

Data Visualization Portfolio

Alan Copa

Data Analytics for Finance I & II, 2024-2025



Contents

1	Introduction	3
2	Visualization Tasks	4
2.1	Task 1: Bad/Manipulative Visualization	4
2.1.1	Critique	4
2.2	Task 2: Improved Bad Visualization	6
2.3	Task 3: Good Visualization Critique	7
2.3.1	Critique	8
2.4	Task 4: Climate Change Visualization for Social Media	9
2.4.1	The Ozone Layer: Evidence of Global Environmental Action	9
2.4.2	LinkedIn Post: Healing the Ozone – A Lesson for Climate Change	10
2.5	Task 5: Black-and-White Visualization	11
2.5.1	Correlation between GDP of Nigeria and U.S. Crude Oil Price Over Time	11
2.6	Task 6: Visualization with Color as a Key Aesthetic	12
2.6.1	Global Energy Consumption in 2023	12
2.7	Task 7: Data-Ink Optimization Showcase	13
2.7.1	Energy Balance by Carrier in Switzerland in 2023	13
2.8	Task 8: Uncommon Visualization	14
2.8.1	Average American Expenditure in 2023	14
2.9	Task 9: Visualization by hand	15
2.10	Ionization Energy	15
2.11	Task 10: Visualization with ChatGPT Pro	16
2.12	Number of Free-to-use vs Pay-to-Use ATMs Over Time	16
2.13	Task 11: Data Map	17
2.13.1	Inflation Rate in Europe in 2023	17
2.14	Task 12: Interactive Visualization	18
2.15	Internet Fixed Broadband Subscriptions Over Time (1998-2022)	18
2.16	Task 13: Creation Process of "GDP of Nigeria and U.S. Crude Oil Price Over Time"	21
2.16.1	Sketch	21
2.16.2	Version 1	22
2.16.3	Version 2	23
2.16.4	Version 3	24
2.16.5	Final Version	24
2.17	Task 14: More Visualizations	25
2.17.1	Measures history of Gravitational Constant G	25
2.18	My Favourite Tools	26

3	Bibliography	26
3.1	Data Sources	27
3.2	Software and Tools Used	27
3.3	Generative AI Disclosure	27
4	Appendix	27

1 Introduction

This portfolio presents a collection of data visualizations created as part of the course Data Analytics for Finance I & II for the 2024-2025 academic year.

2 Visualization Tasks

2.1 Task 1: Bad/Manipulative Visualization



Figure 1: Bad Visualization [1]

2.1.1 Critique

This data visualization is misleading because the values in the horizontal bar chart are presented as the absolute number of serial killers per country. The proportion relative to the population of each country is not considered, resulting in an overwhelming scale for the US, which overshadows the other values. The addition of the map without a proper legend adds confusion, as the red circles are not proportional to the data shown. This suggests an alternative perspective but fails to clarify it. Consequently, the presence of two visualization styles is redundant, as it does not provide new insights but instead increases cognitive load.

Finally, the inclusion of a mask illustration is purely decorative and does not contribute any functional value to the visualization.

2.2 Task 2: Improved Bad Visualization

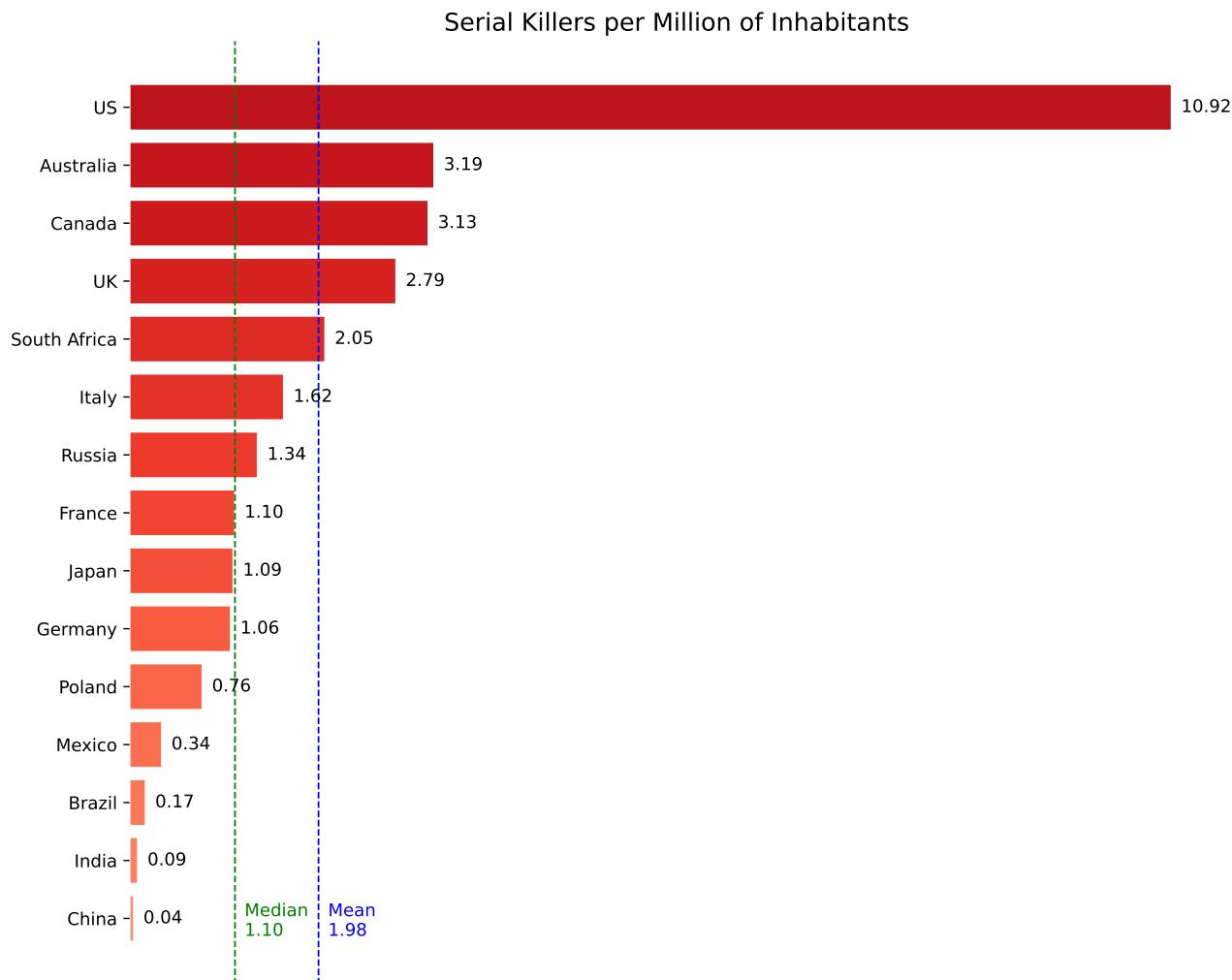


Figure 2: Improved version of the bad visualization, Source[1]

2.3 Task 3: Good Visualization Critique



Figure 3: Good Visualization, Source [11]

2.3.1 Critique

The visualization is clear at first look. The observer does not need to carefully read any axis or legend. Thanks to the prominent title, the theme of the visualization is easily recognizable, and the recognizable logos make its interpretability straightforward. The colors utilized are associated with the brands and differ from one another, emphasizing the state shapes and enhancing engagement. This visualization is perfect for the general public, as it is simple and colorful. More importantly, it avoids complex data such as numbers or units of measurement, making the message elementary and direct. If it were accompanied by another visualization with more specific data or had an interactive element allowing users to overlay and access additional data, the visualization could be perfect.

2.4 Task 4: Climate Change Visualization for Social Media

2.4.1 The Ozone Layer: Evidence of Global Environmental Action

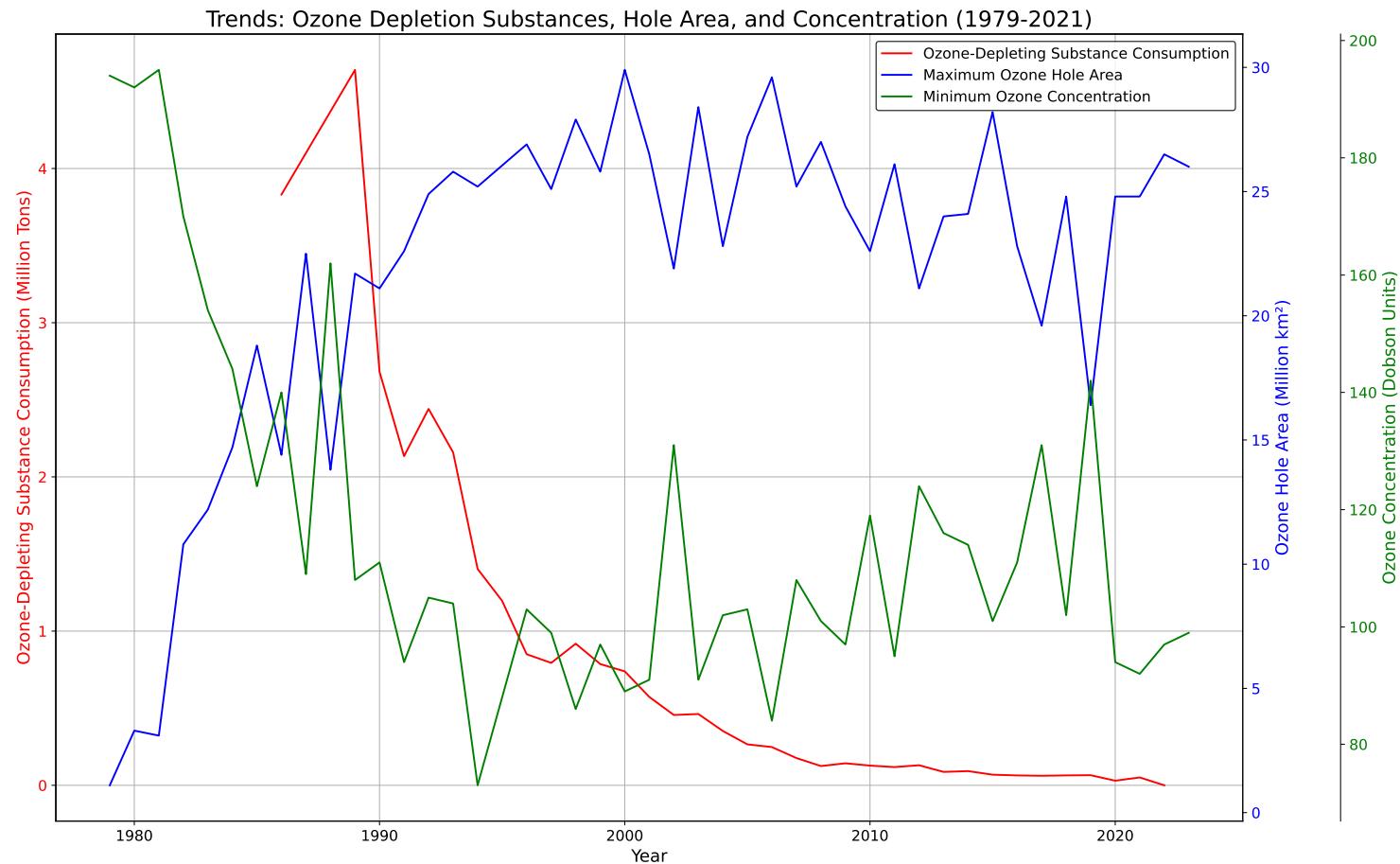


Figure 4: Climate change visualization for social media, Source [8]

2.4.2 LinkedIn Post: Healing the Ozone – A Lesson for Climate Change

What is the ozone layer, and why does it matter?

The ozone layer is the earth's protective shield from the sun, specifically, the sun's dangerous ultraviolet (UV) rays. Without it there would be serious consequences for life on Earth, for example a rise in skin cancer and ecological disruption!

A brief story of the ozone crisis:

Very recently, in the 20th century, a massive hole in the ozone layer appeared primarily over Antarctica. This was caused by chemicals like chlorofluorocarbons (CFCs), more generally called ozone depleting substances. In 1987 the Montreal Protocol brought countries together to combat this dangerous phenomenon. As shown in the graph, after this protocol was introduced, the emissions of ozone harming substances finally declined and the ozone hole stabilized. The Montreal Protocol remained one of the greatest environmental successes, proving that global action can solve even the biggest challenges.

Climate change lessons:

The data in the visualization tells a crucial story! While the ozone layer's destruction was rapid, its healing is significantly slower, and the ozone hole itself persists... This serves as a clear reminder for:

- We alone as a species, we hold the power to both harm and heal environment
- Combating climate crisis like global warming requires unconditional global collaboration
- the healing process of nature is gradual and perhaps irreparable, which makes it even more important we do not make any more mistakes

I call you to action: before the window of opportunity closes, let's take bold actions to address climate change, just as we did with the reduction of ozone depletion substances. We can guarantee a more sustainable and healthy Earth for us and coming generations, if we work together!

2.5 Task 5: Black-and-White Visualization

2.5.1 Correlation between GDP of Nigeria and U.S. Crude Oil Price Over Time

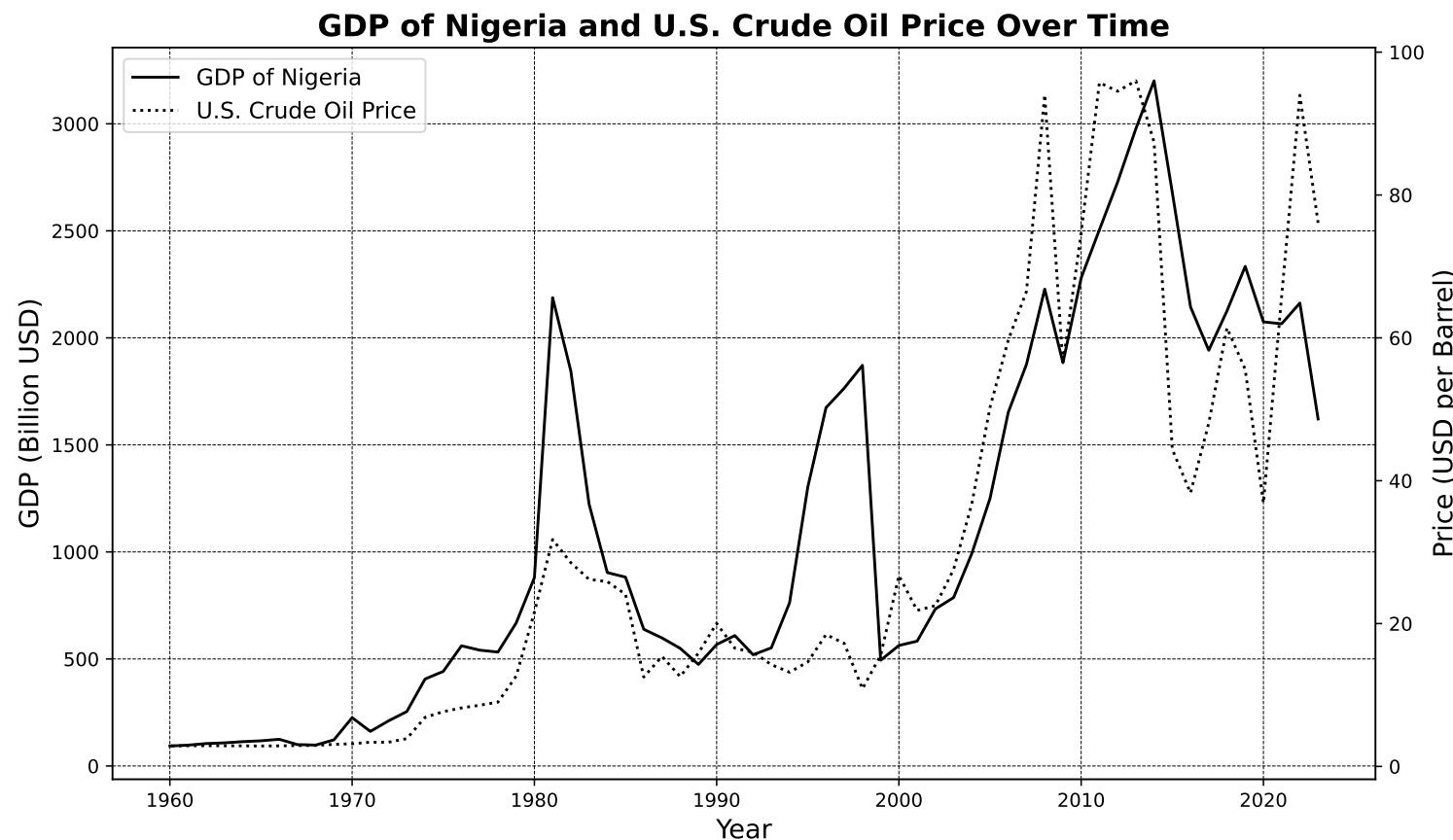


Figure 5: Black-and-white visualization, Sources [3] [2]

2.6 Task 6: Visualization with Color as a Key Aesthetic

2.6.1 Global Energy Consumption in 2023

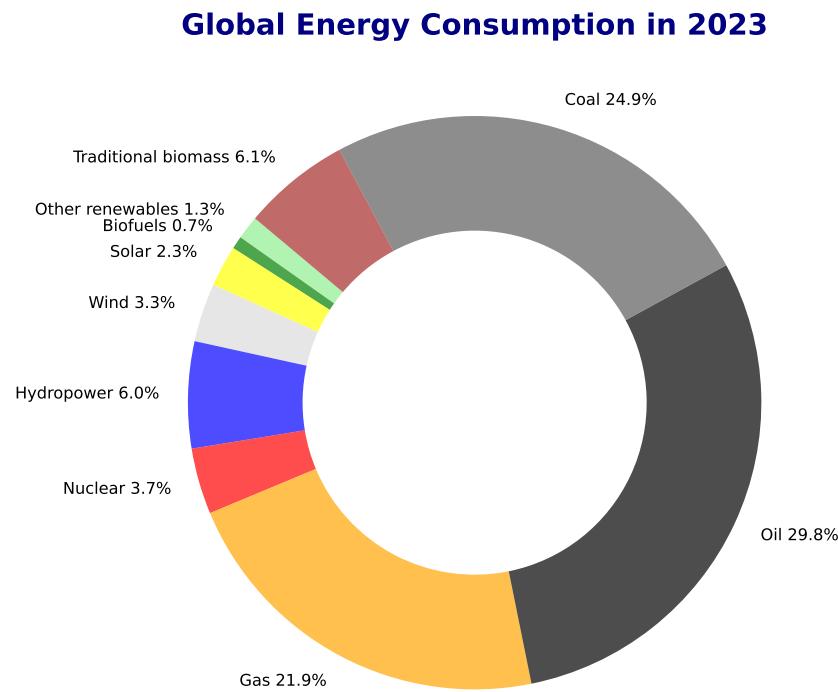


Figure 6: Visualization utilizing color as a key aesthetic, Source [5]

2.7 Task 7: Data-Ink Optimization Showcase

2.7.1 Energy Balance by Carrier in Switzerland in 2023

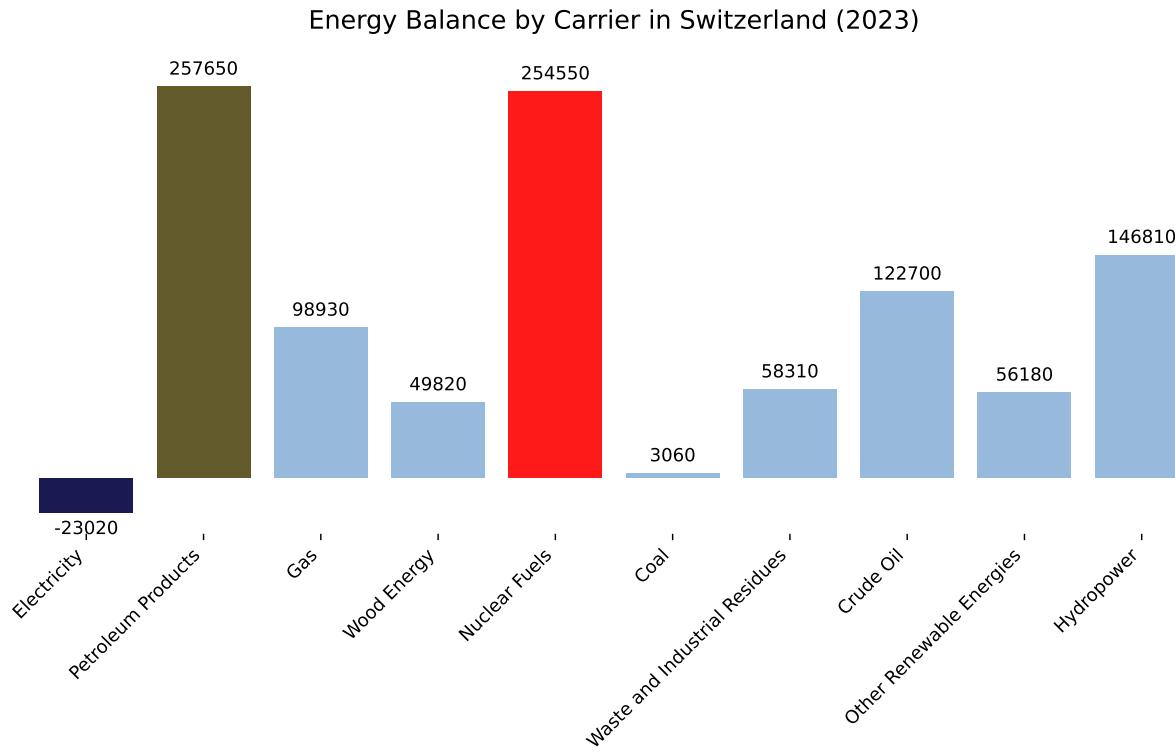


Figure 7: Data-Ink optimization visualization, Source [9]

2.8 Task 8: Uncommon Visualization

2.8.1 Average American Expenditure in 2023

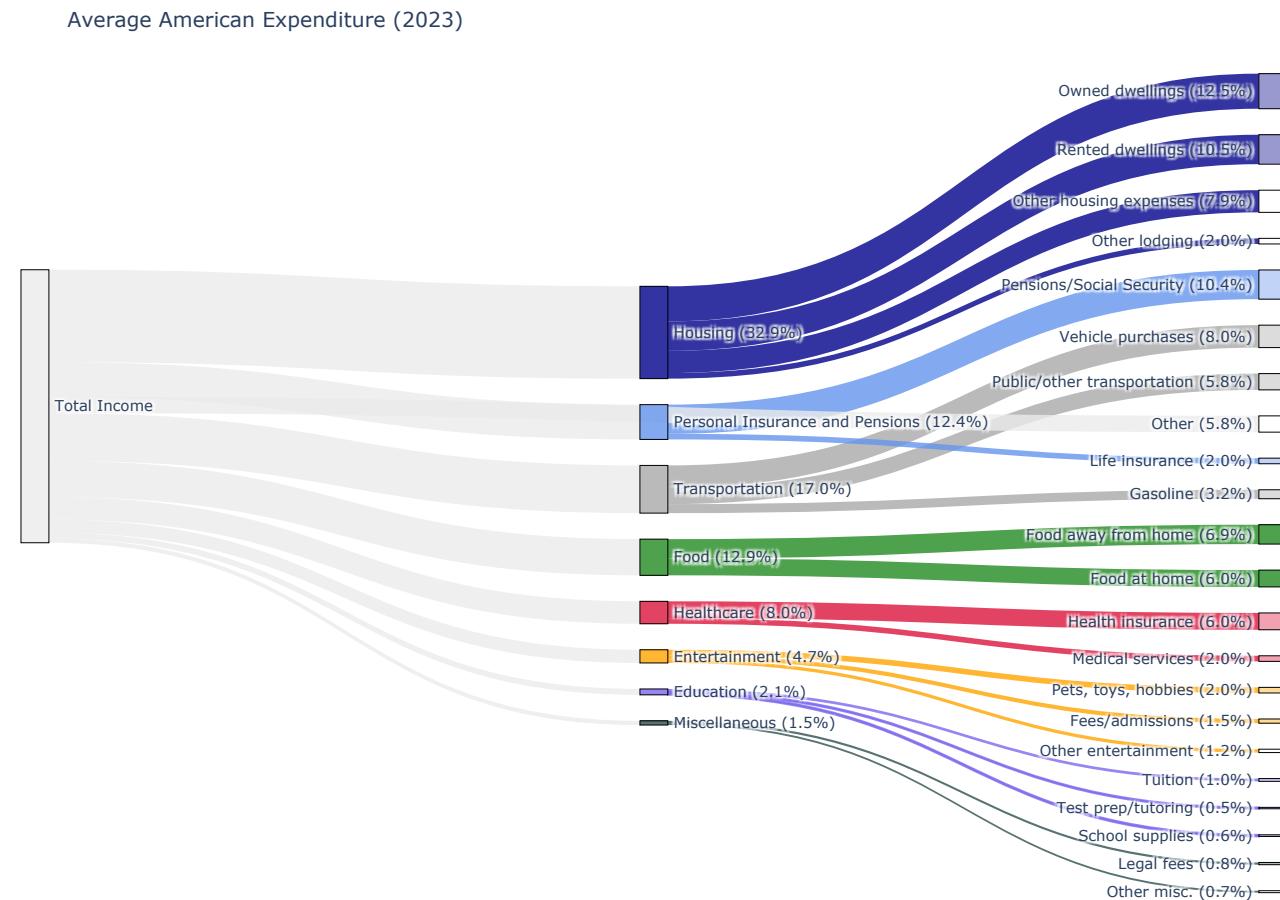


Figure 8: Uncommon visualization, Source [10]

2.9 Task 9: Visualization by hand

2.10 Ionization Energy

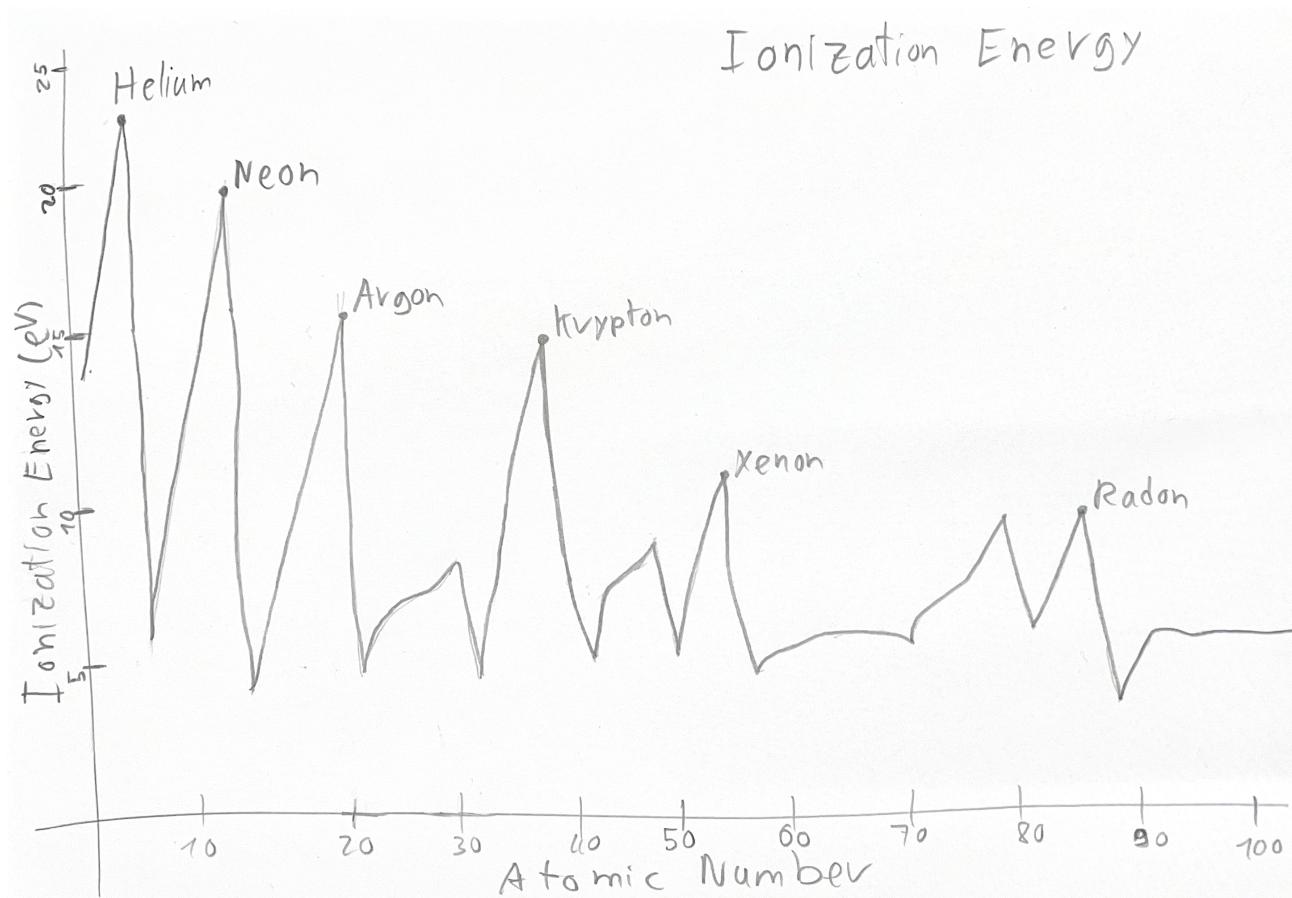


Figure 9: Visualization by hand, Source [12]

2.11 Task 10: Visualization with ChatGPT Pro

2.12 Number of Free-to-use vs Pay-to-Use ATMs Over Time

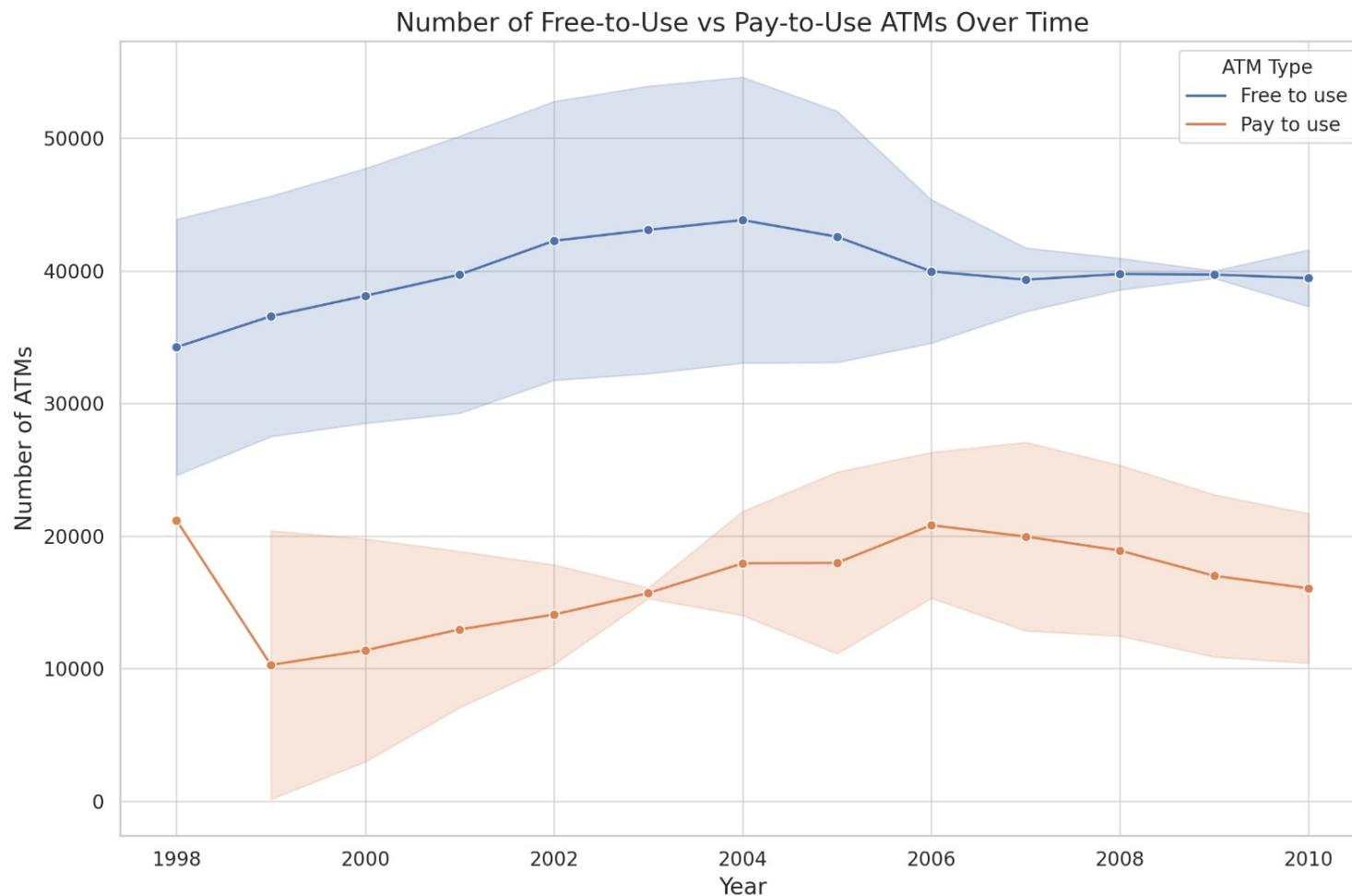


Figure 10: Number of Free-to-use vs Pay-to-Use ATMs Over Time, Source [7]

2.13 Task 11: Data Map

2.13.1 Inflation Rate in Europe in 2023

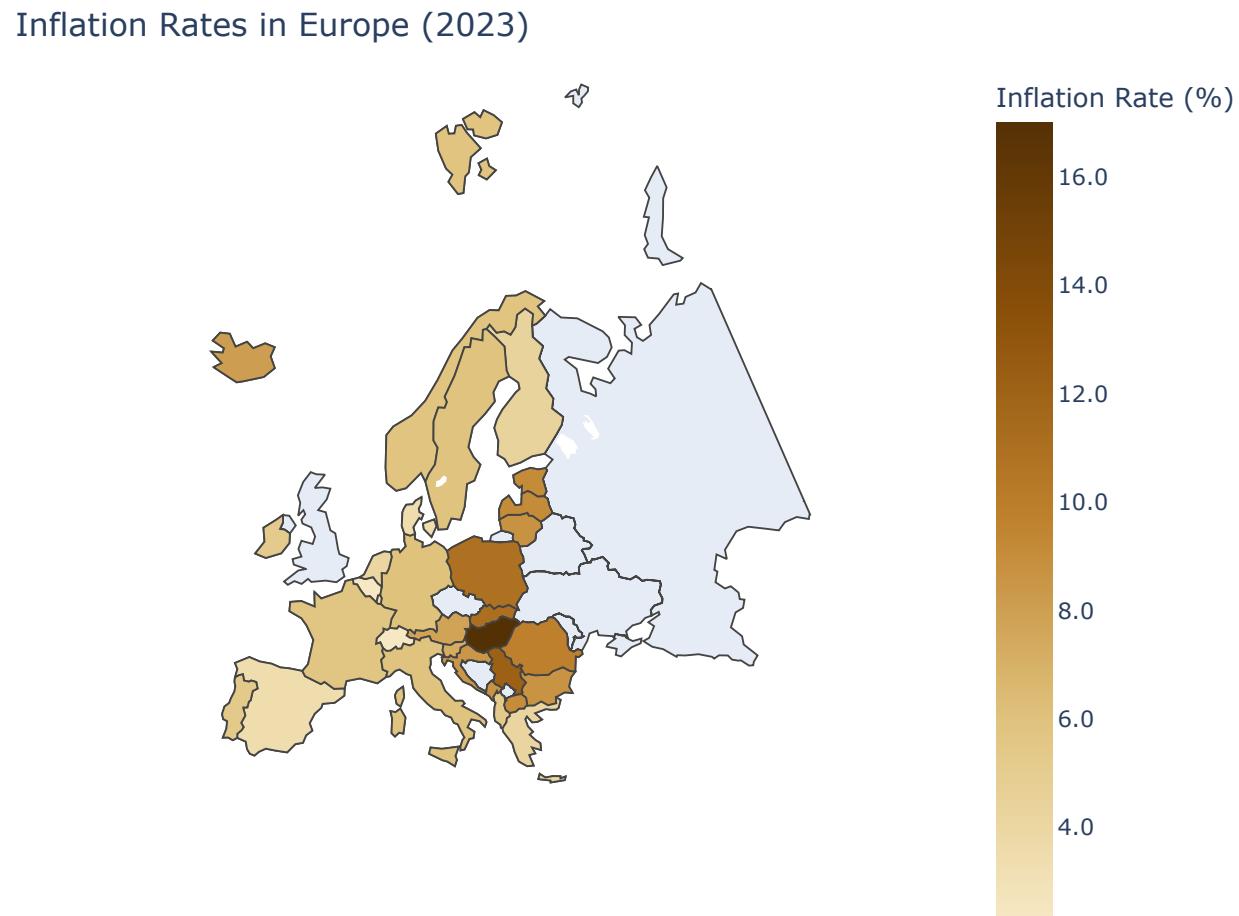


Figure 11: Inflation rate in Europe in 2023, Source [4]

2.14 Task 12: Interactive Visualization

2.15 Internet Fixed Broadband Subscriptions Over Time (1998-2022)

Visit URL to interact!

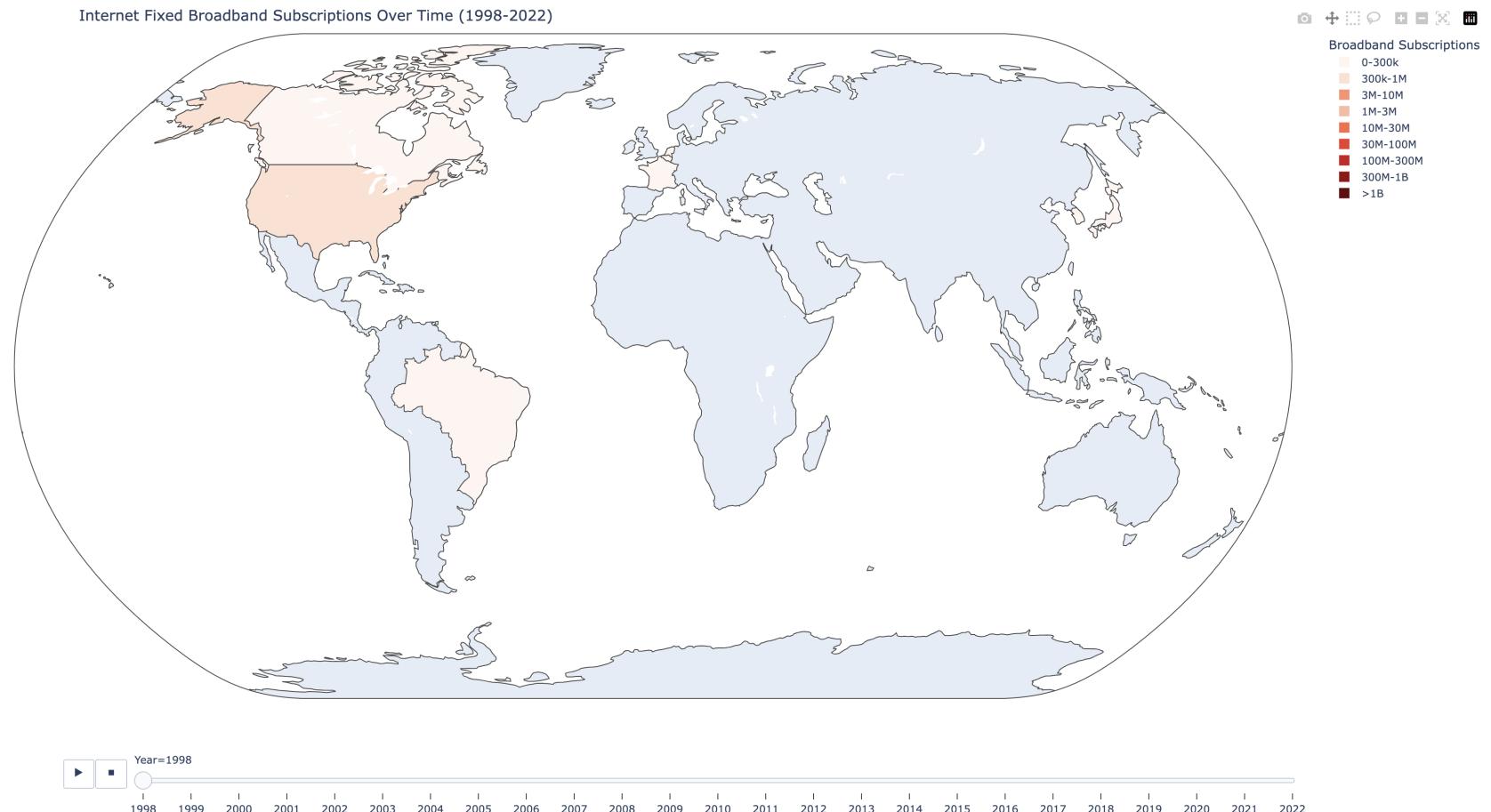


Figure 12: Source [6]

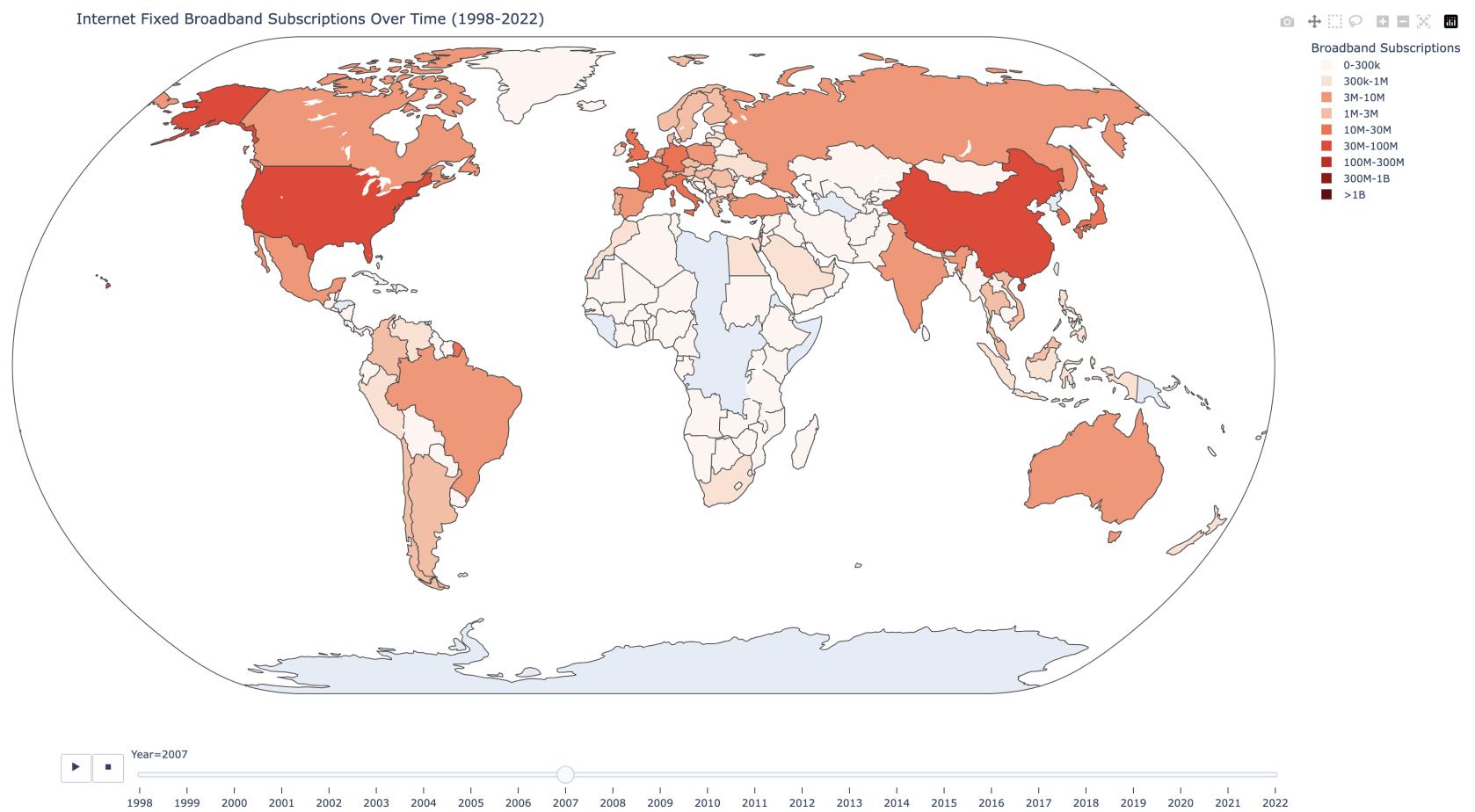


Figure 13: Source [6]

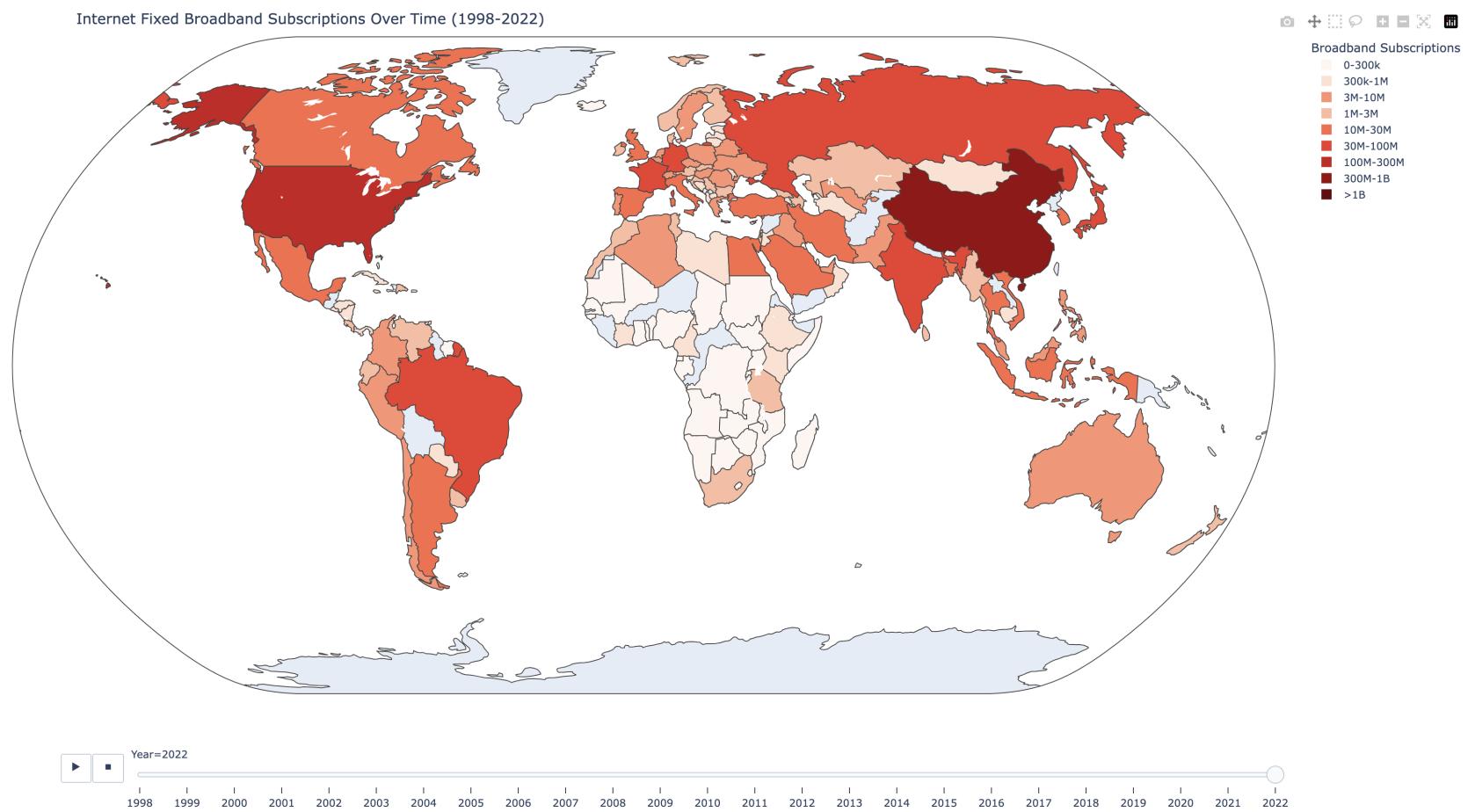


Figure 14: Source [6]

2.16 Task 13: Creation Process of "GDP of Nigeria and U.S. Crude Oil Price Over Time"

2.16.1 Sketch

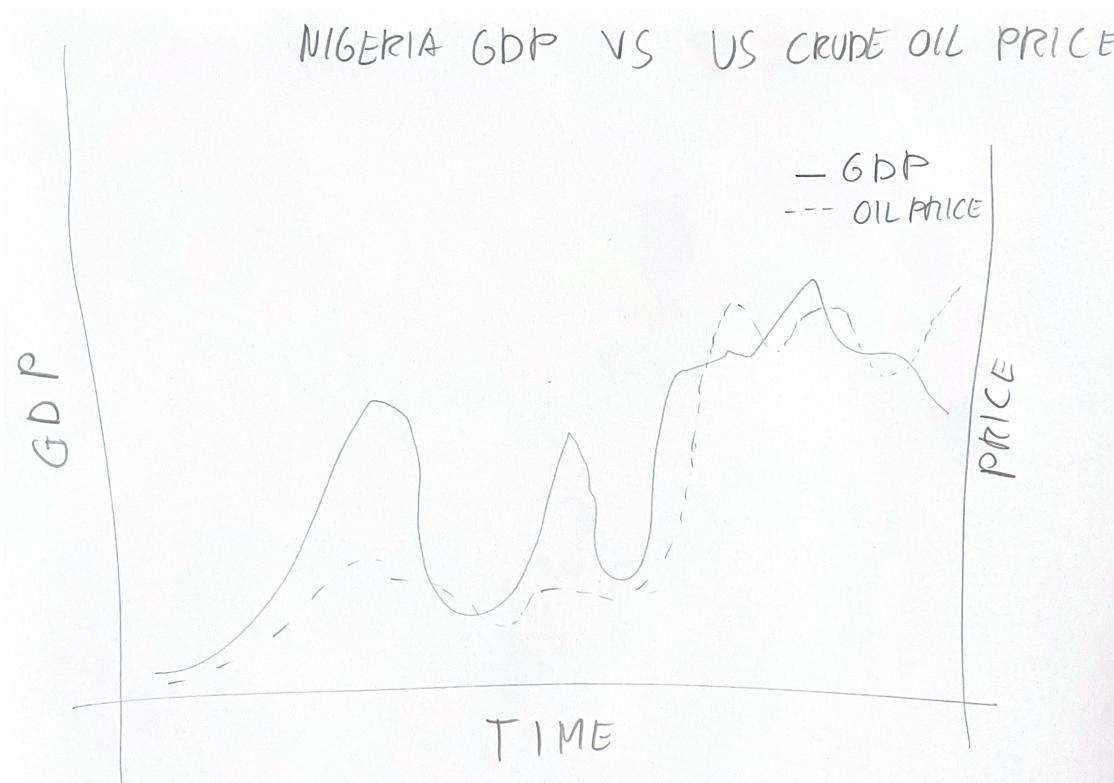


Figure 15: Sketch

My idea was to show the dependency of the Nigerian GDP on the price of the crude oil. The inspiration came from a YouTube video I watched one years ago and I wanted to present the concept in this visualization.

2.16.2 Version 1

I started by searching on the web for the GDP of Nigeria. With assistance of ChatGPT I cleaned the dataset from missing values and typos and plotted it.



Figure 16: Version 1

2.16.3 Version 2

Then I added the crude oil prices after having a little trouble finding a reliable dataset with enough years of data. The result is better than expected, now I have data from the 19th century.

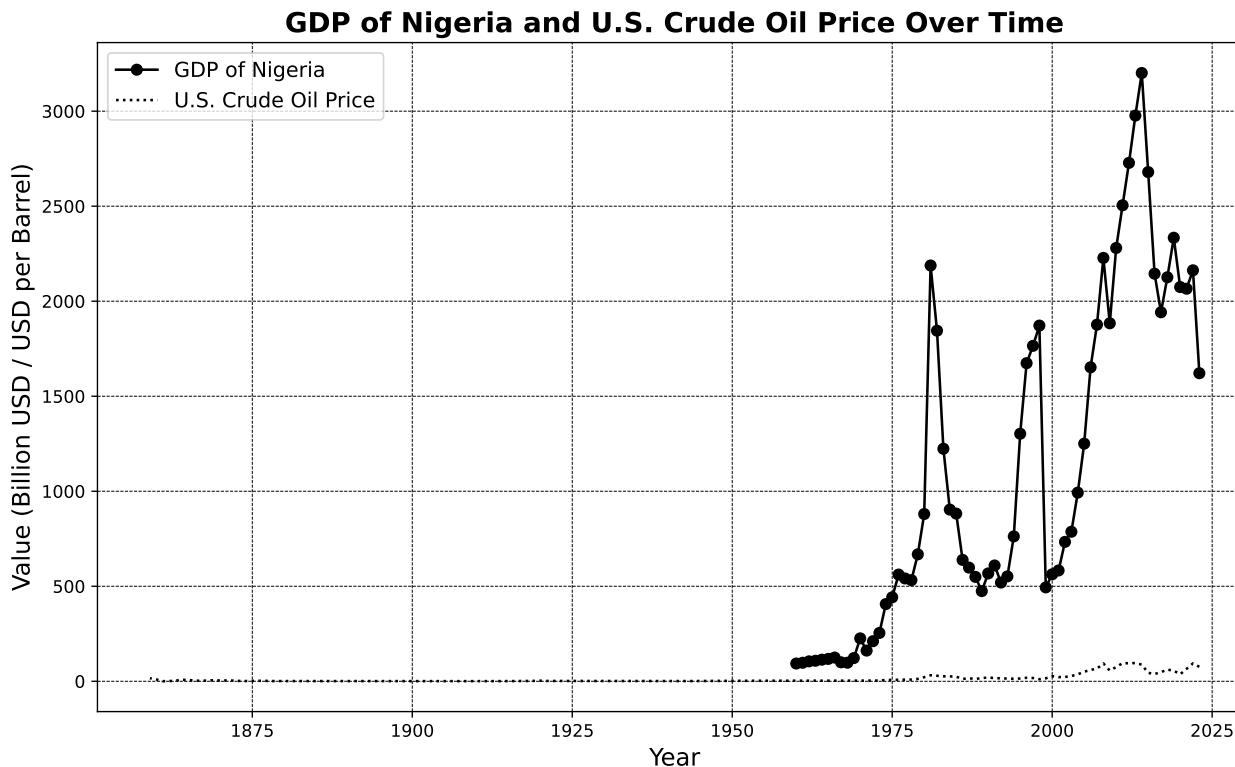


Figure 17: Version 2

2.16.4 Version 3

I limited the time frame to the years where I have data for both lines. I wanted to keep only one scale for axis, thus I played around to fix the scaling of the y-axis, with poor results. I started to realize that I would have been better to sacrifice simplicity to have a better data visualization like figure 4.

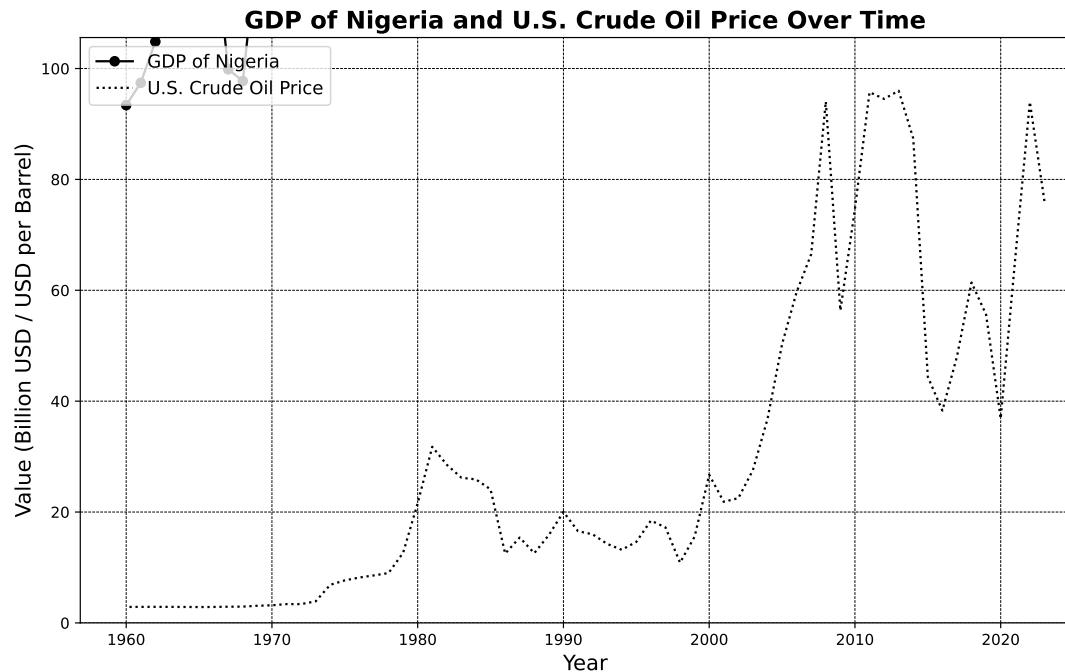


Figure 18

2.16.5 Final Version

For the final version of the visualization I removed the data points to have only lines, added an additional scale on the y-axis on the right of the plot and most importantly fixed the latter in order to have the two lines moving together and not one flat line and another showing the trend. See result at figure 5.

2.17 Task 14: More Visualizations

2.17.1 Measures history of Gravitational Constant G

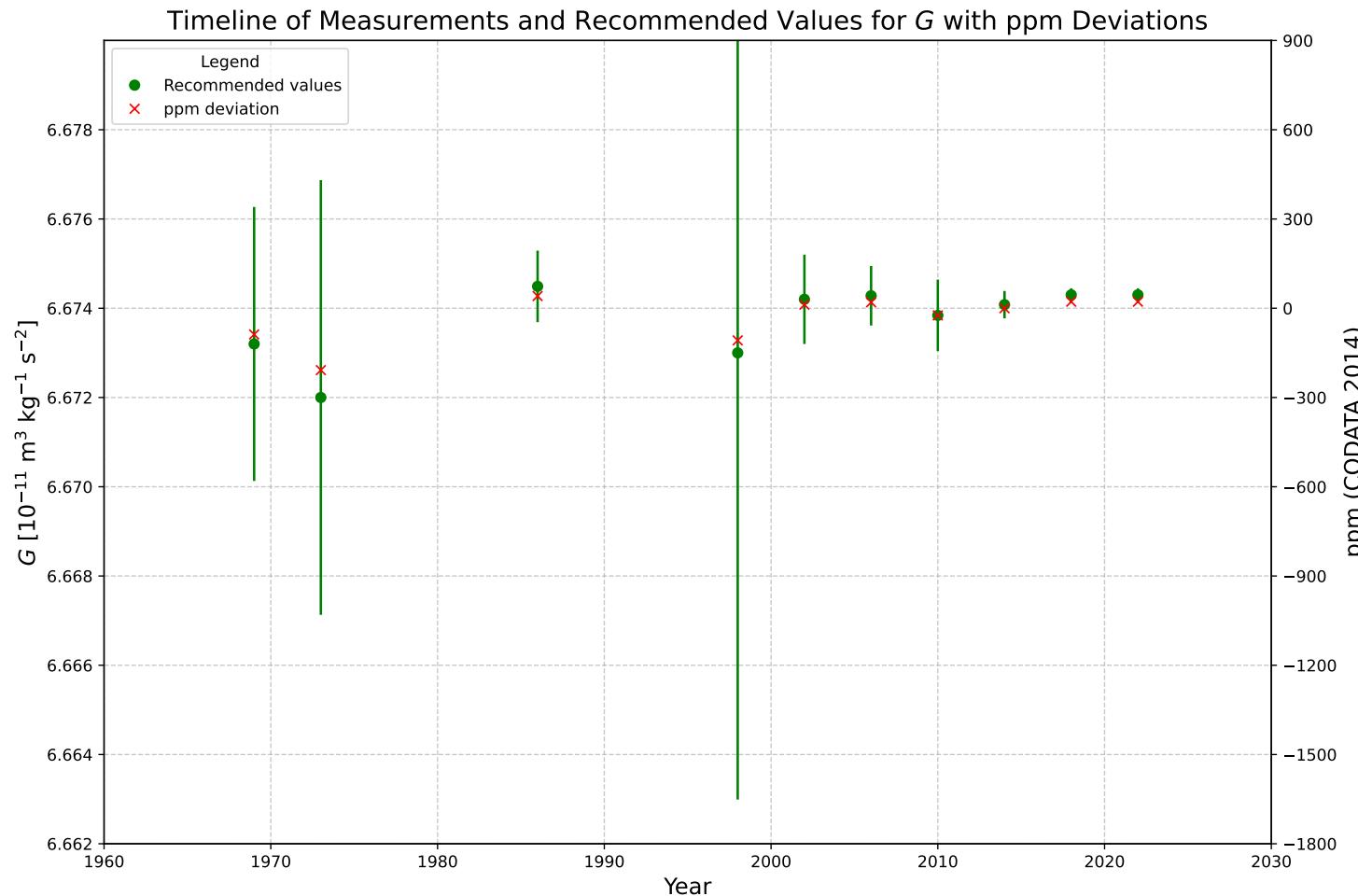


Figure 19: Measures history of Gravitational Constant G, Source [13]

2.18 My Favourite Tools

My favourite tool for creating data visualizations is surely ChatGPT. It helps in the creation of a visualization from A to Z: data cleaning, data interpretation, data plotting and everything feels easier and faster. Quick visualization of an idea helps a lot in the evaluation of the quality of a chart from the beginning, discarding inappropriate type of charts. When the sketch begins to take life, it is straightforward to modify it in order to have a visualization with a personal touch. Clearly the Python language has a role in the simplification of the process, the python packages for data visualization offer a huge variety of charts and plots and are easy to familiarize with. My favourite sources for datasets are the US Government web sites, there is nothing better to ask for: user-friendly, rich in content (data and plot examples) and completely free. The fact that the US makes this amount of scientific/ economic data free of use is incredible. An honorable mention to the NASA website, which is on Google Chrome Bookmark list from now on. Last but not least, netlify.com, the provider where I deployed my interactive visualization through GitHub: simple, fast and reliable... My only regret is to not have a complete interactive portfolio.

3 Bibliography

References

- [1] W. Atlas. Serial killers by country. Available online at World Atlas, 2025. Accessed on 9.1.2025.
- [2] D. Commons. Economic activity in nigeria - gross domestic production. Available online at Data Commons, 2025. Accessed on 16.1.2025.
- [3] U. E. I. A. (EIA). Petroleum consumption historical data. Available online at EIA, 2025. Accessed on 16.1.2025.
- [4] Eurostat. Inflation rates by country. Available online at Eurostat, 2024. Accessed on 12.1.2025.
- [5] O. W. in Data. Energy production and consumption. Available online at Our World in Data, 2025. Accessed on 17.1.2025.
- [6] O. W. in Data. Landline internet subscriptions. Available online at Our World in Data, 2025. Accessed on 9.1.2025.
- [7] LINK. The atm network. Available online at LINK, 2025. Accessed on 18.1.2025.
- [8] NASA. Ozone layer data. Available online at Our World in Data, 2024. Accessed on 9.1.2025.
- [9] S. F. O. of Energy. Energy balance of switzerland. Available online at Swiss Open Data Portal, 2024. Accessed on 17.1.2025.
- [10] B. of Labor Statistics. American expenditure. Available online at BLS, 2025. Accessed on 15.1.2025.
- [11] R. r/dataisbeautiful. The most popular brand on cyber monday by state. Available online at Reddit, 2023. Accessed on 9.1.2025.
- [12] C. Talk. Ionization energy trend. Available online at Chemistry Talk, 2025. Accessed on 17.1.2025.
- [13] Wikipedia. Gravitational constant data. Available online at Wikipedia, 2025. Accessed on 9.1.2025.

3.1 Data Sources

3.2 Software and Tools Used

Python Packages

- matplotlib, numpy, pandas, plotly, dash, os, subprocess

Software

- Visual Studio Code, GPT-4o, My GPTs, netlify.com

3.3 Generative AI Disclosure

As part of the objectives of this portfolio, for every visualization, with the exception of figure 1 and 3, generative AI tools such as ChatGPT have been utilized to assist in their creation. All AI-generated outputs were reviewed and finalized by the author to ensure accuracy and alignment with the intended goals.

4 Appendix

Screenshots of AI Conversations:

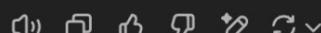


Certainly! Here are some publicly available ATM datasets that you might find interesting:

1. **ATM Cash Demand Forecasting Dataset:** This dataset was used in the NN5 Competition and contains two years of daily cash withdrawals for 111 ATMs in the UK. The primary task associated with this dataset was to forecast ATM cash withdrawals for each ATM for the subsequent 56 days. While the dataset itself isn't directly linked in the provided sources, it is referenced in academic papers, such as "ATM Cash demand forecasting in an Indian Bank with chaos and deep learning." [ARXIV](#)
2. **European Central Bank (ECB) ATM Transactions Data:** The ECB provides datasets related to ATM transactions, including the number of ATM cash withdrawals at terminals acquired by resident Payment Service Providers (PSPs). For instance, the dataset "PTT – ATM, OTC and POS terminal transactions" offers detailed information on such transactions across European countries. [ECB DATA PORTAL](#)
3. **LINK ATM Network Data:** LINK, the UK's largest cash machine network, regularly publishes information about the ATM network, including how people use cash and efforts to support access to cash across the UK. Their datasets cover monthly volumes, ATM numbers, and measures to protect the ATM network. [LINK WEBSITE](#)

These datasets can provide valuable insights into ATM usage patterns, cash withdrawal trends, and the impact of various factors on cash demand. Before using any dataset, ensure you review its licensing and usage terms to comply with any restrictions.

Sources   



↳ **LINK ATM Network Data**: LINK, the UK's largest cash machine network, regularly publishes information about the ATM network, including how people use cash and efforts to support access to cash across the UK. Their datasets cover monthly volumes, ATM numbers, and measures to protect the ATM network

look into this dataset and tell me what interesting visualisations you could do



The LINK ATM Network provides a wealth of data that can be visualized to uncover trends and insights about cash access and usage across the UK. Here are some compelling visualization ideas:

1. Geographical Distribution of ATMs:

- **Heatmaps:** Illustrate the density of ATMs across different regions, highlighting areas with high or low access to cash machines.
- **Choropleth Maps:** Display the number of ATMs per capita in various areas, emphasizing regions that may be underserved.

2. Trends in ATM Usage Over Time:

- **Time Series Analysis:** Plot the monthly or yearly volume of transactions to identify patterns, such as declines in cash usage or seasonal spikes.
- **Impact of External Events:** Analyze how events like the COVID-19 pandemic influenced ATM withdrawals, as noted in LINK's reports. [LINK WEBSITE](#)

3. Comparison of Free vs. Pay-to-Use ATMs:

- **Pie Charts:** Show the proportion of free-to-use versus pay-to-use ATMs within the network.
- **Bar Charts:** Compare the distribution of free and pay-to-use ATMs across different regions or over time.

4. ATM Availability in Relation to Socioeconomic Factors:

- **Overlay Maps:** Combine ATM locations with data on income levels or population density to assess accessibility in various communities.

- **Accessibility Analysis:** Highlight areas where ATM access may be limited, potentially correlating with socioeconomic challenges.

5. Monitoring ATM Network Changes:

- **Animated Maps:** Visualize the opening and closing of ATMs over time to observe how the network evolves.
- **Cluster Analysis:** Identify regions experiencing significant reductions in ATM numbers, which could impact local communities.

6. Consumer Cash Usage Patterns:

- **Histograms:** Depict the distribution of withdrawal amounts to understand common cash usage behaviors.
- **Demographic Analysis:** If data permits, explore how different age groups or professions utilize ATMs.

7. Effectiveness of Financial Inclusion Initiatives:

- **Before-and-After Comparisons:** Assess the impact of initiatives like the installation of new ATMs in previously underserved areas.
- **Accessibility Metrics:** Measure changes in the average distance to the nearest ATM for residents in various regions.

These visualizations can provide valuable insights into the accessibility and usage of cash services, informing decisions by policymakers, financial institutions, and community organizations.

Sources 



↳ **Demographic Analysis**: If data permits,³ explore how different age groups or professions utilize ATMs.



create this graph for me please



To create a graph analyzing ATM usage based on demographics (e.g., age groups or professions), we need a dataset that includes demographic information tied to ATM transactions. If such data is available, it could be uploaded here.

Alternatively, if you're working with a specific dataset and know its structure, please let me know the demographic columns (e.g., age, profession, or similar). With this information, I can create the requested graph.

Feel free to upload or describe the dataset!

you told me you have access to LINK ATM network data so get it and use it please



The LINK ATM Network provides various datasets, including information on ATM locations and monthly cash withdrawal volumes. However, these datasets do not contain demographic details such as age groups or professions of ATM users. Without demographic data, it's not feasible to create a visualization that analyzes ATM usage based on demographics.

If you're interested in other types of analyses or visualizations using the available LINK ATM data, such as:

- **Geographical Distribution of ATMs:** Mapping ATM locations to identify areas with high or low access to cash machines.
- **Trends in ATM Usage Over Time:** Analyzing monthly or yearly cash withdrawal volumes to observe patterns or changes.
- **Comparison of Free vs. Pay-to-Use ATMs:** Examining the distribution and usage of free and pay-to-use ATMs across different regions.¹

Please let me know, and I can assist in creating visualizations based on those aspects.

Sources

↳ **Comparison of Free vs. Pay-to-Use ATMs**: Examining the distribution and usage of free and pay-to-use ATMs across different regions.

yes do this please



To analyze the distribution of free-to-use (FTU) versus pay-to-use (PTU) ATMs across different regions in the UK, we can utilize the data provided by LINK, the UK's main ATM network. LINK regularly publishes information about the ATM network, including the location and type (FTU or PTU) of each ATM.

[LINK WEBSITE](#)

Steps to Create the Visualization:

1. Access the Data:

- Download the latest ATM location dataset from LINK's official website. [LINK WEBSITE](#)
- This dataset includes details such as ATM location, type (FTU or PTU), and other relevant attributes.

2. Data Preparation:

- Load the dataset into a data analysis tool or programming environment (e.g., Python, R).
- Clean the data by handling missing values and filtering out irrelevant information.
- Aggregate the number of FTU and PTU ATMs by region (e.g., counties, cities, or postal codes).

3. Visualization:

32

- Use a visualization library (e.g., Matplotlib, Seaborn, or Plotly in Python) to create a bar chart or stacked bar chart.

ATMs.

- Alternatively, a choropleth map can be created to geographically represent the density and type of ATMs across regions.

Example Code Snippet in Python:

```
python

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
df = pd.read_csv('ATM_Locations.csv')

# Data cleaning and preparation
# Assuming 'Region', 'ATM_Type' columns exist in the dataset
df_clean = df[['Region', 'ATM_Type']].dropna()
atm_counts = df_clean.groupby(['Region', 'ATM_Type']).size().unstack(fill_value=0)

# Plotting
atm_counts.plot(kind='bar', stacked=True, figsize=(14, 8))
plt.title('Distribution of Free-to-Use vs Pay-to-Use ATMs by Region')
plt.xlabel('Region')
plt.ylabel('Number of ATMs')
plt.legend(title='ATM Type')
plt.tight_layout()
plt.show()
```

Note: The actual column names and data structure may vary based on the dataset provided by LINK. Ensure to adjust the code accordingly to match the dataset's schema.

By following these steps, you can create a visualization that highlights the distribution of FTU and

concentrations of fee-charging ATMs versus free ones.

Sources

I want you to create this visualisation for me



To create the requested visualization, I need the specific LINK ATM network dataset containing information on ATM locations, types (free-to-use or pay-to-use), and regions. Unfortunately, I cannot directly access the dataset due to restrictions on downloading files from certain websites.

You can either:

1. **Provide the Dataset:** Upload the dataset here, and I'll process it to create the visualization.
2. **Describe the Dataset:** Let me know the structure and column names, such as region, ATM type, and count, so I can simulate the data and create the visualization.

Let me know how you'd like to proceed!



where can i get the dataset



LINK / The ATM Network
co.uk

To obtain the dataset for analyzing the distribution of free-to-use (FTU) versus pay-to-use (PTU) ATMs across different regions in the UK, you can access the LINK ATM Network's official data repository. LINK provides detailed information about ATM locations and types, which is periodically updated.

To obtain the dataset for analyzing the distribution of free-to-use (FTU) versus pay-to-use (PTU) ATMs across different regions in the UK, you can access the LINK ATM Network's official data repository. LINK provides detailed information about ATM locations and types, which is periodically updated.

Steps to Access the Dataset:

1. Visit the LINK Data Repository:

- Navigate to LINK's official data page:
- Here, you'll find various datasets related to the ATM network.

2. Locate the ATM Snapshot Data:

- On the page, look for sections titled "ATM Snapshot Data" or similar.
- These sections provide datasets detailing the location and type of every ATM on the LINK Network as of specific dates.

3. Download the Relevant Dataset:

- Choose the most recent dataset available to ensure up-to-date analysis.
- Click on the corresponding link to download the dataset, which is typically provided in a CSV format.

Example of Available Data:

• ATM Snapshot Data:

- This dataset includes the location of every ATM and cash-at-the-till counter on the LINK Network as of a particular date.
- It provides details such as ATM ID, location coordinates, type (FTU or PTU), and other relevant attributes.

Snapshot at a moment in time. Ensure to check the dataset's publication date to understand the timeframe of the data.

By accessing and analyzing this dataset, you can visualize the distribution of FTU and PTU ATMs across different regions in the UK, providing valuable insights into cash accessibility.

Sources

Link Atm Numbers Year End 2023 sheet1 ↴				
	Unnamed: 0	Unnamed: 1	Unnamed: 2	Unnamed: 3
1		ATM Numbers - year end		
2			1998	1999
3		Free to use	24574	27500
4		Pay to use		150

here you are



The dataset appears to contain historical ATM data with the following structure:

- **Years:** Spanning from 1998 to 2010 (and potentially beyond in the full dataset).
- **Categories:** Includes "Free to use," "Pay to use," and "Total" ATMs for each year.

The data seems to have some header rows ↓, column names that require cleaning. I'll clean and

The dataset includes free-to-use, pay-to-use, and total ATMs for each year.

The data seems to have some header rows and column names that require cleaning. I'll clean and organize the dataset to focus on the relevant columns and create a visualization comparing free-to-use and pay-to-use ATMs over time. [2]

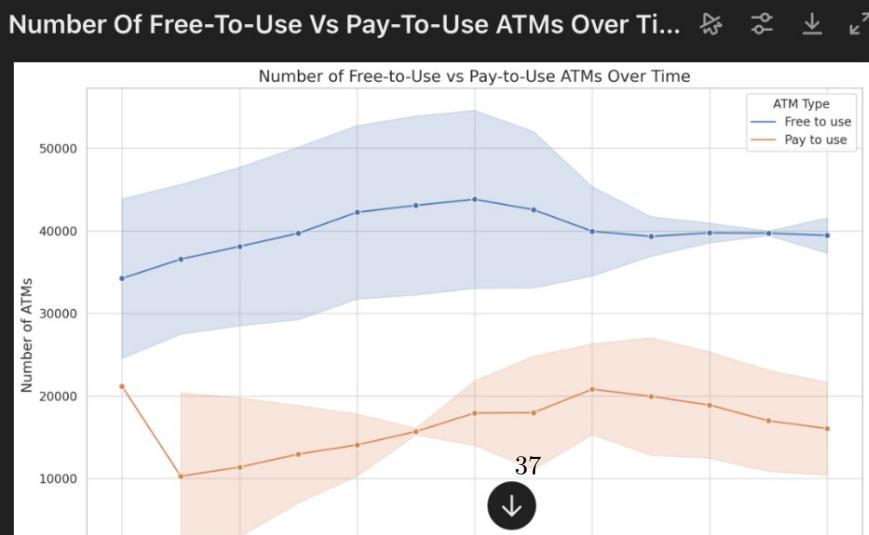
It appears that the dataset has more columns than anticipated, likely including extra empty or mislabeled columns. I'll inspect the column structure and adjust the cleaning process accordingly.

[2]

The dataset includes unexpected columns, such as `Unnamed: 0` and `Unnamed: 1`, along with a mix of labeled and numeric columns. I'll refine the column selection process and exclude unnecessary columns for better clarity. [2]

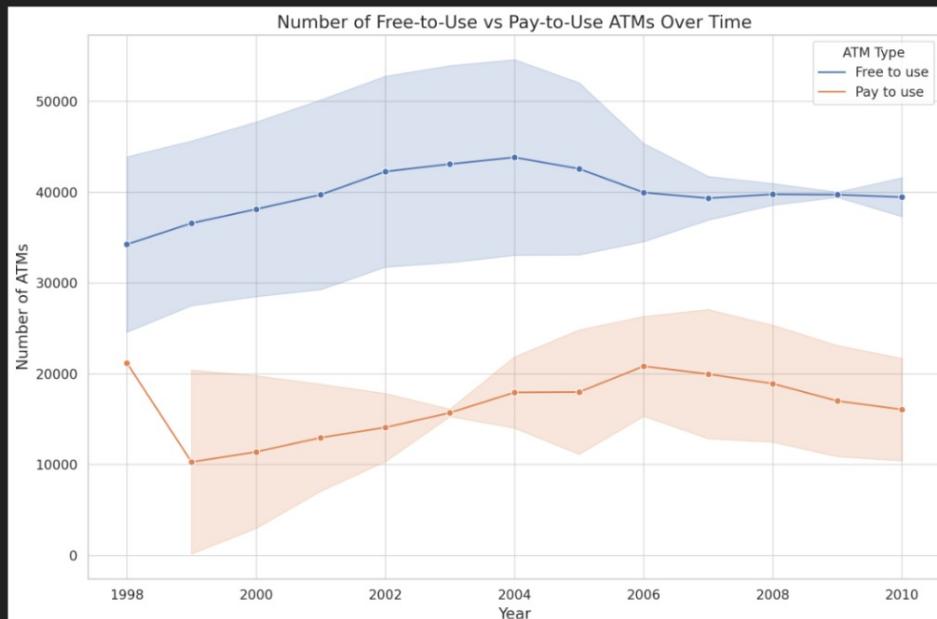
The dataset has been successfully cleaned and reshaped into a format suitable for visualization. Now, I'll create a line chart comparing the number of free-to-use and pay-to-use ATMs over time.

[2]



[i-J]

Number Of Free-To-Use Vs Pay-To-Use ATMs Over Ti... ✖️ ⌂ ⌄ ⌅ ↕



Here is a line chart showing the number of free-to-use (FTU) and pay-to-use (PTU) ATMs over time:

- X-axis: Year
- Y-axis: Number of ATMs
- Lines: Separate trends for FTU and PTU ATMs

This visualization highlights how the availability of free and pay-to-use ATMs has evolved over the years. Let me know if you'd like further customization or additional analysis! [i-J]