

```
##Heart_Failure
import pandas as pd
import numpy as np
import csv
import matplotlib.pyplot as plt
```

```
print('Question 1:')
print('Q1(1):')
```

```
data = pd.read_csv('heart_failure_clinical_records_dataset.csv')
data1 =
data[['creatinine_phosphokinase', 'serum_creatinine', 'serum_sodium', 'pla
telets', 'DEATH_EVENT']]
```

```
df_0 =
data[['creatinine_phosphokinase', 'serum_creatinine', 'serum_sodium', 'pla
telets', 'DEATH_EVENT']].loc[data['DEATH_EVENT'] == 0 ]
df_1 =
data[['creatinine_phosphokinase', 'serum_creatinine', 'serum_sodium', 'pla
telets', 'DEATH_EVENT']].loc[data['DEATH_EVENT'] == 1
]
```

```
print('=====
=====')
```

```
print('Q1(2):')
```

```
plt.matshow(df_0.corr())
plt.show()
```

```
plt.matshow(df_1.corr())
plt.show()
```

```
print('Corr of surviving:')
print(df_0.corr())
print('=====')
print('Corr of deceased:')
print(df_1.corr())
```

```
print('=====
=====')
```

```

print('\n')

##Bakery
import pandas as pd
import math
pd.set_option("display.max.rows",None)
pd.set_option("max_colwidth",None)
df1 = pd.read_csv (r'BreadBasket_DMS.csv')
df2 = pd.read_csv (r'BreadBasket_DMS_output.csv')
print("Q1")
print("The busiest hour
is:",df2.groupby("Hour").sum()["Transaction"].idxmax())
print()
print("The busiest weekday
is:",df2.groupby("Weekday").sum()["Transaction"].idxmax())
print()
print("The busiest period
is:",df2.groupby("Period").sum()["Transaction"].idxmax())
print("#####")
print("Q2")
df_temp = df2[['Hour','Transaction','Item_Price']]
df_temp = df_temp.copy()
df_temp['Profit'] = df_temp['Transaction'].mul(df_temp['Item_Price'])
df_Profit_hr = df_temp [['Hour','Profit']]
print()
print("The most profit hour
is :",df_Profit_hr.groupby('Hour').sum().stack().idxmax(),df_Profit_hr.g
rouphy('Hour').sum().stack().max())

df_temp = df2[['Weekday','Transaction','Item_Price']]
df_temp = df_temp.copy()
df_temp['Profit'] = df_temp['Transaction'].mul(df_temp['Item_Price'])
df_Profit_hr = df_temp [['Weekday','Profit']]
print()
print("The most profit weekday
is :",df_Profit_hr.groupby('Weekday').sum().stack().idxmax(),df_Profit_h
r.groupby('Weekday').sum().stack().max())

```

```

df_temp = df2[['Period','Transaction','Item_Price']]
df_temp = df_temp.copy()
df_temp['Profit'] = df_temp['Transaction'].mul(df_temp['Item_Price'])
df_Profit_hr = df_temp [['Period','Profit']]
print()

```

```

print("The most profit period
is :",df_Profit_hr.groupby('Period').sum().stack().idxmax(),df_Profit_hr.
groupby('Period').sum().stack().max())
print("#####")
print("Q3")
print("The most popular item is ")
print(df2.groupby("Item").sum()["Transaction"].idxmax())
print("The least popular item is ")
print(df2.groupby("Item").sum()["Transaction"].idxmin())
print("#####")
print("Q4")
df_people =
df2.groupby(["Month","Weekday"]).sum()["Transaction"].max()/50
print("Number of barrista needed:")
print(math.ceil(df_people))
print("#####")
print("Q5")
df2['label'] = 'Unknown'
df2.loc[df2["Item"]=="Tea","label"] = 'Drinks'
df2.loc[df2["Item"]=="Soup","label"] = 'Drinks'
df2.loc[df2["Item"]=="Smoothies","label"] = 'Drinks'
df2.loc[df2["Item"]=="My-5 Fruit Shoot","label"] = 'Drinks'
df2.loc[df2["Item"]=="Mineral water","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hot chocolate","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hearty & Seasonal","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee granules ","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coke","label"] = 'Drinks'
df2.loc[df2["Item"]=="Juice","label"] = 'Drinks'
#####
#####
df2.loc[df2["Item"]=="Victorian Sponge","label"] = 'Foods'
df2.loc[df2["Item"]=="Vegan mincepie","label"] = 'Foods'
df2.loc[df2["Item"]=="Vegan Feast","label"] = 'Foods'
df2.loc[df2["Item"]=="Truffles","label"] = 'Foods'
df2.loc[df2["Item"]=="Tartine","label"] = 'Foods'
df2.loc[df2["Item"]=="Tacos/Fajita","label"] = 'Foods'
df2.loc[df2["Item"]=="Spanish Brunch","label"] = 'Foods'
df2.loc[df2["Item"]=="Scone","label"] = 'Foods'
df2.loc[df2["Item"]=="Sandwich","label"] = 'Foods'
df2.loc[df2["Item"]=="Salad","label"] = 'Foods'
df2.loc[df2["Item"]=="Raspberry shortbread sandwich","label"] = 'Foods'
df2.loc[df2["Item"]=="Polenta","label"] = 'Foods'
df2.loc[df2["Item"]=="Pintxos","label"] = 'Foods'

```

```

df2.loc[df2["Item"]=="Pastry","label"] = 'Foods'
df2.loc[df2["Item"]=="Panaotone","label"] = 'Foods'
df2.loc[df2["Item"]=="Olum & polenta","label"] = 'Foods'
df2.loc[df2["Item"]=="Muffin","label"] = 'Foods'
df2.loc[df2["Item"]=="Muesli","label"] = 'Foods'
df2.loc[df2["Item"]=="Mighty Protein","label"] = 'Foods'
df2.loc[df2["Item"]=="Medialuna","label"] = 'Foods'
df2.loc[df2["Item"]=="Lemon and coconut","label"] = 'Foods'
df2.loc[df2["Item"]=="Kids biscuit","label"] = 'Foods'
df2.loc[df2["Item"]=="Jammie Dodgers","label"] = 'Foods'
df2.loc[df2["Item"]=="Jam","label"] = 'Foods'
df2.loc[df2["Item"]=="Honey","label"] = 'Foods'
df2.loc[df2["Item"]=="Granola","label"] = 'Foods'
df2.loc[df2["Item"]=="Gingerbread syrup","label"] = 'Foods'
df2.loc[df2["Item"]=="Fudge","label"] = 'Foods'
df2.loc[df2["Item"]=="Frittata","label"] = 'Foods'
df2.loc[df2["Item"]=="Focaccia","label"] = 'Foods'
df2.loc[df2["Item"]=="Extra Salami or Feta","label"] = 'Foods'
df2.loc[df2["Item"]=="Empanadas","label"] = 'Foods'
df2.loc[df2["Item"]=="Eggs","label"] = 'Foods'
df2.loc[df2["Item"]=="Dulce de Leche","label"] = 'Foods'
df2.loc[df2["Item"]=="Duck egg","label"] = 'Foods'
df2.loc[df2["Item"]=="Crisps","label"] = 'Foods'
df2.loc[df2["Item"]=="Crepes","label"] = 'Foods'
df2.loc[df2["Item"]=="Cookies","label"] = 'Foods'
df2.loc[df2["Item"]=="Chocolates","label"] = 'Foods'
df2.loc[df2["Item"]=="Chimichurri Oil","label"] = 'Foods'
df2.loc[df2["Item"]=="Chicken Stew","label"] = 'Foods'
df2.loc[df2["Item"]=="Chicken sand","label"] = 'Foods'
df2.loc[df2["Item"]=="Cherry me Dried fruit","label"] = 'Foods'
df2.loc[df2["Item"]=="Caramel bites","label"] = 'Foods'
df2.loc[df2["Item"]=="Cake","label"] = 'Foods'
df2.loc[df2["Item"]=="Brownie","label"] = 'Foods'
df2.loc[df2["Item"]=="Brioche and salami","label"] = 'Foods'
df2.loc[df2["Item"]=="Bread Pudding","label"] = 'Foods'
df2.loc[df2["Item"]=="Bread","label"] = 'Foods'
df2.loc[df2["Item"]=="Baguette","label"] = 'Foods'
df2.loc[df2["Item"]=="Alfajores","label"] = 'Foods'

```

```

Drinks_sum=sum(df2.loc[df2['label']=='Drinks']['Item_Price'])/len(df2.l
oc[df2['label']=='Drinks'])
Foods_sum=sum(df2.loc[df2['label']=='Foods']['Item_Price'])/len(df2.loc
[df2['label']=='Foods'])
print("Drinks average price:",Drinks_sum)

```

```
print("Foods average price:",Foods_sum)
print("#####")
```

```
print('Q6')
df2['label'] = 'Unknown'
df2.loc[df2["Item"]=="Tea","label"] = 'Drinks'
df2.loc[df2["Item"]=="Soup","label"] = 'Drinks'
df2.loc[df2["Item"]=="Smoothies","label"] = 'Drinks'
df2.loc[df2["Item"]=="My-5 Fruit Shoot","label"] = 'Drinks'
df2.loc[df2["Item"]=="Mineral water","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hot chocolate","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hearty & Seasonal","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee granules ","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coke","label"] = 'Drinks'
df2.loc[df2["Item"]=="Juice","label"] = 'Drinks'
#####
#####
df2.loc[df2["Item"]=="Victorian Sponge","label"] = 'Foods'
df2.loc[df2["Item"]=="Vegan mincepie","label"] = 'Foods'
df2.loc[df2["Item"]=="Vegan Feast","label"] = 'Foods'
df2.loc[df2["Item"]=="Truffles","label"] = 'Foods'
df2.loc[df2["Item"]=="Tartine","label"] = 'Foods'
df2.loc[df2["Item"]=="Tacos/Fajita","label"] = 'Foods'
df2.loc[df2["Item"]=="Spanish Brunch","label"] = 'Foods'
df2.loc[df2["Item"]=="Scone","label"] = 'Foods'
df2.loc[df2["Item"]=="Sandwich","label"] = 'Foods'
df2.loc[df2["Item"]=="Salad","label"] = 'Foods'
df2.loc[df2["Item"]=="Raspberry shortbread sandwich","label"] = 'Foods'
df2.loc[df2["Item"]=="Polenta","label"] = 'Foods'
df2.loc[df2["Item"]=="Pintxos","label"] = 'Foods'
df2.loc[df2["Item"]=="Pastry","label"] = 'Foods'
df2.loc[df2["Item"]=="Panatone","label"] = 'Foods'
df2.loc[df2["Item"]=="Olum & polenta","label"] = 'Foods'
df2.loc[df2["Item"]=="Muffin","label"] = 'Foods'
df2.loc[df2["Item"]=="Muesli","label"] = 'Foods'
df2.loc[df2["Item"]=="Mighty Protein","label"] = 'Foods'
df2.loc[df2["Item"]=="Medialuna","label"] = 'Foods'
df2.loc[df2["Item"]=="Lemon and coconut","label"] = 'Foods'
df2.loc[df2["Item"]=="Kids biscuit","label"] = 'Foods'
df2.loc[df2["Item"]=="Jammie Dodgers","label"] = 'Foods'
df2.loc[df2["Item"]=="Jam","label"] = 'Foods'
df2.loc[df2["Item"]=="Honey","label"] = 'Foods'
df2.loc[df2["Item"]=="Granola","label"] = 'Foods'
```

```

df2.loc[df2["Item"]=="Gingerbread syrup","label"] = 'Foods'
df2.loc[df2["Item"]=="Fudge","label"] = 'Foods'
df2.loc[df2["Item"]=="Frittata","label"] = 'Foods'
df2.loc[df2["Item"]=="Focaccia","label"] = 'Foods'
df2.loc[df2["Item"]=="Extra Salami or Feta","label"] = 'Foods'
df2.loc[df2["Item"]=="Empanadas","label"] = 'Foods'
df2.loc[df2["Item"]=="Eggs","label"] = 'Foods'
df2.loc[df2["Item"]=="Dulce de Leche","label"] = 'Foods'
df2.loc[df2["Item"]=="Duck egg","label"] = 'Foods'
df2.loc[df2["Item"]=="Crisps","label"] = 'Foods'
df2.loc[df2["Item"]=="Crepes","label"] = 'Foods'
df2.loc[df2["Item"]=="Cookies","label"] = 'Foods'
df2.loc[df2["Item"]=="Chocolates","label"] = 'Foods'
df2.loc[df2["Item"]=="Chimichurri Oil","label"] = 'Foods'
df2.loc[df2["Item"]=="Chicken Stew","label"] = 'Foods'
df2.loc[df2["Item"]=="Chicken sand","label"] = 'Foods'
df2.loc[df2["Item"]=="Cherry me Dried fruit","label"] = 'Foods'
df2.loc[df2["Item"]=="Caramel bites","label"] = 'Foods'
df2.loc[df2["Item"]=="Cake","label"] = 'Foods'
df2.loc[df2["Item"]=="Brownie","label"] = 'Foods'
df2.loc[df2["Item"]=="Brioche and salami","label"] = 'Foods'
df2.loc[df2["Item"]=="Bread Pudding","label"] = 'Foods'
df2.loc[df2["Item"]=="Bread","label"] = 'Foods'
df2.loc[df2["Item"]=="Baguette","label"] = 'Foods'
df2.loc[df2["Item"]=="Alfajores","label"] = 'Foods'

```

```

Drinks_sum=sum(df2.loc[df2['label']=='Drinks']['Item_Price'])*len(df2.loc[df2['label']=='Drinks'])
Foods_sum=sum(df2.loc[df2['label']=='Foods']['Item_Price'])*len(df2.loc[df2['label']=='Foods'])
print("Drinks sum price:",Drinks_sum)
print("Foods sum price:",Foods_sum)
print("#####")

```

```

print("Q7&Q8")
df2['label'] = df2.loc[df2["Weekday"]=="Monday","label"]
df2.loc[df2["Weekday"]=="Monday","label"] = 'Monday'
Monday = df2.groupby(['label','Item']).sum()['Transaction']
print(Monday.nlargest(5))
print()
print(Monday.nsmallest(5))
print("#####")
print()
df2['label'] = df2.loc[df2["Weekday"]=="Tuesday","label"]

```

```

df2.loc[df2["Weekday"]=="Tuesday","label"] = 'Tuesday'
Tuesday = df2.groupby(['label','Item']).sum()['Transaction']
print(Tuesday.nlargest(5))
print()
print(Tuesday.nsmallest(5))
print("#####")
print()
df2['label'] = df2.loc[df2["Weekday"]=="Wednesday","label"]
df2.loc[df2["Weekday"]=="Wednesday","label"] = 'Wednesday'
Wednesday = df2.groupby(['label','Item']).sum()['Transaction']
print(Wednesday.nlargest(5))
print()
print(Wednesday.nsmallest(5))
print("#####")
print()
df2['label'] = df2.loc[df2["Weekday"]=="Thursday","label"]
df2.loc[df2["Weekday"]=="Thursday","label"] = 'Thursday'
Thursday = df2.groupby(['label','Item']).sum()['Transaction']
print(Thursday.nlargest(5))
print()
print(Thursday.nsmallest(5))
print("#####")
print()
df2['label'] = df2.loc[df2["Weekday"]=="Friday","label"]
df2.loc[df2["Weekday"]=="Friday","label"] = 'Friday'
Friday = df2.groupby(['label','Item']).sum()['Transaction']
print(Friday.nlargest(5))
print()
print(Friday.nsmallest(5))
print("#####")

```

```

print()
df2['label'] = df2.loc[df2["Weekday"]=="Saturday","label"]
df2.loc[df2["Weekday"]=="Saturday","label"] = 'Saturday'
Saturday = df2.groupby(['label','Item']).sum()['Transaction']
print(Saturday.nlargest(5))
print()
print(Saturday.nsmallest(5))
print("#####")

```

```

print()
df2['label'] = df2.loc[df2["Weekday"]=="Sunday","label"]
df2.loc[df2["Weekday"]=="Sunday","label"] = 'Sunday'
Sunday = df2.groupby(['label','Item']).sum()['Transaction']

```

```
print(Sunday.nlargest(5))
print()
print(Sunday.nsmallest(5))
print("#####")
```

```
print("Q9")
df2['label'] = 'Unknown'
df2.loc[df2["Item"]=="Tea","label"] = 'Drinks'
df2.loc[df2["Item"]=="Soup","label"] = 'Drinks'
df2.loc[df2["Item"]=="Smoothies","label"] = 'Drinks'
df2.loc[df2["Item"]=="My-5 Fruit Shoot","label"] = 'Drinks'
df2.loc[df2["Item"]=="Mineral water","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hot chocolate","label"] = 'Drinks'
df2.loc[df2["Item"]=="Hearty & Seasonal","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee granules ","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coffee","label"] = 'Drinks'
df2.loc[df2["Item"]=="Coke","label"] = 'Drinks'
df2.loc[df2["Item"]=="Juice","label"] = 'Drinks'
```

```
drinks =sum(df2.loc[df2['label']=='Drinks']['Transaction']) /
df2.sum()["Transaction"]
print(drinks.round(2))
print("#####")
```