**ANALYTICAL STUDY ON UNITED STATES**

**CHRONIC DISEASE INDICATORS USING TABLEAU**

**SUBMITED TO: DR. SHILPA BALAN**

**PRESENTED**

**BY**

**KRITHY NANAIAH ATRANGADA**

**CIS 5270: BUSINESS INTELLIGENCE**

**OBJECTIVE OF STUDY:**

Chronic disease: A disease that persists for a long time. A chronic disease is one lasting three months or more. Chronic diseases generally cannot be prevented by vaccines or cured by medication, nor do they just disappear. Eighty-eight percent of Americans over 65 years of age have at least one chronic health condition. Health damaging behaviors - particularly tobacco use, lack of physical activity, and poor eating habits - are major contributors to the leading chronic diseases. Chronic diseases are important health problem which can result in morbidity, mortality, disability and decreased quality life. Chronic disease represents seven of the top ten causes of death in United States in 2010. These diseases can be prevented by choosing healthy behaviors—like avoiding tobacco, eating and drinking healthy foods and beverages, and getting regular physical activity and enough sleep—people can reduce their chance of getting a chronic disease or improve their health and quality of life if they already have a chronic disease. The Centers for Disease Control and Prevention (CDC) works to prevent chronic diseases and reduce their health and financial costs with an approach that brings together monitoring to understand the problems, communities to work toward solutions and support healthy behaviors, and health care systems to detect problems early and help reduce or eliminate unhealthy behaviors. This project will help me to analyze the reason due to which the chronic disease is increasing every year and among which age group it’s the highest. I can also analyze the location in which the chronic disease is highest.

**About Dataset**

The Chronic Disease Indicators (CDI) is a web-based data system of surveillance indicators developed by consensus among CDC, the Council of State and Territorial Epidemiologists (CSTE), and the National Association of Chronic Disease Directors (NACDD). CDI allows public health professionals and policymakers to retrieve uniformly-defined data for chronic diseases and risk factors with substantial public health burden. These indicators are essential for surveillance, prioritization, and evaluation of public health interventions. The CDI expands to approximately 200 indicators, including existing topics: alcohol, asthma, arthritis, cancer screening, cancer incidence and mortality, cardiovascular disease, chronic obstructive pulmonary disease, diabetes and End Stage Renal Disease, nutrition and physical activity, immunization, oral health, tobacco, and overarching conditions/socioeconomic status; and new topics: maternal and child health, older adults, and school health. For the first time, CDI also include indicators of systems and environmental change. CDC enhances the CDI website with more user-friendly display, navigation, and data retrieval functionality. Because of the close state-CDC partnership in reviewing the CDI, these changes directly reflect the priorities and needs of the states.

**A). DATA-SET:** <https://catalog.data.gov/dataset/u-s-chronic-disease-indicators-cdi-e50c9>

INSIGHTS: CHRONIC DISEASE INDICATORS:

This data is taken from U.S. Department of Health & Human Services. This data set includes:

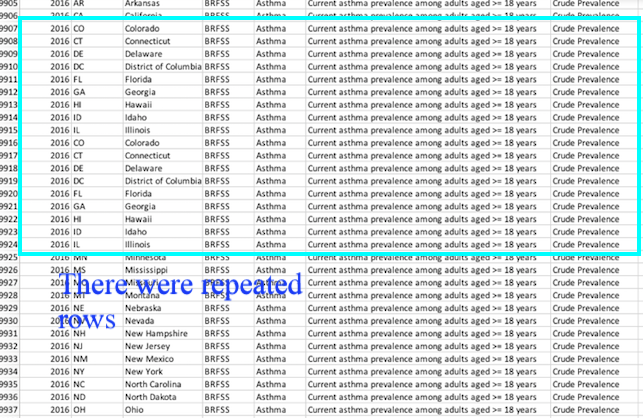
1. The year in which chronic disease started.
2. Increment or decrement in number of chronic disease over the years.
3. Year in which the chronic disease ended.
4. Cause of the chronic disease.
5. The location in which the disease broke out.
6. Average age group where the chronic disease is commonly found.
7. Specific category of race and gender where in its common

This data set will help me analyze the chronic disease indicators in United States.

**B)**. **DATA CLEANING:**

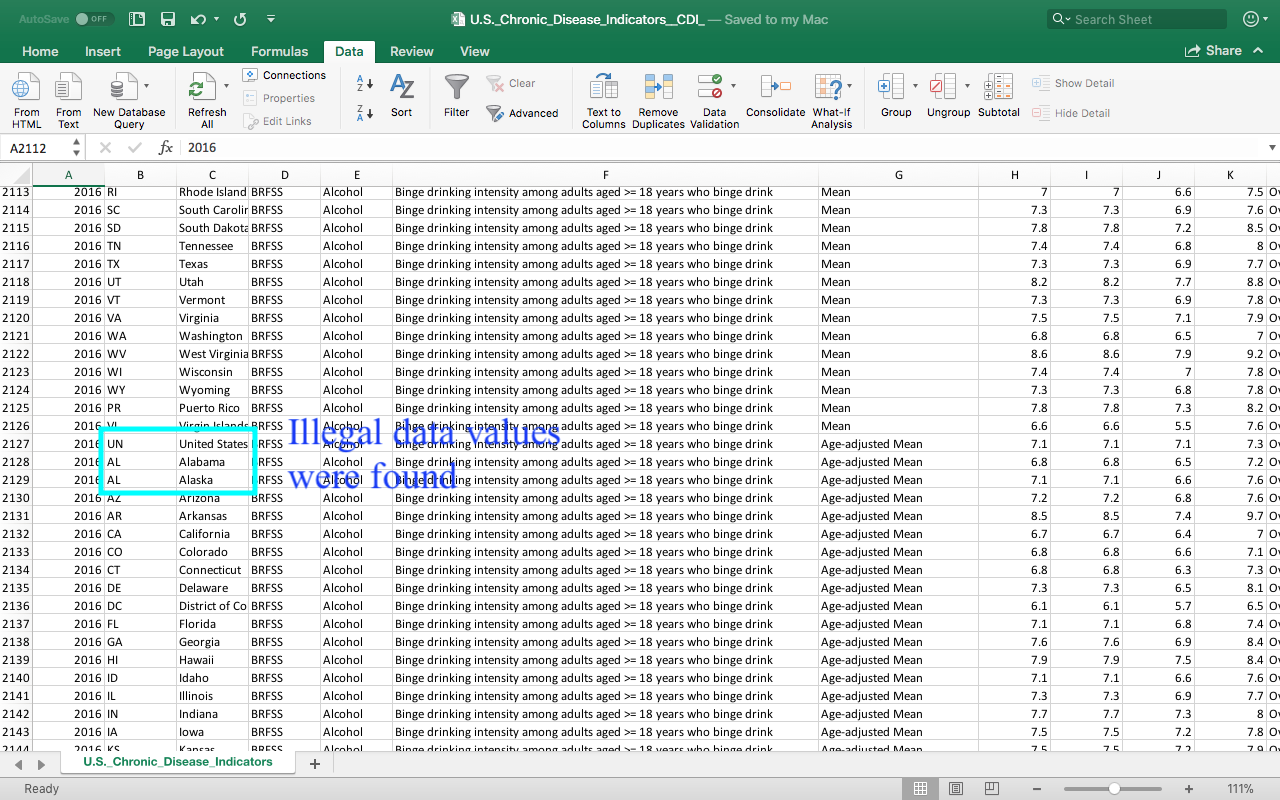
In order to get correct results for analysis it is very crucial that data should be clean enough otherwise results can be manipulated. The following cleaning has been done.

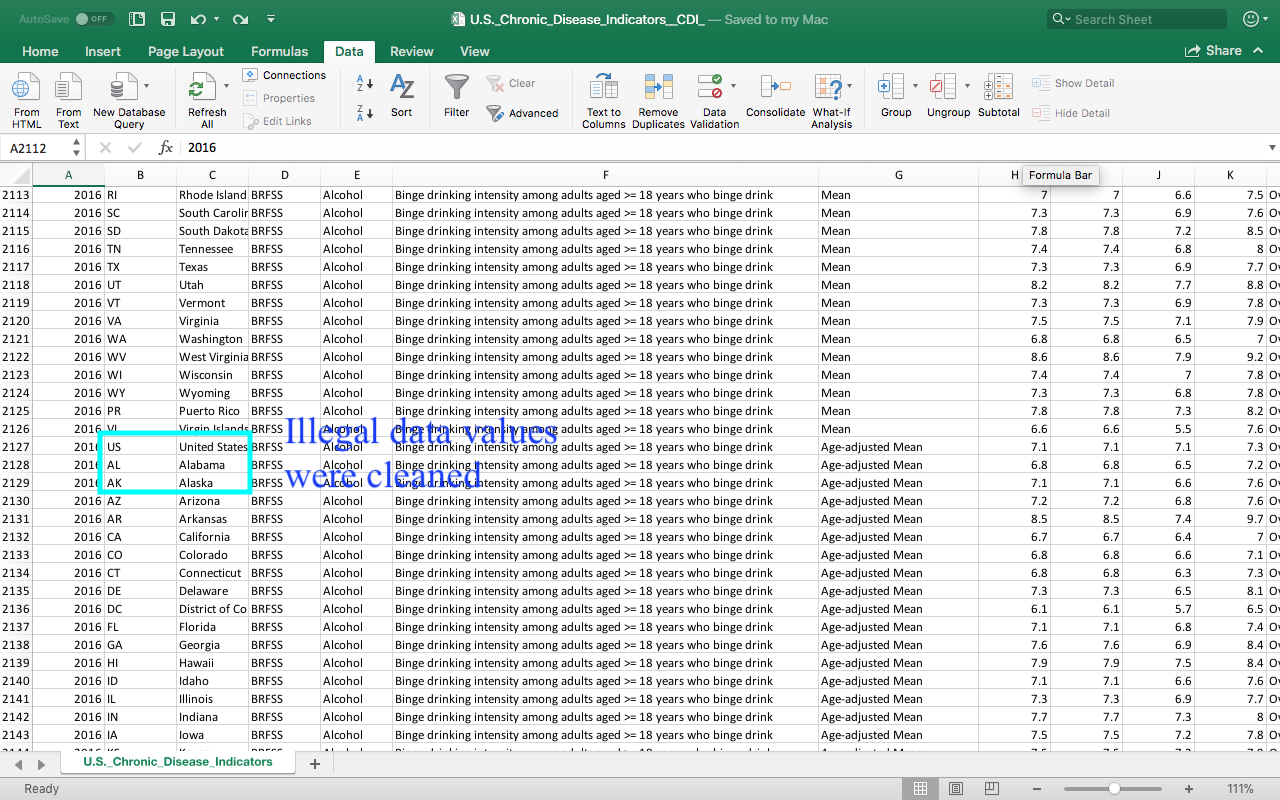
1. UNIQUENESS VIOLATION



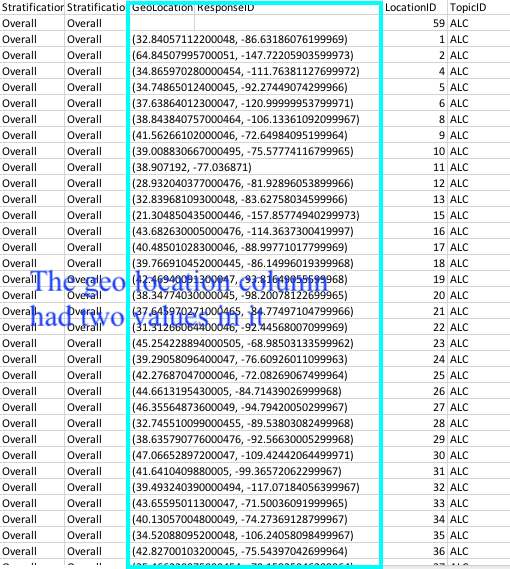


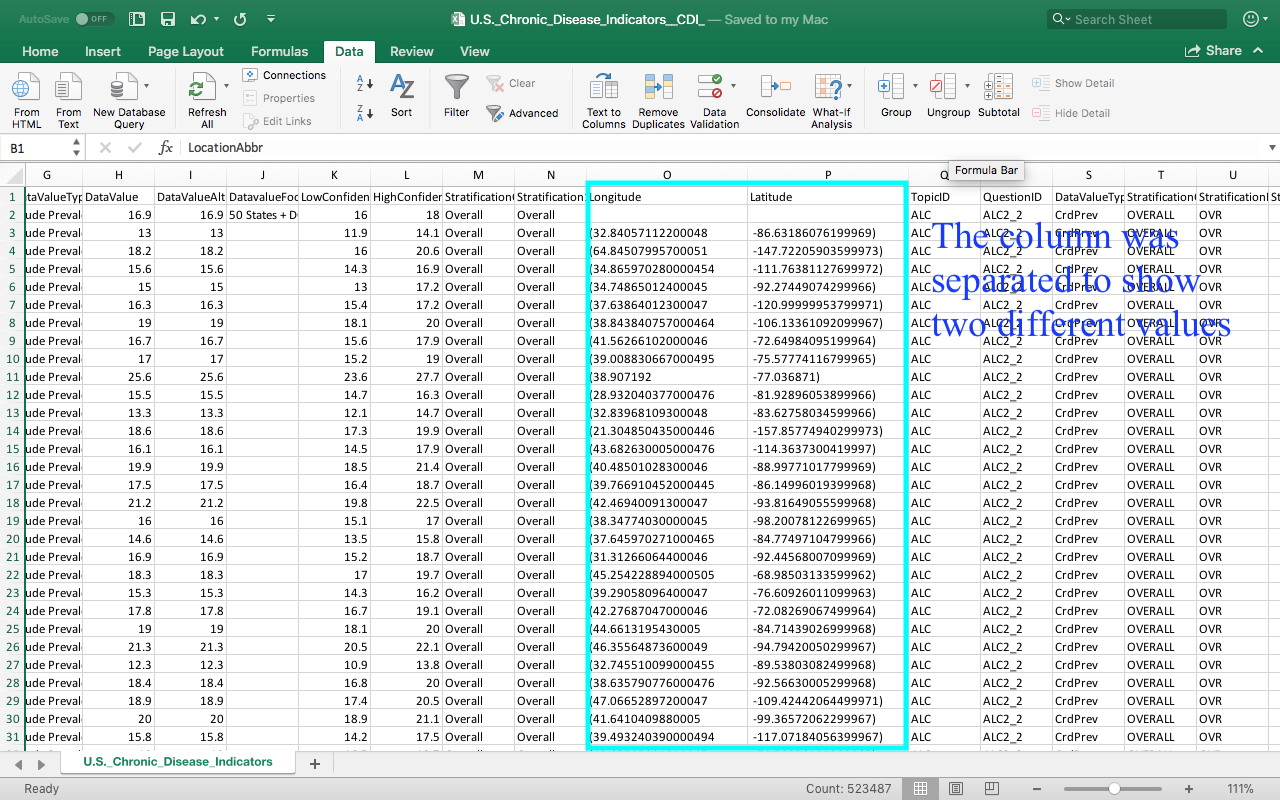
1. ILLEGAL DATA FORMAT





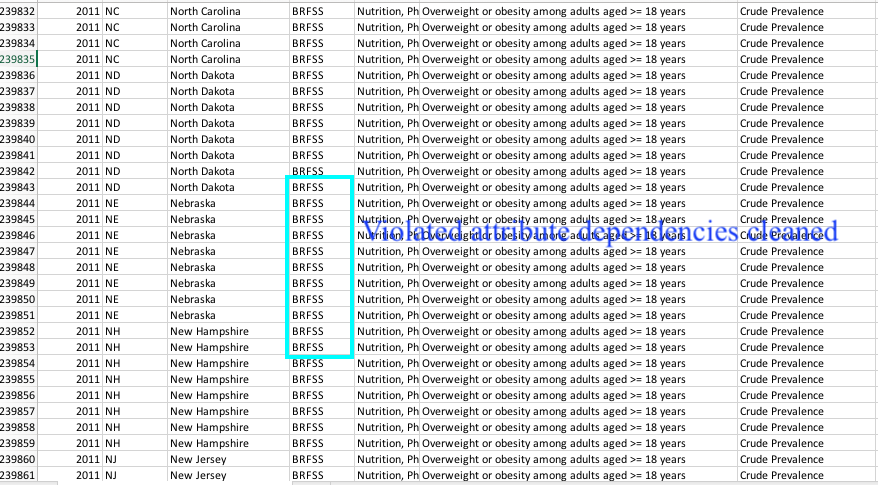
1. EMBEDDED VALUES



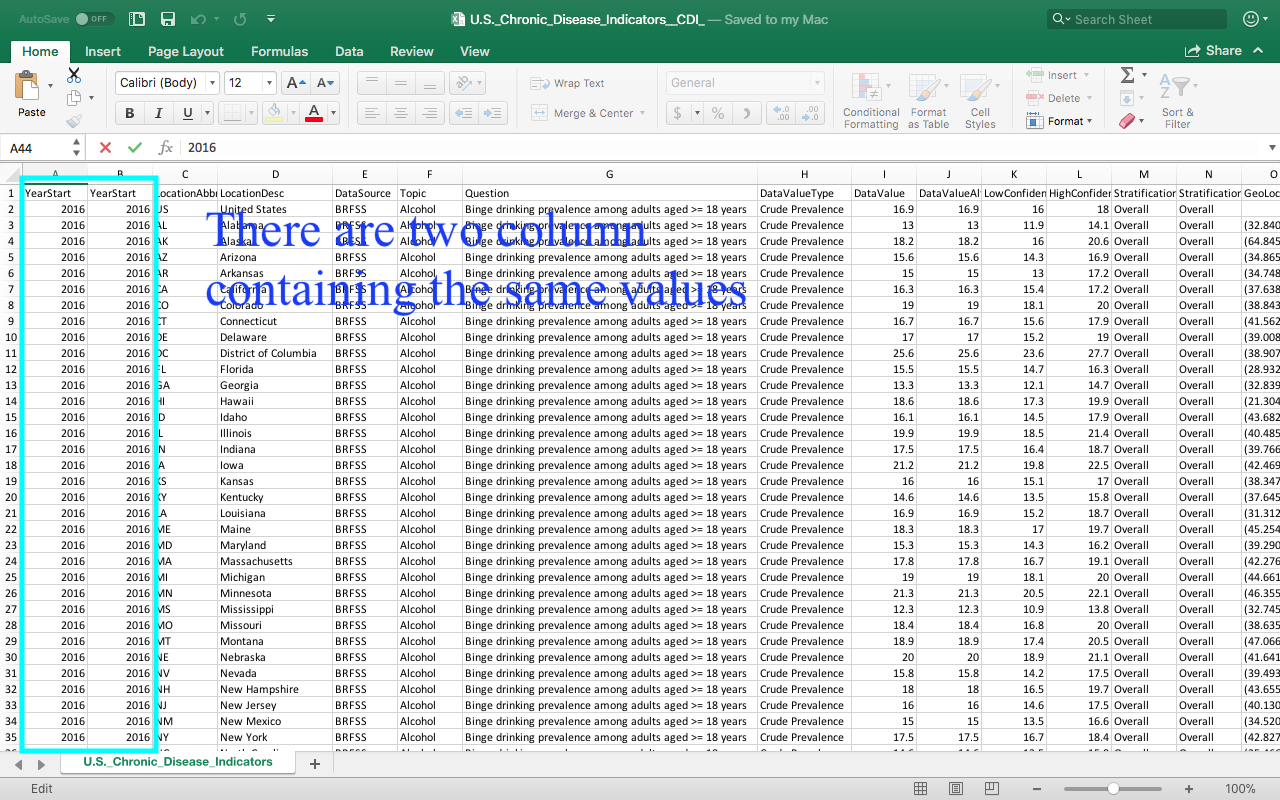


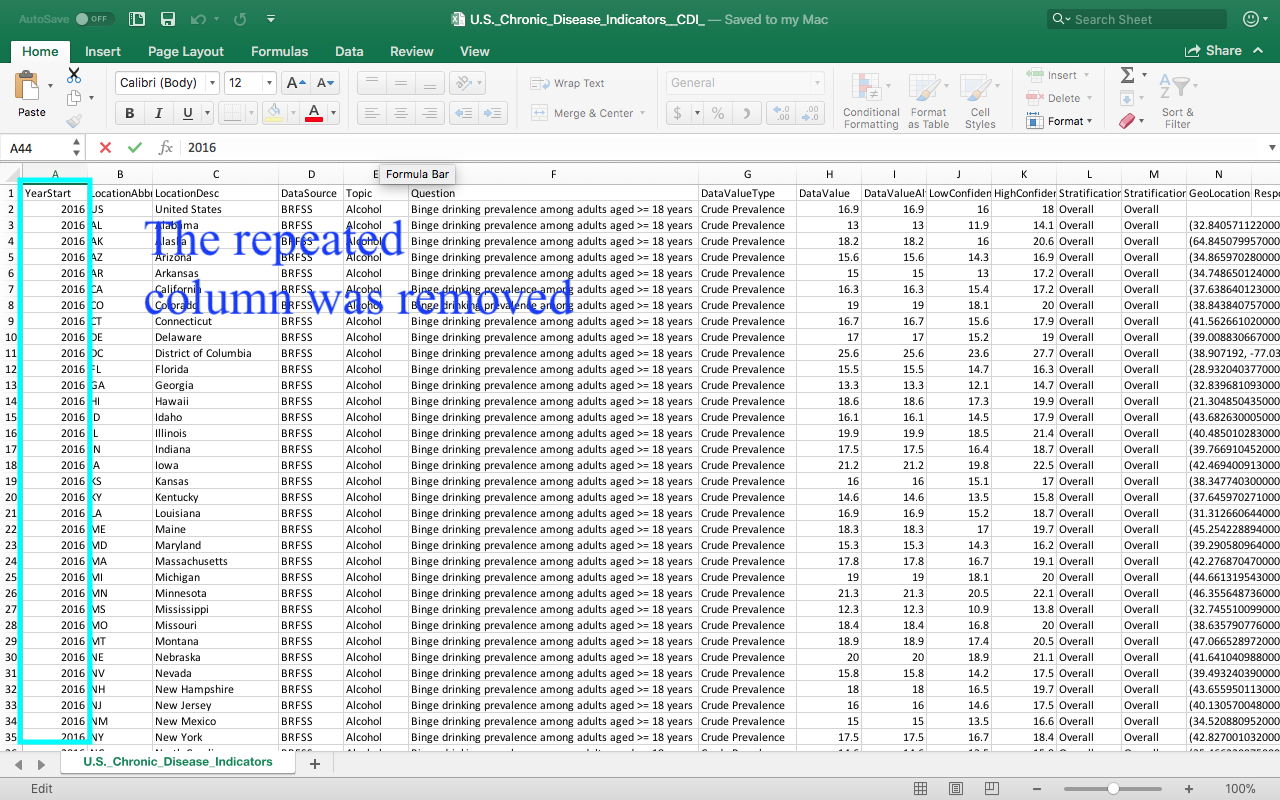
1. VIOLATED ATTRIBUTE DEPENDENCIES





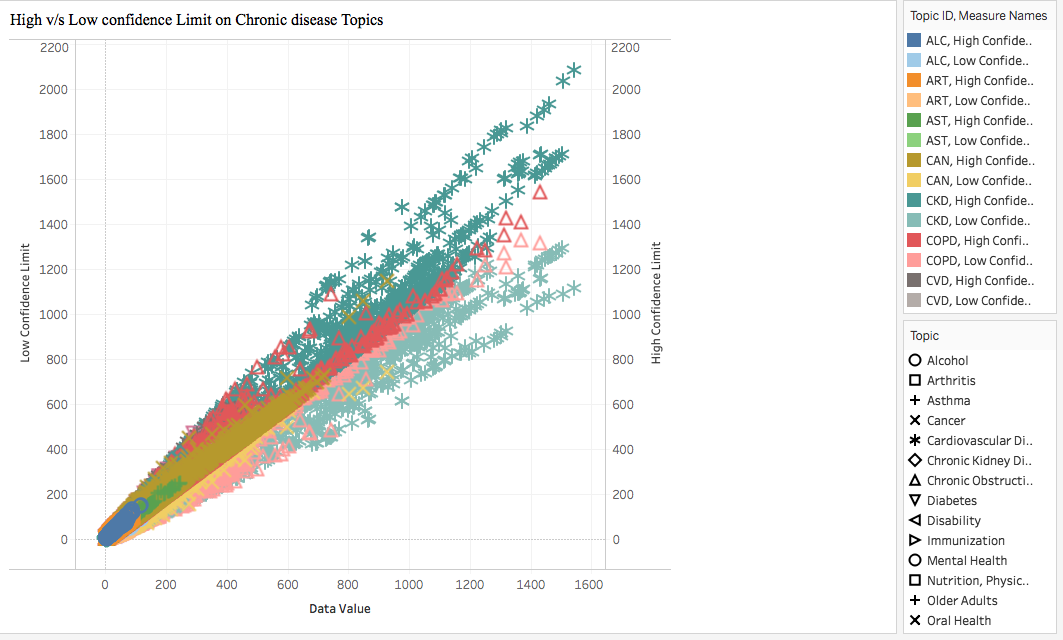
1. DUPLICATE RECORDS





**C). DATA VISUALIZATIONS:**

1. **Insight-1 :** High v/s Low confidence Limit on Chronic disease Topics



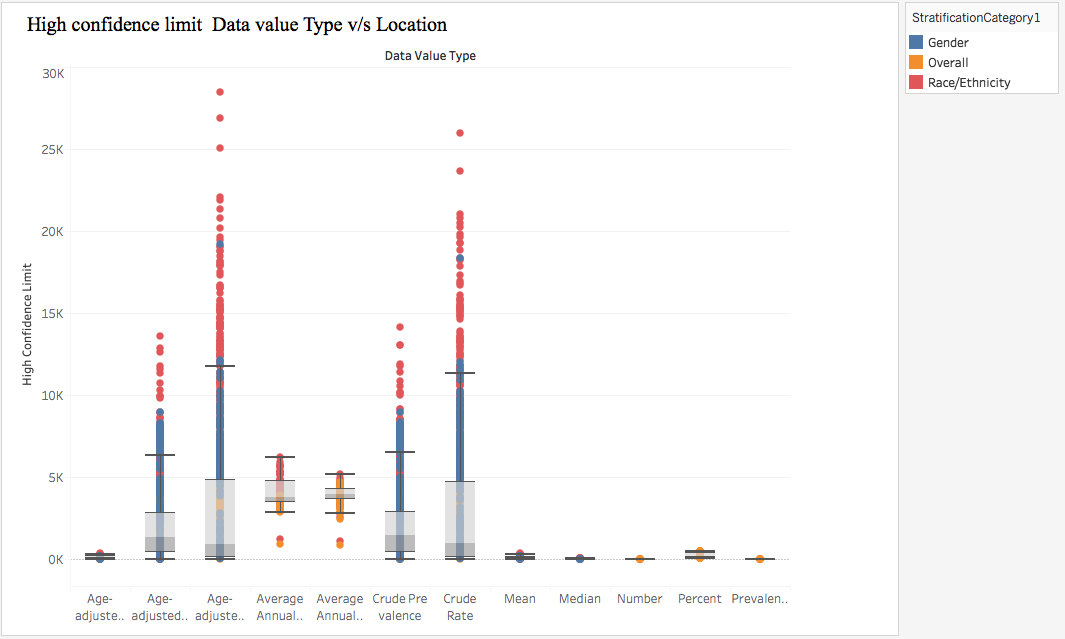
[Used : Scatter Plots]

The scatter plot shows us the higher confidence limit and lower confidence limit of the data values.

It shows different topics within the range of the data values. The chronic diseases like alcohol, overarching conditions, cardio vascular diseases, cancer, chronic obstructive pulmonary disease are at the higher side and asthma, arthritis and nutrition are on the lower side. These reports suggests that the people of United States of America should think wisely before eating and should consume more nutritional added fruits and vegetables in their meals. The study also suggests that people of all group and ages must exercise daily.

The people with higher age groups must get themselves checked regularly and should have proper hours of sleep in their daily life.

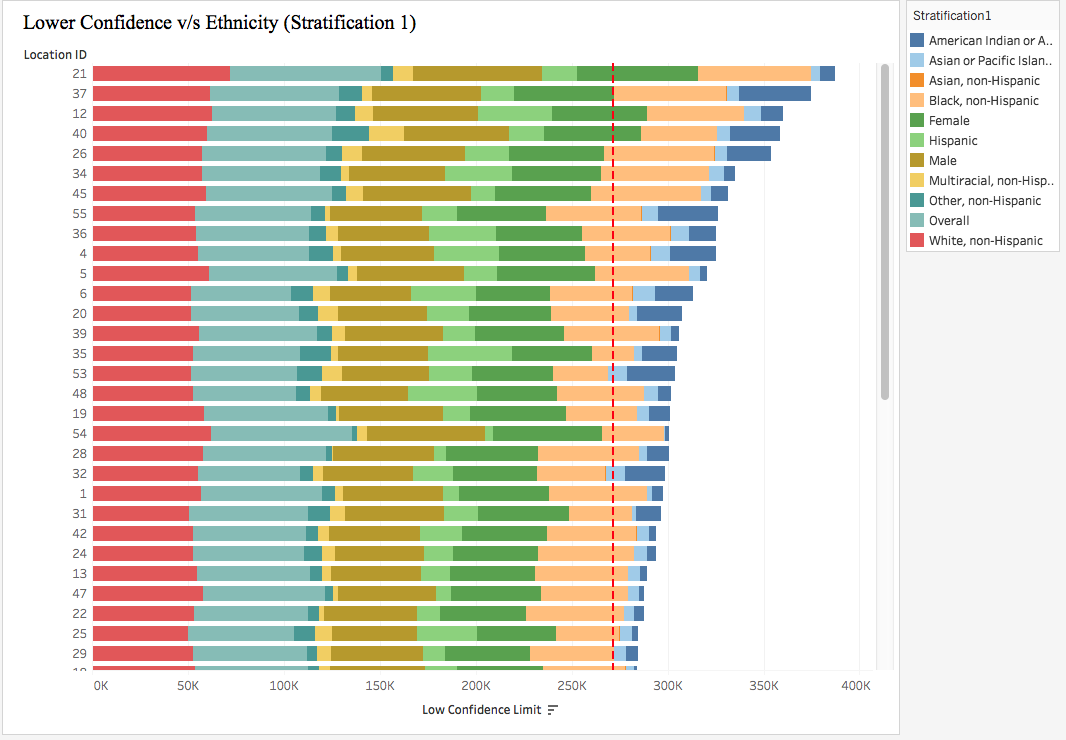
1. **Insight-2 :** High confidence limit Data value Type v/s Location

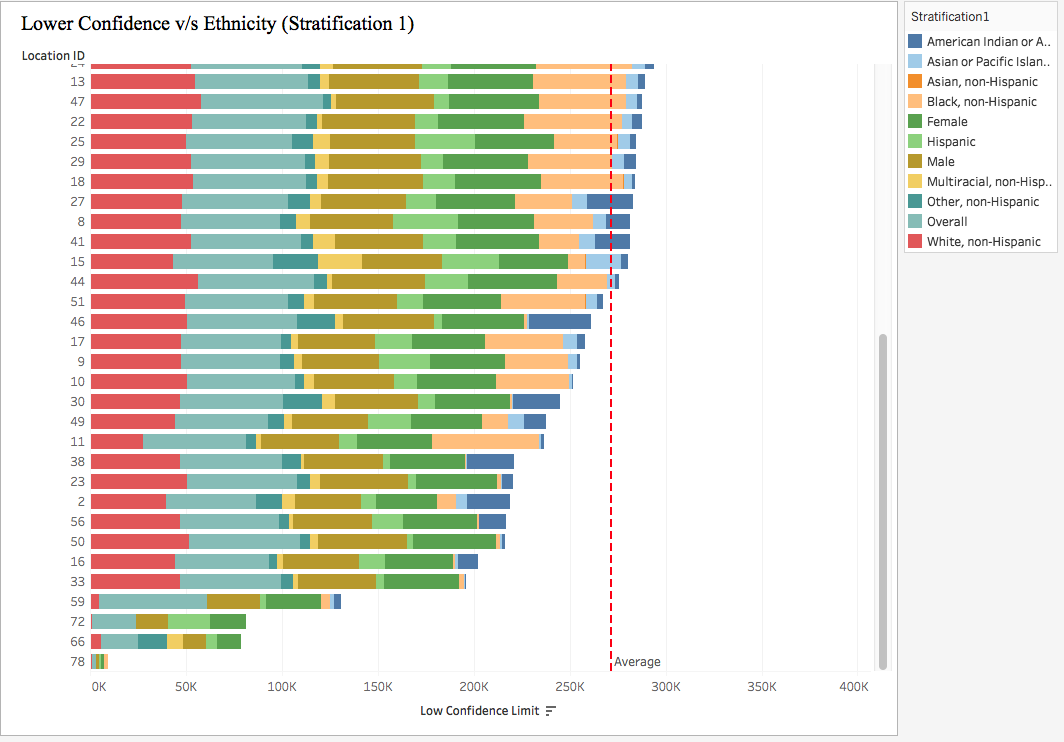


[Used: Box and Whisker Plot]

This insight gives us the high confidence limit values of all the data value types the chronic disease data value has been divided into. If we look closely in age adjusted rate and crude rate we find some of the values are outliers. The upper whisker, lower whisker and the median tells us that with age the possibilities of chronic disease increases and hence the health sectors across United States should educate people about necessary precautions to be taken to prevent themselves from these chronic diseases.

1. **Insight-3 :** Lower Confidence v/s Ethnicity (Stratification 1)

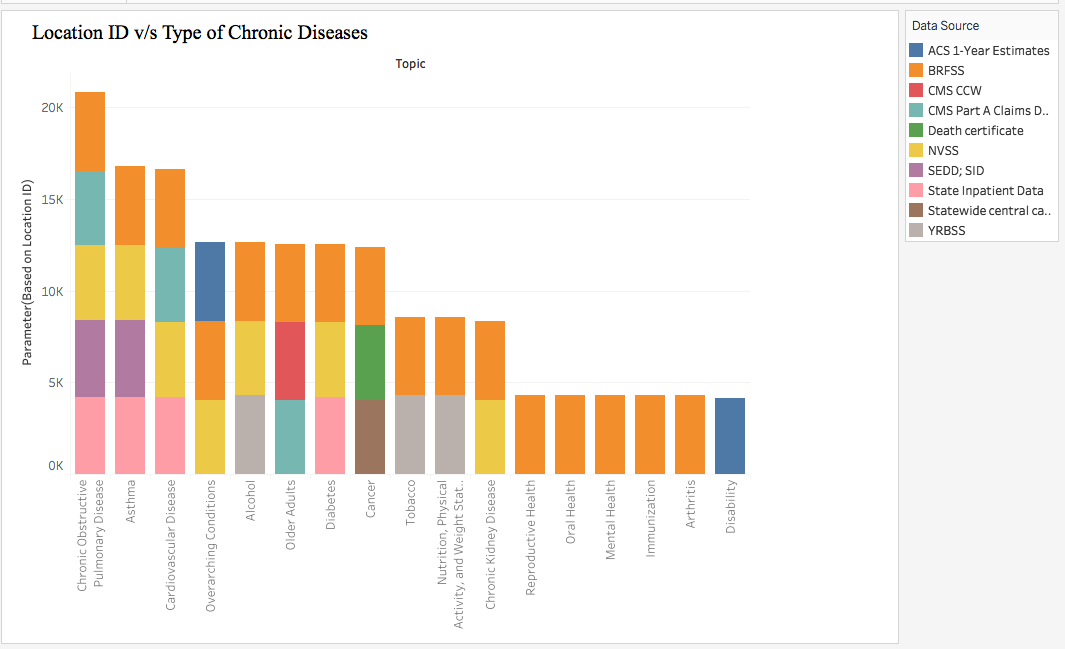




[Used: Reference Lines, Filtering the Location]

This report gives us the highlight of the lower confidence limit based on the people ethnicity. We have taken the average of the low confidence limit which is represented by reference line . In the report we see that Kentucky is on the higher side of the low confidence and overall population of people in Kentucky are suffering through chronic disease problems. This insights can help the government to take necessary steps to fight the chronic disease problems.

1. **Insight-4:** Location ID v/s Type of Chronic Diseases

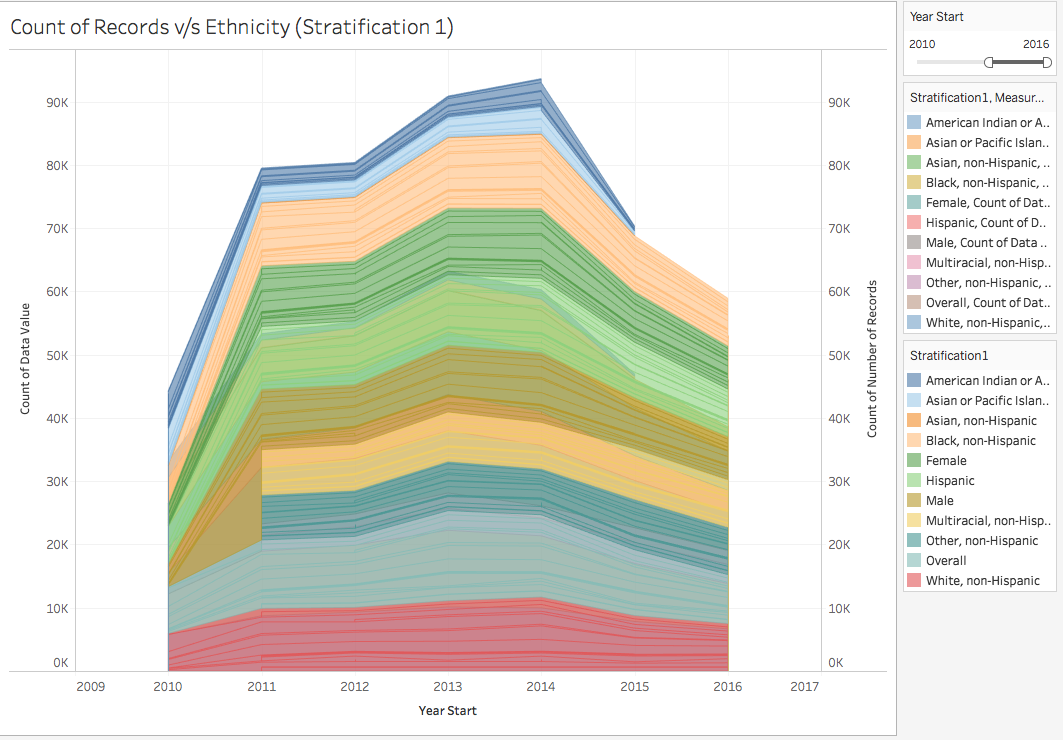


[Used: Parameters (Bar Graph)]

This analysis is primarily done to show that chronic obstructive pulmonary disease is the highest in United States. The different colours indicates the various sources of data it is been collected from. These pulmonary disease is basically due to smoking and using various other substances. The health authorities must take strict actions to supress the use of illegal substance which is causing chronic infections among the people of US.

The location ID is having a parameter of MIN and MAX.

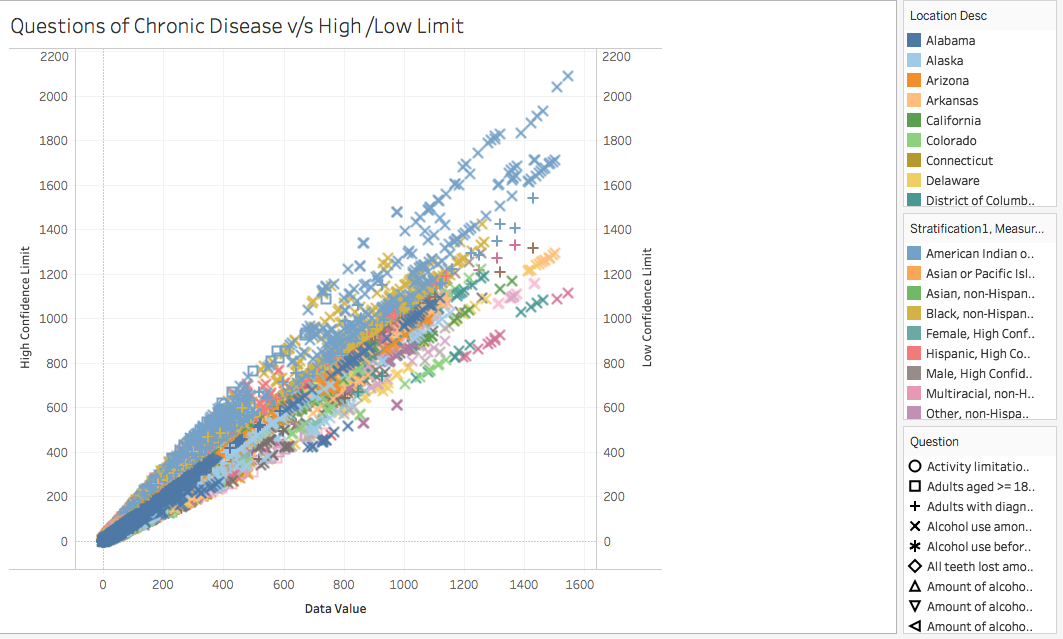
1. **Insight-5:** Count of Records v/s Ethnicity (Stratification 1)



[Used: Area Chart]

In this insight we have the people with different ethnicity and the count of the records of different chronic diseases in US starting from year 2010 till 2016. In this we see that the intake of alcohol is all time high in most of the years. The data value gives us the in depth knowledge of the various disease affecting the population of US.

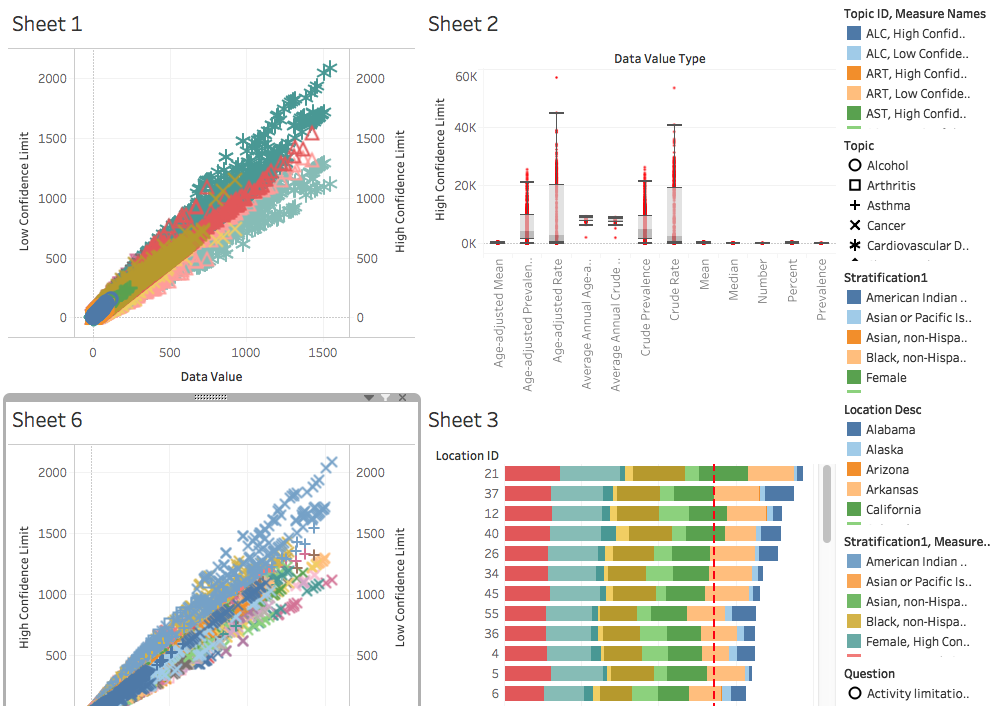
1. **Insight-6 :** Questions of Chronic Disease v/s High /Low Limit



[Used : Dual Axis, Scatter Plot]

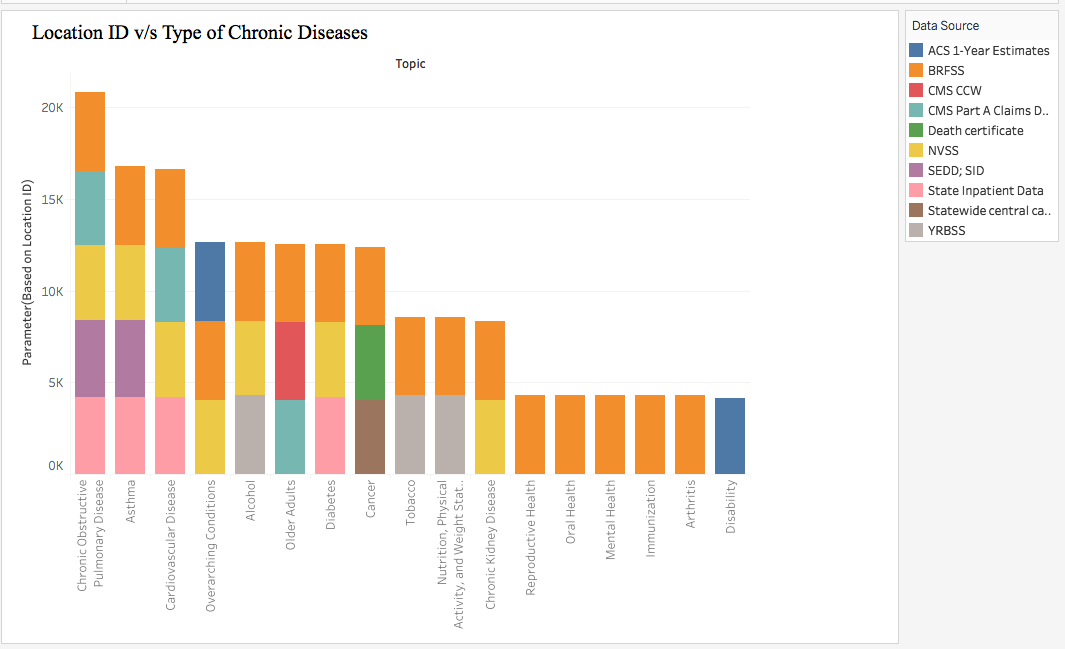
In this insight we analyse the different questions associated with different people in different locations. We group these questions in the low confident limit and high confident limit and populate the scatter plot with based on the data value recorded over a period of time. The insight tells us that some of the most common chronic disease are gradually decreasing.

**D). DASHBOARD:**



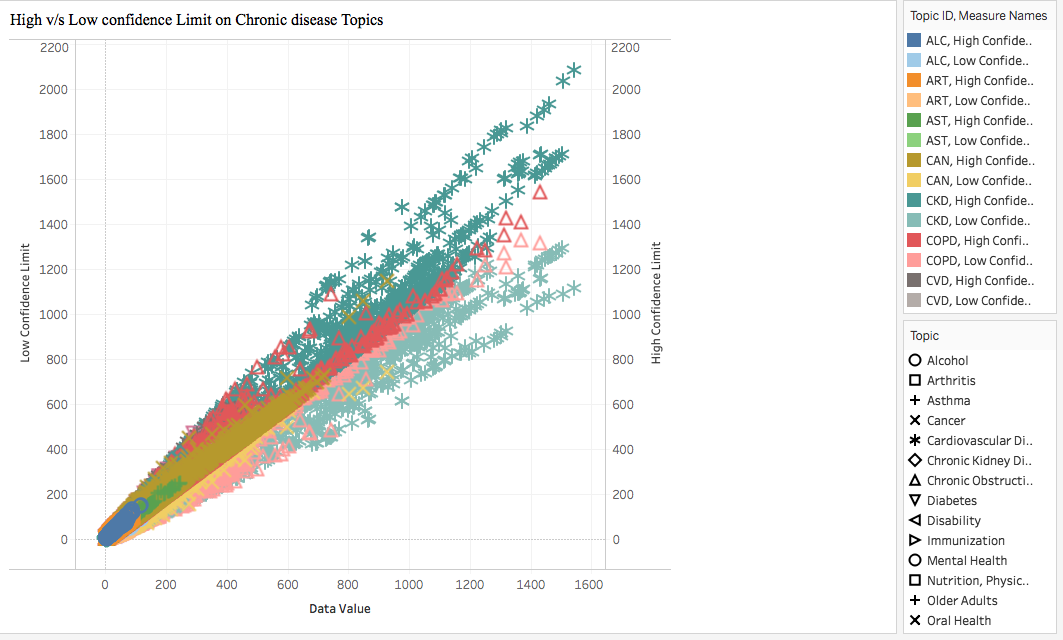
**E). STORY TELLING:**

Chronic diseases are an important public health problem, which can result in morbidity, mortality, disability, and decreased quality of life. Chronic diseases represented seven of the top 10 causes of death in the United States The chronic disease indicators (CDI) are a set of surveillance indicators developed by consensus among CDC, the Council of State and Territorial Epidemiologists (CSTE), and the National Association of Chronic Disease Directors (NACDD). CDI enables public health professionals and policymakers to retrieve uniformly defined state and selected metropolitan-level data for chronic diseases and risk factors that have a substantial impact on public health. These indicators are essential for surveillance, prioritization, and evaluation of public health interventions.  Several of the current chronic disease indicators are available and reported on other websites, either by the data source/custodians or by categorical chronic disease programs. However, CDI is the only integrated source for comprehensive access to a wide range of indicators for the surveillance of chronic diseases, conditions, and risk factors at the state level and for selected large metropolitan areas.  As a result, CDI increased to 124 indicators in the following 18 topic groups: alcohol; arthritis; asthma; cancer; cardiovascular disease; chronic kidney disease; chronic obstructive pulmonary disease; diabetes; immunization; nutrition, physical activity, and weight status; oral health; tobacco; overarching conditions; and new topic areas that include disability, mental health, older adults, reproductive health, and school health. Current trends in US population growth, age distribution, and disease dynamics foretell rises in the prevalence of chronic diseases and other chronic conditions. These trends include the rapidly growing population of older adults, the increasing life expectancy associated with advances in public health and clinical medicine, the persistently high prevalence of some risk factors, and the emerging high prevalence of multiple chronic conditions.



Although preventing and mitigating the effect of chronic conditions requires sufficient measurement capacities, such measurement has been constrained by lack of consistency in definitions and diagnostic classification schemes and by heterogeneity in data systems and methods of data collection. Preventing and mitigating the effect of any single chronic condition, or constellation of conditions, requires improved measurement. The nation is recognizing the emerging high prevalence of multiple chronic conditions (MCC) and related implications for prevention, treatment, public health programs, and planning. People who have MCC may require increased coordination of care from clinicians, public health, and social programs to improve their overall quality of life. As the prevalence of chronic conditions continues to increase in the US population, the United States will face even greater challenges in delivering care to people with MCC.

Current trends in population growth, age distribution, and disease dynamics foretell rises in the prevalence of chronic diseases, other chronic conditions, and combinations of chronic conditions. Such trends threaten both the public and financial health of the United States and include the rapidly growing population of older adults, the increasing life expectancy associated with advances in public health and clinical medicine, and the persistently high prevalence of some risk factors



Future revisions of the CDIs will likely include additional MCC (measuring chronic conditions) indicators as research in this field increases and data become available. In addition, future reviews might address data on health behaviors among younger children and additional data on systems and environmental indicators, should additional data become available. The CDIs are an example of collaboration among CDC and state health departments in building a consensus set of state-based health surveillance indicators. This update will help ensure that the CDI data remain the most relevant and current collection of chronic disease surveillance data for state epidemiologists, chronic disease program officials, and reproductive health and maternal and child health officials. The newly revised indicators are aligned with Healthy People 2020.

**References:**

1. Centre for Disease Control and Prevention

<https://www.cdc.gov/cdi/overview.html>

2. <https://www.cdc.gov/mmwr/pdf/rr/rr6401.pdf>

3. National Library of Medicine National Institute of Health

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3652713/>