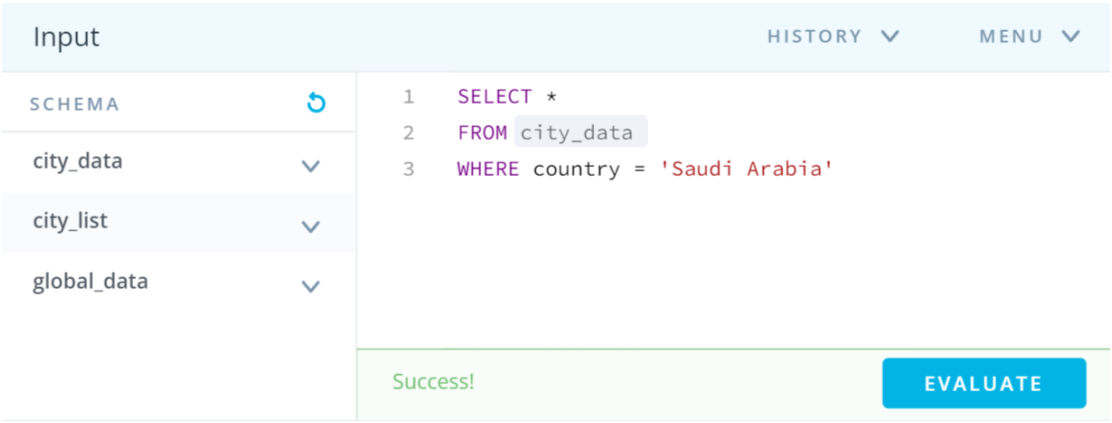
Exploring Weather Trends

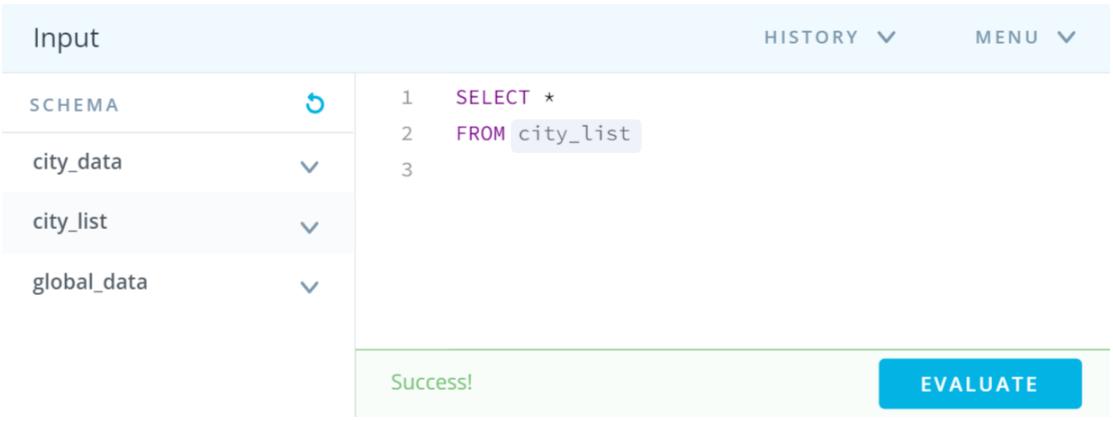
In this project I created a visualization and prepared a write-up describing the similarities and differences between global temperature trends and temperature trends in my city Riyadh

- **Extract the data:**

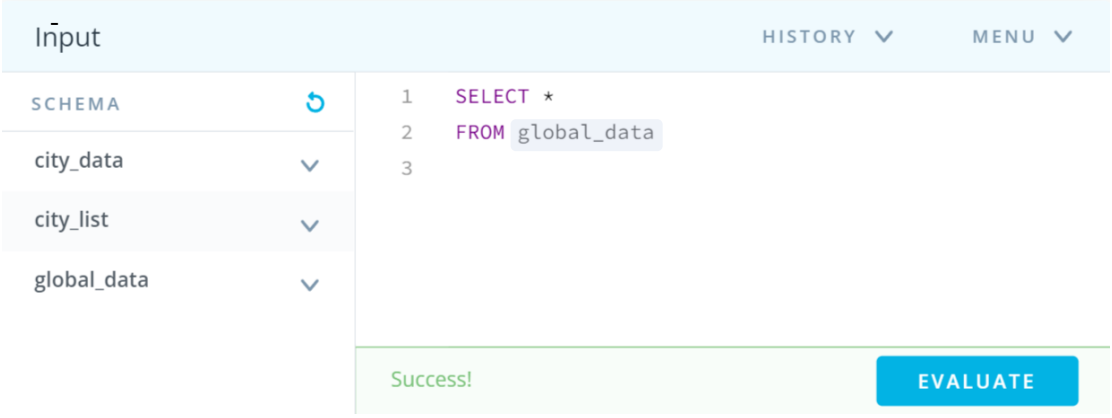
- City_data:



- City_list:



- Global_data:



- **Open up csv:**

In [2]:
```
saudi_data = pd.read_csv('city_data.csv')
global_data = pd.read_csv('global.csv')
```

- **Cleaning & Pre-processing:**

In line number 10 at processing_saudi_data and line number 7 at processing_global_data you can see how I calculated the moving average.

In [7]:
```
def processing_saudi_data(data):

    # Query by My city : RIYADH
    data = data.query('city == "Riyadh" ')

    # To Avoid Nan values in avg_temp columns in my city dataset
    data = data[data['year'].between(1848, 2013)]

    # Moving Average ( 5 years )
    data['avg_temp_riyadh_5years'] = data['avg_temp'].rolling(5).mean()

    # Remove unused columns
    data = data.drop(['city', 'country', 'avg_temp'], axis = 1)

    # Remove null values
    data = data.dropna()

    return data.reset_index(drop = True)
```

In [8]:
```
def processing_global_data(data):

    # To compare global dataset with my city dataset in same years
    data = data[data['year'].between(1848, 2013)]

    # Moving Average ( 5 years )
    data['avg_temp_global_5years'] = data['avg_temp'].rolling(5).mean()

    # Remove unused columns
    data = data.drop('avg_temp', axis = 1)

    # Remove null values
    data = data.dropna()

    return data.reset_index(drop = True)
```

- **After Cleaning & Pre-processing:**

In [10]:
```
cleaned_saudi_data.head()
```

Out[10]:

	year	avg_temp_riyadh_5years
0	1852	24.716
1	1853	24.790
2	1854	24.774
3	1855	24.890
4	1856	24.798

In [12]:
```
cleaned_global_data.head()
```

Out[12]:

	year	avg_temp_global_5years
0	1852	8.028
1	1853	8.040
2	1854	8.086
3	1855	8.128
4	1856	8.092

- **Merge Datasets:**

In [13]:
```
avg = cleaned_global_data.join(cleaned_saudi_data.set_index("year"), on = "year")
```

In [14]:
```
avg
```

Out[14]:

	year	avg_temp_global_5years	avg_temp_riyadh_5years
0	1852	8.028	24.716
1	1853	8.040	24.790
2	1854	8.086	24.774
3	1855	8.128	24.890
4	1856	8.092	24.798
...	...	...	...
157	2009	9.580	26.384
158	2010	9.580	26.604
159	2011	9.578	26.636
160	2012	9.534	26.704
161	2013	9.570	27.018

- **Analysis & Visualization:**

```
In [15]:
1 avg['avg_temp_global_5years'].describe()
```

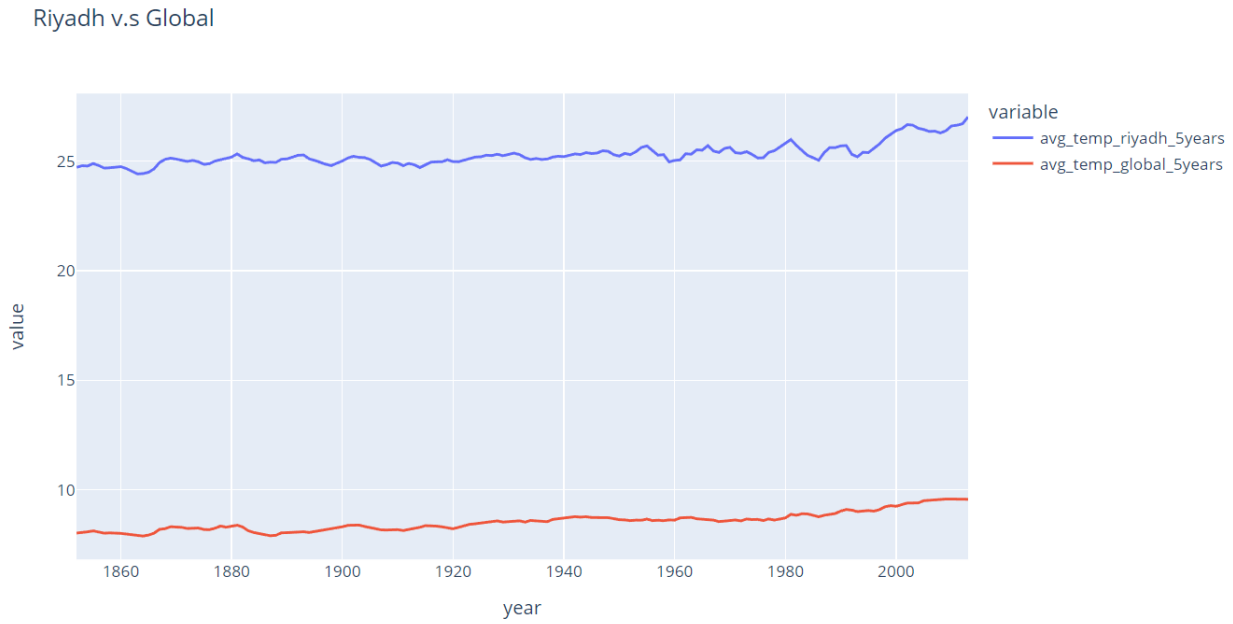
```
Out[15]: count    162.000000
mean       8.545494
std        0.420727
min        7.892000
25%        8.213500
50%        8.550000
75%        8.730500
max        9.580000
Name: avg_temp_global_5years, dtype: float64
```

```
In [16]:
1 avg['avg_temp_riyadh_5years'].describe()
```

```
Out[16]: count    162.000000
mean      25.287370
std        0.489501
min       24.414000
25%       24.972500
50%       25.184000
75%       25.419500
max       27.018000
Name: avg_temp_riyadh_5years, dtype: float64
```

- **Riyadh v.s Global**

```
In [17]:
1 fig = px.line(avg, x = 'year', y = ["avg_temp_riyadh_5years", 'avg_temp_global_5years' ],  title = 'Riyadh v.s Global')
2 fig.show()
```



Observation:

1. During the 5-year MA, the global temperature ranges from 7.98°C to 9.58°C
2. A 5-year MA temperature for Riyadh city ranges from 24.41°C to 27°C.
3. There is a huge difference in temperature between Riyadh and Global in the chart.
4. The average temperature in Riyadh (25.28 °C) is higher than the global average (8.5 °C) temperature
5. Riyadh moving average temperature has been up and down the last few years
6. Both temperatures increased from 1852 to 2013

The final conclusion of this project is Riyadh is hotter than global temperature and temperature is increasing day by day due to changes in the climate.

Tools:

1. **Python Language Programming**
2. **MySQL**

Packages:

1. **Pandas**
2. **Plotly**