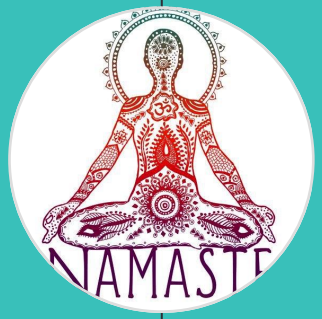
Under the Weather: An Analytical Perspective on Precipitation

Problem Solving with Data Term Project
May 3rd 2017



We are Team Namaste!

1. Nrithya- Solution Architect
2. Akash- Data Analyst
3. Xiqiao- Data Miner
4. Brenden - Information Designer

Will it rain on Rutgers 251st Commencement Day?

- Sun, May 14 2017, New Brunswick.

Life is like Mother Nature, unpredictable
-Artyom Gross

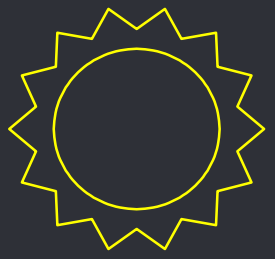
1

Can we predict the occurrence of Rain?

Let’s take a look at why and how we may be able to do this

The Binary Option

It **will rain** if value of PRCP is significantly higher than 0.

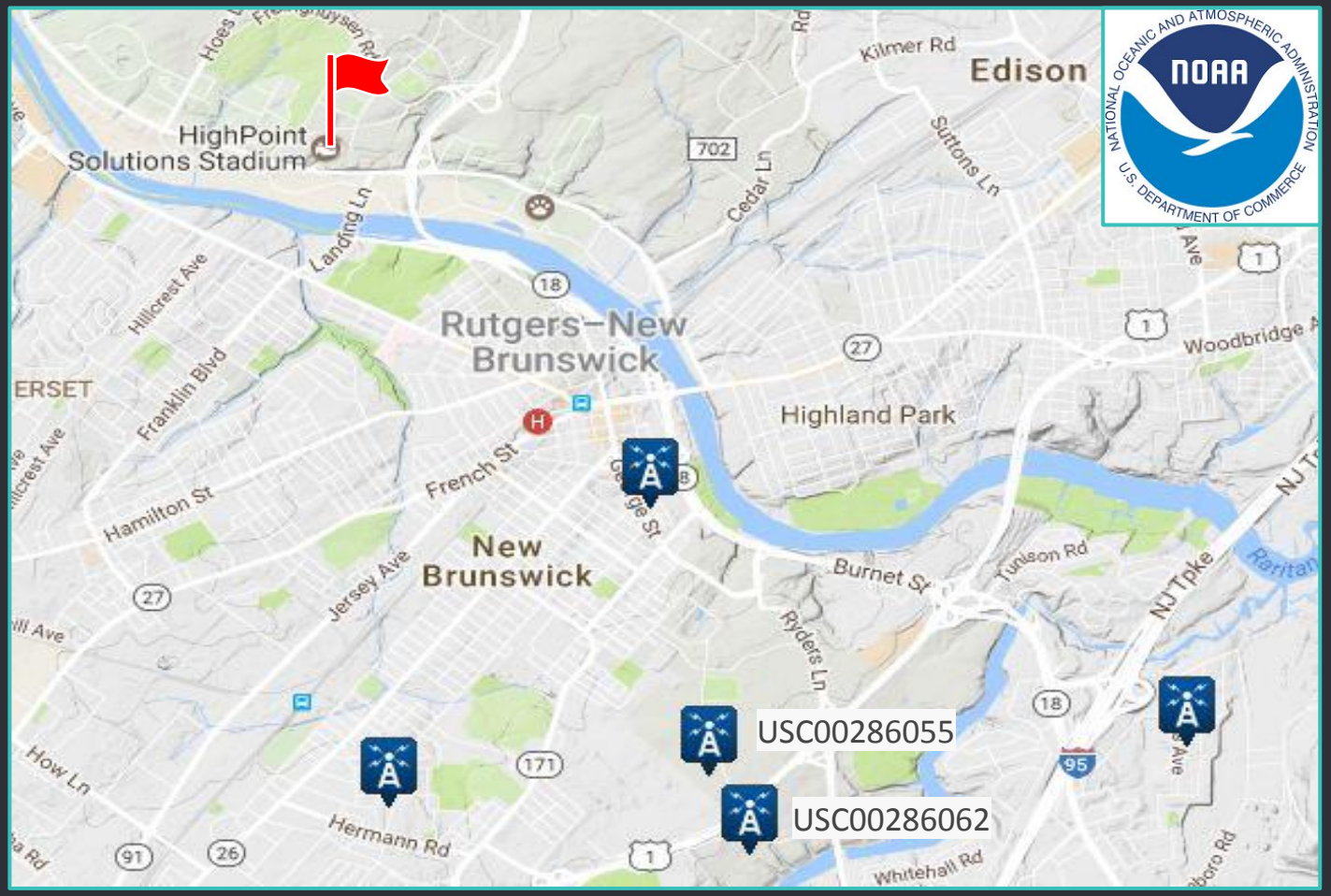


It **will NOT rain** if value of PRCP is closer to 0.

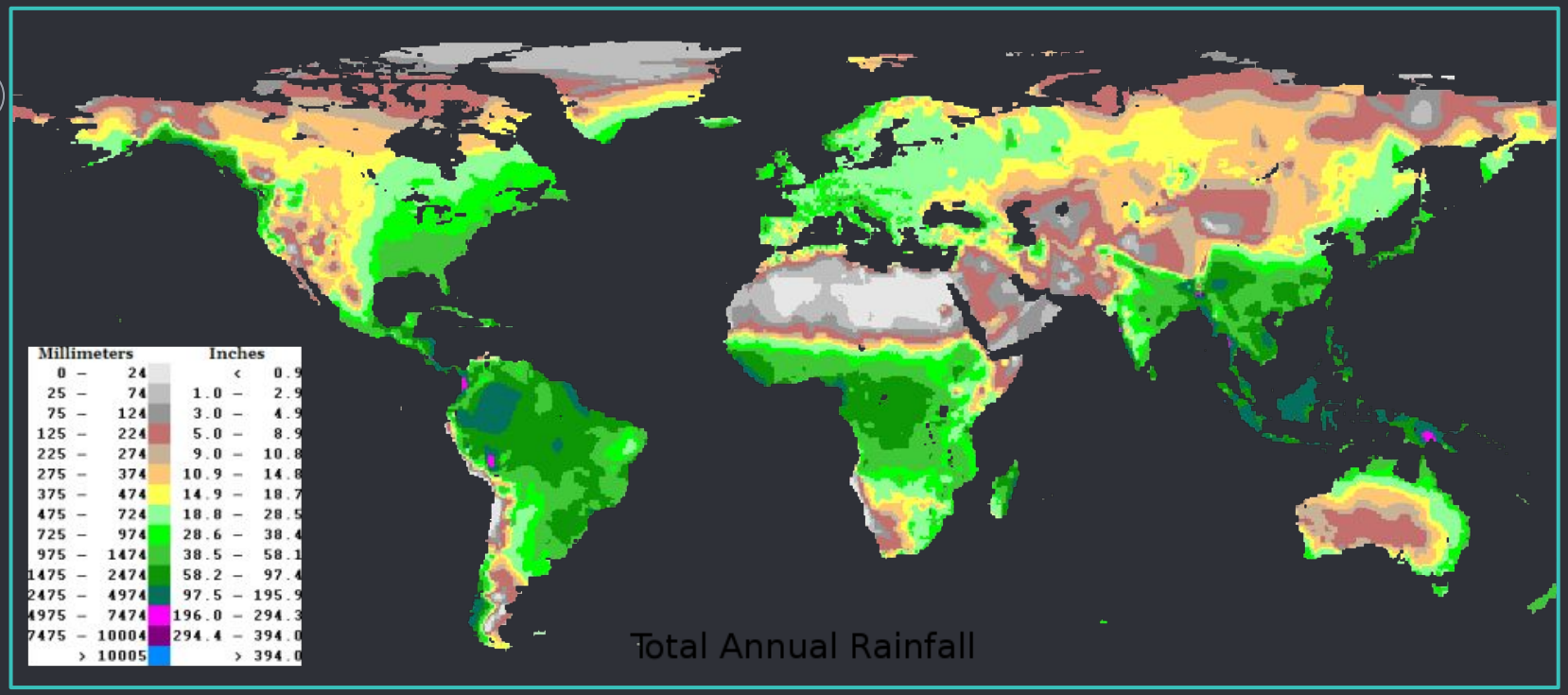
Data source

- “NOAA’-- National Oceanic and Atmospheric Administration
- Global historical climate data and year available from 1763 to 2017.
- Public access to the weather data and information via ftp access
- The Global Historical Climatology Network (GHCN) is an database of climate summaries from weather stations across the globe

Location - New Brunswick



What the experts say

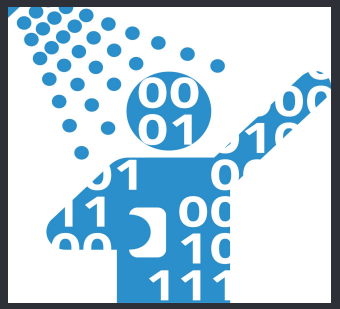


- Latitude
- Altitude
- Temperature
- Humidity

Data Collection and Process

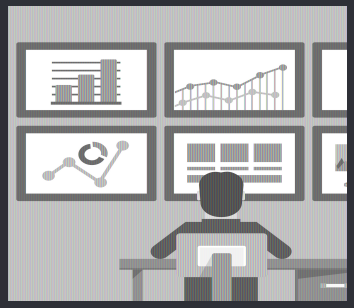
Aggregate Data

- Pulled from NOAA
- Unix to mine station data
- Python to clean



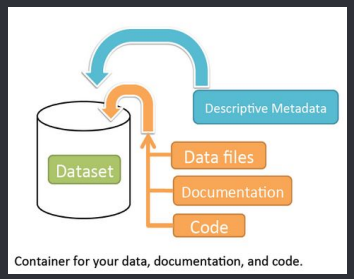
Prioritize Metrics

- PRCP (Precipitation)
- TMAX (Temp. Max)
- TMIN (Temp. Min.)
- EVAP (Evaporation)
- TOBS (Temp. at the time)



Establish Metric Indexes

- Last 100 year data set
- Past 25 year data set



100 years

Weather parameters of the date May 14th for 100 years

25 years

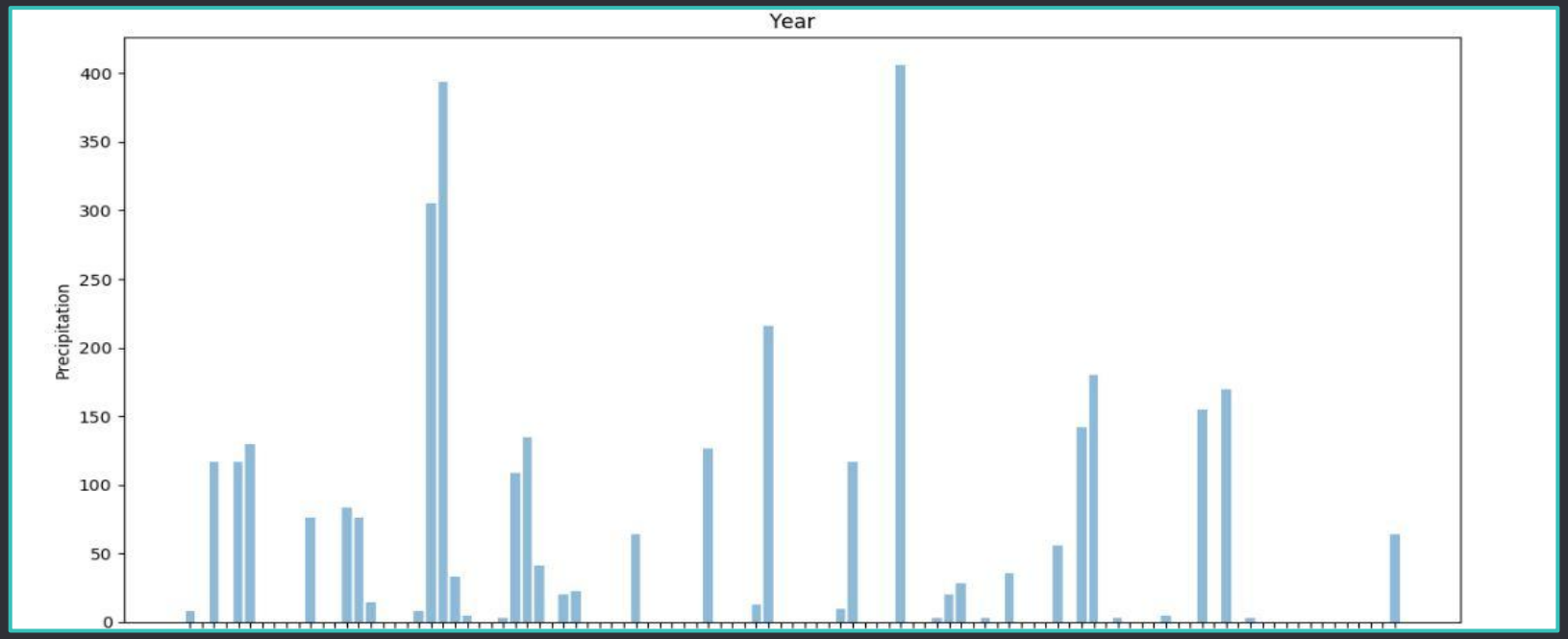
Weather parameters of the most recent 25 years

25 years

Weather parameters of the projected precipitation for 25 years

For the past 100 years, May 14th looks like...

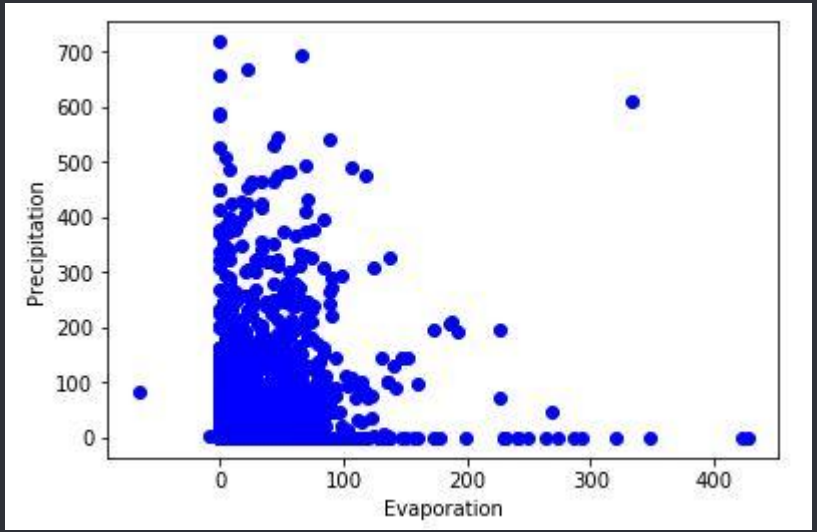
Mean = 3.5mm Median = 0

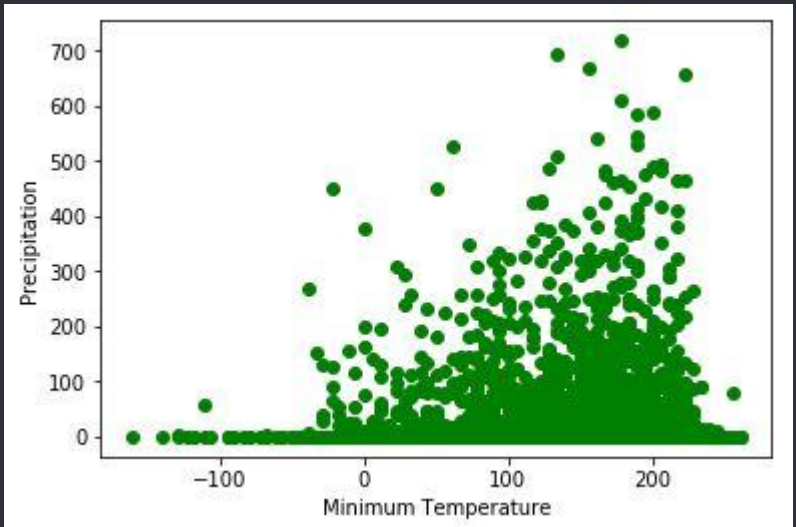


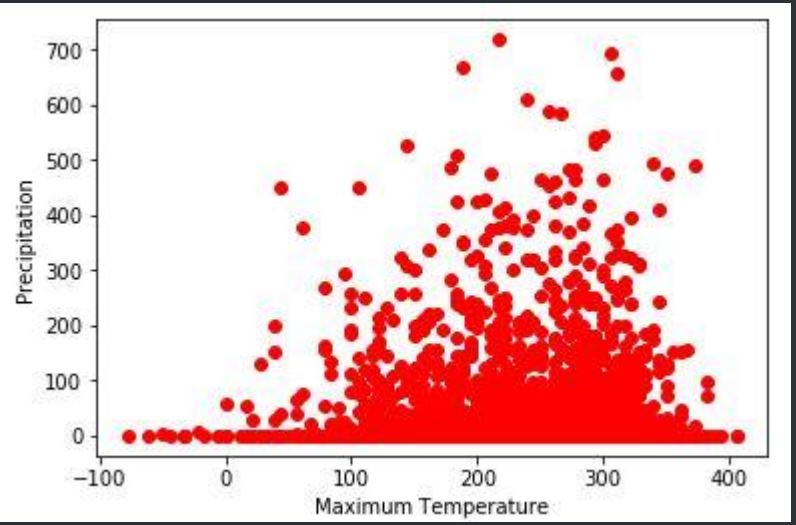
-Unevenly distributed data

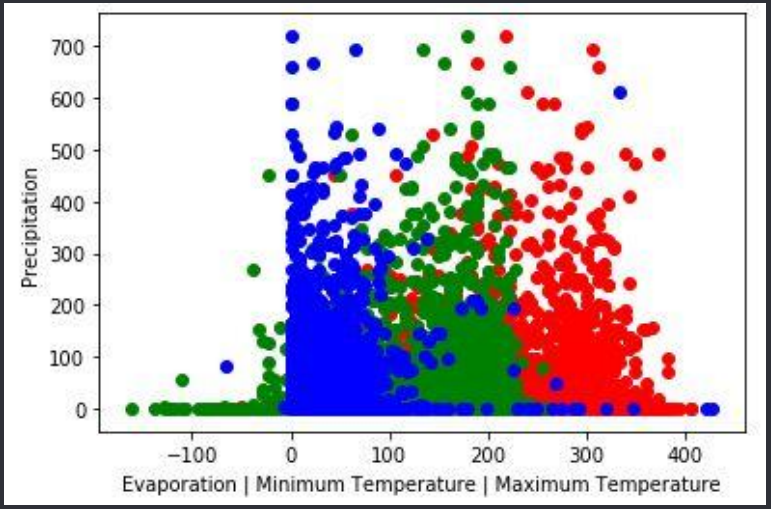
-At least 50% of these years have no precipitation on May 14th

Relationship between the parameters









Feature Selection and Linear Regression

Correlation for 25yrs	EVAP	TMAX	TMIN	TOBS
PRCP (Feature selection)	-0.05	-0.05	0.10	0.04
PRCP (Projected model)	0.03	0.02	0.006	0.007

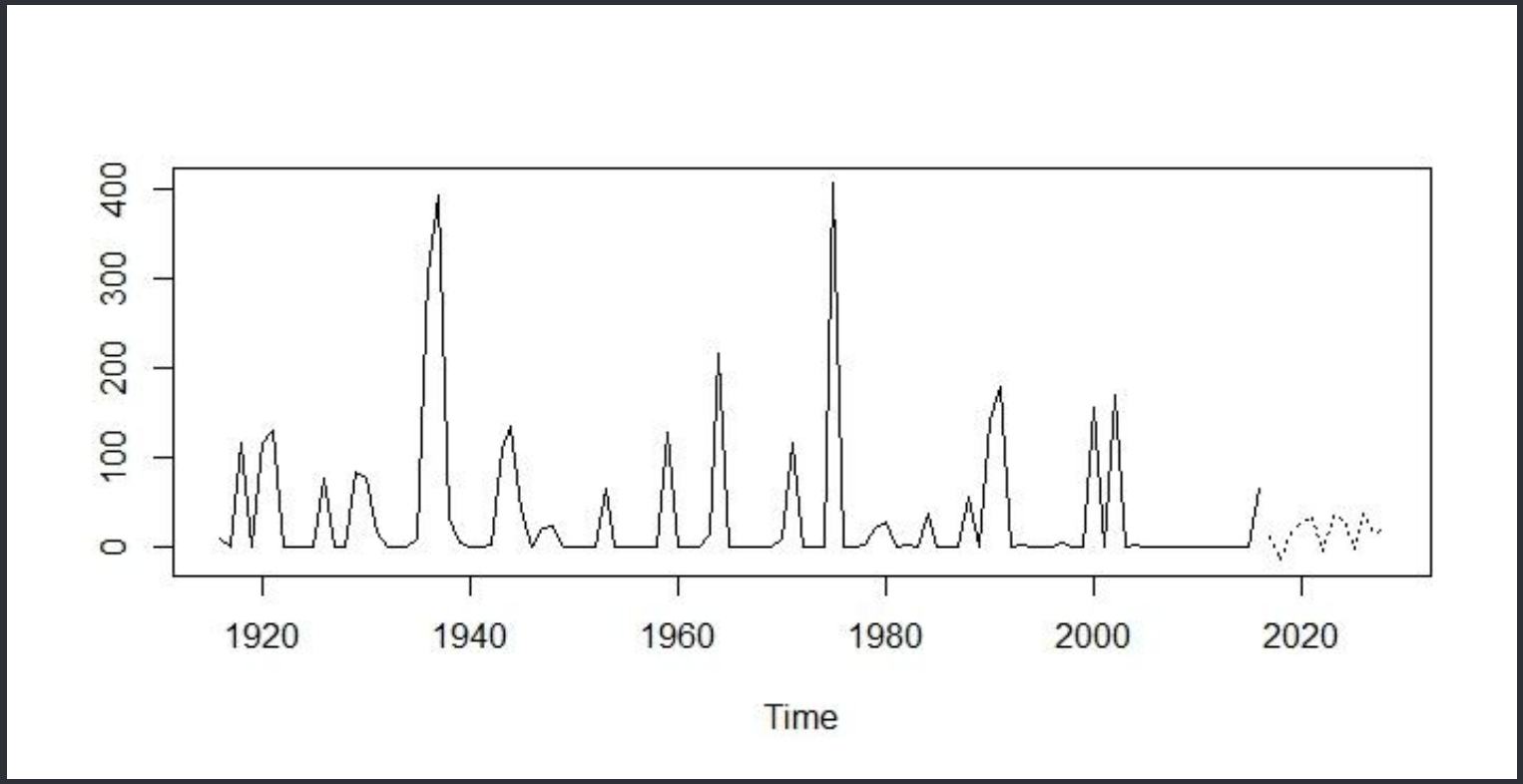
PRCP = **0.07** (EVAP) + **0.06** (TMAX) - **0.05** (TMIN) + **23.67**

Based on the parameters from 30th April...

3.5 mm

of expected precipitation on May 14th 2017.

Time Series Analysis



Past trends

Less than 50% chance for rain

Linear Regression

3.5mm of rain

Time Series Analysis

Closer to 0

Enjoy the sun, Class of 2017!!