## **End of Module Assessment 2: Individual Reflection Introduction**

Although I have a background in statistics, the module would have been a walk in the park. But it wasn't; in this module, I gained valuable data analysis skills and statistical knowledge, particularly in using R for statistical data analysis, interpreting statistical findings using p-values and confidence intervals, producing summary and contingency tables, and creating plots. These skills are essential for practical data analysis, and I am excited to share my reflections on this topic.

**Module Engagement:**

I engaged in several activities and external research work throughout the module, including data exploration, cleaning, hypothesis testing, visualisation, and modelling. One of the critical aspects of my learning was learning how to import data into R, as this is my first time using the R programming language and system. In addition, I learned how to deal with missing values and outliers and format data correctly when trying to get descriptive statistics for the dataset. I also learned how to conduct hypothesis testing using p-values and confidence intervals, interpret the results, and communicate them effectively. Finally, I also learnt how to use R for data visualisation and learned how to create practical summary tables, contingency tables, and plots.

Initially, I needed help with the syntax and formatting of R as it was my first-time exposure to the programming language. However, as the module progressed, I learned how to navigate R more effectively, organise my code and comments, and use various functions to perform statistical analyses. One of my most significant challenges was interpreting statistical findings using p-values and confidence intervals, which I needed clarification on at the beginning of the module. Coming up with the null and alternative hypotheses was also a bit of a learning curve at the early stage of the module. It took a lot of work to understand the concepts of statistical significance and confidence intervals initially. Still, with the help of my tutor and additional research, I gradually understood these concepts and learned how to interpret statistical results more effectively.

Producing summary tables, contingency tables, and plots using R was another area where I initially struggled. However, practising these skills regularly and experimenting with different packages and functions helped further develop my skills. As a result, I now feel more confident in creating tables and visualisations that effectively communicate the results of my analyses.

In interpreting statistical findings, I realised that p-values and confidence intervals provide essential information about the precision and uncertainty of estimates. For instance, when the p-value is low, it suggests that the observed effect is unlikely to have occurred by chance. On the other hand, confidence intervals provide a range of values within which the proper population parameter is likely to lie. Combining these two approaches provides a more robust and comprehensive understanding of the statistical results.

**Challenges and Confident**

Honestly, I found the activities in the module challenging and confusing at the beginning but also rewarding. Working with real-world datasets from the NOS the Health Survey for England, 2011 and applying statistical concepts to real-world problems like alcohol drinking and its impact on hospital admission and death rate in the UK was exciting. It helped me appreciate data analysis's relevance and importance. However, I experienced frustration when my code on the R terminal did not work and difficulty interpreting the results. To overcome these challenges, I had to review the code of others and watch the seminar recording repeatedly to enhance my knowledge and skills. But my confidence in using R for statistical data analysis has significantly increased since starting this module. Initially, I was intimidated by R, but now I feel comfortable navigating R and using various functions to perform statistical analyses. I have also learned the importance of organising my code and comments to make reviewing and replicating the studies easier. Additionally, I am confident in interpreting statistical results, producing summary and contingency tables, and creating plots using R.

And again, One key aspect of data analysis that I have developed a deeper understanding of is the interpretation of statistical findings using p-values and confidence intervals. I have learned that p-values indicate the probability of observing the data or more extreme data if the null hypothesis is true. The smaller the p-value, the more evidence we have against the null hypothesis. Similarly, confidence intervals estimate the range of values within which the correct population parameter is likely to fall. I now better understand how to interpret and communicate these findings to others.

In producing summary tables, contingency tables, and plots using R, I have gained much practical experience through the assignments. For example, creating summary tables allowed me to summarise large datasets and identify trends and patterns quickly. Contingency tables helped me understand the relationship between several variables, and plots enabled me to visualise the data and communicate my findings more effectively. Also, I have learned how to produce summary tables, contingency tables, and plots using R. I can create tables and plots to summarise data, visualise trends, and identify patterns. In addition, I have learned how to use ggplot2, a popular data visualisation package, to create high-quality graphs and charts. I can also create contingency tables and perform statistical tests to assess the relationship between categorical variables.

**Next Steps**

Overall, this module has been a significant learning experience. I am more equipped as I have learnt valuable concepts on drawing insights from data that will be useful as I proceed in my artificial intelligence program. In addition, I have gained practical data analysis skills and statistical knowledge to apply to real-world data analysis projects. Despite my challenges, I am determined to continue practising and refining these skills by learning deep essentials for data analysis to stay up-to-date with the latest techniques and best practices in data analysis.

In conclusion, I thank my module instructor for his guidance and support throughout this module. It is excellent learning to become an artificial intelligence expert through this module's building blocks of data analysis. I am excited to apply what I have learned in this module to future data analysis projects.

References  
  
1. Rolfe, G., Freshwater, D. & Jasper, M. (2001) Critical reflection in nursing and the helping professions: a user’s guide. Basingstoke: Palgrave Macmillan

2. Robin, R., et al. (2005) Alcohol and public health. Available from:

https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)17870-2/fulltext [Accessed 10 April. 2023]

3. Louis, J., et al. (2005) Alcohol use and mental health during COVID-19 lockdown: A cross-sectional study in a sample of UK adults. Available from:

https://www.sciencedirect.com/science/article/pii/S0376871620306530[Accessed 10 April. 2023]

4. NHS. (2019) Statistics on Alcohol, England 2019 [PAS]. Available from:

https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-alcohol/2019[Accessed 10 April. 2023]

5. Office for National Statistic. (2018) Adult drinking habits in Great Britain: 2017. Available from:

https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/drugusealcoholandsmoking/bulletins/opinionsandlifestylesurveyadultdrinkinghabitsingreatbritain/2017[Accessed 10 April. 2023]

6. Thomas, B. (2010) Alcohol: No Ordinary Commodity: Research and Public Policy . 2nd ed. USA: Oxford University Press