**Article:**

Census Income-

This data was extracted from the 1994 Census bureau database by Ronny Kohavi and Barry Becker (Data Mining and Visualization, Silicon Graphics). A set of reasonably clean records was extracted using the following conditions: ((AAGE>16) && (AGI>100) && (AFNLWGT>1) && (HRSWK>0)). The prediction task is to determine whether a person makes over $50K a year.

Description of fnlwgt (final weight)

The weights on the Current Population Survey (CPS) files are controlled to independent estimates of the civilian non-institutional population of the US. These are prepared monthly for us by Population Division here at the Census Bureau. We use 3 sets of controls. These are:

* A single cell estimate of the population 16+ for each state.
* Controls for Hispanic Origin by age and sex.
* Controls by Race, age and sex.

By Data Analysis we can see that after checking the null values we can remove the nulls and we find the result as 0.we can also get the category and variable lists.

By univariate analysis we can get the population distribution, by bivariate analysis we can found the number of count of education by income , by sex , by workclass.

By heatmap we can found out the the precent income distribution by martial status.

Statistical analysis gives an idea of the stat values from age to income.

By checking the skewness we can see the optimal range fom-0.5 to 0.5.

Conclusion:-

It provides data about the recent trend in the population and the situation of the country or any place. The census of India is one of the largest administrative exercises undertaken in the world.