Final Capstone Project

Weather Forecasting using Machine Learning

by

Saeed Manavipour

Weather forecasting is a multi-billion dollar industry with mane applications such as

- Transportation
- Aviation
- Agriculture
- Sport & Entertainment
- Military
- Power Generation
- Etc…

Goal

- To look at weather forecasting from a data-driven perspective
- Use machine learning tools to predict the weather using past data

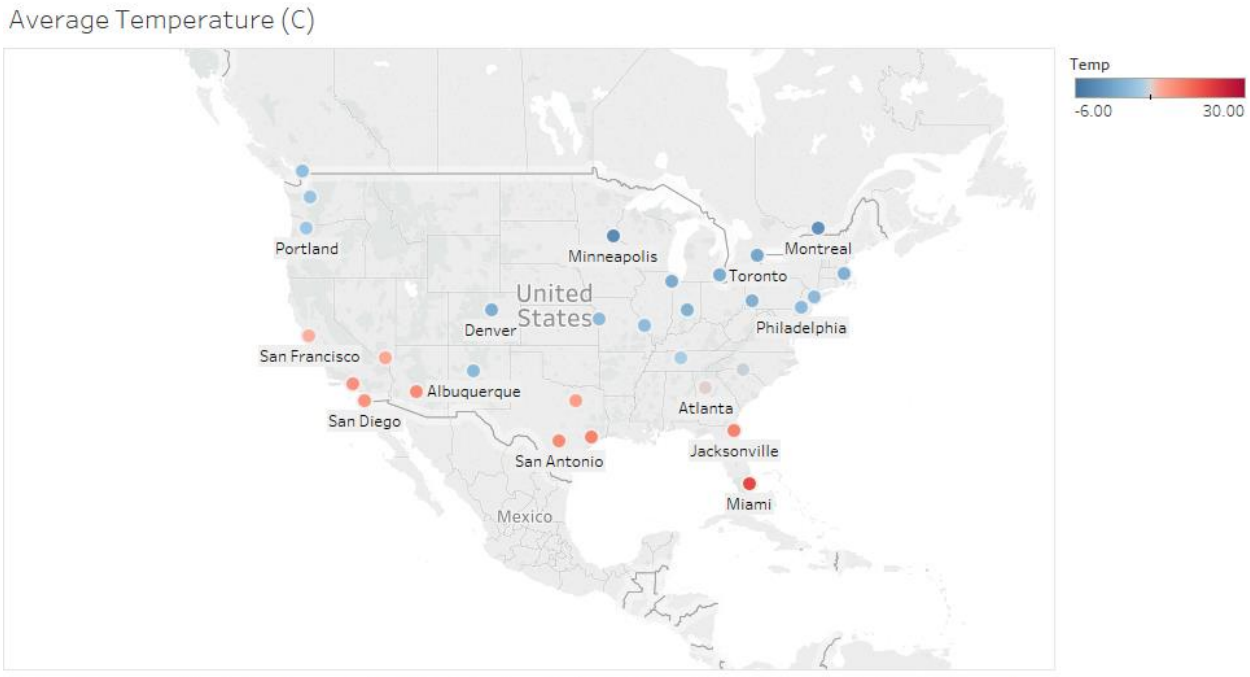
Dataset:

- Five years (2012 to 2017) of hourly data (temperature, pressure, humidity, wind speed, wind direction)
- 30 US & Canadian Cities + 6 Israeli Cities
- 45,000 rows of data
- Hourly weather description (sunny, rainy, clear …)
- It contains geographical coordinates such as latitude, longitude
- Elevation and proximity to water was added later

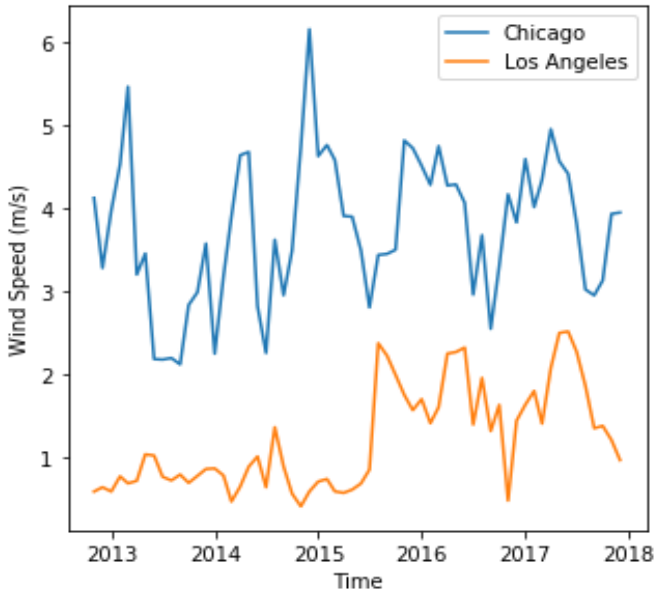
CSV Files:

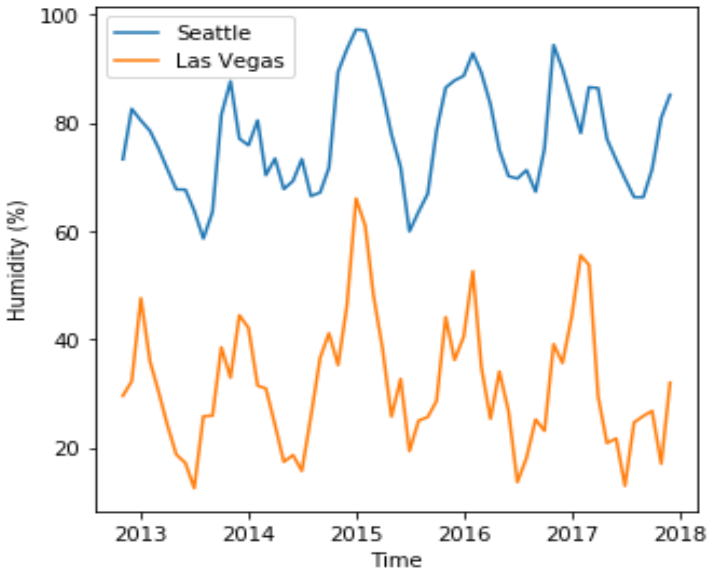
city_attributes.csv
humidity.csv
pressure.csv
temperature.csv
weather_description.csv
wind_direction.csv
wind_speed.csv

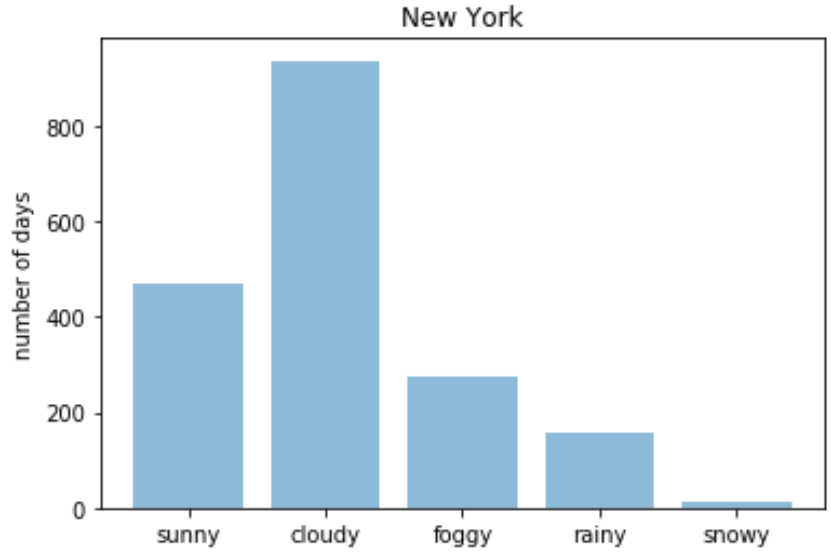
Exploratory Data Analysis (EDA)

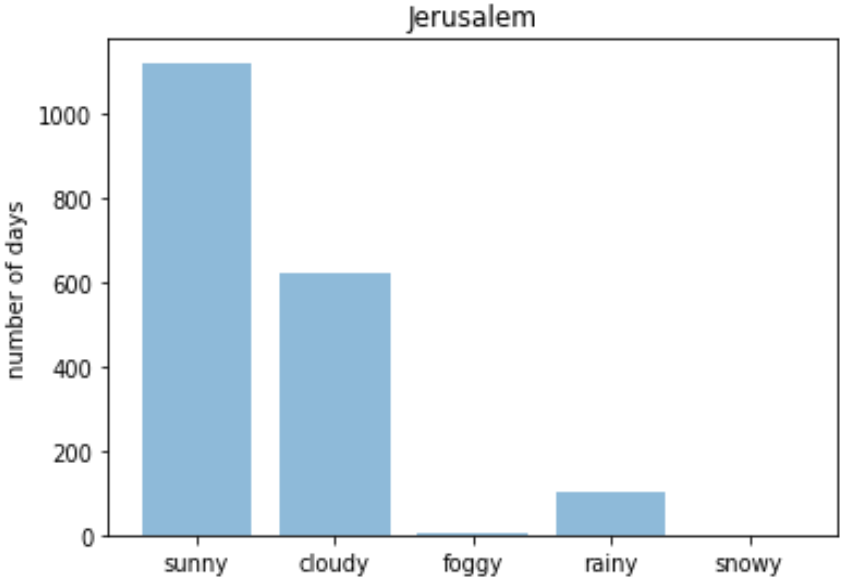


Map based on average of Longitude and average of Latitude. Color shows details about Temp. The marks are labeled by City. Details are shown for Country.


Chicago
Los Angeles
Wind Speed (m/s)
1
2
3
4
5
6
2013
2014
2015
2016
2017
2018
Time


Seattle
Las Vegas
Humidity (%)
20
40
60
80
100
2013
2014
2015
2016
2017
2018
Time





Total: 1850 Days

Approach

Using the data from past two days to predict the weather for the next day!

Time-Series:

	Temperature	Pressure	Humidity	Wind Speed	Wind Direction	Weather Status
Two days ago	✔	✔	✔	✔	✔	✔
Yesterday	✔	✔	✔	✔	✔	✔
Today	?	?	?	?	?	?

Scikit Learn Machine Learning Models:

Linear Regression Model is used to predict continuous variables such as temperature

XGBoost Model is used for classification of weather description and wind direction

Progress so far

Accuracy scores for continuous variables:

```
Linear Regression, Temperature Score: % 92.96
Linear Regression, Pressure Score:    % 77.98
Linear Regression, Humidity Score:    % 67.98
Linear Regression, Wind Score:        % 40.49
```

Accuracy scores for categorical variables:

```
XG Boost, Weather Description score:  % 67.63
XG Boost, Wind Direction score:       % 39.84
```

Sample Results

```
**************************************************************************
**************************************************************************
**************************************************************************
```

Toronto, 2015-05-12

Actual:

```
mostly rainy
Temp: 13.61 C, Pressure: 102.31 kPa, Humidity: 75.5 %,
Wind: 4.67 m/s ,
Wind Direction: SW
************************
```

Predictions:

```
Model predicted Temp for today:      13.43 C
Model predicted Pressure for today:  102.92 kPa
Model predicted Humidity for today:  79.96 %
Model predicted Wind for today:      2.4 m/s

XGB_model predicted status for today:     Rainy
XGB_model predicted wind direction:       North
```

Questions?