# Segment 2

## Presentation

Project Outline

✓ Selected topic

✓ Reason why they selected their topic

✓ Description of their source of data

✓ Questions they hope to answer with the data

✓ Description of the data exploration phase of the project

✓ Description of the analysis phase of the project

Slides (MICHELLE for now)

Presentations are drafted in Google slides (or PPT)

## GITHUB

Main Branch

All code in the main branch is production ready.

The main branch should include:

✓ All code necessary to perform exploratory analysis

✓ Some code necessary to complete the machine learning portion of the project

README.md

README.md must include:

✓ Description of the communication protocols

✓ Outline of the project (this may include images, but should be easy to follow and digest)

Note: The descriptions and explanations required in all other project deliverables

should also be in your README.md as part of your outline, unless otherwise

noted.

Individual Branches

✓ At least one branch for each team member

✓ Each team member has at least four commits for the duration of the second segment (eight total commits per person)

## Machine Learning Model

Team members submit the code for their machine learning model, as well as the

following:

ABE:

✓ Description of preliminary data preprocessing

✓ Description of preliminary feature engineering and preliminary feature selection, including their decision-making process

✓ Description of how data was split into training and testing sets

✓ Explanation of model choice, including limitations and benefits

COURTNEY:

✓ Description of preliminary data preprocessing

✓ Description of preliminary feature engineering and preliminary feature selection, including their decision-making process

✓ Description of how data was split into training and testing sets

✓ Explanation of model choice,

DATABASE

RACHEL:

Team members present a fully integrated database.

✓ Database stores static data for use during the project

✓ Database interfaces with the project in some format (e.g., scraping updates the database, or database connects to themodel)

✓ Includes at least two tables (or collections, if using MongoDB)

✓ Includes at least one join using the database language (not including any joins in Pandas)

✓ Includes at least one connection string (using SQLAlchemy or PyMongo)

Note: If you use a SQL database, you must provide your ERD with relationships.

Dashboard

MICHELLE

blueprint for the dashboard is created and includes all of the following:

✓ Storyboard on Google Slide(s)

✓ Description of the tool(s) that will be used to create final dashboard

✓ Description of interactive element(s)