Instructions

1. For the file Biscobis.xlsx, referring to a sample of 100 client companies of a large supplier in the industrial sector, perform the following analyses:
• Bayesian regression
• K-Nearest Neighbors (KNN)
• Gradient Descent (1,000 iterations, learning rate 0.01)
• Stochastic Gradient Descent (1,000 iterations, learning rate 0.01).

Analyze and compare the obtained results.
Which model performed best considering the training dataset?

Dependent variable:
- X9 = Level of service usage (how much of the company's total product purchases come from Biscobis)

Independent variables: Rated from 0 to 10, representing Biscobis' attributes:
- X1 = Speed of product delivery
- X2 = Price level
- X3 = Price flexibility
- X4 = Supplier image
- X5 = Overall service quality
- X6 = Sales force image
- X7 = Product quality

2. Consider the file student_performance.xlsx with the following layout:

- Hours Studied: Average number of study hours in the previous month
- Previous Scores: Grade on the previous test
- Extracurricular Activities: Indicates whether the student has extracurricular activities (Yes or No)
- Sleep Hours: Average number of sleep hours
- Sample Question Papers Practiced: Number of practice question sets completed
- Performance: Test score

Make the necessary adjustments to the dataset and perform Bayesian regression and KNN analyses.
Compare the results.

Was there a significant difference between the training and testing datasets' outcomes using these techniques?
Persist the best-performing model.

3. A construction materials retail chain (CONSTRUCAO) operating in 52 regions aims to study the quantity sold (qt_vend) of a specific type of material.

Potential influencing factors include:
- Advertising expenditure (gast_prop)
- Number of active accounts (n_cont)
- Number of brands (n_marc)

- Number of stores in the region (n_loj)

Conduct Gradient Descent and Stochastic Gradient Descent regression analyses under the following scenarios:

- Learning rates: 0.01, 0.02, and 0.05
- Number of iterations: 1,000 and 10,000

Which technique and scenario provided the best result?