

MARKET SEGMENT ANALYSIS

ED-TECH MARKET

Team - UTKARSH

UTKARSH,RISHABH,MADHUSMITA,SHRISTI

PREFACE

Artificial Intelligence systems mimics human or animal behaviour, but when we want the machine to improve from its experience then machine learning is employed. Machine Learning imparts the computer the ability to learn from its experience and thereby improve its performance. Today machine learning is employed in almost every field. Its scope can now be seen in Market Segment Analysis as well , by employing various ML techniques it is now possible to segment the market in segments, which helps in increasing profits of business firms.

Market segmentation is a marketing term that refers to aggregating prospective buyers into groups or segments with common needs and who respond similarly to a marketing action. Applying various ML strategies it is now possible to segment the market in clusters. This can be done via various supervised and unsupervised ML techniques.

What is EdTech?

EdTech is short for “education technology.” It uses computers, computer programs, and educational systems to provide learning and training to students and employees.

The broad field of EdTech encompasses not only the hardware and software programs used in what is commonly known as “remote learning,” “distance learning,” or “online education,” but also theories of learning and increasing research into what are the most effective means of teaching people new knowledge and skills.

CFI is one of the pioneer firms in EdTech, providing training and certification for professionals and aspiring professionals throughout the financial services industry worldwide.

The Growth of EdTech

EdTech's been around longer than most people think. Computer-aided learning is simply a natural development arising from improved technology. In the mid-twentieth century, schools and training centers were already using available technology, such as copy machines and media learning through training films.

As early as the 1960s, educational institutions were conducting trials and experiments that involved using computers to help teach students math and spelling. Online courses sprang up soon after the advent of internet connections, and it is estimated that by 2010, there were nearly six million students taking courses online in the U.S. alone.

The United Kingdom's Open University and Canada's University of British Columbia were early innovators in making online education interactive, enabling direct communication for teaching and discussion between teachers and students, and between students and other students.

The development of multimedia sources and better video conferencing computer programs makes EdTech both more accessible and more effective. The industry now includes such innovations as EdTech “robots” that students can use to take notes and [blockchain](#) tools that educators can access to grade student work.

Advantages Offered by EdTech

Educational technology offers numerous significant advantages. Educators believe that different students learn at different paces and learn best in diverse ways.

For example, while some students learn information best from reading, others benefit more from audio-visual presentations of information. EdTech offers students the flexibility to access learning from the sources they find most helpful and effective.

A major benefit hailed by online learning students is that of being able to access teaching and training at their own individual pace and time. The flexibility to access online learning at whatever time is most convenient for the student makes education and ongoing professional training much more available to all types of people, in all types of situations.

It is particularly helpful to people with a full-time job and, therefore, would find it extremely difficult to attend regular courses at a university.

An obvious and substantial benefit of EdTech is the massive potential cost savings in education. Online e-learning can be provided at a fraction of the cost of attending a major university or even a small community college.

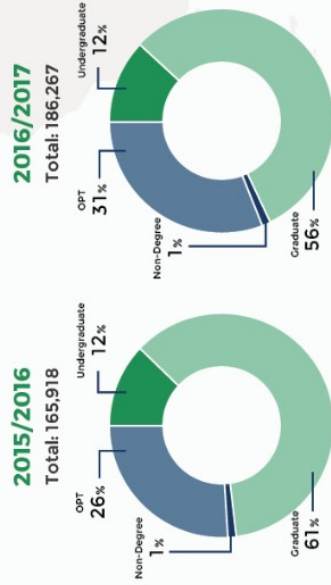
The reduced cost of online learning makes obtaining a degree or additional professional certification much more affordable for anyone interested in expanding their knowledge and skills.

EDUCATION IN INDIA

QUICK FACTS

- Population**
1,340,000,000
(2017, World Bank)
- Compulsory Education**
Grade 1 - 8
(age 6 - 14)
- Languages of Instruction**
Hindi, English
and various Indian languages.
English is the main language of instruction in higher education
- Academic Year**
June/July - March/April
School System
July - May
Higher Education

TOTAL NUMBER OF INDIAN STUDENTS IN THE U.S.



Source: Institute of International Education



INDIA'S INBOUND AND OUTBOUND DEGREE-SEEKING STUDENTS

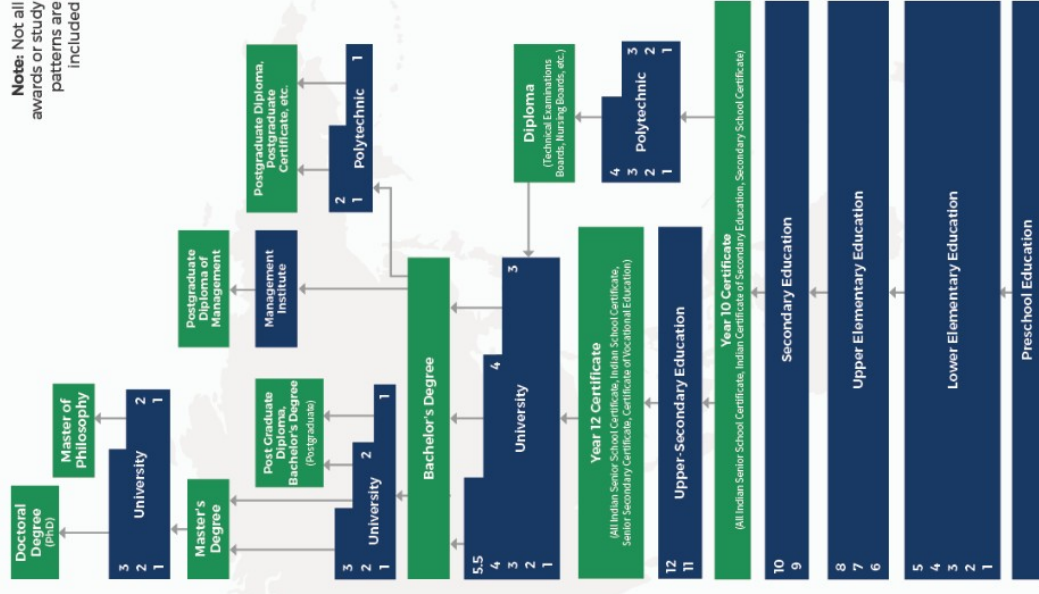
Inbound: **44,766**
Outbound: **277,387**



WES Document Requirements

- Secondary Education**
Final examination certificate (HSC, AISSC or any other standard XII certificate) - sent directly by the examinations board
- State Boards of Technical Education**
Diploma (final or provisional) - submitted by the student
- Statement of marks** - sent directly by the examining board
- Higher Education**
Annual mark sheets (official academic records) - sent directly by the degree-awarding university
- Final or Provisional Degree Certificate - submitted by the student
- Documents must be issued by the university or autonomous college. Documents from affiliated colleges are insufficient and need to be accompanied by university-issued mark sheets.

EDUCATIONAL SYSTEM



Source: Noostr

Overview:

Current Status of Ed Tech market:

Education technology (EdTech) solutions are expected to evolve in line with the advances in the latest technologies, such as the Internet of Things (IoT), Artificial Intelligence (AI), Augmented Reality (AR), and Virtual Reality (VR), and contribute significantly to the market growth. The global education technology market size was valued at USD 89.49 billion in 2020 and is expected to witness a compound annual growth rate (CAGR) of 19.9% from 2021 to 2028. Learners are increasingly shifting toward eBooks that can be accessed online from anywhere across the globe.

Digital technology can improve access to education. Digital content is comparatively easy to generate than printed content, which tends to incur higher production costs. Moreover, digital books are available in different languages and can be easily translated and retrieved by a wider user base. In addition, learners, especially with physical disabilities, can listen to the educational content in an audio format to improve their vocabulary and encourage better interpretive reading.

A boom in Education sector has been witnessed in recent years. This can be attributed to following factors:

1) Increase in EdTech industry: The Indian EdTech industry is believed to have received a \$16.1B in VC funding, a 32X increase from 500M received in 2010. The growth spurt in this industry is driven largely by K-12 Segment, higher education, and upskilling categories. Coupled with the growing popularity of Massive Open Online Courses (MOOCs) and distant education India's EdTech industry is poised to reach \$30 billion in the next 10 years.

2) Government Digital Initiatives: The encouragement and provisions made by the Govt. through initiatives like the SWAYAM (study webs

of active learning for young aspiring minds) Diksha, e-pathshala, etc. to encourage educational institutes even in rural towns to switch to online methods of teaching has helped ensure continuity in academic sessions.

3) Accelerated growth amid pandemic: After the outbreak of Covid-19 pandemic, with closures of institutions and wider adoption of online education, the demand for Edtech products increased multi-fold. In 2016, while the online education market had 1.6 million paid users, there is a projection of 9.6 million users by 2021 (KPMG, 2019).

The data we worked on was based on survey that was conducted in the year 2015-16. It displays how literacy rate is effected by various factors like availability of schools, basic emminities in schools etc.

Use of Machine Learning:

With the help of machine learning we have segmented the states into groups such that states with similar features fall in one group. Also we have tried to describe which segments would be the optimal choice for our project. Through ML we created groups of states which have electricity, water supply and road connectivity to schools. As the supply of electricity , water and road connectivity are basic indicators of development in any region, the states which have these features are an optimal choice for our project.

Data Collection:

Data was scraped from the web site <https://www.kaggle.com>. The dataset we used had State wise primary and secondary education data with lots of features that useful for analyzing the edtech sector.

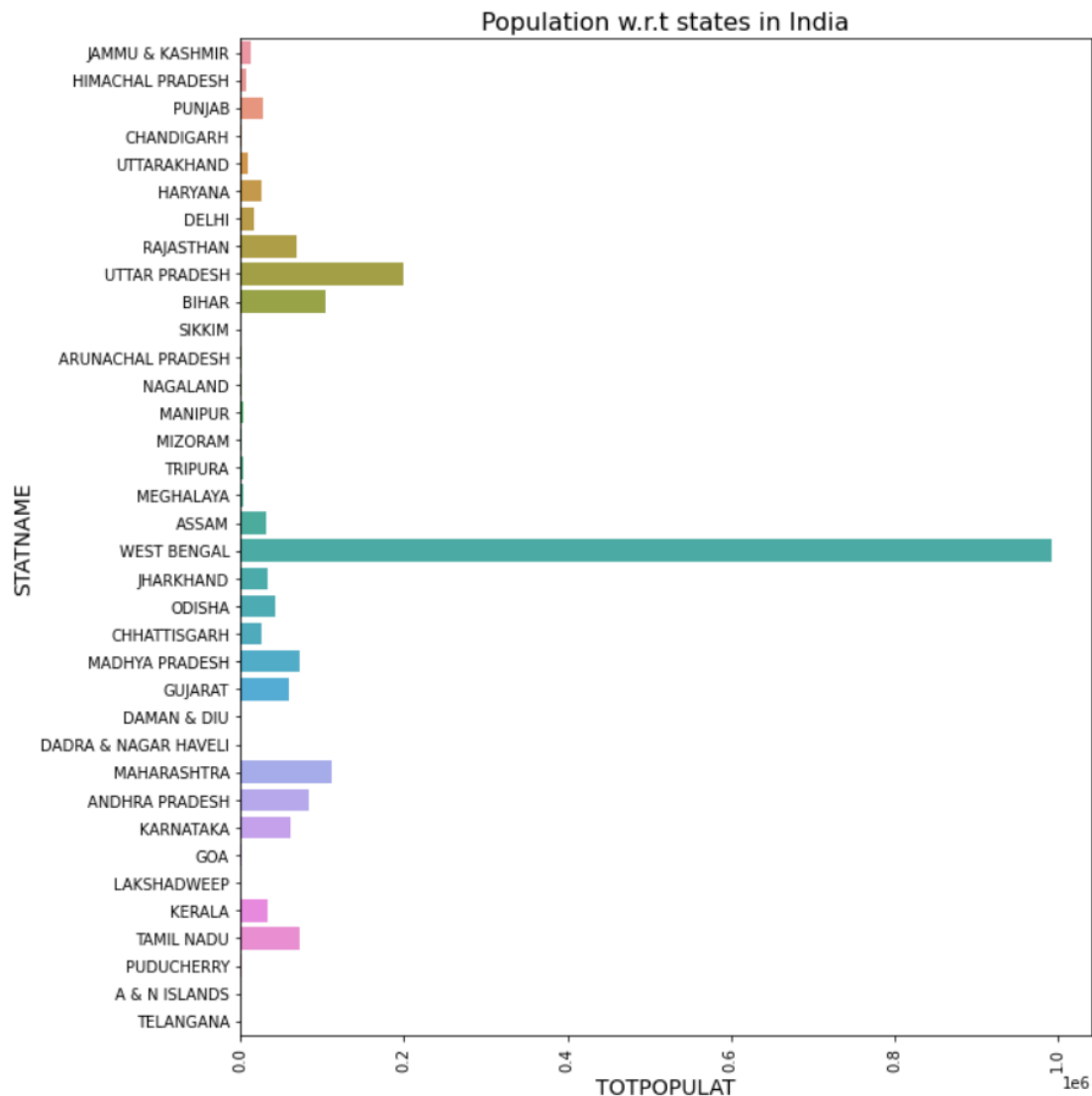
<https://www.kaggle.com/datasets/rajanand/education-in-india>

This dataset contains district wise education data for academic year 2015-16. The file 201516Districtwise_Metadata.csv has all the field details.

Data Pre-processing:

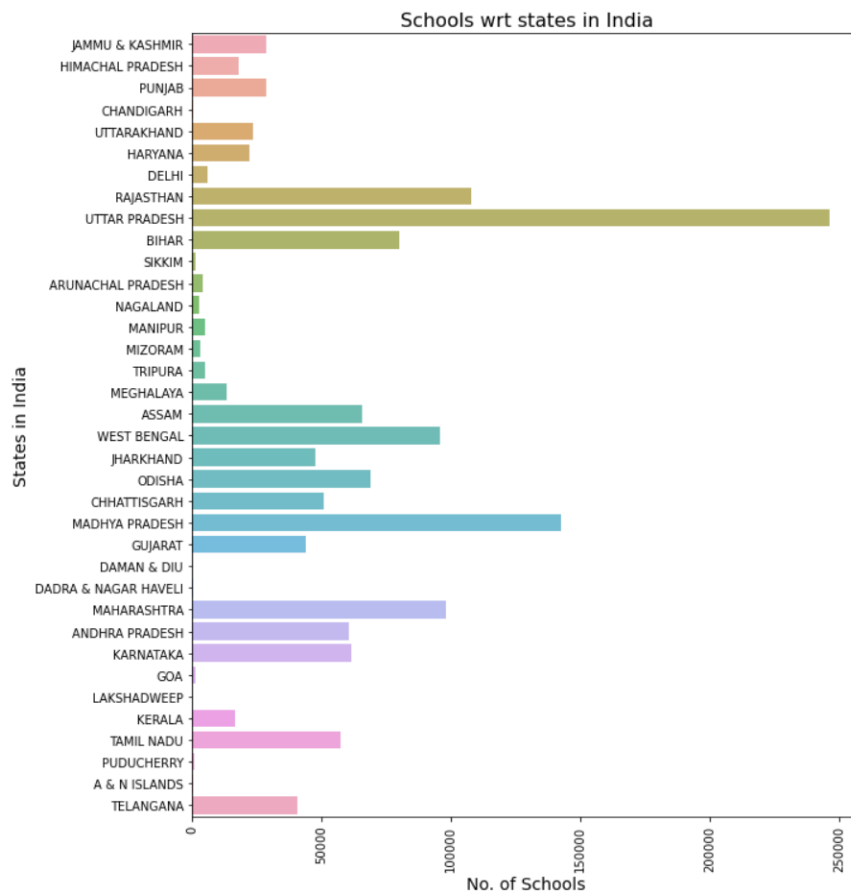
Data pre-processing is **a process of preparing the raw data and making it suitable for a machine learning model**. It is the first and crucial step while creating a machine learning model. For our project we imported the following libraries: pandas , numpy ,sea-born and finally matplotlib for getting plots.

After loading the dataset we start with exploration of data-set. The dataset we are using is based on survey which was conducted in the year 2015-16. We start with the description of school going population in India.



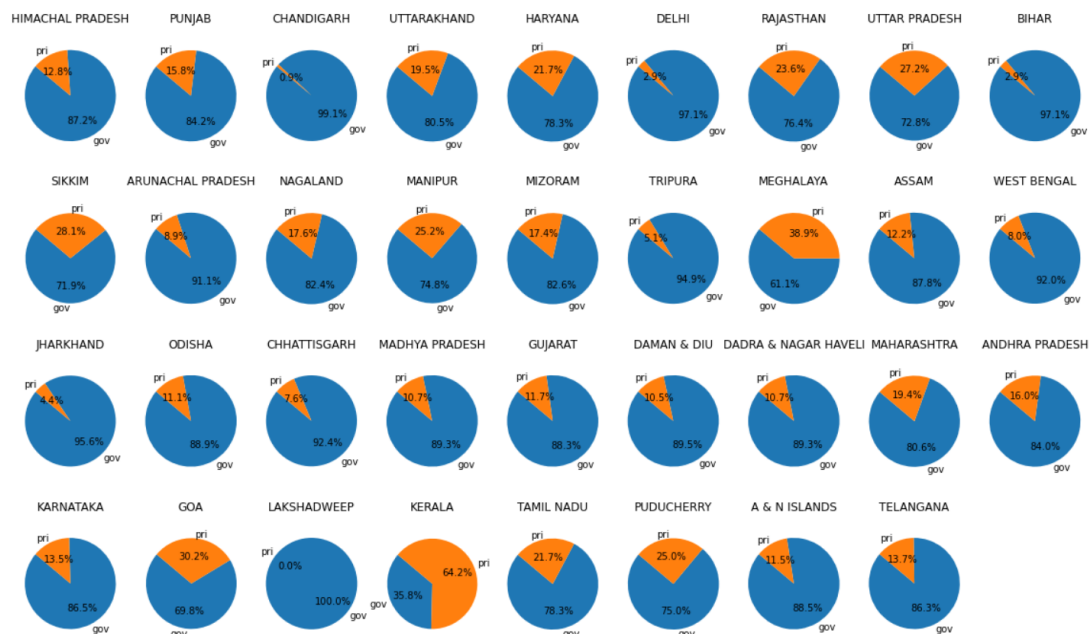
On observing the plot we can see that West Bengal has highest population of school going children followed by Uttar Pradesh, Maharashtra and Bihar. Though states with highest school going population will be optimal for our project but still only having school going population is not what we need. Along with school going population we need to see the infrastructure of available for education system in various states.

Plotting to see number of schools available in various states.



On plotting we see that state of Uttar Pradesh followed by states of Madhya Pradesh, Rajasthan, Maharashtra, West Bengal and Bihar have highest number of schools. Though these states would be a good choice for our project but we also need to have various facilities in schools, these facilities would be an indicator of development in these areas.

Now, considering that private schools would be better than government schools, we plot to see comparison of number of government and private schools in various states.

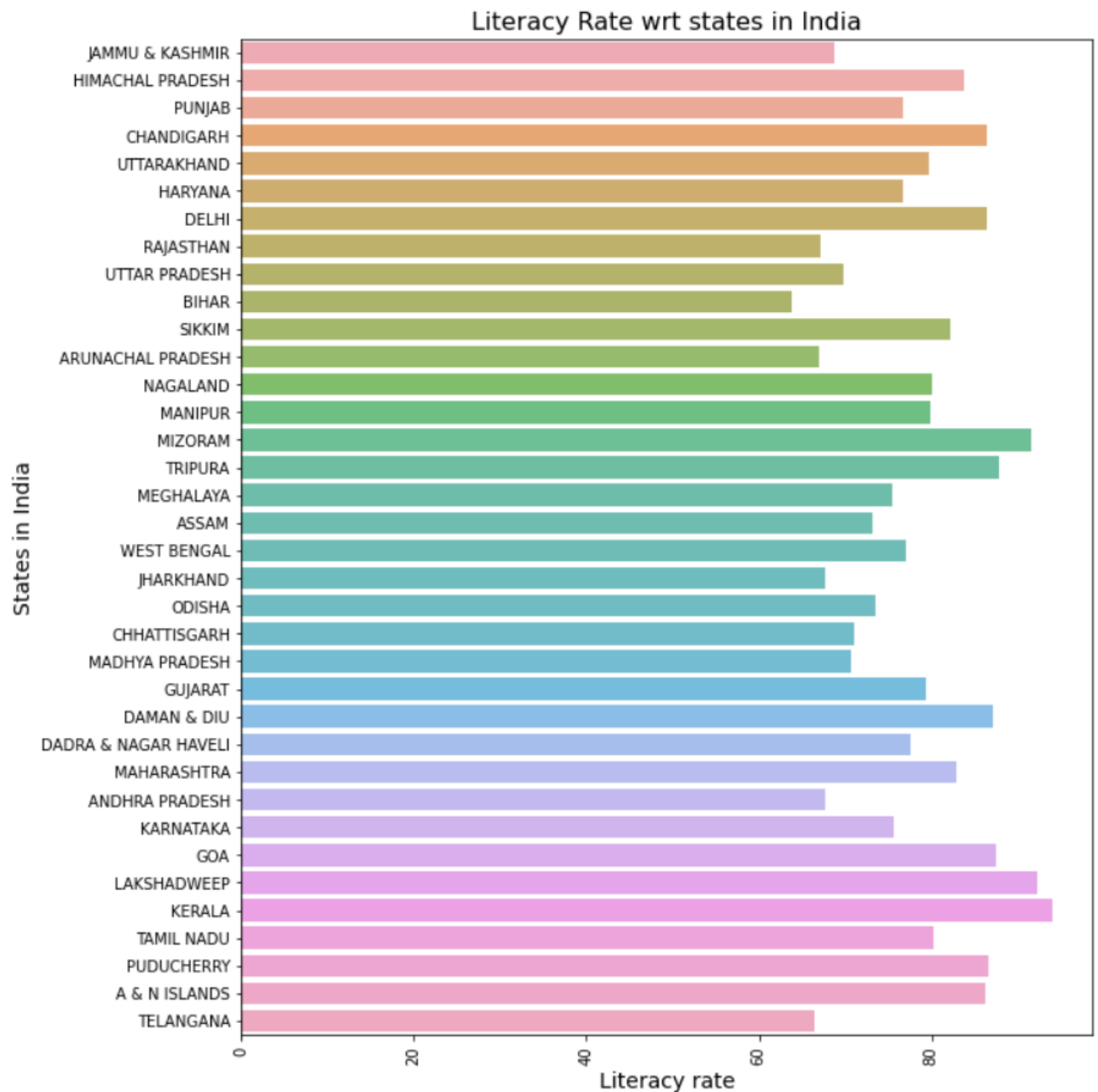


So, state of Kerela has highest number of private schools followed by Meghalay, Sikkim, Uttar Pradesh .We see that most of the states do not have enough number of private schools so number of government and private schools can not be a good factor to be taken into account.

So, now considering Literacy rates in various states in India. This might be a good indicator about the status of educated population. Also, we may assume that educated population would definitely understand the need of education for their children and so such people would prefer that their children should get best education. The states with highly literate population would be best for our start up.

So, plotting to see literacy rate in various states.

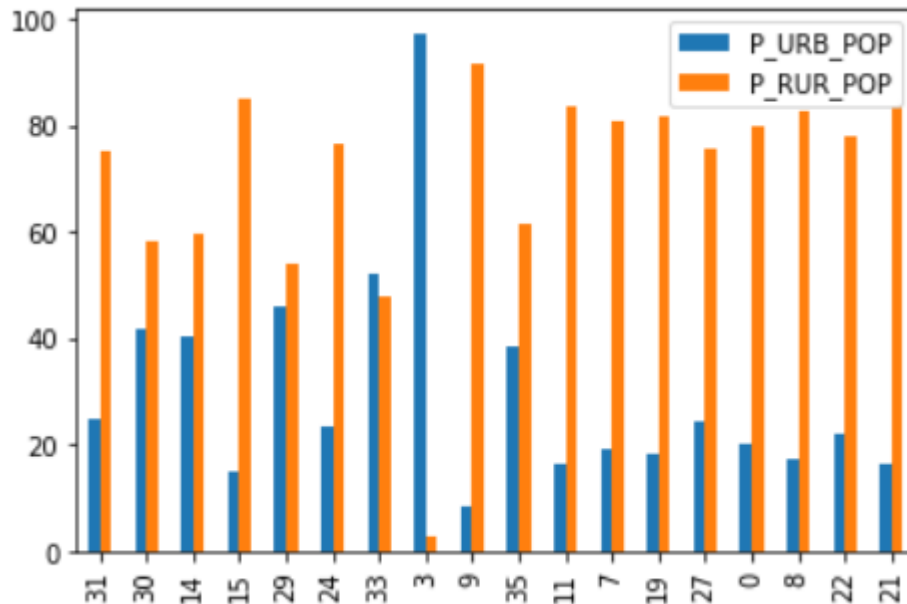
We are using seaborn library for plotting purpose. It is a versatile library and it offers a variety of plots.



Behavioral Analysis:

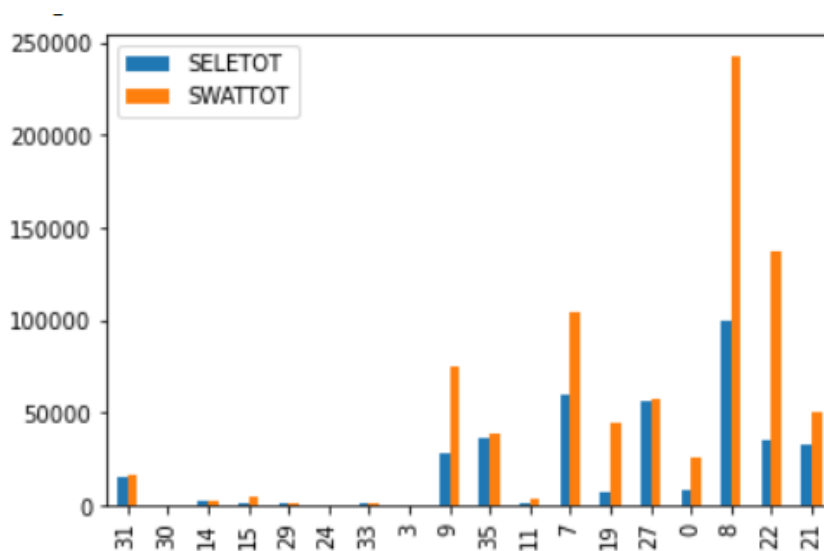
We see that Kerala, Lakshwadeep, Mizoram, Goa, Puducherry , Chandigarh are some of the states with high literacy rates.

Amongst these states now plotting to see the comparison of school going Urban and Rural population.



We see that people living in rural areas are less motivated to go schools even in states with high literacy rates.

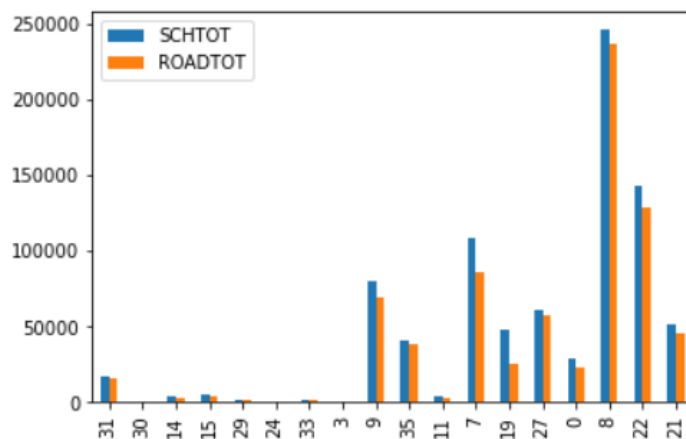
Now, states with electricity and water supply in schools are states where we can expect population to be equipped with technology to use our products. So plotting to see such states.



Andhra Pradesh, Rajasthan, Jharkhand, Bihar , Telengana and Uttar Pradesh have schools with electricity. Thus these states might be sutiable for our project.

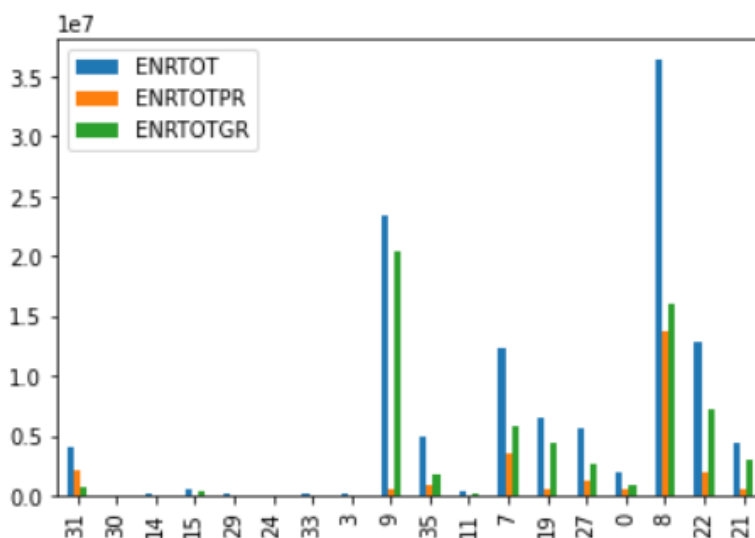
Geographic Analysis:

Road Connectivity can give us good insights into the condition of connections for electricity and water thus plotting to see it's relationship. So, plotting to see connectivity in various states.



Thus, State of Uttar Pradesh, Andhra Pradesh, Rajasthan, Jharkhand and Bihar would be optimal choice for our projects.

Considering private schools better as compared to government schools, elementary enrollemnt in such schools wold have been attracted by facilities avaiable. Thus these schools would be better for our project.



Thus states have more enrollment in government schools than private schools. So, states with considerable enrollment in primary

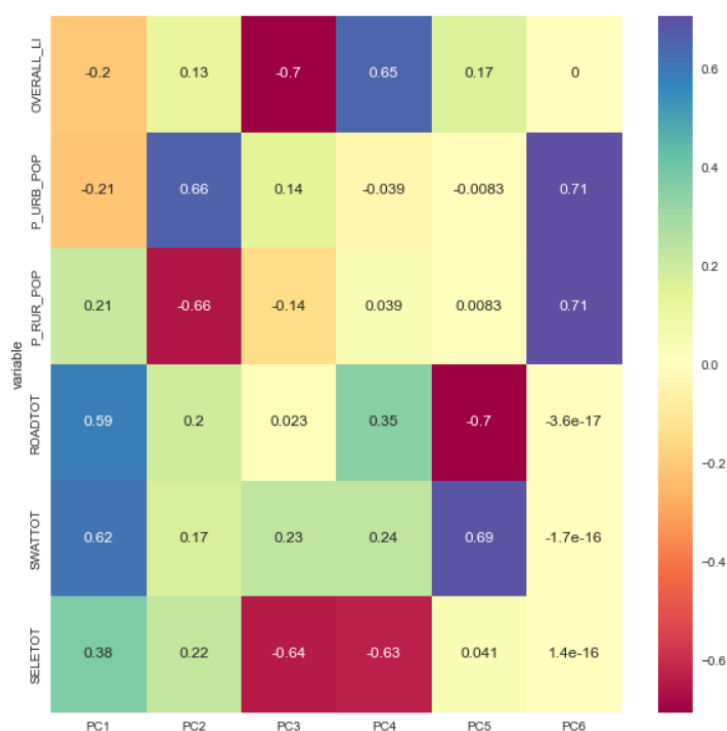
schools would be better for our project. From above plot such states are: Uttar Pradesh, Rajasthan and Madhya Pradesh.

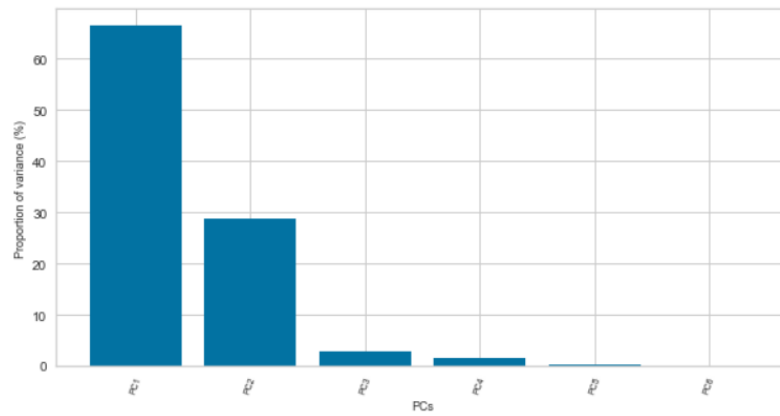
Segment Extraction:

For the purpose of segment extraction we use clustering which is an unsupervised machine learning algorithm. **Clustering** is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them.

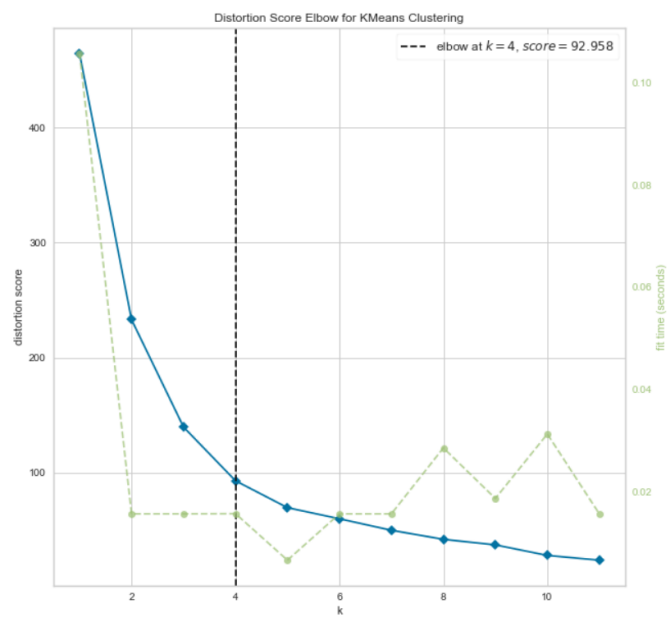
From the data analysis above we see that contributing factors for our project are: urban school going and literate population, rural school going and literate population, road connectivity and availability of water in schools as these factors are indicators for development in any region.

Applying principle component analysis, to reduce number of factors.

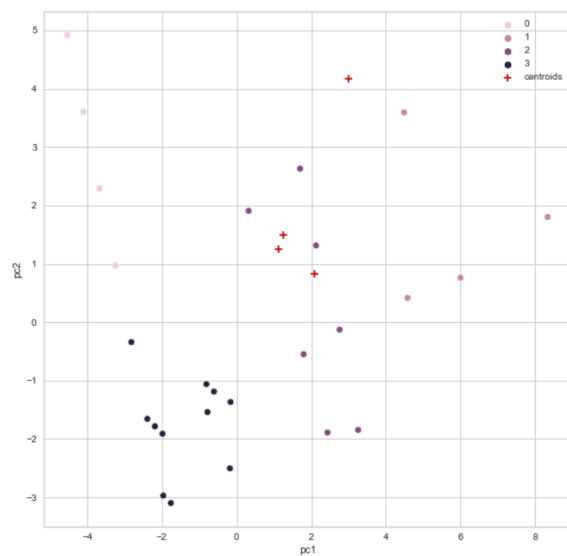




Using clustering on PC1 and PC2.



We thus see that optimal number of clusters is 4. Clusters are:



Profiling:

Market profiling and segmentation generally yields customer profiles that are based on the customers' geographic location (geographic), traits or characteristics (demographic), personality and lifestyle (psychographic), and buying patterns (behavioral). Altogether, profiling and segmentation-related activities will help business owners understand the reason behind under-performing business areas or marketing campaigns that did not fare well among others.

	0	0	1	2	3
OVERALL_LI	1.900000	3.000000	1.25	1.125	2.083333
P_URB_POP	1.700000	4.166667	1.50	1.250	0.833333
P_RUR_POP	4.300000	1.833333	4.50	4.750	5.166667
ROADTOT	1.633333	0.000000	5.50	2.750	0.416667
SWATTOT	1.766667	0.000000	5.75	3.000	0.500000
SELETOT	1.100000	0.000000	3.50	2.000	0.250000
cluster_num	1.866667	0.000000	1.00	2.000	3.000000

Target Segments:

So, from profiling we can conclude that we have now 4 clusters which classify our target segments as:

- 1.States which have good connectivity, schools with power and water supply with high literacy among rural population.
- 2.States which have good connectivity but schools with low power and water supply.
- 3.States which have poor connectivity with poor water and electricity supply at schools.
- 4.States which have poorest connectivity with almost no power and electricity supply at schools.

Out of various clusters we made we see that states in category 1 satisfy our constraints.

Considering that states which have schools with electricity supply, water supply and good road connectivity will be better places for our project . So, from profiling we can see that clusters 1 ,2 and 3 come under the category of feasible solution out of which cluster 1 is a set of optimal solutions to our problem.

Customizing Marketing Mix:

A marketing mix model is a modeling technique used to determine market attribution, the estimated impact of each marketing channel company uses. It is an analysis technique that allow marketers to measure the impact of their marketing and advertising campaigns to determine how various elements contribute to their goal.

Marketing mix modelling uses aggregate data. Marketing mix modeling allows marketers to factor in external influencers such as seasonality, promotions, etc.

Assuming that our ed-tech start up will use TV, radio, newspaper for its advertising we can use it to predict sales.

OLS Regression Results						
Dep. Variable:	sales	R-squared:	0.897			
Model:	OLS	Adj. R-squared:	0.896			
Method:	Least Squares	F-statistic:	570.3			
Date:	Sun, 10 Apr 2022	Prob (F-statistic):	1.58e-96			
Time:	13:01:41	Log-Likelihood:	-386.18			
No. Observations:	200	AIC:	780.4			
Df Residuals:	196	BIC:	793.6			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	2.9389	0.312	9.422	0.000	2.324	3.554
TV	0.0458	0.001	32.809	0.000	0.043	0.049
radio	0.1885	0.009	21.893	0.000	0.172	0.206
newspaper	-0.0010	0.006	-0.177	0.860	-0.013	0.011
Omnibus:	60.414	Durbin-Watson:	2.084			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	151.241			
Skew:	-1.327	Prob(JB):	1.44e-33			
Kurtosis:	6.332	Cond. No.	454.			

The code creates a model which specifies that we are creating a model that predicts sales based on TV, Newspaper, and radio advertising.

From above we can see, The p-values for TV and radio are less than 0.000, but the p-value for newspapers is 0.86, which indicates that newspaper spend has no significant impact on sales.

Potential Customer Base:

Selecting segments 1 and 2 for our analysis, we have:

Customer Base for segment 1: 453,173

Customers Base for segment 2: 327,147

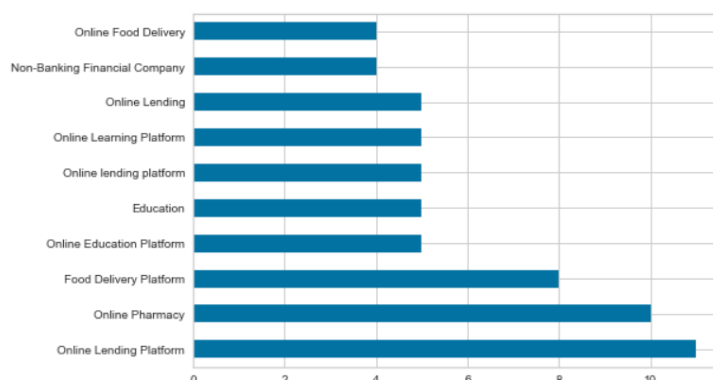
Assuming that our customers can afford up-to 20K, we can see that our potential profit is:

For Segment 1: $\text{Rs.}1.567 \times 10^{10}$

For Segment 2: $\text{Rs.}6.543 \times 10^{10}$

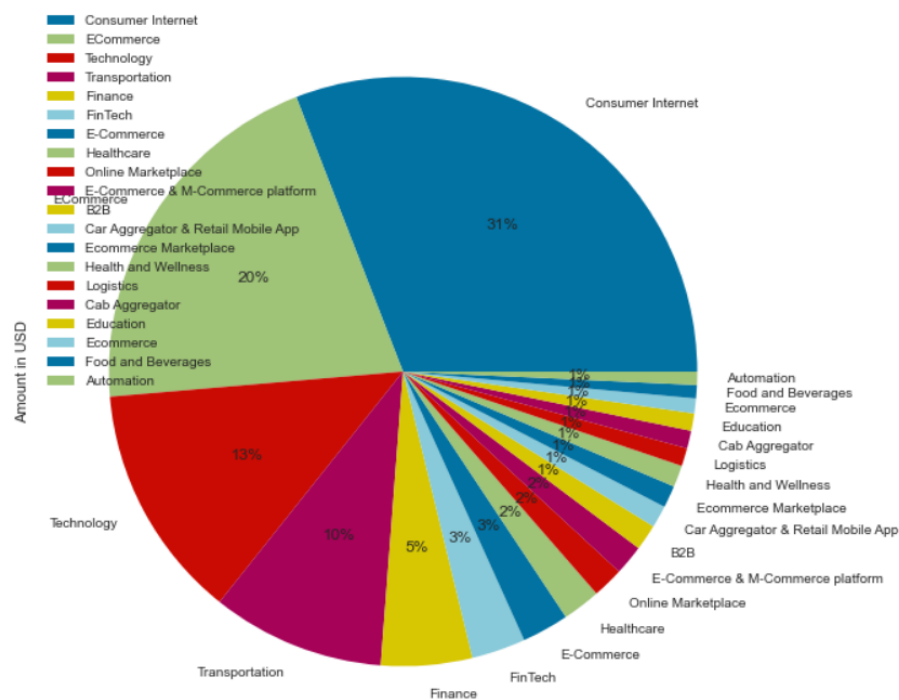
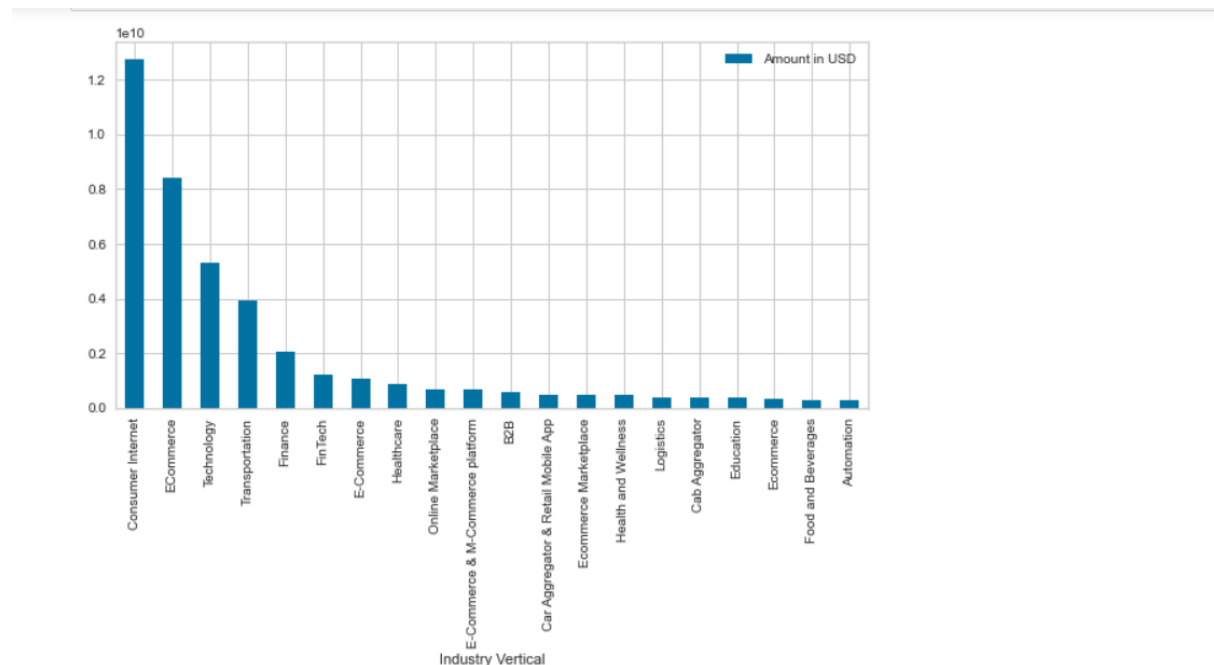
Analyzing Startup funding:

For the purpose of our analysis we import start-up funding dataset from Kaggle. Plotting to see the number of start up with online education .



From above plot we can see that Online learning platform and online education platform are significant in numbers. Thus we can conclude that ed-tech market is a growing industry and can also evolve futhur.

Plotting to see start up funding.



From above plot we can conclude that funding in ed tech start ups is currently very low but as we saw above that online education industry is a growing industry and so we can expect an increase in funding in future.

Strategy for Entering Market:

Target the states as described in the segments created, to choose suitable place for startups . Determine the investment needed and finally by determining the suitable pricing range for product (services that will be provided by use of app in this case, we have estimated in our report that such an amount can be Rs.20K), and adopting suitable advertising strategy it can step into the market.

Code:

https://github.com/Uttu-Joshi/Ed-tech-Market/blob/dev_d1/ed_tech%20starup.ipynb

References:

1. <https://edukemy.com/current-affairs/gazette/2021-12-29/the-edtech-industry-in-india>
2. https://en.wikipedia.org/wiki/Education_in_India
3. https://ruralindiaonline.org/en/library/resource/national-education-policy-2020/?gclid=Cj0KCQjwgMqSBhDCARIsAIIVN1VedNcNntKs9kIt26Z-qFwuXXbkWYUnrAwd7aIXz5YAjt-fPFCwFEMaAuTGEALw_wcB

Datasets Used:

1. https://github.com/Uttu-Joshi/Ed-tech-Market/blob/dev_d1/data1.csv
2. [https://github.com/Uttu-Joshi/Ed-tech-Market/blob/dev_d1/startup_funding\(1\).csv](https://github.com/Uttu-Joshi/Ed-tech-Market/blob/dev_d1/startup_funding(1).csv)
3. https://github.com/Uttu-Joshi/Ed-tech-Market/blob/dev_d1/Advertising.csv