Not So Lonely at the Top: *Billboard* #1s and a New Methodology for Comparing Records, 1958-1975.

William F. Carroll, Jr.

Adjunct Professor of Chemistry, Indiana University, Bloomington, IN 47405

wcarroll@indiana.edu

A Microsoft Excel database was constructed of the weekly entries in the Billboard Hot 100® from its appearance 4 August 1958 through the lifecycle of all records entering the chart before 31 December 1975. The distribution of peak position of records in this era is shown, including the large percentage of #1s; more records peak at #1 than at any other rank. The number of records appearing annually, and consequently, the average chart time varies markedly over the period. The database is also used to compare previously published chart performance methodologies, all of which depend upon chart lifetime. A new system, normalizing individual record score against an internal standard of contemporaneous records, mitigates the effect of the varying average chart lifetime over the period. Thus, it allows improved comparison of chart performance over time, and near consensus on the twenty strongest-charting records of the chosen era.

Keywords

Billboard; Hot 100; chart; life cycle; methodology; history

Objectifying the Subjective

Everyone has favorite records, and for some people, simply appreciating the record is not sufficient. They look for objective means to support the subjective decision. For example, "It went to #1" or "It spent 38 weeks on the chart."

It is impossible to objectify a subjective like "favorite." But one might reasonably ask: "How difficult is it for a record to become #1?" And "Are there ways of comparing chart performance among records over a relatively long period?" These questions can be answered using statistics associated with the most recognized measure of record popularity, the *Billboard* Hot 100®.

This study involves records entering the Hot 100 between 4 August 1958 and 31 December 1975, and compares individual record chart performance against the backdrop of changing chart methodologies, as well as changing music markets and marketing. A process is developed for smoothing some of the background variability, allowing for more level comparison of records across eras. While the period is arbitrary, it begins with the inception of the chart. The end was chosen because the main decade of interest for the author is the 1960s, but the database overlaps sufficiently with the prior and following decades to allow for analysis of the transition periods.

Billboard Hot 100®

The creation of this chart, if not its methodology, is well-documented, growing out of previous *Billboard* popularity lists. (Whitburn 7) The Hot 100 became increasingly important as Top 40 formats grew to dominate radio airplay in the late 1950s and early 1960s.

The *Billboard* Hot 100 was meant to facilitate the marketplace—to let those in the business know what others were playing and selling, and perhaps to suggest records that the reader might play or sell. It was not imagined to be an accurate historical performance record or a consumer tool. As time passed *Billboard* became something of a market maker as well as a market reporter, and labels and promoters developed strategies to leverage the chart's market influence to convince stations to include records.[1] (Stokes)

At the outset, rankings were generated by compiling information on sales strength, jukebox play and radio playlists to create a score for each record. From time to time, *Billboard* would note a change in its methodology openly in the magazine. For example, on 11 May 1968, an article appeared entitled "BB Sheds Airplay Factor in Top Half of Hot 100." ("*Billboard* Sheds Airplay Factor in Top Half of Hot 100" 3) This occurred as a result of a change in the airplay tabulation methods utilized by a number of Top 40 stations. The bottom 50 still reflected dealer sales reports and the "Top 40 stations' printed lists." The article acknowledged the need for airplay for newer entries to make chart progress. The average week-to-week move for records decreased at approximately the same time. The editors may well have made other internal modifications reflecting perceived changes in the marketplace without publicizing them.

The methodology became more directly data-driven with the advent of Nielsen SoundScan, a point-of-sale tracking system, in 1991. The controversy caused by this seemingly rational change to actual sales rather than expert opinion is well-documented elsewhere.   (Anand and Peterson 270-284) These authors note that moving from a qualitative to more quantitative measure of popularity caused a significant discontinuity in the charts and dislocation in the marketplace, especially among stakeholders who appreciated a more hands-on approach to charting. In recent years the charts evolved to incorporate the electronic delivery of music.

The specific method used in the early 1970s is somewhat less precise than implied by a 1-to-100 exact weekly ranking with no ties.[2] Hesbacher described the system in place in 1975 which was first announced as a modification of the existing methodology in a *Billboard* cover story on 9 June 1973. (Zhito 1; Hesbacher, Downing and Berger 74; Hesbacher "Sound Recording")

The changes included increased emphasis on airplay and sales of singles by one-stop distributors, the source for jukeboxes. As a result, sales of singles to jukeboxes were weighted equally with sales of singles to the public, an outcome of increased public preference for albums.

Additionally, this date marked the first computerization of chart compilation. Zhito noted that this would reduce the time from market sample to chart and also reduce "human error." All in all, it seems the changes made the system faster and more responsive to the played record—on radio or in public—rather than the purchased record. In the opinion of the editors, the changes were so significant that the previous week's chart information was not included because it was not comparable to the new system.

After the change, charts were compiled using information from 110 outlets in the 22 largest cities across the US: 63 radio stations, 25 one-stops and 22 retailers. For sales, the latter two entities were asked to judge the relative strength of 200 singles that might make the chart. Points for sales were awarded in two ways. First, he top 15 were to be reported explicitly and points (1-15) given in reverse order. Second, all records were then graded "very good" (20 points), "good" (15 points) or "fair" (5 points) for sales strength; Hesbacher does not report that a record could be given a zero or left off the list. Thus, the number one record at a given sales outlet would receive 20 + 15 = 35 points; the theoretical maximum for the 47 sales outlets would be 1645 points. The apparent minimum for any record was 235.

For radio exposure, 63 stations were weighted 1, 2, 4 or 6 based on listener population. The top 30 records on each station's playlist—which could have been based on airplay or could have simply been a subjective list compiled by station management (Hesbacher "Sound Exposure in Radio")—were assigned points by position. Hesbacher notes that assuming an even distribution of stations by size, the maximum number of airplay points a record could receive was 2040. A #1 ranking calculates to 10 points.[3]

Hesbacher does not detail how points were allocated for airplay ranking beyond that. However, utilizing a calculation of record strength contained in a subsequent article as an example, (Hesbacher "*Record World* and *Billboard*") starting with a #1 score of 10, other points might be awarded as shown in Table 1.

Totaling the national sales and airplay points, then sorting from high to low score, gave a rough draft of the chart. Then four supplementary procedures were applied: new entries, "stars," early deletion and smoothing.

New entries were denoted with an arrow to draw attention. *Stars* were awarded to records showing significant chart movement or sales growth in the week. The star assignment process changed a number of times over the period of study, but after the first four weeks of the Hot 100, stars were never awarded to records in ranks 91-100.

The final two processes had significant impact on the lifecycle of a record. *Early deletion* was applied on a discretionary basis to records deemed to have run their course, regardless of where they were on the charts.

Smoothing was a subjective process employed if sales or airplay seemed out of line, or if there were inordinate up-and-down swings week to week. Smoothing was used to rationalize the chart lifecycle of a record and minimize dislocation in the marketplace that might be caused by wild up-and-down swings.

The State of the Charts and #1s, 1958-1975

The 11,155 records in the database were divided into groups pre- and post- the 1973 change described by Zhito; the first, 4 August 1958 to 3 June 1973; the second 10 June 1973 to 31 December 1975. These were sorted to show the frequency of records peaking at ranks 1 to 100. The results are shown in Figures 1 and 2.

Between about ranks 15 and 90, the records sort themselves into roughly equal bins of approximately 0.85% each. Outside that range, there are two anomalies—regions where such a uniform distribution is perturbed—between ranks 1 and about 15, and also between 91 and 100.

Anomaly at Rank 91.

Inspecting Figures 1 and 2, from the number peaking at #91, it appears records have difficulty "escaping" the 90s. Statistics confirm this observation. (Table 2) Fifty-four percent of records ranked in the 80s advance to the 70s or higher with their next move; only 38% of those ranked in the 90s achieve the 80s or higher. Twenty-seven per cent of records ranked in the 80s stay there; 36% of those in the 90s do. And far fewer records in the 80s exit the chart with their next move (19% vs. 27%).

Perhaps records entering in the 90s simply do not have the momentum to move higher; records entering in the 80s have jumped at least 12 ranks to get there. But because record scores at lower ranks are spaced more closely, at least according to Hesbacher "Record World and Billboard," there is no

**Table 1.** Possible Billboard airplay scoring for ranks 1-30.

Rank	Score
1	10
2	9
3	8
4	7
5	6
6, 7	5
8-10	4
11-15	3
16-20	2
21-30	1

Table 2. Decade of next move of records ranked in the 80s and 90s, %

Rank	Higher	Same	Lower or Off
80s	54	27	19
90s	38	36	27
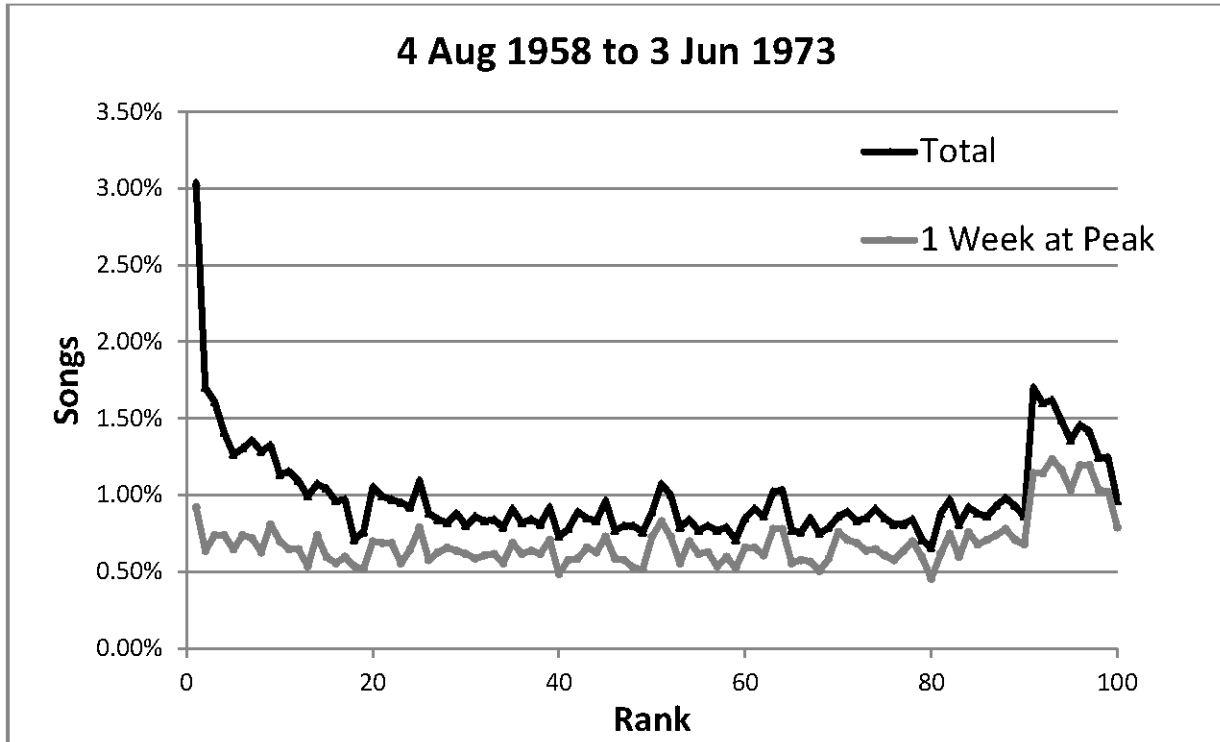
Figure 1. Total records peaking at rank vs. records peaking for one week only, 1958 to June, 1973. No statistical difference in ranks 2-90 for those with 1 week at peak.
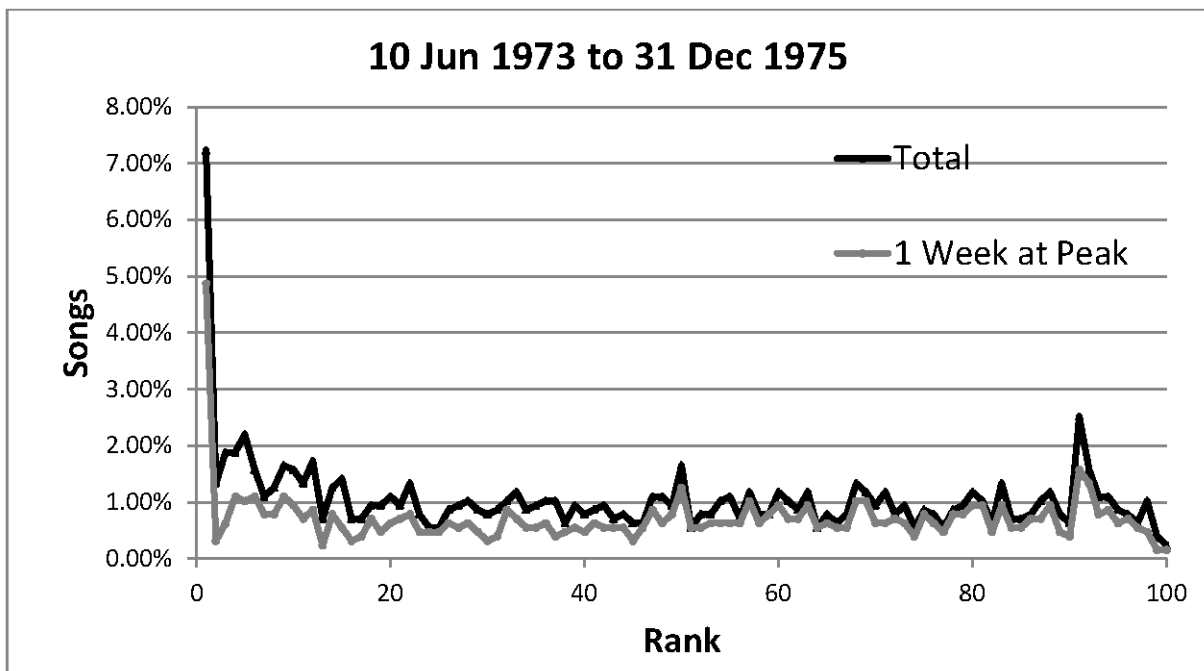
Figure 2. Total records peaking at rank vs. records peaking for one week only, June 1973-end 1975. No statistical difference ranks 2-90 for one week at peak. Note scale is different than Figure 1.

statistical reason that 91 should be so different from 90. Based on the manual nature of chart construction—including smoothing and early deletion—it seems likely that there was also an undocumented impediment preventing records from easily ascending from the 90s.

Anomaly Between #1 and #15 and the Abundance of #1s.

Between 1958 and June 1973, about 3% of records entering the *Billboard* chart reached #1. Between June 1973 and the end of 1975, however, nearly 7% of records reached #1—over double the number in previous years. Analysis of *Cash Box* charts from the same periods, and in the same way showed similar results; in the latter period, nearly 8% of all records entering the *Cash Box* chart went to #1. ("*Cash Box* Magazine Archives") In 1988, a similar behavior was noted without explanation in various charts in Radio & Records. (Barnes 36) It is counterintuitive that the "easiest" peak to achieve is #1, or that one record in 12 could achieve it; however, the paradox can be explained mathematically.

Between 1958 and June, 1973, 70% of #1s spent multiple weeks at that position; 33 records spent five consecutive weeks or more. (Figure 1) However, for the 1973 to 1975 period, most of the #1s spent only one week there and only four spent as many as four consecutive weeks. Fifteen records went to #1 in 1968; 35 did in 1974. (Figure 2) This change in #1 behavior correlates with the time of the methodology change, and the increased emphasis on airplay may be the source of increased turnover velocity at #1.

In this time period, new records on the charts varied between 499 and 744 per year; 1973, to 1975 had the fewest new entries and 1964 to 1967 the most. In each year, there are 52 slots for #1s. Realistically, between 12 and 50 records could achieve #1 in a year (1.6 to 10% of entries). Since records between 1958 and 1972 spent more weeks at #1 there were fewer #1s; at the same time more records were entering the chart. In the latter years, the rate of change at the top increased while there was lower turnover in the chart itself: more #1 records spent fewer weeks there, and fewer records entered. Together these two phenomena account for the swing from 3% to 8%.

Cases can be manually constructed where there are more #2s than #1s in a given year. However, that simply did not happen between 1958 and 1975. High frequency of exchange at the top decreases the average duration and increases the count of #1s. It also seems to decrease the frequency of records peaking at #2, #3 and so on, reaching equilibrium of about 0.85% at approximately #15. (Figure 1, 2)[4]

Perhaps this is because while any record occupying #1 peaks at #1, all other ranks see records "passing through." This pass-through comes to equilibrium at about #15; very few records spend more than one week at ranks 15 or below.

For example, a record occupying #2 will either peak there or be on its way to or from #1. Of the 392 records peaking at #1, 301 spent time at #2 for a total of 558 of the 916 weeks. The more changeover at #1, the more records passing through #2 and not peaking there.

Even with the relatively large number of records peaking at #1, spending as many as nine consecutive weeks at that rank is a transcendent achievement, having occurred only twice in the period; rewards acknowledging multiple consecutive weeks will be included in the charting methodology proposed later in this paper.

Average number of new entries, weeks spent on the chart and average record exit ranks vary markedly over the period but average entry rank (mid-80s) does not. (Figures 3, 4) Records entering the chart and average duration on the chart are obviously inversely correlated[5]; it is difficult to know whether one is the independent variable or whether another factor such as early deletion, or fewer singles released or aggressively promoted is the driver for both.

Previously Published Chart Methodologies

To compare records over time, one would prefer a tool that is both constant and calibrated; the *Billboard* Hot 100 is neither, for reasons given above.   The chart is simply as we find it: records ranked in order, separated by one unit. It is opaque with respect to the underlying calculations, data or changes in methodology. Other tools must be developed to facilitate comparison of records in different times.

At least six documented systems for evaluating chart performance have been published[6]. While some systems reward peak achievement, most rely on the area under the curve (AUC) formed by the record's life cycle. (Figure 5) In many systems, bonus points are awarded for particular types of achievement; their effect is to increase the AUC for a record or add to its score after AUC calculation. While these systems may be derived mathematically, they are all in a sense empirical and arbitrary, particularly given the nature of the underlying data.

These techniques were reproduced from their documentation and applied to the records in the 1958-1975 database. Results were then compared and reviewed.

Area Under the Curve (AUC) Systems

Reverse Rank. It might seem odd to imagine ranks higher than #1 on a *Billboard* chart; on the other hand, if the scale were reversed by subtracting chart rank from 101, wherein the highest ranked record scores 100 rather than 1, total score calculation is facilitated. Adding bonuses to create a score higher than 100 could be proposed without straining credulity.

Starting near the inception of the Hot 100, *Billboard* published an overall ranking for records charting that year. On only one occasion was the methodology referenced:

"based on the weekly chart positioning and length of time records were on the respective charts from the *Billboard* issue of January 4, 1969 through December 13, 1969…. Each disk was given points accordingly for its respective chart."
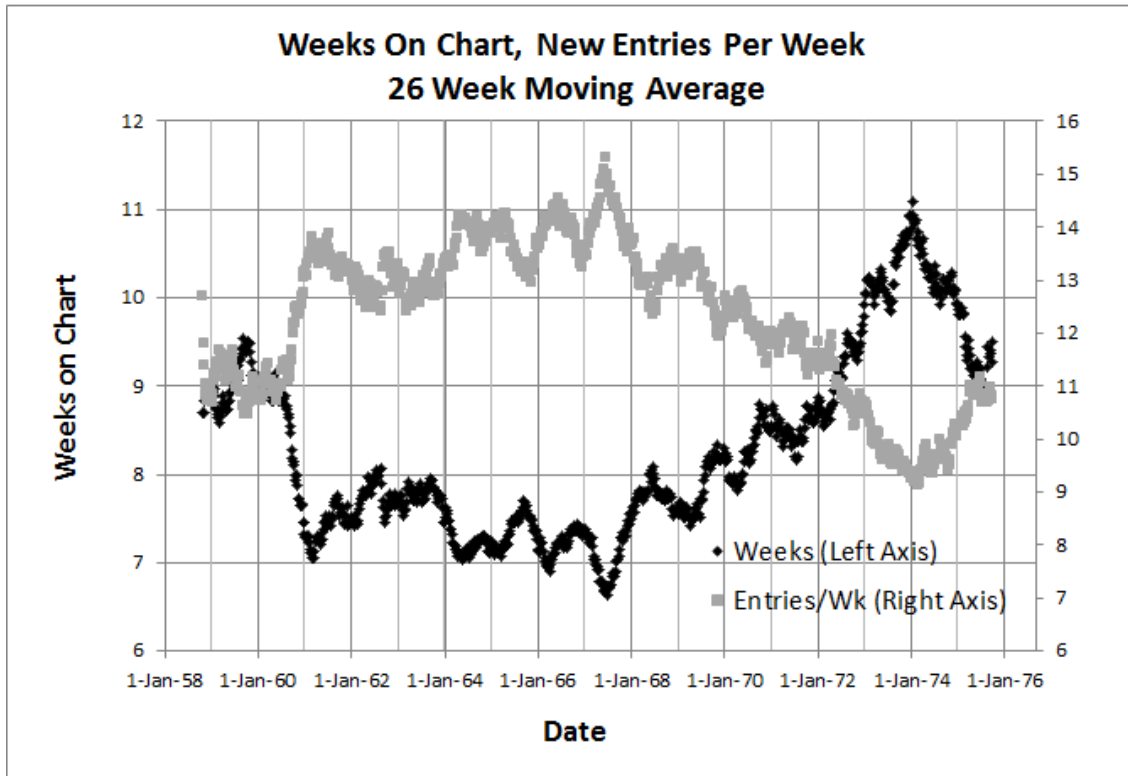
Figure 3. Weeks on Hot 100 and new entries per week, 26 weeks moving average, by week of entry.
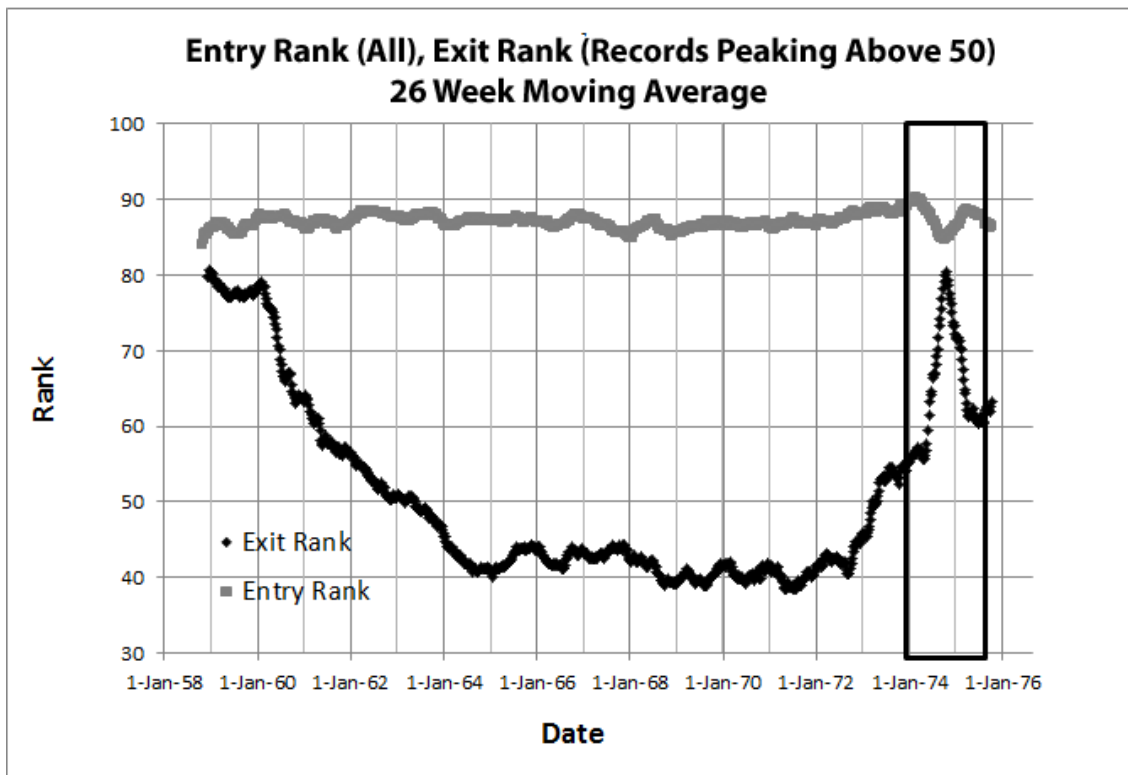
Figure 4. Entry rank for all songs, and last rank at which songs which peaked higher than #50 appeared before exiting the Hot 100, 26 weeks moving average, by week of entry. In 1974-5 the exit rank gapped upward by an average of nearly 20 ranks for six months, then returned.

Area Under the Curve
"Over And Over"
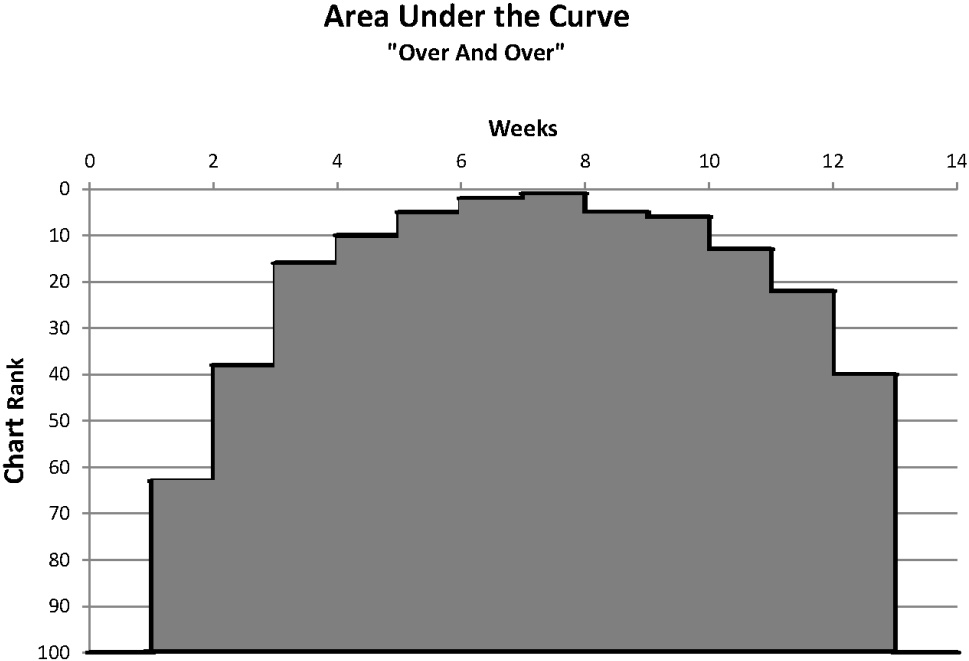
Figure 5. Area under the curve drawing for "Over and Over" by the Dave Clark Five, shown in native rank units.

The Reverse Rank approach—summing 101 minus rank for each week—nearly replicated the list that year. It is a common starting point for this kind of exercise. (Quirin and Cohen 1)

The highest charting record by the Reverse Rank method between 1958 and 1975 was "Mack the Knife," earning 2202 points. The first ten extend to about 1800 points.

Jim Quirin/Barry Cohen, "Rock 100", Chartmasters. "Rock 100" calculates a very technical index for weekly scores and adds bonuses for multiple weeks at #1. (Quirin and Cohen 1)   By studying Hot Movers on the charts, the authors developed a Popularity index ranking #1 to #100 based on a function that includes the logarithm of the inverse of the chart rank; point distances between high ranks are large (#1=100, #2 is about 94) and between low ranks are small (#99 is about 7.7 and #100 is about 7.3). Bonus points are derived from a complex model for consecutive weeks at #1, ranging from approximately 20 for two consecutive weeks to about 1028 for nine consecutive weeks. Two stints at #1 are awarded bonus points for the length of each, not the total length. Thus, "Mack the Knife" receives about 366 bonus points for 6 consecutive weeks plus approximately 63 more for three other consecutive weeks rather than 1028 for nine.

While this AUC method is not explicitly described, it was approximated for this study from graphs and scores included in "Rock 100". High score for the period 1958-1975 is "Hey Jude" at 2619; the top ten scores range downward to 1767.

Peter Hesbacher, University of Pennsylvania. In a series of publications in Popular Music and Society and the Journal of Communications, Hesbacher outlined the methods used by *Billboard* in the early 1970s. In one of those articles he generates a point system for comparing charts because:

> "…differences toward the top of the charts are larger than differences toward the bottom…
> (which) more accurately presents the proportion of popularity represented by the raw point
> totals. Its top 10 selections account for 44% of the total popularity (216 points) while a simple
> inverted point system's top ten selections account for only 19%." (Hesbacher "*Record World*
> and *Billboard*")

His system applies 20, 15 and 12 points for ranks 1, 2, and 3; 10 points decreasing to 4 for ranks 4-10; 3 points for 11-20; 2 for 21-30 and 1 for 31-100, creating weekly scores which were summed to create an AUC. Hesbacher had unique access to the *Billboard* methodology as a consultant, so one suspects this scheme may reflect the field data results. He fitted these weights to an equation he called a "J-Curve."

Hesbacher's equation is incorrectly transcribed in the article and is shown correctly below. The final constant should be added rather than subtracted as it was in the original article.

$$\text{Weight} = -30.404*[\text{Rank}/(\text{Rank}+2)]+30.265$$

The steepness of Hesbacher's points-to-rank plot is apparently driven by the steepness of the radio rating system. His fitting approach is similar to that of Quirin and Cohen; however, he awards no bonus points. Applying his rating system (rather than the fitted equation) to the dataset shows "Mack

the Knife" to be the highest charting record at 264 points, with the top ten extending to 187. Because of the low range of whole-number scores, this system produces a large number of ties.

Dann Isbell, "Ranking the '60s." Isbell begins with the Reverse Rank method, augmented by bonus points awarded for appearance in the top five, top ten, and second ten. It also includes a hierarchy of tiebreakers that allow for an explicit ranking of 6835 records from the 1960s in the cited book. (Isbell 1) The system of bonuses is heavily weighted to reward high rank; the bonus for a week at #1 is 470 points above the Reverse Rank score of 100; bonus for #2 is 295, progressing to #20, for which the bonus is 5. Weekly scores are summed to create the AUC. Sixty-three of the top 64 records in "Ranking the '60s" went to #1.

High score in "Ranking the '60s" is "Hey Jude" at 6960. Applying the methodology to the full dataset, "Mack the Knife" calculates to 7512, and the top ten extend down to 5360.[7]

Figure 6 shows a graph of each of the AUC systems applied to "Hey Jude." Base points are distinguished from bonus points where they are awarded.

Non-AUC Systems

Joel Whitburn, Record Research. For decades, Joel Whitburn has worked closely with *Billboard* creating "aftermarket" products. He developed a scoring system for records from which a career-long artist ranking could be generated. This system and the artist ranking is found in the company's compendium of *Billboard* charted records. (Whitburn 1075)

Points are awarded for peak position, extending from 100 points for #1 to 10 points for #100 and lower point values awarded for "Bubbling Under" records to #130. Additional bonus points are given per week for additional weeks at #1 (10), #2 (5) and #3 (3). Then one point is added for each week on the charts. This methodology does not compute an area under the curve; the data used—peak, weeks of chart life and weeks at #1—are provided in Record Research books. Weekly rankings, needed for AUC, are not.

The published method was applied to the records in the database. "Mack the Knife" has the highest score at 206, and the first ten extends to about 170. Similar to Hesbacher, this method generates many whole-number ties.

Howard Drake, Music VF. This system ranks records from 1900 to the present. (Drake) It is unique in that it integrates performance on most of the US charts as well as international charts. It employs a non-linear rating system for peak achievement similar to Whitburn, and adds bonus points for cumulative weeks at various positions; specifically, 5 points per week at #1; 2.5 points per week in the Top 10; 1 point per week in the Top 40 and 0.5 points per week in the Top 100. The various charts are given different weights, and a global score is generated from aggregating these charts.

Recognizing that chart creation and performance factors have changed over the past century, this system normalizes results year-to-year against a standard total score for the yearly top 100 records.
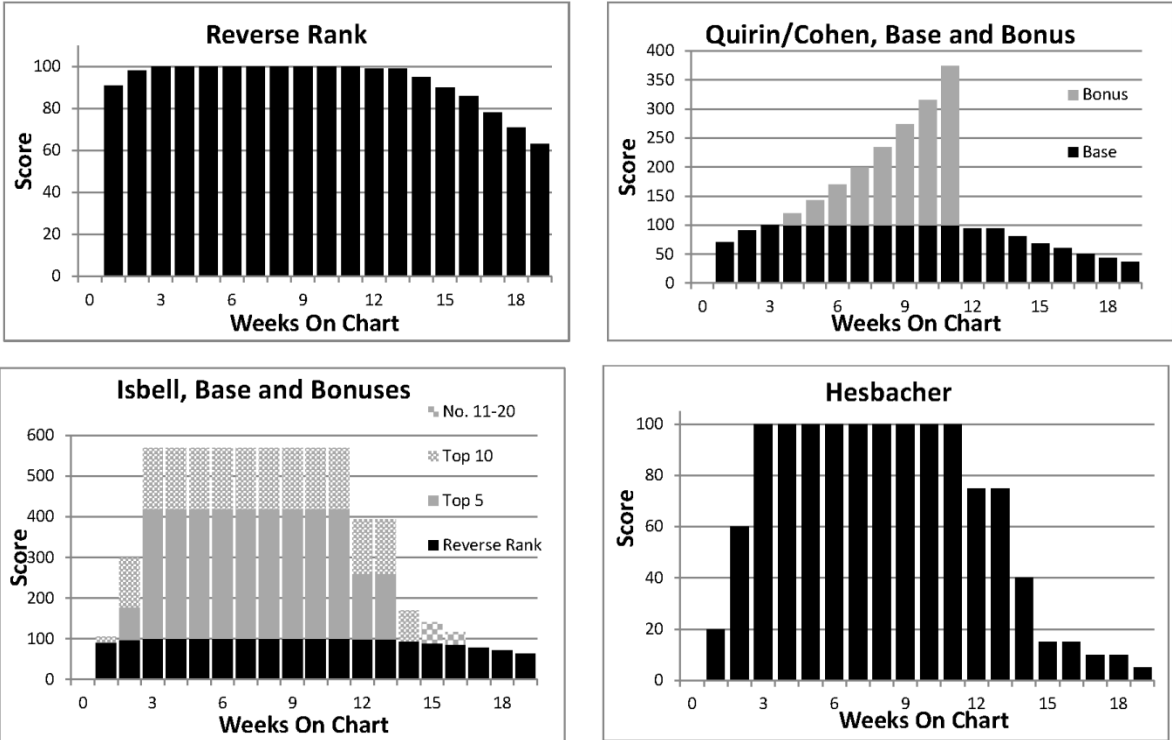
Figure 6. Four approaches to "Hey Jude," showing base points and bonuses. Note different scales. For Quirin/Cohen, the incremental bonus for weeks at #1 is attributed in the week it was accrued.

For the years 1955-2013, that score is 20,000. A factor is generated by dividing 20,000 by the total of each year's top 100 scores, and each record score in the database is multiplied by its annual factor.

 For the purposes of this study, a Hot 100-only version of the Music VF approach was calculated for 1958-1975. High score was "Mack The Knife" with 446.7 points. The top ten extends to 368.4 points.

A New Ranking System

The score sheets described by Hesbacher suggest that there was little differentiation in weekly points awarded for records, particularly below about #50. The non-linear reward system he proposed reinforces this suggestion.

A new AUC system of point rewards has been generated for this study, derived from his description of the *Billboard* polling sheets. For sales outlets, the top 15 records were awarded points, 15 down to 1. The top 100 records were rated "very good" from 1 to 25, "good" from 26 to 50 and "fair" from 51 to 100 and awarded points accordingly. For airplay, the even distribution of stations postulated by Hesbacher was used, and the top 30 records ranked from 10 points to 1 in the manner described earlier in this paper. The scores were combined to create ranks from 1 to 100, as though one record had all the votes for #1 (total 3685), one record all the votes for #2 (total 3434), and so on to #100 (total 235). Then, raw scores were fitted against this equation form: (Hesbacher "*Record World* and *Billboard*")

$$Score = Slope*[Rank/(Rank+n)] + Intercept$$

A complete list of the scores is shown in Table 3. The exercise found its best linear fit where n=10, slope = -4357 and intercept = 4139 (r^2=0.993). The calculated scores and fitted equation, scaled to #1=100, are shown in Figure 7.

Bonus Points

Billboard sometimes awarded a star to a sitting #1. This is evidence of increasing strength even at that rank[8]. A system that awards bonus points according to weeks at #1, as do Quirin and Cohen and Whitburn acknowledges increasing strength even while occupying the top of the chart.

There is an interesting and useful artifact of the proposed scoring system. When it is applied, the average record life cycle ascends linearly, especially approaching the top of the curve. In fact, the lifecycle trajectory resembles a triangle. The ascending slope is lower for records taking longer to reach their peak; the triangle is effectively flatter. An example of the average record lifecycle requiring eight weeks to peak is shown in Figure 8. Multiple weeks at peak is taken as evidence of continued score strength.
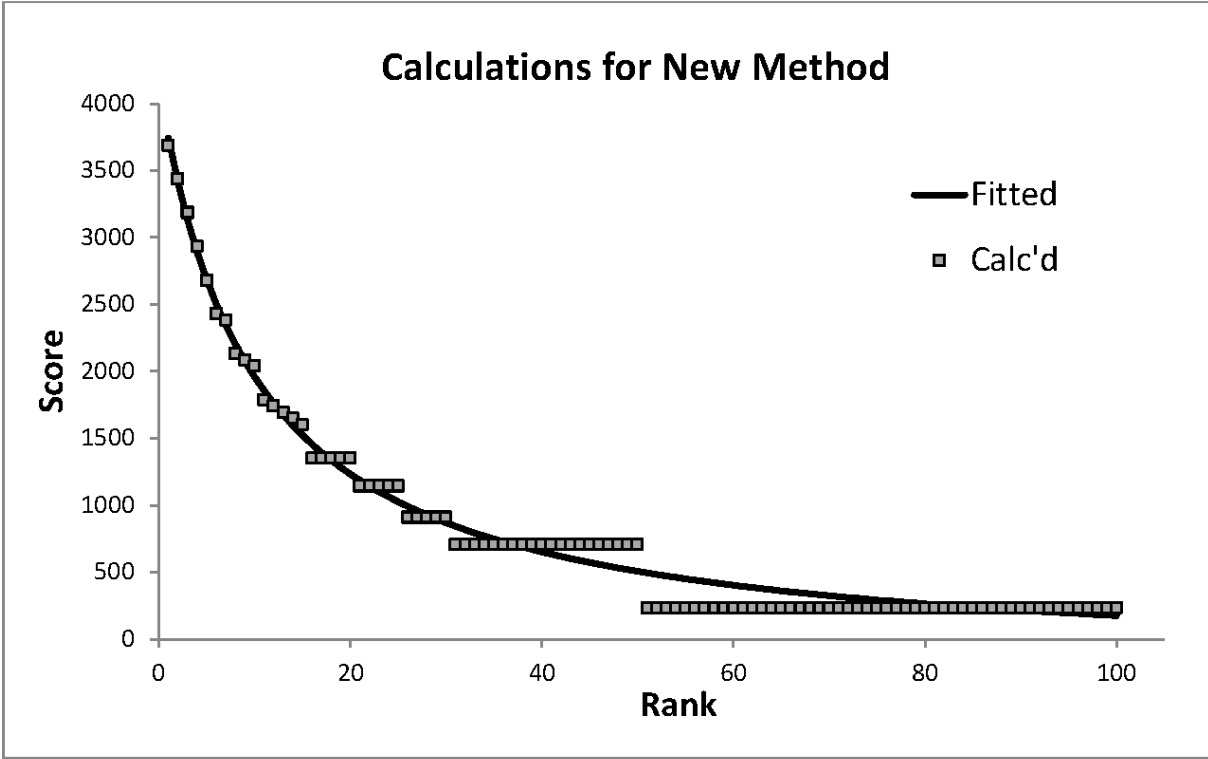
Figure 7. Scores calculated for the new system from the Hot 100 score sheet (Table 2) shown with fitted model. Fitted model is used in chart strength calculations.
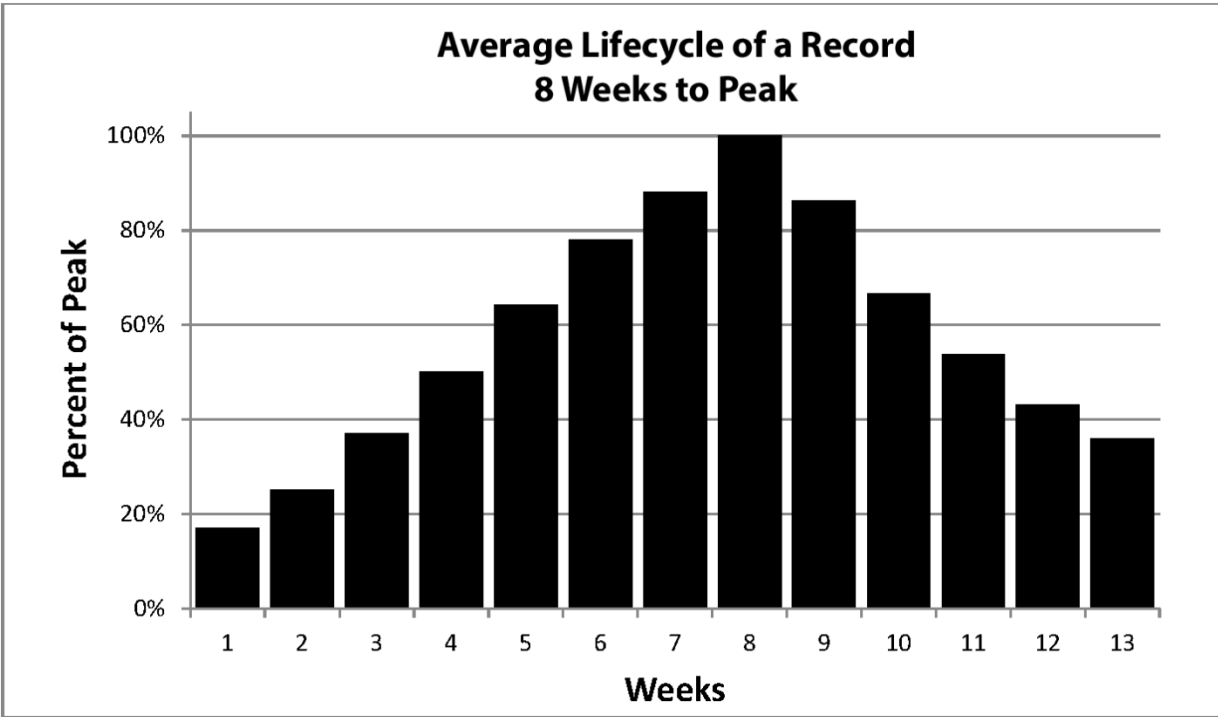

Figure 8. Average lifecycle of 205 songs, peaking at ranks #10 through #20 in 8 weeks.

Table 3. Rank scores prior to fitting derived from score sheet reported in Hesbacher.

Sales Score Sheet						Radio Score					Scale to 100
Rank	Qual	Rank Pts	Qual Pts	Total	Total 47 outlets	Rank	Pts	Total 63 Stations	Rank	Total	
1	VG	15	20	35	1645	1	10	2040	1	3685	100.0
2	VG	14	20	34	1598	2	9	1836	2	3434	93.2
3	VG	13	20	33	1551	3	8	1632	3	3183	86.4
4	VG	12	20	32	1504	4	7	1428	4	2932	79.6
5	VG	11	20	31	1457	5	6	1224	5	2681	72.8
6	VG	10	20	30	1410	6	5	1020	6	2430	65.9
7	VG	9	20	29	1363	7	5	1020	7	2383	64.7
8	VG	8	20	28	1316	8	4	816	8	2132	57.9
9	VG	7	20	27	1269	9	4	816	9	2085	56.6
10	VG	6	20	26	1222	10	4	816	10	2038	55.3
11	VG	5	20	25	1175	11	3	612	11	1787	48.5
12	VG	4	20	24	1128	12	3	612	12	1740	47.2
13	VG	3	20	23	1081	13	3	612	13	1693	45.9
14	VG	2	20	22	1034	14	3	612	14	1646	44.7
15	VG	1	20	21	987	15	3	612	15	1599	43.4
16	G		20	20	940	16	2	408	16	1348	36.6
17	G		20	20	940	17	2	408	17	1348	36.6
18	G		20	20	940	18	2	408	18	1348	36.6
19	G		20	20	940	19	2	408	19	1348	36.6
20	G		20	20	940	20	2	408	20	1348	36.6
21	G		20	20	940	21	1	204	21	1144	31.0
22	G		20	20	940	22	1	204	22	1144	31.0
23	G		20	20	940	23	1	204	23	1144	31.0
24	G		20	20	940	24	1	204	24	1144	31.0
25	G		20	15	940	25	1	204	25	1144	31.0
26	G		15	15	705	26	1	204	26	909	24.7
27	G		15	15	705	27	1	204	27	909	24.7
28	G		15	15	705	28	1	204	28	909	24.7
29	G		15	15	705	29	1	204	29	909	24.7
30	G		15	15	705	30	1	204	30	909	24.7
31-50	G		15	15	705	31-100	0	0	31-50	705	19.1
51-100	F		5	5	235				51-100	235	6.4

Therefore, for this new system, bonus points are awarded for spending more than one week at peak for ranks #1 to #10. One-half bonus unit is awarded for the second week at peak. Full bonus increments are added symmetrically to finish the top of the "cut off" lifecycle curve for three to nine consecutive weeks at peak, yielding a lifecycle trajectory resembling those that spend a single week at peak. (Figure 9)

The size of the bonus is based on the points awarded for advancing from a rank one below the peak to the peak itself. Number 1 is scored 100 and #2 is scored 91.2. The difference is 8.8 and that is the size of the bonus units for #1 records. Similarly, #3 is scored 83.7; the difference between #3 and #2 is 7.5 points, the bonus for multiple weeks at #2. As an example, this system is applied to "Hey Jude" in Figure 10, and can be compared to the other systems shown in Figure 6.

Discussion

All the published systems emphasize one or both of two fundamental quantities: peak performance and longevity. The use of bonus points for sustained high achievement (Quirin and Cohen, Isbell) or the explicit addition of weeks on the chart (Whitburn) acknowledges these two attributes.

Systems with a steep curve of non-linear rewards or those with large bonus awards for consecutive weeks-at-peak emphasize peak performance. For a system such as Reverse Rank, which uses neither of those two approaches, the opposite is true, and longevity is the most important success factor. (Figure 11)

This new methodology has three advantages. First, it is derived directly and transparently from the only detailed description of the *Billboard* polling method used during the relevant era. Second, it charts a more intermediate course between the extremes favoring peak performance or longevity. Third, it applies modest bonuses in a way that acknowledge peak performance while bringing a rational shape to life cycle trajectories for records peaking for multiple weeks.

Table 4 lists the 44 nominees for the top 20 highest-ranked records by raw score in all seven methodologies. With the exception of Reverse Rank, there is substantial agreement on the top three records, each of which spent 9 total weeks at #1. On the other hand, Reverse Rank scores "Why Me?" by Kris Kristofferson as the second strongest record of the period, based on 38 weeks on the charts despite peaking at only #16. Most of the other systems ranked it well down in the hundreds.

Make no mistake: longevity is important. Records that spent a large number of weeks on the chart generally did so because of sequential strength in local areas, even if never achieving Top 20, a traditional mark of national status. However, most people would agree the non-linear methods yield more reasonable candidates for the strongest charting and perhaps most important records of the period; the absence of "I Want To Hold Your Hand" and the inclusion of its contemporary "Hello, Dolly" in the Reverse Rank Top 20 is but one example.
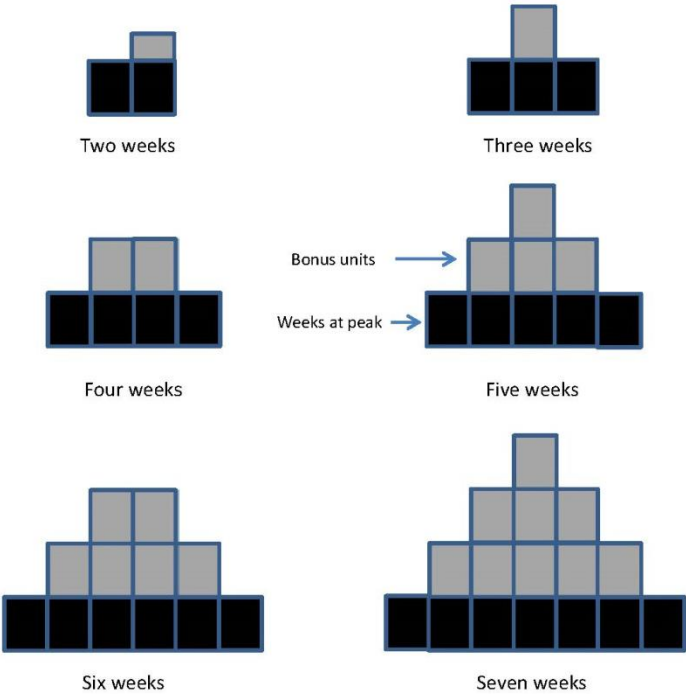
Figure 9. Bonus units are added symmetrically to complete the peak shape.
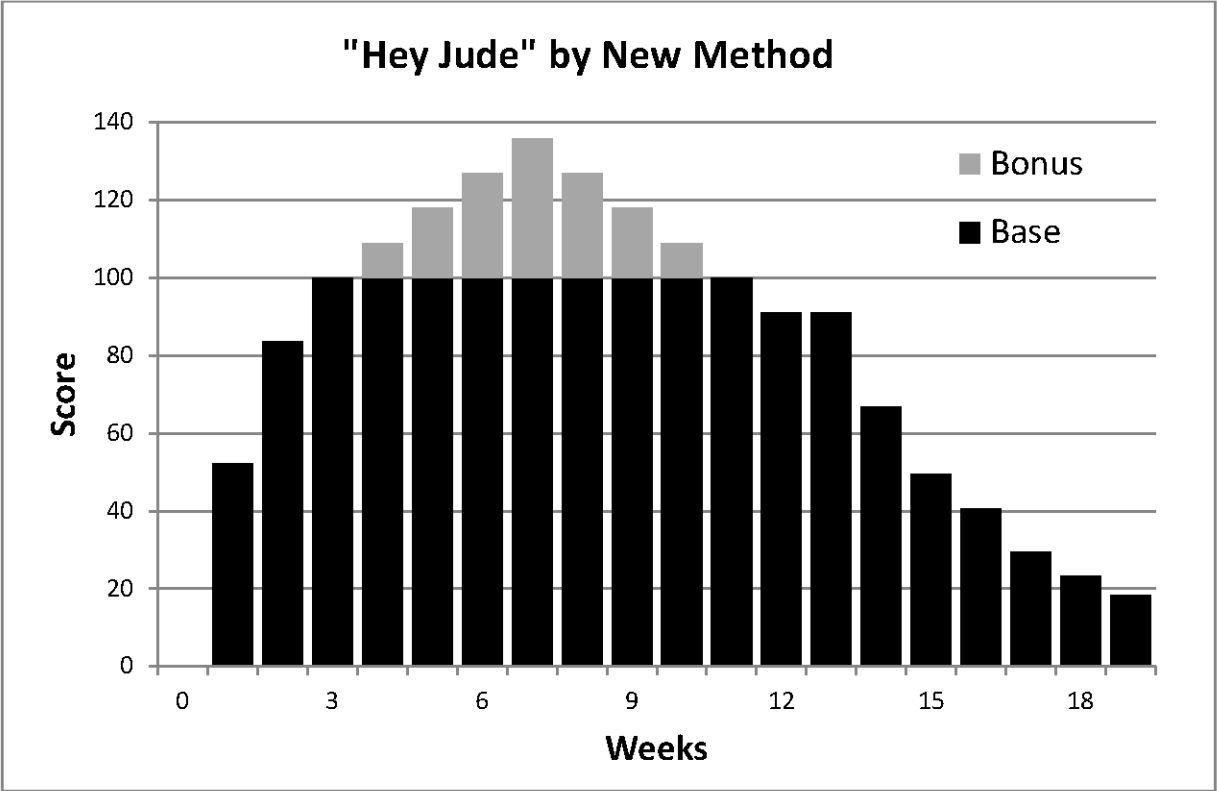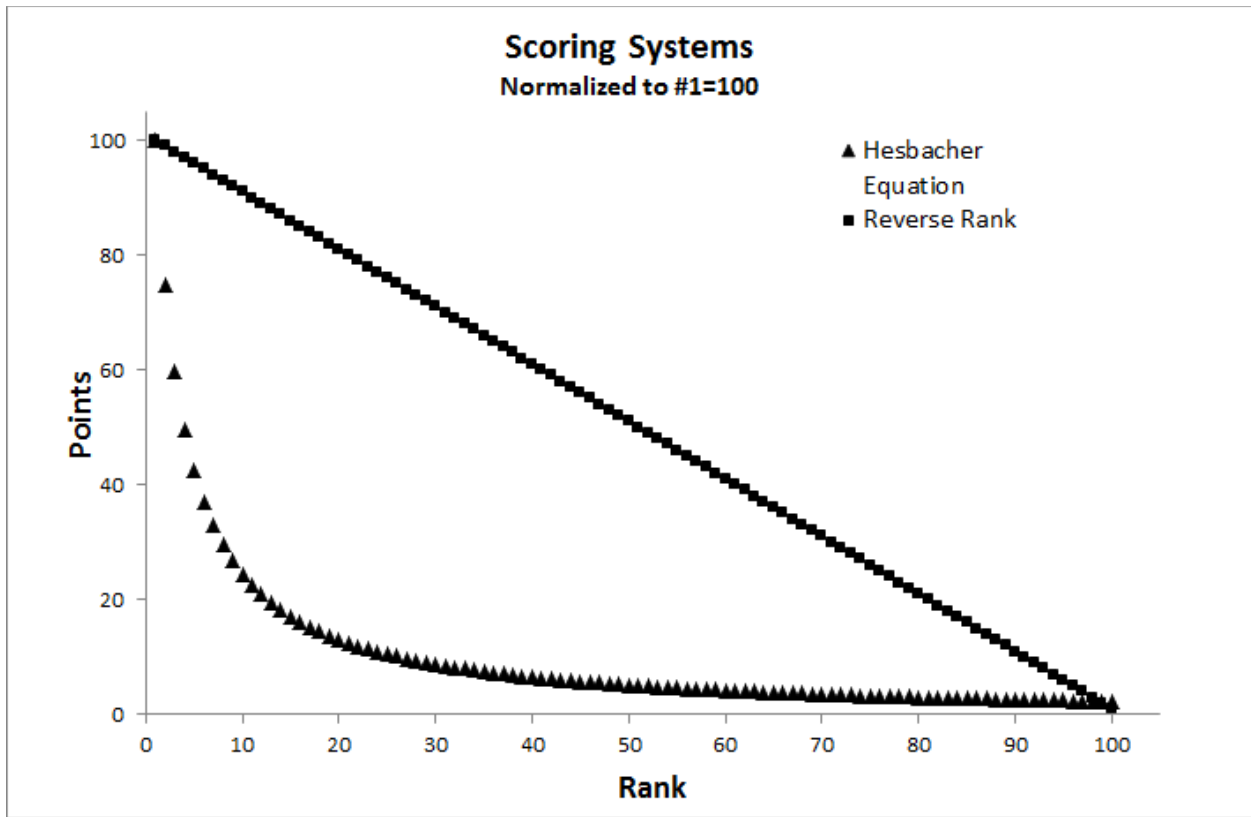

Figure 10. Lifecycle of "Hey Jude", including bonuses, in the proposed system.

Figure 11. Two approaches to scoring. Curves toward the upper right quadrant (Reverse Rank) emphasize longevity; curves toward the lower left quadrant (Hesbacher) emphasize peak performance.

Table 4. Nominees for Top 20 records as determined by the ranking methodologies, Raw Score. Brackets indicate position in that methodology outside the Top 20. Ties are designated by "t."

	Peak (Wks)	Tot Wks	Hes-bacher	Is-bell[a]	Q/C	Car-roll	Whit-burn	Music VF[b]	Rev Rank
Mack The Knife	1 (9)	26	1	1	3	1	1	1	1
Hey Jude	1 (9)	19	2	2	1	2	3	2	17
Theme From A Summer Place	1 (9)	21	3	3	2	3	2	3	[22]
Tossin' And Turnin'	1 (7)	23	4	4	4	4	4	4	10
Battle of New Orleans	1 (6)	21	5	5	6	5	9	5	20
I Want To Hold Your Hand	1 (7)	15	6	6	7	9	5t	7	[161]
It's All In The Game	1 (6)	22	7	7	5	6	8	6	8
Aquarius/Let The Sunshine In	1 (6)	17	8	8	10	11	12t	14	[92]
I'm A Believer	1 (7)	15	9	9	8	14	5t	9	[196]
Joy To The World	1 (6)	17	10	10	12	13	12t	10	[96]
Alone Again (Naturally)	1 (6)	18	11	12	[30]	[29]	10t	20t	[95]
Sugar Sugar	1 (4)	22	12t	13	13	8	[38t]	16	11
Are You Lonesome Tonight?	1 (6)	16	12t	11	15	[25]	14	11	[163]
I Heard It Through The Grapevine	1 (7)	15	14	15	9	[22]	5t	8	[174]
First Time Ever I Saw Your Face	1 (6)	18	15	14	11	15	10t	20t	[73]
Twist (1962)	1 (2)	21	16	20t	19	7	[140t]	[37]	9
I'll Be There	1 (5)	16	17	16	[21]	[27]	[25t]	17	[112]
Raindrops…Fallin' On My Head	1 (4)	22	18	20t	14	10	[38t]	12	12
Volare (Nel Blu Dipinto Di Blu)	1 (5)	16	19	17	[59]	[40]	[25t]	12	[226]
Reason To Believe// Maggie May	1 (5)	21	20t	18	17	[24]	17	13	[63]
I Can't Stop Loving You	1 (5)	18	20t	[22]	[25]	[32]	19t	[22]	[108]
Knock Three Times	1 (3)	18	20t	[25]	[48]	[30]	[84t]	[48]	[86]
It's Now Or Never	1 (5)	20	[23]	19	16	20	18	15	[42]
American Pie	1 (4)	19	[24t]	[24]	[26]	[28]	[44t]	18	[41]
Let's Get It On	1 (2)	19	[24t]	[37]	[40]	16	[155t]	[95]	[37]
I'm Sorry	1 (3)	23	[26t]	[27t]	[23]	17	[68t]	[26]	13
Love Is Blue	1 (5)	18	[26t]	[23]	[27]	[33]	19t	[24]	[125]
Tie A Yellow Ribbon…	1 (4)	23	[28]	[31]	18	[23]	[35t]	[57]	19
Bridge Over Troubled Water	1 (6)	14	[33t]	[30]	[24]	[48]	15	[23]	[323]
It's Too Late//Feel the Earth Move	1 (5)	17	[39t]	[33]	[29]	[39]	[21t]	19	[129]
To Know Him Is To Love Him	1 (3)	23	[39t]	[39]	[22]	18	[68t]	9	15
He'll Have To Go	2 (3)	23	[54]	[62]	20	12	[222t]	[174]	4
Hello, Dolly!	1 (1)	22	[63t]	[80]	[28]	19	[241t]	[46]	6
In The Year 2525	1 (6)	13	[68]	[59]	[54]	[132]	16	[40]	[772]
Rhinestone Cowboy	1 (2)	23	[81t]	[82]	[37]	[44]	[128t]	[182]	16
Limbo Rock	2 (2)	23	[131t]	[165]	[42]	[41]	[287t]	[208t]	18
Love Machine	1 (1)	28	[273t]	[230]	[47]	[101]	[160t]	[269]	5
Feelings	6 (1)	32	[525t]	[420]	[53]	[188]	[753t]	[880]	3
Lover's Question	6 (1)	24	[587t]	[496]	[81]	[199]	[895t]	[639]	14

Chantilly Lace	6 (2)	25	[701t]	[505t]	[91]	[257]	[852t]	[679]	7
Why Me	16 (1)	38	[964t]	[534]	[61]	[440]	[941t]	[1542]	2

[a]Isbell considers the two issues of "The Twist" to be one record; it is his overall #1. In this work the two are considered separately, because they are separated by more than a year. [b]Music VF includes a normalization by years that is described in the text.

The shortcomings of longevity as a primary measure of chart strength can be demonstrated with three other specific examples. First, three weeks at #67 (yielding 100 points) does not seem equivalent in importance to one week at #1 (also yielding 100 points). Second, the average duration for a record on the chart changed markedly over the period (Figure 3). Finally, some records, particularly those in the mid-1960s, were not allowed to play out their full life cycle due to early deletion. Some adjustment is needed to weight peak performance more heavily, without excluding longevity.

Accounting for Changes in Methodology

Some have suggested adjusting for the difference in average chart life over the period. Note that two of the consensus top three records on raw score came from periods of relatively longer average lifetimes. However, normalizing only for life span is difficult because life span and rank achievement are confounded in the total score of a record.

A better result is achieved by dividing a record's score by the average score of records on the chart contemporaneously. This generates a ratio of record strength relative to the competition rather than a raw score for comparison over time. While this does not reduce the longevity of records that were popular in longer-lived eras or lengthen those in shorter-lived times, it normalizes a record's performance against an internal scoring standard of other records in its time of popularity. Use of this ratio should smooth the differences in era, allow more direct comparison of records and methodologies, and virtually eliminate ties.

The average total score for all records entering the charts by week of entry was calculated for each methodology. A 52-week moving average of weekly scores centered on a record's date of entry was then generated. Individual record scores were divided by the moving average and those ratios ordered from highest to lowest.

A ratio of 1 indicates a record with average chart strength for its time. The span of ratios for each methodology starts at zero and ranges as follows: Whitburn, 4.76; MusicVF, 7.51; Carroll, 9.22; Quirin and Cohen, 9.53; Isbell, 12.88 and Hesbacher, 14.15.

This method produces two notable outcomes for the Top 20 charting records of the period (Table 5). Most importantly, it changes the mix: "Tossin' And Turnin'" moves up from consensus 4th to nearly consensus 2nd, and "To Sir With Love"—the #1 record of 1967—moves from outside the list to the mid-teens. These records had their life cycles during a period of significantly shorter longevity, and normalization against their contemporaries puts them on a more level playing field. In the raw score comparison, 34 records were nominated for the Top 20 by the non-Reverse-Rank systems, and 13 (38%) came from the 9 years of shortest average chart duration (1961-1969); after normalization, 15 of 25 (60%) came from that period. Solely on a pro-rata basis, a similar number--approximately 56%--would be expected to come from those years based on the number of charting records.

Secondly, it results in the same or more consensus records—those appearing in all the non-linear methodologies—for the Top 10, 20, 30, 40 and 50, although the impact is greatest for the Top 20.

Table 5. Top 20 records as determined by the ranking methodologies, normalized by moving average. Numbers in brackets are rank in that methodology outside the Top 20. Listed in consensus order.

Name (Top 20 Votes)	Hes-bacher	Isbell[a]	Q/C	Car-roll	Whit-burn	Music VF[b]	Rev Rank
Hey Jude (6)	1	1	1	1	1	1	7
Tossin' And Turnin' (6)	2	2	2	2	4	3	2
Mack The Knife (6)	3	3	4	3	2	2	3
I Want To Hold Your Hand (6)	4	4	5	5	6	7	[51]
Theme From A Summer Place (6)	8	7	3	9	3	6	[45]
Aquarius/Let The Sunshine In (6)	6	6	8	7	8	4	[30]
I'm A Believer (6)	5	5	6	8	7	9	[48]
I Heard It Through The Grapevine (6)	10	8	7	14	5	5	[69]
It's All In The Game (6)	15	15	9	12	9	8	15
Battle of New Orleans (6)	11	11	11	13	15	16	[40]
To Sir With Love (6)	14	12	13	15	13	10	[23]
Are You Lonesome Tonight? (6)	12	10	14	18	14	17	[102]
Love Is Blue (6)	13	14	16	16	16	13	[31]
Joy To The World (6)	17	16	17	20	10	14	[88]
I Can't Stop Loving You (6)	16	17	19	17	19	19	[43]
Sugar Sugar (5)	9	9	10	6	[30]	11	5
Raindrops Keep Fallin' On My Head (5)	18	19	15	11	[36]	12	8
It's Now Or Never (4)	[23]	20	20	19	20	[22]	[32]
Twist (1962) (4)	7	13	12	4	[106]	[21]	4
Big Bad John (3)	19	18	[23]	[23]	18	[24]	[79]
I'll Be There	20	[22]	[25]	[24]	[24]	18	[113]
Bridge Over Troubled Water	[32]	[27]	[22]	[52]	12	20	[245]
In The Year 2525	[50]	[44]	[33]	[97]	11	15	[536]
First Time Ever I Saw Your Face	[42]	[33]	[21]	[49]	17	[54]	[139]
Hello, Dolly!	[41]	[55]	18	10	[190]	[36]	1
Limbo Rock	[103]	[133]	[26]	[21]	[236]	[197t]	6
Sunshine Of Your Love	[689]	[502]	[124]	[362]	[611]	[677]	9
Why Me	[1181]	[714]	[197]	[759]	[1159]	[1657]	11
Exodus	[131]	[161]	[40]	[44]	[378]	[228]	10
Moon River	[1172]	[667]	[208]	[726]	[1071]	[1402]	12
Chantilly Lace	[752]	[516]	[119]	[326]	[879]	[687]	13
Stranger On The Shore	[70]	[71]	[40]	[27]	[193]	[68]	14
He'll Have To Go	[104]	[123]	[48]	[36]	[276]	[264]	16
North To Alaska	[412]	[384]	[78]	[143]	[724]	[466]	17
Feelings	[654]	[533]	[134]	[360]	[893]	[974]	18
Lover's Question	[613]	[508]	[104]	[260]	[922]	[638]	19
To Know Him Is To Love Him	[66]	[66]	[30]	[28]	[66]	[39]	20

[a]Isbell considers the two issues of "The Twist" to be one record; it is his overall #1. In this work the two are considered separately, because they are separated by more than a year. [b]Music VF is calculated using its underlying score system with normalization by moving average rather than year.

By raw score, 34 records were nominated, 11 Top 20 by consensus. After normalization, there were 25 nominees, 15 Top 20 by consensus and two which fell only one vote short.

Variation Between Methodologies at Lower Scores

If the methodologies converged perfectly, all lists would be the same. Obviously, they are not the same; there are two kinds of variation or uncertainty: between methodologies and within a methodology.

It is difficult to compare chart consensus on records further down the list, but some tendencies of the systems can be seen. (Figure 12) Quirin and Cohen scores fall most quickly due to large bonuses for records spending more than three weeks at the top. Their record #20 scores 41% less than their record #1. Whitburn is least steep throughout the curve; this is due to his relatively small bonuses and methodology that extends to weekly #130 rather than #100. The other systems start and end the same places but take different paths otherwise, which will impact records in the middle range.

To test variability between methodologies, ten records ranked 100, 200, 500, 750, 1000, 1500, 2000, 3500, 5000 and 7500 in the six methodologies were retrieved. The ranks of those 60 records in each of the other five methodologies (total 360 data) were also retrieved and a two-way analysis of variance was performed.

No significant systematic scoring differences were found among the methodologies; that is, none was consistently higher or lower than the others. However, they varied greatly among themselves, and variance changes significantly with rank. Standard deviation at #100 is approximately 45%, falling to 12% at #1500, leveling at approximately 6% at the lower scores.

Variation Within Methodologies at Lower Scores

Within a methodology, tie scores are one measure of precision. If five records are tied for #100, all that can be said for the rank of that record is that it is between 100 and 104, absent a tiebreaker system as proposed by Isbell.

Utilizing the normalization process, most exact ties are eliminated; thus, the question becomes: what constitutes a statistical tie? Some separation between scores is so small as to be meaningless for decision-making. There is no obvious objective way to determine what that meaninglessly small separation really is.
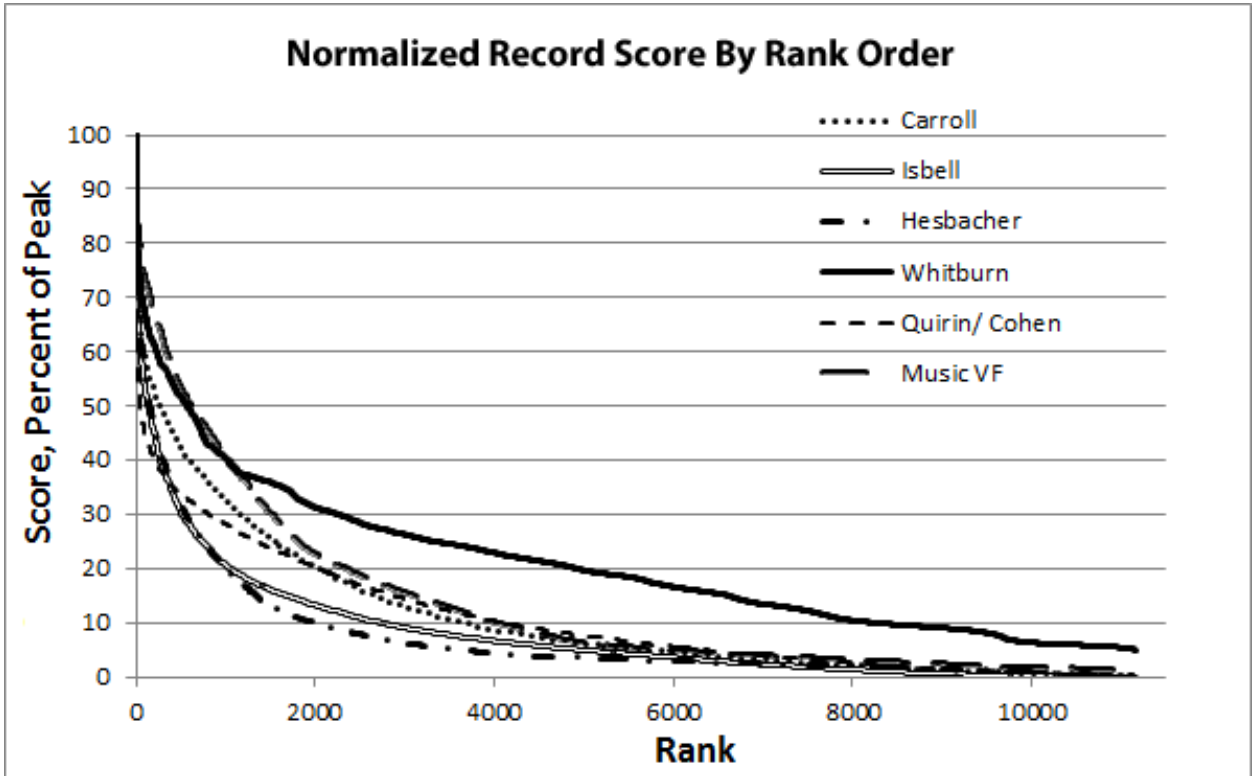
Figure 12. Normalized song scores ranked in order from 1-11,156.

The smallest separation between raw scores in the continuous systems (Carroll, Quirin and Cohen and Isbell) is the difference between 99 and 100 in one week. This translates to approximately 0.0011 +/- 0.0006 ratio units in these three systems, whose maximum ratios are 9.22, 9.53 and 12.88 respectively. Rounding the de minimis level slightly higher at 0.002 for all three systems suggests that the minimum uncertainty for the calculated rankings within any of these systems is approximately   +/- 1 at rank 100, +/- 5 at rank 1,000, +/- 20 at rank 5,000 and +/- 100 at rank 10,000. Increasing this de minimis difference by a factor of 10 raises those uncertainties by about a factor of 5 at the lower end to 10 at the higher end. Given the method for data collection and processing, actual uncertainty is undoubtedly somewhat higher yet, but perhaps not as large as the variation between methodologies.

Summary

In some years up to 7% of records that enter the *Billboard* Hot 100 find their way to #1; more records peak there than at any other rank. While counterintuitive, it is a mathematical result of the number of records to achieve #1 and the number of records entering the charts; both changed over the period and could well be impacted by changes in market, marketing and chart construction.

Hot 100 chart analysis methodologies published to date have been assembled in different ways and differently emphasize peak rank or longevity on the chart. However, with the exception of the Reverse Rank method, the literature methodologies as well as a new methodology derived from a reported *Billboard* scoring procedure can be harmonized for the strongest charting records by normalizing individual record scores versus average scores of contemporaneous records. This offsets the effects of average chart longevity and variability induced by the analysis methodology itself. Thus, all methodologies yield similar results at the top, regardless of construction. The new approach strikes a middle ground between peak performance and longevity. Future studies will utilize this new methodology to compare the careers of artists and songwriters over this same period.

Acknowledgements

References

Anand, N, and Richard A. Peterson. "When Market Information Constitutes Fields: Sensemaking of Markets in the Commercial Music Industry." *Organization Science* 11.3 (2000): 270-84. Print.

Barnes, Ken. "On the Records: Number Ones in the Spotlight." *Radio & Records* 23 December 1988, p 36. Print.

"*Billboard* Sheds Airplay Factor in Top Half of Hot 100." *Billboard* May 11, 1968, p. 3. Print.

"*Cash Box* Magazine Archives". Accessed March 30 2014. <http://www.cashboxmagazine.com/Archives.html>.

Dannen, Frederic. *Hit Men*. New York, NY: Vintage Books, 1991. Print.

Drake, Howard. "Musicvf". Accessed March 30 2014. <http://www.musicvf.com/>.

Hesbacher, Peter. "*Record World* and *Billboard* Charts Compared: Singles Hits, 1970-1979." *Popular Music and Society* 8.2 (1982): 101-12. Print.

---. "Sound Exposure in Radio: The Misleading Nature of the Station Playlist." *Popular Music and Society* 3.3 (1974): 189-201. Print.

---. "Sound Recording Popularity Charts: A Useful Tool for Music Research I." *Popular Music and Society* 4.1 (1975): 3-18. Print.

Hesbacher, Peter, Robert Downing, and David G. Berger. "Record Roulette: What Makes It Spin?" *Journal of Communication* 25.3 (1975): 74-85. Print.

Isbell, Dann. *Ranking the '60s: A Comprehensive Listing of the Top Songs and Acts from Pop's Golden Decade*. Jefrian Books, 2012. Print and Electronic.

Quirin, Jim, and Barry Cohen. *Chartmasters' Rock 100: An Authoritative Ranking of the Most Popular Songs for Each Year, 1954 through 1991*. Covington, LA: Chartmasters, 1992. Print.

Stokes, Geoffrey. *Star-Making Machinery: The Odyssey of an Album*. Indianapolis: Bobbs-Merrill Company, 1976. Paper.

Whitburn, Joel. *Top Pop Singles, 14th Edition, 1955-2012*. Menomonee Falls, WI: Record Research, Inc., 2013. Print.

Zhito, Lee. "*Billboard* Launches Super Singles Chart." *Billboard* June 9, 1973: 1, 9. Print.

Notes on Contributor

Bill Carroll is an industrial chemist and Adjunct Professor of Chemistry at Indiana University, Bloomington, IN. He is a past president of the American Chemical Society, the world's largest single-discipline scientific organization. Thirty-six years after his doctorate and a career which included numerous scientific publications, he returns to his first love: popular music, charts and Top 40 radio, utilizing physical science data reduction techniques to develop music chart analytics.

---

[1]   This leverage of the charts pales compared to the outright manipulation of the industry during the independent promoters scandal, ca. 1980. Frederic Dannen, Hit Men (New York, NY: Vintage Books, 1991).

[2]   In fact, after the first three weeks of the Hot 100, there was only one tie: two #100s on March 7, 1960.

[3]   The average weight for an even distribution is 3.25. 2040 points divided by this weight leaves 627 points for 63 average stations; #1 is apparently worth 10 points.

[4]   Inspection of the New York Times Best Seller List, which is constructed similarly, between 1958 and 1962, seems to confirm this observation. #1s tend to stay for more weeks and there are fewer of them; #2s occupy the ranks for fewer weeks but there appear to be few move-ups. http://www.hawes.com

[5]   Over time, Average Duration in weeks = 100/average weekly entries

[6]   The "100 Biggest Top 100 Hits" chart was published by *Cash Box* and its archivist Randy Price in 2008. A "progressive inverse points system" and a "compensation factor…for each year" were used, and multiple instances of the same record were consolidated. Music Industry News Network http://www.mi2n.com/press.php3?press_nb=113979, accessed October 14, 2014

[7]   Isbell treats the two instances of "The Twist" (1960 and 1962) as one record and totals their scores. Thus, he reports "The Twist" as the strongest charting record of the 1960s. In this work, the two instances are treated separately because they are separate in time by more than a year. There is a similar situation for the two instances of "Light My Fire" by the Doors.

[8]   Fifteen records in this period achieved a star in the second week or later at #1, virtually all in the '70s. However, because star criteria changed over the years it is unreliable to use this as a quantitative measure of growing strength.