

Instructions

GENERAL INSTRUCTIONS: The file found under Midterm Examination Data Set (in Canvas Files) is the data set for patients admitted in a hospital. It contains 10 columns shown under the first row of the said data set. The first column represents the Respondent ID for tracking purposes, and the remaining 9 columns represent the variables under investigation. The variable labels and specified values are shown below:

Each instance represents individual patients and their various medical attributes along with diabetes classification

Number of Instances: 768

Number of Attributes: 9

Pregnancies: Number of pregnancies

PG Concentration: Plasma glucose at 2 hours in an oral glucose tolerance test

Diastolic BP: Diastolic Blood Pressure (mm Hg)

Tri Fold Thick: Triceps Skin Fold Thickness (mm)

Serum Ins: 2-Hour Serum Insulin (μ U/ml)

BMI: Body Mass Index: $(\text{weight in kg} / (\text{height in m})^2)$

DP Function: Diabetes Pedigree Function

Age: Age (years)

Diabetes: Whether or not the person has diabetes (0 = Healthy, 1 = Sick)

Question 1: Read or import the Diabetes Data from Excel to IBM-SPSS and label appropriately. Upload the converted file in SPSS here.

Question 2: Recode the following variables: Pregnancies to Pregnancies1, BMI to BMI1, and Age to Age1. Use the following values:

Pregnancies1: 1 = 0-3, 2 = 4-6, 3 = 7-10, & 4 = Above 10

BMI1: 1 = Underweight (Below 18.5 kg/m²); 2 = Normal Weight (18.5 - 24.9); 3 = Overweight (25.0 - 29.9); 4 = Obese (30.0 & above)

Age1: 1 = Below 25 years old; 2 = 25 - 34; 3 = 35 - 44; 4 = 45 - 54; 5 = 55 years old & above

If you are done, produce the frequencies for the following variables: Pregnancies1, BMI1, Age1, and Diabetes. Output should be in Excel form and upload only the frequencies for the variables specified.

Question 3: Test whether or not there is a significant mean difference in the BMI between the healthy & sick patients. Use the steps discussed in the class. Answers should be in Word document and upload answer (file) here

Question 4: Test whether or not there is/are significant mean differences in the PG Concentration of the patients when they are grouped according to their ages (Age1). If necessary, employ multiple comparisons (post hoc). Use the steps discussed in the class. Formulate conclusions & implications appropriately based on the results of your test & multiple comparisons. Results should be in Word document & should be uploaded here.

Question 5: Test whether or not the Diastolic BP of the patients is significantly different from 90 mm Hg. Use the steps discussed in the class and answer showed uploaded here in Word document.