio*. (Tese de Mestrado, Universidade Católica de São Paulo).
- Watson, J. (1997). Assessing statistical thinking using the media. In I. Gal & J. Garfield (Eds), *The Assessment Challenge in Statistics Education* (pp.107- 121). Amsterdam: IOS Press and International Statistical Institute.
- Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Mathwah, NJ: Lawrence Erlbaum Associates.
- Watson, J. (2008). Exploring beginning inference with novice grade 7 students. *Statistics Education Research Journal*, 7(2), pp. 59-82.
- Watson, J. M., & Donne, J. (2009). TinkerPlots as a research tool to explore student understanding. *Technology Innovations in Statistics Education*, 3(1). Online: <http://repositories.cdlib.org/uclastat/cts/tise/vol3/iss1/art1/>.
- White, A. (2009). Graphing calculator in the secondary mathematics classroom, pedagogical tool or just a gadget?. In *Proceeding of the Conference Redesigning pedagogy: research, policy, practice*. Singapore: National Institute of Education.
- Zbiek, R., & Hollebrands, K. (2007). Chapter 7: A research-informed view of the process of incorporating mathematics technology into classroom practice by inservice and prospective teachers. In K. Heid & G. Blume (Eds), *Research on*

*technology and the teaching of mathematics* , 1 (pp. 287-344).Greenwich, CT: Information Age Publishing.

Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, 8(1), pp. 7-31.

Yin, R. K. (2003). *Case study research: Design and methods* (3.<sup>o</sup> ed.). Thousand Oaks/London: Sage Publications.



## **Anexos**

## **Anexo 1. Pedido de Autorização aos alunos participantes no Estudo-Piloto**

**Exmo. (ª) Sr. (ª) Encarregado(a) de Educação,**

Eu, Patrícia Alexandra Robalo Antunes, professora de Matemática, venho por este meio solicitar a V. EX<sup>ª</sup>. autorização para registar em áudio/vídeo de algumas aulas de Matemática, bem como, a realização de entrevista ao seu educando, no último período, no ano 2012/2013. O nome do seu educando será alterado, de modo a preservar a sua identidade.

O principal objetivo deste trabalho é procurar analisar como o seu educando explora relações entre variáveis e estabelecem a existência da covariação, através do uso do *software* educativo - *TinkerPlots*. É ainda de referir que o trabalho de investigação fará parte de um estudo que desenvolverei no Instituto de Educação, da Universidade de Lisboa, no âmbito de um trabalho científico da disciplina de Didática de Probabilidade e Estatística do Mestrado na área da Didática da Matemática.

Agradeço a colaboração e atenção dispensada

Caso necessite de mais esclarecimentos, não hesite em contactar-me.

Atenciosamente, a professora de Matemática:

\_\_\_\_\_  
(Patrícia Alexandra Robalo Antunes)

-----  
(destacar e entregar à professora)

Eu, \_\_\_\_\_ Encarregado(a) de Educação do(a) aluno(a) \_\_\_\_\_, número \_\_\_\_\_ da turma A do 10º ano.

Tomei conhecimento do assunto referido no documento entregue ao meu educando pela Professora de Matemática e (coloque a x no respetivo espaço):

\_\_\_\_\_ Autorizo o registo nos moldes acima mencionados

\_\_\_\_\_ Não autorizo o registo nos moldes acima mencionados

Assinatura do Encarregado(a) de Educação: \_\_\_\_\_

## **Anexo 2. Pedido de autorização à Direção da Escola para Estudo-Piloto**

Exmo. Sr. Diretor do [REDACTED]

[REDACTED] 2 de abril de 2013

### **Assunto: Pedido de autorização para uma investigação**

No âmbito da realização de um trabalho científico da disciplina de Didática de Probabilidades e Estatística do Mestrado na área da Didática da Matemática do Instituto de Educação da Universidade de Lisboa, pretendo desenvolver uma pequena investigação. Procuo analisar como é que os alunos do 10.º ano exploram relações entre variáveis e estabelecem a existência de covariação, quando utilizam o *software TinkerPlots*, no tópico de Distribuições Bidimensionais, na disciplina de Matemática.

Para esse efeito necessito de realizar uma pequena recolha de dados com a turma A do 10.º ano de escolaridade, pelo que venho junto de V. Ex<sup>a</sup> solicitar a devida autorização.

A recolha de dados, acima referida, irá decorrer no final de maio de 2013 nas aulas de Matemática, e será levada a cabo apenas por mim e implicará a realização de entrevistas a alguns alunos, a gravação áudio e vídeo do trabalho realizado pelos alunos, bem como fotografias dos seus cadernos. Os nomes dos alunos serão alterados, de modo a preservar a sua identidade.

Obrigada pela atenção dispensada,

Atenciosamente, pede deferimento:

---

**Patrícia Alexandra Robalo Antunes**

### **Anexo 3. Guião da entrevista/tarefa**

A tarefa que a seguir apresentada insere-se na unidade da Estatística, onde pretende abranger alguns conteúdos abordados no 10ºano. A tarefa apresentada, refere-se a dados de resultados nos testes de condição física inicial de uma turma de 9ºano de escolaridade. Nesta tarefa, poderás analisar as várias variáveis em estudo e criar as tuas próprias conclusões, tendo como ferramenta auxiliar o *TinkerPlots*.

- **Que informações podes obter a partir deste conjunto de dados?**
- **Que coisas interessantes podes descobrir sobre este conjunto de dados?**
- **Formula hipóteses sobre estes dados.**
- **Que evidência podes produzir para suportar as tuas conjeturas?**  
*Explica-me o que estás a fazer enquanto constróis os gráficos.*
- **Porque é que escolheste este gráfico? O que podes observar nele? Existe outra forma de organizar a informação que também possa ser útil?**
- **Pensas que, neste conjunto de dados, seria interessante procurar algumas diferenças entre alunos do sexo masculino e feminino? Porquê?**
- **Quem teve um melhor desempenho? Rapazes ou raparigas?**  
(Em relação às variáveis que o aluno escolheu)
  - **O resultado vai ao encontro do que estavas à espera?**
  - **O que pensas estar na origem dessas diferenças/resultados?**
- **Os alunos mais altos tendem a ser os mais pesados e os mais rápidos?**
  - **Pensas existir uma explicação plausível para esses resultados?**
- **Sabes o que é o Índice da Massa Corporal (IMC)? Tendo por base esse conhecimento, esperas que os alunos que apresentam índices mais altos sejam também os mais pesados?**  
**Explica porquê. Verifica.**
- Volta a olhar para os gráficos que criaste.  
**Qual dos gráficos mostra a relação mais forte entre os dois atributos? O que é que observaste nos gráficos que te ajudou a tomar essa decisão?**
- **Se recolhermos os dados dos alunos do 11.º ano, esperas obter resultados semelhantes ou diferentes? Semelhantes/Diferentes em que medida? O que justifica/fundamenta essa tua crença?**

## Anexo 4. A tarefa da entrevista (suporte em ficheiro *TinkerPlots*)

The screenshot shows the TinkerPlots software interface. At the top, there is a menu bar with 'File', 'Edit', 'Object', 'Data', 'Window', and 'Help'. Below the menu bar is a toolbar with icons for 'Cards', 'Table', 'Plot', 'Sampler', and 'Text'. The main window is titled 'Collection 1' and shows a table of data for 'case 1 of 21'. The table has two columns: 'Attribute' and 'Value'. The attributes and their values are: Nome (Alcino), Peso (59.9), Altura (164.5), Força\_Superior (27), Resistência (1800), Força\_Abdominal (41), Força\_Inferior (200), Flexibilidade (38), Velocidade (6.5), Impulsão\_Vertical (35), IMC (22.1), and Sexo (M). Below the table, there is a section titled 'Dados de resultados dos testes de condição física inicial de uma turma de 9º ano de escolaridade.' followed by a list of abbreviations and their definitions: NOME, PESO, ALTURA, RESISTÊNCIA, FORÇA SUPERIOR, FORÇA ABDOMINAL, FORÇA INFERIOR, FLEXIBILIDADE, VELOCIDADE, IMPLUSÃO VERTICAL, and IMC. To the right of the data table, there is a section titled 'QUESTÕES:' followed by eight numbered questions in Portuguese. The questions are: 1. Analisa e explica o conjunto de dados extraídos no estudo. Que informações podes obter a partir deste conjunto de dados? 2. Que coisas interessantes podes descobrir sobre este conjunto de dados? 3. Formula hipóteses sobre estes dados. 4. Que evidência podes produzir para suportar as tuas conjecturas? 5. Pensas que, neste conjunto de dados, seria interessante procurar algumas diferenças entre alunos do sexo masculino e feminino? Porquê? 6. Quem teve um melhor desempenho? Rapazes ou raparigas? 7. Os alunos mais altos tendem a ser os mais pesados e os mais rápidos? 8. Sabes o que é o IMC? Tendo por base esse conhecimento, esperas que os alunos que apresentam índices mais altos sejam também os mais pesados? Explica porque.

**QUESTÕES:**

1. Analisa e explica o conjunto de dados extraídos no estudo. Que informações podes obter a partir deste conjunto de dados?
2. Que coisas interessantes podes descobrir sobre este conjunto de dados?
3. Formula hipóteses sobre estes dados.
4. Que evidência podes produzir para suportar as tuas conjecturas?
5. Pensas que, neste conjunto de dados, seria interessante procurar algumas diferenças entre alunos do sexo masculino e feminino? Porquê?
6. Quem teve um melhor desempenho? Rapazes ou raparigas?
7. Os alunos mais altos tendem a ser os mais pesados e os mais rápidos?
8. Sabes o que é o IMC? Tendo por base esse conhecimento, esperas que os alunos que apresentam índices mais altos sejam também os mais pesados? Explica porque.

Dados de resultados dos testes de condição física inicial de uma turma de 9º ano de escolaridade.

NOME: identificação do aluno  
 PESO: de cada aluno  
 ALTURA: de cada aluno  
 RESISTÊNCIA: metros que cada aluno correu, durante 10 minutos  
 FORÇA SUPERIOR: em 30 segundos quantas flexões faz  
 FORÇA ABDOMINAL: 30 segundos quantas abdominais faz  
 FORÇA INFERIOR: salto frontal a pés juntos.  
 FLEXIBILIDADE: é a capacidade de realizar movimentos com a máxima amplitude articular  
 VELOCIDADE: duração da corrida em 60 metros  
 IMPLUSÃO VERTICAL: salto na vertical com pés juntos.  
 IMC: identifica o grau de obesidade  
 SEXO: ise o aluno é feminino ou masculino

TinkerPlots® version 2.1

## Anexo 5. Pedido de autorização ao Diretor da Escola

Exmo. Diretor do [REDACTED]

### **Assunto: Pedido de autorização para uma investigação**

Eu, Patrícia Alexandra Robalo Antunes, docente deste estabelecimento, venho por este meio solicitar a sua autorização para realizar, no ano letivo 2013/2014, com a turma 10.º ano dos cursos profissionais de Técnico Auxiliar de Saúde (TAS) e de Técnico de Apoio à Gestão Desportiva /TAGD), em colaboração com a docente da disciplina de Matemática, [REDACTED], um trabalho de investigação no âmbito do trabalho de projeto do Mestrado em Didática da Matemática que me encontro a concluir no Instituto de Educação da Universidade de Lisboa.

O principal objetivo deste trabalho é compreender as aprendizagens e dificuldades dos alunos, no que respeita à covariação estatística, quando utilizam o software *TinkerPlots* na resolução de tarefas que envolvam esse conceito. A organização de dados a utilizar incluirá a gravação de aulas relativas à unidade de ensino – Distribuições Bidimensionais e a recolha de trabalhos produzidos por eles.

O desenvolvimento da investigação não interfere com o normal funcionamento das atividades letivas e não traz qualquer prejuízo para os participantes, estando garantida a confidencialidade dos dados recolhidos e o anonimato da escola e dos alunos em posteriores divulgações da investigação realizadas no âmbito do Mestrado.

Irei, ainda, proceder ao pedido de autorização dos Encarregados de Educação dos alunos para a referida recolha de dados.

Fico à inteira disposição de V. Ex.º para complementar toda a informação que julgue oportuna.

Agradecendo desde já a sua colaboração, subscrevo-me com os melhores cumprimentos,

Atenciosamente, pede deferimento:

---

(Patrícia Alexandra Robalo Antunes)

[REDACTED], 12 de março de 2014

## **Anexo 6. Pedido de Autorização aos Encarregados de Educação**

**Exmo. (ª) Sr. (ª) Encarregado(a) de Educação,**

Vai ser desenvolvido com os alunos da turma do(a) seu(sua) educando(a), em colaboração com a professora de Matemática, um trabalho de investigação, no ano letivo 2013/2014, no âmbito do trabalho de projeto do Mestrado em Didática da Matemática que me encontro a concluir no Instituto de Educação da Universidade de Lisboa. Este tem como objetivo compreender como é que os alunos trabalhando em grupo e realizando tarefas de exploração e investigação, com o auxílio do programa *TinkerPlots*, desenvolvem a capacidade de comunicação matemática, a exploração entre variáveis e superam dificuldades na aprendizagem da Estatística. Assim, solicito que autorize o(a) seu(sua) educando(a) a participar na recolha de dados. Utilizar-se-ão fotocópias de trabalhos realizados pelos alunos, entrevistas e gravações áudio e vídeo de aulas. Em todas as aulas, a responsável pela turma continuará a ser a professora de Matemática do Módulo.

É garantido, por mim, que os dados recolhidos apenas serão utilizados para os objetivos da investigação. Comprometo-me desde já a garantir o anonimato do seu educando e a confidencialidade dos dados obtidos que serão utilizados apenas no âmbito da referida investigação associada ao projeto e à dissertação do Mestrado.

Fico à inteira disposição de V. Ex.ª para complementar toda a informação que julgue oportuna.

Agradecendo, desde já, a colaboração prestada de V. Ex.ª, solicito que assine a declaração em baixo, devendo depois destaca-la e devolvê-la.

Com os melhores cumprimentos,

\_\_\_\_\_  
(Patrícia Alexandra Robalo Antunes)

-----  
(destacar e entregar à professora)

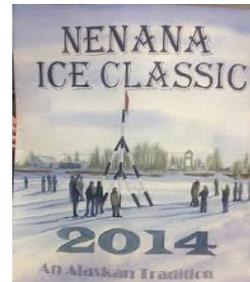
Declaro que autorizo o(a) meu(minha) educando(a) \_\_\_\_\_  
\_\_\_\_\_, n.º \_\_, turma \_\_\_\_ do 10.º ano, que frequenta \_\_\_\_\_  
seja fotografado e/ou filmado e/ou entrevistado com a garantia do respetivo anonimato, a fim de participar no trabalho de investigação conduzido pela professora Patrícia Antunes, docente neste Instituto, no âmbito da sua dissertação de Mestrado e da necessidade da respetiva recolha de dados.

\_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de 2014

O(A) Encarregado(a) de Educação

## Anexo 7. Tarefa: “Nenana Ice Classic”<sup>1</sup>

O *Nenana Ice Classic* é um concurso que ocorre anualmente junto à localidade de Nenana, no Alasca, para adivinhar quando é que o gelo do rio Tanana se quebra. As pessoas apostam sobre o minuto exato em que um tripé de madeira alto irá cair dentro do rio gelado. O concurso foi iniciado pelos inspetores dos caminhos-de-ferro do Alasca em 1917. Tem decorrido todos os anos desde essa data.



Em 2001, dois cientistas de Stanford publicaram um artigo na revista *Science* em que concluíram que, nessa altura, o degelo ocorreu 5.5 dias mais cedo do que quando o concurso começou em 1917. **Será que o degelo tem vindo a ocorrer mais cedo ao longo dos anos?**

**Qual pensas ser a resposta a esta questão?** Explica em que te baseaste para responder.

---

Propomos-te a exploração desta questão a partir de uma base de dados do programa *Tinkerplots* onde podes aceder a informação sobre esta temática. Estão disponíveis dados relativos a alguns atributos (variáveis) que podes visualizar na forma de cartões ou tabelas e cuja descrição se encontra numa caixa de texto no ecrã do programa.

### Parte I

**1. Verifica a tua conjectura** recorrendo a uma representação gráfica dos dados relativos ao número de dias que ocorrem desde 1 janeiro até ao degelo, ao longo dos anos.

Guarda o teu trabalho (File – Save as: nome do grupo\_I\_1).

### Parte II

No *TinkerPlots* encontras também informação sobre o momento do dia em que ocorreu o degelo (atributo "segment") ao longo dos anos. Responde e justifica as tuas respostas às perguntas que se seguem com base em representações gráficas.

1. a) **Em quantos anos** o degelo ocorreu durante a manhã?
- b) **Qual a hora do dia** em que o degelo ocorre mais frequentemente? **Explica** como pensaste.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_1).

**2. Considera a seguinte afirmação: “O degelo no mês de maio tende a ocorrer mais cedo ao longo do dia do que no mês de abril”.**

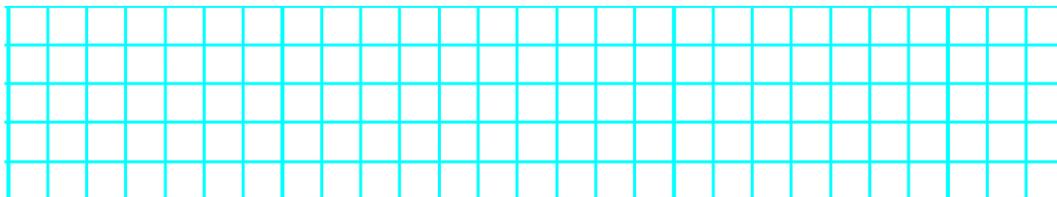
- a) **Diz o que pensas** sobre essa afirmação e apresenta evidência para apoiar a tua resposta.

---

<sup>1</sup> Tarefa elaborada no Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* financiado pela FCT – Fundação para a Ciência e Tecnologia (contrato PTDC/CPE – CED/117933/2010)

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_2a).

**b) Recorre** agora a um diagrama de extremos e quartis (com base na ferramenta Box Plot do *TinkerPlots*) para confirmares a tua resposta. **Desenha** o diagrama no quadriculado.



**c)** Com base na informação que consegues retirar do diagrama de extremos e quartis, **explica** porque concordas ou discordas com a afirmação acima.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_4).

**3. Qual é a tua previsão** sobre a data e o momento do dia em que ocorrerá o degelo do rio Tanana daqui a 10 anos?

**Apresenta argumentos** para a tua previsão.

## Anexo 8. Tarefa: “Uma experiência com peixes”<sup>2</sup>

Aquacultura ou aqüicultura é a produção de organismos aquáticos, como a criação de peixes, moluscos, crustáceos, anfíbios e o cultivo de plantas aquáticas para uso do homem.

Esta atividade é praticada há muito tempo, existindo registos de que os chineses já a realizavam vários séculos antes de nossa era e de que há 4000 anos os egípcios criavam a Tilápia-do-nylo (*Sarotherodon niloticus*). Atualmente, a aquacultura é responsável pela produção de metade do peixe consumido pela população mundial. (<http://pt.wikipedia.org/wiki/Aquacultura>)



Um aqüicultor tem abastecido os seus tanques rede com um novo tipo de peixe, geneticamente modificado, fornecido por uma empresa que lhe assegurou que **“os peixes geneticamente modificados, ao crescerem, atingem o dobro do comprimento dos peixes normais”**.

Pensas que o aqüicultor pode confiar na afirmação da empresa? O que deverá ele fazer para verificar a sua veracidade?

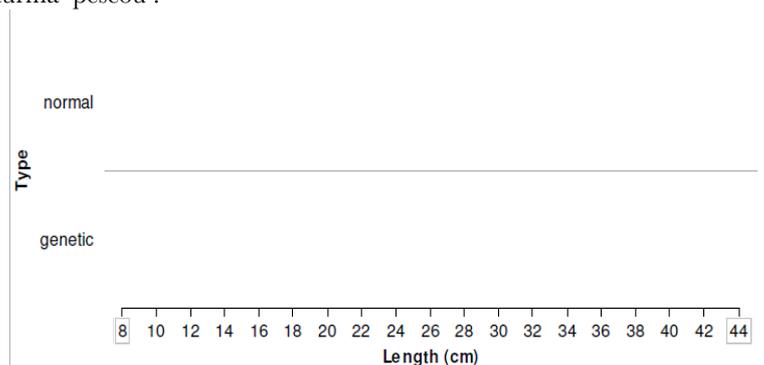
---

Tens agora oportunidade de **simular** uma ‘pescaria’ para responderes à questão: *Os peixes geneticamente modificados são mais compridos do que os normais? Se sim, quanto?*

### Parte I

O aqüicultor decidiu criar num tanque cerca de 625 peixes, alguns de tipo normal e outros geneticamente modificados. Quando os peixes estavam completamente desenvolvidos (fase adulta), o aqüicultor identificou e mediu o comprimento de cada um deles. Na tua experiência, estes peixes são simulados por pequenos cartões (cartões-peixe), com a indicação do seu tipo (normal ou geneticamente modificado) e do seu comprimento (em centímetros).

**1. Regista abaixo** os dados (tipo e comprimento do peixe) relativos a um conjunto de 25 cartões-peixe que a turma ‘pescou’.



**2. Com base na representação gráfica anterior, tenta responder à questão em análise.**

**3. A tua resposta seria a mesma caso tivesses selecionado uma amostra de dimensão 50? Explica porquê.**

---

<sup>2</sup> Tarefa elaborada no Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* financiado pela FCT – Fundação para a Ciência e Tecnologia (contrato PTDC/CPE – CED/117933/2010)

## Parte II

Propomos-te agora a exploração desta questão a partir de uma base de dados do programa *TinkerPlots* - “Fish Experiment”, onde podes simular amostras de dimensões diversas (**SampleSize** = 25, 30, 50, 75, ...).

1. **Simula** amostras de dimensão 25, 50 e 100. Para cada uma delas, com base na representação gráfica obtida, constrói um diagrama de extremos e quartis e regista, na tabela seguinte, algumas das medidas estatísticas que calculaste.

**Para cada amostra**, guarda o teu trabalho.  
(File – Save as: nome do grupo\_II\_#dimensão).

	<b>Amostras de dimensão crescente</b>			
		Amostra 1 (25)	Amostra 2 (50)	Amostra 3 (100)
<b>Geneticamente modificados</b>	Média			
	Mediana			
	Amplitude IQ			
<b>Normais</b>	Média			
	Mediana			
	Amplitude IQ			

2. Com base na informação obtida, **compara as distribuições obtidas** para as várias amostras e responde à questão inicial.

**Apresenta argumentos para justificar** a tua resposta que possam ajudar o aquicultor a decidir se deve manter este negócio com a empresa fornecedora.

## Anexo 9. Tarefa: “Desenvolvimento da criança”

Uma criança é um ser humano no início de seu desenvolvimento. A infância é o período que vai desde o nascimento até aproximadamente o décimo segundo ano de vida de uma pessoa. É um período de grande desenvolvimento físico, marcado pelo gradual aumento da altura e do peso da criança, especialmente nos primeiros três anos de vida e durante a puberdade.

O crescimento é considerado um processo dinâmico e contínuo, sujeito a variações em função do ambiente e, por isso, um indicador mais importante da qualidade de vida de uma criança. Isto justifica a necessidade de o monitorizar ao longo e toda a infância e adolescência de um indivíduo, obtendo-se medidas para diagnosticar o seu crescimento. O peso e a estatura são índices fundamentais na avaliação do crescimento. Destes, o mais utilizado é o peso, por ser de fácil obtenção, no entanto, a estatura é o indicador mais seguro, pois, ao contrário do que ocorre com a estatura, o peso sofre influência de muitos fatores, podendo até diminuir. (<http://pt.wikipedia.org/wiki/Criança>)

**Tuddenham, RD & Snyder, M.M, acompanharam o crescimento físico de rapazes e raparigas da Califórnia, desde o seu nascimento até atingirem os 18 anos. Com base nos dados desse estudo, como caracterizarias estas crianças, no que diz respeito ao seu crescimento físico?**

---

### Parte I

1. Propomos-te agora a exploração dos dados que já se encontram numa base de dados do *Tinkerplots*.
  - a) **Que questões interessantes** poderias colocar sobre estes dados quanto a:
    - altura média das crianças;
    - altura dos rapazes e das raparigas;
    - relação entre a idade, o peso e a estatura das crianças;
    - outro aspeto que consideres relevante estudar.
  - b) **Qual pensas** ser a resposta a essas questões?  
**Explica** em que te baseaste para responder.
  - c) **Responde às questões** que formulaste na questão 1.a) recorrendo a representações gráficas para verificar as tuas conjeturas.

Guarda o teu trabalho (File – Save as: nome do grupo\_I).

### Parte II

Utiliza também, esta base de dados no *TinkerPlots* para responder às questões que se seguem:

1. **Quanto** é que os rapazes diferem das raparigas, em termos de altura, aos 2 anos? E aos 9 anos? E aos 18 anos? **Comenta os resultados** que obtiveste, não esquecendo de incluir evidências e uma justificação para a tua resposta.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_1).

2. **Será que** os mais altos aos 2 anos também são os mais altos aos 18 anos? **Explica** em que baseaste a tua resposta.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_2).

3. Considera as crianças que têm 85 centímetros de altura aos 2 anos. **Que altura esperas** que essas crianças tenha, aos 18 anos? **Explica** como pensaste.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_3).

4. **Qual é a percentagem** de crianças de dois anos de idade que pesam mais do que 14 quilogramas?

**Justifica** a tua resposta.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_4).

5. **Será que** as crianças mais altas têm também mais força?

**Que evidências** tens sobre isso?

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_5).

6. **Comenta** a seguinte afirmação: “As crianças de 9 anos com maior largura na perna tendem a ser as mais pesadas”.

**Explica** em que te baseaste para responder.

Guarda o teu trabalho (File – Save as: nome do grupo\_II\_6).

7. **Será que** se verifica o mesmo com as crianças de 18 anos?

**Comenta** o que observaste.

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_7)

## Anexo 10. Tarefa: “Indicadores demográficos”

Os indicadores demográficos, habitualmente obtidos a partir dos censos populacionais que, em Portugal, são realizados pelo INE (Instituto Nacional de Estatística), permitem obter informações importantes relativas à população, nomeadamente a taxa de natalidade e mortalidade, o seu crescimento, a sua estrutura etária, a sua distribuição por áreas e o número dos que emigram ou imigram.

Os nascimentos em Portugal estão a diminuir e, mais drasticamente nos últimos dois anos, mais de 8% ao ano. Os futuros pais deixam o País e com eles os bebés. E Portugal fica ainda mais velho. Portugal regressou à década de 1966 no que respeita à emigração: partiram mais de 120 mil pessoas num ano. Só que nessa altura tínhamos uma das taxas de natalidade mais elevadas da Europa, mais de três filhos por mulher em idade fértil, agora temos a mais baixa, muito longe dos dois filhos. Estes dois indicadores sociais andam de mãos dadas com o crescimento económico de um país, dizem os especialistas. Se houver retoma económica; se o emprego aumentar; se os fluxos migratórios diminuïrem; haverá mais bebés em 2014. Mas acontece que "as tendências demográficas levam mais tempo a inverter-se".

([http://www.dn.pt/inicio/portugal/interior.aspx?content\\_id=3607779&page=1](http://www.dn.pt/inicio/portugal/interior.aspx?content_id=3607779&page=1))

**A evolução da população de um país, ao contrário da evolução da população mundial não depende só da natalidade e da mortalidade. Como caracteriza e relacionas essa evolução, tendo em conta o comportamento dos vários indicadores demográficos utilizados?**

---

### Parte I

Pensa na informação que tens ao teu dispor e explora os dados, que já se encontram numa base de dados do *TinkerPlots*.

**Responde e justifica** as tuas respostas que se seguem com base em representações gráficas.

1. **Será que** o número de mortes compensa o número de nascimentos? **O que podes concluir** quanto à simetria da distribuição dos dados destas duas variáveis?

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_1)

2. **Constrói um gráfico** para representar simultaneamente as variáveis “idade média da mulher ao 1.º casamento” e “idade média do homem ao 1.º casamento”. **O que podes concluir?**

2.1. No gráfico que construïste, **traça as retas correspondentes às médias de cada** uma das variáveis.

**Explica o que observas em relação ao número de dados que está presente** em cada um dos quadrantes em que o gráfico está dividido.

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_2.1.)

- 2.2. **Constrói agora um outro gráfico** para representar simultaneamente as variáveis relação de masculinidade e a taxa bruta de nupcialidade e **conclui sobre a relação entre elas.**

**2.2.1.** No gráfico que construístes, **traça novamente as retas correspondentes às médias** de cada uma das variáveis.

**Explica o que observas em relação ao número de dados** que está presente em cada um dos quadrantes em que o gráfico está dividido.

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_2.2.1.)

**2.3. Como é que** os resultados obtidos nas alíneas 3.1 e 3.2.1. te permitem avaliar a **existência de uma relação entre duas variáveis?**

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_2.3.)

**3.** Com base em representações gráficas, **comenta** sobre o que acontece à taxa de mortalidade, à taxa de natalidade e à taxa de mortalidade infantil, quando o índice sintético de fecundidade aumenta.

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_3.)

**4.** A partir das representações que construístes, **o que podes concluir sobre a relação**, entre as variáveis representadas?

**Faz corresponder** a cada uma delas um valor possível para o coeficiente de correlação linear, **justificando a tua resposta.**

**4.1.** Em que circunstâncias a relação pode ser devido a outras variáveis?

**5.** Dos gráficos construídos, **identifica as variáveis** que apresentam uma relação entre si, indicando o sentido, a intensidade e a linearidade entre os valores das variáveis?

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_I\_5)

### Parte III

**Elabora um pequeno texto** onde apresentas as tuas conclusões em relação à questão colocada e explicas em que evidências te baseias para responder.

## **Anexo 11. Tarefa: “Bateria de testes Fitnessgram”**

O programa *Fitnessgram* é um programa de educação e avaliação da aptidão física relacionada com a saúde, cuja avaliação funciona como elemento motivador para a atividade física, de forma regular, ou ainda como instrumento cognitivo para informar as crianças e jovens acerca das implicações que a aptidão física e a atividade física têm para a saúde (The Cooper Institute for Aerobics Research, 2002). Este programa avalia três componentes da aptidão física (aptidão aeróbia, a composição corporal e a aptidão muscular) consideradas importantes pela sua estreita relação com a saúde em geral e com o bom funcionamento do organismo, com base, em diversos testes específicos para cada área.

Os objetivos dos testes de condição física são (ACSM, 2009):

- Obter dados que sejam úteis para efetuar prescrições de exercício e acompanhar o progresso dos participantes nos programas de exercício físico;
- Motivar os participantes, estabelecendo objetivos em termos de condição física, razoáveis e possíveis de conseguir;
- Educar os participantes acerca dos conceitos de condição física e estado individual em relação a ela;
- Estratificação de risco.

Segundo Paulo Sena, o problema na promoção da aptidão e atividade física (estabelecimento de metas), para conseguir uma boa condição de saúde: *“O problema não está nos testes de avaliação da condição física, mas sim naquilo que fazemos com os seus resultados.”*

<http://paulosena.com/2010/04/17/testes-de-condicao-fisica-em-ginacios/>

**Nesta tarefa, tens oportunidade de tornar úteis os dados dos alunos do 9.º ano de escolaridade da tua escola, relativos aos resultados obtidos nos testes de condição física inicial, na classificação das provas de aptidão física proposta pela Bateria de Testes Fitnessgram, de acordo com as zonas de desempenho: “necessita de melhorar”, “zona saudável” e “acima da zona saudável”.**

**Tendo em conta as necessidades de manter (ou aumentar) os níveis de atividade física ou desportiva regular ao longo da vida, como caracterizas a condição física dos alunos do 9.ºano?**

---

### **Parte I**

**1. Que questões interessantes** poderias colocar e estudar, sobre estes dados quanto a:

- altura média dos alunos;
- a resistência de rapazes e raparigas;
- relação entre impulsão na vertical e a altura;
- outro aspeto que consideres relevante estudar.

**2. Qual pensas** ser a resposta a essas questões?

Explica em que te baseaste para responder à questão 1.

3. **Responde a duas das questões** que formulaste na questão 1 recorrendo a representações gráficas. **Que evidências** te dá a representação escolhida para suportar as tuas conjecturas?

Guarda o teu trabalho. (File – Save as: nome do grupo\_I)

### Parte II

Nas questões que se seguem, responde e justifica as tuas respostas com base em representações gráficas.

1. Considera a seguinte afirmação: “Os alunos mais altos tendem a ser os mais pesados e os mais rápidos”. **Comenta esta afirmação**, apresentando evidência para apoiar a tua resposta.

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_1)

2. Se recolhermos dados de alunos do 11.º ano, **esperas** que eles sejam mais rápidos? **Porquê?**

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_2)

3. **Encontras alguma relação** entre o peso dos alunos e a sua força? **Justifica** a tua resposta.

4. **Qual é a tua previsão** sobre a força de um aluno com um peso de 98kg? **Apresenta** argumentos para a tua previsão.

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_3\_4)

5. Tendo por base o conhecimento que tens sobre o Índice da Massa Corporal (IMC), **esperas que** os alunos que apresentam índices mais altos sejam também os mais pesados? **Porquê?**

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_5).

6. **Será que** os alunos com maior peso são também os que apresentam mais força? **Justifica** a tua resposta com evidências gráficas.

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_6).

7. **Comenta** sobre o que acontece à resistência quando aumenta/diminui a altura.

Guarda o teu trabalho. (File – Save as: nome do grupo\_II\_7).

8. A partir das representações que construístes nas questões 1, 2, 3, 4, 5 e 6 . **O que podes** concluir sobre a existência de relação entre as variáveis representadas? Quando possível, faz corresponder a cada uma delas um valor para o coeficiente de correlação linear, **justificando** a tua resposta.

Para cada variável, guarda o teu trabalho. (File – Save as: nome do grupo\_II\_8).

### **Parte III**

**Elabora um pequeno texto** onde apresentas as tuas conclusões em relação à questão colocada inicialmente e explicas em que evidencias te baseaste para responder.

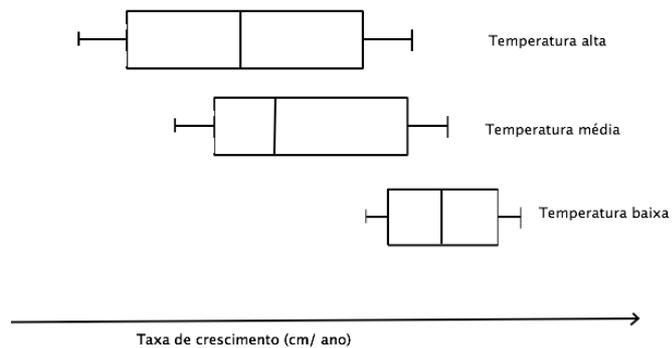
## Anexo 12. Prova de Avaliação Final

1. Foram recolhidos dados sobre a cor de cabelo de uma amostra de crianças de determinada escola. Quando forem analisados os dados, **que tipo de estatística** poderá calcular?

(A) **média**      (B) **mediana**      (C) **moda**      (D) **amplitude**

2. Foi efetuada uma experiência na qual três amostras de pinheiros foram plantadas. Cada amostra continha 50 pinheiros. Uma amostra foi criada a uma temperatura alta, outra a uma temperatura média e a terceira a uma temperatura baixa.

Os diagramas de extremos e quartis paralelos representam a taxa de crescimentos (em centímetros por ano) destas três amostras.



Dos diagramas de extremos e quartis paralelos pode-se concluir que, à medida que a temperatura passa de alta a média e de alta a baixa, a taxa de crescimento destes pinheiros:

- (a) **diminui em média e torna-se menos variável**  
 (b) **diminui em média e torna-se mais variável.**  
 (c) **aumenta em média e torna-se menos variável.**  
 (d) **aumenta em média e torna-se mais variável.**
3. No gráfico da figura 2 estão representados os pontos obtidos por cada jogadora da equipa de basquetebol feminina do Olivais Futebol Clube (Coimbra), bicampeã nacional, relativos ao play-off de 2009 e o seu tempo de jogo em minutos, no conjunto dos três jogos entre os clubes.

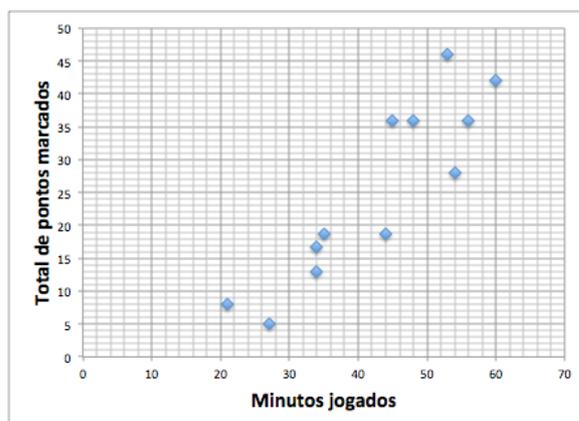


Figura 2

3.1. **Assinala** no gráfico o ponto que representa a jogadora que marcou mais pontos.

**Achas que essa jogadora** também jogou durante mais tempo?

3.2. **Achas que existe uma relação** entre o tempo de jogo (em minutos) da jogadora e o total de pontos marcados?

**Explica** a tua resposta.

3.3. Se uma jogadora joga durante 50 minutos, **quantos pontos** é de esperar que ela marque?

4. Na tabela seguinte estão alguns dados relativos à altura (cm) do Pedro.

Idade (anos)	Altura (cm)
2	78
3	84
4	92
6	110
8	120
12	145
15	160

4.1. **O que podes concluir** relativamente ao **crescimento do Pedro**? **Desenha o diagrama de dispersão** para te ajudar na resposta.

4.2. **Determina a equação da reta de regressão** e faz a respetiva representação gráfica.

4.3. **Determina** os valores aproximados da:

4.3.1. altura do Pedro quando tinha 10 anos.

4.3.2. idade do Pedro quando a altura era de 150 cm.

5. A figura 3 representa o diagrama de dispersão de duas variáveis e a respetiva reta de regressão.

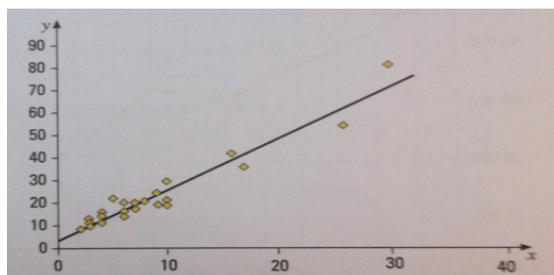


Figura3

**Qual será, aproximadamente, o valor da segunda variável** que corresponderá ao valor 25 para a primeira variável?

(A) 25

(B) 57, 72

(C) 9,99

(D) 12,5

6. **Associa a cada gráfico** de dispersão o tipo de correlação linear que te parece mais adequado.  
**Justifica** a tua resposta.

- (A) positiva fraca
- (B) **positiva forte**
- (C) **nula**
- (D) **negativa forte**
- (E) **negativa fraca**

