

Global Solution – 2º Semestre

Tecnólogo em Inteligência Artificial

Disciplina: Statistical Computing with R

Professor: Me. Eng. Rodolfo M. de Paiva

Nome _____ Turma _____ Data _____

Tema: Soluções tecnológicas para a prevenção de problemas na saúde e melhoria de qualidade de vida, alinhado com a ODS 3 – Saúde e Bem-estar.

Criar ou reunir, no mínimo, uma base de dados relacionada ao tema, composta por pelo menos 5 variáveis e 30 observações. Dentre essas variáveis, é imprescindível que, ao menos, 4 sejam quantitativas. Na sequência, resolver as questões de 01) à 04).

01) **(2,0 pontos)** Assumindo que as variáveis numéricas escolhidas seguem uma Distribuição Normal de Probabilidade, escolher uma variável aleatória da base de dados e determinar:

- a) A probabilidade desta variável ser menor do que a média dos registros da variável. Na sequência classifique o evento quanto ao resultado probabilístico.
- b) A probabilidade desta variável ser maior que a mediana dos registros da variável. Na sequência classifique o evento quanto ao resultado probabilístico.
- c) A probabilidade desta variável estar entre a média menos três vezes o desvio padrão, e a média mais três vezes o desvio padrão. Na sequência classifique o evento quanto ao resultado probabilístico.

02) **(2,0 pontos)** Assumindo que as variáveis numéricas escolhidas seguem uma Distribuição Normal de Probabilidade, escolher uma variável aleatória da base de dados e realizar um teste de hipótese para a média utilizando como dados do problema:

- a) A média amostral é menor do que a média populacional de referência, a um nível de significância de 6%, amostra de tamanho 20 e desvio padrão de referência igual ao desvio padrão populacional dos dados.
- b) A média amostral é maior do que a média populacional de referência, a um nível de significância de 9%, amostra de tamanho 25 e desvio padrão de referência igual ao desvio padrão populacional dos dados.
- c) A média amostral é diferente do que a média populacional de referência, a um nível de significância de 2%, amostra de tamanho 15 e desvio padrão de referência igual ao desvio padrão populacional dos dados.

03) **(2,0 pontos)** Faça uma Análise Bivariada entre 02 variáveis quantitativas contendo:

- a) O Gráfico de Dispersão, com título e nomes nos eixos x e y, e interprete o resultado.
- b) A Covariância e a Correlação Linear de Pearson, e interprete os resultados.
- c) O Gráfico de Correlação Linear de Pearson, e interprete o resultado.
- d) A reta da Regressão Linear Simples (no Gráfico de Dispersão), e a Equação da Regressão Linear Simples.
- e) A verificação da adequação do modelo (R^2 e Teste F), e interprete os resultados.
- f) Faça uma previsão.



04) **(2,0 pontos)** Faça uma Análise Multivariada entre, pelo menos, 03 variáveis quantitativas contendo:

- a) O Gráfico de Dispersão, e interprete os principais resultados.
- b) A Matriz de Covariância e a Matriz de Correlação Linear de Pearson, e interprete os principais resultados.
- c) O Gráfico de Correlação Linear de Pearson, e interprete os principais resultados.
- d) A Equação da Regressão Linear Múltipla.
- e) A verificação da adequação do modelo (R^2 e Teste F), e interprete os resultados.
- f) Faça uma previsão utilizando, pelo menos, 02 variáveis explicativas.

05) **(2,0 pontos)** Elaborar um Relatório Estatístico em Word utilizando os parâmetros e gráficos encontrados nas questões anteriores, para a extração de insights que visem a importância dos seus resultados para o tema proposto.

06) Postar no Portal:

- A(s) base(s) de dados em Excel (formato .xlsx);
- O código em R utilizado (formato .R);
- O Relatório Estatístico em PDF (formato .pdf)

BOA PROVA!!!