



CENTRO UNIVERSITÁRIO RITTER DOS REIS

PÓS-GRADUAÇÃO EM DATA SCIENCE E BIG DATA

DISCIPLINA: *Análise de dados com R*

CARGA HORÁRIA: 36

PROFESSOR: Renan Xavier Cortes

ANO/SEM: 2018/1

Horário das Aulas: *Segundas e Quartas (19h - 22h 25min)*

PROJETO DA DISCIPLINA – Valor de avaliação: 70%

Descrição do Projeto

O projeto final a ser entregue da disciplina deve compreender uma análise de dados completa de uma base de dados real de livre escolha do aluno. A análise de dados a ser conduzida poderá englobar todos os aspectos de uma análise de dados em todas as suas etapas: importação da base para o R, limpeza de dados, organização dos dados, exploração descritiva, pergunta de análise, procedimento inferencial e conclusão.

É importante que o trabalho apresente a explicação das variáveis que estão presentes no banco que serão utilizadas ao longo das análises.

Diretrizes do Trabalho

- Trabalho individual ou em duplas.
- Deverá ser conduzida uma análise descritiva da base de dados (estatísticas de resumo, medidas de tendência central, de dispersão, gráficos, etc.).
- Recomenda-se o uso do ggplot2 para realização dos gráficos da análise descritiva.
- Deverá ser construído ao menos um gráfico interativo usando o htmlwidgets (pode ser com o plotly, leaflet, highcharts, etc.).
- Deverá ser feita, no mínimo, uma pergunta de análise (hipótese de pesquisa).
- Deverá ser feito um procedimento inferencial para responder à pergunta de análise. O procedimento pode ser tanto ser via testes de hipóteses ou via regressão (linear simples, polinomial, múltipla ou GLM) ou ambos.
- Códigos inteiros no R.
- Deve ser entregue o código, a base de dados (ou bases de dados) e, se houverem, documentos suplementares.
- Código reproduzível.
- O trabalho também pode ser entregue em RMarkdown se o (a) aluno (a) ou a dupla preferir.