데이터프레임이란?

Code

배여운

2018년 1월 10일

데이터프레임이란

Hide

```
# Print out built-in R data frame
mtcars

# Call head() on mtcars
head(mtcars)

# Investigate the structure of mtcars
str(mtcars)
```

데이터프레임 만들기

Hide

```
# Definition of vectors
name <- c("Mercury", "Venus", "Earth", "Mars", "Jupiter", "Saturn", "Uranus", "Neptun
e")
type <- c("Terrestrial planet", "Terrestrial planet", "Terrestrial planet",
          "Terrestrial planet", "Gas giant", "Gas giant", "Gas giant", "Gas giant")
diameter <- c(0.382, 0.949, 1, 0.532, 11.209, 9.449, 4.007, 3.883)
rotation <- c(58.64, -243.02, 1, 1.03, 0.41, 0.43, -0.72, 0.67)
rings <- c(FALSE, FALSE, FALSE, FALSE, TRUE, TRUE, TRUE, TRUE)
# Create a data frame from the vectors
planets_df <- data.frame(name, type, diameter, rotation, rings)
planets_df
```

name <fctr>	type <fctr>	diameter <dbl>	rotation <dbl>	rings <lgl>
Mercury	Terrestrial planet	0.382	58.64	FALSE
Venus	Terrestrial planet	0.949	-243.02	FALSE
Earth	Terrestrial planet	1.000	1.00	FALSE
Mars	Terrestrial planet	0.532	1.03	FALSE
Jupiter	Gas giant	11.209	0.41	TRUE
Saturn	Gas giant	9.449	0.43	TRUE
Uranus	Gas giant	4.007	-0.72	TRUE
Neptune	Gas giant	3.883	0.67	TRUE

8 rows

Hide

```
# Check the structure of planets_df
str(planets_df)
```

```
'data.frame':   8 obs. of  5 variables:
 $ name    : Factor w/ 8 levels "Earth","Jupiter",..: 4 8 1 3 2 6 7 5
 $ type    : Factor w/ 2 levels "Gas giant","Terrestrial planet": 2 2 2 2 1 1 1 1
 $ diameter: num  0.382 0.949 1 0.532 11.209 ...
 $ rotation: num  58.64 -243.02 1 1.03 0.41 ...
 $ rings   : logi  FALSE FALSE FALSE FALSE TRUE TRUE ...
```

Selection of data frame elements

Hide

```
# The planets_df data frame from the previous exercise is pre-loaded
planets_df
# Print out diameter of Mercury (row 1, column 3)
planets_df[1,3]

# Print out data for Mars (entire fourth row)
planets_df[4,]

# The planets_df data frame from the previous exercise is pre-loaded
planets_df
# Select first 5 values of diameter column
planets_df[1:5, "diameter"]
```

특정 칼럼만 저장하기

Hide

```
# planets_df is pre-loaded in your workspace
planets_df
# Select the rings variable from planets_df
rings_vector <- planets_df$rings

# Print out rings_vector
rings_vector
```

TRUE, FALSE를 활용한 특정 칼럼 요소 출력

Hide

```
# planets_df and rings_vector are pre-loaded in your workspace
planets_df
# Adapt the code to select all columns for planets with rings
planets_df[rings_vector,]
```

subset()를 활용한 데이터추출

Hide

```
# planets_df is pre-loaded in your workspace
planets_df
# Select planets with diameter < 1
subset(planets_df, subset = diameter < 1)
```

Sorting your data

Hide

```
# planets_df is pre-loaded in your workspace
planets_df
# Use order() to create positions
positions <- order(planets_df$diameter)

# Use positions to sort planets_df
planets_df[positions, ]
```