Figure Skating Scoring Trends and Predictors

Diana Zhu

May 13 2022

Abstract

In the competitive figure skating world, the fate of a figure skater's Olympic dream rests on their performance during a few minutes on the ice. However, behind the artistry there is a quantitative side to the sport that can be leveraged by both athletes and sports analysts to better understand the scoring. Through predictive models and correlation testing, this study aims to identify the most important factors that affect a skater's overall total score and evaluate potential judging bias. From these models, we found that there is a high correlation between the mean Grade of Execution and Program Component Scores with the overall final score. Additionally, skaters likely experience an advantage when competing in their home country at the World Championships whereas judges are not likely to be biased in ranking skaters from their home country.

1 Introduction

This paper seeks to analyze figure skating competition data from the past Olympic cycle in order to understand the 2022 World Championship results and evaluate the significant factors that help identify champions. We also seek to evaluate potential judging bias and explore a potential correlation between individual judge rankings and the final placements.

1.1 The Structure of Figure Skating

There are four main disciplines in figure skating: men's singles, ladies' singles, pairs, and ice dance. Mens and Ladies disciplines are individual events whereas pairs and ice dance teams are partnered events, where each team consists of a man and a woman.

Ladies & Men's Singles

The Ladies and Mens disciplines of figure skating requires an individual to perform jumps, spins, and step sequences. Jumps comprise the majority of points in singles skating. Because of the higher element values of quadruple jumps, it is natural for men's scores to be higher than ladies.

Pairs & Ice Dance

The Pairs and Ice Dance disciplines of figure skating are partnered events. Pair teams perform technical elements such as side-by-side jumps, spins, throw twists, and pair lifts. Many elements of pair skating involve the man throwing or lifting the woman into the air. Unlike ice dance and similar to singles skating, pair skaters perform jumps. The Ice Dancing discipline is loosely based on ballroom dancing, and technical elements of an ice dance program focus on the precision and quality of step sequences.

International Judging System (IJS)

Each skater or team will compete a short program (SP) and a free program or free skate (FP, FS) set to music. Skaters usually perform the same programs throughout a skating season, occasionally opting to switch out elements for more challenging ones as the season progresses. The total score that determines final rankings is the sum of the short and free programs. The two segments of an ice dance competition are called the rhythm dance (RD) (formerly short dance) and the free dance (FD). Under the current IJS scoring system, each program is broken down into two segment scores: the Total Element Score, which represents the Grade of Execution (GOE) of each element performed, and the Total Program Component Score, which represents performance quality such as skating skills and interpretation of the music.

Under the International Judging System, judges grade the quality of each element using a grade of execution score within a range of -5 to +5, which is added to or deducted from the base value. GOEs are proportional to the base value of each element. The highest and lowest scores for each element are omitted, and the remaining scores are averaged to determine the final GOE for each element. The GOE is then added to or subtracted from the base value

Rank	Name	NOC Code	Starting Number	Total Segment Score	Total Element Score	Total Program Component Score (factored)	Total Deductions
1	CHEN Nathan	USA	28	113.97	65.98	47.99	0.00

#	Executed Elements	Info	Base Value	GOE	J1	J2	J3	J4	J5	J6	J7	J8	J9	Ref.	Scores of Panel
1	4F		11.00	4.40	4	4	5	4	4	4	4	4	4		15.40
2	3A		8.00	2.29	3	3	3	3	4	2	2	2	4		10.29
3	CCSp4		3.20	1.05	4	4	3	4	3	3	3	3	3		4.25
4	4Lz+3T		17.27 x	3.94	4	4	4	3	4	2	2	3	4		21.21
5	StSq4		3.90	1.95	5	5	5	5	5	5	4	5	5		5.85
6	FSSp4		3.00	1.03	3	4	4	3	4	3	3	3	4		4.03
7	CCoSp4		3.50	1.45	4	4	4	5	4	4	4	5	4		4.95
			49.87												**65.98**

Program Components	**Factor**	J1	J2	J3	J4	J5	J6	J7	J8	J9	Scores of Panel
Skating Skills	1.00	9.75	9.25	9.50	9.75	9.50	9.75	9.50	9.50	9.50	9.57
Transitions	1.00	9.75	9.25	9.25	9.50	9.50	9.50	9.25	9.00	9.50	9.39
Performance	1.00	10.00	9.50	9.75	10.00	9.50	9.75	9.50	9.50	10.00	9.71
Composition	1.00	9.75	9.50	10.00	9.75	9.75	9.75	9.25	9.50	9.75	9.68
Interpretation of the Music	1.00	10.00	9.75	10.00	9.75	9.50	9.50	9.50	9.50	9.50	9.64
Judges Total Program Component Score (factored)											**47.99**
Deductions:											**0.00**

Figure 1: Nathan Chen's Free Skate scoring breakdown from the 2022 Olympic Winter Games. The top half details the technical score, and the bottom half details the program components score.

for each element, and the sum of the scores for all elements forms the technical score.

The judges will award points on a scale from 0.25 to 10 for five program components to grade overall presentation. The highest and lowest scores for each component are removed, and the remaining scores are averaged. The components score is then multiplied by a factor to ensure that the technical and program components scores are balanced.

1.2 Close-Up of a Protocol

We can examine a protocol from Nathan Chen's free skate performance at the 2022 Olympic Winter Games, as shown in Figure 1 and calculate the method of scoring an element using the U.S. Figure Skating Scoring Cheat Sheet. The fourth element Chen performed was a quadruple lutz-triple toe loop combination jump, denoted as "4Lz+3T". The base value of this jump is the sum of the value of a quadruple lutz and a triple toe loop: $11.5 + 4.2 = 15.7$. Athletes receive a 10 percent bonus on all jumps performed in the second half of the program (a rule designed to reward skaters who perform jumps when they are more fatigues), so Chen receives $15.7 * 1.1 = 17.27$ on the base value of this element.

The GOE mark for a single element is equal to the trimmed mean of the judges' GOE scores multiplied by 10% of the base value. In a jump combination, the GOE mark for each element is weighted based on the unweighted base value of the more difficult jump. In the case of Chen's 4Lz 3T combination, the trimmed average is $24/7 = 3.4285$, resulting in an overall GOE score of $3.4285 * 0.1 * 11.5 = 3.94$. Therefore the "Scores of Panel" for this element of the sum of the base value and the GOE: $17.27 + 3.94 = 21.21$. We will discuss trimmed averages later on in this study.

The World Championships

The World Championships are the most important event in the regular skating season. They first took place in 1896, but it was not until 2005 that the current scoring system was introduced. The number of skaters allowed to compete from each country per discipline is determined by the results of the previous season's World Championships. Each International Skating Union (ISU) member nation automatically receives at least one spot in each discipline, provided they have a skater/team who fulfills the technical minimum scores.

2 Methods

2.1 Data Collection

We scraped our data using the R package "rvest" from `skatingscores.com`, a website that re-publishes scoring data from major international competitions made public by the ISU. We chose the top free skate scores from 15 distinct skaters at the past two senior world championships (2022 and 2021). Due to the volatile nature of the sport and short-lived seasons of skaters, we focus on skaters that performed well in the past two years.

It is also important to note that due to the Russia-Ukraine conflict in recent months, the International Skating Union banned Russian athletes from participating in any international figure skating competitions on or after March 1, 2022, including the 2022 World Championships. Therefore, there are no scores from Russian athletes from this year's World Championships.

2.2 Predictive Models and Analysis

2.2.1 Linear Regression Model

We can start by modeling a skater's score with a simple linear regression, using the skater's previous personal best total score as a predictor. We can also include the mean Grade of Execution (GOE) mark as an additional predictor because well-performed elements generally indicate a higher chance of placing well. We can break the total score into its two segments: the short program and free program.

Each segment is normally distributed based on the predictors, previous personal best (h_{short}) and mean GOE (goe_i).

$$\text{short}_i \sim N(\beta_{0,short} + \beta_{1,short} * h_{short} + \beta_2 * \text{goe}_i, \sigma^2_{short})$$

$$\text{long}_i \sim N(\beta_{0,long} + \beta_{1,long} * h_{long} + \beta_2 * \text{goe}_i, \sigma^2_{long})$$

2.2.2 Mixed Effect Model

The mixed effects model accounts for both fixed effects and random effects, and we choose to deploy this model because the skaters have competed in different numbers of competitions. Because our sample size n for each skater is not consistent across skaters, we can implement a random intercepts model.

The general form of the random intercepts model is:

$$Y_{ij} = \alpha_j +_{ij} + \epsilon_{ij}; \epsilon_{ij} \sim N(0, \sigma^2_Y); i = 1, ..., n_j \text{ and } j = 1, ..., J$$

$$\alpha_j \sim N(\mu_\alpha, \sigma^2_\alpha); j = 1, ..., J$$

where j represents group number and i represents observations number within the jth group. We did not have the chance to implement this method, but this is how we would pursue accounting for the difference in how many times each skater shows up in the data set.

2.2.3 Spearman Rank Correlation

The Spearman rank correlation will allow us to determine to what extent an individual judge's scores align with the final placement of a skater. We want to use the Spearman rank correlation rather than the typical Pearson correlation because the Pearson correlation simply evaluates the linear relationship between two continuous variables, whereas the Spearman rank correlation evaluates the monotonic relationship based on the ranked values for each variable, so we use it when we have two ranked variables and we want to see whether the two variables covary.

A high correlation would indicate a "good" judge, whereas a low correlation would suggest a "biased" or "unusual" judge. Letting d_i be the difference between two ranks of each observation and n be the number of observations, the Spearman's rank correlation coefficient ρ can be expressed as follows:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

This correlation can be calculated for each competition segment and represents the relationship between the ranking of skaters given by one judge and the overall final competition rankings. The overall ranking is a trimmed mean of judges' scores: the highest and lowest scores from the panel of nine judges are dropped.

After calculating the correlation coefficients for each judge at the 2022 World Championships, we can construct a bootstrap distribution to reveal whether any of the coefficients are significantly low.

2.2.4 Home-Town Advantage Regression

We can use a Bayesian generalized linear model to check for a correlation between skaters competing in their home country and their final score. We are modeling a skater's total score as the outcome and factoring in whether they were competing in their home country

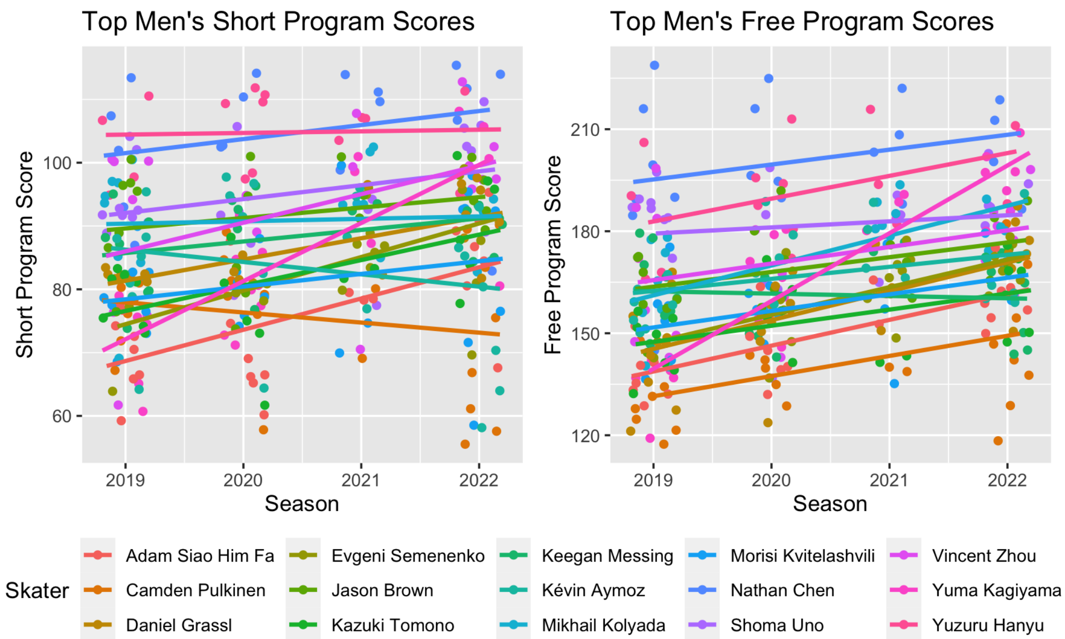

Figure 2: Trends in Men's Scores From 2018-2022

or not, using the following formula:

$$y = f(\beta_0 + \beta_1 * \text{home})$$

3 Results

Initial Data Analysis We begin by examining the overarching trends of male figure skater scores in the past Olympic cycle. Jitter has been applied to the x-axes of Figure 2, in order to give some randomness to the competitions within each season. Nathan Chen and Yuzuru Hanyu dominate the men's discipline, unlike the women's and pairs disciplines where there are no consistent champions. We see that other than Chen and Hanyu, there is significant movement across seasons and high variability in scores. More of the initial visualizations of the skater scores across the other disciplines are described in the appendix.

Regression Models

From the linear regression of men's free program scores over the past four years, we found an intercept coefficient of 32.3652, Mean GOE coefficient of 7.0387, and Mean PCS coefficient of 4.9022. Each of the p-values corresponding to these coefficients are extremely small and much lower than the 0.05 threshold, indicating significance - although we should be hesitant to make any conclusions from this statistic. These coefficients are smaller and approximately halved in magnitude compared to the free program, which is expected because the short program scores are meant to be worth half of what the long program is worth. Again, the p-values are extremely small ($p < 0.05$).

Judging Correlations

As we see in Table 1, Spearman rank correlation calculations indicate that judges generally agree with each other. The Italian judge has the lowest correlation of 0.979, while the Swedish, Turkish, and French judges have the highest correlation at 0.987.

Home-Town Advantage

As seen in Figure 3, the Intercept value of 218 refers to the average World Championships score for skaters not competing in their home country, where the data used is World Championship data in the last 10 years.

The Beta value for the Home variable is 21, which means that men competing in their home country receive, on average, a score 21 points higher than those who are not competing in their home country.

We are 95% confident the true value of the total score for a male athlete competing outside of his home country lies between (214, 221), and the true difference in value of a

Home Country	Coefficient
CAN	0.987
SWE	0.987
TUR	0.987
FRA	0.983
KAZ	0.983
AZE	0.981
EST	0.981
AUT	0.980
ITA	0.979

Table 1: Spearman Rank Correlation between individual judge rankings for skaters and final actual ranking.

Regression of Host Country
How Competing in Their Home Country Affects Mens' Scores

Characteristic	Beta	95% CI[1]
(Intercept)	218	214, 222
Home	21	7.5, 34

[1] CI = Confidence Interval

Regression of Host Country
How Competing in Their Home Country Affects Pairs Scores

Characteristic	Beta	95% CI[1]
(Intercept)	176	172, 180
Home	8.9	-4.3, 22

[1] CI = Confidence Interval

Regression of Host Country
How Competing in Their Home Country Affects Ladies Scores

Characteristic	Beta	95% CI[1]
(Intercept)	162	159, 165
Home	22	11, 34

[1] CI = Confidence Interval

Regression of Host Country
How Competing in Their Home Country Affects Dance Scores

Characteristic	Beta	95% CI[1]
(Intercept)	164	161, 166
Home	9.3	-0.73, 20

[1] CI = Confidence Interval

Figure 3: Summary tables of conducting a regression on whether there is a relationship between skater score and whether they are competing in their home country.

male athlete competing at home is between (8.1,34).

We see a similar trend in the women's discipline. The Intercept value of 162 refers to the average World Championships score for female skaters not competing in their home country. The Beta value for the Home variable is 22, which means that ladies competing in their home country receive, on average, a score 22 points higher than those who are not competing in their home country. Again, the confidence interval does not contain 0, which is a strong indication that there is an advantage for women competing in their home country.

4 Discussion

There are a few sources of potential bias in the design of this study that we should consider before drawing any conclusions. The first source of bias comes from our data collection method. We chose to analyze data from the top 15 performers in the last two World Championships, which introduces a bias into the scores these skaters receive across the Olympic cycle. By cherry picking, our data shows a gradual increase in scores that is not necessarily representative of the scores of 15 randomly selected skaters at the World Championships. In addition to higher average scores, the top skaters are also more likely to have smaller variance within their scores as they are more consistent performers. The linear mixed effects model is particularly susceptible to bias because the multicollinearity and constant variance assumptions for the model were violated. In future studies, this could be addressed by instead randomly selecting a subset of competitors at the World Championships. This method would still allow for a relatively comprehensive data set since we can guarantee that the skater will have qualified for the World Championships, but is able to prevent additional bias that results from only looking at top skaters.

It is also important to consider that there could be a natural score inflation over time due to the gradually increasing difficulty level of elements being performed in competition. Harder jumps are associated with higher point values, which will in turn increase the Total Element Score of a protocol.

Additionally, there is a bias that is introduced when calculating the Spearman Rank Correlation coefficients between judge rankings. A single judge's ranking and the overall ranking are not two independent measures, which implies that we cannot conclude any p-values indicating significance directly calculated from these correlation coefficients. In order to draw conclusions about whether a given judge is partial to their country, future studies should randomly sample two judges from the judging panel and calculate their correlation coefficients with the overall rankings across many competitions. This measurement would put into perspective how extreme our results are and provide a basis to make conclusions.

The advantages of a skater performing in the host country also comes with caveats - top skaters tend to represent wealthy and large countries, which in turn are more likely to host international sporting competitions. Thus, these regressions we run on home-country advantage could represent marginal associates that are confounded with the strength of the host countries.

5 Conclusion

Our results support the idea that mean GOE and PCS are strongly predictive of a skater's final score and ranking. These conclusions must be further supported by additional analysis and modeling, especially taking into account the potential sources of bias mentioned in the Discussion.

6 Appendix

Discipline	Names
Men	Chen, Uno, Kagiyama, Hanyu, Pulkinen, Zhou, Kvitelashvili, Kolyada, Messing, Siao Him Fa, Semenenko, Brown, Grassl, Tomono, Aymoz
Women	Sakamoto, Trusova, Shcherbakova, Hendrickx, Tuktamysheva, Liu, Bell, Chen, Gubanova, You, Lee, Mikutina, Tennell, Kihira, Ryabova
Pairs	Mishina, Sui, Knierim, Tarasova, Boikova, Moore-Towers, James, Peng, Miura, Della Monica, Dafina, Hase, Cain, Ziegler, Walsh
Dance	Papadakis, Sinitsina, Hubbell, Gilles, Chock, Stepanova, Guignard, Fear, Fournier, Smart, Hawayek, Zagorski, Hurtado, Wang, Lajoie

Table 2: All Skaters Chosen for Analysis

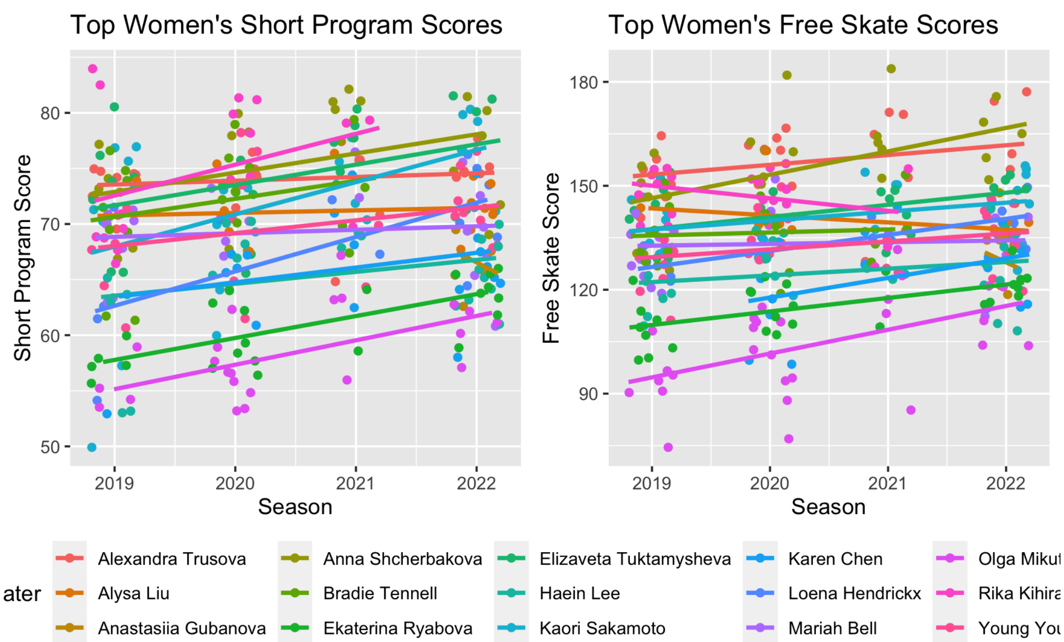

Figure 4: Women's Scores Trends

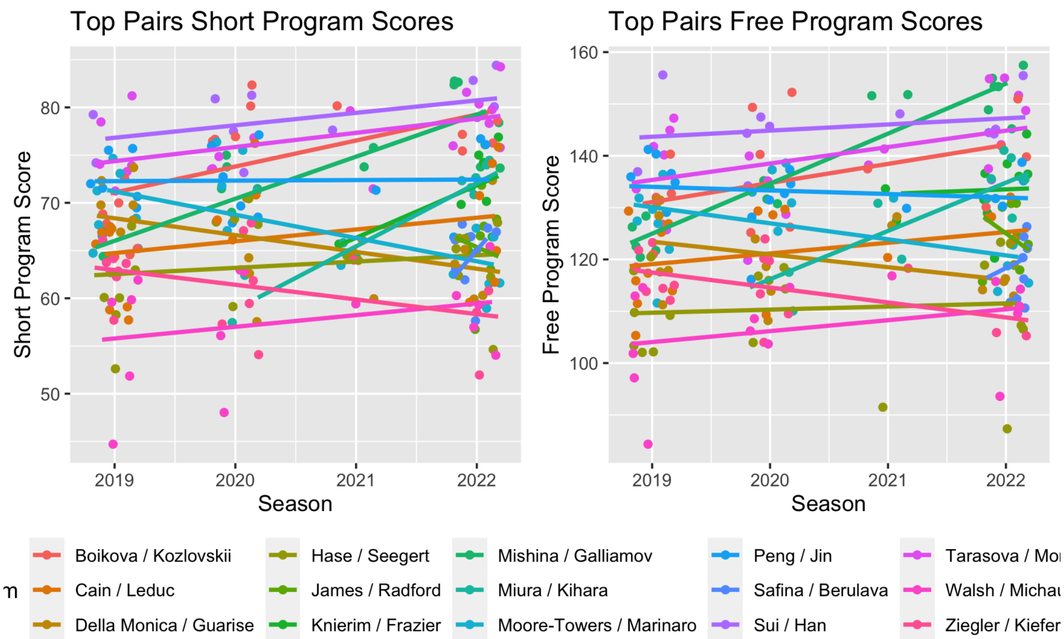

Figure 5: Pairs Scores Trends

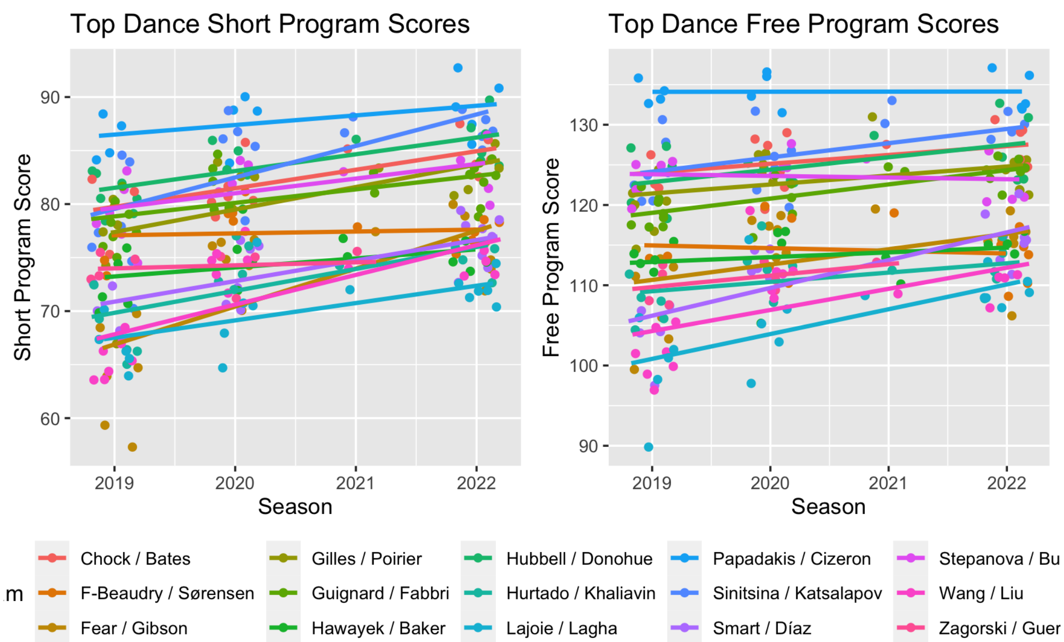

Figure 6: Ice Dance Scores Trends

7 References

https://usfigureskating.org/sites/default/files/media-files/Scoring%20Cheat%20Sheet.pdf

https://www.usfigureskating.org/about/scoring-system#:~:text=In%20the%20IJS%2C%20competitors%20accumulate,level%20through%20program%20component%20scores.

https://dash.harvard.edu/bitstream/handle/1/39011778/ZHU-SENIORTHESIS-2018.pdf?sequence=3&isAllowed=y