**Introduction:**

* Reframe thesis around new story
* footnote correctly
* new sources/background
* change writing to match Jessica’s/TJ’s

**Lit Review:**

* Include more stuff on peer effects
* Include more stuff on low income mobility
* Include more stuff on development and education
* Reframe question and context
* Actually read papers (especially ones w/ influence on data)

**Data:**

* Read other papers to more accurately summarize data
* Understand and decide on summary charts
* Format charts and footnote correctly
* Explain measures, variables, etc.

**Empirics:**

* Update empirical specifications according to notes
* Rewrite final section and peer effects section to better explain what I’m doing and why
* Read Jessica’s/TJ’s to understand everything

**Results:**

* Finish Pie Charts
* Finish Kaplan Meier Plots
* Finish Log Rank Tests
* Finish Survival Models
* Finish Peer Effects Model
* Write Clustering Algorithm
* Finish Clustering Regression
* Finish Coefficient Stability
* Missingness analysis? Perhaps not needed
* Check Comparison students survival

**Discussion:**

* Writeup, limitations, conclusions
* Policy Impact

List of running questions:

* For survival variable, do I use number of years in star? This is a composite, non time-indexed measure (i.e. it doesn’t care if the student was in star between 1980-1981, only cares aggregate number of years)
  + In principle, my functional form doesn’t change if it is time indexed or not, either way attrition is still monotonically decreasing (i.e. more for 0-1 than 1-2)
  + Also, since I measure by entry cohort, doesn’t necessarily matter for if a kid left btwn 2nd and 3rd
  + Am I still teasing out variation especially due to changes between K and 1 due to re-randomization? (by controlling for class size should be ok right)
* Right now I’m using my own variables, I think the coding of the composite duration is incorrect because I want to understand that if they switched to a small class (i.e. the parent was able to lobby, does that lessen their chance of attrition? Tells smtg about preferences)
  + Can switch to composite metric later, but I believe it is wrong
* Do I need to show hypothesis tests for randomization? I think not b/c already well established, but let me know.
* Do I need to control for exit variables (e.g. teacher gender on exit? Will this somehow confound for students that stayed the whole time?)
* Are my predictors in peer effects model correct? Specifically including peers new test scores (if all simultaneously tested, this is a proxy for new peer ability in the class, right? So I should include it)
* Do I need to control for race/FL status/etc. of peers leaving? (e.g. “% of leavers on free lunch status, % of leavers that are non-white, % of leavers that are on special education?) – I don’t think I do since I assume that the test score is what is going to drive the main effect/what I’m trying to measure but potentially effects along similar race/socioeconomic lines may confound? But these are also elementary kids? Isn’t school info capturing most of that. Concerned with overfitting in this case.
* Do I need to log prop leaving (likely right skewed, e.g. not many classes w/ 75% leaving, but lot with 10-50%)
* Do I need to fit a model for students who don’t switch class types?