

***Business Intelligence: Data Science
aplicado à tomada de decisões em
cinema de pequeno porte***

Bruno Belluzzo

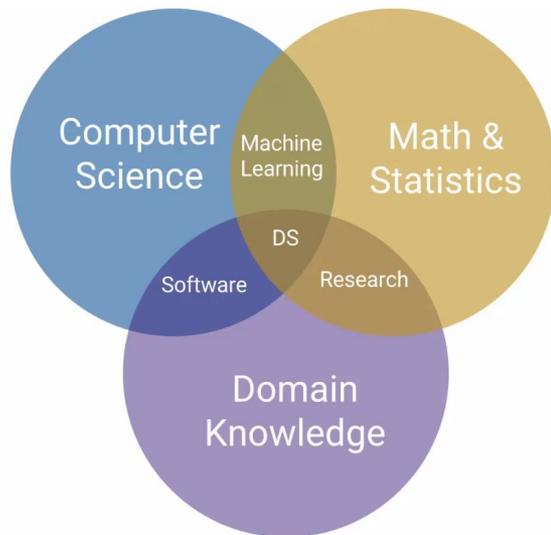


Problemas de cinemas pequenos

- *Indisponibilidade de cópias de filmes*
- *Limitação física*
- *Público limitado*
- *Concorrência de grandes cinemas próximos*
- *Exigências de distribuidoras que não se encaixam a realidade*

Ciência de Dados

A Ciência de Dados busca, a partir de uma quantidade grande e pesada de dados, gerar conhecimentos e informações relevantes para tomar decisões e fazer previsões, e não simplesmente interpretar os números. (L. COELHO, 2018)



Business Intelligence

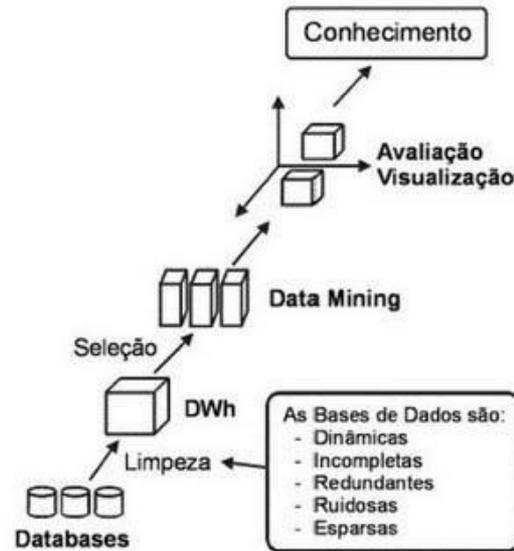
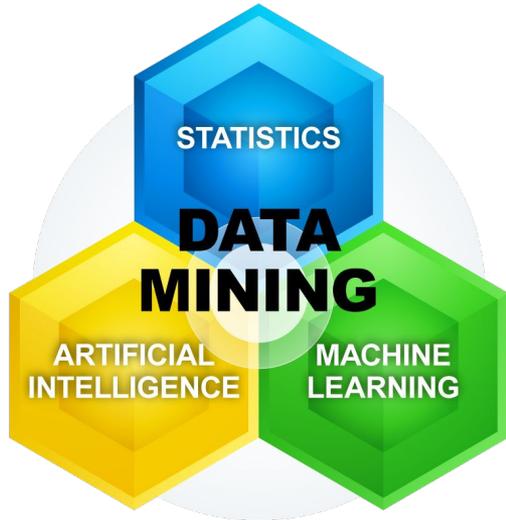
"BI é utilizado especialmente para a gestão da empresa, trabalhando dados do passado para traçar as tendências do futuro, fazendo previsões de vendas, por exemplo, a partir do comportamento do banco de dados de seus próprios clientes."

(CANALCOMSTOR, 2018)

O conhecimento obtido a partir do uso de Data Science é fundamental para apoiar o processo de tomada de decisão. A informação gerada pelas aplicações informáticas disponibiliza aos gestores um conjunto de indicadores sobre o negócio, que lhe dão indicações do que aconteceu no passado e lhe permitem traçar cenários para o futuro. (SANTOS; RAMOS, 2006)

Data Mining

"...o processo não-trivial de identificar, em dados, padrões válidos, novos, potencialmente úteis e ultimamente compreensíveis" (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996)



| Bibliotecas python para Ciência de Dados - Manipulação

NumPy

Menor consumo de memória para armazenamento dos dados, menor tempo de execução de instruções e a otimização de performance.

Pandas

Capacidade de lidar com grandes massas de dados e facilidade de uso, sem perder a eficiência.

| Bibliotecas python para Ciência de Dados - Visualização

➔ Matplotlib

O Matplotlib surgiu como uma alternativa do MATLAB para ser usado com Python. A representação dos dados por meio de gráficos de simples compreensão é importante para que os executivos da empresa consigam compreender como o cientista de dados identificou as soluções inteligentes que ele propôs para serem implementadas.

| Bibliotecas python para Ciência de Dados - Visualização

➔ Seaborn

Como diria Michael Waskom, no site oficial do Seaborn: "Se o Matplotlib tenta tornar as coisas fáceis, fáceis, e coisas difíceis possíveis, o Seaborn tenta tornar um conjunto bem definido de coisas difíceis fáceis também". É exatamente por esse motivo que torna essa biblioteca a favorita para visualização de dados estatísticos em python.

Coleta e preparação dos dados - Formulário

- *Utilização do Google Forms*
- *Formulário partido em 3 seções:*
 - *Perguntas gerais*
 - *Perguntas para quem já havia ido ao cinema*
 - *Perguntas para quem nunca tinha ido ao cinema*
- *Total de 30 perguntas*
- *Divulgação do formulário no Facebook do cinema*
- *Total de 447 respostas coletadas:*
 - *394 já haviam frequentado o cinema*
 - *53 nunca haviam frequentado o cinema*

Coleta e preparação dos dados - Formulário

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from collections import Counter
import regex as re
```

```
In [2]: df = pd.read_excel('../Pesquisa Cine Belluzzo.xlsx')
```

```
In [3]: df = df.iloc[:, 2:]
```

```
In [4]: new_columns = ['Idade',
                        'Sexo',
                        'RendaMensal',
                        'JaAssistiu',
                        'GenerosPreferidos',
                        'GenerosEvitados',
                        'Frequencia',
                        'MotivoFrequencia',
                        'EncorajaCinema',
                        'MelhoresDias',
                        'Companhia',
                        'PrecoAtualIngresso',
                        'PrecoIngressoNaoAdequado',
                        'TotalGasto',
                        'CostumaConsumir',
                        'Dublado/Legendado',
                        'MelhoresHorarios',
                        'BeneficiosFieis',
                        'NotaFilmesExibidos',
                        'NotaPrecoProdutos',
                        'NotaDivulgacao',
                        'NotaPromocoesCombos',
                        'NotaAtendimento',
                        'NotaSessao',
                        'NotaGeral',
                        'Sugestoes1',
                        'PorqueNaoFrequentou',
                        'VontadeDeComecar',
                        'Sugestoes2'
                       ]
```

```
In [5]: df.columns = new_columns
```

Coleta e preparação dos dados - Base de Dados do negócio

- Backup do programa de vendas
 - Unidades vendidas por produto/ingresso
 - Lucro bruto das vendas
- Levantamento de custos com os executivos

	Vendidos	Lucro Bruto	Preço unitário	Custo	Margem de lucro
DISQUETE	28	28	1	1,05	-5,00%
DIAMANTE NEGRO	46	92	2	1,88	6,00%
LACTA AO LEITE	39	78	2	1,88	6,00%
LAKA CHOCOLATE	16	32	2	1,88	6,00%
LAKA OREO	17	34	2	1,88	6,00%
TRENTO	44	88	2	1,87	6,50%
BATON PEQ	68	68	1	0,66	34,00%
LOLLO	21	42	2	1,23	38,50%
HERSHEYS	14	98	7	4,29	38,71%
BIS XTRA	5	20	4	2,32	42,00%
BATATA GD	128	1280	10	5,32	46,80%
DORITOS GD	33	330	10	5,32	46,80%
DORITOS PEQ	51	306	6	3,05	49,17%
SNICKERS	71	284	4	1,93	51,75%
SUCO	58	181	3	1,4	53,33%
SUFLAIR	83	415	5	2,30	54,00%
SENSACAO	52	208	4	1,83	54,25%
REFRIGERANTE	1711	8555	5	2,00	60,00%
BATATA PEQ	74	444	6	2,29	61,83%
FANDANGOS	35	140	4	1,52	62,00%
CHEETOS	69	276	4	1,52	62,00%
CHEETOS LUA	5	20	4	1,52	62,00%
AGUA	734	2936	4	1,13	71,75%

Análise dos Dados - Formulário

Pessoas que nunca frequentaram o cinema

- **Idade**
 - 52,8% entre 15-24 anos
 - 22,6% entre 25-34 anos
- **Sexo**
 - 75,5% mulheres
 - 24,5% homens
- **Motivos de nunca ter ido**
 - 49,1% não tem tempo
 - 30,9% preço inacessível
 - 16,4% não mora na cidade

Análise dos Dados - Formulário

Pessoas que nunca frequentaram o cinema

- **Renda**
 - 43,4% não possui renda mensal
 - 28,3% recebe entre 0-1000 reais
 - 18,9% recebe entre 1000-2000 reais

Análise dos Dados - Formulário

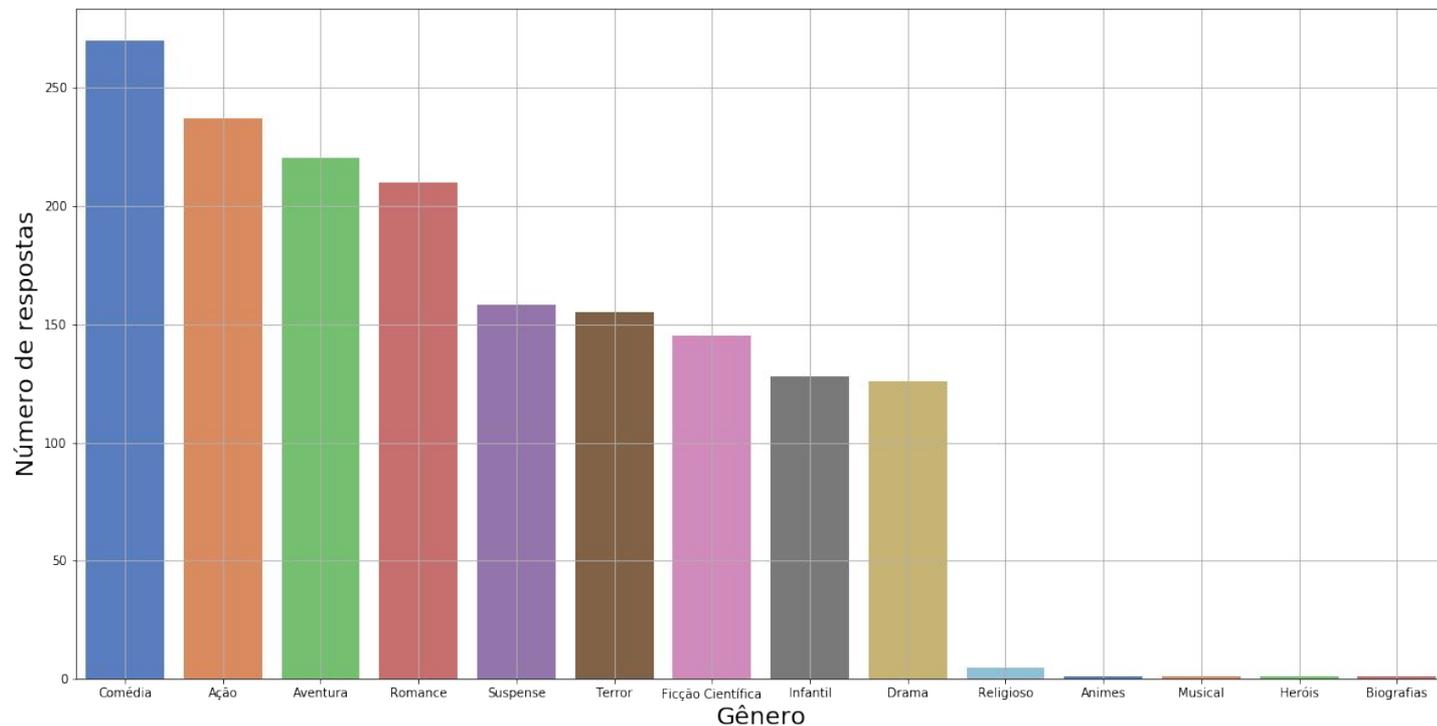
Pessoas que já frequentaram o cinema

- **Idade**
 - 56,1% entre 15-24 anos
 - 20,8% entre 25-34 anos
 - 9,4% entre 35-44 anos
- **Sexo**
 - 72,1% mulheres
 - 27,9% homens
- **Renda**
 - 33,0% não possui renda mensal
 - 22,3% possui entre 0-1000 reais
 - 25,6% possui entre 1001-2000 reais

Análise dos Dados - Formulário

➔ **Pessoas que já frequentaram o cinema**

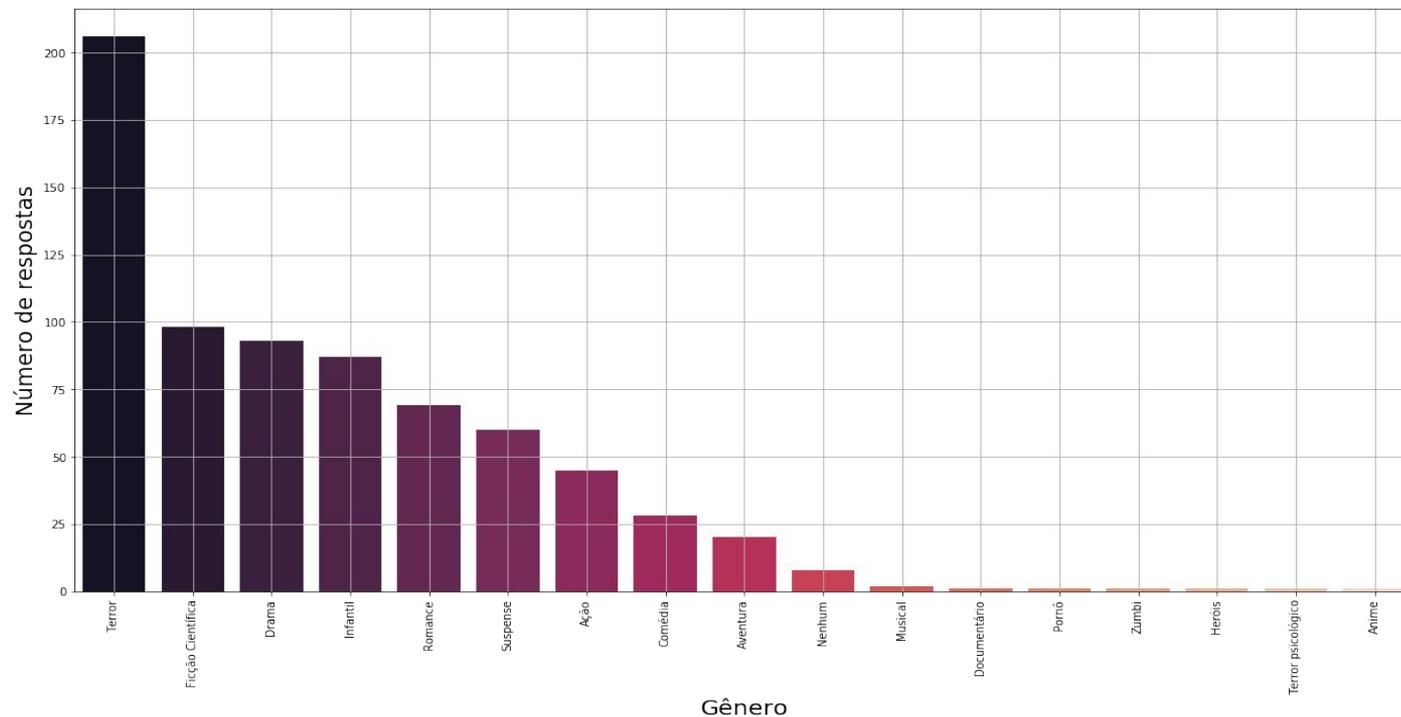
- **Gêneros preferidos**



Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

- Gêneros evitados



Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

- **Frequência com que vai ao cinema**
 - 46,7% vai de vez em quando, não mensalmente
 - 18,8% muito raramente
 - 17,0% uma vez no mês
 - 15,0% 2 ou 3 vezes no mês
 - 2,5% toda semana
- **Motivo da frequência**
 - 36,9% financeiro
 - 23,2% tempo
 - 20,2% não tem interesse no filme

Análise dos Dados - Formulário

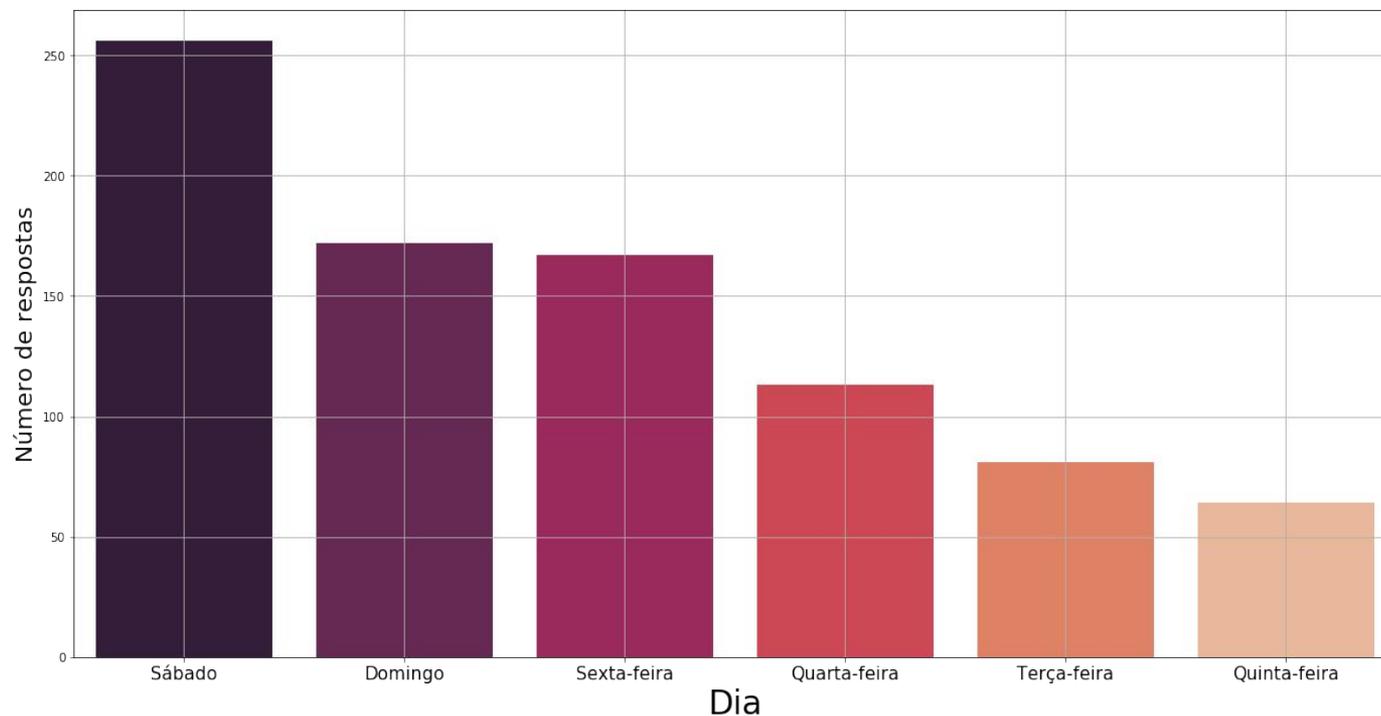
Pessoas que já frequentaram o cinema

- **O que encoraja ir ao cinema**
 - 26,6% novos lançamentos
 - 21,1% melhor qualidade/experiência visual
 - 19,4% assistir junto de amigos/família/namorado(a)
 - 14,2% ambiente da sala
 - 9,9% tamanho da tela
 - 8,6% assistir ao filme sem interrupção

Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

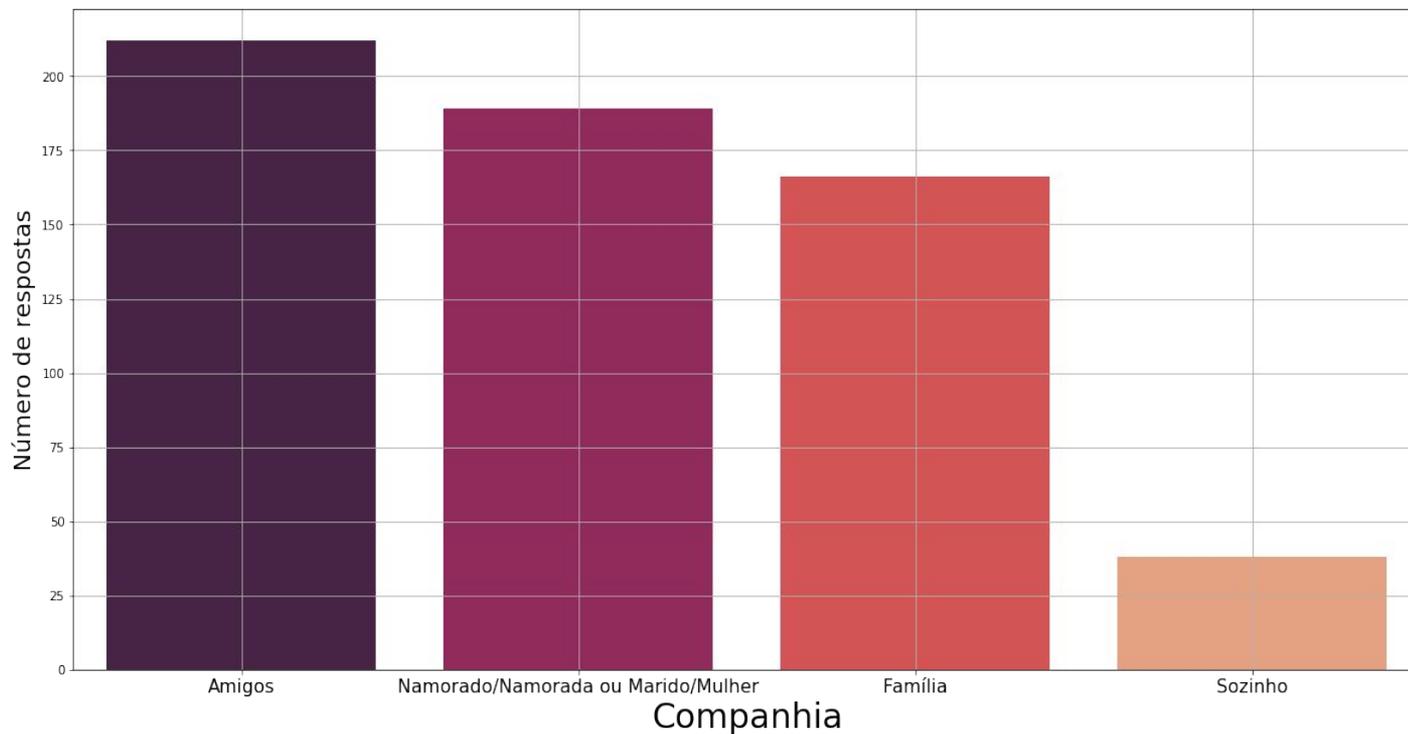
- **Melhores dias**



Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

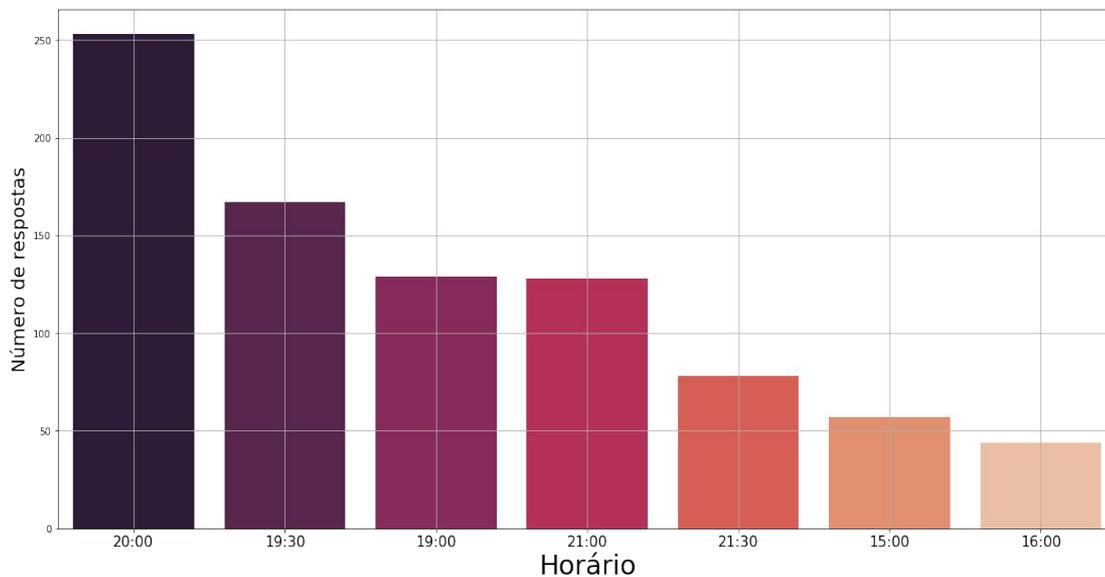
- **Companhia**



Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

- Áudio preferido
 - 76,4% dublado
 - 23,6% legendado
- Horários preferidos



Análise dos Dados - Formulário

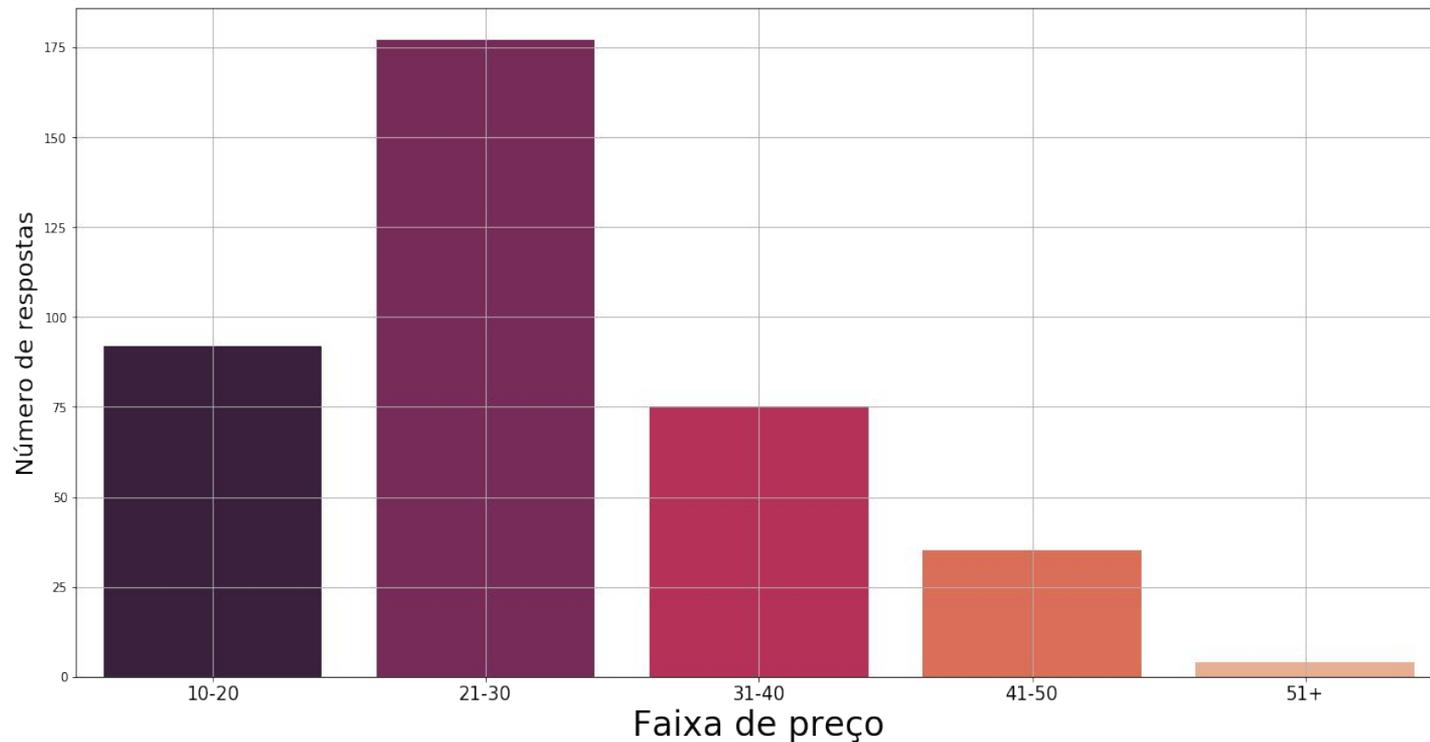
Pessoas que já frequentaram o cinema

- **Considera o preço do ingresso justo**
 - 80,2% consideram justo
 - 19,8% não consideram justo
- **Sugestões para o preço do ingresso**
 - 64,9% R\$15,00 inteira
 - 24,6% R\$10,00 inteira e meia
 - 10,5% R\$10,00 inteira

Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

- Gasto considerado justo por sessão



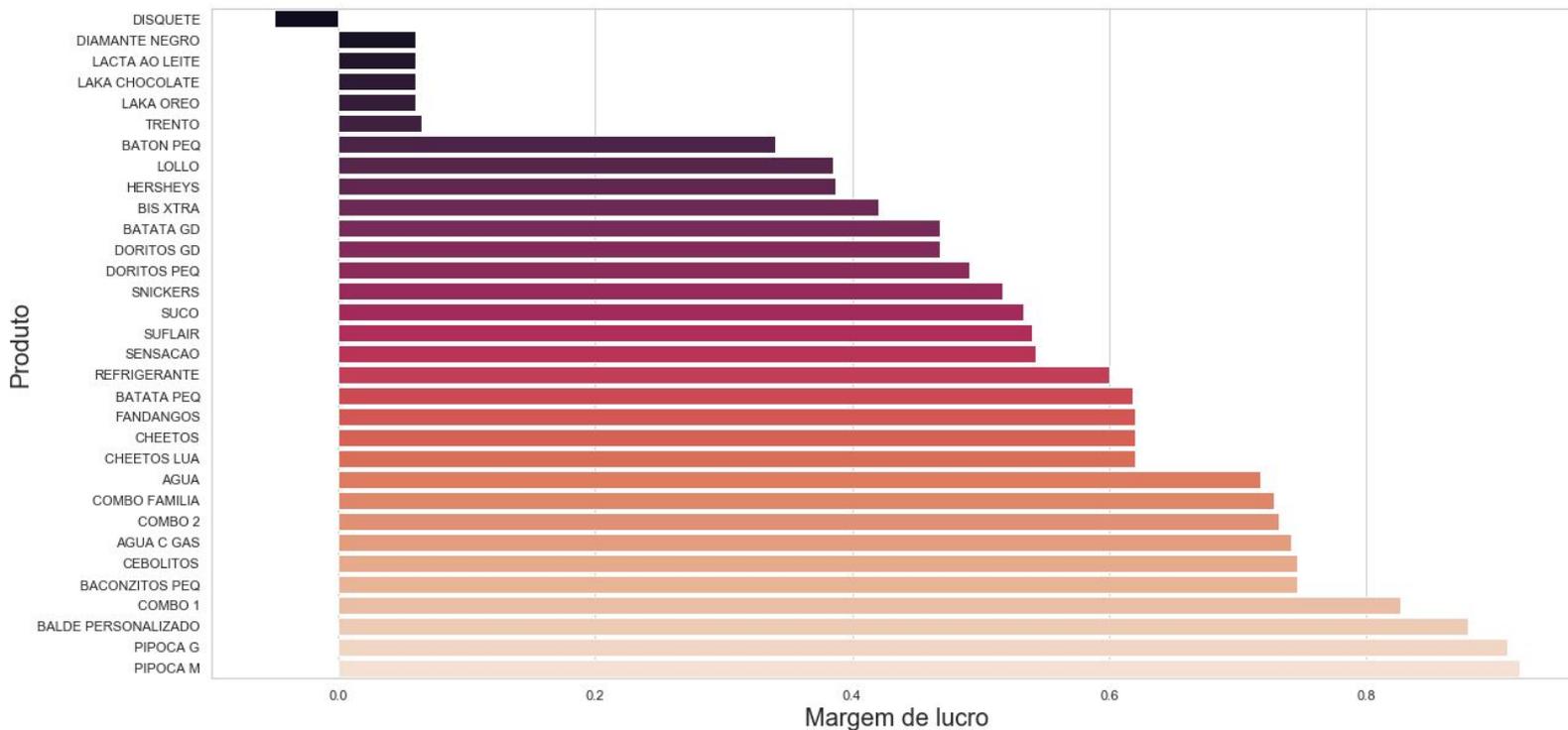
Análise dos Dados - Formulário

Pessoas que já frequentaram o cinema

- **Benefícios para fiéis incentivaria a ir mais ao cinema**
 - 79,9% sim
 - 19,3% talvez
 - 0,8% não
- **Notas gerais**
 - *Nível de satisfação em relação aos filmes exibidos: 4.45*
 - *Preço dos produtos oferecidos para consumo: 3.48*
 - *Divulgação dos filmes: 4.3*
 - *Promoções e combos oferecidos: 3.8*
 - *Atendimento no cinema: 4.51*
 - *Sessão no geral: 4.31*
 - *Nota geral para o cinema: 4.47*

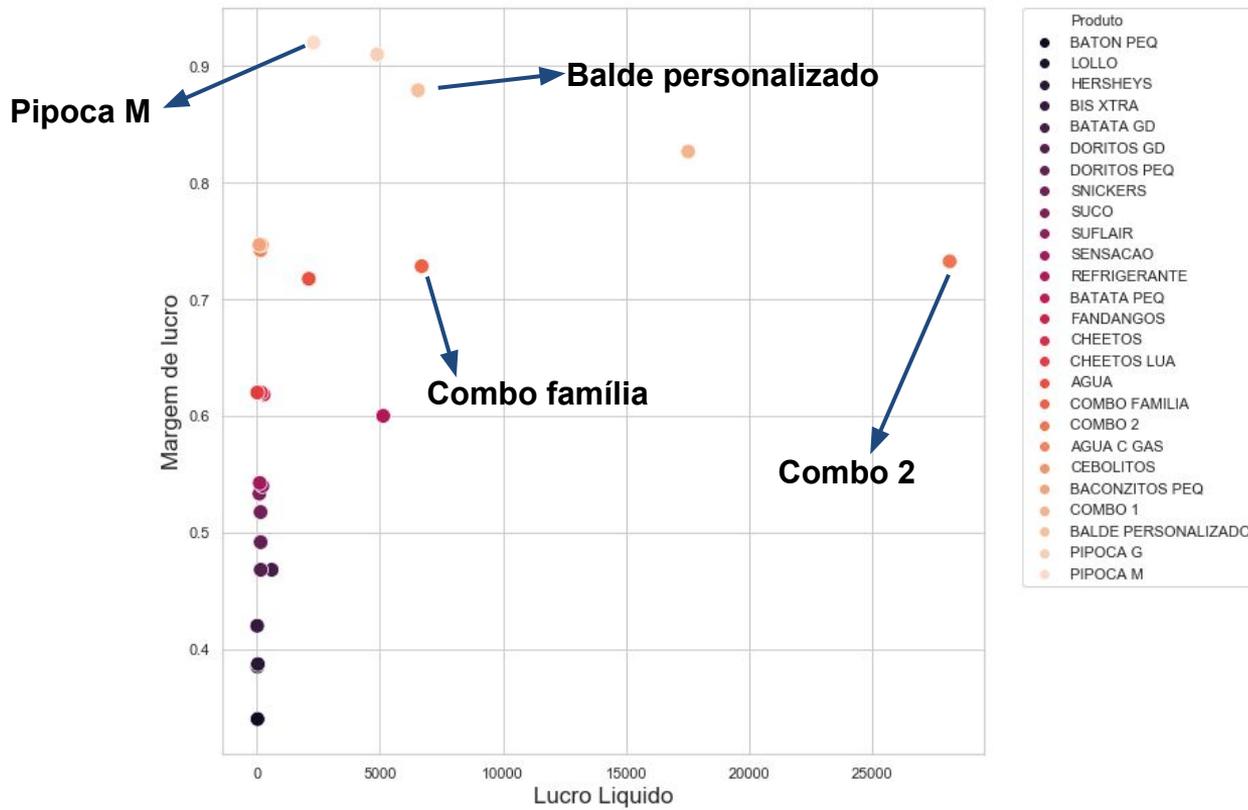
Análise dos Dados - Base de Dados do negócio

Margem de lucro por produto



Análise dos Dados - Base de Dados do negócio

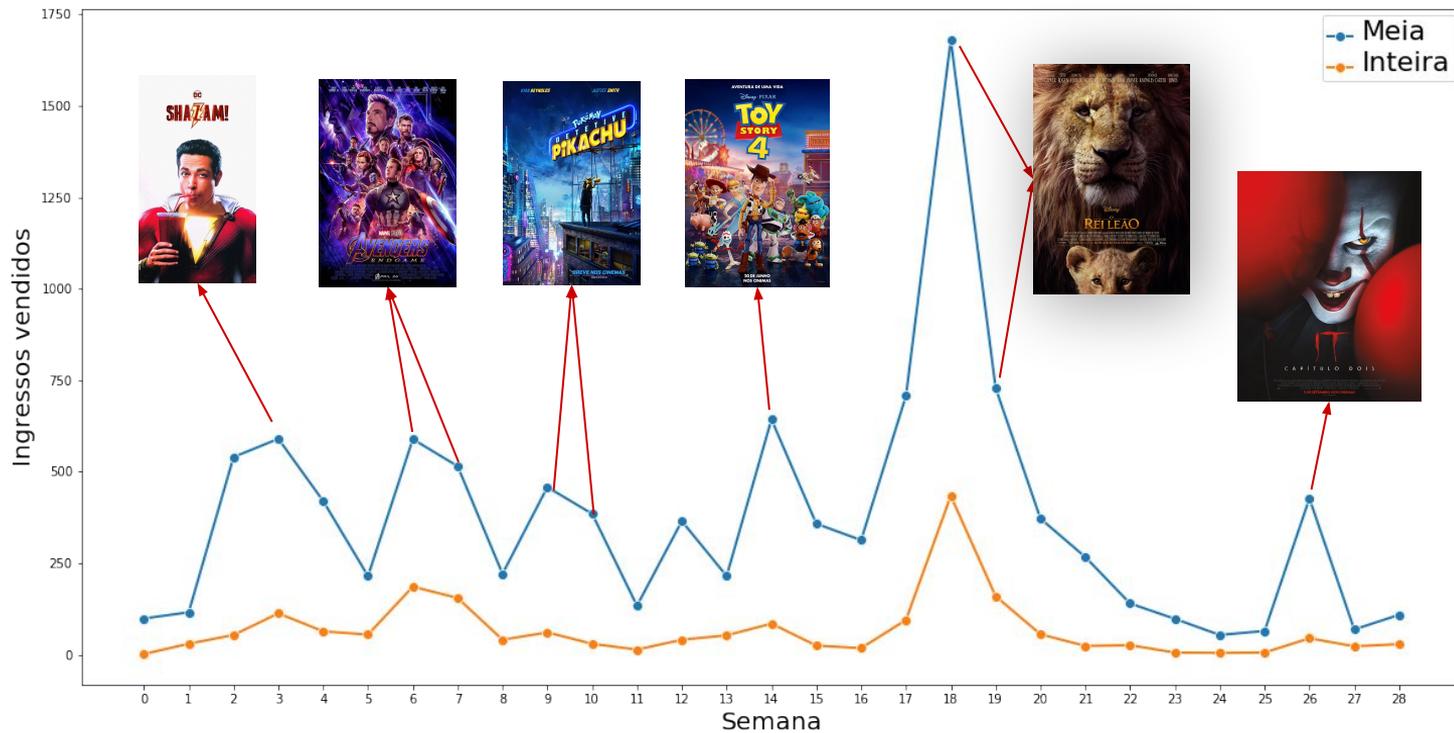
Margem de lucro X Lucro líquido



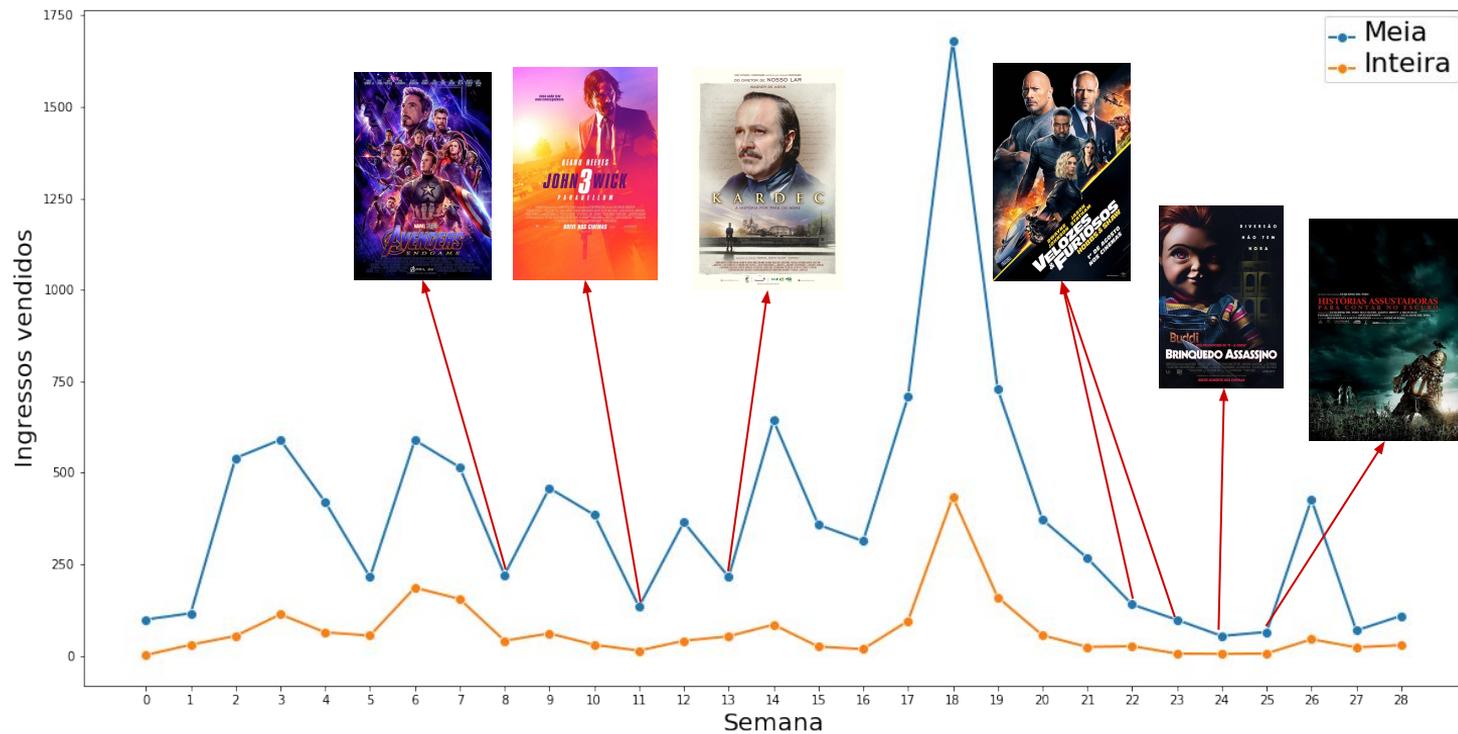
| Resultados

- *Para validar a fidelidade das informações levantadas, foi feita uma coleta de dados de semana por semana em cima do programa de vendas do cinema, e os resultados comprovaram os insights levantados pela pesquisa.*

Resultados

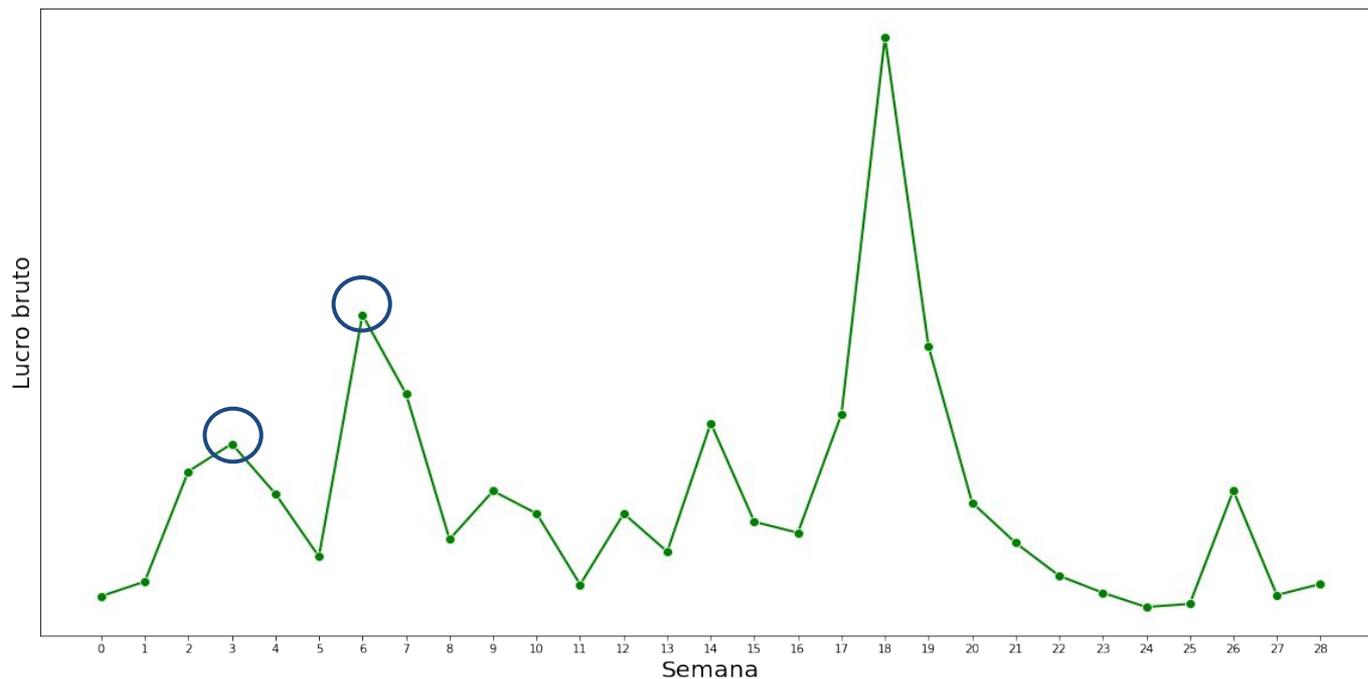


Resultados



Resultados

- *A diferença de lucro entre a semana 3 e 6 mostra a diferença que faz o público que paga inteira, uma pequena diferença na venda desse tipo de ingresso causa uma grande diferença no lucro*



Conclusão

- *Cabe ao cientista de dados coletar, manipular e analisar os dados, minerar possíveis soluções com base nas análises e saber repassar esse conhecimento aos donos do negócio de forma simples e bem estruturada.*
- *Os executivos passaram a possuir as informações necessárias para tomar as decisões no futuro e obtiveram uma visão do perfil dos clientes e do empreendimento que eles não tinham anteriormente.*
- *Os donos do cinema ficaram extremamente satisfeitos com os resultados obtidos na pesquisa, sendo assim foi solicitado a utilização de ciência de dados em uma outra empresa dos mesmos.*

Referências

- L. COELHO. CIÊNCIA DE DADOS: O QUE É, CONCEITO E DEFINIÇÃO. 2018. Disponível em: <<https://www.cetax.com.br/blog/data-science-ou-ciencia-de-dados/>>. Acesso em: 14 mai. 2019.
- CANALCOMSTOR. QUAIS AS DIFERENÇAS ENTRE BUSINESS INTELLIGENCE E DATA SCIENCE? 2018. Disponível em: <<https://blogbrasil.comstor.com/quais-as-diferencas-entre-business-intelligence-e-data-science>>. Acesso em: 14 mai. 2019.
- SANTOS, M. Y.; RAMOS, I. Business Intelligence: Tecnologias da informação na gestão de conhecimento. [S.l.]: FCA-Editora de Informática, Lda, 2006.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. The kdd process for extracting useful knowledge from volumes of data. Communications of the ACM, ACM, v. 39, n. 11, p. 27–34, 1996.

| Obrigado!

