☑ ~~Collect and pre-process a preliminary batch of data~~
☑ ~~Perform a preliminary analysis of the data~~
☑ ~~Answer one key question~~
☑ ~~Refine project scope and list of limitations with data and potential risks of achieving project goal~~
☑ ~~Submit a PR with the above report and modifications to original proposal~~

Team Meeting: 11:00 AM - 2:30 PM 10/15/22

1. The team as a whole spent the first part of the meeting reevaluating our goal of the project to ensure that everyone had the same vision of the project. Once everyone was on the same page, our team started to use some of the given data sets while looking for new data sets.
2. We were able to generalize a couple of the data sets. We took a look at the BDPA and graphed the points to show where small businesses are located. Then we took a look at a data set from kaggle called 7+ Million Company Dataset. We understand that this data set might be slightly out of the scope of the project, however, we thought it would be a good idea to take a look at industry types of businesses that have a small amount of employees. We were able to group by industry in Boston, which helped us answer the question about what businesses exist.
3. Looking at the chart, marking and advertising is the most popular business amongst small businesses in Boston, while legislative offices were the least common.
4. Once again we understand that the 7+ million company dataset might be outside of the district 4 scope, however, we believe that starting here will help us understand what types of small businesses would exist. Right now, some risks are introducing unnecessary data from datasets that would breach the scope, while skewing out analysis of such data. We, as a group, have to carefully look through data and make sure that out analysis is carefully written out.

Future Goals: Find more relevant data sets (use current ones on the document) and dive deeper into analysis and see how that gets us closer to the goal of the project. Moreover, we are trying to connect data sets without introducing bias or extraneous variables that would make our analysis irrelevant or incorrect.