Google Trends Analysis of South Indian Movies 'RRR' and 'KGF 2'

Date of Analysis: 15th May 2022

Indian Movie Industry is a big industry and has a high reach of audience after Hollywood movies. Thanks to Globalization and OTT Platforms South Indian Regional Films are making it big. The recent ones that made film enthusiasts across the world wait for are 'RRR' and 'KGF 2'. Lets see how on Google Trends these two perform

Movie Details: RRR - 24th March 2022,Starring: Ram Charan & NTR Director: SS Rajamouli KGF 2 - 14th April 2022, Starring: Yash, Director: Prashanth Neel Bachchan Pandey - 18th March 2022, Starring: Akshay Kumar, Director: Farhad Samji Jhund - 4th March 2022, Starring: Amitabh Bachchan, Director: Nagraj Manjule

We will see how Hindi releases perform in Google trends from (15th May 2022 till 15th Feb 2022)

In [1]:
```
#pip install pytrends
```

```
Collecting pytrendsNote: you may need to restart the kernel to use updated packages.
  Downloading pytrends-4.8.0.tar.gz (19 kB)
Requirement already satisfied: requests>=2.0 in c:\users\user\anaconda3\lib\site-packa
ges (from pytrends) (2.27.1)
Requirement already satisfied: pandas>=0.25 in c:\users\user\anaconda3\lib\site-packag
es (from pytrends) (1.1.3)
Requirement already satisfied: lxml in c:\users\user\anaconda3\lib\site-packages (from
pytrends) (4.6.1)
Requirement already satisfied: charset-normalizer~=2.0.0; python_version >= "3" in
c:\users\user\anaconda3\lib\site-packages (from requests>=2.0->pytrends) (2.0.12)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\user\anaconda3\lib\si
te-packages (from requests>=2.0->pytrends) (1.25.11)
Requirement already satisfied: idna<4,>=2.5; python_version >= "3" in c:\users\user\an
aconda3\lib\site-packages (from requests>=2.0->pytrends) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\user\anaconda3\lib\site-
packages (from requests>=2.0->pytrends) (2020.6.20)
Requirement already satisfied: python-dateutil>=2.7.3 in c:\users\user\anaconda3\lib\s
ite-packages (from pandas>=0.25->pytrends) (2.8.1)

Requirement already satisfied: numpy>=1.15.4 in c:\users\user\anaconda3\lib\site-packa
ges (from pandas>=0.25->pytrends) (1.19.2)
Requirement already satisfied: pytz>=2017.2 in c:\users\user\anaconda3\lib\site-packag
es (from pandas>=0.25->pytrends) (2020.1)
Requirement already satisfied: six>=1.5 in c:\users\user\anaconda3\lib\site-packages
(from python-dateutil>=2.7.3->pandas>=0.25->pytrends) (1.15.0)
Building wheels for collected packages: pytrends
  Building wheel for pytrends (setup.py): started
  Building wheel for pytrends (setup.py): finished with status 'done'
  Created wheel for pytrends: filename=pytrends-4.8.0-py3-none-any.whl size=16109 sha2
56=ff4887cf1ab69dbd91faf3c94fd5164791f3925955dc63def0876e77e5fdffff
  Stored in directory: c:\users\user\appdata\local\pip\cache\wheels\e8\78\c8\18d4f4804
753e14416809b365773220c48b41fe5387f2bb6b9
Successfully built pytrends
Installing collected packages: pytrends
Successfully installed pytrends-4.8.0
```

In [2]:
```
import pandas as pd
from pytrends.request import TrendReq
import matplotlib.pyplot as plt
import seaborn as sns
trends = TrendReq()
```

'RRR' vs 'Jhund' on Google Trends

In [62]:
```
kw_list=['RRR', 'Jhund']

#search interest per region
#run model for keywords (can also be competitors)
pytrend.build_payload(kw_list, timeframe='today 3-m') ###Google trends from today til
regiondf1 = pytrend.interest_by_region() ## Gives Google Trends by countries
region_df1=regiondf.sort_values('Jhund',ascending=False)
```

In [63]:
```
region_df1
```

Out[63]:

geoName	RRR	Jhund
New Zealand	92	8
Canada	94	6
United Kingdom	94	6
Singapore	95	5
Pakistan	95	5
...	...	...
Guam	0	0
Guatemala	0	0
Guernsey	0	0
Guinea	0	0
Åland Islands	0	0

250 rows × 2 columns

'RRR' vs 'Bachchan Pandey'

In [75]:
```
kw_list2=['RRR', 'Bachchan Pandey']
pytrend.build_payload(kw_list2, timeframe='today 3-m') ###Google trends from today ti
regiondf2 = pytrend.interest_by_region() ## Gives Google Trends by countries
region_df2=regiondf2.sort_values('Bachchan Pandey',ascending=False)
```

In [76]:
```
region_df2
```

Out[76]:

geoName	RRR	Bachchan Pandey
Pakistan	74	26
New Zealand	84	16
Canada	85	15
United Kingdom	86	14
Saudi Arabia	88	12
...	...	...
Guam	0	0
Guatemala	0	0
Guernsey	0	0
Guinea	0	0
Åland Islands	0	0

250 rows × 2 columns

Bachchan Pandey performed better than Jhund on Google Trends, But those numbers are way less to 'RRR'. So moving ahead with 'RRR' vs 'KGF 2'

'RRR' vs 'KGF 2'

In [77]:
```
kw_list=['RRR', 'KGF 2']

#search interest per region
#run model for keywords (can also be competitors)
pytrend.build_payload(kw_list, timeframe='today 3-m') ###Google trends from today til
regiondf = pytrend.interest_by_region() ## Gives Google Trends by countries
region_df=regiondf.sort_values('RRR',ascending=False) ## Sorting the values according
```

In [78]:
```
region_df.head(20)[:18] ##Taking into account top 17 countries
```

Out[78]:

geoName	RRR	KGF 2
United States	69	31
Australia	59	41
France	57	43
Canada	54	46
Germany	52	48
United Kingdom	52	48
Kuwait	49	51
Qatar	48	52
India	47	53
Oman	46	54
Nepal	46	54
Singapore	46	54
Saudi Arabia	42	58
United Arab Emirates	40	60
Malaysia	36	64
Bangladesh	30	70
Sri Lanka	26	74
Pakistan	25	75

In [79]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [80]:
```
region_df.columns
```

Out[80]:
```
Index(['RRR', 'KGF 2'], dtype='object')
```

In [81]:
```
new_df=region_df.head(20)[:18].copy()
```

In [82]:
```
new_df.reset_index(inplace=True)
```

In [83]:
```
new_df.columns
```

Out[83]:
```
Index(['geoName', 'RRR', 'KGF 2'], dtype='object')
```

In [87]:
```
len(new_df)
```

Out[87]:
```
18
```

In [85]:
```
new_df['RRR'].sum()
```

Out[85]:
```
824
```

In [98]:
```
RRR_Max_country = new_df.loc[new_df['RRR'].idxmax()]
RRR_Max_country
```

Out[98]:
```
geoName    United States
RRR                   69
KGF 2                 31
Name: 0, dtype: object
```

In [101…
```
RRR_Min_country = new_df.loc[new_df['RRR'].idxmin()]
RRR_Min_country
```

Out[101…
```
geoName    Pakistan
RRR              25
KGF 2            75
Name: 17, dtype: object
```

In [86]:
```
new_df['KGF 2'].sum()
```

Out[86]:
```
976
```

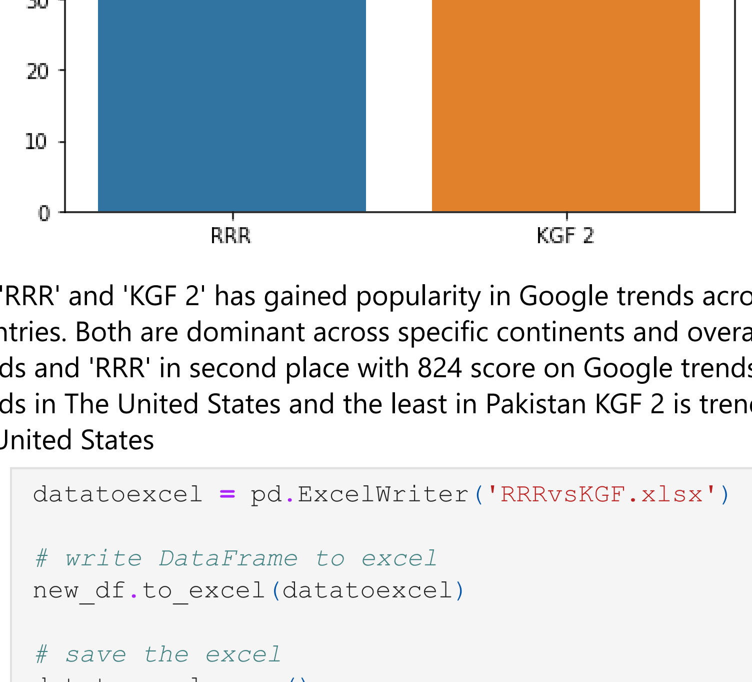
In [100…
```
KGF_Max_Country= new_df.loc[new_df['KGF 2'].idxmax()]
KGF_Max_Country
```

Out[100…
```
geoName    Pakistan
RRR              25
KGF 2            75
Name: 17, dtype: object
```

In [102…
```
KGF_Min_Country= new_df.loc[new_df['KGF 2'].idxmin()]
KGF_Min_Country
```

Out[102…
```
geoName    United States
RRR                   69
KGF 2                 31
Name: 0, dtype: object
```

In [84]:
```
sns.barplot(data=new_df)
plt.show()
```



Conclusion: 'RRR' and 'KGF 2' has gained popularity in Google trends across the world. They appeared in Google trends in over 17 countries. Both are dominant across specific continents and overall 'KGF 2' stood in first place with 976 score on Google Trends and 'RRR' in second place with 824 score on Google trends in over 18 countries. RRR is trending max on Google Trends in The United States and the least in Pakistan KGF 2 is trendinh max on Google Trends in Pakistan and the least in the United States

In [104…
```
datatoexcel = pd.ExcelWriter('RRRvsKGF.xlsx')

# write DataFrame to excel
new_df.to_excel(datatoexcel)

# save the excel
datatoexcel.save()
```

In [ ]: