

Report

To answer the question that why there are missing data, I reviewed the Rate.csv and found that some states may not have data for that year. For instance, for Idaho, there is no data in 2016. Moreover, some data can be confidential, which means that we may not be able to access this data on the table.

In order to find a factor that may be related to the insurance rate, I compared the data with stroke and cancer and drew a conclusion that the median insurance rate does not have a relationship with stroke death rates. However, there may be some positive correlation between medium insurance rates and cancer cases.

For stroke, the rho is -0.249959, which is close to zero. Moreover, since the p-value is 0.1301 and is greater than 0.05, we cannot reject the null hypothesis that is the true rho is equal to 0. The conclusion of the test allows me to conclude that there is no relationship between median insurance rates and stroke death rates.

For the case of cancer, the rho is 0.3253091 and is close to zero. However, the p-value is 0.04679, which is less than 0.05, we can reject the null hypothesis that the true rho is equal to 0. There is enough evidence to conclude that the rho is not equal to 0, and there is a positive correlation between medium insurance rates and cancer cases.