Quantifying Influence

The Beatles vs. The Rolling Stones

DS4
Max Rose

Influence, not popularity

- Generations of musicians and fans have asked a question, so oft asked it has been reduced to “Beatles or Stones”. Perhaps you are not a fan of either, simply don’t care, or you wonder “who”. (Ask your parents…)
- Regardless, the influence of both artists is undeniable, as is their enduring popularity over the years. And that popularity has been demonstrated via sales of their music and countless top N lists. And this is represented in the work of subsequent artist’s recordings
- Can the data say which artist is “*better*” (more impactful)?

Demonstrating Influence with Data

- The MusicBrainz database represents a comprehensive pool of information about recorded music. The two bands I have chosen represent a relatively small subset of this data yet between them exists large catalog of songs and recordings, with decades of history generating a significant data trail that may provide hidden insights demonstrating their impact on artists for over 5 decades.

By pulling key metrics from the source, can this influence be measured?

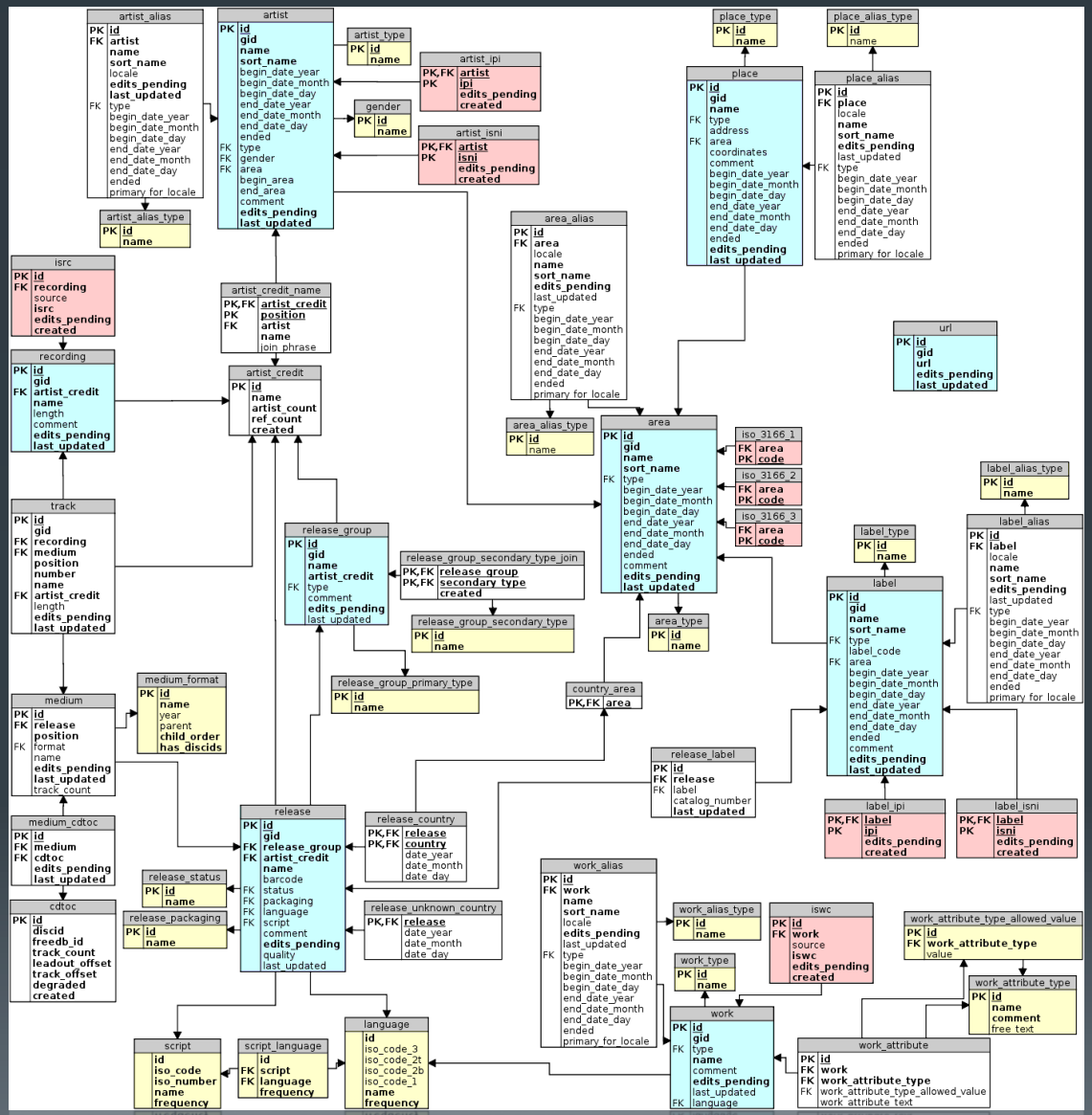
About the Data

- **MusicBrainz is a community-maintained open source encyclopedia of music information. Sim. Wikipedia**
- https://musicbrainz.org/
- Schema Info

- Developers/Users are encouraged to build applications on top of the platform.

Key Data Issues

- Initial exploration of the Musicbrainz API proved challenging to parse the specific relationships I want to explore.
- The entire database is open-source and can be run from a downloadable virtual server.
- SQL queries against the server have been more successful, allowing me to pull a raw dataset with the desired data points for the two artists in question.
- Generally there is too much data which, while clean requires continuing aggregation to reduce to a useful set of features and responses.


artist_alias
PK id
FK artist
name
sort_name
locale
edits_pending
last_updated
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
primary_for_locale
artist
PK id
gid
name
sort_name
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
FK type
FK gender
FK area
begin_area
end_area
comment
edits_pending
last_updated
artist_type
PK id
name
artist_ipi
PK,FK artist
PK ipi
edits_pending
created
gender
PK id
name
artist_isni
PK,FK artist
PK isni
edits_pending
created
artist_alias_type
PK id
name
place_type
PK id
name
place_alias_type
PK id
name
place
PK id
gid
name
FK type
address
FK area
coordinates
comment
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
edits_pending
last_updated
place_alias
PK id
FK place
locale
name
sort_name
edits_pending
last_updated
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
primary_for_locale
area_alias
PK id
FK area
locale
name
sort_name
edits_pending
last_updated
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
primary_for_locale
isrc
PK id
FK recording
source
isrc
edits_pending
created
artist_credit_name
PK,FK artist_credit
PK position
FK artist
name
join_phrase
artist_credit
PK id
name
artist_count
ref_count
created
url
PK id
gid
url
edits_pending
last_updated
recording
PK id
gid
FK artist_credit
name
length
comment
edits_pending
last_updated
area_alias_type
PK id
name
area
PK id
gid
name
sort_name
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
comment
edits_pending
last_updated
iso_3166_1
FK area
PK code
iso_3166_2
FK area
PK code
iso_3166_3
FK area
PK code
label_alias_type
PK id
name
track
PK id
gid
FK recording
FK medium
position
number
name
FK artist_credit
length
edits_pending
last_updated
release_group
PK id
gid
name
artist_credit
FK type
comment
edits_pending
last_updated
release_group_secondary_type_join
PK,FK release_group
PK,FK secondary_type
created
release_group_secondary_type
PK id
name
label_type
PK id
name
label_alias
PK id
FK label
locale
name
sort_name
edits_pending
last_updated
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
primary_for_locale
label
PK id
gid
name
sort_name
FK type
label_code
FK area
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
comment
edits_pending
last_updated
area_type
PK id
name
medium_format
PK id
name
year
parent
child_order
has_discids
release_group_primary_type
PK id
name
country_area
PK,FK area
medium
PK id
FK release
position
FK format
name
edits_pending
last_updated
track_count
release_label
PK id
FK release
FK label
catalog_number
last_updated
label_ipi
PK,FK label
PK ipi
edits_pending
created
label_isni
PK,FK label
PK isni
edits_pending
created
medium_cdtoc
PK id
FK medium
FK cdtoc
edits_pending
last_updated
release
PK id
gid
FK release_group
FK artist_credit
name
barcode
FK status
FK packaging
FK language
FK script
comment
edits_pending
quality
last_updated
release_country
PK,FK release
PK,FK country
date_year
date_month
date_day
work_alias
PK id
FK work
name
sort_name
locale
edits_pending
last_updated
FK type
begin_date_year
begin_date_month
begin_date_day
end_date_year
end_date_month
end_date_day
ended
primary_for_locale
work_alias_type
PK id
name
iswc
PK id
FK work
source
iswc
edits_pending
created
work_attribute_type_allowed_value
PK id
FK work_attribute_type
value
cdtoc
PK id
discid
freedb_id
track_count
leadout_offset
track_offset
degraded
created
release_status
PK id
name
release_packaging
PK id
name
release_unknown_country
PK,FK release
date_year
date_month
date_day
work_type
PK id
name
work_attribute_type
PK id
name
comment
free_text
work
PK id
gid
FK type
name
comment
edits_pending
last_updated
FK language
work_attribute
PK id
FK work
FK work_attribute_type
FK work_attribute_type_allowed_value
work_attribute_text
script
id
iso_code
iso_number
name
frequency
script_language
id
FK script
FK language
frequency
language
id
iso_code_3
iso_code_2t
iso_code_2b
iso_code_1
name
frequency

Metrics to Explore

- # of works (songs) released
- # of times released worldwide
- # *of cover* recordings of each artist's respective catalog
- # of times a given work is covered.
- User ratings original and cover releases.
- Time-based representation of the works and covers.

- *The goal is to build a relationship between the original work and the use by subsequent artists to show and quantify the influence.*