

The background of the slide is a dense, 3D-rendered field of numbers. The numbers are in various shades of blue and white, creating a sense of depth and movement. They are scattered across the entire frame, with some numbers appearing larger and more prominent than others. The overall effect is a vibrant, data-driven aesthetic.

Tidymodels packages in R

Bilikisu Wunmi Aderinto

Bilikisu Wunmi Aderinto

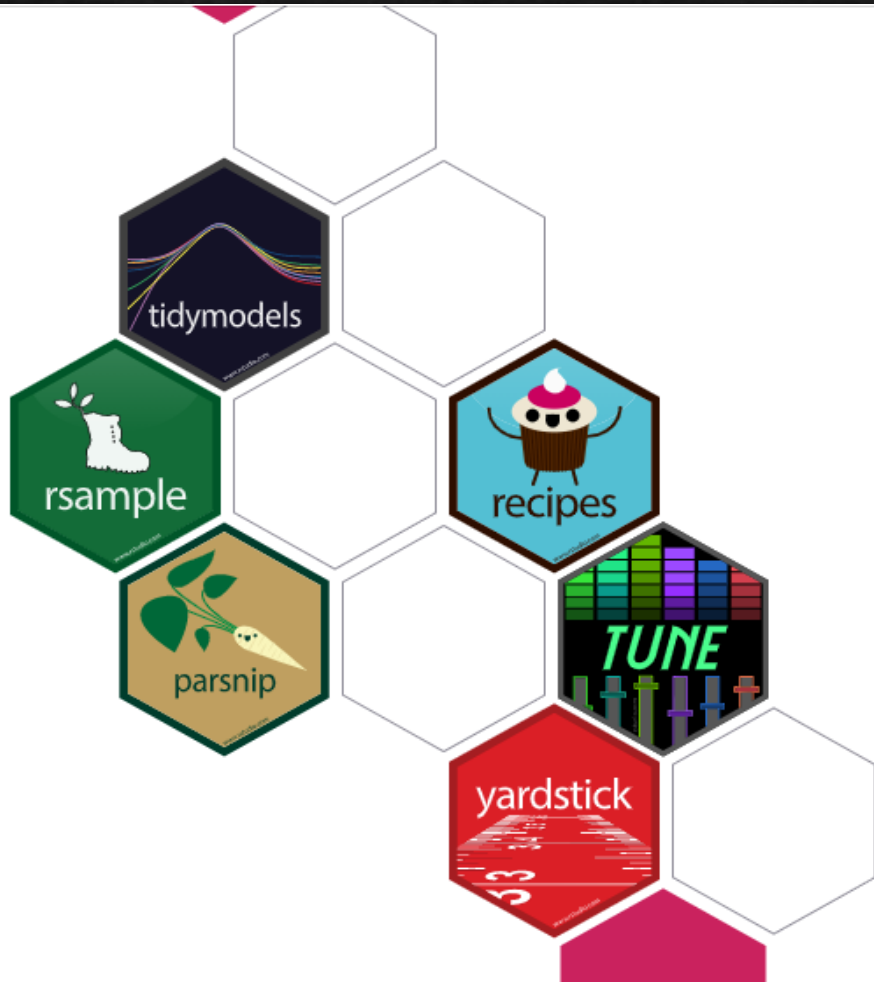
Business Data Analyst, Business Data Lab

<https://www.linkedin.com/in/bilikisuaderinto/>

[Bilikisu W. Adérìntó \(@qbwoa\) / Twitter](#)



Tidymodels



TIDYMODELS

The tidymodels framework is a collection of packages for modeling and machine learning using **tidyverse** principles.

Install tidymodels with:

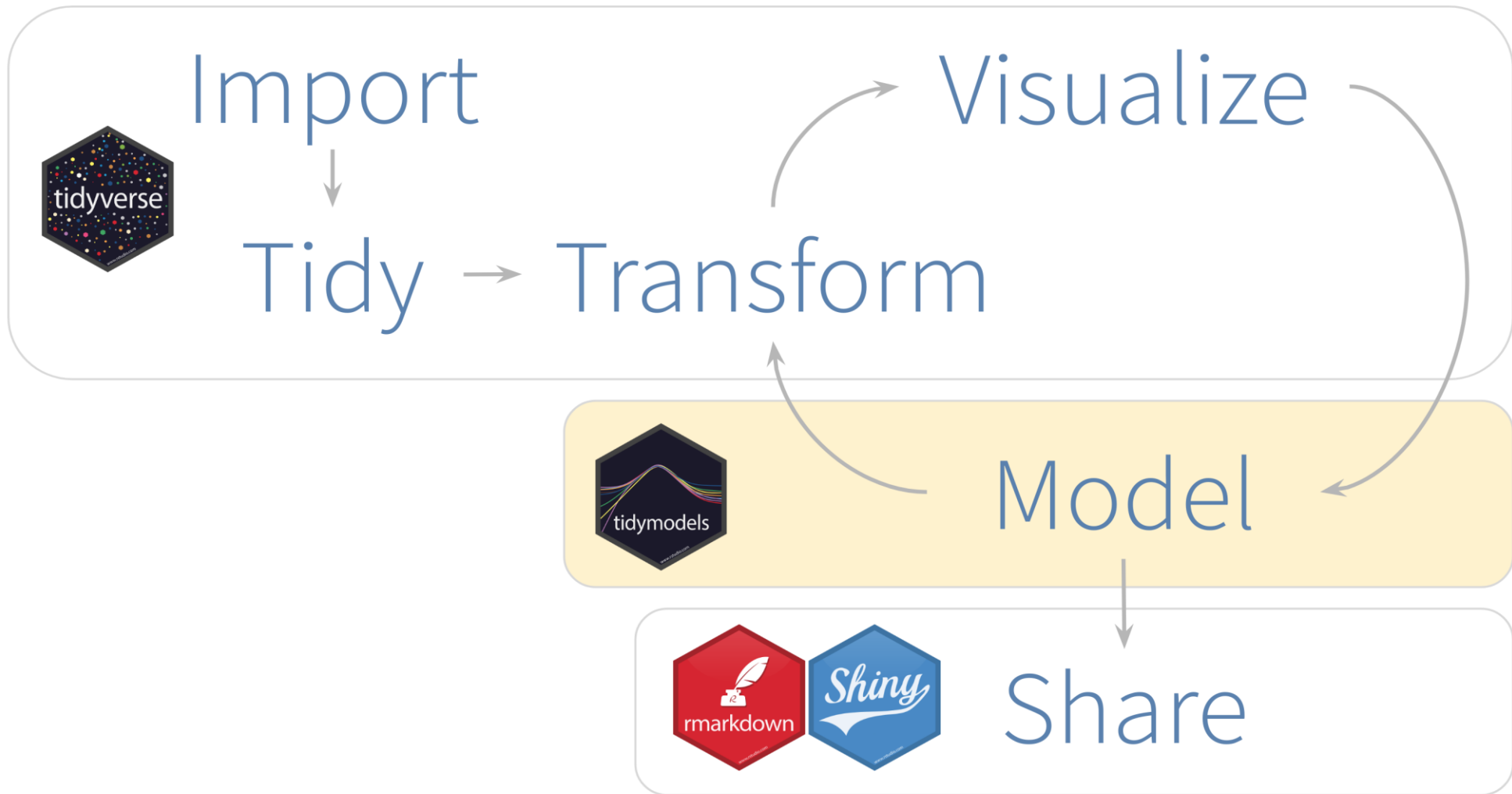
```
install.packages("tidymodels")
```


What is the goal of machine learning?

Goal

Build **models** that
generate **accurate predictions**
for **future, yet-to-be-seen data**.

Tidymodels



Tidymodels

Pre-Process → Train → Validate



Tidymodels

```
library(tidymodels)
### Warning: package 'tidymodels' was built under R version 3.6.2
### — Attaching packages ————— tidymod
### ✓ broom      0.7.0      ✓ recipes    0.1.13
### ✓ dials      0.0.8      ✓ rsample    0.0.7
### ✓ dplyr      1.0.2      ✓ tibble     3.0.3
### ✓ ggplot2    3.3.2      ✓ tidyr      1.1.1
### ✓ infer      0.5.3      ✓ tune       0.1.1
### ✓ modeldata  0.0.2      ✓ workflows  0.1.3
### ✓ parsnip    0.1.3      ✓ yardstick  0.0.7
### ✓ purrr      0.3.4
### Warning: package 'broom' was built under R version 3.6.2
### Warning: package 'dials' was built under R version 3.6.2
### Warning: package 'scales' was built under R version 3.6.2
### Warning: package 'dplyr' was built under R version 3.6.2
### Warning: package 'ggplot2' was built under R version 3.6.2
### Warning: package 'infer' was built under R version 3.6.2
### Warning: package 'modeldata' was built under R version 3.6.2
### Warning: package 'parsnip' was built under R version 3.6.2
```

rsample package

- ◆ Split data into training and test data

```
set.seed(923)  
bankmarketing_split <- initial_split(bankmarketing_cleaned, strata = y)
```

<https://rsample.tidymodels.org/>

recipes package

- ◆ Define the recipe that will be used for preprocessing data needed for modeling

```
[ ]  
bankmarketing_recipe <- bankmarketing_train %>%  
  recipe(y ~.) %>%  
  step_corr(all_numeric()) %>%  
  step_normalize(all_numeric()) %>%  
  step_dummy(all_nominal(), -y) %>%  
  step_smote(y)
```

<https://www.tidymodels.org/find/recipes/>

parsnip package

◆ Specifying a model definition with parsnip

1. Select a **model**
2. Set the **engine**
3. Set the **mode**

parsnip package

```
ranger_spec <- rand_forest() %>%  
  set_mode("classification") %>%  
  set_engine("ranger")
```

<https://www.tidymodels.org/find/parsnip/>

tune package

<https://tune.tidymodels.org/>

yardstick package

<https://yardstick.tidymodels.org/>

Thank you.

<https://www.linkedin.com/in/bilikisuaderinto/>
Bilikisu W. Adérìntó (@qbwoa) / Twitter

References

- ◆ Julia Silge, <https://juliasilge.com/>
- ◆ Alison Hill, <https://alison.rbind.io/post/2020-06-02-tidymodels-virtually/>
- ◆ Rebecca Barter, <http://www.rebeccabarter.com/categories/tidymodels/>
- ◆ Tidy modeling with R, <https://www.tmwr.org/>
- ◆ A Gentle Introduction to Tidymodels, <https://rviews.rstudio.com/2019/06/19/a-gentle-intro-to-tidymodels/>