

```
In [1]: # Import Packages  
import numpy as np  
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
In [2]: # Import Data  
df=pd.read_csv("C:\\\\Users\\\\Dell\\\\OneDrive\\\\Desktop\\\\STATISTICS\\\\Data analysis\\\\project\\\\Netflix\\\\Netflix Dataset.csv")
```

```
In [3]: # EDA  
df.head()
```

Out[3]:

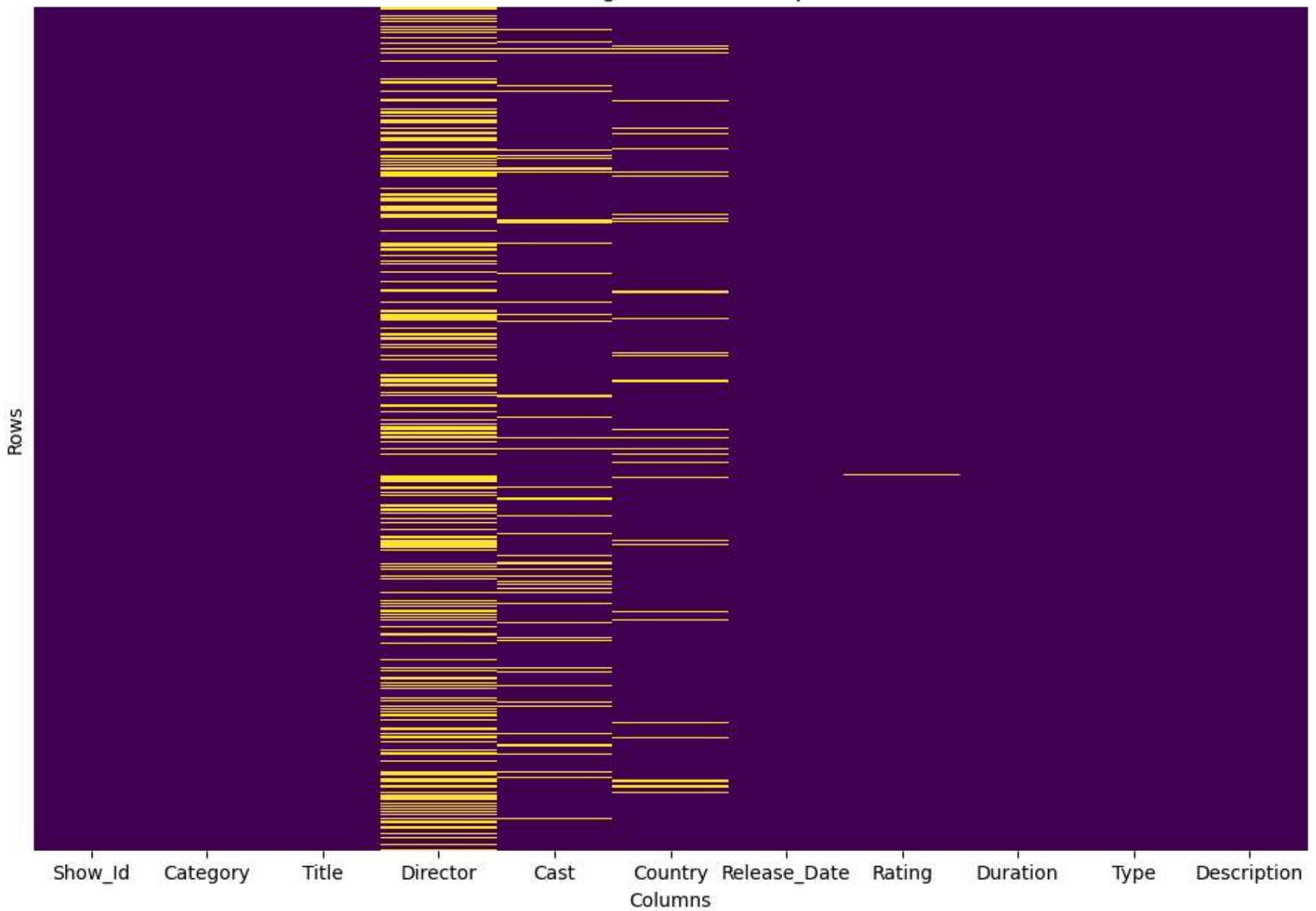
	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...

In [4]: `df.info() # All string and there exists null values`

```
<class 'pandas.DataFrame'>
RangeIndex: 7789 entries, 0 to 7788
Data columns (total 11 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Show_Id     7789 non-null   str    
 1   Category    7789 non-null   str    
 2   Title       7789 non-null   str    
 3   Director    5401 non-null   str    
 4   Cast        7071 non-null   str    
 5   Country     7282 non-null   str    
 6   Release_Date 7779 non-null   str    
 7   Rating      7782 non-null   str    
 8   Duration    7789 non-null   str    
 9   Type        7789 non-null   str    
 10  Description 7789 non-null   str    
dtypes: str(11)
memory usage: 3.3 MB
```

```
In [5]: # heatmap of null values
plt.figure(figsize=(12, 8))
sns.heatmap(df.isnull(), cbar=False, yticklabels=False, cmap='viridis')
plt.title('Missing Values Heatmap')
plt.xlabel('Columns')
plt.ylabel('Rows')
plt.show()
```

### Missing Values Heatmap



```
In [6]: print(f"Missing Values Percentage:\n{round(df.isna().sum()*100/df.shape[0],2)}")
```

```
Missing Values Percentage:  
Show_Id      0.00  
Category     0.00  
Title        0.00  
Director    30.66  
Cast         9.22  
Country      6.51  
Release_Date 0.13  
Rating       0.09  
Duration     0.00  
Type         0.00  
Description   0.00  
dtype: float64
```

```
In [7]: df.shape
```

```
Out[7]: (7789, 11)
```

```
In [8]: df['Director'] = df['Director'].fillna('Not Specified')  
df['Cast'] = df['Cast'].fillna('No Cast Information')  
df['Country'] = df['Country'].fillna('Unknown Country')  
df = df.dropna(subset=['Release_Date', 'Rating']) # Drop rows with missing Release_Date or Rating  
  
# Verify no missing values remain  
print(df.isnull().sum())  
print(f"Final shape: {df.shape}")
```

```
Show_Id      0  
Category     0  
Title        0  
Director     0  
Cast         0  
Country      0  
Release_Date 0  
Rating        0  
Duration      0  
Type          0  
Description   0  
dtype: int64  
Final shape: (7772, 11)
```

```
In [9]: # Duplicate Record check  
df.duplicated().sum()
```

```
Out[9]: 2
```

```
In [10]: df = df.drop_duplicates()  
df.duplicated().sum()
```

```
Out[10]: 0
```

```
In [25]: # For 'House of Cards', what is the Show Id and Who is the Director of this show  
result = df[df['Title']=='House of Cards'][['Show_Id','Director']].values.tolist()  
print(f"Show Id: {result[0][0]}\nDirector: {result[0][1]}")
```

```
Show Id: s2833  
Director: Robin Wright, David Fincher, Gerald McRaney, John David Coles, David Manson
```

```
In [28]: df.head()
```

Out[28]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
0	s1	TV Show	3%	Not Specified	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...

In [37]:

```
# highest number of the TV Shows & Movies release year by graph
# Clean the Release_Date column first, then convert to datetime
df['Release_Date'] = df['Release_Date'].str.strip() # Remove Leading/trailing spaces
df['Release_Year'] = pd.to_datetime(df['Release_Date']).dt.year

plt.figure(figsize=(14, 6))
df['Release_Year'].value_counts().sort_index().plot(kind='bar', color='crimson')
plt.title('Netflix: Number of Titles Released by Year', fontsize=16, fontweight='bold', pad=20)
plt.xlabel('Release Year', fontsize=18)
plt.ylabel('Number of Titles', fontsize=18)
plt.xticks(rotation=45)
plt.grid(axis='y', alpha=0.3)
```

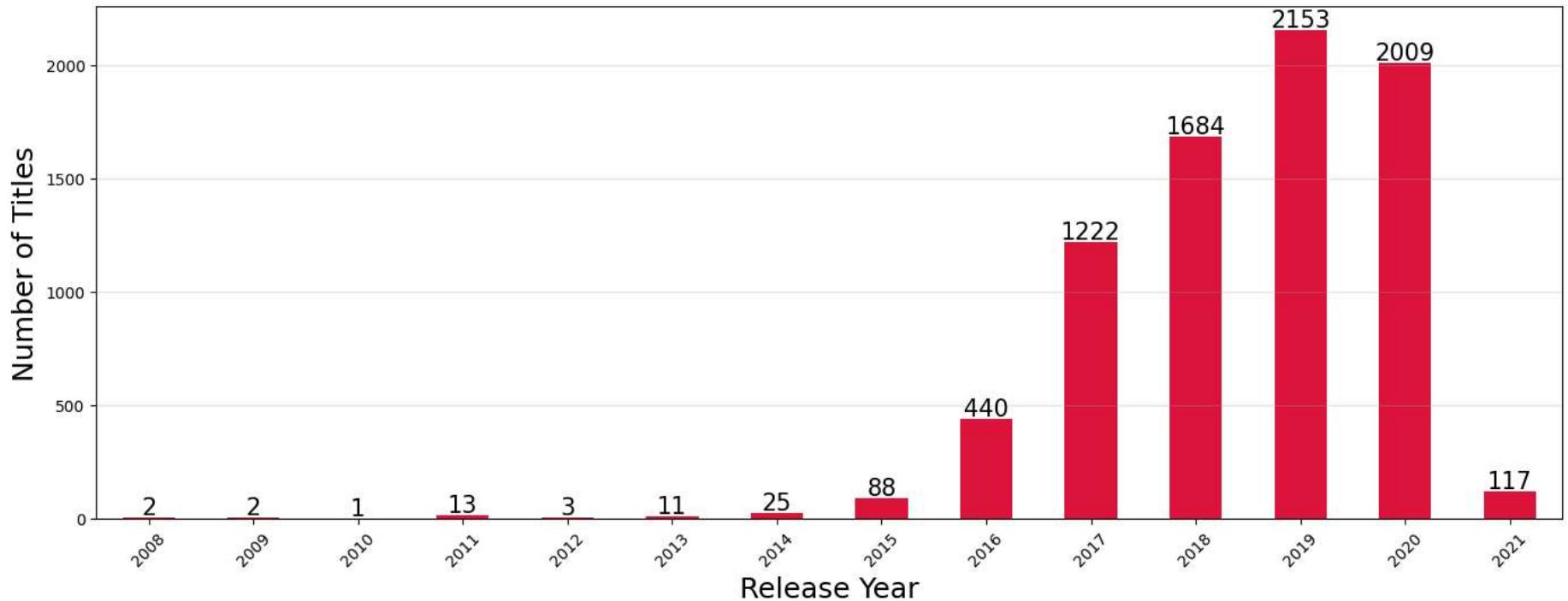
```

# Add value labels on bars
for i, v in enumerate(df['Release_Year'].value_counts().sort_index()):
    plt.text(i, v + 15, str(v), ha='center', fontsize=15)

plt.tight_layout()
plt.show()

```

**Netflix: Number of Titles Released by Year**



```

In [54]: # Number of Movies & TV Shows are in the dataset
category_counts = df['Category'].value_counts()

plt.figure(figsize=(8, 6))
ax = sns.countplot(data=df, x='Category', hue='Category',
                    palette=['Teal', 'Crimson'],
                    order=category_counts.index,
                    legend=False)
plt.title('Number of Movies vs TV Shows on Netflix', fontsize=16, fontweight='bold')

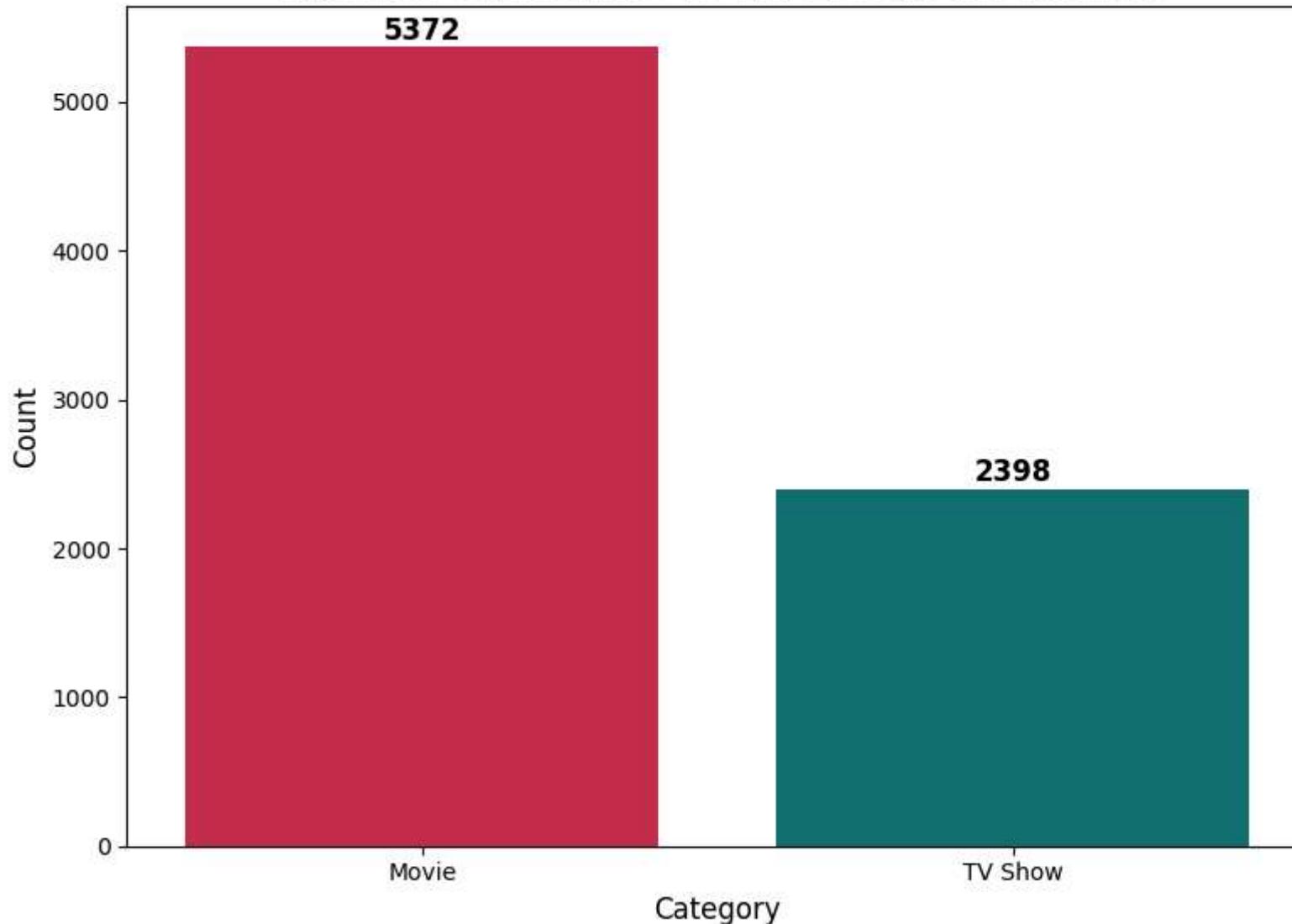
```

```
plt.xlabel('Category', fontsize=12)
plt.ylabel('Count', fontsize=12)

# Clean, professional label placement
for i, (category, count) in enumerate(category_counts.items()):
    plt.text(i, count + 50, str(count), ha='center', fontsize=12, fontweight='bold')

plt.tight_layout()
plt.show()
```

## Number of Movies vs TV Shows on Netflix



```
In [62]: # Movies that were released in year 2012.  
df[df['Release_Year']==2012]
```

Out[62]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Released
818	s819	Movie	Being Elmo: A Puppeteer's Journey	Constance Marks	Kevin Clash, Whoopi Goldberg	United States	February 21, 2012	PG	76 min	Documentaries	Whoopi Goldberg narrates Elmo creator Kevin Cl...	
1229	s1230	Movie	Casa de mi Padre	Matt Piedmont	Will Ferrell, Gael García Bernal, Diego Luna, ...	United States, Mexico	November 14, 2012	R	84 min	Comedies	Will Ferrell stars as a Spanish-speaking cowbo...	
3466	s3467	Movie	Kung Fu Panda: Holiday	Tim Johnson	Jack Black, Angelina Jolie, Dustin Hoffman, Ja...	United States	December 1, 2012	TV-PG	26 min	Children & Family Movies, Comedies	As preparations for the Winter Feast build, Po...	

In [66]:

```
# Released in Bangladesh only
df[df['Country']=='Bangladesh']
```

Out[66]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_
3435	s3436	Movie	Komola Rocket	Noor Imran Mithu	Tauquir Ahmed, Mosharraf Karim, Joyraj, Samia ...	Bangladesh	January 7, 2019	TV-14	95 min	Dramas, Independent Movies, International Movies	As the lives of rich and poor passengers aboar...	2
5621	s5621	Movie	Sincerely Yours, Dhaka	Nuhash Humayun, Syed Ahmed Shawki, Rahat Rahma...	Mostafizur Noor Imran, Orchita Sporshia, Allen...	Bangladesh	December 16, 2019	TV-MA	136 min	Comedies, Dramas, Independent Movies	Eleven emerging Bangladeshi filmmakers present...	2

In [83]:

```
# Top Directors
# Split multiple director names
directors_split = df[df['Director'] != 'Not Specified']['Director'].str.split(', ').explode()
top_individual_directors = directors_split.value_counts().head(10)

plt.figure(figsize=(14, 8)) # Bigger figure size
bars = plt.barh(range(len(top_individual_directors)), top_individual_directors.values, color='Steelblue')
plt.yticks(range(len(top_individual_directors)), top_individual_directors.index, fontsize=14, fontweight='bold') # Larger y-axis ticks
plt.title('Top 10 Individual Directors with Most Collaborations on Netflix', fontsize=18, fontweight='bold')
plt.xlabel('Number of Titles', fontsize=14)
plt.ylabel('Director', fontsize=14)

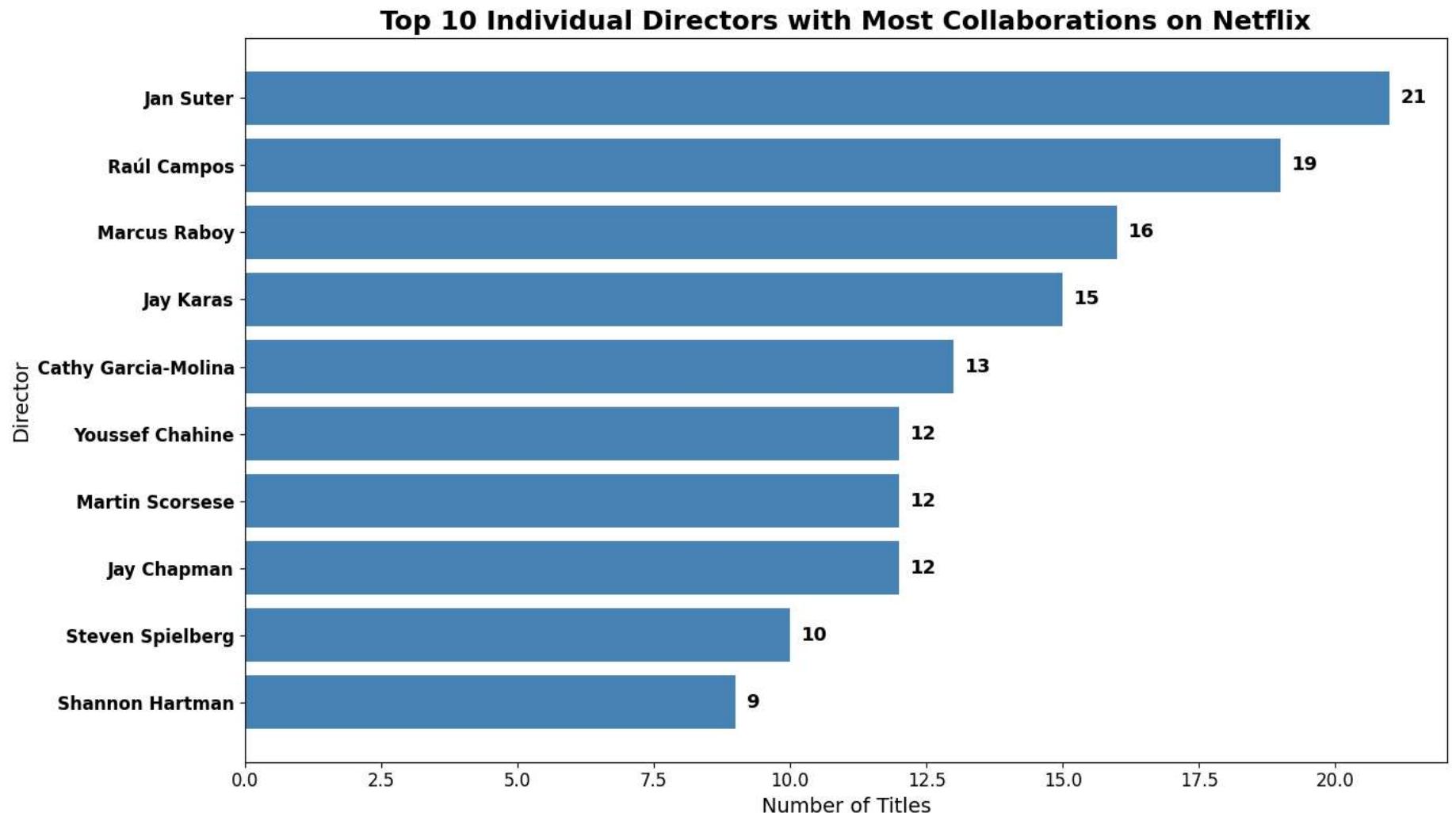
# Make axis tick labels more visible
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)

plt.gca().invert_yaxis()

# Add value labels with larger font
for bar in bars:
    width = bar.get_width()
    plt.text(width+0.2, bar.get_y() + bar.get_height()/2,
```

```
f'{int(width)}', ha='left', va='center', fontsize=13, fontweight='bold')

plt.tight_layout()
plt.show()
```



```
In [84]: # Top cast
# Split multiple cast names
cast_split = df[df['Cast'] != 'No Cast Information']['Cast'].str.split(', ').explode()
top_individual_casts = cast_split.value_counts().head(10)
```

```
plt.figure(figsize=(14, 8)) # Bigger figure size
bars = plt.barh(range(len(top_individual_casts)), top_individual_casts.values, color='teal')
plt.yticks(range(len(top_individual_casts)), top_individual_casts.index, fontsize=14, fontweight='bold') # Larger y-axis labels
plt.title('Top 10 Individual casts with Most Collaborations on Netflix', fontsize=18, fontweight='bold')
plt.xlabel('Number of Titles', fontsize=14)
plt.ylabel('casts', fontsize=14)

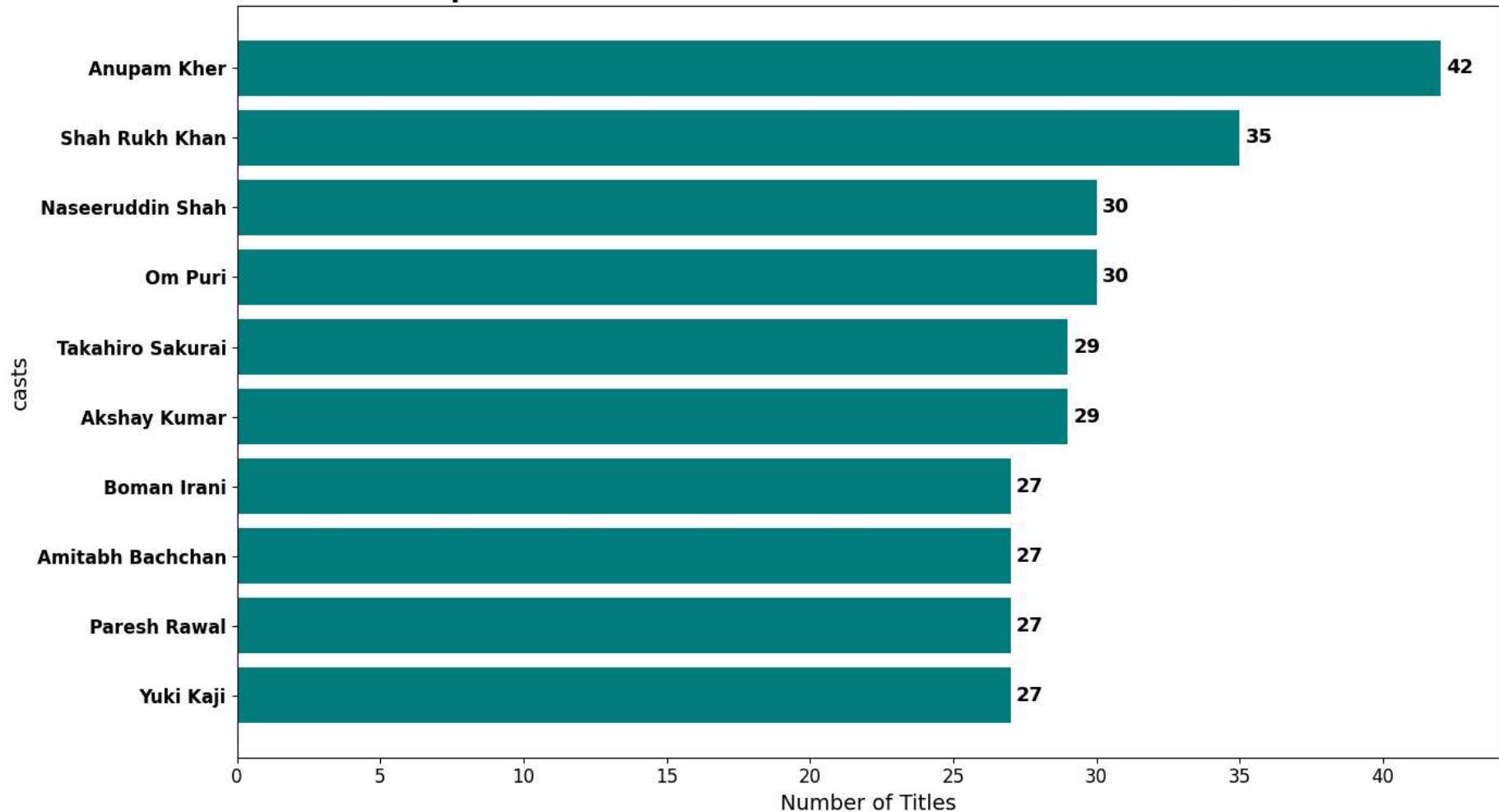
# Make axis tick labels more visible
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)

plt.gca().invert_yaxis()

# Add value labels with larger font
for bar in bars:
    width = bar.get_width()
    plt.text(width+0.2 , bar.get_y() + bar.get_height()/2,
             f'{int(width)}', ha='left', va='center', fontsize=13, fontweight='bold')

plt.tight_layout()
plt.show()
```

## Top 10 Individual casts with Most Collaborations on Netflix



```
In [96]: # In how many movies/shows, Tom Cruise was cast ?
df[df['Cast'].str.contains('Tom Cruise')]
```

Out[96]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year
3860	s3861	Movie	Magnolia	Paul Thomas Anderson	Jeremy Blackman, Tom Cruise, Melinda Dillon, A...	United States	January 1, 2020	R	189 min	Dramas, Independent Movies	Through chance, human action, past history and...	2020
5071	s5071	Movie	Rain Man	Barry Levinson	Dustin Hoffman, Tom Cruise, Valeria Golino, Ge...	United States	July 1, 2019	R	134 min	Classic Movies, Dramas	A fast-talking yuppie is forced to slow down w...	2019

In [100...]

```
# What are the different Ratings defined by Netflix ?
df['Rating'].unique()
```

Out[100...]

```
<ArrowStringArray>
[    'TV-MA',        'R',      'PG-13',      'TV-14',      'TV-PG',        'NR',
    'TV-G',        'TV-Y',      'TV-Y7',       'PG',        'G',      'NC-17',
    'TV-Y7-FV',        'UR']
Length: 14, dtype: str
```

In [117...]

```
# Create crosstab of Rating vs Category
rating_category = pd.crosstab(df['Rating'], df['Category'])
rating_category = rating_category.sort_values('Movie', ascending=True)

# ===== FIGURE 1: MOVIES ONLY =====
plt.figure(figsize=(16, 10))
bars1 = plt.barh(range(len(rating_category)), rating_category['Movie'],
                 color='#E50914', height=0.7, edgecolor='white', linewidth=2)

plt.yticks(range(len(rating_category)), rating_category.index, fontsize=13, fontweight='bold')
plt.xticks(fontsize=12, fontweight='bold')
plt.title('NETFLIX MOVIES - Distribution by Rating', fontsize=22, fontweight='bold', pad=20)
plt.xlabel('NUMBER OF MOVIES', fontsize=15, fontweight='bold')
```

```

plt.ylabel('RATING', fontsize=15, fontweight='bold')
plt.grid(axis='x', alpha=0.2, linestyle='--')

for i, bar in enumerate(bars1):
    count = rating_category['Movie'].iloc[i]
    if count > 0:
        if count > 100:
            plt.text(count/2, bar.get_y() + bar.get_height()/2, f'{count:,}', 
                      ha='center', va='center', fontsize=13, fontweight='bold', color='white')
        else:
            plt.text(count + 20, bar.get_y() + bar.get_height()/2, f'{count:,}', 
                      ha='left', va='center', fontsize=13, fontweight='bold', color="#E50914")

plt.tight_layout()
plt.show()

# ===== FIGURE 2: TV SHOWS ONLY =====
plt.figure(figsize=(16, 10))
bars2 = plt.barh(range(len(rating_category)), rating_category['TV Show'],
                 color='#1E1E1E', height=0.7, edgecolor='white', linewidth=2)

plt.yticks(range(len(rating_category)), rating_category.index, fontsize=13, fontweight='bold')
plt.xticks(fontsize=12, fontweight='bold')
plt.title('NETFLIX TV SHOWS - Distribution by Rating', fontsize=22, fontweight='bold', pad=20)
plt.xlabel('NUMBER OF TV SHOWS', fontsize=15, fontweight='bold')
plt.ylabel('RATING', fontsize=15, fontweight='bold')
plt.grid(axis='x', alpha=0.2, linestyle='--')

for i, bar in enumerate(bars2):
    count = rating_category['TV Show'].iloc[i]
    if count > 0:
        if count > 100:
            plt.text(count/2, bar.get_y() + bar.get_height()/2, f'{count:,}', 
                      ha='center', va='center', fontsize=13, fontweight='bold', color='white')
        else:
            plt.text(count + 15, bar.get_y() + bar.get_height()/2, f'{count:,}', 
                      ha='left', va='center', fontsize=13, fontweight='bold', color="#1E1E1E")

plt.tight_layout()
plt.show()

```

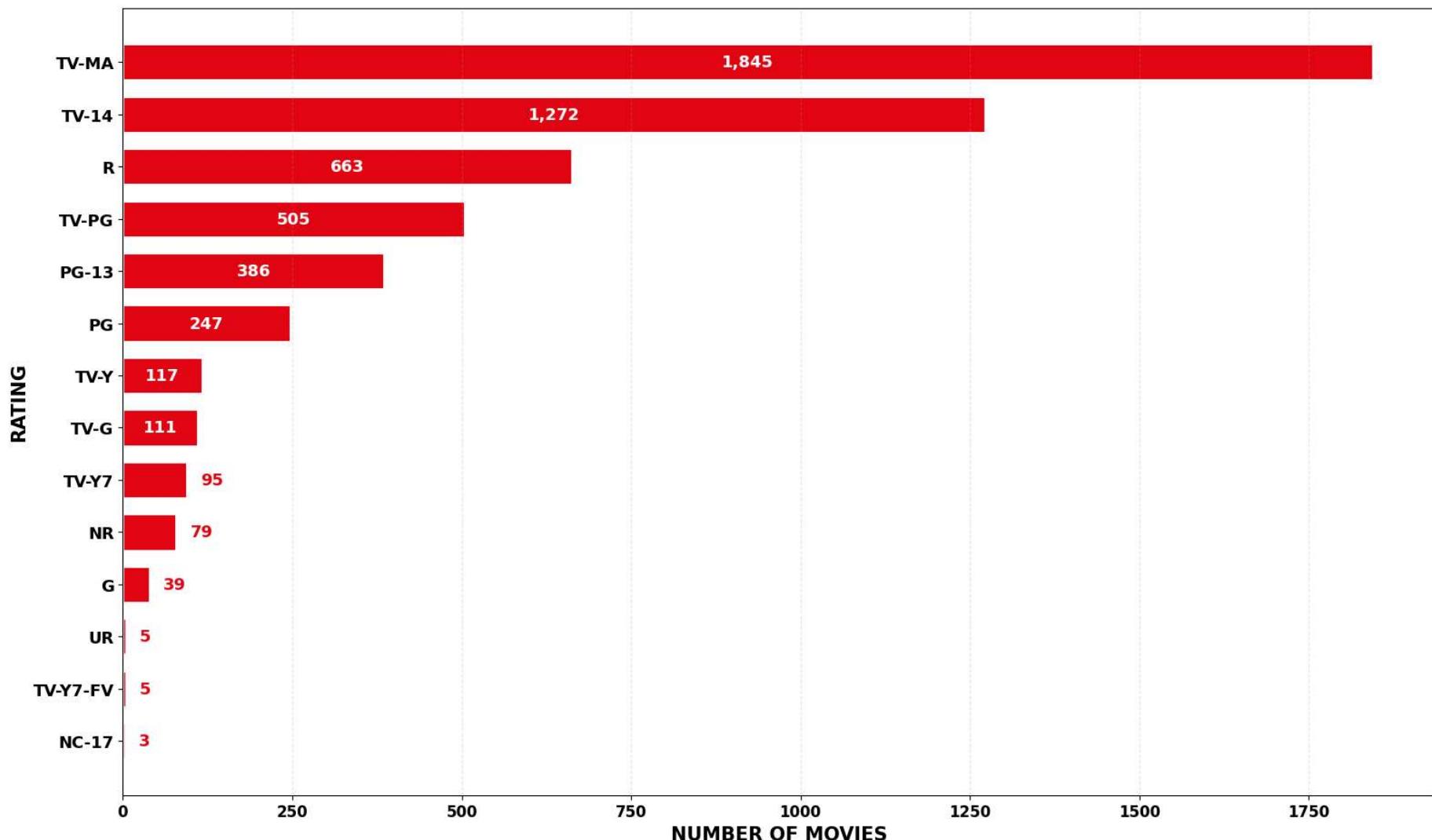
```
# ===== FIGURE 3: STACKED TOGETHER =====
plt.figure(figsize=(16, 10))
plt.barh(range(len(rating_category)), rating_category['Movie'],
         color='#E50914', height=0.7, edgecolor='white', linewidth=1.5, label='Movies')
plt.barh(range(len(rating_category)), rating_category['TV Show'],
         left=rating_category['Movie'], color="#2C2C2C", height=0.7,
         edgecolor='white', linewidth=1.5, label='TV Shows')

plt.yticks(range(len(rating_category)), rating_category.index, fontsize=13, fontweight='bold')
plt.xticks(fontsize=12, fontweight='bold')
plt.title('NETFLIX - Movies vs TV Shows by Rating', fontsize=20, fontweight='bold', pad=20)
plt.xlabel('Number of Titles', fontsize=14, fontweight='bold')
plt.ylabel('Rating', fontsize=14, fontweight='bold')
plt.legend(fontsize=12, loc='lower right')
plt.grid(axis='x', alpha=0.2, linestyle='--')

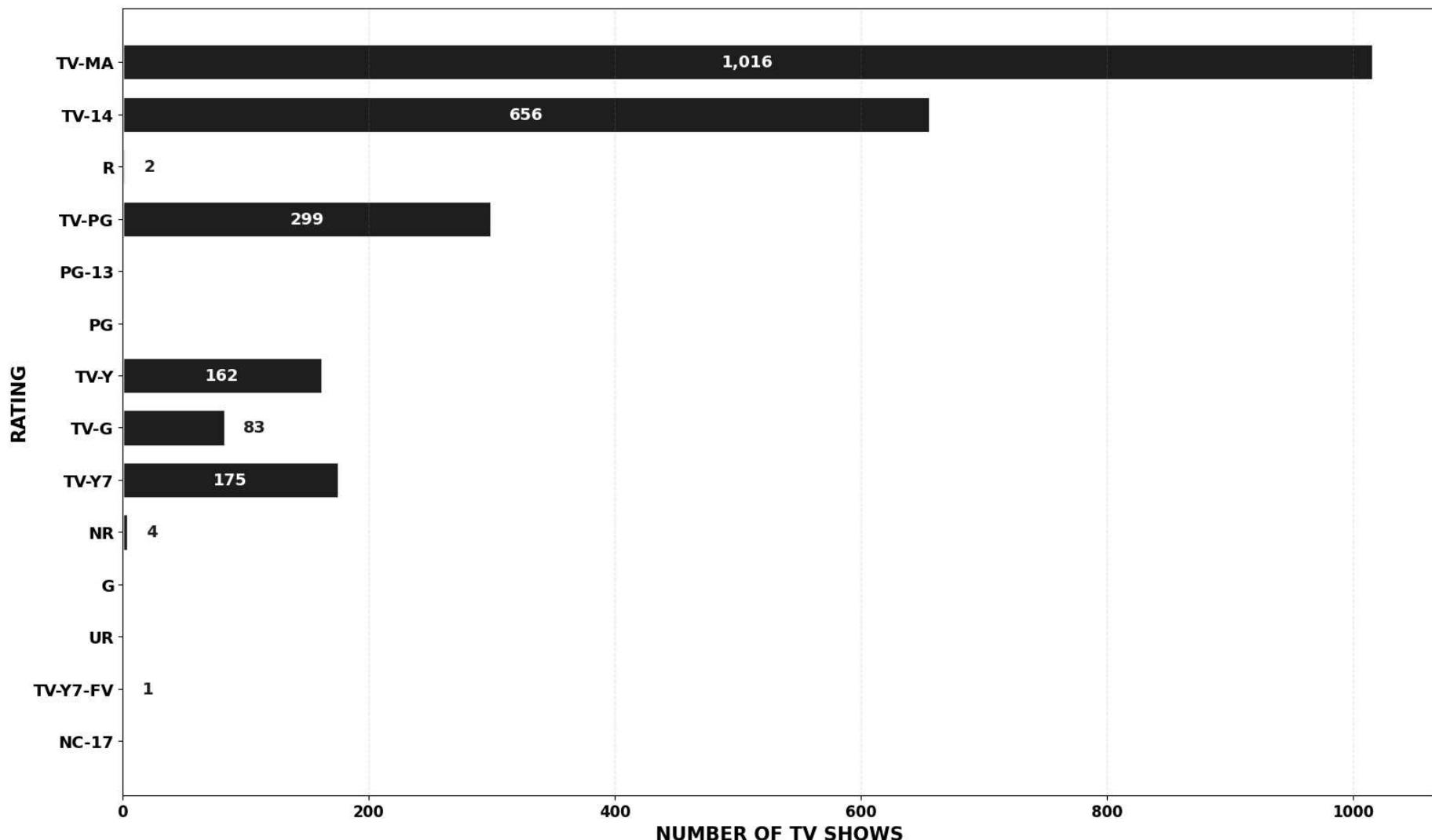
for i, rating in enumerate(rating_category.index):
    total = rating_category.loc[rating].sum()
    plt.text(total + 30, i, f'{total:,}', ha='left', va='center', fontsize=13, fontweight='bold')

plt.tight_layout()
plt.show()
```

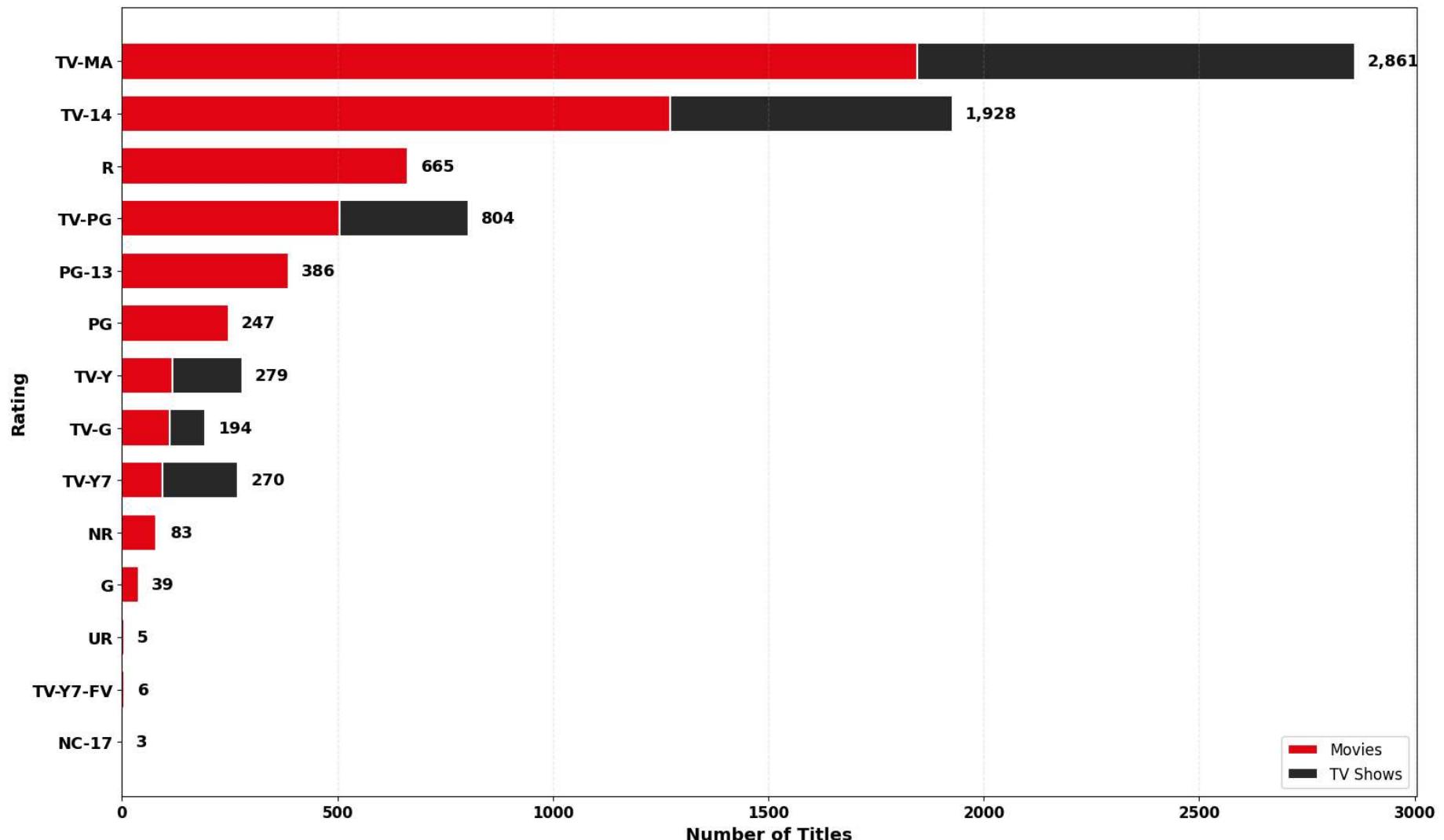
## NETFLIX MOVIES - Distribution by Rating



## NETFLIX TV SHOWS - Distribution by Rating



## NETFLIX - Movies vs TV Shows by Rating



In [129]:

```
# 'R' rated TV Shows after year 2018
df[(df['Rating']=='R') & (df['Release_Year']>2018) & (df['Category']=='TV Show')]
```

Out[129...]

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year
6437	s6436	TV Show	The Hateful Eight: Extended Version	Quentin Tarantino	Samuel L. Jackson, Kurt Russell, Jennifer Jason... ...	Unknown Country	April 25, 2019	R	1 Season	TV Shows	Trapped at a stagecoach stop as a storm rages	2019

In [131...]

```
# 'R' rated Movie and TV Shows plot

# Filter for 'R' rated content
r_rated = df[df['Rating'] == 'R']

# Count 'R' rated content by year and category
r_by_year = r_rated.groupby(['Release_Year', 'Category']).size().unstack().fillna(0)

# LINE CHART - Clean and simple
plt.figure(figsize=(16, 8))

plt.plot(r_by_year.index, r_by_year['Movie'], marker='o', linewidth=3, markersize=8,
          color='Crimson', label='R-Rated Movies')
plt.plot(r_by_year.index, r_by_year['TV Show'], marker='s', linewidth=3, markersize=8,
          color='Teal', label='R-Rated TV Shows')

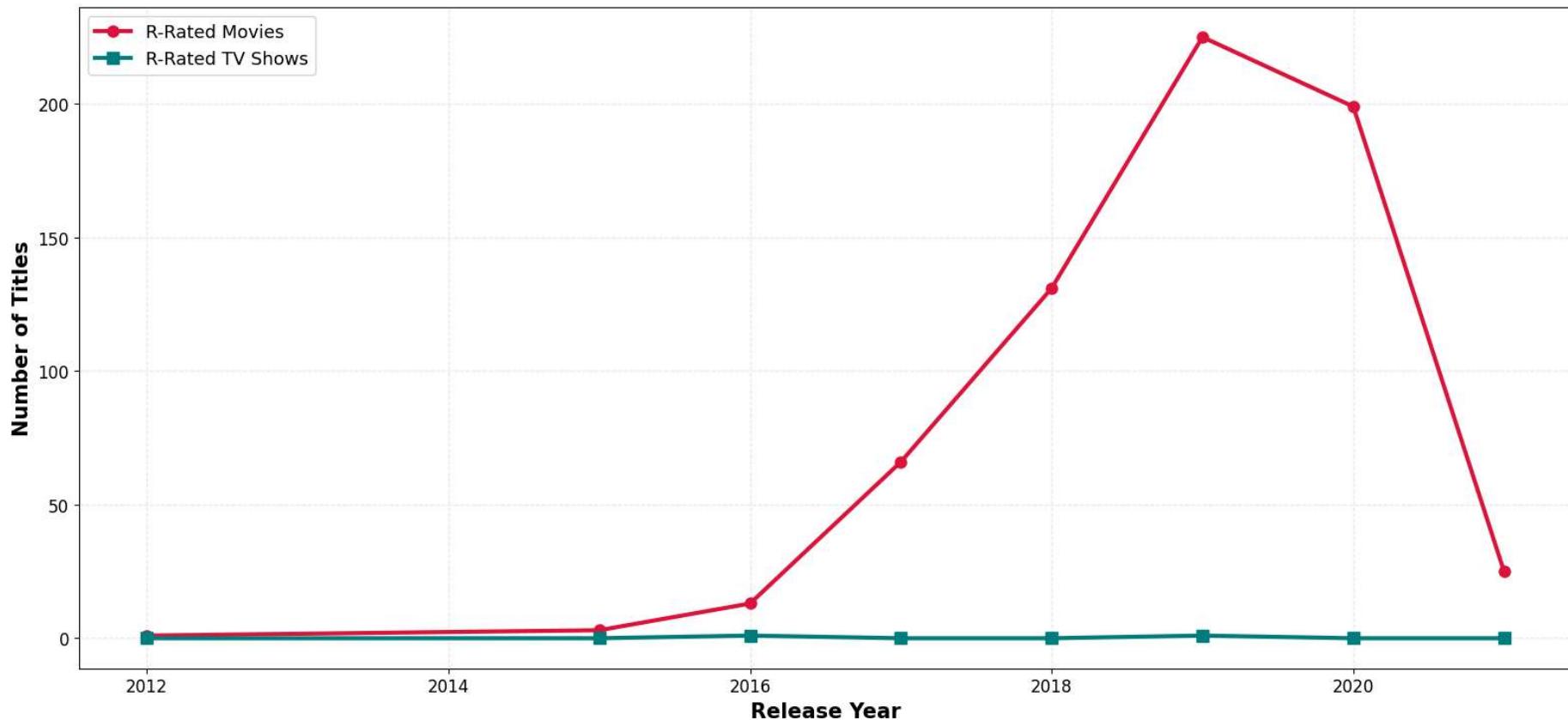
plt.title('R-RATED CONTENT ON NETFLIX OVER TIME', fontsize=22, fontweight='bold', pad=20)
plt.xlabel('Release Year', fontsize=15, fontweight='bold')
plt.ylabel('Number of Titles', fontsize=15, fontweight='bold')
plt.legend(fontsize=13, loc='upper left')
plt.grid(True, alpha=0.2, linestyle='--')
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)

plt.tight_layout()
plt.show()

# Quick stats
print(f"\nTotal R-Rated Movies: {int(r_rated[r_rated['Category']=='Movie'].shape[0])}")
```

```
print(f"Total R-Rated TV Shows: {int(r_rated[r_rated['Category']=='TV Show'].shape[0])}")
print(f"Peak year for R-Rated content: {int(r_by_year.sum(axis=1).idxmax())}")
```

## R-RATED CONTENT ON NETFLIX OVER TIME



Total R-Rated Movies: 663

Total R-Rated TV Shows: 2

Peak year for R-Rated content: 2019

In [135...]

```
# Duration of a Movie>Show on Netflix

# Split into Movies and TV Shows
movies = df[df['Category'] == 'Movie'].copy()
tv_shows = df[df['Category'] == 'TV Show'].copy()

# Extract numeric values
```

```
movies['Duration_min'] = movies['Duration'].str.replace(' min', '').astype(int)

# Handle both 'Season' and 'Seasons' for TV shows
tv_shows['Duration_seasons'] = tv_shows['Duration'].str.replace(' Seasons', '').str.replace(' Season', '').astype(int)

# Find maximums
max_movie = movies.loc[movies['Duration_min'].idxmax()]
max_tv = tv_shows.loc[tv_shows['Duration_seasons'].idxmax()]

print("=*60")
print("MAXIMUM DURATION ON NETFLIX")
print("=*60")

print("\nMOVIES:")
print(f"  Duration: {max_movie['Duration_min']} minutes")
print(f"  Title: {max_movie['Title']}")

print("\nTV SHOWS:")
print(f"  Duration: {max_tv['Duration_seasons']} seasons")
print(f"  Title: {max_tv['Title']}")

print("\n" + "=*60")

print("\nDURATION RANGES:")
print(f"  Movie duration range: {movies['Duration_min'].min()} - {movies['Duration_min'].max()} minutes")
print(f"  TV Shows range: {tv_shows['Duration_seasons'].min()} - {tv_shows['Duration_seasons'].max()} seasons")
```

```
=====
```

MAXIMUM DURATION ON NETFLIX

```
=====
```

MOVIES:

Duration: 312 minutes  
Title: Black Mirror: Bandersnatch

TV SHOWS:

Duration: 16 seasons  
Title: Grey's Anatomy

```
=====
```

DURATION RANGES:

Movie duration range: 3 - 312 minutes  
TV Shows range: 1 - 16 seasons

In [139...]

```
# Minimum Duration Movie
movies[movies['Duration_min'] == movies['Duration_min'].min()]
```

Out[139...]

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year	Di
5606	s5606	Movie	Silent	Limbert Fabian, Brandon Oldenburg	No Cast Information	United States	June 4, 2019	TV-Y	3 min	Children & Family Movies, Sci-Fi & Fantasy	"Silent" is an animated short film created by ...	2019	

In [140...]

```
# Maximum Duration Movie
movies[movies['Duration_min'] == movies['Duration_min'].max()]
```

Out[140...]

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_
957	s958	Movie	Black Mirror: Bandersnatch	Not Specified	Fionn Whitehead, Will Poulter, Craig Parkinson...	United States	December 28, 2018	TV-MA	312 min	Dramas, International Movies, Sci-Fi & Fantasy	In 1984, a young programmer begins to question...	

In [141...]

```
# Minimum Season TV Show
tv_shows[tv_shows['Duration_seasons'] == tv_shows['Duration_seasons'].min()]
```

Out[141...]

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year
5	s6	TV Show	46	Serdar Akar	Erdal Beşikçioğlu, Yasemin Allen, Melis Birkan...	Turkey	July 1, 2017	TV-MA	1 Season	International TV Shows, TV Dramas, TV Mysteries	A genetics professor experiments with a treatm...	20
11	s12	TV Show	1983	Not Specified	Robert Więckiewicz, Maciej Musiał, Michałina O...	Poland, United States	November 30, 2018	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Dramas	In this dark alt-history thriller, a naïve law...	20
12	s13	TV Show	1994	Diego Enrique Osorno	No Cast Information	Mexico	May 17, 2019	TV-MA	1 Season	Crime TV Shows, Docuseries, International TV S...	Archival video and new interviews examine Mexi...	20
16	s17	TV Show	Feb-09	Not Specified	Shahd El Yaseen, Shaila Sabt, Hala, Hanadi Al-...	Unknown Country	March 20, 2019	TV-14	1 Season	International TV Shows, TV Dramas	As a psychology professor faces Alzheimer's, h...	20
26	s27	TV Show	(Un)Well	Not Specified	No Cast Information	United States	August 12, 2020	TV-MA	1 Season	Reality TV	This docuseries takes a deep dive into the luc...	20
...	...	...	...	...	...	...	...	...	...	...	...	...
7768	s7767	TV Show	Zig & Sharko	Not Specified	No Cast Information	France	December 1, 2017	TV-Y7	1 Season	Kids' TV, TV Comedies	Zig, an island-bound hyena, will	20

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year
7769	s7768	TV Show	Zindagi Gulzar Hai	Not Specified	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	December 15, 2016	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...
7777	s7776	TV Show	Zoids Wild	Not Specified	Kensho Ono, Takahiro Sakurai, Mikako Komatsu, ...	Japan	August 14, 2020	TV-Y7	1 Season	Anime Series, Kids' TV	A quest for freedom and legendary treasure beg...
7781	s7780	TV Show	Zona Rosa	Not Specified	Manu NNa, Ana Julia Yeyé, Ray Contreras, Pablo...	Mexico	November 26, 2019	TV-MA	1 Season	International TV Shows, Spanish-Language TV Sh...	An assortment of talent takes the stage for a ...
7787	s7786	TV Show	Zumbo's Just Desserts	Not Specified	Adriano Zumbo, Rachel Khoo	Australia	October 31, 2020	TV-PG	1 Season	International TV Shows, Reality TV	Dessert wizard Adriano Zumbo looks for the nex...

1606 rows × 13 columns

In [142]:

```
# Maximum Season TV Show
tv_shows[tv_shows['Duration_seasons'] == tv_shows['Duration_seasons'].max()]
```

Out[142...]

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release_Year	D
2538	s2539	TV Show	Grey's Anatomy	Not Specified	Ellen Pompeo, Sandra Oh, Katherine Heigl, Just...	United States	May 9, 2020	TV-14	16 Seasons	Romantic TV Shows, TV Dramas	Intern (and eventual resident) Meredith Grey f...	2020	

In [154...]

```
# TOP COUNTRIES WITH MOST CONTENT ON NETFLIX

# Split multiple countries (e.g., "United States, India" becomes two entries)
countries_split = df['Country'].str.split(', ').explode()

# Remove 'Unknown Country' if you want
countries_split = countries_split[countries_split != 'Unknown Country']

# Get overall top countries
top_countries_all = countries_split.value_counts().head(15)

# Split by category
movies_countries = df[df['Category'] == 'Movie']['Country'].str.split(', ').explode()
tv_countries = df[df['Category'] == 'TV Show']['Country'].str.split(', ').explode()

top_movie_countries = movies_countries.value_counts().head(10)
top_tv_countries = tv_countries.value_counts().head(10)

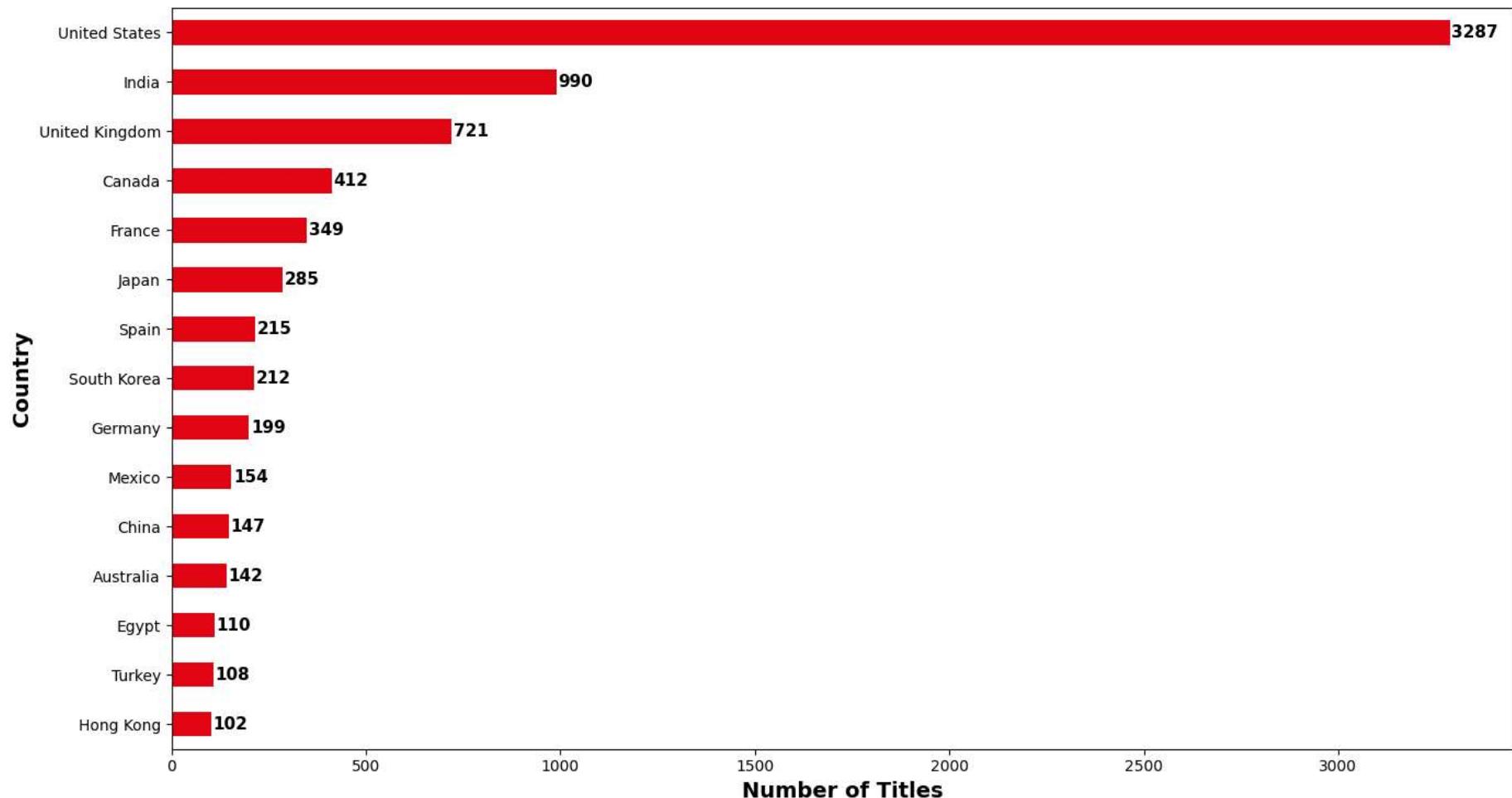
# TOP COUNTRIES OVERALL
plt.figure(figsize=(14, 8))
top_countries_all.plot(kind='barh', color="#E50914")
plt.title('TOP 15 COUNTRIES WITH MOST CONTENT ON NETFLIX', fontsize=18, fontweight='bold', pad=20)
plt.xlabel('Number of Titles', fontsize=14, fontweight='bold')
plt.ylabel('Country', fontsize=14, fontweight='bold')
plt.gca().invert_yaxis()

# Add value labels
for i, v in enumerate(top_countries_all.values):
```

```
plt.text(v + 5, i, str(v), va='center', fontsize=11, fontweight='bold')

plt.tight_layout()
plt.show()
```

## TOP 15 COUNTRIES WITH MOST CONTENT ON NETFLIX



In [ ]: