[**ht**](https://readypipe.com/blog/9-surprises-from-analyzing-spotify-streams-data/)

**Article (Writeup):**

* Sample (for reference) → https://arxiv.org/pdf/1803.09010.pdf
* 1500-3000 words

**Datasheet:**

**Analysis:**

* Clustering (by genre and/or by artist)
* Regression by variables (see if it changes year by year)
  + Popularity
  + Independent variables: danceability, energy, liveness, valence, tempo
  + Correlations,
    - Danceability & energy
    - Danceability & valence
    - Valence & tempo
    - Speechiness & tempo
  + collinearity? etc.
* Networks to track artist collaboration
  + Can rank artists based on # of contributions / collaborations