**Data Collection Memo:**

1. Original Data

Options data was crawled from Nasdaq.com option chain. 125,020 records were obtained. However, since there are many records missing option price or implied volatility value, we trimmed the data rows by selecting implied volatility value that is greater than zero and less than two. There are 18,981 records left.

Files: DirtyData.csv; CleanData.csv

Codes: OptionDataWebGleaner2.py

1. Data Processing

Then we used bisection method to preliminarily run and select European options only, we set a tolerance in advance. There are 2,477 and 9,168 records left for 0.01 and 0.05 tolerance respectively.

Files: europeanOptions\_.01tol.csv; europeanOptions\_.05tol.csv

Codes: OptionSelection.ipynb