

Environment History Connections Tutorial

R - Global Environment

Data

Object	Size
data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b...	5 obs. of 2 variables
mean_emp_by.le...	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

```

1 library(tidyverse)
2 install.packages("tidyverse")
3 library(tidyverse)
4 data <- read_csv("~/Downloads/OccupationEmploymentStatistics (1).csv")
5 colnames(data) <- c("Code", "Title", "Level", "Emp", "EmpRSE", "EmpPer1000",
6   "MedHourlyWage", "MeanHourlyWage", "AnnWage", "MeanRSE")
7 head(data)
8 data <- data %>%
9   mutate(
10     Emp = as.numeric(gsub(" ", "", Emp)),
11     EmpRSE = as.numeric(gsub("%", "", EmpRSE)),
12     AnnWage = as.numeric(gsub("$", "", AnnWage)),
13     MeanRSE = as.numeric(gsub("%", "", MeanRSE))
14   )
15 mean_emp_by_level <- data %>%
16   group_by(Level) %>%
17   summarise(mean_emp = mean(Emp, na.rm = TRUE))
18
19 print(mean_emp_by_level)
20 mean_emp_by_level <- data %>%
21   group_by(Level) %>%
22   summarise(mean_emp = mean(Emp, na.rm = TRUE))
23
24 print(mean_emp_by_level)
25 head(data)
26 mean_emp_by_level <- data %>%
27   group_by(Level) %>%
28   summarise(mean_emp = mean(Emp, na.rm = TRUE))
29
30 print(mean_emp_by_level)
31 data %>%
32   group_by(Level)
33   summarise(mean_emp = mean(emp, na.rm = TRUE))
34 colnames(data)
35 mean_emp_by_level <- data %>%
36   group_by(Level) %>%
37   summarise(mean_emp = mean(Employment, na.rm = TRUE))
38 data_filtered <- data %>%
39   filter(grepl("total", tolower(Title)))
40 max_emp <- data_filtered %>%
41   filter(Emp == max(Emp, na.rm = TRUE)) %>%
42   select(Title, Emp)
43 print(max_emp)
44 mean_emprese_by_level <- data %>%
45   group_by(Level) %>%

```

63:1 (Top Level) : R Script

Console

Environment History Connections Tutorial

R - Global Environment

Data

Object	Size
data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b...	5 obs. of 2 variables
mean_emp_by.le...	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Source

Console Terminal Background Jobs

R - R 4.2.2 - ~/Downloads/

```

The downloaded binary packages are in
/var/folders/kd/2tdwkrq94q1bq50x8twq4c700000gn/T//RtmpP23hpr/downloaded_packages
> library(tidyverse)
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ forcats 1.0.0 ✓ readr 2.1.5
✓ ggplot2 3.5.1 ✓ stringr 1.5.1
✓ lubridate 1.9.4 ✓ tibble 3.2.1
✓ purrr 1.0.4 ✓ tidyr 1.3.1
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag() masks stats::lag()
i Use the conflicted package to force all conflicts to become errors
> data <- read_csv("~/Downloads/OccupationEmploymentStatistics (1).csv")
Rows: 1070 Columns: 10
— Column specification —
Delimiter: ","
chr (8): Occupation code, Occupation title (click on the occupation title...
dbl (1): Employment per 1,000 jobs
num (1): Employment

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(data)
# A tibble: 6 x 10
  Occupation code Occupation title (cli... Level Employment Employment RSE
  <chr> <chr> <chr> <dbl> <chr>
1 00-0000 All Occupations total 144733270 0.10%
2 11-0000 Management Occupations major 7616650 0.20%
3 11-1000 Top Executives minor 2535640 0.30%
4 11-1011 Chief Executives deta_ 195530 0.70%
5 11-1021 General and Operations_ deta_ 2289770 0.30%
6 11-1031 Legislators deta_ 50330 2.30%

# i abbreviated name:
# ^ Occupation title (click on the occupation title to view its profile)
# i 5 more variables: 'Employment per 1,000 jobs' <dbl>,
# 'Median hourly wage' <chr>, 'Mean hourly wage' <chr>,
# 'Annual mean wage' <chr>, 'Mean wage RSE' <chr>
> data <- data %>%
+   mutate(
+     Emp = as.numeric(gsub(" ", "", Emp)),
+     EmpRSE = as.numeric(gsub("%", "", EmpRSE)),
+     AnnWage = as.numeric(gsub("$", "", AnnWage)),
+     MeanRSE = as.numeric(gsub("%", "", MeanRSE))
+   )

```

Source

Console

```
R - R 4.4.2 - ~/
/var/folders/kd/2tdkrg94q1bq50x8twq4c700000gn/T//RtmpP23hpr/downloaded_packages
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union

> movie <- read.csv("movie.csv")
Error in file(file, "rt") : cannot open the connection
In addition: Warning message:
In file(file, "rt") :
  cannot open file 'movie.csv': No such file or directory
> names(movie)
Error: object 'movie' not found
> movie <- read.csv("~/Downloads/movie.csv")
> names(movie)
[1] "title"      "score"      "rating"
[4] "genre"      "box.office" "running.time"
> movie_selected <- select(movie, title, score, rating, genre)
> movie_r <- select(movie, starts_with("r"), contains("I"))
> movie_no_runtime <- select(movie, "running time")
Error in `select()` :
! Can't select columns that don't exist.
✖ Column 'running time' doesn't exist.
Run `rlang::last_trace()` to see where the error occurred.
> movie_pg13 <- filter(movie, rating == "PG-13")
> movie_comedy <- filter(movie, genre == "comedy")
> movie_big_hits <- filter(movie, box.office > 100, score > 50)
> movie <- mutate(movie, running_score = score / 'running time')
Error in `mutate()` :
! In argument: 'running_score = score / 'running
time''.
Caused by error in 'score / 'running time'':
! non-numeric argument to binary operator
Run `rlang::last_trace()` to see where the error occurred.
> tapply(movie$score, movie$genre, mean, na.rm = TRUE)
      action action/adventure    animated
55.50000      56.03214      63.58333
      comedy    documentary      drama
```

Environment

History

Connections

Tutorial

R - Global Environment

Data

movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables
movie_r	140 obs. of 4 variables
movie_selected	140 obs. of 4 variables

Values

score	59.11

Files

Plots

Packages

Help

Viewer

Presentation

Source

```
19 print(mean_emp_by_level)
20 mean_emp_by_level <- data %>%
21   group_by(level) %>%
22   summarise(mean_emp = mean(Emp, na.rm = TRUE))
23
24 print(mean_emp_by_level)
25 head(data)
26 mean_emp_by_level <- data %>%
27   group_by(level) %>%
28   summarise(mean_emp = mean(Emp, na.rm = TRUE))
29
30 print(mean_emp_by_level)
31 data %>%
32   group_by(level) %>%
33   summarise(mean_emp = mean(emp, na.rm = TRUE))
34 colnames(data)
35 mean_emp_by_level <- data %>%
36   group_by(level) %>%
37   summarise(mean_emp = mean(Employment, na.rm = TRUE))
38 data_filtered <- data %>%
39   filter(!grepl("total", tolower(title)))
40 max_emp <- data_filtered %>%
41   filter(Emp == max(Emp, na.rm = TRUE)) %>%
42   select(title, Emp)
43 print(max_emp)
44 mean_emprese_by_level <- data %>%
45   group_by(level) %>%
46   summarise(mean_emprese = mean(Employment RSE, na.rm = TRUE))
47 mean_emprese_by_level <- data %>%
48   group_by(level) %>%
49   summarise(mean_emprese = mean(EmploymentRSE, na.rm = TRUE))
50
51 print(mean_emprese_by_level)
52 mean_annwage_by_level <- data %>%
53   group_by(level) %>%
54   summarise(mean_annwage = mean(AnnWage, na.rm = TRUE))
55
56 print(mean_annwage_by_level)
57
58 max_annwage <- data_filtered %>%
59   filter(AnnWage == max(AnnWage, na.rm = TRUE)) %>%
60   select(title, AnnWage)
61
62 print(max_annwage)
63
```

Environment

History

Connections

Tutorial

R - Global Environment

Data

data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b_	5 obs. of 2 variables
mean_emp_by_le_	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files

Plots

Packages

Help

Viewer

Presentation

Source

Console

```
R - R 4.4.2 - ~/Downloads/
> data <- data %>%
+ mutate(
+   Emp = as.numeric(gsub(",", "", Emp)),
+   EmpRSE = as.numeric(gsub("%", "", EmpRSE)),
+   AnnWage = as.numeric(gsub("$", "", AnnWage)),
+   MeanRSE = as.numeric(gsub("%", "", MeanRSE))
+ )
Error in `mutate()` :
! In argument: `Emp = as.numeric(gsub(",", "", Emp))`.
Caused by error:
! object 'Emp' not found
Run `rlang::last_trace()` to see where the error occurred.
> print(mean_emp_by_level)
Error: object 'mean_emp_by_level' not found
> print(mean_Emp_by_level)
Error: object 'mean_Emp_by_level' not found
>
> print(mean_emp_by_level)
Error: object 'mean_emp_by_level' not found
> head(data)
# A tibble: 6 x 10
  Occupation code Occupation title (click on the occupation title to view its profile)
  <chr> <chr> <chr> <dbl> <chr>
1 00-0000 All Occupations total 144733270 0.10%
2 11-0000 Management Occupations major 7616650 0.20%
3 11-1000 Top Executives minor 2535640 0.30%
4 11-1011 Chief Executives deta_ 195530 0.70%
5 11-1021 General and Operations deta_ 2289770 0.30%
6 11-1031 Legislators deta_ 50330 2.30%
# i abbreviated name:
# 1 Occupation title (click on the occupation title to view its profile)
# i 5 more variables: 'Employment per 1,000 jobs' <dbl>,
# 'Median hourly wage' <chr>, 'Mean hourly wage' <chr>,
# 'Annual mean wage' <chr>, 'Mean wage RSE' <chr>
> print(mean_emp_by_level)
Error: object 'mean_emp_by_level' not found
> data %>%
+ group_by(Level)
# A tibble: 1,070 x 10
# Groups: Level [5]
  Occupation code Occupation title (click on the occupation title to view its profile)
  <chr> <chr> <chr> <dbl> <chr>
1 00-0000 All Occupations total 144733270 0.10%
2 11-0000 Management Occupations major 7616650 0.20%
3 11-1000 Top Executives minor 2535640 0.30%
```

Environment

History

Connections

Tutorial

R - Global Environment

Data

data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b...	5 obs. of 2 variables
mean_emp_by_le...	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files

Plots

Packages

Help

Viewer

Presentation

Source

Console

```
R - R 4.4.2 - ~/Downloads/
> print(mean_emp_by_level)
Error: object 'mean_emp_by_level' not found
> data %>%
+ group_by(Level)
# A tibble: 1,070 x 10
# Groups: Level [5]
  Occupation code Occupation title (click on the occupation title to view its profile)
  <chr> <chr> <chr> <dbl> <chr>
1 00-0000 All Occupations total 144733270 0.10%
2 11-0000 Management Occupations major 7616650 0.20%
3 11-1000 Top Executives minor 2535640 0.30%
4 11-1011 Chief Executives deta_ 195530 0.70%
5 11-1021 General and Operations deta_ 2289770 0.30%
6 11-1031 Legislators deta_ 50330 2.30%
7 Nov-00 Advertising, Marketin_ minor 712220 0.50%
8 Nov-11 Advertising and Promo_ deta_ 25260 3.00%
9 Nov-20 Marketing and Sales M. broad 619500 0.60%
10 Nov-21 Marketing Managers deta_ 240440 0.80%
# i 1,060 more rows
# i abbreviated name:
# 1 Occupation title (click on the occupation title to view its profile)
# i 5 more variables: 'Employment per 1,000 jobs' <dbl>,
# 'Median hourly wage' <chr>, 'Mean hourly wage' <chr>,
# 'Annual mean wage' <chr>, 'Mean wage RSE' <chr>
# Use `print(n = ...)` to see more rows
> summarise(mean_emp = mean(Emp, na.rm = TRUE))
Error: object 'Emp' not found
> summarise(mean_emp = mean(emp, na.rm = TRUE))
Error: object 'emp' not found
> colnames(data)
[1] "Occupation code"
[2] "Occupation title (click on the occupation title to view its profile)"
[3] "Level"
[4] "Employment"
[5] "Employment RSE"
[6] "Employment per 1,000 jobs"
[7] "Median hourly wage"
[8] "Mean hourly wage"
[9] "Annual mean wage"
[10] "Mean wage RSE"
> mean_emp_by_level <- data %>%
+ group_by(Level) %>%
+ summarise(mean_emp = mean(Employment, na.rm = TRUE))
> data_filtered <- data %>%
+ filter(!grepl("total", tolower(title)))
```

Environment

History

Connections

Tutorial

R - Global Environment

Data

data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b...	5 obs. of 2 variables
mean_emp_by_le...	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files

Plots

Packages

Help

Viewer

Presentation

Source

Console

Terminal

Background Jobs

R • R 4.4.2 • ~/Downloads/

```
group_by(Level) %>%
  summarise(mean_emprese = mean(Employment RSE"
> mean_emprese_by_level <- data %>%
+ group_by(Level) %>%
+ summarise(mean_emprese = mean(EmploymentRSE, na.rm = TRUE))
Error in `summarise()`:
! In argument: `mean_emprese = mean(EmploymentRSE, na.rm = TRUE)`.
! In group 1: `Level = "broad"`.
Caused by error:
! object 'EmploymentRSE' not found
Run `rlang::last_trace()` to see where the error occurred.
> print(mean_emprese_by_level)
Error: object 'mean_emprese_by_level' not found
>
> print(mean_annwage_by_level)
Error: object 'mean_annwage_by_level' not found
> print(mean_annualwage_by_level)
Error: object 'mean_annualwage_by_level' not found
> mean_annwage_by_level <- data %>%
+ group_by(Level) %>%
+ summarise(mean_annwage = mean(AnnWage, na.rm = TRUE))
Error in `summarise()`:
! In argument: `mean_annwage = mean(AnnWage, na.rm = TRUE)`.
! In group 1: `Level = "broad"`.
Caused by error:
! object 'AnnWage' not found
Run `rlang::last_trace()` to see where the error occurred.
> mean_annwage_by_level <- data %>%
+ group_by(Level) %>%
+ summarise(mean_annwage = mean(`Annual mean wage`, na.rm = TRUE))
Warning message:
There were 5 warnings in `summarise()`.
The first warning was:
! In argument: `mean_annwage = mean(`Annual mean wage`, na.rm = TRUE)`.
! In group 1: `Level = "broad"`.
Caused by warning in `mean.default()`:
! argument is not numeric or logical: returning NA
! Run `dplyr::last_dplyr_warnings()` to see the 4 remaining warnings.
> max_annwage <- data_filtered %>%
+ filter(AnnWage == max(AnnWage, na.rm = TRUE)) %>%
+ select(Title, AnnWage)
> print(max_annwage)

              Title  AnnWage
1 Political Science Teachers, Postsecondary $99,480
> |
```

Environment

History

Connections

Tutorial

R • Global Environment

Data

data	1070 obs. of 10 variables
data_filtered	1069 obs. of 10 variables
max_annwage	1 obs. of 2 variables
max_emp	1 obs. of 2 variables
mean_annwage_b...	5 obs. of 2 variables
mean_emp_by_le...	5 obs. of 2 variables
movie	140 obs. of 6 variables
movie_big_hits	24 obs. of 6 variables
movie_comedy	39 obs. of 6 variables
movie_pg13	65 obs. of 6 variables

Files

Plots

Packages

Help

Viewer

Presentation

Export