BÀI GIẢI ÔN TẬP

```python
import matplotlib.pyplot as plt
import pandas as pd

# read file, set new columns and skip 1st row
column_names = ['Gender', 'Group', 'Edu', 'Lunch', 'Test', 'Math', 'Reading',
'Writing']
df = pd.read_csv("exams.csv", names=column_names, skiprows=1)

# 1. Clean data: remove rows if value in one of column belows column is NA
df.dropna(subset=['Gender', 'Group', 'Edu', 'Math', 'Reading', 'Writing'],
inplace=True)

# 2. Average of Math, Reading and Writing
print(df[['Math', 'Reading', 'Writing']].mean())

Result:
Math        67.068
Reading     69.783
Writing     68.447
dtype: float64

# 3. Get gender rate then show pie chart
gender_rate = df.groupby(['Gender'])['Gender'].count()
# split if value == max
explode = gender_rate.apply(lambda x: 0.1 if x == gender_rate.max() else 0)
# get label and convert title case
chart_labels = gender_rate.index.str.title()
```
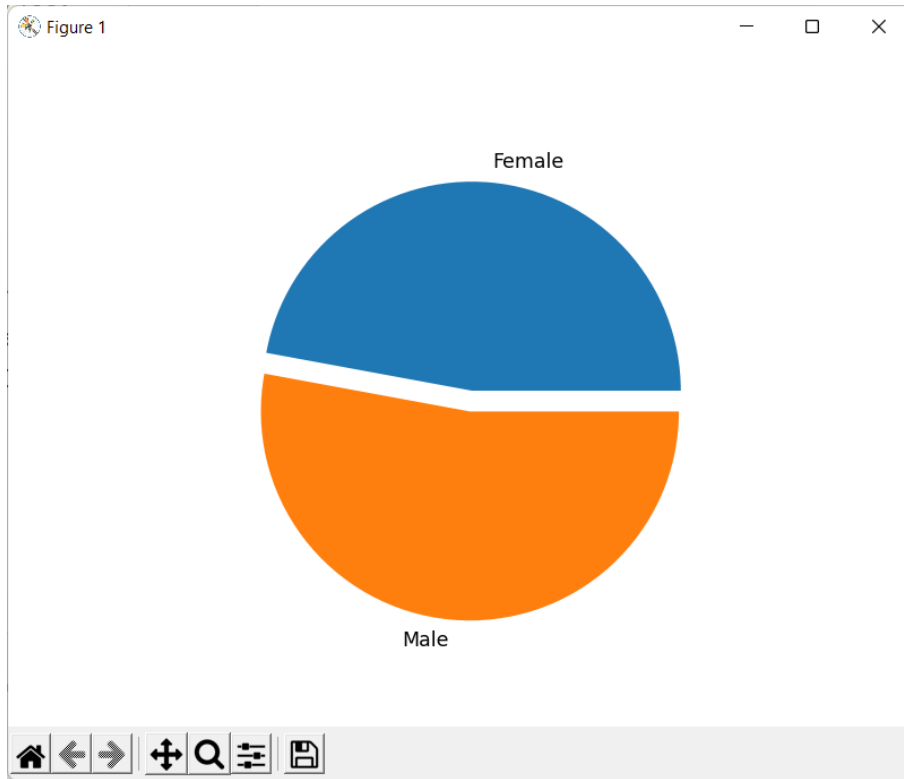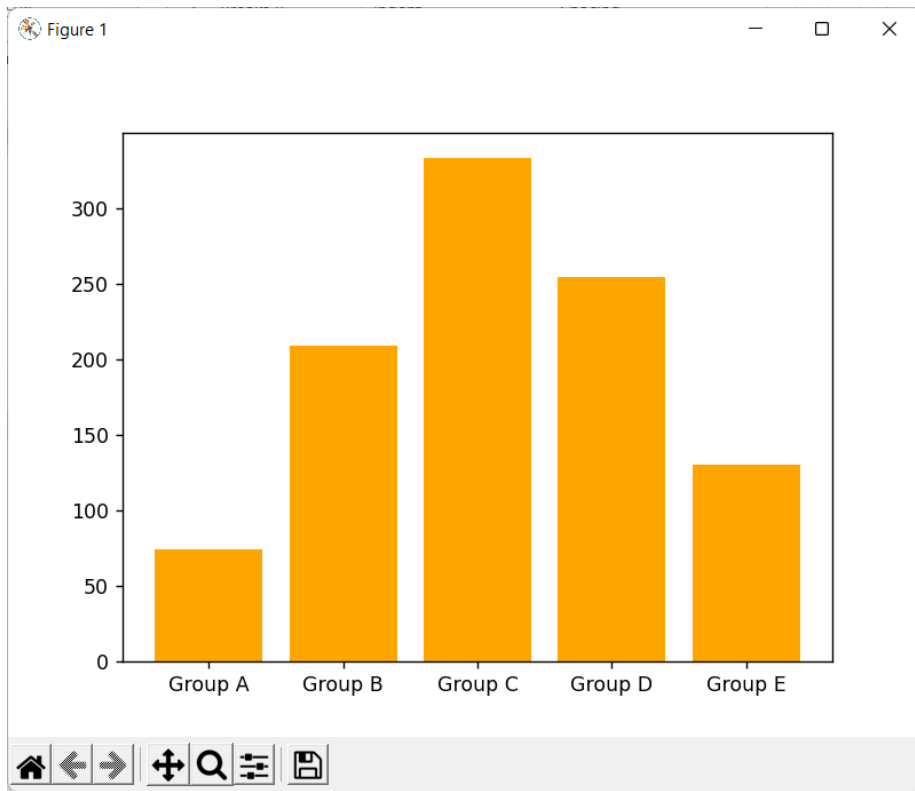
```
plt.pie(gender_rate, explode=explode, labels=chart_labels)
plt.show()
```
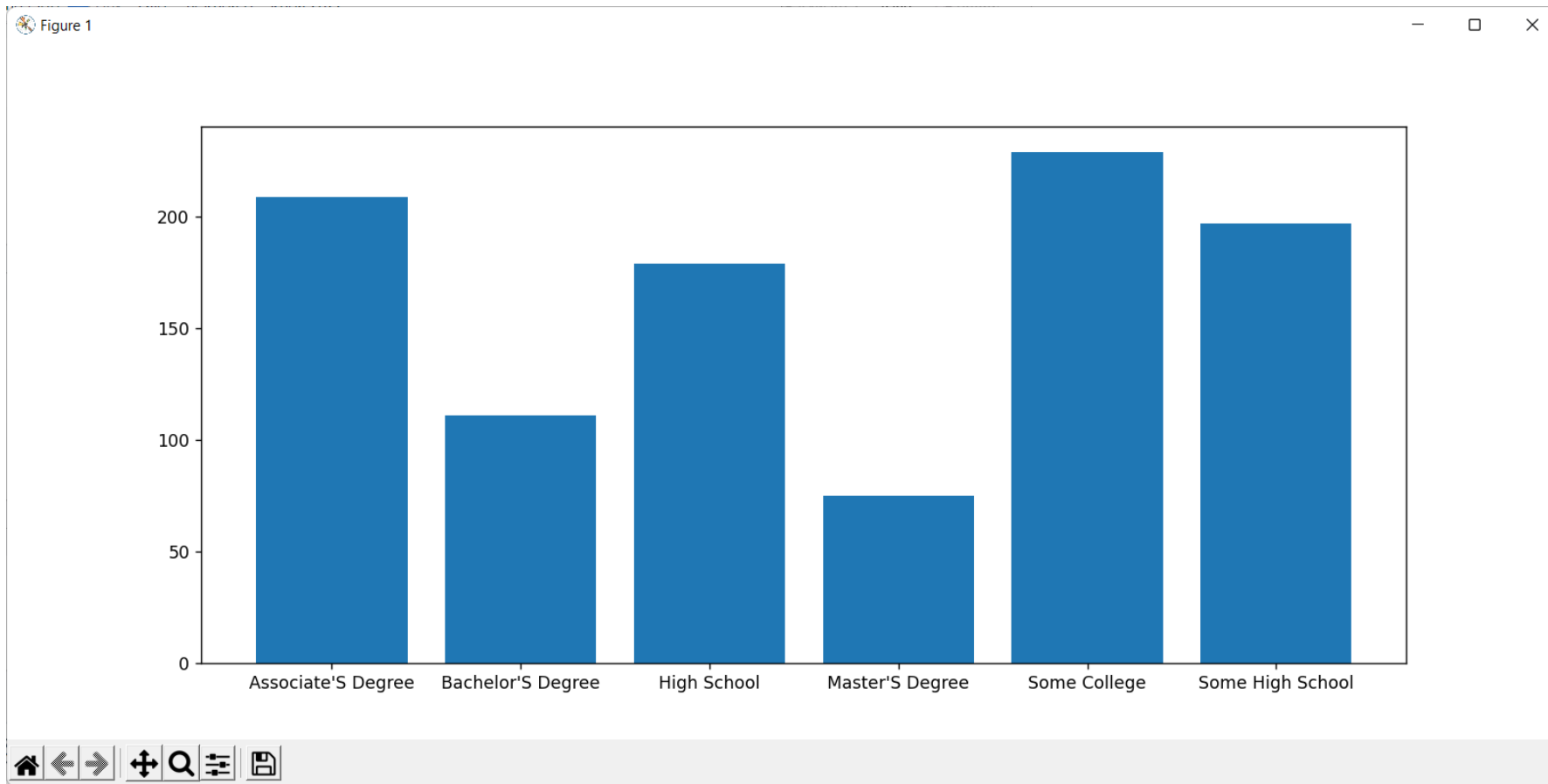
Result:


```
# 4. Count exam-takers by Group
group_rate = df.groupby(['Group'])['Group'].count()
chart_labels = group_rate.index.str.title()
plt.bar(chart_labels, group_rate, color="orange")
plt.show()
```

Result:

```
# 5. Count exam-takers by Edu
edu_rate = df.groupby(['Edu'])['Edu'].count()
chart_labels = edu_rate.index.str.title()
plt.bar(chart_labels, edu_rate)
plt.show()

Result:
```
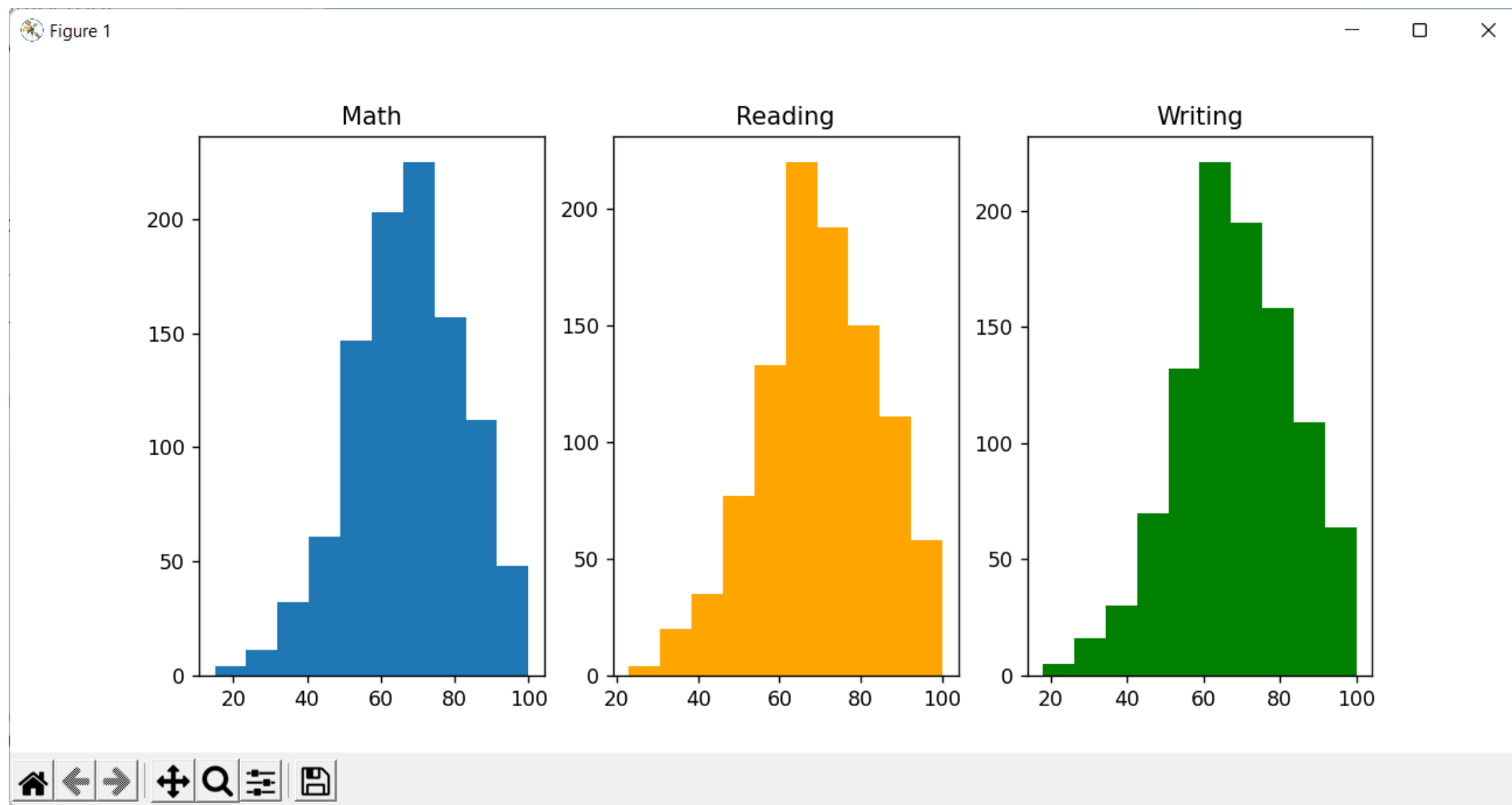
```python
# 6. Histogram of Math, Reading and Writing
plt.subplot(1, 3, 1)
plt.hist(df['Math'])
plt.title('Math')

plt.subplot(1, 3, 2)
plt.hist(df['Reading'], color="orange")
plt.title('Reading')
```

```
plt.subplot(1, 3, 3)
plt.hist(df['Writing'], color="green")
plt.title('Writing')
plt.show()
```

Result:

```python
# 7. Get top 10 of total scores
new_df = df.copy()
new_df['Total'] = new_df['Math'] + new_df['Reading'] + new_df['Writing']
sorted_df = new_df.sort_values(by=['Total'], ascending=False)
top_10 = sorted_df.head(10)
print(top_10)
```

```
Result:
     Gender    Group                  Edu  ... Reading Writing  Total
64     male  group E         some college  ...     100     100    300
101  female  group E         some college  ...     100     100    300
349    male  group E  associate's degree   ...     100     100    300
712  female  group D          high school  ...     100     100    299
142  female  group E     some high school  ...     100     100    298
932    male  group A         some college  ...     100      98    296
735  female  group E         some college  ...     100     100    296
740  female  group E  associate's degree   ...      95      99    294
281    male  group D         some college  ...     100     100    293
931  female  group C      master's degree  ...     100     100    292

[10 rows x 9 columns]
```