```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
from scipy.stats import skew, kurtosis, mode #Python libraries for inferential statistics
import seaborn as sns #This is for generating Histogram with Ker
```

```
df = pd.read_csv('/content/sample_data/hotel_books.csv') #uread the 'hotel_books.csv' file
df.head(5)
```

```
   day  clients  total_bill
0    1       33       23958
1    2       25       26812
2    3        5       24871
3    4       17       17954
4    5       28       29416
```

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
df.dtypes #check for data types
```

```
               0
day        int64
clients    int64
total_bill int64

dtype: object
```

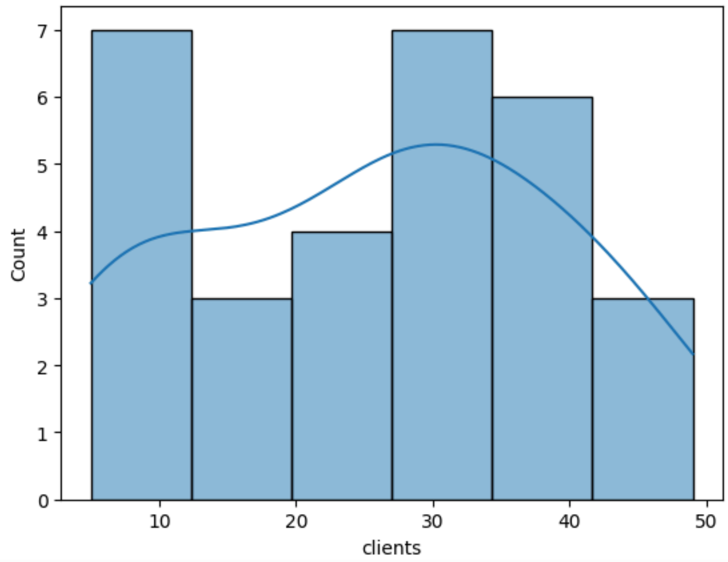
```
df.isnull().sum() #check for missing values
```

```
           0
day        0
clients    0
total_bill 0

dtype: int64
```

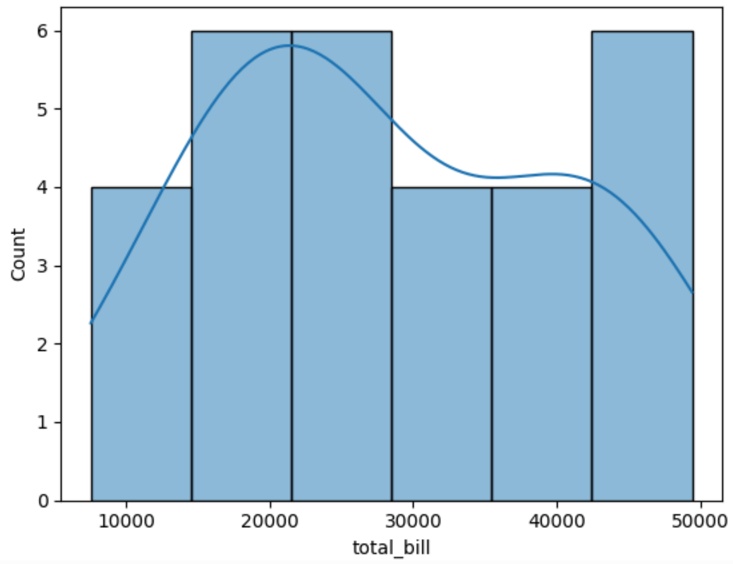
```
sns.histplot(df['clients'], kde=True) #generate histogram with kernel density estimate (KDE) for number of hotel clients
```

```
<Axes: xlabel='clients', ylabel='Count'>
```



```
sns.histplot(df['total_bill'], kde=True) #generate histogram with kernel density estimate (KDE) for total bill collected
```

```
<Axes: xlabel='total_bill', ylabel='Count'>
```



```
#compute for skewness and kurtosis for number of clients
skew1 = df['clients'].skew()
kurt1 = df['clients'].kurt()
print(f'Kurtosis for the number of hotel clients in a day:{kurt1}')
print(f'Skewness for the number of hotel clients in a day:{skew1}')
```

```
Kurtosis for the number of hotel clients in a day:-1.1388703400867874
Skewness for the number of hotel clients in a day:-0.05968808896371035
```

```
#compute for skewness and kurtosis for total number of bill
skew2 = df['total_bill'].skew()
kurt2 = df['total_bill'].kurt()
print(f'Kurtosis for the total bill collected from clients per day:{kurt2}')
print(f'Skewness for the total bill collected from clients per day:{skew2}')
```

```
Kurtosis for the total bill collected from clients per day:-1.130219880444574
Skewness for the total bill collected from clients per day:0.18976914965853053
```

```
df.describe() #generate summary measure and observe the mean and 50% (median)
```

	day	clients	total_bill
count	30.000000	30.000000	30.000000
mean	15.500000	25.666667	28344.233333
std	8.803408	13.557879	12441.769892
min	1.000000	5.000000	7534.000000
25%	8.250000	16.000000	18335.000000
50%	15.500000	28.000000	25841.500000
75%	22.750000	35.750000	39810.250000
max	30.000000	49.000000	49450.000000

```
stats.mode(df['clients']) #compute for mode
```

```
ModeResult(mode=8, count=4)
```

```
stats.mode(df['total_bill']) #compute for mode
```

```
ModeResult(mode=7534, count=1)
```

Start coding or generate with AI.

```
# -*- coding: utf-8 -*-
"""SkewKurt.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1Y-CnLzzLvAx-
Np6zBzvY1oH0hO4TXmhW
"""

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
from scipy.stats import skew, kurtosis, mode #Python libraries for
inferential statistics
import seaborn as sns #This is for generating Histogram with Ker

df = pd.read_csv('/content/sample_data/hotel_books.csv') #uread the
'hotel_books.csv' file
df.head(5)

df.dtypes #check for data types

df.isnull().sum() #check for missing values

sns.histplot(df['clients'], kde=True) #generate histogram with kernel
density estimate (KDE) for number of hotel clients

sns.histplot(df['total_bill'], kde=True) #generate histogram with kernel
density estimate (KDE) for total bill collected

#compute for skewness and kurtosis for number of clients
skew1 = df['clients'].skew()
kurt1 = df['clients'].kurt()
print(f'Kurtosis for the number of hotel clients in a day:{kurt1}')
print(f'Skewness for the number of hotel clients in a day:{skew1}')

#compute for skewness and kurtosis for total number of bill
skew2 = df['total_bill'].skew()
kurt2 = df['total_bill'].kurt()
print(f'Kurtosis for the total bill collected from clients per
day:{kurt2}')
print(f'Skewness for the total bill collected from clients per
day:{skew2}')

df.describe() #generate summary measure and observe the mean and 50%
(median)

stats.mode(df['clients']) #compute for mode

stats.mode(df['total_bill']) #compute for mode
```

{ "nbformat": 4, "nbformat_minor": 0, "metadata": { "colab": { "provenance": [] }, "kernelspec": { "name": "python3", "display_name": "Python 3" }, "language_info": { "name": "python" } }, "cells": [{ "cell_type": "code", "execution_count": 1, "metadata": { "id": "fOSnQwe1IJcw" }, "outputs": [], "source": ["import pandas as pd\n", "import numpy as np\n", "import matplotlib.pyplot as plt\n", "from scipy import stats\n", "from scipy.stats import skew, kurtosis, mode #Python libraries for inferential statistics\n", "import seaborn as sns #This is for generating Histogram with Ker"] }, { "cell_type": "code", "source": ["df = pd.read_csv('/content/sample_data/hotel_books.csv') #uread the 'hotel_books.csv' file\n", "df.head(5)"], "metadata": { "colab": { "base_uri": "https://localhost:8080/", "height": 206 }, "id": "dBt642V6ILgp", "outputId": "b253aae2-9b05-4859-c113-38ce9d49e84a" }, "execution_count": 2, "outputs": [{ "output_type": "execute_result", "data": { "text/plain": [" day clients total_bill\n", "0 1 33 23958\n", "1 2 25 26812\n", "2 3 5 24871\n", "3 4 17 17954\n", "4 5 28 29416"], "text/html": ["\n", "
\n", "
\n", "