STATISTICAL METHODS FOR DECISION MAKING

TOTAL MARKS:30 **DURATION: 2 HOURS**

INSTRUCTIONS

1. *Candidates should answer all the questions in the same order provided in the question paper.*
2. *Any activity that compromises the integrity of the examination will not be permitted.*
3. *Students should complete the examination within the provided timeline.*
4. *Candidates are expected to check and ensure that the correct answer file (in. ipynb format) is uploaded in LMS.*

SECTION A: 5 MARKS

1. Within a school district, students were randomly assigned to one of two Math teachers - Mrs. Smith and Mrs. Jones. After the assignment, Mrs. Smith and Mrs. Jones had 30 students each. At the end of the year, each class took the same standardized test. The scores are of students are given below. Test the hypothesis that weather Mrs. Smith and Mrs. Jones are equally effective teachers. Use a 0.05 level of significance.

Mrs_Smith = [66,68,71,62,63,63,88,87,89,85,98,91,54,53,69,58,70,86,84,87,87,66,61,60,56,77,75,71,70,59,62,63]

Mrs_Jones = [71,80,81,87,64,68,60,72,71,77,39,88,63,55,64,63,71,71,48,63,77,49,67,69,56,44,48,77,63,88,67,67]

A. Frame Null and Alternate Hypothesis (1 Mark)

B. State which test to be performed (1 Mark)

C. Perform the test (2 Mark)

D. Interpret the results (1 Mark)

SECTION B: 10 Marks

2.Read the dataset (Churn.xlsx) it consists of all the information regarding telecom industry for any telecom industry Churning rate of customers causes direct impact on the business so by performing the statistical analysis on the data set we need to find out on what factors the churning rate is dependent on. Perform the following tasks on the data:

A. List out the numerical and categorical features in the dataset and mention the possible statistical hypothesis test for give features in the dataset (2 marks)

B. State the null and alternative hypothesis for any one of the test mention in the question number one (2 marks)

C. Now perform any two required statistical test on the data to determine what are the significant features in the data with respect to target variable (95% C.I)? (4 Marks)

D. Interpret the results? (2 Marks)

SECTION C: 15 Marks

3. Read he data set(Admission.xlsx) consists of complete educational details of students right from their schooling to MBA and previous work experience. Salary is the target variable in the data. Perform the following tasks on the data:

A. List out the numerical and categorical features in the dataset and mention the possible statistical hypothesis test for give features in the dataset (2 marks)

B. State the null and alternative hypothesis for any one of the test mention in the question number one (2 marks)

C. Now perform required statistical tests on the data to determine what are the significant features in the data with respect to target variable (95% C.I). (5 Marks)

D. Interpret the results (2 Marks)

E. Build a model and find the coefficient and intercept value? (4 Marks)