# Lab 1

Q1: In the dataset "data.csv"

# CODE:

```
import pandas as pd

df = pd.read_csv("Data.csv")

print(df)
```

```
In [6]: runfile('C:/Users/KIIT/Desktop/Assignments/TNT/Lab1/q1.py', wdir='C:/Users/KIIT/
Desktop/Assignments/TNT/Lab1')
   Country Age Salary Purchased
   France 44.0 72000.0
    Spain 27.0 48000.0
                                Yes
  Germany 30.0
                                 No
    Spain 38.0 61000.0
                                 No
  Germany 40.0
France 35.0
                  70000.0
                                Yes
                  58000.0
                                Yes
    Spain
           NaN 52000.0
                                 No
   France 48.0 79000.0
                                Yes
  Germany 50.0 83000.0
NaN 37.0 67000.0
                                 No
                                Yes
```

**Q2:** Add a new column: Salary\_class A for loop is implemented and the observations are separated into three categories: o Salary

- greater than 70000 class0
- between 61000-70000 -class1
- between 48000-60000 -class2
- The classes have been stored in a new column 'Salary Class'

# CODE:

```
import pandas as pd
      df = pd.read csv("Data.csv")
      sal class = []
    for i in range(10):
          sal = df['Salary'][i]
10
          if sal>70000:
11
12
              sal class.append('class0')
          elif sal>=61000:
13
14
              sal class.append('class1')
          elif sal >= 48000:
15
              sal class.append('class2')
17
          else:
18
              sal class.append('')
19
      df['Salary class'] = sal class
20
21
      print(df)
22
```

```
In [7]: runfile('C:/Users/KIIT/Desktop/Assignments/TNT/Lab1/untitled0.py', wdir='C:/Users/
KIIT/Desktop/Assignments/TNT/Lab1')
                  Salary Purchased Salary_class
   Country
            Age
           44.0
                 72000.0
                               No
                                        class0
   France
    Spain 27.0
                48000.0
                               Yes
                                        class2
  Germany 30.0
                     NaN
                               No
    Spain 38.0 61000.0
                               No
                                        class1
  Germany 40.0 70000.0
4
                               Yes
                                        class1
                 58000.0
   France 35.0
                               Yes
                                        class2
    Spain
            NaN
                 52000.0
                               No
                                        class2
   France 48.0
                 79000.0
                               Yes
                                        class0
  Germany 50.0 83000.0
                               No
                                        class0
      NaN 37.0 67000.0
                               Yes
                                        class1
```

# **Q3:** Implement above using both for and while loop

# CODE:

```
import pandas as pd
      df = pd.read_csv("Data.csv")
      sal class = []
      i = 0

▼ while i < 10:
</p>
          sal = df['Salary'][i]
11
          if sal>70000:
12
13
              sal_class.append('class0')
          elif sal>=61000:
15
              sal_class.append('class1')
          elif sal >= 48000:
17
              sal_class.append('class2')
          else:
              sal class.append('')
          i += 1
21
22
      df['Salary class'] = sal class
      print(df)
25
```

```
In [8]: runfile('C:/Users/KIIT/Desktop/Assignments/TNT/Lab1/untitled1.py', wdir='C:/Users/
KIIT/Desktop/Assignments/TNT/Lab1')
              Age Salary Purchased Salary_class
   Country
    France 44.0
                    72000.0
                                      No
                                                 class0
1 Spain 27.0
2 Germany 30.0
3 Spain 38.0
4 Germany 40.0
      Spain 27.0 48000.0
                                     Yes
                                                 class2
   Germany 30.0 NaN
Spain 38.0 61000.0
                                      No
                                      No
                                                 class1
                     70000.0
                                                 class1
                                      Yes
    France 35.0 58000.0
                                                 class2
                                      Yes
     Spain NaN 52000.0
                                      No
                                                 class2
7 France 48.0 79000.0
8 Germany 50.0 83000.0
9 NaN 37.0 67000.0
                                     Yes
                                                 class0
                                                 class0
                                      No
                                                 class1
                                     Yes
```

**Q4:** Count the number of each class (class 0, class1,class2) in your dataset.

# CODE:

```
import pandas as pd
      df = pd.read csv("Data.csv")
      sal class = []
      i = 0

▼ while i < 10:
</p>
          sal = df['Salary'][i]
11
12
          if sal>70000:
13
               sal class.append('class0')
14
          elif sal>=61000:
15
              sal class.append('class1')
          elif sal >= 48000:
17
              sal class.append('class2')
          else:
19
              sal class.append('')
21
          i += 1
22
23
      df['Salary class'] = sal class
25
      c0 = len(df[df['Salary_class'] == 'class0'])
      c1 = len(df[df['Salary_class'] == 'class1'])
27
      c2 = len(df[df['Salary class'] == 'class2'])
29
      print(f'class0 = \{c0\}, class1 = \{c1\}, class2 = \{c2\}')
30
```

```
In [9]: runfile('C:/Users/KIIT/Desktop/Assignments/TNT/Lab1/untitled2.py', wdir='C:/Users/
KIIT/Desktop/Assignments/TNT/Lab1')
class0 = 3, class1 = 3, class2 = 3
```

**Q5:** Insert a new column Age\_Converted:

Use function c\_convert to add in the new column the converted values from column "Age":

dataset["Age\_Converted"]=dataset["Age"]\*12

# CODE:

```
import pandas as pd
      df = pd.read_csv("Data.csv")
      sal_class = []
      i = 0
    ▼ while i < 10:</p>
          sal = df['Salary'][i]
          if sal>70000:
              sal_class.append('class0')
          elif sal>=61000:
              sal_class.append('class1')
          elif sal >= 48000:
              sal class.append('class2')
              sal class.append('')
          i += 1
      df['Salary_class'] = sal_class
      age_con = df['Age']*12
      df['Age Converted'] = age con
      print(df)
29
```

```
In [10]: runfile('C:/Users/KIIT/Desktop/Assignments/TNT/Lab1/untitled3.py', wdir='C:/Users/
KIIT/Desktop/Assignments/TNT/Lab1')
                  Salary Purchased Salary_class Age_Converted
   Country
             Age
0
    France
           44.0
                  72000.0
                                No
                                          class0
                                                          528.0
1
     Spain 27.0
                  48000.0
                                Yes
                                          class2
                                                          324.0
                                                          360.0
   Germany
            30.0
                      NaN
                                No
                 61000.0
                                          class1
                                                          456.0
     Spain
           38.0
                                No
4 Germany
                                                          480.0
           40.0
                  70000.0
                                Yes
                                          class1
                                                          420.0
    France
           35.0
                 58000.0
                                Yes
                                          class2
           NaN 52000.0
                                          class2
                                                            NaN
     Spain
                                No
    France 48.0
                 79000.0
                                Yes
                                          class0
                                                          576.0
  Germany 50.0 83000.0
                                                          600.0
                                No
                                          class0
       NaN 37.0 67000.0
                                                          444.0
                                Yes
                                          class1
```