Data Analytics And Reporting

Student Performance Analysis

Bhavika Sharma : 202410101150050
Paridhi Singh: 202410101150036
Submitted to: Ms. Deepika Tiwari

Content:

Using Pandas:

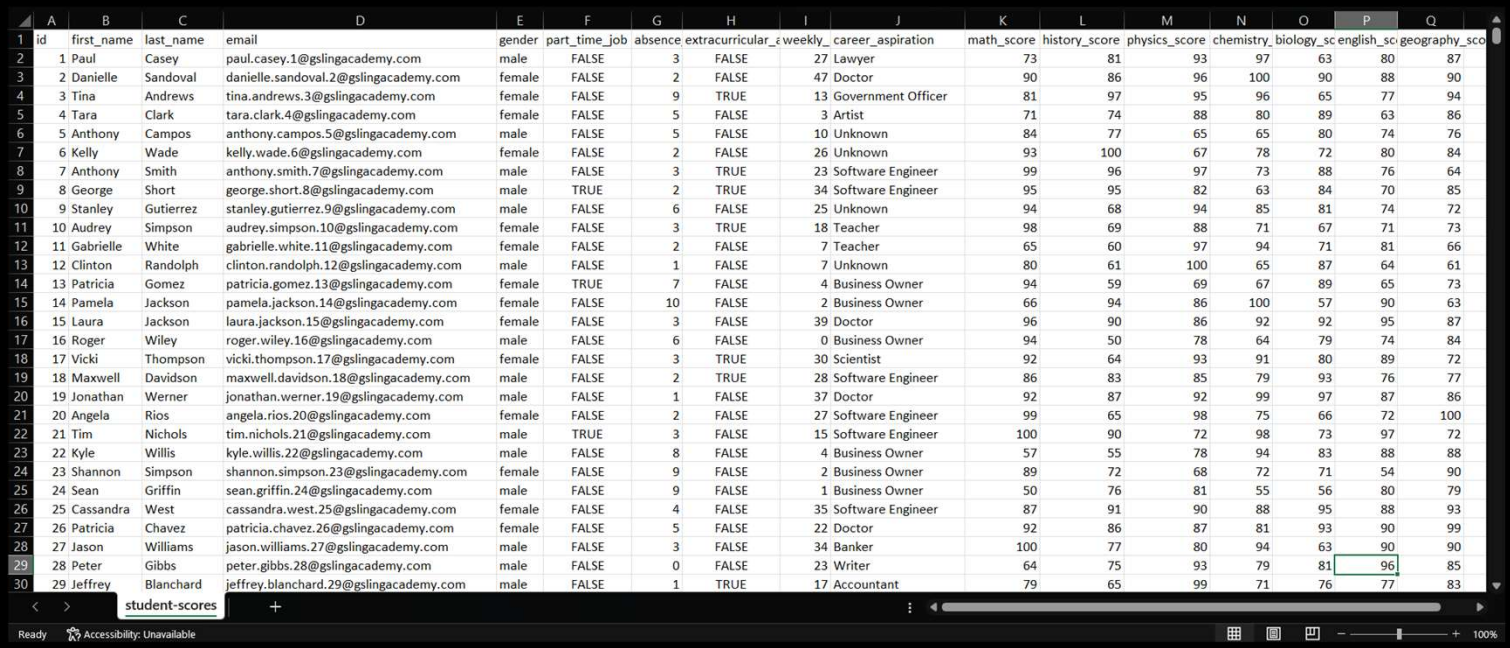
- Data Cleaning
- Summary of Dataset

Using Matplotlib:

- Data Visualization
- Customizing plots

Using Excel:

- Conditional Formatting
- Pivot Tables
- Filtering and sorting



id	first_name	last_name	email	gender	part_time_job	absence	extracurricular_a	weekly_	career_aspiration	math_score	history_score	physics_score	chemistry_	biology_sc	english_sc	geography_sco
1	Paul	Casey	paul.casey.1@gslingacademy.com	male	FALSE	3	FALSE	27	Lawyer	73	81	93	97	63	80	87
2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	FALSE	2	FALSE	47	Doctor	90	86	96	100	90	88	90
3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	FALSE	9	TRUE	13	Government Officer	81	97	95	96	65	77	94
4	Tara	Clark	tara.clark.4@gslingacademy.com	female	FALSE	5	FALSE	3	Artist	71	74	88	80	89	63	86
5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	FALSE	5	FALSE	10	Unknown	84	77	65	65	80	74	76
6	Kelly	Wade	kelly.wade.6@gslingacademy.com	female	FALSE	2	FALSE	26	Unknown	93	100	67	78	72	80	84
7	Anthony	Smith	anthony.smith.7@gslingacademy.com	male	FALSE	3	TRUE	23	Software Engineer	99	96	97	73	88	76	64
8	George	Short	george.short.8@gslingacademy.com	male	TRUE	2	TRUE	34	Software Engineer	95	95	82	63	84	70	85
9	Stanley	Gutierrez	stanley.gutierrez.9@gslingacademy.com	male	FALSE	6	FALSE	25	Unknown	94	68	94	85	81	74	72
10	Audrey	Simpson	audrey.simpson.10@gslingacademy.com	female	FALSE	3	TRUE	18	Teacher	98	69	88	71	67	71	73
11	Gabrielle	White	gabrielle.white.11@gslingacademy.com	female	FALSE	2	FALSE	7	Teacher	65	60	97	94	71	81	66
12	Clinton	Randolph	clinton.randolph.12@gslingacademy.com	male	FALSE	1	FALSE	7	Unknown	80	61	100	65	87	64	61
13	Patricia	Gomez	patricia.gomez.13@gslingacademy.com	female	TRUE	7	FALSE	4	Business Owner	94	59	69	67	89	65	73
14	Pamela	Jackson	pamela.jackson.14@gslingacademy.com	female	FALSE	10	FALSE	2	Business Owner	66	94	86	100	57	90	63
15	Laura	Jackson	laura.jackson.15@gslingacademy.com	female	FALSE	3	FALSE	39	Doctor	96	90	86	92	92	95	87
16	Roger	Wiley	roger.wiley.16@gslingacademy.com	male	FALSE	6	FALSE	0	Business Owner	94	50	78	64	79	74	84
17	Vicki	Thompson	vicki.thompson.17@gslingacademy.com	female	FALSE	3	TRUE	30	Scientist	92	64	93	91	80	89	72
18	Maxwell	Davidson	maxwell.davidson.18@gslingacademy.com	male	FALSE	2	TRUE	28	Software Engineer	86	83	85	79	93	76	77
19	Jonathan	Werner	jonathan.werner.19@gslingacademy.com	male	FALSE	1	FALSE	37	Doctor	92	87	92	99	97	87	86
20	Angela	Rios	angela.rios.20@gslingacademy.com	female	FALSE	2	FALSE	27	Software Engineer	99	65	98	75	66	72	100
21	Tim	Nichols	tim.nichols.21@gslingacademy.com	male	TRUE	3	FALSE	15	Software Engineer	100	90	72	98	73	97	72
22	Kyle	Willis	kyle.willis.22@gslingacademy.com	male	FALSE	8	FALSE	4	Business Owner	57	55	78	94	83	88	88
23	Shannon	Simpson	shannon.simpson.23@gslingacademy.com	female	FALSE	9	FALSE	2	Business Owner	89	72	68	72	71	54	90
24	Sean	Griffin	sean.griffin.24@gslingacademy.com	male	FALSE	9	FALSE	1	Business Owner	50	76	81	55	56	80	79
25	Cassandra	West	cassandra.west.25@gslingacademy.com	female	FALSE	4	FALSE	35	Software Engineer	87	91	90	88	95	88	93
26	Patricia	Chavez	patricia.chavez.26@gslingacademy.com	female	FALSE	5	FALSE	22	Doctor	92	86	87	81	93	90	99
27	Jason	Williams	jason.williams.27@gslingacademy.com	male	FALSE	3	FALSE	34	Banker	100	77	80	94	63	90	90
28	Peter	Gibbs	peter.gibbs.28@gslingacademy.com	male	FALSE	0	FALSE	23	Writer	64	75	93	79	81	96	85
29	Jeffrey	Blanchard	jeffrey.blanchard.29@gslingacademy.com	male	FALSE	1	TRUE	17	Accountant	79	65	99	71	76	77	83

student-scores

Ready Accessibility: Unavailable 100%

Data Cleaning

(Using Pandas)

Loading the Dataset for Data Cleaning:

```
import pandas as pd
#loading data set
student = pd.read_csv('/content/drive/MyDrive/Document from .')
student.head()
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_aspiration
0	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	False	3	False	27	Lawyer
1	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	False	2	False	47	Doctor
2	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	False	9	True	13	Government Officer
3	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	False	5	False	3	Artist
4	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	False	5	False	10	Unknown

Stastical Summery of DataSet:

```
#Stastical summary of data
student.describe()
```

	id	absence_days	weekly_self_study_hours	math_score	history_score	physics_score	chemistry_score	biology_score	english_score	geography_score
count	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.00000	2000.000000	2000.000000
mean	1000.500000	3.665500	17.755500	83.452000	80.332000	81.336500	79.995000	79.58150	81.277500	80.888000
std	577.494589	2.629271	12.129604	13.224906	12.736046	12.539453	12.777895	13.72219	12.027087	11.637705
min	1.000000	0.000000	0.000000	40.000000	50.000000	50.000000	50.000000	30.00000	50.000000	60.000000
25%	500.750000	2.000000	5.000000	77.000000	69.750000	71.000000	69.000000	69.00000	72.000000	71.000000
50%	1000.500000	3.000000	18.000000	87.000000	82.000000	83.000000	81.000000	81.00000	83.000000	81.000000
75%	1500.250000	5.000000	28.000000	93.000000	91.000000	92.000000	91.000000	91.00000	91.000000	91.000000
max	2000.000000	10.000000	50.000000	100.000000	100.000000	100.000000	100.000000	100.00000	99.000000	100.000000

Checking for Null Values :

(Using IsNull())

Removing Null Values from DataSet:

```
#removing all null values
student_Null = student.dropna(axis = 0 , how = "any")
student_Null
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_
0	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	False	3	False	27	
1	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	False	2	False	47	
2	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	False	9	True	13	Goven
3	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	False	5	False	3	
4	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	False	5	False	10	
...	...	...	...	...	...	...	...	...	...	
1995	1996	Alan	Reynolds	alan.reynolds.1996@gslingacademy.com	male	False	2	False	30	
1996	1997	Thomas	Gilbert	thomas.gilbert.1997@gslingacademy.com	male	False	2	False	20	Softw
1997	1998	Madison	Cross	madison.cross.1998@gslingacademy.com	female	False	5	False	14	Softw
1998	1999	Brittany	Compton	brittany.compton.1999@gslingacademy.com	female	True	10	True	5	Bus
1999	2000	Natalie	Smith	natalie.smith.2000@gslingacademy.com	female	False	5	False	27	

2000 rows × 17 columns

```
[ ]
#cleaning dataset by checking for null values
student.isnull().sum()
```

	0
id	0
first_name	0
last_name	0
email	0
gender	0
part_time_job	0
absence_days	0
extracurricular_activities	0
weekly_self_study_hours	0
career_aspiration	0
math_score	0
history_score	0
physics_score	0
chemistry_score	0

Filtering the Data:

```
score_columns = ['math_score', 'history_score', 'physics_score', 'chemistry_score', 'biology_score', 'english_score', 'geography_score']
student['total score'] = student[score_columns].sum(axis=1)
student['total score'] = student['total score'].fillna(0)
student
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_aspi
0	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	False	3	False	27	
1	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	False	2	False	47	
2	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	False	9	True	13	Government
3	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	False	5	False	3	
4	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	False	5	False	10	U
...	...	...	...	...	...	...	...	...	...	
1995	1996	Alan	Reynolds	alan.reynolds.1996@gslingacademy.com	male	False	2	False	30	Cons E
1996	1997	Thomas	Gilbert	thomas.gilbert.1997@gslingacademy.com	male	False	2	False	20	Software E
1997	1998	Madison	Cross	madison.cross.1998@gslingacademy.com	female	False	5	False	14	Software E
1998	1999	Brittany	Compton	brittany.compton.1999@gslingacademy.com	female	True	10	True	5	Business
1999	2000	Natalie	Smith	natalie.smith.2000@gslingacademy.com	female	False	5	False	27	Acc

2000 rows × 18 columns

```
student.columns = student.columns.str.strip()
# Assuming 'total score' is the column intended for filtering
student['total score'] = student['total score'].astype(str).str.replace('$','',regex=False).str.replace(',','',regex=False).str.replace('-','0',regex=False).astype(float)

#filtering data on the basis of score
filter = student['total score']>500
student[filter]
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_aspi
0	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	False	3	False	27	
1	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	False	2	False	47	
2	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	False	9	True	13	Government
3	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	False	5	False	3	
4	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	False	5	False	10	Ur
...	...	...	...	...	...	...	...	...	...	
1995	1996	Alan	Reynolds	alan.reynolds.1996@gslingacademy.com	male	False	2	False	30	Cons Ei
1996	1997	Thomas	Gilbert	thomas.gilbert.1997@gslingacademy.com	male	False	2	False	20	Software Ei
1997	1998	Madison	Cross	madison.cross.1998@gslingacademy.com	female	False	5	False	14	Software Ei
1998	1999	Brittany	Compton	brittany.compton.1999@gslingacademy.com	female	True	10	True	5	Business
1999	2000	Natalie	Smith	natalie.smith.2000@gslingacademy.com	female	False	5	False	27	Acc

Sorting (Using sort_values()):

```
#sorting values
student.sort_values('total score', ascending = False)
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_aspi
796	797	Lisa	Mitchell	lisa.mitchell.797@gslingacademy.com	female	False	3	True	35	
779	780	Todd	Howard	todd.howard.780@gslingacademy.com	male	False	6	False	35	U
1573	1574	Paula	Hernandez	paula.hernandez.1574@gslingacademy.com	female	True	2	False	48	
1419	1420	Mary	Arnold	mary.arnold.1420@gslingacademy.com	female	False	1	False	35	D
551	552	Paul	Shaffer	paul.shaffer.552@gslingacademy.com	male	False	10	False	35	Software E
...	...	...	...	...	...	...	...	...	...	
819	820	Steven	Alvarez	steven.alvarez.820@gslingacademy.com	male	True	2	True	2	Business
527	528	Deborah	Rojas	deborah.rojas.528@gslingacademy.com	female	False	3	True	1	Business
906	907	Alicia	Beltran	alicia.beltran.907@gslingacademy.com	female	False	6	False	4	Business
1661	1662	Thomas	Scott	thomas.scott.1662@gslingacademy.com	male	True	6	False	5	Business
1965	1966	Jennifer	Leblanc	jennifer.leblanc.1966@gslingacademy.com	female	False	6	False	2	Business

2000 rows × 18 columns

Removing Duplicate rows:

```
#removing duplicates
student.drop_duplicates()
```

	id	first_name	last_name	email	gender	part_time_job	absence_days	extracurricular_activities	weekly_self_study_hours	career_aspi
0	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	False	3	False	27	
1	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	False	2	False	47	
2	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	False	9	True	13	Government
3	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	False	5	False	3	
4	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	False	5	False	10	Ur
...	...	...	...	...	...	...	...	...	...	
1995	1996	Alan	Reynolds	alan.reynolds.1996@gslingacademy.com	male	False	2	False	30	Cons Er
1996	1997	Thomas	Gilbert	thomas.gilbert.1997@gslingacademy.com	male	False	2	False	20	Software Er
1997	1998	Madison	Cross	madison.cross.1998@gslingacademy.com	female	False	5	False	14	Software Er
1998	1999	Brittany	Compton	brittany.compton.1999@gslingacademy.com	female	True	10	True	5	Business
1999	2000	Natalie	Smith	natalie.smith.2000@gslingacademy.com	female	False	5	False	27	Acc

2000 rows × 18 columns

Data Visualization

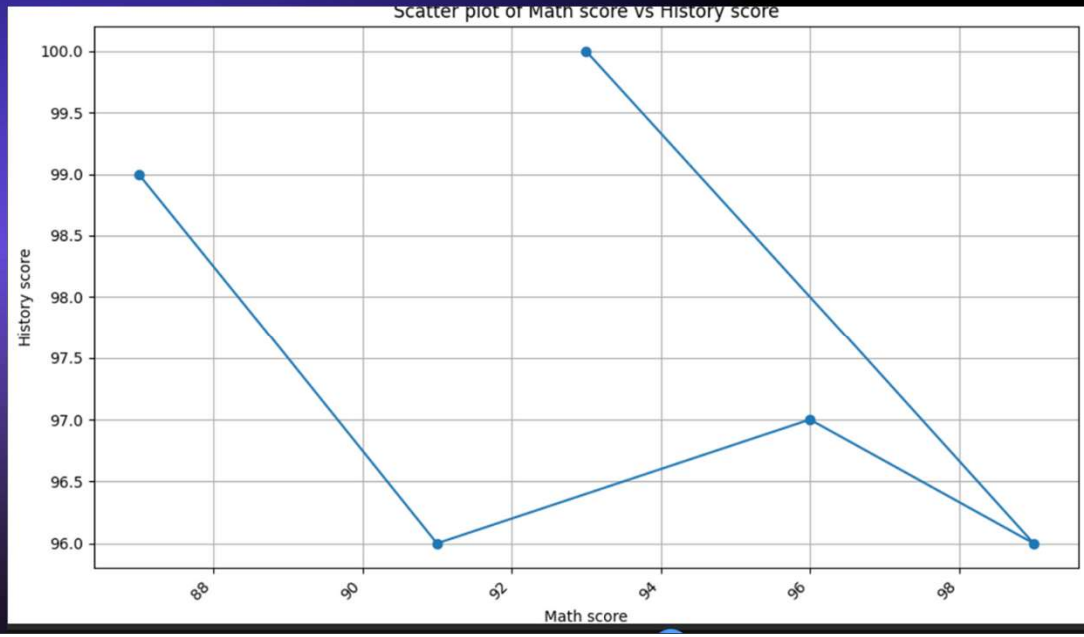
(Using Matplotlib)

Importing required libraries:

```
[ ]  import matplotlib.pyplot as plt

[ ]  # student = pd.read_csv('/content/drive/MyDrive/Document from .')
```

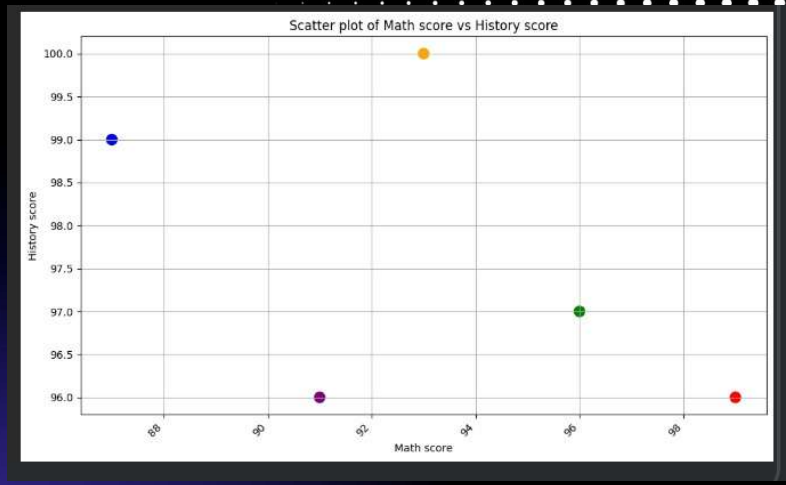
Line Plot:

```
[ ]  import matplotlib.pyplot as plt
     import importlib

     #reload the matplotlib.pyplot module to restore its original functions
     importlib.reload(plt)

     # Define top 5 students based on 'total score' directly before plotting
     top_5_students = student.sort_values(by='total score', ascending=False).head(5)

     plt.figure(figsize=(10, 6))
     plt.plot(top_5_students['math_score'], top_5_students['history_score'], marker='o', linestyle='-')
     plt.xlabel('Math score')
     plt.ylabel('History score')
     plt.title('Scatter plot of Math score vs History score')
     plt.grid(True)
     plt.xticks(rotation=45, ha='right')
     plt.tight_layout()
     plt.show()
```



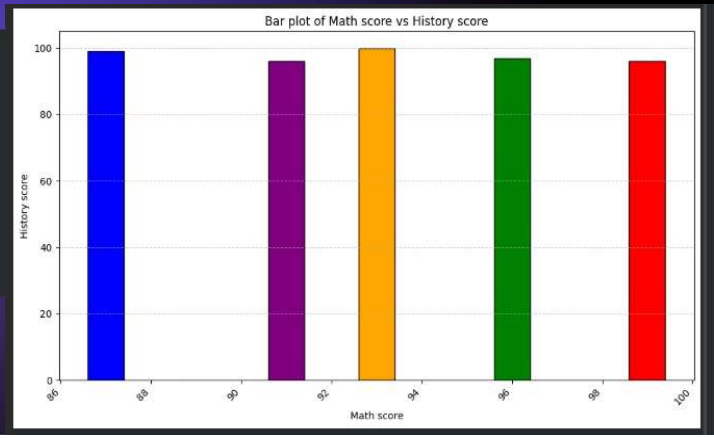
Scatter plot:

```
import matplotlib.pyplot as plt
colours = ['blue', 'purple', 'green', 'red', 'orange']
plt.figure(figsize=(10, 6))
#Assign colours based on the index of top_5_students DataFrame
plt.scatter(x=top_5_students['math_score'], y=top_5_students['history_score'], c=colours[:len(top_5_students)], marker='o', s=100) # s is marker size
plt.xlabel('Math score')
plt.ylabel('History score')
plt.title('Scatter plot of Math score vs History score')
plt.grid(True)
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



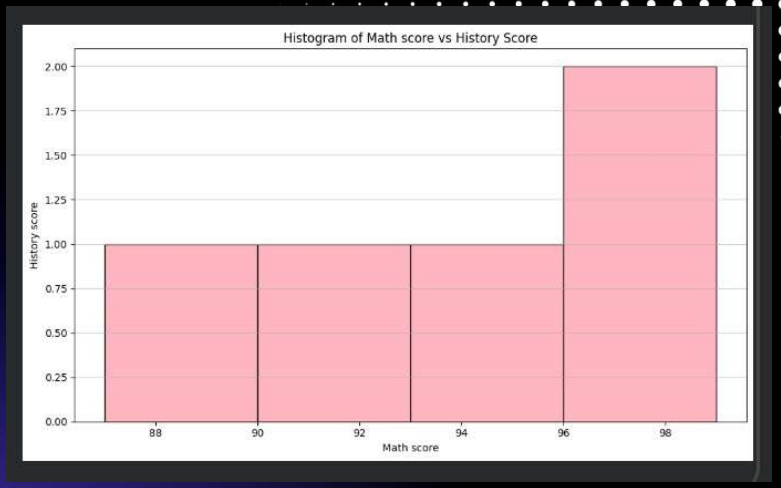
Bar plot:

```
import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))
plt.bar(x=top_5_students['math_score'], height=top_5_students['history_score'], color=['blue', 'purple', 'green', 'red', 'orange'], edgecolor = 'Black')
plt.xlabel('Math score')
plt.ylabel('History score')
plt.title('Bar plot of Math score vs History score')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



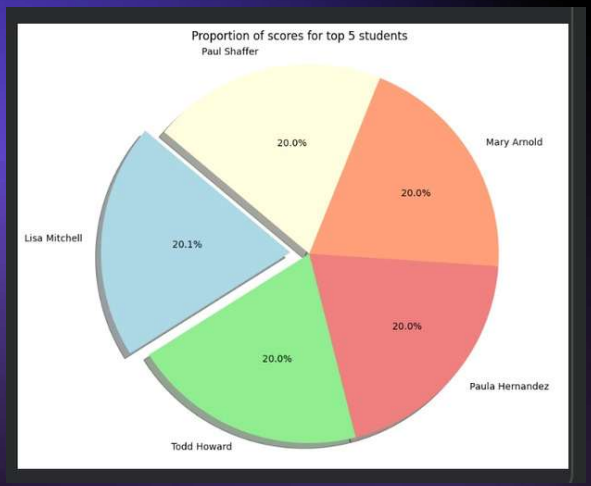
Histogram:

```
# A histogram of 5 values will show 5 bars, one of each score
import matplotlib.pyplot as plt
plt.figure(figsize=(10, 6))
plt.hist(top_5_students['math_score'], bins=len(top_5_students['history_score'].unique()), color='lightpink', edgecolor='black')
plt.xlabel('Math score')
plt.ylabel('History score')
plt.title('Histogram of Math score vs History Score')
plt.grid(axis='y', alpha=0.75)
plt.tight_layout()
plt.show()
```



Pie plot:

```
import matplotlib.pyplot as plt

top_5_students['Name'] = top_5_students['first_name'] + ' ' + top_5_students['last_name']
labels = top_5_students['Name']
sizes = top_5_students['total score']
colors = ['lightblue', 'lightgreen', 'lightcoral', 'lightsalmon', 'lightyellow']
explode = (0.1, 0, 0, 0, 0)
plt.figure(figsize=(10, 8))
plt.pie(sizes, explode=explode, labels=labels, colors=colors,
        autopct='%11.1f%%', shadow=True, startangle=140)
plt.axis('equal')
plt.title('Proportion of scores for top 5 students')
plt.show()
```

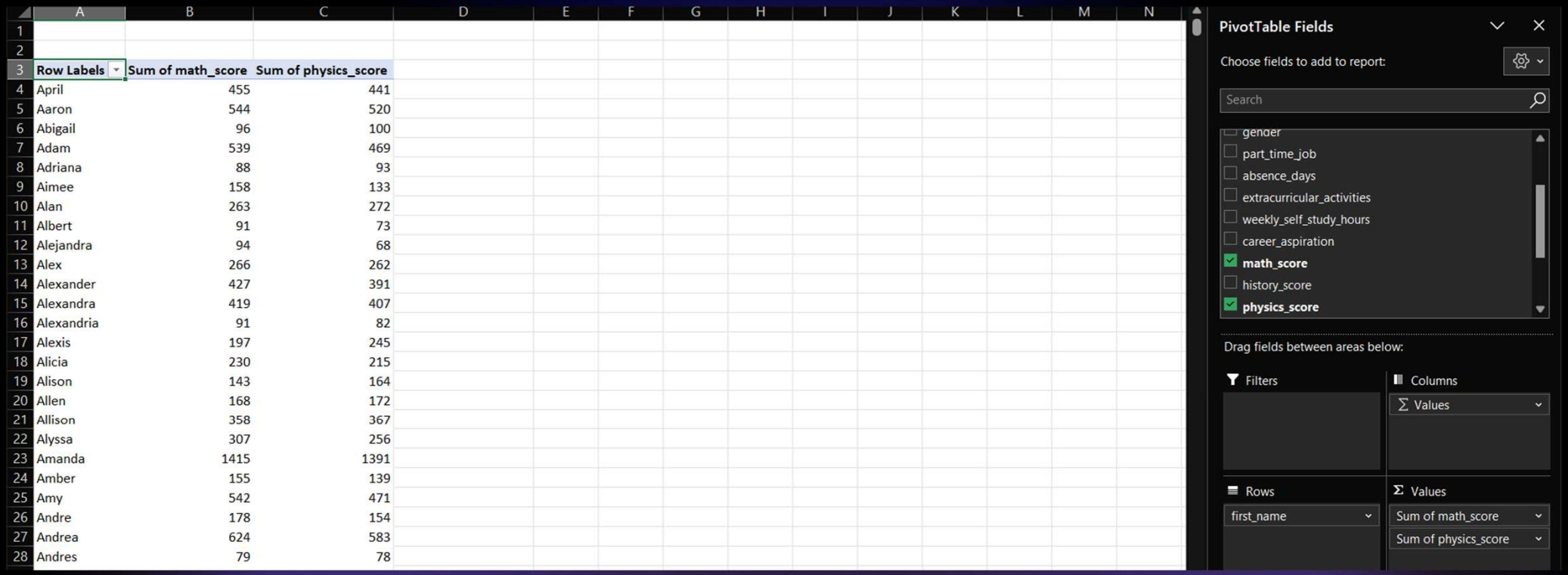


Working on Excel

Conditional Formatting:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	id	first_name	last_name	email	gender	part_time_job	absence	extracurricular_a	weekly_	career_aspiration	math_score	history_score	physics_score	chemistry_	biology_sc	english_sc	geography_sc
2	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	FALSE	3	FALSE	27	Lawyer	73	81	93	97	63	80	87
3	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	FALSE	2	FALSE	47	Doctor	90	86	96	100	90	88	90
4	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	FALSE	9	TRUE	13	Government Officer	81	97	95	96	65	77	94
5	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	FALSE	5	FALSE	3	Artist	71	74	88	80	89	63	86
6	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	FALSE	5	FALSE	10	Unknown	84	77	65	65	80	74	76
7	6	Kelly	Wade	kelly.wade.6@gslingacademy.com	female	FALSE	2	FALSE	26	Unknown	93	100	67	78	72	80	84
8	7	Anthony	Smith	anthony.smith.7@gslingacademy.com	male	FALSE	3	TRUE	23	Software Engineer	99	96	97	73	88	76	64
9	8	George	Short	george.short.8@gslingacademy.com	male	TRUE	2	TRUE	34	Software Engineer	95	95	82	63	84	70	85
10	9	Stanley	Gutierrez	stanley.gutierrez.9@gslingacademy.com	male	FALSE	6	FALSE	25	Unknown	94	68	94	85	81	74	72
11	10	Audrey	Simpson	audrey.simpson.10@gslingacademy.com	female	FALSE	3	TRUE	18	Teacher	98	69	88	71	67	71	73
12	11	Gabrielle	White	gabrielle.white.11@gslingacademy.com	female	FALSE	2	FALSE	7	Teacher	65	60	97	94	71	81	66
13	12	Clinton	Randolph	clinton.randolph.12@gslingacademy.com	male	FALSE	1	FALSE	7	Unknown	80	61	100	65	87	64	61
14	13	Patricia	Gomez	patricia.gomez.13@gslingacademy.com	female	TRUE	7	FALSE	4	Business Owner	94	59	69	67	89	65	73
15	14	Pamela	Jackson	pamela.jackson.14@gslingacademy.com	female	FALSE	10	FALSE	2	Business Owner	66	94	86	100	57	90	63
16	15	Laura	Jackson	laura.jackson.15@gslingacademy.com	female	FALSE	3	FALSE	39	Doctor	96	90	86	92	92	95	87
17	16	Roger	Wiley	roger.wiley.16@gslingacademy.com	male	FALSE	6	FALSE	0	Business Owner	94	50	78	64	79	74	84
18	17	Vicki	Thompson	vicki.thompson.17@gslingacademy.com	female	FALSE	3	TRUE	30	Scientist	92	64	93	91	80	89	72
19	18	Maxwell	Davidson	maxwell.davidson.18@gslingacademy.com	male	FALSE	2	TRUE	28	Software Engineer	86	83	85	79	93	76	77
20	19	Jonathan	Werner	jonathan.werner.19@gslingacademy.com	male	FALSE	1	FALSE	37	Doctor	92	87	92	99	97	87	86
21	20	Angela	Rios	angela.rios.20@gslingacademy.com	female	FALSE	2	FALSE	27	Software Engineer	99	65	98	75	66	72	100
22	21	Tim	Nichols	tim.nichols.21@gslingacademy.com	male	TRUE	3	FALSE	15	Software Engineer	100	90	72	98	73	97	72
23	22	Kyle	Willis	kyle.willis.22@gslingacademy.com	male	FALSE	8	FALSE	4	Business Owner	57	55	78	94	83	88	88
24	23	Shannon	Simpson	shannon.simpson.23@gslingacademy.com	female	FALSE	9	FALSE	2	Business Owner	89	72	68	72	71	54	90
25	24	Sean	Griffin	sean.griffin.24@gslingacademy.com	male	FALSE	9	FALSE	1	Business Owner	50	76	81	55	56	80	79
26	25	Cassandra	West	cassandra.west.25@gslingacademy.com	female	FALSE	4	FALSE	35	Software Engineer	87	91	90	88	95	88	93
27	26	Patricia	Chavez	patricia.chavez.26@gslingacademy.com	female	FALSE	5	FALSE	22	Doctor	92	86	87	81	93	90	99
28	27	Jason	Williams	jason.williams.27@gslingacademy.com	male	FALSE	3	FALSE	34	Banker	100	77	80	94	63	90	90
29	28	Peter	Gibbs	peter.gibbs.28@gslingacademy.com	male	FALSE	0	FALSE	23	Writer	64	75	93	79	81	96	85
30	29	Jeffrey	Blanchard	jeffrey.blanchard.29@gslingacademy.com	male	FALSE	1	TRUE	17	Accountant	79	65	99	71	76	77	83

Pivot Table:

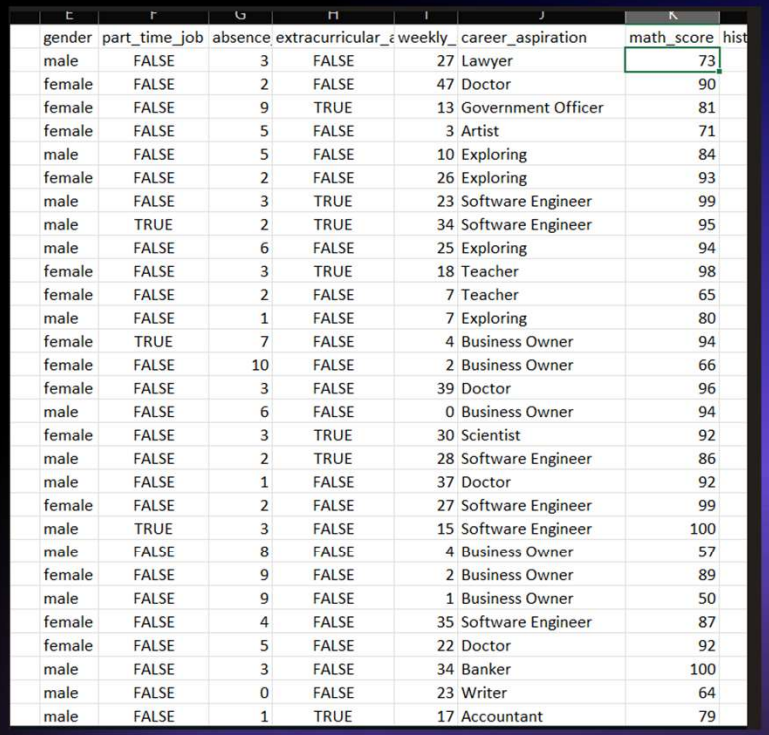


Row Labels	Sum of math_score	Sum of physics_score
April	455	441
Aaron	544	520
Abigail	96	100
Adam	539	469
Adriana	88	93
Aimee	158	133
Alan	263	272
Albert	91	73
Alejandra	94	68
Alex	266	262
Alexander	427	391
Alexandra	419	407
Alexandria	91	82
Alexis	197	245
Alicia	230	215
Alison	143	164
Allen	168	172
Allison	358	367
Alyssa	307	256
Amanda	1415	1391
Amber	155	139
Amy	542	471
Andre	178	154
Andrea	624	583
Andres	79	78

Sorting the names Ascending Order:

1	id	Aaron	last_name	email	gender	part_time_job	absence	extracurricular
2	1	Aaron	Casey	paul.casey.1@gslingacademy.com	male	FALSE	3	FALSE
3	2	Aaron	Sandoval	danielle.sandoval.2@gslingacademy.com	female	FALSE	2	FALSE
4	3	Aaron	Andrews	tina.andrews.3@gslingacademy.com	female	FALSE	9	TRUE
5	4	Aaron	Clark	tara.clark.4@gslingacademy.com	female	FALSE	5	FALSE
6	5	Aaron	Campos	anthony.campos.5@gslingacademy.com	male	FALSE	5	FALSE
7	6	Aaron	Wade	kelly.wade.6@gslingacademy.com	female	FALSE	2	FALSE
8	7	Abigail	Smith	anthony.smith.7@gslingacademy.com	male	FALSE	3	TRUE
9	8	Adam	Short	george.short.8@gslingacademy.com	male	TRUE	2	TRUE
10	9	Adam	Gutierrez	stanley.gutierrez.9@gslingacademy.com	male	FALSE	6	FALSE
11	10	Adam	Simpson	audrey.simpson.10@gslingacademy.com	female	FALSE	3	TRUE
12	11	Adam	White	gabrielle.white.11@gslingacademy.com	female	FALSE	2	FALSE
13	12	Adam	Randolph	clinton.randolph.12@gslingacademy.com	male	FALSE	1	FALSE
14	13	Adam	Gomez	patricia.gomez.13@gslingacademy.com	female	TRUE	7	FALSE
15	14	Adriana	Jackson	pamela.jackson.14@gslingacademy.com	female	FALSE	10	FALSE
16	15	Aimee	Jackson	laura.jackson.15@gslingacademy.com	female	FALSE	3	FALSE
17	16	Aimee	Wiley	roger.wiley.16@gslingacademy.com	male	FALSE	6	FALSE
18	17	Alan	Thompson	vicki.thompson.17@gslingacademy.com	female	FALSE	3	TRUE
19	18	Alan	Davidson	maxwell.davidson.18@gslingacademy.com	male	FALSE	2	TRUE
20	19	Alan	Werner	jonathan.werner.19@gslingacademy.com	male	FALSE	1	FALSE
21	20	Albert	Rios	angela.rios.20@gslingacademy.com	female	FALSE	2	FALSE
22	21	Alejandra	Nichols	tim.nichols.21@gslingacademy.com	male	TRUE	3	FALSE
23	22	Alex	Willis	kyle.willis.22@gslingacademy.com	male	FALSE	8	FALSE
24	23	Alex	Simpson	shannon.simpson.23@gslingacademy.com	female	FALSE	9	FALSE
25	24	Alex	Griffin	sean.griffin.24@gslingacademy.com	male	FALSE	9	FALSE
26	25	Alexander	West	cassandra.west.25@gslingacademy.com	female	FALSE	4	FALSE
27	26	Alexander	Chavez	patricia.chavez.26@gslingacademy.com	female	FALSE	5	FALSE
28	27	Alexander	Williams	jason.williams.27@gslingacademy.com	male	FALSE	3	FALSE
29	28	Alexander	Gibbs	peter.gibbs.28@gslingacademy.com	male	FALSE	0	FALSE
30	29	Alexander	Blanchard	jeffrey.blanchard.29@gslingacademy.com	male	FALSE	1	TRUE

Replacing Text:



gender	part_time_job	absence	extracurricular_a	weekly_	career_aspiration	math_score	hist
male	FALSE	3	FALSE	27	Lawyer	73	
female	FALSE	2	FALSE	47	Doctor	90	
female	FALSE	9	TRUE	13	Government Officer	81	
female	FALSE	5	FALSE	3	Artist	71	
male	FALSE	5	FALSE	10	Exploring	84	
female	FALSE	2	FALSE	26	Exploring	93	
male	FALSE	3	TRUE	23	Software Engineer	99	
male	TRUE	2	TRUE	34	Software Engineer	95	
male	FALSE	6	FALSE	25	Exploring	94	
female	FALSE	3	TRUE	18	Teacher	98	
female	FALSE	2	FALSE	7	Teacher	65	
male	FALSE	1	FALSE	7	Exploring	80	
female	TRUE	7	FALSE	4	Business Owner	94	
female	FALSE	10	FALSE	2	Business Owner	66	
female	FALSE	3	FALSE	39	Doctor	96	
male	FALSE	6	FALSE	0	Business Owner	94	
female	FALSE	3	TRUE	30	Scientist	92	
male	FALSE	2	TRUE	28	Software Engineer	86	
male	FALSE	1	FALSE	37	Doctor	92	
female	FALSE	2	FALSE	27	Software Engineer	99	
male	TRUE	3	FALSE	15	Software Engineer	100	
male	FALSE	8	FALSE	4	Business Owner	57	
female	FALSE	9	FALSE	2	Business Owner	89	
male	FALSE	9	FALSE	1	Business Owner	50	
female	FALSE	4	FALSE	35	Software Engineer	87	
female	FALSE	5	FALSE	22	Doctor	92	
male	FALSE	3	FALSE	34	Banker	100	
male	FALSE	0	FALSE	23	Writer	64	
male	FALSE	1	TRUE	17	Accountant	79	

Duplicate Last Names:

	A	B	C	D	E	F	G	H	I	J
1	id	first_name	last_name	email	gender	part_time_job	absence	extracurricular_a	weekly_	career_aspiration
2	1	Paul	Casey	paul.casey.1@gslingacademy.com	male	FALSE	3	FALSE	27	Lawyer
3	2	Danielle	Sandoval	danielle.sandoval.2@gslingacademy.com	female	FALSE	2	FALSE	47	Doctor
4	3	Tina	Andrews	tina.andrews.3@gslingacademy.com	female	FALSE	9	TRUE	13	Government Officer
5	4	Tara	Clark	tara.clark.4@gslingacademy.com	female	FALSE	5	FALSE	3	Artist
6	5	Anthony	Campos	anthony.campos.5@gslingacademy.com	male	FALSE	5	FALSE	10	Unknown
7	6	Kelly	Wade	kelly.wade.6@gslingacademy.com	female	FALSE	2	FALSE	26	Unknown
8	7	Anthony	Smith	anthony.smith.7@gslingacademy.com	male	FALSE	3	TRUE	23	Software Engineer
9	8	George	Short	george.short.8@gslingacademy.com	male	TRUE	2	TRUE	34	Software Engineer
10	9	Stanley	Gutierrez	stanley.gutierrez.9@gslingacademy.com	male	FALSE	6	FALSE	25	Unknown
11	10	Audrey	Simpson	audrey.simpson.10@gslingacademy.com	female	FALSE	3	TRUE	18	Teacher
12	11	Gabrielle	White	gabrielle.white.11@gslingacademy.com	female	FALSE	2	FALSE	7	Teacher
13	12	Clinton	Randolph	clinton.randolph.12@gslingacademy.com	male	FALSE	1	FALSE	7	Unknown
14	13	Patricia	Gomez	patricia.gomez.13@gslingacademy.com	female	TRUE	7	FALSE	4	Business Owner
15	14	Pamela	Jackson	pamela.jackson.14@gslingacademy.com	female	FALSE	10	FALSE	2	Business Owner
16	15	Laura	Jackson	laura.jackson.15@gslingacademy.com	female	FALSE	3	FALSE	39	Doctor
17	16	Roger	Wiley	roger.wiley.16@gslingacademy.com	male	FALSE	6	FALSE	0	Business Owner
18	17	Vicki	Thompson	vicki.thompson.17@gslingacademy.com	female	FALSE	3	TRUE	30	Scientist
19	18	Maxwell	Davidson	maxwell.davidson.18@gslingacademy.com	male	FALSE	2	TRUE	28	Software Engineer
20	19	Jonathan	Werner	jonathan.werner.19@gslingacademy.com	male	FALSE	1	FALSE	37	Doctor
21	20	Angela	Rios	angela.rios.20@gslingacademy.com	female	FALSE	2	FALSE	27	Software Engineer
22	21	Tim	Nichols	tim.nichols.21@gslingacademy.com	male	TRUE	3	FALSE	15	Software Engineer
23	22	Kyle	Willis	kyle.willis.22@gslingacademy.com	male	FALSE	8	FALSE	4	Business Owner
24	23	Shannon	Simpson	shannon.simpson.23@gslingacademy.com	female	FALSE	9	FALSE	2	Business Owner
25	24	Sean	Griffin	sean.griffin.24@gslingacademy.com	male	FALSE	9	FALSE	1	Business Owner
26	25	Cassandra	West	cassandra.west.25@gslingacademy.com	female	FALSE	4	FALSE	35	Software Engineer
27	26	Patricia	Chavez	patricia.chavez.26@gslingacademy.com	female	FALSE	5	FALSE	22	Doctor
28	27	Jason	Williams	jason.williams.27@gslingacademy.com	male	FALSE	3	FALSE	34	Banker
29	28	Peter	Gibbs	peter.gibbs.28@gslingacademy.com	male	FALSE	0	FALSE	23	Writer
30	29	Jeffrey	Blanchard	jeffrey.blanchard.29@gslingacademy.com	male	FALSE	1	TRUE	17	Accountant

Conclusion:

In conclusion, the student performane analysis successfully hightlighted the significant factors that affect academic results,such as study hours, attendence, and parential involvement. The findings emphasize the importance of continueous assessment and personalized learning approaches. With data-driven strategies, schools and educators can better support students indevidual needs and improve overall academic performance.

THANK YOU.....