

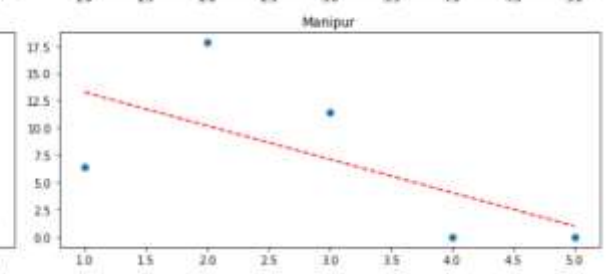
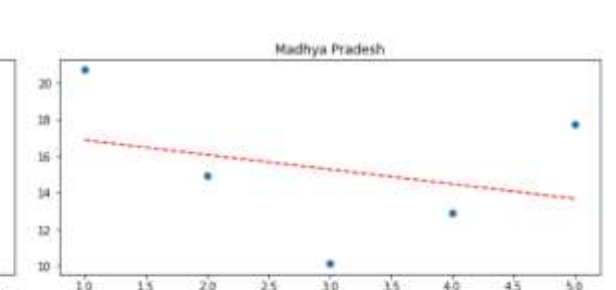
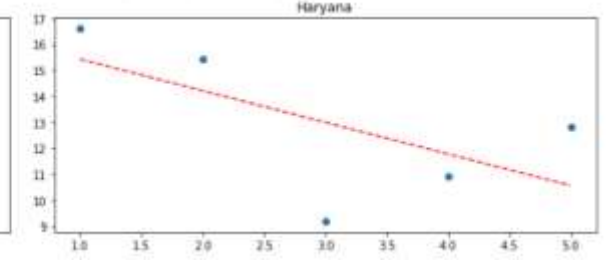
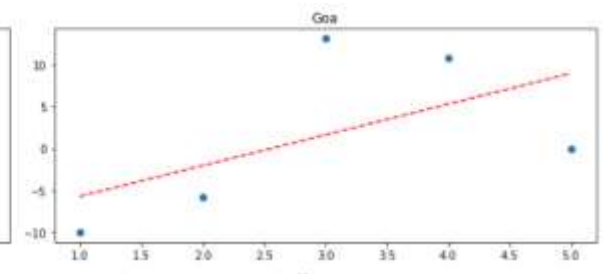
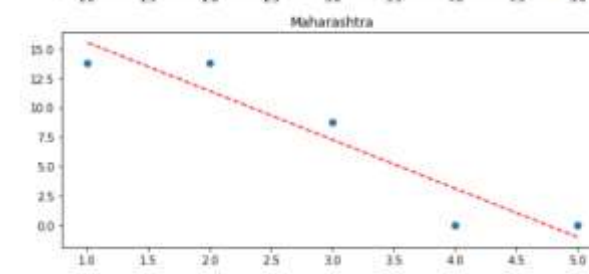
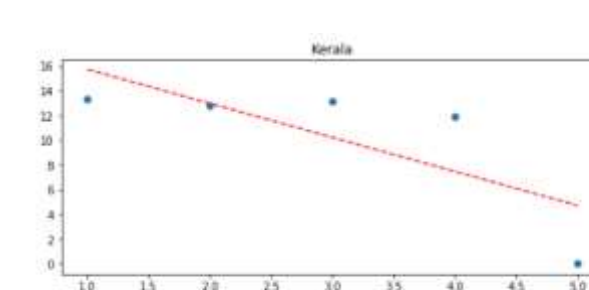
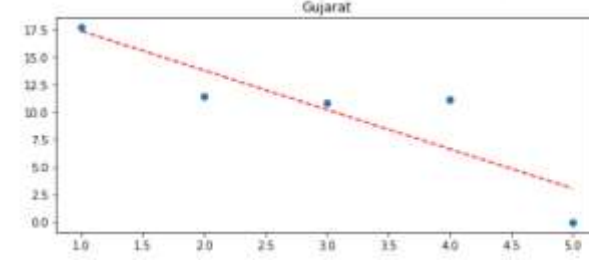
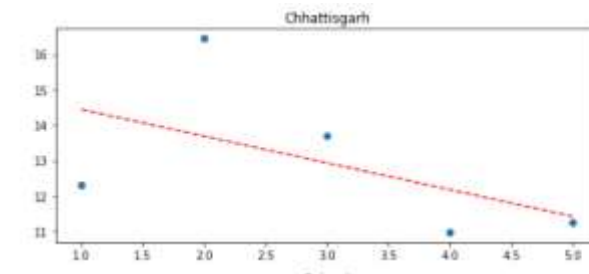
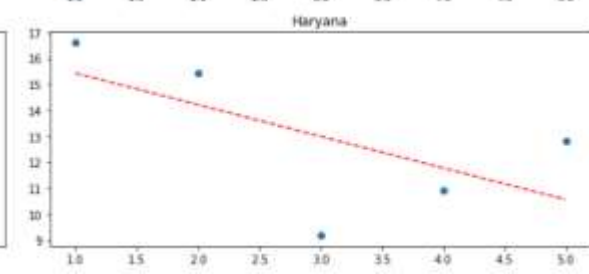
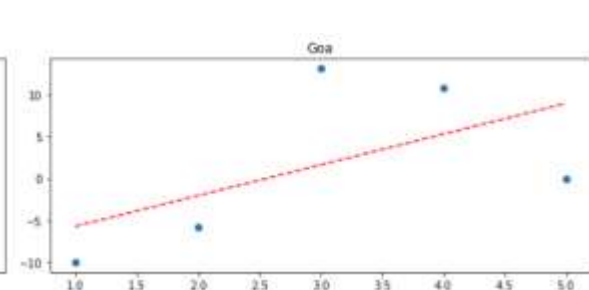
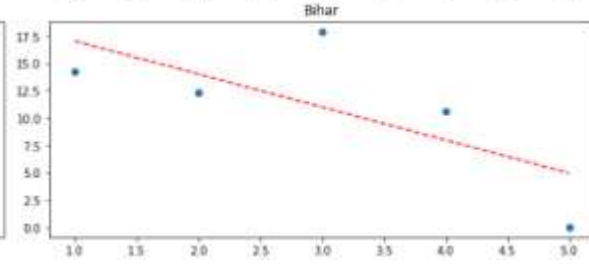
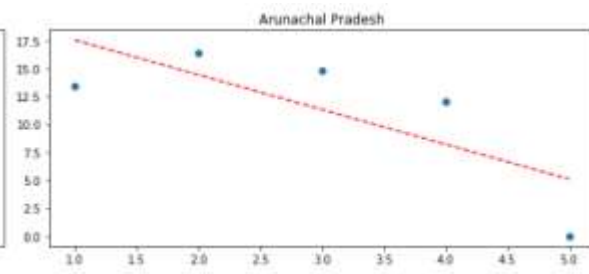
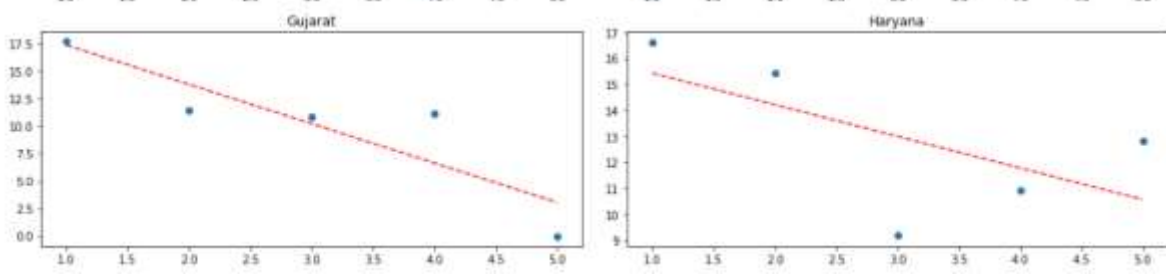
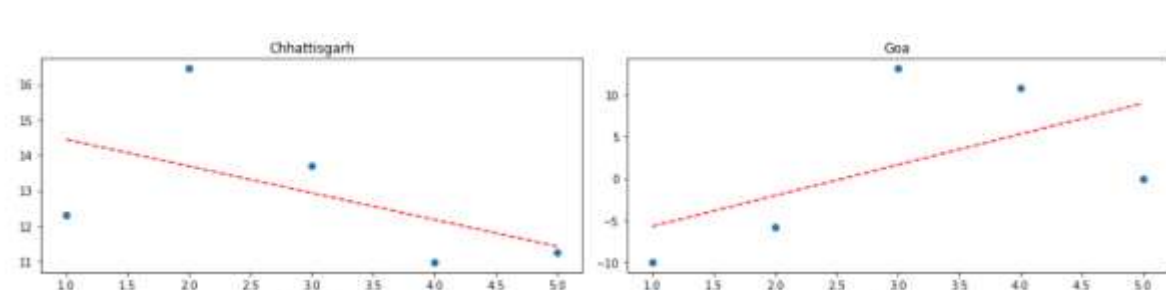
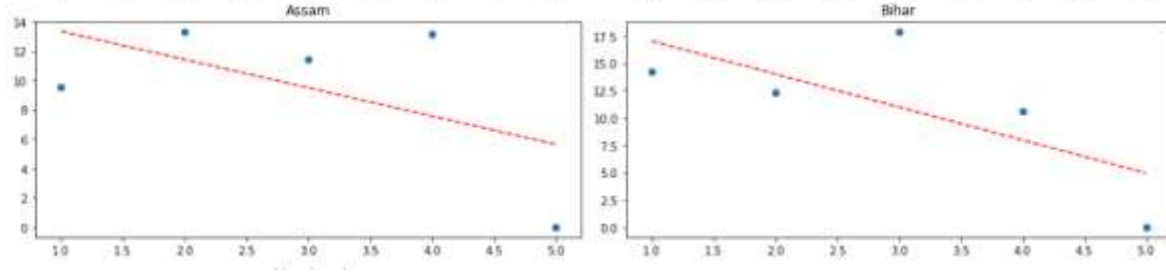
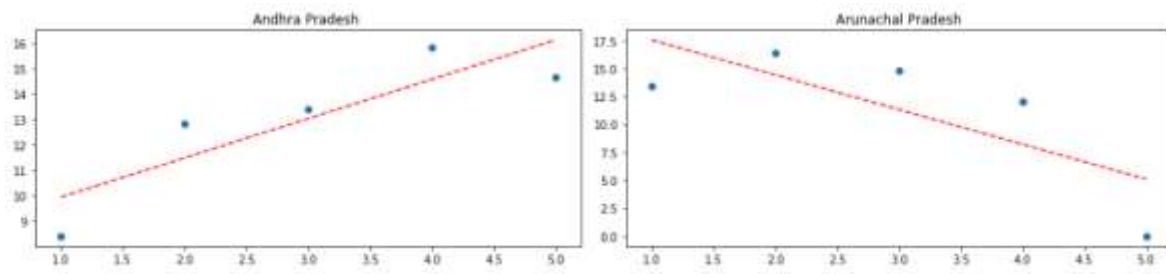
GDP Assignment

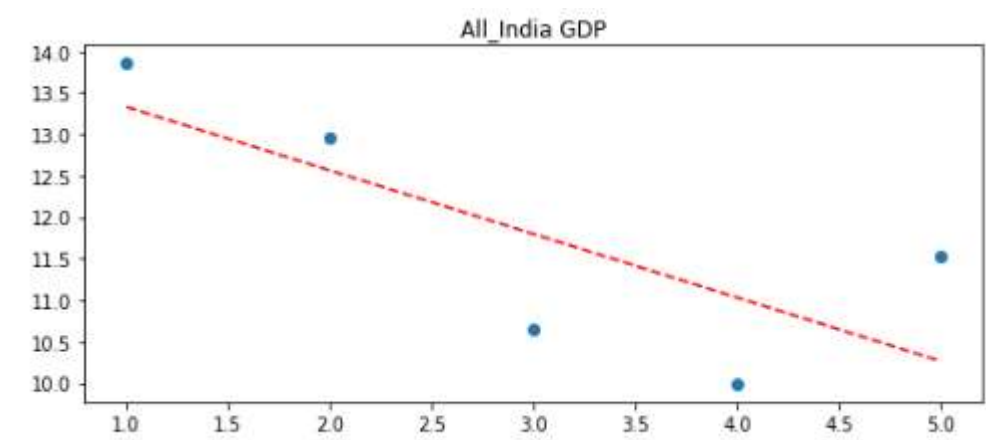
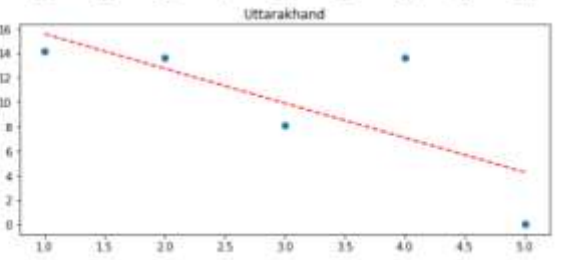
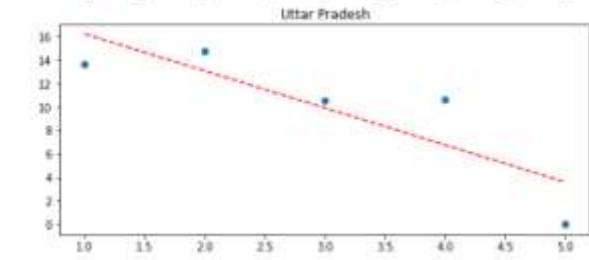
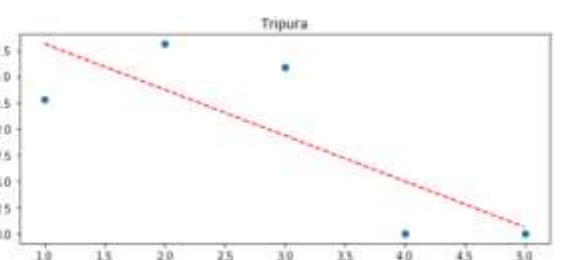
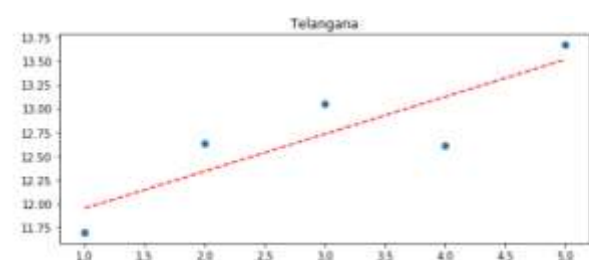
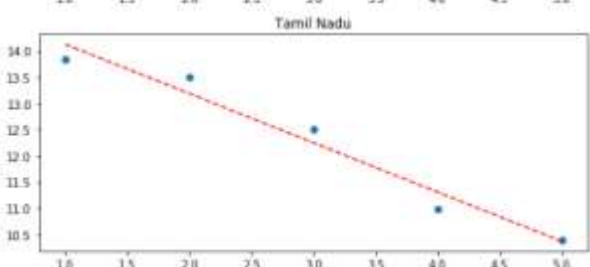
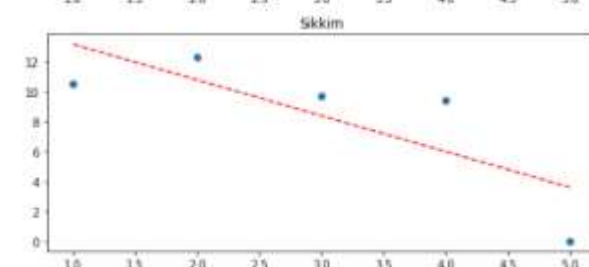
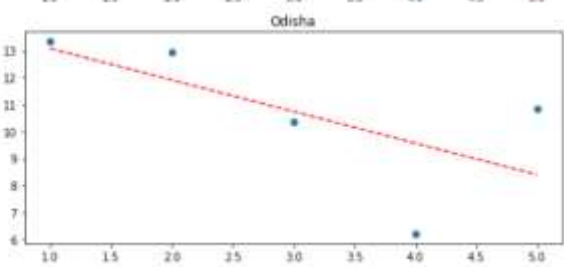
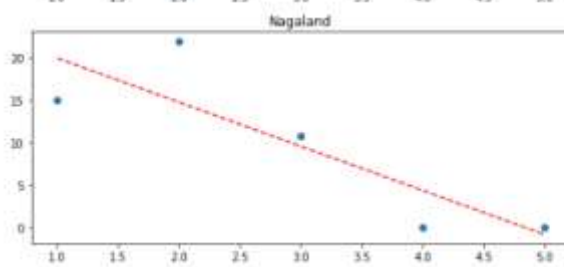
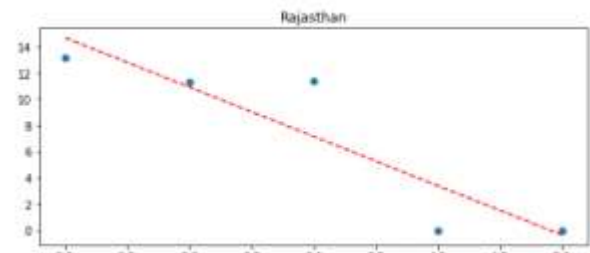
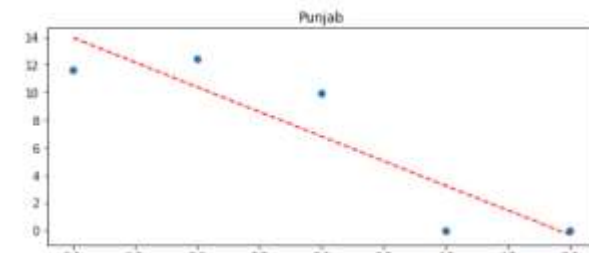
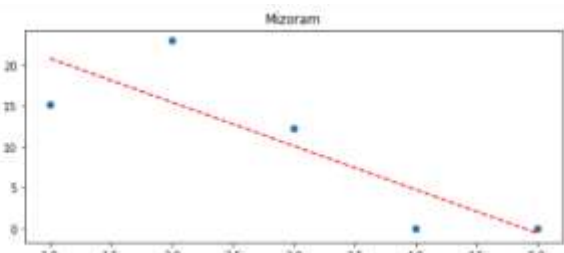
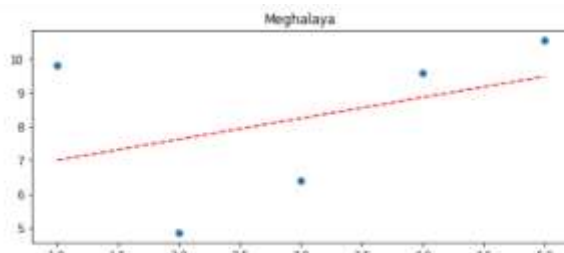
Roadmap and Insights

- Data Gathering :
 - The data is sourced from <https://data.gov.in/>, an Open Government Data (OGD) platform of India.
 - Data 1 A – GSDP data for states and union territories
 - Data 1 B – GSDP data for three sectors (Primary, Secondary, Tertiary)
- Data Cleaning :
 - Data is loaded in to data frames reading the .csv files
 - After loading, all NA values are replaced with 0. Values can be replaced with median also, but here values are replaced with 0 so that no irregular spikes are present in the data.

Part 1 - A:

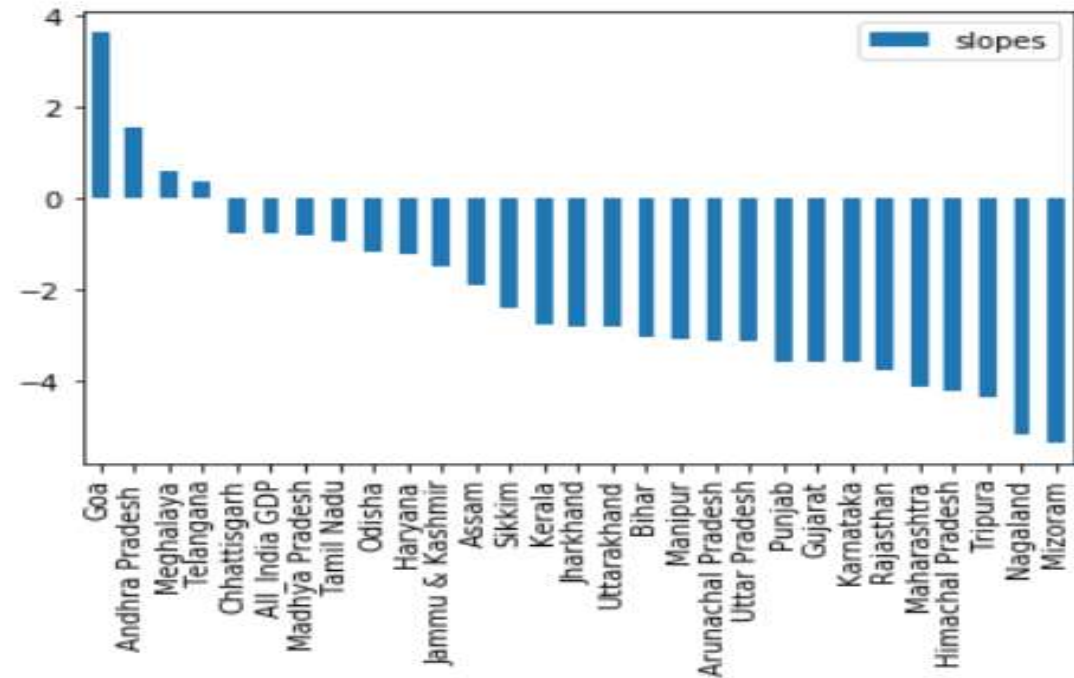
- Data is loaded into the dataframe and is cleaned.
- **Q1:** Plot a graph for rows " % Growth over previous year" for all the states (not union territories) whose data is available, use as much data as possible for this exercise. Use the best fit line to represent the growth for each state. Draw a similar line graph for the nation as well.
- Since data is to be analyzed for “% Growth over previous years”, all other rows are to be removed.
- After it, best fit line graph for the remaining data is plotted.
- A “Best Fit Line” graph is one which provided graphical visualization whether the trend of given data is Positive or Negative.
- Below are the “Best Fit Plots” plotted for all the states.



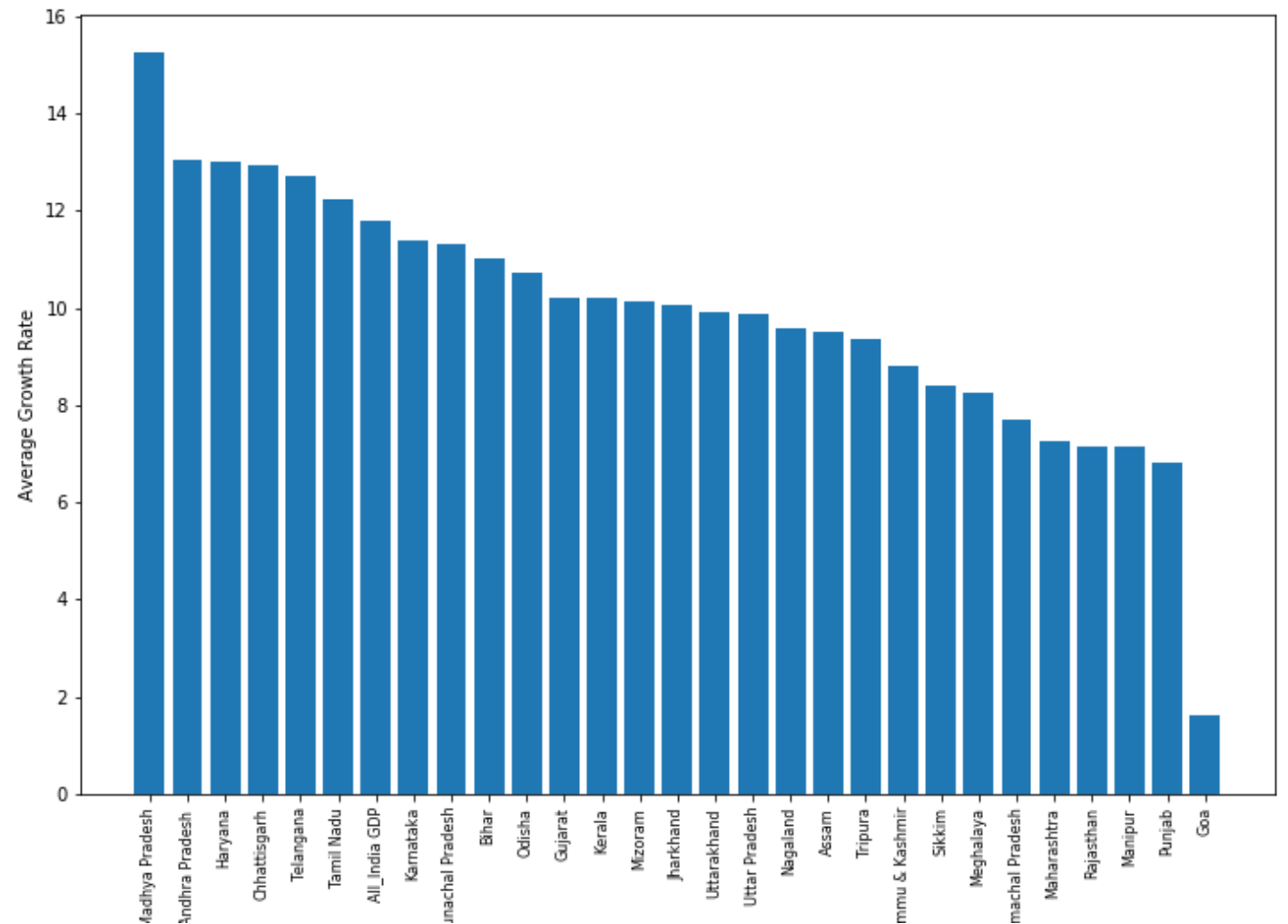


- If the line in the graph is going upwards, then we can say that it's a positive trend.
- In above graph, the line is going downwards, which shows that the trend is Negative. From this we can state that, “% Growth over previous years” i.e. GDP of our nation is coming down over the years and needs attention.

- How will you compare the growth rates of any two states?
 - When plotting the Best Fit line graphs, slopes for each state has been calculated.
 - We can compare slopes of any two states, and reach to conclusion if the state is growing compared to the other or not.
 - Also, by referring the above graphs, we can compare the growth rate of two states.
 - Below is the graph plotted for slopes for each state. From this we can easily compare two states.



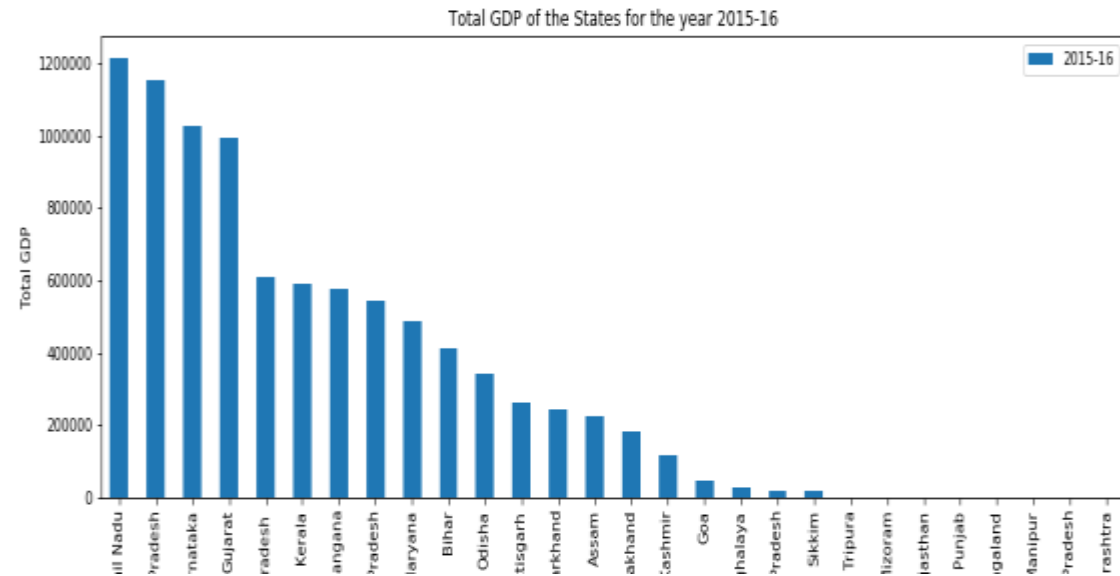
- Which states have been growing consistently fast, and which ones have been struggling? Rank top 3 fastest and 3 slowest-growing states.
 - Calculate the Average of “% Growth over previous years” for each state.
 - Store the result in new column “ Average Growth Rate”
 - Sort the column.
 - After sorting the column in descending order, top 3 states are fastest growing and bottom 3 are slowest growing states.
 - Also this can be represented visually.
- With reference to the above graph
 - Top 3 Fastest Growing States:
 - Madhya Pradesh
 - Andhra Pradesh
 - Haryana
 - Bottom 3 Slowest Growing States:
 - Goa
 - Punjab
 - Manipur



- What is the Nation's growth rate?
 - As observed in the bar graph visual, Nation's growth rate is represented by "All India GDP".
 - The nations growth rate is 11.798%.
- Maharashtra's (Home State) Growth Rate
 - Maharashtra's growth rate over the years is 7.262%.
- Home State VS Nation's Growth Rate
 - When observed in the bar graph visualization, we can say that India's GDP growth rate is better than Maharashtra's GDP growth rate.
 - Maharashtra's GDP growth rate is 7.262 % and India's growth rate is 11.798 %.

Q2: Plot the total GDP of the states for the year 2015-16:

- Filtering out the data for Year 2015 – 16 particularly, plotted the bar graph “Total GDP” vs “States”
- Here "Bar Graph" is plotted which gives distribution of Total GDP of States for Year "2015-16" as it will be easier for anyone to understand just by looking at the visualization.
- Also we can rank the states with decreasing order of total GDP in Bar graph. Due to which "bar graph" is preferred here over other graphs.



- Top 5 States:
 1. Tamil Nadu
 2. Uttar Pradesh
 3. Karnataka
 4. Gujarat
 5. Andhra Pradesh
- Bottom 5 States:
 1. Sikkim
 2. Arunachal Pradesh
 3. Meghalaya
 4. Goa
 5. Jammu And Kashmir
- Since no data is available for below states, these states cannot be ranked as bottom states:
 1. Tripura
 2. Mizoram
 3. Rajasthan
 4. Punjab
 5. Nagaland
 6. Himachal Pradesh
 7. Maharashtra

- We can see that States like Goa, Jammu And Kashmir, Sikkim, Meghalaya, and Arunachal Pradesh are ranked to the bottom as clearly they are not performing well.
- If we see the % Growth over previous years for Goa, we can see that GDP for Goa was performing very bad over the years, and has started recovering.
- Jammu and Kashmir has been ranked at the bottom as well because the travel industry is affected every time there is some situations at the border and affecting the GDP.

Part 1 - B:

- Here data is distributed in different CSV files. Loading data one by one is not a feasible process.
- Hence, data is loaded in to a dataframe using loops.
- All data is present stored in a folder names “GDP Data”. From this code will pick the files and load it into the dataframe.
- Here also, NA values are replaced with 0 to avoid data spikes.
- Union Territories are removed from the dataframe.

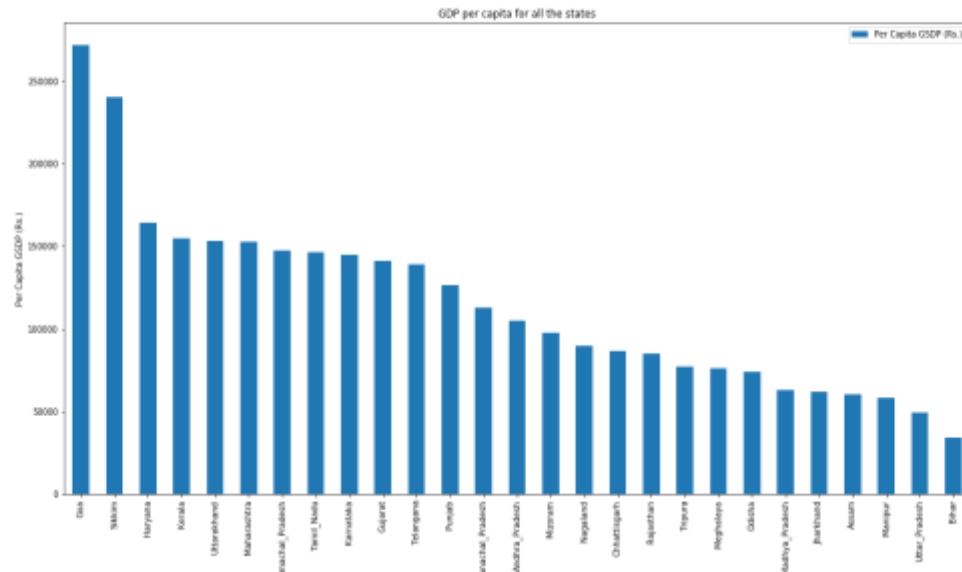
Q1: Plot the GDP per capita for all the states

- Top 5 States:

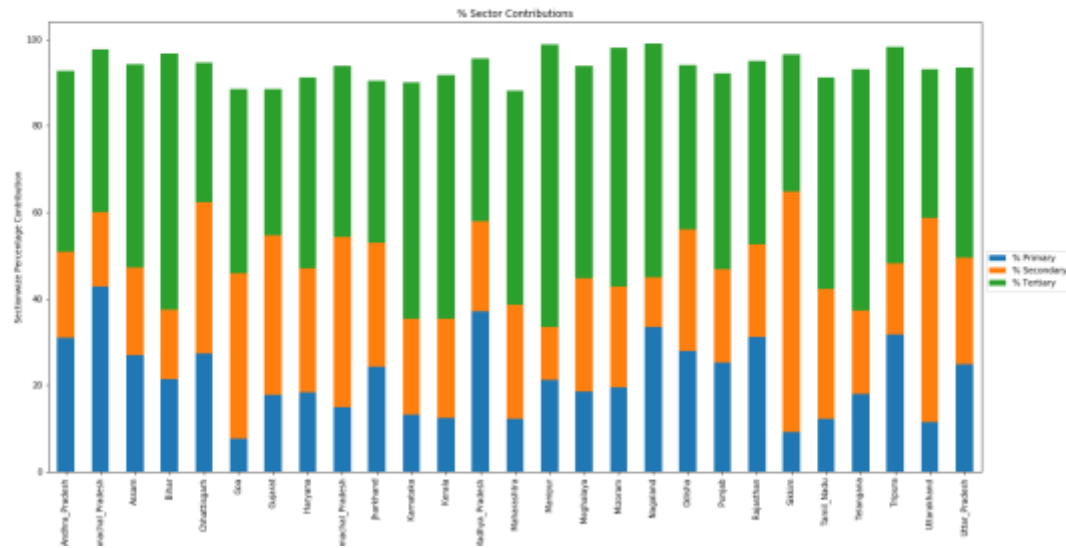
1. Goa
2. Sikkim
3. Haryana
4. Kerala
5. Uttarakhand

- Bottom 5 States:

1. Bihar
2. Uttar Pradesh
3. Manipur
4. Assam
5. Jharkhand



Q2: The ratio of the highest per capita GDP to the lowest per capita GDP is :
8.004741709371503



Q3: Plot the percentage contribution of the primary, secondary and tertiary sectors as a percentage of the total GDP for all the states.

- Bar Stacked Chart is suitable in this scenario as we are comparing 3 parameters of all the states.
- Also the 3 Parameters (% Primary, % Secondary, % Tertiary) are part of the "Total GDP" (part of the same quantity). Hence using Stacked chart will let us easily compare the values with other states.

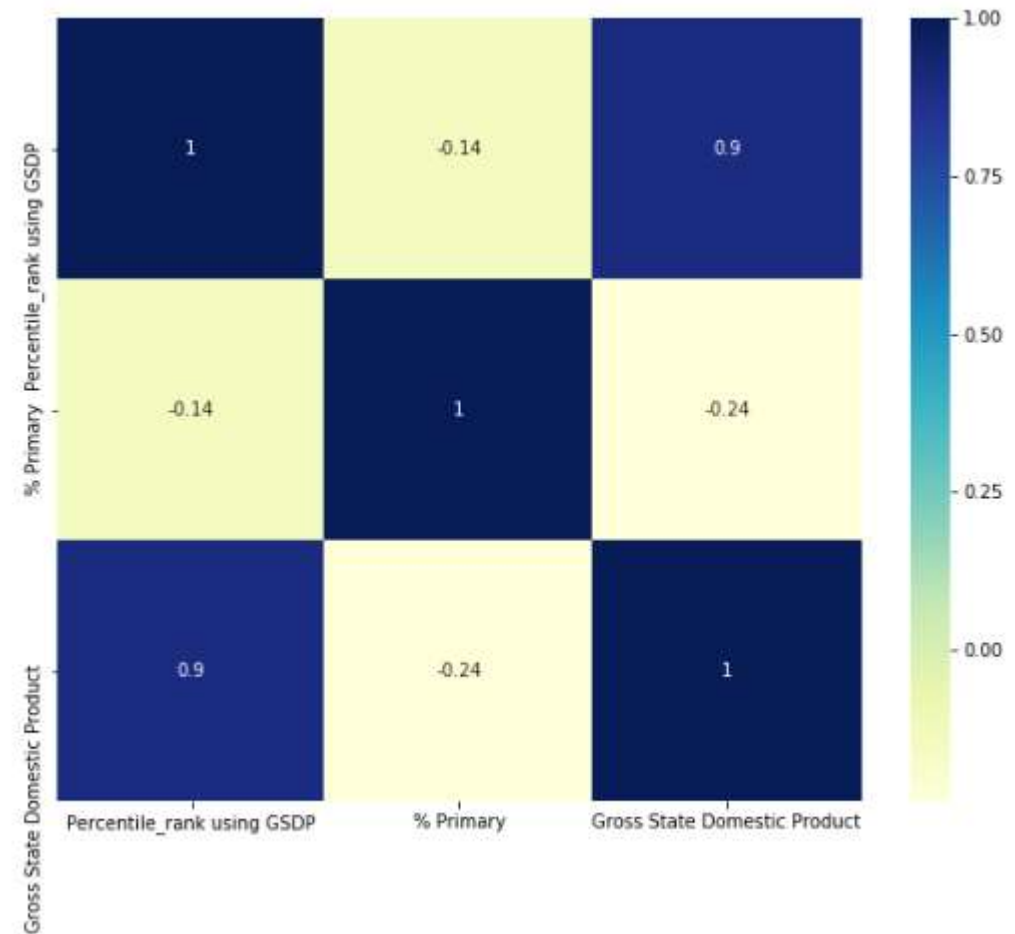
Q4: Categories the states into four groups

- States are categorized in 4 Tags (C1, C2, C3, C4) based on quantile values (0.2, 0.5, 0.85, 1)

correlation of percentile of the state (% of states with lower per capita GDP) and %contribution of Primary sector to total GDP.

Q: Why is (Primary + Secondary + Tertiary) not equal to total GDP?

- Since (Primary + Secondary + Tertiary) is Nation's way of income from industries. There are other factors also those contribute to the total GDP such as Taxes collected, Income from exporting goods as well as debts to global banks.
- Which is why (Primary + Secondary + Tertiary) is not equal to total GDP.

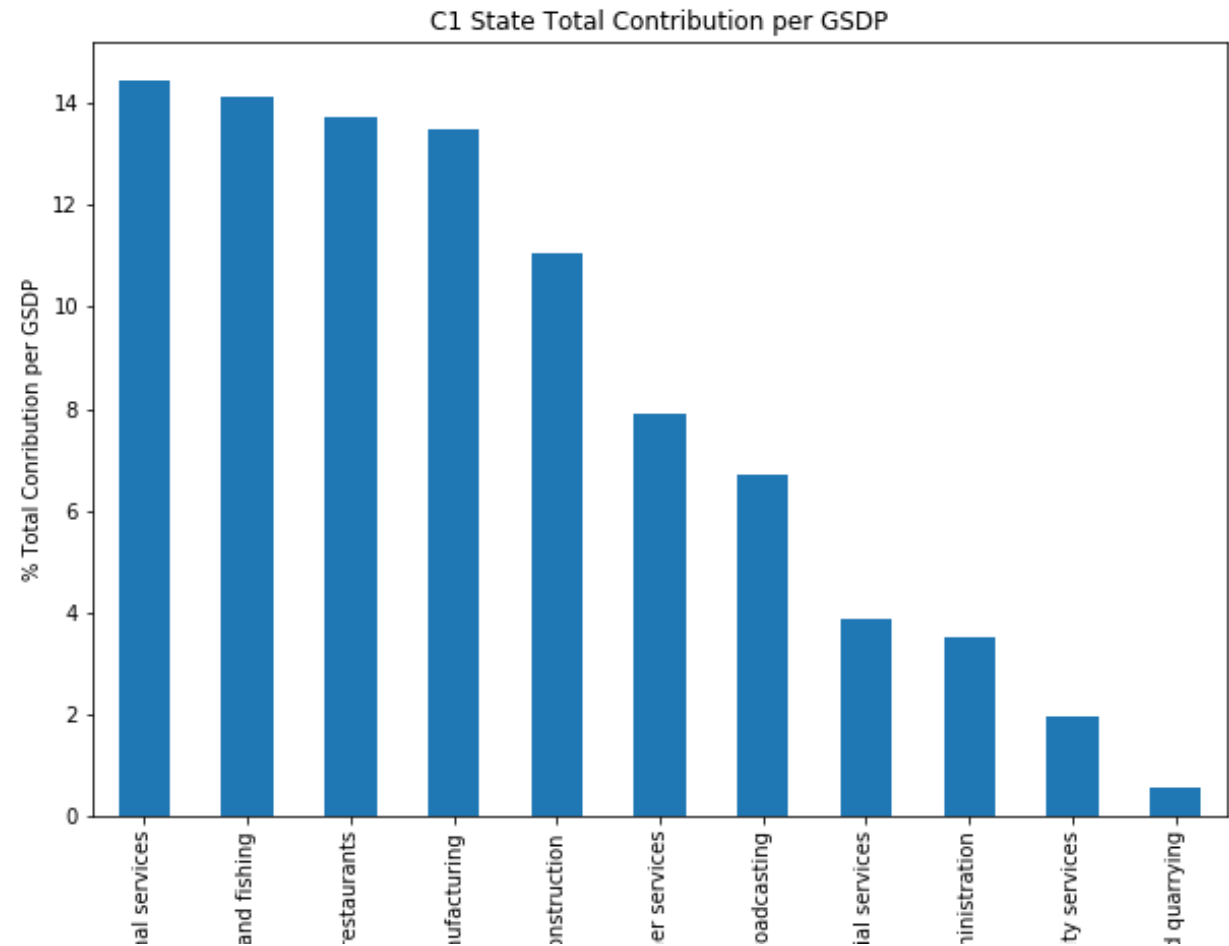


Q5: Find the top 3/4/5 sub-sectors

For C1 States:

1. Below sectors contribute up to 80 % :
2. Real estate, ownership of dwelling & professional services
3. Agriculture, forestry and fishing
4. Trade, repair, hotels and restaurants
5. Manufacturing
6. Construction
7. Other services
8. Transport, storage, communication & services related to broadcasting

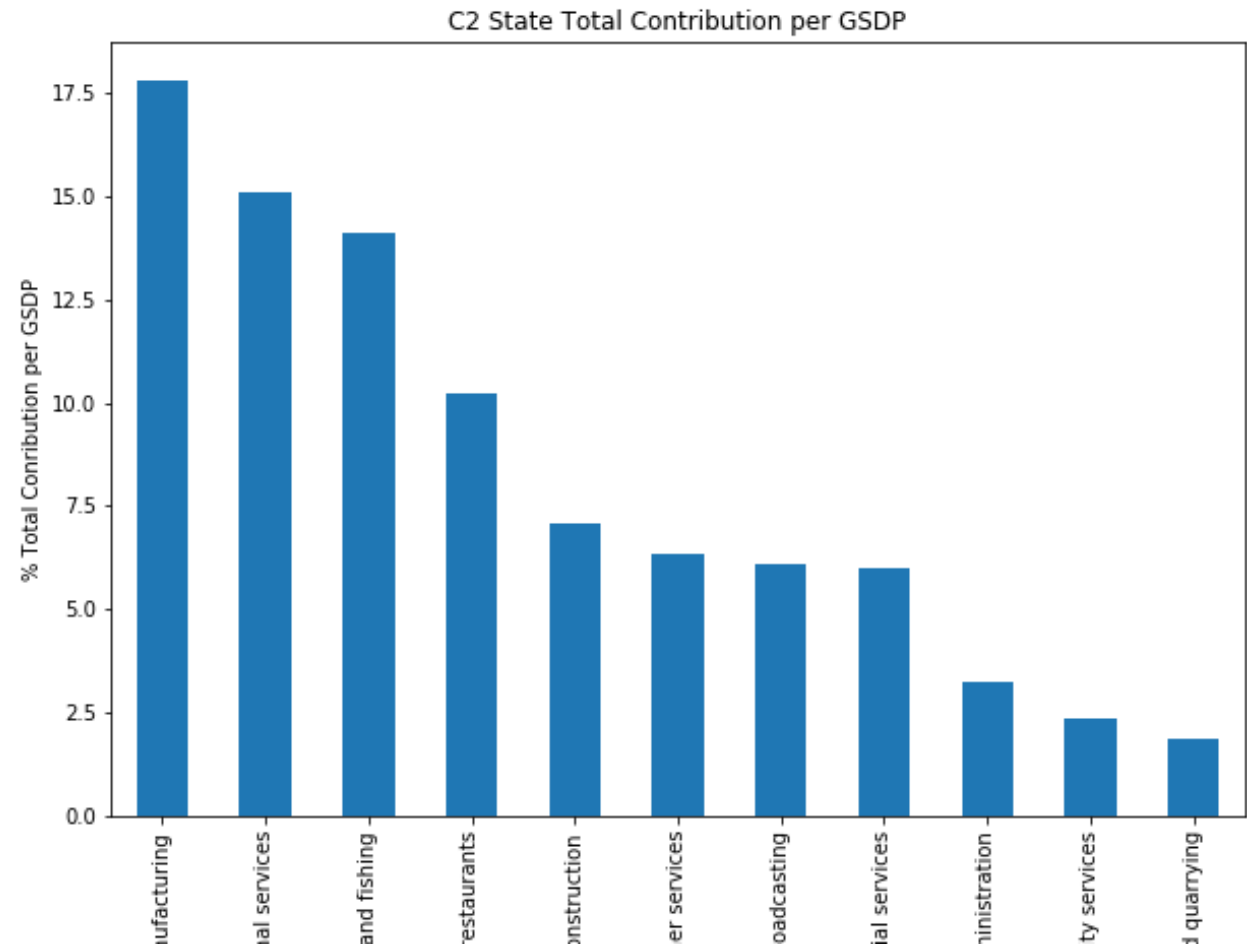
The Bar Chart gives visualization for the different sectors contribution for C1 States.



For C2 States:

1. Below sectors contribute up to 80 %:
2. Manufacturing
3. Real estate, ownership of dwelling & professional services
4. Agriculture, forestry and fishing
5. Trade, repair, hotels and restaurants
6. Construction
7. Other services
8. Transport, storage, communication & services related to broadcasting

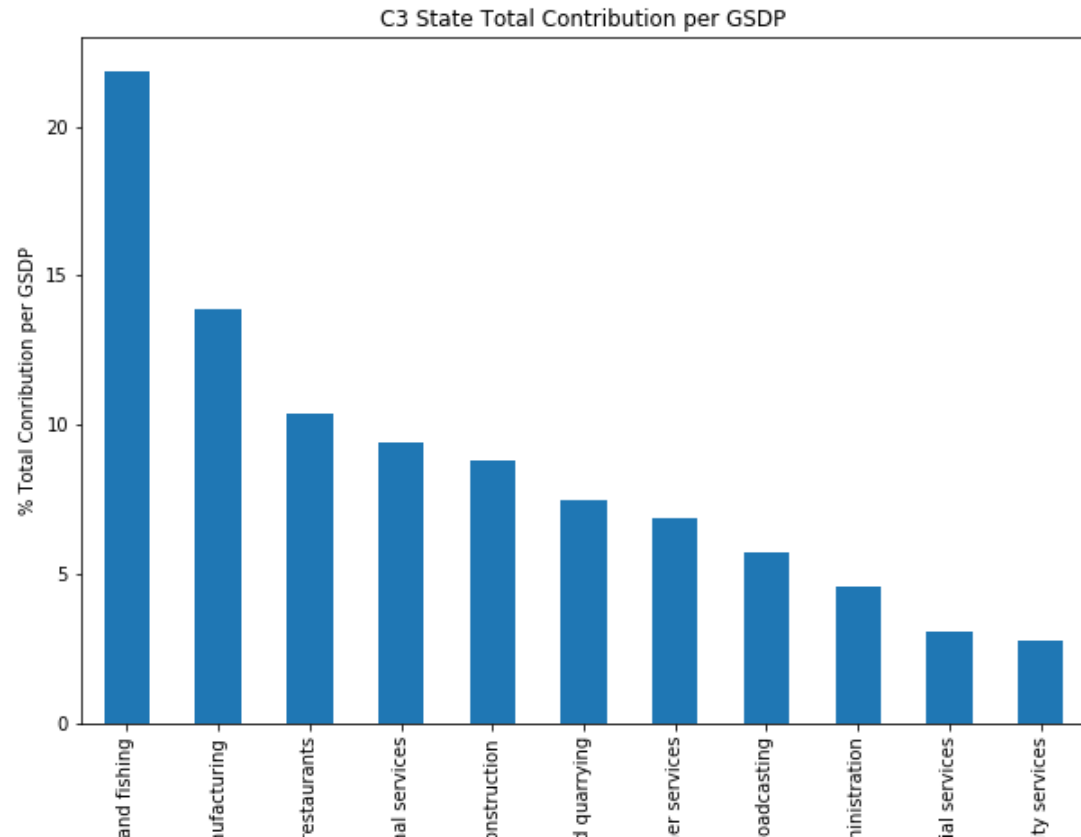
The Bar Chart gives visualization for the different sectors contribution for C2 States.



For C3 States:

1. Below sectors contribute up to 80 %:
2. Agriculture, forestry and fishing
3. Manufacturing
4. Trade, repair, hotels and restaurants
5. Real estate, ownership of dwelling & professional services
6. Construction
7. Mining and quarrying
8. Other services

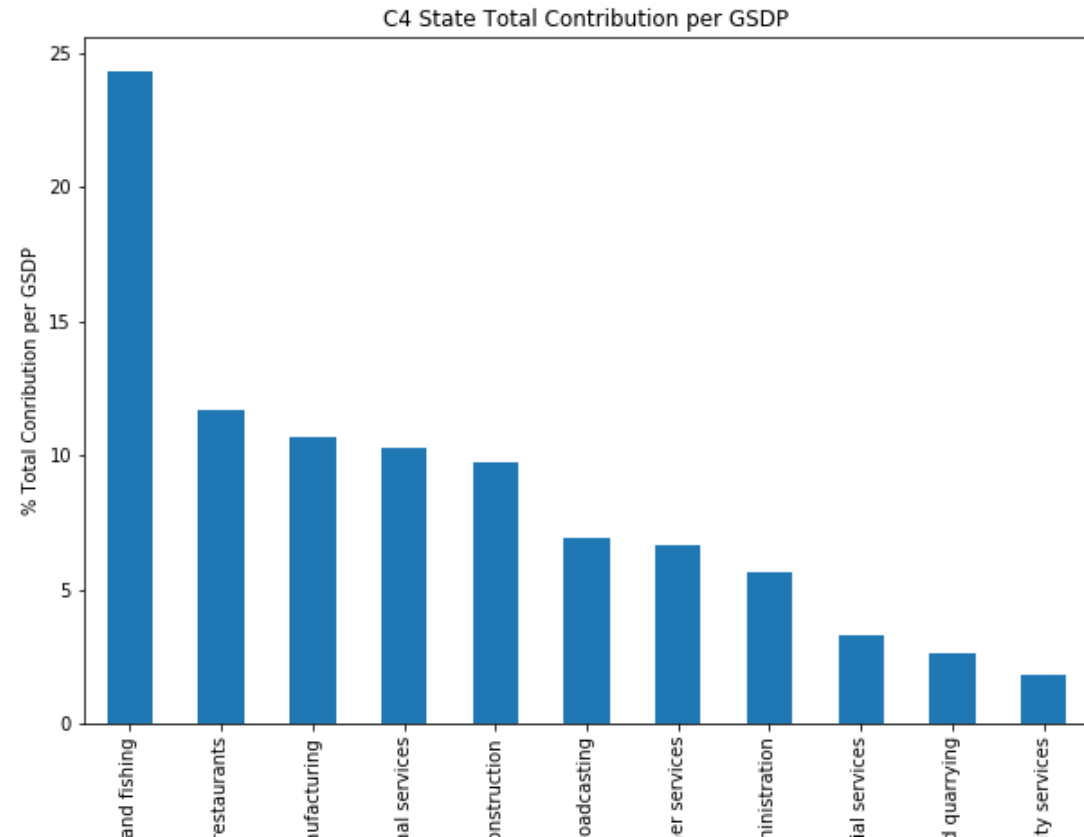
The Bar Chart gives visualization for the different sectors contribution for C3 States.



For C4 States:

1. Below sectors contribute up to 80 %:
2. Agriculture, forestry and fishing
3. Trade, repair, hotels and restaurants
4. Manufacturing
5. Real estate, ownership of dwelling & professional services
6. Construction
7. Transport, storage, communication & services related to broadcasting
8. Other services

The Bar Chart gives visualization for the different sectors contribution for C4 States.



Points to be noted:

- It must be observed that “Agriculture” sector is in top 2 for all the states. This sector contributes to a large amount of GDP for all the sectors. Hence it can be said that if there is drought or limitation of water, it will affect GDP of all the states as well as India’s GDP.

Hence special oversight and funds must be allocated for the “Agriculture” sector.

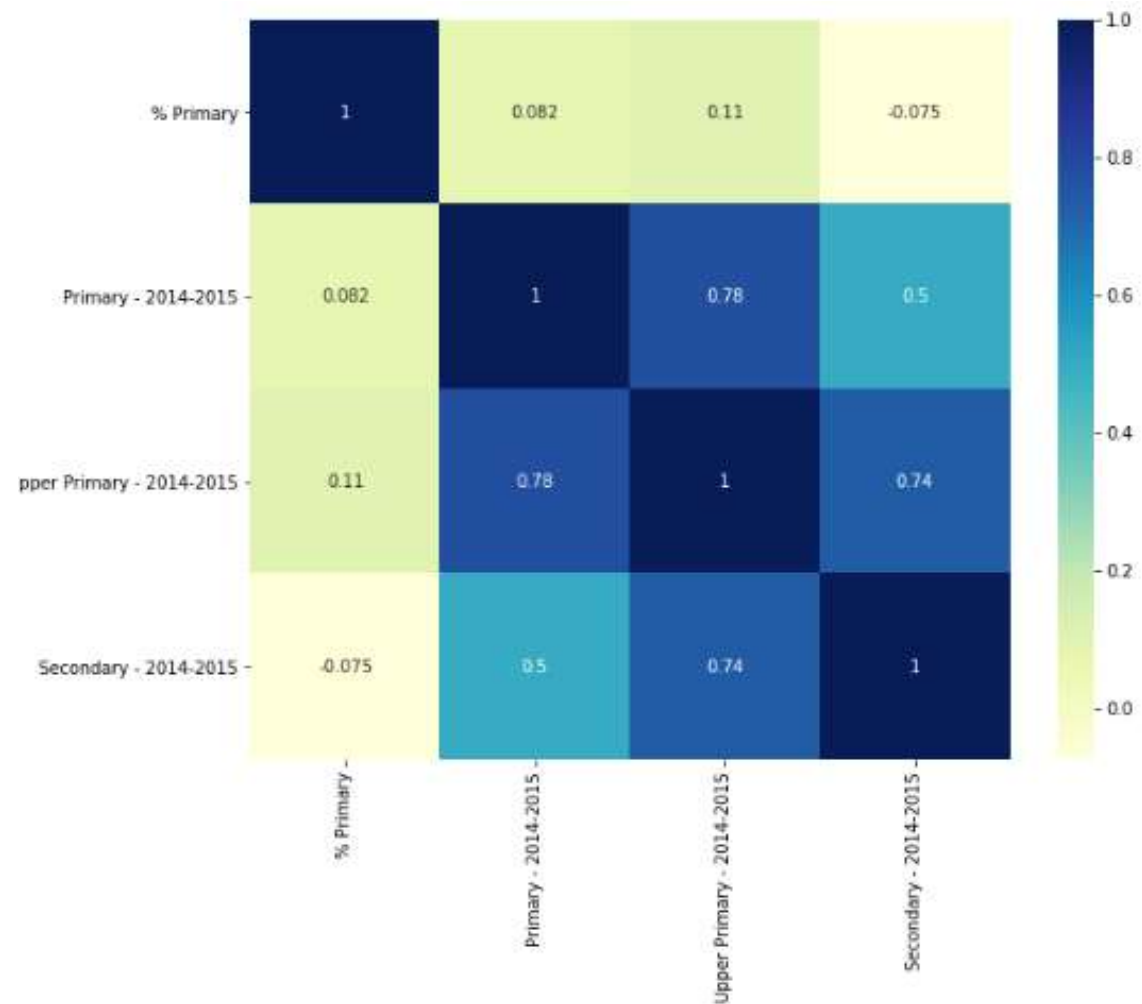
- C1 States also gets special contribution from “Real Estate “sector. From this we can conclude that people are moving to these states and this is giving rise to this sector. Also we can state that population of these states will grow rapidly, hence care must be taken that resources such as water supply are in check. If not then we need to work in that directions.
- For C2 States, “Manufacturing” sector is important while for C3 states its second important. From this we can say that people who work in industries are more likely to come to this city for work. In this case, they should be given proper attention and enough opportunities for job.
- Also C3 States have “ Mining” Sector contributing which is clearly sign that people those who work on daily basis are more in these cities. If more efforts are put here, mining sector can become top sector for C3 States.
- Mining Sector, Construction Sector are among sectors those perform poorly for all states. This condition can be improved by providing more funds for these sectors and creating new job opportunities for people.

Part 2: GDP and Education

- **Q1.** Analyse if there is any correlation of GDP per capita with dropout rates in education (primary, upper primary and secondary) for the year 2014-2015 for each state. Choose an appropriate plot to conduct this analysis.
 - Is there any correlation between dropout rate and %contribution of each sector (Primary, Secondary and Tertiary) to the total GDP?

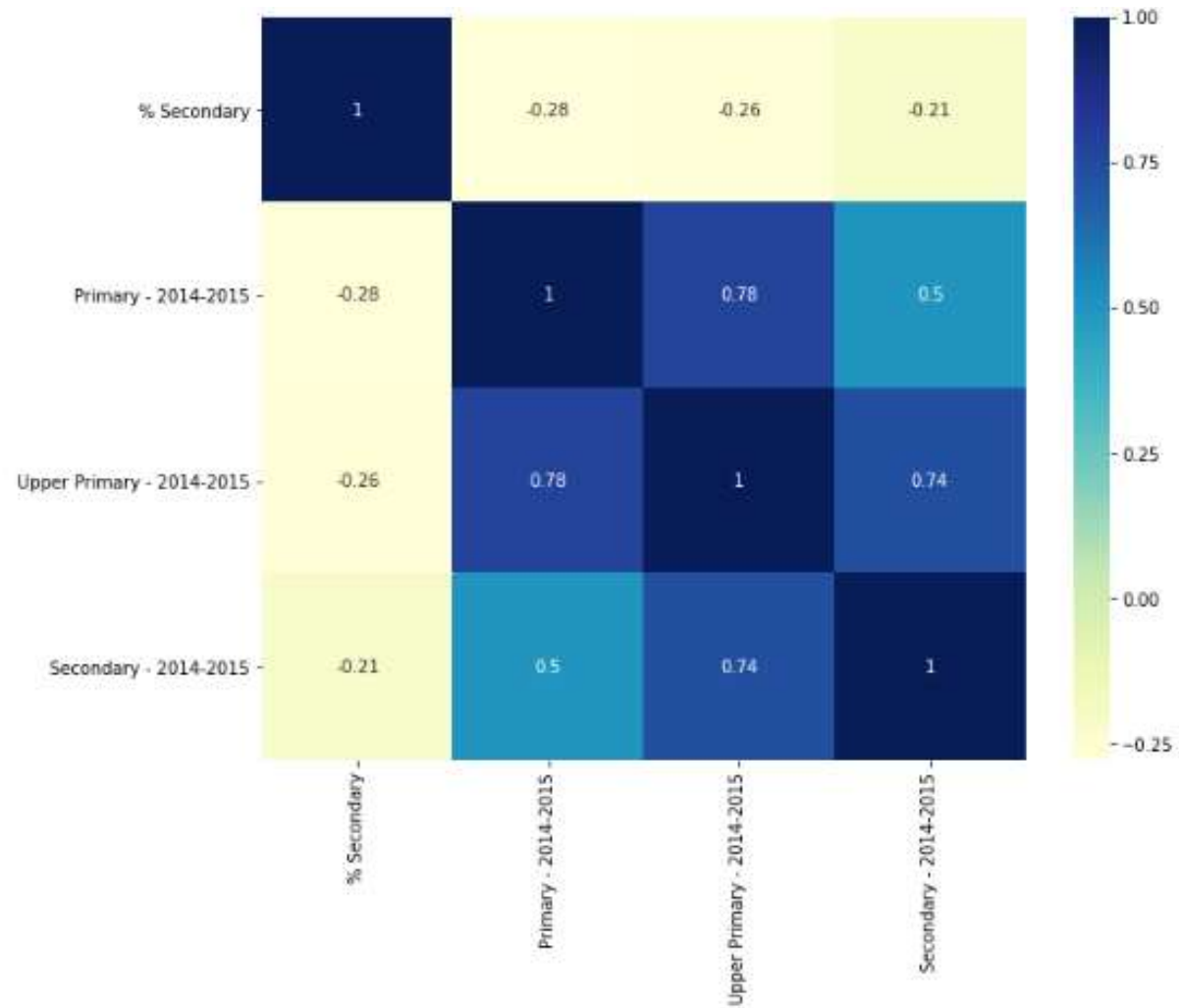
- % Primary Vs Dropout Rates

- Referring above Heatmap plotted for " % Primary " Vs "Dropout Rates (Primary, Upper Primary, Secondary)":
- It can be seen that upper primary is positively correlated with % Primary which means if Primary contribution increases, upper primary drops will increase and vice versa.
- In case of "Secondary", the correlation is negative meaning, increase in % Primary will decrease the dropout rates of Secondary.



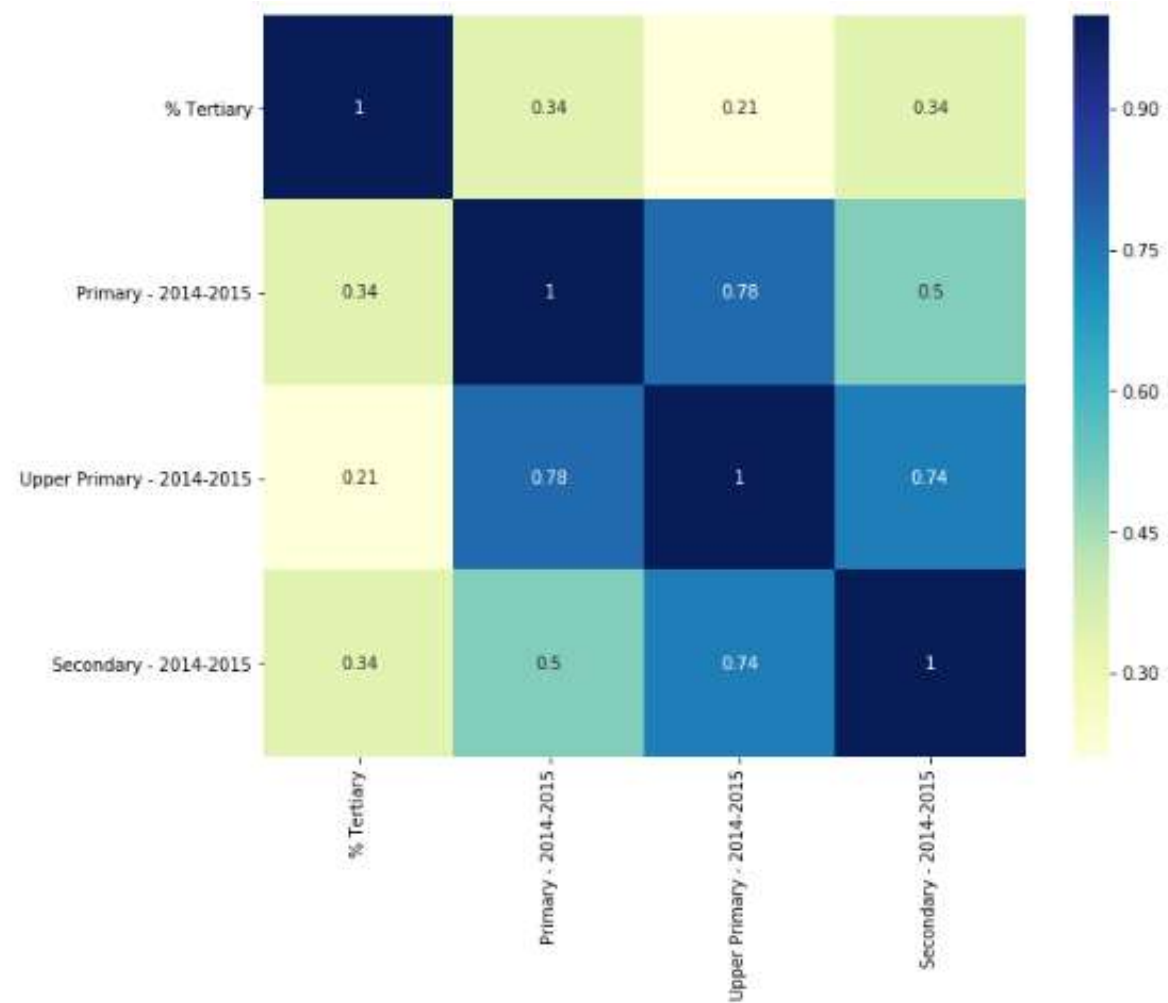
- % Secondary Vs Dropout Rates

- In case of " % Secondary" contributions, dropout will decrease for Primary, Secondary, Upper Primary if secondary contributions increase.

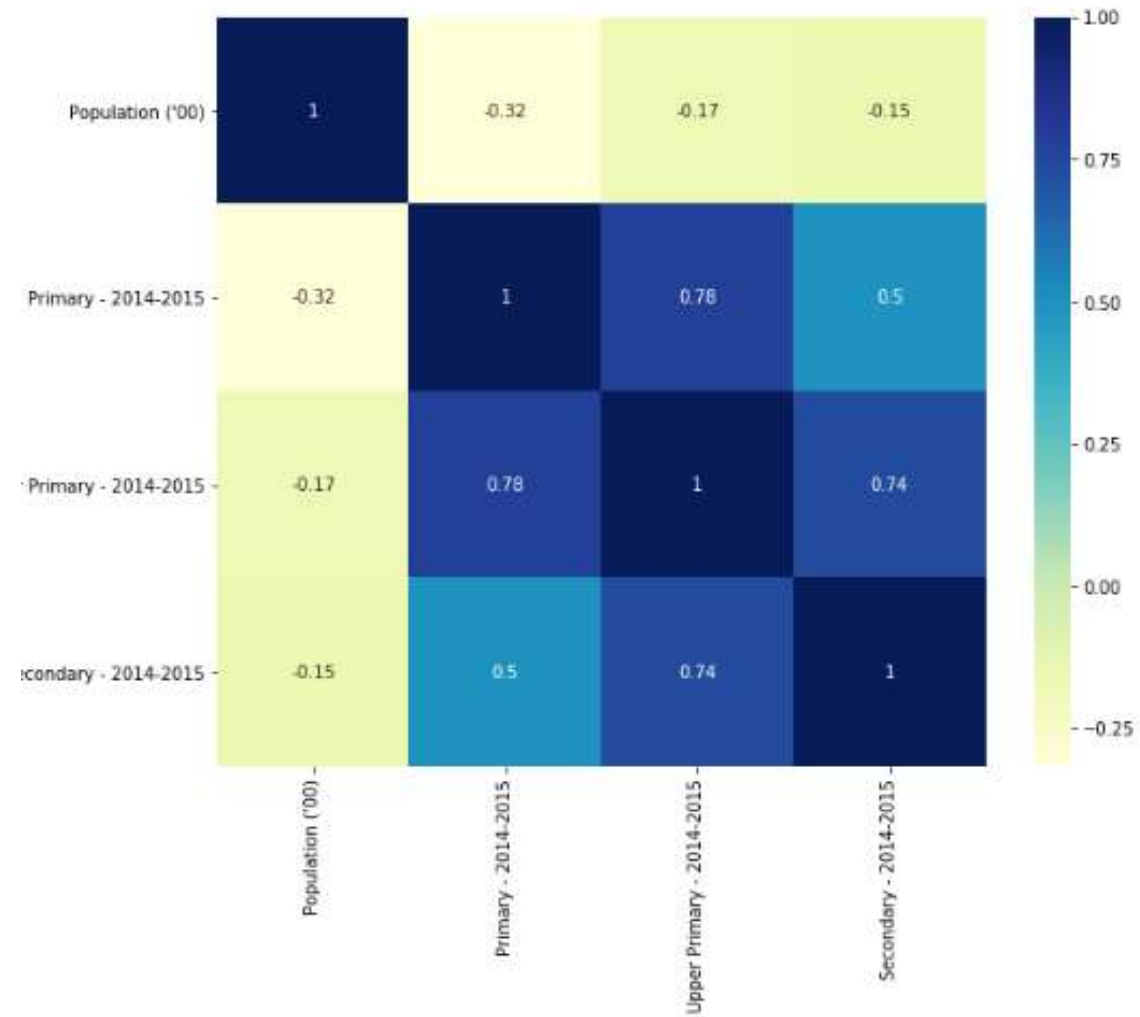


- % Tertiary Vs Dropout Rates

- It can be seen that, behavior of "% Tertiary" contributions is opposite of "% Secondary" contributions.
- In case of secondary, dropouts were decreasing if contribution increased, but in case of tertiary contributions, dropout rate will increase if contributions increases.
- hence it can be said that, tertiary sectors are manual labor sectors where raw manpower is required mainly employing children's above age of 14 due to which dropout rate is increasing.



- Population Vs Dropout Rates
 - Dropout Rates are inversely proportional to the Population
 - As population increases, dropout rates decreases.



Insights

- After analyzing the data for % Primary, % Secondary, % Tertiary contributions, its safe to say that Tertiary Sectors are manual labor sectors. These sectors need attention to be provided as children's working in these sectors are not able to focus on study.
- Due to work pressure and poor financial conditions, dropout rate of Primary, Secondary, Upper Primary is more compared to other sectors.
- Hence more attention needs to be given to this sector which will benefit to create the educated and efficient labor force.