

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud
import datetime as dt
```

```
# Load Dataset
df = pd.read_csv("netflix_titles.csv")
df.head()
```

```
# Basic Info
df.info()
```

```
# Check for nulls
df.isnull().sum()
```

```
# Fill or Drop nulls
df['director'] = df['director'].fillna('Not Available')
df['cast'] = df['cast'].fillna('Not Available')
df.dropna(subset=['country', 'date_added'], inplace=True)
```

```
↩ <class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null  object
1   type            8807 non-null  object
2   title           8807 non-null  object
3   director        6173 non-null  object
4   cast            7982 non-null  object
5   country         7976 non-null  object
6   date_added      8797 non-null  object
```

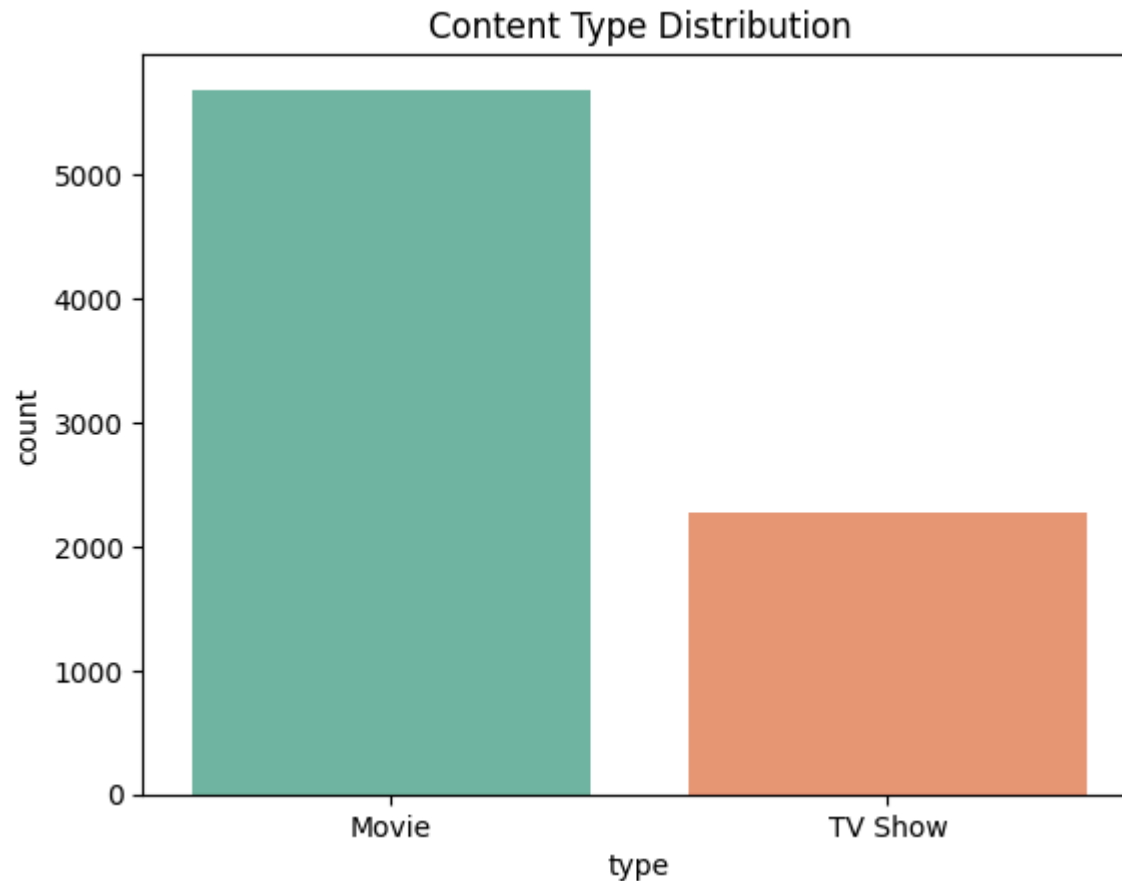
```
7  release_year  8807 non-null  int64
8  rating        8803 non-null  object
9  duration      8804 non-null  object
10 listed_in     8807 non-null  object
11 description   8807 non-null  object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

```
# Movie vs TV Show count
sns.countplot(data=df, x='type', palette='Set2')
plt.title("Content Type Distribution")
plt.show()
```

```
↗ /tmp/ipython-input-3-2424322931.py:2: FutureWarning:
```

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set

```
sns.countplot(data=df, x='type', palette='Set2')
```

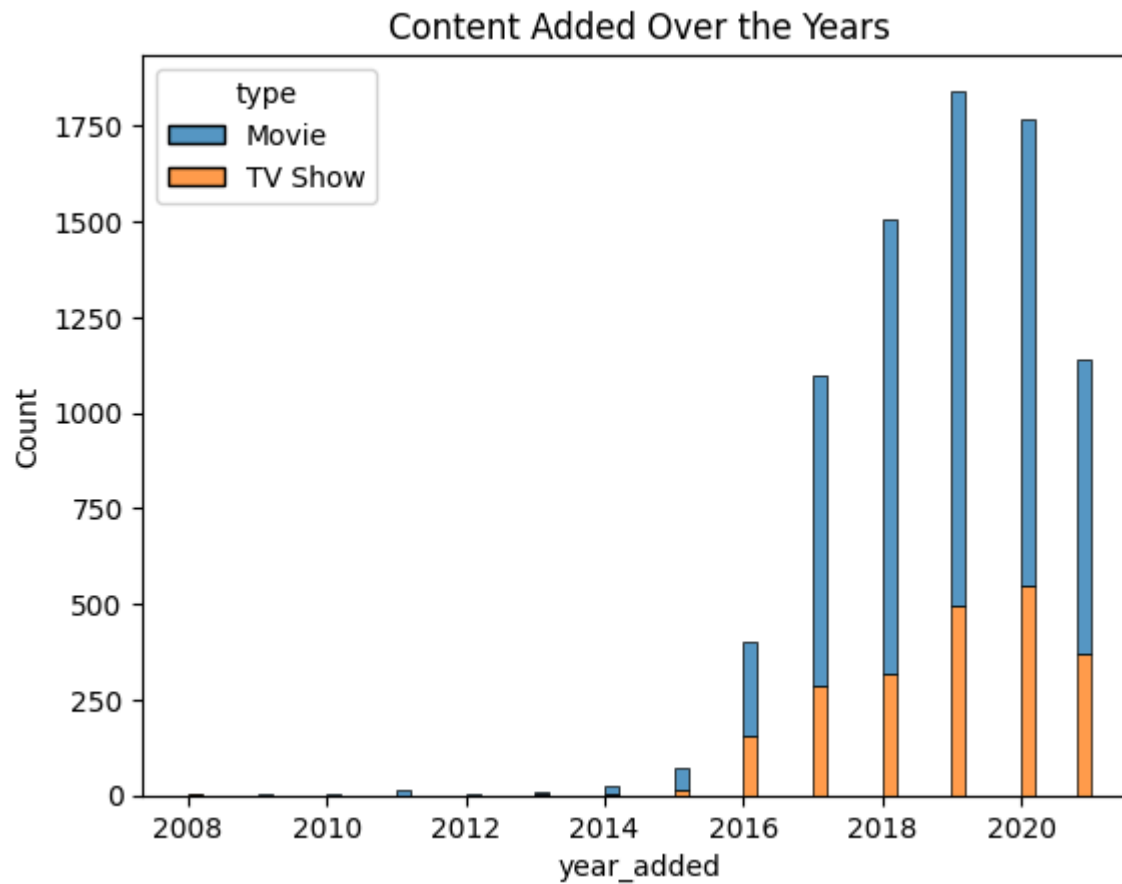


```
df['date_added'] = pd.to_datetime(df['date_added'], errors='coerce')
df['year_added'] = df['date_added'].dt.year
df['month_added'] = df['date_added'].dt.month
```

```
# Content added by year
```

```
sns.histplot(data=df, x='year_added', hue='type', multiple='stack')
```

```
sns.relplot(data=df, x= year_added , hue= type , multiple= stack )
plt.title("Content Added Over the Years")
plt.show()
```



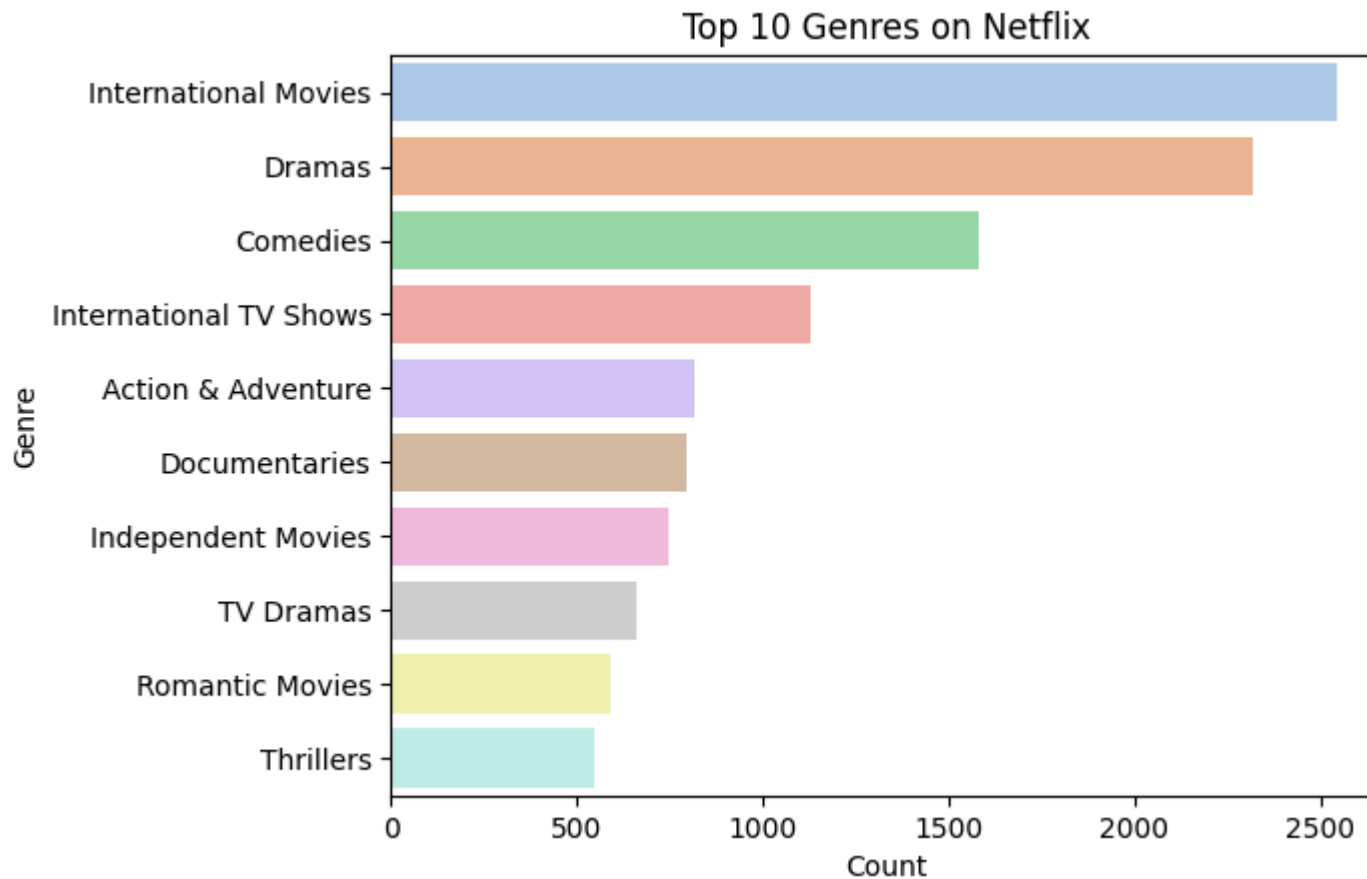
```
# Top 10 genres
from collections import Counter
genres = df['listed_in'].str.split(', ')
genre_counts = Counter([genre for sublist in genres for genre in sublist])
top_genres = pd.DataFrame(genre_counts.most_common(10), columns=['Genre', 'Count'])

sns.barplot(data=top_genres, y='Genre', x='Count', palette='pastel')
```

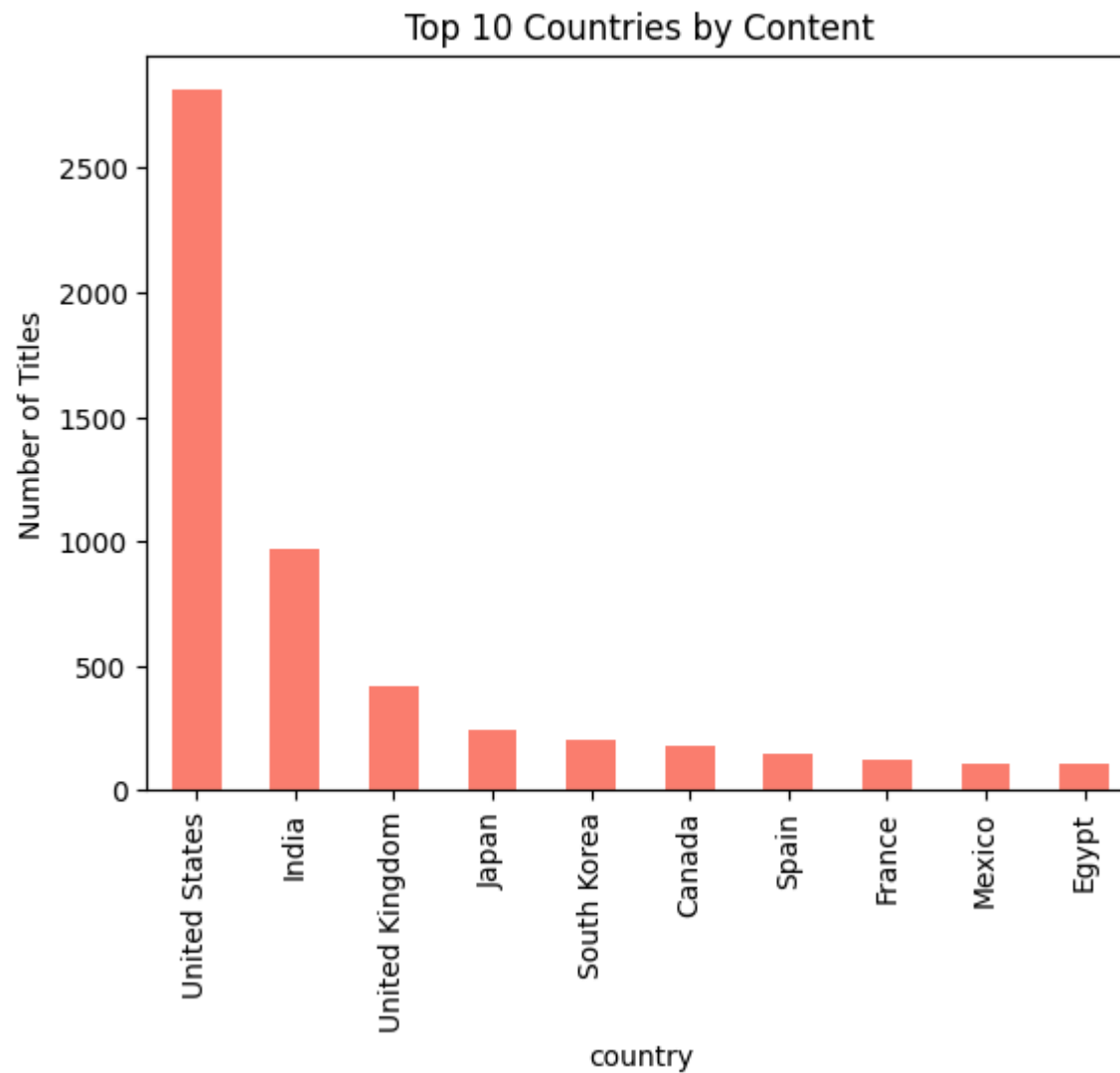
```
plt.title("Top 10 Genres on Netflix")  
plt.show()
```

↗ /tmp/ipython-input-7-2094673239.py:7: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set  
`sns.barplot(data=top_genres, y='Genre', x='Count', palette='pastel')`



```
top_countries = df['country'].value_counts().head(10)
top_countries.plot(kind='bar', color='salmon')
plt.title("Top 10 Countries by Content")
plt.ylabel("Number of Titles")
plt.show()
```



```

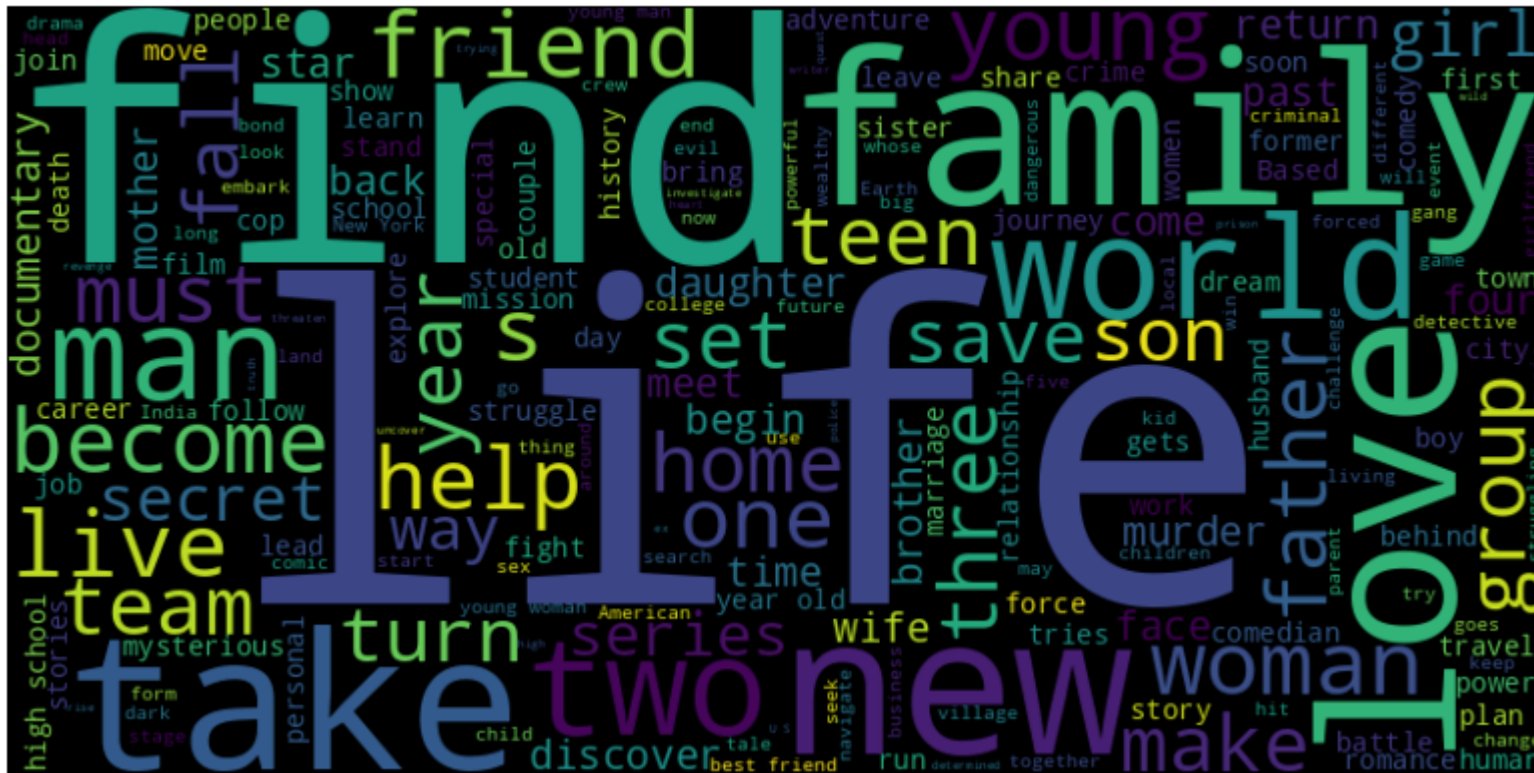
text = " ".join(df['description'])
wordcloud = WordCloud(width=800, height=400, background_color='black').generate(text)

plt.figure(figsize=(10,5))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis('off')
plt.title("Common Words in Descriptions")
plt.show()

```



Common Words in Descriptions



```
!pip install streamlit
```

```
import streamlit as st
```

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

st.set_page_config(page_title="Netflix Analysis", layout="wide")
st.title("📺 Netflix Content Analysis Dashboard")

# Load Dataset
df = pd.read_csv("netflix_titles.csv")

# Fill or Drop nulls
df['director'] = df['director'].fillna('Not Available')
df['cast'] = df['cast'].fillna('Not Available')
df.dropna(subset=['country', 'date_added'], inplace=True)

df['date_added'] = pd.to_datetime(df['date_added'], errors='coerce')
df.dropna(subset=['date_added'], inplace=True) # Drop rows where date_added couldn't be parsed
df['year_added'] = df['date_added'].dt.year

tab1, tab2 = st.tabs(["📈 Trends", "📍 Content Distribution"])

with tab1:
    st.subheader("Content Over Years")
    year_chart = df['year_added'].value_counts().sort_index()
    st.line_chart(year_chart)

    st.subheader("Movie vs TV Shows")
    st.bar_chart(df['type'].value_counts())

with tab2:
    st.subheader("Top Countries")
    st.write(df['country'].value_counts().head(10))

    st.subheader("Top Genres")
```



```

genre_list = df['listed_in'].str.split(', ')
genres = [genre for sublist in genre_list.dropna() for genre in sublist]
genre_series = pd.Series(genres).value_counts().head(10)
st.bar_chart(genre_series)

```

⇒ Requirement already satisfied: streamlit in /usr/local/lib/python3.11/dist-packages (1.47.0)  
Requirement already satisfied: altair<6,>=4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.5.0)  
Requirement already satisfied: blinker<2,>=1.5.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (1.9.0)  
Requirement already satisfied: cachetools<7,>=4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.5.2)  
Requirement already satisfied: click<9,>=7.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (8.2.1)  
Requirement already satisfied: numpy<3,>=1.23 in /usr/local/lib/python3.11/dist-packages (from streamlit) (2.0.2)  
Requirement already satisfied: packaging<26,>=20 in /usr/local/lib/python3.11/dist-packages (from streamlit) (25.0)  
Requirement already satisfied: pandas<3,>=1.4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (2.2.2)  
Requirement already satisfied: pillow<12,>=7.1.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (11.2.1)  
Requirement already satisfied: protobuf<7,>=3.20 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.29.5)  
Requirement already satisfied: pyarrow>=7.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (18.1.0)  
Requirement already satisfied: requests<3,>=2.27 in /usr/local/lib/python3.11/dist-packages (from streamlit) (2.32.3)  
Requirement already satisfied: tenacity<10,>=8.1.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (8.5.0)  
Requirement already satisfied: toml<2,>=0.10.1 in /usr/local/lib/python3.11/dist-packages (from streamlit) (0.10.2)  
Requirement already satisfied: typing-extensions<5,>=4.4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (4.14.1)  
Requirement already satisfied: watchdog<7,>=2.1.5 in /usr/local/lib/python3.11/dist-packages (from streamlit) (6.0.0)  
Requirement already satisfied: gitpython!=3.1.19,<4,>=3.0.7 in /usr/local/lib/python3.11/dist-packages (from streamlit) (3.1.4)  
Requirement already satisfied: pydeck<1,>=0.8.0b4 in /usr/local/lib/python3.11/dist-packages (from streamlit) (0.9.1)  
Requirement already satisfied: tornado!=6.5.0,<7,>=6.0.3 in /usr/local/lib/python3.11/dist-packages (from streamlit) (6.4.2)  
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from altair<6,>=4.0->streamlit) (3.1.6)  
Requirement already satisfied: jsonschema>=3.0 in /usr/local/lib/python3.11/dist-packages (from altair<6,>=4.0->streamlit) (4.  
Requirement already satisfied: narwhals>=1.14.2 in /usr/local/lib/python3.11/dist-packages (from altair<6,>=4.0->streamlit) (1  
Requirement already satisfied: gitdb<5,>=4.0.1 in /usr/local/lib/python3.11/dist-packages (from gitpython!=3.1.19,<4,>=3.0.7->  
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas<3,>=1.4.0->strea  
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas<3,>=1.4.0->streamlit) (202  
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas<3,>=1.4.0->streamlit) (2  
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.27->st  
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.27->streamlit) (3.  
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.27->streamli  
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.27->streamli  
Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.11/dist-packages (from gitdb<5,>=4.0.1->gitpython!=3.  
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->altair<6,>=4.0->stream  
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0-  
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.11/dist-packages (from jsonschem  
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,

```
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas<3,>=1.
2025-07-21 05:53:42.626 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.627 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.629 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.631 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.713 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.714 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.715 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.717 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.718 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.719 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.774 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.775 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.776 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.778 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.779 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.780 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.793 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-21 05:53:42.794 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
```