

# Oh my Lorde

## Project Intro

With the release of Pure Heroine, Lorde singlehandedly redefined the sound of alternative music for many to emulate. Her style pays homage to the era of grunge music whilst introducing elements of youth and pop evident in her lyrics. More so, her unique, dark sound has attained a cult-like following and appeals to youth culture — discussing topics of the ephemeral, love, and what it feels like to grow up. Her second album, Melodrama, follows her journey after an intense love, an even more heartwrenching break-up, and the newfound sense of self that followed.

Growing up listening to Lorde, I've found myself drawn to her music and have decided to delve into her lyrics for my project. I will be discovering her most commonly used lyrics in both Pure Heroine and Melodrama as well as drawing positive or negative sentiments from her work. Finally, I'll be looking at the frequency of some of the words used within the albums

## Questions

- **[Primary Question]** What are the most commonly used words in both Pure Heroine and Melodrama?
- What are the themes / sentiments of both the albums — and does one present differently than the other?
- Knowing that the artist had undergone a massive growth period between her two works, how frequently does the artist use the word "love" or "lovers" and themes of youth in both her albums?

## Dataset Description

The dataset that I will be performing my exploratory data analysis on contains all of Lorde's lyrics from both her Pure Heroine and Melodrama albums. All the lyrics are scraped from genius.com using their API

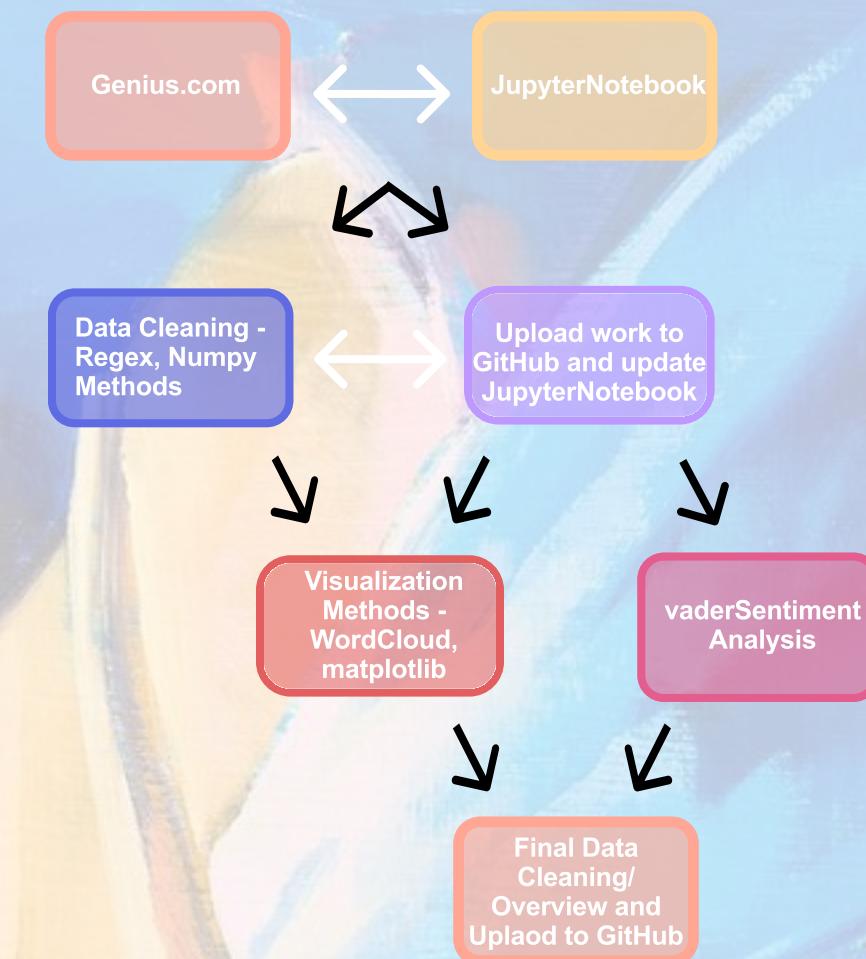
1. Obtain lyrics using Genius.com API in order to import into Jupyter notebook
2. Write a function with the use of pandas clean the dataset of extra descriptive words like "[Verse]"
3. Create visualizations -- WordCloud to analyze most commonly used words in her albums, barcharts to display words like "love" "lovers" and their frequency in the songs
4. Also delve into the vaderSENTIMENT analysis and build a classification model to gauge positivity vs. negativity for each song

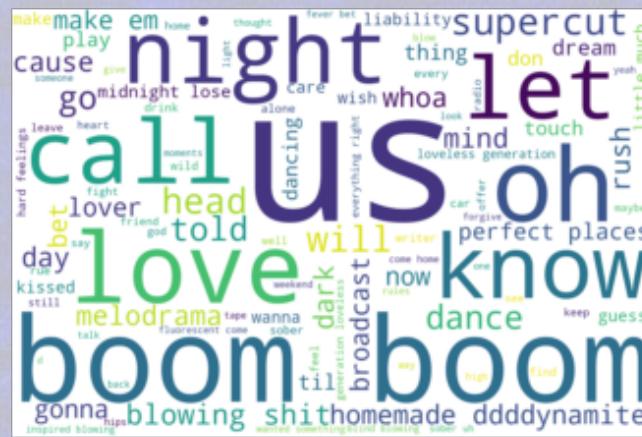
## Hypotheses

- I believe that one of the most common words in Melodrama will for sure be love as Lorde as this album is considered one following her heartbreak from being broken up with
- For Pure Heroine I think some of the more prevalent words will be "never" and "talk" -- this album is about the teenage experience and teenagers gossip and tend to make one another feel excluded based upon superficial notions
- I do believe that Melodrama will perceive as more positive than Pure Heroine -- in exploring her negative emotions with the relationship, there must also come a positive outlook for the future
- Finally, I think that Pure Heroine will contain more instances of youth and less of love -- vice versa for Melodrama

	album	title	lyrics
1	Pure Heroine	400 Lux	we're never done with killing time can i kill ...
2	Pure Heroine	Royals	i've never seen a diamond in the flesh i cut m...
3	Pure Heroine	Ribs	the drink you spilt all over me lover's spit l...
4	Pure Heroine	Buzzcut Season	i remember when your head caught flame it kiss...
5	Pure Heroine	Team	wait till you're announced we've not yet lost ...
6	Pure Heroine	Glory and Gore	there's a humming in the restless summer air a...
7	Pure Heroine	Still Sane	today is my birthday and i'm riding high hair ...
8	Pure Heroine	White Teeth Teens	we wouldn't be seen dead here in the day i gue...
9	Pure Heroine	A World Alone	that slow burn wait while it gets dark bruisin...

## Work Flow



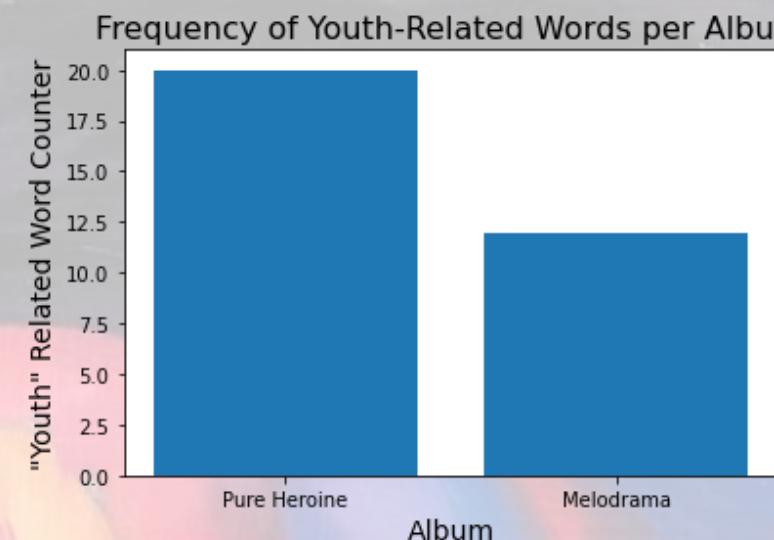
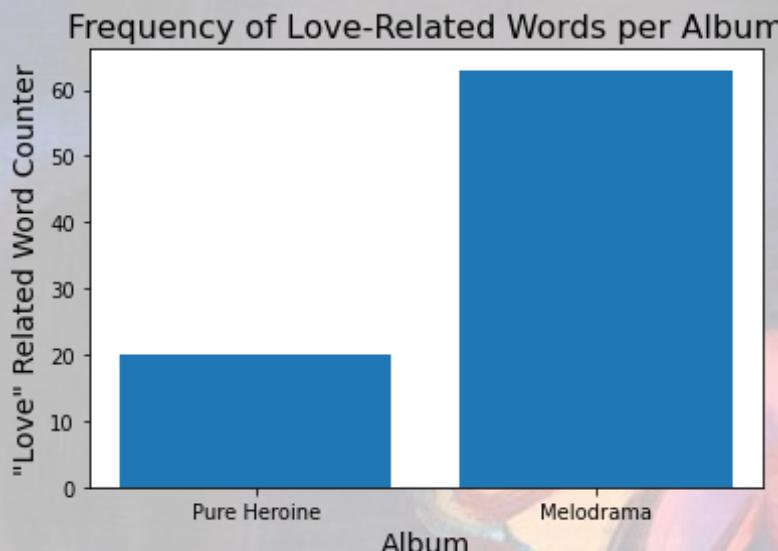


## Interpretation of Results + Limitations

When Lorde wrote Pure Heroine she was 16 -- what I would claim as the peak of teen adolescence and angst. I believe that the high prevalence of the words "people" and "talking" touches on the youth culture. How people value materialistic desires, gossip about social status; how teens experience heartbreak and anxiety for the first time and Lorde is simply delving into the plethora of the human experience .. a suburban teenage fantasy. Frankly the idea of youth is just omnipresent in Pure Heroine as the barchart below confirms

Furthermore, I believe that the most stark difference in the two word clouds is how prevalent the word love appears in Melodrama. In between these two albums Lorde went through a serious break-up and started phasing into early adulthood which dramatically influenced her album's chemistry. For words like "love", cuss words, and "night" to hold as much space in the album, calls to attention the rage, somber moments, and melancholy that is bottled up in this work. This idea is further confirmed by the barcharts, for the use of love-related words skyrocketed in Melodrama indicating a flurry of emotions regarding love, loss, and moving forward.

Interestingly enough, the sentiment analysis concluded that *Melodrama* was a lyrically happier album than *Pure Heroine*. Maybe the angst of our teenage years superseeds that of scorned love or the vaderSentiment analyzer unable to dissect metaphors and unclear speech. For instance, the lyric "I love it here since I've stopped needing you" -- the word love indicates something positive however I believe it's negative as the feeling of love arises only once she moves on from the toxic presence in her life. That proves a limitation as the subjectivity of song lyrics to both the artist and the listener may be interpreted almost too objectively by a computer program.



## Description of Techniques Employed

## 1. Word Cloud

To Answer my Primary EDA Question, I created a Word Cloud that creates a picture with the most prevalent words from each album. Bi-gram and tri-grams -- aka a sequence of two or three adjacent elements from a string like “wireless speaker” -- are in place as well. Furthermore, a Word Cloud analyzes the most common words and therefore makes those words a larger size than less commonly used words. In my case I will be capping the number of words in the Word Cloud at n=40

## 2. Text Scraping and Data Manipulation using matplotlib

I employed an API from Genius to scrape the lyrics from the web and then further cleaned the data myself using regex and pandas methods. The majority of clean-up was to do with the new line breaks. Furthermore, I used the library matplotlib to construct bar charts -- allowing me to analyze the frequency of certain words in both albums and draw more conclusions.

### 3. VaderSentiment Analysis

Finally, I employed vader sentiment analysis to see if her songs hold an overall positive, negative, or neutral emotion in the eyes of an analyzer. This way I will be able to more objectively compare if Pure Heroine or Melodrama proves more positive or negative

```
Overall sentiment dictionary is : {'neg': 0.055, 'neu': 0.674, 'pos': 0.271989}
sentence was rated as 5.5 % Negative
sentence was rated as 67.4 % Neutral
sentence was rated as 27.1 % Positive
Sentence Overall Rated As Positive
```

```
Overall sentiment dictionary is : {'neg': 0.124, 'neu': 0.756, 'pos': 0.1212926}
sentence was rated as 12.4 % Negative
sentence was rated as 75.6 % Neutral
sentence was rated as 12.1 % Positive
Sentence Overall Rated As Negative
```

**Works Cited + GitHub Repo**

- <https://github.com/sumati-w/sumati-DH100>
  - <https://genius.com/albums/Lorde/Pure-heroine>
  - <https://genius.com/albums/Lorde/Melodrama>
  - [youtube video showing me how to use API](#)
  - [why use doc2vec](#)
  - [Professor Anderson's Github](#)
  - [Genius API](#)
  - [past work references](#)

- where I got data visualization pictures
  - lesson on sentiment analysis
  - what is a bigram?
  - why use a word cloud?
  - sentiment analysis using VADER