# Week 12

*Business Research Methods*

**Bhaswar Chakma**

11 May 2021

# Objectives

- Create "Toy Data"
  - `data.frame()`
  - `tibble()`
  - `tribble()`
- Reproducible example
- Export
- RMD:
  - tables: regression; statistics
  - equations

# “Toy Data”

- `data.frame` ?
- `tibble`?

# data.frame

```
class(iris)
```

```
## [1] "data.frame"
```

```
iris
```

```
    Sepal.Length Sepal.Width Petal.Length
1            5.1         3.5          1.4
2            4.9         3.0          1.4
3            4.7         3.2          1.3
4            4.6         3.1          1.5
5            5.0         3.6          1.4
...
```

# tibble

```
iris %>% as_tibble()
```

```
# A tibble: 150 x 5
   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
          <dbl>       <dbl>        <dbl>       <dbl> <fct>
 1          5.1         3.5          1.4         0.2 setosa
 2          4.9         3            1.4         0.2 setosa
 3          4.7         3.2          1.3         0.2 setosa
 4          4.6         3.1          1.5         0.2 setosa
 5          5           3.6          1.4         0.2 setosa
 6          5.4         3.9          1.7         0.4 setosa
 7          4.6         3.4          1.4         0.3 setosa
 8          5           3.4          1.5         0.2 setosa
 9          4.4         2.9          1.4         0.2 setosa
10          4.9         3.1          1.5         0.1 setosa
# ... with 140 more rows
```

More details:

https://r4ds.had.co.nz/tibbles.html

# data.frame()

```
df1 <- data.frame(
  guitarist = c("Joe Satriani", "Eric Johnson"),
  year = c(1987, 1990),
  song = c("Always With Me, Always With You",
           "Cliffs of Dover")
)
```

```
class(df1)
```

```
## [1] "data.frame"
```

```
df1
```

```
##      guitarist year                              song
## 1 Joe Satriani 1987 Always With Me, Always With You
## 2 Eric Johnson 1990                  Cliffs of Dover
```

# tibble()

```
df2 <- tibble(
  guitarist = c("Joe Satriani", "Eric Johnson"),
  year = c(1987, 1990),
  song = c("Always With Me, Always With You",
           "Cliffs of Dover")
)
```

```
class(df2)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

```
df2
```

```
## # A tibble: 2 x 3
##   guitarist     year song
##   <chr>        <dbl> <chr>
## 1 Joe Satriani  1987 Always With Me, Always With You
## 2 Eric Johnson  1990 Cliffs of Dover
```

# tribble()

```
df3 <- tribble(
  ~ guitarist,          ~ year, ~ song,
  "Joe Satriani", 1987, "Always With Me, Always With You",
  "Eric Johnson", 1990, "Cliffs of Dover"
)
```

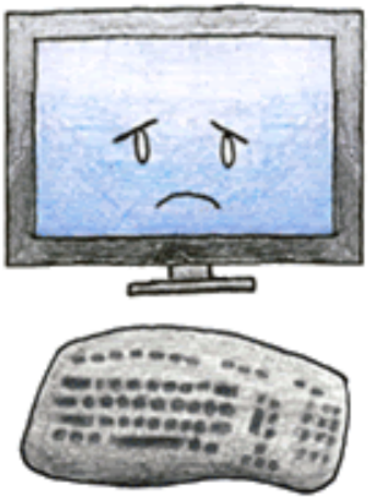
```
class(df3)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

```
df3
```

```
## # A tibble: 2 x 3
##   guitarist     year song
##   <chr>        <dbl> <chr>
## 1 Joe Satriani  1987 Always With Me, Always With You
## 2 Eric Johnson  1990 Cliffs of Dover
```

# Reproducible Example



# RStudio Community

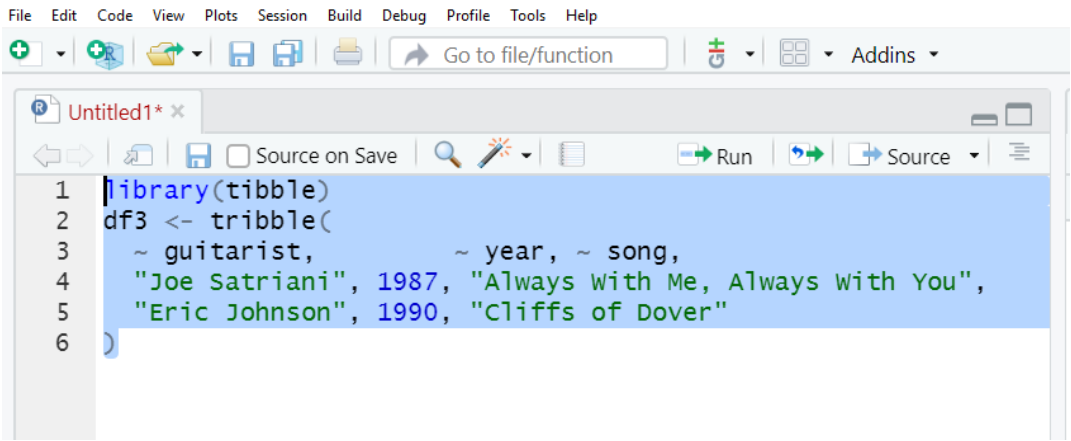
All things RStudio

FIND HELP

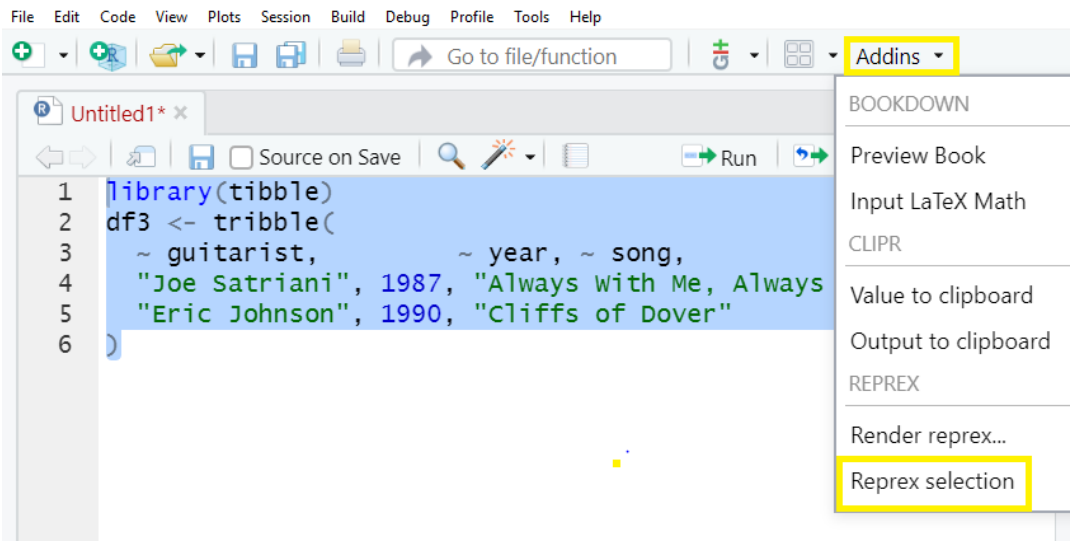
CONTRIBUTE

EXPLORE

1. Select your code



2. Click `Addins` then `Reprex selection`



https://community.rstudio.com/t/test-post-brm-lab-2021-please-do-not-reply/104182

# Test Post: BRM Lab 2021 (Please do not reply)

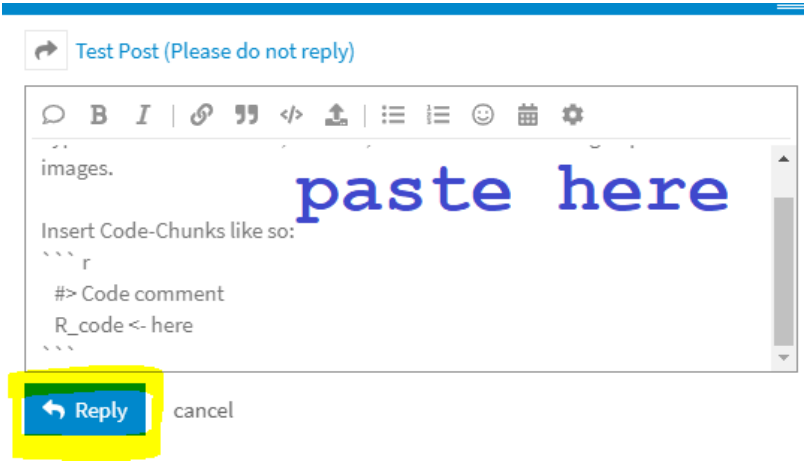
R

r-lab

1 20h

FAQ: What's a reproducible example (`reprex`) and how do I create one? meta

Why reprex? Getting unstuck is hard. Your first step here is usually to create a reprex, or reproducible example. The goal of a reprex is to package your code, and information about your problem so that others can run it and feel your pain. Then, hopefully, folks can more easily provide a solution. What's in a Reproducible Example? Parts of a reproducible example: background information - Describe what you are trying to do. What have you already done? complete set up - include any library() calls and data to reproduce your issue. data for a reprex: Here's a discussion on setting up data for a reprex make it run - include the minimal code required to reproduce your error on the data p…


Test Post (Please do not reply)
images.
paste here
Insert Code-Chunks like so:
```r
#> Code comment
R_code <- here
```
Reply
cancel

# Export("write")

Boss: Give me

- an Excel file.
- a CSV file.

Colleague (who speaks R):

- Give me your data.

```
df <- gapminder::gapminder %>%
  filter(year >= 2000) %>%
  group_by(year, continent) %>%
  summarise(mean_life_exp = mean(lifeExp)) %>%
  pivot_wider(
    names_from = year,
    values_from = mean_life_exp
  )
```

```
df
```

```
## # A tibble: 5 x 3
##   continent `2002` `2007`
##   <fct>      <dbl>  <dbl>
## 1 Africa      53.3   54.8
## 2 Americas    72.4   73.6
## 3 Asia        69.2   70.7
## 4 Europe      76.7   77.6
## 5 Oceania     79.7   80.7
```

```r
# CSV using readr package
readr::write_csv(df, "data/boss.csv")
```

```r
# Excel using the writexl package
writexl::write_xlsx(df, "data/boss.xlsx")
```

```r
# RDS
saveRDS(df, "data/colleague.RDS")
```

# Import("read")

```
readr::read_csv("data/boss.csv")
```

```
## # A tibble: 5 x 3
##   continent `2002` `2007`
##   <chr>      <dbl>  <dbl>
## 1 Africa      53.3   54.8
## 2 Americas    72.4   73.6
## 3 Asia        69.2   70.7
## 4 Europe      76.7   77.6
## 5 Oceania     79.7   80.7
```

```
readxl::read_excel("data/boss.xlsx")
```

```
## # A tibble: 5 x 3
##   continent `2002` `2007`
##   <chr>      <dbl>  <dbl>
## 1 Africa      53.3   54.8
## 2 Americas    72.4   73.6
## 3 Asia        69.2   70.7
## 4 Europe      76.7   77.6
## 5 Oceania     79.7   80.7
```

```
readRDS("data/colleague.RDS")
```

```
## # A tibble: 5 x 3
##   continent `2002` `2007`
##   <fct>      <dbl>  <dbl>
## 1 Africa      53.3   54.8
## 2 Americas    72.4   73.6
## 3 Asia        69.2   70.7
## 4 Europe      76.7   77.6
## 5 Oceania     79.7   80.7
```

# Survey

- S1: between 17:00 and 18:30.
- S2: between 15:30 and 17:00.

# Regression

```
install.packages("stargazer")
```

We will use 25.8 An Example

Get the `Tables-Equations.Rmd` from here

# Table 1: Results

| | *Dependent variable:* | | |
|---|---|---|---|
| | low | | |
| | *OLS* | *logistic* | *probit* |
| | (1) | (2) | (3) |
| smoke | $0.186^{***}$ | $1.041^{***}$ | $0.635^{***}$ |
| | (0.071) | (0.391) | (0.228) |
| race | $0.081^{**}$ | $0.471^{**}$ | $0.281^{**}$ |
| | (0.039) | (0.213) | (0.124) |
| ht | $0.377^{***}$ | $1.851^{***}$ | $1.110^{***}$ |
| | (0.137) | (0.690) | (0.414) |
| ui | $0.188^{**}$ | $0.867^{*}$ | $0.537^{**}$ |
| | (0.092) | (0.451) | (0.274) |
| ftv | 0.005 | 0.056 | 0.026 |
| | (0.031) | (0.169) | (0.100) |
| age | −0.004 | −0.027 | −0.017 |
| | (0.006) | (0.035) | (0.021) |
| lwt | $-0.002^{**}$ | $-0.014^{**}$ | $-0.008^{**}$ |
| | (0.001) | (0.007) | (0.004) |
| Constant | $0.424^{*}$ | −0.118 | −0.087 |
| | (0.230) | (1.264) | (0.744) |

```r
# ols
m0 <- lm(low ~ smoke + race + ht + ui + ftv + age + lwt,
         data = MASS::birthwt)
# logit
m1 <- glm(low ~ smoke + race + ht + ui + ftv + age + lwt,
          data = MASS::birthwt,
          family = binomial(link = "logit"))
# probit
m2 <- glm(low ~ smoke + race + ht + ui + ftv + age + lwt,
          data = MASS::birthwt,
          family = binomial(link = "probit"))
```

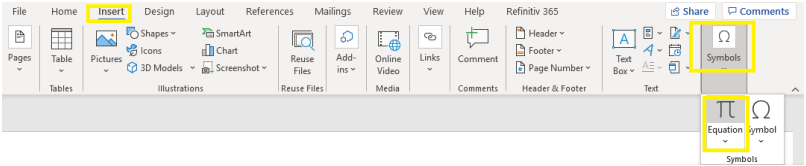
```
```{r regtable, results='asis'}
stargazer::stargazer(m0, m1, m2,
                     type = "latex", # "html" for html
                     title = "Results",
                     header = FALSE,
                     label = "tab:regression1")
```
```

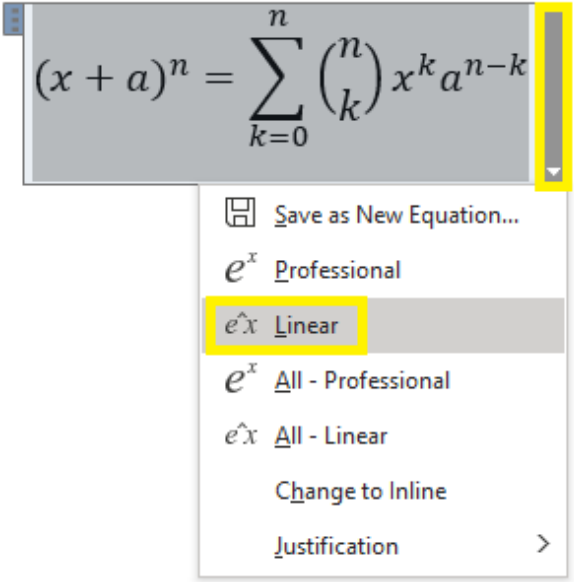
**Important:**

- `results='asis'` must be included
- `label = "tab:regression1"` is needed for cross-referencing
  - You can ~~regression1~~; use something else

# Equations

We need LaTeX!


File
Home
Insert
Design
Layout
References
Mailings
Review
View
Help
Refinitiv 365
Share
Comments
Pages
Table
Tables
Pictures
Shapes
Icons
3D Models
SmartArt
Chart
Screenshot
Illustrations
Reuse Files
Reuse Files
Add-ins
Online Video
Media
Links
Comment
Comments
Header
Footer
Page Number
Header & Footer
Text Box
Text
Symbols
Equation
Symbol
Symbols


$(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$
Save as New Equation...
Professional
Linear
All - Professional
All - Linear
Change to Inline
Justification

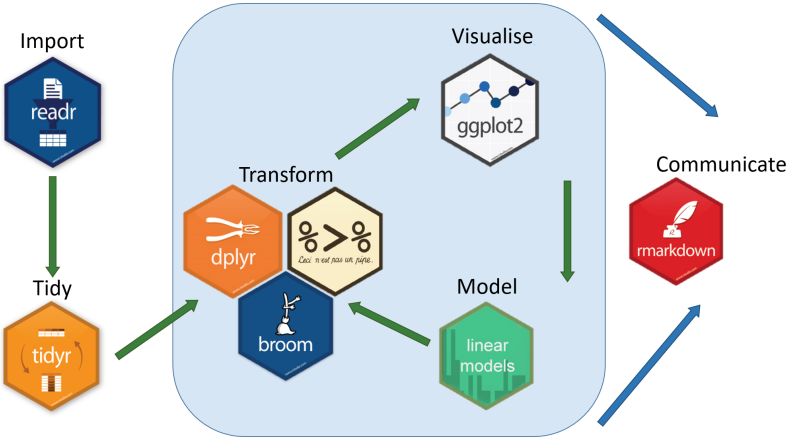
Copy!

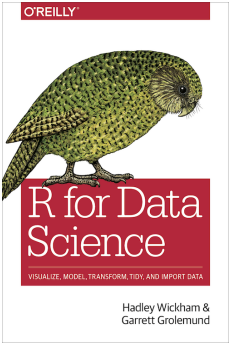
```
\left(x + a\right)^n =\sum_{k = 0}^{n}{\binom{n}{k}x^ka^{n-k}}
```

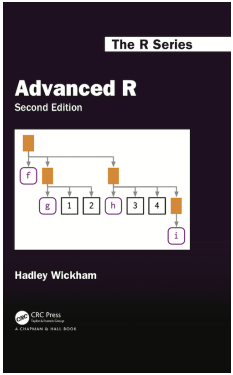
```
\begin{equation}
 \left(x+a\right)^n=\sum_{k=0}^{n}{\binom{n}{k}x^ka^{n-k}}
 (\#eq:binomial)
\end{equation}
```

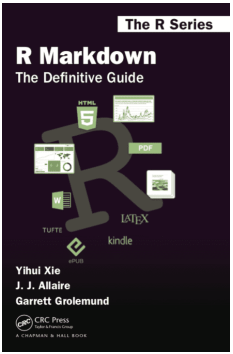
$$(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k} \tag{1}$$

# Wrap-up


Import
readr
Tidy
tidyr
Transform
dplyr
%>%
Ceci n'est pas un pipe.
broom
Visualise
ggplot2
Model
linear models
Communicate
rmarkdown




1. https://r4ds.had.co.nz
2. https://adv-r.hadley.nz
3. https://bookdown.org/yihui/rmarkdown

# Questions?

bhaswar.chakma@ucp.pt