**Thursday March 21 - with Eunice (mentor) & Sining (instructor)**

* Addison: recap on what we did last week

About UI

* Sining: is the UI on a server? Do you plan to make an APP or a web application?

[Naoto: could you add anything here?]

* Eunice: push the UI online will cost , could help with it if share the code with Eunice
* Eunice: show the top 50 bonds in APP, but should be able to show all the bonds recommendations if the user wants to see

About rich and cheap bonds:

* Eunice: have rich and cheap bonds against itself and its peer. When onsite, traders wants to see the comparison.
* Sining: did you resolve the differences between recommendation differences between Hangyu, Addison, Yilin

Visit Vanguard:

* When to visit Vanguard? Can push back for another weeks as long as it is not the end of the month. **The week of April 19th**

Report and Code:

* Upload code on google folder
* Report in Latex

**Next week:**

* **Finish report**
* **Figure out differences between Hangyu, Addison, Yilin**
* **Find a more advanced recommendation system/model**

**Friday March 15 - with Sining**

Sining notes

SteveM showed a preliminary KNN model for recommending similar bonds by utilizing all 27 columns (including categorical columns) and giving them equal weights(without filtering by sector and ticker). Another version filtered by ticker. Also did PCA analysis and used first two principal components.

Sining’s suggestions: upweight ticker instead of strict filtering. Look into interpretation of the principal components as part of the deep EDA. **Also Steve please look into using multi-day data to construct metric** instead of single day. For example incorporate a correlation of spread trajectories over the last 6 months or a sum of squares of the difference (brainstorm time with teammates). I didn’t have time to say this during the meeting! Next step: work with Naoto to incorporate the model into the Django app.

Addison: showed code for implementing the JPM model. Ran out of time.

Jeff: daily excess return column is available. A combination of income and price differences. Vanguard is not particularly interested in evaluating returns. This will be used for portfolio construction rather than return oriented. Sining would like to see excess return used as an evaluation measure for a complete data science project.

Feel free to add below.

**Friday March 8 - with Eunice**

1. Recap on the progress
   1. UI design
   2. JPM model
2. Asked Eunice for
   1. Data on daily trading volume: this data is confidential and won’t be able to share with us.
   2. JPM paper on curve fitting methodology: sent in Slack group.
3. Confirmed that field trip is on **March 29th**.
4. Received two latest datasets (available in Slack).

* The new datasets cover the previous version, removed two columns ‘Security type’ and ‘Breakeven Spd’ . We should use the new datasets for future modelling.

1. Content on the first report (due 3/25):

* Results from JPM model
* KNN baseline model
* UI demo
* Format: Latex/python notebook

1. To-do:
   1. Proceed to recommendations, implement the baseline KNN on the new dataset - Hangyu, Jeff, Addison
   2. Do backtesting on Log fit results to see if it outperforms JPM results - Hangyu?
   3. Implement JPM model, adding the Issuer Curve and Sub-sector Curve. Cross validate with JAZ’s results - Yilin
   4. Send UI demo’s zip files to Eunice for approval - Naoto
   5. In-depth EDA on new dataset - Steve, Yilin
   6. Looking for publicly available data on bond trading volume - consult with Jeff first