EAPR Features

* Data extraction
  + As of now, extraction only works with WRDS repository.
  + Data is taken from CRSP, Compustat, and CRSP-Compustat Merged databases.
  + EAPR code executes SQL statements in the background. There is no option to execute custom SQL queries.
    - I figure if someone knows enough to create custom SQL queries, they should have the ability to extract the data themselves.
  + User will have the option to extract daily, weekly or monthly data with quarterly, semiannual, or annual portfolio rebalancing options.
    - The rebalancing date is the date where variables using Compustat data are recalculated.
    - Annual rebalancing dates will be at the end of June as per the methodology in Fama French (1992).
    - Quarterly and semiannual rebalancing dates will be determined using a reasonable cutoff for the proportion of SEC filings that are still unreported after a given date. For example, if say over 90% of Q1 filings have been released as of April 30th, then this will be chosen as the rebalancing date.
* Calculating variables
  + A preset list of variables will be available for the user to select from
    - There are two distinct types of variables: Fundamental and Technical
    - Fundamental variables include:
      * Market Equity
      * Book Equity
      * Book-to-market
      * Asset-to-market
      * Asset-to-book
      * Earnings-to-price
      * Operating profitability
      * Investment
      * CF-to-price
      * Dividend yield
      * Trading Volume of Stock
    - Technicals variables include:
      * Pre and Post-rank betas
      * CAPM Beta
      * Momentum
      * Moving Average
  + Variables are calculated during the extraction step.
* Data Cleaning and Filtering
  + The main feature for data cleaning will be outlier detection and removal or shrinkage of gross outliers, based on user preference.
    - Note: Need to include fill methods for this
  + The filtering techniques used in Fama-French (1992) are implemented and may be customized by changing the rebalancing periods, number of observations used to calculate technicals, etc.
    - Note: do not assume the reader is acquainted with Fama-French work. Outline the filters used, and then describe the generalized version of the filter.
  + We also plan to add liquidity filters as well. One such filter is the one outlined in Bailer et. al (2011).
    - Note: Expand on this filter as well as ideas for liquidity filters
* Portfolio Analysis
  + Functions to create quantile portfolios sorting on one or two variables
  + Statistics on the “response” variable will be computed and
* Cross-Sectional Regression
  + User will have the ability to perform linear regression over cross-sections of the time series data using robust and/or classical methods.
    - Note: Link description of time series to the extraction step. Also, specify that returns will be the default response variable, but other variables may be used as well
    - Functions for performing cross-sectional regressions using other asset pricing models in the literature (such as FF three and five factor models) will be included as well