GDP Analysis

This will help in providing recommendation on areas that will foster economic development for their respective states. Since the most common measure of economic development is the **GDP**

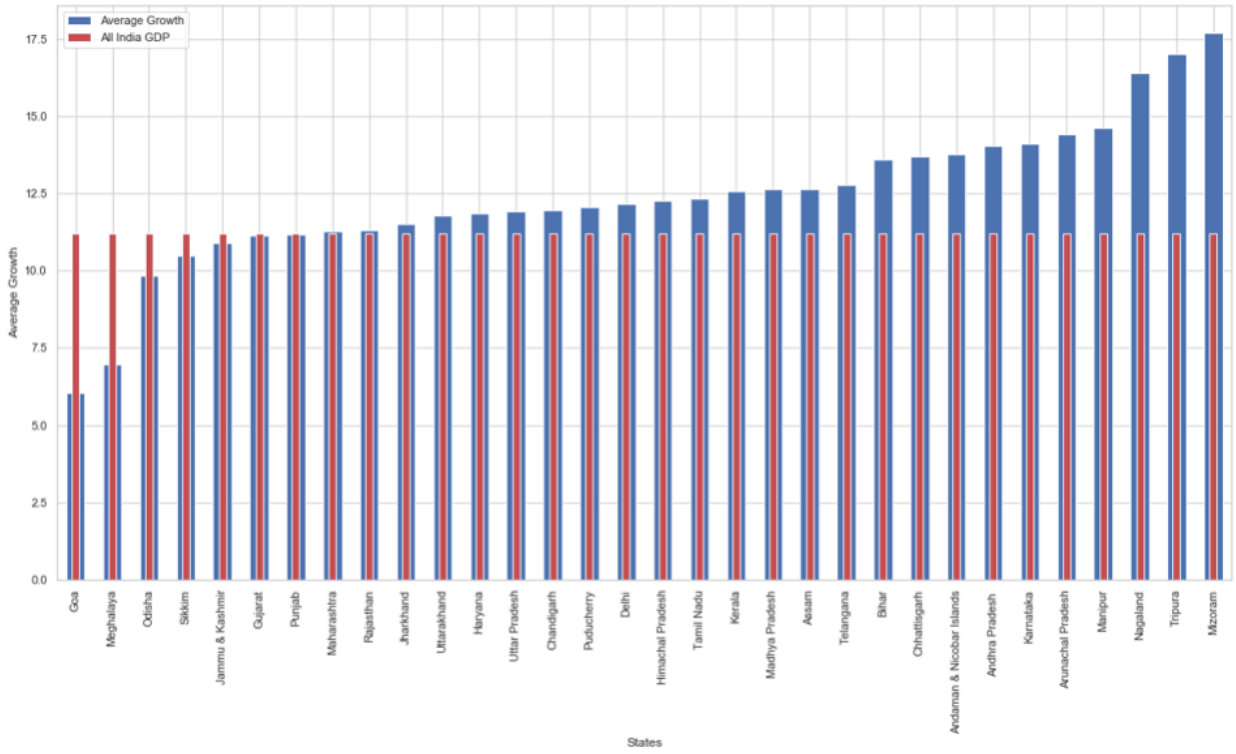
1. The first part was to do Analysis on **Data I-A,** this dataset consists of the GSDP (Gross State Domestic Product) data for the states and union territories.

Data Analysis Phase

- When looking for missing values (across columns), for West Bengal all the columns were empty, hence the column was dropped
- There were no duplicate rows
- Filtered the data to not consider 2016-2017 records,

Analysis of Average of % of Growth (2013-14, 2014-15, 2015-16) vs State

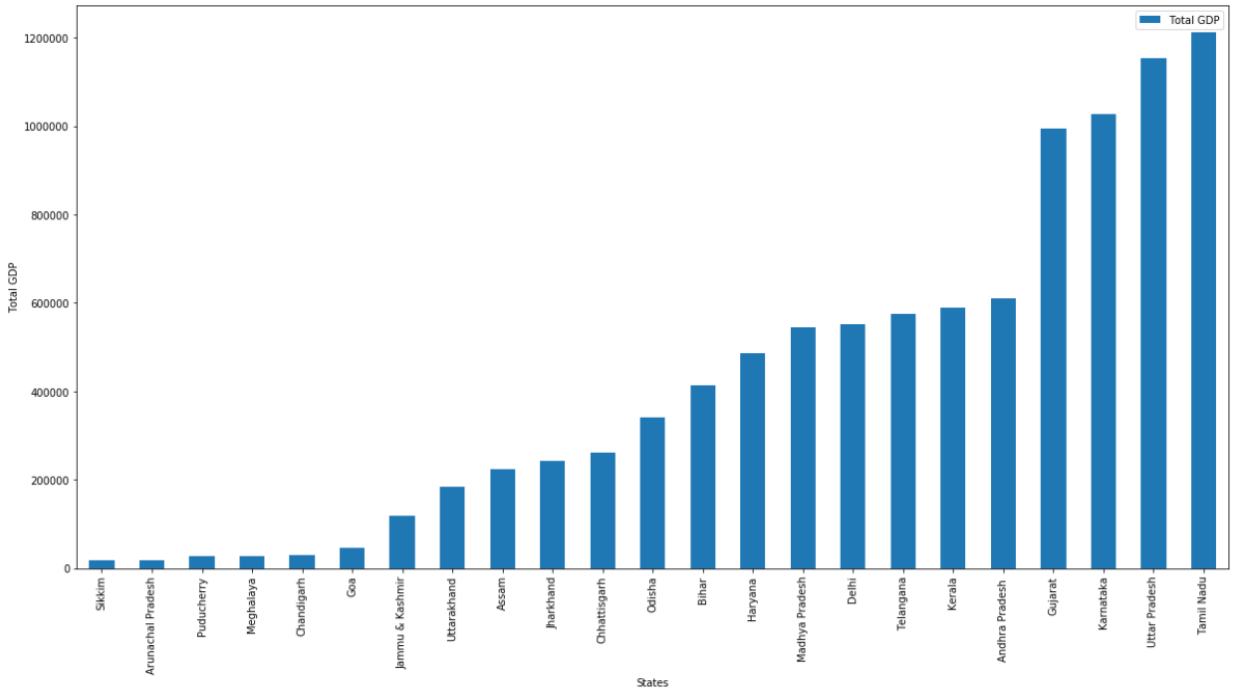
The graph is about Average Growth rate for each of the state and the red bar says how All India GDP is compared to the state



Sorted the graph in the ascending order of Average Growth, as per the plot **Mizoram, Tripura and Nagaland** has the highest average growth.

Analysis Total GDP (2015-16) vs State

There were few states for which the Total GDP was having missing values and after dropping all those states the final plot looks like this.



Total GDP for each **State** in ascending order, as we can see there are few States Tamil Nadu, Uttar Pradesh, Karnataka, Gujrat which contribute **more** towards total GDP, and there are few states which has a very **low** total GDP, Sikkim, Arunachal Pradesh, Puducherry, Meghalaya, Chandigarh & Goa.

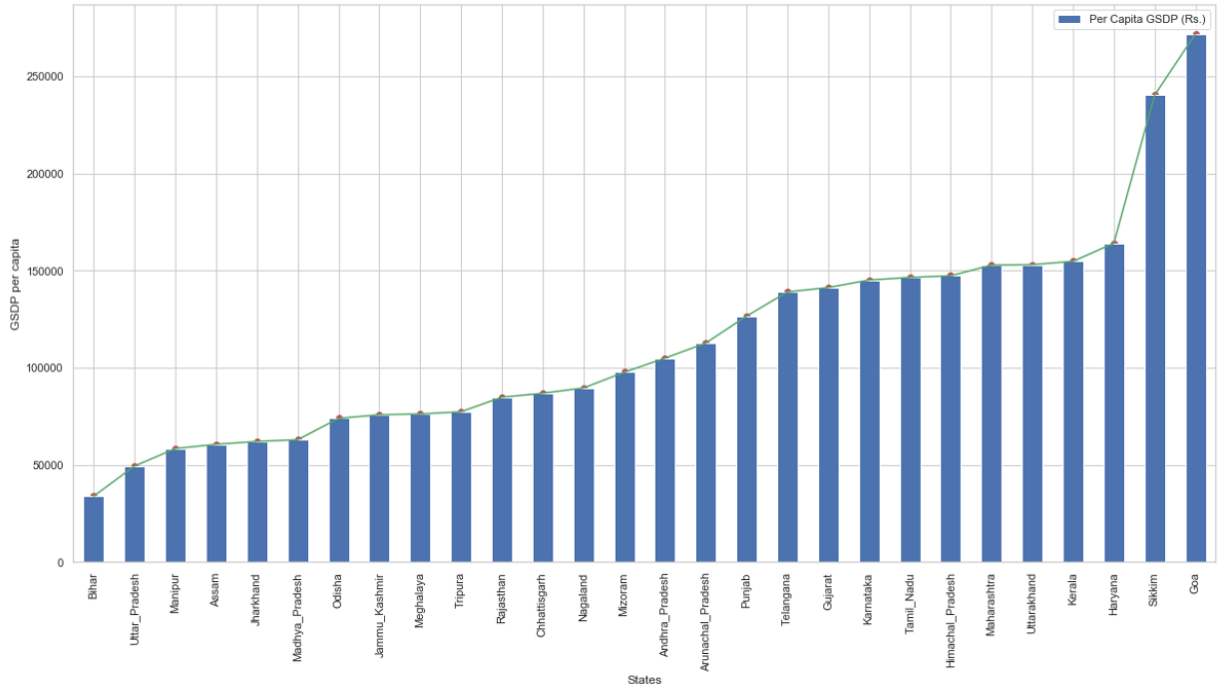
2. The distribution of GSDP among three sectors: the **primary sector** (agriculture), the **secondary sector** (industry) and the **tertiary sector** (services) along with taxes and subsidies for 2014-15

Data Preparation & Data Cleaning

- Collating all the dataset (for all the States) into a single file, for states & Union Territories

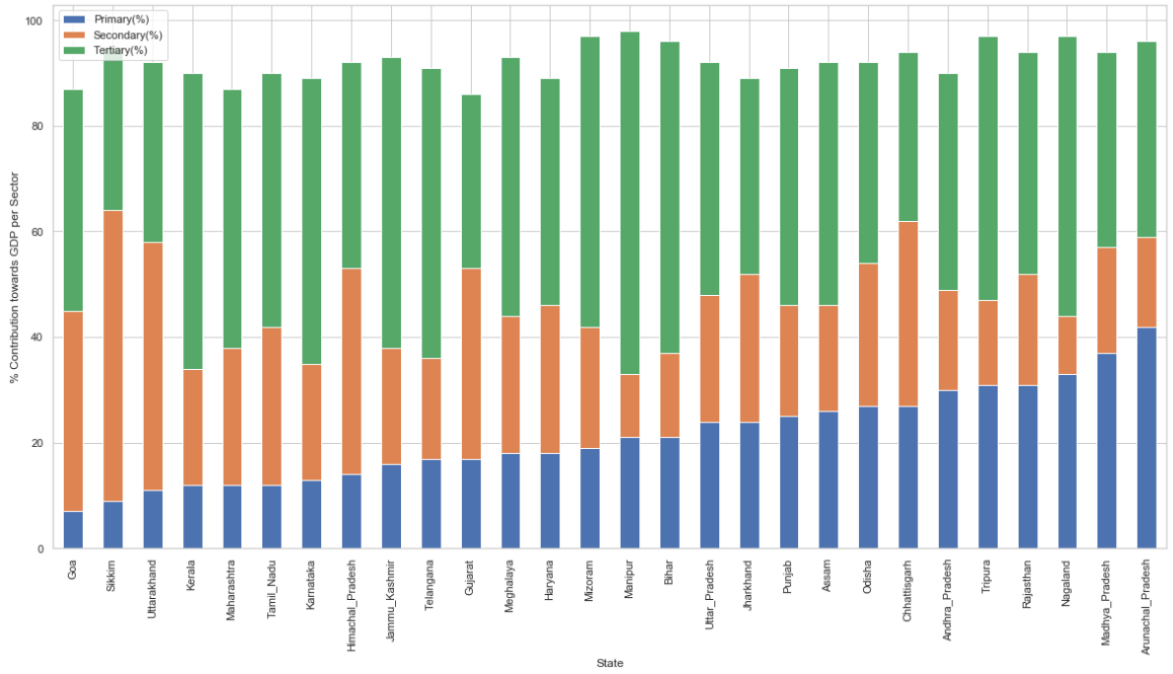
- There are few rows for which Item had some special characters (*), adjust all those items without (*) and put all the values for each of the state to without (*) column
- There are few columns for which there are missing values are present. But as the count was very nominal so I decided to keep as it is.

Analysis of GDP per capita for all the states

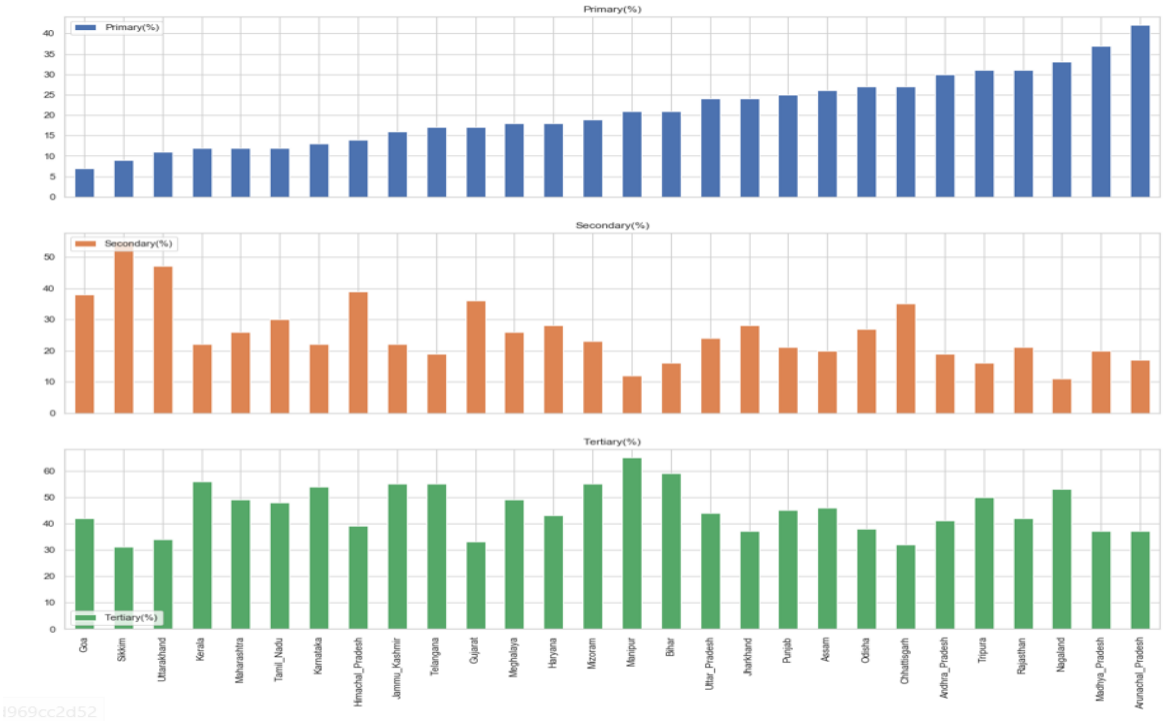


Plotted the graph, GSDP per capita VS States in ascending manner Goa, Sikkim, Haryana seems to highest contributor and for Bihar, Uttar Pradesh, Manipur seems to be the lowest contributor

Percentage contribution of Sectors for each state



When a stacked bar chart was plotted we can see the **Tertiary** has more contribution towards GDP, then comes **Secondary** and then **Primary**. It was also concluded in the split bar chart

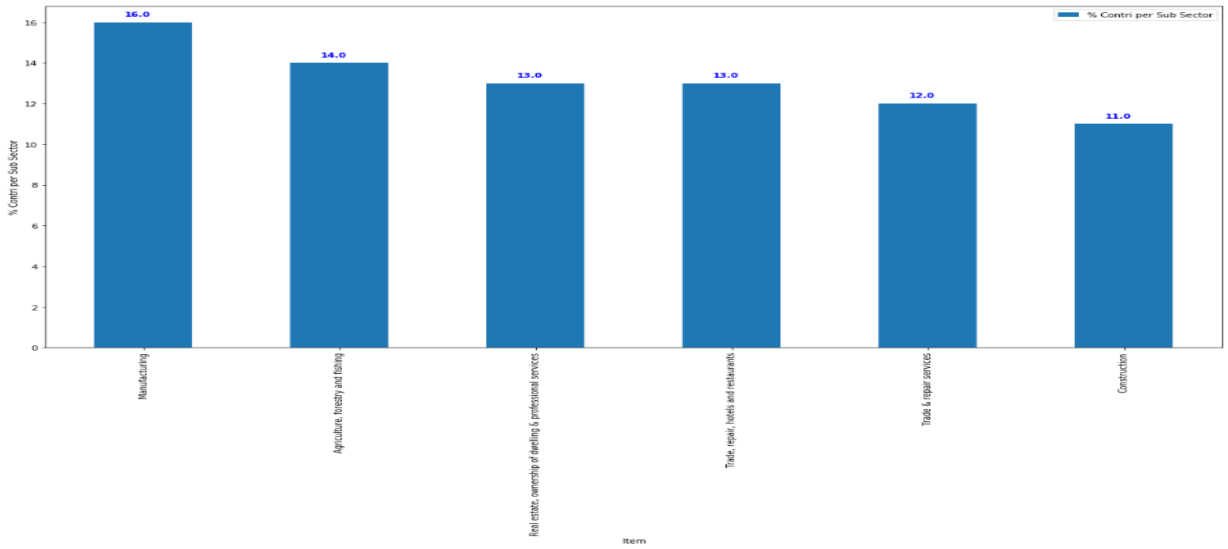


The states were divided into categories on the basis of quantiles, and then GDP contributions by Subsectors were being Compared. The quantile values were (0.20,0.5, 0.85, 1), i.e., the states lying between the 85th and the 100th percentile are in C1; those between the 50th and the 85th percentiles are in C2, and so on.

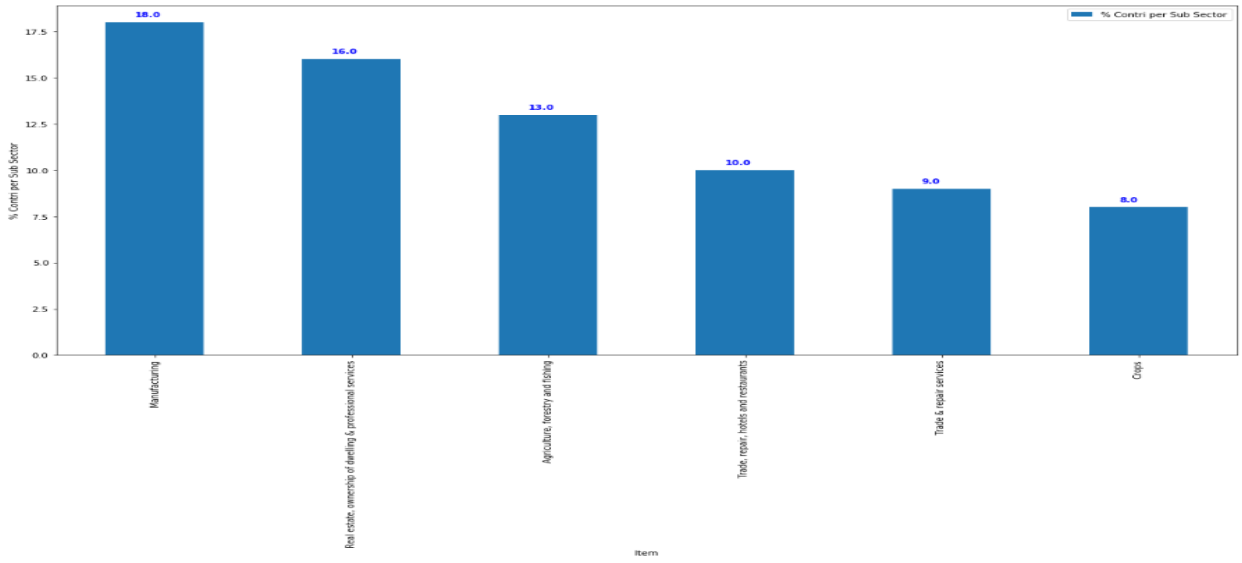
So, the final list of states and the categories are as follows

```
C1 Category States are ['Goa', 'Haryana', 'Kerala', 'Sikkim', 'Uttarakhand']
C2 Category States are ['Andhra_Pradesh', 'Arunachal_Pradesh', 'Gujarat', 'Himachal_Pradesh', 'Karnataka', 'Maharashtra',
'Punjab', 'Tamil_Nadu', 'Telangana']
C3 Category States are ['Chhattisgarh', 'Jammu_Kashmir', 'Meghalaya', 'Mizoram', 'Nagaland', 'Odisha', 'Rajasthan', 'Tripur
a']
C4 Category States are ['Assam', 'Bihar', 'Jharkhand', 'Madhya_Pradesh', 'Manipur', 'Uttar_Pradesh']
```

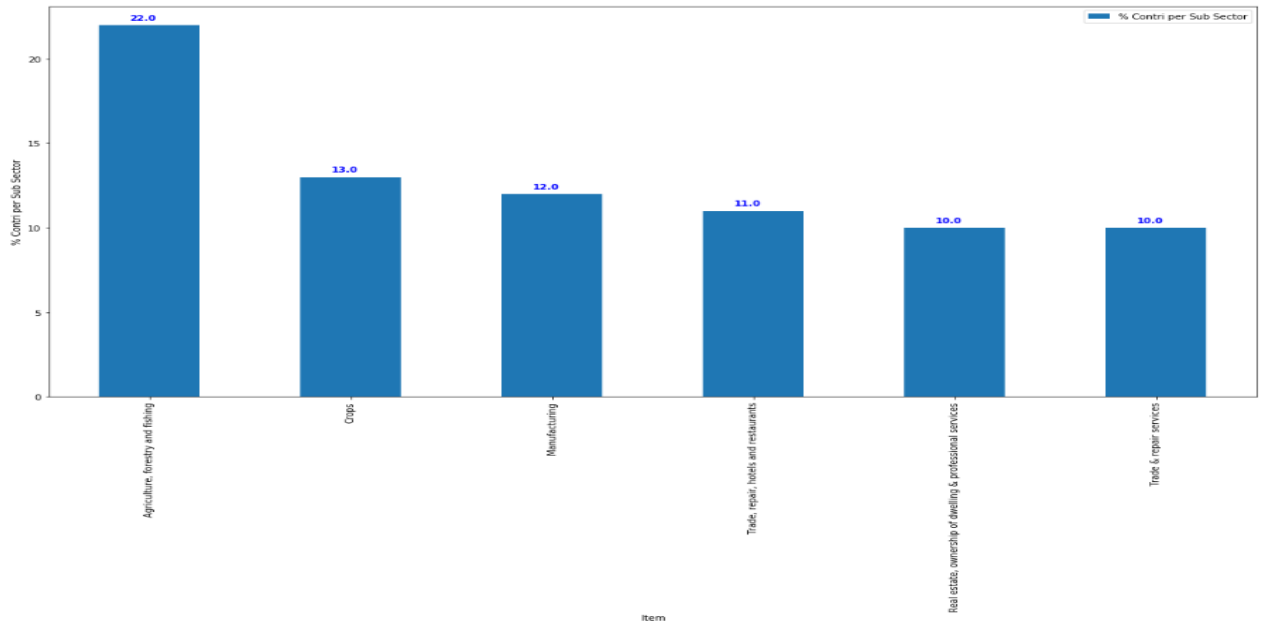
For C1 State Subcategory Contribution Percentage towards GDP



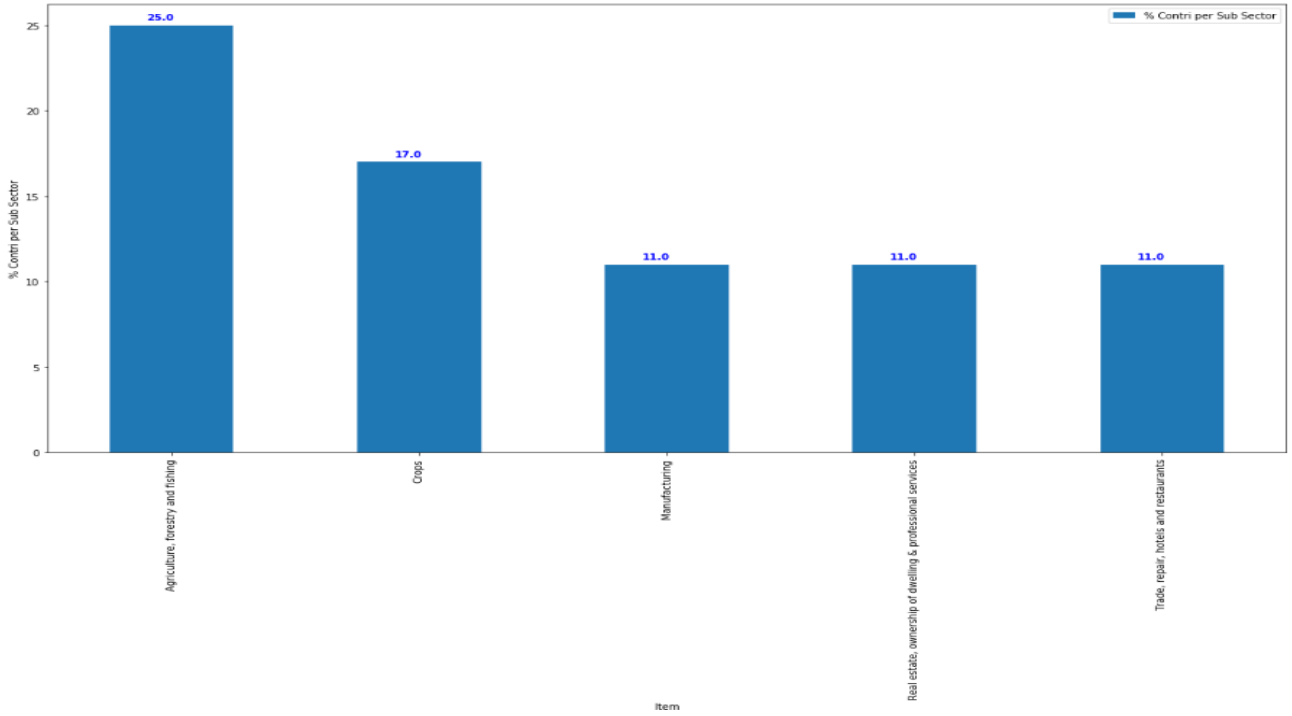
For C2 State Subcategory Contribution Percentage towards GDP



For C3 State Subcategory Contribution Percentage towards GDP



For C4 State Subcategory Contribution Percentage towards GDP



Final Insights which can be drawn from the sub-sectors

1. **Manufacturing**: This sub sector is a high contributor for C1 and C2 sub-sector
2. **Agriculture, forestry and fishing**: This sub-sector is a substantial contributor or a very important one, for each of the categories (For category c1 and c2 it's one of the high contributors and for c3 and c4 it's the highest contributor
3. **Trade, repair, hotels and restaurants**: This sub-sector also is among the top three contributors for each of the category
4. **Trade & repair services**: It's also an sub-sector to focus on because their contribution is (12%, 9% and 10%) for c1, c2 and c3 category state
5. **Real estate, ownership of dwelling & professional services**: this is one of the top two contributor for c2 category state and for c1 state it's among the top contributor (13%) contribution towards GDP, so for c1 and c2 category this sub-sector can also be focused upon
6. **Crops**: This subsector is the second highest contributor towards GDP for c3 and c4 category states, and for C2 category states this subsector contributes (8%) so for c2 states it needs some upliftment.
7. **Construction**: This is contributing around (11%) towards GDP for C1 category State, so for this is also one of the key ingredients for c1 category state to have more GDP

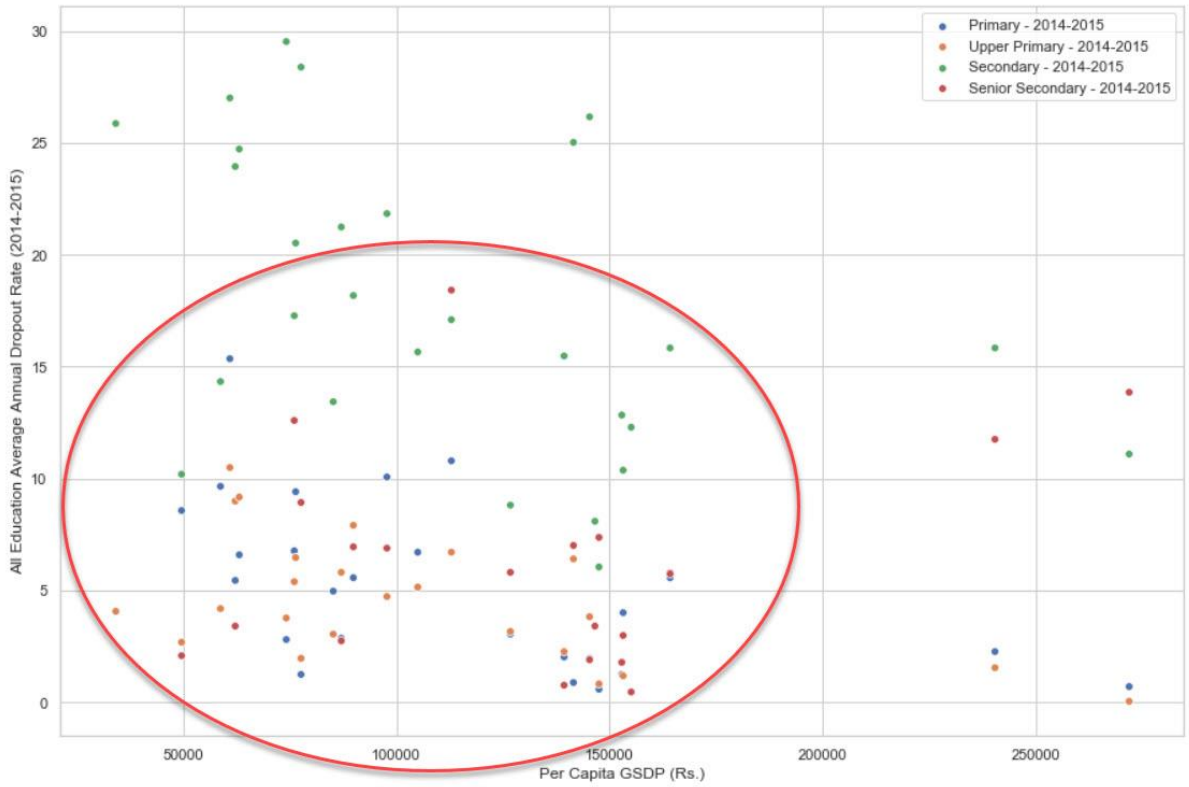
So, a conclusion is being made, as we can see only sub-sector Agriculture, forestry and fishing & Crops are from primary sector, rest all are from Secondary and Tertiary sector. Which we visualize when we plotted between the sector vs % contribution.

3. Comparison between the GDP per capita with dropout rates in education

Data Cleaning and Preparation

- This was only applicable for 2014-15, so all the other data were dropped
- The States name were not identical to Data I-B, distribution of GSDP among three sectors, so renaming all the states.

Analysis of All Education Average Annual Dropout Rate (2014-2015) Vs Per capita GSDP



There are few inferences can be drawn from the figure

- If you see the below plot, most of the points are clustered together, marked inside the circle
- In this cluster, points from different education system are there.
- What we can see below, is for low GDP there is more Dropouts for high GDP there is less drop outs

Final Conclusions

- At the end need drop out rates are indirectly proportional to GDP. If GDP increase Drop rates will decrease
- To increase GDP, there is more investment is requiring in the sub-sectors in Secondary and Tertiary as they are contributing more towards GDP, as well as Primary sector as well
- So, we need to look on few states like Goa, Sikkim or other C1 Category States how we can increase their contribution towards GDP by contributing more towards sectors.