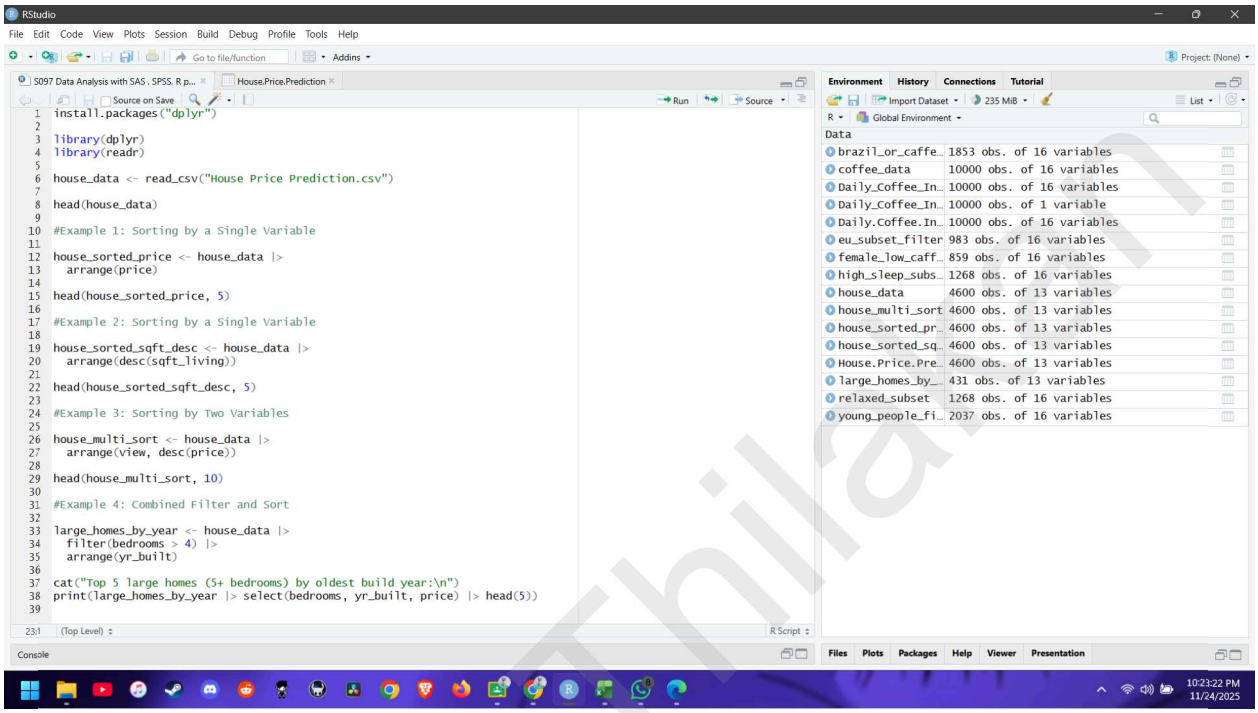
SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS /R

Practical No: 5

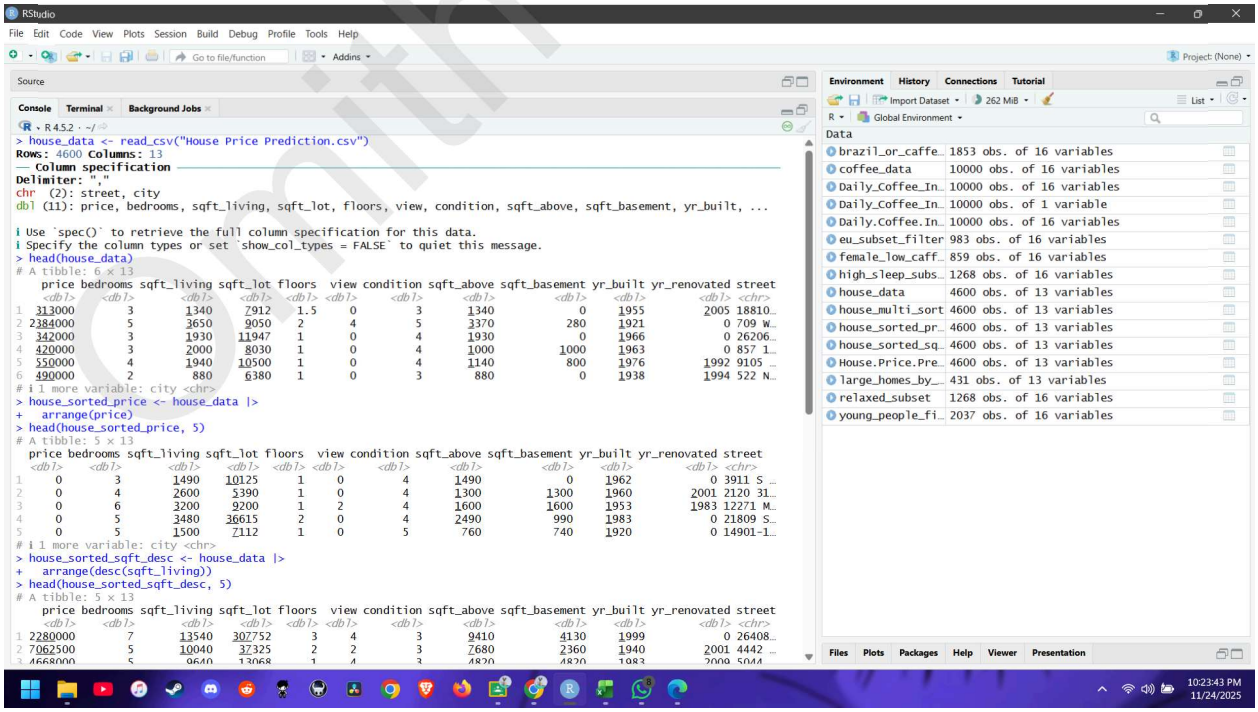
Aim: Sorting data using arrange() in R.

Code:



```
install.packages("dplyr")

library(dplyr)
library(readr)

house_data <- read_csv("House Price Prediction.csv")

head(house_data)

#Example 1: Sorting by a Single Variable

house_sorted_price <- house_data |>
  arrange(price)

head(house_sorted_price, 5)

#Example 2: Sorting by a Single Variable

house_sorted_sqft_desc <- house_data |>
  arrange(desc(sqft_living))

head(house_sorted_sqft_desc, 5)

#Example 3: Sorting by Two Variables

house_multi_sort <- house_data |>
  arrange(view, desc(price))

head(house_multi_sort, 10)

#Example 4: Combined Filter and Sort

large_homes_by_year <- house_data |>
  filter(bedrooms > 4) |>
  arrange(yr_built)

cat("Top 5 large homes (5+ bedrooms) by oldest build year:\n")
print(large_homes_by_year |> select(bedrooms, yr_built, price) |> head(5))
```

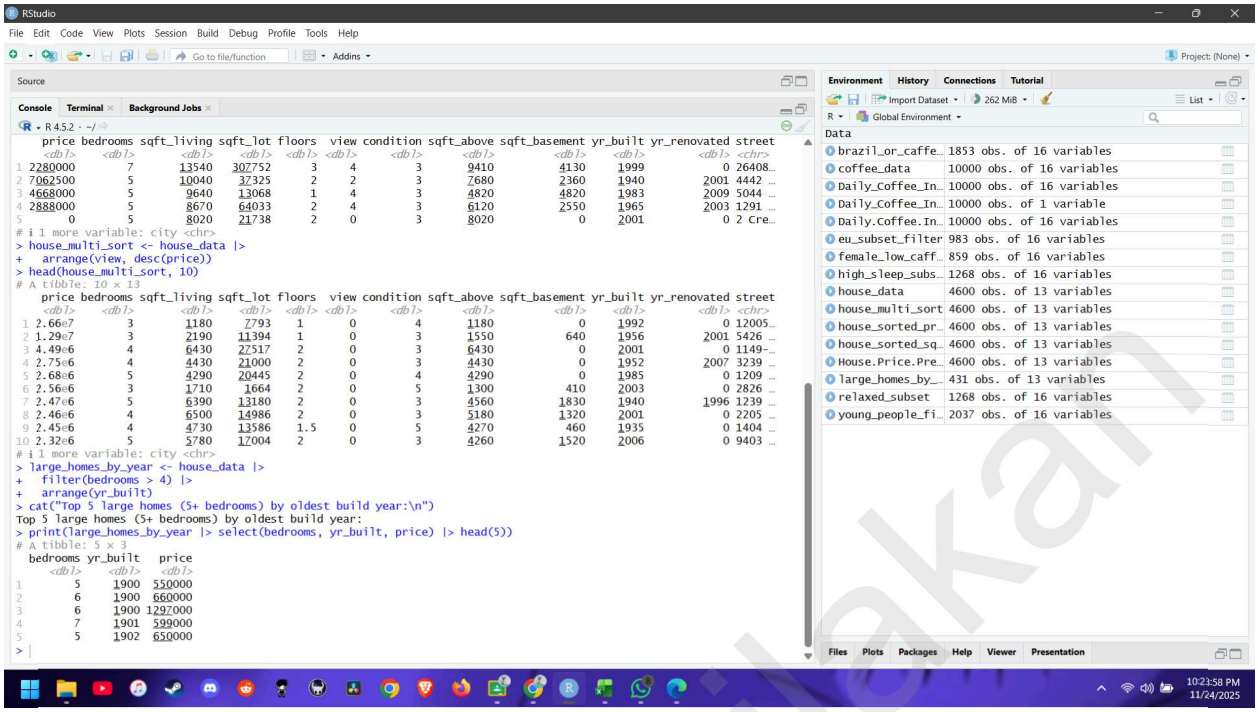
Output:



```
> house_data <- read_csv("House Price Prediction.csv")
Rows: 4600 Columns: 13
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr  (2): street, city
dbl (11): price, bedrooms, sqft_living, sqft_lot, floors, view, condition, sqft_above, sqft_basement, yr_built, ...

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
> head(house_data)
# A tibble: 6 × 13
    price bedrooms sqft_living sqft_lot floors  view condition sqft_above sqft_basement yr_built yr_renovated street
    <dbl>    <dbl>       <dbl>    <dbl>  <dbl> <dbl>     <dbl>      <dbl>        <dbl>    <dbl>        <dbl> <chr>
1  313000        3        1340     7912    1.5     0         3       1340            0     1955         2005 18810…
2 2384000        5        3650     9050    2       4         5       3370          280     1921            0 709 W…
3  342000        3        1930    11947    1       0         4       1930            0     1966            0 26206…
4  420000        3        2000     8030    1       0         4       1000         1000     1963            0 857 1…
5  550000        4        1940    10500    1       0         4       1140          800     1976         1992 9105 …
6  490000        2         880     6380    1       0         3        880            0     1938         1994 522 N…
# ℹ 1 more variable: city <chr>
> house_sorted_price <- house_data |>
+   arrange(price)
> head(house_sorted_price, 5)
# A tibble: 5 × 13
  price bedrooms sqft_living sqft_lot floors  view condition sqft_above sqft_basement yr_built yr_renovated street
  <dbl>    <dbl>       <dbl>    <dbl>  <dbl> <dbl>     <dbl>      <dbl>        <dbl>    <dbl>        <dbl> <chr>
1      0        3        1490    10125      1     0         4       1490            0     1962            0 3911 S …
2      0        4        2600     5390      1     0         4       1300         1300     1960         2001 2120 31…
3      0        6        3200     9200      1     2         4       1600         1600     1953         1983 12271 M…
4      0        5        3480    36615      2     0         4       2490          990     1983            0 21809 S…
5      0        5        1500     7112      1     0         5        760          740     1920            0 14901-1…
# ℹ 1 more variable: city <chr>
> house_sorted_sqft_desc <- house_data |>
+   arrange(desc(sqft_living))
> head(house_sorted_sqft_desc, 5)
# A tibble: 5 × 13
    price bedrooms sqft_living sqft_lot floors  view condition sqft_above sqft_basement yr_built yr_renovated street
    <dbl>    <dbl>       <dbl>    <dbl>  <dbl> <dbl>     <dbl>      <dbl>        <dbl>    <dbl>        <dbl> <chr>
1 2280000        7       13540   307752      3     4         3       9410         4130     1999            0 26408…
2 7062500        5       10040    37325      2     2         3       7680         2360     1940         2001 4442 …
3 4668000        5        9640    13068      1     4         3       4820         4820     1983         2009 5044…
```

Omith Thilakan
S097

SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS /R



Omith Thilakan

Omith Thilakan
S097