University of Wisconsin: Madison

Stressing Over Football

An analysis by the Baltimore Police Competition Group

Sicheng Chu, Scott Lai, Qianyu Liu, Xi He, and Luke Zheng

Professor Karl Rohe

Introduction

Watching a football game is one of the most popular pastimes in the United States. Americans love gathering for festivities in preparation and celebration of a football game with their local team, so fans can get very emotional about the outcome of football games played by their favorite teams. Fans may feel depressed and upset if their local team loses an important game, which can cause unexpected violence and chaos. After a loss, fans with negative emotions and a little too much to drink can cause violence[4] (McGranaghan, 2015, para. 3). By recording the domestic violence calls of 763 police departments, economists David Card and Gordon Dahl report that domestic violence calls can be increased by about 10 percent after a local football team loses[6] (as cited in Park, 2011, para. 3). City of Baltimore has a relatively high amount of 911 calls compared to other cities in the United States, and the city also has a relatively high crime rate. Baltimore Ravens are a local NFL team and have large numbers of fans, mainly in Baltimore. Therefore, we are interested in whether the outcome of a Baltimore Ravens football game will affect the level of distress or chaos in the city of Baltimore during that game day.

This project is using 911 calls from the Baltimore Police Department (BPD)[1] as a measure of the chaos level in Baltimore. Moreover, the results of a Ravens football game and weather conditions during the day day are considered in this project and are used to predict the total number of 911 calls in Baltimore.

Our findings show that despite the festivities and the devotion people have to their home teams, there is no strong evidence to suggest that Baltimore's level of chaos is correlated with the results of a Ravens football game.

Methods
Data Sources and Description

The BPD provided a record of all 911 calls in the past two years. The 911 call dataset[1] mainly contains the calling location (including latitude and longitude), description of the call, time of a 911 call and other useful information. However, some descriptions are hard to recognize because they contain enigmatic codes with no reference table to identify them with. Therefore, it is difficult to analyze each call, so we

decided to only use the aggregate of our calls for our analysis. One category in the dataset contains the priority level of the call. The column is a categorical variable that contains different emergency levels from non-emergency to high-emergency.

While 911 calls are not a perfect measure of the amount of crime in a city, it could be a good measure of the general distress in a city. The unit of analysis is each football game day involving the Baltimore Ravens. Therefore, we also collected data related to the game day such as weather conditions of the game day and game schedule data.

Data on Ravens football games from 2015 to 2017 was collected from various NFL websites. Because the source websites varied, most of the data was collected manually or through web scraping. The collected data contained the actual date of the game, start and end time of a football game, the opposing team with the game name and game location, if the game was a home game, attendance, result of the game, and score gap of that game. According to our datasets[2][3], the Baltimore Ravens have played over 40 games in the NFL since 2015, 20 of which were in their home field M&T Stadium in Baltimore.

Another external dataset is weather data gathered from Weather Underground website[7][8][9]. The weather station is located at Maryland Science Center near the center of Baltimore City. The weather station is also close to the M&T Bank Stadium, which holds home playing football game for Baltimore Raven. The weather data contains useful information including geolocation using latitude and longitude, ID of the weather station, daily average temperature, daily precipitation and other useful information. This analysis mainly uses daily average temperature and precipitation.

Brief samples of our datasets are shown below.

911 Calls

callDateTime	priority	district	description	callNumber	incidentLocation	location
8/27/2015 18:57	Medium	SW	911/NO VOICE	P152192676	3400 BENSON AV	BALTIMORE, MD (39.26726, -76.671516)
8/27/2015 13:51	Medium	CW	Lab Request	P152191557	100 N MONTFORD AV	100 N MONTFORD AV
8/27/2015 14:31	Medium	SW	NARCOTICS ONVIEW	P152191693	1900 BLK HARMAN AV	1900 BLK HARMAN AV
9/1/2015 3:21	Non-Emergency	ED	Repo	P152440313	400 N WOLFE ST	400 N WOLFE ST
8/27/2015 10:44	Non-Emergency	SW	INVESTIGATE AUTO	P152190995	2000 HARMAN AV	2000 HARMAN AV
8/27/2015 13:51	Medium	WD	COMMON ASSAULT	P152191975	2500 PENNSYLVANIA AV	2500 PENNSYLVANIA AV

Football Games

| year	day	date	InMD	location	attendanc	StartTime	StartTime	team	match tea	win	score	gap
2015	Sun	13-Sep	N	Denver, C	76798	16	25	Baltimore	Denver Br	L	19-13	-6
2015	Sun	20-Sep	N	Oakland, (	53500	16	5	Baltimore	Oakland R	L	37-33	-4
2015	Sun	27-Sep	Y	Baltimore	70970	13	2	Baltimore	Cincinnati	L	28-24	-4
2015	Thu	1-Oct	N	Pittsburgh	63929	20	26	Baltimore	Pittsburgh	W	23-20	3
2015	Sun	11-Oct	Y	Baltimore	71046	13	2	Baltimore	Cleveland	L	33-30	-3
2015	Sun	18-Oct	N	Santa Clar	70799	16	25	Baltimore	San Franci	L	25-20	-5

Weather

STATION	NAME	LATITUDE	LONGITUL	ELEVATIO	DATE	PRCP	TAVG
USW00093	MARYLAN	39.2814	-76.6111	6.1	1/1/2015	0	2
USW00093	MARYLAN	39.2814	-76.6111	6.1	1/2/2015	0	4.2
USW00093	MARYLAN	39.2814	-76.6111	6.1	1/3/2015	6.496	3.05
USW00093	MARYLAN	39.2814	-76.6111	6.1	1/4/2015	4.016	9.75
USW00093	MARYLAN	39.2814	-76.6111	6.1	1/5/2015	0	3.1

Analysis Methods:

We used the results of these games to fit a linear model, along with other predictors, to find that there is not much evidence to suggest a correlation between the score gap of a game, and the level of distress in Baltimore measured in 911 calls. We controlled for weather[7][8][9], when the game started, if the game was on a weekend or not, and if the game was at home or not.

Results

Initial Simple Linear Model:
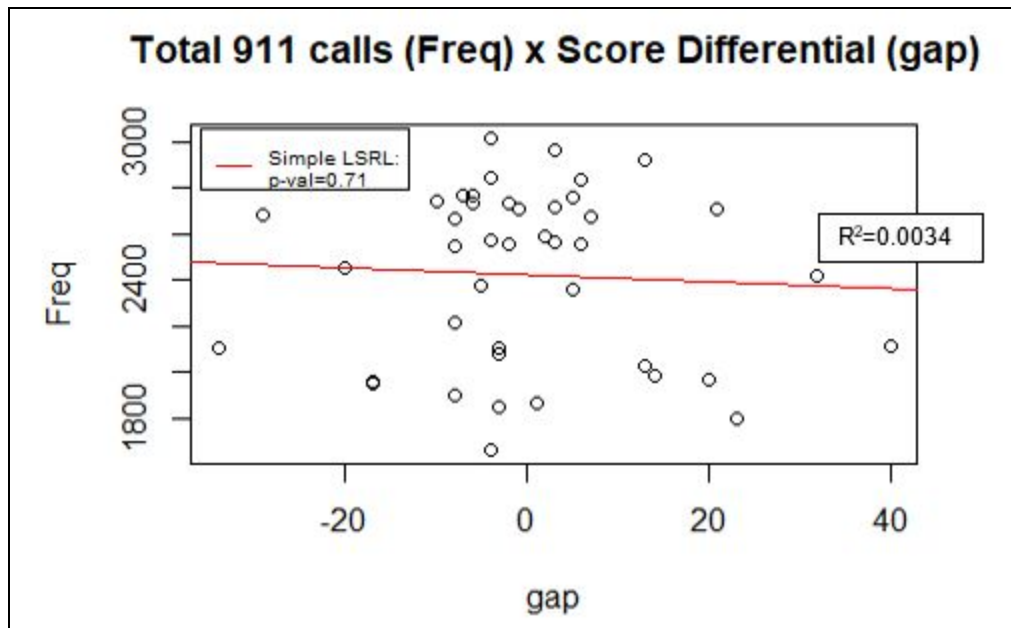
Our analysis is focused on the game results. To measure the results, we used a variable "gap" to indicate the score differential between the Ravens and the opposing team (gap = raven_score - opposing_score). Hence a large negative value for "gap" means a large loss for the Ravens, and a large positive value means a large win for the Ravens. "Freq" is the variable indicating how many 911 calls were received by the BPD on a game day.

A simple linear regression was performed to find any possible correlation between. As can be seen in the above plot with the least squares regression line, there is very little correlation, and there is little evidence to suggest that the score differential actually has any effect on total number of 911 calls.

However, the general trend of a larger loss depicted by a very negative value for "gap" does correspond with a higher 911 call count in "Freq". As the score differential becomes more and more in favor of the Ravens, the number of 911 calls seems to go down.

Multiple Linear Model:

      After running the linear regression through R, there were no significant p-values for any of the variables of interest. To speak even further, the output from R shows that there is almost no predictive power in our model.

      The below table provides a list of the variables included in the multiple linear model and their descriptions. Immediately following is the linear regression performed by R.

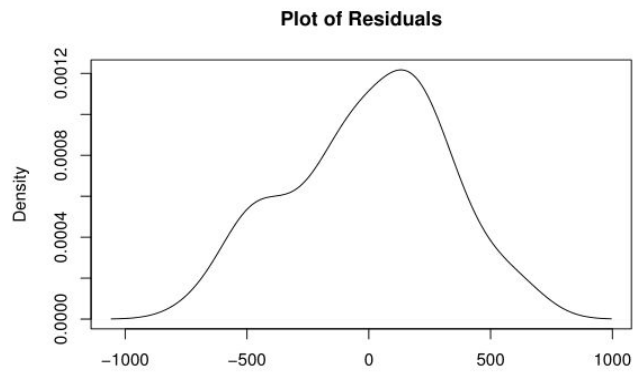| Variable	Description
"gap"	Score differential between Ravens and opposing team.
"StartTimeH"	Starting hour of the game in the day.
"home"	Dummy variable indicating if the game was a home game or not.
"weekend"	Dummy variable indicating if the game was on a weekend or not.
"precip"	Dummy variable indicating if the game was on a day that it rained or not.
"TAVG"	Average temperature of the day.
"Freq"	Total number of 911 calls in a day.

Table 1

```
## lm(formula = Freq ~ gap + StartTimeH + home + weekend + precip +
##     TAVG)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -602.57 -249.91   84.36  249.85  643.06
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3504.410    877.869   3.992  0.00032 ***
## gap           -4.324      4.508  -0.959  0.34404
## StartTimeH   -47.759     44.214  -1.080  0.28745
## home          39.649    128.714   0.308  0.75988
## weekend     -715.781    347.265  -2.061  0.04677 *
## precip      -268.431    152.193  -1.764  0.08650 .
## TAVG           4.591      5.326   0.862  0.39453
## ---
## Signif. codes:  0 '****' 0.001 '***' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 362.6 on 35 degrees of freedom
## Multiple R-squared:  0.2091, Adjusted R-squared:  0.07351
## F-statistic: 1.542 on 6 and 35 DF,  p-value: 0.1935
```

The above linear model gives a p-value of 0.34 for the coefficient for the "gap" variable. This is not significant enough to provide any evidence against the null hypothesis that the results of a Ravens football game has no effect on the number of 911 calls received by the BPD on that day.

Furthermore, none of the variables are particularly predictive except for maybe the weekend dummy variable with a p-value just under the $\alpha = 0.05$ threshold. However In fact, the p-value of 0.19 in the lower right of the above linear regression shows little evidence to suggest that any of the variables are predictive of the number of 911 calls in a day. The F-statistic of 1.542 further supports that the model has no predictive power, and it indicates that the low p-value of the linear regression is not due to collinearity.
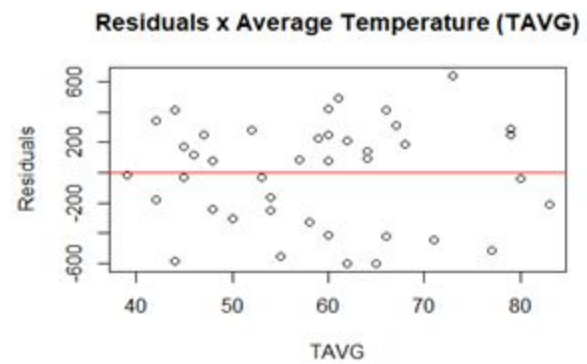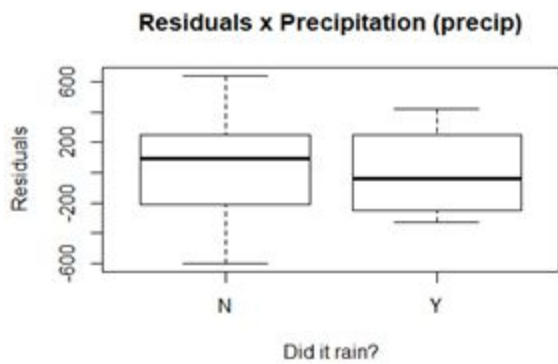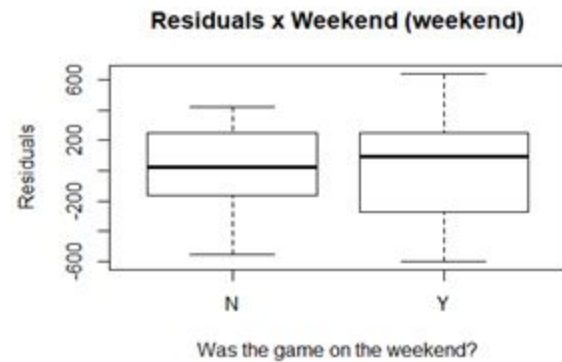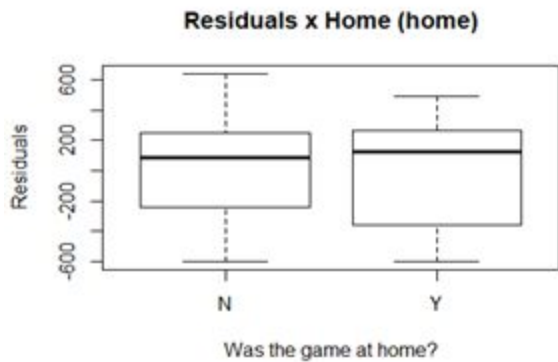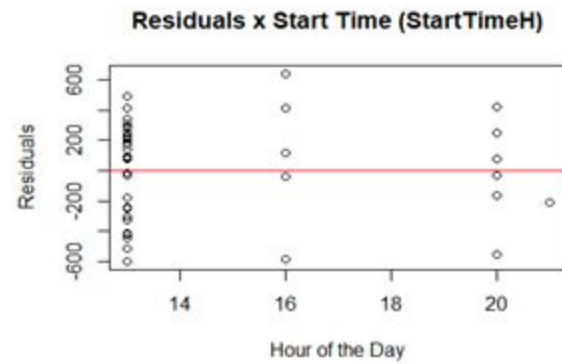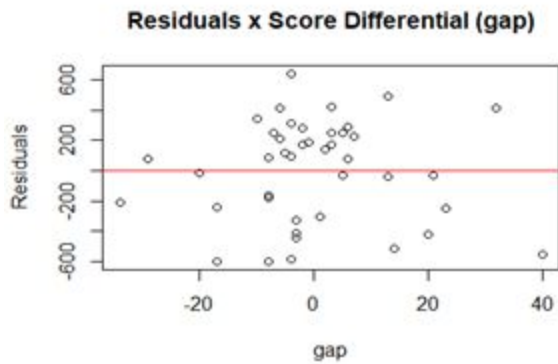
Residual Considerations:

The residuals from the linear model does not pose any particular concerns. All of the residuals seem to follow relatively symmetric distributions centered on zero that are not dependent on any of the predictor variables. Plots of the residuals and their relation to each variable are displayed on the next page.

Plot of Residuals

Mean calls per day:
2425.31

Sample Standard Deviation
of calls per day:
376.7367

Residuals x Score Differential (gap)

gap

Residuals x Start Time (StartTimeH)

Hour of the Day

Residuals x Home (home)

Was the game at home?

Residuals x Weekend (weekend)

Was the game on the weekend?

Residuals x Precipitation (precip)

Did it rain?

Residuals x Average Temperature (TAVG)

TAVG

Conclusion

Interpretation:

It seems that according to the result of the linear regression (Table 1), there is no significant correlation between total number of 911 calls and all of the chosen variables in Baltimore. The result may contradict the traditional convention that emotional football fans may cause chaos after a game depending on its results. When using number of 911 calls as an estimate of chaos level during the game day, the frequency of 911 calls are not significantly related to either the result of the game or any of our other predictors on that day in Baltimore. Therefore we could not reject the null hypothesis that the frequency of 911 calls on a gameday in Baltimore is not related to the score gap of the Baltimore Ravens after controlling for other variables.

Weaknesses:

Our largest weakness is our relatively small sample size of Baltimore Ravens game-days. The police call dataset contains 911 calls from 2015 to 2017, however, football games do not happen frequently, only in certain months during football season. This leads to another weakness. Our analysis does not factor in the time of year and how it can affect 911 calls received by the BPD.

Another issue is that using frequency of 911 calls per day may not be a proper measurement of the chaos level of Baltimore. In this case, the frequency of 911 calls in Baltimore per day is used to estimate chaos level in the city on that day. Conventionally, people call 911 when they are facing some situation that they could not deal with. This implies that when an emergency situation including a gunshot or a serious car accident happens, more than one witness would call 911. Similarly, for other situations including some homicide cases, 911 calls may be delayed when situations are happening or no 911 calls will be received by Baltimore police department. Thus, the definition of "chaos level" here is limited and using the frequency of 911 calls as an estimate of chaos may not be sufficient enough.

Furthermore, we do not value certain kinds of 911 calls any more than others. For example, a 911 call regarding a major traffic accident is counted the same as a call regarding a petty theft or a "no voice" call. As a result, if more serious 911 calls, such as homicide cases, are made more often instead of less serious calls, such as about petty theft or "no voice" calls, then the total count of 911 calls stays about the same. Hence our analysis fails to detect the increase in chaos in Baltimore.

Future Directions:

The weaknesses of our model can be overcome. The important weakness of low sample size can be remedied by waiting until more data becomes available as time passes.

Our analysis also fails to include the fact that the month or day likely has a strong effect on 911 calls in a day, so implementing some sort of time series analysis would likely give us greater power in detecting differences in our variables of interest.

Finally we could also utilize the "priority" variable provided by the BPD in the 911 call dataset. We were unable to get a hold of a representative from the BPD to give us proper descriptions of each variable, so we were unsure of how to use some of the variables. If we were given good documentation regarding the variables, then we could possibly create a proxy for chaos level in Baltimore that is stronger than simply counts for 911 calls in a day.

Other ways to improve the model include the use of cross-validation to see if a polynomial term of some variables could provide more predictive power without overfitting. Another strategy is to compare the frequency of gameday with the following monday for further analysis to see if gameday itself is related to frequency of 911 calls.

Due to the limitations in our weather data, we only used one station's data to represent the weather for all of Baltimore. Further investigation can allow us to use all of the available weather stations along with kriging to generate more precise measurements of weather in different areas of Baltimore.

Other areas of interest for analysis would be to make comparisons between different areas of Baltimore instead of just doing an analysis of Baltimore as a whole. Furthermore, we could also focus more on certain times relative to a Ravens game such as immediately before, during half-time, and after. We are not sure how to implement these into our analysis, but they could definitely be interesting if we did.

References:

[1] Baltimore Police Department. (2017, December 14). 911 Calls for Service | Open Baltimore | City of Baltimore's Open Data Catalog. Retrieved December 14, 2017, from https://data.baltimorecity.gov/Public-Safety/911-Calls-for-Service/xviu-ezkt

[2] Football game data set for Baltimore Ravens:

2015 Baltimore Ravens season. (2017, November 10). Retrieved December 14, 2017, from https://en.wikipedia.org/wiki/2015_Baltimore_Ravens_season

2015 game attendance: Results & Stats. (n.d.). Retrieved December 14, 2017, from http://www.baltimoreravens.com/team/history/all-time-results-and-statistics.html

2015 game schedule & results: 2015 Baltimore Ravens Statistics & Players. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/teams/rav/2015.htm

2016 game attendance: 2016 NFL Attendance Data. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/years/2016/attendance.htm

2016 game schedule & results: 2016 Baltimore Ravens Statistics & Players. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/teams/rav/2016.htm

2016 game schedule details: 2016 Baltimore Ravens Schedule & Game Results. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/teams/rav/2016_games.htm

2017 game attendance: 2017 NFL Attendance Data. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/years/2017/attendance.htm

2017 game schedule & results: 2017 Baltimore Ravens Statistics & Players. (n.d.). Retrieved December 14, 2017, from https://www.pro-football-reference.com/teams/rav/2017.htm

[3] Football game data set for Maryland Terrapins

2015 attendance: 2015 Maryland Terrapins football team. (2017, November 25). Retrieved December 14, 2017, from https://en.wikipedia.org/wiki/2015_Maryland_Terrapins_football_team

2015 game schedule & results: 2015 Maryland Terrapins Football Schedule. (n.d.).

Retrieved December 14, 2017, from
http://www.fbschedules.com/ncaa-15/big-ten/2015-maryland-terrapins-football-schedule.php

2016 attendance: 2016 Maryland Terrapins football team. (2017, November 25).
Retrieved December 14, 2017, from
https://en.wikipedia.org/wiki/2016_Maryland_Terrapins_football_team

2016 game schedule & results: 2016 Maryland Terrapins Football Schedule. (n.d.).
Retrieved December 14, 2017, from
http://www.fbschedules.com/ncaa-16/2016-maryland-terrapins-football-schedule.php

2017 attendance: 2017 Maryland Terrapins football team. (2017, December 01).
Retrieved December 14, 2017, from
https://en.wikipedia.org/wiki/2017_Maryland_Terrapins_football_team

2017 game schedule & results: 2017 Maryland Terrapins Football Schedule. (n.d.).
Retrieved December 14, 2017, from
http://www.fbschedules.com/ncaa-17/2017-maryland-terrapins-football-schedule.php

[4] McGranaghan, J., (2015) Does watching sports affect your mood? Retrieved from:

Sites.psu.edu

[5] Open Baltimore. (2017, February 5). Baltimore City Line | Open Baltimore | City of
Baltimore's Open Data Catalog. Retrieved December 14, 2017, from
https://data.baltimorecity.gov/Geographic/Baltimore-City-Line/fy7v-tvcr/data

[6] Park, A. (2011). Fan rage: How home team losses contribute to domestic violence.

Retrieved from: healthland.time.com

[7] Weather data of 2015:WeatherUnderground. (2017, November 27). Weather History
for Maryland Science Center, MD - January, 2015 - December 2015. Retrieved
November 27, 2017, from
https://www.wunderground.com/history/airport/KDMH/2015/1/1/CustomHistory.html?dayend=31&monthend=12&yearend=2015&req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo=

[8] Weather data of 2016: WeatherUnderground. (2017, November 27). Weather History

for Maryland Science Center, MD - January, 2016 - December 2016. Retrieved November 27, 2017, from https://www.wunderground.com/history/airport/KDMH/2016/1/1/CustomHistory.html?dayend=31&monthend=12&yearend=2016&req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo=

[9] Weather data of 2017: WeatherUnderground. (2017, November 27). Weather History for Maryland Science Center, MD - January, 2017 - November 2017. Retrieved November 27, 2017, from https://www.wunderground.com/history/airport/KDMH/2017/1/1/CustomHistory.html?dayend=31&monthend=12&yearend=2017&req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo=