Lab Assignment 2

Chaudhary Hamdan

1905387

Date: 19-01-2022

Question:

1. Create a copy of the ‘data.csv’ and name the dataframe as dataset1.

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

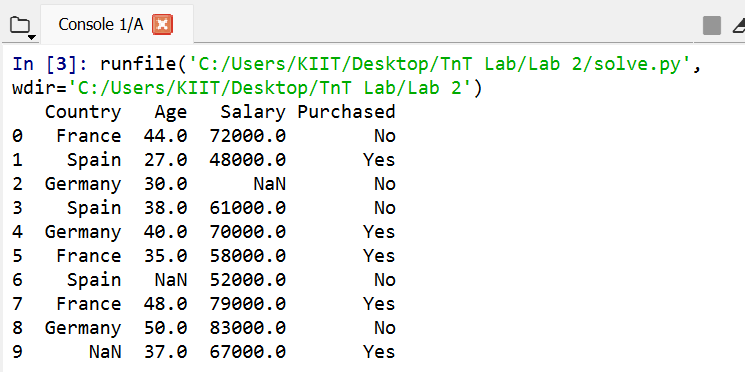
import pandas as pd

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

print(dataset1)

Output:



2. To display the count of each value in the county column.

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

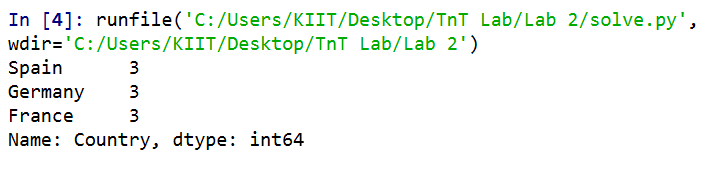
import pandas as pd

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

print(dataset1['Country'].value\_counts())

Output:



3. To display how many individuals from each country are buying the

product and how many aren’t.

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

print(pd.crosstab(

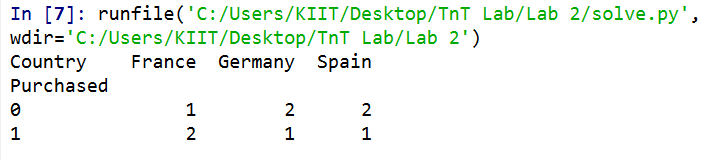
index=dataset1['Purchased'],

columns = dataset1['Country'],

dropna=True

))

Output:



4. Show all probabilities of occurance:

1. Joint

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

print(pd.crosstab(

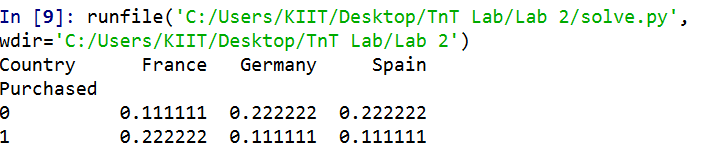
index=dataset1['Purchased'],

columns = dataset1['Country'],

dropna=True,

normalize=True

))

Output:   


1. Marginal

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

print(pd.crosstab(

index=dataset1['Purchased'],

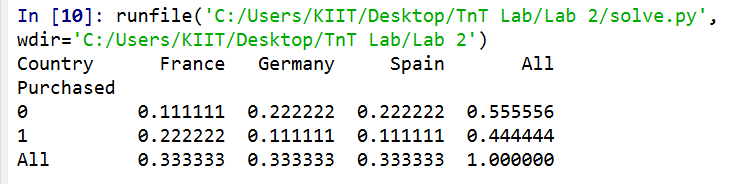
columns = dataset1['Country'],

dropna=True,

normalize=True,

margins=True

))

Output:   


1. Conditional:

A) Country is known, whether the individual will purchase the

product or not

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

print(pd.crosstab(

index=dataset1['Purchased'],

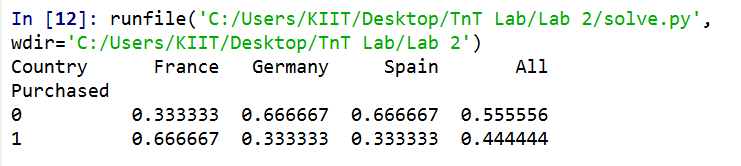
columns = dataset1['Country'],

dropna=True,

normalize='columns',

margins=True

))

Output:   


B) Product has been brought/not brought, find the probability the

individual belongs to which country

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

print(pd.crosstab(

index=dataset1['Purchased'],

columns = dataset1['Country'],

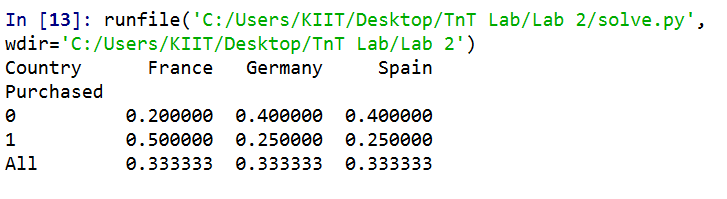
dropna=True,

normalize='index',

margins=True

))

Output:



5. Find out whether there is a correlation between numerical

data(variables) in the dataset.

Code:

import pandas as pd

import numpy as np

df = pd.read\_csv('Data.csv')

dataset1 = df.copy(deep=True)

dataset1['Purchased'].replace('No', 0, inplace=True)

dataset1['Purchased'].replace('Yes', 1, inplace=True)

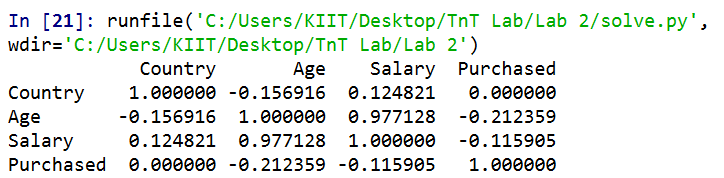
coun = np.array(dataset1['Country'].unique())

for n, i in enumerate(coun):

dataset1['Country'].replace(i, n, inplace=True)

print(dataset1.corr())

Output:



1. Use scatter plot and plot the data given in ‘social\_network\_ad.csv’.

Keep ‘age’ in x-axis and ‘estimated salary’ in y-

axis.

Code:

"""

Created on Wed Jan 19 12:40:43 2022

@author: Chaudhary Hamdan

"""

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_csv('Social\_Network\_Ads.csv')

print(df)

plt.scatter(

df['Age'],

df['EstimatedSalary']

)

plt.title('Graph Ques 6')

plt.xlabel('Age')

plt.ylabel('Estimated Salary')

plt.show()

Output:

