

# Project Proposal

## Banshee

Amish Goyal, Vinay Kola

April 6, 2014

We aim to examine the effect of weather on musical preferences. We would like to explore the effect of weather on various aspects such as genre as well as more latent aspects such as mood of the song, which can be obtained by linguistic analysis of the lyrics.

There has been research into the effect of weather in the fields of finance and psychology. Murray et al. [7] analyzed the effect of weather on consumer spending by examining correlation between a variety of weather variables and the daily sales of a small independent retailer. They also observed that sunlight affects mood. Zimmerman [9] studied the effect of mood on music preference and found that happy music is liked more by people that feel happy than people that feel sad. As mentioned, weather is shown to affect mood and music preference has been linked to a person's mood. Therefore, we hypothesize that the weather conditions could impact one's music preferences.

One of the hypotheses we would like to test is that rainy days cause people to listen to gloomier songs, while sunny days cause people to listen to happier music. We will also investigate whether the long-term climate of a place has a bearing in the popular music at that particular location.

There are various aspects of musical preferences that we would like to investigate. Markus Schedl describes a broad categorization of music descriptors into music content, music context, user context and user properties. [3]

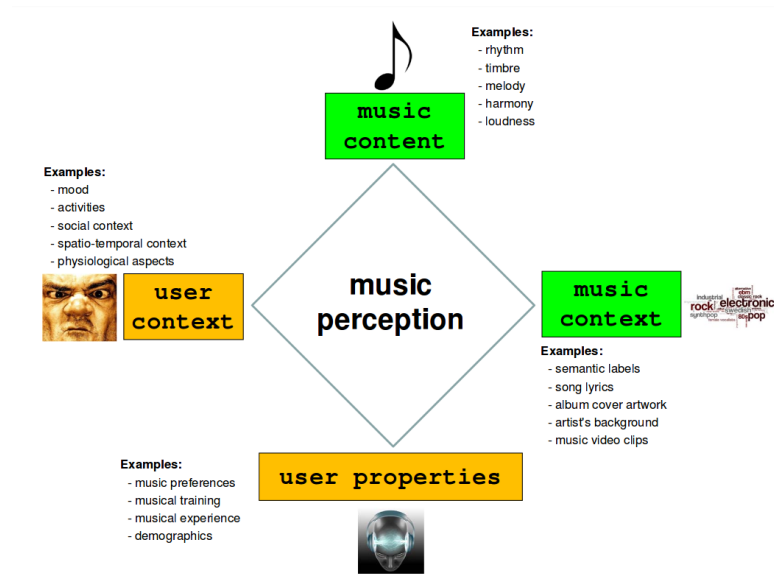


Figure 1: Categorization of music and user descriptors.

Previous studies on music popularity have used Last.fm [2], however the geographic and temporal granularity of the Million Musical Tweets Dataset (henceforth referred to as MMTD) [1] makes it a far more suitable dataset to analyze the influence of weather. Last.fm has information regarding weekly chart lists, but that is too large a time period to observe changes in weather.

The MMTD gives us the geospatial, temporal and social context. We will use the Million Song Dataset [4] to get some metadata regarding music content (eg. beats, danceability, energy measure, tempo, etc.) and use the Lyrics Wikia [5] to get the lyrics of the song. We gather the weather data from National Climatic Data Center [6]. We have daily summaries of weather data for all cities across the world, with data such as snow levels, precipitation and temperature available.

The first step would be to establish whether there is a statistically significant variance in the songs listened to based on climate, using the Kruskal-Wallis ANOVA test. Next, we could estimate the effect of the various aspects of weather (snow, precipitation, temperature, etc.) using a random effects model, similar to the methodology employed in [7]. We plan to group together songs that people listen to during different weather conditions. This would allow us to analyze/compare them based on their lyrics (from word frequencies and topic modelling). The general mood and the preferred genre can also be observed from analysis of the song metadata (gathered through Million Songs Dataset) and linguistic analysis on lyrics.

Music services like Spotify already have recommended playlists for rainy days etc. By studying the preferred type of music for various climatic conditions, these recommendations could be improved. Additionally, from a research perspective, we could not find any prior work linking weather to musical prefer-

ences, so it would establish a precedent if we could prove a statistically significant relation.

Keller et al. [8] found that exposure to higher temperatures predicted increased mood during spring but had opposite effect during summers, especially in southern climates where high temperatures are extremely unpleasant. Hence, it might be important to control for seasons when considering the weather conditions for a particular day. Also, record labels may release different songs for the summer which could be controlled by giving importance to older songs.

Conducting this analysis will not cost anything, as the Million Musical Tweets dataset is free, as is the weather data. The project will take approximately 2 weeks of time between the both of us.

## References

- [1] Hauger, David, et al. "The Million Musical Tweets Dataset: What Can We Learn From Microblogs." Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR), Curitiba, Brazil. 2013.
- [2] Baur, Dominikus, et al. "The streams of our lives: Visualizing listening histories in context." Visualization and Computer Graphics, IEEE Transactions on 16.6 (2010): 1119-1128.
- [3] Schedl, Markus. "Ameliorating Music Recommendation: Integrating Music Content, Music Context, and User Context for Improved Music Retrieval and Recommendation." Proceedings of International Conference on Advances in Mobile Computing & Multimedia. ACM, 2013.
- [4] Bertin-Mahieux, Thierry, et al. "The million song dataset." ISMIR 2011: Proceedings of the 12th International Society for Music Information Retrieval Conference, October 24-28, 2011, Miami, Florida. University of Miami, 2011.
- [5] [http://lyrics.wikia.com/Lyrics\\_Wiki](http://lyrics.wikia.com/Lyrics_Wiki)
- [6] Klein Tank, A.M.G. and Coauthors, 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int. J. of Climatol., 22, 1441-1453.
- [7] Murray, Kyle B., et al. "The effect of weather on consumer spending." Journal of Retailing and Consumer Services 17.6 (2010): 512-520.
- [8] Keller, M. C., Fredrickson, B. L., Ybarra, O., Côté, S., Johnson, K., Mikels, J., ... & Wager, T. (2005). A warm heart and a clear head the contingent effects of weather on mood and cognition. Psychological Science, 16(9), 724-731.
- [9] Zimmerman. The effect of mood on music preferences.