

```

from google.colab import files
uploaded=files.upload()
import pandas as pd
df=pd.read_csv('students_class10_basic_dataset.csv')
print(df.head())
print(df.info())
print(df.describe())
df=df.dropna()
df=df.drop_duplicates()
df.columns=[col.strip().lower().replace(" ","_") for col in df.columns]
#calculating mean marks for each column
mean_scores=df.mean(numeric_only=True)
print(mean_scores.sort_values(ascending=False))
corr=df.corr(numeric_only=True)#correlation matrix
print(corr)
#visualize
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(14,10))
sns.heatmap(corr,annot=True,cmap='coolwarm')
plt.title("subject correlation Heatmap")
plt.show()
#to see higer attendence
plt.scatter(df['Attendance_%'],df['Q_Maths'])
plt.xlabel('attendance %')
plt.ylabel("Maths Marks")
plt.title('Attendance vs Maths Marks')
plt.show()
#top 5 and bottom 5 students
df['Total_Marks']=df[['Q_Maths','Q_Science','Q_English','Q_History','Q_Geography','Q_Civics']].sum(axis=1)
print("Top 5 students:\n", df.nlargest(5, 'Total_Marks')[['RollNo','Total_Marks']])
print("Bottom 5 students:\n", df.nsmallest(5, 'Total_Marks')[['RollNo','Total_Marks']])
#students need improvement
weak_students=df[df['Total_Marks']<df['Total_Marks'].mean()]
print("students below average:\n", weak_students[['RollNo', 'Total_Marks']])
#data analysis report
print("DATA ANALYSIS REPORT")
print("-"*50)

print(f"\nTotal Students: {len(df)}")
print(f"\nAverage Attendance: {df['Attendance_%'].mean():.2f}%")

print("\nSubjects ranked by average marks:")
print(mean_scores.sort_values(ascending=False))

print("\nHighest performing subject:", mean_scores.idxmax())
print("Lowest performing subject:", mean_scores.idxmin())

print("\nCorrelation between Maths and Science:", round(df['Q_Maths'].corr(df['Q_Science']),2))
print("Correlation between English and History:", round(df['Q_English'].corr(df['Q_History']),2))

print("\nAverage total marks:", df['Total_Marks'].mean().round(2))
print("\nTop 5 Students:")
print(df.nlargest(5, 'Total_Marks')[['RollNo','Total_Marks']])

print("\nBottom 5 Students:")
print(df.nsmallest(5, 'Total_Marks')[['RollNo','Total_Marks']])
# Average marks per subject
mean_scores.plot(kind='bar', color='skyblue', figsize=(10,5))
plt.title("Average Marks per Subject")
plt.ylabel("Average Marks")
plt.show()

# Distribution of total marks
plt.hist(df['Total_Marks'], bins=10, color='lightgreen', edgecolor='black')
plt.title("Distribution of Total Marks")
plt.xlabel("Total Marks")
plt.ylabel("Number of Students")
plt.show()

```



Choose Files | students\_cl...dataset.csv  
**students\_class10\_basic\_dataset.csv**(text/csv) - 5085 bytes, last modified: 10/31/2025 - 100% done  
Saving students\_class10\_basic\_dataset.csv to students\_class10\_basic\_dataset (13).csv

	RollNo	Name	Gender	DOB	Attendance_%	House	Q_Maths	\
0	1	Aarav Sharma	Male	12/27/2003	84.2	Blue	69	
1	2	Aditi Nair	Female	2/5/2003	84.7	Red	62	
2	3	Bhavesh Patel	Male	1/24/2003	98.7	Green	55	
3	4	Bhavna Kapoor	Female	12/5/2003	96.5	Yellow	100	
4	5	Charan Reddy	Male	4/27/2003	79.4	Green	62	

	Q_Science	Q_English	Q_History	...	F_History	F_Geography	F_Civics	\
0	70	74	87	...	83	85	56	
1	78	64	69	...	75	86	90	
2	58	75	92	...	93	88	80	
3	74	69	71	...	91	79	86	
4	68	76	65	...	77	56	41	

	Science_Project_20	English_Activity_20	ArtCraft_20	Sports_Points_30	\
0	13	16	12	11	
1	14	12	16	21	
2	10	17	12	23	
3	11	10	11	21	
4	16	13	13	8	

	Cultural_Points_20	Discipline_10	Volunteer_Hours	
0	15	10	11.2	
1	3	6	7.1	
2	3	6	11.7	
3	7	8	7.7	
4	17	7	1.6	

[5 rows x 31 columns]  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 40 entries, 0 to 39  
Data columns (total 31 columns):

#	Column	Non-Null Count	Dtype
0	RollNo	40 non-null	int64
1	Name	40 non-null	object
2	Gender	40 non-null	object
3	DOB	40 non-null	object
4	Attendance_%	40 non-null	float64
5	House	40 non-null	object
6	Q_Maths	40 non-null	int64
7	Q_Science	40 non-null	int64
8	Q_English	40 non-null	int64
9	Q_History	40 non-null	int64
10	Q_Geography	40 non-null	int64
11	Q_Civics	40 non-null	int64
12	H_Maths	40 non-null	int64
13	H_Science	40 non-null	int64
14	H_English	40 non-null	int64
15	H_History	40 non-null	int64
16	H_Geography	40 non-null	int64
17	H_Civics	40 non-null	int64
18	F_Maths	40 non-null	int64
19	F_Science	40 non-null	int64
20	F_English	40 non-null	int64
21	F_History	40 non-null	int64
22	F_Geography	40 non-null	int64
23	F_Civics	40 non-null	int64
24	Science_Project_20	40 non-null	int64
25	English_Activity_20	40 non-null	int64
26	ArtCraft_20	40 non-null	int64
27	Sports_Points_30	40 non-null	int64
28	Cultural_Points_20	40 non-null	int64
29	Discipline_10	40 non-null	int64
30	Volunteer_Hours	40 non-null	float64

dtypes: float64(2), int64(25), object(4)  
memory usage: 9.8+ KB  
None

	RollNo	Attendance_%	Q_Maths	Q_Science	Q_English	Q_History	\
count	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	
mean	20.500000	86.390000	68.825000	71.850000	70.600000	71.375000	
std	11.690452	8.511223	12.691381	12.978384	12.179849	11.226542	
min	1.000000	72.000000	39.000000	43.000000	28.000000	48.000000	
25%	10.750000	79.400000	61.500000	63.000000	62.750000	64.750000	
50%	20.500000	86.400000	67.500000	71.500000	70.000000	70.500000	
75%	30.250000	94.175000	75.000000	81.250000	79.250000	77.000000	
max	40.000000	98.700000	100.000000	100.000000	93.000000	92.000000	

	Q_Geography	Q_Civics	H_Maths	H_Science	...	F_History	\
count	40.000000	40.000000	40.000000	40.000000	...	40.000000	
mean	73.250000	70.900000	70.075000	73.525000	...	76.075000	
std	12.877629	13.318197	15.047595	14.952896	...	18.518719	
min	46.000000	42.000000	38.000000	35.000000	...	19.000000	
25%	63.000000	63.750000	58.750000	64.000000	...	67.000000	
50%	72.500000	69.000000	70.000000	71.000000	...	79.500000	
75%	85.250000	80.250000	80.000000	81.750000	...	90.250000	
max	92.000000	98.000000	100.000000	100.000000	...	100.000000	

```

F_Geography F_Civics Science_Project_20 English_Activity_20 \
count    40.000000 40.000000 40.000000 40.000000
mean     81.050000 76.375000 13.925000 13.750000
std      13.668136 15.836808 3.346544 4.27725
min      47.000000 41.000000 8.000000 6.000000
25%     71.750000 65.000000 11.000000 10.000000
50%     83.500000 78.000000 14.000000 14.000000
75%     89.250000 89.250000 17.000000 17.000000
max     100.000000 100.000000 20.000000 20.000000

ArtCraft_20 Sports_Points_30 Cultural_Points_20 Discipline_10 \
count    40.000000 40.000000 40.000000 40.000000
mean     12.325000 15.725000 10.175000 7.575000
std      4.17801 9.251438 5.687988 1.430214
min      5.000000 0.000000 1.000000 6.000000
25%     9.000000 8.750000 5.500000 6.750000
50%     12.500000 17.500000 11.000000 7.000000
75%     16.000000 21.250000 14.000000 8.000000
max     20.000000 30.000000 20.000000 10.000000

Volunteer_Hours
count    40.000000
mean     7.732500
std      4.209293
min      0.700000
25%     4.700000
50%     7.750000
75%     10.900000
max     14.900000

[8 rows x 27 columns]
Attendance_%          86.3900
F_Geography           81.0500
F_Science              77.5500
H_Geography            76.6500
F_English              76.6000
F_Civics               76.3750
F_History              76.0750
F_Maths                75.1250
H_Civics               74.1250
H_Science              73.5250
Q_Geography            73.2500
H_English              73.1750
H_History              72.3250
Q_Science              71.8500
Q_History              71.3750
Q_Civics               70.9000
Q_English              70.6000
H_Maths                70.0750
Q_Maths                68.8250
RollNo                 20.5000
Sports_Points_30        15.7250
Science_Project_20       13.9250
English_Activity_20      13.7500
ArtCraft_20              12.3250
Cultural_Points_20       10.1750
Volunteer_Hours          7.7325
Discipline_10             7.5750
dtype: float64

          RollNo Attendance_% Q_Maths Q_Science Q_English \
RollNo      1.000000 -0.129442 -0.055216 -0.195701 -0.043219
Attendance_% -0.129442  1.000000  0.253049 -0.234229 -0.148867
Q_Maths     -0.055216  0.253049  1.000000  0.009799 -0.077597
Q_Science   -0.195701 -0.234229  0.009799  1.000000  0.116563
Q_English   -0.043219 -0.148867 -0.077597  0.116563  1.000000
Q_History    -0.230048 -0.053227 -0.183989 -0.057854  0.185270
Q_Geography  -0.022312  0.113041 -0.321974  0.143984 -0.234263
Q_Civics     -0.073615  0.043060  0.180718 -0.087018  0.077359
H_Maths      0.044530  0.079027  0.881245 -0.033290 -0.074260
H_Science    -0.219511 -0.315344 -0.053144  0.875093  0.182659
H_English    0.051880  0.011842 -0.085217  0.033660  0.860585
H_History    -0.165416  0.151476 -0.152142 -0.034290  0.298565
H_Geography  0.009182  0.269676 -0.242101  0.076354 -0.252732
H_Civics     -0.028958  0.138692  0.211315 -0.091992  0.268542
F_Maths      0.131278 -0.133356  0.540742  0.019751  0.084183
F_Science    -0.212620 -0.088351  0.028536  0.639012 -0.102488
F_English    0.114737  0.127706 -0.083644  0.009069  0.558136
F_History    -0.055607  0.105942 -0.089076  0.016478  0.298204
F_Geography  -0.054720  0.285195 -0.174517  0.184917 -0.366296
F_Civics     -0.050759  0.229558  0.142834 -0.032280  0.315578
Science_Project_20 0.020645 -0.141991  0.128274 -0.072290  0.098638
English_Activity_20 0.250241 -0.162419 -0.122220 -0.109702  0.360772
ArtCraft_20      0.077958 -0.034517  0.060095 -0.023194 -0.013504
Sports_Points_30 -0.197369 -0.185617  0.109426  0.150416  0.132345
Cultural_Points_20 -0.168991 -0.137724  0.169864  0.134091  0.063956
Discipline_10      -0.063643  0.051460  0.211928  0.025486  0.224030
Volunteer_Hours   -0.061929 -0.275580 -0.284563  0.247960  0.260679

          Q_History Q_Geography Q_Civics H_Maths H_Science \
RollNo      0.000000  0.000000  0.000000  0.000000  0.000000
Attendance_% -0.000000  0.000000  0.000000  0.000000  0.000000
Q_Maths     0.000000  0.000000  0.000000  0.000000  0.000000
Q_Science   0.000000  0.000000  0.000000  0.000000  0.000000
Q_English   0.000000  0.000000  0.000000  0.000000  0.000000
Q_History   0.000000  0.000000  0.000000  0.000000  0.000000
Q_Geography 0.000000  0.000000  0.000000  0.000000  0.000000
Q_Civics    0.000000  0.000000  0.000000  0.000000  0.000000
H_Maths     0.000000  0.000000  0.000000  0.000000  0.000000
H_Science   0.000000  0.000000  0.000000  0.000000  0.000000
H_English   0.000000  0.000000  0.000000  0.000000  0.000000
H_History   0.000000  0.000000  0.000000  0.000000  0.000000
H_Geography 0.000000  0.000000  0.000000  0.000000  0.000000
H_Civics    0.000000  0.000000  0.000000  0.000000  0.000000
F_Maths     0.000000  0.000000  0.000000  0.000000  0.000000
F_Science   0.000000  0.000000  0.000000  0.000000  0.000000
F_English   0.000000  0.000000  0.000000  0.000000  0.000000
F_History   0.000000  0.000000  0.000000  0.000000  0.000000
F_Geography 0.000000  0.000000  0.000000  0.000000  0.000000
F_Civics    0.000000  0.000000  0.000000  0.000000  0.000000
Science_Project_20 0.000000  0.000000  0.000000  0.000000  0.000000
English_Activity_20 0.000000  0.000000  0.000000  0.000000  0.000000
ArtCraft_20      0.000000  0.000000  0.000000  0.000000  0.000000
Sports_Points_30 0.000000  0.000000  0.000000  0.000000  0.000000
Cultural_Points_20 0.000000  0.000000  0.000000  0.000000  0.000000
Discipline_10      0.000000  0.000000  0.000000  0.000000  0.000000
Volunteer_Hours   0.000000  0.000000  0.000000  0.000000  0.000000

```

RollNo	-0.2730048	-0.0722212	-0.079027	-0.315344
Attendance_%	-0.053227	0.113041	0.043060	0.079027
Q_Maths	-0.183989	-0.321974	0.180718	0.881245
Q_Science	-0.057854	0.143984	-0.087018	-0.033290
Q_English	0.185270	-0.234263	0.077359	-0.074260
Q_History	1.000000	0.030373	0.050847	-0.136016
Q_Geography	0.030373	1.000000	-0.353279	-0.420884
Q_Civics	0.050847	-0.353279	1.000000	0.267187
H_Maths	-0.136016	-0.420884	0.267187	1.000000
H_Science	0.054243	0.207830	-0.078013	-0.044167
H_English	0.167707	-0.016303	-0.013013	-0.088596
H_History	0.837250	0.025028	0.138199	-0.076574
H_Geography	0.028565	0.838353	-0.301492	-0.370106
H_Civics	-0.022169	-0.300678	0.868688	0.242682
F_Maths	-0.278079	-0.426637	0.331737	0.716678
F_Science	-0.198603	0.136003	-0.115241	-0.014616
F_English	0.174239	0.075439	-0.045000	-0.086604
F_History	0.584335	0.025832	0.050973	-0.073264
F_Geography	0.048668	0.575496	-0.104066	-0.303588
F_Civics	-0.101043	-0.215592	0.657991	0.163534
Science_Project_20	0.129757	0.030790	0.005005	0.272017
English_Activity_20	0.260449	0.123595	-0.091374	-0.035954
ArtCraft_20	0.095734	-0.147857	0.112576	0.210460
Sports_Points_30	0.320724	-0.029539	0.125466	0.105323
Cultural_Points_20	-0.153238	-0.108431	0.026300	0.099602
Discipline_10	0.024553	0.079703	0.192901	0.107556
Volunteer_Hours	0.188506	0.253533	-0.143513	-0.233093

	...	F_History	F_Geography	F_Civics	\
RollNo	...	-0.055607	-0.054720	-0.050759	
Attendance_%	...	0.105942	0.285195	0.229558	
Q_Maths	...	-0.089076	-0.174517	0.142834	
Q_Science	...	0.016478	0.184917	-0.032280	
Q_English	...	0.298204	-0.366296	0.315578	
Q_History	...	0.584335	0.048668	-0.101043	
Q_Geography	...	0.025832	0.575496	-0.215592	
Q_Civics	...	0.050973	-0.104066	0.657991	
H_Maths	...	-0.073264	-0.303588	0.163534	
H_Science	...	-0.076539	0.162086	-0.035827	
H_English	...	0.228715	-0.247988	0.247127	
H_History	...	0.796138	-0.008425	-0.006985	
H_Geography	...	0.001937	0.726541	-0.191791	
H_Civics	...	0.024261	-0.144676	0.772099	
F_Maths	...	-0.116545	-0.383058	0.302323	
F_Science	...	-0.347986	0.182093	-0.222934	
F_English	...	0.224500	-0.165146	0.129536	
F_History	...	1.000000	-0.045702	0.033999	
F_Geography	...	-0.045702	1.000000	-0.028518	
F_Civics	...	0.033999	-0.028518	1.000000	
Science_Project_20	...	0.217307	-0.106985	-0.171207	
English_Activity_20	...	0.396143	-0.009430	-0.025456	
ArtCraft_20	...	0.090481	-0.298884	0.067478	
Sports_Points_30	...	0.252156	-0.103304	0.095226	
Cultural_Points_20	...	-0.033964	-0.308490	-0.077887	
Discipline_10	...	-0.000702	0.035218	0.226835	
Volunteer_Hours	...	0.002928	0.262116	-0.045383	

	Science_Project_20	English_Activity_20	ArtCraft_20	\
RollNo	0.020645	0.250241	0.077958	
Attendance_%	-0.141991	-0.162419	-0.034517	
Q_Maths	0.128274	-0.122220	0.060095	
Q_Science	-0.072290	-0.109702	-0.023194	
Q_English	0.098638	0.360772	-0.013504	
Q_History	0.129757	0.260449	0.095734	
Q_Geography	0.030790	0.123595	-0.147857	
Q_Civics	0.005005	-0.091374	0.112576	
H_Maths	0.272017	-0.035954	0.210460	
H_Science	-0.041210	0.012127	0.069435	
H_English	0.171754	0.333604	-0.021988	
H_History	0.204048	0.379617	0.105238	
H_Geography	-0.144367	0.105195	-0.219006	
H_Civics	0.084874	-0.104159	0.015568	
F_Maths	0.273128	-0.073781	0.236105	
F_Science	-0.083955	-0.139287	-0.040569	
F_English	0.080456	0.378543	-0.138411	
F_History	0.217307	0.396143	0.090481	
F_Geography	-0.106985	-0.009430	-0.298884	
F_Civics	-0.171207	-0.025456	0.067478	
Science_Project_20	1.000000	0.150919	-0.271459	
English_Activity_20	0.150919	1.000000	-0.144559	
ArtCraft_20	-0.271459	-0.144559	1.000000	
Sports_Points_30	0.033273	0.326096	-0.001609	
Cultural_Points_20	0.037077	-0.045583	0.036388	
Discipline_10	-0.081831	0.254634	-0.049240	
Volunteer_Hours	-0.091927	0.343972	0.061640	

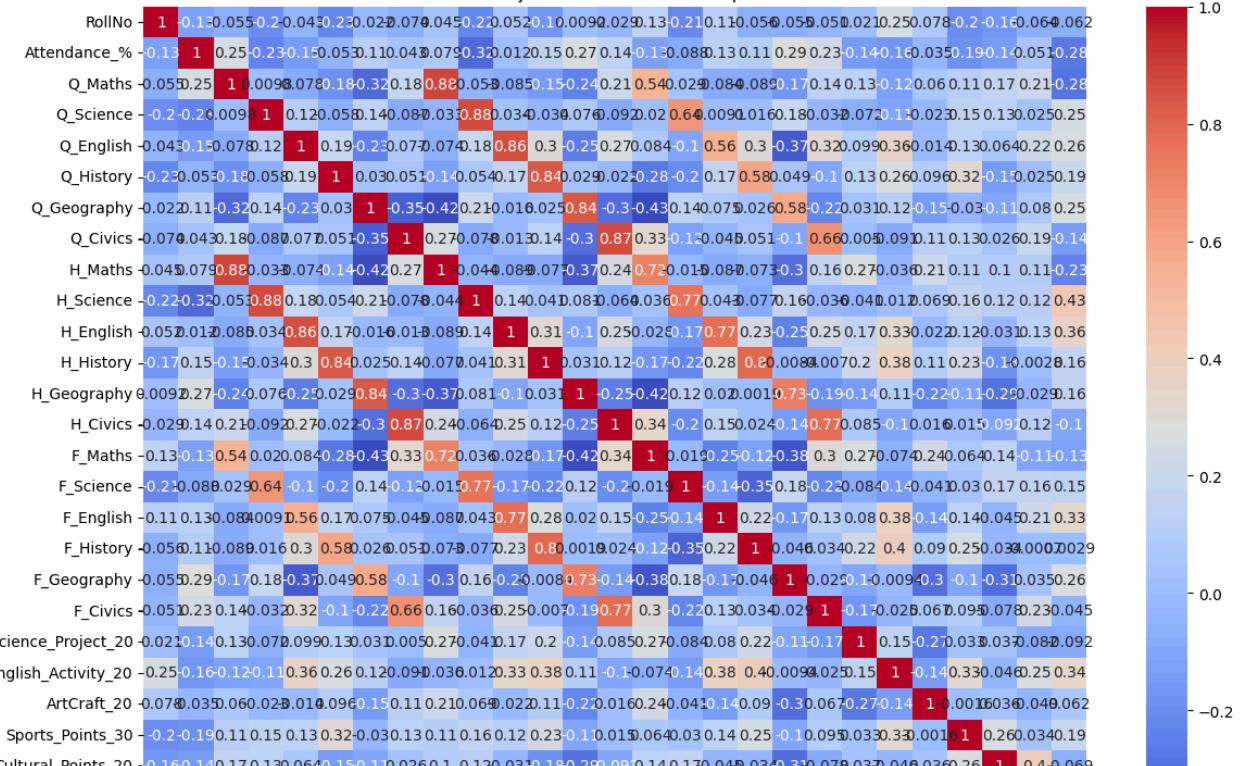
	Sports_Points_30	Cultural_Points_20	Discipline_10	\
RollNo	-0.197369	-0.160991	-0.063643	
Attendance_%	-0.185617	-0.137724	0.051460	
Q_Maths	0.109426	0.169864	0.211928	
Q_Science	0.150416	0.134091	0.025486	

Q_English	0.132345	0.063956	0.224030
Q_History	0.320724	-0.153238	0.024553
Q_Geography	-0.029539	-0.108431	0.079703
Q_Civics	0.125466	0.026300	0.192901
H_Maths	0.105323	0.099602	0.107556
H_Science	0.157879	0.116467	0.122205
H_English	0.121871	-0.030793	0.126281
H_History	0.226832	-0.178306	-0.002820
H_Geography	-0.110004	-0.294888	-0.028708
H_Civics	0.014632	-0.092499	0.115365
F_Maths	0.064465	0.138633	-0.113779
F_Science	0.030129	0.166548	0.162977
F_English	0.142072	-0.045140	0.212855
F_History	0.252156	-0.033964	-0.000702
F_Geography	-0.103304	-0.308490	0.035218
F_Civics	0.095226	-0.077887	0.226835
Science_Project_20	0.033273	0.037077	-0.081831
English_Activity_20	0.326096	-0.045583	0.254634
ArtCraft_20	-0.001609	0.036388	-0.049240
Sports_Points_30	1.000000	0.256753	0.033574
Cultural_Points_20	0.256753	1.000000	0.400215
Discipline_10	0.033574	0.400215	1.000000
Volunteer_Hours	0.191447	-0.069427	0.099036
Volunteer_Hours			
RollNo	-0.061929		
Attendance_%	-0.275580		
Q_Maths	-0.284563		
Q_Science	0.247960		
Q_English	0.260679		
Q_History	0.188506		
Q_Geography	0.253533		
Q_Civics	-0.143513		
H_Maths	-0.233093		
H_Science	0.426494		
H_English	0.358149		
H_History	0.157968		
H_Geography	0.158939		
H_Civics	-0.104558		
F_Maths	-0.125806		
F_Science	0.147905		
F_English	0.332841		
F_History	0.002928		
F_Geography	0.262116		
F_Civics	-0.045383		
Science_Project_20	-0.091927		
English_Activity_20	0.343972		
ArtCraft_20	0.061640		
Sports_Points_30	0.191447		
Cultural_Points_20	-0.069427		
Discipline_10	0.099036		
Volunteer_Hours	1.000000		

[27 rows x 27 columns]

```
/tmp/ipython-input-273181768.py:10: UserWarning: Pandas doesn't allow columns to be created via a new attribute name - see https://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-best-practices
  df.column=[col.strip().lower().replace(" ", "_") for col in df.columns]
```

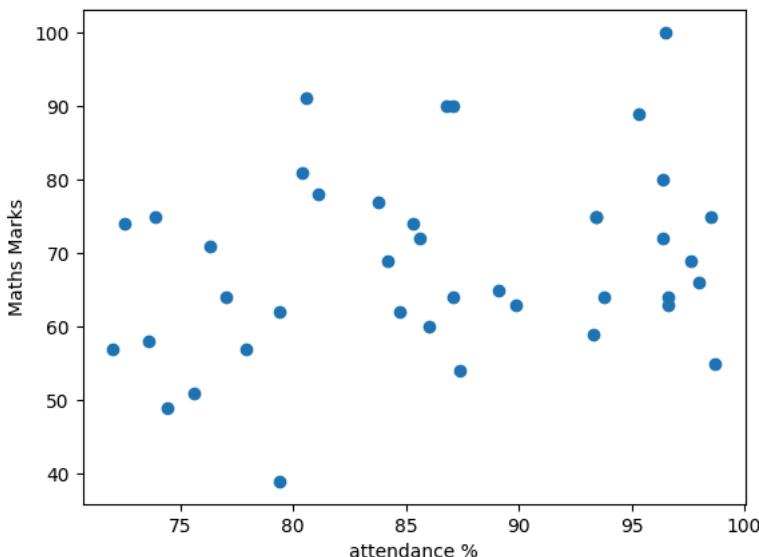
subject correlation Heatmap





-0.4  
1

Attendance vs Maths Marks



**Top 5 students:**

RollNo	Total_Marks
8	476
22	472
6	469
9	460
37	456

**Bottom 5 students:**

RollNo	Total_Marks
26	369
34	374
13	375
25	381
35	381

**Students below average:**

RollNo	Total_Marks
1	417
2	409
4	385
12	417
13	375
15	425
19	399
21	425
24	407
25	381
26	369
29	422
30	415
31	424
34	374
35	381
38	416
39	406

DATA ANALYSIS REPORT

Total Students: 40

Average Attendance: 86.39%

Subjects ranked by average marks:

Attendance_%	86.3900
F_Geography	81.0500
F_Science	77.5500
H_Geography	76.6500
F_English	76.6000
F_Civics	76.3750
F_History	76.0750
F_Maths	75.1250
H_Civics	74.1250
H_Science	73.5250
Q_Geography	73.2500
H_English	73.1750
H_History	72.3250
Q_Science	71.8500

```
Q_History      71.3750
Q_Civics       70.9000
Q_English      70.6000
H_Maths        70.0750
Q_Maths        68.8250
RollNo         20.5000
Sports_Points_30 15.7250
Science_Project_20 13.9250
English_Activity_20 13.7500
ArtCraft_20    12.3250
Cultural_Points_20 10.1750
Volunteer_Hours 7.7325
Discipline_10   7.5750
dtype: float64
```

Highest performing subject: Attendance\_%  
Lowest performing subject: Discipline\_10

Correlation between Maths and Science: 0.01  
Correlation between English and History: 0.19

Average total marks: 426.8

Top 5 Students:

	RollNo	Total_Marks
8	9	476