

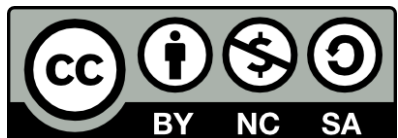
D1INT – Introdução à Ciência de Dados

2021.1

Aula 01 Introdução



Prof. Everton Silva
everton.silva@ifsp.edu.br



Aula de Hoje

- Pensamento Analítico de Dados (Data-Analytic Thinking)
- Exemplo: Mall Costumer Segmentation Data
- Data Science X Data Mining
- Exemplo: Furacão Frances - 2004

Data-Analytic Thinking

- Identificar e definir problemas, extrair informações dos dados e desenvolver soluções;
- Grandes investimentos, nos últimos 15 anos, em infraestrutura de negócios para melhorar a capacidade de coletar dados;
- Ampla disponibilidade de dados levou a um interesse crescente em métodos para extrair informações úteis e conhecimento dos dados;
- As empresas em quase todos os setores estão focadas na exploração de dados para obter vantagem competitiva;



Data-Analytic Thinking

- Antigamente equipes de estatísticos e analistas faziam a exploração manualmente;
- Computadores mais poderosos, a difusão e o desenvolvimento de algoritmos possibilitaram análises mais amplas e profundas;
- Isso deu origem a aplicação comercial dos princípios de ciência de dados e técnicas de aprendizado de máquina;
- Melhorar o pensamento analítico de dados, desenvolve intuição sobre como e onde aplicar a criatividade e o conhecimento do domínio;



Data-Analytic Thinking

- Algumas aplicações mais amplas das técnicas de mineração de dados.



Marketing direcionado
(Target marketing)



Publicidade online
(Online advertising)

Data-Analytic Thinking


- Algumas aplicações mais amplas das técnicas de mineração de dados.

Frequentemente comprados juntos



Preço total: R\$2.725,26

Adicionar os três ao carrinho

 Estes itens são enviados e vendidos por vendedores diferentes. [Ver detalhes](#)

- ✓ **Este item:** Câmera Digital EOS Rebel T100 18-55mm f/3.5-5.6 IS III BR, Canon, Preto R\$2.399,00
- ✓ Cartão de Memória SanDisk Micro SD, 64Gb R\$69,00
- ✓ Porta Camera Média preta DCB304, Case Logic, Acessórios para Câmeras Digitais, Preta R\$257,26

Venda cruzada
(Cross-selling)



Análise de Crédito
(Credit Score)

Data-Analytic Thinking

- Algumas aplicações mais amplas das técnicas de aprendizado de máquina



Detecção de Fraude
(Fraud Detection)



Vendas de Varejistas –
Imagens de Satélites
(Retailers' Sales)

Exemplo – Mall Customer Segmentation Data

- Conjunto de dados criado apenas para fins de aprendizado dos conceitos de segmentação de clientes (análise de cesta de compras).
- São 200 observações no total;
- Suponha que tenhamos que analisar a base de dados separando os clientes em grupos com características semelhantes para entender melhor esses grupos;



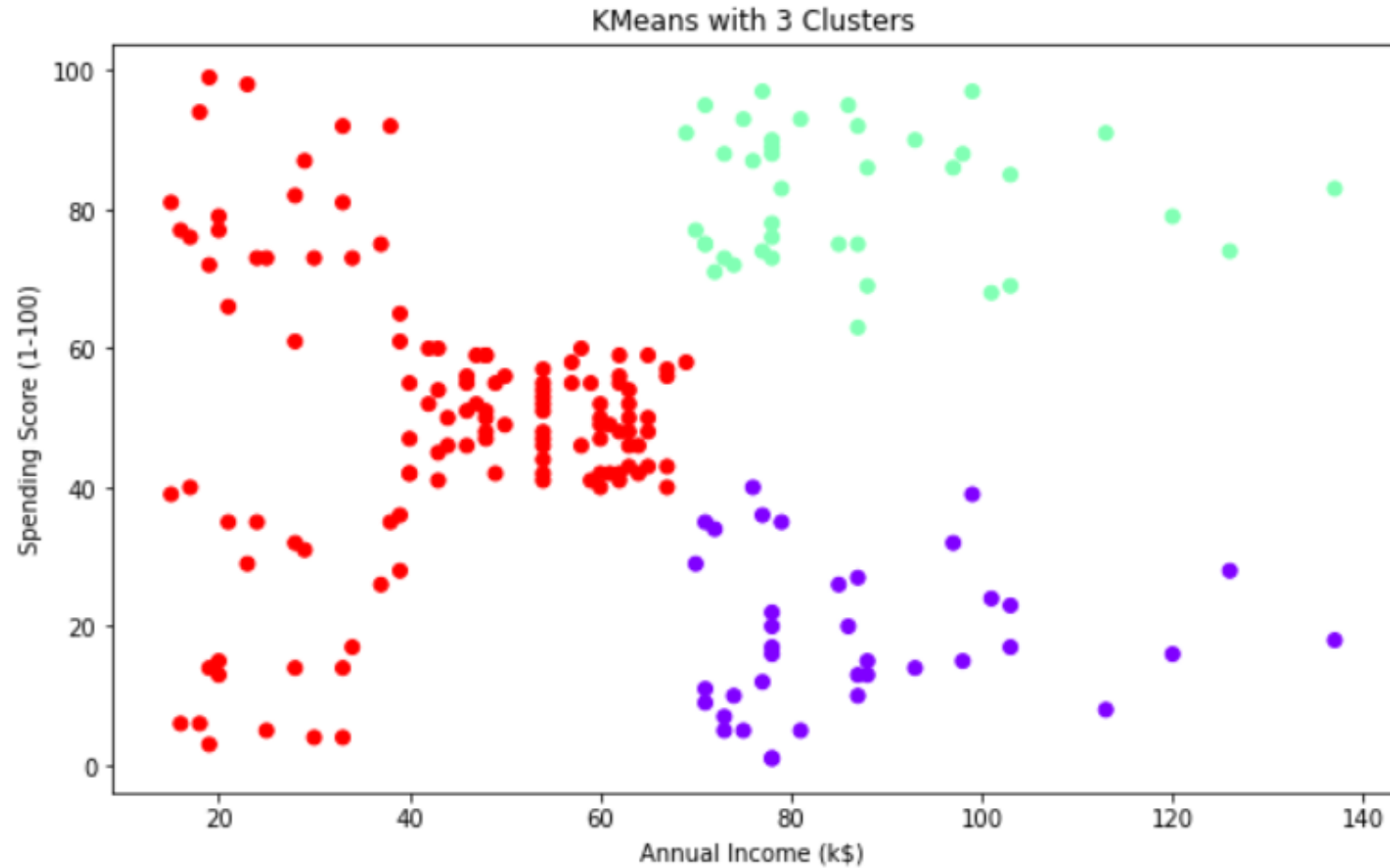
Exemplo – Mall Customer Segmentation Data

	CustomerID	Gender	Age	Annual Income (k\$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40
...
195	196	Female	35	120	79
196	197	Female	45	126	28
197	198	Male	32	126	74
198	199	Male	32	137	18
199	200	Male	30	137	83

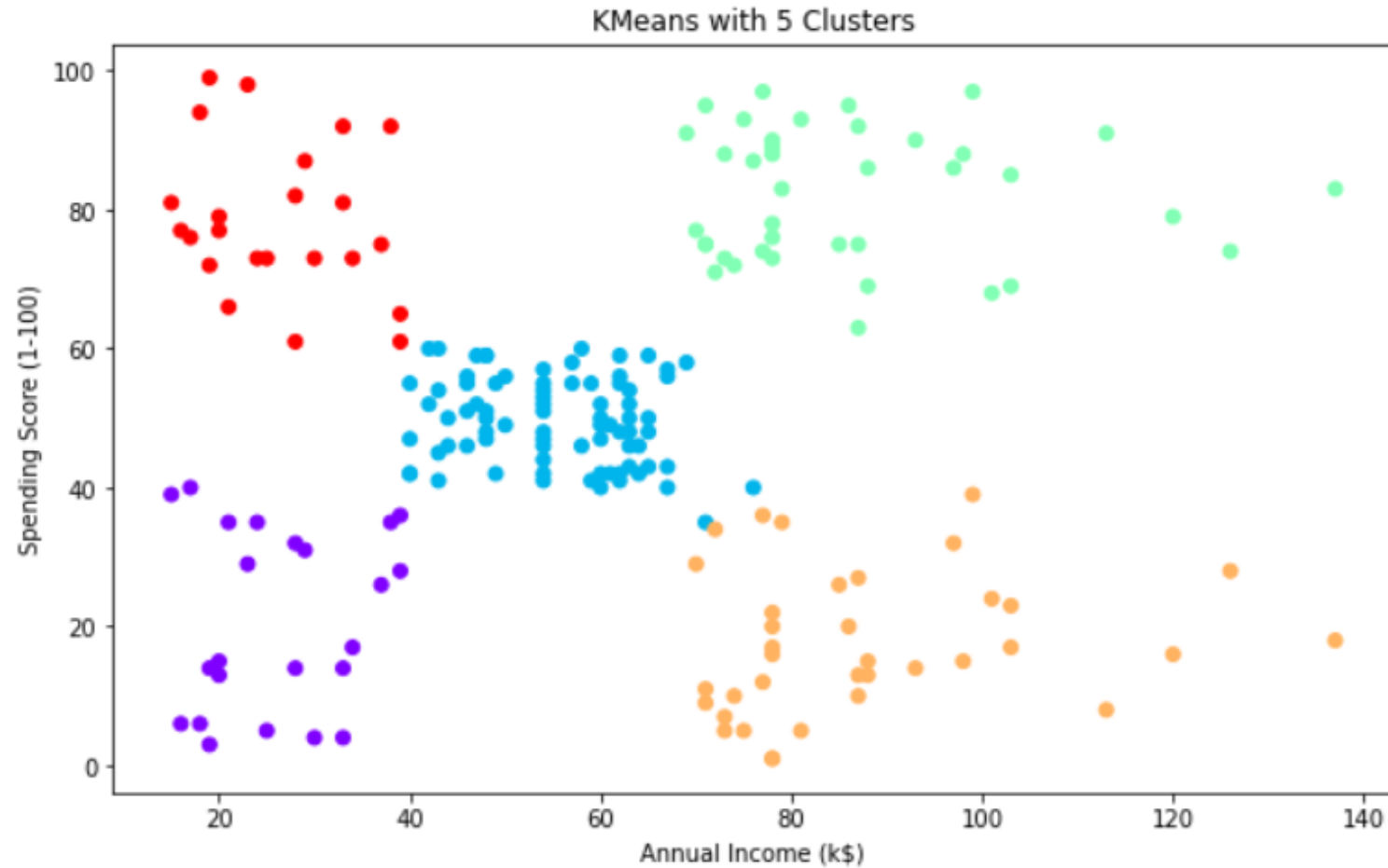
200 rows × 5 columns



Exemplo – Mall Customer Segmentation Data

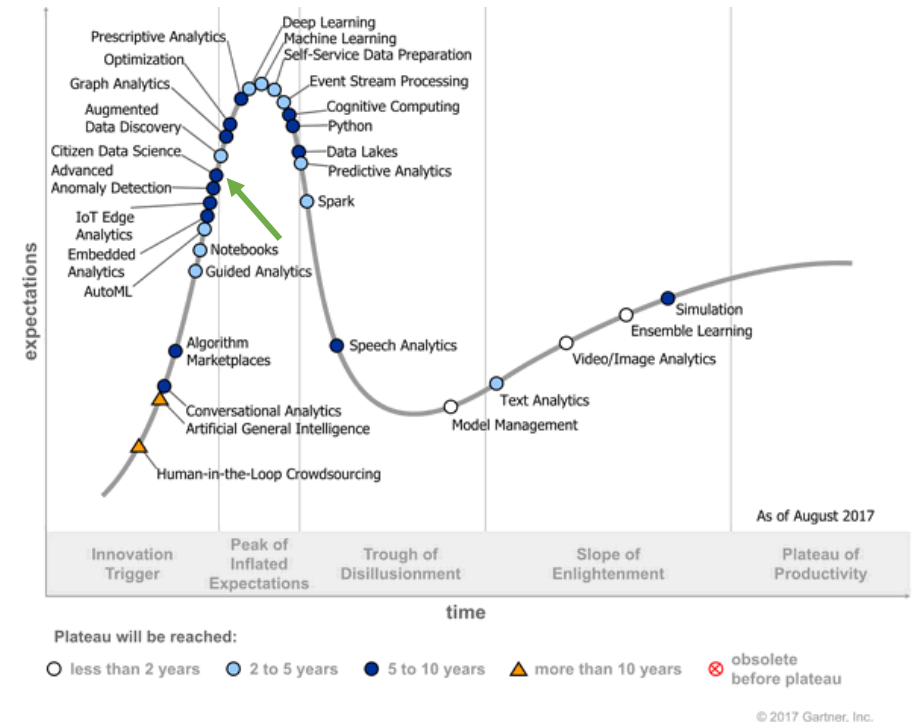


Exemplo – Mall Customer Segmentation Data



Data Science e Data Mining

- Os termos “ciência de dados” e “mineração de dados” costumam ser usados alternadamente;
- Ciência de dados ganhou notoriedade à medida que vários indivíduos e organizações tentam tirar proveito do “*hype*” em torno dela;
- A ciência de dados é um **conjunto de princípios fundamentais** que orientam a extração de conhecimento dos dados;
- A mineração de dados é a **extração de conhecimento dos dados**, por meio de tecnologias que incorporam esses princípios;



Dúvidas?



Bibliografia Básica

- FAWCETT, T.; PROVOST, F. **Data Science para Negócios: O que você precisa saber sobre mineração de dados e pensamento analítico de dados.** Alta Books Editora, 2018.
- MAYER-SCHÖNBERGER, V.; CUKIER, K. **Big data: A revolution that will transform how we live, work, and think.** Houghton Mifflin Harcourt, 2013.
- GRUS J. **Data science from scratch: first principles with python.** O'Reilly Media, 2019.
- CARVALHO, A. et al. **Inteligência Artificial: uma abordagem de aprendizado de máquina.** Rio de Janeiro: LTC. 2011.