W13-submission

Janelle Tan

2023-11-21

W13 submission

A.Diary Entry

1. What is the topic that you have finalised? (Answer in 1 or 2 sentences).

```
# I have decided on the topic of "health" for my data story, namely, Alcohol Abuse and its Publi
c Health Consequences.
```

2. What are the data sources that you have curated so far? (Answer 1 or 2 sentences).

```
# My data sources for now are mainly coming from the Global Health Observatory data repository.

# Some data sets I'm considering are regional prevalence of alcohol dependence in 2016 (%), Regi
onal prevalence, AAFs (15+), liver cirrhosis deaths (%), Alcohol, recorded per capita (15+) cons
umption (in litres of pure alcohol).
```

3. What is the question that you are going to answer? (Answer: One sentence that ends with a question mark that could act like the title of your data story)

```
# "What is the extent of alcohol abuse and its public health consequences in the Americas?" (Not
e: question is still subject to change)
```

4. Why is this an important question? (Answer: 3 sentences, each of which has some evidence, e.g., “According to the United Nations…” to justify why the question you have chosen is important)

```
# Alcohol abuse in the Americas is a significant public health concern with severe consequences. This is evident from the fact that overindulgence in alcohol caused over 140, 000 fatalities annually in the United States alone during 2015-2019—that is, over 380 deaths daily (CDC, 2022). And this isn't even taking into account the other countries in the region. As such, alcohol abuse echoes the WHO's global perspective on general substance abuse, whereby it contributes to a considerable burden on healthcare systems, law enforcement, and productivity loss. This question is hence important, as understanding the extent of alcohol abuse in the Americas is crucial for shaping targeted public health policies and interventions. This aligns with the global imperative to alleviate the negative impacts of substance abuse.
```

5. Which rows and columns of the dataset will be used to answer this question? (Answer: Actual names of the variables in the dataset that you plan to use). Include the challenges and errors

that you faced and how you overcame them.

```
# Note: the below are still subject to change

# Data Sets: Regional prevalence of alcohol dependence & liver cirrhosis, Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol)


# Specific Variables for Regional prevalence of alcohol dependence:
# Columns: WHO region, regional prevalence of alcohol dependence
# Rows: all

# Specific Variables for Regional prevalence of liver cirrhosis:
# Columns: WHO region, regional prevalence of liver cirrhosis
# Rows: all

# Specific Variables for Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol):
# Columns: Regions, BeverageType, Year 2018
# Rows: all


# Challenges and Errors:
# So far, I've encountered some challenges trying to narrow down the scope of my question. It was also difficult trying to find suitable data sets for the scope of my question amidst the sea of data available.

# To solve them, I narrowed down the scope of my question based on what data was available. I learnt how to quickly filter out potential data sources by looking at its size and whether or not its variables were relevant to my question.
```

6. List the visualizations that you are going to use in your project (Answer: What are the variables that you are going to plot? How will it answer your larger question?)

```
# Bar graphs.

#Variables include WHO regions, regional prevalence of alcohol dependence (%) and liver cirrhosis (%) both in 2016.
#For the data set on Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol), variables include: Regions, BeverageType, Year 2018

# The bar graphs show how the alcohol scene and consequences of alcohol abuse in the Americas are the one of the worst among the other regions, highlighting the extent of substance abuse there.
```

7. How do you plan to make it interactive? (Answer: features of ggplot2/shiny/markdown do you plan to use to make the story interactive)

```
# I'm currently experimenting with plotly and ggiraph packages to make my bar graphs interactiv
e. I'm also contemplating using shiny apps.
```

8. What concepts incorporated in your project were taught in the course and which ones were self-learnt? (Answer: Create a table with topics in one column and Weeks in the other to indicate which concept taught in which week is being used. Leave the entry of the Week column empty for self-learnt concepts)

```
library(tibble)

# Create a tibble with columns for topics, weeks, and self-learned
course_concepts <- tibble(
  Topic = c("Visualising data using ggplot", "Making interactive graphs: plotly & ggiraph", "inp
utting the Github logo", "tilting labels of x-axis by 90-degrees for better viewing of labels"),
  Week = c(7, NA, NA, NA),
)

# Print the tibble
print(course_concepts)
```

```
## # A tibble: 4 × 2
##   Topic                                                          Week
##   <chr>                                                         <dbl>
## 1 Visualising data using ggplot                                     7
## 2 Making interactive graphs: plotly & ggiraph                      NA
## 3 inputting the Github logo                                        NA
## 4 tilting labels of x-axis by 90-degrees for better viewing of labels    NA
```

9. Challenges and errors that you faced and how you overcame them.

```
# I had trouble making my bar graph interactive with plotly.

# I overcame them by referring to websites like stack overflow for solutions and trying out diff
erent ways to make my plots interactive, like with ggiraph and shiny apps. I am currently still
experimenting with these methods and looking into other possible methods.
```

10. Challenges and errors that you faced and how you overcame them (Week 12).

```
# I was looking over my question again and realised that I wanted to make my question and answer more meaningful if possible. I started looking for ways to do so, though along the way, I strugg led with coming up with suitable ideas.

#I overcame this challenge by looking at available data sets to see if there was anything I coul d use to achieve my goal. I am currently still on the look out for more ideas, though I have som e rough plans, including adding in data sets that could help me further assess the impacts of al cohol abuse (eg looking at the different types of health consequences that result from alcohol a buse).
```

11. Final Write-Up (Week 13).

 # The general theme of my data story is "Health", namely, Alcohol Abuse and its Public Health Co nsequences.

 # The question that I had decided upon was "What is the extent of alcohol abuse and its public h ealth consequences in the Americas?".

 # Alcohol abuse in the Americas is a significant public health concern with severe consequences. This is evident from the fact that overindulgence in alcohol caused over 140, 000 fatalities ann ually in the United States alone during 2015-2019—that is, over 380 deaths daily (CDC, 2022). An d this isn't even taking into account the other countries in the region. As such, alcohol abuse echoes the WHO's global perspective on general substance abuse, whereby it contributes to a cons iderable burden on healthcare systems, law enforcement, and productivity loss. This question is hence important, as understanding the extent of alcohol abuse in the Americas is crucial for sha ping targeted public health policies and interventions. This aligns with the global imperative t o alleviate the negative impacts of substance abuse.

 # The data sources that I had curated include: Recorded-alcohol-per-capita-consumption-Americas (GHO | By Category | Recorded Alcohol per Capita Consumption, from 2010 - Updated May 2021, n. d.), Regional-prevalence-of-alcohol-dependence (GHO | By Category | Regional Prevalence of Alcoh ol Dependence (%) - by WHO Region, n.d.) and regional-prevalence-of-liver-cirrhosis-deaths (GHO | By Category | Regional Prevalence, AAFs (15+), Liver Cirrhosis Deaths - by WHO Region, n.d.).

 # Recorded alcohol per capita consumption can help answer the first portion of my question, beca use it gives us deeper insights into the alcohol scene in the region of the Americas, helping us have a better picture of the extent of substance abuse there. Understanding the severity of the issue helps us better appreciate the need to address it.

 # Regional prevalence of alcohol dependence and liver cirrhosis can help to answer the second po rtion of the question because it underscores the consequences of alcohol abuse on the public. Th is helps us to better visualise the negative impacts of alcohol on its consumers in the America s, which further highlights the need to address this ongoing issue.

 # The first plot on the Recorded alcohol per capita consumption in the Americas shows how alcoho l consumption over there is actually rather high, especially when contrasted against the global average consumption of alcohol. This information can be inferred from how the points for "All ty pes" of beverages for quite a few countries are depicted to be above the average figure of 6.17 litres per person. The interactivity of the graph also allows users to see the exact consumptio n levels of each country within the region of the Americas.

 # The second and third bar plots on the Regional prevalence of alcohol dependence and liver cirr hosis show how the prevalence of such conditions is one of the highest in the world. This is don e by contrasting the prevalence of such health impacts in the Americas against those of the othe r regions in the world in the bar plots, with America's prevalence being depicted as the mode (s).

 # This entire project was implemented with the help of MUCH patience, both on the part of the mo dule's Teaching Assistants, the professor, and me. I definitely could not have done this on my o wn, so it is worth noting that I had much help from the TAs and the Professor, as well as the be loved "Google" and "ChatGPT".

 # This project was also implemented using Quarto and concepts taught in the lectures, especially the lectures focusing on data visualisations with ggplot.

 # Some new concepts learned included the usage and implementation of "plotly" to make my graphs interactive. I also learned how to tilt the labels of the x-axis at a 90-degree angle for better viewing of labels, and even how to input a cute logo for Github on my webpage to access my actua l Github!

 # References

 # GHO | By category | Recorded alcohol per capita consumption, from 2010—Updated May 2021. (n. d.). WHO. Retrieved 20 November 2023, from https://apps.who.int/gho/data/node.main.A1039?lang=en

 # GHO | By category | Regional prevalence of alcohol dependence (%)—By WHO region. (n.d.). WHO. Retrieved 20 November 2023, from https://apps.who.int/gho/data/node.main.A1071?lang=en

 # GHO | By category | Regional prevalence, AAFs (15+), liver cirrhosis deaths—By WHO region. (n. d.). WHO. Retrieved 20 November 2023, from https://apps.who.int/gho/data/node.main.A1080?lang=en