An Evaluation of Benford's Law in Music

Sybil Prince Nelson

Washington and Lee University

12.02.2021

Full Outline

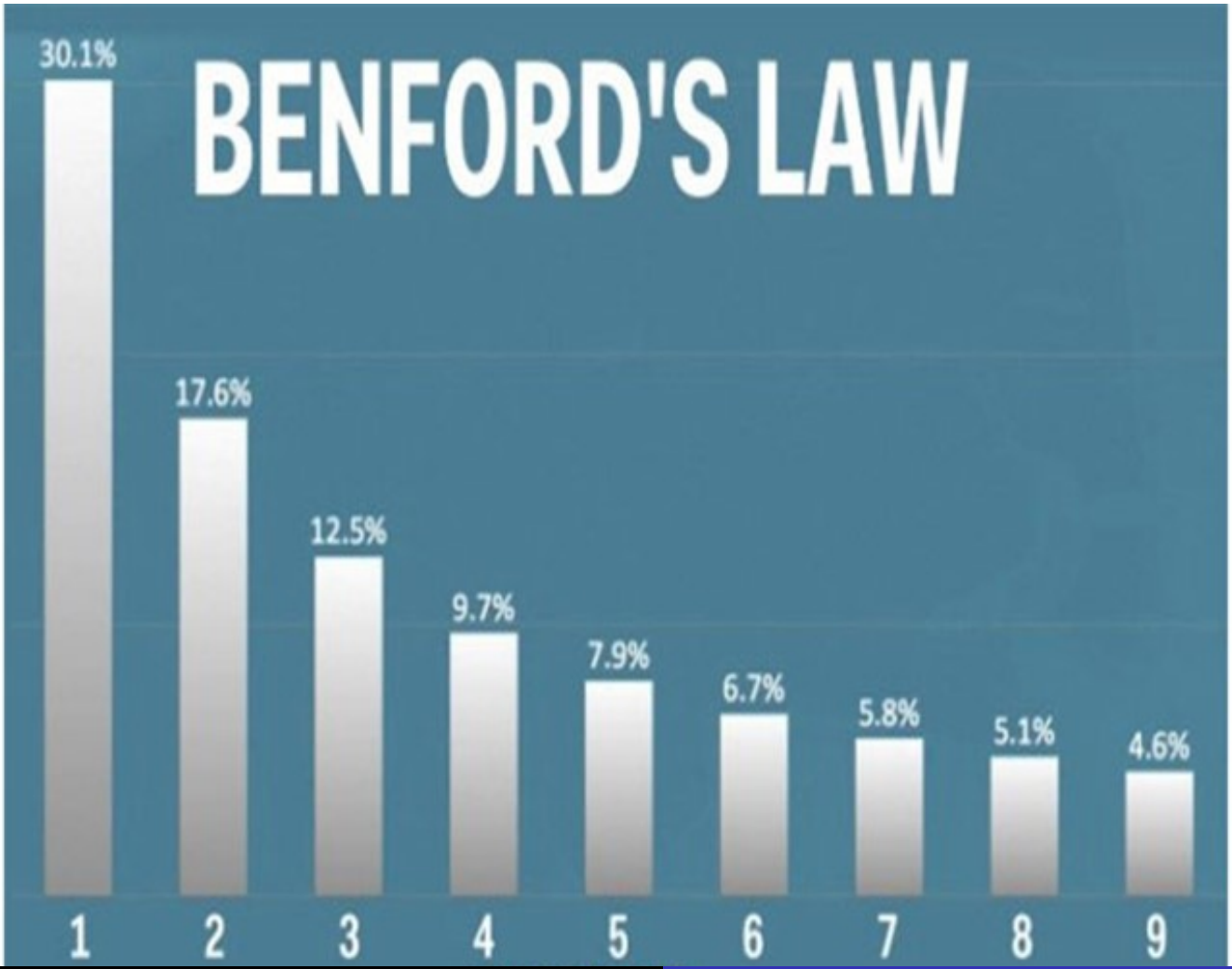
Outline

The Law of First Digits

Benford's Law, also known as the First Digit Law, states that in most naturally occurring data sets, about 30% of the numbers will start with a 1, about 17.5% will start with a 2, and so on logarithmically [1] adhering to the formula $Prob(d) = \log_{10}(1 + \frac{1}{d})$ where $d = 1, 2, \ldots 9$.

[1] Benford F. - *"The Law of Anomalous Numbers." Proceedings of the American Philosophical Society volume 78, number 4 (1938): 551-72.*

The Law of First Digits



The Law of First Digits

The law was proven theoretically by Theodore Hill in 1995 and since then applications have been found in the fields of economics, sociology, physics, computer science and biology among others.[1]

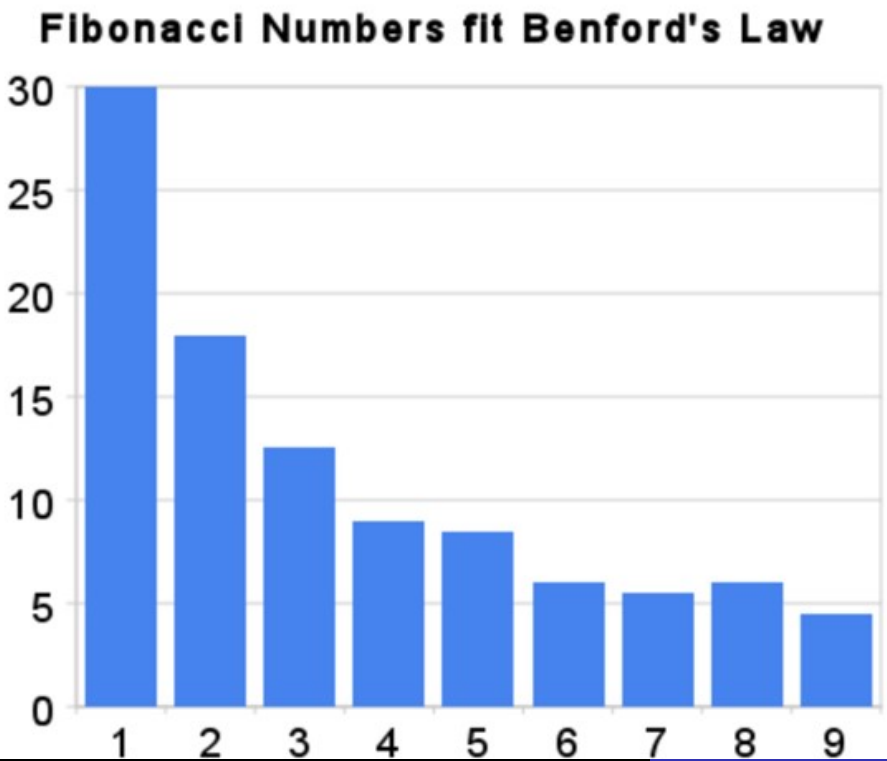
[1] Li et al - *"Application of Benford's Law in Data Analysis." Journal of Physics: Conference Series, volume 1168, (2019)*

Examples of Benford's Law

Believe it or not, the Fibonacci numbers are Benford distributed.
The n^{th} number of the Fibonacci sequence is given by
$f_n = \frac{(\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n}{\sqrt{5}}$

Examples of Benford's Law

Looking at the first digits of the first 200 Fibonacci numbers gives the following distribution.



Applications of Benford's law

- Benford's Law gives an analysis method that helps alert CPAs and officials to possible errors, potential fraud, manipulation, or other irregularities[1] Benford's Law can be used to detect suspicious or fraudulent activity online and to validate datasets[2]

[1] *Journal of Accountancy*
[2] *Benford's Law Applies to Social Media Accounts, PlusOne*

Outline

A Little About Music

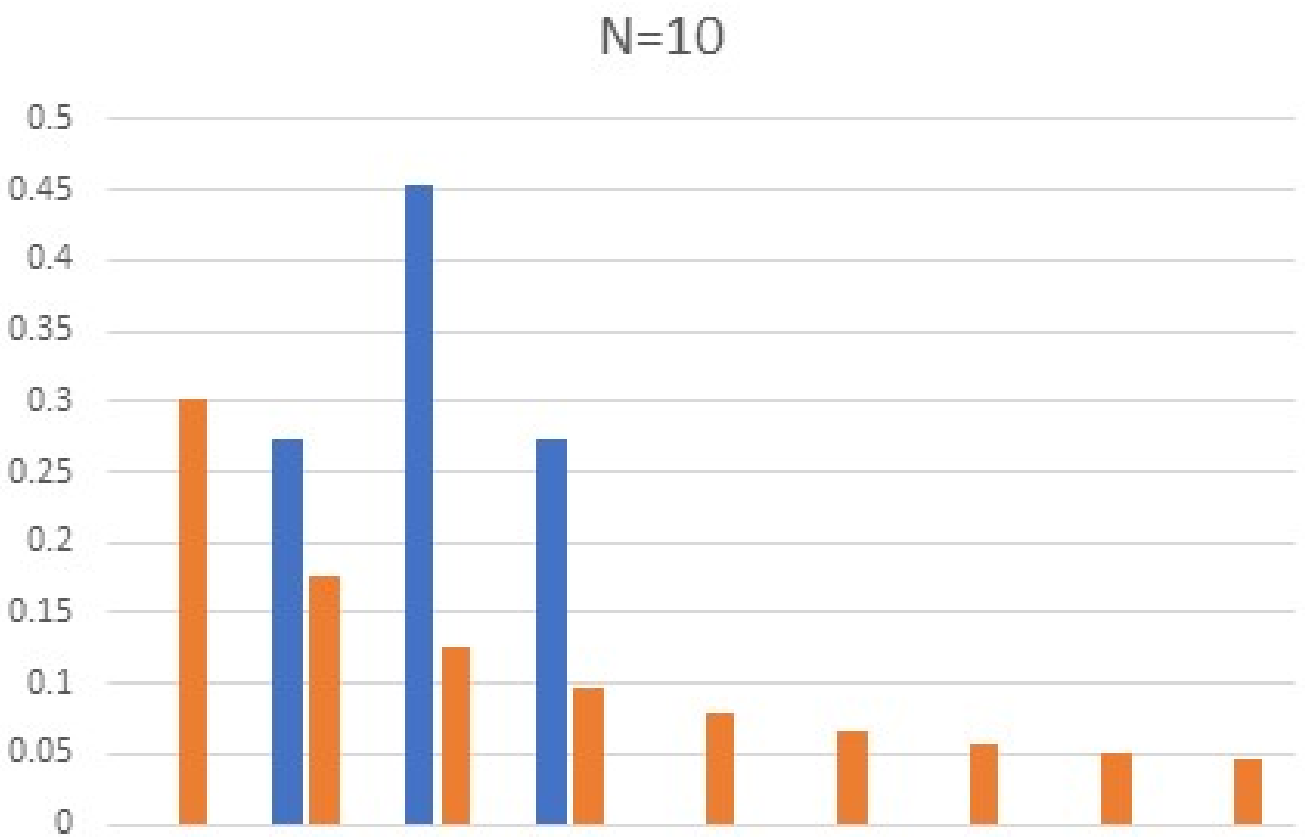
For this project, we considered note durations and note frequencies of classical piano music. Piano note frequencies in hertz can be found with the formula $f_n = 440 \cdot (2)^{\frac{n-49}{12}}$ where n corresponds to the n^{th} key on the piano. The A above middle C, known as tuning A, is the 49^{th} key thus it has a wavelength of 440 Hz. Middle C is the 40^{th} key and has a frequency of 261.63 (Hz). Recall that a wavelength is the distance between successive crests of a sound wave. One wave per second is called a Hertz (Hz)

Piano note frequencies

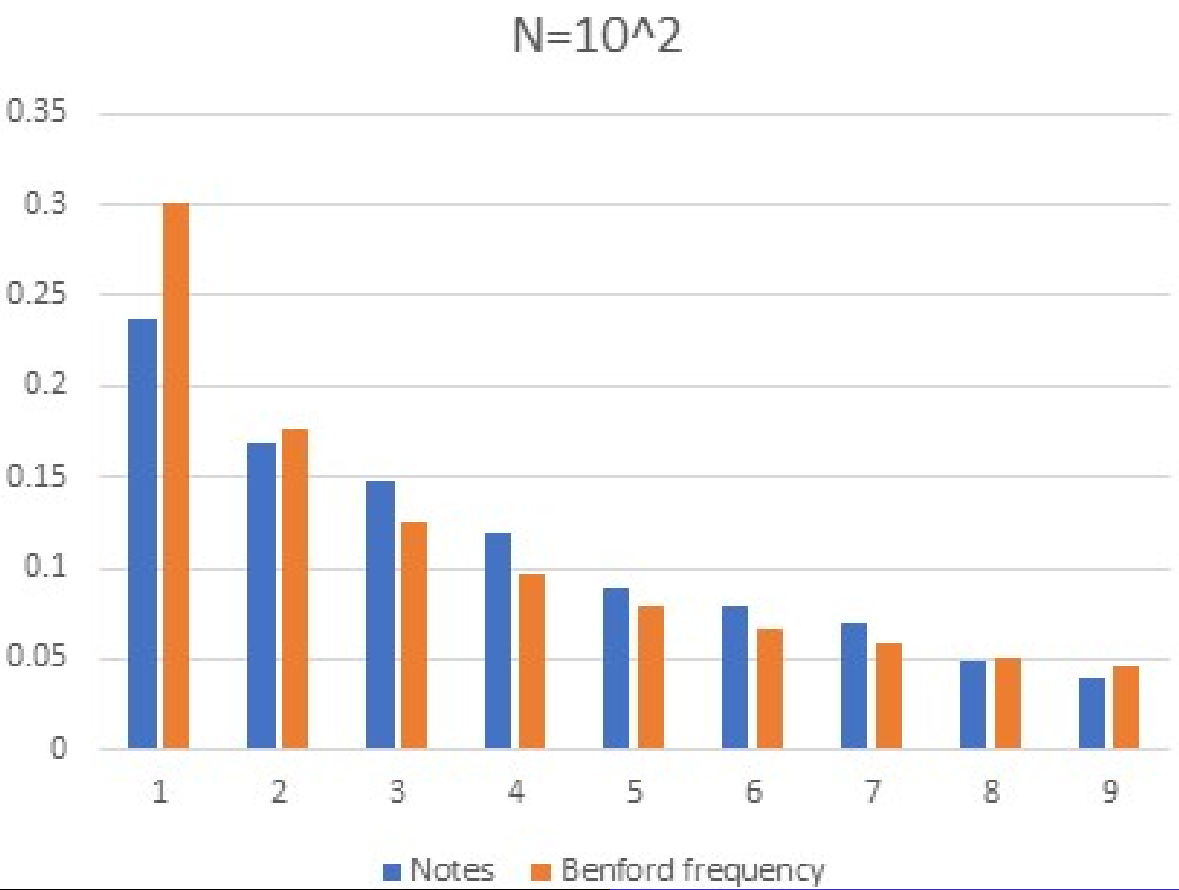
To assess how well a set of numbers fits the Benford distribution, we used the following delta formula.

$$\Delta = 100 \cdot (max)_{d=1}^{9} |Prob(D_1 = d) - \log_{10}(1 + 1/d)|$$

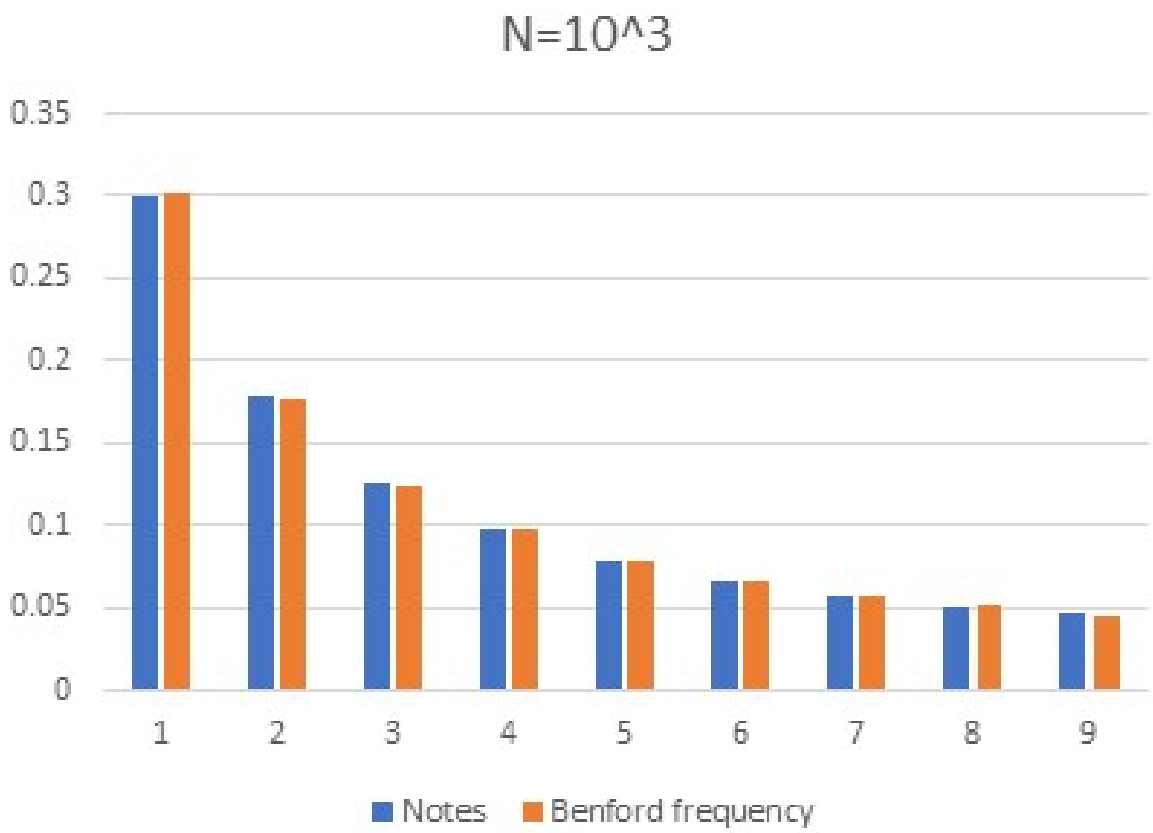
Piano Notes



Piano Notes


N=10^2
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
4
5
6
7
8
9
Notes
Benford frequency

Piano Notes


N=10^3
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
4
5
6
7
8
9
Notes
Benford frequency

Piano Notes

So, we can see that for every natural number m, and $d_1 \in 1, 2, ..., 9$ and all $d_j \in 1, 2, ..., 9, j \geq 2$

$$\lim_{N \to \infty} \frac{total\#1 \leq n \leq N : D_j(f_n) = d_j for j = 1, 2, ..., 9}{N}$$

$$= \log(1 + (\sum_{j=1}^{m} 10^{m-j} d_j)^{-1})$$

Piano Notes

N	P(d=1)	P(d=2)	P(d=3)	P(d=4)	P(d=5)	P(d=6)	P(d=7)	P(d=8)	P(d=9)	Δ
10	0	0.2	0.5	0.3	0	0	0	0	0	37.51
100	0.24	0.16	0.15	0.12	0.09	0.08	0.07	0.05	0.04	6.11
1000	0.3	0.178	0.125	0.097	0.079	0.066	0.058	0.051	0.046	0.19
10000	0.301	0.1756	0.1251	0.097	0.0791	0.0671	0.058	0.0511	0.0459	0.05

Outline

Data Collection

- We used www.kunstderfuge.com to download songs
- In total, we have 320 songs from the Medieval Ages (c. 800-1400) through the Post-War era (1920-Present)
- We used R to obtain the time duration each note was played and the wavelengths of the notes in a given song
- Then, we extracted the leading digit for the cumulative number of seconds each note was played as well as cumulative frequency for every note

Outline

Sample Data

Cumulative Note Length Durations for Mendelssohn's Symphony No. 3

German Note	*Total Length*	*German Note*	*Total Length*	*German Note*	*Total Length*	*German Note*	*Total Length*
C,	0.06	A	45.53	f'	23.94	c#'''	1.92
D,	0.33	A#	4.30	f#'	25.89	d'''	9.33
D#,	0.19	B	18.10	g'	32.73	d#'''	4.96
E,	1.16	c	14.06	g#'	18.23	e'''	14.49
F,	0.76	c#	6.57	a'	81.30	f'''	4.36
F#,	0.15	d	13.69	a#'	9.82	f#'''	3.91
G,	1.38	d#	11.25	b'	38.77	g'''	2.99
G#,	0.84	e	72.14	c''	37.45	g#'''	0.03
A,	13.49	f	11.17	c#''	9.13	a'''	1.86
A#,	0.79	f#	8.80	d''	35.11		
B,	5.02	g	21.86	d#''	19.28		
C	6.56	g#	14.02	e''	70.41		
C#	1.62	a	59.40	f''	16.05		
D	7.79	a#	7.64	f#''	21.83		
D#	3.53	b	50.19	g''	25.76		
E	17.19	c'	55.53	g#''	12.68		
F	6.16	c#'	6.86	a''	36.15		
F#	3.30	d'	44.70	a#''	4.29		

Outline

Music History

- Music history can be divided into six major categories: Medieval, Renaissance, Baroque, Classical, Romantic and Post War
- We analyzed a large collection of music from each time period and found a clear mathematical connection

Music History

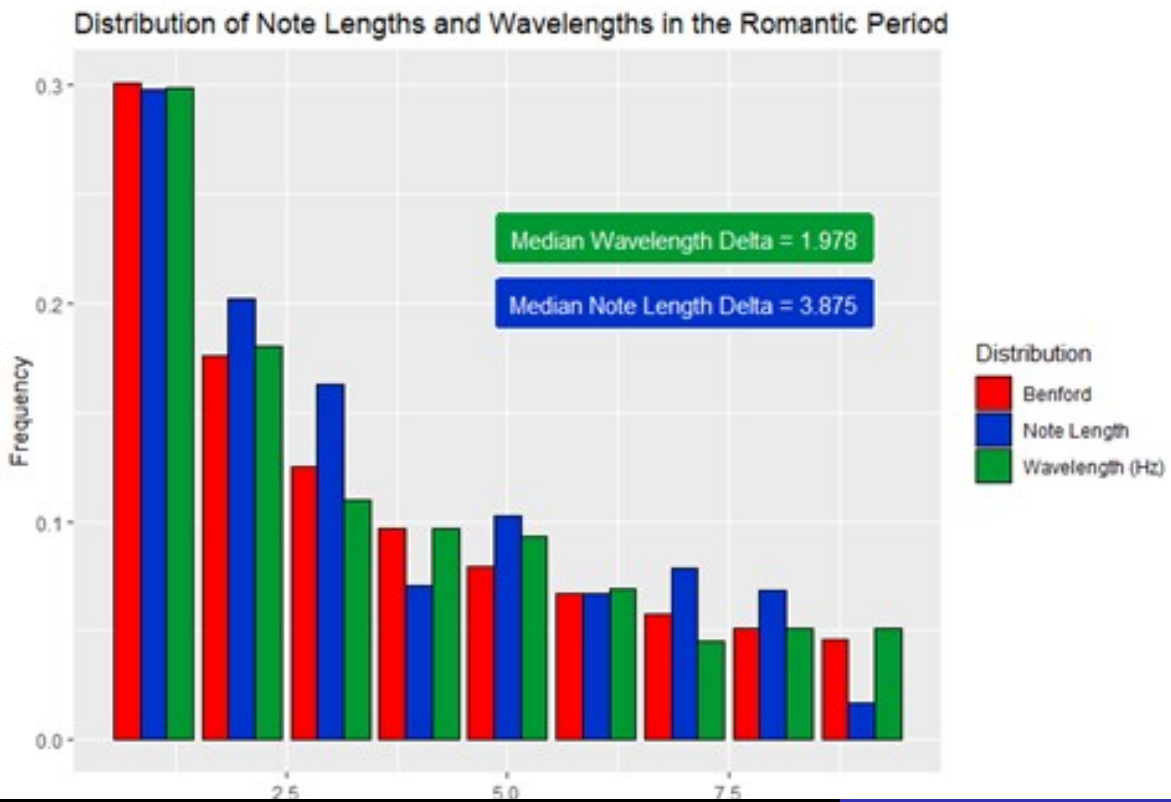
- We found delta values for each song and then compared their distribution between time periods
- An ANOVA test shows that there is a significant difference between time periods for both note length and wavelength

Sample Data

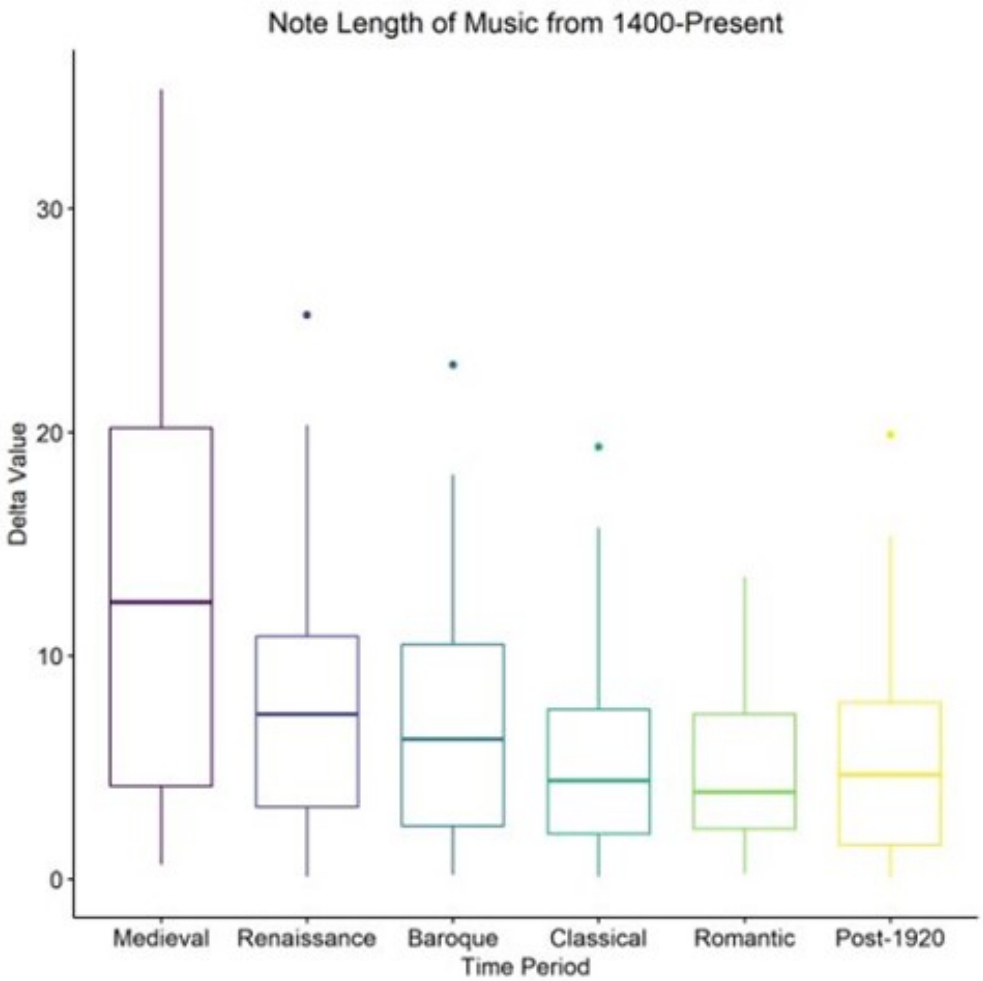
Descriptive Statistics

Time Period	*Years*	*Number of Songs*	*Total Notes Analyzed*	*Composers*
Medieval	1150-1400	31	12,276	Unknown*
Renaissance	1400-1600	72	58,200	King Henry VII, William Byrd, Turlough O'Carolan
Baroque	1600-1750	89	152,600	J.S. Bach, Vivaldi, Monteverdi
Classical	1750-1830	44	176,000	Beethoven, Mozart, Chopin
Romantic	1830-1920	29	207,000	Tchaikovsky, Mendelssohn, Schumann
Post War	1920-Present	55	217,200	Medtner, Prokofiev, Satie

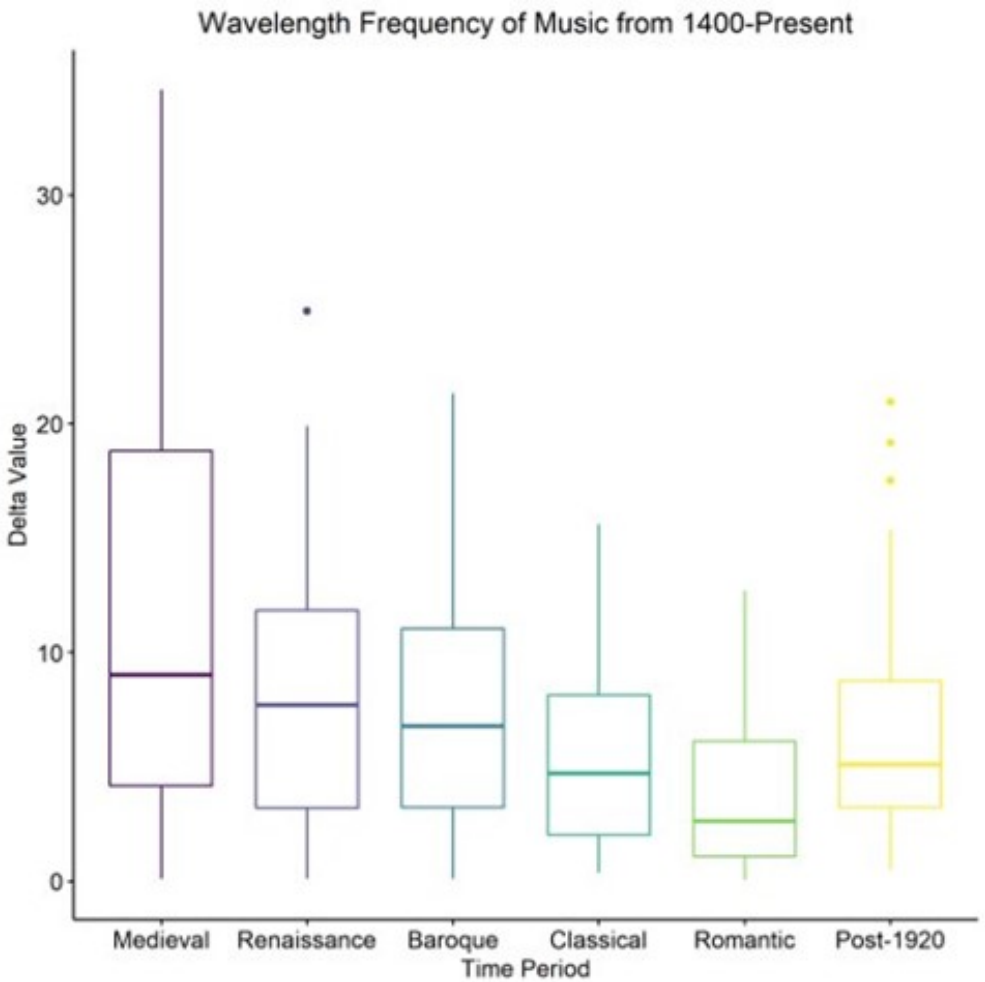
Results


Distribution of Note Lengths and Wavelengths in the Romantic Period
Median Wavelength Delta = 1.978
Median Note Length Delta = 3.875
Frequency
0.3
0.2
0.1
0.0
25
50
75
Distribution
Benford
Note Length
Wavelength (Hz)

Results


Note Length of Music from 1400-Present
Delta Value
30
20
10
0
Medieval
Renaissance
Baroque
Classical
Romantic
Post-1920
Time Period

Results


Wavelength Frequency of Music from 1400-Present
Delta Value
30
20
10
0
Medieval
Renaissance
Baroque
Classical
Romantic
Post-1920
Time Period

Conclusions

- Through history, music adhered closer and closer to the Benford distribution with the exception of the Post-War period.
- There are vast differences in style in the Post-War era so more investigation is needed.
- Future directions for this research include investigating different Benford "scores" of musical genres, popularity of music, and music from different cultures.

Outline

Thank You

Special thanks to Washington and Lee Summer Research Lenfest Grant and my students Brian Wickman, Eric Gazin, and Jack Null.