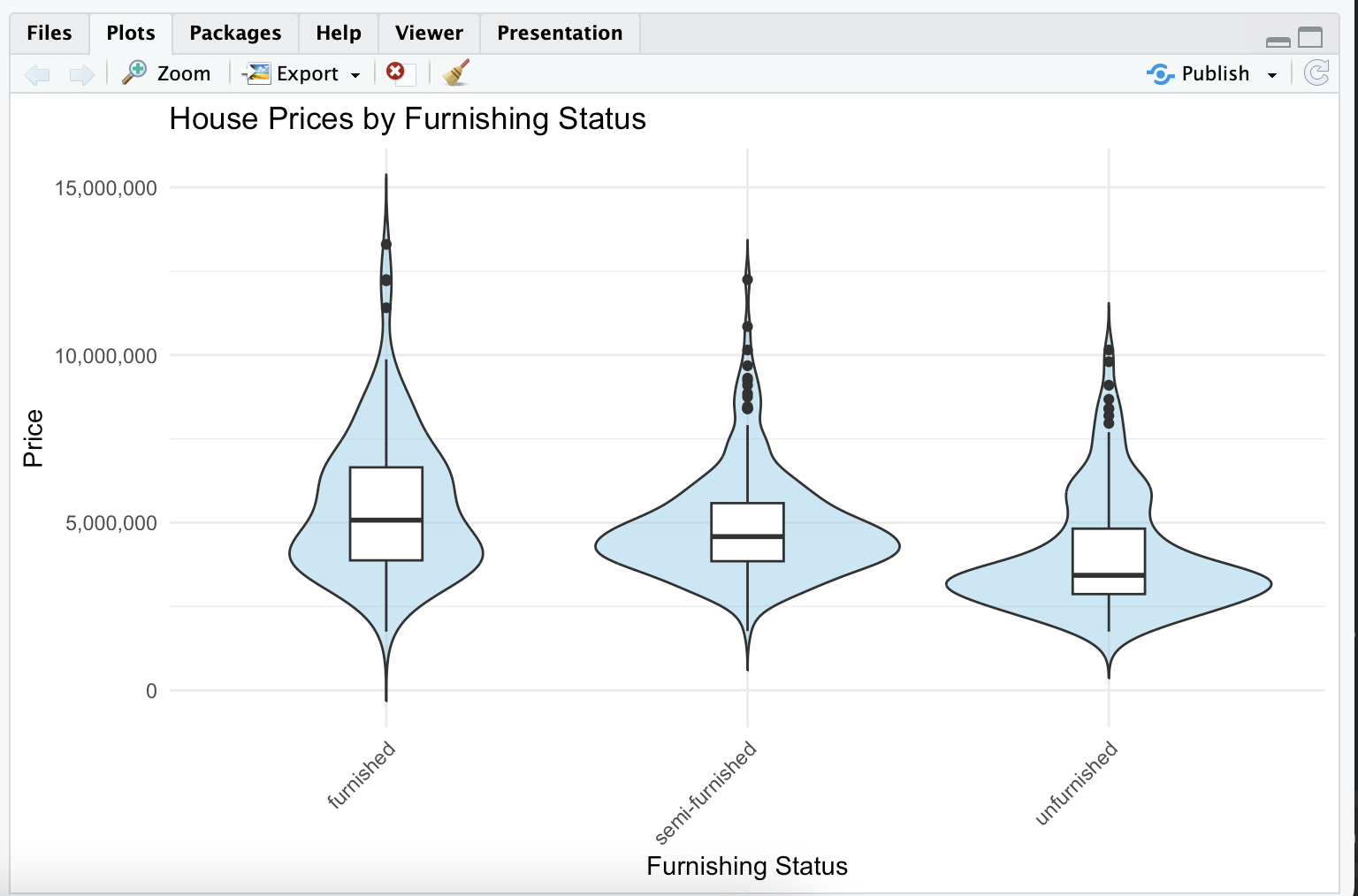
| Name: | Deepika Trivedi |
| --- | --- |
| UID: | 2021700069 |
| Experiment Number: | 5 |

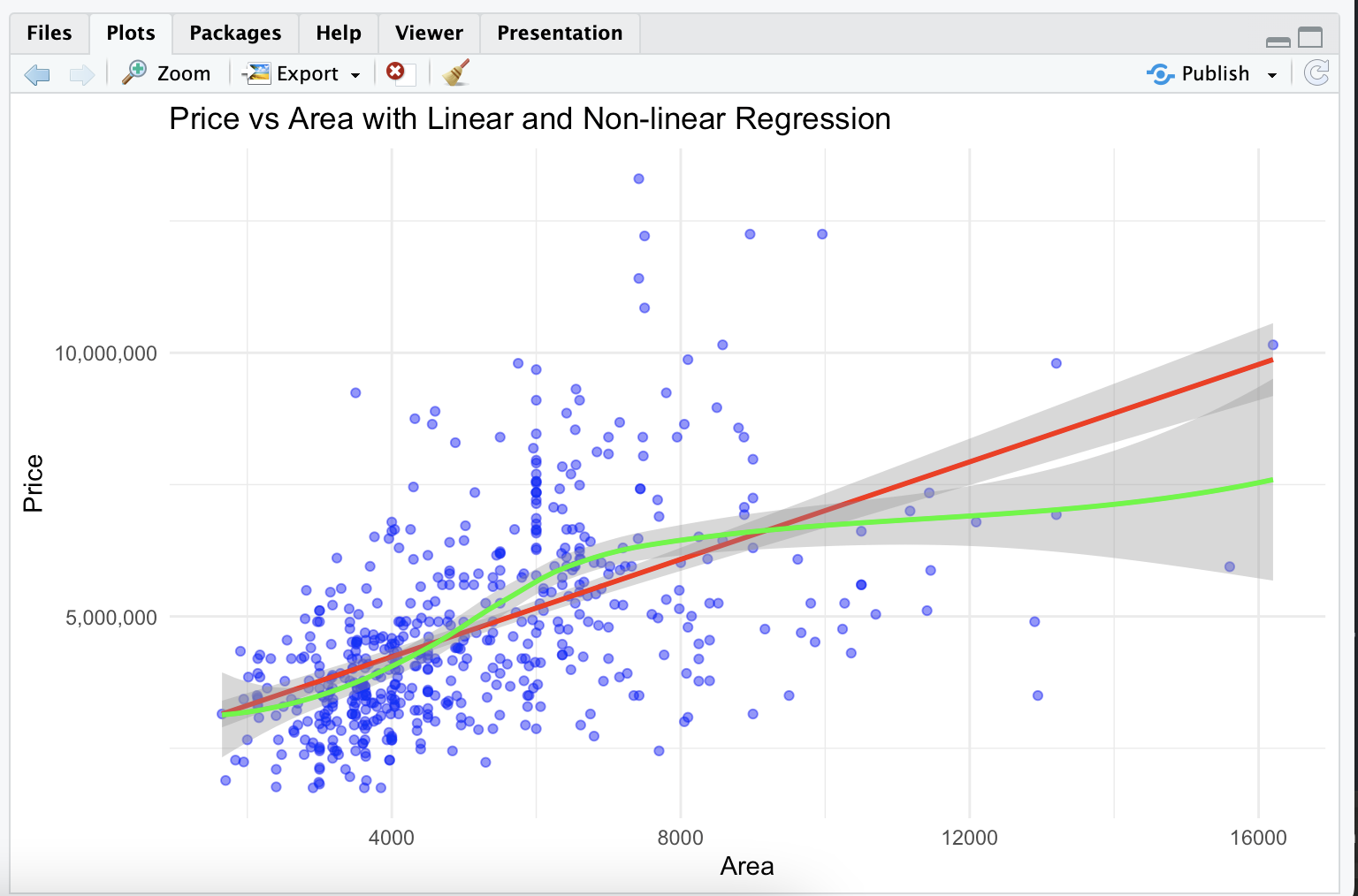
1. **Box and Violin Plot Combination:**

* Shows price distribution across furnishing status
* Violin shape shows density of data points
* Box plot shows median, quartiles, and outliers



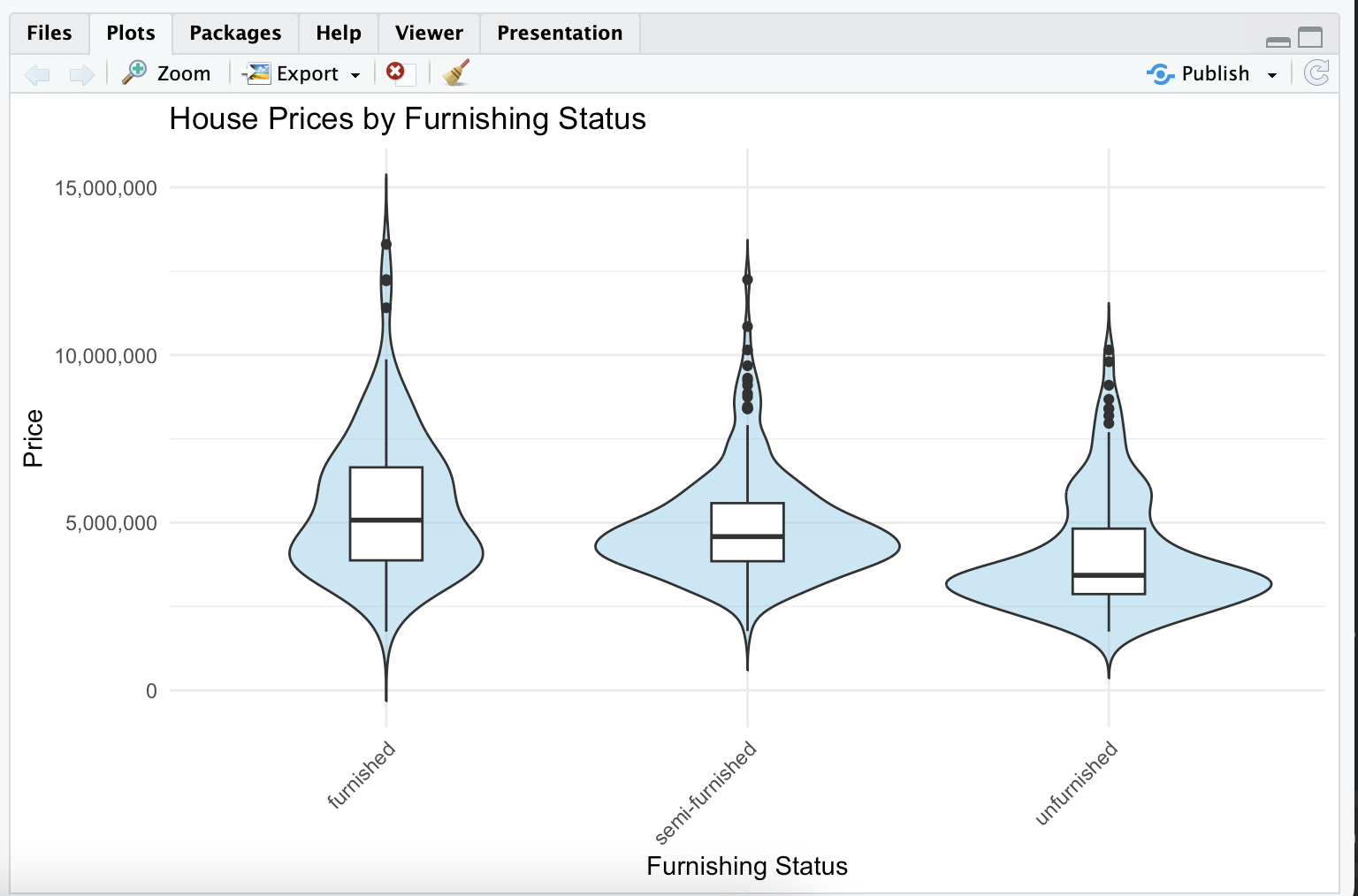
1. **Regression Plot:**

* **Shows relationship between price and area**
* **Red line: linear regression**
* **Green line: non-linear regression (LOESS)**
* **Helps identify if relationship is linear or nonlinear**



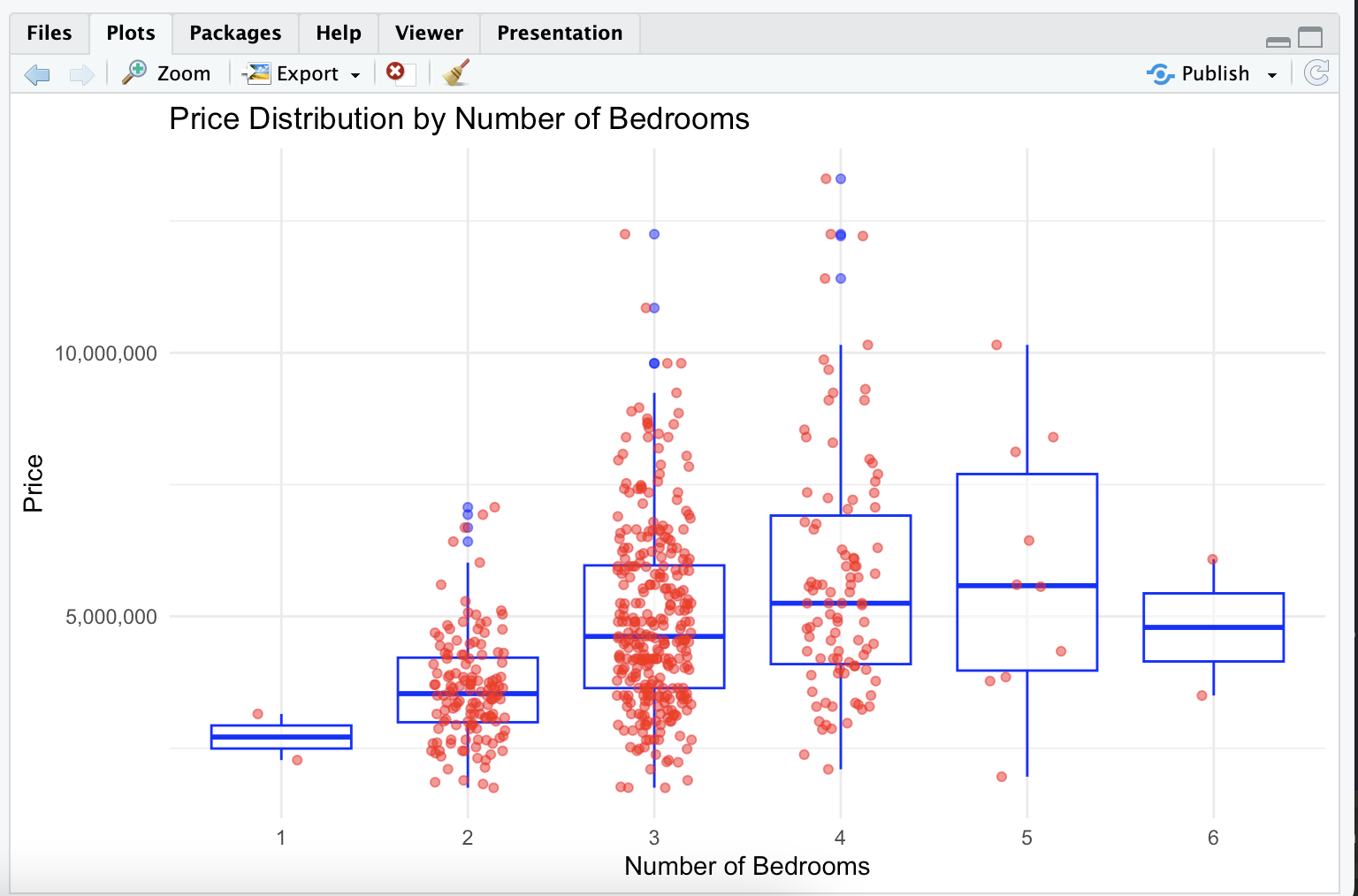
1. House Prices by Furnishing Status (Violin Plot):

* Three categories: furnished, semi-furnished, and unfurnished
* Furnished properties tend to have higher median prices
* All categories show similar price ranges but different distributions
* Furnished properties show a wider distribution in the middle price range
* Unfurnished properties have the lowest median price
* All categories have some high-price outliers



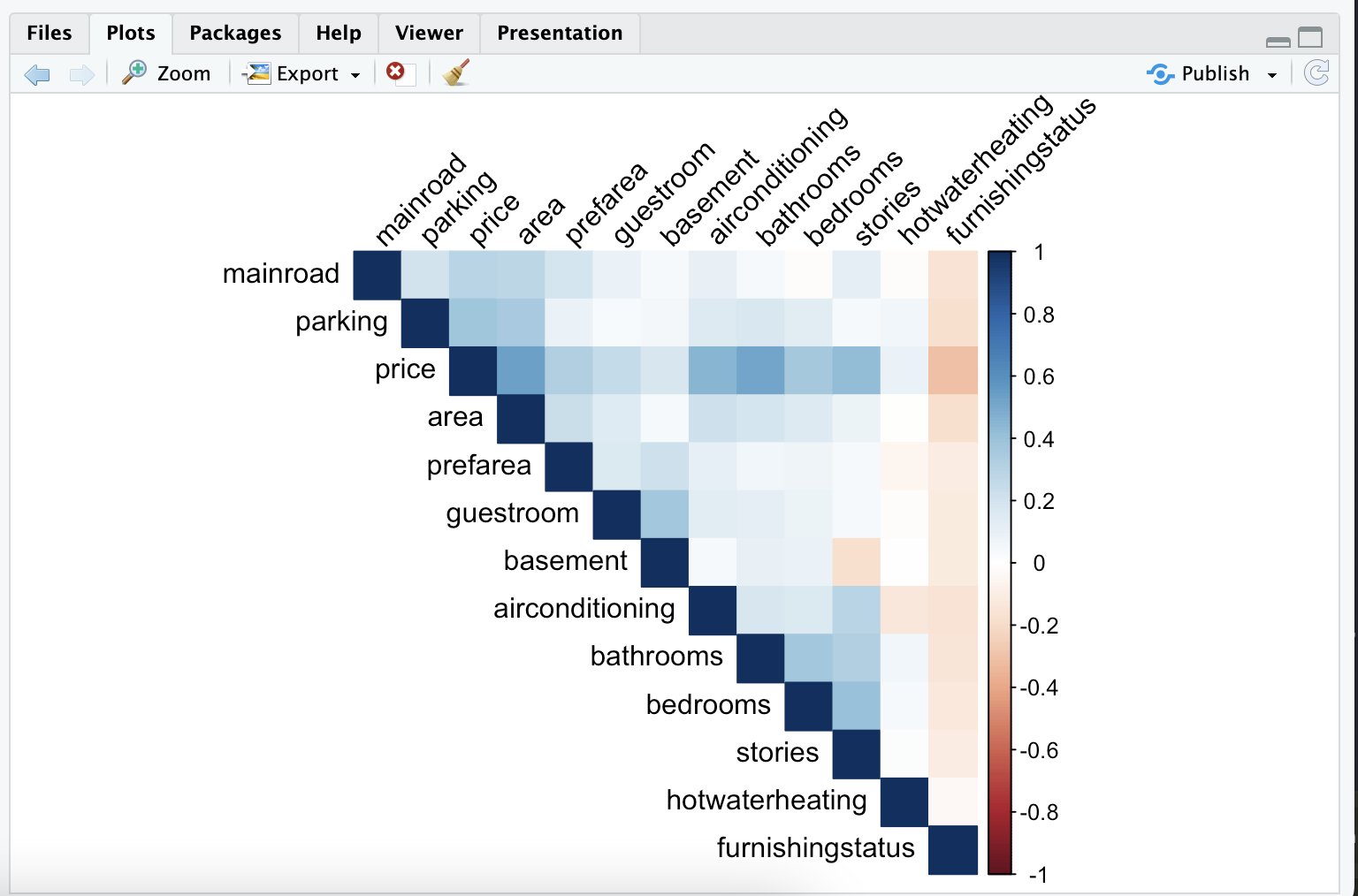
1. Price Distribution by Number of Bedrooms (Box Plot):

* There's a clear positive correlation between number of bedrooms and price
* The median price increases with each additional bedroom up to 5 bedrooms
* Price ranges (variability) also tend to increase with more bedrooms
* Most properties in the dataset have 2-4 bedrooms
* There are numerous outliers, especially in the 3-4 bedroom categories
* 1-bedroom properties show the least price variation
* 6-bedroom properties show a slight decrease in median price compared to 5-bedroom ones



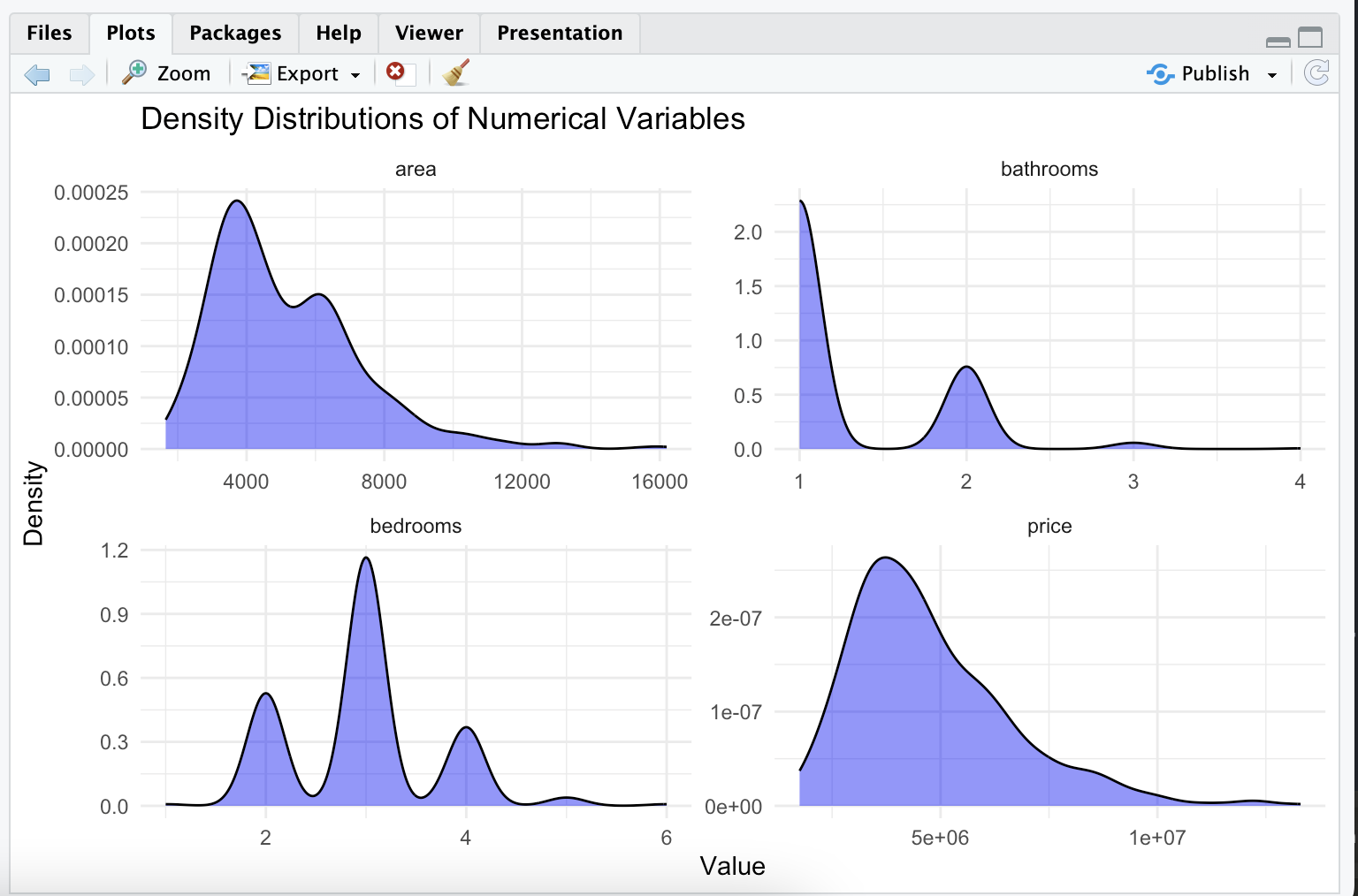
1. Correlation Heatmap:

* Price shows strong positive correlations with:
  + Area
  + Number of bedrooms
  + Number of bathrooms
* Moderate correlations exist between:
  + Area and number of bedrooms
  + Bathrooms and bedrooms
* Some features like hot water heating and furnishing status show weaker correlations with price
* Parking and mainroad show relatively weak correlations with other features



1. Density Distributions:

* Area: Right-skewed distribution with main peak around 4000 sq ft and a smaller peak around 6000 sq ft
* Bathrooms: Bimodal distribution with peaks at 1 and 2 bathrooms
* Bedrooms: Trimodal distribution with major peaks at 2, 3, and 4 bedrooms
* Price: Right-skewed distribution with most properties clustered in the lower price range



1. Price vs Area Scatter Plot with Marginal Distributions:

* Strong positive correlation between price and area
* Most properties are clustered in the 2000-8000 sq ft range
* Price variability increases with larger areas
* The relationship appears somewhat linear but with increasing scatter at higher values
* Both marginal distributions are right-skewed

