In [28]:
```
#Importing Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
#Reading the Given Csv
#Load Dataset
df=pd.read_csv("E:/Student/student-mat.csv")
df.head()
```

Out[28]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3

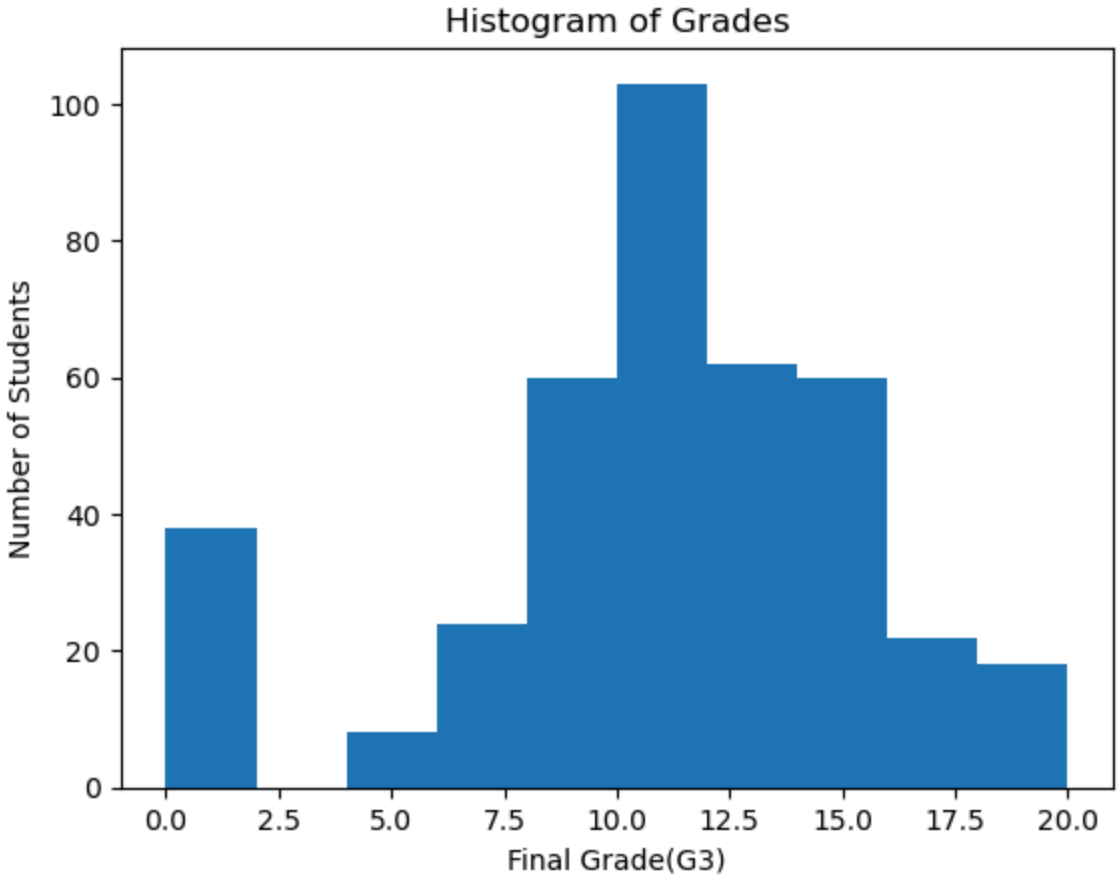
5 rows × 33 columns

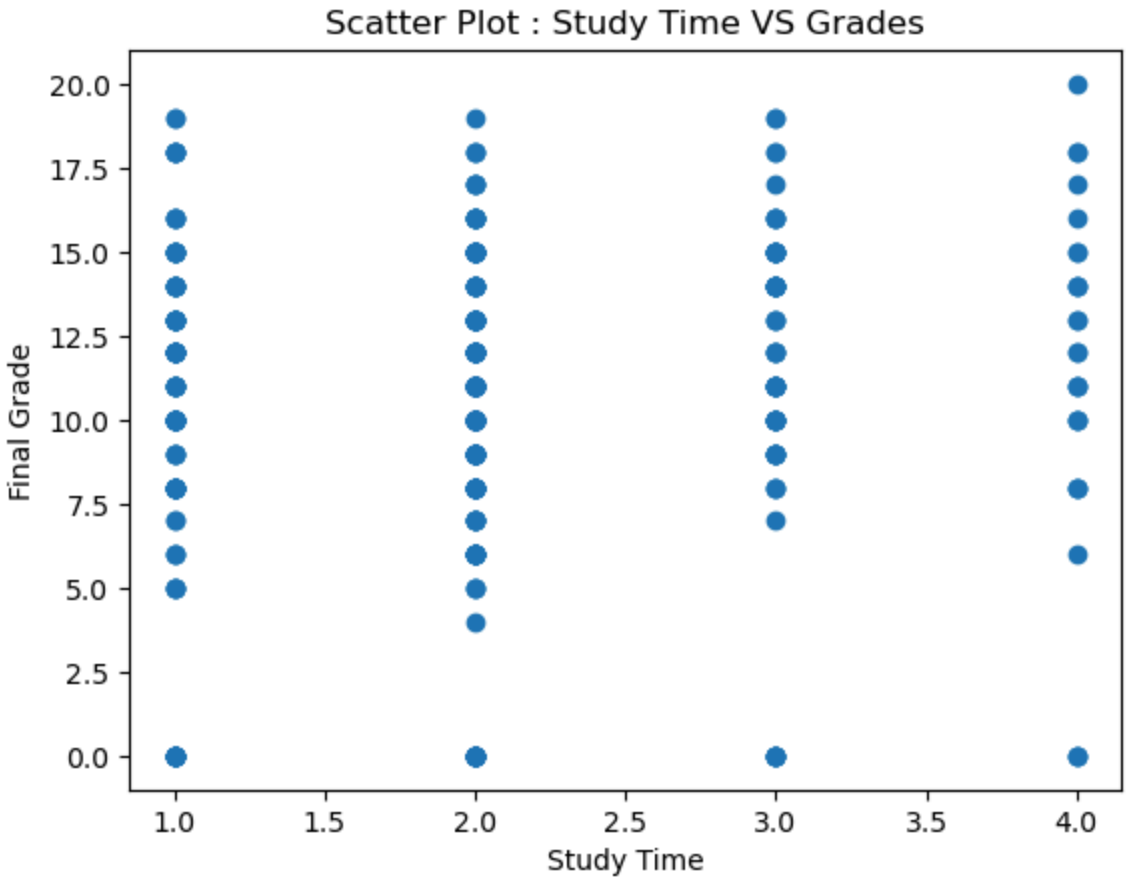
In [29]:
```
#EXPLORE AND CLEAN DATA:
#Check missing values(.isnull())
print("CHECKING FOR MISSING VALUES")
print(df.isnull().sum())

#Remove Duplicates
print("CHECKING FOR DUPLICATES")
print(df.duplicated().sum())
print("If there exists any duplicates, Remove them")
df=df.drop_duplicates()

#Inspect dataset Shape and dtypes
print("INSPECTING DATASET SHAPE")
print(df.shape)

print("INSPECTING DATASET DTYPES")
print(df.dtypes)


#ANALYSIS QUESTIONS:
#Average Final Grade
avg_g3=df['G3'].mean()
print("Average Final Grade :",avg_g3)

#How many students scored above 15?
above_15=df[df['G3']>15].shape[0]
print("Students scored above 15 :",above_15)

#Is study time correlated with performance?
correlation=df['studytime'].corr(df['G3'])
print("The Correlation between the Study Time and Final grade :",correlation)

#Which Gender performs better on Average
```

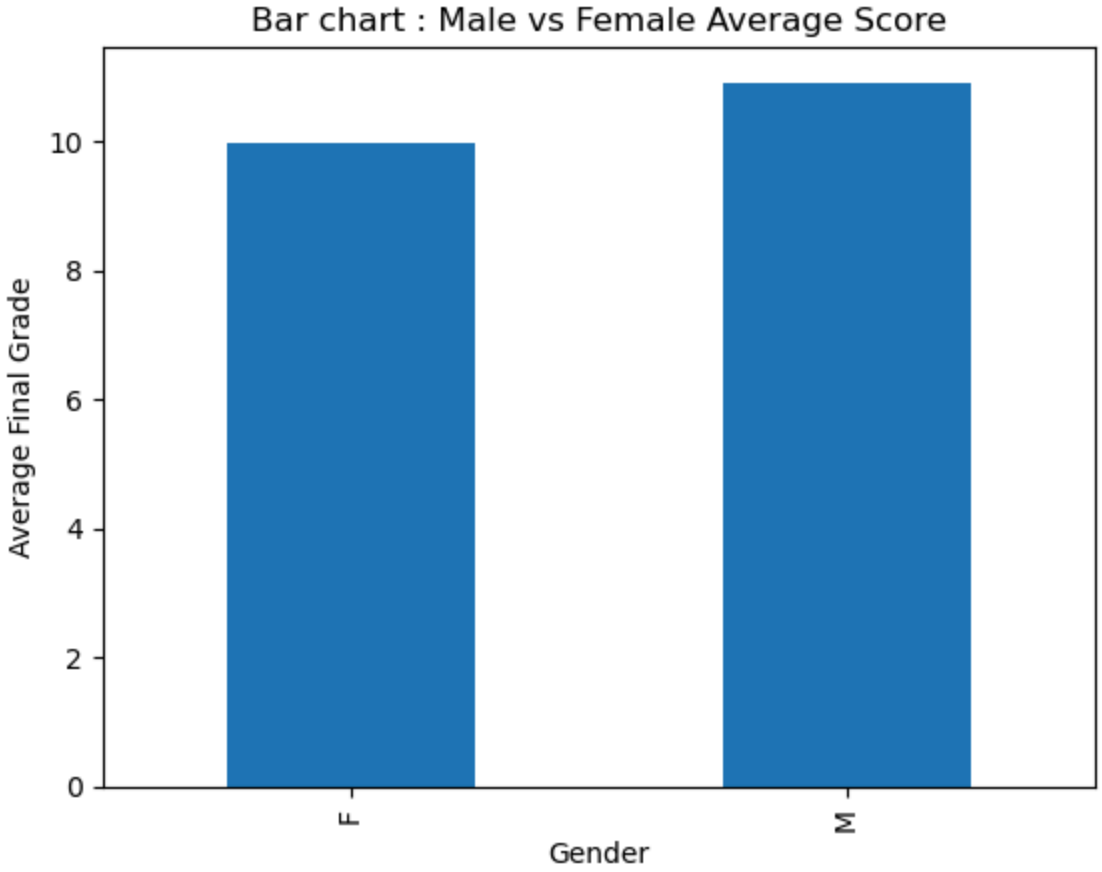
```
print("Average of Performance based upon Gender",df.groupby('sex')['G3'].mean())
print("M denotes MALE and F denotes FEMALE")

#VISUALIZATIONS:
#Histrogram of Grades
plt.figure()
plt.hist(df['G3'],bins=10)
plt.xlabel("Final Grade(G3)")
plt.ylabel("Number of Students")
plt.title("Histogram of Grades")
plt.show()

#Scatter Plot : Study Time VS Grades
plt.figure()
plt.scatter(df['studytime'],df['G3'])
plt.xlabel("Study Time")
plt.ylabel("Final Grade")
plt.title("Scatter Plot : Study Time VS Grades")
plt.show()

#Bar chart : Male vs Female Average Score
gender_avg=df.groupby('sex')['G3'].mean()
plt.figure()
gender_avg.plot(kind='bar')
plt.xlabel("Gender")
plt.ylabel("Average Final Grade")
plt.title("Bar chart : Male vs Female Average Score")
plt.show()
```

```
CHECKING FOR MISSING VALUES
school        0
sex           0
age           0
address       0
famsize       0
Pstatus       0
Medu          0
Fedu          0
Mjob          0
Fjob          0
reason        0
guardian      0
traveltime    0
studytime     0
failures      0
schoolsup     0
famsup        0
paid          0
activities    0
nursery       0
higher        0
internet      0
romantic      0
famrel        0
freetime      0
goout         0
Dalc          0
Walc          0
health        0
absences      0
G1            0
G2            0
G3            0
dtype: int64
CHECKING FOR DUPLICATES
0
If there exists any duplicates, Remove them
INSPECTING DATASET SHAPE
(395, 33)
INSPECTING DATASET DTYPES
school        object
sex           object
age            int64
address       object
famsize       object
Pstatus       object
Medu           int64
Fedu           int64
Mjob          object
Fjob          object
reason        object
guardian      object
traveltime     int64
studytime      int64
failures       int64
schoolsup     object
famsup        object
paid          object
activities    object
```

```
nursery       object
higher        object
internet      object
romantic      object
famrel         int64
freetime       int64
goout          int64
Dalc           int64
Walc           int64
health         int64
absences       int64
G1             int64
G2             int64
G3             int64
dtype: object
Average Final Grade : 10.415189873417722
Students scored above 15 : 40
The Correlation between the Study Time and Final grade : 0.0978196896531963
Average of Performance based upon Gender sex
F     9.966346
M    10.914439
Name: G3, dtype: float64
M denotes MALE and F denotes FEMALE
```




Scatter Plot : Study Time VS Grades
Final Grade
Study Time
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Bar chart : Male vs Female Average Score
Average Final Grade
Gender
10
8
6
4
2
0
F
M