Hey everyone,

**First Slide:** By a show of hands, who here likes music?

* Everyone, good.
* Well I do as well, my name is Cyrus, and personally I like all types of music except country. I know that may be a surprise as you all know I’m from Texas, but I’m also a millennial and us millennials don’t like country music.
  + But getting back on the music train, hip-hop and R&B are my two favorite genres, I’ve always wondered why certain artists have hit the mainstream big, while others didn’t in these two genres, so I decided to build a model predicting if a hip-hop and R&B album will go platinum or not. And the title of this slide is in honor of Mos Def who came out with a song called Mathematics, showing the intersection of hip-hop, R&B, and yes, data science.

**Second Slide:** How I chose my features?

* I was not able to pull every single feature that I was looking for each artist with Spotify’s API, so I limited my model to the top 50 artists
* The way I chose my top 50 artists was using the XXL Annual Freshman List which releases its annual Top 10 Freshman list of artists to watch ranging from unknown/underground artists, as well as artists on the rise which creates buzz among listeners and artists alike, gives many artists their first taste of fame. I also used the XXL list and targeted artists that did not want to be featured on the list.
* I also used prior knowledge and RIAA to see which artists have platinum albums.

(RIAA Freshman List, Spotify API, previous knowledge)

**Switch over to Tableau:** Before going into comparisons of artists, this map shows the rise of artists from 1996 – 2018. Keep in mind, my model leans more toward newer artists with a higher number of them coming out with their first album in 2011 and onward. This viz shows the year each artist in my model has come out with their first album, and the dots represent where in the US they came from. As you can see, Atlanta, NY, and LA have the highest populations of talent in the US, and two points that I did not include are Toronto and Indonesia. Toronto is home to Drake, Tory Lanez, and Nav, while Indonesia is home to Rich Brian.

Through the rest of this presentation, I’ll walk you all through two separate artists and how each features I used stack up against each other. Also, you have to guess which of those two artists are:

* Compare two rappers: Gucci and J Cole
* Platinum or nah? **Third Slide**
  + Gucci has none (0/12), not including mixtapes (show gold image)
  + J Cole has 4 platinum albums, 1 has gone 2X (4/4) (show plat image)
* Spotify followers: **Fourth Slide**
  + Gucci: 1,888,404 (ranking #)
  + J Cole: 5,074,440 (ranking #)
* Spotify pop index **Fifth Slide**
  + Gucci: 88
  + J Cole: 93
* # of genres classified by Spotify **Sixth Slide**
  + Gucci: 7 (dirty south rap, hip hop, pop, pop rap, rap, southern rap, trap music)
  + J Cole: 3 (pop, pop rap, rap)
* # of features on each album **Seventh Slide**
  + Gucci:
  + J Cole:

**Eighth (Tools) slide:** What models did I use:

* Decision Trees were not right for my model because it was overfitting my model, so I used SVMs initially but found that Multinomial Naïve Bayes was a better model for my data set as my features are independent of each other. I was preferring a higher total recall vs a high accuracy, so I decided Multinominal Naïve Bayes was best for my data. Followers do not necessarily correlate with a high Spotify popularity index as an artist could have a high amount of followers, but not have popularity based on them not coming out with an album recently. An example of this would be Post Malone, being #1 on # of popularity index, but #14 for the number of followers.

(Decision Trees, Naïve Bayes, Flask, SVMs)

**Q&A (Nineth) Slide:** This project was extremely fun, and I do look to expand on it. Some features I wish to add are the number of artists that are features, which have gone platinum, meaning in 2016 if post Malone was featured on an artists album then he would be included in this feature because he did not go platinum until 2017. Thus showing a deeper understanding and specificity if features do or do not matter.

Thanks everyone, I hope I gave y’all new content to listen to. Do y’all have any questions