Weather Patterns of Michigan, US and West Bengal, IN

A COMPARATIVE STUDY

Saurav Das | Applied Plotting Charting & Data Representation in Python: Capstone Project | August 1, 2022

Table of Contents

1.0 Region and Domain

Geographical Region A: Michigan, US

Area: 253790 km^2 (approximate)

Containing Latitudes: 41° N - 49°N

Geographical Region B: West Bengal, IN

Area: 88752 km^2 (approximate)

Containing Latitudes: 21° 25 ʻ N - 27° 13ʼ N

Research Domain: Weather Phenomenon.

Research Data Parameters:

Yearly Average Temperature: T_{ave}

Yearly Precipitation: P

Dataset: Time series 120 years [1901 – 2021]

Data Source: See section 3.0 Dataset Links.

2.0 Research Questions.

Research Question I). What is the long-term variation in average temperature and precipitation levels of Michigan, US and West Bengal, IN relative to each other?

To answer this, I calculate basic statistical measures of the data of average temperature and Precipitation levels for 120 years as presented below.

State	$max(T_{ave})$	$min(T_{ave})$	$\overline{T}_{ave}$	$\mu_{1/2}(T_{ave})$	$\sigma(T_{ave})$
Michigan	9.35	4.4	6.99	6.93	0.86
West Bengal	26.85	24.8	25.61	25.56	0.37

Table I.(i) Statistics for average temperature T_{ave} in $°C$ - 120 years.

State	$max(P)$	$min(P)$	$\overline{P}$	$\mu_{1/2}(P_{ave})$	$\sigma(P)$
Michigan	1092.5	595.34	809.58	812.51	83.78
West Bengal	2150.54	1304.62	1758.88	1753.41	175.46

Table I.(ii) Statistics for precipitation P in mm - 120 years.

State	$\sigma(T_{ave})$	$\sigma(P)$
Michigan	0.0013	0.0090
West Bengal	0.0111	0.0094

Table I.(iii) Normalized Standard Deviations for T_{ave} and P - 120 years.

The normalized standard deviation for temperature is calculated by dividing $\sigma(T_{ave})$ by the length of the temperature measurements (treated as a vector) $\sigma(T_{ave}) = \sigma(T_{ave}) / \|\mathbb{T}_{ave}\|$. Here $\mathbb{T}_{ave}$ is the set of all temperature measurements. Similarly, the normalized standard deviation for precipitation is calculated by dividing $\sigma(P)$ by the length of the precipitation measurements (treated as a vector) $\sigma(P) = \sigma(P) / \|\mathbb{P}\|$. Here $\mathbb{P}$ is the set of all precipitation measurements.

State	$R^{\sigma}(T_{ave})$	$R^{\sigma}(P)$
Michigan	7.13	1.04
West Bengal	1.00	1.00

The measure $R^{\sigma}(T_{ave})$ is the relative normalized standard deviation ratio for average temperatures. It is calculated by dividing the normalized standard deviation of temperature by the minimum normalized standard deviation from the set of normalized standard deviations of temperatures. Similarly, the measure $R^{\sigma}(P_{ave})$ is the relative normalized standard deviation ratio for precipitation. It is calculated by dividing the normalized standard deviation of precipitation by the minimum normalized standard deviation from the set of normalized standard deviations of precipitations.

$$R^{\sigma}_{MICH}(T_{ave}) = \sigma_{MICH}(T_{ave}) \Big/ \min(\sigma_{MICH}(T_{ave}), \sigma_{WB}(T_{ave}))$$
$$R^{\sigma}_{MICH}(P) = \sigma_{MICH}(P) \Big/ \min(\sigma_{MICH}(P), \sigma_{WB}(P))$$

$$R^{\sigma}_{WB}(T_{ave}) = \sigma_{WB}(T_{ave}) \Big/ \min(\sigma_{MICH}(T_{ave}), \sigma_{WB}(T_{ave}))$$
$$R^{\sigma}_{WB}(P) = \sigma_{WB}(P) \Big/ \min(\sigma_{MICH}(P), \sigma_{WB}(P))$$

Research Question I). What is the long-term variation in average temperature and precipitation levels of Michigan, US and West Bengal, IN relative to each other?

Research Answer I). Since $R^{\sigma}_{MICH}(T_{ave}) \approx 7, R^{\sigma}_{WB}(T_{ave}) = 1$ the citizens of Michigan, US experience temperature variations approximately 7 times stronger then the citizens of West Bengal, IN. Also as $R^{\sigma}_{MICH}(P) \approx 1, R^{\sigma}_{WB}(P) = 1$ the citizens of Michigan, US experience precipitation variations approximately of same magnitude as the citizens of West Bengal, IN.

Research Question II). Is there any relation between the normalized distributions of average temperature and precipitation levels of Michigan, US and West Bengal, IN?

To answer this question, I have plotted 4 graphs to visually present the answer.

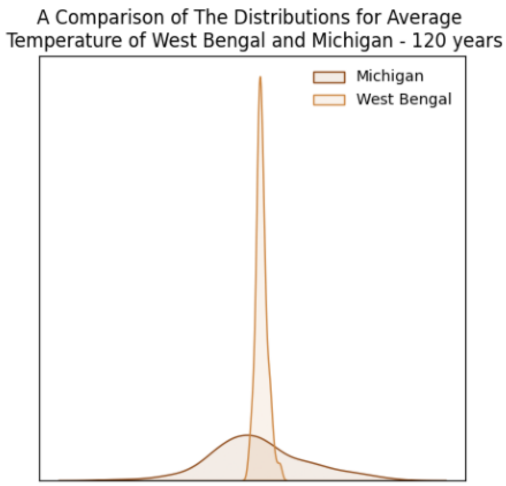

Figure 1.a

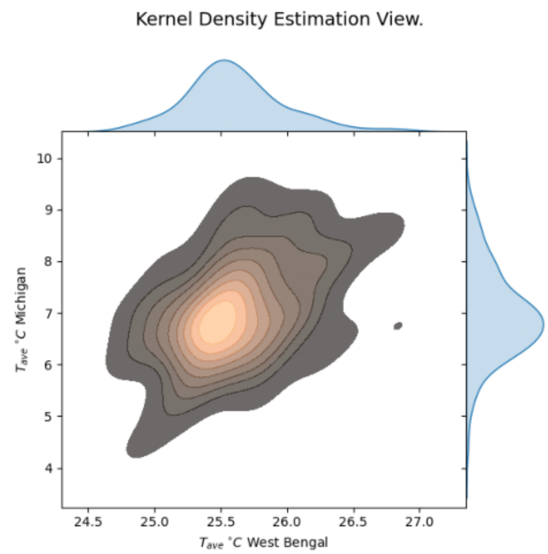

Figure 1.b

The figure 1.a shows the two normalized distributions of T_{ave} for Michigan and West Bengal. We see the shape of the distributions are quite different, while both may be approximated by the normal distribution, the standard deviations for the distribution for West Bengal is much smaller the that of Michigan. The right graph shows a kernel density plot for the two distributions, note the irregular elongated shape which hints a positive correlation between the variables and the small area of high frequency overlap. This is consistent with our earlier calculations

$R^{\sigma}_{MICH}(T_{ave}) \approx 7, R^{\sigma}_{WB}(T_{ave}) = 1$.

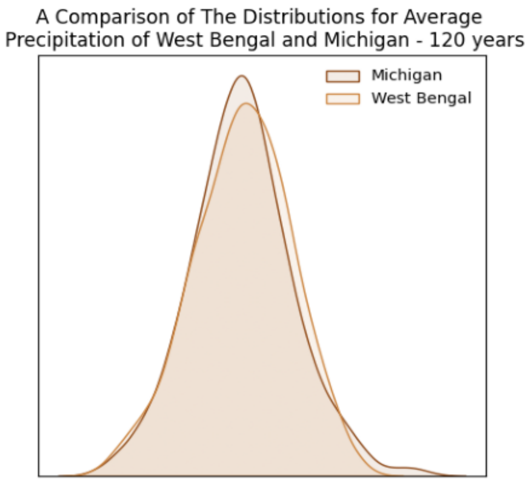

Figure 2.a

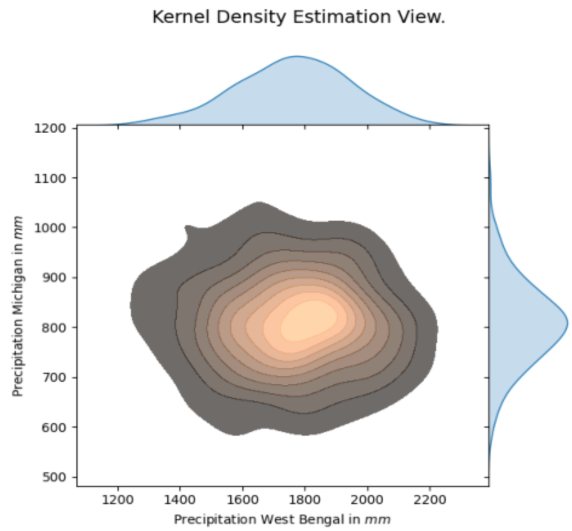

Figure 2.b

The figure 2.a shows the two normalized distributions of precipitation P for Michigan and West Bengal. We see the shape of the distributions are very similar, both distributions may be approximated by the normal distribution, with the standard deviations for the distribution for West Bengal and that of Michigan quite similar in magnitude. The right graph shows a kernel density plot for the two distributions, note the regular shape which hints no correlation between the variables and the large area of high frequency overlap. This is consistent with our earlier calculations $R^{\sigma}_{MICH}(P) \approx 1, R^{\sigma}_{WB}(P) = 1$.

Research Question III). What is the relative impact of global warming on the average temperature and precipitation levels of Michigan, US and West Bengal, IN? Are the changes for the cities in equal proportion?

To answer this question, I have calculated 10 year moving averages on the temperature and precipitation data to smooth the datasets. And I have plotted 4 graphs to visually present the answer.

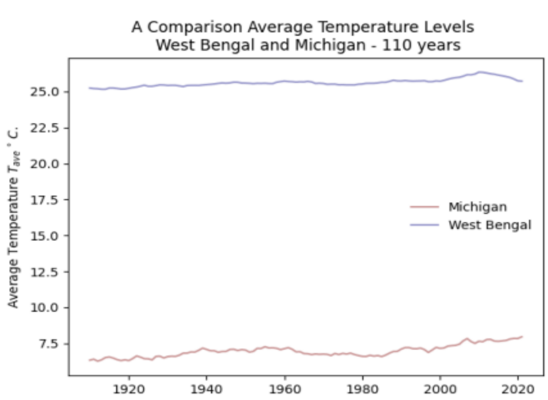

Figure 3.a

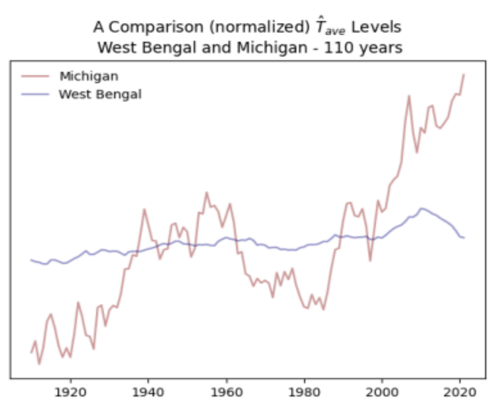

Figure 3.b

The figure 3.a shows the smoothed temperature plot for Michigan and West Bengal for 110 years. Notice the gradual uptrend of the average temperatures of the two states. But strangely for West Bengal the uptrend peaked in about the year 2010 and a gradual downtrend has started. This is more clearly shown in figure 3.b, in which the smoothed normalized temperatures of the two cities are plotted. In this plot, we can clearly conclude that Michigan is more strongly affected by the weather pattern changes due to global warming than West Bengal, by continuously having higher and higher average temperature year after year from the year 2000.

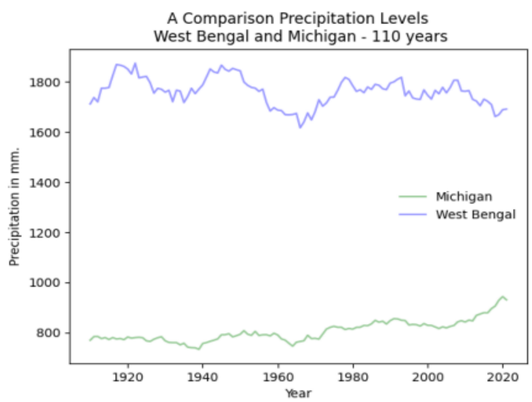

Figure 4.a

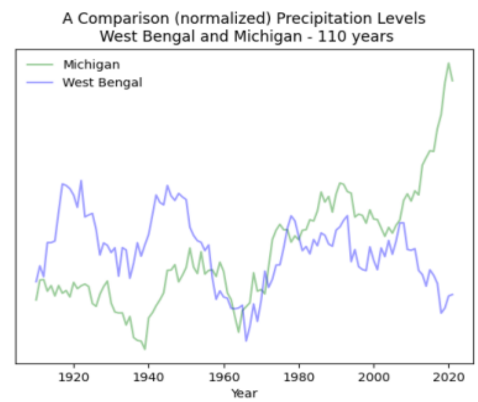

Figure 4.b

The figure 4.a shows the smoothed precipitation plot for Michigan and West Bengal for 110 years. Notice the gradual uptrend of the precipitation levels of Michigan starting from the year 2000. But strangely for West Bengal the precipitation levels are more cyclic in nature and has followed its cyclic pattern without break. This is more clearly shown in figure 4.b, in which the smoothed normalized precipitation levels of the two cities are plotted. In this plot, we can clearly conclude that Michigan is more strongly affected by the weather pattern changes due to global warming than West Bengal, by continually receiving higher precipitation rates year after year from the year 2000.

Research Question IV). In this question, I test a hypothesis:

Hypothesis: 'The measure of [Precipitation / T_{ave}] for a particular geographical location tends to a constant and its variations are tightly contained by a normal distribution with small standard deviation values (5%-10% of the mean value of the distribution). And if we normalize the distributions for the measure of [Precipitation / T_{ave}] for different locations, the mean values of the datasets must converge to a fixed constant."

To understand the reason that leads to this hypothesis, we must intuitively analyze the scenario. For a particular location, if it has not rained for some period the temperature starts rising, creating a low pressure which attracts rain clouds to this location. Again, if it has rained a lot in a specific period, the temperature starts decreasing, creating a high pressure which repels further rain clouds in this location. This should create a balance and a typical constant value (for a particular region) for the measure [**Precipitation** / T_{ave}]**.** Of course, this is a very simplified model actual whether models are a lot more complex, but let us explore and investigate the validity of our simple model and the hypothesis presented above on weather pattern for the two cities Michigan, US and West Bengal, IN. First I present a single graph plotting the normalized measure of [Precipitation / T_{ave}] for the two cities of Michigan and West Bengal as figure 5 below.

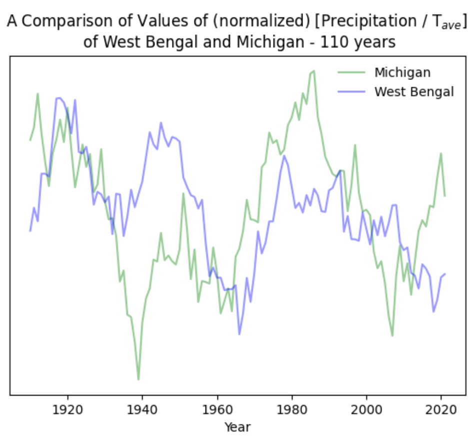

Figure 5

On examining the adjoining graph, we can easily conclude that the variables plotted have very similar range. Also on examining the patterns of the plotted variables we can also conclude that the variables have same or very similar values for their mean values.

Next I present 2 graphs which show us the relation between the distributions of the two variables for the measure of normalized [Precipitation / T_{ave}] for the two cities of Michigan and West Bengal in more detail.

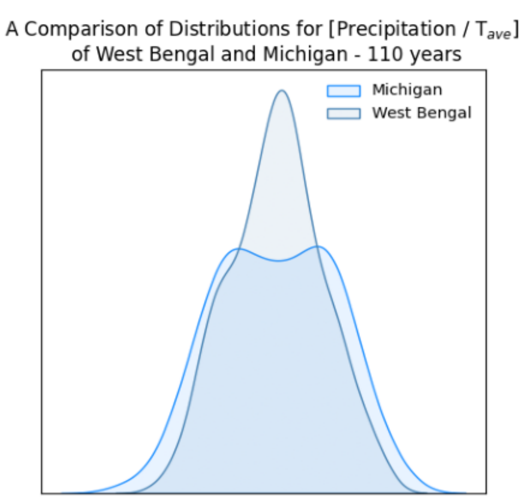

Figure 6.a

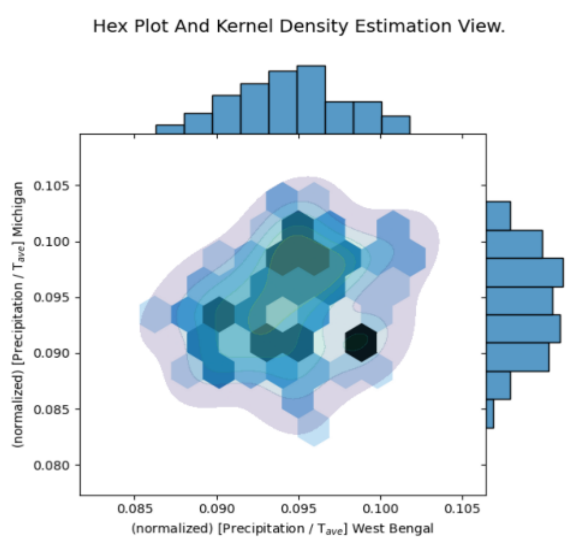

Figure 6.b

The figure 6.a shows the two normalized distributions for the measure of [Precipitation / T_{ave}] for Michigan and West Bengal. We see the shape of the distributions are very similar, both distributions may be approximated by the normal distribution, with the standard deviations for the distribution for West Bengal and that of Michigan quite similar in magnitude also the means of the distributions are very close to each other. The right graph shows a hex plot overlaid with a kernel density plot for the two distributions. The deeper color hex-bins denote more data points in them. Also note the regular shape and the large area of high frequency overlap of the kernel density plot, this hints no correlation between the variables and similar distributions of the variables. Finally observe the fact that most of the deepest color hex-bins reside in the large area of high frequency overlap area of the kernel density plot. This is consistent with part of our hypothesis that 'variations for the measure of [Precipitation / T_{ave}] are tightly contained by a normal distribution with small standard deviation values (5%-10% of the mean value of the distribution)'.

Now I present the statistics for the measure of $K_{P/T} = norm(P/T_{ave})$.

State	$max(K_{P/T})$	$min(K_{P/T})$	$\overline{K}_{P/T}$	$\mu_{1/2}(K_{P/T})$	$\sigma(K_{P/T})$
West Bengal	0.1017693733	0.0863126386	0.0944295464	0.0946112713	0.0034106008
Michigan	0.1035498554	0.0833478019	0.0943884383	0.0943550424	0.0044038783

Table V.(i) Statistics for the measure $K_{P/T}$ (correct to 10 decimal places) - 110 years.

We are interested in the values of $\overline{K}_{P/T}$, specifically according to our hypothesis the values for $\overline{K}_{P/T}$ of different places must be same. Let's check this for Michigan and West Bengal. We have:

$$R^K = \frac{\overline{K}_{P/T,WB}}{\overline{K}_{P/T,MICH}} = 1.0004355205$$

$$\varepsilon(R^K) = R^K - 1 = 0.0004355205$$

$$\varepsilon_{\%}(R^K) = \varepsilon(R^K) \text{ x } 100\% = 0.04355205\%$$

R^K is the ratio of $\overline{K}_{P/T}$ for West Bengal and Michigan. As this must be 1.0 we subtract 1.0 from R^K to get the error or difference from the theoretical predicted value. So, we get $\varepsilon(R^K) = R^K - 1$, it is the error term or difference of the $\overline{K}_{P/T}$ from the theoretical predicted value.

And $\varepsilon_{\%}(R^K) = \varepsilon(R^K) \text{ x } 100\%$, it is the error term represented as a percentage. We see $\varepsilon_{\%}(R^K) = 0.04355205\%$ so the actual data for the states of Michigan, US and West Bengal, IN agree with our hypothesis to a very good error margin of about **0.044%**. But please be aware that though this hypothesis agrees so well with the data from Michigan and West Bengal, it might not agree so well with other data pairs from other geographical locations, more extensive research is needed before we can generalize this hypothesis.

3.0 Dataset Links.

The World Bank Group, Climate Change Knowledge Portal:
https://climateknowledgeportal.worldbank.org/download-data

National Centers for Environmental Information:
https://www.ncei.noaa.gov/access/search/data-search/daily-summaries

Note: To download data you must visit these pages, fill up the appropriate forms and then download the data in CSV format.