

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA



INTEGRAÇÃO DE DADOS DE EMPREGABILIDADE EM SISTEMA DE BI

Ana Patrícia Pinto Espinheira

Mestrado em Informática

Trabalho de Projeto orientado por:
Prof. Doutor António Manuel da Silva Ferreira
Prof. Doutora Maria Dulce Pedroso Domingos

Agradecimentos

Em primeiro lugar, queria agradecer aos meus orientadores Dr. António Ferreira e Dr^a. Dulce Domingos pelo suporte e apoio ao longo do projeto. Também queria agradecer ao Ricardo Miguel Rito e Ana Rute por toda a orientação e apoio na Reitoria da Universidade Lisboa.

A todo o Departamento de Informática da Reitoria da ULisboa pela disponibilidade e ajuda ao longo do projeto, com um especial obrigado à Ana Catarina Silva e Pedro Moita pelas conversas e risadas partilhadas.

Ao Gabinete de Estudo e Planeamento pelo entusiasmo e motivação para este projeto, especialmente à Sandra Fraga por disponibilizar bastante tempo para as minhas dúvidas e por me acompanhar e ensinar sobre a área de empregabilidade.

Obrigada ao Daniel Pires, César Pandini, Nuno Mendes, Ricardo Carvalho e João Batista por terem feito parte desta aventura. Ao Pedro Dias pelos dias de ajuda e de sabedoria sem dúvida que foram essenciais.

Sem faltar ao meu grupo favorito Tasca do Zé; Ana Salvado, Carlos Duarte, Inês de Matos, João Vicente, Raquel Nabais e Tiago Moucho pelas noites de conversas e jogos, pelos jantares e cinemas, literalmente por todos os bons momentos.

A todos os meus amigos, especialmente à Joana Nunes, Marta Castro e Mariana Parente. Um especial obrigado à minha família, pela paciência e por me suportarem ao longo destes anos todos.

Obrigada a todos.

“Stress leads to success”

— Liza Koshy

Resumo

Este trabalho, concretizado no âmbito da disciplina de Dissertação/Projeto em Informática, teve como propósito integrar dados de empregabilidade dos diplomados da Universidade de Lisboa (ULisboa) no sistema de *Business Intelligence* (BI) da Reitoria da ULisboa, motivado pelo processamento, até então, manual deste tipo de dados pelo Gabinete de Estudos e Planeamento (GEP) da Reitoria da ULisboa. Assim, foram considerados três objetivos principais, resumidos em baixo.

O primeiro objetivo incluiu a análise estatística dos dados da 3ª edição do inquérito à empregabilidade dos diplomados da ULisboa, de modo a entender os campos e as suas gamas de valores, bem como a compreensão das necessidades do GEP e recolha de requisitos, tais como: indicadores de empregabilidade e conjuntos de relatórios pertinentes.

O segundo objetivo consistiu no desenvolvimento das componentes de extração, transformação e carregamento (ETL) de dados e a criação de relatórios para análise. Para este efeito foi feita uma aprendizagem das tecnologias e do sistema de BI existente, após o que se passou ao desenho do modelo de dados. Os passos seguintes consistiram na implementação das componentes de ETL, de uma camada que transforma dados não inteligíveis em conceitos de negócio, e por último a criação dos relatórios indicados pelo GEP usando os dados do primeiro objetivo. Cada passo foi validado pelo GEP tendo em conta os requisitos.

O terceiro objetivo foi fazer a avaliação da solução desenvolvida, através da comparação entre os dados obtidos manualmente e com a nova solução, com base na 3ª edição do inquérito à empregabilidade, existindo 0,2% de diferenças entre os resultados. Depois, a solução foi aplicada aos dados de uma nova edição do inquérito de empregabilidade, disponibilizando novos dados para análise na nova solução de BI desenvolvida. Por fim, foi constatada a redução de 4 semanas para alguns minutos entre o método manual e a nova solução.

A solução desenvolvida estruturou e consolidou a informação de empregabilidade, disponibilizando numa interface única os dados para consulta e análise, possibilitou novas análises de futuras edições do inquérito de empregabilidade e reduziu o tempo de trabalho do GEP dando oportunidade para outras análises.

Palavras-chave: *Business Intelligence*, *Data Warehouse*, tomada de decisões, empregabilidade, relatórios.

Abstract

This work, carried out within the scope of the Dissertation/Project in Informatics, aimed to integrate employability data of graduates from University of Lisbon (ULisboa) into the Business Intelligence (BI) system of the Rectory of the ULisboa. The processing and analysis of employability data was a time consuming task that the ULisboa office of studies and planning (GEP) performed manually.

The first objective included statistical analysis of data from the 3rd edition of the ULisboa Graduate Employability Survey, in order to understand the fields and their ranges of values, as well as understanding GEP needs and gathering requirements such as: employability indicators and relevant report sets.

The second objective was the development of data extraction, transformation and loading (ETL) components and the creation of reports for analysis. For this purpose, an appraisal of the technologies and the existing BI system was made, after which was preceded the design of the data model. The next steps were the implementation of ETL components, a layer that transforms unintelligible data into business concepts, and finally the creation of reports indicated by GEP using data from the first objective. Each step was validated by GEP taking into account the requirements.

The third objective was to evaluate the developed solution by comparing the data obtained manually and with the new solution, based on the 3rd edition of the employability survey, with 0.2% differences between the results. The solution was then applied to data from a new edition of the employability survey, making new data available for analysis in the newly developed BI solution. Finally, a reduction from 4 weeks to a few minutes between the manual method and the new solution was found.

The developed solution structured and consolidated the employability information, providing data for consultation and analysis in a single interface, enabling further analysis of future editions of the employability survey and reducing the working time of the GEP, giving opportunity for further analysis.

Keywords: Business Intelligence, Data Warehouse, decision-making, employability, reports.

Conteúdo

Índice de figuras	xv
Índice de tabelas	xvi
Índice de listagens	xviii
1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	2
1.3 Instituição de acolhimento	2
1.4 Contribuições	3
1.5 Notação adotada	4
1.6 Estrutura do documento	4
2 Conceitos e trabalho relacionado	5
2.1 <i>Business intelligence</i>	5
2.1.1 Fontes de dados	5
2.1.2 Sistema de extração-transformação-carregamento	6
2.1.3 Projeto de construção de um <i>data warehouse</i>	7
2.1.4 Modelação dimensional	8
2.1.5 Visualização de dados	9
2.2 <i>Business Intelligence</i> na ULisboa	10
2.2.1 Arquitetura funcional e tecnológica	10
2.2.2 Fontes de dados	11
2.2.3 Armazenamento e processamento	11
2.2.4 Camada semântica	12
2.2.5 Exploração e visualização de dados	12
2.2.6 Colaboração	12
2.2.7 Administração, segurança e auditoria	12
2.2.8 Arquitetura de servidores	13
2.2.9 Principais ferramentas em uso	13
2.3 Sumário	16

3	Análise e desenho	17
3.1	Introdução do trabalho	17
3.1.1	Método de trabalho	17
3.1.2	Tecnologias utilizadas	17
3.1.3	Mapa do trabalho realizado	18
3.2	Análise do problema	20
3.2.1	Análise de requisitos	21
3.2.2	Fluxo de trabalho do GEP	21
3.2.3	Fontes de dados utilizadas	23
3.2.4	Análise de dados	24
3.2.5	Indicadores de empregabilidade	26
3.2.6	Relatórios prioritários	27
3.3	Desenho da solução	27
3.3.1	Modelo de dados inicial	28
3.3.2	Modelo de dados final	28
3.3.3	Arquitetura da solução	33
3.4	Sumário	33
4	Implementação e avaliação	35
4.1	Implementação da solução	35
4.1.1	Desenvolvimento das componentes ETL	35
4.1.2	Desenvolvimento da camada semântica	48
4.1.3	Criação dos relatórios	52
4.1.4	Criação e configuração de utilizadores	53
4.2	Avaliação	55
4.2.1	Comparação entre os dados manuais e da solução desenvolvida	55
4.2.2	Aplicação da solução a um novo ano de inquérito	58
4.2.3	Comparação de tempos de processamento de dados	59
4.3	Sumário	63
5	Conclusão	65
5.1	Principais contribuições	65
5.2	Competências adquiridas	66
5.3	Desafios encontrados	66
5.4	Trabalho futuro	67
	Bibliografia	71
	Abreviaturas	73
A	Manual de instalação das ferramentas SAP Server	75

B	Manual de instalação das ferramentas SAP Client	89
C	Análise de dados da 3ª edição do inquérito à empregabilidade	99
D	Modelo de dados inicial	223

Índice de figuras

2.1	Fluxo de trabalho para a construção de um DW	7
2.2	Representação de um esquema em estrela	9
2.3	Arquitetura da solução de BI da Reitoria da ULisboa	10
2.4	Servidores de desenvolvimento e qualidade	13
2.5	Servidores de produção	14
3.1	Mapa do trabalho realizado	19
3.2	Fluxo de trabalho do GEP	22
3.3	Estrutura geral do inquérito à empregabilidade	24
3.4	Estrutura do modelo final na ferramenta SAP IDT	32
3.5	Arquitetura da solução de BI desenvolvida	33
4.1	<i>Jobs</i> responsáveis pela componente de extração	36
4.2	Objetos do <i>job</i> que extrai os dados da Faculdade de Ciências	37
4.3	<i>Dataflow</i> que extrai os dados da FC	37
4.4	<i>Jobs</i> que agregam os dados das escolas autónomas com os das não autónomas	37
4.5	Objetos do <i>job</i> JOB_UNIAO_ESCOLAS_AUTONOMAS	38
4.6	Objetos do <i>job</i> JOB_LOAD_ESCOLAS_AUTONOMAS	39
4.7	Exemplo de tradução de valores de uma coluna	39
4.8	<i>Job</i> responsável pela união dos dados de todas as escolas da ULisboa . . .	40
4.9	<i>Dataflow</i> responsável pela união dos dados de todas as escolas da ULisboa	40
4.10	Objetos do <i>job</i> JOB_TRANSFORMACAO_DADOS, com cinco <i>dataflows</i> . .	41
4.11	Estrutura do <i>dataflow</i> DF_ETAPA2_AGREGACAO	41
4.12	<i>Dataflow</i> DF_ETAPA3_APL_NAO_RESP, responsável por diferenciar os campos 998 e 999	45
4.13	Exemplo de uma regra aplicada a uma pergunta do inquérito de emprega- bilidade	46
4.14	Estrutura do <i>dataflow</i> DF_ETAPA4_NOMENCLATURA	46
4.15	Estrutura do <i>dataflow</i> DF_ETAPA6_DUPLICADOS	47
4.16	Estrutura do <i>dataflow</i> DF_FINAL	47
4.17	Estrutura do projeto na ferramenta SAP IDT para a criação do universo .	48
4.18	Estrutura do modelo dimensional criado na ferramenta SAP IDT	49

4.19	Painel inicial no WebI para análises de empregabilidade	53
4.20	Relatório predefinido de diplomados por grau, do inquérito de 2017 . . .	54
4.21	Resultados dos diplomados por grau com o método manual e com a solução	57
4.22	Resultados do tipo de vínculo de trabalhadores por conta de outrem por grau com o método manual e com a solução	58
4.23	<i>Boxplot</i> dos tempos de execução de cada <i>job</i>	61
4.24	<i>Boxplot</i> dos tempos de execução dos <i>jobs</i> mais rápidos	62
4.25	<i>Boxplot</i> dos tempos totais de execução de todos os <i>jobs</i>	62
D.1	Esboço do primeiro modelo de dados de empregabilidade	223

Índice de tabelas

3.1	Indicadores de empregabilidade	26
3.1	Indicadores de empregabilidade (continuação)	27
3.2	Descrição das dimensões do modelo de dados inicial	28
3.3	Tabela final da base de dados de empregabilidade	28
4.1	Variáveis auxiliares pedidas pelo GEP e desenvolvidas no SAP IDT	49
4.1	Variáveis auxiliares pedidas pelo GEP e desenvolvidas no SAP IDT (continuação)	50
4.2	Diferença absoluta entre os resultados obtidos pela solução e os dados obtidos pelo processo do GEP no inquérito de 2017	55
4.2	Diferença absoluta entre os resultados obtidos pela solução e os dados obtidos pelo processo do GEP no inquérito de 2017 (Continuação)	56
4.3	Tempo de cada processo realizado manualmente pelo GEP	60
4.4	Tempo médio de execução de cada <i>job</i> da solução de BI desenvolvida	60

Índice de listagens

4.1	<i>Script</i> SQL para consolidar perguntas em comum do inquérito de empregabilidade	42
4.2	Implementação da variável auxiliar correspondente à internacionalização dos diplomados presente na dimensão Emprego	50
4.3	Implementação da métrica de contagem dos diplomados inquiridos	51
4.4	Implementação do indicador de taxa de empregabilidade	51

Capítulo 1

Introdução

O presente relatório descreve o projeto de integração de dados de empregabilidade no sistema de *Business Intelligence* (BI) da Reitoria da ULisboa, desenvolvido no âmbito da disciplina de Dissertação/Projeto em Informática do Mestrado em Informática da Faculdade de Ciências da Universidade de Lisboa.

Neste capítulo descrevem-se a motivação do projeto e os seus objetivos, a instituição onde o trabalho foi realizado, a notação adotada e a estrutura do relatório.

1.1 Motivação

O mercado de hoje em dia é muito dinâmico e a quantidade de dados gerados cresce cada vez mais, pelo que a integração de dados é essencial para obter informações de qualidade. Como o mundo digital evoluiu bastante nos últimos anos, o surgimento de novas tecnologias tem vindo a melhorar o tratamento e armazenamento de informação sem perder integridade e segurança, permitindo às empresas adequar os seus negócios às necessidades atuais.

No que toca ao tratamento de dados, também tem surgido uma evolução na forma de tirar o melhor partido deles, gerando mais e melhor informação, através da chamada *Business Intelligence* (BI), cuja definição pode ser considerada como um conjunto de processos, aplicações e tecnologias desenhadas para apoiar de forma eficiente e eficaz os processos de decisão nas empresas [10], transformando dados em informação útil, de forma a auxiliar a gestão de fluxos de informação de negócio dentro de uma organização [9]. Com a BI é possível gerar relatórios e *dashboards*, que permitem aos utilizadores acompanhar o estado do negócio, analisar tendências com base em dados históricos, realizar simulações de cenários e prever situações.

A mesma necessidade de extrair informação útil a partir de vastos fluxos de dados, motivou a Reitoria da ULisboa a criar uma solução baseada em BI para satisfazer as necessidades dos utilizadores. Em particular, o procedimento na Reitoria da ULisboa responsável pela informação de empregabilidade dos diplomados não era automatizado,

sendo necessária uma transformação a nível tecnológico para facilitar as análises e tomadas de decisões por parte do Gabinete de Estudos e Planeamento (GEP). Consequentemente, foi proposto o desafio de criar uma solução de BI, cujo principal objetivo foi o desenvolvimento de uma extensão do sistema de BI existente que ajudasse o GEP a explorar e analisar de forma automatizada as fontes de dados de empregabilidade, correspondentes a cada uma das 18 escolas que a ULisboa engloba.

1.2 Objetivos

Para este projeto, os objetivos principais foram:

- 1. Análise de dados e requisitos.** Analisar os dados obtidos através da 3^a edição de inquéritos de empregabilidade realizados a alunos que já concluíram os seus estudos, de modo a caracterizar o percurso e situação profissional. Pretendeu-se igualmente compreender as necessidades e expectativas dos utilizadores do sistema de BI face à empregabilidade, incluindo o estudo de novos indicadores de empregabilidade. Este objetivo foi avaliado pelo GEP.
- 2. Desenvolvimento de componentes ETL integradas no sistema de BI e relatórios.** Conhecer as tecnologias usadas no sistema de BI da ULisboa e desenvolver as componentes de extração, transformação e carregamento (ETL) dos mesmos dados utilizados no primeiro objetivo e integrar no sistema de BI existente. Mais do que o tratamento de dados, a integração também inclui a criação de visualizações estruturadas e apelativas, necessárias para mostrar a correção e utilidade dos dados e indicadores sobre empregabilidade e para facilitar as análises do GEP. Este objetivo foi avaliado pelo GEP.
- 3. Avaliação da solução desenvolvida.** Pretendeu-se, numa primeira fase utilizar os dados do primeiro e segundo objetivo e comparar os resultados obtidos manualmente e com solução desenvolvida, e respetivos tempos de execução do trabalho e caso a primeira fase fosse bem sucedida, estabeleceu-se como meta seguinte aplicar a solução desenvolvida a uma segunda edição do inquérito de empregabilidade.

1.3 Instituição de acolhimento

Este projeto foi desenvolvido para a Universidade de Lisboa, considerada como a maior universidade de Portugal e uma das maiores da Europa, tendo atualmente mais de 47 mil alunos inscritos [21]. A ULisboa integra 18 escolas, 8 *campus*, entre outras estruturas e serviços, dando trabalho a cerca de 5800 colaboradores (incluindo docentes, investigadores, bolseiros, técnicos e administradores). Também tem uma boa integração

de alunos estrangeiros, acolhendo cerca de 7000 estudantes internacionais de 106 nacionalidades. Em relação à empregabilidade, os dados permitem concluir que cerca de 84% dos diplomados das escolas da ULisboa encontram-se empregados e com um salário médio bruto de 1128 € [21].

A ULisboa considera a modernização administrativa um fator essencial para a execução eficaz e eficiente de todos os seus processos administrativos, científicos, e educativos, sendo por isso essenciais a renovação, normalização e integração dos sistemas de gestão das escolas que compõem a Universidade. Os alvos principais de integração são [21]:

- Os sistemas de gestão académicos;
- Os sistemas de gestão financeiros;
- Os sistemas de gestão de recursos humanos;
- Os sistemas de gestão científicos.

A Reitoria da ULisboa, tem serviços centrais que coordenam, organizam e apoiam todas as entidades que compõem a ULisboa e prestam suporte ao Reitor e à Equipa Reitoral. A Reitoria está organizada em gabinetes, departamentos e áreas, sendo um desses departamentos o de Informática, o qual tem a seu cargo as matérias associadas às tecnologias de informação e de comunicação da universidade [21]. Este projeto foi realizado nesse departamento e focou-se na gestão académica.

1.4 Contribuições

A solução de BI desenvolvida neste projeto disponibiliza informação útil, estruturada e consolidada através de uma interface única permitindo aos utilizadores criar análises de forma ágil. Antes da solução desenvolvida, o trabalho era realizado manualmente pelo GEP, demorando cerca de dois meses a ser concluído. Todo o processo era custoso em tempo e recursos.

Com esta solução, a execução do trabalho foi significativamente automatizada, diminuindo o tempo de execução bem como a possibilidade de erros humanos, e dando oportunidade ao GEP de fazerem mais ou outras análises.

O processo de automatização desta nova solução passou pelo acompanhamento e sistemáticas validações do GEP para que, as diferenças entre o processo manual e automático fossem reduzidas. Utilizando os dados referentes à 3ª edição do inquérito à empregabilidade, a diferença entre ambos os processos verificou-se ser de 0,2%.

Com a diferença reduzida entre os dois processos, o objetivo seguinte consistiu em aplicar a solução desenvolvida a uma edição seguinte do inquérito à empregabilidade, onde se verificou de forma generalizada que os resultados obtidos entre os dois processos não apresentaram diferenças significativas.

A solução desenvolvida permitiu estender o sistema de BI existente na Reitoria com a informação referente à empregabilidade dos diplomados da ULisboa, tendo havido o cuidado adicional de serem criados e disponibilizados utilizadores para garantir acesso à solução e manuais de instruções para guiar os utilizadores e facilitar futuros desenvolvimentos.

1.5 Notação adotada

Este documento está escrito em português seguindo o novo acordo ortográfico e todas as expressões noutra idioma encontram-se em itálico.

1.6 Estrutura do documento

O documento tem a seguinte estrutura:

Capítulo 1 - Introdução Apresenta a motivação do trabalho e objetivos, a descrição da instituição, principais contribuições, notação adotada e estrutura do documento.

Capítulo 2 - Conceitos e trabalho relacionado Descreve os conceitos teóricos de uma solução de BI, bem como os diferentes componentes já existentes no sistema de BI da Reitoria da ULisboa.

Capítulo 3 - Análise e desenho Apresenta a caracterização do ambiente de trabalho e de um mapa com os diferentes passos do projeto. Inclui a análise do problema e o desenho da solução, contendo a discussão de alternativas.

Capítulo 4 - Implementação e avaliação Apresenta os passos envolvidos na implementação da solução de BI e conclui com a avaliação do trabalho realizado.

Capítulo 5 - Conclusão Descreve as principais contribuições, competências adquiridas, desafios encontrados e trabalho futuro.

Capítulo 2

Conceitos e trabalho relacionado

Este capítulo fornece uma vista de alto nível do projeto de *Business Intelligence* (BI) existente na Reitoria da ULisboa, começando por introduzir e explicar conceitos fundamentais para a compreensão deste. A arquitetura foi um dos pontos focados, tal como a análise da documentação existente, bem como as tecnologias usadas.

2.1 *Business intelligence*

O conceito de *Business Intelligence* pode ser entendido como uma coleção de abordagens de recolha, tratamento, análise e apresentação dos dados aos utilizadores, para uma melhor visão e para boas tomadas de decisão por parte de uma organização [10, 5].

Alguns dos principais benefícios para os utilizadores são:

- Oferece informação transparente, compreensível e a tempo;
- Fácil personalização de relatórios e ferramentas para um utilizador específico;
- Reduz a quantidade de trabalho manual na preparação de relatórios, ficando mais tempo para trabalho analítico;

Existem vários conceitos para a construção de uma solução de BI, sendo explicados ao longo desta secção.

2.1.1 Fontes de dados

O sucesso da informação depende da qualidade das fontes de dados utilizadas, que devem ser bem identificadas de modo a responder às perguntas de negócio. O conhecimento das fontes de dados é relevante para a compreensão das mesmas, devendo ser percebido tanto por elementos de tecnologia de informação, bem como por conhecedores do negócio e que fazem parte do quotidiano da organização [4].

Habitualmente, as fontes de dados dos sistemas de BI são os sistemas operacionais, encarregues de registar dados de rotina gerados diariamente numa organização, juntamente com fontes externas, com dados oriundos de fora da organização [10].

2.1.2 Sistema de extração-transformação-carregamento

Um *Data Warehouse* (DW) é utilizado num sistema de BI para armazenar informação consolidada de uma organização, sendo composto pela *staging area* e a área de apresentação de dados, descritas abaixo.

Staging area consiste tanto numa área de armazenamento de dados extraídos dos sistemas operacionais, bem como um conjunto de processos conhecidos como *extract-transformation-load* (ETL). Todos os elementos pertencentes a esta área encontram-se entre os sistemas operacionais (fontes de dados) e a área de apresentação de dados [8].

Área de apresentação de dados é o local principal onde se encontra a informação necessária para os processos analíticos e a tomada de decisão numa organização. Pode ser definido como um repositório integrado, onde são armazenados dados de fontes internas e externas [10]. Como os *data warehouses* contêm dados consolidados, possivelmente de várias bases de dados operacionais durante períodos de tempo potencialmente longos, tendem a ter uma magnitude maior do que as base de dados operacionais [2, 3].

Esta área é também referida como uma série de *data marts* integrados. *Data marts* são sub-conjuntos lógicos do DW, normalmente organizados segundo os departamentos onde os utilizadores trabalham.

Processos de ETL têm um procedimento bastante custoso e intensivo. É típico na criação de um DW que o maior esforço seja na concretização destes processos (cerca de 70%) [5], descritos em baixo [8]:

- **Extração** é um processo para obter dados para serem armazenados no DW. Extrair significa ler e entender as fontes de dados para que depois os dados necessários sejam copiados para a *staging area*.
- **Transformação** é o processo a iniciar depois dos dados terem sido extraídos para a *staging area*. Aqui são realizadas várias transformações aos dados que podem ser correções de erros ortográficos, lidar com elementos vazios, passagem para formatos *standard*, combinar dados de múltiplas fontes, lidar com dados duplicados e atribuir identificadores.
- **Carregamento** é o último processo, no qual são carregados os dados que foram extraídos e transformados para a área de apresentação dos dados, de forma a estarem prontos para serem utilizados.

2.1.3 Projeto de construção de um *data warehouse*

A construção de um *data warehouse* tipicamente segue o fluxo de trabalho apresentado na Figura 2.1.

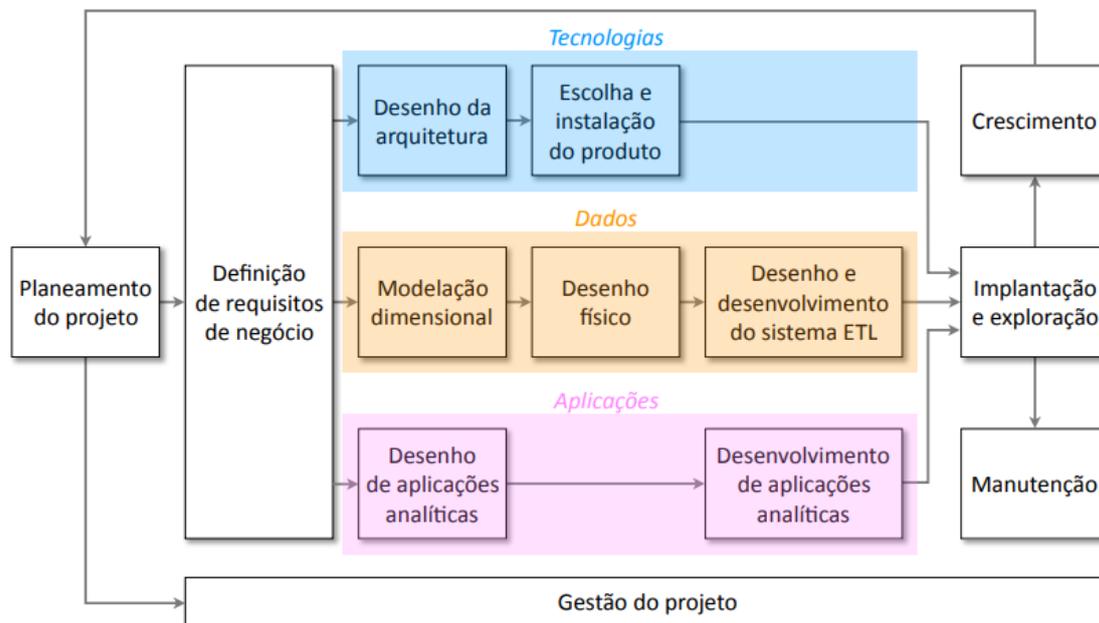


Figura 2.1: Fluxo de trabalho para a construção de um DW [8]

O fluxo contém algumas atividades iniciais como o planeamento do projeto, onde são identificadas e estimadas as tarefas, tempos, recursos e custos, inclui a gestão do projeto que é contínuo ao longo do projeto e é onde são monitorizadas as tarefas bem como a comunicação entre os técnicos e decisores, também inclui a definição de requisitos de negócio, onde são compreendidos os fatores críticos de sucesso do negócio.

A vertente das tecnologias contém o desenho da arquitetura que se baseia na infraestrutura tecnológica existente e a escolha e instalação do produto que inclui a avaliação de alternativas no mercado.

A vertente de dados integra as atividades de modelação dimensional que definem os processos, factos, dimensões, medidas e outros, integra também o desenho físico onde são tomadas decisões sobre índices, compressão de dados, partições, entre outros para otimizar o desempenho e por fim, o desenho e desenvolvimento do sistema ETL, onde são concretizadas a extração, transformação e carregamento dos dados para o DW.

A vertente de aplicações contém o desenho e desenvolvimento de aplicações analíticas onde são desenhados relatórios dinâmicos, *dashboards*, entre outros tendo em consideração as funcionalidades do produto escolhido na vertente das tecnologias.

As atividades finais englobam a implantação e exploração que trata da instalação do DW, documentação e formação aos decisores. A manutenção que trata de monitorizar e

afinar o desempenho, cópias de segurança, entre outros e o crescimento que concretiza a inclusão de novos dados e fontes, bem como a adaptação a novas necessidades.

2.1.4 Modelação dimensional

É uma técnica específica para desenhar *data warehouses*. Este método foi desenvolvido a partir de observações da prática, e em particular, de fornecedores de dados que estão no negócio de proporcionar dados de forma “amigável” aos seus clientes, tendo o seu uso sido um sucesso [11].

O modelo dimensional é constituído por **factos**, **dimensões**, e **medidas**. Uma medida é um valor numérico usado para avaliar um processo de negócio, uma dimensão é uma entidade independente que participa na análise de medidas, e um facto expressa um evento associado ao tema da modelagem, através de dimensões e medidas.

Tabela de factos

A tabela de factos é a tabela principal num modelo dimensional, onde os valores numéricos das medidas são armazenados [8].

As tabelas de factos ocupam a maioria do espaço do DW, por volta de 90%, e normalmente têm um grande número de linhas e poucas colunas.

Após compreender os requisitos de negócio, tendo em conta os dados disponíveis, é então necessário definir a granularidade. O **grão** é o nível de detalhe dos dados que devem ficar disponíveis no modelo dimensional. Todas as medidas numa tabela de factos devem estar ao mesmo nível de granularidade [8].

Tabelas de dimensão

As tabelas de dimensão contêm descrições textuais do negócio. Num modelo dimensional é comum que as tabelas de dimensão contenham muitas colunas e atributos e poucas linhas. Cada dimensão é definida com chaves primárias únicas que, por sua vez, são referenciadas em qualquer tabela de factos à qual estejam ligadas. É significativo para um DW, que os atributos das dimensões sejam bons [8]. Por norma, as tabelas de dimensão apenas representam 10% dos dados e têm como exemplo de atributos:

- Chave substituta;
- Atributos, como elementos descritivos, incluindo códigos, descrições, nomes, estados, entre outros;
- Informação hierárquica;
- A chave natural que os utilizadores funcionais usam para identificar o elemento como único;
- Campos de controlo orientados ao tipo de captura de histórico da dimensão.

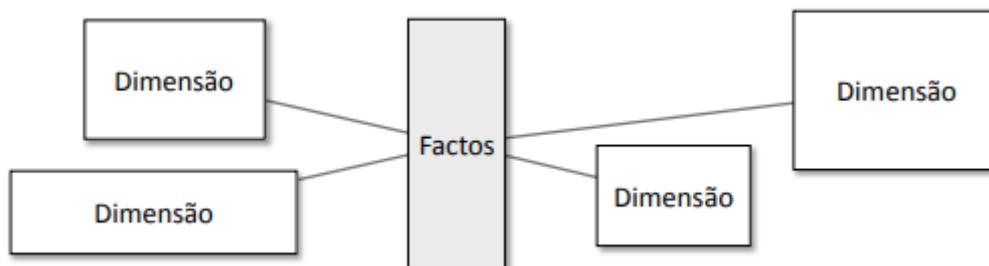


Figura 2.2: Representação de um esquema em estrela

As chaves substitutas são números inteiros atribuídos sequencialmente quando necessários para popular uma dimensão e são independentes das chaves de negócio. Servem para unir as tabelas de dimensão com a tabela de factos [8].

Hierarquias mostram as relações entre grupos de colunas numa tabela de dimensões. Por exemplo, os trimestres contêm meses e os meses contêm dias. Com hierarquias é possível fazer *drill down/up* nos relatórios, permitindo ao utilizador realizar análises com vários níveis de detalhe. Uma tabela de dimensão pode ter uma ou mais hierarquias. Uma hierarquia normalmente começa com um nível geral, seguido de níveis filhos até ao nível mais detalhado e o *drill* é concretizado sobre estes diferentes níveis [7].

Esquema em estrela

O modelo dimensional do DW pode ser representado com um esquema em estrela, no qual várias tabelas de dimensão estão ligadas a uma tabela de factos, como mostrado na Figura 2.2.

2.1.5 Visualização de dados

Este é um componente do ambiente do DW que fornece aos utilizadores uma zona de apresentação visual dos dados para a tomada de decisões analíticas. Por norma são realizadas *queries* aos dados que estão no DW que podem ser relativamente simples de escrever, ou bastante complexas, nomeadamente no âmbito da prospeção de dados [8].

Existem várias aplicações de *front-end* que permitem o acesso e a manipulação da informação através das quais os gestores executam tarefas de BI, como o caso dos portais *web* de pesquisa e das aplicações de gestão de desempenho que permitem aos decisores acompanharem os principais indicadores de desempenho da empresa usando *dashboards* e consultas *ad-hoc* [2].

Dashboards mostram de forma sumariada grandes quantidades de dados, relacionados com o desempenho da organização. Estes podem ser apresentados em apenas um ecrã ou vários, dando capacidade aos utilizadores de rapidamente compararem o desempenho atual com objetivos ou desempenhos anteriores [22].

Relatórios Ad-hoc são relatórios criados pelo utilizador, no momento em que quiser e com o conteúdo que necessitar, sendo por vezes melhores que os relatórios predefinidos [13].

2.2 Business Intelligence na ULisboa

Nesta secção é descrita a arquitetura da solução de BI implementada na Reitoria da ULisboa, bem como explicado cada componente principal da arquitetura. Adicionalmente são apresentadas as tecnologias e ferramentas, tendo adotado várias para o desenvolvimento da solução de BI a que se refere este relatório.

2.2.1 Arquitetura funcional e tecnológica

A arquitetura funcional e tecnológica da solução de BI implementada na Reitoria da ULisboa é composta por 6 componentes, como apresentado na Figura 2.3.

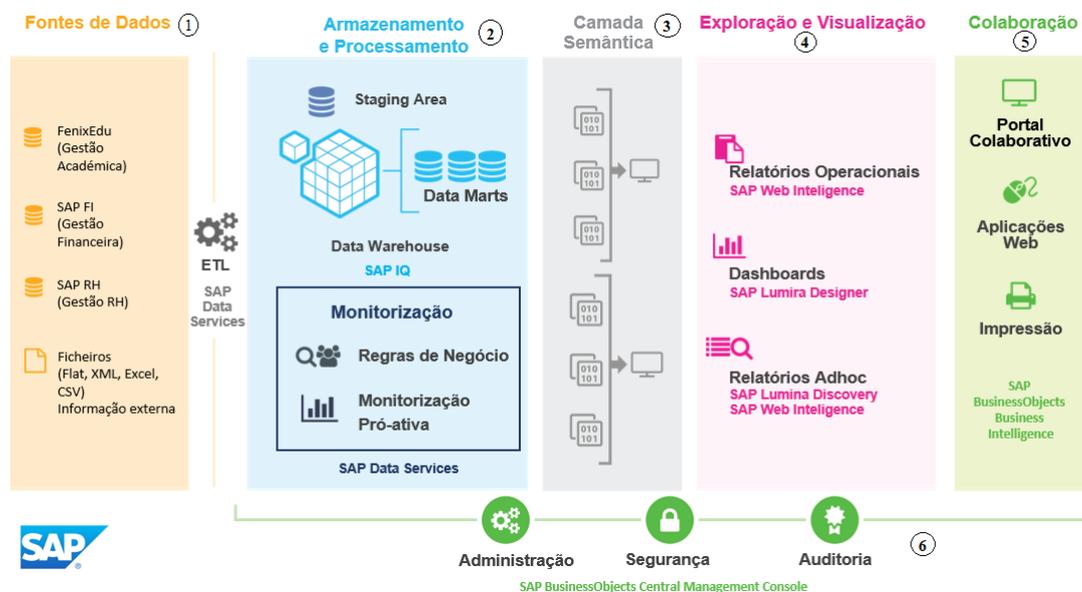


Figura 2.3: Arquitetura da solução de BI da Reitoria da ULisboa [12]

Dois dos principais objetivos da implementação desta solução é que seja dinâmica e ajustável, para que quando alguma alteração seja integrada não haja grande impacto em todos os componentes.

Assim, por exemplo, uma alteração nas fontes de dados (1) apenas deverá ter impacto na camada de *staging* (2) e ETL, e alterações nos relatórios deverão apenas ter impacto na camada de exploração e visualização (4).

As seis componentes da arquitetura de BI da Reitoria são descritas com mais detalhe nos próximos pontos do relatório, adotando a mesma nomenclatura apresentada na Figura 2.3.

2.2.2 Fontes de dados

No projeto de BI da Reitoria da ULisboa, o carregamento de dados é feito por cópia dos dados existentes nos diversos sistemas da ULisboa e que guardam as informações de negócio produzidas diariamente. As fontes de dados dependem da área de gestão, sendo que a gestão académica obtém os dados do FenixEdu, a gestão financeira usa o SAP FI e a gestão de recursos humanos utiliza o SAP HR, também sendo possível obter dados através de informação externa.

2.2.3 Armazenamento e processamento

Os dados provenientes das diferentes fontes devem ser armazenados e tratados, sendo essa a responsabilidade deste componente.

Armazenamento

Os dados obtidos são armazenados num repositório de informação, o qual se encontra dividido em duas áreas:

Staging area Trata-se de uma área de trabalho temporária que integra os dados das diferentes fontes, garante a qualidade, uniformiza os conceitos e aplica regras de negócio para que estes sejam analisados de forma uniforme e integrada. Os utilizadores não têm qualquer acesso aos dados presentes nesta área.

Data warehouse Contém toda a informação uniformizada, sendo a fonte de verdade única para processos de suporte de tomada de decisão e estando acessível para consulta por parte dos utilizadores. O DW armazena uma janela histórica de informação na ordem de diversos anos, dependendo das necessidades da Universidade de Lisboa.

Processamento

Trata-se de um conjunto de processos de extração, transformação e carregamento de dados que devem cumprir as seguintes responsabilidades:

- Transferência de conjuntos de informação, desde as fontes de dados para os diversos componentes da solução;
- Tratamento e uniformização de conceitos, métricas e regras de negócio da ULisboa;
- Envio de alertas sobre problemas de qualidade de dados e mapeamentos de informação;
- Possibilidade de rastreabilidade da informação.

2.2.4 Camada semântica

Nesta camada é realizada a transformação das tabelas e campos existentes na base de dados (DW) para conceitos de negócio de fácil compreensão para os utilizadores da ULisboa, a fim de facilitar a realização de análises na componente de exploração e visualização.

Nesta camada é também possível definir o nível de segurança de acesso à informação, garantindo que os dados só estão disponíveis se o utilizador tiver permissões para tal.

2.2.5 Exploração e visualização de dados

Permite a análise e exploração dos dados de negócio por parte dos utilizadores, sendo fundamental suportar vários tipos de análise, nomeadamente: relatórios web predefinidos, *dashboards* e relatórios *Ad-Hoc*.

Os relatórios web predefinidos, estão formatados de acordo com as necessidades dos utilizadores e não possibilitam manipulações da informação.

Os *dashboards* têm uma apresentação visual da informação mais relevante e necessária para os objetivos de negócio, consolidada num único ecrã de modo a ser monitorizada e percecionada de forma imediata. São normalmente estruturados em objetos gráficos que possibilitam a análise rápida da informação. É possível configurar estes *dashboards* com parâmetros que filtram a informação, seja por período temporal, zona geográfica e outros.

Nos relatórios *Ad-Hoc* é disponibilizada uma interface de fácil utilização, intuitiva, que permite a construção, visualização e exploração de dados pelos utilizadores, sem necessidade de apoio por parte de equipas técnicas.

2.2.6 Colaboração

O objetivo principal desta componente é a partilha de informação, sendo que (relatórios, *dashboards*, entre outras visualizações) podem ser publicadas e partilhadas pela organização, criando uma documentação relevante para o dia-a-dia. Existe a possibilidade de implementar uma camada de segurança que impede o acesso a utilizadores não autorizados.

2.2.7 Administração, segurança e auditoria

A administração do sistema de BI da Reitoria da ULisboa permite a gestão de utilizadores, acessos, conteúdos, perfis, servidores, aplicações, entre outras atividades.

A segurança permite gerir os conteúdos e camadas (dados, semântica e visualização) a que cada utilizador tem acesso e respetivas funcionalidades, como por exemplo, criar, editar e eliminar documentos corporativos ou pessoais.

DESENVOLVIMENTO E QUALIDADE		
 ETL	 DATA WAREHOUSE	 VISUALIZATION
CARACTERÍSTICAS MÍNIMAS: CPU (4 CORES) RAM 32 GB CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 64 GB	CARACTERÍSTICAS MÍNIMAS: CPU (4 CORES) RAM 64 GB CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 64 GB	CARACTERÍSTICAS MÍNIMAS: CPU (8 CORES) RAM 32 GB CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 64 GB
STORAGE: 100 GB OS / 100 GB APP / 120 GB DATA / 60GB CACHE / 16GB PG	STORAGE: 100 GB OS / 100 GB APP / 1 Tb DATA / 32GB PG	STORAGE: 100 GB OS / 100 GB APP / 120 GB DATA / 60GB CACHE / 16GB PG
COMPONENTES: <ul style="list-style-type: none"> • SAP BO IPS 4.2 • SAP BO DS 4.2 • TOMCAT 8 • SQL ANYWHERE 16 	COMPONENTES: <ul style="list-style-type: none"> • SAP IQ 16 	COMPONENTES: <ul style="list-style-type: none"> • SAP BO BI 4.2 • SAP BO DESIGNSTUDIO 2BI • SAP BO LUMIRA 4 BI • SAP BO PREDICTIVE • TOMCAT 8 • SQL ANYWHERE 16

Figura 2.4: Servidores de desenvolvimento e qualidade [12]

Ao nível da auditoria, é possível perceber qual a frequência de utilização dos objetos num determinado período de tempo, com o intuito de identificar os padrões de utilização.

2.2.8 Arquitetura de servidores

A arquitetura de servidores define os ambientes de trabalho e os servidores que os compõem, nomeadamente:

- **Desenvolvimento** - ambiente usado pela equipa de desenvolvimento para implementação de soluções.
- **Qualidade** - ambiente controlado e dedicado à equipa de utilizadores responsáveis pela validação de novas funcionalidades.
- **Produção** - ambiente disponível para os utilizadores finais, após todas as validações nos ambientes anteriores terem decorrido.

Os ambientes de desenvolvimento e qualidade partilham os mesmos servidores, através da criação de duas instâncias diferentes por servidor para as componentes de ETL, *Data Warehouse* e Visualização, como apresentado na Figura 2.4.

O ambiente de produção tem os seus próprios servidores para as componentes de ETL, DW, visualização específica de *dashboards* e outro para visualização, como indica a Figura 2.5.

2.2.9 Principais ferramentas em uso

Nesta secção são descritas as ferramentas SAP essenciais para o desenvolvimento do sistema de BI na Reitoria da ULisboa.

PRODUÇÃO			
 ETL	 DATA WAREHOUSE	 VISUALIZATION (Dashboards)	 VISUALIZATION
CARACTERÍSTICAS MÍNIMAS: CPU (4 CORES) RAM 32 Gb CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 128 Gb	CARACTERÍSTICAS MÍNIMAS: CPU (8 CORES) RAM 128 Gb CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 256 Gb	CARACTERÍSTICAS MÍNIMAS: CPU (8 CORES) RAM 32 Gb CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 128 Gb	CARACTERÍSTICAS MÍNIMAS: CPU (8 CORES) RAM 64 Gb CARACTERÍSTICAS RECOMENDADAS: CPU (8 CORES) RAM 128 Gb
STORAGE: 100 Gb OS / 100 Gb APP / 120 Gb DATA / 60Gb CACHE / 32Gb PG	STORAGE: 100 Gb OS / 100 Gb APP / 1 Tb DATA / 32Gb PG	STORAGE: 100 Gb OS / 100 Gb APP / 120 Gb DATA / 60Gb CACHE / 32Gb PG	STORAGE: 100 Gb OS / 100 Gb APP / 120 Gb DATA / 60Gb CACHE / 32Gb PG
COMPONENTES: <ul style="list-style-type: none"> • SAP BO IPS 4.2 • SAP BO DS 4.2 • TOMCAT 8 • SQL ANYWHERE 16 	COMPONENTES: <ul style="list-style-type: none"> • SAP IQ 16 	COMPONENTES: <ul style="list-style-type: none"> • SAP BO DESIGNSTUDIO 2BI • SAP BO LUMIRA 4 BI • TOMCAT 8 • SQL ANYWHERE 16 	COMPONENTES: <ul style="list-style-type: none"> • SAP BO BI 4.2 • SAP BO PREDICTIVE • TOMCAT 8 • SQL ANYWHERE 16

Figura 2.5: Servidores de produção [12]

SQL Anywhere. É um sistema de Gestão de Bases de Dados (SGBD), isto é, programas ou conjuntos de programas que possibilitam a criação e manipulação de bases de dados, que oferece a facilidade de ser incorporado em aplicações e que tem a capacidade de ser utilizado com pouca gestão de administração, permitindo auto ajustes, para melhorar o desempenho de forma autónoma. É também conhecido por estar disponível em várias plataformas como *Windows, Linux, Oracle Solaris*, entre outras [19].

SAP IQ. Sistema de base de dados relacional colunar otimizado para a análise de grandes volumes de dados não estruturados bem como dados históricos e operacionais de sistemas transacionais numa plataforma de dados unificados [16].

SAP Data Services Designer. Suporta o acesso, transformação e integração de dados, permitindo entregar informação com qualidade, relevante e a tempo, para funções críticas de negócio. Assim, tendo dados de confiança e disponíveis, as tomadas de decisão são consideravelmente melhores [16].

O Data Services Designer é composto por vários objetos, tais como [20]:

- *Project* - É o objeto no topo da hierarquia, tem como propósito organizar os objetos de nível mais baixo, como, *jobs, workflows* e *dataflows*.
- *Job* - É um executável do projeto e o local onde se encontram os objetos de baixo nível, ou seja, os *workflows* e *dataflows*.

- *Work flow* - Especifica a ordem pela qual o SAP Data Services Designer processa os *dataflows*.
- *Data flow* - Local onde as fontes de dados são transformadas em dados finais, processando os dados na ordem pela qual foram alinhados no *workflow*.

SAP IDT. Ferramenta de desenho de informação que permite extrair, definir e manipular metadados de fontes relacionais e de processamento analítico *online* (OLAP), para criar e implementar universos de SAP BusinessObjects. Um universo é uma coleção organizada de objetos metadados que permite ao utilizador analisar e reportar em dados corporativos numa linguagem não técnica. Estes objetos incluem dimensões, medidas, hierarquias, atributos, cálculos predefinidos, funções e *queries*. O utilizador é livre de analisar os dados e criar relatórios de forma simples e sem ajuda técnica [15].

O SAP IDT fornece vários recursos para extrair metadados e construir universos, tais como [15]:

- *Project* - Contém os recursos utilizados para construir um ou mais universos.
- *Connection* - Definição de como um universo acede a uma fonte de dados relacional ou OLAP.
- *Data foundation* - Esquema que define as tabelas e *joins* mais relevantes de uma ou mais bases de dados relacionais.
- *Business layer* - Coleção de objetos metadados que fornecem uma visão abstrata das entidades da base de dados relacional ou dos cubos OLAP, devendo ser compreendidos pelos utilizadores de negócio. Quando a *business layer* é concluída, é compilada com as conexões e a *data foundation*, sendo criado e publicado um universo.

SAP Web Intelligence (SAP WebI). Fornece aos utilizadores de negócio ferramentas de relatórios *ad hoc* flexíveis e intuitivas, bem como análises interativas [18]. O SAP WebI não necessita de ser instalado, pois pode ser acedido com um *browser* à escolha, dando liberdade ao utilizador de gerar os seus próprios relatórios ou consultar relatórios predefinidos.

SAP Lumira Discovery e Designer. O Lumira Discovery é usado por utilizadores que necessitam de flexibilidade para conectar e moldar os dados, criar histórias *ad-hoc* com visualizações de todos os tipos de dados [17]. O Lumira Designer é uma aplicação para analistas profissionais, que tipicamente pertencem aos departamentos de tecnologias e informação, para criarem aplicações de análise corporativa e relatórios.

SAP BI Launchpad. Portal *web* padrão para utilizadores da plataforma SAP BusinessObjects, fornecendo uma interface para partilha de relatórios, *dashboards* e outras

informações, sendo possível configurar permissões para controlar o acesso.

SAP CMC. A consola de gestão central é uma ferramenta baseada na *web* utilizada para tarefas administrativas do dia-a-dia, incluindo gestão de utilizadores, gestão de conteúdo e gestão de servidor. Qualquer utilizador com credenciais válidas para a plataforma de BI tem acesso ao CMC e um conjunto de preferências. Utilizadores que não sejam membros do grupo de administração não podem realizar tarefas de gestão, a não ser que seja garantido acesso para uma tarefa de gestão [14].

2.3 Sumário

Neste capítulo foram descritos conceitos teóricos sobre BI, seguindo-se a descrição do sistema de BI da Reitoria da ULisboa, que serviu de base para o trabalho documentado nos capítulos seguintes.

Capítulo 3

Análise e desenho

Este capítulo descreve o trabalho de análise e desenho realizado no âmbito deste projeto. Primeiramente é apresentada uma introdução com uma perspetiva geral do trabalho realizado, seguindo-se então secções sobre a análise do problema e o desenho da solução.

3.1 Introdução do trabalho

Nesta secção é apresentada uma introdução ao trabalho realizado, onde é descrito como o trabalho foi desenvolvido na organização, bem como alguns intervenientes que participaram no projeto. São descritas as ferramentas e o modo como estas foram utilizadas no âmbito do projeto e também é exposto um mapa com as diferentes fases do projeto.

3.1.1 Método de trabalho

O trabalho foi sendo concretizado com recurso a reuniões organizadas com os intervenientes necessários para as decisões do projeto, nomeadamente elementos do Gabinete de Estudos e Planeamento (GEP) da Reitoria da ULisboa e da equipa da Novabase (responsável pela solução de BI existente).

Como o ambiente de trabalho do Departamento de Informática na Reitoria da ULisboa é de *open space*, a troca de informações e impressões foi mais facilitada e direta. A comunicação com elementos não presentes no mesmo espaço físico foi efetuada através de correio eletrónico, ou então através de reuniões.

3.1.2 Tecnologias utilizadas

A maioria das tecnologias utilizadas para o desenvolvimento da solução proposta neste trabalho estão descritas com detalhe na Secção 2.2.9, sendo aqui apresentada uma síntese, que também inclui ferramentas específicas deste projeto, como o SPSS. Para o

desenvolvimento desta solução foi importante usar as mesmas tecnologias e arquitetura (Figura 2.3) da solução da Novabase, de forma a seguir algo já conhecido pela Reitoria da ULisboa.

SQL Anywhere, foi utilizada na componente de armazenamento dos dados e foi escolhida devido ao facto de vir automaticamente com a instalação do SAP DataServices e a quantidade de dados não era significativa para justificar uma base de dados alternativa.

SAP Data Services Designer foi a ferramenta utilizada para implementar a componente de processamento do projeto, tendo sido usado para os processos de integração dos dados, desde a extração dos dados até à sua transformação e carregamento.

SAP IDT foi utilizado para implementar a componente da camada semântica, tendo sido usado para criar um universo entendido pelos utilizadores e publicado para ser visualizado e analisado.

SAP Web Intelligence foi a ferramenta escolhida para a concretização dos relatórios, bem como para garantir a liberdade ao utilizador para criar os seus próprios relatórios de forma simples e intuitiva. Esta ferramenta implementou a camada de exploração e visualização do projeto.

SAP CMC foi utilizada na componente de administração, segurança e auditoria do projeto e serviu para criar utilizadores e definir níveis de segurança entre utilizadores e grupos, de modo a garantir ou restringir acessos a conteúdos necessários.

SPSS Statistics é um *software* usado pelo GEP para fazer análises estatísticas manualmente. No contexto deste projeto foi utilizado para consulta de dados, entender regras para a transformação dos dados, bem como para fazer comparações de análises estatísticas, mas não faz parte da solução desenvolvida.

3.1.3 Mapa do trabalho realizado

O projeto passou por várias fases de trabalho diferentes, e estas são apresentadas no mapa na Figura 3.1 e são descritas individualmente de forma breve.

1 - Análise de relatórios e necessidades. Consistiu em absorver o máximo de informação possível sobre este projeto e a solução de BI existente na Reitoria. Para tal, foram analisados vários documentos e relatórios e realizadas reuniões com o GEP para entender quais as suas necessidades, requisitos e preocupações.

2 - Modelo dimensional. Foi desenhada uma proposta de modelo de dados para

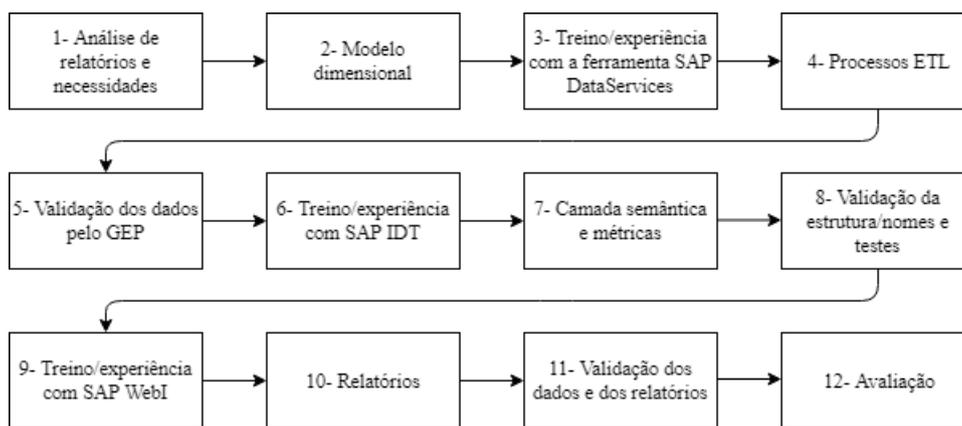


Figura 3.1: Mapa do trabalho realizado

suportar os requisitos e necessidades do cliente.

Após os passos 1 e 2 da Figura 3.1, foram realizadas as instalações e configurações de todas as ferramentas fundamentais para a construção da solução, dando origem à necessidade de manuais de instalação por parte da Reitoria, que podem ser consultados nos Apêndices A e B.

3 - Treino/experiência com a ferramenta SAP DataServices. Foi necessário ganhar conhecimentos sobre a primeira ferramenta a utilizar, neste caso o SAP Data Services Designer para realizar os processos de ETL.

4 - Processos ETL. Consistiu em aplicar os conhecimentos adquiridos e realizar os processos ETL usando o SAP Data Services Designer, começando pela recolha dos dados das diferentes fontes de dados para o sistema, seguida da aplicação de diferentes regras criadas para transformar os dados de acordo com o pretendido pelo cliente. Após a transformação dos dados, estes foram colocados numa tabela de base de dados, prontos para o passo seguinte.

5 - Validação dos dados pelo GEP. Como a solução foi para um cliente, pretendeu-se assegurar a sintonia entre o que estava a ser desenvolvido e o pretendido. Assim, depois de os dados estarem transformados, estes foram objeto de um processo de validação pelo cliente.

6 - Treino/experiência com SAP IDT. Neste passo foi necessário ganhar, conhecimentos sobre a ferramenta SAP IDT, que foi aplicada para tratar da camada semântica do projeto, bem como para criar as métricas, indicadores e variáveis auxiliares.

7 - Camada semântica e métricas. Consistiu na transformação dos campos existentes

na tabela de base de dados resultante do passo 4 para conceitos de negócio compreensíveis aos utilizadores. Foi nesta fase que foram criadas as métricas, indicadores e variáveis auxiliares necessários para mais tarde serem cruzadas com informação útil para análise. Os campos foram organizados em pastas assemelhando-se a um tradicional modelo dimensional, cuja organização foi pensada de forma a facilitar o acesso à informação e para que o seu uso pelos utilizadores fosse o mais ágil possível.

8 - Validação da estrutura/nomes e testes. Depois do passo 7, o GEP testou e validou, em várias iterações, a estrutura do modelo e os nomes atribuídos para que estivessem o mais adequados possível.

9 - Treino/experiência com SAP WebI. Implicou aprender como utilizar o SAP WebI para gerar relatórios de forma eficiente.

10 - Relatórios. Após a aprendizagem da ferramenta SAP WebI foi iniciado o processo de construção de relatórios estáticos para serem diretamente usados nos relatórios de empregabilidade produzidos pelo GEP.

11 - Validação dos dados e dos relatórios. Uma vez criados os relatórios, estes passaram por um processo de validação em que o GEP analisou e verificou se a estrutura dos mesmos seria a mais indicada, bem como o seu conteúdo.

12 - Avaliação. O último passo consistiu em avaliar a qualidade dos dados, realizando vários testes, nomeadamente comparações entre os dados obtidos manualmente e através da solução desenvolvida, aplicação da solução com dados de uma nova edição do inquérito à empregabilidade e diferença de tempos de execução entre os dois métodos.

Percorridas estas fases, procedeu-se à escrita do presente documento.

3.2 Análise do problema

Nesta secção são apresentados os requisitos, expectativas e preocupações das partes interessadas, um estudo do fluxo de trabalho do GEP e descrição das fontes de dados utilizadas, também uma análise de dados de empregabilidade, que já tinham sido processadas pela Reitoria da ULisboa. As últimas secções abordam indicadores e relatórios considerados essenciais para o projeto.

3.2.1 Análise de requisitos

É de extrema relevância que a solução desenvolvida responda às necessidades e requisitos do cliente. Para isso, foi importante compreender as informações necessárias, preocupações que pudessem existir por parte dos utilizadores, bem como as suas expectativas, tais como:

Uniformização dos dados. O GEP necessita que a plataforma de apoio à decisão esteja disponível e tenha os processos e conceitos uniformizados, facilitando a obtenção de informação útil e relevante.

Acesso à informação em tempo útil. Para os decisores, é importante o acesso a dados atualizados num curto espaço de tempo, sendo necessário o acompanhamento da realidade atual e a tomada de decisão de forma consciente e ajustada às necessidades da ULisboa, quer na gestão do dia-a-dia, quer na gestão corporativa da instituição.

Maior disponibilidade para análises estatísticas. Uma das preocupações da ULisboa, especialmente dos colaboradores responsáveis pela análise de dados e estatísticas, é o tempo gasto no processo de recolha de informação, tratamento e uniformização de dados das diferentes escolas, tendo em vista os vários indicadores analisados. Trata-se de tempo de organização e processamento de dados que poderia ser remetido para a análise e cruzamento estatístico de informação com maior detalhe.

Acompanhamento da evolução da solução. Existe um interesse em acompanhar o desenvolvimento desta solução, pois apresenta uma mais-valia para o sistema em caso de manutenção, bem como para quem o utilizar, devido à automatização de processos, agilidade e auxílio nas tarefas do dia-a-dia.

3.2.2 Fluxo de trabalho do GEP

A concretização deste projeto teve como objetivos automatizar, melhorar e facilitar o trabalho do GEP sobre o tema da empregabilidade. À data deste projeto, o trabalho de preparação da base de dados era realizado manualmente, chegando a demorar dois meses a ser concluído. O GEP seguia uma metodologia específica para facilitar o trabalho manual, sendo necessário estudar e entender o fluxo de trabalho do GEP para os dados de empregabilidade, o qual é apresentado na Figura 3.2.

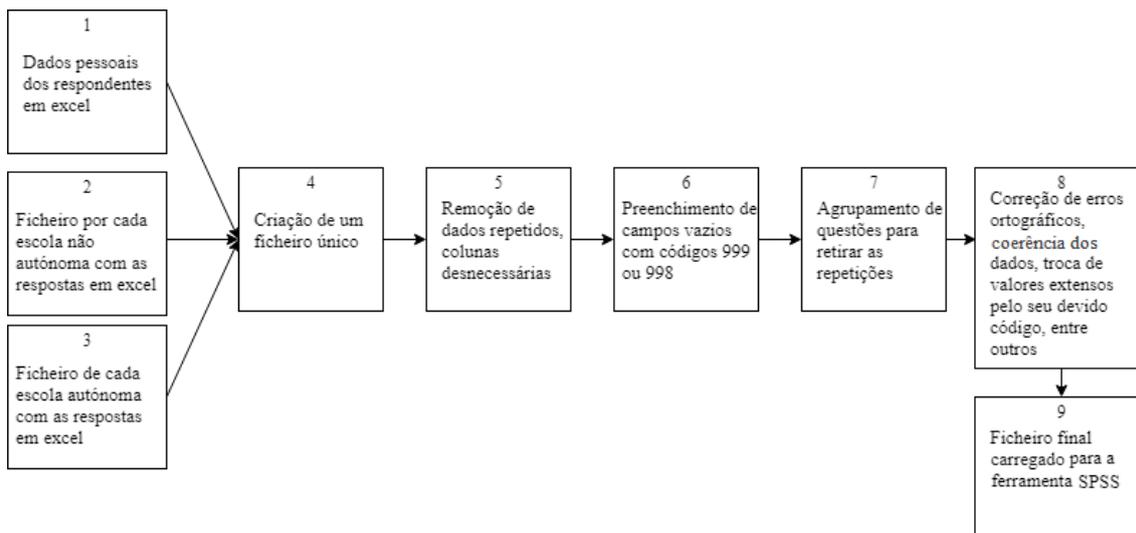


Figura 3.2: Fluxo de trabalho do GEP

1. Ficheiro Excel com informação pessoal de cada diplomado de cursos de Licenciatura, Mestrado Integrado e Mestrado de 2.º Ciclo num determinado ano letivo, que ao ser transposto para a plataforma de inquéritos LimeSurvey [6] gera um *token* específico e pessoal.
2. Ficheiro Excel com os dados brutos obtidos de todos os respondentes ao inquérito lançado em *LimeSurvey*, das várias escolas, exceto as escolas autónomas.
3. As quatro escolas autónomas (FC, ISA, ISEG e IST) preenchem e devolvem um ficheiro Excel fornecido pela Reitoria, contendo as mesmas colunas dos inquéritos da Reitoria referidos no ponto 2, garantindo coerência.
4. Os dados dos diferentes ficheiros são unidos para criar um ficheiro único. Todos os processos avançam mesmo que as escolas autónomas entreguem os seus dados mais tarde, os quais são integrados posteriormente.
5. O primeiro tratamento de dados que o GEP realiza é a remoção dos duplicados, ou seja, dados que decorrem de múltiplos acessos do mesmo *token* do LimeSurvey antes da submissão final do inquérito preenchido que permite que o mesmo seja interrompido. Trata-se de situações que ocorrem quando o respondente sai do inquérito e volta mais tarde para o continuar, reiniciando por vezes o seu processo de resposta. Nesta situação, opta-se pela entrada que contempla um conjunto de dados mais completo, no pressuposto que o respondente dedicou maior tempo e atenção a responder.
6. Os campos em branco são preenchidos com códigos definidos pela Reitoria, como “999”, caso o inquirido não tenha respondido a uma pergunta, e “998”, caso as

perguntas não sejam aplicáveis à situação específica do inquirido. Esta codificação é necessária, para que os valores inicialmente em branco possam ser excluídos do tratamento estatístico posterior de forma diferenciada.

7. O ficheiro obtido pelo LimeSurvey apresenta todas as perguntas dos diferentes perfis de empregabilidade (ainda que possam avaliar aspetos similares e transversais aos perfis), perfis esses que são definidos a partir da resposta a uma questão-filtro, que direciona o respondente para um determinado conjunto de questões. Logo, é necessário colapsar/agrupar as respostas de diplomados de perfis diferentes mas que concorrem para as mesmas perguntas ou indicadores. Passa a estar presente no ficheiro Excel o conjunto de perguntas em comum entre os diferentes perfis de situação de emprego e as perguntas específicas de um certo perfil.
8. Depois de o ficheiro Excel estar organizado segundo critérios sistemáticos definidos pelo GEP, são realizadas correções de erros ortográficos, conversão de valores extensos para códigos definidos pela Reitoria e correção de erros que não tenham sido detetados em iterações anteriores.
9. A versão final do ficheiro Excel é preparada para ser carregada no SPSS, onde são criadas variáveis auxiliares, que correspondem a informação resultante do cruzamento de respostas no inquérito, de modo a agilizar as análises e satisfazer as necessidades dos utilizadores. A tradução dos códigos criados no ponto anterior para texto extenso já está previamente codificada no SPSS. Após o cruzamento de variáveis e extração de tabelas com resultados estatísticos, são gerados os gráficos a colocar no relatório oficial de empregabilidade da ULisboa.

3.2.3 Fontes de dados utilizadas

Os dados necessários para o desenvolvimento deste projeto foram carregados para uma *staging area* para mais tarde serem tratados de acordo com os requisitos do cliente.

Neste projeto todas as fontes de dados utilizadas foram ficheiros Excel e parte dos dados provieram diretamente da Reitoria da ULisboa, sendo estes dados do inquérito à empregabilidade de 2017 no *LimeSurvey*.

Existem quatro escolas, a Faculdade de Ciências, o Instituto Superior Técnico, o Instituto Superior de Agronomia e o Instituto Superior de Economia e Gestão, também conhecidas como escolas “autónomas”, que optam por realizar os inquéritos de forma independente, mas que contêm pelo menos as mesmas perguntas que os inquéritos da Reitoria, para manter a coerência. As questões específicas de cada escola autónoma não são analisadas pelo GEP. Cada escola autónoma prepara e envia um ficheiro Excel com os dados requisitados pelo GEP.

3.2.4 Análise de dados

A informação necessária para a realização deste projeto proveio de inquéritos realizados a diplomados sobre a sua situação profissional atual. Estes inquéritos foram criados através de reunião da Reitoria com as escolas autónomas, consensualizando-se o modelo a adotar, e a informação foi trabalhada pela Reitoria e algumas escolas. A estrutura do inquérito e a análise detalhada está presente no relatório do inquérito à empregabilidade dos diplomados da ULisboa 2014/2015 [6].

Os inquiridos são diplomados que acabaram os seus ciclos de estudos há pelo menos 18 meses e têm grau de licenciatura ou de mestrado. O inquérito é constituído por 8 grupos no total, sendo o primeiro representado como G0, servindo para caracterizar o inquirido, o segundo grupo é o G1 e define a situação face ao emprego do inquirido consistindo numa questão-filtro que dirige posteriormente o respondente para um determinado grupo com questões específicas, ajustadas à caracterização do seu perfil profissional (G2,G3,G4,G5,G6 ou G7, como mostrado na Figura 3.3).

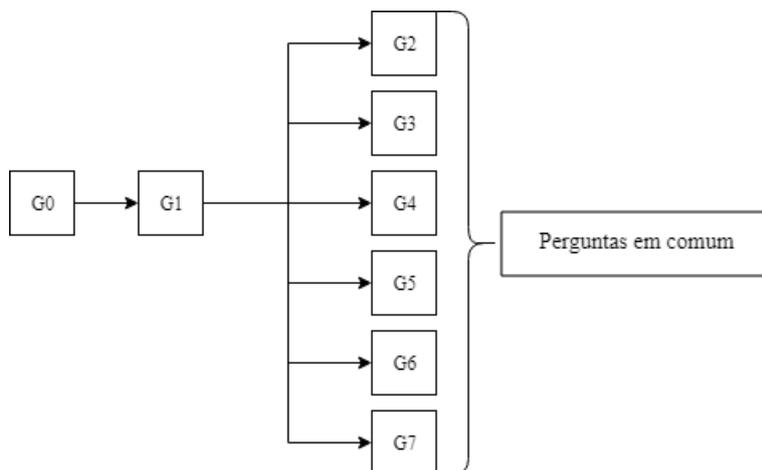


Figura 3.3: Estrutura geral do inquérito à empregabilidade

O inquérito segue uma ordem, criando seqüências de questões e dependências entre grupos. Como o diagrama da Figura 3.3 ilustra, só é possível avançar para o grupo G1 se o G0 for concluído, e o mesmo ocorre do G1 para os restantes grupos. O G1 é fundamental neste inquérito pois tem uma pergunta-filtro designada G1P1 (Grupo 1 pergunta 1) que encaminha para apenas um dos grupos seguintes, permitindo ao inquirido identificar qual a situação face ao emprego. Esta pergunta é obrigatória e só é possível escolher uma das seguintes opções:

- G2 - Trabalhador por conta de outrem;
- G3 - Trabalhador por conta própria sem funcionários a cargo;
- G4 - Trabalhador por conta própria com funcionários a cargo;
- G5 - Estagiário;

- G6 - Bolseiro;
- G7 - Sem atividade profissional remunerada (SAPR).

Como é apresentado no diagrama da Figura 3.3, existem perguntas comuns entre os diferentes grupos, aspeto relevante pois uma das transformações realizada aos dados consiste em agrupar estas perguntas, em vez de repeti-las na base de dados.

Análise do inquérito

Para melhor compreender os dados utilizados para este projeto, o GEP disponibilizou um ficheiro SPSS correspondente à 3ª edição do inquérito à empregabilidade da ULisboa, ou seja, o inquérito lançado aos diplomados de licenciatura e de mestrado que terminaram o seu curso no ano letivo 2014/2015. Sobre estes dados foi realizada uma análise para conhecer as variáveis utilizadas, bem como o tipo de resultados esperados. Esta análise também foi relevante para descobrir possíveis erros ou falhas, como erros ortográficos ou respostas incoerentes possivelmente pela clareza da pergunta.

O estudo do documento fornecido permitiu compreender que cada variável utilizada correspondia maioritariamente a uma questão do inquérito, sendo depois acrescentadas variáveis auxiliares que cruzam resultados de outras variáveis, para facilitar as análises.

O inquérito é partilhado pelas várias escolas da ULisboa e os resultados obtidos pelas escolas não autónomas são centralizados na Reitoria, ao passo que as outras quatro escolas autónomas (FC, ISA, ISEG e IST) lançam os seus inquéritos e depois reencaminham os seus dados para a Reitoria.

Apresentam-se em baixo de forma muito sucinta, parte dos principais resultados da análise dos dados da 3ª edição do inquérito:

- Total de 3478 respondentes;
- A maioria tem nacionalidade portuguesa;
- Grande parte concluiu o grau de licenciatura;
- A classificação média de final de curso é 14,5 valores;
- A idade média à data de obtenção do diploma é 25 anos;
- A maioria identifica-se como trabalhador por conta de outrem;
- A localização do emprego é maioritariamente em Portugal, no distrito de Lisboa;
- Dos respondentes que trabalham no estrangeiro, estes localizam-se maioritariamente em países europeus;
- A maioria das entidades empregadoras são do tipo privada com fins lucrativos;
- A média dos vencimentos brutos mensais é 1143 euros;
- A média dos complementos remuneratórios mensais é 334 euros;
- Dos que se encontram a trabalhar, grande parte dos respondentes indica ter emprego na área do curso em que se diplomou;
- Cerca de um terço obteve o seu primeiro emprego antes de terminar o curso;
- No caso dos que prosseguem estudos, a maioria permanece na mesma escola.

A análise detalhada pode ser encontrada no apêndice C, análise esta que pode apresentar diferenças em relação à versão oficial devido a tomadas de decisão discricionárias que foram humanamente determinadas.

3.2.5 Indicadores de empregabilidade

Estes indicadores são definidos para medir o desempenho de uma organização e utilizam métricas para esse efeito[1]. Conseguem oferecer uma visão mais exata de modo a avaliar a situação de um negócio. Na Tabela 3.1 são descritos os indicadores prioritários de empregabilidade definidos no âmbito do projeto, bem como os seus cálculos.

Tabela 3.1: Indicadores de empregabilidade

Métrica/Indicador	Descrição	Cálculo
Nº de diplomados inquiridos	Número de diplomados inquiridos pelos serviços de acompanhamento ao estudante após a conclusão dos estudos.	Contagem de diplomados inquiridos
Idade média	Idade média dos diplomados até à data de diplomação	Média da idade à data de diplomação
Nº de diplomados em prosseguimento de estudos	Número de diplomados que após a conclusão do seu curso prosseguem estudos noutro curso	Contagem de diplomados que após a conclusão do seu curso prosseguem estudos noutro curso
Nº de diplomados desempregados	Número de diplomados que após conclusão do seu curso estão à procura de emprego	Contagem de diplomados que após a conclusão do seu curso estão à procura de emprego
Nº de diplomados com trabalho	Número de diplomados que após a conclusão do seu curso se encontram a desempenhar uma atividade laboral	Contagem de diplomados empregados
Remuneração média mensal	Remuneração média mensal dos diplomados	Média da remuneração mensal
Taxa de empregabilidade	Rácio entre o número de diplomados empregados e a soma do número de diplomados empregados e o número de diplomados desempregados	$\text{Número de diplomados empregados} / (\sum \text{número de diplomados empregados} + \sum \text{número de diplomados desempregados})$

Tabela 3.1: Indicadores de empregabilidade (continuação)

Métrica/Indicador	Descrição	Cálculo
Nº de diplomados por média etária	Nº de diplomados por idade média à data de diplomação	Contagem de diplomados por idade média à data de diplomação

3.2.6 Relatórios prioritários

Para este projeto foi definida com o GEP uma seleção de relatórios prioritários para apresentar neste trabalho. Assim sendo, foram produzidos vários relatórios estáticos, seguindo a estrutura do relatório do inquérito de empregabilidade [6].

No SAP WebI um documento criado corresponde a uma categoria no relatório do inquérito à empregabilidade e cada uma tem vários relatórios associados. No ambiente de trabalho de projeto, os relatórios foram organizados nas seguintes categorias:

1. Caracterização dos diplomados
 - Diplomados por grau;
 - Diplomados por grau e escola;
 - Diplomados por sexo;
 - Diplomados por sexo e escola;
 - Diplomados por sexo e grau;
 - Média etária por grau.
2. Situação face ao emprego
 - Situação face ao emprego dos diplomados;
 - Situação face ao emprego por grau.
3. Caracterização dos diplomados SAPR
 - Diplomados SAPR, empregados após o curso;
 - Diplomados SAPR, empregados após o curso, por grau;
 - Motivos para não procurar emprego.

Os relatórios prioritários foram selecionados para mostrar como são desenvolvidos com a ferramenta SAP WebI, de modo a mais tarde ser seguida a mesma lógica para a criação de todos os relatórios predefinidos necessários para o negócio da organização.

3.3 Desenho da solução

Esta secção descreve o desenho inicial e final do modelo de dados e concluindo com apresentação e descrição da arquitetura da solução.

Tabela 3.2: Descrição das dimensões do modelo de dados inicial

Dimensão	Descrição
Diplomado	Características do diplomado
Tempo	Tempo, necessário para análise históricas
Local	Local de residência e de trabalho
Curso	Curso no qual o respondente se graduou
Prossegue estudos	Prosseguimento de estudos do respondente, incluindo a área e o local
Emprego	Características do local de emprego do respondente

3.3.1 Modelo de dados inicial

Após a análise dos relatórios e dos dados disponibilizados, foi pensado e desenhado um primeiro modelo dimensional para responder às necessidades do GEP, contendo sete dimensões, sendo cada uma delas descritas na Tabela 3.2. Os detalhes desse modelo podem ser encontrados no Apêndice D.

3.3.2 Modelo de dados final

Devido à complexidade do primeiro modelo de dados e também porque a dimensão diplomado iria conter grande parte dos atributos, a ideia não foi prosseguida.

Após a aprendizagem de conceitos da ferramenta SAP Information Design Tool, foi percebida a versatilidade e facilidade da ferramenta para criar e manter uma organização muito semelhante a um tradicional modelo dimensional de dados a partir de uma única tabela com todos os atributos.

Nesta linha, a Tabela 3.3 mostra a tabela final guardada na base dados, contendo todos os 95 atributos, indicando-se o grupo (G) e a pergunta (p) a que pertence no inquérito ou vazio caso, os atributos sejam obtidos pelo GEP através de sistemas académicos, dependências entre questões e exemplos de respostas.

A reduzida quantidade de dados foi outro fator relevante para a decisão de manter toda a informação numa só tabela, pois nestas circunstâncias um modelo dimensional oferece vantagens pouco significativas.

Tabela 3.3: Tabela final de base de dados com todos os atributos, indicando o grupo e pergunta, as dependências e exemplos de valores

Do inquérito	Atributo	Dependências	Exemplo
	ID_Estudante		18
	ID_SA		18
	Nome		Ana Espinheira
	Email		ana@xyz.pt

Tabela 3.3: Tabela final de base de dados com todos os atributos, indicando o grupo e pergunta, as dependências e exemplos de valores (continuação)

Do inquérito	Atributo	Dependências	Exemplo
	Idade		22
G0p1	Data_nascimento		10/10/1996
	Nacionalidade		Portugal
	Escola		FC
	Curso_cod		6095
	Curso_Designação		Biologia
	Grau		Licenciatura
	CITE_F_1		4-Ciências, matemática e informática
	CITE_F_2		42-Ciências da vida
	CITE_F_3		421-Biologia e bioquímica
	Áreas_ULisboa		Ciências da Vida
	Clas_Final		16
G0p2	Sexo		Feminino
G0p3	Residência_Port_Outro		Portugal
G0p4	Residência_Dist_Port	G0p3	Lisboa
G0p5a	Residência_Conc_Lisboa	G0p4	Sintra
G0p5b	Residência_Conc_Santarém	G0p4	Almeirim
G0p5c	Residência_Conc_Setúbal	G0p4	Moita
G0p6	Residência_Eur_PLOP_Out	G0p3	Europeu
G0p7	Residência_Eur	G0p6	Dinamarca
G0p8	Residência_PLOP	G0p6	Cabo Verde
G0p9	Residência_Out_País	G0p6	Coreia do Sul
G1p1	G1situação		Trabalhador por conta de outrem
G2p1	Emp_Portugal_Outro		Portugal
G2p2	Emp_Dist_Port	G2p1	Lisboa
G2p3a	Emp_Conc_Lis	G2p2	Lisboa
G2p3b	Emp_Conc_Sant	G2p2	Abrantes
G2p3c	Emp_Conc_Setu	G2p2	Alcochete
G2p4	Emp_Eur_PLOP_Out	G2p1	de Língua Oficial Portuguesa
G2p5	Emp_Eur	G2p4	Alemanha
G2p6	Emp_PLOP	G2p4	Brasil
G2p7	Emp_Out_País	G2p4	Japão
G2p8	Proc_Trab_Antes	G2p1	Sim
G2p91	Raz_1_Permaneci	G2p8	Não
G2p92	Raz_2_Experiencia	G2p8	Sim
G2p93	Raz_3_Origem	G2p8	Não
G2p94	Raz_4_Mercado	G2p8	Sim
G2p95	Raz_5_Conjuntura	G2p8	Não
G2p96	Raz_6_Pessoal	G2p8	Não
G2p97	Raz_Outra	G2p8	Sim
G2p10	Raz_Outra_Qual	G2p9	Melhor oferta no estrangeiro

Tabela 3.3: Tabela final de base de dados com todos os atributos, indicando o grupo e pergunta, as dependências e exemplos de valores (continuação)

Do inquérito	Atributo	Dependências	Exemplo
G2p11	Empregador		Privada com fins lucrativos
G2p12	Dimensão		50 a 249 pessoas
G2p13	Vínculo		Efetivo
G2p14	Vínculo_Outro	G2p13	Comissão de serviço
G2p15	Horário		Tempo inteiro
G2p16	Vencimento		1400
G2p17	Complem_SN		Sim
G2p18	Complem_Valor	G2p17	160
G2p19	Emp_Área_SN		Sim
G2p20	Atual_Emp_Antes_Curso_SN		Não
G2p211	Melhor_1_Progredi	G2p20	Não
G2p212	Melhor_2_Compert	G2p20	Não
G2p213	Melhor_3_Habilit	G2p20	Não
G2p214	Melhor_4_Autonomia	G2p20	Não
G2p215	Melhor_5_Funções	G2p20	Não
G2p216	Melhor_6_Reconhecimento	G2p20	Não
G2p217	Melhor_Outra	G2p20	Sim
G2p22	Melhor_Outra_Qual	G2p21	Outra melhoria
G2p23	Emp_1ºemp_após_cuso_sn	G2p20	Sim
G2p24	Tempo_Atual_Emp	G2p23	Antes de terminar o curso
G2p25	Tempo_Primeiro_Emp	G2p23	Até 1 mês após terminar o curso
G2p26	Pross_Est_SN		Sim
G2p27	Nível_Form	G2p26	Mestrado
G2p28	Nível_Form_Outro	G2p27	Especialização
G2p29	Mesma_Fac	G2p26	Não
G2p30	IE_Pub_Priv	G2p29	Pública
G2p30a	Pub_Univ_Politec_Milit	G2p30	Universitário
G2p30a1	Pub_Univ_IE	G2p30a	Universidade de Lisboa
G2p30a1a	Pub_Univ_IE_ULis	G2p30a1	FC
G2p30a2	Pub_Politec_IE	G2p30a	Instituto Politécnico de Leiria
G2p30a3	Pub_Milit_IE	G2p30a	Academia da Força Aérea
G2p30b	Priv_Univ_Politec	G2p30	Politécnico
G2p30b1	Priv_Univ_IE	G2p30b	Instituto de Arte, Design e Empresa
G2p30b2	Priv_Politec_IE	G2p30b	Outro
G2p31	IE_Pub_Priv_Outra	G2p30a1, 30a2, 30a3, 30b1, 30b2	Instituto CRIAP
G2p32	Areas_Estudo	G2p26	Informática
G2p33	Areas_Estudo_Qual	G2p32	Academia Java

Tabela 3.3: Tabela final de base de dados com todos os atributos, indicando o grupo e pergunta, as dependências e exemplos de valores (continuação)

Do inquérito	Atributo	Dependências	Exemplo
G4p12	Faturacao_Clientes		Um dos clientes representa mais de 80% do volume de faturação
G5p13	Estágio_Dur		Entre 6 meses e 1 ano
G6p13	Bolsa_Tipo		Bolsa de gestão de ciência e tecnologia (BGCT)
G6p14	Bolsa_Qual	G6p13	Bolsa da Ação Social
G6p17	Bolsa_Dur		1
G7p1	Emp_após_curso_SN		Não
G7p3	Ativ_remun_SN	G7p1	Não
G7p4	Procura_Emp_SN		Sim
G7p5	Motiv_Nao_Proc	G7p4	Estou a prosseguir estudos
G7p6	Motiv_Nao_Proc_Outro	G7p5	Problemas de saúde
G7p7	Pross_Est_SAPR		Não prosseguir mais estudos
G7p15	Deixar_Estudo_SN	G7p4, p7	Sim
	DT_DATA_PROC		2017

Assim sendo, depois de criar uma conexão entre a tabela única na base de dados e a ferramenta SAP IDT, foi desenhado o um modelo dimensional mostrado na Figura 3.4, que recorre a uma estrutura de pastas para organizar toda a informação proveniente da tabela única.

Em geral, cada pasta pode ser vista como sendo uma dimensão, contendo os respetivos atributos. Adicionalmente, também foram criadas as métricas, indicadores e variáveis auxiliares requisitadas pelo cliente.

Este modelo final foi validado pelo GEP, tendo havido a preocupação de que este fosse intuitivo para os utilizadores identificarem todos os objetos de forma ágil e com clareza.

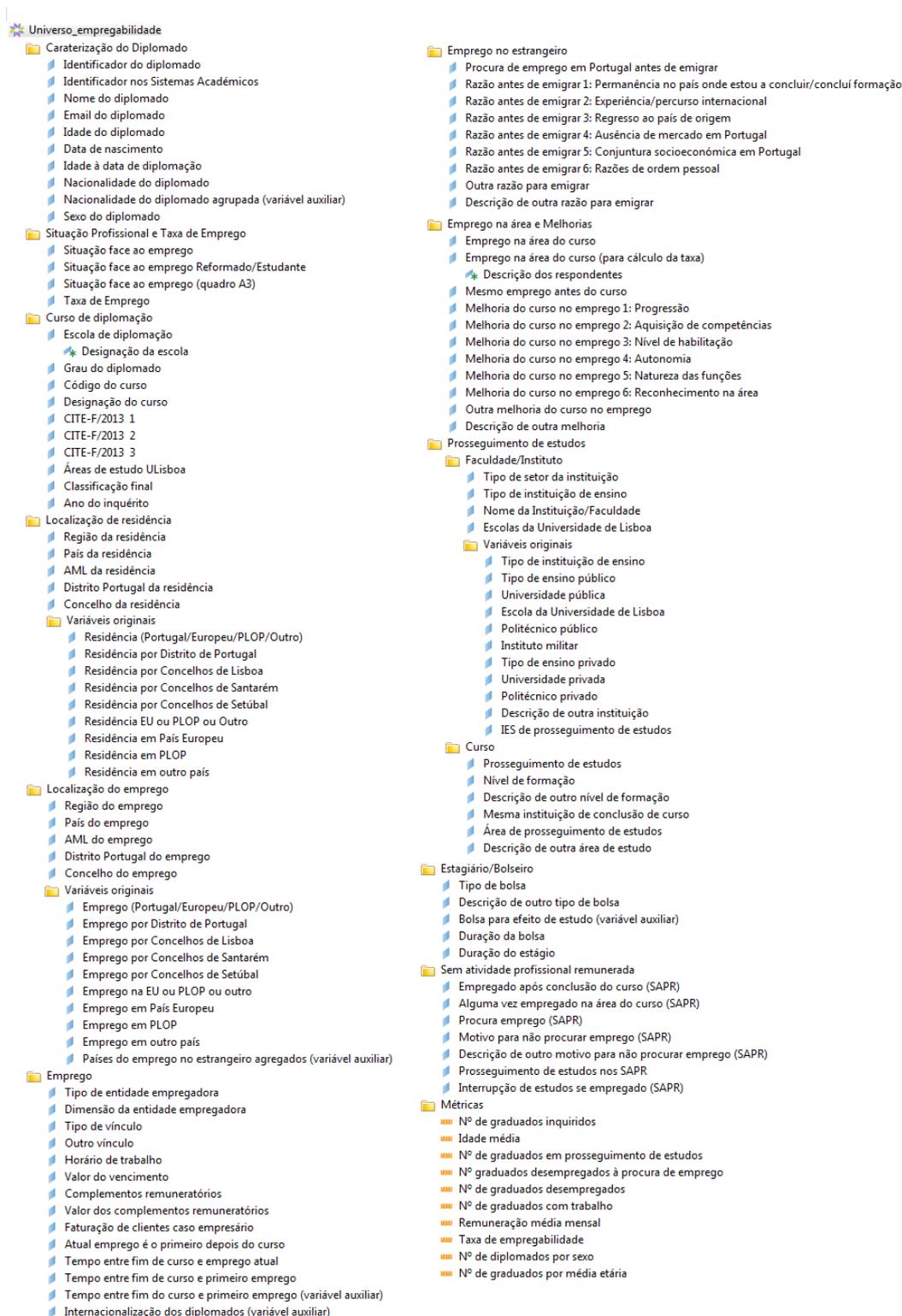


Figura 3.4: Estrutura do modelo final na ferramenta SAP IDT

3.3.3 Arquitetura da solução

A solução desenvolvida foi concebida tendo em consideração a solução já existente na Reitoria de ULisboa, seguindo uma estrutura semelhante, e utilizando as mesmas ferramentas. A Figura 3.5 apresenta a arquitetura da solução de BI desenvolvida.

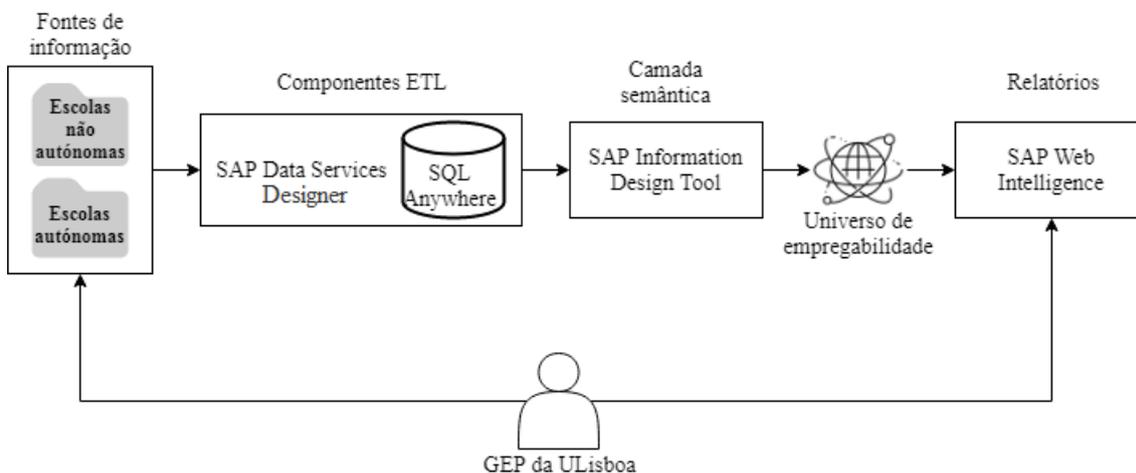


Figura 3.5: Arquitetura da solução de BI desenvolvida

O utilizador do GEP é responsável pela colocação das fontes de dados (ficheiros CSV) nas pastas respetivas seguindo uma nomenclatura estipulada para cada um dos ficheiros.

Com os ficheiros nas pastas devidas, estes devem passar por vários processos definidos no SAP Data Services Designer responsáveis pelas componentes ETL, onde, ao longo de cada processo, os dados devem ir sendo armazenados em diferentes tabelas definidas na base de dados (SQL Anywhere). O último processo deve definir uma tabela final com dados históricos alimentada com novos dados processados e consolidados.

O SAP Information Design Tool (SAP IDT) é usado para definir a estrutura do modelo de dados, transformação de conceitos técnicos em conceitos de negócio e definição de variáveis auxiliares, cálculo de métricas e indicadores.

O universo gerado pelo SAP IDT deve depois ser publicado para a aplicação SAP Web Intelligence (SAP WebI) onde podem ser gerados os relatórios e as análises prioritárias para a organização. Os utilizadores do GEP podem aceder à aplicação SAP WebI para consultar os relatórios gerados ou criar as suas próprias análises.

3.4 Sumário

Neste capítulo foi introduzido o trabalho, análise do problema e desenho e arquitetura da solução. O desenvolvimento da solução descrito no capítulo seguinte foi baseado nesta análise e desenho.

Capítulo 4

Implementação e avaliação

Este capítulo apresenta os processos envolvidos no desenvolvimento da solução, concluindo com a avaliação da mesma.

4.1 Implementação da solução

As subsecções seguintes descrevem o trabalho desenvolvido tendo como recurso as ferramentas SAP, isto é, a elaboração da extração, transformação e carregamento dos dados de empregabilidade, seguido da criação da camada semântica, responsável pela construção do modelo, métricas, indicadores e variáveis auxiliares, bem como a inteligibilidade dos dados. Por fim, apresenta-se a descrição dos relatórios construídos e a criação e configuração de utilizadores para acesso à solução de BI.

4.1.1 Desenvolvimento das componentes ETL

Os processos de ETL foram realizados na camada de armazenamento e processamento da arquitetura, ver Figura 2.3, com recurso à ferramenta SAP Data Services Designer. Os dados foram guardados e processados num local de trabalho temporário conhecido como *staging area*.

O desenvolvimento deste processo foi o mais intensivo e demorado, pois a qualidade dos dados após este processamento é um fator determinante da qualidade da tomada de decisão.

Extração

A primeira fase do processo ETL foi realizar a componente de extração. Para tal, cada ficheiro Excel fornecido pela Reitoria da ULisboa foi convertido para o formato CSV, de valores separados por vírgulas, e depois individualmente foram carregados para a *staging area*.

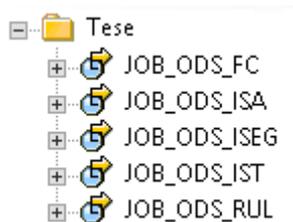


Figura 4.1: *Jobs* responsáveis pela componente de extração, compreendendo *jobs* individuais para cada escola autónoma (FC, ISA, ISEG e IST) e o mesmo *job* para as escolas não autónomas (RUL)

Como referido na Secção 2.2.9, um projeto no SAP Data Services Designer contém *jobs*, *workflows* e *dataflows*. Neste caso, foram criados cinco *jobs* para satisfazer o passo de extração. A Figura 4.1 indica os principais *jobs* responsáveis pelo carregamento dos dados do ficheiro CSV para uma base de dados.

Cada um destes *jobs* processa apenas um ficheiro CSV, ou seja, os dados fornecidos pela FC, ISA, ISEG e IST, bem como, os dados obtidos pela Reitoria para as escolas restantes. A estrutura destes *jobs* é idêntica, com alteração na fonte de dados. Dentro do *job* existe um *try catch* para apanhar erros caso ocorram, e vários *scripts* necessários para o funcionamento do *job*. Conforme apresentado na Figura 4.2 a presença de ciclos é relevante para possibilitar o tratamento de ficheiros com dados de empregabilidade de anos contínuos e preservar os dados de anos anteriores.

Os dados são processados em *dataflows*. Na Figura 4.2, o *dataflow* DF_ODS_FC trata do carregamento dos dados só da FC em ficheiro CSV para uma base de dados relacional SQL Anywhere. A Figura 4.3 mostra a janela de visualização da sequência de passos do carregamento dos dados, iniciando com a leitura do ficheiro a carregar seguido de uma *query transform* que trata da transformação dos dados para os nomes e tipos corretos e, por último o armazenamento do resultado na base de dados.

Transformação

Uma vez concluída a fase de extração, a segunda fase do processo de ETL consistiu em transformar e limpar os dados, isto é, corrigir, padronizar, tratar incoerências e transformar os dados de acordo com as regras do negócio. Seguindo o fluxo de trabalho na Secção 3.2.2, recriaram-se e automatizaram-se as transformações realizadas manualmente pelo GEP.

Após todos os dados estarem armazenados numa bases de dados, estes passaram por mais *jobs* (ver Figura 4.4) de forma a melhorar a sua organização e também para facilitar a futura manutenção da solução de BI.

O *job* JOB_UNIAO_ESCOLAS_AUTONOMAS trata de juntar os dados de cada escola autónoma numa só tabela. A Figura 4.5 mostra a estrutura do respetivo *dataflow*, onde para cada escola é aplicada uma *query transform* (QRY_FC, QRY_ISA, QRY_ISEG e QRY_IST)

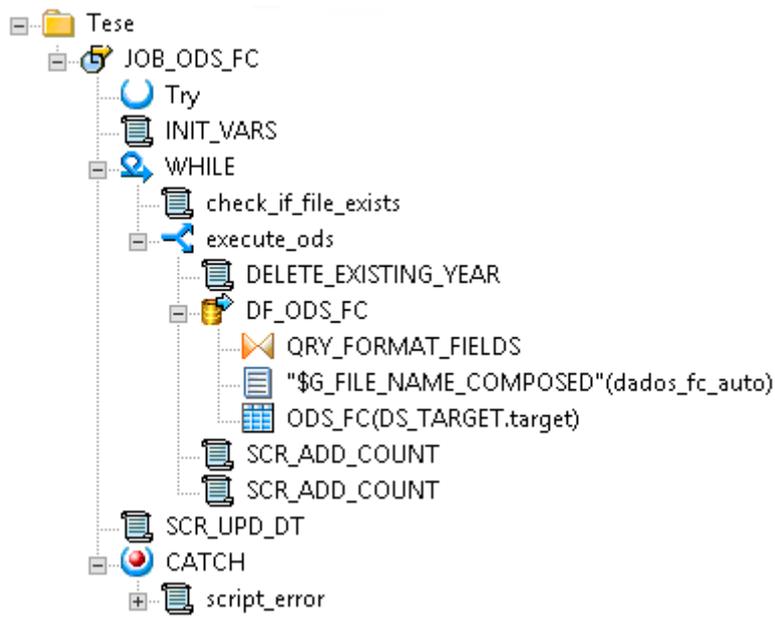


Figura 4.2: Objetos do *job* que extrai os dados da Faculdade de Ciências



Figura 4.3: *Dataflow* que extrai os dados da FC. O ficheiro \$G_FILE_NAME_COMPOSED-(DADOS_FC_AUTO) passa por uma transformação QRY_FORMAT_FIELDS e é carregado para uma tabela na base de dados ODS_FC(DS_TARGET.TARGET)

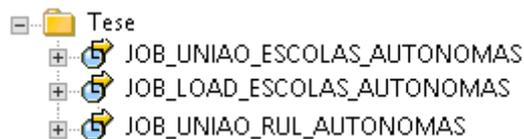


Figura 4.4: *Jobs* que agregam os dados das escolas autónomas com os das não autónomas

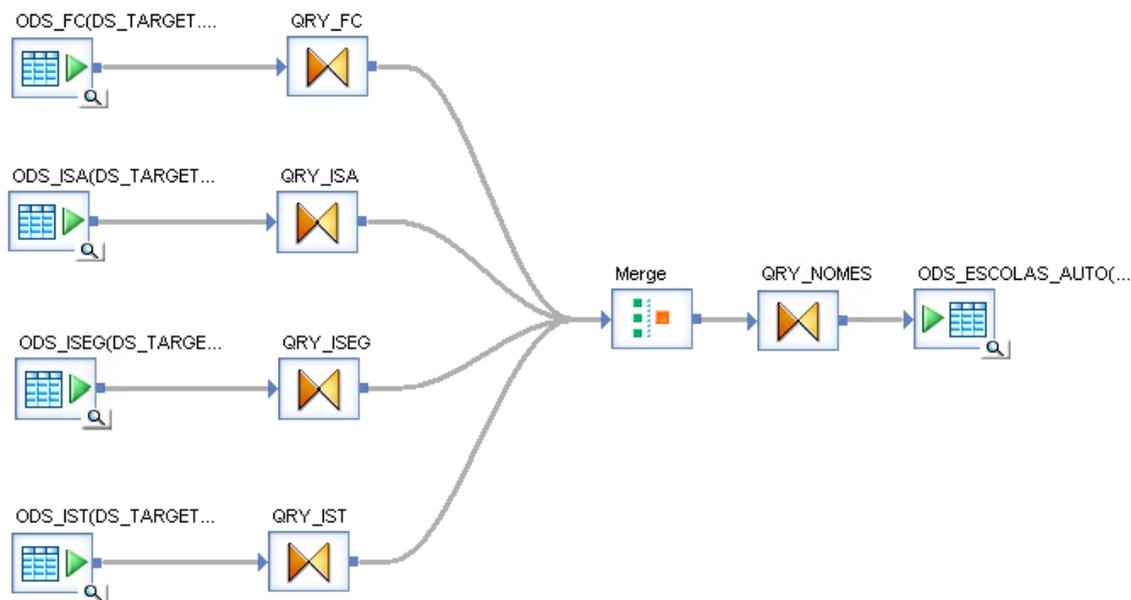


Figura 4.5: Objetos do *job* JOB_UNIAO_ESCOLAS_AUTONOMAS, responsável pela junção dos dados das escolas autónomas

que define os nomes e tipos dos atributos, que devem ser todos iguais para possibilitar o *merge* dos dados, executando-se seguidamente outra *query transform* (QRY_NOMES) que serve para melhorar os nomes e facilitar a sua interpretação. Após estas ações, os dados são carregados para uma nova tabela de base de dados (ODS_ESCOLAS_AUTO), contendo a união dos dados das quatro escolas autónomas.

Apesar de as escolas autónomas terem de seguir um formato específico para os dados nos ficheiros de reporte, indicado pela Reitoria, estes por vezes eram apresentados de forma diferente. Por exemplo, na indicação do sexo do respondente, algumas escolas usam “Feminino”, outras “F”, ou ainda “Mulher”, valores esses que tiveram de ser unificados. O *job* JOB_LOAD_ESCOLAS_AUTONOMAS, na Figura 4.6, trata exatamente deste tipo de traduções, onde os valores são trocados para um formato único, utilizado na Reitoria.

O código na Figura 4.7 é um fragmento da transformação QUERY_CODIFICAÇÃO visível na Figura 4.6, desta feita referente à expansão de siglas de graus académicos. No atributo “Grau” foi utilizada a função *search_replace* que, como o nome sugere, pesquisa uma palavra e substitui por outra. No caso apresentado, foram traduzidos os valores “L” e “L1” por Licenciatura, “MI” por Mestrado Integrado e “M2” por Mestrado (2ª ciclo).

Com os dados das escolas autónomas juntos e unificados, o passo seguinte consistiu em adicionar os dados obtidos diretamente da Reitoria da ULisboa, referentes às restantes escolas. Para tal foi criado o *job* JOB_UNIAO_RUL_AUTONOMAS, na Figura 4.8, onde um dos seus objetos é o *dataflow* DF_UNIAO_TODAS_ESCOLAS, que lê a tabela contendo os dados unidos das escolas autónomas (LOAD_ESCOLAS_AUTO) e a tabela com os dados das restantes escolas (ODS_RUL). Cada uma destas duas tabelas passa por uma *query transform* (QRY_LEA e QRY_RUL) que define os nomes e tipos dos atributos, que têm que

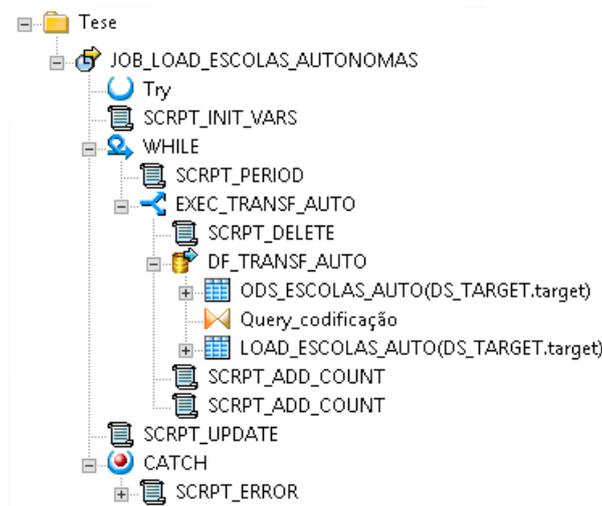


Figura 4.6: Objetos do *job* JOB_LOAD_ESCOLAS_AUTONOMAS, responsáveis pela uniformização dos valores em campos específicos, preenchidos pelas escolas autónomas

```
search_replace([], ['SR_FIELD'], ['no'], [], [ODS_ESCOLAS_AUTO.Grau], []) SET (
"internal_table" = '<?xml version="1.0" encoding="UTF-8"?>
<searchTable>
  <entry>
    <search>L</search>
    <replace>Licenciatura</replace>
  </entry>
  <entry>
    <search>MI</search>
    <replace>Mestrado Integrado</replace>
  </entry>
  <entry>
    <search>M2</search>
    <replace>Mestrado (2º ciclo)</replace>
  </entry>
  <entry>
    <search>L1</search>
    <replace>Licenciatura</replace>
  </entry>
</searchTable>',
"run_as_separate_process"='no')
```

Figura 4.7: Exemplo de tradução de valores de uma coluna, nomeadamente a expansão de siglas de graus académicos

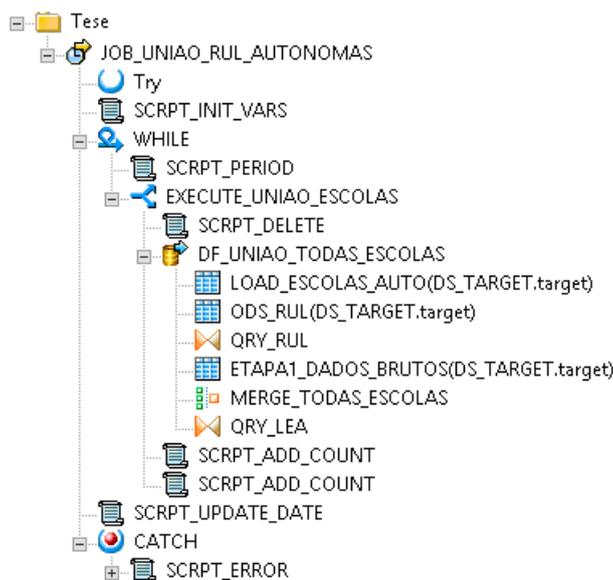


Figura 4.8: Job responsável pela união dos dados de todas as escolas da ULisboa

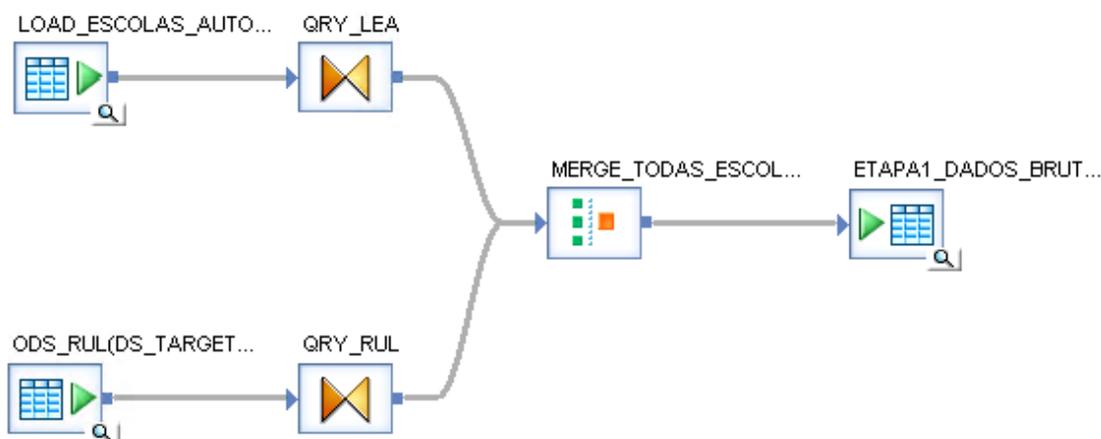


Figura 4.9: Dataflow responsável pela união dos dados de todas as escolas da ULisboa

ser exatamente iguais para ser possível executar o *merge* (MERGE_TODAS_ESCOLAS). Uma vez juntos os dados, estes são carregados para uma nova tabela de base de dados (ETAPA1_DADOS_BRUTOS), como mostrado no *dataflow* da Figura 4.9.

Com os dados organizados e armazenados numa tabela, foi inicializado o processo de replicar as transformações especializadas que o GEP segue no seu fluxo de trabalho. A Figura 4.10 ilustra a estrutura do *job* criado para esse efeito contendo cinco *dataflows* diferentes. Uma das necessidades do GEP era a capacidade de consulta de estados intermédios dos dados ao longo das transformações, tendo a solução passado por criar um ficheiro CSV no final de cada uma das cinco transformações num formato conhecido pelo utilizador.

O primeiro *dataflow*, DF_ETAPA2_AGREGACAO, é responsável pela consolidação de perguntas em comum entre os diferentes perfis de situação face ao emprego (ver Figura

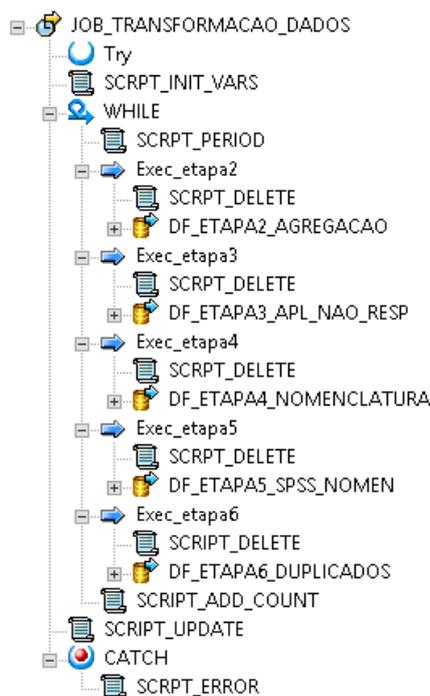


Figura 4.10: Objetos do *job* JOB_TRANSFORMACAO_DADOS, com cinco *dataflows*

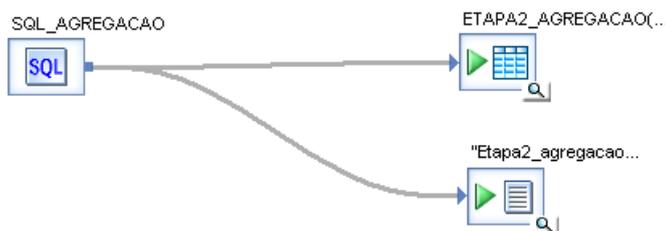


Figura 4.11: *Dataflow* DF_ETAPA2_AGREGACAO, responsável pelo agrupamento de perguntas em comum nos inquéritos. O código SQL_AGREGACAO executa a agregação e carrega o resultado para a tabela ETAPA2_AGREGACÃO, bem como cria um ficheiro ETAPA2_AGREGACAO_DADOS.CSV para consulta pelos utilizadores

3.3), diminuindo drasticamente o número de colunas de 563 para 95. Nesta transformação também são preenchidos todos os campos em branco com o código 998, para mais tarde serem tratados. A Figura 4.11 mostra a estrutura deste *dataflow* que consiste na execução de um *script* SQL cujo resultado é carregado para uma tabela na base de dados relacional e para um ficheiro CSV.

A implementação do *script* SQL, mostrado na Listagem 4.1 requereu uma análise exaustiva do inquérito de empregabilidade para entender quais as perguntas a consolidar, e atenção na escrita do *script* para não cometer erros que pudessem comprometer a qualidade dos resultados.

SELECT

'ID interno', 'ID do diplomado nos SA', Nome, Email, Idade ,
'ID Aluno', Nascimento, Nacionalidade, Estabelecimento,
'Codigo curso', Curso, Grau, CITE_F_1, CITE_F_2, CITE_F_3,
'Areas de estudo ULisboa', 'Media final', 'ID da resposta',

Coalesce(g0p1,'998') as g0p1,
Coalesce(g0p2,'998') as g0p2,
Coalesce(gop3,998) as g0p3,
Coalesce(g0p4,'998') as g0p4,
Coalesce(g0p5a,998) as g0p5a,
Coalesce(g0p5b,998) as g0p5b,
Coalesce(g0p5c,998) as g0p5c,
Coalesce(g0p6,998) as g0p6,
Coalesce(g0p7,998) as g0p7,
Coalesce(g0p8,998) as g0p8,
Coalesce(g0p9,'998') as g0p9,
g1p1,
Coalesce(g2p1,g3p1,g4p1,g5p1,g6p1,998) as g2p1,
Coalesce(g2p2,g3p2,g4p2,g5p2,g6p2,'998') as g2p2,
Coalesce(g2p3a,g3p3a,g4p3a,g5p3a,g6p3a,998) as g2p3a,
Coalesce(g2p3b,g3p3b,g4p3b,g5p3b,g6p3b,998) as g2p3b,
Coalesce(g2p3c,g3p3c,g4p3c,g5p3c,g6p3c,998) as g2p3c,
Coalesce(g2p4,g3p4,g4p4,g5p4,g6p4,998) as g2p4,
Coalesce(g2p5,g3p5,g4p5,g5p5,g6p5,998) as g2p5,
Coalesce(g2p6,g3p6,g4p6,g5p6,g6p6,998) as g2p6,
Coalesce(g2p7,g3p7,g4p7,g5p7,g6p7,'998') as g2p7,
Coalesce(g2p8,g3p8,g4p8,g5p8,g6p8,'998') as g2p8,
Coalesce(g2p91,g3p91,g4p91,g5p91,g6p91,'998') as g2p91,
Coalesce(g2p92,g3p92,g4p92,g5p92,g6p92,'998') as g2p92,
Coalesce(g2p93,g3p93,g4p93,g5p93,g6p93,'998') as g2p93,
Coalesce(g2p94,g3p94,g4p94,g5p94,g6p94,'998') as g2p94,
Coalesce(g2p95,g3p95,g4p95,g5p95,g6p95,'998') as g2p95,
Coalesce(g2p96,g3p96,g4p96,g5p96,g6p96,'998') as g2p96,
Coalesce(g2p97,g3p97,g4p97,g5p97,g6p97,'998') as g2p97,
Coalesce(g2p10,g3p10,g4p10,g5p10,g6p10,'998') as g2p10,
Coalesce(g2p11,g5p11,g6p11,998) as g2p11,
Coalesce(g2p12,g4p11,g5p12,g6p12,998) as g2p12,
Coalesce(g2p13, 998) as g2p13,
Coalesce(g2p14,'998') as g2p14,
Coalesce(g2p15,g3p11,g5p14,g6p18,998) as g2p15,
Coalesce(g2p16,g3p12,g4p13,g5p15,g6p19,998) as g2p16,
Coalesce(g2p17,g4p14,g5p16,'998') as g2p17,
Coalesce(g2p18,g4p15,g5p17,998) as g2p18,
Coalesce(g2p19,g3p13,g4p16,g5p18,g6p20,'998') as g2p19,
Coalesce(g2p20,g3p14,g4p17,g6p21,'998') as g2p20,

Coalesce(g2p211,g3p151,g4p181,g6p221,'998') as g2p211,

```

Coalesce(g2p212,g3p152,g4p182,g6p222,'998') as g2p212,
Coalesce(g2p213,g3p153,g4p183,g6p223,'998') as g2p213,
Coalesce(g2p214,g3p154,g4p184,g6p224,'998') as g2p214,
Coalesce(g2p215,g3p155,g4p185,g6p225,'998') as g2p215,
Coalesce(g2p216,g3p156,g4p186,g6p226,'998') as g2p216,
Coalesce(g2p217,g3p157,g4p187,g6p227,'998') as g2p217,

Coalesce(g2p22,g3p16,g4p19,g6p23,'998') as g2p22,
Coalesce(g2p23,g3p17,g4p20,g5p19,g6p24,'998') as g2p23,
Coalesce(g2p24,g3p18,g4p21,g5p20,g6p25,998) as g2p24,
Coalesce(g2p25,g3p19,g4p22,g5p21,g6p26,g7p2,998) as g2p25,
Coalesce(g2p26,g3p20,g4p23,g5p22,g6p27,'998') as g2p26,
Coalesce(g2p27,g3p21,g4p24,g5p23,g6p28,g7p8,998) as g2p27,
Coalesce(g2p28,g3p22,g4p25,g5p24,g6p29,g7p9,'998') as g2p28,
Coalesce(g2p29,g3p23,g4p26,g5p25,g6p30,g7p10,g6p15a,'998') as g2p29,

Coalesce(g2p30,g3p24,g4p27,g5p26,g6p31,g7p11,g6p15b,998) as g2p30,
Coalesce(g2p30a,g3p24a,g4p27a,g5p26a,g6p31a,g7p11a,998) as g2p30a,
Coalesce(g2p30a1,g3p24a1,g4p27a1,g5p26a1,g6p31a1,g7p11a1,g6p15c,998) as g2p30a1,
Coalesce(g2p30a1a,g3p24a1a,g4p27a1a,g5p26a1a,g6p31a1a,g7p11a1a,g6p15c1,998) as
    ↪ g2p30a1a,
Coalesce(g2p30a2,g3p24a2,g4p27a2,g5p26a2,g6p31a2,g7p11a2,998) as g2p30a2,
Coalesce(g2p30a3,g3p24a3,g4p27a3,g5p26a3,g6p31a3,g7p11a3,998) as g2p30a3,
Coalesce(G2p30b,g3p24b,g4p27b,g5p26b,g6p31b,g7p11b,998) as g2p30b,
Coalesce(g2p30b1,g3p24b1,g4p27b1,g5p26b1,g6p31b1,g7p11b1,g6p15d,998) as g2p30b1,
Coalesce(g2p30b2,g3p24b2,g4p27b2,g5p26b2,g6p31b2,g7p11b2,998) as g2p30b2,
Coalesce(g2p31,g3p25,g4p28,g5p27,g6p32,g7p12,g6p16,'998') as g2p31,
Coalesce(g2p32,g3p26,g4p29,g5p28,g6p33,g7p13,998) as g2p32,
Coalesce(g2p33,g3p27,g4p30,g5p29,g6p34,g7p14,'998') as g2p33,

Coalesce(g4p12,'998') as g4p12,
Coalesce(g5p13,998) as g5p13,
Coalesce(g6p13,998) as g6p13,
Coalesce(g6p14,'998') as g6p14,
Coalesce(g6p17,998) as g6p17,
Coalesce(g7p1,'998') as g7p1,
Coalesce(g7p3,'998') as g7p3,
Coalesce(g7p4,'998') as g7p4,
Coalesce(g7p5,998) as g7p5,
Coalesce(g7p6,'998') as g7p6,
Coalesce(g7p7,998) as g7p7,
Coalesce(g7p15,'998') as g7p15,
DT_DATA_PROC
FROM ETAPA1_DADOS_BRUTOS
WHERE ETAPA1_DADOS_BRUTOS.DT_DATA_PROC = {$G_PERIOD}

```

Listagem 4.1: *Script* SQL para consolidar perguntas em comum do inquérito de empregabilidade

No *dataflow* DF_ETAPA3_APL_NAO_RESP (ver Figura 4.10) a transformação consistiu em diferenciar os campos 998 (que correspondem a perguntas não aplicadas ao inquirido)

dos campos 999 (em que a pergunta era aplicável ao inquirido mas este não respondeu). Estes valores 998 e 999 são utilizados nesta parte da solução de BI porque era a nomenclatura que o GEP seguia para diferenciar o não aplicável do não respondeu.

O desenvolvimento desta transformação foi o mais moroso do projeto, pois foi necessário entender e aplicar várias regras tendo em conta como as perguntas do inquérito se encontravam condicionadas, ou com dependência a uma pergunta prévia, tendo sido realizadas várias sessões de esclarecimento com o GEP.

A Figura 4.12 mostra a estrutura do *dataflow* DF_ETAPA3_APL_NAO_RESP, tendo sido criadas várias *query transform* por motivos de organização, facilidade de manutenção e pelo facto de haver perguntas com dependências. O *dataflow* é executado de acordo com uma sequência, representada na ferramenta com a direção da seta. Atendendo a que várias perguntas do inquérito dependem do resultado da pergunta anterior, foi preciso criar várias *queries* que adotam os resultados de *queries* anteriores.

Este *dataflow* foi organizado tendo em conta a estrutura do inquérito, e os nomes atribuídos às transformações seguem a nomenclatura do GEP, consistindo no número do grupo e a pergunta do inquérito. Por exemplo, a transformação com o nome G0P7_8_9 contém as transformações relativas à pergunta 7, 8 e 9 do Grupo 0. Com esta organização, caso seja necessária alguma alteração, consegue-se com facilidade encontrar o local certo para efetuar a mesma.

A Figura 4.13 apresenta um exemplo de uma regra criada. A variável selecionada, G0p4, corresponde à pergunta 4 do Grupo 0 do inquérito e questiona sobre qual o distrito onde o inquirido reside, tratando-se de uma questão que depende da resposta a uma pergunta anterior, sobre se reside em Portugal ou noutro país. É essencial que a regra a construir especifique esta dependência. A partir da Figura 4.13 é possível observar que se a pergunta G0p3 for igual a 1 (Portugal) e se o campo da pergunta estiver preenchido com 998, então esta deve ser transformada em 999, pois a pergunta era aplicável ao inquirido mas este não a respondeu. Caso o inquirido não responda à pergunta G0p3 e se o campo G0p4 estiver preenchido com 998, este também deve ser transformado em 999, pois não respondeu a nenhuma das duas perguntas. Se nenhum destes casos ocorrer então, deverá manter a resposta original.

Os *dataflows* designados por DF_ETAPA4_NOMENCLATURA e DF_ETAPA5_SPSS_NOMEN tratam da nomenclatura dos dados. O primeiro altera nomes dos valores nos atributos, ao passo que o segundo foi solicitado pelo GEP, que gostaria de dispor de um ficheiro de dados com nomes de atributos iguais aos nomes dos atributos do ficheiro SPSS, utilizado no tratamento manual dos dados de empregabilidade. O *dataflow* DF_ETAPA5_SPSS_NOMEN tem uma estrutura semelhante à do DF_ETAPA4_NOMENCLATURA representada na Figura 4.14: a tabela inicial passa por uma

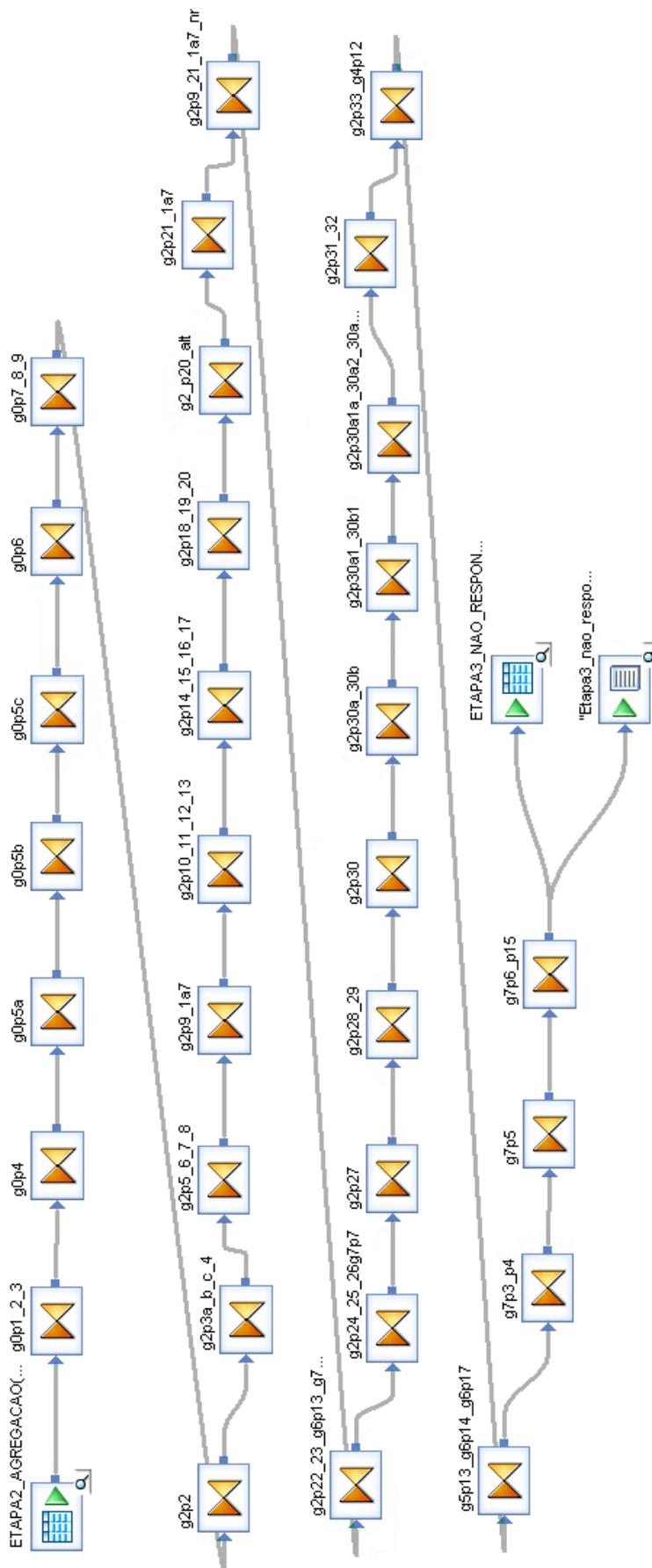


Figura 4.12: Dataflow DF_ETAPA3_APL_NAO_RESP, responsável por diferenciar os campos 998 e 999

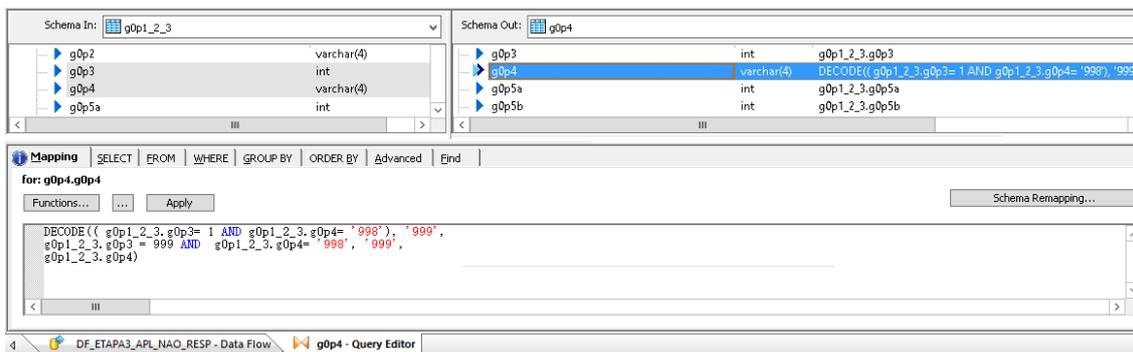


Figura 4.13: Regra aplicada aos dados provenientes do g0p1_2_3 para o g0p4 presentes no início da Figura 4.12

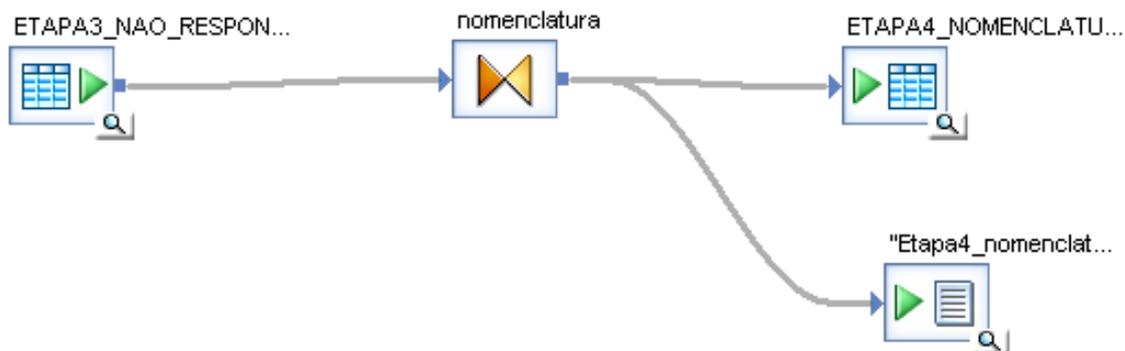


Figura 4.14: Dataflow DF_ETAPA4_NOMENCLATURA, responsável pela alteração dos valores dos atributos

query transform que mapeia os dados necessários e de seguida são armazenados numa nova tabela, bem como, num ficheiro CSV.

O último dataflow implementado neste job, DF_ETAPA6_DUPLICADOS, resolve os registos em duplicado, que ocorrem quando um respondente sai de um inquérito sem o submeter e volta a acedê-lo posteriormente. A solução passou por remover os registos que forneciam menos informação. De referir, ainda, que independentemente de se tratarem ou não de situações de inquiridos duplicados, são removidos todos aqueles que não respondam à questão-filtro, ou seja, que não indicam a sua situação profissional.

A Figura 4.15 mostra a estrutura da implementação do dataflow. Inicialmente são selecionados todos os dados exceto os que não têm valor na pergunta obrigatória (G1p1) do inquérito. Também são selecionados os dados agrupados por nome do respondente e por identificador da resposta atribuído pelas escolas, que por vezes é repetido, sendo escolhidos aqueles que têm o identificador mais elevado. Isto porque, quando os dados são duplicados, o identificador da resposta aumenta, fazendo sentido escolher o valor mais alto, pois corresponde ao inquérito mais completo em questão de respostas preenchidas. Depois de selecionados estes dois conjuntos de dados, é concretizado um inner join, isto é,

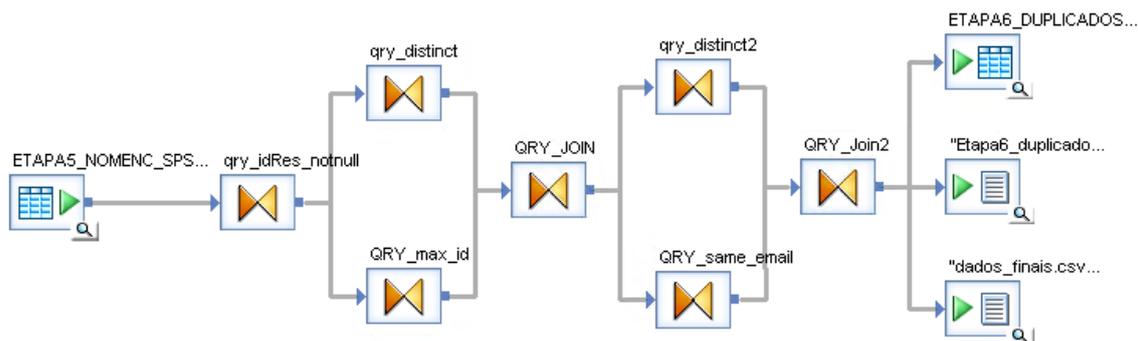


Figura 4.15: Dataflow DF_ETAPA6_DUPLICADOS, responsável pela remoção de registos duplicados

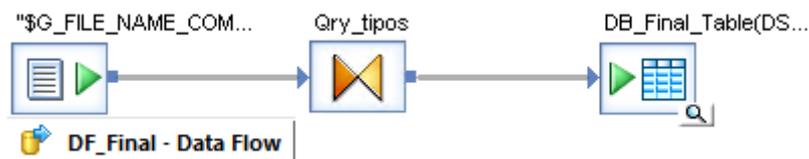


Figura 4.16: Dataflow DF_FINAL, responsável pelo carregamento do ficheiro CSV com os dados finais para uma DW DB_FINAL_TABLE

são cruzados os dados que contêm valores coincidentes nos dois conjuntos.

Ainda a propósito do *dataflow* na Figura 4.15, também foram verificados casos excepcionais em que existiam *emails* iguais com nomes ligeiramente diferentes devido à dificuldade de codificação de caracteres, tendo sido acrescentado um novo *inner join* que, em casos com *email* igual, apenas seleciona os que têm o identificador de resposta mais elevado.

Carregamento

Após a concretização das fases de extração e de transformação, a última fase do processo ETL passou pelo carregamento dos dados tratados para um *data warehouse*, isto é, persistir os dados numa base consolidada.

Para responder às necessidades do cliente, nomeadamente permitir o controlo sobre os dados em situações excepcionais, em que trabalho manual é requerido o último *job* (JOB_FINAL), é iniciado com o carregamento de um ficheiro CSV com as alterações manuais, se existirem. Só então os dados são armazenados numa tabela consolidada designada DB_FINAL_TABLE. A Figura 4.16 apresenta a estrutura do *dataflow* (DF_FINAL) responsável por este último *job*.

Uma vez os dados finalmente carregados na *data warehouse*, estes estão preparados e disponibilizados para serem utilizados. Assim sendo, o desenvolvimento da camada semântica pôde ser iniciado, sendo descrito na secção seguinte.

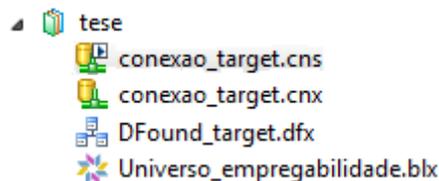


Figura 4.17: Estrutura do projeto na ferramenta SAP IDT para a criação do universo

4.1.2 Desenvolvimento da camada semântica

O desenvolvimento da camada semântica consistiu na transformação da tabela de dados e dos campos existentes para um conceito de negócio tendo em vista um melhor entendimento por parte dos utilizadores. Este pertence à camada semântica da arquitetura e teve como recurso a ferramenta SAP Information Design Tool (SAP IDT).

Esta ferramenta possibilita a criação de um modelo dimensional, com a definição de dimensões e de atributos inteligíveis. Adicionalmente também foram criadas as variáveis auxiliares requeridas pelo GEP, bem como as métricas e indicadores necessários para cruzar com a informação.

Para tal, começou-se por criar uma nova sessão na ferramenta para conectar ao servidor do projeto e, assim, ter acesso ao trabalho desenvolvido no processo ETL. Com a sessão aberta, foi necessário criar um novo projeto e abrir uma conexão relacional para extrair os dados da base de dados alocados no servidor. Foram gerados dois tipos de conexões relacionais: uma conexão local que cria um ficheiro .cnx (*local unsecured connection*) e que não pode ser publicado para um repositório, e uma conexão segura que provém da conexão local e cria um ficheiro .cns (*secured repository connection*), o qual permite publicar um universo para um repositório (ver Figura 4.17).

Uma vez a conexão concluída, foi necessário criar uma camada *data foundation* para importar a tabela da base de dados resultante do processo ETL. Com o acesso à tabela da base de dados estabelecido, o próximo passo consistiu em criar uma *business layer*.

Na *business layer*, o primeiro trabalho a ser desenvolvido foi a organização de todos os atributos presentes na tabela única importada da base de dados, tendo sido criadas dimensões e preenchidas com os atributos correspondentes. A Figura 4.18 mostra a estrutura final a que se chegou depois da validação com o GEP. Numa analogia, cada uma dessas pastas é uma dimensão e no interior de cada uma destas dimensões estão os seus atributos, exceto nas métricas (que são apenas métricas e indicadores). Esta estrutura, mais completa, pode ser revisitada na Figura 3.4, que contém todas as dimensões e os respetivos atributos, ou métricas e indicadores. Em paralelo com a criação do modelo, os nomes dos atributos também foram tornados mais inteligíveis consoante as decisões do GEP.

Com a estrutura e os nomes decididos e validados, o próximo passo consistiu em criar

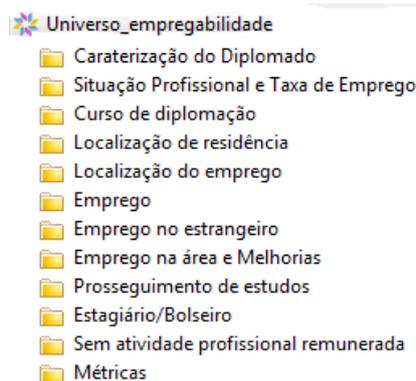


Figura 4.18: Estrutura do modelo dimensional criado na ferramenta SAP IDT

variáveis auxiliares pedidas pelo GEP, descritas na Tabela 4.1. Para criar cada variável auxiliar foi necessário adicionar um novo atributo na dimensão adequada e implementar uma expressão SQL de modo a obter o resultado pretendido.

Tabela 4.1: Variáveis auxiliares pedidas pelo GEP e desenvolvidas no SAP IDT

Nome	Descrição
Nacionalidade do diplomado agrupada	Agrupa as nacionalidades em Portugal, PLOP, Europa e Outra
Internacionalização dos diplomados	Diferencia o respondente em Português/estrangeiro que trabalha em Portugal/estrangeiro
Residência na AML	Diferencia os residentes que pertencem ou não à área metropolitana de Lisboa
Situação face ao emprego reformado/estudante	Distingue dentro dos sem atividade profissional remunerada, os que são estudantes e reformados
Situação face ao emprego	Distingue dentro dos sem atividade profissional remunerada, os que são estudantes à procura de emprego, estudantes que não procuram emprego e os que são de facto sem atividade profissional remunerada
Taxa de emprego	Classifica os respondentes como, CAPR (Com atividade profissional remunerada), SAPR (Sem atividade profissional remunerada), Estudante e Reformado
Emprego na AML	Distingue se os locais de emprego pertencem ou não à área metropolitana de Lisboa
Países do emprego no estrangeiro, agregados	Agrega os países estrangeiros em grupos definidos pelo GEP, como, Outro África, Ásia, Macau, entre outros
Emprego na área	Classifica os respondentes como, CAPR (Com atividade profissional remunerada), SAPR (Sem atividade profissional remunerada), CAPRAF (Com atividade profissional remunerada na área de formação) e ENPE (Estudante que não procura emprego)

Tabela 4.1: Variáveis auxiliares pedidas pelo GEP e desenvolvidas no SAP IDT (continuação)

Nome	Descrição
Tempo entre fim de curso e primeiro emprego	Categoriza todas as respostas sobre tempos em: Antes de terminar o curso, até 1 mês, até 3 meses, até 6 meses, até 1 ano, mais de 1 ano, sem 1º emprego, estudante sem 1º emprego
Instituições de ensino superior de prosseguimento de estudos	Para efeito de apresentação de informação em relatório, foram identificadas macro categorias de instituições de ensino superior, como, politécnico público/privado, outra universidade pública/privada, entre outros
Bolsa para efeito de estudo	Identifica consoante o tipo de bolsa se esta é para efeitos de estudo ou não

A Listagem 4.2 mostra a expressão SQL utilizada para criar a variável auxiliar “internacionalização dos diplomados”, onde o respondente que tem nacionalidade portuguesa e local de emprego em Portugal é considerado um “Português a trabalhar em Portugal”. A mesma lógica é seguida para caracterizar os restantes respondentes com a categoria correta.

```

case
when
@Select(Caraterização do Diplomado\Nacionalidade do diplomado agrupada $(variável auxiliar
    ↪ $)) = 'Portugal' AND @Select(Localização do emprego\Variáveis originais\Emprego $(
    ↪ Portugal/Europeu/PLOP/Outro$)) = 'Portugal'
then
'Português_a_trabalhar_em_Portugal'

when
@Select(Caraterização do Diplomado\Nacionalidade do diplomado agrupada $(variável auxiliar
    ↪ $)) = 'Portugal'
AND @Select(Localização do emprego\Variáveis originais\Emprego $(Portugal/Europeu/PLOP
    ↪ /Outro$)) = 'Outro_país'
then
'Português_que_foi_trabalhar_para_o_estrangeiro'

when
@Select(Caraterização do Diplomado\Nacionalidade do diplomado agrupada $(variável auxiliar
    ↪ $)) IN ('Europa', 'PLOP', 'Outra')
AND @Select(Localização do emprego\Variáveis originais\Emprego $(Portugal/Europeu/PLOP/
    ↪ Outro$)) = 'Portugal'
then
'Estrangeiro_que_ficou_a_trabalhar_em_Portugal'

when

```

```

@Select(Caraterização do Diplomado\Nacionalidade do diplomado agrupada §(variável auxiliar
  ↳ §)) IN ('Europa', 'PLOP', 'Outra')
AND @Select(Localização do emprego\Variáveis originais\Emprego §(Portugal/Europeu/PLOP/
  ↳ Outro§)) = 'Outro_país'
then
'Estrangeiro_que_não_ficou_a_trabalhar_em_Portugal'

when
@Select(Localização do emprego\Variáveis originais\Emprego §(Portugal/Europeu/PLOP/Outro
  ↳ §)) = 'na'
then
'na'

else
'nr'
end

```

Listagem 4.2: Implementação da variável auxiliar correspondente à internacionalização dos diplomados presente na dimensão Emprego

Regressando às métricas e indicadores, todos os que estão na Secção 3.2.5 foram implementados no SAP IDT. A métrica utilizada com maior frequência é referente à contagem dos diplomados inquiridos, cuja concretização em forma de expressão SQL é mostrada na Listagem 4.3.

```
isnull(count(distinct @Select(Caraterização do Diplomado\Identificador do diplomado)),0)
```

Listagem 4.3: Implementação da métrica de contagem dos diplomados inquiridos

A implementação de indicadores teve como objetivo oferecer uma visão mais detalhada de situações tipificadas de forma a melhorar a sua avaliação. O cálculo da taxa de empregabilidade foi um dos indicadores concretizados e indica o rácio entre o número de diplomados empregados e a soma do número de diplomados empregados e o números de diplomados desempregados (ver Listagem 4.4).

```
CAST(@Select(Métricas\Nº de graduados com trabalho)as DECIMAL) / CAST(@Select(Mé
  ↳ tricas\Nº de graduados inquiridos) as DECIMAL) * 100
```

Listagem 4.4: Implementação do indicador de taxa de empregabilidade

Com a transformação da tabela e dos campos existentes em conceitos de negócio para compreensão de todos utilizadores e com o desenvolvimento do modelo de dados e a implementação das variáveis auxiliares, métricas e indicadores o universo gerado (Universo_empregabilidade.blx) ficou concluído e preparado para ser publicado. Uma vez publicado, os dados ficaram disponíveis para os utilizadores.

4.1.3 Criação dos relatórios

O SAP Web Intelligence faz parte da plataforma de *BusinessObjects* e foi utilizado para gerar relatórios analíticos e *ad hoc* que respondessem aos requisitos de negócios do GEP. Mais precisamente para facilitar a escrita do relatório do inquérito à empregabilidade realizado todos anos.

Para utilizar o SAP WebI foi necessário aceder ao SAP BI Launch pad com um *browser*; uma vez autenticado, o procedimento para criar um relatório predefinido consistiu em abrir a aplicação Web Intelligence e criar um documento com o universo gerado no SAP IDT.

A Figura 4.19 mostra o painel inicial para a construção de relatórios, estando do lado esquerdo o *Universe outline* que contém todos os objetos do universo selecionado, incluindo as dimensões e respetivos atributos implementados no SAP IDT. O utilizador deve arrastar os objetos disponíveis no *Universe outline* para o *Result Objects* para serem adicionados ao relatório a gerar. O *Query Filters* permite criar filtros com os objetos do universo, de modo a limitar os resultados finais.

A organização do modelo dimensional foi relevante para este momento: o utilizador quando se deparar com este painel, deve de forma intuitiva e ágil encontrar os objetos necessários para criar uma análise.

Para a criação dos relatórios predefinidos deste projeto foram produzidos vários documentos, tendo por base a estrutura do relatório do inquérito de empregabilidade[6]. Tomando o primeiro documento da caracterização dos diplomados na Secção 3.2.6 como exemplo, foram inicialmente selecionados os atributos, métricas e indicadores necessários para desenhar os gráficos requeridos, incluindo uma interrogação de filtro correspondente ao ano de inquérito a analisar (ver Figura 4.19).

Com os objetos selecionados o próximo passo consistiu em executar a *query* e seleccionar o ano do inquérito pretendido. Para gerar os gráficos foram criadas tabelas arrastando os objetos para o relatório, as quais foram transformadas em gráficos.

A Figura 4.20 é um exemplo de um relatório predefinido, em que a estrutura é estática exceto os valores que variam consoante o ano de inquérito selecionado. Apenas o utilizador com permissão de administrador possui autorização para alterar estes relatórios, enquanto os restantes apenas podem visualizar e exportar, para evitar erros ou múltiplas alterações.

Se for necessária uma análise que não se encontra presente nos relatórios predefinidos, o utilizador necessita de criar o seu próprio, de forma *ad-hoc*, com a informação pretendida ou realizar um requerimento ao administrador com as alterações desejadas.

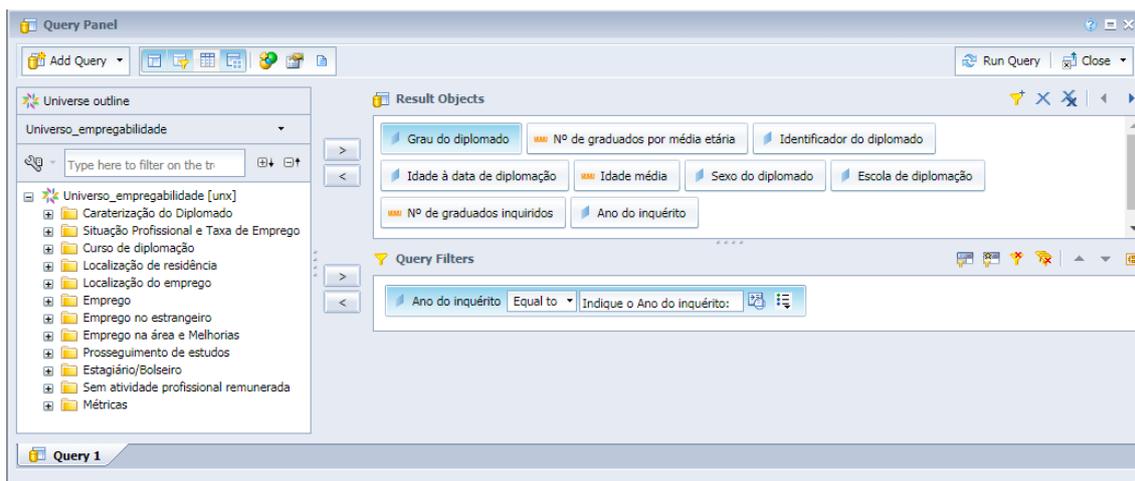


Figura 4.19: Painel de interrogação do WebI com o universo de empregabilidade: em cima estão os objetos selecionados para a criação de um relatório de caracterização dos diplomados; em baixo está a definição de um filtro correspondente ao ano de inquérito, limitando os resultados ao ano selecionado

4.1.4 Criação e configuração de utilizadores

Para este projeto foi necessário configurar contas de utilizadores para que estes tivessem acesso à plataforma SAP BI Launch pad e usufríssem da informação disponível na aplicação SAP WebI. Para isso foi utilizada a SAP CMC que é uma consola de gestão central para realizar tarefas administrativas, tendo como uma das suas funcionalidades a configuração de utilizadores e a atribuição de permissões para controlar acessos.

Para criar um utilizador de teste para o GEP, foi necessário aceder ao SAP CMC como administrador e criar um grupo para facilitar a atribuição de permissões semelhantes entre vários utilizadores. A atribuição de permissões ao grupo criado teve de ser aplicada a todos os níveis, ou seja, foi necessário permitir ao utilizador do grupo ter acesso à aplicação SAP WebI na sua área no SAP BI Launch pad, seguido de permissões sobre as conexões e universo criados no SAP IDT. Por fim, foi necessário selecionar as pastas que os utilizadores podem visualizar ou não.

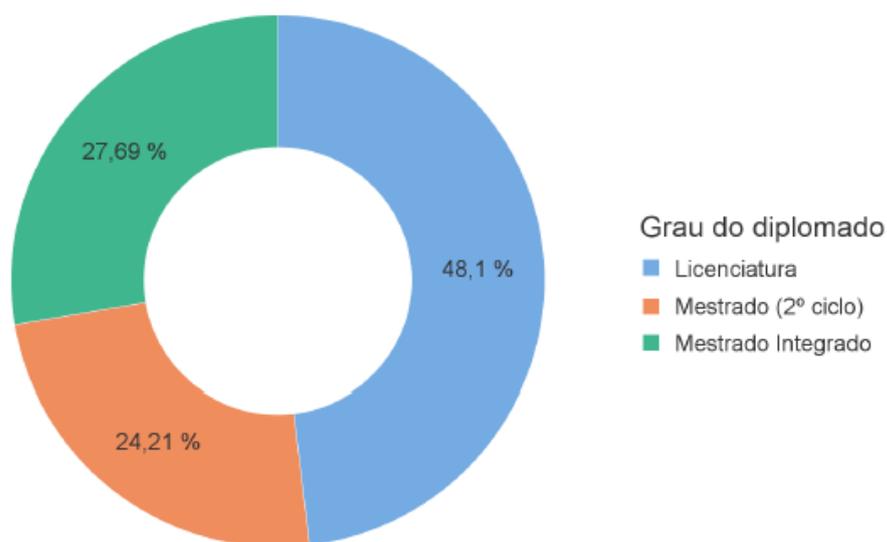
Com o grupo criado e com as permissões configuradas o último passo consistiu na criação de um utilizador e respetiva associação ao grupo. Deste modo, se for necessário adicionar mais utilizadores com as exatas permissões, basta apenas criar um novo utilizador e associar ao mesmo grupo.

Este utilizador foi partilhado com o GEP para os utilizadores terem acesso ao projeto, conhecerem a informação disponível, validarem os relatórios gerados e testarem a resolução de problemas.

Em simultâneo, foi criado outro grupo com permissões avançadas que permite um utilizador com mais responsabilidades executar todos os *jobs* implementados no SAP Data Services Designer. Para tal, foi seguida a mesma lógica anterior usando o SAP CMC, onde

Diplomados por Grau

Figura 1.1 - Diplomados da ULisboa, por Grau (n=3.478)



Grau do diplomado	Nº de graduados inquiridos	Percentagem
Licenciatura	1.673	48,10%
Mestrado (2º ciclo)	842	24,21%
Mestrado Integrado	963	27,69%
Total	3.478	100,00%

Figura 4.20: Relatório predefinido de diplomados por grau, do inquérito de 2017

o novo grupo criado poderá aceder ao SAP Data Services Management Console e executar os *jobs* seguindo uma ordem específica a ser documentada num manual, para guiar as equipas a executar o processo da solução corretamente.

4.2 Avaliação

Nesta secção são apresentados três diferentes avaliações: primeiramente é realizada uma comparação de resultados entre os dados manuais e da solução, seguido-se da aplicação da solução a um novo ano de inquérito, e concluindo com uma comparação de tempos de execução entre o trabalho realizado manualmente e usando a solução.

4.2.1 Comparação entre os dados manuais e da solução desenvolvida

Neste ponto foram concretizadas comparações entre os resultados obtidos pela solução desenvolvida e os resultados obtidos pelo GEP. Estas comparações permitiram analisar a qualidade da solução e se esta foi ao encontro dos requisitos definidos inicialmente.

Assim sendo, a Tabela 4.2 indica os valores totais absolutos entre os resultados obtidos pela solução e pelo GEP, bem como, a diferença entre eles para facilitar a comparação entre os dois métodos. Estes resultados foram referentes ao inquérito à empregabilidade da ULisboa em 2017 dirigido aos diplomados do ano letivo 2014/15, onde o total de respostas válidas corresponderam a uma amostra de 3478 diplomados.

Aquando do desenvolvimento dos gráficos e tabelas com os resultados obtidos pela solução e comparados estes com os resultados obtidos pelo GEP, sempre que foram observadas discrepâncias estas auxiliaram na identificação de possíveis defeitos no projeto, melhorando a qualidade e fiabilidade da solução.

Tabela 4.2: Diferença absoluta entre os resultados obtidos pela solução e os dados obtidos pelo processo do GEP no inquérito de 2017

Nome do gráfico	Valor da solução de BI	Valor do GEP	Diferença
Diplomados por grau	3478	3478	0
Diplomados por grau e escola	3478	3478	0
Diplomados por sexo	3453	3453	0
Diplomados por sexo e escola	3453	3453	0
Diplomados por sexo e grau	3453	3453	0
Média etária por grau	3448	3448	0
Situação face ao emprego dos diplomados	3478	3478	0

Tabela 4.2: Diferença absoluta entre os resultados obtidos pela solução e os dados obtidos pelo processo do GEP no inquérito de 2017 (Continuação)

Nome do gráfico	Valor da solução de BI	Valor do GEP	Diferença
Situação face ao emprego por grau	3478	3478	0
Diplomados SAPR, empregados após curso	588	587	1
Diplomados SAPR, empregados após curso, por grau	588	587	1
Motivos para não procurar emprego	535	539	4
Taxa de emprego, por grau	2978	2974	4
Taxa de emprego, por áreas de estudo	2978	2974	4
Tipo de vínculo, por conta de outrem, por grau	1722	1703	19
Tempo de espera para obter emprego, por grau	2829	2795	34
Prosseguiram estudos após curso, por grau	3286	3288	2
Emprego na área de formação	2878	2878	0
Emprego na área de formação, por grau	2878	2878	0
Dimensão da entidade empregadora	2122	2105	17
Nacionalidade dos estrangeiros por área geográfica	76	76	0
Total		51103	86

Com a divisão da soma dos valores absolutos obtidos pelo processo manual do GEP com a soma dos valores obtidos pela diferença absoluta da Tabela 4.2 obteve-se um resultado de 0,2% de casos em que se verificaram diferenças entre os dois métodos, sendo uma percentagem relativamente baixa.

Uma razão para as diferenças de valores é a transformação responsável pela remoção dos duplicados que resolve a situação numa perspetiva global, sem analisar em detalhe casos particulares, enquanto o GEP consegue identificar estes casos especiais e resolvê-los. A solução para esta situação passou por dar acesso ao ficheiro com dados, permitindo ao

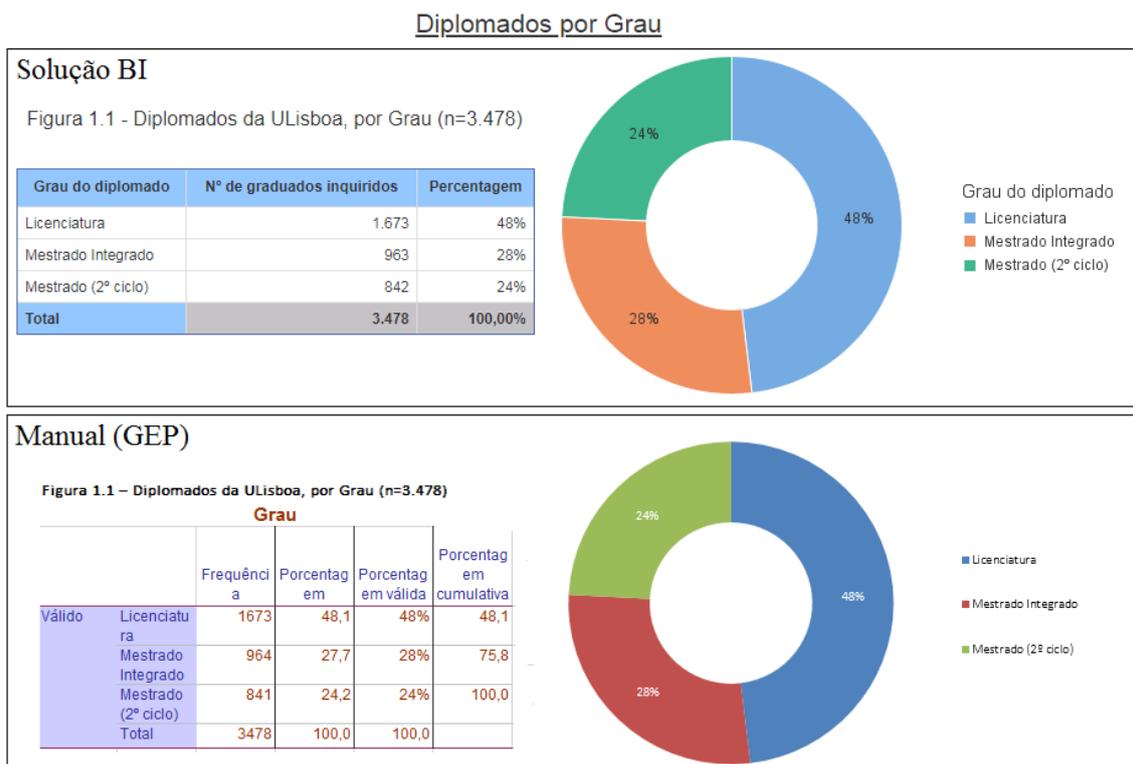


Figura 4.21: Diplomados por grau do inquérito de 2017 com o método manual e com a solução

GEP analisar as situações em duplicado abstraindo-se das restantes transformações.

A Figura 4.21 apresenta com maior detalhe um exemplo de um resultado obtido manualmente e com a solução. O valor total da amostra e as percentagens foram iguais em ambos os métodos, apenas diferenciando uma unidade entre o grau de Mestrado Integrado e Mestrado (2º ciclo).

A Figura 4.22 apresenta a comparação detalhada dos resultados relativamente ao tipo de vínculo de trabalhadores por conta de outrem por grau. Neste caso a diferença de valores deveu-se ao facto de grande parte dos inquiridos identificarem outro tipo de vínculo em resposta aberta como “Recibos verdes” mas tinham declarado pertencer ao perfil de “Trabalhador por conta de outrem”, quando realmente pertenciam ao perfil “Trabalhador por conta própria sem funcionários a cargo”. O GEP identifica estes casos específicos e faz correções manuais nos dados, mas isso não é feito na solução automática.

Em suma, a solução desenvolvida indica melhorias, permite automatizar o processamento da grande maioria dos dados de empregabilidade podendo por vezes ser necessário um toque humano para tratar de casos particulares que requerem a análise de respostas abertas dos inquiridos, sendo difícil identificar incoerências quando o conteúdo das respostas é vasto e sem limitações.

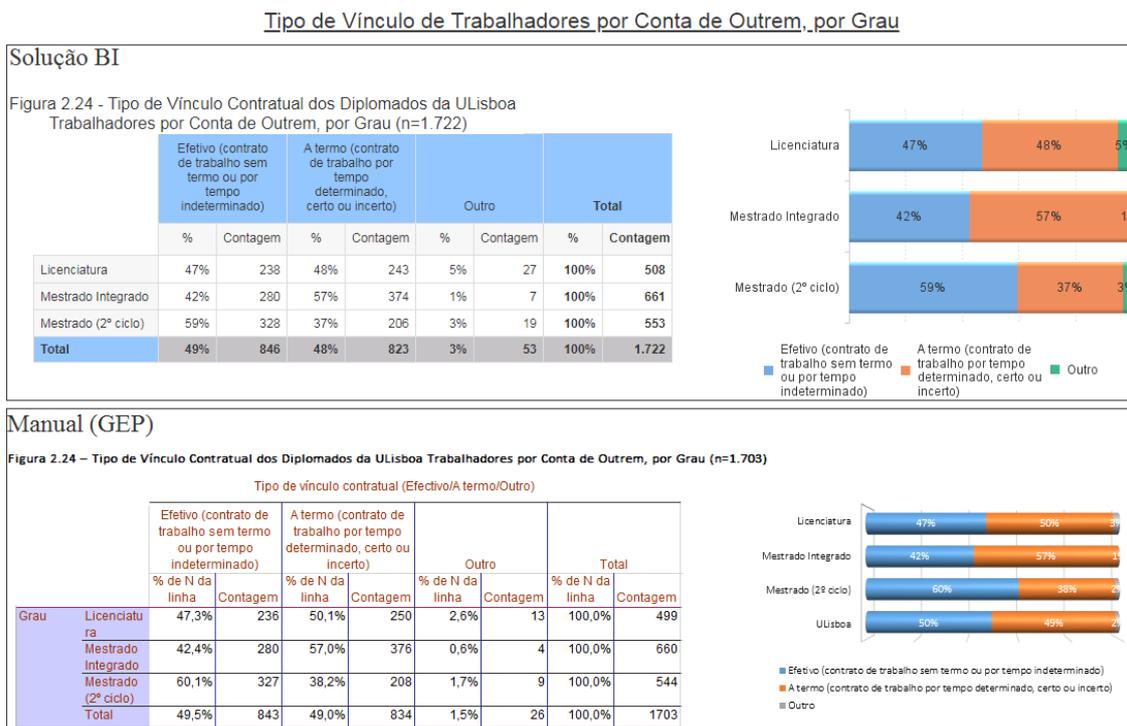


Figura 4.22: Tipo de vínculo de trabalhadores por conta de outrem por grau com o método manual e com a solução

4.2.2 Aplicação da solução a um novo ano de inquérito

Esta secção descreve os diferentes passos aquando da aplicação da solução a um novo ano de inquérito. Os dados fornecidos para este teste correspondem aos resultados do inquérito do ano de 2018 dos diplomados do ano letivo de 2015/16. Os dados do inquérito de 2018 à data deste trabalho de projeto ainda não se encontram disponíveis para o público, sendo aqui realizada uma descrição generalizada dos dados.

O processo iniciou-se com a junção manual dos resultados das escolas não autónomas num único ficheiro Excel, o qual deve seguir uma estrutura específica para que cada coluna do ficheiro corresponda a uma coluna da base de dados. Este ficheiro foi depois convertido para o formato CSV. Para cada escola autónoma (FC, ISA, ISEG, IST) foi criado um ficheiro CSV seguindo uma estrutura idêntica. Com os quatro ficheiros CSV devidamente estruturados, estes foram colocados nas suas pastas respetivas, com nomes específicos, como por exemplo dentro da pasta “LRUL” deve ser colocado um ficheiro CSV com o nome de “dados_rul_2017”. Uma vez que o SAP Data Services Designer lê os ficheiros pelo seu nome e pela sua pasta, é obrigatório seguir a nomenclatura estipulada, isto é, DADOS_[ESCOLA]_[ANO].

O próximo passo consistiu em executar os *jobs* implementados por ordem. Nesta fase foram encontrados alguns desafios, especificamente problemas novos de uniformização de

certos valores nos ficheiros CSV, os quais foram tratados manualmente a lista de valores a serem procurados e traduzidos para o valor respetivo, de forma semelhante à mostrada na Figura 4.7.

Uma vez ultrapassados os desafios, os dados foram automaticamente processados pelo SAP IDT, tendo sido corretamente organizados de acordo com o modelo de dados, assim como foram geradas as variáveis auxiliares, métricas e indicadores. O universo foi atualizado com o novo ano e publicado para uso da organização.

Na aplicação SAP WebI passaram a existir dois anos de escolha para análise, quando escolhido o ano de 2018, o número de respondentes a este novo inquérito foi semelhante ao do inquérito de 2017 e os resultados obtidos e apresentados nos relatórios predefinidos assemelham-se com os resultados obtidos de forma manual pelo GEP.

4.2.3 Comparação de tempos de processamento de dados

Esta secção compara estimativas de tempos de processamento de dados entre o trabalho desenvolvido de forma manual pelo GEP e a solução de BI desenvolvida.

O GEP anualmente realiza um inquérito de empregabilidade aos diplomados correspondentes a um determinado ano letivo. O trabalho manual inicia-se quando os colaboradores recebem os resultados dos inquéritos.

Uma vez com os resultados disponíveis, os processos manuais de transformação e tratamento aos dados começam, demorando cerca de quatro semanas a serem concluídos. Os processos incluem a junção dos resultados num só ficheiro, seguido da remoção de registos duplicados ou errados.

O preenchimento dos campos vazios com códigos “999” e “998” é a transformação mais custosa de todo o processo, seguindo-se a consolidação das repostas dos diplomados com perfis diferentes mas que concorrem para as mesmas perguntas ou indicadores, e concluindo com a correção de erros encontrados. Todos os registos são tratados individualmente, custando tempo e não sendo eficiente com quantidades significativas. Este fluxo de trabalho pode ser visto com detalhe na Secção 3.2.2.

Com os dados tratados e guardados num ficheiro único de Excel, estes são carregados para o SPSS onde o seu processamento demora cerca de um semana, devido à configuração de rótulos e variáveis auxiliares, e outros detalhes.

Com os dados no SPSS, são geradas as tabelas de análise e gráficos necessários para completar o relatório do inquérito à empregabilidade dos diplomados da ULisboa, que é reutilizado de ano para ano com as devidas alterações e revisão. A geração das tabelas e dos respetivos gráficos e a conclusão do relatório final, tem uma duração total de cerca três semanas.

A Tabela 4.3 indica resumidamente os processos descritos acima e as respetivas estimativas de tempos de execução, sendo que, em geral, o trabalho manual demora

Tabela 4.3: Tempo de cada processo realizado manualmente pelo GEP

Processo	Tempo
Transformação e tratamento de dados	4 semanas
Carregamento no SPSS e configuração	1 semana
Criação de análises e relatório final	3 semanas

Tabela 4.4: Tempo médio de execução de cada *job* da solução de BI desenvolvida

Job	Tempo médio
JOB_ODS_FC	17s
JOB_ODS_ISA	16s
JOB_ODS_ISEG	16s
JOB_ODS_IST	17s
JOB_UNIAO_ESCOLAS_AUTONOMAS	26s
JOB_LOAD_ESCOLAS_AUTONOMAS	24s
JOB_ODS_RUL	20s
JOB_UNIAO_RUL_AUTONOMAS	23s
JOB_TRANSFORMACAO_DADOS	3m45s
JOB_FINAL	23s
Total	7m17s

dois meses a ser concluído.

Um dos principais objetivos da solução de BI visa diminuir o tempo gasto com a realização do trabalho de forma manual e permitir à organização disponibilidade para novos objetivos.

Com a solução desenvolvida existem vários passos a seguir em determinada ordem. Inicialmente foi necessário realizar alguns processos de forma manual, isto é, quando os resultados dos inquéritos ficaram disponíveis, estes foram convertidos para ficheiros CSV e colocados em pastas próprias num formato específico, tomando meio dia de trabalho.

Com os ficheiros nas pastas devidas, o passo seguinte consistiu em executar os *jobs* implementados na ferramenta SAP Data Services Designer. Para tal, o utilizador com acesso ao SAP Data Services Management Console executa os *jobs* ordenadamente. Os tempos de execução de cada *job* podem ser consultados na consola onde são executados. No total existem dez *jobs* e a Tabela 4.4 indica o tempo médio que estes demoraram a executar numa máquina virtual com o sistema operativo Windows Server 2012 R2 STD, RAM de 8GB, HDD de 120GB e processador Intel(R) Xeon(R) CPU E5-2660 v2 @ 2.20GHz, 2200Mhz (4 cores).

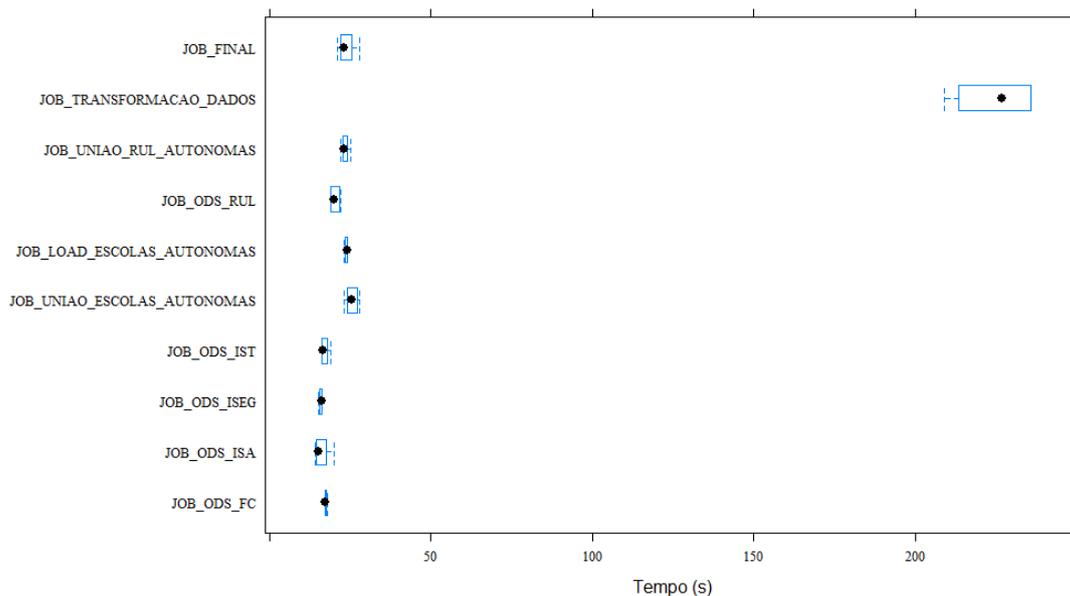


Figura 4.23: *Boxplot* dos tempos de execução de cada *job*

Com a Tabela 4.4 é possível observar que o JOB_TRANSFORMACAO_DADOS foi o que demorou mais tempo, sendo compreensível, pois os dados passaram por muitos processos, de modo a ficarem uniformizados e tratados. Semelhante ao trabalho manual do GEP, a atribuição dos códigos “999” foi a transformação mais longa do *job* pois envolveram muitas regras.

Cada *job* foi executado várias vezes para se obterem os tempos médios apresentados na Tabela 4.4, e dessas diferentes execuções foram criados *boxplots* para cada um dos *jobs* de modo a observar as diferenças de tempos e as suas discrepâncias. A Figura 4.23 mostra que a maioria dos *jobs* executa em menos de 30 segundos, não existindo valores muito distantes em cada *job*, com exceção do *job* JOB_TRANSFORMACAO_DADOS, que demorou bastante mais tempo a executar.

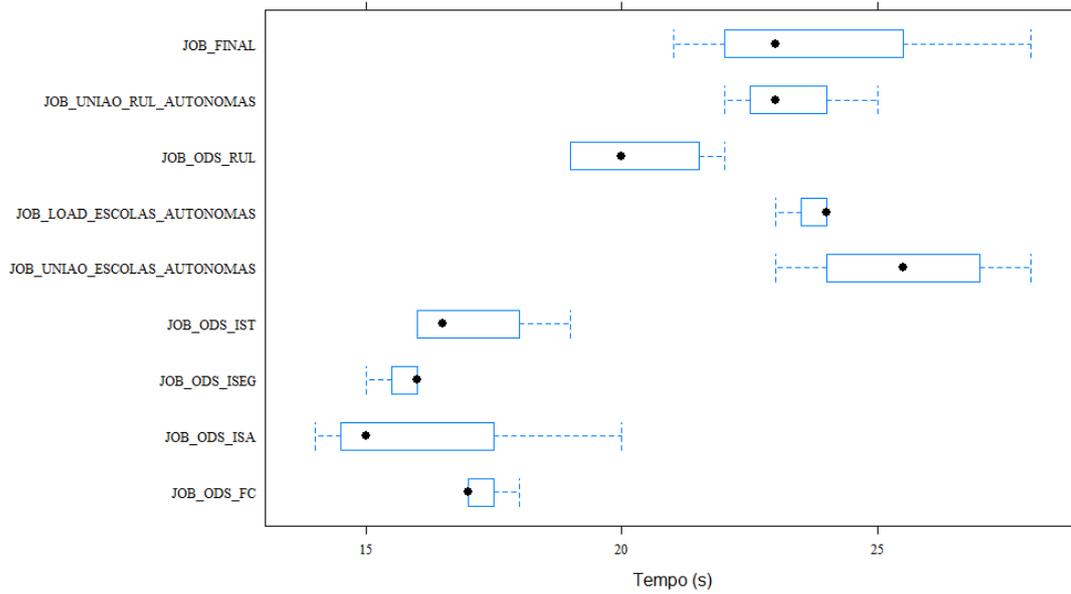


Figura 4.24: *Boxplot* dos tempos de execução dos *jobs* mais rápidos

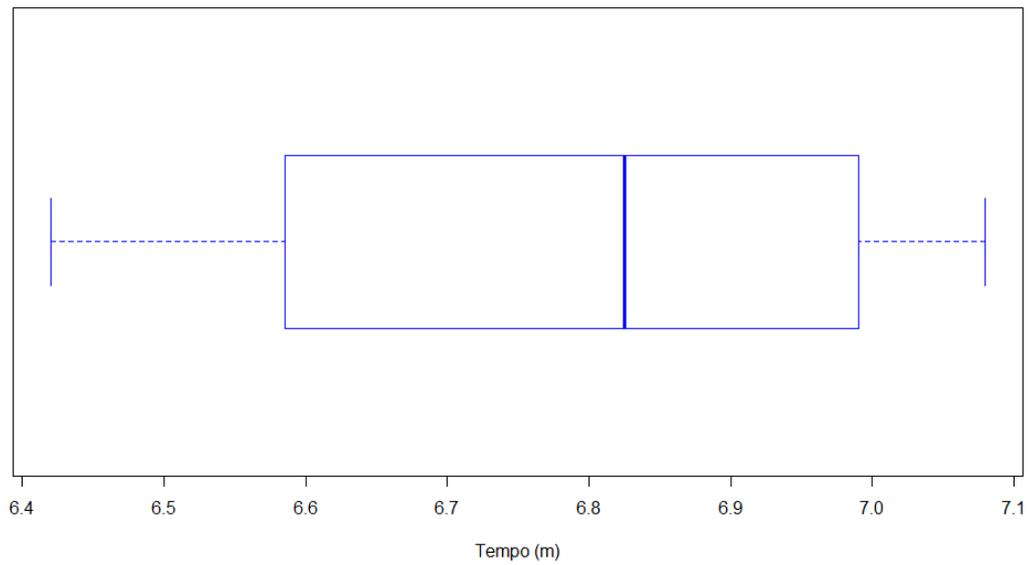


Figura 4.25: *Boxplot* dos tempos totais de execução de todos os *jobs*

A Figura 4.24 representa uma versão ampliada dos resultados, só com os tempos de execução mais baixos, isto é, excluindo o JOB_TRANSFORMACAO_DADOS, para uma melhor análise. É então, possível observar que os *jobs* responsáveis pelo carregamento das fontes de dados das escolas autónomas são o que executam mais rápido, tendo sido obtidos resultados inferiores a 20 segundos.

O *boxplot* da Figura 4.25 sintetiza os tempos totais em minutos de cada execução de todos os *jobs*, tendo os tempos ficado contidos entre os 6 e 7 minutos.

É possível afirmar que a solução reduz significativamente o tempo de trabalho principalmente porque a transformação dos dados passa de quatro semanas para um dia, permitindo à organização apresentar resultados com rapidez e ter disponibilidade para outras análises.

4.3 Sumário

Neste capítulo foi realizada uma descrição detalhada do desenvolvimento da solução de BI, incluindo o desenvolvimento das componentes ETL, da camada semântica, a criação de relatórios e a criação e configuração de utilizadores.

No final foi feita a avaliação da solução que permite afirmar que esta oferece melhorias para a organização como a diminuição significativa do tempo de processamento de dados, na consolidação da informação e na capacidade de manter histórico anual com os resultados.

Capítulo 5

Conclusão

Neste capítulo faz-se uma síntese do trabalho desenvolvido, incluindo as principais contribuições do projeto para a organização, as competências adquiridas, os desafios encontrados, e alguns possíveis desenvolvimentos futuros.

5.1 Principais contribuições

A principal contribuição deste trabalho foi o desenvolvimento de uma solução de BI que disponibiliza dados consolidados de empregabilidade de diplomados e que permite ao utilizador criar análises dinamicamente, facilmente e com coerência através de uma interface única e intuitiva. Como a solução armazena informação de anos anteriores, também é possível fazer análises com dados antigos e comparar com dados atuais, facilitando-se, por exemplo, o estudo de tendências.

Antes desta solução, o trabalho do GEP era concretizado de forma manual, com geração de quantidades consideráveis de ficheiros Excel com os diferentes passos do processo, até se chegar à criação de um documento consolidado. Este serviço era dispendioso tanto em tempo como recursos, com grande exigência de disponibilidade por parte dos elementos da organização. A nova solução permitiu mitigar grande parte do trabalho manual e aumentar o tempo disponível para novas tarefas.

A solução foi aceite com entusiasmo pelo GEP, mas com algum receio devido à perda de controlo sobre o conteúdo da informação e tomadas de decisão, por vezes de natureza discricionária. Para satisfazer as necessidades do GEP, a solução implementada automatiza os processos necessários para transformar os dados em informação útil, mas existe a possibilidade de a organização modificar a informação, preservando algum do controlo anterior.

Em suma, o contributo focou-se em automatizar um trabalho longo e desenhar um modelo de dados para suportar a informação, tendo sido criados processos de ETL que, a partir das fontes de dados, implementam regras de transformação de modo a replicar o trabalho de longas semanas do GEP. Com os dados tratados, desenvolveu-se o modelo

de dados e criaram-se as variáveis auxiliares, métricas e indicadores de acordo com os requisitos.

Outra contribuição importante foi a construção de relatórios predefinidos de modo a apresentar com rapidez análises do quotidiano, poupando tempo e esforço ao GEP. Para acesso às interfaces dos relatórios, criaram-se utilizadores e configuraram-se os mesmos com métodos de segurança para permitir o acesso. Com a finalidade de guiar os utilizadores e facilitar futuros desenvolvimentos foram gerados manuais de instruções.

5.2 Competências adquiridas

Este trabalho foi realizado no Departamento de Informática da Reitoria da ULisboa, envolvendo diferentes departamentos e equipas, dando a oportunidade de melhorar as competências de *soft skills*.

Ao longo do projeto foram aprofundados e adquiridos novos conhecimentos relacionados com BI, também foram aprendidas as ferramentas, nomeadamente o SAP Data Services Designer para desenvolver os processos de ETL, a SAP IDT para tornar os dados mais inteligíveis, incluindo a criação de variáveis auxiliares, métricas e indicadores, e a SAP Web Intelligence para criar relatórios com as principais análises. Parte destes conhecimentos adquiridos, deveram-se a uma formação para *power users* efetuada pela Novabase nas instalações da Reitoria da ULisboa.

Para além dessas ferramentas da SAP, aprendeu-se a criar e configurar utilizadores e a atribuir níveis de segurança com a Central Management Console, conhecimento este obtido através de outra formação realizada pela Novabase.

5.3 Desafios encontrados

No desenvolvimento deste projeto surgiram alguns desafios que eventualmente foram ultrapassados, que aumentaram o tempo da implementação da solução. O que possivelmente custou mais tempo do projeto foi o acesso às ferramentas pois foi necessário resolver algumas burocracias envolvidas, como por exemplo, acesso à chave de licença. Com o acesso às ferramentas o processo de instalação das mesmas levou tempo, tendo sido necessário esperar por uma máquina específica devido à compatibilidade entre as versões das ferramentas e o sistema operativo.

Outro desafio foi conseguir ter acesso aos dados de empregabilidade, devido à confidencialidade dos mesmos, tendo sido necessárias quase 3 semanas até ter autorização.

Por último, a aprendizagem e compreensão da melhor abordagem de utilização das ferramentas, principalmente com a aplicação SAP Data Services Designer, mas rapidamente este desafio foi superado, através de tutorias e de interações com conhecedores da ferramenta.

5.4 Trabalho futuro

A principal ideia de trabalho futuro é melhorar a capacidade de comparação entre dados de empregabilidade de anos diferentes, não fez parte dos requisitos da solução mas pode ser concretizada com uma ferramenta complementar da SAP.

Com a flexibilidade e evolução da ferramenta responsável pelos relatórios, a criação de novos relatórios e análises complexas, poderá ser necessária futuramente.

Se vierem a ser necessários *dashboards*, estes podem ser integrados com a solução desenvolvida com recurso a ferramentas da SAP .

Bibliografia

- [1] Mika Aho. The distinction between Business Intelligence and corporate performance management - a literature study combined with empirical findings. In *Proceedings from the sixth annual Mini Conference on Scientific Publishing (MCSP)*. Citeseer, 2010.
- [2] Surajit Chaudhuri and Umeshwar Dayal. An overview of data warehousing and OLAP technology. *ACM Sigmod Record*, 26(1):65–74, 1997.
- [3] Surajit Chaudhuri, Umeshwar Dayal, and Vivek Narasayya. An overview of business intelligence technology. *Communications of the ACM*, 54(8):88–98, 2011.
- [4] Marcelo Costa. Sobre business intelligence e dados. <https://www.infoq.com/br/articles/bi-e-dados>, 2018. (Acedido em 18/12/2018).
- [5] Umeshwar Dayal, Malu Castellanos, Alkis Simitsis, and Kevin Wilkinson. Data integration flows for business intelligence. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, pages 1–11. ACM, 2009.
- [6] Gabinete de Estudos e Planeamento. Inquérito à Empregabilidade dos Diplomados da ULisboa 2014/15 (Documento de divulgação restrita), 2018.
- [7] Rosie Harvey. Preparing data in Oracle business intelligence cloud service. <https://docs.oracle.com/en/cloud/paas/bi-cloud/bilpd/preparing-data-oracle-business-intelligence-cloud-service.pdf>, 2018. (Acedido em 22/07/2019).
- [8] Ralph Kimball and Margy Ross. *The data warehouse toolkit: the complete guide to dimensional modeling*. John Wiley & Sons, 3 edition, 2011.
- [9] Antti Lönnqvist and Virpi Pirttimäki. The measurement of business intelligence. *Information Systems Management*, 23(1):32, 2006.
- [10] M. Miškuf, I. Zolotová, and M. Nemčík. Application of business intelligence solutions from Microsoft and IBM on manufacturing data. In *Proceedings of the IEEE 14th*

International Symposium on Applied Machine Intelligence and Informatics (SAMI), pages 41–45, 2016.

- [11] Daniel L Moody and Mark AR Kortink. From enterprise models to dimensional models: a methodology for data warehouse and data mart design. In *Proceedings of the International Workshop on Design and Management of Data Warehouses*, pages 5–2, 2000.
- [12] Novabase. Solução BI na ULisboa - Caderno de especificação técnica. Technical report, 2018.
- [13] Elana Roth. What is ad hoc reporting - and do i need it? <https://www.sisense.com/blog/ad-hoc-reporting-need/>, 2017. (Acedido em 21/12/2018).
- [14] SAP. About the central management console - SAP help portal. <https://help.sap.com/viewer/2710d05c6f7f452baddb4ca72b72fdc2/4.2.5/en-US/46f328776e041014910aba7db0e91070.html>. (Acedido em 26/11/2018).
- [15] SAP. Information design tool user guide. https://help.sap.com/http.svc/rc/businessobject_product_guides_sbo42_en_sbo42_info_design_tool_en_pdf/4.2/en-US/sbo42_info_design_tool_en.pdf. (Acedido em 26/11/2018).
- [16] SAP. SAP BusinessObjects - Business Intelligence (BI) Suite. <https://www.sap.com/products/bi-platform.html>. (Acedido em 16/11/2018).
- [17] SAP. SAP BusinessObjects Lumira Discovery - SAP help portal. <https://help.sap.com/viewer/69ab7e78763a4d929126440ab1b7deal/2.0.3.0/en-US/30330f0f3c164bc1b8cfb9fcd269948d.html>. (Acedido em 26/11/2018).
- [18] SAP. SAP BusinessObjects Web Intelligence user's guide. https://help.sap.com/doc/4f6c0512515d48e884444d0671980060/4.2.4/en-US/webi42sp4_user_guide_en.pdf. (Acedido em 22/07/2019).
- [19] SAP. SQL Anywhere Introduction. <https://help.sap.com/doc/a729028300694e19aacfcdd2aea457c7/17.0/en-US/Introduction-en.pdf>. (Acedido em 22/07/2019).
- [20] SAP. Tutorial. https://help.sap.com/doc/83d1c9b908d548c090d5d509ad27d964/4.2.11/en-US/ds_42_tutorial_en.pdf. (Acedido em 22/07/2019).

-
- [21] ULisboa. Sobre nós ULisboa. <https://www.ulisboa.pt/sobre-nos#video>. (Acedido em 08/11/2018).
- [22] Hugh J Watson and Barbara H Wixom. The current state of business intelligence. *Computer*, 40(9), 2007.

Abreviaturas

BI *Business Intelligence*

DW *Data Warehouse*

ETL Extração-Transformação-Carregamento

GEP Gabinete de Estudos e Planeamento

SAP Sistemas, Aplicações e Produtos

SAP CMC SAP Central Management Console

SAP IDT SAP Information Design Tool

SAP WebI SAP Web Intelligence

SAPR Sem atividade profissional remunerada

SPSS Statistical Package for the Social Sciences

ULisboa Universidade de Lisboa

Apêndice A

Manual de instalação das ferramentas SAP Server

Instalação Ferramentas SAP

SAP BO BI PLATFORM – SAP DATA SERVICES – SAP IQ

Introdução

Para este tutorial foi utilizada uma máquina virtual com o sistema operativo Windows Server 2012 R2 64x.

As instalações foram realizadas na mesma máquina, para usar várias máquinas é necessário ler a documentação para uma melhor compreensão. Esta documentação está associada a links partilhados ao longo do documento.

Para iniciar as instalações é necessário uma conta oficial SAP, do qual, irá permitir acesso à página de Downloads e de Licenças de chave dos produtos SAP.

A página de acesso para os produtos, entre outras funcionalidades encontra-se no link:
<https://launchpad.support.sap.com/>

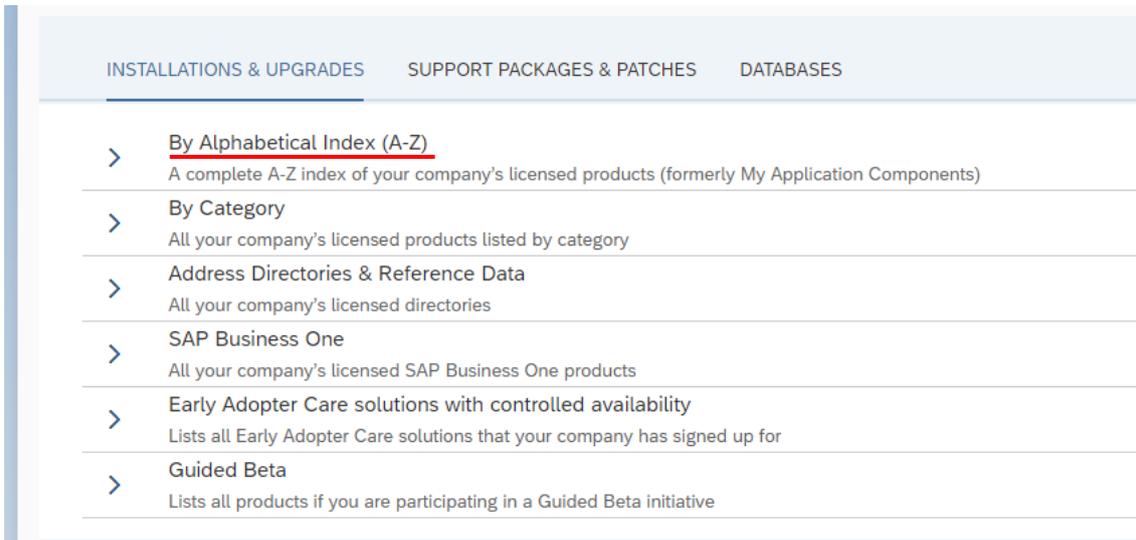
Instalação SAP BusinessObjects BI Platform

Para melhor informação consultar:

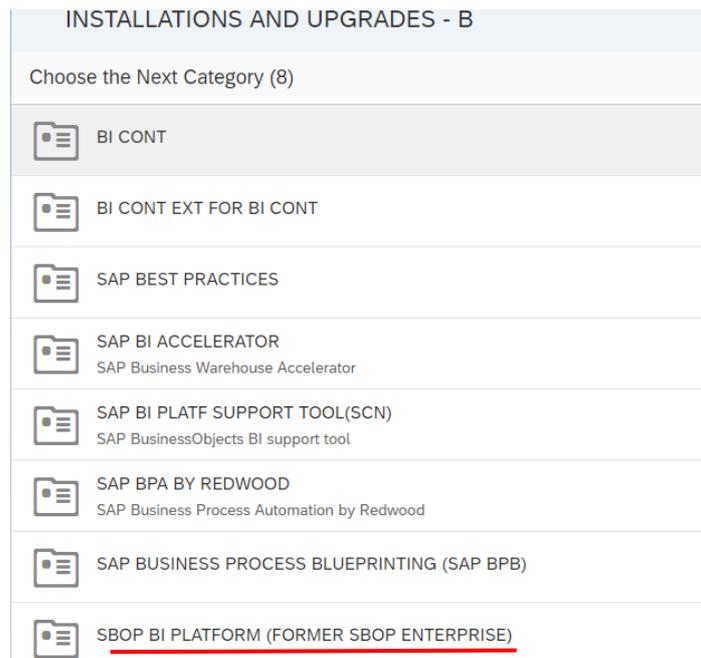
https://help.sap.com/viewer/product/SAP_BUSINESSOBJECTS_BUSINESS_INTELLIGENCE_PLATFORM/4.2.7/en-US

No link acima tem um guia de instalação para os diferentes sistemas operativos e versões do produto, qualquer dúvida deverá consultar os documentos disponibilizados.

Para fazer o Download do SAP BO BI Platform é necessário fazer login na página, <https://launchpad.support.sap.com/#/softwarecenter>. Clicar em *By Alphabetical Index (A-Z)*.



De seguida clicar na letra **B**, procurar pela pasta **SBOP BI PLATFORM (FORMER SBOP ENTERPRISE)**

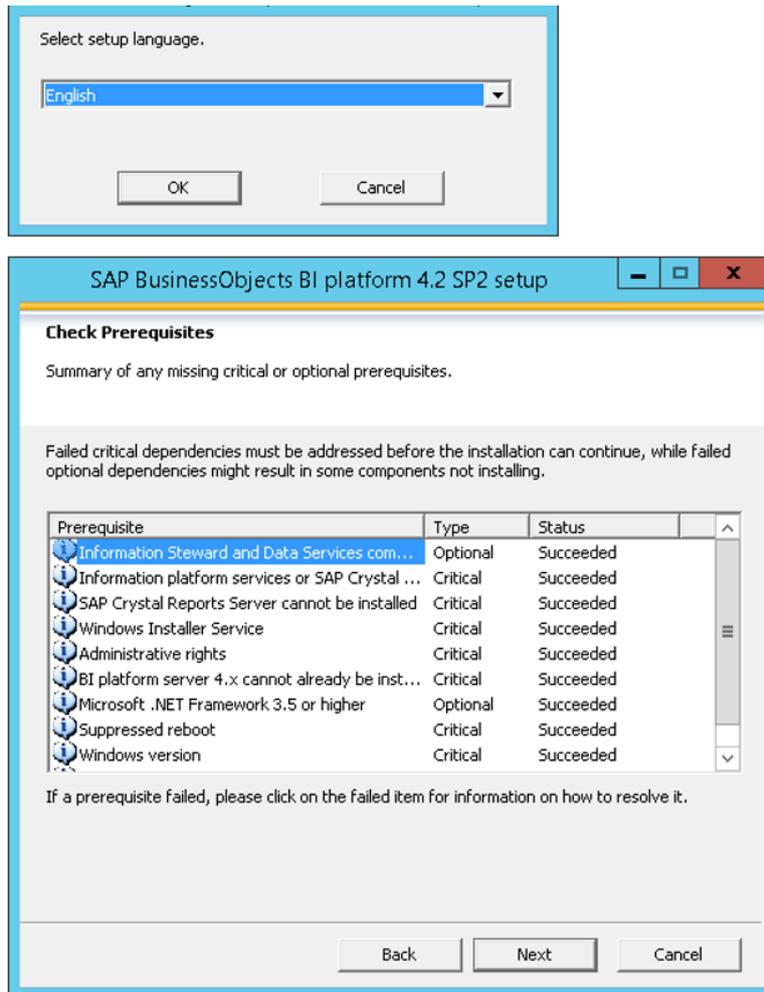


Clicar em *SBOP BI PLATFORM (ENTERPRISE)*, para este tutorial foi escolhido o *SBOP BI PLATFORM 4.2*, clicar em *INSTALLATION AND UPGRADE*.

Filtrar pelo sistema operativo, neste caso Windows, escolher o ficheiro com o nome *SBOP BI PLATFORM 4.2 SP07 SERVER WINDOWS 64B*. Este ficheiro foi escolhido pois era o mais atualizado na altura, entretanto novas versões SPXX podem aparecer.

Depois de seleccionar o produto, fazer o download.

Fazer a extração do produto e iniciar o setup.exe como administrador.



Aceitar o License Agreement.

Para preencher o campo do *Product key* é necessário voltar ao launchpad e realizar o pedido de uma chave.

<https://launchpad.support.sap.com/#/licensekey/wizard>

Preencher os campos como está na imagem em baixo e clicar em Search caso não vá automaticamente para próxima página.

Select installation number using filters

Customer:
 Product:
 Installation Number:

Or enter the Installation/System number directly

Installation Number or System Number:



Escolher a versão do produto desejada e colar a *License Key* na instalação do produto.

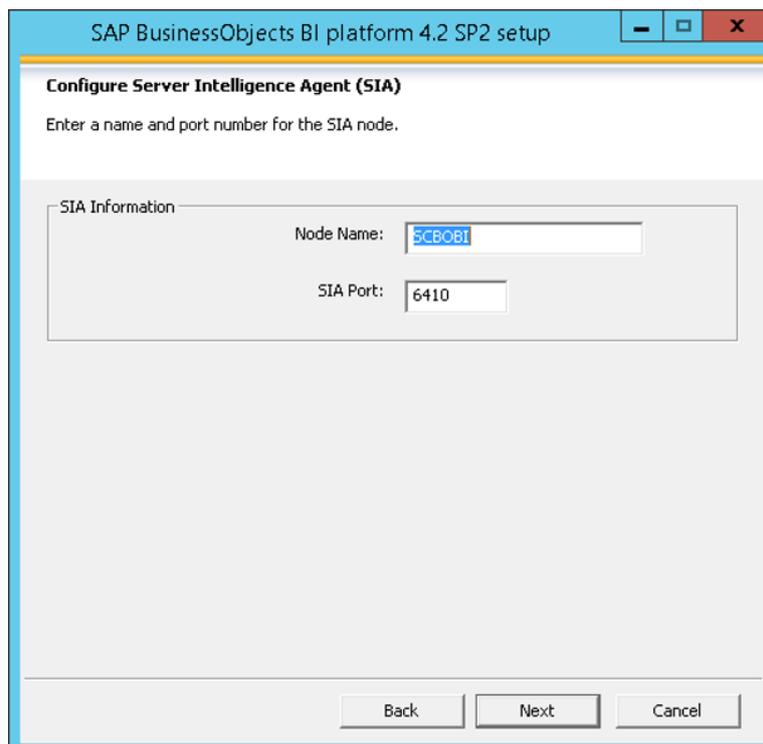
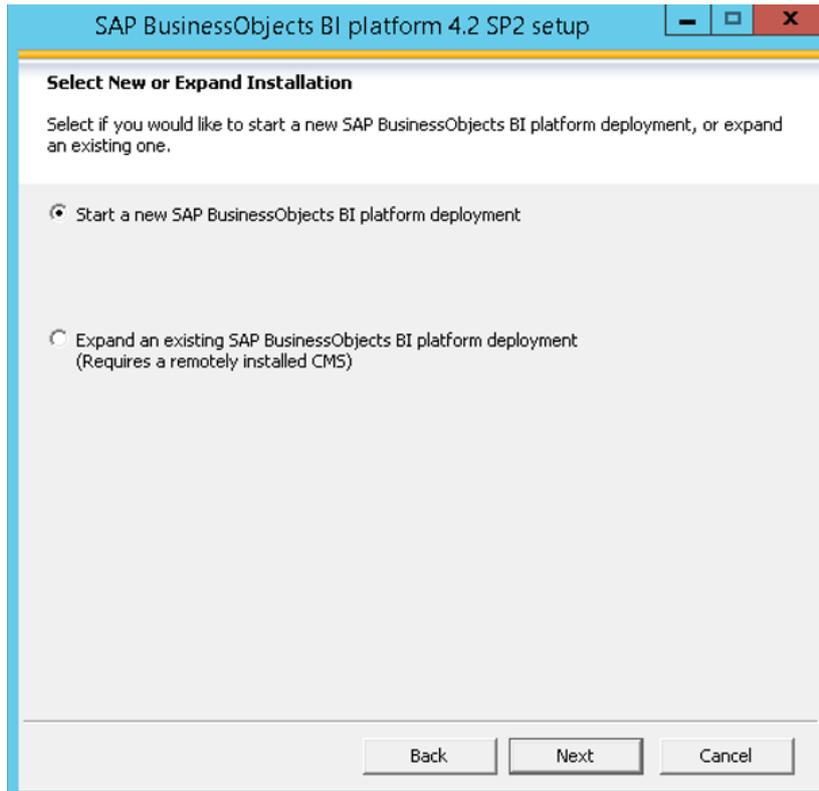
1- Installations 2- Systems 3- Details

> Installation Licenses

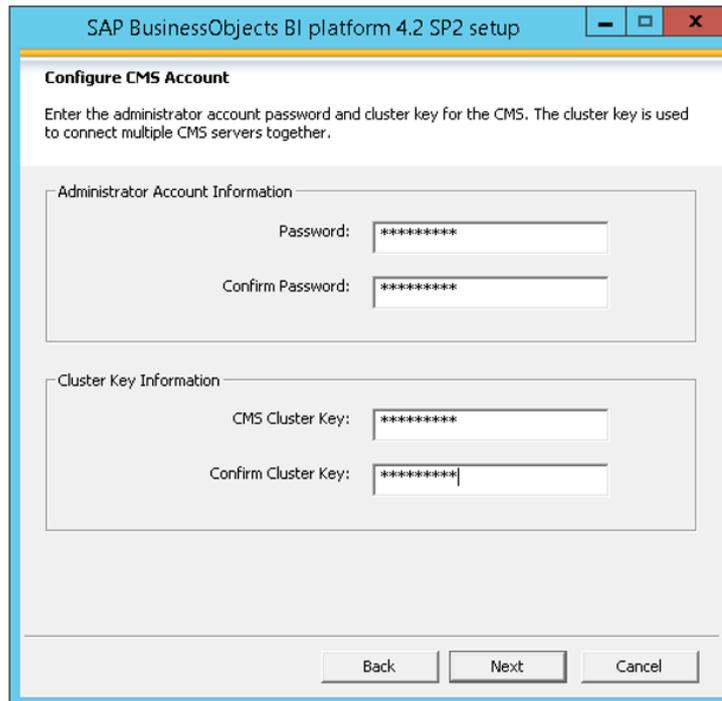
Systems (6) + ▾ ↕

SID	System Number	System Name	Product Version	System Type	
BOD			SAP BusinessObjects Business Intelligence platform 4.2	Development system	>
BOP			SAP BusinessObjects Business Intelligence platform 4.2	Productive System	>
DSD			SAP Data Services 4.2	Development system	>
DSP			SAP Data Services 4.2	Productive System	>
IQD			SAP IQ 16.1	Development system	>
IQP			SAP IQ 16.1	Productive System	>

Continuar com a instalação do produto e escolher o tipo de instalação como *FULL*, continuar com os procedimentos que vão aparecendo. Selecionar sempre a opção de criar algo novo caso seja a primeira instalação.

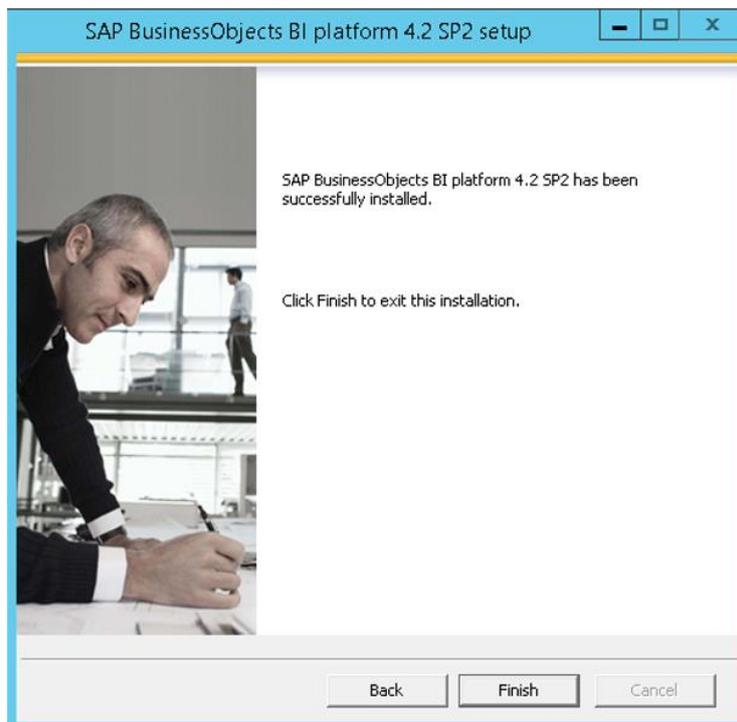


Deixar os ports como default a não ser que estejam a ser utilizados.



Criar passwords para configurar a conta CMS (Central Management Server)

Continuar com o processo de instalação até ao fim.



Instalação SAP Data Services 4.2

Para melhor informação consultar:

https://help.sap.com/viewer/product/SAP_DATA_SERVICES/4.2.11/en-US

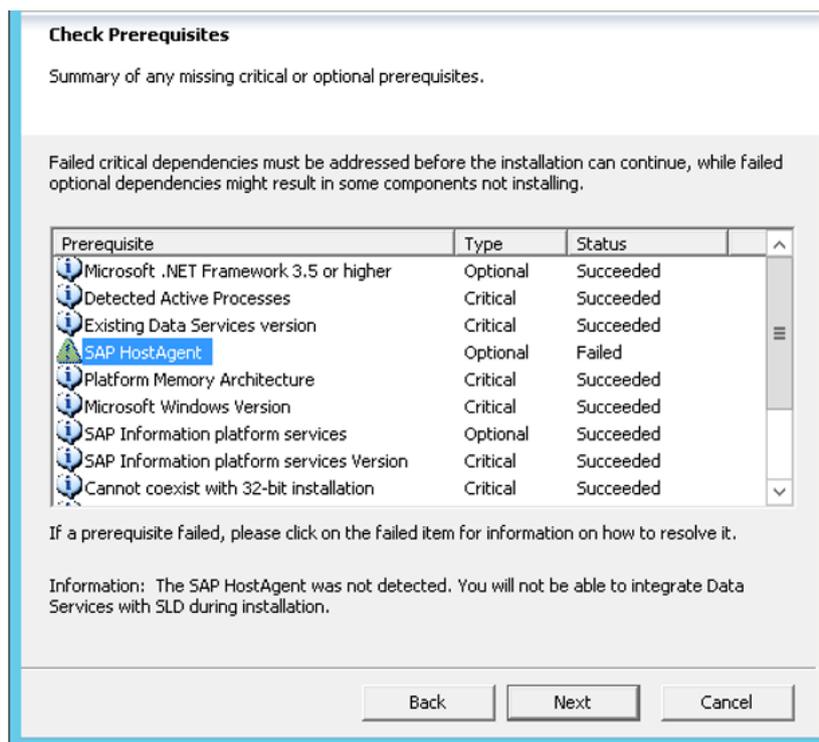
No link acima tem um guia de instalação para os diferentes sistemas operativos e versões do produto, qualquer dúvida deverá consultar os documentos disponibilizados.

Para fazer o Download do SAP Data Services é necessário fazer login na página, <https://launchpad.support.sap.com/#/softwarecenter>. Clicar em *By Alphabetical Index (A-Z)*.

Clicar na letra **D** e escolher *SAP DATA SERVICES -> SAP DATA SERVICES 4.2 -> INSTALLATION AND UPGRADE -> Filtrar por Windows -> SAP DATA SERVICES 4.2 SP11 WINDOWS (64B)*. Foi escolhida esta versão pois era a mais atualizada.

Fazer download do ficheiro e realizar a extração depois.

Correr o setup.exe como administrador



Seguir a instalação aceitando os acordos de licença e escolhendo as opções de default. Sempre que aparecer a opção de criar algo novo, esse deverá ser o caminho a seguir. Isto para a instalação do SQL Anywhere e Tomcat.

Para obter a chave de licença deverá realizar o mesmo procedimento que na instalação anterior (*SBOP BI PLATFORM*).

Escolher o sistema de login com a conta de sistema.

Login Information
Specify login information for the account that Data Services should use for system services.

Login information

Account used by the Data Services system service.

Use system account

Use this account: (Recommended)

User:

Password:

Back Next Cancel

Instalação SAP IQ

Para melhor informação consultar:

https://help.sap.com/viewer/product/SAP_IQ/16.1.2.14/en-US

No link acima tem um guia de instalação para os diferentes sistemas operativos e versões do produto, qualquer dúvida deverá consultar os documentos disponibilizados.

Para fazer o Download do SAP IQ é necessário fazer login na página, <https://launchpad.support.sap.com/#/softwarecenter>. Ir para *SUPPORT PACKAGES & PATCHES*. Clicar em *By Alphabetical Index (A-Z)*.

Clicar na letra **I** -> SYBASE IQ -> SYBASE IQ 16.1 -> SYBASE IQ SERVER 16.1 -> Filtrar por sistema operativo Windows -> SAP IQ 16.1 SP02 PL14.

Esta foi a versão escolhida por ser a mais atualizada na altura, entretanto é possível que novas versões estejam disponíveis.

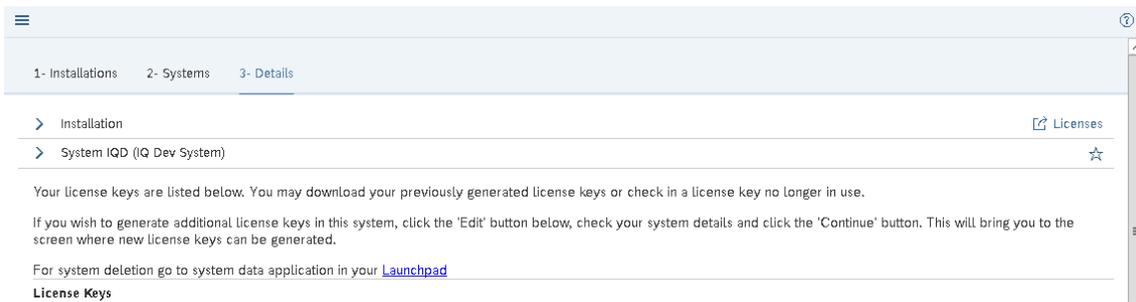
Depois de concretizado o download fazer a extração do mesmo.

Fazer o setup.exe como administrador seguir a instalação como default, deixar os ports de origem e alterar apenas se estes já estão a ser utilizados.



A instalação está feita para apenas durar 90 dias, é necessário criar chaves para que a sua utilização seja ilimitada.

Para tal, e especificamente para esta instalação foi criada uma licença *unserved*. Seguindo o mesmo processo de requisição de chave como as outras ferramentas.



Clicar em *Edit* no fundo da página para criar uma chave, preencher os campos com as necessidades pretendidas.

De uma lista de possíveis chaves a gerar, é necessário escolher a opção a desejar.

Gerar a chave e clicar em *Unsaved License*, escolher o número de máquinas. Documentação disponível no site anteriormente partilhado onde é explicado a diferença entre licenças *Served* e *Unsaved*, deverá escolher o que se adequa para os seus objetivos.

Preencher os campos consoante a máquina. Ter atenção no formato com que se preenche o Host ID. Perceber o número de licenças, para isto deverá seguir os links apresentados neste passo para uma melhor compreensão.

Depois de gerar a chave, deve fazer download das chaves e colocar na pasta devida.

... \SapIQ\SYSAM-2_0\licenses

Generate License for SAP IQ 16.1

Previous Selections

Step	Selection
License Model	Unserviced License
License Model	Quantity = 1

Step 3: Host Machines

Licenses available: 8

[What is my Host ID?](#)

[What is Number of Licenses?](#)

Host ID*:

Host Name*:

Number of Licenses*:

[Back](#) [Generate](#) [Cancel](#)

Apêndice B

Manual de instalação das ferramentas SAP Client

Instalação Ferramentas SAP Client

SAP BO BI PLATFORM – SAP IQ – SAP DATA SERVICES

Instalação do SBOP PLATFORM 4.2 SP05 Client Tools Windows 32B

Correr o setup.exe

 Collaterals	05-04-2019 09:22	Pasta de ficheiros	
 dunit	05-04-2019 09:27	Pasta de ficheiros	
 langs	05-04-2019 09:34	Pasta de ficheiros	
 autorun.inf	29-11-2017 05:33	Informações de c...	1 KB
 ProductId.txt	29-11-2017 05:33	Ficheiro TXT	1 KB
 response.ini	29-11-2017 05:33	Definições de con...	1 KB
 setup.exe	29-11-2017 05:33	Aplicação	355 KB
 setup.ini	29-11-2017 05:33	Definições de con...	1 KB
 SIGNATURE.SMF	08-12-2017 00:20	Ficheiro SMF	3.713 KB

Deixar maioritariamente tudo por default. Na situação de escolher a autenticação, o utilizador deverá escolher a autenticação usando o System Account.

As aplicações abaixo deveram de aparecer após a instalação.

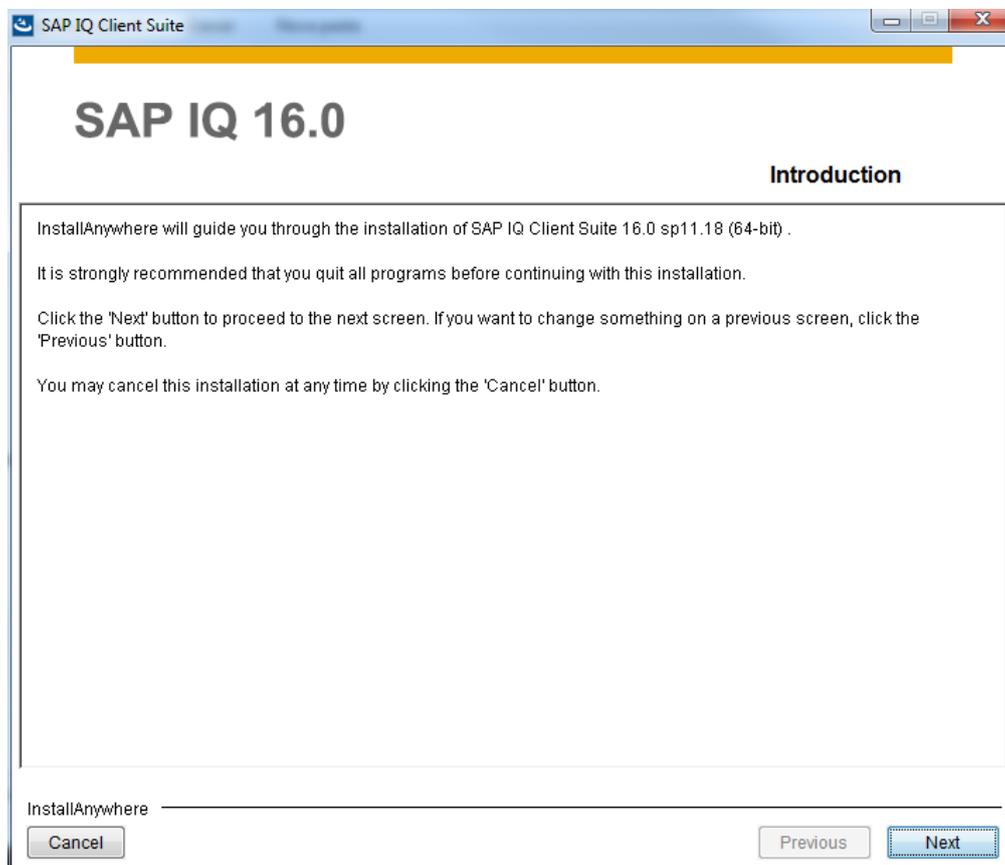
 Business View Manager	03/04/2019 10:17	Atalho	2 KB
 Data Federation Administration Tool	03/04/2019 10:17	Atalho	2 KB
 Information Design Tool	03/04/2019 10:18	Atalho	2 KB
 Query As a Web Service Designer	03/04/2019 10:17	Atalho	2 KB
 Report Conversion Tool	03/04/2019 10:17	Atalho	3 KB
 Translation Management Tool	03/04/2019 10:18	Atalho	2 KB
 Universe Design Tool	03/04/2019 10:18	Atalho	3 KB
 Web Intelligence Rich Client	03/04/2019 10:17	Atalho	2 KB
 Widgets	03/04/2019 10:18	Atalho	2 KB

Instalação Sap IQ Client

Correr o setup.exe que está presente na pasta IQNC160011P_18-20011253.

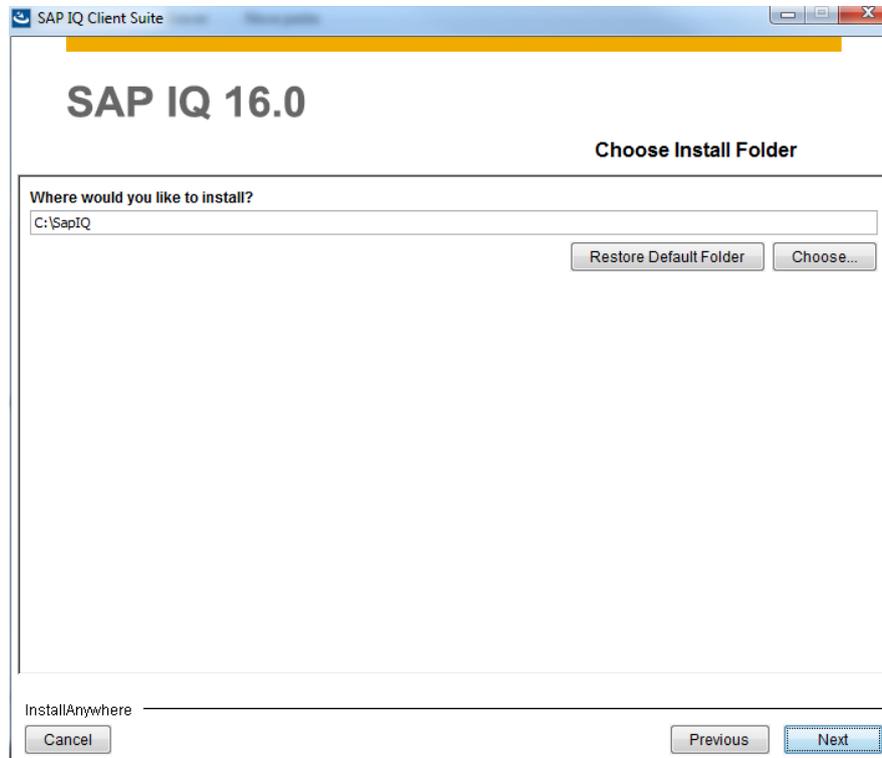
📁 archives	05-04-2019 16:21	Pasta de ficheiros	
📄 Cover.SP11.18.txt	05-03-2018 21:46	Ficheiro TXT	4 KB
📄 installer.properties	05-03-2018 21:46	Ficheiro PROPERT...	1 KB
📄 readme.txt	05-03-2018 21:46	Ficheiro TXT	1 KB
📄 readme_ja.txt	05-03-2018 21:46	Ficheiro TXT	1 KB
📄 readme_zh.txt	05-03-2018 21:46	Ficheiro TXT	1 KB
📄 setup.exe	05-03-2018 21:46	Aplicação	86.334 KB
📄 setupConsole.exe	05-03-2018 21:46	Aplicação	86.334 KB

Clicar em Next

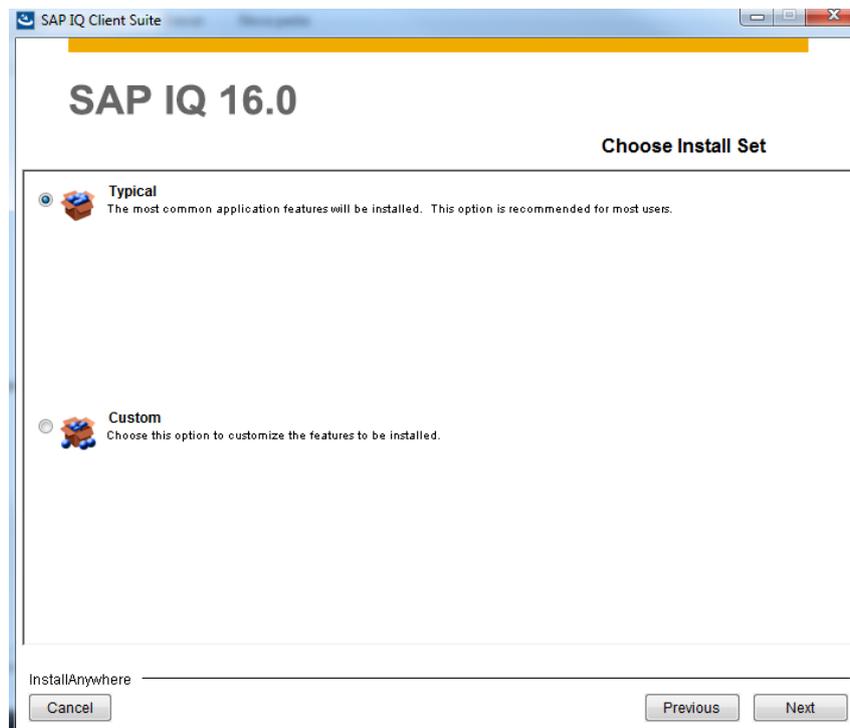


Continuar a clicar em Next.

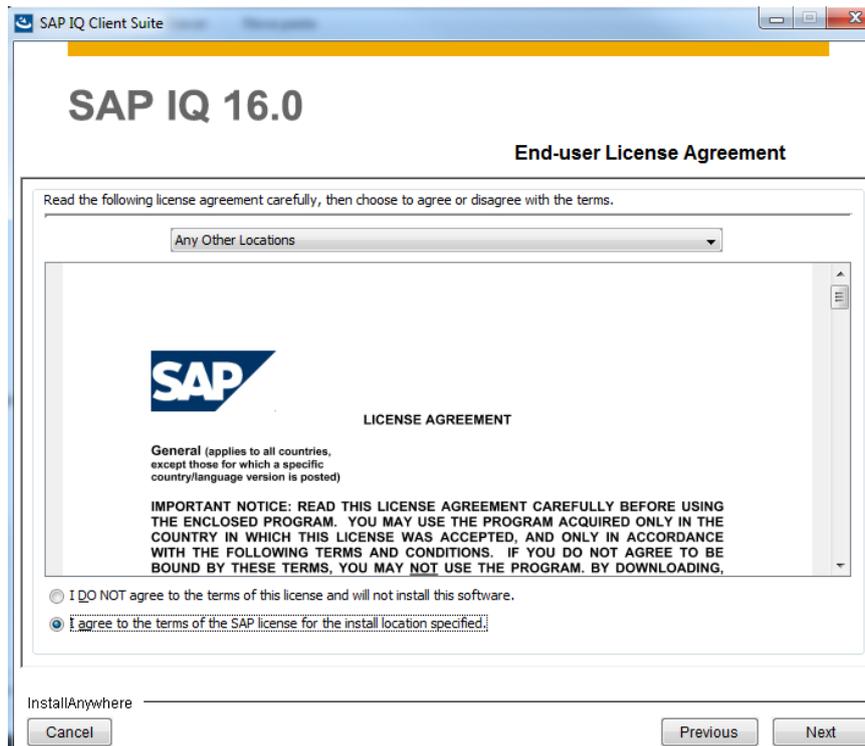
Um pop-up vai aparecer a dizer que a diretoria não existe e se deseja criar. Neste passo deve clicar em Yes.



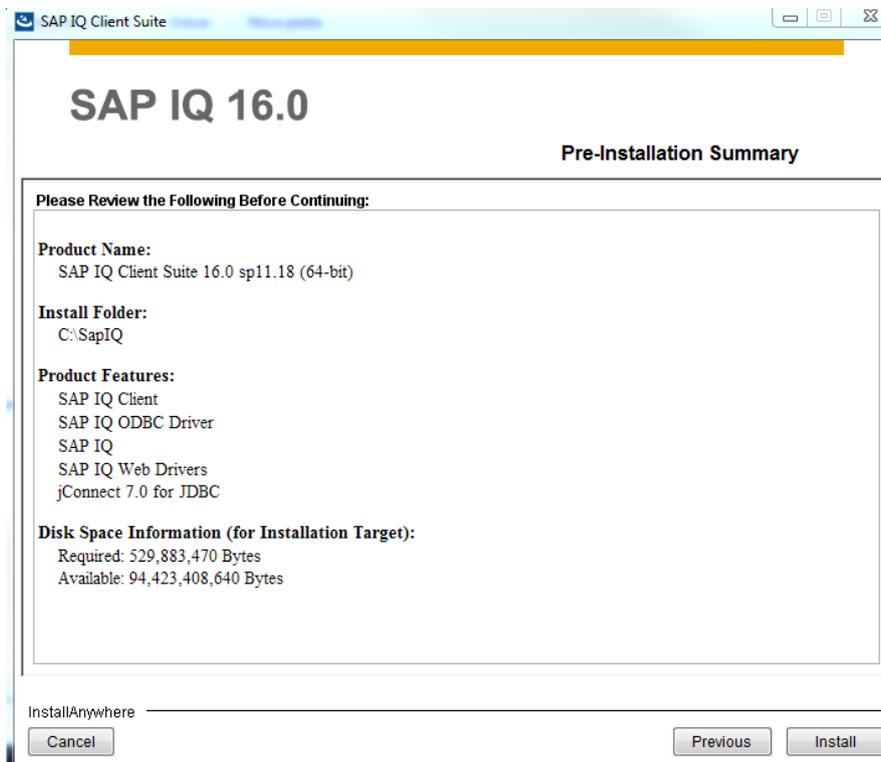
Escolher a opção de Typical e clicar em Next.



Escolher a localização, aceitar os termos e clicar em Next.



Iniciar a instalação.

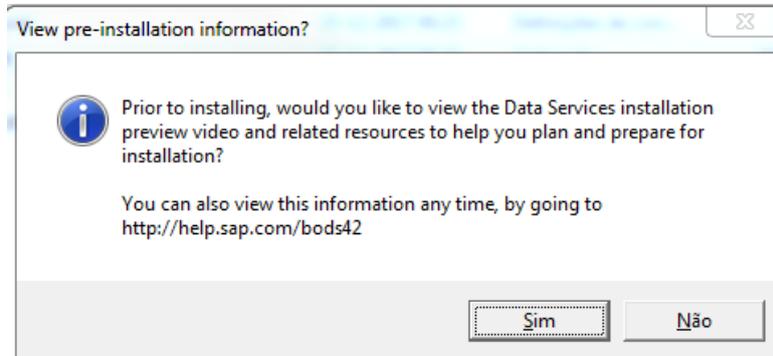


A instalação deverá responder com Successful e clicar em Done.

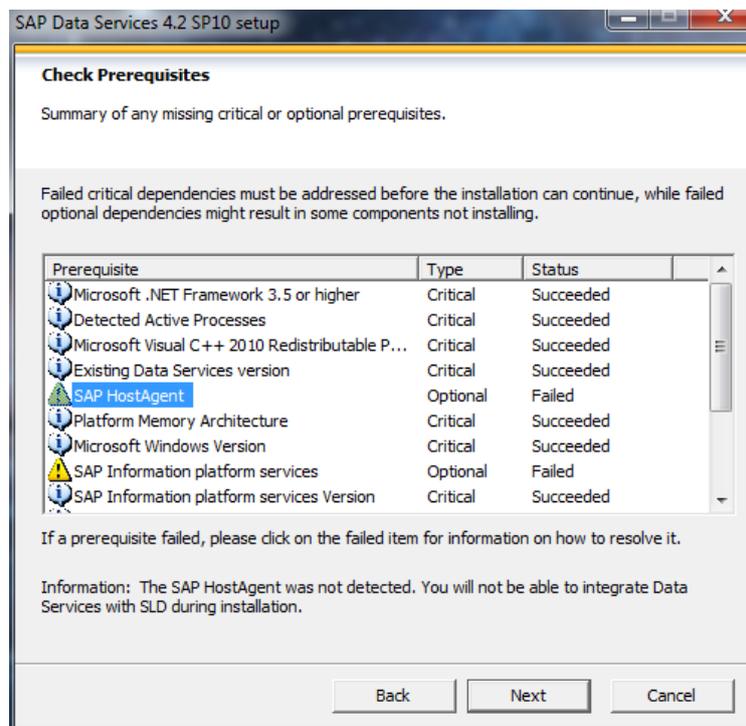
Instalação da ferramenta SAP Data Services 4.2 SP 10 Multi Platform 64B

Executar o setup.exe em modo administrador.

Caso queira visualizar informação antes da instalação pode clicar em Sim no pop-up que irá surgir e este levará para o site. Para iniciar a instalação rapidamente foi clicado Não no pop-up.

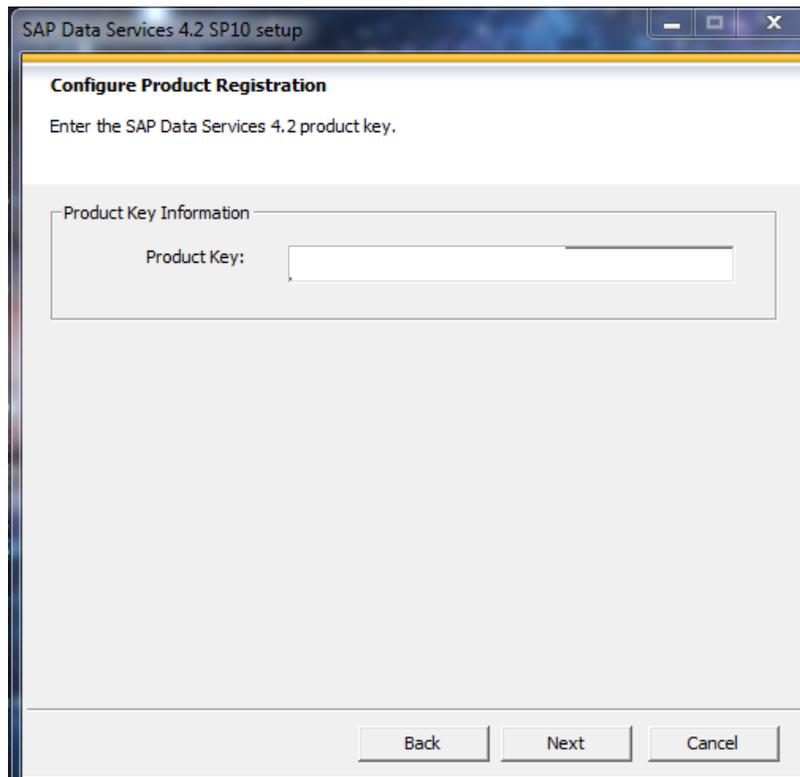


Clicar em Next.

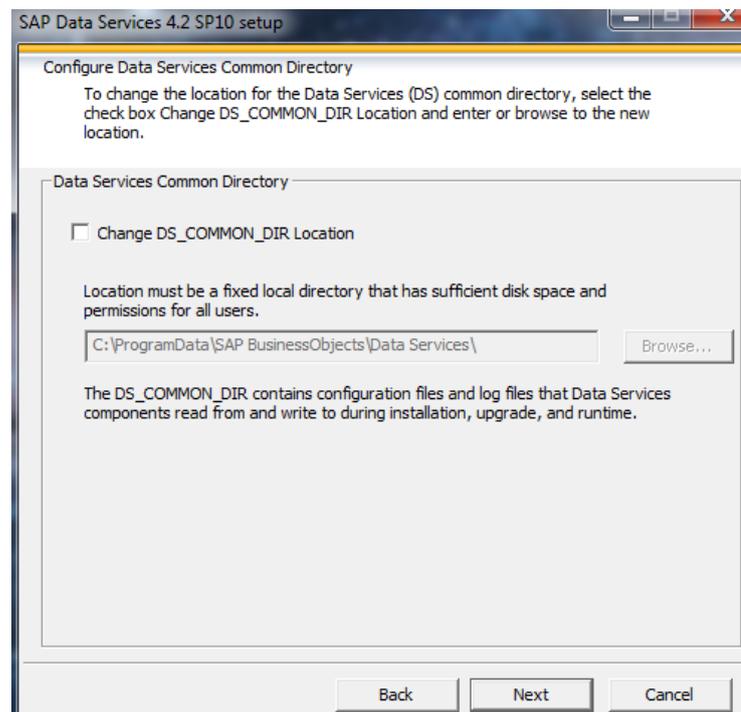


Clicar em Next e aceitar o acordo de licença.

No próximo ponto é necessário adicionar a chave do produto.

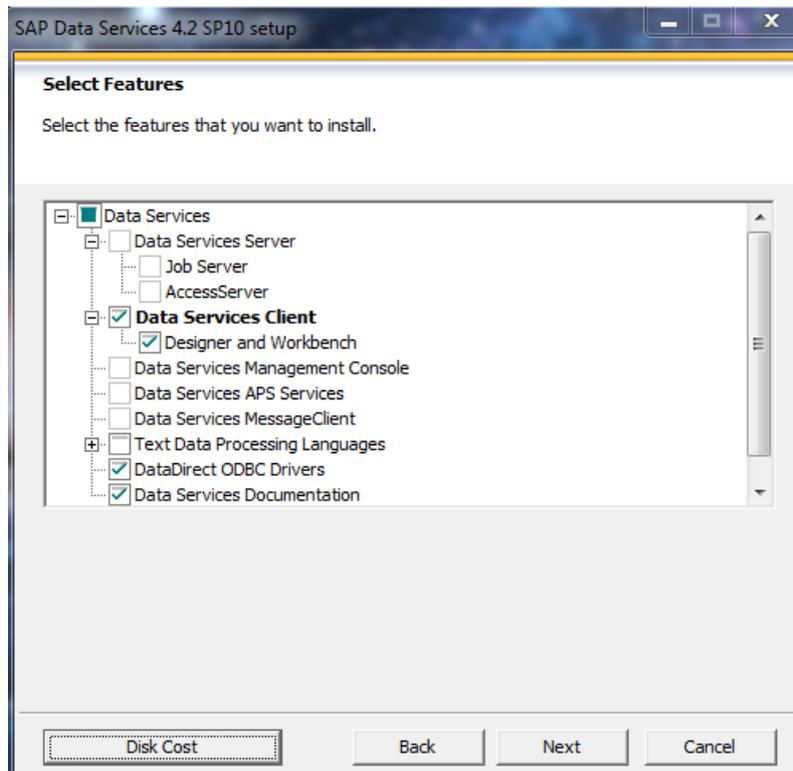


Clicar em Next

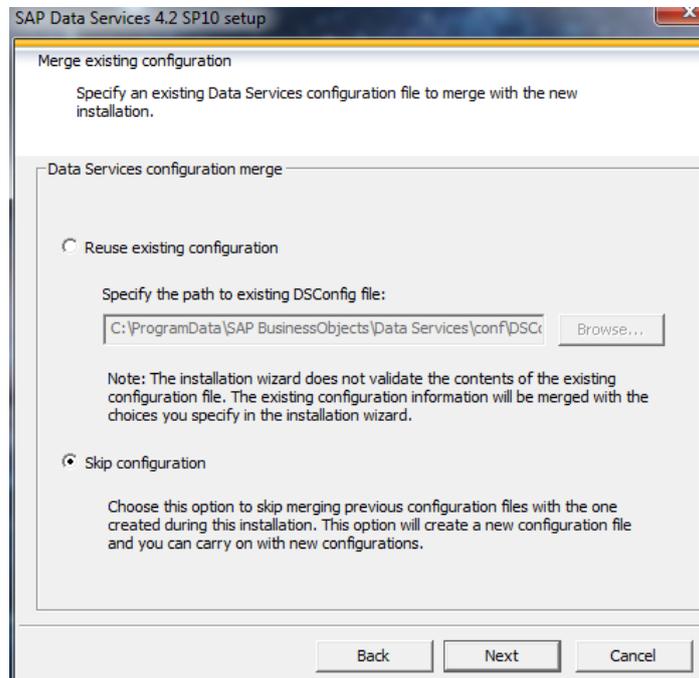


Escolher o pacote de linguagens caso tenha preferências, senão continuar a clicar Next.

Instalar os itens selecionados



Fazer Skip configuration.



Apêndice C

Análise de dados da 3^a edição do inquérito à empregabilidade dos diplomados da ULisboa

Análise de dados

3ª Edição do inquérito à empregabilidade da Universidade de Lisboa

ANA PATRÍCIA PINTO ESPINHEIRA

Índice

Análise dos dados.....	5
ID do respondente (ID interno e não de aluno)	8
ID serviços académicos	8
Nacionalidade.....	8
Nacionalidade.....	9
Nacionalidade (PT/Europeu/PLOP/Outro)	10
Internacionalização	10
Internacionalização dos diplomados (variável auxiliar)	11
Internacionalização dos diplomados/discriminada (variável auxiliar)	11
Unidade orgânica de formação (dentro da ULisboa)	13
Grau de formação	14
Código de curso.....	15
Classificação final de curso.....	16
Classificação Nacional de Áreas de Educação e Formação (CNAEF)	17
CNAEF 1	17
CNAEF 2	18
CNAEF 2 (com 311-Psicologia e 314-Economia).....	20
CNAEF 3	21
Designação do curso	24
Área de estudo	25
Data de nascimento	26
Idade à data de fim do curso/da resposta ao inquérito.....	27
Idade à data da diplomação	27
Idade à data da resposta	29
Género.....	31
Local de residência	32
Residência (Portugal/Europeu/PLOP/Outro) (variável auxiliar)	33
Residência (Portugal/Outro país)	33
Residência por Distrito de Portugal	33
Localização da residência (AML) (variável auxiliar).....	34
Residência por Concelhos de Lisboa	35
Residência por Concelhos de Santarém.....	36
Residência por Concelhos de Setúbal	37
Residência fora de Portugal (Europeu/PLOP/Outro)	38
Residência em País Europeu	38

Residência em PLOP	39
Residência em Outro País (não Europeu e não PLOP)	40
Situação face ao emprego	40
Situação face ao Emprego	40
Situação face ao Emprego 2	41
Situação face ao Emprego com estudante e reformado.....	42
Taxa de emprego.....	43
Local de emprego	44
Emprego (Portugal/Europeu/PLOP/Outro) (variável auxiliar)	45
Emprego (Portugal/Outro País).....	46
Localização do emprego, por Distrito	46
Localização do Emprego (AML) (variável auxiliar)	47
Emprego distrito Lisboa/Setúbal/Outro e Estrangeiro (variável auxiliar).....	48
Emprego por Concelhos de Lisboa	49
Emprego por Concelhos de Santarém.....	50
Emprego por Concelhos de Setúbal	51
Emprego fora de Portugal	52
Emprego em País Europeu	53
Emprego em PLOP.....	54
Emprego em Outro País (não Europeu e não PLOP)	55
Países do emprego no estrangeiro (variável auxiliar)	56
Antes de ir para o estrangeiro, procurou trabalho em Portugal? (Se não, qual a razão?)	57
Procurou emprego antes de ir para o estrangeiro?	57
Razão 1	58
Razão 2	59
Razão 3	59
Razão 4	60
Razão 5	60
Razão 6	61
Outra razão.....	61
Outra razão, qual?.....	62
Entidade empregadora.....	63
Tipologia	63
Dimensão.....	64
Vínculo.....	65
Tipo de vínculo contratual.....	65

Outro tipo de vínculo contratual.....	66
Volume de faturação (Empresário)	67
Horário	68
Vencimento	69
Vencimento com outliers	69
Vencimentos sem outliers.....	69
Complementos remuneratórios.....	70
Complementos remuneratórios (Sim/Não)	70
Valor dos complementos remuneratórios	71
Valor dos complementos remuneratórios sem outliers	72
Duração de estágio ou bolsa (estagiário/bolseiro)	73
Duração do estágio.....	73
Duração da bolsa	74
Tipo de bolsa (bolseiros)	75
Tipo de bolsa	75
Outro tipo de bolsa	76
Bolsa para efeito de estudo	76
Alguma vez já exerceu atividade profissional remunerada na área do curso? (SAPR)	77
Emprego na área de curso	78
Emprego na área (Sim/Não).....	78
Emprego na área (para cálculo da taxa).....	79
Mesmo emprego antes e depois de terminar o curso (benefícios do curso)	79
Mantém mesmo emprego depois de concluir o curso	79
Melhoria 1	80
Melhoria 2	81
Melhoria 3	82
Melhoria 4	83
Melhoria 5	84
Melhoria 6	85
Outra melhoria	86
Outra melhoria, qual?	87
Se já esteve empregado após conclusão do curso (SAPR)	89
Primeiro emprego	90
Tempo de obtenção de emprego.....	91
Tempo entre a conclusão do curso e o atual emprego.....	91
Tempo entre a conclusão do curso e o primeiro emprego	92

Tempo decorrido entre a conclusão do curso e o primeiro emprego	93
Tempo decorrido entre conclusão do curso e o primeiro emprego	94
Continuação dos estudos	95
Intenção.....	96
Nível de formação	97
Tipo de nível de formação.....	97
Outro tipo de nível de formação.....	98
Instituição.....	101
Instituições de Ensino Superior (IES) de prosseguimento de estudos	101
Mesma Faculdade/Instituição da diplomação inicial.....	103
Pública/Privada	103
Instituição pública	104
Universitário, Politécnico ou Militar	105
Universidade pública.....	105
Escola ULisboa.....	106
Politécnico público	107
Instituição Militar pública	108
Instituição privada.....	109
Universitário ou Politécnico	109
Universidade privada	109
Politécnico privado.....	110
Outra Instituição.....	110
Área de estudos.....	111
Área de estudos do curso em que prosseguiu estudos	111
Outra área de estudos do curso em que prosseguiu estudos.....	112
Emprego e estudos de diplomados SAPR	114
Procura de emprego em diplomados SAPR	115
Procura emprego.....	115
Motivo principal para não procurar emprego	116
Outro motivo principal para além dos listados.....	116
Continuação de estudos nos diplomados SAPR.....	117
Diplomado SAPR deixar de estudar.....	118
Situação face ao emprego (distinguir nos SAPR os estudantes e reformados)	119
Situação face ao emprego (estudante detalhado).....	120
Tabelas completas (com muitas entradas)	121

Análise dos dados

Esta análise é sobre os resultados da 3ª edição do inquérito à empregabilidade da Universidade de Lisboa, do qual 3.478 diplomados de licenciatura e de mestrado (integrado e de 2º ciclo) das diversas escolas que concluíram os seus cursos no ano letivo 2014/15.

Este inquérito esteve disponível entre 23 de maio a 28 de dezembro de 2017 na maioria das escolas.

Após o inquérito, foi gerado uma tabela com 116 colunas que correspondem aos resultados dos inquéritos.

Todas as variáveis dos dados adquiridos estão listadas de uma forma generalizada e agrupada para uma melhor compressão.

Existem variáveis específicas para um certo grupo de diplomados, estes grupos estão apresentados com cor encarnada ao longo da lista.

- ID do respondente (não é número de aluno)
- ID serviços académicos
- Nacionalidade
 - Europa
 - Lista de países Europeus
 - PLOP
 - Lista de PLOP
 - Portugal
 - Outra
 - Outros países
- Internacionalização
- Unidade orgânica de formação (dentro da UL)
- Grau de formação
- Código de curso
- Classificação final do curso (10 a 20)
- Classificação Nacional de Áreas de Educação e Formação
 - CNAEF 1
 - CNAEF 2
 - CNAEF 3
- Designação do curso
- Área de estudo
- Data de nascimento
- Idade à data de fim do curso/da resposta ao inquérito
- Género
- Local de residência
 - País
 - Europeu
 - PLOP
 - Portugal
 - Outros

- Área Metropolitana de Lisboa
 - Fora
 - Dentro
- Distrito de Portugal
- Concelhos
 - Lisboa
 - Santarém
 - Setúbal
- Situação face ao emprego (várias categorizações)
- Taxa de emprego (Diplomado com/sem atividade profissional remunerada, estudante ou reformado)
- Local de emprego
 - País
 - Europeu
 - PLOP
 - Portugal
 - Outros
 - Área Metropolitana de Lisboa
 - Fora
 - Dentro
 - Distrito de Portugal
 - Concelhos
 - Lisboa
 - Santarém
 - Setúbal
- Antes de ir para o estrangeiro, procurou trabalho em Portugal? (Qual a razão?)
- Entidade empregadora (tipologia, dimensão)
- Volume de faturação
 - **Empresário**
- Vínculo contratual, horário de trabalho, salário bruto, complementos remuneratórios
- Estágios e bolsas (duração, tipo de bolsa, remuneração)
 - **Estagiário**
 - **Bolseiro**
- Emprego na área do curso
- Já exerceu atividade profissional remunerada na área do curso?
 - **SAPR**
- Mesmo emprego antes e depois de terminar o curso (benefícios do curso com lista de opções e resposta livre)
- Se já esteve empregado após conclusão do curso
 - **SAPR**
- Primeiro emprego (antes, durante, depois do curso, tempo de obtenção)
- Continuação dos estudos (intenção, nível de formação adicional, escola, área)
 - Pública
 - Universitário
 - Lista de IES pública universitária
 - Universidade de Lisboa
 - Lista de Escolas da ULisboa
 - Politécnico

- Lista de IES pública Politécnico
 - Militar
 - Lista de IES militares
 - Outra
 - Privada
 - Universitário
 - Lista de IES privado universitário
 - Politécnico
 - Lista de IES privado politécnico
 - Outra
- Procura de emprego (motivos para não procurar)
 - SAPR
 - Prosseguiu ou está a prosseguir estudos?
 - SAPR
 - Deixar de estudar se encontrar emprego
 - SAPR
-
- Sem Atividade Profissional Remunerada (SAPR)
 - Empresário (Trabalhador por conta própria com funcionários a cargo)

ID do respondente (ID interno e não de aluno)

Este ID é associado internamente ao estudante diplomado não correspondendo ao seu número de aluno. Os ids têm o mínimo 1 e máximo 3513 mas no total são 3478.

Tabela 1: Estatística descritiva dos IDs dos respondentes

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
ID do diplomado atribuído internamente	3478	1	3513	1744,65	1011,678
N válido (de lista)	3478				

ID serviços académicos

Esta variável corresponde ao identificador dos serviços académicos (Escolas com expedição autónoma).

O ID mínimo é 1 e o máximo é 79349 tendo 1738 entradas em branco.

Tabela 2: Estatística descritiva dos IDs serviços académicos

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
ID Serviços Académicos (Escolas com expedição autónoma)	1740	1	79349	23206,30	27553,944
N válido (de lista)	1740				

Tabela 3: Total de campos preenchidos e campos em branco

ID Serviços Académicos (Escolas com expedição autónoma)

N	Válido	1740
	Ausente	1738

Nacionalidade

Neste grupo existem 2 variáveis onde uma estuda a nacionalidade do diplomado e a outra agrupa estas nacionalidades em grupos de interesse.

Nacionalidade

A variável correspondente à nacionalidade do diplomado contém 28 nacionalidades como resposta. Os 'nr' correspondem aos inquiridos que não responderam à questão.

Grande parte dos inquiridos têm nacionalidade portuguesa, seguido de pessoas que não responderam e depois indivíduos com nacionalidade brasileira.

Tabela 4: Frequência de diplomados por nacionalidade

Nacionalidade		Frequência
Válido	Argentina	1
	Bélgica	1
	Camarões	1
	China	1
	Dupla Nacionalidade (PT/AUSTRALIA)	1
	Dupla Nacionalidade (PT/FR)	1
	Equador	1
	Eslovénia	1
	Espanha	1
	Guiné-Bissau	1
	Hungria	1
	Irão (República Islâmica)	1
	Kosovo	1
	México	1
	Nigéria	1
	São Tomé e Príncipe	1
	Suíça	1
	Guiné	2
	India	2
	Moçambique	2
	Roménia	2
	Ucrânia	2
	Alemanha	3
	Cabo Verde	7
	Itália	8
	Angola	10
	Brasil	23

nr	45
Portugal	3355
Total	3478

Nacionalidade (PT/Europeu/PLOP/Outro)

A outra variável agrupa as nacionalidades dos inquiridos em 4 grupos diferentes. Sendo os grupos constituídos por Portugal, Europa, PLOP (Países de Língua Oficial Portuguesa) e Outra.

Com a tabela a baixo é possível observar que 3433 são valores válidos e 45 não têm resposta.

A maioria têm nacionalidade portuguesa seguido do grupo PLOP.

Tabela 5: Total de valores válidos e ausentes

Estatísticas		
Nacionalidade (PT/Europeu/PLOP/Outro)		
N	Válido	3433
	Ausente	45

Tabela 6: Frequência de diplomados por grupos de nacionalidades

Nacionalidade (PT/Europeu/PLOP/Outro)		
		Frequência
Válido	Outra	11
	Europa	21
	PLOP	44
	Portugal	3357
	Total	3433
Ausente	nr	45
Total		3478

Existe uma possível hierarquia de agrupamento de nacionalidades para as nacionalidades detalhadas.

Internacionalização

Este grupo tem 2 variáveis onde uma apresenta a internacionalização dos diplomados e outra também indica a internacionalização dos diplomados mas discriminada, ambas as variáveis são dadas como variáveis auxiliares.

Internacionalização dos diplomados (variável auxiliar)

A primeira variável apresenta 4 tipos de internacionalização, onde 2336 são dados válidos e 1142 não são.

Grande parte dos diplomados são portugueses a trabalhar em Portugal o que é espectável, seguido de 231 portugueses que foram trabalhar para o estrangeiro.

Tabela 7: Total de dados válidos e ausentes na variável

Estatísticas		
Internacionalização dos diplomados (variável auxiliar)		
N	Válido	2336
	Ausente	1142

Tabela 8: Frequência de diplomados por internacionalização

Internacionalização dos diplomados (variável auxiliar)		Frequência
Válido	Estrangeiro que ficou a trabalhar em Portugal	29
	Estrangeiro que não ficou a trabalhar em Portugal	30
	Português que foi trabalhar para o estrangeiro	231
	Português a trabalhar em Portugal	2046
	Total	2336
Ausente	na	935
	nr	207
	Total	1142
Total		3478

Internacionalização dos diplomados/discriminada (variável auxiliar)

A segunda variável deste grupo também é relacionada com a internacionalização do diplomado mas discriminada.

A tabela a baixo indica o número de campos preenchidos ou não. Na mesma é possível observar que grande parte dos diplomados são portugueses a trabalhar em Portugal seguido de portugueses que foram trabalhar para a Europa.

Tabela 9: Total de valores válidos e ausentes da internacionalização discriminada

Estatísticas

Internacionalização dos diplomados/discriminada (variável auxiliar)

N	Válido	2333
	Ausente	1145

Tabela 10: Frequência de diplomados por internacionalização discriminada

**Internacionalização dos diplomados/discriminada
(variável auxiliar)**

		Frequência
Válido	Outro estrangeiro que não ficou a trabalhar em Portugal	3
	Outro estrangeiro que ficou a trabalhar em Portugal	5
	Europeu que ficou a trabalhar em Portugal	7
	Europeu que não ficou a trabalhar em Portugal	7
	Português que foi trabalhar para um PLOP	15
	PLOP que ficou a trabalhar em Portugal	17
	PLOP que não ficou a trabalhar em Portugal	20
	Português que foi trabalhar para outro país	24
	Português que foi trabalhar para a Europa	189
	Português a trabalhar em Portugal	2046
	Total	2333
Ausente	na	935
	nr	210
	Total	1145
Total		3478

Unidade orgânica de formação (dentro da ULisboa)

Neste grupo estão presentes 2 variáveis que correspondem à escola onde o diplomado concluiu o seu curso. A primeira apresenta as siglas das escolas e a segunda apenas associa um ID a cada uma destas siglas.

Todas as linhas foram preenchidas com a escola onde cada diplomado se graduou, e com a tabela a baixo é possível observar que grande parte dos inquiridos concluíram os seus cursos no Instituto Superior Técnico (IST), seguido da Faculdade de Ciências. As escolas onde se verificaram menor número de diplomados neste inquérito foi na Reitoria, seguido da Faculdade de Medicina Dentária.

Tabela 11: Total de valores válidos e ausentes na escola

Estatísticas		
Escola		
N	Válido	3478
	Ausente	0

Tabela 12: Frequência de diplomados por escolas

Escola		Frequência
Válido	Reitoria	16
	FMD	40
	IGOT	51
	FMV	73
	FP	90
	IE	93
	FBA	110
	FA	114
	FMH	150
	FD	170
	FF	181
	ISA	198
	FM	203
	FL	211
	ISCSP	226
	ISEG	327
	FC	390
IST	835	
Total	3478	

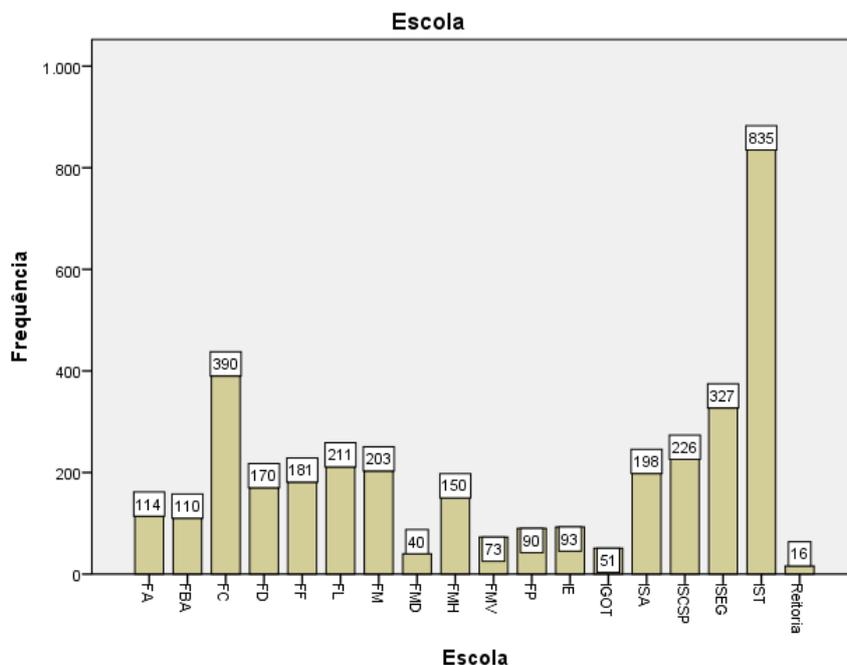


Figura 1: Gráfico de barras com a frequência de cada escola

Grau de formação

Esta variável estudo o grau do diplomado, isto é, se tem licenciatura, mestrado integrado ou mestrado 2º ciclo.

Todas as linhas da coluna foram preenchidas e grande parte dos diplomados inquiridos têm o grau de licenciatura.

Tabela 13: Total de valores válidos e ausentes no grau de formação

Estatísticas		
Grau		
N	Válido	3478
	Ausente	0

Tabela 14: Frequência de diplomados por grau de formação

Grau		Frequência
Válido	Mestrado (2º ciclo)	841
	Mestrado Integrado	964
	Licenciatura	1673
	Total	3478

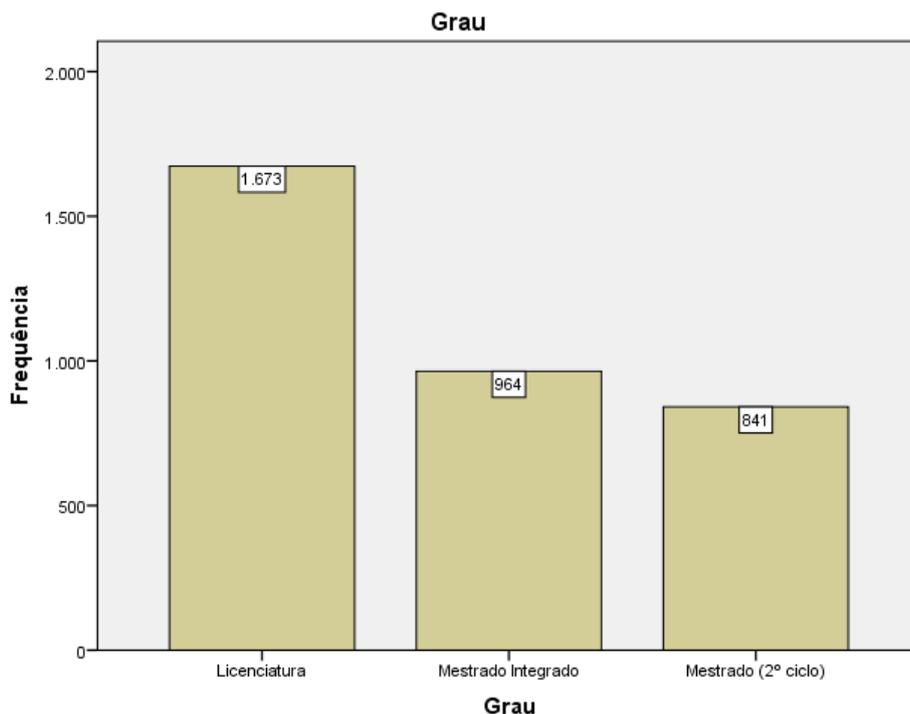


Figura 2: Gráfico de frequências por grau de formação

Código de curso

Esta variável corresponde ao código do curso e a tabela a baixo indica com a frequência de diplomados em cada código de curso.

6008	6011	6012	6030	6031	6038	6057	6062	6067	6077	6078	6080	6081	6093	6095	6097	6106	6111	6125	6127	6150	
1	3	3	11	14	1	3	3	4	3	5	12	6	9	6	1	1	1	4	1	6	
6449	6451	6452	6453	6454	6455	6457	6485	6487	6620	6631	6633	6653	6676	6679	6680	6681	6688	6689	6698	6721	
4	7	4	1	21	2	6	1	8	3	4	1	1	1	1	4	12	1	2	2	3	
8014	8102	8104	8106	8109	8111	8258	8358	8362	8363	8364	8377	8411	8413	8415	8434	8458	9002	9006	9011	9015	
12	9	8	6	7	38	10	12	1	6	11	4	9	10	5	3	7	23	6	69	34	
9133	9135	9139	9140	9141	9143	9146	9147	9162	9181	9182	9204	9209	9210	9212	9223	9226	9229	9238	9240	9250	
2	14	7	9	9	29	20	70	11	16	11	35	3	12	4	6	6	19	24	12	11	
9396	9402	9409	9410	9417	9442	9448	9449	9494	9518	9520	9548	9555	9556	9567	9657	9665	9707	9753	9754	9790	
10	16	6	5	9	6	7	10	162	2	13	15	86	13	27	1	3	46	2	12	16	
M147	M382	M385	M418	M437	M469	M478	M479	M533	M596	M629	M677	M699	M780								
5	3	5	2	1	2	4	4	1	4	3	1	1	1								
6151	6153	6157	6158	6166	6180	6193	6199	6234	6244	6255	6259	6264	6267	6340	6378	6391	6425	6426	6436	6437	6438
14	2	2	3	1	1	5	13	1	2	2	4	4	5	1	2	11	3	2	5	12	12
6723	6724	6726	6728	6730	6731	6745	6746	6747	6749	6766	6786	6787	6876	6888	6891	6933	6935	6941	6955	6961	6998
1	2	1	1	7	2	1	1	3	1	3	1	4	2	10	4	1	2	3	4	5	34
9019	9023	9025	9026	9040	9068	9069	9070	9071	9072	9078	9081	9086	9087	9098	9099	9104	9115	9119	9121	9129	9131
15	17	15	24	7	7	17	21	12	17	155	50	25	39	21	12	58	8	27	183	8	7
9252	9257	9265	9267	9279	9280	9281	9286	9294	9299	9300	9311	9314	9315	9329	9345	9368	9381	9385	9388	9390	9393
15	48	7	4	3	1	19	9	14	19	13	2	24	1	3	17	1	5	20	4	5	13
9791	9811	9813	9822	9841	9845	9847	9885	9903	9904	9911	9912	9913	9914	9917	A001	G004	G008	M112	M117	M118	M131
12	31	162	16	32	15	68	3	7	21	13	13	14	13	7	3	10	10	4	2	2	1

Figura 3: Frequência por cada código de curso

Classificação final de curso

Esta variável corresponde à classificação final de curso do diplomado e deve estar no intervalo entre 10 e 20 valores.

É possível observar que 3437 das linhas foram preenchidas, 39 foram preenchidas com não respondeu e 2 não têm qualquer entrada, sendo considerado como sistema (possivelmente um erro a corrigir).

Outro erro possível é a classificação máxima, nas tabelas abaixo indicam ser de 27 valores, do qual, deveria ser 20 ou 17. (O que fazer neste situação?).

Em média os diplomados tiveram uma classificação de 14.53, onde o mínimo foi 10 valores e o máximo foi 19 se excluirmos o 27.

A frequência mais alta na classificação final é de 15 valores com 714 diplomados e a mais baixo é de 10 valores só tendo 1 diplomado.

Tabela 15: Total de valores válidos e ausentes na classificação final

Estatísticas		
Classificação final de curso		
N	Válido	3437
	Ausente	41

Tabela 16: Frequência por classificação final de curso

Classificação final de curso		Frequência
Válido	10	1
	27	1
	19	20
	11	54
	18	142
	17	305
	12	384
	16	517
	13	594
	14	705
	15	714
	Total	3437
Ausente	nr	39
	Sistema	2
	Total	41
Total		3478

Tabela 17: Estatística descritiva da classificação final de curso

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Classificação final de curso	3437	10	27	14,53	1,718
N válido (de lista)	3437				

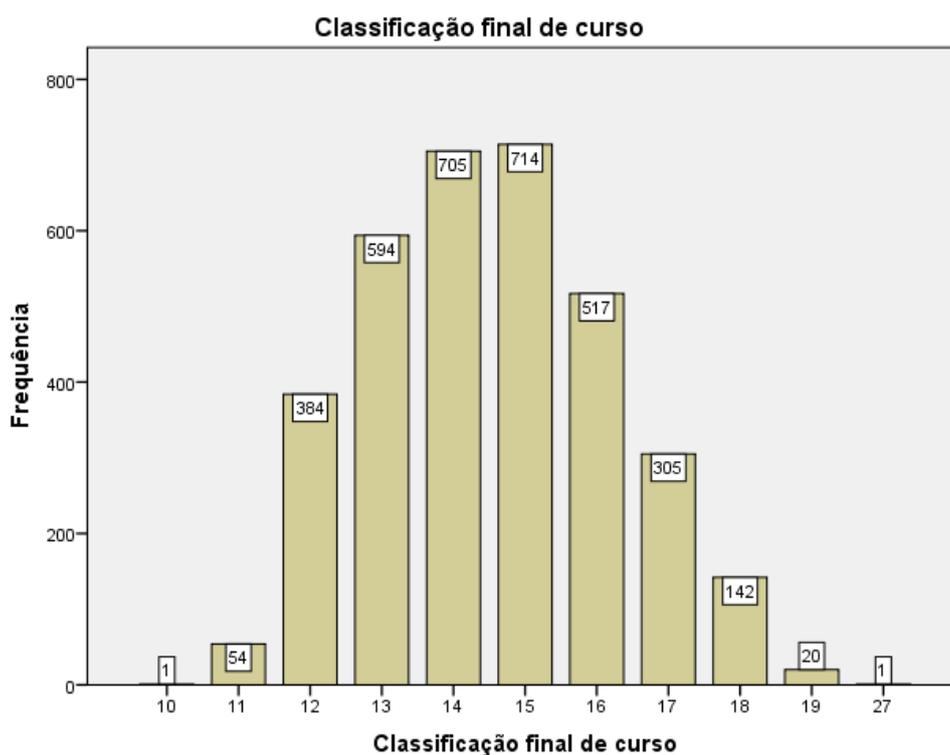


Figura 4: Gráfico das frequências por classificações finais de curso

Classificação Nacional de Áreas de Educação e Formação (CNAEF)

Este grupo tem 4 variáveis, a CNAEF 1, CNAEF 2, CNAEF 2 incluindo 311- Psicologia e 314 Economia e CNAEF 3. Estas variáveis podem constituir uma hierarquia pois cada uma destas classificações é mais detalhada que a outra.

CNAEF 1

A primeira variável corresponde a CNAEF 1 que tem a informação mais generalizada. Com estudo da tabela a baixo.

Todas as linhas foram preenchidas, e grande parte dos diplomados pertencem à área de Engenharia, indústrias transformadoras e construção, a área com menos diplomados é a de Serviços, excluindo o desconhecido ou não especificado.

Tabela 18: Total de valores válidos e ausentes na CNAEF 1

Estatísticas		
CNAEF 1		
N	Válido	3478
	Ausente	0

Tabela 19: Frequência por CNAEF 1

CNAEF 1		Frequência
Válido	9-Desconhecido ou não especificado	7
	8-Serviços	97
	1-Educação	119
	6-Agricultura	131
	2-Artes e Humanidades	309
	4-Ciência, matemática e informática	467
	7-Saúde e proteção social	506
	3-Ciência sociais, comércio e direito	810
	5-Engenharia, indústrias transformadoras e construção	1032
	Total	3478

CNAEF 2

A segunda variável corresponde à CNAEF 2, onde esta é mais detalhada que CNAEF 1.

Através da tabela com as frequências a área com mais diplomados é a Engenharia e técnicas afins com 710 pessoas. A área com menos diplomados é a Serviços de transporte com 1 pessoa.

Tabela 20: Total de valores válidos e ausentes de CNAEF 2

Estatísticas		
CNAEF 2		
N	Válido	3478
	Ausente	0

Tabela 21: Frequência de diplomados por CNAEF 2

CNAEF 2		Frequência
Válido	84-Serviços de transporte	1
	99-Desconhecido ou não especificado	7
	86-Serviços de segurança	8
	32-Informação e jornalismo	25
	85-Protecção do ambiente	29
	48-Informática	32
	76-Serviços sociais	38
	81-Serviços Pessoais	59
	62-Agricultura, silvicultura e pescas	63
	64-Ciências veterinárias	68
	46-Matemática e estatística	87
	54-Indústrias transformadoras	97
	14-Formação de professores/formadores e ciências da educa	119
	22-Humanidades	124
	44-Ciências físicas	130
	38-Direito	170
	21-Artes	185
	42-Ciências da vida	218
	58-Arquitectura e construção	225
	34-Ciências empresariais	290
	31-Ciências sociais e do comportamento	325
	72-Saúde	468
	52-Engenharia e técnicas afins	710
Total	3478	

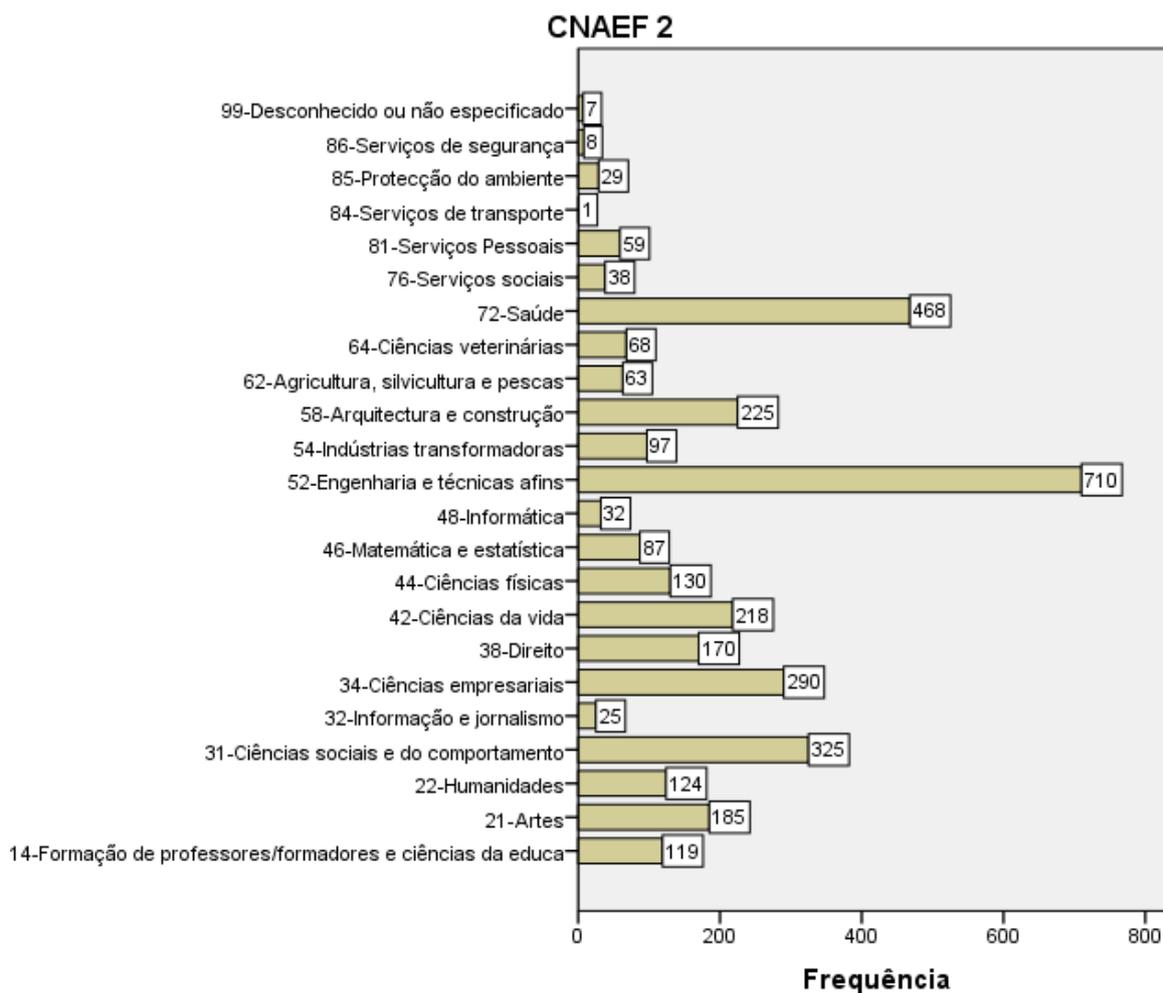


Figura 5: Gráfico de diplomados por CNAEF 2

CNAEF 2 (com 311-Psicologia e 314-Economia)

A variável seguinte apenas acrescenta a área de Psicologia e de Economia à CNAEF 2 separando os diplomados entre essas duas áreas e a de Ciências sociais e do comportamento.

Tabela 22: Frequência por CNAEF 2 com 311 e 314

CNAEF 2 (com 311-Psicologia e 314-Economia)		Frequência
Válido	84-Serviços de transporte	1
	99-Desconhecido ou não especificado	7
	86-Serviços de segurança	8
	32-Informação e jornalismo	25
	85-Protecção do ambiente	29
	48-Informática	32

76-Serviços sociais	38
81-Serviços Pessoais	59
62-Agricultura, silvicultura e pescas	63
64-Ciências veterinárias	68
46-Matemática e estatística	87
311-Psicologia	90
54-Indústrias transformadoras	97
314-Economia	99
14-Formação de professores/formadores e ciências da educação	119
22-Humanidades	124
44-Ciências físicas	130
31-Ciências sociais e do comportamento	136
38-Direito	170
21-Artes	185
42-Ciências da vida	218
58-Arquitectura e construção	225
34-Ciências empresariais	290
72-Saúde	468
52-Engenharia e técnicas afins	710
Total	3478

CNAEF 3

A variável seguinte corresponde à CNAEF 3 e esta é mais detalhada que CNAEF 2.

Com a tabela a baixo é possível observar que a área com mais diplomados é a de Eletrónica e automação sendo 408 dos inquiridos, seguindo da área de Gestão e administração com 223 diplomados. As áreas com menos diplomados são a Serviços de transporte e Turismo e lazer com apenas 1 diplomado cada.

Tabela 23: Freqüência por CNAEF 3

CNAEF 3		Freqüência
Válido	812-Turismo e lazer	1
	840-Serviços de transporte	1
	863-Segurança militar	2
	322-Biblioteconomia, arquivo e documentação (bad)	3
	543-Materiais (indústrias da madeira, cortiça, papel, plástico, vidro e outr	3
	723-Enfermagem	4
	725-Tecnologias de diagnóstico e terapêutica	6
	862-Segurança e higiene no trabalho	6
	226-Filosofia e ética	7
	623-Silvicultura e caça	7
	999-Desconhecido ou não especificado	7
	220-Humanidades	10
	422-Ciências do ambiente	11
	344-Contabilidade e fiscalidade	13
	442-Química	13
	729-Saúde - programas não classificados noutra área de formação	16
	223-Língua e literatura materna	18
	321-Jornalismo e reportagem	22
	212-Artes do espetáculo	23
	342--Marketing e publicidade	24
146-Formação de professores e formadores de áreas tecnológicas	25	

462-Estatística	25
544-Indústrias extrativas	28
851-Tecnologia de proteção do ambiente	29
343-Finanças, banca e seguros	30
145-Formação de professores de áreas disciplinares específicas	32
441-Física	32
481-Ciências informáticas	32
522-Electricidade e energia	32
225-História e arqueologia	35
762-Trabalho social e orientação	38
724-Ciências dentárias	40
525-Construção e reparação de veículos a motor	41
521-Metalurgia e metalomecânica	47
726-Terapia e reabilitação	51
213-Áudio-visuais e produção dos media	52
214-Design	52
222-Línguas e literaturas estrangeiras	54
621-Produção agrícola e animal	56
211-Belas-artes	58
313-Ciência política e cidadania	58
813-Desporto	58
142-Ciências da Educação	62
461-Matemática	62
541-Indústrias alimentares	66
640-Ciências veterinárias	68

529-Eng. e técnicas afins - programas não class. Noutra área de formação	70
582-Construção civil e engenharia civil	77
312-Sociologia e outros estudos	78
443-Ciências da terra	85
311-Psicologia	90
314-Economia	99
524-Tecnologia dos processos químicos	112
581-Arquitectura e urbanismo	148
727-Ciências farmacêuticas	169
380-Direito	170
721-Medicina	182
421-Biologia e bioquímica	207
345-Gestão e administração	223
523-Electrónica e automação	408
Total	3478

Designação do curso

Esta variável corresponde à designação do curso do diplomado. Foi verificado quantas designações haviam disponíveis, tendo sido observados 226 no total.

```
> teste <- unique(emprega$Curso_Designação)
> length(teste)
[1] 226
```

Figura 6: Comando e total de designações de curso

A designação com mais inquiridos foi a de Engenharia informática e de computadores com 183 pessoas.

Tabela 24: Frequência mais alta por designação do curso

Designação do Curso com frequência mais alta		Frequência
Válido	Psicologia	86
	Direito	158
	Ciências Farmacêuticas	162
	Medicina	162
	Engenharia Informática e de Computadores	183

Para consultar a tabela completa, abrir o ficheiro “Tabelas completas” (Tabelas completas (com muitas entradas)).

Área de estudo

Esta variável representa a área de estudo na ULisboa. Todos os dados foram preenchidos, a área de estudo com mais diplomados é a Engenharia excluindo informática, com 913 dos inquiridos. A área de estudo com menos inquiridos tirando o valor desconhecido ou não específico é de informática com apenas 32 pessoas.

Tabela 25: Frequência por áreas de estudo ULisboa

Áreas de Estudo ULisboa		Frequência
Válido	Desconhecido ou não especificado	7
	Informática	32
	Agricultura, Silvicultura e Pescas	63
	Serviços de Desporto, Turismo e Segurança	66
	Veterinária	68
	Matemática e Estatística	87
	Psicologia	90
	Saúde (excl. Medicina)	117
	Educação	119
	Humanidades	124
	Ciências Físicas e da Terra	130
	Arquitetura	148
	Farmácia	169

Direito	170
Medicina	182
Artes	185
Ciências Sociais	201
Ciências da Vida	218
Economia e Gestão	389
Engenharia (excl. Informática)	913
Total	3478

Data de nascimento

Esta variável corresponde à data de nascimento do inquirido, foram verificadas 1819 datas diferentes. Dessas 1819, existem 6 diplomados com a mesma data de nascimento, sendo esta 10-05-1994. Do total de inquiridos 2643 responderam a esta pergunta e 835 não.

Tabela 26: Frequência mais alta por data de nascimento

Data de nascimento com frequência mais alta

		Frequência
Válido	31-07-1991	4
	01-03-1994	5
	07-10-1992	5
	13-07-1991	5
	16-05-1991	5
	23-09-1993	5
	25-06-1992	5
	28-11-1994	5
	28-12-1992	5
	31-03-1991	5
	10-05-1994	6
		835
Total	3478	

Para consultar a tabela completa, abrir o ficheiro “Tabelas completas” (Tabelas completas (com muitas entradas)).

Idade à data de fim do curso/da resposta ao inquérito

Este grupo contém 2 variáveis, onde a primeira estuda a idade até à data da diplomação e a segunda estuda a idade do diplomado quando respondeu ao inquérito.

Idade à data da diplomação

A variável que indica a idade até à data da diplomação do estudante apresenta 3448 valores válidos e 30 sem resposta.

Dos 3448 inquiridos que responderam, 624 diplomaram-se com 23 anos, seguido de 594 que se diplomaram com 24 anos. A idade mais baixa de diplomação foi com 14 anos (possivelmente um erro), o inquirido mais velho a obter a sua diplomação tem 82 anos, mas em média as pessoas têm a sua diplomação aos 25 anos.

Tabela 27: Estatística descritiva da idade à data da diplomação

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Idade à data da diplomação	3448	14,00	82,00	25,6505	5,83985
N válido (de lista)	3448				

Tabela 28: Frequência por idade até à data da diplomação

Idade à data da diplomação		
		Frequência
Válido	14,00	1
	48,00	1
	63,00	1
	67,00	1
	68,00	1
	82,00	1
	46,00	2
	57,00	2
	58,00	2
	61,00	2
	20,00	5
	50,00	5
	53,00	5
	54,00	5
	59,00	5

	45,00	6
	55,00	6
	49,00	7
	51,00	7
	47,00	9
	52,00	11
	39,00	12
	41,00	13
	42,00	13
	44,00	13
	40,00	15
	37,00	16
	43,00	16
	33,00	17
	35,00	17
	36,00	20
	38,00	27
	32,00	31
	34,00	32
	31,00	40
	30,00	57
	29,00	64
	28,00	88
	27,00	137
	26,00	268
	21,00	274
	25,00	484
	22,00	491
	24,00	594
	23,00	624
	Total	3448
Ausente	nr	30
Total		3478

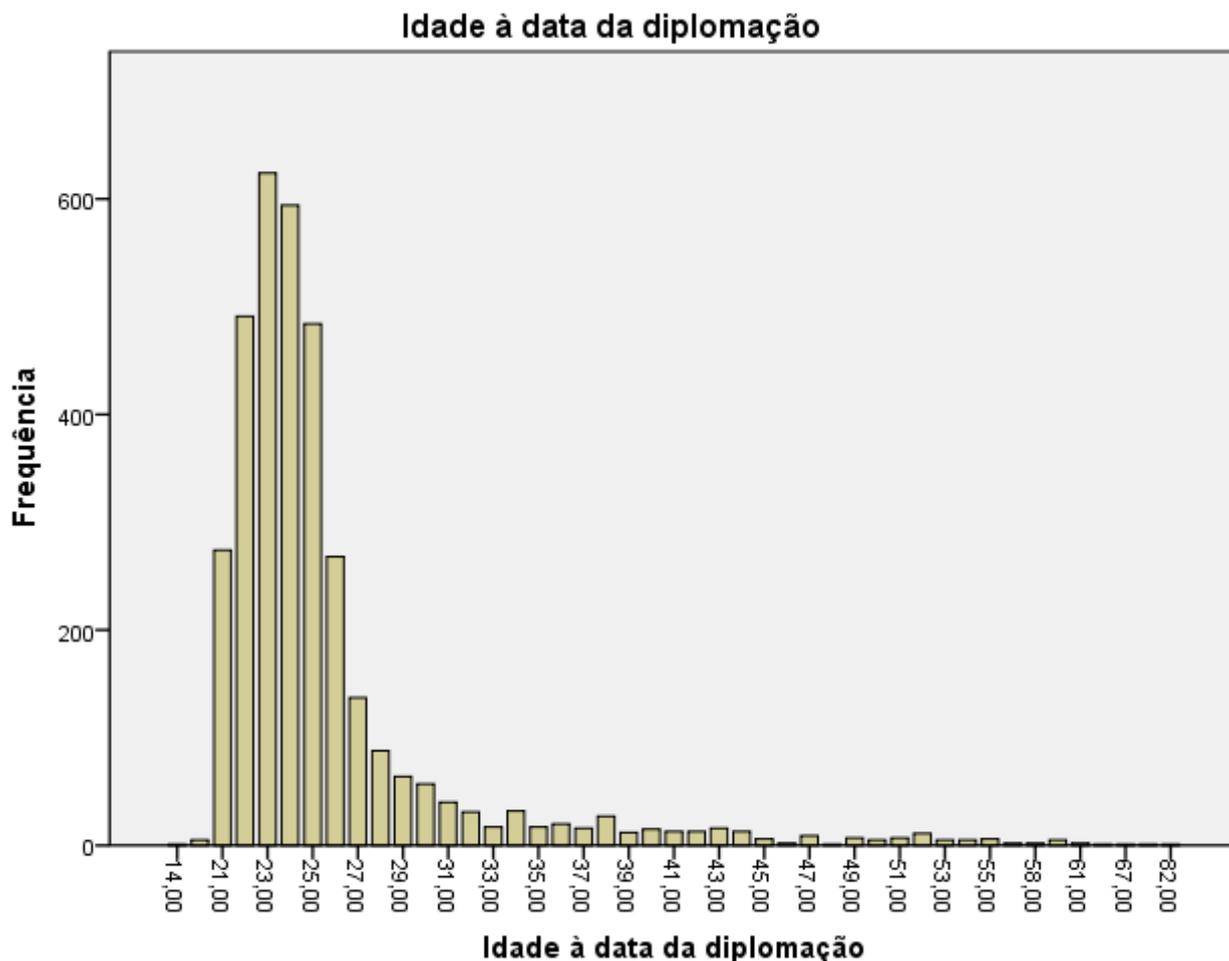


Figura 7: Gráfico com a frequência por idade à data da diplomação

Idade à data da resposta

De seguida é apresentado a idade do inquirido à data da resposta, dos 3478 inquiridos apenas 1041 têm dados válidos. Onde a idade mais baixa é de 16 anos (não sei se é erro ou não) e a máxima é de 63 anos. Em média a idade à data da resposta é de 26 anos.

No entanto acontece em muitos casos terem a mesma idade quando se diplomaram e quando responderam a este inquérito (inconsistência). Sabendo que este inquérito esteve disponível em 2017 para diplomados no ano letivo 2014/15.

Tabela 29: Estatística descritiva da idade à data da resposta

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Idade à data da resposta	1011	16	63	25,98	4,634
N válido (de lista)	1011				

Tabela 30: Frequência por idade à data da resposta

Idade à data da resposta		Frequência
Válido	16	1
	39	1
	47	1
	49	1
	54	1
	57	1
	63	1
	44	2
	51	2
	52	2
	35	3
	36	3
	40	4
	42	4
	20	5
	34	5
	38	5
	41	5
	43	6
	37	7
	32	8
	31	11
	33	13
	30	23
	29	32
	28	53
	22	66
	21	84
	23	84
	27	104
	24	134
26	153	
25	186	
Total	1011	

Ausente	nr	30
	Sistema	2437
	Total	2467
Total		3478

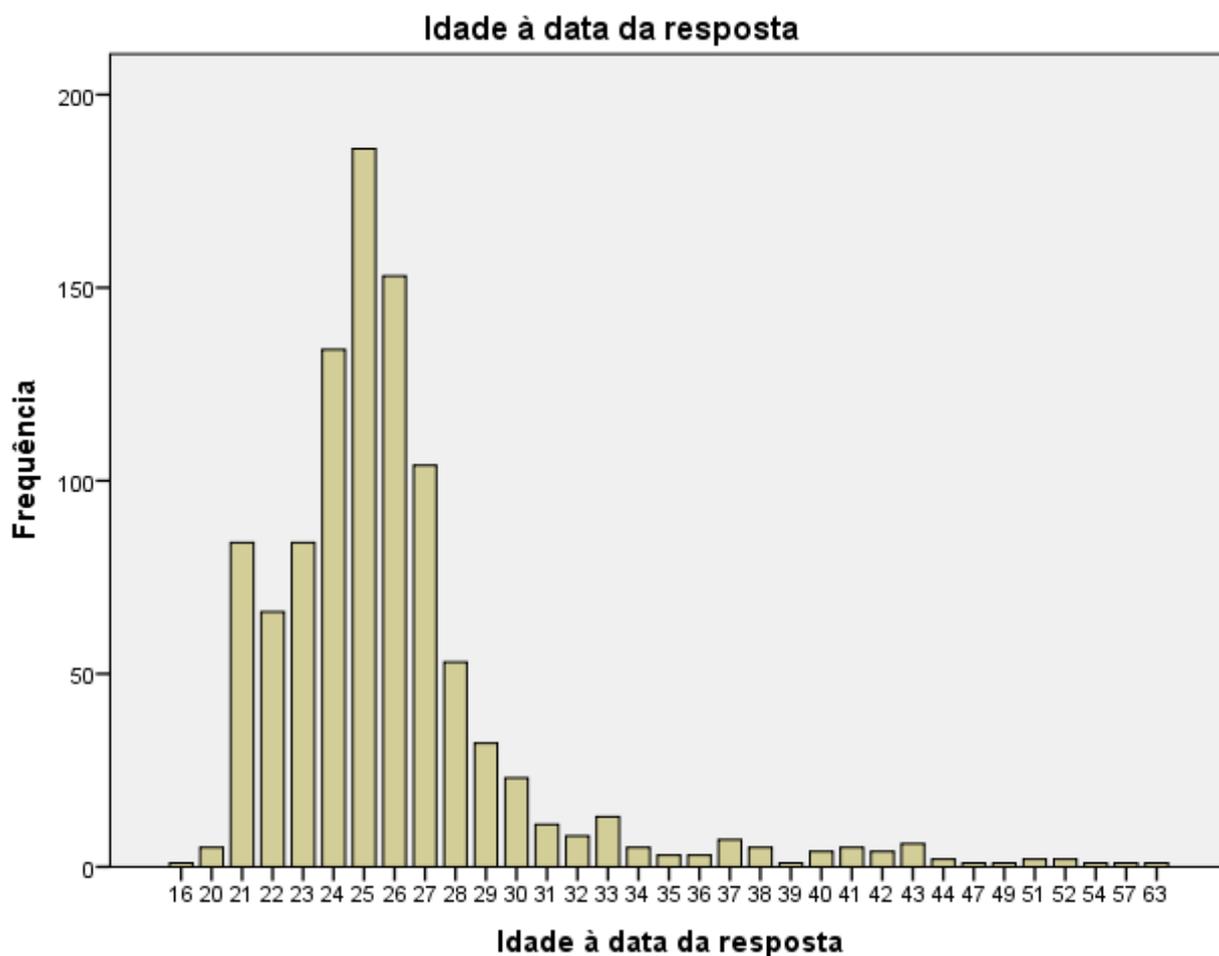


Figura 8: Gráfico de frequência por idade à data da resposta

Género

Nesta variável é apresentado o género dos diplomados, existem no total 3453 dados válidos. 1931 São do sexo feminino e 1522 do sexo masculino. O “999” é outra representação para os que não responderam.

Tabela 31: Frequência por género

Sexo		Frequência
Válido	M	1522
	F	1931
	Total	3453
Ausente	999	25
Total		3478

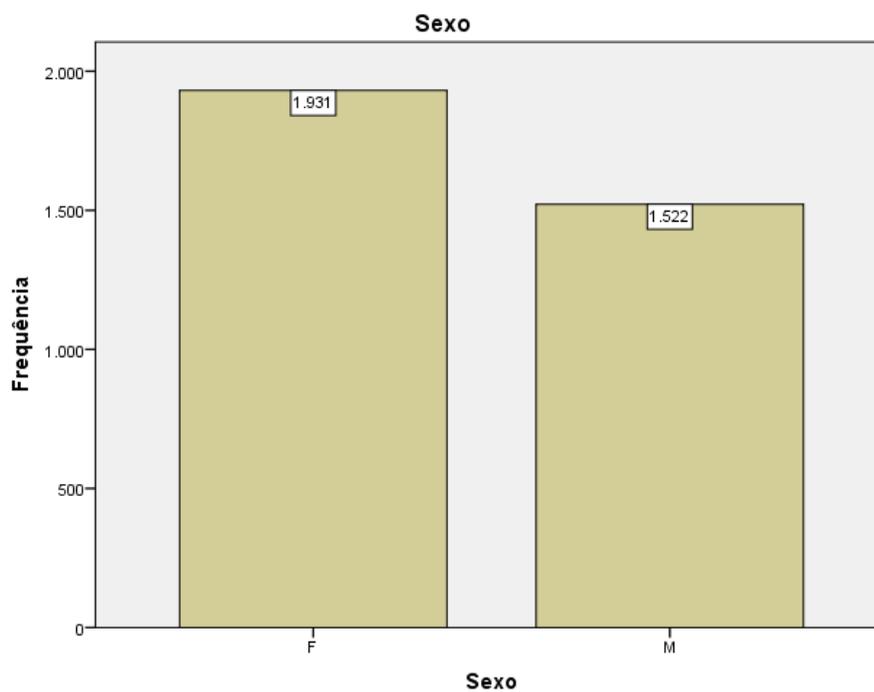


Figura 9: Gráfico da frequência por género

Local de residência

Este grupo tem 11 variáveis relacionadas com a localização da residência do diplomado. É possível criar hierarquias entre as variáveis.

- País
 - Europeu
 - PLOP
 - Portugal
 - Área Metropolitana de Lisboa
 - Fora
 - Dentro
 - Distrito de Portugal
 - Concelhos

- Lisboa
 - Santarém
 - Setúbal
- Outros

Residência (Portugal/Europeu/PLOP/Outro) (variável auxiliar)

A primeira variável corresponde à localização agrupada da residência do diplomado, ou seja se, reside em Portugal, Europa, Países de Língua Oficial Portuguesa ou outro.

Grande parte dos inquiridos reside em Portugal, seguido da Europa.

Tabela 32: Frequência por residência agrupada

Residência (Portugal/Europeu/PLOP/Outro) (variável auxiliar)		Frequência
Válido	Outro	28
	PLOP	35
	nr	88
	Europa	183
	Portugal	3144
	Total	3478

Residência (Portugal/Outro país)

A variável seguinte corresponde à residência, onde apenas diferencia entre Portugal e outro país.

Do total de inquiridos 3411 têm dados válidos e 67 não responderam. Grande parte dos inquiridos residem em Portugal e estes dados são consistentes com a variável anterior.

Tabela 33: Frequência por residência (Portugal/ Outro)

Residência (Portugal/Outro país)		Frequência
Válido	Outro país	267
	Portugal	3144
	Total	3411
Ausente	nr	67
Total		3478

Residência por Distrito de Portugal

A variável residência por distrito de Portugal indica quantos diplomados residem em cada distrito.

A maioria dos diplomados que residem em Portugal, residem no distrito de Lisboa. O distrito onde se verificou menos diplomados foi no distrito de Bragança.

Tabela 34: Frequência de residência por distrito de Portugal

Residência por Distrito de Portugal		Frequência
Válido	Bragança	2
	Guarda	5
	Viana do Castelo	6
	Vila Real	7
	Braga	11
	Viseu	11
	Portalegre	14
	Aveiro	17
	Castelo Branco	18
	Beja	19
	Coimbra	19
	Évora	23
	Madeira	32
	Açores	34
	Porto	36
	Faro	44
	Santarém	91
	Leiria	106
		319
	Setúbal	326
Lisboa	1989	
Total	3129	
Ausente	na	267
	nr	82
	Total	349
Total		3478

Localização da residência (AML) (variável auxiliar)

A variável seguinte corresponde à localização da residência tendo por base a Área Metropolitana de Lisboa (AML).

Dos dados válidos é possível observar que 1949 têm a residência na AML, seguido de 574 que têm a residência fora da AML e 267 no estrangeiro.

Tabela 35:Frequência por localização da residência (AML)

Localização da residência (AML)
(variável auxiliar)

		Frequência
Válido	Estrangeiro	267
	Fora da AML	574
	AML	1949
	Total	2790
Ausente	nr	369
	Sistema	319
	Total	688
Total		3478

Residência por Concelhos de Lisboa

Os resultados seguintes são relacionados com as residências pelos diferentes concelhos de Lisboa.

Através da frequência é possível observar que grande parte reside no concelho de Lisboa, seguido de Sintra. Cadaval é o concelho que tem o menor número de diplomados a residir.

O número de valores válidos e omissos é muito semelhante, dos valores omissos 319 são considerados de 'Sistema', possivelmente um erro a ser corrigido.

Tabela 36:Frequência de residências por concelhos de Lisboa

Residência por Concelhos de Lisboa

		Frequência
Válido	Cadaval	4
	Arruda dos Vinhos	5
	Azambuja	7
	Lourinhã	7
	Sobral de Monte Agraço	8
	Alenquer	11
	Torres Vedras	25
	Mafra	35
	Odivelas	61
	Vila Franca de Xira	65
	Amadora	94
	Loures	97
	Cascais	142

	Oeiras	152
	Sintra	202
	Lisboa	827
	Total	1742
Ausente	na	1087
	nr	330
	Sistema	319
	Total	1736
Total		3478

Residência por Concelhos de Santarém

De seguida é verificado quantos diplomados residem por cada concelho de Santarém.

Apenas 79 dos valores são válidos, 3402 são omissos ou porque não responderam à pergunta ou porque não fazem parte da amostra que reside no concelho de Santarém.

Pela estatística realizada, dos 76 diplomados que responderam, 16 residem no concelho de Santarém sendo este o valor mais alto, o concelho com menos residentes são Chamusca, Coruche, Mação e Sardoal.

Residência por Concelhos de Santarém

		Frequência
Válido	Chamusca	1
	Coruche	1
	Mação	1
	Sardoal	1
	Alcanena	2
	Alpiarça	2
	Entroncamento	2
	Ferreira do Zêzere	2
	Golegã	2
	Rio Maior	2
	Almeirim	3
	Cartaxo	4
	Abrantes	5
	Benavente	5
	Salvaterra de Magos	5
	Torres Novas	5
	Ourém	8
Tomar	9	
Santarém	16	

	Total	76
Ausente	na	2986
	nr	97
	Sistema	319
	Total	3402
Total		3478

Residência por Concelhos de Setúbal

Depois é verificado que concelhos de Setúbal é que os diplomados residem.

287 dos diplomados indicaram valores válidos para esta variável, e grande parte dos diplomados residem no concelho de Almada, seguido do Seixal. O concelho com menos residentes é Alcácer do Sal só com 1 diplomado.

Alcochete não está escrito corretamente, sendo necessário corrigir. Existem 319 resultados considerados de Sistema, podendo ser um erro para alterar.

Tabela 37: Frequência de residência por concelhos de Setúbal

Residência por Concelhos de Setúbal		Frequência
Válido	Alcácer do Sal	1
	Grândola	2
	Sines	4
	Santiago do Cacém	6
	Alcochete	7
	Sesimbra	13
	Palmela	16
	Moita	19
	Montijo	21
	Setúbal	35
	Barreiro	38
	Seixal	40
	Almada	85
	Total	287
	Ausente	na
nr		121
Sistema		319
Total		3191
Total		3478

Residência fora de Portugal (Europeu/PLOP/Outro)

A variável seguinte explora os diplomados que residem fora de Portugal, ou seja, se estes residem na Europa, PLOP ou outro local.

Dos 246 que têm valores válidos 183 indicam ter residência num país Europeu, seguindo de 35 que residem em PLOP e 28 que residem noutros países.

Grande parte dos valores são omissos porque grande parte da amostra reside em Portugal.

Tabela 38: Frequência por residências fora de Portugal

Residência fora de Portugal (Europeu/PLOP/Outro)		Frequência
Válido	Outro	28
	PLOP	35
	Europeu	183
	Total	246
Ausente	na	2825
	nr	88
	Sistema	319
	Total	3232
Total		3478

Residência em País Europeu

A variável seguinte explora os diplomados que residem num país Europeu.

Sobre um conjunto de países Europeus e de 165 respostas válidas, é possível observar que grande parte dos diplomados que residem num país Europeu, residem no Reino Unido. Os países europeus com menos diplomados a residir são Eslovénia, Islândia e Polónia com apenas um diplomado cada.

O valor Suíça não se encontra escrito corretamente sendo necessário corrigir.

Tabela 39: Frequência de residência por país Europeu

Residência em País Europeu		Frequência
Válido	Eslovénia	1
	Islândia	1
	Polónia	1
	Dinamarca	2
	Hungria	2
	Áustria	5

	Itália	5
	França	6
	Luxemburgo	7
	República da Irlanda	7
	Espanha	9
	Suíça	9
	Bélgica	10
	Suécia	10
	Alemanha	13
	Países Baixos (Holanda)	16
	Reino Unido	61
	Total	165
Ausente	na	2888
	nr	88
	Sistema	337
	Total	3313
Total		3478

Residência em PLOP

É explorado os diplomados residentes em Países de Língua Oficial Portuguesa, estes são menos que os que residem num país Europeu. São 35 os diplomados que respondem residir num PLOP. Destes 35, são 13 os que residem no Brasil, seguido de Angola. O PLOP com menos diplomados a residir é Timor-Leste.

Tabela 40: Frequência de residência por PLOP

Residência em PLOP		Frequência
Válido	Timor-Leste	1
	São Tomé e Príncipe	2
	Moçambique	3
	Cabo Verde	5
	Angola	11
	Brasil	13
	Total	35
	Ausente	na
nr		88
Sistema		319
Total		3443
Total		3478

Residência em Outro País (não Europeu e não PLOP)

Os diplomados que dizem ter residência em outro país para além de países Europeus ou de Língua oficial Portuguesa também são avaliados.

Neste inquérito não foram encontrados nenhuns diplomados que tivessem respondido residir em outro país, logo foram encontrados zero dados válidos para a avaliação.

Tabela 41:Frequência de residência por Outro país

Residência em Outro País (não Europeu e não PLOP)		Frequência
Ausente	nr	88
	Sistema	347
	na	3043
	Total	3478

Situação face ao emprego

Este grupo contém 3 variáveis que estudam a situação de emprego do diplomado, com algumas diferentes entre as variáveis.

Situação face ao Emprego

A primeira variável estuda a situação do diplomado em relação ao emprego. Se este, trabalha por conta de outrem, própria ou entre outras opções listadas.

A tabela a baixo indica sobre o número total de inquiridos qual a sua situação face ao emprego, grande parte diz trabalhar por conta de outrem, seguido de diplomados sem atividade profissional remunerada (SAPR). Trabalhar por conta própria com funcionários a cargo é a situação menos comum encontrada com apenas 23 dos diplomados inquiridos.

Tabela 42:Frequência de diplomados por categorias de emprego

Situação face ao Emprego		Frequência
Válido	T. conta própria c/ func. a cargo	23
	Bolseiro	223
	T. conta própria s/ func. a cargo	248
	Estagiário	266
	Sem atividade profissional remunerada	932
	T. conta de outrem	1786
	Total	3478

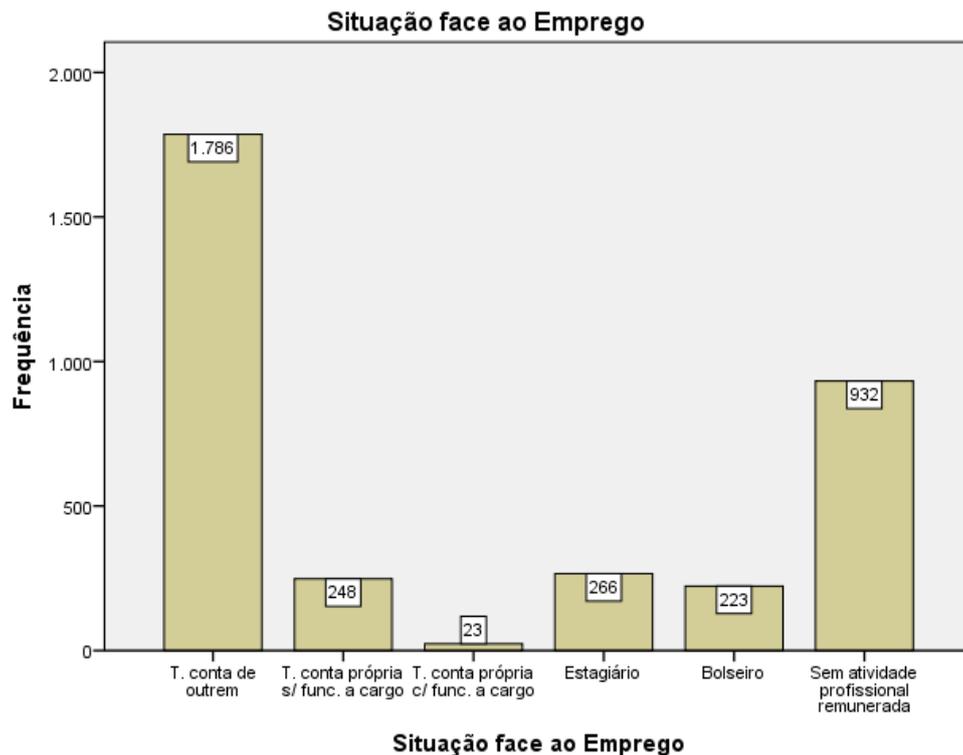


Figura 10: Gráfico de frequência de diplomados por categorias de emprego

Situação face ao Emprego 2

A próxima variável verifica também a situação face ao emprego por parte do diplomado, mas acrescentando as respostas de “Estudante”, “Reformado” e “SAPR, estuda mas deixaria”, este último não está presente na estatística pois tem zero diplomados associados. O valor de “Trabalha por conta própria” está agrupada, não diferenciando se tem ou não funcionários a cargo.

Sobre esta variável é possível retirar que grande parte trabalha por conta de outrem, seguido de estudantes, o valor com menos diplomados é o de reformado contendo apenas 3 do total de inquiridos.

Tabela 43: Frequência de diplomados face ao emprego

Situação face ao Emprego		Frequência
Válido	Reformado	3
	Bolseiro	223
	SAPR	245
	Estagiário	266
	T. conta própria	271
	Estudante	684
	T. conta de outrem	1786
	Total	3478

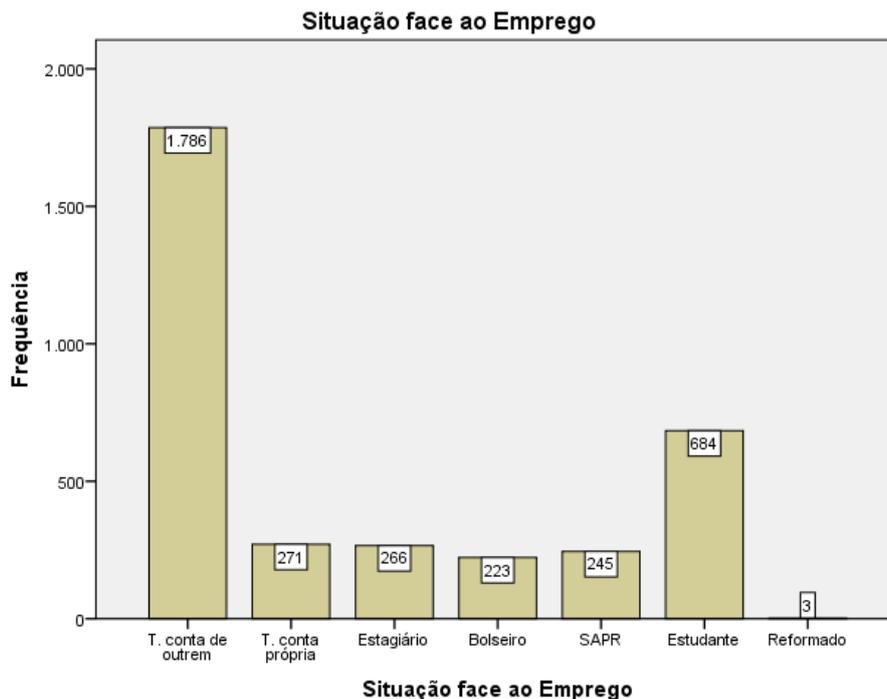


Figura 11: Gráfico de frequências de diplomados por situação de emprego

Situação face ao Emprego com estudante e reformado

Continuando a estudar a variável relacionada com a situação do diplomado face ao emprego, mas agora separando os que trabalham por conta própria com ou sem funcionários a cargo.

Os resultados estão de acordo com as variáveis apresentadas anteriormente.

Tabela 44: Frequência de diplomados face ao emprego

Situação face ao Emprego		Frequência
Válido	Reformado	3
	T. conta própria c/ func. a cargo	23
	Bolseiro	223
	SAPR	245
	T. conta própria s/ func. a cargo	248
	Estagiário	266
	Estudante	684
	T. conta de outrem	1786
	Total	3478

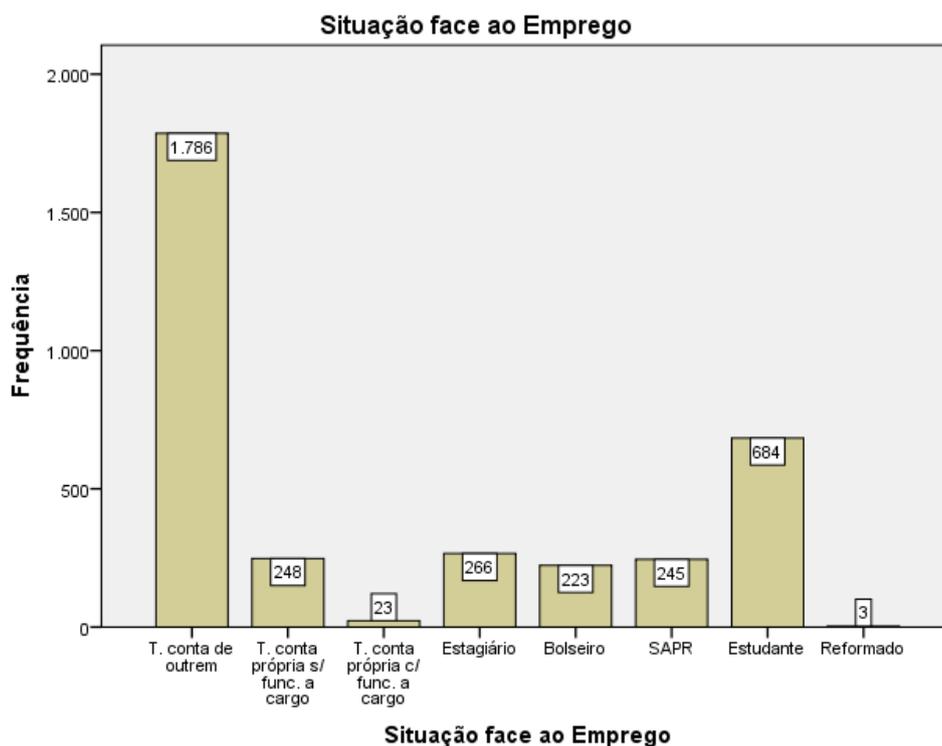


Figura 12: Gráfico de frequências da situação face ao emprego

Taxa de emprego

De seguida é estudado a taxa de emprego, verificando os diplomados sem atividade profissional remunerada (SAPR), com atividade profissional remunerada (CAPR), estudante e reformado. Os valores de estudante e reformado são considerados como omissos, visto que não exercem qualquer tipo de atividade.

Dos diplomados válidos grande parte tem atividade profissional remunerada (CAPR), dos valores omissos, a maioria dos diplomados continuaram com os estudos.

Tabela 45: Frequência categoria de emprego

Taxa de Emprego		Frequência
Válido	SAPR	428
	CAPR	2546
	Total	2974
Ausente	Estudante	501
	Reformado	3
	Total	504
Total		3478

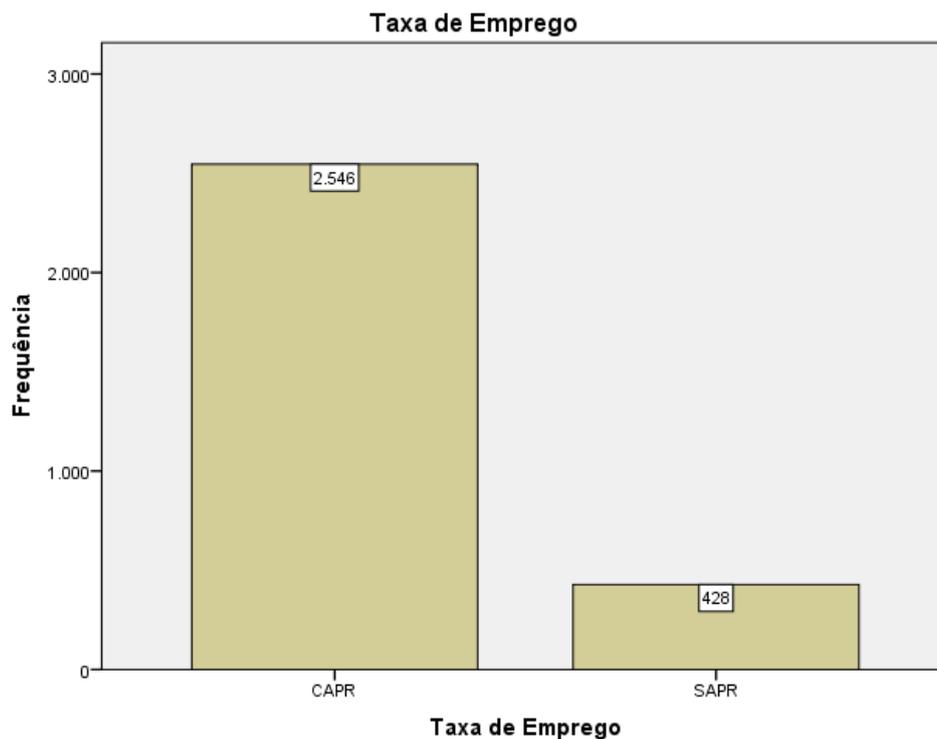


Figura 13: Gráfico com frequência de taxa de emprego

Local de emprego

Este grupo contém 13 variáveis que estudam a localização do emprego. Como a localização da residência, também é possível criar hierarquias semelhantes.

- País
 - Europeu
 - PLOP
 - Portugal
 - Área Metropolitana de Lisboa
 - Fora
 - Dentro
 - Distrito de Portugal
 - Concelhos
 - Lisboa
 - Santarém
 - Setúbal
 - Outros

Emprego (Portugal/Europeu/PLOP/Outro) (variável auxiliar)

A primeira variável deste grupo corresponde ao local onde se encontra o emprego do diplomado, neste caso se é em Portugal, Europa, PLOP ou Outro. Esta variável é apresentada como uma variável auxiliar.

Grande parte dos diplomados têm emprego em Portugal, seguido da Europa. Outro país é a opção que aparece com menos frequência.

Tabela 46: Frequência da localização de emprego por grupo de países

Emprego (Portugal/Europeu/PLOP/Outro) (variável auxiliar)		Frequência
Válido	Outro	28
	PLOP	35
	nr	172
	Europa	198
	na	934
	Portugal	2111
	Total	3478

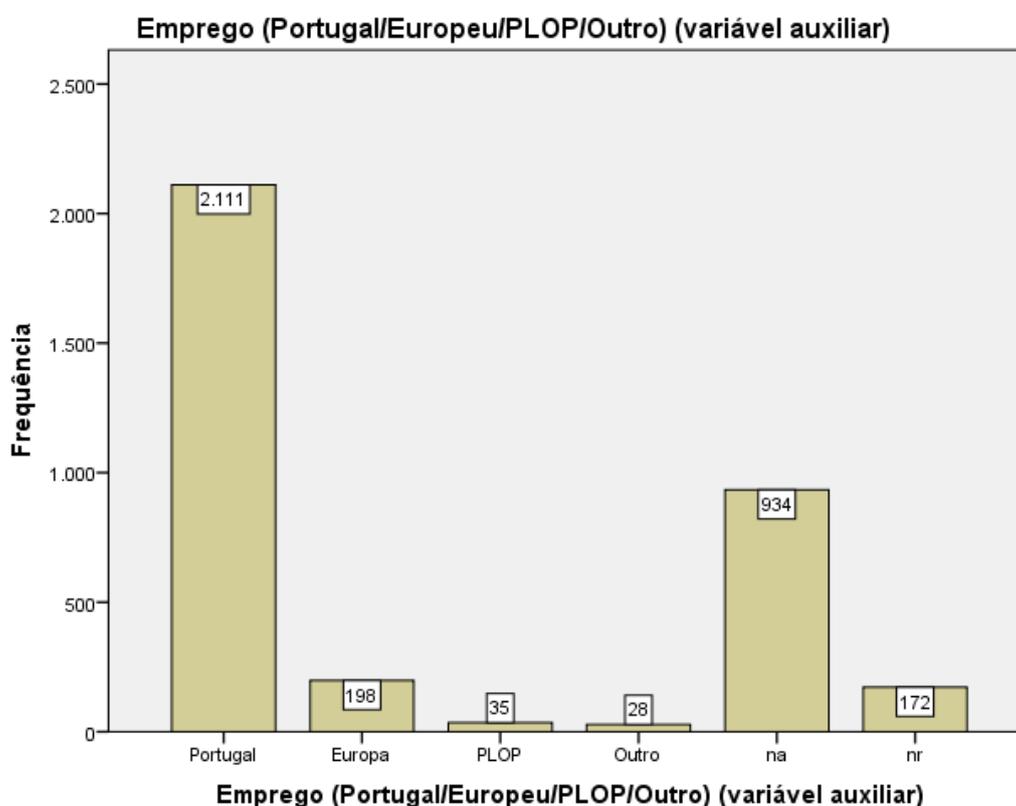


Figura 14: Gráfico de frequências da localização do emprego por grupo de países

Emprego (Portugal/Outro País)

De seguida é estudado o local do emprego dos diplomados por Portugal ou outro país.

É igual à variável anterior, onde 2111 têm emprego em Portugal, e 267 indicam trabalhar em outro país. Se for comparado com a tabela anterior, somando os que trabalham fora de Portugal são $198+35+28 = 261$, faltando 6 respostas em relação aos resultados desta variável, estes 6 valores em falta foram considerados como “não respondeu” na tabela anterior.

Tabela 47: Frequência de local de emprego por Portugal/Outro país

Emprego (Portugal/Outro País)		Frequência
Válido	Outro país	267
	Portugal	2111
	Total	2378
Ausente	na	934
	nr	166
	Total	1100
Total		3478

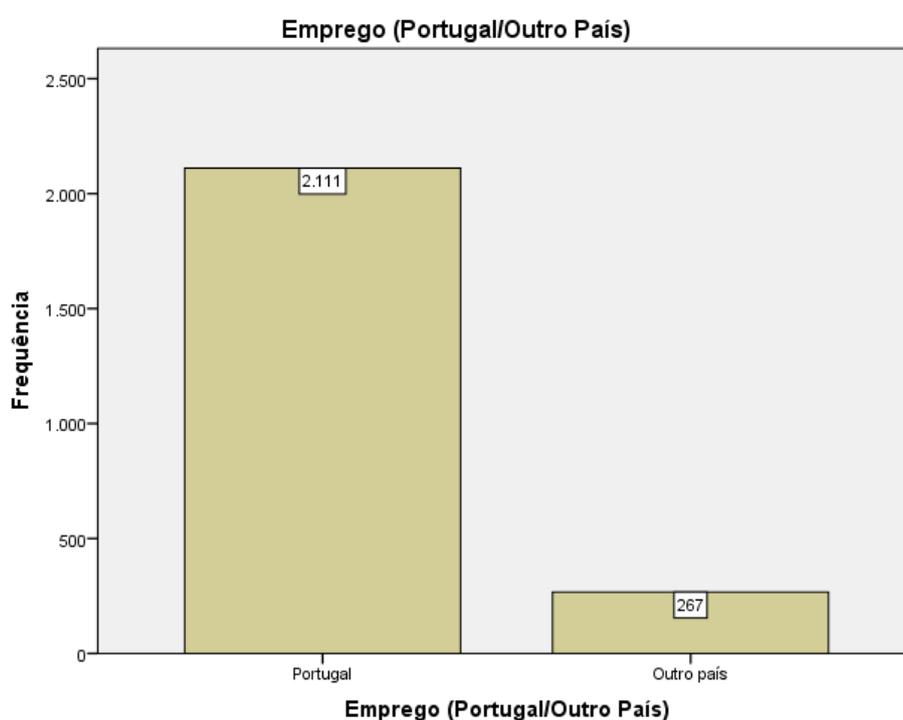


Figura 15: Gráfico de frequência de local de emprego por Portugal/Outro país

Localização do emprego, por Distrito

Para os diplomados que têm emprego em Portugal, é estudado a localização do emprego por distritos de Portugal.

Grande parte dos diplomados têm emprego em Lisboa, seguido de Setúbal. Os distritos com menor frequência são Guarda e Viana do Castelo.

Tabela 48: Frequência de local de emprego por distrito de Portugal

Localização do emprego, por Distrito		Frequência
Válido	Guarda	2
	Viana do Castelo	2
	Viseu	3
	Portalegre	4
	Vila Real	4
	Braga	6
	Aveiro	10
	Coimbra	10
	Castelo Branco	10
	Beja	13
	Madeira	17
	Évora	24
	Açores	25
	Faro	29
	Porto	32
	Leiria	37
	Santarém	42
	Setúbal	129
	Lisboa	1657
	Total	2056
Ausente	na	1201
	nr	221
	Total	1422
Total		3478

Localização do Emprego (AML) (variável auxiliar)

A variável seguinte é uma variável auxiliar que indica a localização do emprego do diplomado por Área Metropolitana de Lisboa (AML) ou no Estrangeiro.

Através da tabela a baixo é possível observar que grande parte dos empregos encontram-se na AML.

Tabela 49: Frequência de local de emprego por AML

Localização do Emprego (AML) (variável auxiliar)

		Frequência
Válido	Estrangeiro	267
	Fora da AML	321
	AML	1698
	Total	2286
Ausente	na	932
	nr	260
	Total	1192
Total		3478

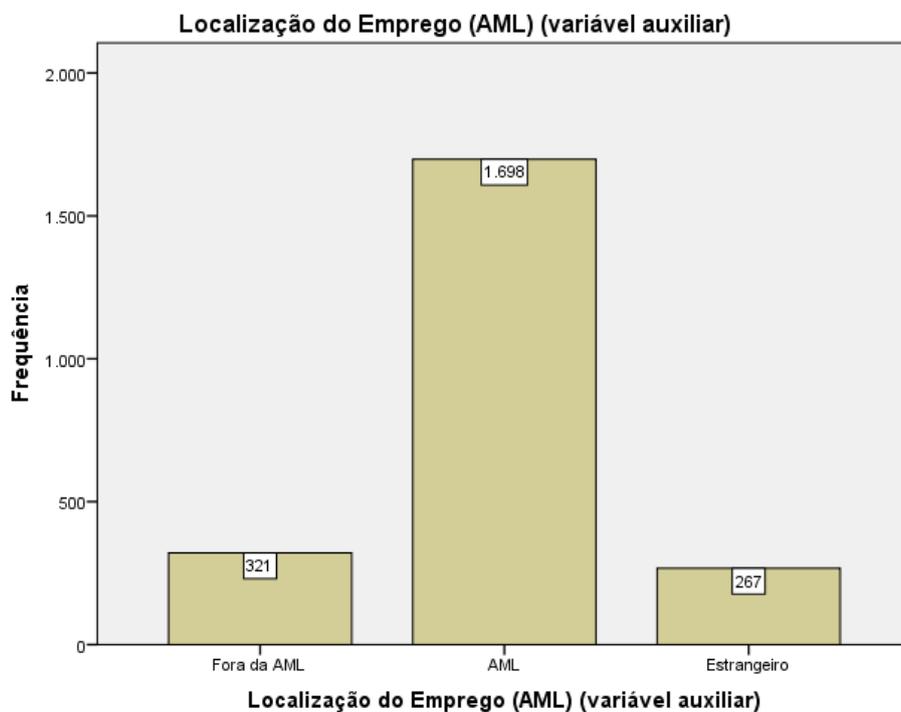


Figura 16: Gráfico com frequência de local de emprego por AML

Emprego distrito Lisboa/Setúbal/Outro e Estrangeiro (variável auxiliar)

De seguida é criada uma variável auxiliar que identifica a localização do emprego por distrito de Lisboa, Setúbal, outros e estrangeiro.

Grande parte dos empregos encontram-se no distrito de Lisboa, existindo valores muito próximos entre outros distritos e estrangeiro.

Tabela 50: Frequência de local de emprego por distrito e estrangeiro

Emprego dist Lisb_Setub_outro dist e Estrangeiro
(variável auxiliar)

		Frequência
Válido	Portugal-distrito Setúbal	129
	Estrangeiro	267
	Portugal-outros distritos	270
	Portugal-distrito Lisboa	1657
	Total	2323
Ausente	na	934
	nr	221
	Total	1155
Total		3478

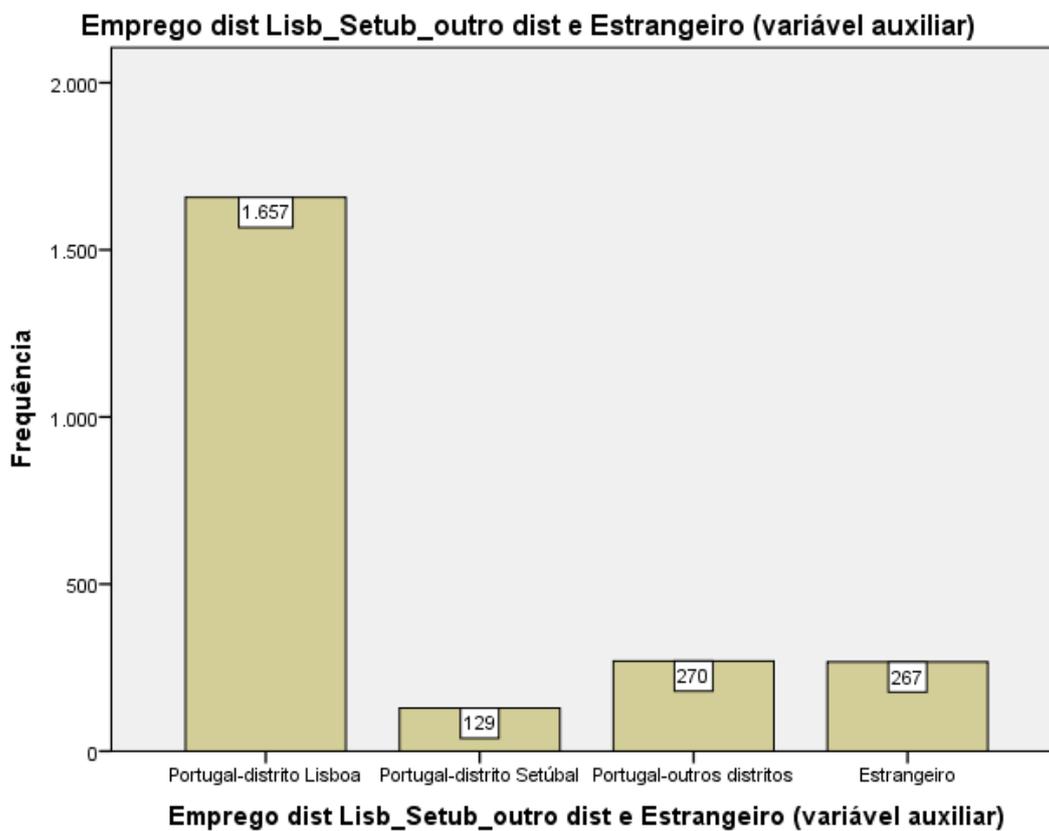


Figura 17: Gráfico de frequências de local de emprego por distrito e estrangeiro

Emprego por Concelhos de Lisboa

A variável seguinte corresponde à localização do emprego por concelho de Lisboa.

O concelho com maior frequência é Lisboa tendo 1122, os concelhos com menor frequência são de Arruda dos Vinhos e Cadaval.

Tabela 51: Frequência de emprego por concelhos de Lisboa

Emprego por Concelhos de Lisboa		Frequência
Válido	Arruda dos Vinhos	1
	Cadaval	1
	Sobral de Monte Agraço	2
	Lourinhã	3
	Alenquer	5
	Azambuja	12
	Mafra	12
	Torres Vedras	16
	Odivelas	25
	Vila Franca de Xira	35
	Amadora	49
	Loures	52
	Cascais	60
	Sintra	87
	Oeiras	138
	Lisboa	1122
Total	1620	
Ausente	na	1600
	nr	258
	Total	1858
Total		3478

Emprego por Concelhos de Santarém

Também são apresentados os empregos que se localizam pelos diferentes concelhos de Santarém.

Sendo que dos 41 valores válidos, 10 dos empregos encontram-se no concelho de Santarém, seguido de Ourém.

Tabela 52: Frequência de emprego por concelhos de Santarém

Emprego por Concelhos de Santarém		Frequência
Válido	Abrantes	1
	Almeirim	1
	Cartaxo	1
	Chamusca	1

	Constância	1
	Golegã	1
	Vila Nova da Barquinha	1
	Coruche	2
	Ferreira do Zêzere	2
	Rio Maior	2
	Benavente	3
	Tomar	3
	Torres Novas	6
	Ourém	6
	Santarém	10
	Total	41
Ausente	na	3215
	nr	222
	Total	3437
Total		3478

Emprego por Concelhos de Setúbal

A variável seguinte corresponde à localização dos empregos por concelhos de Setúbal.

O concelho que apresenta uma maior frequência é o de Almada com 30 empregos. O concelho de Alcácer do Sal é o que apresenta uma menor frequência.

De novo, o concelho de Alcochete não se encontra corretamente escrito, sendo necessário corrigir o erro nos dados (Alcochete em vez de Alchochete).

Tabela 53: Frequência de emprego por concelhos de Setúbal

Emprego por Concelhos de Setúbal		Frequência
Válido	Alcácer do Sal	1
	Grândola	2
	Santiago do Cacém	2
	Sines	4
	Alcochete	5
	Sesimbra	5
	Moita	6
	Palmela	11
	Montijo	13
	Barreiro	14
	Seixal	15

	Setúbal	19
	Almada	30
	Total	127
Ausente	na	3125
	nr	226
	Total	3351
Total		3478

Emprego fora de Portugal

É estudado a localização do emprego se este se encontrar fora de Portugal, ou seja, se está na Europa, PLOP ou outro país.

Dos valores válidos é possível observar que grande parte dos empregos se encontram na Europa, seguido de PLOP e outro país.

172 dos possíveis valores válidos não responderam, sendo um valor alto visto que não são muitos os empregos que se encontram noutra país para além de Portugal.

Tabela 54: Frequência de emprego por grupo de países

Emprego fora de Portugal		Frequência
Válido	Outro	28
	PLOP	35
	Europeu	198
	Total	261
Ausente	na	3045
	nr	172
	Total	3217
Total		3478

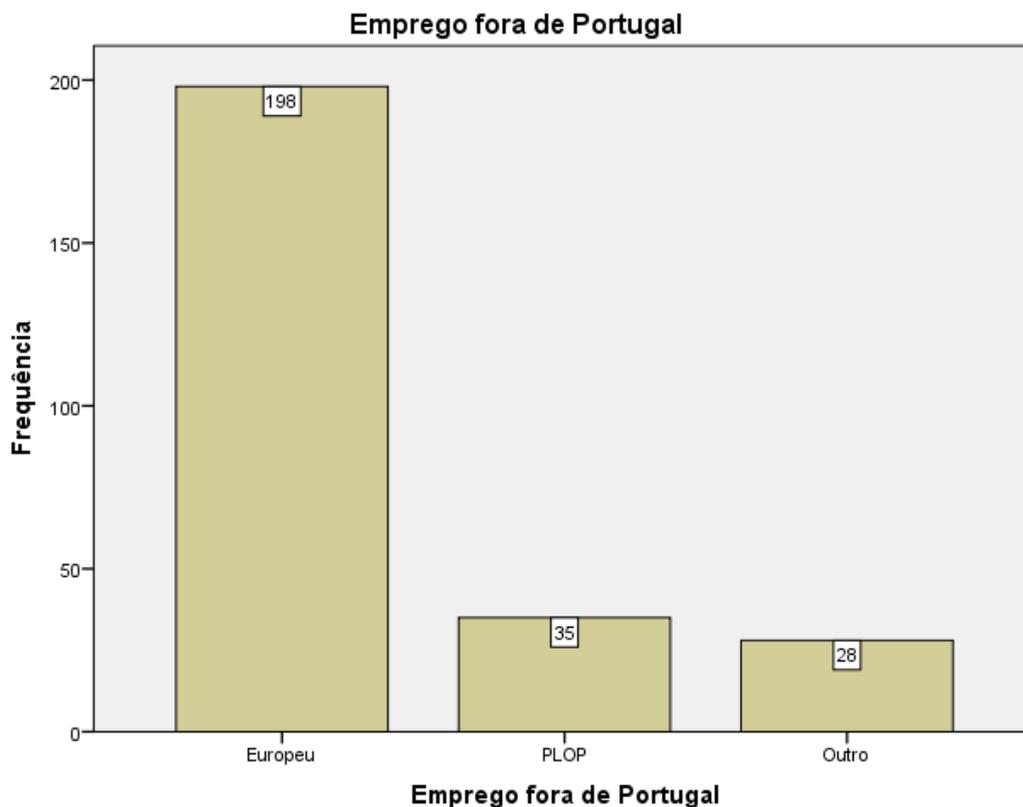


Figura 18: Gráfico de frequências de emprego por grupos de países

Emprego em País Europeu

Caso a localização dos empregos seja na Europa, estes são indicados com maior detalhe. Sendo, indicado em que país europeu se encontra o emprego.

Em comparação com a tabela anterior, existe uma diferença de 1 resposta que não foi apresentada, tendo sido considerada como “não respondeu”. O país Suíça não se encontra bem escrito, sendo necessário corrigir o erro ortográfico (Suiça para Suíça).

É possível observar que existe com maior frequência empregos localizados no Reino Unido (60), seguido da Alemanha (26). Os países em que os empregos estão com menor frequência são Eslovénia, Hungria e Islândia.

Tabela 55: Frequência de emprego por país Europeu

Emprego em País Europeu		Frequência
Válido	Eslovénia	1
	Hungria	1
	Islândia	1
	Itália	4
	Áustria	5
	Dinamarca	5

	Luxemburgo	7
	República da Irlanda	9
	França	10
	Suécia	10
	Espanha	11
	Bélgica	14
	Países Baixos (Holanda)	16
	Suíça	17
	Alemanha	26
	Reino Unido	60
	Total	197
Ausente	na	3108
	nr	173
	Total	3281
Total		3478

Emprego em PLOP

Caso a localização do emprego seja em Países de Língua Oficial Portuguesa (PLOP), estes também são indicados.

Existe consistência entre os valores válidos nesta variável e o total de empregos em PLOP indicados na Figura 49.

Com a variável que estuda a localização do emprego em PLOP é possível observar que Brasil é o país que aparece com maior frequência nas respostas. São Tomé e Príncipe é o país que aparece com menor frequência.

Tabela 56: Frequência de emprego por PLOP

Emprego em PLOP		Frequência
Válido	São Tomé e Príncipe	1
	Timor-Leste	2
	Cabo Verde	3
	Moçambique	3
	Angola	11
	Brasil	15
	Total	35
	Ausente	na
nr		172
Total		3443
Total		3478

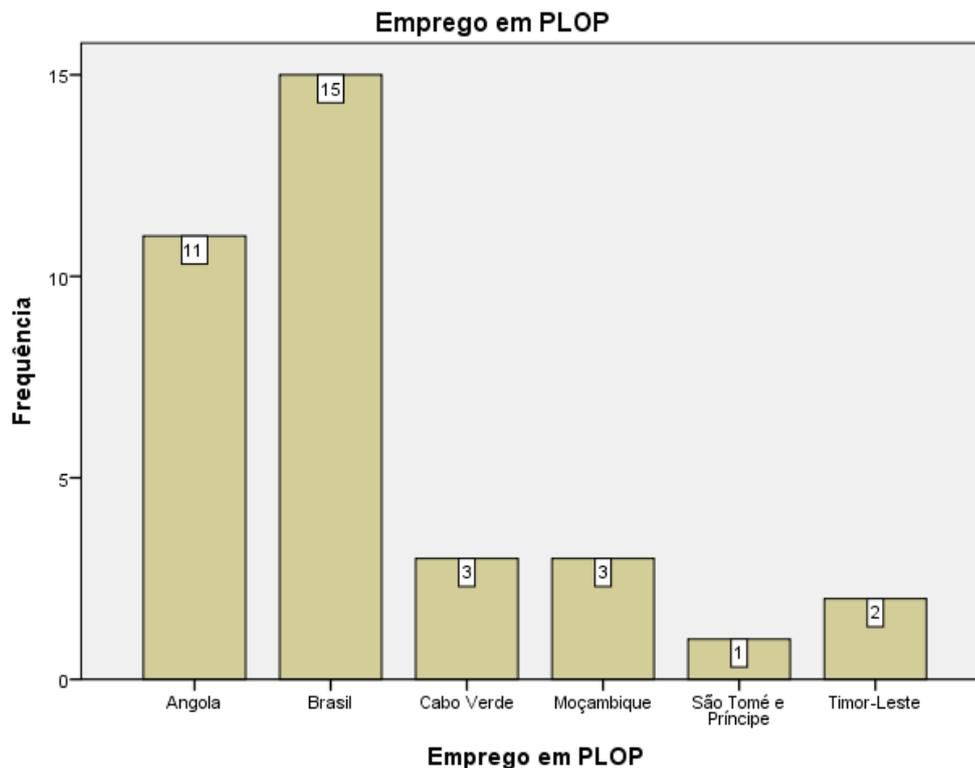


Figura 19: Gráfico de frequências de emprego por PLOP

Emprego em Outro País (não Europeu e não PLOP)

A variável seguinte corresponde à localização dos empregos noutros países (não Europeus e não PLOP).

De acordo com a figura 49 é possível observar que são 28 o número de vezes que a opção “Outro” aparece nos dados. Na tabela a baixo também são observados 28 dados válidos, logo existe consistência entre estas variáveis.

Os resultados são muito semelhantes entre os diferentes países, a Austrália, China e Estados Unidos da América apresentam uma frequência de 4 resultados em cada.

Tabela 57: Frequência de emprego em outros países

Emprego em Outro País (não Europeu e não PLOP)		Frequência
Válido	Colômbia	1
	Dubai, Emiratos Árabes Unido	1
	Equador	1
	Irão	1
	Japão	1
	Macau	1
	Singapura	1
	Tailândia	1

	Vários	1
	Canadá	2
	Emirados Arabes Unidos	2
	Qatar	3
	Austrália	4
	China	4
	Estados Unidos da América	4
	Total	28
Ausente	na	3278
	nr	172
	Total	3450
Total		3478

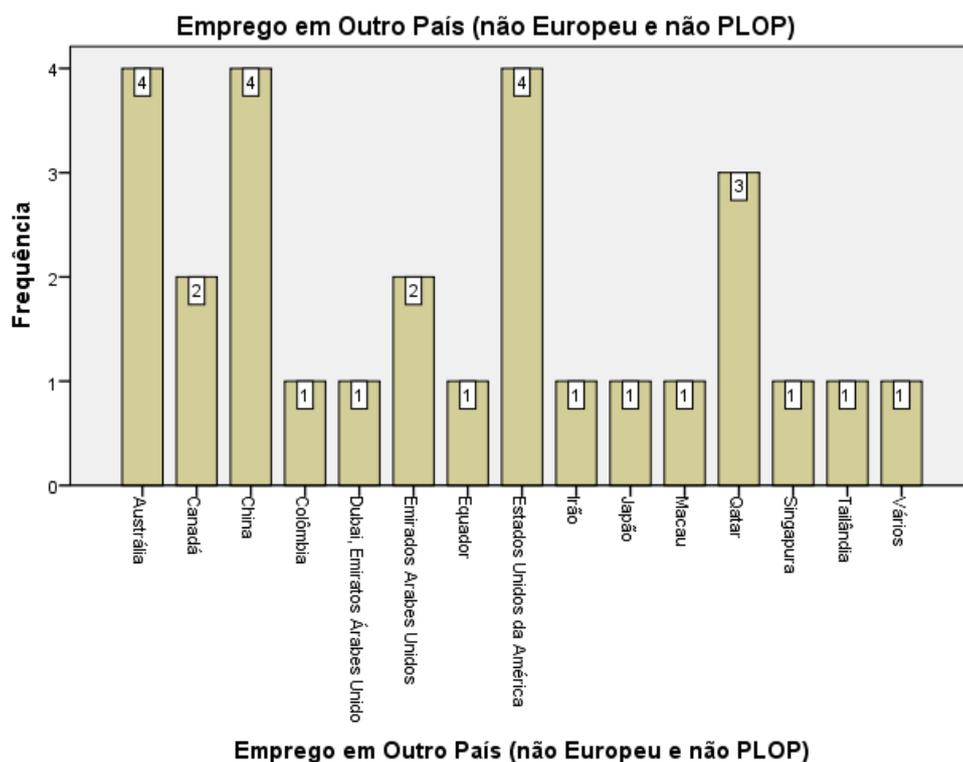


Figura 20: Gráfico de frequência de emprego por outro país

Países do emprego no estrangeiro (variável auxiliar)

A variável seguinte é uma variável auxiliar que corresponde ao quadro dos países onde o emprego se encontra.

Como de esperado, grande parte dos empregos dos diplomados encontram-se em Portugal, seguido de Europa. Macau é o país onde se encontrou empregos com menor frequência.

Existe um erro de escrita do qual precisa de ser corrigido, onde está escrito Oceania deverá ser substituído por Oceânia.

Tabela 58: Frequência de emprego por países no estrangeiro Quadro

Países do emprego no estrangeiro Quadro
(variável auxiliar)

		Frequência
Válido	Macau	1
	América Sul	2
	Oceânia	4
	América Norte	6
	Asia	7
	Outro	7
	PLOP	35
	Europa	198
	Portugal	2111
	Total	2371
	Ausente	na
nr		173
Total		1107
Total		3478

Antes de ir para o estrangeiro, procurou trabalho em Portugal? (Se não, qual a razão?)

Este grupo contém 9 variáveis e estuda se o diplomado procurou emprego em Portugal antes de ir para o estrangeiro iniciar as suas atividades profissionais, caso não tenha procurado são estudadas as razões para tal.

Procurou emprego antes de ir para o estrangeiro?

A próxima variável corresponde a uma pergunta realizada durante o inquérito, do qual, a resposta é de “Sim” ou “Não”.

“Antes de iniciar a atividade profissional no estrangeiro procurou trabalho em Portugal?”

Com a tabela a baixo é possível verificar que dos diplomados que foram para o estrangeiro iniciar a sua atividade profissional, 148 respondeu que tentou encontrar emprego primeiro em Portugal e 106 respondeu que não tentou.

Tabela 59: Frequência de diplomados que tentou encontrar emprego em Portugal antes de ir para o estrangeiro

Antes de iniciar a atividade profissional no estrangeiro procurou trabalho em Portugal?

		Frequência
Válido	Não	106
	Sim	148
	Total	254
Ausente	na	3045
	nr	179
	Total	3224
Total		3478

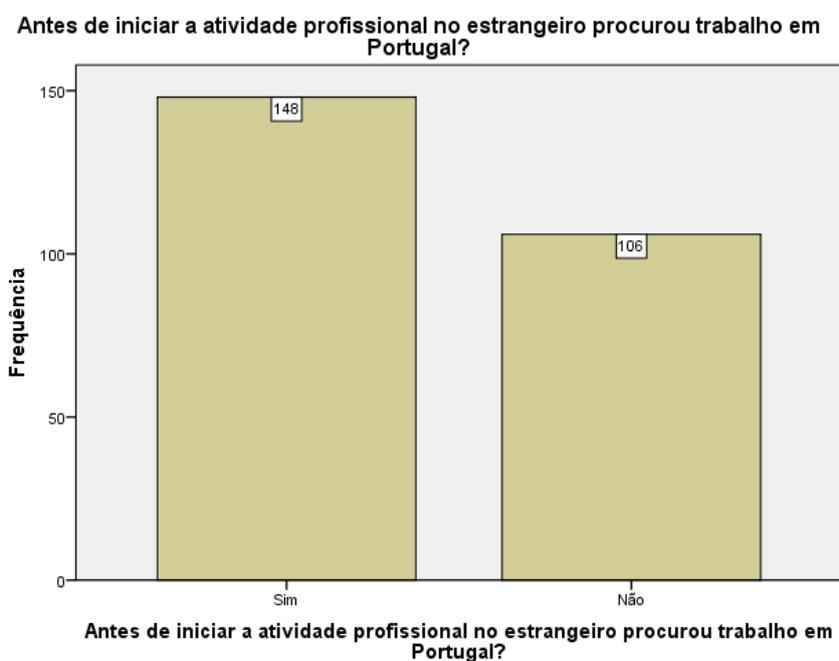


Figura 21: Gráfico de diplomados que tentou encontrar emprego em Portugal antes de ir para o estrangeiro

Razão 1

Depois é verificado dos diplomados que não tentaram procurar emprego em Portugal, qual a razão.

Uma das razões é: “Optei por permanecer no país onde estou a concluir ou concluí formação”

Do qual, 11 diplomados tiveram esta razão para não tentar encontrar emprego em Portugal.

Tabela 60: Frequência de diplomados que escolheram a razão 1

**Razão para não procura: Optei
por permanecer no país onde
estou a concluir ou concluí
formação**

		Frequência
Válido	Sim	11
	Não	95
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Razão 2

Outra razão para a qual os diplomados não tentaram encontrar emprego em Portugal: “Optei por experiência/percurso internacional”.

Onde 62 diplomados indicaram que esta foi a sua razão para não tentar encontrar emprego em Portugal.

Tabela 61: Frequência de diplomados que escolheram a razão 2

**Razão para não procura: Optei
por experiência/percurso
internacional**

		Frequência
Válido	Não	44
	Sim	62
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Razão 3

Outra razão é: “Regressei ao meu país de origem”. Onde 10 diplomados indicaram ser esta a razão para o qual não tentaram arranjar emprego em Portugal.

Tabela 62: Frequência de diplomados que escolheram a razão 3

Razão para não procura:
Regressei ao meu país de origem

		Frequência
Válido	Sim	10
	Não	96
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Razão 4

Outra razão é: “Na área do meu curso não existe mercado de emprego em Portugal”. Onde 15 diplomados indicaram ser esta a sua razão para não procurar emprego em Portugal.

Tabela 63: Frequência de diplomados que escolheram a razão 4

Razão para não procura: Na área
do meu curso não existe
mercado de emprego em
Portugal

		Frequência
Válido	Sim	15
	Não	91
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Razão 5

Outra razão é “Em geral, a conjuntura socioeconómica em Portugal não é favorável (ofertas de emprego/nível salarial/estabilidade)”. Onde 49 diplomados indicaram esta ser uma das razões pelo qual não tentaram procurar emprego em Portugal.

Tabela 64: Frequência de diplomados que escolheram a razão 5

Razão para não procura: Em geral, a conjuntura socioeconómica em Portugal não é favorável (ofertas de emprego/nível salarial/estabilidade)

		Frequência
Válido	Sim	49
	Não	57
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Razão 6

Outra razão é: “Por razões de ordem pessoal”. Onde 15 diplomados indicaram esta ser uma das razões pelo qual não tentaram procurar emprego em Portugal.

Tabela 65: Frequência de diplomados que escolheram a razão 6

Razão para não procura: Por razões de ordem pessoal

		Frequência
Válido	Sim	15
	Não	91
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Outra razão

Pode também ser “Outra” razão para além daquelas que foram apresentadas anteriormente. Onde 9 diplomados indicaram haver outra razão por não tentarem procurar emprego em Portugal.

Tabela 66: Frequência de diplomados que escolheram a "Outra" razão

Razão para não procura: Outra		Frequência
Válido	Sim	9
	Não	97
	Total	106
Ausente	na	3192
	nr	180
	Total	3372
Total		3478

Outra razão, qual?

Se o diplomado escolheu outra razão para não tentar procurar emprego em Portugal, que razões são essas?

Esta variável indica as outras razões que os diplomados apresentaram para não tentarem procurar emprego em Portugal.

No total foram 8 os diplomados que apresentaram as suas outras razões, como é possível observar na tabela a baixo.

Tabela 67: Outras razões indicadas pelos diplomados

Razão para não procura: Se "Outra", qual?		Frequência
Válido	Baixos salários/Demasiadas horas extra sem pagamento/ falta de empresas da área.	1
	Em Portugal não existem mestrados na área de systems biology.	1
	Já não vivia em Portugal	1
	Melhor oferta no estrangeiro	1
	Não sou Português	1
	Nunca quis viver nem trabalhar em Portugal. O Mundo e demasiado grande e a minha área de estudos sempre se focou na China.	1
	Penso que as oportunidades na Dinamarca são melhores	1

	Surgiu a oportunidade de me candidatar a uma bolsa numa universidade em Londres	1
	Total	8
Ausente	na	3289
	nr	181
	Total	3470
Total		3478

Entidade empregadora

Este grupo contém 2 variáveis, onde a primeira estuda a tipologia da entidade e a segunda estuda a dimensão.

Tipologia

Esta variável procura explorar a tipologia da entidade empregadora onde o diplomado exerce as suas atividades profissionais. Ou seja, se esta é pública, privada com fins lucrativos, privada sem fins lucrativos ou parceira público-privada.

Dos dados válidos, 1385 indicam que a sua entidade empregadora é privado com fins lucrativos, seguido de 555 que indicam ser pública.

Tabela 68: Frequência da entidade empregadora por tipologia

Tipologia da entidade empregadora (Pública/Privada com e sem fins lucrativos)

		Frequência
Válido	Parceria público-privada	83
	Privada sem fins lucrativos	97
	Pública	555
	Privada com fins lucrativos	1385
	Total	2120
Ausente	na	1202
	na	156
	Total	1358
Total		3478

Tipologia da entidade empregadora (Pública/Privada com e sem fins lucrativos)

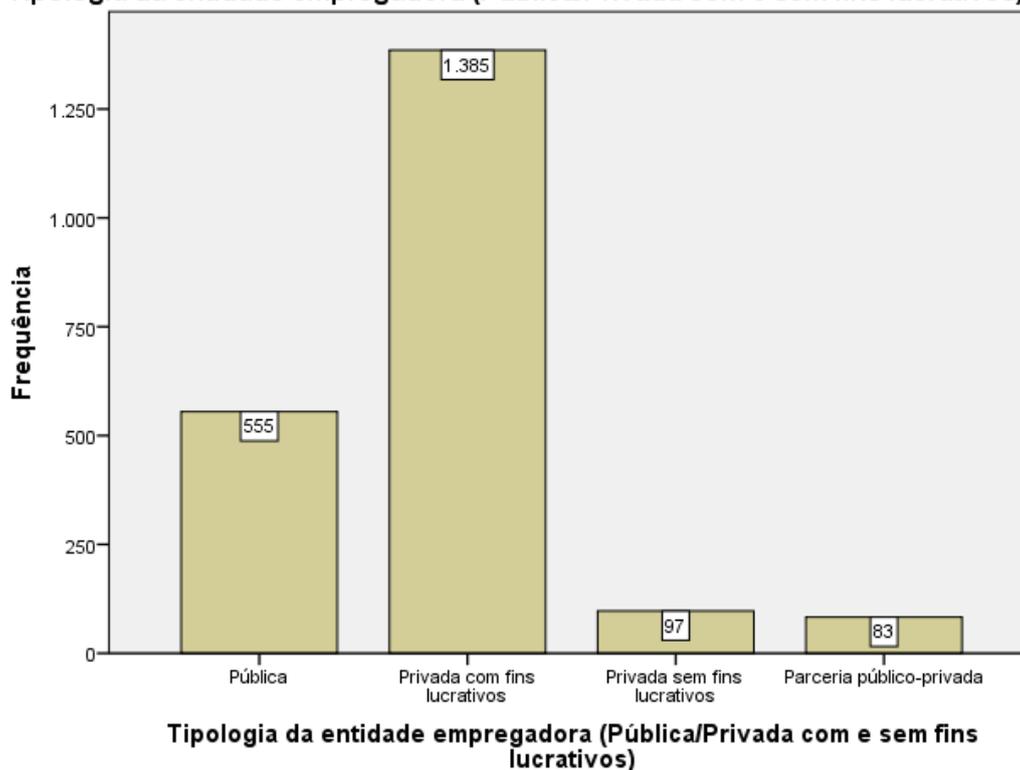


Figura 22: Gráfico de frequência de entidade empregadora por tipologia

Dimensão

Esta variável corresponde à dimensão da entidade empregadora, do qual, é apresentada por intervalos de pessoas, de 1 a 9 pessoas, 10 a 49, 50 a 249 ou 250 ou mais pessoas.

Com a tabela a baixo é possível verificar que 1113 dos diplomados indicaram que a dimensão da empresa está entre 250 ou mais pessoas. Logo grande parte dos diplomados trabalha em empresas de grande escala.

Tabela 69: Frequência da dimensão por intervalo

Dimensão da entidade empregadora		Frequência
Válido	1 a 9 pessoas	280
	10 a 49 pessoas	341
	50 a 249 pessoas	371
	250 ou mais pessoas	1113
	Total	2105
Ausente	na	1180
	nr	193
	Total	1373
Total		3478

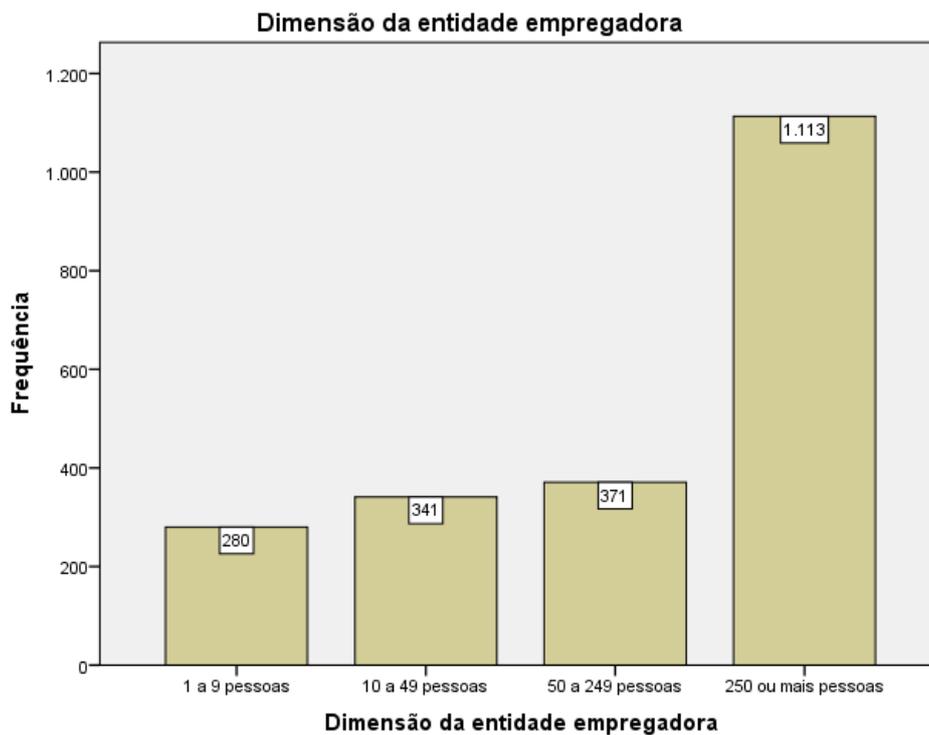


Figura 23: Gráfico de frequências da dimensão da entidade empregadora

Vínculo

Este grupo contém 2 variáveis, onde a primeira variável estuda o tipo de vínculo contratual que o diplomado tem, isto é, se é efetivo, a termo ou outro tipo. A segunda variável estuda o caso de ser outro tipo de vínculo.

Tipo de vínculo contratual

Esta variável estuda o tipo de vínculo contratual que o diplomado tem. Se é efetivo, a termo ou outro.

Os resultados entre efetivo e a termo são muito semelhantes, tendo apenas uma diferença de 9. No entanto, é o efetivo que se encontra com uma frequência mais elevada em relação aos outros.

Tabela 70: Frequência de vínculo por tipo

Tipo de vínculo contratual (Efetivo/A termo/Outro)		Frequência
Válido	Outro	26
	A termo (contrato de trabalho por tempo determinado, certo ou incerto)	834

	Efetivo (contrato de trabalho sem termo ou por tempo indeterminado)	843
	Total	1703
Ausente	na	1692
	nr	83
	Total	1775
Total		3478

Outro tipo de vínculo contratual

Se o vínculo contratual for outro, é pedido ao diplomado que indique que tipo de contrato tem. Houve 13 diplomados que indicaram que outro tipo de vínculo contratual têm.

Nesta variável o valores apresentados com 998 significam 'NA' e 999 significam "Não respondeu".

Tabela 71: Frequência de vínculos por outros tipos

Tipo de vínculo contratual (Se "Outro", qual?)

		Frequência
Válido	Bolsa de Gestão Ciência e Tecnologia (falsa bolsa) - precária do Estado	1
	Comissão de Serviço	1
	Contrato de freelance, em vias de passar a contrato de trabalho.	1
	Desenhador CAD	1
	Estágio IEFP	1
	Neste momento estou em part-time (1x por semana) porque ainda estou a terminar a dissertação de mestrado. No final do ano passarei a tempo inteiro, em Estágio do Instituto de Empre	1
	Nomeação política	1
	Sazonal	1
	Sem contrato de momento	1
	Sem contrato de trabalho	1

	Substituição	1
	Temporário	1
	Trabalho temporário	1
	Total	13
Ausente	998	3367
	999	98
	Total	3465
Total		3478

Volume de faturação (Empresário)

É estudado o volume de faturação da empresa dos diplomados que indicam ser trabalhadores por conta própria com funcionários a cargo (Empresário).

É pedido que indiquem se um dos clientes representa mais de 80% do volume de faturação ou se dois dos clientes representam, isoladamente, um volume de faturação superior a 30% ou se diversos clientes representam volumes de faturação semelhantes.

Dos diplomados que apresentam ser empresários, a maioria indica que diversos clientes representam volumes de faturação semelhantes.

Tabela 72: Frequência do volume de faturação da empresa

Volume de faturação da empresa (apenas para "Empresário")

		Frequência
Válido	Dois dos clientes representam, isoladamente, um volume de faturação superior a 30%	1
	Um dos clientes representa mais de 80% do volume de faturação	3
	Diversos clientes representam volumes de faturação semelhantes	12
	Total	16
Ausente	Outro	2
	na	3455
	nr	5
	Total	3462
Total		3478

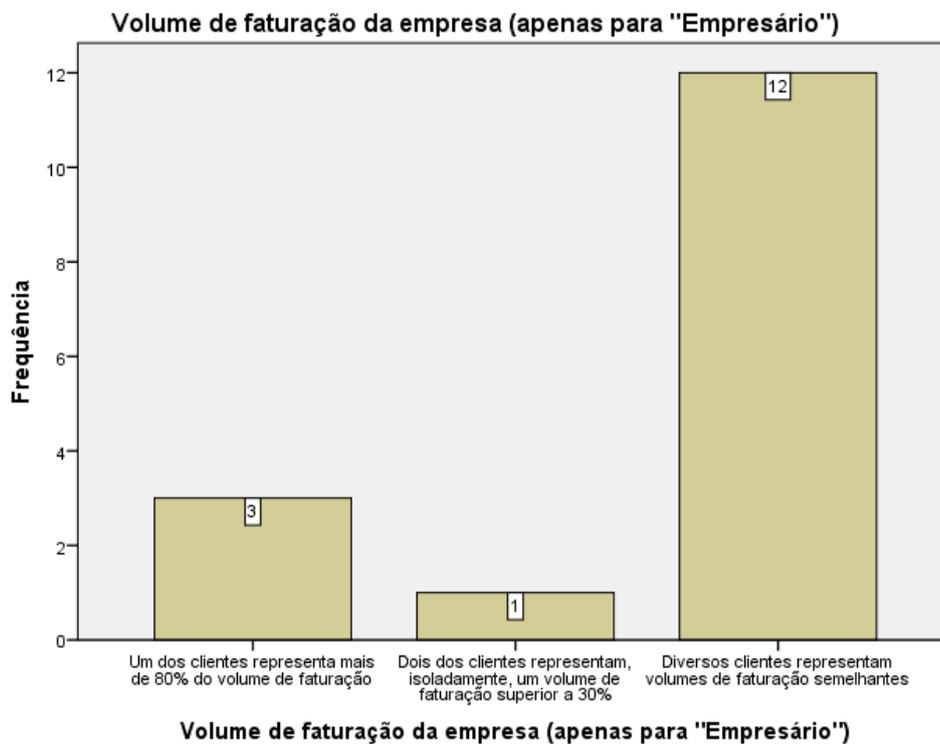


Figura 24: Gráfico de frequência do volume de faturação da empresa

Horário

De seguida é estudado o horário de trabalho do diplomado, ou seja, se este trabalha em tempo parcial ou inteiro.

Na tabela a baixo verificamos que grande parte dos diplomados trabalham a tempo inteiro.

Tabela 73: Frequência por horários de trabalho

Horário de trabalho (Tempo parcial/Tempo inteiro)

		Frequência
Válido	Tempo parcial	240
	Tempo inteiro	2068
	Total	2308
Ausente	na	955
	nr	215
	Total	1170
Total		3478

Vencimento

Este grupo contém 2 variáveis, onde a primeira estuda o salário bruto mensal do diplomado, e a segunda representa o mesmo que a primeira mas retirando os outliers.

Vencimento com outliers

A próxima variável corresponde ao valor aproximado do salário mensal bruto do diplomado.

Esta variável tem 336 valores de salários diferentes, do qual o valor mínimo encontrado foi de zero e o máximo de 500.000.

O valor indicado como 500.000 é considerado um outlier, após análise deste caso específico, foi verificado que pertence a um diplomado de Angola, podendo existir a possibilidade de o valor indicado ser da moeda do país e não em euros.

No entanto o salário mensal bruto com mais registos é de 1000 euros, seguido de 1200 euros, sendo valores plausíveis.

Tabela 74: Estatística descritiva do vencimento

Estatísticas descritivas			
	N	Mínimo	Máximo
Vencimento (Valor)	2305	0	500000
N válido (de lista)	2305		

Tabela 75: Frequências mais altas de vencimentos

Vencimento (Valor)		
		Frequência
Válido	1800	75
	600	80
	700	93
	800	94
	900	103
	1500	107
	980	121
	1200	128
	1000	198
	Total	2305
Ausente	na	932
	nr	241
	Total	1173
Total		3478

Para consultar a tabela completa, abrir o ficheiro "Tabelas completas" (Tabelas completas (com muitas entradas)).

Vencimentos sem outliers

A próxima variável volta a indicar os vencimentos mensais brutos dos diplomados mas excluindo os valores considerados como outliers.

Foram considerados 227 dos valores da tabela como outliers, valores entre 0 e 410 euros, bem como, valores entre 2855 e 500.000.

O valor mínimo passou a ser de 418 euros e o máximo de 2800, dando uma média de 1143,67 euros de salário mensal bruto.

Tabela 76: Estatística descritiva do vencimento sem outliers

Estatísticas descritivas				
	N	Mínimo	Máximo	Média
Vencimento (Valor)	2078	418	2800	1143,67
N válido (de lista)	2078			

Complementos remuneratórios

Este grupo contém 3 variáveis, onde a primeira identifica os diplomados que recebem complementos remuneratórios, a segunda variável estuda os valores desses complementos e a terceira estuda de novo os valores dos complementos mas retirando os outliers.

Complementos remuneratórios (Sim/Não)

Esta variável estuda se o diplomado recebe ou não outros complementos remuneratórios, como por exemplo, subsídio de alimentação, de transporte e outros.

Dos valores válidos, 1065 indicou não receber outros complementos remuneratórios e 824 indicaram que recebiam.

Tabela 77: Frequência de diplomados com/sem complementos

Complementos remuneratórios (Sim/Não)		
		Frequência
Válido	Sim	824
	Não	1065
	Total	1889
Ausente	na	1403
	nr	186
	Total	1589
Total		3478

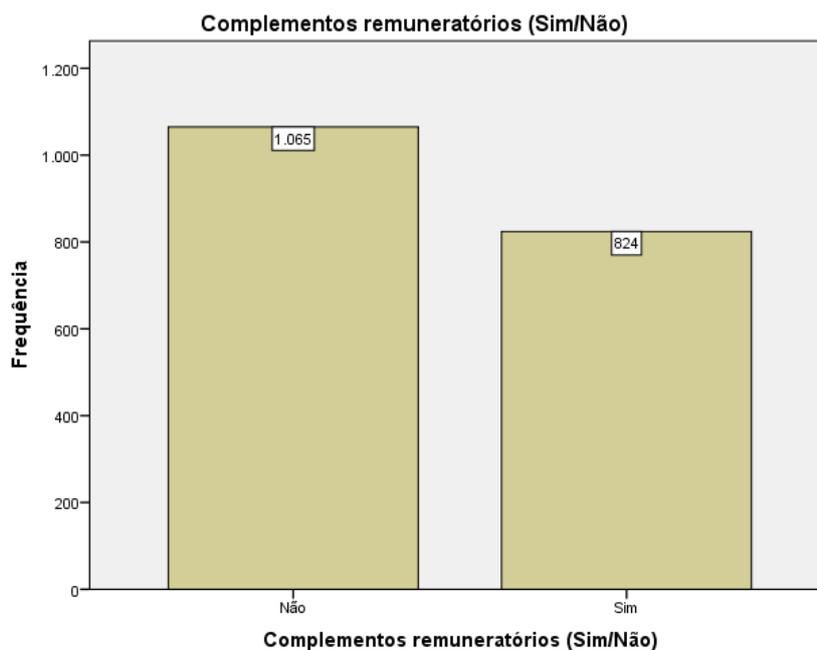


Figura 25: Gráfico de frequências de complementos remuneratórios

Valor dos complementos remuneratórios

Esta variável corresponde ao valor mensal aproximado para o conjunto dos complementos remuneratórios, caso o diplomado receba complementos.

Foram indicados 131 valores diferentes onde o valor mínimo apresentado foi de zero e o máximo de 3800 euros. Respostas como zeros euros ou 1 euro, não fazem muito sentido.

O valor de 100 euros é o mais respondido, tendo 80 dos 713 dos diplomados, seguido de 200 euros com 78 diplomados.

Tabela 78: Frequência de valores de complementos remuneratórios

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Complementos remuneratórios (Valor)	713	0	3800	410,38	536,616
N válido (de lista)	713				

Tabela 79:Frequências mais altas de valores de complementos remuneratórios

Complementos remuneratórios (Valor)		Frequência
Válido	80	10
	350	10
	700	10
	1200	11
	2000	14
	120	18
	1000	18
	400	20
	250	25
	50	29
	500	34
	300	38
	150	61
	200	78
	100	80
	Total	713
	Ausente	na
nr		297
Total		2765
Total	3478	

Para consultar a tabela completa, abrir o ficheiro “Tabelas completas” (Tabelas completas (com muitas entradas)).

Valor dos complementos remuneratórios sem outliers

A próxima variável indica de novo os valores em euros dos complementos remuneratórios mas excluindo os valores considerados como outliers.

Valores entre 0 e 37 euros são considerados como outliers, bem como, valores entre 1550 e 3800 euros. Tendo sido considerados no total 70 valores outlier.

Assim, o novo mínimo é de 40 euros e o máximo de 1500 euros, dando uma média de 334,25 euros.

Tabela 80: Estatística descritiva do valor de complementos remuneratórios

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Complementos remuneratórios (Valor)	643	40	1500	334,25	340,023
N válido (de lista)	643				

Duração de estágio ou bolsa (estagiário/bolseiro)

Este grupo contém 2 variáveis, onde a primeira estuda exclusivamente os estagiários e a duração dos seus estágios. A segunda variável estuda exclusivamente os bolsseiros e as durações das suas bolsas.

Duração do estágio

Esta variável é exclusivamente para os estagiários e corresponde à duração do estágio. Esta questão é apresentada com intervalo de meses e anos, isto é, se tem duração de menos de 6 meses, entre 6 meses a 1 ano ou mais de 1 ano.

Dos estagiários que apresentaram uma resposta para esta variável, grande parte indicou que a duração da sua bolsa está entre 6 meses e 1 ano.

Tabela 81: Frequência da duração do estágio

Duração do estágio (apenas para "Estagiário")		Frequência
Válido	Menos de 6 meses	10
	Mais de 1 ano	77
	Entre 6 meses e 1 ano	163
	Total	250
Ausente	na	3212
	nr	16
	Total	3228
Total		3478

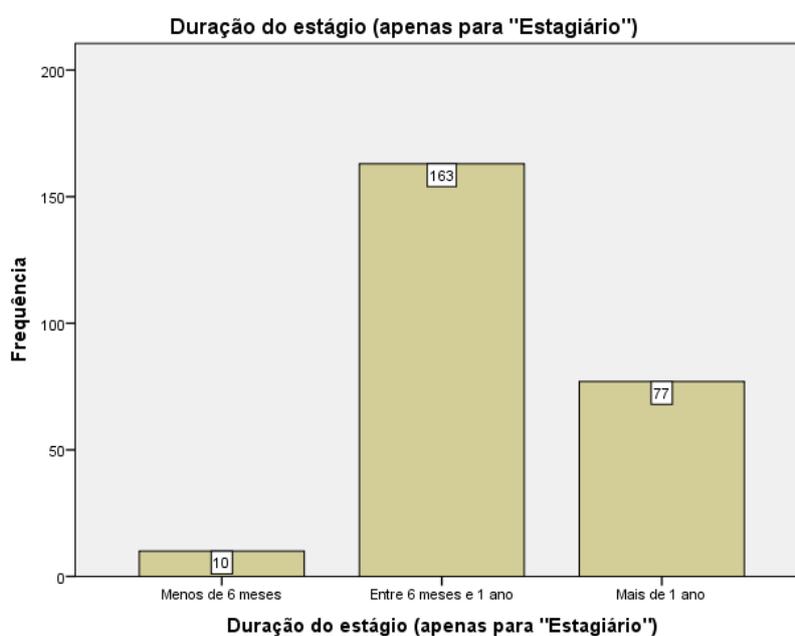


Figura 26: Gráfico de frequências da duração do estágio

Duração da bolsa

Esta variável é destinada para os bolsеiros, onde é estudado a duração das bolsas.

O inquérito pede ao diplomado que insira a duração da sua bolsa, caso seja bolsеiro, em anos e meses.

Observando os resultados desta variável é possível que os inquiridos tenham respondido só com meses ou só com anos. Ficando na dúvida se 1 significa 1 ano ou 1 mês, se o 12 é 12 meses (o mais lógico) ou 12 anos, 3,5 se são 3 meses e meio ou 3 anos e meio.

O valor 998 corresponde ao 'na'. Analisando a tabela a baixo e assumindo que os valores são todos em meses, grande parte dos bolsеiros têm bolsas de 48 meses (4 anos), seguido de bolsas de 12 meses (1 ano).

Tabela 82: Frequência da duração da bolsa

Duração da bolsa (apenas para "Bolsеiro")		Frequência
Válido	1	1
	2	1
	3	3
	4	4
	6	21
	9	3
	10	2
	12	54
	13	1
	16	1
	18	5
	19	1
	20	2
	24	6
	3,5	1
	5	1
	8	1
	36	39
	39	1
	41	1
42	1	
48	57	
60	1	
72	2	

	96	1
	980	1
	Total	212
Ausente	na	3255
	nr	11
	Total	3266
Total		3478

Tipo de bolsa (bolseiros)

Este grupo é composto por 3 variáveis, onde a primeira estuda o tipo de bolsa que o diplomado bolseiro possui, a segunda estuda o tipo de bolsa possui, caso não seja das que estão listadas na primeira variável. A última variável estuda se a bolsa que o diplomado possui é para efeito de estudos ou não.

Tipo de bolsa

A próxima variável é focada nos bolseiros, do qual estudo o tipo de bolsa que o diplomado tem.

Com a tabela a baixo, 112 dos diplomados bolseiros têm bolsa de doutoramento, seguido de 60 que têm bolsa de investigação. A bolsa que aparece com menos frequência é a bolsa de desenvolvimento de carreira científica.

Tabela 83: Frequência de diplomados por tipo de bolsa

Tipo de bolsa (apenas para Bolseiros)		Frequência
Válido	Bolsa de desenvolvimento de carreira científica (BDCC)	1
	Bolsa de doutoramento de empresas (BDE)	4
	Bolsa de iniciação científica (BIC)	4
	Bolsa de técnico de investigação (BTI)	7
	Bolsa de gestão de ciência e tecnologia (BGCT)	10
	Outra bolsa remunerada	20
	Bolsa de investigação (BI)	60
	Bolsa de doutoramento (BD)	112
	Total	218
	Ausente	na
nr		5

	Total	3260
Total		3478

Outro tipo de bolsa

Caso o diplomado tenha outro tipo de bolsa remunerada, é pedido que indiquem esse tipo de bolsa.

Houve 16 diplomados que indicaram as suas bolsas, do qual, 8 disseram ter bolsas de mestrado, seguido de bolsas de estudo do DGES (Direção-Geral de Ensino Superior).

Tabela 84: Frequência de diplomados por outro tipo de bolsa

Tipo de bolsa: Se "Outra", qual? (apenas para Bolseiros)

		Frequência
Válido	Aluno de doutoramento	1
	Bolsa alemã, atribuída pela DFG	1
	Bolsa da Ação Social	1
	Bolsa de estudo	1
	Bolsa de formação profissional em contexto prático	1
	Full Travel Grant, Residency in HKBU Academy of Visual Arts	1
	Bolsa de estudo do DGES	2
	Bolsa de Mestrado	8
	Total	16
	Ausente	na
nr		8
Total		3462
Total		3478

Bolsa para efeito de estudo

A variável seguinte é uma variável considerada como auxiliar, e indica se as bolsas são ou não para efeito de estudo.

Do qual, 133 diplomados indicaram que são para efeito de estudo e 83 responderam que não.

Tabela 85:Frequência de bolsas para efeitos de estudos

**Bolsa para efeito de estudo
(Sim/Não; variável auxiliar)**

		Frequência
Válido	Não	83
	Sim	133
	Total	216
Ausente	na	3255
	nr	7
	Total	3262
Total		3478

Alguma vez já exerceu atividade profissional remunerada na área do curso? (SAPR)

Esta variável é destinada apenas para os diplomados que estão sem atividade profissional remunerada (SAPR). Onde é estudado se o diplomado já exerceu alguma atividade profissional remunerada na área de formação do curso.

Dos diplomados SAPR, 29 não respondeu, 99 indicou que já exerceu atividade profissional remunerada na área do curso e 83 indicou que não.

Tabela 86:Frequência de diplomados SAPR que exerceram ou não atividade profissional na área de curso

**Já exerceu atividade profissional
remunerada na área de formação do
curso? (Sim/Não) (apenas para SAPR)**

		Frequência
Válido	Não	83
	Sim	99
	Total	182
Ausente	na	3267
	nr	29
	Total	3296
Total		3478

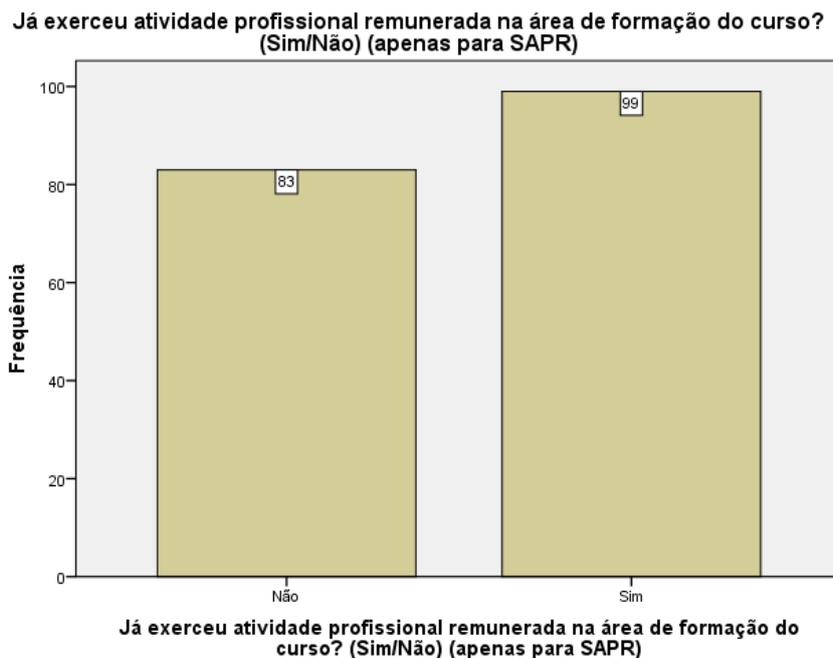


Figura 27: Gráfico de frequência de diplomados SAPR que exerceram trabalho na área

Emprego na área de curso

Este grupo contém 2 variáveis, onde a primeira estuda se o diplomado tem emprego na área de curso que tirou. A segunda tem como objetivo calcular a taxa de diplomados com emprego na área do curso.

Emprego na área (Sim/Não)

Esta variável corresponde ao diplomado e se este tem emprego na área do curso ou não, no entanto esta variável está indicada como “não considerar”.

Os resultados indicam que 1877 responderam que trabalham na área do curso e 573 respondeu que não, tendo havido 96 diplomados que não responderam.

Tabela 87: Frequência de diplomados com emprego na área

Emprego na área (não considerar) do curso (Sim/Não)

		Frequência
Válido	Não	573
	Sim	1877
	Total	2450
Ausente	na	932
	nr	96
	Total	1028
Total		3478

Emprego na área (para cálculo da taxa)

Esta variável também é relacionada com o emprego do diplomado e se este é da área do curso ou não. No entanto, esta variável está indicada como “a utilizar” e serve para cálculo da taxa.

Neste caso, os valores válidos são diferenciados entre diplomados Com Atividade Profissional Remunerada na Área de Formação (CAPRAF), Com Atividade Profissional Remunerada (CAPR) e Sem Atividade Profissional Remunerada (SAPR). Os valores ausentes são apresentados como nr (Não Respondeu) ou Estudante que Não Procura Emprego (ENPE).

Dos valores válidos, 1877 indicou ser um diplomado CAPRAF, 573 indicou ser CAPR e 428 como SAPR. Dos valores ausentes, 504 são ENPE e 96 não respondeu.

Tabela 88: Freqüência de diplomados com emprego na área (para cálculo de taxa)

Emprego na área (a utilizar) do curso (para cálculo da taxa)		Frequência
Válido	SAPR	428
	CAPR	573
	CAPRAF	1877
	Total	2878
Ausente	ENPE	504
	nr	96
	Total	600
Total		3478

Mesmo emprego antes e depois de terminar o curso (benefícios do curso)

Este grupo contém 9 variáveis, onde a primeira estuda se o diplomado mantém o mesmo emprego depois de terminar o curso. As 8 variáveis seguintes estudam as melhorias que o curso deu para o trabalho que tinham antes de concluir o curso.

Mantém mesmo emprego depois de concluir o curso

A variável seguinte estuda se o diplomado exercia o trabalho atual antes de iniciar o curso.

Dos resultados válidos, 1920 indicou que não exercia o trabalho atual antes de iniciar o curso e 297 indicou que sim.

Tabela 89: Frequência de diplomados que têm o mesmo emprego antes e depois de terminar o curso

**Antes de iniciar o curso já exercia o seu atual emprego?
(Sim/Não)**

		Frequência
Válido	Sim	297
	Não	1920
	Total	2217
Ausente	na	1197
	nr	64
	Total	1261
Total		3478

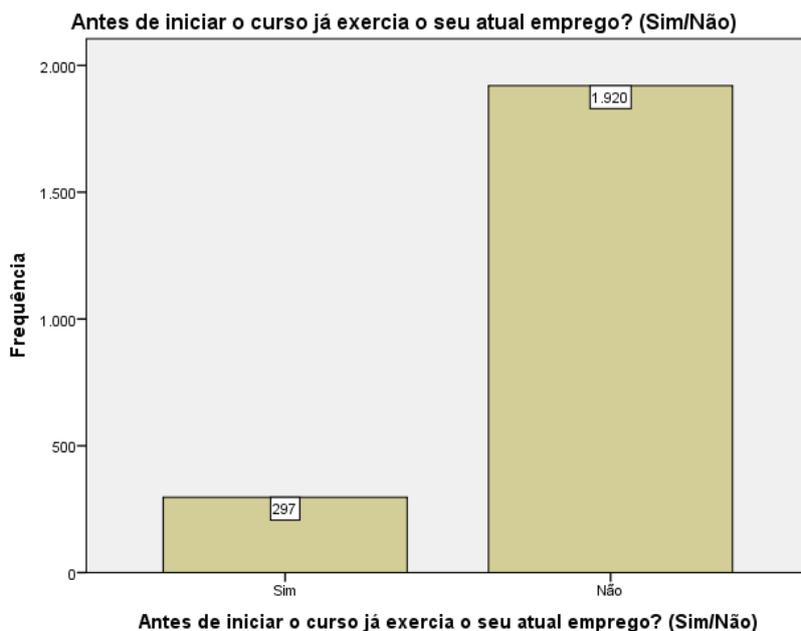


Figura 28: Gráfico de frequências por diplomados que exercem o mesmo emprego antes e depois de terminar o curso

Melhoria 1

De seguida é apresentado um conjunto de melhorias que o curso poderá ter trazido para a atividade profissional do diplomado, caso ainda exerça o mesmo trabalho após conclusão do curso (se respondeu sim à questão apresentada anteriormente).

Uma das melhorias apresentadas é: “Progredi profissionalmente (p. ex., posição mais elevada, melhoria de salários ou de complementos) ”.

Dos valores válidos 214 dos 293 diplomados indicou que o curso não trouxe progressão profissional no emprego que exercia antes e 79 indicou que o curso de facto melhorou na progressão profissional.

Tabela 90: Frequência de diplomados que escolheram a melhoria 1

**Melhorias: Progredi
profissionalmente (p. ex.,
posição mais elevada, melhoria
de salários ou de complementos)**

		Frequência
Válido	Sim	79
	Não	214
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

Melhorias: Progredi profissionalmente (p. ex., posição mais elevada, melhoria de salários ou de complementos)

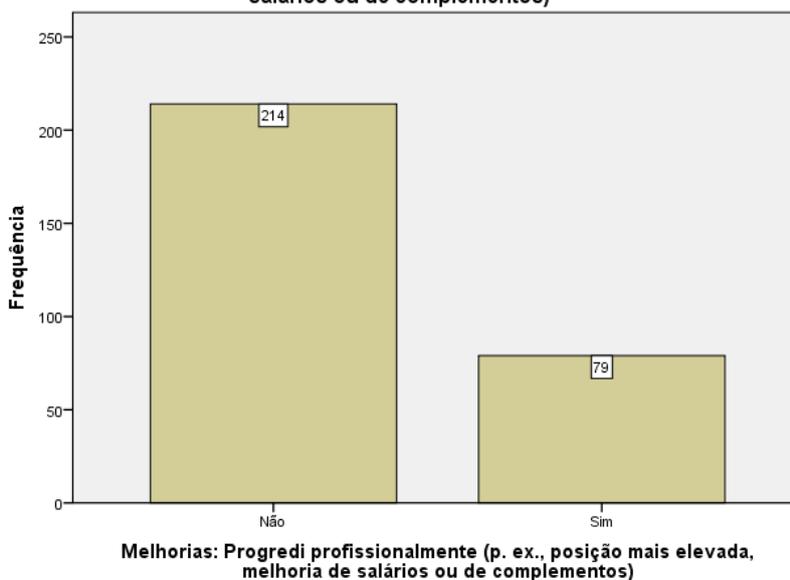


Figura 29: Gráfico de frequências de diplomados que indicaram a melhoria 1

Melhoria 2

A próxima melhoria é: “Adquiri competências/conhecimentos úteis para o meu desempenho”.

Nesta melhoria, 203 diplomados indicaram que adquiriram competências/conhecimentos úteis para o seu desempenho no emprego e 90 indicaram que não.

Tabela 91: Freqüência de diplomados que indicaram a melhoria 2

**Melhorias: Adquiri
competências/conhecimentos
úteis para o meu desempenho**

		Frequência
Válido	Não	90
	Sim	203
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

Melhorias: Adquiri competências/conhecimentos úteis para o meu desempenho

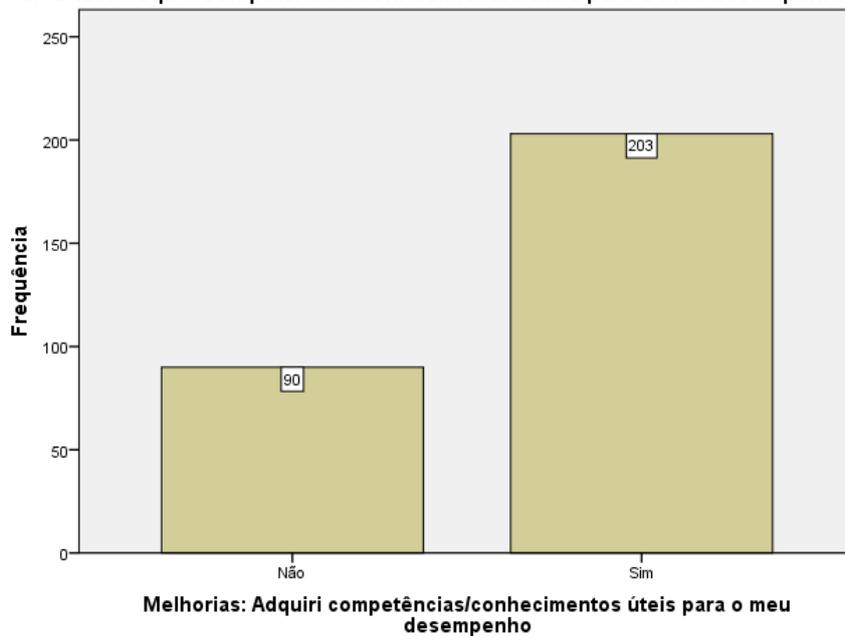


Figura 30: Gráfico de frequência de diplomados que indicaram a melhoria 3

Melhoria 3

A próxima melhoria é: “adquiri um nível de habilitações mais adequado às minhas funções”.

Neste caso, 197 dos diplomados respondeu que “não” a esta melhoria e 96 respondeu que “sim”.

Tabela 92: Frequência de diplomados que indicaram a melhoria 3

Melhorias: Adquiri um nível de habilitações mais adequado às minhas funções

		Frequência
Válido	Sim	96
	Não	197
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

Melhorias: Adquiri um nível de habilitações mais adequado às minhas funções

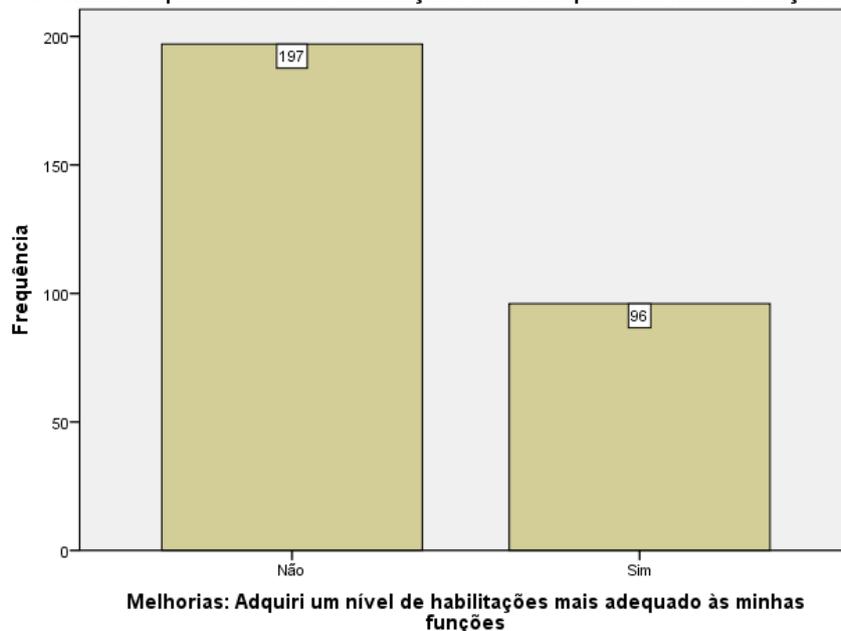


Figura 31: Gráfico de frequência de diplomados que indicaram a melhoria 3

Melhoria 4

A próxima melhoria é: “adquiri maior autonomia nas funções que desempenho”.

Grande parte (237) indicou que que não adquiriu maior autonomia nas funções que desempenha e 56 diplomados indicaram que adquiriram maior autonomia.

Tabela 93: Frequência de diplomados que indicaram a melhoria 4

Melhorias: Adquiri maior autonomia nas funções que desempenho

		Frequência
Válido	Sim	56
	Não	237
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

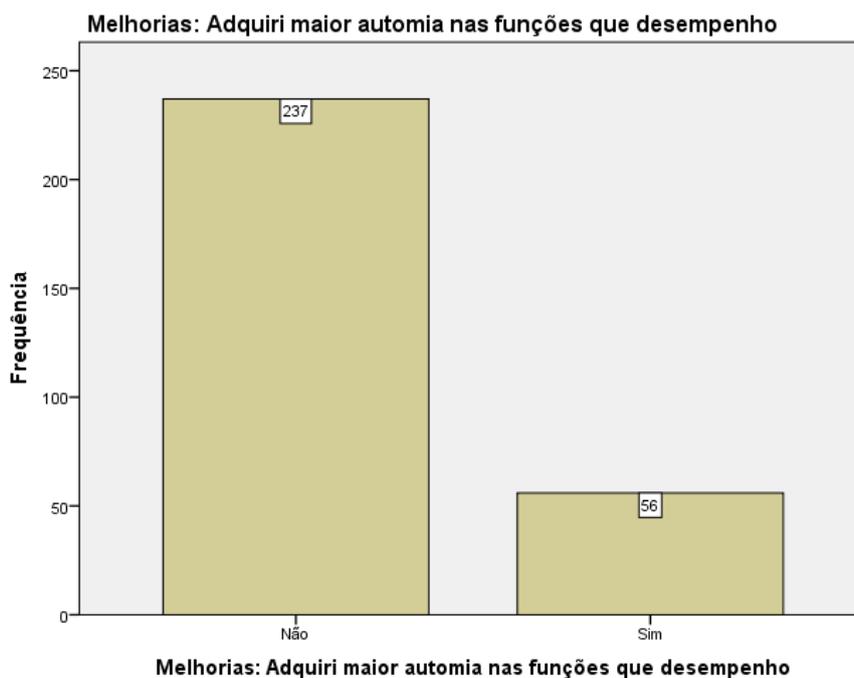


Figura 32: Gráfico de frequência de diplomados que indicaram a melhoria 4

Melhoria 5

A melhoria seguinte é: “As minhas funções mudaram (tipo e/ou diversidade)”.

Do qual, 259 dos diplomados indicaram que o curso não mudou as suas funções e 34 indicaram o oposto.

Tabela 94: Frequência de diplomados que indicaram a melhoria 5

Melhorias: As minhas funções mudaram (tipo e/ou diversidade)

		Frequência
Válido	Sim	34
	Não	259
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

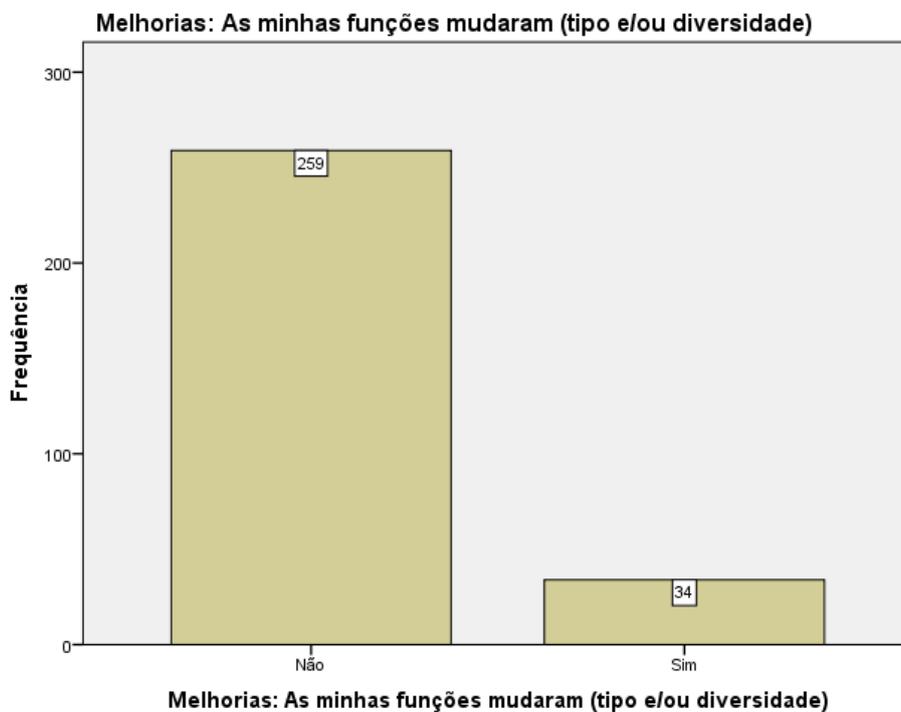


Figura 33: Gráfico de frequência de diplomados que indicaram a melhoria 5

Melhoria 6

A melhoria seguinte é: “O curso/instituição que frequentei trouxe-me maior reconhecimento/prestígio na minha área de trabalho”.

Nesta situação, grande parte dos diplomados indicou que o curso não trouxe maior reconhecimento/prestígio para a área de trabalho que tinha.

Tabela 95: Frequência de diplomados que indicaram a melhoria 6

**Melhorias: O curso/instituição
que frequentei trouxe-me maior
reconhecimento/prestígio na
minha área de trabalho**

		Frequência
Válido	Sim	74
	Não	219
	Total	293
Ausente	na	3117
	nr	68
	Total	3185
Total		3478

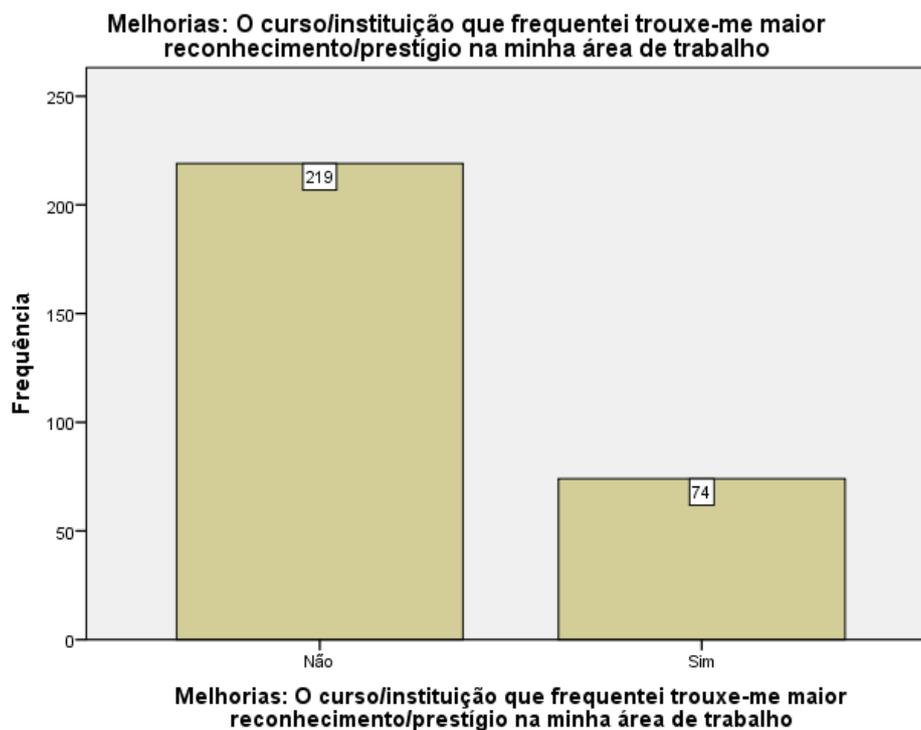


Figura 34: Gráfico de frequência de diplomados que indicaram a melhoria 6

Outra melhoria

A melhoria seguinte é: "Outra".

Grande parte dos diplomados não indicou outra melhoria e 34 diplomados indicaram outra melhoria.

Tabela 96: Frequência de diplomados que indicaram outra melhoria

Melhorias: Outra		Frequência
Válido	Sim	34
	Não	257
	Total	291
Ausente	na	3118
	nr	69
	Total	3187
Total		3478

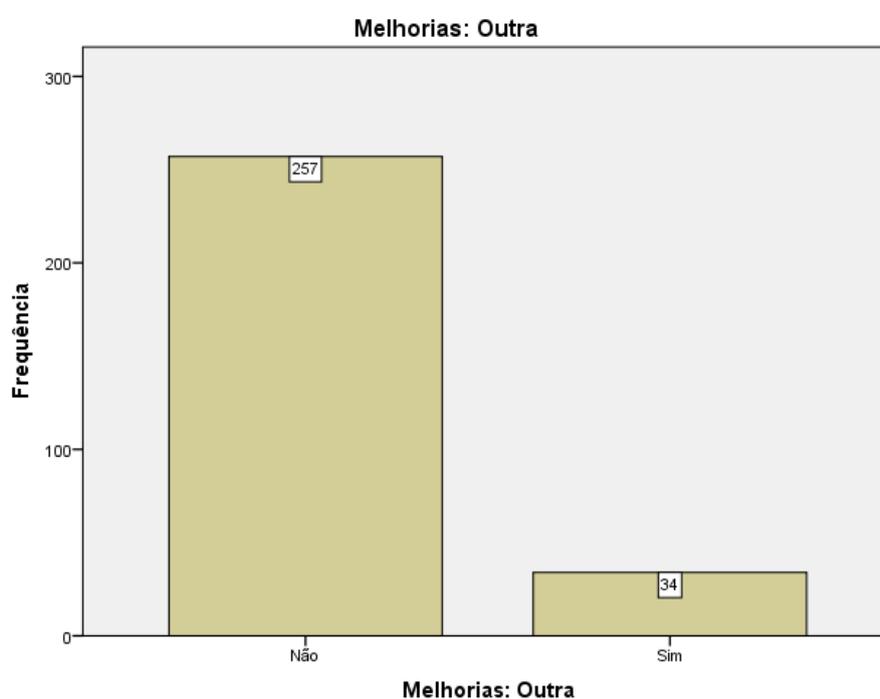


Figura 35: Gráfico de frequência de diplomados que indicaram outra melhoria

Outra melhoria, qual?

Caso a melhoria fosse “outra”, foi pedido que indicassem que outra melhoria o curso trouxe para o trabalho que exercia antes de ter iniciado o curso.

Foram indicadas no total 26 outras melhorias pelos inquiridos. No geral, as respostas dadas pelos inquiridos indicam que não houve quaisquer melhorias no trabalho que exerciam antes, após a conclusão do curso.

Tabela 97: Tabela com outras melhorias indicadas pelos diplomados

Melhorias: Se "Outra", qual?		Frequência
Válido	...	1
	Adquiri competências para a melhoria no meu desempenho para uma área paralela à minha profissão	1
	Adquiri competências/conhecimentos para o meu desenvolvimento pessoal e profissional	1
	Bonificação de 1 ano na carreira	1
	Crescimento/conhecimento/interesse meramente pessoal	1
	Estou a exercer funções docentes no ensino politécnico, a tempo parcial, em acumulação com o emp	1
	fui trabalhar na área do curso	1
	Maiores conhecimentos, não relacionados particularmente com a profissão.	1
	NADA	1
	Não foi fator evolutivo na empresa.	1
	Não houve alteração a nível profissional	1
	Não interferiu com a minha atividade profissional.	1
	Não mudou nada.	1
	não noto diferença, actualmente	1
	não se registaram alterações	1
	Não teve nenhum impacto directo	1
	Não tive qualquer alteração profissionalmente	1
	Não trouxe alterações	1
	não trouxe melhorias	1
	Não trouxe melhorias, a não ser o maior respeito com que as pessoas olham para ti.	1
	Nenhum	1
	Nenhuma	2
	O curso não teve qualquer influência na minha actividade profissional	1
	sem alteração	1
	sem evolução	1
	Total	26

Ausente	na	3376
	nr	76
	Total	3452
Total		3478

Se já esteve empregado após conclusão do curso (SAPR)

A variável seguinte é focada nos diplomados sem atividade profissional remunerada (SAPR) e estuda se o diplomado já esteve empregado após ter concluído o curso.

Uma grande parte dos diplomados (402) indicaram que não tiveram empregados desde que terminaram o curso e 185 diplomados indicaram que já tiveram empregados após conclusão do curso mas neste momento encontram-se SAPR.

Tabela 98: Frequência de diplomados que já tiveram empregados após o curso (Apenas SAPR)

**Já esteve empregado após ter
concluído o curso? (Sim/Não)
(apenas para SAPR)**

		Frequência
Válido	Sim	185
	Não	402
	Total	587
Ausente	na	2546
	nr	345
	Total	2891
Total		3478

Já esteve empregado após ter concluído o curso? (Sim/Não) (apenas para SAPR)



Figura 36: Gráfico de frequência de diplomados que já tiveram empregados após o curso

Primeiro emprego

Esta variável estuda se o atual emprego do diplomado é o seu primeiro emprego desde que terminou o curso.

Com a tabela a baixo é possível observar que os resultados são semelhantes, onde 1137 dos diplomados indicaram que o seu emprego atual é o primeiro emprego após a conclusão do curso e 1017 indicou que não é o mesmo emprego.

Tabela 99: Frequência de diplomados onde o seu emprego atual é o 1º emprego após conclusão do curso

O seu atual emprego é o primeiro emprego desde que terminou o curso?

		Frequência
Válido	Não	1017
	Sim	1137
	Total	2154
Ausente	na	1229
	nr	95
	Total	1324
Total		3478

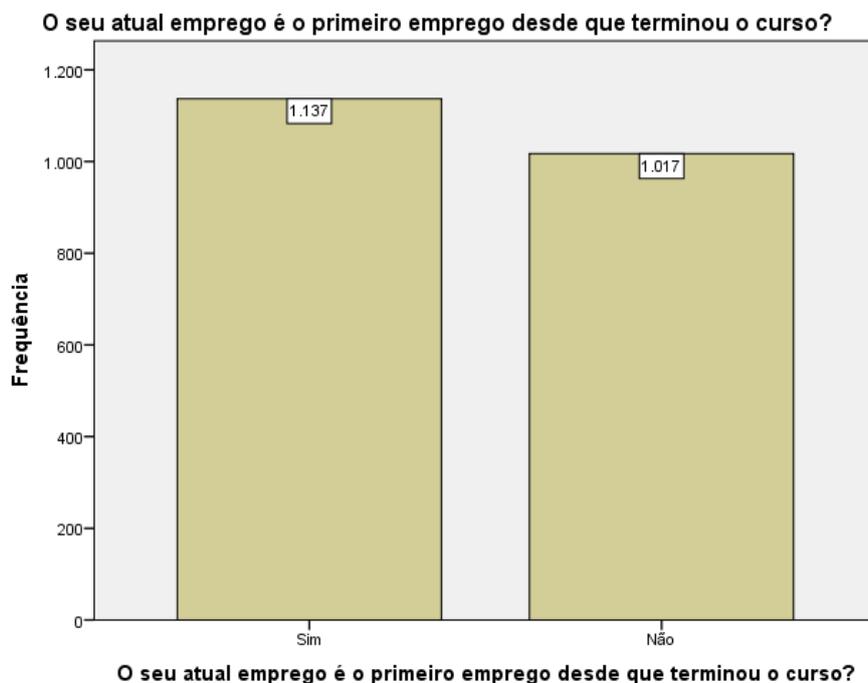


Figura 37: Gráfico de frequência de diplomados onde o seu emprego atual é o 1º emprego após conclusão do curso

Tempo de obtenção de emprego

Este grupo contém 4 variáveis, onde a primeira estuda o tempo entre a conclusão do curso e o primeiro emprego. Caso o diplomado já tenha outro emprego, a segunda variável estuda o tempo entre a conclusão do curso e o primeiro emprego da altura. A terceira variável corresponde ao tempo decorrido entre a conclusão do curso e o primeiro emprego, sendo a variável de interesse para o estudo. A última variável estuda o mesmo mas com tempos aglomerados.

Tempo entre a conclusão do curso e o atual emprego

Esta variável representa o tempo entre a conclusão do curso e o atual emprego, caso o emprego seja o primeiro após a conclusão do curso. Nos dados esta variável está indicada como “não considerar”.

Na tabela a baixo é possível observar que 251 dos diplomados arranhou emprego antes de terminar o curso, seguido de 208 que indicaram ter arranjado o primeiro emprego até 3 meses após terminar o curso.

Os valores indicados com 998 correspondem aos ‘na’ e 999 corresponde a ‘não respondeu’.

Tabela 100: Frequência de diplomados onde o seu emprego atual é o 1º emprego após conclusão do curso

Tempo (Não considerar) entre a conclusão do curso e o atual emprego

		Frequência
Válido	Até 12 meses após terminar o curso	129
	Mais de 12 meses após terminar o curso	159
	Até 1 mês após terminar o curso	163
	Até 6 meses após terminar o curso	205
	Até 3 meses após terminar o curso	208
	Antes de terminar o curso	251
	Total	1115
	Ausente	998
999		117
Total		2363
Total		3478

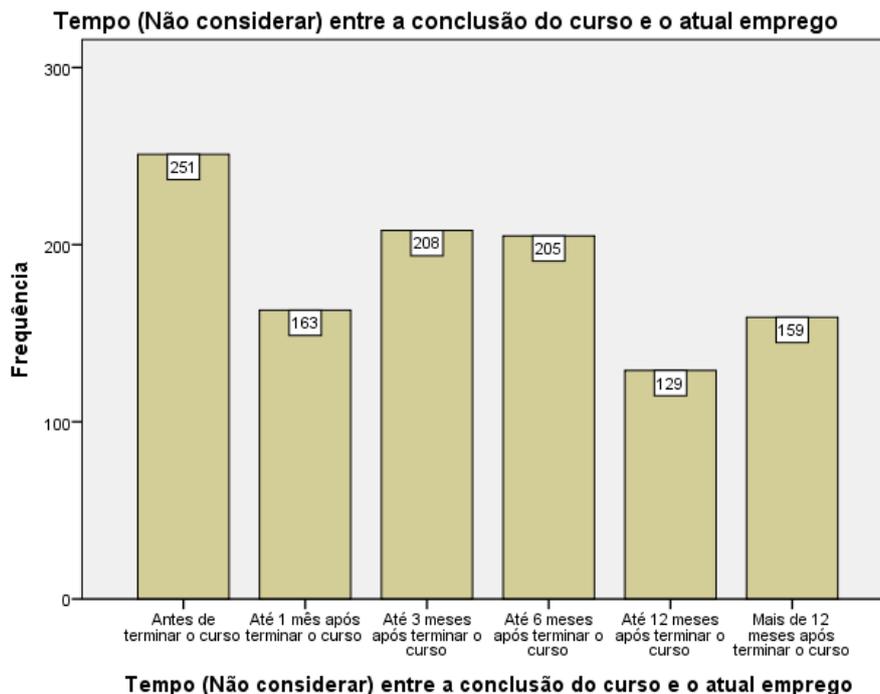


Figura 38: Gráfico de frequência de diplomados onde o seu emprego atual é o 1º emprego após conclusão do curso

Tempo entre a conclusão do curso e o primeiro emprego

Caso o atual emprego não seja o primeiro após ter terminado o curso, é pedido ao inquirido que indique quanto tempo demorou após ter concluído o curso a encontrar o seu primeiro emprego. Esta variável também está indicada como “não considerar”.

Do total de valores válidos (1163), 409 diplomados indicaram que obtiveram o seu emprego antes de concluir o curso, seguido de 231 que indicaram obter emprego até 3 meses após terminar o curso.

Tabela 101: Frequência do tempo entre a conclusão do curso e o primeiro emprego

Tempo (Não considerar) entre a conclusão do curso e o primeiro emprego

		Frequência
Válido	Mais de 12 meses após terminar o curso	84
	Até 12 meses após terminar o curso	105
	Até 1 mês após terminar o curso	156
	Até 6 meses após terminar o curso	178
	Até 3 meses após terminar o curso	231
	Antes de terminar o curso	409
	Total	1163

Ausente	998	2155
	999	160
	Total	2315
Total		3478

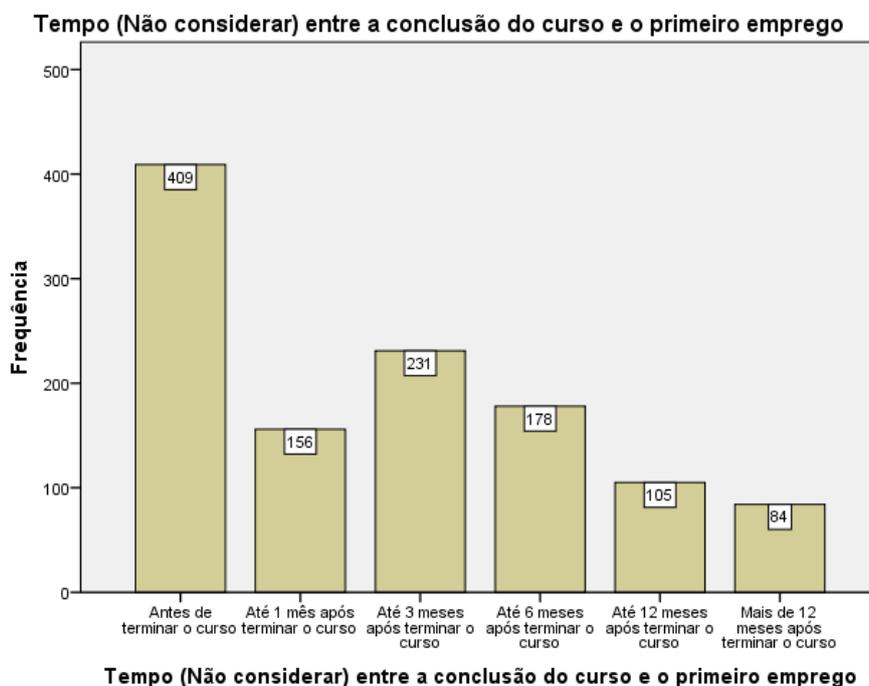


Figura 39: Gráfico de frequência do tempo entre a conclusão do curso e o primeiro emprego

Tempo decorrido entre a conclusão do curso e o primeiro emprego

Esta variável é sobre o tempo decorrido entre a conclusão do curso e o primeiro emprego e é considerada como “interessante”, dando a entender que esta é a variável a usar.

Como as tabelas anteriores e a tabela a baixo indicam, grande parte dos diplomados obteve o seu primeiro emprego antes de concluir o curso, seguido de até 3 meses após terminar o curso.

Os valores ausentes correspondem aos estudantes sem 1º emprego e aos que não responderam.

Tabela 102: Frequência do tempo decorrido entre a conclusão do curso e o primeiro emprego

Tempo (Interessa) decorrido entre a conclusão do curso e o primeiro emprego

	Frequência	
Válido	Até 1 ano	231
	Mais de 1 ano	238
	sem 1º emprego	250
	Até 1 mês	312
	Até 6 meses	376
	Até 3 meses	435
	Antes de terminar o curso	953
	Total	2795

Ausente	Estudante sem 1º emprego	498
	nr	185
	Total	683
Total		3478

Tempo (Interessa) decorrido entre a conclusão do curso e o primeiro emprego

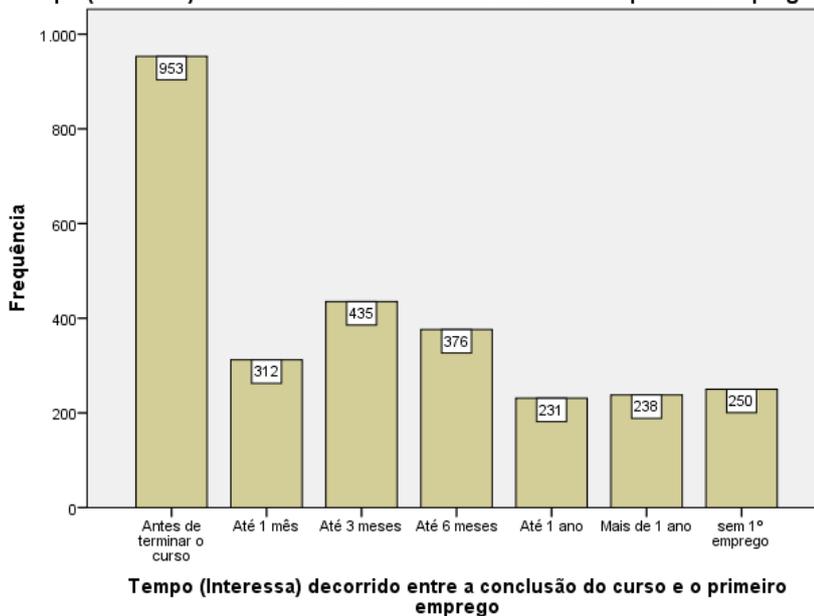


Figura 40:Gráfico de frequência do tempo decorrido entre a conclusão do curso e o primeiro emprego

Tempo decorrido entre conclusão do curso e o primeiro emprego

Esta variável estuda o tempo decorrido entre a conclusão do curso e o primeiro emprego é apresentado aglomeradamente.

Nesta situação, 1123 dos diplomados obtiveram o seu primeiro emprego até 6 meses após terminado o curso, seguido de 953 que obtiveram antes de terminar o curso.

Tabela 103:Frequência do tempo decorrido entre a conclusão do curso e o primeiro emprego (aglomerado)

Tempo decorrido (aglomerado) entre a conclusão do curso e o primeiro emprego total

		Frequência
Válido	Até 1 ano	231
	Mais de 1 ano	238
	sem 1º emprego	250
	Antes de terminar o curso	953
	Até 6 meses	1123
	Total	2795
Ausente	Estudante sem 1º emprego	498
	nr	185
	Total	683
Total		3478

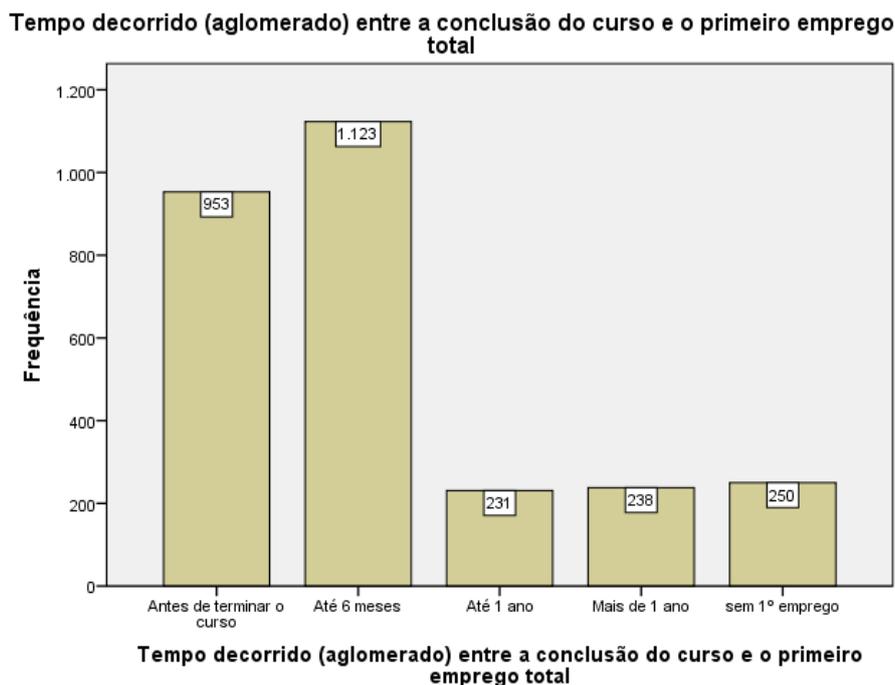


Figura 41: Gráfico de frequência do tempo decorrido entre a conclusão do curso e o primeiro emprego (aglomerado)

Continuação dos estudos

Este grupo tem 7 categorias que no total contem 17 variáveis. A primeira categoria procura saber se o diplomado prosseguiu ou prossegue estudos.

A segunda categoria visa a estudar o nível de formação, do qual contém 2 variáveis. A primeira variável corresponde ao tipo de nível de formação adicional. A segunda corresponde a outros níveis de formação, caso não estejam presentes na primeira variável.

A terceira categoria estuda a instituição onde o diplomado prossegue os estudos, contém 3 variáveis. A primeira corresponde às Instituições de Ensino Superior (IES) onde o diplomado prosseguiu ou prossegue os seus estudos. A segunda variável estuda se o diplomado prossegue os estudos na mesma escola ou não. A última visa a estudar se a instituição é pública ou privada.

A quarta categoria foca na instituição pública, do qual contem 5 variáveis. A primeira variável estuda se a instituição pública é universitário, politécnico ou militar. A próxima estuda a instituição caso esta seja Universitária, seguido de qual a escola da ULisboa. A variável seguinte é focada no Politécnico e a última focada na instituição militar.

A quinta categoria foca na instituição privada, do qual contem 3 variáveis. A primeira variável estuda se a instituição é universitária ou politécnico, as duas variáveis seguintes focam-se com mais detalhe em cada uma delas.

A sexta categoria corresponde a outra instituição, caso nenhuma das opções anteriores esteja de acordo com a escola onde o diplomado está a tirar o seu curso adicional.

A sétima categoria estuda a área de estudos que o diplomado está a prosseguir, esta contem 2 variáveis. A primeira estuda a designação da área de estudos que o diplomado prossegue, a

variável seguinte corresponde a uma designação de área de estudo caso não esteja presente na primeira variável.

Intenção

Esta variável analisa se o diplomado após conclusão do curso, prosseguiu ou se prossegue estudos.

Os resultados são semelhantes, onde 1613 diplomados dizem não estar a prosseguir estudos e 1675 dizem estar a prosseguir.

Tabela 104: Frequência de diplomados que prosseguiu ou prossegue estudos

Prosseguiu ou está a prosseguir estudos (Sim/Não)		Frequência
Válido	Não	1613
	Sim	1675
	Total	3288
Ausente	nr	190
Total		3478

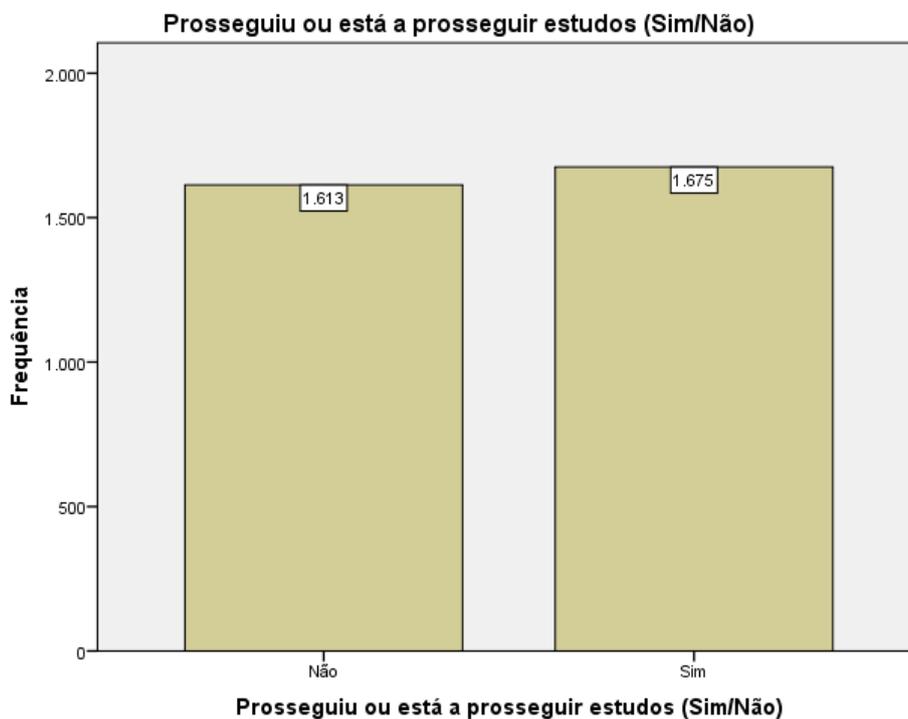


Figura 42: Gráfico de frequência de diplomados que prosseguiu ou prossegue estudos

Nível de formação

Este grupo contém 2 variáveis, onde a primeira estuda o tipo de nível de formação que o diplomado está ou esteve a prosseguir. Caso o tipo de nível de formação não esteja presente na lista do inquirido, é pedido que indiquem que tipo de nível de formação é que prossegue.

Tipo de nível de formação

Caso o diplomado tenha prosseguido ou se ainda prossegue estudos, é pedido que escolham o nível de formação adicional.

Com a tabela a baixo, é possível observar que os diplomados que prosseguem/prosseguiram estudos foi ao nível de mestrado, seguido de pós-graduação.

Tabela 105: Frequência de diplomados por nível de formação adicional

Nível de formação adicional (PG/L/M/D/Outro)		Frequência
Válido	Outro	72
	Licenciatura	83
	Doutoramento	198
	Pós-graduação	210
	Mestrado	1080
	Total	1643
Ausente	na	1613
	nr	222
	Total	1835
Total		3478

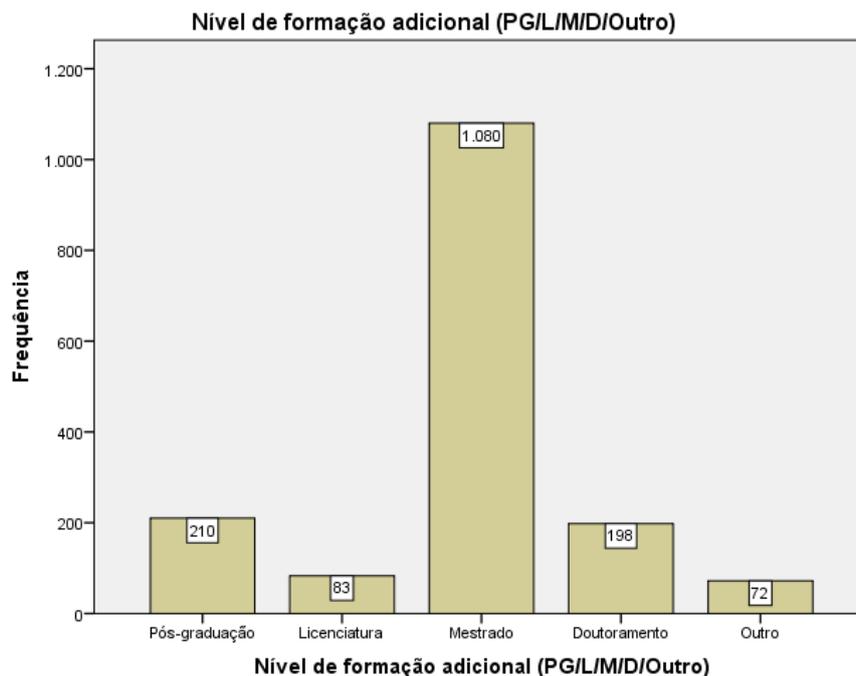


Figura 43: Gráfico frequência de diplomados por nível de formação adicional

Outro tipo de nível de formação

Se o diplomado indicou outro tipo de nível de formação adicional, que tipo de formação é?

Foram 69 diplomados que indicaram outro nível de formação adicional. Algumas das respostas são semelhantes mas escritas de maneira ligeiramente diferente, por exemplo: "Formação Contínua" e "Formação contínua", e quando calculado a frequência, estas serão consideradas como respostas diferentes.

Tabela 106: Outro tipo de nível de formação adicional

Nível de formação adicional: Se "Outro", qual?		Frequência
Válido	"De momento estou a beneficiar do apoio do programa CoopJovem (6 meses de formação) que promove a criação do próprio emprego através da constituição de uma cooperativa a jovens até aos 30 anos. A parte o programa, tenho feito Formações de curto prazo	1
	"Fiz o curso da Academia Java da Rumos, obtendo os seguintes certificados:	1

Bacharel	1
Bolsa de investigação de mestre.	1
CCP	1
Certificação	1
Certificação em PNL	1
CFA	1
Curso de animação	1
Curso de desenvolvimento profissional	1
Curso de Especialização	2
curso de formação em massagem	1
Curso de inglês e de Licenciamento Ambiental com carga horaria de 600 horas	1
Curso de Língua Chinesa	1
Curso de programação JAVA para titulares de cursos superiores na área da biotecnologia	1
Curso diferente	1
Curso Formação Profissional	1
Curso profissional	2
Curso Profissional	2
Curso Profissional de Fotografia	1
Curso profissional de Técnico de Informática - Sistemas	1
Curso profissional prático	1
Curso técnico	1
Curso Técnico	1
Cursos online para completar a formação.	1
EFA	1
Especialidade	1

Especialidade de Oncologia Médica	1
Especialidade em Cirurgia Geral	1
Especialidade medica	1
Especialização	2
Especialização Profissional	1
Estágio na Ordem dos Advogados, quando concluir pretendo prosseguir os estudos - em mestrado - na mesma área.	1
Formação Avançada	4
Formação base - Nível 1 e 2 de Fotografia	1
Formação contínua	1
Formação Contínua	1
formação contínua 50h ou 25h	1
Formação em área diferente.	1
Formação em BIM	1
Formação específica (na área médica)	1
Formação profissional	1
Formação Profissional	1
Formação profissional na área hoteleira	1
Formação profissional Segurança no Trabalho	1
Formação profissional.	1
Formação Profissional.	1
Ife	1
Ingresso no Centro de Estudos Judiciários	1
Internato Médico	4

	Internato médico para especialização em Saúde Pública	1
	Licenciatura com mestrado integrado	1
	Master em Futebol	1
	Membro do IFoA a obter qualificações para ser Actuario certificado.	1
	Ordem Advogados	1
	Ordem dos advogados	1
	Programa Executivo	1
	Qualificacao como Atuario	1
	Taxonomia e Ecologia de zooplâncton marinho; Taxonomia e ecologia de macrofauna bentónica marinha; 2011-2020 Biodiversidade - Educar para Conservar; Auxiliar de Veterinária	1
	Total	69
Ausente	na	3184
	nr	225
	Total	3409
Total		3478

Instituição

Este grupo contém 3 variáveis, onde a primeira estuda as Instituições de Ensino Superior (IES) de prosseguimento de estudos. A segunda questiona se o prosseguimento de estudos foi ou é concretizada na mesma faculdade/instituição da diplomação inicial. A última variável estuda se a instituição onde o diplomado está ou esteve a prosseguir estudos é pública ou privada.

Instituições de Ensino Superior (IES) de prosseguimento de estudos

Esta variável corresponde às Instituições de Ensino Superior (IES) onde o diplomado prosseguiu ou prossegue os seus estudos.

Os diplomados que continuaram com os seus estudos, 940 indicaram que continuaram os estudos na mesma escola, seguido de 151 que continuaram noutra universidade pública. Politécnico privado é a IES com menor frequência, tendo 5 diplomados.

Foram 1613 os diplomados que disseram que não prosseguiram estudos e 362 não responderam.

Tabela 107: Frequência de IES de prosseguimentos de estudos

IES de prosseguimento de estudos		Frequência
Válido	Politécnico Privado	5
	Politécnico Público	41
	Outra Univ. Privada	50
	ISCTE	52
	Univ. Católica	58
	Outra Escola ULisboa	101
	UNL	105
	Outra Univ. Pública	151
	Mesma Escola	940
	Total	1503
Ausente	Não prossegue estudos	1613
	nr	362
	Total	1975
Total		3478

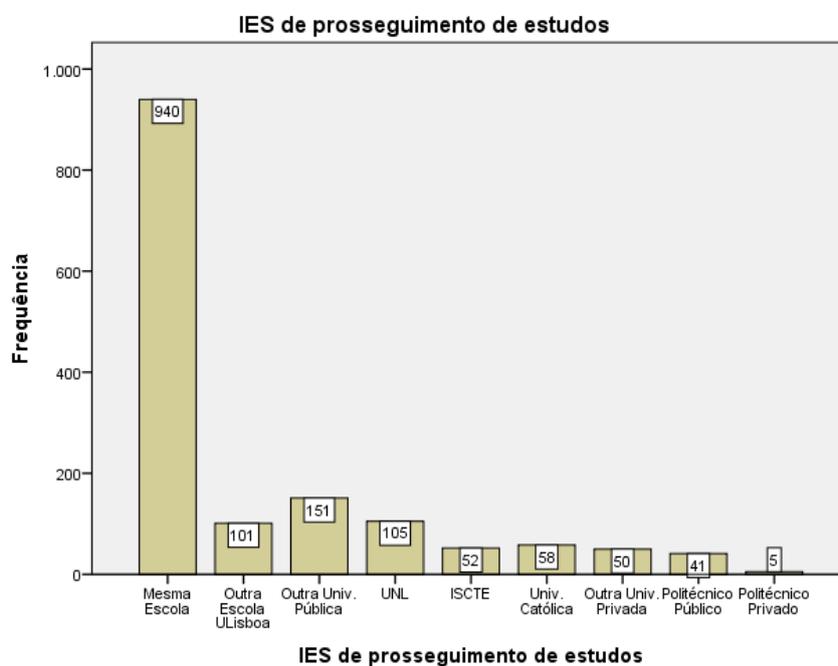


Figura 44: Gráfico frequência de IES de prosseguimentos de estudos

Mesma Faculdade/Instituição da diplomação inicial

Esta variável estuda o diplomado que prosseguiu os seus estudos e se o fez na mesma faculdade/instituição onde concluiu o curso.

Com a tabela a baixo é possível observar que 940 respondeu que sim e 691 que não.

Tabela 108: Frequência de diplomados que prosseguiram os estudos na mesma escola

Na mesma Faculdade/Instituto da diplomação inicial (Sim/Não)		Frequência
Válido	Não	691
	Sim	940
	Total	1631
Ausente	na	1613
	nr	234
	Total	1847
Total		3478

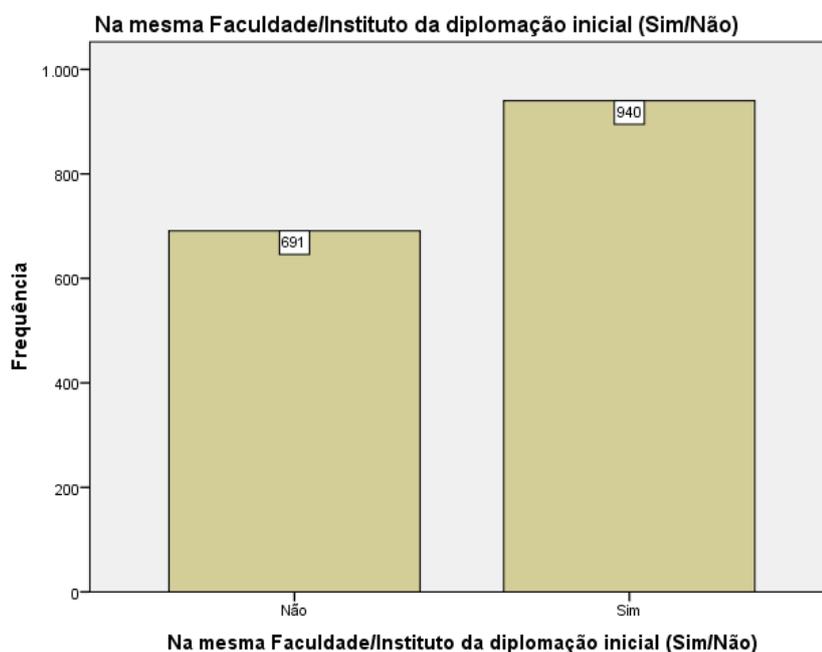


Figura 45: Gráfico frequência de diplomados que prosseguiram os estudos na mesma escola

Pública/Privada

De seguida é estudado se a instituição onde o diplomado prosseguiu os seus estudos, é pública ou privada.

Dos 627 diplomados que responderam, 471 indicou frequentar uma instituição pública e 156 disse ser privada.

Tabela 109: Frequência de diplomados por instituição pública ou privada

Instituição (Pública/Privada)		Frequência
Válido	Privada	156
	Pública	471
	Total	627
Ausente	na	2496
	nr	355
	Total	2851
Total		3478

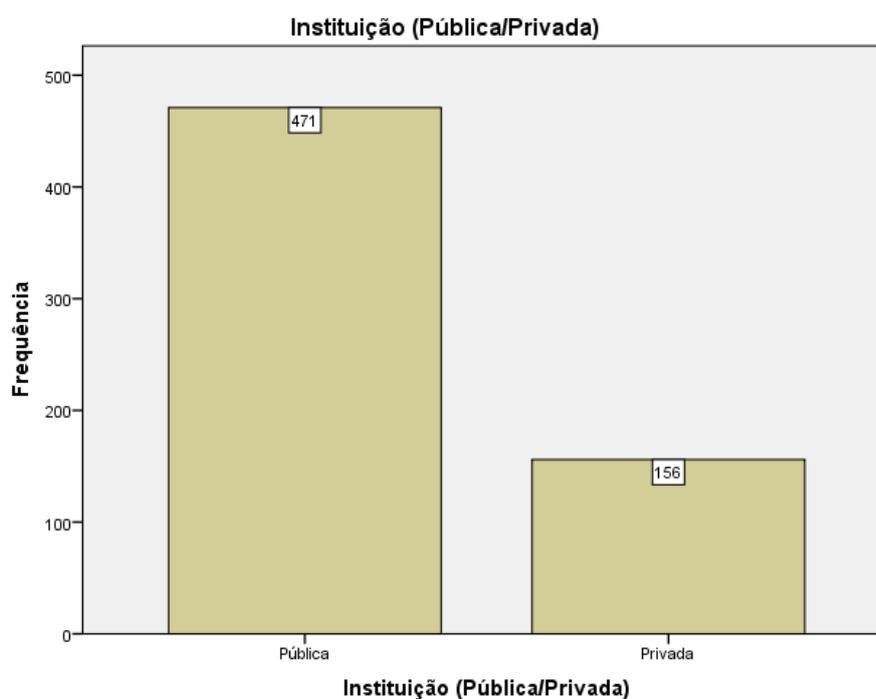


Figura 46: Gráfico frequência de diplomados por instituição pública ou privada

Instituição pública

Este grupo contém 5 variáveis, onde a primeira estuda se o diplomado prosseguiu ou prossegue os seus estudos numa instituição pública Universitário, Politécnico ou Militar. As 3 variáveis seguintes são focadas para cada uma das opções anteriores. Caso o diplomado esteja a estudar numa escola ULisboa é pedido que indicasse essa escola.

Universitário, Politécnico ou Militar

Caso o diplomado tenha prosseguido por uma instituição pública, é estudado se esta instituição é do tipo Universitário, Politécnico ou Militar.

A maioria dos diplomados indicou frequentar numa instituição universitária.

Tabela 110: Frequência de diplomados por tipo de ensino público

Ensino Público:		
Universitário/Politécnico/Militar		
		Frequência
Válido	Militar	3
	Politécnico	41
	Universitário	406
	Total	450
Ausente	na	2654
	nr	374
	Total	3028
Total		3478

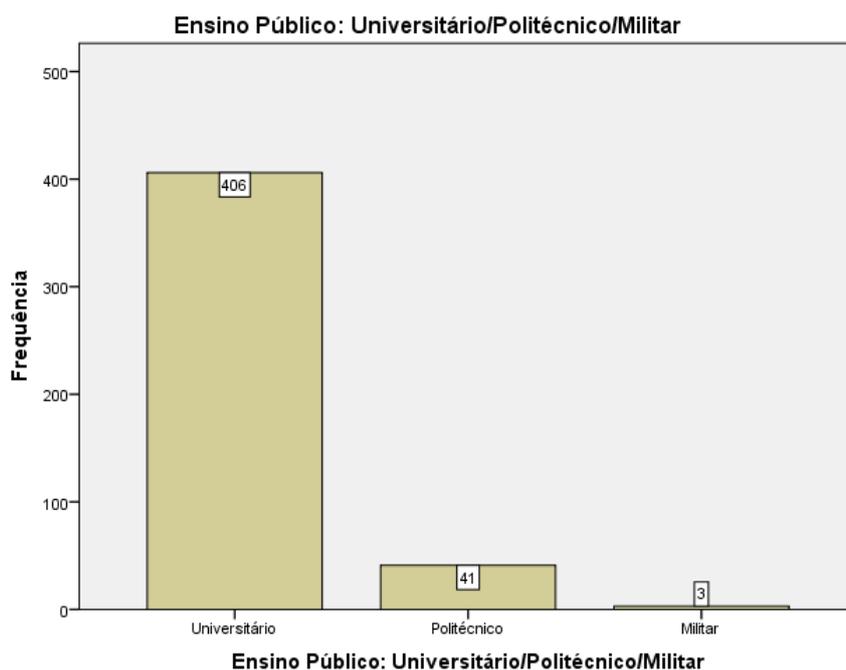


Figura 47: Gráfico frequência de diplomados por tipo de ensino público

Universidade pública

Se o diplomado indicou prosseguir os estudos numa universidade pública, é questionado que universidade é essa.

Existem valores de frequência aproximados entre a Universidade Nova de Lisboa, Universidade de Lisboa e outra universidade. As universidades que apresentam um menor número de

diplomados são a Universidade do Minho, Universidade de Aveiro e Universidade de Trás-os-Montes e Alto Douro.

Tabela 111: Frequência de diplomados por universidades públicas

Prosseguimento de estudos em Universidade Pública		Frequência
Válido	Universidade do Minho	1
	Universidade de Aveiro	2
	Universidade de Trás-os-Montes e Alto Douro	2
	Universidade de Évora	5
	Universidade do Algarve	7
	Universidade de Coimbra	7
	Universidade do Porto	16
	ISCTE	52
	Universidade de Lisboa	101
	Outro	104
	Universidade Nova de Lisboa	105
	Total	402
	Ausente	na
nr		380
Total		3076
Total		3478

Escola ULisboa

Foi questionado em que escola da ULisboa o diplomado prosseguiu os seus estudos.

Do total de respostas, 40 diplomados indicaram que prosseguiram os seus estudos no Instituto Superior de Economia e Gestão da ULisboa (ISEG), seguido da Alameda do Instituto Superior Técnico (IST alameda).

Tabela 112: Frequência de diplomados por escolas da ULisboa

Prosseguimento de estudos em Escola da ULisboa		Frequência
Válido	FA	1
	FMH	1
	FP	1
	IGOT	1
	FBA	1

	IST (Taguspark)	3
	FC	3
	FD	3
	FF	3
	FMV	3
	IE	4
	ISA	4
	FL	4
	ISCSP	5
	FM	7
	IST (Alameda)	15
	ISEG	40
	Total	99
Ausente	na	2998
	nr	381
	Total	3379
Total		3478

Politécnico público

Foi questionado em que escola o diplomado prosseguiu os seus estudos caso tenha sido num Politécnico Público.

Do total de respostas, 15 diplomados indicaram que prosseguiram os estudos no Instituto Politécnico de Lisboa, seguido do Instituto Politécnico de Leiria.

Tabela 113: Frequência de diplomados por Politécnico público

Prosseguimento de estudos em Politécnico Público

		Frequência
Válido	Instituto Politécnico de Portalegre	1
	Instituto Politécnico de Setúbal	1
	Instituto Politécnico do Cávadp e do Ave	1
	Instituto Politécnico do Porto	1
	Universidade do Algarve	1
	Instituto Politécnico da Bragança	1

	Instituto Politécnico de Castelo Branco	1
	Instituto Politécnico de Santarém	2
	Outro	2
	Instituto Politécnico de Tomar	3
	Escola Superior de Hotelaria e Turismo do Estoril	5
	Instituto Politécnico de Leiria	7
	Instituto Politécnico de Lisboa	15
	Total	41
Ausente	na	3063
	nr	374
	Total	3437
Total		3478

Instituição Militar pública

Foi questionado em que escola o diplomado prosseguiu os seus estudos caso tenha sido numa Instituição Militar pública.

Houve 3 respostas para este caso, onde um diplomado seguiu na Academia da Força Aérea, outro na Academia Militar e o último indicou ter prosseguido noutra instituição diferente das apresentadas.

Tabela 114: Frequência de diplomados por Instituição Militar pública

Prosseguimento de estudos em Instituição Militar

Pública

		Frequência
Válido	Academia da Força Aérea	1
	Academia Militar	1
	Outro	1
	Total	3
Ausente	na	3101
	nr	374
	Total	3475
Total		3478

Instituição privada

Este grupo contém 3 variáveis, onde a primeira estuda se a instituição privada é do tipo Universitário ou Politécnico. As duas variáveis seguintes estudam com maior detalhe cada uma das opções anteriores.

Universitário ou Politécnico

Se o diplomado indicou ter prosseguido os seus estudos numa instituição privada, foi questionado o tipo de instituição, se Universitário ou Politécnico.

Do total de respostas, 108 indicaram que prosseguiram os seus estudos numa Universidade Privada e 5 indicaram que prosseguiram num Politécnico Privado.

Tabela 115: Frequência de diplomados por tipo de Ensino privado

Ensino Privado: Universitário/Politécnico		Frequência
Válido	Politécnico	5
	Universitário	108
	Total	113
Ausente	na	2969
	nr	396
	Total	3365
Total		3478

Universidade privada

Foi questionado em que escola o diplomado prosseguiu os seus estudos caso tenha sido numa Universidade Privada.

Do total de respostas, 58 diplomados indicaram que prosseguiram os seus estudos na Universidade Católica Portuguesa, seguido de outra universidade privada.

Tabela 116: Frequência de diplomados por Universidades privadas

Prosseguimento de estudos em Universidade Privada		Frequência
Válido	Instituto Superior de Ciências da Saúde Egas Moniz	1
	Instituto Superior de Educação e Trabalho	1
	Instituto Superior de Gestão	1
	ISPA	3

	Universidade Lusíada	3
	Instituto de Arte, Design e Empresa	4
	Universidade Europeia	6
	Universidade Lusófona de Humanidades e Tecnologias	9
	Outro	21
	Universidade Católica Portuguesa	58
	Total	107
Ausente	na	2972
	nr	399
	Total	3371
Total		3478

Politécnico privado

Foi questionado em que escola o diplomado prosseguiu os seus estudos caso tenha sido num Politécnico Privado.

Para este caso, apenas houveram 4 respostas no total, onde um diplomado indicou ter prosseguido os seus estudos na Escola Superior de Tecnologias e Artes de Lisboa e os outros restantes indicaram terem prosseguido noutra Politécnico privado.

Tabela 117: Frequência de diplomados por politécnico privado

Prosseguimento de estudos em Politécnico Privado		Frequência
Válido	Escola Superior de Tecnologias e Artes de Lisboa	1
	Outro	3
	Total	4
Ausente	na	3077
	nr	397
	Total	3474
Total		3478

Outra Instituição

Caso o diplomado tenha respondido que frequentou “outra” Instituição de Ensino Superior (IES), foi pedido que indicassem essa instituição.

Para este caso houve um total de 124 respostas válidas, na tabela a baixo é possível observar parte das instituições apresentadas pelos inquiridos.

Tabela 118: Parte das outras instituições apresentadas pelos diplomados

Prosseguimento de estudos em outra IES		Frequência
Válido	Vrije University Amsterdam	1
	wimbledon UAL	1
	WUR	1
	Heriot-Watt University	2
	INSTITUTO PORTUGUÊS DE PSICOLOGIA E OUTRAS CIÊNCIAS INSTITUTO	2
	University of Luxembourg	2
	Wageningen University	2
	Denmark Technical University (DTU)	3
	University College London	3
	University of Kent	3
	Total	124
	Ausente	na
nr		425
Total		3354
Total		3478

Para consultar a tabela completa, abrir o ficheiro “Tabelas completas” (Tabelas completas (com muitas entradas)).

Área de estudos

Este grupo contém duas variáveis, onde a primeira estuda a designação da área de estudos do curso em que prosseguiu os estudos. A segunda variável pede ao diplomado para indicar qual o curso que prossegue, caso a designação não esteja presente na lista anterior.

Área de estudos do curso em que prosseguiu estudos

Depois de entender em que instituição o diplomado prosseguiu os seus estudos é de interesse saber a área de estudos do curso.

Numa lista apresentada no inquérito o diplomado selecionava a área do curso. Dos 1584 diplomados que responderam a esta questão, 384 indicaram prosseguir estudos na área de engenharia excluindo informática, seguido de 220 que indicaram seguir a área de Economia e Gestão. As áreas que apresentaram um menor número de diplomados foram a de Veterinária, Arquitetura e outra área que não estava na lista.

Tabela 119:Frequência de diplomados por área de curso de estudo prosseguidos

Área de Estudos do curso em que prosseguiu estudos		Frequência
Válido	Veterinária	13
	Arquitetura	15
	Outra	16
	Farmácia	19
	Psicologia	21
	Agricultura, silvicultura e pescas	30
	Informática	35
	Educação	47
	Saúde (excl. Medicina)	49
	Serviços (Desporto, Segurança, Transportes e Turismo)	49
	Ciências Físicas e da Terra	57
	Matemática e Estatística	58
	Medicina	61
	Humanidades	71
	Artes	100
	Ciências da Vida	103
	Ciências Sociais	111
	Direito	125
	Economia e Gestão	220
	Engenharia (excl. Informática)	384
Total	1584	
Ausente	na	1613
	nr	281
	Total	1894
Total		3478

Outra área de estudos do curso em que prosseguiu estudos

A variável seguinte corresponde a “outra” área de estudos do curso que o diplomado prosseguiu caso a área não estivesse apresentada na questão anterior.

No entanto após análise estatística desta variável foi possível perceber que houve demasiadas respostas para aquelas que se estariam à espera. Pela tabela anterior, seria de esperar pelo menos 16 respostas, visto que foi o número de diplomados que responderam “outra”, mas foram obtidas 981 respostas. Depois de observar os dados é possível perceber que os diplomados escolheram a área do curso na pergunta anterior e depois detalharam o curso na pergunta que representa esta variável.

Foram apresentados 472 designações de curso diferentes, originando uma tabela muito longa. Com análise da frequência de diplomados que indicaram uma designação para o curso, 189 responderam Engenharia Informática e de Computadores, seguido de Engenharia e Gestão Industrial.

G4p29 (Se G4p23=Sim) Em qual das seguintes áreas de estudo se insere o curso?

Caixa Lista das áreas de estudo ULisboa

Outra

G4p30 (Se G4p29=Outra) Qual a designação do curso?

Caixa de resposta

Figura 48: Perguntas do inquérito sobre a área do curso dos estudos que prossegue

Tabela 120: Parte da frequência de diplomados por designação de curso que prosseguiu

**Área de Estudos do curso em que prosseguiu
estudos: Se "Outra", qual?**

		Frequência
Válido	Engenharia Mecânica	4
	Geografia	4
	Gestão e Estratégia Industrial	4
	Sociologia das Organizações e do Trabalho	4
	Arquitetura Paisagista	5
	Biologia da Conservação	5
	Bioquímica Médica	5
	Engenharia Agronómica	5
	Engenharia Informática	5
	Medicina	5
	Engenharia Electrotécnica e de Computadores	7
	Química	7
	Matemática Aplicada à Economia e Gestão	8
	Reabilitação Psicomotora	8

	Tradução	8
	Finance	9
	Mestrado Integrado em Medicina	9
	Gestão	10
	Engenharia e Arquitectura Naval	11
	Marketing	13
	Engenharia Alimentar	14
	Engenharia Geológica e de Minas	15
	Engenharia de Telecomunicações e Informática	20
	Matemática e Aplicações	20
	Engenharia e Gestão Industrial	54
	Engenharia Informática e de Computadores	189
	Total	981
Ausente	na	2210
	nr	287
	Total	2497
Total		3478

Para consultar a tabela completa, abrir o ficheiro “Tabelas completas” (Tabelas completas (com muitas entradas)).

Emprego e estudos de diplomados SAPR

Este grupo contem 3 categorias do qual estudam especificamente os diplomados sem atividade profissional remunerada e as suas situações em relação ao emprego e ao prosseguimento de estudos.

A primeira categoria contem 3 variáveis. A primeira estuda se o diplomado SAPR procura ou não emprego. Caso não procure é pedido que escolha o motivo para tal, se os motivos apresentados não justificam os motivos do inquirido é pedido que indique outra razão que o faz não estar à procura de emprego.

A segunda categoria estuda os diplomados SAPR que prosseguem ou prosseguiram estudos.

A última categoria corresponde aos diplomados que deixariam ou não de estudar se encontrassem emprego.

Procura de emprego em diplomados SAPR

Este grupo contém 3 variáveis que estudam a procura de emprego por parte do diplomado e quais os motivos para não estar à procura, caso não esteja listado o motivo, é pedido que indiquem o outro motivo.

Procura emprego

Esta variável é focada nos diplomados sem atividade profissional remunerada (SAPR), onde é analisado se estes procuram emprego ou não.

Do total de respostas, 547 respondeu que não está à procura de emprego e 353 indicaram que estão.

Tabela 121: Frequência de diplomados SAPR que procuram emprego ou não

Procura de Emprego (Sim/Não) (apenas para SAPR)		Frequência
Válido	Sim	353
	Não	547
	Total	900
Ausente	na	2546
	nr	32
	Total	2578
Total		3478

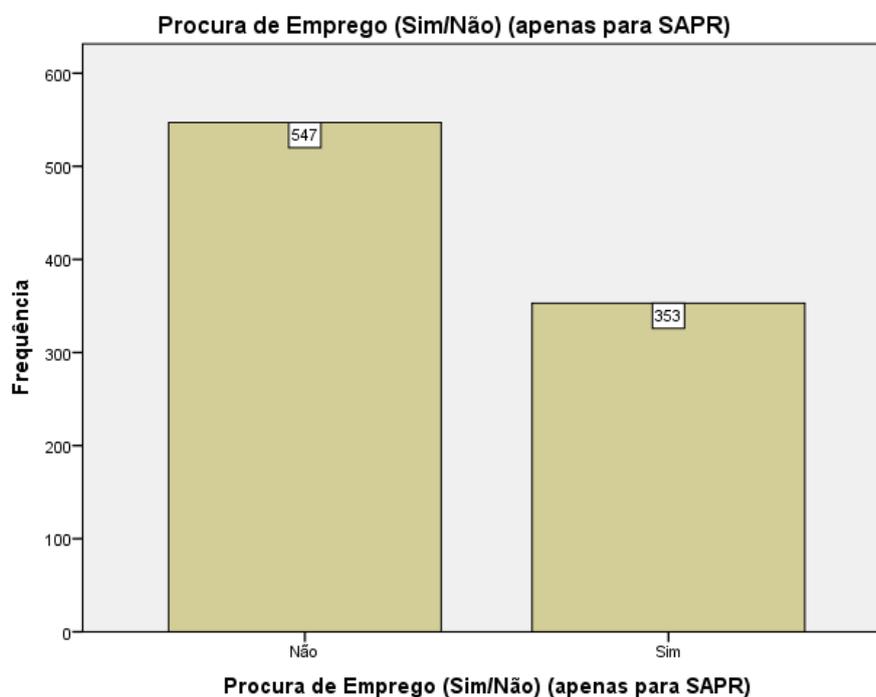


Figura 49: Gráfico de frequência de diplomados SAPR que procuram emprego ou não

Motivo principal para não procurar emprego

Para os diplomados que indicaram não procurar emprego é necessário entender os motivos para não o fazerem.

Dada uma lista de motivos principais para os diplomados não estarem à procura de emprego, estes devem escolher o que melhor se adequa.

Dos 539 diplomados que responderam, 493 indicaram que o principal motivo foi de estarem a prosseguir com estudos, seguido de 23 que indicaram estarem a realizar um estágio não remunerado. O motivo com menos frequência foi o de “Outro” tendo 10 respostas.

Tabela 122: Frequência de diplomados por motivos de não procurarem emprego

Motivo principal para não estar à procura (lista)		Frequência
(apenas para SAPR)		
Válido	Outro	10
	Já tenho uma proposta e/ou aguardo conclusão de prazos/procedimentos	13
	Encontro-me a realizar um estágio não remunerado	23
	Estou a prosseguir estudos	493
	Total	539
Ausente	na	2900
	nr	39
	Total	2939
Total		3478

Outro motivo principal para além dos listados

Se o diplomado indicou “outro” motivo para o qual não procura emprego, foi pedido para escrever esse outro motivo.

Todos os 10 diplomados que responderam como “outro” na questão anterior indicaram esse motivo nesta questão.

Tabela 123: Lista de outros motivos para não procurarem emprego

Motivo principal para não estar à procura (se "Outro" qual?) (apenas para SAPR)

		Frequência
Válido	Complicações de saúde	1
	Cuidar de familiar gravemente doente	1
	De momento estou dedicada a um negócio familiar.	1
	estou a terminar a tese	1
	Fase da criação da própria empresa. Unipessoal Ida	1
	Formação em área diferente.	1
	pretendo tirar uma nova licenciatura	1
	Projecto pessoal	1
	Reformado	1
	Trabalho domestico de cuidar filhos e desenvolvimento de projectos artisticos a base nao renumerada	1
	Total	10
	Ausente	na
nr		39
Total		3468
Total		3478

Continuação de estudos nos diplomados SAPR

É inquirido ao diplomado SAPR se após a conclusão do curso adquiriu formação adicional e não se encontra a prosseguir estudos, se encontra a prosseguir estudos ou se não prosseguiu mais estudos.

Do total de diplomados que responderam, 715 disseram estar a prosseguir com os estudos, seguido de 115 que indicaram não estarem a prosseguir com estudos e 76 disseram que adquiriram formação adicional mas de momento não se encontram a prosseguir estudos.

Tabela 124: Frequência de diplomados SAPR que prosseguem estudos ou não

Prosseguiu ou está a prosseguir estudos (apenas para SAPR)		Frequência
Válido	Adquiri formação adicional e não estou de momento a prosseguir estudos	76
	Não prossegui mais estudos	115
	Estou de momento a prosseguir estudos	715
	Total	906
Ausente	na	2546
	nr	26
	Total	2572
Total		3478

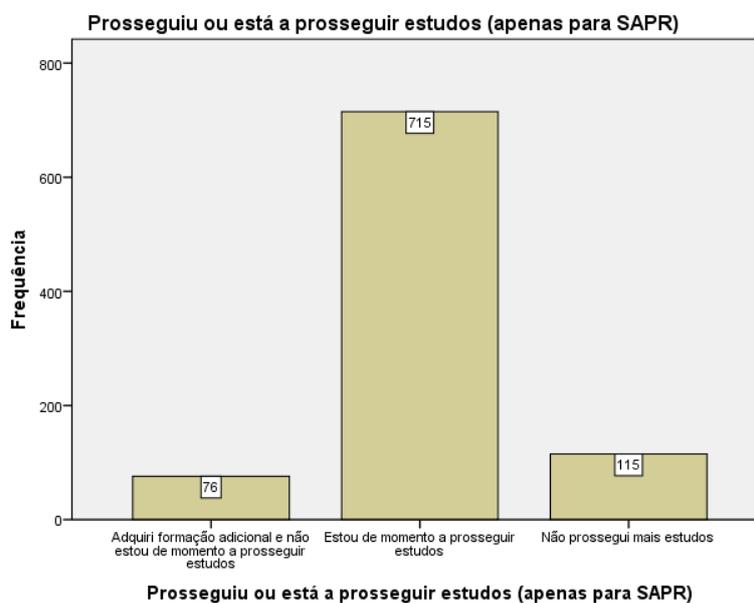


Figura 50: Gráfico de frequência de diplomados SAPR que prosseguem estudos ou não

Diplomado SAPR deixar de estudar

Esta variável tenta avaliar se um diplomado SAPR que se encontra a estudar se deixaria os estudos se uma oportunidade de emprego aparecesse.

Do total de dados válidos, a maioria dos diplomados indicaram que não deixariam de estudar, caso encontrassem emprego.

Tabela 125:Frequência de diplomados que deixariam de estudar se arranjassem emprego

Se encontrar emprego pretende deixar de estudar? (Sim/Não)		Frequência
Válido	Sim	42
	Não	263
	Total	305
Ausente	na	2737
	nr	436
	Total	3173
Total		3478

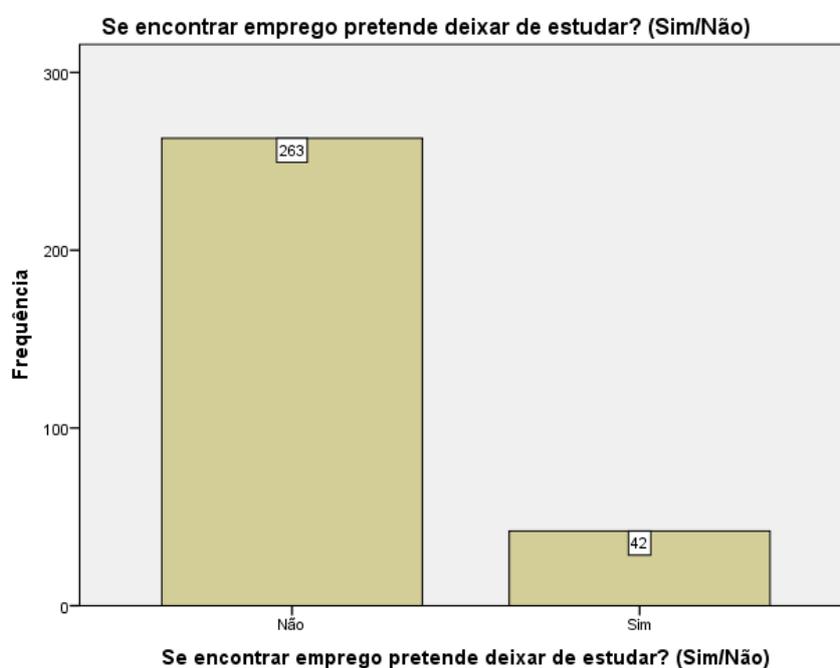


Figura 51:Gráfico de frequência de diplomados que deixariam de estudar se arranjassem emprego

Situação face ao emprego (distinguir nos SAPR os estudantes e reformados)

Esta variável corresponde à situação face ao emprego onde indica “ (distinguir nos SAPR, os estudante e reformados) ”, no entanto esta variável tem exatamente o mesmo resultado que a variável que estuda a situação face ao emprego incluindo os reformados e estudantes (Tabela 44: Frequência de diplomados face ao emprego).

Situação face ao emprego (estudante detalhado)

A última variável do ficheiro de dados disponibilizado corresponde à situação face ao emprego mas o “estudante” é detalhado para “estudante que não procura emprego” e “estudante à procura de emprego”.

Existem mais diplomados que são estudantes que não procuram emprego (514 diplomados) do que os que procuram (170 diplomados).

Tabela 126: Frequência de diplomados por situação face ao emprego, diferenciando os estudantes

Quadro A3

		Frequência
Válido	T. conta própria c/ func. a cargo	23
	Estudante à procura de emprego	170
	Bolseiro	223
	T. conta própria s/ func. a cargo	248
	SAPR	248
	Estagiário	266
	Estudante que não procura emprego	514
	T. conta de outrem	1786
	Total	3478

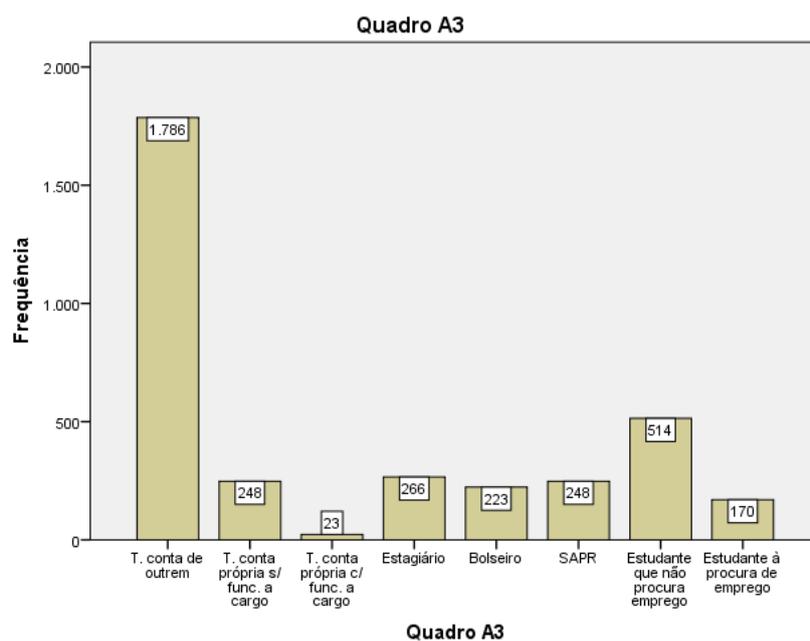


Figura 52: Gráfico de frequência de diplomados por situação face ao emprego, diferenciando os estudantes

Tabelas completas (com muitas entradas)

O documento “Tabelas completas” inclui as tabelas completas para os seguintes pontos:

- Designação do curso
- Data de nascimento
- Valor de vencimentos
- Valor de complementos remuneratórios
- Outras escolas de prosseguimentos de estudos
- Outras áreas de curso



**Tabelas
completas.htm**

