Grading Rubric: Homework 1

Total: 100 points

Problem 1: One-Dimensional Data (20 pts)

1.1 Summary Statistics (8 pts)

- Mean computed correctly (2 pts)
- Median computed correctly (2 pts)
- Variance computed correctly (2 pts)
- Standard deviation computed correctly (2 pts)

1.2 Histogram & Interpretation (6 pts)

- Histogram with ≈ 50 bins (3 pts)
- Correct reasoning about multimodality and inadequacy of mean/median (3 pts)

1.3 Additional Descriptive Metrics (6 pts)

Problem 2: kNN Regression (40 pts)

2.1 Scatter Plot & Relationship (4 pts)

- Correct plot (2 pts)
- Notes periodic or non-linear structure (2 pts)

2.2 kNN Implementation (10 pts)

- Correct distance computation (3 pts)
- Correct neighbor identification (3 pts)
- Correct averaging/prediction (3 pts)
- Clean, runnable code (1 pt)

2.3 MSE for $k = 5$ (6 pts)

- Valid 80/20 split (2 pts)
- Correct MSE computation (2 pts)
- Reasonable reported value (2 pts)

2.4 MSE vs. k Plot (10 pts)

- Correct plot of MSE vs. k (6 pts)
- Correct interpretation (4 pts)

2.5 Bias–Variance Interpretation (10 pts)

Problem 3: Linear Regression (40 pts)

3.1 Regression Without Intercept (5 pts)

3.2 Regression With Intercept (5 pts)

3.3 Centered Regression (8 pts)

- Correct centering transformations (3 pts)
- Correct explanation of intercept removal (5 pts)

3.4 Scatter Plot & Interpretation (5 pts)

- Correct plot (2 pts)
- Correct interpretation of non-linearity (3 pts)

3.5 Regression with $\sin(X_3)$ Feature (10 pts)

- Correct construction of `X3_sin` = $\sin$(`X3`) and fit (5 pts)
- Correct coefficient reporting (5 pts)

3.6 Plot of Transformed Feature (7 pts)

- Correct scatter plot (3 pts)
- Correct explanation of increased linearity (4 pts)