Q4. Inference Procedures (Variant 2)

(a) ***Binary Classification (4 Marks)***
For following binary classification outcome table, calculate the following appraisal metrics.

(i) (1 Mark) accuracy;

(ii) (1 Mark) recall;

(iii) (1 Mark) precision;

(iv) (1 Mark) F-measure.

	Predict Negative	Predict Positive
Observed Negative	9560	100
Observed Positive	270	70

(b) ***Theory of Statistical Inference (4 Marks)***

i. (2 Marks) In the context of hypothesis testing, explain what a p-value is, and how it is used. Support your answer with a simple example.

ii. (2 Marks) What is meant by Type I error and Type II error?

(c) ***Single Sample Proportions (5 Marks)***

Part B

A well-known polling company estimates that 57% of Irish voters support a new constitutional amendment. 800 people were randomly surveyed and asked about their voting preferences. 482 of the 800 people responded positively to the amendment. You are required to:

i. (1 Mark) Obtain a point estimate of the proportion of people supporting the constitutional amendment.

ii. (4 Marks) Construct a 95% confidence interval for the proportion of people in favour of the constitutional amendment.

(d) ***Inference Procedures with `R` (6 Marks)***
The standard deviations of data sets `X` and `Y` are 10 and 9 respectively. An inference procedure was carried out to assess whether or not `X` and `Y` can be assumed to have equal variance.

i. (2 Mark) Formally state the null and alternative hypothesis.

ii. (1 Mark) The Test Statistic has been omitted from the computer code output. Compute the value of the Test Statistic.

iii. (2 Marks) What is your conclusion for this procedure? Justify your answer.

iv. (1 Marks) Explain how a conclusion for this procedure can be based on the 95% confidence interval.

```
        F test to compare two variances

data:  X and Y
F = ......, num df = 13, denom df = 11, p-value = 0.7349
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3639938 3.9475262
sample estimates:
ratio of variances
          .......
```

(c) ***Single Sample Means (6 Marks)***
Two samples of students are randomly selected from two IT training companies; Echelon and Deltatech. The mean and the standard deviation of the number of marks obtained in a well known IT competency exam by both sets of students are described below:

	Number	Mean	Std. Dev.
DeltaTech	14	24	10
Echelon	12	22.5	9

i. (1 Mark) Obtain a point estimate of the difference in mean scores

ii. (5 Marks) Construct a 95% confidence interval for the proportion of people in favour of the constitutional amendment.