Import Liraries and Load the dataset.

In [1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
xls = pd.ExcelFile(r'C:\Users\NEW\Downloads\Data set for BA.xlsx')
sheet_names = xls.sheet_names
dfs = {sheet_name: xls.parse(sheet_name) for sheet_name in sheet_names}
dfs
```

Out[2]:
```
{'Report Snapshot': Empty DataFrame
 Columns: []
 Index: [],
 'User Acquisition':   First user default channel group  New users  Engaged sessions  \
 0                           Display       9957             12008
 1                  Organic Search       7652             18141
 2                     Paid Search       3025              4408
 3                          Direct       1903              4975
 4                      Unassigned        325              1619
 5                  Organic Social         10                13

    Engagement rate  Engaged sessions per user  Average engagement time  \
 0         0.544457                   1.206107                 58.86209
 1         0.813680                   2.367041                534.31280
 2         0.474284                   1.458154                102.23780
 3         0.318808                   2.261364               1128.88100
 4         0.813159                   4.981538                798.34150
 5         0.722222                   1.300000                145.30000

    Event count  Conversions  Total revenue
 0       204820        37434              0
 1       770710       109801              0
 2        81997        14770              0
 3       227434        31093              0
 4        33320          789              0
 5          248           27              0  ,
 'Traffic Aquisition':   Session default channel group  Users  Sessions  Engaged sessions
\
 0                     Unassigned  20263     13448              1481
 1                        Display   9613     18292             10613
 2                  Organic Search   7689     21241             17814
 3                         Direct   4042     13220              7649
 4                     Paid Search   2909      6788              3452
 5                  Organic Social     11        16                12

    Average engagement time per session  Engaged sessions per user  \
 0                             34.11704                   0.073089
 1                             28.52198                   1.104026
 2                            195.94340                   2.316816
 3                            177.17060                   1.892380
 4                             36.65321                   1.186662
 5                             60.06250                   1.090909

    Events per session  Engagement rate  Event count  Conversions  \
 0           18.023130         0.110128       242375       114161
 1            9.069320         0.580199       165896        20031
 2           29.302290         0.838661       622410        33612
 3           17.135850         0.578593       226536        18496
 4            8.989982         0.508544        61024         7595
 5           18.000000         0.750000          288           19
```

```
   Total revenue
0              0
1              0
2              0
3              0
4              0
5              0  ,
'Event Report':                          Event name  Event count  Total users  \
0                       screen_view       694729        23254
1              notification_receive       125146         1700
2                   user_engagement       124836        22699
3              notification_dismiss        70128         1369
4                     session_start        61163        23226
..                              ...          ...          ...
374  Promilo119_myProfile_mediator            1            1
375                  Promilo_feeds            1            1
376                          feeds            1            1
377           my_interests_screen             1            1
378                      (not set)            0        22269

     Event count per user  Total revenue
0               30.865870              0
1              138.896800              0
2                5.622230              0
3              144.000000              0
4                3.121357              0
..                    ...            ...
374              1.000000              0
375              1.000000              0
376              1.000000              0
377              1.000000              0
378              0.000000              0

[379 rows x 5 columns],
'Conversion Report':                              Event name  Conversions  Total users  Tota
l revenue
0               notification_receive        94890         1311              0
1                      session_start        56203        21674              0
2                         first_open        22872        23059              0
3                         app_remove        12468        12538              0
4              Promilo111_otp_screen         1738          855              0
5   Promilo111_Event_Enter_Feed_Page         1594          969              0
6                  Promilo106_login         1458          603              0
7                  Promilo106_feeds          683          185              0
8                         os_update          672          634              0
9                 notification_open          569          308              0
10            Promilo106_feedDetails          195           67              0
11     Promilo106_my_meetings_screen          136           20              0
12             Promilo106_otp_screen          128           90              0
13         Promilo106_resume_builder          127           66              0
14    Promilo106_my_interests_screen          117           23              0
15             Promilo106_dashboard            23           16              0
16    Promilo106_my_profile_learners           21           10              0
17      Promilo106_campaign_interest           20           10              0,
'Pages & Screens Report':             Page path and screen class   Views  Users  Views pe
r user  \
0                              Flutter  156708   8726       17.958740
1                         MainActivity   44326   8978        4.937180
2                                feeds   18514   4358        4.248279
3                                login   16883   7291        2.315595
4                    my_rewards_screen   15381   2045        7.521271
5                           storyboard    8189   5244        1.561594
6                    SignInHubActivity    6650   3778        1.760191
7                 registration_screen     5501   3566        1.542625
8                           feedDetails    3971   1047        3.792741
```

9	otp_screen	3291	1678	1.961263
10	video_viewer_screem	2880	1521	1.893491
11	FacebookActivity	2299	675	3.405926
12	resume_builder	1781	828	2.150966
13	CustomTabMainActivity	1301	193	6.740933
14	notification_store	1062	648	1.638889
15	dashboard	1058	411	2.574209
16	myProfile_mediator	1056	600	1.760000
17	WebViewActivity	878	490	1.791837
18	video_tutorial_view	835	722	1.156510
19	my_profile_learners	804	321	2.504673
20	FlutterViewController	758	155	4.890323
21	my_meetings_screen	715	273	2.619048
22	my_interests_screen	688	375	1.834667
23	discovery_screen	486	225	2.160000
24	calculator_intro	388	281	1.380783
25	campaign_interest	244	58	4.206897
26	calculator_one	199	151	1.317881
27	UIActivityViewSuccessController	99	18	5.500000
28	UIActivityContentViewController	97	17	5.705882
29	calculator_two	88	73	1.205479
30	SFAuthenticationViewController	76	45	1.688889
31	my_profile_professional	67	23	2.913043
32	CustomTabActivity	35	24	1.458333
33	SFSafariViewController	34	21	1.619048
34	CheckoutActivity	24	3	8.000000
35	my_profile_others	20	12	1.666667
36	PHPickerViewController	13	10	1.300000
37	SLComposeViewController	6	4	1.500000
38	campaign_meeting	4	3	1.333333
39	CAMImagePickerCameraViewController	2	2	1.000000
40	UIAlertController	2	2	1.000000
41	(not set)	0	9145	0.000000

	Average engagement time	Event count	Conversions	Total revenue
0	83.412220	203901	328	0
1	78.292160	53374	101	0
2	61.600050	37628	253	0
3	34.881770	40772	435	0
4	94.179950	32910	5	0
5	5.341152	15676	115	0
6	0.003176	6653	0	0
7	45.075720	13496	136	0
8	69.316140	7820	84	0
9	46.864720	10833	32	0
10	28.120970	5256	31	0
11	0.524444	2310	0	0
12	118.043500	3776	17	0
13	0.005181	1302	0	0
14	10.137350	1971	0	0
15	38.114360	2279	4	0
16	57.020000	2276	4	0
17	105.806100	1321	2	0
18	36.303320	1662	2	0
19	127.722700	1856	4	0
20	15.625810	1060	18	0
21	45.765570	1480	15	0
22	29.189330	1340	7	0
23	70.235560	890	7	0
24	6.391459	680	0	0
25	42.379310	459	0	0
26	41.013250	387	0	0
27	0.111111	101	0	0
28	32.411760	177	0	0
29	9.287671	158	0	0
30	12.200000	161	0	0

```
31              184.087000          143             0              0
32                1.666667           36             0              0
33               46.857140           74             0              0
34              304.333300           26             0              0
35              123.833300           49             1              0
36                9.500000           27             0              0
37               15.250000           13             0              0
38               14.333330            7             0              0
39                7.500000            4             0              0
40                1.000000            4             0              0
41                0.001093       109189         88518              0  ,
'Retention Overview': Empty DataFrame
Columns: []
Index: [],
'User Engagement Overview': Empty DataFrame
Columns: []
Index: [],
'Demographics Report':                  Country  Users  New users  Engaged sessions  Enga
gement rate  \
0                    India  23024      22528             41479         0.593626
1            United States    272        213               197         0.491272
2                   Canada     37         18                25         0.416667
3                (not set)     36         36                17         0.459459
4           United Kingdom     20          8                13         0.371429
5                Singapore     17          6                13         0.419355
6                    Japan     11          6                11         0.550000
7                Australia     10          7                 8         0.500000
8               Bangladesh      7          2                10         0.625000
9                  Germany      7          2                 6         0.500000
10               Malaysia      7          7                 7         0.636364
11                  Nepal      7          3                 5         0.357143
12           Saudi Arabia      7          5                 6         0.600000
13  United Arab Emirates      5          3                 3         0.300000
14                 Kuwait      4          3                 5         0.833333
15       Myanmar (Burma)      3          2                 8         0.533333
16                 Qatar      3          2                 3         0.600000
17                 China      2          2                 4         1.000000
18             Indonesia      2          1                 2         1.000000
19               Ireland      2          2                 1         0.500000
20                 Italy      2          2                 2         0.666667
21           Netherlands      2          2                 2         0.666667
22           South Korea      2          0                 4         0.444444
23           Switzerland      2          0                 1         0.500000
24           Afghanistan      1          1                 1         1.000000
25             Argentina      1          1                 1         1.000000
26               Bahamas      1          1                 1         1.000000
27    Dominican Republic      1          1                 1         1.000000
28                 France      1          0                 1         0.333333
29               Guernsey      1          1                 1         1.000000
30                   Iran      1          0                 0         0.000000
31              Kyrgyzstan      1          1                 1         1.000000
32                  Latvia      1          1                 1         1.000000
33                  Norway      1          0                 0         0.000000
34                    Oman      1          1                 1         1.000000
35                  Panama      1          0                 1         1.000000
36                 Romania      1          1                 1         1.000000
37                  Russia      1          1                 1         0.500000
38                  Serbia      1          1                 1         1.000000
39                  Sweden      1          1                 1         1.000000
40                 Czechia      0          0                 0         0.000000
41                 Hungary      0          0                 0         0.000000
42                   Kenya      0          0                 0         0.000000
43                Maldives      0          0                 0         0.000000
44                Pakistan      0          0                 0         0.000000
45               Sri Lanka      0          0                 0         0.000000
46                 Ukraine      0          0                 0         0.000000
```

	Engaged sessions per user	Average engagement time	Event count	\
0	1.801555	334.81660	1312097	
1	0.724265	50.96324	3157	
2	0.675676	43.21622	410	
3	0.472222	24.80556	241	
4	0.650000	61.85000	289	
5	0.764706	70.00000	299	
6	1.000000	51.45455	283	
7	0.800000	26.90000	132	
8	1.428571	49.85714	121	
9	0.857143	15.42857	82	
10	1.000000	536.00000	507	
11	0.714286	13.85714	74	
12	0.857143	25.28571	74	
13	0.600000	7.60000	64	
14	1.250000	36.50000	91	
15	2.666667	47.00000	142	
16	1.000000	20.66667	27	
17	2.000000	61.00000	42	
18	1.000000	8.50000	21	
19	0.500000	72.00000	61	
20	1.000000	14.50000	23	
21	1.000000	156.50000	58	
22	2.000000	45.00000	49	
23	0.500000	2.00000	9	
24	1.000000	17.00000	6	
25	1.000000	12.00000	6	
26	1.000000	13.00000	6	
27	1.000000	160.00000	25	
28	1.000000	14.00000	18	
29	1.000000	30.00000	8	
30	0.000000	1.00000	3	
31	1.000000	20.00000	11	
32	1.000000	16.00000	7	
33	0.000000	1.00000	7	
34	1.000000	2.00000	9	
35	1.000000	6.00000	9	
36	1.000000	13.00000	6	
37	1.000000	152.00000	23	
38	1.000000	32.00000	8	
39	1.000000	9.00000	8	
40	0.000000	0.00000	2	
41	0.000000	0.00000	1	
42	0.000000	0.00000	1	
43	0.000000	0.00000	1	
44	0.000000	0.00000	3	
45	0.000000	0.00000	1	
46	0.000000	0.00000	7	

	Conversions	Total revenue
0	192766	0
1	643	0
2	121	0
3	54	0
4	43	0
5	29	0
6	24	0
7	22	0
8	20	0
9	9	0
10	19	0
11	18	0
12	13	0
13	15	0
14	11	0

```
15           16              0
16            4              0
17            5              0
18            3              0
19            4              0
20            4              0
21           10              0
22           11              0
23            2              0
24            2              0
25            2              0
26            2              0
27            2              0
28            7              0
29            2              0
30            1              0
31            2              0
32            1              0
33            1              0
34            2              0
35            1              0
36            1              0
37            4              0
38            2              0
39            2              0
40            2              0
41            1              0
42            0              0
43            1              0
44            2              0
45            1              0
46            7              0  ,
'Citiwise Report':       Town/City  Users  New users  Engaged sessions  Engagement rate
\
0     Bengaluru   6097       5685             15013         0.769385
1         Patna   1594       1467              2127         0.440646
2     Hyderabad   1038        920              1578         0.569264
3        Indore    983        915              1241         0.426460
4       Lucknow    897        839              1125         0.450180
..          ...    ...        ...               ...              ...
569     Titwala      0          0                 0         0.000000
570     Vagator      0          0                 0         0.000000
571   Vellakoil      0          0                 0         0.000000
572      Wardha      0          0                 0         0.000000
573  Washington      0          0                 0         0.000000

     Engaged sessions per user  Average engagement time  Event count  \
0                     2.462359                762.20550       607200
1                     1.334379                 98.22208        38830
2                     1.520231                243.69080        96826
3                     1.262462                 67.89115        21383
4                     1.254181                 83.40580        21041
..                         ...                      ...          ...
569                   0.000000                  0.00000            1
570                   0.000000                  0.00000           16
571                   0.000000                  0.00000            1
572                   0.000000                  0.00000            1
573                   0.000000                  0.00000            1

     Conversions  Total revenue
0          62939              0
1           6980              0
2          34103              0
3           4121              0
4           3650              0
..           ...            ...
```

```
569           1              0
570          16              0
571           0              0
572           1              0
573           1              0

[574 rows x 10 columns],
'Gender Report':     Gender  Users  New users  Engaged sessions  Engagement rate  \
0  unknown  13142      12691             23161         0.564077
1     male   7218       5877             10467         0.543091
2   female   4944       4304              7877         0.637710

   Engaged sessions per user  Average engagement time  Event count  \
0                   1.762365                 439.5776       761771
1                   1.450125                 128.2319       282504
2                   1.593244                 208.7407       274254

   Conversions  Total revenue
0        93180              0
1        65651              0
2        35083              0  ,
'User By Interest':                                          Interests  Users  New use
rs  \
0                                          Shoppers  10950       9256
1       Media & Entertainment/Comics & Animation Fans  10946       9247
2                         Technology/Mobile Enthusiasts  10934       9239
3                    Food & Dining/Cooking Enthusiasts   8410       6970
4               Sports & Fitness/Health & Fitness Buffs   5844       4580
..                                                 ...    ...        ...
84                                      Food & Dining     15          4
85                                      Home & Garden     15          5
86  Sports & Fitness/Sports Fans/Racquetball Enthu...     11         11
87                           Vehicles & Transportation     11          3
88  Sports & Fitness/Sports Fans/Fans of American ...     10          4

    Engaged sessions  Engagement rate  Engaged sessions per user  \
0              15652         0.581534                   1.429406
1              15680         0.583008                   1.432487
2              15619         0.582451                   1.428480
3              12332         0.602325                   1.466350
4               8226         0.588328                   1.407598
..               ...              ...                        ...
84                24         0.489796                   1.600000
85                12         0.631579                   0.800000
86                21         0.840000                   1.909091
87                 9         0.450000                   0.818182
88                18         0.782609                   1.800000

    Average engagement time  Event count  Conversions  Total revenue
0                 162.83470       490664        86846              0
1                 165.17720       491025        86845              0
2                 162.69450       489353        86742              0
3                 176.95670       409713        73814              0
4                 155.14510       257831        43074              0
..                      ...          ...          ...            ...
84                 70.86667          460           58              0
85                133.86670          453          107              0
86                487.45450          736           39              0
87                 71.54545          161           27              0
88                201.40000          375           81              0

[89 rows x 10 columns],
'User by Language':      Language  Users  New users  Engaged sessions  Engagement rate  \
0     English  22495      21990             40639         0.595147
1       Hindi    586        552               798         0.406314
2     Marathi     85         84                98         0.426087
```

3	Gujarati	78	77	100	0.448430
4	Telugu	43	42	56	0.455285
5	Tamil	36	36	43	0.518072
6	Malayalam	17	15	36	0.654545
7	Bengali	14	11	18	0.600000
8	Chinese	13	13	13	1.000000
9	Kannada	13	12	31	0.500000
10	Panjabi	9	9	17	0.708333
11	Persian	8	8	6	0.400000
12	Spanish	6	6	8	0.470588
13	Finnish	4	3	4	0.571429
14	Japanese	4	4	3	0.428571
15	Oriya	4	4	2	0.666667
16	Afrikaans	1	1	1	1.000000
17	Assamese	1	0	1	1.000000
18	German	1	1	0	0.000000
19	Malay	1	1	1	1.000000
20	Nepali	1	1	1	1.000000
21	Russian	1	1	0	0.000000
22	Urdu	1	1	0	0.000000
23	Sanskrit	0	0	0	0.000000

	Engaged sessions per user	Average engagement time	Event count	\
0	1.806579	341.36350	1297970	
1	1.361775	60.03413	13523	
2	1.152941	38.48235	1589	
3	1.282051	46.53846	1794	
4	1.302326	36.65116	812	
5	1.194444	45.86111	615	
6	2.117647	161.94120	548	
7	1.285714	50.07143	217	
8	1.000000	136.76920	138	
9	2.384615	249.07690	680	
10	1.888889	92.44444	229	
11	0.750000	28.25000	99	
12	1.333333	22.16667	113	
13	1.000000	89.25000	64	
14	0.750000	9.25000	49	
15	0.500000	7.50000	29	
16	1.000000	37.00000	12	
17	1.000000	42.00000	6	
18	0.000000	0.00000	5	
19	1.000000	2.00000	7	
20	1.000000	5.00000	7	
21	0.000000	70.00000	12	
22	0.000000	1.00000	7	
23	0.000000	0.00000	4	

	Conversions	Total revenue
0	189946	0
1	2699	0
2	323	0
3	327	0
4	170	0
5	115	0
6	71	0
7	39	0
8	20	0
9	75	0
10	35	0
11	23	0
12	22	0
13	11	0
14	12	0
15	10	0
16	2	0

```
17            1              0
18            2              0
19            2              0
20            2              0
21            1              0
22            4              0
23            2              0  ,
'User By Age':        Age  Users  New users  Engaged sessions  Engagement rate  \
0  unknown  14303      13636             24976         0.569098
1    18-24   4282       3678              7291         0.695308
2    25-34   2920       2161              3749         0.504780
3      65+   1422       1081              1640         0.539829
4    55-64   1403        979              1552         0.519411
5    35-44   1202        785              1420         0.510424
6    45-54    810        552               881         0.561862

   Engaged sessions per user  Average engagement time  Event count  \
0                   1.746207                422.22330       817501
1                   1.702709                251.16300       309328
2                   1.283904                 97.24144        90074
3                   1.153305                 52.30661        24780
4                   1.106201                 55.37063        25169
5                   1.181364                 96.08236        33016
6                   1.087654                 84.54321        18661

   Conversions  Total revenue
0        99310              0
1        53661              0
2        20172              0
3         4891              0
4         4823              0
5         8111              0
6         2946              0  ,
'Google Ads Report':                    Session Google Ads campaign  Users  Sessions  \
0            App Installation for May --Shahid   5429     10936
1          App Install-States-A200Inst-20Jun22    842      1655
2      App Install-States-B100Installs-22Jun22    742      1332
3              App Install for April -- Shahid    473       976
4      Video-AppInstall-PS-Internships-11Jul22    510       966
5                        App promotion-App-3    732       945
6     App Instal-States-B200 &A100Inst-22Jun22    373       742
7         App Install-1to5NC-StateA200-07Jul22    370       610
8   App Instal-6to10NC-States-A200Inst-07Jul22    242       432
9         Video-AppInstall-PS-Browsing-11Jul22     91       188
10         Video-AppInstall-PS-Webinar-11Jul22     78       124
11        Video-AppInstall-PS-Colleges-11Jul22     46        77
12          Video-AppInstall-PS-Videos-11Jul22     38        75
13            Video-AppInstall-PS-Jobs-11Jul22     25        49
14      App installation for May 06-05-2022 T1      2         5

    Engaged sessions  Google Ads clicks  Google Ads cost  \
0               6276             147100     179175.00000
1                968              28742      24309.13000
2                780              17809      22374.58000
3                546              19302      20525.18000
4                515               9831       6377.83300
5                763               5793      12084.04000
6                425              10595      11993.01000
7                462               3659       8839.72300
8                296               4475       9204.69600
9                112               1899       1535.27000
10                81                893       1903.41800
11                50               1851       1263.62500
12                39               1706        935.72270
13                32               1649        800.09510
14                 5                 14         16.62396
```

	Google Ads cost per click	Conversions	Cost per conversion	Event count	\
0	1.218049	12257	14.618180	97802	
1	0.845770	1794	13.550240	15311	
2	1.256363	1422	15.734580	11640	
3	1.063370	1115	18.408230	8001	
4	0.648747	1032	6.180071	10323	
5	2.085972	922	13.106330	10825	
6	1.131950	851	14.092850	7504	
7	2.415885	709	12.467870	10864	
8	2.056915	630	14.610630	6092	
9	0.808462	206	7.452768	2172	
10	2.131487	150	12.689450	1600	
11	0.682671	101	12.511140	1268	
12	0.548489	76	12.312140	728	
13	0.485200	53	15.096130	731	
14	1.187426	5	3.324793	163	

	Total revenue	Return on ad spend
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0 }

Lets Check the sheets & their names.

In [3]:
```
sheet_names
```

Out[3]:
```
['Report Snapshot',
 'User Acquisition',
 'Traffic Aquisition',
 'Event Report',
 'Conversion Report',
 'Pages & Screens Report',
 'Retention Overview',
 'User Engagement Overview',
 'Demographics Report',
 'Citiwise Report',
 'Gender Report',
 'User By Interest',
 'User by Language',
 'User By Age',
 'Google Ads Report']
```

In [4]:
```
df1 = dfs['User Acquisition']
df2 = dfs['Traffic Aquisition']
df3 = dfs['Event Report']
df4 = dfs['Conversion Report']
df5 = dfs['Pages & Screens Report']
df6 = dfs['Demographics Report']
df7 = dfs['Citiwise Report']
df8 = dfs['Gender Report']
df9 = dfs['User By Interest']
```

```
df10 = dfs['User by Language']
df11 = dfs['User By Age']
df12 = dfs['Google Ads Report']
```

Data cleaning.

Let's find if there is any null values in the dataframe

In [5]:

```
dfs_list = [df1, df2, df3, df4, df5, df6, df7, df8, df9, df10, df11, df12]

for i, df in enumerate(dfs_list, start=1):
    print(f"DataFrame {i}:")
    print(df.isnull().sum())
    print("\n---\n")
```

```
DataFrame 1:
First user default channel group    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

---

DataFrame 2:
Session default channel group          0
Users                                  0
Sessions                               0
Engaged sessions                       0
Average engagement time per session    0
Engaged sessions per user              0
Events per session                     0
Engagement rate                        0
Event count                            0
Conversions                            0
Total revenue                          0
dtype: int64

---

DataFrame 3:
Event name              1
Event count             0
Total users             0
Event count per user    0
Total revenue           0
dtype: int64

---

DataFrame 4:
Event name       0
Conversions      0
Total users      0
Total revenue    0
dtype: int64

---
```

```
DataFrame 5:
Page path and screen class    0
Views                         0
Users                         0
Views per user                0
Average engagement time       0
Event count                   0
Conversions                   0
Total revenue                 0
dtype: int64

---

DataFrame 6:
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

---

DataFrame 7:
Town/City                     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

---

DataFrame 8:
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

---

DataFrame 9:
Interests                     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

---

DataFrame 10:
Language                     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

---

DataFrame 11:
Age                          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

---

DataFrame 12:
Session Google Ads campaign    0
Users                          0
Sessions                       0
Engaged sessions               0
Google Ads clicks              0
Google Ads cost                0
Google Ads cost per click      0
Conversions                    0
Cost per conversion            0
Event count                    0
Total revenue                  0
Return on ad spend             0
dtype: int64

---
```

In [6]:

```
df3 = df3.dropna()
df3.isnull().sum()
```

Out[6]:

```
Event name              0
Event count             0
Total users             0
Event count per user    0
Total revenue           0
dtype: int64
```

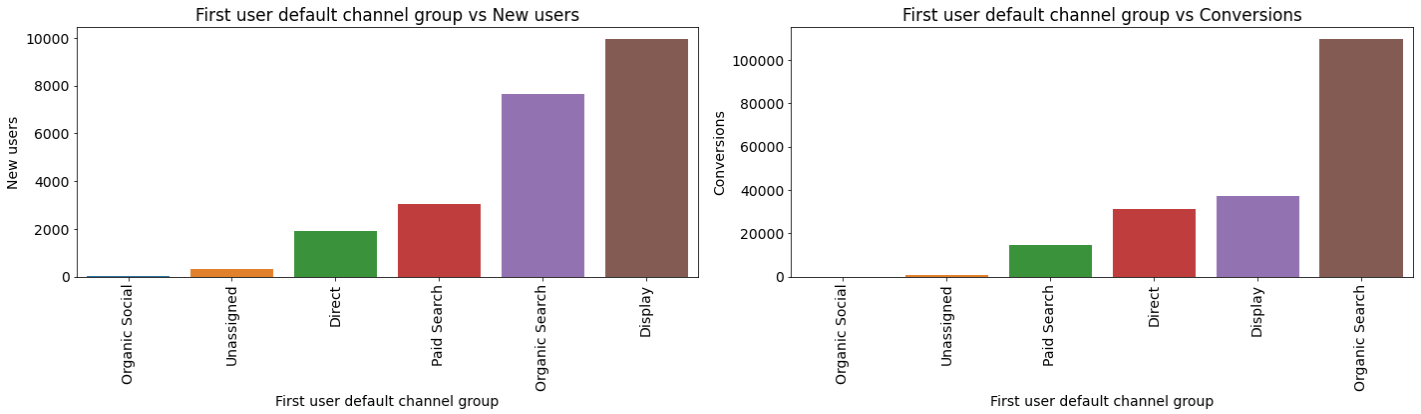
(Now, there is no null value & seems fine)

For a Data Visualization, I choose the 'Seaborn' & 'matplot' library...

User Acquisition

Let's analyse in the User Acquisition Data & let's which has to be improve...

In [7]:
```
plt.rcParams.update({'font.size': 14})

fig, axs = plt.subplots(nrows=1, ncols=2, figsize=(20, 6))

# channel group vs New users...
df1 = df1.sort_values('New users')
sns.barplot(x='First user default channel group', y='New users', data=df1, ax=axs[0])
axs[0].set_title('First user default channel group vs New users')
axs[0].set_xlabel('First user default channel group')
axs[0].set_ylabel('New users')
axs[0].tick_params(axis='x', rotation=90)

# channel group vs Conversions...
df1 = df1.sort_values('Conversions')
sns.barplot(x='First user default channel group', y='Conversions', data=df1, ax=axs[1])
axs[1].set_title('First user default channel group vs Conversions')
axs[1].set_xlabel('First user default channel group')
axs[1].set_ylabel('Conversions')
axs[1].tick_params(axis='x', rotation=90)

plt.tight_layout()
plt.show()
```

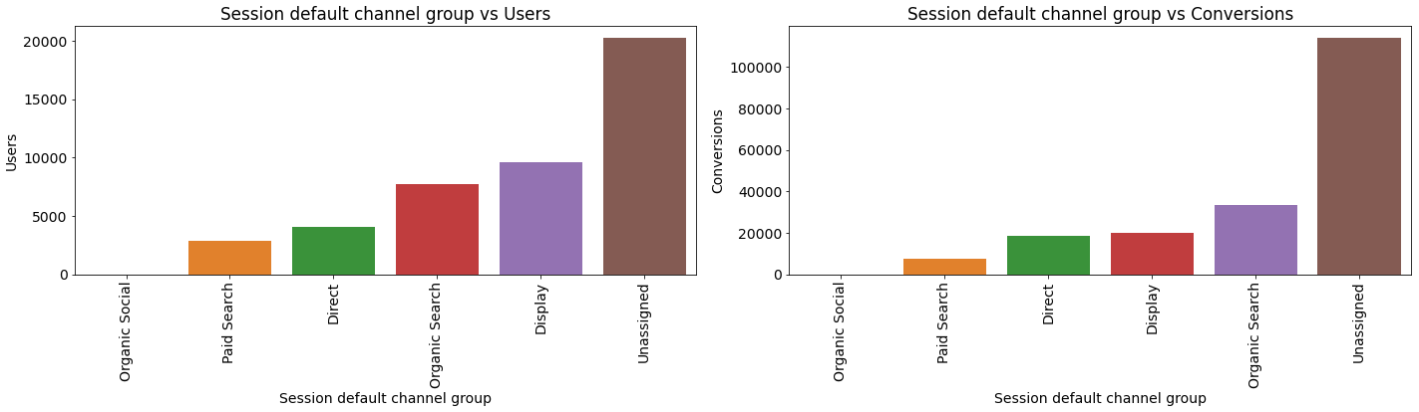


from the data, the Organic search is play a vital role for a conversion.

Traffic Acquisition

In [8]:
```
plt.rcParams.update({'font.size': 14})

fig, axs = plt.subplots(ncols=2, figsize=(20, 6))

df2 = df2.sort_values('Users')
sns.barplot(x='Session default channel group', y='Users',data=df2, ax=axs[0])
axs[0].set_title('Session default channel group vs Users')
axs[0].set_xlabel('Session default channel group')
axs[0].set_ylabel('Users')
axs[0].tick_params(axis='x', rotation=90)
```

```
df2 = df2.sort_values('Conversions')
sns.barplot(x='Session default channel group', y='Conversions', data=df2, ax=axs[1])
axs[1].set_title('Session default channel group vs Conversions')
axs[1].set_xlabel('Session default channel group')
axs[1].set_ylabel('Conversions')
axs[1].tick_params(axis='x', rotation=90)

plt.tight_layout()
plt.show()
```



"Here, the Unassigned category beats organic search when it comes to conversions."

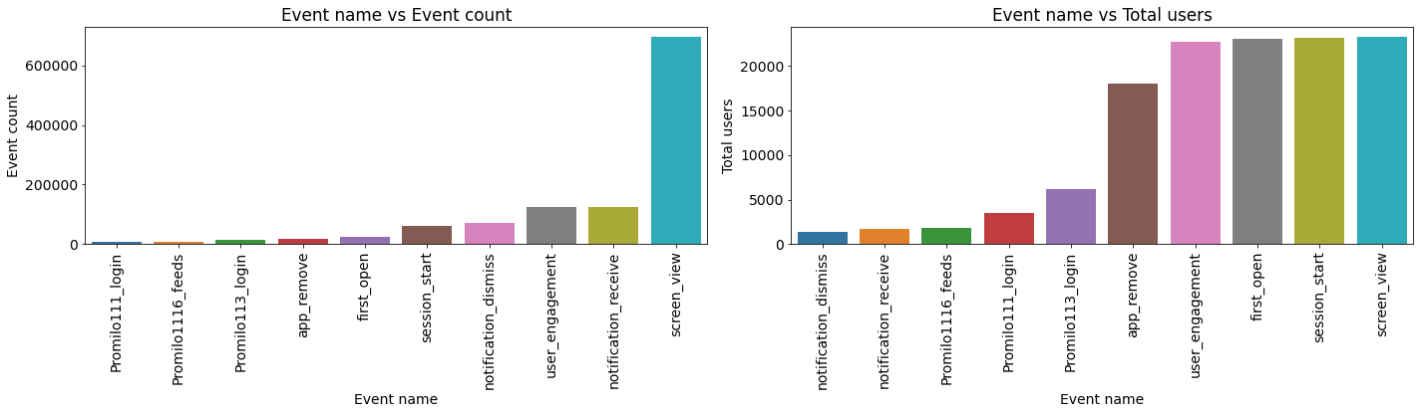
"Paid search should be improved for the conversion rate."

"The conversion rate is low in paid search."

Event Report

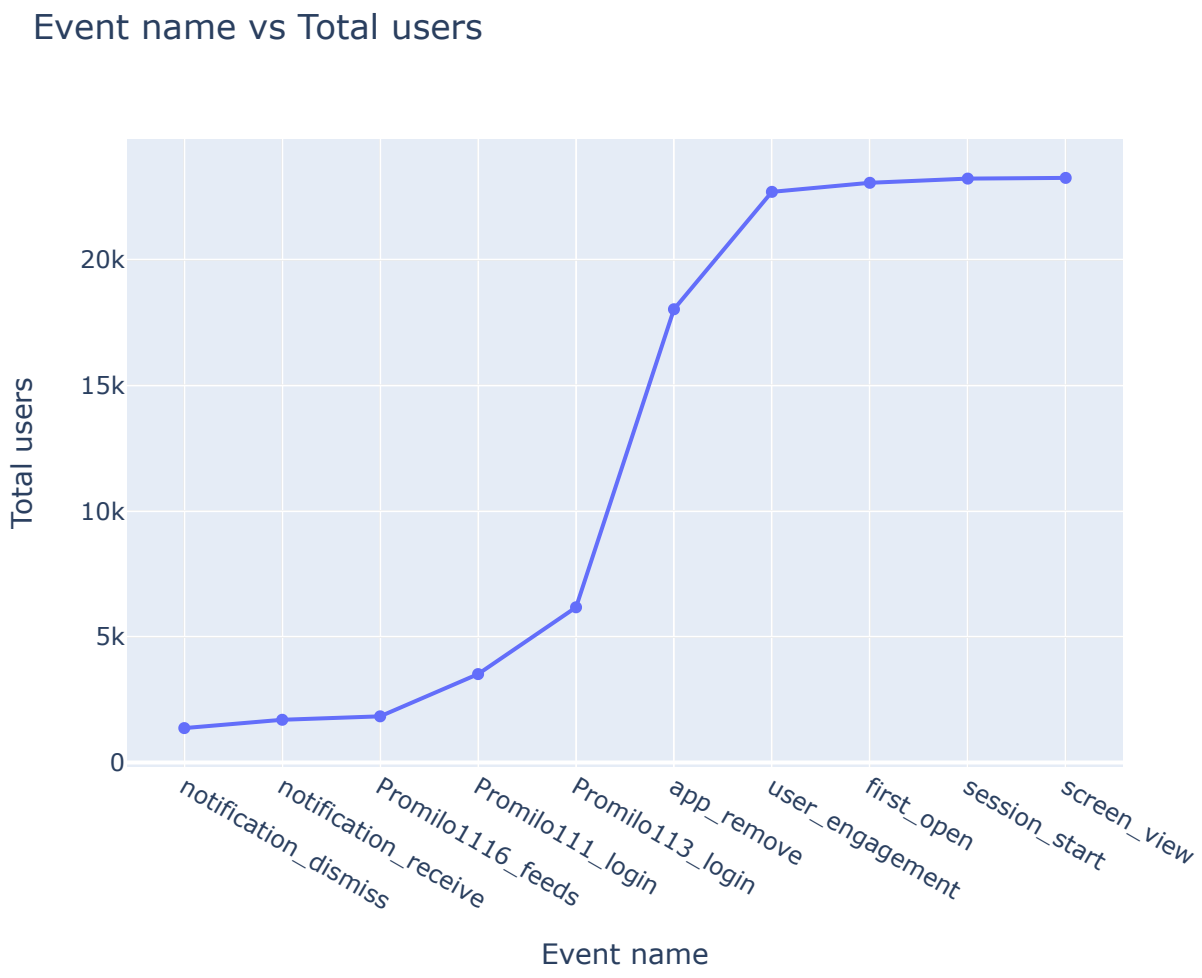
In [9]:

```
top_events = df3.nlargest(10, 'Event count')['Event name']

fig, axs = plt.subplots(ncols=2, figsize=(20, 6))

df_top = df3[df3['Event name'].isin(top_events)].sort_values('Event count')
sns.barplot(x='Event name', y='Event count', data=df_top, ax=axs[0])
axs[0].set_title('Event name vs Event count')
axs[0].set_xlabel('Event name')
axs[0].set_ylabel('Event count')
axs[0].tick_params(axis='x', rotation=90)

df_top = df3[df3['Event name'].isin(top_events)].sort_values('Total users')
sns.barplot(x='Event name', y='Total users', data=df_top, ax=axs[1])
axs[1].set_title('Event name vs Total users')
axs[1].set_xlabel('Event name')
axs[1].set_ylabel('Total users')
axs[1].tick_params(axis='x', rotation=90)

plt.tight_layout()
plt.show()
```



In [10]:

```
import plotly.express as px

# For Better Understanding... (Event name vs Total users)
df_top = df3[df3['Event name'].isin(top_events)].sort_values('Total users')
fig1 = px.line(df_top, x='Event name', y='Total users', title='Event name vs Total users')
fig1.update_traces(mode='markers+lines')
fig1.show()
```



When it comes to Promilo, the user's time spent on the app is minimal, and screen views predominantly take precedence, with the user login rate also being very low.

Conversion Report

In [11]:

```
df4 = dfs['Conversion Report']
```

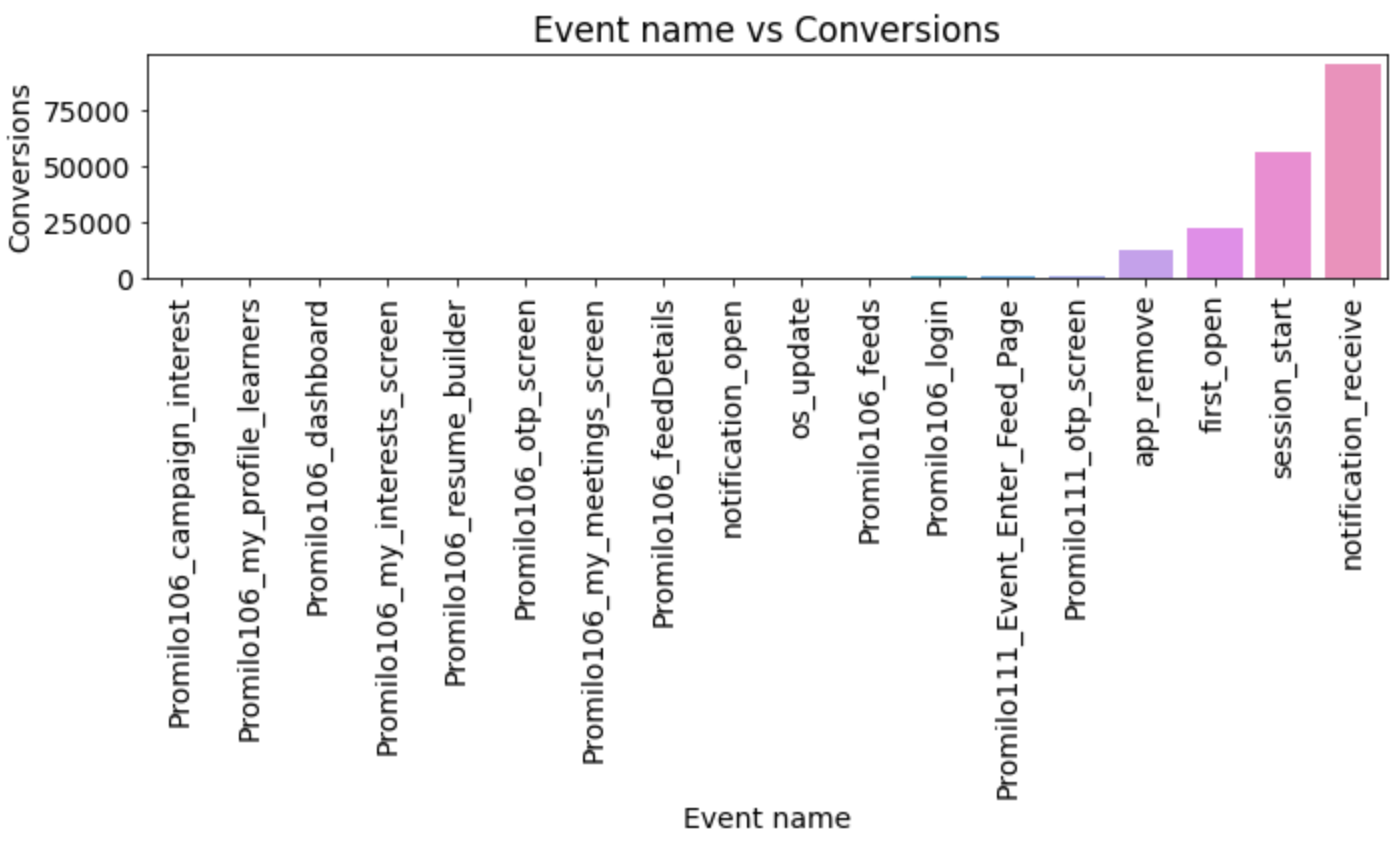
```
df4.head()
```

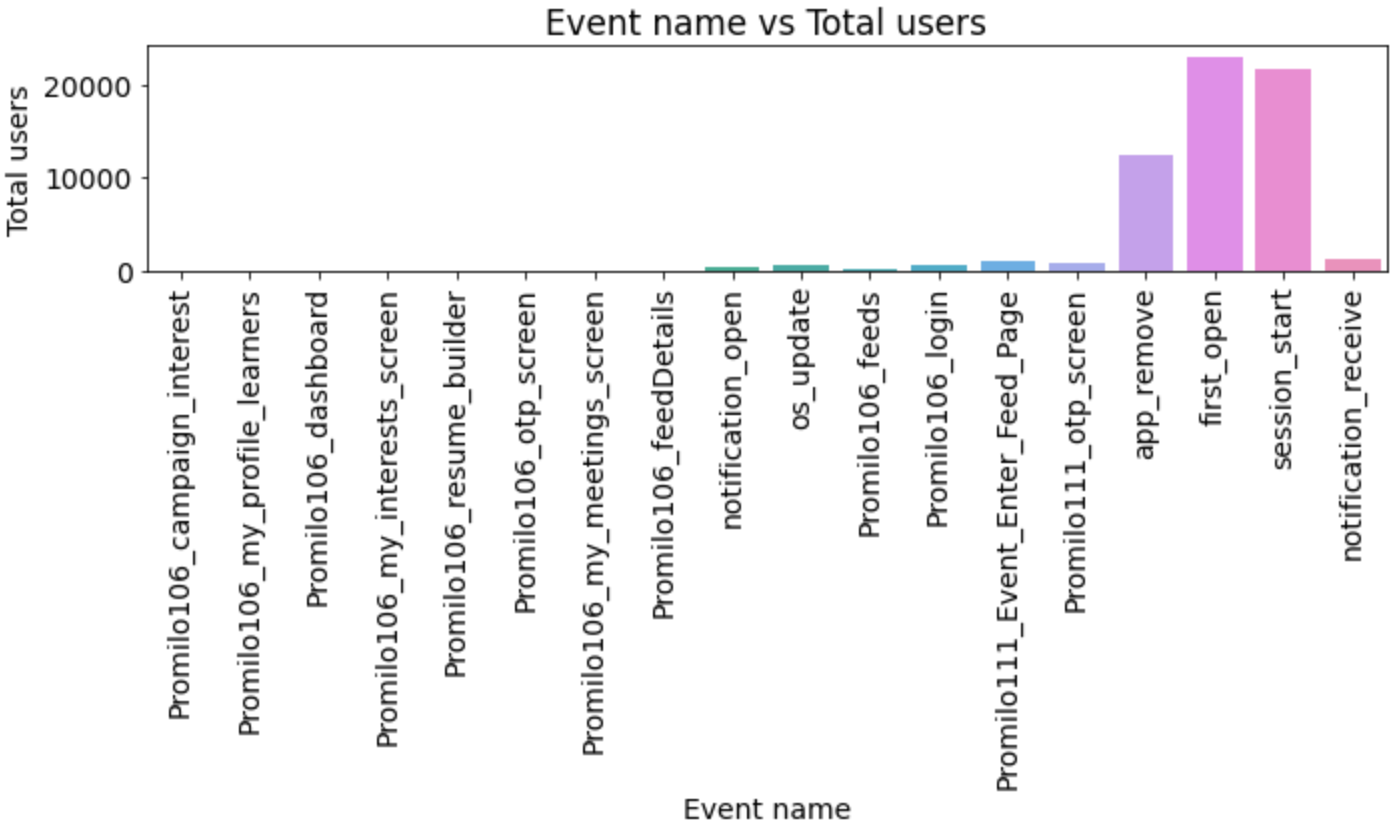
Out[11]:

	Event name	Conversions	Total users	Total revenue
0	notification_receive	94890	1311	0
1	session_start	56203	21674	0
2	first_open	22872	23059	0
3	app_remove	12468	12538	0
4	Promilo111_otp_screen	1738	855	0

In [12]:

```
fig, axs = plt.subplots(nrows=2, figsize=(10, 12))

# Event name Vs Conversion...
df4_sorted = df4.sort_values('Conversions')
sns.barplot(x='Event name', y='Conversions', data=df4_sorted, ax=axs[0])
axs[0].set_title('Event name vs Conversions')
axs[0].set_xlabel('Event name')
axs[0].set_ylabel('Conversions')
axs[0].tick_params(axis='x', rotation=90)

# Event name vs Total Users...
sns.barplot(x='Event name', y='Total users', data=df4_sorted, ax=axs[1])
axs[1].set_title('Event name vs Total users')
axs[1].set_xlabel('Event name')
axs[1].set_ylabel('Total users')
axs[1].tick_params(axis='x', rotation=90)

plt.tight_layout()
plt.show()
```

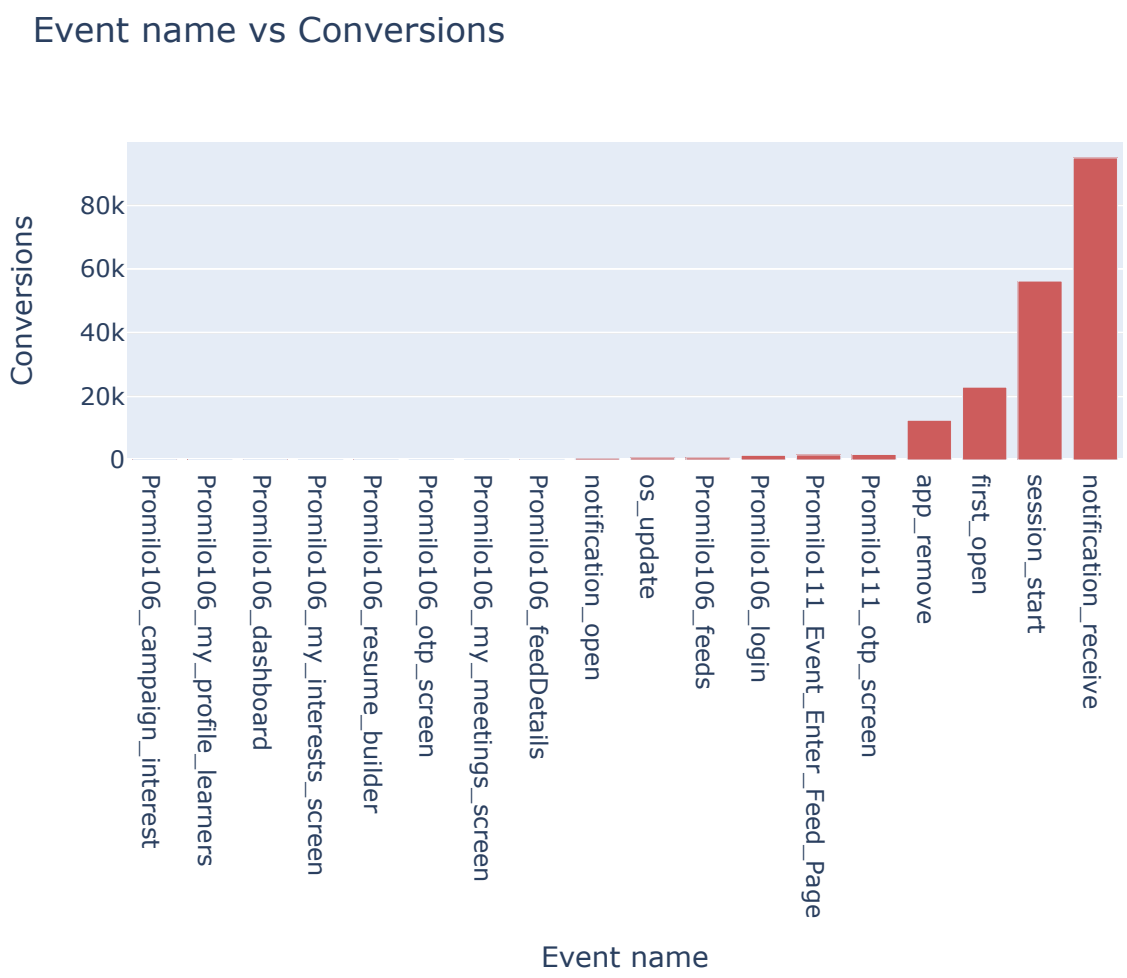


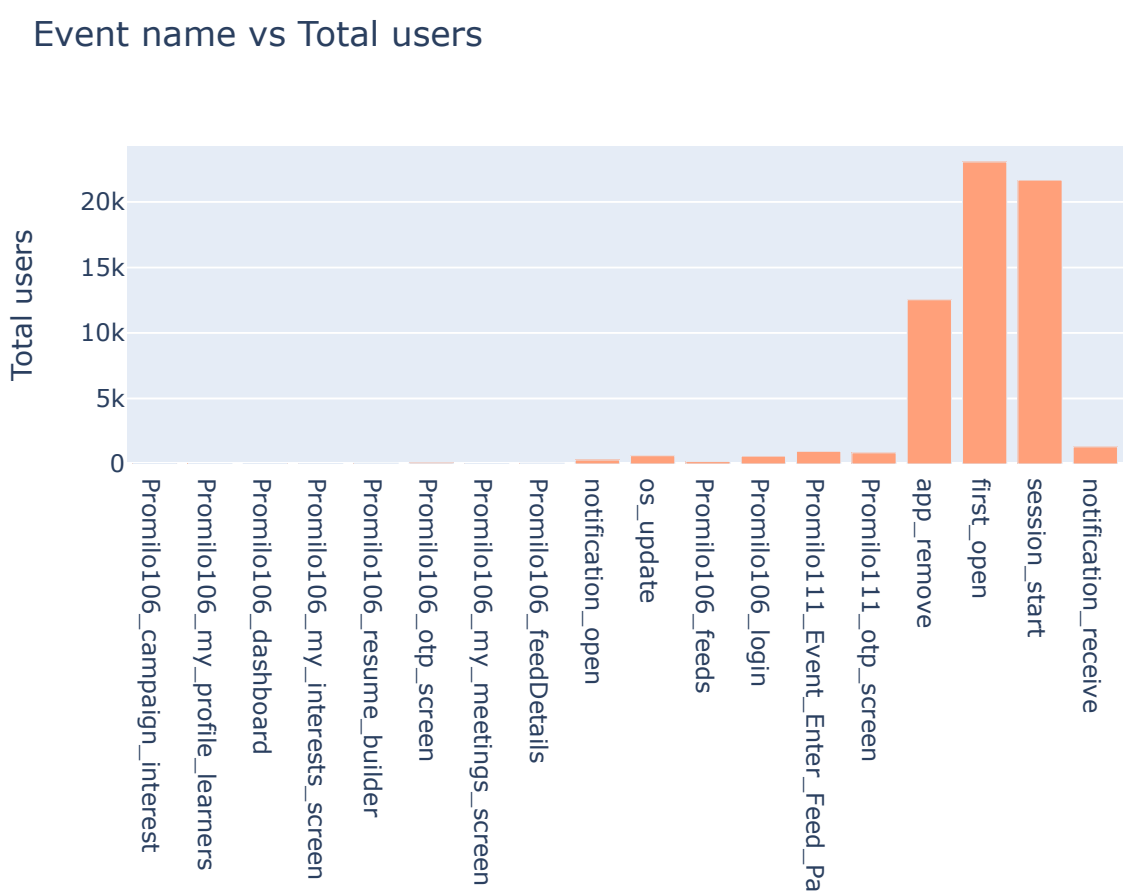


In [13]:

```
import plotly.express as px

# Event name Vs Conversions...
df4_sorted = df4.sort_values('Conversions')
fig2 = px.bar(df4_sorted, x='Event name', y='Conversions', title='Event name vs Conversion
fig2.update_traces(marker_color='indianred')

# Event name vs Total Users...
fig3 = px.bar(df4_sorted, x='Event name', y='Total users', title='Event name vs Total user
fig3.update_traces(marker_color='lightsalmon')
```

```
fig2.show()
fig3.show()
```





Event name

In the Conversion Vs Event Name chart, notifications play a vital role in conversions. Perhaps, when a customer sees a notification, it triggers the user to engage with it.

Page path and screen class

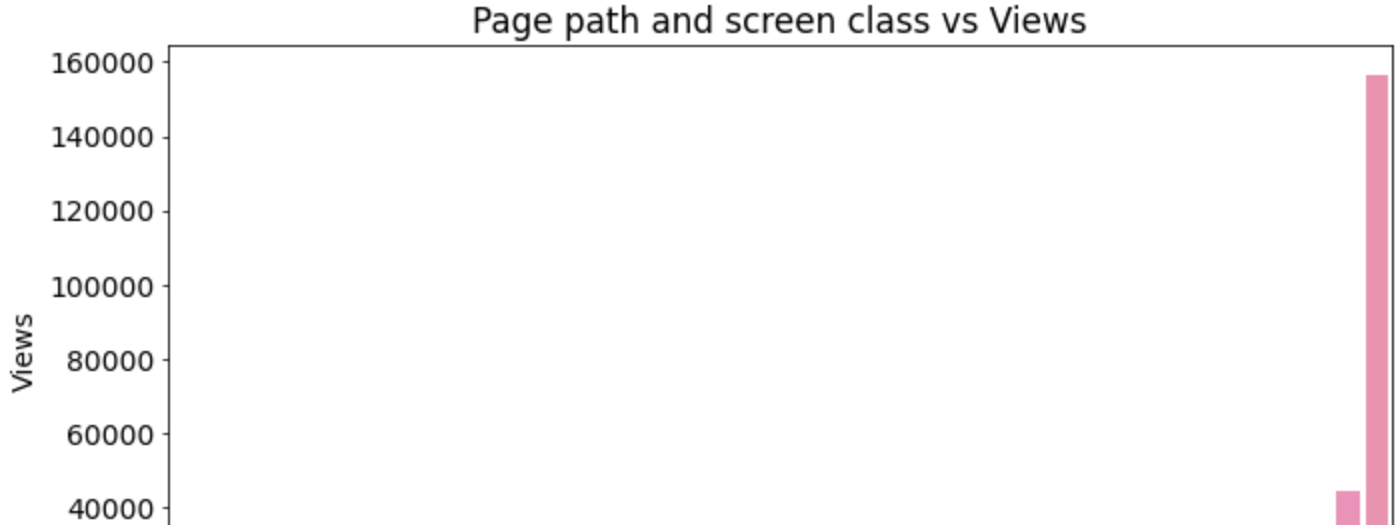
In [14]:
```
print(df5.columns)
df5.head(2)
```

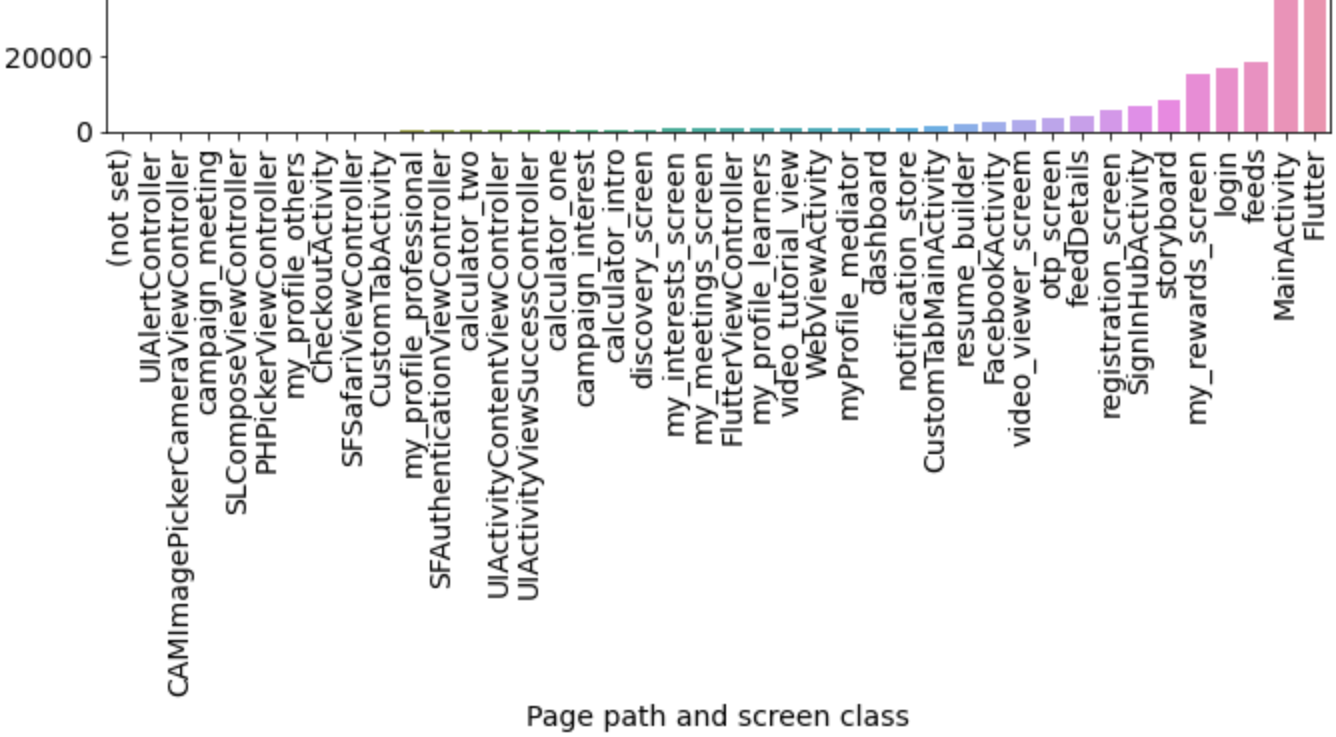
```
Index(['Page path and screen class', 'Views', 'Users', 'Views per user',
       'Average engagement time', 'Event count', 'Conversions',
       'Total revenue'],
      dtype='object')
```

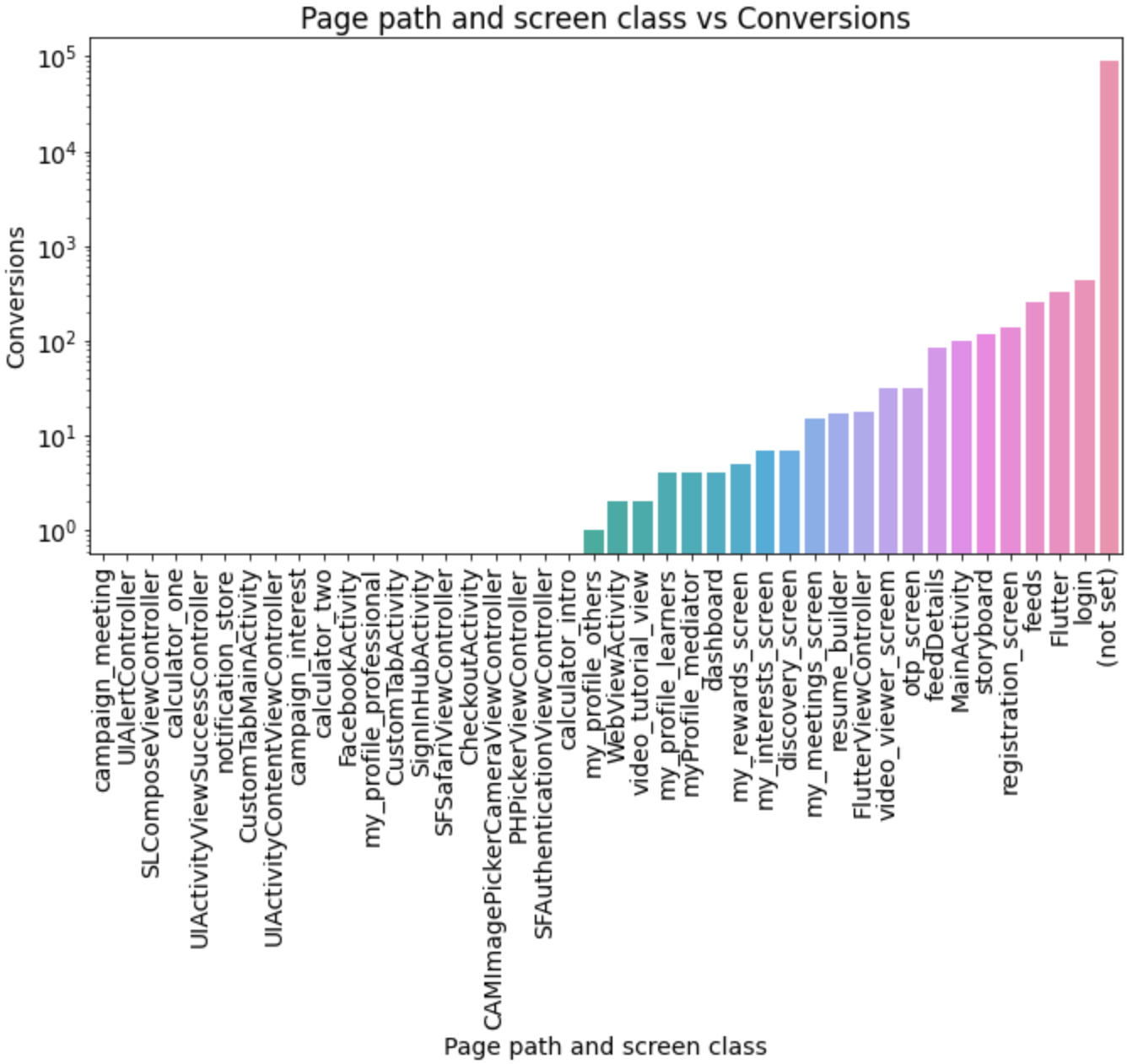
Out[14]:

	Page path and screen class	Views	Users	Views per user	Average engagement time	Event count	Conversions	Total revenue
0	Flutter	156708	8726	17.95874	83.41222	203901	328	0
1	MainActivity	44326	8978	4.93718	78.29216	53374	101	0

In [15]:
```
fig, axs = plt.subplots(nrows=2, figsize=(10, 18))

df5_sorted = df5.sort_values('Views')
sns.barplot(x='Page path and screen class', y='Views', data=df5_sorted, ax=axs[0])
axs[0].set_title('Page path and screen class vs Views')
axs[0].set_xlabel('Page path and screen class')
axs[0].set_ylabel('Views')
axs[0].tick_params(axis='x', rotation=90)

df5_sorted = df5.sort_values('Conversions')
sns.barplot(x='Page path and screen class', y='Conversions', data=df5_sorted, ax=axs[1])
axs[1].set_title('Page path and screen class vs Conversions')
axs[1].set_xlabel('Page path and screen class')
axs[1].set_ylabel('Conversions')
axs[1].tick_params(axis='x', rotation=90)

# Set y-axis to logarithmic scale
axs[1].set_yscale('log')

plt.tight_layout()
plt.show()
```







In these pages, such as the main activity, Flutter, login, and storyboard, are highly viewed by the users.

Demographics Report

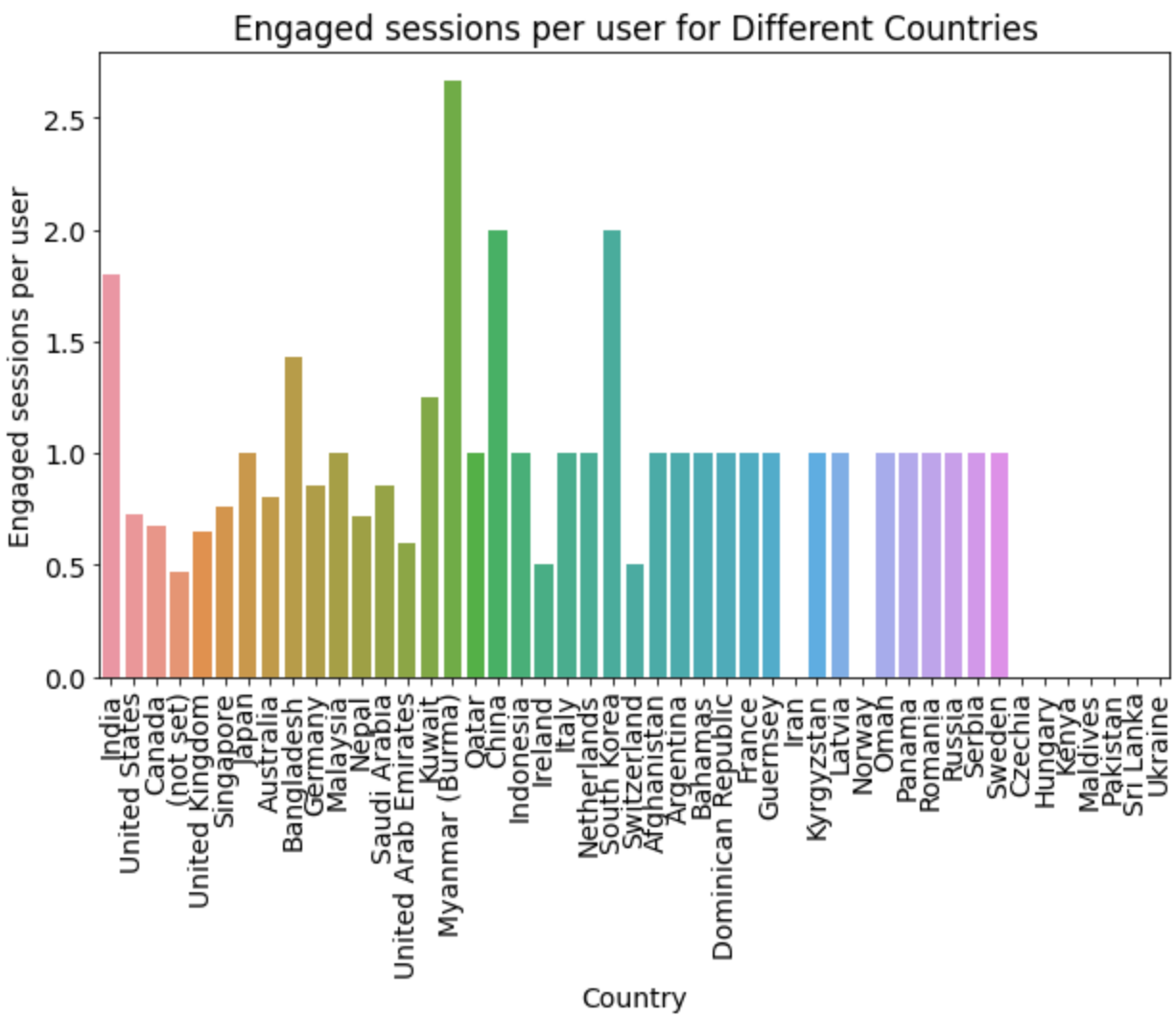
In [16]:
```
df6 = dfs['Demographics Report']
print(df6.columns)
df6.head()
```

```
Index(['Country', 'Users', 'New users', 'Engaged sessions', 'Engagement rate',
       'Engaged sessions per user', 'Average engagement time', 'Event count',
       'Conversions', 'Total revenue'],
      dtype='object')
```

Out[16]:

	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

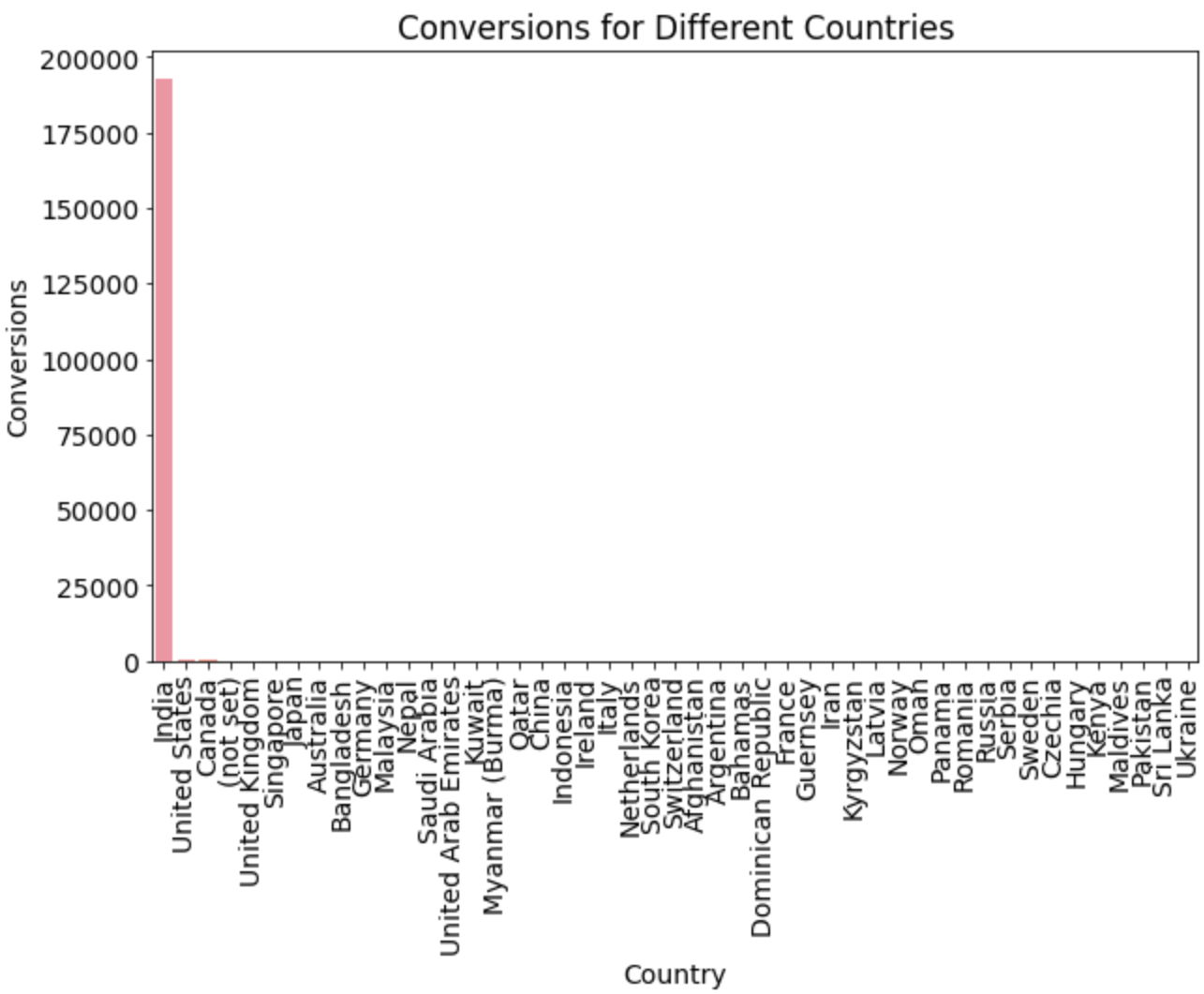
In [17]:
```
df6 = dfs['Demographics Report']

plt.figure(figsize=(10, 6))
sns.barplot(x=df6['Country'], y=df6['Engaged sessions per user'])
plt.title('Engaged sessions per user for Different Countries')
plt.xlabel('Country')
plt.ylabel('Engaged sessions per user')
plt.xticks(rotation=90)
plt.show()
```



The Users from the Myanmar(Burma) are more engaging in this session...

In [18]:

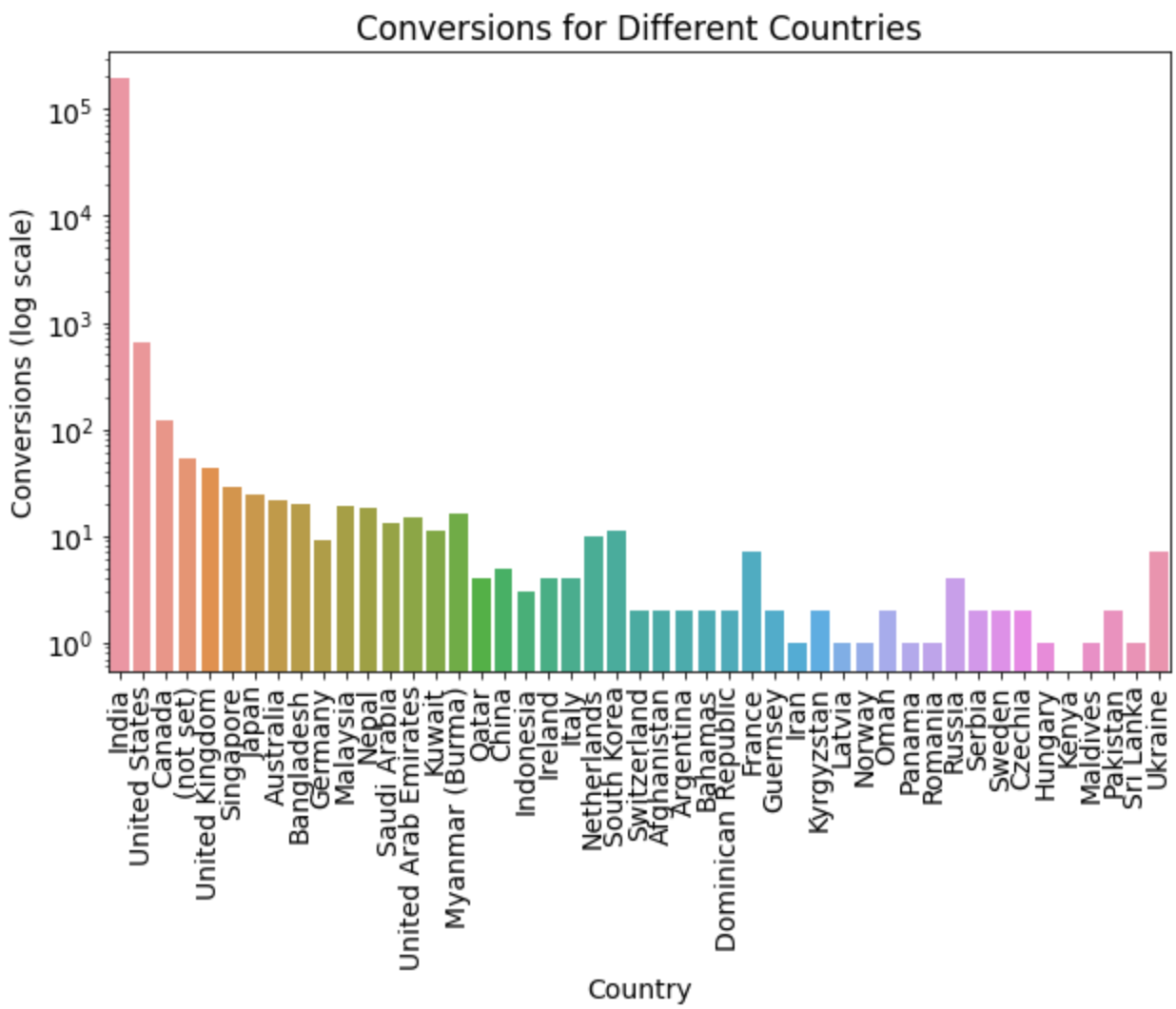
```
plt.figure(figsize=(10, 6))
sns.barplot(x=df6['Country'], y=df6['Conversions'])
plt.title('Conversions for Different Countries')
plt.xlabel('Country')
plt.ylabel('Conversions')
plt.xticks(rotation=90)
plt.show()
```



India has a higher conversion rate

In [19]:

```
plt.figure(figsize=(10, 6))
sns.barplot(x=df6['Country'], y=df6['Conversions'])
plt.title('Conversions for Different Countries')
plt.xlabel('Country')
plt.ylabel('Conversions (log scale)')
plt.yscale('log')  # Set the y-axis to log scale
plt.xticks(rotation=90)
plt.show()
```



when it ccomes to contry, india play a vital in the conversion.

Citywise Report

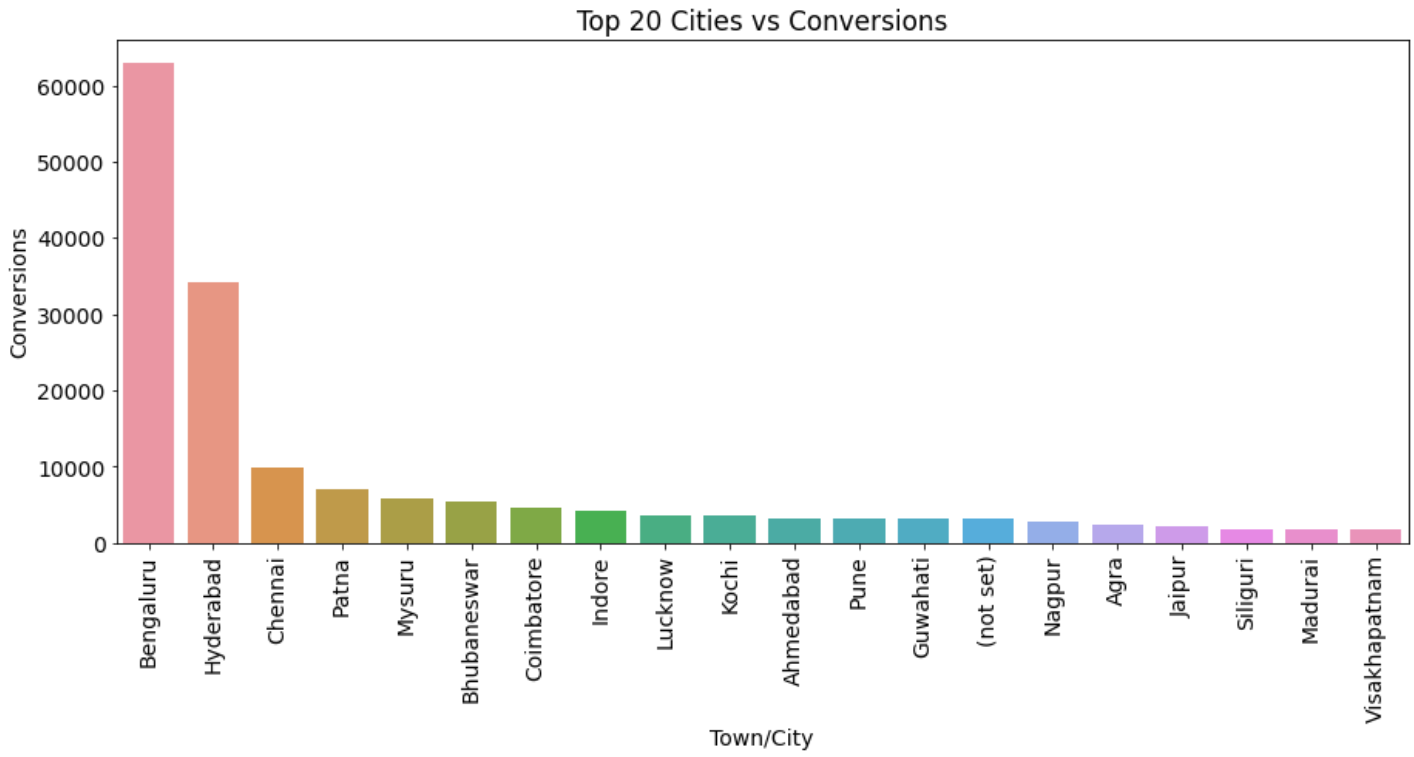
In [20]:
```
df7 = dfs['Citiwise Report']
df7.head()
```

Out[20]:

	Town/City	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	Bengaluru	6097	5685	15013	0.769385	2.462359	762.20550	607200	62939	0
1	Patna	1594	1467	2127	0.440646	1.334379	98.22208	38830	6980	0
2	Hyderabad	1038	920	1578	0.569264	1.520231	243.69080	96826	34103	0
3	Indore	983	915	1241	0.426460	1.262462	67.89115	21383	4121	0
4	Lucknow	897	839	1125	0.450180	1.254181	83.40580	21041	3650	0

In [21]:
```
top_cities = df7.nlargest(20, 'Conversions')

plt.figure(figsize=(15, 6))
sns.barplot(x='Town/City', y='Conversions', data=top_cities)
plt.title('Top 20 Cities vs Conversions')
plt.xlabel('Town/City')
plt.ylabel('Conversions')
```

```
plt.xticks(rotation=90)  # Rotate x-axis labels by 90 degrees
plt.show()
```



In citiwise, the bengaluru has the highest conversion rate, second highest is hyderabad.

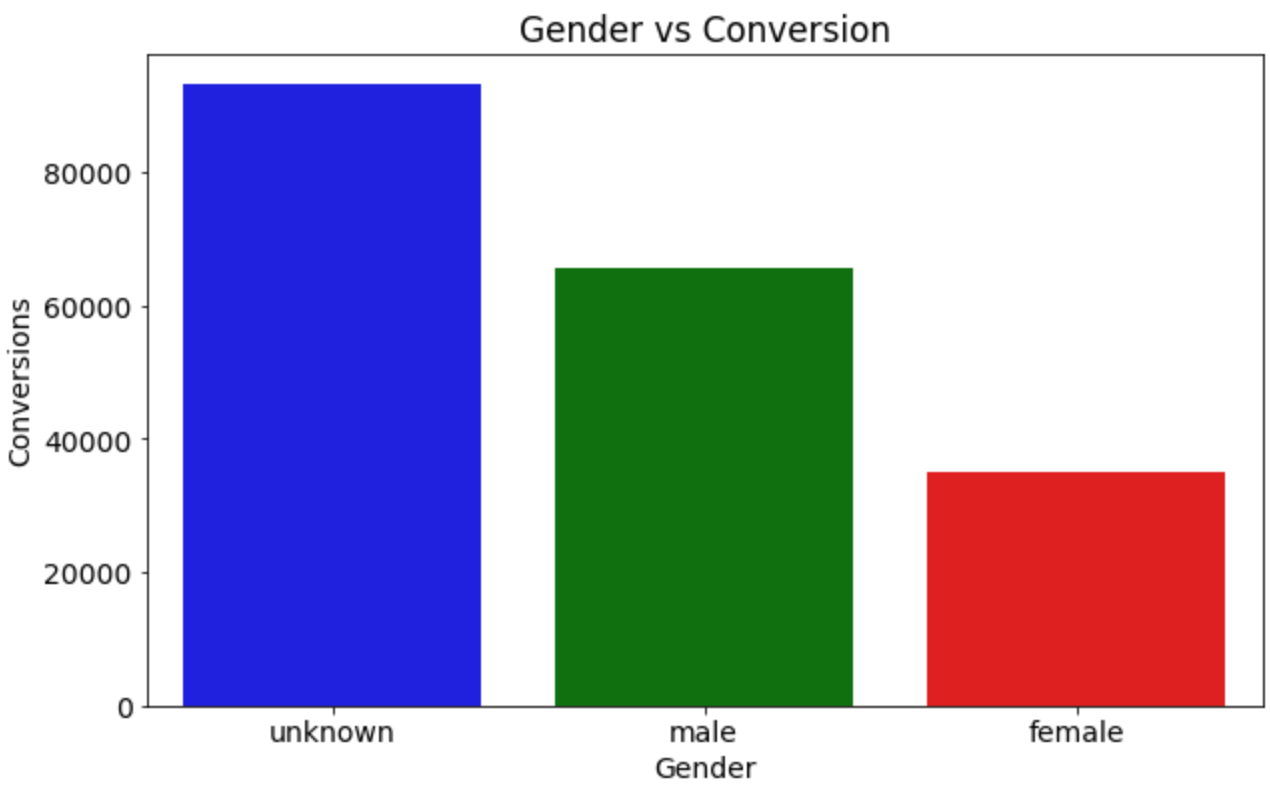
Gender Report

In [22]:
```
df8 = dfs['Gender Report']
df8.head()
```

Out[22]:

	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 [23]:
```
genders = df8['Gender']
conversions = df8['Conversions']

plt.figure(figsize=(10,6))
sns.barplot(x=genders, y=conversions, palette=['blue', 'green', 'red'])
plt.title('Gender vs Conversion')
plt.xlabel('Gender')
plt.ylabel('Conversions')
plt.show()
```



Users hide their gender when logging in, so it may be considered as unknown. Apart from that, males have a higher conversion rate.

User By Instrest

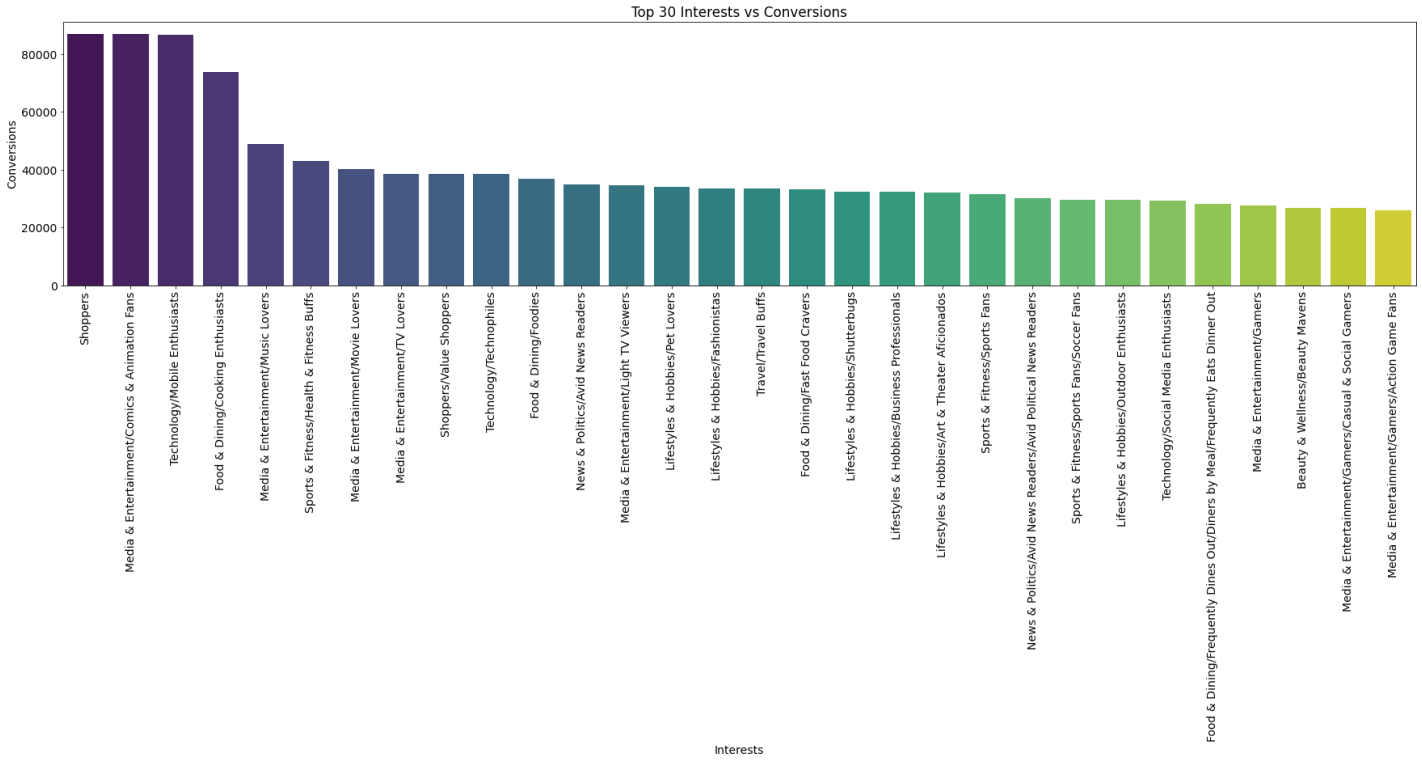
In [24]:
```
df9 = dfs['User By Interest']

df9.head()
```

Out[24]:

	Interests	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	Shoppers	10950	9256	15652	0.581534	1.429406	162.8347	490664	86846	0
1	Media & Entertainment/Comics & Animation Fans	10946	9247	15680	0.583008	1.432487	165.1772	491025	86845	0
2	Technology/Mobile Enthusiasts	10934	9239	15619	0.582451	1.428480	162.6945	489353	86742	0
3	Food & Dining/Cooking Enthusiasts	8410	6970	12332	0.602325	1.466350	176.9567	409713	73814	0
4	Sports & Fitness/Health & Fitness Buffs	5844	4580	8226	0.588328	1.407598	155.1451	257831	43074	0

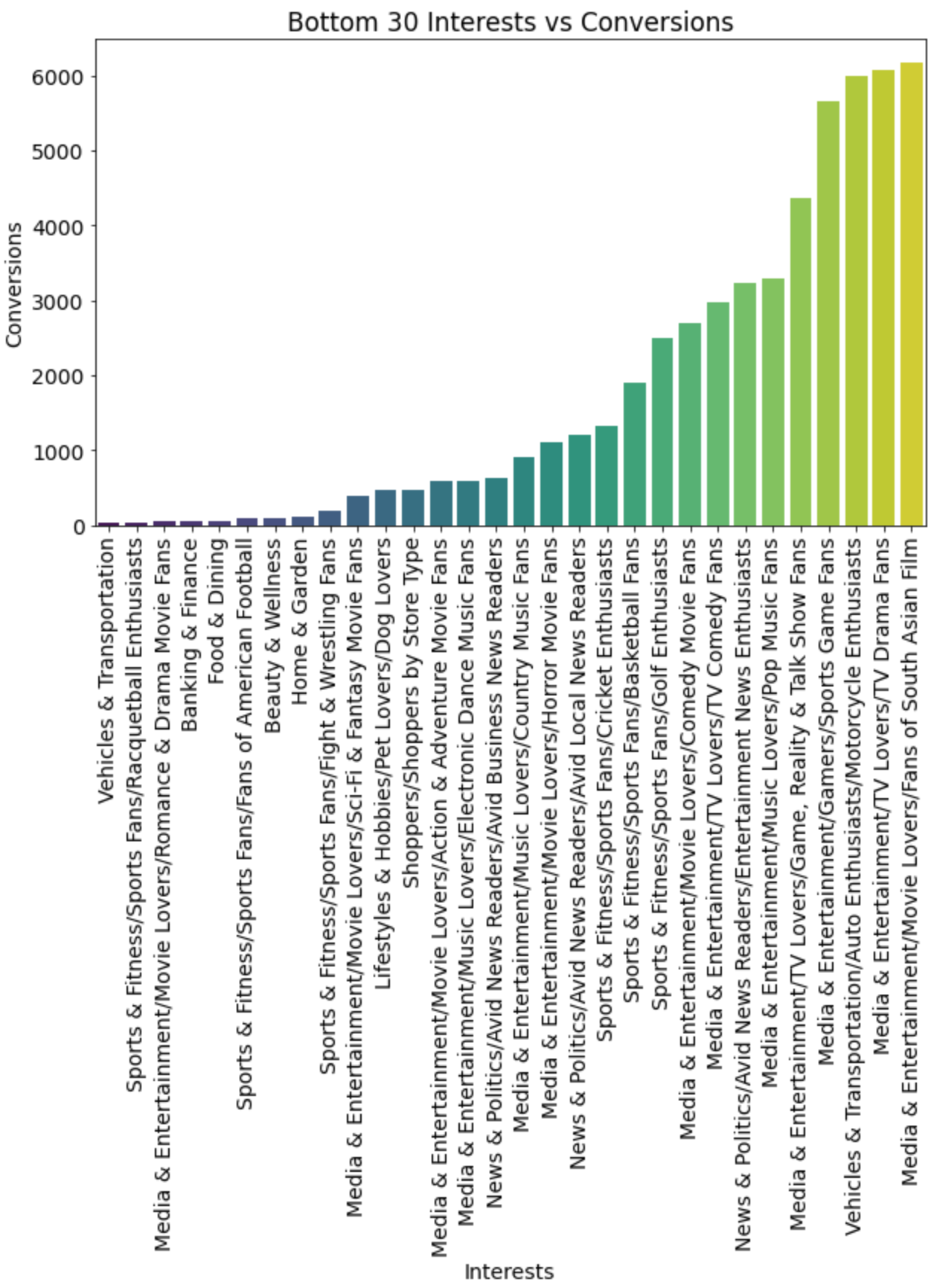
In [25]:
```
df_top30 = df9.sort_values('Conversions', ascending=False).head(30)

interests = df_top30['Interests']
conversions = df_top30['Conversions']

plt.figure(figsize=(30,6))
sns.barplot(x=interests, y=conversions, palette='viridis')
```

```
plt.title('Top 30 Interests vs Conversions')
plt.xlabel('Interests')
plt.ylabel('Conversions')
plt.xticks(rotation=90)  # Rotate x-axis labels by 90 degrees
plt.show()
```



In [26]:

```
# Sort by 'Conversions' and take bottom 30
df_bottom30 = df9.sort_values('Conversions', ascending=True).head(30)

interests = df_bottom30['Interests']
conversions = df_bottom30['Conversions']

plt.figure(figsize=(10,6))
sns.barplot(x=interests, y=conversions, palette='viridis')
plt.title('Bottom 30 Interests vs Conversions')
plt.xlabel('Interests')
plt.ylabel('Conversions')
plt.xticks(rotation=90)  # Rotate x-axis labels by 90 degrees
plt.show()
```



from the top & bottom analysis of User's Intrest, User highly focusing the shopping, media entertainment and technology and mobile enthusiasts.

and users are not focusing on vehicles, fitness, & banking category

User by Language.

In [27]:
```
df10 = dfs['User by Language']
```

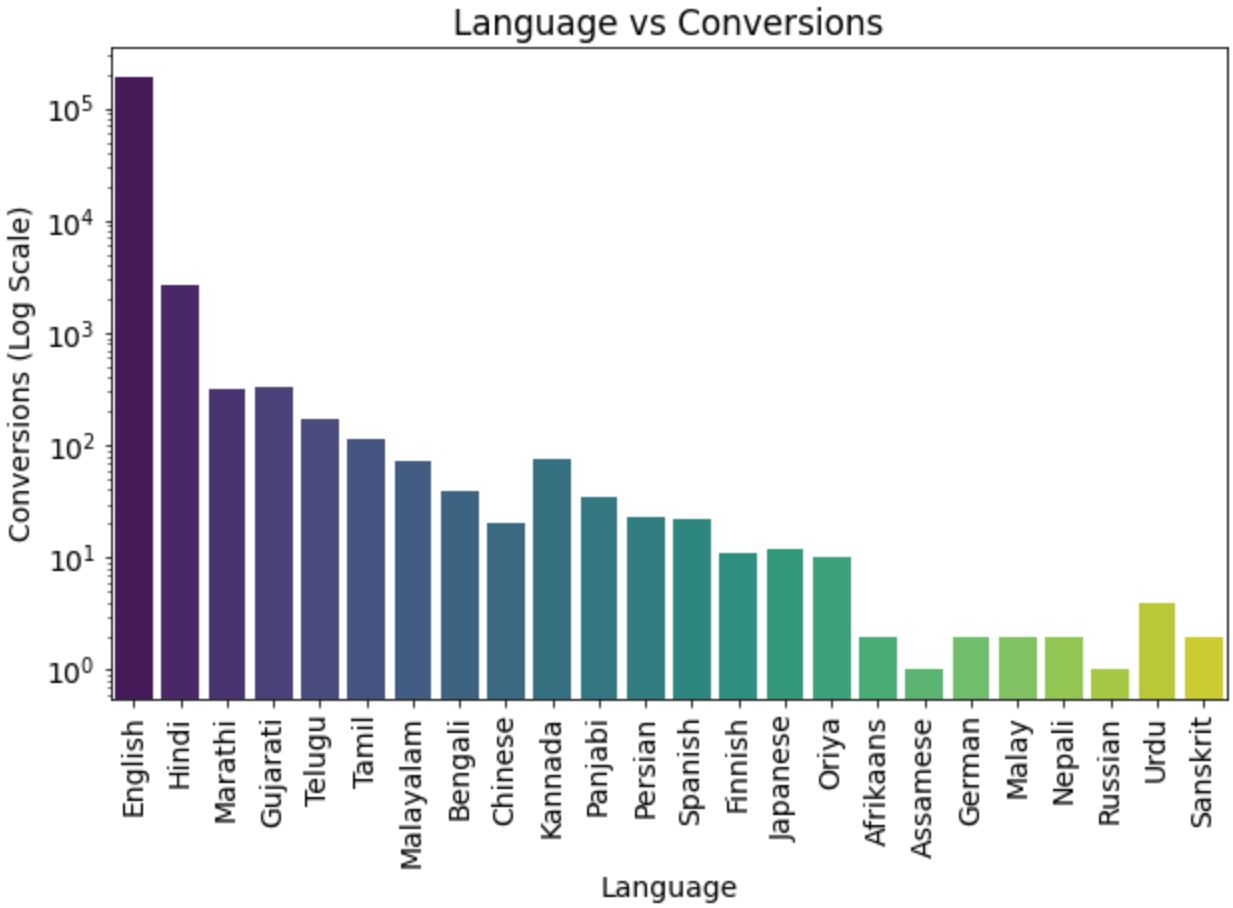
```
df10.head(3)
```

Out[27]:

	Language	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	English	22495	21990	40639	0.595147	1.806579	341.36350	1297970	189946	0
1	Hindi	586	552	798	0.406314	1.361775	60.03413	13523	2699	0
2	Marathi	85	84	98	0.426087	1.152941	38.48235	1589	323	0

In [28]:

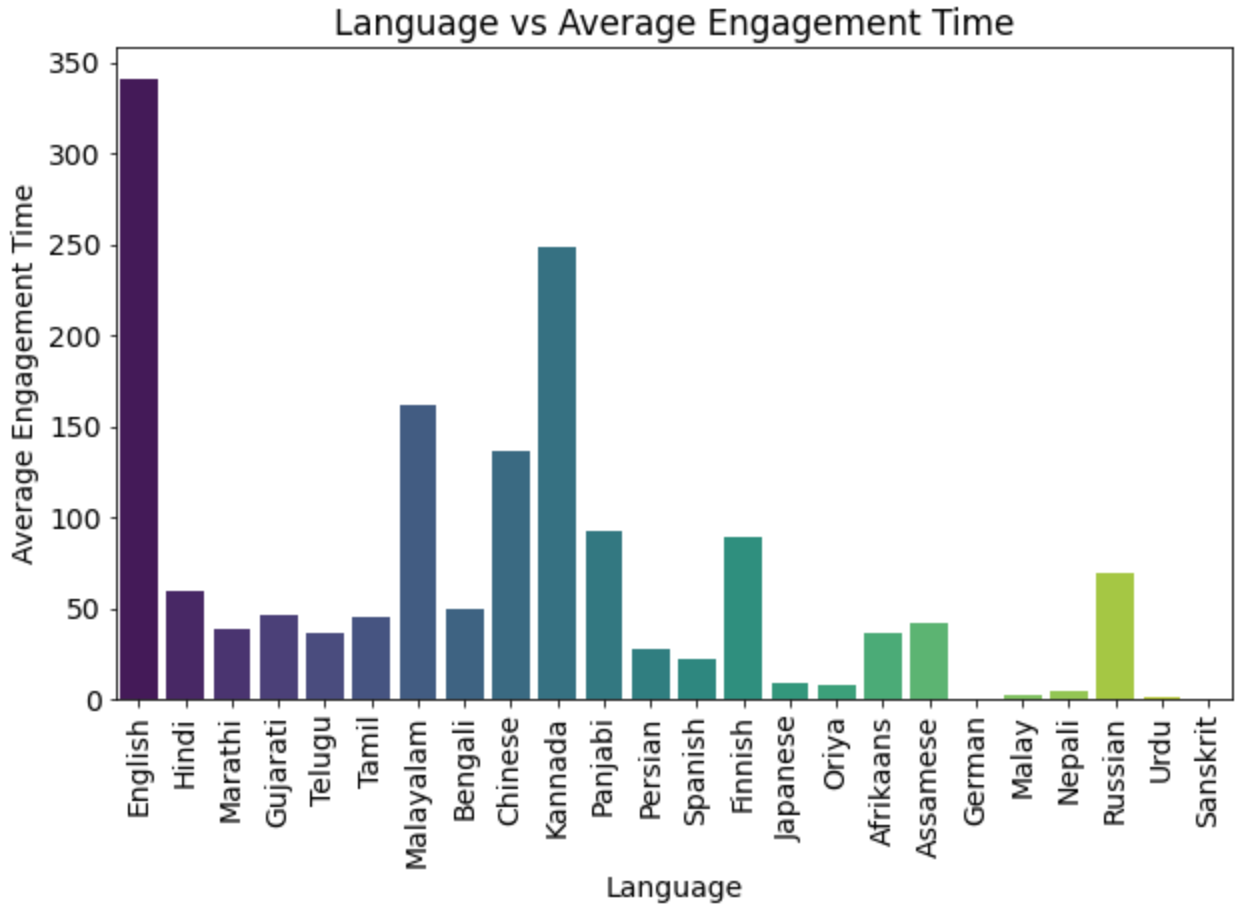
```
languages = df10['Language']
conversions = df10['Conversions']

plt.figure(figsize=(10,6))
sns.barplot(x=languages, y=conversions, palette='viridis')
plt.title('Language vs Conversions')
plt.xlabel('Language')
plt.ylabel('Conversions (Log Scale)')
plt.xticks(rotation=90)  # Rotate x-axis labels by 90 degrees

# Set y-axis to logarithmic scale
plt.yscale('log')

plt.show()
```



In [29]:

```
languages = df10['Language']
avg_engagement_time = df10['Average engagement time']

plt.figure(figsize=(10,6))
sns.barplot(x=languages, y=avg_engagement_time, palette='viridis')
plt.title('Language vs Average Engagement Time')
```

```
plt.xlabel('Language')
plt.ylabel('Average Engagement Time')
plt.xticks(rotation=90)  # Rotate x-axis labels by 90 degrees
plt.show()
```



English Language is play mojor role when it comes to conversion and engagement time.

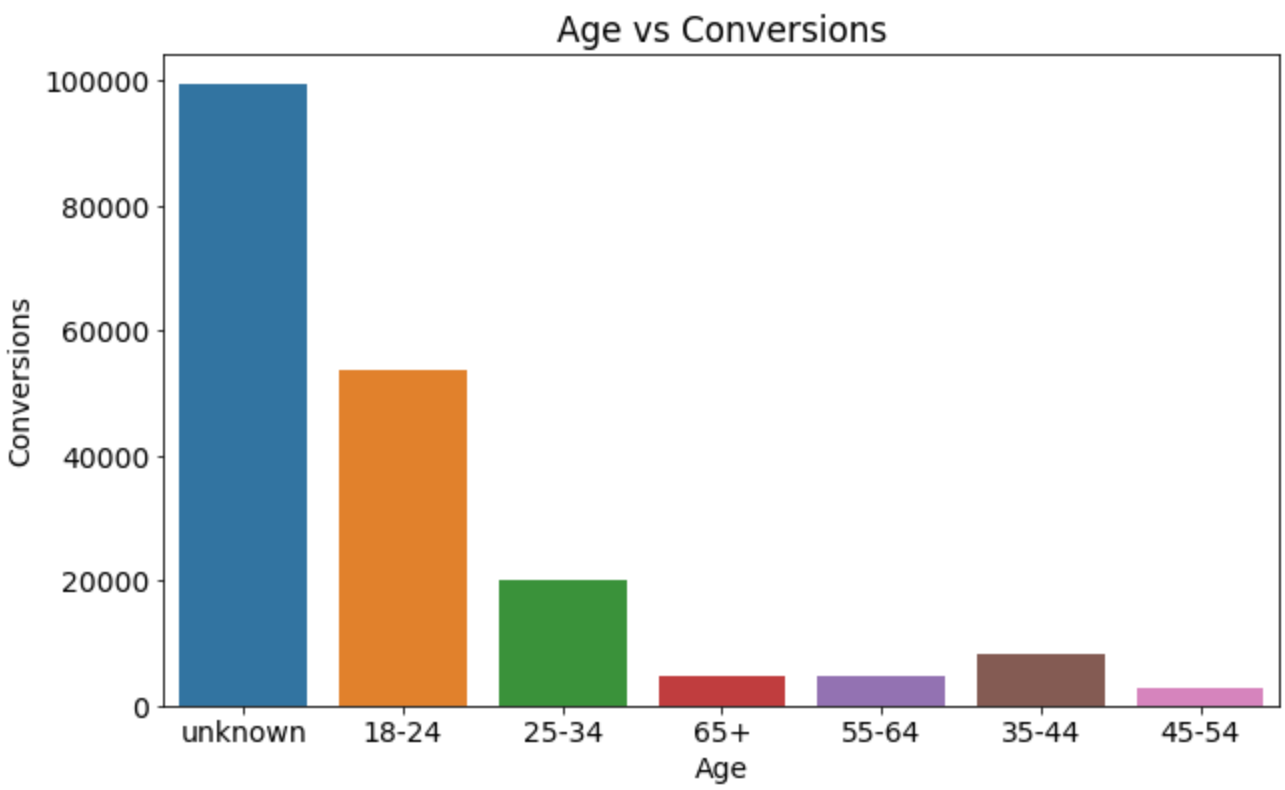
User by Age

In [30]:
```
df11 = dfs['User By Age']
df11.head(2)
```

Out[30]:

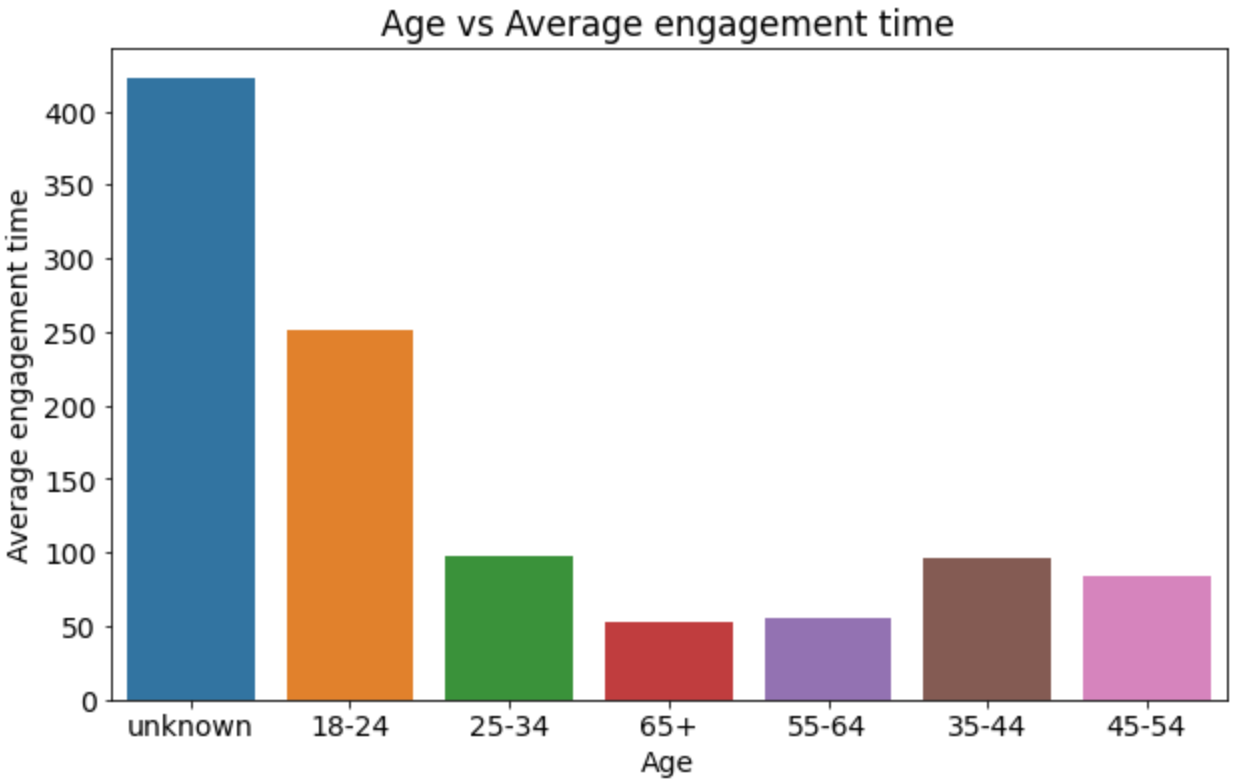
	Age	Users	New users	Engaged sessions	Engagement rate	Engaged sessions per user	Average engagement time	Event count	Conversions	Total revenue
0	unknown	14303	13636	24976	0.569098	1.746207	422.2233	817501	99310	0
1	18-24	4282	3678	7291	0.695308	1.702709	251.1630	309328	53661	0

In [31]:
```
plt.figure(figsize=(10, 6))
sns.barplot(x='Age', y='Conversions', data=df11)
plt.title('Age vs Conversions')
plt.show()
```



In [32]:

```
plt.figure(figsize=(10, 6))
sns.barplot(x='Age', y='Average engagement time', data=df11)
plt.title('Age vs Average engagement time')
plt.show()
```



people hide their age, while log in. so its consider as a unknown. The unknown category are in higher conversion.

18-24 age people are second highest in conversion

45-54 age people are less when comes to conversion

65+ age people are heving less engagement time

Google Ads Report

We need to more focus paid marketing campaign.

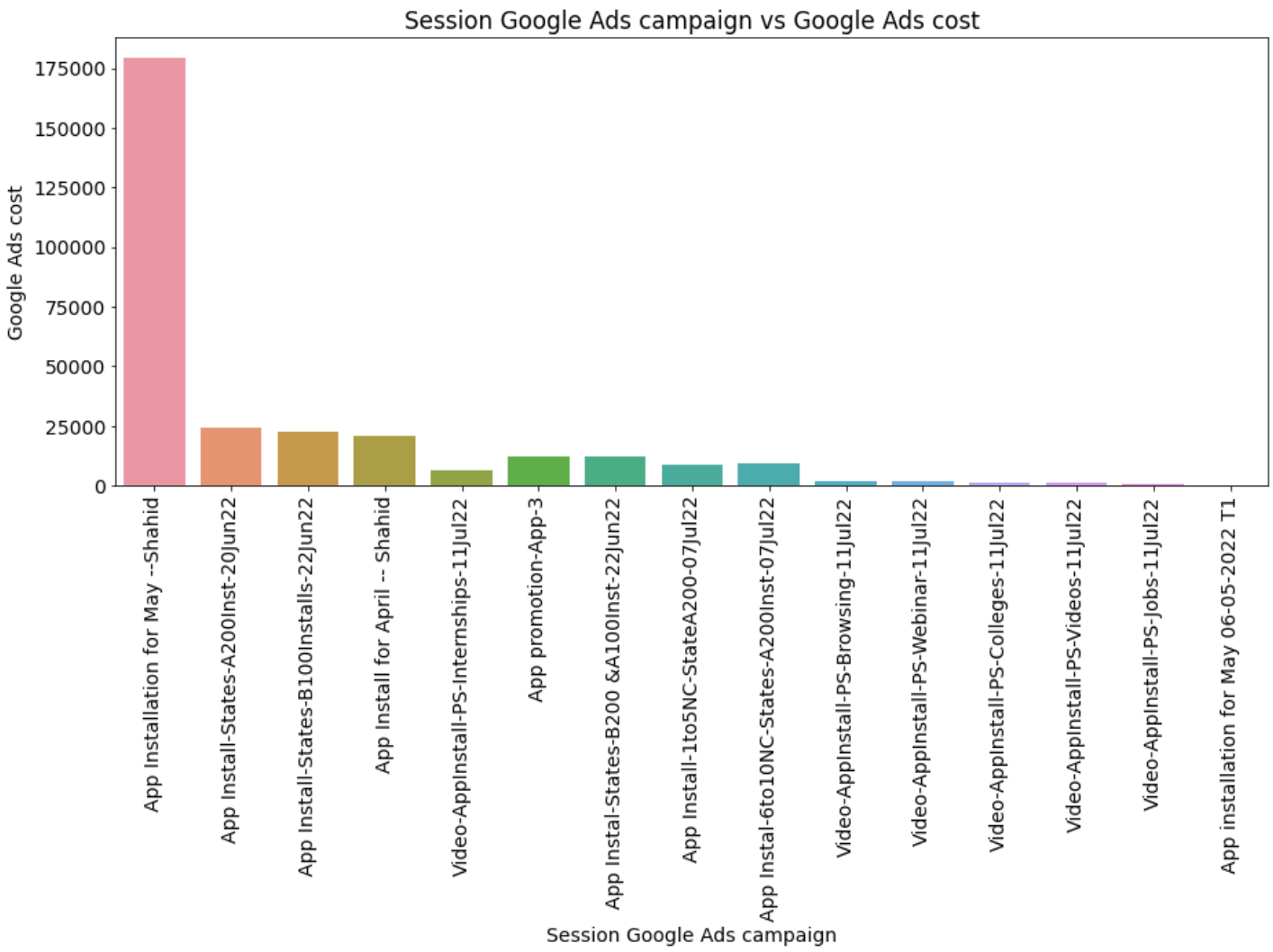
Let see what are the thing working and not working for a converison the Users

In [33]:
```
df12 = dfs['Google Ads Report']
df12.head(2)
```

Out[33]:

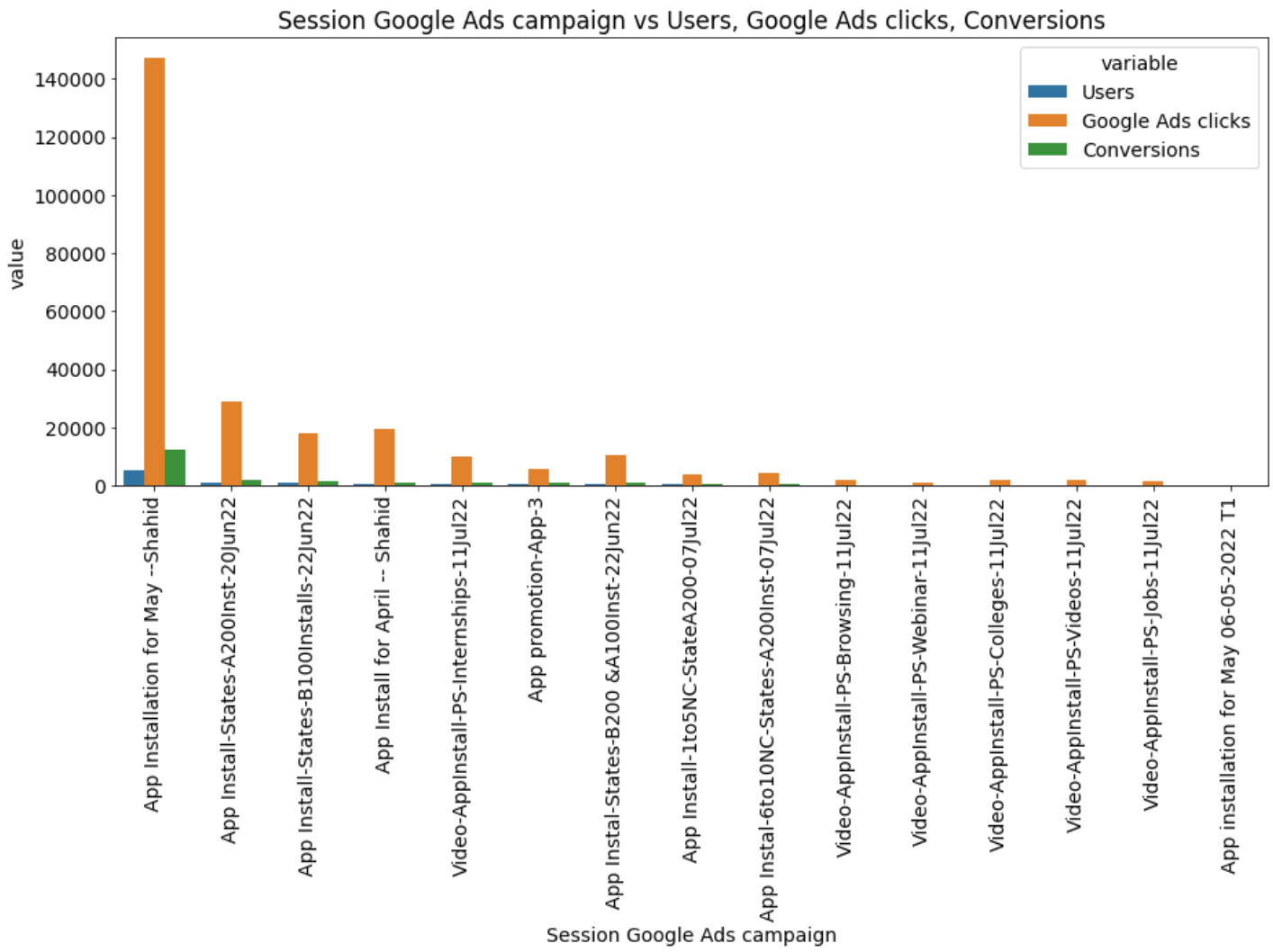
	Session Google Ads campaign	Users	Sessions	Engaged sessions	Google Ads clicks	Google Ads cost	Google Ads cost per click	Conversions	Cost per conversion	Event count	Total revenue	Re o sp
0	App Installation for May -- Shahid	5429	10936	6276	147100	179175.00	1.218049	12257	14.61818	97802	0	
1	App Install-States-A200Inst-20Jun22	842	1655	968	28742	24309.13	0.845770	1794	13.55024	15311	0	

In [34]:
```
plt.figure(figsize=(15, 6))
sns.barplot(x='Session Google Ads campaign', y='Google Ads cost', data=df12)
plt.title('Session Google Ads campaign vs Google Ads cost')
plt.xticks(rotation=90)
plt.show()
```



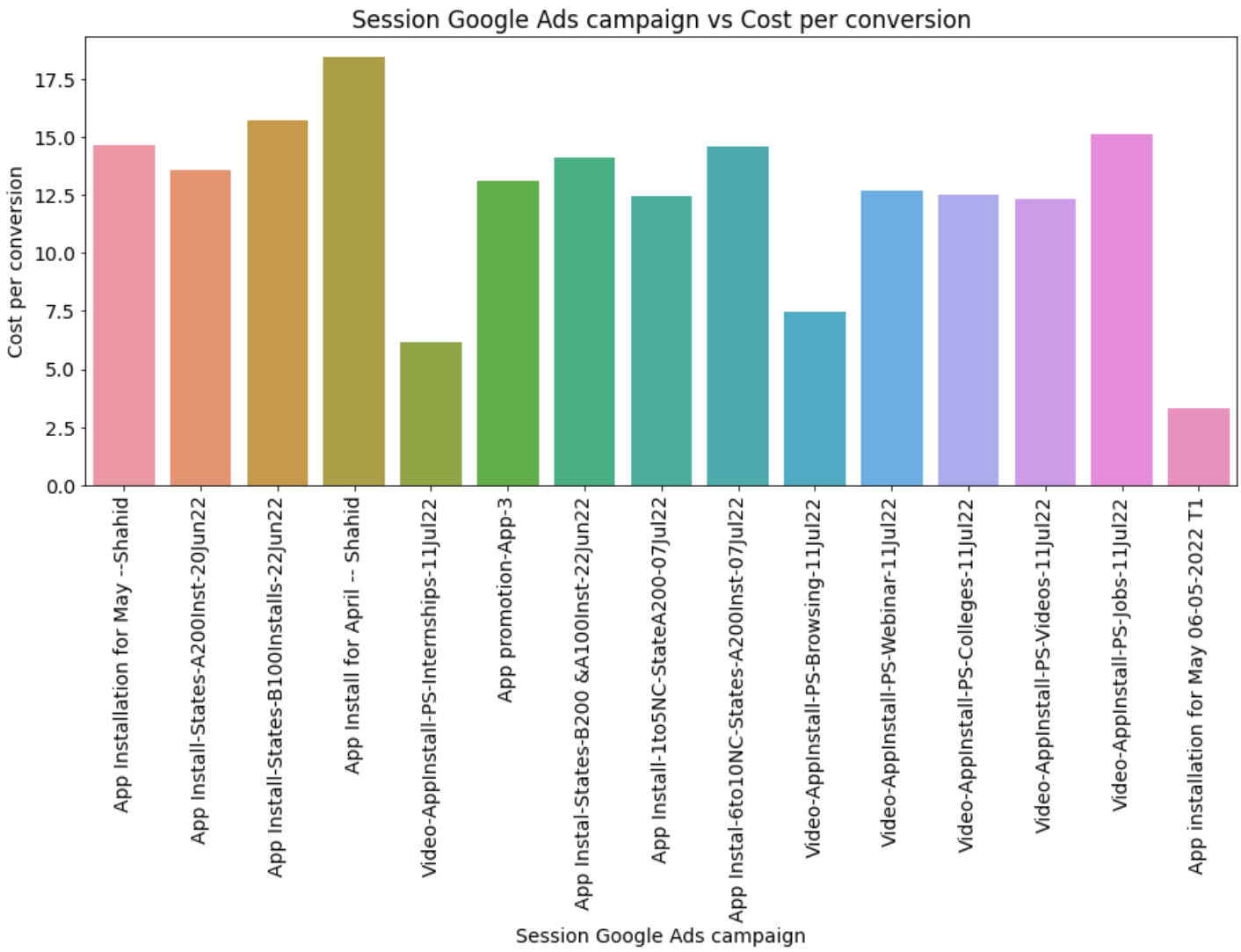
In [35]:

```
df_melted = df12.melt(id_vars='Session Google Ads campaign', value_vars=['Users', 'Google

plt.figure(figsize=(15, 6))
sns.barplot(x='Session Google Ads campaign', y='value', hue='variable', data=df_melted)
plt.title('Session Google Ads campaign vs Users, Google Ads clicks, Conversions')
plt.xticks(rotation=90)
plt.show()
```



In [36]:

```
plt.figure(figsize=(15, 6))
sns.barplot(x='Session Google Ads campaign', y='Cost per conversion', data=df12)
plt.title('Session Google Ads campaign vs Cost per conversion')
plt.xticks(rotation=90)
plt.show()
```



---------- IMPORTANT THINGS FROM THE ANALYSIS ----------

1. In User Acquisition: Organic Search is the most important factor for user conversion.
2. In Traffic Acquisition: The Unassigned Category is crucial for conversions.
3. In Event Report: Screen Views, Session Starts, First Opens, and User Engagements occur more frequently during user usage.
4. In Conversion Report: Notifications play a vital role in conversions. When a customer sees a notification, it triggers user engagement.
5. In Page Path and Screen Class: If a conversion occurs, the Login Page appears more frequently. Login Flutter, Feeds, Storyboard, etc., are major components here.
6. In Demographics Report: India has the highest conversion rate, while Kenya, Panama, Maldives, Sri Lanka, and Romania have the lowest conversions.
7. In Citywise Report: Bengaluru has the highest conversion rate, followed by Hyderabad.
8. In Gender Report: Users often do not specify their gender during login, resulting in it being considered unknown. Additionally, males have a higher conversion rate.
9. In User by Interest: Users are interested in shopping, media entertainment, technology, and mobile usage.

10. In User by Language: Users primarily use English, followed by Hindi, which has the highest conversion rate.
11. In User by Age: People aged 18-24 have the second-highest conversion rate, while those aged 45-54 have lower conversion rates. Those aged 65+ have less engagement time.
12. In Google Ads Report: The cost per conversion is higher for April app installations. Despite the amount paid for Google Ads, the conversion rate is very low.

---------- NEEDS TO IMPROVRE ----------

1. Users are required to provide their correct age and gender during login; only then can we analyze which categories people interact with our products most.

1. In the Indian region, we need to focus on tier-2 and tier-3 cities.

1. In the 25-34 age category, the conversion rate is very low. Despite users being independent in that category, conversions are low. We need to focus more on that age group.

1. Vehicle, Food & Dining, Finance & Banking, Home & Garden, and Wellness are the least interesting to users. Perhaps we should focus on improving or discontinuing them.

1. We must be cautious about investing in Google Ads because the conversion rate is very low considering the money we allocate to the ads.

In [ ]:

In [ ]:

In [ ]: