Journal of Applied Mathematics and Computation, 2021, 5(4), 361-372
https://www.hillpublisher.com/journals/JAMC/
ISSN Online: 2576-0653
ISSN Print: 2576-0645

A Comparative Statistical Analysis of Two Popular Indian Patriotic Songs

Arshi Rahman, Soubhik Chakraborty*

Department of Mathematics, Birla Institute of Technology, Mesra Ranchi-835215, Jharkhand, India.

How to cite this paper: Arshi Rahman, Soubhik Chakraborty. (2021) A Comparative Statistical Analysis of Two Popular Indian Patriotic Songs. *Journal of Applied Mathematics and Computation*, 5(4), 361-372.
DOI: 10.26855/jamc.2021.12.015

Received: November 11, 2021
Accepted: December 8, 2021
Published: December 27, 2021

***Corresponding author:** Soubhik Chakraborty, Department of Mathematics, Birla Institute of Technology, Mesra Ranchi-835215, Jharkhand, India.
Email: soubhikc@yahoo.co.in

Abstract

Pressing the keys of a piano randomly does not create music. *Music is created only when there is a pattern in the musical notes*. This pattern gets embedded in the corresponding musical data, numerical in character. Given that *statistics is the science of exploring and studying patterns in numerical data*, we are motivated to perform a statistical analysis, with special emphasis on modeling wherein lies the strength of statistics, of one of India's patriotic songs '*Aye Mere Watan Ke Logo*'. There is also a comparative study done with India's National Anthem *Jana Gana Mana*. The experimental results are quite interesting. Scatter plot plotted between these two patriotic songs shows that the two songs are entirely different from each other. Apart from modeling, there are other comparative studies done through Information Content and Box plot. For India's National Anthem, note A has the highest Information content of 4.7116. In case of '*Aye Mere Watan Ke Logo*', note E has the highest Information content of 6.0371. The note E has the highest probability for India's National Anthem as compared to the note G for *Aye Mere Watan Ke Logo*. From Box plot we get that India's National Anthem is positively skewed whereas *Aye Mere Watan Ke Logo* is approximately symmetric. But it is quite interesting to note that, despite the aforesaid differences, both the songs can be approximated by the same time series model to a fair degree of accuracy.

Keywords

Time series modeling**,** Single exponential smoothing, Information content, scattered plot, Box plot

1. Introduction

Pressing the keys of a piano randomly does not create music. *Music is created only when there is a pattern in the musical notes*. This pattern gets embedded in the corresponding musical data, numerical in character. Given that *statistics is the science of exploring and studying patterns in numerical data*, we are motivated to perform a statistical analysis, with special emphasis on modeling wherein lies the strength of statistics, of one of India's patriotic songs '*Aye Mere Watan Ke Logo*'.

Aye Mere Watan Ke Logo ("O people of my country") is a Hindi patriotic song written by the lyricist Pradeep, the tune being composed by C. Ramchandra and was sung by the legendary playback singer Lata Mangeshkar. The song commemorates the Indian soldiers who died during the Sino-Indian War in 1962. It is a tribute to the martyrs; the song enabled the young and the old to control and overcome the anger of defeat. This song is often presented at patriotic occasions in India [1]. An interesting single exponential model (also called simple exponential model) is fitted to this song which captures the note progression (pattern). The fit is having a smoothing factor Alpha = 0.6998330 which explains the note progression well enough.

1.1 Modeling

Statistical modeling is one of the most powerful and interesting tools for any statistical analysis. A statistical model captures the pattern in the numerical data. Although approximate, they help in intelligently guessing the true model with fewer parameters (the true complex model may have multiple parameters an explicit idea of which we may not have and we also do not have the knowledge about the exact functional forms in which these parameters enter the model). Of course there is an element of error in such approximation but it is possible to minimize such error (e.g. by using the principle of least squares) which is again a statistical issue. Statistical models are not unbiased either, but it should be kept in mind that (i) we can make the data unbiased (ii) behind the statistical models are elegant mathematical theorems supporting the models and their properties and these theorems are unbiased and (iii) what the statistician expects from the model is what justifies its fitting. Quantitative analysis of music is done using various techniques and here we are using statistical modeling for the analysis of the music. *Although the decision process of a music composer is deterministic (as music is planned and not random) yet from a listener's or an analyst's point of view this deterministic process can be realized as the outcome of a stochastic process that can be depicted by a suitable time series model*. Musical structures are organized and hence there is a pattern. It is this pattern which the statistical model attempts to capture [2].

1.2 Single Exponential Smoothing

Time series is a series if data in a chronological order. Time series analysis can be done by a number of methods. Here, we have taken the musical notes as a time series. Single exponential smoothing model is used for the smoothing purpose of the time series data.

The simplest of the exponentially smoothing methods is naturally called single or simple exponential smoothing. This method is suitable for forecasting data with no clear trend or seasonal pattern.

The basic formula for achieving Simple Exponential Smoothing is:-

$F_{t+1} = \alpha Y_t + (1- \alpha) F_t$, where α (Alpha) is called the smoothing factor and lies between 0 and 1 that is, $0< \alpha <1$. Here, F_{t+1} is the smoothing statistic which is a simple weighted average of the previous observations Y_t and the corresponding previously smoothed (predicted) statistic F_t. Initially $F_0 = Y_0$.

The nomenclature "smoothing factor" for α is somewhat a misnomer since larger values of α actually reduces the level of smoothing and in the limiting case where $\alpha = 1$ the output series is just the same as the original series (with delay of one time unit). Single exponential smoothing is easy to apply and as soon as two observations are available it produces a smoothed statistic. Values of α close to one have lesser smoothing effect and give greater weight to the most recent changes in the data whereas values of α close to zero have greater smoothing effect and are less responsive to the recent changes in the data. There is no particular way of choosing α, it is sometimes the statistician's judgement. Apparently, a statistical technique can be used to optimize the value of α[7]. In our case, α is determined by minimizing the error or residual sum of squares, i.e., by minimizing $\Sigma (Y_t – F_t)^2$ [2-4].

2. Research Methodology

Time series is a series of observations in chronological order. Musical data can also be taken as a time series in which a note characterized by pitch Y_t is the entry corresponding to the argument time t which may mean time of clock in actual performance or just the instance at which phenomenon (in our case the occurrence of the musical note) is realized. Since here we are modeling only the structure of India's patriotic song '*Aye Mere Watan Le Logo*' so the argument will simply be the instances 1, 2, 3….. [2]. The desired note sequence is given in table1 in western notations. Table 2 gives the western notations and the corresponding Indian notations. The tonic (Sa in Indian music) is taken at natural C. Analyzing the structure of a musical piece helps in giving an approximate model that captures the note progression in general without bringing the style of a particular artist into play. It also helps in classification of music from a statistical perspective. In Table 3, pitches of notes in three octaves are represented by corresponding integers, C of the middle octave being assigned the value 0. We are motivated by the works of Adiloglu, Noll and Obermayer [5]. In Indian notations, the letters S, R, G, M, P, D and N stand for Sa, Sudh Re, Sudh Ga, Sudh Ma, Pa, Sudh Dha and Sudh Ni respectively. The letters r, g, m, d, n represent Komal Re, Komal Ga, Tibra Ma, Komal Dha and Komal Ni respectively. The terms Sudh, Komal and Tibra imply, respectively, natural, flat and sharp. In Table 2, the notes without any asterisk sign (*) are in the middle octave. Those with the asterisk sign (*) before them are in the lower octave while the notes in which the asterisk sign (*) are after the notes imply that the notes are in the higher octave. Performance analysis gives additional features like pitch movements between the notes, note duration, inter onset interval, timbre, pitch velocity etc. Readers interested in performance analysis of Indian music are referred to [6].

For the understanding of the patriotic song '*Aye Mere Watan Ke Logo*' with reference to the sequence of notes given in Table 1, the non Hindi reader should refer to Appendix.

Table 1. Note Sequence of "*Aye Mere Watan Ke Logo*"

Instance of realization of note t	Musical Note	(Logarithms of) Pitch
1	D	2
2	G	7
3	G	7
4	G	7
5	G	7
6	A	9
7	G	7
8	F	5
9	F	5
10	F	5
11	D	2
12	G	7
13	G	7
14	G	7
15	G	7
16	G	7
17	A	9
18	G	7
19	F	5
20	F	5
21	D	2
22	F	5
23	A	9
24	A	9
25	A	9
26	Bb	10
27	A	9
28	G	7
29	G	7
30	A	9
31	B	11
32	B	11
33	A	9
34	A	9
35	G	7
36	G	7
37	F	5
38	A	9
39	G	7
40	G	7
41	G	7
42	D	2
43	G	7
44	G	7
45	G	7
46	G	7
47	A	9
48	G	7
49	F	5
50	F	5
51	F	5
52	D	2
53	G	7
54	G	7
55	G	7
56	G	7
57	A	9
58	G	7
59	F	5
60	F	5
61	D	2
62	F	5
63	A	9
64	A	9
65	A	9
66	Bb	10
67	A	9
68	G	7
69	G	7
70	A	9
71	B	11
72	B	11
73	A	9
74	A	9
75	G	7
76	G	7
77	F	5
78	A	9
79	G	7
80	G	7
81	G	7
82	D*	14
83	D*	14
84	C*	12
85	D*	14
86	D*	14
87	C*	12
88	B	11
89	A	9
90	G	7
91	A	9
92	D*	14
93	D*	14
94	D*	14
95	C*	12
96	D*	14
97	E*	16
98	E*	16
99	D*	14

100	C*	12
101	C*	12
102	D*	14
103	C*	12
104	B	11
105	G	7
106	A	9
107	A	9
108	C*	12
109	C*	12
110	B	11
111	A	9
112	G	7
113	B	11
114	A	9
115	A	9
116	G	7
117	A	9
118	A	9
119	Bb	10
120	Bb	10
121	A	9
122	Bb	10
123	A	9
124	F	5
125	A	9
126	G	7
127	G	7
128	D	2
129	G	7
130	G	7
131	G	7
132	G	7
133	A	9
134	G	7
135	F	5
136	F	5
137	D	2
138	G	7
139	G	7
140	G	7
141	G	7
142	A	9
143	G	7
144	F	5
145	F	5
146	D	2
147	F	5
148	A	9
149	A	9

150	A	9
151	Bb	10
152	A	9
153	G	7
154	G	7
155	A	9
156	B	11
157	B	11
158	A	9
159	A	9
160	G	7
161	G	7
162	F	5
163	A	9
164	G	7
165	G	7
166	G	7
167	D	2
168	C	0
169	D	2
170	D	2
171	C	0
172	F	5
173	F	5
174	E	4
175	G	7
176	F	5
177	F	5
178	D	2
179	C	0
180	D	2
181	D	2
182	C	0
183	G	7
184	G	7
185	F	5
186	A	9
187	G	7
188	G	7
189	F	5
190	F	5
191	G	7
192	F	5
193	F	5
194	G	7
195	F	5
196	F	5
197	G	7

Table 2. Representation of musical notes as Indian and Western notes

S	r	R	g	G	M	m	P	d	D	n	N	Indian Notation
C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B	Western Notaion

Table 3. Numbers representing pitches of the musical notes in the three octaves (Western notation)

*C	*Db	*D	*Eb	*E	*F	*F#	*G	*Ab	*A	*Bb	*B	Notes (Lower Octave)
-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	Numbers For (Logarithms of) Pitch
C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B	Notes (Middle Octave)
0	1	2	3	4	5	6	7	8	9	10	11	Numbers For (Logarithms of) Pitch
C*	Db*	D*	Eb*	E*	F*	F#*	G*	Ab*	A*	Bb*	B*	Notes (Upper Octave)
12	13	14	15	16	17	18	19	20	21	22	23	Numbers For (Logarithms of) Pitch

3. Results and Discussion

3.1 Fitting a Single Exponential Smoothing to *Aye Mere Watan Ke Logo*

Single Exponential Smoothing for C1 (Zero values of Y_t exist; MAPE calculated inly for non zero Y_t) Data C1 (C1 in Minitab represents pitch Y_t) is of length 197 with Smoothing constant Alpha, α = 0.698330

These results of Single Exponential Smoothing are obtained using Minitab 16 statistical package.

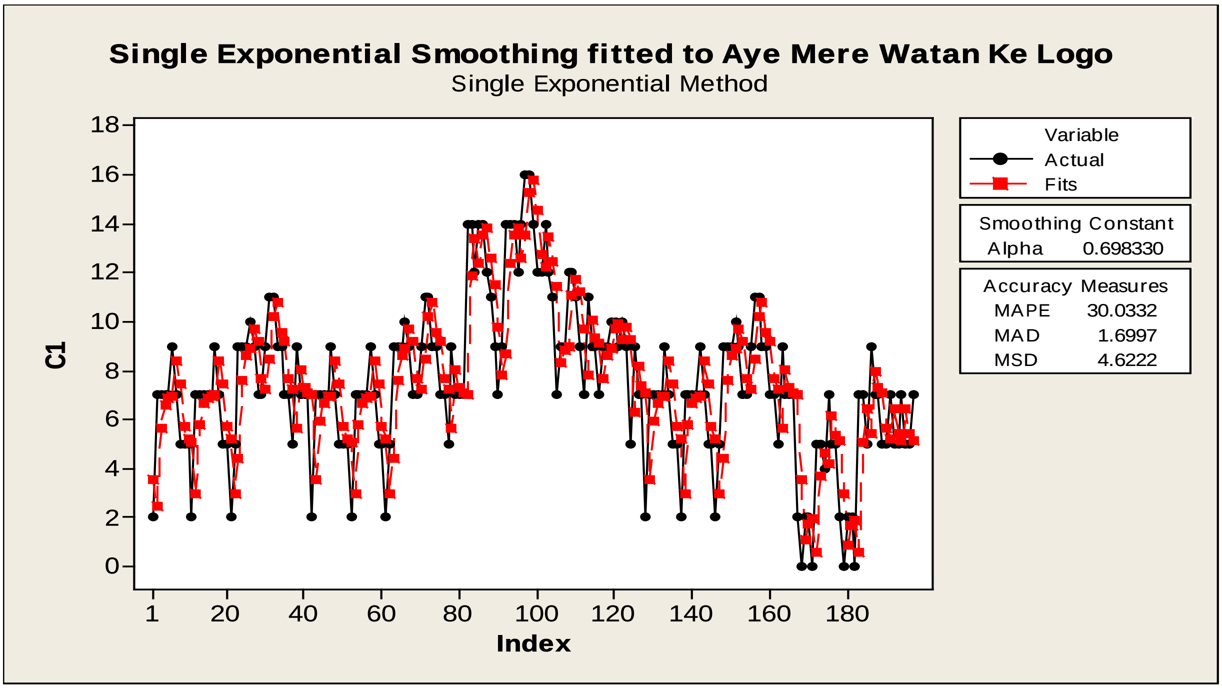

Figure 1. Single Exponential Smoothing plot for C1.

A detailed discussion from the findings from Figure 1 is provided next.

MAPE (Mean Absolute Percent Error) measures the accuracy of fitted time series values. It expresses accuracy as a percentage. MAD (Mean Absolute Deviation) measures the accuracy of fitted time series values. It expresses accuracy in the same units as the data, which helps conceptualize the amount of error. MSD (Mean Squared Deviation) measures the accuracy of fitted time series values. MSD is always computed using the same denominator (the number of forecasts) regardless of the model, so one can compare MSD values across models and hence compare the accuracy of two different models [2].

For all three measures, smaller values generally indicate a better fitting model. So from Figure 1, the accuracy measures are:

MAPE (Mean Absolute Percentage Error) = 30.0332
MAD (Mean Absolute Deviation) = 1.6997
MSD (Mean Squared Deviation) = 4.6222

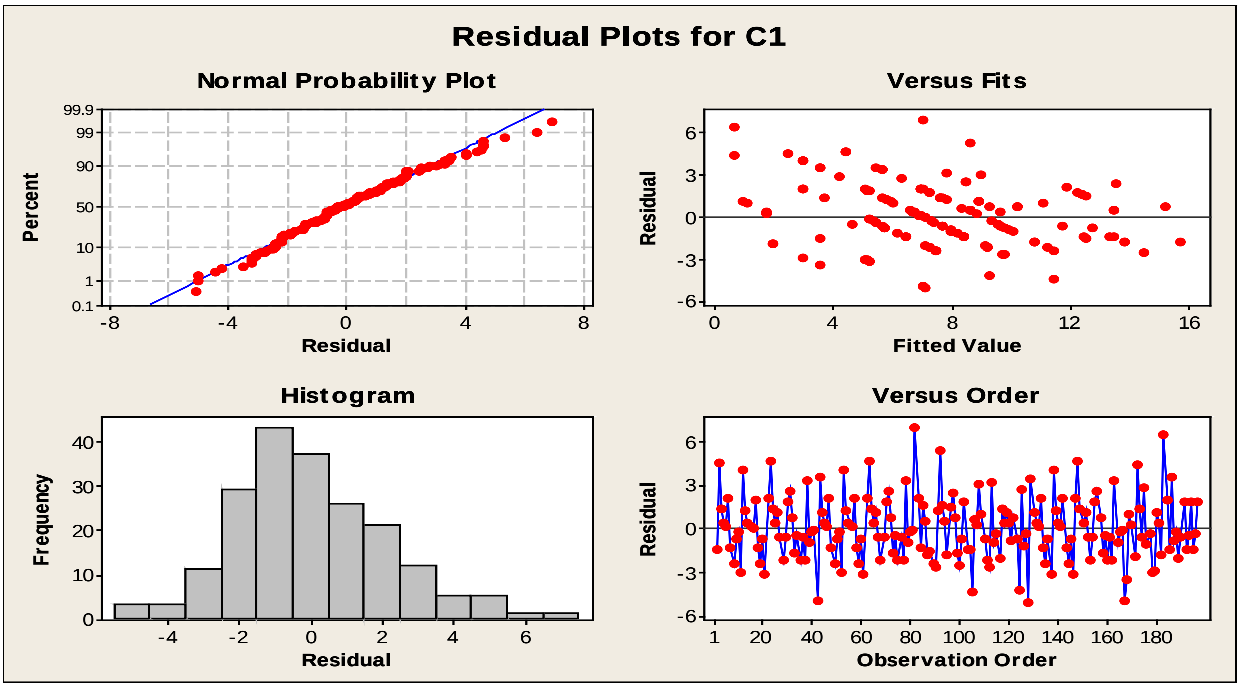

Figure 2. Residual plots for C1.

The normal probability graph plots the residuals versus their expected values when the distribution is normal. The residuals from the analysis should be normally distributed. Generally, for data with a large number of observations, moderate shifts from normality do not affect the results much. The normal probability plot of the residuals should be roughly following a straight line. Following are the interpretation for the graph [2]:

This pattern...	Indicates...
Not a straight line	Nonnormality
Curve in the tails	Skewness
A point far away from the line	An outlier
Changing slope	An unidentified variable

It is quite clear from Figure 2 that the plot roughly follows a straight line and contains no outliers.

The next graph plots the residuals versus the fitted values. The residuals should be scattered randomly about zero [2]. One can use this plot to look for the following:

This pattern...	Indicates...
Fanning or uneven spreading of residuals across fitted values	Non-constant variance
Curvilinear	A missing higher-order term
A point far away from zero	An outlier

Figure 2 clearly shows that the points are scattered randomly.

A histogram of the residuals shows the distribution of the residuals for all observations. One can use the histogram as an exploratory tool to learn about the following characteristics of the data: i. Typical values, spread or variation, and shape, ii. Unusual values in the data [2].

The histogram of the residuals should be bell-shaped which is quite clear in this case. One can use this plot to look for the following:

This pattern...	Indicates...
Long Tails	Skewness
A bar far away from the other bars	An Outlier

Because the appearance of the histogram can change depending on the number of intervals used to group the data, use the normal probability plot and goodness-of-fit tests to assess whether the residuals are normal [2]. We have already given the normal probability plot for residuals.

The graph residuals versus order plots the residuals in the order of the corresponding observations. The plot is useful when the order of the observations may influence the results, which can occur when data are collected in a time sequence (as in this case) or in some other sequence, such as geographic area. This plot can be particularly helpful in a designed experiment in which the runs are not randomized [2].

The residuals in the plot should fluctuate in a random pattern around the center line as in Figure 2. One can examine the plot to see if any correlation exists between error terms that are near each other. Correlation among residuals may be signified by (1) an ascending or descending trend in the residuals or (2) rapid changes in signs of adjacent residuals [2].

Interpretations from Figures 1 and 2: The random pattern of the residuals (Figure 2) together with the closeness of smoothed data with the observed one (Figure 1) justifies the Single Exponential Smoothing.

Remark: Single exponential smoothing is useful when: (1) there is no trend; (3) there is no seasonal variation; (3) there is no missing value; (4) we want short term forecast.

3.2 Comparing *Aye Mere Watan Ke Logo* with India's National Anthem [7] Scattered Plot

A scatter plot [8] is a type of plot using Cartesian coordinates to display values for typically two variables for a set of data. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis. A scatter plot can be used either when one continuous variable that is under the control of the experimenter and the other depends on it or when both continuous variables are independent. If a parameter exists that is systematically incremented and/or decremented by the other, it is called the *control parameter* or independent variable and is customarily plotted along the horizontal axis. The measured or dependent variable is customarily plotted along the vertical axis. If no dependent variable exists, either type of variable can be plotted on either axis and a scatter plot will illustrate only the degree of correlation between two variables. Figure 3 gives the scatter plot with notes of India's national anthem *Jana Gana Mana* along X-axis (variable C1) and those of *Aye Mere Watan Ke Logo* along Y-axis (variable C2). Variable C1 represents the pitch of the India's National anthem and variable C2 represents the pitch of the song '*Aye Mere Watan Ke Logo*'.

From Figure 3, it is observed that the points on the graph are randomly distributed which means there is no correlation between the two songs and two songs are quite different from each other.

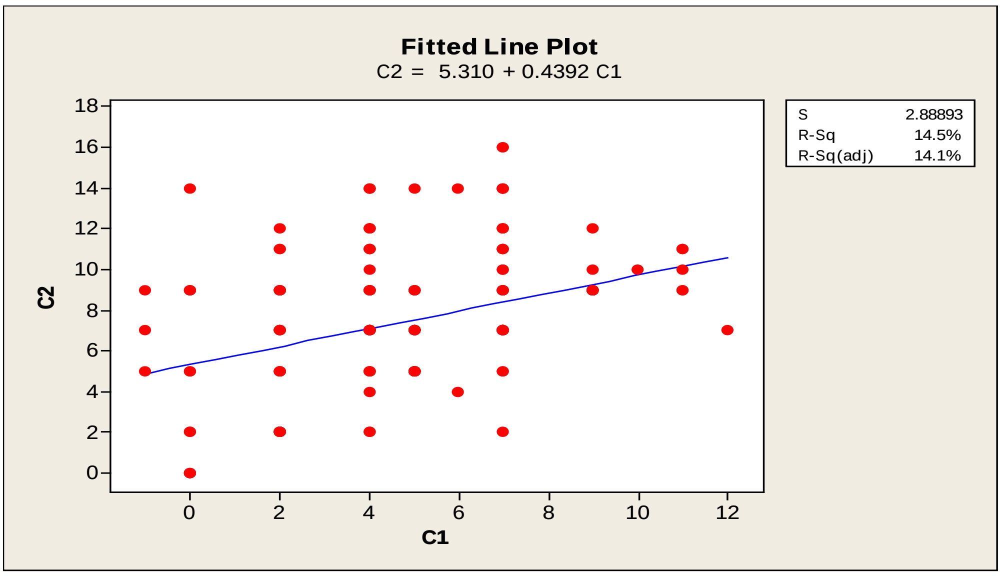

Figure 3. Scatters plot for C1 and C2.

3.3 Comparing the two songs with Single Exponential Smoothing

Comparison Table

	India's National Anthem [7]	*Aye Mere Watan Ke Logo*
Alpha (smoothing factor)	0.716404	0.698330
MAPE	35.3967	30.0332
MAD	1.2554	1.6997
MSD	3.6897	4.6222

It is very interesting to note that while the two songs are quite different from each other as evident from the scattered plot but they can still be captured by the same time series model fairly well. This shows how modeling helps in classification of music from a statistical perspective. The explanation for this finding is that although the two songs appear to be uncorrelated (Figure 3), it is only suggestive of an absence of linear relationship in the notes of the songs. However, some strong and complex non-linear relationship might be present in the note patterns which is why the same time series model is working well for both the songs.

3.4 Information Content

Information content (IC) measures the surprise an event E (in our case the occurrence of a particular note) carries when it appears. Higher the probability of the event E, lower will be the information content and hence lower the surprise. The formula is

IC (note) =$-\log_2\{P(E)\}$ = $-\log_2\{\text{Probability(note)}\}$ in our case. Table 4 and Table 5 give the information content (IC) of the common notes between India's National Anthem and '*Aye Mere Watan Ke Logo*' respectively.

Table 4. Probability of common notes and their information content for India's National Anthem

Sl. No.	Common notes	Number of Occurrence of the note= f in India's National Anthem	Probability=f/131	Information Content
1	C	9	0.0687	3.8635
2	D	22	0.1679	2.5740
3	E	43	0.3282	1.6072
4	F	17	0.1298	2.9460
5	G	27	0.2061	2.2785
6	A	5	0.0382	4.7116
7	B	6	0.0458	4.4485

Table 5. Probability of common notes and their information content for '*Aye Mere Watan Ke Logo*'

Sl. No.	Common notes	Number of Occurrence of the note =f in Aye Mere WatanKe Logo	Probability=f/197	Information Content
1	C	12	0.0609	4.0371
2	D	25	0.1269	2.9782
3	E	3	0.0152	6.0371
4	F	32	0.1624	2.6220
5	G	66	0.3350	1.5777
6	A	43	0.2183	2.1958
7	B	10	0.0508	4.3001

3.4.1 Interpretation (for common notes only in the two songs)

Information content is maximum for the note having lowest probability. For India's National Anthem least probability is of the note A and hence it has the highest Information content of 4.7116. In case of '*Aye Mere Watan Ke Logo*' least probability is of note E and it has highest Information content of 6.0371. The note E has the highest probability for India's National Anthem as compared to the note G for *Aye Mere Watan Ke Logo*.

3.5 Box Plot

BOX PLOT, also known as BOX WHISKER PLOT in Descriptive Statistics, is a method for graphically depicting groups of numerical data through their quartiles. Box plots may also have lines extending vertically from the boxes (*whiskers*) indicating variability outside the upper and lower quartiles, hence the terms box-and-whisker plot and box-and-whisker diagram. Outliers may be plotted as individual points. Box plots are non-parametric: they display variation in samples of a statistical population without making any assumptions of the underlying statistical distribution. The spacings between the different parts of the box indicate the degree of dispersion (spread) and skewness in the data, and show outliers. Box plots received their name from the box in the middle [9, 10].

The Quartiles along with min and max give a very sophisticated comprehensive summary of the data that measures central tendency as well as dispersion (that is, the amount of discrepancies among the data). The box plot is a graphical sketch of the data with the help of the five numbers (min, Q_1, Q_2, Q_3, max). The box whisker plot is often used to compare two or more datasets when there is a common quantitative variable under study [9, 10]. Figure 4 and Figure 5 give the box plot model of India's National Anthem and '*Aye Mere Watan Ke Logo*' respectively.

Interpretations of box plot [9, 10]:

1) If the intervals (Q_1, Q_2) and (Q_2, Q_3) are almost equal and the whiskers are almost equal then it roughly implies that the histogram is approximately symmetric.
2) If the intervals (Q_1, Q_2) is smaller than the (Q_2, Q_3) and the box is closer to the minimum then roughly the histogram is positivity skewed (that is, it has longer right tail).
3) If the intervals (Q_1, Q_2) is larger than the (Q_2, Q_3) and the box is closer to the Maximum then roughly the histogram is negatively skewed (that is, it has longer left tail).

Descriptive statistics for the India's National Anthem are given in Table 6. Descriptive statistics for the India's patriotic song '*Aye Mere Watan Ke Logo*' are given in Table 7.

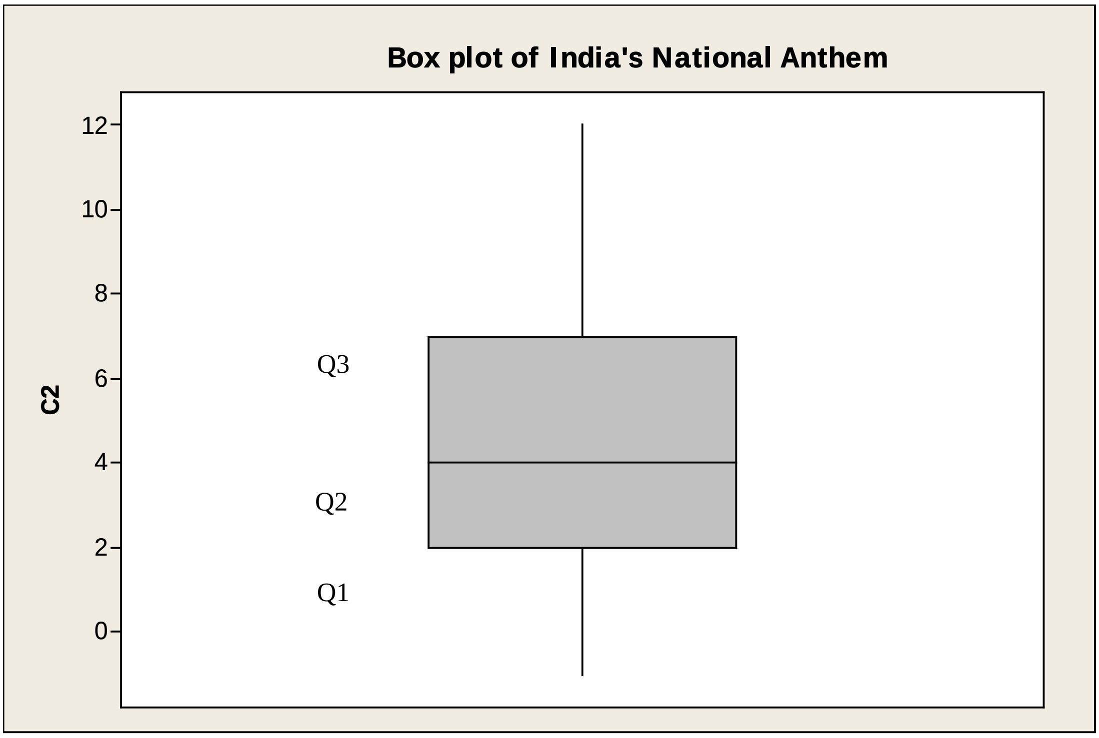

Figure 4. Box plot for India's National Anthem.

Table 6. Descriptive statistics for the India's National Anthem

Variable	N	Mean	SEMean	St Dev	Minimum	Q1	Median	Q3	Maximum
C2	131	4.496	0.225	2.576	-1.000	2.000	4.000	7.000	12.000

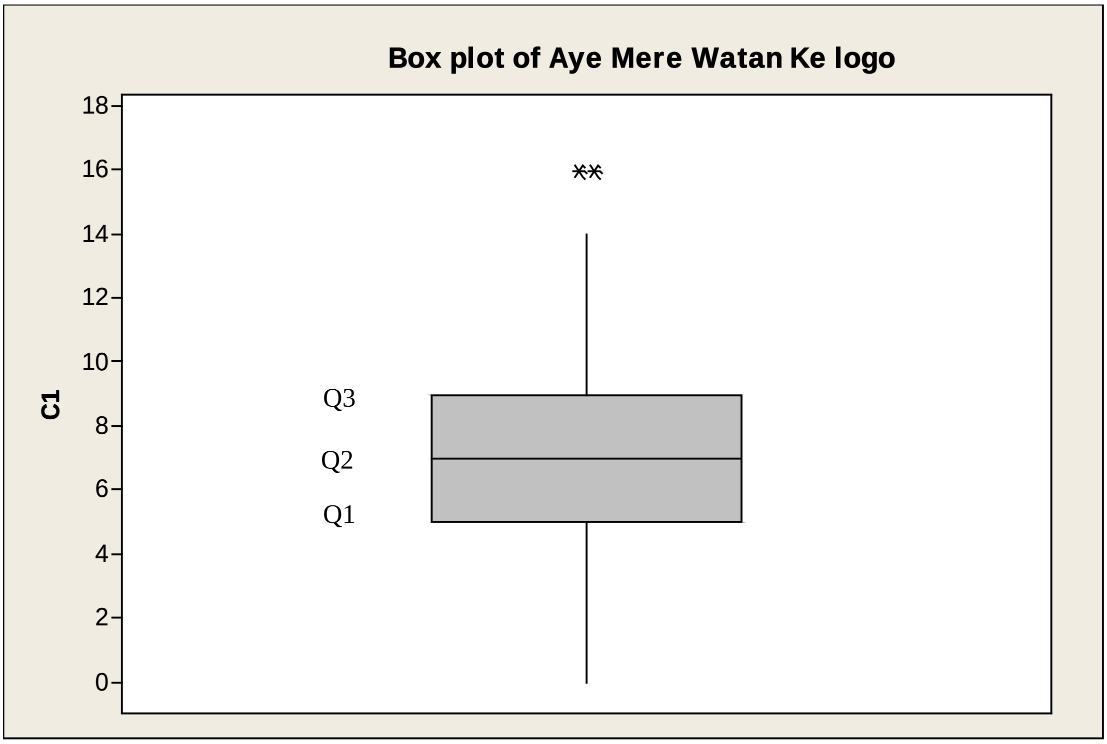

Figure 5. **Box Plot of '*Aye Mere Watan Ke Logo*'.**

Table 7. Descriptive statistics for the India's patriotic song '*Aye Mere Watan Ke Logo*'

Variable	N	Mean	SE(Mean)	St Dev	Minimum	Q1	Median	Q3	Maximum
C1	197	7.518	0.221	3.108	0.000	5.000	7.000	9.000	16.000

3.5.1 Interpretation

From Figure 4, it is clear that the interval (Q_1, Q_2) is smaller than the interval (Q_2, Q_3). Hence, this roughly implies that the histogram for India's National anthem is positively skewed. For Figure 5, for '*Aye Mere Watan Ke Logo*', the intervals (Q_1, Q_2) and (Q_2,Q_3) are almost equal. Hence the histogram is approximately symmetric. Comparing tables 5 and 6, we find that the mean pitch as well as the standard deviation of pitch values is higher in *Aye Mere Watan Ke Logo* as compared to those in India's National anthem.

Remark: Readers interested in further literature in Indian music and scientific research therein are referred to [12], [13] and [14].

4. Concluding remarks

Single exponential smoothing fitted to one of India's patriotic songs '*Aye Mere Watan Ke Logo*' explains the note progression well enough with smoothing factor alpha 0.698330. In Single Exponential Smoothing, the inner mechanism is quite straight forward, hence can be easily understood by everyone. There is also a comparative study done with India's National Anthem *Jana Gana Mana*. The results are quite interesting. Scatter plot plotted between the two songs shows that the two songs are entirely different from each other.

There are other comparing studies done like Information Content and Box plot. For India's National Anthem, note A has the highest Information content of 4.7116. In case of '*Aye Mere Watan Ke Logo*', note B has the highest Information content of 4.3001. The note E has the highest probability for National Anthem as compared to the note G for *Aye Mere Watan Ke Logo*.

From Box plot we get that India's National Anthem is positively skewed whereas *Aye Mere Watan Ke Logo* is approximately symmetric.

But it is quite interesting to note that, despite the aforesaid differences, both the songs can be approximated by the same time series model to a fair degree of accuracy. As mentioned earlier, this could be due to the presence of some complex non-linear relation in the note progression of the two songs which could not be depicted by the scatter plot between C1 and C2. *This strong and complex non-linear relationship in the note patterns reveals an interesting commonality which is why the same time series model is working well for both the songs.*

As a final comment, there is no unique way of analyzing a musical piece [11] whence statistics and probability have a definite role to play. This is well demonstrated by the present study.

References

[1] https://en.wikipedia.org/wiki/Aye_Mere_Watan_Ke_Logo accessed on 2.9.18.

[2] Chakraborty, S., Mazzola, G., Tewari, S. and Patra, M. (2014). Computational Musicology in Hindustani Music, Springer.

[3] https://en.wikipedia.org/wiki/Exponential_smoothing accessed on 2.9.1.

[4] https://otexts.org/fpp2/ses.html accessed on 3.9.18.

[5] Adiloglu, K., Noll, T., and Obermayer, K. (2006). A Paradigmatic Approach to Extract the melodic Structure of a Musical Piece. Journal of New Music Research, 35(3), 221-236 (2006).

[6] Chakraborty, S., Ranganayakulu, R., Chauhan, S., Solanki, S. S., and Mahto, K. (2009). A Statistical Analysis of Raga Ahir Bhairav. Journal of Music and Meaning, Vol. 8, sec. 4, http://www.musicandmeaning.net/issues/showArticle.php?artID=8.4.

[7] Chakraborty, S., Sarkar, S., Tewari, S., and Pal, M. (2012). Modeling India’s National Anthem: A Statistical Approach. International Research Journal of Social Sciences, Vol. 1(2), 1-7, October (2012).

[8] https://en.wikipedia.org/wiki/Scatter_plot accessed on 3.9.18.

[9] Pal, N. and Sarkar, S. (2008). Statistics: Concepts and Applications. Prentice-Hall Of India Private Limited.

[10] https://en.wikipedia.org/wiki/Box_plot accessed on 4.9.18.

[11] Beran, J. and Mazzola, G. (1999). Analyzing musical structure and performance—a statistical approach. Statistical Science, Vol. 14, No.1, 47-79.

[12] Raja, D. (2016). The Raga-ness of Ragas: Ragas Beyond the Grammar. D. K. Print World Ltd., New Delhi, India.

[13] Datta, A. K., Solanki, S. S., Sengupta, R., Chakraborty, S., Mahto K., and Patranabis, A. (2017). Signal Analysis of Hindustani Classical Music, Springer (2017).

[14] Chakraborty, S., Tewari, S., Rahman, A., Jamal, M., Lipi, A., Chakraborty, A., Nanda, A., and Shukla, P. (2021). Hindustani Classical Music: A Historical and Computational Study. Sanctum Books.

Appendix

Understanding the song (for non-Hindi readers) '*Aye Mere Watan Ke Logo*' through a free translation [1]

Understanding the song *Aye Mere Watan Ke Logo* with the help of its translation and with reference to the sequence of notes as provided in Table 1.

Line 1:*Ae Mere Watan Ke Logo, zaraa ankh me bhar lo paani,* [Translation**:** O' people of my country! Fill your eyes with tears! This line corresponds to notes 1- 20 in Table 1].

Line 2: *Jo shaheed huye hain unki, zaray aad karo qurbaani* [Translation :Remember the sacrifice, of those who became martyrs! This line corresponds to notes 21-41 in Table 1].

Line 3: *Tum bhoo lnaa jao unko, isliye suno ye* kahaani [Translation: And lest you forget them, listen to this story! This line corresponds to notes 42-60 in Table 1].

Line 4*: Jo shaheed huye hain unki, zaray aad karo qurbaani* [Translation: Remember the sacrifice, of those who became martyrs! This line corresponds to notes 61-81 in Table 1].

Line 5: *Jab ghayal huwa himalaya, khatre me padi azaadi, Jabtak thi saans lade woh, phir apni lash bichha di,* **[**Translation: When the Himalayas were injured and our freedom was threatened, They fought until their last breath, and then they laid down their bodies. Line 5 corresponds to notes 82- 127 in that sequence of Table 1].

Line6: *Sangeen pe dharkar maatha, so gaye amar balidaani Jo shaheed huye hain unki, zara yaad karo qurbaani* [Translation: With their faces on their bayonets, the immortal martyrs went to sleep. Remember the sacrifice, of those who became martyrs! Line 6 corresponds to notes 128-166 in that sequence of Table 1.]

Line 7: *Jay Hind, jay Hind ki sena* [Translation: Jai Hind! Glory to the Indian Army! Line 6 corresponds to notes 167-188 in that sequence of Table 1].

Line 8: *Jay Hind, jay Hind, jay Hind* **[**Translation: Victory to India, Victory to India, Victory to India! Line 7 corresponds to notes 189-197 in that sequence of Table 1].