```
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

```
!ls "/content/drive/My Drive/Colab Notebooks/cs131"
```

```
replies_highinfluence_count.txt  retweets_highinfluence_count.txt
replies_lowinfluence_count.txt   retweets_lowinfluence_count.txt
```

```
import pandas as pd
# Load high influence replies
df_high_replies = pd.read_csv('/content/drive/My Drive/Colab Notebooks/cs131/replies_highinfluence_count.txt', sep='\s+', header=None, names
```

```
# Load low influence replies
df_low_replies = pd.read_csv(
    '/content/drive/My Drive/Colab Notebooks/cs131/replies_lowinfluence_count.txt', sep='\s+', header=None, names=['Count', 'ID'])

# Load high influence retweets
df_high_retweets = pd.read_csv('/content/drive/My Drive/Colab Notebooks/cs131/retweets_highinfluence_count.txt', sep='\s+', header=None, nam

# Load low influence retweets
df_low_retweets = pd.read_csv('/content/drive/My Drive/Colab Notebooks/cs131/retweets_lowinfluence_count.txt', sep='\s+', header=None, names
```

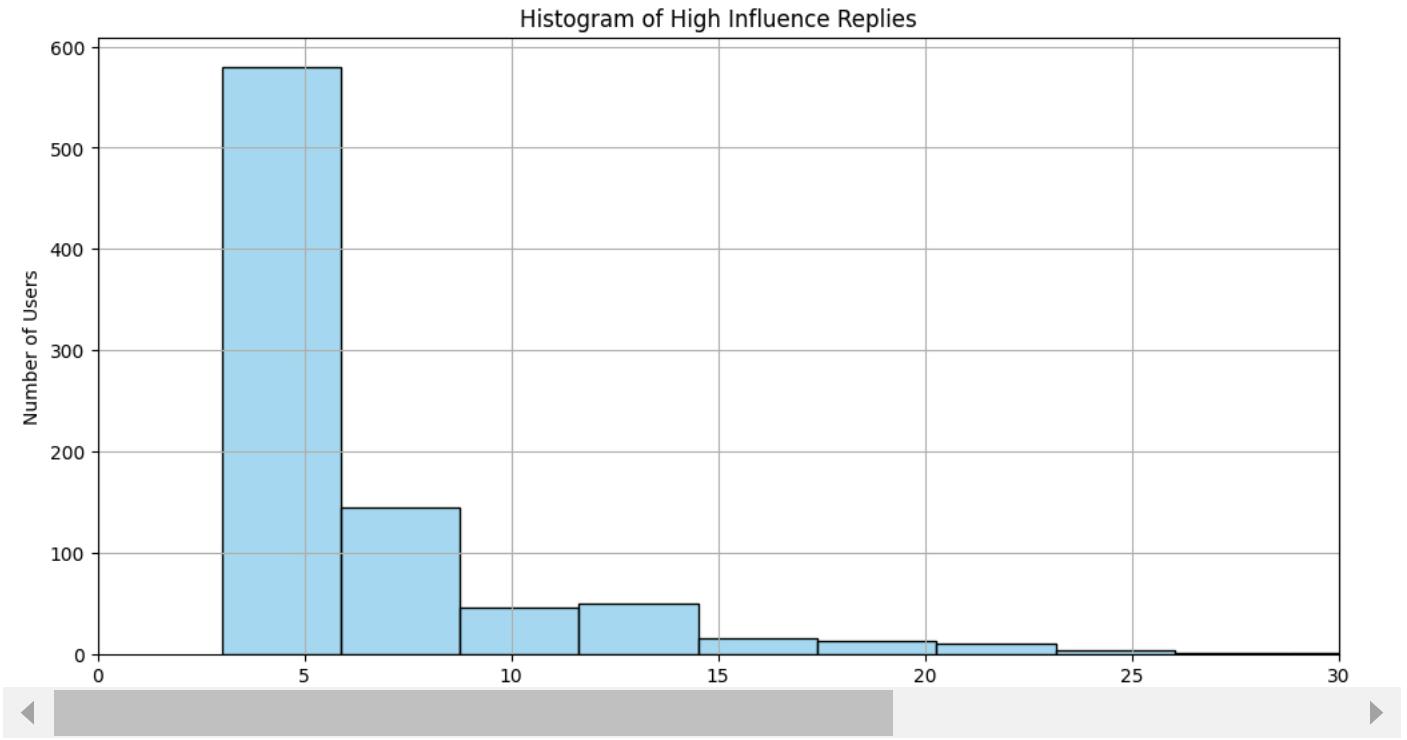
```
#check if the 4 files are uploaded
# print("Low influence replies")
# print(df_low_replies.head())
# print("\nHigh influence replies")
# print(df_low_retweets.head())
# print("\nLow influence retweets")
# print(df_high_replies.head())
# print("\nHigh influence retweets")
# print(df_high_retweets.head())
```

New Section

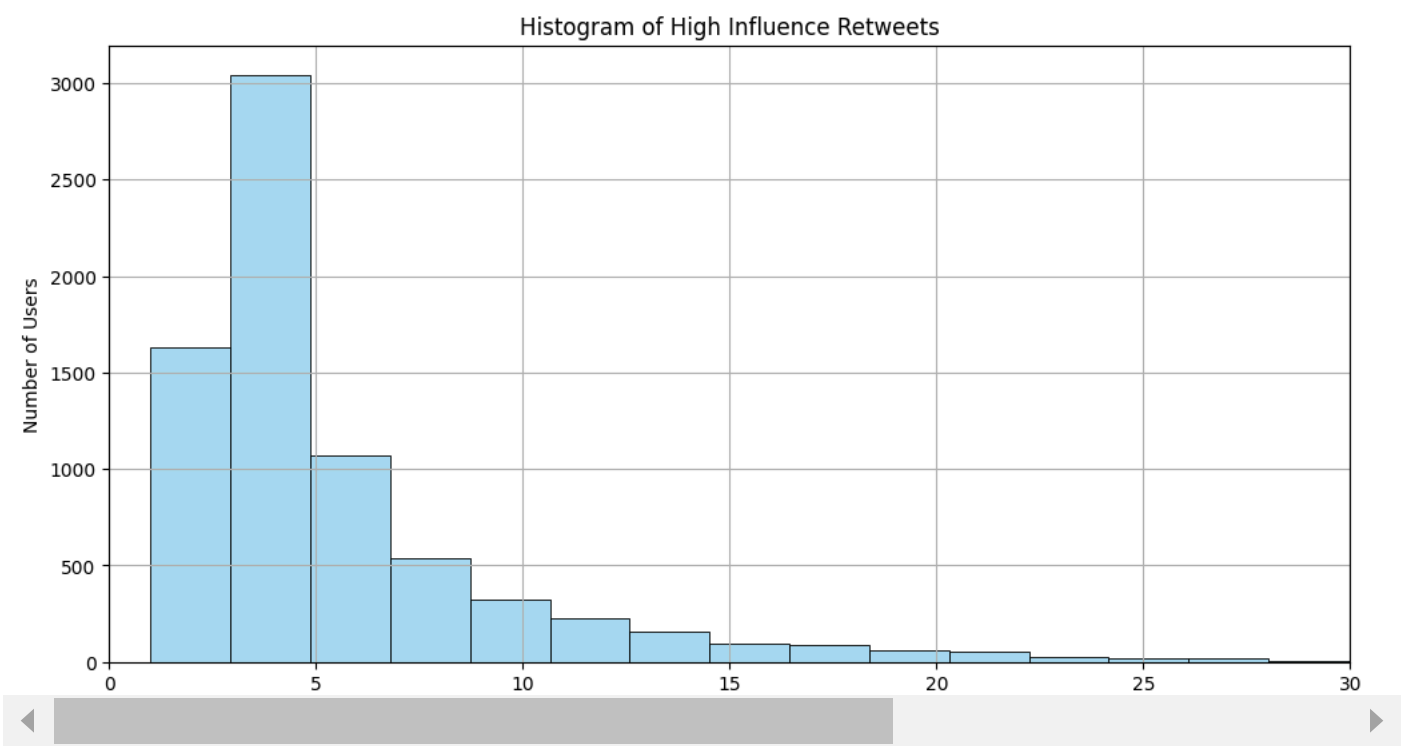
```
import seaborn as sns
import matplotlib.pyplot as plt

# Function to plot histograms
def plot_influence_histogram(df, column, title, bins, x_limit):
    plt.figure(figsize=(12, 6))
    sns.histplot(df[column], bins=bins, color='skyblue')
    plt.title('Histogram of ' + title)
    plt.xlabel('Number of ' + title)
    plt.ylabel('Number of Users')
    plt.xlim(0, x_limit)  # Set x-axis limit for data
    plt.grid(True)
    plt.show()
```

```
# High influence replies data visualization
plot_influence_histogram(df_high_replies, 'Count', 'High Influence Replies', bins=50, x_limit=30)
```



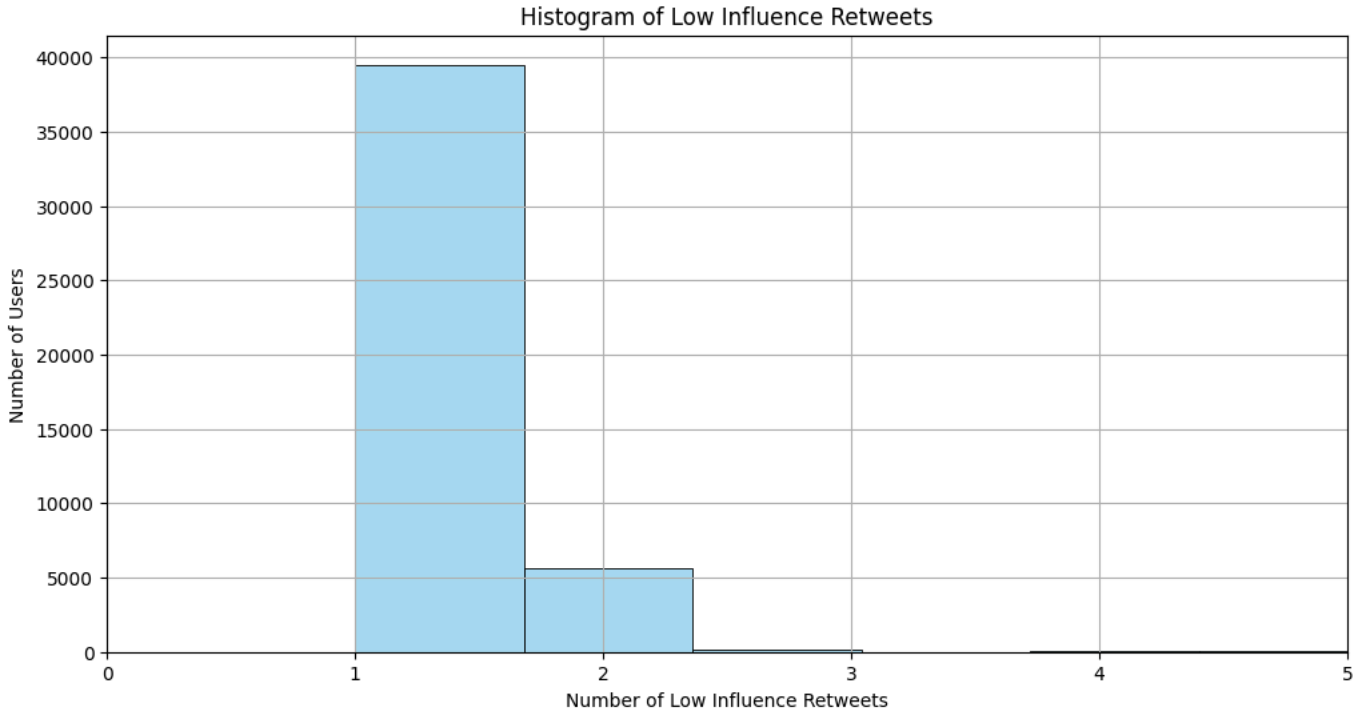

```
# High influence retweets data visualization
plot_influence_histogram(df_high_retweets, 'Count', 'High Influence Retweets', bins=100, x_limit=30)
```



Start coding or generate with AI.

```
# Low influence retweets data visualization
```

```
plot_influence_histogram(df_low_retweets, 'Count', 'Low Influence Retweets', bins=100, x_limit=5)
```




```
# Low influence replies data visualization
plot_influence_histogram(df_low_replies, 'Count', 'Low Influence Replies', bins=3, x_limit=5)
```