

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
```

```
In [8]: #task 1.0
#Read the dataset (transactions.csv) as a Pandas dataframe.
#Note that the first row of the CSV contains the column names.
#Reading the data set
file = pd.read_csv('transactions.csv')
df = pd.DataFrame(file)
#df.head()
df.head(1)
```

```
Out[8]:
```

	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDes
0	1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	0.



```
In [13]: #task 1.1
#Return the column names as a List from the dataframe.
col_names = df.columns

col_names
```

```
Out[13]: Index(['step', 'type', 'amount', 'nameOrig', 'oldbalanceOrg', 'newbalanceOrig',
               'nameDest', 'oldbalanceDest', 'newbalanceDest', 'isFraud',
               'isFlaggedFraud'],
              dtype='object')
```

```
In [14]: #task1.2
#Return the first k rows from the dataframe.
k = 6 #Assumption
df.head(k)
```

```
Out[14]:
```

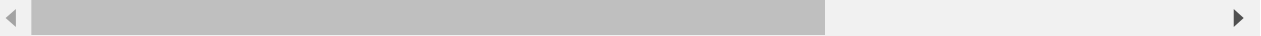
	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceD
0	1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	
1	1	PAYMENT	1864.28	C1666544295	21249.0	19384.72	M2044282225	
2	1	TRANSFER	181.00	C1305486145	181.0	0.00	C553264065	
3	1	CASH_OUT	181.00	C840083671	181.0	0.00	C38997010	2118
4	1	PAYMENT	11668.14	C2048537720	41554.0	29885.86	M1230701703	
5	1	PAYMENT	7817.71	C90045638	53860.0	46042.29	M573487274	



```
In [15]: #task1.3
#Return a random sample of k rows from the dataframe
df.sample(n=k)
```

Out[15]:

	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbal
162465	12	CASH_OUT	369666.62	C60334083	0.00	0.00	C976144053	19
127174	11	CASH_OUT	311374.26	C117519162	0.00	0.00	C111425078	6
123279	11	CASH_IN	121677.36	C1286519011	11303714.83	11425392.19	C292255799	3
170697	12	CASH_IN	93682.57	C1956185629	4617.00	98299.57	C147135576	1
84349	10	PAYMENT	38000.83	C1281732879	0.00	0.00	M95533648	
86786	10	PAYMENT	21606.74	C658110083	0.00	0.00	M262576834	



In [18]:

```
#task1.4
#Return a list of the unique transaction types.
uniq_transaction = df.type.unique()
uniq_transaction
```

Out[18]: array(['PAYMENT', 'TRANSFER', 'CASH_OUT', 'DEBIT', 'CASH_IN'],
dtype=object)

In [19]:

```
#Return a Pandas series of the top 10 transaction destinations with frequencies.
```

In [32]:

```
#task1.6
#Return all the rows from the dataframe for which fraud was detected.
fraud = df.isFraud[:]
kk= []
for i in range(6):
    if fraud[i] == 1:
        kk.append(pd.DataFrame(df.iloc[i]))
kk
```

```
Out[32]: [
      step      2
      type      TRANSFER
      amount    181.0
      nameOrig  C1305486145
      oldbalanceOrg  181.0
      newbalanceOrig  0.0
      nameDest    C553264065
      oldbalanceDest  0.0
      newbalanceDest  0.0
      isFraud      1
      isFlaggedFraud  0,
      step      3
      step      1
      type      CASH_OUT
      amount    181.0
      nameOrig  C840083671
      oldbalanceOrg  181.0
      newbalanceOrig  0.0
      nameDest    C38997010
      oldbalanceDest  21182.0
      newbalanceDest  0.0
      isFraud      1
      isFlaggedFraud  0]
```

In []: