Submission-Week-11

Hariz Emran

2023-11-03

Week-9: Webpage + Diary Entry

What is the topic that you have finalised?

I have finalised on the topic of drug abuse and rehabilitation trends in Singapore over the years.

What are the data sources that you have curated so far?

I have curated some data sets from https://beta.data.gov.sg/, Singapore's national open data collection site. The data sets that I have collated thus far are namely the demographics of inhalant abusers arrested and the inmates undergoing rehabilitation in the Drug Rehabilitation Centre (DRC) as well as the quantity and type of controlled drugs that have been seized.

Week-10: Diary Entry

What is the question that you are going to answer?

How effective have current efforts been in addressing the problem of drug abuse in Singapore and how can they be improved?

Why is this an important question?

Drug abuse and addiction can have serious adverse effects on an individual's life. Tackling drug abuse requires a multifaceted approach given the nuanced factors involved. Recognising current trends related to drug abuse could prove crucial in designing more targeted and effective policies to resolve the issue in the long-term.

Which rows and columns of the dataset will be used to answer this question?

Multiple variables of the different collated data sets will be used, namely the quantities and types of drugs that have been seized over the years, the demographics of the inhalant abusers arrested by age, ethnicity, gender and status over the years as well as the demographics of inmates in the DRC by age, education level and gender. A challenge that I faced was the inconsistency of some variables such as the unit of measurement (kilograms, grams, tablets) in the data set on quantities and types of drugs seized. To overcome this, I separated the drugs that had their quantities in tablets while converting the drugs that had their quantities in grams to be reflected as kilograms.

Week 11: Diary Entry + Placeholders Integrated with Webpage

List the visualisations that you are going to use in your project.

I am going to use grouped bar plots to visualise the following data:

- Controlled Drugs Seized by Measurements, i.e. Kilograms, Grams, Tablets

- Inhalant Abusers Arrested by Status, Age, Gender & Ethnicity

- Population of Inmates in the DRC by Age (Before & After 2020), Gender & Education Level

How do you plan to make it interactive?

I plan to make the visualisations interactive by using "plotly" so that users can zoom in and explore the specific values of each bar in all the grouped bar plots. I also plan to use "shiny" so that users can choose which bar chart to look at on the specific page that they are on.

What concepts incorporated in your project were taught in the course and which ones were self-learnt?

Week 3 - Explicit Coercion of Variable Types

Week 4 - Selection & Creation of Columns; Slicing of Rows

Week 7 - Customisation of Bar Plots

Week 8 - "shiny"

Self-Learnt - Renaming of variables and values; "plotly"; Grouped Bar Plots

While I had converted the drugs that had their quantities in grams to be reflected as kilograms earlier, I decided to revert this so as to present the drugs which had their original quantities in grams on their own plot. I realised that comparing them directly with the drugs that had their original quantities in kilograms would indirectly reduce their significance due to their extremely small values by comparison. Another hurdle I faced was the inability to generate proper and accurate plots, for which I later realised the issue was due to the variables of the original data sets being of the incorrect type, e.g. character or double, and I utilised explicit coercion to address this.