Exercise 5 – There are 2 Parts, **please submit both**

**PART 1: General questions**

1. What is a software life cycle?
2. Describe Software Engineering and the reasons why it is relevant.
3. Define Machine Learning and provide examples of its application.

**PART 2: Coding**

SPECIFICATION

Description: This program uses the marketingdata.txt data to print totals and create a bar and pie charts showing the number of men and women at the mall, by social class. See the description of the survey questions. Income is question 1 and Gender is question 2 in the survey.

Input:

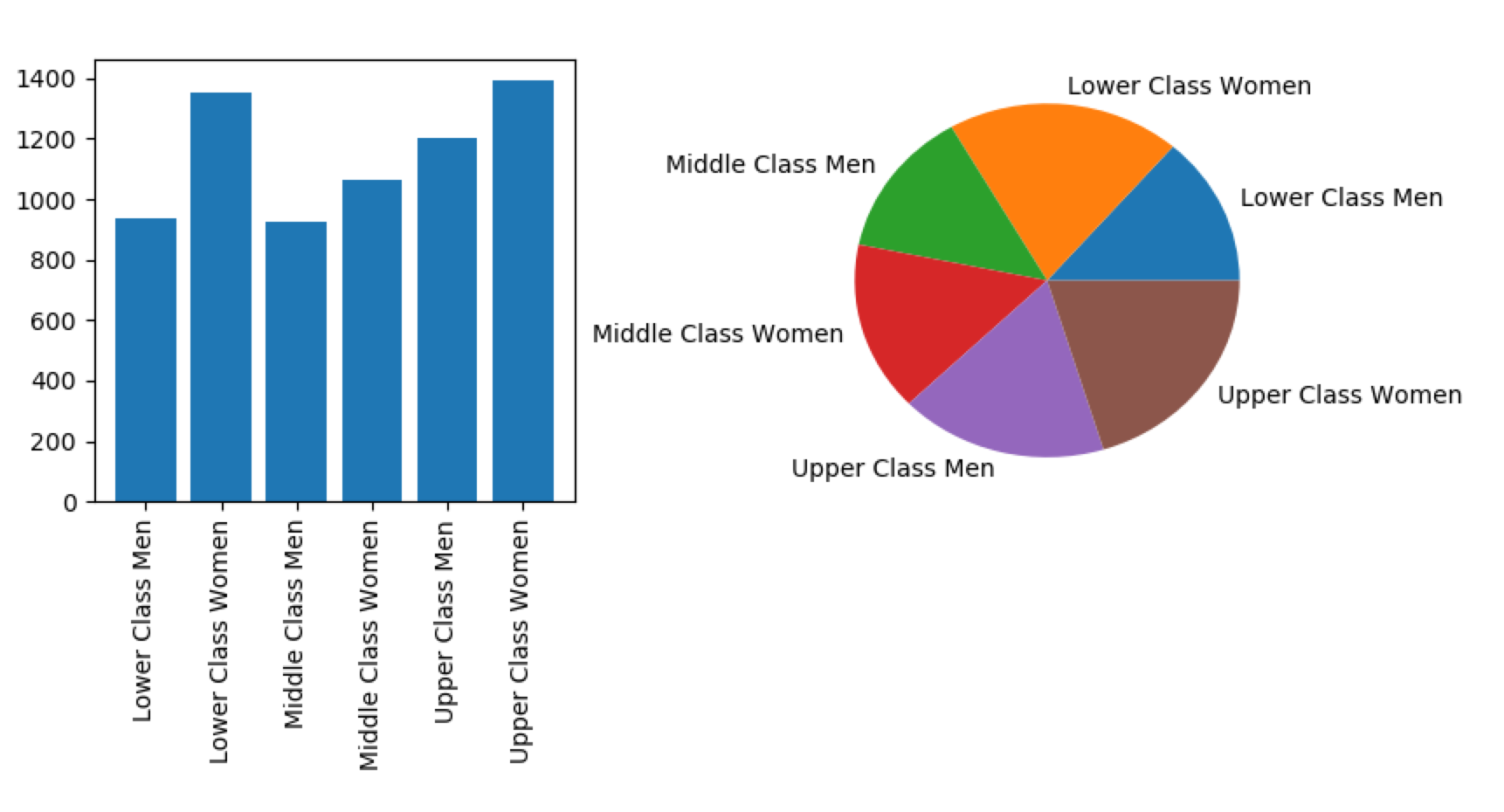
marketingdata.txt

Output:

Count of people at the mall, by income and gender (see details in procedure)

Bar chart (see below)

Pie chart (sample below – **NOT** the actual graph)



Procedure notes:

You will need to import matplotlib.pyplot. The preferred way to do this is: import matplotlib.pyplot as plt

The program will use a function (to be created as part of this exercise) get\_index(parts) to process each relevant line of marketing data and return an index into a count list.

Procedure:

1. import matplotlib

**Function**

1. Define a function named get\_index(parts) that accepts a parts list created from a line of the marketing data file and returns an index that is used to update the counts list.
   1. parts[1] contains the gender: ‘1’ = male, ‘2’ = female
   2. parts[0] contains the annual household income: ‘1’, ‘2’ or ‘3’ = Lower income, ‘4’, ‘5’ or ‘6’ = Middle income, anything else = Upper income.
   3. use parts[0] and parts[1] to determine the index to return

**Main Program**

1. Create a list of counts with six elements, initialized at zeros. Counts will contain the total number of “Lower Income Men” in count[0], for example.
2. Open the marketingdata.txt file for reading. Since this is a program specific to this file, there is no reason to use input to request the file name.
3. Read each line in the file. **Skip any line with ‘NA’** in any of the fields. For the other lines:

* create the list ***parts*** to use as input to the function: parts = line.strip().split()
* call the function to get the appropriate index for the ***counts*** list: index = get\_index(parts)
* update the appropriate count.

1. Now that you have collected the numbers for the six gender/social groups, you should print the results, with your totals in the *nnn*:

Who is at the mall?

Lower income males: *nnn* Lower income females: *nnn* Total lower income: *nnn*

Middle income males: *nnn* Middle income females: *nnn* Total middle income: *nnn*

Upper income males: *nnn* Upper income females: *nnn* Total upper income: *nnn*

1. Create the charts, as subplots. Labels for both the charts should be a list:

names = ['group 1','group 2', …]

* + Subplot 221 (left): bar chart
* bar plot
* x = [0,1,2,3,4,5]
* y = count list for each one of the 6 groups. Name y axis as “Count”
* do not overlap labels
  + Subplot 222 (right): pie chart. This **is not** as in the sample above and it is representing the cumulative distribution of lower/middle/upper income (meaning, for each one it would be females + males)
* pie plot
* label the wedges
* Add percentage values as second label to your pie chart
  + Insert a proper title to the main plot

Submit your program .py file via Canvas