

sdmay19-06: Quantitative Research Modeling Library

Week 2 Report

January 28 - February 3

Client: Joseph Byrum

Advisor: Srikanta Tirthapura

Team MembersJosiah Anderson — *Meeting Facilitator*Doh Yun Kim — *Scribe*Gabriel Klein — *Report Manager*Drake Mossman — *Communication Manager*Nathan Schaffer — *Overseer*Jacob Richards — *Quality Assurance Manager*

Summary of Progress this Report

During this week of progress each of the three subteams began progress on their individual portions of the pipeline. We have a few different people working for our client who are in charge of these same portions, so we had some initial meetings with them to make more concrete plans and to discuss inputs and outputs. In addition, we made the decision to abandon using Jupyter Notebooks within our repository in favor of regular Python files, we continued trying to get our continuous testing setup working, and we investigated whether packages such as modin or ray would be useful for parallelizing our code.

Past Week Accomplishments

Josiah Anderson

- Started working with Jacob on storing the output data from each module.
- Merged my code with master and fixed conflicts
- Set up meeting time with Josh at Principal
- Prepared our presentation for meeting with Principal

Doh Yun Kim

- Worked on the modeling step in regards with inputs and outputs standardization
- Tried and get example model working

Gabriel Klein

- Updated master branch with new database credentials
- Investigated feasibility of using modin or ray for parallelization

Drake Mossman

- Got rid of Jupyter Notebook files in git master branch
- Made Python files conform to consistent formatting
- Planning / communication for optimization prototype v1

Jacob Richards

- Researched CI testing problems
- Messed with python tests on local computer

Nathan Schaffer:

- Communication with Fikri about optimization part of pipeline
 - Planned out optimization part of pipeline
-

Pending Issues

We are still having issues getting our tests to run as mentioned in the previous report. Jacob will continue to investigate this problem this week. The research into modin and ray led to the conclusion that the transition to using those packages might be worth it for large enough datasets and without using certain operations, but we won't know if these conditions will be present until we have a prototype of the pipeline complete, so we'll set the idea aside for the time being.

Plans for Upcoming Reporting Period

Josiah Anderson

- Create table for storing the data created in the factor portfolios module
- Meet with Josh to figure out what data he needs for the initial diagnostic database

Doh Yun Kim

- Schedule meeting with Vishnu to get input on abstract models and inputs, outputs
- Work on modeling step more

Gabriel Klein

- Work on asynchronous database calls
- Plan roadmap out for the modeling step with Doh Yun Kim and begin work

Drake Mossman

- Initial Meeting with Fikri and optimization team
- Finish optimization ABC prototype v1
- Make example extension from optimization ABC

Jacob Richards

- Meet with Josh and Josiah
- Update CI server to fix testing
- Work on databases

Nathan Schaffer:

- First optimization meeting with Fikri and Drake
 - Start building prototype of optimization step
-

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Josiah Anderson	Dashboard/DB, Documentation, Presentation	6.5	94.5
Doh Yun Kim	Standardization of inputs/outputs on models	5	82.5
Gabriel Klein	DB Credentials, Parallelization	5	97
Drake Mossman	Solved.ipynb problems; improved .py format	6.25	96
Nathan Schaffer	Optimization planning, Communication	5	73
Jacob Richards	Fixing CI Tests	6	77.5