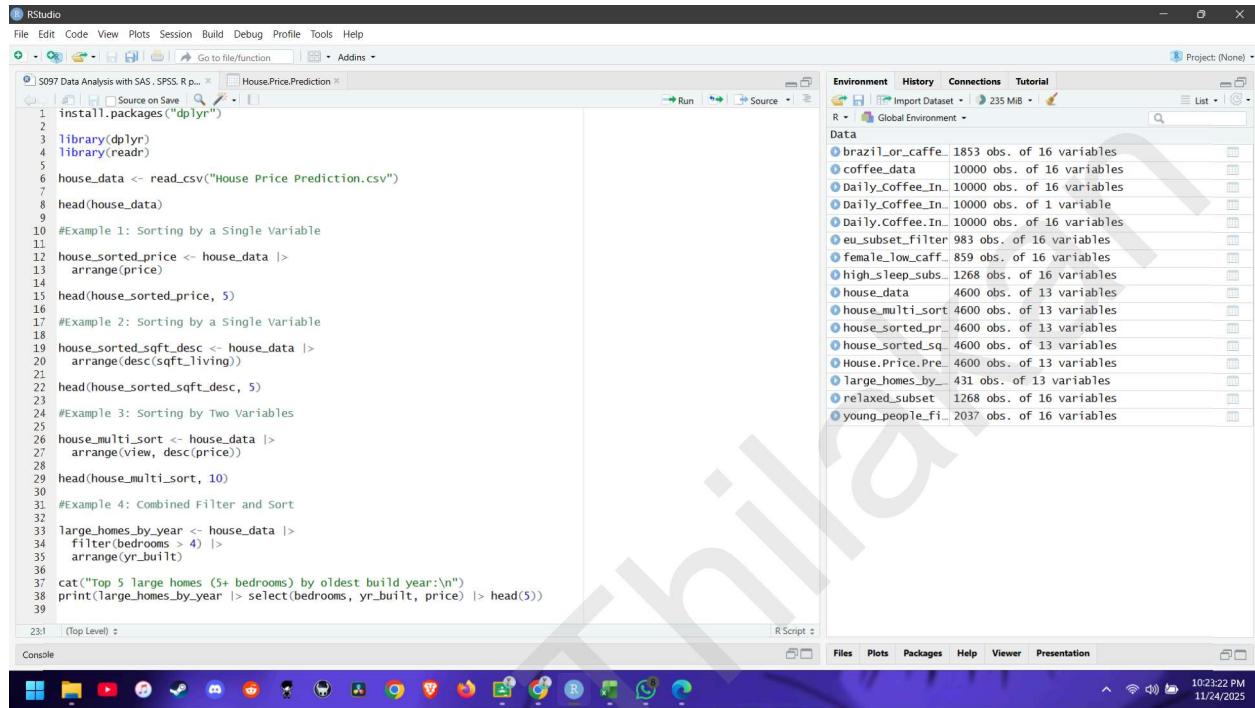


SHETH LUJ AND SIR MV COLLEGE
Subject: Data Analysis with SAS / SPSS /R

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)

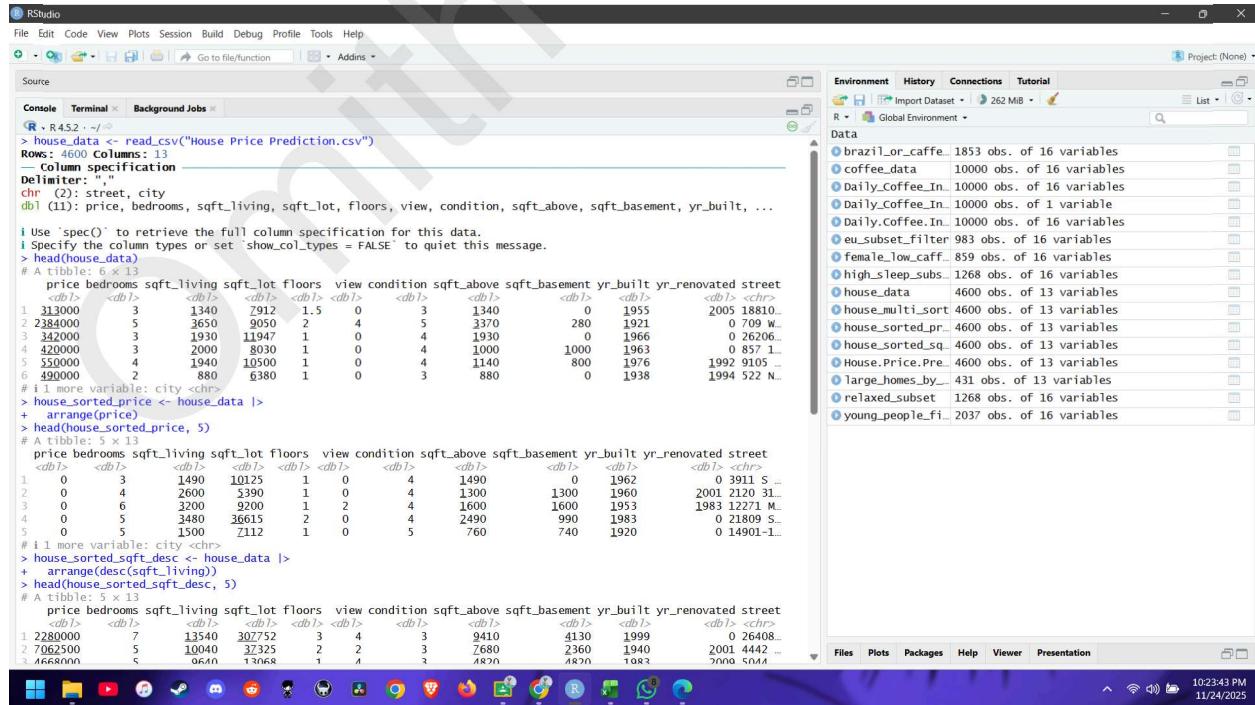
#Example 1: Sorting by a Single Variable
house_sorted_price <- house_data |>
  arrange(price)
head(house_sorted_price, 5)

#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))
  
```

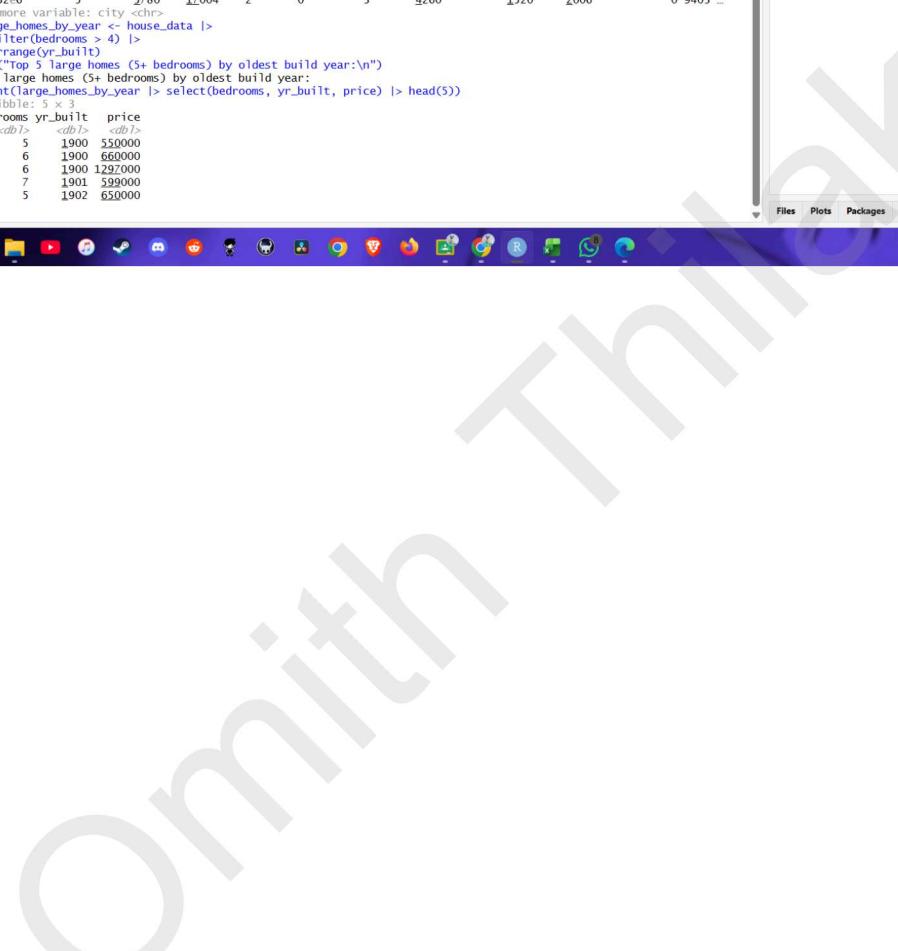
Output:



```

> 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, ...
Use 'spec()' to retrieve the full column specification for this data.
Specify the column types or set 'show_col_types = FALSE' to quiet this message.
> head(house_data)
# A tibble: 6 x 13
  price bedrooms sqft_living sqft_lot floors view condition sqft_above sqft_basement yr_built yr_renovated street
  <dbl>     <dbl>      <dbl>    <dbl>   <dbl> <chr>      <dbl>      <dbl>       <dbl>      <dbl> <chr>
1 313000     3        1340    2912  1.5     0      3        1340      0      1955    2005 18810...
2 3284000    5        3650    9050   2      4      5        3370      280     1921     0      709 W...
3 3420000    3        1930   11947   1      0      4        1930      0      1966     0      26206...
4 4200000    3        2000    8030   1      0      4        1000      1000    1963     0      857 L...
5 5500000    4        1940   10500   1      0      4        1140      800     1976    1992 9105...
6 4900000    2        880     6380   1      0      3        880      0      1938    1994 522 N...
# ... with 4594 variables: city <chr>
> house_sorted_price <- house_data |>
+   arrange(price)
> head(house_sorted_price, 5)
# A tibble: 5 x 13
  price bedrooms sqft_living sqft_lot floors view condition sqft_above sqft_basement yr_built yr_renovated street
  <dbl>     <dbl>      <dbl>    <dbl>   <dbl> <chr>      <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    2112    1     0      5        760      740     1920     0      14901 L...
# ... with 1 more variable: city <chr>
> house_sorted_sqft_desc <- house_data |>
+   arrange(desc(sqft_living))
> head(house_sorted_sqft_desc, 5)
# A tibble: 5 x 13
  price bedrooms sqft_living sqft_lot floors view condition sqft_above sqft_basement yr_built yr_renovated street
  <dbl>     <dbl>      <dbl>    <dbl>   <dbl> <chr>      <dbl>      <dbl>       <dbl>      <dbl> <chr>
1 2280000    7        13540   302752   3     4      3        9410      4130     1999     0      26408...
2 7062500    5        10040    32325   2     2      3        7680      2360     1940    2001 4442 ...
3 4668000    5        9640    13068   1     4      2        4820      4820     1082     0      5044 ...
  
```

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



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

Source

```

R • 4.5.2 - ~/r
price bedrooms sqft_living sqft_lot floors view condition sqft_above sqft_basement yr_built yr_renovated street
<dbl> <dbl>
1 2280000 7 13540 30752 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 ...
4 2888000 5 8670 64033 2 4 3 6120 2550 1965 2003 1291 ...
5 0 3 8020 21738 2 0 3 8020 0 2001 0 2 Cre...
# i 1 more variable: city <chr>
> house_multi_sort <- house_data |>
+ arrange(view, desc(price))
> head(house_multi_sort, 10)
# A tibble: 10 x 13
  price bedrooms sqft_living sqft_lot floors view condition sqft_above sqft_basement yr_built yr_renovated street
  <dbl> <dbl>
1 2.66e7 3 1180 2793 1 0 4 1180 0 1992 0 12005...
2 1.29e7 3 2190 11394 1 0 3 1550 640 1956 2001 5426 ...
3 4.49e6 4 6430 27517 2 0 3 6430 0 2001 0 1149...
4 2.75e6 4 4430 21000 2 0 3 4430 0 1952 2007 3239 ...
5 2.68e6 5 4290 20445 2 0 4 4290 0 1985 0 1200...
6 2.56e6 3 1710 1654 2 0 5 1300 410 2003 0 1200...
7 2.46e6 5 6390 1280 2 0 3 4560 1830 1940 1996 1239 ...
8 2.46e6 4 6500 14986 2 0 3 5180 1320 2001 0 2205 ...
9 2.45e6 4 4730 13586 1.5 0 5 4270 460 1935 0 1404 ...
10 2.32e6 5 5780 12004 2 0 3 4260 1520 2006 0 9403 ...
# i 1 more variable: city <chr>
> large_homes_by_year <- house_data |>
+ filter(bedrooms > 4) |>
+ arrange(yr_built)
> cat("Top 5 large homes (5+ bedrooms) by oldest build year:\n")
Top 5 large homes (5+ bedrooms) by oldest build year:
> print(large_homes_by_year |> select(bedrooms, yr_built, price) |> head(5))
# A tibble: 5 x 3
  bedrooms yr_built price
  <dbl> <dbl> <dbl>
1 5 1900 550000
2 6 1900 660000
3 6 1900 1297000
4 7 1901 599000
5 5 1902 650000
> |

```

Environment History Connections Tutorial

Data

- brazi..._or_caffe_ 1853 obs. of 16 variables
- coffee_data 10000 obs. of 16 variables
- Daily_Coffee_In_ 10000 obs. of 16 variables
- Daily_Coffee_In_ 10000 obs. of 1 variable
- Daily.Coffee.In_ 10000 obs. of 16 variables
- eu_subset_filter 983 obs. of 16 variables
- female_low_caff_ 859 obs. of 16 variables
- high_sleep_subsc_ 1268 obs. of 16 variables
- house_data 4600 obs. of 13 variables
- house_multi_sort 4600 obs. of 13 variables
- house_sorted_pr_ 4600 obs. of 13 variables
- house_sorted_sq_ 4600 obs. of 13 variables
- House.Price.Pre_ 4600 obs. of 13 variables
- large_homes_by_ 431 obs. of 13 variables
- relaxed_subset 1268 obs. of 16 variables
- young_people_fi_ 2037 obs. of 16 variables

Files Plots Packages Help Viewer Presentation

10:23:58 PM
11/24/2025