Business Modeling and Analysis

Objectives

In this assignment you are expected to develop a business report that will be presented to a senior manager of a human capital management company. The report should be informative but concise and follows a specific structure that allows the document to be easily read and understood. For this purpose, please ensure that you have studied supplementary material about the academic report writing, available on e-tivity 5.4 before you start your assignment. This assignment is worth 30% of your final mark.

This assignment provides students with opportunity to:

- combine statistical analysis and report writing skills to prepare a concise, non-technical business report,
- draw a random sample to complete statistical analysis,
- develop and enhance computing skills, specifically the use of the built-in functions in MS Excel and MS Word,
- apply statistical techniques to a data set, and
- reinforce taught concepts including descriptive statistics, sampling and estimation, hypothesis testing, correlation and regression.

Background

Margaret is the CEO of a Human Capital Management company in Melbourne. She has recently attended a TED Talk in which Michael Green discussed the importance of the 2030 United Nation Sustainable Development Goals. Michael presented challenges that countries around the world would face to meet the UN Sustainable Development Goals and elaborated on the importance of the Social Progress Index in that process (please watch two videos posted on e-tivity 7.5, *What are the UN Sustainable Development Goals?* and *How we can make the world a better place by 2030?*). After the seminar, Margaret decided to further study the importance of the Social Progress Index and compare countries based on their performance at each sub category of this index.

For that purpose, Margaret asked Jason, a newly appointed research officer, to perform the analysis. Margaret downloaded 2017 Social Progress Index data set (available on e-tivitiy 7.5), passed it to Jason and asked him to run several analyses listed below. There are 50 variables under 12 categories for 182 countries in this data set. For the purpose of analysis, Margaret asked Jason to concentrate on categories with even numbers (which are highlighted in yellow in the dataset). They are:

Category 2: Water and Sanitation
Category 4: Personal Safety
Category 6: Access to Information and Communications
Category 8: Environmental Quality
Category 10: Personal Freedom and Choice
Category 12: Access to Advanced Education

Presentation of a business report

You are going to help Jason in developing a report that addresses Margaret's concerns. Since Margaret is not very familiar with statistical terms, try to not just quote statistics or analysis results, but explain what they mean. In general, do not include formulae, calculations, definitions of statistical terms or discussions on how graphs are constructed in the report. Where appropriate, these may be included in the appendices. Follow the Harvard Style Guide (available on e-tivity 5.4) for all in text references, reference lists and bibliographies.

It is important that the values that have been calculated are correctly analysed, discussed and interpreted. The written description of the main features of the tables and graphs should be presented. The emphasis will be on the computation of statistics and construction of graphs, as well as on interpretation of analysis. The assessment should be in one single Microsoft Word compatible document and should be written in 12 point font. It is expected that all work submitted, will have been edited for spelling, grammar and clarity.

Tasks

Most of your statistical calculations should be carried out using Excel only and you will use Microsoft Word and Excel to complete this assignment.

1. Select a Random Sample

Select a random sample of size 100 from the given 182 countries Remember that 182 countries constitute the population of interest and ***do not*** calculate any parameters such as mean and standard deviation from the population. Then from each category with an even number (categories 2, 4, 6, 8, 10 and 12), **select one variable only**. For example, under category 2, Water and Sanitation, there are three variables. They are:

- Access to piped water
- Rural access to improved water source
- Access to improved sanitation facilities

You need to select only one of them based on your preference for this category. As a result, each student will have a different sample that includes different variables. After this process, you will have a data set with a **total of eight variables** (including Countries Names and Continent) and 100 observations corresponding to each variable. If you have variables with missing values in your sample, you **are allowed** to remove those observation (delete the entire row that has a missing value). In that case your sample will be below 100 observations, which is **acceptable**. You will use this ***sample data*** to complete tasks 2 to 5.

2. Descriptive Statistics

Use appropriate data summary methods to describe the variables in your sample. At this stage, you only concentrate on **seven variables and exclude Countries Names.** For each of the seven remaining variables, you need to have a graph and a table and descriptions of the graph

and the table. Use an appropriate graphical and/or summary statistical technique based on the type of variable. These techniques include:

Tabular Techniques: e.g. Frequency tables and Grouped (joint) frequency tables.

Graphical Techniques: e.g. Pie chart, Bar graph, Histogram (avoid three dimensional graphs).

Summary Statistics: e.g. Mode, Median, Mean, Standard Deviation, Range, Coefficient of Variation and Interquartile Range (you do not have to report all of these summary statistics. Choose the most appropriate measures based on your personal judgement).

You will need to choose the most appropriate technique(s) for each analysed variable. Less appropriate/inappropriate techniques will receive fewer/no marks.

For a nominal or an ordinal (discrete) variable, draw a graph **and** present a frequency table in percentages.

For a ratio or an interval variable (continuous), draw a graph **and** a summary statistics table.

Try to use different graphs, e.g. pie chart/bar chart or histogram as much as possible for different variables. Do not draw two different graphs for the same variable.

3. Confidence intervals

Estimate the following quantities using 95% confidence intervals. Explain the meaning of your confidence intervals.

i. The average of the variable that you have chosen from Category 2 (Water and Sanitation).
ii. The average of the variable that you have chosen from Category 12 (Access to Advanced Education).

4. Hypothesis Testing

Margaret has some concerns about the common assumptions regarding social development across countries in different continents. As such, she asked Jason to carry out the following hypothesis tests based on the chosen variables in these categories.

i. It has been argued that the level of Access to Advanced Education is higher among ***European*** countries than ***African*** countries?
ii. It there any difference in terms of Personal Safety between ***Asian and American*** countries?
iii. It there any difference in terms of Environmental Quality between ***European and American*** countries?

Only present the main elements of your analysis and your important findings in the main section of the report. The computations and Excel outputs should be placed in an appendix.

5. Correlation and Regression

Margaret is interested in the relationships between some of these social variables. Margaret asked Jason to conduct regression analysis on the following variables:

- Environmental Quality related variables (Category 8) and Water and Sanitation variables (Category 2).
- Access to Information and Communications related variables (Category 6) and Personal Freedom and Choice variables (Category 10).

Use these variables to develop two regression models and make sure to provide full discussions on each test. Use your chosen sample for these analyses. You discussions for each test should include:

- a scattergram and full interpretation
- an estimate of the linear regression model
- the coefficients of correlation and determination
- a test of the hypothesis that there is no linear relationship between dependent and independent variables.

Assignment structure:

Presentation

Presentation is an important feature of a business report. The guide to presentation that follows includes an Executive summary, Introduction, Analysis and Appendices (please follow the instructions given in e-tivity 5.4). Your assignment should contain the following sections:

Executive summary

This is your first page and **not** to be **included in the page count**. Executive summary should be a combined form of Introduction and Conclusions.

- Report only the highlights of the findings
- Entice an Executive to read on
- Essentially a lively summary of the main conclusions
- No longer than one page since not counted in the page count, must be on a separate page from the rest of the report.

Introduction (about half a page)

This should contain information about what we expect to read in the project.

- State the purpose of the report, e.g. what you will discuss in the report

- Outline the contents of the report
- Provide a brief description of the methodology
- Describe the source of the data and state its location in the report.

Analysis

This section contains a thorough yet non-technical description of all the findings (graphs and tables will be included only where they help this discussion). It details the results that were highlighted in the Executive Summary. No calculation is required here but appropriate graphs and tables which are needed to support your discussions should be included. Your analysis section should contains the following divisions:

- DESCRIPTIVE STATISTICS (about two pages)

There are 8 variables in your data sets. Depending on the level of measurement (Nominal, Ordinal or Interval/Ratio measures), you will need to provide an appropriate graph, a summary statistics table and descriptions related to **only 7 of those variables** (exclude Countries Names). **Take one variable at a time.**

- CONFIDENCE INTERVAL (about half a page)

"We are 95% confident that....."

Do not show the method of calculating a confidence interval (calculation details should be reported in appendices). Do not draw a graph. Report the sample data used for analysis of this section in the appendix.

- HYPOTHESIS TESTING (about one page)

Do not write H_0, H_A etc. Only your result such as "hypothesis test was performed on… ....." Also report the type of hypothesis (i.e., upper, lower, two-tail), the test statistic, *p*-value, t-value and degrees of freedom. Report the sample data used for these analyses in the appendix.

- CORRELATION and REGRESSION (about one and half pages)

Present your scatter plots with a line of best fit included in the graphs. Clearly specify the dependent and independent variables and describe the strength of relationships using R and R squared values. A full linear regression model should be stated on both graph and in the discussion.

Interpret the slope and intercept coefficients and hypothesis test result (t-value, *p*-value, etc.). Report the sample data used for these analyses in the appendix.

Conclusion and limitations (about one page)

All the important findings and results of your work should be presented here in a concise format. Make sure not to give tables or graphs here. This section you need to present

- The final conclusion of your analysis of all sections. It is essentially an expansion of the Executive Summary written from the point of view that the Executive Summary has not been read.
- Limitations of your analysis. In this section, you need to critically evaluate the method and data used in this analysis. What are the potential shortcomings of your conclusions?
- This section ends with a discussion on the ethical implications of statistical methods. What are the ethical elements that needed to be considered when reporting results of a statistical analysis? You need to consult with the literature for this section.