Demographics_Wise

In [18]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib as plt
```

In [19]:
```
xls = pd.ExcelFile(r"C:\Users\dell\Downloads\App Analytics Report-06.05.2023 (1).xlsx")
Dg_wise = pd.read_excel(xls, 'Demographics Report')
```

In [20]:
```
Dg_wise
```

Out[20]:

	Country	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	India	23024	22528	41479	0.593626	1.801555	334.81660	1312097	192766	0
1	United States	272	213	197	0.491272	0.724265	50.96324	3157	643	0
2	Canada	37	18	25	0.416667	0.675676	43.21622	410	121	0
3	(not set)	36	36	17	0.459459	0.472222	24.80556	241	54	0
4	United Kingdom	20	8	13	0.371429	0.650000	61.85000	289	43	0
5	Singapore	17	6	13	0.419355	0.764706	70.00000	299	29	0
6	Japan	11	6	11	0.550000	1.000000	51.45455	283	24	0
7	Australia	10	7	8	0.500000	0.800000	26.90000	132	22	0
8	Bangladesh	7	2	10	0.625000	1.428571	49.85714	121	20	0
9	Germany	7	2	6	0.500000	0.857143	15.42857	82	9	0
10	Malaysia	7	7	7	0.636364	1.000000	536.00000	507	19	0
11	Nepal	7	3	5	0.357143	0.714286	13.85714	74	18	0
12	Saudi Arabia	7	5	6	0.600000	0.857143	25.28571	74	13	0
13	United Arab Emirates	5	3	3	0.300000	0.600000	7.60000	64	15	0
14	Kuwait	4	3	5	0.833333	1.250000	36.50000	91	11	0
15	Myanmar (Burma)	3	2	8	0.533333	2.666667	47.00000	142	16	0
16	Qatar	3	2	3	0.600000	1.000000	20.66667	27	4	0
17	China	2	2	4	1.000000	2.000000	61.00000	42	5	0
18	Indonesia	2	1	2	1.000000	1.000000	8.50000	21	3	0
19	Ireland	2	2	1	0.500000	0.500000	72.00000	61	4	0
20	Italy	2	2	2	0.666667	1.000000	14.50000	23	4	0
21	Netherlands	2	2	2	0.666667	1.000000	156.50000	58	10	0
22	South Korea	2	0	4	0.444444	2.000000	45.00000	49	11	0
23	Switzerland	2	0	1	0.500000	0.500000	2.00000	9	2	0
24	Afghanistan	1	1	1	1.000000	1.000000	17.00000	6	2	0
25	Argentina	1	1	1	1.000000	1.000000	12.00000	6	2	0
26	Bahamas	1	1	1	1.000000	1.000000	13.00000	6	2	0
27	Dominican Republic	1	1	1	1.000000	1.000000	160.00000	25	2	0
28	France	1	0	1	0.333333	1.000000	14.00000	18	7	0
29	Guernsey	1	1	1	1.000000	1.000000	30.00000	8	2	0
30	Iran	1	0	0	0.000000	0.000000	1.00000	3	1	0
31	Kyrgyzstan	1	1	1	1.000000	1.000000	20.00000	11	2	0
32	Latvia	1	1	1	1.000000	1.000000	16.00000	7	1	0
33	Norway	1	0	0	0.000000	0.000000	1.00000	7	1	0

	Country	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
34	Oman	1	1	1	1.000000	1.000000	2.00000	9	2	0
35	Panama	1	0	1	1.000000	1.000000	6.00000	9	1	0
36	Romania	1	1	1	1.000000	1.000000	13.00000	6	1	0
37	Russia	1	1	1	0.500000	1.000000	152.00000	23	4	0
38	Serbia	1	1	1	1.000000	1.000000	32.00000	8	2	0
39	Sweden	1	1	1	1.000000	1.000000	9.00000	8	2	0
40	Czechia	0	0	0	0.000000	0.000000	0.00000	2	2	0
41	Hungary	0	0	0	0.000000	0.000000	0.00000	1	1	0
42	Kenya	0	0	0	0.000000	0.000000	0.00000	1	0	0
43	Maldives	0	0	0	0.000000	0.000000	0.00000	1	1	0
44	Pakistan	0	0	0	0.000000	0.000000	0.00000	3	2	0
45	Sri Lanka	0	0	0	0.000000	0.000000	0.00000	1	1	0
46	Ukraine	0	0	0	0.000000	0.000000	0.00000	7	7	0

In [21]:
```
Dg_wise.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 47 entries, 0 to 46
Data columns (total 10 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Country                   47 non-null     object 
 1   Users                     47 non-null     int64  
 2   New users                 47 non-null     int64  
 3   Engaged sessions          47 non-null     int64  
 4   Engagement rate           47 non-null     float64
 5   Engaged sessions per user  47 non-null     float64
 6   Average engagement time   47 non-null     float64
 7   Event count               47 non-null     int64  
 8   Conversions               47 non-null     int64  
 9   Total revenue             47 non-null     int64  
dtypes: float64(3), int64(6), object(1)
memory usage: 3.8+ KB
```

In [22]:
```
Dg_wise.isnull().sum()
```

Out[22]:
```
Country                     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 [23]:
```
Dg_wise["Country"].nunique()
```

Out[23]:
```
47
```

In [ ]:
```
Dg_wise.describe().transpose()
```

Out[24]:

	count	mean	std	min	25%	50%	75%	max
Users	47.0	500.148936	3357.099816	0.0	1.000000	2.000000	7.00000	2.302400e+04
New users	47.0	486.638298	3285.103773	0.0	0.000000	1.000000	3.00000	2.252800e+04
Engaged sessions	47.0	890.340426	6049.238031	0.0	1.000000	1.000000	6.00000	4.147900e+04
Engagement rate	47.0	0.561662	0.358970	0.0	0.364286	0.533333	1.00000	1.000000e+00
Engaged sessions per user	47.0	0.835367	0.568228	0.0	0.550000	1.000000	1.00000	2.666667e+00
Average engagement time	47.0	48.376626	93.511799	0.0	6.800000	17.000000	48.42857	5.360000e+02
Event count	47.0	28053.808511	191369.184392	1.0	7.000000	23.000000	86.50000	1.312097e+06
Conversions	47.0	4125.829787	28114.328609	0.0	2.000000	4.000000	15.50000	1.927660e+05
Total revenue	47.0	0.000000	0.000000	0.0	0.000000	0.000000	0.00000	0.000000e+00

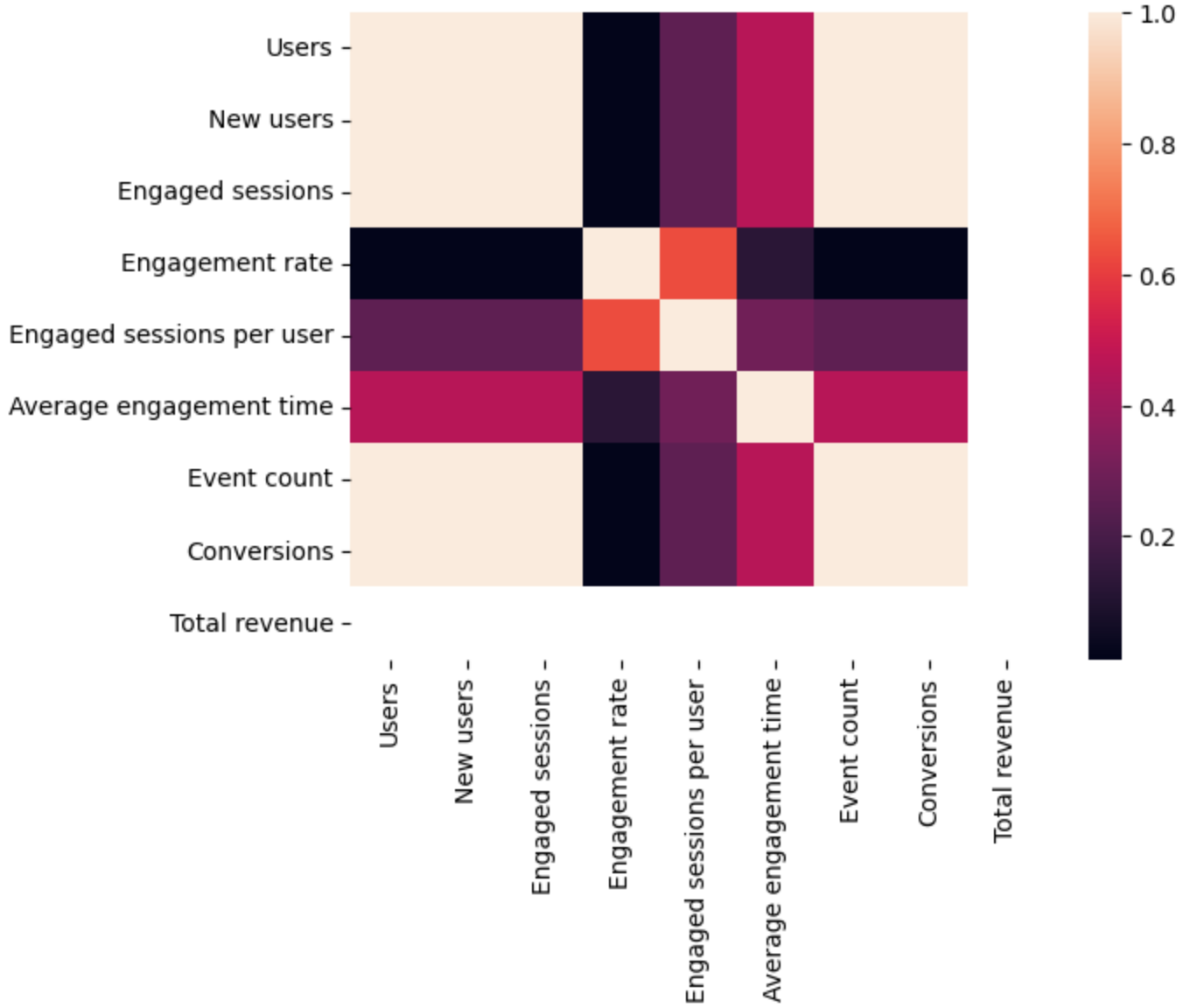
In [25]:
```
Dg_wise.corr()
```

Out[25]:

	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Tota revenu
Users	1.000000	0.999997	0.999975	0.012727	0.253096	0.456856	0.999955	0.999964	Nal
New users	0.999997	1.000000	0.999988	0.012950	0.253141	0.456841	0.999975	0.999981	Nal
Engaged sessions	0.999975	0.999988	1.000000	0.013110	0.253468	0.456710	0.999997	0.999999	Nal
Engagement rate	0.012727	0.012950	0.013110	1.000000	0.629911	0.124109	0.013168	0.013094	Nal
Engaged sessions per user	0.253096	0.253141	0.253468	0.629911	1.000000	0.294999	0.253441	0.253361	Nal
Average engagement time	0.456856	0.456841	0.456710	0.124109	0.294999	1.000000	0.456871	0.456638	Nal
Event count	0.999955	0.999975	0.999997	0.013168	0.253441	0.456871	1.000000	0.999999	Nal
Conversions	0.999964	0.999981	0.999999	0.013094	0.253361	0.456638	0.999999	1.000000	Nal
Total revenue	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Nal

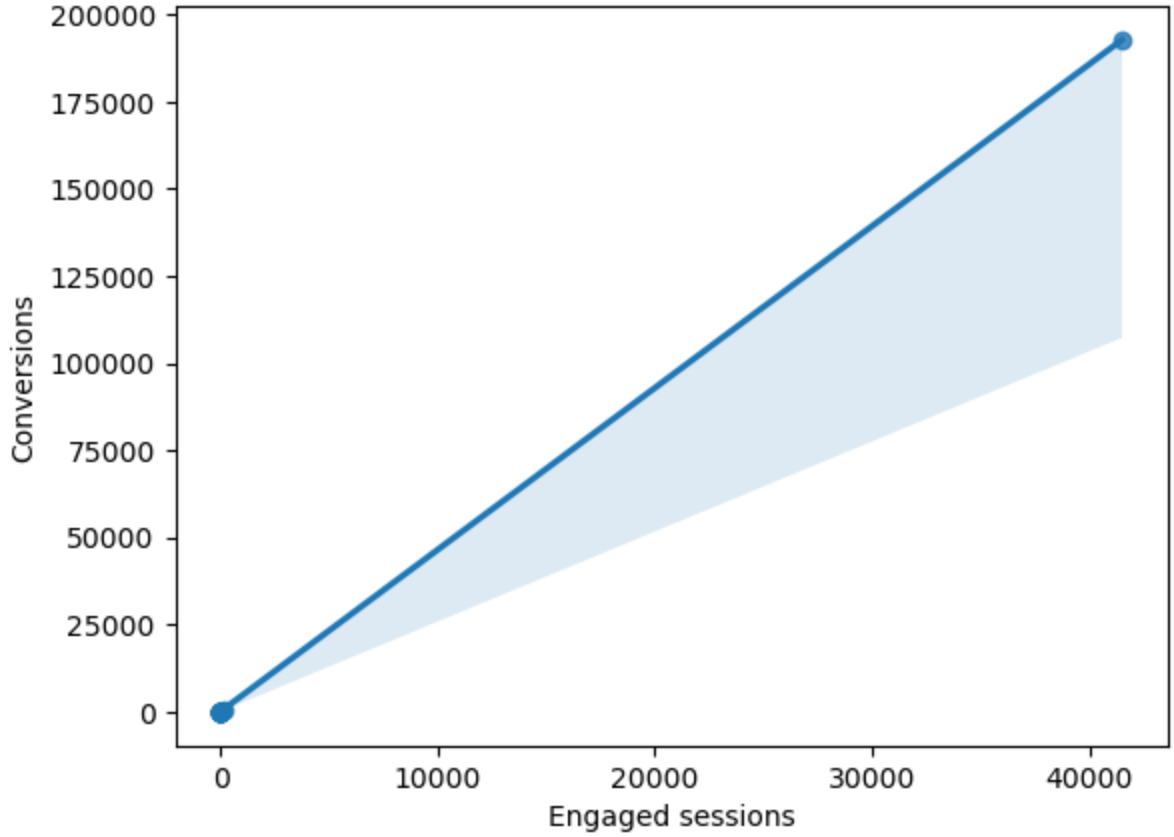
In [26]:
```
sns.heatmap(Dg_wise.corr())
```

Out[26]:
```
<AxesSubplot:>
```



In [27]:
```
sns.regplot(x = "Engaged sessions" , y = "Conversions" , data = Dg_wise)
```

Out[27]:
```
<AxesSubplot:xlabel='Engaged sessions', ylabel='Conversions'>
```



Observations

analysis and findings

1) There are total 47 countries where 'XYZ' company is running, out of which India uses it's website maximum.

2) No null values present in the given data.

3) The country where the 'Users','New users','Engaged sessions' and 'Event count' is more convertions rate are higher and 'Engagement rate','Engaged sessions per user' and Average engagement time is less convertions rate are also poor.