Document-Entries

Tan Ee Xuan

2023-10-18

Week 9

Answer the following questions in an R Markdown file,

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

The topic I have chosen to finalise on would be on deforestation in Brazil and how agriculture and livestock produce contributed to deforestation.

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

I have decided to use the datasets from Our World in Data and from Climate Watch Data. To be more specific, the datasets I intend to use would be "sugarcane.csv", "soybean.csv", "maize.csv", "cattle.csv", "pigmeat.csv", and "poultry.csv" from Our World in Data and "historical_emissions.csv" from Climate Watch Data.

Week 10

(1) 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)

How has Agriculture and Livestock Production Contributed to Deforestation in Brazil?

(2) 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)

According to the United Nations (n.d.), the process of climate change is accelerating as a result of greenhouse gas emissions that humans produce and are responsible for. The effects of climate change will have a devastating impact on countries and communities all around the world, especially for the poor who may depend on the environment for their livelihoods (MercyCorps, 2021). Brazil is one of the largest emitters of greenhouse gasses due to "deforestation, agriculture and other land-use" (Gratten, 2022).

(3) 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).

Our World in Data

I have decided to use the datasets from Our World in Data and from Climate Watch Data. To be more specific, the datasets I intend to use would be "sugarcane.csv", "soybean.csv", "maize.csv", "cattle.csv", "pigmeat.csv", and "poultry.csv" from Our World in Data.

For the "sugarcane.csv", "soybean.csv", "maize.csv", "cattle.csv", "pigmeat.csv", and "poultry.csv" datasets, I have filtered the column "Country" for all the rows containing "Brazil" and I filtered the column "Year" such that it only shows rows with "2011", "2012", …, "2019", and "2020" from its respective raw datasets that I obtained from Our World in Data.

For each of the "sugarcane.csv", "soybean.csv", and "maize.csv" datasets, I will use the columns "year", "population", "production_kg", "production_kgpercapita", "landuse_ha", "animalfeed_kg", "animalfeed_kgpercapita".

For each of the "cattle.csv", "pigmeat.csv", and "poultry.csv" datasets, I will use the columns "year", "population", "production_t", "production_kgpercapita", "producing_or_slaughtered_animals", "producing_or_slaughtered_animals_percapita".

Climate Watch Data

For the "historical_emissions.csv" dataset, I filtered the column "Gas" for all the rows containing "All GHG". I also filtered the column "Country" for all the rows containing "BRA" which stands for my country of focus, Brazil. I will only look at the years from "2011", "2012", …, "2019", and "2020". I will use the columns "year", "agriculture", "land_use_change_and_forestry", and "total."

Include the challenges and errors that you faced and how you overcame them.

I struggled to find specific datasets that would provide me with the information that I need to answer my question. Some of the datasets available only helps answer part of the question. I also realised that I had to come up with specific indicators or variables that I want to focus on such that I would not be overwhelmed with the large amount of data available.

Week 11

(1) 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?)

There will be different graphs shown for each relationship I want to observe. I will be able to clearly visualise and make comments on how certain trends have influenced each other or suggest what might happen in the future.

Some of the relationships I hope to study would be between the soybean, maize, sugar cane, cattle, poultry, and pigmeat production with respect to time. The land use over time and percentage of crop used for animal feed over time for the soybean, maize, and sugar cane datasets.

(2) 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 will be using shiny and plotly to make the story interactive. There will be a drop-down feature for the user to interact with and learn more about the trends of the variable that they are interested in.

(3) 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)

Concept	Week
Include graphics	Week-1
ggplot	Week-2
Variables	Week-3
Manipulate data (filter, select, group_by, etc)	Week-4
Make a simple function	Week-5

Concept	Week
strsplit	
ifelse	
nchar	
Loops	Week-6
Sys.sleep	
substr	
ggplot graphs (geom_bar, geom_point, geom_violin, facet_wrap, facet_grid)	Week-7
Shiny App	Week-8
tribble	Week-9
pivot_longer, pivot_wider	Week-9
data.frame	
plotly, add_trace	
fluidrow	
wellPanel, tabsetPanel	
toggleImages()	

Include the challenges and errors that you faced and how you overcame them.

I had some difficulty in coming up with ideas on what I could do to make the story interactive. To overcome this I went to search up some information and draw some inspiration from others.