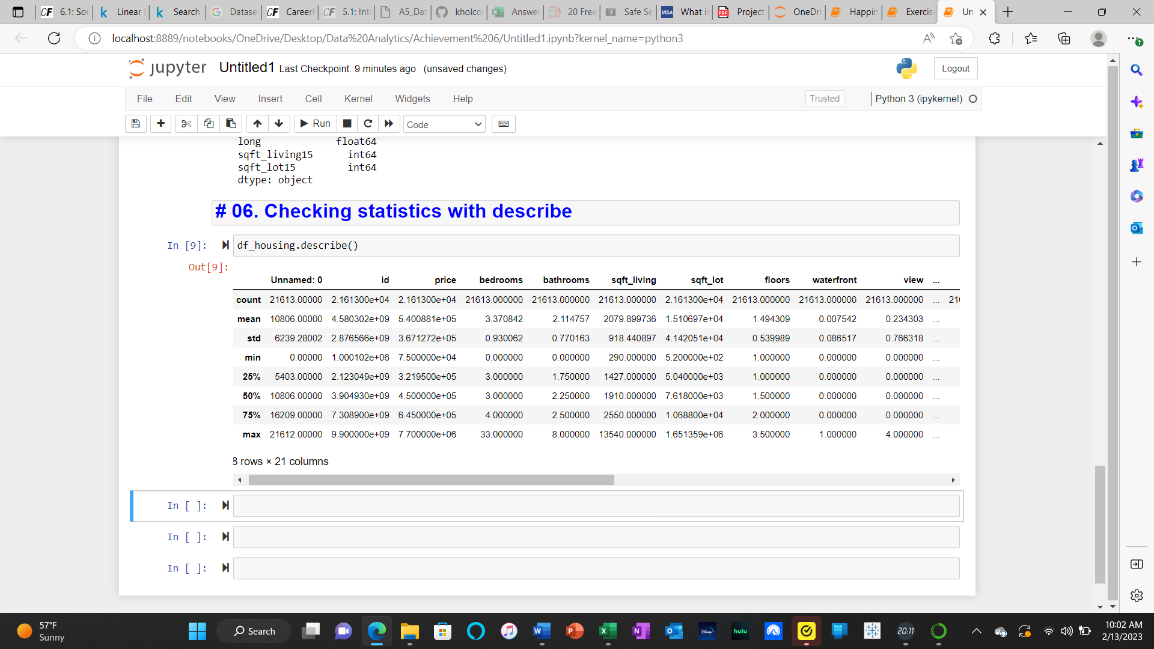
1. **Data Source**
   1. The dataset collected contains information about the King County, Washington housing market. This dataset in particular originates from the UCI Machine Learning Repository, so it is reliable. The dataset contains 22 columns and 21613 rows. Most columns are self-explanatory, but columns like ‘view’, ‘condition’, and ‘grade’ have scales. View is an index from 0-4 of how good the view of the property is. Condition is an index from 1-5 on the condition of the apartment, and grade is an index from 1-13, with 1 being the lowest and 13 being the highest level of construction and design. This variable is somewhat ambiguous, so it may be removed from the dataset for analysis.
2. **Data Profile**
   1. I took the data into Jupiter and conducted basic data cleaning and cleansing. No column names needed changing as they are self-explanatory. No missing values or duplicate entries were found. I did a check on the datatypes as well and none needed to be altered. I calculated basic statistics on each column with the describe feature.
   2. ****
   3. **A screenshot of a computer

      Description automatically generated**
   4. The contents of the data may pose a threat to security. It contains the geographical location of the homes and zipcodes, and as this data is easily accessible, anyone could find out where these homes are. But, there is no personally identifiable information so there is no information on the individual who owns the home.
3. **Questions to Explore**
   1. What factors contribute to the higher priced homes?
   2. In what areas of King County do the homes have the best views?
   3. What area of King County has the most expensively priced homes?