THEORY ACTIVITY 1

DATASET: HOUSE PRICE

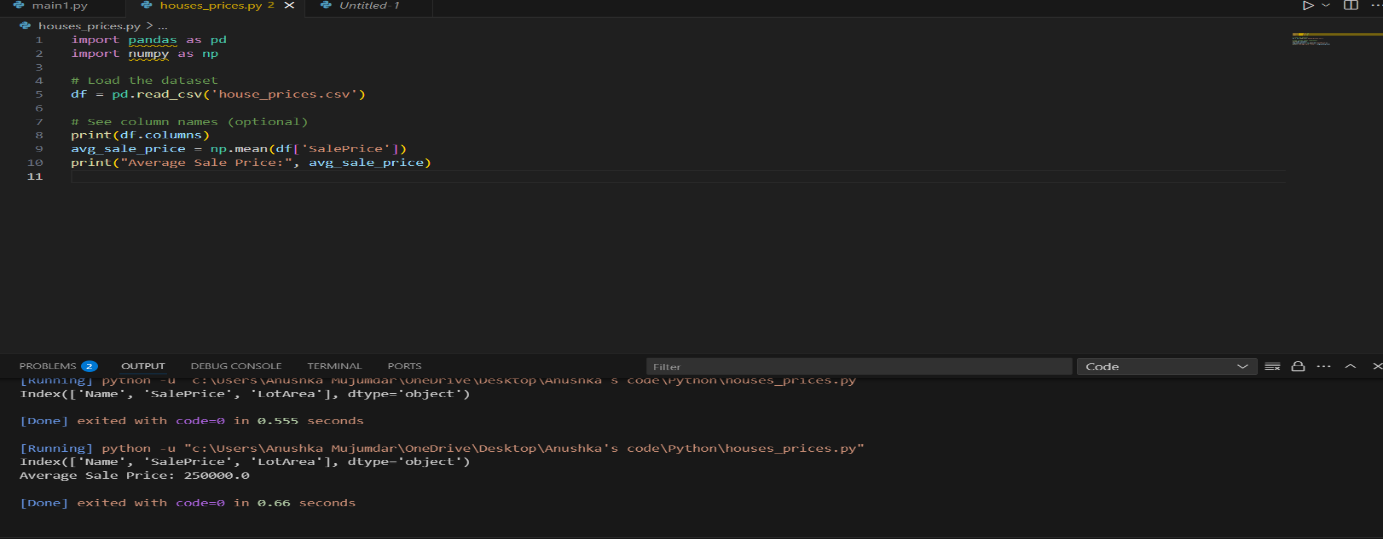
NAME: ANUSHKA MUJUMDAR

ROLL NO.: CC3-68

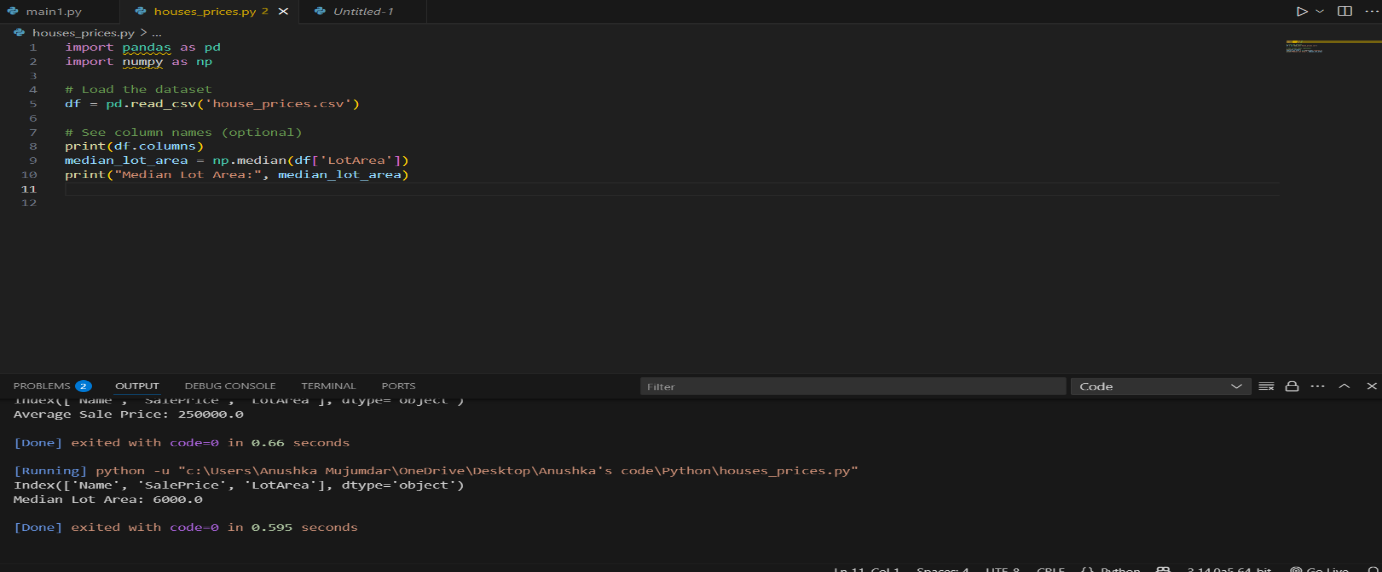
PRN: 202401050007

PROBLEM STATEMENTS

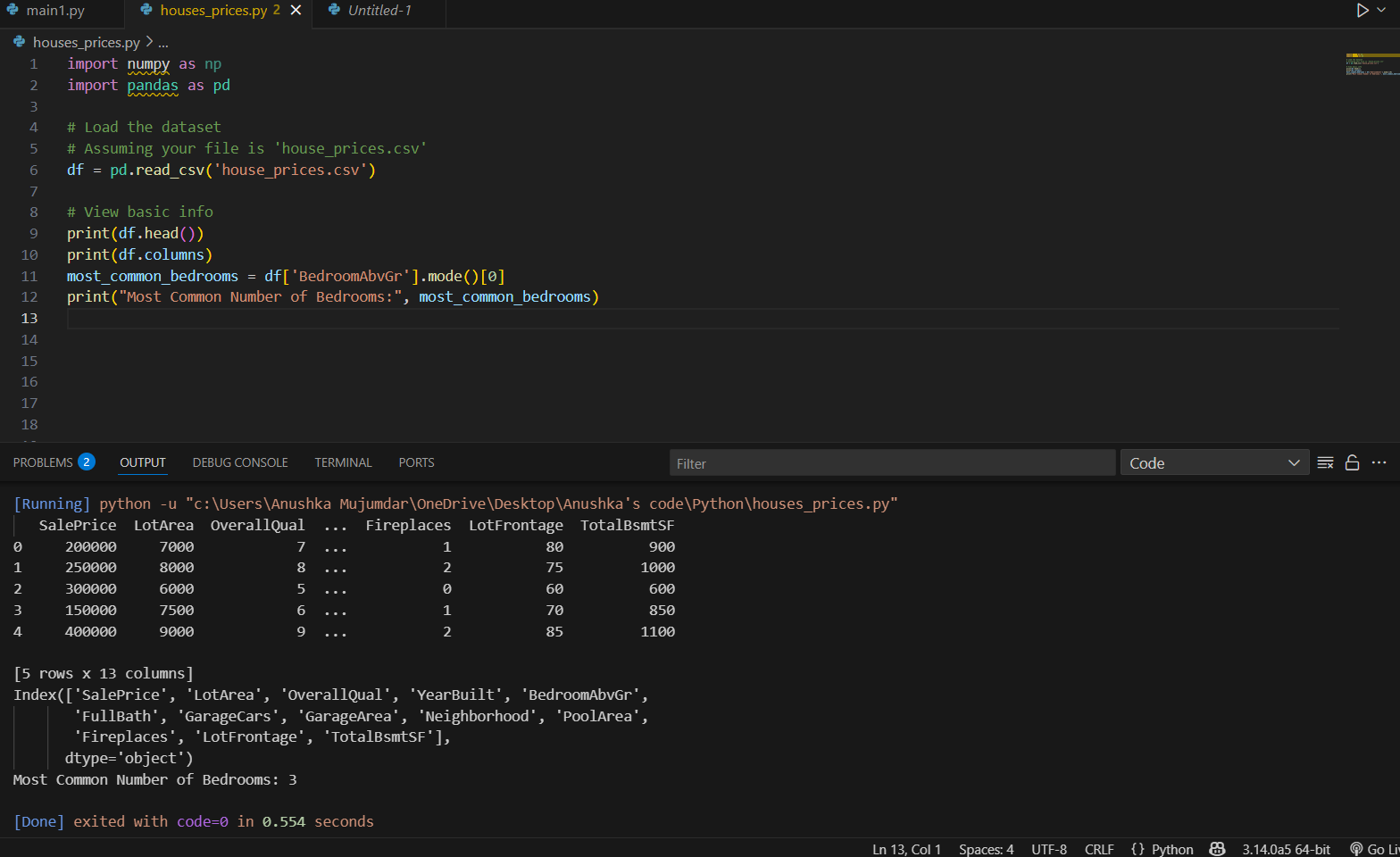
1. Find the average (mean) house price.



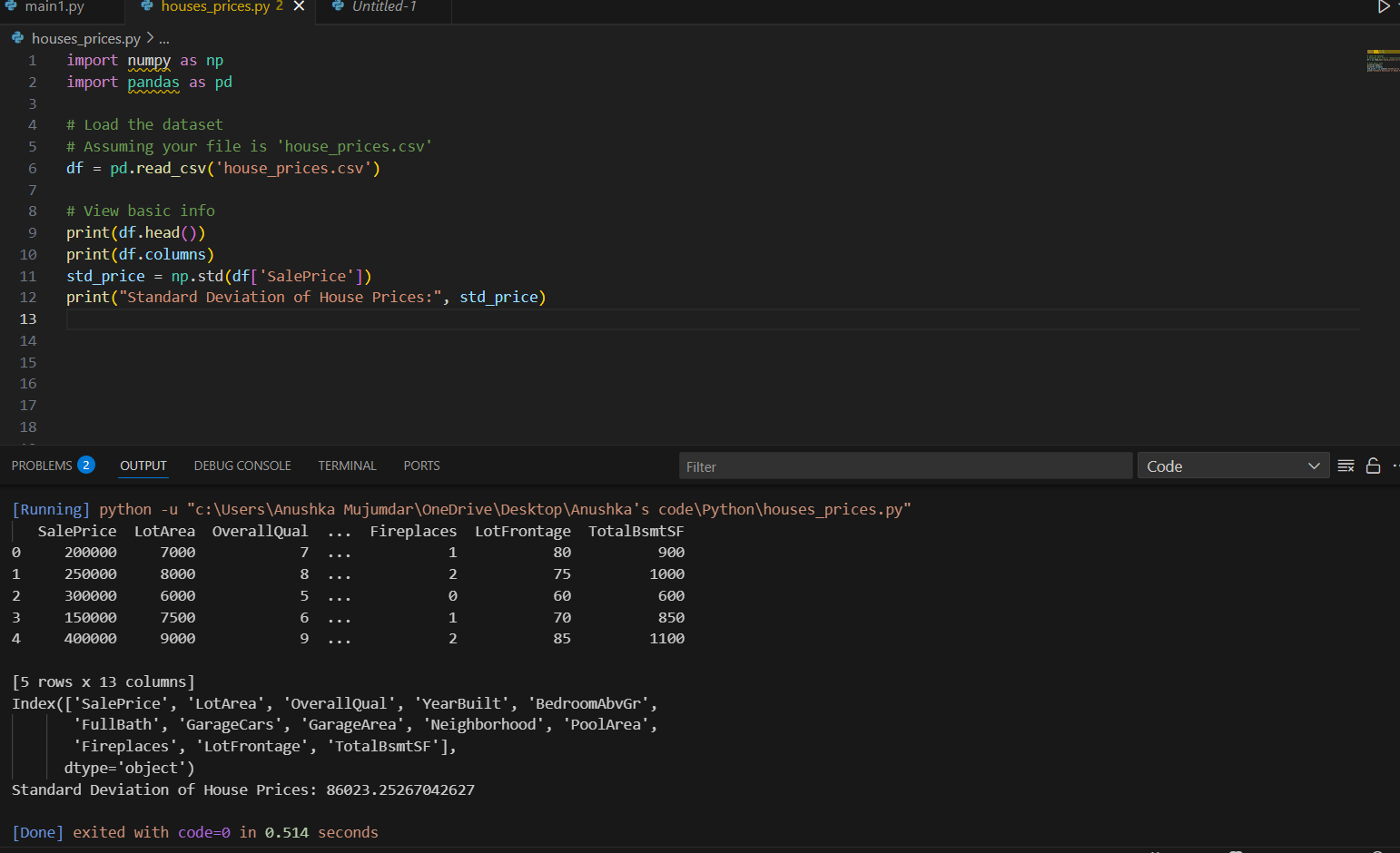
1. Find the median lot area.



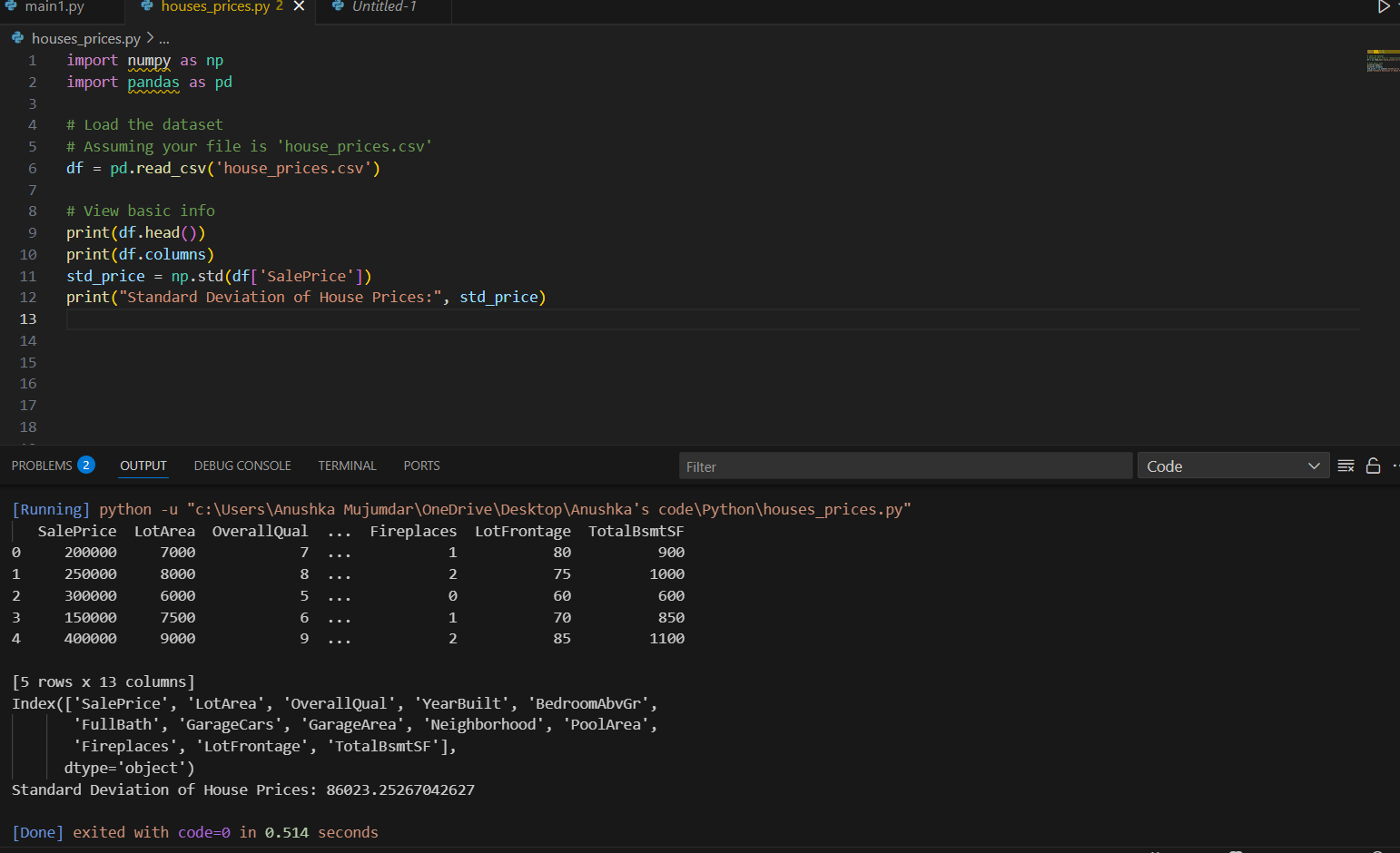
1. Find the maximum living area (GrLivArea) of the houses.



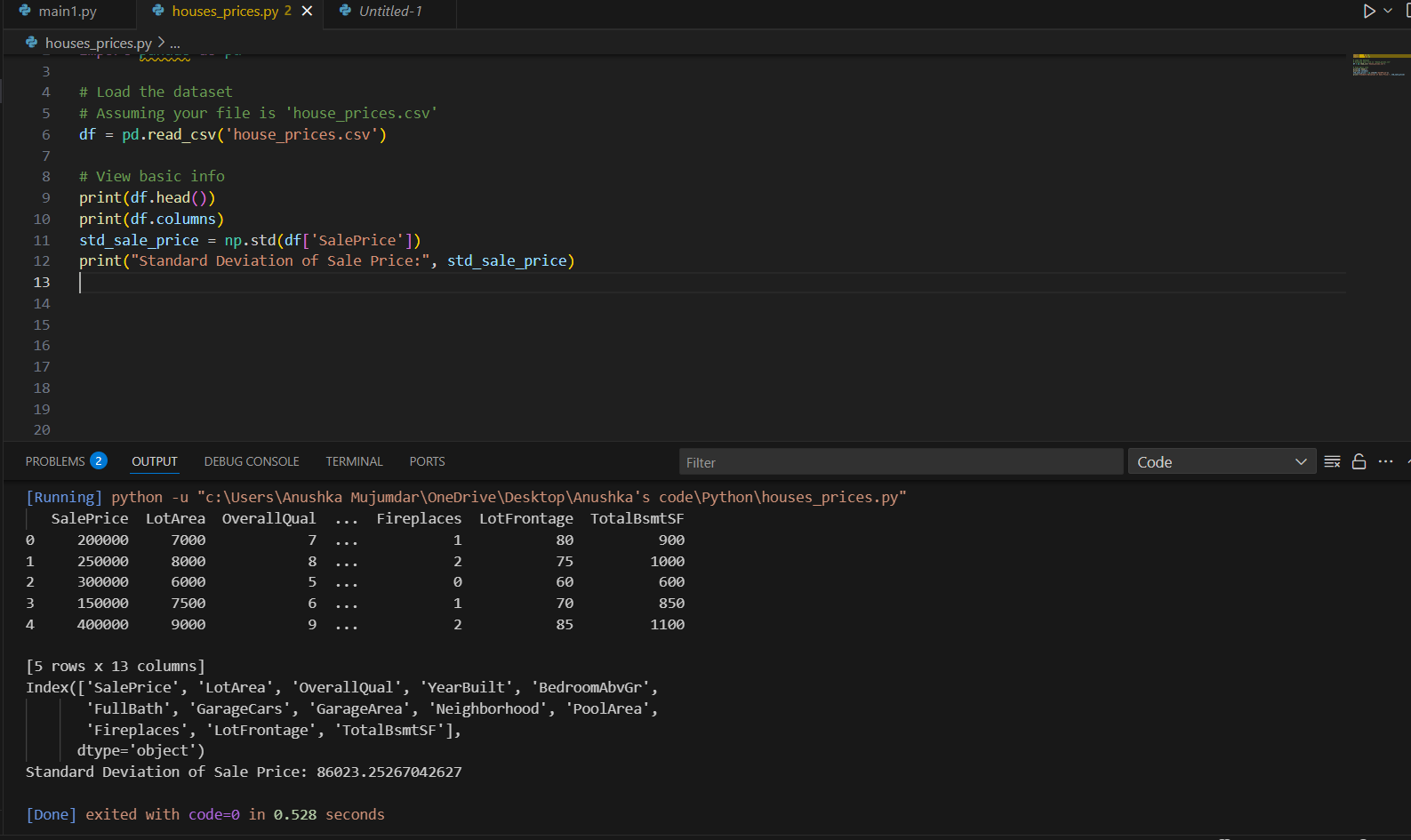
1. Find the minimum number of bedrooms.



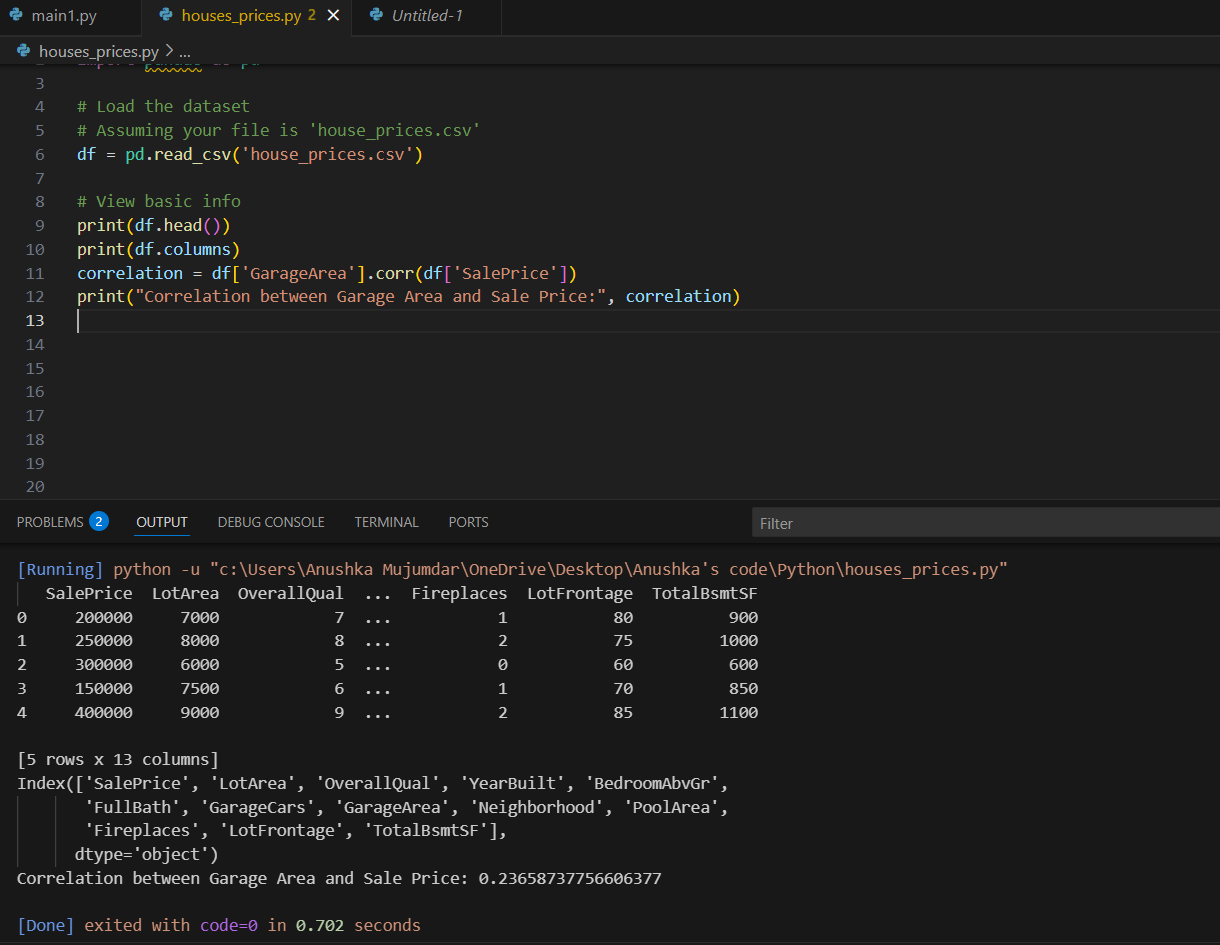
1. **How many houses have more than 4 bedrooms?**



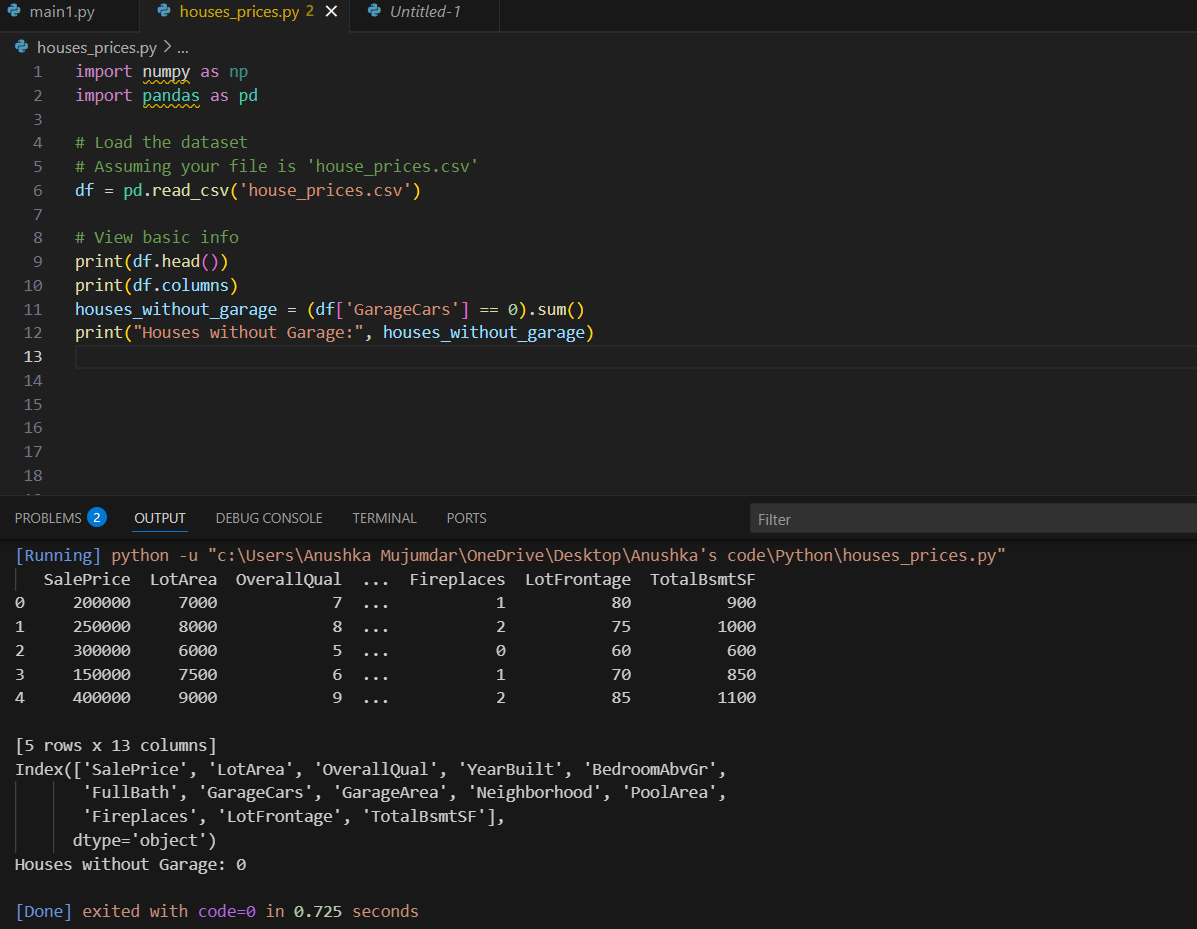
1. **Find the standard deviation of house sale prices.**



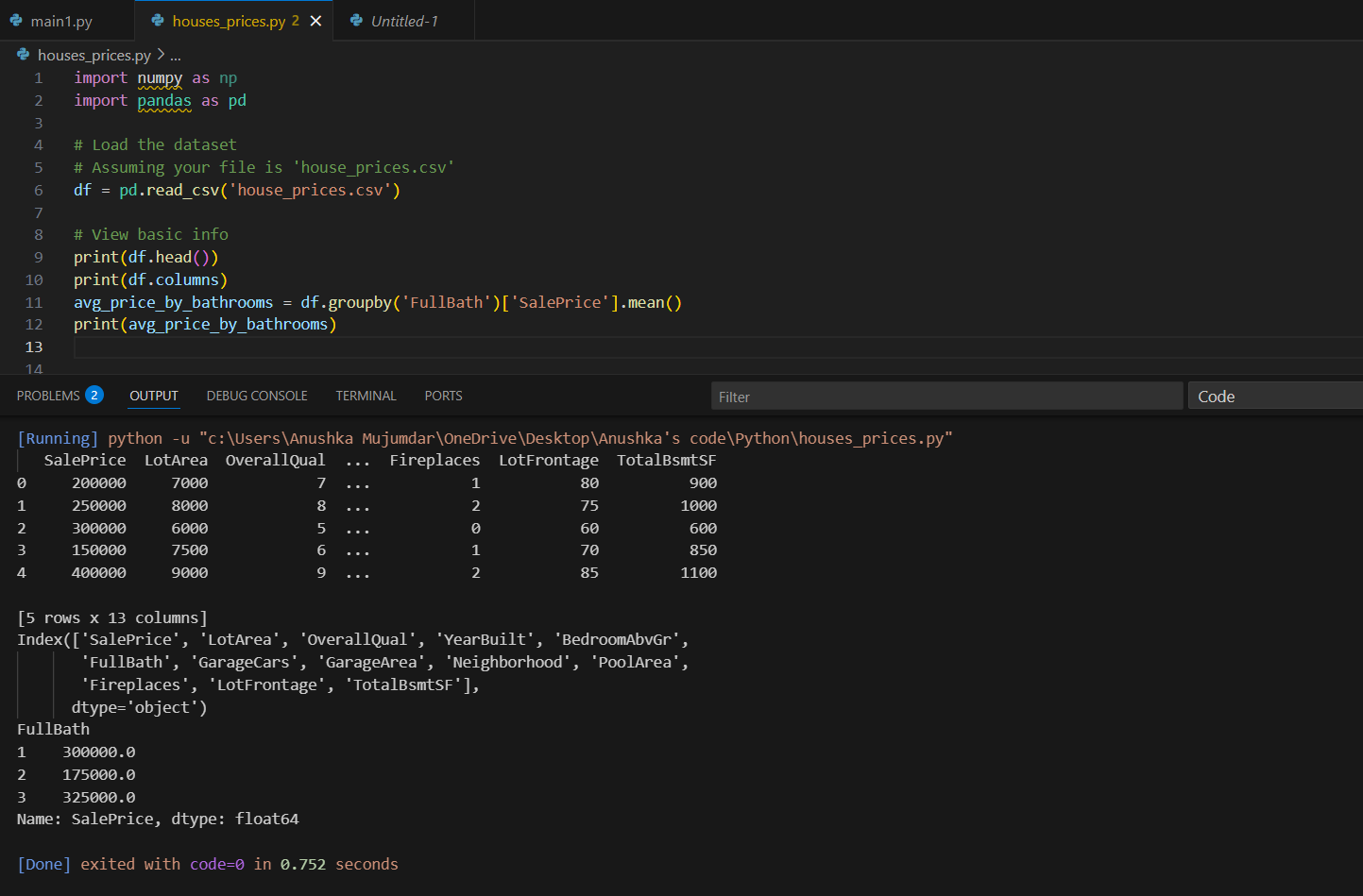
1. **Find the correlation between garage area and sale price.**



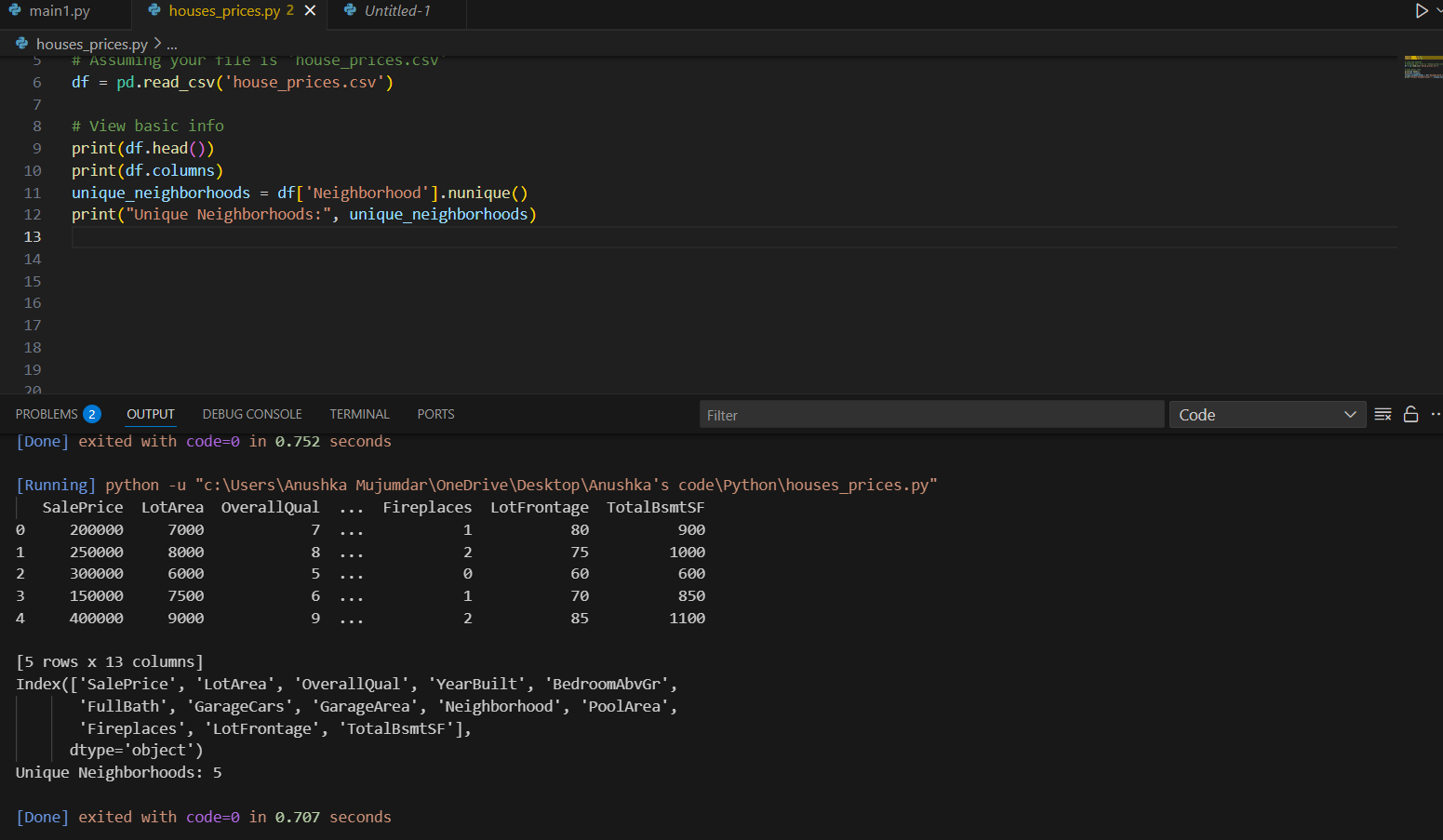
1. **Find the number of houses without a garage.**



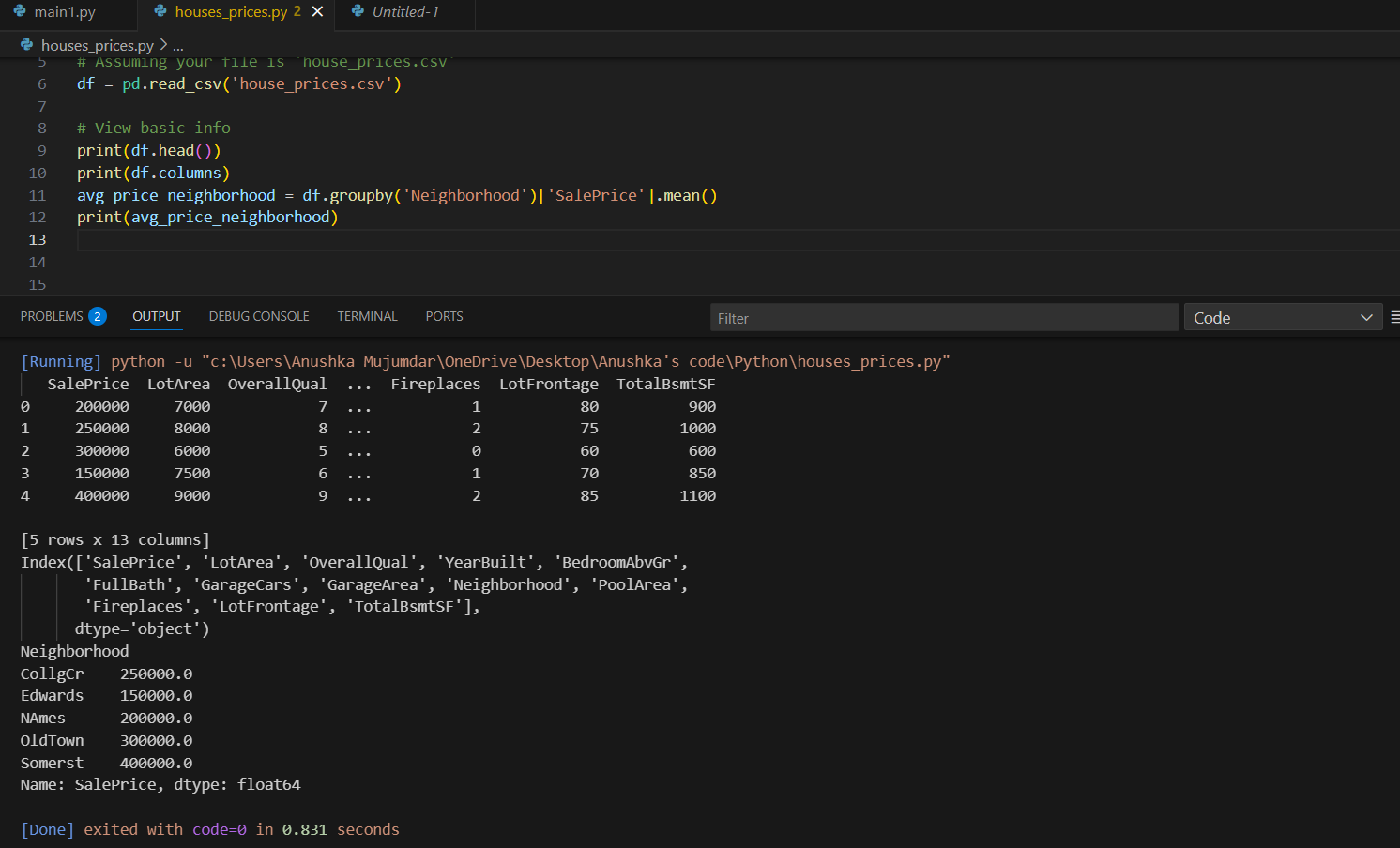
1. **Group houses by number of full bathrooms and calculate the average sale price.**



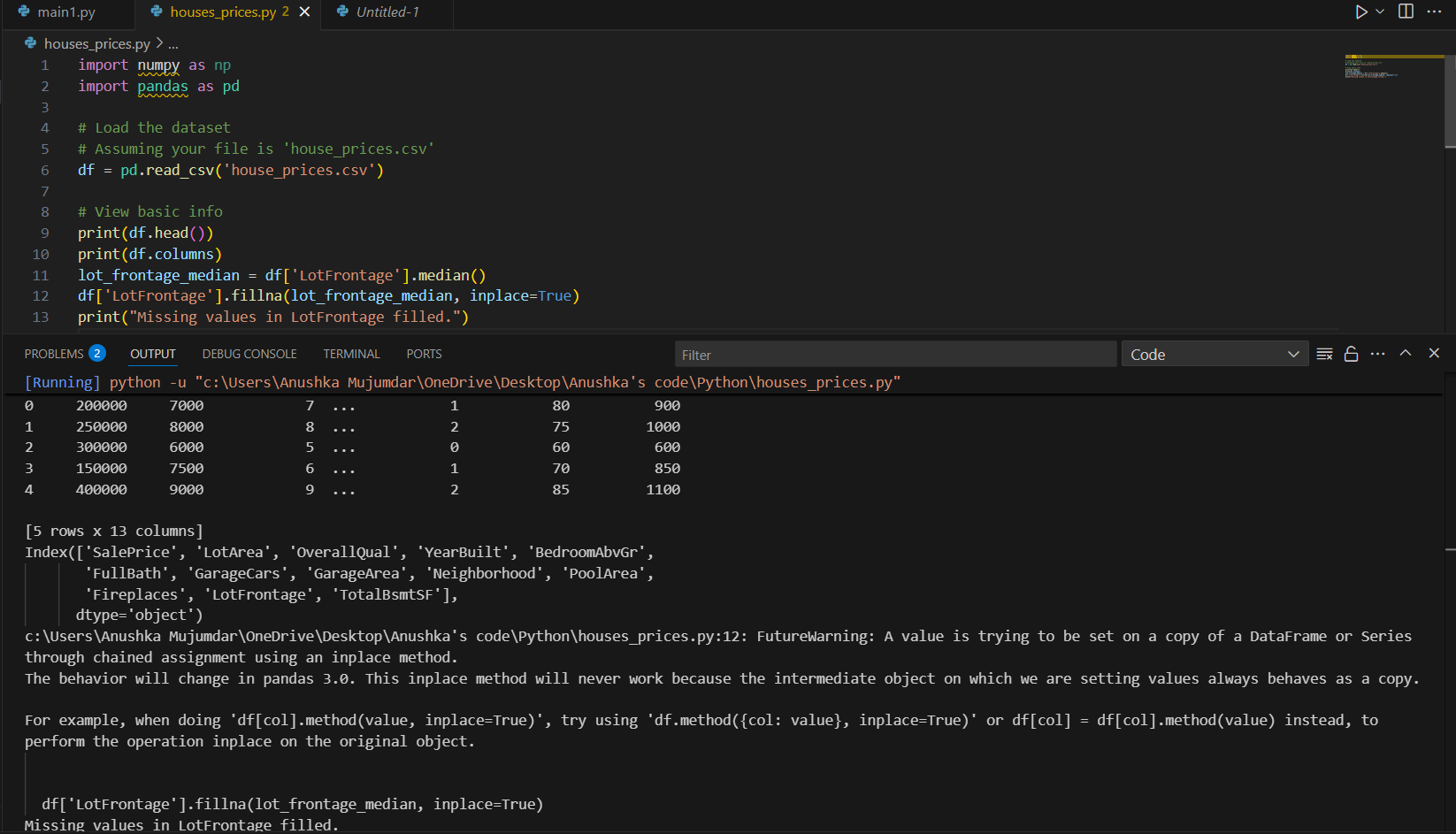
1. **How many unique neighborhoods are there?**



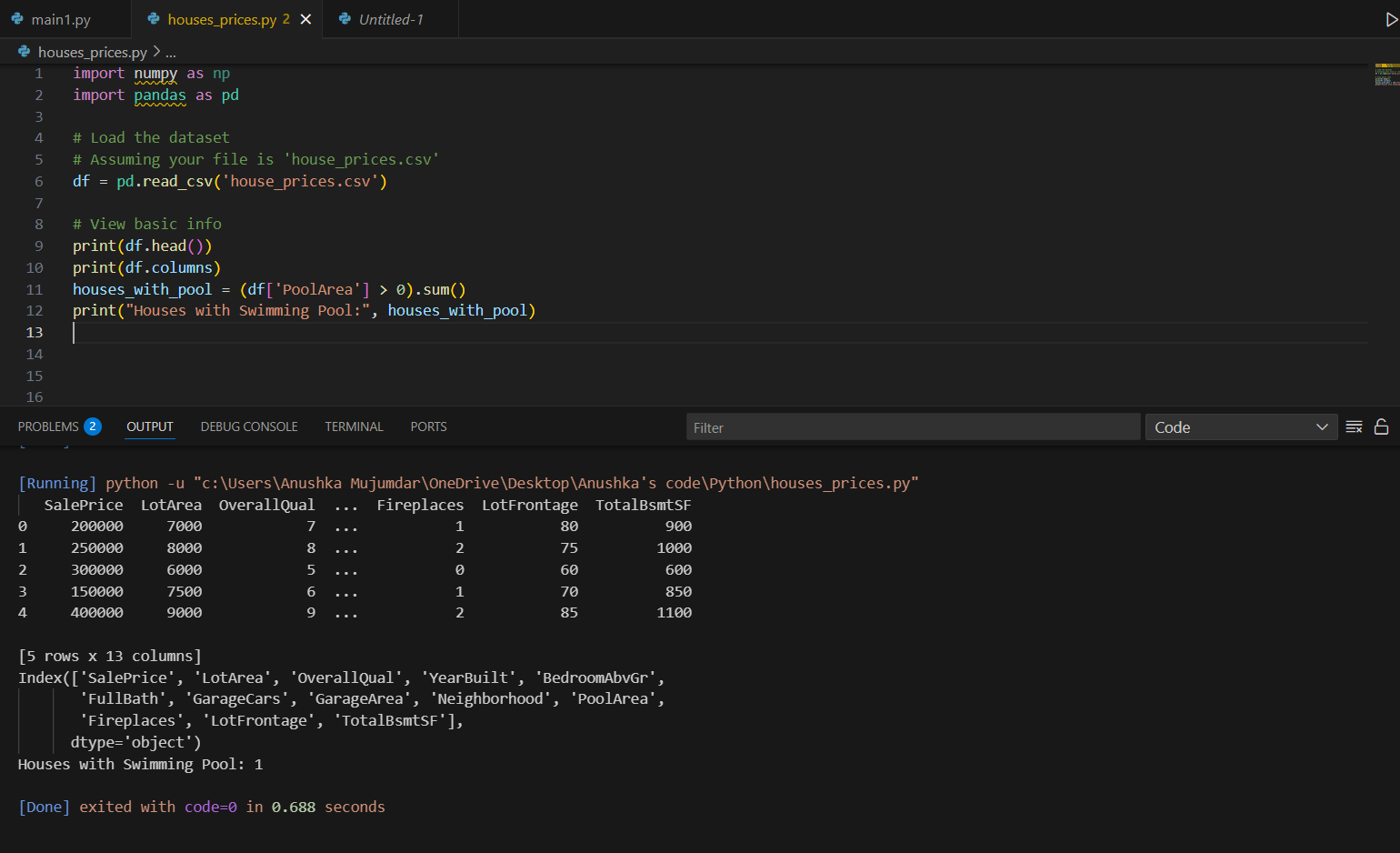
1. **Find the average sale price for each neighborhood.**



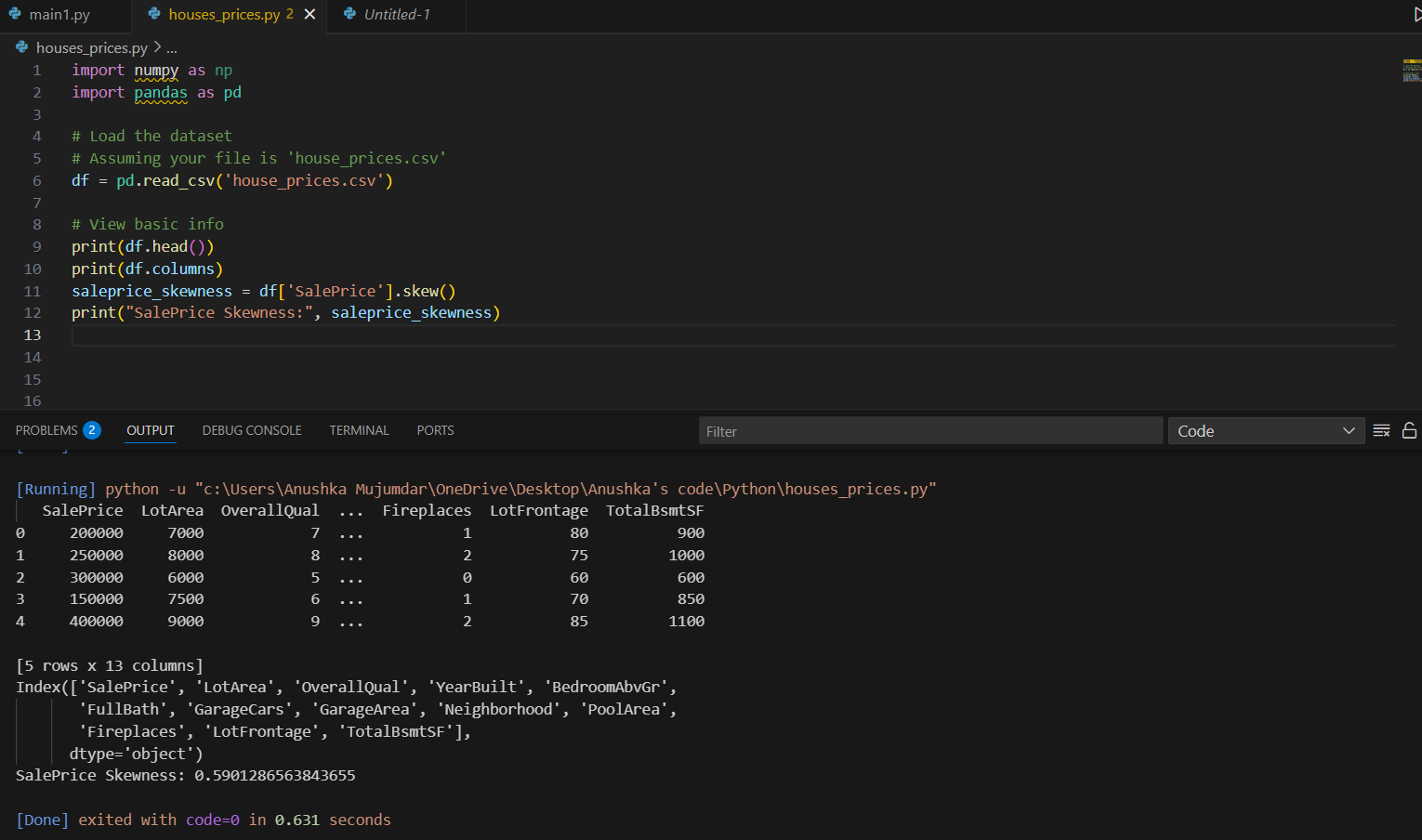
1. **Fill missing values in LotFrontage with the column’s median value.**



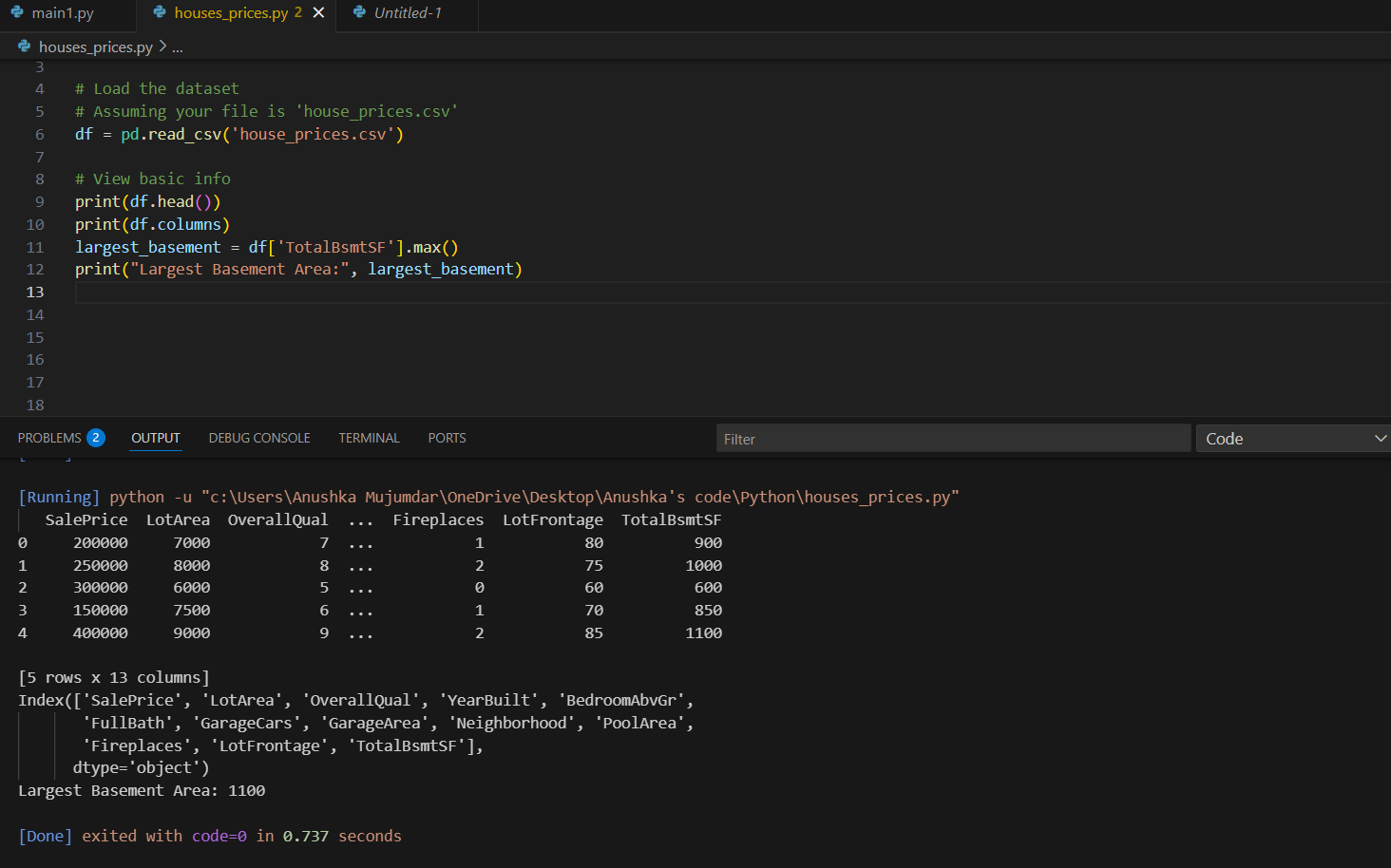
1. **How many houses have a swimming pool?**



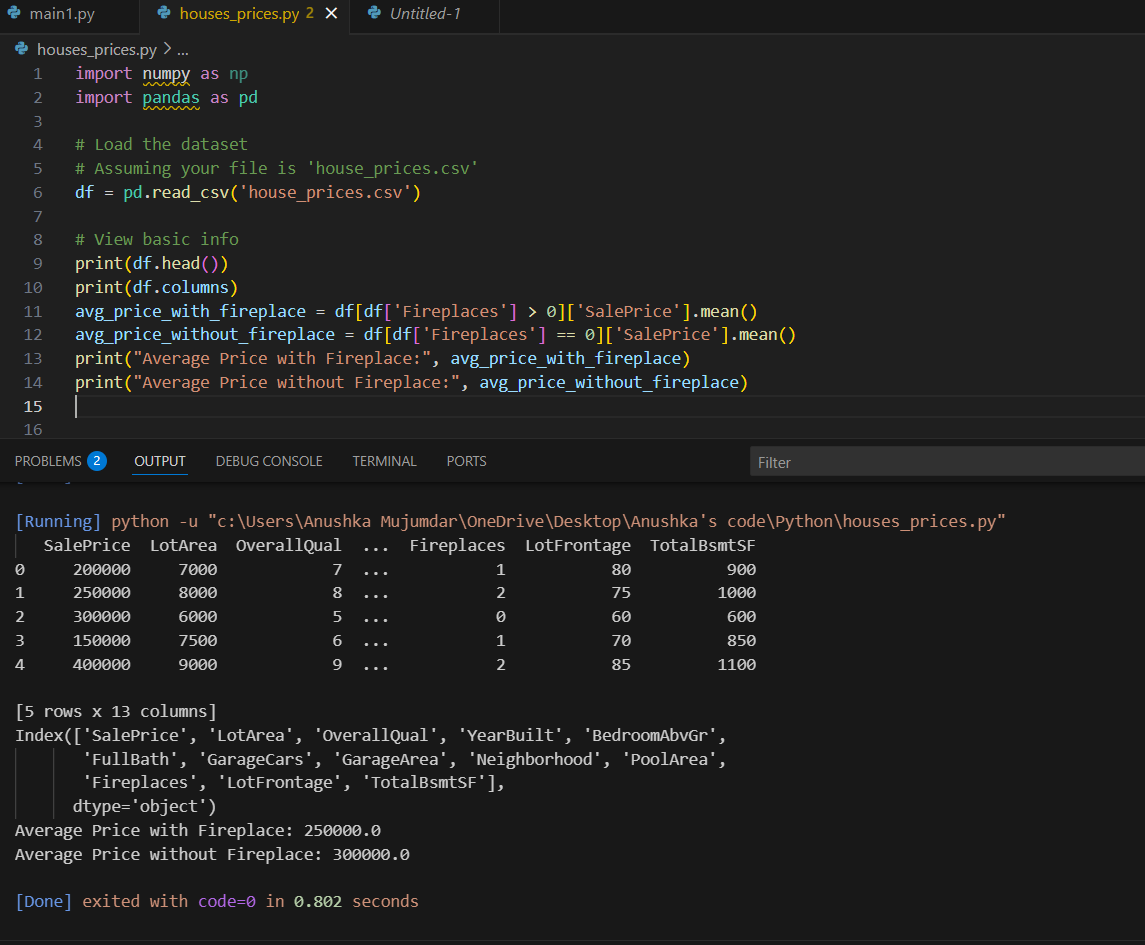
1. **What is the skewness of the SalePrice?**



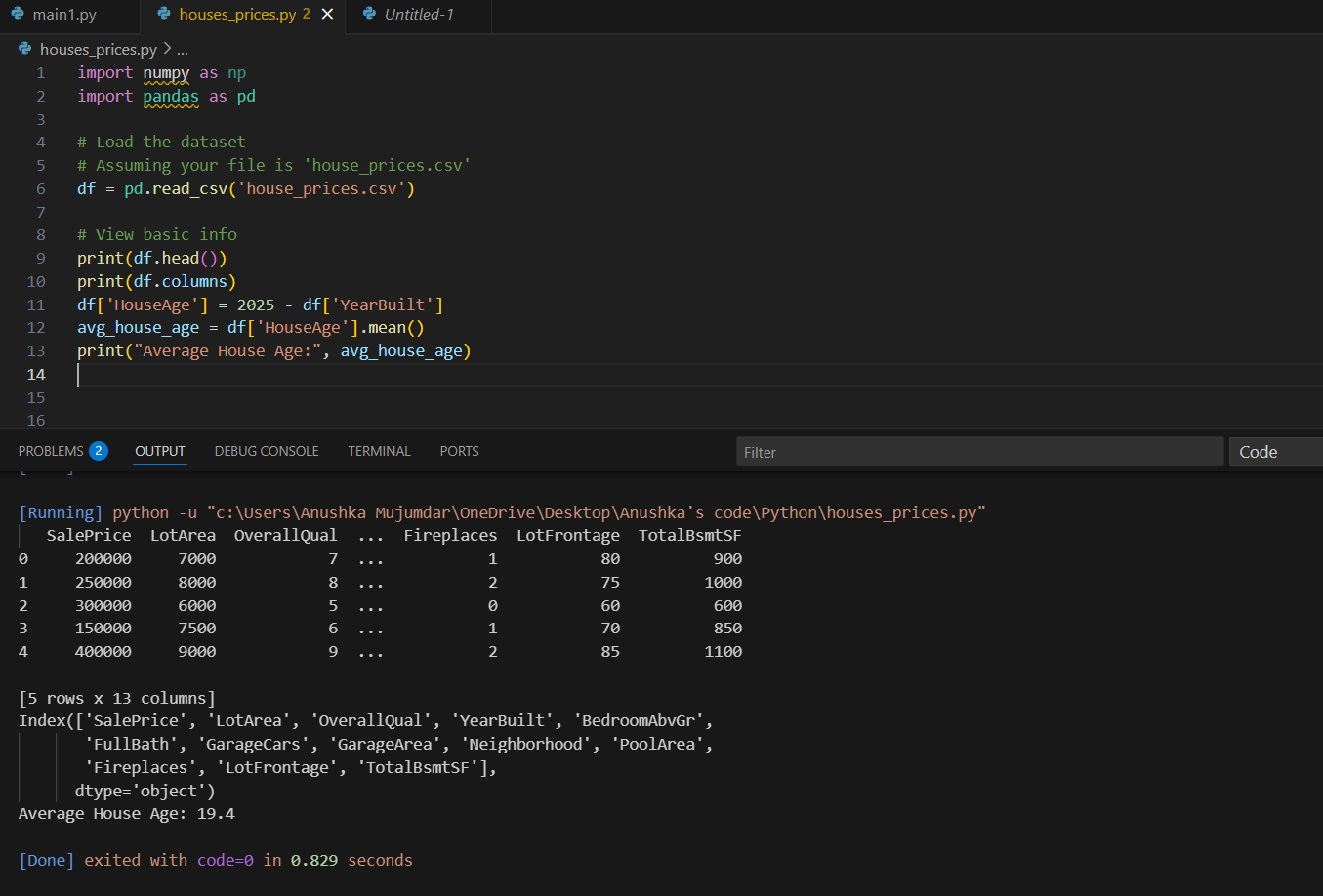
1. **Find the house with the largest total basement area (TotalBsmtSF).**



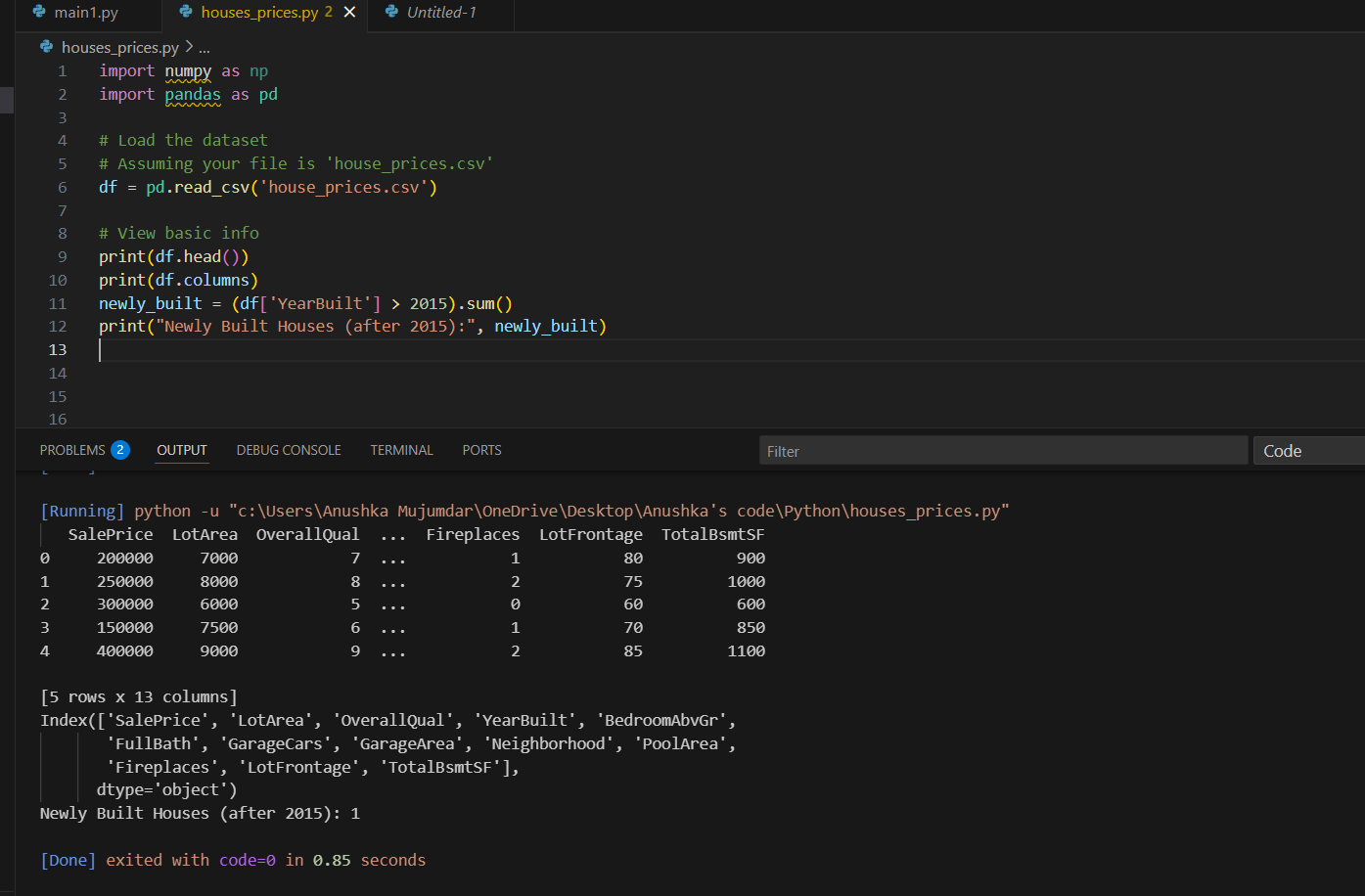
1. **Find the average sale price for houses with and without fireplaces.**



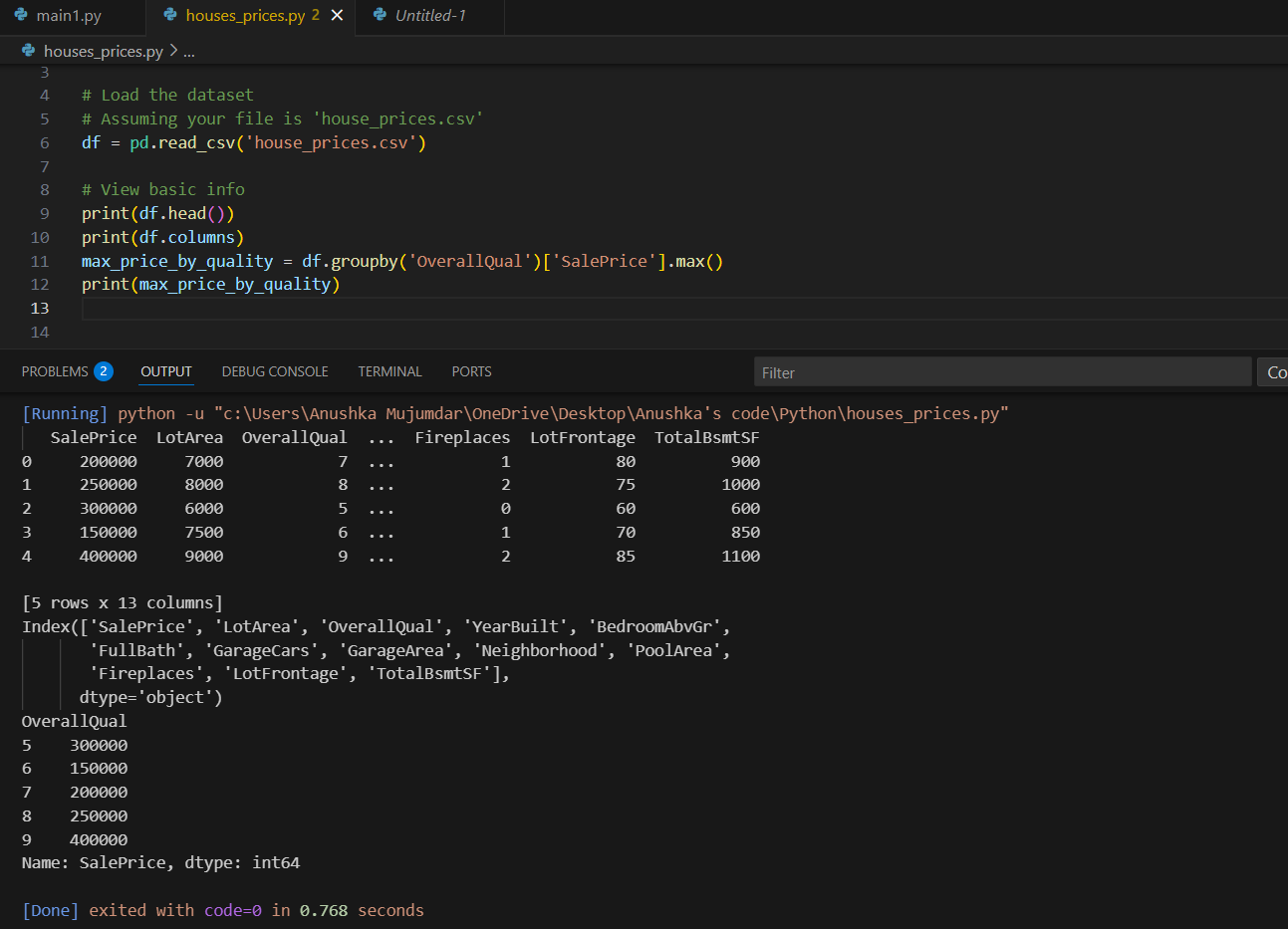
1. **Create a new column for the house age and find the average age.**



1. **How many houses are newly built (after 2015)?**



**19. Group houses by overall quality (OverallQual) and find maximum sale price in each group.**



**20. Find outliers in SalePrice using the IQR method.**

