r

o

t

c

e

r

i

D

Top 5 Directors on Netflix

Rajiv Chilaka

Jan Suter

Raúl Campos

Marcus Raboy

Suhas Kadav

The Top 5 successful actor

0 5 10 15 20

Total Content

dff['cast']=dff['cast'].fillna('No Cast Specified')

filtered\_cast=pd.DataFrame()

filtered\_cast=dff['cast'].str.split(',',expand=True).stack() filtered\_cast=filtered\_cast.to\_frame()

filtered\_cast.columns=['Actor']

actors=filtered\_cast.groupby(['Actor']).size().reset\_index(name='Total Content') actors=actors[actors.Actor !='No Cast Specified']

actors=actors.sort\_values(by=['Total Content'],ascending=False) actorsTop5=actors.head()

actorsTop5=actorsTop5.sort\_values(by=['Total Content'])

actorsTop5

**Actor Total Content**

**23624** Om Puri 27

**15541** Julie Tejwani 28

**30303** Takahiro Sakurai 30

**26941** Rupa Bhimani 31

**2612** Anupam Kher 39

Analysing content on Netflix:

df1=dff[['type','release\_year']]

df1=df1.rename(columns={"release\_year": "Release Year"})

df2=df1.groupby(['Release Year','type']).size().reset\_index(name='Total Content') df2=df2[df2['Release Year']>=2010]

df2

**Release Year type Total Content**

**95** 2010 Movie 154

**96** 2010 TV Show 40

**97** 2011 Movie 145

**98** 2011 TV Show 40

**99** 2012 Movie 173

**100** 2012 TV Show 64

**101** 2013 Movie 225

**102** 2013 TV Show 63

**103** 2014 Movie 264

**104** 2014 TV Show 88

**105** 2015 Movie 398

**106** 2015 TV Show 162

**107** 2016 Movie 658

**108** 2016 TV Show 244

**109** 2017 Movie 767

df1=dff[['type','release\_year']]

**110** 2017 TV Show 265

df1=df1.rename(columns={"release\_year": "Release Year"})

**111** 2018 Movie 767

df2=df1.groupby(['Release Year','type']).size().reset\_index(name='Total Content')

df2=df2[df2['Release Year']>=2010]

**112** 2018 TV Show 380

fig3 = px.line(df2, x="Release Year", y="Total Content", color='type',title='Trend of content produced over the years on Netflix') fig3.show()

**113** 2019 Movie 633

**114** 2019 TV Show 397

**115** 2020 Movie 517

**116** 2020 TV Show 436

**117** 2021 Movie 277

Trend of content produced over the years on Netflix

**118** 2021 TV Show 315

800type

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

700

600

Movie

TV Show

t

n

e

t

n

o

C

l

a

t

o

T

TextBlob import nltk

500

400

300

200

100

2010 2012 2014 2016 2018 2020 0 Release Year

from textblob import TextBlob

dfx=dff[['release\_year','description']]

dfx=dfx.rename(columns={'release\_year':'Release Year'}) for index,row in dfx.iterrows():

z=row['description']

testimonial=TextBlob(z)

p=testimonial.sentiment.polarity

if p==0:

sent='Neutral'

elif p>0:

sent='Positive'

else:

sent='Negative'

dfx.loc[[index,2],'Sentiment']=sent

dfx=dfx.groupby(['Release Year','Sentiment']).size().reset\_index(name='Total Content')

dfx=dfx[dfx['Release Year']>=2010]

fig4 = px.bar(dfx, x="Release Year", y="Total Content", color="Sentiment", title="Sentiment of content on Netflix") fig4.show()

Sentiment of content on Netflix

1200Sentiment Negative

1000

800

t

n

e

t

n

600

o

C

l

a

t

o

T

400

200

2010 2012 2014 2016 2018 2020 0 Release Year

Colab paid products - Cancel contracts here

check 18s completed at 4:47 PM

Neutral Positive