Analysis of School and School Teacher data of Gujarat(2019-20)

A COURSE PROJECT REPORT

MASTER OF TECHNOLOGY

INDUSTRIAL ENGINEERING AND MANAGEMENT

SUBMITTED BY:

Gohel Hardik Mohanbhai(M220494ME)

UNDER THE GUIDANCE OF

DR. ANUP APREM (ASSISTANT PROFESSOR)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

CALICUT - 673601

Content

Dataset And Database

Description of Data

The data for this subject is taken from the data.gov.in. This data is the primary data for this analysis. The data was divided as per the district of Gujarat. Therefore , there were 33 dataset for 33 districts of Gujarat. Data collected is covering wide range of information from school management to total number of teachers. There are total 36 attributes in the data and 31231 records.

#	Column	Non-Null Count	Dtype
0	Academic_Year	31231 non-null	object
1	State_Code	31231 non-null	int64
2	State_Name	31231 non-null	object
3	District_Code	31231 non-null	int64
4	District_Name	31231 non-null	object
5	Block_Code	31231 non-null	int64
6	Block_Name	31231 non-null	object
7	School_Management_Id	31231 non-null	int64
8	School_Management_Name	31231 non-null	object
9	School_Category_Id	31231 non-null	int64
10	School_Category_Name	31231 non-null	object
11	Academic_Qualification_Id	31231 non-null	int64
12	Academic_Qualification_Name	31231 non-null	object
13	Professional_Qualification_Id	31231 non-null	int64
14	Professional_Qualification_Name	31231 non-null	object
15	Only_Pre_Primary_Male	31231 non-null	int64

In this 36 attributes, 8 have object as data type and 28 have integer as data type.

Dtype in the table show the datatype of the attribute.

Converting Data to 3NF:

First check the Primary key:

In sqlite the attributes are given column number.

In this dataset the primary key is form by combing the following attributes:

Column no	Attributes(Primary Keys)
6	District_Name
8	Block_Name
10	School_Management_Name
12	School_Category_Name
14	Academic_Qualification_Name
16	Professional_Qualification_Name

Lets start from 1NF

The data in the data set is already in 1NF because all the attributes are containing atomic value and there are no duplicate rows in the dataset.

	Academic_Year	State_Code	State_Name	District_Code	District_Name	Block_Code	Block_Name
0	2019-20	24	Gujarat	2407	AHMEDABAD	240706	DASCROI
1	2019-20	24	Gujarat	2407	AHMEDABAD	240703	VIRAMGAM
2	2019-20	24	Gujarat	2407	AHMEDABAD	240704	SANAND
3	2019-20	24	Gujarat	2407	AHMEDABAD	240707	DHOLKA
4	2019-20	24	Gujarat	2407	AHMEDABAD	240710	BAVLA

For 2NF

To convert 1NF to 2NF , we have to remove partial dependency of attributes . In the given dataset the functional dependency of primary key are as follow.

Academic_Qualification_Name ---> Academic_Qualification_Id

Block_Name ---> Block_Code

District_Name ---> District_Code

Professional_Qualification_Name ---> Professional_Qualification_Id

School_Category_Name ---> School_Category_Id

School_Management_Name ---> School_Management_Id

To remove this dependency we have to make new table for each of this functional dependency:

column13	column14
Academic_Qualification_Id	Academic_Qualification_Name
3	Higher secondary
4	Graduate
4	Graduate
4	Graduate
4	Graduate
4	Graduate
4	Graduate

Above table is just an example, we have to do it for all six functional dependency.

3NF table is formed:

Now removing all partial dependency and not having any transitive dependency, 3NF table is created.

column6	column8	column10	column12	column14
District_Name	Block_Name	School_Management_Name	School_Category_Name	Academic_Qualification_Name
AHMEDABAD	DASCROI	Department of Education	SS (IX-X)	Higher secondary
AHMEDABAD	VIRAMGAM	Department of Education	SS (IX-X)	Graduate
AHMEDABAD	SANAND	Department of Education	SS (IX-X)	Graduate
AHMEDABAD	DHOLKA	Department of Education	SS (IX-X)	Graduate
AHMEDABAD	BAVLA	Department of Education	SS (IX-X)	Graduate
AHMEDABAD	DHANDHUKA	Department of Education	SS (IX-X)	Graduate
AHMEDABAD	DHOLERA	Department of Education	SS (IX-X)	Graduate

In this table we can observe that there is no district code with district name, no block code with block name ,etc because only primary keys are there in the table .

Above table is just part of the 3NF table. New table have 27 attributes after removing dependent variables and attributes which are not required.

Data Visualization

School Managements

Question : What is the percentage/ number of schools under each school management in Gujarat?

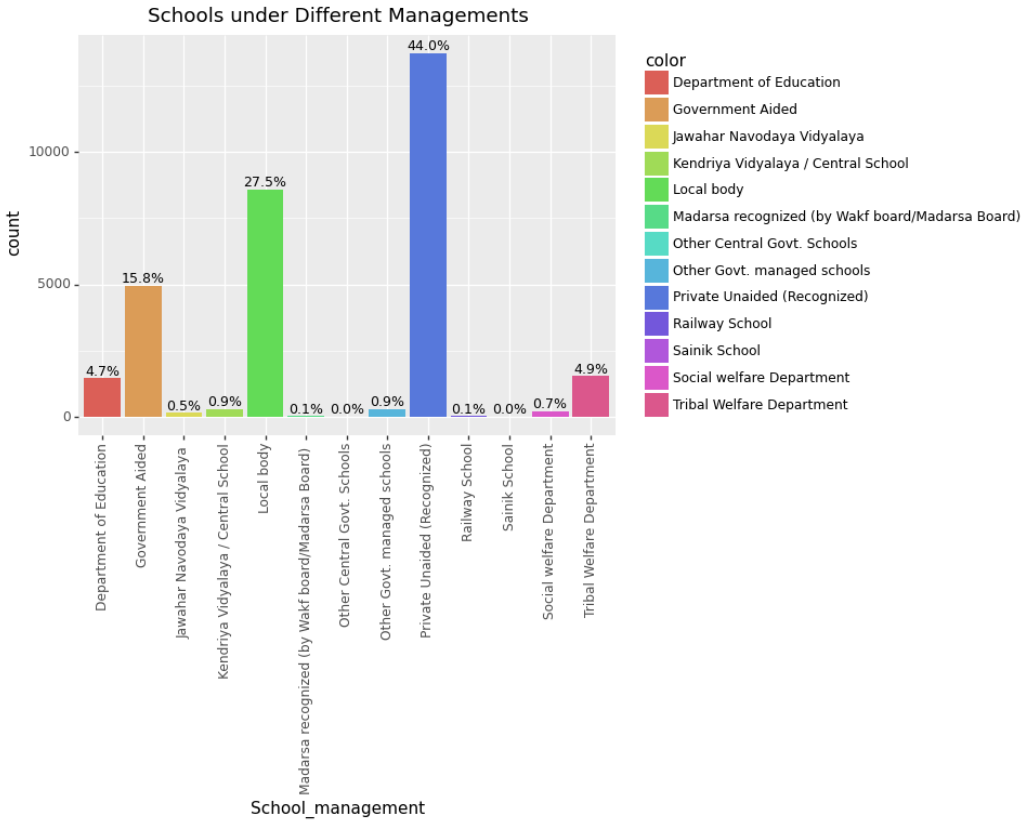
- **Data:**
  Total number of schools in Gujarat is 31230 which can be known by using *shape()* of the table number of
  To get the required data for total number of school under each school management, use groupby function with size(). For total number of schools in Gujarat use sum() function .

	School_Management_Name	No_Of_Schools
0	Department of Education	1458
1	Government Aided	4940
2	Jawahar Navodaya Vidyalaya	146
3	Kendriya Vidyalaya / Central School	281
4	Local body	8589
5	Madarsa recognized (by Wakf board/Madarsa Board)	19
6	Other Central Govt. Schools	12
7	Other Govt. managed schools	276
8	Private Unaided (Recognized)	13739
9	Railway School	18
10	Sainik School	3
11	Social welfare Department	211
12	Tribal Welfare Department	1539

- **Graph:** As data is about counting the total number of school, bar graph is used. X axis represent the categories of school management and Y axis is representing the count.

To plot this graph plotnine library is used. First the data is categorize by using categorical function. geom_bar is use for ploting the bar graph. For labeling the title labs is use.

For writing the percentage on bar geom_text is use. After_ stat is use for converting count into proportion or percentage.



Inference

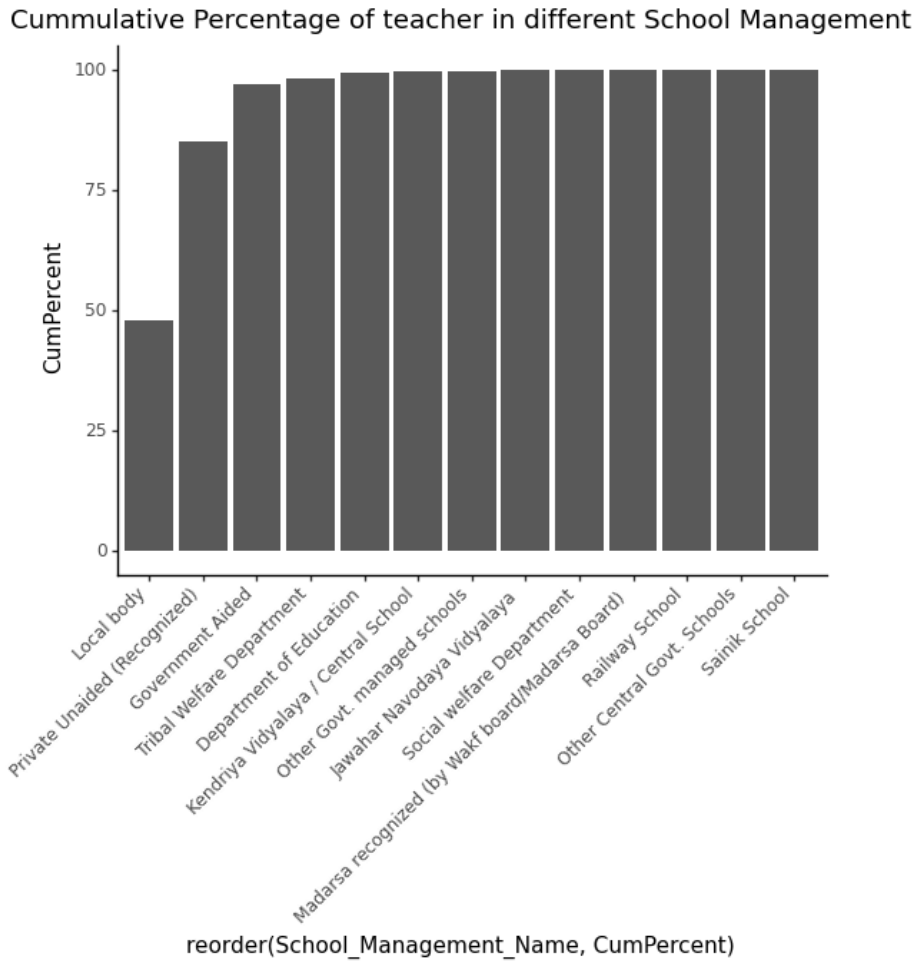
There are total 13 school management working in Gujarat. From the given graph it can be inferred that about 71 % of total schools are local body or private unaided schools. Rest 29% are government schools, which include Government aided, Tribal welfare, central school,etc.

Question: what is the number and percentage of teacher working in different school management? Top 3 school management .

- **Data:**
  step1: Using groupby, the data of school_Mannagement_Name column was taken with total male , total female and total teacher as column.
  Step2: Created a list of percentage by using *for loop* and assigned that list to the 'percent' column.
  Step3: Added another column for cumulative percentage as 'cumpercent' . It is calculated by using numpy function *cumsum().*

	School_Management_Name	Total_Male	Total_Female	Total_Teacher	percent	CumPercent
0	Local body	99976	91555	191531	47.893526	47.893526
1	Private Unaided (Recognized)	44606	104607	149213	37.311645	85.205171
2	Government Aided	32984	14716	47700	11.927684	97.132855
3	Tribal Welfare Department	3033	1570	4603	1.151009	98.283864
4	Department of Education	2372	1605	3977	0.994474	99.278338
5	Kendriya Vidyalaya / Central School	687	538	1225	0.306319	99.584657
6	Other Govt. managed schools	7	576	583	0.145783	99.730439
7	Jawahar Navodaya Vidyalaya	360	122	482	0.120527	99.850966
8	Social welfare Department	294	142	436	0.109025	99.959991
9	Madarsa recognized (by Wakf board/Madarsa Board)	13	34	47	0.011753	99.971744
10	Railway School	21	23	44	0.011002	99.982746
11	Other Central Govt. Schools	27	16	43	0.010752	99.993499
12	Sainik School	20	6	26	0.006501	100.000000

- **Graph:** Below bar chart is plotted using plotnine. Here the X-axis represents the school management name and Y-axis represents the cumulative percentage of teachers.
  To keep the order of the bar in increasing order, here *reorder* function is use. By default the values or characters will follow the ascending order in value or alphabet.
  Stat='identity' is use to provide the value for Y axis , which is cumulative percentage. By default in geom_bar *stat='bin'* is present which take Y axis value as number of rows or count.



Inference: From the above table and graph it can be infer that about 97% of the teachers in Gujarat are working in Local body, Private unaided and Government aided schools. Rest 3% are working in other government school.

It can be seen that 85% of the teacher are working in private schools which include local body and private unaided school.

School Category

Question: What is the Number/ percentage of school by category in Gujarat?

Data:

Step1: use the groupby function with with size() and reset_index for School_category_Name.

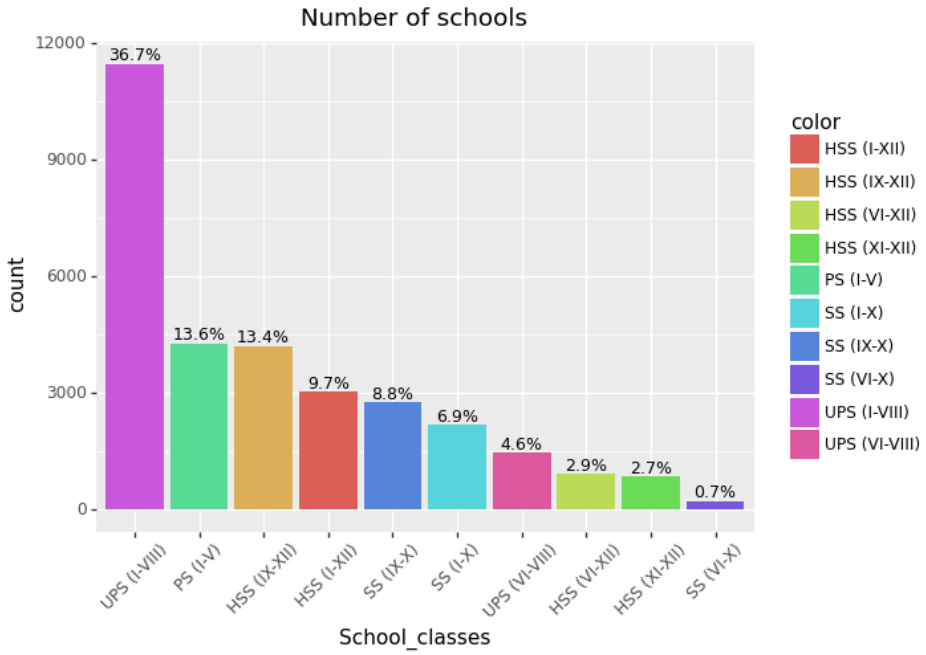
Step2: Dataframe is created and by using *for loop* percentage is calculated .

Step3: list of percentage is assigned to the dataframe.

	School_Category_Name	Count	Percent
0	HSS (I-XII)	3023	9.679485
1	HSS (IX-XII)	4194	13.428965
2	HSS (VI-XII)	903	2.891358
3	HSS (XI-XII)	836	2.676828
4	PS (I-V)	4258	13.633889
5	SS (I-X)	2161	6.919407
6	SS (IX-X)	2748	8.798950
7	SS (VI-X)	204	0.653197
8	UPS (I-VIII)	11458	36.687906
9	UPS (VI-VIII)	1446	4.630015

Graph:

Below bar chart is plotted using plotnine. X-axis represents school category and Y-axis represents count. The bars in the graph are arranged in descending order. This is done by using categorical function of pandas.On each bar percentage are shown to get the idea of school category percentage.



Inference: From the above graph it is clear that about 36% of total schools are upper primary school than next 2 are primary school and higher secondary school.

Upperprimary schools are more because primary education is very important and everybody have the right of primary education.

Question: which 5 district are having the highest and lowest number of schools in Gujarat and how much?

Data:

Step1: Use groupby for District_Name and count() the number of district. Count will show the number of school in each district.

	District_Name	School_count
4	BANAS KANTHA	1683
27	SURAT	1658
0	AHMEDABAD	1641
25	RAJKOT	1614
15	KACHCHH	1474

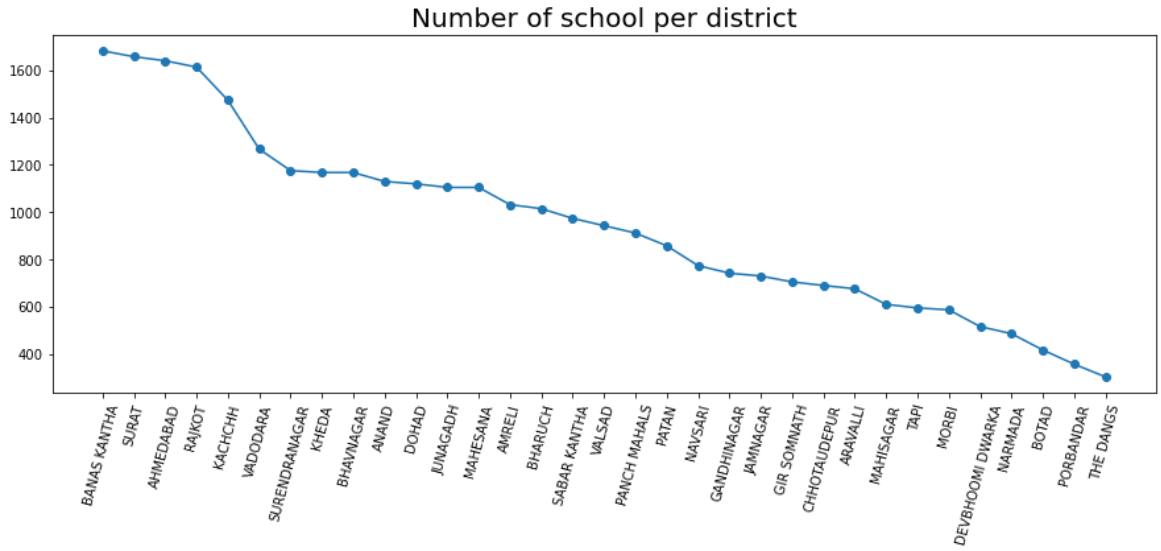
	District_Name	School_count
30	THE DANGS	303
24	PORBANDAR	357
7	BOTAD	417
20	NARMADA	486
9	DEVBHOOMI DWARKA	516

		School_Category_Name
District_Name	School_Category_Name	
BANAS KANTHA	UPS (I-VIII)	647
	PS (I-V)	282
	HSS (IX-XII)	233
	SS (IX-X)	152
	UPS (VI-VIII)	104
	HSS (I-XII)	79
	HSS (VI-XII)	76
	SS (I-X)	65
	HSS (XI-XII)	30
	SS (VI-X)	15

Graph:

To plot the below graph , matplotlib library is use. Here X axis is number of schools and Y axis is district in Gujarat.

First of all the data was arranged in descending order , therefore graph is showing the negative slop. Plt.plot is use for ploting the graph.



Inference:

From the graph and table it can be infer that highest number of school are in banas kantha district which is 1683, while the least school are found in the Dang district of Gujarat, Dang is also the most backward district in Gujarat because it is a tribal area. Dang has only 303 schools.

In Banas Kantha most of the schools are upper primary school which account 647 schools out of 1683 schools.

Academic Qualification of Teachers

Question: What is the number of male and female teachers as per qualification in Gujarat?

Data: The data is as below:

Step1: By using groupby function sum is calculated for male and female for each academic qualification.

Step2: Both table of male and female are merged using concat function

Step3: percentage is calculated and round off to 2 decimal point using *round()*

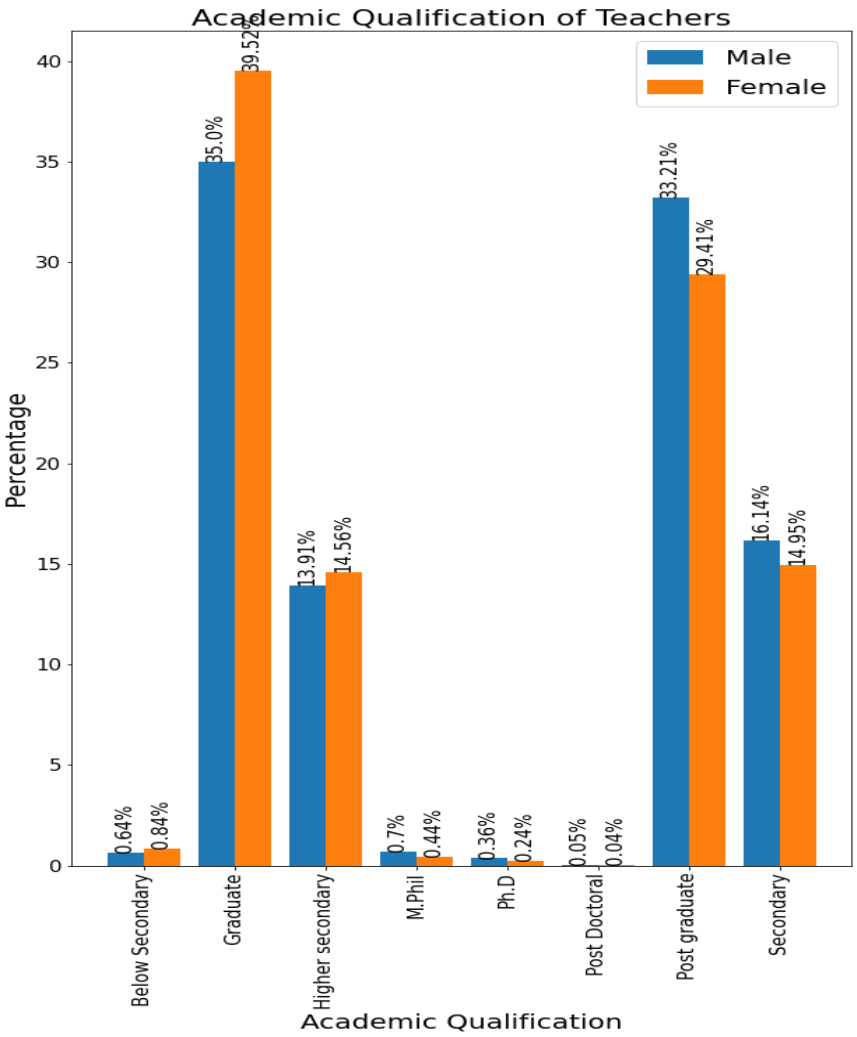
Academic_Qualification_Name	Total_Male	Total_Female	Male_Teacher_percent	Female_Teacher_percent
Below Secondary	1189	1821	0.64	0.84
Graduate	64531	85167	35.00	39.52
Higher secondary	25654	31368	13.91	14.56
M.Phil	1286	944	0.70	0.44
Ph.D	668	524	0.36	0.24
Post Doctoral	86	85	0.05	0.04
Post graduate	61233	63378	33.21	29.41
Secondary	29753	32223	16.14	14.95

Graph:

Matplotlib library is used for ploting this graph. Two bars are made individually and plotted side by side. Here X axis represents Academic qualification and Y axis represents count.

Plt.bar is use for plotting this bar graph.

Count above each bar is executed by using plt.text(). For loop is used to execute this command on each bar of the plot.



Inference: From the graph it can be infer that about 35% of male teacher and 39% of female teacher are graduates which is highest as compared to other academic qualification. In this female percentage is higher than male percentage.

Second highest academic qualification by percentage is postgraduates , in this male percentage is higher than female percentage.

The least percentage is of post doctoral because it is highest degree among all this.

Professional qualification

Professional Qualification of Teachers

Question: What is percentage of professional qualification of teacher in Gujarat?

Data: Groupby the data by professional qualification and sum the total number of male and female . calculated the percentage using formula.

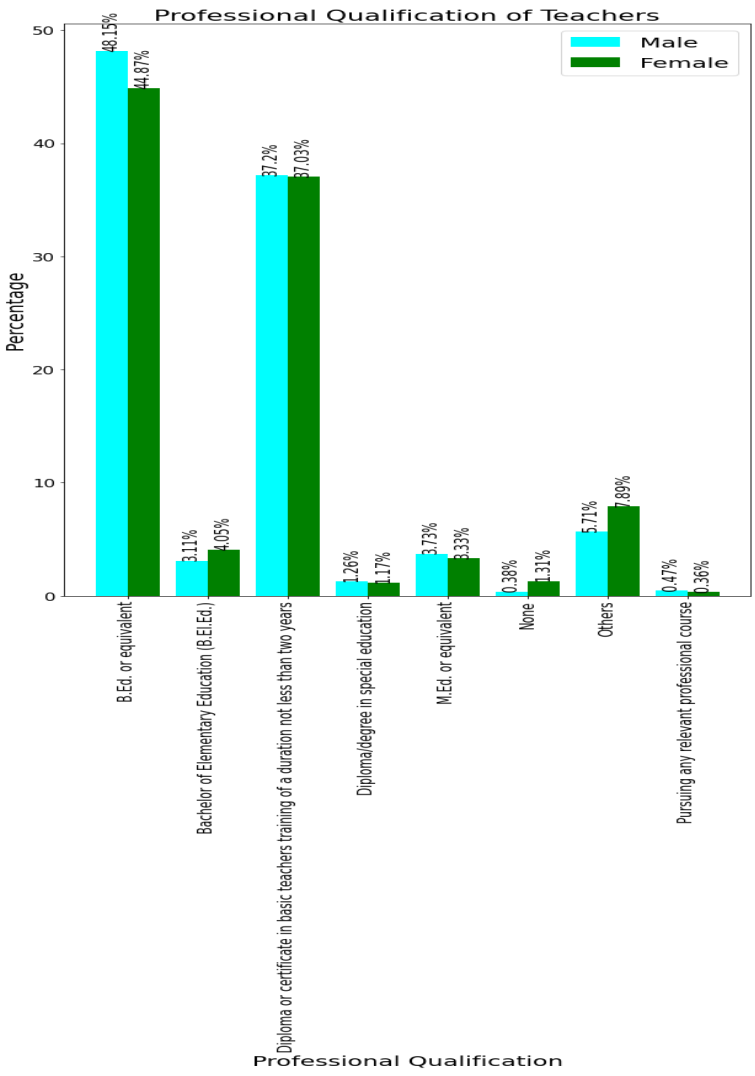
Professional_Qualification_Name	Total_Male	Total_Female	Male_Teacher_percent	Female_Teacher_percent
B.Ed. or equivalent	88782	96698	48.15	44.87
Bachelor of Elementary Education (B.El.Ed.)	5733	8724	3.11	4.05
Diploma or certificate in basic teachers training of a duration not less than two years	68599	79796	37.20	37.03
Diploma/degree in special education	2319	2517	1.26	1.17
M.Ed. or equivalent	6871	7169	3.73	3.33
None	702	2827	0.38	1.31
Others	10524	17003	5.71	7.89
Pursuing any relevant professional course	870	776	0.47	0.36

Graph: X axis in the bar graph represents professional qualification of teachers and Y axis represents the percentage. Both male and female percentage is compared.

The graph is plotted using matplotlib.

Inference: From the graph it can be infer that most of teacher have done B.Ed or equivalent course and in that male percentage is higher than female .

Second highest is Diploma certificate in teacher training, male and female percentage is almost equal. Therefore from this it can be infer that the B.Ed and diploma constitute more than 80% of the teachers.

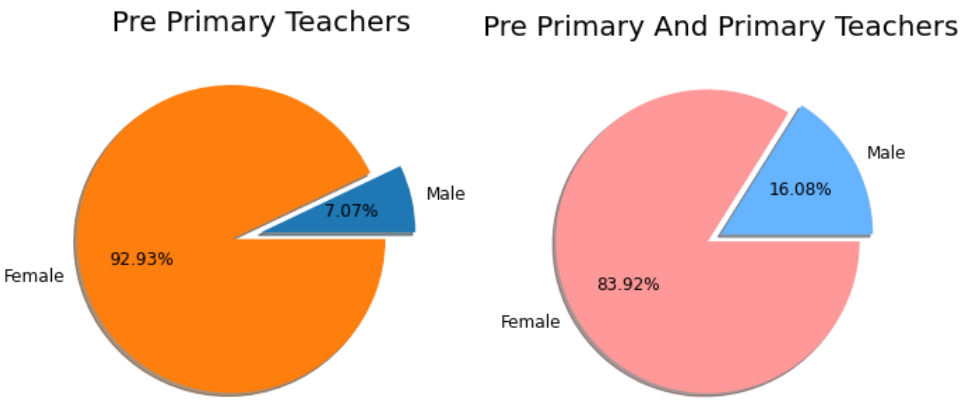

Professional Qualification of Teachers
Male
Female
Percentage
0
10
20
30
40
50
48.15%
44.87%
3.11%
4.05%
37.2%
37.03%
1.26%
1.17%
3.73%
3.33%
0.38%
1.31%
5.71%
7.89%
0.47%
0.36%
B.Ed. or equivalent
Bachelor of Elementary Education (B.El.Ed.)
Diploma or certificate in basic teachers training of a duration not less than two years
Diploma/degree in special education
M.Ed. or equivalent
None
Others
Pursuing any relevant professional course
Professional Qualification

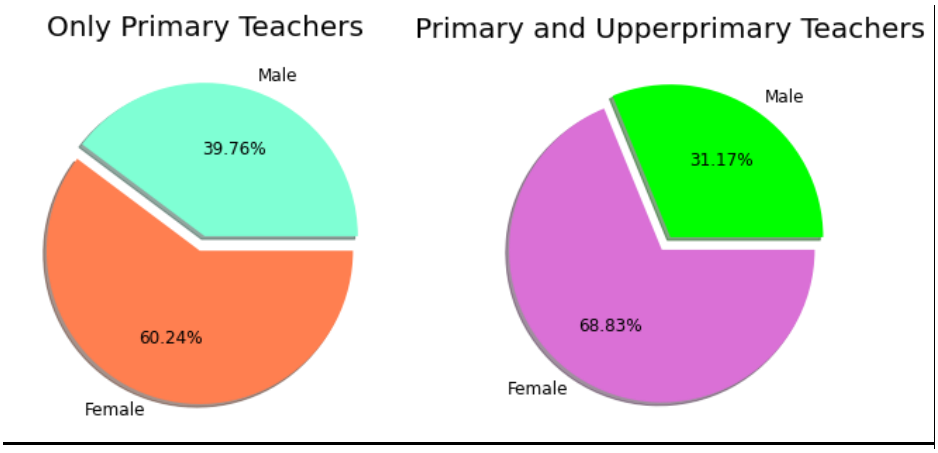
Teachers Data as per Schools

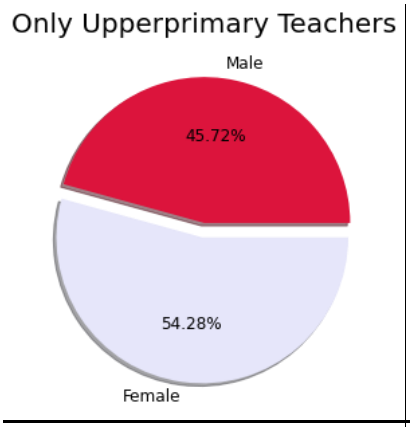
Question: What is the male female ratio for different school category teachers?

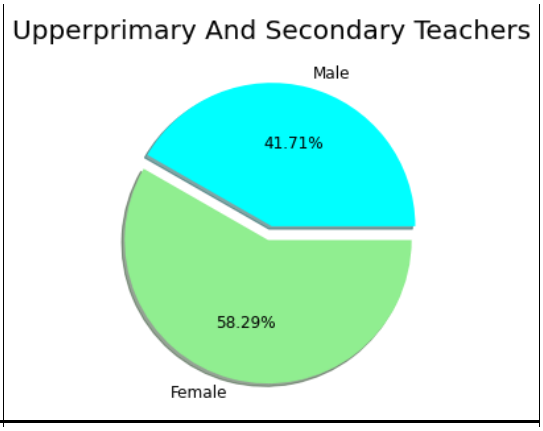
Data: Data for this is directly taken from the attributes of the teacher teaching in different school category.

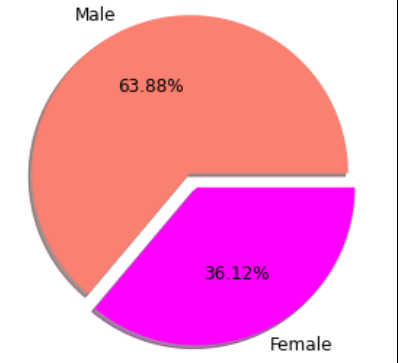
Graph : Graph for this are plot with matplotlib using plt.pie.

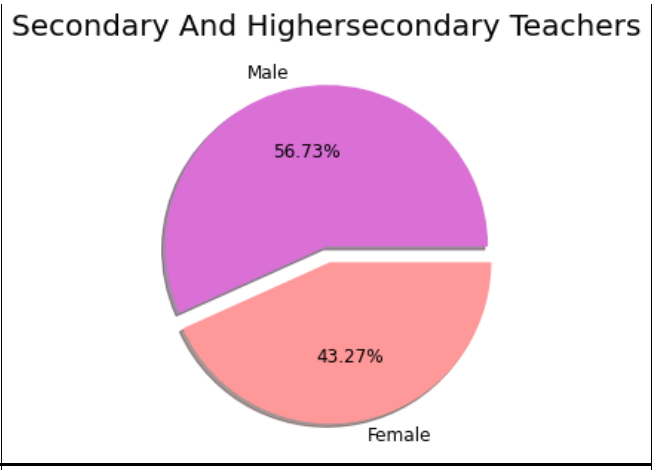


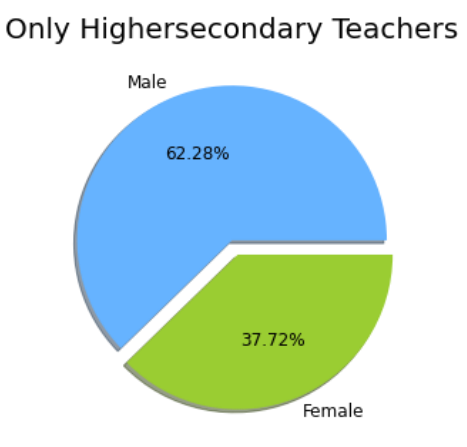


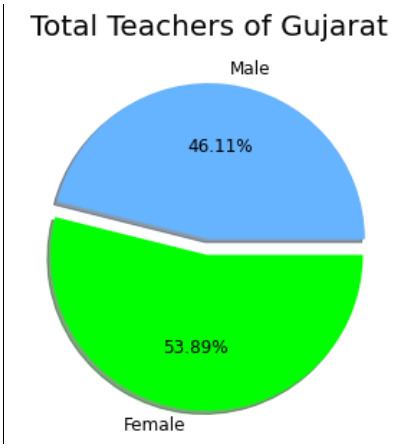












Inference: From the above graphs it can be infer that female teacher ratio is higher in pre primary to upper primary classes while male teacher ratio is higher in secondary to higher secondary classes.

Question: What is the male and female teacher count as per district?

Data:

Step1: Using group by function ,data is group by district name. By using sum() function total male and female data is summed.

Step2: concate the both data of male and female using concate function.

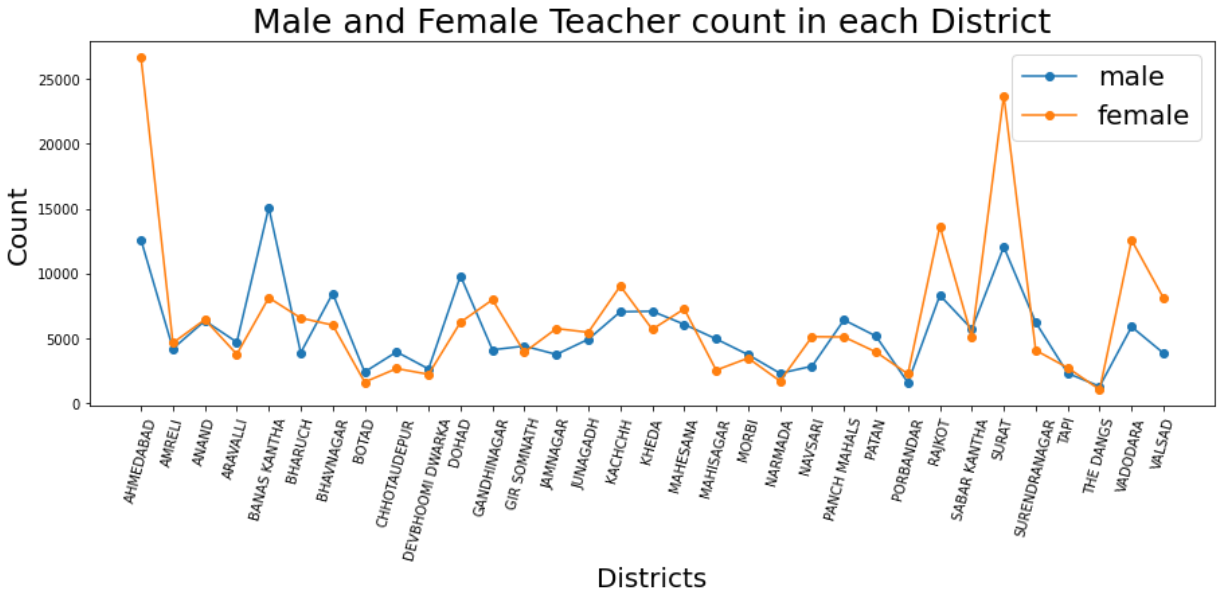
	District_Name	Total_Male	Total_Female
0	AHMEDABAD	12614	26693
27	SURAT	12060	23725
25	RAJKOT	8322	13658
31	VADODARA	5896	12582
15	KACHCHH	7054	9045

	District_Name	Total_Male	Total_Female
30	THE DANGS	1299	1053
7	BOTAD	2410	1632
20	NARMADA	2307	1711
9	DEVBHOOMI DWARKA	2644	2229
24	PORBANDAR	1592	2256

Graph:

The graph is plotted using matplotlib. Plt.plot is use for plotting the graph.

Here X axis represent the district of Gujarat and Y axis represents the count or number of teachers.



Inference: From the graph it can be inferred that women teachers are higher in main city like ahmedabad,surat, vadodara and rajkot while women teachers are less in less develop area like Dang, Botad, Narmada.

Relation Between School Management and School Category

Question: How many number of school of different category are under different management?

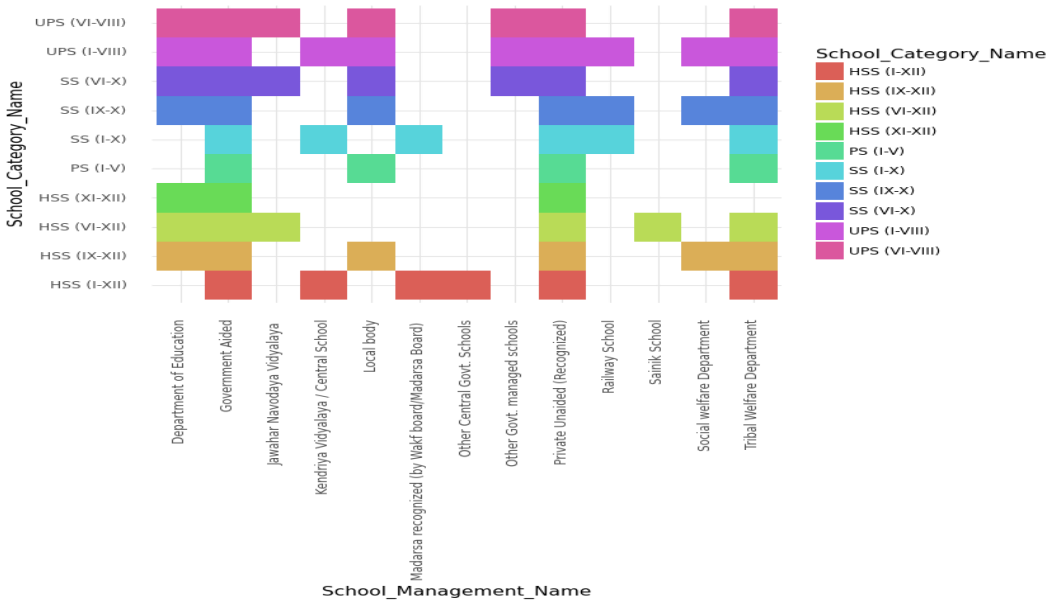
Data: Using crosstab function , relation between school management and school category is establish.

School_Category_Name School_Management_Name	HSS (I-XII)	HSS (IX-XII)	HSS (VI-XII)	HSS (XI-XII)	PS (I-V)	SS (I-X)	SS (IX-X)	SS (VI-X)	UPS (I-VIII)	UPS (VI-VIII)
Department of Education	0	325	177	229	0	0	702	21	1	3
Government Aided	68	2368	222	60	138	7	1120	45	430	482
Jawahar Navodaya Vidyalaya	0	0	126	0	0	0	0	7	0	13
Kendriya Vidyalaya / Central School	240	0	0	0	0	34	0	0	7	0
Local body	0	94	0	0	2358	0	71	6	5923	137
Madarsa recognized (by Wakf board/Madarsa Board)	15	0	0	0	0	4	0	0	0	0
Other Central Govt. Schools	12	0	0	0	0	0	0	0	0	0
Other Govt. managed schools	0	0	0	0	0	0	0	16	4	256
Private Unaided (Recognized)	2661	1213	194	547	1760	1996	711	46	4080	531
Railway School	0	0	0	0	0	12	2	0	4	0
Sainik School	0	0	3	0	0	0	0	0	0	0
Social welfare Department	0	54	0	0	0	0	20	0	137	0
Tribal Welfare Department	27	140	181	0	2	108	122	63	872	24

Graph1:

Using plotnine, tile graph is plotted to show which school categories are present in each school management.

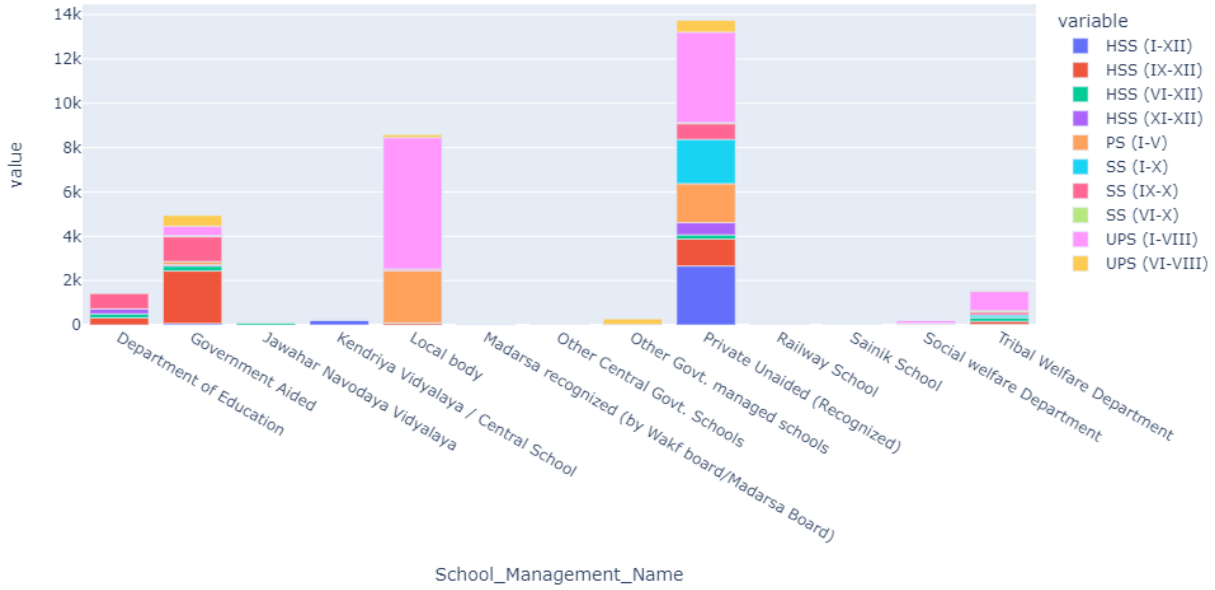
Here X axis represent the school management name and Y axis represents school category.



Graph2:

Using plotly, the below graph is made. Px.bar is used for plotting the graph.

Here X axis represents school management name and y axis represent the count. Stacks in the barplot shows school category.



Inference: From Graph 1 it can be infer that which school category are present in each school management. Government aided and private unaided are having all categories of school while other are not having all categories of school. Least category are found in madarsa board having only one category of school (I-XII).

Graph 2 represents count of school in each school management. By using plotly graphs we can get the count just by howering the cursor on the desired block of which we want information.

Relation Between Academic Qualification And Professional Qualification of Teacher

Question: What is the professional qualification of teacher with respect to their academic Qualification?

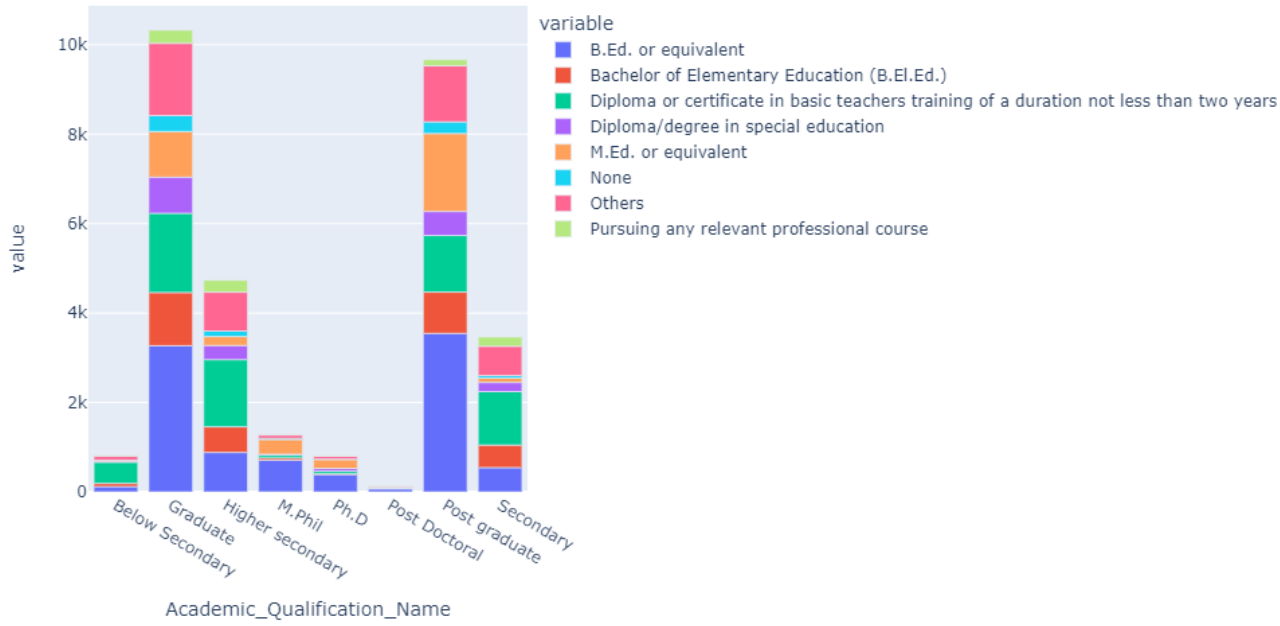
Data: To get the required data, crosstab() function is use for getting cross tabulation between Academic qualification and Professional qualification.

Professional_Qualification_Name Academic_Qualification_Name	B.Ed. or equivalent	Bachelor of Elementary Education (B.El.Ed.)	Diploma or certificate in basic teachers training of a duration not less than two years	Diploma/degree in special education	M.Ed. or equivalent	None	Others	Pursuing any relevant professional course
Below Secondary	120	72	475	19	16	9	92	10
Graduate	3272	1188	1772	809	1018	357	1617	293
Higher secondary	884	573	1502	317	199	123	870	264
M.Phil	711	49	64	20	330	23	77	9
Ph.D	390	21	54	61	201	8	64	7
Post Doctoral	69	6	12	11	14	1	15	12
Post graduate	3546	923	1266	537	1743	261	1252	138
Secondary	545	505	1193	206	96	57	656	207

Graph:

The below graph is plotted using plotly library.

Here X-axis represents academic qualification and Y- axis represents count of the professional qualification of teacher with respect to academic qualification.



Inference: From the graph it can be infer that most of the Graduates and post graduates have done B.Ed or equivalent course. Teacher who have higher secondary and secondary as their academic qualification have mostly done diploma or certificate in basic teacher training.

References:

https://data.gov.in/catalog/teachers-gender-academic-qualification-professional-qualification-classes-taught-school?page=1

https://data.gov.in/catalog/teachers-gender-academic-qualification-professional-qualification-classes-taught-school?page=2#

https://data.gov.in/catalog/teachers-gender-academic-qualification-professional-qualification-classes-taught-school?page=3