[Data Analyst Skill Bootcamp Portfolio]

[Laura Bunel]

Introduction

Enrolling in the Skills Bootcamp program stemmed from my aspiration to transition into a Data Analyst role and forge a lifelong career in Data Analytics. With a background in Sales and Customer Services, I've grown to understand the pivotal role of data within organizations and developed a passion for leveraging it to drive business value.

My proficiency in tools like Excel served as a foundation, and through the course I have learned much more about different data tools such as Python, Pandas and data visualisation tools.

I'm enthusiastic about the prospect of deepening my skills and knowledge. Ultimately, my long-term goal is to continually evolve as a Data Analyst, staying up to date with industry trends and making meaningful contributions to data-centric decision-making processes.

How does data add value to organisations?

As it is the case in my current role, data has proven to be an invaluable tool that allows the business to enhance its performance through informed decisions by drawing insights from historical data but also looking into the future with predictive models. Additionally, using data from existing customers allows to tailor the products and services offered, which improves the overall satisfaction of customers. Keeping informed with the data within our business but also across the broader landscape in a rapidly evolving world is the key to staying ahead of competitor and succeeding.

Analysis of FTSE Data

I was tasked to analyse the data of The Financial Times Stock Exchange 100 (FTSE 100) Index.
Firstly, I work on cleaning the data by deleting a column only containing N/A values and deleting a duplicate row.

```
clean_df = df.copy()
clean_df = clean_df.drop('Strong Buy', axis = 1)
delete_row = clean_df[clean_df['Ticker'] == 'RDSA'].index
clean_df = clean_df.drop(delete_row)
clean_df
```

I then notice that some numerical value are shown as string, I change them to a float datatype.

```
price_df = clean_df.copy()
price_df['Mid-price (p)'] = price_df['Mid-price (p)'].str.replace(',', '')
price_df['Mid-price (p)'] = price_df['Mid-price (p)'].astype(float)
price_df
```

Analysis of FTSE Data

Some of the values needed tidying, I created a function to work through the data.
The function is set to remove '%' and to multiply positive values by 100.

```
def format_change(string):
    for i in string:
        if i == '%':
            string = string.replace('%', '')
    string = float(string)
    if string >= 0:
        string *= 100
    return string
```

Based on a list called watchlist, I created a smaller list based on 2 criteria, want to add it to the list if:

- The prices are equal to or lower to the given target price.
- The Buy Ratio is 0.5 or greater.

```
companies_list = []
for company, price in watchlist:
    if comparison_df['Mid-price (p)'][comparison_df['Company'] == company].item() <= price\
    or comparison_df['Buy Ratio'][comparison_df['Company'] == company].item() >= 0.5:
            companies_list.append(company)
```

Analysis of FTSE Data

This project has helped me understand the steps of data cleaning and the importance of clean data for analysis.

As a first project, I have learned that attention to detail is very important working on a project as a singular error in the code could mean it is not running or not returning the right result.

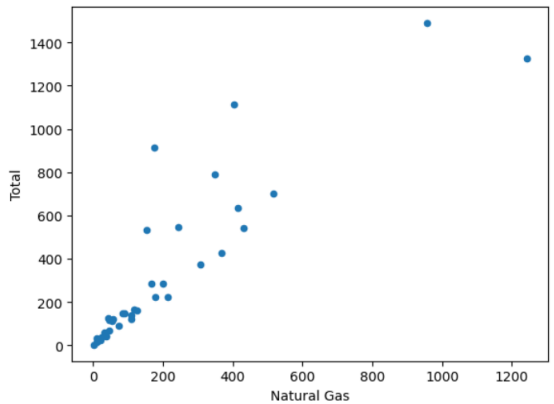
Overall, this project has helped me develop my python and pandas skill and allowed me to understand the reasons why python and pandas are widely use to manipulate data.

Analysis of Energy Demand

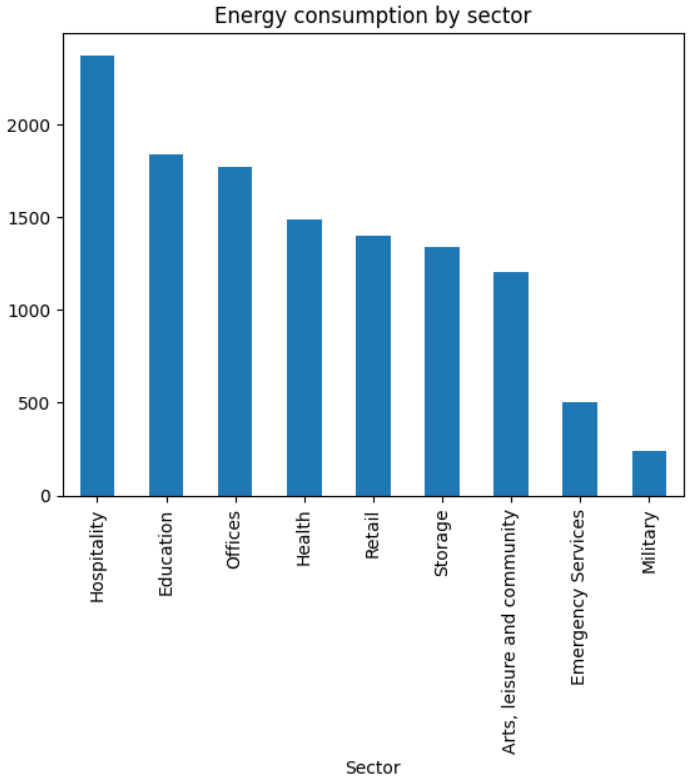
In this project, I had a chance to work with matplotlib to create some visualisation, which helps find insights within the data.
After cleaning the data, I created different visualisations.
In the below visualisation, we study the relationship between the column for Natural Gas and our Total column previously calculated:

```
In [14]: plt.figure()
         gas_total = ss.plot(kind='scatter', x='Natural Gas', y='Total')
```

```
<Figure size 640x480 with 0 Axes>
```



Analysis of Energy Demand



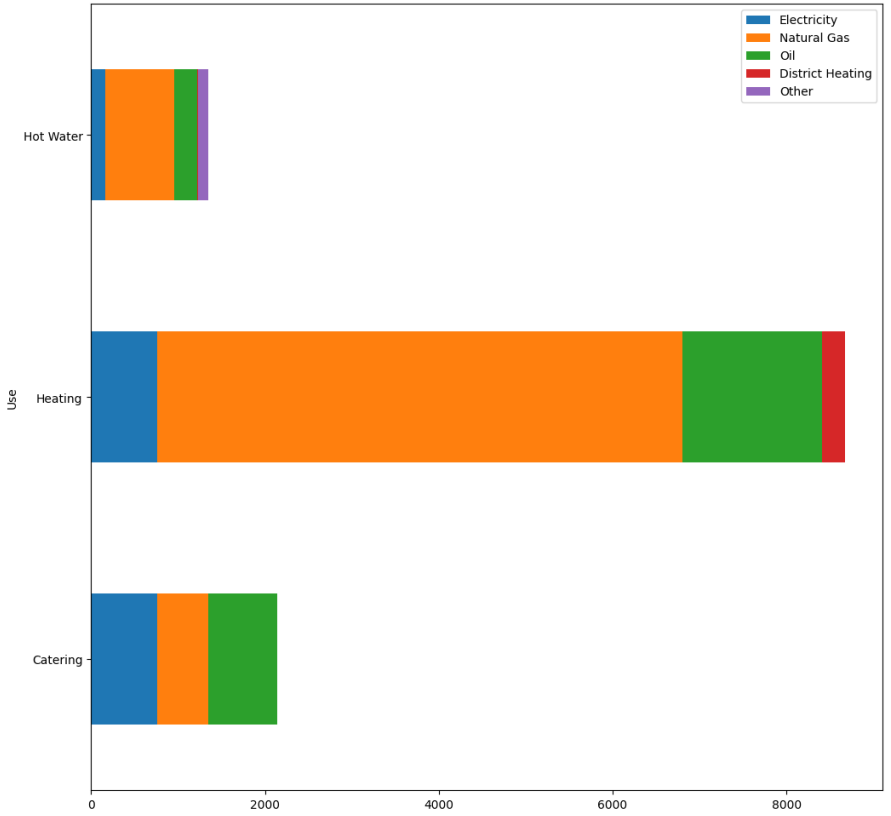
I then looked at the energy consumption by sector.

From this bar chart, it is evident that hospitality has the highest consumption of energy and that military has the lowest consumption.

Analysis of Energy Demand

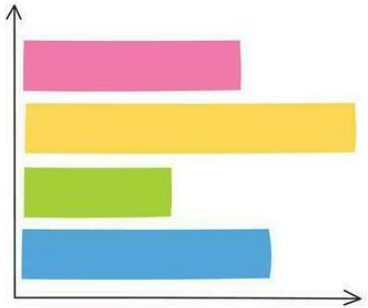
Finally, we look at the amount of energy utilised for each use.

Natural Gas is a very popular solution for heating and hot water particularly.



Analysis of Energy Demand

What I learned wit this project is that incorporating visualization is an essential tool when it comes to understanding the dataset . Using charts, graphs, and other visual aids, I have been able to transform raw data into meaningful insights that cannot be seen with raw data. This not only streamlines the decision-making process but also allows to make informed choices based on a deeper understanding of the dataset's nuances.



Bar

Area

Column - Line

Streaming Service Hackathon

In this group task, we had to analyse Netflix data to understand what makes it a successful streaming service and how a competitor service could optimise its service and compete with Netflix.

This task has highlighted how essential teamwork can be when having to work to a deadline. With each person having different strength, I undertook most of the data cleaning on this dataset, which then allowed the team to work on exploring the data and drawing conclusion.

My data cleaning code can be found under my portfolio.

Streaming Service Hackathon

Once the data was cleaned, the group discussed what questions should be answered with this dataset. We worked on the following:

1. Split films and series, by country (country of origin)
2. How many series do shows run for
3. Ages of series and films
4. Split of ratings
5. Date added - when most content is added; how long things stay on platform
6. Average duration of films / shows - show mean
7. Top genres - adventure, action, drama, comedy, thriller
8. Which countries produce the most content
9. Top directors globally
10. Top actors

With a strict deadline to adhere to, each person participated in creating visualisation that was later used in our presentation.

Conclusion

Overall, this course has taught me essential skills to be able to work with data.

I have learned technical skills such as Python, Pandas, Matplotlib and seaborn which are essential to manipulate data and understand trends, but I have also worked on my problem-solving skills, resilience and collaboration within a team.

This course has equipped me with technical knowledge but also deepened my existing interpersonal skill, providing a great foundation for a successful career in data analysis.