

BEATS OF BIAS: ANALYZING LYRICS WITH TOPIC MODELING AND GENDER BIAS MEASUREMENTS



Chen Danqing, Adithi Satish, Rasul Khanbayov, Carolin M. Schuster, Georg Groh

Technical University of Munich, Germany

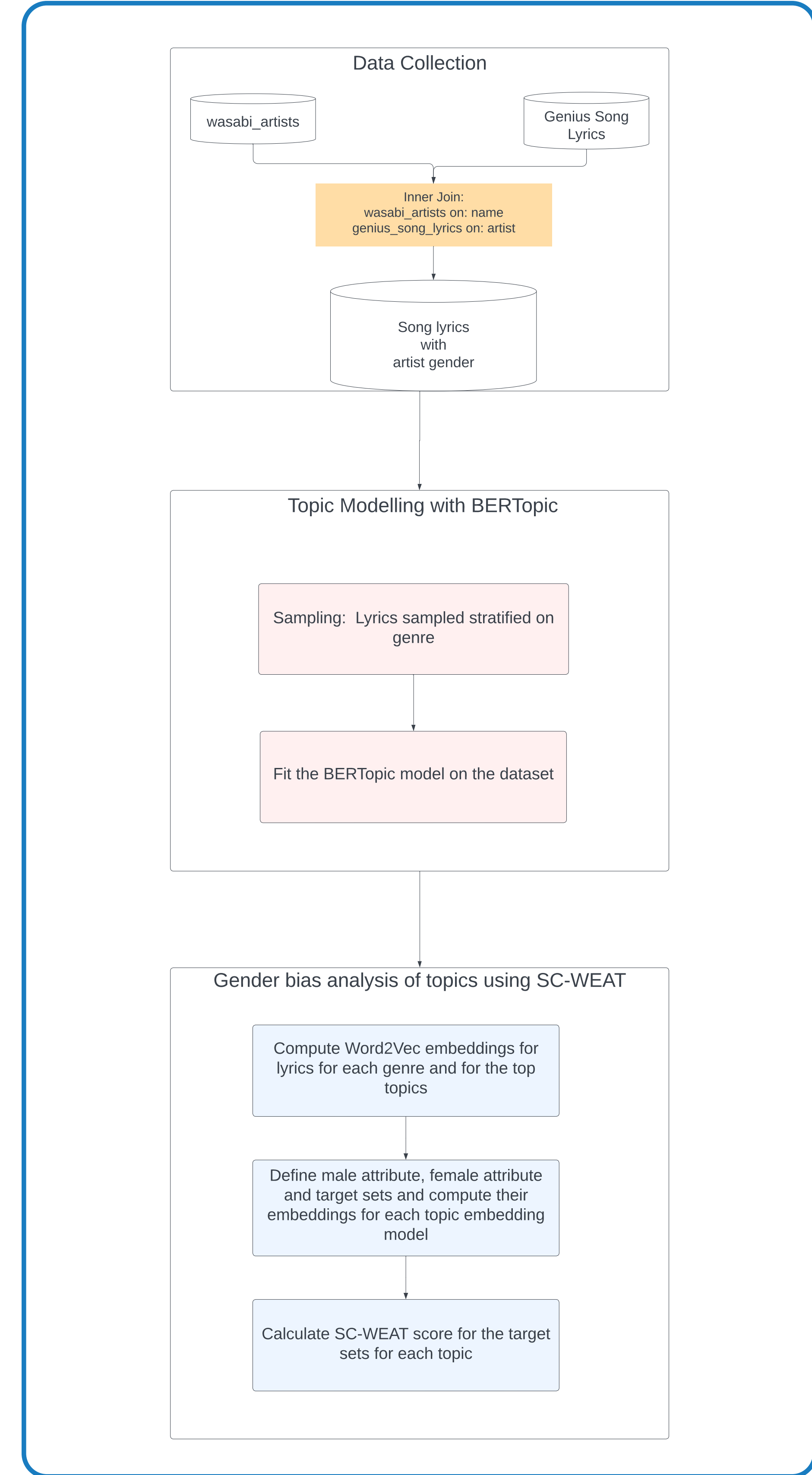
CONTRIBUTIONS

- Topic analysis, using a stratified sample, to determine cross-genre topics, recurrent themes and charting their historical development.
- Analysis of the prevalence and variation of the gender bias in lyrics across topics and genres, using SC-WEAT scores.

DATA

- WASABI (Artist Info) + Genius Song Lyrics
- 537k records & 5 genres (pop, R&B, rock, country and rap)

PIPELINE



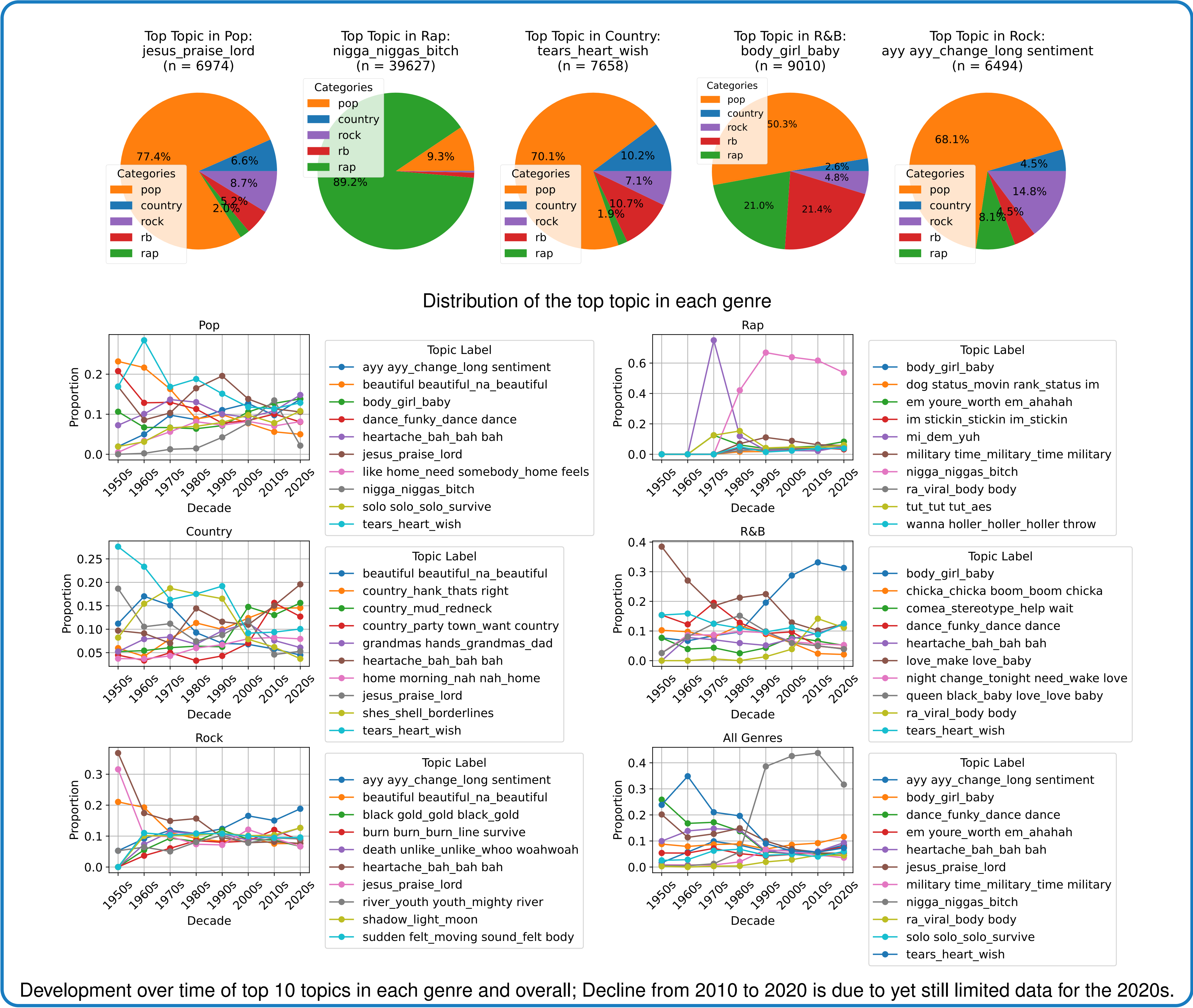
TECHNICAL DETAILS

- **Topic Modelling: BERTopic**
 - BERTopic uses transformers and c-Tf-Idf to create clusters and identify recurrent themes in lyrics
 - Embedding Model: All-MiniLM-L6-V2
- **SCWEAT Analysis**
 - SCWEAT measures the association of a target set of words to two attribute sets to determine bias
 - Word Embeddings: Word2Vec
 - WEAT Score Formulas:

$$s(w, A, B) = \text{mean}_{a \in A} \cos(\vec{w}, \vec{a}) - \text{mean}_{b \in B} \cos(\vec{w}, \vec{b})$$
$$SCWEAT(X, A, B) = \sum_{x \in X} s(x, A, B)$$
$$d = \frac{\text{mean}_{x \in X} s(x, A, B)}{\text{stddev}_{x \in X} s(x, A, B)}$$

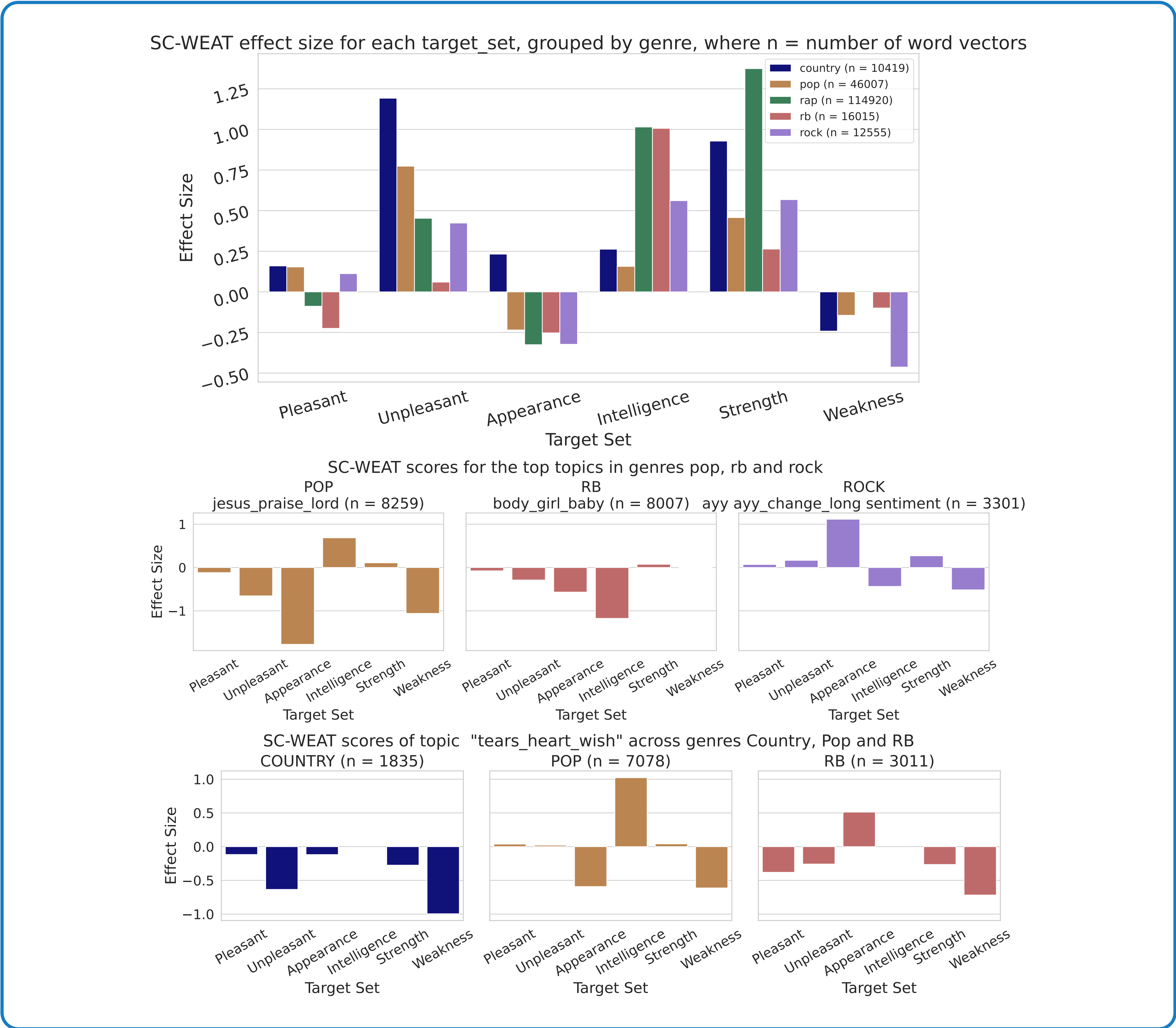
TOPIC ANALYSIS

- Total of 541 topics with 1.5% outliers; Pop songs have a diverse range of topics; An overwhelming 89.2% of the songs in the topic "nigga_niggas_bitch" are rap songs.
- "nigga_niggas_bitch" contains explicit and misogynistic language, highlighting negative female stereotypes
- Shift in music themes from romance and emotional content in 1950s - 1980s (top topics: "wonderful_sweeter_years_sweeter", "tears_heart_wish") to profanity and sexualization in 1990s - 2020s (top topic: "nigga_niggas_bitch")



SCWEAT ANALYSIS

- Attribute Sets: Male and Female attributes; 6 target word sets: Pleasant, Unpleasant, Appearance, Intelligence, Strength, Weakness.
- Men more likely to be associated with competence and intelligence whereas women more likely to be associated with weakness
- Appearance words show female bias in 4/5 genres, supporting trend of objectification and sexualization of women
- Not only do different topics show different bias characteristics but even the same topic shows differences across genres



CONCLUSION

- Noticeable shift in themes from romance to increased sexualization and profanity in songs: 1950s-1980s included "wonderful_sweeter_years_sweeter" "tears_heart_wish"; 1990s onwards: "nigga_niggas_bitch"
- SC-WEAT results show implicit gender bias in lyrics: Words related to Weakness and Appearance are biased towards women whereas Intelligence words show a male bias.
- Combining topic analysis with bias measurements gives insights into lyrics, highlighting how lyrics reinforce gender roles and stereotypes.