

Data Manipulation with R: Basics: Takeaways



by Dataquest Labs, Inc. - All rights reserved © 2021

Syntax

- Install package

```
install.packages("name_of_the_package")
```

- Load a package

```
library(name_of_the_package_without_quotation_marks)
```

- Import a csv-file into R with the read_csv() function from readr package

```
library(readr_r)
new_data_frame <- read_csv("name_of_the_dataset.csv")
```

- Find the number of columns in a dataset

```
n_col <- ncol(new_data_frame)
```

- Find the number of rows in a dataset

```
n_row <- nrow(new_data_frame)
```

- Find the name of columns in a dataset

```
data_frame_names <- colnames(new_data_frame)
```

- Display the first six rows in a dataset

```
head_rows <- head(new_data_frame)
```

- Display the last six rows in a dataset

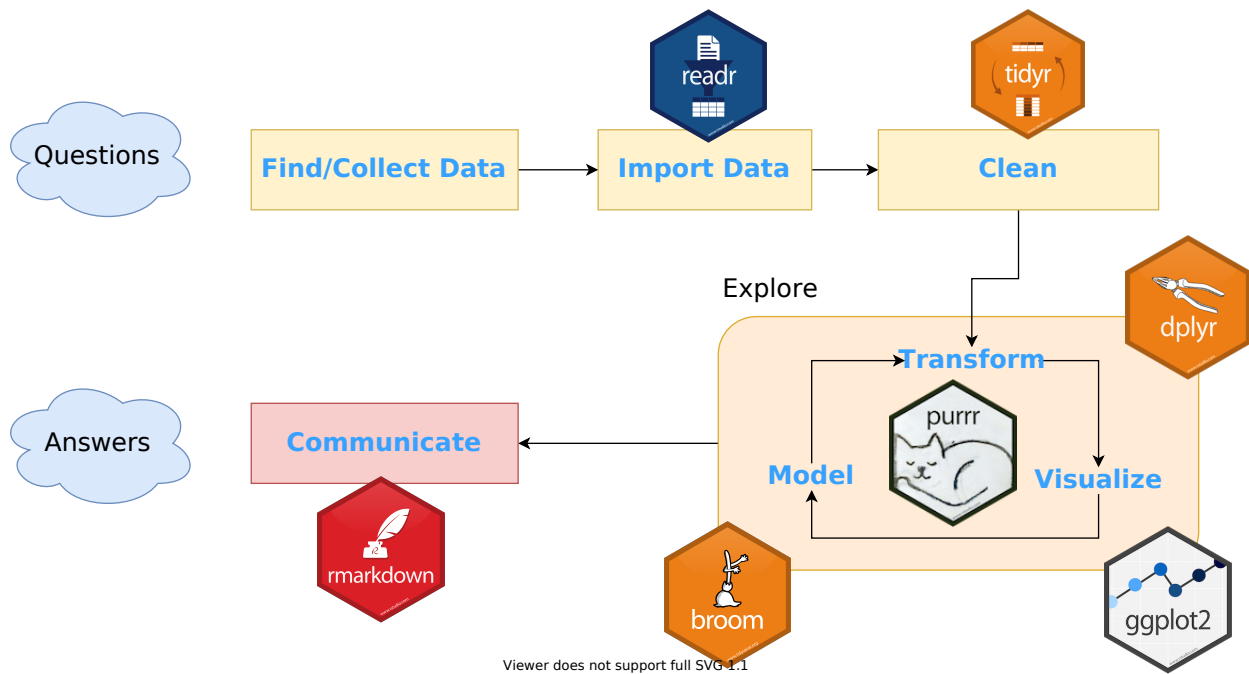
```
tail_rows <- tail(new_data_frame)
```

- Visualize a dataset with the qplot() function from ggplot2 package

```
library(ggplot2)
qplot(x = job_id,
      y = salary_max,
      color = job_type,
      data = monster_jobs_clean)
```

Concepts

- Data analysis workflow and tidyverse collection of packages.



- We can import several types of datasets into R using the

`readr`

package namely CSV (Comma-separated Values) and TSV (Tab-separated Values) files.

- We can use the package

`ggplot2`

to visualize our dataset.

Resources

- [CRAN repository](#)
- [readr documentation](#)
- [ggplot2 documentation](#)