1. Data: provides access to the data for all the other modules of the system both to read and write.
   1. Data Base: encapsulates the DB functionality in case the DB changes. Could provide data from files, a relational DB or whatever.
   2. Data Sets: provides access to these data sets: Toy, Exploration, Full, Test.
2. Data Preparation: imports, cleans, transforms, extracts features and generally speaking prepares the data for further processing. Includes acquiring the data by scrapping for instance. It could also include dimensionality reduction.
   1. Acquisition
   2. Preprocessing
   3. Feature Extraction
3. Data Exploration: calculates descriptive statistics, builds exploration plots, produces correlation matrices, etc.
4. Models: contains models to make predictions also the code to train them, select them, etc. It could include complex aggregations of models