Week 10 Final Project Submission

Wu Xingyi (Kyra)

2023-11-03

Week 11 Updated Topic Choice

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?)

Nutritional Value Comparison: I will create bar charts or radar charts to compare the nutritional content between different menu items

This visualization will help answer the question of how the average price of McDonald's menu items varies across different countries or regions. It can reveal insights into economic disparities, currency exchange rates, and local market factors that influence pricing.

Perhaps, another visualization that showcases the average price of McDonald's menu items across various countries, using a world map color-coded by price ranges to provide a quick overview of pricing disparities.

I also plan to create a visualization that combines Big Mac prices with the number of people per outlet at McDonald's locations to identify the cheapest and most accessible places to eat. To do this, I'm considering two types of visualizations: *Geographical Scatterplot*: I'll use a scatterplot where the x-axis will show the number of people served by each outlet, and the y-axis will display the Big Mac price at that location. Each point's color will indicate the price range or accessibility level; for example, green for affordable and accessible and red for expensive and less accessible outlets. The size of the points could represent the popularity of each outlet.

Heatmap: Another option is to create a heatmap that showcases regions or areas using different shades of color. The intensity of the color will reflect a combination of the Big Mac price and the number of people served per outlet. Darker shades on the map could signify cheaper and more accessible areas, while lighter shades could indicate pricier and less accessible locations.

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

Inputs: Use Shiny's input functions like selectInput(), sliderInput(), and textInput() to allow users to filter or customize the data displayed in the visualizations. For example, create a dropdown menu for selecting regions or a slider for choosing a price range.

Tooltips: Integrate tooltips using Shiny's hoverText() or hoverInfo() to display additional information when users hover over data points in the visualizations. This can include specific price values, number of people served, or other relevant data.

Linked Visualizations: [This one might be harder] Use Shiny's linking feature to establish interactivity between different plots. When a user interacts with one plot (e.g., selecting a region), it can update or filter data in other linked plots, providing a more holistic view of the data.

ggiraph Package: I am also considering using the ggiraph package, which extends ggplot2 to create interactive graphics. This can include interactive bar charts, scatterplots, or other visualizations where users can click on elements for more information.

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)

```r
library(knitr)
concepts_table <- data.frame(
  Topics = c(
    "Data Visualization with ggplot2",
    "Interactive Web Applications with Shiny",
    "Markdown for Document Creation",
    "Data Cleaning and Pre-processing",
    "Exploratory Data Analysis",
    "Data Filtering and Transformation",
    "Statistical Analysis and Plotting",
    "Linking Visualizations in Shiny",
    "Embedding Shiny Widgets in R Markdown"
  ),
  Weeks = c(
    "Week 2 & 7",
    "Week 2",
    "Week 1",
    "Week 4",
    "Week 9 + Self Learned",
    "Week 4 + Self Learned",
    "Week 7",
    "Self Learned",
    "Announcements: Integration of Shiny dashboard with Quarto"
  )
)

kable(concepts_table, format = "markdown")
```

Topics	Weeks
Data Visualization with ggplot2	Week 2 & 7
Interactive Web Applications with Shiny	Week 2
Markdown for Document Creation	Week 1
Data Cleaning and Pre-processing	Week 4
Exploratory Data Analysis	Week 9 + Self Learned
Data Filtering and Transformation	Week 4 + Self Learned
Statistical Analysis and Plotting	Week 7
Linking Visualizations in Shiny	Self Learned
Embedding Shiny Widgets in R Markdown	Announcements: Integration of Shiny dashboard with Quarto

Week 10 Updated Topic Choice

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 should you order next at Mcdonalds? (Depending on Nutritional Value and Price)

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)

This question is important because McDonald's is one of the most popular fast food restaurants in the world, and many people are looking for ways to make healthier choices when eating there. Additionally, many people are on a budget, and want to get the most value for their money.

https://edition.cnn.com/2018/10/03/health/fast-food-consumption-cdc-study/index.html

A study by The Journal of Preventive Medicine and Hygiene (JPMH) found that people who eat fast food more than twice a week are more likely to be overweight or obese. Additionally, fast food is often high in unhealthy fats, sugar, and sodium.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6196377/

Lastly, cost is a crucial factor for many consumers, especially those on tight budgets. Understanding the trade-off between nutrition and price at a popular fast-food restaurant like McDonald's can help individuals and families make more informed decisions when dining out, contributing to financial well-being. Addressing this question can ultimately promote healthier eating habits while keeping affordability in mind. **https: //egrove.olemiss.edu/cgi/viewcontent.cgi?article=1593&context=etd**

Data

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).

Big Mac (Ala Carte) in APAC (Row), Territory (Column) - https://github.com/schmwong/ APAC-McDelivery-Menu-Logger

Category, Item, Calories, Total Fat, Saturated Fat, Trans Fat - https://www.kaggle.com/ datasets/mcdonalds/nutrition-facts

Country, Number of Restaurants, Number of Restaurants per Million Population - https: //openaxis.com/data/3740

Challenges: The sheer volume of data sets can be overwhelming. It becomes difficult to manage and analyze them comprehensively, especially if they are diverse in terms of format, structure, or content.With numerous data sets, I often find myself losing sight of the main narrative or research question. This can lead to a lack of clarity and direction in my analysis, making it challenging to derive meaningful insights. Furthermore, they might not naturally intersect or provide a clear link to build a cohesive narrative. Finding the connections between these sets can be time-consuming and demanding.

Old Topic Choice

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)

Can our world sustain the paradox of abundance and scarcity – where some are overweight, and food is wasted, while others suffer from hunger and environmental pollution?

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)

Each year, the Food and Agriculture Organization (FAO) calculates that approximately **one-third of the food produced for human consumption globally is lost or wasted,** not only squandering the chance to enhance global food security but also to reduce the environmental footprint and resource utilization within food chains. (Link: https://www.fao.org/3/i3347e/i3347e.pdf) The UN reports that global hunger crisis worsened in 2020 due to the COVID-19 pandemic, where the percentage of undernourished people increased to approximately 9.9% in 2020, up from 8.4% the previous year. Considering the statistical margin of error, it is estimated that between 720 and 811 million individuals worldwide experienced hunger in 2020. (Link: https://www.un.org/sustainabledevelopment/hunger/) According to IPCC research, despite there being over 800 million people experiencing undernourishment — 151 million children under the age of five are stunted, 613 million women and girls between the ages of 15 to 49 are afflicted by iron deficiency — there is a staggering 2 billion adults are grappling with issues related to being overweight or obese. (Link: https://www.ipcc.ch/srccl/chapter/chapter-5/)

Data

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).

Per capita kilocalorie supply from all foods per day

This measures the quantity that is available for consumption at the end of the supply chain. It does not account for consumer waste, so the quantity that is actually consumed may be lower than this value. Columns:Country, Year, Population, Food supply (kcal per capita per day) Rows: Continent Rows (Africa, Asia), United States, China, World https://ourworldindata.org/explorers/global-food?time=earliest&facet= none&pickerSort=asc&Food=Total&Metric=Production&Per+Capita=false

Food Loss Index

The Food Loss Index measures the percentage of food lost from the farm level up until retail. It is compared to percentage losses in 2015. Values greater than 100 show increased waste since 2015; lower values indicate

a decrease. Columns Entity Code Year 12.3.1 - Global food loss index - AG_FLS_INDEX Rows: Continent Rows (Africa, Asia), United States, China, World https://ourworldindata.org/grapher/global-food-loss-index

Share of the population that is undernourished

Share of individuals that have a daily food intake that is insufficient to provide the amount of dietary energy required to maintain a normal, active, and healthy life. Columns: Entity, Code, Year, Prevalence of undernourishment (% of population) Rows: Continent Rows (Africa, Asia), United States, China, World https://ourworldindata.org/hunger-and-undernourishment

Global Hunger Index

The Global Hunger Index tries to capture the multidimensional nature of hunger by calculating an index score from four key hunger indicators. It is measured on a 100-point scale where 0 is the best score (no hunger) and 100 is the worst. Columns: Entity, Code, Year, Global Hunger Index Rows: Continent Rows (Africa, Asia), United States, China, World https://ourworldindata.org/grapher/global-hunger-index

Challenges: Combining all the data set into 1 might be rather tough, and also for the first data set, I want to be able to include the codes for the countries also so that is another challenge I have. Combining the data based on their country codes is what I wanted to do as well but I have yet to figure out a way to do it so I will be doing more research into that. Furthermore, trying to incorporate the Shiny applicaion to this has been rather challenging.

Week 9 Updated Topic Choice

What Is The Topic That You Have Finalized?

The topic I have finalized for my project is an in-depth analysis of McDonald's meals, with a specific focus on their nutritional content, pricing, and the geographical distribution of McDonald's stores. This research aims to examine the nutritional value of popular menu items, such as Big Macs, and how their prices vary across different locations, providing insights into the relationship between menu offerings, nutrition, and cost-effectiveness.

What Are The Data Sources That You Have Curated So Far?

Nutrition Facts for McDonald's Menu

https://www.kaggle.com/datasets/mcdonalds/nutrition-facts

Big Mac Prices

https://www.kaggle.com/datasets/vittoriogiatti/bigmacprice/data

APAC-McDelivery-Menu-Logger

https://github.com/schmwong/APAC-McDelivery-Menu-Logger

Data-mcdonalds-nutrition

https://github.com/tsterbak/data-mcdonalds-nutritionfacts/blob/master/workbooks/pffy-data-mcdonalds-nutrition-facts.xlsx

Old Topic Choice

What Is The Topic That You Have Finalized? The topic that I have decided on is the Global Food Paradox. In a world where vast feasts and debilitating famines coexist, the global food paradox challenges us. We're awash in food production yet mired in hunger. The paradox highlights the stark contrast between regions and communities with an overabundance of food, often leading to issues like over consumption and food waste, while other areas suffer from chronic hunger and malnutrition. As I do not want the data to be too infinitely large, I am thinking of utilising continental data. I am also not too sure how much I can include, so I may do away with the environmental impact and animal welfare part if there is too much. Also, I will definitely reduce the data size as I go along the project because I forsee that I would be trying to pick and choose the more relevant ones.

What Are The Data Sources That You Have Curated So Far?

So far, I have managed to get the following data sources which I think are quite interesting to look at. Of course I do not think I will be using everything for the final project but the story I hope to tell would be along the lines of

Food Production

Per capita kilocalorie supply from all foods per day, 1961

https://ourworldindata.org/explorers/global-food?time=earliest&facet=none&pickerSort=asc&Food=Total&Metric=Production&Per+Capita=false

Environmental Impact

Global emissions from food by life-cycle stage, Total GHGs

https://ourworldindata.org/grapher/food-emissions-life-cycle

Freshwater withdrawals per kilogram of food product

https://ourworldindata.org/grapher/water-withdrawals-per-kg-poore

Carbon opportunity costs per kilogram of food

https://ourworldindata.org/grapher/carbon-opportunity-costs-per-kilogram-of-food

Land use per kilogram of food product

https://ourworldindata.org/grapher/land-use-per-kg-poore

Greenhouse gas emissions per kilogram of food product

https://ourworldindata.org/grapher/ghg-per-kg-poore

Scarcity-weighted water use per kilogram of food product

https://ourworldindata.org/grapher/scarcity-water-per-kg-poore

Animal Welfare

Number of land animals slaughtered for meat per year, 1961 to 2021

https://ourworldindata.org/explorers/animal-welfare?facet=none&country=OWID_WRL~USA~CHN~IND~BRA~GBR&Metric=Animals+slaughtered&Animal=All+land+animals+(total)&Per+person=false

Projected habitat loss for all species in 2050: Business-as-usual

https://ourworldindata.org/explorers/habitat-loss

World Hunger

Share of population that cannot afford a healthy diet, 2021

https://ourworldindata.org/food-prices

Number of people that cannot afford a nutrient adequate diet, 2017

https://ourworldindata.org/grapher/number-nutritional-diet-unaffordable

Number of people that cannot afford a calorie sufficient diet, 2017

https://ourworldindata.org/grapher/number-calorie-diet-unaffordable