

# A Internet como repositório de dados reais para o ensino de Estatística em cursos universitários

## Internet as a repository of real data for the teaching of statistics in university courses

Fernando Frei<sup>1</sup>

<sup>1</sup> Departamento de Ciências Biológicas – Laboratório de Estatística Aplicada, Faculdade de Ciências e Letras de Assis, Universidade do Estado de São Paulo - UNESP, Assis, São Paulo, Brasil.

[ffrei@assis.unesp.br](mailto:ffrei@assis.unesp.br)

Recibido: 26/07/2017 | Corregido: 13/06/2018 | Aceptado: 23/07/2018

Cita sugerida: F. Frei, “A Internet como repositório de dados reais para o ensino de Estatística em cursos universitários,” *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, no. 22, pp. 18-27, 2018. doi:10.24215/18509959.22.e02

Esta obra se distribuye bajo **Licencia Creative Commons CC-BY-NC 4.0**

### Resumo

A alfabetização estatística é uma habilidade chave esperada dos cidadãos em sociedades carregadas de informação. Estudos mostram que os cursos de serviço de estatística são ainda direcionados para a aprendizagem de técnicas matemáticas. Por outro lado, a integração da tecnologia digital aliada a utilização de dados reais nos cursos de serviço de estatística pode contribuir para o aprendizado do raciocínio estatístico e sua aplicação nas mais diversas áreas. Esse estudo descreve estratégias didáticas para o ensino da Análise Exploratória de Dados em dois cursos de serviços universitários no Brasil. As estratégias são baseadas no desenvolvimento de atividades práticas com a aplicação de técnicas estatísticas em dados reais provenientes de repositórios governamentais da internet. O resultado da prática didática se mostrou promissor dado que dos 81 estudantes matriculados nos cursos de serviço de estatística, 80 avaliam positivamente o uso de dados reais provenientes da internet. Foi também possível avaliar pelas respostas dos alunos que a proposta didática atingiu seu propósito ao motivá-los a aplicar e interpretar técnicas e resultados estatísticos. A percentagem de alunos reprovados foi reduzida e as médias finais aumentaram quando comparamos as turmas que utilizaram e as turmas que não utilizaram as bases de dados reais provenientes dos repositórios.

*Palavras chave:* Alfabetização estatística; Prática de ensino; Dados reais; Integração da tecnologia; Internet.

### Abstract

Statistical literacy is an essential skill for citizens who live in societies brimming with information. Research shows that statistics service courses are still directed towards learning out of context mathematical techniques. On the other hand, the integration of digital technology and the use of real data in statistics service courses may contribute to the learning of statistical thinking and its application in different areas. This research paper describes didactic strategies for teaching Exploratory Data Analysis in two university service courses in Brazil. The strategies are based on the development of practical activities with the application of statistical techniques to real data, obtained from government repositories online. The result of the teaching practice seemed promising, given that 80 out of the 81 students enrolled in the statistics service courses assessed the use of real data from the internet positively. It was also possible to evaluate, by the students' answers, that the didactic proposal has achieved its objective of motivating them and interpreting statistical techniques and results. The percentage of failed students was reduced and the final

averages increased when we compared the classes they used and the classes that did not use the actual databases coming from the repositories.

*Keywords:* Statistical literacy; Teaching practice; Real data; Integration of technology; Internet.

## 1. Introdução

O ensino da estatística em cursos universitários de serviço é um desafio para docentes que procuram fazer com que a aprendizagem seja significativa e relevante, mas, ao mesmo tempo, se deparam com estudantes de diversas áreas que não desejam se envolver com qualquer estudo que seja percebido como matemático [1], [2].

Cursos de serviço em estatística são aqueles dirigidos para estudantes que irão se formar em qualquer outra área, como Biologia, Medicina, Engenharias, Propaganda e demais cursos.

Hoje, o estudo das estatísticas e suas aplicações para a análise de dados é parte do currículo para estudantes em muitas e diversas áreas e estes alunos têm diferentes bases de conhecimento, uma gama de interesses profissionais e objetivos e habilidades matemáticas variáveis [3].

A educação estatística foi transformada pela disponibilidade e uso generalizado da tecnologia, principalmente pelos programas computacionais e internet [3]. No entanto, como Nicholls [4] destaca, embora as universidades tenham registado progressos significativos nos aspectos teóricos da estatística, infelizmente, esse avanço foi em detrimento da formação adequada de utilizadores ocasionais de estatísticas. Isto está intimamente relacionado com a área problemática da estatística para o ensino de serviço para não especialistas.

O foco principal da literatura na educação estatística de graduação é sobre cursos de estatística oferecidos pelos departamentos de matemática ou estatística. Os dados publicados sobre cursos de serviço em estatística, ministrados por outros departamentos, são mais difíceis de acessar, existem poucos dados sobre a aprendizagem em cursos de serviços em estatística [5].

A despeito da baixa publicação de trabalhos na área da educação em estatística para cursos de serviço, autores como Gordon [5], Petocz e Reid [6], Gordon, Petocz e Reid [1], Gordon e Nicholas [2], e März e Kelchtermans [7] mostraram que grande parte dos estudantes dos cursos de serviços de estatística aprendem conceitos e técnicas matemáticas descontextualizadas e informações irrelevantes a serem dominadas em curto prazo de tempo. Não é raro que esses mesmos alunos, distantes do campo da estatística, revelem experiências negativas as quais ocasionam altos níveis de ansiedade enquanto estão matriculados nesses cursos [8]. Essa situação apresenta claros resultados no desempenho dos estudantes. São poucos aqueles que entendem a estatística como uma

ferramenta para o trabalho profissional e crescimento pessoal.

Por outro lado, como as demandas por lidar com dados em uma era da informação continuam a crescer, os avanços em tecnologia e software tornam ferramentas e procedimentos mais fáceis de usar e mais acessíveis a mais pessoas, diminuindo assim a necessidade de ensinar a mecânica dos procedimentos, mas aumentando a importância de uma compreensão sólida dos conceitos estatísticos necessários para usar e interpretar essas ferramentas de forma inteligente [9].

Como sugere Neumann, Neumann e Hood [10], conjuntos de dados reais apresentam interesse intrínseco e captam a atenção, bem como podem estimular o aluno na pesquisa por respostas a perguntas que foram feitas quando os dados foram coletados. Os docentes devem nesse caso, tomar os devidos cuidados para que essa estratégia não torne o uso de dados reais inconvenientes visto que a alta carga de trabalho na busca dos mesmos possa causar desânimo nos estudantes.

Atualmente, diversas áreas da ciência fornecem um rico acervo de dados, diferentes dos dados utilizados em muitos dos cursos que insistem em apresentar moedas, dados, bolas em urnas e outros exemplos. Brooks [11], Barnett [12] e Mvududu e Kanyongo [13] sugerem o uso de dados ambientais para apresentar técnicas estatísticas de análise exploratória de dados, correlação e outras. As áreas sociais são também um repositório farto de dados que podem contribuir para exemplos reais e suscitar o interesse dos estudantes das diversas áreas que frequentam cursos de serviços de estatística.

Existe um consenso entre os estatísticos que a educação estatística deve ser centrada nos dados e no raciocínio estatístico, e minimamente na apresentação da teoria matemática. De forma geral, o objetivo da educação estatística para estudantes de outras áreas é responder perguntas do mundo real [14], [15], [13]. Neste sentido, o uso de dados reais associado a utilização de programas computacionais pode trazer diversos benefícios na aprendizagem dos estudantes.

Para Neumann, Neumann e Hood [16], a disciplina de estatística é adequada à integração da tecnologia que propicia a melhora na aprendizagem e o envolvimento dos alunos. A tecnologia pode ser útil para simular conceitos estatísticos, criar exercícios de aprendizagem interativos e ilustrar aplicações do mundo real.

A integração da tecnologia é percebida pelos educadores como uma ferramenta muito importante para a prestação efetiva de ensino para todos os níveis de educação. Desta forma, o uso de planilhas eletrônicas e programas computacionais podem minimizar o ônus de elaborar fórmulas estatísticas e assim utilizar de forma mais efetiva o tempo para analisar, interpretar e justificar suas conclusões com base nos dados [17]. Soma-se a essas tecnologias a rede Internet, a qual permite a disseminação de informação e a construção do

conhecimento pelas experiências em que a comunicação e colaboração possam predominar [18].

A internet é amplamente utilizada para armazenar e disseminar informações. Neste aspecto, diversas agências desempenham importante papel na disseminação de dados estatísticos, entre elas as agências nacionais de estatísticas, as bases de dados de estatística das Nações Unidas e Banco Mundial e arquivos de dados das sociedades de estatística. No que se refere as agências nacionais, estas tem ampliado consideravelmente suas bases o que reflete a importância que as estatísticas assumiram na atual sociedade [19].

### 1.1. Agências nacionais de dados Latino-americanas e Caribe

Um sítio na internet que agrega diversos dados dos países latino-americanos e do caribe é mantido pela Comissão Econômica para América Latina e Caribe – CEPAL. A seção denominada Dados e Estatísticas agrega informações nas mais diversas áreas, entre as quais, demografia, educação, saúde, pobreza, economia, meio ambiente, e muitas outras. Esses dados podem ser descarregados para computadores pessoais por qualquer usuário.

### 1.2. Agência Internacional

A Divisão de Estatística das Nações Unidas (*United Nations Statistics Division*), compila e divulga informações globais e congrega dados de várias agências e instituições entre as quais Organização Mundial de Saúde e Banco Mundial. Da mesma forma que outras instituições, pode-se filtrar informações de diversos países e períodos diferentes e também descarregar os dados em planilhas eletrônicas.

## 2. Práticas Educacionais Estatísticas

Em 2005, a Associação Estatística Americana (ASA) endossou o Relatório GAISE de Educação Estatística como um manual de boas práticas e objetivos para o ensino da estatística. Este relatório teve um impacto profundo no ensino de estatísticas introdutórias em todo o mundo, e as seis recomendações apresentadas no relatório resistiram ao teste do tempo. Assim, este estudo tem como base didática as seguintes recomendações do relatório GAISE:

- a) Enfatizar a alfabetização estatística e desenvolver o pensamento estatístico;
- b) Utilizar dados reais;
- c) Concentrar a aprendizagem nos conceitos em vez do mero conhecimento dos procedimentos;
- d) Promover o aprendizado ativo em sala de aula.

- e) Utilizar a tecnologia para desenvolver a compreensão conceitual.

## 3. Objetivo

Como base, portanto, nas novas tecnologias e de acordo com as diretrizes da Associação Estatística Americana este trabalho apresenta os seguintes objetivos:

- 1) Descrever atividades práticas de Análise Exploratória de Dados desenvolvidas em cursos de serviços universitários utilizando dados reais obtidos de repositórios na internet.
- 2) Avaliar o impacto que essas atividades produziram na motivação dos alunos para o desenvolvimento e aprendizagem dos conceitos estatísticos apresentados.
- 3) Avaliar o desempenho dos alunos submetidos ao ensino da análise exploratória de dados utilizando dados reais obtidos de repositórios na internet em comparação a cursos sem o uso de dados reais.

## 4. Método

### 4.1. Sujeitos do Estudo

O curso de estatística foi desenvolvido com 41 estudantes de graduação em Ciências Biológicas (CB-Novo) e 42 de Engenharia Biotecnológica (Eng. Biotec. Novo) da Faculdade de Ciências e Letras de Assis da Universidade Estadual Paulista – UNESP, São Paulo, Brasil. A carga horária para cada curso foi de 60 horas e 30 horas respectivamente durante um semestre. Para verificar o progresso das turmas com a metodologia proposta, duas outras turmas dos mesmos cursos de anos anteriores, sem o uso do método proposto, foram comparadas em relação à suas notas finais – Turma de Engenharia Biotecnológica (Eng. Biotec.) com total de 44 alunos e turma de Ciências Biológicas (CB) com 45 alunos. Foi tomado o cuidado para que todas as turmas apresentassem número de alunos semelhantes e que todos tivessem frequência semelhantes aos cursos.

### 4.2. Conceitos estatísticos abordados

Para o programa curricular de análise exploratória de dados foram selecionados os tópicos envolvendo:

- 1) Distribuição de Frequência
- 2) Representação Gráfica
- 3) Medidas de Tendência Central e Variabilidade e Gráfico Box Plot.

Para cada um dos tópicos, foi realizada uma aula clássica em que o professor apresentava os principais conceitos da técnica estatística. Posteriormente as atividades práticas com dados reais obtidos de várias instituições foram distribuídas para grupos de quatro alunos que dispunham de duas semanas para realizar e apresentar

suas análises por escritos. Os alunos podiam recorrer ao professor para sanar dúvidas e discutir a temática estatística abordada. Posteriormente as análises eram avaliadas e comentadas pelo professor e devolvidas para os alunos.

Ao final do curso, os alunos avaliaram as atividades quanto ao uso de dados reais dos temas propostos utilizando questionário confeccionado pelo Google Forms®.

### 4.3. Base de dados e temas abordados

Para compor as atividades práticas a serem desenvolvidas pelos estudantes dos dois cursos, foram selecionados alguns temas das áreas de ecologia e temas sociais. Posteriormente buscamos dados associados aos temas em algumas bases de dados descritas a seguir.

#### 4.3.1. Base de dados brasileiras

Diversos temas podem suscitar interesse entre os alunos dos diversos cursos. Entre eles, temas relacionados ao meio ambiente, renda da população e AIDS. Como temática foram escolhidas a degradação das matas e florestas do Brasil (focos de queimadas), lista das aves brasileiras ameaçadas de extinção, a renda per capita nos estados brasileiros e mortalidade por Síndrome da Imunodeficiência Adquirida – SIDA nos estados brasileiros.

##### 4.3.1.1 Focos de Incêndio

Os dados apresentados aos alunos foram dispostos em uma planilha eletrônica, coletados junto ao Instituto Nacional de Pesquisa Espacial – INPE (Brasil) para focos de incêndios por estados – municípios, no mês de dezembro de 2016.

##### 4.3.1.2 Aves Ameaçadas

A lista de aves ameaçadas de extinção foi confeccionada em uma planilha eletrônica com o nome de cada uma das espécies de aves do Brasil (mais de 1800) e sua respectiva classificação quanto à extinção. O objetivo dos dados é proporcionar a oportunidade de avaliar a situação do país em relação às condições das aves brasileiras. Os dados reais podem ser obtidos no Ministério do Meio Ambiente no Brasil, mas também em outras instituições internacionais.

##### 4.3.1.3 Renda per capita

Para apresentação dos dados relativos à renda per capita, utilizou-se como fonte de dados o Instituto Brasileiro de Geografia e Estatística – IBGE. Foram disponibilizados dados da renda per capita dos 5565 municípios brasileiros do ano de 2010.

Para os dados de desmatamento e renda, o objetivo foi o de introduzir distribuição de frequência (tabelas). Desta forma, pode-se inserir conceitos de níveis de mensuração associados ao uso de tabelas com dados reais.

#### 4.3.1.4 Mortalidade por SIDA

Os dados de mortalidade por SIDA foram coletados junto ao IBGE para os anos de 2007 a 2014. Também foram apresentados os dados de população para cada estado para que os alunos pudessem discutir o conceito epidemiológico de taxas.

No caso do estudo da SIDA o objetivo é apresentar formas de representação gráfica do evento que ocorre ao longo de um período

O INPE e IBGE contam com uma base dados na Internet com acesso livre e gratuito. Os dados podem ser descarregados em formatos de planilhas eletrônicas o que facilita seu manuseio.

#### 4.3.2. Base de dados Internacionais

Temas que envolvem a comparação de culturas diferentes também podem ser utilizados para despertar o interesse nas análises estatísticas por parte dos estudantes. Nesse sentido, utilizamos dois temas que se mostraram interessantes pois apresentavam, antes de qualquer análise, alguns conhecimentos equivocados, por parte dos alunos, o que foi modificado após o uso das técnicas estatísticas apresentadas. A temática estudada através dos dados foi a inserção da mulher no mercado de trabalho em comparação aos homens, e a idade do primeiro matrimônio das pessoas do sexo feminino. Os dados foram obtidos junto ao Banco Mundial por meio de sua página na internet.

##### 4.3.2.1 Emprego Feminino

Nas últimas duas décadas, o mundo reduziu a divisão entre homens e mulheres, especialmente na educação primária e na saúde. Ainda subsistem lacunas críticas. Para analisar dados que podem mostrar diferenças entre homens e mulheres em várias partes do mundo, foi utilizado o indicador da relação da % empregados do sexo feminino em relação a % empregados do sexo masculino. Assim, resultados iguais a 100% indicam que existe uma igualdade entre gêneros. Percentagens menores que 100% indicam que existem menos mulheres no mercado de trabalho do que homens.

##### 4.3.2.2 Casamentos

O segundo conjunto de dados estudadas refere-se a data do primeiro casamento de mulheres em diversos países. O que estamos tentando medir é a transição que efetivamente está mudando os direitos e obrigações legais ou sociais do sexo feminino. Também é importante observar que a data do primeiro casamento pode definir condições futuras para as mulheres.

Para os dados relativos ao indicador da percentagem de empregados do sexo feminino e data do primeiro casamento foram realizadas análises por meio de medidas de tendências central e de variabilidade e gráfico *box-plot*.

Os dados obtidos no repositório foram agrupados em grupos de países que representam América Latina, Europa, África e Oriente Médio entre os anos de 2006 e 2014.

É importante salientar que todos os temas foram contextualizados e apresentados e distribuídos em layouts atraentes no formato *Portable Document Format* – PDF.

#### 4.4. Análise

Para aferir a motivação dos estudantes nas atividades desenvolvidas de análise exploratória de dados com o uso de dados reais obtidos da internet, foram avaliadas as respostas de um questionário com questões fechadas e abertas ao final do curso. As notas médias para as duas turmas de Engenharia Biotecnológica e para as duas turmas do curso de Ciências Biológicas foram comparadas utilizando medidas descritivas, teste t student para duas amostras independentes e intervalo de confiança para diferença entre médias (nível de significância de 5%). A Suposição de normalidade para as amostras foi avaliada pelo teste de Shapiro -Wilk (5%).

#### 4.5. Programa Computacional Adotado

Para ambos os cursos foi utilizado o programa computacional Instat versão 3.36.

### 5. Resultados

#### 5.1. Respostas ao questionário

A figura 1 apresenta respostas múltiplas ao questionário sobre como os dados reais motivaram os alunos que puderam escolher mais do que uma alternativa a questão.

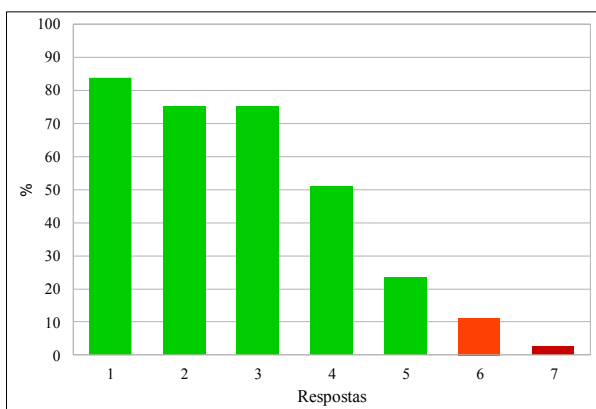


Figura 1. Respostas dos estudantes quanto aos dados utilizados no curso

#### Legenda

- 1 – Achei interessante trabalhar com dados referentes a temas brasileiros.
- 2 – Achei interessante trabalhar com dados referentes a diversos países.

3 – Os dados reais me ajudaram a conhecer temas que eu não conhecia.

4 – Os dados reais me motivaram a usar técnicas estatísticas.

5 – Os dados reais me ajudaram a perder um pouco o medo de lidar com estatística.

6 – Não achei fácil trabalhar com dados reais.

7 – Os dados reais não me motivaram a usar técnicas estatísticas.

Não há diferenças significantes entres as percentagens para cada item de reposta quando comparamos os dois cursos analisados neste trabalho.

Uma segunda questão sobre o conhecimento e uso de planilhas eletrônicas foi realizada (figura 2).

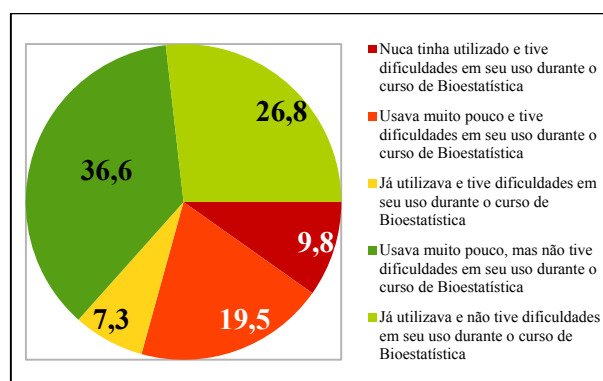


Figura 2. Respostas dos estudantes quanto ao uso de planilhas eletrônicas

#### 5.2. Respostas espontâneas sobre as atividades propostas

As respostas do quadro 1 espelham tanto a opinião dos estudantes do curso de Ciências Biológicas quanto dos estudantes de Engenharia Biotecnológica.

##### Quadro 1. Respostas espontâneas

- 1) *Eu acho que todas as atividades de forma geral foram interessantes porque tinham dados reais, com fontes seguras.*
- 2) *O caso em questão que mais me chamou atenção foi sobre o trabalho feminino. A atividade me fez pesquisar o por que de tão poucas mulheres trabalharem no oriente médio.*
- 3) *Porque foi possível ver temas com dados reais relacionados ao país e fazer as análises e tirar conclusões.*
- 4) *... ela despertou-me interesse quando eu e meu grupo elaboramos análises gráficas e, eu me surpreendi com os resultados e sua interpretação como, por exemplo, a porcentagem de aves ameaçadas e em extinção, entre outros assuntos que me fizeram pesquisar mais sobre eles no decorrer do trabalho.*

5) <i>Achei interessante os exemplos utilizados, junto com o contexto da questão, é útil para que além dos conhecimentos da atividade.</i>
6) <i>Possibilidade de avaliar a relação de dados de forma mais prática.</i>
7) <i>Fiquei mais interessado e aprendi mais.</i>
8) <i>Gostei bastante da atividade envolvendo Box-Plot, principalmente pela aplicação dele. ... acredito ser muito bom para atuação em uma empresa ou até mesmo no ramo da pesquisa.</i>
9) <i>...adorei a atividade de medidas de tendência central e variabilidade e do box plot, os enunciados dos exercícios me entretiveram muito! Adorei o tema do trabalho feminino, acho que essa atividade abordou temas importantes!</i>
10) <i>Achei esta parte da matéria menos mecânica e mais lúdica.</i>

### 5.3. Resultados do desempenho dos estudantes

A tabela 1 apresenta a comparação do desempenho final dos estudantes das turmas com o uso das bases de dados e sem o uso das bases.

Tabela 1. Comparação de desempenho. Notas Finais (escala de zero a dez)

Turmas	média	Desvio	$p$
		Padrão	
Eng. Biotec. Novo	6,9	1,3	0,001
Eng. Biotec.	5,7	1,7	
CB Novo	6,2	1,4	0,003
CB	5,3	1,1	

O intervalo de confiança de 95% para a diferença entre as médias da turma Eng. Biotec. Novo em relação à turma Eng. Biotec. apresenta resultado igual a  $[0,50$  a  $1,80]$ , enquanto para a diferença entre médias das turmas CB Novo em relação a CB é de  $[0,30$  a  $1,41]$ .

Os resultados de reprovação foram reduzidos com o uso de dados provenientes das bases de dados da Internet. No caso da engenharia a redução foi de cerca de 23%, de um total de 30% sem o uso das bases para 7% com o uso das bases. No caso da turma de Ciências Biológicas, a redução foi de 11%, de 29% de reprovação para 18% com o uso das bases.

## 6. Discussão

Como destacam Foresti e Teixeira [20], a aprendizagem na sociedade contemporânea da era digital apresenta grande fluxo de informação que afeta a tomada de decisões no cotidiano das pessoas. Neste sentido, a partir da apropriação pedagógica dos recursos tecnológicos a atuação de educadores e educandos pode ser ressignificada.

A estatística como instrumento de análise de dados está disseminada nas mais diversas áreas e pode contribuir para a formação de estudantes universitários. Por outro lado, para que essa importante ferramenta possa contribuir para a formação de jovens estudantes, é necessário a construção de diferentes formas de integrar as Tecnologias da Informação e Comunicação no contexto educativo.

A internet tem sido um repositório para dados e informações desde sua criação e o uso para fins educacionais já se provou eficaz e enriquecedor.

Quanto ao ensino da estatística, vários softwares estão sendo introduzidos para estudar como analisar e interpretar os dados com os quais estamos trabalhando. No entanto, o enfoque em técnicas e métodos numéricos e teóricos para analisar dados isoladamente não é suficiente.

Os dados reais utilizados para fins didáticos nos cursos apresentados neste trabalho são provenientes de organizações idôneas e conhecidas pelos alunos. Esses mesmos dados são recentes e espelham a realidade atual dos locais em que foram coletados. Assim, a experiência de trabalhar com dados reais provenientes de sítios de instituições disponíveis na Internet se mostrou efetiva para com os objetivos dos cursos, qual seja a motivação dos estudantes para o uso das técnicas e a discussão, interpretação e o aprendizado. Observamos que dos 81 estudantes, 80 responderam afirmativamente a pelo menos uma das alternativas positivas relacionadas ao uso de dados reais provenientes da internet (figura 1).

As respostas de que os alunos não acharam fácil trabalhar com dados reais, dez no total, pode estar ligada a conhecimentos anteriores, visto que desses, 6 assinalaram que tiveram dificuldades como o manuseio de planilhas eletrônicas ao longo do curso (figura 2). Mesmo esses alunos, também reportaram o interesse em trabalhar com dados reais.

Por outro lado, pelas respostas apresentadas no quadro 1, fica evidente que os dados reais atingiram seu propósito ao motivar os estudantes a aplicar e interpretar técnicas e resultados estatísticos. Chama a atenção as respostas 2 e 4 do quadro 1 em que os alunos buscaram outras informações além das apresentadas nas atividades. Também merece destaque as respostas 9 e 10 em que existe a conotação do lúdico. Essas afirmações mostram que as atividades puderam ser tratadas de forma mais agradável.

Outra resposta apresentada por um dos estudantes no quesito “outros” na questão sobre a motivação é particularmente interessante e merece destaque:

*“Por meio dos dados reais nas atividades pude compreender a extrema relevância que a estatística possui no momento de extração de uma informação em meio ao colossal mar de informações.”*

Assim, ao congregarmos as diferentes respostas dos alunos dos dois diferentes cursos, pode-se verificar que as atividades envolvendo os dados reais obtidos junto a organizações governamentais na internet alcançaram o objetivo de facilitar e motivar o ensino da estatística.

São escassos os trabalhos que adotam o método denominado de quase-experimental para avaliar o desenvolvimento de estudantes com novas práticas no ensino de estatística descritiva no ensino superior e mesmo os exemplos relacionados ao uso de dados on-line com avaliações qualitativas não são frequentes [21]. No entanto, autores como Zeleke e Lee [22], Chance et al. [23] e Vaughan [24] apresentam sugestões do uso de base de dados on-line para o ensino de estatística.

Os resultados das comparações do final dos cursos (tabela 1) mostram um aumento das notas médias em torno de 1,2 ponto quando observamos as duas turmas de engenharia biotecnológica. Resultado semelhante é obtido quando comparamos as duas turmas do curso de ciências Biológicas, ou seja, o crescimento da média em torno de 0,9 ponto (resultados estatisticamente significativos).

Experimento randomizado conduzido por Lesser [25] mostrou que estudantes universitários de curso introdutório de estatística classificaram um exemplo contextualizado na área de justiça social mais interessante do que um exemplo convencional. Esse resultado mostrou-se estatisticamente significativo ( $p \leq 0,05$ ).

Outro resultado importante obtido na pesquisa é a redução do percentual de estudantes reprovados ao final do curso. Em ambos os cursos o percentual foi reduzido.

Importante observar que no curso de engenharia biotecnológica a melhora do desempenho e a redução das reprovações foi maior, o que pode ser explicado pela área em que se insere o curso, ou seja, um curso da área de exatas, enquanto o curso de ciências biológicas apresenta maiores resistências a essa mesma área.

Resultado semelhante foi obtido por Tchantchane et al. [26] com a introdução de tecnologia como a inclusão de planilhas eletrônicas para trabalhar os dados. Segundo o autor a redução das reprovações foram reduzidas de 34% para 14%.

Trabalhar com exemplos reais auxilia os alunos a ganhar experiência prática, o que os torna mais capazes de enfrentar o mundo após a universidade. A estatística é parte da vida cotidiana, quer se refira a pesquisas de opinião, resultados de estudos médicos, informações

demográficas, ou outras informações. Usar exemplos reais no ensino é um trabalho árduo. O docente necessita escolher os exemplos cuidadosamente. Eles devem ser interessantes, atualizados e variados.

Para Unwin [27], usar dados reais no ensino de estatística não é uma tarefa fácil para o professor. Nossa experiência mostra que são necessários conjuntos de dados variados para diferentes situações e contextos, o que aumenta a carga horária para preparar as aulas. Ao contrário, um fator persuasivo importante para o uso de dados artificiais é que um docente pode criar mais facilmente e rapidamente conjuntos de dados com as características necessárias.

No entanto, as vantagens superam as dificuldades de aplicar uma abordagem de ensino baseada na aplicação de dados reais que podem ser obtidos em sites da internet.

É fato também que devido à influência das novas tecnologias, o ensino e aprendizagem da estatística nas universidades estão em constante mudanças [28]. Assim, a criação de um ambiente de aprendizagem tecnológico terá um impacto considerável sobre os conteúdos da estatística e deve ser acompanhada do desenvolvimento da alfabetização e pensamento estatístico bem como na ampliação da aprendizagem dos conceitos. Nesse sentido, deve-se promover a utilização da Internet e outras formas tecnológicas em sala de aula [29].

## Conclusão

A experiência descrita neste artigo nos ensinou muito sobre a aprendizagem dos alunos, sobre maneiras de pensar e perceber conceitos e ideias em estatística. Os resultados obtidos indicam que o uso de dados de fontes da internet melhorou o ensino dos estudantes e os processos de aprendizagem. Ao longo do curso foi também possível observar que os alunos estavam mais envolvidos nas atividades, mas ainda é necessário repensar quem são os alunos de cursos de serviço, quais conhecimentos estão disponíveis para que possam ter maior chance de sucesso.

Ao ensinar nossos alunos desta forma, acreditamos que estamos fornecendo-lhes um contexto mais real e rico para desenvolver a alfabetização estatística e tornar o ensino significativo.

## Sugestões

A tecnologia levou a inúmeras mudanças no conteúdo, na pedagogia e no formato dos cursos de estatística. Muitos problemas que antes apresentam difíceis análises agora apresentam soluções. Essas mudanças na prática estatística têm um impacto direto sobre o conteúdo que deve ser ensinado, mesmo em cursos introdutórios. Foi possível verificar com a experiência relatada como a associação da tecnologia de repositórios de dados na

internet podem auxiliar os alunos de diferentes cursos na aprendizagem da estatística. No entanto, o processo de ensino e aprendizagem é contínuo e desafiador.

Como sugestões gerais para cursos de serviço de estatística que tem como objetivo o ensino da análise exploratória de dados apresentamos as seguintes práticas:

### **Realizar atividades iniciais que possam melhorar o conhecimento em planilhas eletrônicas**

A diversidade de alunos nos cursos de serviço de estatística é conhecida e suscita cuidados no que se refere ao conhecimento de alguns conceitos e ferramentas. Assim, o uso de planilhas eletrônicas pode ser um dos obstáculos ao bom desenvolvimento dos cursos de estatística. Nossa experiência mostra que diversos estudantes não conhecem algumas das várias ferramentas geralmente incorporadas em suas funções. Entre elas podemos destacar algumas que devem figurar em um rápido curso introdutório:

a) Funções matemáticas simples – O cálculo das quatro operações básicas da matemática com o uso de planilhas é também desconhecido por parte de vários alunos de área biológicas e de humanidades. Um simples cálculo pode ser desenvolvido e automatizada de forma simples e rápida em diversas planilhas eletrônicas. Por exemplo, o cálculo de uma taxa de mortalidade infantil pode ser custoso do ponto de vista mecânico quando os dados são referentes a todos os 5561 municípios brasileiros. No entanto, é possível realizar o cálculo em uma planilha eletrônica com uma simples fórmula matemática. Desta maneira, essa deve ser uma das primeiras atividades para aproximar o aluno do uso de planilhas eletrônicas.

b) Filtro – Essa função fornece um mecanismo muito simples para obter subconjuntos de dados. Auto filtros podem ser configurados para escolher valores ou atributos de uma variável. Vários filtros podem ser usados, cada um agindo sobre os dados de diversas variáveis. Esta é uma ferramenta que pode facilitar várias tarefas.

c) Classificar – A função classificar é útil para ordenar os valores de uma variável do menor para maior valor ou vice e versa. Essa função pode ser ampliada para mais de uma variável de forma hierárquica e traz benefícios para realizar gráficos e tabelas por exemplo.

### **Realizar atividades iniciais que possam melhorar o conhecimento do programa estatístico adotado**

A organização dos dados é mais facilmente desenvolvida em planilhas eletrônicas, no entanto o uso de programas computacionais estatísticos são instrumentos indispensáveis nos cursos de serviço. Propiciam a rápida realização de análises através de menus e botões preestabelecidos, o que torna os cálculos simples e rápidos. De maneira geral, tanto os programas computacionais comerciais quanto os livres apresentam layout semelhante, mas para a grande maioria dos estudantes são instrumentos tecnológicos desconhecidos.

a) Importação de dados – A importação de dados por parte dos programas estatísticos é uma tarefa fácil, mas que necessita de uma pequena apresentação.

b) Onde estão as análises estatísticas? - Alguns minutos são suficientes para que o professor possa apresentar as principais ferramentas disponíveis no programa estatístico.

c) Transferir os resultados – De maneira geral os resultados das análises são organizados em algum programa de edição de texto para que possa combinar tabelas, gráficos e estatísticas com os textos das análises efetuados pelos estudantes. Desta forma, realizar algumas simulações de transposição de informações entre os programas computacionais é uma tarefa produtiva.

### **Escolher temas variados para cada item do programa de estatística**

A diversidade de temas é importante para que a maioria dos estudantes em uma ou mais atividades possa sentir maior afinidade em relação ao assunto abordado. A inclusão de dados regionais, nacionais e internacionais, disponíveis em várias organizações governamentais nos diversos sítios da internet podem também estimular novos e mais aprofundados estudos.

### **Seja prudente, prefira abordagens iniciais simples**

Em nossa primeira tarefa, apresentamos uma única variável para desenvolver a atividade referente a distribuição de frequência. Buscamos avaliar junto aos alunos a situação de risco de extinção de 1807 espécies de aves no Brasil. Cada espécie é classificada segundo uma das cinco categorias - pouco preocupante, quase ameaçada, vulnerável, ameaçadas de extinção e criticamente em perigo. A planilha eletrônica de dados disponibilizada aos estudantes possuía duas únicas colunas. Os alunos puderam observar que sem o uso de uma simples tabulação, seria impossível conhecer a situação das aves brasileiras.

### **Contextualizar os dados**

A ausência de informações sobre o tema estudado não estimula o aluno que em muitas ocasiões pode não conhecer o assunto. Desta forma, deve-se preparar material que contenha informações chaves sobre as questões tratadas.

### **Procure apresentar as atividades de forma visualmente atrativa**

Os atuais estudantes pertencem a era digital em que a cultura da imagem apresenta grande influência no desenvolvimento educacional. Dessa forma, materiais atraentes podem, de forma inicial, chamar a atenção dos alunos para as atividades de estatística. Utilize imagens sobre os temas, crie layout atraentes.



### **Apresentar os endereços da internet de cada uma das instituições responsáveis pelos dados utilizados nas atividades**

É importante que os estudantes possam consultar os sítios da internet que disponibilizam os dados reais, o que permitirá credibilidade e veracidade nos resultados alcançados.

### **Avaliar cada uma das atividades de maneira periódica**

Durante o curso de análise exploratória de dados para os cursos de Ciências Biológicas e Engenharia Biotecnológica, a cada um, dos ou três tópicos do programa, desenvolvemos uma avaliação que propiciou ao professor e aos estudantes o direcionamento de estratégias para sanar as dúvidas e melhorar a aprendizagem. Todas as avaliações foram discutidas com alunos, o que proporcionou um melhor engajamento nas futuras avaliações. Esse processo, mesmo que exaustivo, em conjunto com a utilização dos dados reais apresentou bons resultados ao final do curso.

## **Referencias**

[1] S. Gordon, P. Petocz, A. Reid, Teachers' Conceptions of Teaching Service Statistics Courses. *International Journal for the Scholarship of Teaching and Learning*, 1(1), 1- 15, 2007.

[2] S. Gordon, J. Nicholas, Teaching with examples and statistical literacy: views from teachers in statistics service courses. *International Journal of Innovation in Science and Mathematics Education*, 18(1), 14-25, 2010.

[3] S. Gordon, P. Petocz, A. Reid, What makes a “good” statistics student and a “good” statistics teacher in service courses? *The Mathematics Enthusiast*, 6(1), 25-40, 2009.

[4] F. Nicholls, Future Directions for the Teaching and Learning of Statistics at the Tertiary Level. *International Statistical Review*, 69(1), 11-15, 2001.

[5] S. Gordon, Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, 3(1), 40-59, 2004.

[6] P. Petocz, A. Reid, Something strange and useless: service students' conceptions of statistics, learning statistics and using statistics in their future profession. *International Journal of Mathematical Education in Science and Technology*, 36(7), 789–800, 2005.

[7] V. März, G. Kelchtermans, Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education*, 29, 13 – 24, 2013.

[8] A. Onwuegbuzie, V. Wilson, Statistics anxiety: Nature, etiology, antecedents, effects, and treatments—a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), 195–209, 2003.

[9] M. Everson, A. Zieffler, J. Garfield, Implementing New Reform Guidelines in Teaching Introductory College Statistics Courses. *Teaching Statistics*, 30(3), 66–70, 2008.

[10] D. L. Neumann, M. M. Neumann, M. Hood, The development and evaluation of a survey that makes use of student data to teach statistics. *Journal of Statistics Education*, 18(1), 1-19, 2010.

[11] R. Brooks, Using carbon emissions data to ‘heat up’ descriptive statistics. *Teaching Statistics*, 34(1), 25–30, 2011.

[12] V. Barnett, Green Statistics. *Teaching Statistics*, 23(2), 35 – 37, 2001.

[13] N. Mvududu, G. Y. Kanyongo, Using Real Life Examples to Teach Abstract Statistical Concepts. *Teaching Statistics*, 33(1), 12–16, 2011.

[14] S. Peiris, E. J. Beh, Where statistics teaching can go wrong. *International Journal of Innovation in Science and Mathematics Education*, 15(1), 21-23, 2006.

[15] Z. Libman, Integrating Real-Life Data Analysis in Teaching Descriptive Statistics: A Constructivist Approach. *Journal of Statistics Education*, 18(1), 1-22, 2010.

[16] D. L. Neumann, M. M. Neumann, M. Hood, Evaluating computer-based simulations, multimedia and animations that help integrate blended learning with lectures in first year statistics. *Australasian Journal of Educational Technology*, 27(2), 274-289, 2011.

[17] A. Tchanchane, P. C. A. Fortes, Multivariate Survey Analysis: Evaluation of Technology Integration in Teaching Statistics. Proceedings ascilite, University of Tasmania, Australia, 1227-1238, 2011.

[18] A. C. Teixeira, E. J. R. Brandão, Internet e democratização do conhecimento: repensando o processo de exclusão social. *Novas Tecnologias na Educação*, 1(1), 1-7, 2003.

[19] B. Phillips, Overview of online teaching and internet resources for statistics education. 10 p. Obtido em 04 de março de 2017. <https://iase-web.org/documents/papers/sat2003/Phillips.pdf>

[20] A. Foresti, A.C. Teixeira, Proposta de um conceito de aprendizagem para a era digital. *Revista*

*Latinoamericana de Tecnología Educativa*, 11(2), 55 – 68, 2012.

[21] L. Taylor, K. Doehler, Using Online Surveys to Promote and Assess Learning. *Teaching Statistics*. 2014 Teaching Statistics Trust, 36(2), 34–40, 2014.

[22] A. Zeleke, C. Lee, Teaching introductory statistics using student generated data in a large class. Proceedings of International Conference on Teaching Statistics 8, 2010.

[23] T. S. Vaughan, Teaching Statistical Concepts with Student-Specific Datasets, *Journal of Statistics Education*, 11:1, 2003.

[24] B. Chance, D. Ben-Zvi, J. Garfield, E. Medina, The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education Journal* 1(1), 2007. Obtido em 15 de maio de 2018. <https://cloudfront.escholarship.org/dist/prd/content/qt8sd2t4rr/qt8sd2t4rr.pdf?t=kro7mj>

[25] L. M. Lesser, Insight into Motivation of Mathematics Learners via Double-Blind Experiment. Paper presented at annual National Council of Teachers of Mathematics Research Pre-session, St. Louis, MO. 2006

[26] A. Tchantchane, C. P. Fortes, S. Koshy, *International Journal of Web Based Learning and Teaching Technologies*, 7(2), 16-27, 2012.

[27] A. Unwin, Datasets on the web: a resource for teaching Statistics? *MSOR Connections*, 10(3), 38-41, 2010.

[28] N. Ramesh, Use of technology to enhance teaching and learning in mathematics and statistics. *MSOR Connections*, 11(1), 34–36, 2011.

[29] W. Xu, Y. Zhang, C. Su, Z. Cui, X. Qi, Roles of technology in student learning of University level biostatistics. *Statistics education research journal*, 13(1), 66 – 76, 2014.

*Información de contacto de los autores:*

**Fernando Frei**  
FCLAssis  
Depto. Ciências Biológicas - UNESP  
Assis - SP  
Brasil  
ffrei@assis.unesp.br