Experiment-1

IMPLEMENT PROGRAM FOR TIME SERIES DATA CLEANING, LOADING AND HANDLING TIME SERIES DATA AND PREPROCESSING TECHNIQUES

AIM:
TO WRITE A TO IMPLEMENT PROGRAM FOR TIME SERIES DATA CLEANING, LOADING AND HANDLING TIME SERIES DATA AND PREPROCESSING TECHNIQUES

PROCEDURE:
1) Import necessary libraries.
2)Load the necessary libraries
3) view the data & get the information about the features
4)Remove the missing values and duplicate values.
5) split the date and time from date column and convert it to datetime type
6) find the different categories where money is spend
7)Use a bar plot to check category with highest expense

CODE:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

df = pd.read_csv("expense_data_1.csv")

df.head()

df.drop_duplicates()

df.isna().sum()

df.drop(["Subcategory", "Note.1"], axis=1, inplace=True)

df.head()

df.describe()

df.info()

df['time'] = df['Date'].str.split(" ").str[1]
df['Date'] = df['Date'].str.split(" ").str[0]

df.head()
```

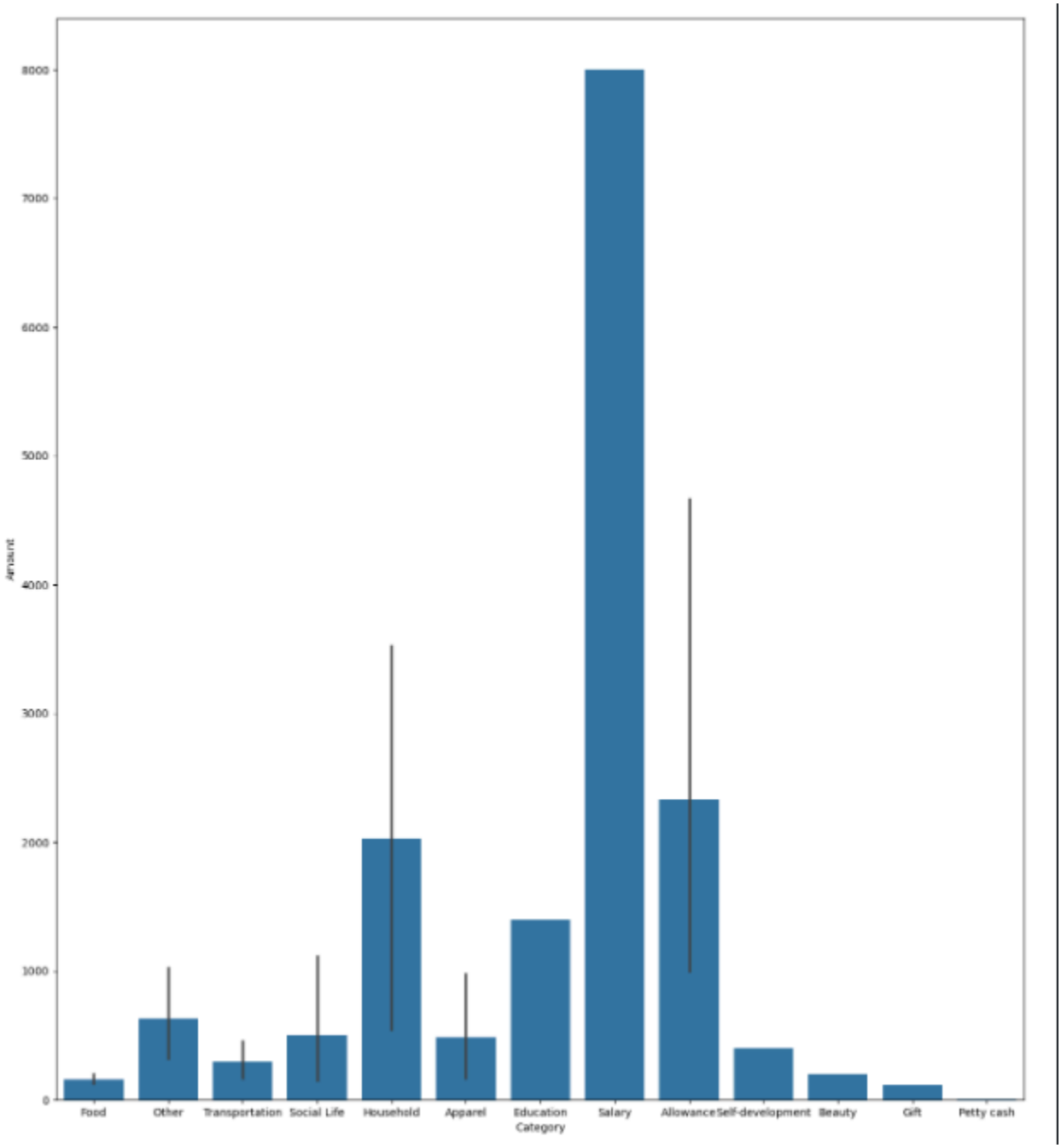
```
df['Currency'].unique()

df['Amount'] = df.apply(lambda row: row['Amount'] * 93 if row['Currency'] == 'USD' else
row['Amount'], axis=1)

df.drop(['Currency', 'Account.1'], axis=1, inplace=True)

df.head()

df['Category'].unique()

df['Account'].unique()

df['Account'] = df.apply(lambda row: 'online' if row['Account'] == 'CUB - online payment' else
row['Account'], axis=1)

df['Account'].unique()

plt.figure(figsize=(16,18))
sns.barplot(data=df, x='Category', y='Amount')

sns.countplot(df['Income/Expense'])

plt.hist(df['Amount'])
```

OUTPUT:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 277 entries, 0 to 276
Data columns (total 9 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Date            277 non-null    object
 1   Account         277 non-null    object
 2   Category        277 non-null    object
 3   Note            273 non-null    object
 4   INR             277 non-null    float64
 5   Income/Expense  277 non-null    object
 6   Amount          277 non-null    float64
 7   Currency        277 non-null    object
 8   Account.1       277 non-null    float64
dtypes: float64(3), object(6)
memory usage: 19.6+ KB
```

Out [23]:

	Date	Account	Category	Note	INR	Income/Expense	Amount	Currency	Account.1	time
0	3/2/2022	CUB - online payment	Food	Brownie	50.0	Expense	50.0	INR	50.0	10:11
1	3/2/2022	CUB - online payment	Other	To lended people	300.0	Expense	300.0	INR	300.0	10:11
2	3/1/2022	CUB - online payment	Food	Dinner	78.0	Expense	78.0	INR	78.0	19:50
3	3/1/2022	CUB - online payment	Transportation	Metro	30.0	Expense	30.0	INR	30.0	18:56
4	3/1/2022	CUB - online payment	Food	Snacks	67.0	Expense	67.0	INR	67.0	18:22



RESULT:
The program to implementing different preprocessing techniques is implemented successfully.