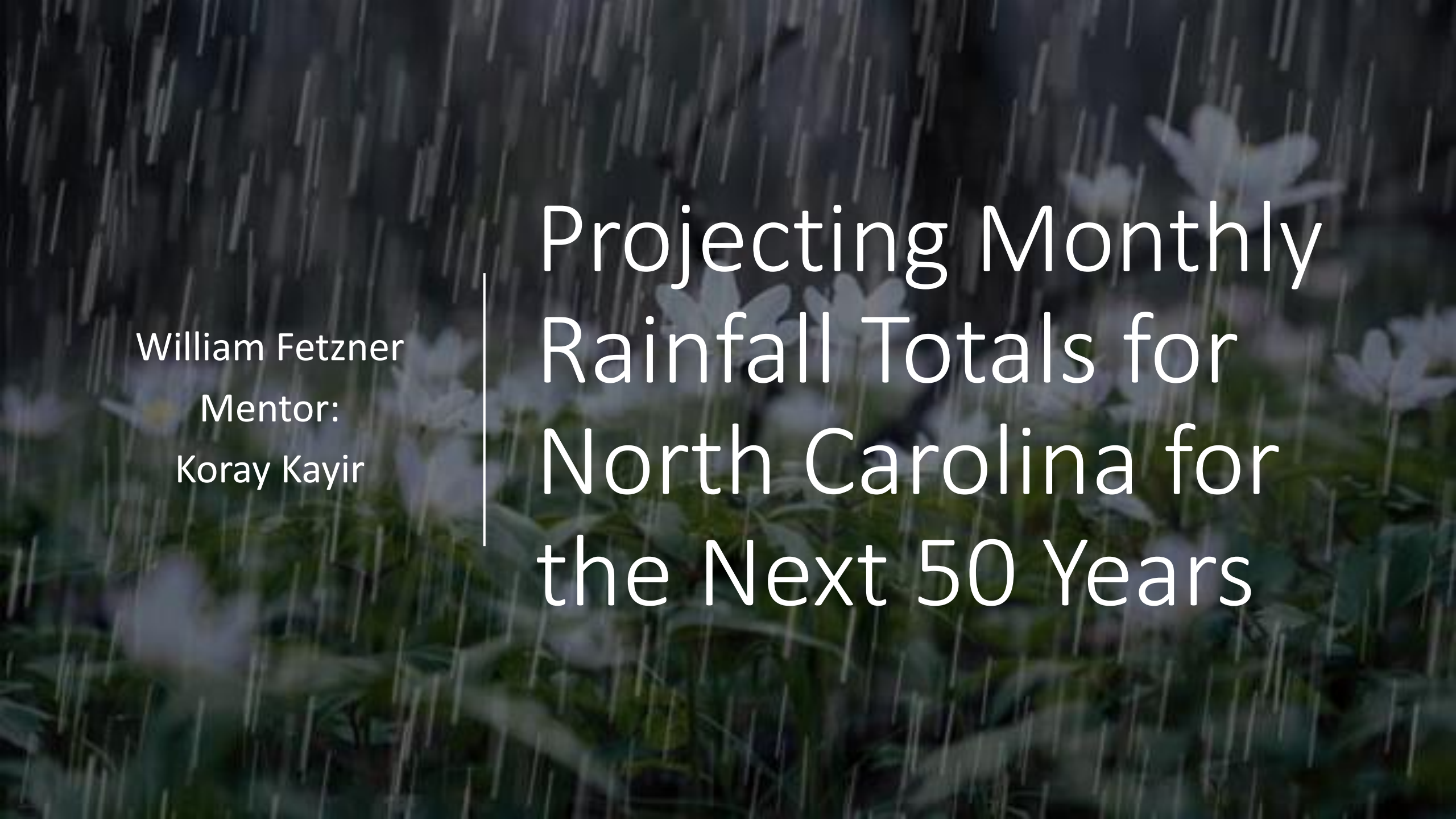


Projecting Monthly Rainfall Totals for North Carolina for the Next 50 Years

William Fetzner
Mentor:
Koray Kayir

Introduction: Question

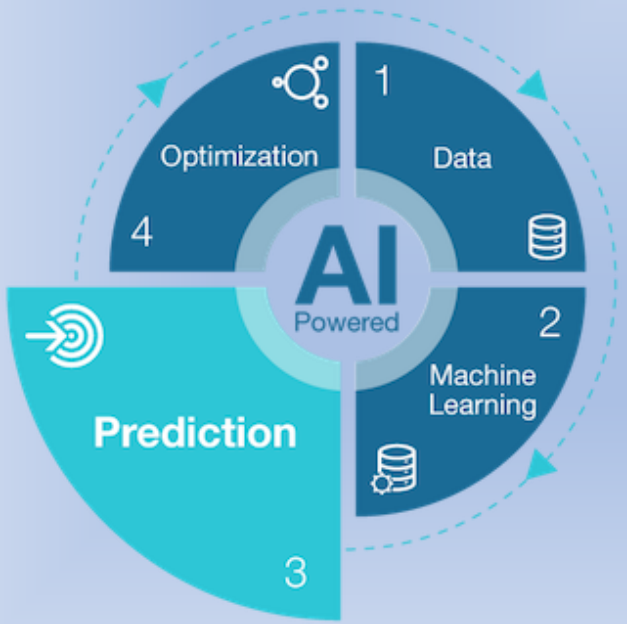
- Is the average amount of rain changing?
- Will it change over the next 50 years?

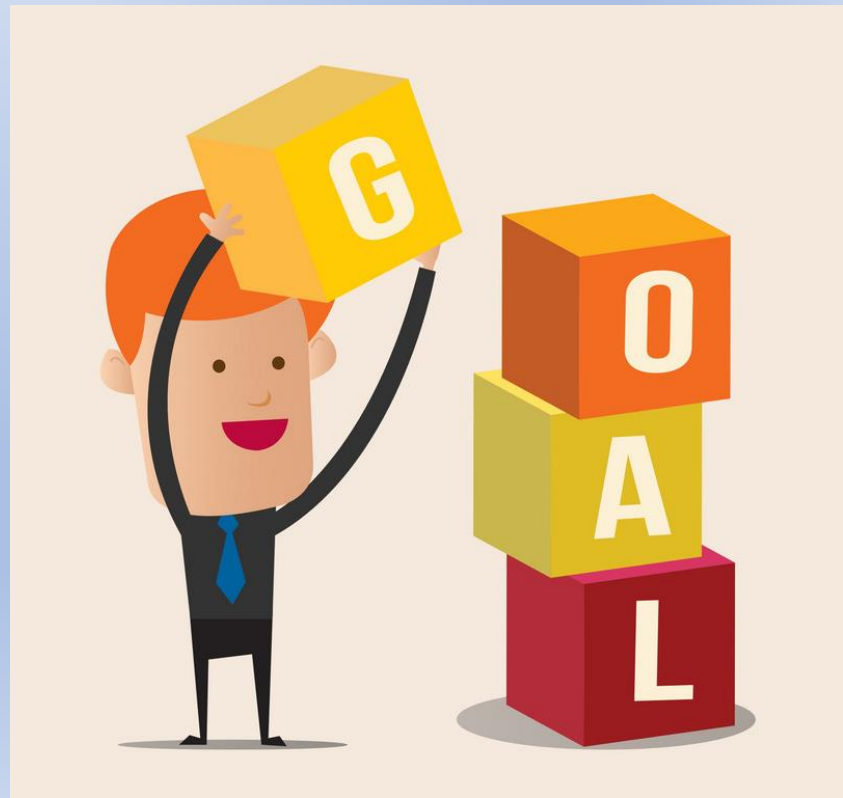
Why is this important?

- Agricultural industry
  - What crops to grow during a season
  - Whether to switch to a different crop
- Governmental planning
  - Building an additional reservoir
  - Implementing water restrictions

Goal

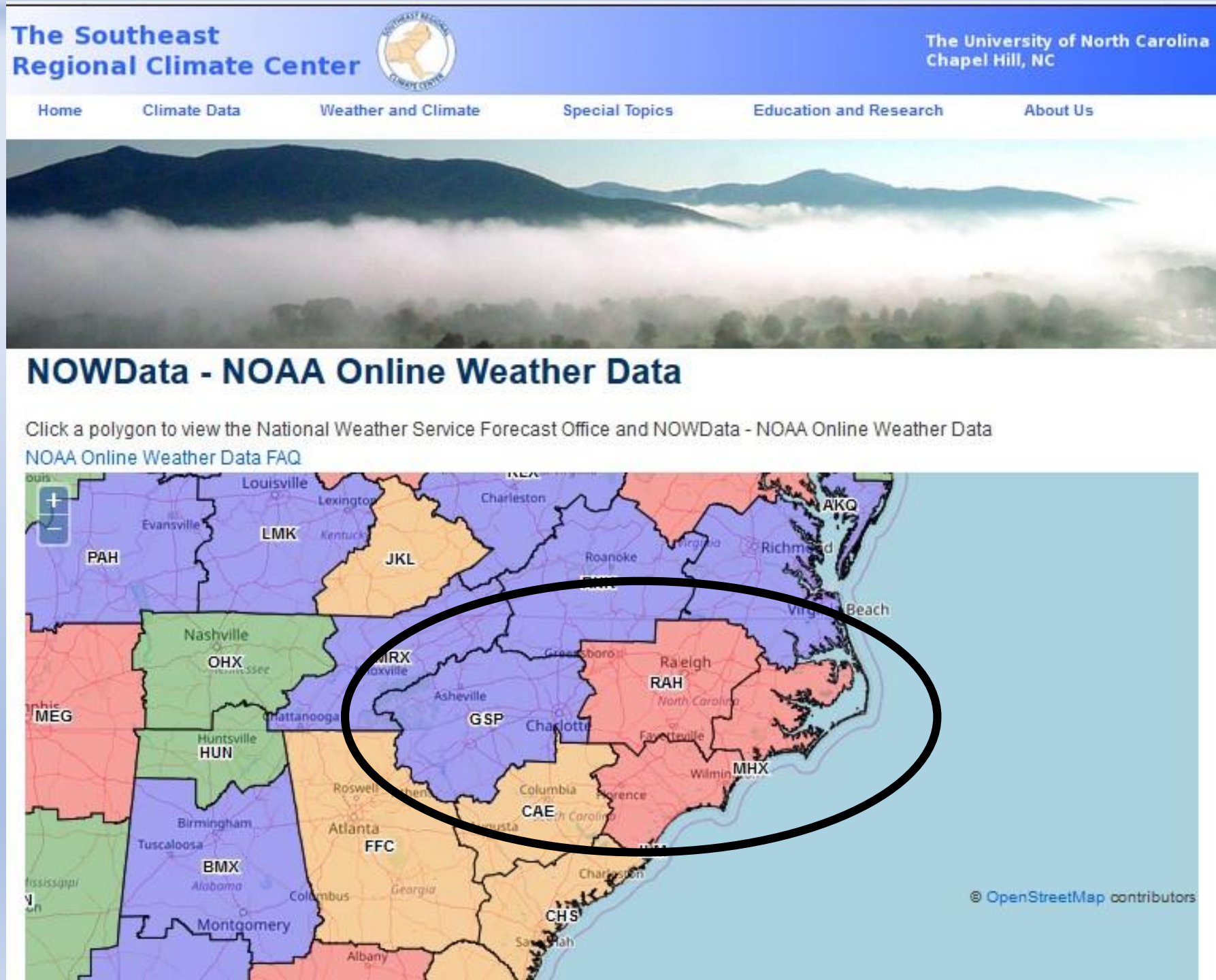
- To predict rainfall amounts for North Carolina for the next 50 years





Data Source

- Gathered from NOWData NOAA website
- The circled area on the map is where the data was collected from
- 112 NC locations and 124 locations from SC, GA, TN, and VA



Data Cleaning

- Poor original format for analysis

Monthly Total Precipitation for ATLANTIC BEACH WATER PLANT, NC

Click column heading to sort ascending, click again to sort descending.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2004	M	M	M	M	M	M	4.89	9.88	5.70	1.71	4.06	1.94	M
2005	2.06	2.38	2.23	4.90	9.29	10.57	8.71	5.53	9.58	18.34	2.40	4.67	80.66
2006	3.17	3.01	1.66	2.33	3.22	6.80	2.96	5.24	8.36	7.67	5.93	3.74	54.09
2007	4.93	2.36	1.04	1.68	3.09	3.99	5.60	3.49	7.56	6.46	1.63	6.36	48.19
2008	2.09	4.40	2.75	7.11	4.52	3.31	2.89	2.11	3.58	1.54	8.57	5.29	48.16
2009	2.66	2.21	1.21	2.77	4.82	2.93	3.36	6.31	14.99	2.48	6.20	7.65	57.59
2010	5.23	3.47	5.93	1.61	1.94	1.45	5.19	8.34	12.44	3.87	1.21	4.09	54.77
2011	4.89	3.36	3.37	2.52	0.58	2.84	3.24	8.74	M	M	2.66	1.24	M
2012	2.74	2.36	M	3.04	4.96	1.27	3.26	9.69	4.82	3.05	0.92	5.64	M
2013	1.90	5.89	1.99	6.07	1.38	4.95	5.71	4.34	4.76	4.22	M	4.89	M
2014	4.08	3.62	4.29	4.74	2.13	4.62	9.95	16.55	14.16	2.54	3.35	1.90	71.93
2015	5.09	5.34	4.33	1.52	3.48	3.14	10.11	8.79	12.32	9.42	12.17	4.96	80.67
2016	5.98	7.20	3.08	2.06	6.37	8.09	2.83	3.16	15.53	4.29	1.24	6.80	66.63
2017	4.31	2.12	3.47	5.06	5.38	2.69	3.76	6.92	6.43	4.14	1.30	4.41	49.99
2018	7.58	1.21	2.69	6.81	6.48	7.17	10.46	2.93	21.50	3.93	7.92	6.87	85.55
2019	3.08	5.83	3.17	3.70	1.93	4.68	4.45	7.24	M	M	M	M	M
Mean	3.99	3.65	2.94	3.73	3.97	4.57	5.46	6.83	10.12	5.26	4.25	4.70	63.48
Max	7.58 2018	7.20 2016	5.93 2010	7.11 2008	9.29 2005	10.57 2005	10.46 2018	16.55 2014	21.50 2018	18.34 2005	12.17 2015	7.65 2009	85.55 2018
Min	1.90 2013	1.21 2018	1.04 2007	1.52 2015	0.58 2011	1.27 2012	2.83 2016	2.11 2008	3.58 2008	1.54 2008	0.92 2012	1.24 2011	48.16 2008

Data Cleaning

- Poor original format for analysis
- Removal of totals

Own column for months

Monthly Total Precipitation for ATLANTIC BEACH WATER PLANT, NC

Click column heading to sort ascending, click again to sort descending

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2004	M	M	M	M	M	M	4.89	9.88	5.70	1.71	4.06	1.94	M
2005	2.06	2.38	2.23	4.90	9.29	10.57	8.71	5.53	9.58	18.34	2.40	4.67	80.66
2006	3.17	3.01	1.66	2.33	3.22	6.80	2.96	5.24	8.36	7.67	5.93	3.74	54.09
2007	4.93	2.36	1.04	1.68	3.09	3.99	5.60	3.49	7.56	6.46	1.63	6.36	48.19
2008	2.09	4.40	2.75	7.11	4.52	3.31	2.89	2.11	3.58	1.54	8.57	5.29	48.16
2009	2.66	2.21	1.21	2.77	4.82	2.93	3.36	6.31	14.99	2.48	6.20	7.65	57.59
2010	5.23	3.47	5.93	1.61	1.94	1.45	5.19	8.34	12.44	3.87	1.21	4.09	54.77
2011	4.89	3.36	3.37	2.52	0.58	2.84	3.24	8.74	M	M	2.66	1.24	M
2012	2.74	2.36	M	3.04	4.96	1.27	3.26	9.69	4.82	3.05	0.92	5.64	M
2013	1.90	5.89	1.99	6.07	1.38	4.95	5.71	4.34	4.76	4.22	M	4.89	M
2014	4.08	3.62	4.29	4.74	2.13	4.62	9.95	16.55	14.16	2.54	3.35	1.90	71.93
2015	5.09	5.34	4.33	1.52	3.48	3.14	10.11	8.79	12.32	9.42	12.17	4.96	80.67
2016	5.98	7.20	3.08	2.06	6.37	8.09	2.83	3.16	15.53	4.29	1.24	6.80	66.63
2017	4.31	2.12	3.47	5.06	5.38	2.69	3.76	6.92	6.43	4.14	1.30	4.41	49.99
2018	7.58	1.21	2.69	6.81	6.48	7.17	10.46	2.93	21.50	3.93	7.92	6.87	85.55
2019	3.08	5.83	3.17	3.70	1.93	4.68	4.45	7.24	M	M	M	M	M
Mean	3.99	3.65	2.94	3.73	3.97	4.57	5.46	6.83	10.12	5.26	4.25	4.79	65.48
Max	7.58 2018	7.20 2016	5.93 2010	7.11 2008	9.29 2005	10.57 2003	10.46 2018	16.55 2014	21.50 2018	18.34 2005	12.17 2015	7.65 2009	85.55 2018
Min	1.90 2013	1.21 2018	1.04 2007	1.52 2015	0.58 2011	1.27 2012	2.83 2016	2.11 2008	3.58 2008	1.54 2008	0.92 2012	1.24 2011	48.16 2008

Data Cleaning

- Poor original format for analysis
- Removal of totals
- MISSING DATA!!!

Monthly Total Precipitation for ATLANTIC BEACH WATER PLANT, NC

Click column heading to sort ascending, click again to sort descending.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2004	M	M	M	M	M	M	4.89	9.88	5.70	1.71	4.06	1.94	M
2005	2.06	2.38	2.23	4.90	9.29	10.57	8.71	5.53	9.58	18.34	2.40	4.67	80.66
2006	3.17	3.01	1.66	2.33	3.22	6.80	2.96	5.24	8.36	7.67	5.93	3.74	54.09
2007	4.93	2.36	1.04	1.68	3.09	3.99	5.60	3.49	7.56	6.46	1.63	6.36	48.19
2008	2.09	4.40	2.75	7.11	4.52	3.31	2.89	2.11	3.58	1.54	8.57	5.29	48.16
2009	2.66	2.21	1.21	2.77	4.82	2.93	3.36	6.31	14.99	2.48	6.20	7.65	57.59
2010	5.23	3.47	5.93	1.61	1.94	1.45	5.19	8.34	12.44	3.87	1.21	4.09	54.77
2011	4.89	3.36	3.37	2.52	0.58	2.84	3.24	8.74	M	M	2.66	1.24	M
2012	2.74	2.36	M	3.04	4.96	1.27	3.26	9.69	4.82	3.05	0.92	5.64	M
2013	1.90	5.89	1.99	6.07	1.38	4.95	5.71	4.34	4.76	4.22	M	4.89	M
2014	4.08	3.62	4.29	4.74	2.13	4.62	9.95	16.55	14.16	2.54	3.35	1.90	71.93
2015	5.09	5.34	4.33	1.52	3.48	3.14	10.11	8.79	12.32	9.42	12.17	4.96	80.67
2016	5.98	7.20	3.08	2.06	6.37	8.09	2.83	3.16	15.53	4.29	1.24	6.80	66.63
2017	4.31	2.12	3.47	5.06	5.38	2.69	3.76	6.92	6.43	4.14	1.30	4.41	49.99
2018	7.58	1.21	2.69	6.81	6.48	7.17	10.46	2.93	21.50	3.93	7.92	6.87	85.55
2019	3.08	5.83	3.17	3.70	1.93	4.68	4.45	7.24	M	M	M	M	M
Mean	3.99	3.65	2.94	3.73	3.97	4.57	5.46	6.83	10.12	5.26	4.25	4.70	63.48
Max	7.58 2018	7.20 2016	5.93 2010	7.11 2008	9.29 2005	10.57 2005	10.46 2018	16.55 2014	21.50 2018	18.34 2005	12.17 2015	7.65 2009	85.55 2018
Min	1.90 2013	1.21 2018	1.04 2007	1.52 2015	0.58 2011	1.27 2012	2.83 2016	2.11 2008	3.58 2008	1.54 2008	0.92 2012	1.24 2011	48.16 2008

Data Cleaning

Monthly Total Precipitation for Cape Hatteras Area, NC (ThreadEx)

Click column heading to sort ascending, click again to sort descending.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1893	[illegible]	2.49	4.51	3.14	4.51	5.23	3.66	M	4.63	7.16	8.32	3.86	M
1894	[illegible]	4.97	2.92	1.32	4.52	1.38	9.65	7.36	3.73	5.77	3.02	M	M
1895	6.14	[illegible]	7.14	7.72	7.61	4.15	7.62	5.32	5.89	2.84	7.78	3.44	[illegible]
1896	4.05	4.96	[illegible]	1.71	1.50	2.34	6.25	1.44	3.75	5.28	3.86	5.40	44.07
1897	2.00	5.21	3.79	[illegible]	4.08	5.76	9.23	5.46	1.00	9.83	0.98	[illegible]	57.99
1898	1.82	1.75	2.45	4.39	[illegible]	3.27	4.36	6.48	1.59	12.45	[illegible]	3.38	47.87
1899	4.52	8.07	2.18	3.99	2.55	3.16	9.50	14.19	0.61	9.38	2.24	3.47	63.86
1900	4.50	M	4.74	4.84	4.24	2.83	5.42	1.56	0.56	[illegible]	3.81	3.33	M
1901	2.90	2.35	3.81	4.16	3.99	4.84	[illegible]	3.87	[illegible]	0.81	2.17	6.53	50.16
1902	1.41	3.62	2.78	M	M	2.66	M	[illegible]	M	4.85	M	M	M
1903	M	4.82	5.30	1.74	4.43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.52	3.27	M
1904	5.11	3.52	3.85	2.84	2.80	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.37	2.72	40.97
1905	3.04	3.73	M	4.17	M	2.28	4.22	1.65	3.42	3.05	0.80	5.28	M
1906	4.29	7.87	5.39	1.78	0.55	7.94	6.33	5.45	3.00	6.25	1.38	3.71	53.94
1907	1.40	2.89	1.81	3.99	3.10	6.31	2.48	7.96	6.06	0.40	4.28	2.93	43.61
1908	2.85	2.60	5.77	2.46	7.34	6.22	7.55	14.93	3.37	6.98	0.99	4.54	65.60
1909	3.05	3.30	1.80	7.95	1.42	4.17	2.50	4.99	1.35	0.41	2.24	2.15	35.33
1910	2.35	2.60	2.16	0.67	3.02	2.95	1.44	13.14	2.60	1.45	0.38	2.39	35.15
1911	1.08	0.88	5.63	1.64	0.70	2.75	1.03	3.61	5.08	2.43	1.05	3.76	29.64
1912	4.54	4.38	3.74	2.55	4.04	7.83	1.55	2.80	2.74	8.64	3.04	3.13	48.98
1913	3.62	2.54	3.83	2.28	3.08	1.62	3.78	4.97	7.28	6.93	0.53	2.94	43.40
1914	1.93	4.32	2.93	4.14	0.67	2.58	2.88	2.66	3.33	5.78	1.63	4.71	37.56
1915	6.05	1.98	1.11	3.99	4.51	4.62	4.75	3.30	1.26	3.40	1.22	1.80	37.99
1916	1.26	5.05	2.30	2.82	1.98	3.89	4.93	4.54	2.73	4.16	2.12	2.10	37.88
1917	3.12	1.23	2.39	2.08	0.85	8.02	3.83	2.48	14.19	1.23	2.43	3.07	44.92
1918	1.78	1.73	1.34	5.34	1.48	4.69	7.10	3.10	2.41	1.12	2.30	4.58	36.97

Dates are way off

Monthly Total Precipitation for ATLANTIC BEACH WATER PLANT, NC

Click column heading to sort ascending, click again to sort descending.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2004	M	M	M	M	M	M	4.89	9.88	5.70	1.71	4.06	1.94	M
2005	2.06	2.38	2.23	4.90	9.29	10.57	8.71	5.53	9.58	18.34	2.40	4.67	80.66
2006	3.17	3.01	1.66	2.33	3.22	6.80	2.96	5.24	8.36	7.67	5.93	3.74	54.09
2007	4.93	2.36	1.04	1.68	3.09	3.99	5.60	3.49	7.56	6.46	1.63	6.36	48.19
2008	2.09	4.40	2.75	7.11	4.52	3.31	2.89	2.11	3.58	1.54	8.57	5.29	48.16
2009	2.66	2.21	1.21	2.77	4.82	2.93	3.36	6.31	14.99	2.48	6.20	7.65	57.59
2010	5.23	3.47	5.93	1.61	1.94	1.45	5.19	8.34	12.44	3.87	1.21	4.09	54.77
2011	4.89	3.36	3.37	2.52	0.58	2.84	3.24	8.74	M	M	2.66	1.24	M
2012	2.74	2.36	M	3.04	4.96	1.27	3.26	9.69	4.82	3.05	0.92	5.64	M
2013	1.90	5.89	1.99	6.07	1.38	4.95	5.71	4.34	4.76	4.22	M	4.89	M
2014	4.08	3.62	4.29	4.74	2.13	4.62	9.95	16.55	14.16	2.54	3.35	1.90	71.93
2015	5.09	5.34	4.33	1.52	3.48	3.14	10.11	8.79	12.32	9.42	12.17	4.96	80.67
2016	5.98	7.20	3.08	2.06	6.37	8.09	2.83	3.16	15.53	4.29	1.24	6.80	66.63
2017	4.31	2.12	3.47	5.06	5.38	2.69	3.76	6.92	6.43	4.14	1.30	4.41	49.99
2018	7.58	1.21	2.69	6.81	6.48	7.17	10.46	2.93	21.50	3.93	7.92	6.87	85.55
2019	3.08	5.83	3.17	3.70	1.93	4.68	4.45	7.24	M	M	M	M	M
Mean	3.99	3.65	2.94	3.73	3.97	4.57	5.46	6.83	10.12	5.26	4.25	4.70	63.48
Max	7.58 2018	7.20 2016	5.93 2010	7.11 2008	9.29 2005	10.57 2005	10.46 2018	16.55 2014	21.50 2018	18.34 2005	12.17 2015	7.65 2009	85.55 2018
Min	1.90 2013	1.21 2018	1.04 2007	1.52 2015	0.58 2011	1.27 2012	2.83 2016	2.11 2008	3.58 2008	1.54 2008	0.92 2012	1.24 2011	48.16 2008

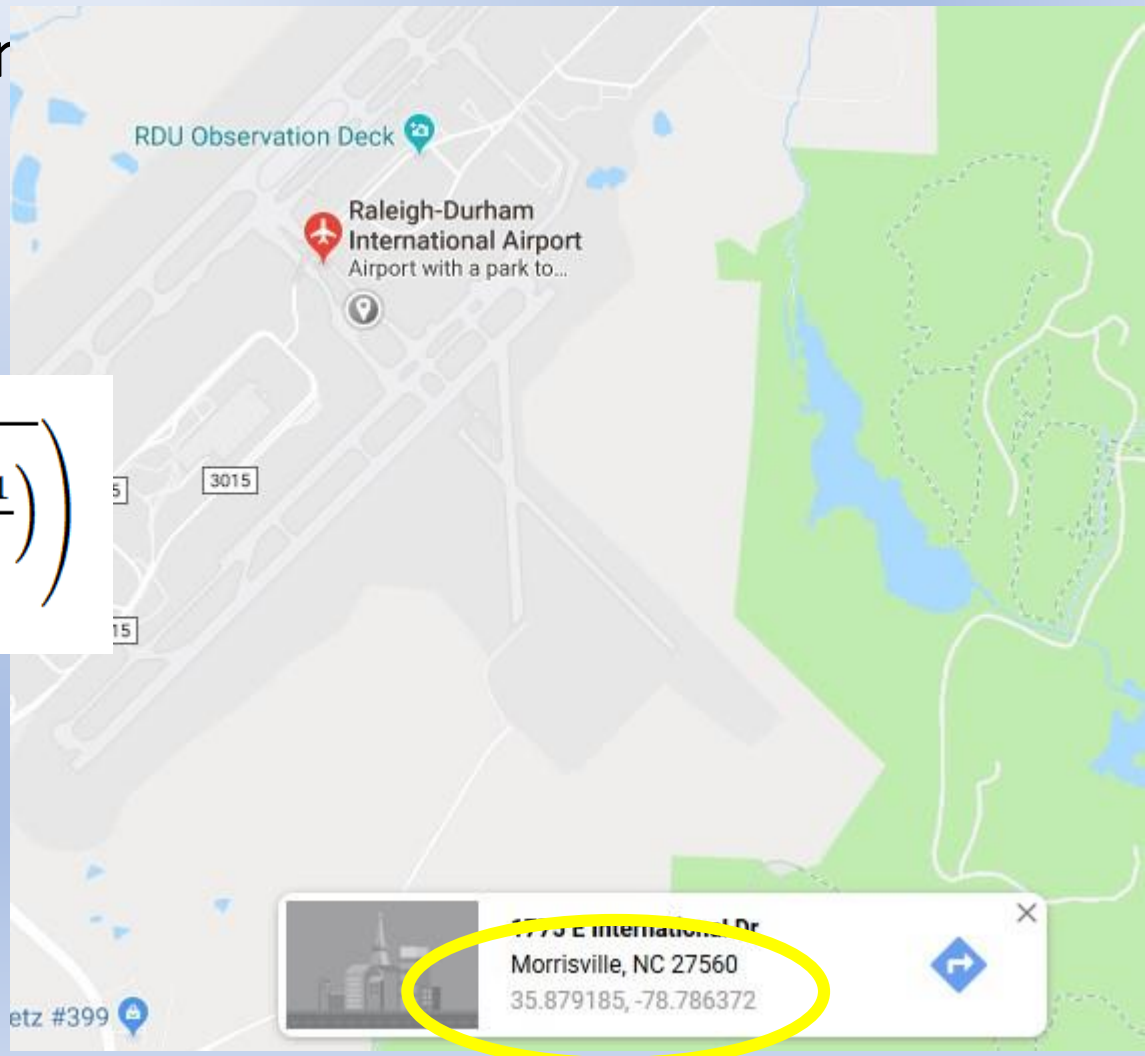
Data Cleaning

- 3 ways of dealing with missing data
  1. Only included data from 1980-2019 (80% of locations had full data)
  2. Estimation of missing values
     1. Utilized the rainfall total from the previous month, the next month and the corresponding month of the previous year and averaged them
     2. Averaged the rainfall for the target month of all locations within 85 kilometers (~ 50 miles)

Where’d the distance come from?

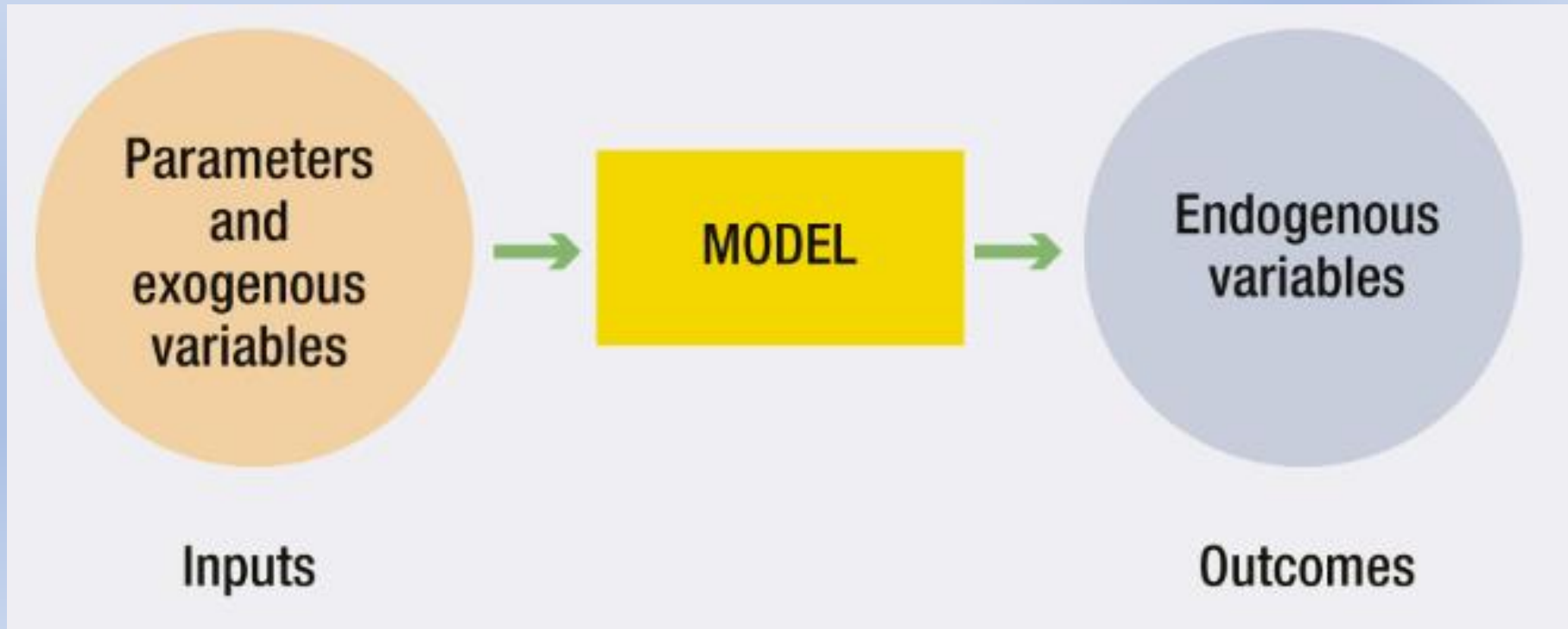
- Latitude and Longitude of each location
- Haversine Formula

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)$$



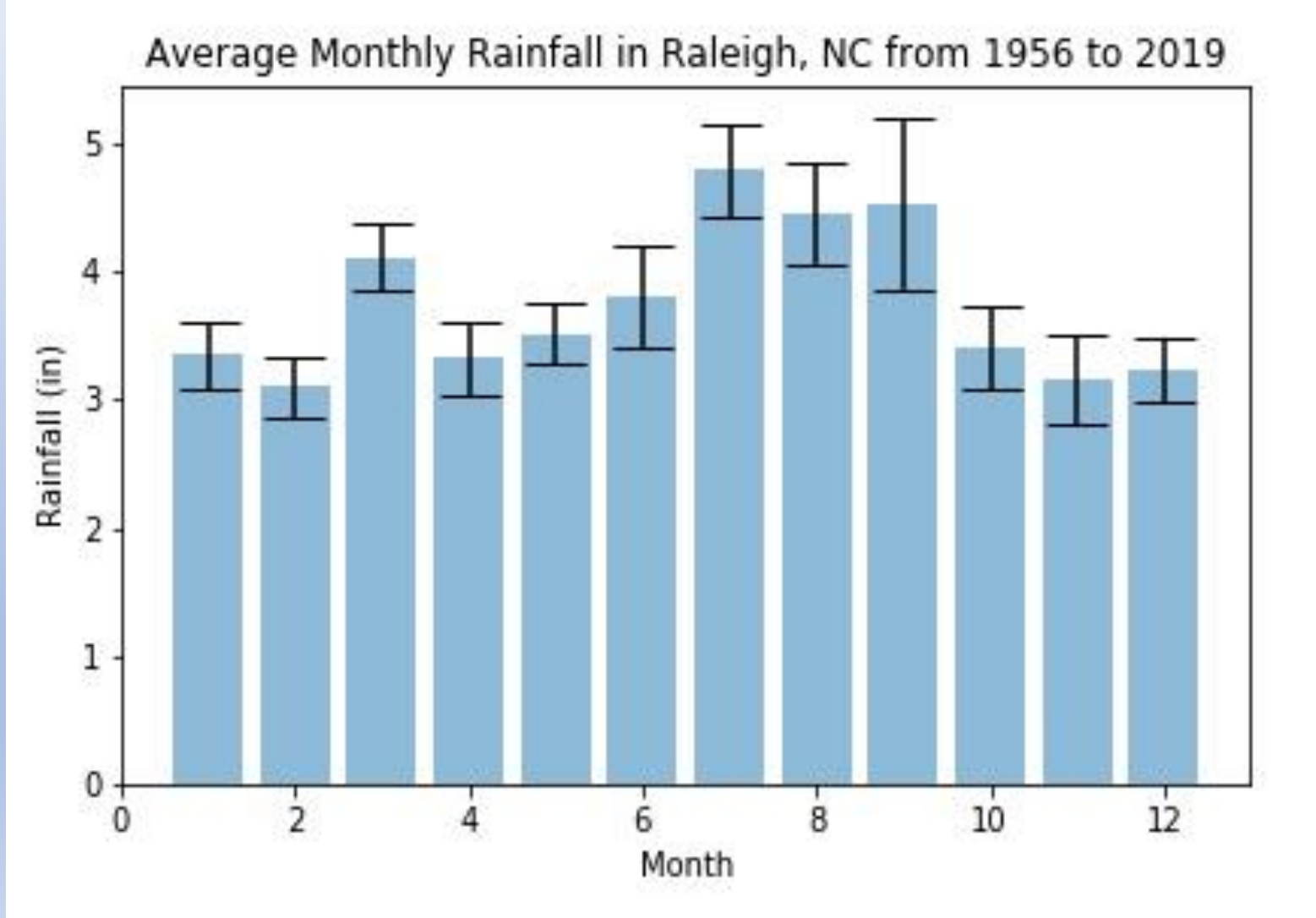
External location pairing

- 46 target locations
- External locations within 50 kilometers
- Target and corresponding external placed into dictionary



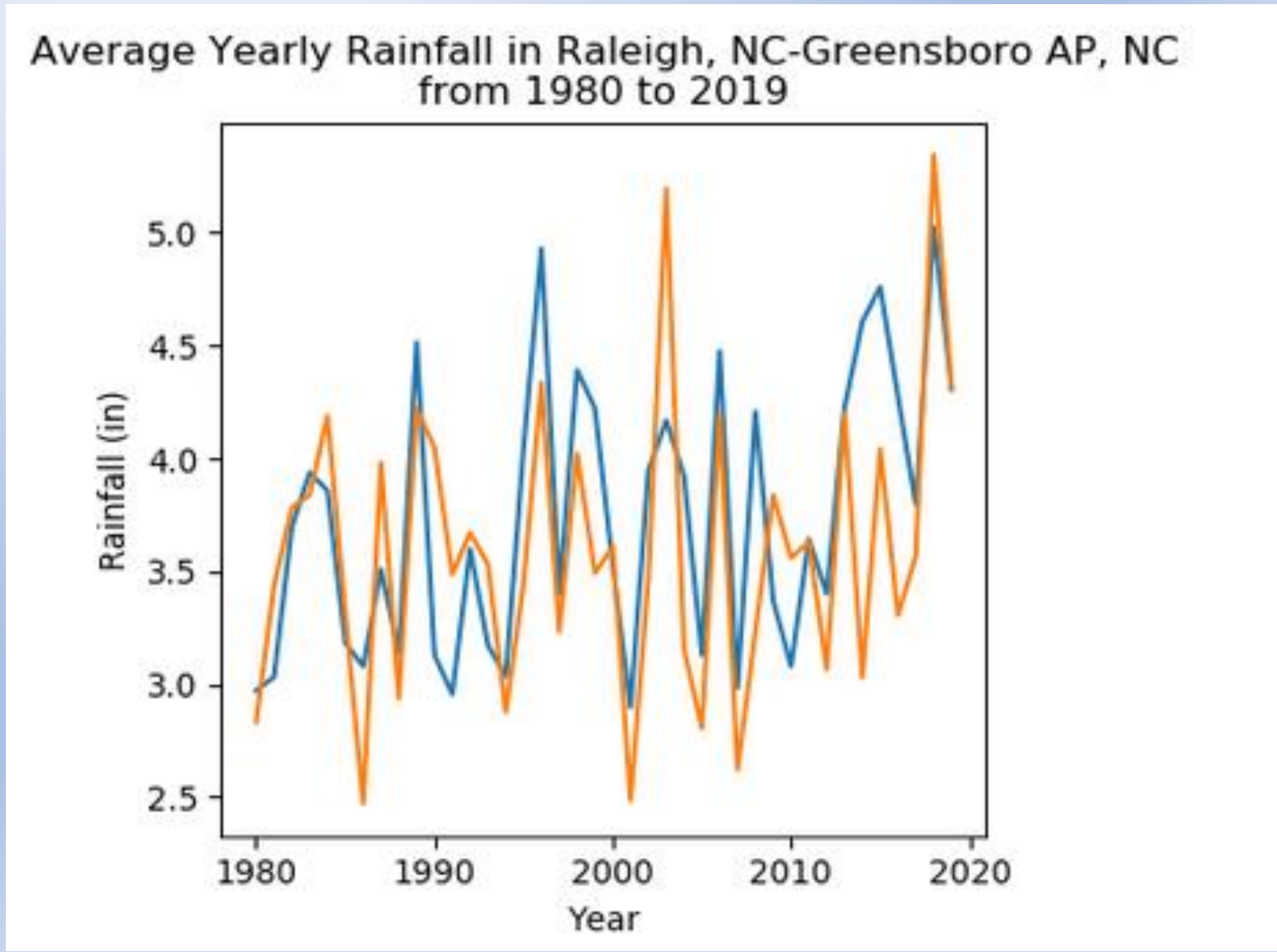
Exploratory Data Analysis

- What does it look like?



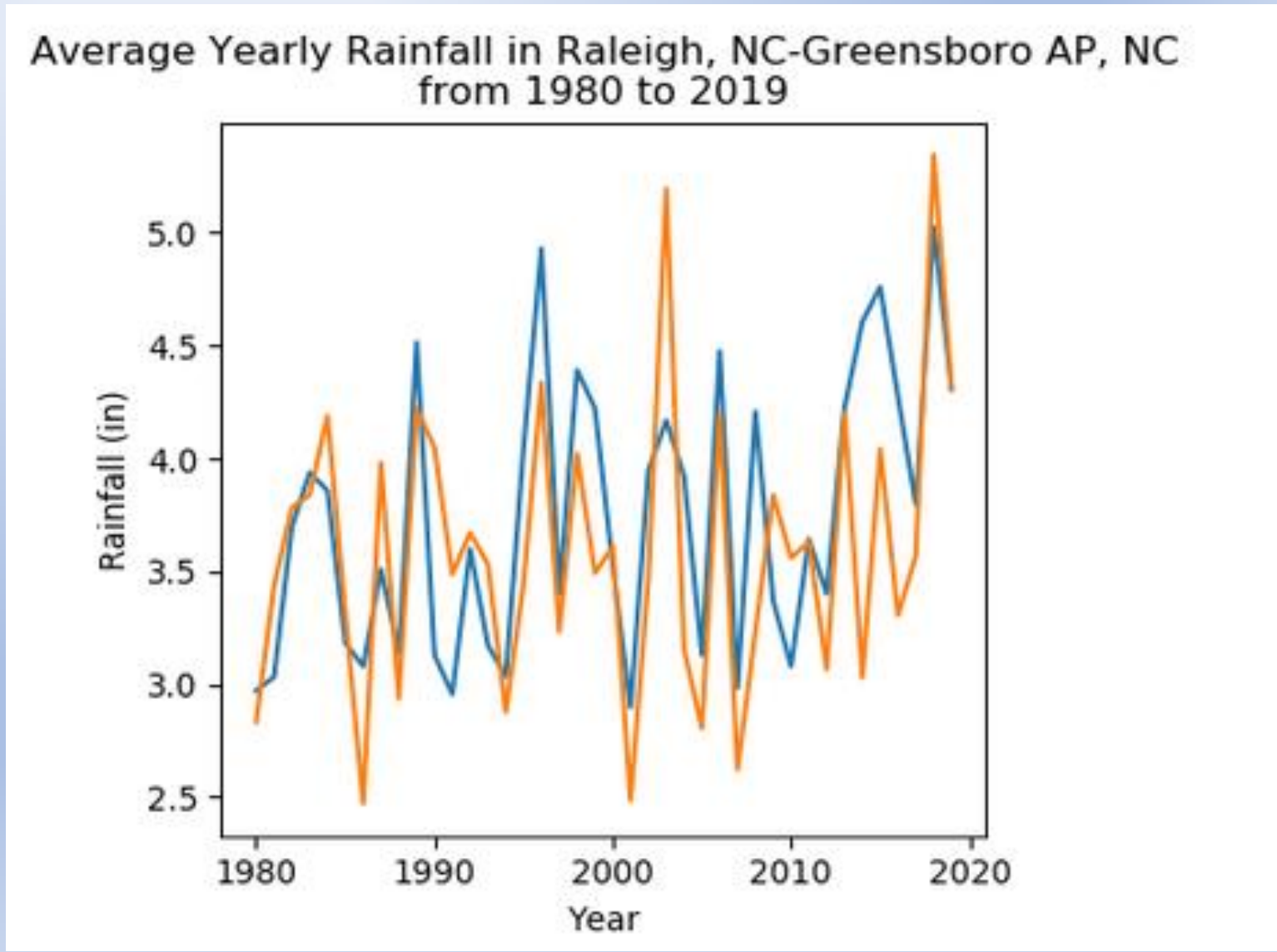
Exploratory Data Analysis

- What does it look like?



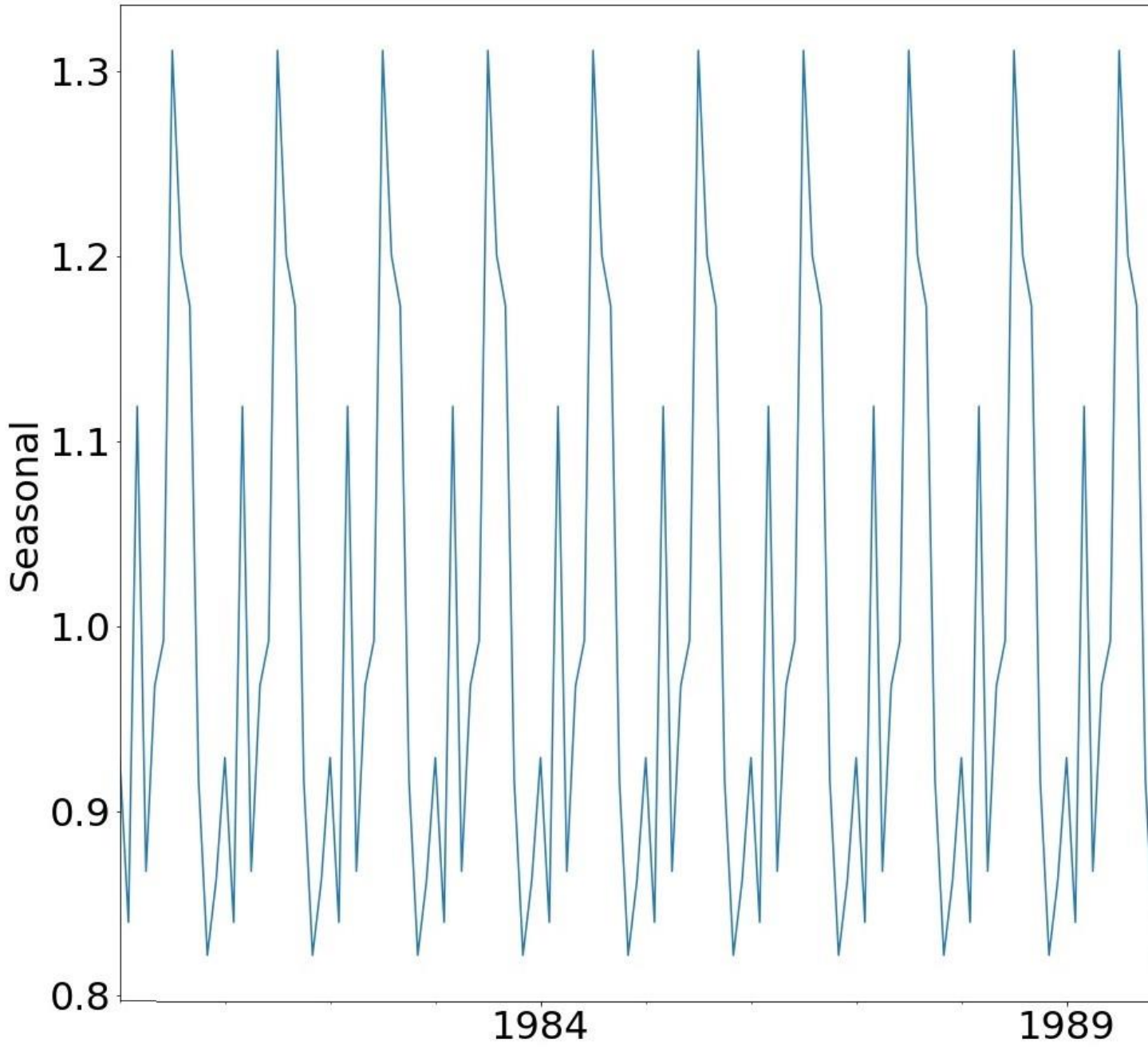
Exploratory Data Analysis

- What does it look like?



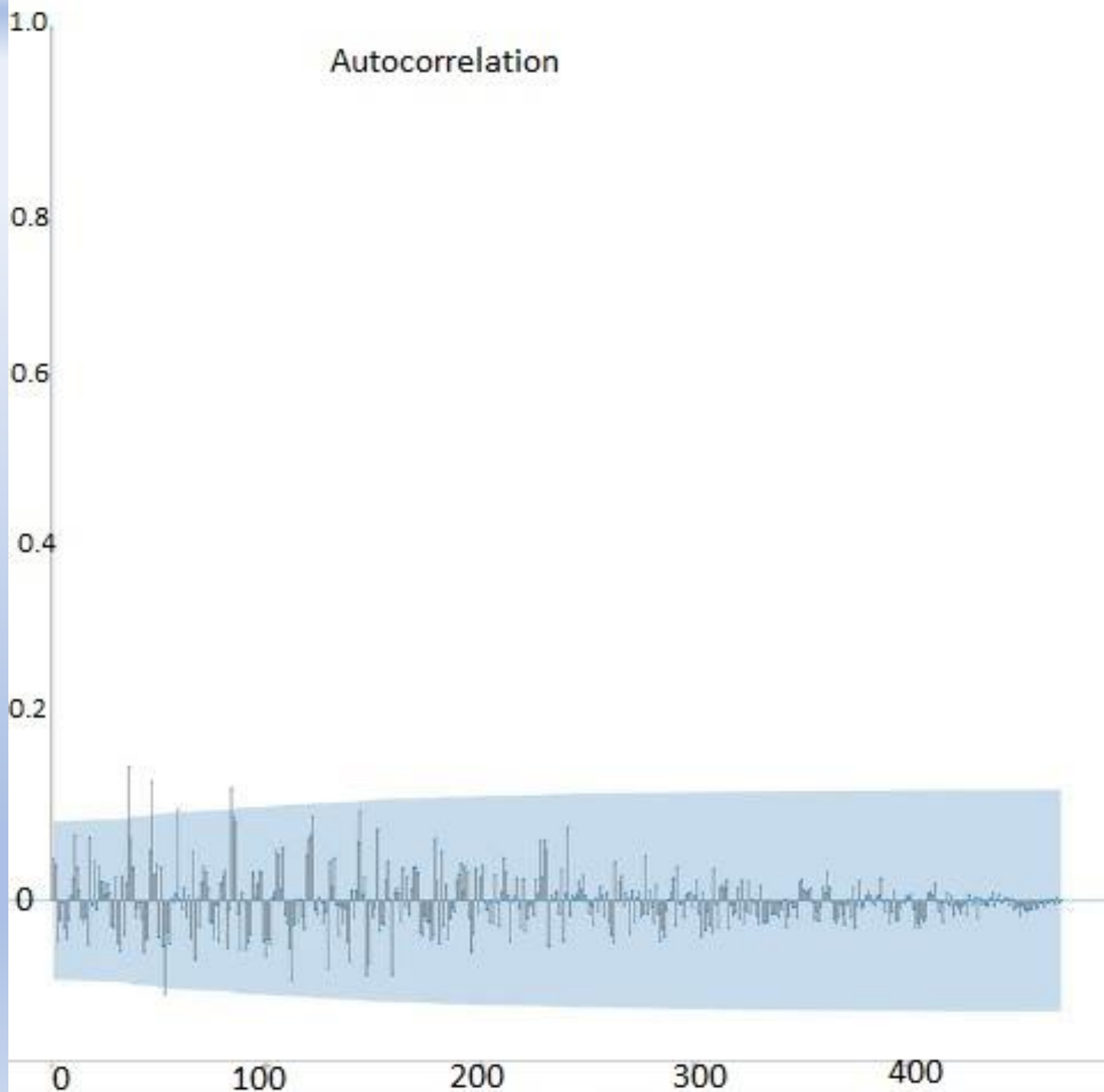
Seasonal Decomposition

- Periodicity was 12 months



Autocorrelation

- Correlation of the location with itself



Autocorrelation

- Lag 1 & lag 12

Date	Raleigh, NC	lag 1	lag 12
1/1/1980	4.39	1.91	0.87
2/1/1980	1.91	5.87	3.02
3/1/1980	5.87	1.97	2.35
4/1/1980	1.97	2.33	1.03
5/1/1980	2.33	4.89	4.28
6/1/1980	4.89	2.11	0.55
7/1/1980	2.11	1.87	5.69
8/1/1980	1.87	3.76	5.34
9/1/1980	3.76	2.25	2.7
10/1/1980	2.25	2.87	4.64
11/1/1980	2.87	1.42	0.95
12/1/1980	1.42	0.87	4.96
1/1/1981	0.87	3.02	3.43
2/1/1981	3.02	2.35	4.97
3/1/1981	2.35	1.03	3.02
4/1/1981	1.03	4.28	3.33
5/1/1981	4.28	0.55	4.2
6/1/1981	0.55	5.69	8.39
7/1/1981	5.69	5.34	3.34
8/1/1981	5.34	2.7	1.83
9/1/1981	2.7	4.64	1.55
10/1/1981	4.64	0.95	3.93

Machine Learning

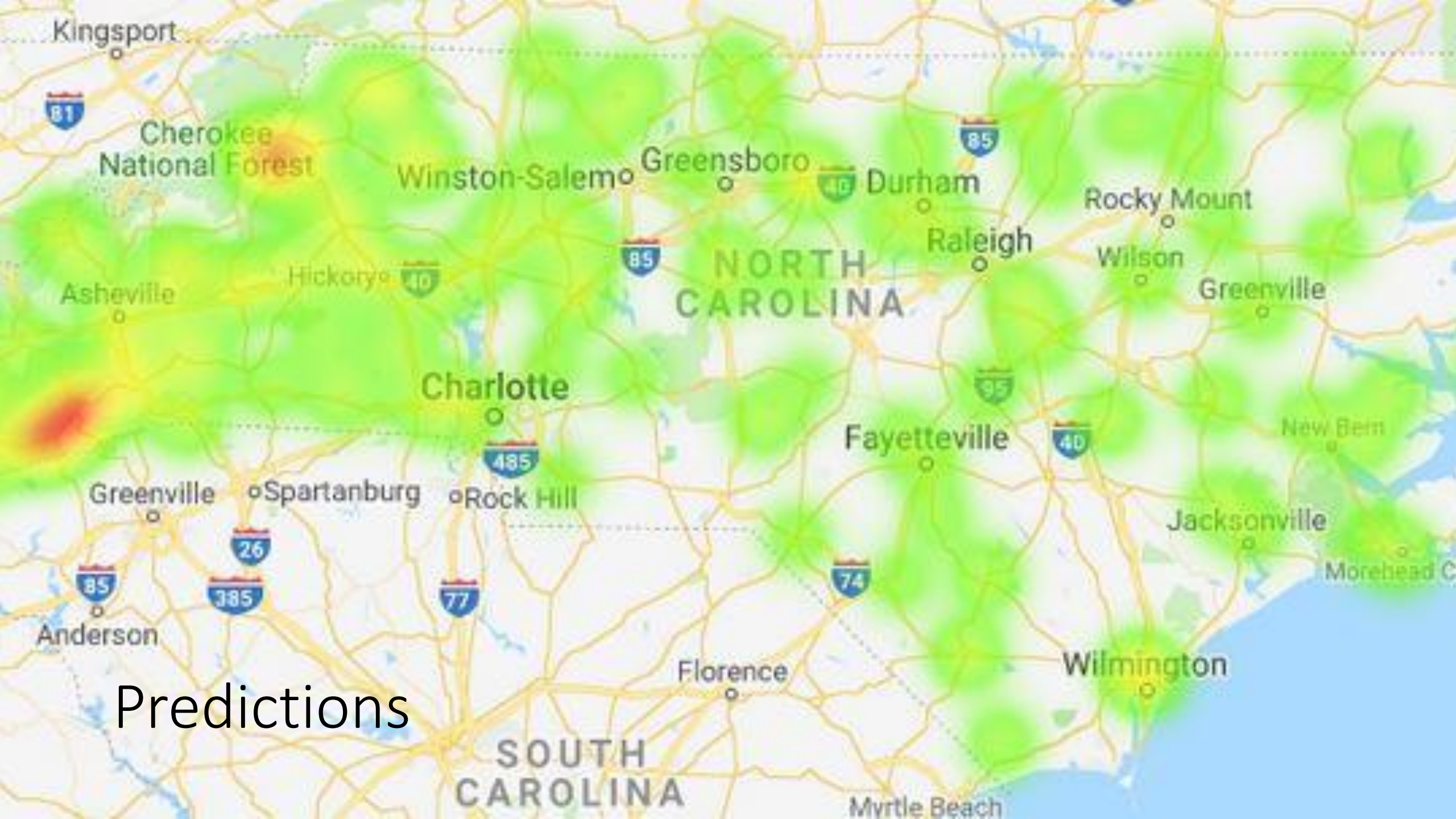
- Seasonal Autoregressive Integrated Moving Average
- SARIMA (p,d,q)(P,D,Q,m)
  - p = order or number of time lags of the autoregressive model
  - d = degree of differencing (number of times the data has had past values subtracted)
  - q = order of the moving-average model
  - P, D, Q = similar just for seasonal part of model
  - m = periodicity of the data

Hyperparameters of SARIMA

- D, d and m were determined from the dataset
  - Trend = 0 thus, D and d = 0
  - m = 12, period shown in seasonal decomposition and autocorrelation
- P, Q, p, and q were tuned
  - Training (80%), Validation (10%), Testing (10%)
  - Mean absolute error
    - Determined performance of model
- RESULT: p = 4, d = 0, q=3, P=3, D=0, Q=4, m=12

Exogenous help?

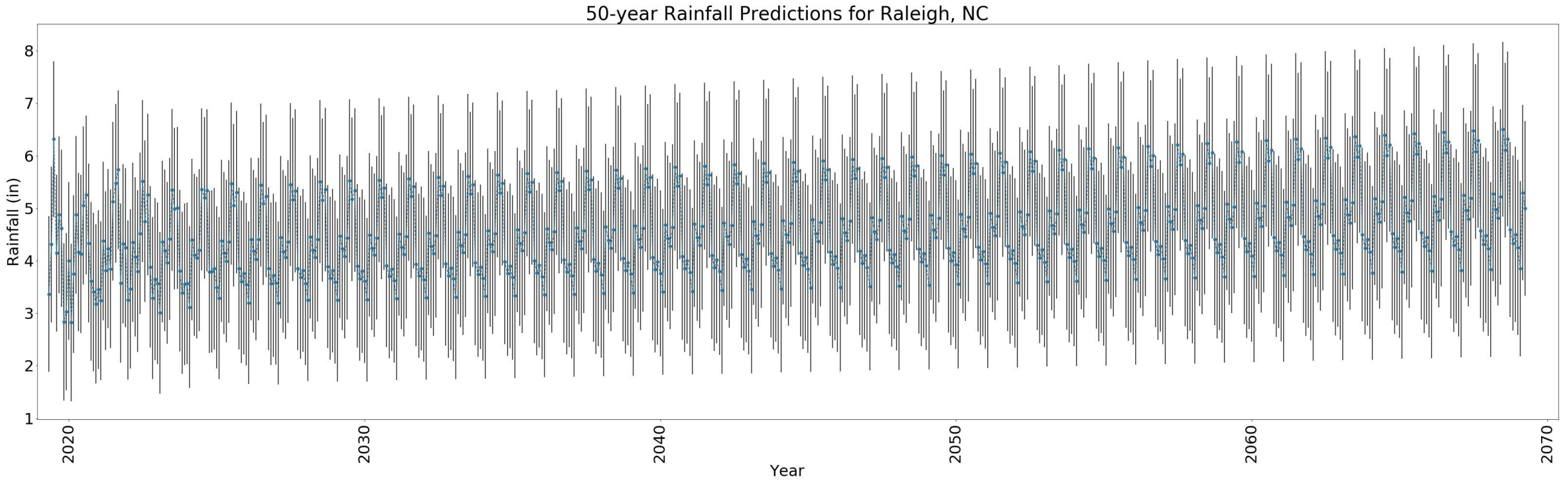
- 46 locations with external locations
- 8 of the 46 with a better model when using exogenous variables
  - Whiteville, NC
  - Casar, NC
  - Forest City, NC
  - Gastonia, NC
  - Lake Lure, NC
  - Elizabethtown, NC
  - Mount Holly, NC
  - Grandfather Mtn, NC



Predictions

Conclusions

- Not a significant change, but rainfall is predicted to increase over next 50 years
- Implement higher regulations for protections against flooding in developing areas



Future Research

- Include other influential variables
  - Climate change predictions
  - Historic climate data based upon cores and other areas of research besides just historic records
- Larger more complete dataset