In [1]:
```
#This one is analyzing combination of date_created column
# + country location + user friends into one facet count graph.
```

In [100]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import os
print(os.getcwd())
```

```
C:\Users\spark\Desktop\Indesign Print
```

In [101]:
```
os.chdir('C:\\Users\\spark\\Desktop\\Data Science Projects\\Squid Game Analysis')
```

In [102]:
```
print(os.getcwd())
```

```
C:\Users\spark\Desktop\Data Science Projects\Squid Game Analysis
```

In [103]:
```
data = pd.read_csv('tweets_v83.csv', parse_dates= ["user_created"])
```

In [104]:
```
1 data
```

Out[104]:

	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	user_verified
0	the _??nd??r-rat??d nigg??h????????	NaN	@ManUtd die hard??????????????????????????? \n...	2019-09-06 19:24:57+00:00	581	1035	8922	False
1	Best uncle on planet earth	NaN	NaN	2013-05-08 19:35:26+00:00	741	730	8432	False
2	marcie	NaN	animal crossing. chicken nuggets. baby yoda. s...	2009-02-21 10:31:30+00:00	562	1197	62732	False
3	YoMo.Mdp	Any pronouns	Where the heck is the karma\nI'm going on my s...	2021-02-14 13:21:22+00:00	3	277	1341	False
4	Laura Reactions	France	I talk and I make reactions videos about shows...	2018-12-19 20:38:28+00:00	330	152	2278	False
...	...	...	...	...	...	...	...	...
80014	Sale X Low	USA	Sale X Low, the cheapest low price online stor...	2021-04-03 20:49:35+00:00	12	49	7	False
80015	RevAAA	NaN	Review Anything Anyone Anywhere	2010-11-24 21:48:53+00:00	6907	0	0	False
80016	Omo K-Pop News	NaN	OMO K-Pop\nNews and Updates	2021-09-22 12:01:24+00:00	152	807	593	False
80017	???? Pumpkin??? Queen ????	Halloweentown ????	Creator of Stuff | Horror fan | Traveler | Bad...	2009-02-19 17:09:28+00:00	8048	8822	68980	False
80018	levi do lay	NaN	NaN	2020-06-17 11:55:56+00:00	16310	14368	108325	False

80019 rows × 12 columns

In [105]:
```
1 data2 =pd.DataFrame(data, columns = ['user_followers','user_friends','source','user_created','user_location'])
```

In [106]:
```
1 data2
```

Out[106]:

	user_followers	user_friends	source	user_created	user_location
0	581	1035	Twitter for Android	2019-09-06 19:24:57+00:00	NaN
1	741	730	Twitter for Android	2013-05-08 19:35:26+00:00	NaN
2	562	1197	Twitter Web App	2009-02-21 10:31:30+00:00	NaN
3	3	277	Twitter Web App	2021-02-14 13:21:22+00:00	Any pronouns
4	330	152	Twitter Web App	2018-12-19 20:38:28+00:00	France
...	...	...	...	...	...
80014	12	49	SocialRabbit Plugin	2021-04-03 20:49:35+00:00	USA
80015	6907	0	Twitter for iPhone	2010-11-24 21:48:53+00:00	NaN
80016	152	807	Twitter Web App	2021-09-22 12:01:24+00:00	NaN
80017	8048	8822	Twitter for Android	2009-02-19 17:09:28+00:00	Halloweentown ????
80018	16310	14368	Twitter for iPhone	2020-06-17 11:55:56+00:00	NaN

80019 rows × 5 columns

In [107]:
```
1 data2.dropna(subset = ["user_location","source","user_created"], inplace=True)
```

In [108]:
```
1 data2
```

Out[108]:

	user_followers	user_friends	source	user_created	user_location
3	3	277	Twitter Web App	2021-02-14 13:21:22+00:00	Any pronouns
4	330	152	Twitter Web App	2018-12-19 20:38:28+00:00	France
5	546	318	Twitter for Android	2018-01-27 12:07:31+00:00	United Kingdom
7	1877	2057	Twitter for iPhone	2011-03-28 18:56:28+00:00	South Africa
9	1027	278	Twitter Web App	2009-05-11 20:14:51+00:00	Calgary, Canada
...	...	...	...	...	...
80008	188	244	Twitter for Android	2017-06-01 15:56:21+00:00	Whitehaven, England
80009	681095	38	TweetDeck	2010-05-29 12:45:22+00:00	Karachi, Pakistan
80013	1153	165	Twitter Web App	2013-04-07 16:56:34+00:00	Murcia Spain
80014	12	49	SocialRabbit Plugin	2021-04-03 20:49:35+00:00	USA
80017	8048	8822	Twitter for Android	2009-02-19 17:09:28+00:00	Halloweentown ????

56149 rows × 5 columns

In [109]:
```
1 #Find locations that are within the United States including the states
2 findUSA = data2
```

In [110]:
```
1 findUSA =findUSA.loc[findUSA['user_location'].str.contains("America|United States|USA|U.S.|Alabama|Alaska|Arizona|Ca
```

In [111]:
```
1 data2.loc[findUSA.index,"user_location"]="USA"
2
```

In [112]:
```
findUK = data2
findUK= findUK.loc[findUK['user_location'].str.contains("United Kingdom|Kingdom|Lond|England|Wales|Scotland|Northern
data2.loc[findUK.index,"user_location"]="UK"
```

In [113]:
```
#Find locations that are in India
findIndia = data2
data2
```

Out[113]:

	user_followers	user_friends	source	user_created	user_location
3	3	277	Twitter Web App	2021-02-14 13:21:22+00:00	Any pronouns
4	330	152	Twitter Web App	2018-12-19 20:38:28+00:00	France
5	546	318	Twitter for Android	2018-01-27 12:07:31+00:00	UK
7	1877	2057	Twitter for iPhone	2011-03-28 18:56:28+00:00	South Africa
9	1027	278	Twitter Web App	2009-05-11 20:14:51+00:00	Calgary, Canada
...	...	...	...	...	...
80008	188	244	Twitter for Android	2017-06-01 15:56:21+00:00	UK
80009	681095	38	TweetDeck	2010-05-29 12:45:22+00:00	Karachi, Pakistan
80013	1153	165	Twitter Web App	2013-04-07 16:56:34+00:00	Murcia Spain
80014	12	49	SocialRabbit Plugin	2021-04-03 20:49:35+00:00	USA
80017	8048	8822	Twitter for Android	2009-02-19 17:09:28+00:00	Halloweentown ????

56149 rows × 5 columns

In [114]:
```
findIndia=data2
```

In [115]:
```
findIndia= findIndia.loc[findIndia['user_location'].str.contains("India|inda", case=True)]
data2.loc[findIndia.index,"user_location"]="India"
```

In [116]:
```
1 findIndia
```

Out[116]:

	user_followers	user_friends	source	user_created	user_location
39	177	124	Twitter for Android	2010-12-28 06:27:05+00:00	Chennai, India ????????
54	34	143	Twitter for iPhone	2021-06-21 19:11:40+00:00	India
97	591	431	Twitter for iPhone	2017-07-28 11:24:42+00:00	India
109	1080	69	Twitter for Android	2014-07-18 12:47:10+00:00	Gurgaon, India
127	44501	1319	Twitter for Android	2010-08-15 08:25:21+00:00	Mumbai, India
...	...	...	...	...	...
79702	49810	43	Twitter for iPhone	2014-04-28 03:47:09+00:00	India
79714	57	495	Twitter for iPhone	2018-09-04 16:32:20+00:00	Bengaluru South, India
79754	772	937	Twitter for Android	2013-09-05 08:27:04+00:00	Kurnool, India
79817	4791	1715	Twitter for iPhone	2010-05-05 19:13:43+00:00	New Delhi, India
79878	0	0	Twitter Web App	2021-10-25 15:29:08+00:00	India

2352 rows × 5 columns

In [117]:
```
1
2
3 indexes = data2[ (data2['user_location'] != "India") & (data2['user_location']!="UK" )&(data2['user_location']!="USA
4
```

In [118]:
```
1
2
3 data2.drop(indexes,inplace=True)
4
```

In [119]:
```
#Next, user friends of twitters more than 1000 means  that the account is not a private but a public platform.
#So I chose to eliminate these twitter accounts that have friends higher than 1000 in analysis
```

In [120]:
```
data2= data2[data2['user_friends'] <= 1000] #It's not possible to have more than 1000 user friends unless the id is
#a business platform.
data2= data2[data2['user_followers'] <= 1000]
```

In [121]:
```
data2.dtypes
```

Out[121]:
```
user_followers                 int64
user_friends                   int64
source                        object
user_created      datetime64[ns, UTC]
user_location                 object
dtype: object
```

In [122]:
```
data2['user_created']= pd.to_datetime(data2['user_created'])
```

In [123]:
```
data2.user_created = data2.user_created.dt.year
```

In [124]:
```
data2['new']= data2["user_followers"] / data2["user_friends"]
```

In [125]:
```
1 data2
```

Out[125]:

	user_followers	user_friends	source	user_created	user_location	new
5	546	318	Twitter for Android	2018	UK	1.716981
31	24	47	Twitter Web App	2020	UK	0.510638
39	177	124	Twitter for Android	2010	India	1.427419
47	561	700	Twitter Web App	2013	USA	0.801429
54	34	143	Twitter for iPhone	2021	India	0.237762
...	...	...	...	...	...	...
79942	34	247	Twitter Web App	2014	USA	0.137652
79965	555	224	Twitter for iPhone	2017	USA	2.477679
79971	12	49	SocialRabbit Plugin	2021	USA	0.244898
80008	188	244	Twitter for Android	2017	UK	0.770492
80014	12	49	SocialRabbit Plugin	2021	USA	0.244898

8991 rows × 6 columns

In [126]:
```
1 data2['new_percentage'] = data2['new']*100
```

In [127]:
```
1 c=data2.loc[(data2.source == 'Twitter for Android') |
2               (data2.source == 'Twitter Web App') |
3               (data2.source == 'Twitter for iPhone') ,
4    ['source','user_created','new_percentage','new','user_friends','user_followers','user_location'],]
```

In [128]:
```
1 c.dtypes
```

Out[128]:
```
source             object
user_created        int64
new_percentage    float64
new               float64
user_friends        int64
user_followers      int64
user_location      object
dtype: object
```

In [129]:
```
1 c
```

Out[129]:

	source	user_created	new_percentage	new	user_friends	user_followers	user_location
5	Twitter for Android	2018	171.698113	1.716981	318	546	UK
31	Twitter Web App	2020	51.063830	0.510638	47	24	UK
39	Twitter for Android	2010	142.741935	1.427419	124	177	India
47	Twitter Web App	2013	80.142857	0.801429	700	561	USA
54	Twitter for iPhone	2021	23.776224	0.237762	143	34	India
...	...	...	...	...	...	...	...
79890	Twitter for Android	2018	29.629630	0.296296	351	104	UK
79907	Twitter Web App	2011	45.600000	0.456000	125	57	UK
79942	Twitter Web App	2014	13.765182	0.137652	247	34	USA
79965	Twitter for iPhone	2017	247.767857	2.477679	224	555	USA
80008	Twitter for Android	2017	77.049180	0.770492	244	188	UK

8036 rows × 7 columns

In [130]:
```
#c=c.drop(80015)
data2
```

Out[130]:

	user_followers	user_friends	source	user_created	user_location	new	new_percentage
5	546	318	Twitter for Android	2018	UK	1.716981	171.698113
31	24	47	Twitter Web App	2020	UK	0.510638	51.063830
39	177	124	Twitter for Android	2010	India	1.427419	142.741935
47	561	700	Twitter Web App	2013	USA	0.801429	80.142857
54	34	143	Twitter for iPhone	2021	India	0.237762	23.776224
...	...	...	...	...	...	...	...
79942	34	247	Twitter Web App	2014	USA	0.137652	13.765182
79965	555	224	Twitter for iPhone	2017	USA	2.477679	247.767857
79971	12	49	SocialRabbit Plugin	2021	USA	0.244898	24.489796
80008	188	244	Twitter for Android	2017	UK	0.770492	77.049180
80014	12	49	SocialRabbit Plugin	2021	USA	0.244898	24.489796

8991 rows × 7 columns

In [131]:
```
c.loc[(c.source == 'Twitter for Android'),'source']='Android'
c.loc[(c.source == 'Twitter Web App'),'source']='Web App'
c.loc[(c.source == 'Twitter for iPhone'),'source']='iPhone'
```

In [132]:
```
c.rename(columns = {'user_location':'location'}, inplace = True)
```

In [133]:
```
1 c
```

Out[133]:

	source	user_created	new_percentage	new	user_friends	user_followers	location
5	Android	2018	171.698113	1.716981	318	546	UK
31	Web App	2020	51.063830	0.510638	47	24	UK
39	Android	2010	142.741935	1.427419	124	177	India
47	Web App	2013	80.142857	0.801429	700	561	USA
54	iPhone	2021	23.776224	0.237762	143	34	India
...	...	...	...	...	...	...	...
79890	Android	2018	29.629630	0.296296	351	104	UK
79907	Web App	2011	45.600000	0.456000	125	57	UK
79942	Web App	2014	13.765182	0.137652	247	34	USA
79965	iPhone	2017	247.767857	2.477679	224	555	USA
80008	Android	2017	77.049180	0.770492	244	188	UK

8036 rows × 7 columns

In [134]:
```
1 os.chdir('C:\\Users\\spark\\Desktop\\Indesign Print')
```

In [135]:
```
1 c
```

Out[135]:

	source	user_created	new_percentage	new	user_friends	user_followers	location
5	Android	2018	171.698113	1.716981	318	546	UK
31	Web App	2020	51.063830	0.510638	47	24	UK
39	Android	2010	142.741935	1.427419	124	177	India
47	Web App	2013	80.142857	0.801429	700	561	USA
54	iPhone	2021	23.776224	0.237762	143	34	India
...	...	...	...	...	...	...	...
79890	Android	2018	29.629630	0.296296	351	104	UK
79907	Web App	2011	45.600000	0.456000	125	57	UK
79942	Web App	2014	13.765182	0.137652	247	34	USA
79965	iPhone	2017	247.767857	2.477679	224	555	USA
80008	Android	2017	77.049180	0.770492	244	188	UK

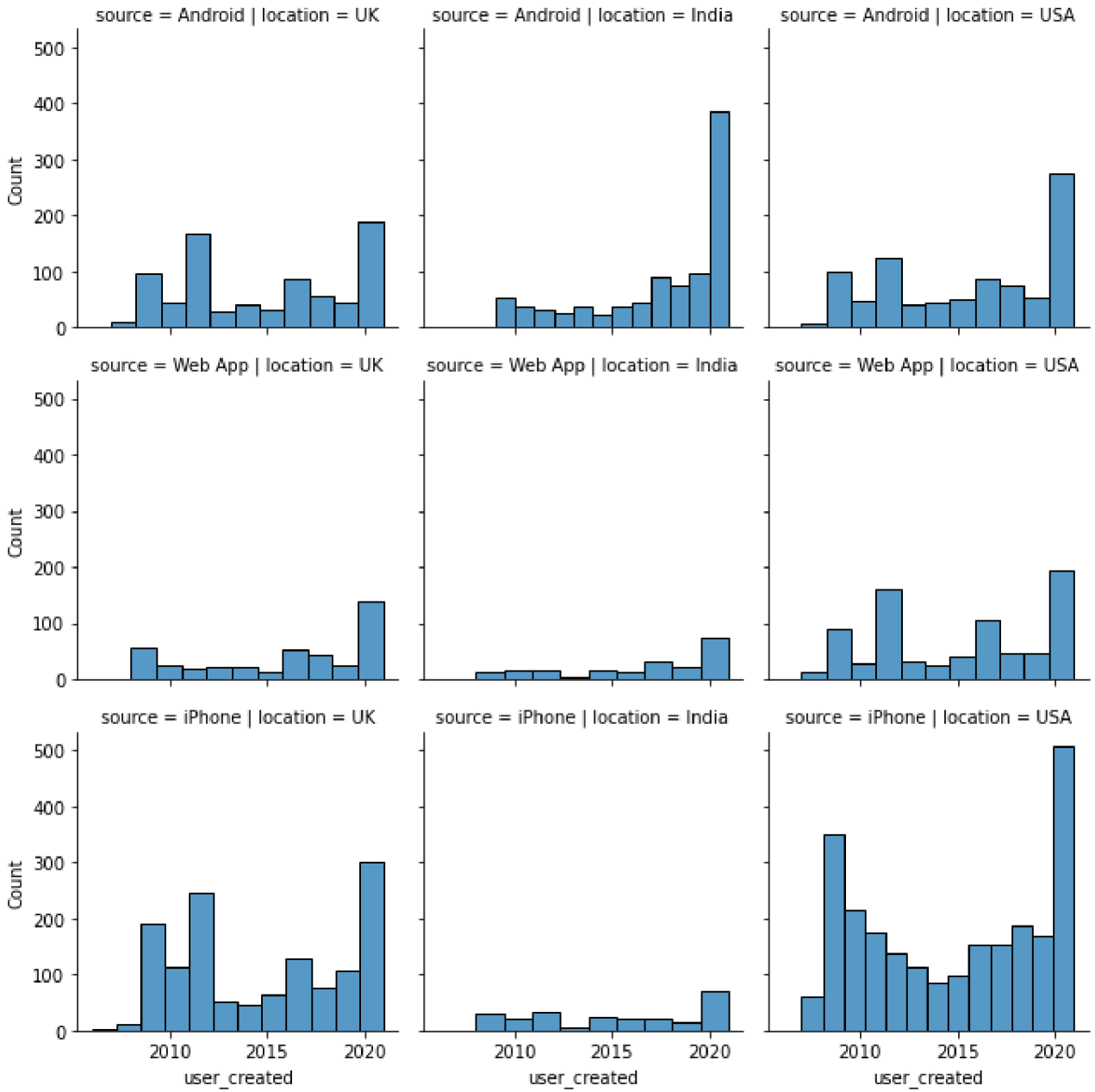
8036 rows × 7 columns

In [142]:

```
plt.tight_layout()
plt.subplots_adjust(hspace=1, wspace=1)
g = sns.FacetGrid(c, col="location",  row="source")
g.map_dataframe(sns.histplot, x="user_created")
plt.savefig('test14.png',dpi=200);

```

```
<Figure size 432x288 with 0 Axes>
```

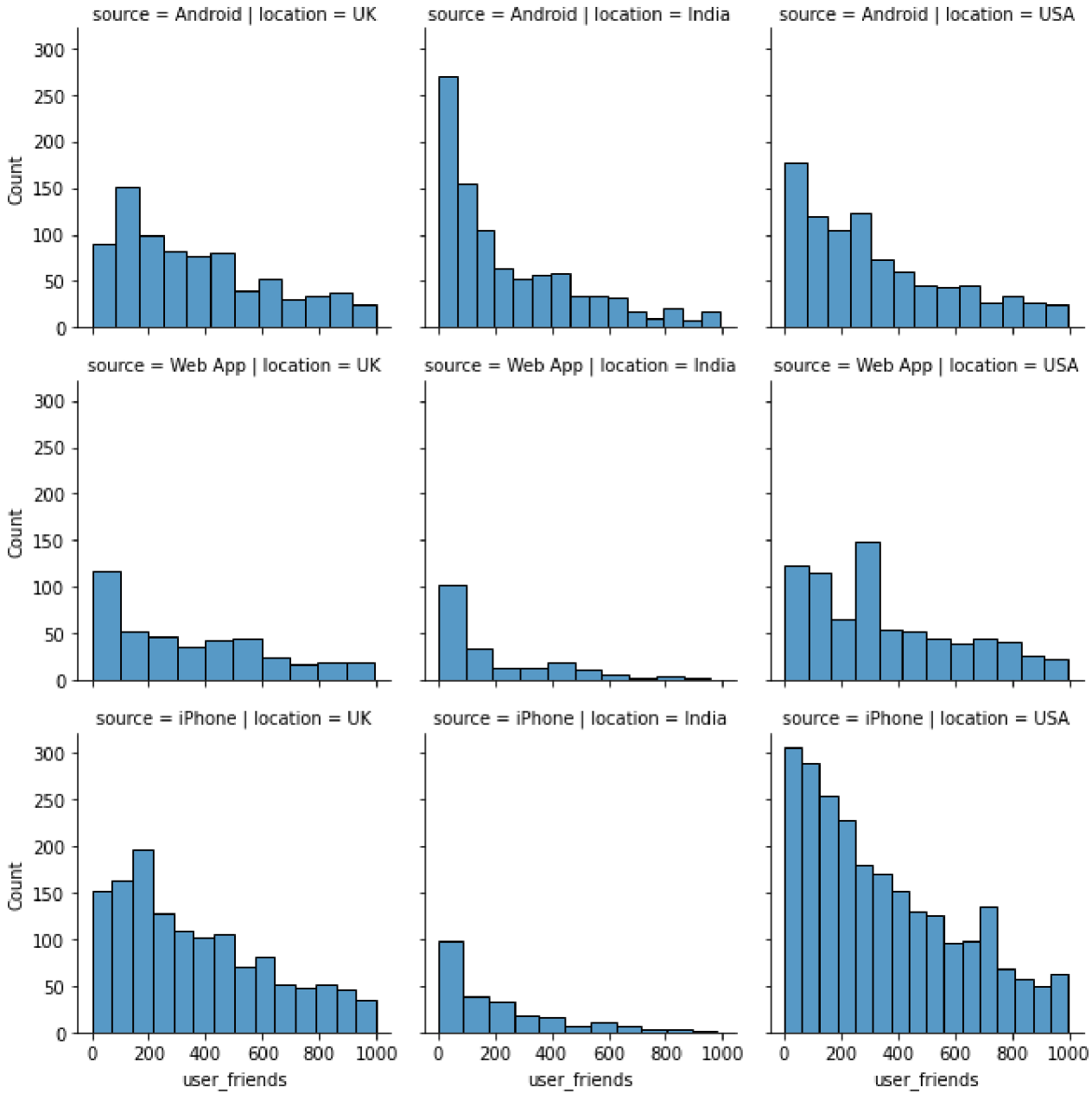

source = Android | location = UK
source = Android | location = India
source = Android | location = USA
source = Web App | location = UK
source = Web App | location = India
source = Web App | location = USA
source = iPhone | location = UK
source = iPhone | location = India
source = iPhone | location = USA
Count
0
100
200
300
400
500
2010
2015
2020
user_created

In [143]:

```
#g = sns.FacetGrid(c, col="location",  row="source")
#g.map(sns.scatterplot, "user_created", "user_friends")

g = sns.FacetGrid(c, col="location",  row="source")
g.map_dataframe(sns.histplot, x="user_friends")

plt.savefig('test15.png',dpi=200);
```


source = Android | location = UK
source = Android | location = India
source = Android | location = USA
source = Web App | location = UK
source = Web App | location = India
source = Web App | location = USA
source = iPhone | location = UK
source = iPhone | location = India
source = iPhone | location = USA
Count
user_friends
0
50
100
150
200
250
300
400
600
800
1000

In [144]:

```
plt.figure(figsize=(2,1))
g = sns.FacetGrid(c, col="user_created",row="source")
g.map(sns.scatterplot, "user_followers", "user_friends")
plt.savefig('test16.png',dpi=200);
#If you have over 4000 friends, then you are most likely some sort of a business related coporations that can be eli
#There tends to be more bussiness related users commenting about squid game from android.
#The graph below uses a similar properties using scatterplot instead of countplot
```

```
<Figure size 144x72 with 0 Axes>
```