

Bluefield Blue Jays



By:

Drew DiSanto

For this research project I was assigned the Bluefield Blue Jays. They are the Toronto Blue Jays affiliate, who play in the Pioneer League. Bluefield represents the twin cities from West Virginia and Virginia. Bluefield, WV has about double the population of Bluefield, VA (9,869 to 4,714). Both cities see a higher population of females, with around a 55-45 split. Also, both cities have a median age of 42 years old. Median household income in Virginia exceeded West Virginia by around \$9,000. Another interesting number was that the poverty rate in Bluefield, WV was around 21.6%, compared to 10.7% in Bluefield, VA.

The Bluefield Blue Jays compete at Bowen field in Peters park. The park is located in Virginia but operated by the West Virginia city. Bowen field stadium has a capacity of only 3,000, and is also shared with the Bluefield College Rams' baseball team. Bluefield College is a 5-min drive away from the stadium and about 150 feet from the Virginia state border. The tuition for Bluefield College is around \$25k and enrollment is just shy of 700 students. Just across the state line in West Virginia is the Bluefield school district. All of Bluefield's elementary, middle, and high school buildings are within 10 mins from Bowen field.

The Bluefield Blue Jays went through an affiliation swap during this last decade. Previously, they were affiliates with the Baltimore Orioles (1958-2010) and called themselves the Bluefield Orioles. These players were known as the "Baby Birds" or the "Baby O's" in reference to their big league team. Two of the best Baby O's to ever play were Boog Powell (1959) and Cal Ripken Jr. (1978). Both of these players debuted in Bluefield as 17 year olds. Following the 2010 season, the Baltimore Orioles cut ties with Bluefield, and Bluefield got picked up by Toronto before the 2011 season. In 2013, Kevin Pillar became the first Bluefield alumni to play in the majors for Toronto.

The data collection for this research began by scraping the schedule and results for all Bluefield Blue Jays' home games from MiLB.com. This list included wins, losses, teamscore, opponent, opponent score, day, date, and attendance. Next, dummy variables were created for each weekday and month. Following that, columns that tallied wins/losses were developed and

used to calculate the Blue Jays' win percentage. Then, Bluefield weather data was scraped from wunderground.com. Columns for temperature, humidity, wind speed, and precipitation were generated. Finally, dummy variables were produced for different promotions the Blue Jays put on: fireworks, free admission, free hats, big prize (such as a car), various prizes (small miscellaneous prizes), free health screenings. After all the data was collected, the research was able to move forward.

Following data collection, a regression was run to reveal which variables affected attendance. Attendance was the dependent variable, and the independent variables were chosen based on their potential to affect attendance at Blue Jays' games. These are the results of the regression:

Variable	Estimate	Std. Error	T-value	P-value	
opponentBurlington	11.687	129.625	0.09	0.92978	
opponentDanville	-12.782	283.794	-0.045	0.96488	
opponentElizabethton	65.933	106.757	0.618	0.54942	
opponentGreeneville	59.79	94.875	0.63	0.54143	
opponentJohnsonCity	40.74	124.992	0.326	0.75058	
opponentKingsport	-5.406	119.702	-0.045	0.96479	
opponentPrinceton	268.993	92.944	2.894	0.0146	**
opponentPulaski	55.802	201.559	0.277	0.78703	
weekday	9.459	14.743	0.642	0.53427	
month	-58.555	65.493	-0.894	0.39043	
winpct	294.719	1138.745	0.259	0.80056	
temp	-2.31	11.397	-0.203	0.8431	
precip	309.92	302.103	1.026	0.32698	
homeopen	46.964	283.472	0.166	0.87142	
doubleheader	136.28	83.324	1.636	0.1302	
fireworks	293.071	94.292	3.108	0.00996	***
freeadmin	192.212	82.159	2.34	0.03921	**
free_hats	-60.764	95.939	-0.633	0.53945	
big_prize	-74.893	88.7	-0.844	0.41647	
various_prizes	209.542	107.585	1.948	0.07742	*
health_screening	-116.546	81.577	-1.429	0.18087	

As shown in the regression, if the games were against Princeton,

had fireworks, had free admission, or had various prizes, then they were statistically significant at the 10% level or better. Since all of the estimates on these variables were positive, all variables positively impacted attendance at these games. Next, this regression was used to test for heteroskedasticity and autocorrelation. These tests reveal whether the results are accurate or if they need adjustments. Using the Breusch Pagan test for heteroskedasticity, the P-value was found to be statistically insignificant. Therefore, the null hypothesis of homoscedasticity

cannot be rejected. Following Breusch Pagan, the Durbin-Watson test for autocorrelation was ran on the same regression. The Durbin-Watson test produced a statistically significant P-value, therefore the null hypothesis can be rejected and the data may contain autocorrelation. This makes sense, given that the data is a time series. Following all corrections and tests, the final output regression looked very similar:

Variable	Estimate	Std. Error	T-value	P-value	
opponentBurlington	11.6873	120.685	0.0968	0.924595	
opponentDanville	-12.7817	271.4525	-0.0471	0.963288	
opponentElizabethton	65.9326	69.7963	0.9446	0.365128	
opponentGreeneville	59.7901	86.4716	0.6914	0.503619	
opponentJohnsonCity	40.7404	108.6914	0.3748	0.714916	
opponentKingsport	-5.4063	86.9865	-0.0622	0.951557	
opponentPrinceton	268.993	81.7696	3.2896	0.007209	***
opponentPulaski	55.8023	180.4325	0.3093	0.762896	
weekday	9.4594	13.6717	0.6919	0.503345	
month	-58.5549	61.1641	-0.9573	0.358963	
winpct	294.719	1078.762	0.2732	0.78976	
temp	-2.3099	10.7904	-0.2141	0.834413	
precip	309.92	251.6249	1.2317	0.243753	
homeopen	46.9641	220.9227	0.2126	0.835541	
doubleheader	136.28	80.0682	1.702	0.116803	
fireworks	293.071	94.6867	3.0952	0.010192	**
freeadmin	192.212	67.5562	2.8452	0.015932	**
free_hats	-60.7637	77.2014	-0.7871	0.447862	
big_prize	-74.8925	121.409	-0.6169	0.54988	
various_prizes	209.542	91.5705	2.2883	0.042905	**
health_screening	-116.546	80.5821	-1.4463	0.175978	

As in the original regression, Princeton, fireworks, free admission, and various prizes are statistically significant. The promotions are significant at the 5% level, whereas Princeton is significant at the 1% level. Furthermore, each significant variable has a positive impact on attendance. According to this regression, fireworks had the greatest impact on attendance, bringing in an excess of 293 fans.

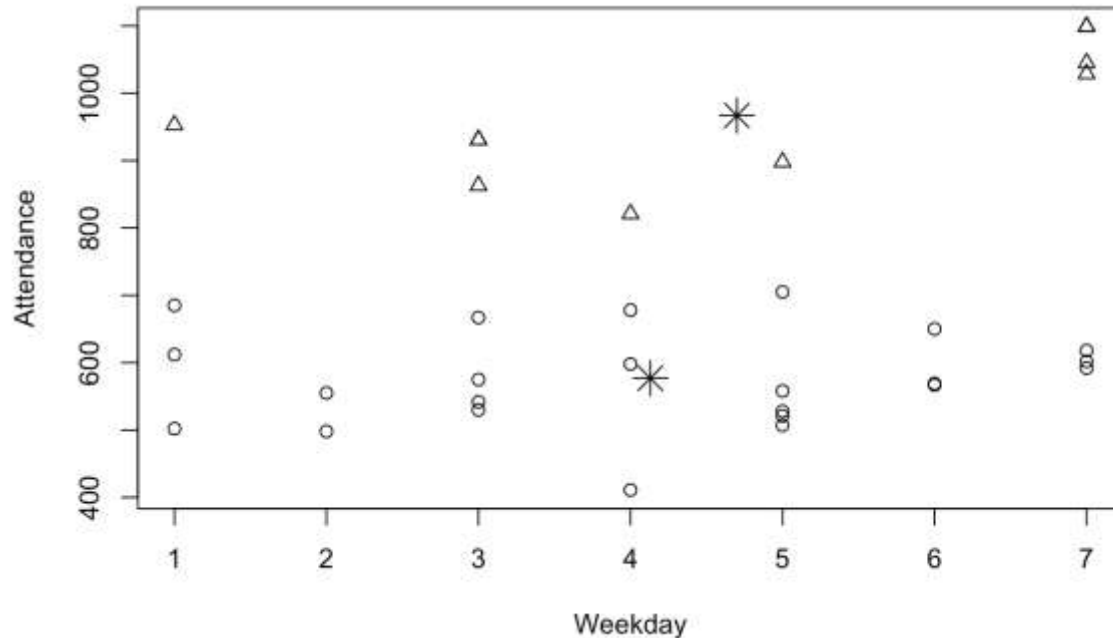
The next portion of this research project was an attempt to look at the significant variables that affect attendance. It also drew a few conclusions as to why/how these variables

positively impacted attendance. When looking at the opponent Princeton, it was initially assumed that there is a rivalry. Upon further research, this assumption was proved to be true. The Princeton Rays are located in WV, only a 12 mile trip to Bluefield. Bluefield's twitter page discusses the Mercer Cup, an annual competition between Bluefield and Princeton that began in 1992. Despite losing the cup this past season, Bluefield leads the competition with a record of 14-12-1. Five out of the seven games Bluefield played against Princeton produced above average attendance (average attendance is usually around 700 fans). So, how did Bluefield take advantage of this rivalry to maximize fans?

Bluefield hosted Princeton 7 times last season, with one doubleheader on 7/16/19. The most attended game was on Saturday, August 10th, where the Blue Jays hosted a firework night as well. As demonstrated above, fireworks had the greatest impact on increasing attendance. Another record high attendance of 1000 fans was the game on July 13th. Bluefield's twitter page mentioned that Mefcor Outdoors hosted a Shark night on July 13th. Mefcor Outdoors sells anything from hunting equipment to kayaks and canoes. This shows how the Blue Jays and Mefcor used the above average attendance at Princeton games to maximize their promotion nights.

Given that Princeton draws large crowds, how can the Blue Jay's obtain above average attendance versus other opponents as well? The answer to this question comes from three other promotion variables. Fireworks were at the #1 and #2 highest attended games. Games with free admission and games with various prizes also drew in significantly higher crowds than the average. The graph below is a cluster model developed in Rstudio. Attendance is on the y-axis, and the days of the week, starting with Sunday, is on the x-axis. The two clusters are represented by Δ and \circ . Triangles are the games with above average attendance and circles are below average games. The $*$ symbol represents the midpoint for each cluster. Looking at

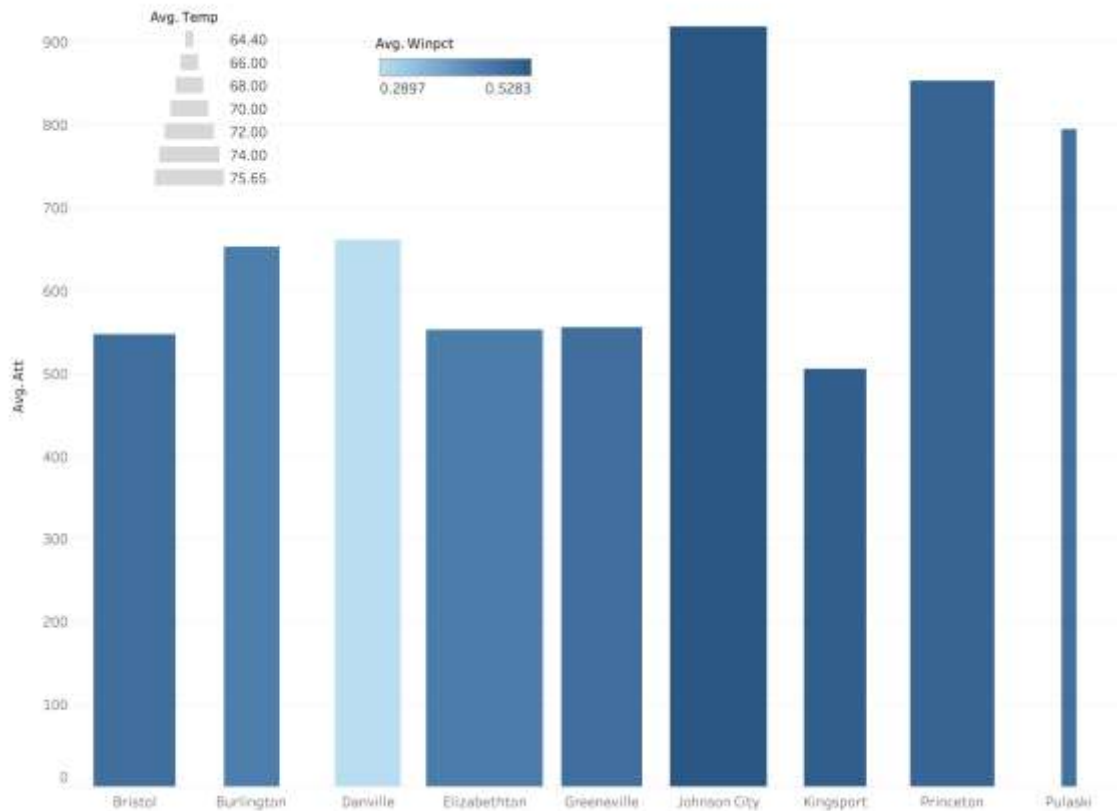
weekday 6 (Friday), it can be observed that not a single game was above average. The only promotion held on a Friday game was a free health screening for fans. Sounds fun right?



In defense of the Blue Jays, one of their Friday games got rained out. There were supposed to be fireworks that Friday, but they got postponed to Saturday after a double-header. However, Fridays have great potential for family activities, and only one promotion night seems like a waste of the Blue Jays' Friday games.

Even though the win percentage variable appeared to be insignificant in the regression above, further research was done to evaluate the accuracy of this insignificance. In the visualization below, the y-axis represents average attendance while the x-axis represents each of the Blue Jays' opponents. The color of each bar signifies the Blue Jays' win percentage versus respective opponents. Finally, the width of the bars represents the average temperature for these games.

Average Attendance, Win Percentage, and Temperature versus each opponent



Although win percentage was not significant in the regression ran earlier, this bar graph demonstrates the opposite. Teams that Bluefield has a higher win percentage against appear to draw larger crowds. This is expected for Princeton regardless of win percentage due to the Mercer Cup rivalry. However, since Danville beat Bluefield on Bowen field quite often, fans didn't seem to enjoy attending those games as much. One way to combat a poor win percentage is to use big promotion nights against the better teams of the Appalachian League. That way, fans come to enjoy the promotions and are entertained regardless of the outcome on the field.

Following these findings and evaluations, it is evident that there are ways Bluefield can increase attendance for future seasons. First is by utilizing the local schools. Bluefield's twitter talks about having college nights where tickets and food are discounted. However, it is known

that a price tag immediately scares off students. A better use of a college night might be free admission, without any discounts on concessions. To further research this idea, financial reports would need to be evaluated to see if concessions are a driving force of revenue.

Another promotion variable that improved attendance was when the team gave out small prizes to their fans. Fans seemed to enjoy being surprised with little giveaways. That being said, Friday games need more fans. So, combining Friday games with little giveaways would help improve this downfall in attendance. In both Bluefield, WV and Bluefield, VA, females make up more of the population. If the team was able to tailor their giveaways for females on certain nights, this may also improve overall attendance. The Blue Jays could give out softball bows, prewrap, change purses, or stickers. Furthermore, since Mother's Day is not during the minor league season, the Blue Jays could have a mother/daughter night and implement an entire feminine theme for the night.

Location also plays a big role in a team's fanbase. Not many teams have it better than the Bluefield Blue Jays; they represent twin cities from neighboring states and intermediate through college level schools, all within 10 minutes of the stadium. Despite this, Bluefield ranks second to last in average attendance amongst all teams in the Appalachian league. A regression model was used to display which variables impact attendance. It showed that Bluefield fans tend to attend games against the Blue Jays' rival Princeton, or when fireworks, free admission, or various giveaway prizes are provided. If the team combined their free admission games with local student nights, they would likely see a spike in attendance. Also, since females represent 55% of the population in WV and VA, targeting females with various prize giveaways might bring more families into the stadium. One final idea could be a Mother's themed night in July/August to incorporate more of the female population. These are just a few suggestions that may help the Bluefield Blue Jays to maximize their average attendance. There are many more components to attendance that may reveal better options if more time and access to additional data is provided.

References:

<https://www.milb.com/bluefield/schedule/2019/fullseason?homeGame=true>

<https://twitter.com/BluefieldJays>

<https://www.wunderground.com/weather/us/va/bluefield>

<http://www.city-data.com/city/Bluefield-West-Virginia.html>

<http://www.city-data.com/city/Bluefield-Virginia.html>