PRODUCTION VALUE
Identifying Record Producers from Audio Data
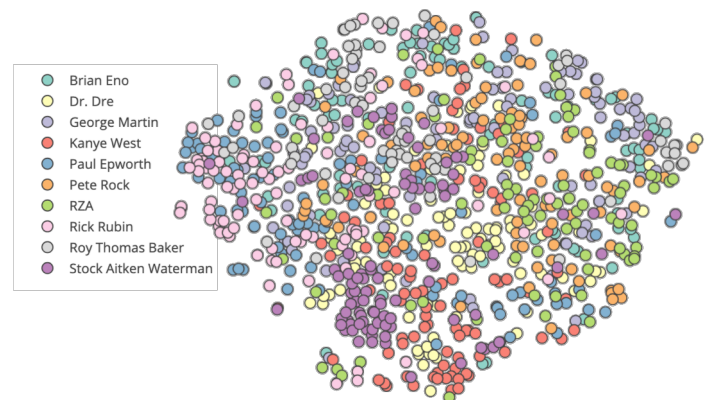### Maxwell Fisch

Background
A record producer controls the creation of a music album. Each music producer leaves a sonic fingerprint on every album they produce. Production Value seeks to identify and quantify that signature. This model could be used for:
- **Music discovery:** aiding Spotify and Pandora users in finding music they like.
- **Music publishing:** helping record labels identify and distribute royalties to song collaborators.

Approach
- 1000 songs from 10 producers (100 songs each) were chosen.
- 30-second snippets of audio from Spotify's API were processed and transformed into Mel-Cepstral Frequency Coefficients (MFCCs) – a featurization that roughly translates to timbre.
- Principal Component Analysis (PCA) reduced dimensionality (24k features → 12).
- K-Nearest Neighbors (KNN) classification was used to classify the record producers.

2D t-SNE Plot of Producers

- Brian Eno
- Dr. Dre
- George Martin
- Kanye West
- Paul Epworth
- Pete Rock
- RZA
- Rick Rubin
- Roy Thomas Baker
- Stock Aitken Waterman

Results and Analysis
- KNN model multiclass accuracy (10 balanced classes) was 44% compared to a baseline of 10%.
- Producers and songs cluster in MFCC-space. Some producers have characteristic sound (e.g. Stock Aitken Waterman), others are diverse (e.g. George Martin).
- Future improvements:
    - Deconvolution of variables (e.g. Artist, Album, Genre) via nuanced feature engineering.
    - Neural Network Model with Increased Scale (>100k songs)

Confusion Matrix

True label \ Predicted	Brian Eno	Dr. Dre	George Martin	Kanye West	Paul Epworth	Pete Rock	RZA	Rick Rubin	Roy Thomas Baker	Stock Aitken Waterman
Brian Eno	13	2	7	2	4	1	2	0	0	2
Dr. Dre	0	11	4	3	1	3	2	0	1	6
George Martin	2	1	8	0	1	2	4	0	1	3
Kanye West	0	4	2	14	1	0	0	0	0	4
Paul Epworth	0	0	4	1	16	0	0	2	1	4
Pete Rock	3	4	5	0	0	12	6	0	1	0
RZA	4	5	1	1	1	3	9	0	1	3
Rick Rubin	4	2	2	0	11	2	0	12	1	3
Roy Thomas Baker	1	0	6	1	1	1	1	0	8	2
Stock Aitken Waterman	0	1	1	1	1	1	1	0	5	24

Tech Stack: python, mongoDB, NumPy, matplotlib, Pandas, Spotify, librosa, plotly, scikit learn

linkedin.com/in/maxwellhfisch | github.com/mhfisch | @Maxwell_Fisch
maxwellfisch@gmail.com | (805) 982-0955

Project Links: http://github.com/mhfisch/production-value | http://www.production-value.com