T&T Lab 1

12/01/2022

Himanshu Lohani
1905393

Code:

```
import pandas as pd

dataset=pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");

print (dataset)
```

Code Snippet:

```
import pandas as pd
dataset =pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");
print (dataset)
```

Output Snippet:

```
((IPdb [1])): runfile('/Users/himanshulohani/.spyder-py3/temp.py', wdir='/Users/himanshulohani/.spyder-py3')
   Country   Age   Salary Purchased
0   France  44.0  72000.0        No
1    Spain  27.0  48000.0       Yes
2  Germany  30.0      NaN        No
3    Spain  38.0  61000.0        No
4  Germany  40.0  70000.0       Yes
5   France  35.0  58000.0       Yes
6    Spain   NaN  52000.0        No
7   France  48.0  79000.0       Yes
8  Germany  50.0  83000.0        No
9      NaN  37.0  67000.0       Yes
```

Tasks

1. In the dataset "data.csv", in google classroom:
   i) 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 -classl
- between 48000-60000 -class2
- The classes have been stored in a new column 'Salary Class'

Code:

```
import pandas as pd

dataset=pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&
T_Lab_PPTs/lab1/Data.csv.xlsx");



sal_class =[]

for i in range(10):

    sal = dataset['Salary'][i]



    if sal>70000:

        sal_class.append('class0')

    elif sal>=61000:

        sal_class.append('class1')

    elif sal>=48000:

        sal_class.append('class2')

    else:

        sal_class.append('')

dataset['Salary_class']=sal_class

print (dataset)
```

Code Snippet:

```
import pandas as pd
dataset =pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");

sal_class =[]
for i in range(10):
    sal = dataset['Salary'][i]

    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')
dataset['Salary_class']=sal_class

print (dataset)
```

Output Snippet:

```
(((IPdb [5]))): runfile('/Users/himanshulohani/.spyder-py3/t
himanshulohani/.spyder-py3')
   Country   Age   Salary Purchased Salary_class
0   France  44.0  72000.0        No       class0
1    Spain  27.0  48000.0       Yes       class2
2  Germany  30.0      NaN        No
3    Spain  38.0  61000.0        No       class1
4  Germany  40.0  70000.0       Yes       class1
5   France  35.0  58000.0       Yes       class2
6    Spain   NaN  52000.0        No       class2
7   France  48.0  79000.0       Yes       class0
8  Germany  50.0  83000.0        No       class0
9      NaN  37.0  67000.0       Yes       class1
```

ii) lmplement above using both for and while loop

Code:

```
i=0

while i<10:


    sal = dataset['Salary'][i]
```

```
    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

i += 1
```

Code Snippet:

```
import pandas as pd
dataset =pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");

sal_class =[]

i=0
while i<10:

    sal = dataset['Salary'][i]

    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

    i += 1

dataset['Salary_class']=sal_class

print (dataset)
```

Output Snippet:

```
(((IPdb [6]))): runfile('/Users/himanshulohani/.spyder-py3/temp.py',
himanshulohani/.spyder-py3')
   Country   Age   Salary Purchased Salary_class
0   France  44.0  72000.0        No       class0
1    Spain  27.0  48000.0       Yes       class2
2  Germany  30.0      NaN        No
3    Spain  38.0  61000.0        No       class1
4  Germany  40.0  70000.0       Yes       class1
5   France  35.0  58000.0       Yes       class2
6    Spain   NaN  52000.0        No       class2
7   France  48.0  79000.0       Yes       class0
8  Germany  50.0  83000.0        No       class0
9      NaN  37.0  67000.0       Yes       class1
```

iii) Count the number of each class (class 0, classl,class2) in your dataset.

Code:

```
import pandas as pd

dataset=pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");


sal_class =[]


i=0

while i<10:


    sal = dataset['Salary'][i]
```

```
    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

    i += 1

dataset['Salary_class']=sal_class

c0 = len(dataset[dataset['Salary_class'] =='class0'])
c1 = len(dataset[dataset['Salary_class'] =='class1'])
c2 = len(dataset[dataset['Salary_class'] =='class2'])

print(f'class0 ={c0},class1 ={c1},class2={c2}')
```

Code Snippet:

```
import pandas as pd
dataset =pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");

sal_class =[]

i=0
while i<10:

    sal = dataset['Salary'][i]

    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

    i += 1

dataset['Salary_class']=sal_class

c0 = len(dataset[dataset['Salary_class'] =='class0'])
c1 = len(dataset[dataset['Salary_class'] =='class1'])
c2 = len(dataset[dataset['Salary_class'] =='class2'])

print(f'class0 ={c0},class1 ={c1},class2={c2}')
```

Output Snippet:

```
(((IPdb [8]))): runfile('/Users/himanshulohani/.spyder-py
wdir='/Users/himanshulohani/.spyder-py3')
class0 =3,class1 =3,class2=3
```

iv) Insert a new column Age_Converted:

Use function c_convert to add in the new column the converted values fromcolumn "Age" : dataset["Age_Converted"]=dataset["Age"]*12

Code:

```
import pandas as pd

dataset=pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");
```

```
sal_class =[]

i=0
while i<10:

    sal = dataset['Salary'][i]

    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

    i += 1

dataset['Salary_class']=sal_class

age_con = dataset['Age']*12
```

dataset['Age_Converted']= age_con

print(dataset)

Code Snippet:

```
import pandas as pd
dataset =pd.read_excel("/Users/himanshulohani/Desktop/KIIT/T&T_Lab_PPTs/lab1/Data.csv.xlsx");

sal_class =[]

i=0
while i<10:

    sal = dataset['Salary'][i]

    if sal>70000:
        sal_class.append('class0')
    elif sal>=61000:
        sal_class.append('class1')
    elif sal>=48000:
        sal_class.append('class2')
    else:
        sal_class.append('')

    i += 1

dataset['Salary_class']=sal_class

age_con = dataset['Age']*12

dataset['Age_Converted']= age_con

print(dataset)
```

Output Snippet:

```
(((IPdb [11]))): runfile('/Users/himanshulohani/.spyder-py3/
temp.py', wdir='/Users/himanshulohani/.spyder-py3')
   Country   Age   Salary Purchased Salary_class  Age_Converted
0   France  44.0  72000.0        No       class0          528.0
1    Spain  27.0  48000.0       Yes       class2          324.0
2  Germany  30.0      NaN        No                       360.0
3    Spain  38.0  61000.0        No       class1          456.0
4  Germany  40.0  70000.0       Yes       class1          480.0
5   France  35.0  58000.0       Yes       class2          420.0
6    Spain   NaN  52000.0        No       class2            NaN
7   France  48.0  79000.0       Yes       class0          576.0
8  Germany  50.0  83000.0        No       class0          600.0
9      NaN  37.0  67000.0       Yes       class1          444.0
```