



UNIVERSIDADE FEDERAL DE SERGIPE  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS

Aluno (a): \_\_\_\_\_ Data: \_\_\_\_\_  
Disciplina: \_\_\_\_\_ Prof.: Esdras Santos

*Atividade Unidade I*

## 1 A Atividade

Esta atividade prática visa apresentar ao situações reais no cotidiano de sua profissão e com consolidar o conteúdo teórico apresentado em sala de aula. Neste o discente aplicará técnicas estatístico-computacionais na resolução de um problema e a escrita de um relatório explicando os métodos e cálculos realizados.

Os alunos deve escolher um banco de dados e realizar uma análise estatística usando o que foi exposto em sala de aula. Uma análise não é apenas o resumir dados ou cálculo de medidas estatística. É uma, investigação na busca de conhecer o comportamento da pesquisa que gerou os dados. Assim, o inquérito deve ser desenvolvido de forma detalhada escrevendo todos os detalhes dos cálculos e explicando seus resultados. O aluno deve tomar com base os modelos disponíveis no Google Sala de aula. Os bancos de dados devem ser baixados da internet. A seguir, temos alguns sites de onde podem ser baixados os dados para a realização da atividade:

### Objetivo

O objetivo deste trabalho é avaliar o banco de dados da Escolha do aluno usando os conhecimentos adquiridos em sala de aula. As peças do trabalho são:

1. Construir a Distribuição de Frequência de pelo menos uma Variável Qualitativa e uma Variável Quantitativa.
  2. Calcular as Medidas descritivas de Posição:
    - Média;
    - Mediana;
    - Moda.
    - Calcular as Medidas Descritivas de Dispersão:
    - Variância;
    - Desvio Padrão;
    - Coeficiente de Variação.
  3. Como item adicional construir gráficos
1. Mais de 70 sites para obter grandes repositórios de dados gratuitamente: <https://www.crayondata.com/70-websites-to-get-large-data-repositories-for-free/>
  2. Dados do livro Noções de Probabilidade e Estatística (MAGALHÃES; LIMA, 2010), a saber <<https://www.ime.usp.br/~noproest/doku.php?id=dados>>.

3. UCI Machine Learning Repository (DUA; GRAFF, 2017).
4. A Kaggle é uma comunidade on-line de cientistas de dados e aprendizes de máquinas, de propriedade da Google LLC. (GOLDBLOOM; HAMNER, 2010).
5. Site onde o autor apresenta vários repositórios de onde os bancos de dados podem ser coletados (PARUCHURI, 2020).

As questões desenvolvidas devem ser expostas em um relatório com o máximo de detalhes possível

## 1.1 O Relatório

O relatório deve conter a seguintes partes

1. **Introdução:** esta parte deve apresentar uma visão geral sobre os temas sob estudo e a relevância do trabalho.
2. **Objetivo:** esta parte deve informar ao leitor de forma concisa o(s) objetivo(s) do trabalho.
3. **Materiais e Métodos:** esta parte apresenta ao leitor uma explicação das peças do trabalho. Estas peças são a informação das técnicas usadas, o ou os Softwares usados no calculo dos resultados
4. **Resultados:** Este deve apresentar umas lista com o máximo de detalhes dos cálculos e seus resultados.
5. **Conclusões:** esta parte deve apresentar os detalhes da trajetória do trabalho de forma geral.
6. **Relato:** o aluno deve relatar o que achou do trabalho.
7. **Referência:** esta parte exhibe as referências das citações expostas ao longo do trabalho e devem seguir as normas da ABNT.

## 1.2 A data de entrega do trabalho

A data de entrega desta atividade é a data limite que esta determinada no SIGAA.

## 1.3 A notas do trabalho

A condição para que o aluno receba a nota é a entrega relatório via SIGAA na data determinada. As notas serão divulgadas após a correção que terá inicio no dia posterior a entrega da atividade.

## Referências Bibliográficas

DUA, D.; GRAFF, C. *UCI Machine Learning Repository*. 2017. Disponível em: <<http://archive.ics.uci.edu/ml>>. 2

GOLDBLOOM, A.; HAMNER, B. *Kaggle*. 2010. Disponível em: <<https://www.kaggle.com/>>. 2

MAGALHÃES, M. N.; LIMA, A. C. P. de. *Noções de Probabilidade e Estatística*. 7. ed. São Paulo: EDUSP, 2010. 1

PARUCHURI, V. *21 Places to Find Free Datasets for Data Science Projects*. 2020. Disponível em: <<https://www.dataquest.io/blog/free-datasets-for-projects/>>. 2