In [ ]:

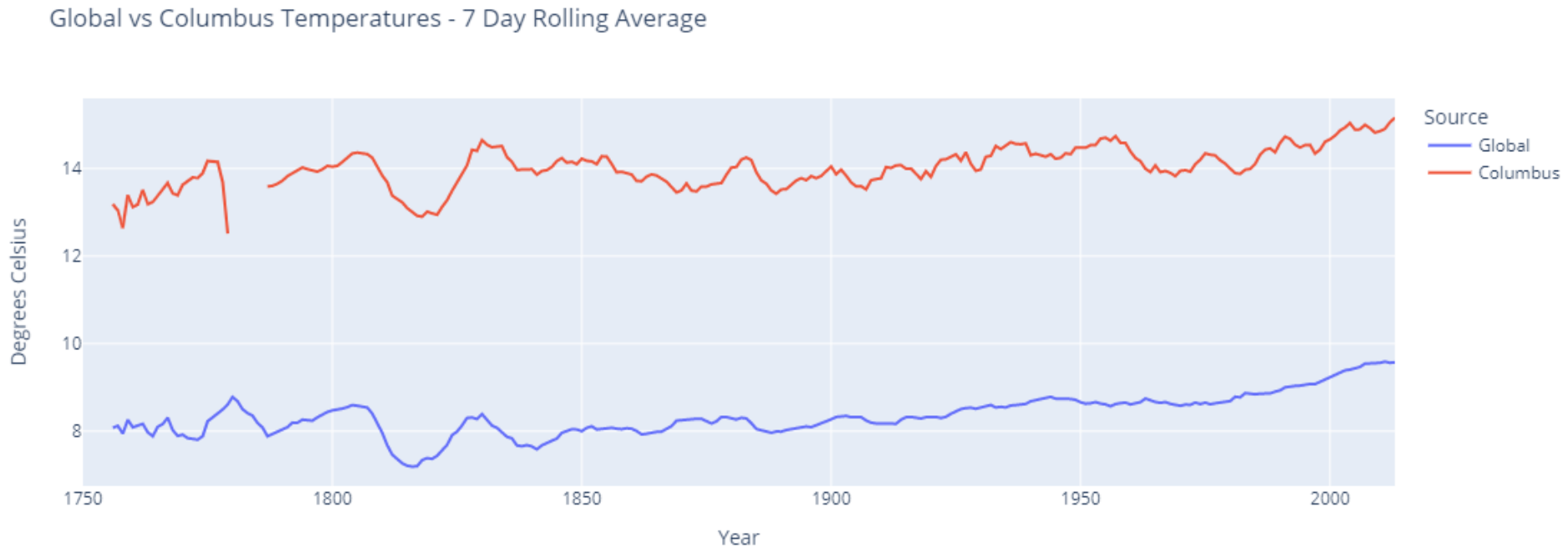
```
import pandas as pd
import plotly.express as px

# Load Data
df = pd.read_csv('temperatures.csv')

# Build a Rolling Average
df['Columbus'] = df['avg_temp'].rolling(7).mean()
df['Global'] = df['global_avg_temp'].rolling(7).mean()

df.tail(1)
```

Out[ ]:

	year	city	country	avg_temp	global_avg_temp	Columbus	Global
263	2013	Columbus	United States	16.05	9.61	15.158571	9.572857

In [ ]:

```
# Display rolling averages
labs = {'value': 'Degrees Celsius', 'year': 'Year', 'variable': 'Source'}
px.line(df, x='year', y=['Global', 'Columbus'], title='Global vs Columbus Temperatures - 7 Day Rolling Average', labels=labs)
```



Steps

1. Querying Data using SQL

```sql
SELECT
    cd.*,
    gd.avg_temp AS global_avg_temp
FROM
    city_data cd
    INNER JOIN global_data gd ON cd.year = gd.year
WHERE
    country = 'United States'
    AND city = 'Columbus'
ORDER BY year ASC
```

1. Exporting results as a CSV file
2. Created a Jupyter Notebook
3. Imported the CSV data into a Pandas DataFrame
4. Added Rolling Average Columns
5. Graphed the Rolling Averages using Plotly Express

Observations

1. Both global and Columbus temperatures seem to be gradually trending upwards
2. There seems to be more of a variance in Columbus temperatures vs global temperature measurements
3. Columbus temperatures are generally around 5 or 6 degrees Celsius warmer than the global average
4. Columbus tends to follow major changes present in the global temperature
5. Columbus has some bad or missing values around 1780 that contribute to lower overall temperature averages and some missing rolling average values