Q1 (a)

- We calculate the sum of squared errors for each Model using the formula

$$\textit{sum of squared errors} = \frac{1}{2}\sum_{i=1}^{n}(t_i - \mathbb{M}(\mathbf{d}_i))^2$$

The relevant aggregate figures are as follows:

Model 1 SSE: 16750

Model 2 SSE: 47369

Q1 (b)

- The R^2 measure is calculated using the following formula

$$R^2 = 1 - \frac{sum\ of\ squared\ errors}{total\ sum\ of\ squares}$$

- We already have the sum of squared errors from part (a). However, we need to calculate the total sum of squares, which can be done using the formula

$$total\ sum\ of\ squares = \frac{1}{2}\sum_{i=1}^{n}(t_i - \bar{t})^2$$

Q1 (b)

- Using the formula, the total sum of squares for

Model 1 = 234064.4

Model 2 = 274579.6

Putting these values into the R^2 formula, we get

R^2 for Model 1 = $1 - \frac{16750}{234064.4} = 0.928$

R^2 for Model 2 = $1 - \frac{47369}{274579.6} = 0.827$

Based on the above, Model 1 is better able to capture patterns than Model 2.

Q2

- The contingency table below shows the evaluation results for a binary classifier applied to a set of 768 test examples, which are annotated with the class labels (A, B). From this table calculate:
  a) The precision score for both of the classes.
  b) The recall score for both of the classes.
  c) The F1-measure score for both of the classes.
  d) The overall classification accuracy for all the data.

	Predicted Class: A	Predicted Class: B	
Real Class: A	407	93	A
Real Class: B	108	160	B

Q2(a,b)

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN} = \text{Sensitivity}$$

Predicted

	Pos	Neg	
P		FN	Pos
N	FP	TN	Neg

Real

Predicted

	Pos	Neg	
P		FN	Pos
N	FP	TN	Neg

Real

Note: These measures are always relative to one class!

Predicted Class

A	B	Real Class
407	93	A
108	160	B

Class	Precision	Recall
A	407/(407+108) = 0.79	407/(407+93) = 0.814
B	160/(93+160) = 0.632	160/(108+160) = 0.597

Q2(c)

- F1-Measure: harmonic mean of precision and recall

$$\mathrm{F1} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

Also relative to one class!

Class	Precision	Recall	F1
A	407/(407+108) = 0.79	407/(407+93) = 0.814	(2*0.79*0.814)/(0.79+0.814) = 0.802
B	160/(93+160) = 0.632	160/(108+160) = 0.597	(2*0.632*0.597)/(0.632+0.597) = 0.614

Q2(d)

- Accuracy: Number of predictions correct / all predictions

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

Predicted Class

A	B	
407	93	A
108	160	B

Real Class

Accuracy score is relative to the overall dataset, often reported as a percentage.

OVERALL ACCURACY:

(407+160)/(407+93+108+160)

= 73.8281%

Q3

- The table below shows the true classes for 12 example emails, which are labelled as “spam” or “non-spam”. The table also reports the labels predicted by three different binary classifiers

Example	True Class Label	KNN	DecisionTree	LogisticRegression
1	spam	spam	spam	spam
2	non-spam	non-spam	spam	non-spam
3	spam	non-spam	non-spam	spam
4	non-spam	non-spam	non-spam	non-spam
5	spam	spam	spam	spam
6	non-spam	non-spam	non-spam	non-spam
7	non-spam	spam	spam	non-spam
8	non-spam	non-spam	spam	spam
9	spam	spam	non-spam	spam
10	spam	spam	non-spam	non-spam
11	spam	non-spam	non-spam	spam
12	spam	spam	spam	spam

Q3 (a,b)

Example	True Class Label	KNN	Decision Tree	Logistic Regression
1	spam	spam	spam	spam
2	non-spam	non-spam	spam	non-spam
3	spam	non-spam	non-spam	spam
4	non-spam	non-spam	non-spam	non-spam
5	spam	spam	spam	spam
6	non-spam	non-spam	non-spam	non-spam
7	non-spam	spam	spam	non-spam
8	non-spam	non-spam	spam	spam
9	spam	spam	non-spam	spam
10	spam	spam	non-spam	non-spam
11	spam	non-spam	non-spam	spam
12	spam	spam	spam	spam

#Correct	9/12	5/12	10/12
Accuracy	75%	41.7%	83.3%

"Spam"	**TP**	5	3	6
	FP	1	3	1
	Precision	5/6 = 0.833	3/6 = 0.5	6/7 = 0.857

Overall Accuracy: Number of predictions correct / all predictions

Precision for spam: Number of correct spam predictions / all predictions of spam

Logistic Regression is most accurate

Logistic Regression has highest precision for spam