READ ME

Team Members:
Surabhi Singh:50208675
Vanshika Nigam: 50208031

1. **Featured activity: Analysis of Latin documents for word-co-occurrence**

- **Source Code : Written in Python**
    1. bigram.py
    2. trigram.py

- **Input :** Different number of Latin files for each run
- **Processing :** The word co-occurrence calculated on SPARK
- **Output of the files is of the format:**
    - \<Word1, Word2 \>\<location\>
    - \<Word1, Word2, Word3 \>\<location\>

- **Explanation of the output :** As shown via plots and data in the Jupyter Notebook attached , the time taken for an output to be generated varies directly with the number of files.