

```
library(dplyr)
```

Loading a sample dataset

```
df = mtcars
```

Print top five rows

```
head(df, 5)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2

Print bottom five rows

```
tail(df, 5)
```

```
      mpg  cyl  disp  hp  drat    wt  qsec  vs  am  gear  carb
Lotus Europa  30.4   4  95.1 113 3.77 1.513 16.9  1  1    5    2
Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.5  0  1    5    4
Ferrari Dino  19.7   6 145.0 175 3.62 2.770 15.5  0  1    5    6
Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.6  0  1    5    8
Volvo 142E    21.4   4 121.0 109 4.11 2.780 18.6  1  1    4    2
```

Print the names of the columns

```
names(df)
```

```
[1] "mpg"  "cyl"  "disp" "hp"   "drat" "wt"   "qsec" "vs"
[9] "am"   "gear" "carb"
```

Print names of rows

```
rownames(df)
[1] "Mazda RX4"           "Mazda RX4 Wag"       "Datsun 710"
[4] "Hornet 4 Drive"     "Hornet Sportabout"  "Valiant"
[7] "Duster 360"        "Merc 240D"          "Merc 230"
[10] "Merc 280"          "Merc 280C"          "Merc 450SE"
[13] "Merc 450SL"        "Merc 450SLC"        "Cadillac Fleetwood"
[16] "Lincoln Continental" "Chrysler Imperial"  "Fiat 128"
[19] "Honda Civic"       "Toyota Corolla"     "Toyota Corona"
[22] "Dodge Challenger" "AMC Javelin"        "Camaro Z28"
[25] "Pontiac Firebird" "Fiat X1-9"          "Porsche 914-2"
[28] "Lotus Europa"     "Ford Pantera L"     "Ferrari Dino"
[31] "Maserati Bora"    "Volvo 142E"
```

Removing column names

```
unname(df)
head(df, 5)

Mazda RX4           21.0 6 160.0 110 3.90 2.620 16.46 0 1 4 4
Mazda RX4 Wag      21.0 6 160.0 110 3.90 2.875 17.02 0 1 4 4
Datsun 710         22.8 4 108.0  93 3.85 2.320 18.61 1 1 4 1
Hornet 4 Drive     21.4 6 258.0 110 3.08 3.215 19.44 1 0 3 1
Hornet Sportabout 18.7 8 360.0 175 3.15 3.440 17.02 0 0 3 2
```

Print the position of a column by name

```
match("cyl",names(df))  
[1] 2
```

Or

```
which(colnames(df)=="disp" )  
[1] 3
```

Check the columns names with types

```
glimpse(df)  
  
Rows: 32  
Columns: 11  
$ mpg <dbl> 21.0, 21.0, 22.8, 21.4, 18.7, 18.1, 14.3, 24.4, 22.8, 19.2,  
17.8, 16.4, 1...  
$ cyl <dbl> 6, 6, 4, 6, 8, 6, 8, 4, 4, 6, 6, 8, 8, 8, 8, 8, 8, 4, 4, 4,  
4, 8, 8, 8, 8...  
$ disp <dbl> 160.0, 160.0, 108.0, 258.0, 360.0, 225.0, 360.0, 146.7,  
140.8, 167.6, 167...  
$ hp <dbl> 110, 110, 93, 110, 175, 105, 245, 62, 95, 123, 123, 180, 180,  
180, 205, 2...  
$ drat <dbl> 3.90, 3.90, 3.85, 3.08, 3.15, 2.76, 3.21, 3.69, 3.92, 3.92,  
3.92, 3.07, 3...  
$ wt <dbl> 2.620, 2.875, 2.320, 3.215, 3.440, 3.460, 3.570, 3.190,  
3.150, 3.440, 3.4...  
$ qsec <dbl> 16.46, 17.02, 18.61, 19.44, 17.02, 20.22, 15.84, 20.00,  
22.90, 18.30, 18...
```

```
$ vs <dbl> 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1,  
1, 0, 0, 0, 0...  
$ am <dbl> 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1,  
0, 0, 0, 0, 0...  
$ gear <dbl> 4, 4, 4, 3, 3, 3, 3, 4, 4, 4, 4, 3, 3, 3, 3, 3, 4, 4, 4,  
3, 3, 3, 3, 3...  
$ carb <dbl> 4, 4, 1, 1, 2, 1, 4, 2, 2, 4, 4, 3, 3, 3, 4, 4, 4, 1, 2, 1,  
1, 2, 2, 4, 2...
```

Renaming columns with dplyr

```
df<- df |> rename(  
  col1= mpg,  
  col2= cyl,  
  col3= disp)
```

Renaming with base R by column position

```
names(df)[1] <- "col1"  
names(df)[2] <- "col2"  
names(df)[3] <- "col3"
```

Sorting columns

- By name

```
neworder <- c("mpg", "cyl", "disp", "hp", "drat", "wt")
```

```
df <- df[, neworder]
```

```
df
```

	mpg	cyl	disp	hp	drat	wt
Mazda RX4	21.0	6	160.0	110	3.90	2.620
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875
Datsun 710	22.8	4	108.0	93	3.85	2.320
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215

- By column position

```
neworder <- c(7, 3, 5, 6, 2, 3)
```

```
df <- df[, neworder]
```

```
df
```

	qsec	disp	drat	wt	cyl	disp.1
Mazda RX4	16.46	160.0	3.90	2.620	6	160.0
Mazda RX4 Wag	17.02	160.0	3.90	2.875	6	160.0
Datsun 710	18.61	108.0	3.85	2.320	4	108.0
Hornet 4 Drive	19.44	258.0	3.08	3.215	6	258.0

- Alphabetically

```
df[,order(colnames(df))]
```

Subsetting

- Keeping selected columns

```
subset(df, select = c(cyl, carb, hp, wt, vs, am))
```

```
head(df, 5)
```

	cyl	carb	hp	wt	vs	am
Mazda RX4	6	4	110	2.620	0	1
Mazda RX4 Wag	6	4	110	2.875	0	1
Datsun 710	4	1	93	2.320	1	1
Hornet 4 Drive	6	1	110	3.215	1	0
Hornet Sportabout	8	2	175	3.440	0	0

- Dropping selected columns

```
df <- subset(df, select = -c(cyl, carb, hp, wt, vs, am))
```

```
head(df, 5)
```

	mpg	disp	drat	qsec	gear
Mazda RX4	21.0	160.0	3.90	16.46	4
Mazda RX4 Wag	21.0	160.0	3.90	17.02	4
Datsun 710	22.8	108.0	3.85	18.61	4
Hornet 4 Drive	21.4	258.0	3.08	19.44	3
Hornet Sportabout	18.7	360.0	3.15	17.02	3

- Keeping columns matching a certain string in the name

```
df <- select(df,matches("m"))
```

```
names(df)
```

```
[1] "mpg" "am"
```

- Keeping columns matching certain strings in the name

```
df <- select(df,contains(c('m', 'c')))
```

```
head(df, 5)
```

	mpg	am	cyl	qsec	carb
Mazda RX4	21.0	1	6	16.46	4
Mazda RX4 Wag	21.0	1	6	17.02	4
Datsun 710	22.8	1	4	18.61	1
Hornet 4 Drive	21.4	0	6	19.44	1
Hornet Sportabout	18.7	0	8	17.02	2

- Keeping columns matching either of the strings in the name


```
df <- select(df,matches('mlc'))
```

```
head(df, 5)
```

	mpg	cyl	qsec	am	carb
Mazda RX4	21.0	6	16.46	1	4
Mazda RX4 Wag	21.0	6	17.02	1	4
Datsun 710	22.8	4	18.61	1	1
Hornet 4 Drive	21.4	6	19.44	0	1
Hornet Sportabout	18.7	8	17.02	0	2

- **Keeping columns not matching a string in the name**

```
df <- select(df,!matches("c"))
```

```
head(df, 5)
```

	mpg	disp	hp	drat	wt	vs	am	gear
Mazda RX4	21.0	160.0	110	3.90	2.620	0	1	4
Mazda RX4 Wag	21.0	160.0	110	3.90	2.875	0	1	4
Datsun 710	22.8	108.0	93	3.85	2.320	1	1	4
Hornet 4 Drive	21.4	258.0	110	3.08	3.215	1	0	3
Hornet Sportabout	18.7	360.0	175	3.15	3.440	0	0	3

- **Keeping columns if name start with a certain string**

```
df |> select(starts_with('c'))
```

```
head(df, 5)
```

	cyl	carb
Mazda RX4	6	4
Mazda RX4 Wag	6	4
Datsun 710	4	1
Hornet 4 Drive	6	1
Hornet Sportabout	8	2

- **Keeping columns if name ends with a certain string**

```
df |> select(ends_with('t'))
```

```
head(df, 5)
```

	drat	wt
Mazda RX4	3.90	2.620
Mazda RX4 Wag	3.90	2.875
Datsun 710	3.85	2.320
Hornet 4 Drive	3.08	3.215
Hornet Sportabout	3.15	3.440

- **Subsetting based on a condition**

```
df |> filter(cyl==4)
```

```
df
```

	qsec	disp	drat	wt	cyl	disp.1
Datsun 710	18.61	108.0	3.85	2.320	4	108.0
Merc 240D	20.00	146.7	3.69	3.190	4	146.7

Merc 230	22.90	140.8	3.92	3.150	4	140.8
Fiat 128	19.47	78.7	4.08	2.200	4	78.7
Honda Civic	18.52	75.7	4.93	1.615	4	75.7
Toyota Corolla	19.90	71.1	4.22	1.835	4	71.1
Toyota Corona	20.01	120.1	3.70	2.465	4	120.1
Fiat X1-9	18.90	79.0	4.08	1.935	4	79.0
Porsche 914-2	16.70	120.3	4.43	2.140	4	120.3
Lotus Europa	16.90	95.1	3.77	1.513	4	95.1
Volvo 142E	18.60	121.0	4.11	2.780	4	121.0

- **Subsetting based on multiple conditions**

```
df |> filter(cyl==4 & drat>=4)
```

```
df
```

	qsec	disp	drat	wt	cyl	disp.1
Fiat 128	19.47	78.7	4.08	2.200	4	78.7
Honda Civic	18.52	75.7	4.93	1.615	4	75.7
Toyota Corolla	19.90	71.1	4.22	1.835	4	71.1
Fiat X1-9	18.90	79.0	4.08	1.935	4	79.0
Porsche 914-2	16.70	120.3	4.43	2.140	4	120.3
Volvo 142E	18.60	121.0	4.11	2.780	4	121.0

Three conditions

```
df |> filter(cyl<=5 & drat==4 | qsec >19)
```

```
df
```

	qsec	disp	drat	wt	cyl	disp.1
Hornet 4 Drive	19.44	258.0	3.08	3.215	6	258.0
Valiant	20.22	225.0	2.76	3.460	6	225.0
Merc 240D	20.00	146.7	3.69	3.190	4	146.7
Merc 230	22.90	140.8	3.92	3.150	4	140.8
Fiat 128	19.47	78.7	4.08	2.200	4	78.7
Toyota Corolla	19.90	71.1	4.22	1.835	4	71.1
Toyota Corona	20.01	120.1	3.70	2.465	4	120.1