

Top 50 Spotify Songs 2019

Gabrielle Isaak
Clara Brown
Grace Okamoto
Ian Miller



Outline

- Dataset
- Processing
- Analysis
 - Pivot Chart
 - Correlation
 - Scatter Plot
 - Word analysis
 - Histogram
- Conclusion
- Final remarks

Big Question

Was there any
relationship in the data
that linked which songs
ranked higher in the top
50?

Columns of the data set

- Track.Name
- Artist.Name
- Genre
- Beats.Per.Minute
- Energy
- Danceability
- Loudness..dB..
- Liveness
 - The higher the value, the more likely the song is a live recording.
- Valence.
 - The higher the value, the more positive mood for the song.
- Length.
- Acousticness..
- Speechiness.
 - The higher the value the more spoken word the song contains.
- Popularity

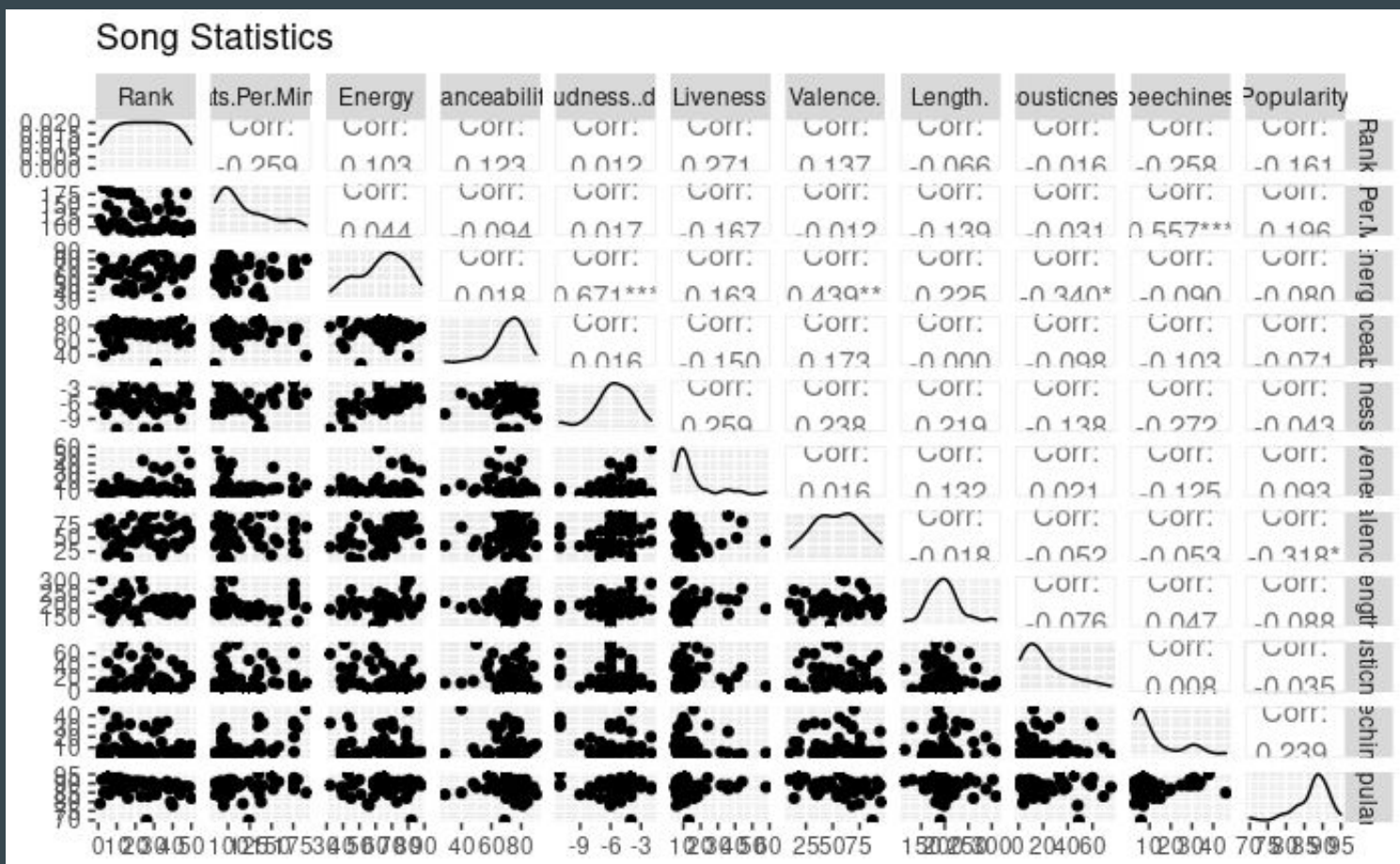
Pivot Chart

Row Labels	Count of Artist.Name
Ed Sheeran	4
Marshmello	2
Ariana Grande	2
The Chainsmokers	2
J Balvin	2
Shawn Mendes	2
Billie Eilish	2
Sech	2
Post Malone	2
Lil Nas X	2
Jonas Brothers	1
ROSALÍA	1
Drake	1
MEDUZA	1
Ali Gatie	1
Martin Garrix	1
Y2K	1
Tones and I	1
Daddy Yankee	1
Maluma	1

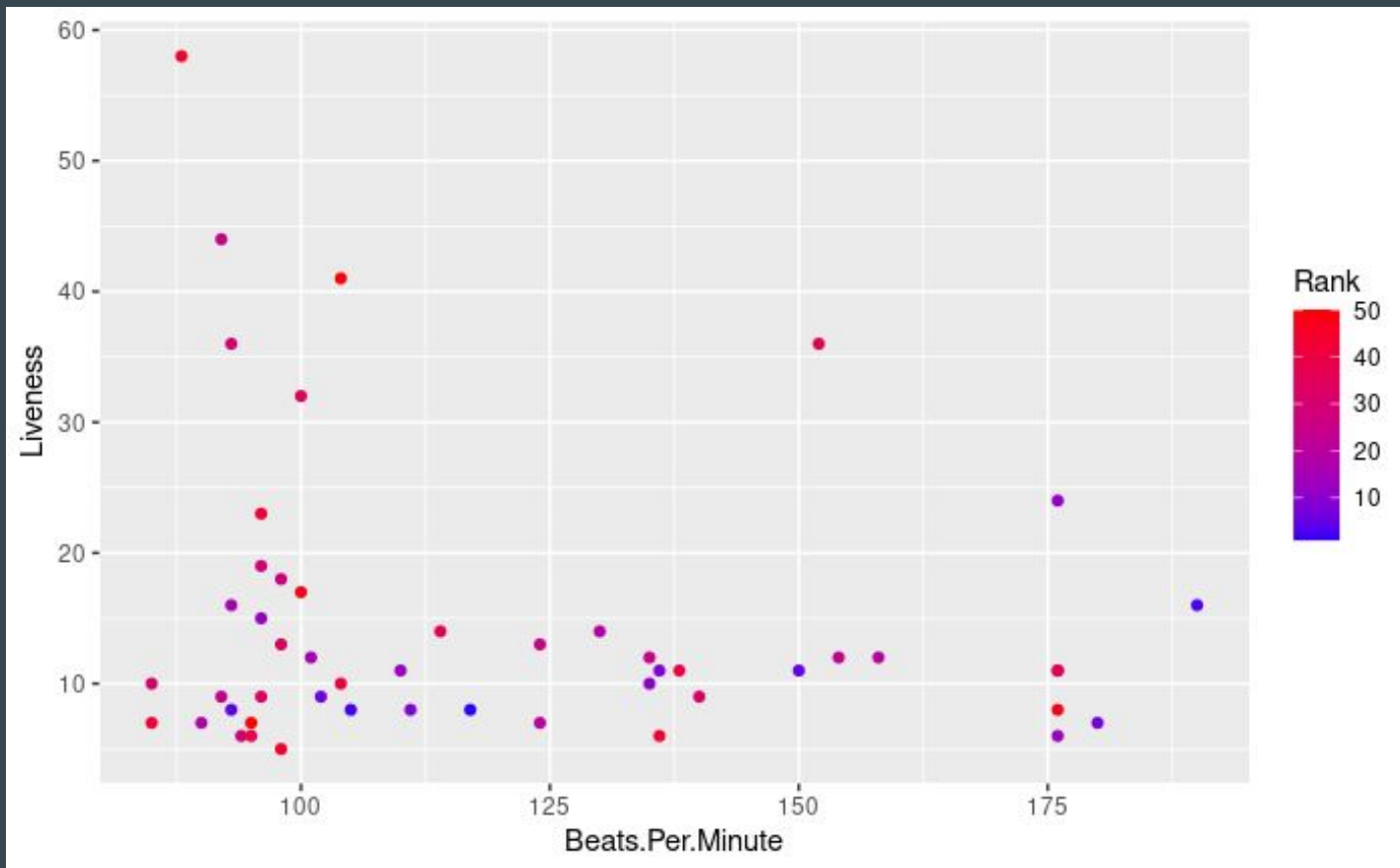
Row Labels	Count of Genre
dance pop	8
pop	7
latin	5
canadian hip hop	3
edm	3
reggaeton	2
panamanian pop	2
reggaeton flow	2
canadian pop	2
brostep	2
country rap	2
dfw rap	2
electropop	2
pop house	1
trap music	1
r&b en espanol	1
atl hip hop	1
boy band	1
big room	1
australian pop	1
escape room	1
Grand Total	50

Row Labels	Count of Genre2.0
pop	23
latin	5
rap	4
reggaeton	4
hip hop	4
edm	3
brostep	2
escape room	1
boy band	1
trap	1
big room	1
r&b en espanol	1
Grand Total	50

Correlation



Scatter plot



Spanish vs English

- 14 Songs Contained Spanish Lyrics
- 28% of Dataset

Did the Spanish and English Songs have a different formula or recipe for getting on the list?

Exploring the Genres

English Songs

- Pop
- Rap
- Trap
- Hip Hop
- EDM
- Dance
- Brostep
- Boyband

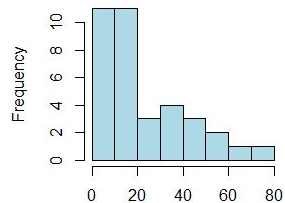
Spanish Songs

- Pop
- Reggaeton
- Latin

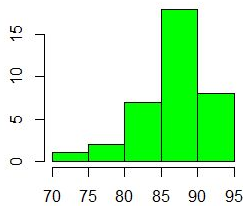
Looking at the Averages

Variable	English	Spanish
BPM	117.8	125.9
Energy	60.6	72.85
Danceability	69.9	75.2
Loudness	-6.2	-4.3
Liveness	14.8	14.2
Valence	50.8	64.3
Length	189.8	229.7
Acousticness	22.9	20.4
Speechiness	11.1	16
Popularity	87.2	88.35

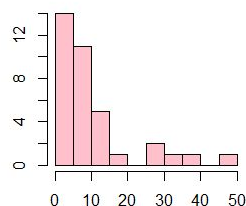
English Song Acousticness



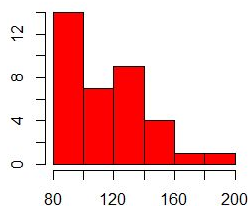
English Song Popularity



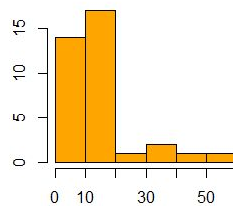
English Song Speechiness



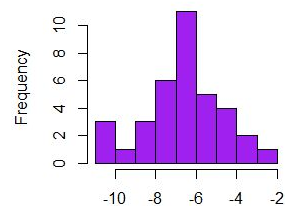
English Songs BPM



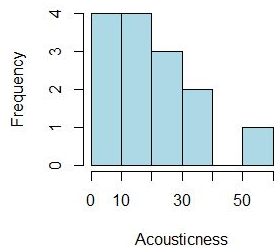
English Song Liveness



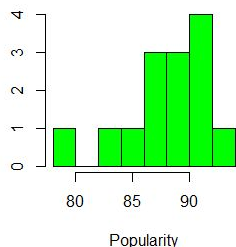
English Song Loudness Distribut



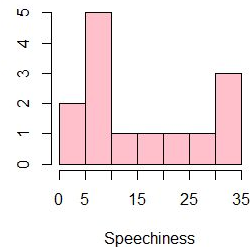
Spanish Song Acousticness



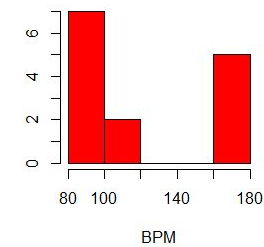
Spanish Song Popularity



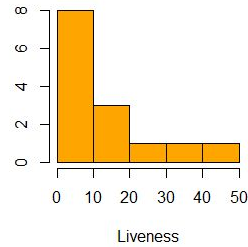
Spanish Song Speechiness



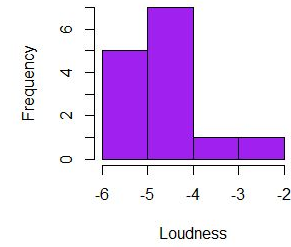
Spanish Songs BPM



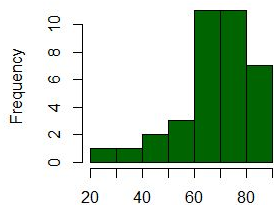
Spanish Song Liveness



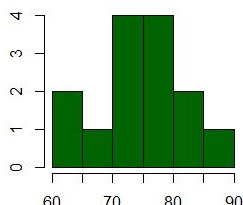
Spanish Song Loudness Distribut



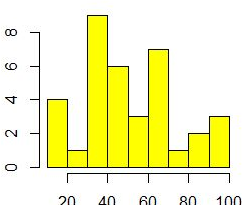
English Song Danceability



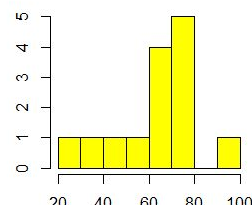
Spanish Song Danceability



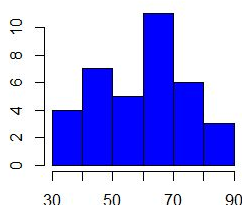
English Song Valence



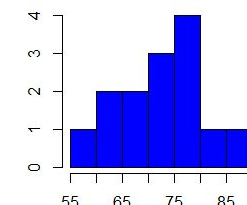
Spanish Song Valence



English Songs Energy



Spanish Songs Energy



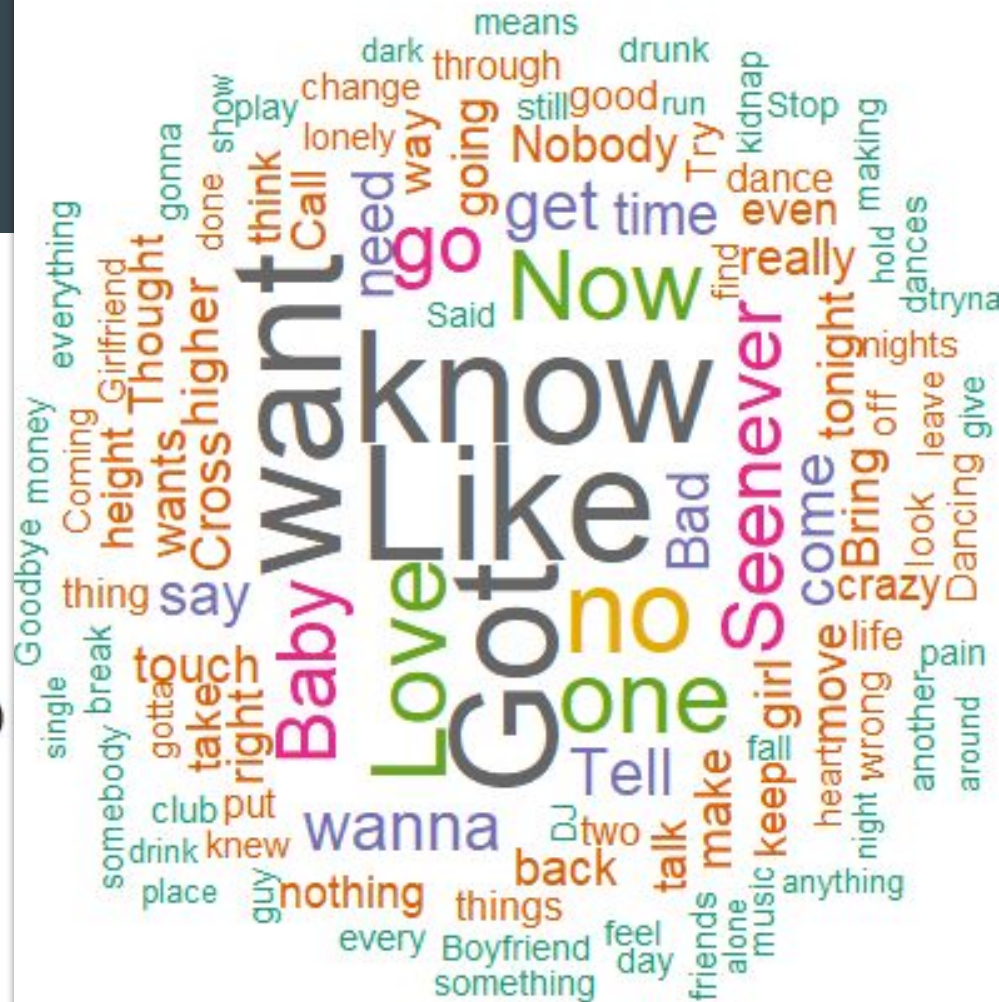
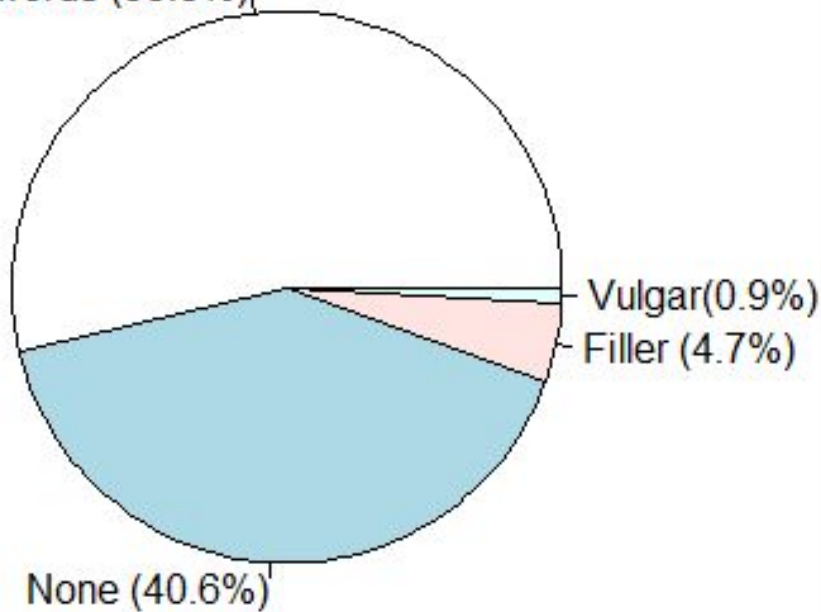
How to make the Cut?

- Genre: Pop, Latin, Reggaeton
- Danceability: 60-90 Range
- Higher Valence, therefore more positive lyrics and vibe
- Higher Energy
- BPM: 80-90 or 160-180
- Not many bell curves, therefore rules are more rigid and strict

Overall Less Room for Variety as English Songs

Keyword Analysis

Pie Chart of Lyrics



Keyword Analysis

- Most common keywords
- Common themes reflected in popular lyrics
- What does this infer about society as a whole?

Row Labels	Count of Keyword	Average of Ranking
Boyfriend	19	2.947368421
girlfriend	21	4.047619048
ya	27	4.925925926
club	20	5.3
Dancing	24	7.666666667
Goodbye	20	8
Nobody	34	8.352941176
uah	16	9
Eh	31	9.870967742
She's	21	11
Who	20	12.95
was	63	13.96825397
ooh	94	14.03191489
somebody	18	14.72222222
Ain't	56	15.26785714
baby	73	15.36986301
guy	19	15.68421053
Try	26	16.11538462
Yeh	19	16.31578947
bad	44	16.31818182
up	54	16.68518519
feel	19	17
gotta	21	17.0952381
fall	17	18.41176471
way	31	18.70967742
keep	30	18.86666667
very	21	19
here	21	19.0952381
Drunk	18	19.22222222
he	20	19.3
by	20	19.35
'Cause	40	19.375
take	33	19.45454545
play	18	19.55555556

Conclusion

- If we were to write a song we would have Ed Sheeran sing it live, classify it as pop and only have english and use the the top keywords.

Final Thoughts

- What we wanted to know about the dataset and how they got of of the numbers
- What else we wanted to do:

Analyze the Chord Composition, but had difficulty finding the correct version, when multiple versions with different chords were uploaded online

Thank you!

Peace out 2021