In [2]:
```
import pandas as pd
import matplotlib.pyplot as plt

# Load the dataset
file_path = 'C:\\Users\\tejas\\Downloads\\disney_plus_titles.csv'
disney_data = pd.read_csv(file_path)
disney_data.head()
```

Out[2]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	du
0	s1	Movie	A Spark Story	Jason Sterman, Leanne Dare	Apthon Corbin, Louis Gonzales	NaN	September 24, 2021	2021	TV-PG	
1	s2	Movie	Spooky Buddies	Robert Vince	Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...	United States, Canada	September 24, 2021	2011	G	
2	s3	Movie	The Fault in Our Stars	Josh Boone	Shailene Woodley, Ansel Elgort, Laura Dern, Sa...	United States	September 24, 2021	2014	PG-13	1
3	s4	TV Show	Dog: Impossible	NaN	Matt Beisner	United States	September 22, 2021	2019	TV-PG	S
4	s5	TV Show	Spidey And His Amazing Friends	NaN	Benjamin Valic, Lily Sanfelippo, Jakari Fraser...	United States	September 22, 2021	2021	TV-Y	1

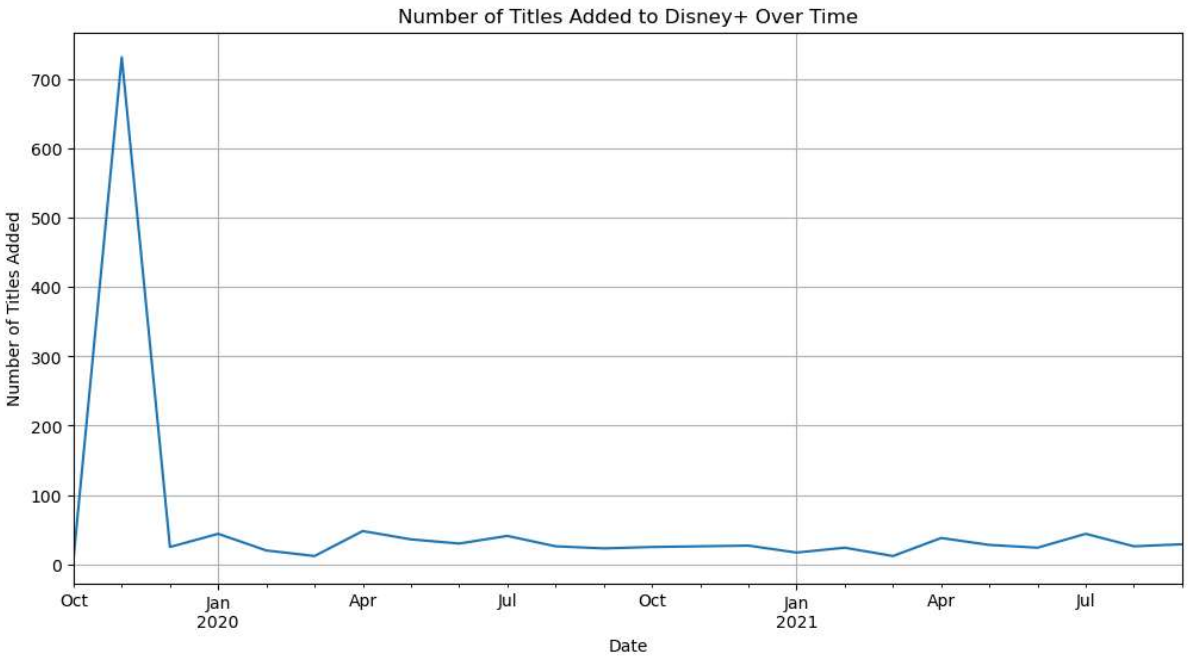
In [3]:
```
# Convert 'date_added' to datetime format
disney_data['date_added'] = pd.to_datetime(disney_data['date_added'])

# Group by year and month to analyze trends
monthly_trend = disney_data['date_added'].groupby(disney_data['date_added'].dt.to_p
disney_data.head()
```

Out[3]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	du
0	s1	Movie	A Spark Story	Jason Sterman, Leanne Dare	Apthon Corbin, Louis Gonzales	NaN	2021-09-24	2021	TV-PG	
1	s2	Movie	Spooky Buddies	Robert Vince	Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...	United States, Canada	2021-09-24	2011	G	
2	s3	Movie	The Fault in Our Stars	Josh Boone	Shailene Woodley, Ansel Elgort, Laura Dern, Sa...	United States	2021-09-24	2014	PG-13	1
3	s4	TV Show	Dog: Impossible	NaN	Matt Beisner	United States	2021-09-22	2019	TV-PG	S
4	s5	TV Show	Spidey And His Amazing Friends	NaN	Benjamin Valic, Lily Sanfelippo, Jakari Fraser...	United States	2021-09-22	2021	TV-Y	1

In [4]:

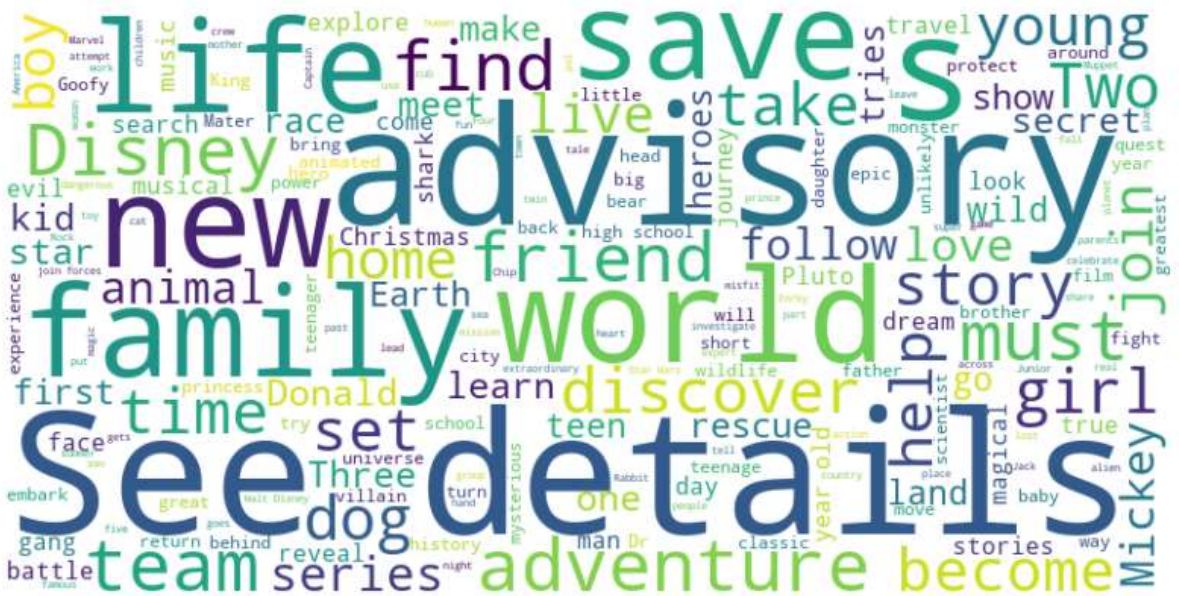
```
# Plot the trend
plt.figure(figsize=(12, 6))
monthly_trend.plot()
plt.title('Number of Titles Added to Disney+ Over Time')
plt.xlabel('Date')
plt.ylabel('Number of Titles Added')
plt.grid(True)
plt.show()
```



In [ ]:

In [7]:
```
#Perform sentiment analysis or text mining on unstructured data
from wordcloud import WordCloud
import matplotlib.pyplot as plt
from nltk.sentiment import SentimentIntensityAnalyzer
import nltk

# Ensure nltk resources are downloaded
nltk.download('vader_lexicon')

# Text preprocessing (optional: cleaning, tokenization)
descriptions = disney_data['description'].dropna()

# Generate a word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(' '

# Display the word cloud
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()

# Sentiment Analysis
sia = SentimentIntensityAnalyzer()
disney_data['sentiment_score'] = disney_data['description'].apply(lambda x: sia.pol

# Visualize the sentiment scores
plt.figure(figsize=(10, 5))
plt.hist(disney_data['sentiment_score'], bins=20, color='blue', edgecolor='black')
plt.title('Distribution of Sentiment Scores')
plt.xlabel('Sentiment Score')
plt.ylabel('Number of Titles')
plt.show()
```

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     C:\Users\tejas\AppData\Roaming\nltk_data...
```

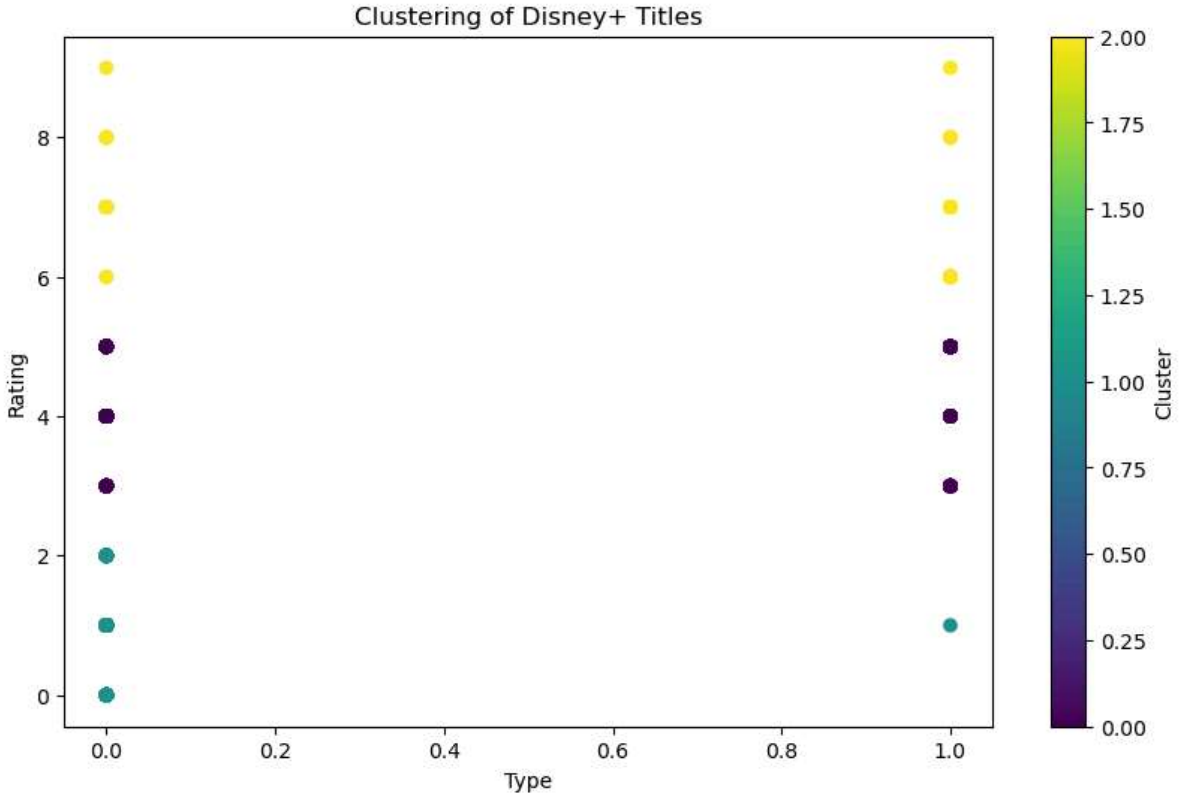


In [8]:

```
#Explore clustering or classification techniques for segmentation and pattern recog

from sklearn.preprocessing import LabelEncoder
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Encode categorical features for clustering
le = LabelEncoder()
disney_data['type_encoded'] = le.fit_transform(disney_data['type'])
disney_data['rating_encoded'] = le.fit_transform(disney_data['rating'].fillna('Unkr

# Prepare data for clustering (e.g., using type, rating, duration)
features = disney_data[['type_encoded', 'rating_encoded']]

# Apply KMeans clustering
kmeans = KMeans(n_clusters=3, random_state=42)
disney_data['cluster'] = kmeans.fit_predict(features)

# Visualize the clusters
plt.figure(figsize=(10, 6))
plt.scatter(disney_data['type_encoded'], disney_data['rating_encoded'], c=disney_da
plt.title('Clustering of Disney+ Titles')
plt.xlabel('Type')
plt.ylabel('Rating')
```

```
plt.colorbar(label='Cluster')
plt.show()
```

```
C:\Users\tejas\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: FutureWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
```



In [ ]: