

Ultimate Analytics

Mid-Project Update

Highlights

- Data Cleaning/Ingestion - Developed a pipeline from source data through a cleaning step into ingestion onto s3 and preparation for hosting on RDS
- Back-end development of AWS infrastructure and connection for use in making the data available for the app

Progress

- Completed stories 1 & 2 of Theme1.Epic1 - Determined the appropriate data sources and considered model choice, will focus on productionizing model choice in coming sprint.
- Completed stories 1 & 2 of Theme2.Epic1 - Built out a data pipeline that will serve the app with data from the source, focus will move to exposing that data through the machine learning model
- Completed Theme2.Epic2.Story1 - Built out the necessary backend on AWS in order to host the app and data, gaining useful experience for future projects. Focus will move to building out a front end that a user can interact with.

Demo/Analysis

Copydata_bash script

throw_cleaning.Rmd (to be translated to Python)

Lessons Learned

I've learned a lot about building connections between cloud-based services and the importance of thinking about privacy/security early on and how that affects who will be able to access different parts of the app.

I've learned the intricacies of data availability even for data that is very easily accessible for me as the developer.

Recommendations

The focus of the coming sprint will be to build off of the back end work from the previous sprint.

This sprint will be user-focused, with large stories related to the front-end design of the app and incorporating user input to drive engagement with the app.