Project Report

GitHub URL

https://github.com/rupaligs/UCDPA_Rupali_Murkumbi

Abstract

This project is about the analysis of travel trends and impact due to weather. This project will look at if there is any variation in travel because of change in weather (rainfall, temperature, and sunshine) in Ireland.

Introduction

- Personal motivation:
  I have experienced that the weather in Ireland is very unpredictable and I wanted to see if people decide to travel based on the weather.
- Why is the business problem important:
  This analysis can be useful for the tourism industry to prepare for tourists based on weather conditions.
- Can it be converted to data analytics problem:
  If trends show that there is an impact on travel due to weather conditions, this can convert into data analytics problem to predict and prepare.

Dataset

- This dataset is available on website of central statistics office data.cso.ie. The travel data has been collected by CSO office as part of the surveys and weather data is provided by Met Eireann to CSO.

Dataset ID	Dataset Name	Source
HTQ01	Travel by Irish residents	https://data.cso.ie/table/HTQ01
MTM01	Rainfall	https://data.cso.ie/table/MTM01
MTM02	Temperature	https://data.cso.ie/table/MTM02
MTM03	Sunshine	https://data.cso.ie/table/MTM03

- Description of the variable(columns) in the dataset
  HTQ01

Field	Description	Contents example
Statistic		Number of Trips by Irish Residents (Thousand) Number of Nights by Irish Residents (Thousand) Average Length of Stay by Irish Residents (Nights per Trip) Estimated Expenditure by Irish Residents (Euro Million)
Quarter	Quarter of the year	2022Q1 2021Q2
Domestic	Type of journey	Domestic Outbound

Reason for journey		All reasons for journey Holiday Business Visiting friends/relatives Other reasons

MTM01 (Rainfall)

Field	Description	Contents example
Statistics		Total Rainfall (Millimetres) Most Rainfall in a Day (Millimetres) Raindays (0.2mm or More) (Number)
Month		2022M12 2022M11 2022M10 2022M09
Weather station		Cork airport Dublin airport Galway Kilkenny

MTM02 (Temperature)

Field	Description	Contents example
Statistics		Average Maximum Temperature (Degrees C) Average Minimum Temperature (Degrees C) Mean Temperature (Degrees C) Highest Temperature (Degrees C) Lowest Temperature (Degrees C)
Month		2022M12 2022M11 2022M10 2022M09
Weather station		Cork airport Dublin airport Galway Kilkenny

MTM03 (Sunshine)

Field	Description	Contents example
Statistics		Total Sunshine Hours (Number) Most Sunshine Hours in a Day (Number)
Month		2022M12 2022M11 2022M10 2022M09

Weather station		Cork airport Dublin airport Galway Kilkenny	

Implementation Process

- Data Importing
  - All the csv files were downloaded and imported into panda frames
  - Checked for number of rows for each file.

    e.g.
    ```
    rainfall=pd.read_csv('dataset/MTM01.20230430T170443.csv')
    ```
  - Created database travel_ireland in SQLite
  - Imported csv file of travel details into table travel_ireland

- Data Cleaning
  - Checked for data types of imported csv files
  - Checked for null or missing values.
  - Fixed the missing values using *dropna* function
  - Merged the data frames rainfall, sunshine, and temperature on month of the year
  - Separated year and month from column Month
  - Created extra column Quarter using the values from Month column to join with travel data
  - Created new dataframe to group by on values of rain, sunshine and temperature on year and quarter.
  - Created year column on travel data frame
  - Merged travel and weather datasets

    ```
    climate_travel = pd.merge(grouped, travel_3, on=['YY','Quarter'])
    ```

Figure 1

- Data Exploration
  - Exported each data frame in csv file to analyze the results after each operation (e.g., Fixing missing values, merging)
  - Applied various filters to explore and analyze data
  - Created function in order to produce graphs/charts by passing year parameter
  - Executed SQL queries in SQLite database to analyze travel data

```sql
SELECT Reason_Journey,Quarter,ROUND(SUM(VALUE)/24) AS Days_Traveled
FROM travel_ireland
WHERE substr(Quarter,1,4) ='2019'
and Domestic='Domestic'
GROUP BY Reason_Journey,Quarter;
```

```sql
SELECT Reason_Journey,Quarter,ROUND(SUM(VALUE)/24) AS Days_Traveled
FROM travel_ireland
WHERE substr(Quarter,1,4) ='2021'
and Domestic='Domestic'
GROUP BY Reason_Journey,Quarter;
```

Figure 2

- Data Visualization
    - Generated various charts (e.g., Weather trends for month, quarterly weather trends with travel)
- Data Modelling
    - Regression model will suit this data analytics problem

Results

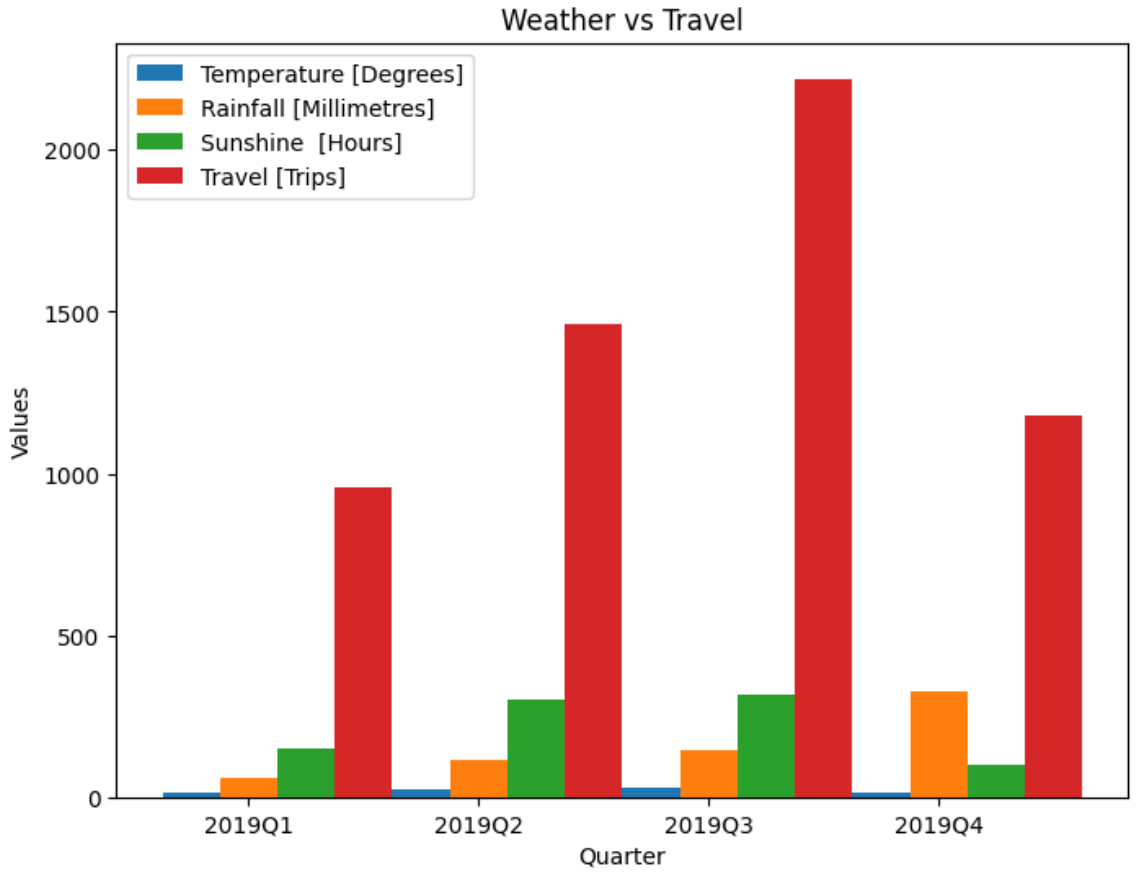

Figure 3 - Bar chart for 2019

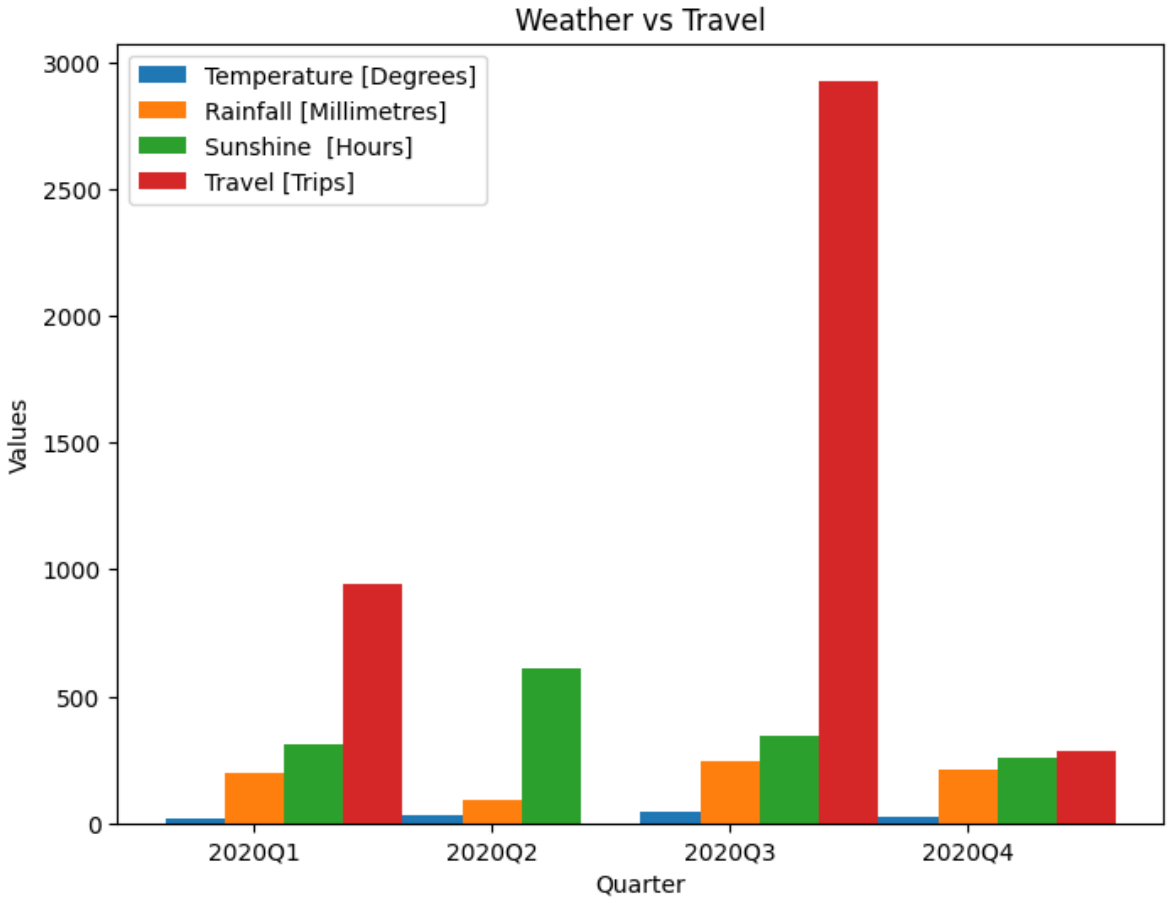

Figure 4 - Bar chart for 2020

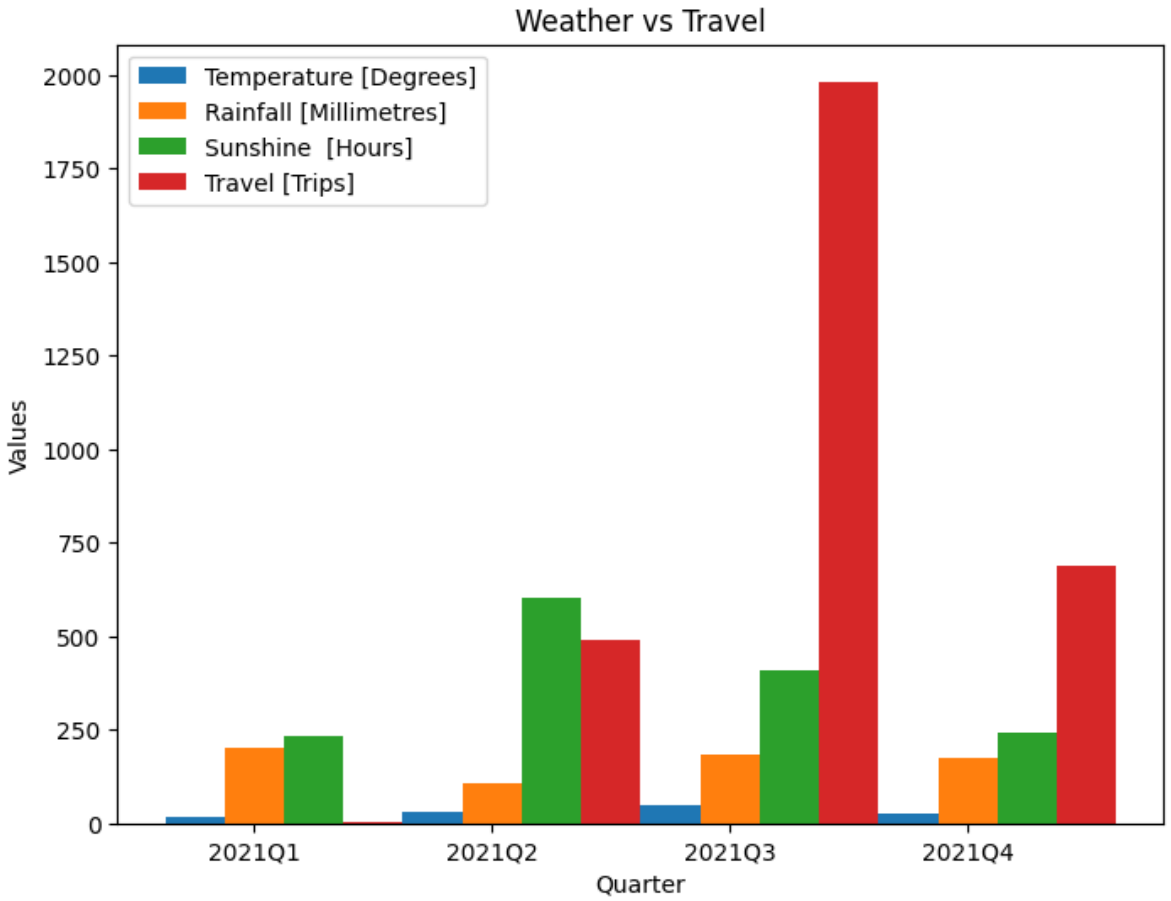

Figure 5 - Bar chart for 2021

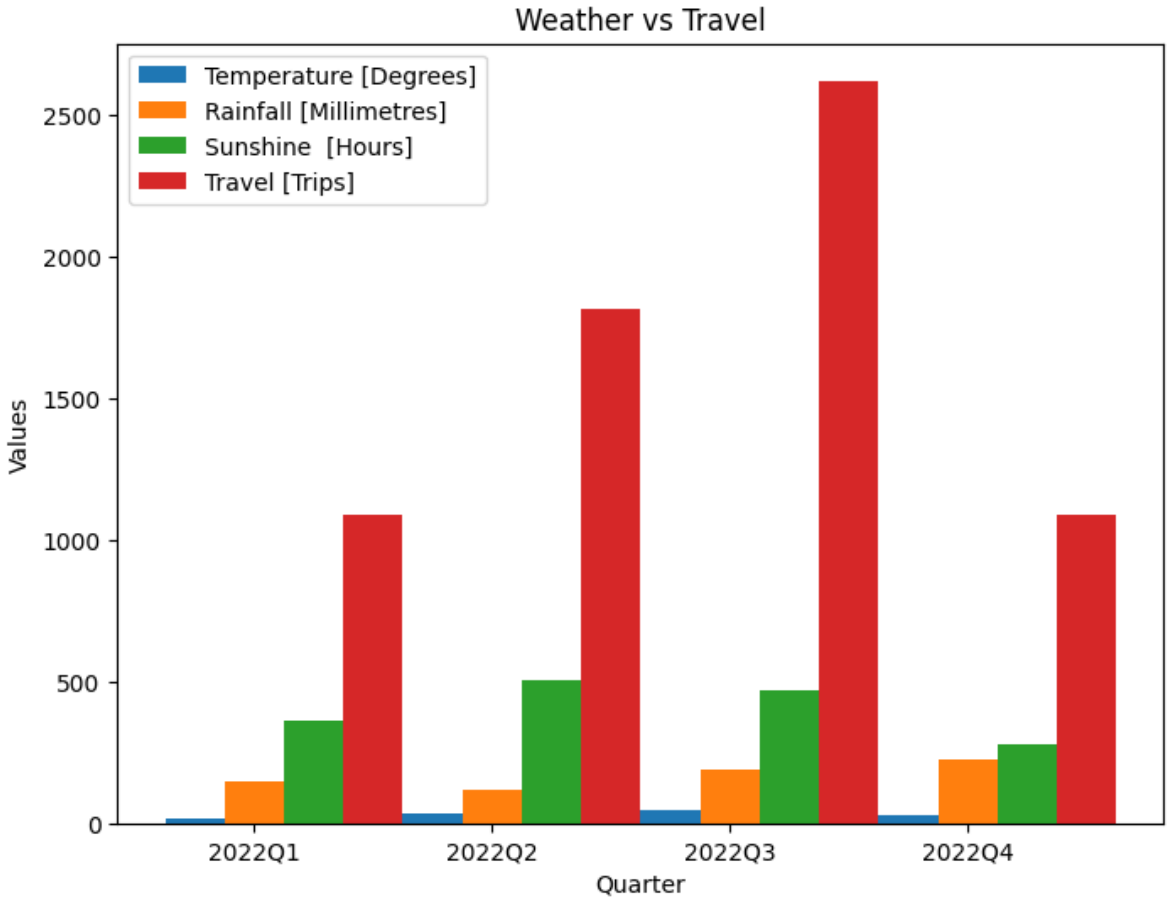

Figure 6 - Bar chart for 2022

Add snapshots(picture) of your code where relevant
Results of data exploratory
Results of data visualization

Results of modelling (visualization)

```
SELECT Reason_Journey,Quarter,ROUND(SUM(VALUE)/24) AS Days_Traveled
FROM travel_ireland
WHERE substr(Quarter,1,4) ='2019'
and Domestic='Domestic'
GROUP BY Reason_Journey,Quarter;
```

```
SELECT Reason_Journey,Quarter,ROUND(SUM(VALUE)/24) AS Days_Traveled
FROM travel_ireland
WHERE substr(Quarter,1,4) ='2021'
and Domestic='Domestic'
GROUP BY Reason_Journey,Quarter;
```

Figure 7

Statistic_Label	Domestic	Quarter	Expenditure_in_million
Estimated Expenditure by Irish Residents	Domestic	2000Q1	228
Estimated Expenditure by Irish Residents	Outbound	2000Q1	1120
Estimated Expenditure by Irish Residents	Domestic	2000Q2	312
Estimated Expenditure by Irish Residents	Outbound	2000Q2	1752
Estimated Expenditure by Irish Residents	Domestic	2000Q3	572
Estimated Expenditure by Irish Residents	Outbound	2000Q3	2096
Estimated Expenditure by Irish Residents	Domestic	2000Q4	303
Estimated Expenditure by Irish Residents	Outbound	2000Q4	1408
Estimated Expenditure by Irish Residents	Domestic	2001Q1	296
Estimated Expenditure by Irish Residents	Outbound	2001Q1	1410
Estimated Expenditure by Irish Residents	Domestic	2001Q2	318
Estimated Expenditure by Irish Residents	Outbound	2001Q2	1898
Estimated Expenditure by Irish Residents	Domestic	2001Q3	748
Estimated Expenditure by Irish Residents	Outbound	2001Q3	2362
Estimated Expenditure by Irish Residents	Domestic	2001Q4	399
Estimated Expenditure by Irish Residents	Outbound	2001Q4	1431
Estimated Expenditure by Irish Residents	Domestic	2002Q1	328
Estimated Expenditure by Irish Residents	Outbound	2002Q1	1468

Figure 8

Reason_Journey	Quarter	Days_Traveled
Business	2020Q1	17
Business	2020Q3	7
Business	2020Q4	9
Holiday	2020Q1	127
Holiday	2020Q3	607
Holiday	2020Q4	41
Other reasons	2020Q1	43
Other reasons	2020Q3	14
Other reasons	2020Q4	48
Visiting friends/relatives	2020Q1	107
Visiting friends/relatives	2020Q3	203
Visiting friends/relatives	2020Q4	95

Figure 9 - 2020

Reason_Journey	Quarter	Days_Traveled
Business	2019Q1	14
Business	2019Q2	17
Business	2019Q3	12
Business	2019Q4	23
Holiday	2019Q1	126
Holiday	2019Q2	232
Holiday	2019Q3	447
Holiday	2019Q4	160
Other reasons	2019Q1	37
Other reasons	2019Q2	60
Other reasons	2019Q3	88
Other reasons	2019Q4	44
Visiting friends/relatives	2019Q1	114
Visiting friends/relatives	2019Q2	115
Visiting friends/relatives	2019Q3	180
Visiting friends/relatives	2019Q4	135

Figure 10 - 2019

Reason_Journey	Quarter	Days_Traveled
Business	2021Q1	4
Business	2021Q2	4
Business	2021Q3	4
Business	2021Q4	11
Holiday	2021Q1	2
Holiday	2021Q2	85
Holiday	2021Q3	441
Holiday	2021Q4	98
Other reasons	2021Q1	21
Other reasons	2021Q2	19
Other reasons	2021Q3	19
Other reasons	2021Q4	24
Visiting friends/relatives	2021Q1	34
Visiting friends/relatives	2021Q2	64
Visiting friends/relatives	2021Q3	103
Visiting friends/relatives	2021Q4	103

Figure 11 -2021

Reason_Journey	Quarter	Days_Traveled
Business	2022Q1	10
Business	2022Q2	31
Business	2022Q3	21
Business	2022Q4	15
Holiday	2022Q1	161
Holiday	2022Q2	282
Holiday	2022Q3	538
Holiday	2022Q4	148
Other reasons	2022Q1	41
Other reasons	2022Q2	58
Other reasons	2022Q3	86
Other reasons	2022Q4	67
Visiting friends/relatives	2022Q1	121
Visiting friends/relatives	2022Q2	135
Visiting friends/relatives	2022Q3	177
Visiting friends/relatives	2022Q4	212

Figure 12 - 2022

Insights

1) In the year 2020Q2, there is no travel, neither domestic nor outbound due to travel restrictions.
   As the 2020Q3 include the summer and school holidays, families had more time to spend together as students have more free time. Since 2020 was the year when travel abroad restrictions were still in place, people were encouraged to partake in 'staycations' which caused people to still travel around the country.
   In the year 2020Q1, 2020Q3, 2020Q4 in 2020Q3 most of the travel is for holidays and to visit friends and relatives.

2) The graph in figure 4 shows that very few people travelled in 2020 in the second quarter. Although the rainfall for this quarter is less than the other three years and the temperature is roughly the same as other years, less people travelled in 2020 within the second quarter than the other three years. In this case, the COVID 19 restrictions can be the most likely reason as people were advised to decrease their travelling to avoid the spread of COVID 19.

3) In figure 4, 2020Q4 shows that a lot less people travelled in comparison to other years. However, COVID restrictions were relaxed during this quarter to allow people to celebrate Christmas despite the rising COVID 19 cases which explains why some people still travelled during this time. In the following year, 2021Q1 displays that very few people travelled during this time compared to any other year even though temperature and rainfall are similar to other years. This is because COVID 19 restrictions were placed again after an increase in cases.

4) In general, many people prefer to travel within the 3rd quarter as there is a rise in temperature and sunshine.

5) The trends for people travelling in the graphs shown in figure 3 and figure 6 are similar. The number of people that travel increases steadily from the 1st quarter to the 3rd quarter, and then it falls noticeably in both of these graphs. This is because in 2019 COVID 19 restrictions had not come into place yet, and therefore, people were free to travel any time during the year. In 2022 COVID 19 restrictions in Ireland were fully lifted which allowed people to freely travel once again.

References

Dataset references: data.cso.ie

Dataset ID	Dataset Name	Source

HTQ01	Travel by Irish residents	https://data.cso.ie/table/HTQ01
MTM01	Rainfall	https://data.cso.ie/table/MTM01
MTM02	Temperature	https://data.cso.ie/table/MTM02
MTM03	Sunshine	https://data.cso.ie/table/MTM03