```
import pandas as pd
import matplotlib.pyplot as plt


import pandas as pd

# قراءة ملف Excel
df = pd.read_excel('/content/data-sell4all.xlsx')

# عرض أول خمس صفوف
print(df.head())
```

```
       Pays  Age    Genre Dépenses des clients
0   France    32  Female               150.50
1  Germany    45    Male               200.75
2    Spain    28  Female                75.25
3    Italy    39    Male               180.00
4       UK    52  Female               250.30
```

```
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 65 entries, 0 to 64
Data columns (total 4 columns):
 #   Column                Non-Null Count  Dtype 
---  ------                --------------  ----- 
 0   Pays                  65 non-null     object
 1    Age                  62 non-null     object
 2   Genre                 64 non-null     object
 3   Dépenses des clients  63 non-null     object
dtypes: object(4)
memory usage: 2.2+ KB
None
```

```
print(df.columns)
```

```
Index(['Pays', ' Age ', 'Genre', 'Dépenses des clients'], dtype='object')
```

```
df.columns = df.columns.str.strip()


# تحويل القيم غير الرقمية إلى NaN
df['Age'] = pd.to_numeric(df['Age'], errors='coerce')
df['Dépenses des clients'] = pd.to_numeric(df['Dépenses des clients'], errors='coerce')


# إزالة الصفوف التي تحتوي على NaN
df = df.dropna(subset=['Age', 'Dépenses des clients'])


# حساب الوسيط والمتوسط للعمود "Age"
age_median = df['Age'].median()
age_mean = df['Age'].mean()
print(f"Age - Median: {age_median}, Mean: {age_mean}")

# حساب الوسيط والمتوسط للعمود "Dépenses des clients"
expenses_median = df['Dépenses des clients'].median()
expenses_mean = df['Dépenses des clients'].mean()
print(f"Dépenses des clients - Median: {expenses_median}, Mean: {expenses_mean}")
```

```
print(f"Dépenses des clients - Median: {expenses_median}, Mean: {expenses_mean}")
```

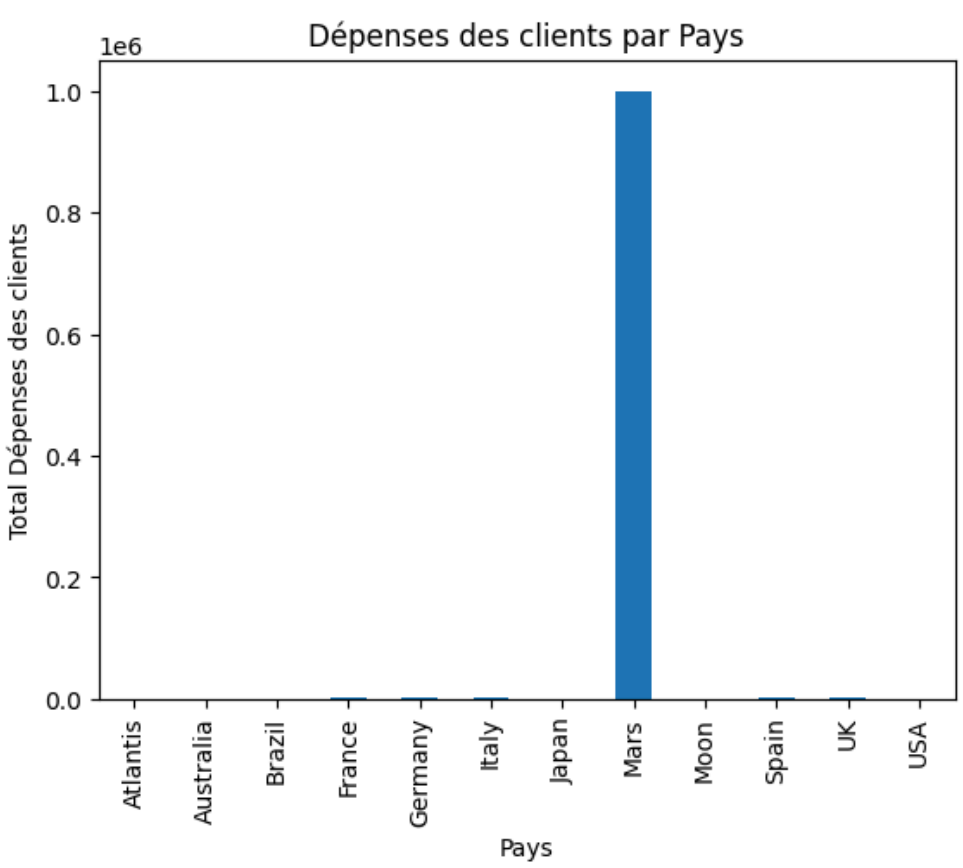
```
Age - Median: 36.5, Mean: 54.232142857142854
Dépenses des clients - Median: 167.5, Mean: 18014.616964285717
```

```
import matplotlib.pyplot as plt

# إنشاء الرسم البياني
df.groupby('Pays')['Dépenses des clients'].sum().plot(kind='bar')

# إضافة العنوان والتسميات
plt.title('Dépenses des clients par Pays')
plt.xlabel('Pays')
plt.ylabel('Total Dépenses des clients')

# عرض الرسم البياني
plt.show()
```




```
# إزالة الصفوف ذات النفقات الأقل من 10 يورو
df = df[df['Dépenses des clients'] >= 10]

# حفظ البيانات النظيفة إلى ملف CSV جديد
df.to_csv('cleaned_data.csv', columns=['Pays', 'Age', 'Genre', 'Dépenses des clients'], index=Fals

# قراءة الملف CSV الجديد للتحقق منه
cleaned_df = pd.read_csv('cleaned_data.csv')
print(cleaned_df.head())  # يعرض أول 5 صفوف من البيانات الجديدة
```

```
      Pays   Age   Genre  Dépenses des clients
0   France  32.0  Female                150.50
```

1	Germany	45.0	Male	200.75
2	Spain	28.0	Female	75.25
3	Italy	39.0	Male	180.00
4	UK	52.0	Female	250.30

Commencez à coder ou à générer avec l'IA.