

# MINERAÇÃO DE DADOS COMPLEXOS

## Curso de Extensão



#### TRABALHO FINAL (EM DUPLA OU EM TRIO) INF-0612 – Análise de Dados

#### Descrição

O objetivo deste trabalho é exercitar os conceitos apresentados na disciplina. Você deve baixar os dados climatológicos da cidade de Campinas, disponíveis em https://www.ic.unicamp.br/~zanoni/cepagri/cepagri.csv, e fazer um conjunto de análises, gerando gráficos e tabelas. Você deve analisar apenas os dados compreendidos no intervalo de 01/01/2015 a 31/12/2021. Como resultado, você deverá criar um relatório, descrevendo e comentando as análises feitas, e um código R, com todos os comandos necessários para obter os resultados descritos em seu relatório.

As colunas do conjunto de dados contém as seguintes informações, nesta ordem: uma string informando o dia e hora em que o dado foi coletado; a temperatura (em  $^{\circ}$ C); a velocidade do vento (em km/h); a umidade (em porcentagem); a sensação térmica (em  $^{\circ}$ C).

#### Requisitos

Seu relatório deve ser submetido em formato PDF e deve conter:

- Tratamento dos dados: descrição de como foi realizado o tratamento inicial dos dados para análise.
- Análises: pelo menos quatro análises distintas, sendo pelo menos uma delas comparando os dados de pelo menos 4 anos diferentes.
- Tabelas: pelo menos quatro e no máximo oito tabelas.
- Gráficos: pelo menos quatro e no máximo oito gráficos. Além disso, devem existir pelo menos três tipos distintos de gráficos, e todos devem ser criados com o pacote ggplot2.
- Páginas: pelo menos oito e no máximo doze páginas (formato livre).

### Considerações Finais

- $\bullet\,$ Suas tabelas e gráficos deverão conter legendas, e os rótulos devem ser representativos.
- Devem ser entregues dois arquivos: o relatório em formato PDF, explicando as análises realizadas bem como
  as informações obtidas (gráficos e tabelas), e um arquivo R, com a implementação para cada uma delas. Você
  não deve compactar seus arquivos, envie-os no Moodle de forma separada (são aceitas submissões contendo 2
  arquivos para esta tarefa).
- Você não será penalizado caso utilize funções ou estruturas não vistas em sala de aula.
- Seu código R deve estar organizado e minimamente comentado, explicando o que cada função ou bloco de comandos está fazendo. Códigos mal documentados ou com erros de execução sofrerão desconto de nota.
- Apenas um membro do grupo deve enviar o trabalho final. Os nomes dos membros devem constar tanto na capa do relatório PDF quanto no cabeçalho do arquivo R.
- Envie os dois arquivos no sistema Moodle, clicando no link "Trabalho Final" da Seção "Avaliações". Clique em "Adicionar tarefa", anexe os arquivos e, por fim, clique em "Salvar mudanças". Você voltará para a tela da atividade e deverá constar o status "Enviado para avaliação". A qualquer momento, antes do prazo final de submissão, você pode alterar sua submissão clicando em "Editar envio".

Prazo de entrega: 28 de Agosto de 2022 (Domingo) às 23h55.

Forma de entrega: via sistema Moodle:

• https://moodle.lab.ic.unicamp.br/moodle/course/view.php?id=502

**Pontuação:** Este relatório será pontuado de 0 a 10, e corresponderá 40% da nota final.