School of Technology
Dr. Yuanji Cheng

MA660E

Lab assignment

Part one

In this part, you make use of the job survey data from the book: *Quantitative data analysis with SPSS*.
(the spss file **jss13_ht22.sav** or excel file Data_source) .
ethnicgp = ethnic group, 1 = White, 2 = Asian, 3 = West Indian, 4 = African, 5 = other
gender (male =1, female =2),
income = gross annual income before tax in 1000£
age in years
years = number of years working at this firm
commit = organisational commitment (scale 1 ~ 5)
satis = job satisfaction
autonom = job autonomy
routine = job routine
attend = attendance at meeting (yes = 1, no = 2)
skill = rated skill (unskilled, semi-skilled, fairly skilled, highly skilled, 1 ~ 4)
prody = rated productivity (very poor, poor, average, good, very good, 1 ~ 5)
qual = rated quality (very poor, poor, average, good, very good, 1 ~ 5)
absence = days of being absent in last 12 months

You may use any computer program for calculations (MATLAB, Python, Excel, SPSS, R……)

Exercise 1.1

a) Make plot of gender in Bar chart and ethnic group in pie diagram.
b) Make a five number summary (max, min, median, the first and third quartiles) of Age then a box-plot
c) Find the mean and standard deviation of income, and as well as histogram of it.

Exercise 1.2 Consider the variables income and absence.

a) Make scatter plot to visualize the relationship between them
b) Find the simple regression model where income is dependent variable and rated absence is independent variable. What is your determination coefficient?

Exercise 1.3 Study the multiple regression model with satis as dependent variable and commit, autonom, income, skill, rated quality, age, years as independent variables.

a) Which variables among them do NOT have any significant impact on satis ?

b) Find a simpler multiple regression model with satis as dependent variable by deleting all those non-impact variables.

Exercise 1.4 Find confidence interval of job satisfaction and also confidence interval of *difference* in job satisfaction between men and women.

Exercise 1.5 Using the Mann-Whitney-Wilcoxon test to see if there is any significance in skill between man and woman, and compare the result with confidence interval for the difference.

Exercise 1.6 Using the Kruskal-Wallis test to see if there is any significance in absence among ethnic group, and compare the result with One-Way ANOVA analysis.

Exercise 1.7 Re-code the variable “income” into “income_class “ with proper choice of limits of class classification (suggestion: low income class [Min, Q1], middle income class (Q1, Q3], high income class (Q3, Max]) thereafter investigate if there is any significant relationship between income_class and skill.

Part two

In this part you have chance to work on your own data and thus have free choice of topic.

Exercise 2.1 Descriptive statistics analysis for at least two qualitative and quantitative variables

Exercise 2.2 Confidence interval for one quantitative variable; Confidence interval for difference between two groups.

Exercise 2.3 Carry out a T-test to check if the difference in characteristics between two groups is significant, or conduct an ANOVA to see if all groups have the same mean value in some characteristics.

Exercise 2.4 Non-parametric test for same variable as in 3) and even compare the conclusion(s) with ANOVA

Exercise 2.5 Correlation analysis, thereafter, identify the strongest correlation and statistically not significant relation(s)

Exercise 2.6 Make a linear multiple regression analysis.

If you are interested in COVID-19 and you may find the data from either WHO website: https://covid19.who.int/ or John Hopkins university coronavirus resource centre: https://coronavirus.jhu.edu/

Lab report

You may work individually or in team with max two students.

Remark on Lab report
General rules: Please follow the template on next page.
For each question the answer should contain both table / figure and the associated interpretation.
Only table or figure will be treated as **"fail**".

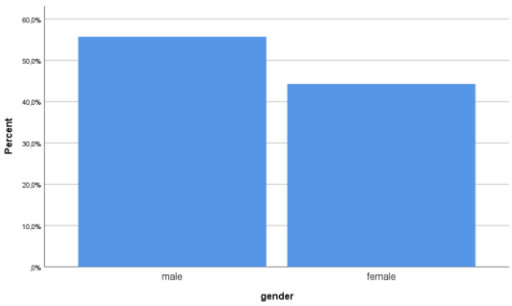
Deadline for hand-in of lab report via Canvas: Sunday 2024-10-27.

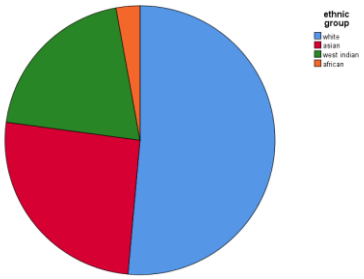
MA660E, Lab Report

Yuanji Cheng, Martin Svensson

Part One

Exercise 1.1 a)





From the bar chart for gender, we can see that there are more male participants than female. The pie diagram shows that the majority of participants are white (more than 50%), and very few are African.

b) …

c) …

Exercise 1.2 a)

Part Two

First give a short description of your data by introducing the variables and their units, then answer the questions 2.1) to 2.6) in a similar way as in part one.

Exercise 2.1 a) Qualitative variable