

METU, Department of Economics

Instructor: Hakan Ercan & Kağan Parmaksız

T.A.: Oğuzhan Çelik

ECON 106 – ASSIGNMENT 2

Note: Don't waste time entering the dataset manually; set the location/path of the downloaded Excel data as the working directory, and then use the **read_excel** package to directly pull the data into R.

Q1) A random sample of 100 business majors was asked a series of demographic questions including major, gender, age, year in school, and current grade point average (GPA). Other questions were also asked for their levels of satisfaction with campus parking, campus housing, and campus dining. Responses to these satisfaction questions were measured on a scale from 1 to 5, with 5 being the highest level of satisfaction.

Finally, these students were asked if they planned to attend graduate school within five years of their college graduation (0: no, 1: yes). These data are contained in the sheet “Q1”.

- i) Construct the cluster bar chart of the respondents’ major and gender.
- ii) Construct a pie chart of their majors.
- iii) Construct a histogram of their GPAs depending on bin numbers. Set the bin numbers (upper levels) as 0.5 1, 1.5, 2, 2.5, 3, 3.5, 4.

Q2) The information about the salaries and the consumption levels are stored in the sheet “Q2”. By using this data, construct a scatter plot.

Q3) There is a panel dataset stored in the sheet “Q3”. The dataset consists of observations for 70 countries over the time period from 1992 to 2018. The first three columns are country name, country code (in three character), and year. The variables are cereal yield, carbon dioxide emissions per capita, total carbon dioxide emissions, land under cereal production, rural population, temperature, and precipitation, respectively.

- i) Create a country subset by selecting one of the countries in the dataset.

- ii) Fit a linear regression model by using the data frame obtained from part 1. Use 'c_yield' as dependent variable and 'co2_per', 'co2_kt', 'land', 'rural_pop', 'temp' and 'precip' as the independent variable. Summarize the fitted model.