

Nome: \_\_\_\_\_ matrícula: \_\_\_\_\_

### **Etapa 1: Informações da base de dados**

(obrigatório utilizar uma base com dados reais)

Link da fonte: \_\_\_\_\_

Descrição da base (do que se trata esses dados)

Variável alvo (o que pretende prever/classificar)

Descrição sucinta das variáveis da base

### **Etapa 2: Preparação da base de dados**

(obrigatório justificar e interpretar)

Identificar e tratar nans

Identificar e tratar outliers

Aplicar normalização

Aplicar discretização supervisionada e não supervisionada

### **Etapa 3: Reduzir a dimensionalidade da base**

Aplicando PCA variáveis quantitativas

### **Etapa 4: Regressão linear múltipla**

### **Etapa 5: Regressão logística**

### **Etapa 6: KNN**

### **Etapa 7: Clusters (hierárquico e não hierárquico)**

### **Etapa 8: Arvore de decisão**

### **Etapa 9: Resultados**

Compare todos os resultados, apontando os melhores e piores modelos com foco no seu objetivo de estudo

### **Etapa 10: Apresentação**