**Project Requirements**

 Use Pandas to clean and format your data set(s)

Group

 Create a Jupyter Notebook describing the **data exploration and cleanup** process

Amber- where to find?

 Create a Jupyter Notebook illustrating the **final data analysis (Conversational)**

Amber

 Create at least 6-8 visualizations (ideally 2 per "question" asked)

Group

 Save images of your visualizations (png or svg) to include in a markdown document (?) with findings of your project (and to distribute to the class and instructional team), and for inclusion in your presentation = Read Me document and Presentation

Amber

 Use at least one API

Steve

 Create a write-up summarizing your major findings. This should include a heading for each "question" you asked of your data, and under each heading, a short description of what you found and any relevant plots. This should be continually updated as part of your repository's readme.md file.

Amber

 Use at least one statistical test to accept or reject a hypothesis described in your project objective

Steve/Christian

***\*\* Distinguish listener versus play count***

**IDEAS for Tasks**

Replay ability stats per artist- Amber (?)----Ask about this – where to find?

* ***Question 1- Which countries’ users engage the most with the app?***
  + Top 10 countries by listener average – Steve- Done
  + Type of chart: bar
  + Type of chart: Stacked bar
* ***Question 2- Which artist are most popular based on their country play count?***
  + Top 20 artists by count of countries- Steve- Done
  + Type of chart: Bar
* ***Question 3- Which 5 artists are played most often globally? \*\*\*\*Why does this matter?\*\*\****
  + 5 top artists (+ other) globally by sum of play counts- Amber
  + Type of chart: Pie
  + Look at Steve’s dataframe
* ***Question 4- Who are the top artists and listener count in Hungary (top play count or listener Country)?***
  + Top artist and the listeners by (1) top country chart (Hungary): Scatter- Ishamael (?)
* ***Question 5- Who are the most popular artists by play count/listener per country?***
  + Top 5 artist per country (geomap with click action)- Christian
  + Type of visualization: Geo map
* ***Question 6- What are the most popular genres by play count/listener per country?*** 
  + Top 5 genres per country – Steve- dataframe/Christian - geomap
  + Type of visualization: geomap with click action- group effort
* ***Question 7- What is the GNP and play count of each country?***
  + GNP and playcount- Christian
  + R2 – 0.01, so zero correlation
  + We can accept the null hypothesis, there seems to be no correlation between GNP and play counts/listeners (?)

**Hypothesis**

By the data set, lastfm will not be used (Listener or Play Count?) less in lower GNP countries.

(Creates null and alternative Hypothesis)

Type of analysis: Statistical

**Limitations-**

* No time frame or trending ability
* Only Last.fm data
* Redundant in data (example?)
* API documentation is not up to date
* Listeners versus population (could be a GNP correlation)

Note for Amber: Look up Last Fm listener count and Spotify Listener count