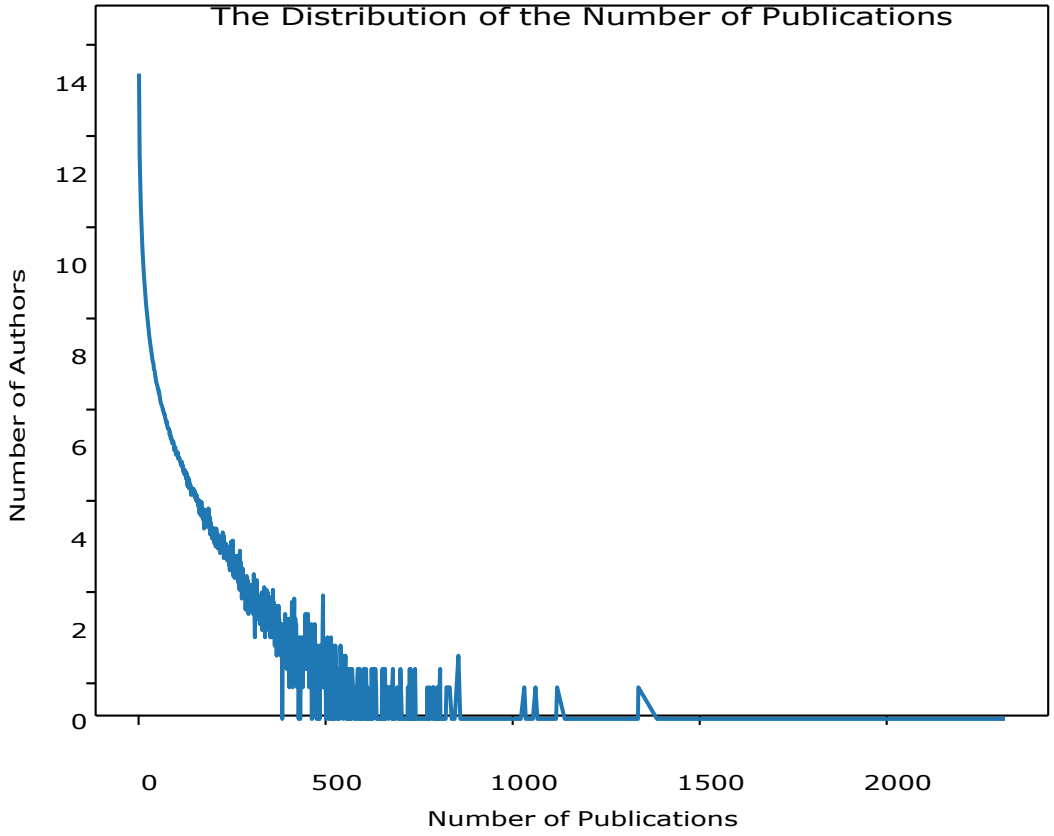

- We used a log scale, to determine best the law. As we see, the scatter plot has an exponential law distribution.

-First I installed psycopg2 package to be able to connect to a postgres databse.

-I run in my terminal python vis.py to execute the following script and get the graph of the distribution of the number of publications(This script is on vis.py file):

```
import psycopg2
import matplotlib.pyplot as plt
import numpy as np
def main():
    # Connect to the 'dblp' database
    try:
        conn = psycopg2.connect("dbname='dblp' user='postgres' host='localhost' password='Zineb.03SENANE'")
    except:
        print "I am unable to connect to the database"
    cur = conn.cursor()

    # Define the queries to be implemented
    queries = {
    'Number of Publications':
        '''
        SELECT NumPublications, COUNT(ID) AS NumAuthors
            FROM (SELECT ID, COUNT(PubID) AS NumPublications
                FROM published
                GROUP BY ID) AS AuthorPub
            GROUP BY NumPublications
            ORDER BY NumPublications;
        '''
    }
    # Draw the graph
    fig, axes = plt.subplots(1, 1)
    plt.subplots_adjust(hspace=0.8)
    for (name, query) in queries.items():
        cur.execute(query)
        rows = cur.fetchall()
        x = [row[0] for row in rows]
        y = np.log([row[1] for row in rows])
        axes.plot(x, y)
        axes.set_title('The Distribution of the ' + name)
        axes.set_xlabel(name)
        axes.set_ylabel('Number of Authors')

# Output the file
    file_name = 'graphOfDistributionOfNumberOfPublication.pdf'
    plt.savefig(file_name)
if __name__ == '__main__':
    main()
```