SDG-4: Quality Education

Team- Edu-Force

Title- **Education in India: A Decade of Growth and impact of COVID-19**

ES1101: Computational Data Analysis

Submitted By Team 11(B) –

Vivek Pareek (2022BTech112)

Shrashti Chouhan (2022BTech098)

Aryan Lunawat (2022BTech021)

Shridarshan Mishra (2022BTech100)

Institute of Engineering and Technology (IET)
JK LAKSHMIPAT UNIVERSITY, JAIPUR

TABLE OF CONTENT

LIST OF TABLES

ABSTRACT

The goal of this study is to analyze the growth and development of education in India over the past few decades, focusing on the availability and quality of educational resources and infrastructure, as well as the growth of education for people with disabilities and the supply of qualified teachers in the country. Additionally, the study aims to evaluate the impact of COVID-19 on these factors and predict future trends in the education system. To achieve these objectives, we employed both inferential and descriptive statistics such as mean, standard deviation, and statistical tests like single mean tests, difference of means, linear regression, and correlation. Furthermore, hypothesis tests were used to compare means, whether they were intra-sample or not. Through the analysis, the objective of quality education is emphasized.

INTRODUCTION

Goal 4: Quality Education

Sustainable Development Goal (SDG) 4 is one of the 17 goals established by the United Nations in 2015 as part of the 2030 Agenda for Sustainable Development. The goal aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all."

The SDG 4 targets include:

1. increasing the number of children and young people who have access to quality early childhood development, care and pre-primary education.
2. increasing the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs, and entrepreneurship.
3. increasing the number of people who have access to affordable and quality technical, vocational, and tertiary education, including university.
4. eliminating gender disparities in education and ensuring equal access to affordable and quality technical, vocational, and tertiary education, including university.
5. increasing the number of adults who have relevant skills, including technical and vocational skills, for employment, decent jobs, and entrepreneurship.
6. increasing the number of people who have access to education, training, and lifelong learning opportunities throughout their lives.
7. increasing the number of teachers, including those from marginalized communities, trained to international standards.
8. increasing the number of people who benefit from international cooperation for education.

Ensuring access to quality education for all is seen as crucial for achieving other SDGs, such as reducing poverty and inequality, promoting sustainable economic growth and development, and fostering peaceful and inclusive societies.

In addition to the targets outlined above, achieving SDG 4 also requires addressing various challenges and barriers to access to education. These can include lack of infrastructure and resources, financial constraints, cultural and societal attitudes, and discrimination and marginalization of certain groups.

To achieve SDG 4, a range of actions are needed including:

1. Investing in education infrastructure and resources, including building schools, and training facilities, and providing teachers with the necessary resources and support.
2. Providing financial assistance and scholarships to support access to education for disadvantaged and marginalized groups, such as girls and children from low-income families.
3. Developing and implementing inclusive education policies and practices that consider the diverse needs of all learners and address discrimination and marginalization.
4. Promoting lifelong learning opportunities, including vocational training and adult education programs, to help people acquire the skills needed for the changing labor market.
5. Encourage international cooperation and sharing of best practices in education to support the sharing of knowledge, expertise, and resources to achieve SDG 4.

Overall, achieving SDG 4 requires a commitment from governments, international organizations, the private sector, and individuals to invest in education and work together to address the barriers and challenges to access to quality education for all.

LITERATURE REVIEW

Objective1:
The article provides on school enrolment in Africa, which shows that the region is performing poorly compared to other regions of the world in terms of achieving SDG 4.1, which aims to ensure that all girls and boys have access to free and quality primary and secondary education on an equal footing.

The data presented in the article shows that net primary enrolment rates and primary completion rates in Africa are lower than the average for developing countries and like those of countries affected by conflict and low-income countries. Additionally, the article notes that there is a strong selection-elimination pyramid in place in Africa, with the net primary school enrolment rate of 80% dropping to 66% in secondary 1, 43% in secondary 2 and 8% in higher education.

The article also highlights that Africa still has a long way to go before achieving target 4.1 of SDG 4, and that there is a need for efforts to be made to improve access to education in the region, particularly for disadvantaged and marginalized groups. This can be done through increased investments in education infrastructure, provision of financial assistance and scholarships, and implementation of inclusive education policies. *source: Article- Quality and Inclusive Education in the Context of Education for Sustainable Development in Nigeria on ijelict.acu.edu.ng*

The challenges facing education in Kenya as stated earlier include regional and gender disparities in access and participation, poverty, insecurity in some areas, negative cultural practices that affect girls, a lack of adequate secondary schools, low quality of education, high pupil-teacher ratios, and budget constraints that have led to a freeze on hiring civil servants. Additionally, there are also issues related to teacher training and deployment, low enrolment rates in some regions, and inadequate implementation and supervision of education programs. These challenges need to be addressed to improve the overall quality and accessibility of education in Kenya. Despite the introduction of free primary and secondary education in Kenya, there are still many challenges facing education in the country. These include regional and gender disparities in access and participation, poverty, insecurity in some areas, negative cultural practices

that affect girls, a lack of adequate secondary schools, low quality of education that have led to a freeze on hiring civil servants. Additionally, there are also issues related to teacher training and deployment, low enrolment rates in some regions, and inadequate implementation and supervision of education programs. These challenges need to be addressed to improve the overall quality and accessibility of education in Kenya. *Source: Article on Policies on Free Primary and Secondary Education In East Africa: Are Kenya And Tanzania On Course To Attain Education For All (Efa) Goals By 2015? from ir-library.ku.ac.ke.*

Objective 2:

The Centre for Space Science and Technology Education in Asia and the Pacific (CSSTEAP) is a training and research institution in India that is supported by the government's Department of Space. The centre provides comprehensive facilities for students, including expert teaching staff, modern teaching methods, and laboratory and fieldwork opportunities. The courses are taught using a combination of classroom lectures, video lectures, computer-based training, and laboratory experiments. Students also can visit national facilities and meet with experts in the field. The faculty at CSSTEAP is made up of scientists from various institutions and agencies, and the centre also hosts visiting international experts. The centre is in the cities of Dehradun and Ahmedabad and offers students access to excellent libraries, internet facilities, and hostel accommodations with modern amenities.

In addition to providing a high-quality education, CSSTEAP also offers a diverse and multicultural student community, as it attracts students from the Asia Pacific region. The centre provides a variety of opportunities for students to engage in both academic and extracurricular activities, and encourages them to develop new skills and explore their interests. The Central library of the centre possess the best collection of books in the field of space science, technology, and its application areas. The library also has access to e-journals, e-books, online databases, and other sources of open access information. With excellent internet facilities and dedicated computer resources, students have ample opportunity to explore online resources and conduct research. The centre also provides hostel facilities, with accommodation for up to 60 students, and offers a range of indoor and outdoor sports and recreational activities. Overall, CSSTEAP is a well-

equipped and well-staffed institution that provides students with a comprehensive education in space science and technology.

Source: CSSTEAP

In recent years, many virtual schools have emerged that offer online education as an alternative to traditional in-school education. In the United States, many states support individual online schools and district-level online courses, in which a significant portion of students' education can be delivered through the internet. However, according to the 2016 Annual Survey of Education Report (ASER), there are still significant challenges in terms of basic infrastructure in many schools in India. The report found that 3.5% of schools in India do not have toilet facilities, and only 68.7% of schools have usable toilet facilities. Additionally, the report found that the percentage of schools with a library has decreased from 78.1% in 2014 to 75.5% in 2016. On a positive note, the percentage of schools with separate girls' toilets has increased from 32.9% in 2010 to 61.9% in 2016. The report also found that 74.1% of schools had drinking water facilities, and 64.5% of schools had playgrounds. *Source-Wikipedia, Elsevier*

Objective:3

Differing combinations of structural factors (such as caste, gender, religion, poverty etc.) intersect with disability resulting in varied individual experiences, but the broad commonalities that shape the lives of people with disabilities in India transcend these divisions. Their lives are largely marked by poverty and marginalization from mainstream social processes. A recent study by the World Bank (2007), for example, noted that children with disability are five times more likely to be out of school than children belonging to scheduled castes or scheduled tribes (SC or ST). Moreover, when children with disability do attend school, they rarely progress beyond the primary level, leading ultimately to lower employment chances and long-term income poverty. *Source: UNESCO*

People with disabilities face barriers in both the external and internal environments. These barriers can make it difficult for children with disabilities and their parents to access education. It is crucial to address these barriers, which include accessible transportation and safe and accessible ramps from home to school. Governments can

support this effort by investing in the infrastructure of education and training institutions, as well as in the continuing professional development of teachers in relation to disabilities, safety, inclusive and effective education. This would enable them to provide the necessary support and accommodations for students with disabilities and ensure a safe and accessible learning environment for all. *Source– UNICEF*

Ensuring that all individuals have the confidence and belief that they can succeed in education, regardless of their personal characteristics or background, is essential. This can be achieved by making local education and training initiatives more accessible by catering to the needs of vulnerable groups and attracting staff who are qualified to work with these groups. Adequate funding for schools and training institutions that serve vulnerable populations, as well as providing training for teachers who work with these learners, is also crucial for success. This will ensure that everyone has the opportunity to succeed in their educational pursuits, regardless of their background or personal characteristics. *Source – researchgate*

Objective 4:

Teacher education is a vital program for enhancing the quality of school education. It focuses on equipping prospective teachers with the necessary skills and knowledge to be effective educators. As the social, cultural, economic, and political landscape continues to evolve, it is essential for teacher education to adapt and change to meet the changing needs of society. This will enable teachers to maintain high standards and be responsive to the evolving needs of their students and communities.

In recent years, India has made efforts to improve the supply of qualified teachers in the education system. This has been done through a variety of initiatives, such as:

The Sarva Shiksha Abhiyan (SSA) program, which aims to universalize elementary education and improve the quality of education by providing teacher training and resources. The National Council for Teacher Education (NCTE) has been established to regulate and standardize teacher education

in the country. The Right to Education Act (RTE) has made it mandatory for all teachers in primary and upper primary schools to possess the minimum qualification as prescribed by the NCTE. The National Policy on Education (NPE) has emphasized the need for in-service teacher education and training to improve the quality of teaching. The National Curriculum Framework (NCF) has emphasized the need to improve the quality of teachers by providing them with proper training and resources.

Despite these efforts, the education system in India still faces challenges in terms of providing a sufficient supply of qualified teachers. According to a report by the National University of Education Planning and Administration (NUEPA), there is a shortage of teachers in the country, especially in rural areas and for subjects like mathematics, science, and English. Additionally, the quality of teachers is also a concern, as many teachers in the country lack proper qualification and training.

To address these challenges, there is a need for further investment in teacher education and training programs, as well as efforts to improve working conditions and career advancement opportunities for teachers. Additionally, policies and programs that aim to improve teacher retention and motivation, especially in rural and underprivileged areas, may also by beneficial.

There is a decrease in the number of teachers by academic qualification after covid-19. However, the pandemic has had a significant impact on the education system, and many teachers have faced challenges and changes in their role and responsibilities. Some teachers may have left the profession due to the added stress and uncertainty caused by the pandemic, while others may have been laid off or furloughed due to budget cuts. Additionally, online and remote learning has also put pressure on teachers and made it difficult for some to maintain the same level of qualification.

In India the development of professional qualification for teachers has been a focus of the government in recent years. The National Council of Teacher Education (NCTE) and the University Grants Commission (UGC) have implemented various initiatives and programs to improve the quality of teacher education and professional development.

One such initiative is the introduction of the Bachelor of Education (B.Ed.) program as a mandatory qualification for teachers at the primary and secondary level. This program, which focuses on the development of

pedagogical skills and subject knowledge, is aimed at improving the quality of teaching and learning in schools.

Another initiative is the National Professional Development Program for Teachers (NPDPT) which provides opportunities for in-service teachers to upgrade their professional qualifications through online and offline courses. The program focuses on areas such as classroom teaching, assessment, and curriculum development.

In addition to these initiatives, the government has also introduced several schemes and programs to improve the quality of teacher's education institutions and to provide financial assistance to teachers for their professional development.

In recent years, the government of India has also taken steps to implement technology in teacher education. The use of digital technologies and online platforms to provide professional development opportunities for teachers is becoming increasingly common. This is expected to help bridge the gap between urban and rural areas in terms of access to professional development opportunities.

It is important to note that while these initiatives and programs have the potential to improve the professional qualifications of teachers in India, the success of these efforts will depend on various factors such as the availability of resources, the quality of implementation, and the willingness of teachers to participate in such programs. *Source: Writingcenter. Unc. Edu*

OBJECTIVES

Main Objective:

"The objective of this study is to analyse the growth and development of education in India over the past decades, with a focus on investigating the availability and quality of educational resources and infrastructure, as well as the growth of education for people with disabilities and the supply of qualified teachers in the country. The study also aims to examine the impact of COVID-19 on these factors and make predictions for future trends in the education system."

Objective 1:

To analyse the growth of education in India in last decades.

1.1 To analyse the average growth rate in the number of schools in India from 2000 to 2020 and forecast growth in the number of schools for the following five years.

1.2 To check below claim that gross enrolment is increased in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) education level in India during covid-19 pandemic based on previous year.

According to article of times of India, Total enrolment in schools (Class 1 to 12) has increased by 0.76 % in 2021-22 as compared to 2020-21. As per the report, the overall enrolment in school education from primary to higher secondary levels was slightly more than 25.57 crore. *Source: timesofindia.indiatimes.com*.

1.3 To check if there is mean difference between literacy rate of male and female in all states of India for year 2011 and 2001 is not equal.

Objective 2:

The study aims to investigate the educational facilities in India, including the availability and quality of resources and infrastructure in schools.

2.1 The purpose of the study is to examine the growth of internet facilities in schools after the outbreak of COVID-19 occurred in India.

2.2 To determine whether Covid-19 has had an impact on school libraries.

2.3 Compare the data of the educational facilities of year 2021(after covid) and year 2018(before covid-19).

Objective 3:

To analyze and study the education growth in terms of facilities and provisions made for disabled children (CWSN) in schools in India.

3.1 To determine the rank of the total number of enrolment of children with different types of disabilities in India of year 2021.

3.2 To predict the number of schools having adapted infrastructure for students with disabilities for the next 10 years by analyzing current infrastructure for students with disabilities from year 2014 to 2021 of different states of India.

3.3 To analyze the relation between the number of male and female teachers trained for teaching Children with special needs (CWSN) from 2014 to 2021 of India.

Objective 4:

To study the education growth in terms of supply of qualified teachers in India.

4.1 To check the claim whether there is a decrease in number of teachers by academic qualification after covid-19.

4.2 To compare the development of professional qualification of teachers in the last five years in India.

4.3 To predict the classes taught for the next Ten years by analyzing current situations from year 2017 to 2021.

DATA AND GRAPHS

Table 1.1: Total number of schools in India from 2000 to 2020

Year	Total Number of Schools
2000	994731
2001	1033758
2002	1069916
2003	1106621
2004	1142867
2005	1176789
2006	1220007
2007	1252858
2008	1288018
2009	1314868
2010	1342471
2011	1366068
2012	1389352
2013	1410484
2014	1430562
2015	1448846
2016	1469100
2017	1486064
2018	1499914
2019	1507387
2020	1507708

Source: *https://dashboard.udiseplus.gov.in/#/reportDashboard/sReport*

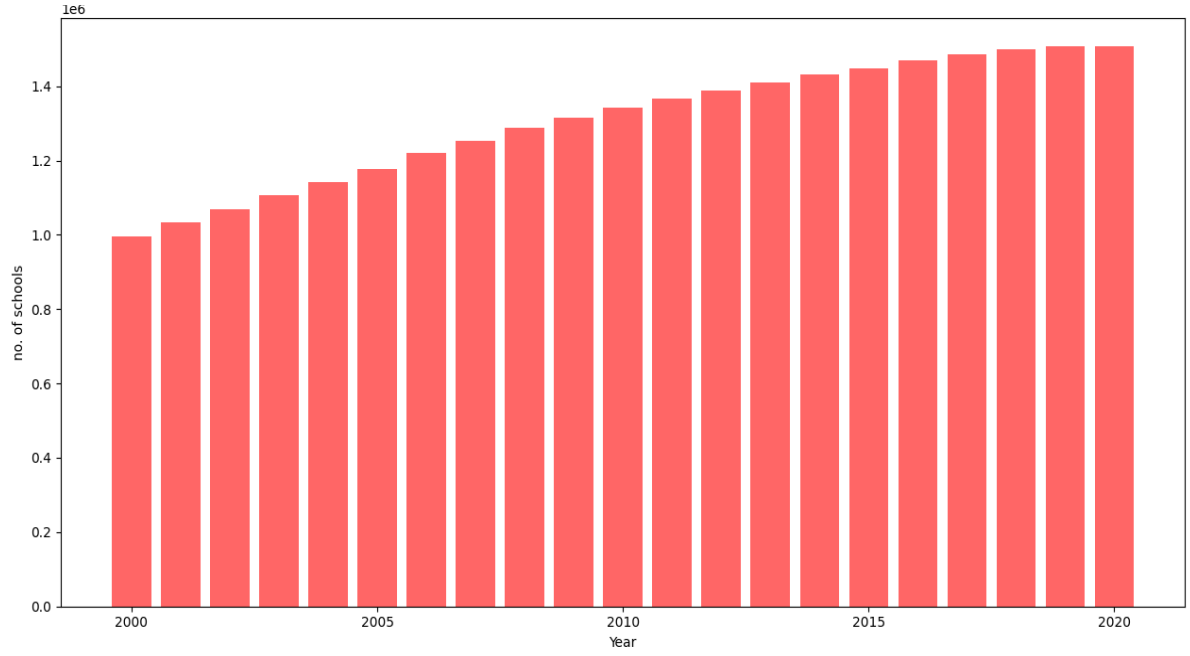


Table 1.2.1: Gross Enrolment Ratio in Elementary(I-VIII) from year 2014 to 2021

STATES	2014	2015	2016	2017	2018	2019	2020	2021
Andaman & Nicobar Islands	92.58	89.88	86.81	84.26	81.23	78.96	73.42	69.22
Andhra Pradesh	89.94	86.92	88.08	90.99	92.91	97.52	99.8	100.13
Arunachal Pradesh	119.23	118.68	102.94	104.35	101.64	105.19	108.18	109.87
Assam	107.12	99.63	101.91	99.94	104.47	107.38	110.32	109.86
Bihar	98.59	103.79	95.67	92.11	87.58	88.1	92.83	96.23
Chandigarh	102.32	98.96	96.51	97.32	94.16	91.61	87.55	88.51
Chhattisgarh	104.69	103.29	101.5	100.05	96.25	95.44	95.04	95.85
Delhi	120.94	122.46	122.58	120.72	120.19	121.26	119.32	121.15
Goa	106.3	104.55	101.54	100.9	98.21	96.36	92.49	91.1
Gujarat	98.09	98.21	98.03	96.37	94.43	93.79	92.19	92.36
Haryana	98.55	93.81	97.03	100.3	101.63	102.76	100.16	103.19
Himachal Pradesh	101.63	101.99	102.14	102.66	103.62	104.65	104.35	106.14
Jammu & Kashmir	82.34	81.83	74	78.33	79.55	81.9	85.4	90.13
Jharkhand	102.4	102.04	90.51	96.3	96.34	95.05	95.94	97.04
Karnataka	98.41	99.48	100.92	101.43	104.51	105.01	104.34	107.09
Kerala	97.48	97.33	97.19	97.4	97.88	99.14	99.38	101.02
Lakshadweep	80.91	76.49	72.36	70.64	69.86	71.1	70.61	73.23
Madhya Pradesh	105.08	100.42	97.93	94.99	92.05	92.35	90.39	88.66
Maharashtra	101.06	101.09	101.57	102.34	102.55	103.48	102.78	104.31
Manipur	107.95	107.23	99.11	105.89	112.15	116.07	119.58	117.58
Meghalaya	125.87	129.6	120.21	137.91	135.07	143.48	151.56	155.65
Mizoram	116.96	117.51	111.18	115.1	120.01	124.36	129.1	137.52
Nagaland	93.86	93.02	78.56	84.08	82.06	84.39	85.72	87.33
Odisha	99.73	99.89	98.92	96.56	93.9	94.04	95.77	95.36
Puducherry	97.48	94.8	92.96	88.97	86	84.65	82.84	77.03
Punjab	107.1	107.59	107.84	106.01	104.56	109.29	111.22	109.61
Rajasthan	95.65	98.96	98.27	99.21	98.05	99.7	101.6	101.78
Sikkim	114.53	109.29	98.78	100.04	93.79	93.56	93.43	92.89
Tamilnadu	103	104.27	104.3	98.3	98.26	98.01	97.9	98.75
Telangana	98.66	102.53	100.86	101.74	102.43	106.31	107.95	110.21
Tripura	106.68	105.69	102.13	104.66	105.32	107.26	107.41	109.11
Uttarakhand	102.48	104.33	103.85	105.72	106.29	108.42	109.92	113.15
Uttar Pradesh	97.41	98.07	95.18	94.13	91.58	95.11	96.64	98.07
West Bengal	99.52	100.48	94.59	98.32	99.29	103.61	109.5	108.45

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRe

Graph of table- 1.2.1(This graph shows that mean of state wise data of before covid (2014 to 2019) and after covid (2020 to 2021)

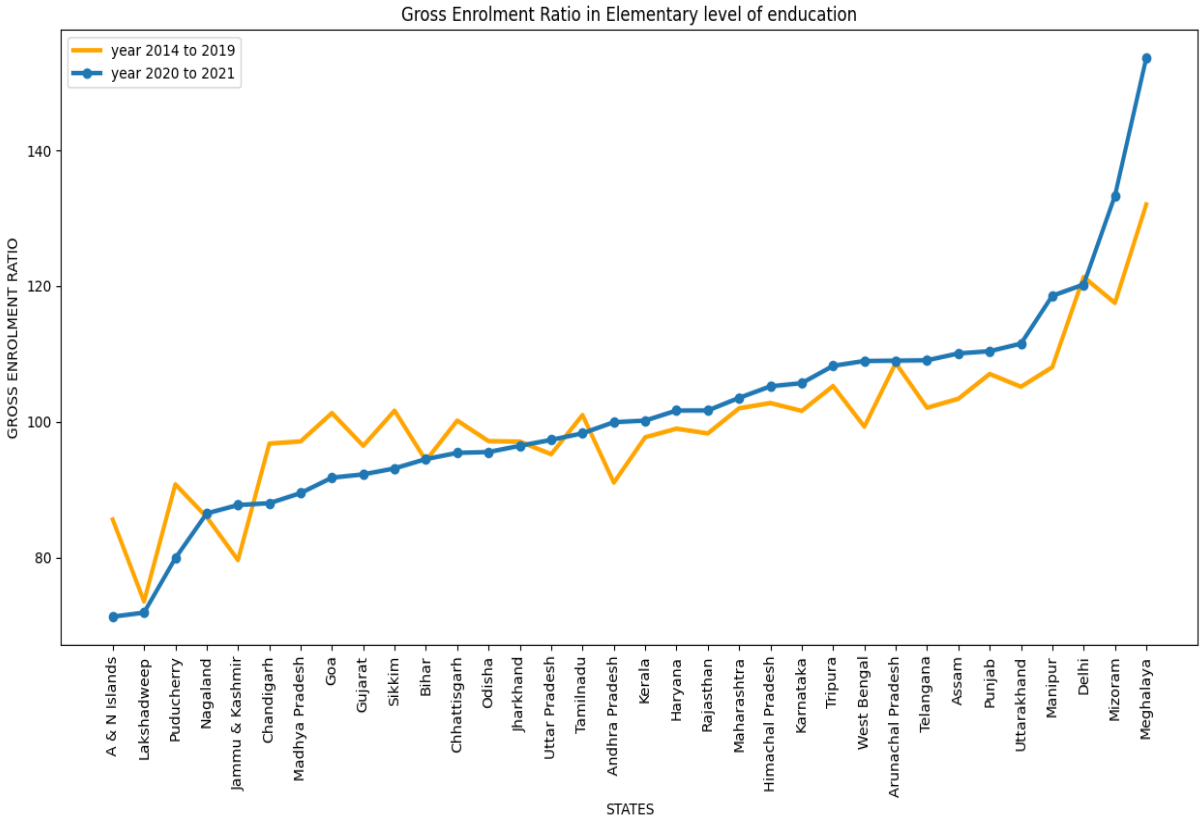


Table 1.2.2: Gross Enrolment Ratio in Secondary (IX-X) from year 2014 to 2021

STATES	2014	2015	2016	2017	2018	2019	2020	2021
A & N Islands	87.42	85.58	82.02	78.77	74.31	78.33	78.24	68.62
Andhra Pradesh	71.09	71.89	73.6	76.44	79.69	81.42	84.24	85.38
Arunachal Pradesh	75.64	76.74	73.42	72.63	68.68	62.32	68.18	66.51
Assam	70.13	71.35	71.68	70.51	72.76	74.04	75.59	74.48
Bihar	59.22	64.85	61.78	59.7	57.67	58.5	63.52	64.94
Chandigarh	91.55	95.2	96.74	95.43	92.71	87.55	89.67	90.06
Chhattisgarh	90.32	89.74	85.25	86.1	86.56	86.86	86.24	78.3
Delhi	97.71	103.51	110.75	105.1	110.39	110.31	116.33	111.24
Goa	105.56	103.05	97.14	95.81	91.86	90.39	91.07	82.96
Gujarat	72.71	72.31	72.69	75.27	76.1	77.43	78.56	75.16
Haryana	82.11	83.75	86.09	92.69	93.92	93.88	95.15	94.74
Himachal Pradesh	106.89	103.94	101.73	102.97	105.76	103.27	100.41	94.1
Jammu & Kashmir	63.25	62.52	57.07	60.6	59.34	58.66	59.83	60.54
Jharkhand	66.05	65.69	55.7	62.46	60.87	60.71	62.97	68.41
Karnataka	78.84	79.94	81.9	80.93	83.7	86.35	90.62	94.73
Kerala	103.72	103.47	100.77	99.13	97.97	97.16	97.55	97.85
Lakshadweep	115.76	97.79	95.36	83.59	77.28	63.1	77.73	63.32
Madhya Pradesh	78.84	79.43	78.64	79.71	77.42	75.5	71.3	69.95
Maharashtra	87.28	88.18	90.48	91.63	91.14	91.41	92.64	93.65
Manipur	68.93	69.82	64.77	69.31	69.88	73.3	75.66	75.99
Meghalaya	70.62	74.7	71.1	74.36	79.46	83.23	84.9	85.12
Mizoram	91.67	93.32	81.9	84.29	93.29	94.68	91.58	93.36
Nagaland	55.01	61.31	52.79	59.76	58.38	58.14	59.71	62.22
Odisha	74.74	76.33	76.68	77.07	80.68	79	84.52	80.36
Puducherry	99.99	94.28	91.69	87.1	83.12	80.3	78.88	76.06
Punjab	85.16	87.33	88.59	90.88	93.15	103.05	109.17	95.06
Rajasthan	73.61	73.61	73.81	76.77	81.6	84.21	84.82	79.23
Sikkim	95.88	102.47	95.58	96.26	99.73	99.4	89.96	89.07
Tamilnadu	93.32	95.75	96.81	89.43	90.46	90.54	92.59	95.59
Telangana	76.7	79.27	79.46	81.49	84.75	87.97	92.32	94.06
Tripura	105.81	101.41	95.87	92.09	90.16	89.47	78.95	81.25
Uttarakhand	90.37	88.44	87.36	86.53	89.01	91.54	91.64	89.85
Uttar Pradesh	68.78	69.67	69.46	65.9	65.39	65.79	66.41	69.26
West Bengal	75.26	79.75	75.77	81.63	81.54	86.25	91.18	88.2

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sReport

Graph of table-1.2.2: (This graph shows that mean of state wise data of before covid (2014 to 2019) and after covid (2020 to 2021)

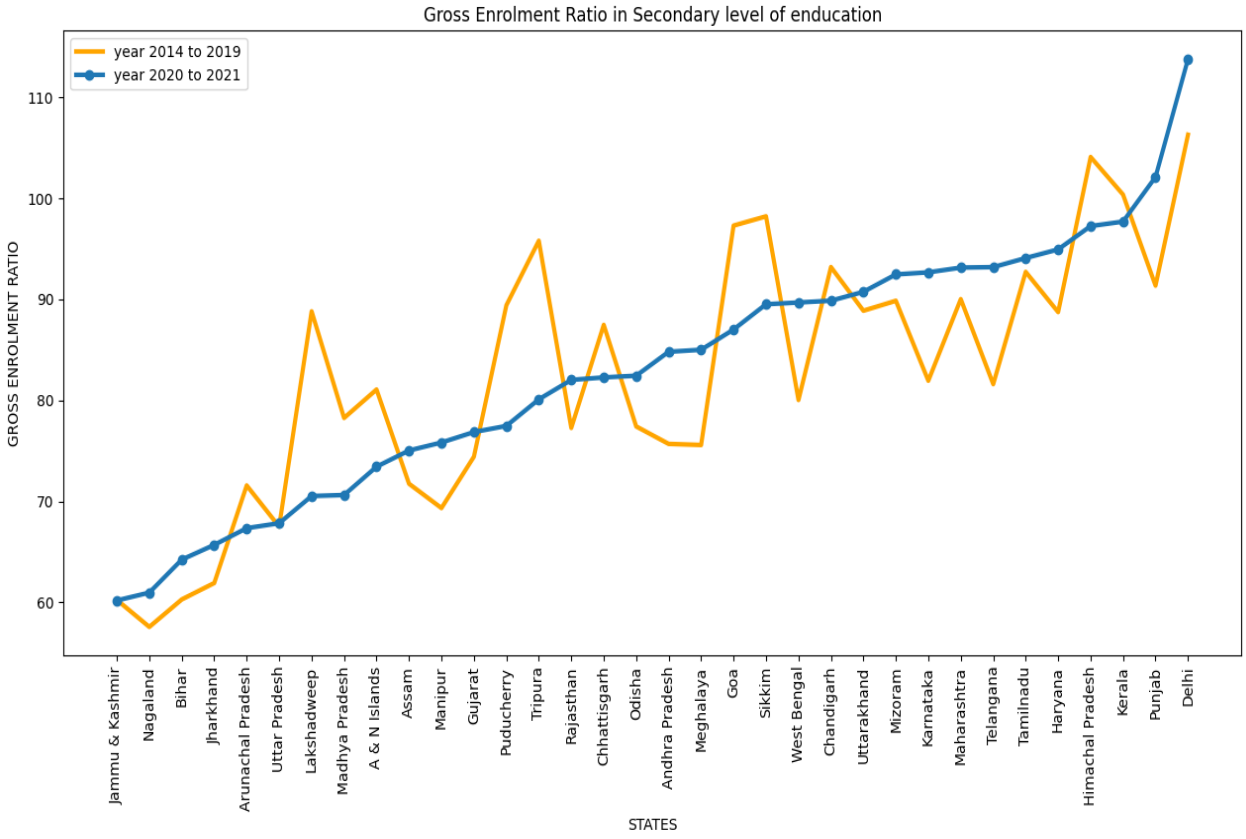


Table 1.2.3: Gross Enrolment Ratio in Higher Secondary (XI-XII) from year 2014 to 2021

STATES	2014	2015	2016	2017	2018	2019	2020	2021
Andaman & Nicobar Islands	79.43	70.93	67.86	66.48	55.55	45.13	49.88	65.84
Andhra Pradesh	44.54	51.18	0	40.18	46.88	52.19	53.37	56.7
Arunachal Pradesh	54.67	54.42	44.73	47.6	38.32	35.82	41.09	53.71
Assam	28.55	32.01	32.31	31.33	30.93	30.9	32.29	40.04
Bihar	22.38	24.46	18.97	20.55	26.39	30.8	34.01	35.88
Chandigarh	87.09	87.61	85.81	91.68	83.79	78.02	79.2	81.71
Chhattisgarh	50.61	49.53	49.37	49.19	52.08	54.3	57.62	68.11
Delhi	81.37	74.29	70.26	73.07	70.19	72.92	82.19	95.01
Goa	72.29	72.3	73.53	74.55	71.94	68.55	69.86	73.66
Gujarat	40.62	38.96	38.52	41.47	41.2	43.15	41.78	48.19
Haryana	58.95	55.38	56.38	55.21	56	61.74	66.79	75.54
Himachal Pradesh	86.37	89.85	86.92	85.97	81.79	83.43	85.55	94.08
Jammu & Kashmir	48.58	46.22	41.08	46.57	43.18	38.37	50.07	53.15
Jharkhand	37.65	36.38	27.1	41.2	38.89	40.79	43.95	46.44
Karnataka	29.68	35.79	1.18	45.05	44.4	52.13	55.64	56.6
Kerala	71.4	74.74	76.43	77	80.27	82.99	84.15	85.04
Lakshadweep	75.75	89.51	87.52	78.09	76.75	67.58	67.57	62.35
Madhya Pradesh	39.86	39.48	40.52	41.83	43.72	45.01	45.43	51.33
Maharashtra	58.6	64.08	67.02	68.04	68.91	66.97	68.19	71.48
Manipur	49.42	51.65	48.58	54.37	54.77	57.1	61.18	69.85
Meghalaya	31.88	38.12	35.42	38.91	41.7	43.36	41.13	45.96
Mizoram	48.34	48.89	47.6	46.46	52.13	53.7	54.06	61.3
Nagaland	29.07	32.04	31.7	31.4	33.96	34.02	33.67	35.83
Odisha	1.5	30.01	32.71	35.07	65.86	47.56	46.39	43.58
Puducherry	71.94	76.42	74.34	66.88	69.68	69.3	67.77	68.7
Punjab	65.01	65.51	68.23	66.17	68.17	71.28	77.76	82.02
Rajasthan	49.99	52.18	52.34	54.27	56.51	58.44	62.06	70.33
Sikkim	58.43	59.93	56	61.05	57.64	53.74	59.49	64.2
Tamilnadu	70.2	73.66	75.83	74.41	72.31	73.21	76.51	81.45
Telangana	48.2	50.31	41.64	51.41	56.52	57.18	61.88	64.84
Tripura	38.18	38.18	36.54	39.19	38.65	40.66	45.77	56.28
Uttarakhand	69.53	68.29	69.1	68.63	66.32	69.39	72.73	78.77
Uttar Pradesh	53.91	52.19	49.79	46.9	46.12	46.88	48.78	50.65
West Bengal	44.61	45.94	45.57	50.56	51.73	55.21	58.5	62

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table- 1.2.3(This graph shows that mean of state wise data of before covid (2014 to 2019) and after covid (2020 to 2021)

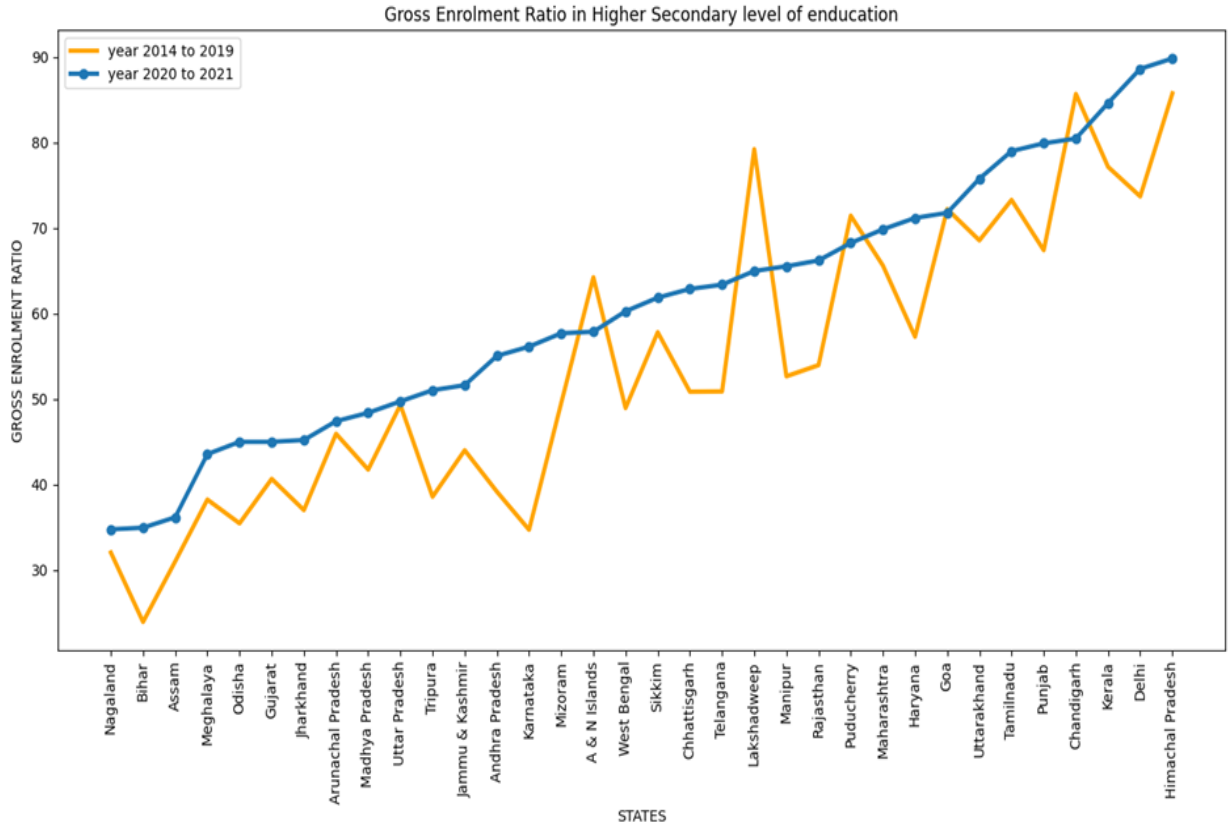
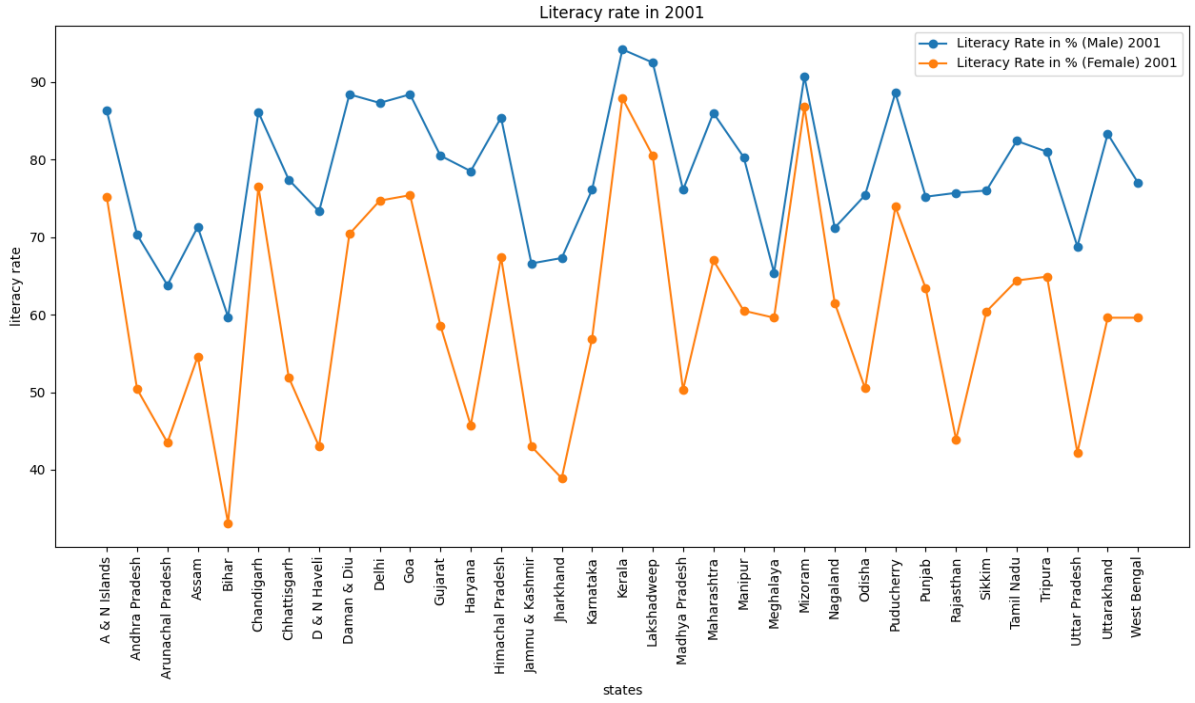


Table 1.3: Male and Female Literacy rate in India in year 2011 and 2021

State/Union Territory	Female_literacy_2001	Male_literacy_2001	Female_literacy_2011	Male_literacy_2011
A & N Islands	75.2	86.3	82.4	90.3
Andhra Pradesh	50.4	70.3	59.1	74.9
Arunachal Pradesh	43.5	63.8	57.7	72.6
Assam	54.6	71.3	66.3	77.8
Bihar	33.1	59.7	51.5	71.2
Chandigarh	76.5	86.1	81.2	90
Chhattisgarh	51.9	77.4	60.2	80.3
D & N Haveli	43	73.3	64.3	85.2
Daman & Diu	70.4	88.4	79.5	91.5
Delhi	74.7	87.3	80.8	90.9
Goa	75.4	88.4	84.7	92.6
Gujarat	58.6	80.5	69.7	85.8
Haryana	45.7	78.5	65.9	84.1
Himachal Pradesh	67.4	85.4	75.9	89.5
Jammu & Kashmir	43	66.6	56.4	76.8
Jharkhand	38.9	67.3	55.4	76.8
Karnataka	56.9	76.1	68.1	82.5
Kerala	87.9	94.2	92.1	96.1
Lakshadweep	80.5	92.5	87.9	95.6
Madhya Pradesh	50.3	76.1	59.2	78.7
Maharashtra	67	86	75.9	88.4
Manipur	60.5	80.3	72.4	86.1
Meghalaya	59.6	65.4	72.9	76
Mizoram	86.8	90.7	89.3	93.3
Nagaland	61.5	71.2	76.1	82.8
Odisha	50.5	75.4	64	81.6
Puducherry	73.9	88.6	80.7	91.3
Punjab	63.4	75.2	70.7	80.4
Rajasthan	43.9	75.7	52.1	79.2
Sikkim	60.4	76	75.6	86.6
Tamil Nadu	64.4	82.4	73.4	86.8
Tripura	64.9	81	82.7	91.5
Uttar Pradesh	42.2	68.8	57.2	77.3
Uttarakhand	59.6	83.3	70	87.4
West Bengal	59.6	77	70.5	81.7

Source: censusofindia

Graph of above table-1.3 (literacy rate of male and female in 2001)



Graph of above table-1.3 (literacy rate of male and female in 2011)

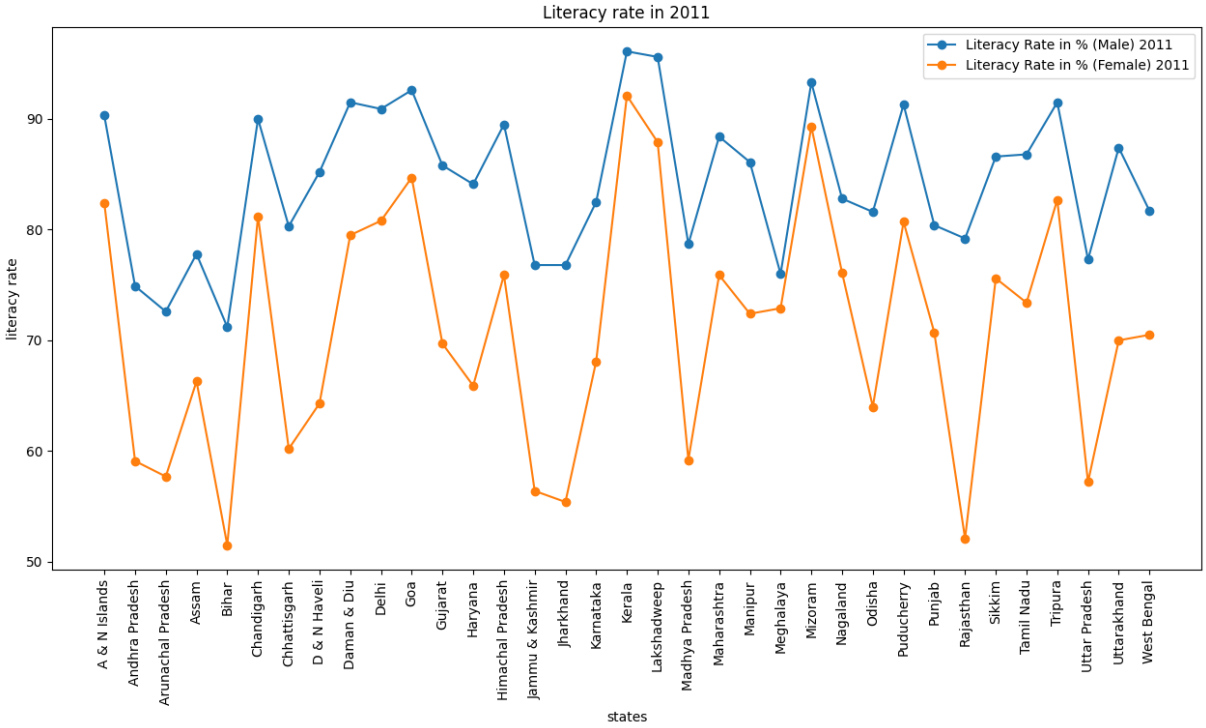
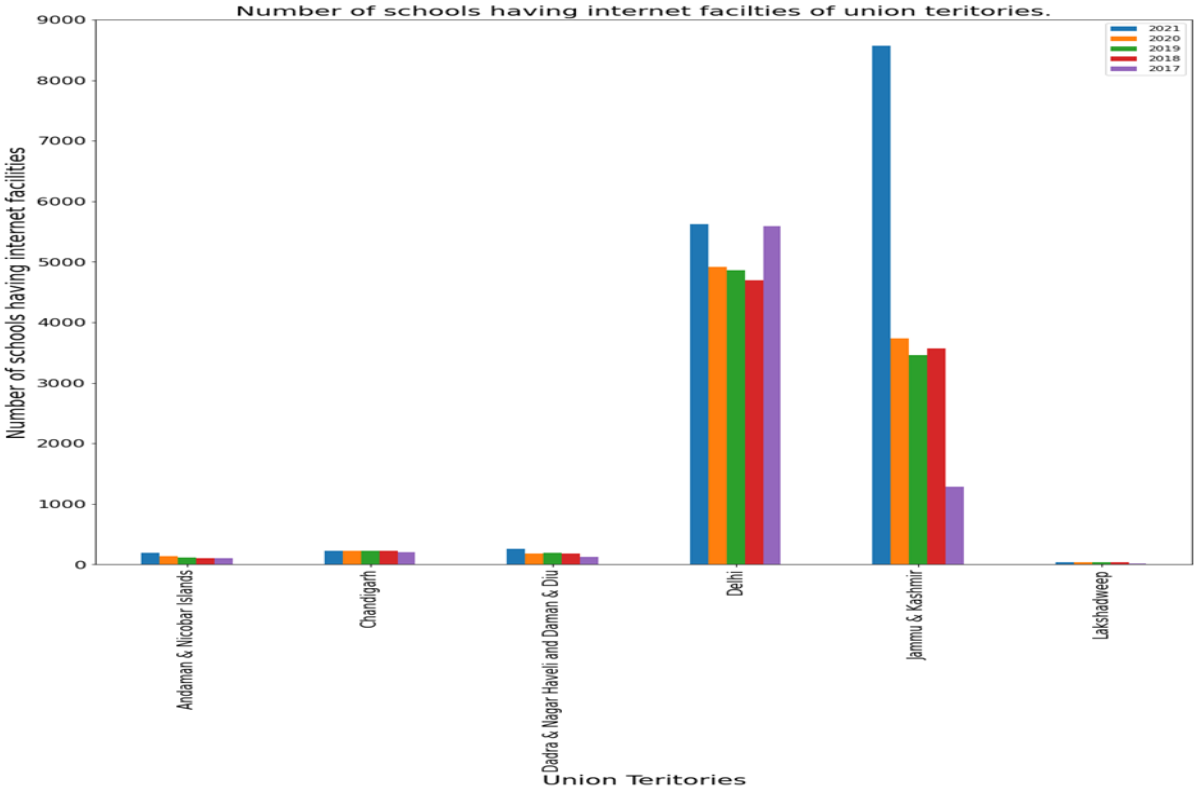


Table 2.1: Number of schools having internet facilities in India from year 2012 to 2021.

Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	total
Andaman & Nicobar Islands	187	140	113	107	106	73	76	78	77	73	414
Andhra Pradesh	34744	14975	14514	10954	7337	6410	8114	7887	14907	14639	62702
Arunachal Pradesh	794	350	307	272	235	164	164	158	145	118	4047
Assam	7126	4206	3836	3662	1604	1367	1227	895	562	458	71042
Bihar	10381	8359	7671	5991	4623	1049	1066	1099	1153	939	84236
Chandigarh	230	224	223	222	206	161	160	155	151	138	201
Chhattisgarh	20735	7922	4891	3078	1603	1605	1545	1313	1020	754	53781
Dadra & Nagar Haveli and Daman & Diu	264	176	188	177	124	89	82	71	62	51	492
Delhi	5619	4917	4858	4694	5590	3962	3789	2919	2048	1907	5755
Goa	879	577	575	601	624	483	471	447	418	318	1554
Gujarat	49522	40956	38658	36351	28924	8498	8244	7982	7950	6790	52424
Haryana	12158	10545	10172	9705	8701	8037	6720	6200	5055	3190	22315
Himachal Pradesh	6215	4216	4146	3938	3056	2260	2126	1946	1774	1664	18039
Jammu & Kashmir	8566	3740	3460	3570	1287	989	883	829	768	699	29092
Jharkhand	16726	14723	15001	13579	3695	1579	1074	999	874	734	48528
Karnataka	22590	19105	19020	9246	6804	6861	6841	6277	5530	4558	75489
Kerala	15459	14670	14638	14633	10832	4581	4502	4353	4091	3622	17130
Lakshadweep	37	42	42	38	14	14	13	13	13	12	45
Madhya Pradesh	34562	20115	18805	17386	15235	13366	11944	5735	8665	6472	150762
Maharashtra	52553	40259	39014	37870	30481	17480	16245	14898	12760	10156	107624
Manipur	1065	767	719	647	332	335	334	287	267	229	4993
Meghalaya	2460	645	572	550	356	342	332	255	173	145	14514
Mizoram	307	359	282	272	335	218	222	224	148	135	3825
Nagaland	1383	454	434	414	334	242	241	215	207	190	2826
Odisha	9284	4660	4328	4300	4318	2806	2331	1221	573	502	70300
Puducherry	724	719	484	492	457	387	321	297	276	248	731
Punjab	16429	27047	14087	13303	4833	10378	10099	9615	6763	6132	28988
Rajasthan	63674	42365	38561	31978	19445	16154	14548	13054	11574	9972	108428
Sikkim	434	267	247	241	216	100	100	96	82	77	1279
Tamilnadu	22086	18865	18819	14083	16907	11180	10883	10499	9996	8774	57583
Tripura	896	274	190	169	128	106	86	74	72	63	4844
Uttarakhand	6245	4185	3883	3663	2503	1393	1380	1365	1191	1026	24026
Uttar Pradesh	54554	38388	34638	29751	16097	11315	10327	9643	8545	5697	255969
West Bengal	15796	10442	9590	6754	7179	4968	4512	3932	2851	2011	95736

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-2.1(India's union territories graph)



Graph of table-2.1(Indian states graph)

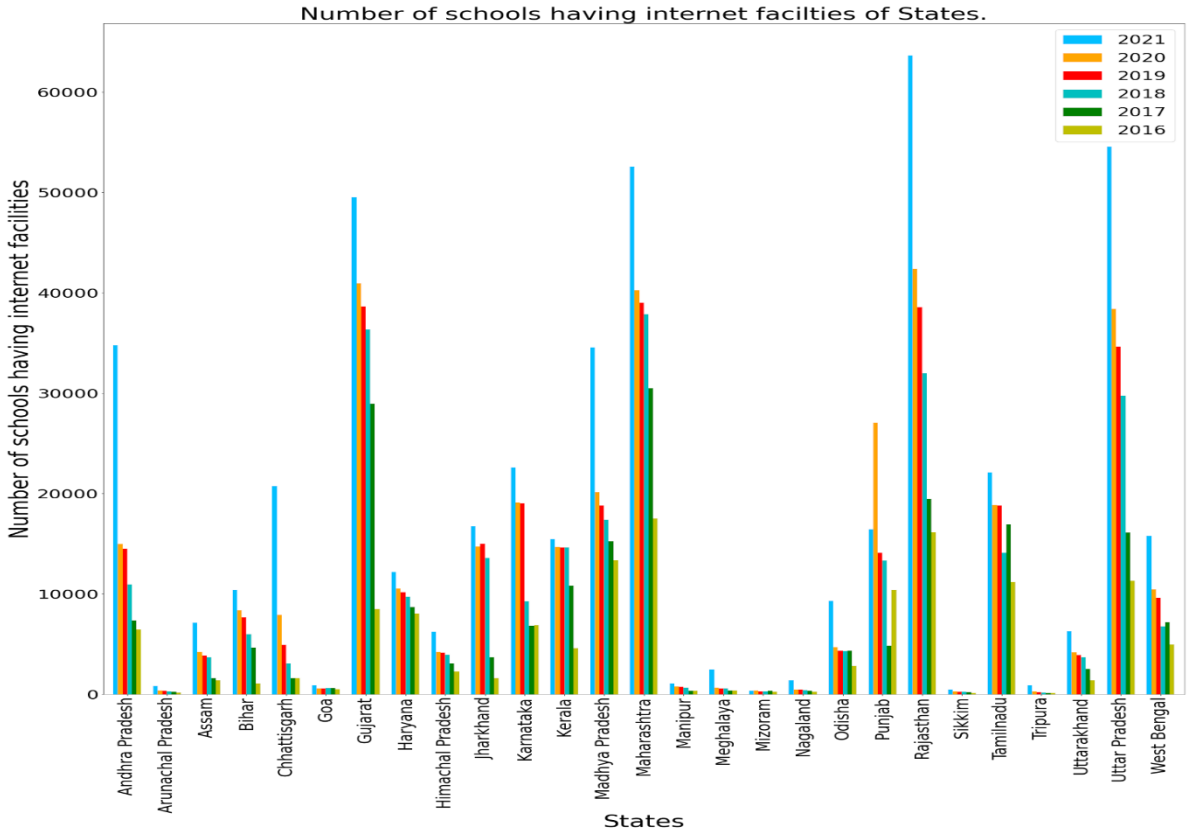
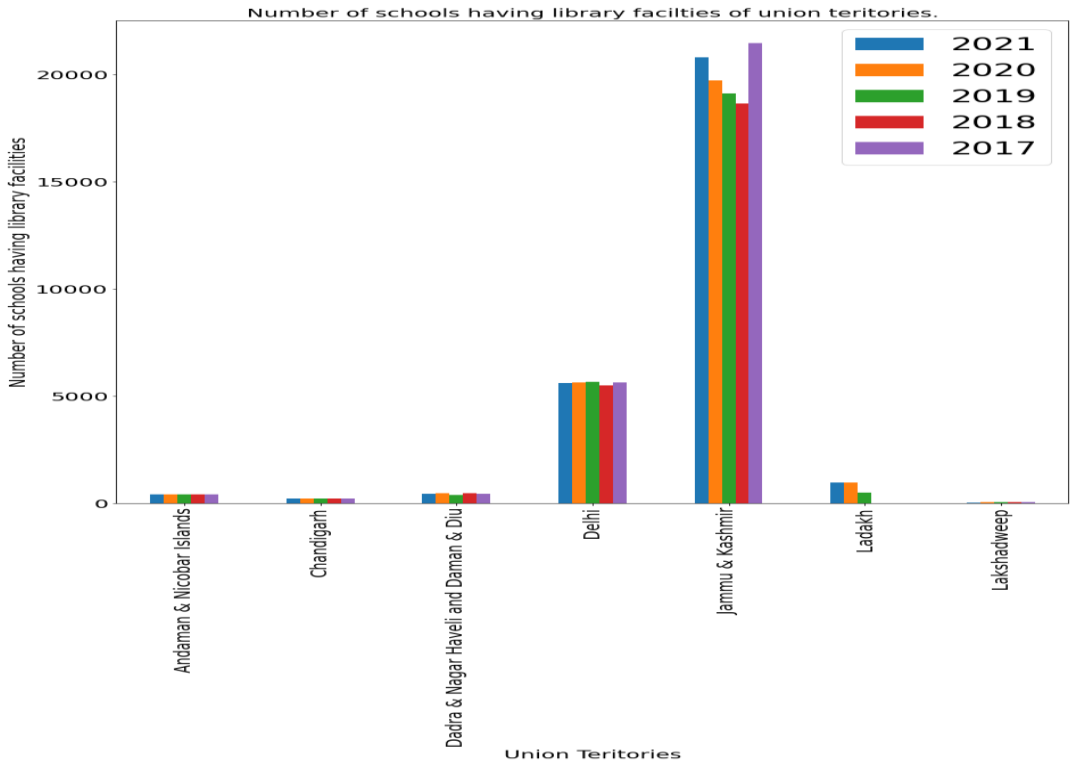


Table 2.2: Number of schools having library facilities in India from year 2012 to 2021.

Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Andaman & Nicobar Islands	414	415	416	407	410	402	399	393	396	357
Andhra Pradesh	60444	57218	56817	56173	24696	58912	60218	61691	98750	100171
Arunachal Pradesh	1621	1565	1554	914	1701	1153	1064	1012	745	424
Assam	53168	54859	52115	50244	48799	44216	41541	39480	31437	20979
Bihar	59021	56647	55699	52740	63301	59041	59019	58245	48608	42575
Chandigarh	229	226	223	220	224	197	199	196	190	182
Chhattisgarh	55489	54786	51598	48913	52118	50248	49795	51478	46975	45334
Dadra & Nagar Haveli and Daman & Diu	454	475	378	480	456	468	477	475	444	419
Delhi	5619	5642	5669	5492	5644	5612	5649	5626	5216	4623
Goa	1506	1468	1467	1468	1465	1538	1542	1527	1504	1445
Gujarat	51693	52157	52186	49984	51717	50461	48710	47746	46042	41757
Haryana	23026	22900	22721	22031	22593	22008	21632	21307	21574	18173
Himachal Pradesh	17294	17263	17148	16839	17431	17308	17263	17219	17056	15055
Jammu & Kashmir	20791	19703	19108	18644	21442	18761	18313	17035	13737	9176
Jharkhand	42119	42213	42119	42018	45312	43818	44564	43966	41655	36456
Karnataka	73678	73085	71657	70435	72705	73245	73829	73177	72328	70949
Kerala	15987	16098	16077	15885	16522	16585	16558	16452	15681	13812
Ladakh	954	980	505							
Lakshadweep	38	45	45	45	45	45	44	44	45	39
Madhya Pradesh	119074	124880	123847	142792	142553	137494	135706	133790	121113	85538
Maharashtra	107155	107242	106723	104989	106987	104474	103356	102451	98712	86587
Manipur	1187	1185	1156	1054	1431	1430	1382	1420	1240	1098
Meghalaya	3468	2078	1781	1282	3771	1797	1734	1660	1187	904
Mizoram	3230	3210	3152	3032	3130	3016	2959	2791	2588	1306
Nagaland	1966	1513	1081	1060	1310	1138	1155	1007	926	537
Odisha	58960	58868	63411	65045	65287	65545	64867	62634	57900	51579
Puducherry	733	741	741	739	731	736	730	730	722	677
Punjab	27577	28695	28775	28361	27730	28055	27895	27900	27807	25803
Rajasthan	82439	81082	79045	75802	82369	76995	77544	74029	72623	65514
Sikkim	1132	1139	1141	1142	986	776	772	603	454	342
Tamilnadu	58801	58896	58897	56210	56111	57632	57129	56688	55321	53982
Telangana	39510	39444	39382	39154	39720	39793	39117	42606	1445	548
Tripura	3493	3253	3003	2654	3288	2134	2094	1986		
Uttarakhand	20886	20782	20313	19364	22204	21628	21534	21348	20101	18831
Uttar Pradesh	203962	198365	186016	173456	215019	192940	190010	188625	181533	185041
West Bengal	82811	82459	81635	80650	76523	74304	72365	69489	54197	37696

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-2.2 (number of schools having library facilities of union territories)



Graph of table-2.2(number of schools having library facilities of states)

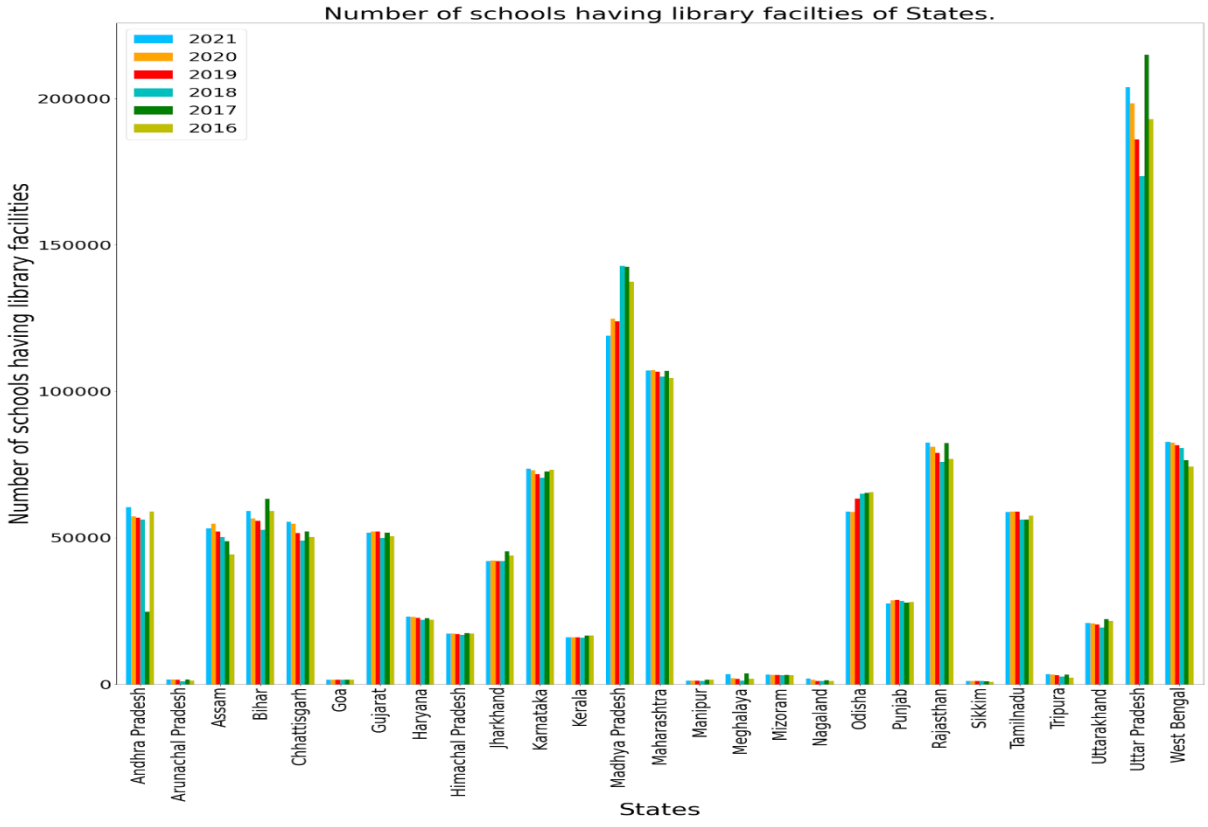


Table 2.3.1: Number of Schools by Availability of Infrastructure and Facilities in India in 2018-19.

Location	Library or Reading Corner or Book Bank	Functional Boy's Toilet	Functional Girl's Toilet	Drinking Water	cility(Drinking Water, Toilet and H	omplete Medical Checku	Internet	Computer Available
Andaman & Nicobar Islands	407	405	408	414	402	384	107	216
Andhra Pradesh	56173	38057	43209	56918	36430	46518	10954	18371
Arunachal Pradesh	914	2430	2467	2354	789	669	272	583
Assam	50244	45829	50844	58948	42430	32265	3662	9239
Bihar	52740	85471	86440	88832	71806	12412	5991	9803
Chandigarh	220	226	228	229	229	136	222	225
Chhattisgarh	48913	45653	54610	55853	51963	28997	3078	52402
Dadra & Nagar Haveli	340	344	339	346	338	341	75	252
Daman & Diu	140	133	132	140	137	115	102	130
Delhi	5492	4822	4731	5703	5701	2305	4694	5270
Goa	1468	1471	1473	1486	1485	557	601	725
Gujarat	49984	52519	53095	54573	50811	24704	36351	38341
Haryana	22031	21815	22362	23476	22883	12990	9705	11272
Himachal Pradesh	16839	17517	17700	18197	17666	12529	3938	5915
Jammu & Kashmir	18644	24086	24764	28460	24844	11052	3570	6451
Jharkhand	42018	42782	43611	44398	36057	10258	13579	25766
Karnataka	70435	69423	71976	77189	69125	36734	9246	71989
Kerla	15885	16158	16453	16626	16159	2824	14633	14682
Lakshadweep	45	45	45	45	45	38	38	41
Madhya Pradesh	142792	139773	142783	146947	124505	45699	17386	10102
Maharashtra	104989	98732	101752	108888	101145	55694	37870	76686
Manipur	1054	3280	3295	4338	2893	3115	647	1331
Meghalaya	1282	9212	8729	5531	2377	2432	550	1796
Mizoram	3032	3288	3267	3664	2267	2733	272	1825
Nagaland	1060	2415	2427	1747	995	770	414	1095
Orissa	65045	64772	66397	68717	65749	23296	4300	6763
Puducherry	739	701	709	739	725	660	492	659
Punjab	28361	26659	27164	28611	27042	21858	13303	17985
Rajasthan	75802	84111	87813	100528	82789	64303	31978	46946
Sikkim	1142	1266	1193	1288	1215	480	241	736
Tamilnadu	56210	55447	57303	57033	53287	34784	14083	30056
Telangana	39154	35970	39161	39688	33560	13048	7756	13862
Tripura	2654	3668	3764	3954	2926	2564	169	703
Uttarakhand	19364	20645	21240	21933	19927	5597	3663	7880
Uttar Pradesh	173456	247997	251780	264546	211279	105477	29751	34108
West Bengal	80650	92031	93637	93284	85041	8833	6754	11264

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Table 2.3.2: Number of Schools by Availability of Infrastructure and Facilities in India in 2020-21.

Location	Total No. of Schools	Library or Reading Corner or Boo	Functional Boy's Toilet	Functional Girl's Toilet	Drinking Water	l Facility(Drinking Water, Toilet and Hand	Complete Medical Checkup	Internet	Computer Available
Andaman & Nicobar Islands	416	414	413	414	416	413	244	187	267
Andhra Pradesh	61948	60444	51787	59277	61915	59639	50507	34744	29108
Arunachal Pradesh	3603	1621	2423	2475	2951	1390	814	794	1031
Assam	60859	53168	46556	50160	57129	46583	29768	7126	8622
Bihar	93165	59021	89002	90796	92911	79856	9030	10381	11830
Chandigarh	233	229	230	232	233	233	49	230	233
Chhattisgarh	56512	55489	51970	54324	56500	54679	24470	20735	34987
Dadra & Nagar Haveli and Daman & Diu	460	454	456	455	460	459	343	264	409
Delhi	5619	5619	4776	4693	5619	5619	1094	5619	5619
Goa	1510	1506	1496	1498	1510	1510	340	879	798
Gujarat	53851	51693	50669	51555	53836	51159	34937	49522	50379
Haryana	23726	23026	22018	22618	23670	23296	11425	12158	21596
Himachal Pradesh	18028	17294	17592	17741	18024	17758	2557	6215	6599
Jammu & Kashmir	28805	20791	22688	23728	28045	25107	7226	8566	8713
Jharkhand	44855	42119	42357	43287	44082	39729	9519	16726	26731
Karnataka	76450	73678	71403	74349	76230	67580	48952	22590	33919
Kerala	16240	15987	15863	16084	16218	16084	5901	15459	15845
Ladakh	978	954	909	879	973	612	287	418	403
Lakshadweep	38	38	38	38	38	38	28	37	37
Madhya Pradesh	125582	119074	116525	119032	123599	113260	42793	34562	27814
Maharashtra	109605	107155	99670	102703	109026	103475	52642	52553	81778
Manipur	4617	1187	3453	3475	4564	3269	776	1065	1401
Meghalaya	14600	3468	10668	10163	7743	3820	1966	2460	2315
Mizoram	3911	3230	3402	3396	3587	2498	971	307	971
Nagaland	2718	1966	2071	2095	1796	1115	219	1383	1344
Odisha	62291	58960	54591	56141	61566	57444	22938	9284	14776
Puducherry	736	733	703	710	736	733	461	724	720
Punjab	27701	27577	26819	27322	27699	27527	14445	16429	26521
Rajasthan	106373	82439	93591	98485	104188	95996	35332	63674	48494
Sikkim	1259	1132	1240	1160	1253	1220	666	434	1079
Tamilnadu	58801	58801	57628	58080	58801	58801	53548	22086	29193
Telangana	43083	39510	29137	33428	41105	32474	11476	9887	14898
Tripura	4929	3493	3571	3671	4137	3263	2936	896	1443
Uttarakhand	22815	20886	20087	20598	22118	20574	7949	6245	12299
Uttar Pradesh	258054	203962	245441	249626	255731	233102	86643	54554	50880
West Bengal	94744	82811	91838	93617	94261	92035	11454	15796	17416

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Table 3.1 -The total number of enrolment of children with different disabilities of year 2021 in India.

Category	2021
Blindness	114752
Low Vision	327517
Hearing impairment	228988
Speech and Language	197186
Locomotor Disability	330487
Mental illness	137577
Specific Learning Disabilities	314825
Cerebral palsy	71302
Autism Spectrum Disorder	20902
Multiple Disability incl. deaf, blindness	89978
Leprosy Cured students	3023
Dwarfism	11160
-Intellectual Disability	363289
Muscular Dystrophy	21247
Chronic Neurological conditions	5116
Multiple Sclerosis	6225
Thalassemia	5996
Haemophilia	1947
Sickle Cell disease	12708
Acid Attack victim	761
Parkinson's disease	1806

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-3.1.1 (the line chart shows the number of enrollment of different types of children with disabilities in year 2021 of India)

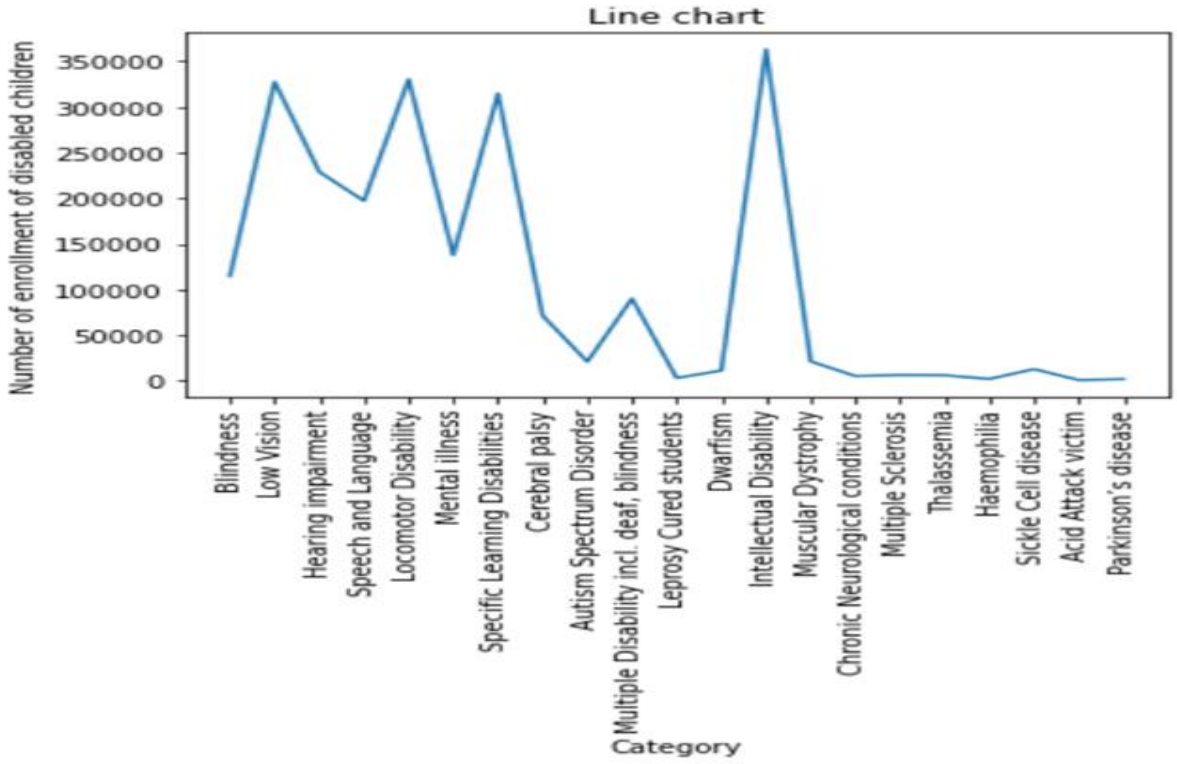


Table 3.2.1 - Number of schools having ramps for disabled students from 2014-2021 of different states of India.

Location	2021	2020	2019	2018	2017	2016	2015	2014
Andaman & Nicobar Islands	261	234	223	196	134	122	113	91
Andhra Pradesh	33306	27737	27331	24006	20634	15098	16263	20849
Arunachal Pradesh	884	781	781	776	937	918	885	833
Assam	42188	43656	38410	37291	38295	39675	38869	38270
Bihar	66179	65953	64276	62687	56368	55289	51676	45745
Chandigarh	204	200	171	163	198	141	120	108
Chhattisgarh	47177	47624	37485	35694	36704	35915	35203	36644
Daman & Diu and Dadra & Nagar Have	427	445	447	286	272	261	244	225
Delhi	5619	5642	5669	113	106	99	96	85
Goa	917	855	817	4240	4242	4170	4220	4176
Gujarat	44126	44471	44528	801	783	753	701	624
Haryana	17236	16581	16111	44129	43862	43017	41407	40774
Himachal Pradesh	13987	13282	12753	14836	14773	13985	13889	14009
Jammu & Kashmir	11334	10976	10413	12255	13913	13887	13397	12481
Jharkhand	28803	28832	29282	10120	7015	6488	6247	5566
Karnataka	53834	48763	37936	28574	29284	28183	22535	20499
Kerala	12675	12161	11610	37323	37067	37347	49136	48457
Ladakh	770	874	353	11284	10842	10523	10119	9922
Lakshadweep	36	45	42	41	28	28	27	27
Madhya Pradesh	100824	106987	99317	114405	105632	102699	100057	96531
Maharashtra	101637	100924	99626	97659	95200	92640	90656	88313
Manipur	2291	2370	2398	2316	2117	1852	1849	1870
Meghalaya	4436	3975	3707	3053	5017	4837	4705	4171
Mizoram	1732	1734	1664	1735	1704	1557	1590	1590
Nagaland	1029	1048	1034	963	871	813	791	614
Odisha	50128	52341	54522	47911	47040	46467	45067	42759
Puducherry	481	485	479	467	464	466	467	466
Punjab	24849	28695	28475	17894	17927	19958	19929	20075
Rajasthan	69459	66234	63812	59711	57758	52324	52355	49442
Sikkim	366	250	241	230	250	173	174	96
Tamilnadu	43520	47128	43364	37183	41154	38831	38029	36604
Telangana	32393	34867	36830	16371	14329	12998	12642	12946
Tripura	3105	2627	2575	2499	2432	2408	2407	2402
Uttarakhand	14072	13528	13183	12902	12163	11882	11747	11628
Uttar Pradesh	166323	162603	158262	180934	191245	189470	185483	184784
West Bengal	73187	72844	71837	67049	57585	56224	55474	52917
Total	1069795	1067752	1019964	988097	968345	941498	928569	906593

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-3.2.1 (stacked area chart shows number of schools having ramps from year 2014-2021 off different states of India)

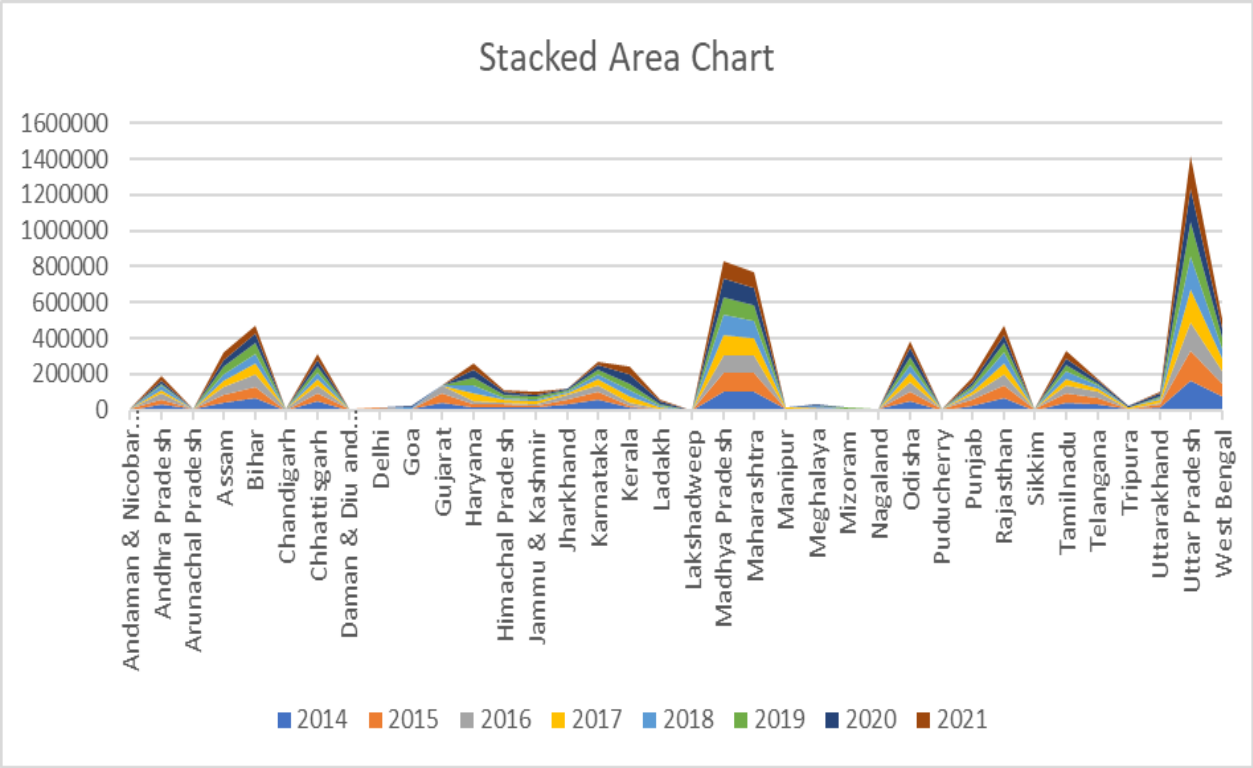


Table 3.3- the total number of male and female teachers trained to teach children with disabilities from 2014 to 2021 in India.

Year	No.of male teachers	No.of female teachers
2014	73569	51501
2015	74542	53492
2016	74886	54490
2017	75688	56997
2018	193760	193235
2019	152120	154942
2020	148198	152621
2021	2167727	1728878

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-3.1 (no. of male and no. of female teachers trained for teaching children with disabilities from 2014 to 2021 in India)

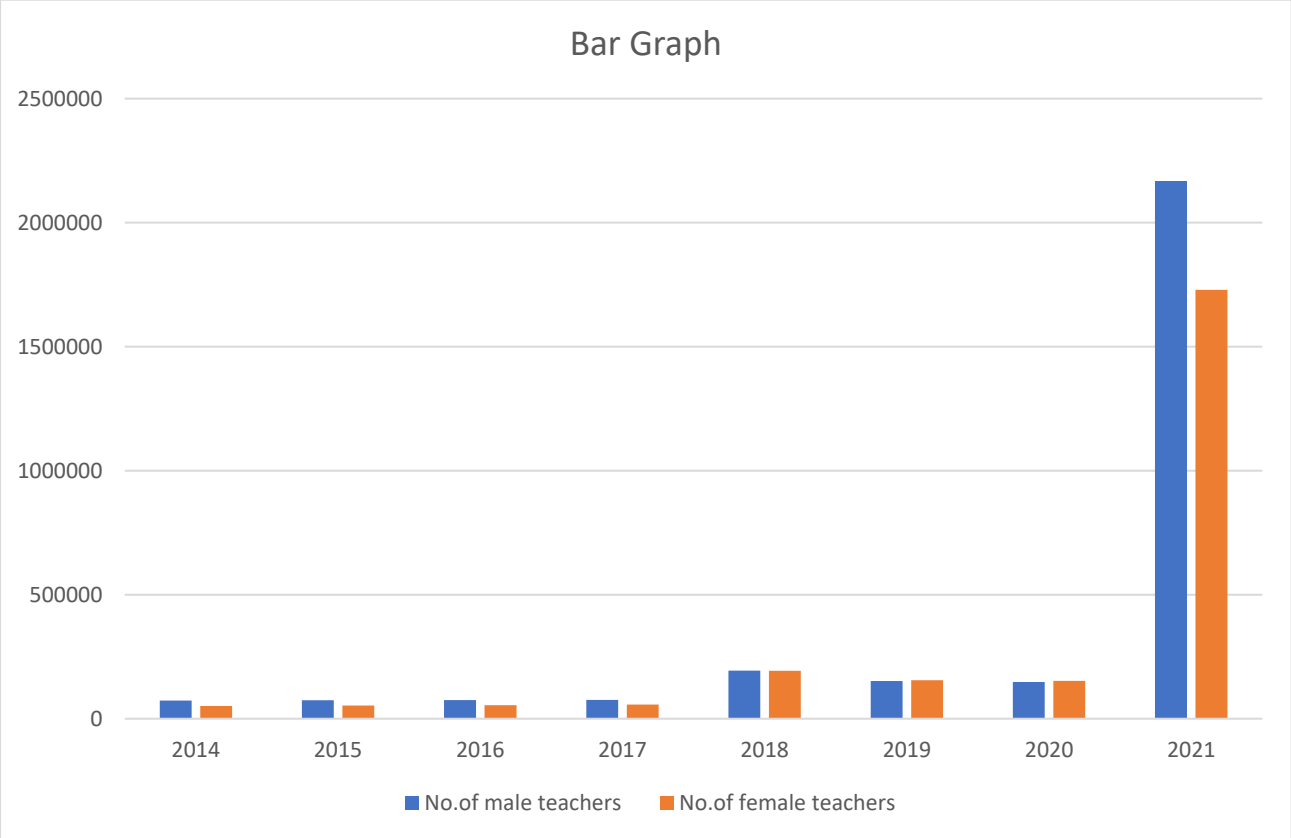


Table 4.1- Total Number of Teachers by Academic Qualification from year 2017 to 2021

states	total(year2021)	total(year 2020)	total(year 2019)	total(year 2018)	total(year 2017)
Andaman & Nicobar Islands	13288	13143	13746	24509	22504
Andhra Pradesh	860881	81274372	777706	1546606	1357108
Arunachal Pradesh	55767	59184	57645	109744	100244
Assam	825053	872000	864281	1741164	1572660
Bihar	1282514	1355994	1349389	2517608	2475984
Chandigarh	25476	27070	28168	54946	53936
Chhattisgarh	615135	639857	642095	1285028	1022594
Daman & Diu and Dadra & Nagar Haveli	14380	10060	6921	13136	12356
Delhi	303200	303866	3757	7258	7182
Goa	33613	32011	315294	622672	557032
Gujarat	801296	822627	31950	61992	53932
Haryana	613336	630859	830199	1634558	1630744
Himachal Pradesh	257690	263658	663923	1212774	1146436
Jammu & Kashmir	354001	364705	265005	496372	404056
Jharkhand	482142	521919	363388	733758	690124
Karnataka	1024353	1216147	519003	939820	932748
Kerala	552054	570353	1208400	1999742	1699761
Ladakh	1921	13899	570509	1050174	1094976
Lakshadweep	1337679	2330	13479	5358	5716
Madhya Pradesh	1658885	3070000000	3147	2345390	2446712
Maharashtra	102688	1715551	1314962	3333046	3073032
Manipur	118484	111427	1738114	213536	206176
Meghalaya	47892	120582	110004	234972	222160
Mizoram	70411	49217	120527	96148	94900
Nagaland	782860	72104	47798	138346	133724
Odisha	31812	916329	71073	1650580	1445160
Puducherry	594548	33874	892762	65968	65424
Punjab	1905691	572227	35800	1089432	1084972
Rajasthan	30217	2056173	576796	3956130	2757908
Sikkim	1401710	30919	2087227	61930	60528
Tamilnadu	837842	1466713	31579	2601568	2307372
Telangana	79361	873080	1444226	1424736	1347344
Tripura	287132	79377	833067	189978	195756
Uttarakhand	3294068	286319	95653	529372	517920
Uttar Pradesh	1707421	3289844	272839	5583644	4901331
West Bengal	9590	1689358	3117216	2523604	2500844

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Table 4.2- Total Number of Teachers by Professional Qualification from year 2017 to 2021

year	Diploma/degree in s	Diploma or certificate	Bachelor of Elementary	B.Ed. or equivale	None	Others	Pursuing any re	M.Ed. or e
2017	116732	2097218	712687	3809929	1700180	615193	41605	243923
2018	133741	2491290	380996	4705722	1591341	698853	148151	281323
2019	161764	2992296	426929	5369117	1368845	879532	131699	277893
2020	169219	3040632	403733	5530483	1274956	916206	123374	278195
2021	163745	2913972	399738	5701968	1069555	918939	107068	302025

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-4.2 (Total Number of Teachers by Professional Qualification from year 2017 to 2021)

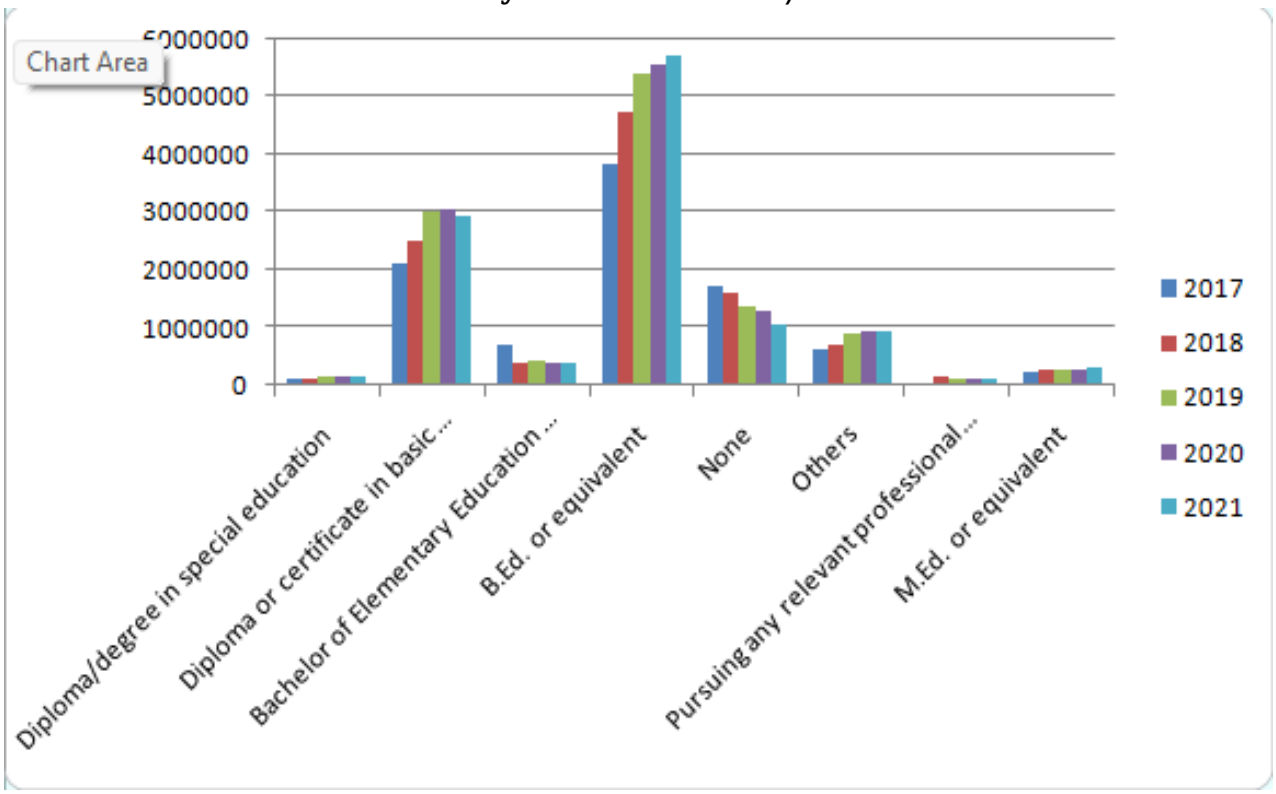
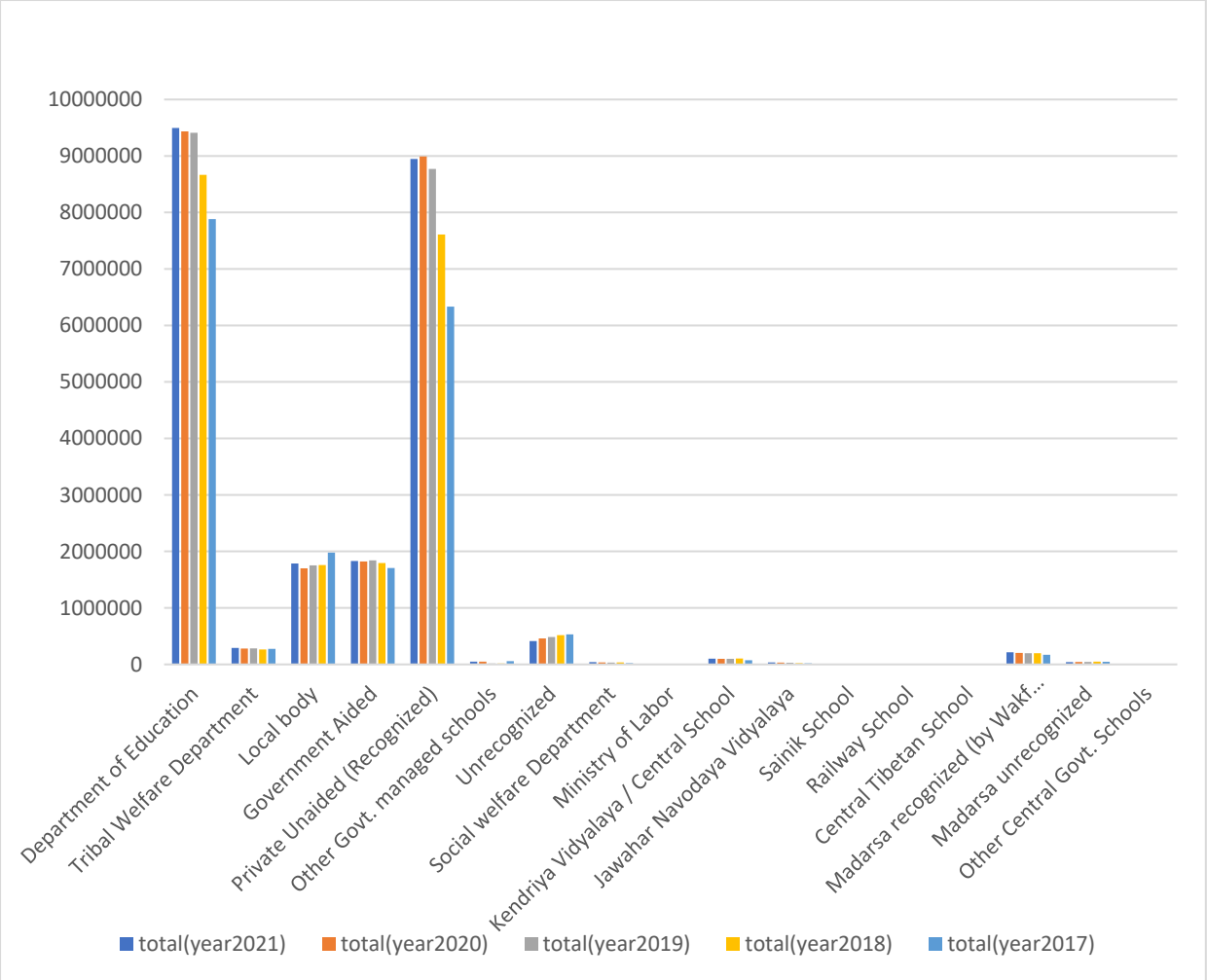


Table 4.3- Number of Teachers by School Management and Classes Taught from year 2017 to 2021.

MANAGEMENT NAME	total(year2021)	total(year2020)	total(year2019)	total(year2018)	total(year2017)
Department of Education	9495945	9435055	9409248	8665149	7881609
Tribal Welfare Department	291825	280558	284227	264616	274638
Local body	1785219	1701922	1754392	1759913	1977966
Government Aided	1829213	1823256	1840491	1795033	1706162
Private Unaided (Recognized)	8946210	8991805	8771248	7610719	6334138
Other Govt. managed schools	48709	47841	16351	17060	57746
Unrecognized	412035	460853	483437	517519	530660
Social welfare Department	40850	34930	32978	34338	20812
Ministry of Labor	734	1355	1448	1640	918
Kendriya Vidyalaya / Central School	101546	97678	99983	105521	73992
Jawahar Navodaya Vidyalaya	35714	31472	30166	22292	19188
Sainik School	3818	4188	4332	4174	4204
Railway School	4204	4282	3764	3292	2828
Central Tibetan School	943	969	804	712	354
Madarsa recognized (by Wakf board/Madarsa Board)	214026	203419	199240	197691	170464
Madarsa unrecognized	41128	43730	43181	48175	45358
Other Central Govt. Schools	7770	8320	6476		

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph of table-4.3 (Number of Teachers by School Management and Classes Taught from year 2017 to 2021.)

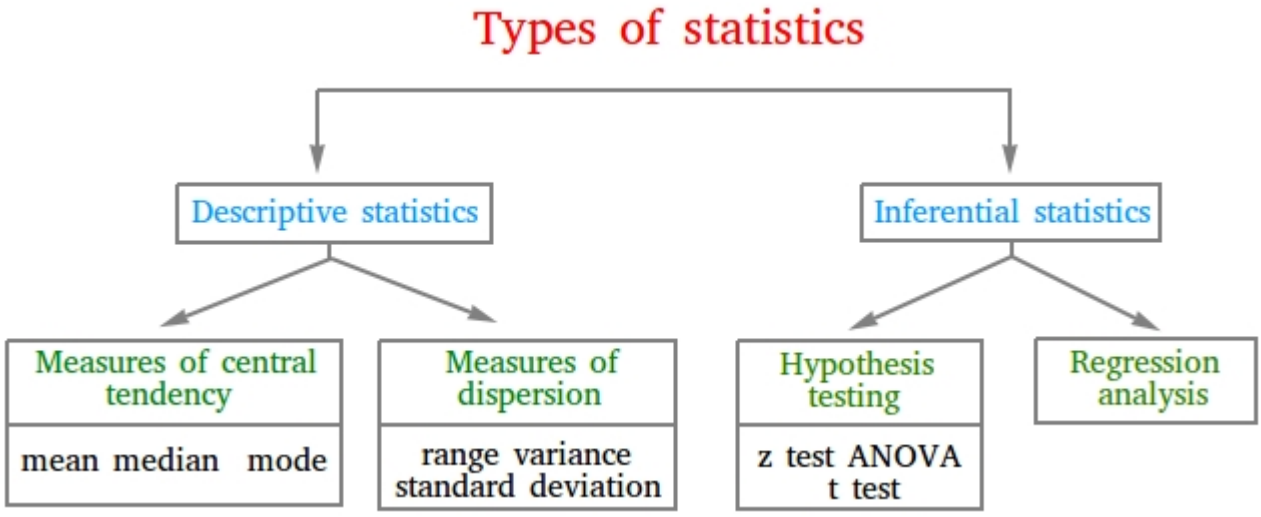


METHODOLOGY

STATISTICS:

Statistics is the field of study that deals with the collection, analysis, interpretation, and organization of numerical data. It is a science that helps to process and make sense of large amounts of data. Using statistics, one can measure, control and communicate the level of uncertainty associated with the data. It allows for making inferences about a larger population based on a smaller sample. In simpler terms, statistics is used to describe the characteristics of a sample and is typically used to estimate the values of parameters in a population.

TYPES OF STATISTICS:



- **DESCRIPTIVE STATISTICS:**

Descriptive statistics is a way to summarize and present data. It is performed on a sample of a population using various measures such as the mean or standard deviation. The purpose of descriptive statistics is to organize, represent, and explain a set of data by using charts, graphs, and summary measures. This can be done in various ways such as using tables, histograms, pie charts, bar graphs, or scatter plots. Descriptive statistics is simply a method of describing the data, it does not involve any normalization beyond the data that was collected.

I. MEASURE OF CENTRAL TENDENCY:

The mode, median, and mean are the three main measures of central tendency. Each of these measures describes a different indication of the distribution's typical or central value.

1. **Mean**: The arithmetic average of a Data set is calculated by adding the numbers in the set and dividing by the number of observations in the Data set.
   For n number of observations, we have

$$\text{Mean} = \frac{\text{Sum of Observations}}{\text{Total Number of Observations}}$$

2. **Median**: The Median is the middle number in the Data set when listed in ascending or descending order.
   For n number of observations, we have

$$\text{If 'n' is odd:} \quad \text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{term}$$

$$\text{If 'n' is even:} \quad \text{Median} = \frac{\left(\frac{n}{2}\right)^{th}\text{term} + \left(\frac{n}{2}+1\right)^{th}\text{term}}{2}$$

3. **Mode**: The Mode is the number that appears the most in a Data set and ranges between the highest and lowest value.
   For n number of observations, we have

$$\text{Mode} = L + h\,\frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)}$$

II. MEASURES OF DISPERSION:

The measures of central tendency are insufficient to describe all the information about a given dataset. As a result, the variability is described by a number known as the measure of dispersion.

1. **Variance:** To calculate the variance, subtract the mean from each data point in the set, square each one, add each square, and then divide the total number of values in the data set.

	Sample Variance
Ungrouped Data	$\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n - 1}$
Grouped Data	$\frac{\sum_{i=1}^{n} f(m_i - \overline{x})^2}{N - 1}$

2. **Standard Deviation:** The square root of Variance is used to calculate Standard Deviation. To calculate the standard deviation of any data, first calculate the variance. The standard deviation is thought to be the best measure of dispersion.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}$$

- **INFERENTIAL STATISTICS:**

We attempt to interpret the meaning of descriptive statistics in Inferential Statistics. We use Inferential Statistics to describe the Meaning of the collected Data after it has been collected, analysed, and summarised.

I. Hypothesis Testing:

Hypothesis is an assumption that we make about the population parameters. Hypothesis testing is a way for testing a claim about a particular parameter in a population, using the sample data. It consists of mainly two attributes which are- Null and Alternate hypothesis. We validate our assumptions on the basis of these two assumptions.

1. **The null hypothesis (H0)-** It is a statistical hypothesis that states that there is no difference between a parameter and a specific value. (Always has the = sign)

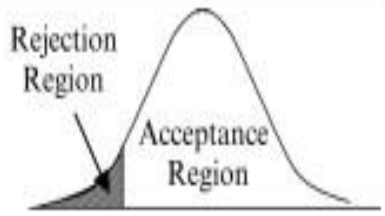
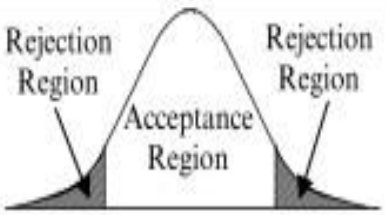
2. **Alternative Hypothesis (H1) -** A statistical hypothesis indicating that there is a difference between a parameter and a specific value.

3. **Level of significance (α) -** The maximum probability of making an error in which the null hypothesis is rejected when it is true.

4. **The critical or rejection region-** It depicts the range of values for the test value that include the significant difference and infer that the null hypothesis should be rejected.

5. **Non-critical or Non-rejection Region -**It depicts the range of values for the test value that include the significant difference and infer that the null hypothesis should not be rejected.

6. **Critical Value:** This value separates the critical region from the non-critical.

One-Tailed Test (Left Tail)	Two-Tailed Test	One-Tailed Test (Right Tail)
$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X < \mu_0$	$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X \neq \mu_0$	$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X > \mu_0$
		

- **T- Test:** It is a parametric statistical test used to compare and infer whether the means of two populations differ or not when the standard deviation is unknown.

Types of t-tests:

Test	Number of Variables	Type of Variable	Purpose of Test	Estimate of population mean	Population standard deviation	Degrees of freedom
One-sample t-test	One	Continuous measurement	Decide if the population mean is equal to a specific value or not	Sample average	Unknown, use sample standard deviation	n-1
Two-sample t-test	Two	Continuous measurement, Categorical or Nominal to define groups	Decide if the population means for two different groups are equal or not	Sample average for each group	Unknown, use sample standard deviations for each group	n1 + n2 – 2
Paired t-test	Two	Continuous measurement, Categorical or Nominal to define pairing within group	Decide if the difference between paired measurements for a population is zero or not	Sample average of the differences in paired measurements	Unknown, use sample standard deviation of differences in paired measurements	n-1

- **Z-Test:** A z-test is a statistical method used to evaluate if the means of two populations are equal. It is used when the population has a normal distribution, the data points are independent, and the sample size is at least 30. When the population variance is known, it is used to calculate the z test statistic. If the calculated z test statistic is statistically significant when compared to the critical value, it means that the null hypothesis of the test can be rejected.
- **One sample z-test:**
  When the population standard deviation is known, a one-sample z test can be used to determine if there is a difference between the sample mean and the population mean. The formula for the z test statistic is:

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$\bar{x}$	= sample mean
μ	= population mean
σ	=population standard deviation
n	=sample size

The process to conduct a one-sample z test, utilizing the z test statistic, is as follows:

Left Tailed Test:

Null Hypothesis: H0: μ=μ0=0

Alternate Hypothesis: H1: μ<μ0<0

Decision Criteria: If the z statistic < z ,critical value then reject the null hypothesis.

Right Tailed Test:

Null Hypothesis: H0: μ=μ0=0

Alternate Hypothesis: H1: μ>μ0>0

Decision Criteria: If the z statistic > z critical value then rejects the null hypothesis.

Two Tailed Test:

Null Hypothesis: H0: μ=μ0=0

Alternate Hypothesis: H1: μ≠μ0≠0

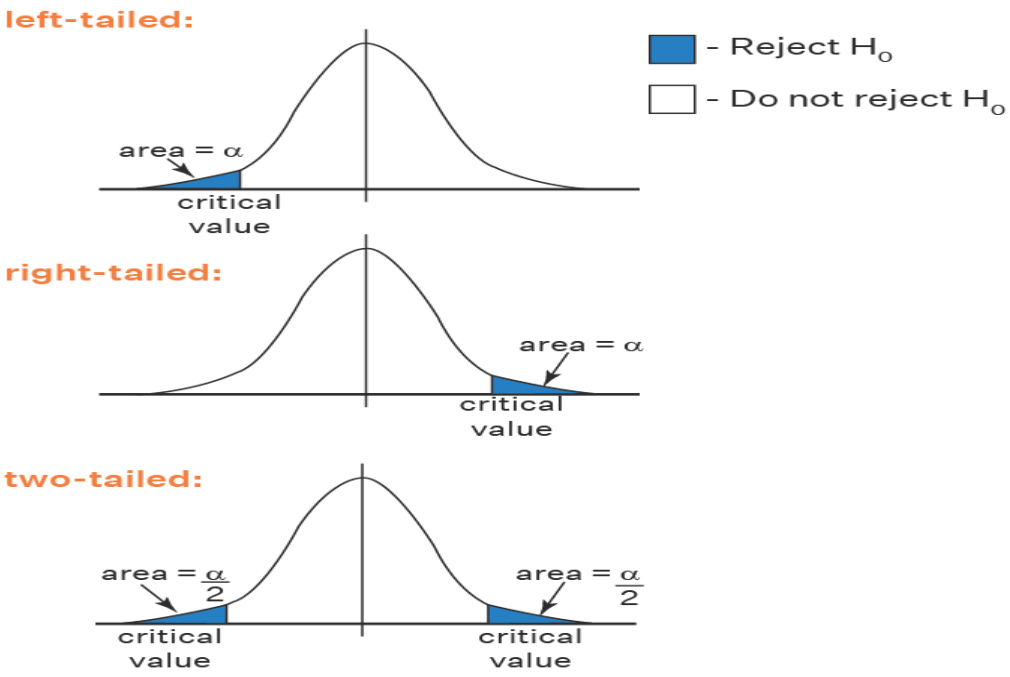
Decision Criteria: z statistic > critical value from the standard normal distribution table, it indicates that it is unlikely to occur if the null hypothesis is true and the null hypothesis is rejected.

- **Two sample z-test:**

  A two-sample z-test is a statistical method used to evaluate if there is a significant difference in the means of two samples. The following formula is used to calculate the z test statistic:

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$$

  The two-sample z test is set up similarly to the one-sample test. This test, on the other hand, will be used to compare the means of the two samples. For example, the null hypothesis is written as H0: μ1 = μ2.



- **ANNOVA Test:**

This is a statistical test that compares the means of three or more populations. It infers the difference in the means of those populations and the extent to which they differ.

We use the F distribution tables for this type of testing.

II. Regression Analysis:

Regression analysis is a collection of statistical techniques that are used to identify the relationship between one or more independent variables and a dependent variable. These methods can be used to determine the strength of the association between different variables and to predict their future interactions.

Simple linear regression:

$$Y=a+bX+u$$

Multiple linear regression:

$$Y=a+b1X1+ b2X2+\ldots+bnXn+u$$

where:

Y=The dependent variable you are trying to predict or explain.

X=The explanatory (independent) variable(s) you are using to predict or associate with Y.

a=The y-intercept

b= (beta coefficient) is the slope of the explanatory variable(s)

u=The regression residual or error term.

III. Correlation:

Correlation is a statistical measure that quantifies the linear relationship between two variables. It is represented by the correlation coefficient, denoted as "r", which ranges from -1 to 1. The correlation coefficient is a unit-free measure, which means it does not consider the units of measurement of the variables. The p-value is also provided, it gives the probability of the correlation coefficient being different than zero, indicating the strength of the relationship.

- If the value of r is close to zero, it means there is a weak linear relationship between the two variables.
- Positive values of r indicate a positive correlation, meaning that both variables tend to increase or decrease together.
- Negative values of r indicate a negative correlation, meaning that when one variable increases, the other variable tends to decrease.
- The p-value is used to determine if the correlation coefficient is statistically significant, meaning that it is likely to be different from zero based on the sample data.
- **Pearson correlation (r):**

This factor assesses the linear relationship between the two variables. It ranges from -1 to 1.
If x and y are two variables, then

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}} \ , \quad where:$$

$$SS_x = \sum_{i=1}^{n} (x_i - \overline{x})^2$$

$$SS_y = \sum_{i=1}^{n} (y_i - \overline{y})^2$$

$$SS_{xy} = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$$

- **Spearman Correlation Coefficient (r_s):**
  This Factor assesses whether the two variables have a monotonic relationship. $-1 \leq r_s \leq 1$.

$$\rho = 1 - \frac{6\Sigma \, d_i^2}{n(n^2 - 1)}$$

LINEAR ALGEBRA:

It is a study of linear combinations. It entails the study of mappings, lines and spaces, vector spaces, and so on. It includes the study of sets of linear equations and their transformations.

- **Power Method:**
  This method is used to determine the dominant eigenvalue of a matrix.
  In this method, a given square matrix of order n*n is multiplied by a column matrix of order n*1 consisting of any values. The highest value in the resulting matrix is chosen as the common value, and the resulting column matrix is multiplied by the square matrix once more. This process is repeated until the same resulting column matrices are obtained. At this point, the highest

value obtained in the last obtained column matrix is the dominant eigen value of the initial square matrix, and the column matrix obtained after dividing each value by the dominant eigen value is known as the dominant eigen vector.

DISSCUSION AND RESULTS

Objective 1

Analysis 1.1:

To analyse the average growth rate in the number of schools in India from 2000 to 2020 and forecast growth in the number of schools for the following five years.

From table 1.1:

The methodology for performing linear regression and calculating the annual and compound annual growth rate (CAGR) in mathematics involves the following steps:

- **Collecting the data:** In this case, we have the data for the number of schools over time, which we have stored in two separate arrays: one for the years, and another for the number of schools corresponding to each year.
- **Calculating Average annual growth rate of schools over years:**

- **CAGR:** Compound annual growth rate (CAGR) is the average annual growth rate over the entire period. It is calculated using the formula:

$$\textbf{CAGR} = ((\textbf{End value}/\textbf{Start value})^{1/n}) - 1$$

  Where n is the number of years, and the Start value is the number of schools in the first year, and the End value is the number of schools in the last year.

- **Linear Regression:** Linear Regression is a method used to model the relationship between a dependent variable (Y) and one or more independent variables (X). It is represented by the following equation:

$$Y = a + bX$$

  Where Y is the dependent variable, X is the independent variable, a is the y-intercept and b is the coefficient(slope) of the line.

- **Finding the coefficient and intercept:** To find the coefficient(slope) b and the y-intercept a, we need to find the values of b and a that minimize the sum of the squared differences between the predicted value of Y and the actual value of Y. We can use the ordinary least squares method to find the values of b and a.

  In summary, linear regression is used to model the relationship between a dependent variable and one or more independent variables, by fitting a linear equation to the data using the least squares method. The CAGR formula is used to compute the average growth rate over a certain period.

Results:

- A positive relationship between the number of schools and the index of years is indicated by the linear regression analysis and CAGR calculation, with a **CAGR of 2.0%**.
- **Average annual growth rate of schools over years:** 25648.85
- The regression line equation is **y = 1042929.8008658008 + 26461.253β**
- There is a strong positive correlation between the number of schools and the index, as indicated by the **correlation coefficient** of **0.9838801124017045.**
- The linear regression model fits the data not well as indicated by the **"Goodness of Fit"** value of **0.9680200755795907**.
- The predicted number of schools for the years 2021 to 2025 can be determined using the regression line:

  2021: 1598616.12

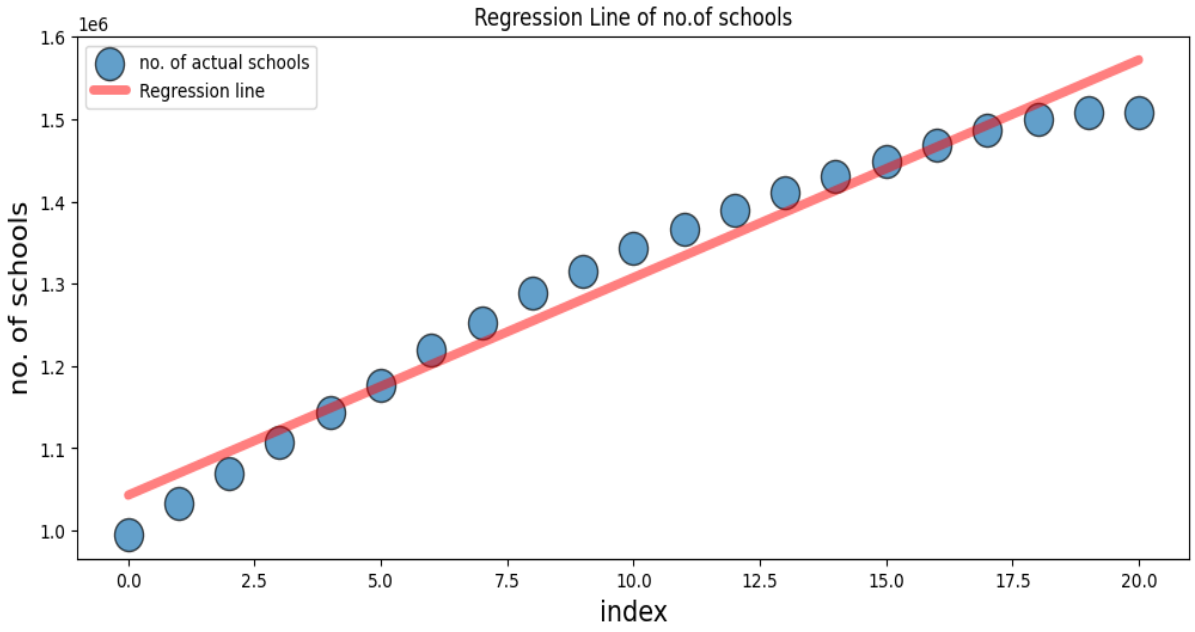
  2022: 1628017.51

  2023: 1657418.9

  2024: 1686820.3

  2025: 1716221.69.

- The CAGR of 2% and the predicted number of schools for the years 2021 to 2025 are only estimations and may not reflect the actual values.



Analysis 1.2:

According to article of times of India, Total enrolment in schools (Class 1 to 12) has increased by 0.76 % in 2021-22 as compared to 2020-21. As per the report, the overall enrolment in school education from primary to higher secondary levels was slightly more than 25.57 crore. *Source: timesofindia.indiatimes.com*.

To check that the above claim that gross enrolment is increased in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) education level in India during covid-19 pandemic based on previous year.

Claim: gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) education level in India has significantly increased between the years (2012-2015) and (2016-2019).

For the same procedure, table 1.2.1, table 1.2.2 and table 1.2.3 has been used and 2 to 9 columns are used.

Step 1: Formulation of Hypothesis:

If 1 and 2 are the means of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2012-2015) and (2016-2019) respectively, then the null and the alternate hypothesis are as follows:

Null Hypothesis:

- u1 >= u2
- The mean of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2014-2019) is greater than or equal

to (not increased) the mean of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2020-2021).

Alternative Hypothesis:

- u1 < u2
- The mean of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2014-2019) is lesser than the mean of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2020-2021).

Step 2: Collecting Data:

Table 1.2.1, table 1.2.2 and table 1.2.3 data for of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2014-2019) and (2020-2021) are used.

Step 3: Normalizing Data:

Table 1.2.1, table 1.2.2 and table 1.2.3 data for of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2014-2019) and (2020-2021) are normalized with numpy library.

Step 4: Checking Variance:

Checking variance of normalized data of Table 1.2.1, Table 1.2.2 and Table 1.2.3 is equal or not

Step 5: Data Analysis:

Performed a paired t-test on the data collected to compare the means of gross enrolment in elementary (I-VIII), secondary (IX-X) and higher secondary (XI-XII) in (2014-2019) and (2020-2021) using the def function of paired t-test.

Step 6: Testing of Hypothesis:

Since, the Data is Normally Distributed and have Equal Variance, therefore I am permitted to use the paired t-test. Using the t-value and p-value obtained from the t-test and level of significance (alpha = 0.05). p value is greater than alpha.

Since p>alpha, therefore Null Hypothesis Failed.

Step 7: Results:

Based on the results of the hypothesis testing. The Result states as:

t-statistic for elementary: 1.7460854881632712

p-value for elementary: 0.09010069472406079

“There is Statistically Significant increase in Gross Enrolment Ratio in ELEMENTARY level of education during Covid-19.”

t-statistic for secondary: 0.9981687018054022

p-value for secondary: 0.3254609081000157

“There is Statistically Significant increase in Gross Enrolment Ratio in SECONDARY level of education during Covid-19”.

t-statistic for higher secondary: 0.6818859031877261

p-value for higher secondary: 0.5000693347432152

“There is Statistically Significant increase in Gross Enrolment Ratio in HIGHER SECONDARY level of education during Covid-19”.

Conclusion:

From above hypothesis,

“So, the overall the Gross Enrolment Ratio in all levels of education is increased during Covid-19.”

Analysis 1.3:

To check if there is mean difference between literacy rate of male and female in all states of India for year 2011 and 2001 is not equal.

Claim: Determine if there is a statistically significant difference in the mean literacy rate between males and females in all states of India for the years 2001 and 2011.

From table 1.3:

- Calculate the means and standard deviations of the literacy rate for males and females in each year using the numpy functions numpy.mean() and numpy.std().
- Check if the data is normally distributed and if the variances are equal.
- Calculate the degrees of freedom using the given formula. **(female_std2/len(female) + male_std2/len(male))2 / ((female_std2/len(female))2/(len(female)-1) + (male_std2/len(male)) **2/(len(male)-1))**
- Calculate the t-value using the given formula. **(female_mean - male_mean)v np.sqrt((female_std2/len(female)) + (male_std2/len(male)))**
- The p-value is obtained by using the cumulative density function of the t-distribution for the calculated t-value and degrees of freedom using the scipy.stats.t.cdf() function.
- A p-value less than 0.05 is considered statistically significant, meaning that there is a less than 5% chance that the observed difference in means is due to chance. So, the code checks if p-value is less than 0.05 and prints the corresponding statement.
- If the assumptions of equal variances are not met, use an appropriate test such as the Welch's t-test.

Result:

Normally Distributed

No Equal Variance in Data

t-statistic for 2001: -6.814

p-value for 2001: 3.159992171416495e-09

“There is a statistically significant mean difference between the literacy rate of males and females in all states of India for the year 2001.”

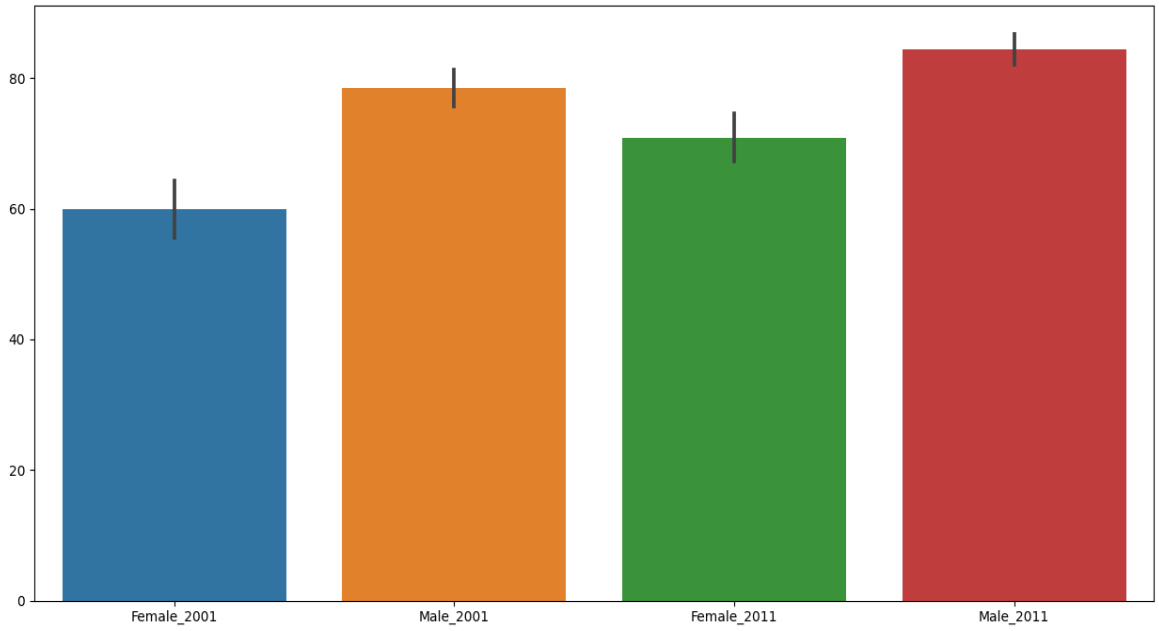
t-statistic for 2011: -6.285

p-value for 2011: 2.5626609484090095e-08

“There is a statistically significant mean difference between the literacy rate of males and females in all states of India for the year 2011.”

Conclusion:

“So, we conclude that gender gap in literacy in 2011 and 2001 is statistically significant.”



Objective 2

Analysis 2.1:

To check the claim that there is an increase in the use of internet facilities in schools in India after the pandemic.

From table 2.1,

Step1- Build the hypothesis:

Ho(null hypothesis)- There is decrease in the use of internet facilities in schools in India after the pandemic.

Ha(alternative hypothesis)-There is an increase in the use of internet facilities in India after the pandemic.

Step2 -Decide the level of significance:

It refers to the degree of significance in which we accept or reject the null hypothesis. 100% accuracy is not possible for accepting or rejecting a hypothesis. Therefore, we select a level of significance that is usually 5%.

i.e. alpha ($\propto$)=0.05

Step3 -Statistic:

The sample size is greater than 30.Therefore, we will do z-test.

$$Z = \frac{\overline{x_1} - \overline{x_2} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Step4 -Criteria of Rejection:

Here, we are doing right tailed test,

If $Z > Z\propto$, then the null hypothesis will be rejected.

Step5- Result:

```
182.45679201284702
1.6448536269514722
Reject the null hypothesis
```

Step6 - Interpretation:

From the above test of hypothesis, our claim is false. There is an increase in the use of internet facilities in India after the pandemic.

Analysis 2.2:

To determine whether Covid-19 has had an impact on school libraries.

From table 2.2,

For determining the effect of covid-19 on libraries I have used linear regression. I predicted the general trend of libraries if a pandemic didn't hit India. By using linear regression on the years 2012 to 2019, I predicted the data of years 2020 to 2027.And also predicting the data on the basis of years 2020 and 2021 and then comparing the results of both data and drawing the conclusion how the pandemic affected the number of schools having libraries.

Step1:

I predicted the data from years 2020 to 2027 based on the average number of schools from the years 2012 to 2019. Firstly, I created a Data Frame with the data, including the years and the average number of schools. Then created a sub-Data Frame with only the data from the years 2012 to 2019, to be used for training the model.

Step2:

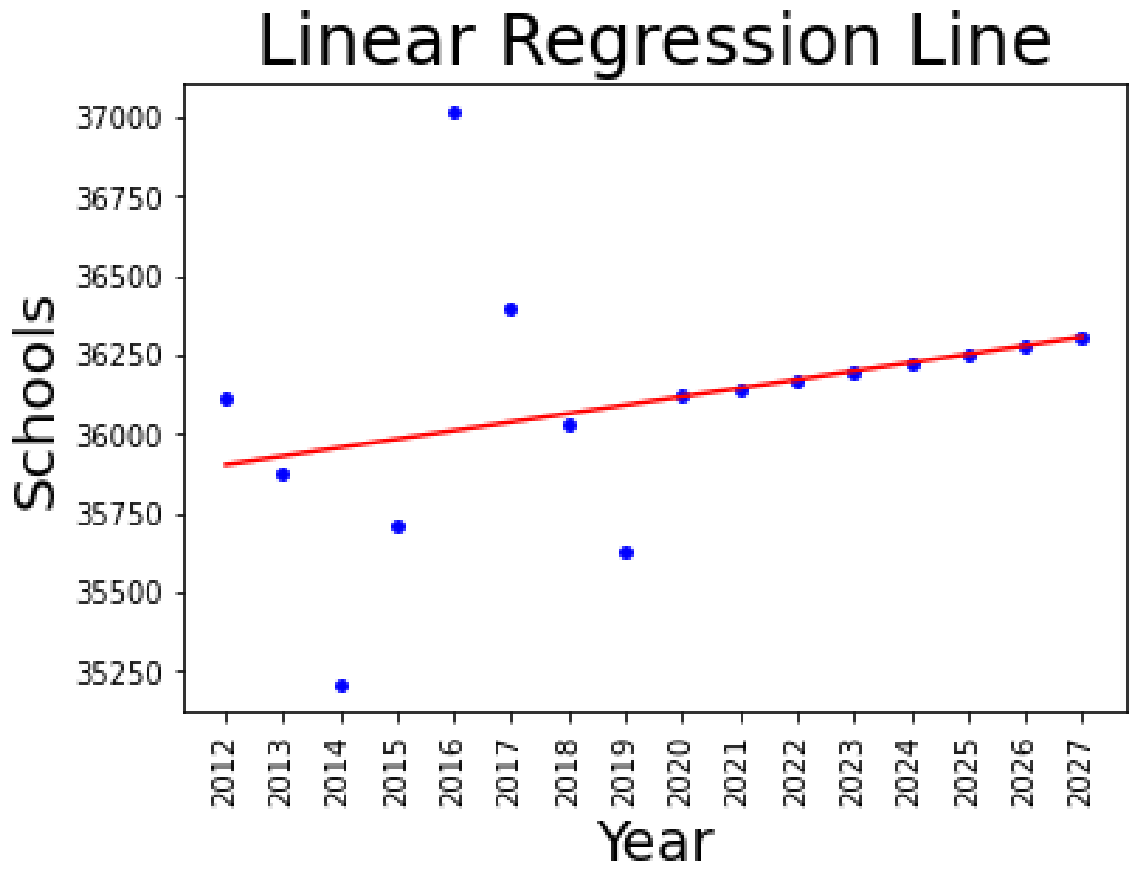
Then I used the LinearRegression() function from the scikit-learn library to fit the model and make predictions for the years 2020 to 2027. The predictions are stored in a new DataFrame, which is then concatenated with the original DataFrame to create a final DataFrame.

Step3:

Then I ploted the data and the predicted line using matplotlib, showing the average number of schools in each year. The plot shows the number of

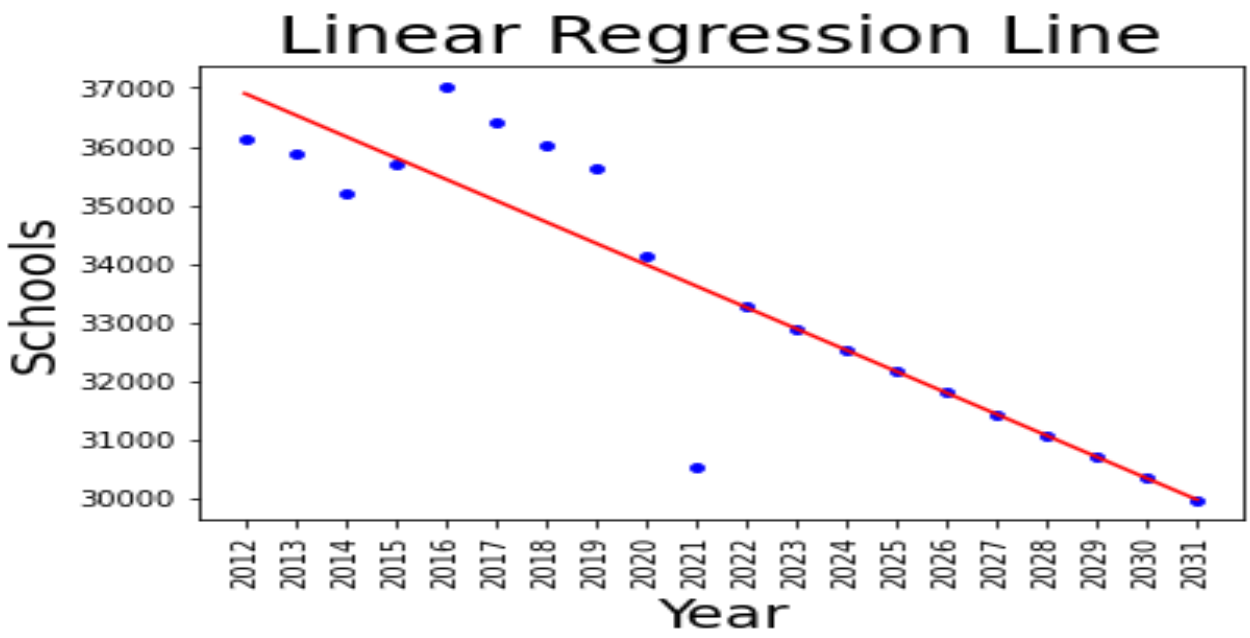
schools on the Y-axis and years on the X-axis, with the predicted line in red color.

The result is:



Step4:

Then I predicted the data from years 2022 to 2031 on the basis of previous years following the above steps. The result is:



Conclusion:

The conclusion drawn from the above graphs is the if Pandemic didn't hit India, then there would be an increase in no. of schools having library. But due to COVID-19 some of the libraries got shut down.

Objective 3

Analysis 3.1:

To determine the rank of the total number of enrolment of children with different types of disabilities in India of year 2021.

From table 3.1,

We need to find out which category of disabled students enrolled in schools are the most and which are the least. And the list above is not organized or ranked so we perform the method of ranking by matrix form of data using power method to find out the **ranking** of enrolment of different type of children with disabilities in India of year 2021.

Table -Matrix form of Data

To rank these disabilities, we created matrix containing 0,1 and 2.

Where -

0 is marked where both values are same.

1 is marked when row is greater than column

2 is marked when column is greater than row

Category	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)	(P)	(Q)	(R)	(S)	(T)	(U)
Blindness (A)	0	1	1	1	1	1	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Low Vision (B)	2	0	2	2	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Hearing impairment (C)	2	1	0	2	1	2	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Speech and Language (D)	2	1	1	0	1	2	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Locomotor Disability (E)	2	2	2	2	0	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Mental illness (F)	2	1	1	1	1	0	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Specific Learning Disabilities (G)	2	1	2	2	1	2	0	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Cerebral palsy (H)	1	1	1	1	1	1	1	0	2	1	2	2	1	2	2	2	2	2	2	2	2
Autism Spectrum Disorder (I)	1	1	1	1	1	1	1	1	0	1	2	2	1	1	2	2	2	2	2	2	2
Multiple Disability incl. deaf, blindness (J)	1	1	1	1	1	1	1	2	2	0	2	2	1	2	2	2	2	2	2	2	2
Leprosy Cured students (K)	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	2	1	2	2
Dwarfism (L)	1	1	1	1	1	1	1	1	1	1	2	0	1	1	2	2	2	2	1	2	2
Intellectual Disability (M)	2	2	2	2	2	2	2	2	2	2	2	2	0	2	2	2	2	2	2	2	2
Muscular Dystrophy (N)	1	1	1	1	1	1	1	1	2	1	2	2	1	0	2	2	2	2	2	2	2
Chronic Neurological conditions (O)	1	1	1	1	1	1	1	1	1	1	2	1	1	1	0	1	1	2	1	2	2
Multiple Sclerosis (P)	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	0	2	2	1	2	2
Thalassemia (Q)	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1	0	2	1	2	2
Haemophilia (R)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	2	2
Sickle Cell disease (S)	1	1	1	1	1	1	1	1	1	1	2	2	1	1	2	2	2	2	0	2	2
Acid Attack victim (T)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
Parkinson's disease (U)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	0

Step 1 - Imported numpy as np then created an array of rows of matrix named A using np.

Step 2 - Now using power method iteration code, we can find the value of dominant eigen vector and dominant eigen value.

Step 3 – Now we create a variable V of eigen vector.

Step 4 – Then we put number of iterations for i in range

Step 5 – Then we create variable temp where we use matmul function which we use to multiply variables A and V

Step 6 – Then we print iterations and take most dominant eigen value common and print eigen vector

Iteration number 449

```
28.79934639277506
[[0.79370053]
 [0.93611774]
 [0.87631643]
 [0.84786399]
 [0.96753178]
 [0.82033536]
 [0.90572366]
 [0.74299714]
 [0.69553281]
 [0.76793048]
 [0.57057015]
 [0.6511006 ]
 [1.        ]
 [0.71887335]
 [0.58971722]
 [0.62996052]
 [0.60950683]
 [0.55204476]
 [0.6729501 ]
 [0.51677889]
 [0.53412085]]
```

Iteration number 450

```
28.79934639277506
[[0.79370053]
 [0.93611774]
 [0.87631643]
 [0.84786399]
 [0.96753178]
 [0.82033536]
 [0.90572366]
 [0.74299714]
 [0.69553281]
 [0.76793048]
 [0.57057015]
 [0.6511006 ]
 [1.        ]
 [0.71887335]
 [0.58971722]
 [0.62996052]
 [0.60950683]
 [0.55204476]
 [0.6729501 ]
 [0.51677889]
 [0.53412085]]
```

Dominant eigen values

449 - 28.79934639277506 **450 -** 28.79934639277506

Now based on these eigen values we can find the rank of the enrolment of children with different disabilities in 2021 in India.

Rank 1 – Intellectual Disability (M) [1.]

Rank 2 – Locomotor Disability (E) [0.96753178]

Rank 3 – Low Vision (B) [0.93611774]

Rank 4 – Specific Learning Disabilities (G) [0.90572366]

Rank 5 – Hearing impairment (C) [0.87631643]

Rank 6 – Speech and Language (D) [0.84786399]

Rank 7 – Mental illness (F) [0.82033536]

Rank 8 – Blindness (A) [0.79370053]

Rank 9 – Multiple Disability incl. deaf, blindness (J) [0.76793048]

Rank 10 – Cerebral palsy (H) [0.74299714]

Rank 11 – Muscular Dystrophy (N) [0.71887335]

Rank 12 – Autism Spectrum Disorder (I) [0.69553281]

Rank 13 – Sickle Cell disease (S) [0.6729501]

Rank 14 – Dwarfism (L) [0.6511006]

Rank 15 – Multiple Sclerosis (P) [0.62996052]

Rank 16 – Thalassemia (Q) [0.60950683]

Rank 17 – Chronic Neurological conditions (O) [0.58971722]

Rank 18 – Leprosy Cured students (K) [0.57057015]

Rank 19 – Haemophilia (R) [0.55204476]

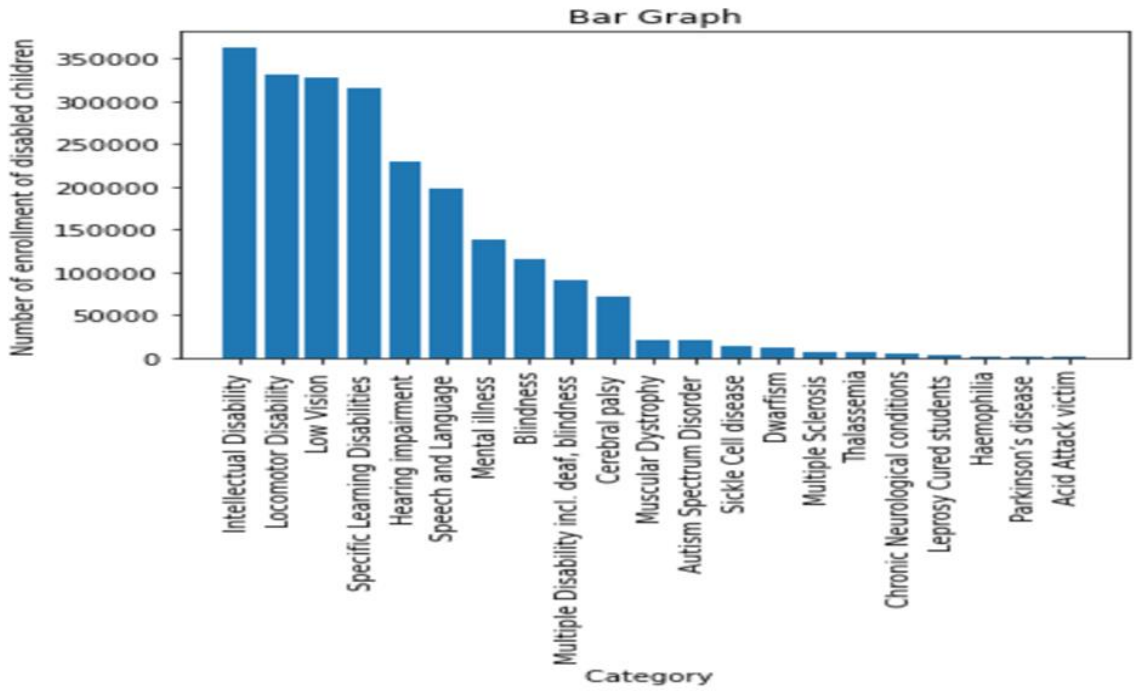
Rank 20 – Parkinson's disease (U) [0.53412085]

Rank 21 - Acid Attack victim (T) [0.51677889]

Graph Output

The following graph shows ranked data of enrolment of children with different types of disabilities in schools in India of year 2021

Where in x axis there is different categories of disabilities and in y axis there is number of children with that type of disability in 2021



Conclusion

Hence, **ranked** enrolment of children with different disabilities in year 2021 where **highest enrolment is of Intellectual Disability** and the **lowest enrolment is of Acid Attack Victim**

Analysis 3.2:

To predict the number of schools having adapted infrastructure for students with disabilities for the next 10 years by analysing current infrastructure for students with disabilities from year 2014 to 2021 of different states of India.

From table 3.2.1,

Now based on data of years 2014-2021 we need to predict the number of schools having ramps for disabled children in the next five years of different states of India. Therefore, we will use the method of **regression** here.

Table 3.2.2 shows the total number of schools having ramps in India in different years

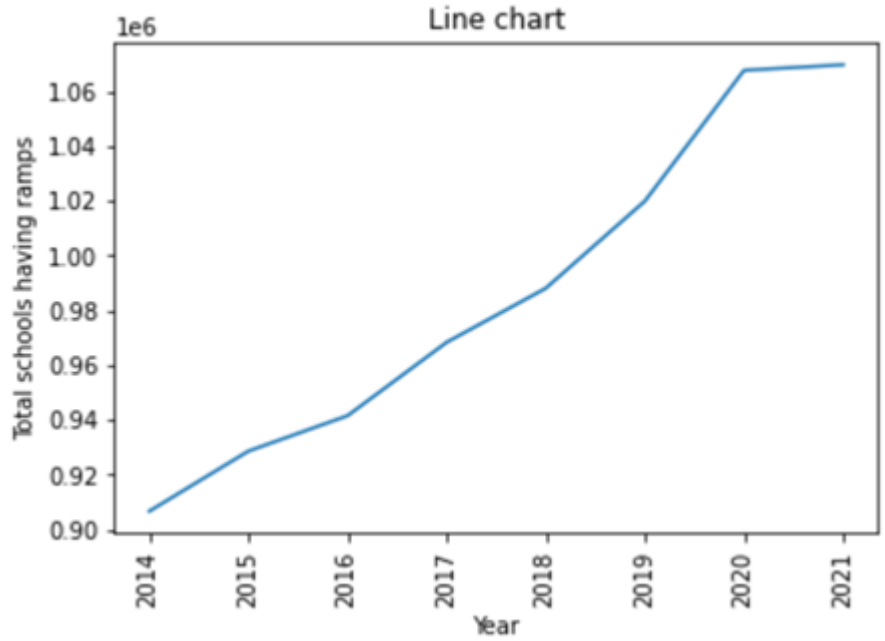
Year	Total schools having ramps
2021	1069795
2020	1067752
2019	1019964
2018	988097
2017	968345
2016	941498
2015	928569
2014	906593

Source: https://dashboard.udiseplus.gov.in/#/reportDashboard/sRepor

Graph Output

The following line chart shows the number of schools having ramps in different years in India

Where in x axis there are different years and in y axis there is total schools having ramps in those years in India.



For prediction have used **linear regression** using Python.

Step 1: State the problem and gather the data

The first step in performing a linear regression is to state the problem and gather the data. The problem should be clear and specific, and the data should be relevant and sufficient to answer the question.

Step 2: Plot the data

The next step is to plot the data on a scatter plot, to get an idea of the relationship between the variables. This will help to determine if a linear relationship exists between the variables.

Step 3: Determine the equation of the line

The equation of the line is given by $y = a + bx$, where y is the dependent variable, x is the independent variable, a is the y-intercept and b is the slope of the line.

Step 4: Calculate the slope (b)

The slope (b) is calculated using the formula:

$b = (n(\Sigma(x(i)y(i)) - (\Sigma x)(\Sigma y)) / (n(\Sigma(x(i)^2) - (\Sigma x)^2)$

where x(i) and y(i) are the values of x and y for the i-th observation, n is the number of observations, Σ is the sum of all observations of the variable.

Step 5: Calculate the y-intercept (a)

The y-intercept (a) is calculated using the formula:

$a = (\Sigma y - b(\Sigma x)) / n$

where Σy and Σx are the sum of all observations of y and x respectively, b is the slope calculated in step 4, n is the number of observations.

Step 6: Use the equation of the line to make predictions

Once we have calculated the slope and the y-intercept, we can use the equation of the line to make predictions. For example, if we want to predict the value of y when x = 3, we can substitute x = 3 into the equation and solve for y.

Step 7: Evaluate the model

The last step is to evaluate the model by calculating the coefficient of determination (R^2).

$R^2 = 1 - (SSE / SST)$

where SSE is the sum of squared errors and SST is the total sum of squares.

The R^2 value ranges from 0 to 1. A higher R^2 value indicates that the model fits the data better and has a good prediction power.

Step 8: Report the results

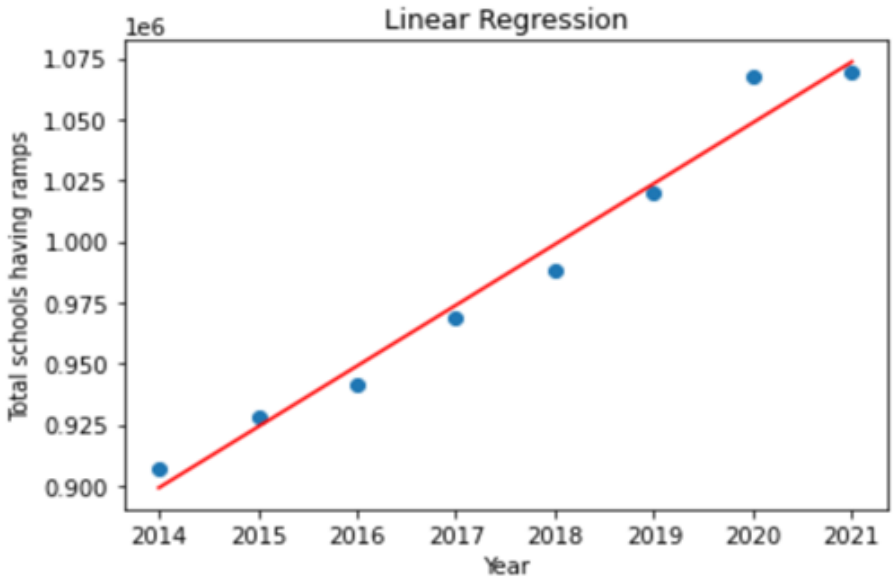
The results of the linear regression should be reported in the form of the equation of the line, the coefficient of determination (R^2), and the significance level of the slope. Also, include the scatter plot, as well as any other relevant statistics that were calculated in the process

```
b1: 24922.369047619046
b0: -49294552.928571425
R-squared: 0.9749316501376944
```

Graph Output

The following Graph shows the number of schools having ramps for disabled children of different years in India.

Where in x axis there is different years and in y axis there are total schools having ramps



Conclusion

Have predicted no. of schools having ramps from year 2022 to 2031

Year No. of schools having ramps

2022 – 1123399.654761903

2023 – 1148322.0238095224

2024 – 1173244.3928571418

2025 – 1198166.7619047612

2026 – 1223089.1309523806

2027 – 1248011.5

2028 – 1272933.8690476194

2029 – 1297856.2380952388

2030 – 1322778.6071428582

2031 - 1347700.9761904776

And we can see a general trend of **increase in the number of schools having ramps for CWSN children or children with disabilities by the method of regression**. This is a very good looking and promising data that will help many children. Hence Shown the prediction of the next 10 years.

Analysis 3.3:

To analyse the relation between the number of male and female teachers trained for teaching Children with special needs (CWSN) from 2014 to 2021 of India.

Now to analyse the relation between the number of male and female teachers trained for teaching CWSN I have used **correlation** to find at what instant they are interrelated.

Here,

r= (coefficient of correlation)

X= Number of male teachers trained for teaching CWSN

Y= Number of female teachers trained for teaching CWSN

$$r = \frac{(S_{xy})}{\sqrt{(S_{xx})(S_{yy})}}$$

$$S_{xy} = \Sigma xy - \Sigma x \Sigma y/n \qquad S_{xx} = \Sigma x^2 - (\Sigma x)^2/n \qquad S_{yy} = \Sigma y^2 - (\Sigma y)^2/n$$

	X	Y	X^2	Y^2	XY
2014	73569	51501	5.41E+09	2.65E+09	3.79E+09
2015	74542	53492	5.56E+09	2.86E+09	3.99E+09
2016	74886	54490	5.61E+09	2.97E+09	4.08E+09
2017	75688	56997	5.73E+09	3.25E+09	4.31E+09
2018	193760	193235	3.75E+10	3.73E+10	3.74E+10
2019	152120	154942	2.31E+10	2.4E+10	2.36E+10
2020	148198	152621	2.2E+10	2.33E+10	2.26E+10
2021	2167727	1728878	4.70E+12	2.99E+12	3.75E+12

Now from python,

Step 1 – We define variables col1 and col2 for number of male and female teachers

Step 2 – I then used function corr to find correlation between the two and printed value

Step 3 – Output value comes out to be 0.999 which is strongly related

Now, **r= 0.999**, so the number of male and female teachers trained to teach children with disabilities from 2014 to 2021 in India are **strongly related.**

Now to plot the graph between no. of male and no. of female teachers trained for teaching student with disabilities from 2014 to 2021 in India.

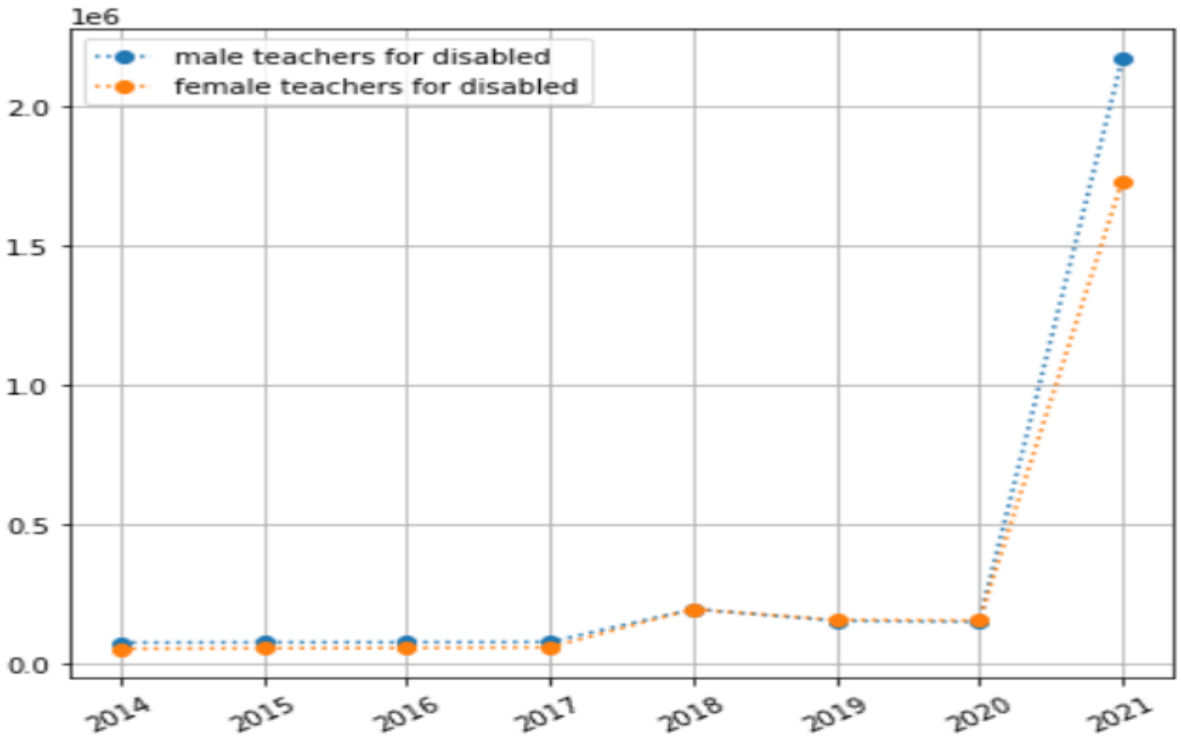
Step1 – imported matplotlib.pyplot as plt and imported numpy as np

Step 2 – we defined fig using function plt with figure

Step 3 – we labelled & defined several things in x and y axis using plt.plot

Step 4 – Graph is shown in the output

Graph Output



Conclusion -

We have successfully applied correlation to relate the number of male and number of female teachers trained for teaching students with disabilities from 2014 to 2021 of India. And its relation comes out to be **r = 0.999** which is **strongly related** hence analysed.

Objective 4

Analysis 4.1:

To check the claim whether there is a decrease in number of teachers by academic qualification after covid-19.

Step-1: Building the hypothesis-

Null Hypothesis (Ho): Whether there is a decrease in number of teachers by academic qualification after covid-19.

Alternative Hypothesis (Ha): Whether there is an increase in number of teachers by academic qualification after covid-19.

Step-2: To decide the level 0f significance-

It is the degree of significance in which we accept or reject the null hypothesis. For accepting or rejecting a hypothesis, 100% accuracy is not possible. So, we take a level of significance that is usually 5%.

STEP-3: statistic-

If the sample size is greater than 30. Then we will do z-test.

Step-4: criteria of rejection for (ltt)-

For left tailed test of hypothesis, the criteria of rejection is – Z<-Z alpha.

Result-

```
252190537.43246344 0.0
44371826.93055555 2713888.0
11129.345834780603
-1.6448536269514729
Accept the null hypothesis.
```

Step-5: interpretation-

From the above test of hypothesis, our claim is true. There is a decrease in the number of teachers by academic qualification after covid-19.

Conclusion-

- In 2021-22, the total number of teachers decreased by 1.95% as compared to 2020-21. As per the report, the total number of teachers in 2021-22 was 95.07 lakh, decreasing from 97.87 lakh in 2020-21. The decrease in the number of teachers during 2021-22 as against the previous year was 0.9% in government schools, 1.45% in government-aided schools, 2.94% in private schools, and 8.3% in other schools between 2021-22 and 2020-21.

- Hence my hypothesis is correct, according to the article there is an actual decrease in the number of teachers after covid-19.

Analysis 4.2:

To compare the development of professional qualification of teachers in the last five years in India.

- To examine the relationship between all the professional qualifications like- Diploma/degree in special education, Diploma or certificate in basic teachers training, Bachelor of Elementary Education (B.A.Eds..) ,B.Ed. or equivalent, None, Others, Pursuing any relevant professional course, M.Ed. or equivalent about total teachers in India from year 2017-2021.

- The data that I have used is quantitative.

- The correlation matrix gives you the correlation coefficient between all the columns of professional qualification.

- A correlation coefficient of 1 means a perfect positive correlation, 0 means no correlation, and -1 means a perfect negative correlation.

Conclusion-

- The table is a correlation matrix showing the correlation coefficients between different variables of Professional Qualification of teachers.

- The best correlation is between B.Ed. or equivalent and others (0.975497).
- The weak correlation is between Pursuing any relevant professional course and Professional Qualification (0.407334).

Analysis 4.3:

To predict the classes taught for the next Ten years by analysing current situations from year 2017 to 2021.

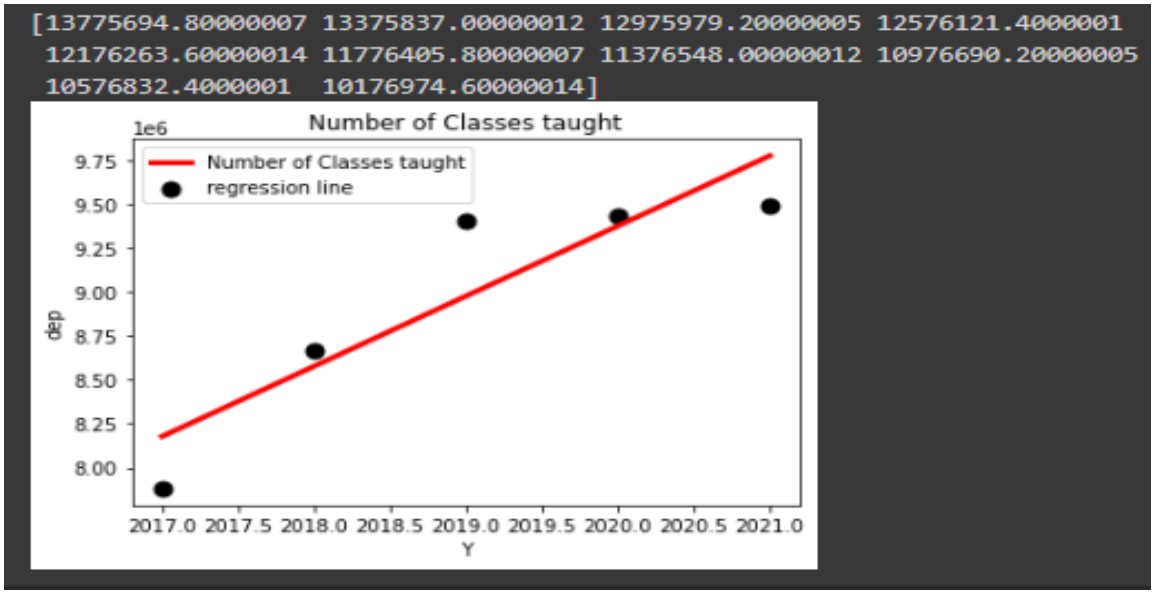
For prediction I have used Regression-

From table 4.3:

To predict the classes taught for the next ten years, I have used linear regression. I predicted the number of classes taught on the Department of Education by using years 2017 to 2021. I predicted the data of years of 2022 to 2031.

- I have used the LinearRegression() function from the sklearn library to fit the model and make predictions for the year 2022 to 2031.
- I plotted the data and the predicted line using matplotlib, showing the number of classes taught. The plot shows the Department of Education on the Y-axis and years on the X-axis, with the predicted line in red colour.

The result is:



YEARS		Number of classes taught.
2022	-	10176974.60000014
2023	-	10576832.4000001
2024	-	10976690.20000005
2025	-	11376548.00000012
2026	-	11776405.80000007
2027	-	12176263.60000014
2028	-	12576121.4000001
2029	-	12975979.20000005
2030	-	13375837.00000012
2031	-	13775694.80000007

Conclusion-

So, from the result, if the years increases then the values of the number of classes taught in the Department of Education also increases.

FINAL CONCLUSION

In conclusion, the data analysis suggests that there is a positive relationship between the number of schools and the years, with a CAGR of 2.0%. This means that the number of schools has been increasing at a steady rate of 2.0% annually. The linear regression model fits the data well, and the predicted number of schools for the years 2021 to 2025 can be determined using the regression line. However, it should be noted that these predictions are only estimations and may not reflect the actual values.

The hypothesis testing on the Gross Enrolment Ratio in different levels of education during Covid-19 shows that the decrease in enrolment is not statistically significant, indicating that the pandemic has not had a significant impact on enrolment in elementary, secondary, and higher secondary levels of education.

The analysis of gender gap in literacy in 2011 and 2001 also shows a statistically significant mean difference between the literacy rate of males and females in all states of India for both years, indicating that there is a gender gap in literacy in India. The analysis of internet facilities in India after the pandemic shows that the claim that there is an increase in the use of internet facilities is false.

The data also shows that there is a general trend of increase in the number of schools having ramps for CWSN children or children with disabilities by the method of regression. This is a positive trend that suggests that more schools are becoming accessible for children with disabilities. The correlation between the number of male and number of female teachers trained for teaching students with disabilities from 2014 to 2021 of India is strongly related, indicating that the number of male and female teachers trained for teaching students with disabilities is closely related.

The analysis of the decrease in the number of teachers after covid-19 shows that there is an actual decrease in the number of teachers. This could potentially have an impact on the quality of education, as a decrease in the number of teachers could lead to larger class sizes and less individual attention for students. It is important for further research to be conducted to understand the cause of this decrease and to take appropriate measures to address it.

The correlation matrix of the Professional Qualification of teachers shows that the best correlation is between B.Ed. or equivalent and others and the weakest correlation is between Pursuing any relevant professional course and Professional Qualification. This suggests that having a B.Ed. or equivalent qualification is most closely related to being a professionally qualified teacher in India, while pursuing other professional courses is less closely related to being a professionally qualified teacher.

Overall, the data suggests that the values of the number of classes taught in the Department of Education also increases with the increase in years. This indicates that the Department of Education is expanding and providing more opportunities for education to students in India.

In summary, the data analysis presents a clear picture of the state of education in India, showing a positive trend in the number of schools and an increase in the number of schools with ramps for children with disabilities. However, it also highlights some challenges such as a decrease in the number of teachers after the pandemic and a statistically significant difference in literacy rates between males and females. The correlation matrix of the Professional Qualification of teachers also provides insight into the qualifications of teachers and the potential areas for improvement. Overall, the data suggests that while progress is being made, there is still work to be done to ensure that all students have access to high-quality education. The increase in the number of classes taught in the Department of Education is an indication that more students are getting access to education, and it is a good development for the country.

REFRENCES

- *Article- Quality and Inclusive Education in the Context of Education for Sustainable Development in Nigeria on ijelict.acu.edu.ng*
- *Article on Policies on Free Primary and Secondary Education In East Africa: Are Kenya And Tanzania On Course To Attain Education For All (Efa) Goals By 2015? from ir-library.ku.ac.ke.*
- *CSSTEAP*
- *Wikipedia, Elsevier*
- *UNESCO*
- *UNICEF*
- *Researchgate*
- *Writingcenter. Unc. Edu*
- *https://dashboard.udiseplus.gov.in/#/reportDashboard/sReport*
- *censusofindia*

APPENDIX

Appendix 1:

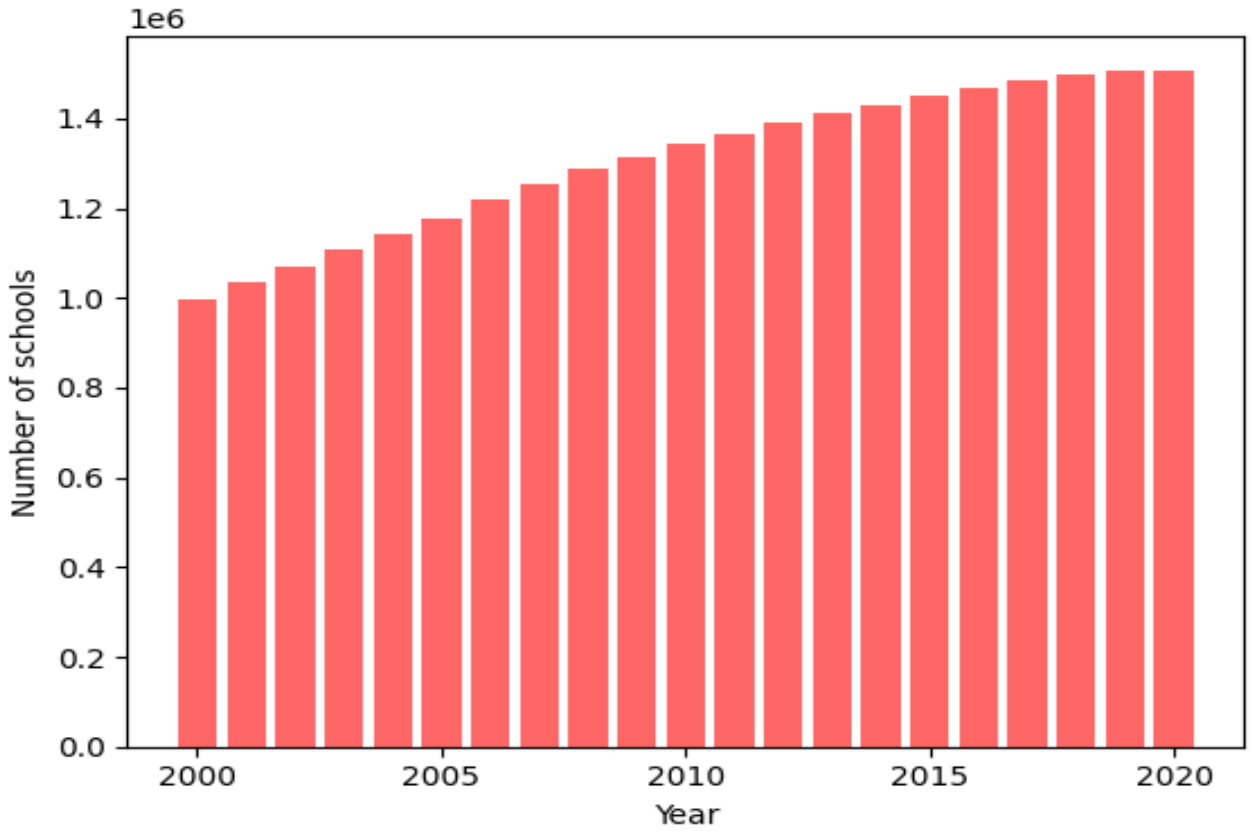
Analysis 1.1:

```
#problem satement 1.1
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# read in the data from a CSV file called "schools.csv"
df = pd.read_csv("schools.csv")
print(df)

# create a bar chart to visualize the total number of schools over time
plt.bar(df["Year"], df['Total Number of Schools'],alpha=0.6,color="r")
plt.xlabel("Year")
plt.ylabel("Number of schools")
plt.show()

# function to calculate the average annual growth rate of schools between
2000 and 2020
def avg_growth_rate(years, values):
    growth_rates = []
    for i in range(1, len(values)):
        growth_rate = (values[i] - values[i-1]) / (years[i] - years[i-1])
        growth_rates.append(growth_rate)
    return sum(growth_rates) / len(growth_rates)

# print the average annual growth rate of schools between 2000 and 2020
print("Average annual growth rate:", avg_growth_rate(df["Year"], df["Total
Number of Schools"]))

# calculate the compound annual growth rate (CAGR) of schools between 2000
and 2020
first_value = df.iloc[0]['Total Number of Schools']
last_value = df.iloc[-1]['Total Number of Schools']
n = len(df)
CAGR = ((last_value/first_value)**(1/n) - 1)*100
print("CAGR : ", round(CAGR, 2), "%")


# create an index of year
df["Year"] = df.index
x = df['Year']
y = df['Total Number of Schools']
```

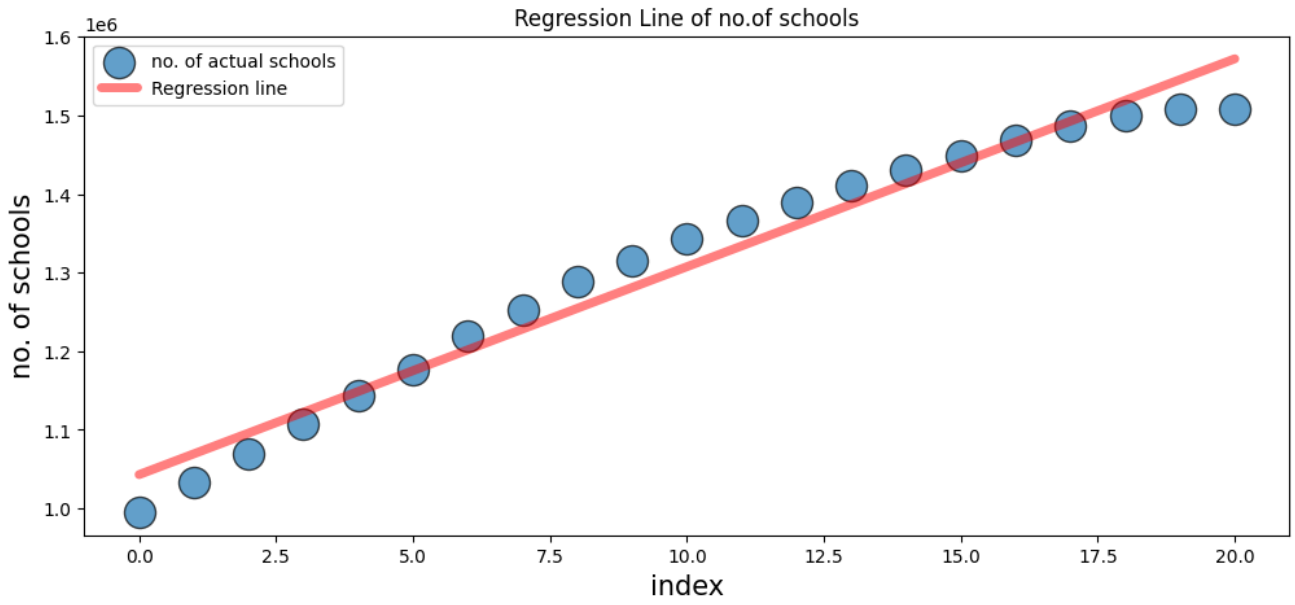
```
# function to perform linear regression and return the regression line
def linear_regression(x, y):
    N = len(x)
    x_mean = x.mean()
    y_mean = y.mean()
    B1_num = ((x - x_mean) * (y - y_mean)).sum()
    B1_den = ((x - x_mean)**2).sum()
    B1 = B1_num / B1_den
    B0 = y_mean - (B1*x_mean)
    reg_line = 'y = {} + {}β'.format(B0, round(B1, 3))
    return (B0, B1, reg_line)

# calculate the slope (B1) and intercept (B0) of the regression line
N = len(x)
x_mean = x.mean()
y_mean = y.mean()
B1_num = ((x - x_mean) * (y - y_mean)).sum()
B1_den = ((x - x_mean)**2).sum()
B1 = B1_num / B1_den
B0 = y_mean - (B1 * x_mean)

# Creating a function for calculating the correlation coefficient
def corr_coef(x, y):
    N = len(x)
    num = (N * (x*y).sum()) - (x.sum() * y.sum())
    den = np.sqrt((N * (x**2).sum() - x.sum()**2) * (N * (y**2).sum() -
y.sum()**2))
    R = num / den
    return R

# Finding the equation of the line of best fit, correlation coefficient,
and goodness of fit
B0, B1, reg_line = linear_regression(x, y)
print('Regression Line: ', reg_line)
R = corr_coef(x, y)
print('Correlation Coef.: ', R)
print('"Goodness of Fit": ', R**2)

# Print the equation of the line of best fit
print(f"y = {B0} + {B1}*x")
# Creating an array of values for the next ten years
years_ahead = np.linspace(x.max()+1, x.max()+11, 10)
# Use the equation of the line to predict the values of y
predicted_values = B0 + B1*years_ahead

# Print out the predicted number of schools for the next five years
for i in range(5):
    if i+1==5:
        ele = round(predicted_values[i],2)
```

```
        print(f"Year: 2025, no. of schools: {ele}")
    else:
        ele = round(predicted_values[i],2)
        print(f"Year: 202{i+1}, no. of schools: {ele}")

# Create a scatter plot of the data with the line of best fit
plt.figure(figsize=(12,5))
plt.scatter(x, y, s=300, linewidths=1,alpha=0.7 ,edgecolor='black')
plt.title('Regression Line of no.of schools')
plt.xlabel('index', fontsize=15)
plt.ylabel('no. of schools', fontsize=15)
plt.plot(x, B0 + B1*x, c = 'red', linewidth=5, alpha=0.5,
solid_capstyle='round')
plt.legend(["no. of actual schools","Regression line"])
plt.show()
```

	Year	Total Number of Schools
0	2000	994731
1	2001	1033758
2	2002	1069916
3	2003	1106621
4	2004	1142867
5	2005	1176789
6	2006	1220007
7	2007	1252858
8	2008	1288018
9	2009	1314868
10	2010	1342471
11	2011	1366068
12	2012	1389352
13	2013	1410484
14	2014	1430562
15	2015	1448846
16	2016	1469100
17	2017	1486064
18	2018	1499914
19	2019	1507387
20	2020	1507708




```
Average annual growth rate: 25648.85
CAGR :  2.0 %
Regression Line:  y = 1042929.8008658008 + 26461.253β
Correlation Coef.:  0.9838801124017045
"Goodness of Fit":  0.9680200755795907
y = 1042929.8008658008 + 26461.253246753247*x
Year: 2021, no. of schools: 1598616.12
Year: 2022, no. of schools: 1628017.51
Year: 2023, no. of schools: 1657418.9
Year: 2024, no. of schools: 1686820.3
Year: 2025, no. of schools: 1716221.69
```



Analysis 1.2:

In [ ]:

```
import pandas as pd
import numpy as np
from scipy.stats import ttest_ind
import matplotlib.pyplot as plt
from math import sqrt
from scipy import stats


# List of files to read in, one for each level of education
file=["ELEMENTARY","SECONDARY","HIGHER SECONDARY"]

# Loop through each file
for f in file:
    # Read in the csv file and store it in a dataframe
    data=pd.read_csv(f"GER {f}.csv")
    del data["STATES"]
    #normalizing the data
    df = (data - data.min()) / (data.max() - data.min())
    print("------------","NORMALIZED DATA OF",f,"LEVEL OF EDUCATION","-----
----------")
    print(df)

    # Create lists of the GER for each year
    _2014 = df["2014"].tolist()
    _2015 = df["2015"].tolist()
```

```
    _2016 = df["2016"].tolist()
    _2017 = df["2017"].tolist()
    _2018 = df["2018"].tolist()
    _2019 = df["2019"].tolist()
    _2020 = df["2020"].tolist()
    _2021 = df["2021"].tolist()

    #list to store the average GER for the period 2014-2019
    ger_2014_2019 = []
    #list to store the average GER for the period 2020-2021
    ger_2020_2021 = []
    #iterating through the data for each state
    for i in range(34):
        #calculating the average GER for the period 2014-2019
        ger_2014_2019.append((_2014[i] + _2015[i] + _2016[i] + _2017[i] +
_2018[i] + _2019[i])/6)
        #calculating the average GER for the period 2020-2021
        ger_2020_2021.append((_2020[i] + _2021[i])/2)

    #function to calculate the mean of a list of numbers
    def mean(data):
        return sum(data) / len(data)

    #function to calculate the standard deviation of a list of numbers
    def stddev(data):
        m = mean(data)
        return sqrt(sum((x - m) ** 2 for x in data) / (len(data) - 1))

    def t_test(sample1, sample2):
        # calculate mean and standard deviation for each sample
        x1 = mean(sample1)
        x2 = mean(sample2)
        s1 = stddev(sample1)
        s2 = stddev(sample2)
        n = len(sample1)
        # calculate t-value
        t = (x1 - x2) / sqrt((s1 ** 2 + s2 ** 2) / n)
        df = n - 1
        # calculate p-value
        p = (1 - stats.t.cdf(abs(t), df)) * 2
        return t, p

    # Check if the variances of two samples are equal
    def equal_variances(data1, data2):
        # Check if the variances of two samples are equal
        n1 = len(data1)
        n2 = len(data2)
        # calculate variance for each sample
        var1 = sum([(i - sum(data1)/n1)**2 for i in data1])/(n1-1)
```

```
        var2 = sum([(i - sum(data2)/n2)**2 for i in data2])/(n2-1)
        # calculate F-value
        f = var1/var2
        return f

    # Making Variable to go for Paired T-test
    check = False
    # Checking Equal Variances
    p_value = equal_variances(ger_2014_2019, ger_2020_2021)
    if p_value > 0.05:
        print("Equal Variance")
        check = True
    else:
        print("No Equal Variance")
        check = False
    if check == True :
        # perform the paired t-test
        t_stat, p_val = t_test(ger_2014_2019,ger_2020_2021)
    else:
        #Wlech's T-test
        t_stat, p_val = ttest_ind(ger_2014_2019,ger_2020_2021,equal_var =
False)
    # print the results
    print(f"t-statistic for {f}: ", t_stat)
    print(f"p-value for {f}: ", p_val)

    print("-----------Result---------------")
    # checking the p-value against a significance level (0.05)
    if p_value < 0.05:
        print(f"There is not Statistically Significant increase in Gross
Enrolment Ratio in {f} level of education during Covid-19")
    else:
        print(f"There is Statistically Significant increase in Gross
Enrolment Ratio in {f} level of education during Covid-19")


------------ NORMALIZED DATA OF ELEMENTARY LEVEL OF EDUCATION -------------
--
        2014      2015      2016      2017      2018      2019      2020  \
0   0.259564  0.252118  0.287734  0.202468  0.174360  0.108594  0.034713
1   0.200845  0.196385  0.313023  0.302512  0.353473  0.365018  0.360593
2   0.852313  0.794389  0.608921  0.501115  0.487349  0.470986  0.464114
3   0.582963  0.435699  0.588411  0.435558  0.530747  0.501243  0.490550
4   0.393238  0.514027  0.464158  0.319162  0.271737  0.234872  0.274490
5   0.476201  0.423084  0.480884  0.396611  0.372642  0.283366  0.209265
6   0.528915  0.504613  0.580247  0.437193  0.404693  0.336281  0.301791
7   0.890347  0.865562  1.000000  0.744463  0.771814  0.693009  0.601729
8   0.564724  0.528337  0.581043  0.449829  0.434749  0.348991  0.270290
9   0.382117  0.408963  0.511151  0.382488  0.376783  0.313484  0.266584
```

10	0.392349	0.326116	0.491239	0.440910	0.487195	0.437414	0.365040
11	0.460854	0.480136	0.592991	0.475992	0.517712	0.463526	0.416800
12	0.031806	0.100546	0.032656	0.114315	0.148597	0.149212	0.182705
13	0.477980	0.481077	0.361410	0.381448	0.406073	0.330893	0.312909
14	0.389235	0.432875	0.568698	0.457708	0.531360	0.468500	0.416677
15	0.368550	0.392393	0.494425	0.397800	0.429689	0.387400	0.355405
16	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
17	0.537589	0.450574	0.509160	0.361974	0.340285	0.293589	0.244348
18	0.448176	0.463190	0.581641	0.471235	0.501303	0.447361	0.397406
19	0.601423	0.578799	0.532656	0.524008	0.648520	0.621304	0.604941
20	1.000000	1.000000	0.952808	1.000000	1.000000	1.000000	1.000000
21	0.801824	0.772359	0.772999	0.660919	0.769054	0.735839	0.722545
22	0.288034	0.311241	0.123457	0.199792	0.187088	0.183614	0.186658
23	0.418594	0.440595	0.528873	0.385313	0.368655	0.316938	0.310809
24	0.368550	0.344756	0.410195	0.272484	0.247508	0.187206	0.151081
25	0.582518	0.585577	0.706491	0.525792	0.532127	0.527632	0.501668
26	0.327847	0.423084	0.515930	0.424706	0.432296	0.395137	0.382829
27	0.747776	0.617586	0.526085	0.437045	0.366968	0.310307	0.281902
28	0.491326	0.523065	0.636002	0.411179	0.435516	0.371788	0.337122
29	0.394795	0.490303	0.567503	0.462316	0.499463	0.486460	0.461272
30	0.573176	0.549802	0.592792	0.505723	0.543782	0.499586	0.454602
31	0.479760	0.524195	0.627041	0.521481	0.558657	0.515612	0.485608
32	0.366993	0.406326	0.454401	0.349190	0.333078	0.331721	0.321557
33	0.413923	0.451704	0.442652	0.411476	0.451311	0.449157	0.480420

	2021
0	0.000000
1	0.357630
2	0.470323
3	0.470207
4	0.312507
5	0.223186
6	0.308111
7	0.600833
8	0.253153
9	0.267731
10	0.393035
11	0.427166
12	0.241930
13	0.321879
14	0.438158
15	0.367928
16	0.046396
17	0.224922
18	0.405993
19	0.559528
20	1.000000
21	0.790235
22	0.209534

```
23  0.302441
24  0.090362
25  0.467315
26  0.376721
27  0.273863
28  0.341664
29  0.474257
30  0.461530
31  0.508273
32  0.333796
33  0.453893

Equal Variance
t-statistic for ELEMENTARY:  1.7460854881632712
p-value for ELEMENTARY:  0.09010069472406079
-----------Result---------------
There is Statistically Significant increase in Gross Enrolment Ratio in
ELEMENTARY level of education during Covid-19

------------ NORMALIZED DATA OF SECONDARY LEVEL OF EDUCATION --------------
-
        2014      2015      2016      2017      2018      2019      2020  \
0   0.533498  0.569317  0.504313  0.420044  0.315630  0.387004  0.327270
1   0.264691  0.248182  0.359041  0.368722  0.417678  0.446233  0.433239
2   0.339588  0.361952  0.355935  0.284802  0.208839  0.080123  0.149594
3   0.248889  0.235515  0.325914  0.238106  0.286229  0.304773  0.280466
4   0.069300  0.083040  0.155107  0.000000  0.000000  0.006901  0.067291
5   0.601481  0.794980  0.758282  0.787004  0.664643  0.563734  0.529142
6   0.581235  0.666901  0.560041  0.581498  0.547989  0.550508  0.468562
7   0.702881  0.989913  1.000000  1.000000  1.000000  1.000000  1.000000
8   0.832099  0.979123  0.765183  0.795374  0.648520  0.618171  0.553868
9   0.291358  0.258034  0.343340  0.342952  0.349583  0.369753  0.332921
10  0.446091  0.526390  0.574534  0.726652  0.687595  0.685068  0.625927
11  0.853992  1.000000  0.844375  0.953084  0.912178  0.865057  0.718827
12  0.135638  0.028384  0.073844  0.019824  0.031677  0.009967  0.002119
13  0.181728  0.102745  0.050207  0.060793  0.060698  0.049262  0.057577
14  0.392263  0.437016  0.502243  0.467621  0.493741  0.540732  0.545920
15  0.801811  0.988975  0.827812  0.868502  0.764416  0.747939  0.668315
16  1.000000  0.855735  0.734472  0.526211  0.371965  0.095074  0.318262
17  0.392263  0.425053  0.445997  0.440749  0.374621  0.332758  0.204698
18  0.531193  0.630307  0.650276  0.703304  0.634863  0.637723  0.581597
19  0.229136  0.199625  0.206694  0.211674  0.231601  0.290588  0.281703
20  0.256955  0.314098  0.315908  0.322907  0.413316  0.480928  0.444896
21  0.603457  0.750880  0.502243  0.541630  0.675645  0.700403  0.562875
22  0.000000  0.000000  0.000000  0.001322  0.013467  0.000000  0.000000
23  0.324774  0.352334  0.412181  0.382599  0.436457  0.399847  0.438184
24  0.740412  0.773399  0.671153  0.603524  0.482739  0.424765  0.338573
25  0.496296  0.610368  0.617667  0.686784  0.672989  0.860840  0.873543
26  0.306173  0.288529  0.362664  0.375991  0.453907  0.499712  0.443483
```

```
27  0.672757  0.965517  0.738268  0.805286  0.797800  0.790876  0.534264
28  0.630617  0.807882  0.759489  0.654846  0.621965  0.621047  0.580714
29  0.357037  0.421300  0.460145  0.479956  0.513657  0.571785  0.575945
30  0.836214  0.940652  0.743271  0.713436  0.616275  0.600537  0.339809
31  0.582058  0.636406  0.596446  0.590969  0.594461  0.640215  0.563935
32  0.226667  0.196106  0.287612  0.136564  0.146434  0.146636  0.118333
33  0.333333  0.432559  0.396480  0.483040  0.452769  0.538815  0.555811

        2021
0   0.159369
1   0.489941
2   0.117751
3   0.274951
4   0.086785
5   0.582249
6   0.350296
7   1.000000
8   0.442209
9   0.288363
10  0.674556
11  0.661933
12  0.000000
13  0.155227
14  0.674359
15  0.735897
16  0.054832
17  0.185602
18  0.653057
19  0.304734
20  0.484813
21  0.647337
22  0.033136
23  0.390927
24  0.306114
25  0.680868
26  0.368639
27  0.562722
28  0.691321
29  0.661144
30  0.408481
31  0.578107
32  0.171992
33  0.545562
Equal Variance
t-statistic for SECONDARY:  0.9981687018054022
p-value for SECONDARY:  0.3254609081000157
-----------Result---------------
There is Statistically Significant increase in Gross Enrolment Ratio in
SECONDARY level of education during Covid-19
```

---------- NORMALIZED DATA OF HIGHER SECONDARY LEVEL OF EDUCATION -------
----

	2014	2015	2016	2017	2018	2019	2020	\
0	0.910504	0.710659	0.775366	0.645719	0.508014	0.272278	0.330267	
1	0.502862	0.408625	0.000000	0.275974	0.356969	0.406422	0.395794	
2	0.621217	0.458174	0.511083	0.380290	0.207840	0.095383	0.165227	
3	0.316042	0.115461	0.369173	0.151553	0.079094	0.001900	0.000000	
4	0.243954	0.000000	0.216750	0.000000	0.000000	0.000000	0.032294	
5	1.000000	0.965744	0.980462	1.000000	1.000000	0.897207	0.880774	
6	0.573782	0.383392	0.564100	0.402643	0.447561	0.446513	0.475591	
7	0.933170	0.762043	0.802788	0.738366	0.763066	0.800304	0.936913	
8	0.827083	0.731610	0.840151	0.759173	0.793554	0.717272	0.705407	
9	0.457063	0.221746	0.440128	0.294109	0.258014	0.234657	0.178183	
10	0.671223	0.472855	0.644196	0.487277	0.515854	0.587878	0.647766	
11	0.991588	1.000000	0.993144	0.919724	0.965157	1.000000	1.000000	
12	0.550064	0.332773	0.469378	0.365809	0.292509	0.143834	0.333834	
13	0.422362	0.182291	0.309644	0.290314	0.217770	0.189816	0.218926	
14	0.329244	0.173268	0.013483	0.344440	0.313763	0.405282	0.438415	
15	0.816684	0.768925	0.873286	0.793617	0.938676	0.991640	0.973714	
16	0.867508	0.994800	1.000000	0.808941	0.877352	0.698841	0.662411	
17	0.448183	0.229699	0.462980	0.299171	0.301916	0.269998	0.246714	
18	0.667134	0.605903	0.765768	0.667651	0.740767	0.687251	0.674052	
19	0.559878	0.415813	0.555073	0.475467	0.494425	0.499715	0.542433	
20	0.354948	0.208900	0.404707	0.258119	0.266725	0.238647	0.165978	
21	0.547260	0.373605	0.543876	0.364263	0.448432	0.435113	0.408750	
22	0.322117	0.115920	0.362203	0.152538	0.131882	0.061182	0.025911	
23	0.000000	0.084875	0.373743	0.204133	0.687631	0.318450	0.264739	
24	0.822993	0.794617	0.849406	0.651343	0.754181	0.731522	0.666166	
25	0.742026	0.627772	0.779593	0.641361	0.727875	0.769143	0.853736	
26	0.566538	0.423918	0.598035	0.474062	0.524739	0.525176	0.558956	
27	0.665148	0.542438	0.639854	0.569380	0.544425	0.435873	0.510702	
28	0.802664	0.752409	0.866431	0.757205	0.800000	0.805814	0.830267	
29	0.545624	0.395320	0.475777	0.433854	0.524913	0.501235	0.555576	
30	0.428555	0.209818	0.417505	0.262055	0.213589	0.187346	0.253098	
31	0.794836	0.670286	0.789534	0.675945	0.695645	0.733232	0.759294	
32	0.612338	0.424071	0.568899	0.370448	0.343728	0.305529	0.309613	
33	0.503680	0.328491	0.520681	0.421904	0.441463	0.463804	0.492114	

	2021
0	0.507097
1	0.352653
2	0.302129
3	0.071139
4	0.000845
5	0.775262
6	0.545455
7	1.000000

```
8   0.639236
9   0.208854
10  0.671004
11  0.984285
12  0.292666
13  0.179284
14  0.350963
15  0.831531
16  0.448124
17  0.261913
18  0.602399
19  0.574856
20  0.171173
21  0.430382
22  0.000000
23  0.130956
24  0.555424
25  0.780500
26  0.582967
27  0.479385
28  0.770869
29  0.490199
30  0.345556
31  0.725583
32  0.250422
33  0.442210
Equal Variance
t-statistic for HIGHER SECONDARY:  0.6818859031877261
p-value for HIGHER SECONDARY:  0.5000693347432152
-----------Result---------------
There is Statistically Significant increase in Gross Enrolment Ratio in
HIGHER SECONDARY level of education during Covid-19
```

Analysis 1.3:

```
#problem satement 1.3
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import shapiro,t,stats
import numpy as np
import math

# load data into a pandas DataFrame
data = pd.read_csv("lite.csv")
print("--------------------literacy Rate Data-----------------------\n")
print(data)
```

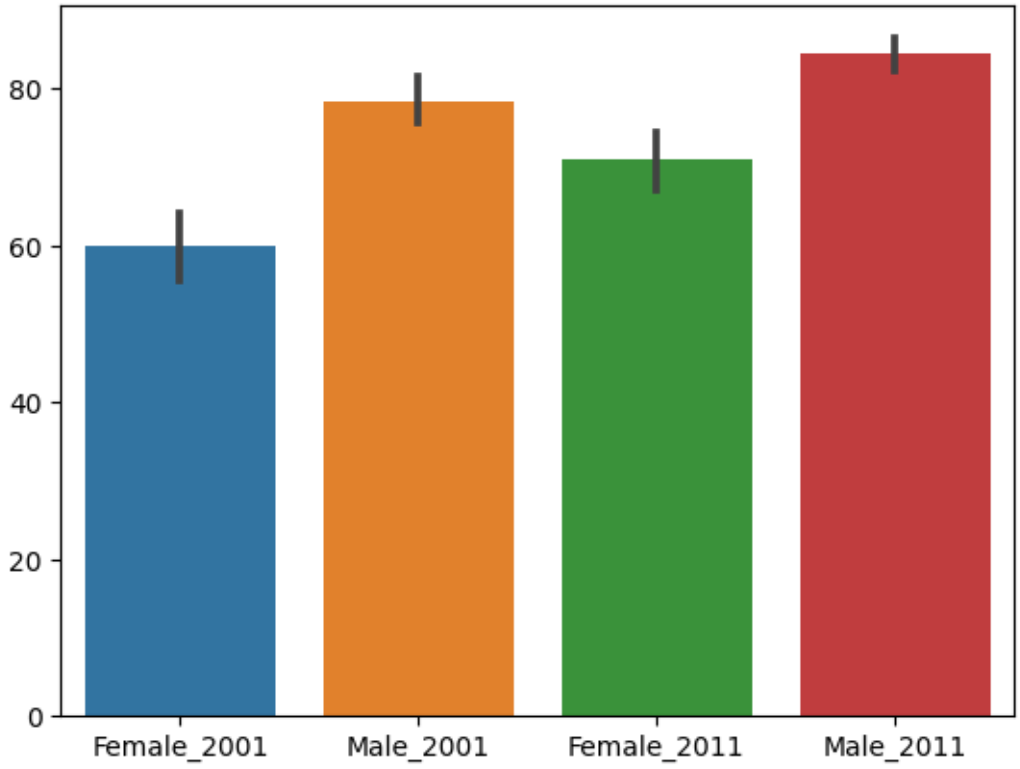
```
def equal_variances(data1, data2):
        n1 = len(data1)
        n2 = len(data2)
        var1 = sum([(i - sum(data1)/n1)**2 for i in data1])/(n1-1)
        var2 = sum([(i - sum(data2)/n2)**2 for i in data2])/(n2-1)
        f = var1/var2
        return f

print("----------------RESULT----------------")
# Checking Normal Distribution
w_,p2_value = shapiro(data["Female_2011"])
if p2_value > 0.05:
    print("Data is Normally Distributed")
else:
    print("Data is Not Normally Distributed")

years = [2001, 2011]
for year in years:
    # Checking Equal Variances
    f = equal_variances(data[f"Female_{year}"],data[f"Male_{year}"])
    if round(f) == 1:
        print("Equal Variance in Data")
    else:
        print("No Equal Variance in Data")

    # creating code of Welch"s T-Test
    female = data[f"Female_{year}"]
    male = data[f"Male_{year}"]
    female_mean = np.mean(female)
    male_mean = np.mean(male)
    female_std = np.std(female)
    male_std = np.std(male)
    degrees_of_freedom = (female_std**2/len(female) +
male_std**2/len(male))**2 / ((female_std**2/len(female))**2/(len(female)-1)
+ (male_std**2/len(male))**2/(len(male)-1))
    t_value = (female_mean - male_mean) /
np.sqrt((female_std**2/len(female)) + (male_std**2/len(male)))
    p_value = t.cdf(t_value,df=degrees_of_freedom)

    print(f"t-statistic for {year} :",t_value)
    print(f"p-value for {year} :",p_value)

     # check the p-value against a significance level (0.05)
    if p_value < 0.05:
        print(f"There is a statistically significant mean difference
between the literacy rate of males and females in all states of India for
the year {year}.")
    else:
```

```
        print(f"There is a not statistically significant mean difference
between the literacy rate of males and females in all states of India for
the year {year}.")

#plotting graph of data to visualize the hypothesis
import seaborn as sns
sns.barplot(data)
plt.show()

--------------------literacy Rate Data-----------------------
```

	State/Union Territory	Female_2001	Male_2001	Female_2011	Male_2011
0	A & N Islands	75.2	86.3	82.4	90.3
1	Andhra Pradesh	50.4	70.3	59.1	74.9
2	Arunachal Pradesh	43.5	63.8	57.7	72.6
3	Assam	54.6	71.3	66.3	77.8
4	Bihar	33.1	59.7	51.5	71.2
5	Chandigarh	76.5	86.1	81.2	90.0
6	Chhattisgarh	51.9	77.4	60.2	80.3
7	D & N Haveli	43.0	73.3	64.3	85.2
8	Daman & Diu	70.4	88.4	79.5	91.5
9	Delhi	74.7	87.3	80.8	90.9
10	Goa	75.4	88.4	84.7	92.6
11	Gujarat	58.6	80.5	69.7	85.8
12	Haryana	45.7	78.5	65.9	84.1
13	Himachal Pradesh	67.4	85.4	75.9	89.5
14	Jammu & Kashmir	43.0	66.6	56.4	76.8
15	Jharkhand	38.9	67.3	55.4	76.8
16	Karnataka	56.9	76.1	68.1	82.5
17	Kerala	87.9	94.2	92.1	96.1
18	Lakshadweep	80.5	92.5	87.9	95.6
19	Madhya Pradesh	50.3	76.1	59.2	78.7
20	Maharashtra	67.0	86.0	75.9	88.4
21	Manipur	60.5	80.3	72.4	86.1
22	Meghalaya	59.6	65.4	72.9	76.0
23	Mizoram	86.8	90.7	89.3	93.3
24	Nagaland	61.5	71.2	76.1	82.8
25	Odisha	50.5	75.4	64.0	81.6
26	Puducherry	73.9	88.6	80.7	91.3
27	Punjab	63.4	75.2	70.7	80.4
28	Rajasthan	43.9	75.7	52.1	79.2
29	Sikkim	60.4	76.0	75.6	86.6
30	Tamil Nadu	64.4	82.4	73.4	86.8
31	Tripura	64.9	81.0	82.7	91.5
32	Uttar Pradesh	42.2	68.8	57.2	77.3
33	Uttarakhand	59.6	83.3	70.0	87.4
34	West Bengal	59.6	77.0	70.5	81.7

```
----------------RESULT----------------
Data is Normally Distributed
```

```
No Equal Variance in Data
t-statistic for 2001 : -6.813969822982744
p-value for 2001 : 3.159992171416495e-09
There is a statistically significant mean difference between the literacy
rate of males and females in all states of India for the year 2001.
No Equal Variance in Data
t-statistic for 2011 : -6.284880058213214
p-value for 2011 : 2.5626609484090095e-08
There is a statistically significant mean difference between the literacy
rate of males and females in all states of India for the year 2011.
```



Appendix 2:

Analysis 2.1:

```
#importing the data of in
from google.colab import files
upload=files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.

```
Saving internetfacility_data1.csv to internetfacility_data1.csv
```

```
# reading the dataset

import pandas as pd
df=pd.read_csv("internetfacility_data1.csv")
df
```

Out[ ]:

	Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	total
0	Andaman & Nicobar Islands	187	140	113	107	106	73	76	78	77	73	414
1	Andhra Pradesh	34744	14975	14514	10954	7337	6410	8114	7887	14907	14639	62702
2	Arunachal Pradesh	794	350	307	272	235	164	164	158	145	118	4047
3	Assam	7126	4206	3836	3662	1604	1367	1227	895	562	458	71042
4	Bihar	10381	8359	7671	5991	4623	1049	1066	1099	1153	939	84236
5	Chandigarh	230	224	223	222	206	161	160	155	151	138	201
6	Chhattisgarh	20735	7922	4891	3078	1603	1605	1545	1313	1020	754	53781
7	Dadra & Nagar Haveli and Daman & Diu	264	176	188	177	124	89	82	71	62	51	492
8	Delhi	5619	4917	4858	4694	5590	3962	3789	2919	2048	1907	5755
9	Goa	879	577	575	601	624	483	471	447	418	318	1554

	Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	total
10	Gujarat	49522	40956	38658	36351	28924	8498	8244	7982	7950	6790	52424
11	Haryana	12158	10545	10172	9705	8701	8037	6720	6200	5055	3190	22315
12	Himachal Pradesh	6215	4216	4146	3938	3056	2260	2126	1946	1774	1664	18039
13	Jammu & Kashmir	8566	3740	3460	3570	1287	989	883	829	768	699	29092
14	Jharkhand	16726	14723	15001	13579	3695	1579	1074	999	874	734	48528
15	Karnataka	22590	19105	19020	9246	6804	6861	6841	6277	5530	4558	75489
16	Kerala	15459	14670	14638	14633	10832	4581	4502	4353	4091	3622	17130
17	Lakshadweep	37	42	42	38	14	14	13	13	13	12	45
18	Madhya Pradesh	34562	20115	18805	17386	15235	13366	11944	5735	8665	6472	150762
19	Maharashtra	52553	40259	39014	37870	30481	17480	16245	14898	12760	10156	107624
20	Manipur	1065	767	719	647	332	335	334	287	267	229	4993
21	Meghalaya	2460	645	572	550	356	342	332	255	173	145	14514
22	Mizoram	307	359	282	272	335	218	222	224	148	135	3825
23	Nagaland	1383	454	434	414	334	242	241	215	207	190	2826
24	Odisha	9284	4660	4328	4300	4318	2806	2331	1221	573	502	70300
25	Puducherry	724	719	484	492	457	387	321	297	276	248	731
26	Punjab	16429	27047	14087	13303	4833	10378	10099	9615	6763	6132	28988

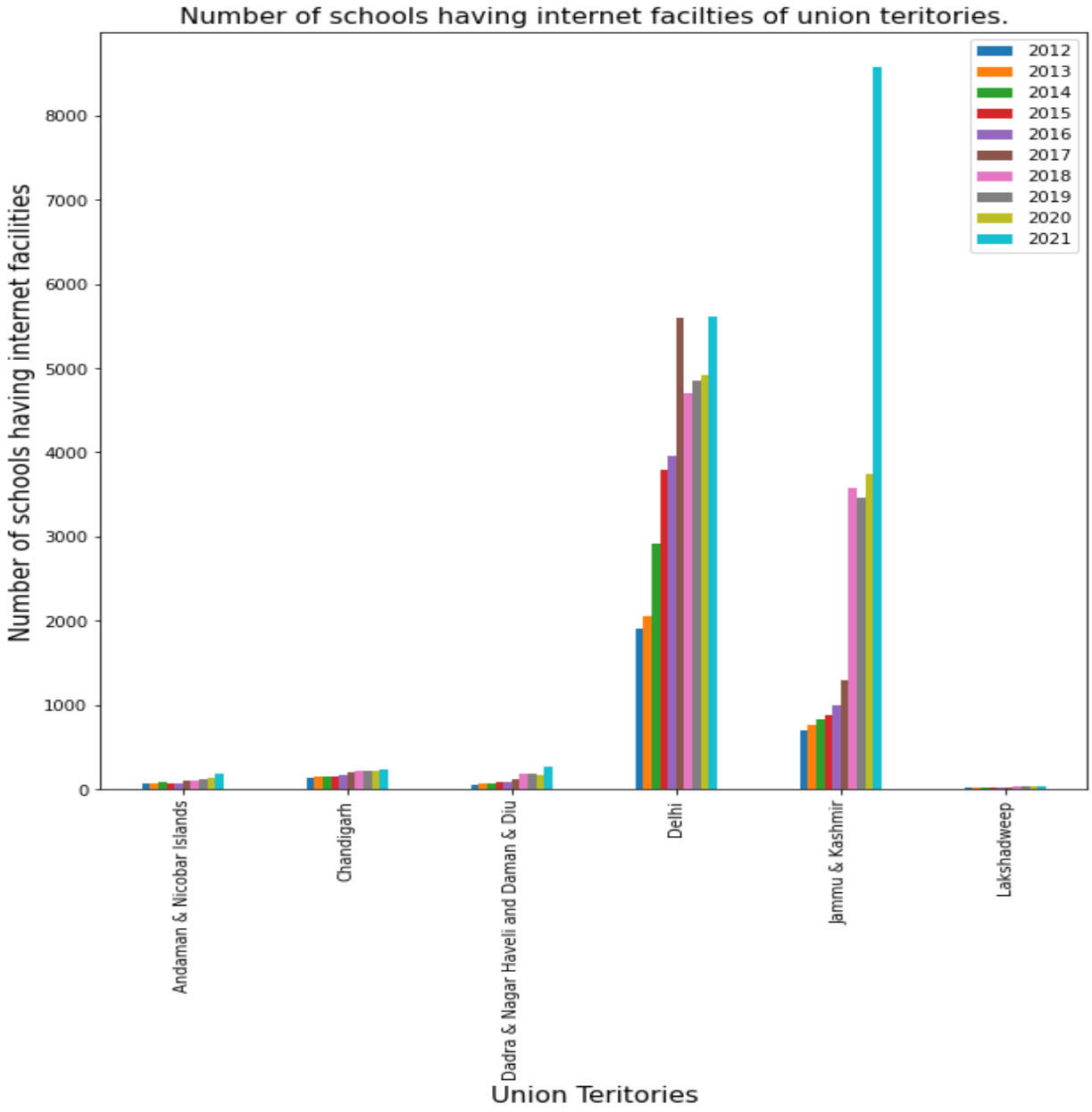
	Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	total
27	Rajasthan	63674	42365	38561	31978	19445	16154	14548	13054	11574	9972	108428
28	Sikkim	434	267	247	241	216	100	100	96	82	77	1279
29	Tamilnadu	22086	18865	18819	14083	16907	11180	10883	10499	9996	8774	57583
30	Tripura	896	274	190	169	128	106	86	74	72	63	4844
31	Uttarakhand	6245	4185	3883	3663	2503	1393	1380	1365	1191	1026	24026
32	Uttar Pradesh	54554	38388	34638	29751	16097	11315	10327	9643	8545	5697	255969
33	West Bengal	15796	10442	9590	6754	7179	4968	4512	3932	2851	2011	95736

In [ ]:

```
#Visualising the data of union teritories
import matplotlib.pyplot as plt

#soting the data of union teritories
df_1=df.loc[(df["Location"]=="Delhi")|
          (df["Location"]=="Andaman & Nicobar Islands")|
          (df["Location"]=="Chandigarh")|
          (df["Location"]=="Dadra & Nagar Haveli and Daman & Diu")|
          (df["Location"]=="Jammu & Kashmir")|
          (df["Location"]=="Ladakh")|
          (df["Location"]=="Lakshadweep")|
          (df["Location"]=="puducherry")]

#ploting the graph between union teritories and the years
df_1.plot(x="Location",y=["2012","2013","2014","2015","2016","2017","2018",
"2019","2020","2021"],kind="bar",figsize=(10,10))
plt.xlabel("Union Teritories",fontsize=15)
plt.ylabel("Number of schools having internet facilities",fontsize=15)
plt.title("Number of schools having internet facilties of union
teritories.",fontsize=15)
plt.xticks(fontsize=10)
plt.yticks(fontsize=10)
plt.legend()
```

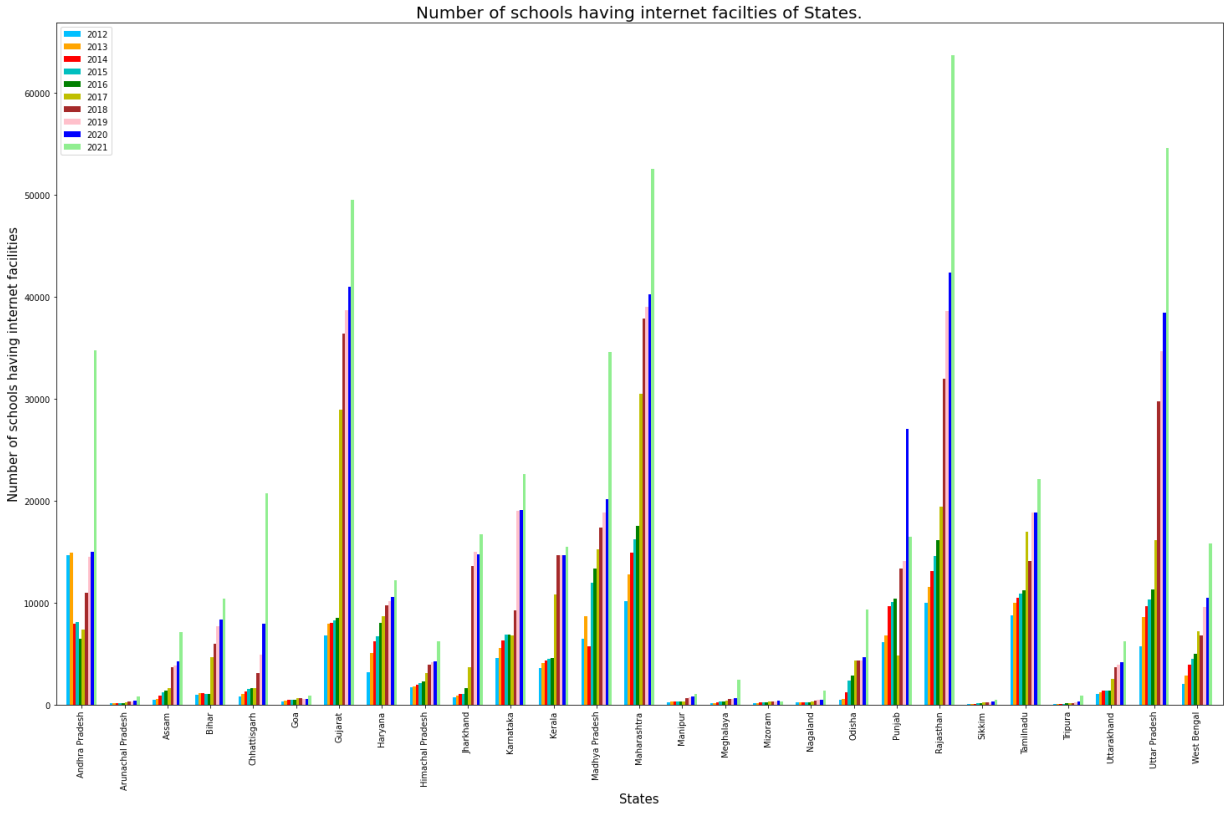
Out[ ]:

```
<matplotlib.legend.Legend at 0x7fd0bb37ff10>
```




```
#visualising data of other states

df_2=df.loc[(df["Location"]=="Andhra Pradesh")|
            (df["Location"]=="Arunachal Pradesh")|
            (df["Location"]=="Assam")|
            (df["Location"]=="Bihar")|
            (df["Location"]=="Chhattisgarh")|
            (df["Location"]=="Goa")|
            (df["Location"]=="Gujarat")|
            (df["Location"]=="Haryana")|
            (df["Location"]=="Himachal Pradesh")|
            (df["Location"]=="Jharkhand")|
            (df["Location"]=="Karnataka")|
            (df["Location"]=="Kerala")|
            (df["Location"]=="Madhya Pradesh")|
```

```
            (df["Location"]=="Maharashtra")|
            (df["Location"]=="Manipur")|
            (df["Location"]=="Meghalaya")|
            (df["Location"]=="Mizoram")|
            (df["Location"]=="Nagaland")|
            (df["Location"]=="Odisha")|
            (df["Location"]=="Punjab")|
            (df["Location"]=="Rajasthan")|
            (df["Location"]=="Sikkim")|
            (df["Location"]=="Tamilnadu")|
            (df["Location"]=="Tripura")|
            (df["Location"]=="Uttarakhand")|
            (df["Location"]=="Uttar Pradesh")|
            (df["Location"]=="West Bengal")]

my_color=["deepskyblue","orange","r","c","green","y","brown","pink","blue",
"lightgreen"]
df_2.plot(x="Location",y=["2012","2013","2014","2015","2016","2017","2018",
"2019","2020","2021"],kind="bar",figsize=(25,15),width=0.7,color=my_color)
plt.xlabel("States",fontsize=15)
plt.ylabel("Number of schools having internet facilities",fontsize=15)
plt.title("Number of schools having internet facilties of
States.",fontsize=20)
plt.xticks(fontsize=10)
plt.yticks(fontsize=10)
plt.legend(fontsize=10)
<matplotlib.legend.Legend at 0x7fd0bc2d47c0>
```



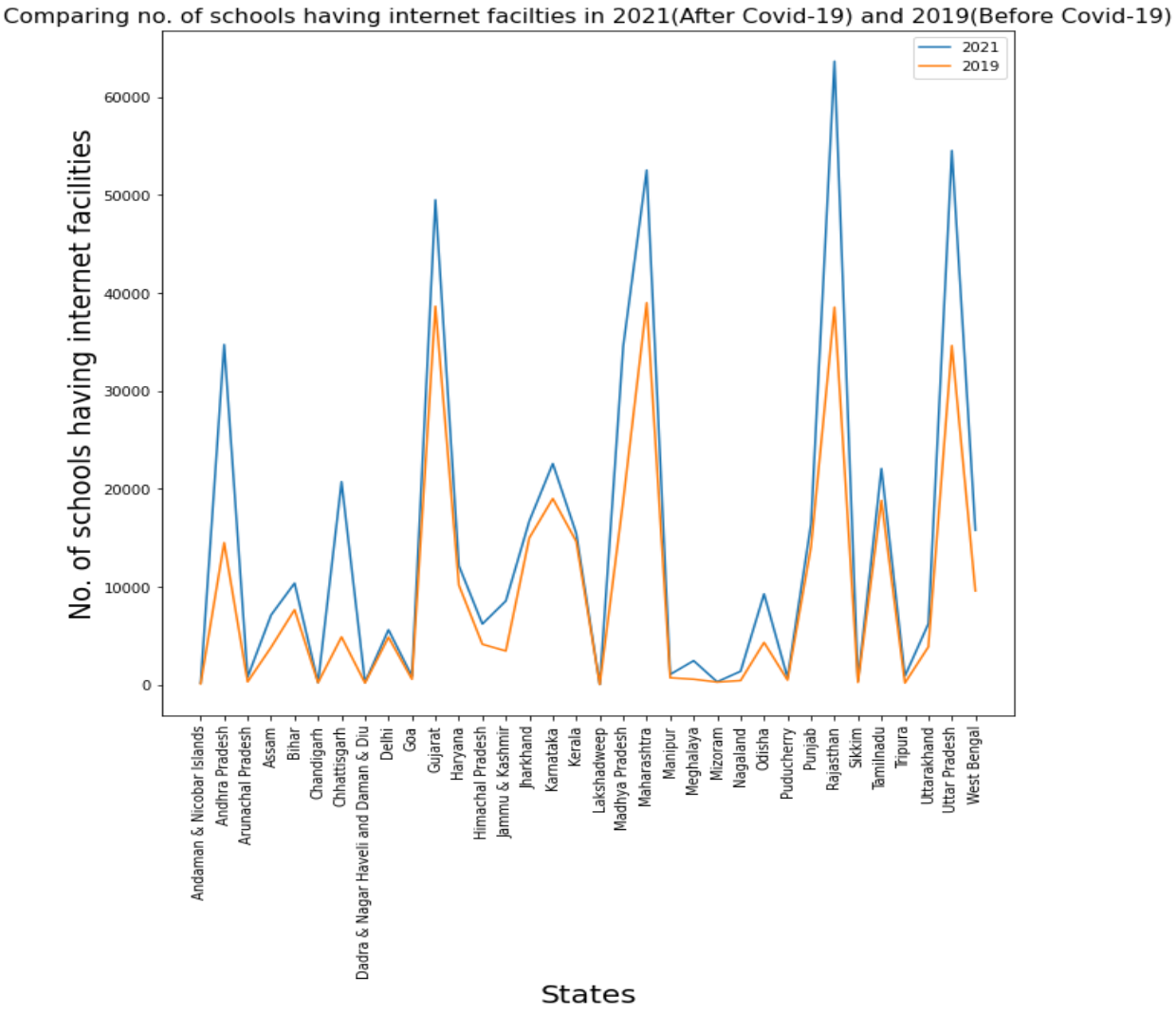
In [ ]:

```
#Test of hypothesis

import scipy.stats as stats
import numpy as np
a=df["2021"].tolist()
b=df["2020"].tolist()

# creating a list of mean of no. of schools of every state of years(2021
and 2020) i.e. after Covid-19.

l1=[]
for i in range(len(a)):
  l1.append((a[i]+b[i])/2)
a1=np.array(l1)
c=df["2019"].tolist()
d=df["2018"].tolist()
e=df["2017"].tolist()
f=df["2016"].tolist()

# creating a list of mean of no. of schools of every state of
years(2019,2018,2017,and 2016) i.e. before Covid-19.

l2=[]
for j in range(len(b)):
  l2.append((c[i]+d[i]+e[i]+f[i])/4)
a2=np.array(l2)

 # calculating the standard deviation

sd1=a1.std()
sd2=a2.std()

# calculating the mean

m1=a1.mean()
m2=a2.mean()

# sample size is more than 30, therefore ztest should be done,

z = (m1-m2) / ((sd1*2 /len(a1)) + (sd2*2 /len(a2)))**0.5
print("zvalue:",z)



#calculating the z_alpha(right tailed test)

z_alpha=stats.norm.ppf(1-.05)
```

```
print("z_alpha:",z_alpha)

#Criteria of rejection

if z>z_alpha:
  print("Reject the null hypothesis")
else:
  print("Accept the null hypothesis")
zvalue: 182.45679201284702
z_alpha: 1.6448536269514722
Reject the null hypothesis
```

In [ ]:

```
#Conclusion graph
#Comparing the number of schools having internet facilties in year 2021 and
2019.

import matplotlib.pyplot as plt
x =df["Location"].tolist()
y1 =df["2021"].tolist()
y2 =df["2019"].tolist()
plt.figure(figsize=(10,10))
plt.plot(x, y1,label="2021")
plt.plot(x, y2,label="2019")
plt.legend(fontsize=10)
plt.xticks(rotation=90,fontsize=10)
plt.yticks(fontsize=10)
plt.xlabel("States",fontsize=20)
plt.ylabel("No. of schools having internet facilities",fontsize=20)
plt.title('Comparing no. of schools having internet facilties in 2021(After
Covid-19) and 2019(Before Covid-19)',fontsize=15)
plt.show()
```



Analysis 2.2:

```
#importing the data of number of schools having library
from google.colab import files
upload=files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.

```
Saving library_data1.csv to library_data1.csv
#reading the dataset

import pandas as pd
df=pd.read_csv("library_data1.csv")
df
```

	Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
0	Andaman & Nicobar Islands	414	415	416	407	410	402	399	393	396	357
1	Andhra Pradesh	60444	57218	56817	56173	24696	58912	60218	61691	98750	100171
2	Arunachal Pradesh	1621	1565	1554	914	1701	1153	1064	1012	745	424
3	Assam	53168	54859	52115	50244	48799	44216	41541	39480	31437	20979
4	Bihar	59021	56647	55699	52740	63301	59041	59019	58245	48608	42575
5	Chandigarh	229	226	223	220	224	197	199	196	190	182
6	Chhattisgarh	55489	54786	51598	48913	52118	50248	49795	51478	46975	45334
7	Dadra & Nagar Haveli and Daman & Diu	454	475	378	480	456	468	477	475	444	419
8	Delhi	5619	5642	5669	5492	5644	5612	5649	5626	5216	4623
9	Goa	1506	1468	1467	1468	1465	1538	1542	1527	1504	1445
10	Gujarat	51693	52157	52186	49984	51717	50461	48710	47746	46042	41757
11	Haryana	23026	22900	22721	22031	22593	22008	21632	21307	21574	18173
12	Himachal Pradesh	17294	17263	17148	16839	17431	17308	17263	17219	17056	15055
13	Jammu & Kashmir	20791	19703	19108	18644	21442	18761	18313	17035	13737	9176
14	Jharkhand	42119	42213	42119	42018	45312	43818	44564	43966	41655	36456
15	Karnataka	73678	73085	71657	70435	72705	73245	73829	73177	72328	70949
16	Kerala	15987	16098	16077	15885	16522	16585	16558	16452	15681	13812

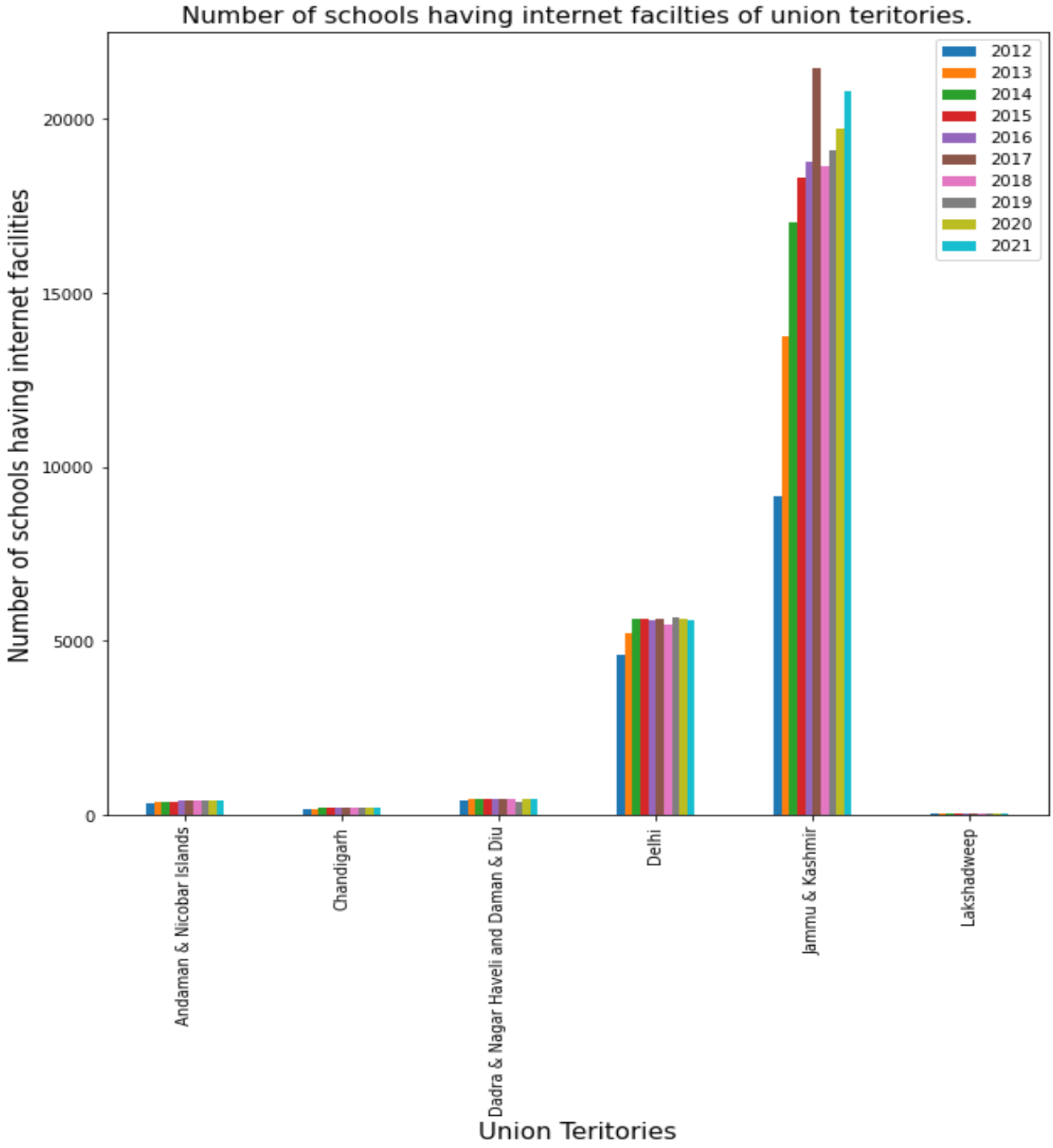
	Location	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
17	Lakshadweep	38	45	45	45	45	45	44	44	45	39
18	Madhya Pradesh	119074	124880	123847	142792	142553	137494	135706	133790	121113	85538
19	Maharashtra	107155	107242	106723	104989	106987	104474	103356	102451	98712	86587
20	Manipur	1187	1185	1156	1054	1431	1430	1382	1420	1240	1098
21	Meghalaya	3468	2078	1781	1282	3771	1797	1734	1660	1187	904
22	Mizoram	3230	3210	3152	3032	3130	3016	2959	2791	2588	1306
23	Nagaland	1966	1513	1081	1060	1310	1138	1155	1007	926	537
24	Odisha	58960	58868	63411	65045	65287	65545	64867	62634	57900	51579
25	Puducherry	733	741	741	739	731	736	730	730	722	677
26	Punjab	27577	28695	28775	28361	27730	28055	27895	27900	27807	25803
27	Rajasthan	82439	81082	79045	75802	82369	76995	77544	74029	72623	65514
28	Sikkim	1132	1139	1141	1142	986	776	772	603	454	342
29	Tamilnadu	58801	58896	58897	56210	56111	57632	57129	56688	55321	53982
30	Telangana	39510	39444	39382	39154	39720	39793	39117	42606	1445	548
31	Uttarakhand	20886	20782	20313	19364	22204	21628	21534	21348	20101	18831
32	Uttar Pradesh	203962	198365	186016	173456	215019	192940	190010	188625	181533	185041
33	West Bengal	82811	82459	81635	80650	76523	74304	72365	69489	54197	37696

In [ ]:

```
#Visualising the data of union teritories
import matplotlib.pyplot as plt

#soting the data of union teritories
df_1=df.loc[(df["Location"]=="Delhi")|
          (df["Location"]=="Andaman & Nicobar Islands")|
          (df["Location"]=="Chandigarh")|
          (df["Location"]=="Dadra & Nagar Haveli and Daman & Diu")|
          (df["Location"]=="Jammu & Kashmir")|
          (df["Location"]=="Ladakh")|
          (df["Location"]=="Lakshadweep")|
          (df["Location"]=="puducherry")]

#ploting the graph between union teritories and the years
df_1.plot(x="Location",y=["2012","2013","2014","2015","2016","2017","2018",
"2019","2020","2021"],kind="bar",figsize=(10,10))
plt.xlabel("Union Teritories",fontsize=15)
plt.ylabel("Number of schools having internet facilities",fontsize=15)
plt.title("Number of schools having internet facilties of union
teritories.",fontsize=15)
plt.xticks(fontsize=10)
plt.yticks(fontsize=10)
plt.legend()
```

Out[ ]:

```
<matplotlib.legend.Legend at 0x7ff7034be760>
```

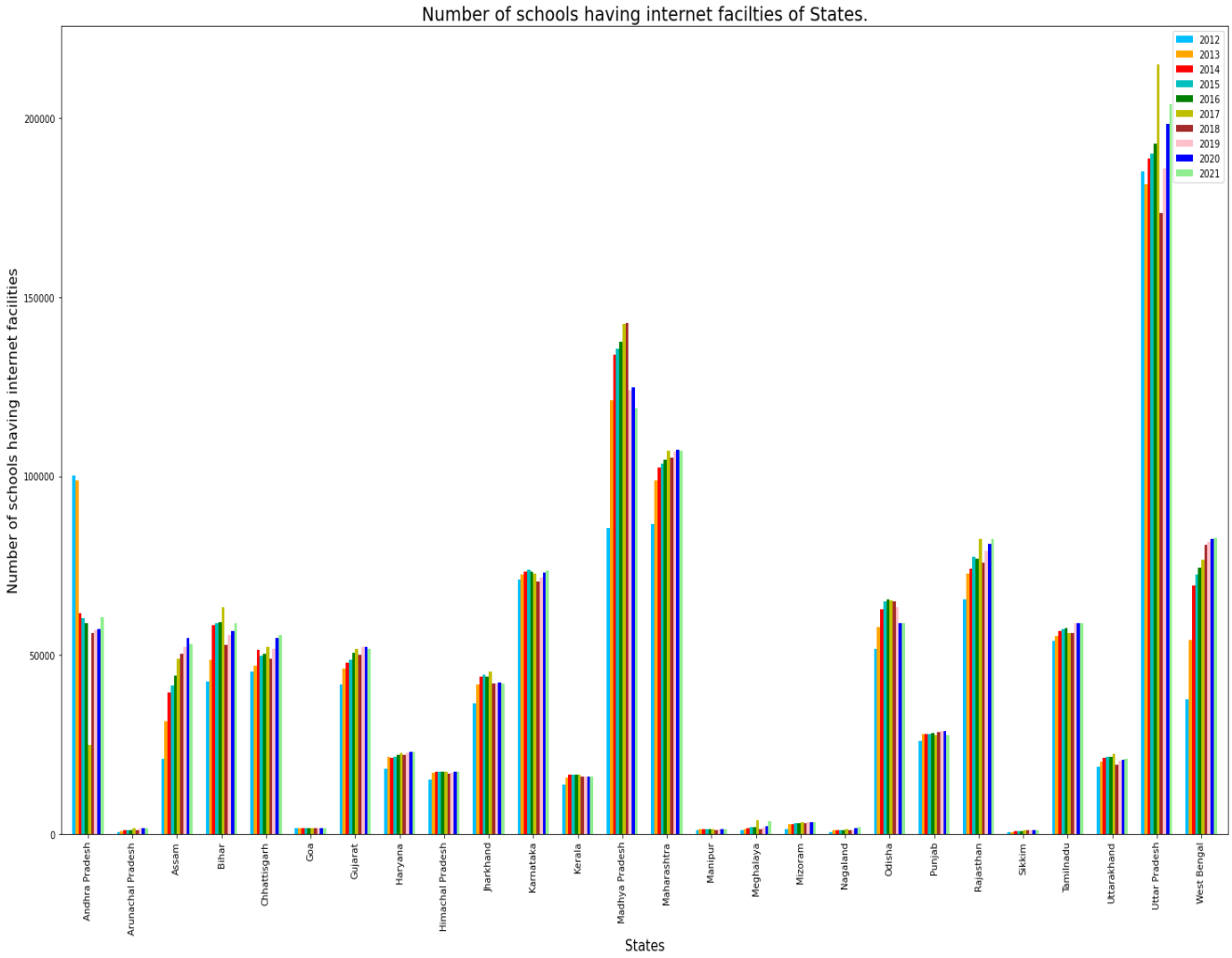



```
#visualising data of other states

df_2=df.loc[(df["Location"]=="Andhra Pradesh")|
            (df["Location"]=="Arunachal Pradesh")|
            (df["Location"]=="Assam")|
            (df["Location"]=="Bihar")|
            (df["Location"]=="Chhattisgarh")|
            (df["Location"]=="Goa")|
            (df["Location"]=="Gujarat")|
            (df["Location"]=="Haryana")|
```

```
            (df["Location"]=="Himachal Pradesh")|
            (df["Location"]=="Jharkhand")|
            (df["Location"]=="Karnataka")|
            (df["Location"]=="Kerala")|
            (df["Location"]=="Madhya Pradesh")|
            (df["Location"]=="Maharashtra")|
            (df["Location"]=="Manipur")|
            (df["Location"]=="Meghalaya")|
            (df["Location"]=="Mizoram")|
            (df["Location"]=="Nagaland")|
            (df["Location"]=="Odisha")|
            (df["Location"]=="Punjab")|
            (df["Location"]=="Rajasthan")|
            (df["Location"]=="Sikkim")|
            (df["Location"]=="Tamilnadu")|
            (df["Location"]=="Tripura")|
            (df["Location"]=="Uttarakhand")|
            (df["Location"]=="Uttar Pradesh")|
            (df["Location"]=="West Bengal")]

my_color=["deepskyblue","orange","r","c","green","y","brown","pink","blue",
"lightgreen"]
df_2.plot(x="Location",y=["2012","2013","2014","2015","2016","2017","2018",
"2019","2020","2021"],kind="bar",figsize=(25,15),width=0.7,color=my_color)
plt.xlabel("States",fontsize=15)
plt.ylabel("Number of schools having internet facilities",fontsize=15)
plt.title("Number of schools having internet facilties of
States.",fontsize=20)
plt.xticks(fontsize=10)
plt.yticks(fontsize=10)
plt.legend(fontsize=10)
```

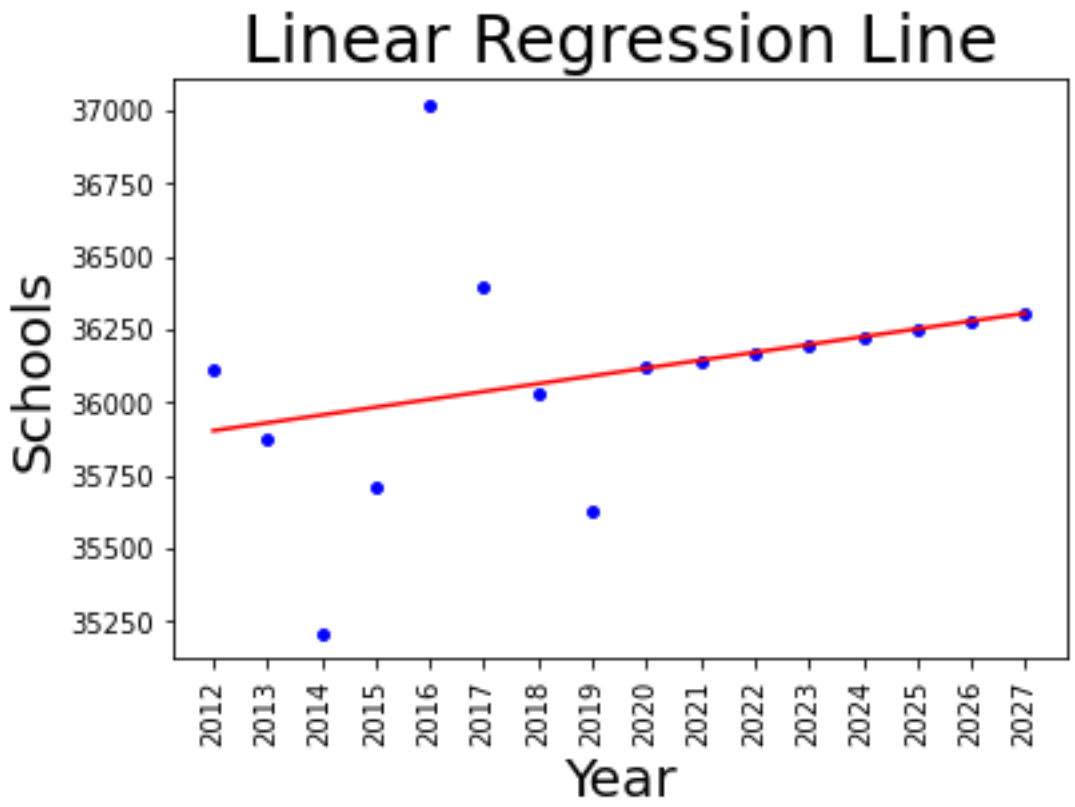
Out[ ]:

```
<matplotlib.legend.Legend at 0x7ff703566f40>
```



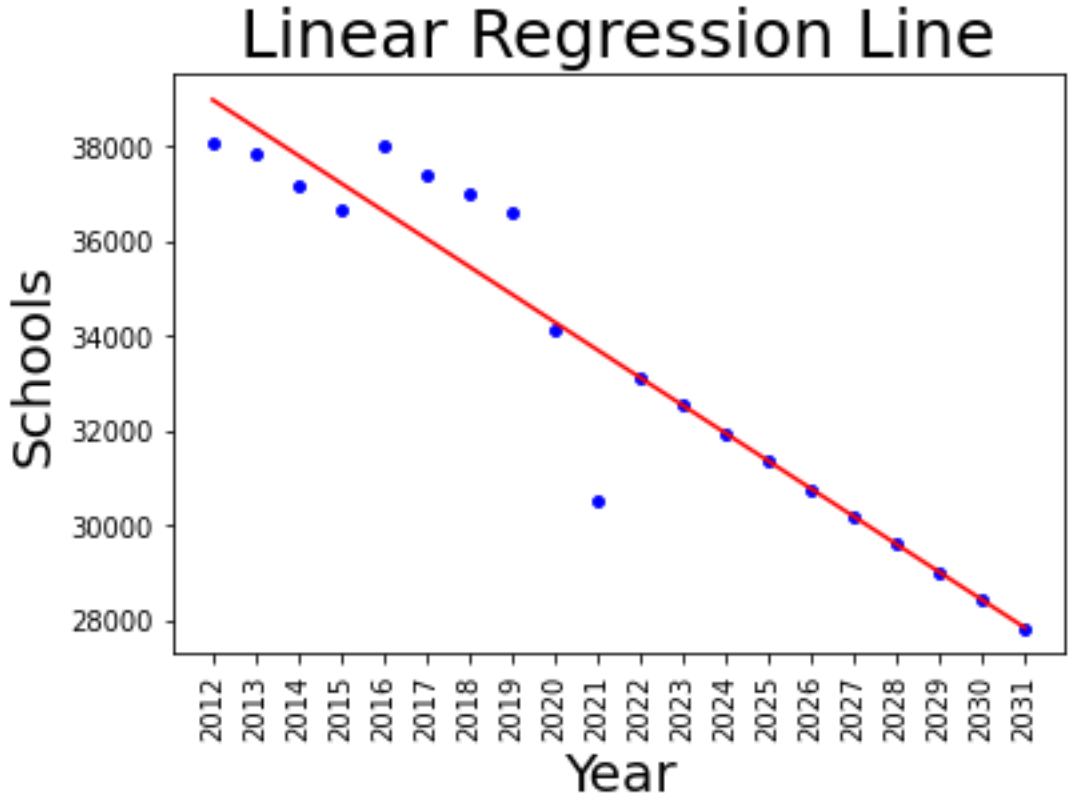

```
import pandas as pd
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
import numpy as np



# Create a pandas DataFrame with the data set
numeric_columns=df.select_dtypes(include=['float64', 'int64'])
schools =numeric_columns.mean(axis=0)
data={"years":list(range(2012,2022)),
      "schools":list(schools)}
df1=pd.DataFrame(data)


#Sorting data of years 2012 to 2019 i.e. before COVID-19 and creating its
Dataframe.
data={"years":list(range(2012,2020)),

"schools":[36109.13888888889,35877.13888888889,35211.13888888889,35706.2285
7142857,37020.885714285716,36397.28571428572,36033.28571428572,35623.6]}
```

```
df2=pd.DataFrame(data)


# Check if the data is of correct data types
df2.dtypes


# Convert data types if necessary
df2['years'] = df2['years'].astype(int)
df2['schools'] = df2['schools'].astype(float)


# Finding regression
x = (df2[['years']]).values
y = (df2['schools']).values
model = LinearRegression()
model.fit(x, y)


# Prediction
years = list(range(2020, 2028))
predictions = model.predict([[year1] for year1 in years])


# Creating dataframe of predictions
pred={'years':list(range(2020,2028)),
     'schools':list(predictions)}
df3=pd.DataFrame(pred)

# Joing two datasets
df4=pd.concat([df2,df3])


#Getting the values
X = (df4[['years']]).values
Y = (df4['schools']).values


#Plotting the graph
plt.scatter(X, Y, color='blue',s=15)
plt.plot(X, model.predict(X), color='red')
plt.xlabel('Year',fontsize=20)
plt.xticks(df4['years'], rotation=90)
plt.ylabel('Schools',fontsize=20)
plt.title('Linear Regression Line',fontsize=25)
plt.show()
```




```
import pandas as pd
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
import numpy as np



# Create a pandas DataFrame with the data set
numeric_columns=df.select_dtypes(include=['float64', 'int64'])
schools =numeric_columns.mean(axis=0)
data={"years":list(range(2012,2022)),
      "schools":list(schools)}
df1=pd.DataFrame(data)


# Check if the data is of correct data types
df1.dtypes

# Convert data types if necessary
df1['years'] = df1['years'].astype(int)
df1['schools'] = df1['schools'].astype(float)


# Finding regression
x = (df1[['years']]).values
y = (df1['schools']).values
```

```
model = LinearRegression()
model.fit(x, y)

# Prediction
years = list(range(2022, 2032))
predictions = model.predict([[year1] for year1 in years])

#Creating dataframe of the prdicted data
pred={'years':list(range(2022,2032)),
     'schools':list(predictions)}
df2=pd.DataFrame(pred)

#Join two datasets
df3=pd.concat([df1,df2])

#Getting values from the dataframe
X = (df3[['years']]).values
Y = (df3['schools']).values

#Plotting graph
plt.scatter(X, Y, color='blue',s=15)
plt.plot(X, model.predict(X), color='red')
plt.xlabel('Year',fontsize=20)
plt.xticks(df3['years'], rotation=90)
plt.ylabel('Schools',fontsize=20)
plt.title('Linear Regression Line',fontsize=25)
plt.show()
```



Appendix 3:

Analysis 3.1:

```
#PROBLEM STATEMENT 1 (RANKING) -To determine the rank of the total number
of enrolment of children with different types of disabilities in India of
year 2021.

import numpy as np
A=np.array([[0,1,1,1,1,1,1,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,0,2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,1,0,2,1,2,1,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,1,1,0,1,2,1,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,2,2,2,0,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,1,1,1,1,0,1,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [2,1,2,2,1,2,0,2,2,2,2,2,1,2,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,0,2,1,2,2,1,2,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,1,0,1,2,2,1,1,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,2,2,0,2,2,1,2,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,2,1,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,2,0,1,1,2,2,2,2,1,2,2],
            [2,2,2,2,2,2,2,2,2,2,2,2,0,2,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,1,1,2,1,2,2,1,0,2,2,2,2,2,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,0,1,1,2,1,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,0,2,2,1,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,1,0,2,1,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,2,2,1,1,2,2,2,2,0,2,2],
            [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1],
            [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,0]])
V=np.array([[1],[0],[1],[0],[1],[0],[1],[0],[1],[0],[1],[0],[1],[0],[1],[0]
,[1],[0],[1],[0],[1]])
for i in range (10):
    print("iteration no.",i+1)
    temp=np.matmul(A,V)
    print (temp)
    V=temp/np.max(temp)
    print(np.max(temp))
    print(V)
    print()
iteration no. 1
[[16]
 [20]
 [17]
 [18]
 [19]
 [18]
 [18]
 [17]
```

```
 [15]
 [17]
 [11]
 [15]
 [20]
 [17]
 [12]
 [15]
 [13]
 [12]
 [14]
 [11]
 [10]]
20
[[0.8 ]
 [1.  ]
 [0.85]
 [0.9 ]
 [0.95]
 [0.9 ]
 [0.9 ]
 [0.85]
 [0.75]
 [0.85]
 [0.55]
 [0.75]
 [1.  ]
 [0.85]
 [0.6 ]
 [0.75]
 [0.65]
 [0.6 ]
 [0.7 ]
 [0.55]
 [0.5 ]]

iteration no. 2
[[24.4 ]
 [28.55]
 [26.95]
 [26.  ]
 [29.6 ]
 [25.1 ]
 [27.75]
 [22.65]
 [21.15]
 [23.5 ]
 [17.35]
 [19.7 ]
```

```
 [30.5 ]
 [21.8 ]
 [17.85]
 [18.95]
 [18.4 ]
 [16.7 ]
 [20.5 ]
 [15.7 ]
 [16.3 ]]
30.5
[[0.8       ]
 [0.93606557]
 [0.88360656]
 [0.85245902]
 [0.9704918 ]
 [0.82295082]
 [0.90983607]
 [0.74262295]
 [0.69344262]
 [0.7704918 ]
 [0.56885246]
 [0.64590164]
 [1.        ]
 [0.7147541 ]
 [0.5852459 ]
 [0.62131148]
 [0.60327869]
 [0.54754098]
 [0.67213115]
 [0.5147541 ]
 [0.53442623]]

iteration no. 3
[[22.80491803]
 [26.93770492]
 [25.19672131]
 [24.37540984]
 [27.83934426]
 [23.58196721]
 [26.05409836]
 [21.34918033]
 [19.99016393]
 [22.06393443]
 [16.41803279]
 [18.71967213]
 [28.78032787]
 [20.66229508]
 [16.9704918 ]
 [18.12295082]
```

```
 [17.53770492]
 [15.89180328]
 [19.33934426]
 [14.87540984]
 [15.3704918 ]]
28.780327868852456
[[0.79237867]
 [0.9359763 ]
 [0.87548416]
 [0.84694691]
 [0.96730463]
 [0.81937799]
 [0.90527455]
 [0.74179768]
 [0.69457735]
 [0.76663249]
 [0.57046024]
 [0.6504329 ]
 [1.        ]
 [0.71793119]
 [0.58965596]
 [0.62969925]
 [0.60936432]
 [0.55217589]
 [0.671964  ]
 [0.51686033]
 [0.53406243]]

iteration no. 4
[[22.84159262]
 [26.93745728]
 [25.2171907 ]
 [24.39878104]
 [27.84210526]
 [23.60697198]
 [26.06288448]
 [21.38374345]
 [20.01845523]
 [22.10070631]
 [16.42099567]
 [18.74020278]
 [28.77671451]
 [20.68967874]
 [16.9722602 ]
 [18.13123718]
 [17.54220779]
 [15.88710412]
 [19.36910458]
 [14.87149692]
```

```
 [15.37115516]]
28.77671451355662
[[0.79375262]
 [0.93608522]
 [0.87630541]
 [0.84786542]
 [0.96752203]
 [0.8203498 ]
 [0.90569354]
 [0.7430919 ]
 [0.6956477 ]
 [0.76800659]
 [0.57063483]
 [0.65122802]
 [1.        ]
 [0.71897293]
 [0.58979145]
 [0.63006627]
 [0.60959731]
 [0.55208193]
 [0.67308256]
 [0.51678926]
 [0.53415254]]

iteration no. 5
[[22.860108  ]
 [26.96174219]
 [25.23952304]
 [24.42009762]
 [27.86639061]
 [23.62726344]
 [26.08644033]
 [21.39967023]
 [20.0324938 ]
 [22.11784744]
 [16.43310623]
 [18.75260291]
 [28.80143467]
 [20.70481627]
 [16.98458445]
 [18.14369839]
 [17.55457004]
 [15.8995772 ]
 [19.38197639]
 [14.88392808]
 [15.38335405]]
28.801434667975705
[[0.79371421]
 [0.93612497]
```

```
 [0.87632867]
 [0.84787782]
 [0.96753481]
 [0.82035023]
 [0.90573406]
 [0.74300709]
 [0.69553805]
 [0.76794256]
 [0.57056554]
 [0.65109961]
 [1.        ]
 [0.71888142]
 [0.58971314]
 [0.62995815]
 [0.60950332]
 [0.55204115]
 [0.67295177]
 [0.51677732]
 [0.53411763]]

iteration no. 6
[[22.85814406]
 [26.95973829]
 [25.23747186]
 [24.41804489]
 [27.86445342]
 [23.62522226]
 [26.08439515]
 [21.39790153]
 [20.03095109]
 [22.11597316]
 [16.43213208]
 [18.75133816]
 [28.79952304]
 [20.70314578]
 [16.98355003]
 [18.14252148]
 [17.55347298]
 [15.89861531]
 [19.38058561]
 [14.88298421]
 [15.38242121]]
28.799523043723283
[[0.7936987 ]
 [0.93611753]
 [0.87631562]
 [0.84786282]
 [0.96753177]
 [0.8203338 ]
```

```
 [0.90572316]
 [0.742995  ]
 [0.69553065]
 [0.76792845]
 [0.57056959]
 [0.65109891]
 [1.        ]
 [0.71887113]
 [0.58971636]
 [0.6299591 ]
 [0.60950568]
 [0.5520444 ]
 [0.67294815]
 [0.51677884]
 [0.53412069]]

iteration no. 7
[[22.85801863]
 [26.95953391]
 [25.23729704]
 [24.41788701]
 [27.8642372 ]
 [23.62508223]
 [26.08420511]
 [21.39779888]
 [20.03086144]
 [22.11586042]
 [16.43202471]
 [18.75124612]
 [28.79930073]
 [20.70305161]
 [16.98344753]
 [18.14242684]
 [17.55337457]
 [15.89850551]
 [19.3804958 ]
 [14.88287152]
 [15.38230852]]
28.799300733176832
[[0.79370047]
 [0.93611766]
 [0.87631631]
 [0.84786389]
 [0.96753173]
 [0.82033527]
 [0.90572356]
 [0.74299717]
 [0.69553291]
 [0.76793047]
```

```
 [0.57057027]
 [0.65110074]
 [1.        ]
 [0.71887341]
 [0.58971736]
 [0.62996067]
 [0.60950697]
 [0.55204484]
 [0.67295022]
 [0.51677892]
 [0.5341209 ]]

iteration no. 8
[[22.85805813]
 [26.95958045]
 [25.23734192]
 [24.41793046]
 [27.86428403]
 [23.6251238 ]
 [26.08425099]
 [21.3978338 ]
 [20.03089173]
 [22.11589767]
 [16.43204814]
 [18.75127294]
 [28.79934749]
 [20.70308414]
 [16.98347132]
 [18.14245234]
 [17.55339906]
 [15.89852872]
 [19.3805242 ]
 [14.88289482]
 [15.38233176]]
28.79934748670809
[[0.79370056]
 [0.93611775]
 [0.87631645]
 [0.84786402]
 [0.96753178]
 [0.82033538]
 [0.90572368]
 [0.74299717]
 [0.69553283]
 [0.76793051]
 [0.57057015]
 [0.65110062]
 [1.        ]
 [0.71887338]
```

```
 [0.58971723]
 [0.62996053]
 [0.60950683]
 [0.55204475]
 [0.67295011]
 [0.51677889]
 [0.53412084]]

iteration no. 9
[[22.85805679]
 [26.95957968]
 [25.23734086]
 [24.41792927]
 [27.86428341]
 [23.62512252]
 [26.08425008]
 [21.39783249]
 [20.03089062]
 [22.11589632]
 [16.43204782]
 [18.7512721 ]
 [28.79934697]
 [20.70308291]
 [16.9834709 ]
 [18.14245165]
 [17.55339852]
 [15.89852846]
 [19.38052322]
 [14.8828946 ]
 [15.38233153]]
28.799346972184512
[[0.79370052]
 [0.93611774]
 [0.87631643]
 [0.84786399]
 [0.96753178]
 [0.82033535]
 [0.90572366]
 [0.74299714]
 [0.69553281]
 [0.76793048]
 [0.57057015]
 [0.6511006 ]
 [1.        ]
 [0.71887335]
 [0.58971722]
 [0.62996052]
 [0.60950682]
 [0.55204476]
```

```
 [0.67295009]
 [0.51677889]
 [0.53412084]]

iteration no. 10
[[22.85805632]
 [26.95957907]
 [25.23734029]
 [24.41792873]
 [27.86428277]
 [23.62512202]
 [26.08424948]
 [21.39783208]
 [20.03089027]
 [22.11589589]
 [16.4320475 ]
 [18.75127178]
 [28.79934633]
 [20.70308253]
 [16.98347059]
 [18.14245134]
 [17.55339821]
 [15.89852815]
 [19.38052289]
 [14.88289427]
 [15.38233121]]
28.799346330518308
[[0.79370053]
 [0.93611774]
 [0.87631643]
 [0.84786399]
 [0.96753178]
 [0.82033536]
 [0.90572366]
 [0.74299714]
 [0.69553281]
 [0.76793048]
 [0.57057016]
 [0.6511006 ]
 [1.        ]
 [0.71887335]
 [0.58971722]
 [0.62996053]
 [0.60950683]
 [0.55204476]
 [0.6729501 ]
 [0.51677889]
 [0.53412085]]
```

Analysis 3.2:

```
#PROBLEM STATEMENT 2 (REGRESSION)-To predict the number of schools having
adapted infrastructure for students with disabilities for the next 10 years
by analyzing current infrastructure for students with disabilities from
year 2014 to 2021 of different states of India.
from google.colab import files
uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.

```
Saving new ramps.xlsx to new ramps.xlsx
```

In [ ]:

```
import pandas as pd
df = pd.read_excel('new ramps.xlsx')
df
```

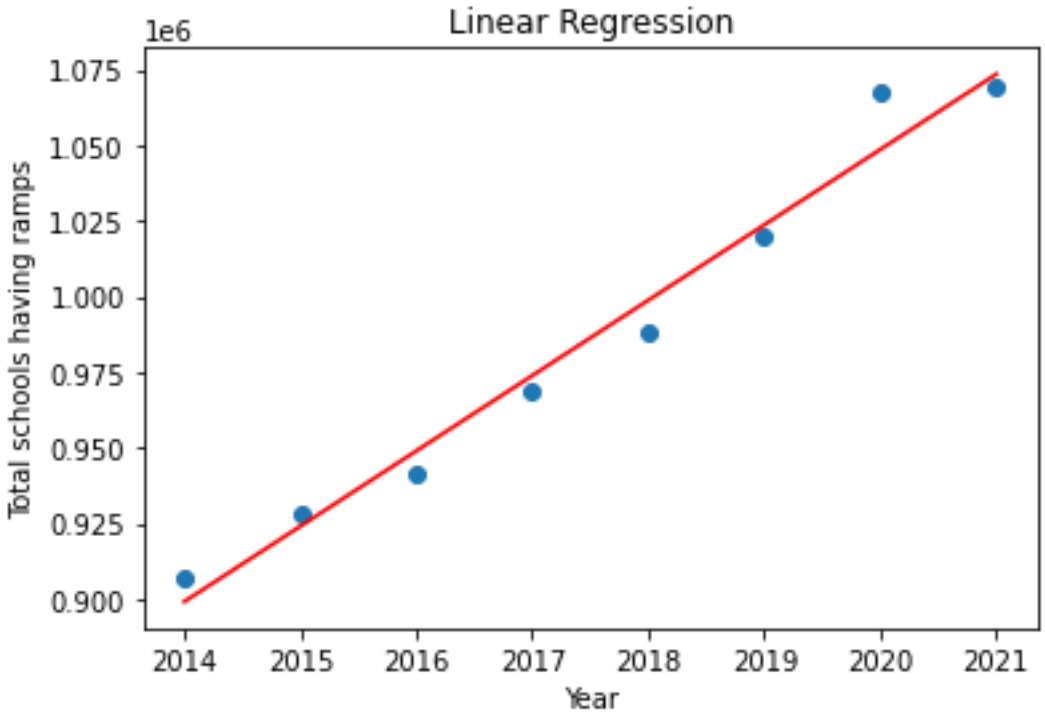
Out[ ]:

	Year	Total schools having ramps
0	2021	1069795
1	2020	1067752
2	2019	1019964
3	2018	988097
4	2017	968345
5	2016	941498
6	2015	928569
7	2014	906593

In [ ]:

```
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
import matplotlib.pyplot as plt

# Read the data
years = [year for year in range(2014, 2022)]
Ramps = [906593,928569,941498,968345,988097,1019964,1067752,1069795]

# Calculate mean of x and y
x_mean = np.mean(years)
y_mean = np.mean(Ramps)
```

```
# Calculate values of b1 and b0
numerator=sum([(years[i]-x_mean)*(Ramps[i]-y_mean)for i in
range(len(years))])
denominator = sum([(years[i] -x_mean)**2 for i in range(len(years))])
b1 = numerator/denominator
b0 = y_mean-(b1*x_mean)

# Print coefficients
print("b1:", b1)
print("b0:", b0)

# Predict values of next 10 years
future_years = [year+1 for year in range(2022, 2032)]
future_predictions = [b0 + b1*year for year in future_years]
y_pred = [b0 + b1*year for year in years]
SSE = sum([(Ramps[i] - y_pred[i])**2 for i in range(len(years))])
SST = sum([(Ramps[i] - y_mean)**2 for i in range(len(years))])

#Calculate R-squared
r_squared = 1 - (SSE/SST)
print("R-squared:", r_squared)

print(future_predictions)
plt.scatter(years, Ramps)
plt.plot(years, [b0 + b1*year for year in years], color='red')
plt.xlabel('Year')
plt.ylabel('Total schools having ramps')
plt.title('Linear Regression')
plt.show()
b1: 24922.369047619046
b0: -49294552.928571425
R-squared: 0.9749316501376944
[1123399.654761903, 1148322.0238095224, 1173244.3928571418,
1198166.7619047612, 1223089.1309523806, 1248011.5, 1272933.8690476194,
1297856.2380952388, 1322778.6071428582, 1347700.9761904776]
```



Analysis 3.3:

```
#PROBLEM STATEMENT 3 (CORRELATION) - To analyze the relation between the
number of male and female teachers trained for teaching Children with
special needs (CWSN) from 2014 to 2021 of India.
from google.colab import files
uploaded =files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.

```
Saving CORRELATION 1.5.xlsx to CORRELATION 1.5.xlsx
import pandas as pd
df=pd.read_excel("CORRELATION 1.5.xlsx")
df
```

Out[ ]:

	Year	No.of male teachers	No.of female teachers
0	2014	73569	51501
1	2015	74542	53492
2	2016	74886	54490
3	2017	75688	56997
4	2018	193760	193235

	Year	No.of male teachers	No.of female teachers
5	2019	152120	154942
6	2020	148198	152621
7	2021	2167727	1728878

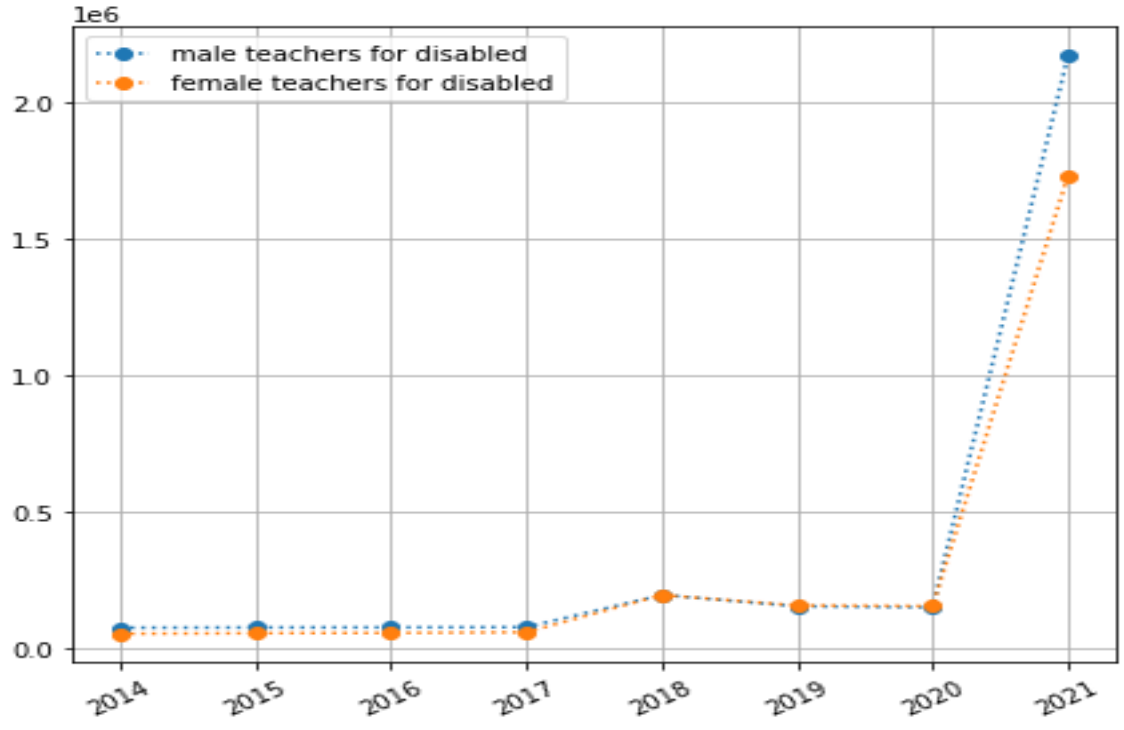
In [ ]:

```
col1 = "No.of male teachers"
col2 = "No.of female teachers"
correlation = df[col1].corr(df[col2])
print(correlation)
```

```
0.9993640854586129
```

In [ ]:

```
import matplotlib.pyplot as plt
import numpy as np
fig = plt.figure(figsize=(7,6))
plt.plot(df['Year'],df['No.of male
teachers'],linestyle=':',marker='o',label='male teachers for disabled')
plt.plot(df['Year'],df['No.of female
teachers'],linestyle=':',marker='o',label='female teachers for disabled')
plt.xticks(df['Year'],rotation=30)

plt.grid(True)
plt.legend()
plt.show()
```



Appendix 4:

Analysis 4.1:

```
#CODES FOR PROBLEM STATEMENT-1
from google.colab import files
uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving sorted data of numbers of teachers by academic qualifications1.xlsx
to sorted data of numbers of teachers by academic qualifications1.xlsx
```

In [ ]:

```
import pandas as pd
df = pd.read_excel('sorted data of numbers of teachers by academic
qualifications1.xlsx')
df
```

Out[ ]:

	states	total(year2021)	total(year 2020)	total(year 2019)	total(year 2018)	total(yaer 2017)
0	Andaman & Nicobar Islands	13288	13143	13746	24509	22504
1	Andhra Pradesh	860881	81274372	777706	1546606	1357108
2	Arunachal Pradesh	55767	59184	57645	109744	100244
3	Assam	825053	872000	864281	1741164	1572660
4	Bihar	1282514	1355994	1349389	2517608	2475984
5	Chandigarh	25476	27070	28168	54946	53936
6	Chhattisgarh	615135	639857	642095	1285028	1022594
7	Daman & Diu and Dadra & Nagar Haveli	14380	10060	6921	13136	12356
8	Delhi	303200	303866	3757	7258	7182
9	Goa	33613	32011	315294	622672	557032
10	Gujarat	801296	822627	31950	61992	53932
11	Haryana	613336	630859	830199	1634558	1630744
12	Himachal Pradesh	257690	263658	663923	1212774	1146436
13	Jammu & Kashmir	354001	364705	265005	496372	404056

	states	total(year2021)	total(year 2020)	total(year 2019)	total(year 2018)	total(yaer 2017)
14	Jharkhand	482142	521919	363388	733758	690124
15	Karnataka	1024353	1216147	519003	939820	932748
16	Kerala	552054	570353	1208400	1999742	1699761
17	Ladakh	1921	13899	570509	1050174	1094976
18	Lakshadweep	1337679	2330	13479	5358	5716
19	Madhya Pradesh	1658885	3070000000	3147	2345390	2446712
20	Maharashtra	102688	1715551	1314962	3333046	3073032
21	Manipur	118484	111427	1738114	213536	206176
22	Meghalaya	47892	120582	110004	234972	222160
23	Mizoram	70411	49217	120527	96148	94900
24	Nagaland	782860	72104	47798	138346	133724
25	Odisha	31812	916329	71073	1650580	1445160
26	Puducherry	594548	33874	892762	65968	65424
27	Punjab	1905691	572227	35800	1089432	1084972
28	Rajasthan	30217	2056173	576796	3956130	2757908
29	Sikkim	1401710	30919	2087227	61930	60528
30	Tamilnadu	837842	1466713	31579	2601568	2307372
31	Telangana	79361	873080	1444226	1424736	1347344
32	Tripura	287132	79377	833067	189978	195756
33	Uttarakhand	3294068	286319	95653	529372	517920
34	Uttar Pradesh	1707421	3289844	272839	5583644	4901331
35	West Bengal	9590	1689358	3117216	2523604	2500844

In [ ]:

```
df.head()
```

Out[ ]:

	states	total(year2021)	total(year 2020)	total(year 2019)	total(year 2018)	total(yaer 2017)
0	Andaman & Nicobar Islands	13288	13143	13746	24509	22504
1	Andhra Pradesh	860881	81274372	777706	1546606	1357108
2	Arunachal Pradesh	55767	59184	57645	109744	100244
3	Assam	825053	872000	864281	1741164	1572660
4	Bihar	1282514	1355994	1349389	2517608	2475984

In [ ]:

```
import scipy.stats as stats
import numpy as np
a=df["total(year2021)"].tolist()
b=df["total( year 2020)"].tolist()
l1=[]
for i in range(len(a)):
  l1.append((int(a[i])+int(b[i]))/2)
a1=np.array(l1)
c=df["total(year 2019)"].tolist()
d=df["total(year 2018)"].tolist()
e=df["total(yaer 2017)"].tolist()
l2=[]
for j in range(len(c)):
  l2.append((c[i]+d[i]+e[i])/3)
a2=np.array(l2)
sd1=a1.std()          # calculating the standard deviation
sd2=a2.std()
print(sd1,sd2)
m1=a1.mean()          # calculating the mean
m2=a2.mean()
print(m1,m2)


# sample size is more than 30, therefore ztest should be done,
z = (x1-x2)-(m1-m2) / ((sd1*2 /len(a1)) + (sd2*2 /len(a2)))**0.5
print(z)



#calculating the z-value
z_alpha=stats.norm.ppf(.05)
print(z_alpha)
```

```
#level of significance
alpha=0.05


if z<-z_alpha:
  print("Reject the null hypothesis.")
else:
  print("Accept the null hypothesis.")
252190537.43246344 0.0
44371826.93055555 2713888.0
41646809.58472077
-1.6448536269514729
Accept the null hypothesis.
```

Analysis 4.2:

```
#CODES FOR PROBLEM STATEMENT-2
from google.colab import files
uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving Professional Qualification 1.csv to Professional Qualification 1.csv
import pandas as pd
df = pd.read_csv('Professional Qualification 1.csv')
df
```

Out[ ]:

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
0	2017	116732	2097218	712687	3809929	1700180	615193	41605	243923
1	2018	133741	2491290	380996	4705722	1591341	698853	148151	281323

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
2	2019	161764	2992296	426929	5369117	1368845	879532	131699	277893
3	2020	169219	3040632	403733	5530483	1274956	916206	123374	278195
4	2021	163745	2913972	399738	5701968	1069555	918939	107068	302025

In [ ]:

```
df.head()
```

Out[ ]:

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
0	2017	116732	2097218	712687	3809929	1700180	615193	41605	243923
1	2018	133741	2491290	380996	4705722	1591341	698853	148151	281323

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
2	2019	161764	2992296	426929	5369117	1368845	879532	131699	277893
3	2020	169219	3040632	403733	5530483	1274956	916206	123374	278195
4	2021	163745	2913972	399738	5701968	1069555	918939	107068	302025

In [ ]:

```
corr_matrix = df.corr()
```

In [ ]:

```
corr_matrix
```

Out[ ]:

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
Professional Qualification	1.000000	0.901598	0.853623	-0.683526	0.938716	-0.992802	0.932788	0.407334	0.857051

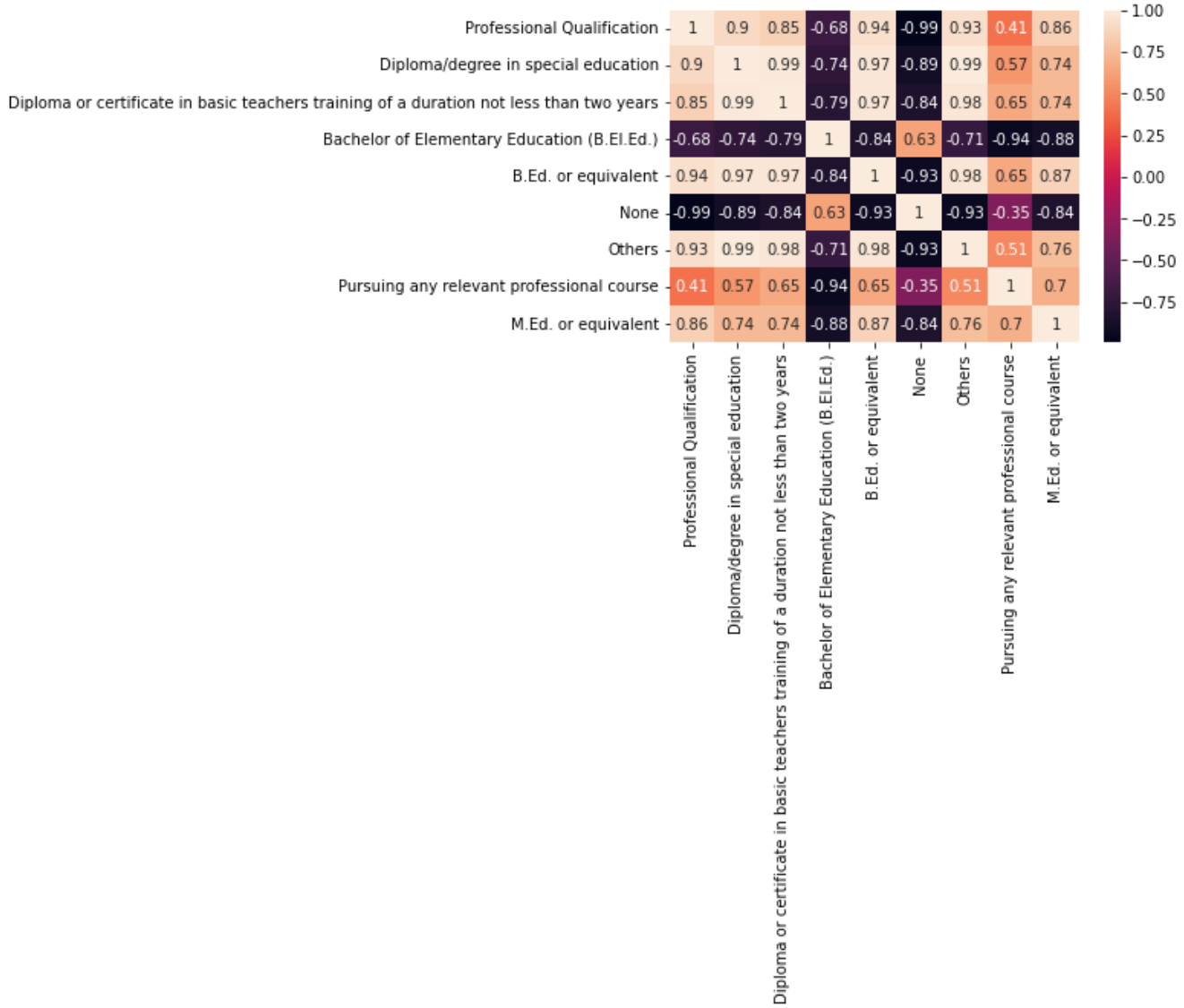
	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
Diploma/degree in special education	0.901598	1.000000	0.991189	-0.741419	0.973799	-0.893944	0.993893	0.568745	0.741902
Diploma or certificate in basic teachers training of a duration not less than two years	0.853623	0.991189	1.000000	-0.787608	0.969128	-0.844882	0.975505	0.653110	0.743085
Bachelor of Elementary Education (B.El.Ed.)	-0.683526	-0.741419	-0.787608	1.000000	-0.836449	0.632495	-0.707179	-0.939276	-0.884405
B.Ed. or equivalent	0.938716	0.973799	0.969128	-0.836449	1.000000	-0.927277	0.975497	0.650041	0.874702
None	-0.992802	-0.893944	-0.844882	0.632495	-0.927277	1.000000	-0.933461	-0.352866	-0.836938
Others	0.932788	0.993893	0.975505	-0.707179	0.975497	-0.933461	1.000000	0.507784	0.760392
Pursuing any relevant	0.407334	0.568745	0.653110	-0.939276	0.650041	-0.352866	0.507784	1.000000	0.695087

	Professional Qualification	Diploma/degree in special education	Diploma or certificate in basic teachers training of a duration not less than two years	Bachelor of Elementary Education (B.El.Ed.)	B.Ed. or equivalent	None	Others	Pursuing any relevant professional course	M.Ed. or equivalent
professional course									
M.Ed. or equivalent	0.857051	0.741902	0.743085	-0.884405	0.874702	-0.836938	0.760392	0.695087	1.000000

In [ ]:

```
# Select the columns to calculate the correlation for
column1 = df["Diploma/degree in special education"]
column2 = df["Diploma or certificate in basic teachers training of a
duration not less than two years"]
column3 = df["Bachelor of Elementary Education (B.El.Ed.)"]
column4 = df["B.Ed. or equivalent"]
column5=df["None"]
column6=df["Others"]
column7=df["Pursuing any relevant professional course"]
column8=df["M.Ed. or equivalent"]

# Create a list of columns
columns = [column1, column2, column3, column4,column5]

# Create a dictionary to store the correlations
correlations = {}

# Iterate through the columns
for i in range(len(columns)):
    for j in range(i+1, len(columns)):
        column1 = columns[i]
        column2 = columns[j]

        # Calculate the means of the columns
        mean_column1 = column1.mean()
        mean_column2 = column2.mean()
```

```
        # Calculate the standard deviations of the columns
        std_column1 = column1.std()
        std_column2 = column2.std()

        # Calculate the covariance between the columns
        covariance = ((column1 - mean_column1) * (column2 -
mean_column2)).sum() / (len(column1) - 1)
# Calculate the correlation
        correlation = covariance / (std_column1 * std_column2)

        # Add the correlation to the dictionary
        key = "column" + str(i+1) + " and column" + str(j+1)
        correlations[key] = correlation

# Print the correlations
print(correlations)

corr_matrix = df.corr()
import seaborn as sns
sns.heatmap(corr_matrix,annot=True)
plt.show()
{'column1 and column2': 0.9911892875503429, 'column1 and column3': -
0.7414189382339003, 'column1 and column4': 0.9737986600710081, 'column1 and
column5': -0.8939437830223227, 'column2 and column3': -0.7876081275993746,
'column2 and column4': 0.969127511124176, 'column2 and column5': -
0.844882134367895, 'column3 and column4': -0.8364487474758922, 'column3 and
column5': 0.6324946191628111, 'column4 and column5': -0.9272765535523655}
```



Analysis 4.3:

```
#CODES FOR PROBLEM STATEMENT-3
from google.colab import files
uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.

```
Saving book21.csv to book21.csv
```

In [ ]:

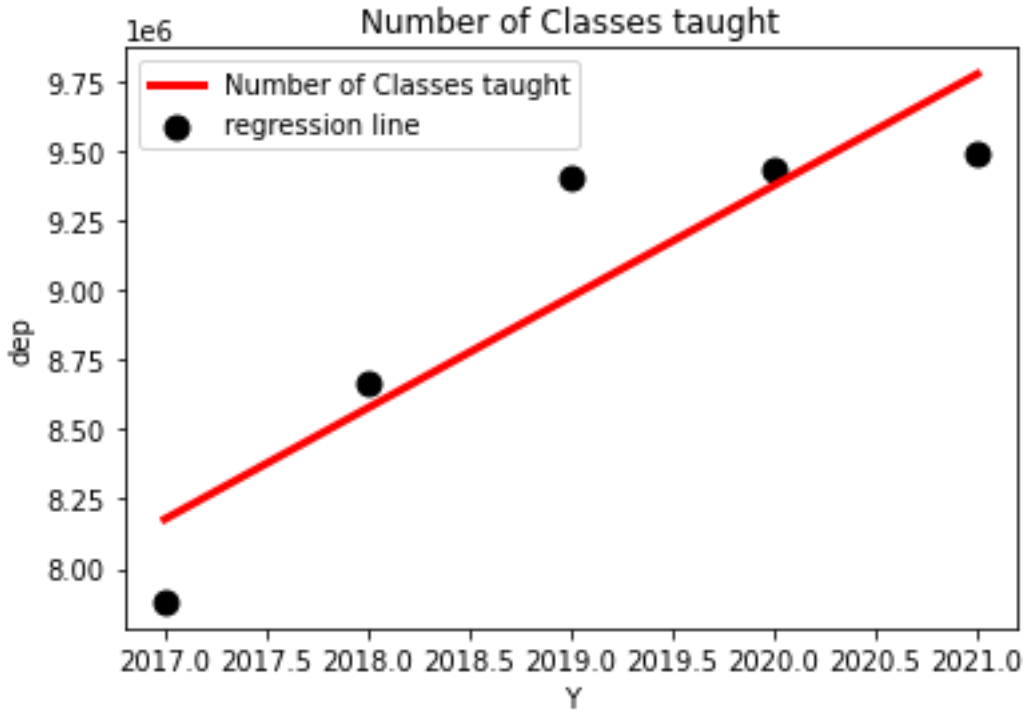
```
import pandas as pd
df = pd.read_csv('book21.csv')
df
```

Out[ ]:

	Year	Department of Education
0	2017	7881609
1	2018	8665149
2	2019	9409248
3	2020	9435055
4	2021	9495945

In [ ]:

```
import numpy as np
from sklearn.linear_model import LinearRegression
y=np.array(df["Year"].tolist()).reshape(-1,1)
dep=df['Department of Education'].tolist()
reg = LinearRegression().fit(y, dep)

# Predict number of Schools for the next five years
next_years = np.array([2031, 2030, 2029, 2028, 2027, 2026, 2025, 2024,
2023, 2022]).reshape(-1,1)
predictions = reg.predict(next_years)

# Print predictions
print(predictions)

#importing matplotlib. python for plotting graphs
import matplotlib.pyplot as plt

# Plot data and regression line
plt.scatter(y, dep,s=80,color="black")
plt.plot(y, reg.predict(y),linewidth=3,color="red")

# Add labels and title
plt.xlabel('Y')
plt.ylabel('dep')
plt.title('Number of Classes taught')
plt.legend(["Number of Classes taught","regression line"])
# Show plot
plt.show()
[13775694.80000007 13375837.00000012 12975979.20000005 12576121.4000001
 12176263.60000014 11776405.80000007 11376548.00000012 10976690.20000005
 10576832.4000001  10176974.60000014]
```


Number of Classes taught
1e6
9.75
9.50
9.25
9.00
8.75
8.50
8.25
8.00
dep
Number of Classes taught
regression line
2017.0
2017.5
2018.0
2018.5
2019.0
2019.5
2020.0
2020.5
2021.0
Y