

“I Just Can’t Wait to Be Queen”

An Investigation of Predictive Factors in the SAI Quartet Competition

Adam B. Zufall¹

University of California, Davis; Davis, California, 95616, USA

Members of Sweet Adelines International, a barbershop community, all love singing with their friends, ringing dominant seventh chords, and, for many SAI members, chasing the impossible dream of becoming “Queens of Harmony” by winning the annual International Quartet Competition. Informally, every member of SAI has observations and theories about what makes a winning quartet, but this paper presents the first data-driven analysis of the competition. In my analysis, I will confirm some commonly-held notions, such as having past winners in a quartet is correlated with higher scores, and challenge other common beliefs, such as showing that song choice is uncorrelated with competition score.

I. Nomenclature

df	= Degrees of Freedom
F	= Finals
FQ	= Future Queen
NQ	= Non-Queen
Q	= Queen
SAI	= Sweet Adelines International
SF	= Semi-Finals
Δ	= Change from first appearance at the International Competition
μ	= Mean

II. Introduction and Research Questions

Sweet Adelines International (SAI) is a women’s barbershop harmony society founded in 1945 with approximately 30,000 members across the United States and a few other (predominantly English-speaking) countries. Each year, quartets compete regionally to advance to the international competition; the winners of the international competition are crowned “Queens of Harmony” and are inducted into the prestigious Coronet Club. After earning their crowns, a winning quartet is no longer permitted to compete, encouraging Queens to form new quartets and sing with other members of the organization.

The format of the competition, first held in 1947, has undergone changes through the years, but has largely been unchanged during the years for which data is publicly available. Around 40 quartets compete at the Semifinals, the top 10 quartets perform again in the Finals, and the winner is the quartet with the most points summed across all stages of the competition. A standard competition “package” consists of two songs, typically a slower ballad and a faster up-tune. Packages are scored across four categories (Sound, Music, Expression, and Visual Communication) which assess their adherence to barbershop benchmarks. Although scores did increase across through the decades (see Fig. 1 of Section IV), scores have stabilized in the last two decades, allowing for the comparisons contained in this paper. As an artistic competition, the judging is inherently subjective; however, extensive effort has been expended by members of the organization to standardize the scoring criteria and train judges to ensure reliable scoring. As evidence of this, two judges are used for the international competition, with scores typically only varying by one half of a percent.

¹ Graduate Student, Department of Aerospace and Mechanical Engineering

The scoresheets from the judges provide direct feedback to competitors for how they can improve their scores; for example, having each member of a quartet pronounce vowels the same way to improve tuning. Quartets prepare for competition as you would expect: rehearsing, performing for other barbershoppers, and receiving instruction from trusted coaches. From this perspective, the scores earned by each quartet are a reflection of how long the quartet has been singing together, how much time they can dedicate to rehearsal, and the musical skill of the members. My own interest, however, is in the meta-game: what are the patterns that predict success apart from personal preparation? While an individual should focus on factors within their control, e.g. rehearsal and coaching, I would like to look for patterns on the macro level. My research questions stem from my own thoughts while listening to several competitions and speaking with competitors afterwards. I hope that this analysis can be, at the very least, interesting to members of SAI and, if the results are insightful, illuminate factors that influence quartets' scoring on the international stage. My primary research questions are:

1. *Do quartets containing a Queen of Harmony tend to score higher?*
2. *Does performing a common song (a barbershop “standard”) influence score?*
3. *Do quartets who “make a splash”, i.e. score highly at their first international competition, achieve more success at future competitions?*
4. *Is there a relationship between rank and the longevity of a quartet?*

This paper will be divided into the following sections: in Section III, I will describe some of the limitations of my observations and resulting analysis; in Section IV, I will describe the dataset, how it was cleaned, and give descriptive statistics; in Section V, I will describe and perform my statistical tests; and in Section VI, I will offer my interpretations of the statistical tests and the limitations of my interpretations.

III. Disclaimers

There are a few disclaimers to address before diving into the statistical analysis. First, a logistical clarification: I will refer to winners by the year in which the competition was held; this is different from SAI's tradition of referring to Queens by the “year of their reign”, which is the calendar year following the competition. For example, SAI lists the “2025 International Champion Quartet” as “Clever Girl”, who won their crowns on November 1, 2024.

Second, while this paper will focus on the competitive nature of SAI, the goals and benefits of the organization reach far beyond the competition stage. SAI is primarily an educational service organization; members all share their love of the barbershop musical tradition with their friends and communities. Some members may never compete at all, instead focusing on arranging music, choreographing dance, holding music workshops for singers of all ages, and, above all, enjoying the life-long friends they have made in Sweet Adelines. Also worth noting: due to the nature of overtones in the human voice, ringing chords in the barbershop style is extremely powerful and many musicians have found barbershop to be highly addictive. In light of this, SAI also serves as a support network for singers who need a regular dose of vocal opioids.

The third disclaimer is that past trends do not necessarily preclude an event from happening in the future. For example, there is no one with the name “Rachel” in the Coronet Club, yet it would be illogical to conclude that quartets with a member named Rachel are unable to win the competition. While this example is laughable, it's a small stretch to other statements, such as:

- “No quartet wins during their first year at Internationals”, which was true until Ambiance became queens at their 1st competition in 1986.
- “All Queens earn their crown within 7 appearances at Internationals”, which was true until Brava! became queens at their 11th competition in 2003.
- “No quartet from Australia can progress to the finals”, which was true until Skylark placed 9th in 2025.

The story behind every Queen quartet is unique; every winning quartet consists of 4 singers who kept competing until they earned their crown.

My final disclaimer is that all the data is observational, not experimental; correlative, not causal. While finding these patterns is interesting and, I hope, insightful, I cannot definitively state that certain choices cause certain outcomes. For example, suppose I find that quartets who perform the song “My Foolish Heart” outscore quartets who perform the song “I've Got a Feeling I'm Falling”. First, it could simply be by random chance that one song outscores another; this will be addressed when assessing the statistical significance of my tests. What still remains is that I cannot say whether more skilled quartets choose certain songs or whether certain songs are better vehicles for a quartet to score highly. While this distinction is easy to overlook, I will do my best to use careful language when discussing the results of the analyses.

IV. Dataset

The dataset consists of the Semifinal and Final scores from 1979 to 2025, excluding 2020 and 2021 when the contest was cancelled due to the COVID-19 pandemic; the compiled dataset is available on Github at [1] along with the R code used to perform the cleaning and statistical tests. While the documentation publicly available on the SAI website is not perfectly complete, the rank, quartet name, and score for each stage of the competition was generally recorded. Song titles started to be recorded in 2001, but the scoring of individual songs was not regularly recorded until 2008. In total, the dataset consists of 1772 quartet appearances, although 50 of those appearances are quartets who withdrew from the competition and did not compete that year. In total, there are 696 differently-named quartets, with 45 Queens.

From this dataset, we must begin by getting oriented with typical scores, so that later discussion of the difference in score can be taken in the appropriate context. As shown in Fig. 1, the scores did increase through the 1980s and 1990s. Whether this is due to changes in judging standards or rising quality of competitors is not known by the author. The latter hypothesis is supported by the fact that the 10th place quartets have increased in score more than the 1st place quartets, implying that the competition is becoming more competitive. Regardless of the reason, for certain analyses scores prior to the year 2009 will be excluded to minimize the confounding effect of time and increase applicability of results to today's competition. Throughout this paper, translating between Rank and Score will also be very helpful. For recent competition scores, the relationship between score and rank is shown in Fig. 2. Referring back to this figure will be helpful, for instance, to know that a quartet generally needs to score near 1400 points in the semifinals to win the competition or that a score of 1200 points places them near the middle of the pack.

As the competition score sheets are generated by hand, numerous inconsistencies were cleaned before analysis began. For quartet names, errors were typically punctuation ("Brava", "Braval"), foreign words ("Lady Acapella", "Lady A Cappella"), or abbreviations ("Jazzmin", "Jazzmin, Inc."). As quartets must register their name, there is no risk of repeating names at the same competition. If a quartet name appeared in subsequent years, I assumed it is the same quartet. If the same name

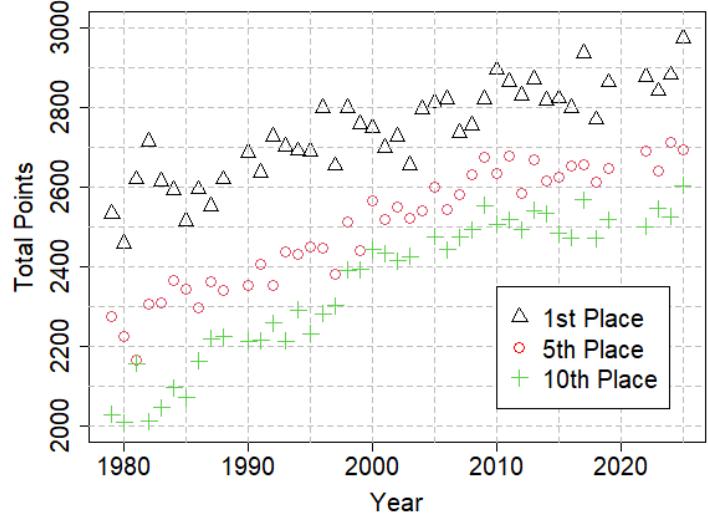


Fig. 1 Combined SF and F Scores over Time

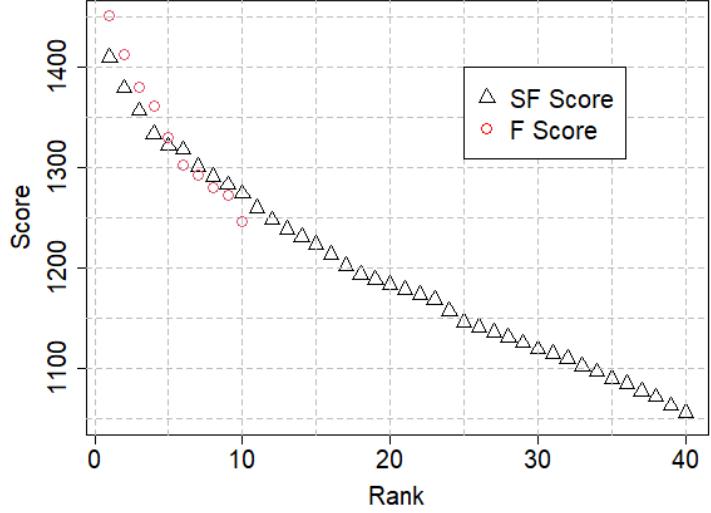


Fig. 2 Average Score by Rank for Years after 2009

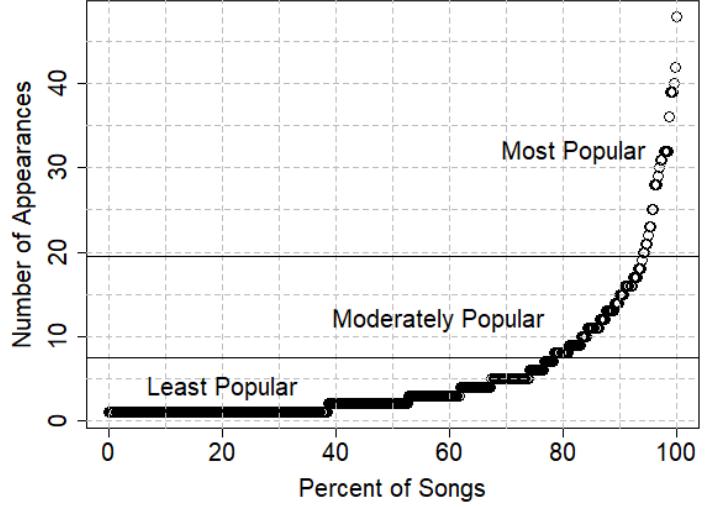


Fig. 3 Song Popularity across all unique titles

appeared after more than 5 years of absence from the competition, I treat it as a new quartet. In total, 696 unique quartets appear across the dataset. Song titles were significantly more difficult to clean, first due to the volume of songs (2437 entries) and second due to the more intricate ways the song name could be recorded. Since some songs are very popular, titles are routinely abbreviated, for example the song “I’m Dancing This Dance in the Wrong Romance” could be recorded as “Wrong Romance”, “Dancing This Dance”, or some other version that is easily identifiable to fellow barbershoppers, but would appear as completely different via simple text matching. Over 200 regular expressions were used to match misspellings and abbreviations to the full titles, resulting in a song list of 444 unique titles. The songs were then tagged with a “popularity” group for later analysis; group “A” for more than 19 appearances, group “B” for between 8 to 19 appearances, and group “C” for less than 8 appearances. These arbitrary cutoffs were selected to have an equal number of appearances in each group, as shown in Fig. 3.

The list of members of the Coronet Club is available on the SAI website and the members of each competing quartet is printed in the convention program each year; by cross-referencing these lists, I tabulated how many Queens were competing in new quartets. As convention programs are only available online for recent years, a relatively small number of quartet appearances (5 years, $n = 232$) include data for the number of queens. Names were sometimes changed between sources (e.g., “Cindy”, “Cynthia”), but using last names helped to identify those situations. Changes to last name were more difficult to identify, except for two cases where the individual’s new last name was hyphenated with their previous last name.

V. Analyses

One distinction needs to be made clear before diving into the data analysis. I use the term “Queen” to refer to past winners of the competition and will specify “Future Queens” as quartets that go on to win the competition. Quartets that have never won the competition will be referred to as “non-Queens”, even if one of their members is a past winner. When displaying plots, I will use the corresponding abbreviations (Q, FQ, NQ) to refer to each population. I will also use the term “number of Queens” or N_{Queens} to refer to how many members of a quartet are past winners, regardless of whether that quartet is a FQ or NQ.

A. Patterns in Future Queen Quartets

Fig. 4 shows the distribution of rank the year before FQ won their crowns. Unsurprisingly, the quartet which scores 2nd place is the most likely quartet to win the next competition. Despite the obvious nature of this observation, seeing the distribution in this way validates the assumption and allows more precise conclusions. For half of the competitions, the winner placed 2nd in the previous year and 90% of the time, the winning quartet was in the Top 5 the previous year. This plot does not show the two times that a quartet won during their first appearance at competition, “Ambiance” in 1986 and “the BUZZ” in 2004.

Fig. 5 shows how many times a FQ competed. The weakness of this plot is that it cannot account for quartets competing under a different name. For example, “Lady A Cappella” competed twice before winning their crowns in 2023; however, 3 of the 4 members of that quartet competed together under the name “Glamour” for 4 years preceding the name change. Despite this limitation, we can still learn that 50% of queen quartets earn their crown within 4 appearances on the international stage and 90% are within 7 appearances. This benchmark, 4 to 7 appearances, is a useful metric for considering the longevity of quartets.

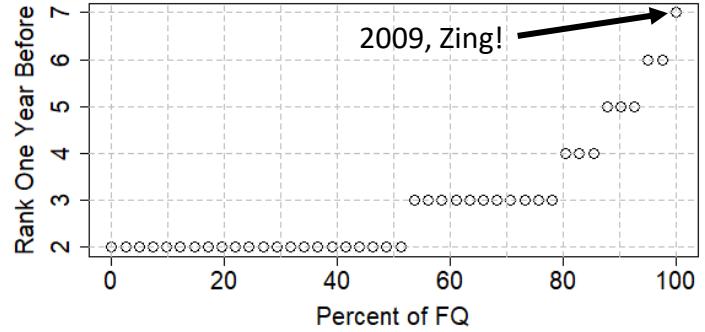


Fig. 4 Rank of FQ the Year Before Winning Crown

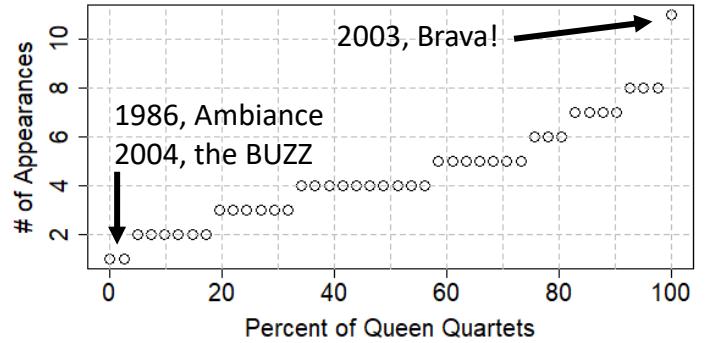


Fig. 5 Total Number of Appearances for FQ

B. Observations between Future Queens and Non-Queens

The first “analysis” is simply displaying some trends that show how quartets that go on to win the competition differ from quartets that have not won the competition. Two trends, Figs. 6 and 7, were immediately apparent: FQ scored higher and ranked higher at their initial appearance. Both figures show a startling difference between the two populations: 71% of FQ proceeded to the finals at their first appearance at the competition. The distribution of rank for NQ, as would be expected, is fairly uniform, with most new quartets placing between 20th and 40th place. Fig. 7 tells the same story, that FQ quartets start with a median score 211 points above other new competitors. A visual inspection clearly shows that FQ and NQ are pulled from different populations; however, for completeness, a two-sided t-test was performed on the SF score data. The results of that test are:

$$t = 5.3, df = 9.5, p = 0.00038$$

Given this dissimilar scores between FQ and NQ, I was driven to investigate how quartets’ scores changed with each subsequent performance to see if this deficit could be overcome. The results of that investigation will be explained in Sections V-C and V-D.

Figure 8 shows the max number of appearances between FQ and NQ, i.e., the total number of times a quartet competed. For FQ, this is an alternative way of displaying the same data in Fig. 4. The majority of non-winning quartets (82%) only compete 1, 2, or 3 times. However, even if a quartet stops competing, each member of that quartet may go on to join new quartets. Tracking individual singers would be a fantastic next step to analyzing the intricate re-shuffling of competitors that occurs as new quartets are formed. However, such an analysis is outside the scope of this paper. Another observation from Fig. 4 is that, while most NQ quartets only appear briefly on the International stage, some do go on for lengthy “careers”. 27 different quartets have at least 8 appearances.

C. Change in Each Quartet’s Performance Over Time

Fig. 9 traces the path of each quartet’s change in score from their first appearance. While the plot is difficult to draw firm conclusions, it does give some insight. First, the overall trend of the cloud is upward, matching the intuition that quartets would use the year to practice, improve, and then score higher at the next competition. The maximum observed improvement, around 350 points, is unsurprising given the scores in Fig. 2: if a quartet started near the bottom of the competition (rank 40 ~ 1000 points) they would ascend to the top (rank 3 ~ 1350 points) with

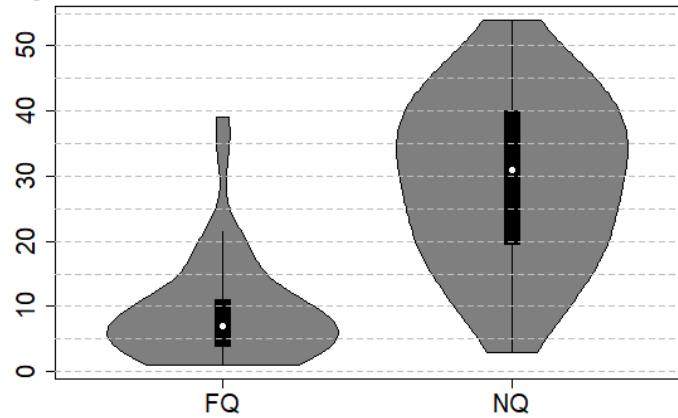


Fig. 6 Quartet Rank at First Appearance

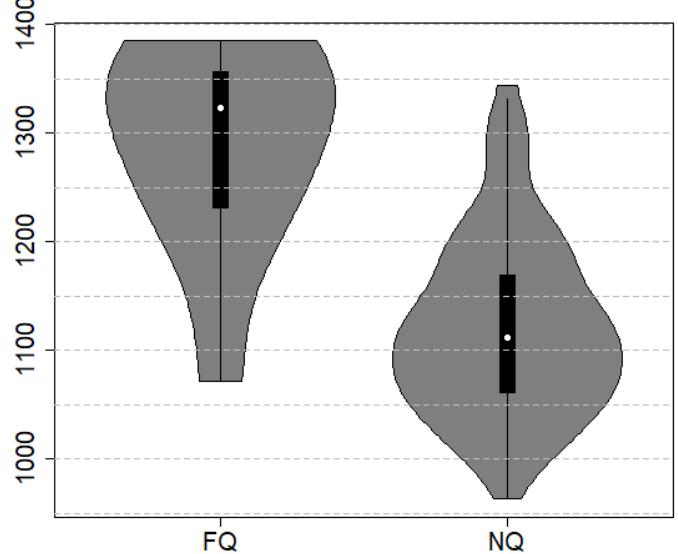


Fig. 7 Quartet SF Score at First Appearance, 2010+

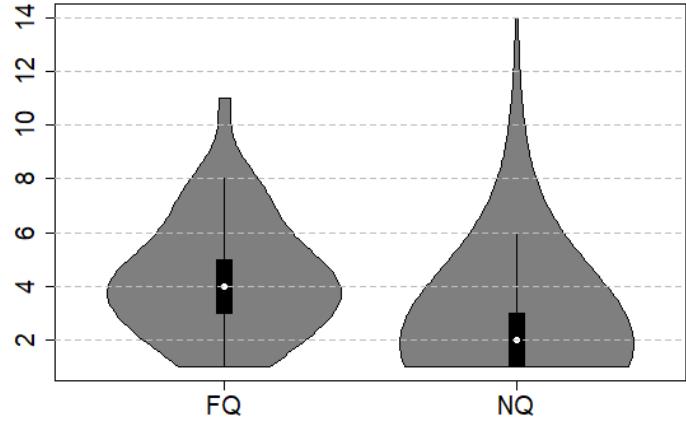


Fig. 8 Number of Appearances by each quartet

a Δ of that magnitude. Second, the longevity of quartets is displayed – very few lines trace out further than 8 appearances, concurring with Fig. 8. Both initial observations can be refined with better plots, Figs. 8 and 10. For Fig. 10, the Queen quartets were removed, as they were outliers in the 200~300 range, but their removal has very little impact on the quartile values plotted in the box plot. After 5 or 6 appearances, the scores for quartets stop improving; however, trends beyond 7 or 8 years are difficult to determine because the sample size for that many appearances is so small. The median improvement for all quartets at their 6th appearance is 94 points, with 56 different quartets having competed at least 6 times.

Similarly, the change in rank can be plotted, shown in Fig. 11. This plot tells a similar story as Fig. 10; that after a quartet's first appearance in the competition, the majority never improve more than 5 places in rank. The median value for appearance number 5, 6, and 7 is -5, with only the top 25% of quartets improving more than 10 places from their initial score. Given this ceiling on improvement, quartets must typically start with a score of 1300 and rank within the top 10 to eventually win the competition (SF score of 1400, rank of 1).

D. Modeling the Effect of Previous Winners

Before fitting a model to the data, one more interaction needs to be examined. Fig 12 shows the semifinal score for quartets that have no Queens as members and quartets that do. The difference between the medians of the two populations is 103 points. Unsurprisingly, having past winners (i.e., highly skilled members) in a quartet improves scores. As before, visual inspection alone is sufficient to say that the two populations are distinct, and a two-tailed t-test can confirm this assertion. The results of the t-test are:

$$t = 5.9, df = 24.8, p = 4.2E - 6$$

Strictly speaking, a t-test is not appropriate, as repeated quartet appearances are treated as independent by this test. However, removing repeated appearances would not significantly change the outcome.

By fitting a model to the data, more precise statements can be tested than merely looking at plots that display one-dimensional trends, as was done previously. I fit a linear model that predicts a quartet's change in score (ΔSF) based on several parameters: the score at the quartet's first appearance (SF_{first}), how many times the quartet has competed ($N_{appearance}$), and

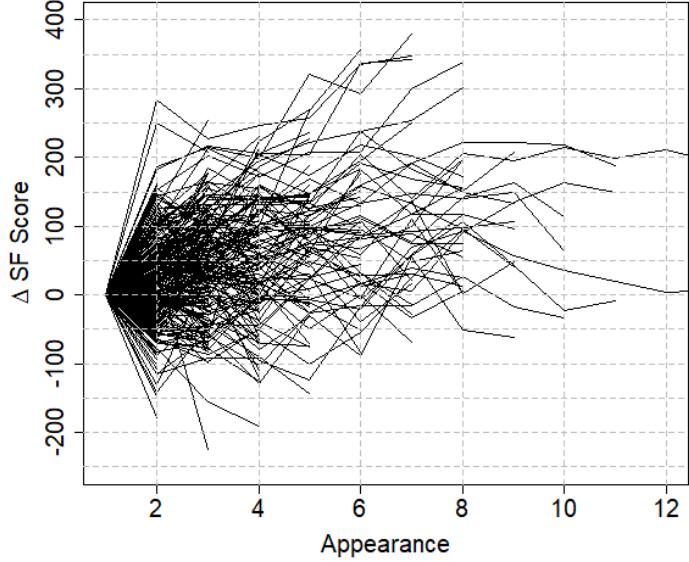


Fig. 9 Change in SF Score for all quartets by appearance

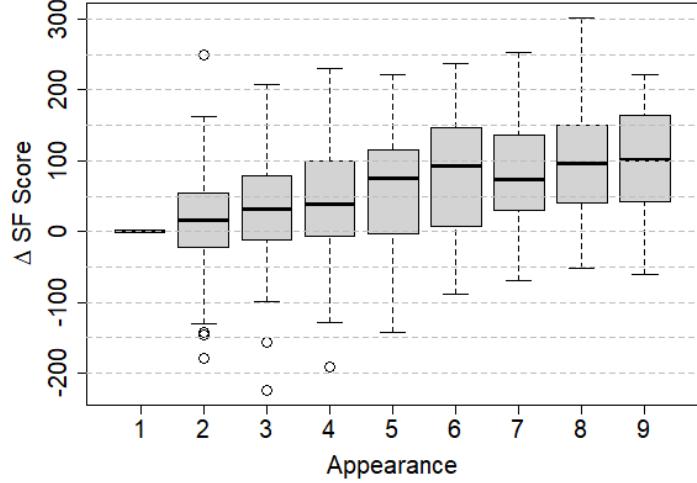


Fig. 10 Change in SF Score for NQ quartets by appearance

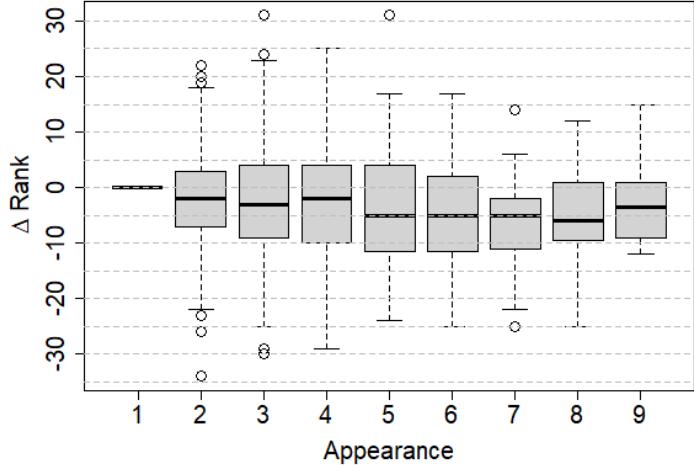


Fig. 11 Change in Rank for NQ quartets by appearance

whether the quartet has a Queen as a member ($N_{Queen} > 0$). Furthermore, I added an interaction term for appearances and queen membership, as I believe that quartets with queens will differ in their response over time. That is, I believe that quartets with past winners will have already reached higher levels of potential, so will not improve as much over time. The model coefficients are listed in Table 1, along with the significance of each coefficient. Interpreting these coefficients, we see:

- $\beta_0 = 352$, generally quartets increase from their initial score.
- $\beta_1 = -0.3$, increasing initial score is associated with smaller (more negative) increases in future scores. This matches our intuition of having a ceiling, so quartets with high initial scores have less room to improve.
- $\beta_2 = 16.5$, with each subsequent appearance, a quartet improves their score by about 16 points.
- $\beta_3 = 46.6$, quartets with queens ($N_{Queens} > 0$) have higher increases in score; however, this coefficient is on the boundary of having statistical significance.
- $\beta_4 = -13.4$, quartets with queens do not improve as much with subsequent appearances, however this coefficient does not have statistical significance.

Using these coefficients, predictions can be made to compare two imaginary quartets in similar situations. For example, suppose two quartets return to competition after their debut the previous year. The first quartet had an initial score of 1250 and has one queen as a member. The second quartet had an initial score of 1150 and does not have any queens. The predicted ΔSF for each quartet would be +41.3 and +37.0, for scores of 1291 and 1187. If those same quartets returned for a third year, the predicted ΔSF would be +44.3 and +53.4, for scores of 1294 and 1203. An additional note on the interaction term: the presence of a Queen is an initial boost (+33) to the second appearance, but over time, the effect diminishes. At 4 appearances, the two effects nearly cancel each other out. These effects can be better seen by the predictions made in Table 3.

Restricting model to only terms with statistical significance at the 0.05 level results in dropping any mention of the predictor with the number of queens in a quartet. This is partially due to the fact that the number of observations is so much smaller in the former, as I have individual membership data for only the most recent 5 contests. The new model (Table 2) shows the same trend in Initial Score and Appearance and a slightly smaller intercept.

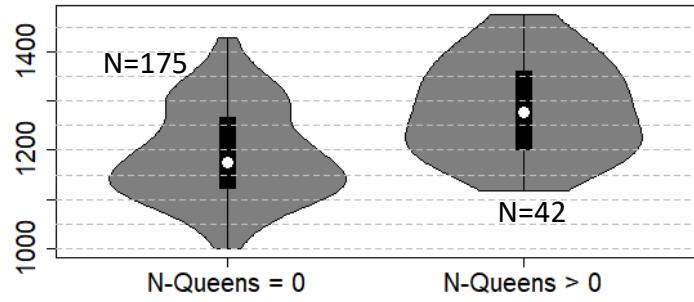


Fig. 12 SF Score for all quartets by number of queens, 2010+

Table 1. Model Fit to Predict ΔSF

$df = 144, R^2 = 0.22$	Coefficient, β	Std. Error, σ	p value
Intercept	352.45	87.46	9E-5
Initial Score	-0.289	0.07	1E-4
Appearance	16.46	3.61	1E-5
$N_{Queens} > 0$	46.59	26.43	0.08
Appearance·($N_Q > 0$)	-13.42	8.95	0.14

Table 2. Model Fit to Predict ΔSF

$df = 991, R^2 = 0.19$	Coefficient, β	Std. Error, σ	p value
Intercept	288.09	28.14	< 2E-16
Initial Score	-0.25	0.03	< 2E-16
Appearance	13.61	1.19	< 2E-16

Table 3. Predicted SF Scores (and Associated Rank) According to Model Fit

	Q1, $N_Q = 0$	Q1, $N_Q > 0$	Q1, $N_Q = 0$	Q1, $N_Q > 0$
Appearance 1	1150, 25 th	1150, 25 th	1250, 12 th	1250, 12 th
Appearance 2	1187, 19 th	1220, 15 th	1258, 11 th	1291, 8 th
Appearance 3	1203, 17 th	1223, 15 th	1275, 10 th	1294, 8 th
Appearance 4	1220, 15 th	1226, 15 th	1291, 8 th	1297, 7 th

E. Determining the “Pop Song” Effect

To investigate whether the song performed by a quartet influences its score, the most important step is eliminating the bias caused by quartet skill. Eliminating this bias was done by choosing instances where a quartet performed songs of differing popularity in the same set and taking the difference between the two songs’ scores. This occurred over four hundred times; the distribution of the difference in score is shown in Fig. 13. The less popular score was always subtracted from the more popular, and as the plot shows, the average difference is zero. In other words, the second song is just as likely to score higher as score lower, there is no correlation between score and song popularity. We can confirm that assertion with a paired sample t-test, with results shown in Table 4. With such high p-values, there is no evidence to support the claim that the populations differ. Although this result is “unexciting” in that there is no statistically significant difference, the result is quite exciting for quartets who agonize over song selection in the hope of swaying the judges.

This lack of effect of song popularity can also be displayed by looking at quartets who performed the same song popularity for both songs – if there was an effect of popularity on score, we would expect it to be amplified when both songs from a set are from the same group. However, as before, there is no difference between the distributions (Fig. 14) and the lack of difference is supported by another round of t-tests (Table 5). This test is not as solid evidence as the previous test (Fig. 13 and Table 4), as there could be a bias between quartet skill and song selection, i.e. higher performing quartets could (purposefully or accidentally) choose certain songs. Both tests tell the same story: the popularity of songs does not influence quartet scores.

There are two major sources of potential error in this analysis. First is the data cleaning, which was done by hand and could have mislabeled some songs, muddying the waters and reducing the effect size. If songs which should have gone into a more popular group went uncorrected, they would likely be in the C group. For example, there was one appearance of a song recorded as “Georgia”, but it is unknown if the title should be corrected to “Sentimental Gentleman From Georgia”, a popular song with 37 appearances, corrected to “Georgia May”, a low popularity song with 5 appearances, or kept as-is. If either of the latter two cases, then the song was correctly placed into the C group, but if the first case, then one data point was mislabeled. However, if this was a significant source of error, we would still expect to see some

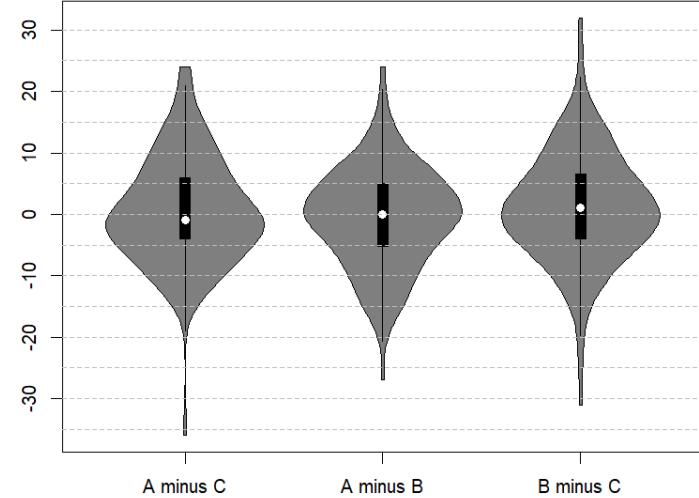


Fig. 13 Difference in Score for Sets with Differing Popularities

Table 4. Results of Paired Sample t-Test for Sets with Differing Popularity

Null Hypothesis: $\mu_{AC} = \mu_{AB} = \mu_{BC} = 0$			
Population	Number of Data Points	t-value	p-value
AC	110	0.56	0.58
AB	176	0.89	0.37
BC	123	1.39	0.17

Table 5. Results of Independent Samples t-test on Single Popularity Pairs

Null Hypothesis: $\mu_{AA} = \mu_{BB} = \mu_{CC} = 0$			
Population	Number of Data Points	t-value	p-value
AA	108	1.26	0.21
BB	71	0.64	0.52
CC	85	0.61	0.55

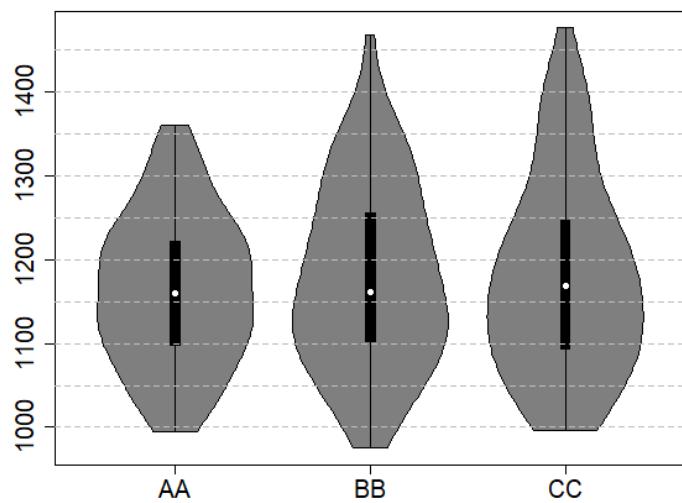


Fig. 14 Scores for Sets with Single Popularity Pairs

8

effect, especially between the A and B groups, as most mistakes would occur with few appearances and be placed in the C group. The second source of error in the analysis is that there could still be correlations between songs and score, but that the common denominator is not popularity. For example, there could be a connection between tempo of the song and score, but as I do not have the beats per minute (BPM) of each song in my dataset, I would not be able to test that hypothesis. I also could look for particular songs that are over-achievers, but, given the large number of songs in the dataset, odds are that some of them would randomly achieve above-average scores. To avoid p-hacking, I chose not to test songs on the individual level. As previously mentioned, the dataset is available online to anyone who wishes to tabulate other factors (tempo, up tune or ballad, composer) and run their own analyses.

F. Quartet Longevity

My final research question was regarding any relationship between quartet longevity (the total number of appearances) and the rank. As previously shown (Fig. 11), most quartets change their rank by only a few places over the course of all their international appearances, so analyzing quartets based on their initial rank is a good place to start. Figure 15 has binned all quartets into groups of 3 (initial ranks of 1-3, 4-6, 7-9, etc.) and the associated distribution of total appearances of those quartets. The groups on the far right (initial rank > 30) are mayflies – they might not qualify at their regional competition the following year, so it's a coin-flip to determine if they will have a second or third appearance at the international competition. The middle group (initial rank 15 to 30) are solid competitors who are likely to win their regional contests and return the following year. The high ranks are split into two separate groups: moderately high (initial rank 7 to 15) are the longest-lived quartets, and the highest (initial rank ≤ 6) are shorter lived (since they are the population that goes on to win and leave the competition!).

Another way of looking at quartet longevity is simply asking “what is the probability that the quartet returns to the competition?” The observed probability of future performances for each rank is shown in Fig. 16. Of the quartets that score in the Top 10, nearly 90% of them return to compete again, with that percentage falling steadily with rank, hitting nearly 30% for ranks 35 and beyond. The trend line, plotted in blue in Fig. 16, has the equation of:

$$p = 0.958 - 0.0109 \cdot Rank - 0.000389 \cdot Rank^2 + 0.00000680 \cdot Rank^3$$

Together, Fig. 16 shows that high ranking quartets are more likely to return to the competition and Fig. 15 shows how that trend plays out for quartets over their entire lifespan.

The next step in this analysis is adding other factors, such as change in score or rank, to predict if a quartet returns. For example, we might expect that if a quartet earns a lower score than the year before, they would be more discouraged and less likely to return. Likelihood of an event occurring is a classic binomial situation, which can be fit with a Generalized Linear Model. After looking at various predictors, the strongest was certainly Rank, followed by Appearance (number of times appearing at internationals), followed by Change in Score from Previous Appearance. For a quartet's first appearance, the change in score is set to zero. As many of the predictors are

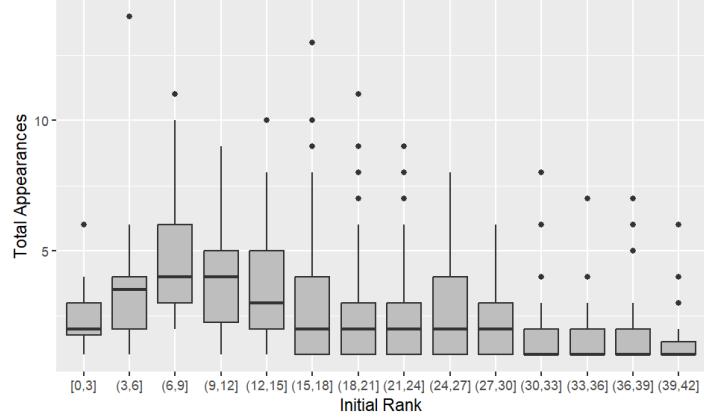


Fig. 15 Distribution of Quartet Appearances by Initial Rank

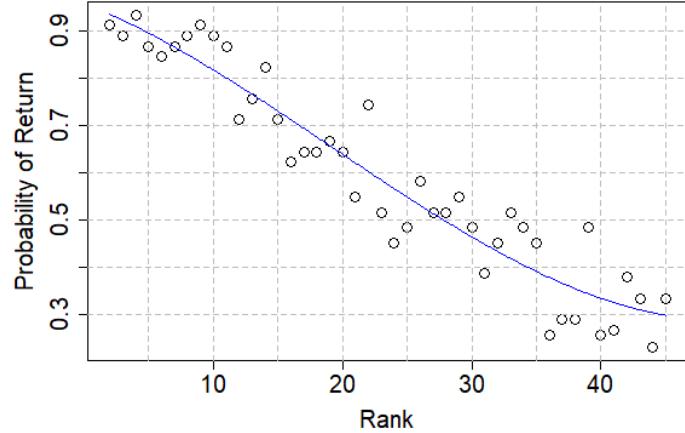


Fig. 16 Observed Probability of Return by Rank

correlated (e.g., rank and score are directly related), the minimum number of predictors was used in the model. The model has the form:

$$\hat{p} = \frac{1}{1 + e^{-X_i\beta}}$$

Where \hat{p} is the predicted probability of a quartet's return, X_i is the quartet's predictors, and β is the model coefficients listed in Table 6. As with the previous model, using it to compare hypothetical quartets allows us to see the effect of the interactions between predictors. More generally, the value of the predictors tell us the trends:

- $\beta_0 = 2.76$, the “default” probability (all predictors set to zero, which is non-sensical for Rank and Appearances) is 94%. According to the trendline for Fig. 16, a quartet with a Rank of zero would have a return probability of 96%, so this is a good sign of matching our observed data.
- $\beta_1 = -0.080$, as Rank increases, the probability of return decreases.
- $\beta_2 = -0.134$, as number of Appearances increases, the probability of return decreases. Although the magnitude of this coefficient is larger than β_1 , that does not mean that it is a larger contributor than Rank, since Rank varies over a larger range. Appearances is typically between 1 to 4, but Rank can be anywhere between 2 and 50.
- $\beta_3 = 0.004$, as change in score from the previous year increases, the probability of return increases. The same caveat regarding coefficient magnitude applies.

To demonstrate the trends, the probability of return was used to make predictions on fictitious quartets; the predictions are listed in Table 7.

Table 7. Predicted Probability of Return for Fictitious Quartets

Rank	Appearance	SF-Delta-Previous	Probability of Return, \hat{p}
2	2	+50	92.7%
2	2	-50	89.3%
25	2	+50	67.0%
25	2	-50	57.1%
5	1	0	90.2%
35	1	0	45.9%
10	1	0	86.1%
10	5	0	78.4%

VI. Conclusion

I will begin my concluding remarks by returning to my research questions and a short summary of my findings:

1. *Do quartets containing a Queen of Harmony tend to score higher?*

Yes. Based on recent data (2019 to 2025, $n = 262$), the median score of quartets that contain at least one Queen is 110 points (15 ranks) higher than the median score for quartets without a Queen.

2. *Does performing a common song (a barbershop “standard”) influence score?*

No. I found no statistical difference in score between songs that were performed often, performed moderately often, or performed rarely. There are potentially factors other than popularity that could influence score, but those analyses could not be performed with my current dataset.

3. *Do quartets who “make a splash”, i.e. score highly at their first international competition, achieve more success at future competitions?*

Yes. At a future Queen’s first international appearance, 50% are rank 7 or higher and 90% are rank 19 or higher. Considering all quartets’ first appearance, a rank of 7 or higher is only achieved by 6.6% of competitors and of those 6.6%, half go on to win the competition. For first international appearances, the median score of quartets that eventually earn a crown is 211 points (25 ranks) higher than the median score of quartets that have not earned a crown.

4. *Is there a relationship between rank and the longevity of a quartet?*

Yes. The longest-lived quartets are within the Top 15 in their first appearance. The probability of a quartet returning to competition is most strongly predicted with rank, with some effect due to change in score from the previous year and total number of appearances.

One major limitation of the analyses contained in this paper is that all data is considered at the quartet level, not the individual level, which skews some conclusions. For example, considering the performance at first appearance (Fig. 6 and Fig. 7), there could be a confounding factor of age or years competing that would further explain the

discrepancy; it is by no means the first time each quartet member has competed at Internationals. Continuing this line of reasoning, it is feasible that a quartet whose members have more years of competitive barbershop experience would perform better; however, the data available is insufficient to investigate this effect. Furthermore, it would be dishonest to point at the dynasty of a high-profile performer, such as Debbie Cleveland who won the competition in 1993, 2004, and 2025, and make claims that the singers on top stay on top and individuals have no hope of climbing the ladder. The data shows that quartets tend to increase their score by 16 points per year and plateaus after 5 or 6 years (Fig. 10), but I simply do not have the data to make any claims for how individual's scores increase or decrease across the years.

I now return to my disclaimers in Section III and remember that the goals of SAI reach far beyond deciding who earns the honor of wearing crowns for the next year. Emphasizing the journey over the destination is cliché, but is very relevant in this competition where everyone is truly an “amateur” (from Latin “amator”, one who loves); all members of SAI are bound together by their shared passion for the artform. 1 in every 15 quartets who performs on the international stage will be immortalized as Queens of Harmony, but every single quartet forms indelible memories with lifelong friends.

Acknowledgments

I would like to thank all the members of Saving Grace: Elise Zufall, Kieley Morales, Marisa Garza, Kelly Leary, Mikayla Wallis, and Andréa Walker for sharing their beautiful music and love of barbershop with me. Cheering you on from the sidelines has been a joy and I deeply respect the authenticity you bring to both your songs and your relationships within the barbershop community. Wishing you all the best for your future musical endeavors both on and off the stage!

References

- [1] The Github repository with my sources, dataset, and R code: https://github.com/azufall/SAI_Data_Analysis/tree/main
- [2] Score sheets were pulled from the SAI website: <https://sweetadelines.com/competition-results>
- [3] Additional data was found on the Barbershop wiki: https://www.barbershopwiki.com/wiki/Barbershop_Wiki_Project

Errata

For those who wish to continue the work of archiving competition data, I have compiled a list of missing scores:

- Results from 1947 to 1979
- 1989, 2009, 2010, 2015 SF results for the top 10
 - However, the SF totals were included in the F results for all except the 1989 contest
- 1979 and 1981 through 1984 SF for ranks after 21

For those who wish to extend this work to tracking individual-level rather than quartet-level performance, I would recommend starting with acquiring competition programs for more years. Contest years 2019 through 2025 can be found online and I have also included copies on my Github repository.

Fun Facts

I cannot help but share additional observations from the dataset, please enjoy.

The number of past winners in the champion quartet, for many years, was zero due to the low number of individuals in the Coronet Club. The first woman to win a second crown was Jarmela Speta, a member of “The Nota-belles” in 1955 and “Rarities” in 1970. The first woman to win a third crown was Connie Noble, a member of “4th Edition” in 1972, “Savvy” in 1988, and “Fanatix” in 2001.

The blue line of Fig. 17 is a rolling average of the number of past winners in the champion quartet. Although this could be a case of seeing patterns in random data, there does appear to be a cyclic rise and fall in the number of queens. If this trend continues, there could be an argument for “generations” of competitors that join together, win together, and retire together.

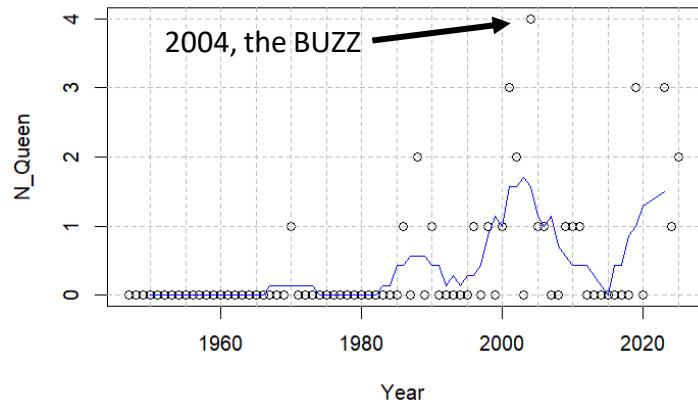


Fig. 17 Number of Past Winners in Champion Quartet by Year

Before I made the plot in Fig. 17, I was expecting to see the trend line rise over time (as more members in the Coronet Club meant more past winners to join quartets) and then eventually level off to some constant value (as the number of actively competing members of the Coronet Club reached steady state). More data is needed to determine if the pattern will continue to be cyclic, if an equilibrium will be reached, or if the process is stochastic.

All songs with more than 30 appearances are listed in Table 8. Assuming each song is an independent event (which is not true, since a quartet never repeats a song in their set) and assuming that you, as a member of the audience, listens to 100 total songs during the competition (45 SF sets with 2 songs each, 10 F sets with 1 contestable song each), the expected number of times you will hear the song is listed in the third column of the table. Additional data mining would need to be done to tabulate similar statistics for the chorus competition.

There have been 3 quartets with the name “Fusion”, which competed in 1989, 2009, and 2016-2017 and “Viva!” which competed in 2001, 2007, and 2018-2019. It

is unknown to the author if any of these quartets share members or if they are completely unrelated.

The quartets with the most appearances are listed in Table 9. Some notable members of the list are “Brava!”, which is a champion quartet, “Chord Company”, which is the lowest ranked (bucking the trend), and “Voce” which is the only member of the list that is still actively competing.

Table 9. Quartets with 9 or more appearances

Quartet Name	Appearances	Years*	Highest Rank
Starshine	14	1979-1989, 1993-1995	2, 1986
Razzcals	13	1997-2010	2, 2004
Brava!	11	1992-1994, 1996-2003	1, 2003
Night Magic	11	1993, 1995-2004	3, 1997 & 2000
Shimmer!	11	1997-1998, 2001, 2004-2012	9, 2007
Blue Razzberry Rhythm	10	1991-2000	3, 1992 & 1993
Symmetry	10	1995-2004	7, 1998
Remix	10	2000-2005, 2008-2013	4, 2011
Harmony Highlights	9	1979-1987	6, 1986
Tivoli	9	1980, 1982-1989	6, 1983
Rainbow Collection	9	1985, 1987-1994	5, 1991 & 1992
Grand Prix	9	1990-1998	7, 1990
Chord Company	9	1998-2003, 2006-2008	19, 2006
Spritzer	9	2009-2018	4, 2017
Voce	9	2015-2019, 2022-2025	5, 2025

*note, the numbers of years may be greater than the number of appearances if a quartet withdrew from competition. Withdrawals are not counted as an “Appearance”.