

Python Dataviz: The New Generation

Stephanie Kirmer
@data_stephanie

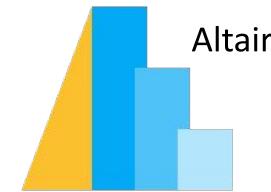
<https://github.com/skirmer/new-py-dataviz>

Our Contenders

The Old Standards



The New Generation



Criteria

- Easy learning curve
- Sensible, consistent grammar
- Flexibility
- Beautiful, readable output

Nice to have: Interactivity

Priority Perspectives

New user
(not power user)

Busy person
(not unlimited time to waste)

Care how it looks
(Want to be visually compelling)

The Tests

- Histogram
- Scatterplot
- Faceted Scatterplot
- Grouped Bar
- Time Series Line
- BONUS: 3D Scatterplot

Dataset: [Spotify Tracks](#)

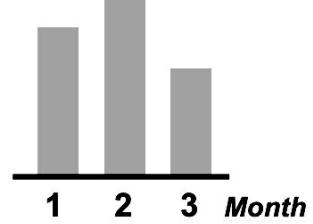
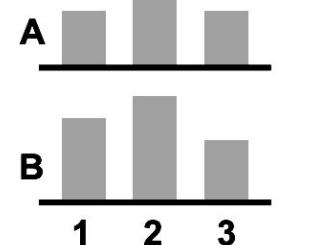
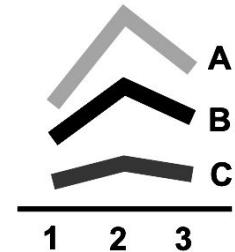
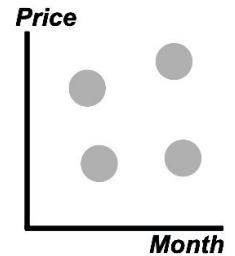
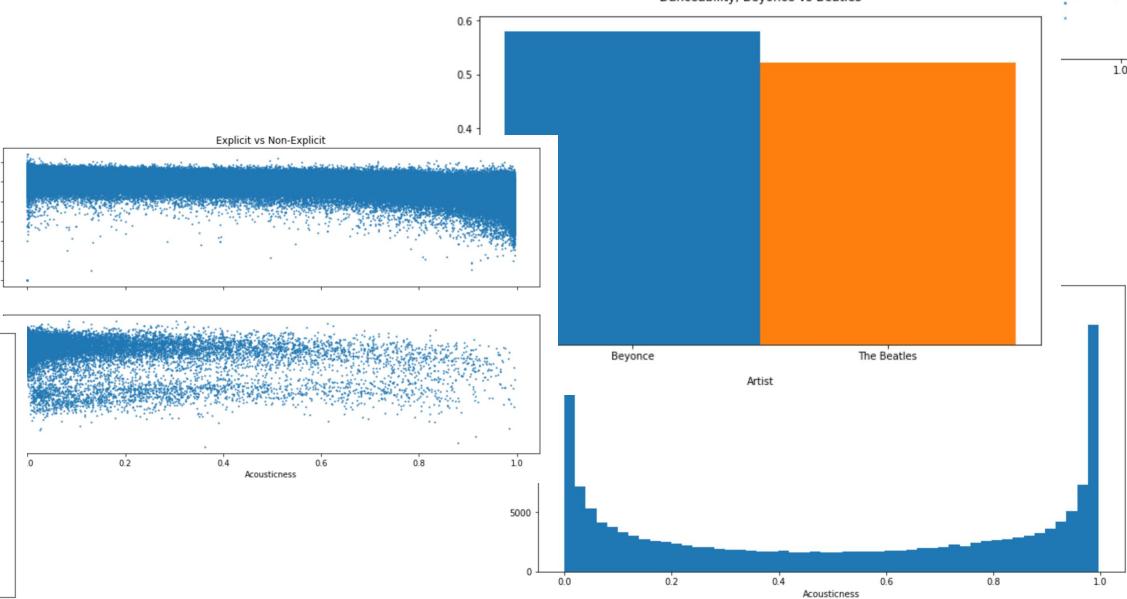
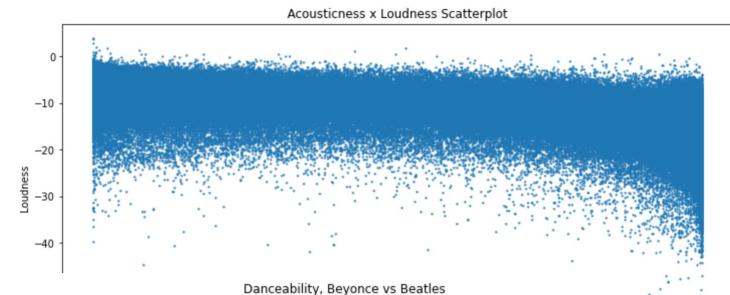
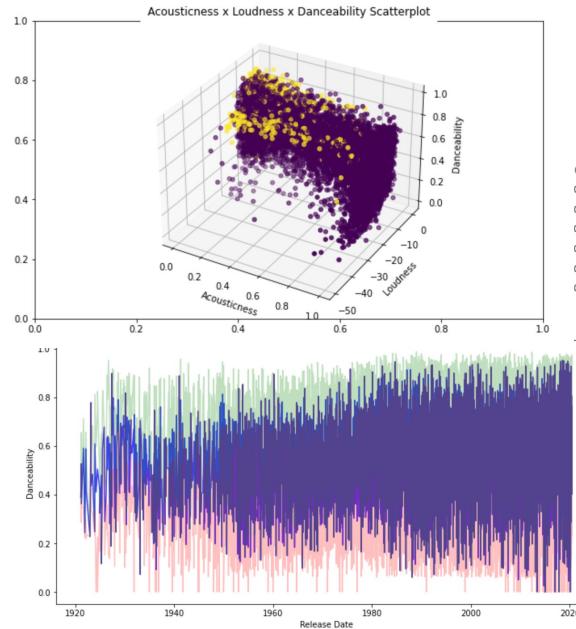
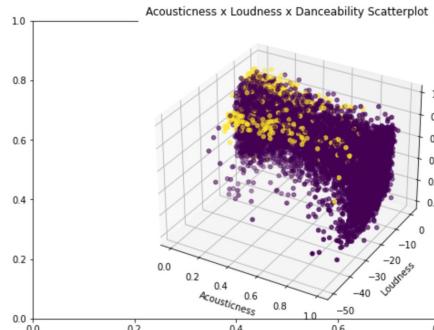


Image credits: [Steve Franconeri, 2019](#)

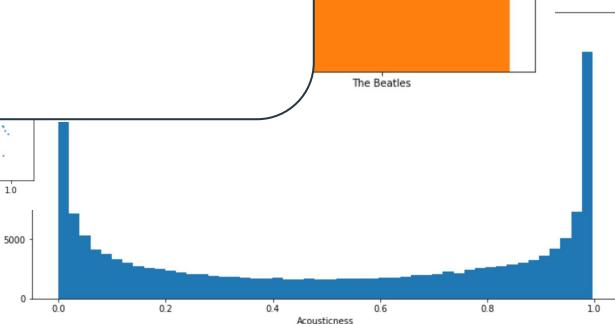
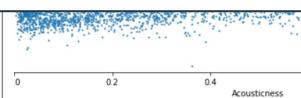
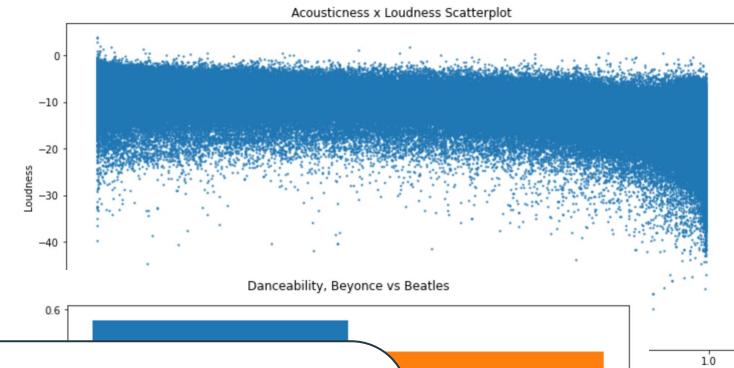
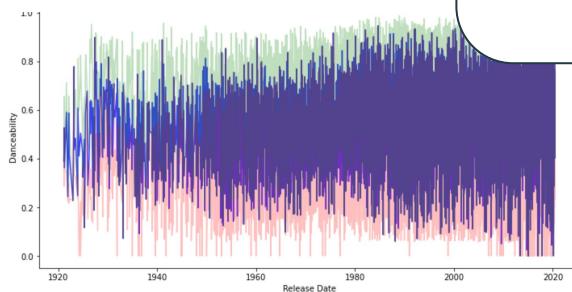
matplotlib



matplotlib

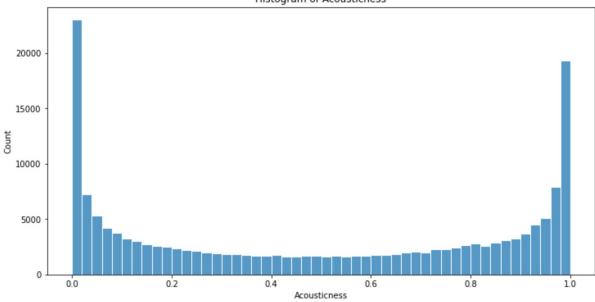
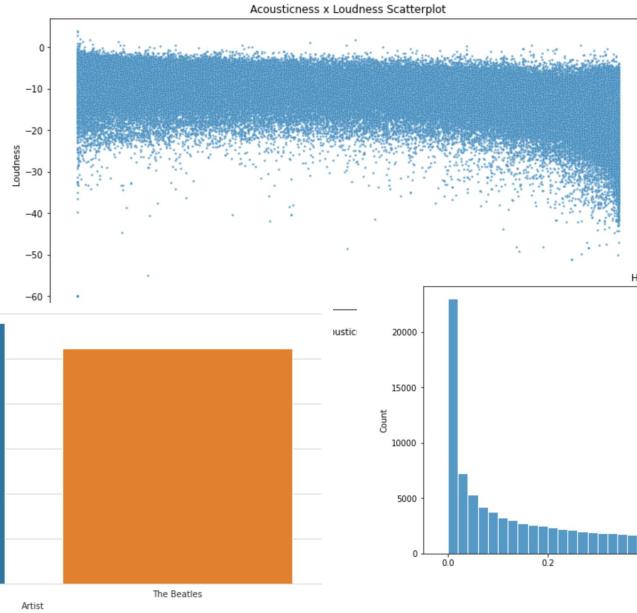
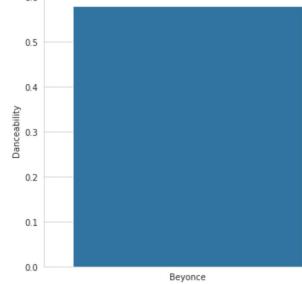
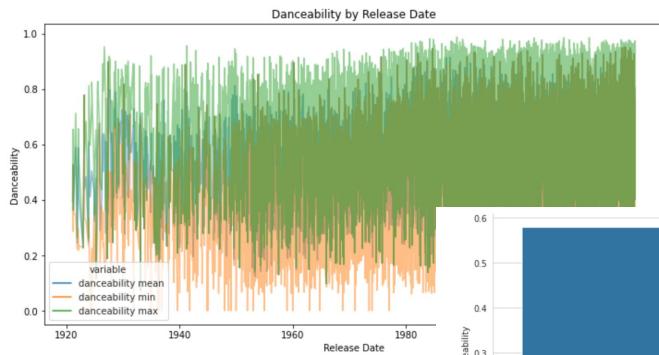


- Easy learning curve: No.
- Sensible, consistent grammar: Mixed.
- Flexibility: Mixed.
- Beautiful output: No.
- Interactivity: No.



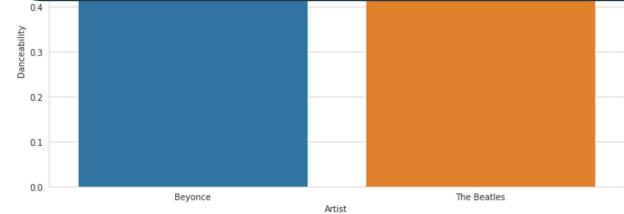
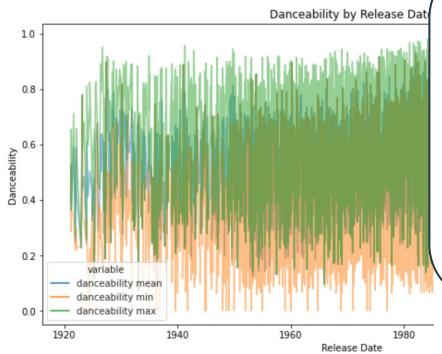


seaborn

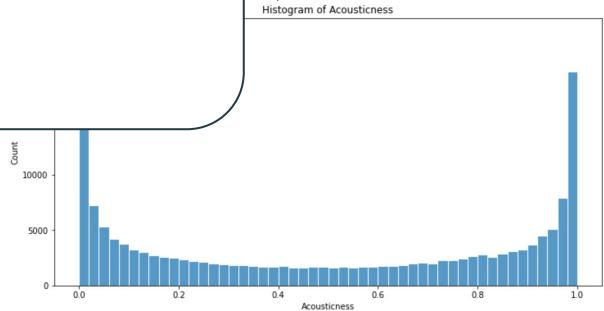
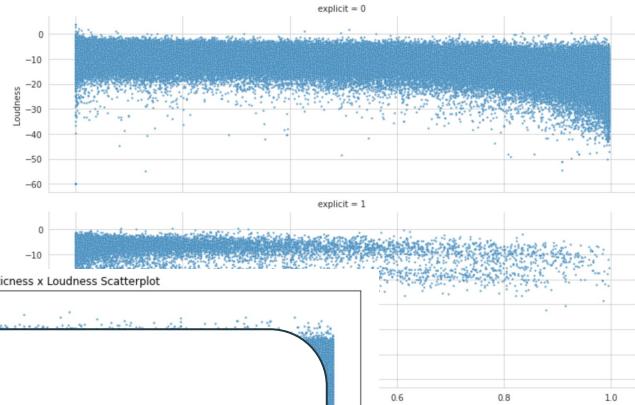


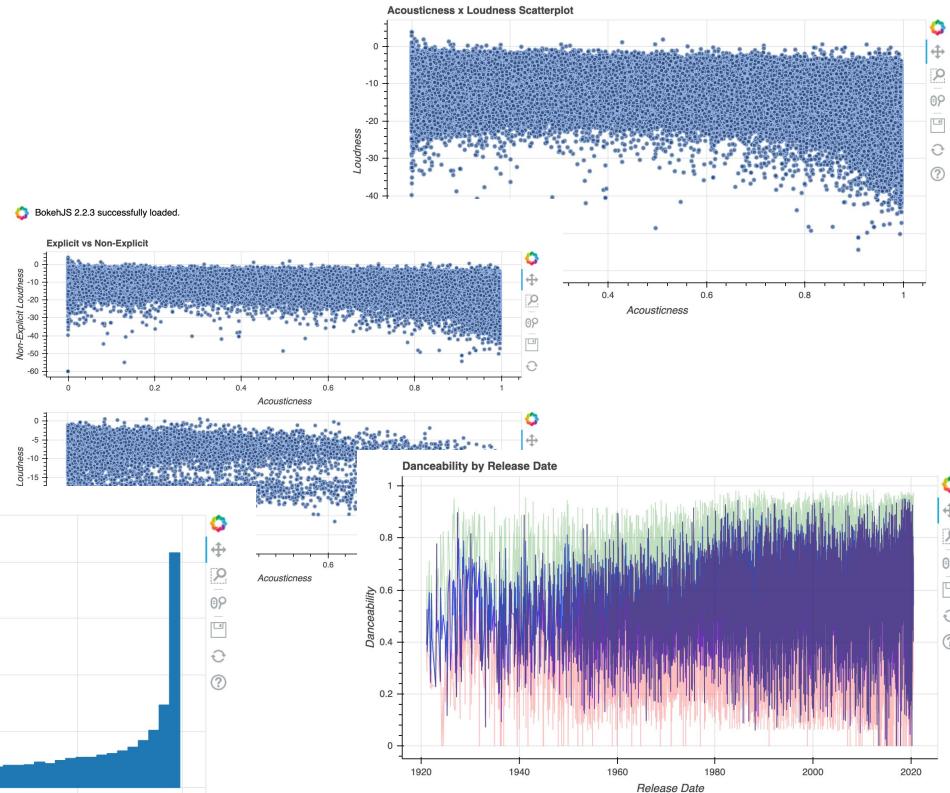
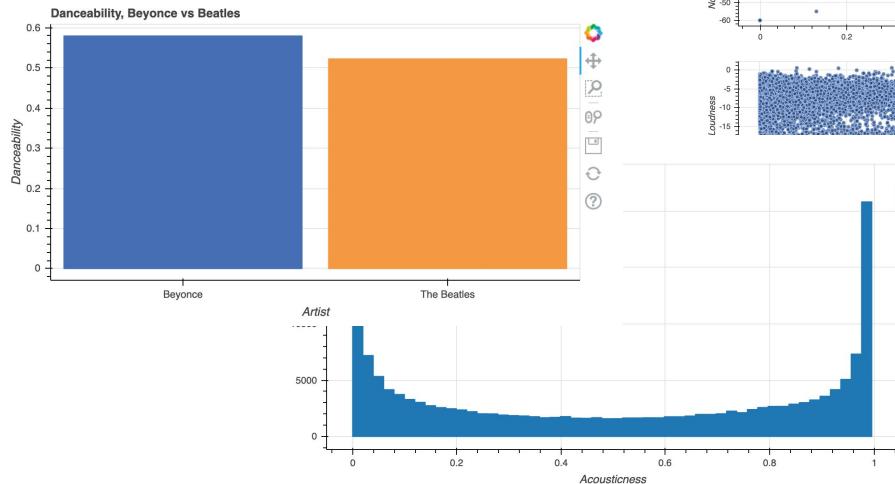


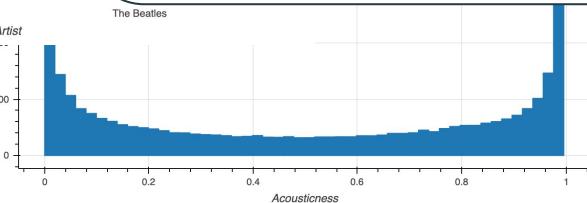
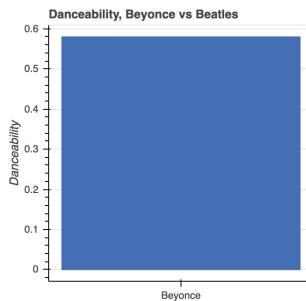
seaborn



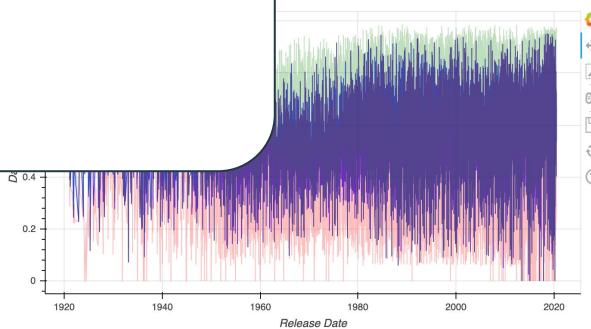
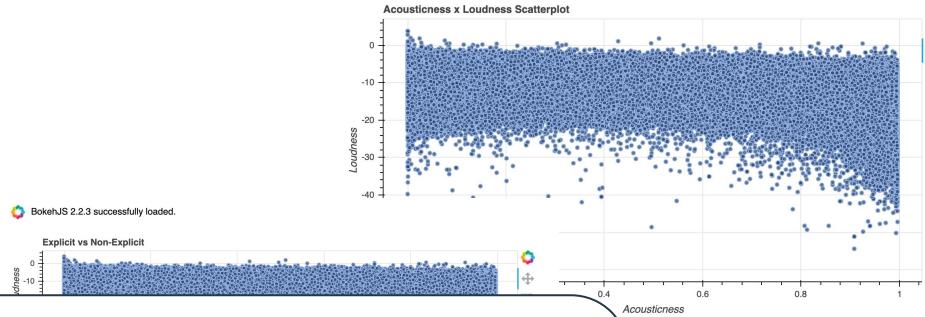
- Easy learning curve: No.
- Sensible, consistent grammar: No.
- Flexibility: Better.
- Beautiful output: Better.
- Interactivity: No.

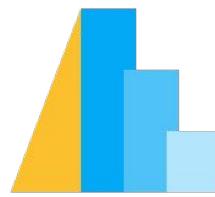




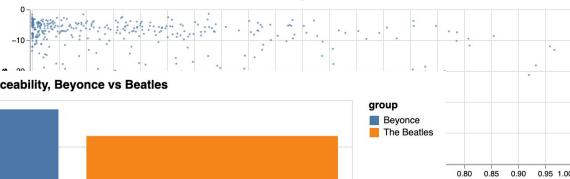
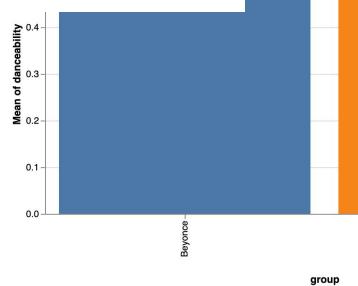
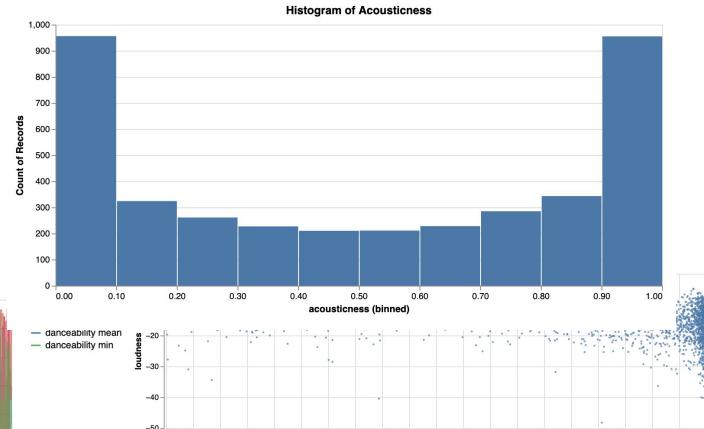
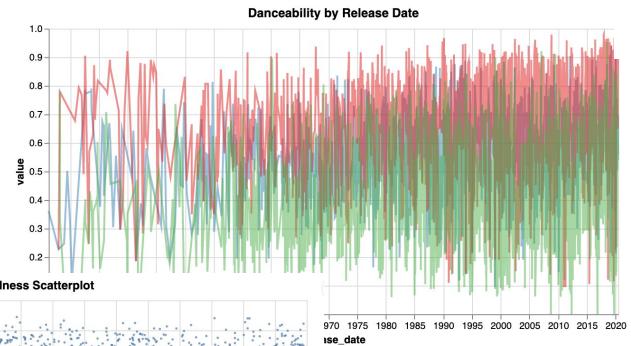
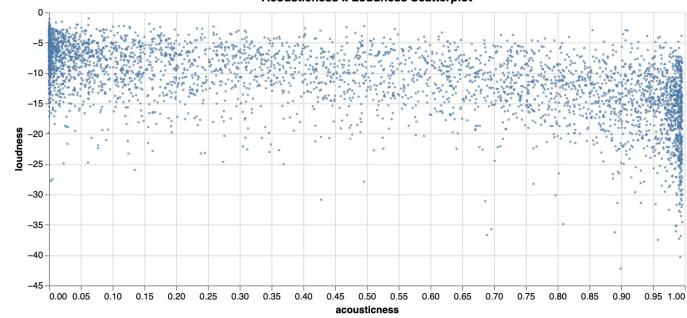


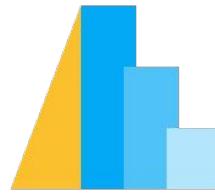
- Easy learning curve: Yes.
- Sensible, consistent grammar: Mixed.
- Flexibility: Yes.
- Beautiful output: Yes.
- Interactivity: Yes.



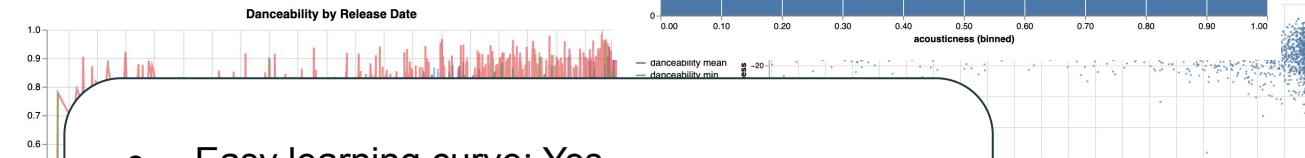
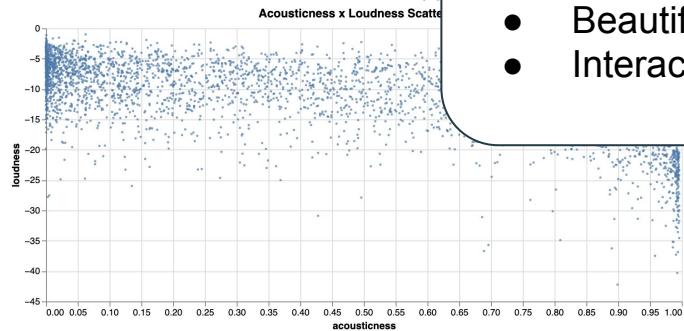


Altair

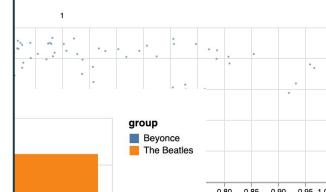
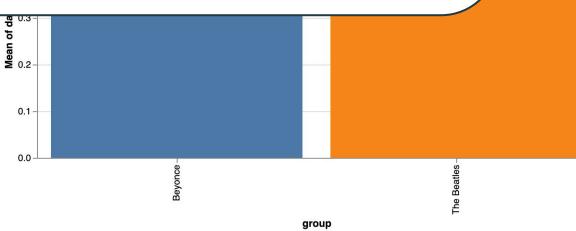




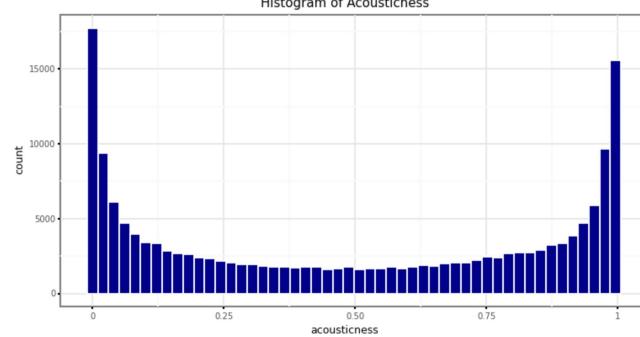
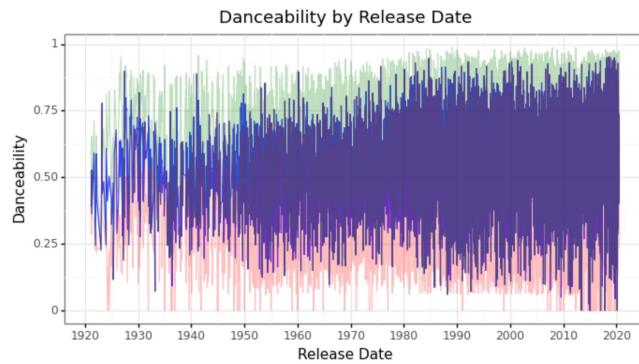
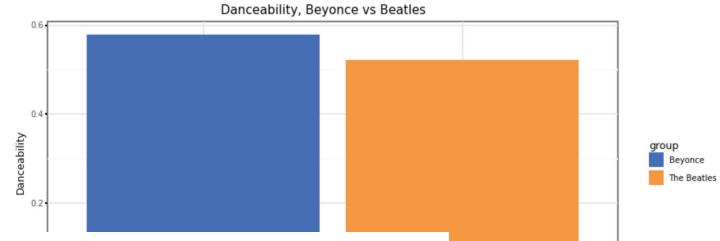
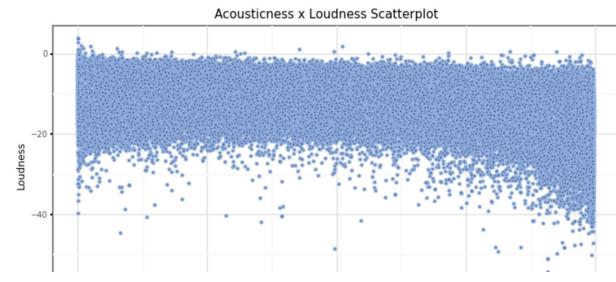
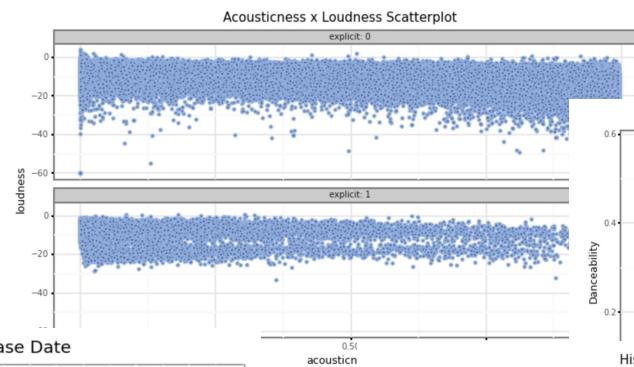
Altair

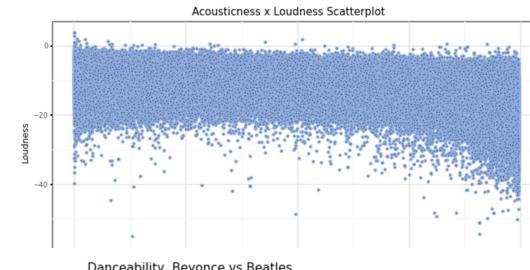
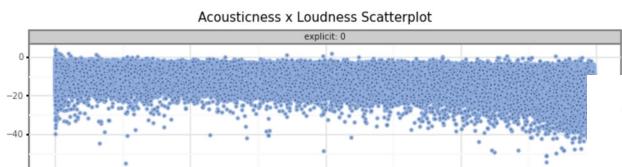


- Easy learning curve: Yes.
- Sensible, consistent grammar: Yes.
- Flexibility: No - data limitation.
- Beautiful output: Yes.
- Interactivity: No.

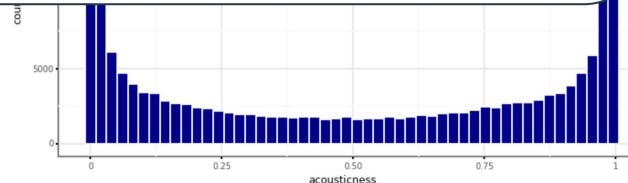
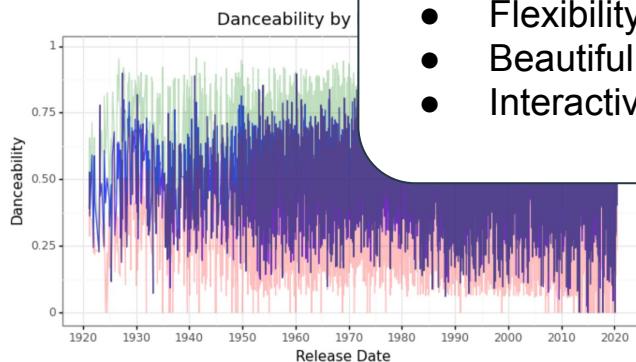


Histogram of Acousticness





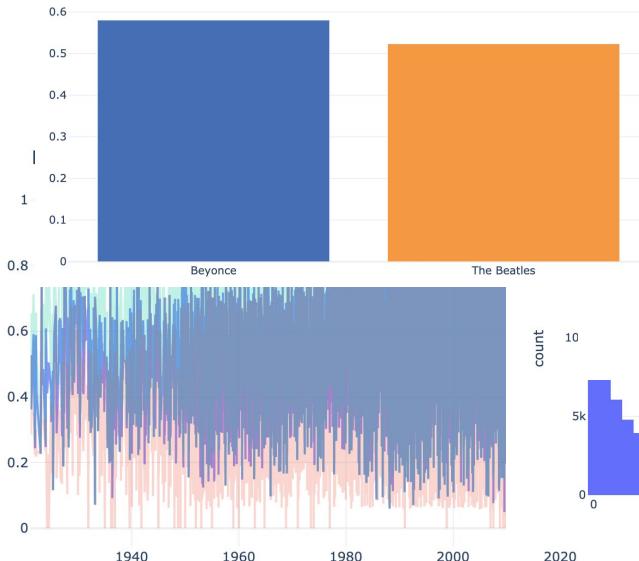
- Easy learning curve: Yes, especially if you know R.
- Sensible, consistent grammar: Yes.
- Flexibility: Yes.
- Beautiful output: Mixed.
- Interactivity: No.



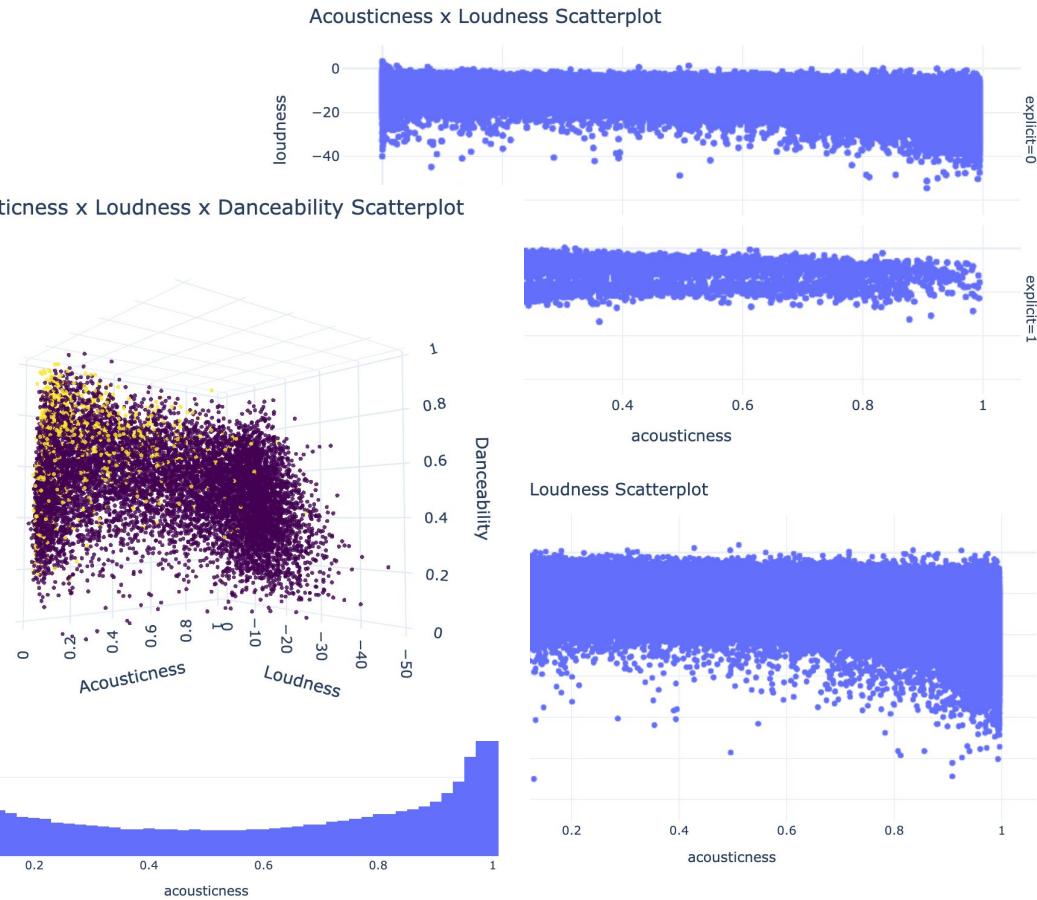


plotly

Danceability, Beyonce vs Beatles

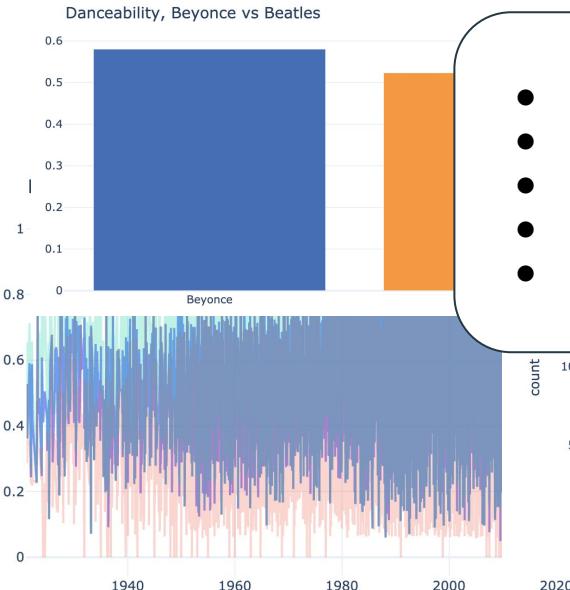


Acousticness x Loudness x Danceability Scatterplot

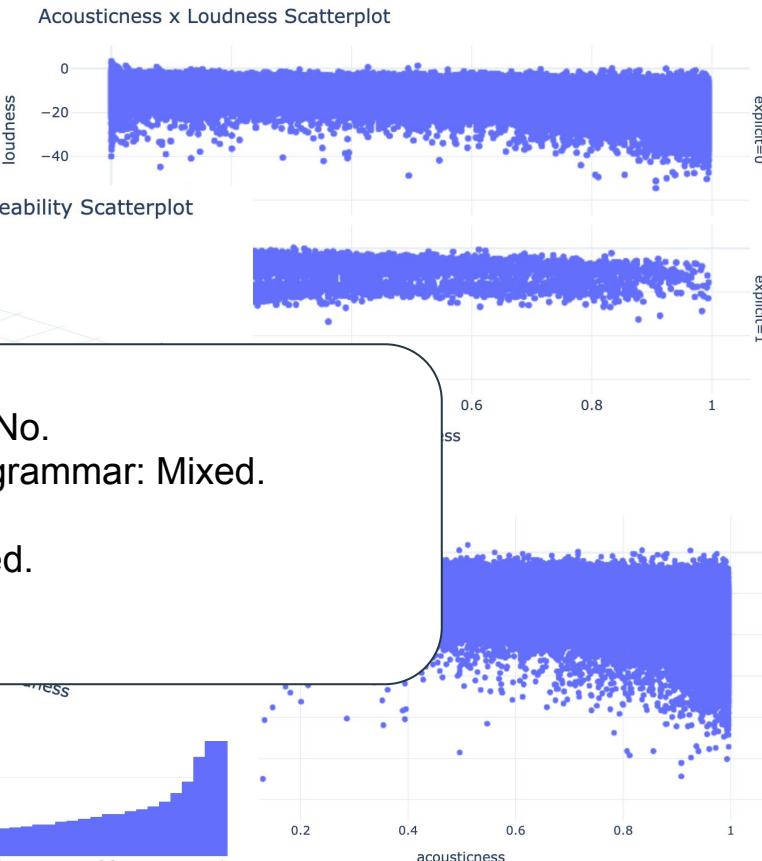




plotly



- Easy learning curve: No.
- Sensible, consistent grammar: Mixed.
- Flexibility: Yes.
- Beautiful output: Mixed.
- Interactivity: Yes.



Choose Your Fighter

Sensible grammar

Plotnine or Altair*

Easy learning curve

Plotnine or Altair*

Interactivity

Plotly or Bokeh

Beautiful images

Altair* or Bokeh

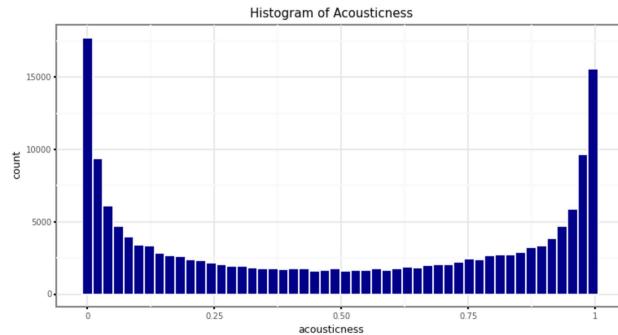
Flexibility

Plotnine or Bokeh

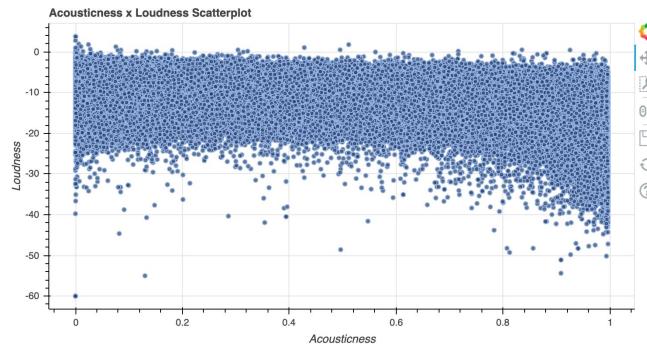
*provided you don't have big datasets.

Best-Of Awards

Histogram: Plotnine



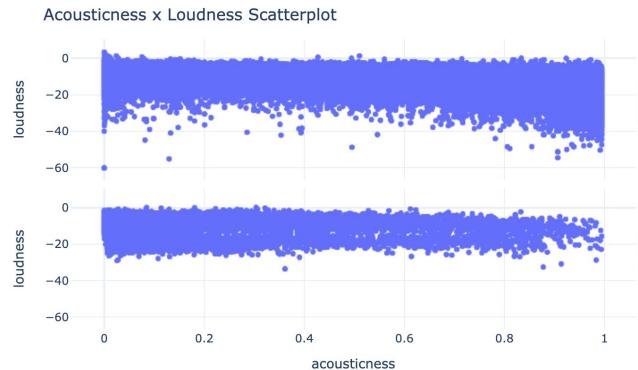
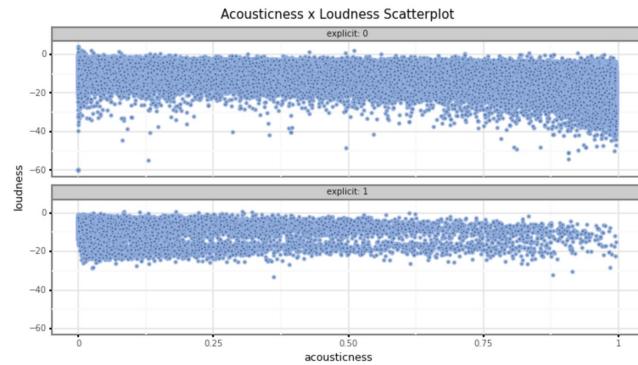
Plain Scatter: Bokeh



According to me, unscientifically.

Best-Of Awards

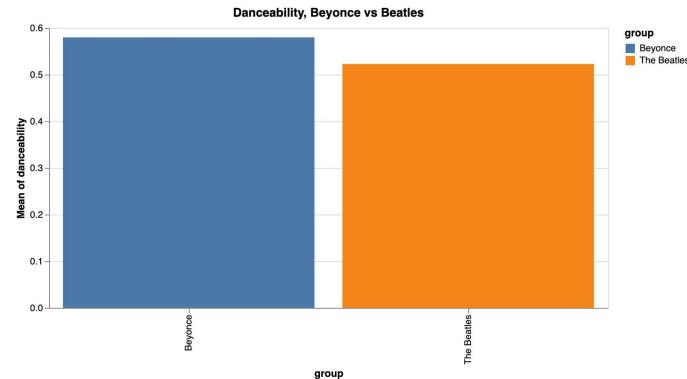
Faceted Scatter: Plotnine/Plotly tie



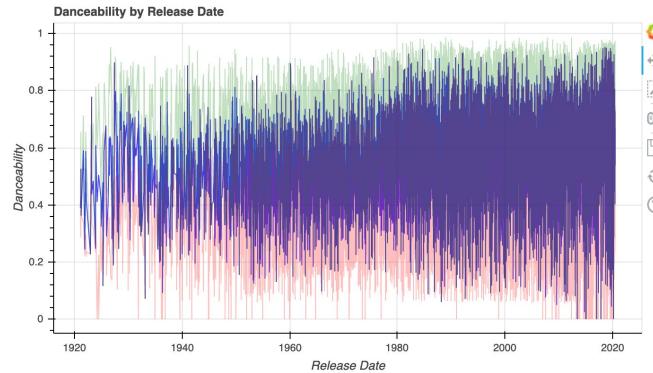
According to me, unscientifically.

Best-Of Awards

Grouped Bar: Altair



Timeline: Bokeh



According to me, unscientifically.

Conclusions

There is not a perfect Python dataviz library.

You have lots of wonderful choices.

Choose what works for you!

Thank you!

github.com/skirmer/new-py-dataviz

www.stephaniekirmer.com
@data_stephanie

Saturn Cloud
saturncloud.io