

Roteiro Aula Prática

PROBABILIDADE E ESTATÍSTICA PARA ANÁLISE DE DADOS

ROTEIRO DE AULA PRÁTICA

NOME DA DISCIPLINA: PROBABILIDADE E ESTATÍSTICA PARA ANÁLISE DE DADOS

Unidade: U1_Fundamentos de Probabilidade e Estatística para Análise de Dados

Aula: A4_Ferramentas para Análise de Dados

OBJETIVOS

Definição dos objetivos da aula prática:

Apresentar ferramentas e bibliotecas estatísticas disponíveis no R para análise de dados.

Demonstrar como manipular dados e implementar cálculos probabilísticos em R.

Capacitar os alunos para realizar análises básicas de conjuntos de dados estruturados.

SOLUÇÃO DIGITAL: RStudio ou Google Colab

- Se for um software será necessário a inserção do link no roteiro:

LINK SOLUÇÃO DIGITAL (EXCETO ALGETEC): <https://www.rstudio.com/> ou
<https://colab.google/>

PROCEDIMENTOS PRÁTICOS E APLICAÇÕES

Procedimento/Atividade nº 1

Inserir o nome do experimento: *Manipulação e análise de dados com R*

Atividade proposta: Utilizar bibliotecas estatísticas no R para realizar manipulação de dados e cálculos probabilísticos.

Procedimentos para a realização da atividade:

- Acessar a plataforma RStudio instalada no computador ou acessar a versão online.
- Baixar os dados <https://github.com/AndersonSalata/projeto-integrado-ciencia-de-dados/blob/main/datatran2024.csv>
- Carregar o conjunto de dados em formato .csv (como o exemplo fornecido).
Exemplo de comando: `dados <- read.csv("datatran2024.csv", sep = ";", fill = TRUE, check.names = FALSE)`
- Explorar o conjunto de dados para identificar variáveis e valores principais:
Exemplo: `summary(dados)` e `str(dados)`.
- Utilizar a biblioteca dplyr para realizar operações básicas de manipulação de dados (ex.: filtragem, agrupamento e sumarização).
Exemplo: `library(dplyr)`
- Aplicar conceitos probabilísticos para responder questões sobre os dados.
Exemplo: Calcular a probabilidade de um acidente ocorrer em uma determinada hora do dia.
- Visualizar os resultados usando a biblioteca ggplot2

Avaliando os resultados:

1. Introdução

Contextualização sobre o objetivo da análise.

Breve descrição do conjunto de dados utilizado (ex.: número de observações, variáveis principais).

2. Métodos

Explicação das etapas realizadas para manipulação e análise dos dados, incluindo:

Ferramentas utilizadas (R, pacotes como dplyr e ggplot2).

Breve descrição dos comandos aplicados.

3. Resultados e responder as seguintes questões.

Qual foi o estado com o maior número de acidentes?

Qual a probabilidade de um acidente ocorrer em condições climáticas claras?

Como a fase do dia afeta a ocorrência de acidentes?

Que insights podem ser gerados sobre os tipos de acidentes predominantes e suas causas?

4. Conclusão

Resumo dos principais insights obtidos com a análise.

Considerações sobre a aplicação de conceitos estatísticos na análise de dados reais.

5. Anexos

Código completo em R utilizado para a análise.

Capturas de tela ou links para gráficos gerados.

Checklist:

- ✓ Baixar e acessar a plataforma RStudio.
- ✓ Carregar o conjunto de dados no R.
- ✓ Realizar manipulações básicas com o conjunto de dados.
- ✓ Explorar os dados com funções do R e bibliotecas.
- ✓ Aplicar cálculos probabilísticos no conjunto de dados.
- ✓ Visualizar resultados com gráficos.
- ✓ Responder às perguntas com base nos resultados.

RESULTADOS

Resultados do experimento:

Ao final dessa aula prática, você deverá enviar um arquivo em word contendo as informações obtidas no experimento, os cálculos realizados, em conjunto com um texto conclusivo a respeito das informações obtidas. O arquivo não pode exceder o tamanho de 2Mb.

- Referências bibliográficas ABNT (quando houver).

Resultados de Aprendizagem:

Ao final da aula prática, o aluno será capaz de:

- Carregar e explorar conjuntos de dados em R.
- Utilizar bibliotecas estatísticas como dplyr e ggplot2 para manipulação e visualização.
- Aplicar conceitos de probabilidade em dados reais.
- Interpretar insights baseados na análise exploratória e probabilística de dados.