

Tópicos Avançados de Automação Industrial

TAAE9

Profº José W. R. Pereira

Ciência de dados

Análise de Dados

A análise de dados é o **processo de manipulação de dados** através de **ferramentas computacionais e estatísticas**, de modo a buscar **informações relevantes** que auxiliam à **tomada de decisão**.

Tipos de análise de dados

Analista de Dados

Cientista de Dados

Descritiva

Diagnóstica

Preditiva

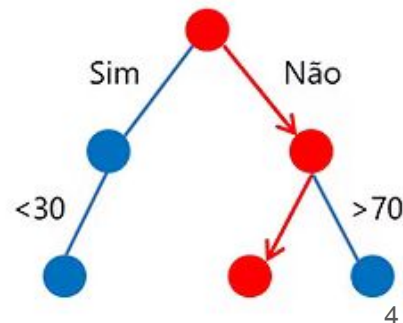
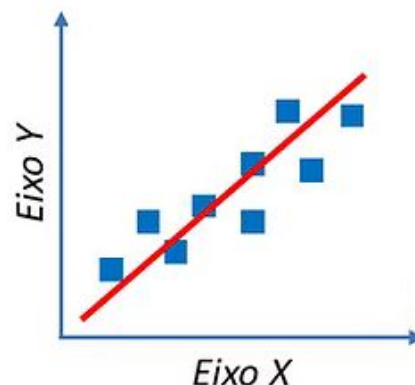
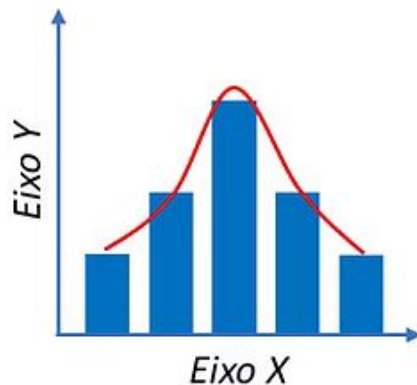
Prescritiva

Quem?

Quando?

Onde?

Por quê?



Desenvolvimento da Análise



1. Problema de negócio:

1. O resultado da ação foi positivo ou negativo?
2. Por que os resultados foram esses?
3. Como tomar uma decisão?

Desenvolvimento da Análise

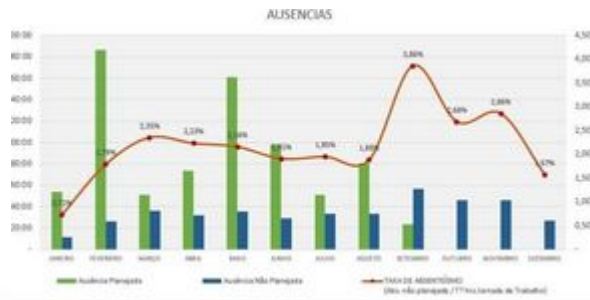


2. **Descrever** o comportamento dentro de um escopo;
3. **Diagnosticar** os motivos que produziram os comportamentos;
4. **Prever** o comportamento baseado em histórico;
5. **Prescrever/orientar** ações para alcançar as metas do negócio.

Análise Descritiva

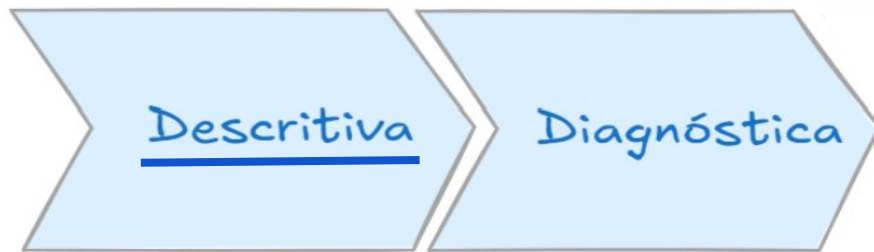


Fornece um resumo simples de uma planilha de dados, através de **indicadores, gráficos e tabelas.**

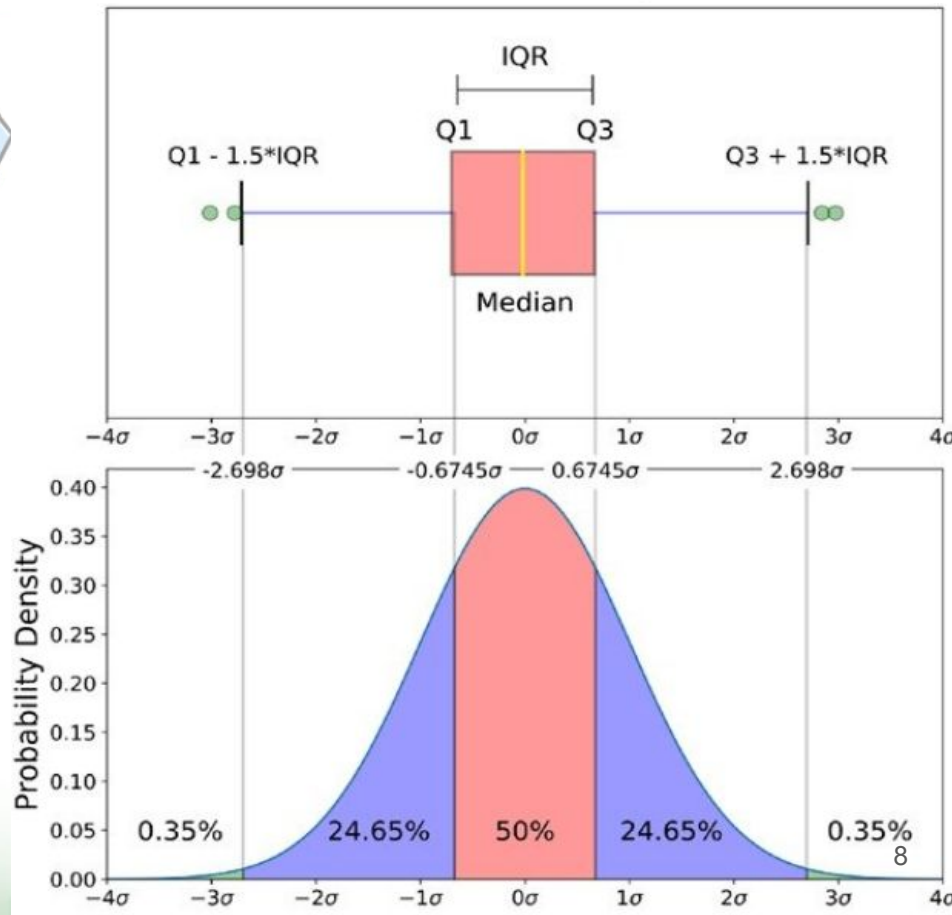


Cia Aérea	Horário
Latam	11:10
Gol	11:15
Azul	11:20
Gol	12:40
Latam	12:45

Análise Descritiva



Realiza a exploração inicial dos dados, permitindo a compreensão da **distribuição, valor central e dispersão dos dados**, além da presença de possíveis **outliers**.



Características

Descritiva

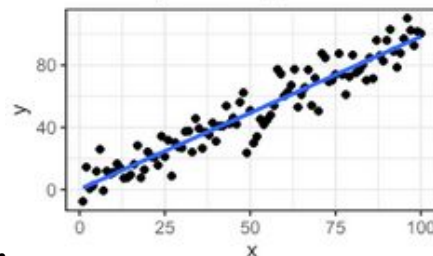
Diagnóstica

Preditiva

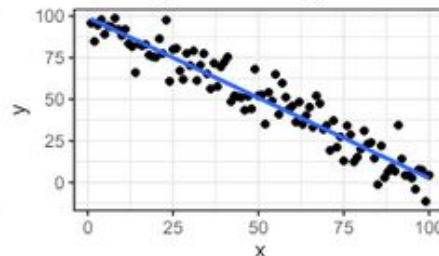
Prescritiva

- Resumo de dados:
 - Medidas de tendência central:
 - Média, Moda e Mediana;
 - Medidas de dispersão:
 - Desvio padrão e Variância.
- Visualização de dados:
 - Relação entre variáveis.
- Identificação de padrões.

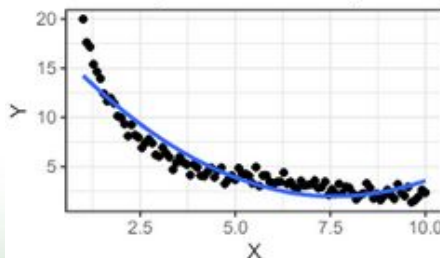
Correlação linear positiva



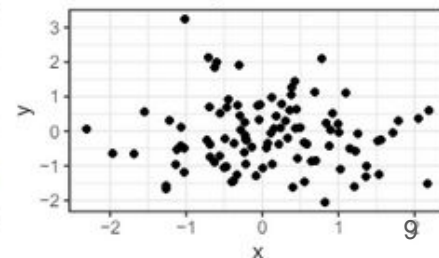
Correlação linear negativa



Correlação não linear negativa



Sem correlação



Ferramentas



1. Planilha de cálculos: Excel, Google Sheets;
2. Softwares estatísticos: SPSS, Stata;
3. Software de visualização de dados: Tableau, Power BI
4. Linguagem de programação: Python, R



Google Sheets

STATA



Análise Diagnóstica



Busca:

- **entender os motivos ou causas** por trás de determinado fenômeno ou resultado.
- **identificar padrões e relações** nos dados que podem explicar mudanças ou tendências observadas.



Características



- Investigativa:
 - Buscando entender as causas subjacentes de um fenômeno ou problema.
- Baseada em Hipóteses:
 - Uma hipótese é proposta e a análise é usada para testar ela.



Características



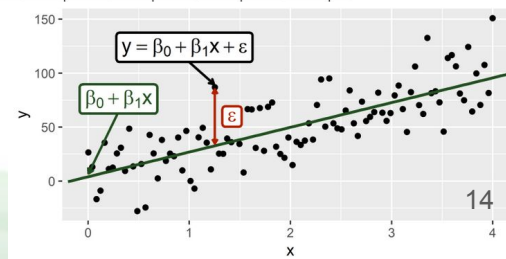
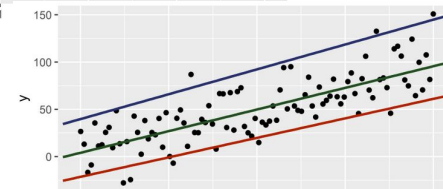
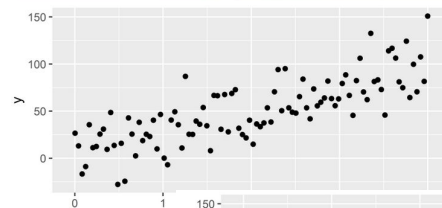
- Uso de dados históricos:
 - Identificar tendências, padrões e relações que podem explicar os resultados atuais.
- Análise Multivariada:
 - Pode envolver múltiplas variáveis ao mesmo tempo, com interações complexas que podem estar influenciando um resultado.
- Interpretação cuidadosa.



Ferramentas



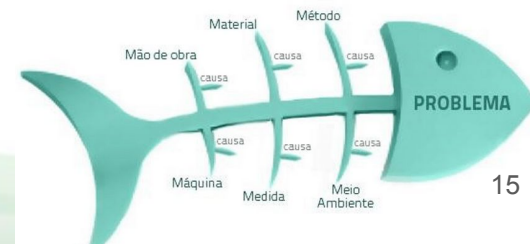
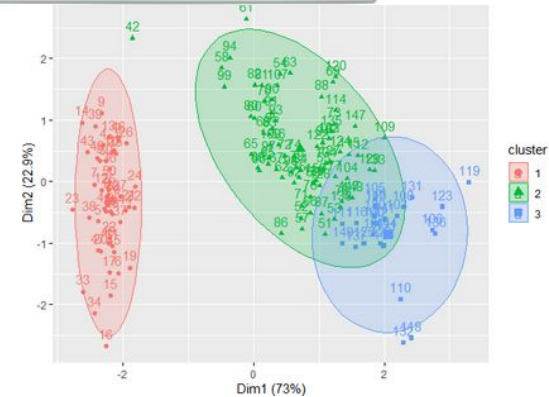
- Análise de regressão:
 - Identificar a relação entre variáveis dependentes e independentes;
- Análise de correlação:
 - Determinar a relação entre duas ou mais variáveis;



Ferramentas



- Análise de *cluster*:
 - Agrupa por características semelhantes;
- Análise de causa e efeito (Diagrama de Ishikawa):
 - Proporciona uma visão ampla e identificação de possíveis causas de um problema.



Análise Preditiva



- Análise estatística que tem como objetivo prever resultados futuros com base em dados históricos e técnicas de modelagem.
- Coleta de dados >> modelo estatístico >> Previsão

Características



1. Baseada em modelos: Estatísticos e de **Machine Learning**.
2. Usa histórico de dados: Treinamento dos modelos.
3. Multivariável:
Consideram a interação entre muitas variáveis.
4. Preditiva, mas não infalível:
Sempre há um grau de incerteza.
5. Orientada para a ação:
Resultados orientam a tomada de decisão.

Ferramentas



1. Python: Pandas, NumPy, Sci-kit Learn.
2. R: Análise estatística e gráficos.
3. SAS: *Statistical Analysis System* (financeiro e saúde).
4. SPSS: Software de fácil usabilidade e com ampla gama de ferramentas disponíveis.
5. Excel: Aplicado quando não se requer técnicas estatísticas avançadas.

Análise Prescritiva



- Fornece recomendações sobre o que deve ser feito, utilizando **técnicas avançadas** que levam em consideração uma **variedade de cenários** e **possíveis resultados**, para ajudar os **tomadores de decisão** a entender as implicações de diferentes cursos de ação.

Características



1. Recomendações de ação;
2. Consideração de diferentes cenários;
3. Otimização de operações;
4. Auxílio à tomada de decisão;
5. Adaptação a mudanças nas condições: pode incorporar novos dados à medida que se tornem disponíveis.

Ferramentas



1. Softwares de modelagem: Python, R, SAS, SPSS
2. Ferramentas de otimização: Gurobi, CPLEX
3. Ferramentas de simulação: Simul8, AnyLogic
4. Ferramentas de visualização de dados: Tableau, Power BI
5. Plataformas de machine learning: Azure ML, Amazon SageMaker

Exemplo

Forecasting Future Passengers

using Auto ML (PyCaret)

Tópicos Avançados de Automação Industrial

Prof^o José W. R. Pereira

jose.pereira@ifsp.edu.br

josewrpereira.github.io/docs