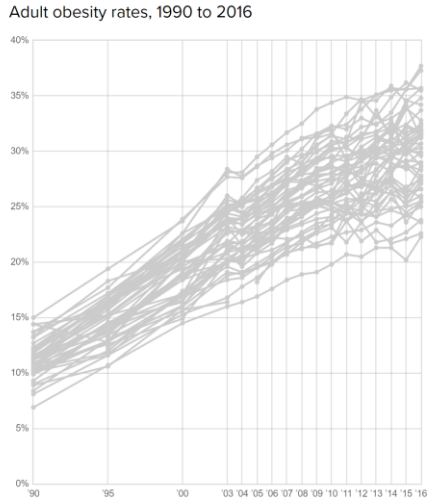
The Correlation Between Internet Speeds and Obesity Rates

Akshath Jain

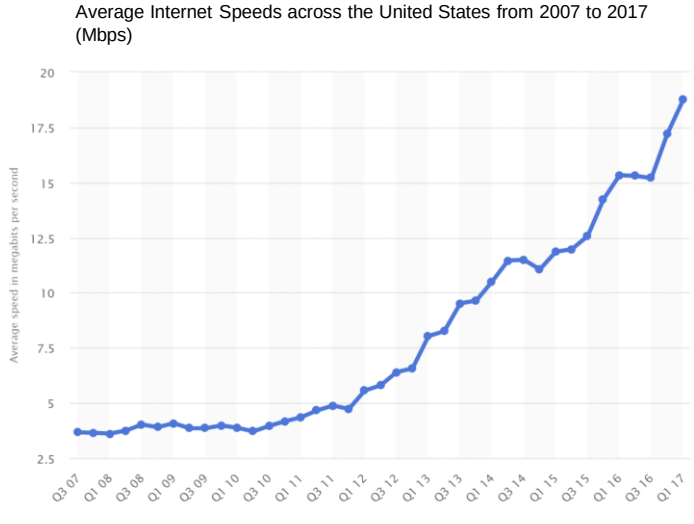
North Allegheny

Introduction

- 36.5% of American adults are obese (CDC)
- Internet speeds and usage among Americans is on the rise
- In this project, we explored any possible correlation between Internet speeds and obesity rates in Pittsburgh and in the US



[src: State of Obesity]



[src: Statista]

Hypothesis

Null Hypothesis (H_0):

There is no correlation between obesity and internet speeds.

$$\rho_{XY} = 0$$

Alternative Hypothesis (H_a):

There is a significant correlation between obesity and internet speeds.

$$\rho_{XY} \neq 0$$

Extraneous Variables

Confounding Variables

Incorporated into the design of the study

- Income

Lurking Variables

Variables not incorporated into the design of the study

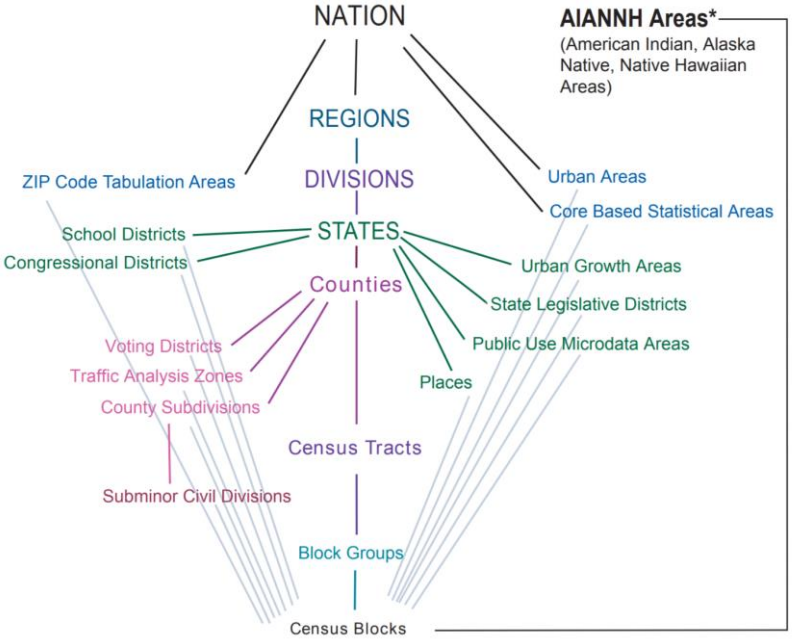
- Age
- Education
- Technical Literacy

Pittsburgh

Data Sources - Pittsburgh

	Name	Source	Purpose
Pittsburgh	Allegheny County Obesity Rates	Western PA Regional Data Center	Provides obesity rates as a percentage of obese in a specific Census block group.
	Pittsburgh ISPs by Block	FCC / Western PA Regional Data Center	Provides maximum advertised consumer and business upload and download speeds by Census block.
	2017 TIGER / Line Shapefiles	US Census Bureau	Matches PA block codes to shapefile metadata. Used for heat map construction.
	PGH Snap Census Data—Education and Income	Western PA Regional Data Center	Provides median income for Pittsburgh Neighborhoods in 2010.
	PGH Snap Census—Housing	Data.gov	Provides number of occupied housing units per Pittsburgh neighborhood. Used for income normalization.
	PGH Snap Census—Neighborhood Census Data	Data.gov	Provides population per Pittsburgh neighborhood. Used for income normalization.

Census Breakdown

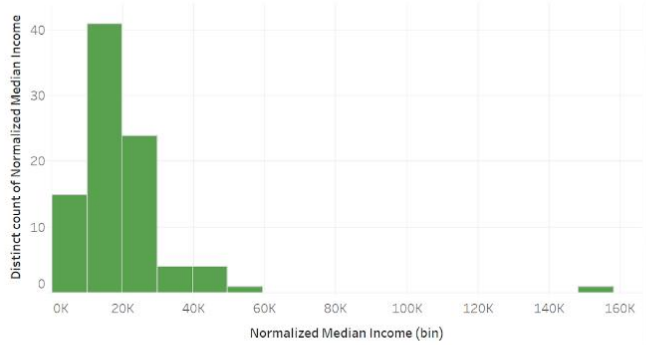


[src: US Census Bureau]

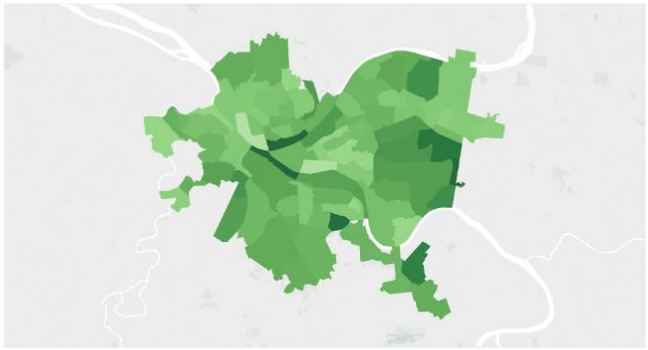
Pittsburgh Census Blocks
>30,000 ft^2

Pittsburgh - Income Distribution

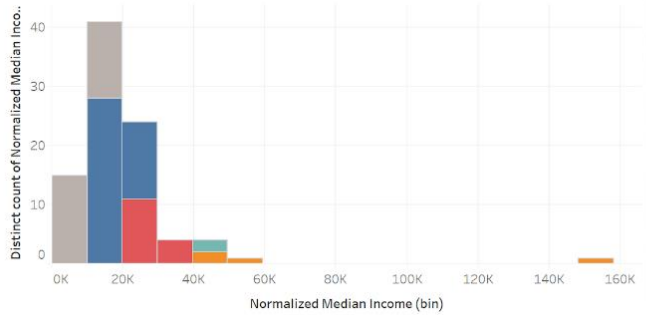
Income Distribution for Pittsburgh Neighborhoods



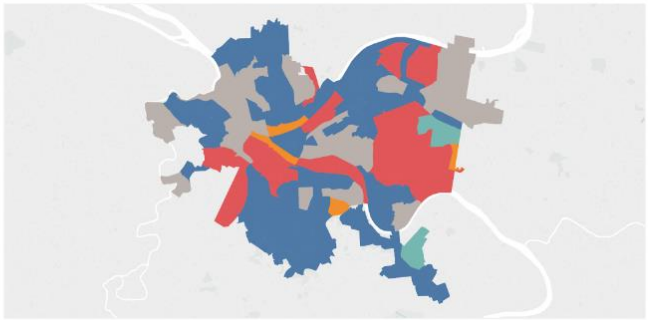
Income Heat Map for Pittsburgh Neighborhoods



Income Distribution for Pittsburgh Neighborhoods (with Income Grouping)



Income Heat Map for Pittsburgh Neighborhoods (with Income Grouping)



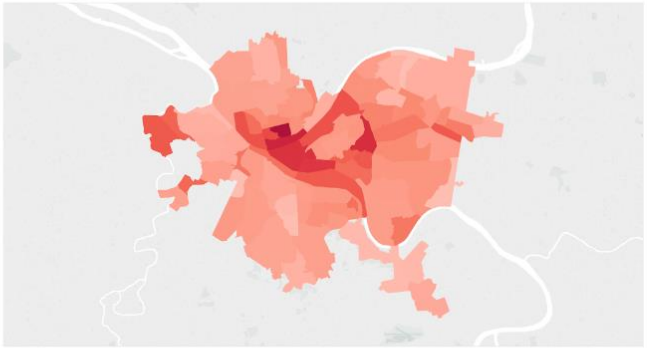
Data Sets

1. PGH Snap Census Data—Education and Income
2. PGH Snap Census—Housing
3. PGH Snap Census—Neighborhood Census Data
4. 2017 TIGER / Line Shapefiles

Pittsburgh - Internet Speeds

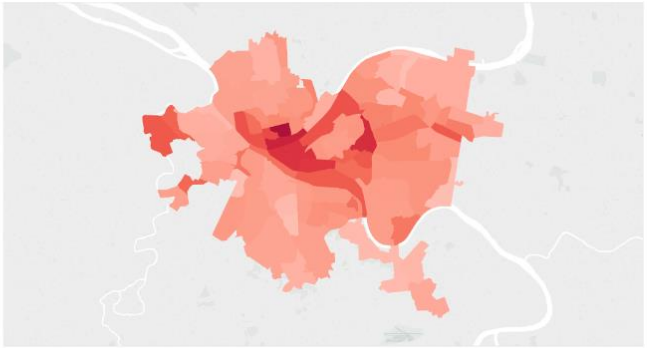
Consumer Upload Speeds (Mbps) Heat Map

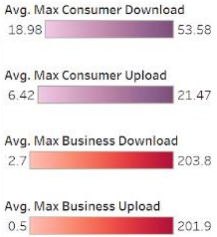
Business Upload Speeds (Mbps) Heat Map



Consumer Download Speeds (Mbps) Heat Map

Business Download Speeds (Mbps) Heat Map

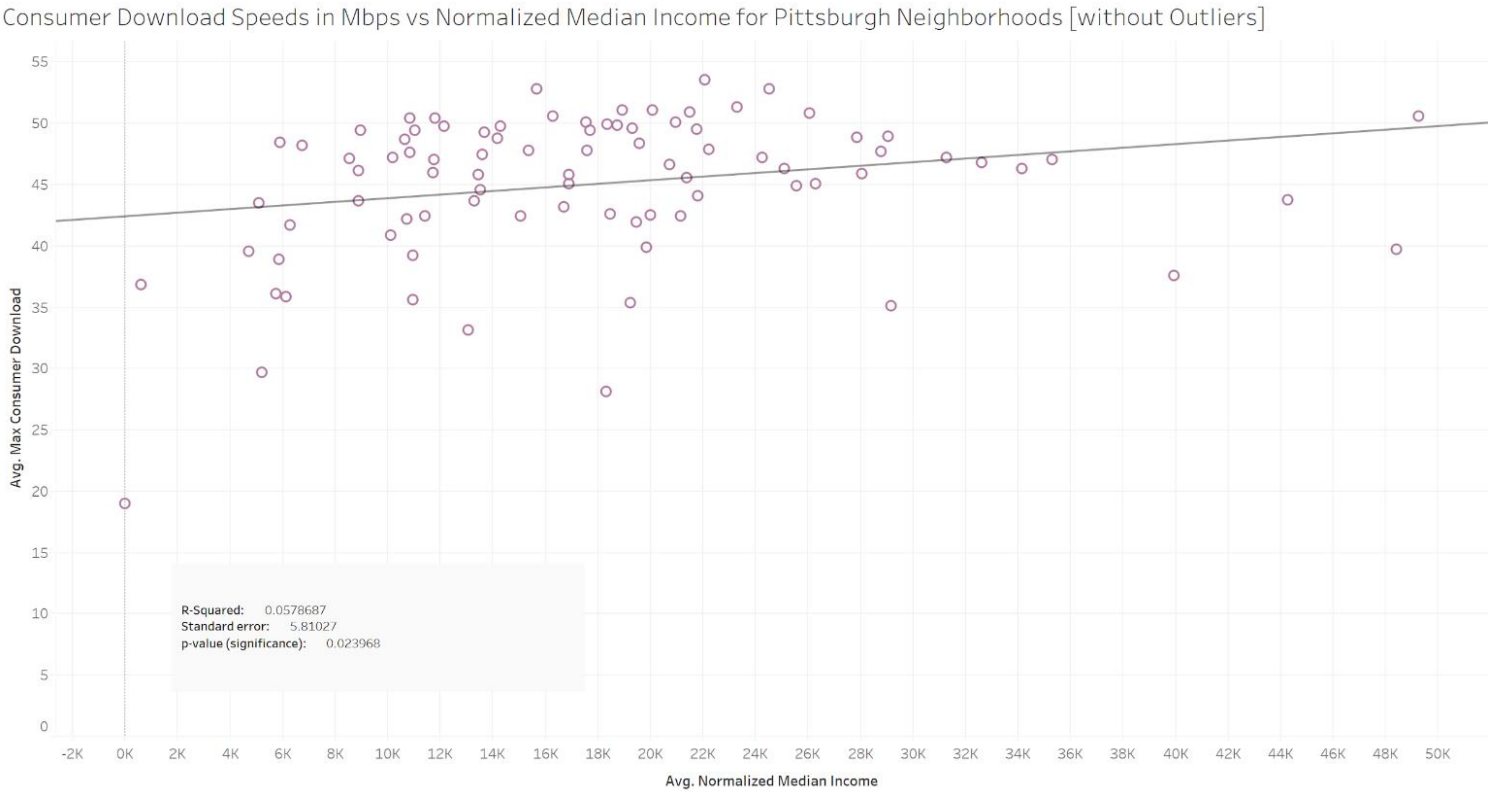




Data Sets

1. Pittsburgh ISPs by Block
2. 2017 TIGER / Line Shapefiles

Pittsburgh - Consumer Download Speeds vs Income

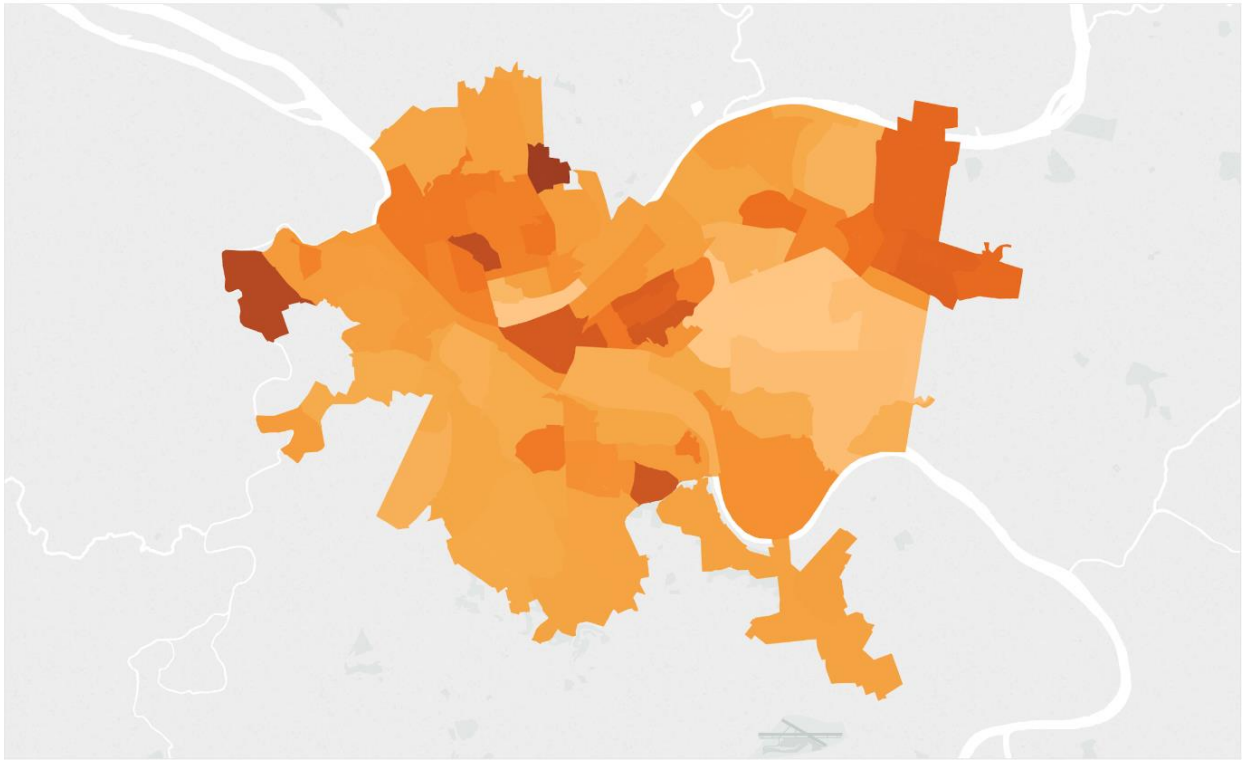


Data Sets

1. Pittsburgh ISPs by Block
2. PGH Snap Census Data—Education and Income

Pittsburgh - Obesity

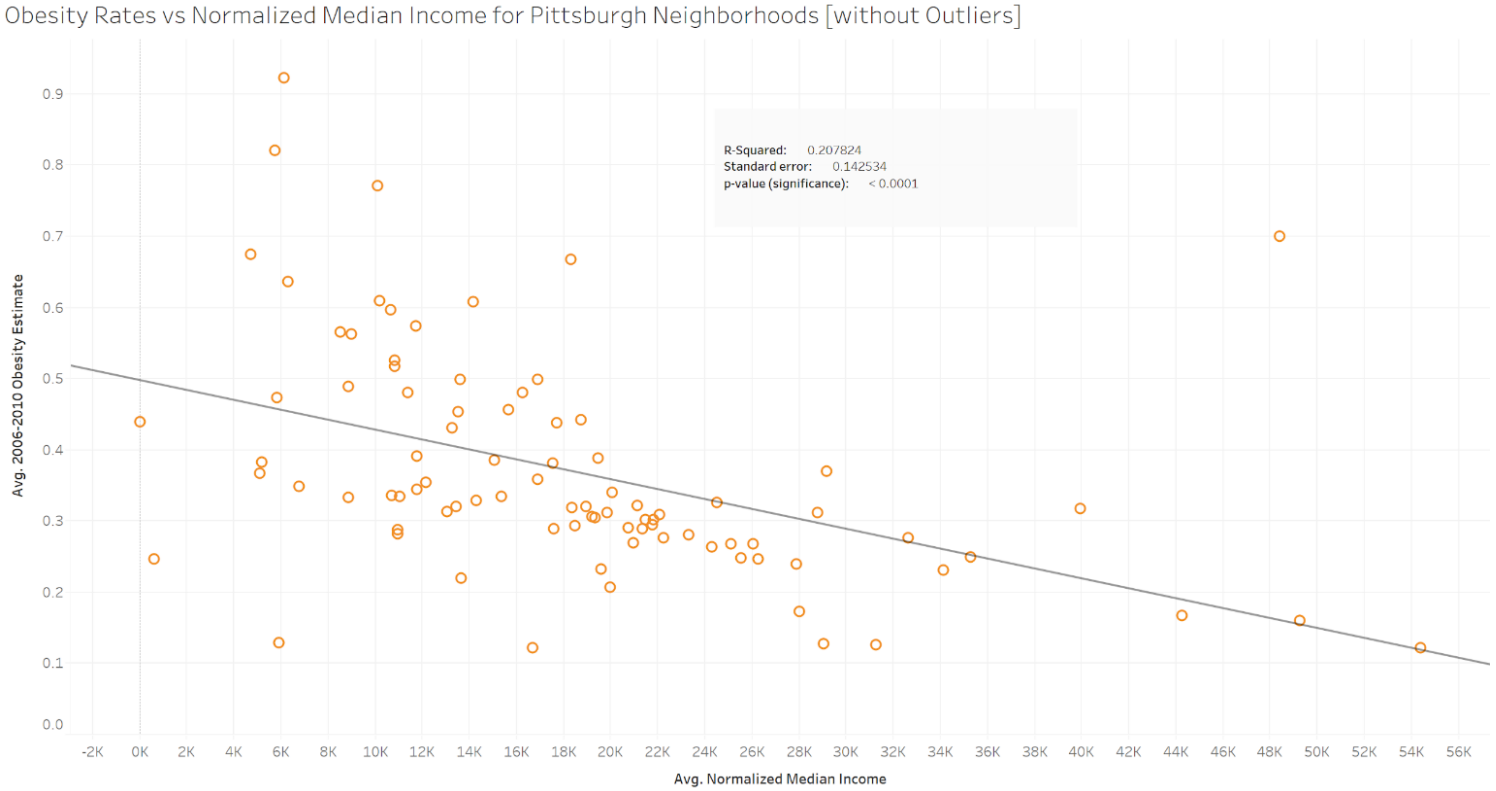
Pittsburgh Neighborhoods Obesity Heat Map



Data Sets

1. Allegheny County Obesity Rates
2. 2017 TIGER / Line Shapefiles

Pittsburgh - Obesity Rates vs Income

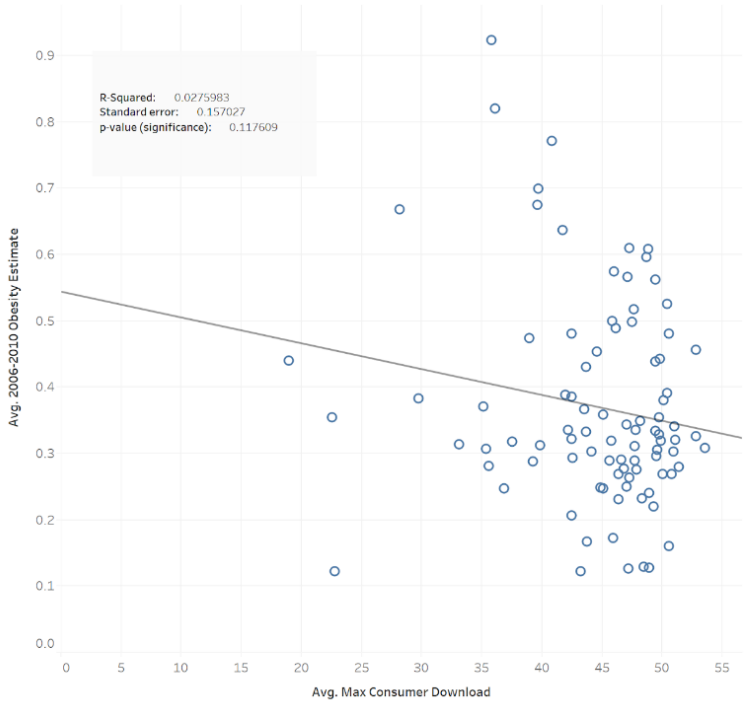


Data Sets

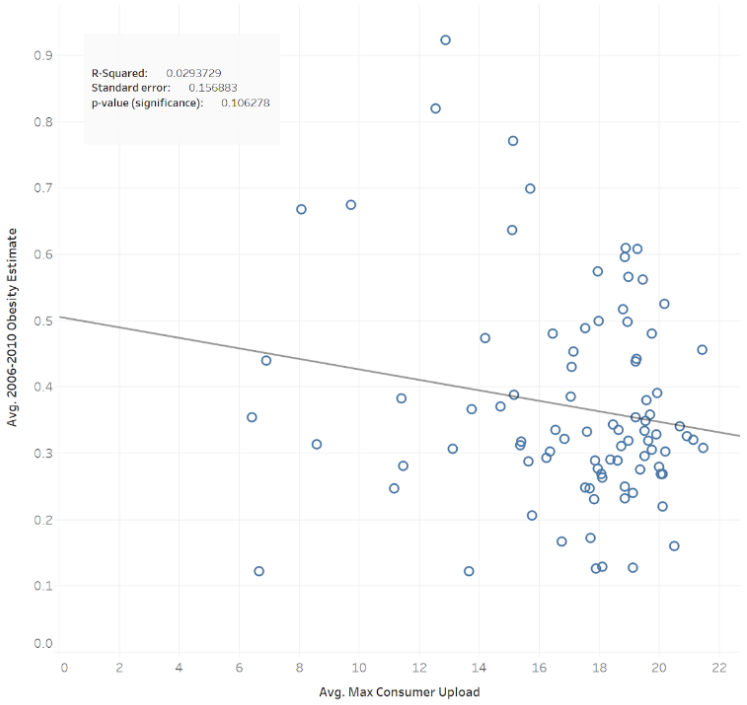
1. Allegheny County Obesity Rates
2. PGH Snap Census Data—Education and Income

Pittsburgh - Obesity vs Consumer Internet Speeds

Obesity vs Max Consumer Download Speeds (Mbps) for Pittsburgh Neighborhoods



Obesity vs Max Consumer Upload Speeds (Mbps) for Pittsburgh Neighborhoods



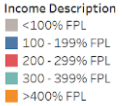
Data Sets

1. Allegheny County Obesity Rates
2. Pittsburgh ISPs by Block

Pittsburgh - Obesity vs Consumer Internet Speeds [with Income Grouping]

Obesity vs Max Consumer Download Speeds (Mbps) for Pittsburgh Neighborhoods (with Income Grouping)

Obesity vs Max Consumer Upload Speeds (Mbps) for Pittsburgh Neighborhoods (with Income Grouping)

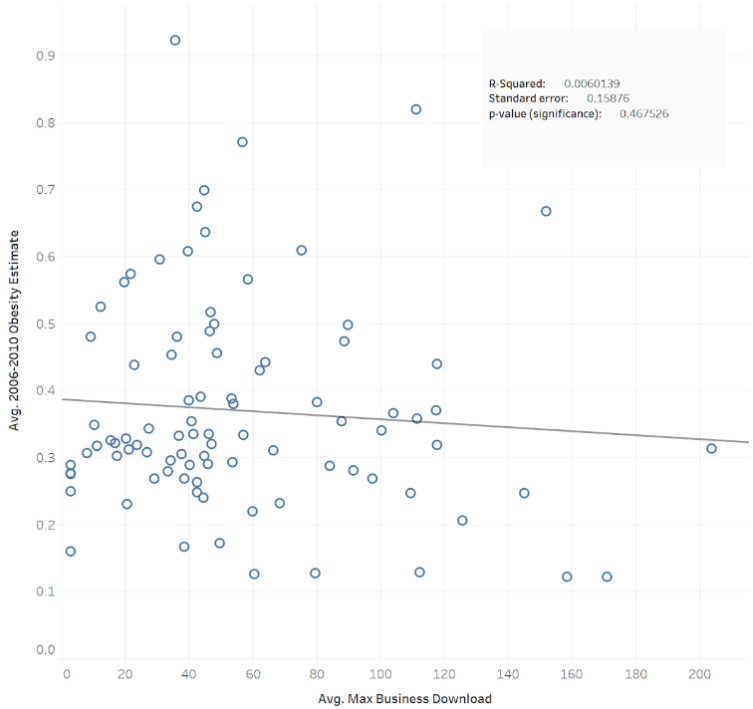


Data Sets

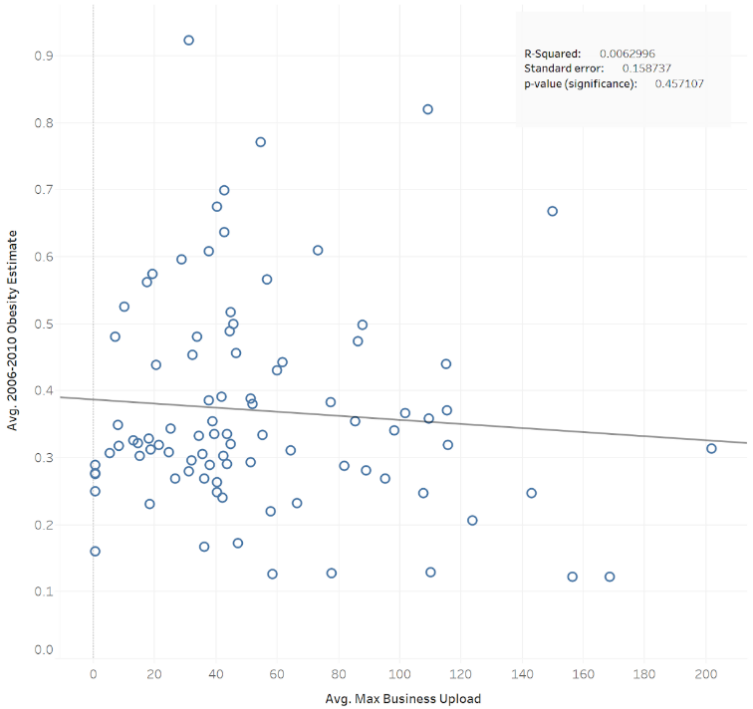
1. Allegheny County Obesity Rates
2. Pittsburgh ISPs by Block
3. PGH Snap Census Data—Education and Income
4. PGH Snap Census—Housing
5. PGH Snap Census—Neighborhood Census Data

Pittsburgh - Obesity vs Business Upload/Download Speeds

Obesity vs Max Business Download Speeds (Mbps) for Pittsburgh Neighborhoods



Obesity vs Max Business Upload Speeds (Mbps) for Pittsburgh Neighborhoods

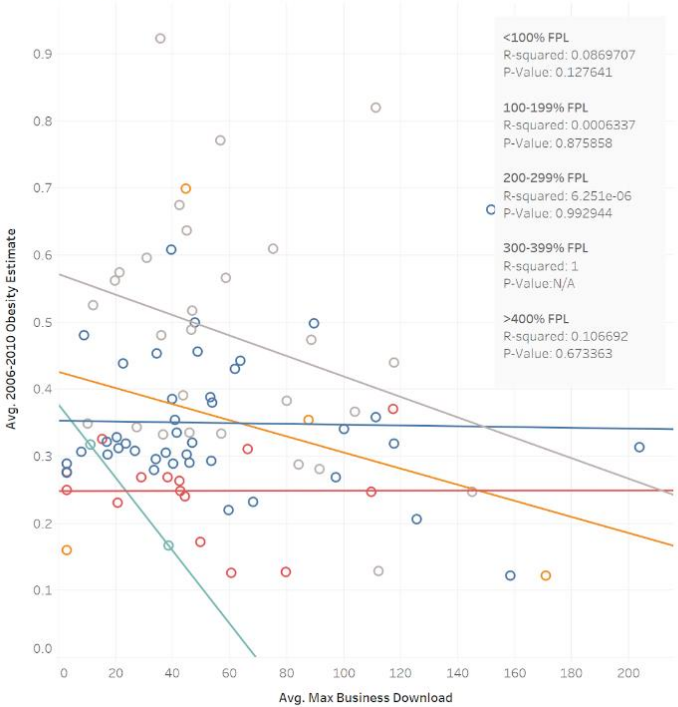


Data Sets

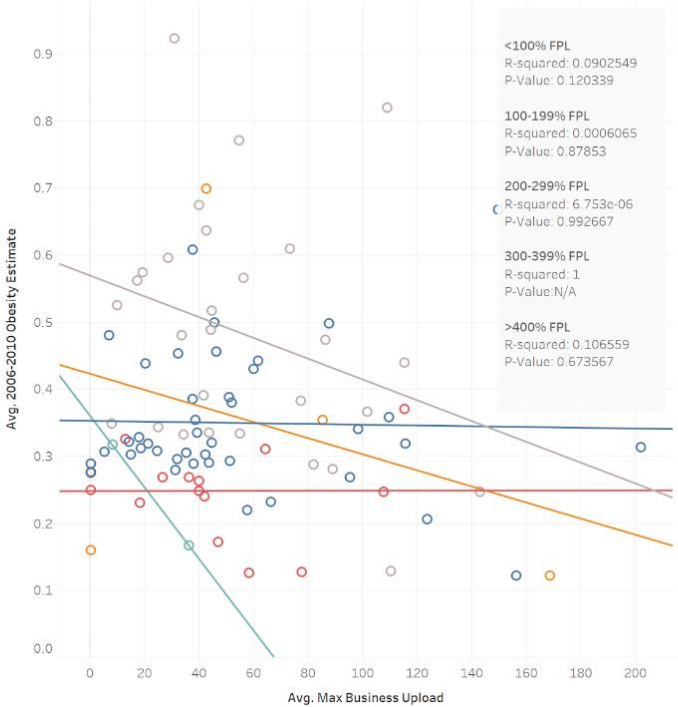
1. Allegheny County Obesity Rates
2. Pittsburgh ISPs by Block

Pittsburgh - Obesity vs Business Speeds [with Income Grouping]

Obesity vs Max Business Download Speeds (Mbps) for Pittsburgh Neighborhoods (with Income Grouping)



Obesity vs Max Business Upload Speeds (Mbps) for Pittsburgh Neighborhoods (with Income Grouping)



Data Sets

1. Allegheny County Obesity Rates
2. Pittsburgh ISPs by Block
3. PGH Snap Census Data—Education and Income
4. PGH Snap Census—Housing
5. PGH Snap Census—Neighborhood Census Data

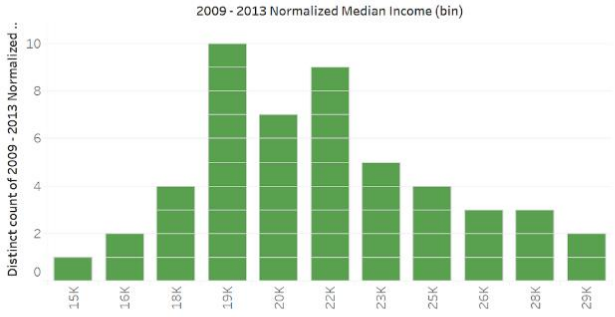
United States

Data Sources - United States

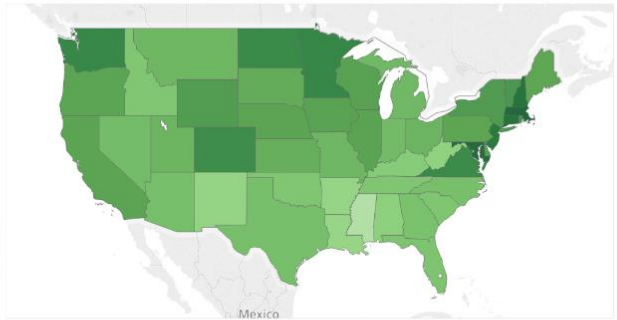
	Name	Source	Purpose
United States	Adult Obesity in the United States	State of Obesity	Provides adult obesity rates as a percentage of state population.
	USA Average Peak Internet Speeds	Fastmetrics	Provides average internet speeds for each state.
	US Median Income—2016	US Census Bureau	Provides median income by state.
	United States - Average household size, 2009-2013 by State	Index Mundi	Provides average household size by state. Used for income normalization.

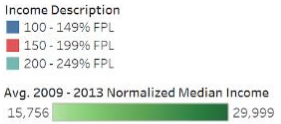
United States - Income Distribution

Distribution of Normalized Median Income per Capita (United States)

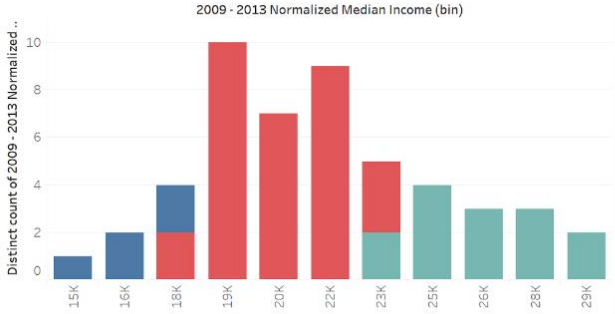


Normalized Median Income per Capita Heat Map (United States)

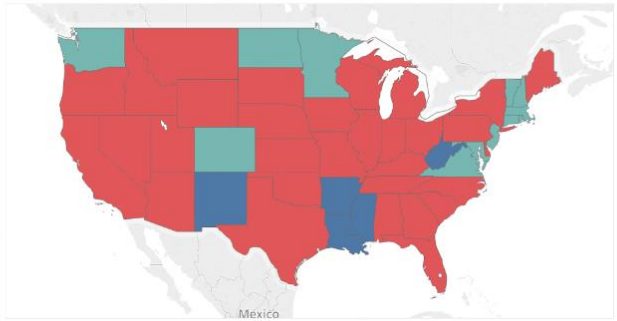




Distribution of Normalized Median Income per Capita (United States) (with Income Grouping)



Normalized Median Income per Capita Heat Map (United States) (with Income Blocking)

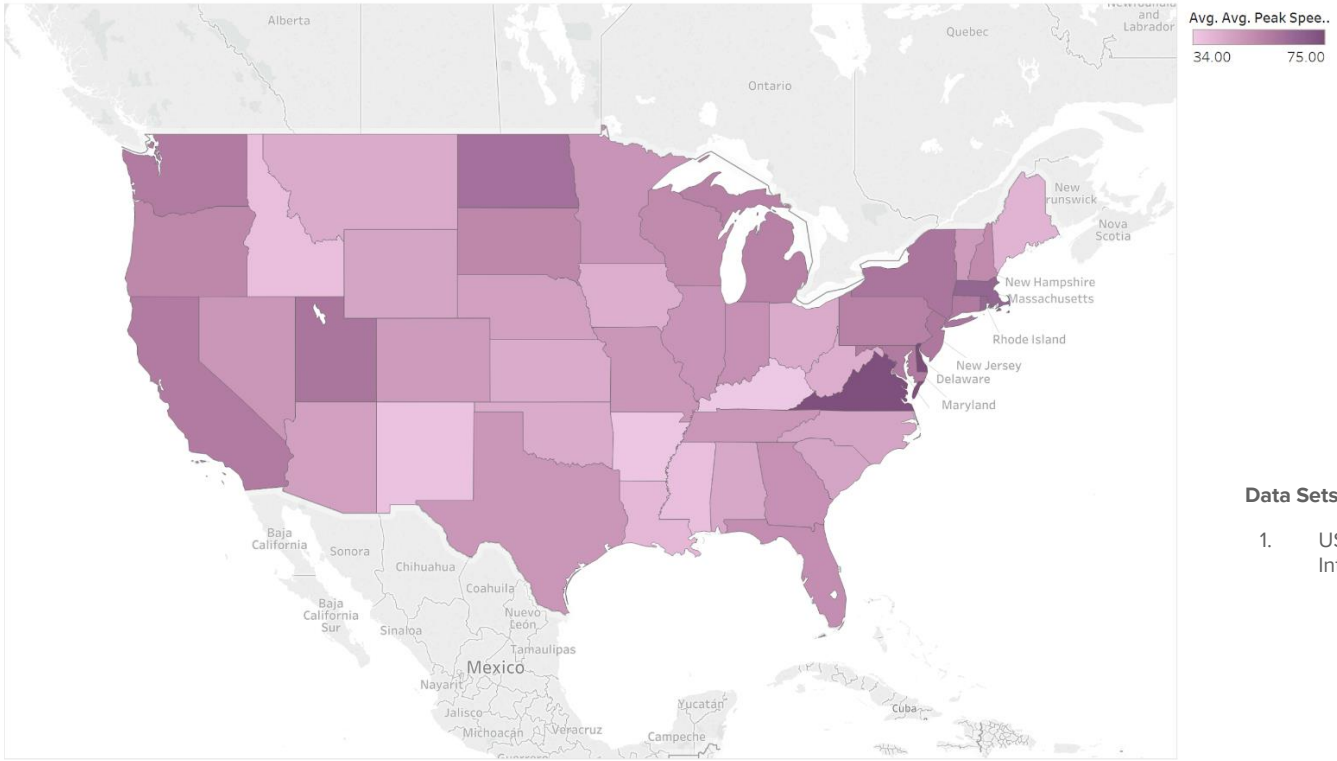


Data Sets

1. United States - Average household size, 2009-2013 by State
2. US Median Income—2016

United States - Internet Speeds

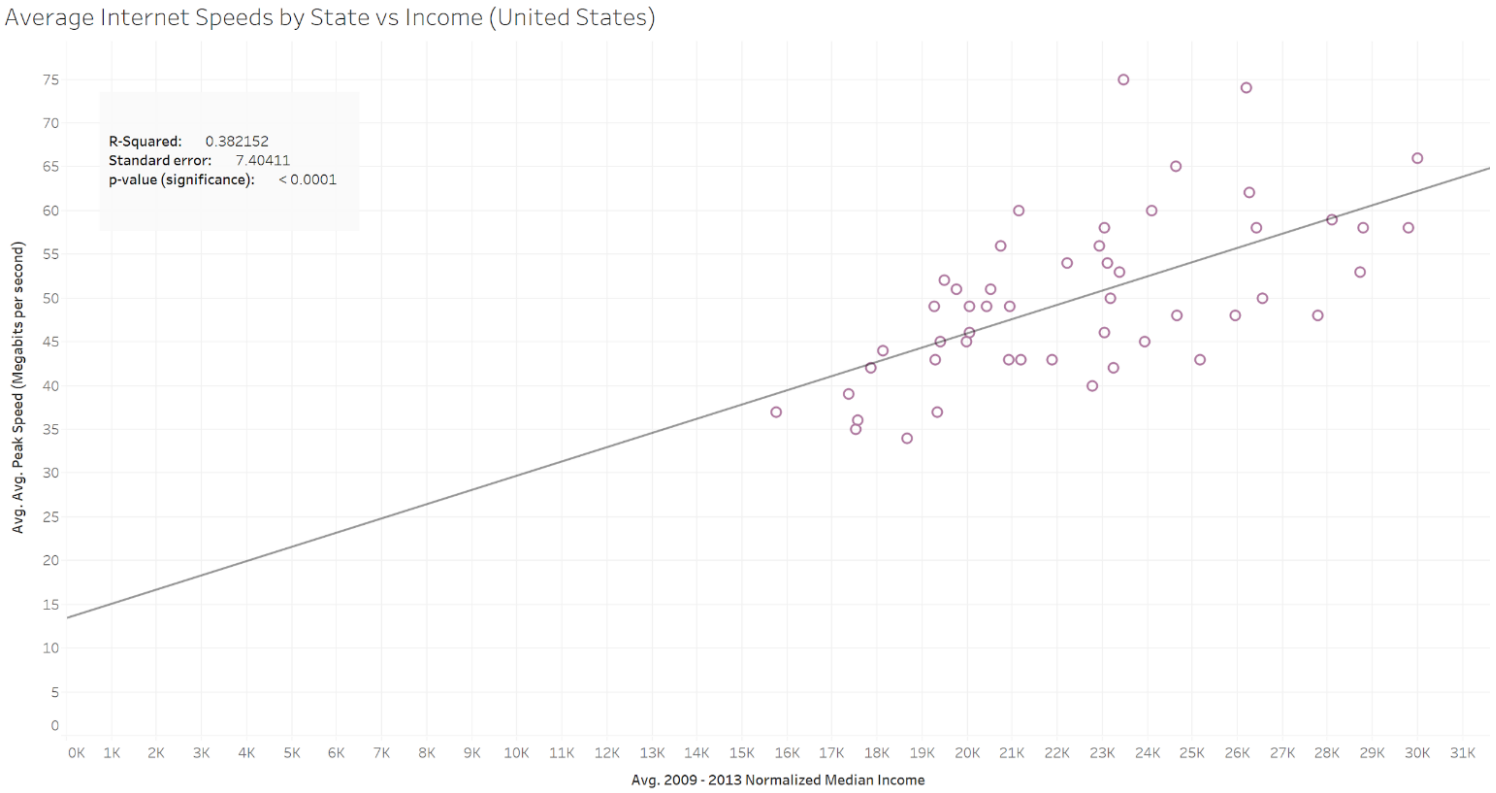
United States Internet Speeds Heat Map



Data Sets

1. USA Average Peak Internet Speeds

United States - Internet Speeds vs Income

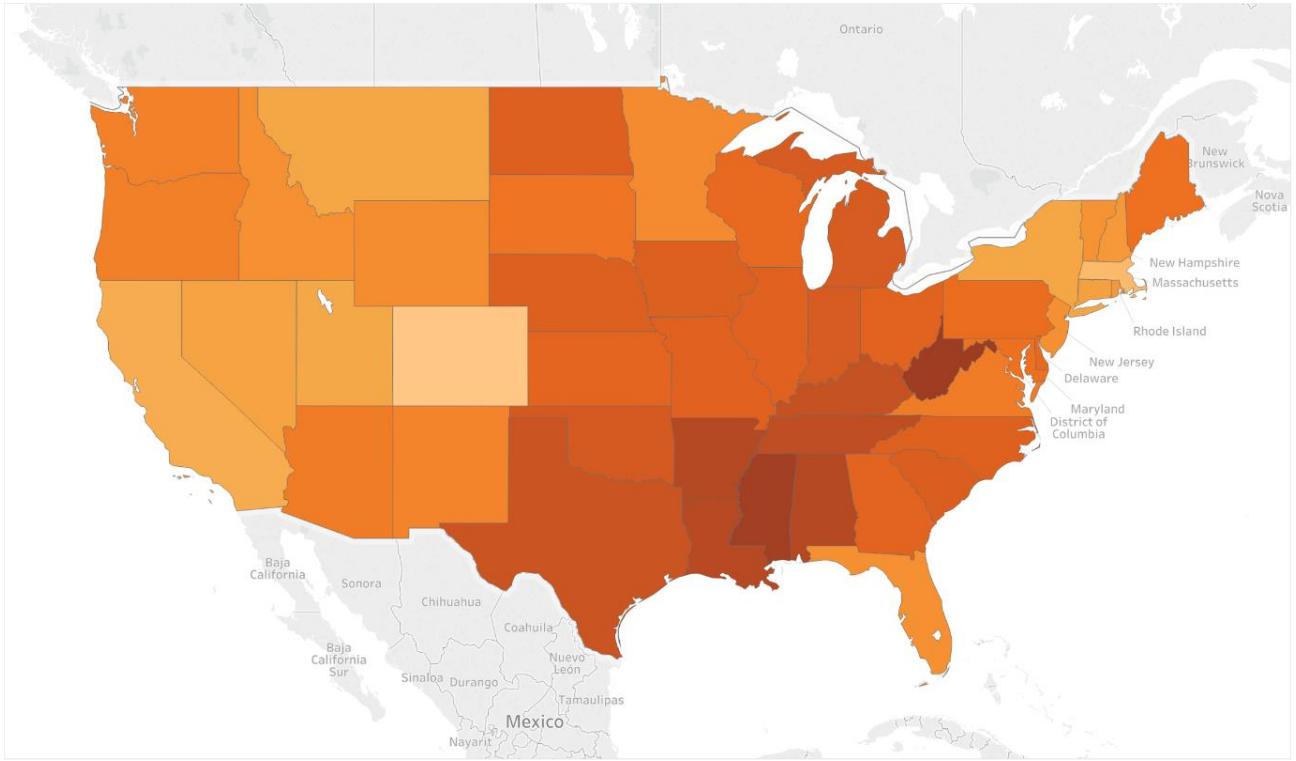


Data Sets

1. USA Average Peak Internet Speeds
2. US Median Income—2016
3. United States - Average household size, 2009-2013 by State

United States - Obesity

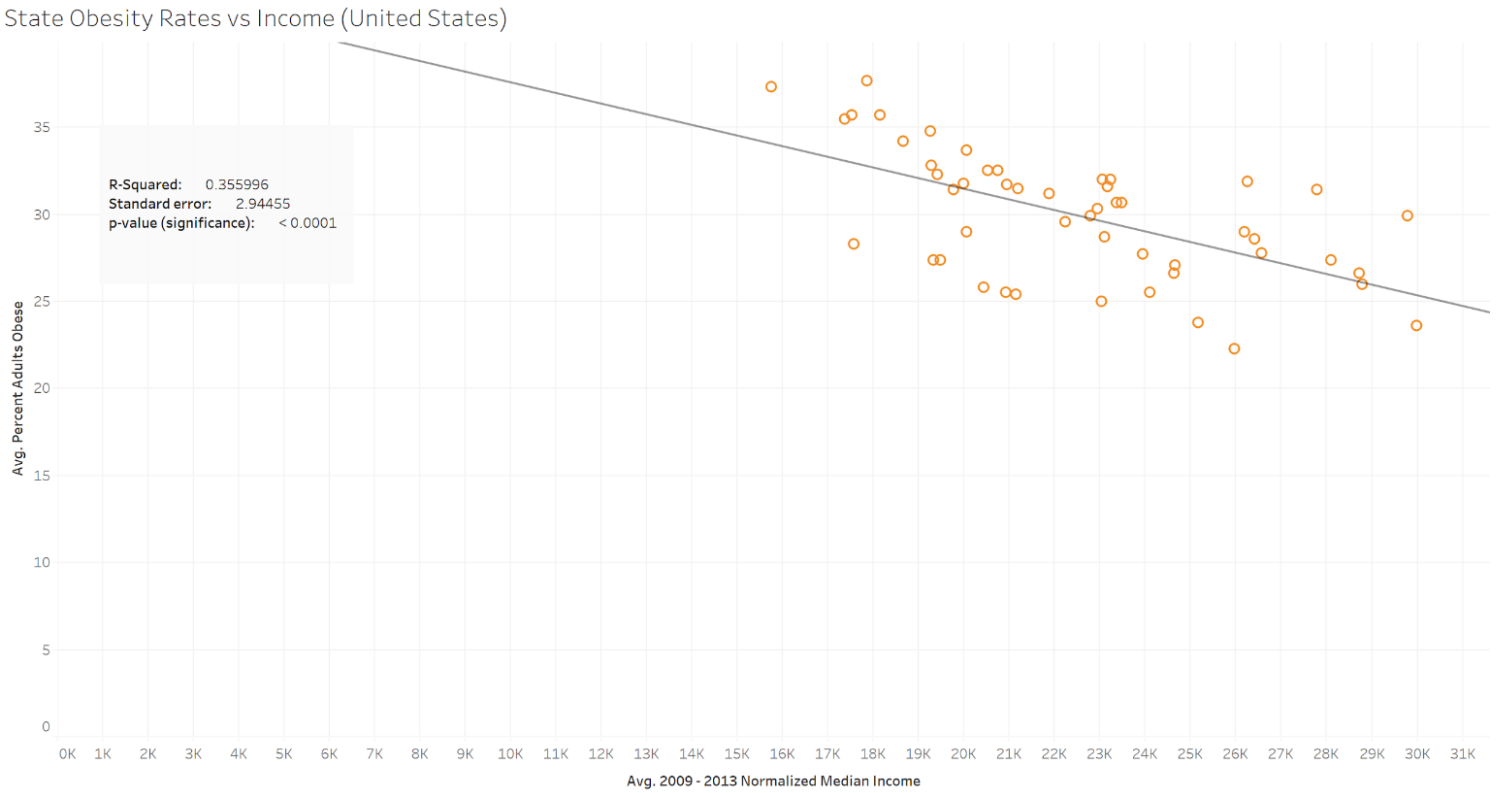
United States Obesity Heat Map



Data Sets

1. Adult Obesity in the United States

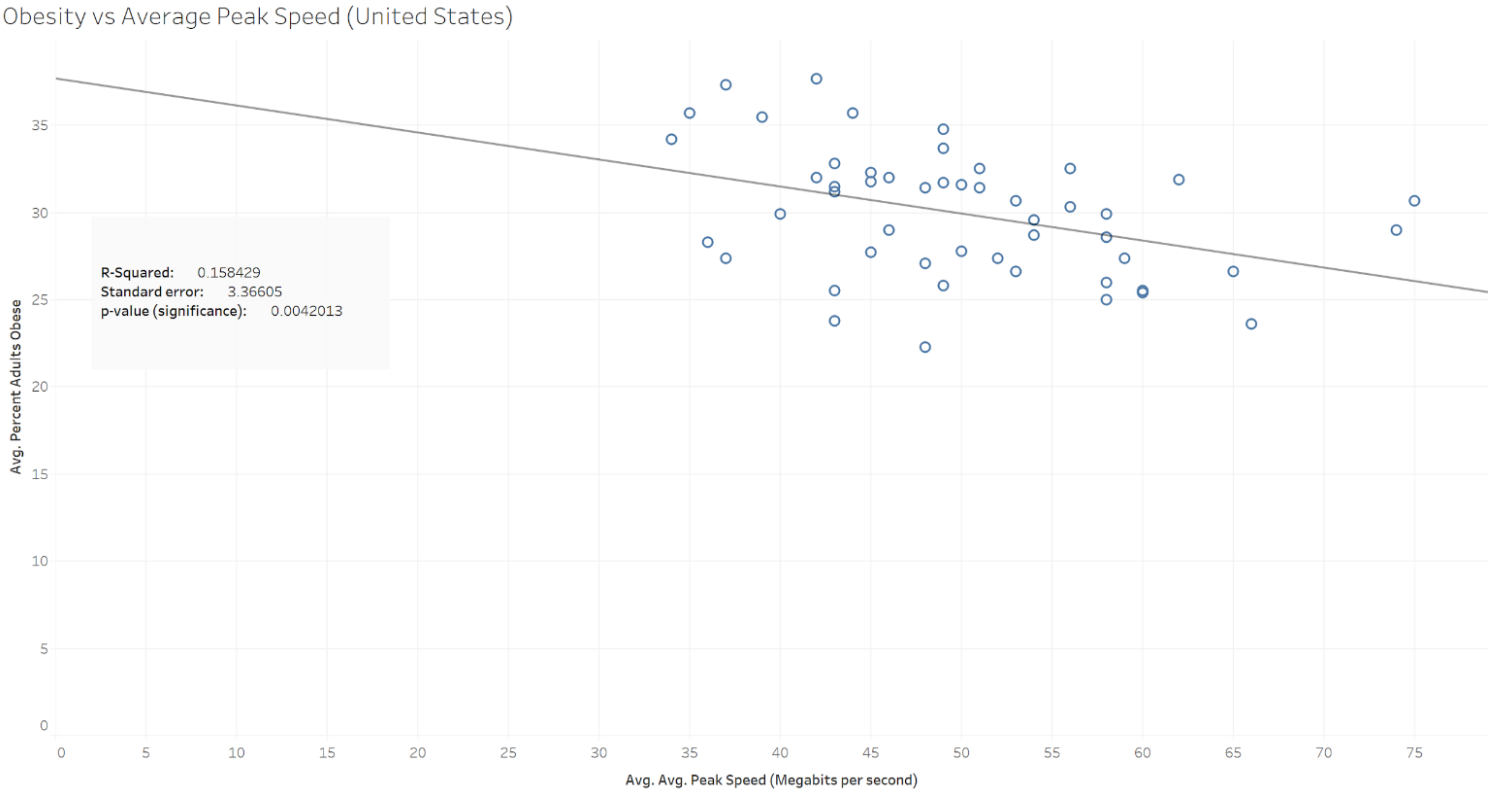
United States - Obesity Rates vs Income



Data Sets

1. Adult Obesity in the United States
2. US Median Income—2016
3. United States - Average household size, 2009-2013 by State

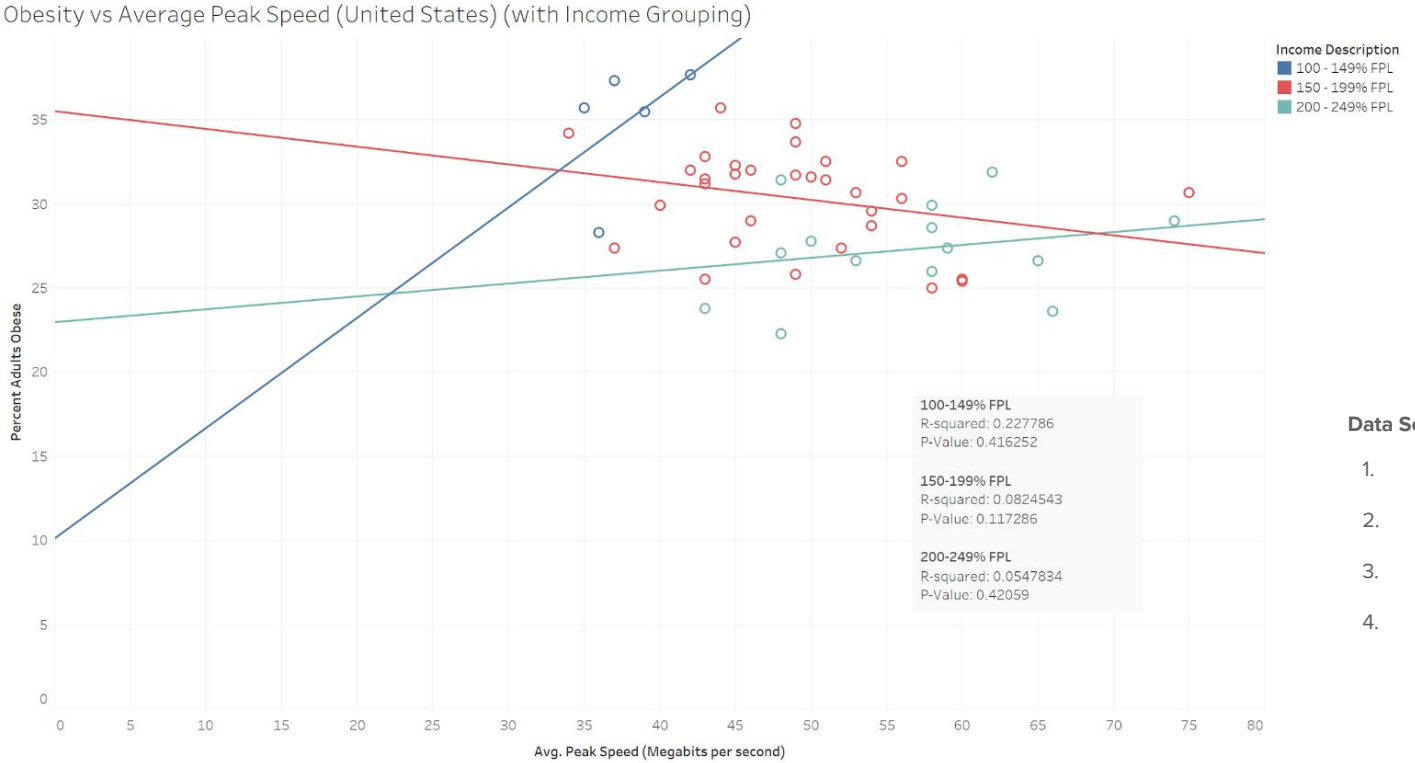
United States - Obesity vs Internet Speeds



Data Sets

1. Adult Obesity in the United States
2. USA Average Peak Internet Speeds

United States - Obesity vs Internet Speeds [with Income Grouping]



Data Sets

1. Adult Obesity in the United States
2. USA Average Peak Internet Speeds
3. US Median Income—2016
4. United States - Average household size, 2009-2013 by State

Analysis

Pittsburgh - Consumer Download

- Weak ($r = -0.1661$) and statistically insignificant ($p = 0.12 > \alpha = 0.05$) correlation in aggregate data
- When divided into distinct income groups:

Income Group	r	p-value	Analysis
<100% FPL	-0.0846	0.6686	Weak, insignificant
100-199% FPL	-0.1625	0.3101	Weak, insignificant
200-299% FPL	-0.2971	0.2822	Weak, insignificant
300-399% FPL*	—	—	—
>400% FPL	0.1246	0.8754	Weak, insignificant

* The 300-399% FPL group only had 2 data points, so no accurate r and p values could be determined for the group.

United States

- Moderate ($r = -0.3980$) yet statistically significant ($p = 0.0042 < \alpha = 0.05$) correlation in aggregate data
- When divided into distinct income groups:

Income Group	r	p-value	Analysis
100-149% FPL	0.4773	0.4163	Moderate, insignificant
150-199% FPL	-0.2871	0.1173	Weak, insignificant
200-249% FPL	0.2341	0.4206	Weak, insignificant

Discussion

Similar Studies

- Central Queensland University, 2009:
  - Participants with high leisure-time Internet and computer usage are 1.46 times more likely to be overweight than participants with less leisure time Internet usage
- Milken Institute, 2012:
  - Each 10% increase in national spending on technology is correlated with a 1% increase in that nation’s obesity rate

Possible Sources of Error

- Pittsburgh Data
  - Internet speeds data from internet providers (max advertised download/upload speeds provided for an area)
    - May not be representative of actual internet speeds experienced by consumers
  - Obesity rate data compiled from demographically similar neighborhoods from around nation
    - May not be representative
    - Assumed to be negligible
- US Data
  - Internet speed tests were conducted voluntarily (introduced voluntary response bias)

Conclusion

Conclusion

- Cannot reject the null: p-values are not significant at any reasonable α
- No correlation between internet speeds and obesity rates
- Significant correlation between:
  - Obesity and Income
  - Internet Speed and Income

Recommendations

- Internet Service Providers (ISPs) should reduce prices
  - Increase high speed Internet access in lower-income neighborhoods
- Further research in area to find more definitive conclusions
  - Understand when Internet access benefits public health and when it harms it

Challenges

- Finding credible and representative datasets for both the local and national levels
- Determining how to judge a household’s income
- Deciding cutoffs for income grouping
- Data visualization using Tableau
- Dealing with several 500,000+ line files

Future Studies

- Comparing internet speeds over time for each neighborhood / state
- Comparing obesity rates over time for each neighborhood / state
- Analyzing internet speeds vs. internet usage

Acknowledgements

- Mrs. Cris Ruffolo, *North Allegheny Senior High School*
- Dr. Raja Sooriamurthi, *Carnegie Mellon University*
- Pitt Data Science Students, *University of Pittsburgh*