Usha Bhanu Komaragiri, Vishwa Shrirame, Ajinkya Pahinkar, Ayush Sekhsaria
Professor: Yong-Yeol Ahn
INFO - I 590 Data Visualization.

India's Unemployment Crisis

Final Report

Abstract

In this fast-paced, advancing world where innovation and execution are limitless, one could say there is a vast scope of skilled/unskilled employment in the present and the future. The reality is that every country suffers from the problem of unemployment as there are fewer jobs than the number of applicants in every sector. In our data visualization project, we first focus on how this unemployment crisis has affected India and what factors are responsible for this crisis. Through data, we would like to address India's unemployment rate over the years and relate it with the scenarios leading to the same. We would consider factors like literacy, GDP per capita of the state, external factors like COVID-19, etc that have contributed to this increase and also discuss what possible measures can be taken to decrease this alarming rate in the future. Using data-driven analysis and visualizations, we aim to derive meaning from these datasets and address this ongoing problem.

Introduction

The problem of unemployment has aggravated in recent times and the economic slowdown is taking up many jobs with the majority of the leading companies having declared a hiring freeze. Many capable youths are not able to put their skills to the best use while there are tons of hourly wage people who are losing their jobs. This is happening all over the world, and it is only visible when conditions become severe. As we mentioned above, unemployment is a serious problem and so after looking at India's state-wide data, we aim to bring to notice how serious the problem is and correlate it with the factors which caused them to happen in the first place.

Unemployment is one of the major socio-economic problems in India. The number of unemployed divided by the total number of people in the civilian labor force is the unemployment rate. Not everyone who is jobless is unemployed. To be included in the unemployment rate, a person must not only be jobless but also have actively sought employment over the previous four weeks. A person is still included even if they were just temporarily fired and are waiting to be called back to their position. In the unemployment rate, people are not included if they have given up seeking work.

Structural unemployment: Unemployment occurs when an educated individual is unable to find employment. Despite having the qualifications, this unemployment is caused by the absence of work prospects in the Indian economy. In the Indian economy, structural

unemployment is also known as educated unemployment. Educated unemployment is a growing issue in the country. Particularly during the pandemic, the country's educated unemployment has undergone a paradigm shift.

Motivation

The consequences of unemployment are an intriguing reason to study this. One of the most crucial measures of a nation's health is its ability to offer individuals who want jobs, a place to work. As a result, the unemployment rate which reflects the inability to create such employment becomes the most crucial sign of how well a modern economy is doing.

The well-being of society increases with the proportion of the labor force that is productively engaged in a population. A society where the majority of people are employed, offers relief from vulnerability and poverty. Households are also motivated to spend more to enhance their quality of life. Households promote economic expansion and more employment in this way.

On the other side, a rise in unemployment weakens overall consumer spending, slows the economy, and most significantly, makes people more vulnerable to changes in the economy. In recent times, we have come across several LinkedIn posts where people are either seeking employment or laid off by their current employer. It's unfortunate we are living in this dire times where thousands of people are struggling to seek employment. It makes us wonder what could be the reason behind this. The year 2020 was a period when the world experienced severe disruption in daily activities with the rise of the COVID-19 Pandemic. Thousands of people lost their jobs and a lot of families were affected by it. Due to the pandemic, the unemployment rate rose to 20.9% during the April-June quarter of 2020 which was more than double the unemployment rate in the same quarter the previous year. [8] Another aspect we want to research is whether literacy has any impact on the unemployment rate in India. Are people losing out on opportunities based upon the number of degrees they have or are there any lack of facilities for education provide by the government?

Using the data and with the help of visualizations we can identify which regions are most affected by unemployment, thus such indicators must be available so that correct actions can be taken promptly. We want to understand what factors are responsible for the rise and decline of unemployment rate's in Indian States. Our main motivation is to enlighten these factors and represent our results using various visualization techniques.

Existing visualizations

We will look into some existing visualizations that have contributed to the unemployment crisis analyzes and state our suggestions and improvements on them.
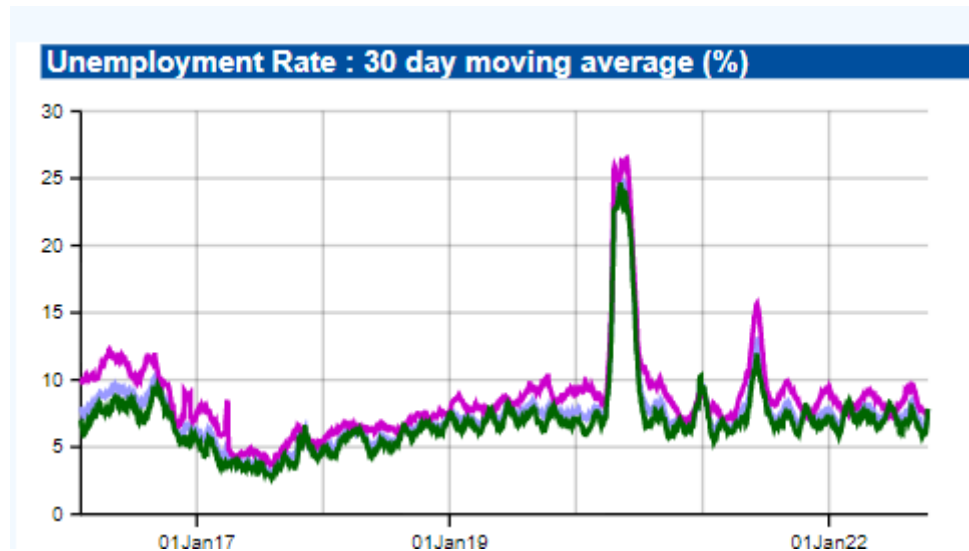

Fig 1: 30-day moving average of unemployment rate in India

The daily and monthly unemployment rates in India, in rural India, in urban India, as well as for the Indian states, are provided by CMIE. [5] Every day, a 30-day moving average of the unemployment rate in India is calculated by the CMIE using the data gathered over the previous 30 days. According to the May-August 2022 report from the Center for Monitoring Indian Economy (CMIE), the nation's unemployment rate rose to 8.3% in August of that year. This rate is the highest it has been in the previous year. More than 20 lakh individuals lost their jobs in August 2022. In Figure 1 we can see a huge spike around the year 2020 in the unemployment rate which tells us the effect of COVID-19 on industries. An observation could be that the X-axis labels are not defined in a detailed manner as we had to estimate the spike to be around the year 2020. There's also use of 3 different colours in this distribution but it would have been more insightful with the presence of a legend. While this is a good indicator of depicting the unemployment rate of the whole country over a 30 day time period, we have emphasized more the State-wise statistics and spread across India. In our project, we aim to identify the regions where unemployment has impacted the most and figuring out the why behind such metrics.
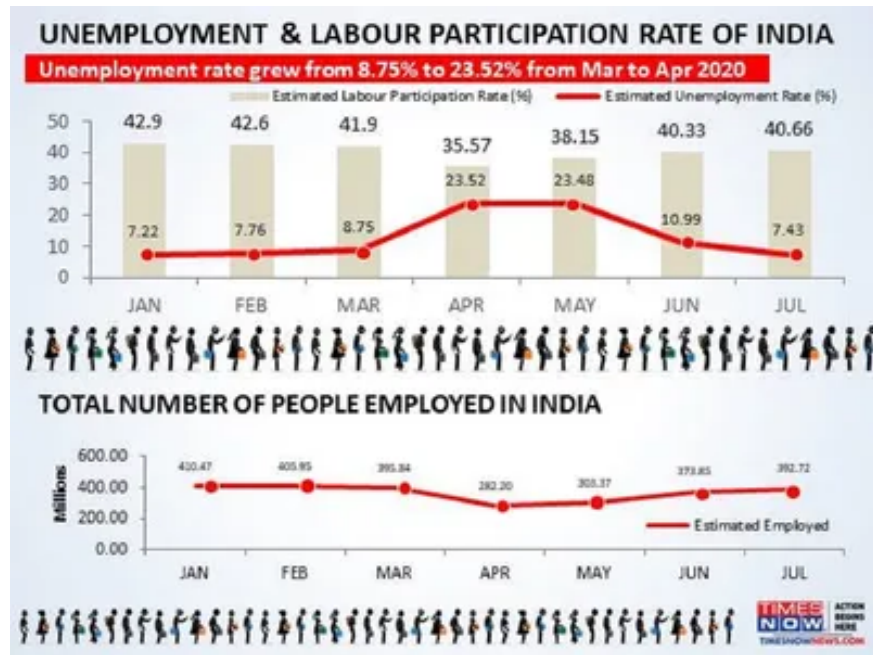
Fig 2: Unemployment and Labour participation rate in 2020

In figure 2 we see a visualization posted by Times Now that depicts how the Covid-19 outbreak has affected the unemployment rate in India. [7]. The rate grew from 8.75% to 23.52% from March to April 2020 which is very shocking. So many people lost their jobs in that period which led to this increase but the rate soon came back to normal in 3 months which explains the lockdown period. However, certain things seem unnecessary in this visualization. From the first graph, we understand that unemployment has gone through a huge change but the below graph is just an opposite way of expressing the first graph. And the images in between also seem unnecessary which depicts a cluttered visualization as a whole. We have used both the unemployment and labour participation rate data across all states in our implementation to depict the necessary information. Analysis on certain state's statistic and their respective reasons are been provided below. We aim to derive conclusions from a general trend among the States of India.

Objectives & Contribution

The contribution of our work will focus on the factors listed below that we think are responsible for the unemployment rates across India. We have listed down our main objectives in this project that we have answered with various visualization techniques.

1. How has covid-19 affected the unemployment rate in India?
2. Which states do we observe a high/low unemployment and employment rate? What could be the factors?
3. Is education an important factor that affects the unemployment of the state?
4. Does the GDP per Capita affect the unemployment rate of that particular state?
5. How's labor participation spread in India? Is it more in rural areas or urban areas?

We wanted to make better use of the data to represent visualizations effectively. Different factors needed to be considered to analyze the unemployment rates across the country. Our aim was to address these factors and understand if there's any relation with the unemployment rates. Initially, we started with GDP-per capita data of all the States from 1996 to 2020. The idea was to get the top performing states from this period and compare with the mean unemployment rate for the year 2019. It's important that we considered only 2019 data for unemployment rates as 2020 which was the year of the pandemic would be an inaccurate assumption. Later on, we wanted to understand the growth of unemployment rates and decline of labour participation rates due to COVID-19 and we had some interesting observations which we have explained later in the report. In order to understand the impact of education on unemployment rates, we considered literacy rate of each State and compared with the unemployment rate of 2019. We got results on both ends i.e, for a higher literacy rate, we got higher and lower unemployment rates and vice versa. Thus, it would be very difficult to come with a conclusion here and thus we counted that as a failed experiment and as well as something to work towards in the future.

The idea behind our contributions was to improve the existing visualizations by using interesting/intuitive visualizations and make better use of the dataset. Another aspect was to research about various factors that could affect the unemployment rates of any nation. We chose to start our analysis from India.

Dataset & core ideas

We have identified multiple datasets that can help us in our implementation. We will focus on state-wise statistics of unemployment, education, and GDP, and thus some data cleaning would be required to join the data.

Unemployment statistics -

	Region	Date	Frequency	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	Area
0	Andhra Pradesh	31-05-2019	Monthly	3.65	11999139.0	43.24	Rural
1	Andhra Pradesh	30-06-2019	Monthly	3.05	11755881.0	42.05	Rural
2	Andhra Pradesh	31-07-2019	Monthly	3.75	12086707.0	43.50	Rural
3	Andhra Pradesh	31-08-2019	Monthly	3.32	12285693.0	43.97	Rural
4	Andhra Pradesh	30-09-2019	Monthly	5.17	12256762.0	44.68	Rural

Table 1: Unemployment statistics of 2019 and 2020

This dataset would help us with the unemployment rate and labour participation rates across the Indian states. We have only used 2019 and 2020 data as to analyze the impact of COVID-19 on the unemployment rates.

Region-wise distribution of unemployment statistics -

	State	Region_in_India	Estimated Unemployment Rate (%)
0	Andhra Pradesh	South	7.477143
1	Assam	Northeast	6.428077
2	Bihar	East	18.918214
3	Chandigarh	North	15.991667
4	Chhattisgarh	West	9.240357
5	Delhi	North	16.495357
6	Goa	West	9.274167
7	Gujarat	West	6.663929
8	Haryana	North	26.283214
9	Himachal Pradesh	North	18.540357
10	Jammu & Kashmir	North	16.188571
11	Jharkhand	East	20.585000
12	Karnataka	South	6.676071
13	Kerala	South	10.123929
14	Madhya Pradesh	West	7.406429
15	Maharashtra	West	7.557500
16	Meghalaya	Northeast	4.798889
17	Odisha	East	5.657857

Table 2: Region-wise unemployment rate statistics

This data would help us to analyze the spread of unemployment rates across various regions of India.

GDP statistics -

State Name	Bihar	Gujarat	Himachal Pradesh	Karnataka	Madhya Pradesh	Maharashtra	Tamil Nadu	Telangana	Uttar Pradesh	West Bengal
1996-97	4001	16153	11960	11202	8819	17844	11670	13269	7476	9857
2000-01	6415	18392	22795	17195	11862	22777	18344	20972	9828	16583
2005-06	8223	37780	36949	28987	16631	41965	31239	35243	14221	24720
2010-11	19111	77485	68297	66951	32453	84858	62251	78473	26698	47245
2011-12	21750	87481	87721	90269	38551	99564	91121	92984	32002	51543
2012-13	24487	102826	99730	102314	44931	111980	101007	104943	35812	58195
2013-14	26948	113139	114095	118829	52169	125035	112162	116236	40124	65932
2014-15	28671	127017	123299	130024	56069	132476	124104	128372	42267	68876
2015-16	30404	139254	135512	148108	62626	146258	140840	140441	47118	75592
2016-17	34156	156295	150290	170133	74787	162005	159395	154272	52744	82291
2017-18	38631	173079	167044	187649	82941	176102	180494	171583	58821	93711
2018-19	43822	197447	179188	210887	90998	191736	204488	193750	66512	109491
2019-20	46292	213936	183407	223175	98418	202130	213396	233325	65431	121463

Table 3: GDP-per capita statistics of all States from 1996 - 2020

We have made use of GDP per capita data from 1996 to 2020. This would help us in analyzing the growth of GDP of each state and compare the data with the unemployment statistics to understand its impact. The main idea here is to get the States who have performed consistently better than the other States. We found 5 states namely, Gujarat, Karnataka, Maharashtra, Tamil Nadu and Telangana would contributes to the majority of nation's GDP per capita and compared their unemployment rates to the mean unemployment rate of the country.

Education statistics -

statname	tot_population	urban_population	sexratio	literacy_rate	male_literacy_rate	female_literacy_rate	schools
Punjab	27704.24	29.82	893.0	76.68	81.48	71.34	9171
Uttarakhand	10116.75	21.54	963.0	79.63	88.33	70.70	3512
Haryana	25353.08	24.12	877.0	76.64	85.38	66.77	7663
Rajasthan	68621.01	19.26	926.0	67.06	80.51	52.66	28195
Uttar Pradesh	199581.48	17.31	908.0	69.72	79.24	59.26	24569
Bihar	103804.64	8.36	916.0	63.82	73.39	53.33	7627
Meghalaya	2964.01	15.32	986.0	75.48	77.17	73.78	1555
Assam	31169.27	11.03	954.0	73.18	78.81	67.27	9062
West Bengal	91347.74	24.55	947.0	77.08	82.67	71.16	10194
Jharkhand	32966.24	18.18	947.0	67.63	78.45	56.21	4601
Odisha	41947.36	13.15	978.0	73.45	82.40	64.36	10745

Table 4: Education statistics of all States

This is the education data we have collected to analyze the literacy rate across the Indian states. We will again compare it with the unemployment statistics to understand if education/literacy is a factor for unemployment.

We have also made use of shape files to plot geospatial data.

	State_Name	geometry
0	Andaman & Nicobar	MULTIPOLYGON (((10341718.474 1449533.161, 1034...
1	Chandigarh	POLYGON ((8546255.616 3606050.813, 8546315.400...
2	Daman and Diu and Dadra and Nagar Haveli	MULTIPOLYGON (((8122247.822 2312434.407, 81223...
3	Delhi	POLYGON ((8583390.570 3359116.190, 8583476.212...
4	Haryana	POLYGON ((8524318.539 3516490.865, 8524451.392...

The geometry column in the shape file is a vector that would help to represent the boundary and exact location of a particular State of India. We have used this extensively to project the spread of various metrics across India.

The Fig 4 represents a sample prototype of working with shape files. It represents the GDP per capita spread across Indian states. The intensity of the shade represents high/low GDP per capita in that State.
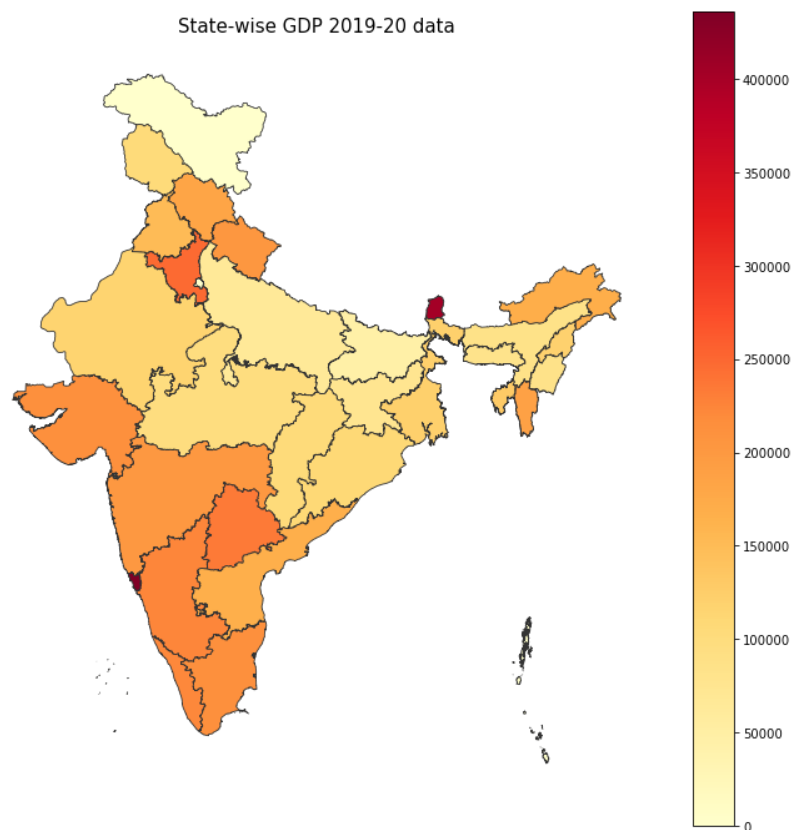
Fig 3: State-wise GDP per capita for 2019-20

Visualization Methods

In our visualizations, we aim to visualize graphs that will represent the distribution of the factors like GDP per capita, Unemployment rate, Labour Participation rate, Literacy rate etc. and compare them state-wise to understand the impact of every State towards unemployment statistics.

A straightforward method that makes it possible to see how a variable varies throughout a geographic area where states are observed for a high/low unemployment and employment rates. Choropleth maps are excellent for highlighting distinct geographic trends in the data. Choropleth maps are of most use with the way it uses color to show how our data changes from state to state. A similar plot can be made for State-wise labour participation rate to observe the geographical trends in the data. A location like a state or country, as well as the number of submissions from that location, are required for putting up a choropleth map for use.

To compare data among the States of India, we went ahead with bar charts. Bar charts can be applied when presenting categories of data. For time series data, they are also particularly effective. When you want to see how a variable changed over time or compare variables to one another, such as unemployment growth or a decline in labor participation, a bar chart is the best option. We chose the horizontal bar plot to see how the labour participation and unemployment rate have changed in different states over time. The use of a horizontal bar plot has three key benefits. Long category labels are simpler to show and read. The display of many categories is simpler. It is simpler to display labels for multiple bars without a collision.

Our data's traditional bar chart uses bars to represent a single categorical variable. Two categorical variables are shown in a stacked bar chart. The first variable is displayed throughout the full bar, and stacks for the second variable are shown inside each categorization bar. In our data we intend to check what is the literacy rate out of hundred for different states in India. The bars are divided into colored bar segments and piled on top of one another in a 100% stacked bar chart. The colored bar segments show the components' varying percentage contributions to the entire bar, with each bar's height being 100%.

For data visualization and simple and straightforward data interpretation, the Python Plotly package can be quite helpful. Plotly graph objects are a high-level, user-friendly interface for plotly. The scatter geo function in Plotly is used to map geographic information into the maps. A slider for this plot is to observe the monthly statistics for unemployment rates. Longitude/latitude pairs in "lon" and "lat," respectively, or geographic place IDs or names in "locations," are the two ways that data that is represented as scatter points or lines on a geographic map are provided.

Methodology

We used matplotlib as the primary visualization tool in our implementation and certain other interesting tools such as plotly, bar_chart_race for a couple of visualizations.

For observing the geographical pattern for the unemployment rates and GDP-per capita over all states, the choropleth map visualizations are plotted. Similar plot was plotted to observe Labor participation rates across the States. In terms of data pre-processing, we merged the GDP/unemployment dataset with the shape files dataset on State-name and plotted the merged dataset to get the choropleth map visualization.

Based on the summary statistics of the dataset, we manually gave the index limits for the choropleth maps. This was important to get accurate visualizations. For bar charts we have chosen Set 2 cmap as we had many categorical variables (States) to work with. In regards to the bar chart race we had to perform a couple of data wrangling steps to transform the dataset that is accepted by the bar_chart_race library. Every row represents the period of time and the colum holds the value for that categorical variable. The index contains the time component.[12] These were the constraints that were accepted by the bar_chart_race libary. Stacked bar charts was important to represent literacy rate in % for every state. We used that data and compared with the unemployment rates to understand their relationship.

We observed many null values in the dataset and the GDP per capita data for most of the years from 2001 to 2010 were missing.

Since our analysis covered all the 28 states, the bar charts were not that intuitive as we wanted it to be, but we tried to provide the necessary information about the insights in this report. There were 400 columns in the education dataset and thus we just chose the important ones that were necessary for our problem. The columns we used were the total population, urban population, sex ratio, literacy rate, male literacy rate, female literacy rate and the number of schools for our analysis. In terms of data cleaning, conversion of string formats to integers, grouping by and aggregating on unemployment and labour participation rates of State variables, joining multiple dataframes and conversion of string to datetime constituted our implementation.

Failed Experiments

Initially, we wanted to understand whether education is a factor in the rising unemployment in India. We chose Literacy rate as a metric to analyze and compare among all the States. A strong question arises that - Are people unemployed because they don't have the necessary degrees/diplomas or are there any lack of education facilities provided by the government that leads to illiteracy and becomes a factor for unemployment?
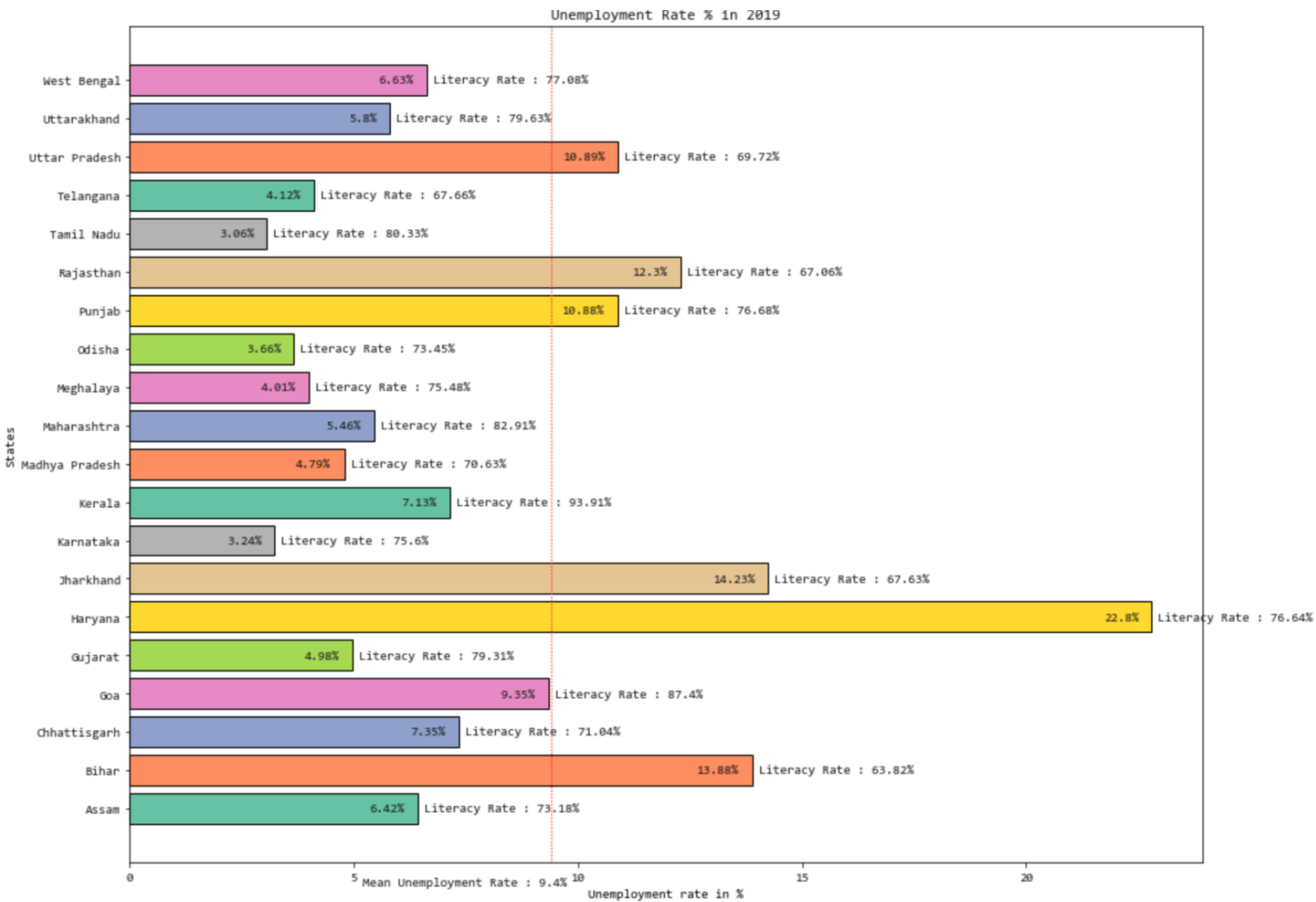
Fig 4: Literacy rate analysis with Unemployment rates in year 2019

We visualized the unemployment rates in 2019 for all the states and added the literacy rate statistic against it. We also added a red dotted line that indicates the mean unemployment rate of the whole country for that year. There was a mix-bag of observations after doing so.

The states like Uttar Pradesh, Rajasthan, Jharkhand and Bihar have a low literacy rate and high unemployment rates. We can infer that literacy among people may be a factor for high unemployment rates. However, states like Punjab and Haryana who have a high literacy rate also have high unemployment rates. We can also see that for states like Madhya Pradesh and Telangana, who have a low literacy rate and a low unemployment rate.

It becomes difficult to infer from such insights as there are observations for both sides and thus we count this experiment in failed experiments. We can say that we didn't get enough evidence to prove that literacy rate had any impact on the unemployment rates in India. Although, it was interesting to learn from this comparison. As a future work we plan to consider other educational factors to understand their impact towards unemployment in India.

Results and Discussion

Impact of GDP per capita on Unemployment

We started with visualizing the GDP per capita growth of states from 1996 to 2020. This would help us identify the top Indian states that share the majority of Indian economy. We made use of an interesting tool called bar chart race which animates the use of bar charts. This gives us the option to visualize time series bar charts and monitor the growth of Indian States over time.
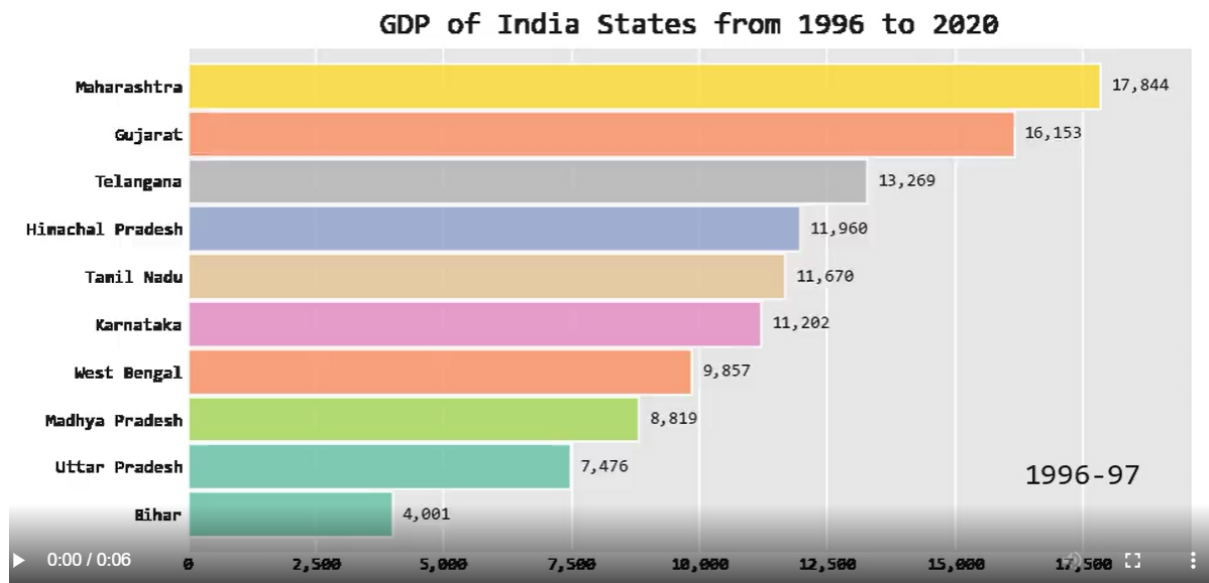

Fig 5: State-wise GDP per capita (in Rs.) for 1996-97
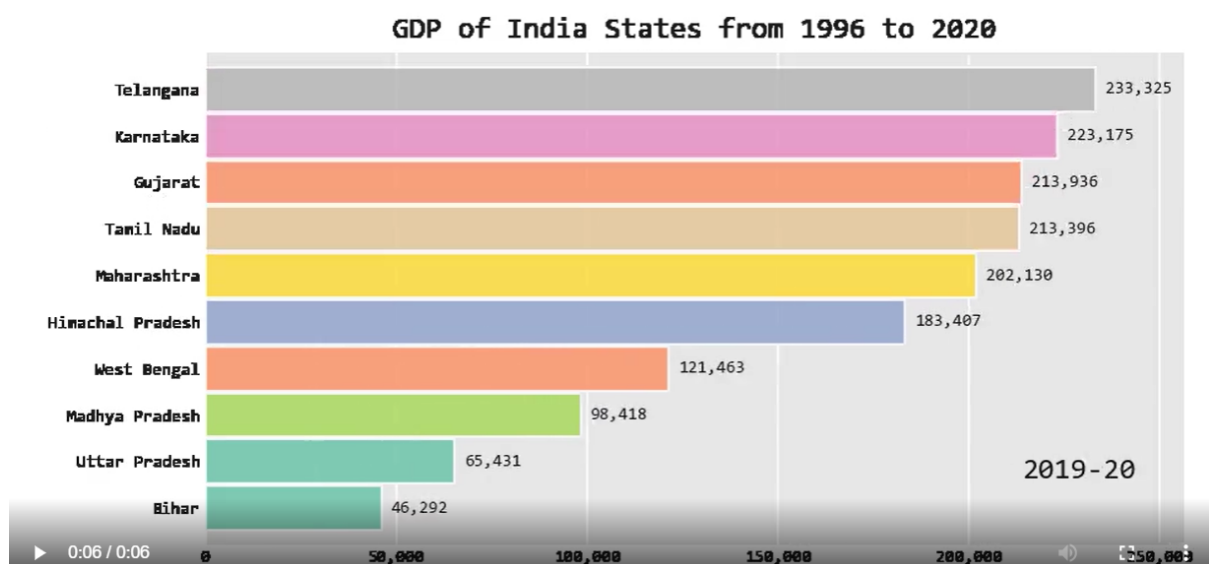
Fig 6: State-wise GDP per capita (in Rs.) for 2019-20

We have identified states Maharashtra, Gujarat, Karnataka, Tamil Nadu and Telangana as the top states that have been consistently performing better as compared to the other states.

Now, we'll use the unemployment statistics of these states and compare with the overall unemployment rate of India.
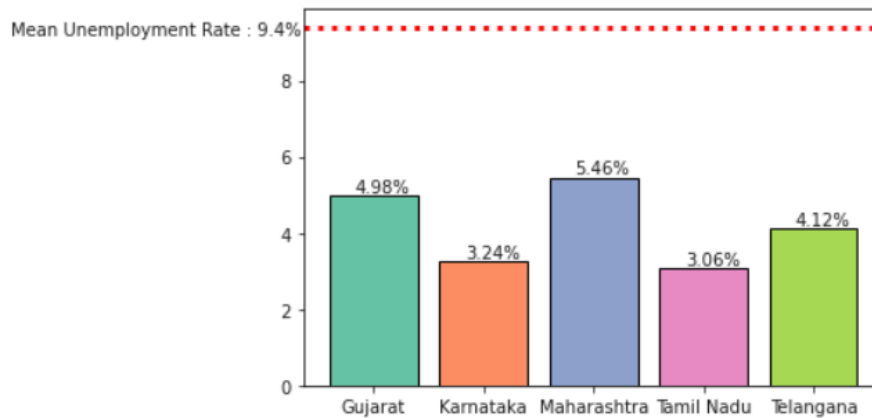
Fig 7: Comparison of unemployment rates with the mean unemployment rate

The mean unemployment rate of the whole country for the year 2019 was 9.43% (marked by the red dotted line). We observe that the unemployment rate for the top performing states are well below the the mean rate. We have only considered the year 2019 because it was pre-covid period and thus we wanted to see the impact of only one external factor i.e GDP with respect to the unemployment rate. Based on our analysis we will infer that unemployment rate is less on an average for top performing states in terms of GDP per capita.

Impact of COVID-19 on Unemployment

Now, we'll look into the impact of COVID-19 on the unemployment rates in India. COVID-19 started in March 2020, so we compared the years 2019 and 2020 to understand its impact. We started with comparing the choropleth maps as it would give us the whole perspective of Indian States. The intensity of the shade tells us how much the area is affected by unemployment due to COVID-19. The index at the right acts as a legend which denotes the unemployment rate in %.
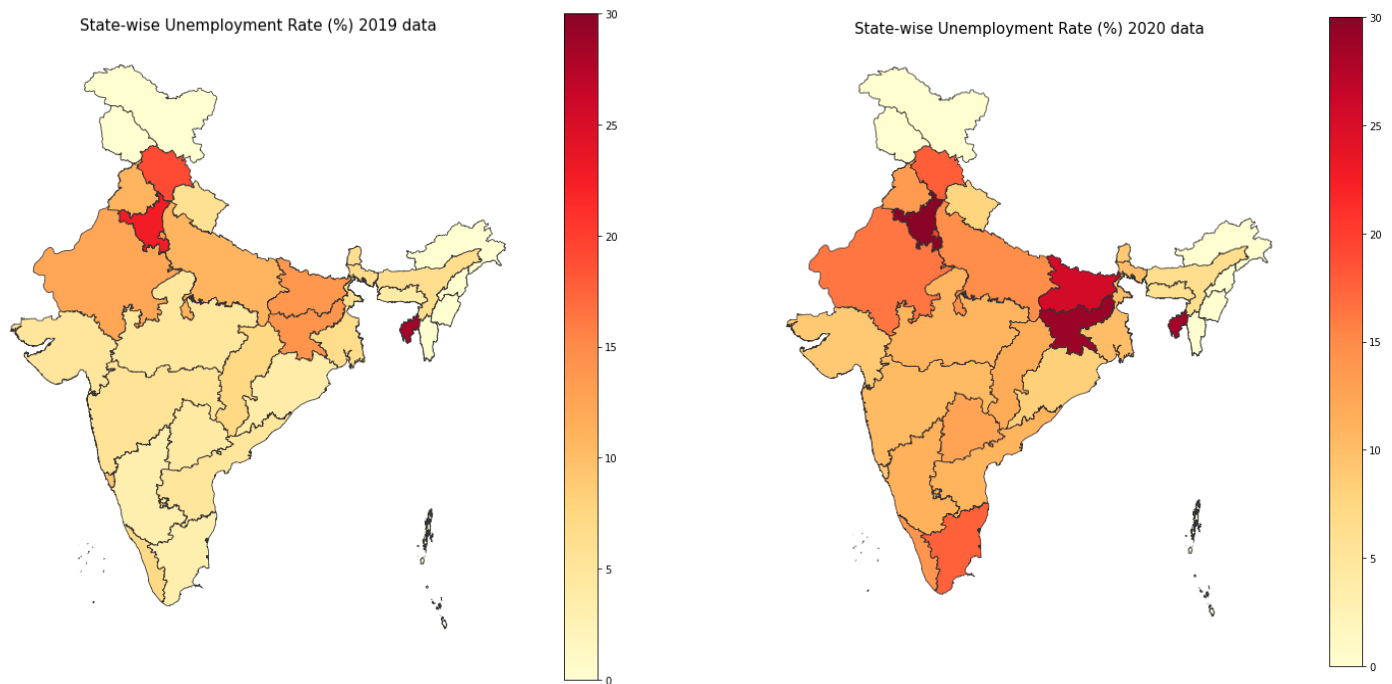
We have plotted two choropleth maps indicating the spread of unemployment rates across the country. We observe that most of the states have a darker shade in the 2020 choropleth map which indicates that COVID-19 was a huge factor in the unemployment increase in the country. The parts that were already high in terms of unemployment rates in 2019 has a even more darker shade in 2020. Comparison among choropleth maps gives a good overview of the rise of unemployment rates in 2020. However, we can understand better if we have certain metrics indicating how much unemployment rates grew from 2019 to 2020. In order to understand those metrics, we made use of bar charts to visualize the growth in % of unemployment rates across all states.
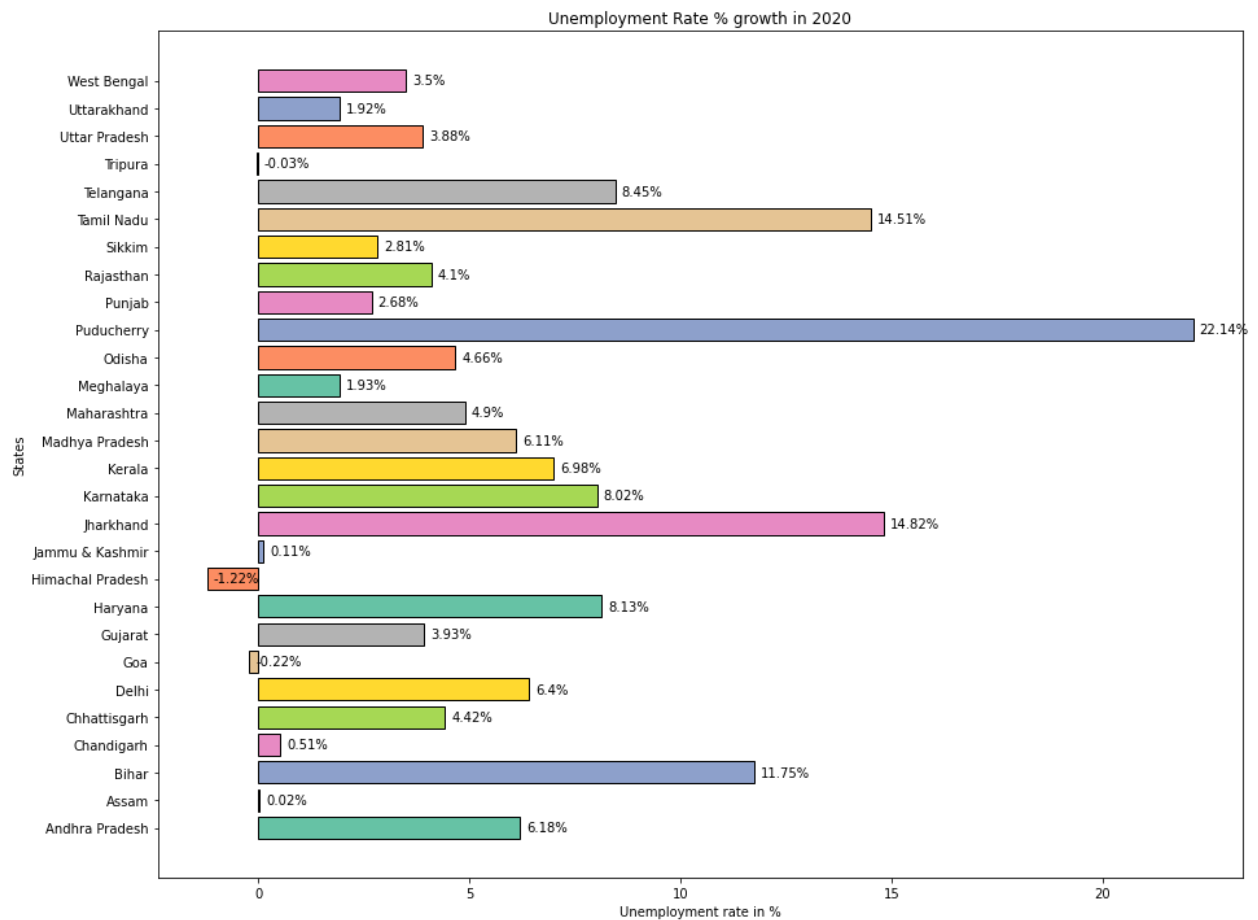
Fig 9: Growth in Unemployment Rates (%) in year 2020 across all States

In this visualization, we have represented the growth of unemployment numerically across all States. Apart from states like Tripura, Jammu & Kashmir, Himachal Pradesh, Assam, Goa and Chandigarh, we see a significant rise in the unemployment rates across India. As industrialization is not widespread across the country, states such as Jammu & Kashmir, Himachal Pradesh, Assam that lie towards the north experience less impact due to the COVID-19 pandemic. States such as Telangana, Tamil Nadu, Maharashtra, Haryana and Jharkhand that houses majority of the industries have had a significant impact on unemployment rates due to the pandemic. There was a nation-wide lockdown imposed during March 2020 to prevent the spread of COVID-19 which resulted in loss of jobs. Many industries had to cut costs by numerous lay offs which impacted the households of millions of people.

On a further note, we look into the unemployment rates of certain states in the year 2019 and we found that the states having very less impact due to COVID-19 such as Tripura (28.36%), Jammu & Kashmir (16.14%), Himachal Pradesh (19.06%), Assam (6.42%), Goa (9.35%) and Chandigarh (15.82%) etc. already had a large unemployment rate in 2019. Due to the presence of less industries we deduced that the unemployment rates within these states stayed the same for 2020.

The unemployment rate in Puducherry was 1.7% in 2019 and as per Fig 9, this state was impacted the most due to the pandemic. It recorded a rise of 22.14% in it's unemployment rate. On further research, we found that Puducherry's vital sectors was not given priority by the government which explains the meteoric rise. The mean unemployment rate across India in the year 2020 was 15.1% which was the most since decades.

Impact of COVID-19 on Labour participation

Labour participation in India is widespread across the country and millions of people are dependent on it for their livelihood. Labour participation comprises of people who supply labour for the production of goods and services in various sectors [9]. We wanted to check how much labour participation rate was affected by the COVID-19 pandemic.

Firstly, we made a comparison between the years 2019 and 2020 on choropleth maps to understand the spread of labour participation across India and how much it was impacted in 2020 by the pandemic.
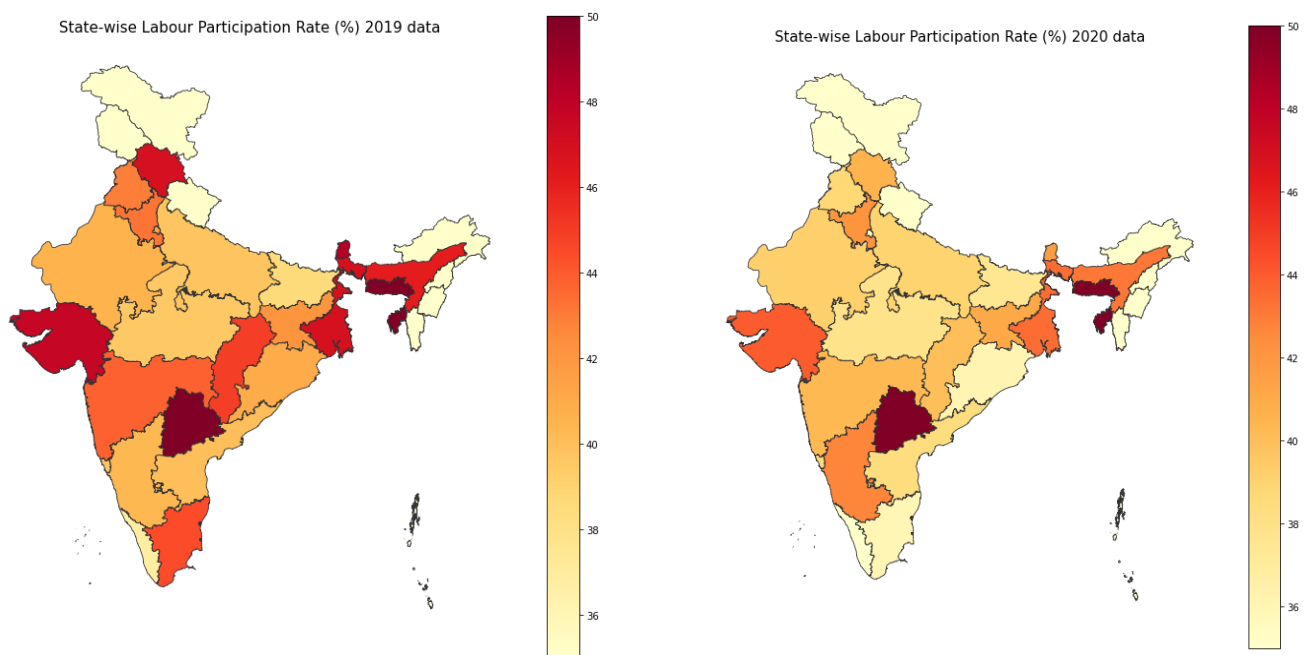

Fig 10: Comparison of Labour Participation between years 2019 and 2020

In the year 2019, labour participation rate was widespread across the country as seen by the darker shades in majority of the states. However, most of the states have a lighter shade in the year 2020 due to the COVID-19 pandemic. Since there was a nationwide lockdown imposed in the country, labour activities had to come to a stop for a brief period. A significant decline in the labour participation rates adds to the unemployment rates impacted by the pandemic. We will again use bar charts to understand the decline of labour participation numerically across all the states.
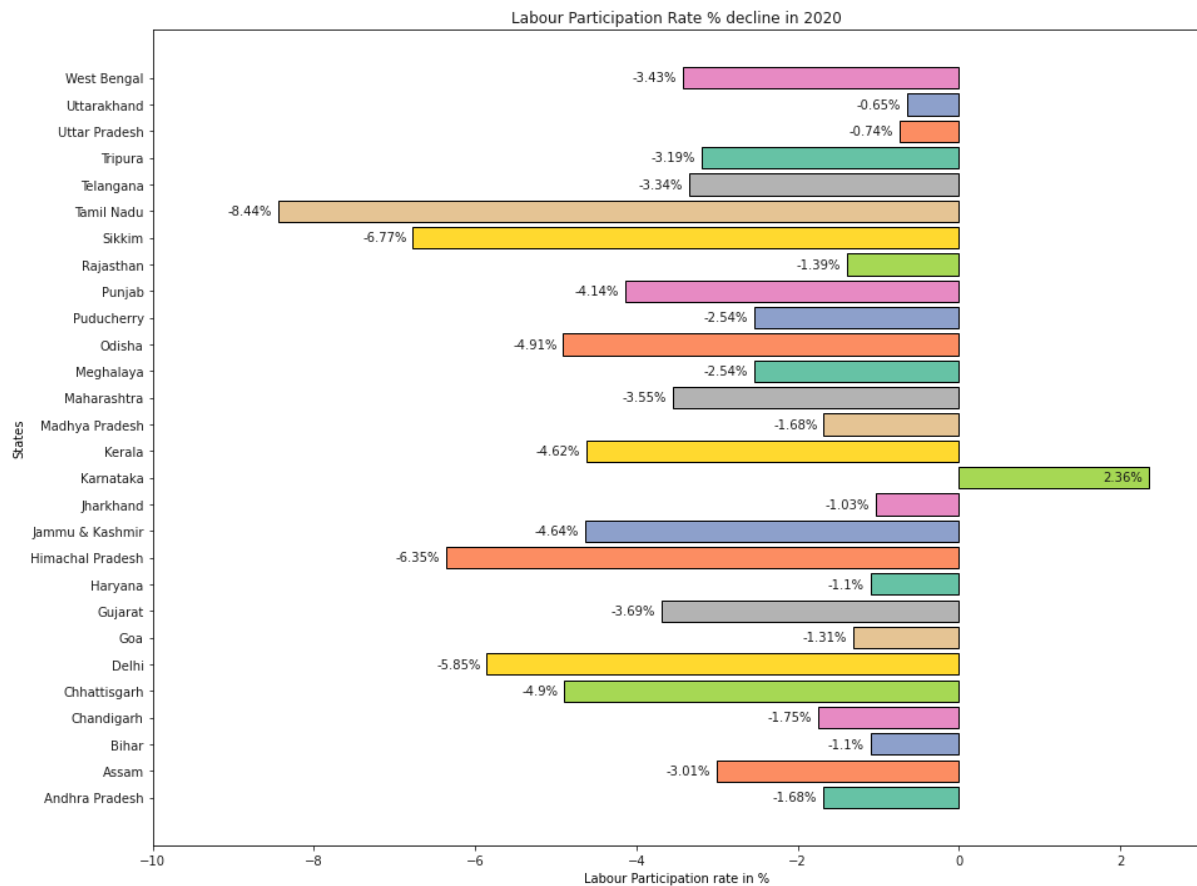
Fig 11: Decline in Labour Participation Rates (%) in year 2020 across all States

In the above bar chart, we have represented labour participation rate for every State in India and there is a common trend among all the states. It's certain that due to the nation-wide lockdown because of the COVID-19 virus, governments had to take harsh measures by stopping all the outdoor activities. It has been a major impact on people's lives who's livelihood depended on labour work. India accounts for 501 million people [10] who are working as labourers and such an impact due to the pandemic has been atrocious. However, there's an outlier in the visualization as the labour participation rate in Karnataka increased during the pandemic. On further research, we found that Urban Karnataka has seen a rise in women's participation in labour force which increased from 22.1% to 27.2%. This could be the reason in the rise in labour participation rate. [11] However, this increase did not align with the unemployment figures.

Impact of COVID-19 on Urban and Rural Areas

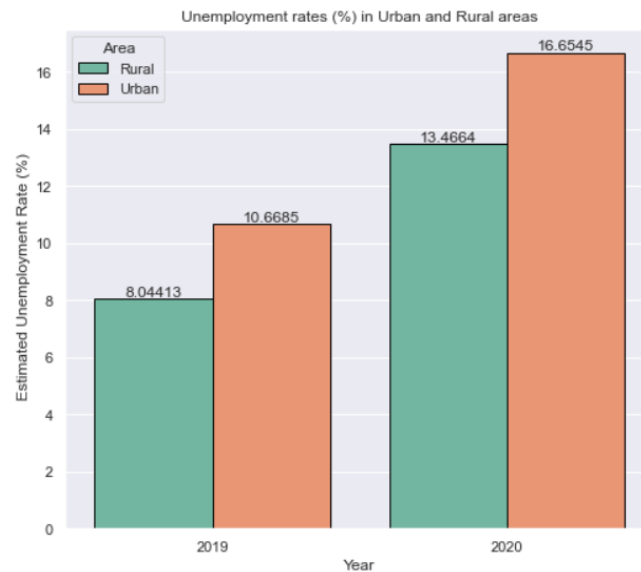We did a minor analysis to understand the impact of COVID-19 on Urban and Rural areas.

Fig 12: Unemployment rates in Urban and Rural areas

We observed that Urban areas were impacted more in unemployment rate due to the COVID-19 pandemic than Rural areas. This is expected as most of the employment opportunities are present in Urban areas.
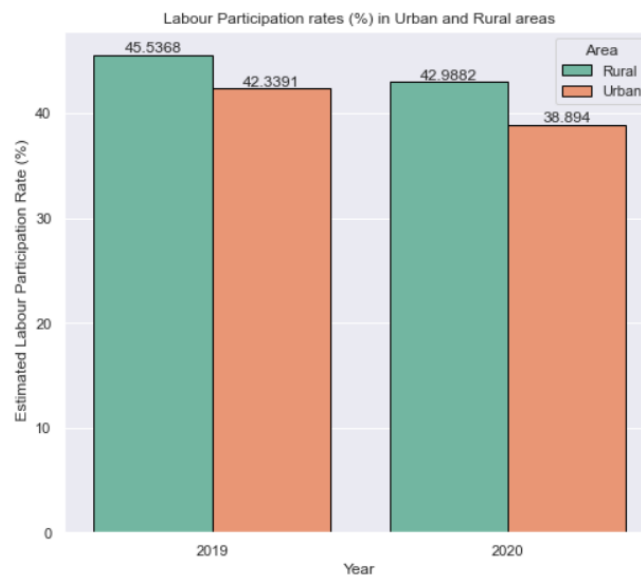

Fig 13: Labour Participation rates in Urban and Rural areas

However, we observed that labour participation rates were impacted approximately equally in urban and rural areas. Labour participation is more in rural areas than Urban areas.

Conclusion

In order to determine the impact of multiple variables on the unemployment rate, we researched various visualization techniques and libraries. As a result, we learnt about visualizations like the bar chart race, choropleths, and scattergeo plots, among others. There are a lot of parameters that can be used with the bar chart race library. We tweaked some to get the desired speed of the bar charts. It was also a great opportunity to work with matplotlib extensively. In our project, we have identified the effects of variables like GDP per capita, literacy rate, and labor participation rate on the unemployment crisis in India using the datasets we have chosen from Kaggle.

On interpretation of our visualizations, we found a relation between GDP-per capita of States that constitutes the majority of Indian economy and the unemployment rates in the country. We learnt that their unemployment rates were less as compared to the mean unemployment rate across all the States. Later on, we understood the magnitude of the impact of the COVID-19 pandemic to the unemployment in the country. However, there were some anomalies and on further research, those anomalies made sense in the context of visualization. We wanted to consider education also as a factor for unemployment rates but we couldn't come to a concrete conclusion using it. Overall, we think there are a lot more factors that can explain the unemployment crisis of a country.

Future work

Only a few factors that are or could be significant in affecting India's unemployment rate have been taken into account in this project. A few further improvements can be considered:

1. Deeper analysis on Male and Female unemployment rates and factors leading to the rise/decline of them.
2. Studying other important factors like increase in population, poverty in different States.
3. Identifying other educational factors like the number of schools in States, academic graduation statistics of population that relates to the unemployment rates across all States.
4. Consider more factors that could affect unemployment rates like government initiatives on employment programs, sector-wise industry layoffs etc.

Appendix

The data and code can be accessed for this project from this GitHub repository link -

https://github.com/vishwa3011/Data-Visualization-IU-Project

References

1. https://www.kaggle.com/datasets/gokulrajkmv/unemployment-in-india?select=Unemployment_Rate_upto_11_2020.csv
2. https://www.kaggle.com/datasets/rajanand/education-in-india
3. https://www.kaggle.com/datasets/varunnagpalspyz/india-statewise-gdp-per-capita-geometry
4. https://www.kaggle.com/datasets/mukeshmanral/income-data
5. https://unemploymentinindia.cmie.com/
6. http://udise.in/src.htm
7. https://www.timesnownews.com/business-economy/economy/article/how-the-covid-19-outbreak-has-affected-the-joblessness-rate-in-india-explained-in-4-charts/634284
8. https://prsindia.org/theprsblog/impact-of-covid-19-on-unemployment-in-urban-areas
9. https://tradingeconomics.com/india/labor-force-total-wb-data.html#:~:text=Labor%20force%2C%20total%20in%20India,Bank%20on%20December%20of%202022.
10. https://en.wikipedia.org/wiki/Labour_in_India
11. https://bengaluru.citizenmatters.in/womens-unemployment-has-worsened-in-urban-karnataka-due-to-the-lockdown-plfs-april-june-2020-58861
12. https://www.dexplo.org/bar_chart_race/