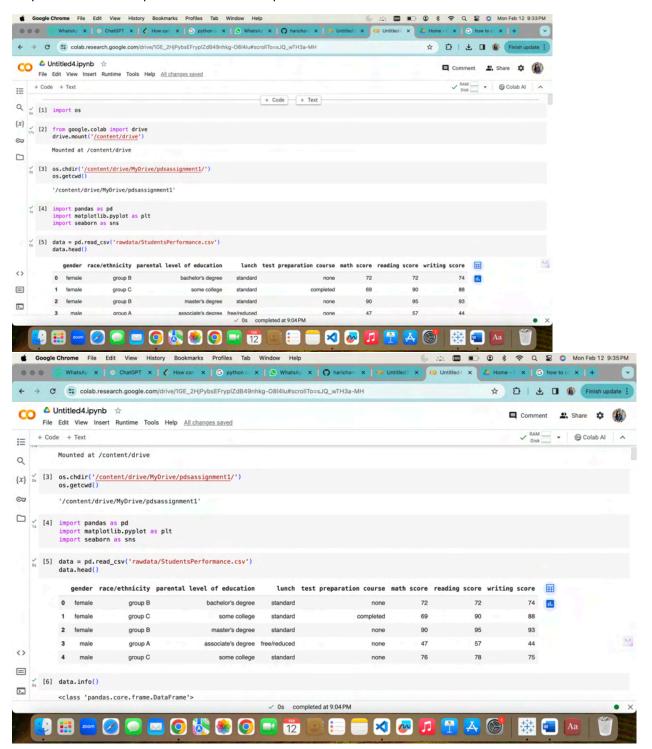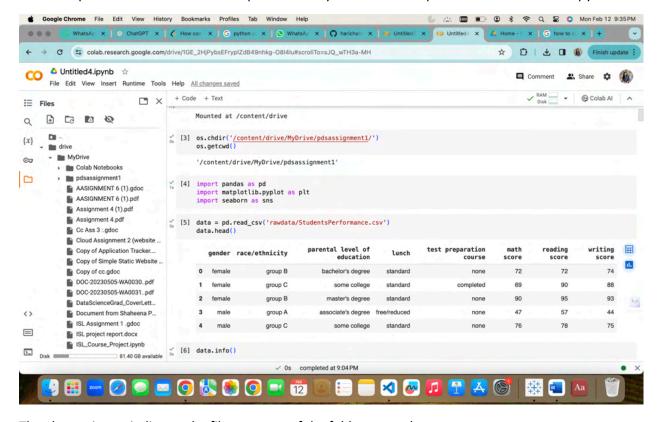2) Perform 5 data visualization tasks on the student performance dataset given in the link below (create 5 different visualizations). Explain what kind analysis has become easier with each of the visualizations. Create the folder structure for this question like question 1. (15 points)
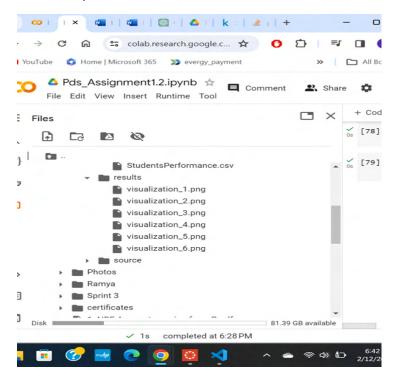
Step 1:Collection of input data from csv file provided.

Above step indicates the utilization of pandas library to load studentperformance.csv file into python
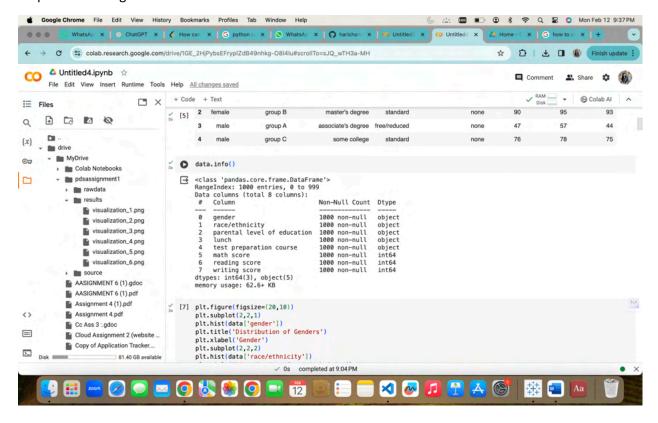

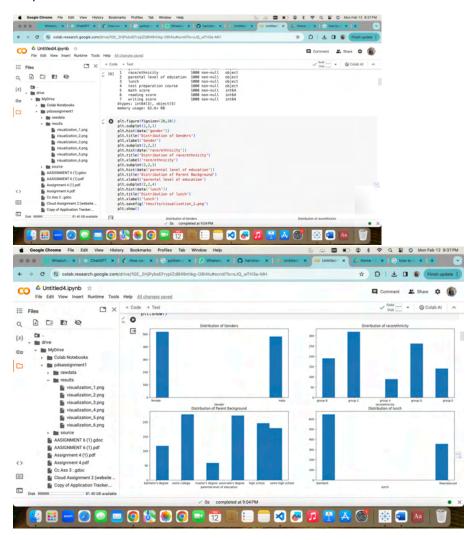The above picture indicates the file structure of the folder created.

This is the file structure indicates the final folder structure with all data visualization images obatined.

Step 2: Processing of the data

	2	female	group B	master's degree	standard	none	90	95	93
[5]	3	male	group A	associate's degree	free/reduced	none	47	57	44
	4	male	group C	some college	standard	none	76	78	75

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 8 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   gender                       1000 non-null   object
 1   race/ethnicity               1000 non-null   object
 2   parental level of education  1000 non-null   object
 3   lunch                        1000 non-null   object
 4   test preparation course      1000 non-null   object
 5   math score                   1000 non-null   int64
 6   reading score                1000 non-null   int64
 7   writing score                1000 non-null   int64
dtypes: int64(3), object(5)
memory usage: 62.6+ KB
```

```
[7] plt.figure(figsize=(20,10))
    plt.subplot(2,2,1)
    plt.hist(data['gender'])
    plt.title('Distribution of Genders')
    plt.xlabel('Gender')
    plt.subplot(2,2,2)
    plt.hist(data['race/ethnicity'])
```
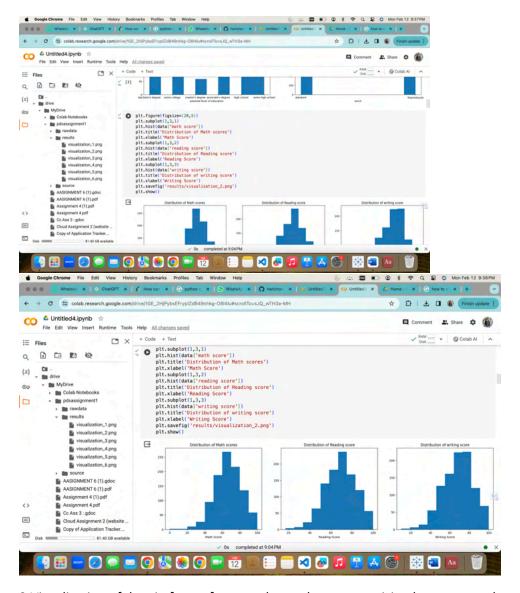
✓ 0s   completed at 9:04 PM
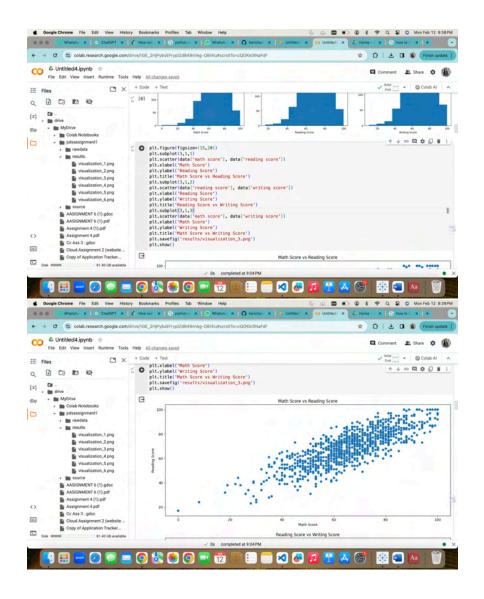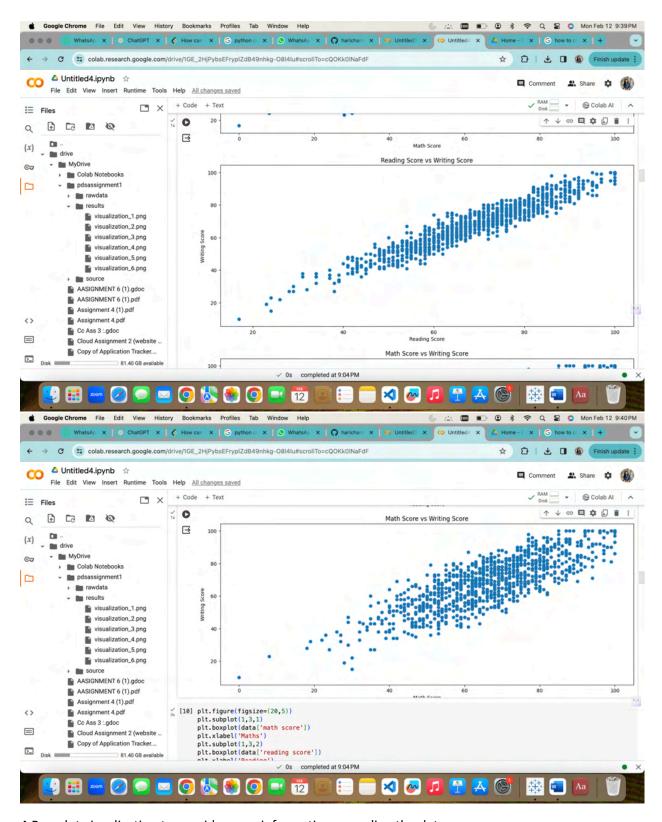
Step 3:Visualizations of data

2.Visualization to show distribution between the maths,reading and writing scores on a bar graph

3.Visualization of data in form of scatterplot to show comparision between maths vs reading vs writing scores.

```python
plt.figure(figsize=(15,20))
plt.subplot(3,1,1)
plt.scatter(data['math score'], data['reading score'])
plt.xlabel('Math Score')
plt.ylabel('Reading Score')
plt.title('Math Score vs Reading Score')
plt.subplot(3,1,2)
plt.scatter(data['reading score'], data['writing score'])
plt.xlabel('Reading Score')
plt.ylabel('Writing Score')
plt.title('Reading Score vs Writing Score')
plt.subplot(3,1,3)
plt.scatter(data['math score'], data['writing score'])
plt.xlabel('Math Score')
plt.ylabel('Writing Score')
plt.title('Math Score vs Writing Score')
plt.savefig('results/visualization_3.png')
plt.show()
```

Math Score vs Reading Score

```python
plt.xlabel('Math Score')
plt.ylabel('Writing Score')
plt.title('Math Score vs Writing Score')
plt.savefig('results/visualization_3.png')
plt.show()
```


Math Score vs Reading Score

Reading Score vs Writing Score

4.Box plot visualization to provide more information regarding the data.

5.Bar chart visualization to provide relation between number of students and test preparation course.

```
prep_course_count = data['test preparation course'].value_counts()
plt.bar(prep_course_count.index, prep_course_count.values)
plt.xlabel('Test Preparation Course')
plt.ylabel('Number of Students')
plt.title('Number of Students Who Completed Test Preparation Course')
plt.savefig('results/visualization_5.png')
plt.show()
```
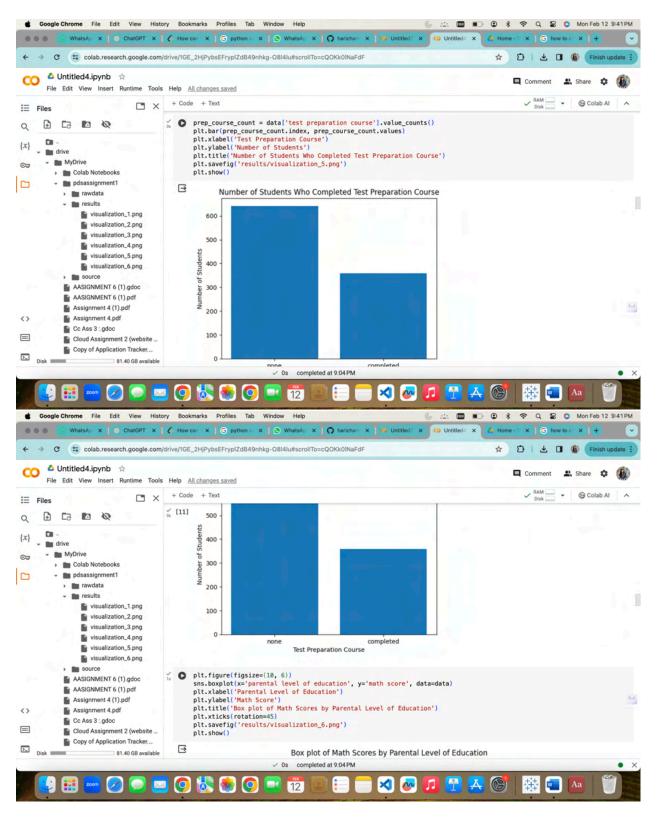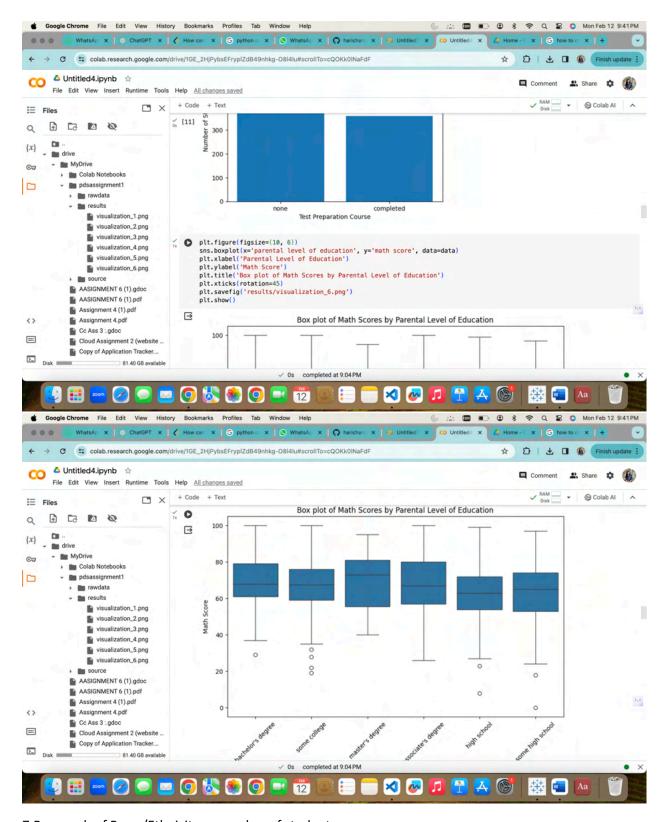
Number of Students Who Completed Test Preparation Course


```
plt.figure(figsize=(10, 6))
sns.boxplot(x='parental level of education', y='math score', data=data)
plt.xlabel('Parental Level of Education')
plt.ylabel('Math Score')
plt.title('Box plot of Math Scores by Parental Level of Education')
plt.xticks(rotation=45)
plt.savefig('results/visualization_6.png')
plt.show()
```

Box plot of Math Scores by Parental Level of Education

6.Box Plot Visualization ,this visualization can help identify whether there is a differnce in test scores between students whose parents have different levels of education.

```python
plt.figure(figsize=(10, 6))
sns.boxplot(x='parental level of education', y='math score', data=data)
plt.xlabel('Parental Level of Education')
plt.ylabel('Math Score')
plt.title('Box plot of Math Scores by Parental Level of Education')
plt.xticks(rotation=45)
plt.savefig('results/visualization_6.png')
plt.show()
```


7.Bar graph of Race /Ethnicity vs number of students

```python
counts = data['race/ethnicity'].value_counts()

plt.bar(counts.index, counts.values)
plt.xlabel('Race/Ethnicity')
plt.ylabel('Number of Students')
plt.title('Number of Students in each Race/Ethnicity Category')
plt.show()
```



```python
plt.title('Number of Students in each Race/Ethnicity Category')
plt.show()
```

```python
[15] plt.hist(data['writing score'], bins=10)
     plt.xlabel('Writing Score')
```

8.histogram of writing score and frequency

9.Box Plot of reading scores by parental level of Education

```python
plt.figure(figsize=(10, 6))
sns.boxplot(x='parental level of education', y='reading score', data=data)
plt.xlabel('Parental Level of Education')
plt.ylabel('Reading Score')
plt.title('Box plot of Reading Scores by Parental Level of Education')
plt.show()
```


10.pairplot where hue is gender.

11. volin plot of parental level of education and reading score.

12. Box plot of writing scores by parental level of Education

```
plt.figure(figsize=(10, 6))
sns.boxplot(x='parental level of education', y='writing score', data=data)
plt.xlabel('Parental Level of Education')
plt.ylabel('writing Score')
plt.title('Box plot of Writing Scores by Parental Level of Education')
plt.show()
```


13.Distrubution of math ,reading and writing scores.

Untitled4.ipynb ☆
File Edit View Insert Runtime Tools Help   All changes saved

+ Code   + Text

[21]   20

bachelor's degree   some college   master's degree   associate's degree   high school   some high school
Parental Level of Education

[ ] Start coding or generate with AI.

```python
plt.figure(figsize=(20,5))
plt.subplot(1,3,1)
plt.hist(data['math score'])
plt.title('Distribution of Math scores')
plt.xlabel('Math Score')
plt.subplot(1,3,2)
plt.hist(data['reading score'])
plt.title('Distribution of Reading score')
plt.xlabel('Reading Score')
plt.subplot(1,3,3)
plt.hist(data['writing score'])
plt.title('Distribution of writing score')
plt.xlabel('Writing Score')
plt.savefig('results/visualization_2.png')
plt.show()
```

Distribution of Math scores   Distribution of Reading score   Distribution of writing score

✓ 0s   completed at 9:04 PM

Untitled4.ipynb ☆
File Edit View Insert Runtime Tools Help   All changes saved

+ Code   + Text

```python
plt.title('Distribution of Reading score')
plt.xlabel('Reading Score')
plt.subplot(1,3,3)
plt.hist(data['writing score'])
plt.title('Distribution of writing score')
plt.xlabel('Writing Score')
plt.savefig('results/visualization_2.png')
plt.show()
```

Distribution of Math scores   Distribution of Reading score   Distribution of writing score

✓ 0s   completed at 9:04 PM