Lec 02: In-class Exercise: Scatterplots

WRITE YOUR NAME HERE

Today's Dataset

This dataset comes from Pioneer Valley Data and documents estimates of population characteristics for each municipality in the Pioneer Valley.

Step 1: Load Packages

```
# Load the ggplot library
```

Step 2: Read the Data

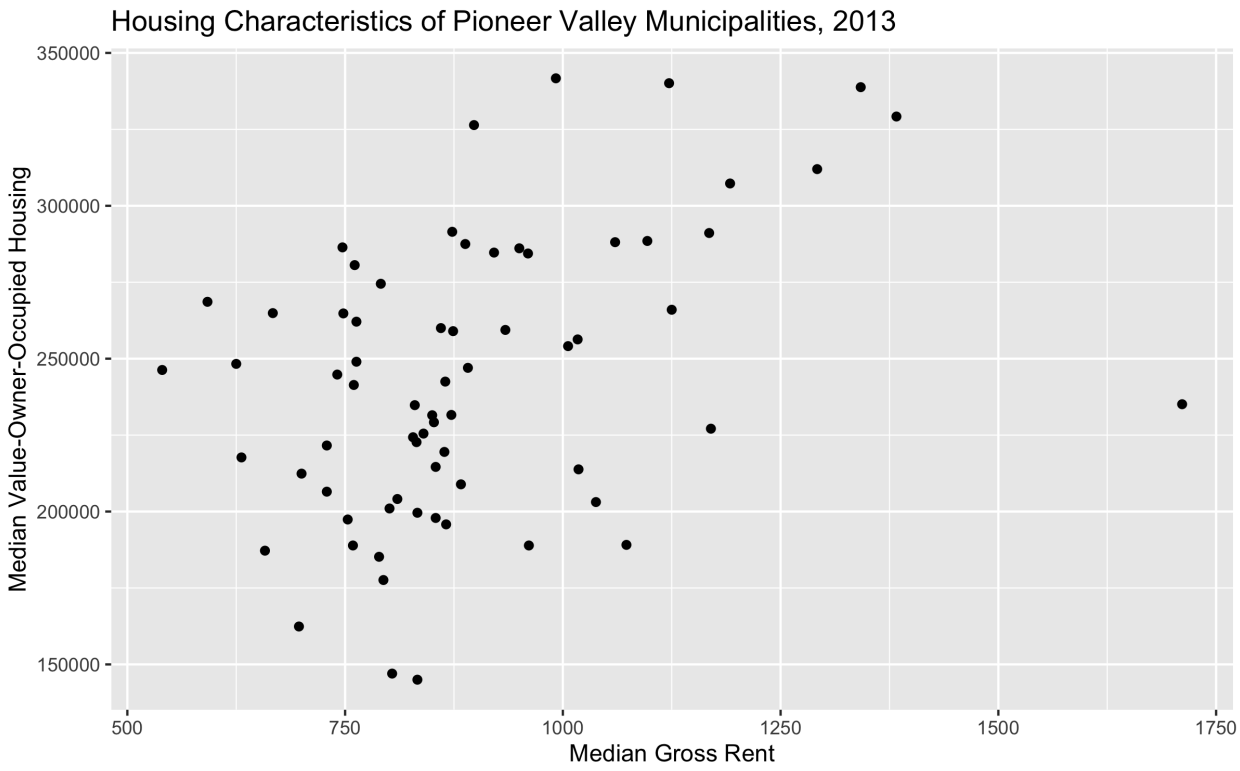
```
pioneer_valley_2013 <- read.csv("https://raw.githubusercontent.com/SDS-192-Intro/sds-192-labs/main/Day7
pioneer_valley_2013_dictionary <- read.csv("https://raw.githubusercontent.com/SDS-192-Intro/sds-192-labs
```

Step 3: View the Data

```
# Write your code below to check the column names for pioneer_valley_2013
```

Step 4: Recreate This Image Using the `ggplot()` Function

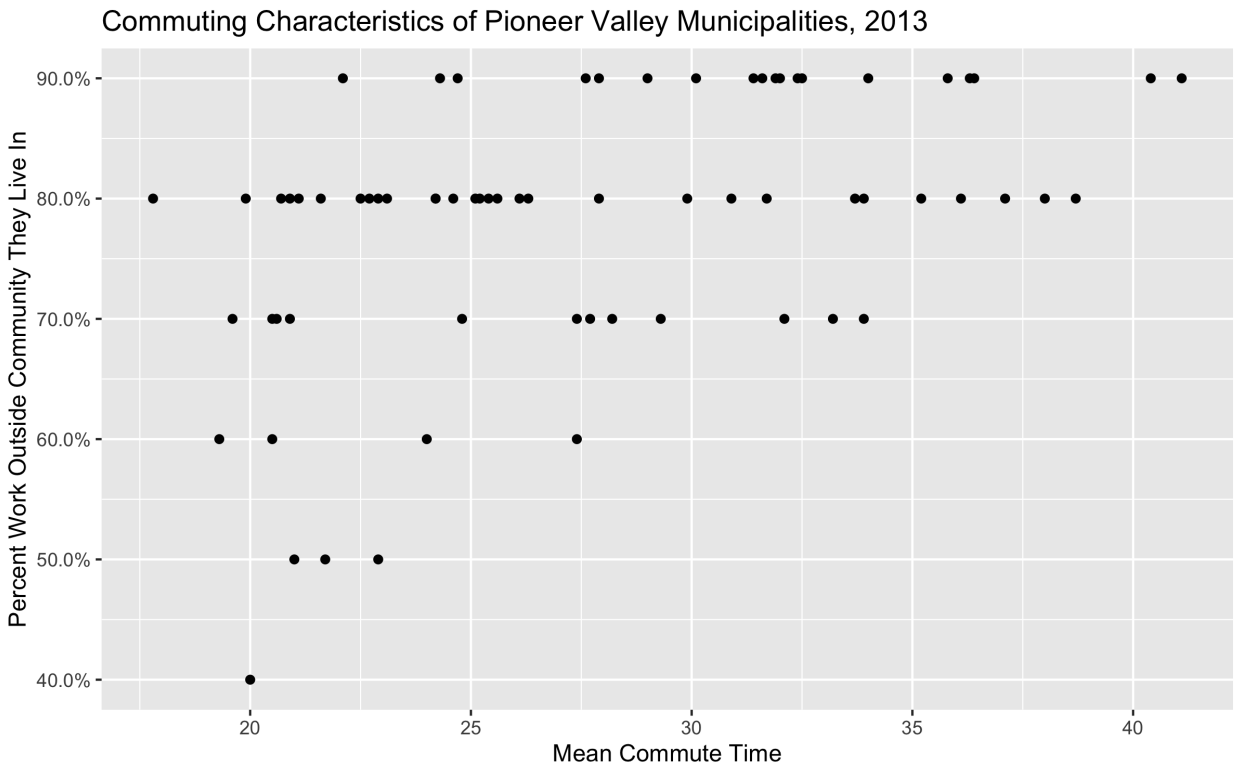
(Full size image in your downloaded folder/img folder.)




```
# Write your code below
```

Step 5: Recreate This Image Using the `ggplot()` Function

(Full size image in your downloaded folder/img folder.)



Hint: We need the function scale_y_continuous(), and the labels argument needs to be set to `scales::percent`. Check the help pages for this function to see how to format this!

```
# Write your code below
```

Discussion: Which of the following does each point on this plot indicate?

(1) A person
(2) A municipality
(3) A commute
(4) A county