Vlad Lunch Notes

To get from Vlad:

* full essays and need statements for rejected projects
  + will send essay data for rejected, verify if everything is identical to approved essays online
  + if there are differences, check pre/post 2008
* prezi on the process of recruiting and posting project

Insights:

* biggest project: Set of playground for $60K
* Average days of projects: 30 days, 4m period
  + New project: passing savings to teachers for their next project
  + 4m limit can be manually extended
* point system: new teachers can’t post projects that are too expensive

Create Account:

* Certain teachers get emailed and recruited, they register on the site, and post drafts
* Third party verification for account creation
  + In cases no verification, can manually verify by calling the principle
* The supply is shipped to school

Posting and Verification:

* Teachers draft and submit
  + Add itemized resources (resource request)
  + Items not available from vendors: manually vetted by donorschoose
* Volunteers read and approve/reject projects
* Vendor system verifies all resource requested

Requirements:

* there’s a max limit on essay word count
* are all fields required to submit?

Funding:

* if a project doesn’t get funded, money go back to donors
* if items are not available anymore or teacher doesn’t need anymore, manually reconcile

If project gets rejected by vendor verification, not volunteers, what happens?

* don’t think so, verification happens before volunteers read and approve project

concern:

* data missing of rejected data
  + teachers get notified of rejected essays, Q: do they get why it was rejected?
* Need statement
  + ‘[fullfillment]’ 🡪 what is this?
* If see awkward strings:
  + Is it null/empty?
  + Anything predating 2007: not the same website and was a hack, use 2008-2014
  + Replace awkward system generated strings as empty strings
* Don’t look at last month
  + Delete all data after 10/21/14
* No categorical level data for resource ID

Feature to add:

* # of words present in the need statement that’s also in resources

On complexity

* Treat as engineering problem (KDD)
* if there’s same measurement for features
* online models: for sure need simplicity
* don’t think in dozens or hundreds