Spotify Insights

How can data influence creativity?

Questions to consider

- Are there subsets of music that have measurably different attributes?
- Are there trends present in the way users create their playlists?
- If so, how different are they from the general population?
- How can musicians use the differences to influence creative decisions?

Hypotheses

- Songs that are grouped together for an express purpose have a definable set of attributes when compared to the population.
- We can use these differences to create a template from which to draw when creating music.

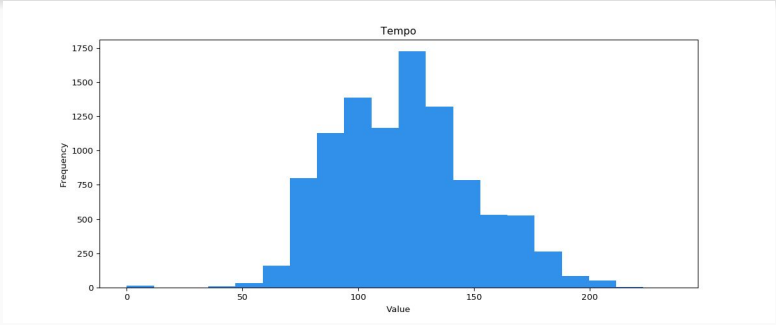
Targets

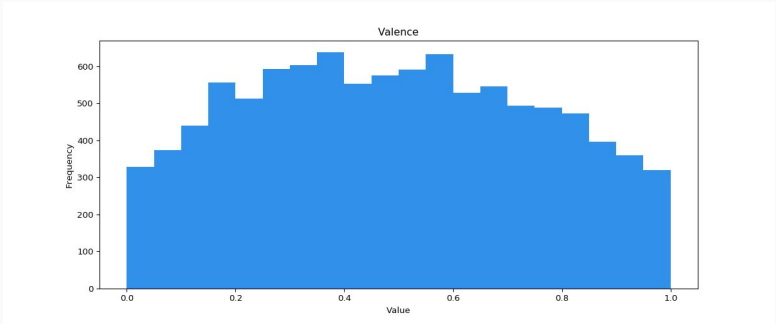
- Sample of the general population
- Songs taken from user-created playlists with “running” in the title
- Songs taken from user-created playlists with “study” in the title
- N ≈ 5,000 for each set
- Playlist songs were gathered by me using the API and a very useful library, Spotipy
- Population data sampled from a Kaggle user
- Metrics generated by Spotify for every song

Data in Music

- Objective
  - Tempo
  - Amplitude (volume)
  - Instrumentalness
- Subjective
  - Energy
  - Valence (mood)

[Images from spotify api]





Target Playlists: Running

Expectations

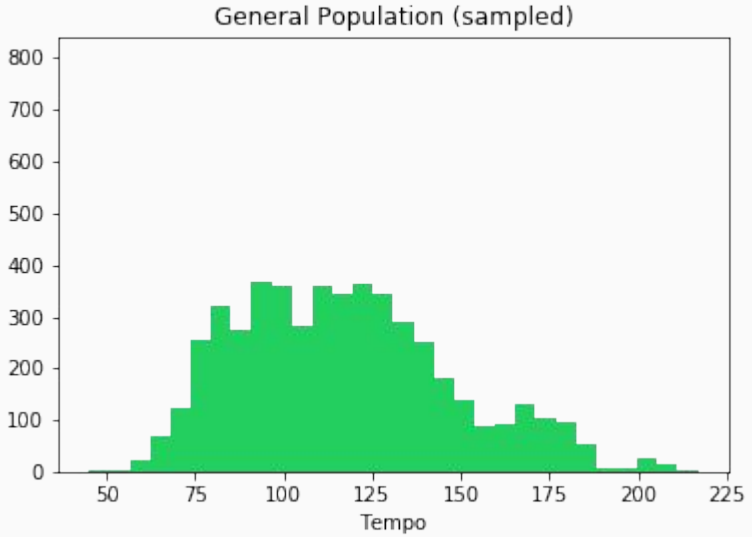
- Fast
- Upbeat / energetic
- Uplifting

Target Playlists: Running

Tempo

“...the speed or pace of a given piece.”

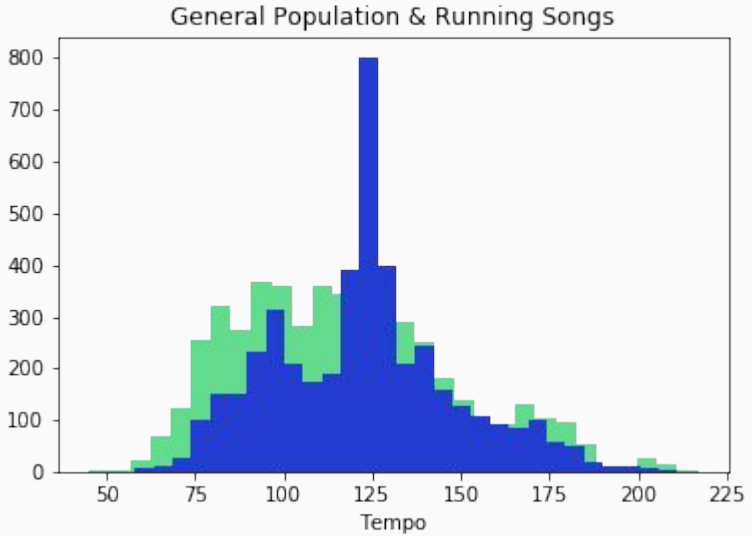
Measured in beats per minute (BPM)



Target Playlists: Running

Tempo

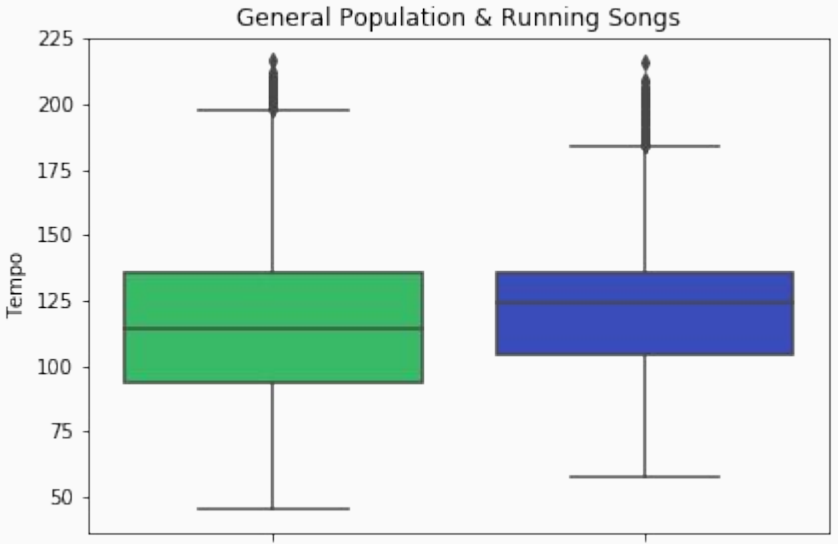
- 4 - 7 BPM higher



Target Playlists: Running

Tempo

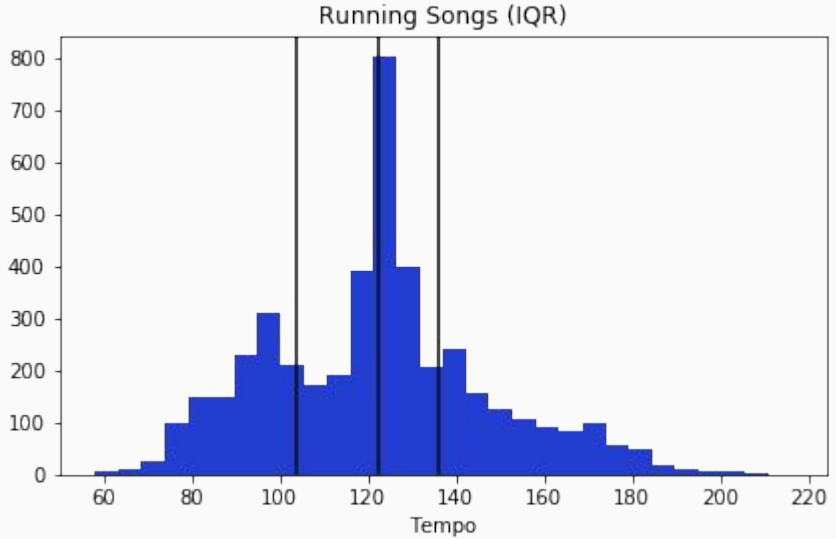
- 4 - 7 BPM higher
- Narrower range



Target Playlists: Running

Tempo

- 4 - 7 BPM higher
- Narrower range
- Emphasis around 122 BPM
- Examples
  - “How Deep Is Your Love” - Calvin Harris
  - “Animal I Have Become” - Three Days Grace



Musical Guidelines

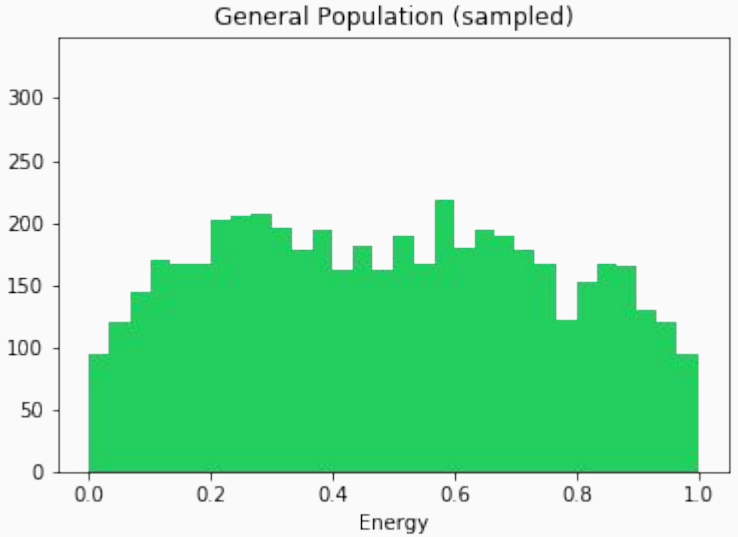
Running Playlists

- **Tempo around 122 BPM**

Target Playlists: Running

Energy

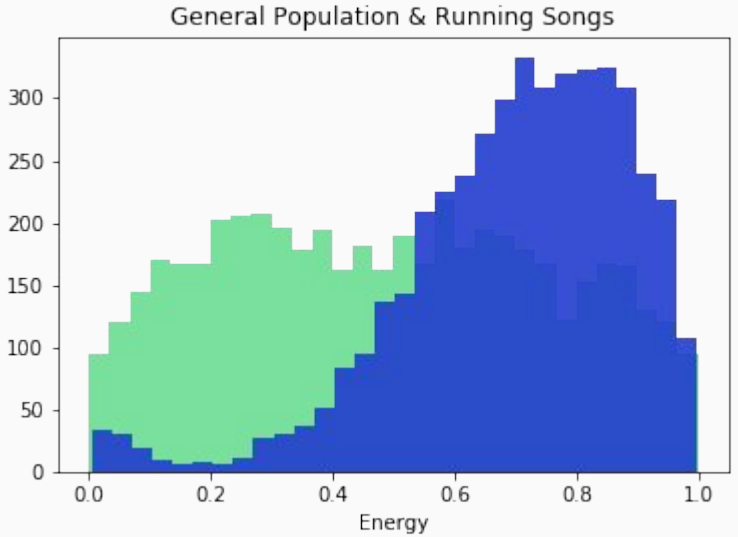
“...perceptual measure of intensity and activity.”



Target Playlists: Running

Energy

- Louder
- Busier
- More ‘energetic’
- Example
  - “Blinding Lights” - The Weeknd



Musical Guidelines

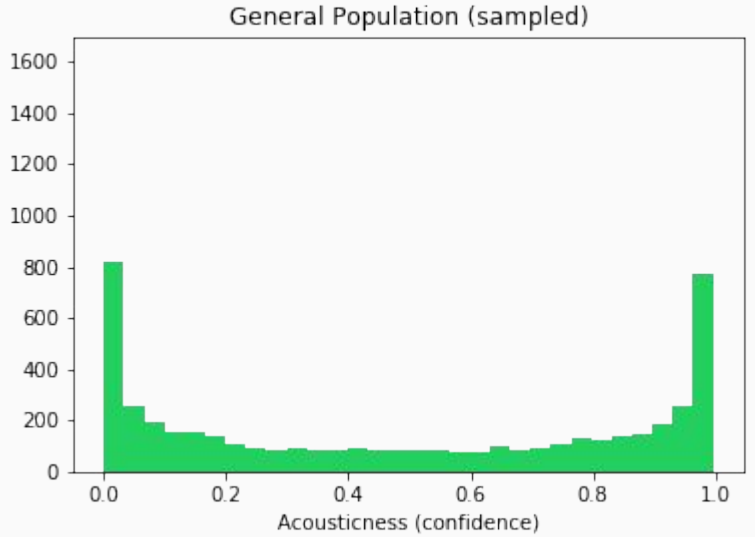
Running Playlists

- Tempo around 122 BPM
- **Higher energy level (louder / more dynamic)**

Target Playlists: Running

Acousticness

“A confidence measure from 0.0 to 1.0 of whether the track is acoustic.”

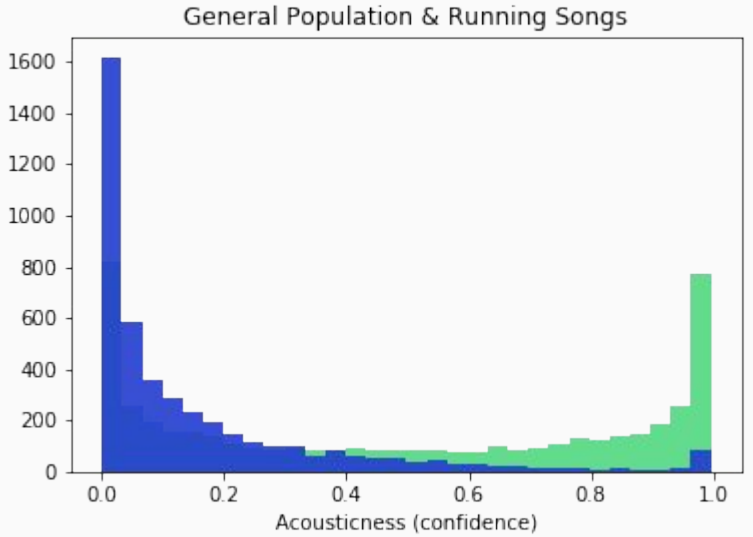


*note, distribution differs significantly from what is listed in the documentation

Target Playlists: Running

Acousticness

- Hardly any acoustic songs in running playlists



Musical Guidelines

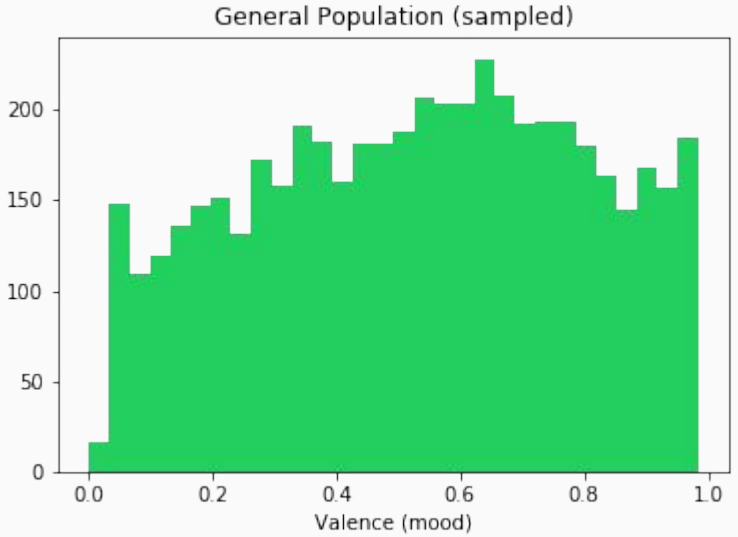
Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- **Not acoustic**

Target Playlists: Running

Valence

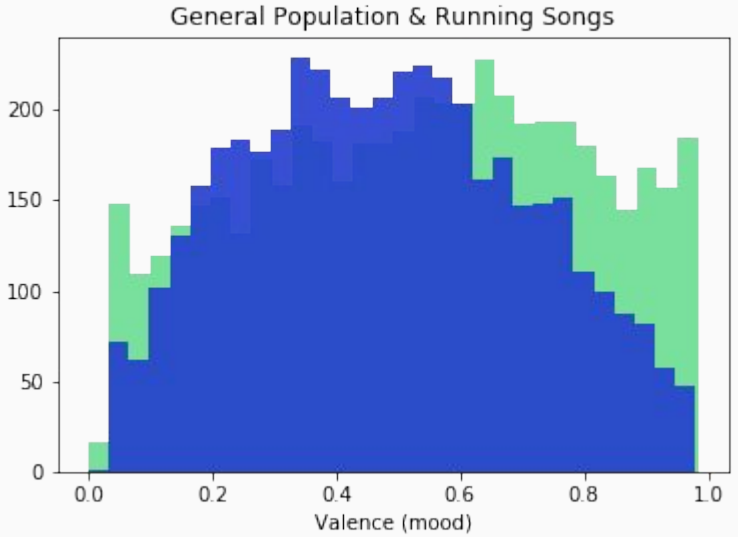
“...the musical positiveness conveyed by a track”



Target Playlists: Running

Valence

- Technically lower by statistical significance
- Small difference and effect size
- Example
  - “ROXANNE” - Arizona Zervas



Musical Guidelines

Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- **Less positive mood**

Target Playlists: Studying

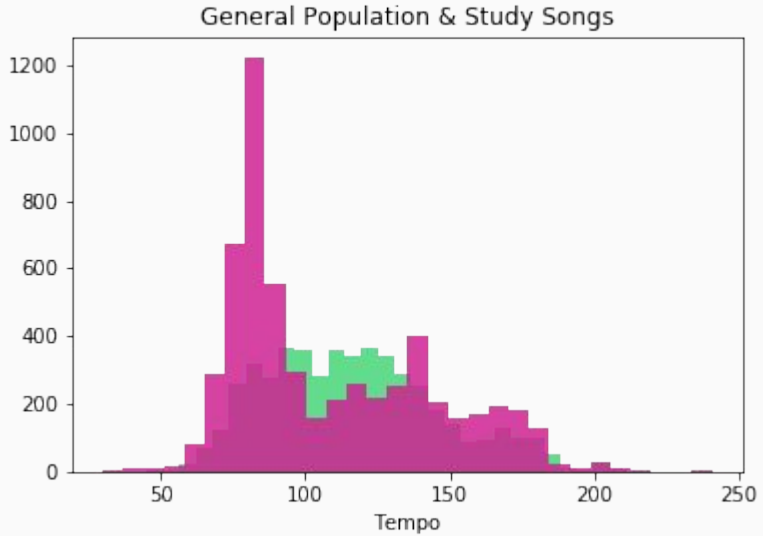
Expectations

- Slower
- Less distracting
- Neutral

Target Playlists: Studying

Tempo

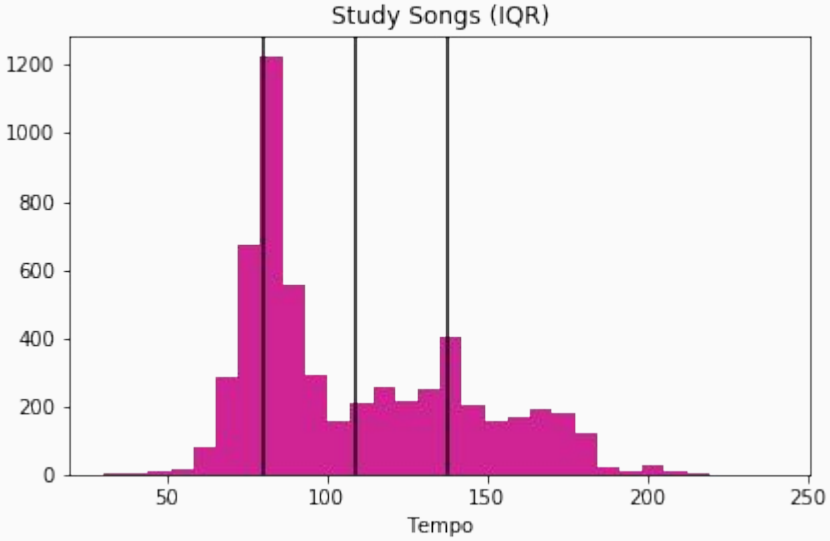
- 7 - 9 BPM lower



Target Playlists: Studying

Tempo

- 7 - 9 BPM lower
- Bottom of IQR better choice than the average
- Lines up with the mode
- 80 BPM
- Example
  - “How To Save a Life” - The Fray



Musical Guidelines

Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- Less positive mood

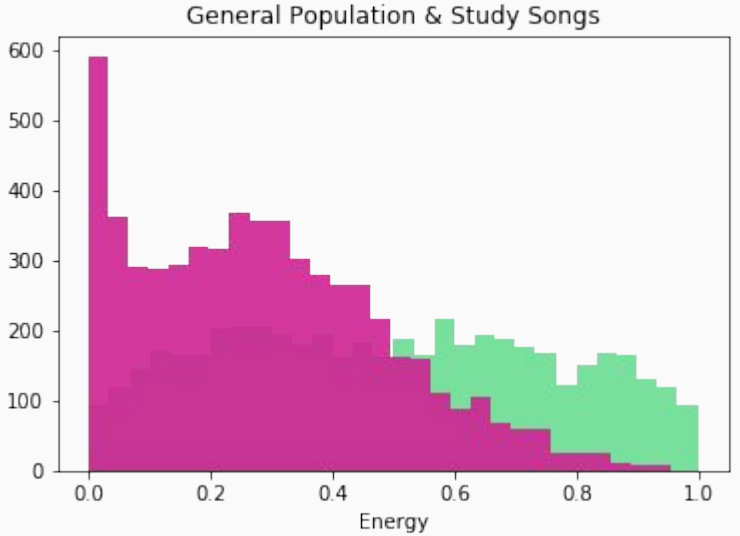
Study

- **Tempo around 80 BPM**

Target Playlists: Studying

Energy

- Much lower energy
- Softer, simpler songs to avoid distracting the listener
- Example
  - “All of Me” - John Legend



Musical Guidelines

Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- Less positive mood

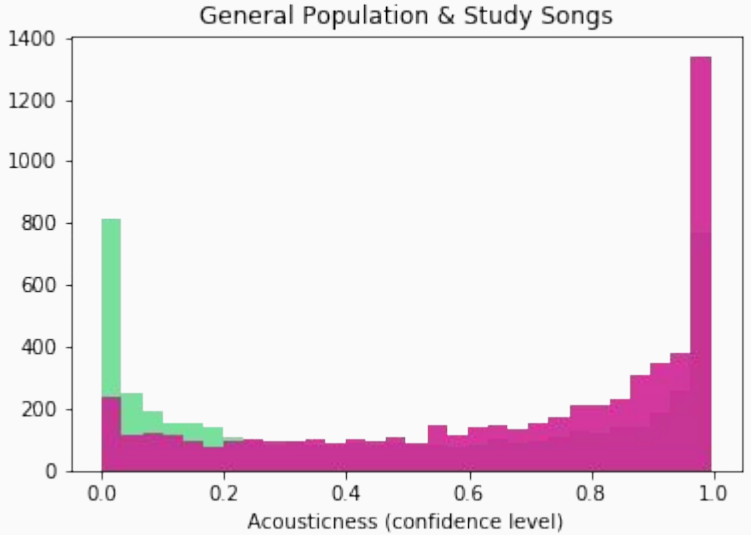
Study

- Tempo around 80 BPM
- **Low energy level (simple / less dynamic)**

Target Playlists: Studying

Acousticness

- One playlist was very large and contained only solo piano songs
- Maybe better. Would require further analysis



Musical Guidelines

Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- Less positive mood

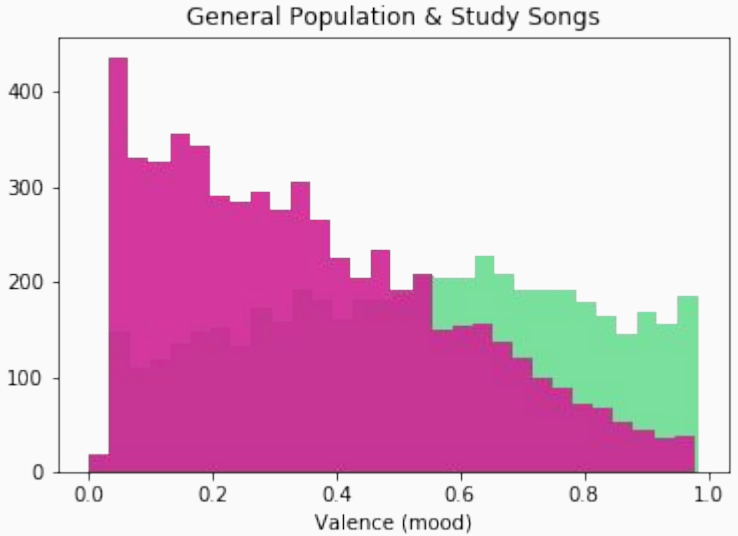
Study

- Tempo around 80 BPM
- Low energy level (simple / less dynamic)
- ***Maybe* acoustic**

Target Playlists: Studying

Valence (mood)

- Decidedly less positive
- Music as a tool, not an escape
- Example
  - “Falling” - Harry Styles
  - “Perfect” - Ed Sheeran



Musical Guidelines

Running Playlists

- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- Less positive mood

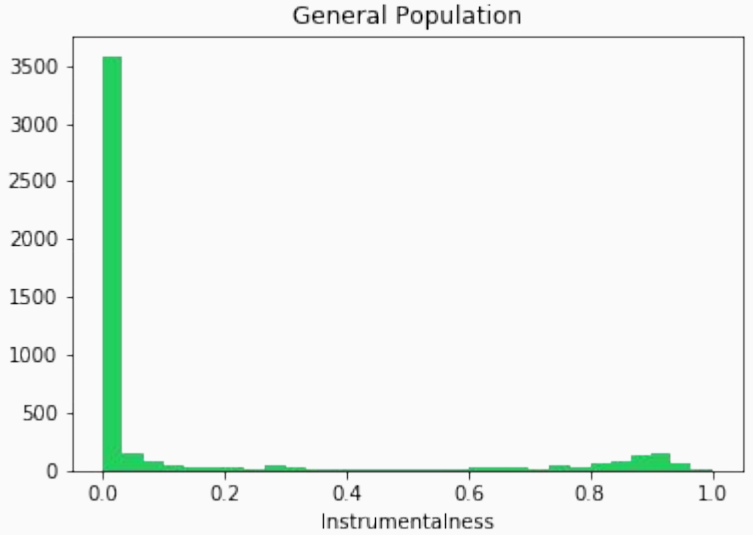
Study

- Tempo around 80 BPM
- Low energy level (simple / less dynamic)
- Maybe acoustic
- **Emphasis on less happy**

Target Playlists: Studying

Instrumentalness

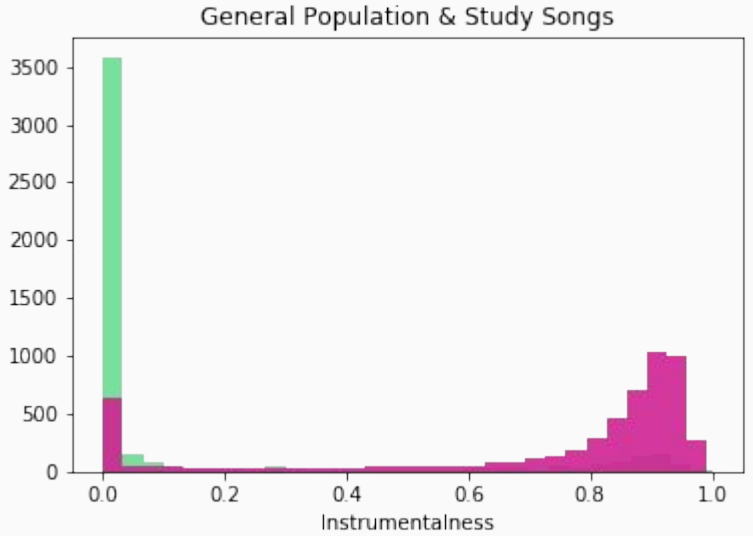
“The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content.”



Target Playlists: Studying

Instrumentalness

- Overwhelming amount of instrumental songs
- Statistically, this represented the largest deviation from the population



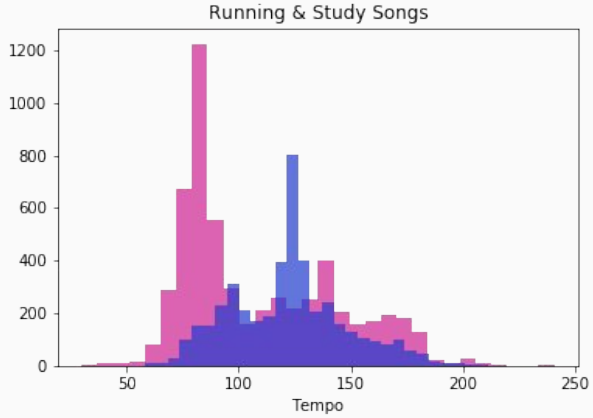
Musical Guidelines

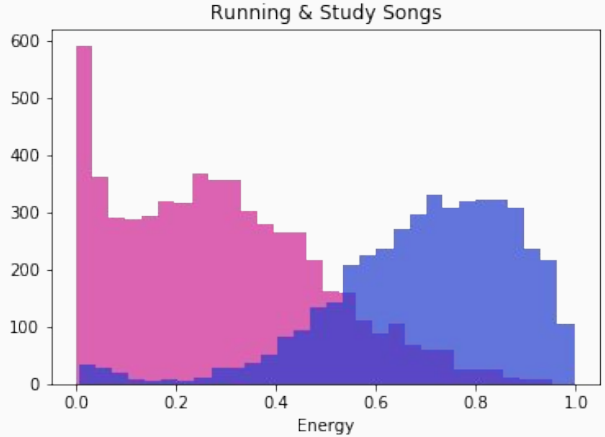
Running Playlists

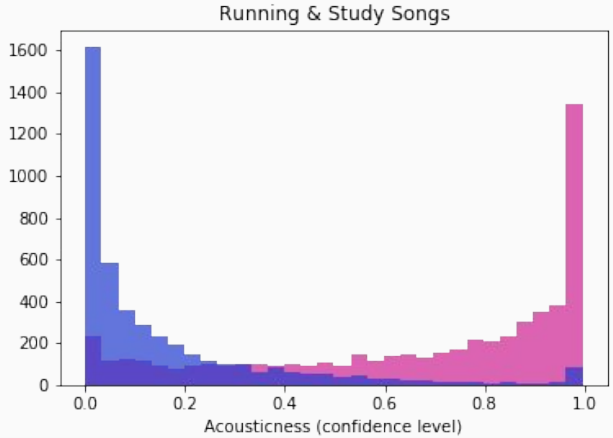
- Tempo around 122 BPM
- Higher energy level (louder / more dynamic)
- Not acoustic
- Less positive mood

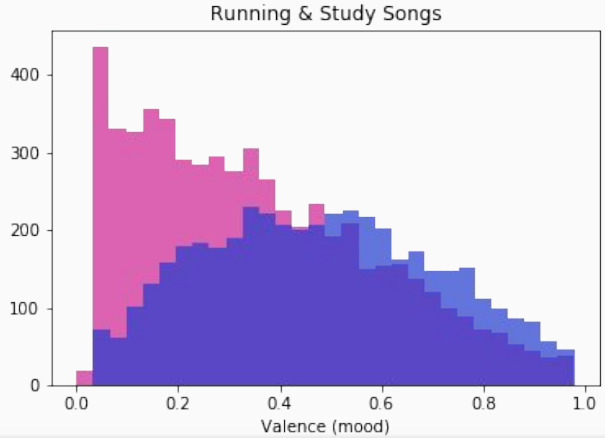
Study

- Tempo around 80 BPM
- Low energy level (simple / less dynamic)
- Maybe acoustic
- Emphasis on less happy
- **Minimal lyrical content**









= Running

= Study

Conclusions

- We are able to establish that certain types of playlists carry attributes that are significantly different than the general population.
- With objective measures like tempo, we can set forth a clear target.
- Looking at subjective measures, we can set some loose overall goals for what the music should feel like
- All of these are established in our final set of “Musical Guidelines”

Improvements

- Sample the population in a way that better represents it.
- Limit playlist size and draw from more playlists
- Instead of dropping duplicates, use them to weight the data

Additional Use Cases

- Artists can look through their discography and try to market old songs in a fresh way
- Labels can make their own running or study playlists
- Guide algorithmically generated playlists

Questions?

Thank You