In [1]:

```python
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
```

In [2]:

```python
data = pd.read_excel(r"D:\python data\facebook_user_data1.xlsx")
```

In [3]:

```python
data.shape
```

Out[3]:

```
(99003, 15)
```

In [4]:

```python
data.head()
```

Out[4]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	likes	likes_received	mobile_likes	mobile_likes_received	www_likes	www_likes_received
0	2094382	14	19	1999	11	male	266.0	0	0	0	0	0	0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0	0	0	0	0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0	0	0	0	0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0	0	0	0	0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0	0	0	0	0	0	0

In [5]:

```python
data.isnull().sum()
```

Out[5]:

```
userid                     0
age                        0
dob_day                    0
dob_year                   0
dob_month                  0
gender                   175
tenure                     2
friend_count               0
friendships_initiated      0
likes                      0
likes_received             0
mobile_likes               0
mobile_likes_received      0
www_likes                  0
www_likes_received         0
dtype: int64
```

In [6]:

```python
data['gender'] = data['gender'].fillna(data['gender'].mode()[0])
```

In [9]:

```python
data.isnull().sum()
```

Out[9]:

```
userid                   0
age                      0
dob_day                  0
dob_year                 0
dob_month                0
gender                   0
tenure                   2
friend_count             0
friendships_initiated    0
likes                    0
likes_received           0
mobile_likes             0
mobile_likes_received    0
www_likes                0
www_likes_received       0
dtype: int64
```

In [11]:

```python
data['tenure'] = data['tenure'].fillna(data['tenure'].median())
```

In [12]:

```python
data.isnull().sum()
```

Out[12]:

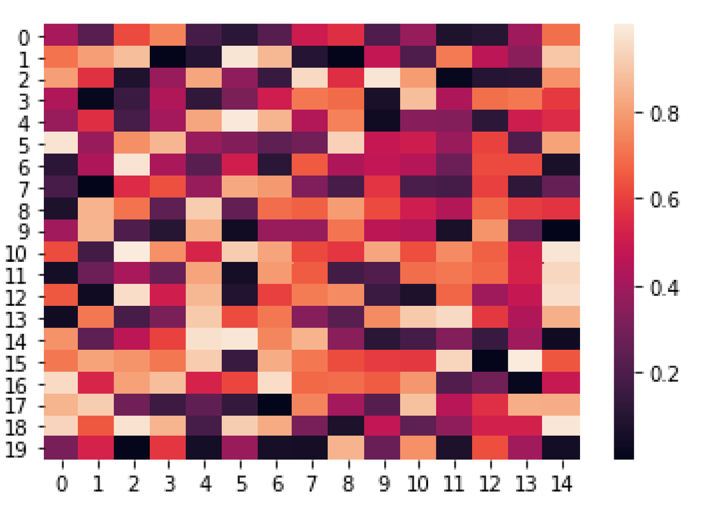
```
userid                   0
age                      0
dob_day                  0
dob_year                 0
dob_month                0
gender                   0
tenure                   0
friend_count             0
friendships_initiated    0
likes                    0
likes_received           0
mobile_likes             0
mobile_likes_received    0
www_likes                0
www_likes_received       0
dtype: int64
```

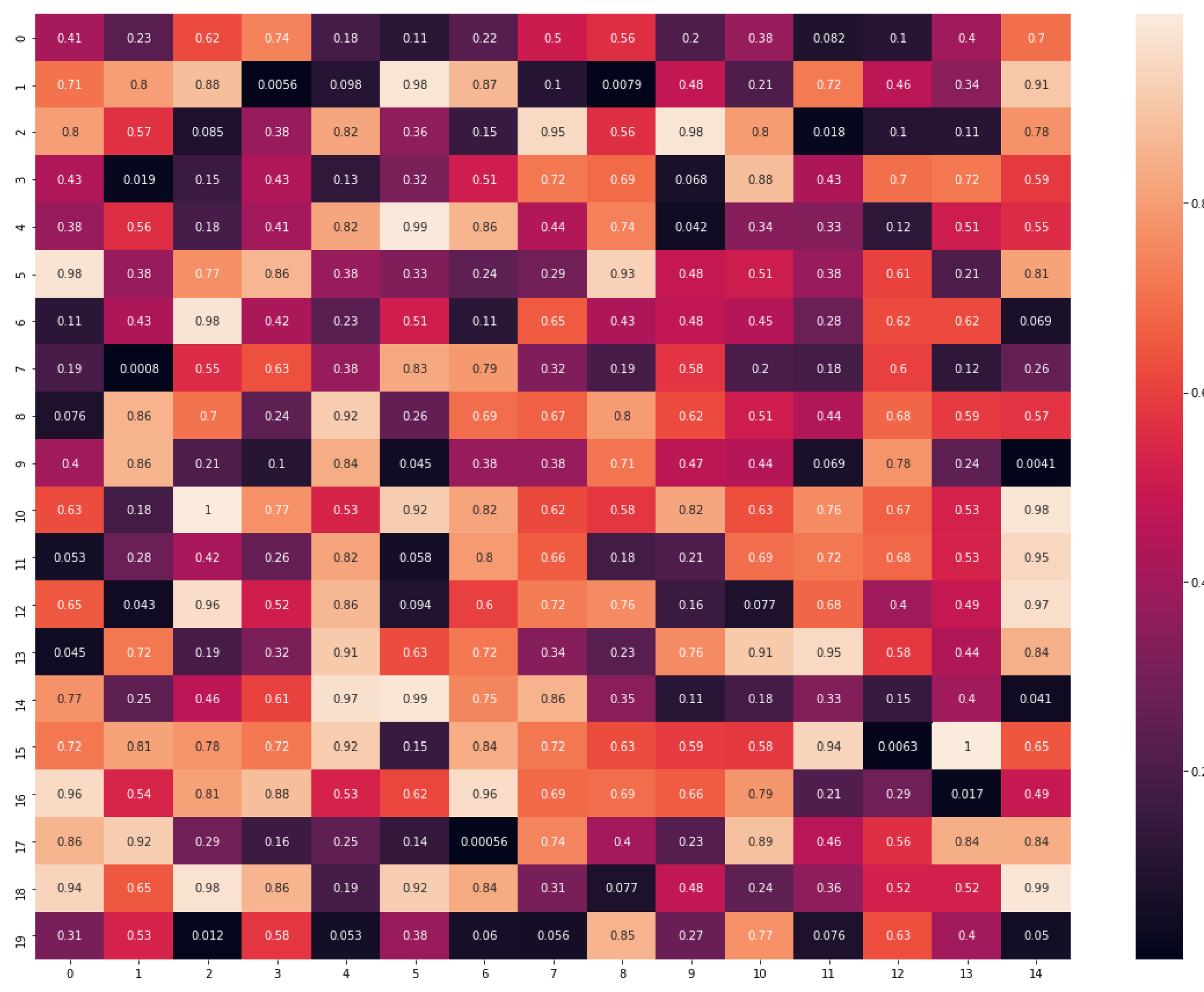
In [13]:

```python
data=np.random.rand(20,15)
data

sns.heatmap(data)

plt.figure(figsize=(20,15))
sns.heatmap(data, annot=True)
```

Out[13]:

```
<AxesSubplot:>
```





In [15]:

```python
data = pd.read_excel(r"D:\python data\facebook_user_data.xlsx")
data.groupby(['gender']).size().reset_index(name='count')
```

Out[15]:

	gender	count
0	female	40254
1	male	58574

In [8]:

```python
data.groupby('gender').agg({'friend_count':['sum']})
```

Out[8]:

	friend_count
	sum
gender	
female	9740258
male	9699059

In [16]:

```python
data.groupby('gender').agg({'friendships_initiated':['sum']})
```

Out[16]:

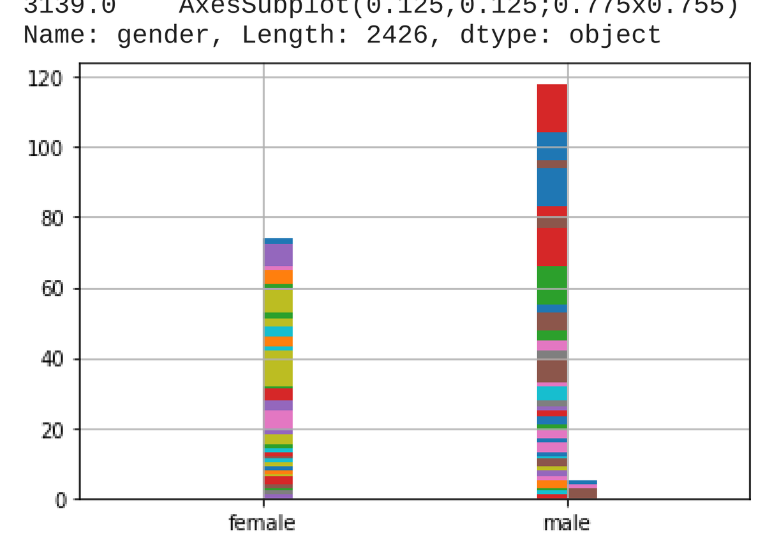
	friendships_initiated
	sum
gender	
female	4584894
male	6037023

In [23]:

```python
data.groupby('tenure').gender.hist()
```

Out[23]:

```
tenure
0.0       AxesSubplot(0.125,0.125;0.775x0.755)
1.0       AxesSubplot(0.125,0.125;0.775x0.755)
2.0       AxesSubplot(0.125,0.125;0.775x0.755)
3.0       AxesSubplot(0.125,0.125;0.775x0.755)
4.0       AxesSubplot(0.125,0.125;0.775x0.755)
                          ...
3019.0    AxesSubplot(0.125,0.125;0.775x0.755)
3101.0    AxesSubplot(0.125,0.125;0.775x0.755)
3128.0    AxesSubplot(0.125,0.125;0.775x0.755)
3129.0    AxesSubplot(0.125,0.125;0.775x0.755)
3139.0    AxesSubplot(0.125,0.125;0.775x0.755)
Name: gender, Length: 2426, dtype: object
```



In [18]:

```python
data[data.friend_count < 1].shape[0]
len(data[data.friend_count < 1])
data[data.friend_count < 1]
```

Out[18]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	likes	likes_received	mobile_likes	mobile_likes_received	www_likes	www_likes_received
0	2094382	14	19	1999	11	male	266.0	0	0	0	0	0	0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0	0	0	0	0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0	0	0	0	0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0	0	0	0	0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1957	1395843	90	4	1923	7	male	0.0	0	0	95	78	95	56	0	22
1958	1306409	23	24	1990	1	male	48.0	0	0	164	0	164	0	0	0
1959	1838179	18	29	1995	10	female	0.0	0	0	246	100	0	9	246	91
1960	1594192	63	1	1950	1	female	57.0	0	0	271	1	271	0	0	1
1961	1609579	20	23	1993	1	female	257.0	0	0	848	1	848	0	0	1

1962 rows × 15 columns

In [19]:

```python
data[data.likes < 1].shape[0]
len(data[data.likes < 1])
data[data.likes < 1].count()
```

Out[19]:

```
userid                   22308
age                      22308
dob_day                  22308
dob_year                 22308
dob_month                22308
gender                   22285
tenure                   22308
friend_count             22308
friendships_initiated    22308
likes                    22308
likes_received           22308
mobile_likes             22308
mobile_likes_received    22308
www_likes                22308
www_likes_received       22308
dtype: int64
```

In [20]:

```python
data[data.likes_received < 1].shape[0]
len(data[data.likes_received < 1])
data[data.likes_received < 1]
```

Out[20]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	likes	likes_received	mobile_likes	mobile_likes_received	www_likes	www_likes_received
0	2094382	14	19	1999	11	male	266.0	0	0	0	0	0	0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0	0	0	0	0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0	0	0	0	0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0	0	0	0	0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
98005	1883270	18	14	1995	11	female	341.0	2050	467	0	0	0	0	0	0
98016	1993020	21	8	1992	1	male	1262.0	1954	1050	0	0	0	0	0	0
98061	1680023	17	13	1996	10	female	736.0	2415	1795	3	0	0	0	3	0
98081	2022131	16	14	1997	8	male	743.0	2294	1157	6	0	6	0	0	0
98082	1867596	41	14	1972	3	male	412.0	1936	1488	6	0	6	0	0	0

24428 rows × 15 columns

In [21]:

```python
data.groupby("gender")[["mobile_likes","www_likes"]].mean()
```

Out[21]:

	mobile_likes	www_likes
gender		
female	172.912928	87.138297
male	60.261328	24.416550

In [22]:

```python
data.groupby("gender")[["mobile_likes_received","www_likes_received"]].mean()
```

Out[22]:

	mobile_likes_received	www_likes_received
gender		
female	147.100884	104.334451
male	40.833015	27.078533

In [ ]:

```python
data.plot(x="gender", y=["tenure"], kind="bar")
```

In [ ]:

```python

```