## **Addition notes**

Below are some observations that we made during developing this project. Since most of them are irrelevant with classification itself, I just put them in additional notes. Also keep brief.

## 1. Most frequent 3-gram by genre and artist

It's interesting to see how an artist or a genre is unique in wording. Below are tables of 3-gram stats.

Genres				
Electronic	Нір-Нор	Jazz	Rock	
Boom boom boom	Stay with me	I love you	I don't want	
In the air	La la la	Fall in love	I want to	
Are you ready	I'm the man	Somebody to love	I just can't	
I want to	I want to	Doo wah doo	It's just a	
Clap your hands	Tear it off	Let it be	I want you	
Oh oh oh	Up in the	Fool of me	I know that	

Artists				
Celine Dion	Beyoncé	Elvis Presley	The Carpenters	
You were my	Oh oh oh	Baby baby baby	And I know	
When I couldn't	I love your	Didja ever get	Can't smile without	
Little bit of	Eh eh eh	No no no	I love you	
You make me	Know bout me	Come back baby	In love with	
I love you	Run the world	Ever get one	Over the world	
You gave me	To the left	Gonna study war	I believe you	

TABLE. MOST FREQUENT 3-GRAMS FOR SELECTED GENRES AND ARTISTS

These 3-grams are picked from each genre or artist's top 10 frequent 3-gram results. First, let's look at genres. Jazz is much different than the others, talking more about love and romance. Rock, Hip-Hop and Electronic are more energetic, or more about intriguing people's passion and ego, etc. Surprisingly, I found most of the 3-grams of Rock music start with "I", talking a lot about oneself, while Electronic also try to activate friends and the others. That might be the reason why Rock music was popular for a long period of time back in 70s and 80s - people can gain sort of empathy when they resonate with the singer, trying to let out self-ego with singing Rock songs. Just some groundless conjecture, hehe...

As for artists, just like what I said in the report - Beyoncé loves to "Oh", which is a surprising and funny fact . And for the others, we can clearly see some style in each one's lyrics. I'd like to describe The Carpenters as loving, slow and peaceful, Elvis as romantic, and Celine Dion as gentle or tender. Or you can interpret from your own angle.

## 2. Word Cloud attempts

Word Cloud was a fancy visualizing method of textual data. The idea was simple - given certain

weight for a word or phrase (usually its word count), place this word randomly in a graph with font size and font weight, sometimes even hex depth of color corresponding to its weight. However, despite the naivety of this method, it really gives us an intuitive way of learning the metrics of a group of word or a piece of text.



FIGURE. A WORD CLOUD BASED ON ELVIS PRESLEY'S LYRIC WORD COUNT

Yep. In memory of The King.