

STREAMINGDIENST VERGLEICH

von: Lilli Elsner, Julia Gebel, Alexandra
Geworsky, Mirja Jordan & Daniel Lock



GLIEDERUNG

Projektgruppe 4

Programmieren



01

**Content
Analyse**



03

**Rating
Analyse**



02

**Geographische
Analyse**



04

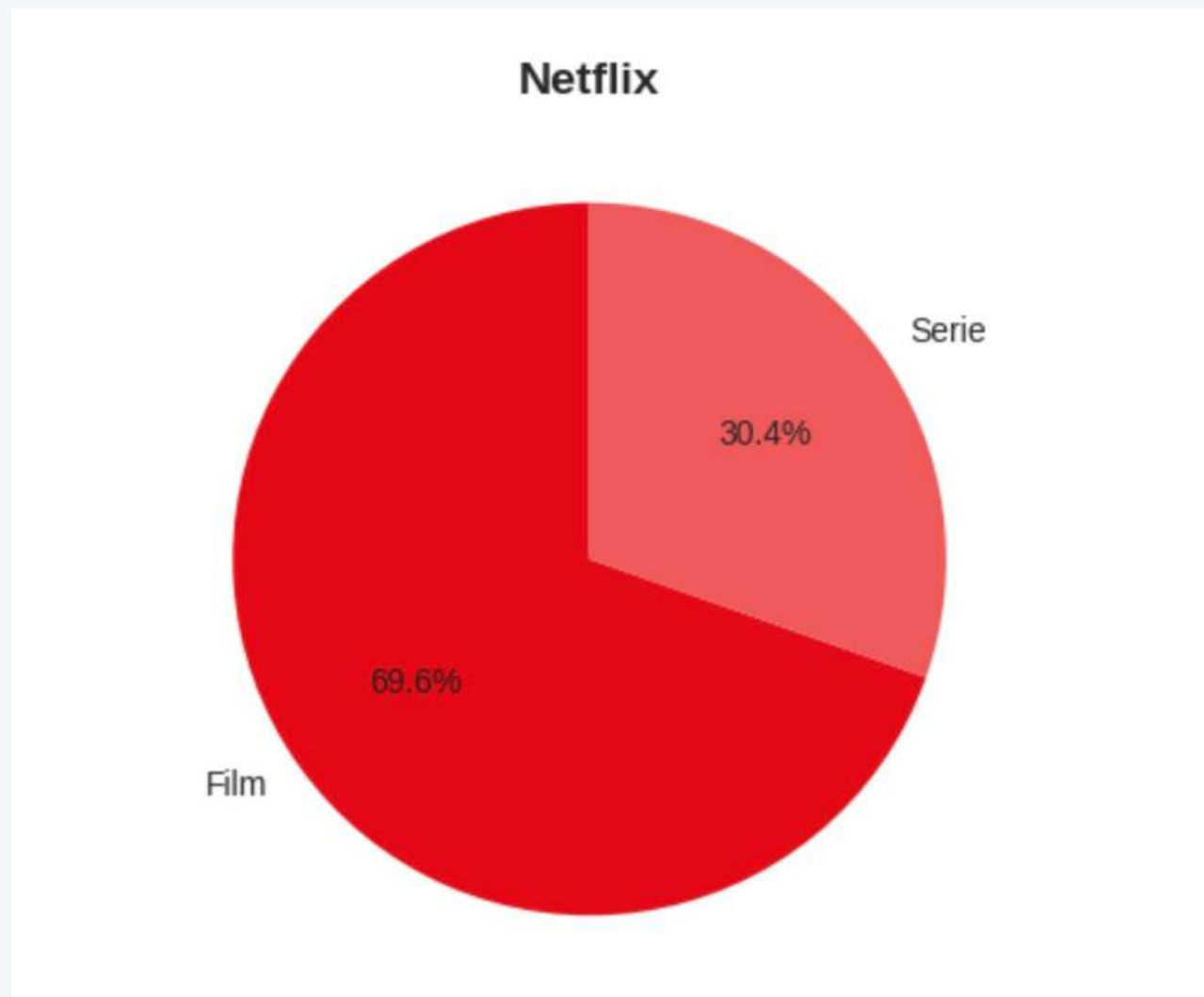
**Genre
Analyse**



NETFLIX



CONTENT



Inhalts Typ Analyse

```
fig, axes = plt.subplots(1, 3, figsize=(18, 5))  
fig.suptitle("Verteilung der Inhaltstypen auf den Plattformen",  
             fontsize=16, fontweight="bold")
```

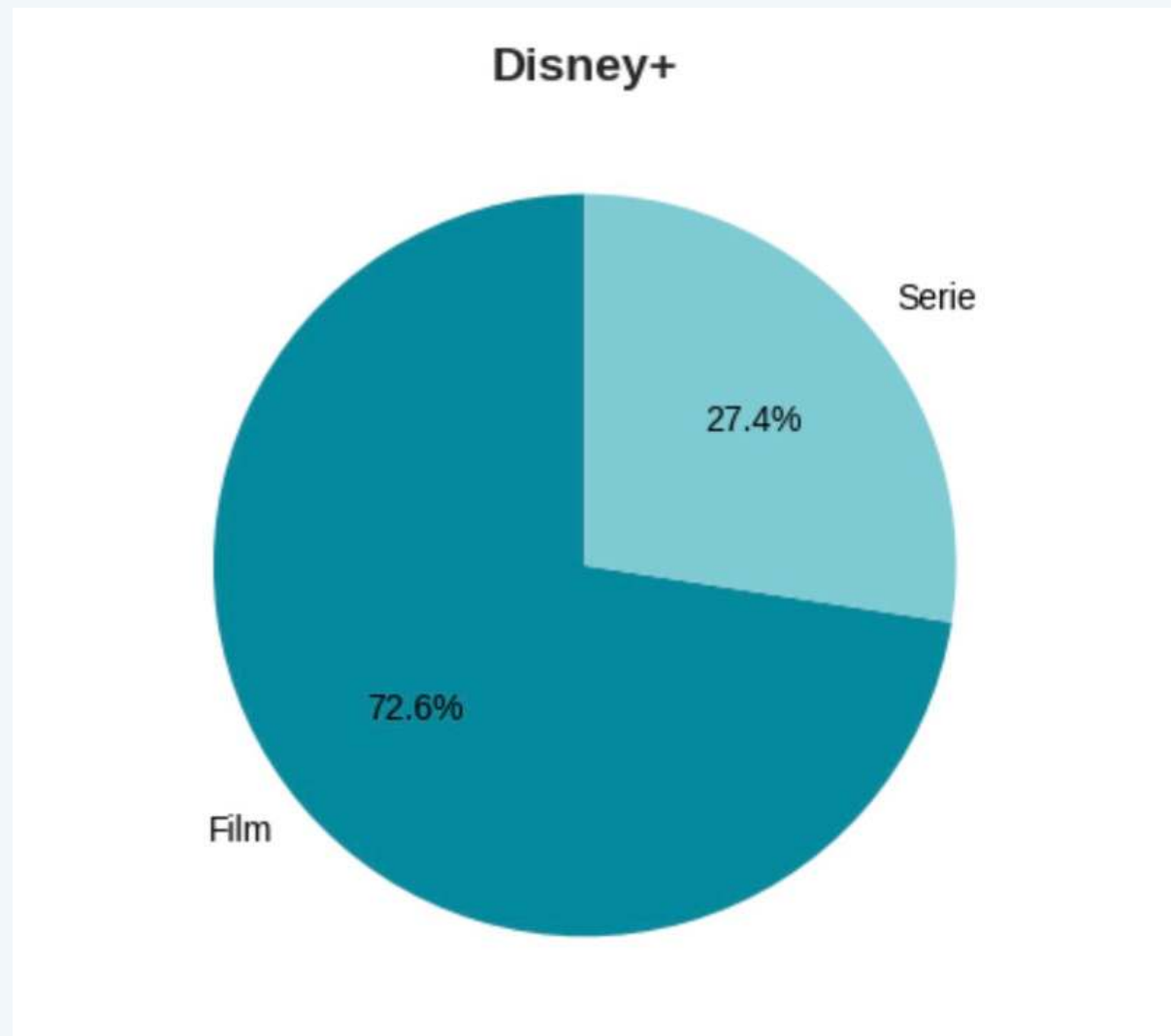
Netflix Inhaltstyp

```
netflix_type = netflix["type_de"].value_counts()  
axes[0].pie(  
    netflix_type.values,  
    labels=netflix_type.index,  
    autopct="%1.1f%%",  
    startangle=90,  
    colors=["#E50914", "#F05A5A"],  
)  
axes[0].set_title("Netflix", fontsize=14, fontweight="bold")
```

DISNEY+



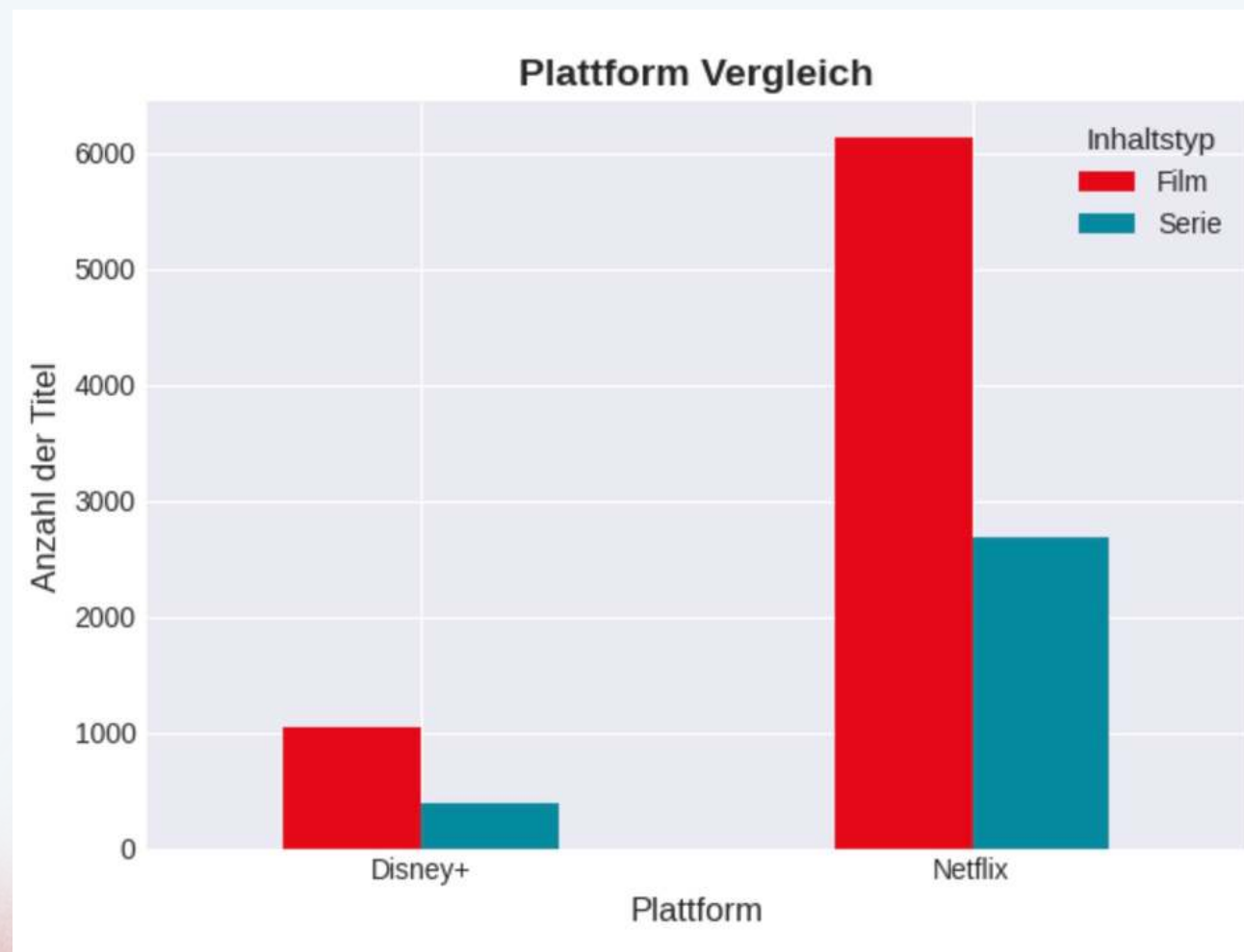
CONTENT



```
# Disney+ Inhaltstyp
disney_type = disney["type_de"].value_counts()
axes[1].pie(
    disney_type.values,
    labels=disney_type.index,
    autopct="%1.1f%%",
    startangle=90,
    colors=["#0a899c", "#7fcbd4"],
    textprops={"color": "black"},
)
axes[1].set_title("Disney+", fontsize=14, fontweight="bold")
```

DIE INHALTSTYPEN VON NETFLIX UND DISNEY+ IM DIREKTEN VERGLEICH

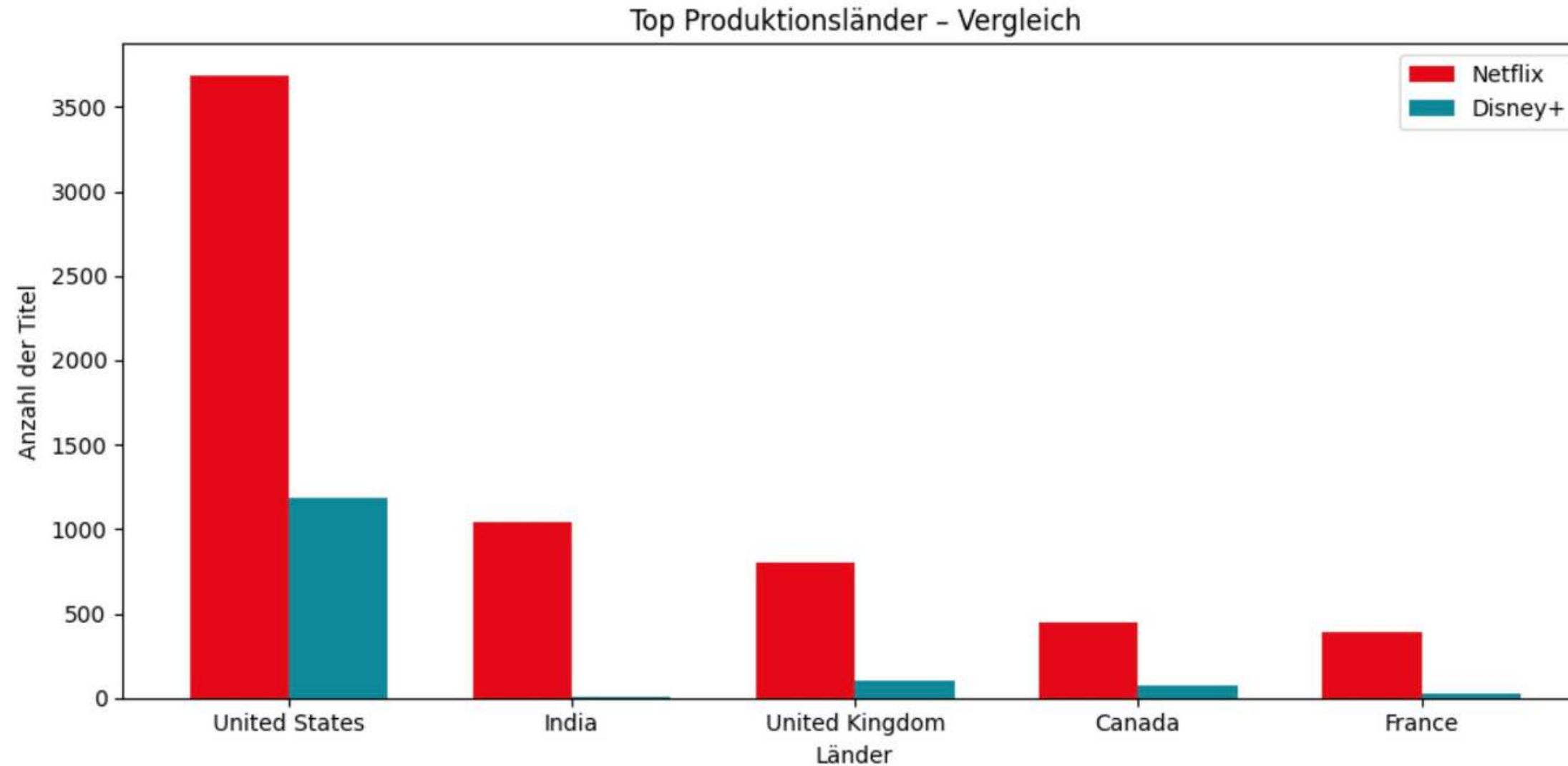
CONTENT



```
# Vergleich Balken Diagramm
combined = pd.concat([netflix, disney], ignore_index=True)
type_comparison = combined.groupby(["platform", "type_de"]).size().unstack(fill_value=0)
type_comparison.plot(kind="bar", ax=axes[2], color=["#E50914", "#0a899c"])
axes[2].set_title("Plattform Vergleich", fontsize=14, fontweight="bold")
axes[2].set_xlabel("Plattform", fontsize=12)
axes[2].set_ylabel("Anzahl der Titel", fontsize=12)
axes[2].legend(title="Inhaltstyp", title_fontsize=11)
axes[2].tick_params(axis="x", rotation=0)

plt.tight_layout()
plt.savefig("content_type_analysis.png", dpi=300, bbox_inches="tight")
plt.show()
```

DIE TOP PRODUKTIONSLÄNDER VON NETFLIX UND DISNEY+ IM VERGLEICH COUNTRIES



```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Daten laden
netflix = pd.read_csv("netflix_titles.csv")
disney = pd.read_csv("disney_plus_titles.csv")

# Farben
NETFLIX_RED = "#E50914"
DISNEY_BLUE = "#0a899c"

n_länder = netflix["country"].dropna().str.split(", ").explode()
d_länder = disney["country"].dropna().str.split(", ").explode()

n_counts = n_länder.value_counts()
d_counts = d_länder.value_counts()

top_länder = (n_counts + d_counts).sort_values(ascending=False).head(5).index

n_values = n_counts.reindex(top_länder, fill_value=0)
d_values = d_counts.reindex(top_länder, fill_value=0)

# Plot
x = np.arange(len(top_länder))
width = 0.35

plt.figure(figsize=(10,5))

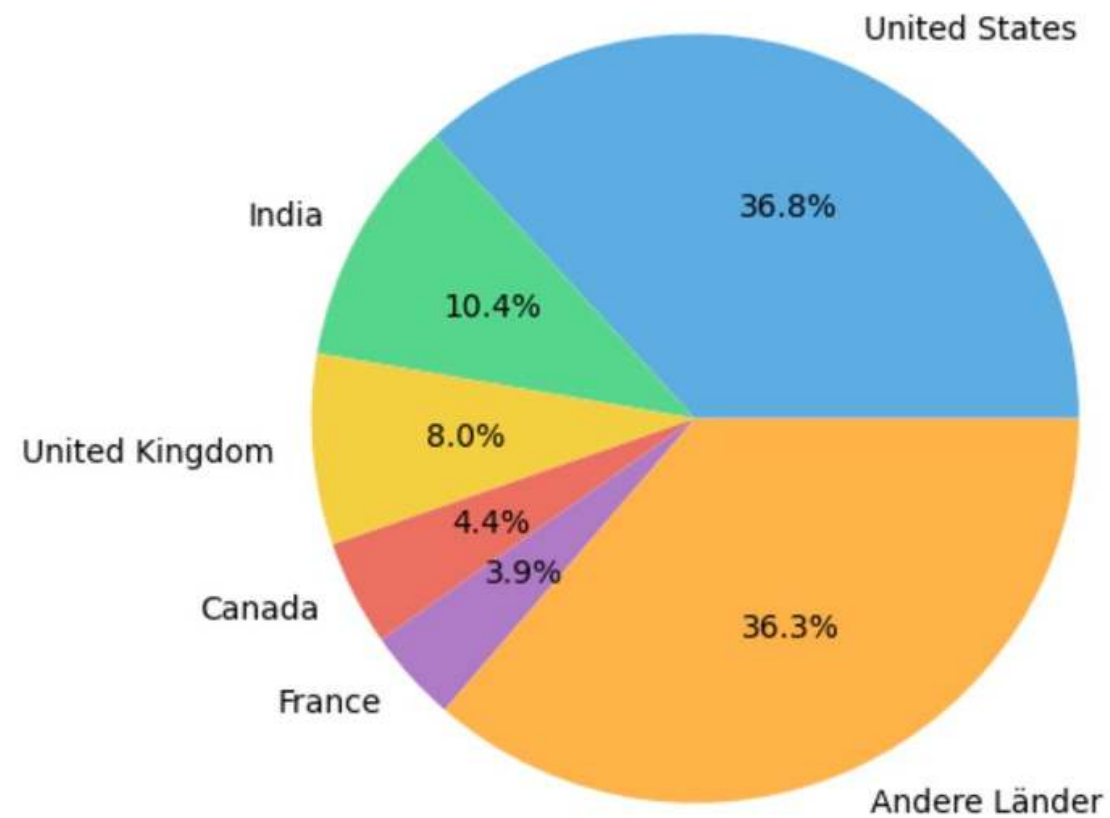
plt.bar(x - width/2, n_values, width, label="Netflix", color=NETFLIX_RED)
plt.bar(x + width/2, d_values, width, label="Disney+", color=DISNEY_BLUE)

plt.xticks(x, top_länder)
plt.title("Top Produktionsländer - Vergleich")
plt.xlabel("Länder")
plt.ylabel("Anzahl der Titel")
plt.legend()
plt.tight_layout()
plt.show()
```

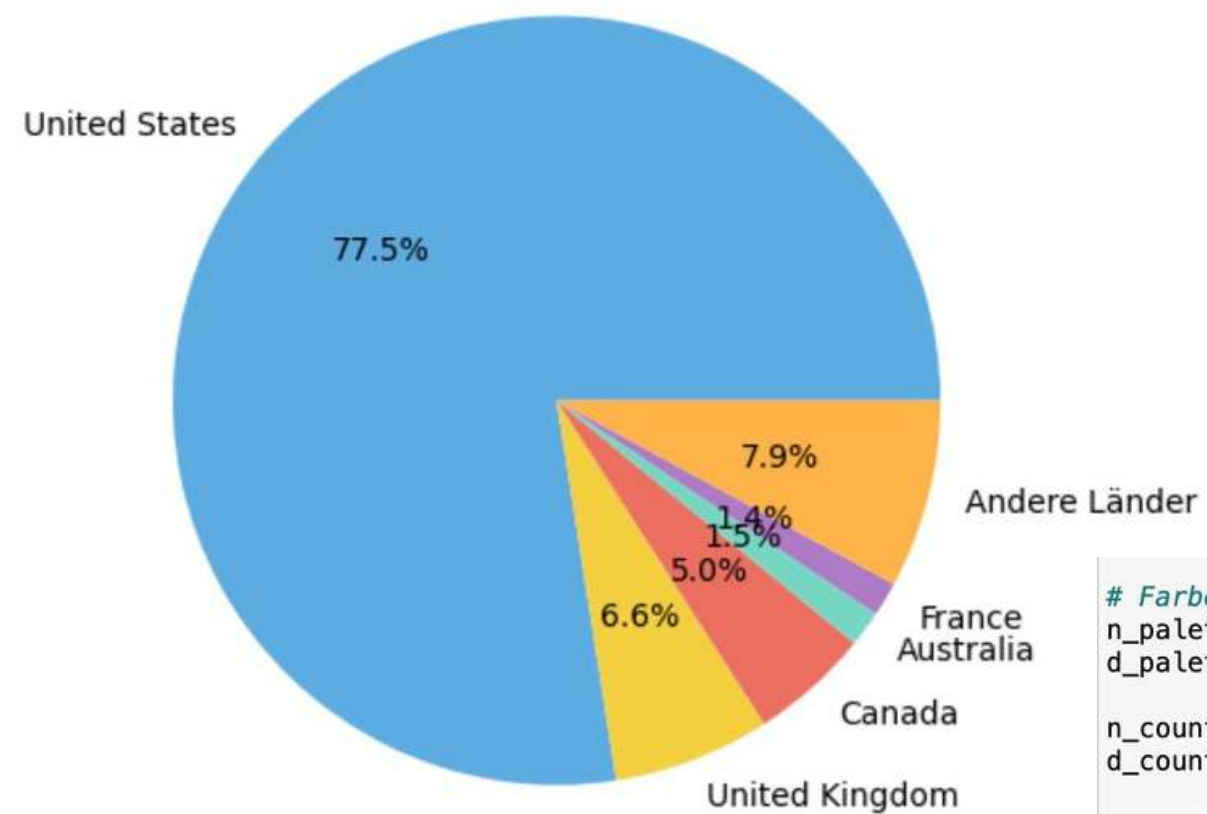
LÄNDERANTEILE VON NETFLIX UND DISNEY+ IM VERGLEICH

COUNTRIES

Netflix - Länderanteile



Disney+ - Länderanteile



```
# Farben
n_palette = ["#5DADE2", "#58D68D", "#F4D03F", "#EC7063", "#AF7AC5", "#FFB347"]
d_palette = ["#5DADE2", "#F4D03F", "#EC7063", "#76D7C4", "#AF7AC5", "#FFB347"]

n_counts = n_länder.value_counts()
d_counts = d_länder.value_counts()

n_top5 = n_counts.head(5)
n_rest = pd.Series({"Andere Länder": n_counts.iloc[5:].sum()})
n_pie = pd.concat([n_top5, n_rest])

d_top5 = d_counts.head(5)
d_rest = pd.Series({"Andere Länder": d_counts.iloc[5:].sum()})
d_pie = pd.concat([d_top5, d_rest])

# Plot
plt.figure(figsize=(12,6))

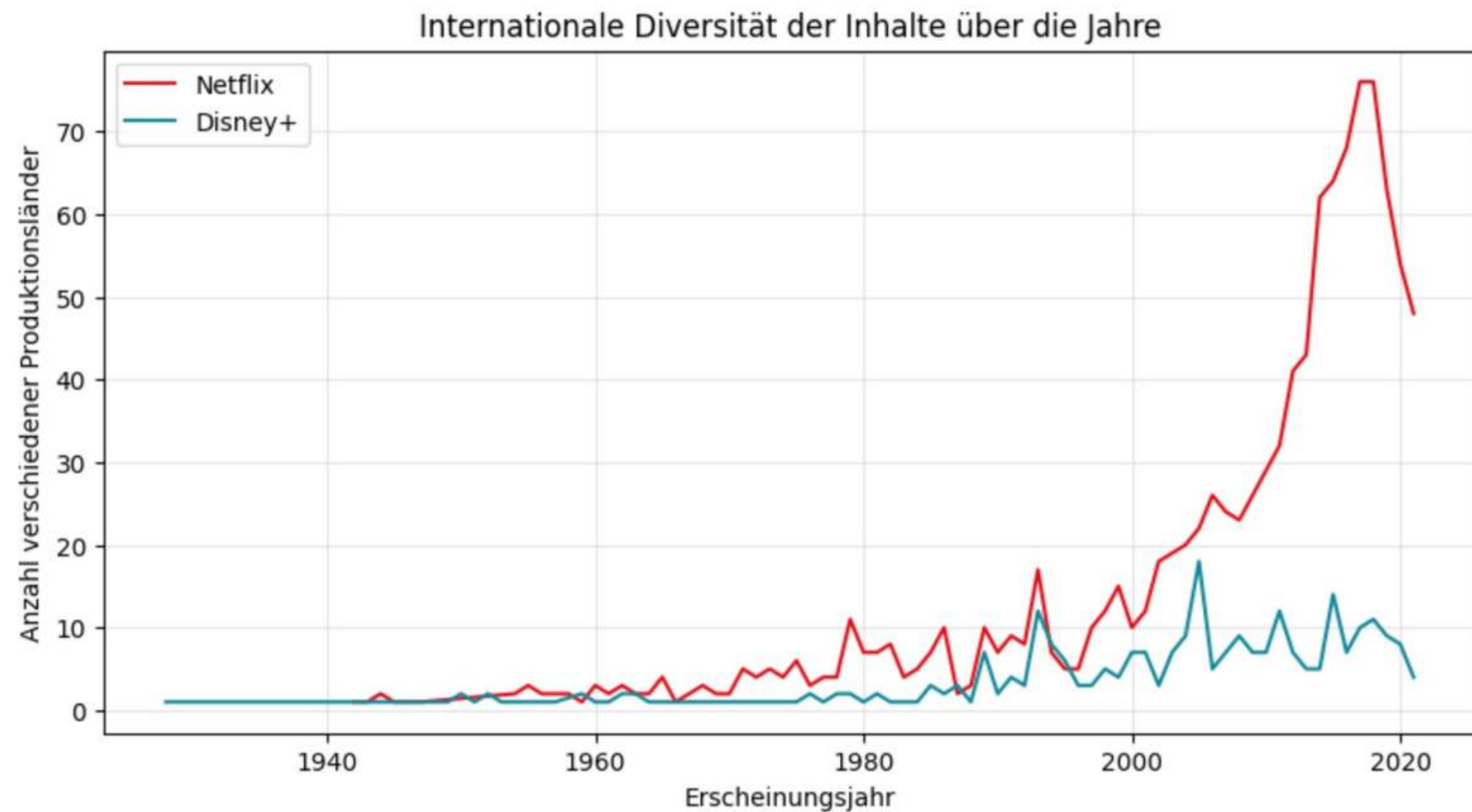
plt.subplot(1,2,1)
plt.pie(n_pie.values, labels=n_pie.index, autopct='%1.1f%%', colors=n_palette)
plt.title("Netflix - Länderanteile")

plt.subplot(1,2,2)
plt.pie(d_pie.values, labels=d_pie.index, autopct='%1.1f%%', colors=d_palette)
plt.title("Disney+ - Länderanteile")

plt.show()
```

ENTWICKLUNG DER INTERNATIONALEN DIVERSITÄT DER INHALTE VON NETFLIX UND DISNEY+

COUNTRIES



```
netflix_länder = netflix.dropna(subset=["country"]).copy()
disney_länder = disney.dropna(subset=["country"]).copy()

netflix_länder["country"] = netflix_länder["country"].str.split(", ")
disney_länder["country"] = disney_länder["country"].str.split(", ")

n_explode = netflix_länder.explode("country")
d_explode = disney_länder.explode("country")

n_länderdiversität = n_explode.groupby("release_year")["country"].nunique()
d_länderdiversität = d_explode.groupby("release_year")["country"].nunique()

# Plot
plt.figure(figsize=(10,5))

plt.plot(n_länderdiversität.index, n_länderdiversität.values, label="Netflix", color=NETFLIX_RED)
plt.plot(d_länderdiversität.index, d_länderdiversität.values, label="Disney+", color=DISNEY_BLUE)

plt.xlabel("Jahr")
plt.ylabel("Anzahl verschiedener Produktionsländer")
plt.title("Internationale Diversität der Inhalte über die Jahre")
plt.legend()
plt.grid(alpha=0.3)

plt.show()
```

RATINGS DISNEY+



```
#Disney+
disney_bewertungen = disney['rating'].value_counts().head(10)

plt.figure(figsize=(8,5))
disney_bewertungen.plot(kind='bar', color='#0a899c')

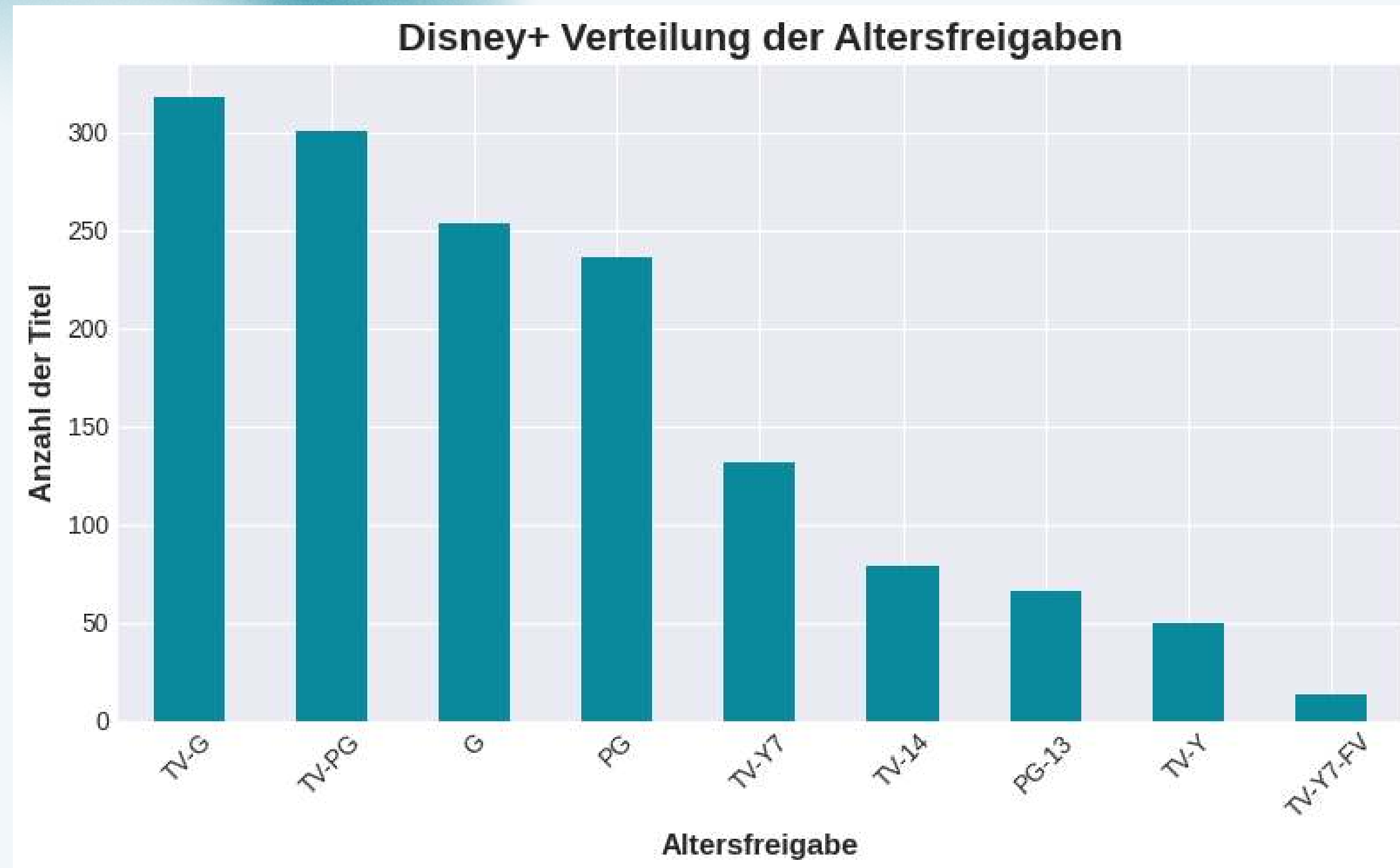
plt.title('Disney+ Verteilung der Altersfreigaben', fontsize=16, fontweight='bold')
plt.xlabel('Altersfreigabe', fontsize=12, fontweight='bold')
plt.ylabel('Anzahl der Titel', fontsize=12, fontweight='bold')

plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

„RATING“ AUS
DATASET
SELEKTIEREN

PLOT
ERSTELLEN

RATINGS DISNEY+



ZIEL :
**LEGENDE
EINFÜGEN**
zur Erklärung der
Begriffe

RATINGS DISNEY+



```
rating_erklärung = {  
    "TV-14": "Ab 14 Jahren",  
    "G": "für alle geeignet",  
    "TV-PG": "Empfehlung elterlicher Aufsicht",  
    "PG-13": "Eltern empfohlen ab 13",  
    "TV-Y7": "ab 7 Jahren",  
    "PG": "Eltern empfohlen",  
    "TV-G": "Für alle geeignet",  
    "TV-Y": "Für Kinder geeignet",  
    "TV-Y7-FV": "harmlose Fantasy Gewalt \n(ab 7 Jahren)"  
}
```

```
balken=plt.bar(disney_bewertungen.index, disney_bewertungen.values, color='#0a899c')  
  
# Jedem Balken ein Label geben  
for bar, rating in zip(balken, disney_bewertungen.index):  
    bar.set_label(f"{rating}: {rating_erklärung.get(rating, 'Keine Info')}")  
  
plt.legend(title="Rating Erklärung", bbox_to_anchor=(1.05, 1), loc='upper left')
```

LEGENDE
EINFÜGEN
zur Erklärung der
Begriffe

Dictionary mit
Erklärungen

jeden Balken
als Objekt speichern

jedem Balken ein
Label zuweisen

RATINGS DISNEY+



DIAGRAMM
MIT LEGENDE

ERGEBNISSE :

größte
Kategorien
→ für alle
geeignete Inhalte
→ mit elterlicher
Aufsicht

kleinste
Kategorien
→ für Kinder
geeignet

RATINGS NETFLIX



```
#Netflix
netflix_bewertungen = netflix['rating'].value_counts().head(10)

plt.figure(figsize=(8, 5))
netflix_bewertungen.plot(kind='bar', color='#E50914')

plt.title('Netflix - Verteilung der Altersfreigaben', fontsize=16, fontweight='bold')
plt.xlabel('Altersfreigaben', fontsize=12, fontweight='bold')
plt.ylabel('Anzahl der Titel', fontsize=12, fontweight='bold')
plt.xticks(rotation=45)
plt.tight_layout()

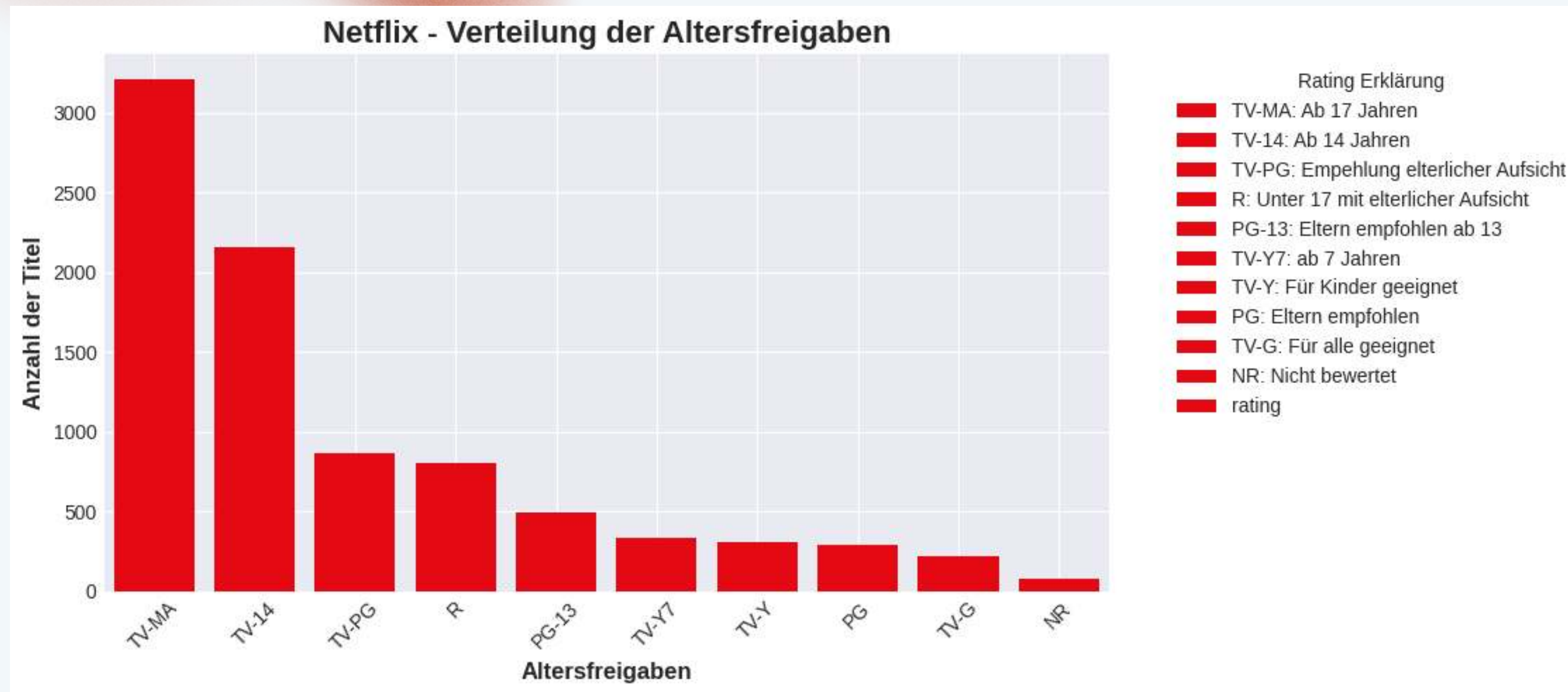
rating_erklärung = {
    "TV-MA": "Ab 17 Jahren",
    "TV-14": "Ab 14 Jahren",
    "TV-PG": "Empfehlung elterlicher Aufsicht",
    "R": "Unter 17 mit elterlicher Aufsicht",
    "PG-13": "Eltern empfohlen ab 13",
    "TV-Y7": "ab 7 Jahren",
    "TV-Y": "Für Kinder geeignet",
    "PG": "Eltern empfohlen",
    "TV-G": "Für alle geeignet",
    "NR": "Nicht bewertet"
}

balken=plt.bar(netflix_bewertungen.index, netflix_bewertungen.values, color='#E50914')

# Jedem Balken ein Label geben
for bar, rating in zip(balken, netflix_bewertungen.index):
    bar.set_label(f"{rating}: {rating_erklärung.get(rating, 'Keine Info')}")
plt.legend(title="Rating Erklärung", bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```

CODE MIT
LEGENDE

RATINGS NETFLIX



ERGEBNISSE

größte
Kategorien
→ Erwachsene
& ab 14 Jahren

kleinste
Kategorien
→ nicht bewertet
→ unter
elterlicher
Aufsicht
→ für alle
geeignet
→ für Kinder
geeignet

DISNEY+

Disney+ Verteilung der Altersfreigaben

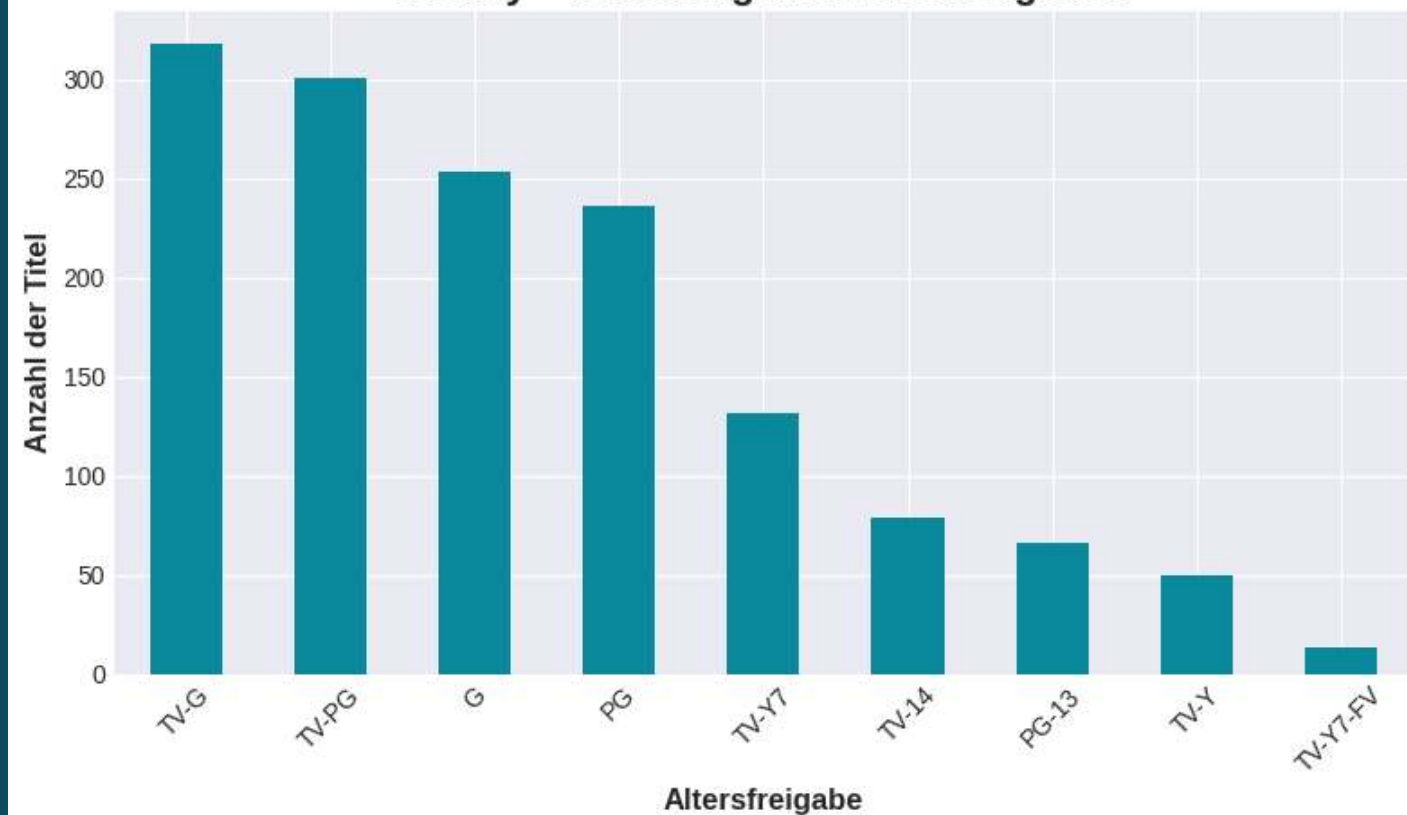
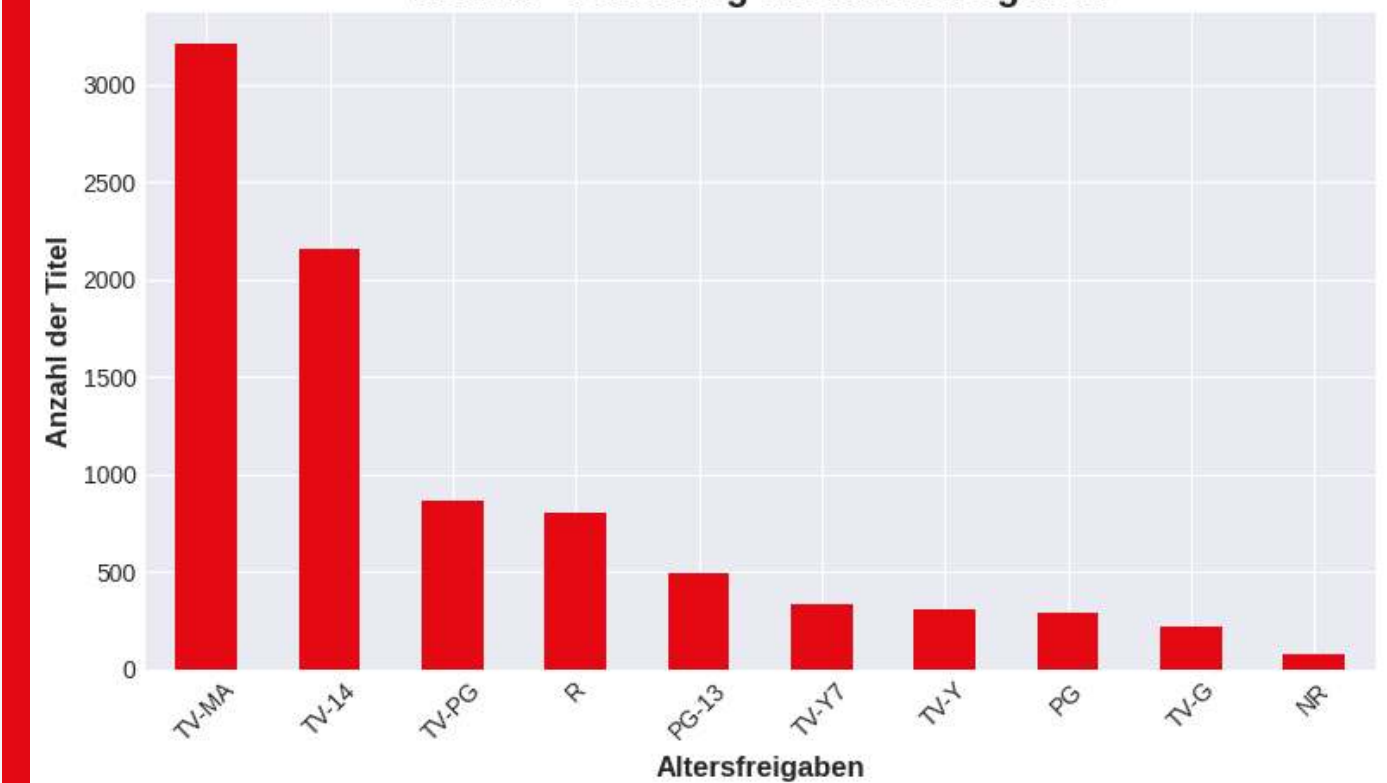


DIAGRAMME VERGLEICHEN

⚡ KEINE DARSTELLUNG IN
RELATION

NETFLIX

Netflix - Verteilung der Altersfreigaben



RATINGS VERGLEICH



```
# Werte mit minuten rausfiltern
netflix = netflix[~netflix['rating'].str.contains('min', na=False)]
disney = disney[~disney['rating'].str.contains('min', na=False)]

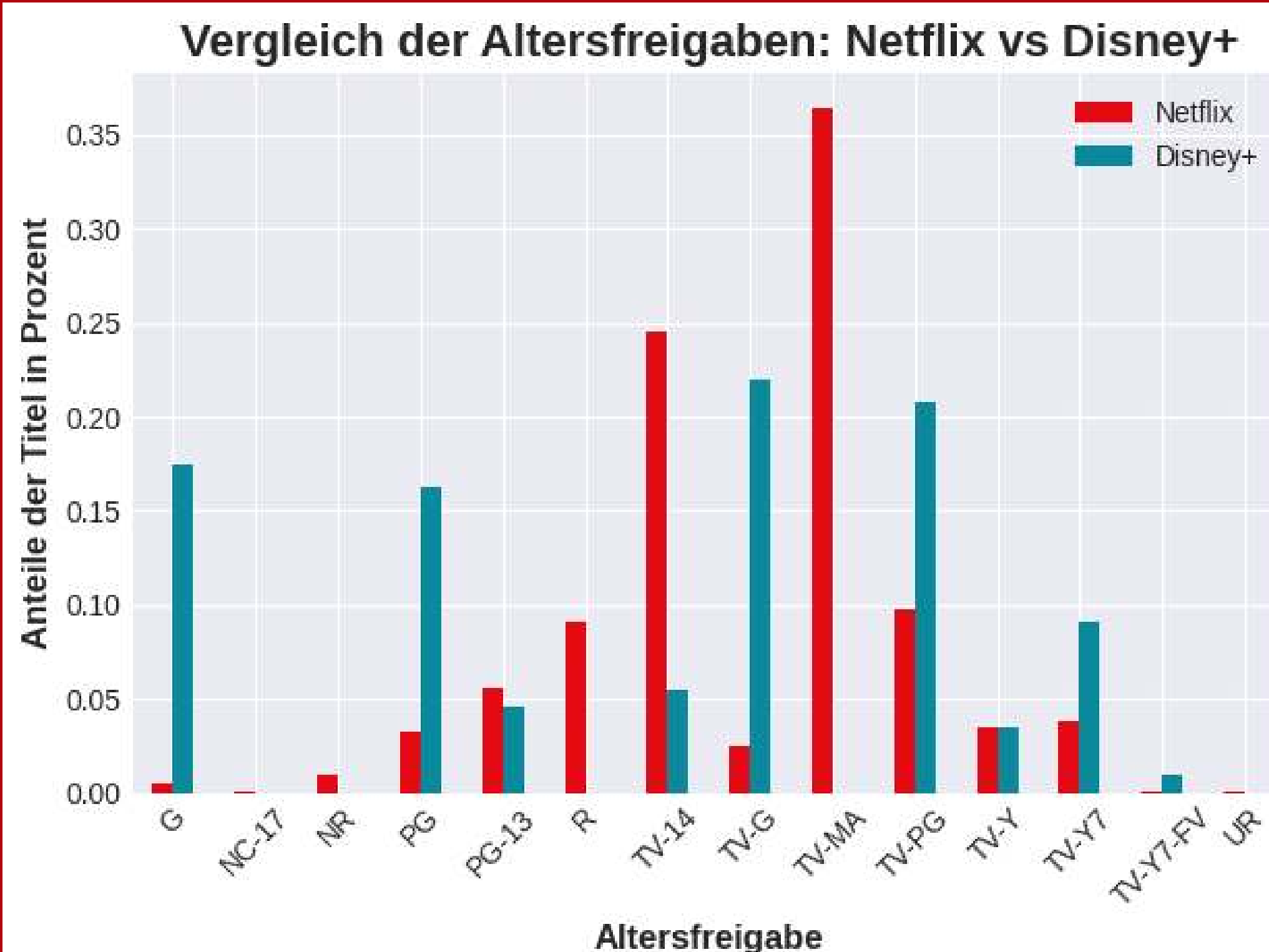
#Vergleich der Ratings
vergleich_bewertung = pd.DataFrame({
    'Netflix': netflix['rating'].value_counts(normalize=True),
    'Disney+': disney['rating'].value_counts(normalize=True)
}).fillna(0)

plt.figure(figsize=(10,6))
vergleich_bewertung.plot(kind='bar', color=['#E50914', '#0a899c'])

plt.title('Vergleich der Altersfreigaben: Netflix vs Disney+', fontsize=16, fontweight='bold')
plt.xlabel('Altersfreigabe', fontsize=12, fontweight='bold')
plt.ylabel('Anteile der Titel in Prozent', fontsize=12, fontweight='bold')

plt.xticks(rotation=45)
plt.tight_layout()
plt.legend()
plt.show()
```

RATINGS VERGLEICH



X-ACHSE

Kategorien der Altersfreigabe

Y-ACHSE

Anteile der Titel in Prozent

VERGLEICH

- Unterschiedliche Bewertungen:
 - Netflix größte Kategorien
= für Erwachsene & ab 14 Jahren
 - Inhalte nur für Erwachsene
gibt es nicht auf Disney+
 - Disney+ Inhalte für alle
geeignet & unter elterlicher Aufsicht
 - größere Kategorien-Diskrepanz
bei Netflix Inhalten



→ **familienfreundlicheres Angebot**



→ bietet überwiegend
Content für erwachsenere Zuschauer an

DIE TOP GENRES BEI NETFLIX UND DISNEY+ IM VERGLEICH

GENRES

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

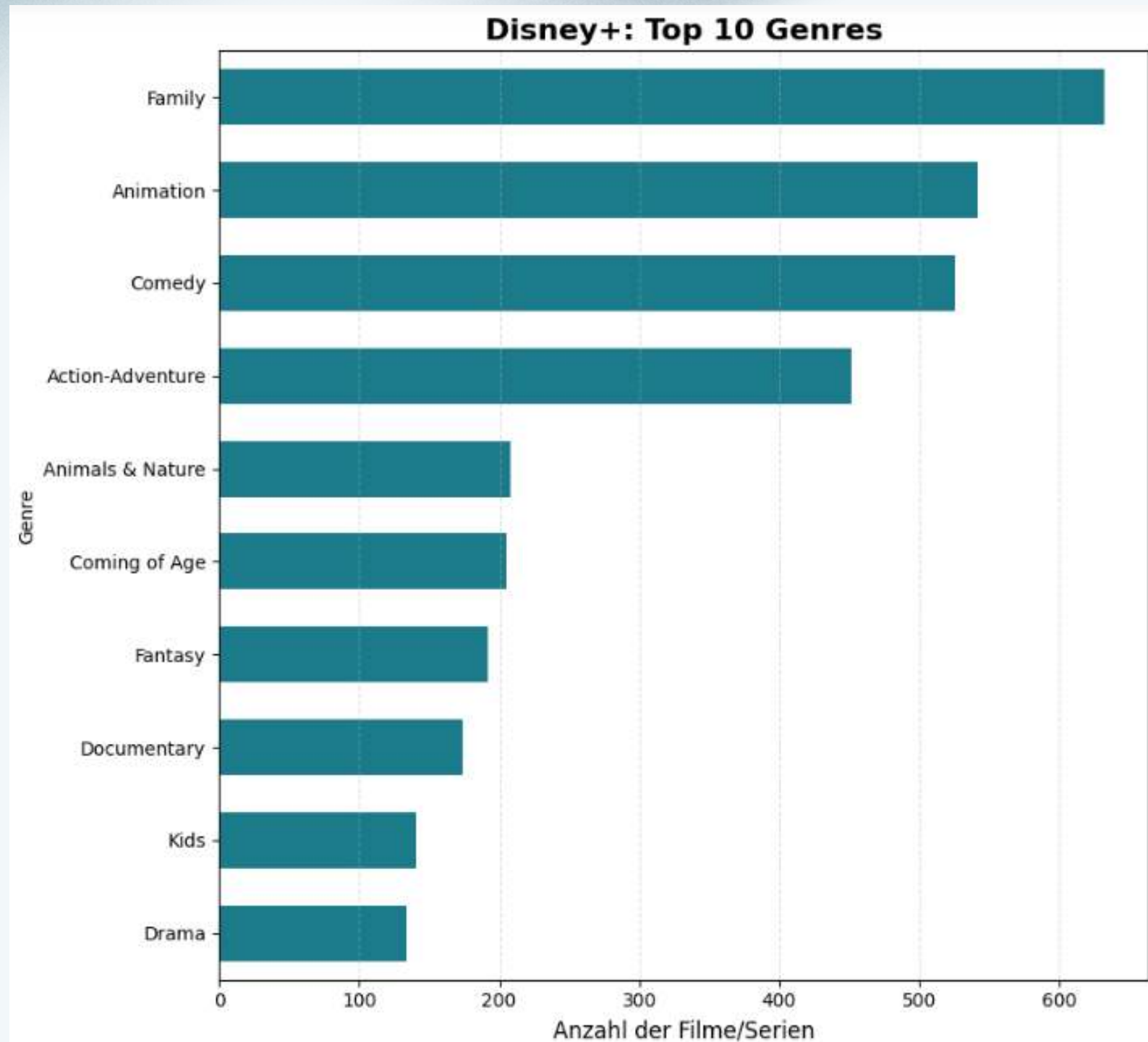
# Daten Laden
netflix_df = pd.read_csv('netflix_titles.csv')
disney_df = pd.read_csv('disney_plus_titles.csv')

def get_top_genres(df, top_n = 10):
    all_genres = df['listed_in'].str.split(',').explode()
    genre_counts = all_genres.value_counts().head(top_n).reset_index()
    genre_counts.columns = ['Genre', 'Anzahl']
    return genre_counts

# Top 10 Genres werden jeweils extrahiert:
top_10_netflix = get_top_genres(netflix_df, 10)
top_10_disney = get_top_genres(disney_df, 10)
```

DIE TOP GENRES BEI NETFLIX UND DISNEY+ IM VERGLEICH

GENRES



```
# Visualisierung
```

```
fig, axes = plt.subplots(1, 2, figsize=(18, 8))
```

```
# Gitterlinien für bessere Lesbarkeit
```

```
axes[0].grid(axis = 'x', linestyle = '--', alpha = 0.4)
```

```
axes[1].grid(axis = 'x', linestyle = '--', alpha = 0.4)
```

```
# Disney+ Plot:
```

```
sns.barplot(
```

```
    data = top_10_disney,
```

```
    x = 'Anzahl',
```

```
    y = 'Genre',
```

```
    ax = axes[1],
```

```
    color = '#0a899c',
```

```
    width = 0.6
```

```
)
```

```
axes[1].set_title('Disney+: Top 10 Genres', fontsize = 16, fontweight = 'bold')
```

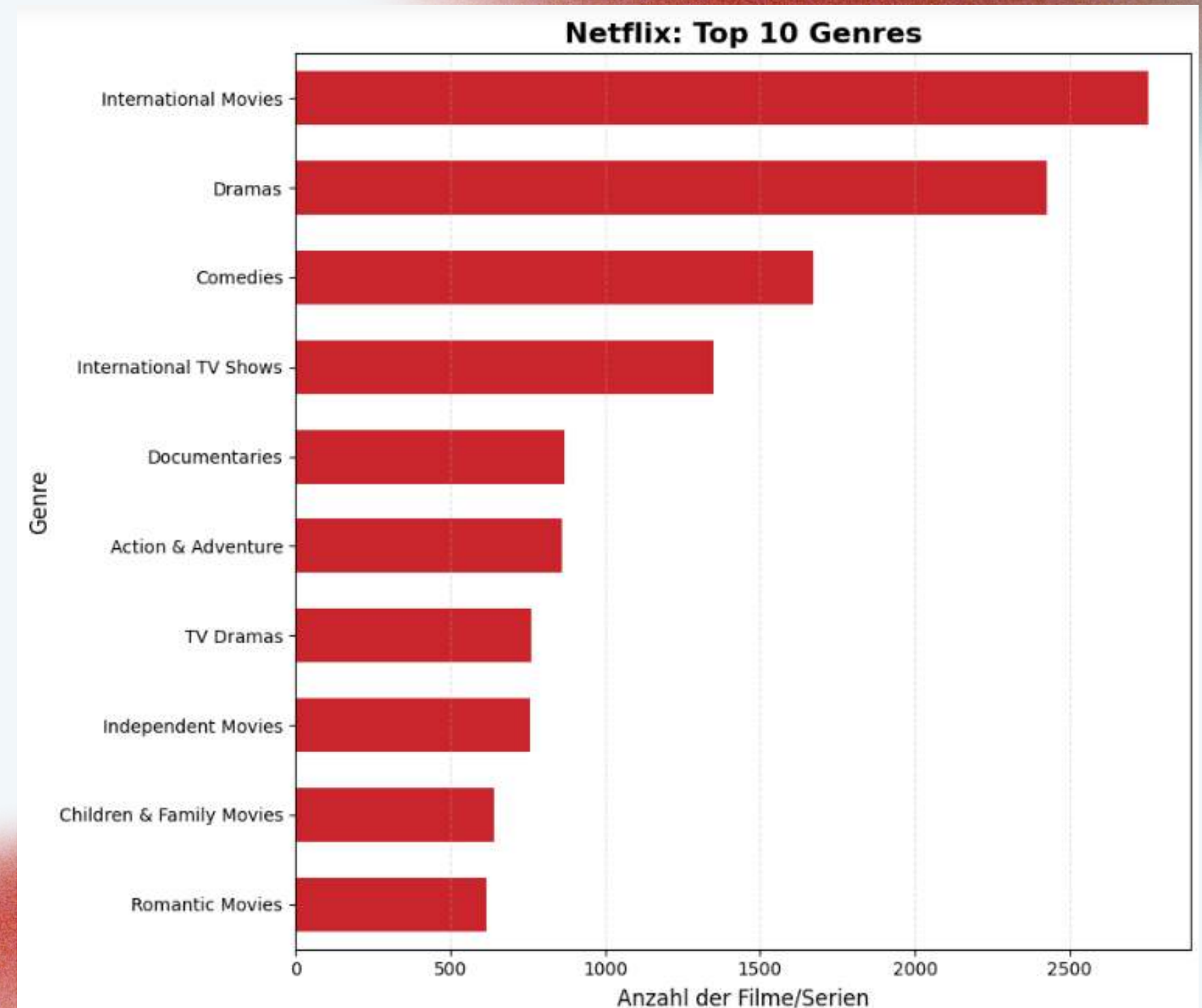
```
axes[1].set_xlabel('Anzahl der Filme/Serien', fontsize = 12)
```

DIE TOP GENRES BEI NETFLIX UND DISNEY+ IM VERGLEICH

GENRES

```
# Netflix Plot:
sns.barplot(
    data = top_10_netflix,
    x = 'Anzahl',
    y = 'Genre',
    ax = axes[0],
    color = '#E50914',
    width = 0.6
)
axes[0].set_title('Netflix: Top 10 Genres', fontsize = 16, fontweight = 'bold')
axes[0].set_xlabel('Anzahl der Filme/Serien', fontsize = 12)
axes[0].set_ylabel('Genre', fontsize = 12)

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



GENRES

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Daten Laden
netflix_df = pd.read_csv('netflix_titles.csv')
disney_df = pd.read_csv('disney_plus_titles.csv')

def get_top_5_data(df, name):
    genres = df['listed_in'].str.split(',').explode()
    counts = genres.value_counts().head(5).reset_index()
    counts.columns = ['Genre', 'Anzahl']
    counts['Plattform'] = name
    return counts
```

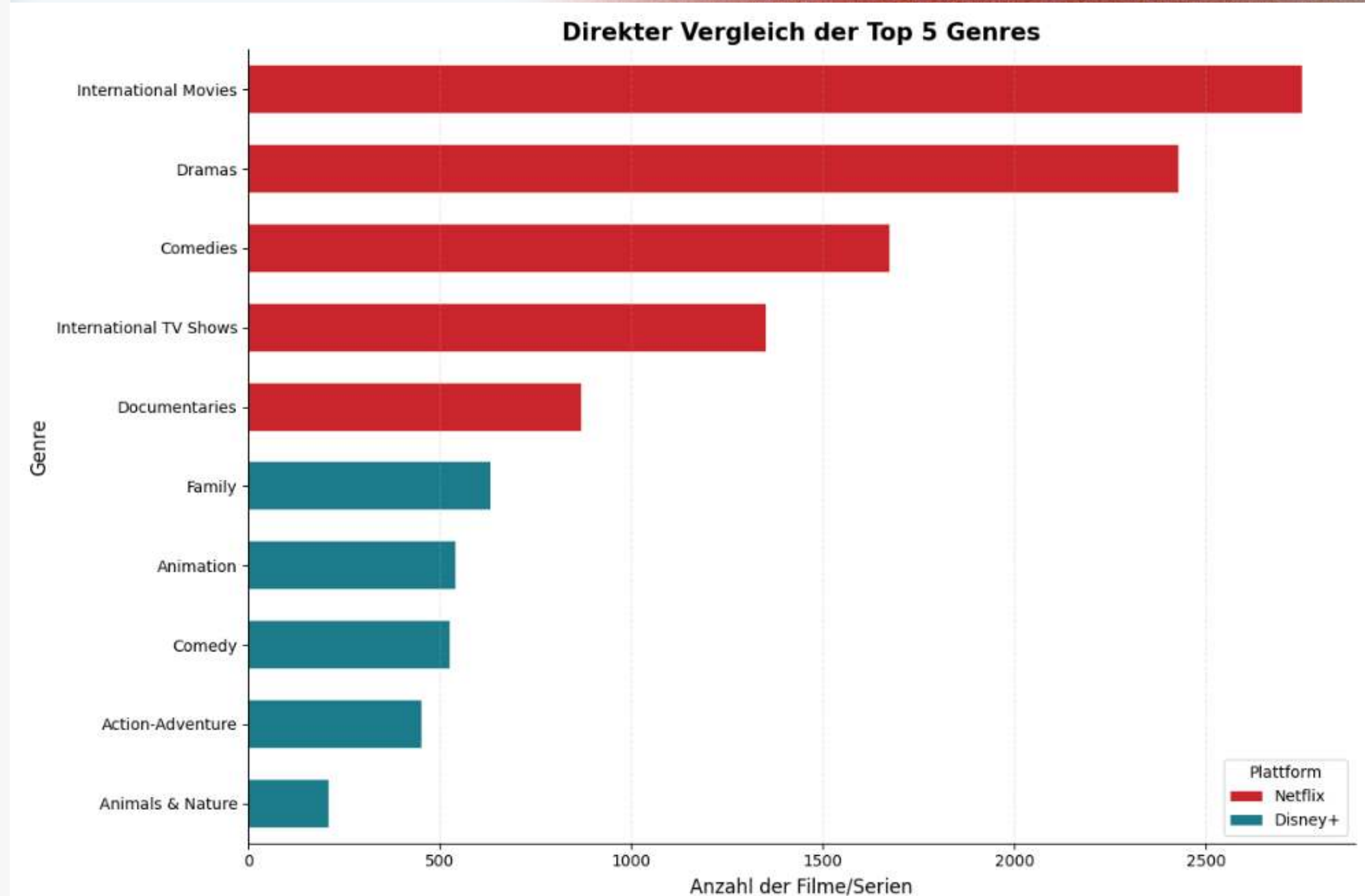
```
# Jeweilige Top 5 extrahieren und kombinieren
top_5_n = get_top_5_data(netflix_df, 'Netflix')
top_5_d = get_top_5_data(disney_df, 'Disney+')
combined_top_5 = pd.concat([top_5_n, top_5_d])
```

```
# Visualisierung
plt.figure(figsize = (12, 8))

sns.barplot(
    data=combined_top_5,
    y = 'Genre',
    x = 'Anzahl',
    hue = 'Plattform',
    palette = {'Netflix': '#E50914', 'Disney+': '#0a899c'},
    width = 0.6
)
```

```
plt.title('Direkter Vergleich der Top 5 Genres', fontsize = 15, fontweight = 'bold')
plt.xlabel('Anzahl der Filme/Serien', fontsize = 12)
plt.ylabel('Genre', fontsize = 12)
plt.grid(axis = 'x', linestyle = '--', alpha = 0.3)
```

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



THANK
YOU

