

# Sistemas de Apoio às Decisões

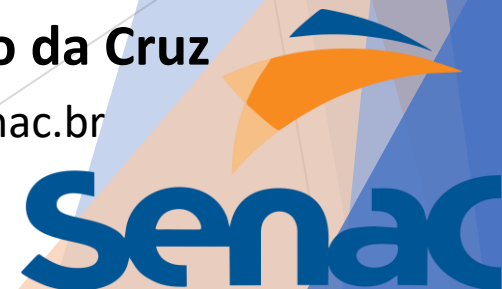
Aula 08 – Análise Preditiva de Dados

---

**Prof. Esp. Guilherme Jorge Aragão da Cruz**

 [guilherme.jacruz@sp.senac.br](mailto:guilherme.jacruz@sp.senac.br)

 [linkedin.com/in/guijac](https://linkedin.com/in/guijac)



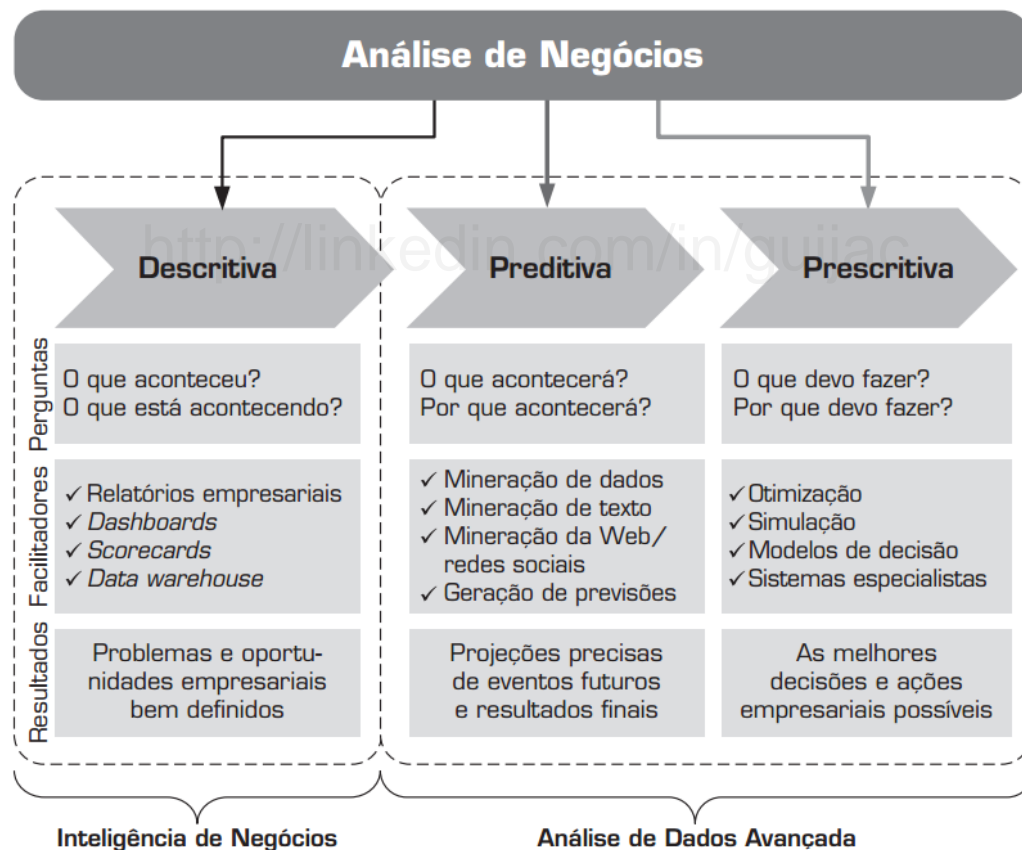
# Roteiro

---

- A Análise Preditiva;
- A Mineração de Dados;
- Funcionamento da Mineração de Dados;
  - Técnicas de Agrupamento;
  - Técnicas de Previsões;
- Aplicações da Mineração de Dados;
- Atividade Discente Orientada 04;
- Referências Bibliográficas.

# A Análise Preditiva

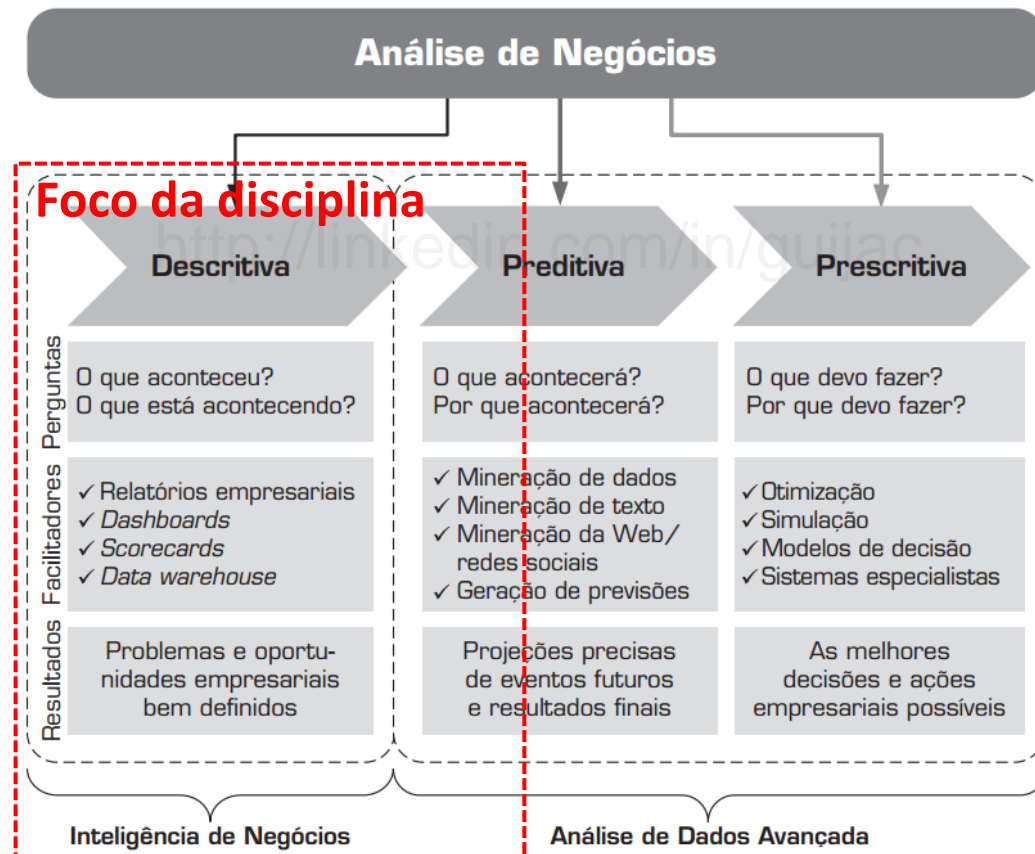
- Análise de dados preditiva (ou de **extração de relatórios**);
- Responde às perguntas: “O que vai acontecer” e “Por que isso vai acontecer?”.



Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

# A Análise Preditiva

- Análise de dados preditiva (ou de **extração de relatórios**);
- Responde às perguntas: “O que vai acontecer” e “Por que isso vai acontecer?”.



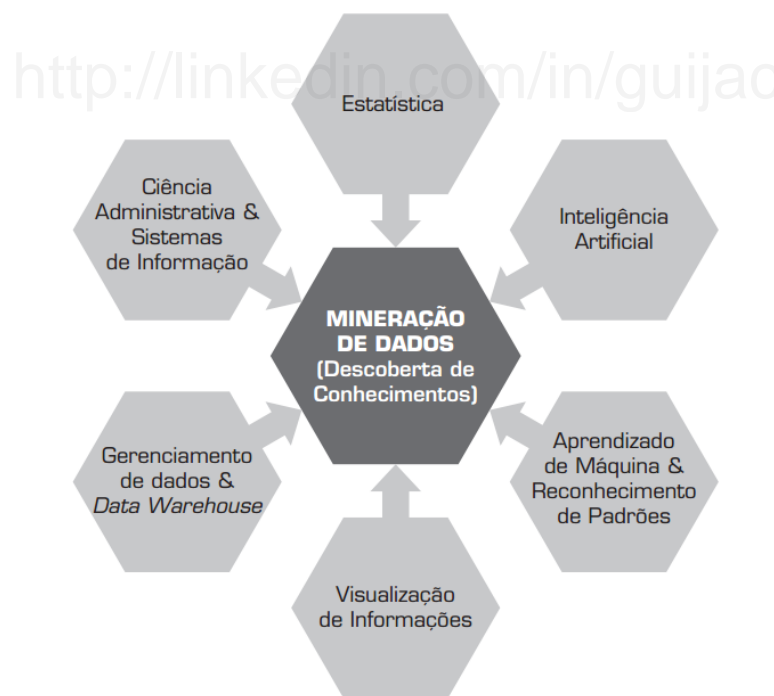
Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

# Mineração de Dados?

---

<http://linkedin.com/in/guijac>

# Mineração de Dados?



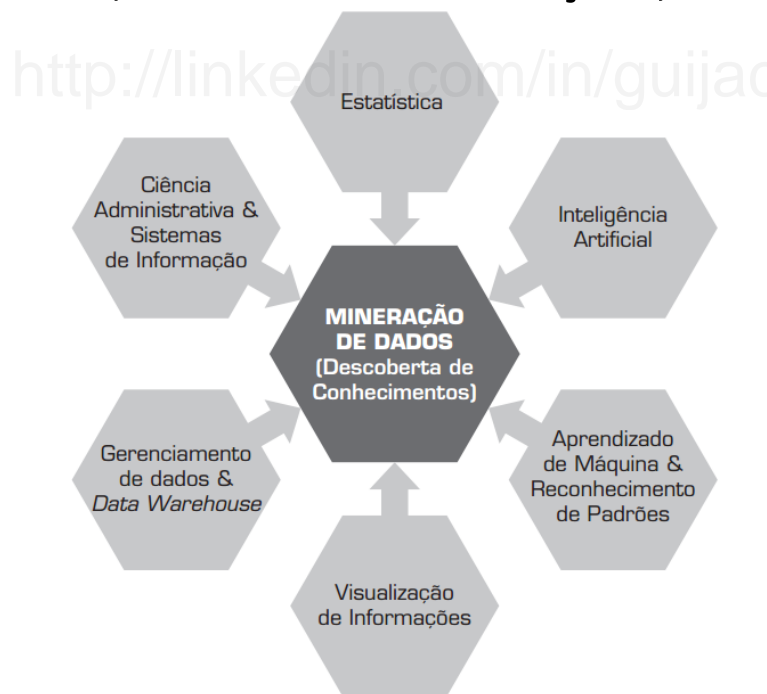
Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

# A Mineração de Dados

“ Processo não trivial de **identificar padrões** válidos, novos, potencialmente úteis e **ulteriormente compreensíveis** junto a bases de dados estruturados. ”

FAYYAD et. al (1998)

- Outros nomes: extração de conhecimentos, análise de padrões, arqueologia de dados, colheita de informações, busca de padrões.



Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

# Funcionamento da Mineração de Dados

- A mineração de dados constrói **modelos** para **revelar padrões** dentre os atributos apresentados no conjunto de dados.
  - Modelos são as **representações matemáticas** que identificam os padrões dentre os **atributos** dos aspectos descritos no **conjunto de dados**.

Tarefas & Métodos de Mineração de Dados	Algoritmos de Mineração de Dados	Tipo de Aprendizado
Previsão		
Classificação	Árvores de Decisão, Redes Neurais, Máquinas de Vetores de Suporte, kNN, Naive Bayes, GA	Supervisionado
Regressão	Regressão Linear/Não linear, ANN, Árvores de Regressão, SVM, kNN, GA	Supervisionado
Série temporal	Métodos Autorregressivos, Métodos de Extração de Médias, Suavização Exponencial, ARIMA	Supervisionado
Associação		
Cesta de mercado	Apriori, OneR, ZeroR, Eclat, GA	Não supervisionado
Análise de elos	Maximização de Expectativa, Algoritmo Apriori, Correspondência Baseada em Gráficos	Não supervisionado
Análise sequencial	Algoritmo Apriori, FP-Growth, Correspondência Baseada em Gráficos	Não supervisionado
Segmentação		
Agrupamento	k-means, Maximização de Expectativa (ME)	Não supervisionado
Análise de discrepâncias	k-means, Maximização de Expectativa (ME)	Não supervisionado

Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)



# Funcionamento da Mineração de Dados

- A mineração de dados constrói **modelos** para **revelar padrões** dentre os atributos apresentados no conjunto de dados.
  - Modelos são as **representações matemáticas** que identificam os padrões dentre os **atributos** dos aspectos descritos no **conjunto de dados**.

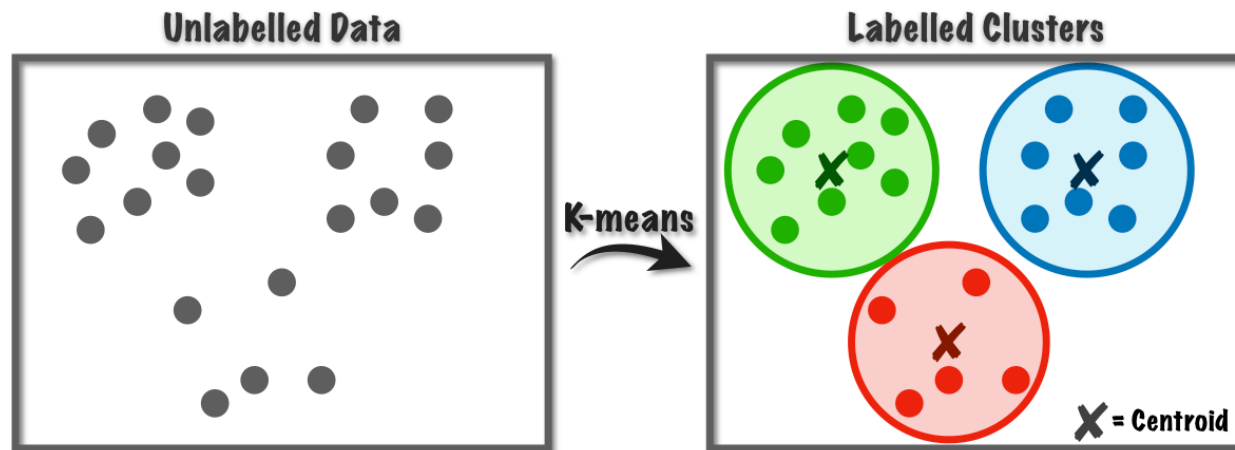
Tarefas & Métodos de Mineração de Dados	Algoritmos de Mineração de Dados	Tipo de Aprendizado
Previsão		
Classificação	Árvores de Decisão, Redes Neurais, Máquinas de Vetores de Suporte, kNN, Naive Bayes, GA	Supervisionado
Regressão	Regressão Linear/Não linear, ANN, Árvores de Regressão, SVM, kNN, GA	Supervisionado
Série temporal	Métodos Autorregressivos, Métodos de Extração de Médias, Suavização Exponencial, ARIMA	Supervisionado
Associação		
Cesta de mercado	Apriori, OneR, ZeroR, Eclat, GA	Não supervisionado
Análise de elos	Maximização de Expectativa, Algoritmo Apriori, Correspondência Baseada em Gráficos	Não supervisionado
Análise sequencial	Algoritmo Apriori, FP-Growth, Correspondência Baseada em Gráficos	Não supervisionado
Segmentação		
Agrupamento	k-means, Maximização de Expectativa (ME)	Não supervisionado
Análise de discrepâncias	k-means, Maximização de Expectativa (ME)	Não supervisionado

Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

# Técnicas de Agrupamento

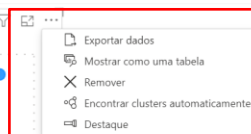
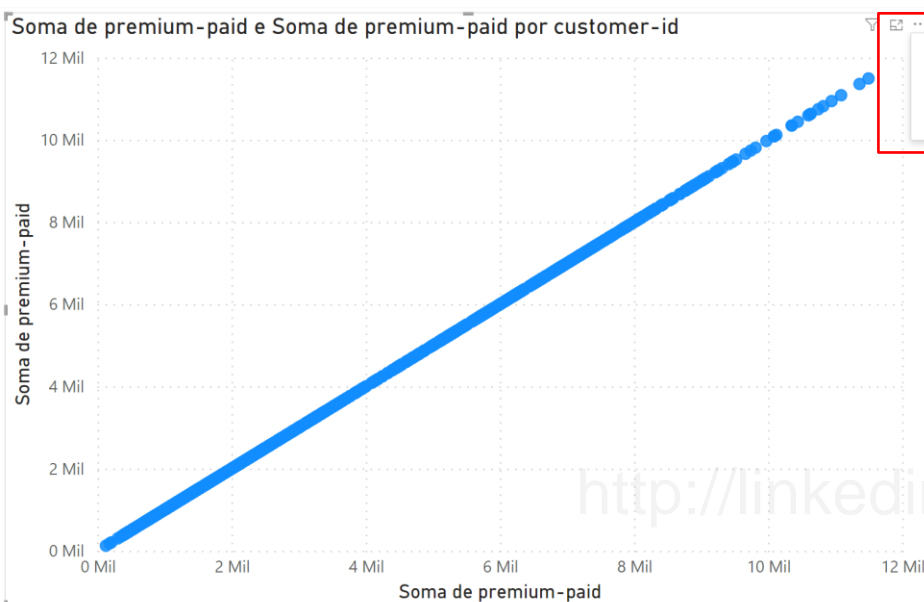
## ▪ Algoritmo de agrupamento por *k-means*

- $k$  = quantidade pré-determinada de agrupamentos.
- Algoritmo (**Etapas de inicialização**: escolha o valor de  $k$ )
- **Etapas 1**: Gere aleatoriamente  $k$  pontos como os centros iniciais dos agrupamentos.
- **Etapas 2**: Designe cada ponto ao centro de agrupamento mais próximo.
- **Etapas 3**: Recompute os novos centros de agrupamentos.
- **Etapas de repetição**: Repita as etapas 2 e 3 até que algum critério de convergência seja alcançado.



Fonte: [K-means: A Complete Introduction](#). K-means is an unsupervised clustering

# Técnicas de Agrupamento no Power BI



Clusters

Nome \*  
customer-id (clusters)

Campo  
customer-id

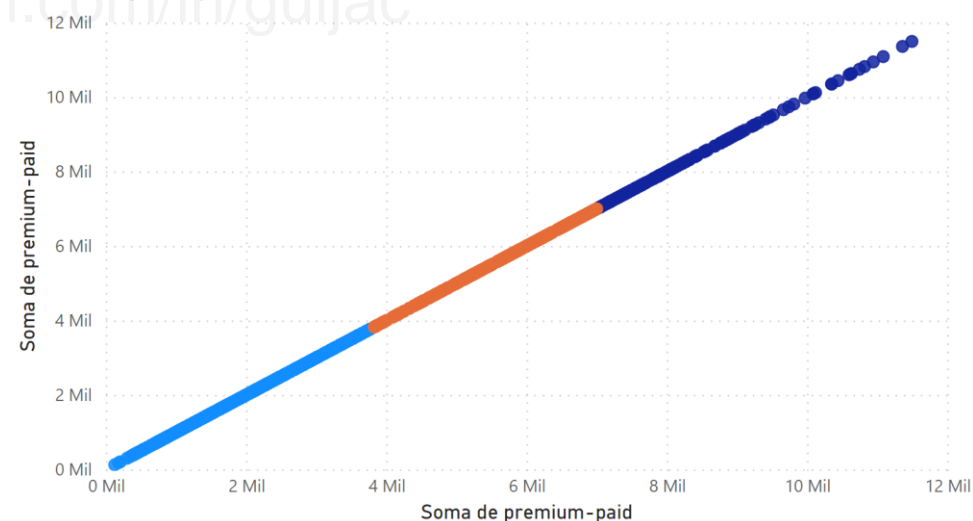
Descrição  
Clusters para customer-id

Número de clusters  
3

OK Cancelar

Soma de premium-paid e Soma de premium-paid por customer-id e customer-id (clusters) 8

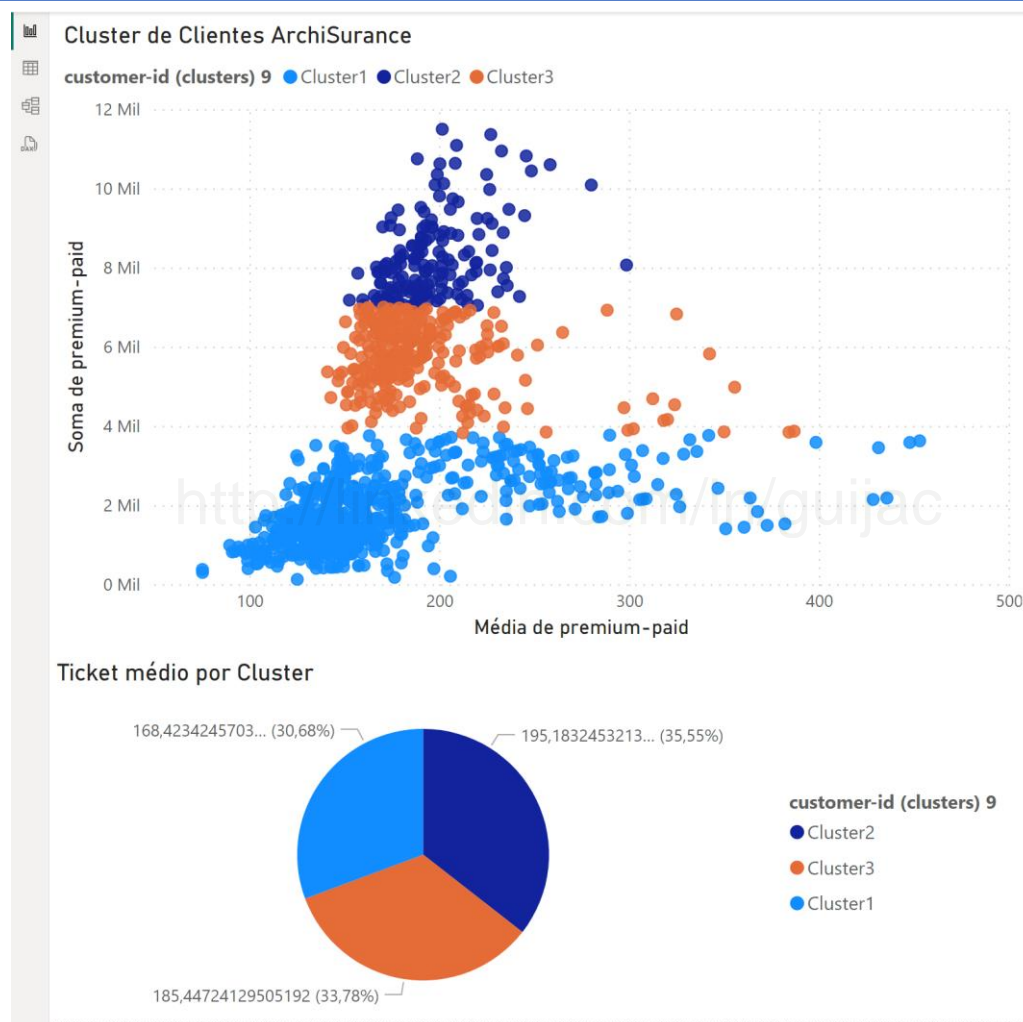
customer-id (clusters) 8 ● Cluster1 ● Cluster2 ● Cluster3



Exemplo de clusterização realizada no Power BI<sup>1</sup>, levando em consideração a soma do prêmio pago por segurado, ideal para segmentação de clientes por renda/gastos.

<sup>1</sup> [Algoritmo de clustering da Microsoft | Microsoft Learn](#)

# Técnicas de Agrupamento no Power BI



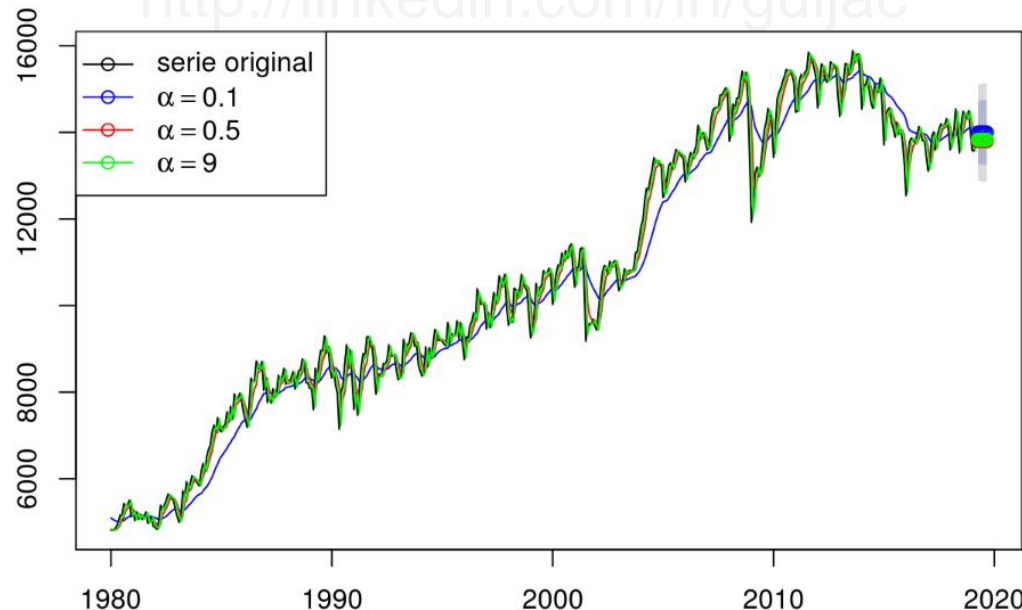
Exemplo de clusterização realizada no Power BI<sup>1</sup>, levando em consideração a soma do prêmio pago por segurado, ideal para segmentação de clientes por renda/gastos.

<sup>1</sup> [Algoritmo de clustering da Microsoft | Microsoft Learn](#)

# Técnicas de Previsões

## ▪ Suavização Exponencial Simples

- **Etapa 1:** Define a previsão anterior como o primeiro valor na série temporal.
- **Etapa 2:** Itera sobre cada valor na série temporal, começando pelo segundo valor.
- **Etapa 3:** Calcula a previsão atual utilizando a fórmula do SES, que é uma combinação ponderada do valor atual e da previsão anterior, com uma variável alfa de suavização.

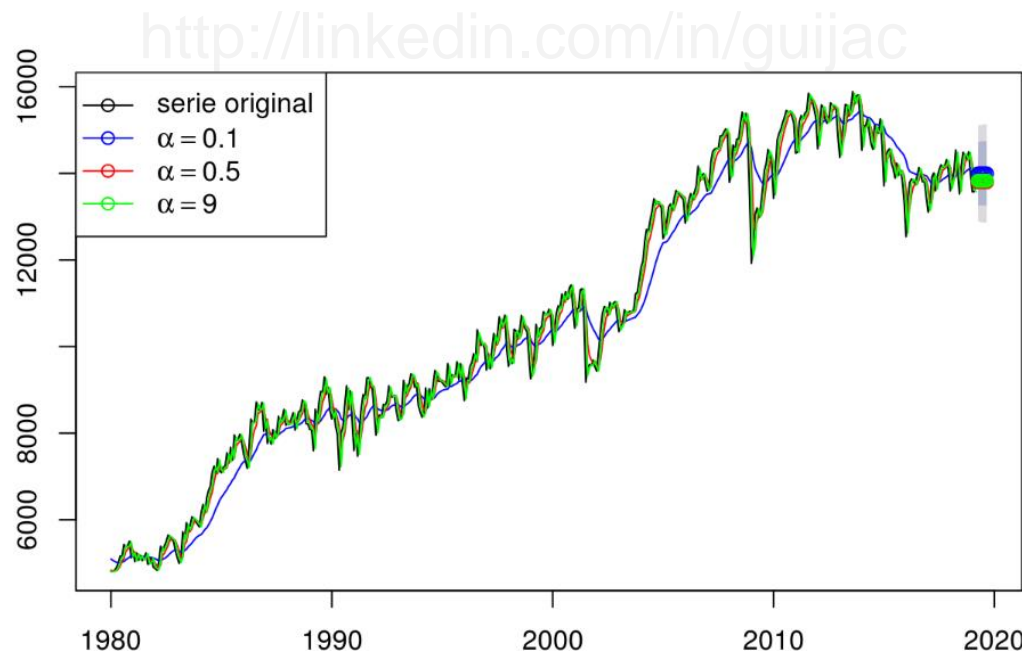


Fonte: [Capítulo 5 Suavização exponencial | Curso de Séries Temporais \(sillasgonzaga.com\)](#)

# Técnicas de Previsões

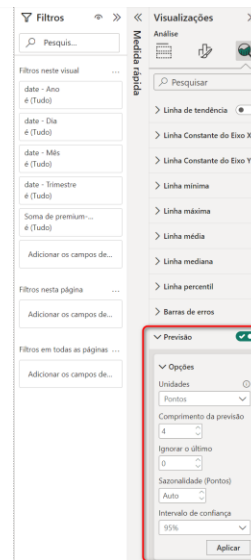
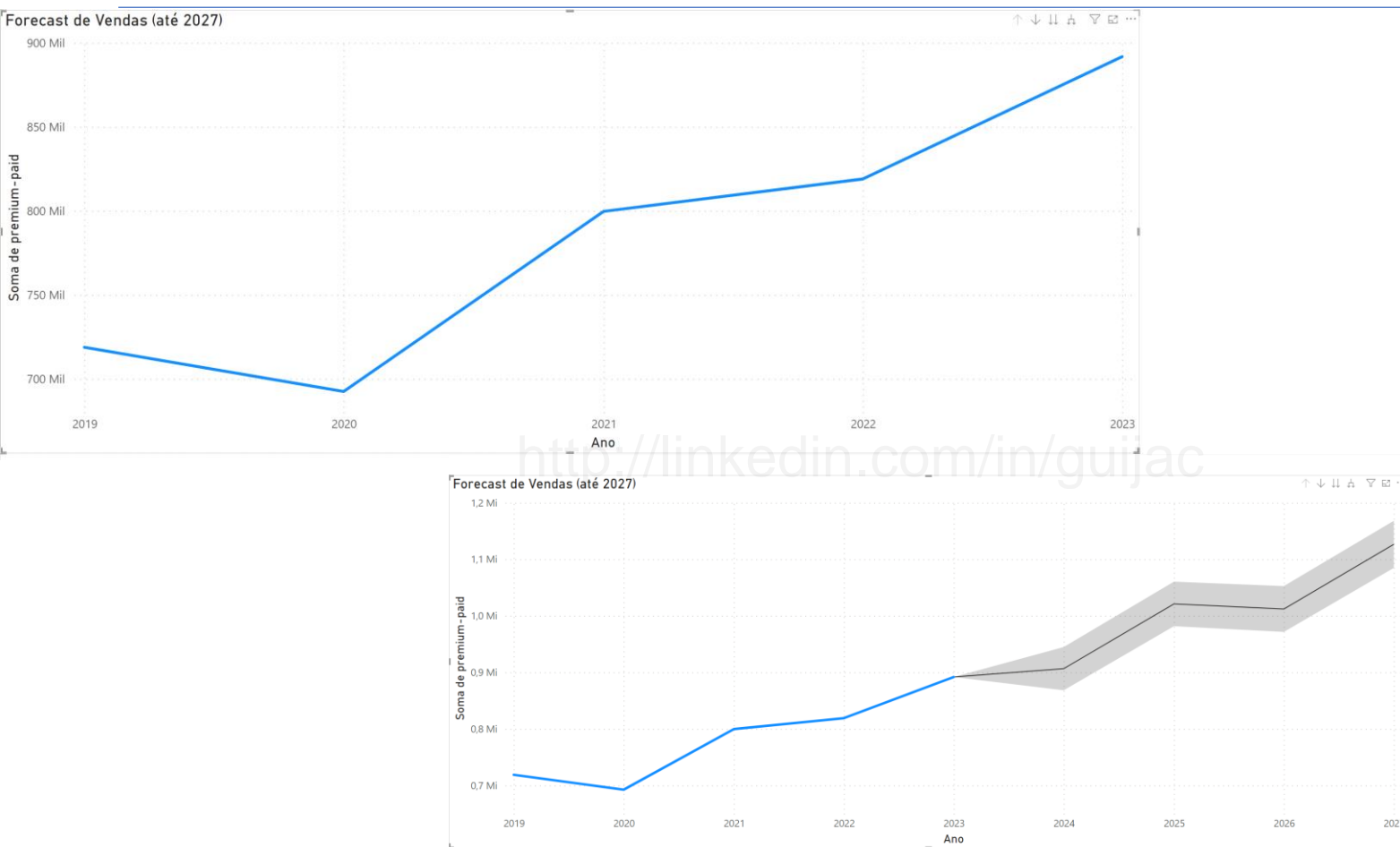
## ▪ Suavização Exponencial Simples

- Um **alfa menor** resulta em uma **suavização maior**, o que significa que o modelo é **mais lento** para reagir a **mudanças** nos dados;
- Um **alfa maior** torna o modelo mais **sensível às mudanças** recentes nos dados, resultando em menos suavização;
- É **mais adequada** para dados com **padrões estáveis** e **mudanças graduais** ao longo do tempo.



Fonte: [Capítulo 5 Suavização exponencial | Curso de Séries Temporais \(sillasgonzaga.com\)](#)

# Técnicas de Previsões no Power BI



Exemplo de previsão por série temporal realizada no Power BI<sup>1</sup>. A região sombreada mostra o intervalo de valores possíveis, ou seja, valores futuros devem estar dentro da área sombreada com uma probabilidade de 95%.

<sup>1</sup> [Algoritmo de Série Temporal da Microsoft | Microsoft Learn](https://learn.microsoft.com/pt-br/azure/machine-learning/tutorial-time-series-forecasting)

# Aplicações da Mineração de Dados

## ▪ Gestão de relacionamento com o cliente

- Identificar os respondentes/compradores mais propensos de novos produtos/serviços;
- Entender as causas-raiz da evasão (*churn*) de clientes a fim de promover sua retenção;
- Descobrir associações variáveis no tempo entre produtos e serviços a fim de maximizar as vendas e o valor dos clientes;
- Identificar os clientes mais lucrativos e suas necessidades preferenciais a fim de reforçar relacionamentos e maximizar vendas.



Fonte: [Valor do CRM: quanto custa a licença das principais ferramentas? \(ploomes.com\)](https://ploomes.com/pt-br/valor-do-crm-quantos-custos-tem-uma-licenca-de-crm/)



# Aplicações da Mineração de Dados

## ▪ Setor bancário

- Automatização do processo de pedido de empréstimos;
- Detecção de transações fraudulentas;
- Identificação de maneiras de maximizar o valor dos clientes ao vender-lhes produtos e serviços mais propensos a serem por eles adquiridos;
- Otimização do retorno monetário ao prever com precisão o fluxo de caixa em entidades bancárias.



Fonte: [Em constante evolução, setor bancário abraça a inovação; veja 5 tendências \(febraban.org.br\)](http://febraban.org.br)

# Aplicações da Mineração de Dados

## ▪ Seguros

- Prever quantias de custos envolvendo sinistros de propriedades e médicos para planejar melhor os negócios;
- Determinar planos de taxas ideais com base na análise de sinistros e em dados dos clientes;
- Prever quais clientes são mais propensos a comprar novas apólices com itens especiais;
- Identificar e prevenir pagamentos incorretos de sinistros e atividades fraudulentas.



Fonte: [Mercado Segurador Brasileiro – 1º Semestre de 2019 - Cresça com a Liberty \(crescacomaliberty.com.br\)](https://crescacomaliberty.com.br/)

# Atividade Discente Orientada 04

- **Em grupos**, a partir do Caso de Negócio da ArchiSurance **ou** através da base selecionada na ADO 01:
  - Realizar Análise Preditiva da Base de Dados.
- **Critérios de Aceite:**
  - Gráficos de Análise Preditiva da Base de Dados (Execução de Análise de Forecast/Previsão e/ou Agrupamento, 2.0 por gráfico, totalizando até 4.0);
  - Discussão de cada gráfico gerado (3.0 por discussão, totalizando até 6.0);
  - **OU**
    - Análise Preditiva da ArchiSurance (4.0);
    - Discussão de cada métrica obtida (2.0);
- Entregar arquivo PDF com a análise preditiva e discussão das medidas, contendo o nome dos integrantes até 13/04/2024 às 23h59;
- Quaisquer dúvidas, usem a aula de hoje!

# Por hoje é só!

Tarefas & Métodos de Mineração de Dados	Algoritmos de Mineração de Dados	Tipo de Aprendizado
<div>Previsão</div> <div>Classificação</div> <div>Regressão</div> <div>Série temporal</div>	<div>Árvore de Decisão, Redes Neurais, Máquinas de Vetores de Suporte, kNN, Naive Bayes, GA</div> <div>Regressão Linear/Não linear, ANN, Árvores de Regressão, SVM, kNN, GA</div> <div>Métodos Autorregressivos, Métodos de Extração de Médias, Suavização Exponencial, ARIMA</div>	<div>Supervisionado</div> <div>Supervisionado</div> <div>Supervisionado</div>
<div>Associação</div> <div>Cesta de mercado</div> <div>Análise de elos</div> <div>Análise sequencial</div>	<div>Apriori, OneR, ZeroR, Eclat, GA</div> <div>Maximização de Expectativa, Algoritmo Apriori, Correspondência Baseada em Gráficos</div> <div>Algoritmo Apriori, FP-Growth, Correspondência Baseada em Gráficos</div>	<div>Não supervisionado</div> <div>Não supervisionado</div> <div>Não supervisionado</div>
<div>Segmentação</div> <div>Agrupamento</div> <div>Análise de discrepâncias</div>	<div>k-means, Maximização de Expectativa (ME)</div> <div>k-means, Maximização de Expectativa (ME)</div>	<div>Não supervisionado</div> <div>Não supervisionado</div>

Fonte: SHARDA, R. ; DELEN, D. ; TURBAN (2019)

**Prof. Esp. Guilherme Jorge Aragão da Cruz**

[guilherme.jacruz@sp.senac.br](mailto:guilherme.jacruz@sp.senac.br)

[linkedin.com/in/guijac](https://www.linkedin.com/in/guijac)

# Referências Bibliográficas

- CENTRO DE INFORMÁTICA UFPE. Análise e Previsão de Séries Temporais. Disponível em [https://www.cin.ufpe.br/~psgm/Series%20Temporais/Aula\\_01.pdf](https://www.cin.ufpe.br/~psgm/Series%20Temporais/Aula_01.pdf). Acesso em 01 Abr 2024;
- MICROSOFT LEARN. **Referência técnica do algoritmo MTS**. Disponível em <https://learn.microsoft.com/pt-br/analysis-services/data-mining/microsoft-time-series-algorithm-technical-reference?view=asallproducts-allversions>. Acesso em 01 Abr 2024;
- QLIK. **Previsão de séries temporais**. Disponível em [https://help.qlik.com/pt-BR/cloud-services/Subsystems/Hub/Content/Sense\\_Hub/Visualizations/LineChart/timeseries-forecast.htm](https://help.qlik.com/pt-BR/cloud-services/Subsystems/Hub/Content/Sense_Hub/Visualizations/LineChart/timeseries-forecast.htm). Acesso em 01 Abr 2024;
- SHARDA, R. ; DELEN, D. ; TURBAN, E. **Business intelligence e análise de dados para gestão do negócio**. 4. ed. Porto Alegre: Bookman, 2019;
- SILAS GONZAGA. **Capítulo 5 Suavização exponencia**. Disponível em [http://sillasgonzaga.com/material/curso\\_series\\_temporais/suaviza%C3%A7%C3%A3o](http://sillasgonzaga.com/material/curso_series_temporais/suaviza%C3%A7%C3%A3o). Acesso em 01 Abr 2024.