



Universidade de Brasília (UnB)

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas
Públicas (FACE)

Departamento de Ciências Contábeis e Atuariais (CCA)

Curso de Graduação em Ciências Contábeis

ARTHUR RIBEIRO LIMA

Título:

Importância de disciplinas de análise de dados no currículo do curso de Ciências
Contábeis

Brasília – DF
2022

ARTHUR RIBEIRO LIMA

Importância de disciplinas de análise de dados no currículo do curso de Ciências Contábeis

Trabalho de Conclusão de Curso (Monografia ou Artigo) apresentado ao Departamento de Ciências Contábeis e Atuariais da Faculdade de Economia, Administração e Contabilidade da Universidade de Brasília como requisito parcial de obtenção do grau de Bacharel em Ciências Contábeis.

Prof. Responsável:

Dr. Paulo Roberto Barbosa Lustosa

Linha de pesquisa:

Contabilidade para Tomada de Decisão

Área:

Educação em Contabilidade

Professora Doutora Márcia Abrahão Moura
Reitora da Universidade de Brasília

Professor Doutor Enrique Huelva Unternbäumen
Vice-Reitor da Universidade de Brasília

Professor Doutor Diêgo Madureira de Oliveira
Decano de Ensino de Graduação

Professor Doutor José Márcio Carvalho
**Diretor da Faculdade de Economia, Administração, Contabilidade e Gestão de
Políticas Públicas**

Professor Doutor Sérgio Ricardo Miranda Nazaré
Chefe do Departamento de Ciências Contábeis e Atuarias

Professora Doutora Fernanda Fernandes Rodrigues
Coordenador de Graduação do curso de Ciências Contábeis - Diurno

Professor Mestre Wagner Rodrigues dos Santos
Coordenador de Graduação do curso de Ciências Contábeis - Noturno

ARTHUR RIBEIRO LIMA

Importância de disciplinas de análise de dados no currículo do curso de Ciências Contábeis

Trabalho de Conclusão de Curso (Monografia ou Artigo) apresentado ao Departamento de Ciências Contábeis e Atuariais da Faculdade de Economia, Administração e Contabilidade da Universidade de Brasília como requisito parcial de obtenção do grau de Bacharel em Ciências Contábeis.

Prof. Dr. Paulo Roberto Barbosa Lustosa
Orientador

Departamento de Ciências Contábeis e Atuariais
Universidade Brasília (UnB)

Prof. Dr. José Antonio de França
Examinador

Departamento de Ciências Contábeis e Atuariais
Universidade de Brasília (UnB)

Brasília – DF
2022

Dedico à toda comunidade de desenvolvedores, professores ou demais profissionais de TI que ensinam novos estudantes puramente por bondade.

AGRADECIMENTO

Meus mais amorosos agradecimentos à minha esposa por motivar esse reinício de estudos. Meus mais orgulhosos agradecimento à minha mãe e minha família por me apoiarem incondicionalmente e entenderem os momentos de minha ausência, mas certos que seguimos os valores de nossos avós que sempre apoiaram a educação. Agradeço também à comunidade acadêmica da UnB, em especial aos professores que não medem esforços para se atualizarem e ensinarem aquilo que há mais inovador no mercado, eternamente buscando para os alunos a excelência. Por fim, mas incrivelmente importante, todos os amigos e colegas que viveram comigo essa experiência acadêmica.

Muito obrigado!

RESUMO

O presente estudo tem como objetivo destacar a importância da inclusão de disciplinas de análise de dados no currículo do curso de Ciências Contábeis, apresentando métodos de análise de dados digitais consolidados no meio acadêmico e no mercado, bem como realizando uma análise de dados norteada por tais métodos em conjunto com linguagens e ferramentas abertas de tecnologia da informação, evidenciando a possibilidade que o estudante de contabilidade detém de se tornar um profissional altamente demandado na atualidade.

Palavras-chaves: atualização de currículo, data analytics, modernização da contabilidade.

LISTA DE TABELAS E GRÁFICOS

Figura 1: Diagrama de Venn - adaptado de Jéssica (2022).....	13
Figura 2: Adaptado de Richardson et al. (2022).....	14
Figura 3: Elaborado pelo autor.....	20
Figura 4: Elaborado pelo autor.....	22

SUMÁRIO

1. Introdução.....	10
2. Revisão de literatura.....	12
Literatura teórica.....	12
Literatura empírica.....	14
3. Metodologia.....	16
Fase 1 – Identify the questions.....	16
Fase 2 – Master the data.....	17
Fase 3 – Perform test plan.....	17
Fase 4 – Address and refine results.....	18
Fase 5 e 6 – Communicate insights and track outcomes.....	18
4. Resultado.....	19
5. Considerações finais.....	23
Referências.....	25

1. INTRODUÇÃO

Adaptação: tópico tão estudado pelas ciências biológicas e tão praticado pelas ciências contábeis em função da evolução tecnológica. O que outrora requeria um andar inteiro de um prédio para abrigar equipamento de informática, hoje, cabe na palma da mão. As fronteiras dos países – praticamente intransponíveis para grande parte da população – já não desempenham este efeito: da comodidade do lar qualquer pessoa conversa em tempo real com outra do globo terrestre. O registro da vida de um ser humano, que antes exigia deter um status ou realizar um grande feito para merecer ser documentado, ocorre para praticamente todos os indivíduos nos dias atuais, gerando um detalhado rastro digital de cada indivíduo da população.

Semelhantemente, o registro de dados está permeando praticamente todas as cadeias de produção existentes. Como exemplo, a entrada de insumos que antes coletava poucos atributos (data, quantidade, valor, descrição e fornecedor) da transação e, atualmente, coleta dezenas de campos exigidos pelo governo, além de muitos outros de interesse gerencial. Justo mencionar que diversos outros dados podem ser coletados e analisados pela empresa: o acompanhamento em tempo real da entrega pode ser registrado pelo ente empresarial e direcionar a alocação de recursos deste com a previsão de entrega, também pode-se levantar dados de redes sociais públicas para aferir a demanda com base em uma análise de sentimento de efetivos ou potenciais consumidores.

Frente a estas diariamente renováveis possibilidades, a grande maioria dos gestores buscam bases de informações mais críveis para a tomada de decisão (Bajoux-Besbainou, 2022), afastando-se da mera intuição (muitas vezes emocional e/ou subjetiva) quanto ao mercado, aprimorando também a confiança na própria gerencia, que passa a contar com decisões técnicas, justificáveis, auditáveis, evolutivas, escalonáveis e geradoras de conhecimento.

Esta demanda encontra arrimo naqueles que já foram conhecidos como guarda livros. É fato que o produto ofertado pelo profissional contador é a informação patrimonial. Deter tal informação, que outrora era mero diferencial entre as aziendas, atualmente figura como base para eficiência e eficácia que se desdobram ao fim como sobrevivência da empresa como um todo.

Ao longo do tempo, diversas técnicas foram acrescentadas à rotina do profissional contador visando melhorar a qualidade da informação prestada, destaca-se o uso do método de Veneza, um marco que contribuiu para a segurança, controle, entendimento e diversos outros atributos da informação. Desde então, inúmeros estudiosos debruçaram-se sobre a matéria desenvolvendo, ramificando, adaptando ao momento e expandindo a ciência. Hoje, os dados digitais, além de exigir nova adaptação, requerem do profissional

um foco diferenciado, pois mostram-se como uma fonte com potencial confiável para gerar respostas relevantes a nortear as decisões mais favoráveis para a organização como um todo.

Contudo, o processo de “lapidação” destes dados demanda uma análise técnica especializada apta a extrair a melhor informação possível dentro da necessidade da empresa, vencendo os diversos obstáculos existentes. Como exemplo destes últimos, podem ser considerados: a grande quantidade de dados continuamente armazenados, conclusões superficiais/equivocadas, ausência de padronização, nivelamento dos profissionais de contabilidade com os de outras áreas (especialmente tecnologia da informação) e demais desafios. Acrescenta-se que a demanda por profissionais de informática cresce continuamente e o aprimoramento de um terreno comum de entendimento entre este e aquele profissional otimiza os resultados possíveis.

Em linha com esse contexto, o objetivo do presente estudo é **destacar a importância de disciplinas de análise de dados no currículo do curso de Ciências Contábeis**. Para tanto, serão apresentados métodos de análise de dados digitais consolidados no meio acadêmico e no mercado, utilizando linguagens de programação de alto nível (mais próxima da linguagem humana) e programas de fácil acesso para a comunidade em geral. Com o intuito de ampliar a acessibilidade, foram selecionadas ferramentas de uso livre (isentas de pagamento).

O presente trabalho tem potencial para inquietar os gestores da educação quanto à necessidade de um maior contato com ferramentas tecnológicas e também de motivar estudos posteriores mais específicos quanto à matéria.

A estrutura é composta por introdução, revisão de literatura, metodologia, resultados, considerações finais e referências.

2. REVISÃO DE LITERATURA

Literatura teórica

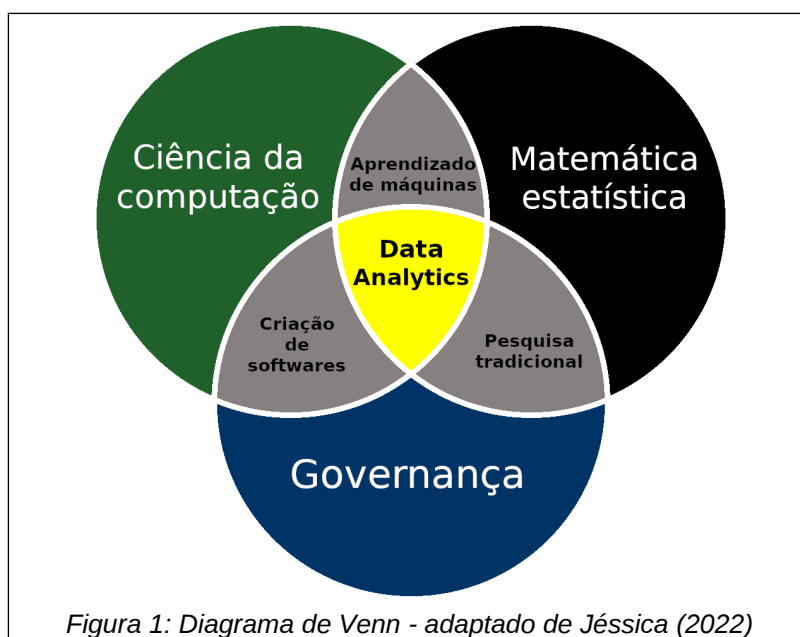
O uso de tecnologia da informação começou a ser percebida como aliada da gestão já na década de 70 (Sharda et al., 2019). Naquele momento, o conjunto de ferramentas tecnológicas usadas pela administração eram chamados de sistemas de apoio à decisão. Nos anos 80, momento dos sistemas internos empresariais, a evolução da programação de computadores permitiu que o conhecimento dos especialistas fosse simulado pelo computador utilizando essencialmente estruturas condicionais. Ampliava-se o uso decisões do sistema.

Dos anos 90 até o momento houve o crescimento exponencial da tecnologia de informação (Sharda et al., 2019). O uso de informática passou a ter protagonistas sucessivos: inteligência de negócios, *data warehouse*, *big data* e outros. Estes dois últimos mencionados tiveram por base a alta capacidade de armazenamento de dados em proporções que desafiam a análise destes pelos métodos existentes. Ao longo desta evolução, a guerra por mercado na oferta de sistemas gerou diversos títulos com maior ou menor peso, podendo alguns destes nomes populares serem usados para denominar a época.

Frente ao desafio gerado pelo *big data*, entendido como conjunto de dados com características exacerbadas (Filatro, 2021): volume (quantidade), variedade (formatos), velocidade (quão rápido são gerados), veracidade (acurados), valência (conexões internas) e valor (significado), houve o desenvolvimento do termo *data analytics*. Este visa gerar criar/aprimorar técnicas e metodologias eficientes e eficazes para extrair valor da gigantesca quantidade de dados disponíveis e, continuamente, criados.

Para tanto, um novo ramo de estudos foi inaugurado já com alta demanda de profissionais. O analista de dados de *big data* e o cientista de dados utilizam conhecimentos de diversas áreas (Jessica, 2022), destacando-se a ciência da computação, a matemática/estatística e a área de conhecimento (Figura 1). Para o presente estudo a área de conhecimento será ciências contábeis.

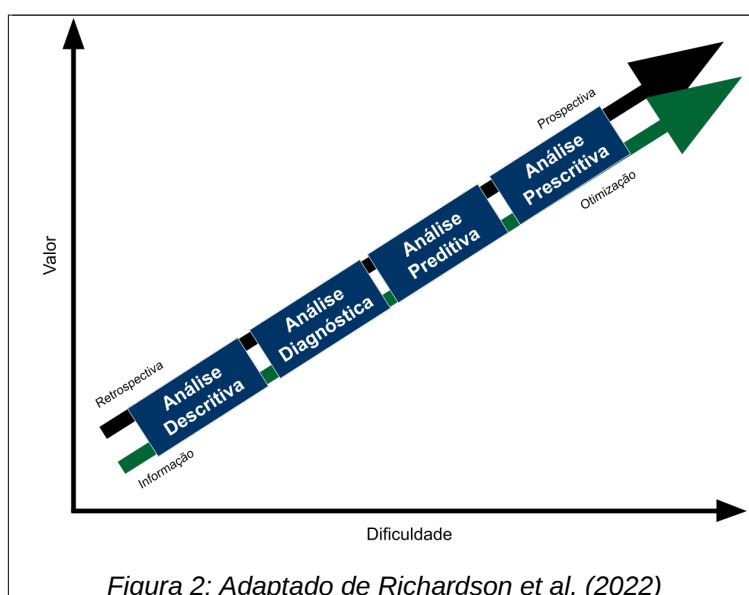
Em virtude disto, é prática comum o uso de equipes multidisciplinares para os estudos de análise de dados (Bajeux-Besbainou, 2022). Assim, pode-se contar com profissionais especializados em cada ramo, maior segurança contra erros, maior número de colaboradores e pluralidade de opiniões sob óticas diversas. Como requisito para essas vantagens do trabalho em grupo, é imperioso que tais profissionais entendam o potencial de colaboração que as demais áreas podem trazer para o projeto, além de compartilharem uma linguagem comum e visão ampla das possibilidades do estudo.



Para o caso do contador, verifica-se que, durante a graduação atual, este já adquire os conhecimentos próprios da gestão de negócios e, parcialmente, conhecimentos de matemática e estatística, qualificando-se para realizar pesquisa tradicional. Contudo, esse método dificilmente atenderá os requisitos de quantidade de dados e velocidade de resposta (Filatro, 2021), perdendo espaço de aplicação na rotina do contador. O próprio uso da estatística sem o aprofundamento em ferramentas tecnológicas tende a perecer.

Para mitigar essa perda de atuação, é importante que o estudante internalize certa gama de conhecimentos ligados à ciência da computação, em especial, aqueles que detenham sinergia com os conhecimentos já ensinados na graduação frente às tecnologias atuais. Explica-se que não se trata de transformar o contador em um programador de grande desempenho, mas sim de aprimorar o conhecimento para que o contador tenha um entendimento amplo das possibilidades desta junção de disciplinas (Shinn, 2022), e, com isso, atue mais proativamente com essa nova realidade de dados.

O principal objetivo disto seria dotar o estudante de contabilidade com um leque maior de patamares de valor da informação produzida. Hoje, grande parte da informação produzida pela contabilidade tradicional é classificada (Richardson et al., 2022) como descritiva (aquela que reflete o passado com os dados disponíveis). Contudo, existe uma ordem de gradação de valor e dificuldade para outras classes (Figura 2): diagnóstica (busca explicar com ajuda de outros dados o porquê do ocorrido), preditiva (estima o comportamento do futuro com base no comportamento do passado) e prescritiva (indica o que deve ser feito para atingir um almejado resultado).



Essa gradação encontra base na busca de padrões dentro de uma base de big data que só são possíveis com o uso de técnicas e ferramentas computacionais modernas em conjunto com os conhecimentos de gestão, exigindo mais de cada ramo a cada degrau. É importante ter em mente que estes avanços já se encontram em uso e aprimoramento fora do Brasil (Bajeux-Besbainou, 2022; Ciampi et al., 2021; Shinn, 2022).

Literatura empírica

Os professores Claypton e Chopton (2018) apresentaram um artigo esclarecedor defendendo a integração das disciplinas de data analytics nos currículos de contabilidade. Para tanto, destacaram a revolução trazida pelo uso de software *Tableau* para a tomada de decisões no mundo corporativo e acrescentaram farta literatura que ampararam a implementação. Tudo isso levou à decisão de, no período de 2018-2019, acrescer ao currículo de contabilidade as seguintes matérias: introdução à data analytics, princípios de comunicação de dados e visualização e aplicações de análise de data analytics. Também foram colocadas matérias não obrigatórias que permite galgar o certificado de data analytics ainda na graduação. Esse estudo mostra a tendência mundial de ampliar os estudantes de contabilidade com tais competências.

Recente estudo realizado na maior economia do mundo (Richardson & Shan, 2019) buscou a opinião dos chefes de cadeira de contabilidade para o seguinte quesito: um curso de data analytics deve ser parte do curso de graduação de contabilidade. Como resultado, 90.7% das 150 respostas adquiridas afirmaram que sim. Também foi perguntado se os atuais cursos ministram tal disciplina no momento e, surpreendentemente, apenas 31.8% responderam afirmativamente. Essa interseção díspare reflete a distância entre o

considerado melhor para o aluno e a realidade. Como conclusão, os autores relataram que existe pouca margem para trabalhar a extensa carga horária dos cursos atuais.

No Brasil, um estudo (Nogueira et al., 2020) foi conduzido aplicando questionários em dois momentos para uma turma de 42 alunos de dois cursos de pós-graduação, utilizando uma gradação de 1 a 5 para os conhecimentos ligados ao trabalho de data analytics. Buscou-se aferir o incremento da sensação de ganho de conhecimento após uma aula de três horas utilizando *software* de análise de *big data*. Destacou-se uma sensível variação positiva para a maioria dos questionamentos após a ministração do estudo de caso, que teve duração diminuta. O estudo, apesar de contar com uma amostra pequena, foi eficaz para apresentar um indício que o contato com análise de dados apresenta favorecimento proporcionalmente grande para os participantes.

3. METODOLOGIA

Para cumprir o objetivo geral deste trabalho, será usada a sequência de procedimentos da metodologia IMPACT, detalhada a seguir, em um caso específico. Também serão utilizadas ferramentas tecnológicas de apoio, na apresentação do caso, a plataforma Economática, o notebook do Google, linguagem de programação Python e bibliotecas para Python¹ (pandas, numpy, plotly e sklearn).

Durante toda essa linha do tempo, diversas metodologias de tratamento de dados foram sugeridas e testadas no mundo dos negócios. Para o presente estudo, foi selecionado o método IMPACT (Richardson et al., 2022), nome que se trata de um anagrama formado pela letra inicial em inglês da premissa de cada fase, por se tratar de um modelo desenvolvido especificamente para gestão de negócios. Apresenta-se abaixo as seis fases deste ciclo defendido pelos autores como contínuo para o aprimoramento do conhecimento com base nos dados.

É importante mencionar que a separação do método em fases trata-se de um recurso meramente didático, sendo requisito ter uma visão ampla de todo o processo para participação eficiente no estudo, mesmo para aqueles colaboradores pontuais. Participações ativas com visão parcial da metodologia tendem a gerar descompasso, colocando em risco a eficácia do estudo.

Fase 1 – Identify the questions²

Destarte, os autores sugerem definir o escopo do projeto, determinando as perguntas a serem respondidas pelo estudo. Trata-se de um procedimento determinante para o andamento do trabalho, pois deste emanará a escolha dos dados relevantes, os processos necessários de ajuste da informação, os modelos matemáticos/estatísticos possíveis e, por fim, a resposta adquirida.

Para a eficácia do trabalho, alguns pontos devem ser considerados nesta fase: a quem o estudo será enviado (conselho diretor, equipe de auditoria, governo, usuários externos, entre outros); se o alcance do estudo é muito amplo ou muito específico; e como os resultados serão utilizados pelos gestores. Estes fatores dificultam em parte a escolha das perguntas, mas, por se tratar de procedimento inaugural dos trabalhos, requer certo empenho pois este irá conduzir doravante todo o estudo.

Neste ponto, o conhecimento do contador é de suma importância. Deste podem advir a escolha dos dados (financeiros ou não) relevantes, os índices a serem calculados, a

1 Pacotes de funcionalidades prontos para linguagem de programação.

2 Identificar as perguntas (em tradução livre).

existência de dados externos relevantes, entre outros. São contribuições que encontram fundamento na atividade base do contador, sendo de difícil conhecimento por profissionais puramente de outras áreas.

Fase 2 – Master the data³

Definidos os quesitos, passa-se a juntar os dados necessários. Alguns desafios podem ser encontrados: inexistência dos dados necessários, ausência de padronização, ausência de integralidade, dados não confiáveis, entre outros. A minoração destes problemas requer o empenho dos diversos gestores que podem ser motivados pela visão da importância dos dados ou pelo controle rígido, sendo a primeira opção menos onerosa.

Para a execução serão necessários alguns conhecimentos da forma como os dados são armazenados pelas empresas e, caso o profissional esteja disposto a se aprofundar, algum contato com linguagens de programação específicas de acesso e manipulação de bancos de dados.

Esta fase, que também pode ser entendida como ETL (*extraction, transformation and loading⁴*), consiste em obter os dados, aferir a integralidade, limpar registros indesejados e inserir os dados na ferramenta para análise. Essa ordem de ações foi determinada pela limitação dos computadores à época, sendo que hoje as ações podem ocorrer em ordem diversa.

Fase 3 – Perform test plan⁵

De posse do objeto e ciente do objetivo, inicia-se a execução dos testes com base no modelo matemático planejado. Trata-se de uma fase bem prática e o resultado encontrado pode influenciar (ou mesmo determinar) o próximo teste a ser feito. Neste ponto, as técnicas computacionais serão mais utilizadas e o acompanhamento sob a ótica da matemática e estatística dotará de confiança a pesquisa.

Diversas técnicas podem ser utilizadas na busca de padrões existentes na quantidade massiva de dados: classificação, regressão, pesquisa de similaridade, agrupamento, entre outras. A escolha deve ser feita caso a caso e conta com a experiência e conhecimento dos pesquisadores.

3 Dominar os dados (em tradução livre).

4 Extração, transformação e leitura (em tradução livre).

5 Executar plano de teste (em tradução livre).

Fase 4 – Address and refine results⁶

Galgados os primeiros resultados, deve-se iniciar a revisão e aprimoramento destes. Os dados comportam diversas abordagens que podem ser realimentadas pelos resultados adquiridos. Também podem ocorrer novas discussões para explorar (ou especificar) outras questões a serem respondidas em complementação com base na opinião da equipe ou colaboradores, comportando reanálise e nova aplicação de testes.

Essa fase, além de acrescentar mais segurança ao estudo, aprimora os resultados possíveis, otimizando portanto o ganho de valor oriundo dos trabalhos.

Fase 5 e 6 – Communicate insights and track outcomes⁷

Tratam-se das fases finais do ciclo. A escolha da forma de apresentação das percepções demanda um conhecimento do público-alvo, bem como as prováveis interpretações que serão dadas. O ato de transformar o enorme conjunto de dados em uma informação resumida, mas que transpareça as peculiaridades relevantes para os tomadores de decisão exige experiência e análise do conteúdo informacional.

É fato que demais gestores podem colaborar nesta fase, contudo o profissional contador, pela sua própria forma de atuação, tende a ter papel de destaque neste planejamento. Ele consegue estimar quais variáveis são mais relevantes para o gestor, inclusive indicar dependência que precisa ser apresentada também.

Também é relevante o acompanhamento dos resultados. Seria contra produtivo utilizar as percepções atingidas apenas para uma decisão pontual. O conhecimento gerado pode ser usado como norteador durante considerável intervalo de tempo extraindo maior valor do estudo. O acompanhamento se mostra importante tanto para indicar novas pesquisas, quanto para confirmar se ele ainda está adequado para embasar decisões. As respostas destas perguntas podem ensejar a interrupção do uso ou mesmo um novo ciclo de estudos a partir do conhecimento já existente.

Para esse acompanhamento, os autores recomendam rotina de relatórios indicando se os novos dados estariam dentro do previsto/esperado, enfatizando que o uso de *dashboards* atualizados em tempo real ajudam nesta última fase.

Esta seria a descrição sucinta das fases do ciclo *IMPACT* (Richardson et al., 2022), tratando-se de uma rotina didática apta a ser implementada na maioria dos projetos de análise de dados para as mais diversas empresas, podendo inclusive ser adaptado às especificidades apresentadas pelo ente.

6 Abordar e refinar os resultados (em tradução livre).

7 Comunique percepções e acompanhe os resultados (em tradução livre).

4. RESULTADO

Desta forma, para ilustrar a gama de possibilidades que o *data analytics* pode trazer para a atuação do profissional de contabilidade, será apresentada uma análise de dados norteada pelo método *IMPACT*, usando o ambiente de programação Notebook do Google e a linguagens de programação Python. Pretende-se mostrar que o uso de tais tecnologias encontram-se mais facilitadas atualmente, possibilitando que o estudante aprofunde-se no conhecimento e uso de técnicas computacionais gerais, não se limitando à condição de usuário final imposta por softwares presentes no mercado.

Para a fase 1, optou-se por analisar as demonstrações de resultado da Petrobras (PETR4 – Bovespa), mais especificamente, as contas de Receitas Operacionais Líquidas, Custo Produtos Vendidos e LAIR, do período de 2004 a 2021. Pretende-se visualizar eventuais oscilações intercorrentes, bem como projetar simplificada os próximos anos.

Para a fase 2, foram importados dados da plataforma Economatica e transformados para o correto reconhecimento e manipulação necessárias. Trata-se da fase mais trabalhosa do procedimento, pois, mesmo em uma base de dados própria para tais pesquisas, não existe uma padronização realmente aplicada para registro de datas, numerais e demais campos. Mesmo utilizando ferramentas não comerciais, foram necessários diversos ajustes:

- Filtrar as linhas que realmente eram necessárias para a pesquisa;
- Ajustar os nomes das colunas;
- Converter os dados para o formato certo;
- Substituir os dados ausentes por valor zero (pesquisando se isso afetaria os resultados).

Todos esses ajustes foram realizados por linhas de comandos em Python, permitindo assim que a execução pudesse ser aplicada inclusive para uma massiva quantidade de dados, fato esse que não é possível para as tradicionais planilhas eletrônicas. Segue o exemplo de alguns comandos:

```
# Leitura dos dados
df = pd.read_csv("Petrobras_PETR4_Bovespa.csv")
df = df[183:209]

# Ajuste dos nomes da coluna
df.columns = df.iloc[0]
df = df[1:]
```

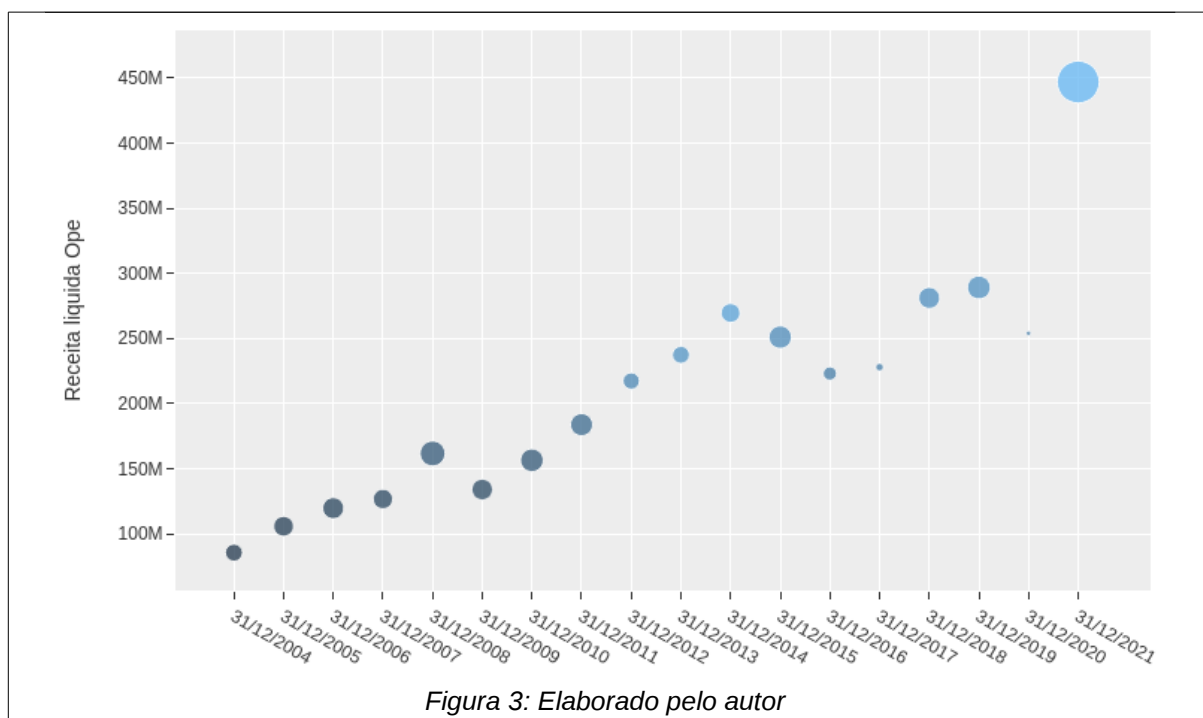
```
# Substituição de caracteres e troca de formato em laço de
repetição
for coluna in lista_colunas_float:
    df[coluna] = [x.replace(",",".") for x in df[coluna]]
    df[coluna] = [x.replace("-","0") for x in df[coluna]]
```

O conjunto de comandos que melhor apresenta essa capacidade de trabalhar com big data foi o último acima, introduzido pela anotação “substituição de caracteres e troca de formato em laço de repetição”. Ele consistia em percorrer todos os valores de colunas listadas na variável `lista_colunas_float` e executar as ações determinadas abaixo. Este é um recurso poderoso e que pode ser aplicado em conjuntos pequenos de dados ou em conjuntos com dezenas de milhares de linhas e/ou colunas.

Para o profissional, ter essa visão computacional, seja pela prática direta ou pela participação em estudo que faz uso, amplia e aprimora a interação que este terá com as mais diversas tecnologias existentes em uso.

A fase 3 iniciou-se com a criação de um gráfico (figura 3) com quatro variáveis sendo apresentadas: data (eixo x), Receitas Operacionais Líquidas (eixo y), Custo Produtos Vendidos (cor) e LAIR (tamanho da marcação). O comando para criação contou com poucas linhas de código:

```
fig_pontos = px.scatter(df_2004_dez, y="Receita liquida Ope",
    color="Custo Produtos Vendidos (-)", size="LAIR (=)")
fig_pontos.show()
```



Essa visualização permitiu entender que existe uma gradação positiva com algumas oscilações para as variáveis em análise.

Como fase 4, decidiu-se mostrar os pontos separados de cada variável e, com base nestes, calcular a regressão linear para comparar as retas de cada uma entre as demais. Para isso, as seguintes linhas de código foram suficientes para calcular a reta das Receitas Operacionais Líquidas, sendo repetido para as duas outras:

```
reg_receita = linear_model.LinearRegression()
reg_receita.fit(df_2004_dez[["Ano"]],df_2004_dez["Receita
liquida Ope"])
print(reg_receita.coef_[0])
print(reg_receita.intercept_)
```

Após o cálculo de cada uma, foi executado o seguinte código para juntar em uma única figura todos os valores das variáveis e as retas de regressão calculadas. Será disposto abaixo apenas o componente de uma variável e uma reta, pois seria mera repetição para as demais:

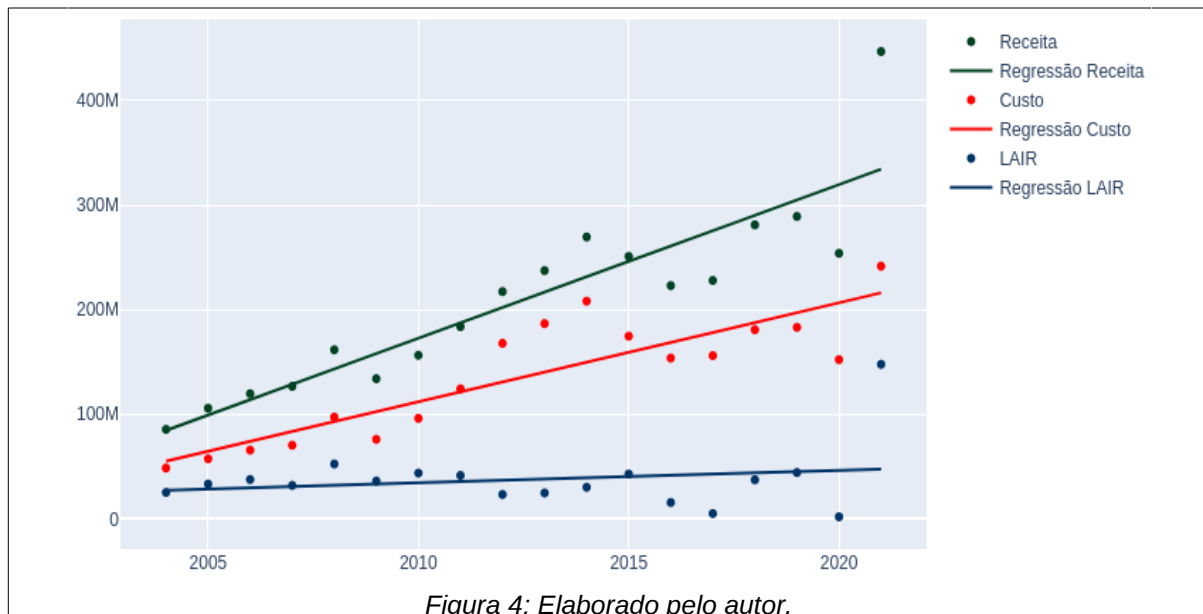
```
# Criando o objeto figura
fig_final = go.Figure()

# Adicionando pontos
fig_final.add_trace(go.Scatter(x = df_2004_dez["Ano"],
                               y = df_2004_dez["Receita liquida Ope"],
                               mode="markers",
                               marker_color="rgba(00, 66, 33, 1)",
                               name="Receita"))

fig_final.add_trace(go.Scatter(
    x = df_2004_dez["Ano"],
    y = reg_receita.coef_[0] * df_2004_dez["Ano"] +
    reg_receita.intercept_,
    name='Regressão Receita',
    mode='lines',
    marker_color='rgba(00, 66, 33, 1)'
))

fig_final.show()
```

Após realizar o mesmo para todas as colunas em estudo, foi publicado o gráfico final (figura 4), que facilita a visualização tanto de pontos que destoam dos demais, quanto da inter-relação entre as retas calculadas. Caberia, neste momento, elaborar um relatório, apresentação ou outro meio de divulgação para entregar os resultados (fase 5).



Também é importante acompanhar se os resultados se confirmam com o passar do tempo (fase 6), podendo ser fonte de futuras pesquisas validando, criticando, especificando ou mesmo aprimorando o conhecimento construído. Durante os trabalhos, diversas peculiaridades foram observadas: pontos muito fora da curva, variáveis que se mantiveram constantes frente às demais e oscilações positivamente bruscas das três em conjunto, todas podem motivar mais aprofundamento ou pesquisas paralelas.

Coloca-se que a quantidade de dados utilizados neste exercício não reflete um conjunto massivo de dados. Tal opção diminuta foi feita para manter-se alinhado com o objetivo do trabalho de destacar as possibilidades que o estudo de data analytics traria aos estudantes e, em segundo argumento, para utilizar ferramentas o mais livre de custo possível. Acrescenta-se que as rotinas aqui descritas podem ser aplicadas em big data sem maiores ajustes, apenas demandando maquinário de geração recente e muito mais tempo de processamento para cada rotina.

5. CONSIDERAÇÕES FINAIS

Percebe-se que o contabilista, munido dos conhecimentos destacados, estaria apto a participar da extração do valor otimizado das informações digitais com que tem contato nos diversos momentos da cadeia produtiva destes, simplesmente, pela boa gestão: decidindo quais dados devem ser armazenados para pesquisa posterior; inferindo quais quesitos empresariais podem ser respondidos; solicitando mais acertadamente estudos favoráveis e possíveis de serem implementados com as informações disponíveis; planejar o organograma da empresa e o fluxo de informações internas com base neste importante ativo; participar ativamente em equipes de pesquisa nas decisões de modelagem matemática da questão; e diversas outras.

Importante mencionar que o elencado anteriormente não exige necessariamente a implementação de rotinas computacionais por escrita em linguagem de programação, requerendo apenas o conhecimento do que pode ser realizado com as informações digitais internas ou externas, eminentemente uma atividade de gestão. É fato que aquele que aprofunde nos estudos agregando conhecimentos computacionais e contábeis terá um ramo de atuação específico e promissor, mas todos os gestores da azienda aprimorariam a entrega de valor de seus serviços com os conhecimentos propostos neste trabalho.

Aponta-se também que, mesmo atuando afastado dos estudos de dados, o contato com a matéria de dados eleva a familiaridade que o aluno tem ao usar calculadoras financeiras, planilhas eletrônicas, sistemas de informática do governo e, entre outros, softwares especializados. Sendo quase impossível citar uma atuação de gestão que não utilize ferramentas tecnológicas, denota-se que este efeito secundário mencionado é mais que bem vindo para o estudante de graduação.

Para tais profissionais, o contato ainda na fase de graduação com disciplinas de análise de dados digitais alavanca a contribuição destes para a gestão em seus mais diversos ramos e em qualquer dos níveis de atuação. Ademais, participar de uma instituição atenta para essa digitalização e, por ser um assunto com foco global, com potencial para internacionalização (tanto em parcerias acadêmicas como intercâmbio estudantes/professores) eleva o peso de seu currículo no mercado.

O estudo apresenta como limitação a ausência de aprofundamento em diversos tópicos mencionados. Ocorre que para uma exposição do conjunto, foi necessário limitar explicações pontuais. Também é verdade que a análise sugerida poderia contemplar temas mais complexos, contudo o objetivo de apresentar a facilidade de interação com ferramentas de análise de dados determinou a escolha de prioridades.

O trabalho se mostra suficiente para levantar discussões quanto à necessidade de incluir os estudos mencionados na competência do contador do futuro, além de motivar um estudo em continuidade e mais aprofundado de temas específicos.

Por fim, coloca-se que a digitalização das mais diversas aziendas está ocorrendo agora, com diversas discussões, criação de normatização, desenvolvimento de novos indicadores, definição de padrões, entre outros. Não instruir o aluno para participar imediatamente nisto irá atrasar (senão tolher) a colaboração que estes podem trazer à mudança.

REFERÊNCIAS

- Bajeux-Besbainou, I. (2022). Business Schools Are the Key to the Intelligent Future | AACSB. <https://www.aacsb.edu/insights/articles/2022/01/business-schools-are-the-key-to-the-intelligent-future>
- Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, 123, 1–13. <https://doi.org/10.1016/j.jbusres.2020.09.023>
- Clayton, P. R., & Clopton, J. (2018). Business curriculum redesign: Integrating data analytics. *Journal of Education for Business*.
- Filatro, A. (2021). *Data Science Na Educação—Presencial, A Distância E Corporativa* (10 ed).
- Jessica, S. (2022, maio 6). Statistics for Data Science. Medium. <https://pub.towardsai.net/statistics-for-data-science-79584e7d6d9c>
- Nogueira, D. R., Albertin, E. A., Nasu, V. H., & Marques, C. (2020). Big Data no Ensino de Contabilidade: Uma experiência com estudantes de Pós-Graduação Lato Sensu. *Revista de Contabilidade do Mestrado em Ciências COntábeis da UERJ*, 25.
- Richardson, V. J., & Shan, Y. (2019). Data Analytics in the Accounting Curriculum. *Advances in Accounting Education: teaching and Curriculum Innovation*, 23.
- Richardson, V. J., Terrell, K., & Teeter, R. (2022). *Data Analytics for Accounting* (30 ed). Mc Graw Hill.
- Sharda, R., Delen, D., & Turban, E. (2019). *Business intelligence e análise de dados para gestão do negócio* (4aed). Bookman.
- Shinn, S. (2022). Developing Tech-Savvy Leaders | AACSB. <https://www.aacsb.edu/insights/articles/2022/05/developing-tech-savvy-leaders>