# INTRODUCTION TO R (1995- Ross and Robert)

## Brief overview

* Programming language for statistical analysis
* Designed for data analysis, statistical and mathematical modelling.
* Flexible (Open source)
* Core/base packages
  + Preselected and installed by default when R studio is installed
* Other packages
  + Developed by users and installed by the user either through:
    - CRAN
    - GITHUB
* Integrated with other tools
  + Tinymodels (For modeling and ML
  + Tinydiverse (for data science)

## Things to note

* R is case sensitive
* It’s an OOP
* Objects can have names (can not start with numbers and special characters)
* # Used to make comments
* Data types in R

## R Datatypes

* Integers
  + Whole numbers
* Double
  + Decimals (64 bits)
  + Different with Floats (32 bits)
* Character /String
  + ‘’Anne Kariithi’’
* Factor
  + Assign roles in a categorical dataset
* Logic
  + AND (&)
    - Gives a True if both are the same
  + OR (|)
    - Gives a False if both are the same
* Dates
* Boolean
  + Trues / False

|  |  |  |
| --- | --- | --- |
|  | T | F |
| T | T | F |
| F | F | T |

## Shortcuts

* CTRL + ENTER (Run)
* Table(ideal3$CalfSex) -> to see the frequency in the dataset
* Distinct () -> to only select the unique data from the dataset

## Install packages

* Click on packages and write the package name
* Using code:
  + Install.packages(“tidyverse”)

## Setting the working directory

* setwd – this sets the working directory

## Importing libraries

* Using code
  + require (tidyverse)
  + library (tidyverse)

## Data Wrangling

* Means to clean data and set it in the right format

## Piping (|>, %>%)

* This is used to run more lines of code concurrently

## Stacking

* You can use a Column to specify the stacking method

## Combining two datasets

* Right Join
  + Picks all data from the right table and all data that matches in both tables
* Left Join
  + Picks all data from the left table and all data that matches in both tables
* Inner Join
  + Picks only data that matches both tables
* Full Join / Outer Join
  + Joins the datasets regardless of whether they appear in both tables