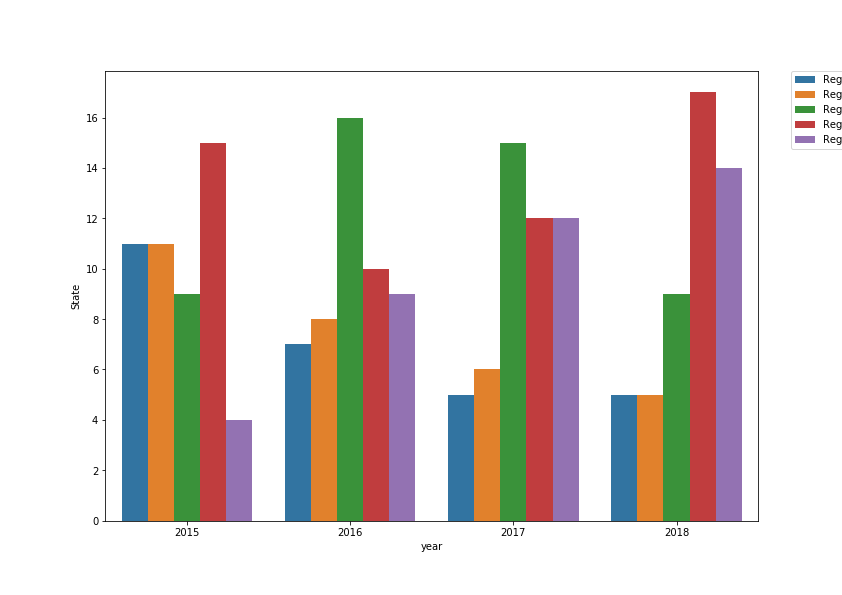
**FINAL REPORT**

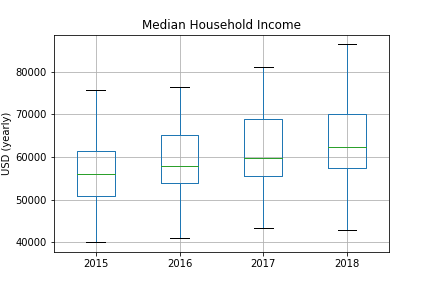
**Does the Median Household Income has increased through time?**



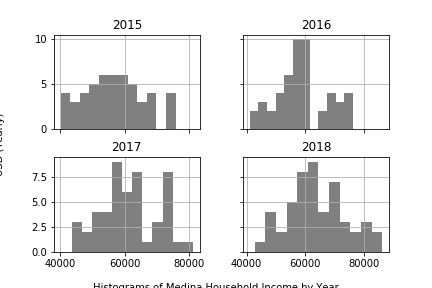
The barplot shows that Region 1 and Region 2, which are the states with the lowest Median Household Income, have decreased considerably being in 2018 only 5 states compared to the 11 states in 2015. Region 5 presents the states that have the highest Median Household Income. It is clear that the tendency is increasing year by year.

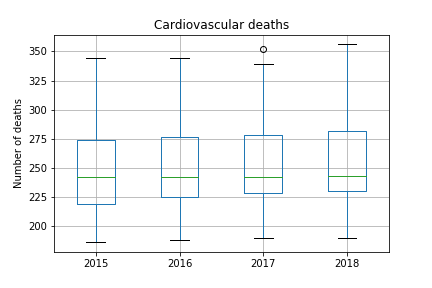
In the case of Region 3, in 2015 was the region with the highest number of states belonging to this classification. However, it shows a considerably decreased, but recover in the following 2 years.

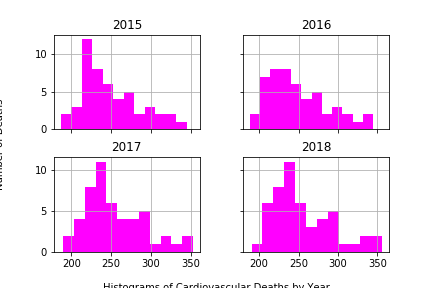
Descriptive analysis



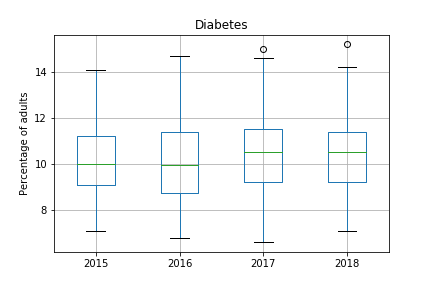
The median of each year is close to 60,000 USD yearly. However, 2017 and 2018 have the major variance compared to 2015 and 2016. No outliers are present.



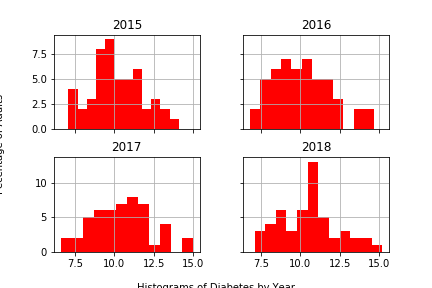


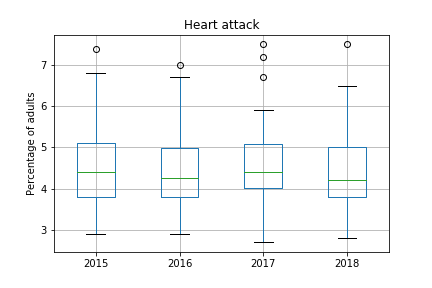


In all years, it is clear that there is a bias to the right side. The variance is similar in each year and there is one outlier present in year 2017.

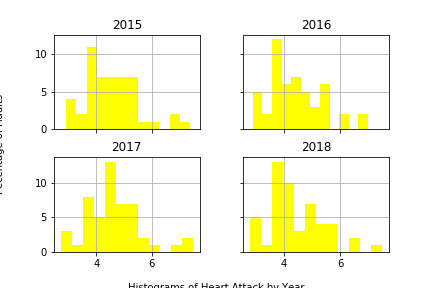


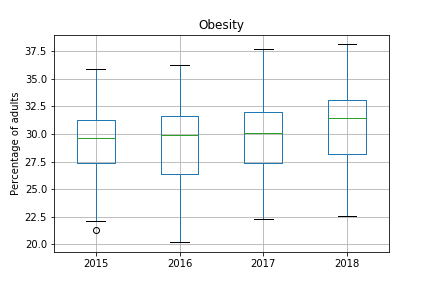
The year that has the lowest variance compared to the rest of the years is 2015. Nevertheless, it has a slight bias to the right. The years 2017 and 2018 have an outlier each one.



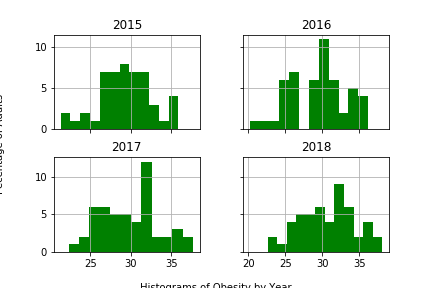


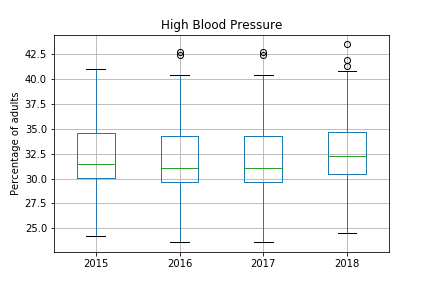
All the years have a bias to the right side and outliers are present in each of it. The year that might have the smallest variance is 2017.



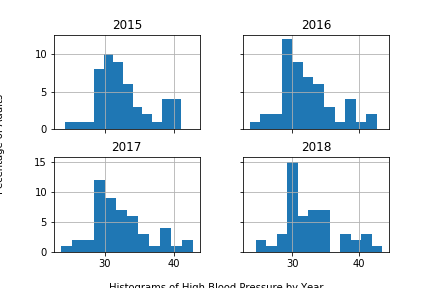


The year that has the biggest variance is 2016, followed by 2017 and 2018. The only year that presents an outlier is 2015.





In High Blood Pressure, it is clear that is the variable that has more outliers in compared to the rest the variables to be analyzed. All years have a bias to the right side.



**Is there any correlation between Income and Cardiovascular Deaths, Diabetes, Heart Attack, Obesity and High Blood Pressure?**

H0: The higher the Median Household Income, the lower the Cardiovascular Deaths.

Ha: The higher the Median Household Income, the higher the Cardiovascular Deaths.

H0: The higher the Median Household Income, the lower rate of Diabetes cases.

Ha: The higher the Median Household Income, the higher rate of Diabetes cases.

H0: The higher the Median Household Income, the lower rate of Heart Attack cases.

Ha: The higher the Median Household Income, the higher rate of Heart Attack cases.

H0: The higher the Median Household Income, the lower indexes of body mass.

Ha: The higher the Median Household Income, the higher indexes of body mass.

H0: The higher the Median Household Income, the lower rate of High Blood Pressure cases declared.

Ha: The higher the Median Household Income, the higher rate of High Blood Pressure cases declared.

**A screenshot of a cell phone

Description automatically generated**

**A close up of a map

Description automatically generated**

**A close up of a map

Description automatically generated**

**A close up of a map

Description automatically generated**

**A close up of a map

Description automatically generated**

**Is there any correlation between Obesity and Diabetes, Heart Attack and High Blood Pressure?**

H0: Obesity might increase the cases of Diabetes.

Ha: Obesity might not increase the cases of Diabetes.

H0: Obesity might increase the cases of Heart Attack.

Ha: Obesity might not increase the cases of Heart Attack.

H0: Obesity might increase the cases of High Blood Pressure.

Ha: Obesity might not increase the cases of High Blood Pressure.

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlation with Obesity** | | | |
| **Year** | **Diabetes** | **Heart Attack** | **High Blood Pressure** |
| **2015** | 0.655191802 | 0.660504672 | 0.73863101 |
| **2016** | 0.754042394 | 0.714703693 | 0.74607346 |
| **2017** | 0.69362846 | 0.674159094 | 0.777385131 |
| **2018** | 0.66079473 | 0.646125068 | 0.770837544 |

As we can see in the above table, there is a strong correlation between Obesity and Diabetes, Heart Attack and High Blood Pressure. Undoubtedly, health politics with regards to the reduction of obesity must be strength with the objective of reducing the indexes of body mass to prevent this type of diseases.

**Do the 6 variables to study have a normality distribution in order to be analyzed as it?**

H0= The data has a normal distribution.

Ha= The data does not have a normal distribution.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Median Household Income** | **Cardiovascular Death** | **Diabetes** | **Heart Attack** | **Obesity** | **High Blood Pressure** |
| **2015** | 0.444445 | 0.093089 | 0.679934 | 0.036112 | 0.753147 | 0.272607879 |
| **2016** | 0.668052 | 0.088783 | 0.361914 | 0.099508 | 0.637354 | 0.103498354 |
| **2017** | 0.568593 | 0.056833 | 0.693967 | 0.012007 | 0.730023 | 0.103498354 |
| **2018** | 0.634505 | 0.061685 | 0.596367 | 0.056738 | 0.48818 | 0.091950114 |

The variable Heart Attack in the years 2015 and 2017 reject the null hypothesis with p-values of 0.036112 and 0.012007, respectively and with a significance level of 0.05. The rest of the variable accept the null hypothesis with a significance level of 0.05.

In the case of the raw data that rejected the null hypothesis, deep statistical analysis should be done in order to get a normal distribution. Data transformation with Box Transformation is an option.