

# Tidyverse Cheat Sheet

## Definition:

A powerful collection of R packages that are data tools for transforming and visualizing data.

## Core Tidyverse packages:



**ggplot2:** Use for data visualization.

Implements the grammar of graphics

**dplyr:** Implements the grammar of data manipulation.

**tidyr:** Creates tidy data or data where each variable is in a column, each observation is a row, and each value is a cell.

**readr:** Reads rectangular data

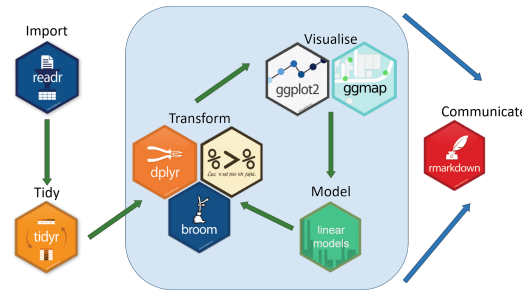
**purrr:** enhances R's functional programming toolkit.

**tibble:** Reimages the data frame.

**stringr:** Simplifies the manipulation of strings.

**forcats:** Solves problems with factors.

## Workflow:



## Install & Load Tidyverse:

```
> install.packages("tidyverse")
```

Or

```
> devtools::install_github("tidyverse/tidyverse")
```

```
> library(tidyverse)
```

## Update Tidyverse packages:

```
> tidyverse_update()
```

## ggplot2:

```
> ggplot(data = <DATA>) +  
<GEOM_FUNCTION>(mapping=aes(<MAPPINGS>))
```

## Typical ggplot2 types & examples:

### Scatter plot with Facets:

```
> ggplot(data = mpg)  
+ geom_point(mapping = aes(x = displ, y = hwy))  
+ facet_wrap(~ class, nrow = 2)
```

### Line Plot:

```
> ggplot(by_year, aes(x=year, y=median))  
+ geom_line()
```

### Barplot:

```
> ggplot(data=diamonds) +  
+  
geom_bar(aes(x=cut,fill=factor(Expensive  
)))
```

## Histograms:

```
> ggplot(iris_small, aes(x=Petal.Length))  
+ geom_histogram()
```

## Box Plot:

```
> ggplot(iris_small, aes(x=Species,  
y=Sepal.Width))  
+ geom_boxplot()
```

## dplyr:

### Typical dplyr functions:

> mutate():

Adds new variables that are functions of existing variables

> select():

Picks variables based on their names.

> filter():

Picks cases based on their values.

> summarise():

Reduces multiple values to a single summary.

> arrange():

Changes the ordering of the rows.

## readr:

### Typical functions to read data:

> read\_csv():

Reads comma separated (CSV) files

> read\_delim():

Reads general delimited files

> read\_fw():

Reads fixed width files

> read\_table():

Reads tabular files where columns are separated by white-space

> read\_log():

Reads web log files