A Beginner's Guide to Tidyverse

The Most Powerful Collection of R Packages for Data Science



Ready to explore the tidyverse? Go ahead and install it directly from within RStudio: install.packages("tidyverse")

What is Tidyverse?

Tidyverse is a collection of essential R packages for *data science*. The packages under the tidyverse umbrella help us in performing and interacting with the data. There are a whole host of things you can do with your data, such as *subsetting*, *transforming*, *visualizing*, etc.

Core R Packages in Tidyverse

Data Wrangling and Transformation

Data Import and Management

Functional Programming

- . d.b./v
- dplyr
- tidyr
- stringrforcats

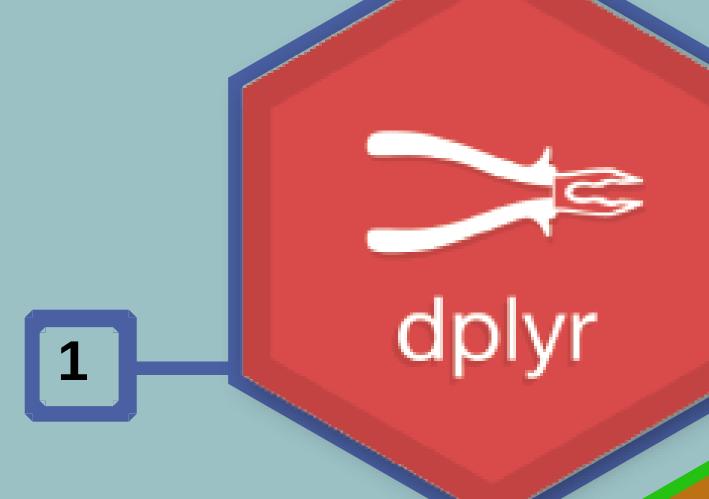
tibble

readr

- purrr
- Dete

Data Visualization and Exploration

ggplot2



dplyr is one of my all-time favorite packages. It is simply the most useful package in R for data manipulation. One of the greatest advantages of this package is you can use the pipe function "%>%" to combine different functions in R. From filtering to grouping the data, this package does it all.



The **tidyr** package complements dplyr perfectly. It boosts the power of dplyr for data manipulation and pre-processing.



stringr is used for strings. The stringr package provide a cohesive set of functions designed to make working with strings as easy as possible.



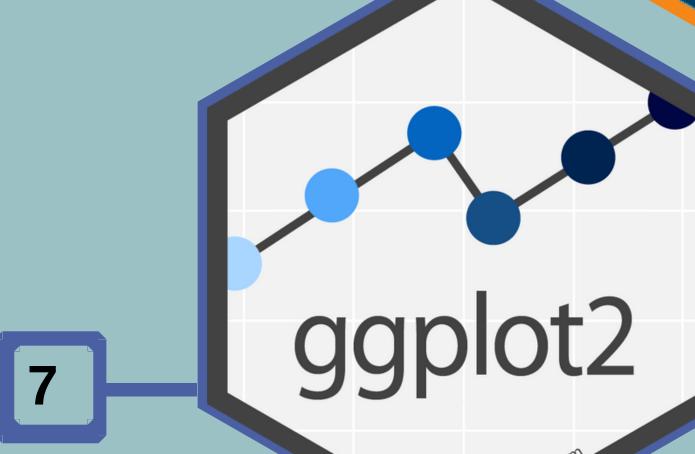
The **forcats** package is dedicated to dealing with categorical variables or factors. Anyone who has worked with categorical data knows what a nightmare they can be. forcats feels like a godsend.



Tibble is a type of dataframe in R. It truly stands out when we're trying to detect anomalies in our dataset. How? Tibble does not change variable names or types. It certainly doesn't throw up errors when a variable does not exist or a value is missing.



This package is very useful when you want to import Excel sheets in R



6

Data scientists universally love using **ggplot2** to produce their charts and visualizations. It's such a useful and popular package that they've integrated it into the Python language!



The **purrr** package in R provides a complete toolkit for enhancing R's functional programming. We can use the functions provided by purrr to avoid many loops with just one line of code.



Magrittr let's you transform nested function calls into a simple pipeline of operations that's easier to write and understand.

The **magrittr** package offers a set of operators which make your code more readable by:

- structuring sequences of data operations left-to-right (as opposed to from the inside and out),
- avoiding nested function calls,

All the implementation part has been given in the article itself Link: https://bit.ly/2YwUp6O

For more amazing infographics, Visit www.analyticsvidhya.com

