We Are All Trying to Describe Something And We Don't Have the Words*

The Jack Antonoffification of Pop Music?

Michaela Drouillard

4/17/23

In this essay, I will be exploring Jack Antonoff's production style using data from the Spotify API. Drawing from an interview with Caleb Gamman, a guy who went viral for guessing which song on Taylor Swift's Midnights album was produced by Antonoff within seconds, I will delve into the specific characteristics that define an Antonoff-y song, and assess whether or not the audio features defined by Spotify can capture this. I plan to lasso Glen McDonald,the Spotify engineer behind these features, into explaining them to me. There will be a Bayesian inference model. It will be a logistic regression. We do not know whether or not it will be good. Finally, I will armchair a little bit about the cultural significance of Antonoff's music and the reasons behind his ubiquity in the music industry.

1 Introduction

Sometime in early October of 2022 there was a chemical shift in my brain and I couldn't stop listening to Taylor Swift. I couldn't listen to anything other than Fearless or Red when deadlifting, walking to a friend's house, folding laundry, etc. I listened to so much Taylor in that month alone that she ended up as my #4 artist on Spotify Wrapped, which stops collecting its data for the year on October 31st. It didn't even capture what happened to me after Midnights dropped in November.

I was standing outside of the Garrison eating a slice of pizza with some friends, and one of them asked what music we had been listening to lately.

ME: "Midnights!"

ZACH: "Is it good?"

---

*Code and data are available at: LINK.

[I glitched.]

ME: "No!"

["*No*?"]

A STRANGER WALKING BY WHO HEARD EVERYTHING: "EXACTLY!!"

Midnights was produced by Swift and her longtime collaborator Jack Antonoff, who is the subject of this paper, and, I suspect, the reason behind both my blacking out, and the stranger's "Exactly".

Antonoff is prolific, to the extent that it is starting to bother people. He has produced and co-produced music for Lorde, Taylor Swift, Florence and the Machine, The Chicks, Clairo, The 1975, Grimes, Zayn, Pink, Lana Del Rey, Olivia Rodrigo, the Minions: The Rise of Gru Soundtrack, and more. All of this and his own bands too (Bleachers, and fun.). When I was getting brunch with a friend one Sunday and a Swift song came on, and she said "Jack Antonoff's gonna kill us all", we both laughed. What my friend was referring to, of course, is the fact that Jack Antonoff and his sound seems to be everywhere these days. As Andrew Marantz writes in the New Yorker: "When his band releases an album, the world responds politely. When he produces one by Lorde or Lana Del Rey or Taylor Swift, the world wobbles on its axis".

What I am trying to get to here is that I don't actually really mind Jack Antonoff's music, I kind of like it, but I still say his voice with the same tone that everyone else does. There's just something about him… this sense that he's everywhere… and that he's just such a nice guy… Why do music people seem to be a little annoyed with him? What is it about Antonoff music that is so Antonoff-y? Is the Jack Antonoff-fication of pop something we should be worried about?

To find an answer to this question, I turned to statistics, people who know more than me about describing music, and the overlord of engineered virality itself – the Spotify API. There will be a Bayesian inference model. It will be a logistic regression. We do not know whether or not it will be good.

2 An Interview with Caleb Gamman

Around the time that Midnights came out, I came across a video on Twitter of a guy, Caleb Gamman, guessing which songs on Midnights were produced by Antonoff within seconds. The tweet reads: "im able to instantly detect if jack antonoff worked on a song due to a visceral hatred of his production style". He nails it. Only one song on the album wasn't produced by Antonoff, and he guesses it.

I reached out to Gamman to ask him to explain Antonoff to me, only to find out, on the morning of the interview, that Gamman had released a 29 minute video essay on his Youtube Channel covering:

(you can skim the quote chunks)

1. what makes an Antonoff song so Antonoff-y

> I'm literally not an expert and of course I don't actually claim to always be able to 100 identify a Jack Antonoff song it's not like I've been practicing the way that I can tell if he's produced something is basically if it sounds to me like he's produced it I know that's annoying I will clarify first I think it's mostly in the vocal treatment and that's the thing that is the most off-putting to me but just every decision uh feels a little bit wrong to me I don't know of course if I wanted to be cringe in CinemaSins I could point out the Juno 6 Bass the perfectly stacked vocals the Juno 6 pad the wide and tight plate Reverb the Juno 6 sword the goo goo gaga the man hates treble unless it's just random crap swishing around everything gets low passed with thin 707 drum samples thin 808 drum samples then 909 drum samples thin Lin drum samples thin thin drum samples all the resonances the harmonics get removed everything is sort of squished down to the fundamental reduction in Timbre everything exists in a small frequency range everything becomes a tone it's an extremely consonant sound with no interplay between the different elements now that is a skill to do and it produces a cleaner sound and the only problem is I viscerally hate that sound

2. his relationship to music growing up (some gems in here)

> hing I got into drumming at the age oflike six so maybe as a drummer the drumsamples are of a fence I wasanachronistically raised on cassettes soyou know maybe I like the sound ofaudio and maybe that explains I tend tolike dissonance and in harmonic soundinterplay of elements Distortion noisethe Clashing of elements I imagine I'velistened to more SoundCloud in my lifethan radio so uh maybe I only want tohear amateur internet loser I'veliked over 10 000 songs on my Spotifyaccount uh and I can identify lot ofthem off the bat but I'm not that greatat identifying music generally so it'sjust uh compartmentalized mediaconsumption habit I first did musicproduction in primary school I mademashups and dubstep in it as a lateteenager I extensively used Juno 6emulators to make embarrassing synthwaveadmittedly that's probably not thesubconscious Association that he's goingfor lyrics aren't of much interest to meit's just a sound in the mix and somaybe I don't like the focus on vocalI've previously described it as a fightor flight sort of thing so maybe itliterally is tapping into some memory oftrauma I have an unusually small earcanal so maybe I'm just hearing the highend more acutely I was hired as ateenager

to play drums on uh a lot ofJack Antonoff songs for a universityGlee Club performer so maybe I don'tlike it because of the embarrassment ofhaving ever been associated with a Gleecloth I liked the album 1989 but to makethe song out of the woods palette ofwool to me I had to mash it up with M83Midnight City and that got a dmcatakedown so maybe it's a personalVendetta and of course I have known anddo know people who like this music somaybe I secretly hate my friends andfamily but I don't know if any of thattells us anything uh I have a taste inmusic w

3. An elegant and simple defense of taste

   "i did use the phrase visceral hatred. but i did so only to describe my visceral hatred. i cant stand the sound of it. i hate it. and the way in which i hate it is viscerally. but i didnt go beyond that because it doesnt go beyond that. nothing against the guy, i like taylor swift's stuff. it's just subjective taste…

4. His tongue in cheek response to some of Twitter's negative response to his video

   …. [it's just subjective taste]. there's just nothing to engage with, so ive had to fabricate a reason for controversy. and what ive come up with is this. *looks at phone* the real reason iive been saying for years that i have a visceral reaction to jack antonoffs production style, the REAL reason is this -- cause during the rollout cycle of the album midnights, i wanted to ascribe taylor swifts authorship to a man and diminish women's voices. so actually, i'm a sexist. so now thats out there, publishhed as facts, in the news. with all the other facts…"

5. Basically a refutation to the entire premise of my interview

   "what makes you so special and unique that you can tell the production of a song?" "nothing"

6. Which was refuted even harder when I found a screenshot he tweeted of his tweet drafts in June 2021, one of them reading: "people often say to me:"caleb, what's wrong with jack antonoff¿'. don't ask questions"

caleb gamman
@calebgamman

collected unpublished jack antonoff hate 2016-2020

people often say to me: "caleb, what's wrong with jack antonoff?" don't ask questions

i don't actually have any ill will towards jack antonoff, all i'm saying is he wouldn't be able to make that music if he was dead

at the crj concert fantasising about glassing jack antonoff in the face

a lot of people have been around jack antonoff, and every single one of them has made the active decision to not kick him to the ground, spit on him and...

whenever an artist makes a genuinely good album jack antonoff pops his head around the corner "now let's do a silly one"

i will enjoy literally any piece of music

It is a perfect video essay. I pestered him anyway.

5

"In the treble mix of the song, like in the very upper end, like sort of above where the vocals are. He often really crushes that down. Like it's lower in volume than anything else, which is sort of a weird effect. Like, when you hear someone speaking, you hear a lot of that, like, noise in their voice. When you yourself are speaking, you hear less of it, right? You hear more of your own voice bouncing around in your head. And so it sort of creates this effect of – **like he mixes his vocals the way it would sound to the person who's singing them. Which is sort of strange.** He often has no vocals on that upper end, which is very unusual, and then often there are random little bits of noise happening up there which is the sort of thing her likes to do."

Gamman is a fan of messier music.

"Jack Antonoff does a very good job of making stuff very clean. So you know, all of the notes that are happening simulaneously are perfectly stacked. They're like mathematically stacked .. Like, you have the harmonic spectrum where everything is like a multiple on top of each other. So that's like your notes on a keyboard. And if you played a keyboard note that was a little bit off where it's supposed to be, you're gonna create this interesting effect where these two notes sort of clash against each other and they wiggle around in terms of their tone because you've got these two harmonics happening that don't match. And Jack Antonoff avoids that quite strictly"

"His drums are low past, like the vocals, so you don't get like the tinny high end in the drums or whatever because that tinny high end in the drums is like a bunch of random harmonics all over the place. So if you cut that out, you just get the tone of the drum. And then he achieves those effects using the same sounds like literally the same synthesizer on the same drum machine and stuff."

"If you used a cscope spectrogram sort of thing to look at a Jack Antonoff song, you're gonna see a lot of big lines. And then if you're listening to a song by someone else, you're gonna see like a bunch of sort of like static wiggling. It's gonna be really jagged because you've got a bunch of harmonics that don't exactly mix with each other, so they're stacking in all sorts of weird different ways"

3 Numbers

TLDR of the Gamman interview is: Taste is subjective. People who can play instruments tend to be better at identifying and describing sounds. Cultural criticism is a hall of smoke and mirrors. Antonoff is everywhere.

Can we translate any of this into something observable? Can we quantify the Jack Antonoff-ification of pop music?

4 Data

I pulled the data using the Spotify API, which I accessed using the `spotifyr` package in R. I used the `get_artist_audio_features` and `get_track_audio_analysis functions` function to acquire data on Lorde, Taylor Swift, St. Vincent, Lana Del Rey, The Chicks, Florence and the Machine, and Bleachers discographies.

4.0.1 Artist Selection

I chose these artists because they've produced multiple albums, and at least one album with Antonoff, as opposed to just individual songs. I chose the Chicks because I figured that collaborating with Antonoff after an extended gap in a longer spanning career than the other artists might offer some interesting results. The data points at beginning of their career reflect music that existed in a different time from the other artists in this dataset.

4.0.2 Data Cleaning

I joined these data sets using the `rbind` function in base R, and created an additional variable called "jack", which contains a 1 if Jack Antonoff produced or co-produced the song, and a 0 if he had nothing to do with it. I got the information on which songs Antonoff did or did not produce from the "Jack Antonoff Production Discography" Wikipedia page.

4.0.3 Spotify's API

So what do any one these features mean?

The Spotify API returns JSON metadata directly from the Spotify Data Catalogue.

4.0.3.1 Audio Features Dataset

The `get_artist_audio_features` provides a general profile of each song, according to variables developed by spotify, which makes for simple comparisons across artists.

However, it's hard to gauge exactly what variables like "danceability" or "valence" mean, even using the documentation. For instance, "valence" is described as "A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry)". "Danceability" is described as "how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable". Neither of these definitions fully cover what metrics and thresholds go into something like "danceability".

However, whatever these features capture is easier to compare across artists in visualizations than more granular details about bars, beats, duration, and tatums[1].

This is pretty vague. Especially valence – is the positivity related to lyrics, or something else? I compared the valence scores of classical music tracks to see if this score is based on lyrics, or the audio itself. Jupiter's valence score is 326.78% higher than Mars. This not only indicates the valence is based on audio features other than lyrical content, but that it is in fact capturing a difference between the two songs' profiles.

```
#| echo: false
#| warning: false
library(dplyr)
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':

    filter, lag
```

```
The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
sub_planets<- read.csv(here::here("inputs/data/subset_planets.csv"))
sub_planets |>
  select(artist_name, valence, danceability, track_name) |>
  head() |>
  knitr::kable(
    col.names = c("Artist Name",
                  "Valence",
                  "Danceability",
                  "Track Name")
)
```

---

[1] -- a word I just learned, which refers to "the lowest regular pulse train that a listener intuitively infers from the timing of perceived musical events". The full documentation for each variable included in this study is available in the Appendix

Artist Name	Valence	Danceability	Track Name
Gustav Holst	0.0527	0.204	The Planets, Op. 32, H. 125: I. Mars, the Bringer of War
Gustav Holst	0.2250	0.298	The Planets, Op. 32, H. 125: IV. Jupiter, the Bringer of Jollity

I looked at the valence measures on Flight of the Bumblebee, a song that's much faster and "upbeat" than Mars and Jupiter. What I found complicated my understanding of valence – different performances of the same song had different valence scores. Some were arias, played by different instruments. These tend to have higher valence scores. The lowest recorded valence is 0.2410, and the highest is .9230. What gives?

```
sub_bees<- read.csv(here::here("inputs/data/subset_bumblebee.csv"))
sub_bees |>
    select(artist_name, valence, danceability, album_release_year, track_name) |>
    head(17) |>
    knitr::kable(
      col.names = c("Artist Name",
                    "Valence",
                    "Danceability",
                    "Album Release Year",
                    "Track Name")
    )
```

Artist Name	Valence	Danceability	Album Release Year	Track Name
Nikolai Rimsky-Korsakov	0.360	0.448	2023	Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.657	0.374	2021	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Double Reed Ensemble)
Nikolai Rimsky-Korsakov	0.742	0.476	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Flute & Orchestra)
Nikolai Rimsky-Korsakov	0.534	0.509	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Violin & Orchestra)

Artist Name	Valence	Danceability	Album Release Year	Track Name
Nikolai Rimsky-Korsakov	0.742	0.476	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Flute & Orchestra)
Nikolai Rimsky-Korsakov	0.534	0.509	2018	Tale of Tsar Saltan: Flight of the Bumblebee (Arr. for Violin & Orchestra)
Nikolai Rimsky-Korsakov	0.241	0.429	2017	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.923	0.663	2012	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.241	0.429	2011	The Tale of Tsar Saltan: The Tale of Tsar Saltan: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.245	0.429	2011	The Tale of Tsar Saltan: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.407	0.413	2006	The Tale of Tsar Saltan: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.359	0.226	2004	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.510	0.425	1999	Rimsky-Korsakov: The Tale of Tsar Saltan, Act III: The Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.466	0.457	1998	The Tale of Tsar Saltan: The Flight of the Bumblebee (Arr. Rachmaninoff)
Nikolai Rimsky-Korsakov	0.444	0.440	1997	Tale of Tsar Saltan, Op. 57: Flight of the Bumblebee: Flight of the Bumble-Bee
Nikolai Rimsky-Korsakov	0.481	0.458	1995	Tsar Saltan, Op. 57: Flight of the Bumblebee
Nikolai Rimsky-Korsakov	0.481	0.458	1995	Tsar Saltan, Op. 57: Flight of the Bumblebee

I asked Glen McDonald, a "data alchemist" (see: Principal Engineer) who worked at The Echo Nest, a machine listening start-up that was acquired by Spotify in [YEAR]. He still works at Spotify now.

Glen told me that Valence and Danceability we created by giving thousands of examples to college interns and asking them to tag whether a song was positive or gloomy, or danceable or un-danceable. Variables like energy or instrumentalness, on the other hand, were determined through machine listening techniques. The human subjectivity part only came in when it came to tweaking the training data. (My favourite example Glen gave was that the ML treated banjos like a human voice, giving bluegrass songs with banjos high Speechiness ratings. Because there were no banjos in the training data, they were registered as human).

Glen said that the way he's made use of these variables is by combing Energy and Valence to get a quadrant, which usually turns out "pretty good".

[discuss the convo we had about how even two humans can have disagreement about a song's positivity, let alone human vs machine. i.e. an elliot smith song where the lyrics are depressing but the melody is happy. or a happy song but you know that the band died in a plane crash yesterday. etc.]

The final audio features in this study data set contains the variables: "artist_name", "track_name", "energy", "danceability", "energy", "key", "loudness", "mode", "speechiness", "acousticness", "instrumentalness", "liveness", "valence", "tempo", "jack".

5 "EDA"

```
#| label: fig-temp
#| fig-cap: Tempo Variable
#| echo: false
#| warning: false


df %>%
  filter(album_release_date_precision == "day") %>%
  mutate(album_release_date = ymd(album_release_date)) %>%
  ggplot(aes(y = tempo, x = album_release_date, color = jack)) +
    geom_point(alpha = 0.3) +
  geom_smooth(method = lm, se = FALSE) +
  theme_minimal() +
  scale_color_manual(values = c("#808080","#FFB6C1")) +
  labs(color = "Has 'y' in jack?") +
  facet_wrap(~ artist_name, ncol = 3)
```
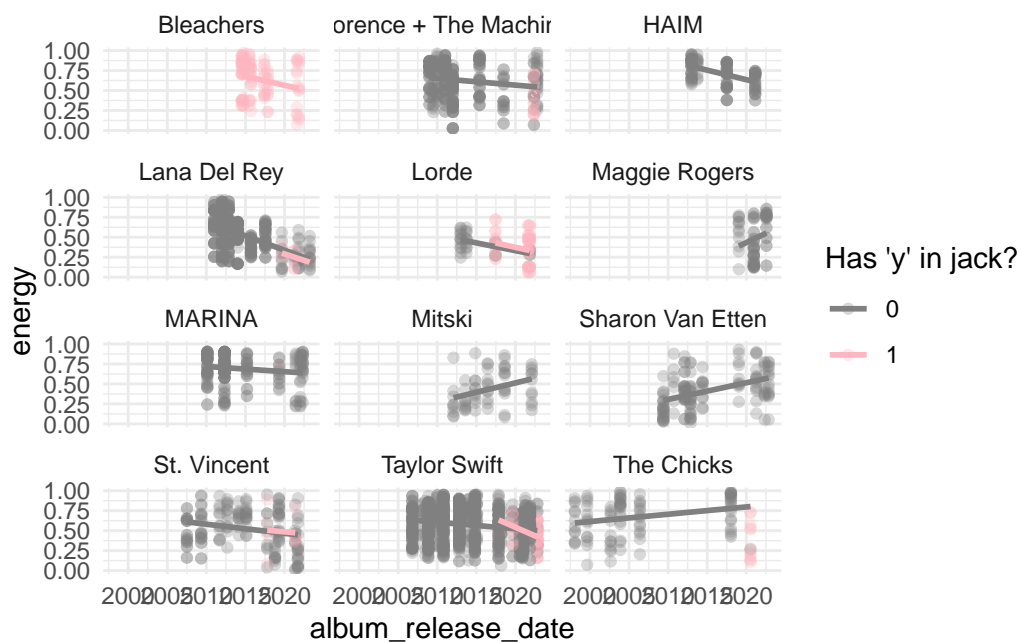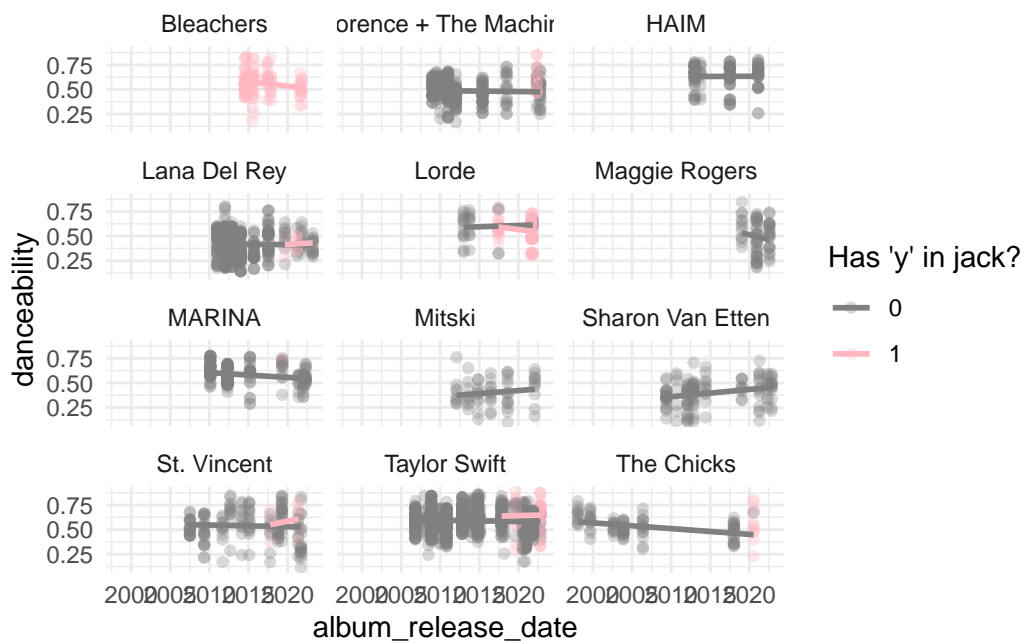
Figure 1: Energy Variable

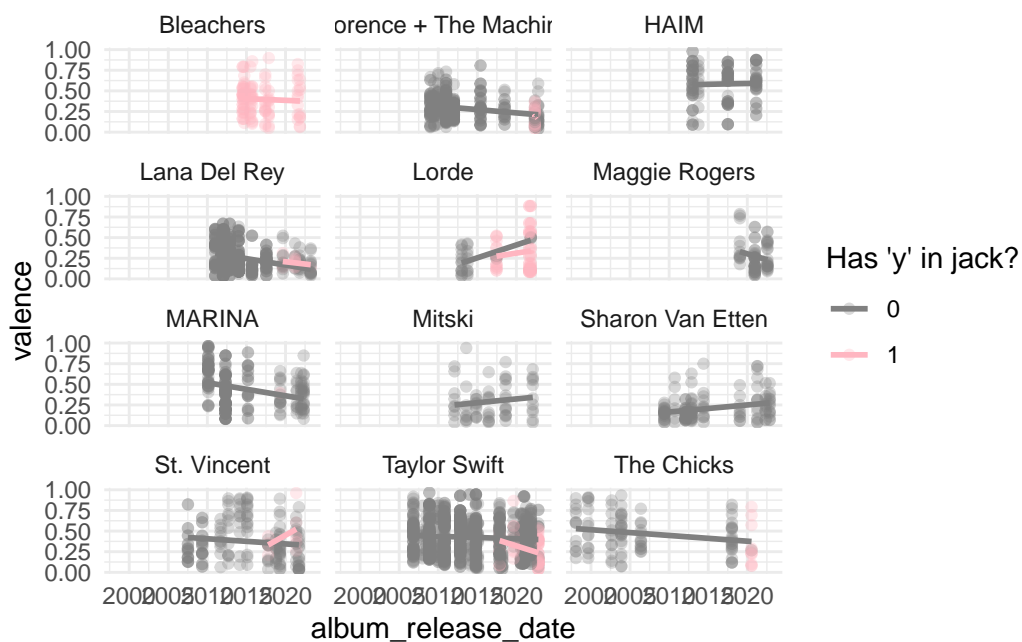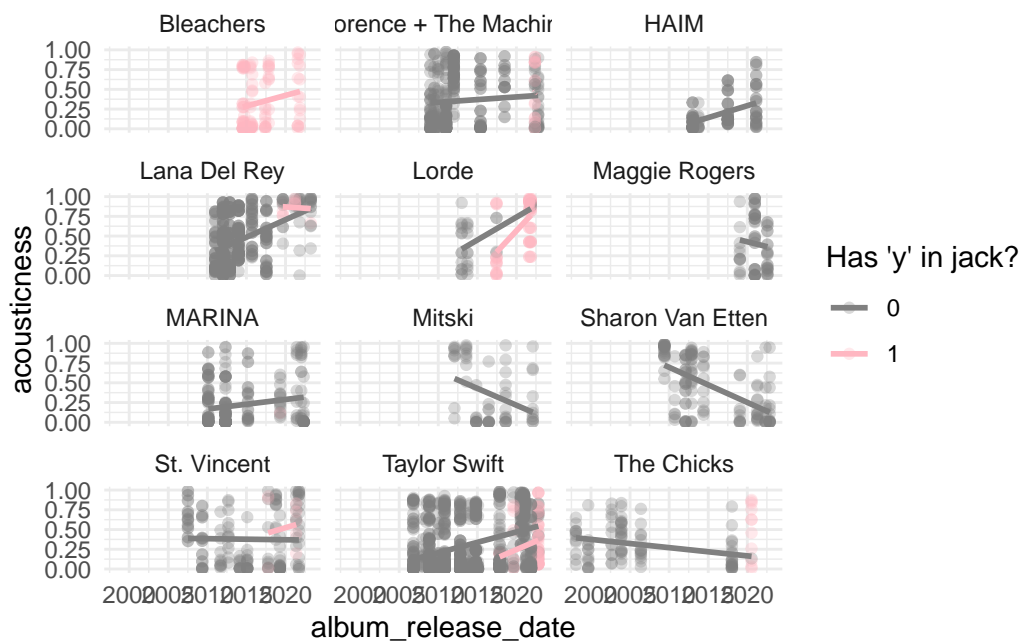Figure 2: Danceability Variable

Figure 3: Valence Variable
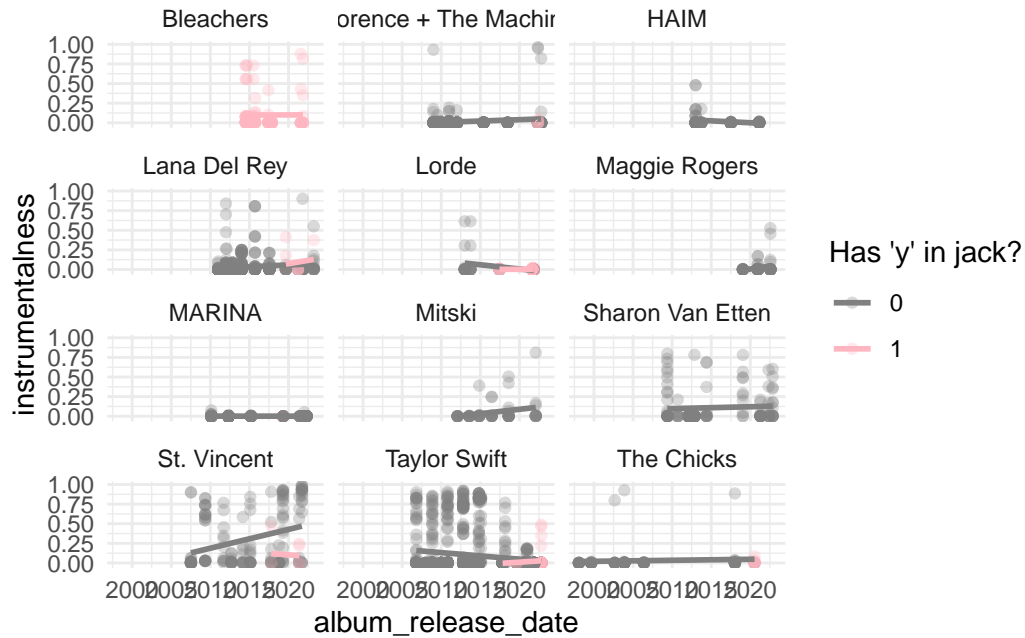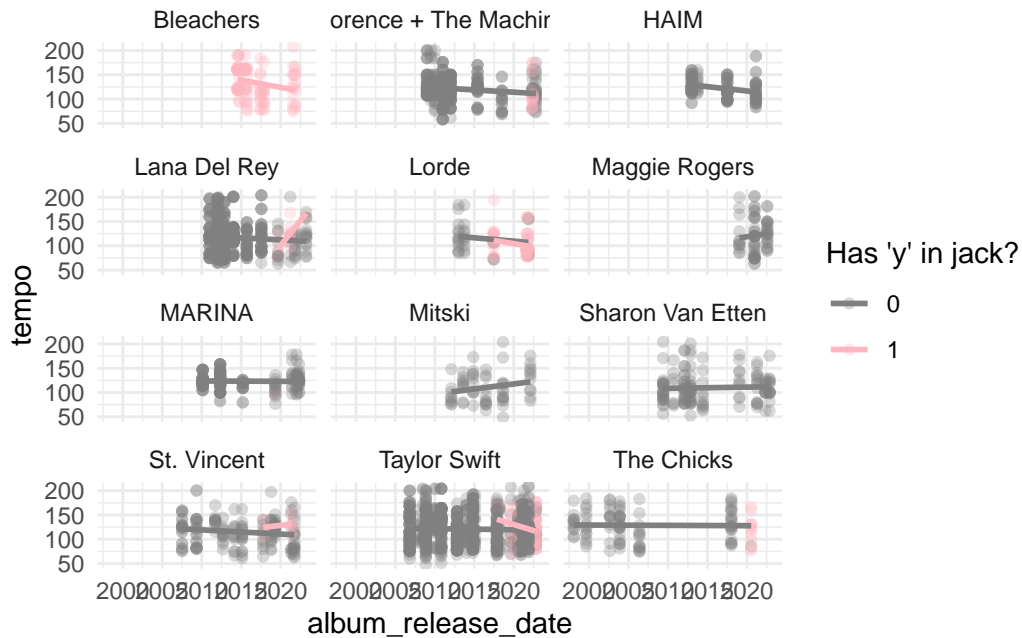

Figure 4: Acousticness Variable

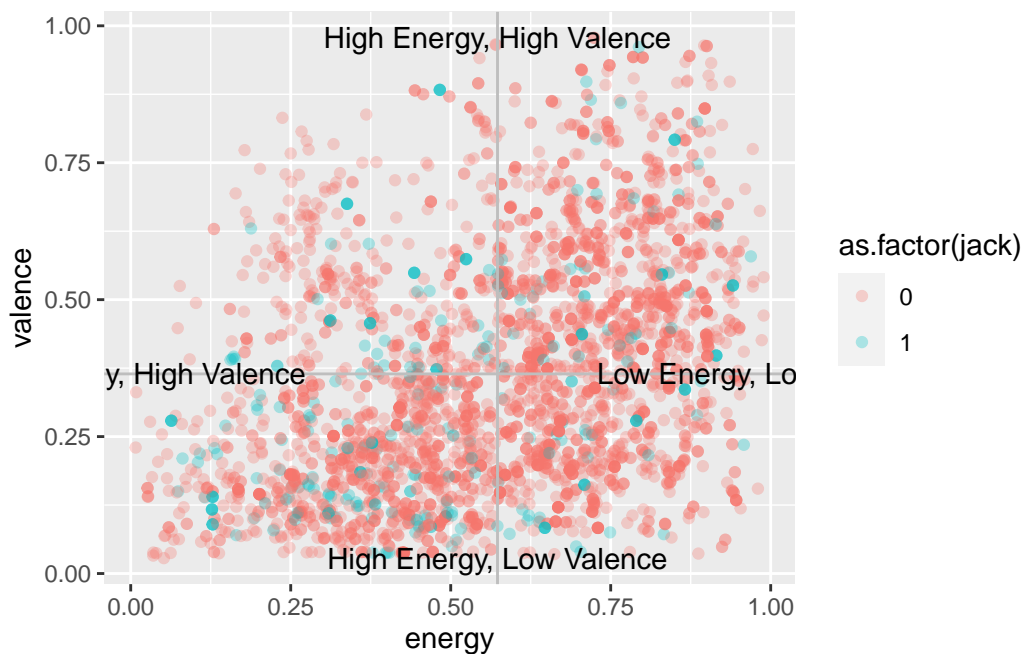Figure 5: Instrumentalness Variable

```
`geom_smooth()` using formula 'y ~ x'
```

I'm gonna fix the axis labels later

```
#| label: fig-quadvalence1
#| fig-cap: Valence and Energy Quadrant
#| echo: false
#| warning: false
```

```
ggplot(data = df, aes(x = energy, y = valence)) +
  geom_point(aes(color = as.factor(jack)), alpha = 0.3) +
  geom_vline(xintercept = mean(df$energy), color = "gray") +
  geom_hline(yintercept = mean(df$valence), color = "gray") +
  annotate("text", x = mean(df$energy), y = max(df$valence),
           label = "High Energy, High Valence") +
  annotate("text", x = mean(df$energy), y = min(df$valence),
           label = "High Energy, Low Valence") +
  annotate("text", x = min(df$energy), y = mean(df$valence),
           label = "Low Energy, High Valence") +
  annotate("text", x = max(df$energy), y = mean(df$valence),
           label = "Low Energy, Low Valence")
```
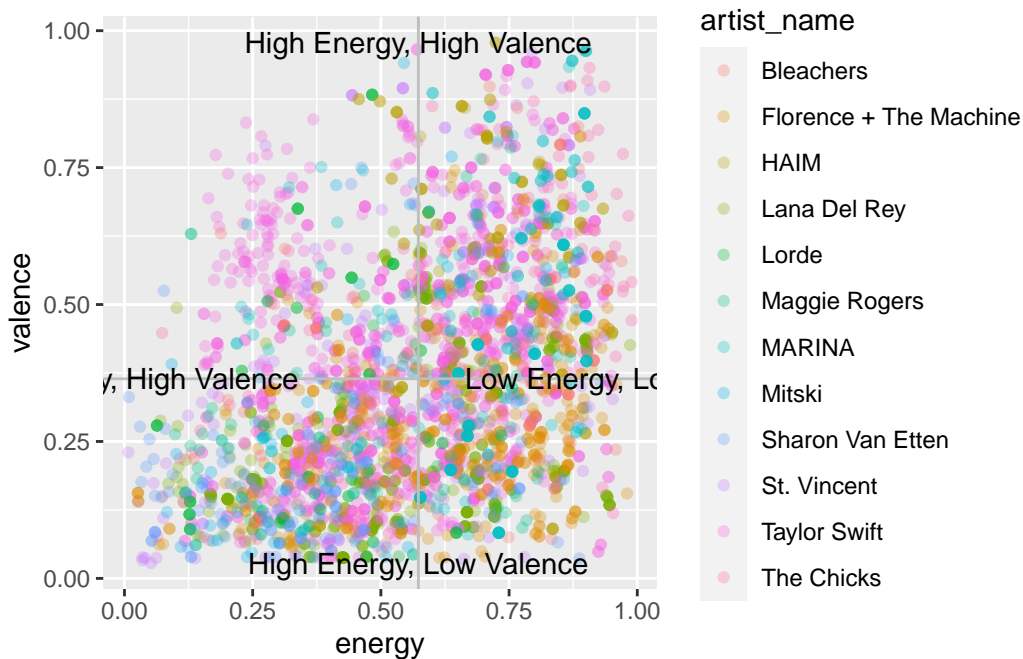
```
#| label: fig-quadvalence2
#| fig-cap: Valence and Energy Quadrant
#| echo: false
#| warning: false



ggplot(data = df, aes(x = energy, y = valence)) +
  geom_point(aes(color = artist_name), alpha = 0.3) +
  geom_vline(xintercept = mean(df$energy), color = "gray") +
  geom_hline(yintercept = mean(df$valence), color = "gray") +
  annotate("text", x = mean(df$energy), y = max(df$valence),
           label = "High Energy, High Valence") +
  annotate("text", x = mean(df$energy), y = min(df$valence),
           label = "High Energy, Low Valence") +
  annotate("text", x = min(df$energy), y = mean(df$valence),
           label = "Low Energy, High Valence") +
  annotate("text", x = max(df$energy), y = mean(df$valence),
           label = "Low Energy, Low Valence")
```
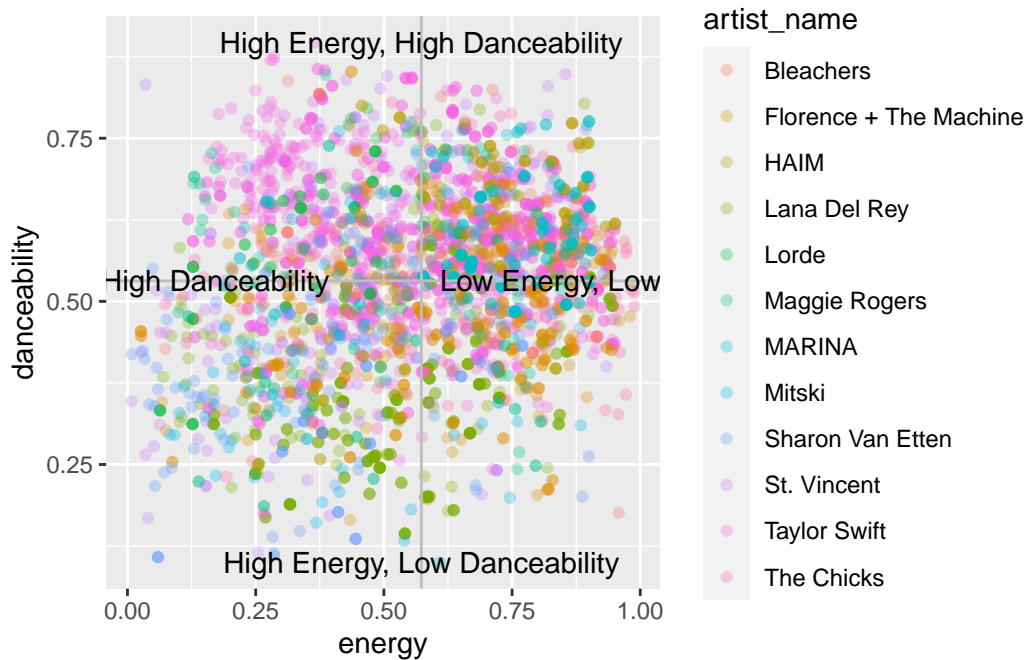

Valence and Energy Quadrant

```
#| label: fig-quaddanceability
#| fig-cap: Danceability and Energy Quadrant
#| echo: false
#| warning: false

ggplot(data = df, aes(x = energy, y = danceability)) +
  geom_point(aes(color = artist_name), alpha = 0.3) +
  geom_vline(xintercept = mean(df$energy), color = "gray") +
  geom_hline(yintercept = mean(df$danceability), color = "gray") +
  annotate("text", x = mean(df$energy), y = max(df$danceability),
           label = "High Energy, High Danceability") +
  annotate("text", x = mean(df$energy), y = min(df$danceability),
           label = "High Energy, Low Danceability") +
  annotate("text", x = min(df$energy), y = mean(df$danceability),
           label = "Low Energy, High Danceability") +
  annotate("text", x = max(df$energy), y = mean(df$danceability),
           label = "Low Energy, Low Danceability")
```



```
#| label: fig-quaddanceability2
#| fig-cap: Danceability and Energy Quadrant
#| echo: false
```

```
#| warning: false


ggplot(data = df, aes(x = energy, y = danceability)) +
  geom_point(aes(color = as.factor(jack)), alpha = 0.3) +
  geom_vline(xintercept = mean(df$energy), color = "gray") +
  geom_hline(yintercept = mean(df$danceability), color = "gray") +
  annotate("text", x = mean(df$energy), y = max(df$danceability),
           label = "High Energy, High Danceability") +
  annotate("text", x = mean(df$energy), y = min(df$danceability),
           label = "High Energy, Low Danceability") +
  annotate("text", x = min(df$energy), y = mean(df$danceability),
           label = "Low Energy, High Danceability") +
  annotate("text", x = max(df$energy), y = mean(df$danceability),
           label = "Low Energy, Low Danceability")
```
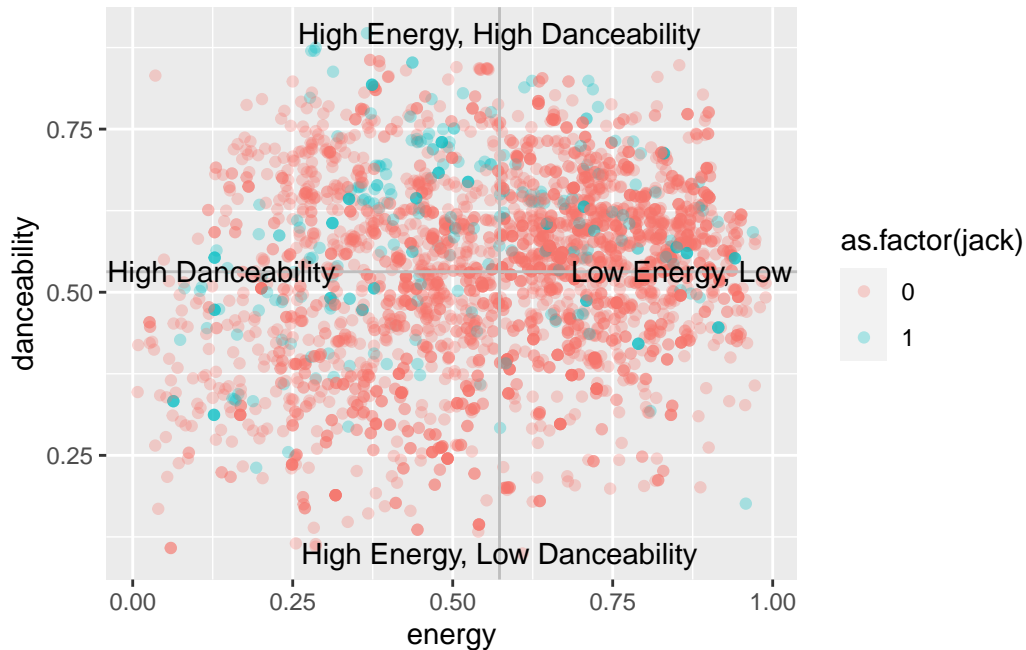

Sure. These are fun quadrants.But it doesn't solve the Jack-iness underlying each of those points, which are scattered pretty broadly.

Let's throw a logistic regression at the problem.

6 Model

Since we're trying to figure out whether these variables can tell us anything about what Jack Antonoff's sound is all about, a good way to figure that out is by seeing if any of these variables can help us predict what songs are produced by Jack. Logistic regressions like questions with binary outcomes.

[add the formula]

This model estimates the probability P of the outcome variable Y using predictor variables Xc. In other words, it estimates the probability that a song was produced by Jack Anotnoff using the Spotfiy variables. Which gives us information about the "research" "question": are these characteristics related to Jack's music?

The chart below gives us an idea of the mean value of each of these features for Jack and non-Jack song. A logistic regression will give us a better idea of how to weigh each of these features.

```
#|: echo: false
#|: warning: false

data_reduced <- read.csv(here::here("inputs/data/data_reduced.csv"))

data_reduced %>%
  group_by(jack) %>%
  summarise(across(-c(tempo, loudness), mean)) %>%
  pivot_longer(cols = -jack) %>%
  ggplot(aes(value, name, fill = factor(jack))) +
  geom_col(alpha = 0.8, position = "dodge") +
  labs(x = "Mean Value", y = "Feature")
```
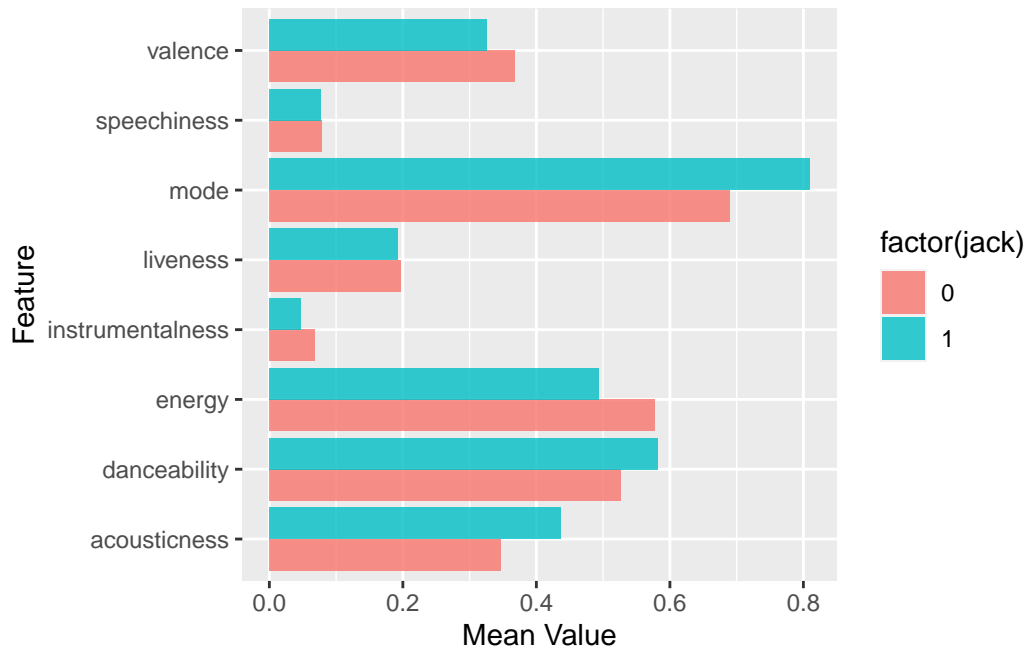
One issue is that in our dataset, there are a lot more songs that aren't produced by Jack than songs that are:

```
#|: echo: false
#|: warning: false

data_reduced %>%
  count(jack) %>%
  knitr::kable(col.names = c("Is it Jack?", "Number of Songs"))
```

Is it Jack?	Number of Songs
0	3110
1	294

So, we have to balance the classes.

Following Julia Silge's tutorial [TK], I used the tidymodels metapackage to split the training and testing data and make cross-validation samples. The `workflow_set()` function lets us hold different sets of feature engineering together and use tuning to evaluate them all at once. I need to evaluate two elements: a basic model, and a model that was downsampled using the `themis` package.

7 Results

0.743 accuracy, 0.841 roc_auc. Interesting. Check it out:

```
#jackmatrix


data_reduced |>
  pivot_longer(cols = c(danceability, energy, loudness, mode,
                speechiness, acousticness, instrumentalness, liveness, valence,
                tempo),
                names_to = "name",
                values_to = "value"
  ) |>
  ggplot(aes(x = value, color = jack)) +
  stat_ecdf() +
  facet_wrap(vars(name),
            nrow = 3,
            ncol = 4,
            scales = "free") +
  theme_minimal() +
  labs(
    x = "Value",
    y = "Proportion"
  )
```
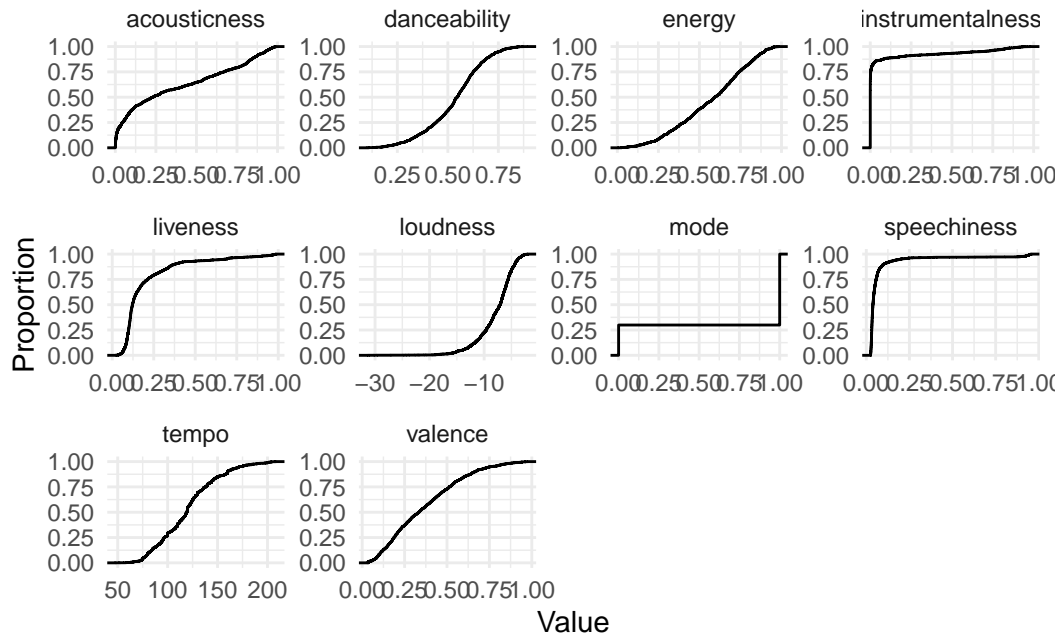
And here are the different variables' importance on determining whether a song was had any Antonoff involvement.

Variable	Importance	Sign
danceability	4.3200694	POS
instrumentalness	2.5899482	NEG
valence	2.0246212	NEG
speechiness	1.7422846	NEG
energy	1.3839382	POS
liveness	1.0386332	POS
mode	0.5269857	POS
acousticness	0.3696266	NEG
loudness	0.3531411	NEG
tempo	0.0072614	POS
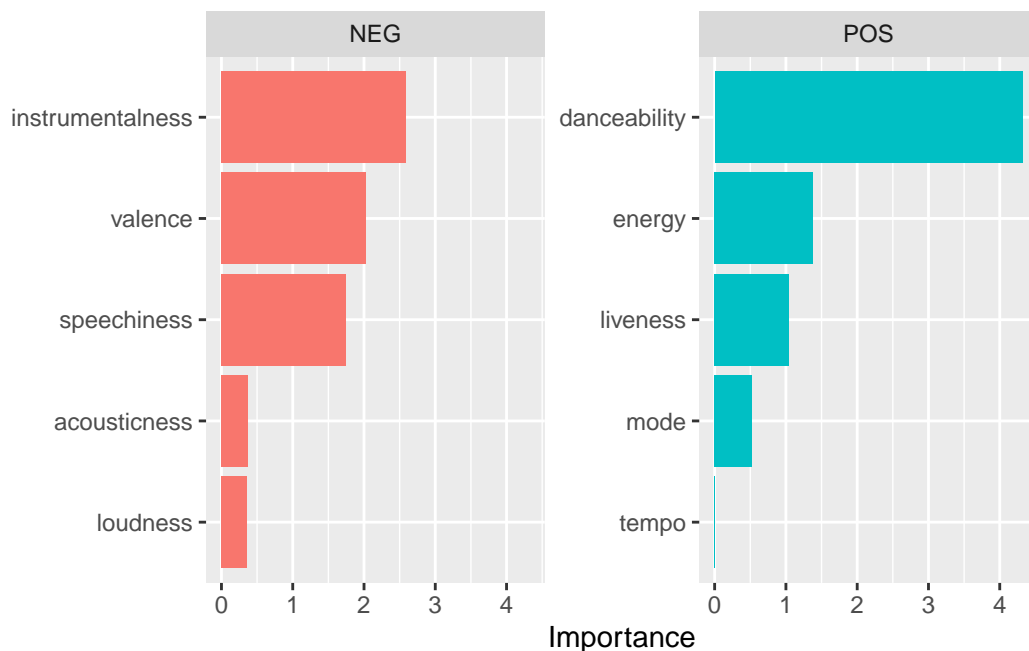
Danceability, with a score of 4.32, had the highest level of importance in determining whether or not a song was produced or co-produced by Jack Antonoff.

8 Discussion

8.1 Danceability Predicts Jack-iness

Using the Spotify Audio Features, you can't really gauge what Jack Antonoff's music sound like. You get a rough idea, but you don't get the same granularity that you would by simply DM'ing Caleb Gamman on Twitter and asking him to describe it you. [add gamman quote]

According to this model, though, danceability is the most important feature in predicting whether or not a song was produced by Jack Antonoff. This is so hilarious. This was a feature that was invented in a lab 15 years ago. This number refers to the time a bunch of undergraduates were paid (hopefully?) to listen to a bunch of songs and tag whether they were danceable or not. It is pretty much based on intuition. It captures this je-ne-sais-quoi essence of music that connects with listeners on a visceral level. Even if it doesn't make YOUR body want to move, you could imagine it making SOMEBODY'S body want to move.

Artist Name	Danceability	Track
Taylor Swift	0.897	I Think He Knows
Taylor Swift	0.875	Vigilante Shit

Artist Name	Danceability	Track
Florence + The Machine	0.852	Heaven Is Here
Bleachers	0.838	I Wanna Get Better - Demo Version
Taylor Swift	0.824	Cornelia Street
Bleachers	0.824	Hate That You Know Me - MTV Unplugged
Bleachers	0.818	Wake Me
Bleachers	0.814	Shadow
Taylor Swift	0.811	Paper Rings
The Chicks	0.805	Julianna Calm Down
Lorde	0.796	Sober
The Chicks	0.784	Texas Man
Bleachers	0.776	Hate That You Know Me
St. Vincent	0.774	Daddy's Home
Bleachers	0.769	All My Heroes
Taylor Swift	0.766	Look What You Made Me Do
Taylor Swift	0.751	Question...?
Taylor Swift	0.743	Lavender Haze
Taylor Swift	0.739	False God
Bleachers	0.732	Goodbye
Florence + The Machine	0.731	King
Bleachers	0.731	I Wanna Get Better
Lorde	0.730	Solar Power
MARINA	0.724	Karma
Taylor Swift	0.719	Dress
St. Vincent	0.713	Down
The Chicks	0.710	Sleep at Night
Taylor Swift	0.700	Bejeweled
Taylor Swift	0.700	Paris
Taylor Swift	0.696	You're On Your Own, Kid
St. Vincent	0.696	Masseduction
Taylor Swift	0.695	London Boy
Lorde	0.683	Mood Ring
Taylor Swift	0.681	Dear Reader
The Chicks	0.681	Gaslighter
Taylor Swift	0.675	Glitch
Lorde	0.669	Secrets from a Girl (Who's Seen it All)
Bleachers	0.669	Stop Making This Hurt
Bleachers	0.665	I Wanna Get Better - Vince Clarke Remix
Lorde	0.664	The Louvre
Taylor Swift	0.662	Mastermind

Artist Name	Danceability	Track
Taylor Swift	0.658	Maroon
Bleachers	0.655	Take Me Away
Bleachers	0.653	Goodmorning - MTV Unplugged
Bleachers	0.653	I Miss Those Days
Lorde	0.644	California
Bleachers	0.644	Wild Heart
Taylor Swift	0.643	Midnight Rain
Lorde	0.643	Dominoes
St. Vincent	0.641	The Laughing Man
Florence + The Machine	0.640	Choreomania
Taylor Swift	0.639	Anti-Hero
Florence + The Machine	0.634	The Bomb
Bleachers	0.631	Reckless Love
Bleachers	0.630	Don't Take The Money
Bleachers	0.629	Goodmorning
St. Vincent	0.625	Sugarboy
Bleachers	0.623	Everybody Lost Somebody - MTV Unplugged
Lorde	0.622	Perfect Places
Bleachers	0.621	Nothing Is U
Bleachers	0.618	You're Still a Mystery
Bleachers	0.615	Chinatown (feat. Bruce Springsteen) - Recorded at Electric Lady Studio
Bleachers	0.611	I'm Ready To Move On/Mickey Mantle Reprise
Florence + The Machine	0.610	Girls Against God
Bleachers	0.608	Big Life - Recorded at Electric Lady Studio
Bleachers	0.606	Let's Get Married
Bleachers	0.606	Take Me Away (feat. Grimes)
Lorde	0.605	The Path
Taylor Swift	0.598	Call It What You Want
Bleachers	0.598	All My Heroes - MTV Unplugged
Lorde	0.594	Green Light
Bleachers	0.592	Wild Heart - Demo Version
Lorde	0.587	Liability
Bleachers	0.584	Nothing Is U - MTV Unplugged
Bleachers	0.579	Don't Go Dark - Recorded at Electric Lady Studio
Bleachers	0.579	I Wanna Get Better - RZA Mix
Bleachers	0.573	Reckless Love - Demo Version

Artist Name	Danceability	Track
Florence + The Machine	0.571	Cassandra
Bleachers	0.568	Secret Life (feat. Lana Del Rey)
Taylor Swift	0.567	This Is Why We Can't Have Nice Things
Florence + The Machine	0.565	Dream Girl Evil
The Chicks	0.564	Everybody Loves You
Bleachers	0.563	I Miss Those Days - MTV Unplugged
Taylor Swift	0.562	Getaway Car
Lorde	0.562	Supercut
Bleachers	0.560	Like a River Runs
Taylor Swift	0.557	Daylight
Bleachers	0.554	Dream Of Mickey Mantle
Lorde	0.553	The Man with the Axe
Lorde	0.553	Liability (Reprise)
Taylor Swift	0.552	Cruel Summer
Bleachers	0.543	45
Bleachers	0.543	Rollercoaster
Lorde	0.540	Sober II (Melodrama)
The Chicks	0.537	Young Man
Bleachers	0.534	Everybody Lost Somebody
Lana Del Rey	0.531	Mariners Apartment Complex
Bleachers	0.527	How Dare You Want More
St. Vincent	0.525	Somebody Like Me
Bleachers	0.525	Strange Behavior
Bleachers	0.512	Don't Take The Money (feat. Lorde) - MTV Unplugged
The Chicks	0.508	Set Me Free
St. Vincent	0.506	Happy Birthday, Johnny
Bleachers	0.506	Who I Want You to Love
Bleachers	0.503	Don't Go Dark
Lana Del Rey	0.502	For Free
Bleachers	0.502	Drive - Recorded at Electric Lady Studio
Florence + The Machine	0.500	Prayer Factory
Bleachers	0.498	How Dare You Want More - Recorded at Electric Lady Studio
Lana Del Rey	0.493	White Dress
Bleachers	0.492	Big Life
Lorde	0.491	Fallen Fruit
Florence + The Machine	0.490	Morning Elvis
The Chicks	0.488	My Best Friend's Weddings

Artist Name	Danceability	Track
Taylor Swift	0.487	Labyrinth
Lorde	0.487	Hard Feelings/Loveless
Bleachers	0.487	I'm Ready to Move On / Wild Heart Reprise (feat. Yoko Ono)
Bleachers	0.485	Shadow (feat. Carly Rae Jepsen) - MTV Unplugged
Bleachers	0.474	Strange Behavior - Recorded at Electric Lady Studio
Lorde	0.473	Stoned at the Nail Salon
Lorde	0.473	Oceanic Feeling
The Chicks	0.469	March March
Florence + The Machine	0.467	Free
Bleachers	0.466	Rollercoaster - MTV Unplugged
Bleachers	0.455	I'm Ready to Move On / Wild Heart Reprise
Lana Del Rey	0.445	Not All Who Wander Are Lost
Bleachers	0.440	45 - Recorded at Electric Lady Studio
Bleachers	0.436	Chinatown (feat. Bruce Springsteen)
Taylor Swift	0.435	Sweet Nothing
Bleachers	0.427	What'd I Do With All This Faith?
Bleachers	0.427	Let's Get Married - MTV Unplugged
Lana Del Rey	0.421	A&W
The Chicks	0.401	For Her
Bleachers	0.391	Foreign Girls
St. Vincent	0.390	Candy Darling
Bleachers	0.390	I Wanna Get Better - MTV Unplugged
Lana Del Rey	0.387	Tulsa Jesus Freak
St. Vincent	0.373	New York
Lana Del Rey	0.370	Breaking Up Slowly
Lana Del Rey	0.368	Venice Bitch
Lana Del Rey	0.359	Dance Till We Die
Taylor Swift	0.359	Lover
Bleachers	0.344	What'd I Do With All This Faith? (feat. St. Vincent) - Recorded at Electric Lady Studio
Lorde	0.333	Big Star
Bleachers	0.333	91
Bleachers	0.318	Rollercoaster - Live in Boston
Lorde	0.312	Leader of a New Regime
Lana Del Rey	0.294	Cinnamon Girl
Taylor Swift	0.292	The Archer
The Chicks	0.231	Hope It's Something Good
Bleachers	0.176	I Wanna Get Better - Live in Boston

8.1.1 What do I think:

- I Think He Knows by Taylor Swift:

  - pretty danceable tbh

- Vigilante Shit by Taylor Swift:

  - Not danceable. But it is trying to coerce you into some sort of body feeling.

- Heaven is Here by Florence and the Machine

  - DEFINITELY not danceable. But it is trying to get your body to do something. Kind of sounds like it's suggesting that you join a Macedonian phalanx or something. But no. Not danceable.
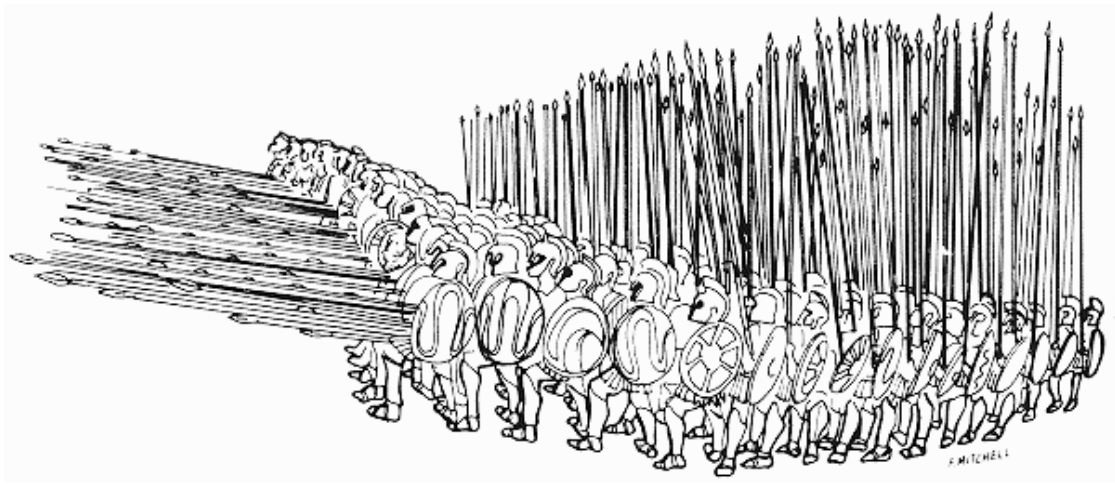
- I Wanna Get Better by Bleachers

  - I don't know if "danceability" is the word, but I'll give it "anthem"-y. Like, I can't picture a group of friends dancing to this in the kitchen, but I can picture someone listening to it in their car when it comes on the radio on their way to give a presentation or negotiate a promotion or something.

- Cornelia Street by Taylor Swift

  - Yeah I'm dancing. Or at least bobbing my head.

8.2 But Danceability isn't even about Dancing! I propose: Visceral-ness

I am thinking about what Gamman said in his video.

> I did use the phrase "visceral hatred". But i did so only to describe my visceral
> hatred. I can't stand the sound of it. I hate it. And the way in which I hate it is
> viscerally. But I didn't go beyond that because it doesn't go beyond that. Nothing
> against the guy, I like Taylor swift's stuff. It's just subjective taste.

It's just his subjective taste, but in one way Gamman was actually totally on the money
about this. The danceability variable, which is the strongest predictor of Jack-ness, seems to
actually be more about "visceral"-ness when you do a close-listen of the songs with the highest
danceability scores. The reason Jack is getting under people's skin could just be because his
music... gets underneath your skin. Right to the viscera.

8.3 Weaknesses and next steps

One fundamental weakness is cooked into the dataset. Since we're not in the studio, we can't
know the degree of Jack-iness in each individual song. Especially when he takes on the role of
a more backseat co-producer, as he did with Lana Del Rey's Did You Know There's a Tunnel
Under Ocean Blvd? [tk].

8.3.0.1 Could our model just be predicting Jack via Taylor Swift?

Boxplot of Danceability by Artist

29

Artists with Most Jack = 1 Values
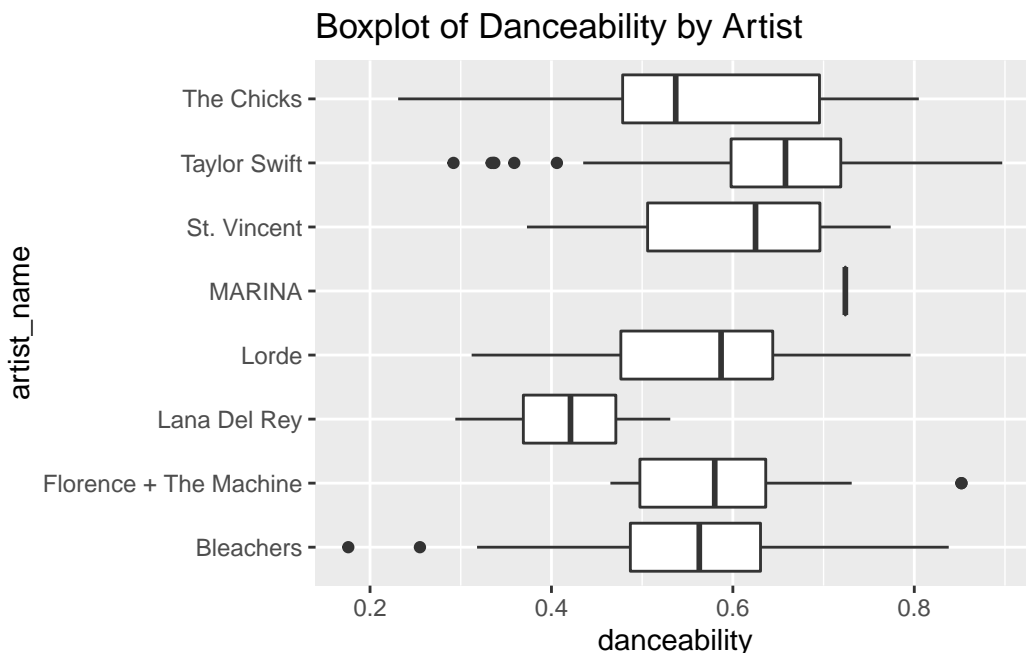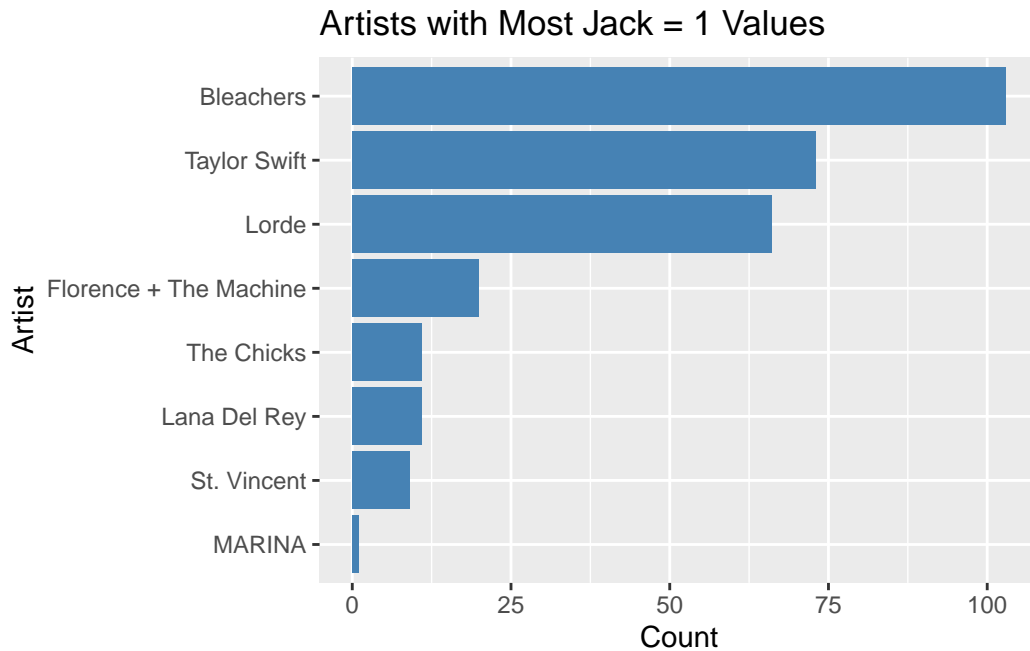
8.3.0.2 Next Steps

It only makes sense, of course, to hire a bunch of undergrads and get them to tell us whether or not a few thousand songs are visceral, and then build our own visceral-ness metric, and see if we get the same results – that danceability is just visceral-ness in a trench coat, and that visceral-ness is the characteristic that helps us best understand what makes Jack, Jack.

Appendix

.0.1 `get_artist_audio_features` documentation:

- **acousticness**: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

- **danceability**: Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

- **energy**: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

- **instrumentalness**: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

- **key**: The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C /D , 2 = D, and so on. If no key was detected, the value is -1.

- **liveness**: Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.

- **loudness:** The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

- **mode:** Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.

- **tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

- **valence:** A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

A Additional details

B References