From Gentrification to “Within-Trification”

Big Data and Data Analytics at Bethel Park High School

Bethel Park High School Team Two
Sarah Steeb, Emma Ratti, Zak Gorman, Cassidy DeLeo, Mike Musciano

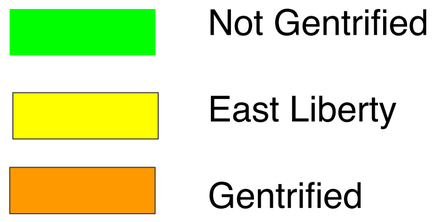
PROBLEM & IMPORTANCE:

What are the impacts (economic, social, demographic) of gentrification in East Liberty and how can we compare them to surrounding communities? According to the United Nations, Community Development is a process designed to create conditions of economic and social progress for the **whole** community within its active participation and fullest possible reliance upon the community’s initiative.

THE DATA

- There is an excerpt of the data we used in our project. We obtained most of our data from PGH Snap.
- We also did extensive research on news articles from the Pittsburgh area & other cities that went through gentrification, such as Brooklyn, NY.
- We picked the neighborhoods to compare to East Liberty that are in Pittsburgh. South Side, Shadyside, and Squirrel Hill North are known as trendy,up-and-coming areas. Knoxville, Carrick, and Homewood North are in the works of developing their communities to thrive.

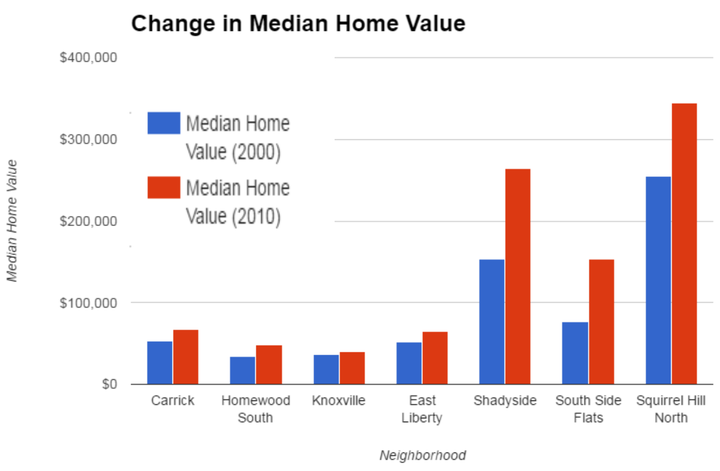
Neighborhood	Pop. 1960	Pop. 1970	Pop. 1980	Pop. 1990	Pop. 2000	Pop. 2010	% Pop. Change, 60-70	% Pop. Change, 70-80	% Pop. Change, 80-90	% Pop. Change, 90-00	% Pop. Change, 00-10
Carrick	16,480	15,855	12,930	11,625	10,685	10,113	-3.79%	-18.45%	-10.09%	-8.09%	-5.35%
Homewood South	22,463	8,876	6,228	4,811	3,647	2,344	-60.49%	-29.83%	-22.75%	-24.19%	-35.73%
Knoxville	7,353	6,527	5,524	4,971	4,432	3,747	-11.23%	-15.37%	-10.01%	-10.84%	-15.46%
East Liberty	12,005	8,647	8,741	7,973	6,871	5,869	-27.97%	1.09%	-8.79%	-13.82%	-14.58%
Shadyside	18,177	15,848	13,945	13,385	13,754	13,915	-12.81%	-12.01%	-4.02%	2.76%	1.17%
South Side Flats	12,586	9,260	7,894	6,177	5,726	6,597	-26.43%	-14.75%	-21.75%	-7.30%	15.21%
Squirrel Hill North	13,778	13,576	12,353	11,471	10,408	11,363	-1.47%	-9.01%	-7.14%	-9.27%	9.18%

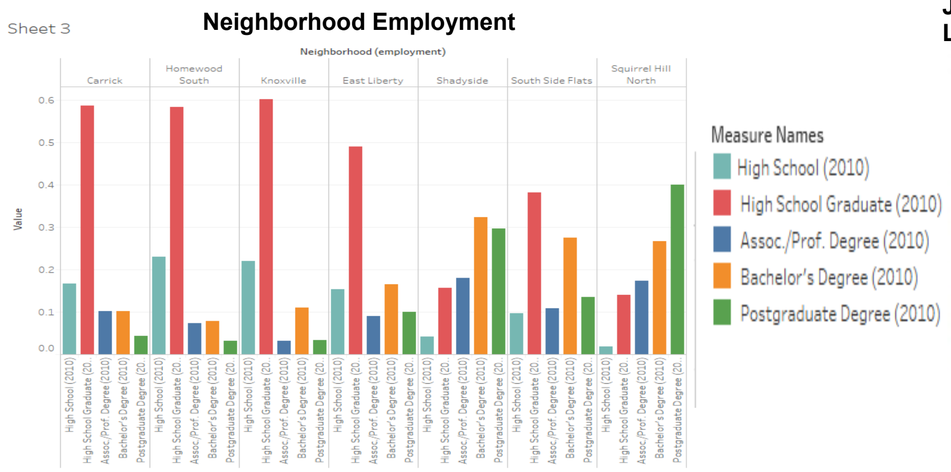


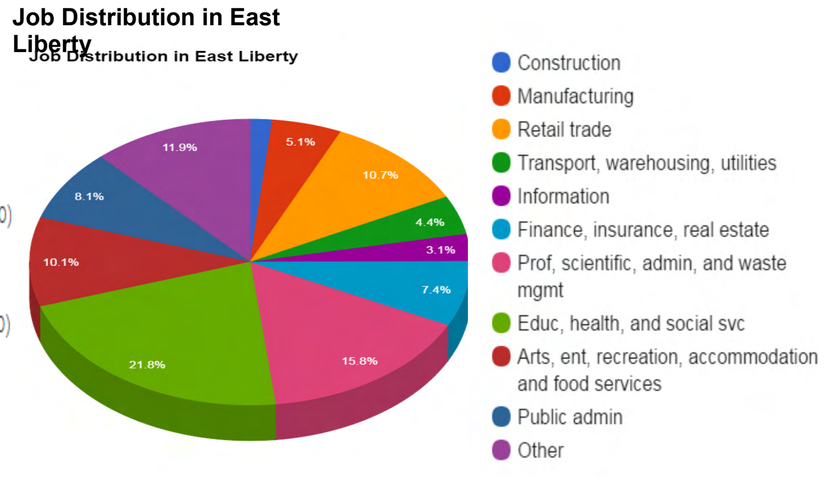
CHALLENGES

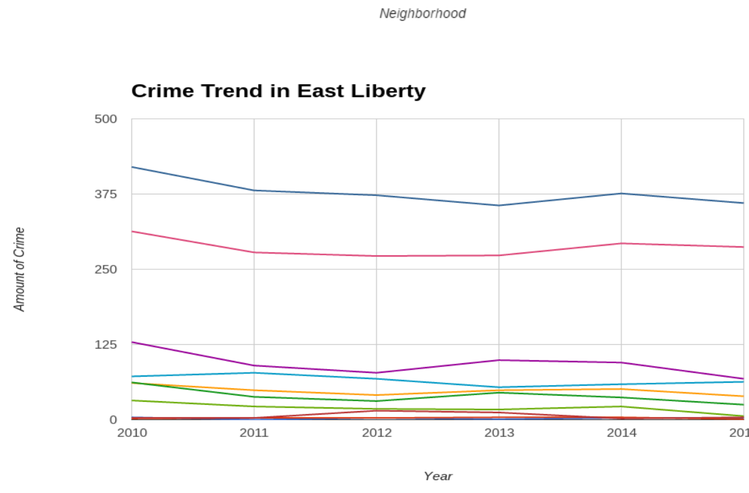
- One challenge we faced was finding a way to portray gentrification in a meaningful way
- We had difficulties deciding which type of graph to best display our data.
- There were other variables we had to account for, especially since we do not reside in East Liberty.

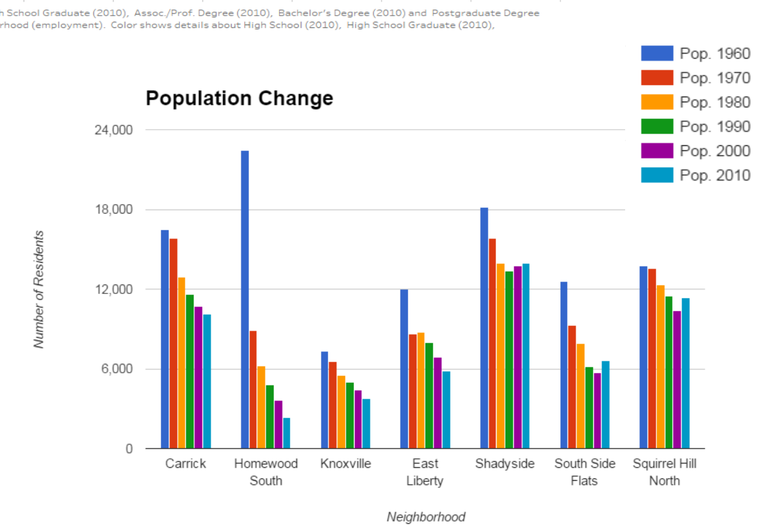
VISUALIZATIONS

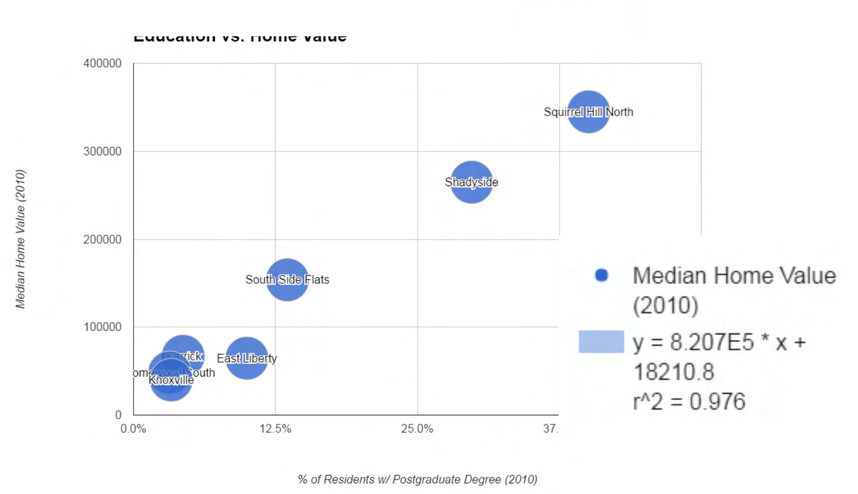












SUMMARY

- Our visualizations and graphs provide evidences that East Liberty qualifies to go through the renovation process known as gentrification. The trends show that East Liberty is between the gentrified and non-gentrified communities when compared to home value, crime, population, education level, and employment.
- Residents who have a postgraduate degree have a higher median home value, which falls in the category with the three gentrified areas: Shadyside, South Side Flats, and Squirrel Hill North.
- From the years 2010 to 2015, crime in East Liberty has decreased in almost every aspect, providing evidence that gentrification is improving the wellbeing of the community.
- The population for gentrified communities decreased steadily from 1960 to 1990, but increased between 1990 and 2000 (likely due to gentrification). The other communities have decreased drastically in population every year from 1960 to 2010.

- From 1999 to 2009, the median income has gone up for gentrified communities, but stayed the same or decreased in areas with no action. East Liberty remained the same during the ten year period.

POLICY PLAN

- We are just starting here with East Liberty as our main focus. If we had more time and data, we imagine furthering research to other Pittsburgh neighborhoods.
- Our policy is for neighborhoods facing gentrification to get in contact with their local churches, organizations, and government to bring the citizens in the neighborhood together as a community.
- When new businesses come in, they can minimize outsourcing and offer job positions first to the locals before reaching out to the rest of the city. Since education, health, and social services is what most people are working in, those types of businesses should go in East Liberty.
- Moving forward in East Liberty, affordable housing needs to become a priority.
- When going through the gentrification process, developers and community leaders need to cherish the wants and opinions of citizens already living there. Gentrification is a great thing, as long as everyone gets a fair share of the benefits.

- As shown in the video, gentrification should be shifted towards the process of “within-trification”.

Coal Conflicts

Bethel Park High School Team 3
Tarush Bahl, Rachel McClaine, Aubree Stewart, Hannah Tower, Maggie Wolf

The Question

How was the coal industry impacted by the economy and the environment, and what are the future implications for regulations on coal?

Challenges

We faced many challenges when finding data. For example, some of our data sets were from PA and others were from the whole country. It was also difficult to determine the impact of exclusively coal on the environment since there are many other factors affecting the environment (global warming).

Data Set Examples

YEAR	CO2 Residential Emissions million metric tons of CO2	CO2 Commercial Sector million metric tons of CO2	CO2 Industrial Sector million metric tons of CO2	CO2 electric pow	COal total emmission million metric tons of CO2	% Ash content COKE PLANTS in PA	%Ash Content Commercial Institutional in PA	% Ash Content Electricity power in PA	%Ash Content Independent Power producers in PA	average coal mine workers	averag
1980	0.7	2.7	52.4	96.3	136.5					253,007	
1981	0.7	3.3	45.7	89.5	152.0					249,738	
1982	0.7	3.2	25.8	91.0	139.2					241,454	
1983	0.6	2.8	29.5	92.0	120.7					200,199	
1984	0.7	2.9	39.2	93.8	125.0					208,160	
1985	0.6	2.2	33.2	95.9	136.5					197,049	
1986	0.8	2.5	26.8	93.4	131.9					185,167	
1987	0.9	2.9	30.6	95.1	123.5					172,780	
1988	0.8	2.6	34.7	99.4	129.4					166,278	
1989	0.8	2.8	34.2	102.1	137.5					164,929	
1990	0.6	2.5	35.9	99.3	139.9					168,625	
1991	0.6	2.7	31.8	99.1	138.3					158,677	
1992	0.7	3.1	34.3	100.2	134.2					153,128	
1993	0.5	2.4	35.8	101.0	138.3					141,183	

Coal mining jobs

State	Totla Coal Production	Underground mining jobs	Surface Mining Jobs	Total Jobs	Annual Production per miner
Alabama	18,620.00	3,077.00	1,135.00	4,212.00	4.40
Alaska	1,632.00	-	125.00	125.00	13.10
Arizona	7,603.00	-	405.00	405.00	18.80
Arkansas	59.00	52.00	2.00	54.00	1.10
Colorado	24,236.00	1,705.00	470.00	2,175.00	11.10
Illinois	52,147.00	3,660.00	504.00	4,164.00	12.50
Indiana	39,102.00	1,986.00	1,626.00	3,612.00	10.80
Kansas	22.00	-	7.00	7.00	3.10
Kentucky	80,380.00	8,938.00	3,967.00	12,905.00	6.20
Louisiana	2,810.00	-	280.00	280.00	10.00
Maryland	1,925.00	184.00	221.00	405.00	4.80
Mississippi	3,575.00	-	309.00	309.00	11.60
Missouri	414.00	-	24.00	24.00	17.30

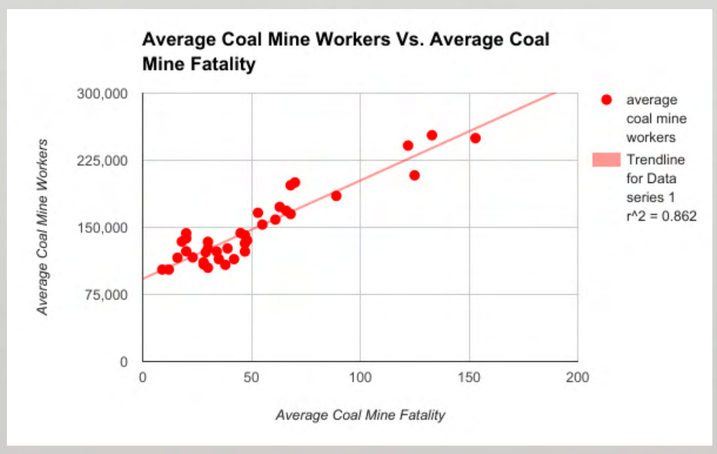
% renewable energy produced

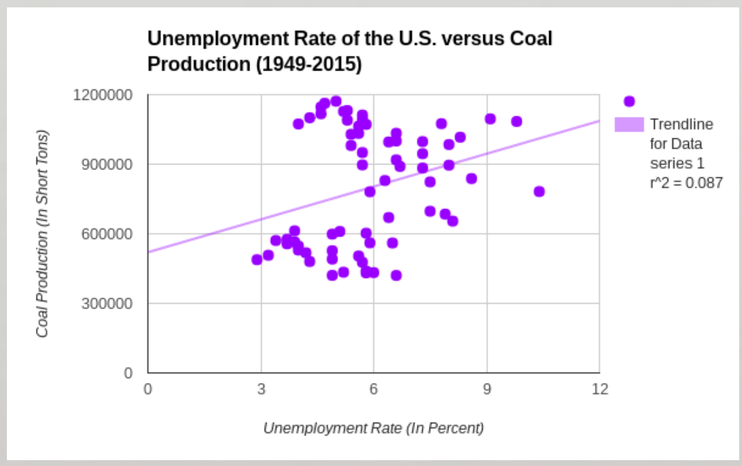
State	Renewable Energy	Wind	Solar	Hyd
Alabama	16.15			7
Alaska	0.77			>1%
Arizona	15.51		>1%	11
Arkansas	12.01			3
California	24.38	2%	1%	10
Colorado	3.11	1%		>1%
Connecticut	13.02		1%	2
Delaware	100		2%	
Florida	40.99		8%	>1%
Georgia	36.36			6
Hawaii	100	15%	18%	7
Idaho	100	2%		73
Illinois	11.54	1%		

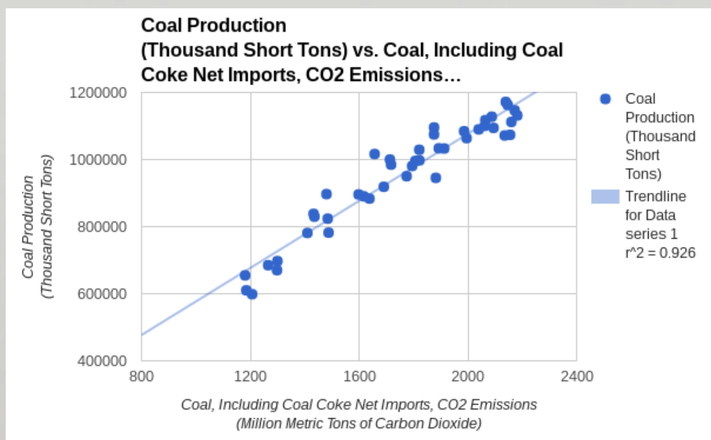
Data Sources

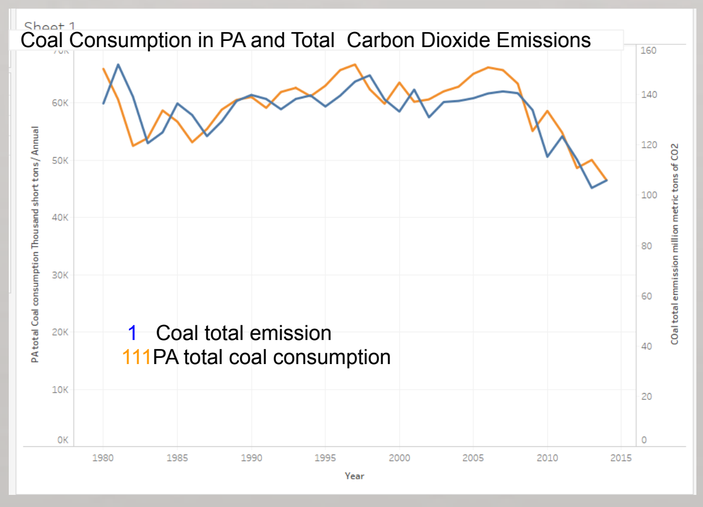
To find accurate data we looked at many different websites that we thought were trustworthy. We looked for the most recent and reliable data that we could find, even though it was difficult at times to find data specific to our topic. In particular, some data available was not coal specific and included other energy sources.

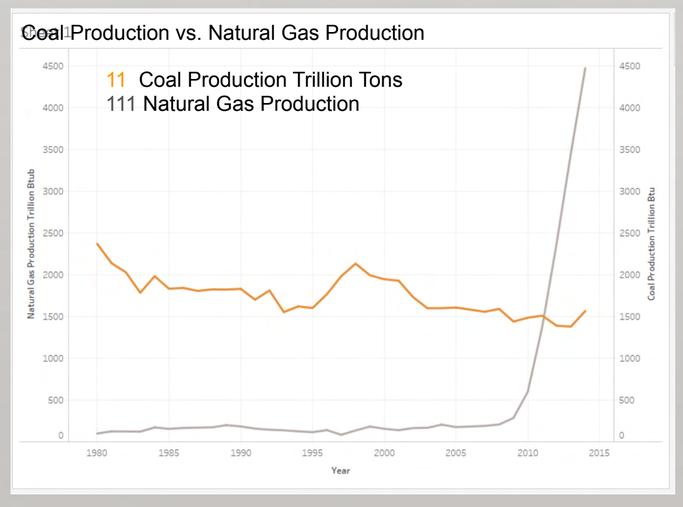
Analysis

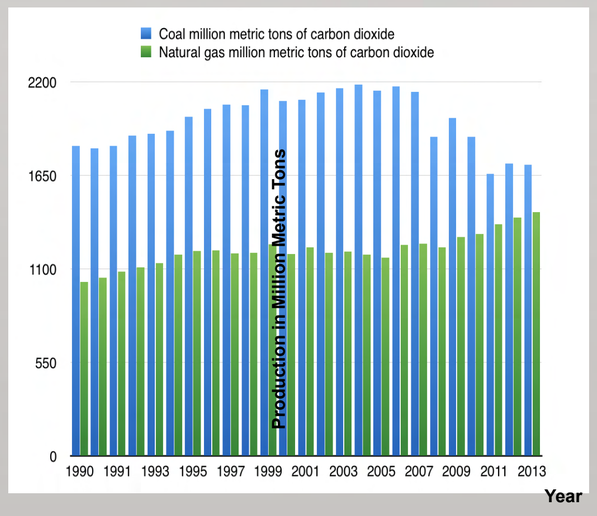












Conclusion

There is a strong positive correlation between total carbon dioxide emissions and coal production. Additionally, as more workers are employed in the coal industry, the number of fatalities increases. However, there is virtually no significant correlation between the number of coal miners employed and the overall unemployment rate of the United States. Natural gas production has increased significantly since 2010, but carbon dixide emissions have only gone up slightly, still significantly less than the emissions from coal.

Potential Policy

The EPA can use this data in order to set regulation to begin phasing out coal. Based on our data analysis, the regulations should increase. This is based on the positive correlation between coal production and carbon dioxide emissions and the increase in mine fatalities when there are more workers in the industry. Natural gas produces less carbon dioxide than coal and therefore has less harmful effects on the environment, even with increased production. Hopefully with increased sanctions, we will be able to improve and save our planet.

Does Crime Rate Affect Property Value?

Brashear High School
Michaela Koch,
Luis Gil-Chacon,
Kayli Short,
Taiwo Lawal,
Amanda Adams,
Isabella Romano

Introduction

When areas have higher crime rates people are less likely to purchase property in that area. When people are sent to jail, taxes increase to cover the inmates cost of living.

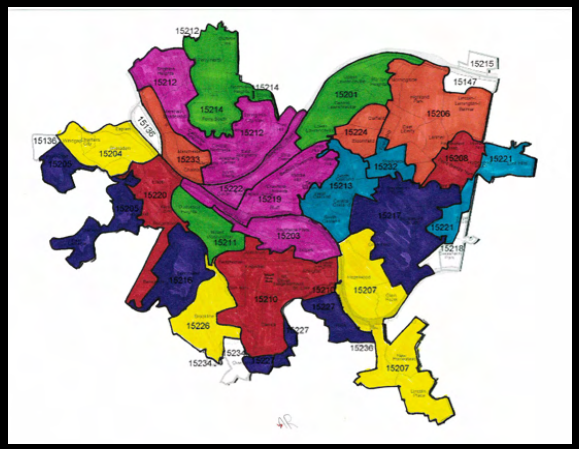
Conclusion

Based on our data, our hypothesis was false. We found in our data maps that in some areas like 15203 the property values are very high, while the arrest rate is also very High.

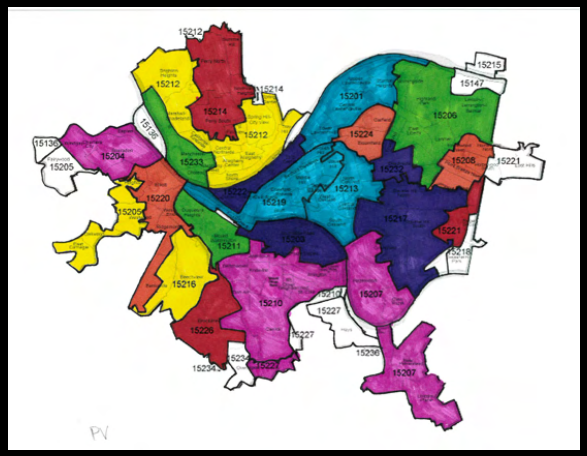
Zip Code	Count of ARREST	Population	Percent of Arrests
15222	236	3,294	7.16%
15203	305	9,940	3.07%
15212	811	27,895	2.91%
15219	362	16,696	2.17%
15220	168	17,718	2.17%
15208	154	10,406	1.86%
15210	380	25,954	1.47%
15233	65	4,451	1.46%
15224	117	10,141	1.15%
15206	326	28,615	1.14%
15207	122	11,268	1.08%
15226	142	13,214	1.07%
15204	131	14,352	0.91%
15214	131	14,352	0.91%
15201	92	12,713	0.72%
15211	66	11,081	0.60%
15221	148	31,602	0.48%
15232	52	11,374	0.46%
15213	139	30,844	0.45%
15216	86	23,359	0.37%
15217	88	27,220	0.32%
15205	67	21,865	0.31%
15227	39	28,156	0.14%

Average Property Value	Zip Code
31,067.28	15201
31,495.59	15210
42,168.28	15227
44,078.59	15207
46,108.63	15214
64,035.34	15236
73,321.20	15221
90,674.60	15220
92,309.77	15208
92,545.37	15224
108,155.00	15205
115,370.60	15216
133,334.70	15212
136,113.00	15206
144,919.80	15211
153,475.20	15233
158,873.07	15201
214,408.60	15213
239,105.30	15219
259,661.20	15217
266,073.40	15203
291,667.60	15232
813,055.10	15222

Arrest Rates



Property Value



Resources:
https://data.wprdc.org/dataset/real-estate-sales
https://data.wprdc.org/dataset/arrest-data
city.pittsburgh.pa.us

RYAN ESKENASY, EZEKIEL STUART
CENTRAL CATHOLIC HIGH SCHOOL

UNDERSTANDING THE CONNECTION BETWEEN FINANCE AND CRIMINAL ACTIVITY

RESEARCH QUESTION

MORE CONCENTRATED NEIGHBORHOODS SHOULD HAVE HIGHER PROPERTY AND VIOLENT CRIME RATES THAN LESS CONCENTRATED NEIGHBORHOODS. HOWEVER, FINANCIAL DEMOGRAPHICS COULD IMPACT THESE DIFFERENT CRIME RATES MORE THAN THE ASPECT OF POPULATION DENSITY. IS THERE A SIGNIFICANT CORRELATION BETWEEN THE TYPES OF CRIME IN AFFLUENT VERSUS DESTITUTE NEIGHBORHOODS IN ALLEGHENY, PA?

PROBLEMS WITH DATA COLLECTION/USAGE

1. COLLECTING DATA FOR CONSISTENT TIME PERIODS FOR CRIME, MEDIAN HOUSEHOLD INCOME, AND DENSITY

2. IDENTIFYING ERRORS AND ERRONEOUS CONTENT IN OUR DATA SETS

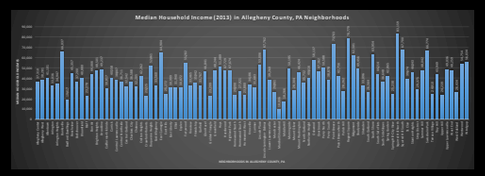
3. INITIALLY STRUGGLED WITH DETERMINING WHICH FINANCIAL STATISTIC WOULD BE MOST APPROPRIATE FOR WORKING WITH OUR TOPIC.

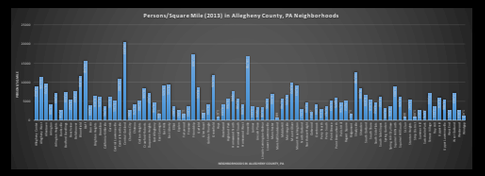
4. UNDERSTANDING AND ANALYZING UNEXPECTED RESULTS

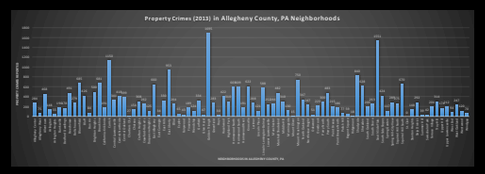
DATA REFERENCES

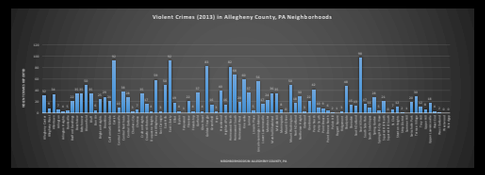
1. CITY-DATA
   - CITY-DATA.COM
2. UNITED STATES CENSUS BUREAU
   - FACTFINDER.CENSUS.GOV
3. WESTERN PENNSYLVANIA REGIONAL DATA CENTER
   - WPRDC.ORG
4. PGHSNAP
   - PITTSBURGHPA.GOV/DCP/SNAP/
5. TABLEAU PUBLIC
   - PUBLIC.TABLEAU.COM

VISUALIZATION OF COLLECTED DATA







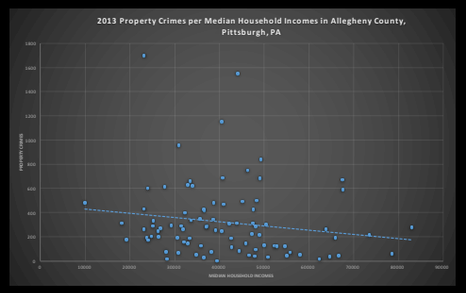


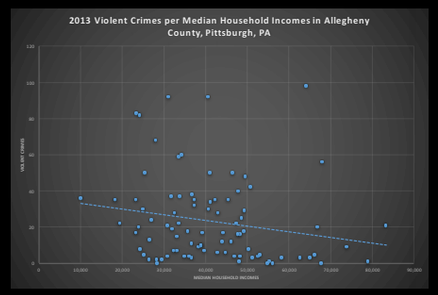
RESULTS

After organizing, graphing, and analyzing our data, our conclusions have not supported our initial claim, that there will be a correlation between the property and violent crimes, in affluent versus destitute neighborhoods.

- In neighborhoods that had higher median household incomes, we predicted that they would have higher property crimes. This was invalid, as there was an exiguous amount of correlation. Our R^2 value was .0264, which means this model suggests almost no correlation, as a 1.0 R^2 value is a complete positive linear correlation

- Throughout neighborhoods that had lower median household incomes, our group predicted that the violent crimes would increase. Also proving invalid by our collected data, our analysis representing violent crime rates per median household income had a R^2 value of .0375. This suggests that there is almost no correlation between violent crimes and neighborhoods that had lower median household incomes.





CONCLUSION

- After analyzing all of our data, our team's expectation of results was invalid. Although this was the case, we were still able to come up with conclusions as to what this data represents, and how the results can still portray a direct impact on society. One major conclusion that we reached, is that the lack of correlation can help prevent people searching for a new home to make rash, ill-advised decisions. Often times, buyers tend to look at houses in areas that are financially more adept, because they assume that crime rates will be lower. This would be incorrect, as for example the plethora of neighborhoods that contradicted our theories.

- This can be extremely significant, because it would inform potential home-buyers to thoroughly look at all aspects of their home before making a final decision of purchase. They must not look at just the financial demographics, but rather the totality that could cover employment rates, security levels, etc.

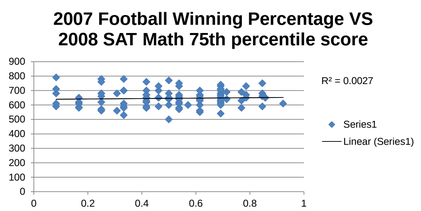
The Relationship Between the Performance of a University's Football and Basketball Programs and its Admissions Selectivity

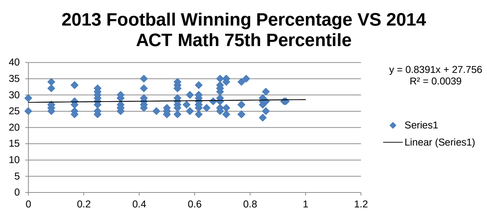
Introduction

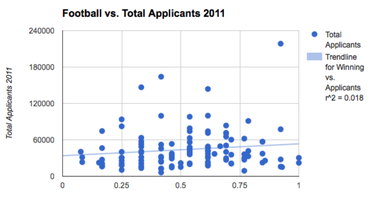
College and Sports are intrinsically tied; The first university level organized sports club, a boat club, was founded in 1843. Since then, the average number of sports programs at universities has increased drastically, and the purpose for them has changed too. Now, sports at the university level are a platform for the athlete and a revenue stream for the schools. Schools with large, well-established athletic programs can justify the continuation of their basketball and football programs (the pillar NCAA athletic programs for advertising revenue) with the revenue they generate. While the revenue a college can generate from a successful sports program is enticing, some colleges still decide that the money isn't enough to warrant creating or investing more into basketball and football programs. What if there was a quantitative relationship between the performance of a school's basketball and football and the quality of non-transfer first-time degree-seeking undergraduates applicants to the school for the following school year? Such a connection could warrant that small colleges or universities with primarily academic reputations, begin or invest more into their basketball and football programs. The goal of our project was to investigate our null hypothesis: There is a statistically insignificant relationship between the performance of a college's basketball and football programs and the quality of first-time degree-seeking undergraduate applicants for the following school year. To test our null hypothesis we collected football, basketball, and applicant data for about 2000 schools; Organized the data, examined and compared it identifying little correlation

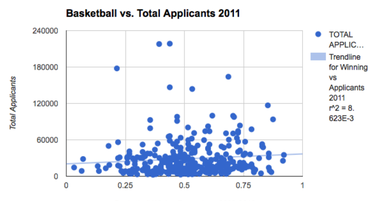
Challenges

- One of our biggest challenges was collecting the data. The education data came in a friendly spreadsheet, but required us to manually sift through almost 2000 institutions to find the ones that pertained to our comparison. The sports data did not come in a spreadsheet, which meant that we had to collect 20 years of winning percentages for both sports for over 300 schools.
- The collection of the data took up a lot of time; It caused us to add more meeting dates to our schedule, so that we could analyze the data to the level we desired.
- Another problem was the volume of the data we had. Our final spreadsheet had over 550 columns of information for our institutions spread over many years.

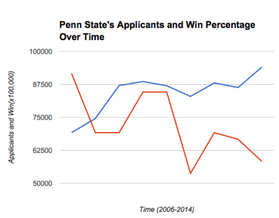


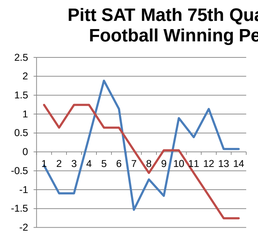


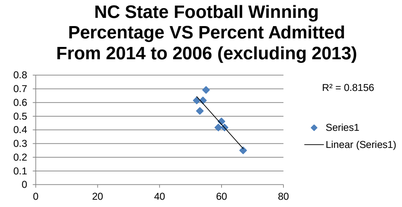




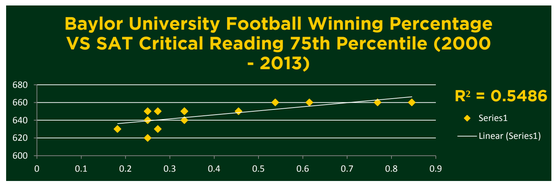
The general comparisons among our admission variables and winning percentages provided almost no evidence that the two were related in any way.

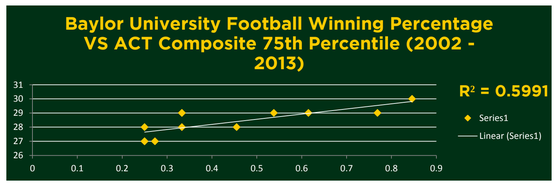


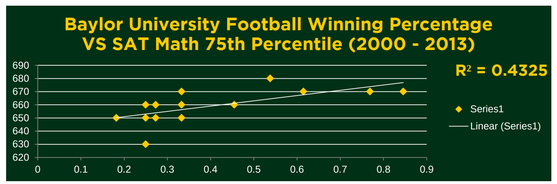


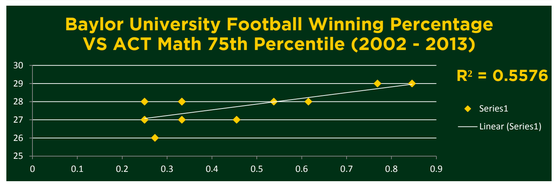


When we began to look at individual schools, patterns would sometimes occur that intrigued us and led us to investigate for a pattern.









Four correlations shown of Baylor University's Football Winning Percentage over time compared to different standardized test statistics. These correlations provide evidence that schools should be looking at this data themselves to see if the success of their sports' teams benefits their institution.

Method

- We first began by trying a establish a general pattern with our data, looking specifically at one year's winning percentage for a sport and then looking at some indication of college performance for the subsequent year.
- Some of the indicators of performance we looked at were: SAT 25th and 75th Quartiles for Math and Critical Reading, Admissions Yield, Graduation Rate, and Enrollment
- After our general analysis yielded results that supported our null hypothesis, we decided to shrink our scope and began looking at universities on divisional basis then a case by case basis.

Conclusions

- Test for correlation repeatedly reported insignificant correlation, supporting our null hypothesis.
- We did find some data that suggests that for certain schools, there is a pattern that should be explored further. Therefore, we advise that colleges run statistical analysis on their institutions to see if it would benefit them to begin or to invest more into basketball and football programs. Because of the results of our testing at a national level we must accept our null hypothesis but further examination could and did yield different results on a smaller school by school basis.

Resources

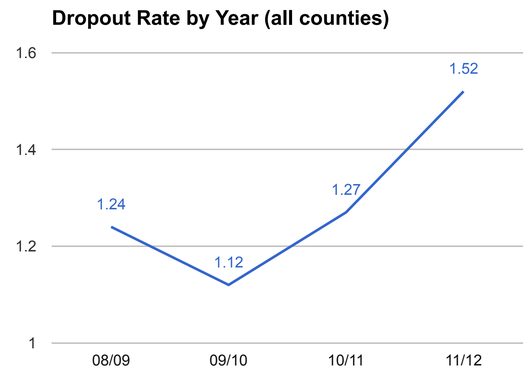
- Our data set for all of our education statistics was Integrated Postsecondary Education Data System (IPEDS), from the National Center for Education Statistics
- Our data set for all the sports' winning percentages came from SportsReference.com for both football and basketball.

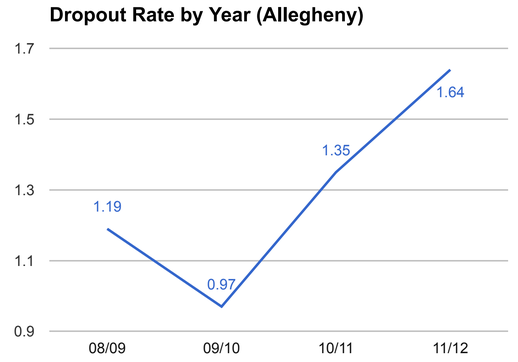
Central Catholic Team 1
John Lynch
Jackson Kaib
Adam Wipprecht
Tristian Yanalitis

Keystone Crumbling

CORNELL SCHOOL DISTRICT

LIAM WHITE MCSHANE, CIARA TRIMMER, JAYLIN CICCONE, DERRIC DENNISTON, KAYLA PICKENS



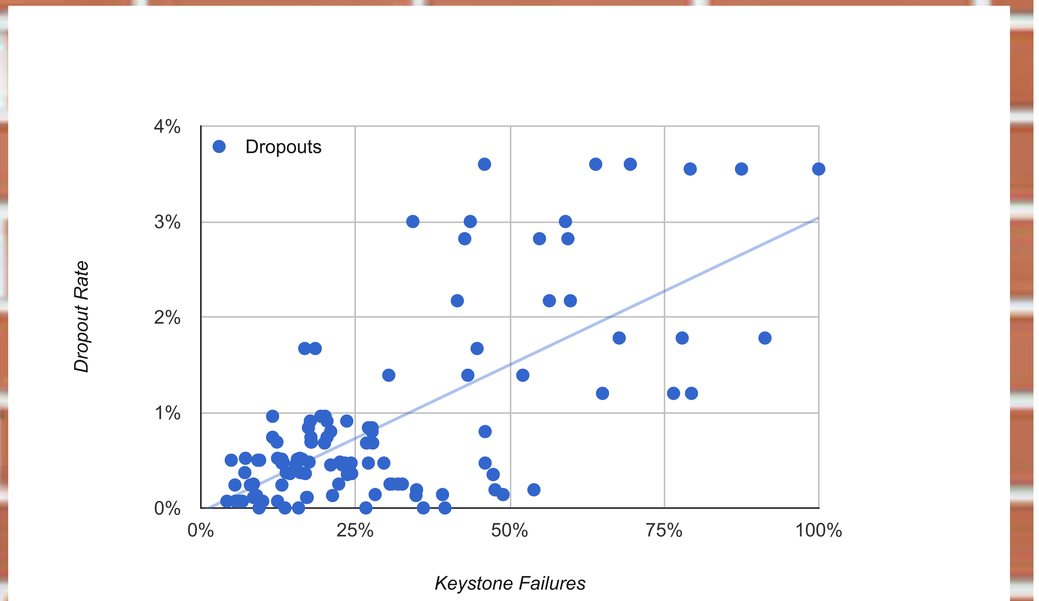


DID THE INTRODUCTION OF THE KEYSTONE EXAMS AFFECT DROPOUTS?

- We collected data from pa.gov on the school dropout rates for Allegheny County districts and state dropout rates. Our school administrator provided us with Keystone scores for Allegheny County.

- We noticed right away that there was a drastic increase in dropout rates the year that the Keystone Exams were implemented across the state.

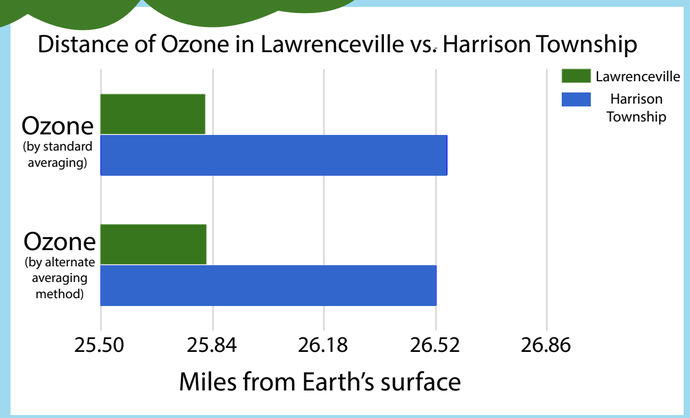
- After analyzing our data from Allegheny County and across the state, we conclude that Keystone Exams are creating a negative atmosphere in the schools and should be discontinued.

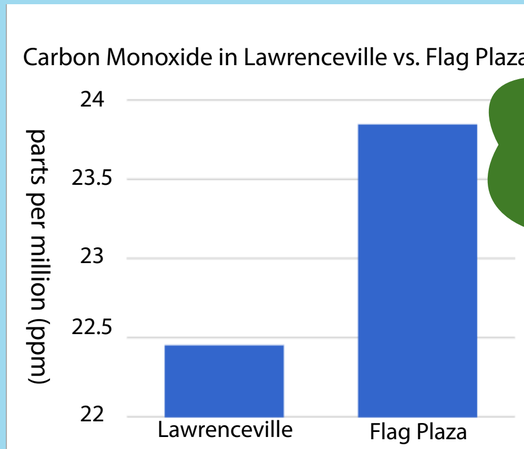


The Effects of Green Spaces on Air Quality in Allegheny County

Question

Do green spaces or parks affect air quality in Allegheny County?





Introduction

Pittsburgh is ranked to have the eighth worst air quality in the U.S. due to the city's history of steel mills and increased car emissions over the years. So, does having outdoor parks or green spaces affect the air quality of Pittsburgh? Our goal was to analyze the air quality data from Allegheny County and compare it to the air quality data from areas with little or no green spaces. We predicted that the areas with more green spaces would have lower levels of pollutants in the air. than the areas with less green spaces. Before analyzing the data, we assumed that the data sets would correlate and the air quality would be affected by green spaces.

Process

In order to correctly evaluate the data, we conducted research to define the chemicals that were tested in the air and what sources emit them into our air. This information gave us a better idea of how the green spaces might play a role in the elimination of these pollutants. The process of analyzing the air polution data was our biggest challenge since the terms used to measure the data were new to us as were the concepts that went along with air quality. Our team gained an immense amount of knowledge about the environment through analyzing the data.

Conclusion and Analysis

We found that green spaces do affect air quality and that air quality is better in areas with green spaces compared to areas without. We compared the distance of the ozone in two areas, one with green spaces and one without, as well as amounts of carbon monoxide. The ozone in the area with less green spaces is closer than the area with more green spaces, degrading the quality of air in that area. When we looked at the amounts of carbon monoxide, which can be a deadly gas, we compared levels recorded in an area populated with multiple parks to downtown Pittsburgh, an area with lots of polution. The downtown area had higher levels of carbon monoxide than the area with more green spaces showing the correlation between green spaces and air quality. Even though the differences are subtle and small, there is a correlation between green spaces and air quality. Areas with more green spaces do have better and less poluted air quality than spaces without. To improve the air quality in Allegheny County, we would suggest based off of our analysis, to make larger and more numerous green spaces.

By Yumika Amemiya, Sydne Ballengee, Sierra Brandegee, Erica Davis, Leah Ewers, Amelia Rosenstock, Alison Taylor, and Yolanda Zheng

from The Ellis School

The Correlation Between Gun Law Leniency and the Prevalence of Violence

Introduction

Does gun law leniency correlate with increased rates of gun violence in America? Is it possible that other factors/aspects are involved in the rates of gun violence? With the increasing number of debates over the Second Amendment, we deem it necessary to analyze the common belief that more guns constitute higher levels of gun-related violence in America. We used the number of background checks to help determine gun law leniency.

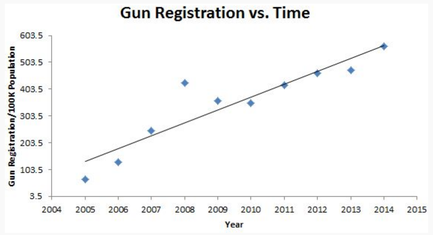
Challenges

1. Finding appropriate data sets for exactly what we wanted to analyze.
2. Determining exactly which relationships we wanted to analyze: urban/rural, various cities/states/countries, homicide/suicide/unintentional deaths, etc.
3. Standardizing data sets from various sources for analysis.
4. Combining different data sets to obtain desired data and still accurately represent the data
5. Finding data sets could be combined with each other (AKA finding data sets with a common variable)

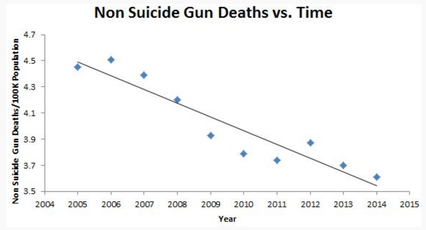
Akshath Jain, Benjamin Cheong, Grant Jiang, Janet Wang, Steven Lu

North Allegheny Senior High School

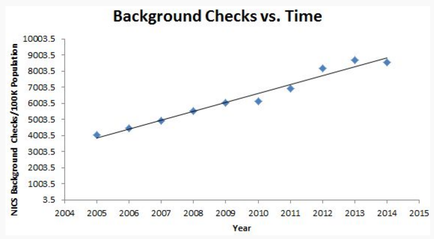
Data



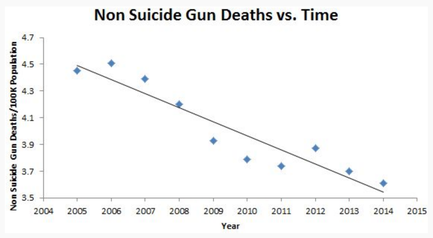
This graph shows the number of newly registered guns registered in the United States as a function of time. This graph also controls for population growth.



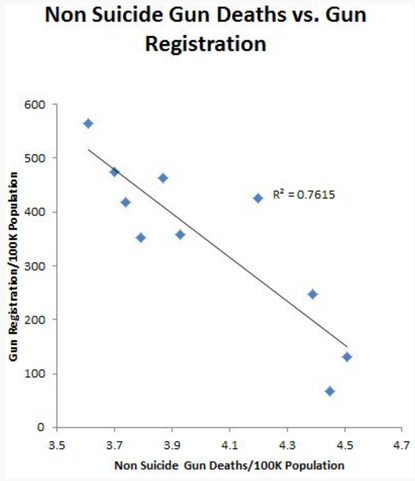
This graph displays the number of non-suicidal guns deaths as a function of time, and it also controls for population growth. A non-suicidal gun deaths is one that is either homicide or accidental gun death.



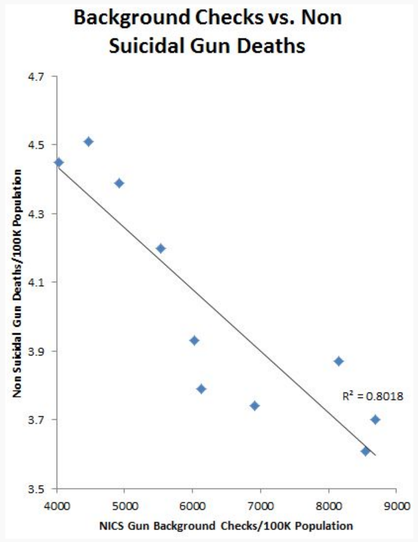
Here we see a plot of NICS (National Instant Criminal Background Check System) Background Checks (per 100,000 people) as a function of time.



Here we have the correlation of Non Suicidal Gun Deaths vs. Time. A non suicidal gun death is classified as either a gun related homicide or an unintentional gun death.



Here we combine the two data sets, eliminating the common variable of time to make a more useful correlation of Non Suicide Gun Deaths vs. Gun Registration. With an R^2 value of 0.76, we see a clear inverse correlation, contrary to our original belief.



Here we combine the two data sets on the left, again eliminating the common variable of time to create a more useful plot of Background Checks vs. Non Suicidal Gun Deaths. With an R^2 of 0.8 we see another clear inverse correlation.

Analysis/Conclusion

The common misconception that increased gun ownership causes an increased rate of gun-related violence is untrue, evidenced by our first plot. In fact, the number of non suicidal gun deaths have decreased over the years as newly registered guns in America continue to increase. After further probing, the rate of gun related violence was found to be correlated to stricter gun laws, demonstrated by our second plot. Thus, it would seem that having more responsible gun owners would be the ideal.

Recommendations

To continue this decreasing trend of gun-violence, continued passing of strict gun laws will prove beneficial. In addition, increased background checks will increase the number of responsible people who owns guns while decreasing the number of the number of irresponsible people. This, in turn, will lead to decreased numbers of unintentional gun deaths in the US. We also analyzed data from Hawaii, a state that is notorious for strict gun laws and thus a small number of guns.

Data Set Sources

"Bureau of Alcohol, Tobacco, Firearms and Explosives." Data & Statistics | Bureau of Alcohol, Tobacco, Firearms and Explosives. Bureau of Alcohol, Tobacco, Firearms, and Explosives, 24 Jan. 2017. Web. 13 Feb. 2017.

Krouse, William J. Gun Control Legislation. Tech. N.p.: Congressional Research Service, 2012. Print.

NICS. “NICS Firearm Background Checks: Month/Year.”

Correlation Between Police Officer Training Hours of Pittsburgh and Crime Rates in the City

Introduction

Although crime rates have been relatively low for the last several years, the media constantly focuses the public's attention on the horrific acts of violence that are occurring on a seemingly frequent basis. As such, we chose to do a study on the impact of police officer training with crime rates to determine if there was a correlation between hours invested and crime rates minimized. If there is a negative correlation, this will indicate that there is a benefit to increasing officer training hours in the city of Pittsburgh as it will decrease crime.

Challenges

1. Finding credible data sets
2. Correctly represent our data through graphs based on our data sets
3. Making a data-driven social policy recommendation

Data Sets

CLASS_NAME	OFFICERS_TRAINED_2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	TOTALS
*TOTALS	2,932	4,054	4,634	5,364	5,908	5,725	7,015	6,790	6,392	6,652	5,561	61,027
FIREARMS	828	808	764	891	1,171	1,270	1,209	1,383	1,360	1,398	1,231	12,313
FIRST AID/MENTAL HEALTH	287	259	313	266	308	306	355	327	1,107	511	1,169	5,208
FTO	17	11	11	0	24	0	49	82	9	89	15	307
IN-SERVICE	65	889	1,318	610	135	1,225	1,393	806	98	1,229	186	7,954
INSTRUCTOR TRAINING	72	54	91	42	81	46	83	63	83	35	38	688
INTELLIGENCE	0	1,000	7	0	1	40	0	662	1,393	6	2	3,111
INVESTIGATIONS	39	22	6	9	670	1,313	49	607	9	39	0	2,763
LEADERSHIP/SUPERVISOR	510	84	59	101	0	76	85	90	92	295	261	1,653
LEGAL UPDATE/LAW	500	514	565	580	707	621	637	695	705	1,517	780	7,821
LESS LETHAL	245	396	361	348	381	394	340	411	440	485	90	3,891
PROCEDUAL JUSTICE/ETHICS	0	0	0	0	0	0	0	0	15	809	847	1,671
SIGNIFICANT INCIDENT	0	0	1	1,746	671	22	636	696	2	0	130	3,904
SPECIAL UNIT	10	0	0	0	0	78	615	102	76	25	67	973
TACTICAL	38	16	1,118	771	1,046	22	1,258	0	749	150	24	5,192
TECHNOLOGY	321	1	20	0	713	312	306	866	254	52	721	3,566
UNION	0	0	0	0	0	0	0	0	0	12	0	12

References

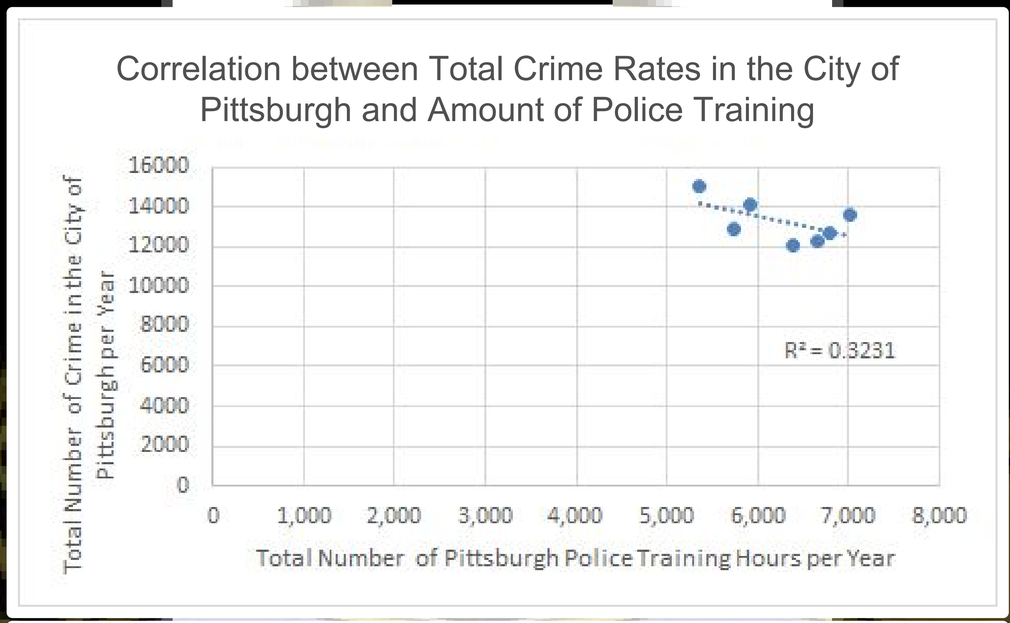
Western PA Regional Data Center

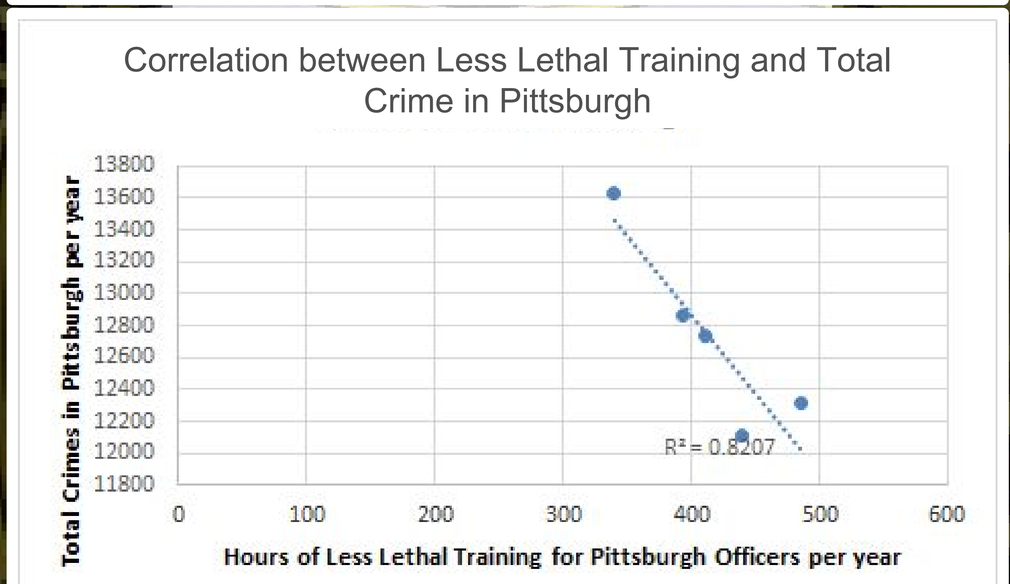
City of Pittsburgh Bureau of Police

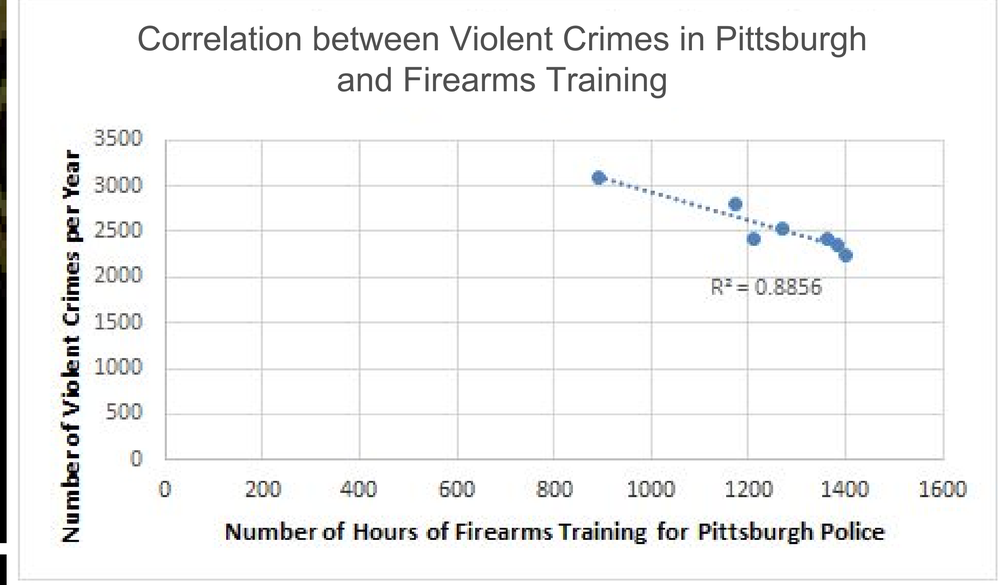
Pittsburgh Department of Public Safety

North Allegheny Team 1: Vivian Shao, Angela Liu, Kimberly DelSignore, Nur Iren, Remi Akindele

Results







Methods

Through the Western PA Regional Data Center and the City of Pittsburgh Bureau of Police annual report, we were able to compile and compare data on officer training and crime rates between the years 2009 and 2015 in Pittsburgh. We compared the occurrences of different categories of crime (i.e. violent crimes and property crimes) to the officer training hours in different specialties (i.e. less lethal training and firearms training) to find a relative correlation to determine if, and how, effective increased officer training is to deterring crime. Once we found a sufficient amount of data, we were able to determine the correlation between various categories of officer training hours and crimes and found the R^2 value to judge if there was a relationship between the two.

Analysis/Conclusions

There is a weak correlation between officer training and crime rates. We believe, however, that this correlation isn't indicative of Police effectiveness as the other graphs show how specific training can be very beneficial to police officers. There is a strong, negative correlation between the amount of firearms training police received and the amount of violent crimes in the city of pittsburgh. This indicates that police, with the proper instruction of firearms are able to make our city safer. However, firearms are not the only form of training that are beneficial.

Recommendation: In order to decrease not only violent crimes but the total crime rates in a city, police need to have training in less lethal methods. This can be seen as there is a strong negative correlation between total crime rates in the city and the amount of less lethal officer training.

Mary Cumpston
&
Kellen Heywood
&
Cassie Moats

IMMIGRATION

On the Fence?

Kara Puszko
&
Erin Sheedy
&
Jess Berger

Using data analysis to determine how immigration affects the Pittsburgh Economy

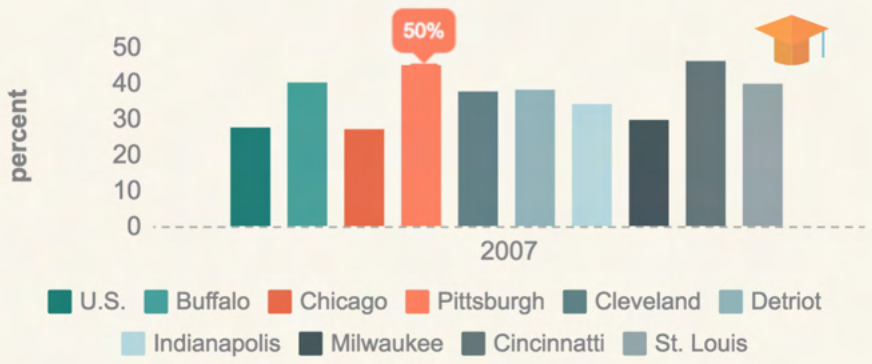
The Question

As the political climate changes drastically, the complaint that immigrants negatively affect the economy has come to the forefront. We set out to see if this was the case in the Pittsburgh MSA.

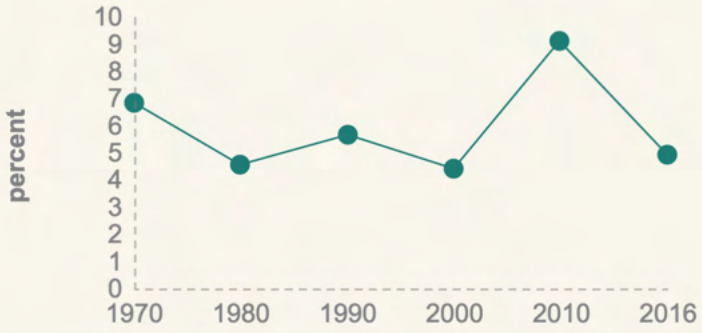
- How do employment rates correlate with immigration rates in the Pittsburgh MSA?
- What types of people immigrate to Pittsburgh and why?

The Data

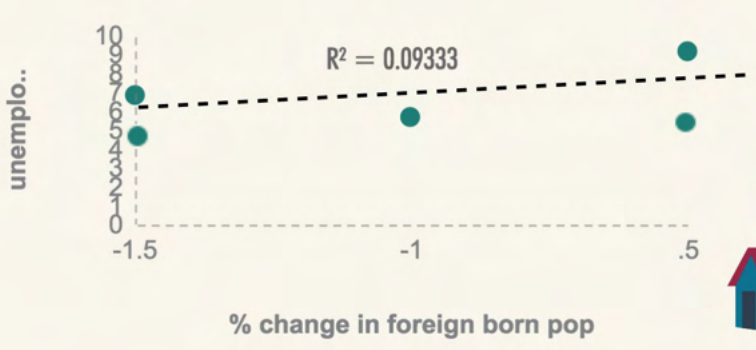
% of Foreign Born in Large Metro Area with at Least a Bachelor's Degree



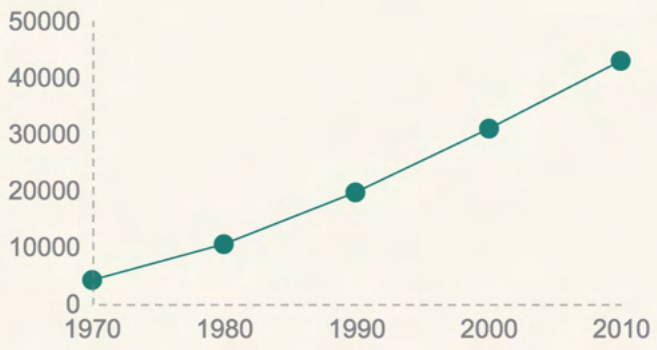
Pittsburgh MSA Unemployment rates



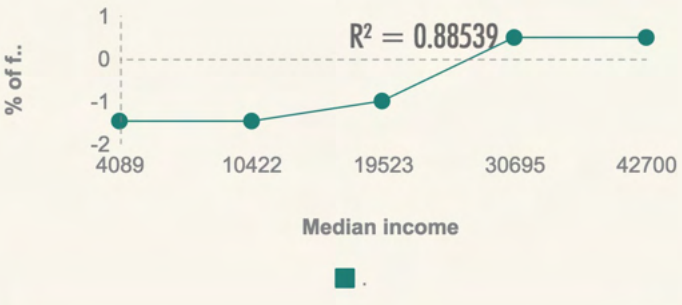
Percent Change in Foreign Born Population vs. Unemployment Rate



Pittsburgh MSA Per Capita Income



Percentage Change of Foreign Born Population vs. Median Income



"Mr. Peduto has sought to grow the city population by 20,000 during the next decade or so. More than 82,000 immigrant residents live in the 10-county Pittsburgh region, with the foreign-born population having increased nearly 8 percent from 2009 to 2014, according to the Partnership for a New American Economy. The overall regional population was stagnant."

Source: Pittsburgh Post Gazette. Feb 1, 2017 by Adam Smeltz

The Conclusion

Our data proved a weakly positive- if any- correlation between foreign-born population and median income in the Pittsburgh Area. However, we did learn that 50% of the immigrants living in Pittsburgh have at least a bachelor's degree, suggesting that our city attracts a highly educated foreign-born population.

With our data, we came to the conclusion that immigration causes no problems in the Pittsburgh MSA and therefore should be encouraged, especially because the people who tend to immigrate here have high levels of education because of our universities and hospitals.

We, as a team, struggled with finding immigration data that depicted the recent upward trend in immigration and learning to use Excel for larger, more complicated data sheets and comparing data sets.

Correlation Between Air Quality and Public Transportation in Pittsburgh

Lexi Jarvie, Kennedi Wade, Jing Li, Luisa Watkins, and Sylvia Li
Oakland Catholic High School Team 2

Research Question: How is air quality in the Pittsburgh area affected by the availability and use of public transportation?

Background information

Pittsburgh's particulate matter air pollution is among the worst 15 percent of cities across the country. Transportation contributes substantially to air pollution in the region. The immediate consequence of transportation on the environment is black carbon in the air. The indirect impacts are the health problems. For instance, the air pollution could accelerate the aging of the lungs and decrease lung function.

Addressing the Problem

We will compare the air qualities and the content of black carbon of different areas of Pittsburgh that have similar population size, but have different access to public transportation (we will calculate the number of stops each day). Then we will determine if there is a correlation between an area's air quality and the availability and frequency of public transportation in that area. (Black carbon is an important constituent of atmospheric aerosol particle matter (PM) with significant effects on the global radiation budget and on human health. Black carbon is produced both naturally and by human activities as a result of the incomplete combustion of fossil fuels, biofuels, and biomass.)

Data Visualizations

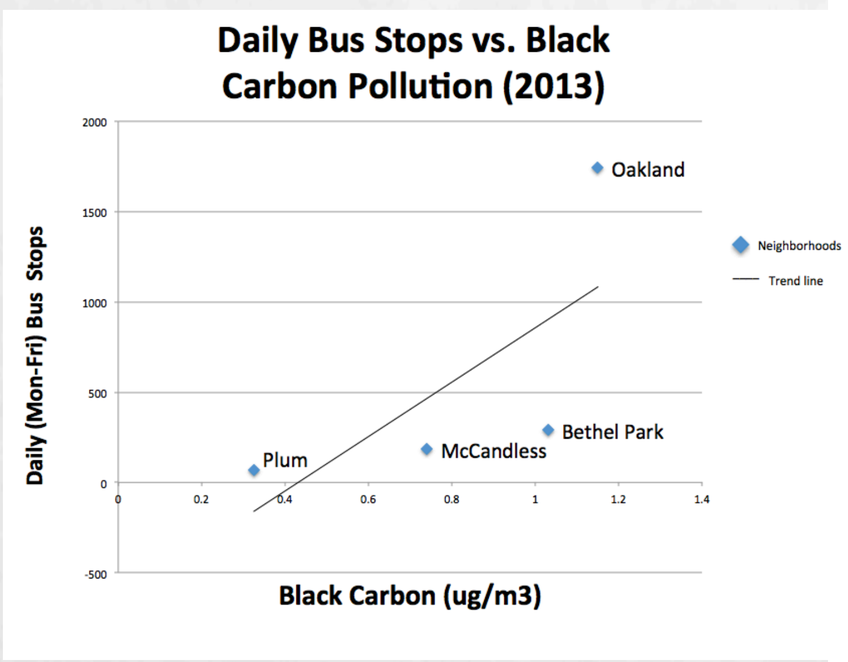
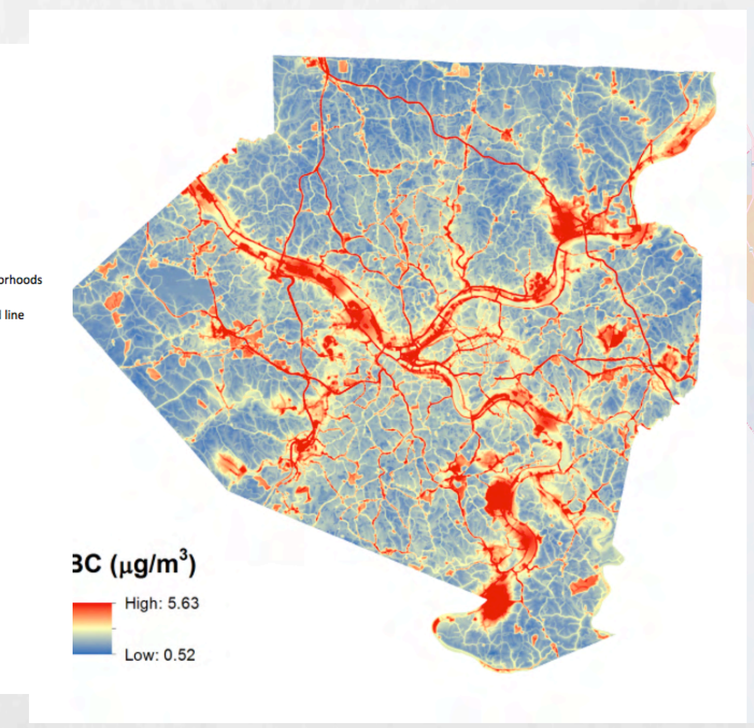
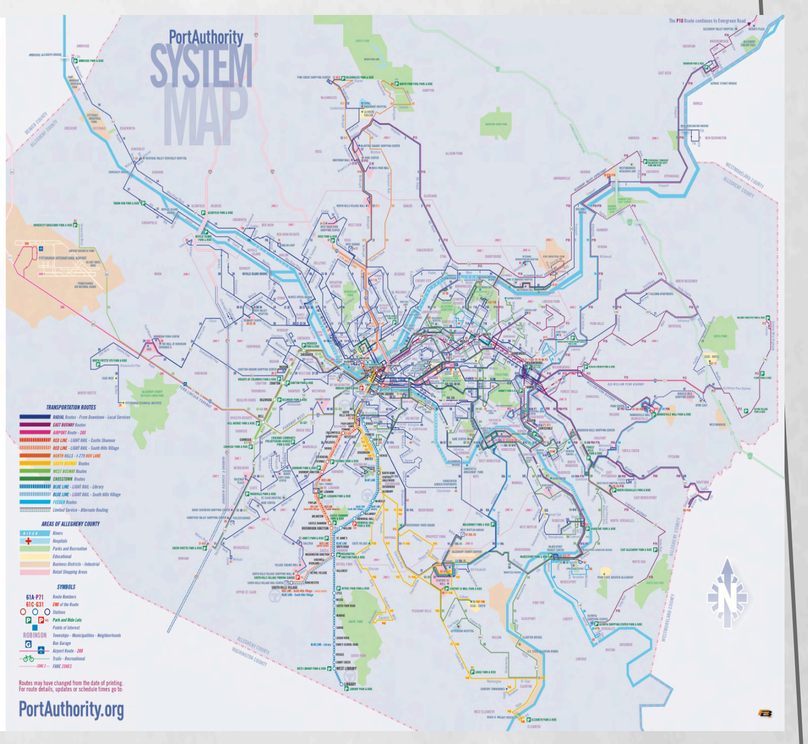


Chart of daily bus stops in different areas that have different content of black carbon.
(micrograms per cubic meter: ug/m3)



Map of the worst air pollution communities (in red) in Allegheny County.



Port Authority Transit System of Allegheny County

Analysis

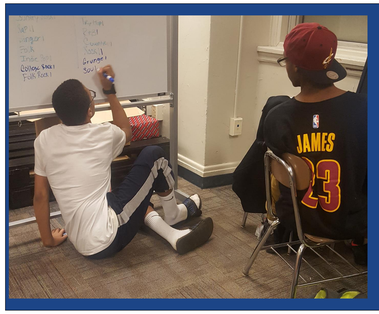
Based on the above chart and our findings, we can conclude that public transportation has an effect on air quality; the pollution is always more serious in those areas which have busier transportation. Even though we chose different areas in Pittsburgh with similar population sizes, most people work in the city, which probably leads to more transportation and worse air quality in the Oakland area.
In addition, air quality could be connected with vegetation coverage; if an area has high vegetation coverage, then it has less air pollution. Emissions from industries and manufacturing activities could also contribute to worse air quality because if industrial exhaust does not get well treated, it can cause air pollution.

Challenges

- The accuracy of air quality from different regions in Pittsburgh is hard to find.
- There are many factors that may affect this air quality, like industries and manufacturing facilities and vegetation coverage.

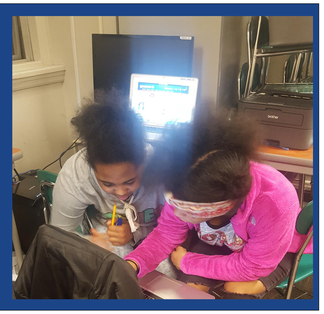
References

https://people.hofstra.edu/geotrans/eng/ch8en/conc8en/ch8c1en.html
http://www.post-gazette.com/news/environment/2015/02/11/Carnegie-Mellon-University-professor-makes-new-pollution-maps-public/stories/201502110023
http://wesa.fm/post/researcher-maps-pittsburghs-worst-air-pollution#stream/0
http://www.portauthority.org/paac/SchedulesMaps/Maps.aspx



You Are The Music In Me: Political Parties & Music Preference Scholar Project

Kyla Gardner, Naketta Williams, Cornel Collins, & Cori Parks

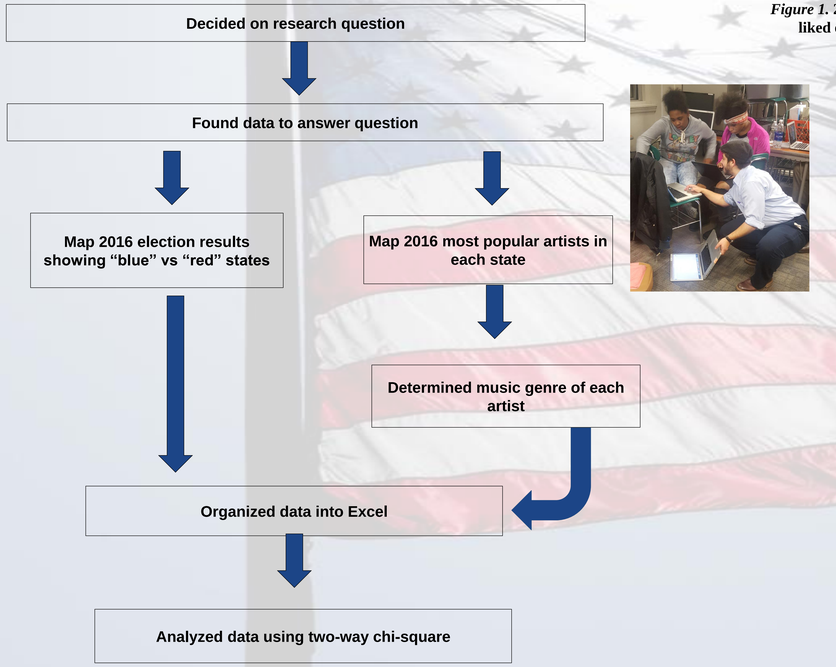


Introduction

Does musical preference predict political preference?

Music can impact which political candidate a person votes for (Pham 2016). Each state has an overall preference for an artist and genre (The Echo Nest Blog). We are interested in the relationship between whether a state voted red or blue, and what music the state prefers. The relationship between music preference and political preference is important because it could be potentially informative to politicians as they court voters. What we learned today is that we did a graph on the states with their music genres. We predicted that Democrats would listen to R&B/Rap/HipHop and that Republicans would listen to Rock and Country.

Method



Results

Genre	Number of Blue States	% Out of 20	Number of Red States	% Out of 30
Indie	3	15	6	20
Electronic	2	10	0	0
R&B/Rap/HipHop	4	20	9	30
Country	2	10	6	20
Rock	8	40	8	27
Worship	0	0	1	3

Figure 1. 20 states voted blue and 30 voted red. The table shows the number of blue states that liked each genre and red states that liked each genre. It also calculates the percentage of blue states that liked each genre and red states that liked each genre.

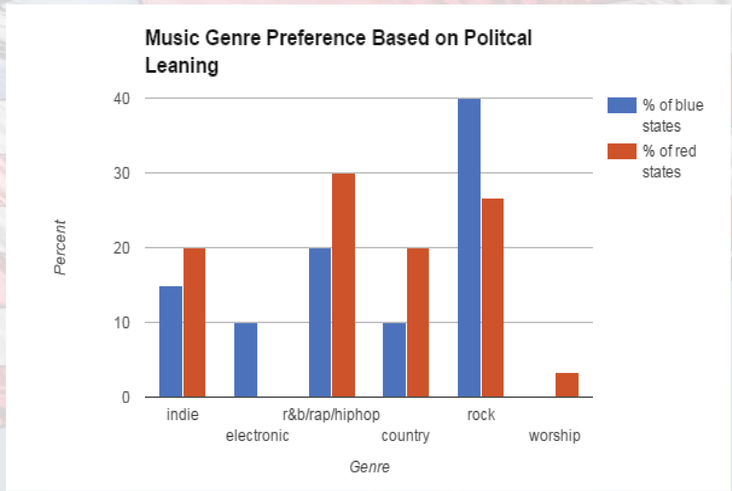


Figure 2. The bar compares the percentage of blue or red states that prefer the genre shown on the y axis.

Two-way chi-square showed no statistical significant differences in musical genre preference by political party preference of state.

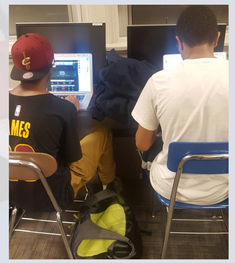
Conclusions

- **Percentage blue states that prefer rock is much higher than any of the others**
  - **Not what we predicted**
- **Higher percentage blue states preferred rock compared to red states.**
  - **Not what we predicted**
- **No red states prefer electronic.**
- **More red states than blue states prefer country**
  - **What we predicted!**
  - **Bar graph shows that this is true**

Possible that findings would differ at city and individual level and over time
Possible that region of country, urban/rural, and other factors matter more than political preference of state

Challenges

- **Deciding research interest**
- **Deciding HOW to address question**
- **Deciding on best data to use**



References

https://musicmachinery.com/2014/02/25/exploring-regional-listening-preferences

http://www.270towin.com/

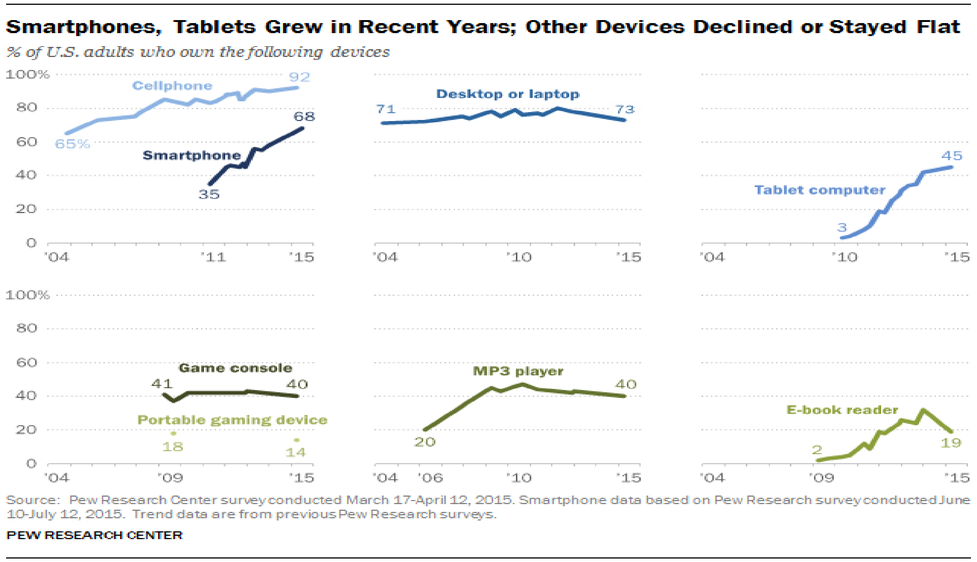
http://www.forbes.com/sites/alexpham/2016/09/06/musical-tastes-of-trump-and-clinton-voters-led-zeppelin-or-linkin-park/#6ae771241f19

The Impact Texting Has Had on Car Crashes Over the Years 2004-2015

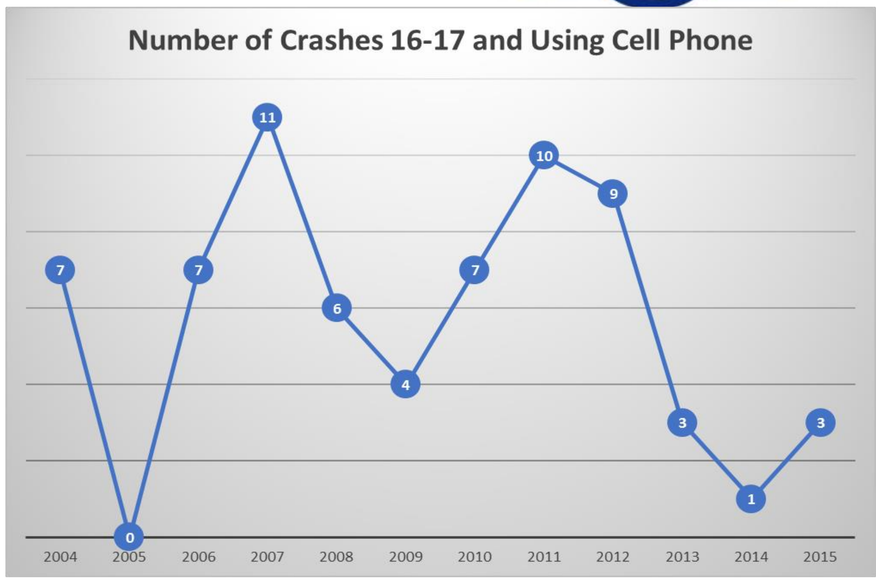
Perry High School: DJ Clark, Brianna Kenney-King, Ethan Shomo, Lillie Harris Briggs, Naomi Ilochi and Elijah Ross

Introduction:

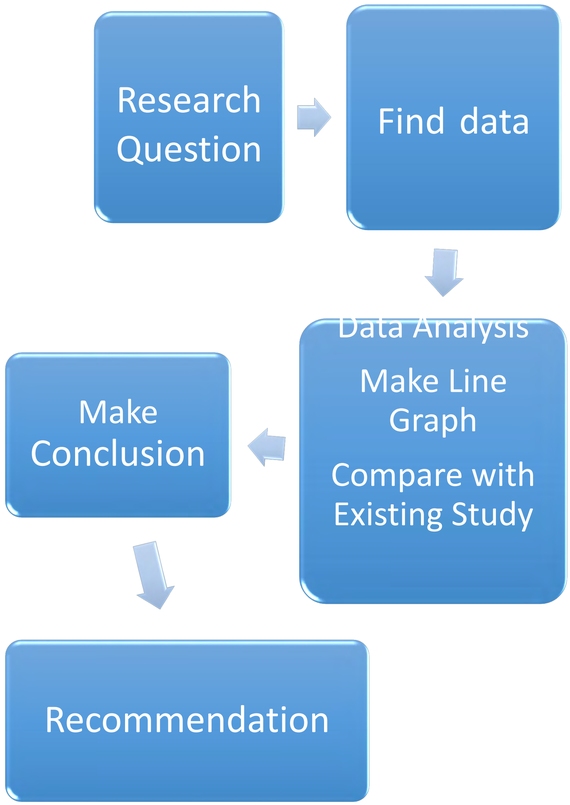
A major issue in today's society is texting while driving. This has caused many accidents and deaths over time. We are trying to reduce the amount of car crashes due to texting while driving with the help of modern data.



A study done by Pew Research Center



Methods:



Conclusion:

Data Analysis helped to create visual representations of car crashes caused by cell phone usage allowing us to be able to convey the data in a way that people will easily be able to understand. We were able to identify the factors that made people more likely to crash. One of the correlating factors is age. Cell phone ownership has increased since 2011 but crashes have decreased. Education on the dangers of cell phone use will driving may be one of the reason for this. In addition, people other than 16- and 17=year-olds may be getting cell phones, contributing to the rise in ownership, but not a rise in crashes.

Policy Recommendation:

We believe that the most efficient way to reduce the number of crashes due to texting is to make people more aware of the dangers. High schools should be educating their students on the dangers because the 16-17 age group are probably one of the most likely groups to text while driving.

X-axis-years
Y-axis-number of crashes

Year	Number of Crashes 16-17 and Using Cell Phone
2004	7
2005	0
2006	7
2007	11
2008	6
2009	4
2010	7
2011	10
2012	9
2013	3
2014	1
2015	3

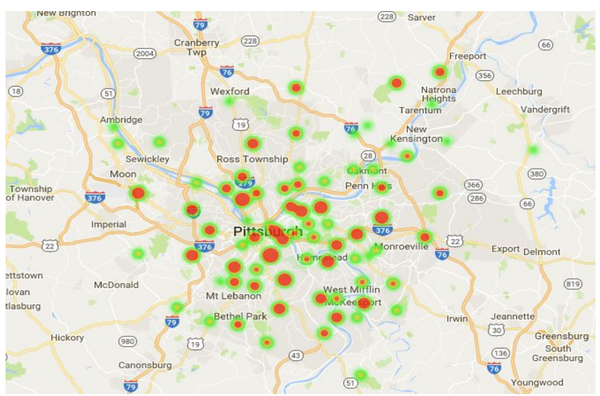
Reference:

https://www.distraction.gov/
http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/pi_2015-10-29_device-ownership_0-01/
https://data.wprdc.org/terms-of-use?came_from=%2Fdataset%2Fallegheny-county-crash-data

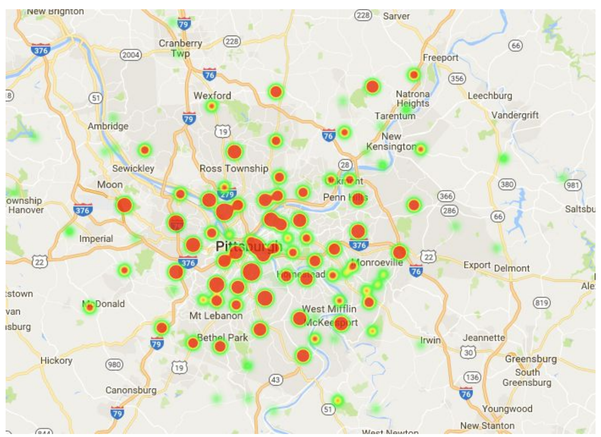
Propel Andrew Street High School: Drug Overdoses in Western PA

Casjmier Malone, I'onna Pitts, Daniel Williams, Taye Halliday

Fentanyl Density Map



Herion Density Map



Research Question

Are there patterns of death by drug overdose over time in different areas of Western PA?

Dataset

Western PA Regional Data Center- Allegheny County Fatal Accidental Overdoses

Our data shows that the most frequent overdosed drugs are heroin, alprazolam, cocaine, alcohol, fentanyl.

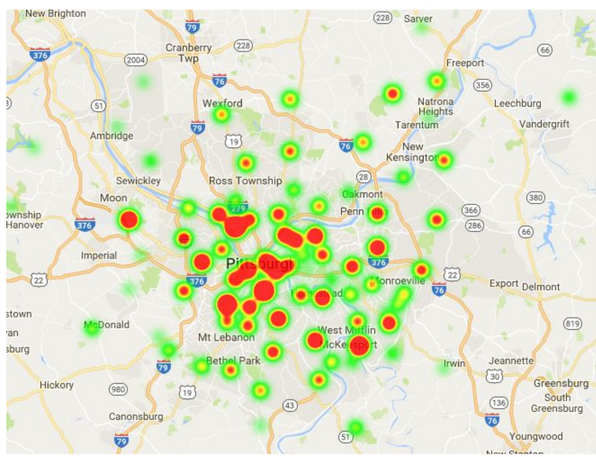
Problems

We mostly had problems with missing or bad data in the dataset. We noticed that there were a limited amount of individuals under 18 years old in our data. We also found that some of the data was inputted incorrectly.

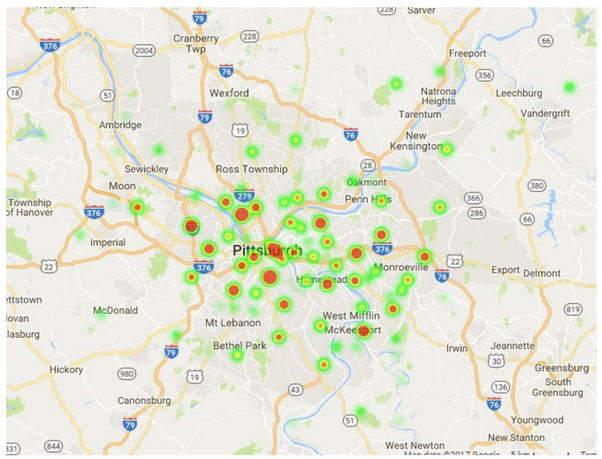
Summary

Our data shows some variation of the amount of overdoses by location and year. We did not find that there was significant change over time.

Alcohol Density Map



Cocaine Density Map



Pittsburgh Science and Technology Academy
TEAM 2

PITTSBURGH FLIGHTS AND AIRPORT TRAVEL COSTS

Alexandra Borelli, David Donehue, Nicholas Naumov, Michael Warren, Sophia Yurkovetsky, and Joshua Zito

RESEARCH QUESTION

What effect has airport travel costs had on flight frequency in Pittsburgh?

HYPOTHESIS

A decrease in airport costs has caused a decrease in flight frequency.

DATA SET EXAMPLE & SOURCE(S)

All data comes from Bureau of Transportation Statistics

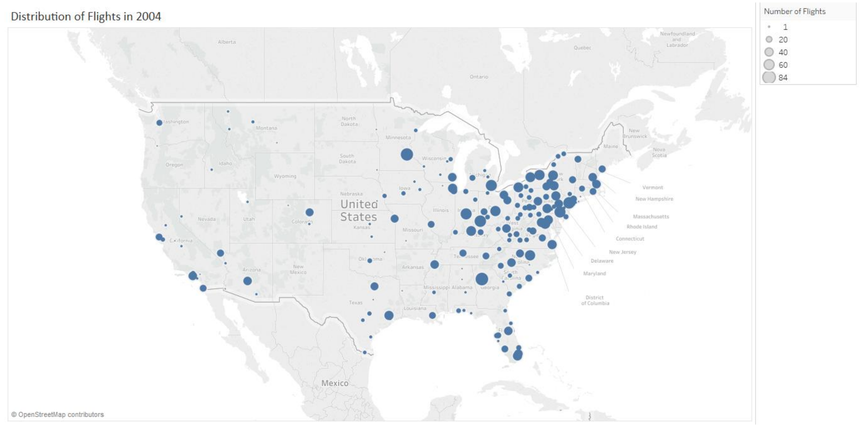
Preview of Data

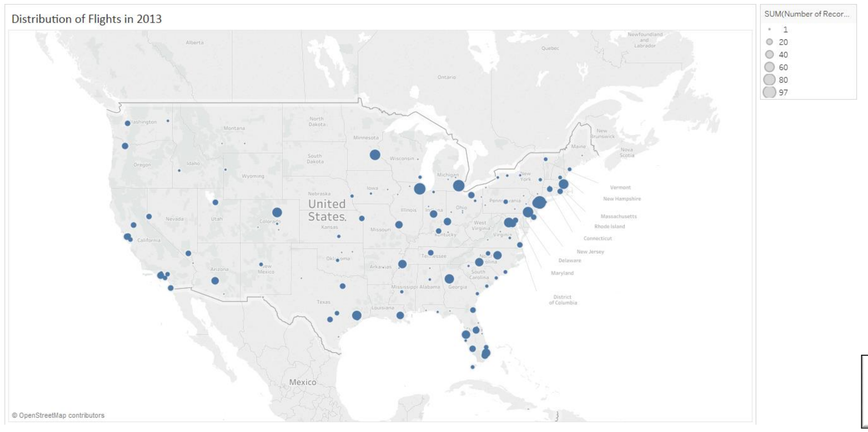
Year	Quarter	U.S. Average (Current $)	U.S. Average (Inflation-Adjusted $)	Pittsburgh, PA (Current $)	Pittsburgh, PA (Inflation-Adjusted $)
1995	1	296.9	473.44	398.37	635.26
1995	2	296.8	469.87	373.34	591.04
1995	3	287.51	453.09	372.05	586.31
1995	4	287.78	452.62	378.99	596.09
1996	1	283.97	440.32	377.38	585.17
1996	2	275.78	424.89	362.73	558.85
1996	3	269.49	412.31	366.34	560.49
1996	4	278.33	423.68	374.8	570.54
1997	1	283.4	427.63	396.77	598.69
1997	2	289.44	435.92	404.15	608.68
1997	3	282.27	422.76	408.7	612.1
1997	4	293.51	439.32	419.76	628.28
1998	1	304.74	453.6	434.26	646.38
1998	2	300.97	445.78	415	614.68
1998	3	315.25	465.23	404.75	597.3
1998	4	316.18	465.74	405.4	597.17

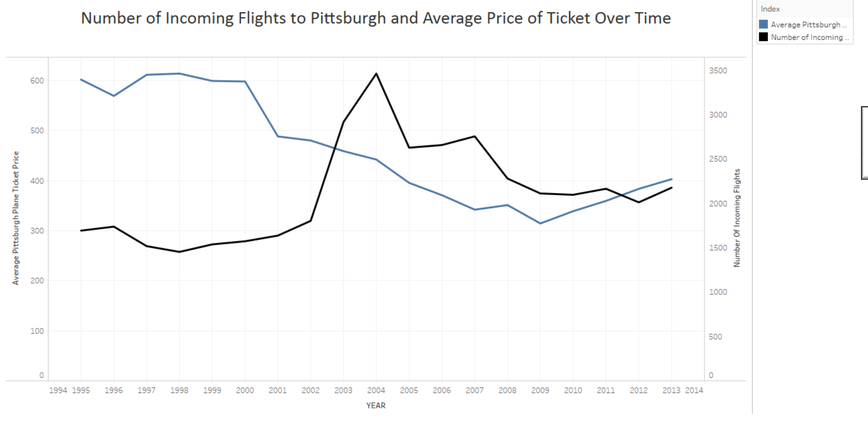
- This dataset comes from a source calculating the average quarterly costs of ticket prices .The quarter values were averaged when imported into graphs.
- The other datasets list every commercial flight in the US with US carriers. This includes information such as Carrier Code, Carrier Name, destination, origin, class, etc.

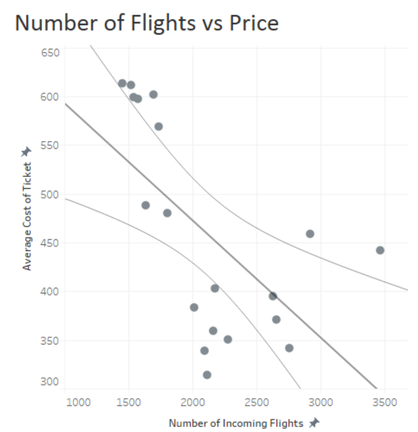
VISUALIZATIONS

Pittsburgh Commercial Flight Origin Distributions









- 37.8% of the variability in the data can be explained by the linear association between average cost of ticket and number of incoming flights. $R^2 = .378$ $R = -.615$

Significance Tests (number of flights)

- Assumptions are met to perform a T-test on all samples.
- $\bar{x}$ = 2116.84 (2004,3464) (2013, 2177)
- H_0: 2004 = $\bar{x}$ / H_a: 2004 ≠ $\bar{x}$ / α = .05
- P = 3.727 * 10^{-9} Reject H_0
- H_0: 2013 = $\bar{x}$ / H_a: 2013 ≠ $\bar{x}$ / α = .05

Significance Tests (average cost of ticket)

- Assumptions are met to perform a T-test on all samples.
- $\bar{x}$ = 458.92 (2004,442) (2013,402.8)
- H_0: 2004 = $\bar{x}$ / H_a: 2004 ≠ $\bar{x}$ / α = .05
- P = .5048 Fail to Reject H_0
- H_0: 2013 = $\bar{x}$ / H_a: 2013 ≠ $\bar{x}$ / α = .05
- P = .0367 Reject H_0

CHALLENGES

- This data is not an explicit google search, which required contacting officials from websites, and performing advanced database searches, very little data was available.

SUMMARY

- As price of travel for Pittsburgh has generally declined over the years, this could have been caused by the decrease in incoming flights.
- Recommendation to get an increase in flights such as in 2004 would be to make Pittsburgh another major airline hub (US Airways hub had moved from Pittsburgh to Charlotte, NC). This is ideal considering prices being low.

High School Choices

Should students go to their home district or choose to attend a different type of school?

Propel Andrew Street High School

Zachary Grimm
Jayla Moore
Laraun Josey
Kearrah Barlow
Eddie Banks-Hicks

Diversity

Row Labels	Average of Most Diverse School in Pittsburgh Ranking	Average of diversity GPA
Charter	61	2.75
Magnet	33	3.25
Public	42	2.48
Private	N/A	2.272727273

Career and College Readiness

Row Labels	Average of Best Public High school in Pittsburgh Ranking	Average of STEM Ranking in Pennsylvania	Average of College Readiness Ranking
Charter	28	85	76
Magnet	46	51	40
Public	47	43	45

Extracurricular Activities

Row Labels	Average of Best High school for Athletes Ranking	Average of Club and Activity GPA
Charter	36	3
Magnet	42	3
Public	40	3
Private	N/A	3

Test Scores

Row Labels	Average Math Academics	Average Reading Academics	Average SAT Scores
Charter	16%	36%	1135
Magnet	41%	59%	1083
Public	63%	75%	1090
Private	N/A	N/A	1174.545455

We created our own dataset using two sources. First, we used data found on the Niche.com website because of the comprehensiveness of the information found there. We also included data from the PA Department of Education for test scores in 2016.

Problems

We had some problems with finding data about Private and Catholic Schools. They are not held to the same reporting standards as Public Schools. Also, the SAT Scores were self reported and may skew high.

Summary

The data that we compiled would be somewhat helpful to a family trying to decide what type of schooll to attend.

John Rogers
Angeleia Hall
Kayla Klink
Mekala Kaiser
Team 1 , Pittsburgh SciTech

Homicides In Allegheny County

Research Question:

What are some contributing factors of homicide in Allegheny County?

Hypotheses:

1)In Allegheny County, an increase in high school dropout rates will cause an increase in homicide rates because with more teenagers out on the streets and without guidance, more violent crime is likely to occur.

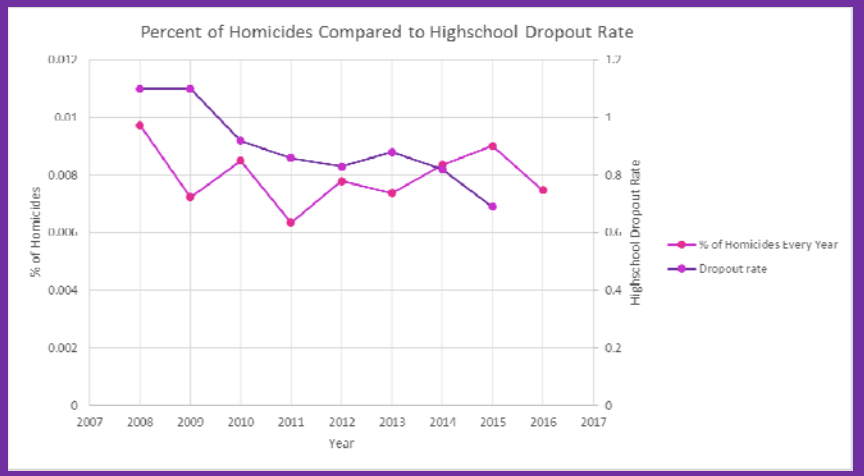
2)We anticipate discovering previously unknown factors that contribute to crime problems that may exist in Allegheny County communities

Challenges Faced:

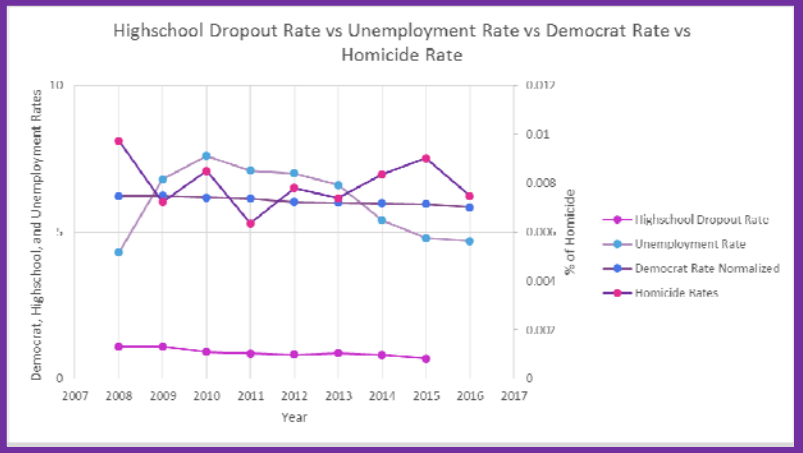
Hard to get race populations per year, as sometimes hard to collect data from the same source per year.

Hard to decode U.S. Census Excel sheets

Difficult getting race statistics by county, since data collected per year varied in their representation.



This graph shows the correlation between percent of homicides and high school dropout rates per year. Overall, the graph shows a trend that there is no correlation between dropout rate and homicides .For example, in 2009, we see a .011% of dropout rates but only .007% of homicides in Allegheny County. Only in 2015, in fact, are the percentages of homicides greater than the dropout rates. This does not support the hypothesis that as dropout rates increase, homicide rates increase.



Conclusions:

There is no strong correlation to show that an increase in dropout rates cause an increase in homicide rates. One main reason is that over the span of 2006-2015, dropout rates have steadily decreased with no visible spikes . Interestingly, however, we do notice that as the number of registered Democrats in Allegheny County decreases, homicide rates increase. This is an odd observation, as this could suggest that to reduce homicide rates, the number of registered Democrats needs to go up. We observe that although homicide rates have varied, in the past decade they have decreased.

Dataset used: American Census, Education PA, PA Unemployment Rates, Bureau of Labor Statistics

Comparing Pittsburgh and The Rust Belt

Alec Helbling, Joseph Flot, Sree Mekala, Liam Hainsworth, Gibran Biswas, Meredith Harrison

How do education and infrastructure relate to economic growth and livability?

Introduction

Background:

- Rust Belt is the region extending from the Great Lakes to the upper Midwest States .
- Known for its economic decline, population loss, and urban decay due to the shrinking of its once - powerful industrial sector.
- Some Rust Belt cities have recovered from this industrial collapse; others have not been so lucky.
- Pittsburgh has recovered immensely from economic misfortune.

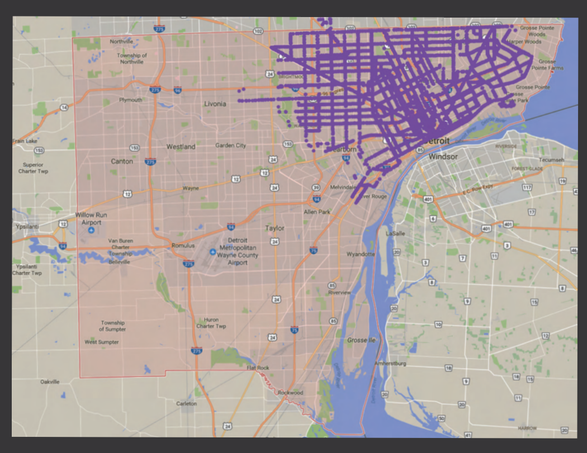
Hypothesis:

- Allegheny County, serving as a representation of Pennsylvania, will have better education and infrastructure, than other Rust Belt cities/counties as it has improved from enconomic decline and is generally considerable highly livable.

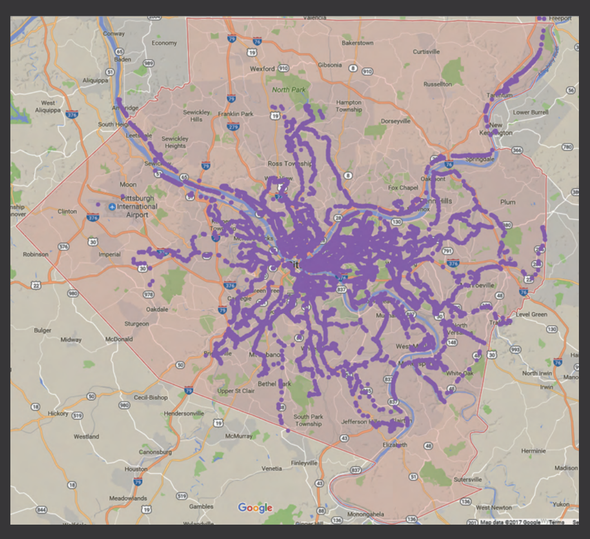
Infrastructure

- The left shows Wayne County's (Detroit) Public Transportation System.
- The right is Allegheny County's PAT system.
- Detroit's transportation system covers considerably less area.
- This could be an indicator of denser population or of a less extensive and passenger friendly transportation system.

Detroit Transportation
Wayne County Area: 673 sqmi



Pittsburgh Transportation
Wayne County Area: 745 sqmi



Analysis and Conclusions

Analysis:

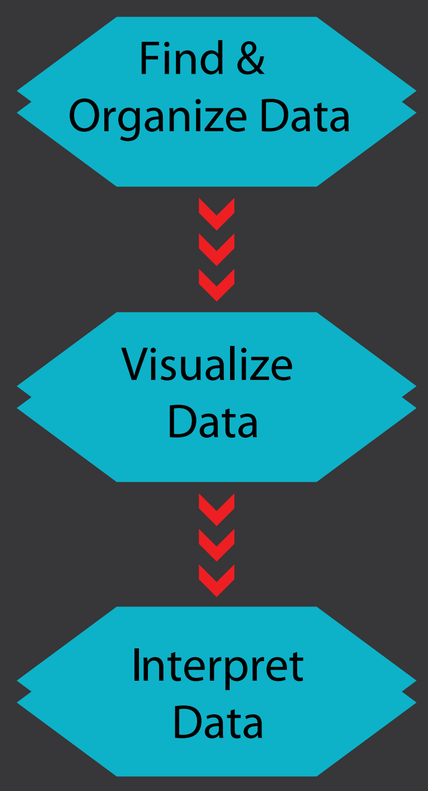
- Unemployment data shows Allegheny County (Pittsburgh) to have the lowest unemployment whereas Wayne County (Detroit) had the highest.
- The dropout data shows that Allegheny has the lowest dropuout rate as well as the lowest unemplloyment rate
- The linear regression plot with a R-Squared value of 0.399 (.698 when excluding Wayne County and Lake County)

Signifigance of Findings and Recommendations:

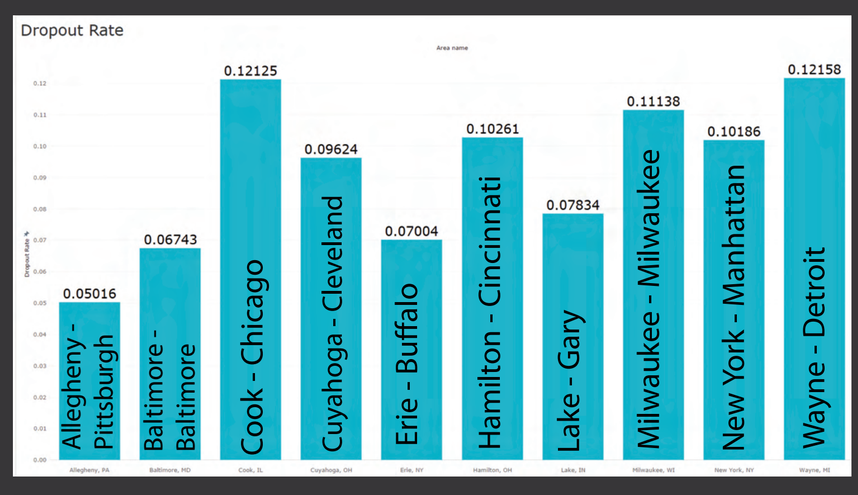
- Highschool and Unemployment data suggests that Allegheny county is justifiably improved compared to other Rust Belt counties.
- Dropout rates and unemployment rates are an excellent way to analyze livability as they are practical improvements which require a good education system to maintain, indicating funding to support them which shows economic growth.
- On a similar note, the diagrams of bus transportation show strength of the infrastructure in Detroit and Pittsburgh. Pittsburgh, while only one city, can cover more of its (larger) county than Detroit can; indicating the economic ability to support such a system.

Methods

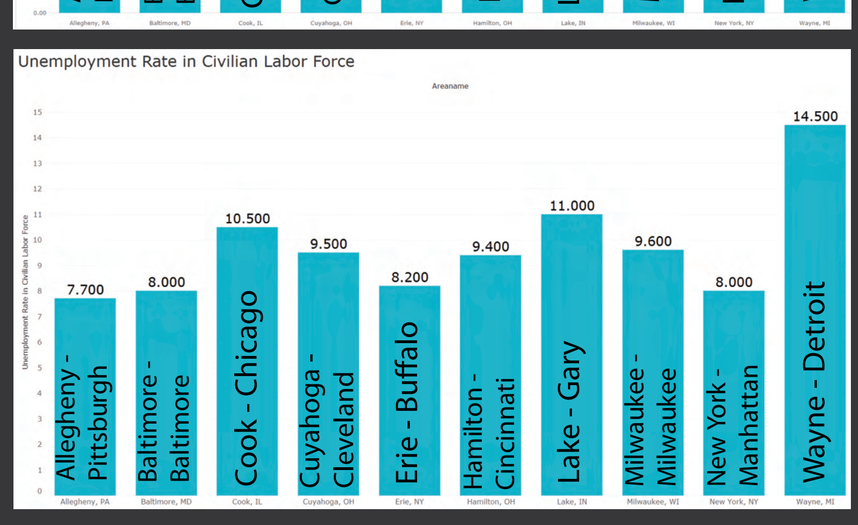
1. We first found the US Census Counties Dataset and constructed a legend for using it.
2. We then visualized the dataset using Tableau Public based on various variables.
3. We finally interpreted and analyzed the data based on the vizualizations.



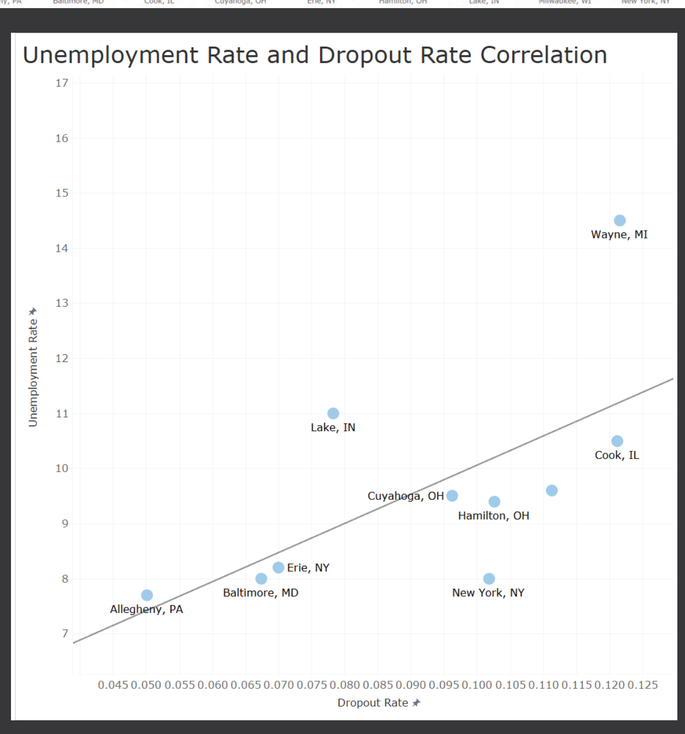
Education



•Allegheny county seems to have the lowest dropout rate.



•Allegheny county has the lowest unemeployment rate as well.



•The linear regression plot with an R-Squared value of 0.399 (.698 when excluding Wayne County and Lake County)
• r =.631 or .836

- The correlation between dropout rates and unemployment is posi - tive and relatively strong
- The high dropout rate could be indicative of the high unemployment or the opposite could be justified

Challenges:

- Finding relevant and proper data
- Applying this data effectively to our research question.
- We found a dataset with all of our desired cities for the year 2000, which is not as recent as we wished.

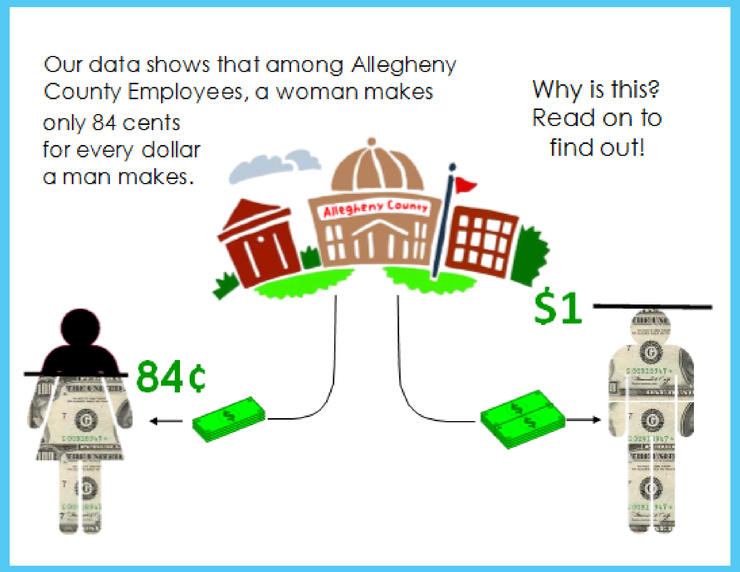
References:

- http://geography.about.com/od/urbaneconomicgeography/a/Rust-Belt.htm
- https://www.census.gov/support/USACdataDownloads.html
- http://www.portauthority.org/paac/CompanyInfoProjects
- http://www.detroitmi.gov/How-Do-I/Locate-Transportation/Bus-Schedules

Are Women Underpaid in Allegheny County?

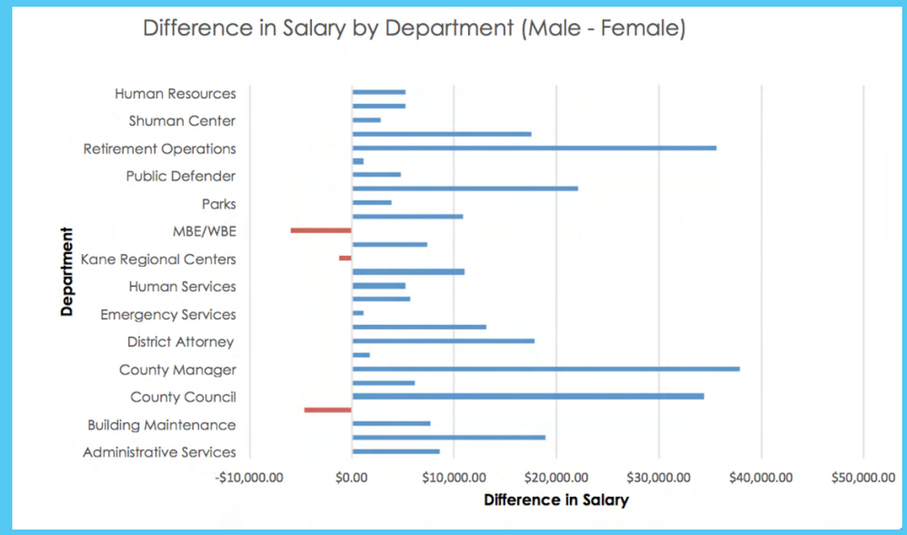
Lucas Lang, Griffin Mackey, Daniel Phillips, Pasha Sachivichik, Aydin Turkay

Sewickley Academy

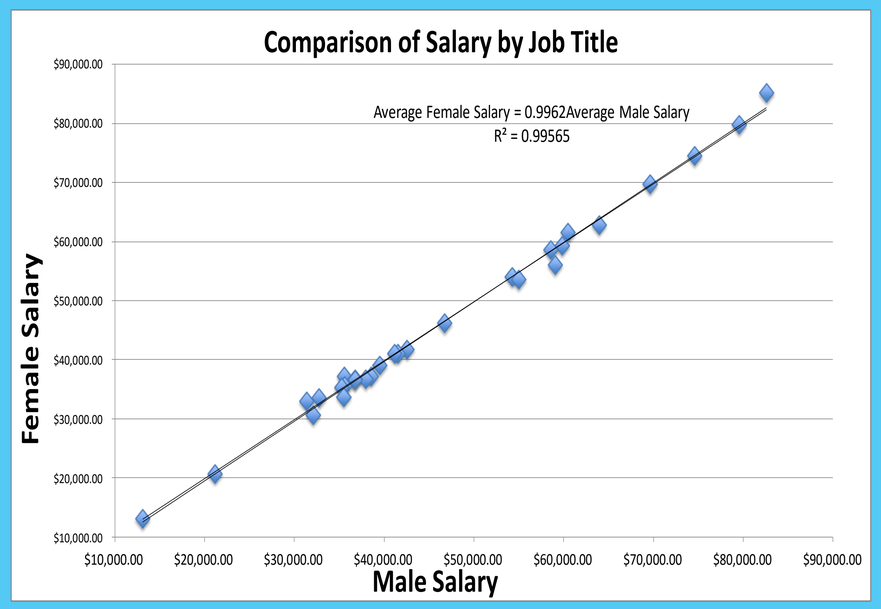


=AVERAGEIFS(E2:E6461, D2:D6461, "Agriculture",B2:B6461, "F")

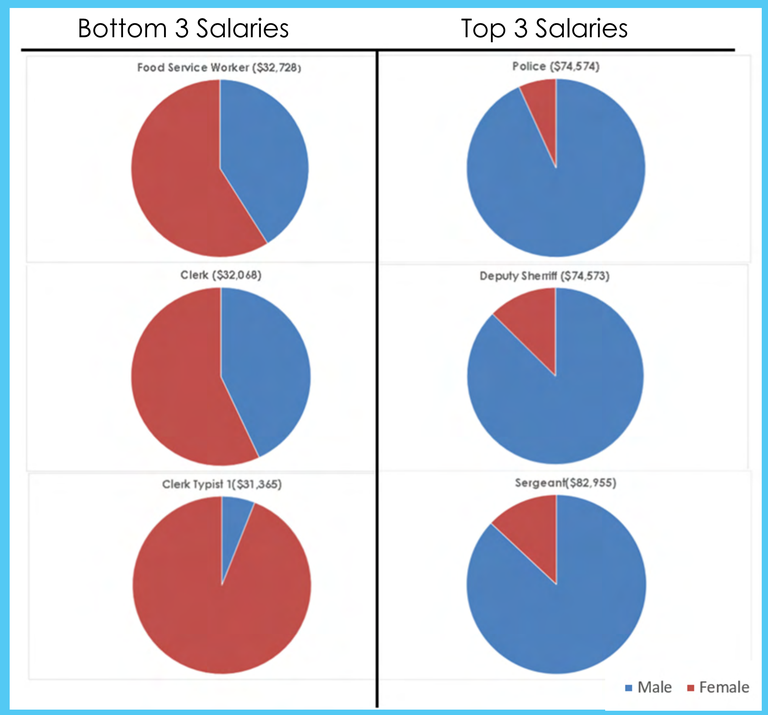
	B	C	D	E	
	Gender	**Job Title**	**Department Name**	**Annual Salary**	**Reg**
	F	NURSING ASSISTANT	Kane Regional Centers	32,931.81	31
	M	BLDG AUTOMATED SYS TECH	Building Maintenance	48,022.42	52
	F	CLERK	Court Records	29,173.46	15
	F	COOK	Kane Regional Centers	35,381.63	34
	M	FOREMAN BUILDING MAINTE	Building Maintenance	39,315.95	38
	F	FORENSIC INVESTIGATOR	Medical Examiner	40,245.41	31
	F	MANAGER REGISTRATION	Administrative Services	44,044.56	42
	F	CHILD WELFARE AIDE	Human Services	36,413.07	35



A majority of departments show that men earn more than women on average.



As the graph shows, there is little to no difference in salary when sorted by position.



Men hold the majority of the three highest paying positions, whereas women hold a majority of the three lowest paying positions.

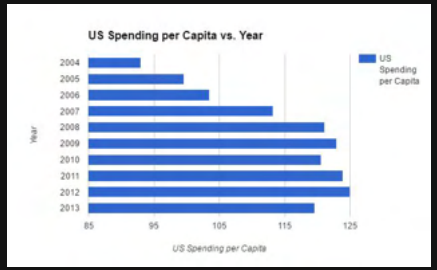
Recommendations

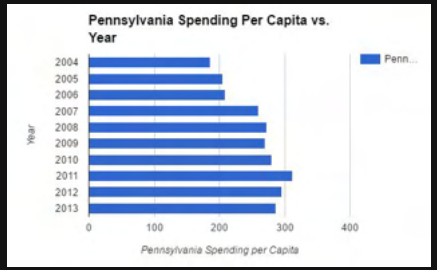
- Unequal distributions among positions are the underlying reason for a gender gap.
- More education to the public about salaries for different positions.
- Recruiting more women into law enforcement positions.
- Rather than focusing on increasing pay for women, the county should work on employing an equal number of men and women in each position.

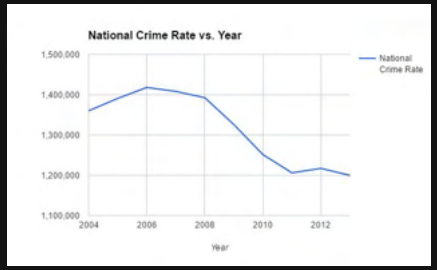
Spending on Mental Health Affect on Crime Rates

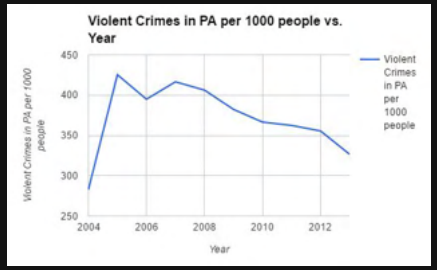
Introduction

When individuals receive help with disabilities involving their mental state, they are less likely to commit violent crimes. Studies show that countries with higher crime rates experience a lack of economic growth that could be detrimental to society. This rising crime rate also lowers the trust of citizens in the government which could even lead to an economic recession.









References:
Data.gov
Kff.org
Disastercenter.com

Nina Dorfner, Hannah Finestone, Haley Marinack, Joey Black, Joseph Froetschel, David Gubinsky

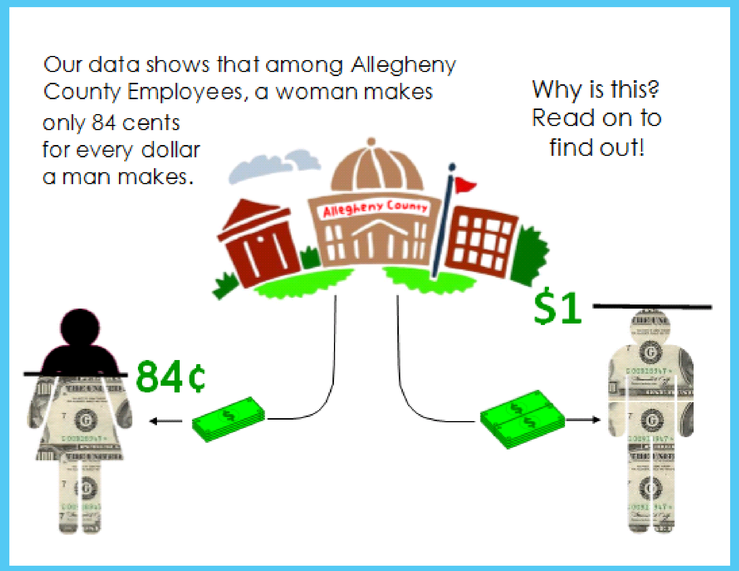
Analysis

With the aid of graphs depicting the government spending of mental health versus the crime rate in both the US and state of Pennsylvania, it was easy to conclude that the general trend of increased mental health spending directly correlates to a decreased crime rate. A specific instance of this is between the years of 2008 and 2010 in which the US government continued to increase the budget for mental health. However, this data does not account for a change in population for either location.

Are Women Underpaid in Allegheny County?

Lucas Lang, Griffin Mackey, Daniel Phillips, Pasha Sachivichik, Aydin Turkay

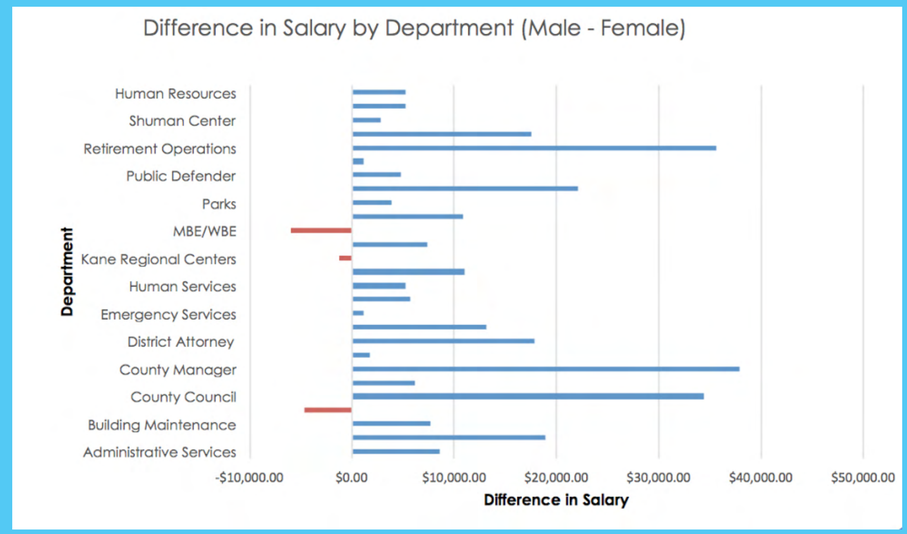
Sewickley Academy



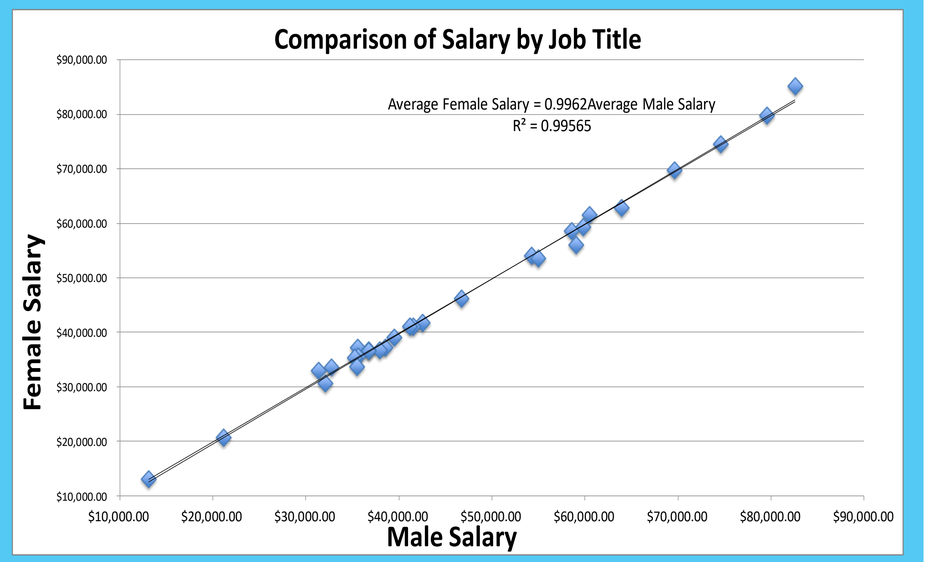
=AVERAGEIFS(E2:E6461, D2:D6461, "Agriculture",B2:B6461, "F")

	B	C	D	E	
	Gender	**Job Title**	**Department Name**	**Annual Salary**	**Reg**
	F	NURSING ASSISTANT	Kane Regional Centers	32,931.81	31
	M	BLDG AUTOMATED SYS TECH	Building Maintenance	48,022.42	52
	F	CLERK	Court Records	29,173.46	15
	F	COOK	Kane Regional Centers	35,381.63	34
	M	FOREMAN BUILDING MAINTE	Building Maintenance	39,315.95	38
	F	FORENSIC INVESTIGATOR	Medical Examiner	40,245.41	31
	F	MANAGER REGISTRATION	Administrative Services	44,044.56	42
	F	CHILD WELFARE AIDE	Human Services	36,413.07	35

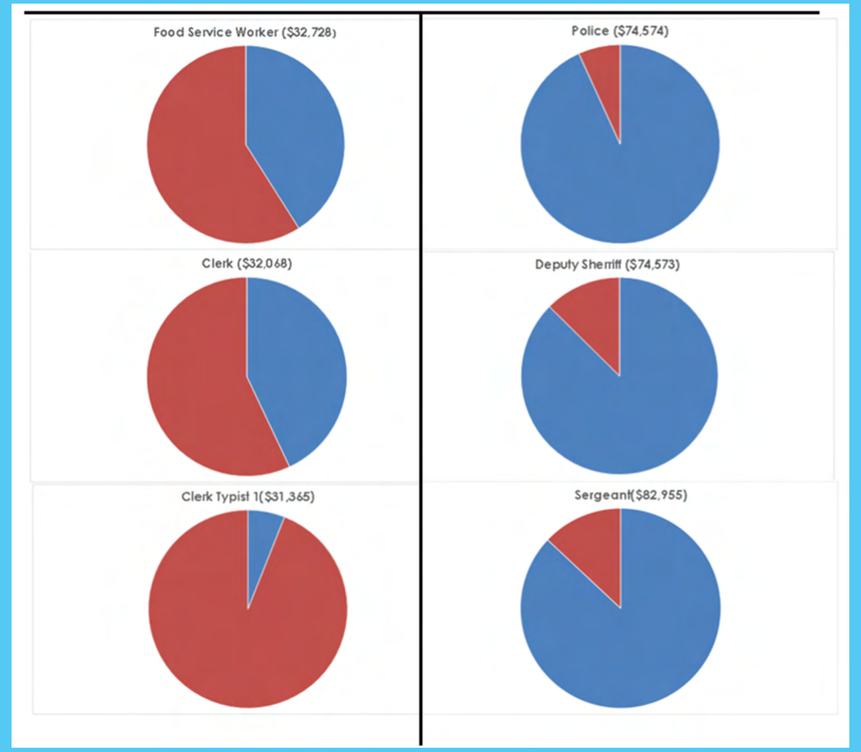
Allegheny County Employee Salary Data, 2015



A majority of departments show that men earn more than women on average.



As the graph shows, there is little to no difference in salary when sorted by position.



Men hold the majority of the three highest paying positions, whereas women hold a majority of the three lowest paying positions.

Recommendations

- Unequal distributions among positions are the underlying reason for a gender gap.
- More education to the public about salaries for different positions.
- Recruiting more women into law enforcement positions.
- Rather than focusing on increasing pay for women, the county should work on employing an equal number of men and women in each position.

Impacts Environment Has on Obesity

Research Question: Does the conditions of each state impact obesity?

Helina VanBibber, Savannah Vetterly, Emily Veltri, Rachel Weis, Taylor Maida
South Fayette Group 2

Introduction

Ever since the 1970s, obesity has been heavily increasing. Our group wanted to delve into the possible reasons for this spike of obesity. Since an original obvious factor we thought would cause an increase in obesity was fast food restaurants, we included this into our data set. Additionally, because global warming has also been showing a rapid increase lately, we also wanted to incorporate the average temperature in each state into our date set.

Methods

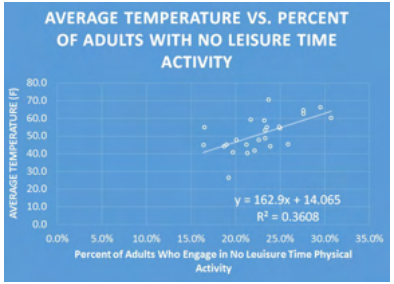
1) Our data was gathered from a variety of of trustworthy sources, such as the Center for Disease Control, the United States Environmental Protection Agency, and Time.

2)We analyzed our data by making a variety of graphs and seeing the correlation between the X and Y axis. We also compared the results of each graph with each other, so we could see how different the factors' data was from one another.

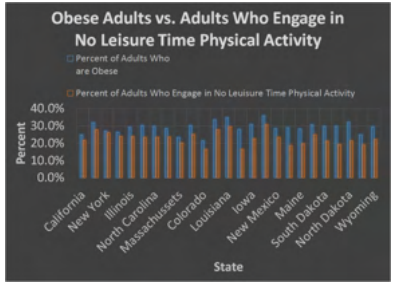
3) Excel helped us visualize our data because we were able to see how each of the different data sets correlated with each other.

Challenges

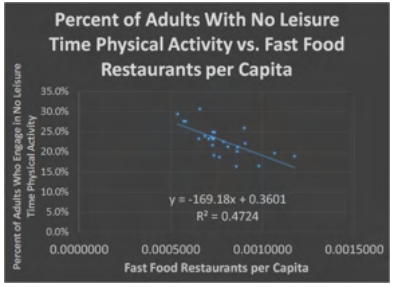
Finding all of the original necessary data- such as the number of gyms per state (the information was not available for free)

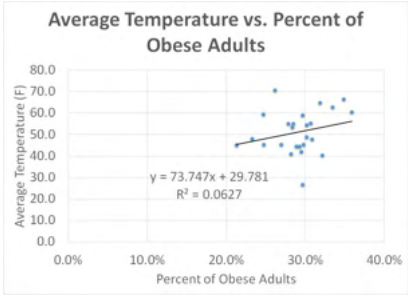


States with high percentages of adults who engage in no leisure time activity also have high percentages of obese adults



The more fast food restaurants in a state, the more adults engaged in physical activity





The higher the temperature, the lower amount of active adults and the higher the obesity rate

Conclusion/Analysis

There is a correlation between more fast food restaurants and higher numbers of adults that engage in physical activity. This graph displayed our highest correlation with an r^2 value of .47. This is potentially because when people eat fast food more often, they could be more motivated to work out. Generally, when people consume fast food they may feel unhealthy, causing them to workout.

A higher average temperature in states also correlate with people engaging in less activity. When the weather is extremely hot, people will most likely want to stay inside with their air conditioning instead of participate in any activity. Thus, the higher the temperature, the percent of the obese adults increases. The r^2 value of the data set was about .36 so it was not as strong as the correlation between fast food and amounts of activity. Therefore, the r^2 between temperature and adults who do not engage in any physical activity leisure time means there is only a slight correlation in which temperature impacts exercise.

sources

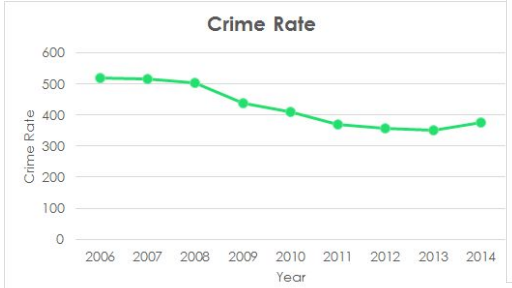
- Time
- CDC
- Current results weather and Science Facts
- Business Insider
- THe Internet Library
- Tableau Public
- The United States Environmental Protection Agency

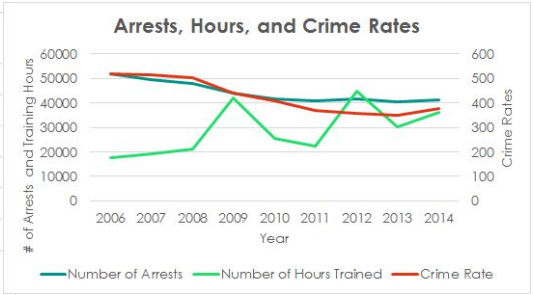
Impact Number of Training Hours has on Arrests

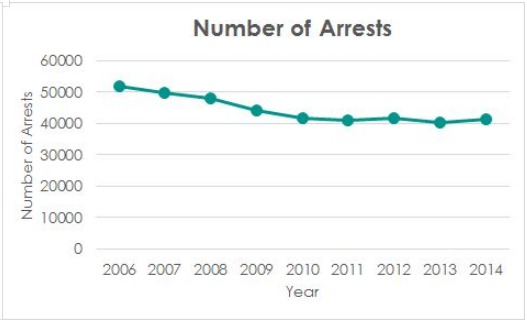
Research Question: Does the number of hours trained by police officers impact the number of arrests made? Does the number of hours trained by officers depend on the crime rates?

Introduction: Police officers train every year to be prepared for what happens on the streets. There is no method to the number of hours trained, however. The numbers are not based even off of crime rates.

Challenges: One of the major challenges was with the number of arrests data. There were over 500,000 arrests spanning 2006 to 2014. Getting through all of it was the biggest challenge





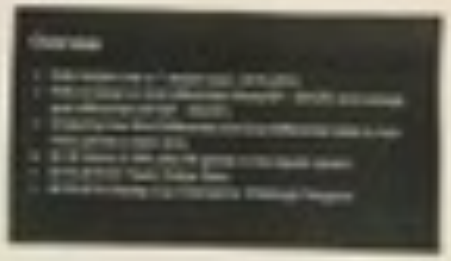


Conclusion: The number of arrests from 2006 to 2014 decreased. The number of hours trained had absolutely no connection to the number of arrests. However, except for a minor difference in 2007 and 2008, the crime rate almost exactly follows the number of arrests. The number of hours trained may have an effect on an officer's performance in the field, but it ultimately comes down to how much crime is actually happening.

Bhavana Kolla, Aubrey Lutz, Anvitha Ramagiri, Manisha Manivannan, Rachel Dorfner

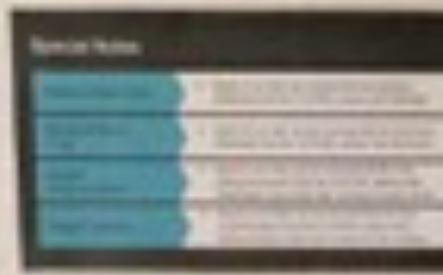
Sources

- WPRDC
- https://www.neighborhoodscout.com/pa/pittsburgh/crime/



Nick Tatarn, 11th Grade

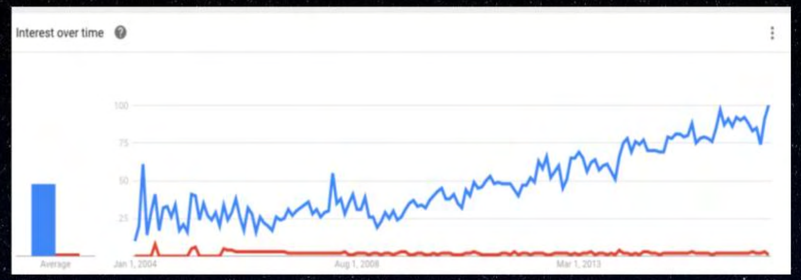
Ethan McFarland, 12th Grade



Does the Number of Insomniacs Correlate with Income Earnings?

Background and Purpose

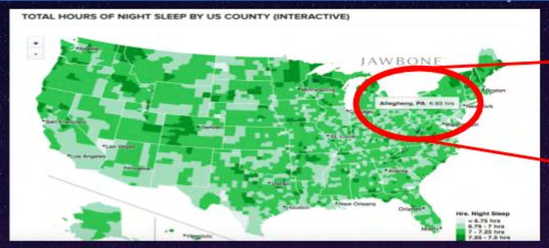
To determine the relationship between the number of insomniacs and income earnings



This graph shows the trend in search interest levels regarding "how to increase salary" and "how to prevent insomnia." **People value income more than sleep.**

According to the American Sleep Association, "Insomnia can be mild to severe depending on how often it occurs and for how long. Chronic insomnia means having symptoms at least 3 nights per week for more than a month. Insomnia that lasts for less time is known as short-term or acute insomnia."

Data



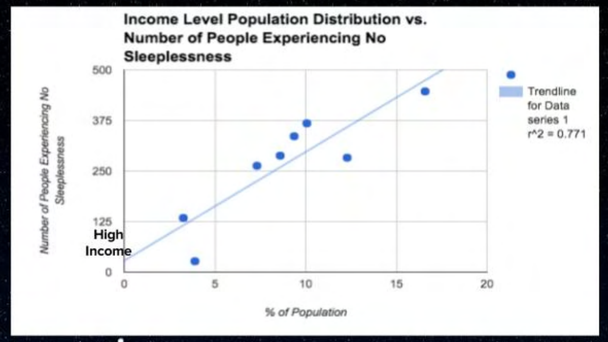
% of Population	Less than $15,000 (c	$15,000 to $24,9	$25,000 to $34,9	$35,000 to $49,9	$50,000 to $74,9	$75,000 to $99,9	$100,000 to $14	$150,000 to $199,999	$250,000 or more
0 nights	20.31	7.32	12.29	10.07	16.59	9.37	8.6	3.26	3.9
1-2 nights	256	263	283	368	447	336	288	134	27
3 or more night	252	241	274	380	457	267	219	89	17
	207	178	182	196	210	145	83	46	17

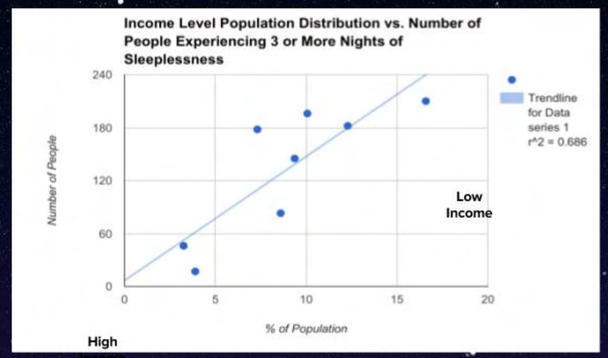
County	Average Hours of Sleep	Average Weekly Wage	Population
Allegheny	6.93	1128	1231255
Beaver	6.96	798	169392
Butler	7	902	185943
Westmoreland	7.02	791	359320
Washington	7.05	1066	208187
Lawrence	6.87	699	88771
Indiana	6.86	796	87706
Jefferson	6.86	688	44638
Cambria	6.98	790	137732
Centre	7.08	876	158742
Clinton	6.99	753	39745
Potter	6.97	749	17206
Huntingdon	7.06	681	45750
Fulton	7	763	14632
Bedford	6.96	627	48946
Somerset	7.14	830	76218
Fayette	6.98	764	134086
Greene	7	800	37843

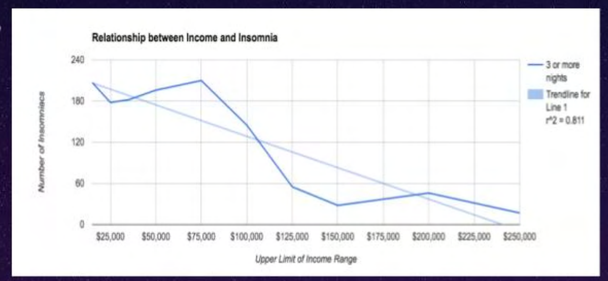
Data on the % of the population suffering from 0, 1-2, 3+ night of sleeplessness and income. Data source: sleepanddreamdatabase.org

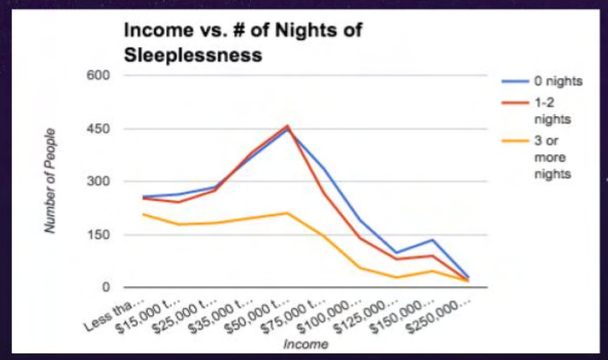
This data shows hours of sleep per county in America and night of insomnia on a national level compared to income.

Data and Results

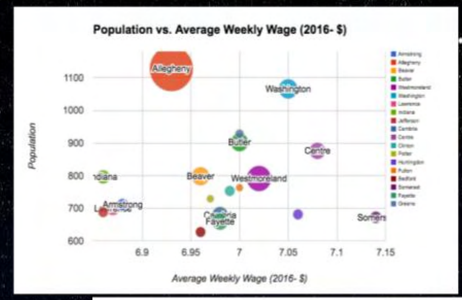


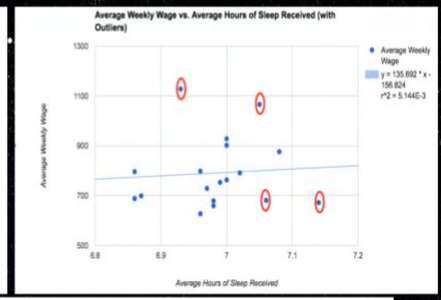


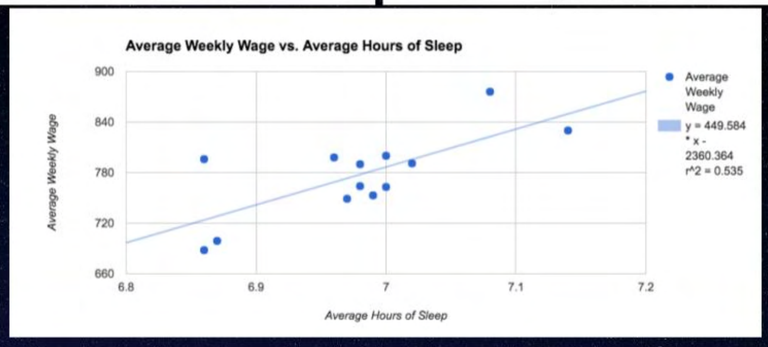




Analysis and Conclusion







Correlation:

- The national data shows a strong, negative correlation between the number of insomniacs and income levels. The correlation coefficient (r) is -.901 and the coefficient of determination (r^2) is .811
- Western PA is an example of Simpson's Paradox: The nation as a whole has a strong, negative correlation between the number of insomniacs and wage levels. In Western PA, there is a moderately strong, positive correlation

Challenges in gathering data:

- Data may not be recent.
- The collection of sleep data is usually biased since voluntary studies typically observe patients concerned about their sleep health.
- People may be reluctant to admit their true salary.
- Lurking variable: stress in a person's job on sleep levels.
  - Wage (fixed) vs. income (variable)

Conclusion/Discussion:

- We observed negative correlation at a national level and positive correlation at a local level, which means that those living in Western PA are more likely to have insomnia if they have a higher income.

Recommendations:

- Business owners should recognize how much sleep affects quality of work.
- The Sleep Center of Greater Pittsburgh offers treatment plans for insomnia.

Upper St Clair Team

Aditi Chattopadhyay, Brooke Christiansen, Mahima Reddy, Kriti Shah, Sanath Boddula, Yash Lahoti