

MG-220 Final Project

From: Travis Kerr

To: Dr. Thomas Craig

Date: May 2nd, 2022

Subject: Analysis of Politics overtaking Sports

Section 1: Introduction/Overview

For as long as I can remember, Baseball has been my “go-to” to clear my mind and help me relax from anything bothering me. Baseball, still being my primary choice, has allowed me to find a greater love for the world of Sports as a whole. Over the past years, Sports seemed to be the outlet for people to chat about daily and wanted to be constantly involved. Nowadays, every topic you talk about requires you to speak upon your political thoughts and ideas that come along with this topic as well. Being a massive Sports fan, this has grown an annoyance to hear Politics talked about every day within a Sports environment when I expect players to speak about their respected sport and not their political views. As a result, collecting data to show that Politics is taking over the Sports should be paid more attention to because these worlds should stay alone in their respective states and should not interfere, along with all other parts of the world.

Section 2: Statement of Hypothesis

In collecting and studying two datasets off the website Statista, with the data collected by Nielsen, 270towin.com, the US Census Bureau, the United States Election Project, ProCon, and FairVote, I am interested in looking to see the number of household TV ratings of Super Bowls from 1990-2022 and the number of people who voted in elections from 1824-2020. Whether you are a die-hard sports fan, everyone always seems to watch the Super Bowl every year in February. The game is between the winner of each conference, and Super Bowl I took place during the 1966-67 NFL season (Carroll). In my other dataset, I would like to mention that Voting-Age Population started being collected in 1932, and Voting-Eligible Population began being collected in 1980. The following hypotheses will be tested and used to determine the trends and relationships in the data I collected.

Hypotheses:

Null Hypothesis (H_0): There is no association between the Total population, the Total votes cast, and the Voting-eligible population in the U.S. Presidential Elections.

Alternative Hypothesis (H_1): There is an association between the Total population, Total votes cast, and the Voting-eligible population in the U.S Presidential Elections.

Null Hypothesis (H_0): There is no association between Views (Millions) and the Year of each Super Bowl from 1990 to 2022.

Alternative Hypothesis (H_1): There is an association between Views (Millions) and the Year of each Super Bowl from 1990 to 2022.

Section 3: Data and Analysis Methodology

The datasets I collected and studied using the website Statista which Nielsen, 270towin.com, the US Census Bureau, the United States Election Project, ProCon, and FairVote collected by survey, include links on the last page of this memo. I used another website called Britannica to gain more information about the Super Bowl itself. The first dataset I found was titled “Voting populations and number of votes cast in U.S. presidential elections from 1824 to 2020,” which displays the Total population, Total votes cast, Voting-age population, Votes cast for winner, Voting-eligible population, and Votes cast for runner-up for each election, in millions, from 1824 to 2020. The Votes cast for runner-up were the votes for the most popular runner-up. The second dataset I found was titled “TV ratings of the Super Bowl in the United States from 1990 to 2022” displaying the number, in millions, of views of each Super Bowl from 1990 to 2022.

To test the hypotheses in Section 2, I will be using a multiple regression analysis test to determine if there is an association between the Total population, Total votes cast, and Voting-eligible population during the most recent presidential election in 2020 and use the same test to analyze if there is an association between the household TV views during the Super Bowl from 1990 to 2020 and each year it happens. In performing these tests, I used the application Minitab, with the data collected and formatted using an Excel worksheet. To complete these tests using the Minitab application, I transferred the data from each dataset to a separate Excel spreadsheet

by adding a column title called Year, replacing “0” for each of the years that Voting-age Population and Voting-eligible Population were not recorded because the total would be 0 regardless. In my other dataset, I in a column title called Year as well as a “Views (Millions)” column to correspond to the views for each Super Bowl. For each dataset as well, in each Year column, I converted each pasted “year” to a number and typed the full year of each to help it filter its category on Minitab.

Section 4: Results

Multiple Regression Analysis: Total population vs. Total votes cast, Voting-eligible population

In discovering an association between Total population, Total votes cast, and Voting-eligible population, it was found that there is significantly specific data to show that there is no association between Total population, Total votes cast, and Voting-eligible population. While the P-Value for the Total votes cast is 0.00, the P-Value for Voting-eligible population is 0.223, greater than 0.05, indicating that we fail to reject the null hypothesis.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	30.10	3.42	8.80	0.000	
Total votes cast	2.2819	0.0940	24.26	0.000	3.65
Voting-eligible population	-0.0622	0.0504	-1.24	0.223	3.65

With this, 97.66% of the variability of this in the Total population is explained by the Total votes cast and the Voting-eligible population. This stat was very shocking because I was expected to have an R-sq of about 60% of influence from the Total population, but the data below suggests otherwise.

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
15.4052	97.66%	97.57%	97.29%

To conclude, the calculated regression equation is shown below:

Regression Equation

$$\text{Total population} = 30.10 + 2.2819 \text{ Total votes cast} - 0.0622 \text{ Voting-eligible population}$$

This regression equation appears to show a positive correlation between the Total votes cast and Total population, providing that as the Total votes cast in each election increase, so does the Total population, and a negative correlation between the Voting-eligible population and Total population, indicating that the Voting-eligible population decreases, as the Total population increases.

Multiple Regression Analysis: Views (Millions) vs. Year

In discovering an association between Views (Millions) and Year, it was found that there is significantly specific data to show that there is no association between Views (Millions) and Year. The P-Value for the Views (Millions) equaled to 0.945 and the P-Value for Year equaled 0.629, both of which are greater than 0.05, indicating that we fail to reject the null hypothesis.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-7	102	-0.07	0.945	
Year	0.0249	0.0510	0.49	0.629	1.00

With this, 0.76% of the variability of this in the Views (Millions) is explained by the Year. It was surprising to see this because I expected to see that each year brings a Super Bowl early in the year which people are excited to watch, but the data suggests otherwise.

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.79224	0.76%	0.00%	0.00%

To conclude, the calculated regression equation is shown below:

Regression Equation

$$\text{Views (Millions)} = -7 + 0.0249 \text{ Year}$$

This regression equation appears to show a positive correlation between the Views (Millions) and Year, providing that as the Views increase steadily, as each year passes.

Section 5: Discussion and Conclusion

The Sports world is seemingly being taken over by the Political world and it is not something to take nicely. In exploring this issue, I collected data on electoral populations and votes from each election from 1824 to 2020 and household TV views of each Super Bowl from 1990 to 2022. I conducted two multiple regression analysis tests for the two hypotheses I developed in my analysis. After the tests were conducted, there was significant data showing that there was no association between each of my hypotheses. I thought using presidential election data was a good idea to find my hypothesis because the political world revolves around the presidential election every 4 years. If there is an association here, it shows that politics are trying to take over the Sports world and keep increasing its Voting-eligible population. I thought using the Super Bowl was the best way to help my hypothesis because, as I mentioned earlier, even if you are not a major sports fan, everyone enjoys spending most of their Sunday watching NFL football. If there is an association, the views happen from people feeling like they must watch the Super Bowl. The associations were not statistically supported by these multiple regression tests. This was surprising and unexpected because I believed strongly there was an association between these variables but, the stats suggested no significant association. To further a future analysis, more research along with data collection will be needed. Some other examples could be presidential candidate debate views, and views of other major Sports championships like the World Series, NBA Finals, and Stanley Cup. Overall, each of these worlds has its own authority and respect in saying what they want to say if it stays relevant to its own side. Politics needs to keep asking questions to politicians and candidates, not to professional male and female athletes.

Works Cited

- Carroll, Bob. "Super Bowl." Encyclopedia Britannica, 26 Jan. 2022, <https://www.britannica.com/sports/Super-Bowl>. Accessed 28 April 2022.*
- Gough, Christina. "US Super Bowl TV Ratings 2022." Statista, 16 Feb. 2022, <https://www.statista.com/statistics/216538/super-bowl-us-tv-ratings/>.*
- O'Neill, Aaron. "U.S. Presidential Elections: Number of Votes Cast 1824-2020." Statista, 17 Feb. 2021, <https://www.statista.com/statistics/1139763/number-votes-cast-us-presidential-elections/>.*