Project 1 Report
Determining Correlations Between United States' Health and Income Data

ECEN 689: Applied Information Science Practicum - Team 7

Course Instructor: Prof. Jean-Francois Chamberland

Team 7 Members

1) Khaled Nakhleh
2) Shabarish Prasad
3) Fazia Batool
4) Venkata Pydimarri

Due Date: October 30th, 2018

Abstract

The goal of this project is to support or refute the claim, ”Income level is positively correlated to health”. We are given two well-being indicators namely, diabetes and obesity rates for the years 2008 and 2013. Income data was gathered for the year 2013, since the 2008 dataset was offered per state, and not for the entire United States.

Data Sets

We used two datasets in this project.

The first dataset is “**Food Environment Atlas”.** Food environment factors interact to influence food choices and diet quality. The objectives of the Food Atlas, made available by the U.S. Department of Agriculture, are to assemble statistics on food environment indicators and to provide a spatial overview of access to healthy food. The atlas contains health and well-being indicators such as diabetes and obesity rates.
The current version of the Food Environment Atlas has over 275 variables, including new indicators on access and proximity to a grocery store for sub populations; an indicator on the SNAP Combined Application Project for recipients of Supplemental Security Income (at the State level); and indicators on farmers' markets that report accepting credit cards or report selling baked and prepared food products. Data was used for the years 2008 and 2013.

The second dataset is “**County Income Tax Statistics”.** The Internal Revenue Service (IRS) is the revenue service of the United States federal government. It is responsible for collecting taxes and administering the Internal Revenue Code. Public information includes individual tax statistics for 2016, grouped by FIPS code.

Introduction

Project 1 for ECEN 689 course involved gathering health and income data for the United States. The project also involved parsing through the data, and ensuring accurate representation of U.S. counties. The project’s objective was to determine if health parameters (obesity, diabetes, etc) were correlated with income data collected by the IRS (Internal Revenue Service).

The project goal was also to better understand the relationship between health measurements between the broader US social classes, and the classes’ income disparity proven in previous academic work [4]. We want to acknowledge other factors that we assume affect overall health. Education level was also seen to affect health quality in US social classes [6]. More discussion is offered in the next sections.

Going through the literature, one can see an established truth about income and health: the higher an individual's income, the better his or her health [1]. However, a broader scope is needed when

considering health and income metrics. Not to mention the perceived health evaluation individuals have about their life. We highlight the importance of rationally choosing variables for measuring health and income levels. Previous work recommends the variables should be dependent on the relevance of the variable to the model being studied [5].

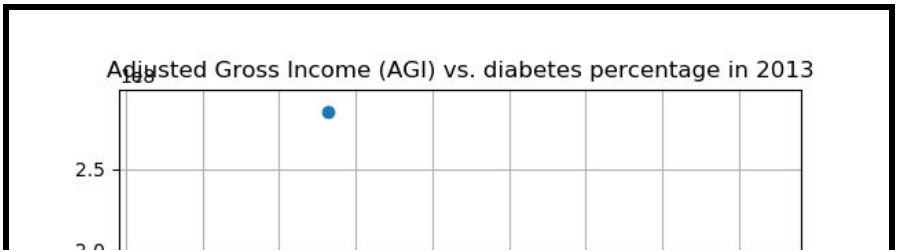
Life evaluation refers to the thoughts that people have about their life when they think about it [2]. While previous academic work did prove a correlation between perceived mental health and income, it also proved no tangible progress beyond an annual income of ~$75,000 [2]. Effects of mental health were not analyzed in this work due to time constraints. We recommend future studies to incorporate mental health evaluation into their models.

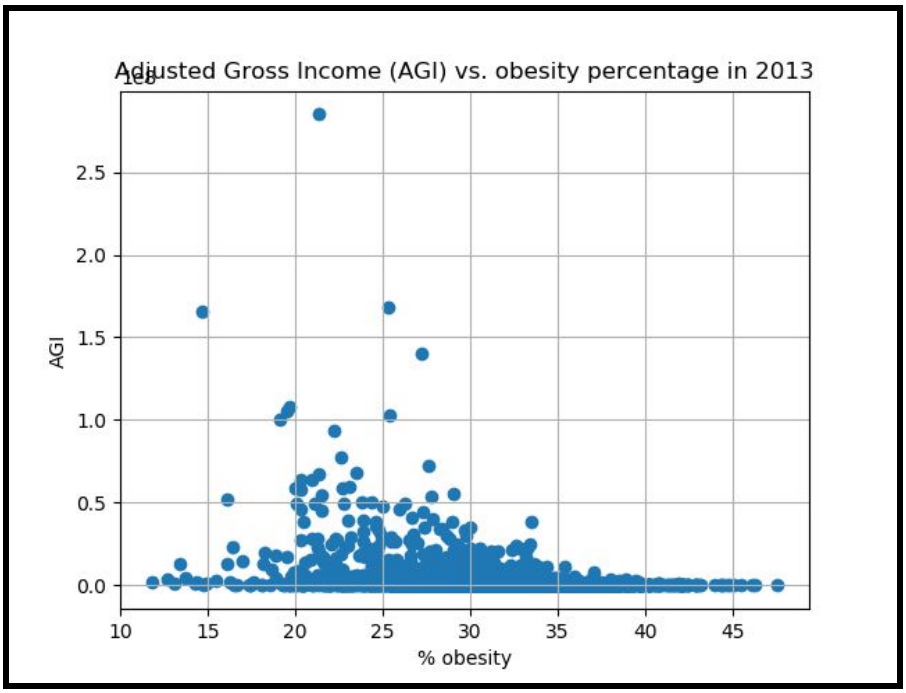
The health dataset was downloaded from the U.S. department of agriculture website. The health data includes both parameters from the years 2008 and 2013. The income data was downloaded from the IRS website. County income data was used, and not the provided individual income data. A combined health and data file was generated from the previous downloaded files. All US counties were represented in the generated combined file, except for Hawaii's Kalawao county, since no income data was listed in the income file.

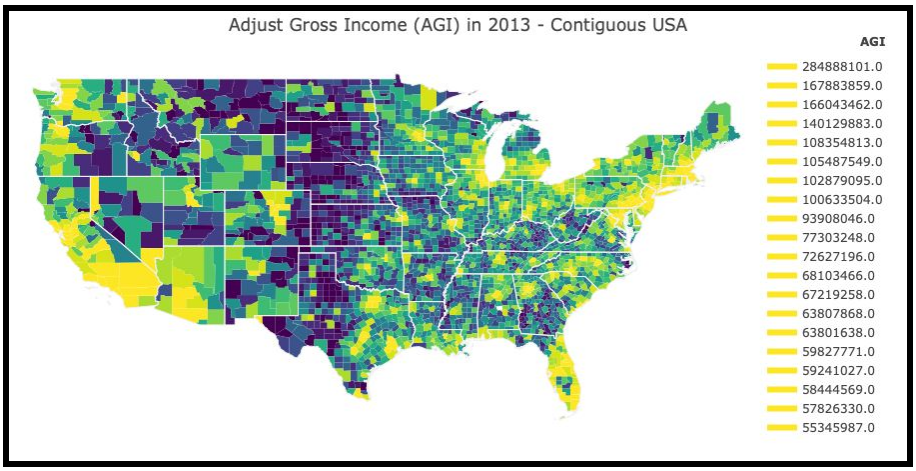
We highlight how previous work showed that it cannot be decisively concluded how income inequality affects individual health from population-level studies [3]. Our used data is classified as a population-level data, hence, the results here cannot firmly prove a relationship between social class income inequality and health levels.
The next sections of this report provide graphs, illustrations, and sources that give better understanding of health-income relationship in the US. References are provided at the end of this report.

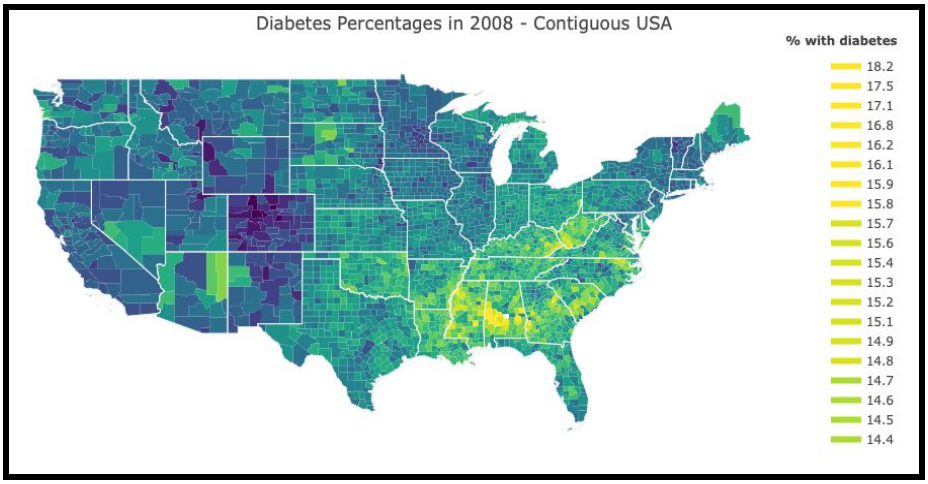
Methodology

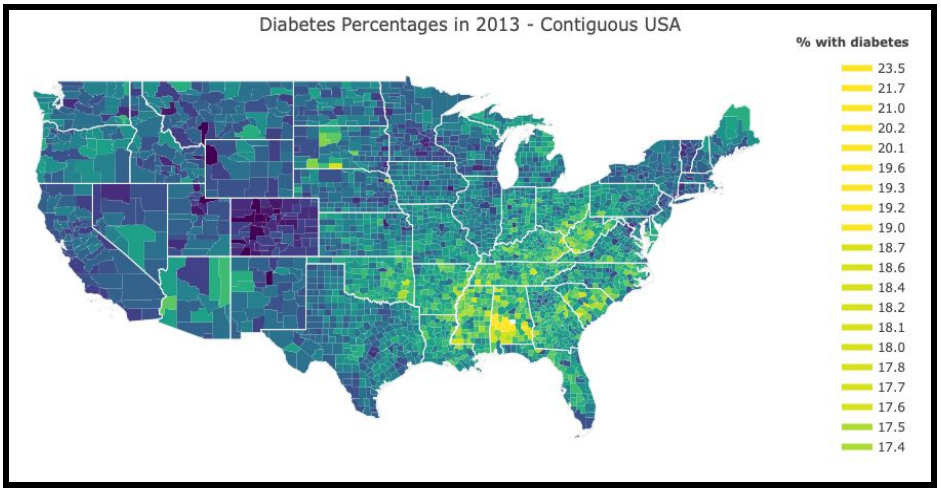
Results

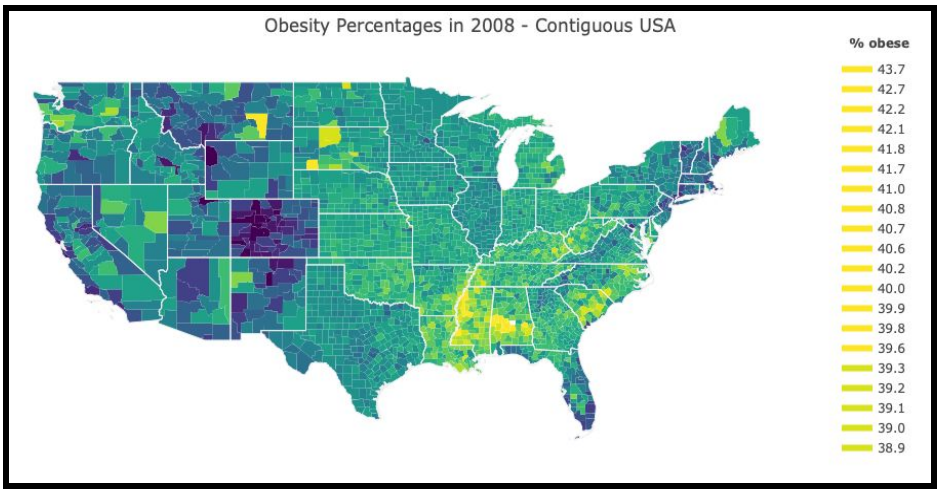


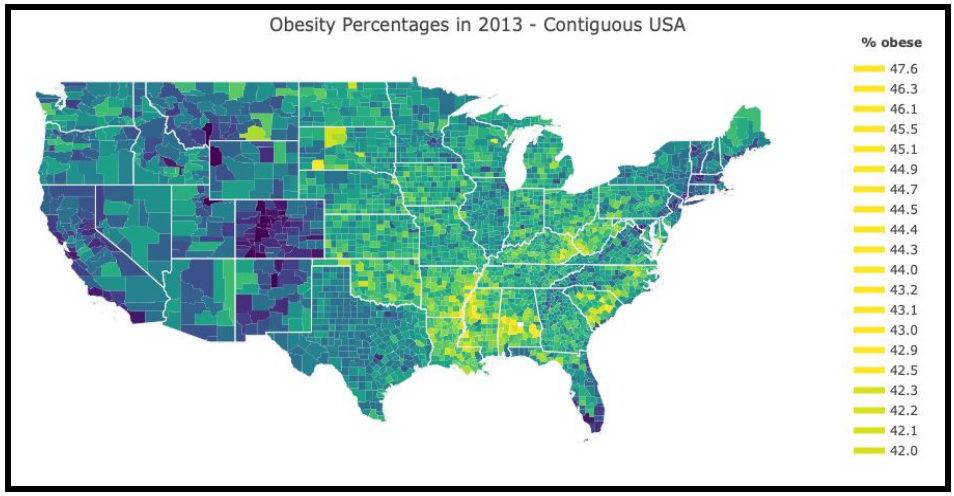

Adjusted Gross Income (AGI) vs. obesity percentage in 2013
1e8
2.5
2.0
1.5
1.0
0.5
0.0
AGI
10
15
20
25
30
35
40
45
% obesity

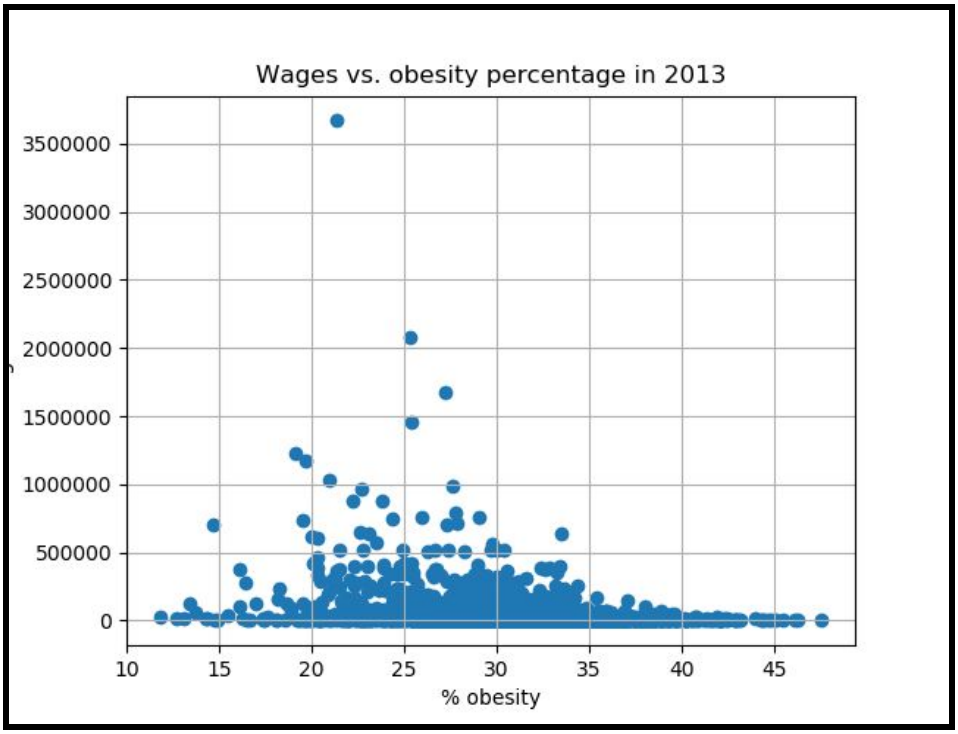

Adjust Gross Income (AGI) in 2013 - Contiguous USA
AGI
284888101.0
167883859.0
166043462.0
140129883.0
108354813.0
105487549.0
102879095.0
100633504.0
93908046.0
77303248.0
72627196.0
68103466.0
67219258.0
63807868.0
63801638.0
59827771.0
59241027.0
58444569.0
57826330.0
55345987.0

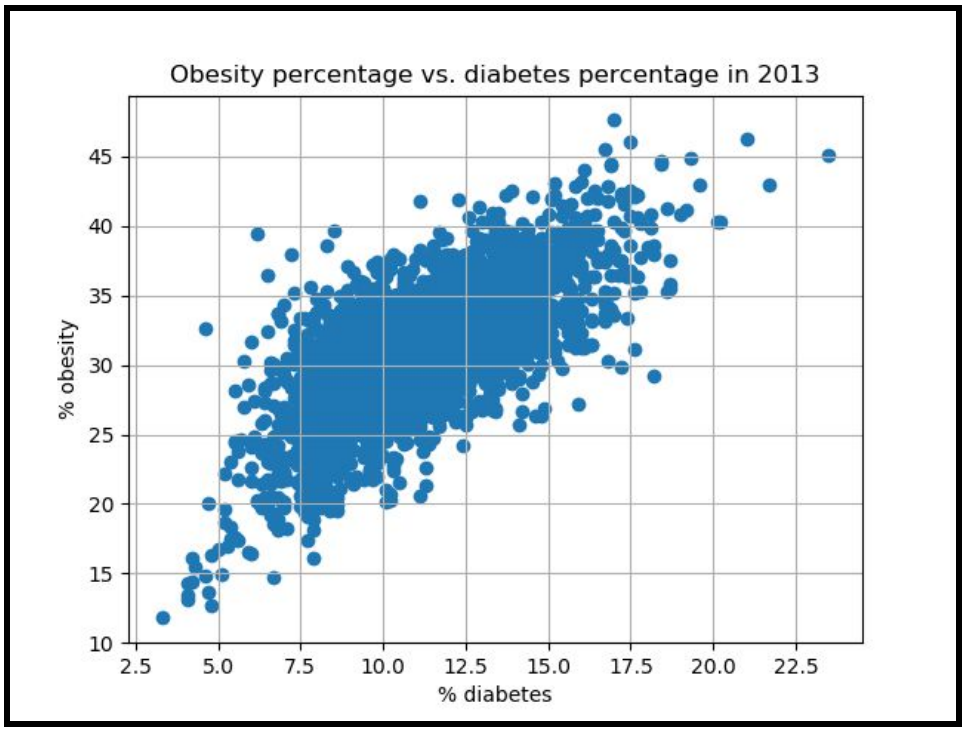

Diabetes Percentages in 2008 - Contiguous USA
% with diabetes
18.2
17.5
17.1
16.8
16.2
16.1
15.9
15.8
15.7
15.6
15.4
15.3
15.2
15.1
14.9
14.8
14.7
14.6
14.5
14.4


Diabetes Percentages in 2013 - Contiguous USA
% with diabetes
23.5
21.7
21.0
20.2
20.1
19.6
19.3
19.2
19.0
18.7
18.6
18.4
18.2
18.1
18.0
17.8
17.7
17.6
17.5
17.4


Obesity Percentages in 2008 - Contiguous USA
% obese
43.7
42.7
42.2
42.1
41.8
41.7
41.0
40.8
40.7
40.6
40.2
40.0
39.9
39.8
39.6
39.3
39.2
39.1
39.0
38.9


Obesity Percentages in 2013 - Contiguous USA
% obese
47.6
46.3
46.1
45.5
45.1
44.9
44.7
44.5
44.4
44.3
44.0
43.2
43.1
43.0
42.9
42.5
42.3
42.2
42.1
42.0


Wages vs. obesity percentage in 2013
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
10
15
20
25
30
35
40
45
% obesity


Obesity percentage vs. diabetes percentage in 2013
% obesity
45
40
35
30
25
20
15
10
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
% diabetes

Conclusion

References

[1] Kawachi, I., & Kennedy, B. P. (1999). Income inequality and health: pathways and mechanisms. *Health service research*, 34(1 Pt 2), 215-27.

[2] Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38), 16489-93.

[3] Wagstaff, Adam & Doorslaer, Eddy. (2000). Income Inequality and Health: What Does the Literature Tell Us?. *Annual review of public health,* 543-67.

[4] Bryan, Kevin A. and Martinez, Leonardo, On the Evolution of Income Inequality in the United States (2008). *FRB Richmond Economic Quarterly*, Vol. 94, No. 2, Spring 2008, 97-120.

[5] L Shavers, Vickie. (2007). Measurement of Socioeconomic Status in Health Disparities Research. *Journal of the National Medical Association*. 99, 1013-23.

[6] Marmot, Michael. (2002). The Influence Of Income On Health: Views Of An Epidemiologist. *Health affairs (Project Hope)*. 21, 31-46.

[7] Steven H. Woolf, Laudan Aron, Lisa Dubay, Sarah M. Simon, Emily Zimmerman, Kim X. Luk. (April 2015). How Are Income and Wealth Linked to Health and Longevity? *Center on Society and Health.*

[8] Serah Hyde. (February 2017) Income and health outcomes, *BLS*,

[9] Michaela Benzeval, Lyndal Bond, Mhairi Campbell, Mathew Egan, Theo Lorenc, Mark Petticrew and Frank Popham. (March 2014). How does Money influence Health? *Joseph Rowntree Foundation.*

[10] Karien Stronks, H Dike van de Mheen, Johan P Mackenbach, A higher prevalence of health problems in low income groups: does it reflect relative deprivation? *J Epidemiol Community Health* 1998;52:548–557.