# **Data Merging Process**

## **CRSP**

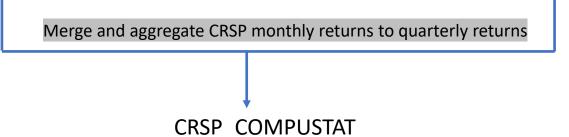
- Total number of monthly obs. in CRSP: 2,477,547
- Number of monthly obs. in CRSP that contains numeric returns: 2,427,687

#### **COMPUSTAT**

- Total number of quarterly obs. (10-Q + 10-K) in Compustat: 1,240,141
- Number of quarterly obs. in Compustat with 9-digits cusips: 1,239,632

### **EDGAR**

- Total number of 10-Qs in Edgar: 594,017
- Number of 10-Qs successfully parsed and downloaded: 575,579



- Number of cusip-quarter observations after
  - merging
  - dropping obs. with missing gvkey and return: 707,333

Merge

CRSP COMPUSTAT EDGAR

# 6.6. \_66.... 66.... \_12.6.

- Number of cik-quarter observations after
  - merging: 303,843
  - dropping obs. with total number of 10-Q words less than 1% threshold of the merged sample (1185 words): 300,806

#### Time Period

- Compustat & Edgar: 1993Q1 2020Q1
- CRSP: 1992 Nov. 2019 Dec.