INT217 PROJECT REPORT

(Project Semester August-December 2021)
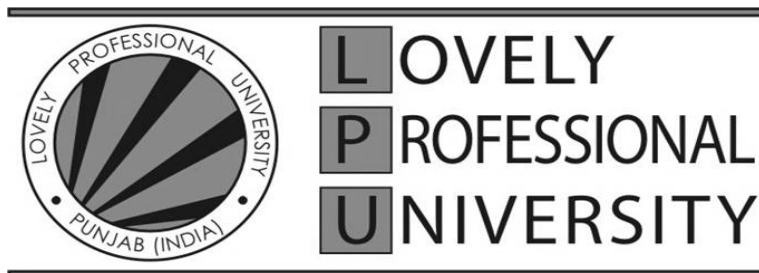

EDUCATION IN INDIA

Submitted by

Asutosh Patra

Registration No. 11907074

Programme: Bachelor's in technology (Computer Science and Engineering)

Section: KM002

Course Code: INT217

Under the Guidance of

Ms. Maneet Kaur (ID - 15709)

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Asutosh Patra bearing Registration no. 11907074 has completed INT217 project titled, **"EDUCATION IN INDIA"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort, and study.

Maneet Kaur

School of Computer Science & Engineering

Lovely Professional University

Phagwara, Punjab.

Date:

<u>DECLARATION</u>

I, Asutosh Patra, student of Bachelor's in technology (Computer Science and Engineering) under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: Signature

Registration No. 11907074   Asutosh Patra

<u>ACKNOWLEDGEMENT</u>

I would like to express my special thanks of gratitude to my teacher Ms. Maneet Kaur who gave me the golden opportunity to do this wonderful project of analysis of the data of a superstore namely "EDUCATION IN INDIA" which also helped me in doing a lot of research and I came to know about so many new things. I am thankful to them. Secondly, I would also like to thank my parents and friends who helped me a lot in finalizing this project within the limited time frame.

TABLE OF CONTENT

INTRODUCTION

Data Analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, while being used in different business, science, and social science domains.

The analytics team of GOVERNMENT regulatory and SOCIAL ORGANIZATIONS anywhere in the India would love to check us through data analysis of every objective leading to a well-organized and fruitful information. My analysis contains data on Top 5 Populated States, Top 5 Populated Districts, State wise Education, District wise Education, Literacy and Literacy Rate.

Education is the process of facilitating learning, or the acquisition of knowledge, skills, values, morals, beliefs, habits, and personal development. Educational methods include teaching, training, storytelling, discussion and directed research. Education frequently takes place under the guidance of educators; however, learners can also educate themselves. Education can take place in formal or informal settings, and any experience that has a formative effect on the way one thinks, feels, or acts may be considered educational. The methodology of teaching is called pedagogy.

Formal education is commonly divided formally into stages such as preschool or kindergarten, primary school, secondary school and then college, university, or apprenticeship. In most regions, education is compulsory up to a certain age.

There are movements for education reforms, such as for improving quality and efficiency of education towards relevance in students' lives and efficient problem solving in modern or future society at large, or for evidence-based education methodologies. A right to education has been recognized by some governments and the United Nations. Global initiatives aim at achieving the Sustainable Development Goal 4, which promotes quality education for all.

Education in India is primarily managed by state-run public education system, which fall under the command of the government at three levels: Central, state and local. Under various articles of the Indian Constitution and the Right of Children to Free and Compulsory Education Act, 2009, free and compulsory education is provided as a fundamental right to children aged 6 to 14. The approximate ratio

of public schools to private schools in India is 7:5. Major policy initiatives in Indian education are numerous. Up until 1976, education policies and implementation were determined legally by each of India's constitutional states. The 42nd amendment to the constitution in 1976 made education a 'concurrent subject'. From this point on the central and state governments shared formal responsibility for funding and administration of education. In a country as large as India, now with 28 states and eight union territories, this means that the potential for variations between states in the policies, plans, programs, and initiatives for elementary education is vast. Periodically, national policy frameworks are created to guide states in their creation of state-level programs and policies. State governments and local government bodies manage most primary and upper primary schools, and the number of government-managed elementary schools is growing. Simultaneously the number and proportion managed by private bodies is growing. In 2005-6 83.13% of schools offering elementary education (Grades 1-8) were managed by government and 16.86% of schools were under private management (excluding children in unrecognised schools, schools established under the Education Guarantee Scheme and in alternative learning centres). Of those schools managed privately, one third are 'aided' and two thirds are 'unaided'. Enrolment in Grades 1-8 is shared between government and privately managed schools in the ratio 73:27. However in rural areas this ratio is higher (80:20) and in urban areas much lower (36:66).

In the 2011 Census, about 73% of the population was literate, with 81% for males and 65% for females. National Statistical Commission surveyed literacy to be 77.7% in 2017–18, 84.7% for male and 70.3% for female. This compares to 1981 when the respective rates were 41%, 53% and 29%. In 1951 the rates were 18%, 27% and 9%.India's improved education system is often cited as one of the main contributors to its economic development. Much of the progress, especially in higher education and scientific research, has been credited to various public institutions. While enrolment in higher education has increased steadily over the past decade, reaching a Gross Enrolment Ratio (GER) of 26.3% in 2019, there still remains a significant distance to catch up with tertiary education enrolment levels of developed nations, a challenge that will be necessary to overcome in order to continue to reap a demographic dividend from India's comparatively young population.

Poorly resourced public schools which suffer from high rates of teacher absenteeism may have encouraged the rapid growth of private (unaided) schooling in India, particularly in urban areas. Private schools divide into two types: recognised and unrecognised schools. Government 'recognition' is an official stamp of approval and for this a private school is required to fulfil a number of conditions,

though hardly any private schools that get 'recognition' actually fulfil all the conditions of recognition. The emergence of large numbers of unrecognised primary schools suggests that schools and parents do not take government recognition as a stamp of quality.

At the primary and secondary level, India has a large private school system complementing the government run schools, with 29% of students receiving private education in the 6 to 14 age group. Certain post-secondary technical schools are also private. The private education market in India had a revenue of US$450 million in 2008 but is projected to be a US$40 billion market.

As per the Annual Status of Education Report (ASER) 2012, 96.5% of all rural children between the ages of 6-14 were enrolled in school. This is the fourth annual survey to report enrolment above 96%. India has maintained an average enrolment ratio of 95% for students in this age group from year 2007 to 2014. As an outcome the number of students in the age group 6-14 who are not enrolled in school has come down to 2.8% in the academic year 2018 (ASER 2018). Another report from 2013 stated that there were 229 million students enrolled in different accredited urban and rural schools of India, from Class I to XII, representing an increase of 2.3 million students over 2002 total enrolment, and a 19% increase in girl's enrolment. While quantitatively India is inching closer to universal education, the quality of its education has been questioned particularly in its government run school system. While more than 95 per cent of children attend primary school, just 40 per cent of Indian adolescents attend secondary school (Grades 9-12). Since 2000, the World Bank has committed over $2 billion to education in India. Some of the reasons for the poor quality include absence of around 25% of teachers every day. States of India have introduced tests and education assessment system to identify and improve such schools.

Although there are private schools in India, they are highly regulated in terms of what they can teach, in what form they can operate (must be a non-profit to run any accredited educational institution) and all the other aspects of the operation. Hence, the differentiation between government schools and private schools can be misleading. However, in a report by Geeta Gandhi Kingdom entitled: The emptying of public Schools and growth of private schools in India, it is said that for sensible education policy making, it is vital to take account of the changing trends in the size of the private and public schooling sectors in India. Ignoring these trends involves the risk of poor policies/legislation, with attendant adverse consequences for children's life chances.

In January 2019, India had over 900 universities and 40,000 colleges. In India's higher education system, a significant number of seats are reserved under affirmative action policies for the historically disadvantaged Scheduled Castes and Scheduled Tribes and Other Backward Classes. In universities, colleges, and similar institutions affiliated to the central government, there is a maximum 50% of reservations applicable to these disadvantaged groups, at the state level it can vary. Maharashtra had 73% reservation in 2014, which is the highest percentage of reservations in India.

SCOPE OF ANALYSIS

This project on Education in India provides the overall Statistics details of the population and literacy in various aspects from the year 2001.

Objectives of this project:

• To good hand on excel.

• To use different feature and get friendly with the excel.

• To learn the ETL process in the Tableau Prep.

• How to link one sheet to another and traverse between different sheets.

• How to use pivot table and pivot chart.

• Learn to make dashboard in excel.

• To make different type of graphs in excel.

• To learn how to fetch data from other source to excel in different formats.

Aim of this project is to answer the above objectives in the form of visualization by creating a dashboard to convey the answers effectively and efficiently.

SOURCE OF DATASET

The data is being taken from the Kaggle. Kaggle is an Airbnb for Data Scientists – this is where they spend their nights and weekends. It's a crowd-sourced platform to attract, nurture, train, and challenge data scientists from all around the world to solve data science, machine learning and predictive analytics problems. It has over 536,000 active members from 194 countries, and it receives close to 150,000 submissions per month. Started from Melbourne, Australia Kaggle moved to Silicon Valley in 2011, raised some 11 million dollars from the likes of Hal Varian (Chief Economist at Google), Max Levchin (PayPal), Index and Khosla Ventures and then ultimately been acquired by the Google in March of 2017. Kaggle is the number one stop for data science enthusiasts all around the world who compete for prizes and boost their Kaggle rankings. There are only 94 Kaggle Grandmasters in the world to this date.

Kaggle, a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

Kaggle got its start in 2010 by offering machine learning competitions and now also offers a public data platform, a cloud-based workbench for data science, and Artificial Intelligence education. Its key personnel were Anthony Goldbloom and Jeremy Howard. Nicholas Gruen was founding chair succeeded by Max Levchin. Equity was raised in 2011 valuing the company at $25 million. On 8 March 2017, Google announced that they were acquiring Kaggle.

ETL PROCESS

In computing, extract, transform, load (ETL) is a process in database usage to prepare data for analysis, especially in data warehousing. Data extraction involves extracting data from homogeneous or heterogeneous sources, while data transformation processes data by transforming them into a proper storage format/structure for the purposes of querying and analysis; finally, data loading describes the insertion of data into the final target database such as an operational data store, a data mart, or a data warehouse. A properly designed ETL system extracts data from the source systems, enforces data quality and consistency standards, conforms data so that separate sources can be used together, and finally delivers data in a presentation-ready format so that application developers can build applications and end users can make decisions.

Precisely, ETL is defined as a process that extracts the data from different RDBMS source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system. ETL stands for Extract, Transform and Load.

Before ETL, the dataset looked like this. **This data is taken from Kaggle.**
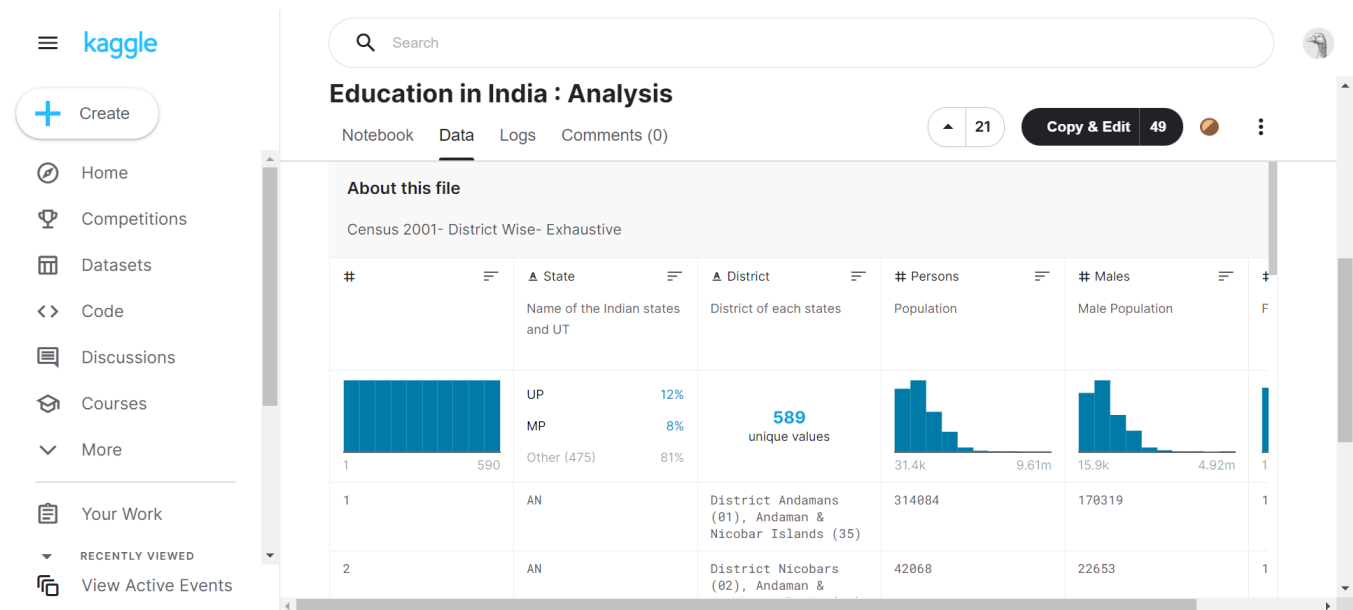
Through the process of ETL, we are going to clean the dataset and bring all the entities to their proper data format.

Step 1: Removing the blank cells from the dataset.

For this, select the whole dataset. Go to Find and Select in the Home tab of excel. Select Go to Special from the drop-down menu and then tick the blank option. All the blank cells will be selected. Then go to Delete option in the home tab again and select Delete Rows from the drop-down menu. This will remove any rows with blank cells.

Step 2: Removing columns which are not properly defined or not crucial to our analysis.

For this we will columns which are redundant like the column with just the index numbers. For this we will select that particular column and then go to delete option in the home tag and then select Delete Columns from the drop-down menu.

Step 3: Giving proper and appropriate column names.

The dataset does not have proper columns so our next step would be to giver proper column names to the columns wherever required.

Step 4: Excluding the NULL values from the data.

We'll be using Tableau prep for this work as it'll make the work simple and faster because we might not know how many null values could be there in this huge data set. Tableau helps us doing one step cleaning with ease.

Step 5: Improvising Proper Data Formatting

Without proper Data Formatting, proper analysis will not take place. So, we will bring down certain columns to their proper format. For example, the dates should be in the date format and price and sales should be in currency format for better results.

Step 6: Removing Duplicate Values

It might be possible that our data may be containing duplicate values which may hinder in precise analysis. So, our last task in ETL will be removing duplicate values and making our data perfect for analysis.

Top screen:

Alerts (0)

all — Clean 1 — Output

100%

Output   82 fields

Save output to
File
Browse

Name
Final Dataset

Location
E:

Output type
Microsoft Excel (.xlsx)

Worksheet
Final

Write Options
Select an option to create or update your output table.

Run Flow

Save to Final Dataset.xlsx

F1	State	District	Persons	Males	Females	Gro
1	Andaman & Nicobar Islands	District Andamans (01), Andaman & Nicobar Islands (35	3,14,084	1,70,319	1,43,765	3
2	Andaman & Nicobar Islands	District Nicobars (02), Andaman & Nicobar Islands (35)	42,068	22,653	19,415	7
3	Andhra Pradesh	District Adilabad (01), Andhra Pradesh (28)	24,88,003	12,50,958	12,37,045	1
4	Andhra Pradesh	District Nizamabad (02), Andhra Pradesh (28)	23,45,685	11,62,905	11,82,780	1
5	Andhra Pradesh	District Karimnagar (03), Andhra Pradesh (28)	34,91,822	17,47,968	17,43,854	1
6	Andhra Pradesh	District Medak (04), Andhra Pradesh (28)	26,70,097	13,52,446	13,17,651	1
7	Andhra Pradesh	District Hyderabad (05), Andhra Pradesh (28)	38,29,753	19,81,173	18,48,580	1
8	Andhra Pradesh	District Rangareddi (06), Andhra Pradesh (28)	35,75,064	18,39,227	17,35,837	3
9	Andhra Pradesh	District Mahbubnagar (07), Andhra Pradesh (28)	35,13,934	17,82,340	17,31,594	1
10	Andhra Pradesh	District Nalgonda (08), Andhra Pradesh (28)	32,47,982	16,51,990	15,95,992	1
11	Andhra Pradesh	District Warangal (09), Andhra Pradesh (28)	32,46,004	16,44,895	16,01,109	1
12	Andhra Pradesh	District Khammam (10), Andhra Pradesh (28)	25,78,927	13,05,543	12,73,384	1
13	Andhra Pradesh	District Srikakulam (11), Andhra Pradesh (28)	25,37,593	12,60,020	12,77,573	8

Bottom screen:

Alerts (0)

100%

Clean 1   82 fields 590 rows   Filter Values...   2 Recommendations

Changes (1)

Group Values
[State]
2 values replaced

F1 590

0
200
400
600

State 35
Search
Karnataka
Kerala
Lakshadweep
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Orrisa
Pondicherry
Punjab

District 589
District Bishnupur (04)...
District Bokaro * (13), ...
District Bongaigaon (0...
District Budaun (19), U...
District Bulandshahar
District Buldana (04), ...
District Bundi (23), Raj...
District Buxar * (30), B...
District Cachar (21), A...
District Central
District Chamarajanag...
District Chamba (01), ...

Persons 590

0
20,00,000
40,00,000
60,00,000
80,00,000
1,00,00,000

| F1 | State | District | Persons | Males | Females | G |

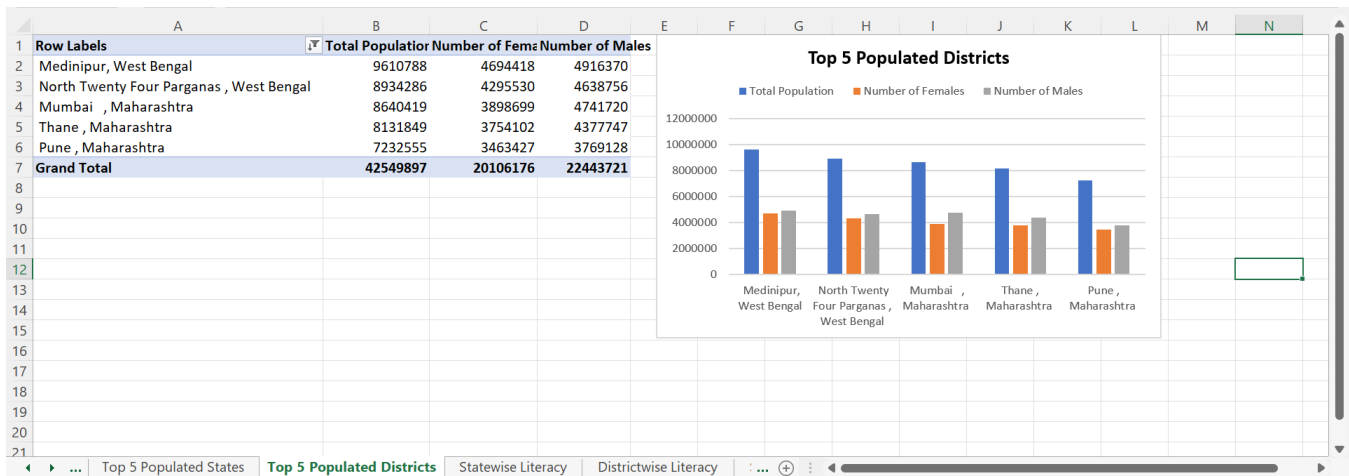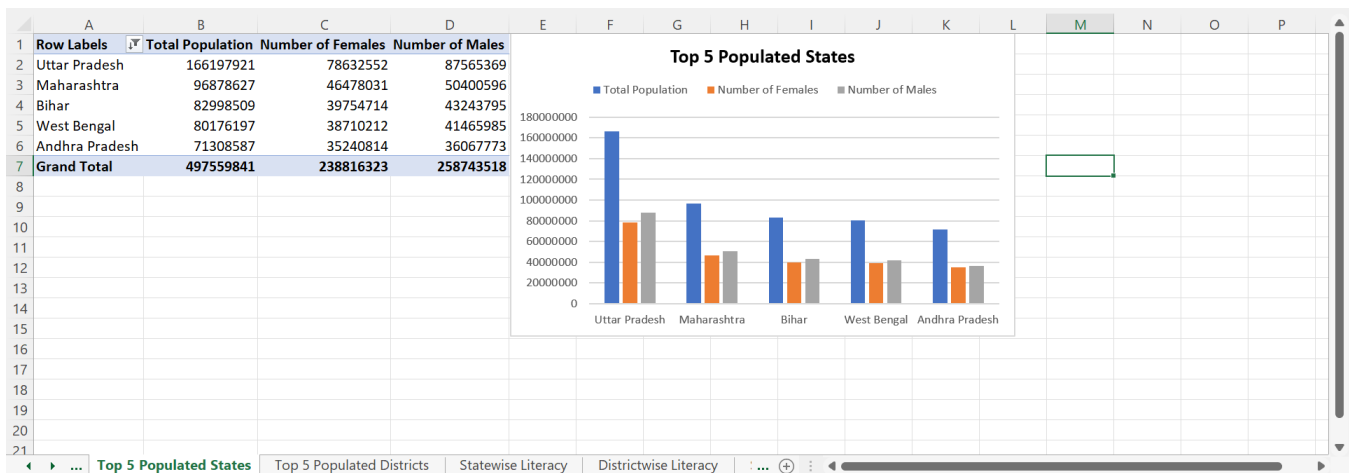# ANALYSIS OF DATASET

1. **Top 5 Populated States and Districts:**

Description:

In this objective we will be finding the top 5 populated States and Districts for this we will be using the total population of each state and district.

Specific function and requirements

We have to create a pivot table to determine the populated states and districts, and then visualize it on graph.

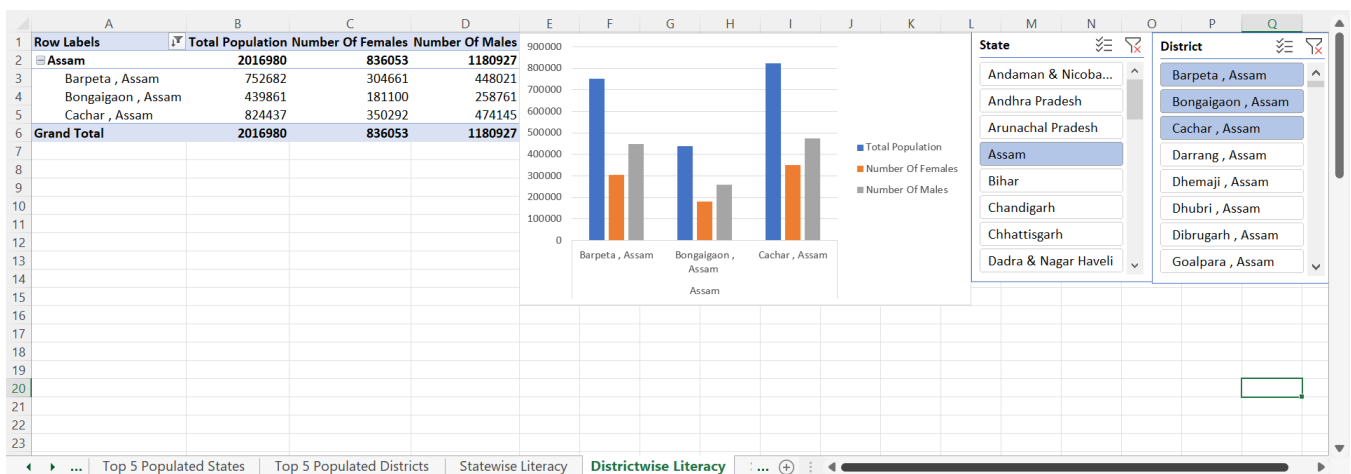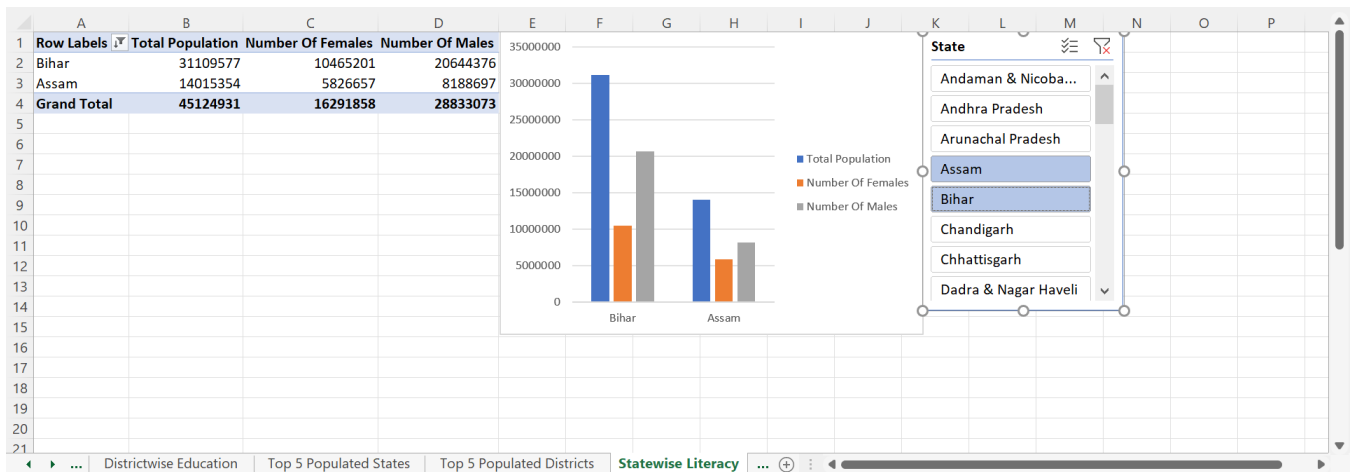**Results and Visualization:**

2. State and District wise Literacy:

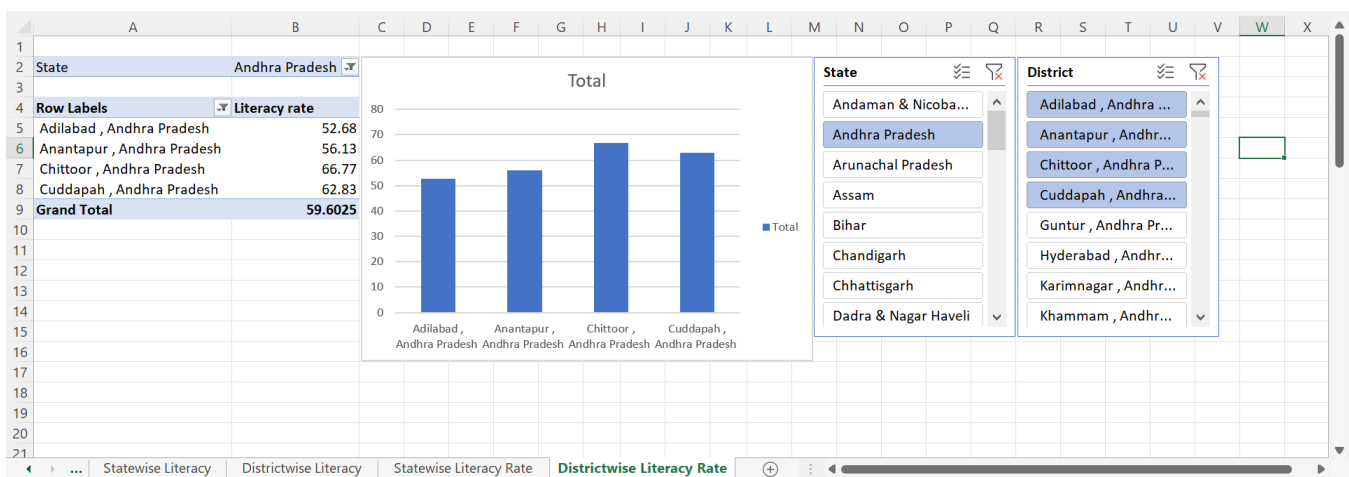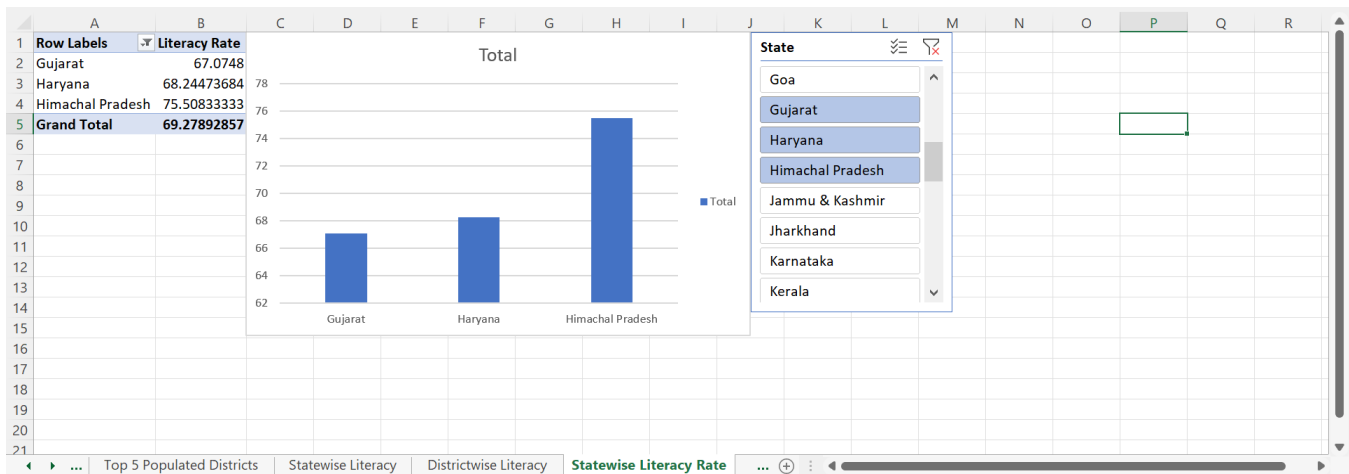**Description:**

In this objective we will be finding the literate population of each States and Districts for this we will be using the total literate population of each state and district.

Specific function and requirements

We have to create a pivot table and slicer to determine the literate states and districts, and then visualize it on graph.

Results and Visualization:

3. State and District wise Literacy Rate:

Description:

In this objective we will be finding the literacy rate of each States and Districts for this we will be using the literacy rate of each state and district.

Specific function and requirements

We have to create a pivot table and slicer to determine the literacy rate of each states and districts, and then visualize it on graph.

Results and Visualization:

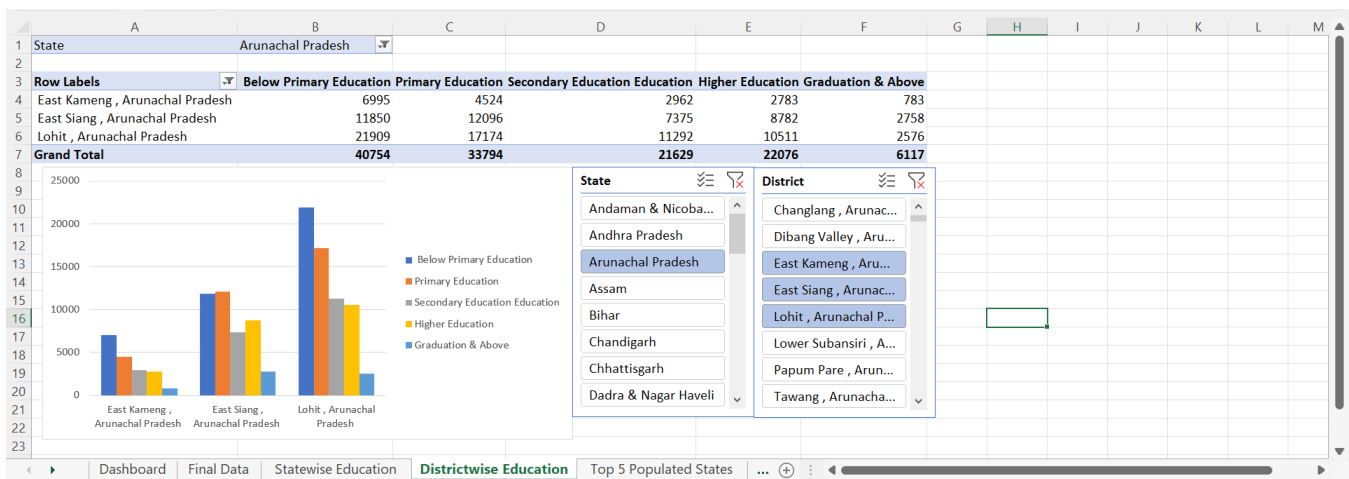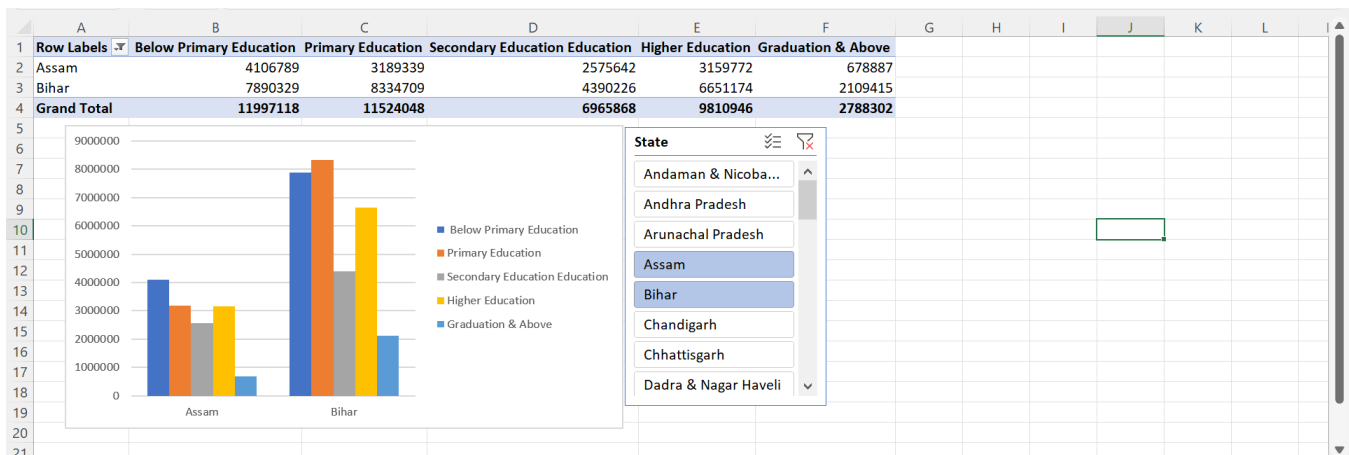## 4. State and District wise Education:

Description:

In this objective we will be finding the education of each States and Districts (**Below Primary Education, Primary Education, Secondary Education, Higher Education, Graduation & Above)** for this we will be using the no of students in each state and district in all categories.
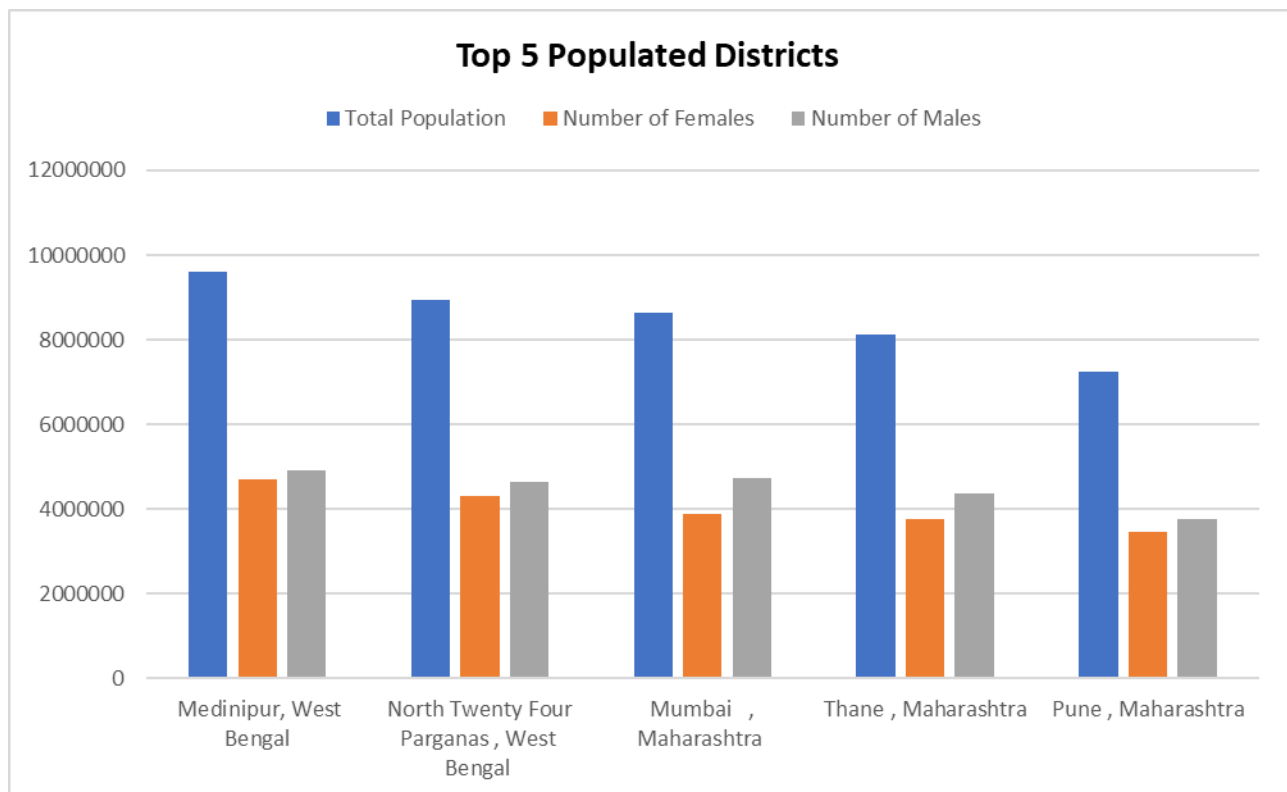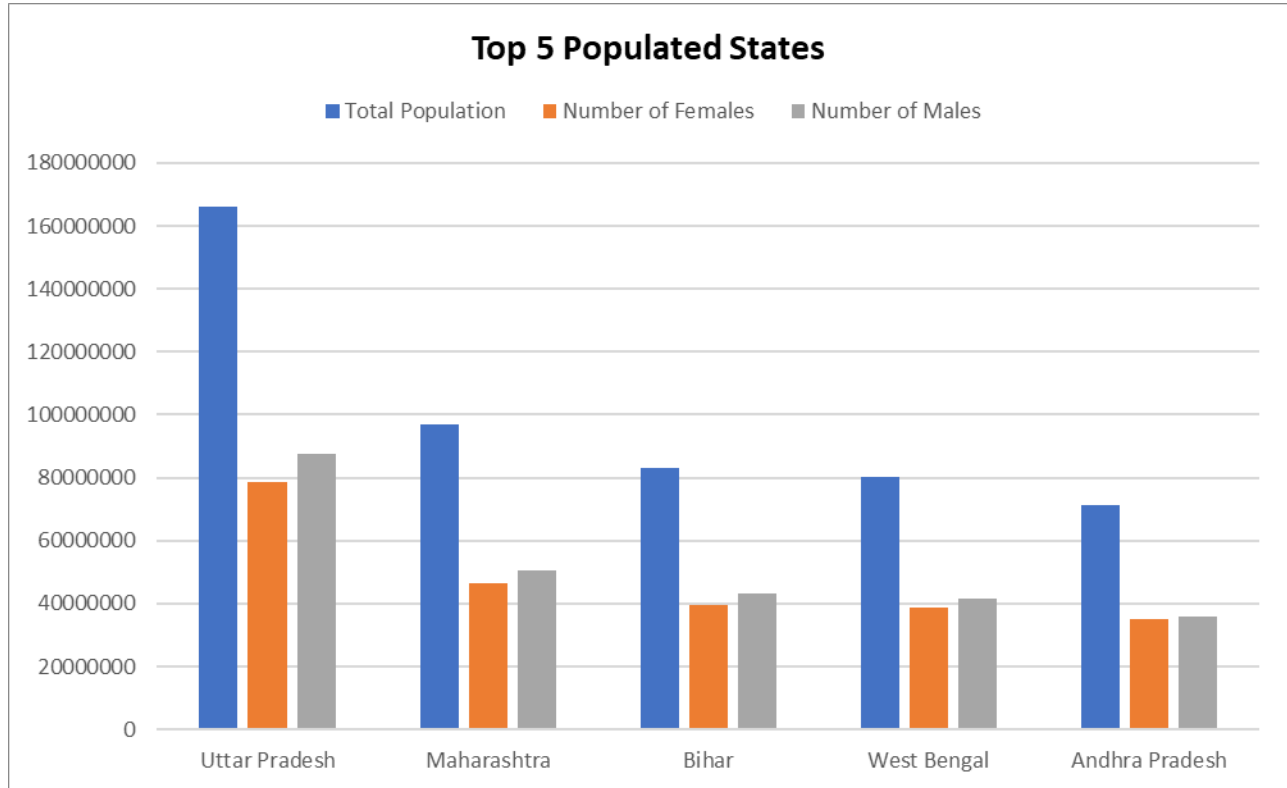
Specific function and requirements

We have to create a pivot table and slicer to determine the no of students of different categories in each states and districts, and then visualize it on graph.
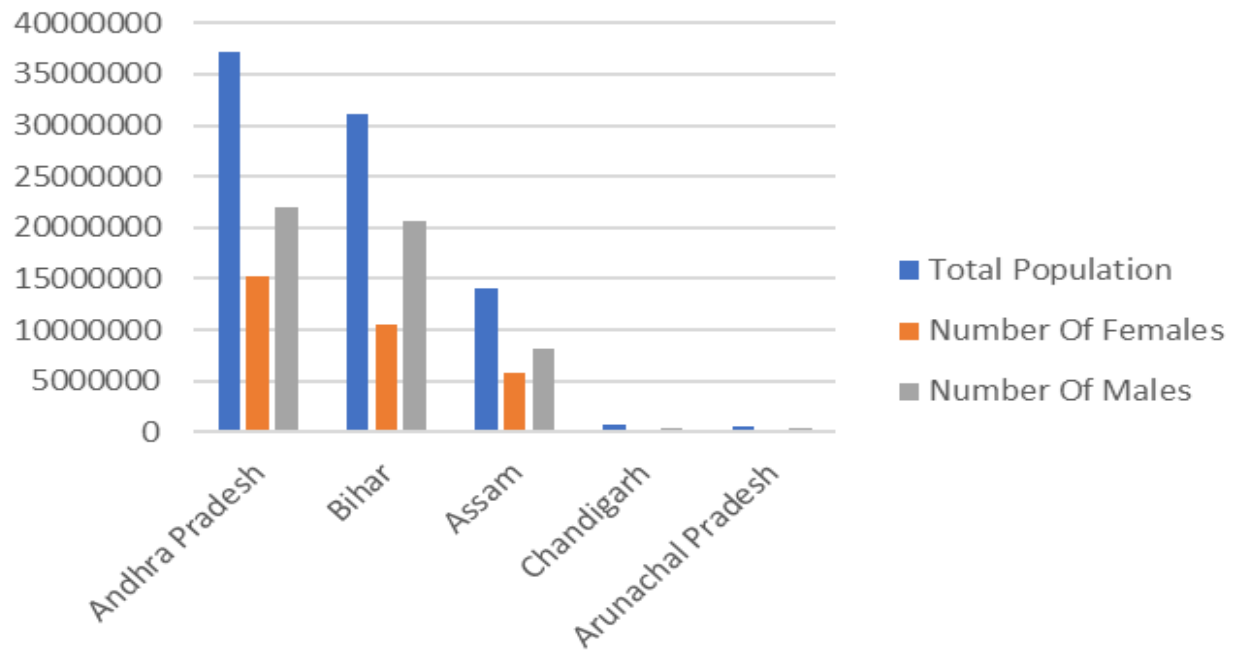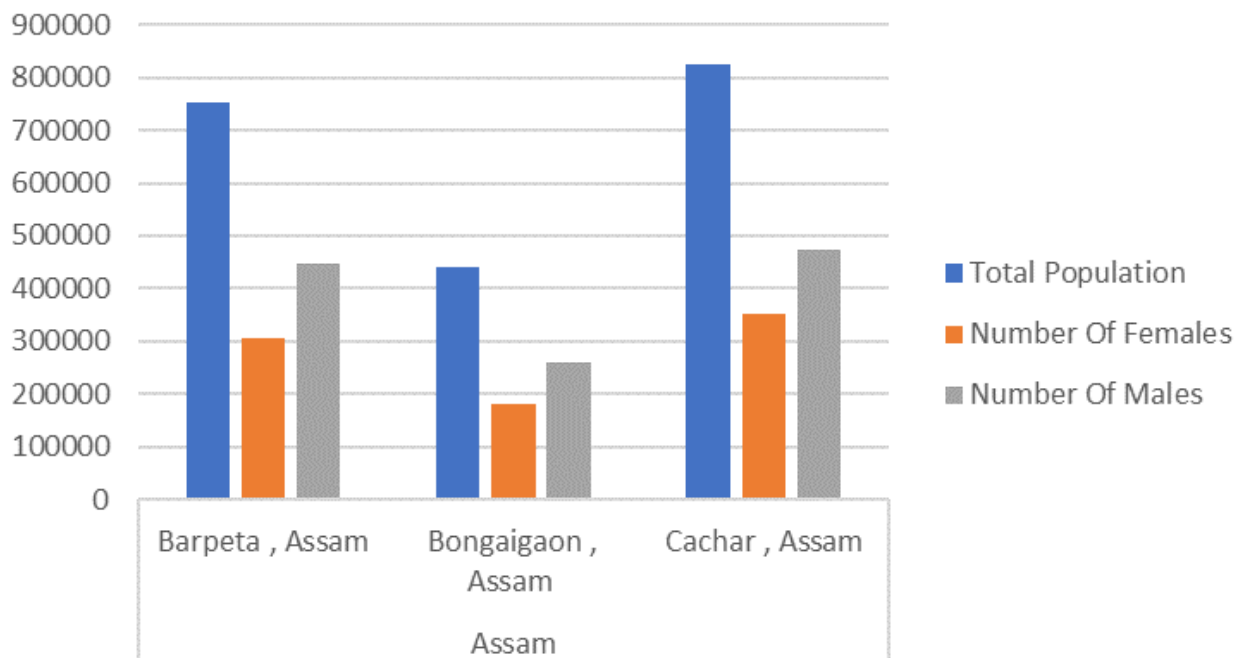
Results and Visualization:

ANALYSIS RESULTS

Top 5 Populated States

■ Total Population ■ Number of Females ■ Number of Males

Categories: Uttar Pradesh, Maharashtra, Bihar, West Bengal, Andhra Pradesh

(Y-axis: 0 to 180000000)

Top 5 Populated Districts

■ Total Population ■ Number of Females ■ Number of Males

Categories: Medinipur, West Bengal; North Twenty Four Parganas, West Bengal; Mumbai, Maharashtra; Thane, Maharashtra; Pune, Maharashtra

(Y-axis: 0 to 12000000)

Statewise Literacy

Total Population / Number Of Females / Number Of Males

Andhra Pradesh, Bihar, Assam, Chandigarh, Arunachal Pradesh

Districtwise Literacy

Total Population / Number Of Females / Number Of Males

Barpeta , Assam; Bongaigaon , Assam; Cachar , Assam

Assam

Statewise Literacy Rate

	Total
Andaman & Nicobar Islands	~77
Andhra Pradesh	~59
Arunachal Pradesh	~53

Districtwise Literacy Rate

	Total
Adilabad, Andhra Pradesh	~52
Anantapur, Andhra Pradesh	~56
Chittoor, Andhra Pradesh	~67
Cuddapah, Andhra Pradesh	~63

Districtwise Education

- Below Primary Education
- Primary Education
- Secondary Education Education
- Higher Education
- Graduation & Above

Statewise Education

- Below Primary Education
- Primary Education
- Secondary Education Education
- Higher Education
- Graduation & Above

Final Dashboard

Dashboard will be giving all the brief details of the work done on the project.

Reference and Bibliography

- Kaggle.com
- Wikipedia.com
- Google.com

Tools Used

- Tableau prep
- Microsoft Excel