

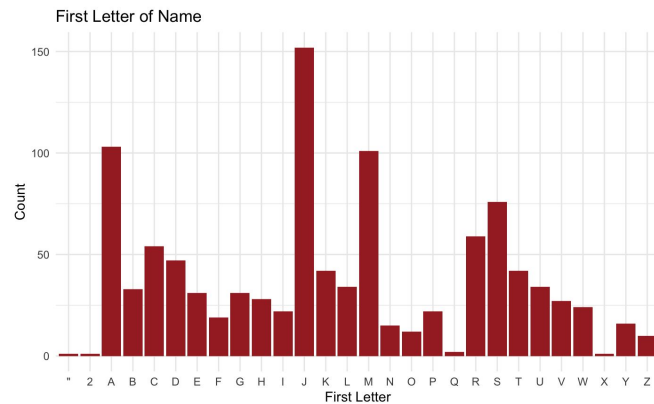
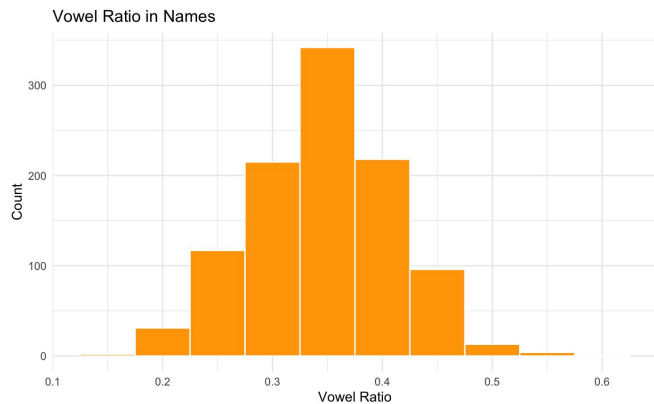
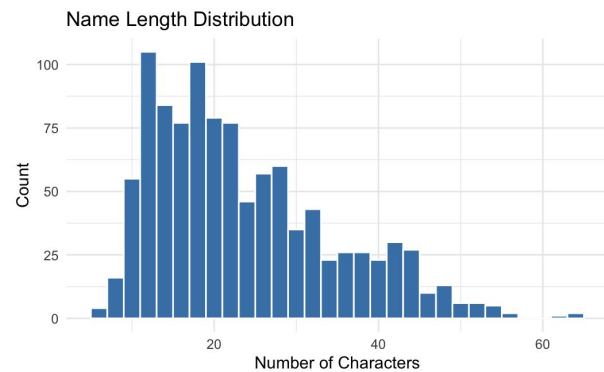
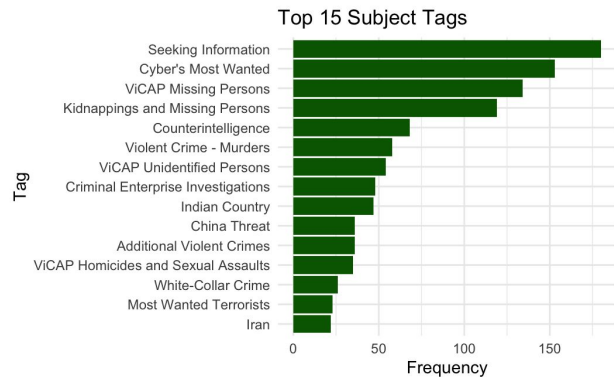
Project Proposal

Yuki Urata

Dataset Used

- FBI Wanted API
 - Public API provided by the Federal Bureau of Investigation (FBI), gives access to the current list of wanted fugitives and persons of interest.
 - Created to Increase Public Awareness, enable developers to create tools, and to support transparency.
 - API returns information on
 - Title (name)
 - Sex
 - Race
 - Nationality
 - Dates of birth
 - Subjects (description of the case)
 - >1000 rows

Some EDA



Model making

- Use name features to predict subjects
 - Random Forest
 - Is it ethically acceptable?
 - “Patterns in name structure that are associated with the way FBI categorizes wanted Individuals”
- Use name features to predict gender or the date born
 - No reason to use FBI database
 - Use [Genderize.io](#) (API) or [Agify.io](#) (API) and compare results
 - Predicts gender and age from a first name, based on global data