GDP Analysis of Indian States

Prioritize areas of development for different states in India.

Index

Project Brief :-

- The overall goal of this project is to help the CMs of different states focus on areas which will foster economic development for their respective states.
- Since the most common measure of economic development is the GDP, Analyze will done on GDP of the various states of India and suggest ways to improve it.
- There are two parts to this project : Part 1 (A,B), Part 2

Note:-

- We will consider the financial year 2015-16 as the base year for our analysis.

India ranks 11th in the world in terms of total GDP though it lies at the 139th position in terms of per capita GDP.

Part-I: GDP Analysis of Indian States

Note : All the tables shown in the upcoming slides are having just glimpse of the same, To see full table, a hyperlink is provided along with them

Part I-A:

a. Compare the Calculated Average growth for various states (for 2013-16).
b. Analyze and compare the GDPs of various Indian states (both total and per capita).

Important points used in following analysis:-

- The GDP of a state is referred to as the GSDP (Gross State Domestic Product).
- For the analysis Data I-A is used.
  - This dataset contains the GSDP (Gross State Domestic Product) data for the different states and union territories.

Note:-

- File directory : E:\Upgrad\C3 M6 GDP analysis

Look on data received :

Items Description	Duration	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Goa	Gujarat	Haryana
GSDP - CURRENT PRICES (` in Crore)	2011-12	379402	11063	143175	247144	158074	42367	615606	297539
GSDP - CURRENT PRICES (` in Crore)	2012-13	411404	12547	156864	282368	177511	38120	724495	347032
GSDP - CURRENT PRICES (` in Crore)	2013-14	464272	14602	177745	317101	206690	35921	807623	400662
GSDP - CURRENT PRICES (` in Crore)	2014-15	526468	16761	198098	373920	234982	40633	895027	437462
GSDP - CURRENT PRICES (` in Crore)	2015-16	609934	18784	224234	413503	260776	45002	994316	485184
GSDP - CURRENT PRICES (` in Crore)	2016-17	699307	NA	NA	NA	290140	NA	NA	547396
(% Growth over previous year)	2012-13	8.43	13.41	9.56	14.25	12.3	-10.02	17.69	16.63
(% Growth over previous year)	2013-14	12.85	16.38	13.31	12.3	16.44	-5.77	11.47	15.45
(% Growth over previous year)	2014-15	13.4	14.79	11.45	17.92	13.69	13.12	10.82	9.18
(% Growth over previous year)	2015-16	15.85	12.07	13.19	10.59	10.98	10.75	11.09	10.91
(% Growth over previous year)	2016-17	14.65	NA	NA	NA	11.26	NA	NA	12.82

Approach:

- Read data from file.
- Remove rows: '(% Growth over the previous year)' and 'GSDP – CURRENT PRICES (in Crore)' for the year 2016-17.
- Modification Of data for better usage and better understanding.
  - Combine columns 'Items Description' & 'Duration' as 'Description for year‘.
  - Removal of 'Items Description' & 'Duration' columns.
  - Take Transpose of whole data frame.
- Checking for rows and columns for having Null values.
  - As some of the states having Null values are not be computed by either of mathematical means (mean, median, sum etc.) as of there complete unpredictibility,.

Approach continuum :

- As evaluating it is found that (excluding West Bengal) all the Null values are present for the year 2015-16 and for both analysis of Part 1:
  - ✓ For Calculate the average growth of states over the duration 2013-14, 2014-15 and 2015-16

Note:

- There is an option of computing average growth by skipping these NaN values but as % Growth for 1 year makes a lot difference.
  - ✓ For Comparing the total GDP of the states for the year 2015-16.
- Therefore, Data for 2015-16 is required. Therefore all the states having Null values are being drop, States included :-

✓ Himachal Pradesh
✓ Maharashtra
✓ Manipur
✓ Mizoram
✓ Nagaland
✓ West Bengal1
✓ Punjab
✓ Rajasthan
✓ Tripura
✓ Andaman & Nicobar Islands

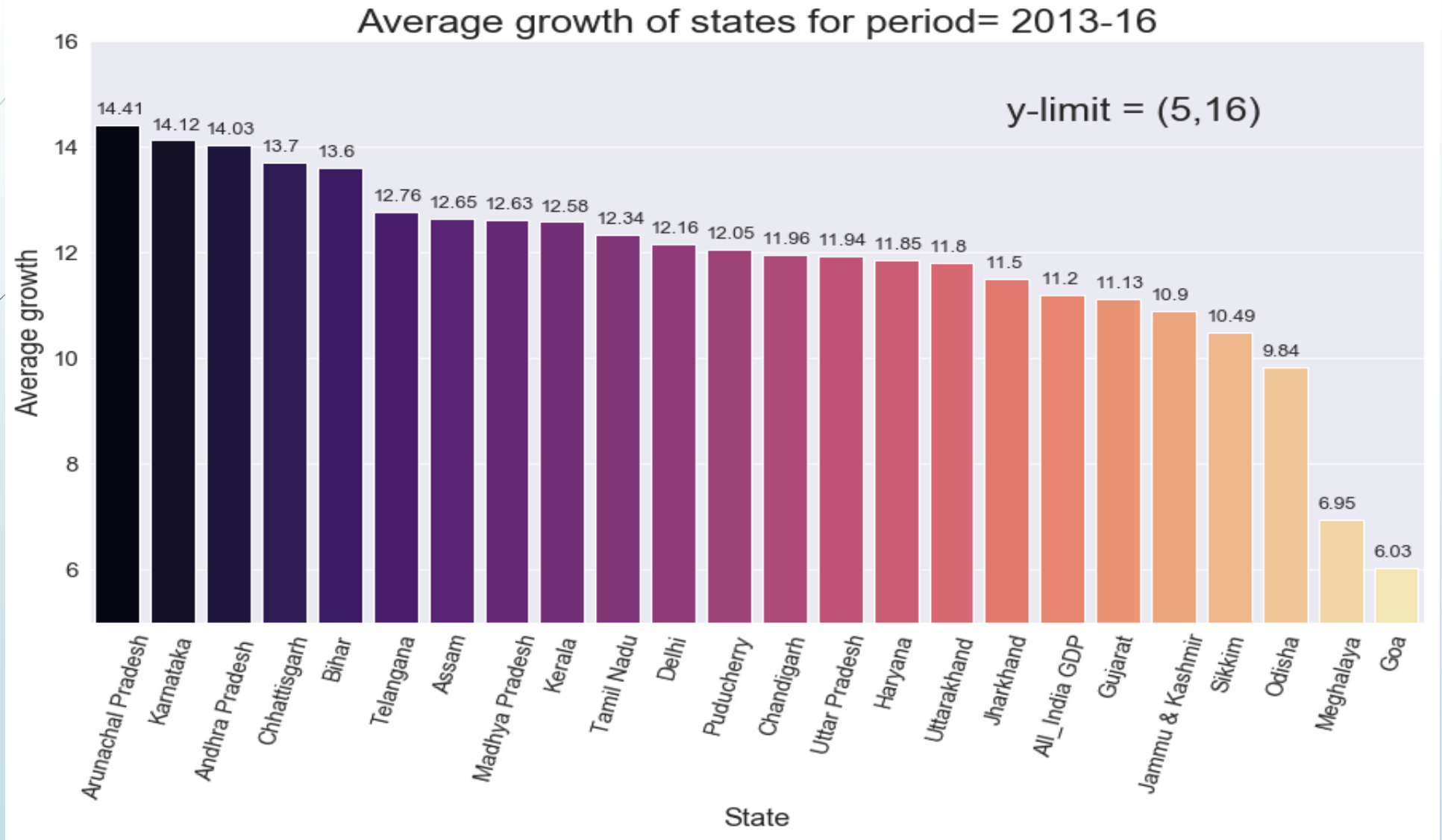
Approach Continuum:

- Removal of rows containing above states.
- Addition of column to data frame defining ‘The average growth of states over the duration 2013-14, 2014-15 and 2015-16.
- Sort the table to decreasing order with respect to Average growth.

Look on new table formed:-

	GSDP - CURRENT PRICES (` in Crore) 2011-12	GSDP - CURRENT PRICES (` in Crore) 2012-13	GSDP - CURRENT PRICES (` in Crore) 2013-14	GSDP - CURRENT PRICES (` in Crore) 2014-15	GSDP - CURRENT PRICES (` in Crore) 2015-16	(% Growth over previous year) 2012-13	(% Growth over previous year) 2013-14	(% Growth over previous year) 2014-15	(% Growth over previous year) 2015-16	Average growth though 2013-16
Arunachal Pradesh	11063	12547	14602	16761	18784	13.41	16.38	14.79	12.07	14.41333333
Karnataka	603778	691700	817886	921788	1027068	14.56	18.24	12.7	11.42	14.12
Andhra Pradesh	379402	411404	464272	526468	609934	8.43	12.85	13.4	15.85	14.03333333
Chhattisgarh	158074	177511	206690	234982	260776	12.3	16.44	13.69	10.98	13.70333333
Bihar	247144	282368	317101	373920	413503	14.25	12.3	17.92	10.59	13.60333333

Plot the comparison of average growth of all states:



Insights from Plot:

a. Which states have been growing consistently fast, and which ones have been struggling ?
   - Consistently fast --> Arunachal Pradesh
   - Most Struggling one --> Goa

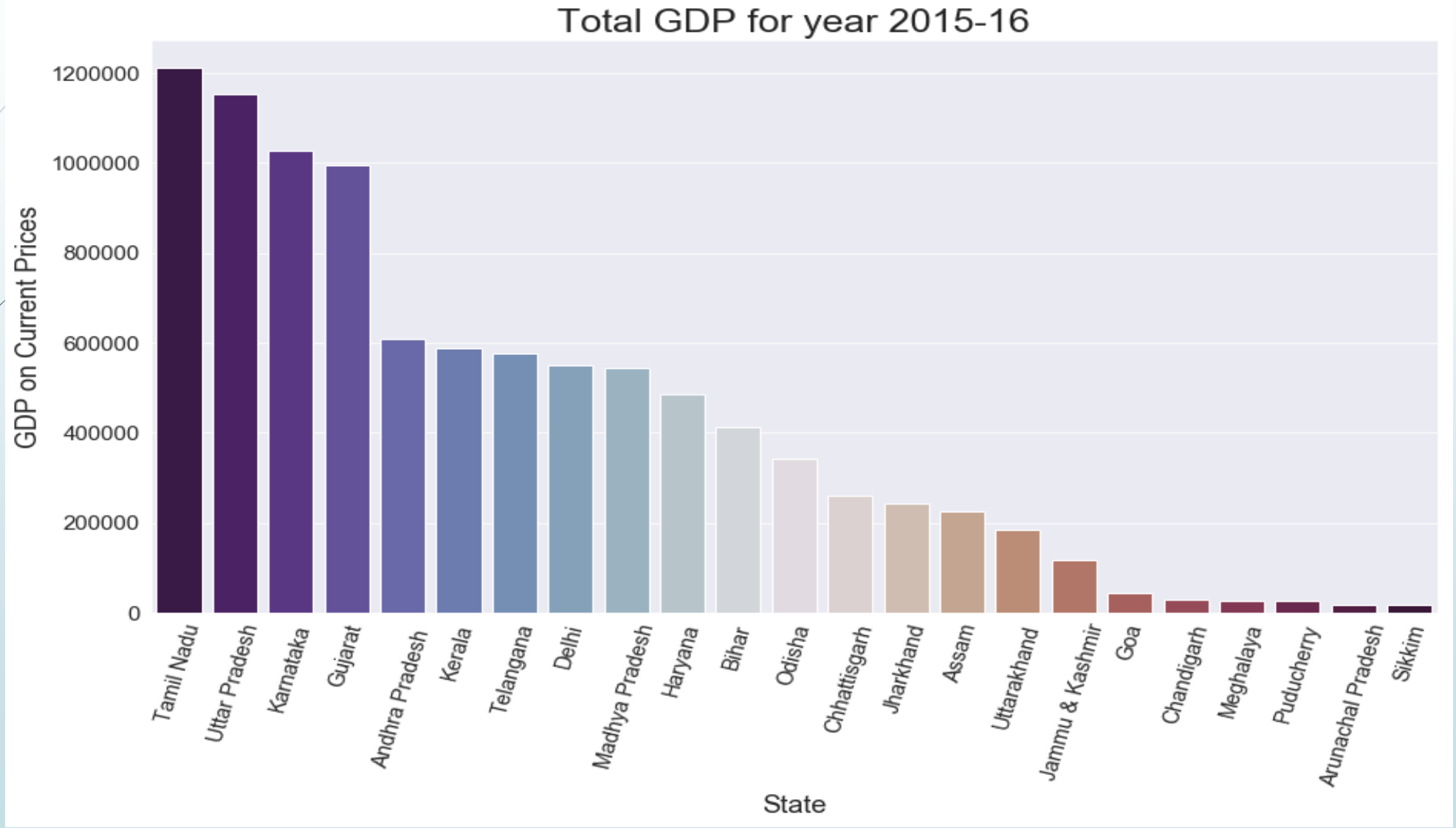
b. What has been the average growth rate of your home state, and how does it compare to the national average over this duration?
   - Home State --> Punjab

   Note: Punjab Data is being discarded due to incomplete...

Approach Continuum:

- Data preparation for comparing GDP for various states.
  - Sorting all the data in decreasing order as respect to GSDP prices for 2015-16.
  - Inclusion All India GDP value due to its very high value.

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



Insights from Plot:

a. Identify the top-5 and the bottom-5 states based on total GDP ?

- The top - 5 :
  - ✓Tamil Nadu
  - ✓Uttar Pradesh
  - ✓Karnataka
  - ✓Gujarat
  - ✓Andhra Pradesh
- The Bottom - 5:
  - ✓ Chandigarh
  - ✓ Meghalaya
  - ✓ Puducherry
  - ✓ Arunachal Pradesh
  - ✓ Sikkim

Part 1B:

a. Analyze and compare the GDP per capita for all the states.

b. Analyze the percentage contribution of primary, secondary and tertiary sectors as a percentage of total GSDP for all the states.

c. Categorize the states into four categories based on GDP per capita (C1 contains states with highest GDP per Capita) and for each of these four categories:

- Analyze the sub-sectors which contribute most to the GDP (such as agriculture, real estate, manufacturing etc.).

Note:- The data uses for analysis:

- There is separate dataset for each state**.**
- Analysis is only done for the duration : 2014-15.
- Data is being filter out for the Union Territories as there data is not required.

Look of the data received :

S.No.	Item	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17
9	Real estate, ownership of dwelling & professional services	2833805	3434098	3897345	4405409	5092310	5942431
10	Public administration	1393752	1552379	1738971	2200897	2594904	3039676
11	Other services	2730376	3063920	3538298	4215389	5248604	6180240
Total	Tertiary	14308520	16592526	19035788	22032942	25445252	29518654
12	TOTAL GSVA at basic prices	34975303	37947671	42672154	48825542	55721862	64169719
13	Taxes on Products	4243900	4656500	5263500	5512100	6719300	7236500
14	Subsidies on products	1279000	1463800	1508500	1690800	1447800	1475500
15	Gross State Domestic Product	37940203	41140371	46427154	52646842	60993362	69930719
16	Population ('00)	492750	495660	498570	501510	504460	507430
17	Per Capita GSDP (Rs.)	76997	83001	93121	104977	120908	137814

Approach :

- Read all csv files into python.

Note: Two Data Frames are Being built from the collected data.

a. Data Frame: gsdp(gross state domestic product):
   - Primary
   - Secondary
   - Tertiary
   - Total GSVA across various sectors
   - taxes
   - Subsidies
   - GSDP
   - Population
   - Per Capita GSDP

b. Data Frame : gsdp_sec(sub-sector wise) :
   - Sub-Sector
   - State
   - Sector
   - GPV (Gross Product value)

Look Upon the formed tables:

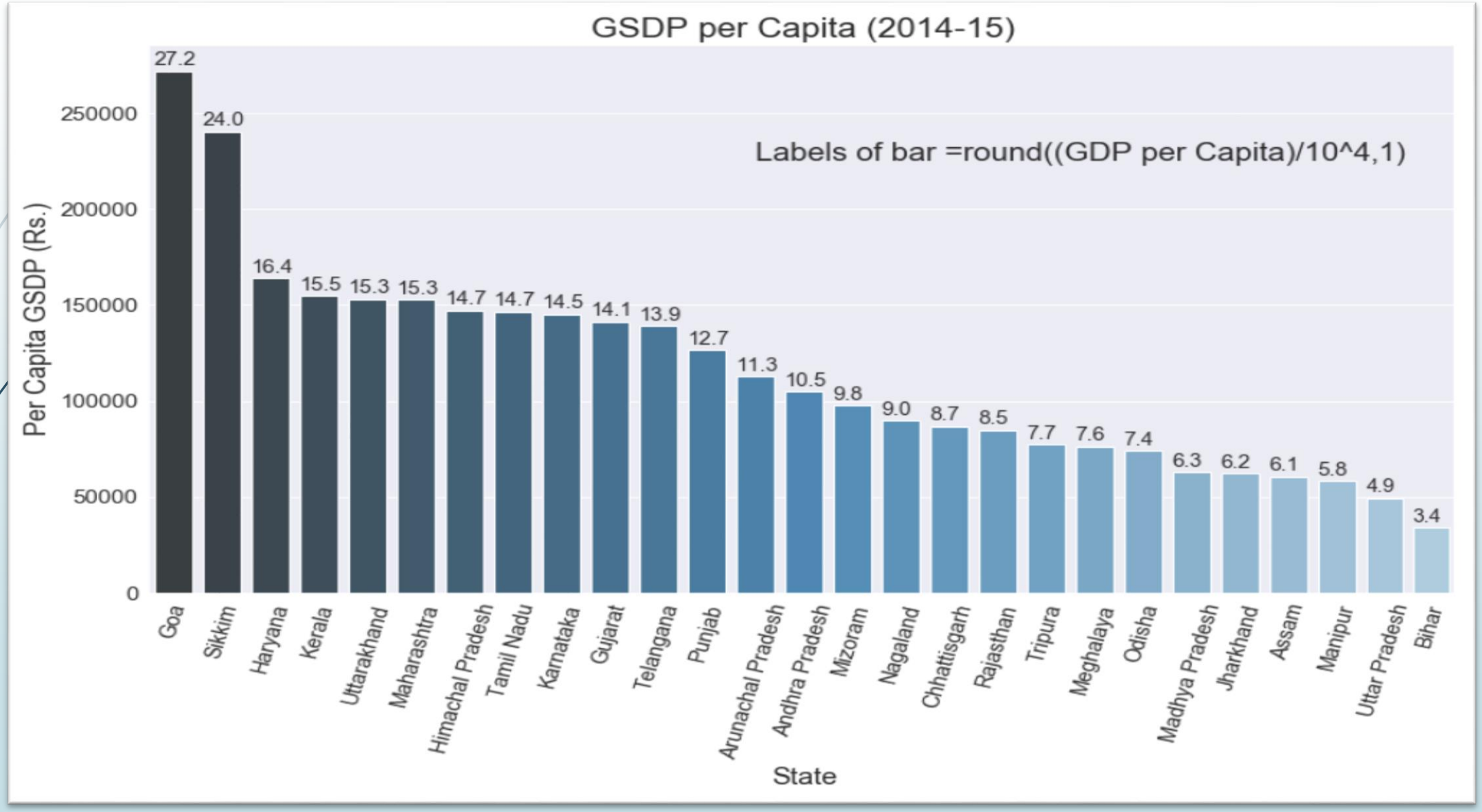
	State	Primary	Secondary	Tertiary	TOTAL GSVA at basic prices	Taxes on Products	Subsidies on products	Gross State Domestic Product	Population ('00)	Per Capita GSDP (Rs.)
0	Andhra Pradesh	16303716	10488884	22032942	48825542	5512100	1690800	52646842	501510	104977
1	Arunachal Pradesh	716959	287489	631844	1636292	70099	30272	1676119	14870	112718
2	Assam	5326697	4033091	9307109	18666897	1725309	582406	19809800	326780	60621
3	Bihar	8019997	5984896	22179969	36184863	3213546	2006421	37391988	1101240	33954
4	Chhattisgarh	6400817	8238886	7588778	22228481	2601791	1332092	23498180	270530	86860
5	Goa	312129	1547536	1738217	3597882	527279	61854	4063307	14950	271793
6	Gujarat	15887187	33023538	30220377	79131102	12353171	1981546	89502727	633590	141263
7	Haryana	8040424	12561411	19226568	39828404	4985670	1067867	43746207	266620	164077
8	Himachal Pradesh	1548366	4119162	4133326	9800853	895317	259291	10436879	70840	147330
9	Jharkhand	5248354	6241471	8133341	19623166	2499171	411619	21710718	349660	62091

INDEX	Sub-Sectors	Sector	State	GPV (Gross Product Value)
0	Crops	Primary	Andhra Pradesh	7893514
1	Livestock	Primary	Andhra Pradesh	4309078
2	Forestry and logging	Primary	Andhra Pradesh	346160
3	Fishing and aquaculture	Primary	Andhra Pradesh	2270664
4	Mining and quarrying	Primary	Andhra Pradesh	1484300
5	Manufacturing	Secondary	Andhra Pradesh	4672266
6	Electricity, gas, water supply & other utility services	Secondary	Andhra Pradesh	1151729
7	Construction	Secondary	Andhra Pradesh	4664889
8	Trade & repair services	Tertiary	Andhra Pradesh	3716000
9	Hotels & restaurants	Tertiary	Andhra Pradesh	517400
10	Railways	Tertiary	Andhra Pradesh	424228
11	Road transport	Tertiary	Andhra Pradesh	2816000

Approach Continuum :

- Data Preparation :
  - Removal and Addition of required columns.
  - Reorder and Renaming of columns name.
  - Merging of all data frames.
  - Sorting the data as per requirement.
  - Sorting out and set index of data frames.
  - Categorize data into four categories as per requirement based on given info.
  - Evaluation of any NaN values in formed data frames.
  - Searching for any disperancy in data (For eg. GDP value can’t be –ve, error while I/O of data) and resolve them either by correcting them or by removing them.

Plot the GDP per capita for all the states :



Insights from the Plot:

a. Identify the top-5 and the bottom-5 states based on GDP per capita (From Highest to lowest) ?

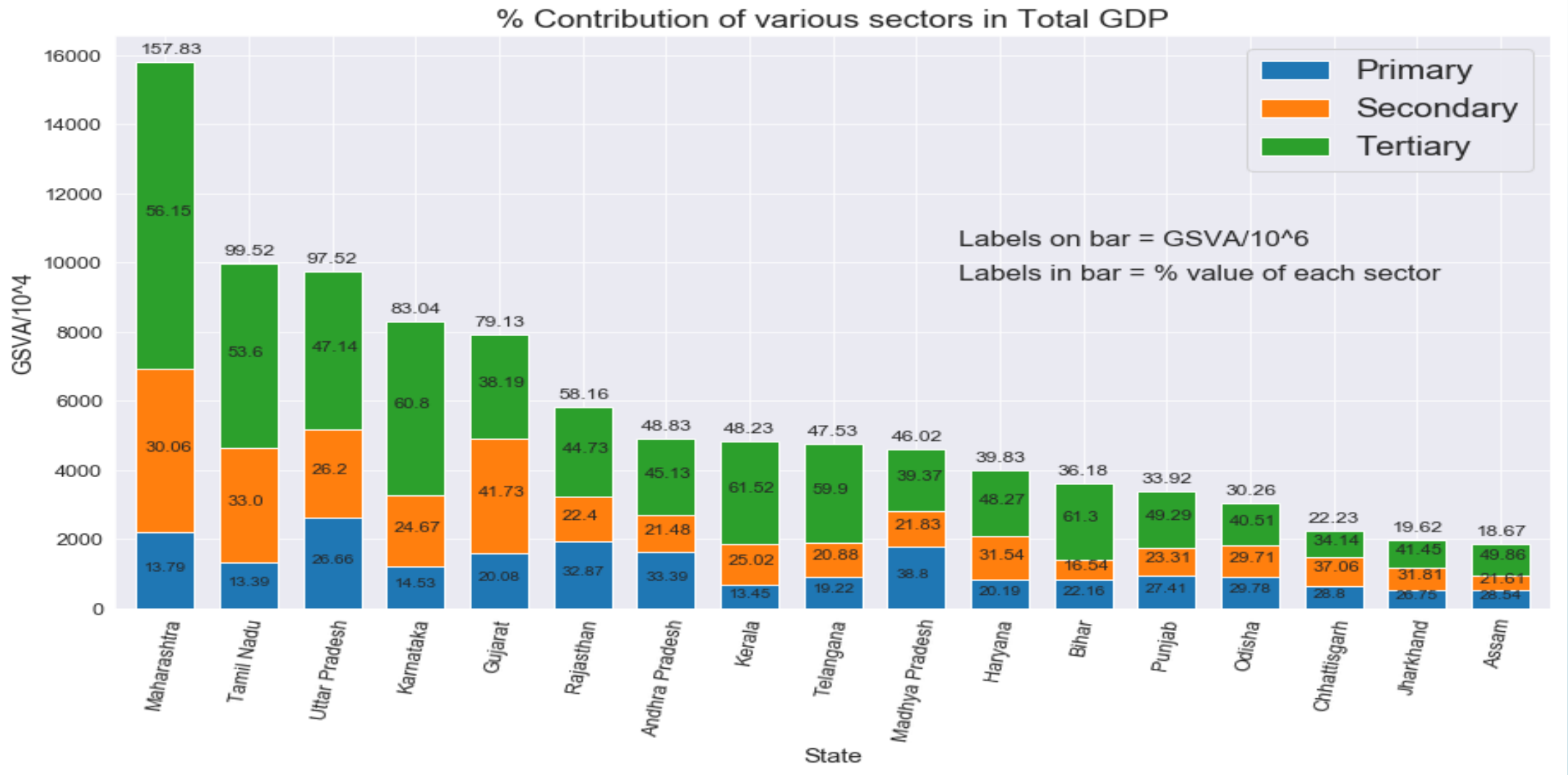
- The top - 5 :
  - ✓ Goa
  - ✓ Sikkim
  - ✓ Haryana
  - ✓ Kerala
  - ✓ Uttarakhand
- The Bottom - 5:
  - ✓ Jharkhand
  - ✓ Assam
  - ✓ Manipur
  - ✓ Uttar Pradesh
  - ✓ Bihar

b. The ratio of highest per capita GDP to the lowest per capita GDP?

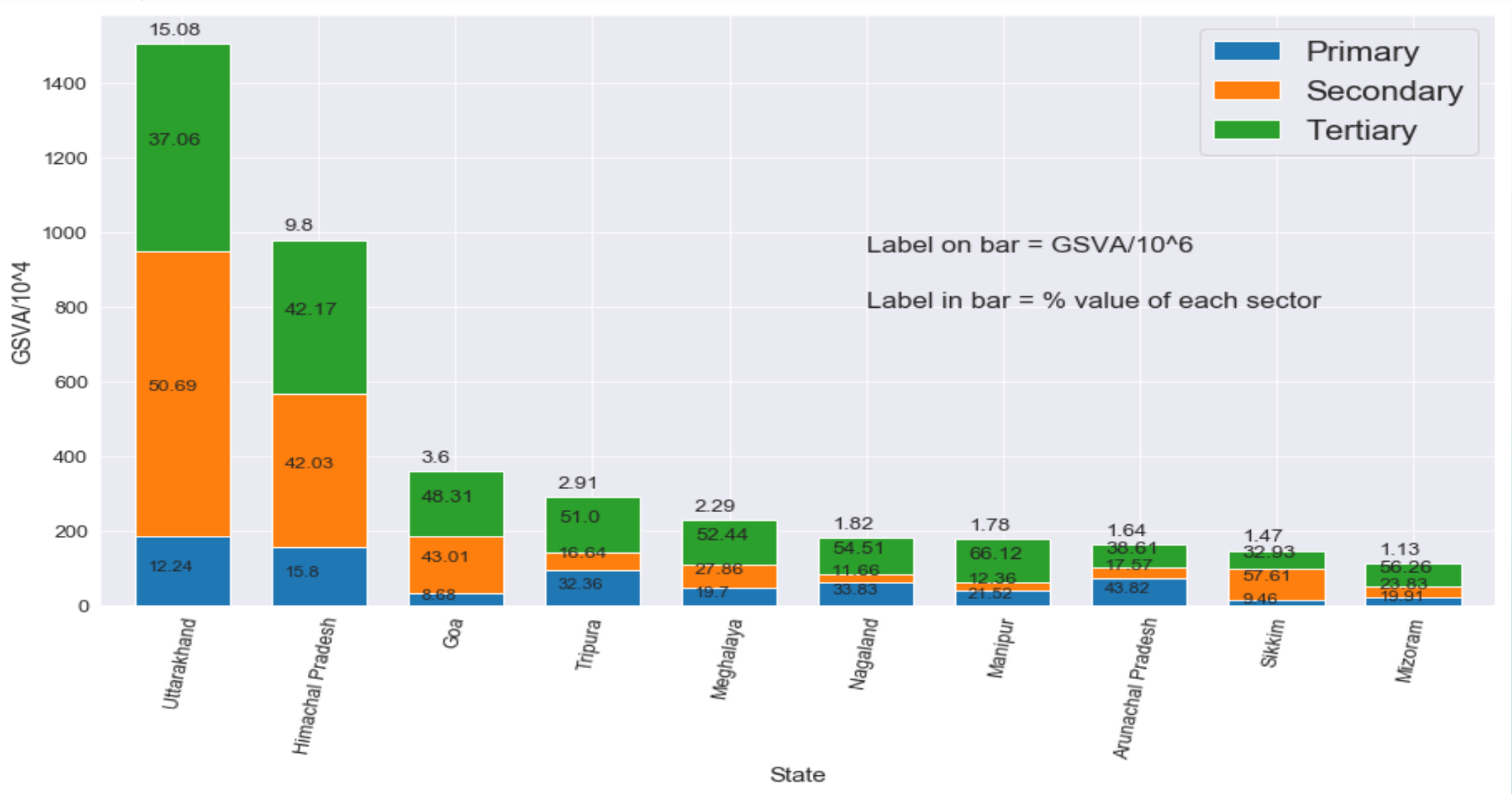
- A (Goa GDP per capita) = 271793
- B (Bihar GDP per capita) = 33954

➢ *A/B* = 8.004741709371503

Plot the % contribution of Primary, Secondary and Tertiary sectors as a % of total GSVA for all the states :



Note : As the difference between Total GSVA values between higher GSVA states and lower GSVA is very high, therefore the plot is divided into two parts.



Top 3/4/5 Sub Sectors which contribute to approx. 80% of the GSDP of category

For Category 1 :

Sub-Sectors	GPV	% of total GPV	cum
Real estate, ownership of dwelling & professional services	14740245	15.8281658	15.8281658
Manufacturing	13758793	14.7742766	30.60244241
Trade & repair services	12989747	13.94847028	44.55091268
Construction	11264451	12.09583681	56.6467495
Other services	8059922	8.65479385	65.30154335
Crops	7962514	8.55019654	73.85173988
Livestock	4619756	4.960722426	78.81246231

For Category 2:

Sub-Sectors	GPV	% of total GPV	cum
Manufacturing	1.14E+08	19.75827	19.75827
Real estate, ownership of dwelling & professional services	96526855	16.74913	36.50739
Trade & repair services	58968334	10.23206	46.73945
Crops	55443609	9.620454	56.35991
Construction	45318451	7.863559	64.22346
Other services	40469046	7.022101	71.24557
Financial services	38197505	6.627948	77.87351
Livestock	23672969	4.107682	81.9812

For Category 3:

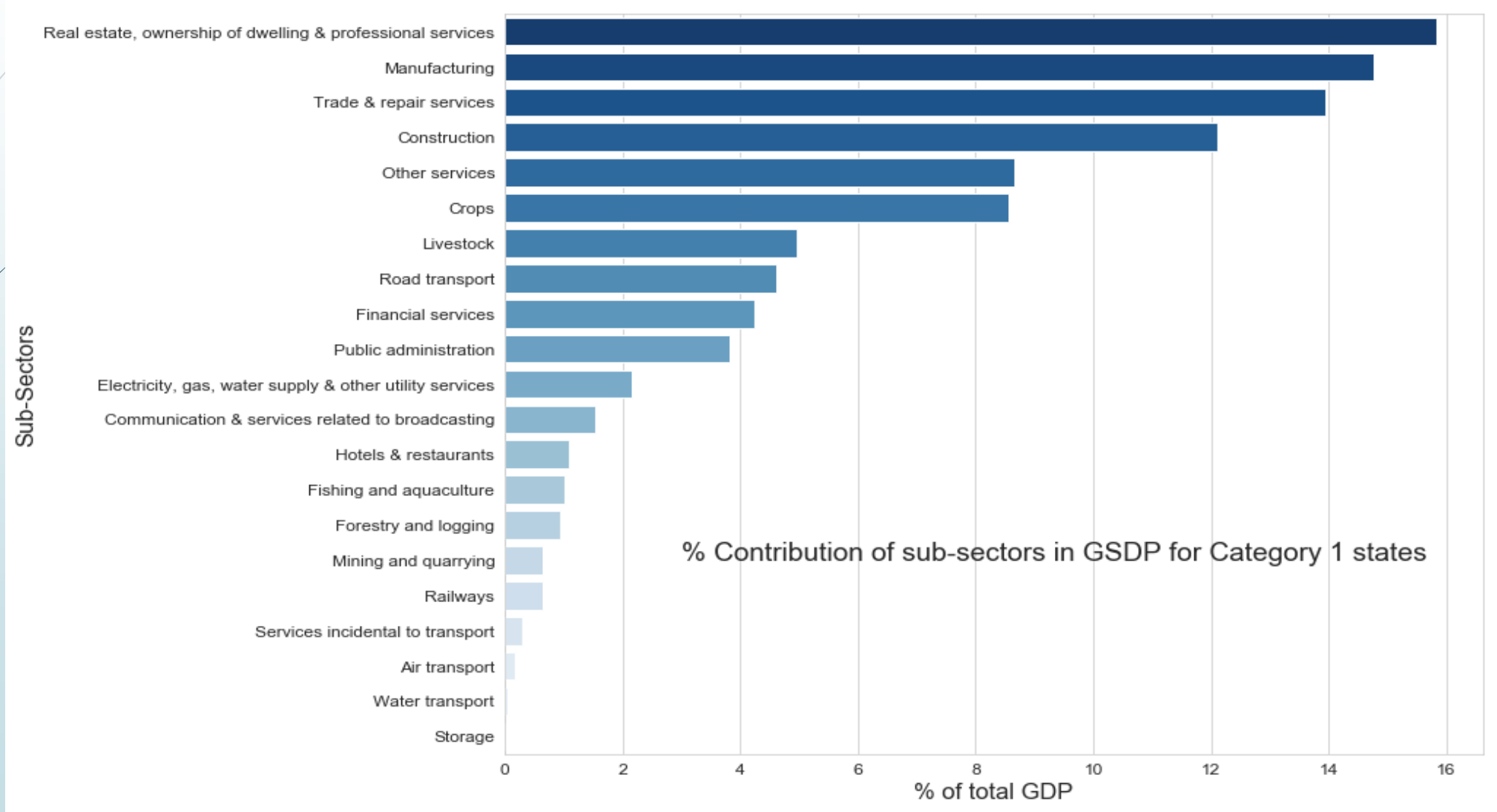
Sub-Sectors	GPV	% of total GPV	cum
Manufacturing	17366065	14.61816	14.61816
Crops	15782803	13.28543	27.90359
Trade & repair services	12237349	10.30098	38.20457
Real estate, ownership of dwelling & professional services	11818709	9.948587	48.15316
Construction	11043032	9.295648	57.44881
Mining and quarrying	9351471	7.87175	65.32056
Other services	8600028	7.239211	72.55977
Livestock	6826716	5.746497	78.30626

For Category 4 :

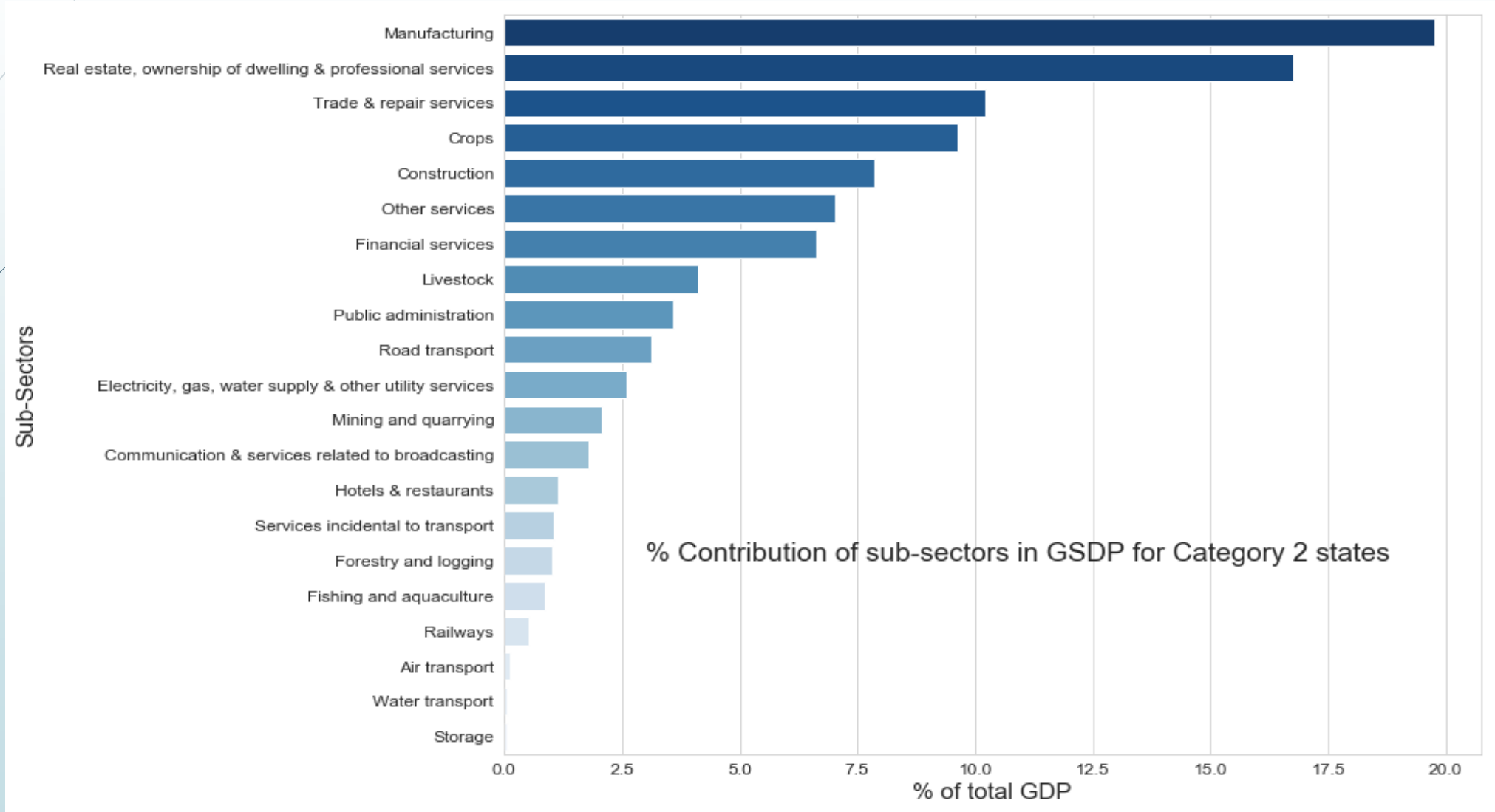
Sub-Sectors	GPV	% of total GPV	cum
Crops	39331055	17.89493	17.89493
Trade & repair services	25581239	11.63901	29.53394
Manufacturing	24987032	11.36866	40.9026
Real estate, ownership of dwelling & professional services	24177534	11.00035	51.90295
Construction	22775948	10.36265	62.2656
Other services	15859015	7.215571	69.48117
Public administration	13486630	6.136178	75.61735
Livestock	11901405	5.414928	81.03228

Plot the contribution of the sub-sectors as a percentage of the GSDP of each category

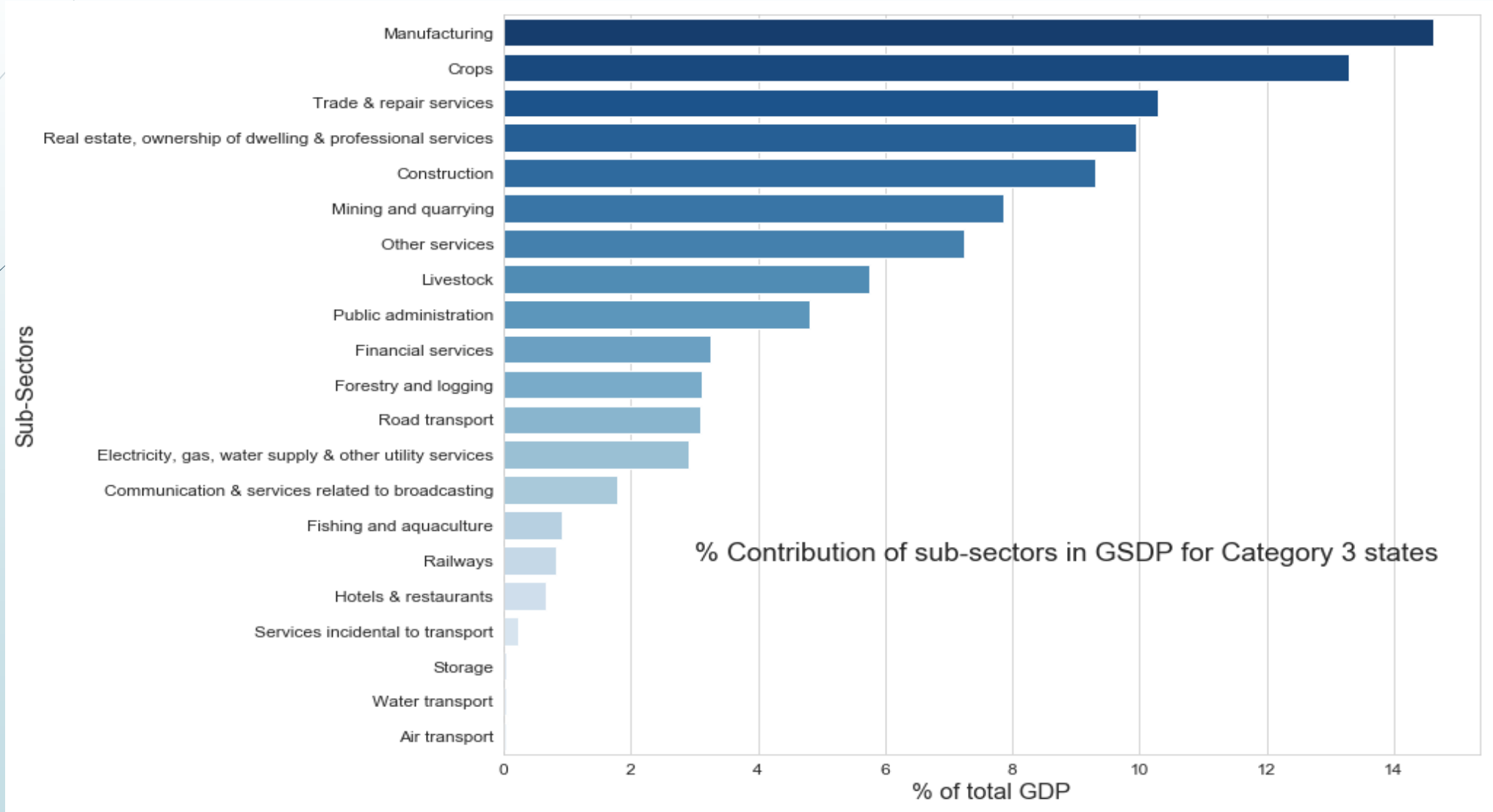
For Category 1 :



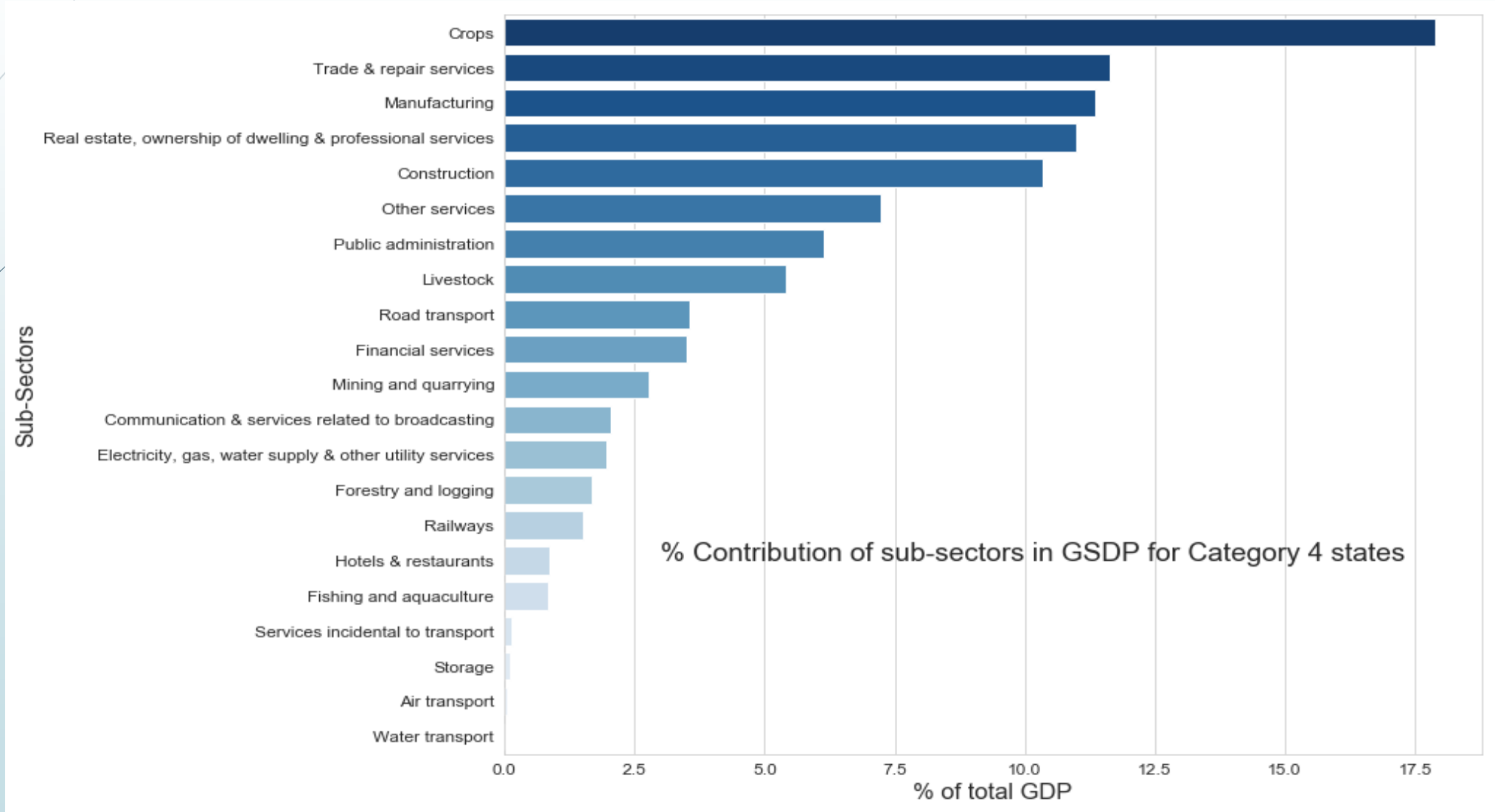
For Category 2 :



For Category 3:



For Category 4:



Insights From plot :

- In all four categories, sub-sectors :
  - Storage
  - Water Transport
  - Air Transport
  - Services incidental to transport

are doing really bad and need of continuously focus.

- In all categories sub-sectors:
  - Real estate, ownership of dwelling & professional services
  - Crops
  - Manufacturing
  - Trade And Repair Services
  - Construction

are doing really well.

Insights Continuum :

- From Category 1 to Category 4, It is clear that % contribution of ***CROPS*** is increasing and in category 4 it stands for most contribution to GDP.
- Where as contribution for tertiary sub-sectors like Estate, ownership, Trade and services are decline as move from category 1 to category 2.
- Among the transport services, Road Transport has the maximum % contribution.
- In case of Secondary sectors, Mining and quarrying contribute much less than the other sub-sectors of secondary sector in all categories.
- Forestry and logging is much popular in category 3 than other categories.

Part-II: GDP and Education Drop-out Rates:

Brief:

In this part of the analysis, we will investigate whether there is any relationship between per capita GDP with drop-out rates in education.

Take a look on data received :

Sl. No.	Level of Education - State	Primary - 2014-2015	Upper Primary - 2014-2015	Secondary - 2014-2015	Senior Secondary - 2014-2015
1	A & N Islands	1.21	1.69	9.87	16.93
2	Andhra Pradesh	4.35	5.2	15.71	NA
3	Arunachal Pradesh	10.89	6.71	17.11	18.42
4	Assam	7.44	10.51	27.06	NA
5	Bihar	2.09	4.08	25.9	NA
6	Chandigarh	NA	0.44	NA	10.55
7	Chhatisgarh	1.42	5.85	21.26	2.76
8	Dadra & Nagar Haveli	1.57	4.02	16.77	9.47

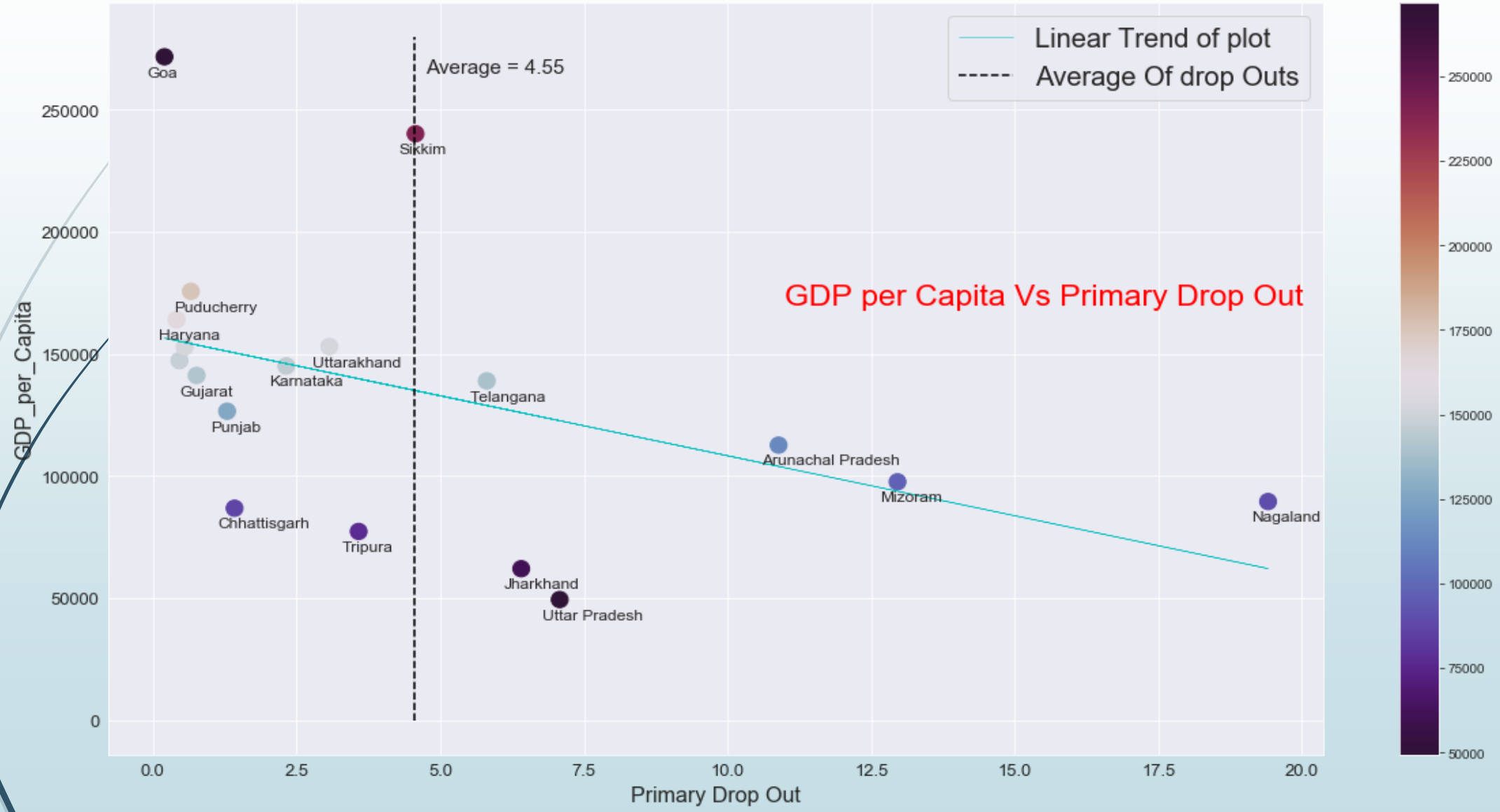
Approach :

- Read CSV file into python.
- Filter Out the data as analyze is done for year 2014-15.
- Renaming of the column names for better understanding.
- Evaluation of data frame for any disperancy in data or any NaN values.
- Correction of disperancy of data and remove of data having NaN values.
- Resetting of index of data frame.
- Add column to existing data frame consisting of GDP per capita from previous made data frames.
- Checking of data missing(GDP per capita value for Puducherry) and correct them.

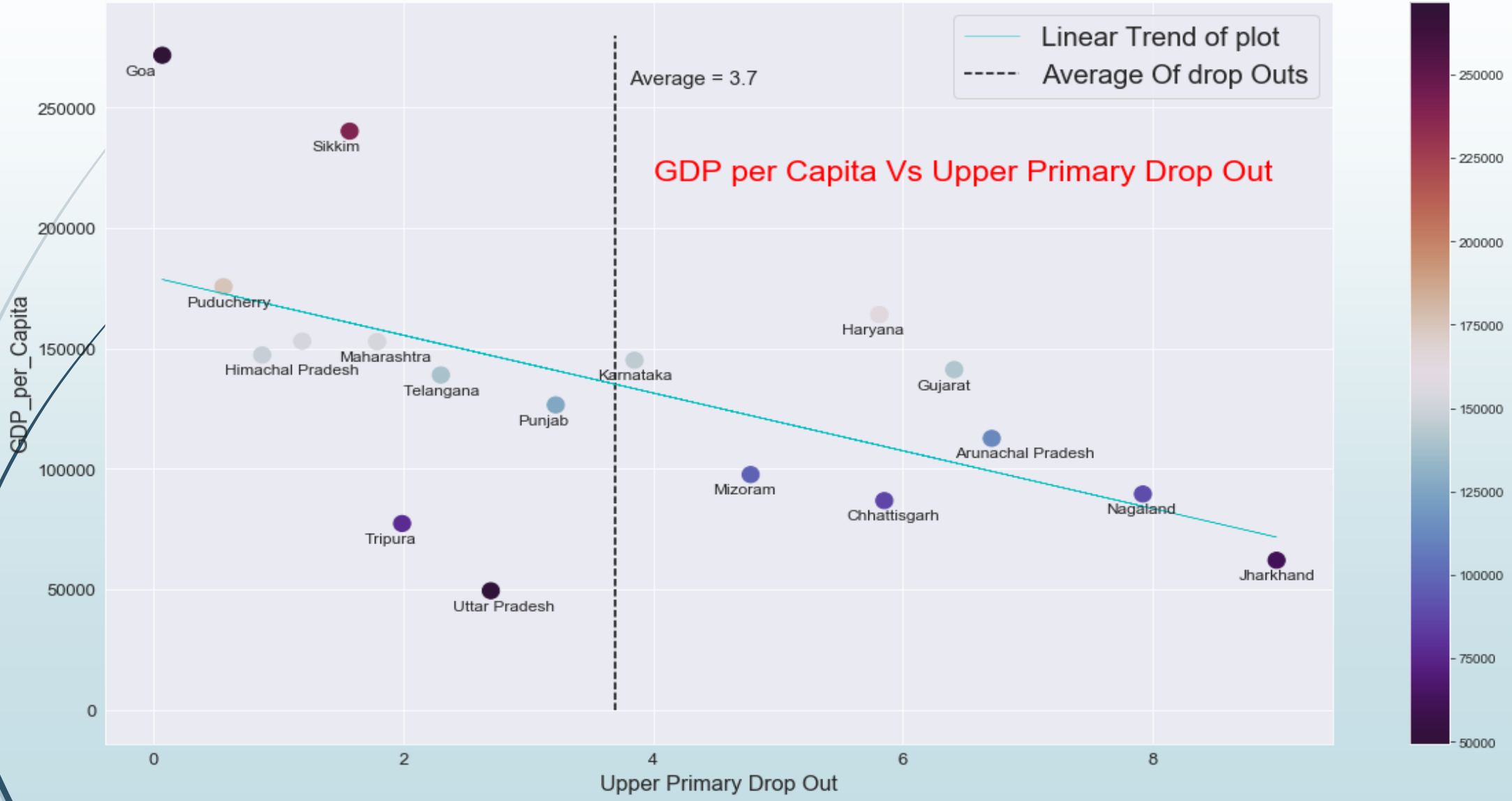
Look on the formed table:

State	Primary	Upper Primary	Secondary	Senior Secondary	GDP_per_Capita
Arunachal Pradesh	10.89	6.71	17.11	18.42	112718
Chhattisgarh	1.42	5.85	21.26	2.76	86860
Goa	0.2	0.07	11.15	13.91	271793
Gujarat	0.76	6.41	25.04	7.04	141263
Haryana	0.41	5.81	15.89	5.75	164077
Himachal Pradesh	0.46	0.87	6.07	7.41	147330
Jharkhand	6.41	8.99	24	3.41	62091
Karnataka	2.32	3.85	26.18	1.96	145141
Maharashtra	0.55	1.79	12.87	1.83	152853

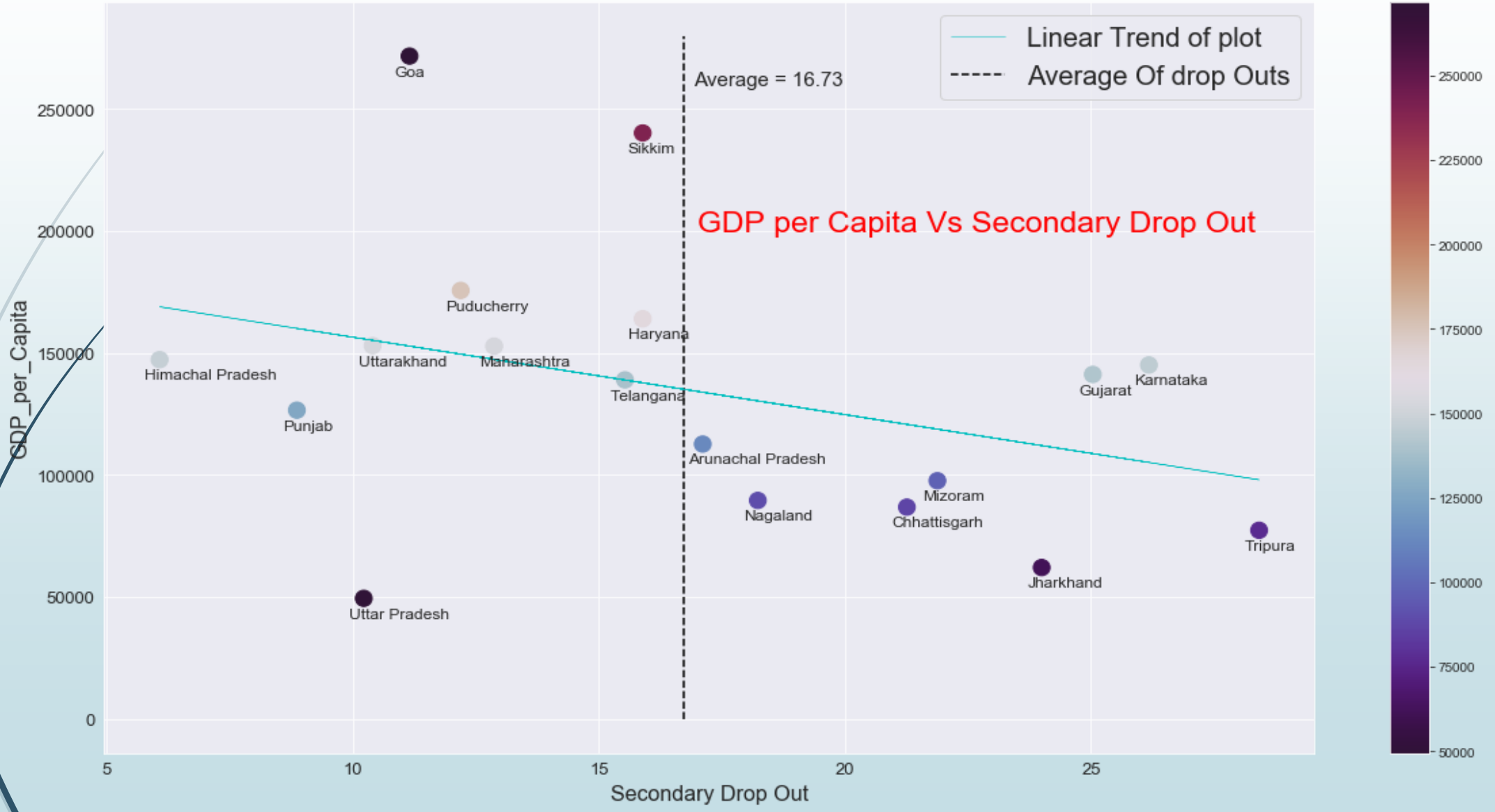
GDP per Capita vs Primary Drop Out :



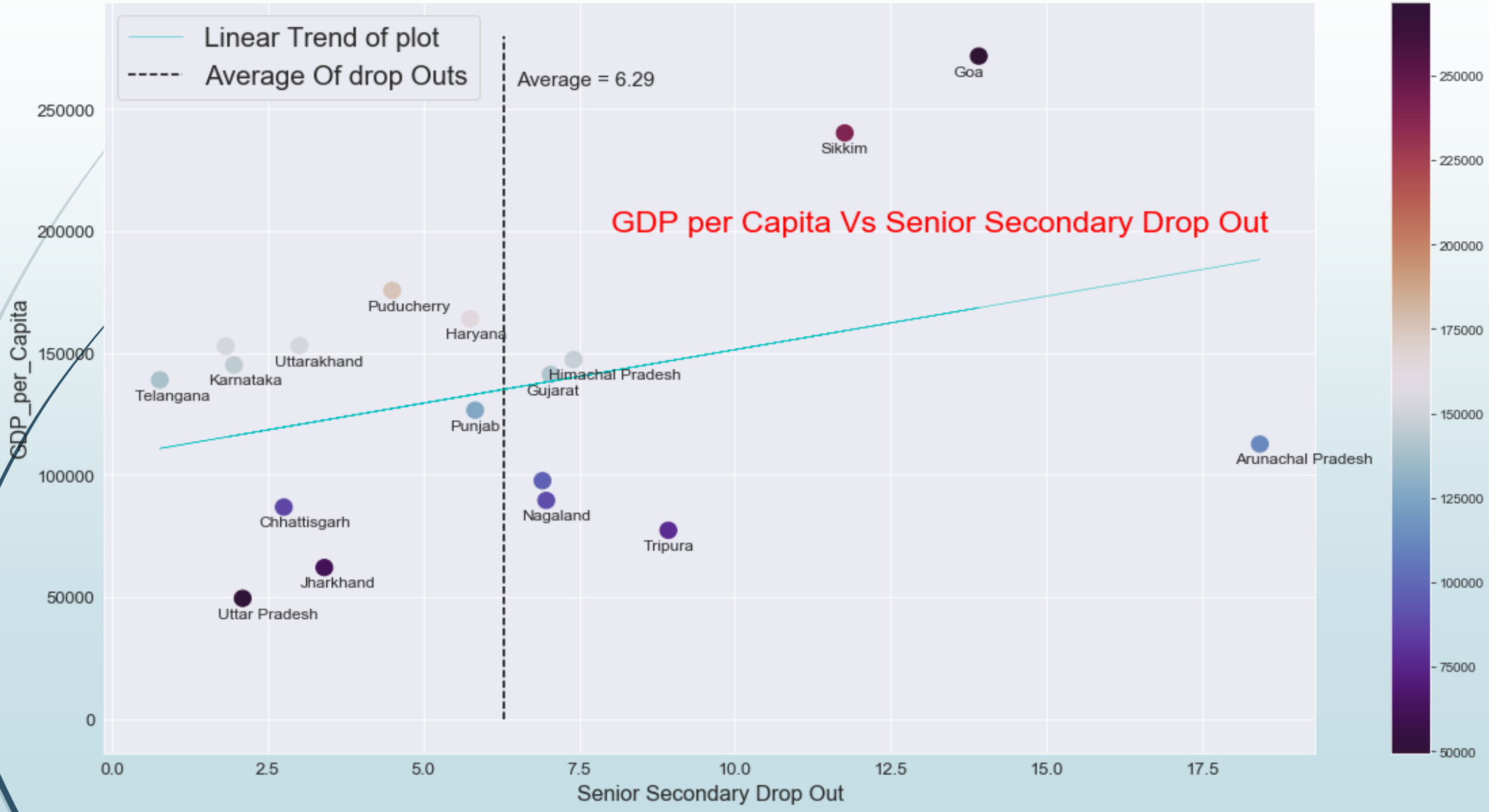
GDP per Capita vs Upper Primary Drop Out :



GDP per Capita vs Secondary Drop Out :



GDP per Capita vs Senior Secondary Drop Out :



Insights From Plot:

- GDP per Capita is decreases with increasing Drop out but in case of Senior Secondary education, GDP per Capita of is increasing with drop out.
- Decrease in GDP per Capita is decreasing as we move on from lower education to higher education drop outs.

Hypothesis :

- A student who drop out in there early stages of education (For eg. Primary, Upper Primary) had not learn or achieved the position of self development (They need someone to nurture them and gets the best from them) consequently not able to successful in there ambition and contribute to there state GDP.
- whereas in case drop outs at higher level of education (Senior Secondary), Students are be in the position of self capable and the probably the universities don't capable of doing same or provide the culture, quantity of knowledge students wants.

Thank you….