

# lab1\_DBSYS

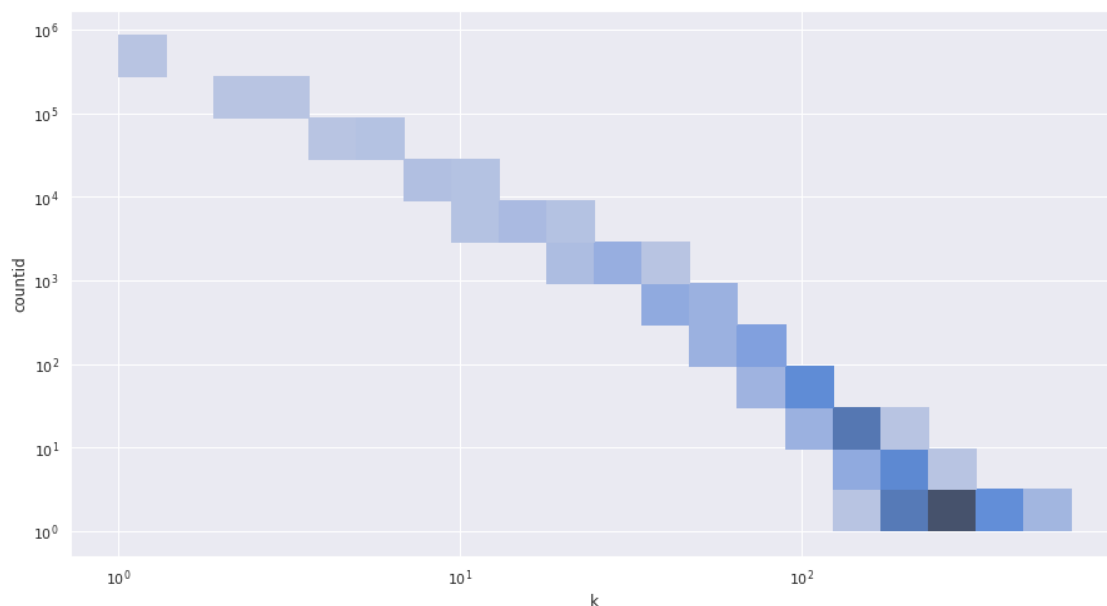
October 31, 2021

```
[ ]: import pandas as pd
[3]: import seaborn as sns
[ ]: import psycopg2
    # Connect to postgres database with simone user
    conn = psycopg2.connect("dbname=dblp user=simone")
    # Open a cursor to perform operation
    cur = conn.cursor()
    # Run the vis.sql
    cur.execute(open('vis.sql', 'r').read())
    results = cur.fetchall()
    cur.close()

[ ]: # store the results in a dataframe
    df = pd.DataFrame(results, columns=['k', 'countId'])

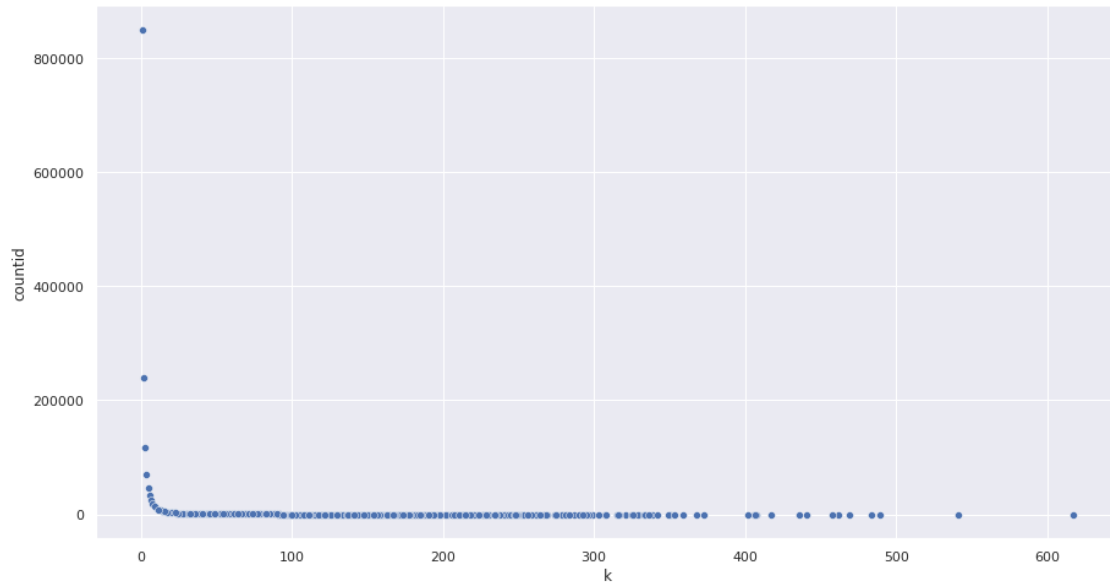
[7]: sns.set(rc = {'figure.figsize':(15,8)})
    sns.histplot(x=df.k, y=df.countId, log_scale=True)

[7]: <matplotlib.axes._subplots.AxesSubplot at 0x7fceefee4590>
```

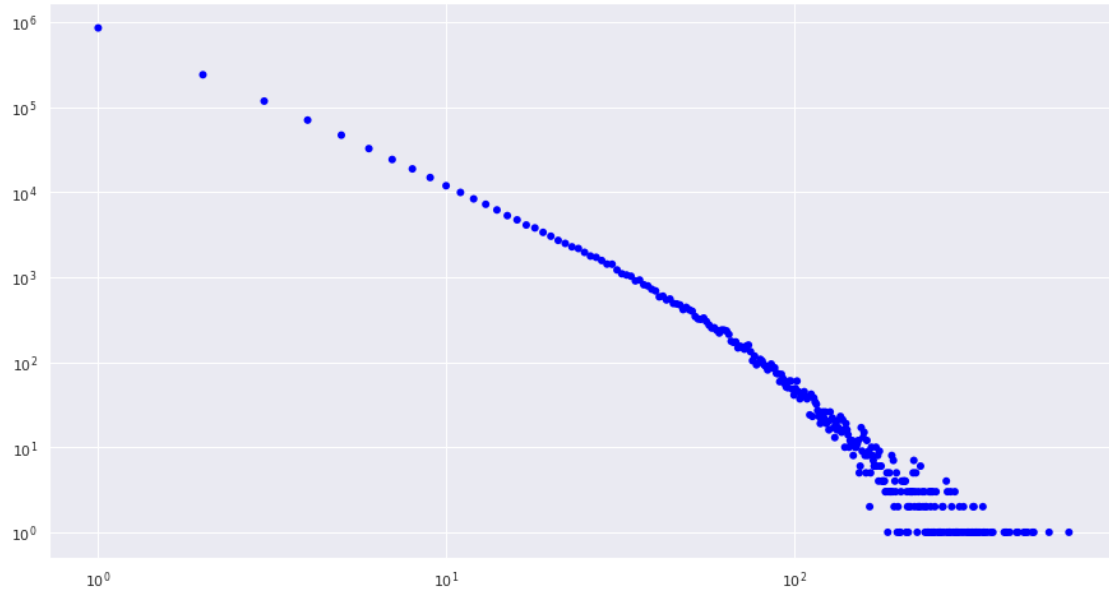


```
[ ]: sns.scatterplot(x=df.k, y=df.countid)
```

```
[ ]: <matplotlib.axes._subplots.AxesSubplot at 0x7f42a43675d0>
```



```
[ ]: import matplotlib.pyplot as plt
fig = plt.figure()
ax = plt.gca()
ax.scatter(x=df.k ,y=df.countid , c='blue', alpha=1, edgecolors='none')
ax.set_yscale('log')
ax.set_xscale('log')
```



Since the elevated number of bins the scatter plot does not give us the possibility to understand what is going on, neither with the log-log scale. For this reason I moved to scatter plot. It is possible to see that we have an exponential decrease trend (first figure) which is confirmed by the log-log scatter plot (second plot) which in the first part contains a decreasing straight line