**Repo structure**

1. reference\_papers (contains all the 11 papers with online appendix)
2. README (explains what the repo contains and how to use each file)
3. src (11 py files + utility files)
   * 1 py file for raw data generation and feature creation
   * Utility (data cleaning, statistics (charts/visualization), data quality checks (checking null values, feature descriptions, etc.)
4. raw\_data (WRDS, Yahoo Finance, etc.) -> Based on data sources
5. result (11 csv files; unique?) -> Based on papers
   * + 11 sub-folders