Gender_Wise

In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib as plt
```

In [2]:
```
xls = pd.ExcelFile(r"C:\Users\dell\Downloads\App Analytics Report-06.05.2023 (1).xlsx")
gn_wise = pd.read_excel(xls, 'Gender Report')
```

In [3]:
```
gn_wise
```

Out[3]:

	Gender	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	unknown	13142	12691	23161	0.564077	1.762365	439.5776	761771	93180	0
1	male	7218	5877	10467	0.543091	1.450125	128.2319	282504	65651	0
2	female	4944	4304	7877	0.637710	1.593244	208.7407	274254	35083	0

In [5]:
```
gn_wise.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 10 columns):
 #   Column                     Non-Null Count  Dtype  
---  ------                     --------------  -----  
 0   Gender                     3 non-null      object 
 1   Users                      3 non-null      int64  
 2   New users                  3 non-null      int64  
 3   Engaged sessions           3 non-null      int64  
 4   Engagement rate            3 non-null      float64
 5   Engaged sessions per user  3 non-null      float64
 6   Average engagement time    3 non-null      float64
 7   Event count                3 non-null      int64  
 8   Conversions                3 non-null      int64  
 9   Total revenue              3 non-null      int64  
dtypes: float64(3), int64(6), object(1)
memory usage: 368.0+ bytes
```

In [6]:
```
gn_wise.isnull().sum()
```

Out[6]:
```
Gender                     0
Users                      0
New users                  0
Engaged sessions           0
Engagement rate            0
Engaged sessions per user  0
Average engagement time    0
Event count                0
Conversions                0
Total revenue              0
dtype: int64
```

In [9]:
```
gn_wise["Gender"].nunique()
```

Out[9]:
```
3
```

In [10]:
```
gn_wise.describe().transpose()
```

Out[10]:

	count	mean	std	min	25%	50%	75%
Users	3.0	8434.666667	4232.258184	4944.000000	6081.000000	7218.000000	10180.000000
New users	3.0	7624.000000	4458.076828	4304.000000	5090.500000	5877.000000	9284.000000
Engaged sessions	3.0	13835.000000	8179.714665	7877.000000	9172.000000	10467.000000	16814.000000
Engagement rate	3.0	0.581626	0.049691	0.543091	0.553584	0.564077	0.600893
Engaged sessions per user	3.0	1.601911	0.156300	1.450125	1.521685	1.593244	1.677805
Average engagement time	3.0	258.850067	161.608316	128.231900	168.486300	208.740700	324.159150
Event count	3.0	439509.666667	279116.984160	274254.000000	278379.000000	282504.000000	522137.500000
Conversions	3.0	64638.000000	29061.744253	35083.000000	50367.000000	65651.000000	79415.500000
Total revenue	3.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

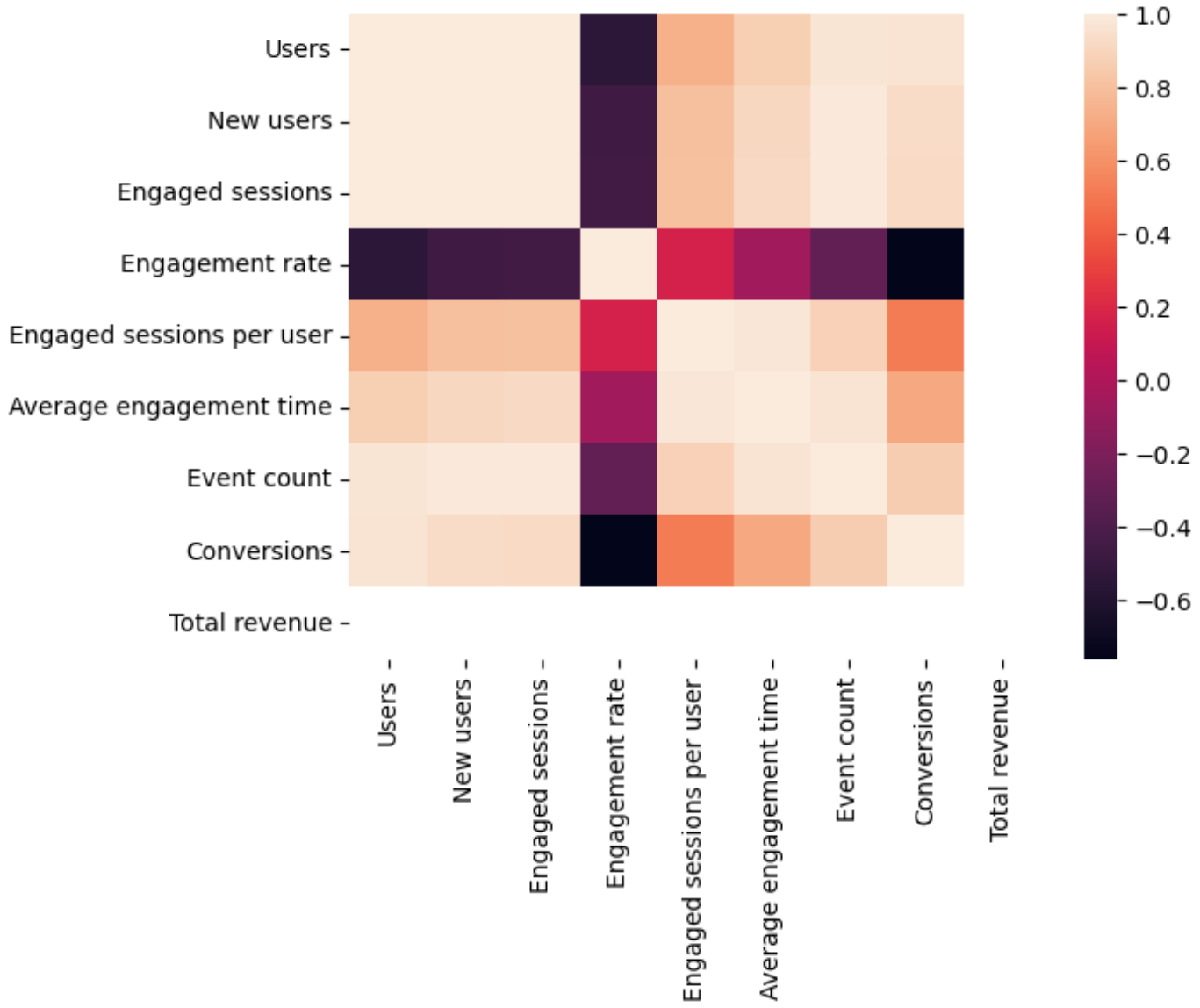
In [11]:
```
gn_wise.corr()
```

Out[11]:

	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	T reve
Users	1.000000	0.995525	0.993622	-0.550383	0.733358	0.865960	0.967103	0.960557	N
New users	0.995525	1.000000	0.999831	-0.469020	0.794322	0.909346	0.986815	0.929979	N
Engaged sessions	0.993622	0.999831	1.000000	-0.452724	0.805342	0.916832	0.989620	0.923072	N
Engagement rate	-0.550383	-0.469020	-0.452724	1.000000	0.163982	-0.059060	-0.319887	-0.760849	N
Engaged sessions per user	0.733358	0.794322	0.805342	0.163982	1.000000	0.975057	0.882175	0.515379	N
Average engagement time	0.865960	0.909346	0.916832	-0.059060	0.975057	1.000000	0.964694	0.692731	N
Event count	0.967103	0.986815	0.989620	-0.319887	0.882175	0.964694	1.000000	0.858217	N
Conversions	0.960557	0.929979	0.923072	-0.760849	0.515379	0.692731	0.858217	1.000000	N
Total revenue	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

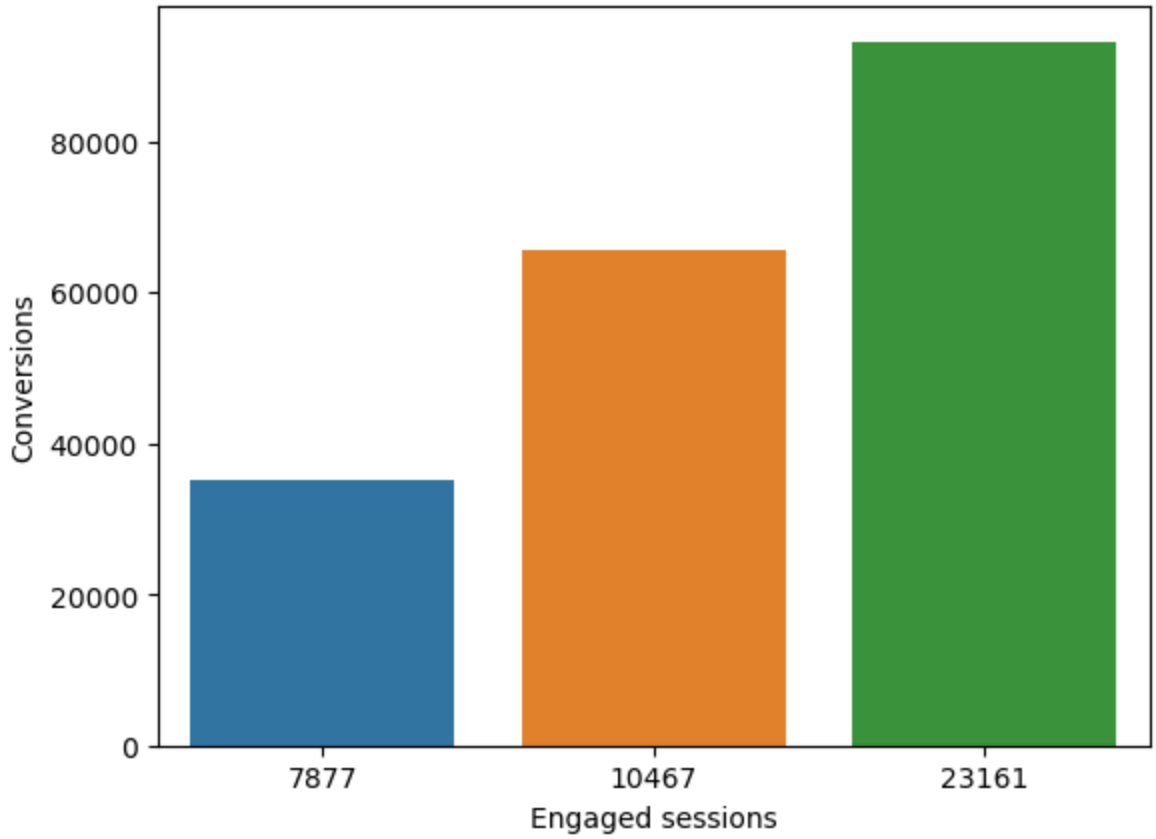
In [12]:
```
sns.heatmap(gn_wise.corr())
```

Out[12]:
```
<AxesSubplot:>
```



In [14]:
```
sns.barplot(x = "Engaged sessions" , y = "Conversions" , data = gn_wise)
```

Out[14]:
```
<AxesSubplot:xlabel='Engaged sessions', ylabel='Conversions'>
```



Observations

analysis and findings

1) No null values are present.

2) 'Users','New users' and 'Engaged sessions' are highly co-relative with conversions while 'Engagement rate' is negatively co-related with it.

3) There are 3 types of gender are given(male,female and unknown), if we devide the unknown gender equally in two parts(male and female) to get a better insights, female engage time is more than the male but the no. of users and conversion rate of male is high.