tidyverse（再）入門

7th December 2019, Japan.R 2019
Yuta Kanzawa @yutakanzawa

SFE Senior Analyst at Janssen Pharmaceutical K.K., Tokyo
A Family Company of Johnson & Johnson

I am...

- 神沢雄大 Yuta Kanzawa (twitter: @yutakanzawa)
- Data scientist at Janssen Japan, Tokyo
- Opera & wine lover
  - Wagner
  - Bourgogne
- 7 languages
  - Human: Japanese, English, German
  - Computer: R, Python, SAS, SQL

クッキーをデザインしました🍪

Japan.R @Japan_R_net · Nov 20

Japan.R オリジナルのクッキーを作りました。会場に早く来た方限定です！（全員分は用意してないです。）写真を取るだけでもテンション上がるかも。１年に1度のJapan.R を楽しみましょう！

@yutakanzawa さんがデザインしてくれました。

「クッキー製作・写真撮影：（株）デコレ」

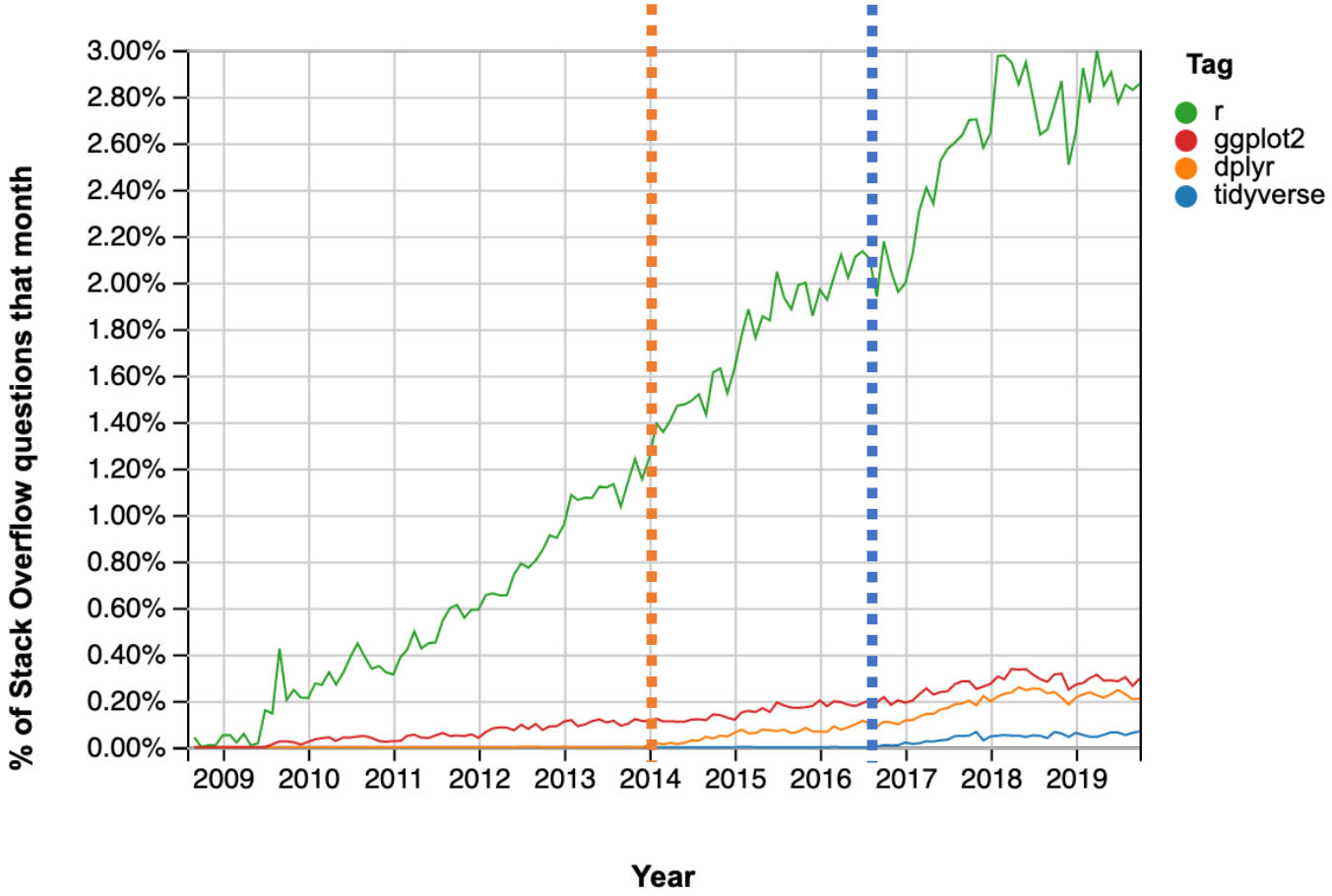
1　19　47

アジェンダ

- 今日話すこと
  - tidyverseと私たち

- 今日話さないこと
  - tidyverseの詳細（参照：宇宙本）
  - Python

tidyverse native



* https://insights.stackoverflow.com/trends?tags=tidyverse%2Cdplyr%2Cr%2Cggplot2

Base R native

- [演算子、$演算子
- sort(), melt(), dcast()
- aggregate()

Things are changing...

e.g. *Analyzing Baseball Data with R* (Marchi, Albert)*1

1st Edition (October 2013)*2

2 Introduction to R 29

2.1 Introduction . . . 30
2.2 Installing R and RStudio . . . 30
2.3 Vectors . . . 31
2.3.1 Career of Warren Spahn . . . 31
2.3.2 Vectors: defining and calculations . . . 32
2.3.3 Vector functions . . . 34
2.3.4 Vector index and logical variables . . . 35
2.4 Objects and Containers in R . . . 36
2.4.1 Character data and matrices . . . 37
2.4.2 Factors . . . 38
2.4.3 Lists . . . 40
2.5 Collection of R Commands . . . 41
2.5.1 R scripts . . . 41
2.5.2 R functions . . . 42
2.6 Reading and Writing Data in R . . . 43
2.6.1 Importing data from a file . . . 43
2.6.2 Saving datasets . . . 45
2.7 Data Frames . . . 45
2.7.1 Introduction . . . 45
2.7.2 Manipulations with data frames . . . 47
2.7.3 Merging and selecting from data frames . . . 49
2.8 Packages . . . 50
2.9 Splitting, Applying, and Combining Data . . . 50
2.9.1 Using `sapply` . . . 51
2.9.2 Using `ddply` in the `plyr` package . . . 52
2.10 Getting Help . . . 54
2.11 Further Reading . . . 55
2.12 Exercises . . . 55

2nd Edition (November 2018)*3

Chapter 2 ▪ Introduction to R 33

2.1 Introduction 33
2.2 Installing R and RStudio 33
2.3 The Tidyverse 34
2.3.1 `dplyr` 35
2.3.2 The pipe 35
2.3.3 `ggplot2` 36
2.3.4 Other packages 36
2.4 Data Frames 36
2.4.1 Career of Warren Spahn 36
2.4.2 Introduction 37
2.4.3 Manipulations with data frames 39
2.4.4 Merging and selecting from data frames 41
2.5 Vectors 42
2.5.1 Defining and computing with vectors 42
2.5.2 Vector functions 43
2.5.3 Vector index and logical variables 45
2.6 Objects and Containers in R 46
2.6.1 Character data and data frames 47
2.6.2 Factors 49
2.6.3 Lists 51
2.7 Collection of R Commands 52
2.7.1 R scripts 52
2.7.2 R functions 53

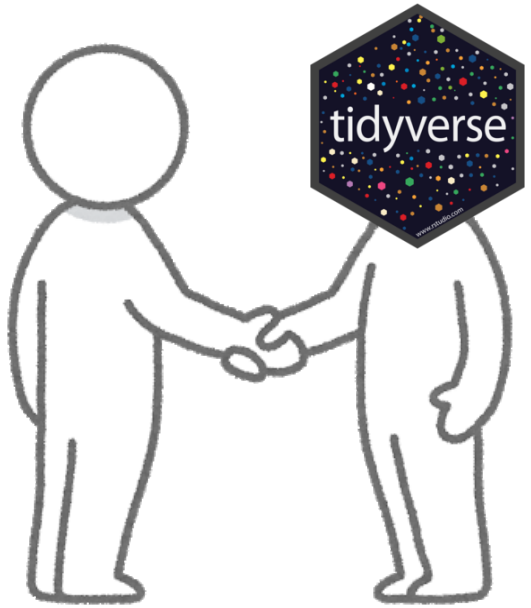
*1 https://shinyorke.hatenablog.com/#fn-bf977ec1
*2 https://www.amazon.com/Analyzing-Baseball-Data-Chapman-Hall/dp/1466570229
*3 https://www.amazon.com/Analyzing-Baseball-Data-Second-Chapman/dp/0367024861

To be tidyverse adaptive

- tidyverseの便利なパッケージや関数を使ってみる。
  - e.g.1 dplyr::coalesce()
  - e.g.2 tidyr::separate()

- パイプ演算子%>%の記法に慣れる。

- レガシーコードを書き換えてみる。
  - 正解があるので、取り組みやすいはず。



詳しくは
R Advent Calendar 2019で！
https://qiita.com/advent-calendar/2019/rlang

Enjoy!