

# A6 Review

*Ayushi & Manthan Thakar*

*10/30/2017*

Commit hash for the review: 4fa900456ac1b048b438c0c2bc1468177b8a6f71

## Building & Running The Project

- The project doesn't build with the existing Makefile
- The Makefile seems to be missing targets for building the jar file and compiling the project
- There's no target for submitting the application to spark
- Since it would require major changes to the Makefile, we didn't run the project using the Makefile and actually resorted to running it from the IDE
- The project contains a `lib` folder with about **391 (446M)** jar files. The authors should consider shipping only the jar files required to build the project or taking the jar directory path from the user

## Code

- The author creates `SongRecord` and `ArtistTermRecord` to store the parsed record from the `song_info.csv` and `artist_terms.csv` files. The parsed fields are stored as class attributes each being mutable `var`. But since, these records are never mutated, it would be better to define them as `val`.

## Performance

- For the big corpus, we measured the results and found matching results to ours.
- It takes about 11 mins to run the program in the IDE, which seems reasonable.

## Report

- The report doesn't contain any results, it's hard to get anything out of it.