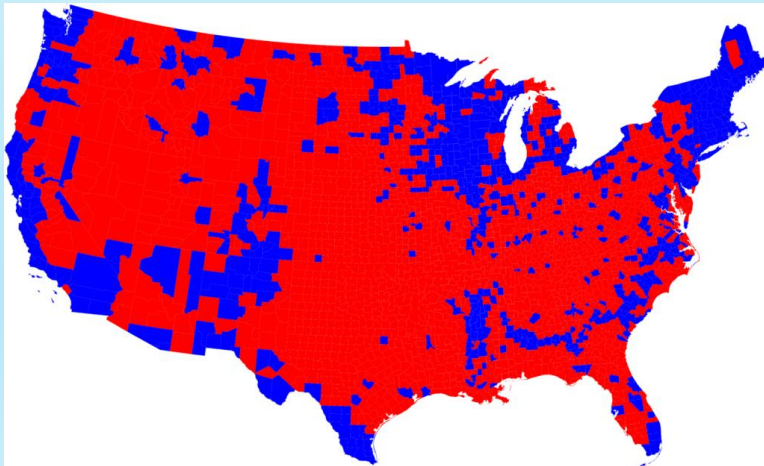


USA STATISTICS



Group 03:

Jayalakshmi Vaidyanathan

Krutika Ambavane

Neha Narayankar

Dataset On USA Statistics

- The dataset is composed of homeowners financial, family and population related information from 2012-2016

Important coulmnns in the dataset are :

- hi_mean: Household Income
- debt: Percentage of home with some type of debt
- rent_mean: The mean gross rent of the specified graphic location
- type: The place type reported by the U.S. Census Bureau for the specified geographic location

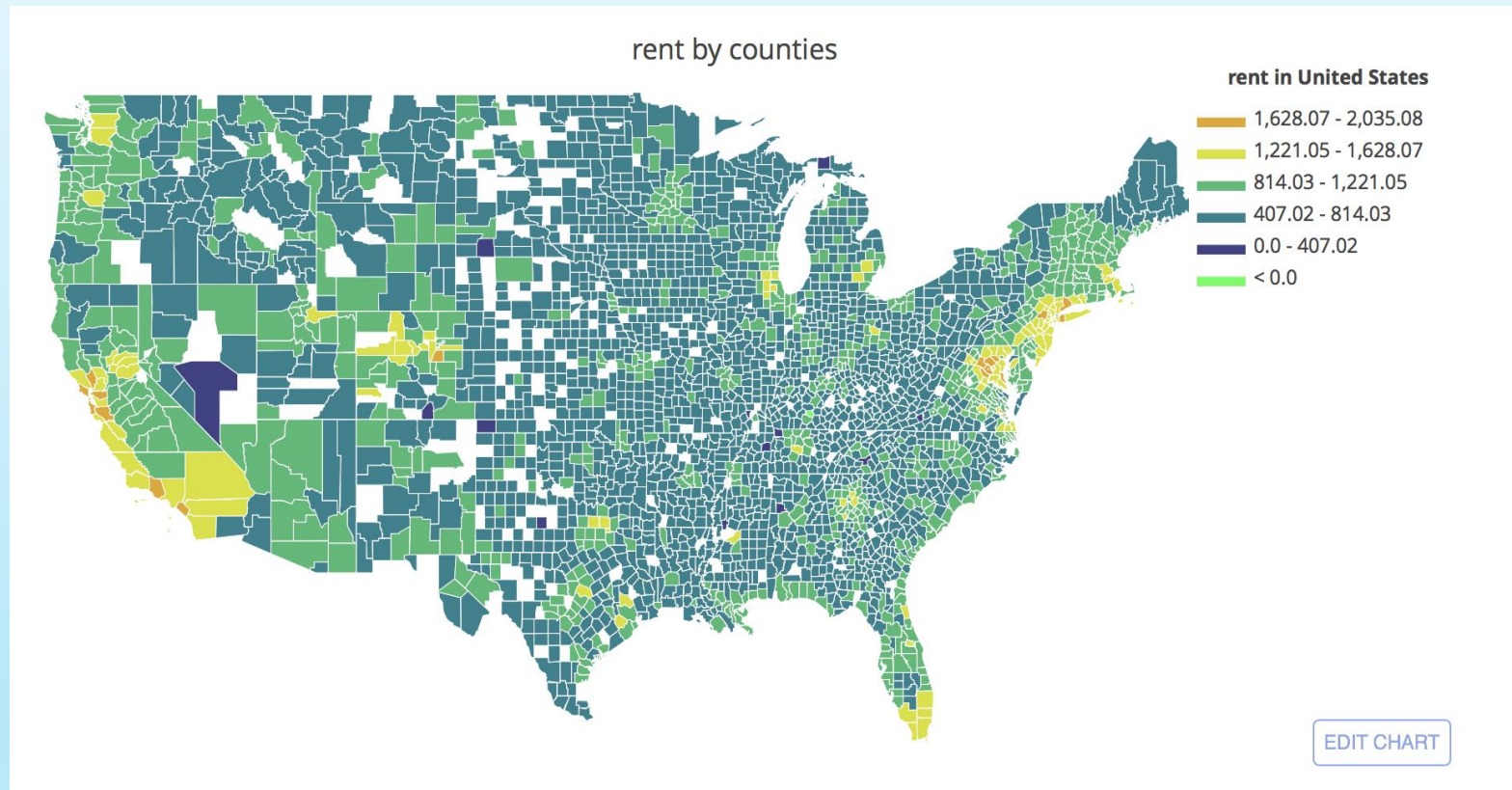
Data Cleansing

- Dropped unnecessary and redundant columns
- Handled NaN values
- Converted all the columns to lower case
- Used the binning concept
- Adjusted the precision of the values

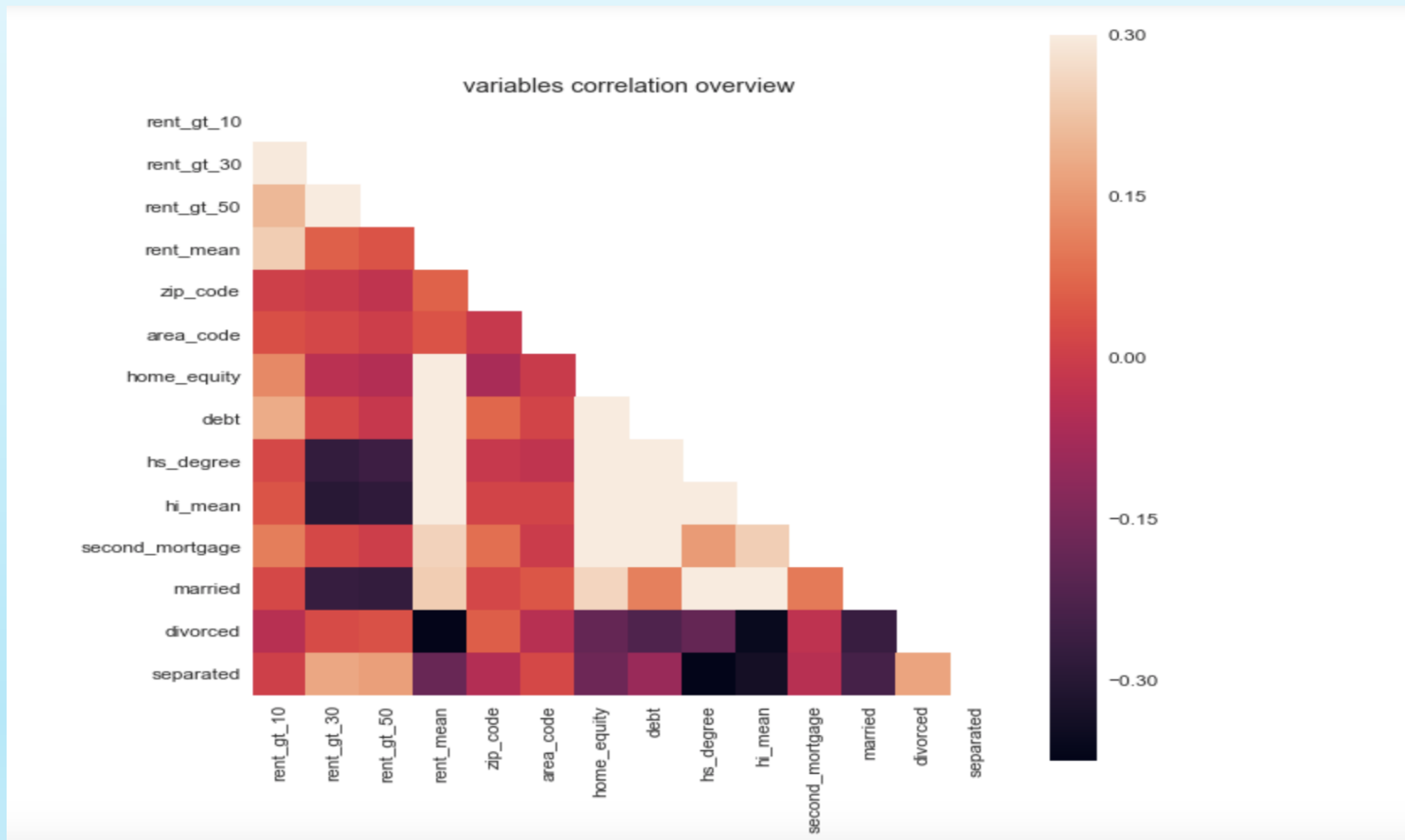
Questions To Analyze

- ▶ How is the rent variation across entire United States?
- ▶ What are the factors that affect the rent prices?
- ▶ What kind of people prefer higher rent?
- ▶ How are income, rent and debt related?
- ▶ What do people do to handle their debts?

US Rent Distribution



Heatmap : Variable Correlation



Jupyter Notebook