DATA ANALYSIS OF AUTOTUNE TREND
IN HIP-HOP SONGS FROM THE PERIOD 2009-2020

STUDENT NAMES: Leon Tjandra, Mihir Chhiber

STUDENT NUMBER: 1004353, 1004359

COURSE NAME: Introduction to Digital Humanities

PILLAR NAME: Information Systems Technology and Design (ISTD)

COURSE CODE: 02.137DH

TABLE OF CONTENTS

INTRODUCTION & BACKGROUND

Overview

The main topic we are focusing on in this research is the trend of autotune usage in popular hip-hop songs from the year 2009-2020, where this phenomenon has been a breakthrough in the music industry that invites both praise and backlash from the community. The songs selected here are limited to the hip-hop genre since the usage of the autotune software is most prominent in this genre and the fact that hip hop stands as the mainstream of the music industry in the United States starting from the 2000s.

The release of autotune brought a new dimension in the world of hip hop music, allowing flawless and exquisite music performance due to the sophisticated algorithm behind it, however as the number of artists who worshipped autotune increased rapidly, there is an increasing concern of overreliance towards the software, degrading the pristine sound generated from both the vocal and instrument. Anti-autotune movements began to emerge since the richness and texture of the music is stripped off in an autotuned song, however, the success and popularity of the software allows it to remain in demand since some musicians prioritize creating a perfect and flawless performance more than originality and ethics. The autotune trend is believed to experience a rise from its birth in 1996 up until 2009, before the start of its downfall in the same year. The research revolves around the extraction of audio elements focused on the frequency of a song, to help distinguish between autotuned and non-autotuned songs to analyze the trends over the period and whether autotune increases the homogeneity amongst songs.

Aims and Objectives

The research aims to provide an argument on the trend of autotune usage in hip hop and the music industry from 2009 to 2020 which was expected to be the year of the downfall of autotune due to criticism received from musicians and newspaper reporters. There was a downwards trend expected after a major backlash in 2009 but there is not as much information about the trends of usage of autotune over the period 2010-2020. Answering a variety of questions of interest related to the main topic which includes finding the usage trend of autotune in hip hop genre and the factors affecting it over the period of 2009-2020 and whether autotune has increased the homogeneity of the songs in the hip-hop genre. The results intend to provide arguments on the general belief of the autotune trend from 2009-2020 to support or contest the claim based on the audio information extracted and analyzed statistically with elements of interpretation and traditional humanities research.

Research Question

Questions of interest relies on audio information extracted from the dataset of hip hop songs from the years 2009-2020 involving interpretation:

- How does the trend of autotune usage in hip hop music change from the year 2009 until 2020?
- Did autotune experience a downward trend starting from 2009 until 2020?
- What are the factors that may affect the trend of autotune usage in hip hop songs from 2009 until 2020?
- Has Autotune increased the homogeneity of the songs in the hip-hop genre?

LITERATURE REVIEW

The idea of autotune was initiated by Dr. Andy Hildebard who worked in the oil industry who pioneered a breakthrough in his field of expertise by developing an algorithm that utilizes seismic waves to create accurate subsurface maps of potential oil drilling sites. This algorithm is better known as autocorrelation and further development allowed it to perform pitch correction for music since it is highly related to sound waves (Chris Perez, 2012). Auto-correlation was later presented at the NAMM (National Association of Music Merchants) conference as an auto-tune software under the company Antares Audio Technology in 1996 (Larry Fitzmaurice, 2018). The auto-tune software sees its first release in 1997. The first song to utilize this software was *Believe*, a single by Cher and soon the technology spread rapidly in the music industry which allows a more flawless track to be recorded (Chris Perez, 2012).

Autotune expanded over the years, gaining favors from artists such as Madonna and Radiohead, even when a backlash was expressed by Neil McCormick from the Daily Telegraph in 2004 labeling it as a "sinister invention" (Larry Fitzmaurice, 2018). One major push in the usage of autotune was due to the contribution of singer-songwriter T-Pain who used the software in his debut album *Rappa Ternt Sanga*, and he claimed that the autotune software allowed him to sing uniquely and differently from every other singer and praised the existence of it. Autotune finds glory in the billboard charts due to the efforts of Snoop Dog's *Sexual Eruption* which hits top 10 in the billboard hot 100 in 2007 and Lil Wayne's *Lollipop* that survived for five weeks at the top of the charts in 2008, moreover, Kanye West's *808* and *Heartbreak* shows the artist's interest in using autotune.

The major autotune backlash started in 2009, where the Death Cab for Cutie band attended the Grammy awards with blue ribbons as a sign of protest towards the autotune movement and condemning it as a "digital manipulation of the human voice". Jay-z also clearly expressed his disgust regarding autotune in the form of his single *Death of Autotune* (D.O.A). The concern of autotune usage in 2010 is attributed to the fact that artists used it as a real-time process rather than a fix-in after a performance. MC figures like Future, Chief Keef, and Quavo have blended with autotune effect and intuitively sing flatly to trigger the correction capabilities of the software (Simon Reynolds, 2018). To counter the incoming backlash, T-Pain and Antares tried to keep autotune in power by releasing the "I am T-Pain" app which acquired over 3 million downloads, however, the negative view persisted, and *Time* magazine listed autotune as one of the 50 worst inventions. Autotune is viewed as software which only benefits "bad singers" since those who sing in the correct pitch will experience almost no change to the outcome. In 2010, after the revelation of autotune usage in The X Factor, Simon Cowell banned the usage of the technology for future broadcasts.

Despite the negative views and pressure, autotune remains widely used by musicians. In 2013, Michael Buble use autotune on his single *It's a Beautiful Day* where he claimed that autotune is needed to compete against songs on the pop radio. In an interview in 2013, Thomas Bangalter from Daft Punk stated that autotune "created a musical landscape that is very uniform" (Simon Reynolds, 2018). A campaign initiated by David Mindel in 2014 titled "Live Means Live" tried to promote an autotune-free environment by creating a symbol that indicates no involvement of autotune in concerts, however, it did not gather sufficient support. Lady Gaga expressed her views on autotune through *Perfect Illusion* in 2016, describing it as a fake, altered, filtered take on songs.

The anti-autotune sentiments associated it with many defects which include the lack of talent in music production, the lack of quality, and homogenization of music (Kyle Kramer, 2014). The sound it produces is often dismissed as unnatural, inauthentic, and gimmicky. Autotune changes the way music is produced by allowing subtle pitch correction through changing the adjustment time, it also provides a graphical mode of manual note-by-note correction and an automatic mode that converts sharps and flats to the nearest correct note. Aside from changing the production of music it also changed the way music is heard which threatens to homogenize music to be perfectly in tune.

RESEARCH METHODS AND RESULTS

The first step in the research was to select a reasonable period and dataset size to get valuable results. We decided to observe the period 2009 to 2020 as 2009 was a period where autotune faced criticisms from music artists and magazine reporters as mentioned in the literature review. For each year, we observed the top 20 R&B/Hip-Hop billboard songs. Around 2537 songs are released per year out of which hip-hop constitutes about 20% of the total songs (the 2018 US music database), therefore the number of average hip-hop songs is estimated to be 507 songs per year, and our sample size represents approximately 4% of this number. Even though this constitutes 4% of the total songs released per year, these songs are highly influential and liked by the audience. These songs influence the artists in the music industry as every artist wants to create good music. We extracted the list of top 20 R&B/Hip-Hop songs from the Billboard website for each year. The Billboard charts give a representation of the popularity of songs based on the US population. After extracting the list, we manually downloaded all the songs and stored them in the folder of the respective year. This is the dataset (Dataset 1) we are going to examine for our study.

Autotune is an audio processing tool that alters the pitch in the vocals in a musical recording a performance. Autotune works by fixing the frequency of the vocals by either replacing the audio patch or adding an audio patch to match the frequency rhythm. Some music artists use autotune to alter their vocal frequency to a very high or very low pitch as they cannot attain that pitch with their voice. As autotune focuses on altering the frequency, we decided to analyze the song frequency, mean of frequency, and standard deviation of the frequency of the songs.

To answer our research question, we had to understand the difference between autotuned songs and non-autotuned songs in terms of frequency data and how we can differentiate these songs. We created another dataset (Dataset 2) of 60 songs where 35 songs are autotuned and 25 songs are non-autotuned. We extracted the frequency of these songs using the python library *Essentia* which extracts the frequency of every frame of the song and stores it in a list. After extracting the frequencies of the songs, we calculated the mean and standard deviation of the songs and plotted them on a scatter plot. As we can observe from Graph.1, songs with no autotune are consolidated in the center whereas autotuned songs are spread out throughout the graph. The reason for this is because music artists use autotune to fix their rhythm, reducing the standard deviation or they sing with simulate different voices by changing their vocal pitch which increases the standard deviation.
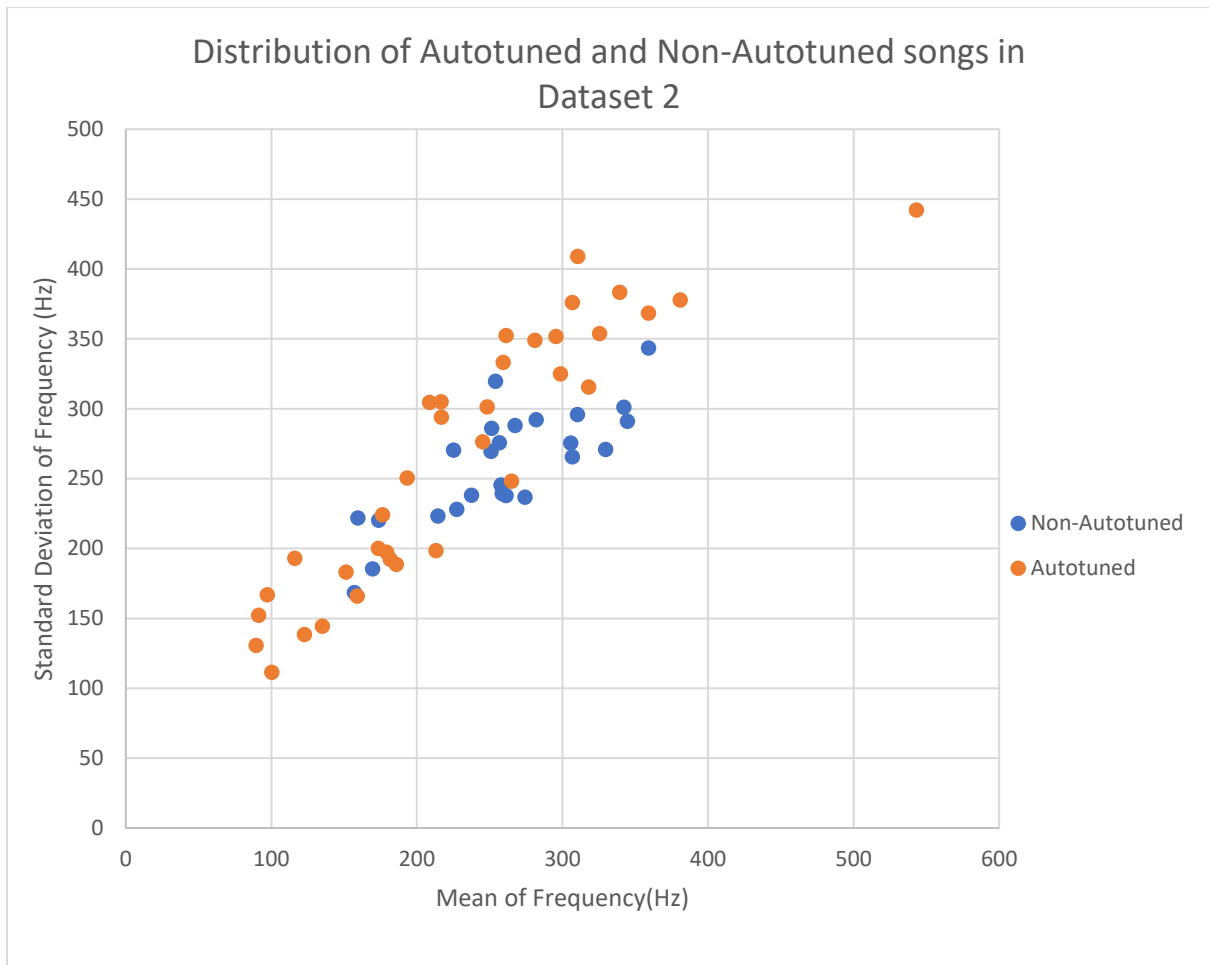
Graph.1 Scatterplot of standard deviation and mean of frequency of autotuned and not autotuned songs

As it is too hard for us to differentiate between autotuned and not autotuned songs, we created a random forest machine learning model that takes the standard deviation of songs as the input and predicts whether a song is autotuned or not. This model has an accuracy of 80%.

Using dataset 1, the python program extracts the frequency data for each song, computes the standard deviation of the frequencies for each song, and stores it in a csv file. Later we run our random forest model to find out how many songs in each year are autotuned.
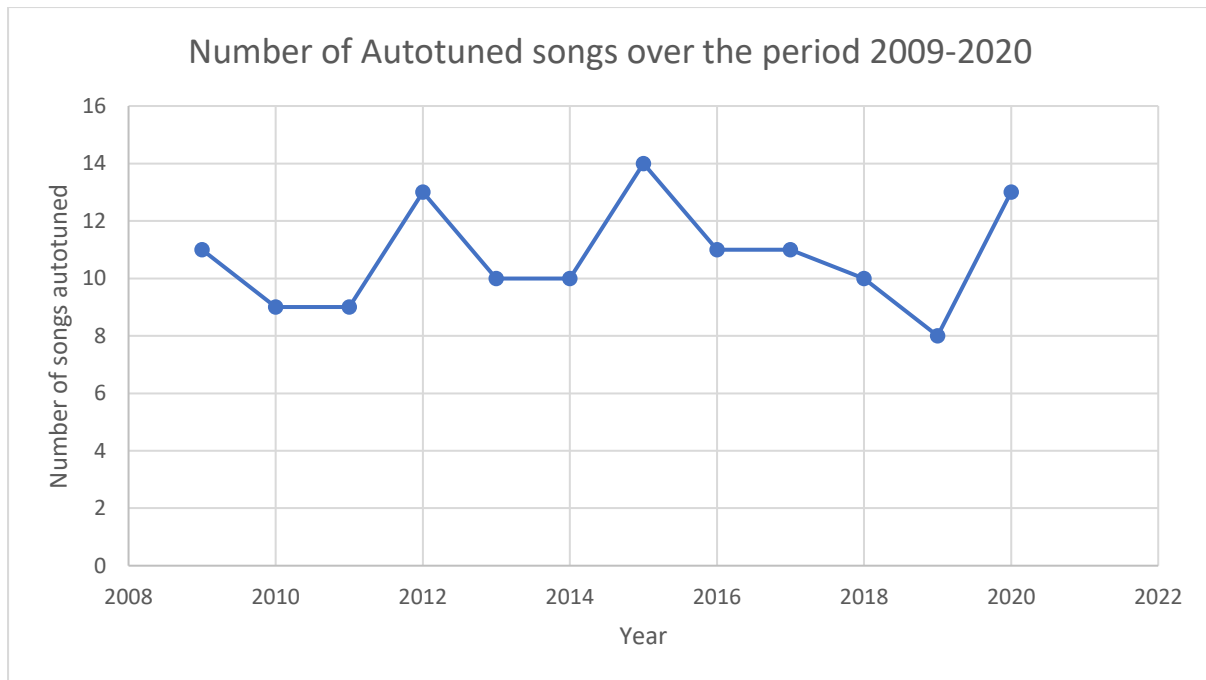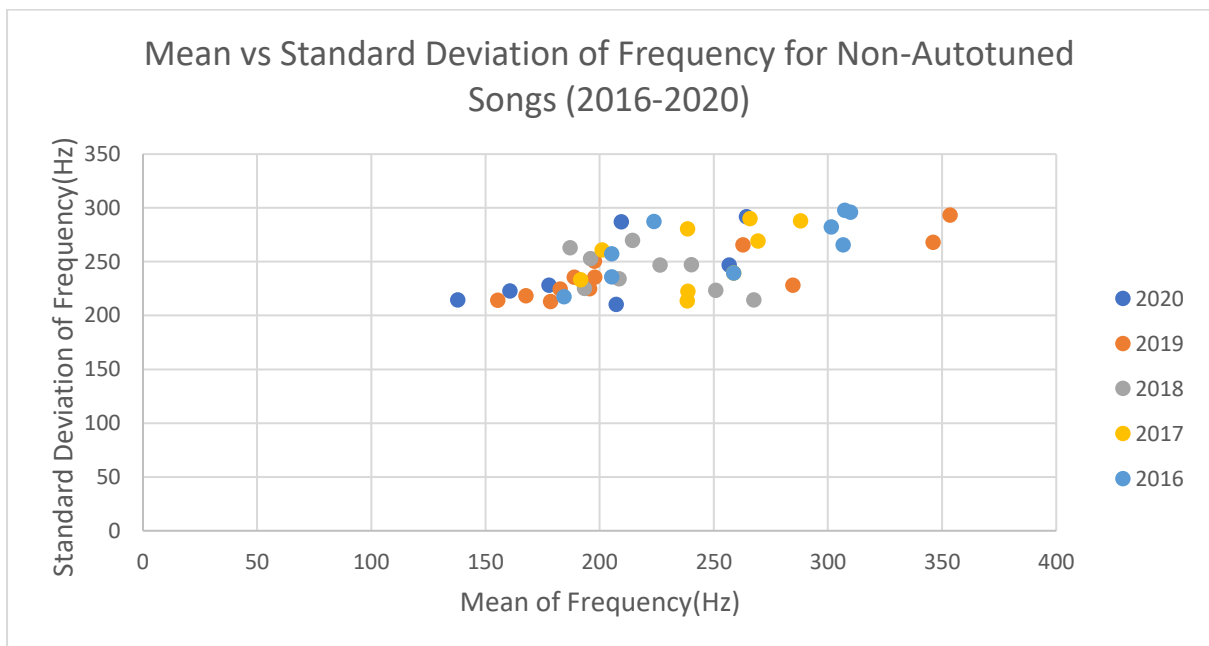
Graph.2 Number of Autotuned songs over the period 2009-2020

As observed from graph 2, autotuned songs did not experience a decrease in popularity from 2009 to 2020 even after the criticism the audio tool received, the graph featured three peaks in autotune usage in 2012, 2015, and 2020 and it experienced a fluctuating trend rather than a downfall. There can be two reasons why autotune is still widely used amongst music artists. The first reason could be the improvement in technology over the period. The most significant change with autotune in the songs released in the last 5 years is autotuned songs sound more polished. There have been many discussions on the recent songs released by artists such as Bruno Mars and Lady Gaga and whether their songs are autotuned (James Jordan, 2014). The second reason could be getting accustomed to autotuned songs. As autotuned songs have been in the industry for 2 decades now, the listeners to the hip-hop genre have got used to the autotune effect in the songs, generating a culture of appreciation towards autotuned songs, making them the new normal in the music industry.
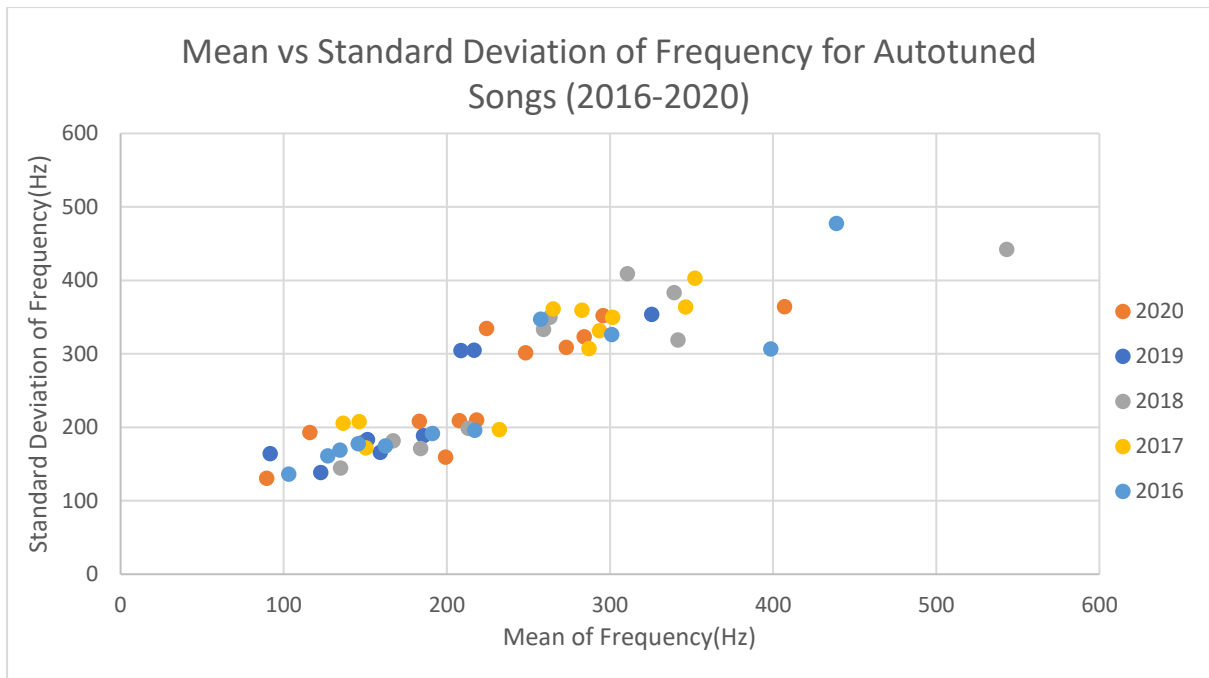
Graph.3 Mean vs Standard Deviation for Autotuned Songs (2016-2020). The songs for each year are spread out evenly and they do not cluster for a specific year.

Graph.4 Mean vs Standard Deviation for Non-Autotuned Songs (2016-2020). The songs for each year are spread out evenly but the songs over the period of 2016 to 2020 are concentrated and are not as spread out as compared to the autotuned songs.
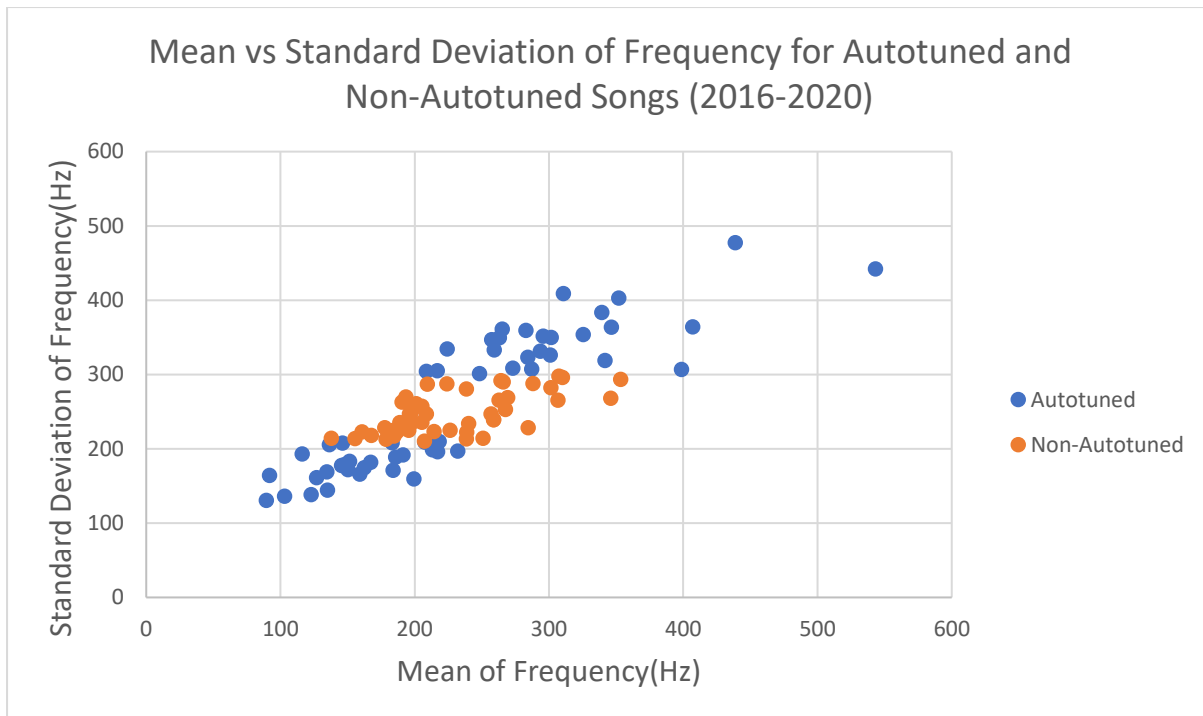
Graph.5 Mean vs Standard Deviation of Frequency for Autotuned and Non-Autotuned Songs (2016-2020)

To study the homogeneity of songs, we plotted the mean and standard deviation of the frequency of autotuned songs and non-autotuned songs over the last 5 years in graph 5. We can observe that non-autotuned songs are more clustered and most of the songs have a mean between 150 to 300. Whereas songs with autotune have means between the value of 100 to 350 and the standard deviation values are either very small or very high due to which the data points are scattered more than the data points in songs which do not have autotune. Therefore, songs that do not make use of autotune are more homogenous than autotuned songs contrary to the belief that autotune contributed to an increase in homogeneity. Autotune allows musicians to achieve sounds they will not be able to achieve with their voice. This allows the musicians to either increase or decrease the pitch of their voice to add variety to the song and gives the musician more options while creating music whereas a musician who does not use autotune has to rely on his voice and the instruments used for creating music.

LIMITATIONS AND FUTURE PROSPECTS

Our research paper makes use of musical elements of the songs to get an insight into the song dataset and analyze the data differently. However, our methodology has its limitations. The songs used to build the dataset are all taken from the billboard charts which is relevant only in the United States but may not represent the whole world, since the country constitutes about 4.2% of the world's population. To represent the whole world, we can create a list of the top 100 hip-hop songs using song sales worldwide.

The random forest model is not a specialized model to identify whether a song is autotuned or not. Our model was able to achieve a high prediction accuracy possibly due to the small size of our dataset. A specialized neural network model would be more accurate and would require a larger dataset to have a good prediction rate. As autotune alters the pitch of the songs, our research thoroughly analyzed the frequency of the songs. Analysis of other musical elements such as rhythm and tempo could give us some interesting insights about the dataset. With music data being generated rapidly and an increase in the performance of computers and machine learning algorithms, there is a huge scope of improvement in such studies being conducted.

CONCLUSION

The research focuses on analyzing how the trend of autotune usage changes over the period from 2009 until 2020, with hip-hop as the genre of interest. The autotune phenomenon is influential in the United States music industry as well as the whole world in general, reshaping the way music is created and perceived, inviting positive and negative reception. The general belief of the autotune which expected a fall from 2009 was contradicted by the usage trend as more than 50% of the top 20 songs used autotuned in the last decade. Autotune has been widely used in the hip/hop genre in the last 2 decades which has led to getting accustomed to the sound which was aided by the technological improvements of the audio tool. Autotuned was blamed for increasing the homogeneity of the songs but autotuned songs seemed to have more varying frequency compared to non-autotuned songs contradicting the general view.

By conducting this research project, we were able to bridge the gap between humanities and computational tools with the help of digital humanities tools. We made use of the programming language Python with the help of library *Essentia* and *Pandas* to extract the frequency of songs, compute mean and standard deviation of frequency and store them in an excel file. Later we made use of scatter plot and line plot on excel to analyze the data extracted by the programming. We wrote the possible hypothesis to answer the humanities question supported by the analysis of the trends in the data. By applying the technology with humanities research and interpretation we were able to understand a cultural phenomenon which is autotuning.

References

Larry Fitzmaurice (2018). Great Moments in Auto-Tune History https://www.vulture.com/2018/12/great-moments-in-auto-tune-history.html (accessed 9th November 2020).

Simon Reynolds (2018). How Auto-Tune Revolutionized the Sound of Popular Music https://pitchfork.com/features/article/how-auto-tune-revolutionized-the-sound-of-popular-music/ (accessed 9th November 2020).

Kyle Kramer (2014). The T-Pain Effect: How Auto-Tune Ruined Music and Saved Hip-Hop https://www.complex.com/music/2014/04/the-t-pain-efffect-how-auto-tune-ruined-music-and-saved-hip-hop (accessed 9th November 2020).

Adam R. Gold (2011). Death of Auto-tune https://www.thecrimson.com/article/2011/3/11/autotune-pitch-pop-sound/ (accessed 9th November 2020).

Wikipedia (2020). Billboard Hot 100 History https://en.wikipedia.org/wiki/Billboard_Hot_100#History (accessed 9th November 2020).

Wikipedia (2020). Category:2011songsd https://en.wikipedia.org/wiki/Category:2011_songs

Jennifer Max (2019). 13 Music Genres That Every Music Lover Should Know https://www.soundmaximum.com/music-genre-statistics/ (accessed 9th November 2020).

Amy Watson (2020). Music album consumption U.S. 2018, by genre https://www.statista.com/statistics/310746/share-music-album-sales-us-genre/ (accessed 9th November 2020).

Data Commons (2019). United States https://datacommons.org/place/country/USA (accessed 9th November 2020).

Jordan James (2014). Top 10 Singers who don't need Autotune https://watchmojo.com/suggest/Top+10+singers+who+don%27t+need+Autotune (accessed 9th November 2020).

Chris Perez (2012). A Quick History of Auto-Tune https://www.apartmenttherapy.com/the-history-of-auto-tune-178567 (accessed 9th November 2020).

Dan Fletcher (2010). The 50 Worst Inventions https://web.archive.org/web/20110515130413/http://www.time.com/time/specials/packages/article/0,28804,1991915_1991909_1991903,00.html (accessed 9th November 2020).

Tom Scott (2019). How Auto-Tune Works https://www.youtube.com/watch?v=i09V6SkVkz8 (accessed 9th November 2020).

Parul Pandey (2018). Music Genre Classification with Python https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8 (accessed 10th November 2020).

Allen Downey (2015). Basic Sound Processing in Python https://www.youtube.com/watch?v=0ALKGR0I5MA (accessed 10th November 2020).

Essentia (2013). Documentation https://essentia.upf.edu/ (accessed 10th November 2020).