**Descrizione Progetto**

**1-Analisi Dataset**

**-Descrizione delle variabili**

**-Istogrammi delle frequenze**

**-Calcolare media e deviazione standard delle variabili e riportarle in una tabella.**

**-Rimozione degli outliers**

**-Matrice di correlazione**

**-Divisione dataset in training e test**

**2- Regressione Lineare**

1. **Creazione del modello**
   1. **Riportare i Beta con confidence interval**
   2. **Riportare R^2**
   3. **Riportare test RMSE**
2. **Analisi dei residui**
   1. **Plottare**
      1. **Istogramma dei residui**
      2. **Normal Q-Qplot**
      3. **Residual (studentized) vs Fitted Values**
   2. **Test sui residui (?)**
      1. **Shapiro Wilk**
      2. **Kolmogorov-Smirnov**
      3. **bpTest (Heteroschdadistct)**

**3- K-fold + Polynomials Values (?)**

**-Test RMSE e grafico per i cinque modelli creati**

**-Test Anova per confrontare i modelli**

**4- Subset Selection**

**-Divisione dataset in training e test**

**5- Lasso e Ridge Regression**

1. **Creazione del modello**
   1. **Definire un setdi lamnda per le analisi**
   2. **Scelta del miglior lambda**
   3. **Riportare i Beta del modello**
   4. **Riportare test RMSE**
2. **Grafici**
   1. **Andamento dei coefficienti per lambda e l2 norm**
   2. **Andamento di lambda e MSE**
3. **Analisi dei residui (?)**

**6-GAM**

1. **Valori delle spline**
2. **Grafici delle spline**
3. **Test MSE**
4. **Analisi dei residui(?)**

**7-Regression Tree**

**8-Bagging e Boosting**