Exploratory Data Analysis of Indian Rainfall Data

India is an agricultural country and secondary agro based market will be steady with a good monsoon. The economic growth of each year depends on the amount of duration of monsoon rain, bad monsoon can lead to destruction of some crops, which may result in scarcity of some agricultural products which in turn can cause food inflation, insecurity and public unrest. In our analysis we are trying to understand the behavior of rainfall in India over the years, by months and different subdivisions.

Describing the data:

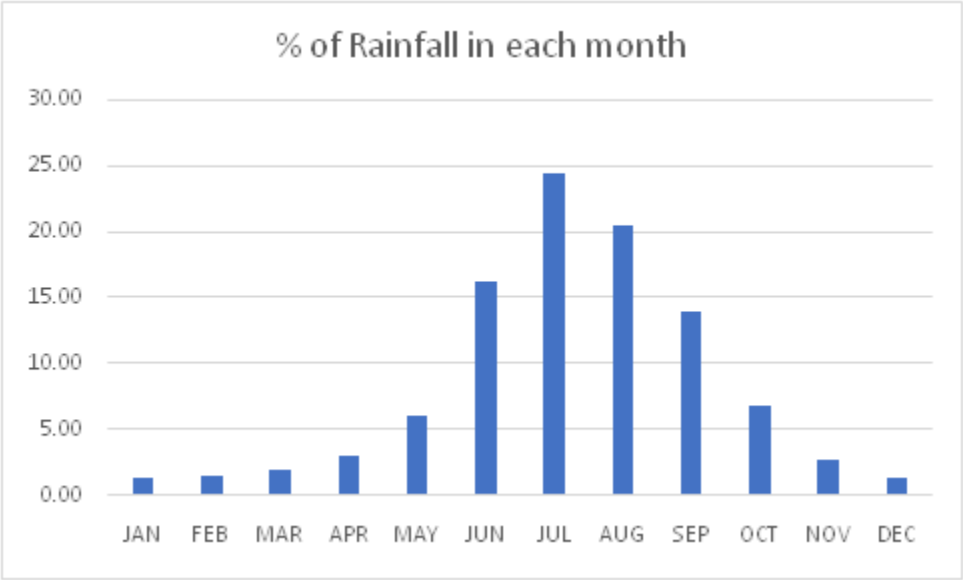
Data-set is downloaded from “data.gov.in” website. It has data for 117 years (1901–2017) consisting of monthly and seasonal data for all 36 meteorological subdivisions of India. So in total we have 117*12*36 = 50,544 observations. Our data-set had 0.7% of missing values. For the subdivision Arunachal Pradesh we had missing values for the first 15 years i.e. 1901 to 1915, so for all subdivisions we have considered data from 1916 to 2017 when we are analyzing as whole India. For the rest of the missing values we have used sequential imputation technique. Below table shows mean rainfall observed for each month over years. We can see that average rainfall is high in July and August followed by June and September.

Month	% of Rainfall
JAN	1.32
FEB	1.50
MAR	1.93
APR	3.06
MAY	6.13
JUN	16.25
JUL	24.45
AUG	20.42
SEP	13.99
OCT	6.81
NOV	2.79
DEC	1.34

Monthly mean annual rainfall

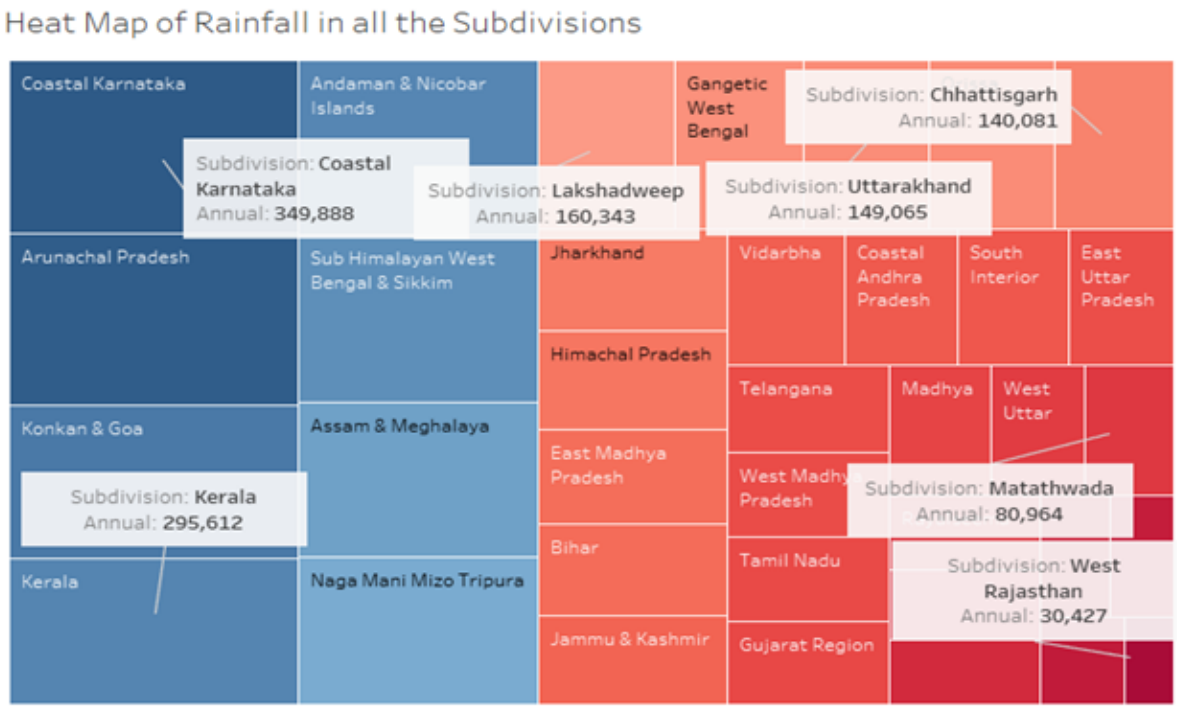
Annual rainfall by months:

The below graph shows the percentage of rainfall each month receives when we consider India as a whole. The rainfall in the months of June, July, August and September together contribute to almost 80% of the annual rainfall.

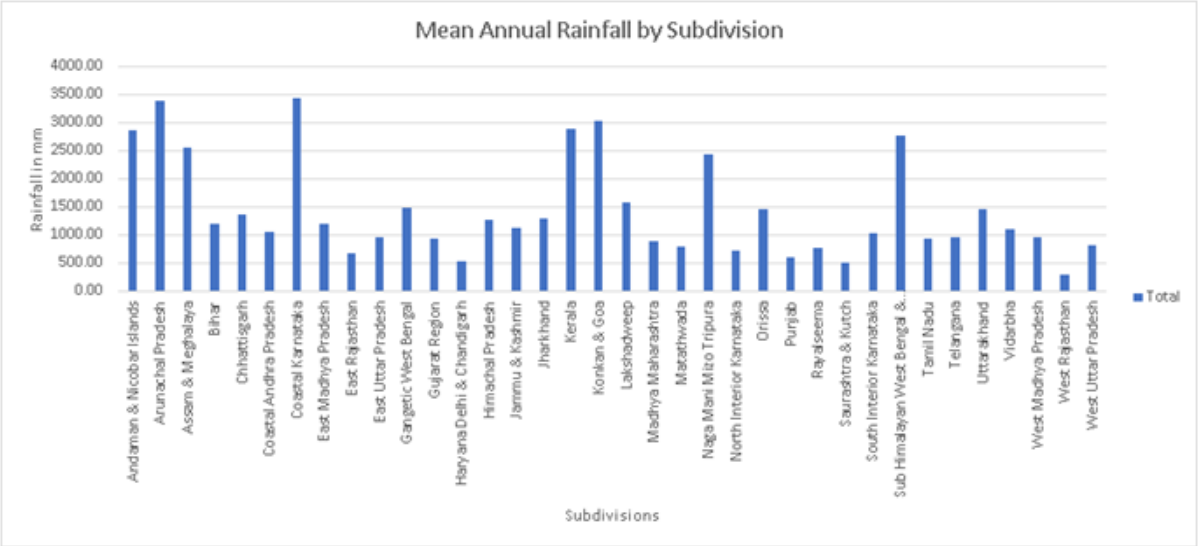


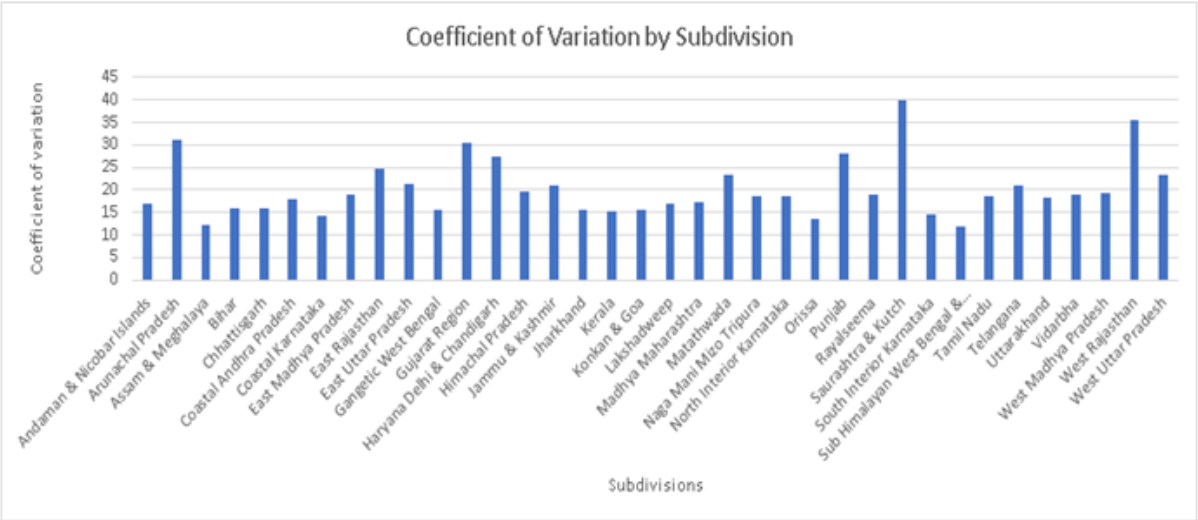
Annual rainfall by subdivision

The following is a heat map plotted based on sum of rainfall received by each subdivision for all these years. The subdivisions with large area represents high rainfall and with small boxes represent less rainfall. We can see that the subdivision located at Southwest and Northeast part of India have received more rainfall compared to central India.



The average rainfall and variation values are plotted for each subdivision on different graphs which are given below. We can see that the subdivision which receive High rainfall have less variation seen over years whereas the subdivisions receiving low rainfall showed more variation over the years.





Understanding the Monsoon in India:

We will now move to a more interesting part of the analysis.

We will see what exactly is monsoon, different types of monsoon winds in India, which subdivisions of India receives rainfall from which monsoon winds and why only particular subdivisions receive highest rainfall during this monsoon season.

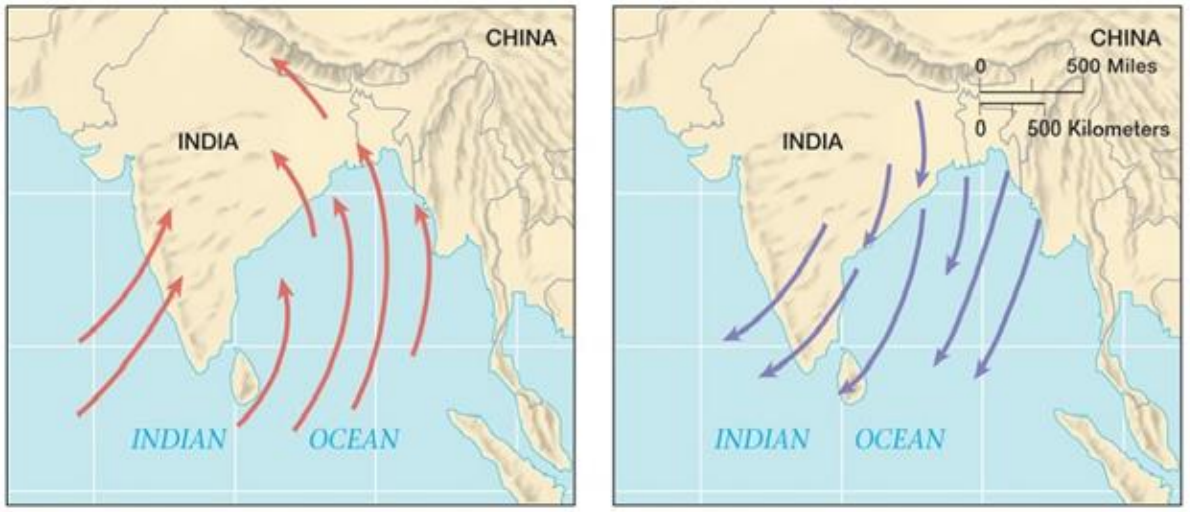
Before addressing these questions, we will see the basic concept of how rainfall actually occurs. So basically, during the summers, the

Indian subcontinent heats up more as compared to the Indian ocean as the sun is directly over the landmass. This creates a low pressure over the Indian subcontinent and a relatively low pressure over the Indian ocean. And as we know, the wind flows from high pressure area to the low pressure are in order to fill the void that was created thanks to the pressure system. So when the wind starts flowing from high pressure area to low pressure are i.e. from sea to land, it picks up the moisture from the sea and while entering the indian subcontinent it comes in contact with the high terrains and hence precipitation occurs.

Now that we have understood how rainfall occurs we will now see what exactly is monsoon? Monsoon is traditionally defined as a seasonal reversal of wind accompanied by corresponding changes in precipitation.

As seen in the image below, India receives rainfall fromSouthwest Monsoon winds (Summer Monsoon or Advancing Monsoon) and Northeast Monsoon winds (Winter Monsoon or Retreating Monsoon)

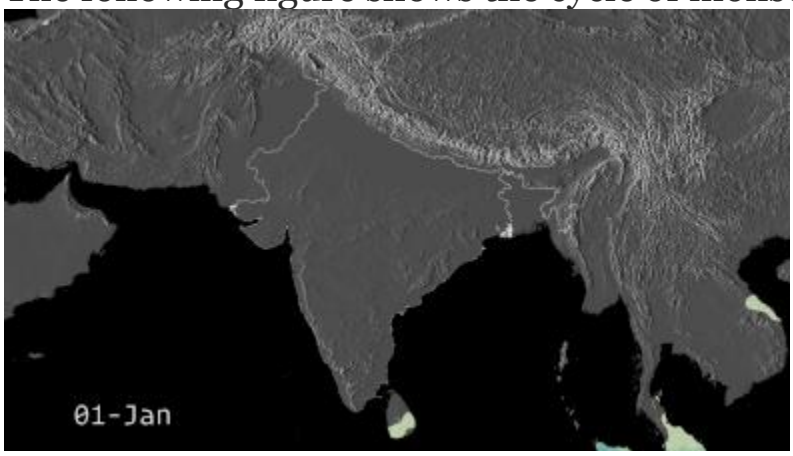
The Southwest Monsoon usually starts in the first week of June and ends by first week of September and monsoon usually starts retreating from the Indian Subcontinent by the start of September and leaves the subcontinent completely by the end of November. And as we have seen in the previous graphs that Southwest monsoon provides almost 80% of the rainfall in India. This Southwest Monsoon has two branches, namely Arabian Sea branch and Bay of Bengal Branch.



Summer Winter

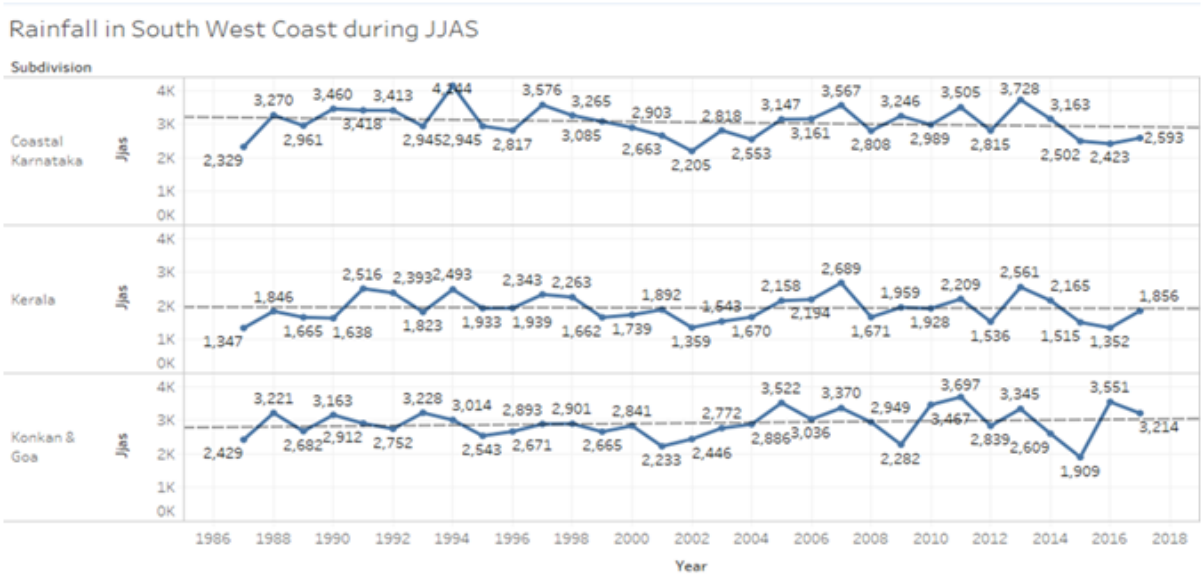
Source: sites.google.com

The following figure shows the cycle of monsoon in India over a year



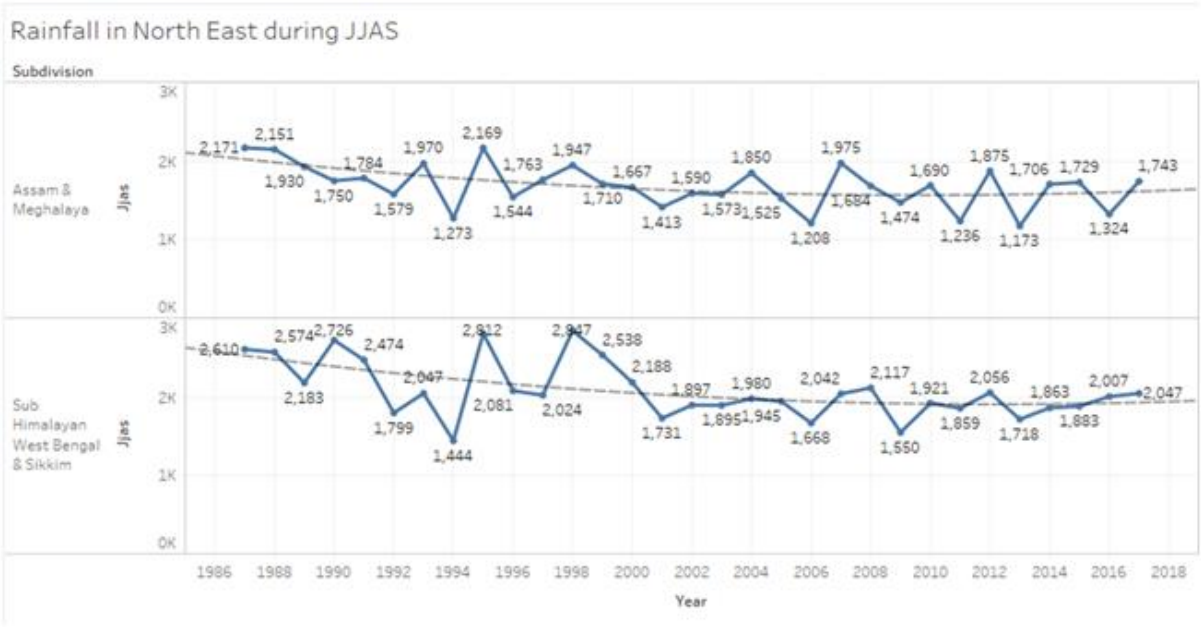
Source: en.wikipedea.org

Now, due to the presence of high rising Western Ghats which runs along the South West coast of India in the states of Kerala, Karnataka, Goa and Maharashtra, they block the Arabian Sea branch of southwest monsoon and hence these regions receive very high rainfall during monsoon season. This is shown in the graph below. (For all the graphs we have considered the last 3 decades i.e. last 30 years data (1987–2017))

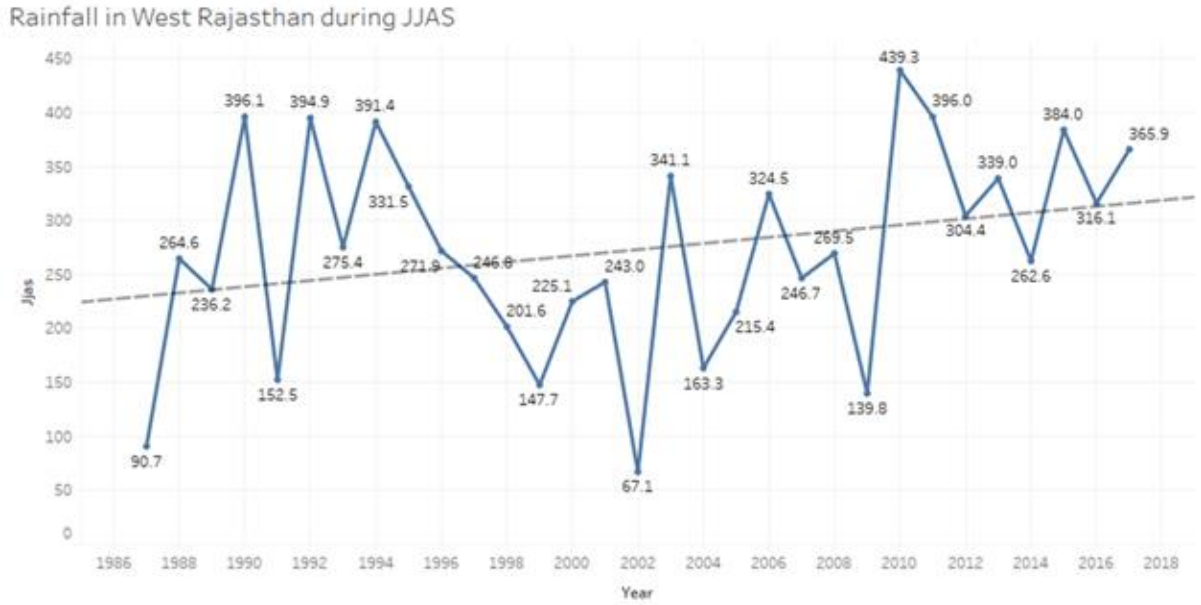


The Bay of Bengal branch of the Southwest Monsoon provides rainfall to the eastern and north eastern regions of India. Due to the presence of the eastern ghats along Orissa and West Bengal, these regions receive high rainfall in the East and when this branch reaches the north eatsern part of India i.e. Sikkim, Arunachal Pradesh, Assam & Meghalaya, due to the presence of Khasi hills and other hilly areas, these subdivisions receive very high rainfall. Cherrapunji in Meghalaya is titled as one of the wettest places on earth.

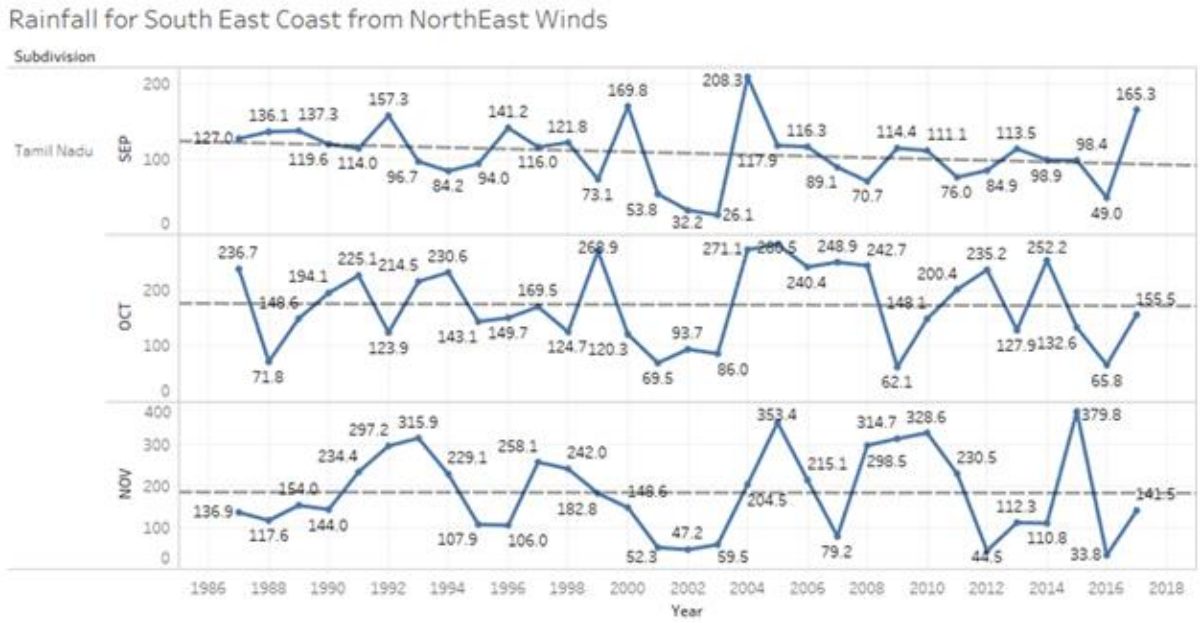
The average annual rainfall in these regions is shown in the graphs below.



Now, when these monsoon winds provide rain to the Southwest, Northeast and most of the central part of India, the moisture in them starts decreasing. And by the time they reach the western part of Rajasthan, the moisture content in these winds decreases to a very great extent and hence this region receives very less rainfall and hence the presence of Thar desert in West Rajasthan. This is shown from the graph as follows:

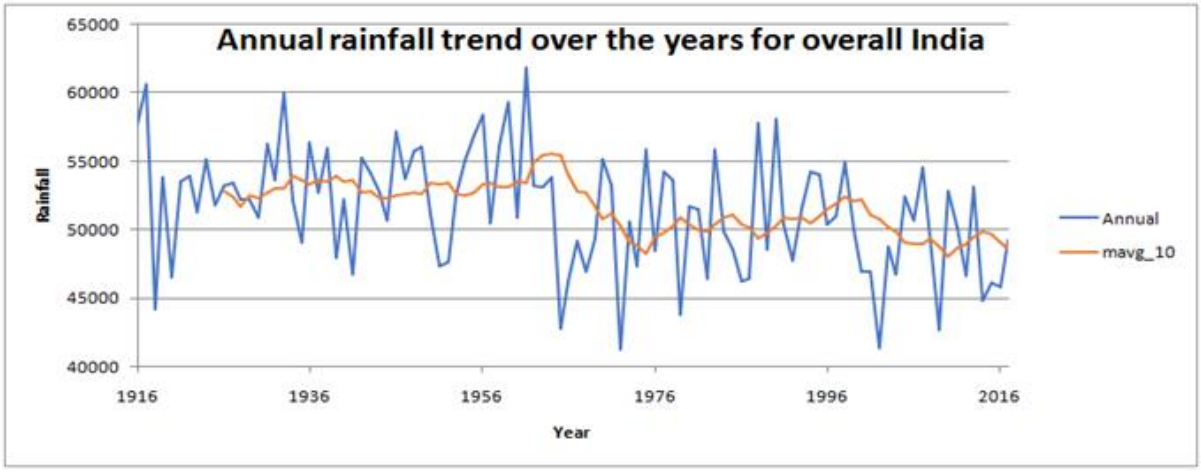


By the start of September, when Southwest monsoon have provided rainfall to most of India, the Sun shifts its position toward the Indian ocean which is actually the start of winter, the sea heats more as compared to the land and hence a high pressure is created on the Indian landmass and a low pressure on the Indian ocean. So the wind starts flowing from Indian subcontinent i.e. from Himalayas to the Indian ocean. These winds are usually dry, but when they come in contact with the Bay of Bengal they carry moisture in them and when they flow over Tamil Nadu and Kerala, it gives rain in those regions and hence Tamil Nadu receives most of the rainfall from the Northeast Monsoon as compared to the Southwest monsoon. Also Kerala receives a considerable amount of rainfall from the retreating monsoon. This is explained in the graph below. We can see that Tamil Nadu receives most rainfall during October and November as compared to September.



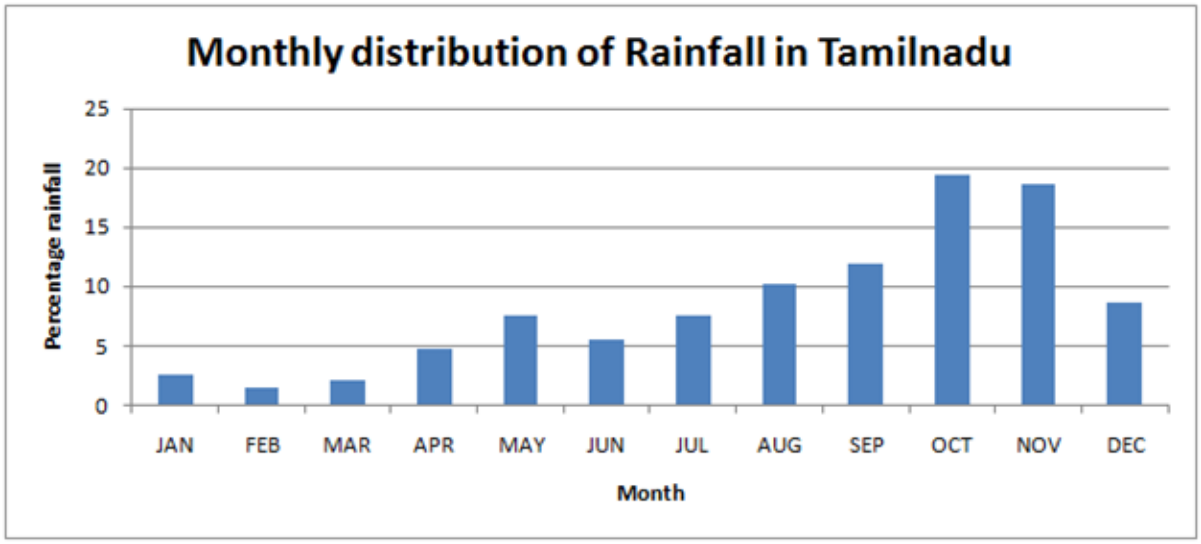
Annual Rainfall trend over the years for whole India

10 years moving average was plotted, we can see that there is a decreasing trend in rainfall in the recent years.



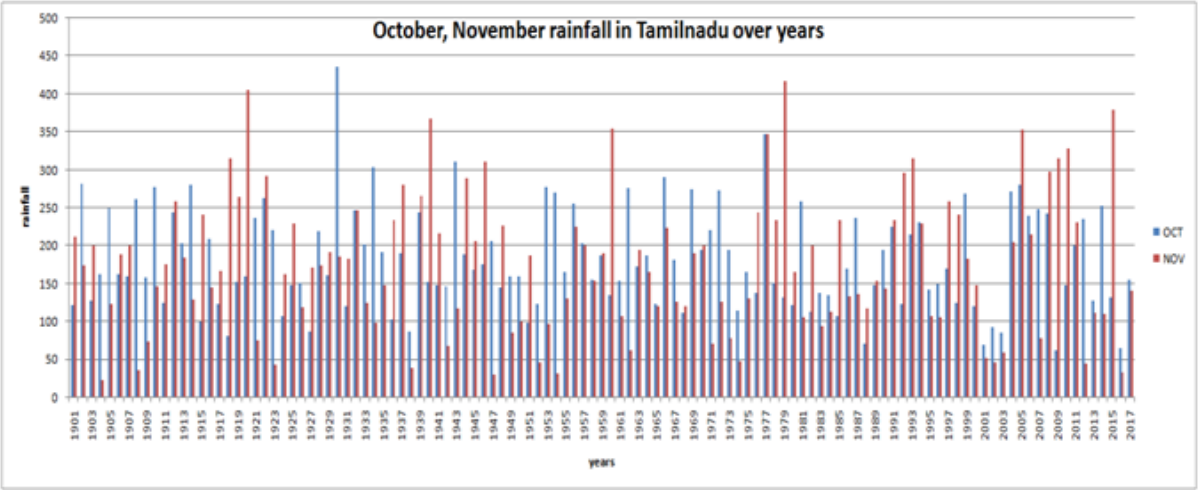
Rainfall Data Analysis for Subdivision TamilNadu

To do rainfall data analysis of Tamil Nadu subdivision, we have considered the data from year 1901 to 2017. First we will see the distribution of rainfall over months and we can see in below figure that rainfall is more in October and November compared to other months as Tamilnadu receives rain during retreating monsoon season because of Northeast trade winds. The climate of Tamil Nadu is tropical in nature with less variation in temperature in summer and winter. This is because of its geographical location.



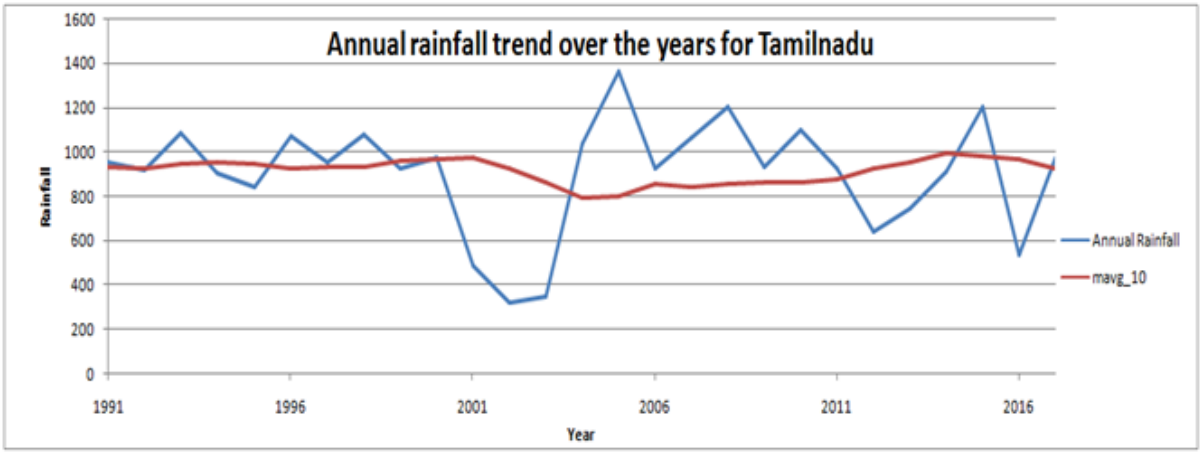
The below graph shows rainfall in Tamil Nadu in months October and November here we can see that most of the years when there is very low rainfall in October there is very high rainfall in November and vice versa. Tamil Nadu also rainfalls from tropical cyclones emerging in the neighborhood of Andaman islands during the retreat monsoon.

In 2015 there were south Indian floods, the flooding in Chennai was described as the worst in a century. In winters of 2017 the worst drought in a century happened in southern India.



Annual Rainfall trend over the years in Tamil Nadu

10 years moving average was plotted, we can see that there is very weak increase in rainfall in the recent years.



Conclusion:

The results show that India has two main rainfall season: one is southwest monsoon(advancing monsoon) and other is Northeast monsoon(retreating monsoon). Advancing monsoon contributes almost 80% of the rainfall. Southwest and Northeast part of India receives most of the rainfall during the advancing monsoon. During the retreating monsoon, Andaman & Nicobar Islands, Kerala, Tamil Nadu receive more rainfall as compared to other subdivisions.

The trend analysis of Annual rainfall considering India as whole show decreasing trend however when trend is analysed for all subdivision individually we can see some division showing increasing trend and some showing decreasing trend. It showed that is is import to study subdivision for better forecasting.

We considered Tamil Nadu as one of the subdivisions to do further analysis. It receives more rainfall during October and November because of retreating monsoon. Since there are only a few months when the Tamil Nadu gets rains and its location at tropical results in high temperature which in turn results in water scarcity problem.

Also because of its geographic location near it is hit sometimes by the cyclones formed in Indian Oceans which results in extreme storms and non normal rainfall.

In an interview, Mrutyunjay Mohapatra, the director general of the IMD, explained how climate change is increasing number of days with heavy rainfall. The season started with 33% deficit rainfall but is ending with 10% higher than normal rainfall, with heavy spells of rain resulting in devastating floods in many states.He said the number of heavy rainfall days was increasing because of climate change, which was making predictions more difficult. Also, this year, the monsoon in India withdrew 40 days later than normal.

Usually, the monsoon withdraws from September 1 from extreme northwest India, that is West Rajasthan, and by September 15, it withdraws from the entire country. But this year, the monsoon withdrew around October 9 or 10.

The monsoon has been quite active in the month of September because of various factors. One important factor is the low-pressure systems that dropped over Bay of Bengal moved towards Rajasthan and under its influence, an East-West oriented low-pressure zone came about. This sustained the monsoon for quite a long time.

Thereafter, a depression formed over the northeast Arabia Sea and it crossed the Gujarat coast and it moved in that direction sustaining the monsoon features.

Authors: Anusha Gajinkar, Vighnesh Tamse-(Article title: Exploratory Data Analysis on Indian Rainfall Data (1901–2017))