Name: Sarthak Gharat

UID: 2021700025

Branch: CSE DS

Expt 5: Housing Data

Problem Statement: Create advanced charts using R programming language on the dataset - Housing data

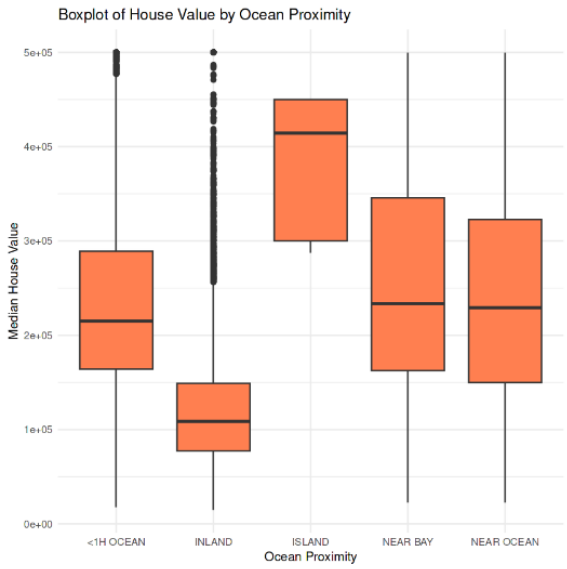
- Advanced - Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, Jitter
- Write observations from each chart

To explore and visualize housing data using advanced charts in R, including Word chart, Box and Whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, and Jitter plot, in order to uncover patterns and insights in the dataset.

Software Used: R
Dataset Used: California Housing Prices

1. Box and Whisker Plot

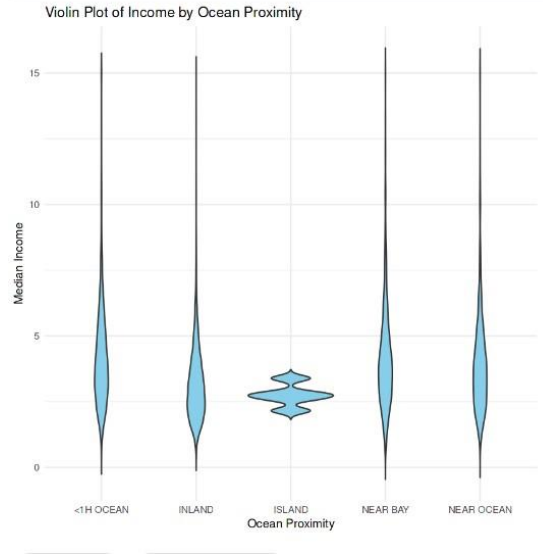
```
ggplot(data, aes(x=ocean_proximity, y=median_house_value)) +
  geom_boxplot(fill='coral') +
  labs(title='Boxplot of House Value by Ocean Proximity', x='Ocean Proximity', y='Median House Value') +
  theme_minimal()
```



Island houses are costliest, inland houses are cheaper.

2. Violin Plot

```
ggplot(data, aes(x=ocean_proximity, y=median_income)) +
  geom_violin(trim=FALSE, fill='skyblue') +
  labs(title='Violin Plot of Income by Ocean Proximity', x='Ocean Proximity', y='Median Income') +
  theme_minimal()
```

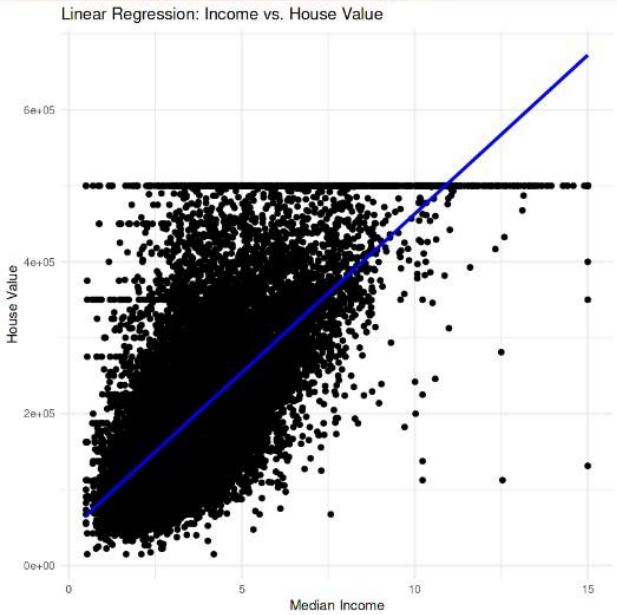


Same as above, but with density distribution.

3. Regression Plot (Linear)

```
ggplot(data, aes(x=median_income, y=median_house_value)) +
  geom_point() +
  geom_smooth(method='lm', se=FALSE, color='blue') +
  labs(title='Linear Regression: Income vs. House Value', x='Median Income', y='House Value') +
  theme_minimal()
```

```
`geom_smooth()` using formula = 'y ~ x'
```

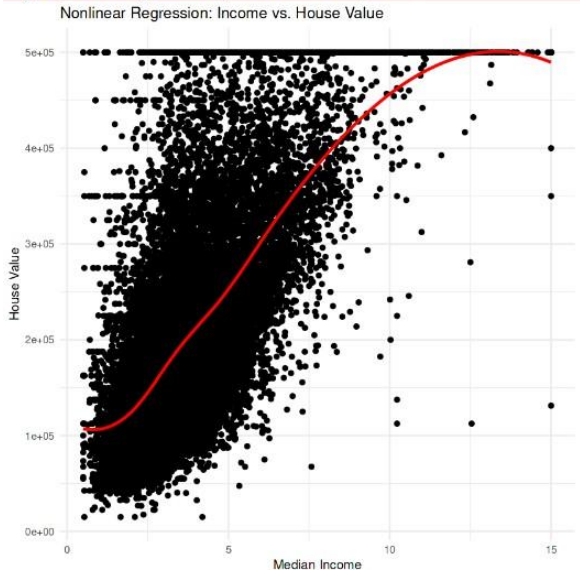


Median Income and House value dont have a perfect linear relationship.

4. Regression Plot (Non-Linear)

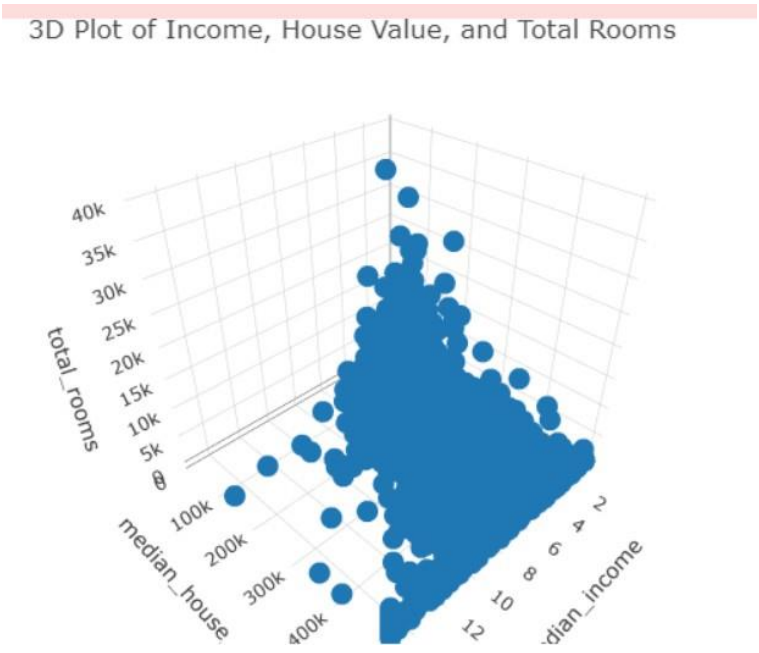
```
    theme_minimal()
```

```
`geom_smooth()` using formula = 'y ~ x'
```



The house value gets saturated with median income.

5. 3D-Chart

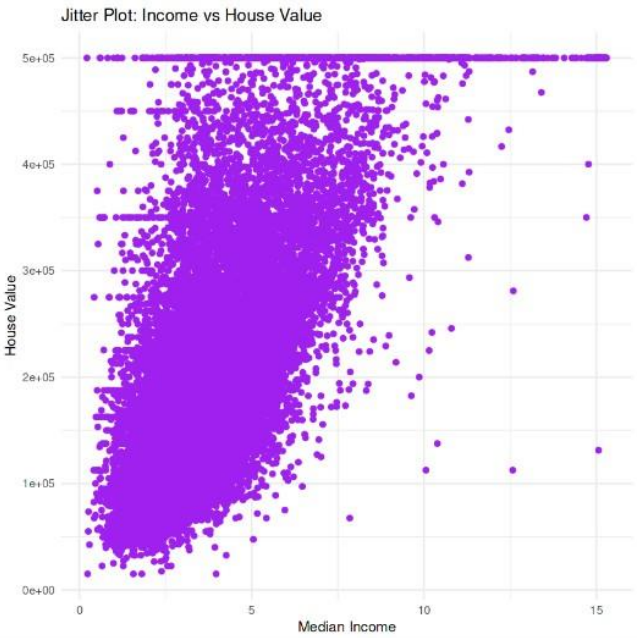


Generally, more median income implies higher median house value and more number of rooms.

6. Jitter Chart

[9]:

```
ggplot(data, aes(x=median_income, y=median_house_value)) +
  geom_jitter(color='purple', width=0.3, height=0.3) +
  labs(title='Jitter Plot: Income vs House Value', x='Median Income', y='House Value') +
  theme_minimal()
```



Same as scatter plot, but with random noise.
