**Dublin Business School**

**Electronic Assignment Cover sheet**

**Student (s) Number as per your student card:** Juliana Salvadori - No: 10521647

**Course Title:** MSc Data Analytics

**Lecturer Name:** Terri Hoare

**Module/Subject Title:** Data Visualisation (B9DA106)

**Assignment Title:** CA2 - Summative assessment - EDA

GitHub Python code:

<https://gist.github.com/jusalvadori/2bf43a415f4a7f41eea490df30359a50>

GitHub R code:

**Data Visualisation - Exploratory Data Analysis**

**Introduction**

In this task we have explored a dataset based on a survey called ‘*What Do Men Think It Means To Be A Man?*’ from FiveThirtyEight website [1], to apply some of the visualisation techniques learned on this module. This study was carried out at a time when lots of questions arose regarding claims of sexual harassment, gender differences in the workplace and the [role of masculinity in society](https://www.nbcnews.com/think/opinion/father-s-day-men-are-experiencing-crisis-masculinity-solution-more-ncna884051), and it aimed to understand how men fell about being a man and what they think about masculinity. The dataset has 1,615 adults’ answers, who identified themselves as men, to nearly 40 questions about masculinity, workplace culture and intimacy, and other related subjects.

Hence, we selected a few questions from the survey to perform an exploratory data analysis of the data using three different visualisation tools: Tableau, R and Python.

**About Tableau**

Tableau is a visualisation tool utilized to help on data exploration and analysis that does not require any type of technical or programming skills as it presents the data in a simplified format that is easy to understand and work with [10]. Tableau is broadly used for data analytics and allows you to build from a simple chart, reports to complex dashboards. It works similarly to Excel at certain points and that is one of the reasons it is easy to operate, as many users already have some familiarity with Excel, however, Tableau is more powerful for visualisation purpose.

**About R**

R is a programming language specific for statists purpose and it has some basic visualisation package by default with a few functions that allows you to create the most common plots [8]. However, there are other more advanced packages such as ggplot2, lattice and plotly, that can be installed and allows you to work with different types of plots and provide more functions to improve plots appearance and interactivity. For this task we have used the ???? library.

**About Python**

Python is also a programming language; however, it is a general-purpose one, and widely used for data analysis due to it is capacity to deal with large datasets [6]. As Python is not a specific statistic or visualisation tool it does not have by default any visualisation package but has several visualisation libraries available such as matplotlib, ggplot and seaborn, that can be easily imported and used. For this task we have used the matplotlib and seaborn libraries.

Both R and Python, unlike Tableau, require some technical and/or programming skills to be operated and can take longer to build a simple chart as you might need to apply some data manipulation to get it in the right format to be able to call a plot function.

Python, there is no direct/easy way to add for instance the amounts or percentage in each piece of the char, you need to add a couple of lines to the code to do that and the percentages need to be calculated previously and then added to the chart.

How data is distributed, what are the inferences/conclusions that we can get from the visualisations?

What was developed on each tool, types of visualisation.

*Justify and appraise the visualisation techniques applied with reference to theory*

Comparison between Tableau, R and Python

*Critically evaluate and apply programming constructs, algorithms and interactive strategies • Critically appraise the capabilities of the three tools*

References

[1] Koeze, E. and Barry-Jester, A. (2018) ‘What Do Men Think It Means To Be A Man?’, FiveThirtyEight. Available at: <https://fivethirtyeight.com/features/what-do-men-think-it-means-to-be-a-man/> (Accessed: 17 Nov 2020)

[2] Hoare, T. (2020) ‘The Art and Science of Data Visualisation’, Dublin Business School. Available at: <https://elearning.dbs.ie/pluginfile.php/1223232/mod_resource/content/0/Tufte%20Design%20Principles.pdf> (Accessed: 17 Nov 2020)

[3] Shaffer, J. (2020) ‘Owning Your Data Story’, Adjunct Faculty, University of Cincinnati. Available at: <https://elearning.dbs.ie/pluginfile.php/1230557/mod_resource/content/0/Tableau%20Resources%20-%20Story%20Telling%20-%20Owning%20Your%20Data%20Story.pdf> (Accessed: 17 Nov 2020)

[4] Shaffer, J. (2020) ‘Pie Charts and Other Chart Types’, Tableau, Adjunct Faculty, University of Cincinnati. Available at: <https://elearning.dbs.ie/pluginfile.php/1230597/mod_resource/content/0/Tableau%20Resources%20-%20Chart%20Types%20%28Jeffery%20Shaffer%29.pdf> (Accessed: 17 Nov 2020)

[5] Hardin, M., Hom, D., Perez, R. and Williams, L. (2020) ‘Which chart or graph is right for you?’, Tableau. Available at: <https://elearning.dbs.ie/pluginfile.php/1230592/mod_resource/content/0/Tableau%20RESOURCE%20-%20Which%20Chart%20or%20Graph%20is%20Right%20for%20You.pdf> (Accessed: 17 Nov 2020)

R vs Python

[6] ‘Python vs R for Data Science: And the winner is..’ (2018), Data-Driven Science, Medium. Available at: <https://medium.com/@datadrivenscience/python-vs-r-for-data-science-and-the-winner-is-3ebb1a968197#:~:text=Python%20has%20caught%20up%20some,ggplot2%2C%20htmlwidgets%2C%20Leaflet).&text=Python%20is%20a%20powerful%2C%20versatile,of%20tasks%20in%20computer%20science.&text=Using%20more%20tools%20will%20only%20make%20you%20better%20as%20a%20data%20scientist>. (Accessed: 25 Nov 2020)

[7] ‘Data Visualization in R vs. Python’ (2019), R-bloggers. Available at: <https://www.r-bloggers.com/2019/12/data-visualization-in-r-vs-python/> (Accessed: 25 Nov 2020)

[8] Golchert, M. (2019) ‘Data Visualization in R vs. Python’, INWT Statistics. Available at: <https://www.inwt-statistics.com/read-blog/data-visualization-R-versus-python.html> (Accessed: 25 Nov 2020)

[9] Mendis, A. (2019) ‘R vs Python for Data Visualization’, KDnuggets. Available at: <https://www.kdnuggets.com/2019/03/r-vs-python-data-visualization.html> (Accessed: 25 Nov 2020).

[10] ‘What is Tableau? Uses and Applications’ (2020), Guru99. Available at: <https://www.guru99.com/what-is-tableau.html#:~:text=Tableau%20is%20a%20powerful%20and,form%20of%20dashboards%20and%20worksheets>. (Accessed: 26 Nov 2020)

[11]