EDA Plan for Recommendations –

What is our overarching question?

* Do we have the data necessary to do what we think? What we make sure we have is a:
  + Do we have a full feature class?
  + Are there missings?
  + Will we trash them if so? (Yes)
* Users –
  + Merge checkin, photos, users, tips
  + At the review level -
  + Do counts on what’s there and what’s missing
* Businesses –
  + Merge business data to reviews

Background stuff for Jay and Maia to do

1. Column Headers
2. Get our user file open, etc

First step:

1. List all variables – and here’s how many observations they have
2. If we don’t have a certain threshold, here’s the ones that pass
3. Then a correlation matrix
4. Conclusion – Here are things we’re dropping, and here’s what we’re keeping

Among the things we’re keeping, what patterns do we have?

1. Basic stats on the types of users we have
2. Basic stats on the types of restaurants we have