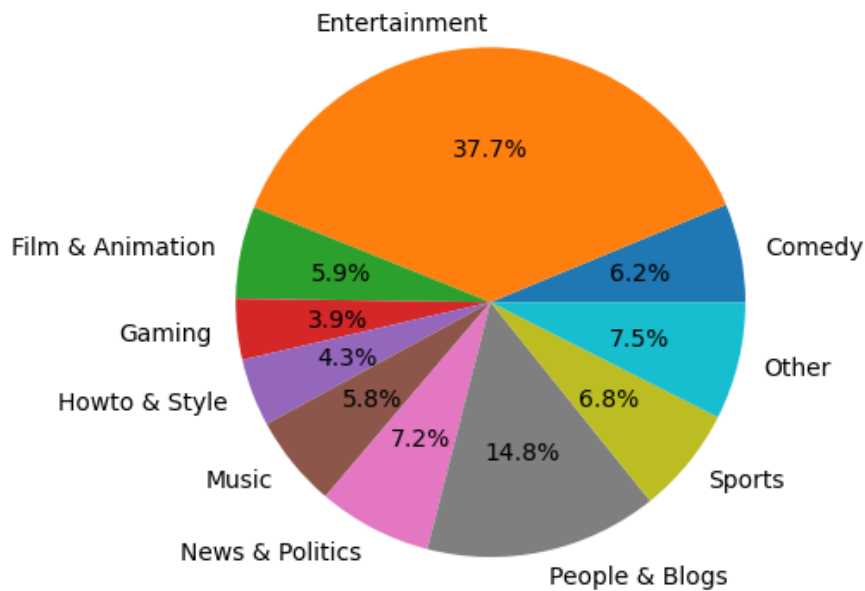# REPORT

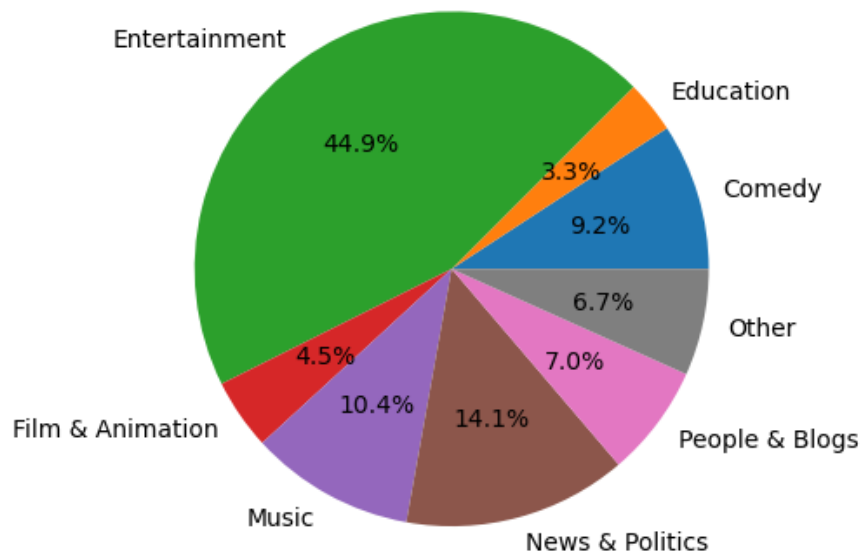**a.** Categorical differences in viewing patterns between Germans and Indians
   DE – Germany, IN - India

Video Categories in DE By Videos Trending

Entertainment 37.7%
Comedy 6.2%
Other 7.5%
Sports 6.8%
People & Blogs 14.8%
News & Politics 7.2%
Music 5.8%
Howto & Style 4.3%
Gaming 3.9%
Film & Animation 5.9%

Video Categories in IN By Videos Trending

Entertainment 44.9%
Education 3.3%
Comedy 9.2%
Other 6.7%
People & Blogs 7.0%
News & Politics 14.1%
Music 10.4%
Film & Animation 4.5%

The first comparison shows the ratio between popular categories by grouping categories together and summing the number of videos per category. In both countries entertainment is the most popular category. Indians are more interested politics and news, while Germans like to follow people and their blogs. Music is also a big trend with Indians, Germans like to focus more on sports.

Video Categories in DE By Views

Video Categories in IN By Views
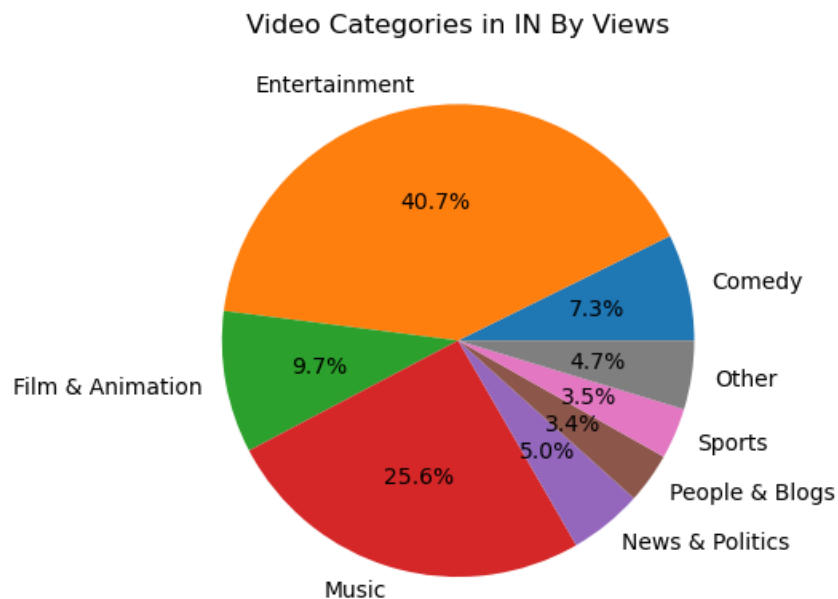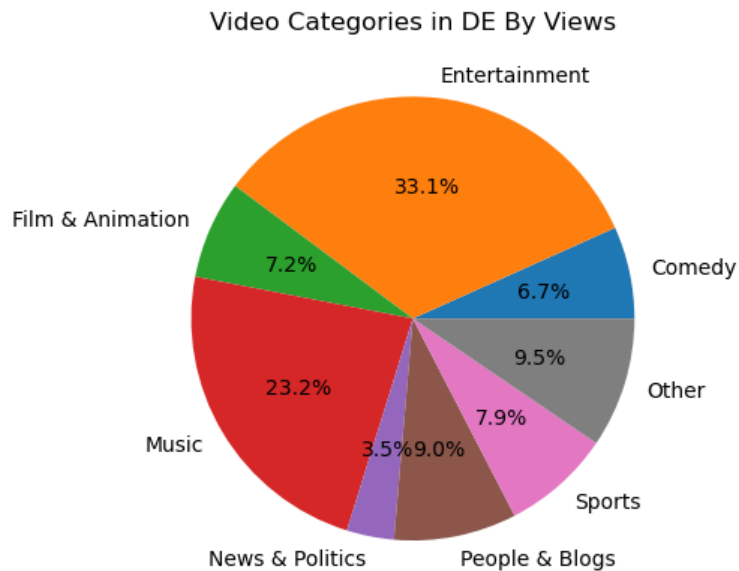
The second comparisons shows the ratio between categories based on the total views in videos that were trending in different categories. The viewing pattern is very similar by these two groups. In India the Entertainment is little higher than Germany, while Sports category is twice as popular in Germany than in India.

b. Ratios between views / likes / dislikes for different categories / countries.

Country	Most Liked Category	Most Disliked Category	Most Liked Category/ Views	Total Likes / (Likes + Dislikes)
Germany	Pets & Animals 97.61 %	Shows 20.68 %	Music 5.28%	94.3 %
India	Pets & Animals 98.23 %	Shows 22.32 %	Pets & Animals 7.43%	94.2 %
Canada	Pets & Animals 97.95 %	Shows 16.97 %	Comedy 5.29 %	95.3%
France	Travel & Events 98.76 %	Shows 26.29 %	Travel & Events 8.94%	95.6 %
Great Britain	Shows 99.24 %	Sports 11.69 %	How-to & Style 4.31 %	94.7 %
Japan	Music 97.8 %	Travel & Events 17.9 %	Music 7.81 %	95.6 %
South Korea	Trailers 100 %	Shows 11.18%	Music 6.33 %	95. 8 %
Mexico	Education 97.42 %	Shows 15.64 %	Comedy 6.95 %	95.6 %
Russia	How to & Style 95.26 %	News & Politics 37.54 %	Music 7.48 %	90.3 %
United States	Shows 97.79 %	News & Politics 18.72 %	Nonprofits & Activism 8.77 %	95.2 %

This is a summary of most liked and disliked category in a country. There is also the most liked category per view and the total like percentage. Following this report you can find like, dislike ratio for every category in every country.

Germany

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	97,11	2,89	3,37
Comedy	96,69	3,31	4,55
Education	96,98	3,02	5,22
Entertainment	92,25	7,75	3,3
Film & Animation	96,05	3,95	2,58
Gaming	93,57	6,43	4,31
How to & Style	95,33	4,67	2,56
Movies	96,1	3,9	2,08
Music	95,86	4,14	5,28
News & Politics	85,86	14,14	1,7
People & Blogs	93,06	6,94	2,78
Pets & Animals	97,61	2,39	3,64
Science & Technology	95,42	4,58	2,69
Shows	79,32	20,68	0,75
Sports	94,59	5,41	1,94
Trailers			0
Travel & Events	96,68	3,32	4,62
Total	94,31	5,69	3,59

India

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	96,03	3,97	1,78
Comedy	96,15	3,85	6,03
Education	96,8	3,2	5,13
Entertainment	93,52	6,48	2,46
Film & Animation	94,13	5,87	1,8
Gaming	95,15	4,85	2,44
How to & Style	90,17	9,83	1,07
Movies	95,09	4,91	1,35
Music	94,42	5,58	2,51
News & Politics	85,03	14,97	0,78
People & Blogs	93,07	6,93	1,7
Pets & Animals	98,24	1,76	7,43
Science & Technology	95,9	4,1	4,48
Shows	77,68	22,32	0,56
Sports	96,16	3,84	2,39
Travel & Events	96,49	3,51	1,23
Total	94,21	5,79	2,55

Canada

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	95,99	4,01	2,37
Comedy	97,33	2,67	5,29
Education	96,35	3,65	3,94
Entertainment	93,76	6,24	2,83
Film & Animation	96,15	3,85	2,65
Gaming	95,48	4,52	4,17
How to & Style	96,73	3,27	3,26
Movies	95,37	4,63	1,48
Music	96,23	3,77	4,28
News & Politics	90,93	9,07	2,03
People & Blogs	91,75	8,25	3,26
Pets & Animals	97,95	2,05	3,75
Science & Technology	95,98	4,02	3
Shows	83,03	16,97	0,82
Sports	94,76	5,24	1,98
Travel & Events	96,65	3,35	2,61
Total	95,27	4,73	3,44

France

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	96,85	3,15	3,67
Comedy	97,47	2,53	6,9
Education	97,63	2,37	6,88
Entertainment	91,95	8,05	3,13
Film & Animation	95,75	4,25	2,68
Gaming	95,35	4,65	5,03
Howto & Style	96,02	3,98	3,31
Movies	95,86	4,14	2,36
Music	96,6	3,4	5,52
News & Politics	92,3	7,7	1,88
People & Blogs	94,26	5,74	2,26
Pets & Animals	96,77	3,23	3,08
Science & Technology	97,31	2,69	4,01
Shows	73,71	26,29	0,56
Sports	95,35	4,65	2,2
Trailers	95,52	4,48	1,48
Travel & Events	98,76	1,24	8,94
Total	95,65	4,35	4,12

Great Britain

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	92,65	7,35	1,65
Comedy	96,87	3,13	3,62
Education	96,73	3,27	3,95
Entertainment	89,42	10,58	2,5
Film & Animation	95,97	4,03	1,76
Gaming	96,82	3,18	4,01
How to & Style	96,97	3,03	4,32
Music	95,92	4,08	2,19
News & Politics	89,02	10,98	2,18
People & Blogs	91,86	8,14	2,35
Pets & Animals	97,66	2,34	2,82
Science & Technology	95,15	4,85	1,94
Shows	99,24	0,76	2,93
Sports	88,31	11,69	2,24
Travel & Events	96,4	3,6	0,7
Total	94,71	5,29	2,27

Japan

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	93,16	6,84	1,26
Comedy	94,41	5,59	2,69
Education	96,34	3,66	5,48
Entertainment	93,78	6,22	2,25
Film & Animation	95,25	4,75	1,9
Gaming	93,13	6,87	1,75
Howto & Style	94,14	5,86	2,24
Music	97,8	2,2	7,81
News & Politics	84,67	15,33	0,89
People & Blogs	91,34	8,66	1,65
Pets & Animals	96,73	3,27	1,96
Science & Technology	95,92	4,08	1,97
Sports	93,63	6,37	0,96
Travel & Events	82,1	17,9	1,19
Total	95,65	4,35	3,08

South Korea

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	92,08	7,92	1,06
Comedy	96,08	3,92	1,77
Education	95,41	4,59	1,52
Entertainment	91,82	8,18	1,93
Film & Animation	96,32	3,68	1,38
Gaming	96,76	3,24	2,5
Howto & Style	96,61	3,39	2,06
Music	97,44	2,56	6,33
News & Politics	91,85	8,15	1,19
People & Blogs	95,53	4,47	1,44
Pets & Animals	98,43	1,57	2,18
Science & Technology	96	4	1,72
Shows	88,82	11,18	0,41
Sports	94,99	5,01	0,75
Trailers	100	0	0,02
Travel & Events	91,94	8,06	0,77
Total	95,77	4,23	2,87

Mexico

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	97,27	2,73	6,06
Comedy	95,84	4,16	6,96
Education	97,43	2,57	5,9
Entertainment	94,37	5,63	4,64
Film & Animation	97,2	2,8	3,88
Gaming	96,24	3,76	5,69
Howto & Style	96,89	3,11	5,29
Music	96,61	3,39	5,46
News & Politics	92,56	7,44	2,42
People & Blogs	93,7	6,3	3,89
Pets & Animals	93,96	6,04	2,17
Science & Technology	94,88	5,12	2,52
Shows	84,36	15,64	0,75
Sports	96,02	3,98	2,22
Travel & Events	97,16	2,84	5,82
Total	95,55	4,45	4,63

Russia

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	93,51	6,49	5,53
Comedy	90,72	9,28	6,85
Education	93,65	6,35	5,26
Entertainment	91,25	8,75	4,74
Film & Animation	94,17	5,83	3,54
Gaming	94,82	5,18	5,63
Howto & Style	95,26	4,74	3,28
Movies	92,62	7,38	1,73
Music	93,71	6,29	7,48
News & Politics	62,46	37,54	2,07
People & Blogs	86,13	13,87	4,68
Pets & Animals	94,95	5,05	3,96
Science & Technology	94,74	5,26	6,08
Shows	85,11	14,89	1,63
Sports	94,2	5,8	2,99
Travel & Events	89,41	10,59	4,14
Total	90,3	9,7	5,11

United States

category_name	likes/(dislikes + likes)[%]	dislikes/(dislikes + likes)[%]	likes/views[%]
Autos & Vehicles	94,59	5,41	0,82
Comedy	96,77	3,23	4,23
Education	97,33	2,67	4,17
Entertainment	92,5	7,5	2,57
Film & Animation	96,47	3,53	2,28
Gaming	88,26	11,74	3,22
Howto & Style	96,75	3,25	3,99
Music	96,51	3,49	3,53
News & Politics	81,28	18,72	1,23
Nonprofits & Activism	81,74	18,26	8,77
People & Blogs	94,82	5,18	3,8
Pets & Animals	97,35	2,65	2,53
Science & Technology	94,78	5,22	2,37
Shows	97,79	2,21	2,1
Sports	95,05	4,95	2,24
Travel & Events	93,42	6,58	1,41
Total	95,24	4,76	3,15

c. Channels that are popular in most countries.

We were searching for channels that are popular in most countries, which means more than half countries. I decided that a popular channel is in the top 50 channels in a country by views.

Number of Countries Popular	Channel_title
10	Marvel Entertainment, TaylorSwiftVEVO, Sony Pictures Entertainment
9	ChildishGambinoVEVO, Maroon5VEVO, ibighit, YouTube Spotlight, 20th Century Fox
8	Ed Sheeran
7	Warner Bros. Pictures, ArianaGrandeVevo, PewDiePie, Universal Pictures, Kylie Jenner, Dude Perfect
6	SpaceX, FoxStarHindi, WWE, MalumaVEVO, Cardi B

This is the end of report. Now you can check the code I used to create this report.

In [74]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import os

# When importing the dataset for every country you have to change the read_csv t

# load the data
COvideos = pd.read_csv('DEvideos.csv')
CO_category_id = pd.read_csv('DE_category_id.csv')
filename = os.path.basename('DEvideos.csv')
title = filename[:2]

# merge the data
CO_category_id = CO_category_id.rename(columns={'items__id': 'category_id'})
COdataset = pd.merge(CO_category_id, COvideos, on='category_id')
COdataset1 = pd.merge(CO_category_id, COvideos, on='category_id')
COdata = COdataset.loc[:, ['channel_title','items__snippet__title', 'category_id
                           'views', 'likes', 'dislikes', 'comment_count', 'comme
                           'ratings_disabled', 'video_error_or_removed']]
COdata = COdata.rename(columns={'items__snippet__title': 'category_name'})

# group by category and count the number of entries
category_count = COdata.groupby(['category_name']).size()

# calculate the percentage of each category and filter out categories with less
category_percent = category_count / len(COdata)
category_percent_filtered = category_percent[category_percent >= 0.03]
category_percent_filtered['Other'] = category_percent[category_percent < 0.03].s

# plot the pie chart
plt.pie(category_percent_filtered, labels=category_percent_filtered.index, autop
plt.title('Video Categories in ' + title + ' By Videos Trending')
plt.show()

total_count = COdata['views'].sum()

# filter out categories with less than 3% of the total count
threshold = 0.03 * total_count
COcategorycount = COdata.groupby(['category_name']).sum()
COcategorycount = COcategorycount[COcategorycount['views'] >= threshold]

# group the remaining categories into an "Other" category
other_count = total_count - COcategorycount['views'].sum()
COcategorycount.loc['Other'] = other_count

# create the pie chart
fig, ax = plt.subplots()
ax.pie(COcategorycount['views'], labels=COcategorycount.index, autopct='%1.1f%%'
ax.set_title('Video Categories in ' + title + ' By Views')

plt.show()

#Creating a total row
COcategorized = COdata.groupby(['category_name'])[['views','likes','dislikes']].
grouped = COcategorized.groupby('category_name').sum()
total_row = pd.DataFrame({'category_name': 'Total',
                          'views': grouped['views'].sum(),
                          'likes': grouped['likes'].sum(),
```

```python
                              'dislikes': grouped['dislikes'].sum()}, index=[0])
COcategorized = pd.concat([COcategorized, total_row])

#Calculating the ratios
COcategorized['likes/(dislikes + likes)[%]'] = ((COcategorized['likes'] / (COcat
COcategorized['dislikes/(dislikes + likes)[%]'] = ((COcategorized['dislikes'] /
COcategorized['likes/views[%]'] = (COcategorized['likes'] / COcategorized['views

#Plot the second pie chart
total_likes = COcategorized['likes'].sum()
total_dislikes = COcategorized['dislikes'].sum()
fig, ax = plt.subplots()
ax.pie([total_likes, total_dislikes], labels=['Likes', 'Dislikes'], autopct='%1.
ax.set_title('Likes vs. Dislikes Ratio')
plt.show()

#Exporting the ratio calculations for every Country
COratios = COcategorized[['category_name','likes/(dislikes + likes)[%]','dislike
COratios.to_excel('USratios.xlsx', sheet_name='Sheet1', index=False)
```
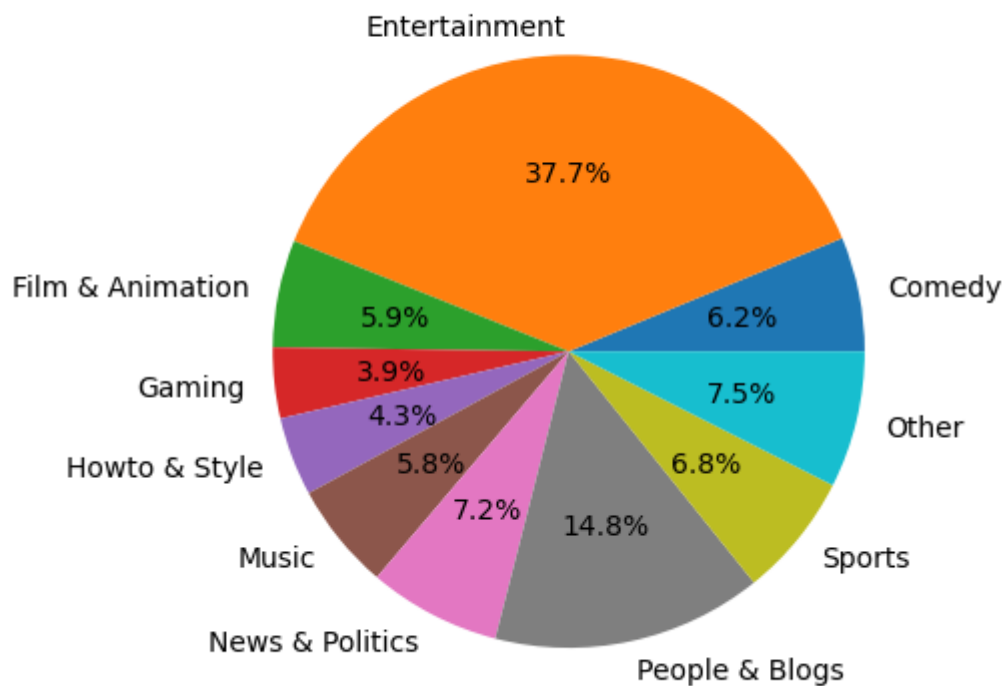
Video Categories in DE By Videos Trending

Entertainment 37.7%
Comedy 6.2%
Other 7.5%
Sports 6.8%
People & Blogs 14.8%
News & Politics 7.2%
Music 5.8%
Howto & Style 4.3%
Gaming 3.9%
Film & Animation 5.9%

```
C:\Users\rokzu\AppData\Local\Temp\ipykernel_14752\3819398493.py:37: FutureWarni
ng: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In
a future version, numeric_only will default to False. Either specify numeric_on
ly or select only columns which should be valid for the function.
  COcategorycount = COdata.groupby(['category_name']).sum()
```

Video Categories in DE By Views
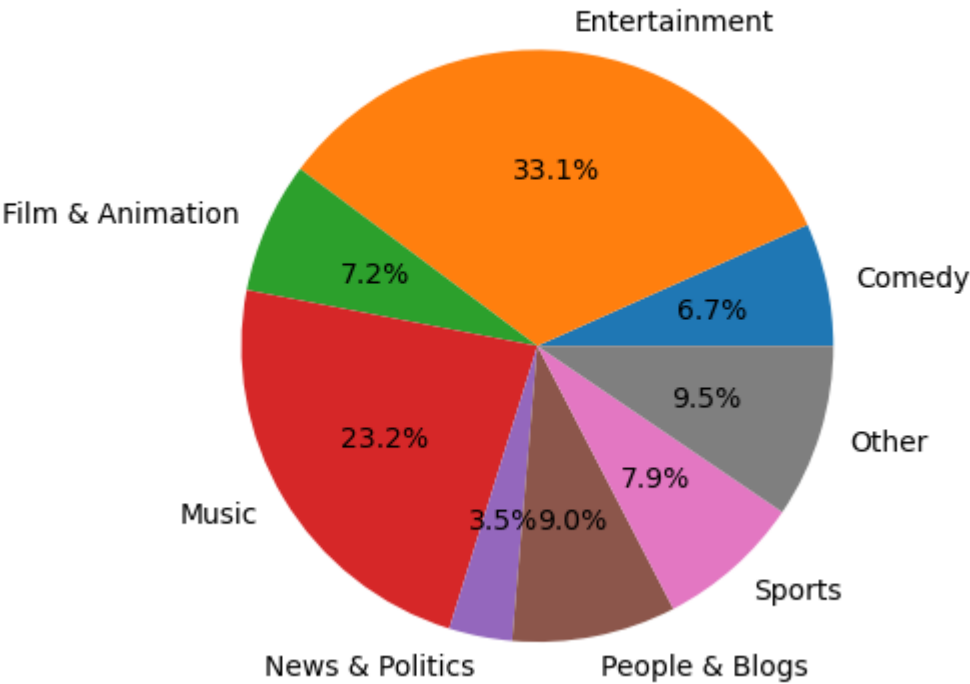
Likes vs. Dislikes Ratio
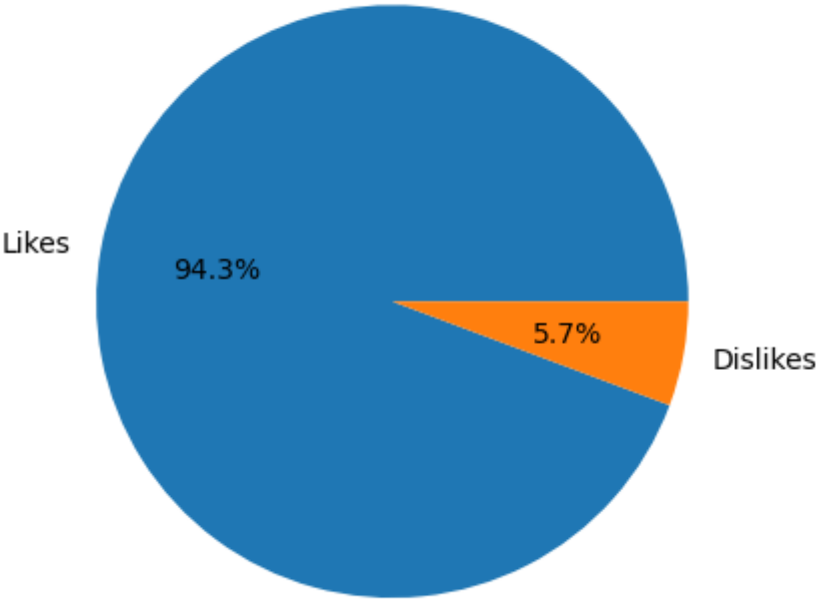
In [2]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import os

# load the data
COvideos = pd.read_csv('USvideos.csv')
CO_category_id = pd.read_csv('US_category_id.csv')


# merge the data
CO_category_id = CO_category_id.rename(columns={'items__id': 'category_id'})
COdataset = pd.merge(CO_category_id, COvideos, on='category_id')
COdataset1 = pd.merge(CO_category_id, COvideos, on='category_id')
COdata = COdataset.loc[:, ['channel_title','items__snippet__title', 'category_id
                           'views', 'likes', 'dislikes', 'comment_count', 'comme
                           'ratings_disabled', 'video_error_or_removed']]
COdata = COdata.rename(columns={'items__snippet__title': 'category_name'})

COchannels = COdata.groupby(['channel_title'])[['views','likes','dislikes']].sum
COchannels = COchannels.sort_values('views', ascending = False)
COchannels = COchannels.head(50).reset_index()
COchannelsall = pd.DataFrame() # Only run one time when importing DE dataset
COchannelsall = COchannelsall.append(COchannels)

# I ran this part of the code and up 10 times to get complete list of 500 channe
```

```
C:\Users\rokzu\AppData\Local\Temp\ipykernel_20440\2241091999.py:23: FutureWarni
ng: The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.
  COchannelsall = COchannelsall.append(COchannels)
```

In [3]:
```python
COchannelsall1 = COchannelsall.groupby(['channel_title']).size().sort_values(asc
COchannelsall1 = COchannelsall1.rename(columns={0: 'Number of Countries Popular'
COchannelsall_test = COchannelsall1.groupby('Number of Countries Popular')['chan
COchannelsall_test.to_excel('PopularChannels.xlsx', sheet_name='Sheet1', index=F
```