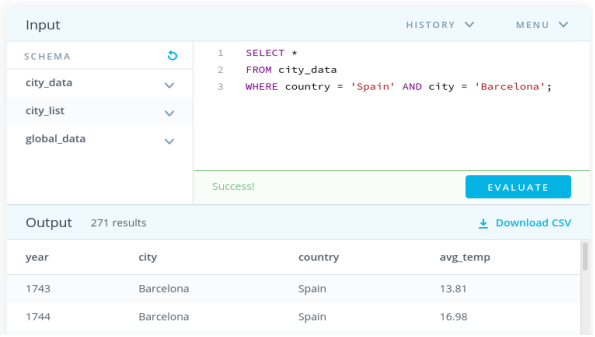
Project 1 – Weather Trends

In this project a exploratory data analysis is performed in dataset of temperatures averages around the world overtime.

Outline

1. Extract data from a relational database using SQL
2. Import data to Python in a Pandas DataFrame
3. Use Pandas Rolling method to calculate moving average with given a window
4. Use PyPlot from Matplotlib to plot the moving averages and compare.
5. Summarize insights from the data

Downloaded a CSV file as output from the SQL query of the portal.



Then imported the csv data into python serializing into a pandas DataFrame using the code below:

```
#!/usr/bin/env python3
"""Temp comparison, Barcelona vs Global trend."""

import csv
import pandas as pd
from matplotlib import pyplot as plt

def get_csv(path):
    df = pd.read_csv(path)
    return df

def moving_average(df, col, window=10):
    df['SMA'] = df.iloc[:,col].rolling(window=window).mean()
    return None


def plot_lines(df1, df2):
    plt.plot(df1["year"], df1["SMA"], label = 'Barcelona')
    plt.plot(df2["year"], df2["SMA"], label = 'World')
    plt.xlabel('year')
    plt.ylabel('average')
    plt.title('Average temperature in Barcelona')
```

```
    plt.legend()
    plt.show()

def main():
    # Get data into pandas DataFrame
    path_barcelona = r'Barcelona_Temp_AVG.csv'
    path_world = r'Global_Data.csv'
    temp_barcelona = get_csv(path_barcelona)
    temp_world = get_csv(path_world)

    # Add Simple Moving Average of given windows at given colum from df
    window = 15
    col_barcelona = 3
    col_world = 1
    moving_average(temp_barcelona, col_barcelona, window)
    moving_average(temp_world, col_world, window)
    # Plot
    plot_lines(temp_barcelona, temp_world)

if __name__ == "__main__":
    main()
```

The output with insights in the graph.