

# Chronic Diseases and Their Relationship with Unhealthy Behaviors and Preventive Measures

*Jeremy James, Ali Mohamud, and Ben Donovan*

*Due on Oct 22, 2018*

## Abstract

Chronic diseases affect millions of Americans every day. They have no cure, nor do they go away with time. In this project, we analyze measures of chronic diseases, including unhealthy behaviors, health outcomes, and use of preventive services of 500 major cities in the United States. Through comparing different communities based on this data, we find which communities thrive or suffer, and report what factors play a role in their respective outcomes. To visualize this information, we use map plots to show the location of healthy or unhealthy cities, and scatter plots to show correlation between different measures and chronic diseases. We do not know any specific challenges we will face, nor do we know of any correlations between specific unhealthy behaviors or preventative services and chronic diseases. Potential future research could include investigating additional data sets for weather or socioeconomic factors to find more factors that play a role in increased rates of chronic diseases. This analysis will allow public health officials to allocate their limited resources based on the needs of each community, to better improve the well-being of the residents.

## 1 Introduction

Medicinenet.com gives the medical definition of chronic disease as “A disease that persists for a long time [(3 months +)].” They are generally not prevented by vaccines nor go away with medication or time[<https://www.medicinenet.com/script/main/art.asp?articlekey=33490>]. Table 1 displays the chronic diseases tracked in our data set. The different kinds of chronic diseases are measured by percentage of population affected on the national, city, and census tract level.

Table 1: Conditions and description of measure

Chronic Disease	Data Set Measures
Arthritis	Arthritis among adults aged $\geq 18$ Years
High Blood Pressure	High blood pressure among adults aged $\geq 18$ Years
Cancer (except skin)	Cancer (excluding skin cancer) among adults aged $\geq 18$ Years
Current Asthma	Current asthma among adults aged $\geq 18$ Years
Coronary Heart Disease	Coronary heart disease among adults aged $\geq 18$ Years
COPD	Chronic obstructive pulmonary disease among adults aged $\geq 18$ Years
Diabetes	Diagnosed diabetes among adults aged $\geq 18$ Years
High Cholesterol	High cholesterol among adults aged $\geq 18$ Years who have been screened in the past 5 Years
Chronic Kidney Disease	Chronic kidney disease among adults aged $\geq 18$ Years

Chronic Disease	Data Set Measures
Mental Health	Mental health not good for $\geq 14$ days among adults aged $\geq 18$ Years
Physical Health	Physical health not good for $\geq 14$ days among adults aged $\geq 18$ Years
Stroke	Stroke among adults aged $\geq 18$ Years
Teeth Loss	All teeth lost among adults aged $\geq 65$ Years

Unhealthy behaviors can lead to the development of chronic diseases. Table 2 displays the unhealthy behaviors tracked in our data set, measured by percentage of population affected on the national, city, and census track level.

Table 2: Unhealthy behaviors and description of measure

Unhealthy Behavior	Data Set Measures
Binge Drinking	Binge drinking among adults aged $\geq 18$ Years
Current Smoking	Current smoking among adults aged $\geq 18$ Years
Physical Inactivity	No leisure-time physical activity among adults aged $\geq 18$ Years
Obesity	Obesity among adults aged $\geq 18$ Years
Sleep $< 7$ hours	Sleeping less than 7 hours among adults aged $\geq 18$ Years

Though chronic diseases have no cure, they can be prevented. Table 3 displays the preventive services tracked in our data set, measured by percentage of population affected on the national, city, and census track level.

Table 3: Preventive service and description of measure The rest of the report will discuss the ‘500 Cities: Local Data for Better Health’ dataset. We will relate what we did to prepare the data for analyses, and display the data products we created from it. We will then make conclusions on what we found.

Preventive Service	Data Set Measures
Health Insurance	Current lack of health insurance among adults aged 18-64 Years
Taking BP Medication	Taking medicine for high blood pressure control among adults aged $\geq 18$ Years with high blood pressure
Annual Checkup	Visits to doctor for routine checkup within the past Year among adults aged $\geq 18$ Years
Cholesterol Screening	Cholesterol screening among adults aged $\geq 18$ Years
Colorectal Cancer Screening	Fecal occult blood test, sigmoidoscopy, or colonoscopy among adults aged 50-75 Years
Pap Smear Test	Papanicolaou smear use among adult women aged 21-65 Years

Preventive Service	Data Set Measures
Core preventive services for older men	Older adult men aged $\geq 65$ Years who are up to date on a core set of clinical preventive services: Flu shot past Year, PPV shot ever, Colorectal cancer screening
Core preventive services for older women	Older adult women aged $\geq 65$ Years who are up to date on a core set of clinical preventive services: Flu shot past Year, PPV shot ever, Colorectal cancer screening, and Mammogram past 2 Years
Dental Visit	Visits to dentist or dental clinic among adults aged $\geq 18$ Years
Mammography	Mammography use among women aged 50-74 Years

## 2 Description of the data.

The data that we have used in this project is the '500 Cities: Local Data for Better Health' data obtained from the CDC's '500 Cities project' [<https://catalog.data.gov/dataset/500-cities-local-data-for-better-health-b32fd>]. This is a publically available data set that includes a model-based small area estimates for 27 measures of chronic disease related to unhealthy behaviors (5), health outcomes (13) and use of preventive services (9) for 500 cities throughout the United States. The primary aim of the data is to make it possible to identify emerging health problems and effective interventions at city level.

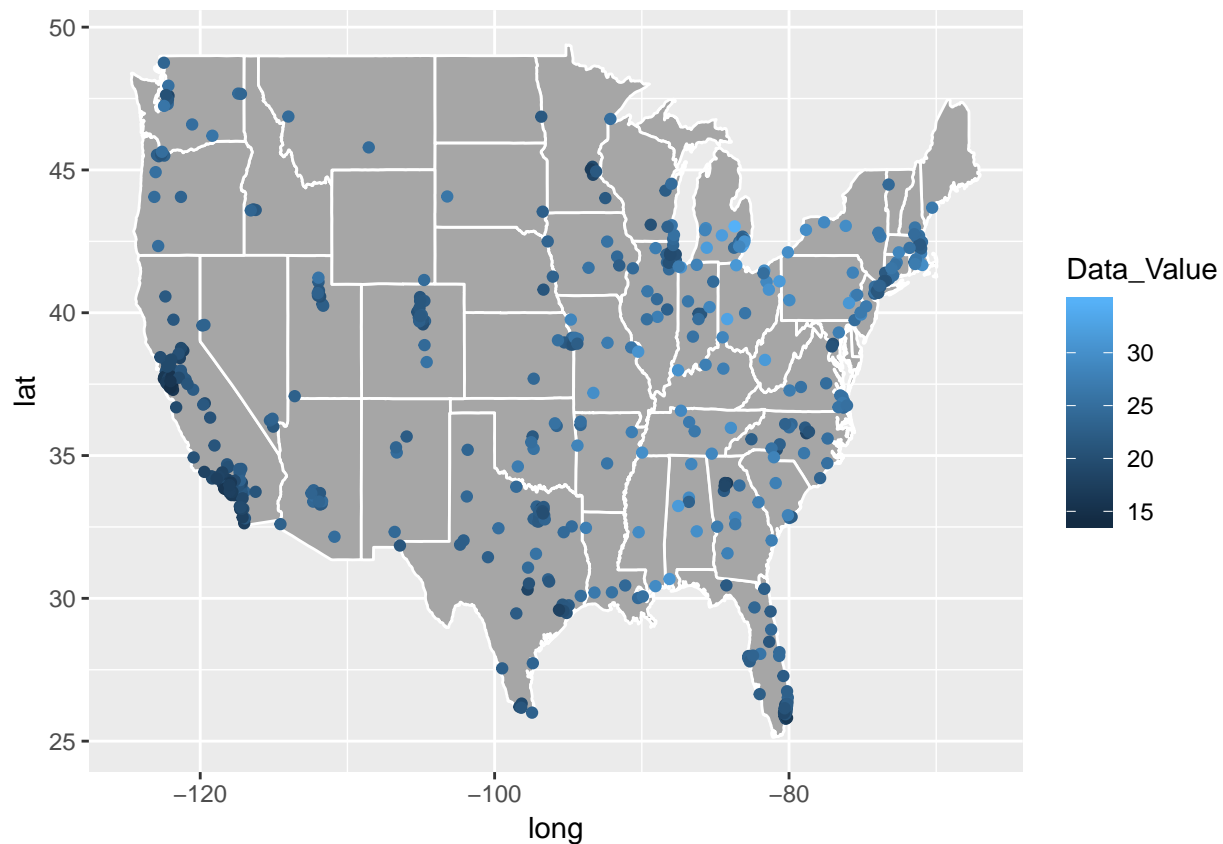
### 2.1 Preparing the Data.

The data was read from the above link into R using the `read.csv` function. This data set for this project did not require data manipulation. However, All the columns of the data set were not necessary for the analysis. Therefore, we subsetting by selecting only the column that were relevant for the analysis. Here are the columns we will use: *UniqueID*, *CityName*, *StateAbbr*, *GeoLocation*, *Year*, *Measure*, *Data\_Value*, *PopulationCount*, *GeographicLevel*, *Short\_Question\_Text*, *Category* and *DataValueTypeID*

## 3 Data Exploration

This section is incomplete. We will be doing more exploration of this dataset as the semester progresses.

### 3.1 Mapping Chronic Disease levels



shiny dashboard? usmap ##3.2 Comparison of large and small cities

This plot visualizes the distribution of percentages of chronic disease in large and small cities, where large cities have populations greater than 100,000.

### 3.3 Unhealthy Behaviours and Health Outcomes

In this section, we will explore the relationship between unhealthy behaviors and chronic diseases.

```
library(reshape2)
```

```
##  
## Attaching package: 'reshape2'  
## The following object is masked from 'package:tidyr':  
##  
## smiths
```

```
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
correlation_matrix_data <- dcast(data[data$GeographicLevel=="Census Tract",], UniqueID+PopulationCount~  
row.names(correlation_matrix_data) <- correlation_matrix_data$UniqueID  
correlation_matrix_data$UniqueID <- NULL  
correlation_matrix_data <- correlation_matrix_data[complete.cases(correlation_matrix_data),]  
correlation_matrix <- cor(correlation_matrix_data)
```

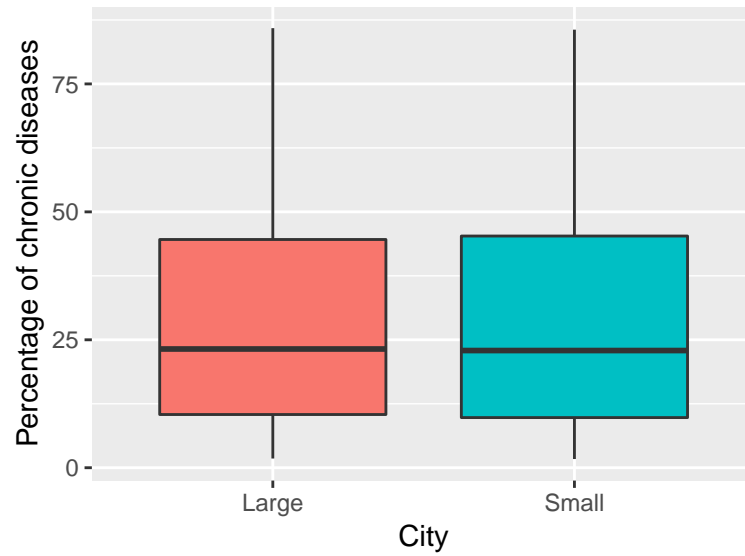


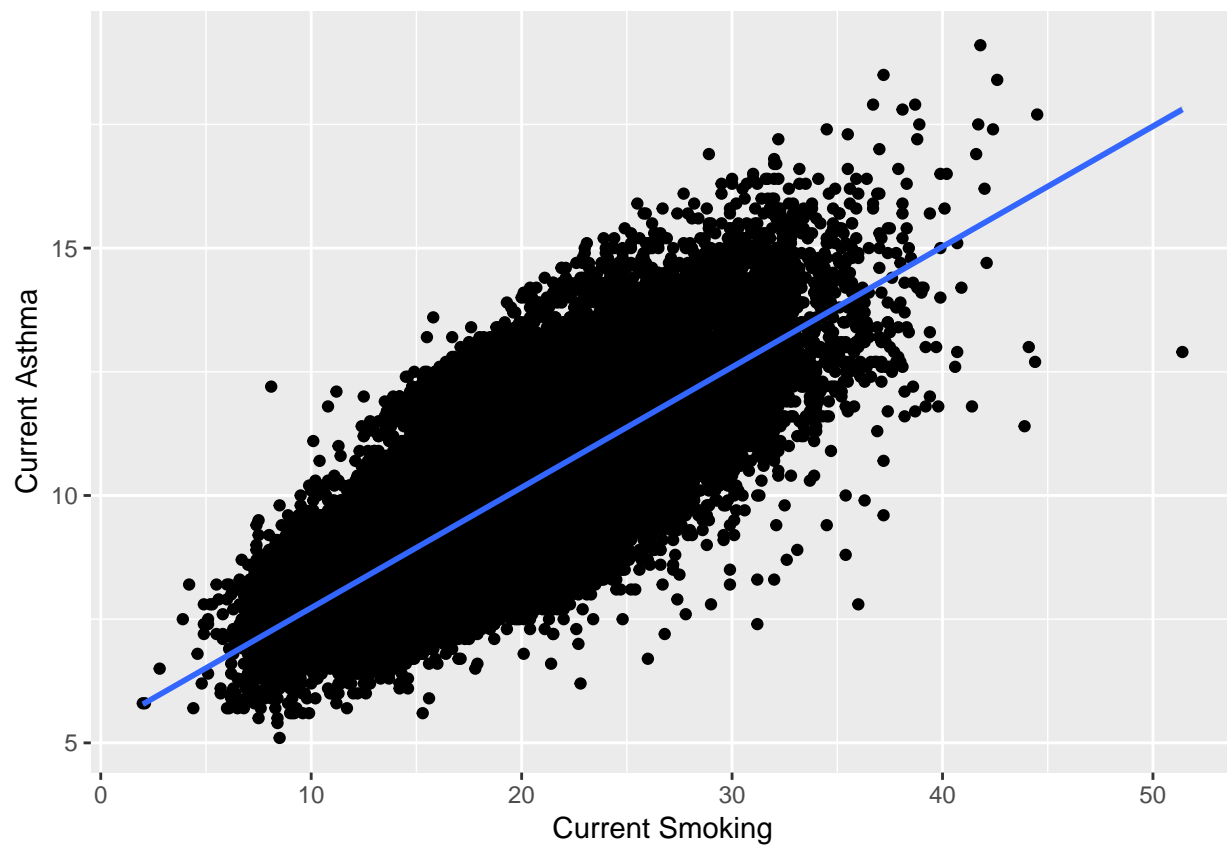
Figure 1: Comparison of chronic disease rates in big and small cities

```
correlation_matrix <- round(correlation_matrix, 2)
corrplot(correlation_matrix, title = "Correlation Plot", insig = "pch", addrect = 3)
```



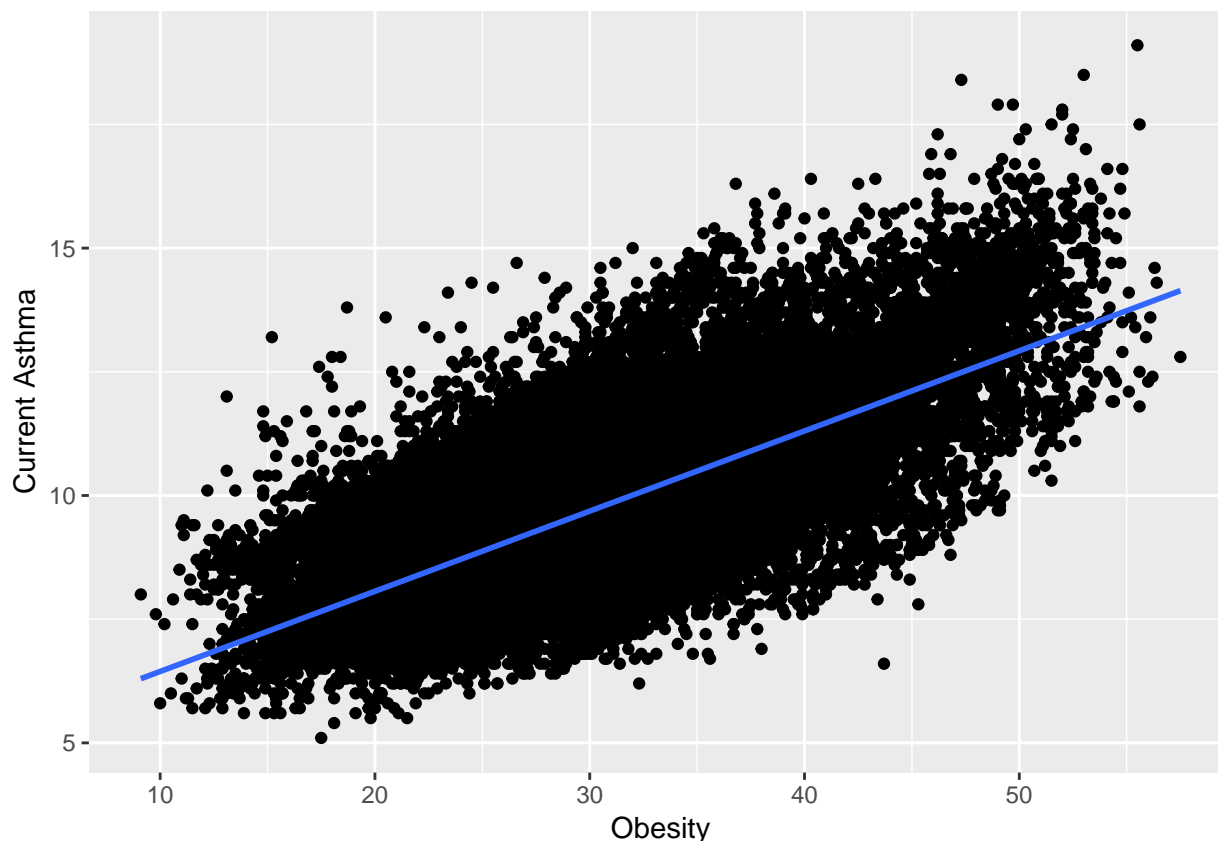
### 3.3.1 Smoking and asthma

According to the ALA, cigarette smoke irritates the airways. Smokers have a high risk of asthma, as well as those who inhale the smoke secondhand. Therefore, we should expect to find that communities with higher smoking rates would also have higher asthma rates.



### 3.3.2 Obesity and Asthma

The ALA has found that though obesity is associated with asthma, it is unclear why. Some believe it is due to inflammation in the body that occur with extra weight. At the census tract level, asthma is clearly associated with obesity.



### 3.4 Preventive Measures and Health Outcomes

In this section, we will explore the relationship between preventive measures and chronic diseases.

## 4 Findings

This section is still in progress. It will contain our interpretations of the data we have analyzed.

### 4.1 Comparison of large and small cities

As we saw from a boxplot, population has little to no affect on overall chronic disease rates. We can also confirm this by using a t-test with the hypothesis  $H_0 : \mu_L - \mu_S = 0$  versus the hypothesis  $H_a : \mu_L - \mu_S \neq 0$ . Here  $\mu_L$  is a mean percentage of chronic diseases in the large cities (count of peoples is more than 100000) and  $\mu_S$  is a mean percentage of chronic diseases in the small cities (count of peoples is less or equal to 100000).

Table 4: Welch Two Sample t-test:  
v1\$ageadjcitydat.Data\_Value by v11 (continued below)

Test statistic	df	P value	Alternative hypothesis
0.4789	13407	0.632	two.sided



mean in group Large	mean in group Small
30.48	30.28

Since the p-value is greater than significance level  $\alpha = 0.05$  then we do not reject a null hypothesis  $H_0$  and we conclude that no significant difference of the mean percentage of chronic diseases in the large and small cities. In the future, we will look at each chronic disease individually to see if any specific disease is correlated with population.

## 5 Conclusion

We have no conclusions to report yet, other than that city population has little to do with city health outcomes. In our final report, this section will have more details on our findings.