Final Project
April 24, 2015

Scoring Summary for Hockey, Comparing 1st, 2nd, and 3rd Periods

Introduction:

This observational study will examine data from all of the hockey games played by the teams in the NHL for a Saturday-Saturday week to determine when the most goals are scored and if there is a relationship between the scoring in different periods.

The goal of this study is to determine if there is a notable relationship between the goals scored in the first, second, and third periods. There are a lot of variables that affect scoring in hockey. However, this study will only focus on the period during which the goals are scored to see if there is a trend.

Methodology:

The population of this study is all professional hockey leagues all over the world. The sample is the 30 teams of the National Hockey League (NHL) that includes teams from America and Canada. This is a representative sample because it includes a wide variety of teams from two different countries. All of the teams have different records from the season, and some of them are playing really well, while other teams are not playing as well as they can be. The observational unit is a single game, and the data that is being studied is the quantitative variable of the total number of goals scored each period by both teams.

To analyze the data and compare the total number of goals between periods, this study will use all of the games played during a single week in the regular season. This gives a good sampling of data because all 30 teams will have played at least once during the week, and all of them will be represented in the data.

Data Exploration and Results:

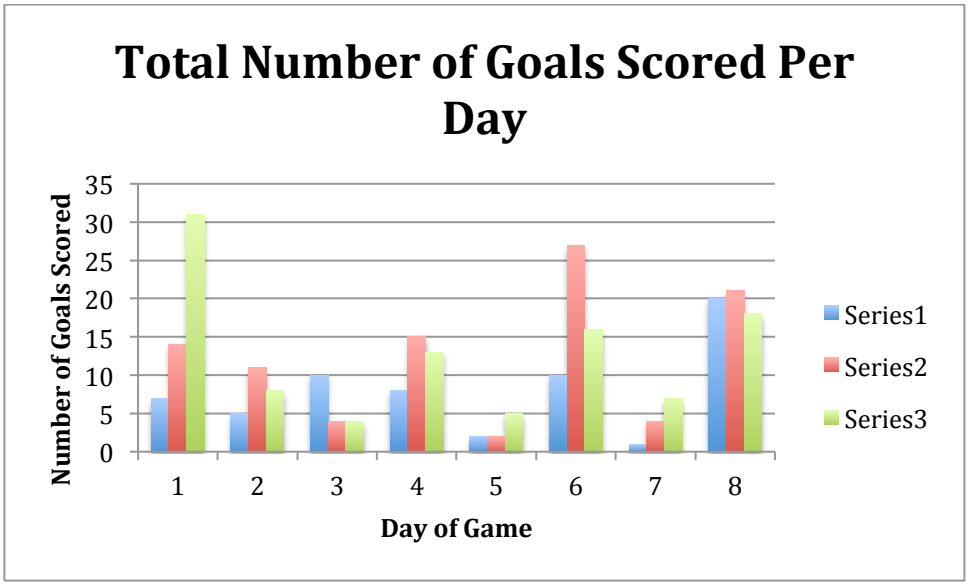

Figure 1 Total Number of Goals Scored Per Day

Figure 1 illustrates the total number of goals scored per day in one week of hockey games. The blue bar shows the number of goals scored in each first period, the red in each second, and the green in each third. Even though the number of games played each game is different, and therefore there are a lower total number of goals scored, there was one day where the most goals were scored in the first period, four days when the most goals were scored in the second, and three days where the most goals were scored in the third. The mean of first period goals is 7.8, the mean of second period goals is 12.25 and the mean of third period goals is 12.75.

	Min	Q1	M	Q3	Max
1st period goals	1	3.5	7.5	10	20
2nd period goals	2	4	12.5	18	27
3rd period goals	4	4	10.5	17	31

Table 1 Five Number Summaries for Goals in Each Period

Table 1 shows the five number summaries for each period, which includes the min, median, max, and 1st and 3rd quartiles. The summaries provide a way to view the trends of each period's goals in comparison to each other in a different visual way than the graph. This summary provides the information from all 8 days into a single visualization.

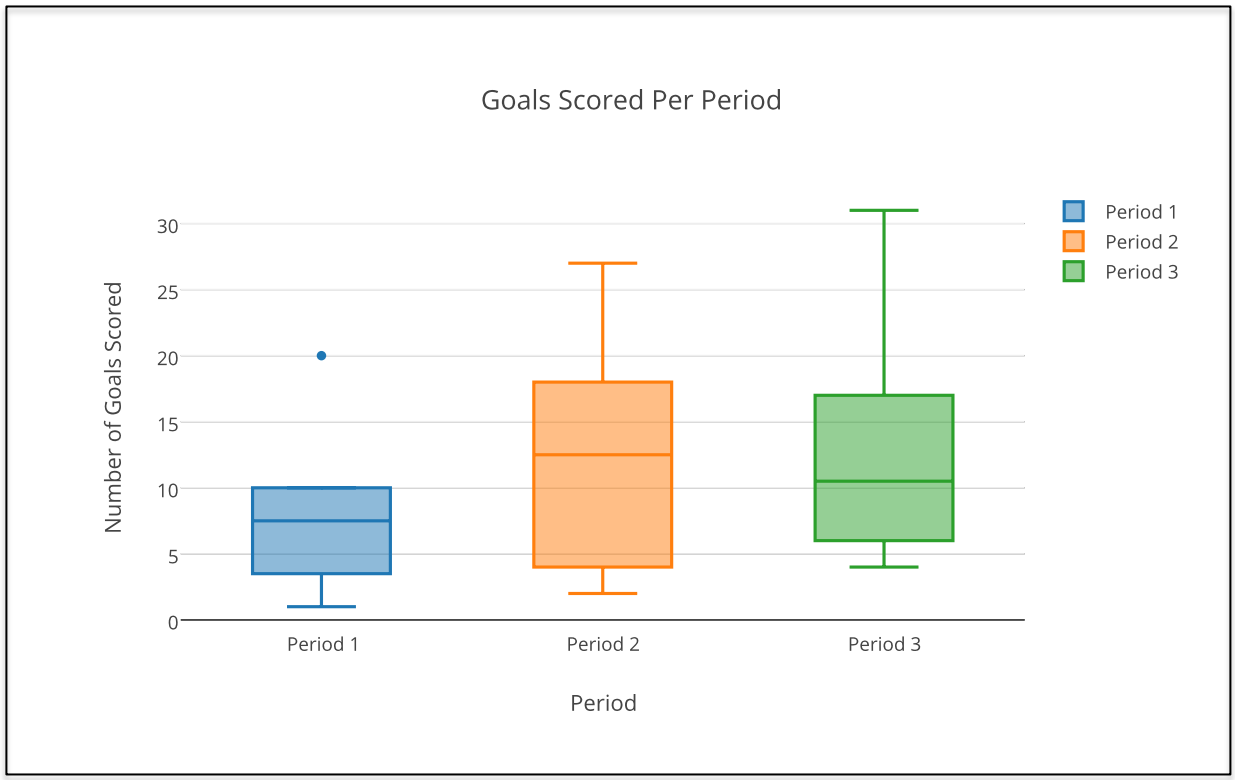

Figure 2 Boxplots Compare the Spreads of the Three Periods

Figure 2 compares the spread of the goals scored in each period for all 61 games played over a period of 8 days. The spread for the first period goals shows that 20, the highest number of goals scored in the first period on one day, is an outlier. The spread for first period goals is much smaller than the spread for second and third period goals. The IQR for first period goals was 6.5, for second period goals 14, and for third period goals 13. The boxplot shows that the spread of third period goals is slightly higher than the

spread of second period goals. This is true, for the total number of goals scored were as follows: 1st period: 63 2nd period: 98 3rd period 107

	Visiting Team	Home Team	Total
1st Period Goals	19	44	63
2nd Period Goals	53	45	98
3rd Period Goals	58	49	107
Total	130	138	268

Figure 3: Total Goals Scored by each team in each period

Figure 3 lays out the total number of goals scored in each period by each team. With this information, we can calculate various probabilities from the data. For example, we know that the probability that a goal is scored in the first period is 63/268= .24. We also know that the probability that a goal is scored in the second period given that it was scored by the visiting team is 53/130= .41. Additionally, figure 3 shows us that both teams are more likely to score in the second or third period, but home teams scored more in the first period than visiting teams did.

Conclusion:

The data shows that most of the goals scored in hockey are scored in the second and third periods. This is not a surprising result for this experiment. While teams have equal opportunities to score in all three periods of play, it is more likely that they will score in the later periods of the game. This is due to many reasons. One reason might be that during the first period, the teams are getting used to playing each other and are not desperate to score because there is plenty of playing time left. During the later two periods, however, a lot of time has passed and they want to start scoring more in order to increase

their chances of winning the game. Additionally, visiting teams score more often in later periods because they may not be used to the home teams stadium and atmosphere.

Future Study:

After discovering that most goals are scored in the third periods of hockey games during a weeklong study, future study might study a larger sample to see if the trend continues. For example, one might study data from an entire season of hockey instead of just one week. Additionally, some of the teams played multiple times in the week. Because of this, their statistics were counted more than once. In a future study, one might only count one game for each team. This would change the outcome because the scoring trend for each team would change with each game they played. Furthermore, one could study data from post-season games or data from hockey leagues all over the world instead of just the NHL. It can be expected that the trend of the majority of goals being scored in the third period will continue.