1. Considering the following data set:

	y_1	y_2	y_3	y_4
x_1	1	1	A	1.4
x_2	2	1	B	0.5
x_3	2	3	B	2
x_4	3	3	B	2.2
x_5	2	2	A	0.7
x_6	1	2	A	1.2

Assuming kNN, with $k = 3$ applied within a leave-one-out schema:

(a) Considering an y_3 categoric output variable and the Euclidean distance, provide the prediction for x_1.

(b) Considering an y_4 numeric output variable and the cosine similarities, provide the mean regression estimate for x_1.

(c) Considering a weighted-distance kNN, with Manhattan distance, identify both the **weighted-mode** estimate of x_1 for a y_3 outcome and the **weighted-mean** estimate of x_1 for a y_4 outcome.

(a)

(b)

(c)

2. Consider the following training data set:

	y_1	y_2	z
x_1	1	1	1.4
x_2	2	1	0.5
x_3	1	3	2
x_4	3	3	2.5

(a) Find the closed form solution for a linear regression, minimizing the sum of squared errors.

(b) Predict the target value for the query vector $x_{new} = \begin{bmatrix} 2 & 3 \end{bmatrix}^T$.

(c) Sketch the predicted three-dimensional hyperplane.

(d) Compute both the MSE and MAE produced by the linear regression.

(e) Are there biases on the residuals against any of the input variables?

(f) Compute the closed form solution, considering Ridge regularization term with $\lambda = 0.2$.

(g) Compare the hyperplanes obtained utilizing ordinary least squares and ridge regression.

(h) Why is the Lasso regression usually preferred over Ridge regression for data spaces with a larger number of features?

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

3. Considering the following training data, with z as an ordinal variable:

	y_1	y_2	z
x_1	1	1	1
x_2	2	1	1
x_3	1	3	0
x_4	3	3	0

(a) Find a linear regression using the closed form solution.

(b) Assuming an output threshold $\theta = 0.5$, provide the predicted class for $x_{new} = \begin{bmatrix} 2 & 2.5 \end{bmatrix}^T$.

(a)

(b)

4. Considering the data below to learn the following model, compare:

$$z = w_1 y_1 + w_2 y_2 + \epsilon, \epsilon \sim \mathcal{N}(0, 0.1)$$

	y_1	y_2	z
x_1	3	-1	2
x_2	4	2	1
x_3	2	2	1

(a) $w = \begin{bmatrix} w_1 & w_2 \end{bmatrix}^T$, using an MLE approach.

(b) w using the Bayesian approach, assuming $p(w) = N(w \mid \mu = [0,0], \Sigma = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix})$.

(a)

(b)

5. Identify a transformation to aid linearly modelling the data set above. Sketch the predicted surface.

6. Consider both logarithmic and quadratic transformations:

$$\phi_1(x_1) = \log(x_1), \quad \phi_2(x_2) = x_2^2$$

(a) Plot both of the closed form regressions.

(b) Which transformation minimizes the sum of squared errors on the original data?

(a)

(b)

7. Select the criteria promoting a smoother regression model:

(a) Applying Ridge and Lasso regularizations to linear regression models.

(b) Increasing the depth of a decision tree regressor.

(c) Increasing the k parameter of a kNN regressor.

(d) Parametrizing a kNN regressor with uniform weights, instead of the default distance-based weights.