# Sprint 4 Progress Report

## GitHub Repository

[Provide link to GitHub repository, which should contain:

1. README.md file that explains how to deploy your code (if you add/change things from usual, otherwise this is probably a copy of Augur's README.md)
2. Progress Report, which will be some kind of indication about whether or not you need to update your scope.

]

GitHub Repository Link: <https://github.com/JacksonHaskamp/augur>

Branch for Sprint 4: sprint4 <https://github.com/JacksonHaskamp/augur/tree/sprint4>

## Scope

[Indicate whether or not the project scope needs to be updated.]

* For sprint 2, we were focused on creating visualizations based off of json data Dr.Goggins provided us. Now we need to shift towards looking at python endpoints and how they work in Augur.
* We thought about what our Minimum Viable Product is.
  + Getting our visualization data to work, even if it’s with JSON file data we already have
  + We wanted to hook up the visualizations to the endpoint. But we weren’t sure how to do that.
* After meeting with Dr. Goggins, we decided that creating a Python API endpoint was beyond the scope of this project, given the time constraint we had.
  + Rather than utilizing a python API endpoint, we are focused on utilizing Google charts for visualizing Java Scripts manipulating JSON records.
* Dr. Goggins preferred us to focus on making nice visualizations.
  + He said that making the visualizations “live” (with an API endpoint) could be a different activity and would plug in nicely.

## Updates to Design Document and Requirements

[Indicate changes made to design document and requirements, if we made them.]

* Since our project scope changed to focus on the visualizations, we did not fully implement the functionality shown in some of the Design Document diagrams.
* Little to no changes to Design Document itself.
* In the requirements spreadsheet, removed “Worker deliverables” requirement because implementing machine learning workers was outside the scope of our project.
* Added a requirement specifically for the visualizations: using JSON data to produce a graph or chart.

## Testing Plan

[What is our testing plan? May depend on project context.]

* If we had gotten the Python endpoint working, we might have tried to test it with different data.
* Pytest unit testing can be used to test Python code, but we would need to have Python code written for this to be useful.
* For the visualizations, they could be checked to see if they make sense. For example, we could check if the axes are labeled correctly.
  + We could ask someone unrelated to the project to see if they can identify pertinent features of the data.
  + Along with this making sure data is not being squished or pulled too far apart visually.
* The visualizations could be checked to see how helpful they are at conveying trends or important information from the data.
* When creating visualizations, we need to make sure that data is easily understood at a first glance analysis.

## Team Reflection

[For this sprint, explanation of obstacles encountered, reflections, and goals.]

* Our team was not experienced with creating Python API endpoints, so we weren’t sure how to do that or how to connect the visualizations to the endpoint.
  + Also, we didn’t have much time left before Sprint 4 was due.
  + We met with Dr. Goggins and decided to change our project scope to focus on the visualizations.
* A goal for this sprint was to produce quality visualizations that are helpful to understanding the data.
* Another goal for this sprint was to produce a Minimum Viable Product. We tried to determine realistically what we wanted to accomplish given the time we had.
* We also had trouble determining which kind of visualization to use for the clustering data. We didn’t know how to interpret the clustering data or what it meant at first, but Dr. Goggins explained it to us.
* Clusters data:
  + cluster\_content is the cluster ID. There are usually from 0-8 clusters. (numbered 0-7).
  + A cluster is a group of repos that are grouped together as similar based on similar message content.
* Since our team members had different schedules and other projects at this time, not all team members were able to attend every team meeting.
  + The option to meet in-person or remotely helped us to be flexible.
  + We attempted to communicate via Discord about whether we’d be attending meetings and about what we were able to accomplish or planned to work on.