

MA289 Intro to Statistical Learning Assignment 1: EDA

MAJ Patrick Kuiper

DUE No Later Than 03 FEB 2026 via Shared Google Colab Notebook.

You will perform **exploratory data analysis (EDA)** on a real dataset. You should demonstrate how EDA informs modeling decisions, including feature selection, scaling, data imputation, and model choice.

Grading Rubric

Criterion	Points	Score
1. Clarity of Explanation	15	
– Clear introduction of the dataset	5	_____
– Clear explanation of EDA goals and purpose	5	_____
– Clear step-by-step reasoning from EDA to model	5	_____
2. Exploratory Data Analysis Quality	15	
– Correct identification of variable types	5	_____
– Appropriate summaries and visualizations	5	_____
– Insights that motivate modeling choices	5	_____
3. Mathematical Understanding	20	
– Correct formulation and reasoning	10	_____
– Links EDA, mathematics, and implementation	10	_____
Total	50	_____