JOM RESEARCH

INTRODUCTION TO R, POSIT CLOUD AND R PACKAGES

Tengku Muhammad Hanis Mokhtar

September 19, 2023

About me

Lead academic trainer at **Jom Research**

Background

- PhD student of Public Health Epidemiology, USM
- MSc (Medical Statistics) from USM, 2019
- MBBCh from Al-Azhar University, 2015

Interest:

- Medical statistics, meta-analysis, bibliometrics, scientometrics, text analysis
- Machine learning and deep learning application in medical sciences
- Application of R on health/medical data

Contact me:

- Email: tengkuhanismokhtar@gmail.com
- Website:
  - https://tengkuhanis.netlify.app/
  - https://jomresearch.netlify.app/

JOM RESEARCH

Outline

- R
- RStudio IDE
- Posit Software, PBC
- Timeline
- Why use R?
- Posit Cloud
- R packages

R

- A language and environment for statistical computing and graphics
- Continuation of S language
- Developed by Ross Ihaka and Robert Gentleman at University of Auckland in 1991

RStudio IDE

- Integrated development environment (IDE) for R
- There are other IDEs for R:
  - JupyterLab
  - PyCharm
  - VS Code
  - Etc.

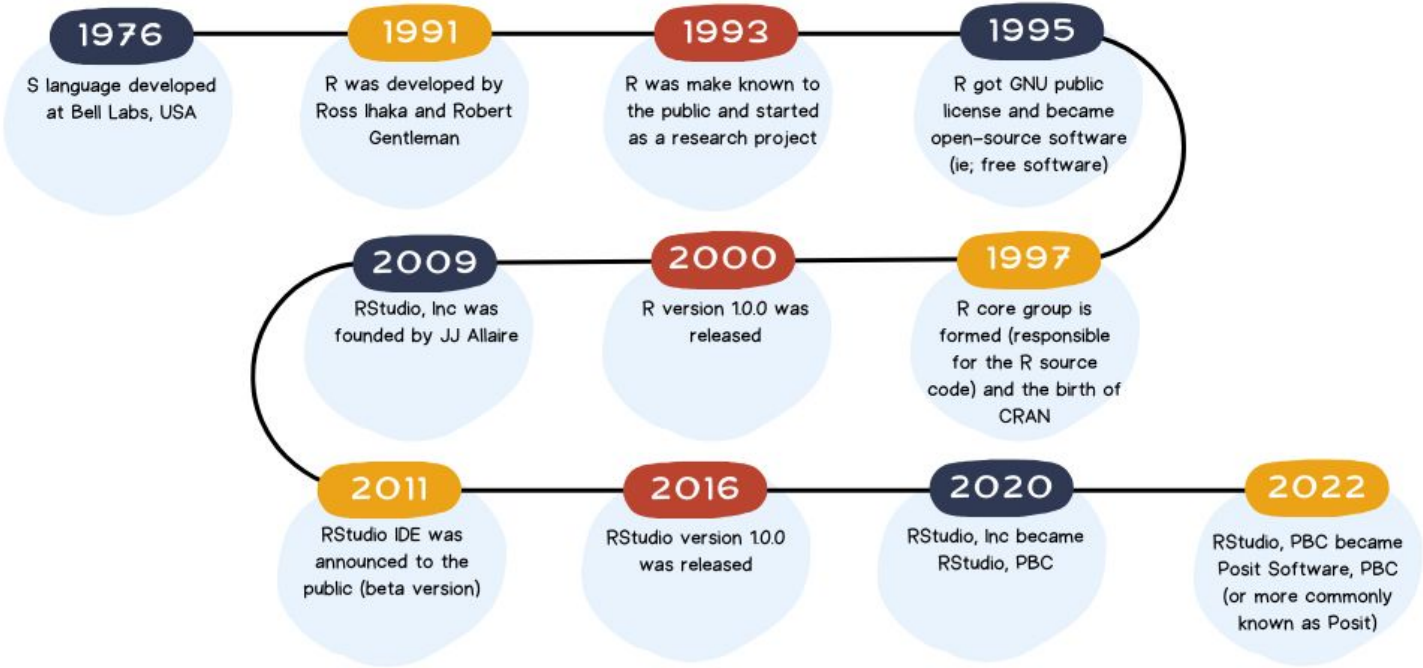
Posit Software, PBC

- More commonly known as Posit
- A company that developed RStudio IDE
- Previously, known as RStudio, Inc → RStudio, PBC → Posit Software, PBC
- PBC is a Public Benefit Corporation:
  - Company need to balance balance between the profit and public benefit
- Basically, by being PBC, Posit strengthen their commitment to open source software
- Why change to Posit? - because Posit support more than just R (ie; Python and Julia)

- Posit support the growth of R by:
  - Develop RStudio IDE
  - Develop and maintain a few vital packages:
    - Tidyverse, tidymodel, tidytext, rmarkdown, shiny, devtools, etc
  - Develop quarto
    - Multi-language version of R Markdown (R, Python, Julia, Observable)
    - Can be run in RStudio, VS Code, Jupyter Lab, Neovim, etc.
  - Develop Posit Cloud

TIMELINE INFOGRAPHIC

R, RSTUDIO, POSIT



Why use R?

- Extensively supported by Posit
- Wide functionality:
  - Statistical analysis
  - Machine learning, deep learning
  - Image processing (to certain extent)
  - Website development
  - Etc
- As a statistical software:
  - Cover wide range of analysis; basic to advanced
  - Early access to new statistical method
  - Free

R packages

- Fundamental units of reproducible R code
- An R package includes:
  - R functions
  - Documentation how to use R functions
  - Sample data
- How to install R packages:
  - CRAN*
  - Bioconductor
    i. For bioinformatics related packages
    ii. Need to install BiocManager (available in CRAN)

- GitHub repository*
  - Need to install remotes or devtools packages and Rtools (for RStudio desktop)
  - Example:
    - dmetar package
    - Dmetatools
- GitLab - similar to GitHub
- Bitbucket - similar to GitHub

**Cover in hands-on today*

Posit Cloud

- Formerly known as RStudio Cloud
- More or less has similar functionalities with RStudio IDE (for desktop/pc)
- Free with limited specs: 1GB RAM, 25 hours per month
- Sign-up/log in: https://posit.cloud/

Any question?

To do (10 minutes)

- Sign up to Posit Cloud
- Install R packages in Posit Cloud:
  - CRAN: remotes package
  - GitHub: dmetatools package

DATA VISUALISATION WITH GGPLOT2

JOM RESEARCH

Outline

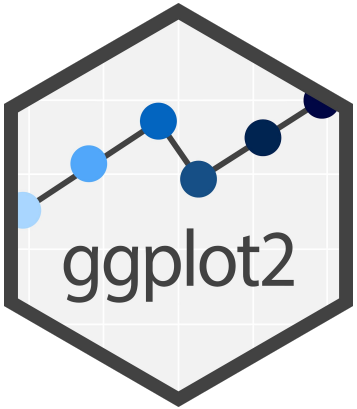
- R packages for visualisation
- Why use ggplot2?
- Good references
- Basic of ggplot2
- Hands-on in R

R packages for visualization

- ggplot2 - static and various plots
- lattice - static and various plots
- highcharter - interactive and various plots
- leaflet - interactive map
- plotly - interactive and various plots
- etc.

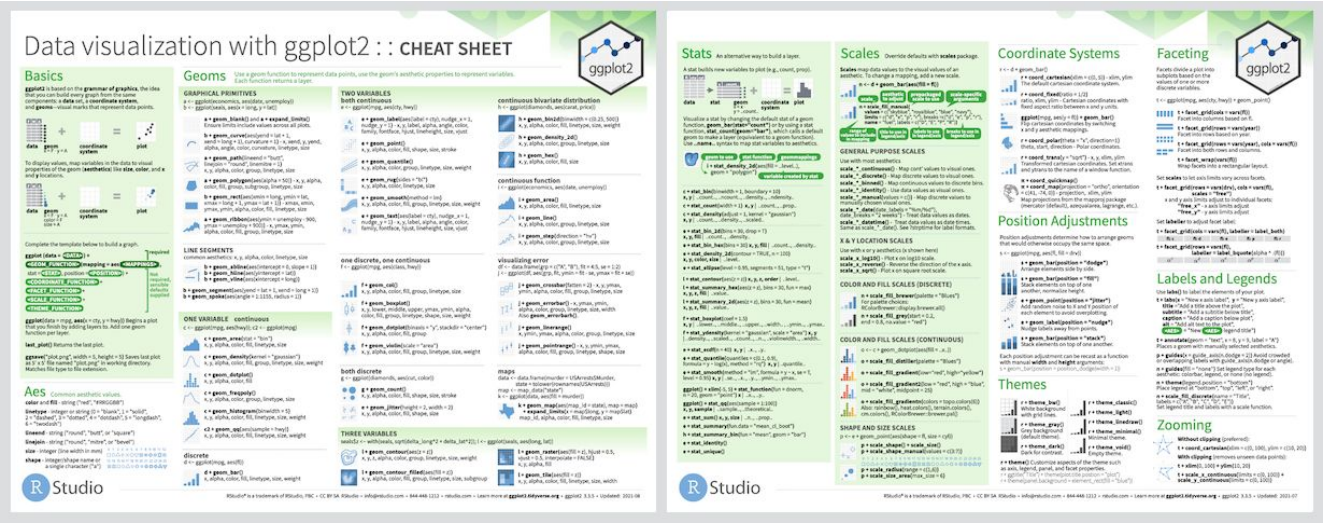
Why use ggplot2?

- Developed and maintained by Posit
- Part of tidyverse world
- Quite popular with R community
- Wide variety of plots - from simple bar charts to complex geospatial visualisations
- ggplot2 extension



Good references

- Posit Cheatsheets
- Just google it
- Stack Overflow



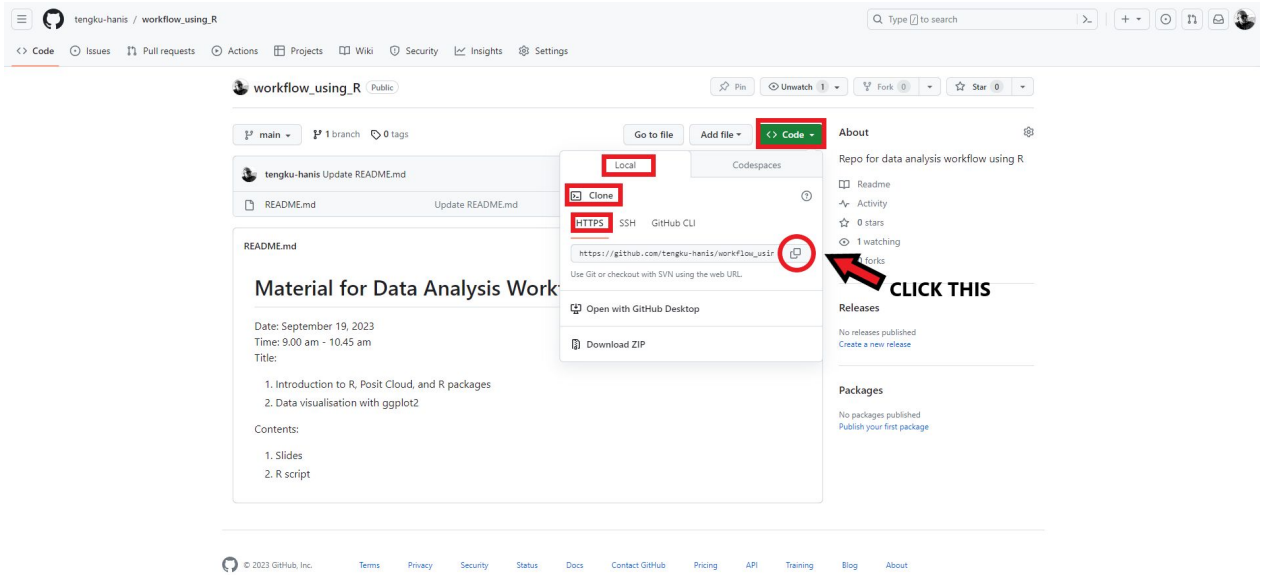
Basic of ggplot2

- Basically, one layer is stacked over another
- Begin with `ggplot()` for data holder
- Each layer can add using `+`
- Add geom layers (type of plot): `geom_point()`, `geom_bar()`, etc.
- Add other elements such as axis label, title, theme, etc.

```
ggplot() +
    geom_point() +
    labs() +
    theme()
```

Hands-on in R

1. Google search tengku hanis github
2. Go to the first website (my GitHub page)
3. Click <> Code → local → Clone → HTTPS → copy the link



4. Log in to Posit Cloud
5. Click New Project → New Project from Git Repository → paste the url
6. Install R packages:
    a. tidyverse
    b. mlbench
    c. gganimate
7. Open:
    a. basic_ggplot.R
    b. exercise.R

Any question?

tengkuhanismokhtar@gmail.com