

R Learners

6-20-2018

Intro

- Mike Feyder
- feyderm@gmail.com
- Brain Modulation Lab
- https://github.com/feyderm/Pitt_R_lectures



Base R vs. Tidyverse

- Base R
 - The R language - does not depend on any packages (i.e. external code)
 - Ex.
 - `df$variable`
 - `df[c(1:4),]`

Base R vs. Tidyverse

- “[An] opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.”



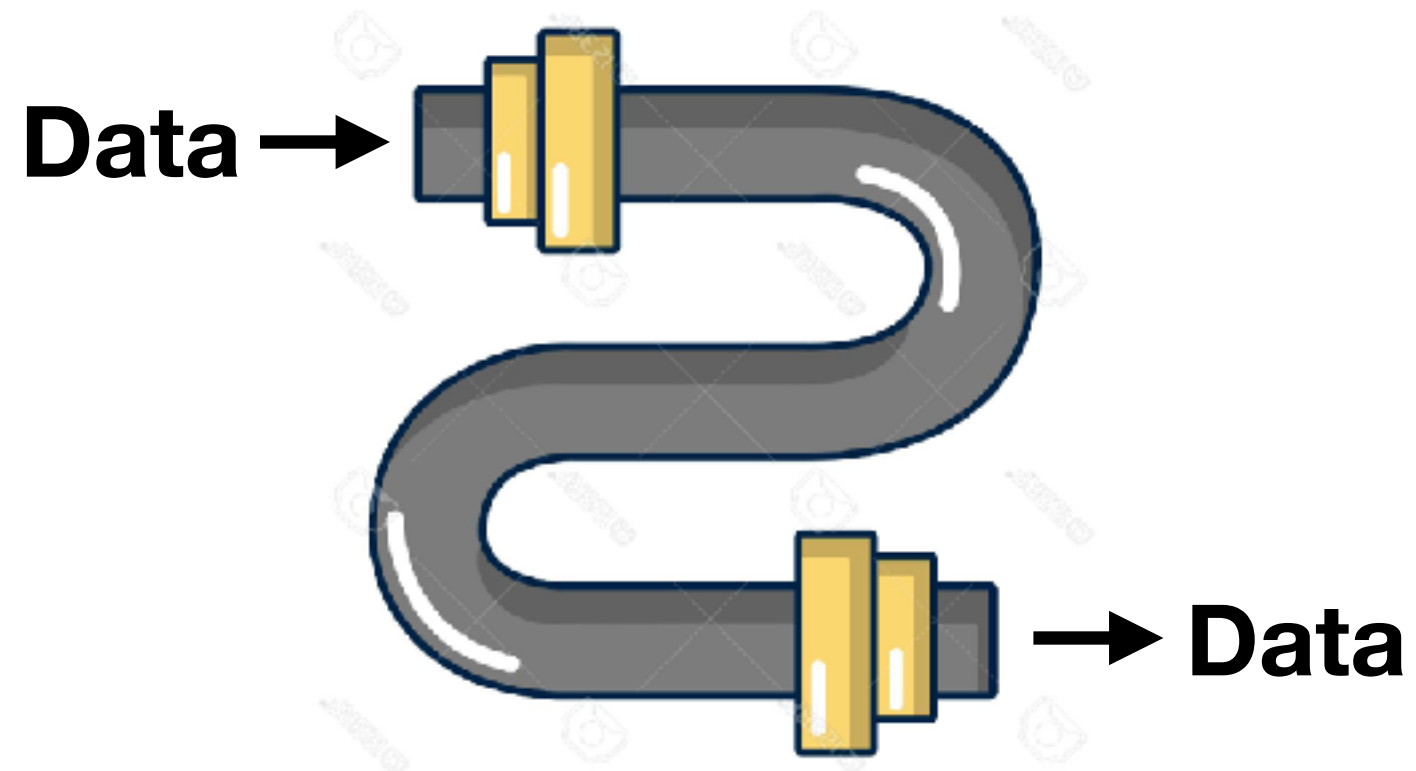
<http://r4ds.had.co.nz/>

Base R vs. Tidyverse

- Selecting a column
 - **Base R:** `df$variable`
 - **Tidyverse:** `df %>% select(variable)`



Piping



Piping

```
df %>%  
  filter(var_1 < 10) %>%  
  select(var_2, var_3)
```

Dplyr

- “a grammar of data manipulation”



Summarizing Data

```
experiment %>%  
  summarize(mean_RT = mean(RT) )
```

Summarizing Data

```
experiment %>%  
  group_by(SerialPosition) %>%  
  summarize(mean_RT = mean(RT))
```

Summarizing Data

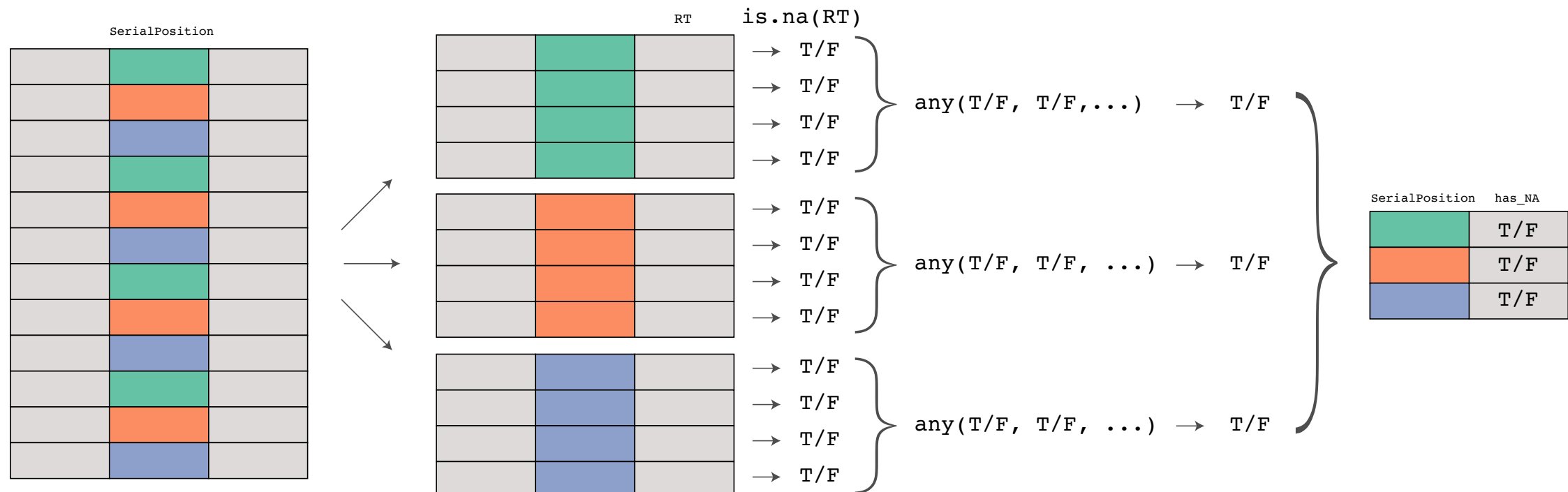
```
experiment %>%  
  group_by(SerialPosition, Condition) %>%  
  summarize(mean_RT = mean(RT))
```

Summarizing Data

```
experiment %>%  
  group_by(SerialPosition) %>%  
  summarize(  
    n = n(),  
    min_RT = min(RT),  
    max_RT = max(RT),  
    mean_RT = mean(RT),  
    median_RT = median(RT),  
    sd_RT = sd(RT),  
    var_RT = var(RT),  
    has_NA = any(is.na(RT))  
  )
```

Summarizing Data

```
Experiment %>% group_by(SerialPosition) %>% summarize(has_NA = any(is.na(RT)))
```



Useful Things

```
experiment <- experiment %>%  
  rename(item_name = ItemName)
```

Useful Things

```
experiment %>% arrange(RT)
```

ggplot2

- data visualization package



<https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf>

ggplot2

- based on *The Grammar of Graphics*



Histogram

```
ggplot(data = experiment, mapping = aes(x = RT)) +  
  geom_histogram()
```



Scatterplot

```
ggplot(data = experiment, mapping = aes(x = Age, y = RT)) +  
  geom_point()
```



Bar Chart

```
experiment %>%  
  group_by(Age) %>%  
  summarize(mean_RT = mean(RT)) %>%  
  ggplot(mapping = aes(x = Age, y = mean_RT)) +  
  geom_col()
```



geoms

```
ggplot(data = experiment, mapping = aes(Age, RT)) +  
  geom_point()
```

```
ggplot(data = experiment, mapping = aes(Age, RT)) +  
  geom_boxplot()
```

```
ggplot(data = experiment, mapping = aes(Age, RT)) +  
  geom_violin()
```

Facets

```
ggplot(data = experiment, mapping = aes(Age, RT)) +  
  geom_point() +  
  facet_wrap(~TestingRoom)
```