

Introduction to statistics using R

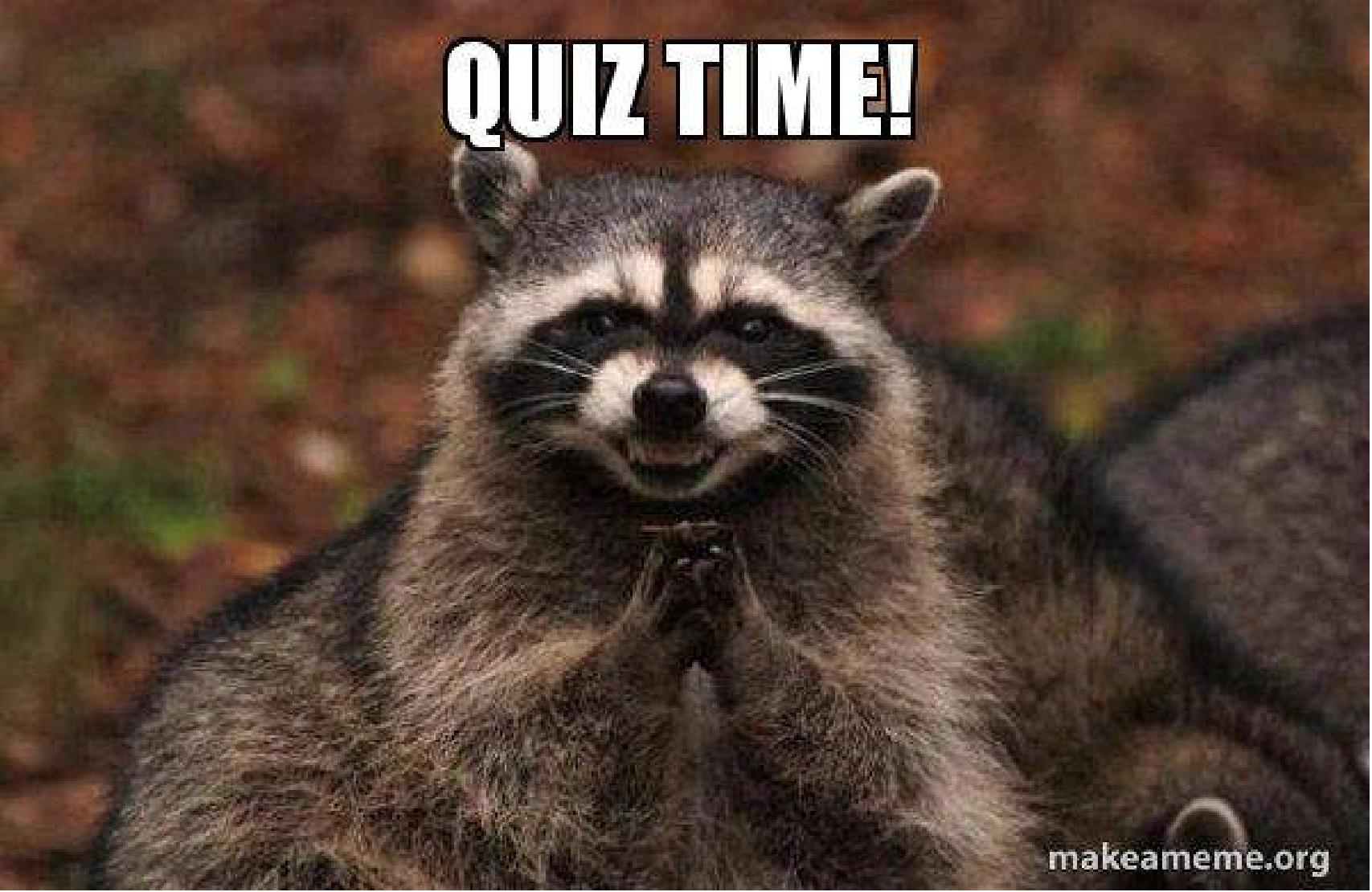
Seminar series - session 2



Session 2 – Learning objectives

- Understand the concept of probability distribution
- A very short intro to R and Rstudio

But first,



QUIZ TIME!

Distribution

The concept of distribution is the soul of data analysis

If you understand distribution, then you'll understand stats

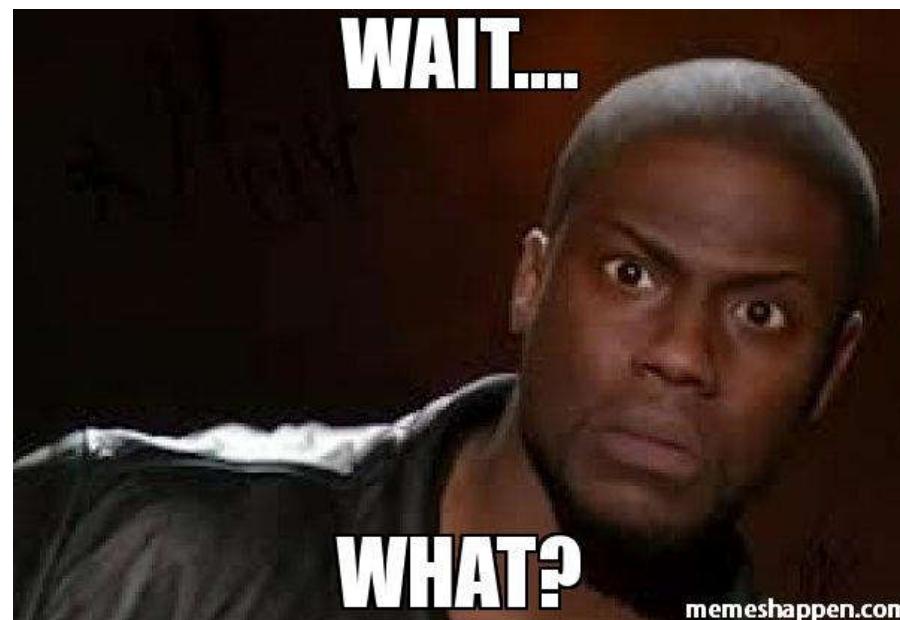
Statistical tests assume data follow specific distributions

What is distribution?

- **Probability** distribution
- Mathematical **function** defining all the possible values of a **random variable**, and how often they occur
- Within a given range
- Describes the long-run behavior of random variables
- How the possible values are plotted depends on:
 - Central tendency
 - Spread
 - Skewness
 - Kurtosis (among others)

Do not confuse

Frequency vs probability distribution



memeshappen.com

Frequency and probability distribution

Frequency:

- Empirical
- Summarizes observed data

Probability:

- Theoretical
- Calculates the observation probabilities



An example

- We captured 5 sandgrouses, 7 ravens, and 3 falcons
- Frequency distribution of species is 5, 7 and 3
- Probability distribution of species is $5/15$, $7/15$ and $3/15$

Why use probability?

- Data come from experiment or complex systems
- Variability and randomness
- Measure uncertainty in the data
- Probability measures uncertainty

The chances of winning
the lottery is 50%. You
either win or don't





Some definitions

- **random variable:** variable which outcome varies from measurement to measurement
- **Probability distribution:** possible values a random variable can take, and how likely they are
- Behavior of random variables must be defined using **probability**.



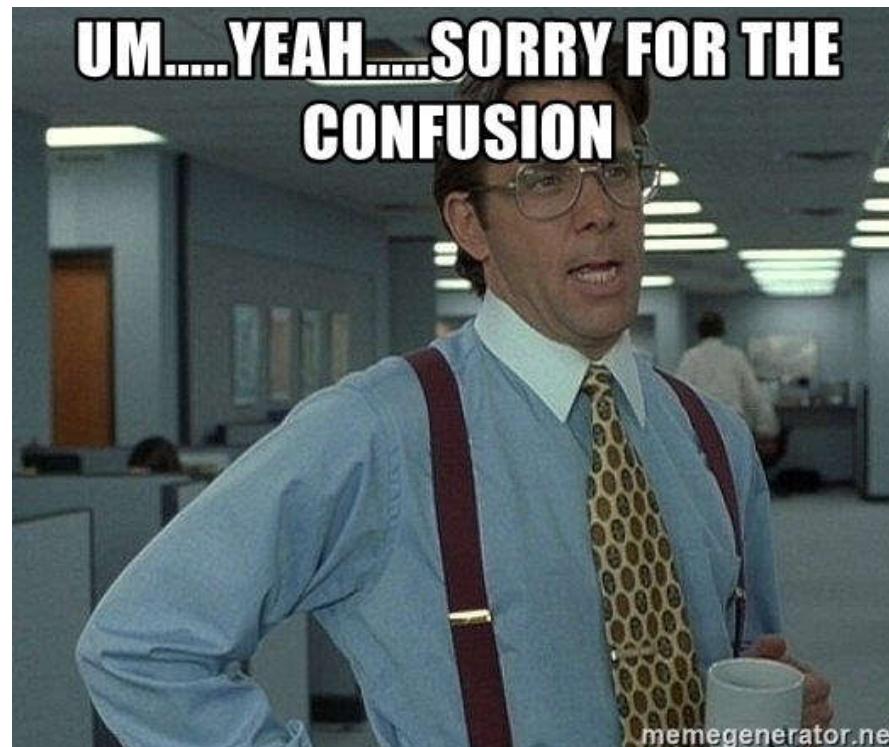
Do not confuse random variable and variable

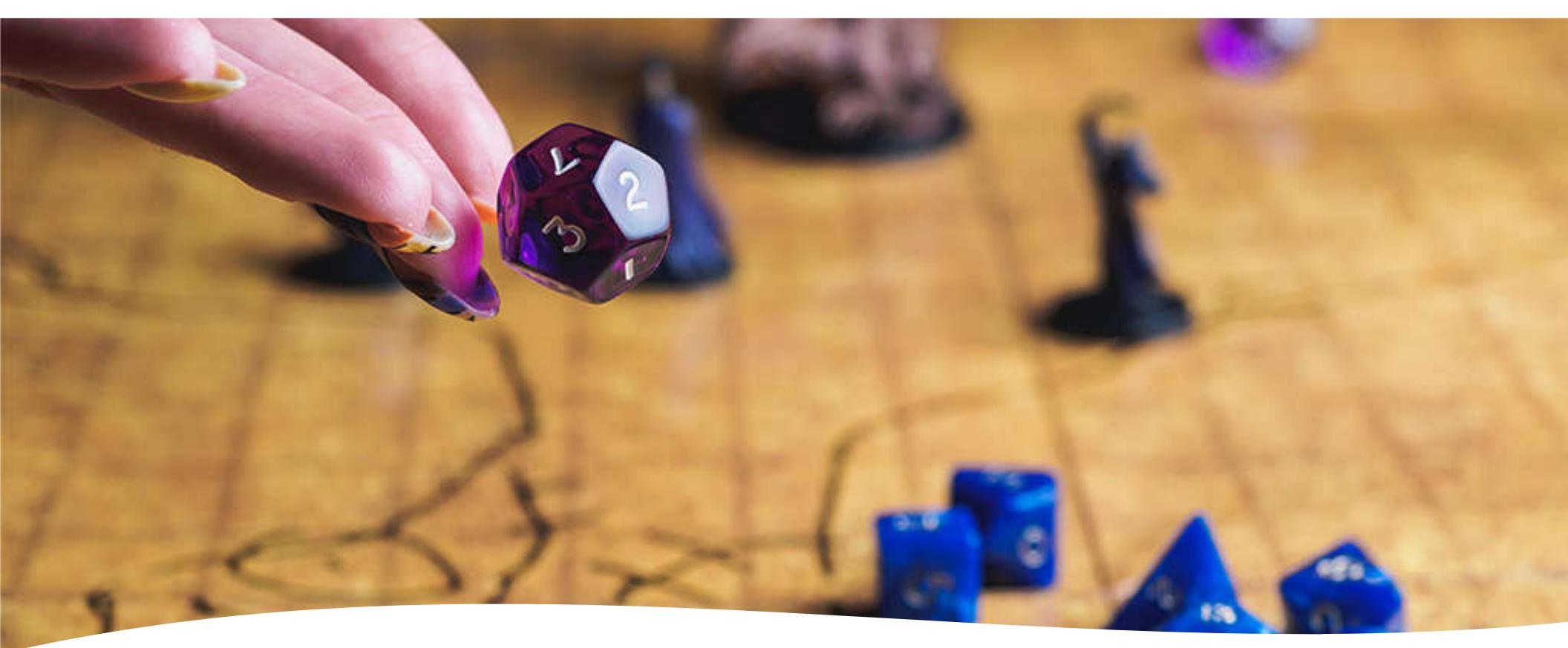
A **random variable** is always numerical and depends on the outcome of a chance experiment.

A **variable** is any property that you can measure or control

Unfortunate choice of words

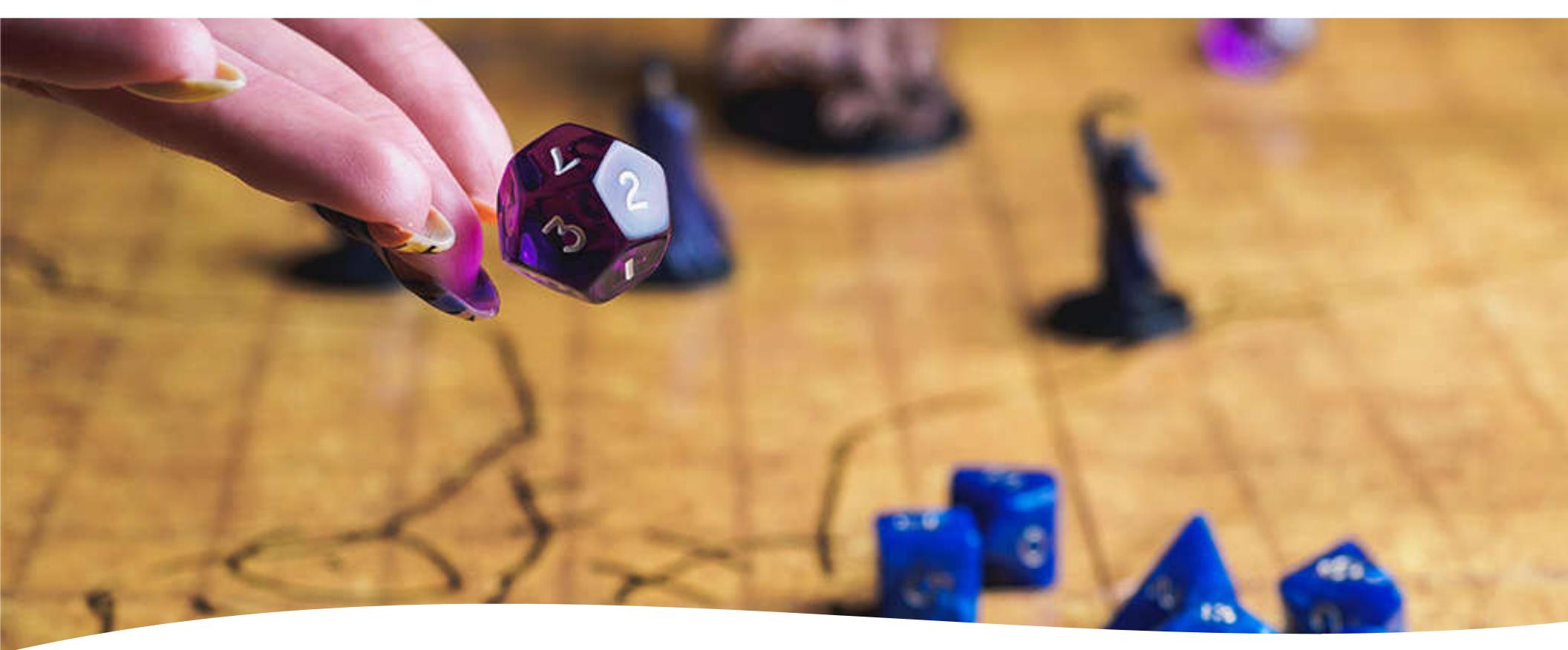
- Random variables: **not random & not variables.** They are functions mapping from possible outcomes of sample space to measurable space
- Categorical variables can also describe random outcomes





Some more definitions

- **Chance experiment**
Uncertain situation with 2 or more possible outcomes



- **Chance experiment**
Uncertain situation with 2 or more possible outcomes
- Your dragonborn barbarian wants to open a closed door in the dungeon using her Sword of Vengeance. When you roll your d20 to perform the deed, you'll have 20 possible outcomes from nat 1 to nat 20.*



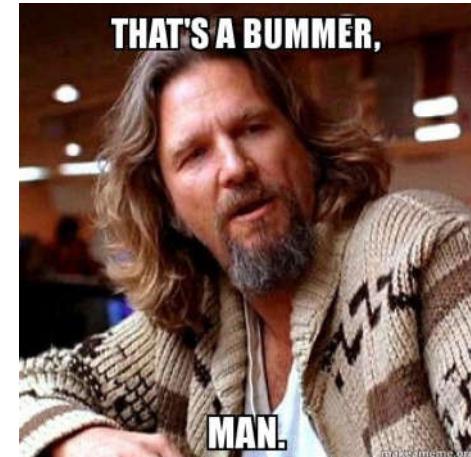


Some more definitions

- **Outcome**

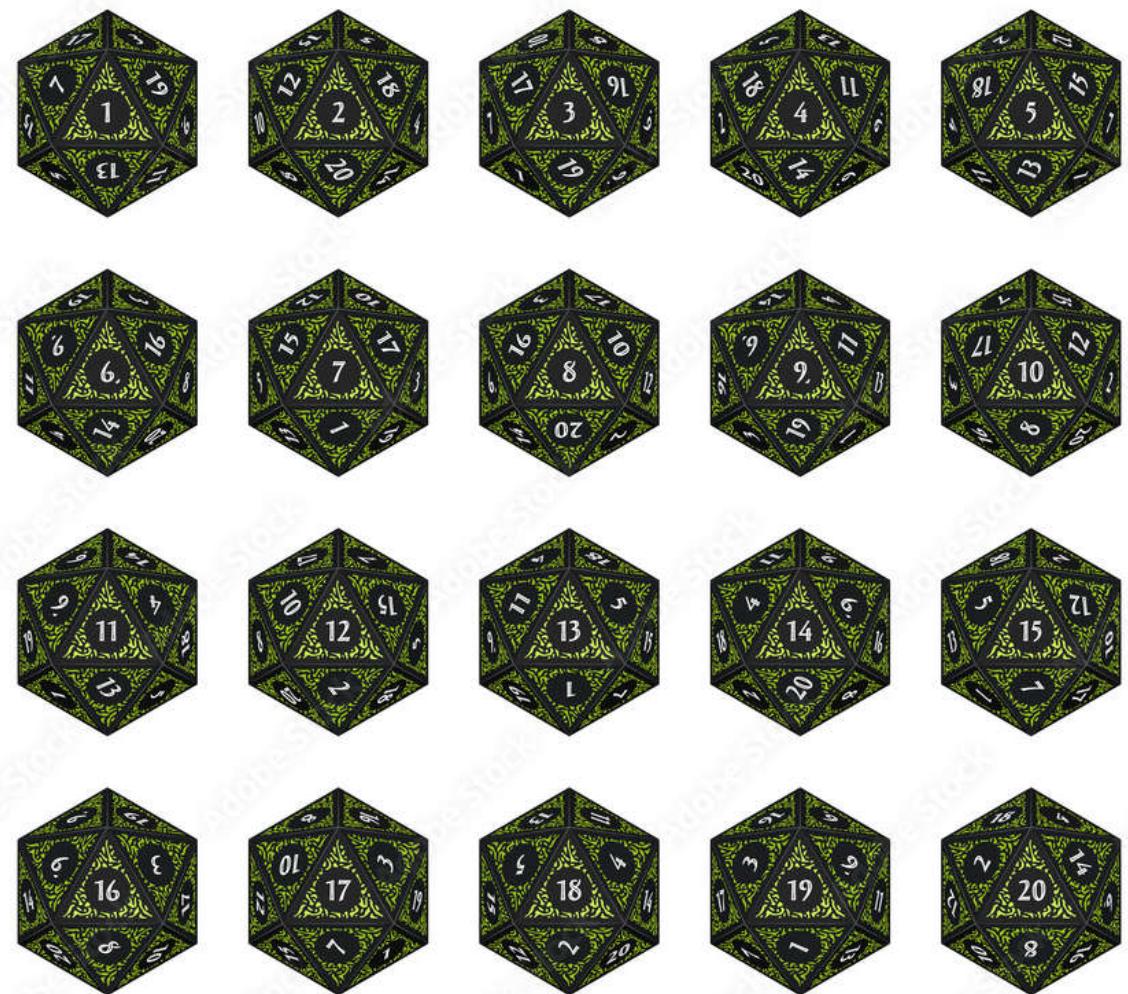
Result of a single trial

You rolled your d20 and got a nat 1.



Some more definitions

- **Sample space**
Collection of all possible outcomes from a chance experiment



The sample space when rolling your d20 is all integers from 1 to 20.



OR



Some more definitions

- **Event**

Set of outcomes from a chance experiment

Event “nat. 20 or 1”: all dice rolls that are either 20 or 1.

Some more definitions

- **Probability**

Measure of likelihood of an event occurrence

The Barbarian dragonborn rolls nat 1 three times in a row, almost killing off the party.

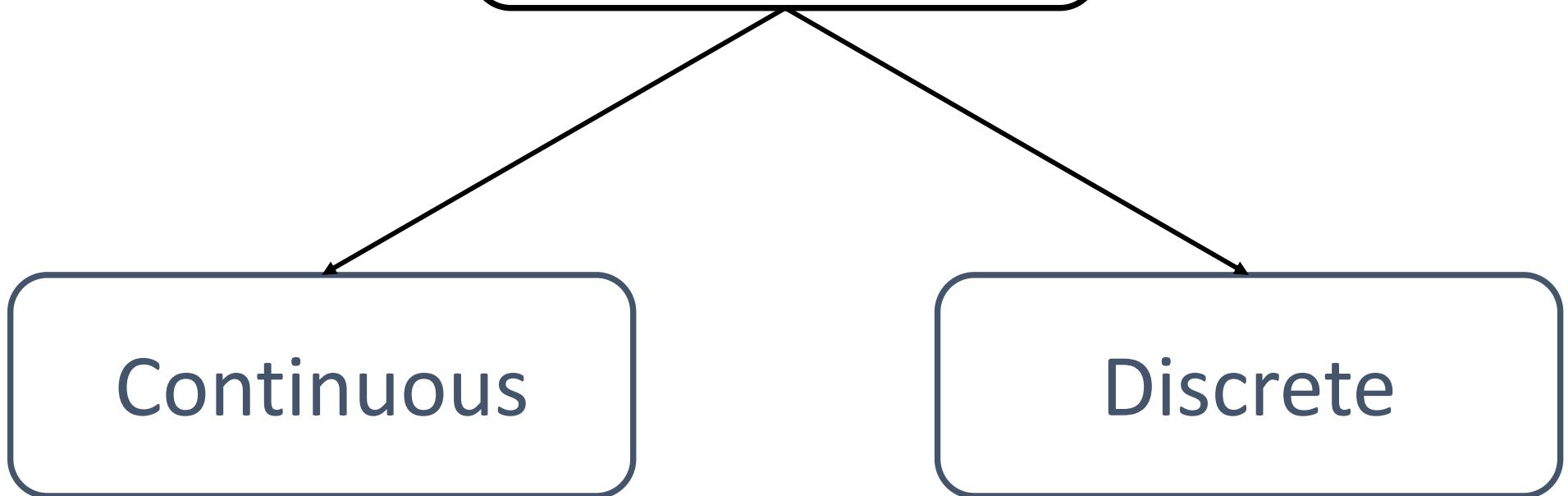
My tabaxi bloodhunter:

What was the probability?

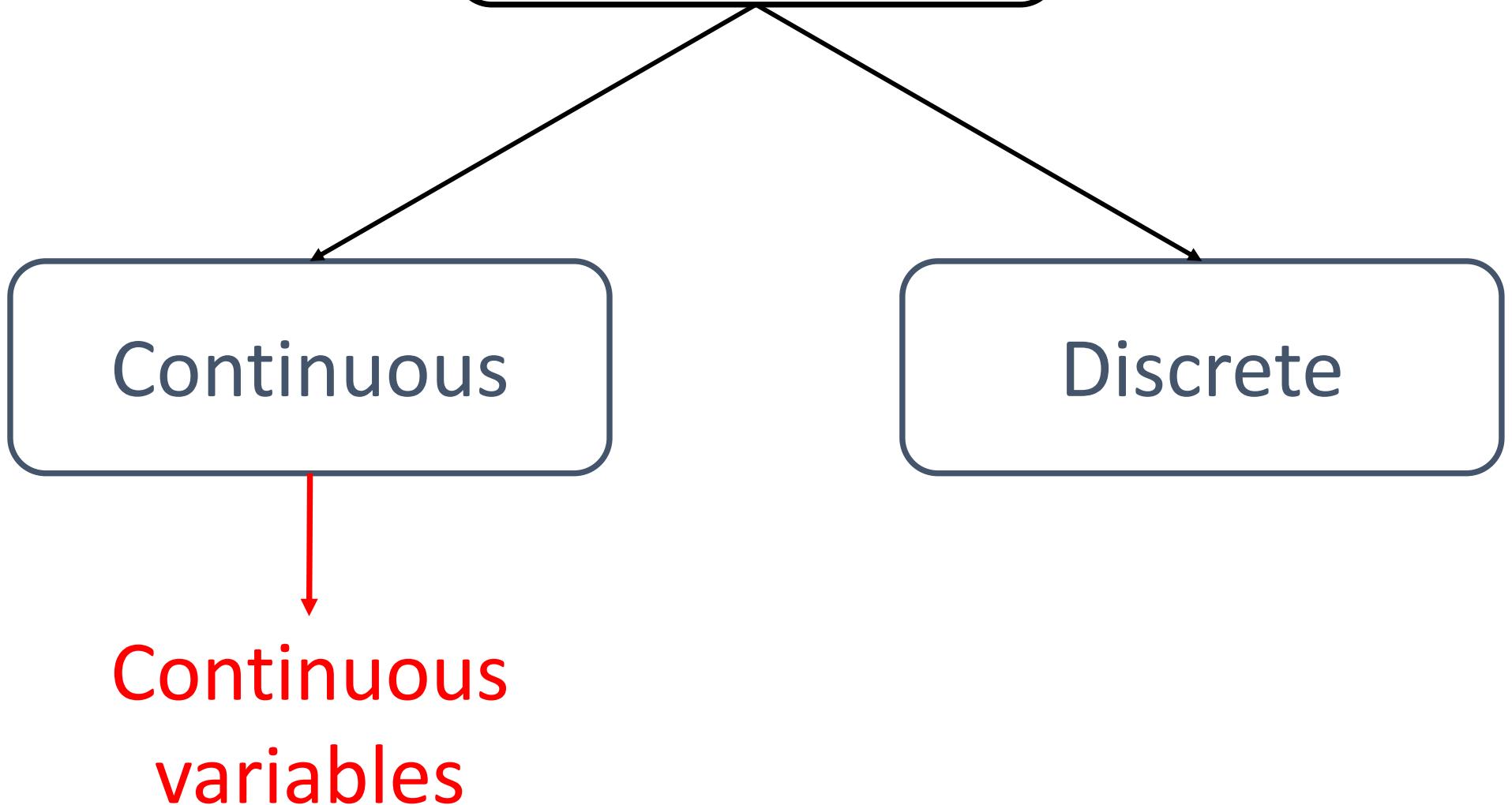
$$1/20 * 1/20 * 1/20 = 1/8000 = 0.000125$$



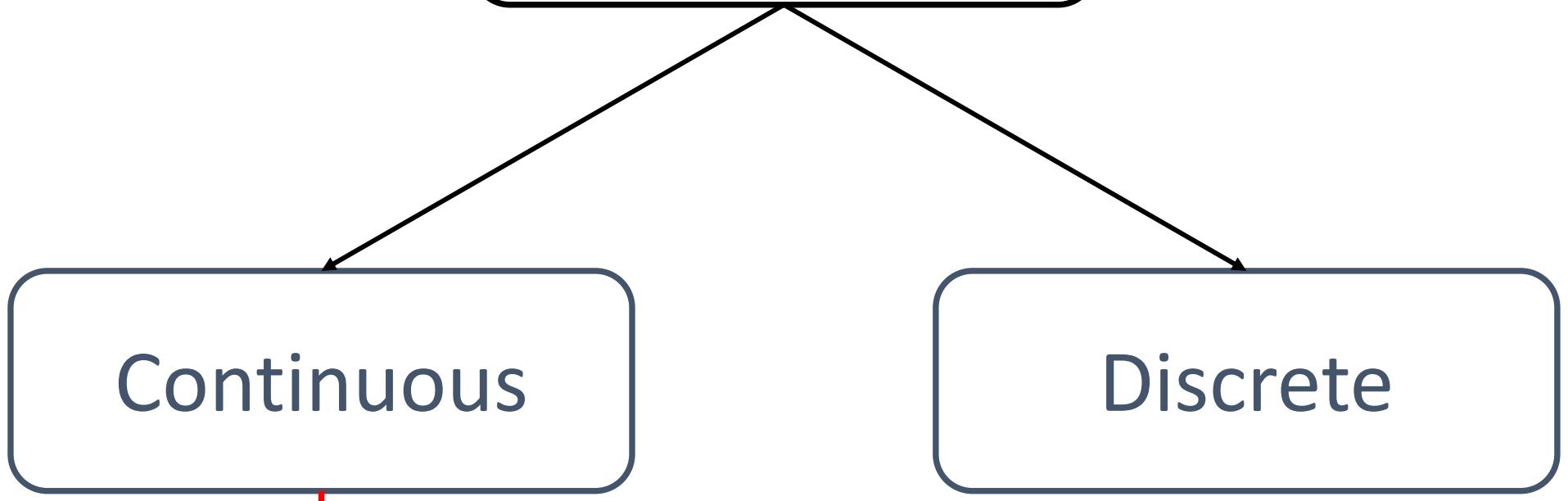
Probability Distribution



Probability Distribution



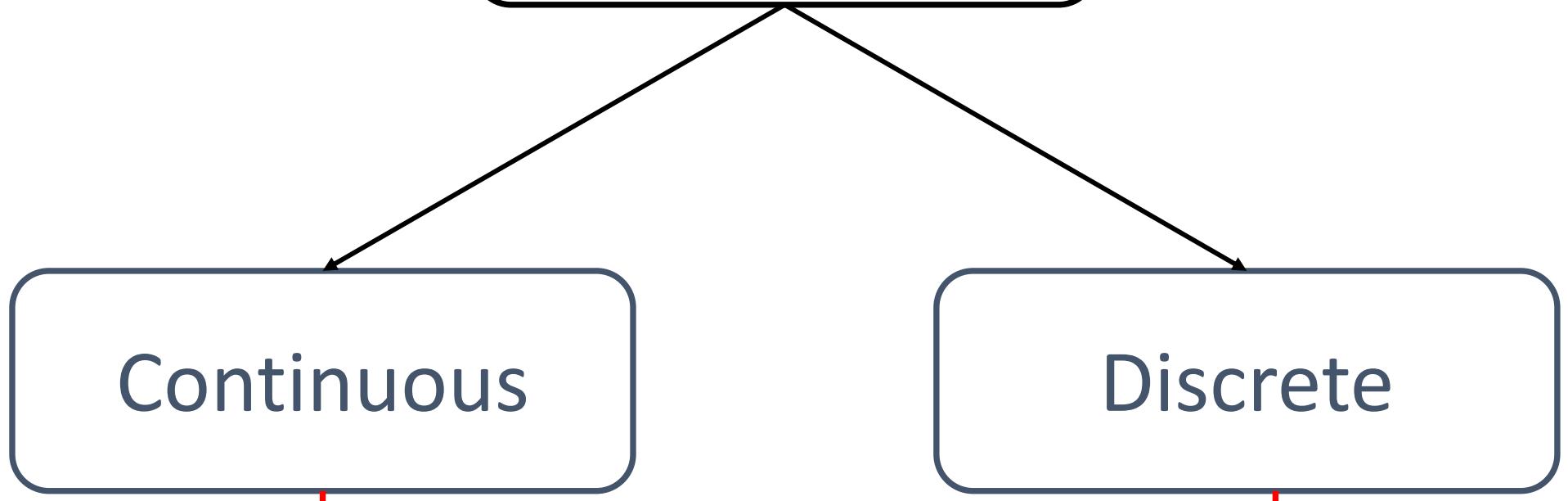
Probability Distribution



**Continuous
variables**

Continuum of value within range

Probability Distribution



Continuous

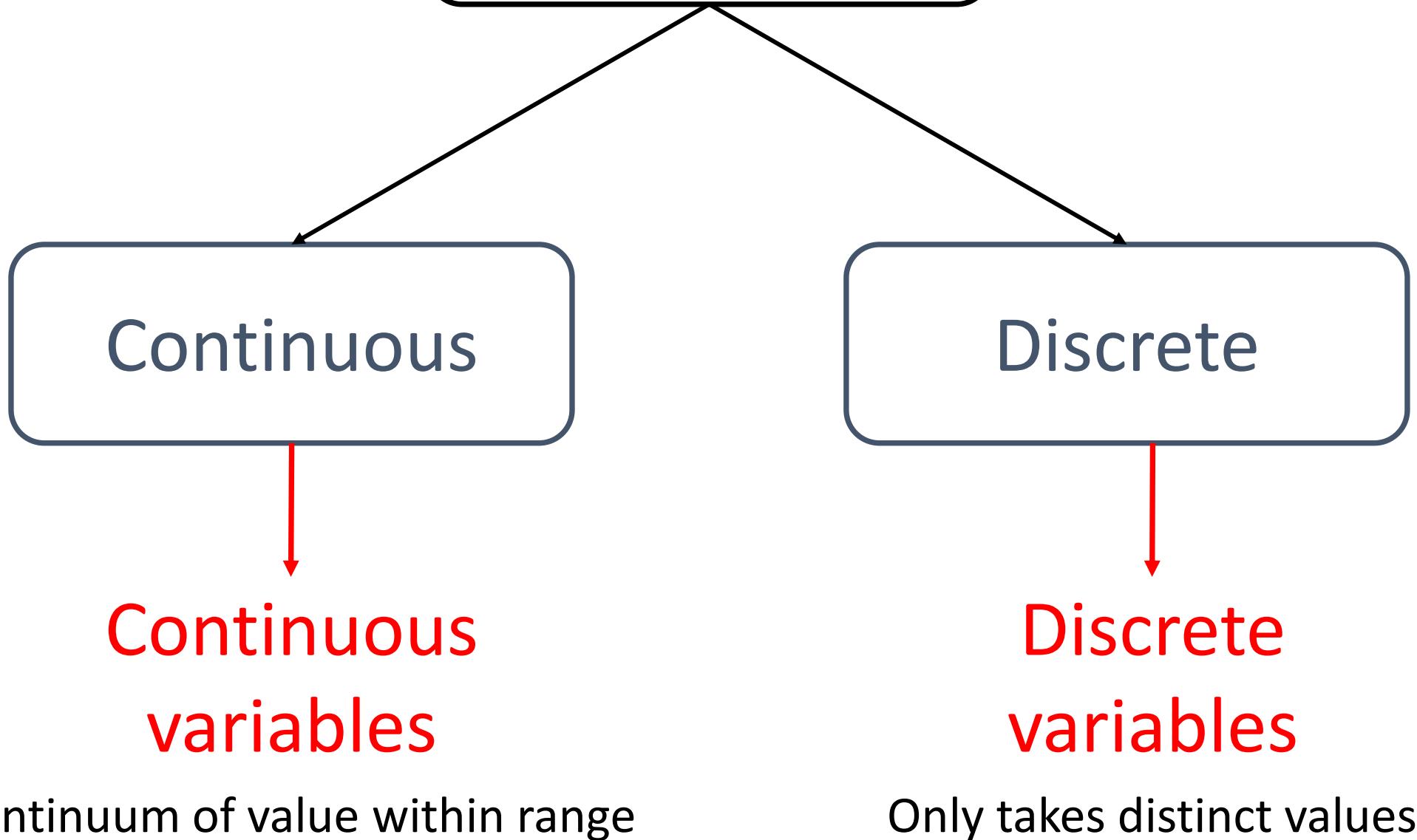
Discrete

Continuous
variables

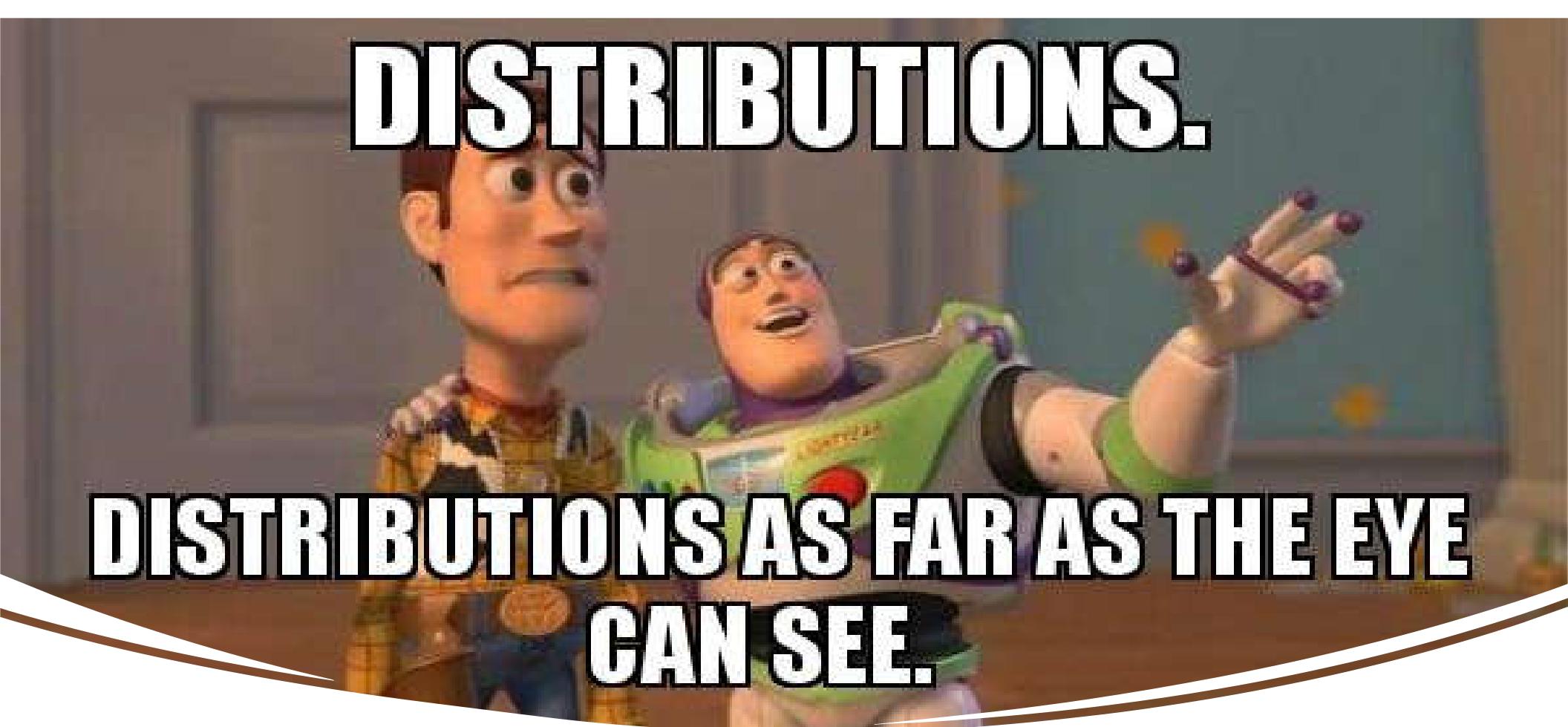
Discrete
variables

Continuum of value within range

Probability Distribution



DISTRIBUTIONS.



DISTRIBUTIONS AS FAR AS THE EYE
CAN SEE.

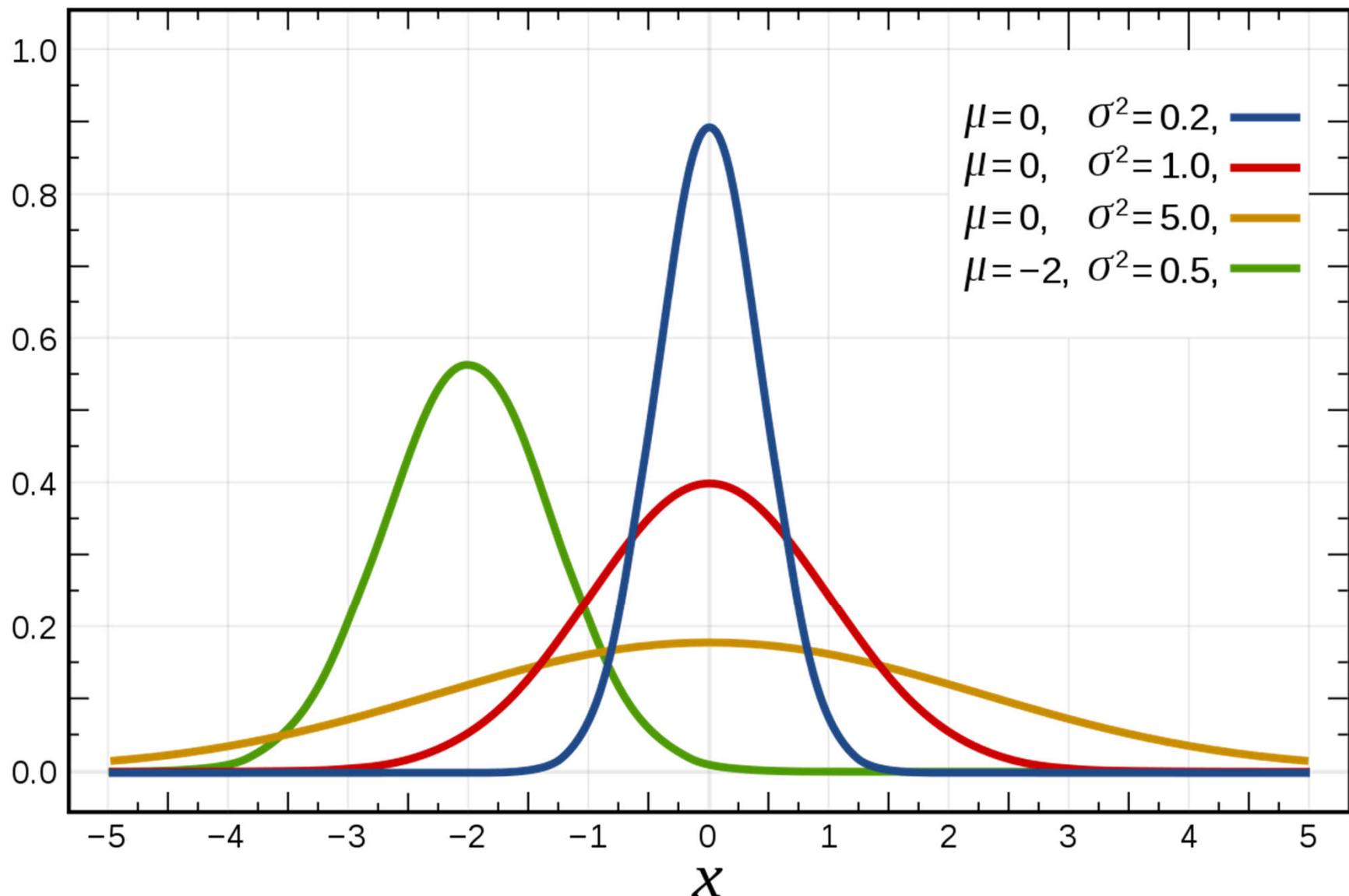
Most
common
distributions

Continuous variable:

- Normal (a.k.a. Gaussian)
- Uniform (a.k.a. rectangular)
- **Beta**
- gamma

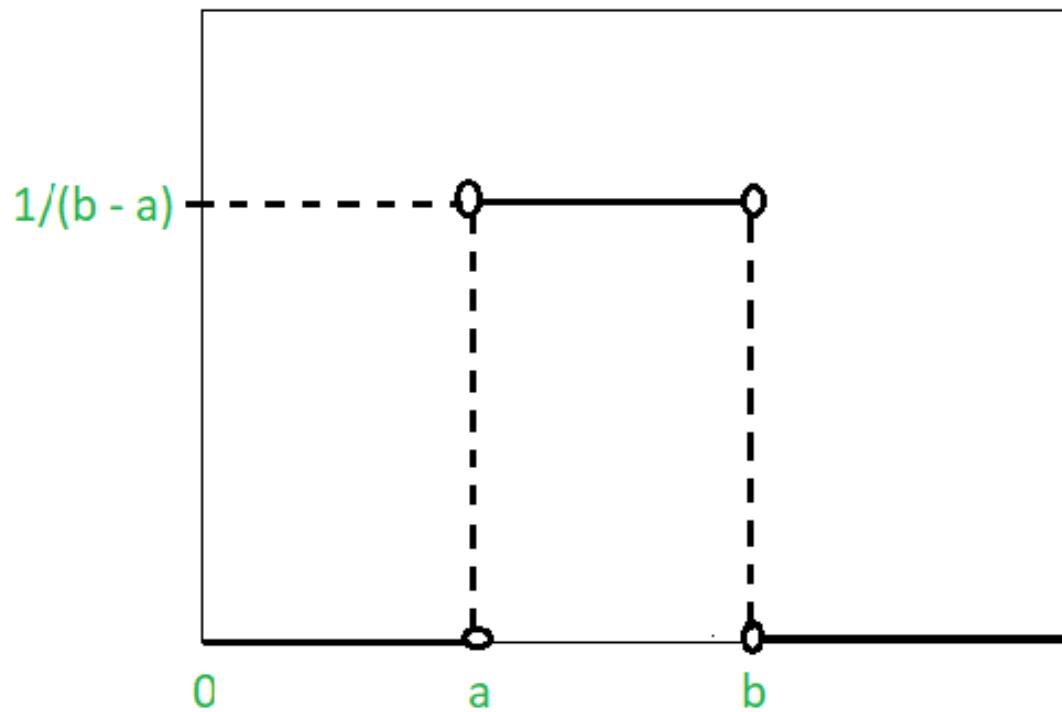
Gaussian or Normal distribution

- Symmetrical
- Defined by μ and σ



Uniform or rectangular distribution

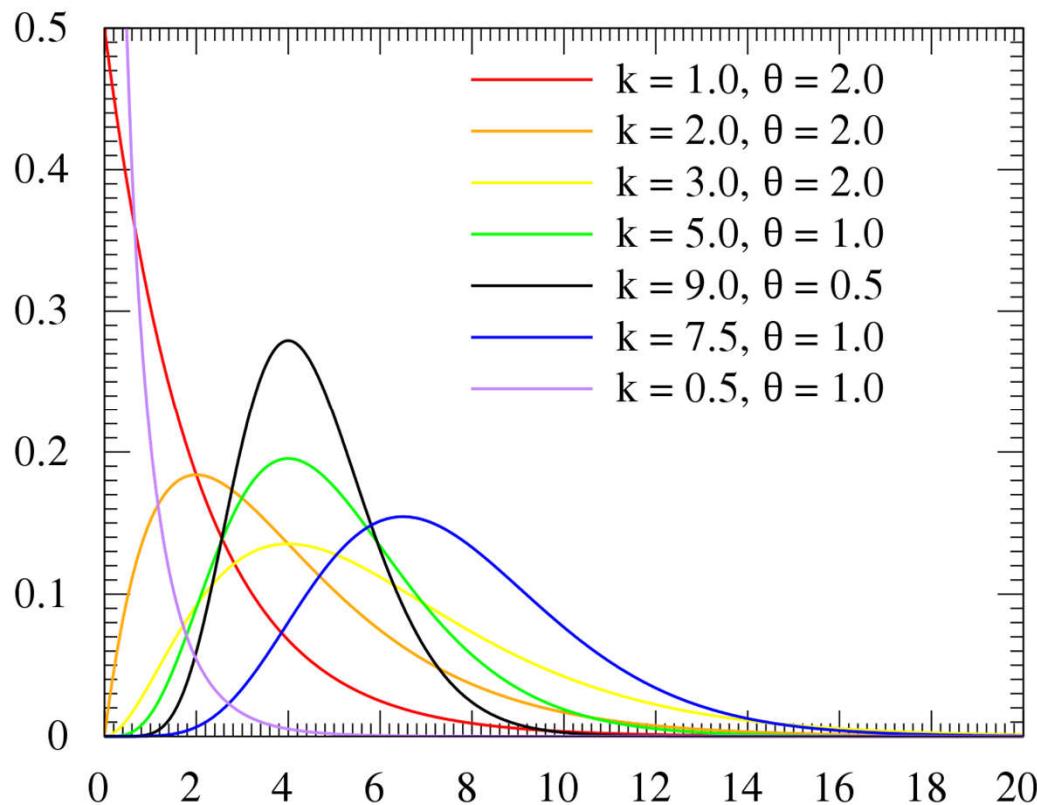
Defined by a (minimum value) and b (maximum value)



Constant probability over interval $[a-b]$

Gamma distribution

- Positive and right-skewed
- Defined by parameters k (shape) and θ (scale)



- Continuous, positive, right-skewed data with constant variance on the log-scale
- Survival data, rainfall,...
- χ^2 distribution: special case of gamma
- Used in χ^2 goodness-of-fit tests

Beta distribution

- Defined on $[0,1]$
- Defined by shape parameters α and β , both >0



Beta distribution

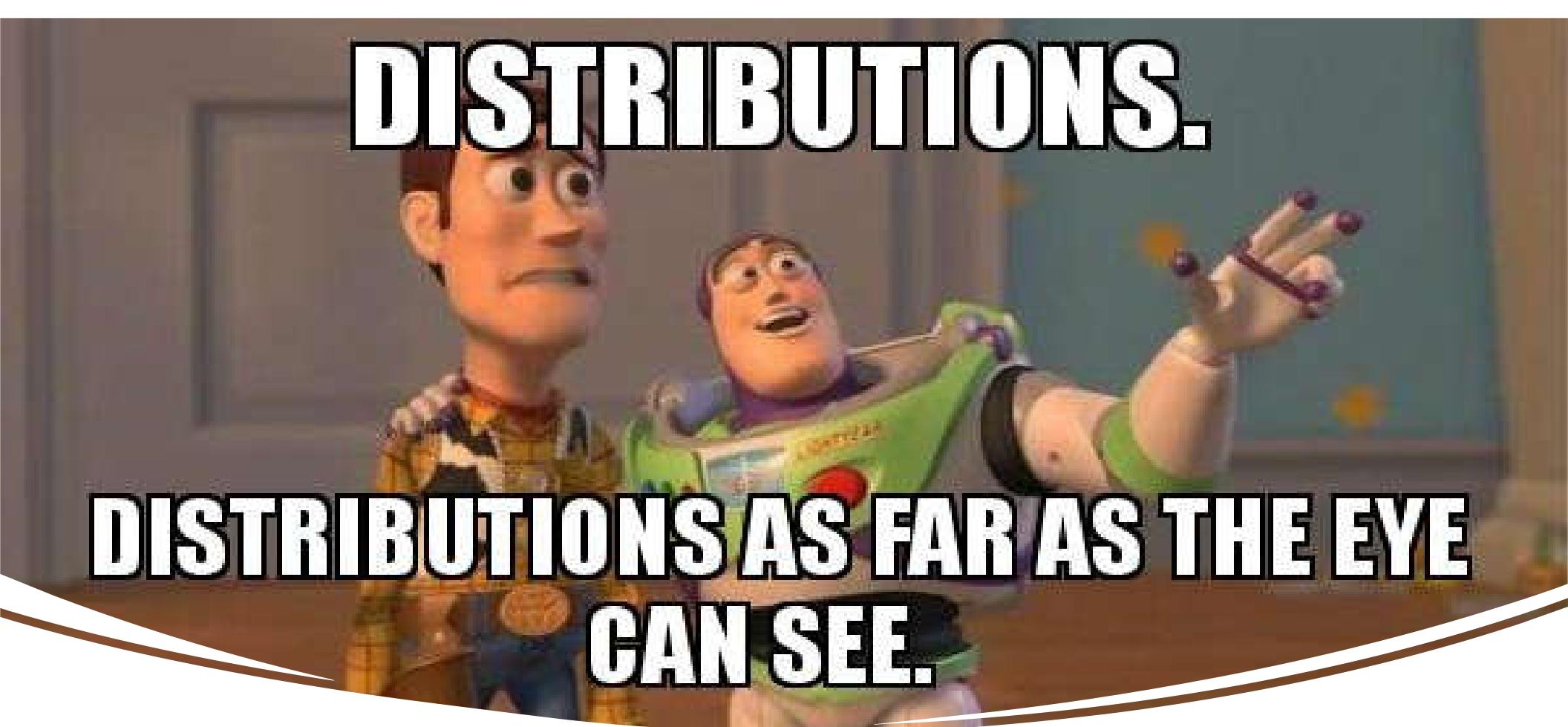
- Derived variables, notably proportions
- Doesn't deal with exact 0s and 1s (*loglikelihood function contains $\ln(x)$ and $\ln(1-x)$, and thus unbounded since $\ln(1) = 0$ and $\ln(0) = -\infty$*)



When 0 and 1 occur:

- Possible transformation: $(y * (n-1) + 0.5) / n$, $n = \text{sample size}$
- Zero-one inflated beta models

DISTRIBUTIONS.



DISTRIBUTIONS AS FAR AS THE EYE
CAN SEE.

Most
common
distributions

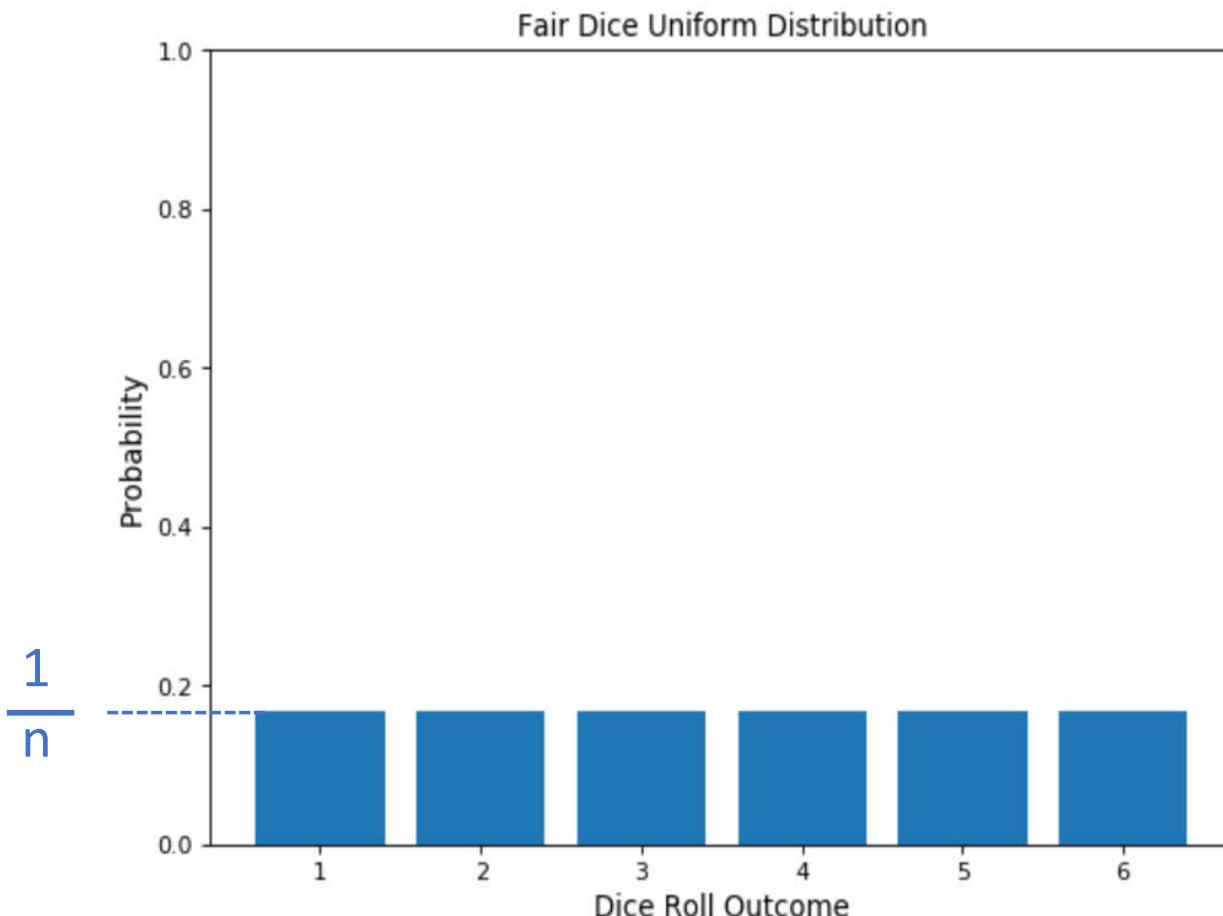
Discrete variables:

- Uniform (Yes, again!)
- Poisson
- Negative Binomial
- Binomial

Discrete uniform distribution

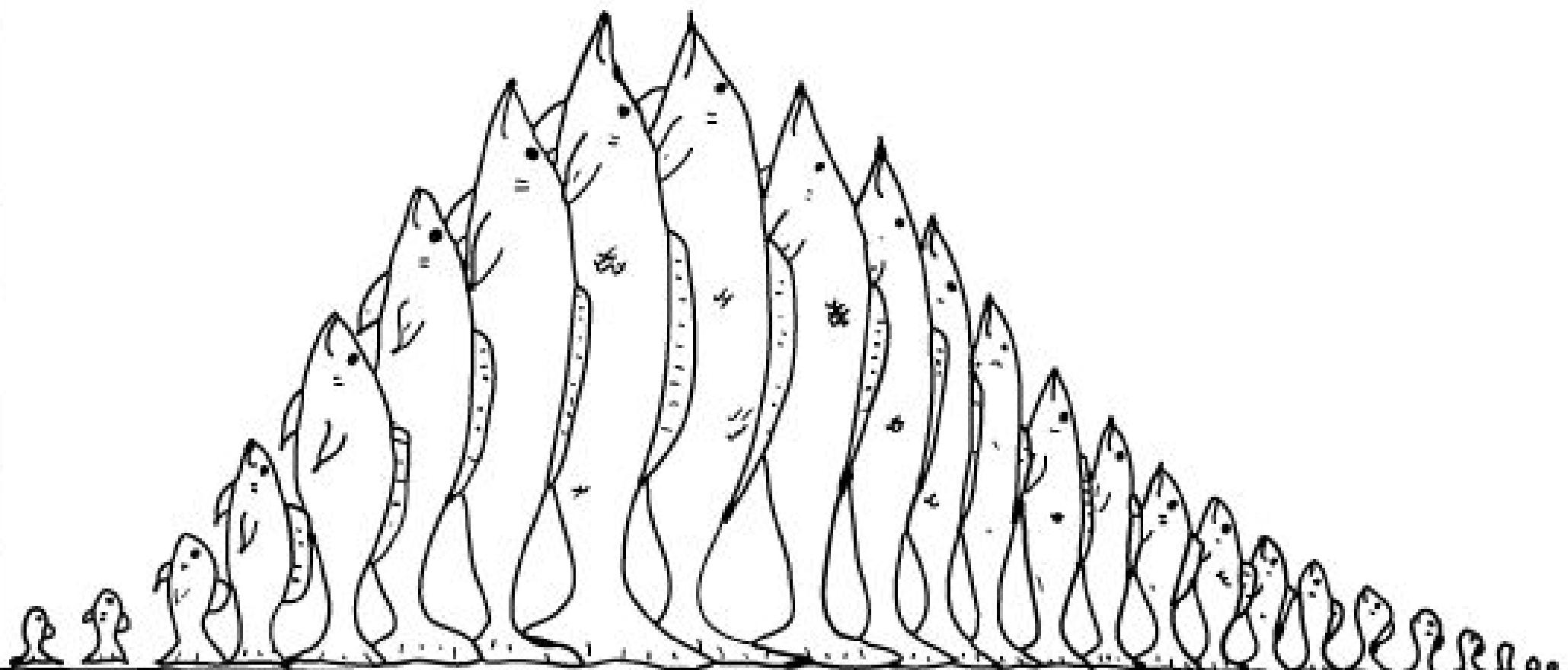
All outcome equally likely

Defined by a (min) and b (max) like the continuous



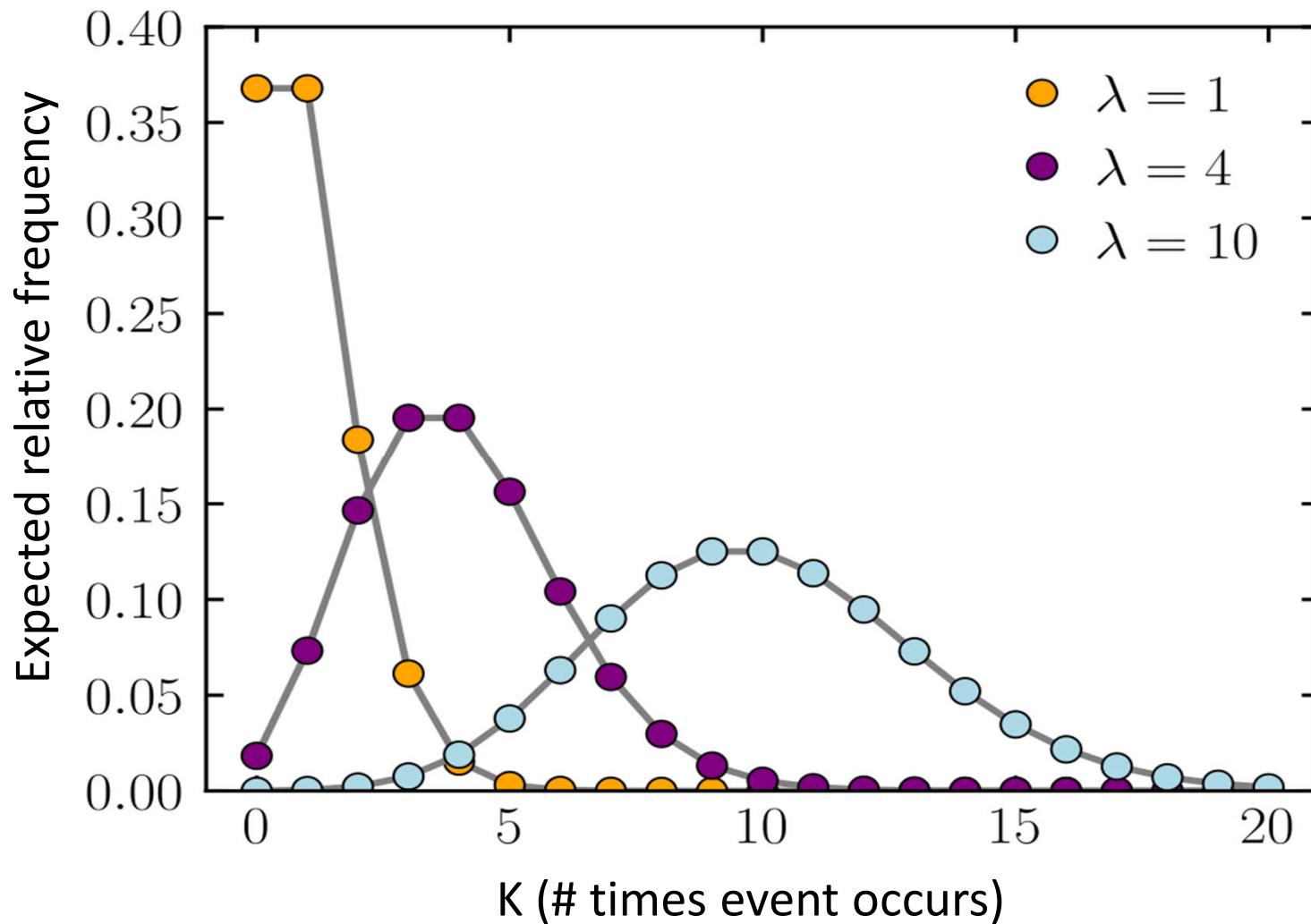
The Poisson distribution

- Count distribution
- Defined by 1 parameter λ mean number of events
- $\lambda = \mu = \sigma^2$ (lambda = mean = var)



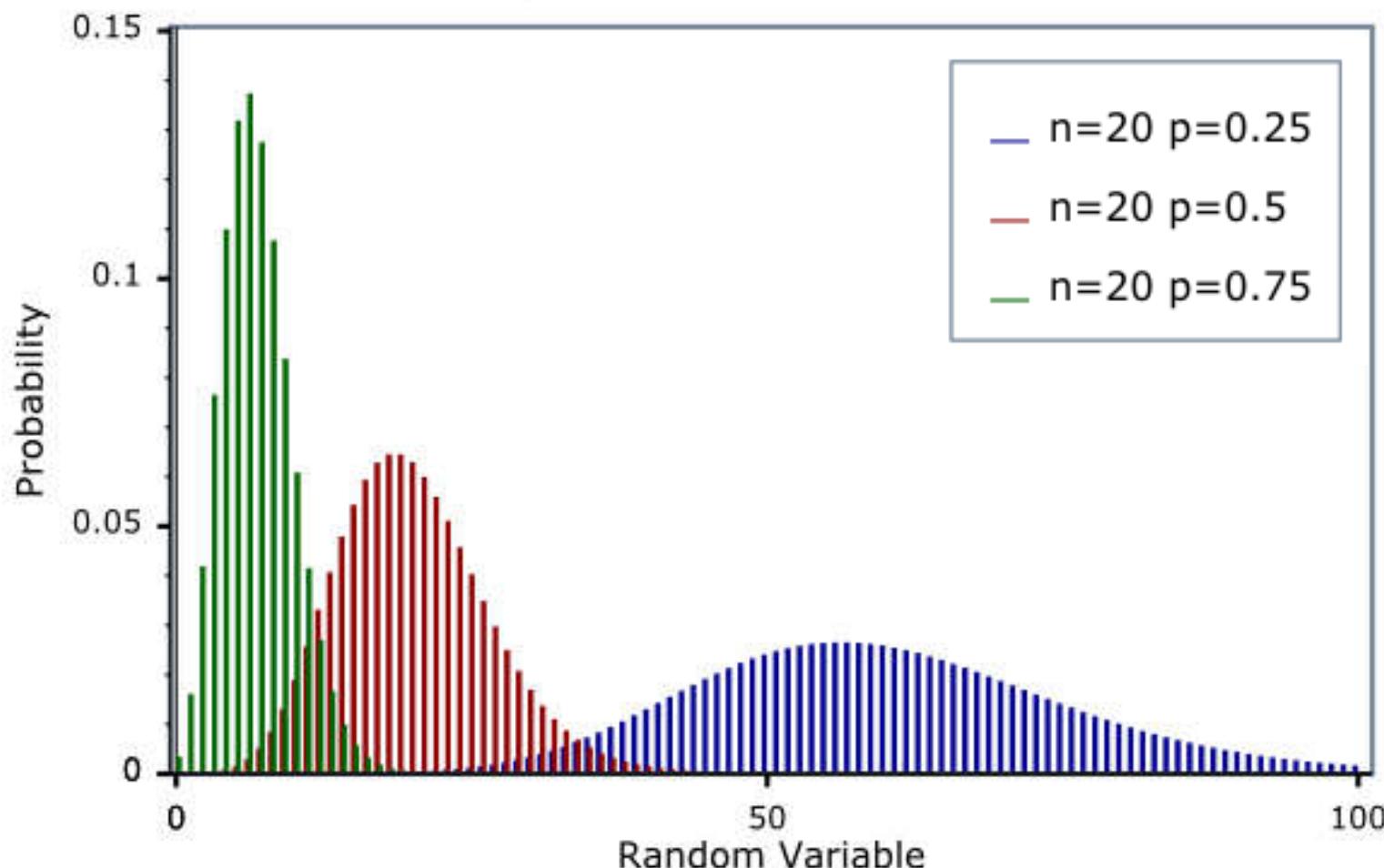
The Poisson distribution

- As mean increases, Poisson distributions approximate normal distribution
- Count data with large mean can often be modelled as continuous



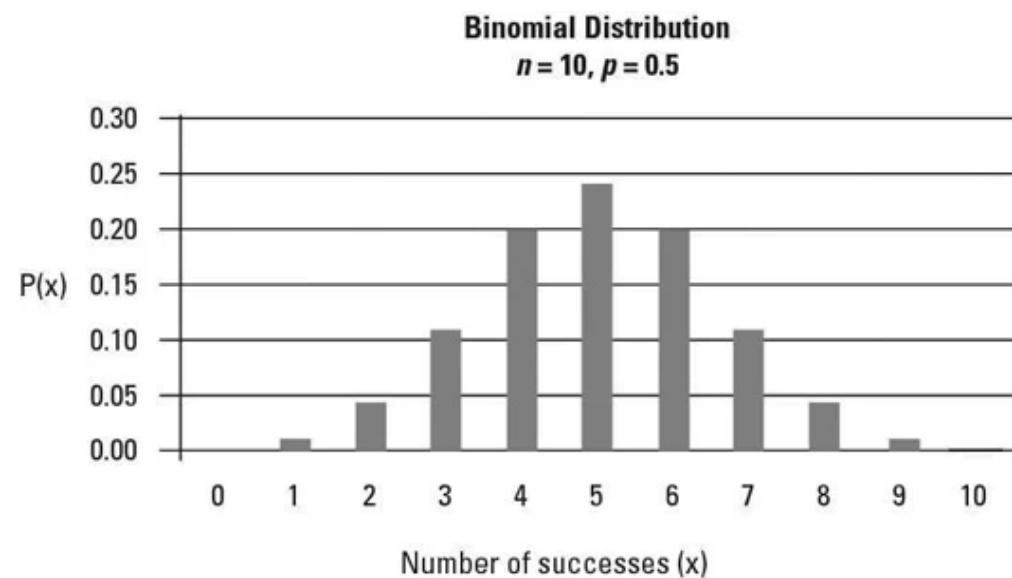
Negative binomial distribution

- In most cases $\mu = \sigma^2$ will not hold
- Overdispersion: $\mu < \sigma^2$
- 1 more parameter θ
- $\lambda = \mu$ and $\sigma^2 = \mu + \frac{\mu^2}{\theta}$



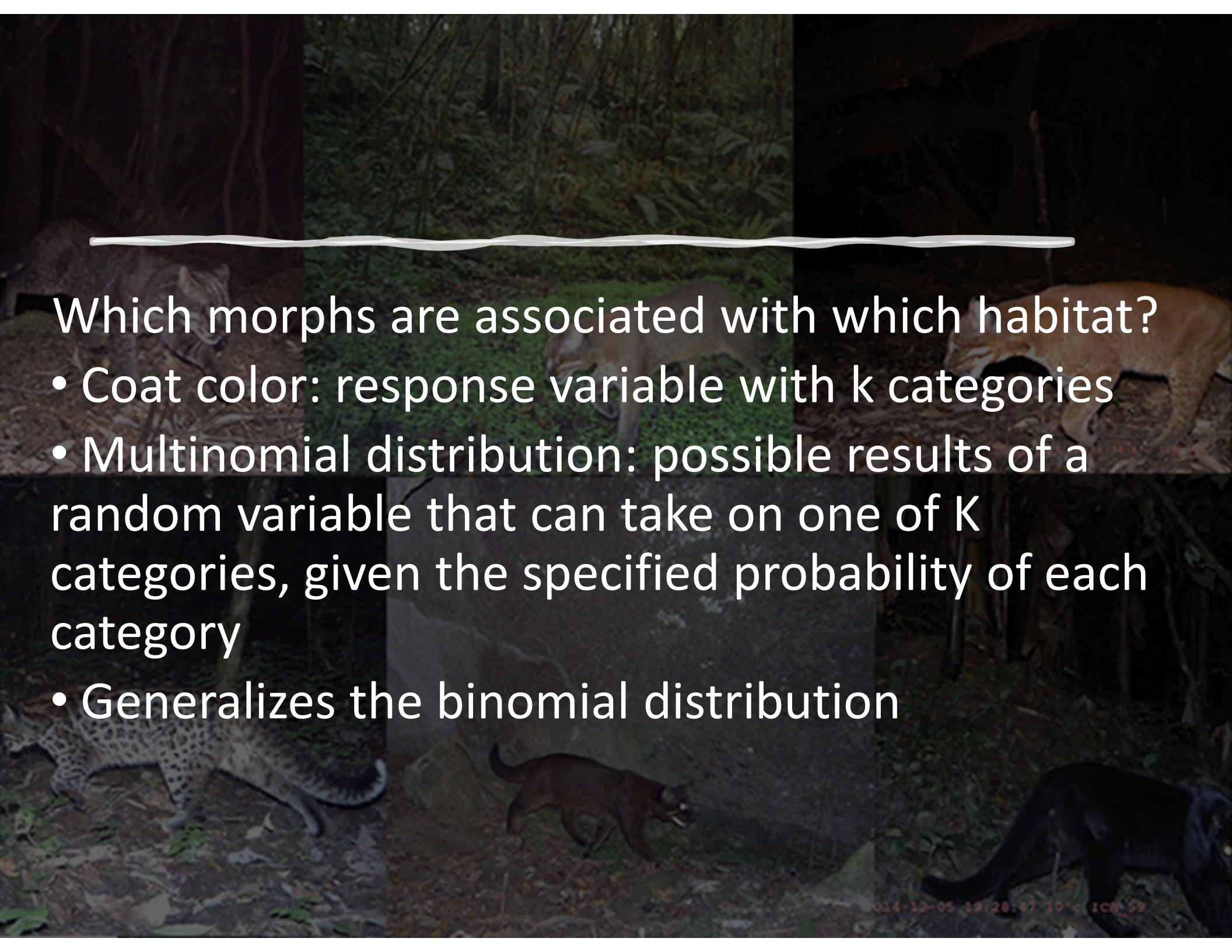
Binomial distribution

- Series of Bernoulli trials
- Bernouilli trials: only 2 outcomes
- Binary variable: 1/0
- # of 1 in n independent Bernoulli trials
- Defined by n (# trials) and p (probability of success in 1 trial)





Categorical distributions

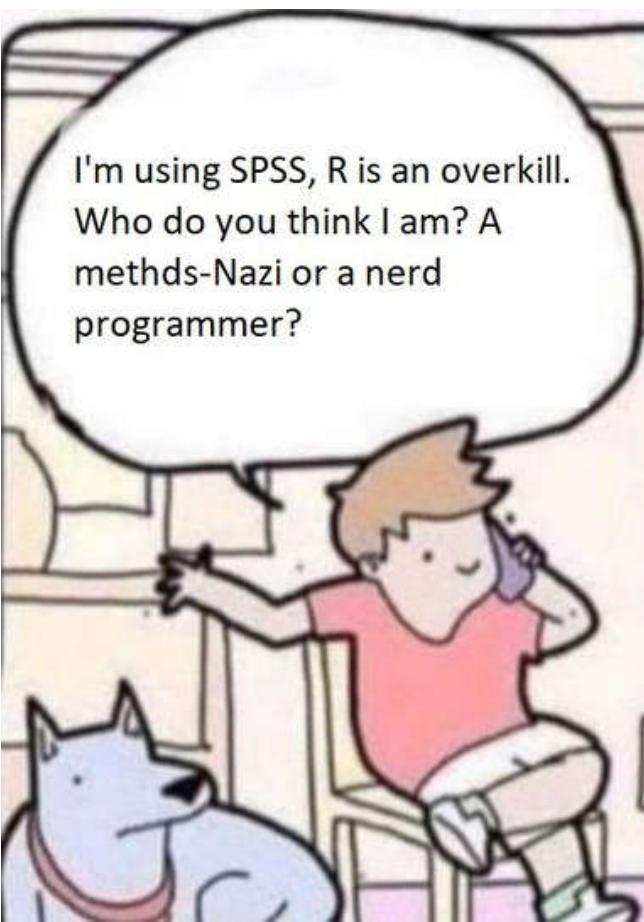


Which morphs are associated with which habitat?

- Coat color: response variable with k categories
- Multinomial distribution: possible results of a random variable that can take on one of K categories, given the specified probability of each category
- Generalizes the binomial distribution

Intro to R and RStudio

What is R?



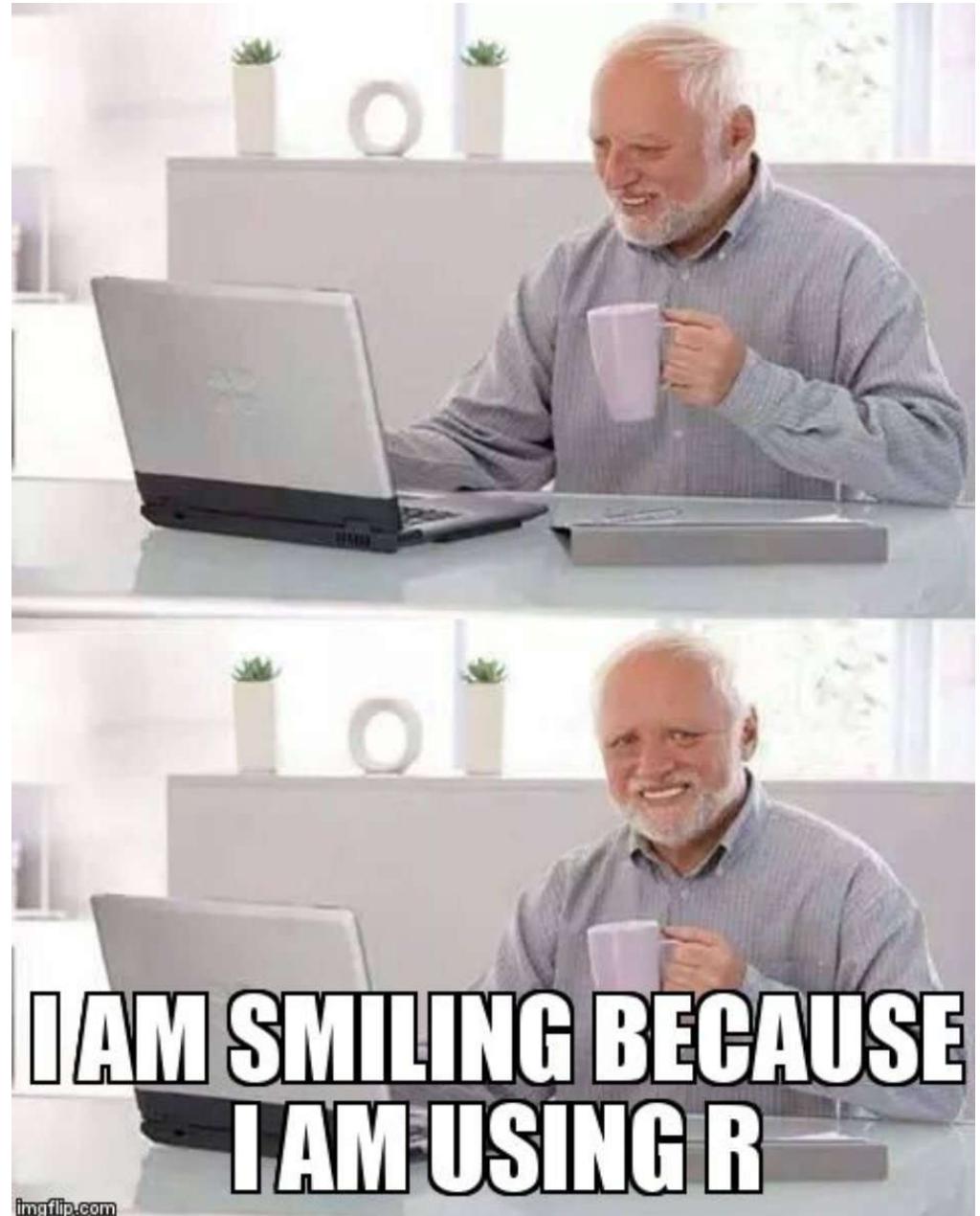
I'm using SPSS, R is an overkill.
Who do you think I am? A
methds-Nazi or a nerd
programmer?

A cartoon illustration of a person with short brown hair, wearing a red t-shirt, sitting at a desk and looking thoughtful. A large thought bubble above them contains the text from the previous slide.

- A programming language
- A calculator
- An environment: integrated suite of software
- Open source
- Get under the hood, unlike most software: you know what you're doing and what's going on

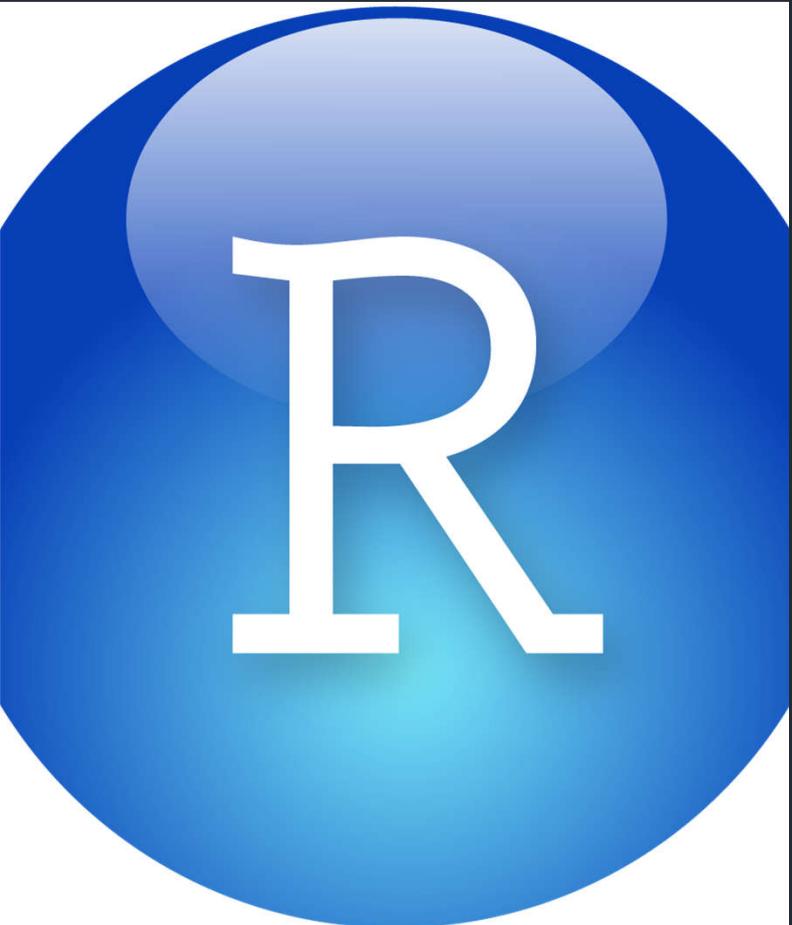
Why is R?

- Data handling
- Computation
- Graphics
- Wide variety of:
 - Statistical techniques
 - Graphical techniques
 - Data-type it can handle



What is RStudio?

WHEN YOUR CODE FINALLY
RUNS AFTER 50 FAILED ATTEMPTS



- integrated development environment for R
- A tool to make your life easier with R

Why is RStudio?

tools for:

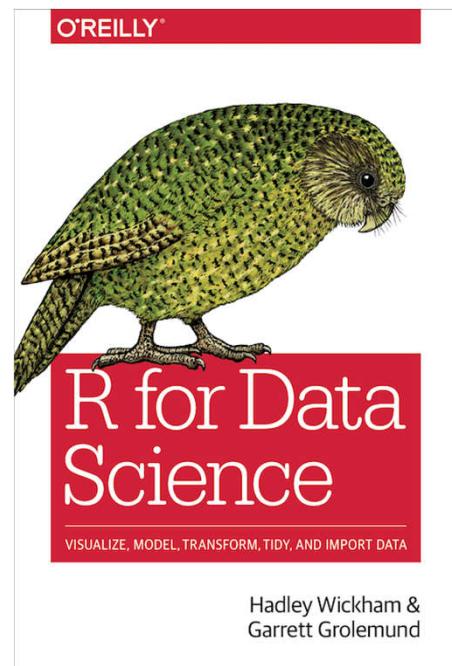
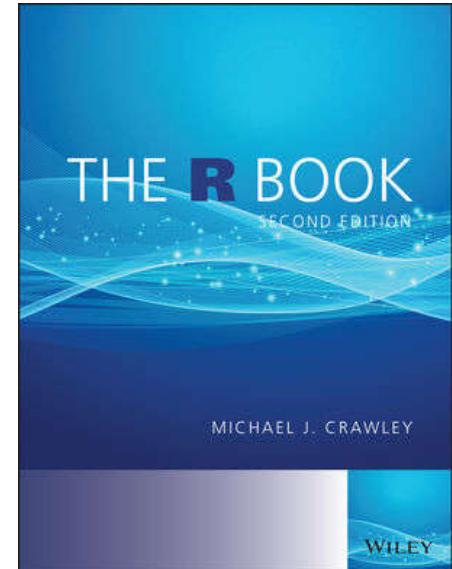
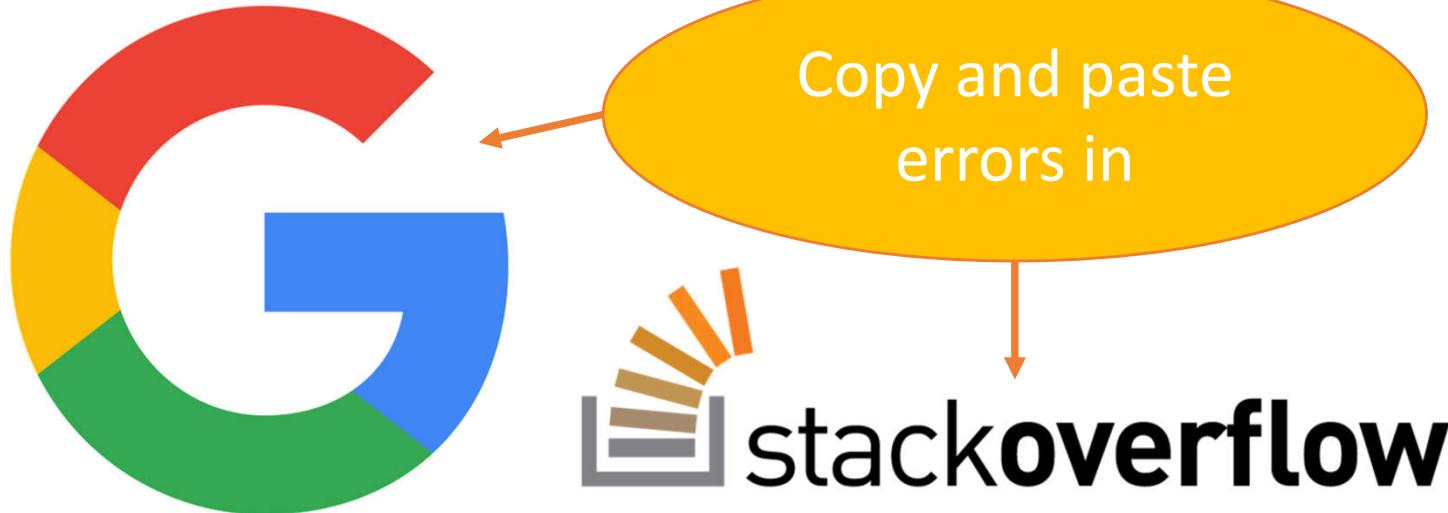
- plotting
- history
- debugging
- workspace management

sometimes you just need the comforting warm embrace of rstudio #rstats



R resources

- <https://r4ds.had.co.nz/> for data wrangling and graphs
- List of sites I provided during the last seminar
- I'm happy to help you trouble shoot your code, but from now on, **I'm expecting you to first try these resources**



Panes in RStudio

The screenshot shows the RStudio interface with four main panes, each highlighted by a large blue circle containing a number:

- Pane 1:** The top-left pane, which is the default workspace for R code. It contains the R console output, showing the R version, license information, and workspace details.
- Pane 2:** The bottom-left pane, which is the R script editor. It displays the R code used in the session.
- Pane 3:** The top-right pane, which is the Environment browser. It lists the objects currently loaded in the R environment, including their names, types, and sizes.
- Pane 4:** The bottom-right pane, which is the Global Environment viewer. It provides a visual representation of the objects in the environment, showing their names and values.

RStudio Menu Bar: File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.

Environment Tab: Shows the global environment with the following objects:

Object	Description
a.co1	1500 obs. of 2 variables
bp.d	List of 7
butterfly	15 obs. of 2 variables
df	500 obs. of 5 variables
df2	1500 obs. of 1 variables
e.co1	1000 obs. of 2 variables
f.co1	2500 obs. of 2 variables
h.co1	5000 obs. of 2 variables
ls	15000 obs. of 1 variable
small1.df	15 obs. of 1 variables
values	
a	chr [1:1500] "agricultural" "agricultural" "agricultural" ...
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.463 2.636 180.772 0.292 3883.884 ...
bimodalpois	num [1:10000] 2 20 29 12 27 6 4 19 ...

Bottom Taskbar: Type here to search, Start button, Taskbar icons (File Explorer, Edge, Google Chrome, Mail, RStudio, Settings, Teams, Power), and system status indicators (Battery, Network, Volume, Language, Date/Time).

1

Source editor

The screenshot shows the RStudio interface with the following components:

- Header Bar:** Contains the RStudio logo, a search bar labeled "Go to file/function", and an "Addins" dropdown.
- Toolbar:** Includes icons for file operations like New, Open, Save, and Print, along with Run, Source, and other tools.
- Left Panel:** Shows a list of open files: "Gout_drivers_in_sg.Rmd", "Salpyngitis_20230330.R", "Intro to statistic using R - session 1.Rmd", "Intro to statistic using R - session 2.Rmd", and "Untitled1".
- Code Editor:** A large central area with the placeholder text "This is where you type your code". Below it, sample R code is shown:

```
# will prevent whatever comments you add to be read as code  
# you need to add the # symbol at the beginning of each line, which  
# you don't want to run
```
- Environment View:** On the right, it shows the current environment with objects like "a.col", "bp.d", "butterfly", "df", "df2", "e.col", "f.col", "h.col", "ls", "small1.df", and "values". It also lists "a", "a1", "a25", "bimodalno", and "bimodalpo".
- Console:** At the bottom, it shows the R console output:

```
1:1 (Top Level) :  
Console Terminal Background Jobs  
R 4.2.2 · ~/R_seminar_series_stats for vets/R codes/  
R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"  
Copyright (c) 2022 The R Foundation for Statistical Computing
```

1

Source editor

The screenshot shows the RStudio interface with the Source editor active. The title bar says "RStudio". The menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The toolbar has icons for file operations like Open, Save, and Run, along with "Go to file/function" and "Addins". The main workspace shows several Rmd files open in tabs: "Gout_drivers_in_sg.Rmd", "Salpyngitis_20230330.R", "Intro to statistic using R - session 1.Rmd", "Intro to statistic using R - session 2.Rmd", and "Untitled1". Below the tabs are icons for "Source on Save", "Search", and "Edit". The status bar at the bottom shows "1:1 (Top Level)" and "R Script". The bottom-left corner shows the R console output: "R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"" and "Copyright (c) 2022 The R Foundation for Statistical Computing". The right side of the interface features the Environment browser, Data browser, and Files/Plots browser.

Source editor color convention: default theme

Black
Objects and anything you define or call

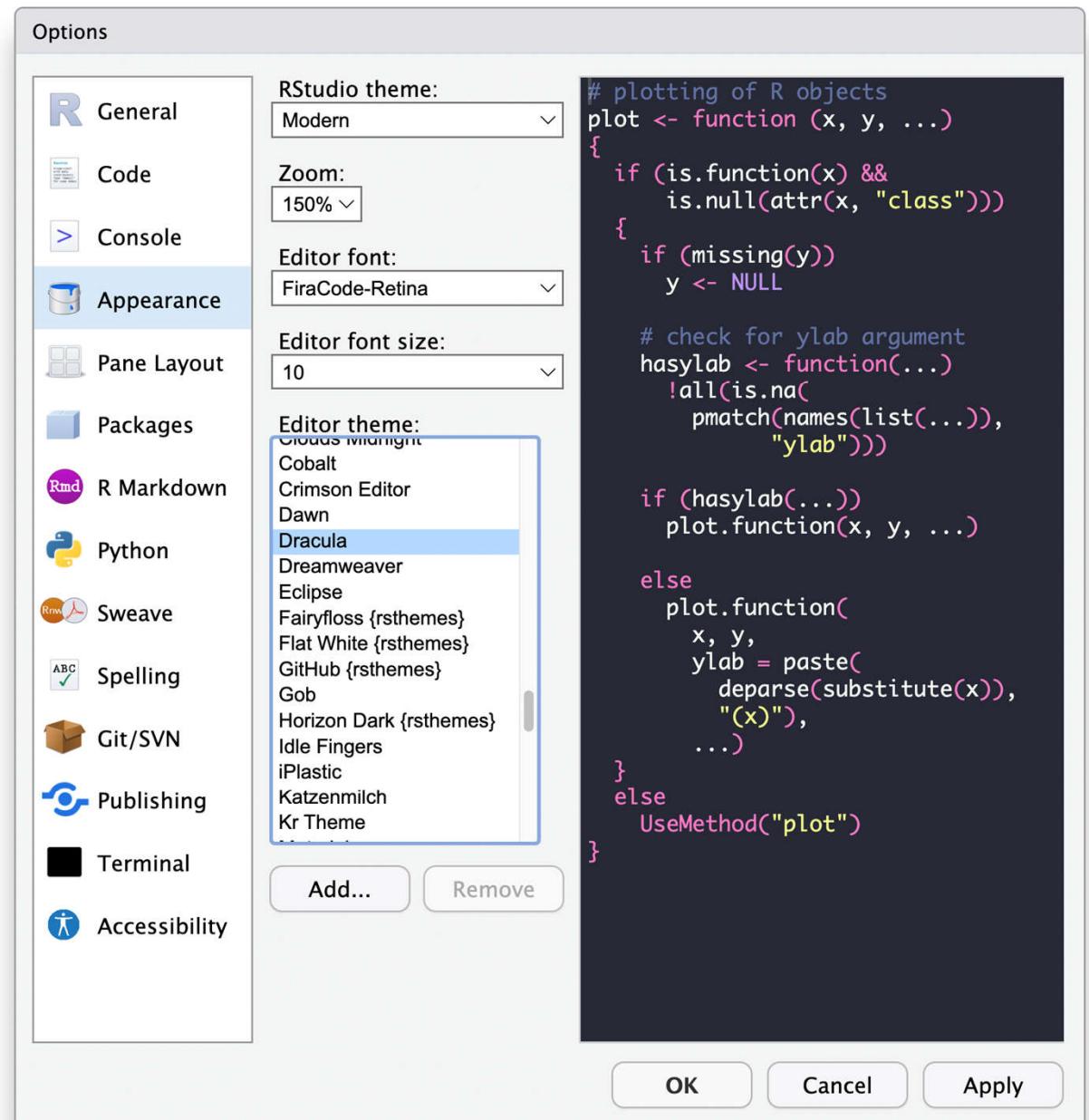
Blue
commands

Green
“characters”
will not be considered as code, thus won’t run

1

Source editor

But you
can change
the theme!



2

R Console

The screenshot shows the RStudio interface. The top bar includes tabs for 'Console' (selected), 'Terminal', and 'Background Jobs'. The status bar at the bottom indicates '4:1 (Top Level)'. The main area displays the R startup message and workspace information. A vertical sidebar on the right contains the 'Environment' pane with variables 'a', 'a1', 'a25', 'bim', and 'bim', and the 'Files' pane showing a file icon.

```
R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

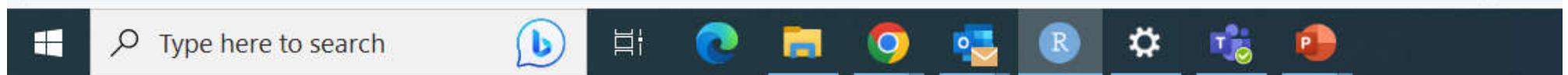
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[workspace loaded from ~/R_seminar_series_stats for vets/R codes/.RData]
```

> |

- > This is actual R
- > Your code runs here
- > All commands are blue preceded by >
- Output is black
- Warnings and errors are red



2

R Console

The screenshot shows the RStudio interface. The top bar includes tabs for 'Console' (which is selected), 'Terminal', and 'Background Jobs'. The status bar at the bottom indicates '4:1 (Top Level)'. On the right side, there's a sidebar titled 'Value' containing variables like 'a', 'a1', 'a25', 'bim', and 'bim'. Below that is a 'Files' section with icons for 'New File' and 'Open'. The main workspace shows the R startup message, workspace loading information, and a command-line history starting with '> |' followed by user input and output.

```
R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[workspace loaded from ~/R_seminar_series_stats for vets/R codes/.RData]

> |
> If you type code in the console
> It will be executed directly. Ex.:

> 3 + 2
[1] 5
```



Type here to search



3

Environment pane

The screenshot shows the RStudio interface with the 'Environment' tab selected in the top menu bar. The main pane displays the 'Global Environment' containing various R objects categorized under 'Data' and 'values'. A large blue callout bubble with white text is overlaid on the right side of the pane, pointing towards the list of objects. The text inside the bubble reads: 'Very useful' on the top line and 'Shows all your objects' on the bottom line.

Project: (None)

Environment History Connections Tutorial

Import Dataset 158 MiB

R Global Environment

Data

a.col	1500 obs. of 2 variables
bp.d	List of 7
butterfly	15 obs. of 5 variables
df	500 obs. of 2 variables
df2	1500 obs. of 2 variables
e.col	15 obs. of 2 variables
f.col	20 obs. of 2 variables
h.col	50 obs. of 2 variables
ls	15000 obs. of 2 variables
small.df	15 obs. of 2 variables

values

a	chr [1:1500] "agricultural" "agricultural" ...
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.463 2.636 180.352 0.292 3883.884 ...
bimodalpois	num [1:10000] 2 20 29 12 27 6 4 19 24 5 ...

Files Plots Packages Help Viewer Presentation

Zoom Export

3

Environment pane

Project: (None)

Environment History Connections Tutorial

Import Dataset 158 MiB

R Global Environment

Data

a.col	1500 obs. of 2 variables
bp.d	List of 7
butterfly	15 obs. of 5 variables
df	500 obs. of 1 variables
df2	1500 obs. of 2 variables
e.col	1500 obs. of 1 variables
f.col	21 obs. of 1 variables
h.col	500 obs. of 1 variables
ls	15000 objects
small.df	15 obs. of 2 variables

values

a	chr [1:1500] "agriculture" "agriculture" ...
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.463 2.636 180.352 0.292 3883.884 ...
bimodalpois	num [1:10000] 2 20 29 12 27 6 4 19 24 5 ...

Files Plots Packages Help Viewer Presentation

Zoom Export

Dataframes, matrices,
arrays and lists are under
“Data”

3

Environment pane

The screenshot shows the RStudio interface with the 'Environment' tab selected. The 'Global Environment' is chosen from the dropdown. The environment is divided into sections: 'Data' and 'values'. The 'values' section contains objects like 'a', 'a1', 'a25', 'bimodalnorm', and 'bimodalpois', each with its type and value.

Vectors or functions are under “values”

Object	Type	Value
a	chr	[1:1500] "agricultural" "agricultural..."
a1		0.01
a25		0.25
bimodalnorm	num	[1:10000] 0.463 2.636 180.352 0.292 3883.884 ...
bimodalpois	num	[1:10000] 2 20 29 12 27 6 4 19 24 5 ...

3

Environment pane

The pane indicates how many columns you have in your data frame, or the number of elements in a list

Variable	Description
a.col	1500 obs. of 2 variables
bp.d	List of 7
butterfly	15 obs. of 2 variables
\$ site	chr "site1" "site1" "site1" "site1" ...
\$ color	chr "orange" "blue" "orange" "yellow" ...
df	500 obs. of 5 variables
df2	1500 obs. of 1 variables
e.col	1000 obs.
f.col	2500 obs.
h.col	5000 obs.
ls	15000
small.df	15 obs.
values	
a	chr [1:1]
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.40

3

Environment pane

The screenshot shows the RStudio interface with the 'Environment' tab selected. The 'Global Environment' dropdown is open, showing a list of variables. A red circle highlights the 'bp.d' entry, which is described as 'List of 7'. A large blue callout bubble points from this highlighted text towards the bottom right of the slide. The environment pane also lists other variables like 'a.col', 'butterfly', 'df', etc., each with its size and type information.

Variable	Description
a.col	1500 obs. of 2 variables
bp.d	List of 7
butterfly	15 obs. of 2 variables
\$ site	site1 "site1" "site1" "site1" ...
\$ color	"orange" "blue" "orange" "yellow" ...
df	500 obs. of 5 variables
df2	1500 obs. of 1
e.col	1000 obs.
f.col	2500 obs.
h.col	5000 obs.
ls	15000
small.df	15 obs.
values	
a	chr [1:1]
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.40

You can also see how many observations you have in your columns

3

Environment pane

The screenshot shows the RStudio interface with the 'Environment' tab selected. The 'Global Environment' dropdown is open, showing a list of objects. A large blue oval callout points to the entry for 'butterfly'. The 'butterfly' object is a list with two elements: '\$ site' and '\$ color'. The 'Data' section below lists various other objects like 'a.col', 'bp.d', 'df', etc., each with their type and size information.

Object	Type	Description
a.col	1500 obs. of 2 variables	
bp.d	List of 2	
butterfly	15 obs.	\$ site : chr "site1" "site1" "site1" "site1" "site1" ... \$ color: chr "orange" "blue" "red" "green" "yellow" ...
df	500 obs. of 1 variable	
df2	1500 obs. of 1 variable	
e.col	1000 obs. of 1 variable	
f.col	2500 obs. of 1 variable	
h.col	5000 obs. of 2 variables	
ls	15000 obs. of 1 variable	
small.df	15 obs. of 1 variable	
values		
a	chr [1:1500]	"agricultural" "agricultural" "agricultural" "agricultural" "agricultural" ...
a1	0.01	
a25	0.25	
bimodalnorm	num [1:10000]	0.463 2.636 180.352 0.292 3883.884 ...

3

Environment pane

The screenshot shows the RStudio interface with the 'Environment' tab selected. The 'Data' section lists various objects:

Object	Type	Description
a.col	1500 obs. of 2 variables	
bp.d	List of 7	
butterfly	15 obs.	\$ site : chr "site1" "site1" "site1" ... \$ color: chr "orange" "blue" "green" ...
df	500 obs.	
df2	1500 obs.	
e.col	1000 obs.	
f.col	2500 obs. of 1 variable	
h.col	5000 obs. of 2 variables	
ls	15000 obs. of 1 variable	
small.df	15 obs. of 1 variable	

A blue callout bubble points to the 'butterfly' object, specifically highlighting the expanded view of its components (\$ site and \$ color). A red circle is drawn around the arrow icon to the right of 'butterfly', indicating where it can be clicked to expand or collapse the object's details.

When you click on the arrow, you can see:

- Columns of data frames
 - Elements of a list

3

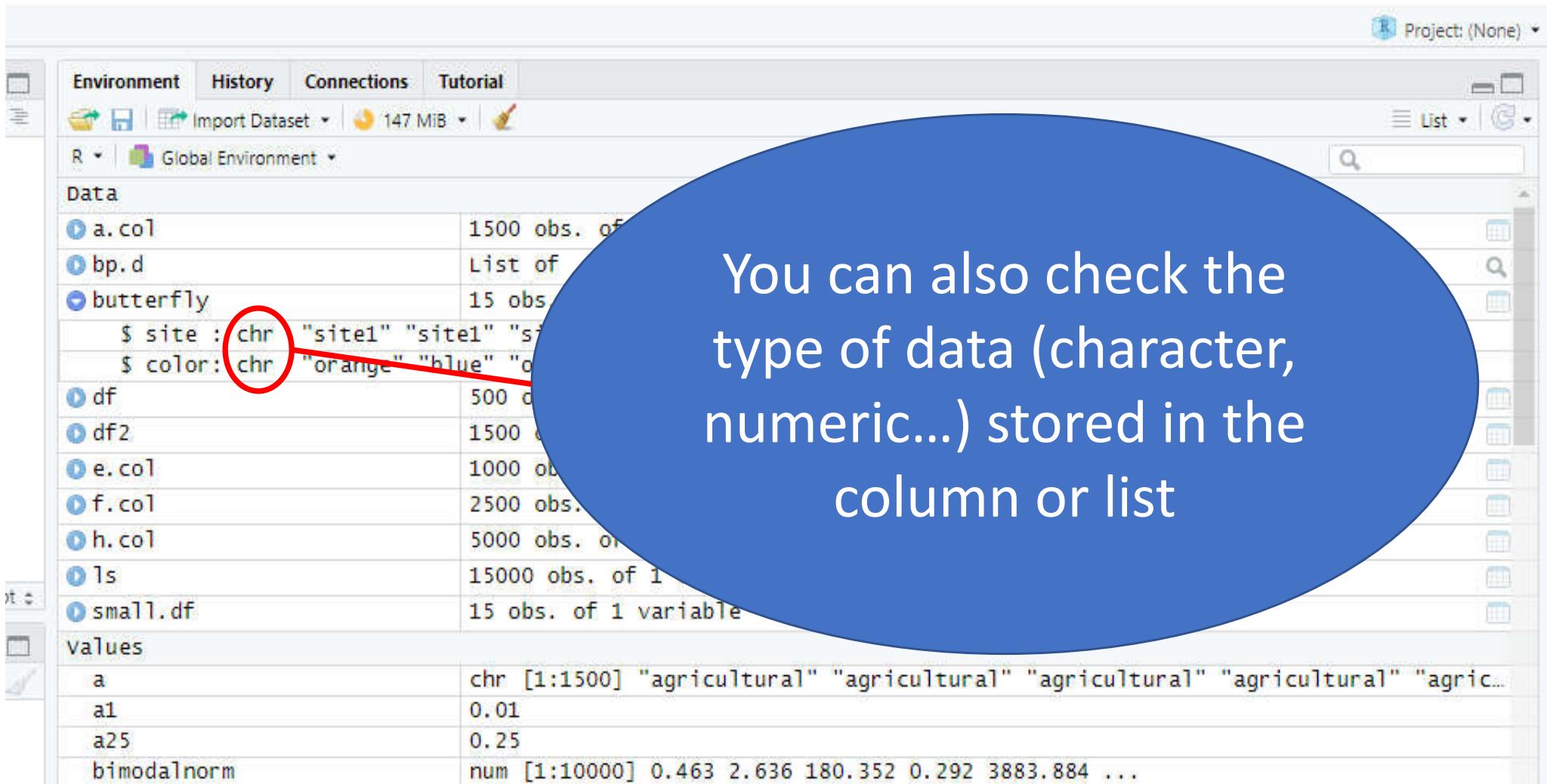
Environment pane

R also indicates what operator you must use to call a particular element. For ex., `butterfly$site` calls the column “site” of data frame “butterfly”

Object	Type	Description
a.col	1500 obs.	
bp.d	List of	
butterfly	15 ob	
\$ site : chr "site1" "site1" "		
\$ color: chr "orange" "blue" "		
df	500	
df2	1500	
e.col	1000 ob	
f.col	2500 ob	
h.col	5000 obs.	
ls	15000 obs. o	
small.df	15 obs. of 1 va	
values		
a	chr [1:1500] "agricultural"	
a1	0.01	
a25	0.25	
bimodalnorm	num [1:10000] 0.463 2.636 180.352 0.292 3883.884 ...	

3

Environment pane



The screenshot shows the RStudio Environment pane. The top menu bar includes 'Environment', 'History', 'Connections', and 'Tutorial'. Below the menu, there are icons for 'Import Dataset', '147 MiB', and a paintbrush. The 'Global Environment' dropdown is set to 'Data'. The main area displays a list of objects:

Object	Type	Description
a.col	1500 obs. of	
bp.d	List of	
butterfly	15 obs.	
\$ site : chr	"site1" "site1" "site1"	
\$ color: chr	"orange" "blue" "orange"	
df	500 obs.	
df2	1500 obs.	
e.col	1000 obs.	
f.col	2500 obs.	
h.col	5000 obs. of	
ls	15000 obs. of 1	
small.df	15 obs. of 1 variable	
values		
a	chr [1:1500]	"agricultural" "agricultural" "agricultural" "agricultural" "agric...
a1	0.01	
a25	0.25	
bimodalnorm	num [1:10000]	0.463 2.636 180.352 0.292 3883.884 ...

A large blue callout bubble points to the 'butterfly' object's 'site' column, with the text: 'You can also check the type of data (character, numeric...) stored in the column or list'. A red circle highlights the 'chr' type in the 'site' column entry for 'butterfly'.

3

Environment pane

Same for vectors, you can check the type of data

Object	Type	Description
a.col	1500 obs.	
bp.d	List of	
butterfly	15 obs.	\$ site : chr "site1" "site1" "site1" ... \$ color: chr "orange" "blue" "orange" ...
df	500 obs.	
df2	1500 obs.	
e.col	1000 obs. of 2 variables	
f.col	2500 obs. of 2 variables	
h.col	5000 obs. of 2 variables	
ls	15000 obs. of 1 variable	
small.df	15 obs. of 1 variable	
values		
a	chr [1:1500]	"agricultural" "agricultural" "agricultural" "agricultural" "agricultural" ...
a1	0.01	
a25	0.25	
bimodalnorm	num [1:10000]	0.463 2.636 180.352 0.292 3883.884 ...

3

Environment pane

Project: (None) ▾

Environment History Connections Tutorial

Import Dataset 147 MiB

R Global Environment

Data

a.col	1500 obs.
bp.d	List of
butterfly	15 obs. \$ site : chr "site1" "site1" "site1" ... \$ color: chr "orange" "blue" "orange" ...
df	500 obs.
df2	1500 obs. of 2 variables
e.col	1000 obs. of 2 variables
f.col	2500 obs. of 2 variables
h.col	5000 obs. of 2 variables
ls	15000 obs. of 1 variable
small.df	15 obs. of 1 variable

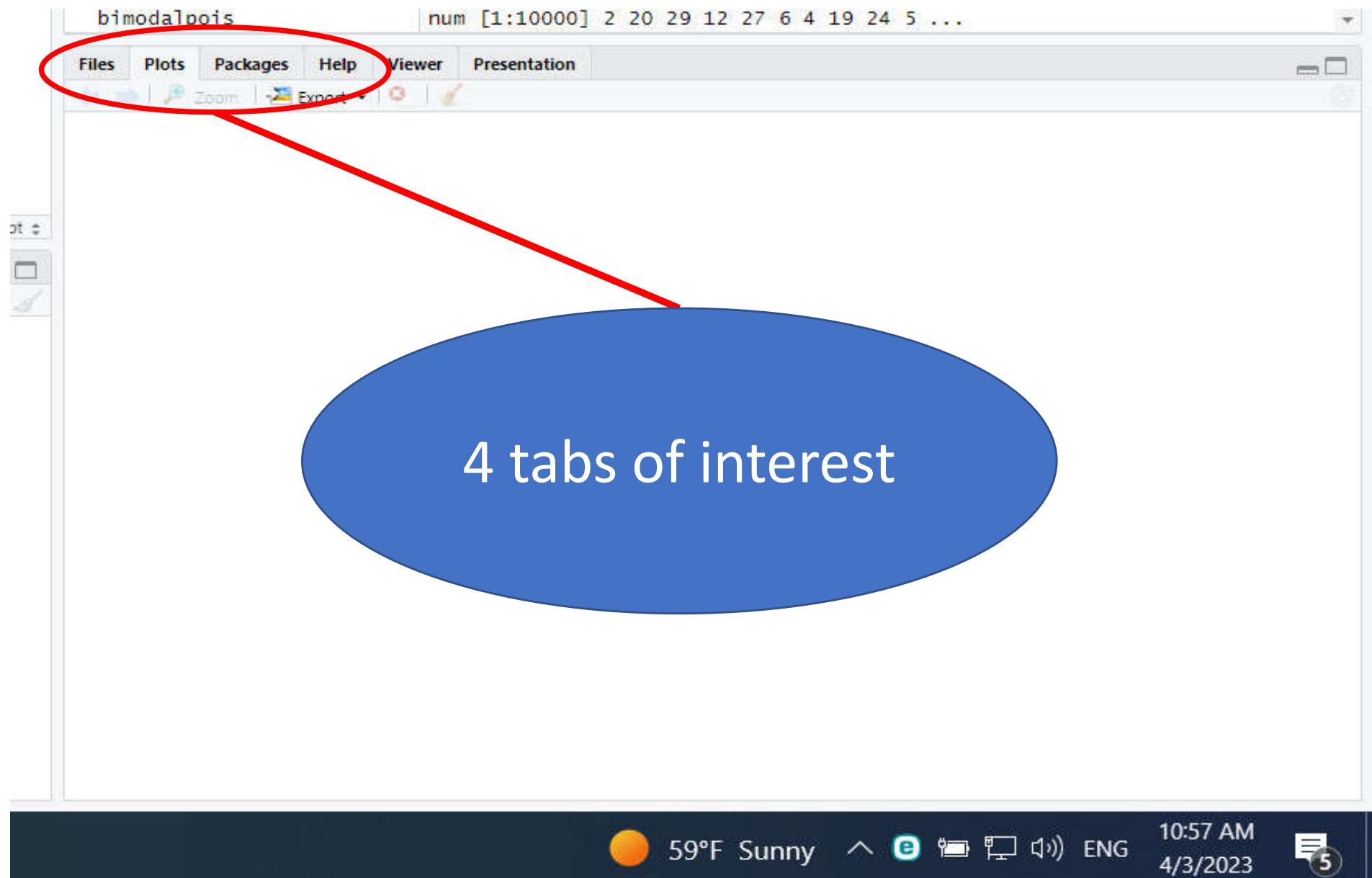
values

a	chr [1:1500] "agricultural" "agricultural" "agricultural" "agricultural" "agric...
a1	0.01
a25	0.25
bimodalnorm	num [1:10000] 0.463 2.636 180.352 0.292 3883.884 ...

Same for vectors, you can check the type of data, and how many observations you have

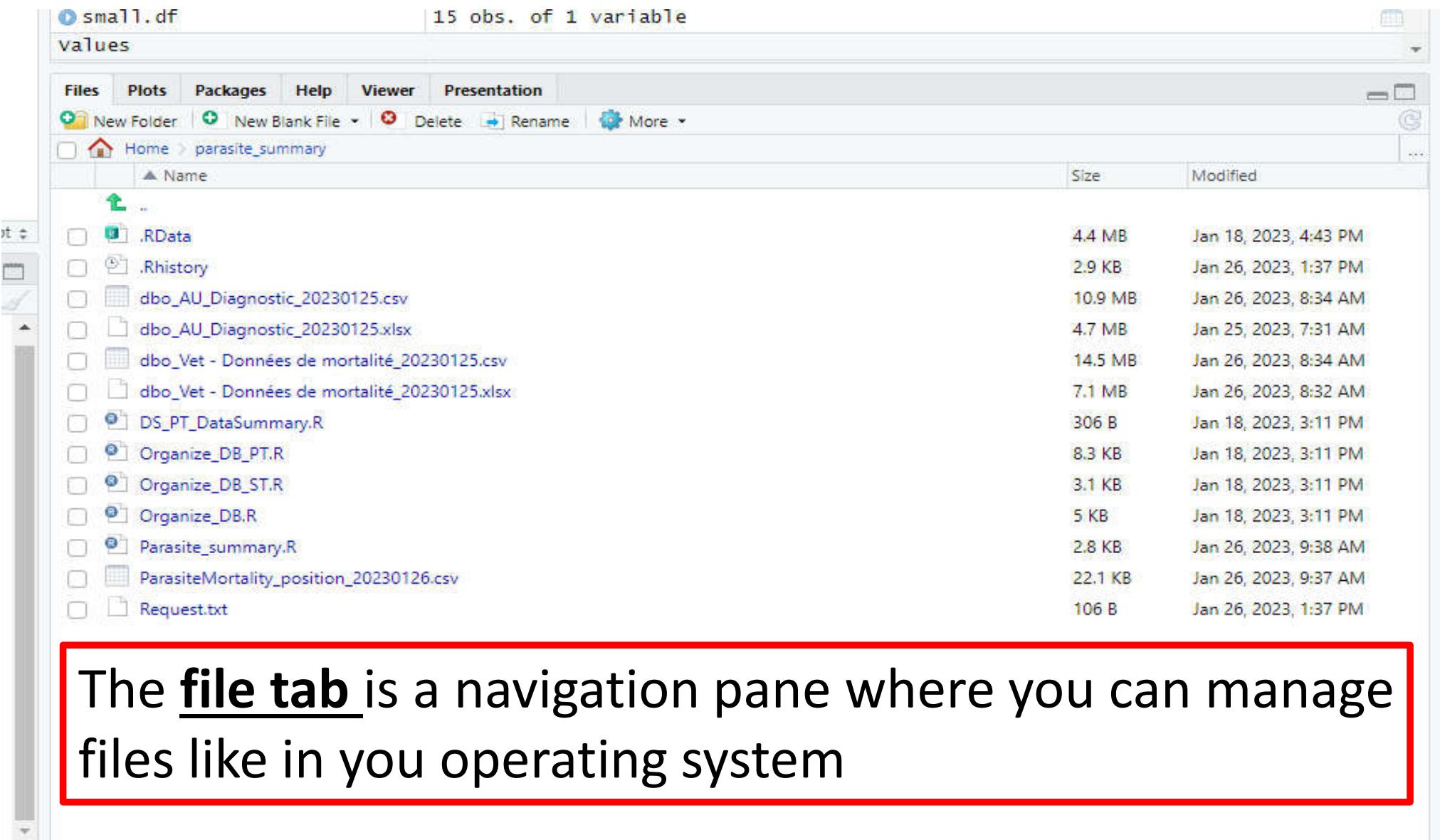
4

Output pane



4

Output pane



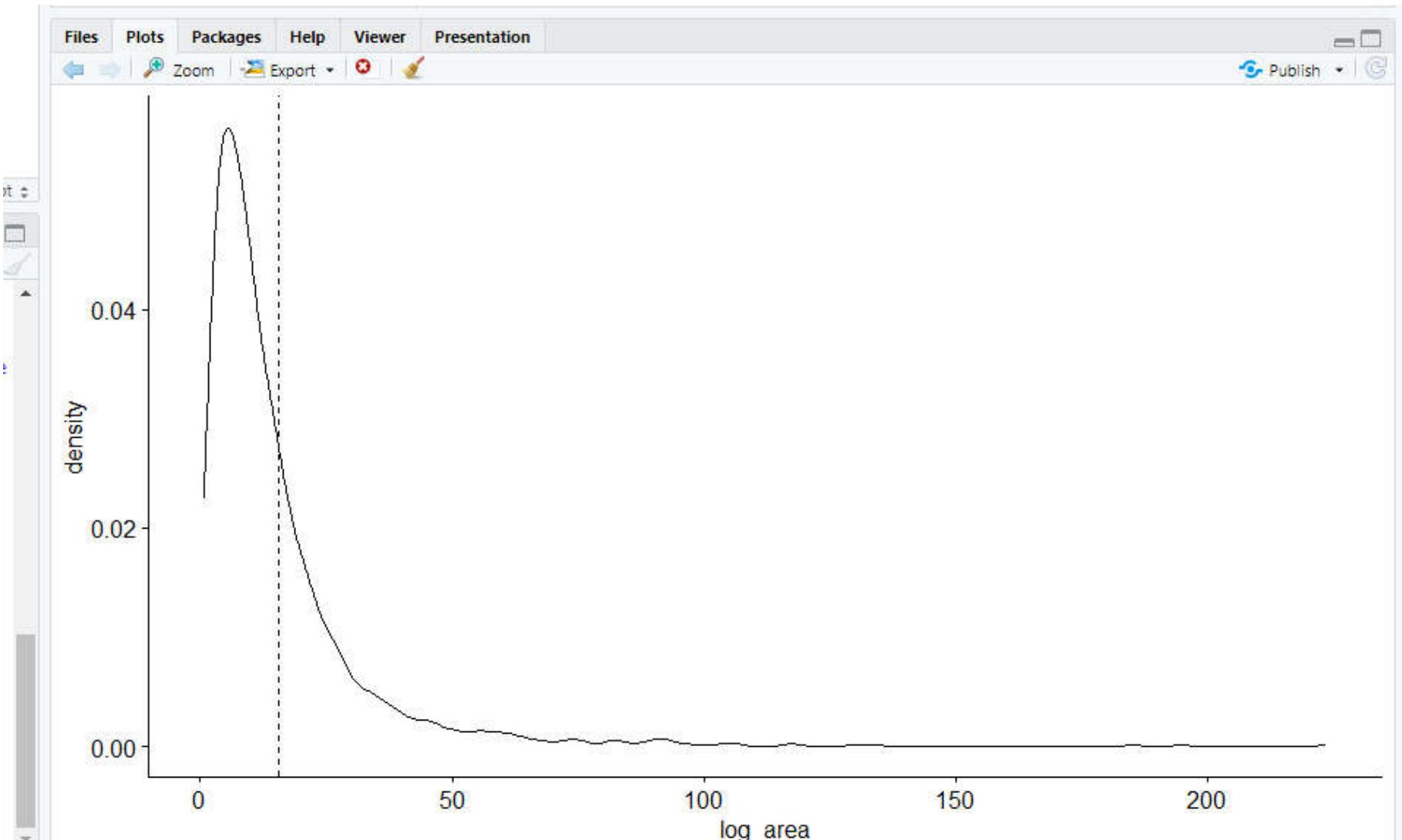
The screenshot shows the RStudio interface with a large blue number 4 in the top-left corner. The main title is "Output pane". Below it is a screenshot of the RStudio environment. The top bar shows a file named "small.df" with "15 obs. of 1 variable". The "values" tab is selected. The menu bar includes "Files", "Plots", "Packages", "Help", "Viewer", and "Presentation". The "File" tab is highlighted with a red box. The file tree shows a directory structure under "parasite_summary". The table below lists files with their names, sizes, and last modified dates.

Name	Size	Modified
..		
.RData	4.4 MB	Jan 18, 2023, 4:43 PM
.Rhistory	2.9 KB	Jan 26, 2023, 1:37 PM
dbo_AU_Diagnostic_20230125.csv	10.9 MB	Jan 26, 2023, 8:34 AM
dbo_AU_Diagnostic_20230125.xlsx	4.7 MB	Jan 25, 2023, 7:31 AM
dbo_Vet - Données de mortalité_20230125.csv	14.5 MB	Jan 26, 2023, 8:34 AM
dbo_Vet - Données de mortalité_20230125.xlsx	7.1 MB	Jan 26, 2023, 8:32 AM
DS_PT_DataSummary.R	306 B	Jan 18, 2023, 3:11 PM
Organize_DB_PT.R	8.3 KB	Jan 18, 2023, 3:11 PM
Organize_DB_ST.R	3.1 KB	Jan 18, 2023, 3:11 PM
Organize_DB.R	5 KB	Jan 18, 2023, 3:11 PM
Parasite_summary.R	2.8 KB	Jan 26, 2023, 9:38 AM
ParasiteMortality_position_20230126.csv	22.1 KB	Jan 26, 2023, 9:37 AM
Request.txt	106 B	Jan 26, 2023, 1:37 PM

The **file tab** is a navigation pane where you can manage files like in your operating system

4

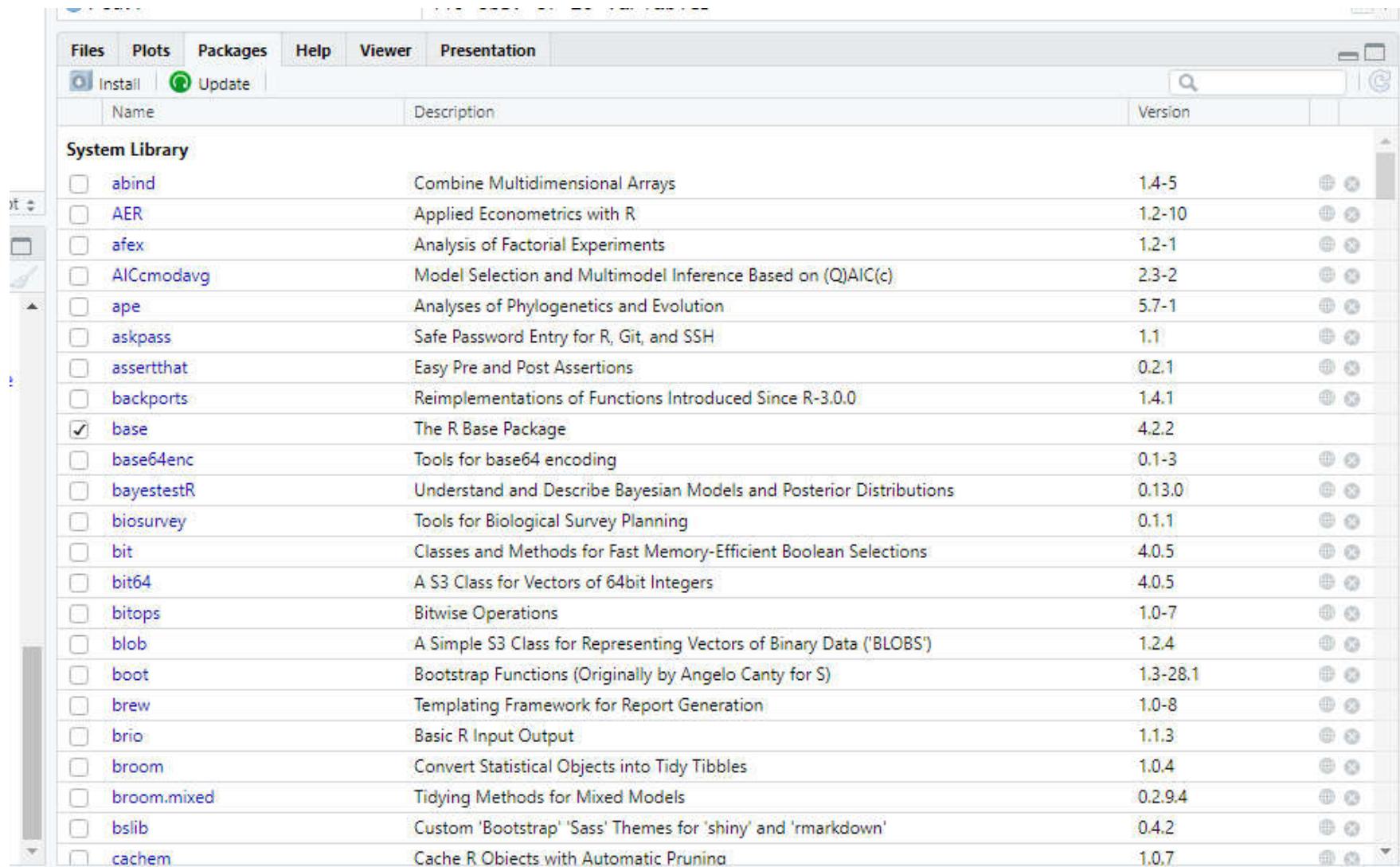
Output pane



The **plot tab** is where you can see your plots

4

Output pane



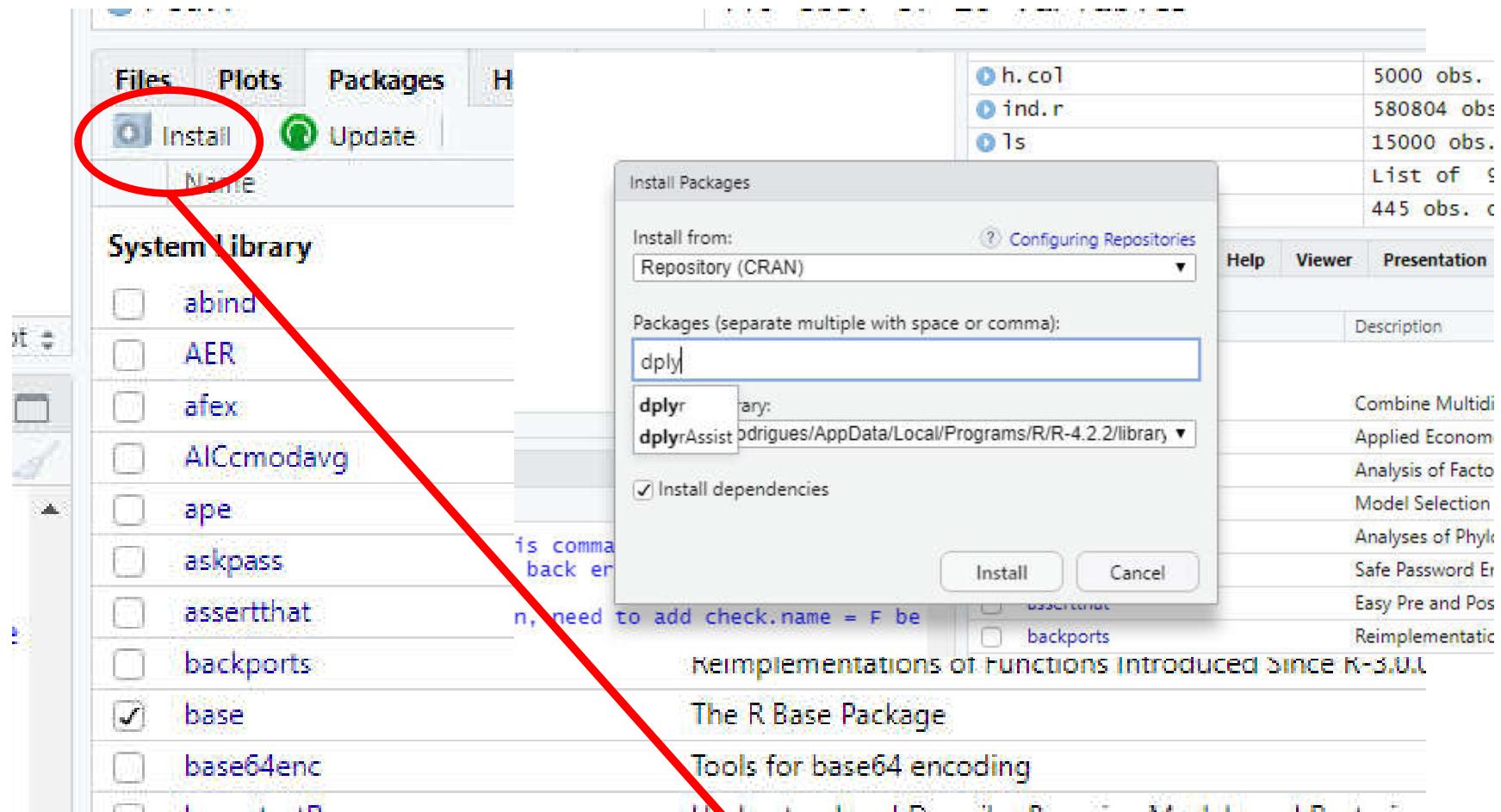
The screenshot shows the RStudio interface with the 'Packages' tab selected in the top menu bar. The main area displays a table titled 'System Library' listing various R packages. The columns in the table are 'Name', 'Description', and 'Version'. Each row includes a checkbox column on the far left. The package 'base' is checked. The table lists numerous packages such as abind, AER, afex, AICmodavg, ape, askpass, assertthat, backports, base64enc, bayestestR, biosurvey, bit, bit64, bitops, blob, boot, brew, brio, broom, broom.mixed, bslib, and cachem.

Name	Description	Version	
System Library			
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5	 
<input type="checkbox"/> AER	Applied Econometrics with R	1.2-10	 
<input type="checkbox"/> afex	Analysis of Factorial Experiments	1.2-1	 
<input type="checkbox"/> AICmodavg	Model Selection and Multimodel Inference Based on (Q)AIC(c)	2.3-2	 
<input type="checkbox"/> ape	Analyses of Phylogenetics and Evolution	5.7-1	 
<input type="checkbox"/> askpass	Safe Password Entry for R, Git, and SSH	1.1	 
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.1	 
<input type="checkbox"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.4.1	 
<input checked="" type="checkbox"/> base	The R Base Package	4.2.2	 
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3	 
<input type="checkbox"/> bayestestR	Understand and Describe Bayesian Models and Posterior Distributions	0.13.0	 
<input type="checkbox"/> biosurvey	Tools for Biological Survey Planning	0.1.1	 
<input type="checkbox"/> bit	Classes and Methods for Fast Memory-Efficient Boolean Selections	4.0.5	 
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	4.0.5	 
<input type="checkbox"/> bitops	Bitwise Operations	1.0-7	 
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.2.4	 
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-28.1	 
<input type="checkbox"/> brew	Templating Framework for Report Generation	1.0-8	 
<input type="checkbox"/> brio	Basic R Input Output	1.1.3	 
<input type="checkbox"/> broom	Convert Statistical Objects into Tidy Tibbles	1.0.4	 
<input type="checkbox"/> broom.mixed	Tidying Methods for Mixed Models	0.2.9.4	 
<input type="checkbox"/> bslib	Custom 'Bootstrap' 'Sass' Themes for 'shiny' and 'rmarkdown'	0.4.2	 
<input type="checkbox"/> cachem	Cache R Objects with Automatic Pruning	1.0.7	 

The **package tab** shows the packages stored in your computer

4

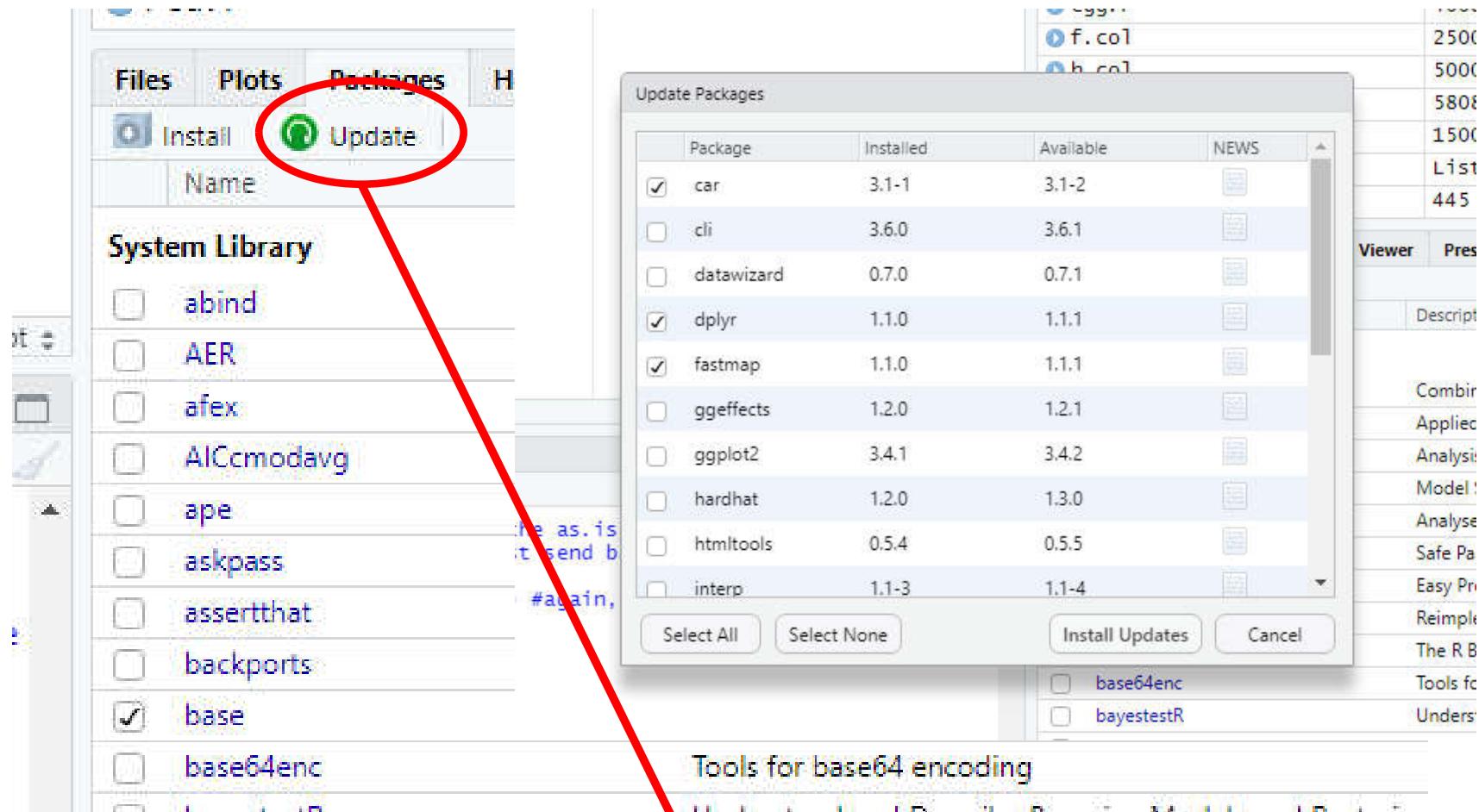
Output pane



The package tab also has the install button from which you can directly install packages

4

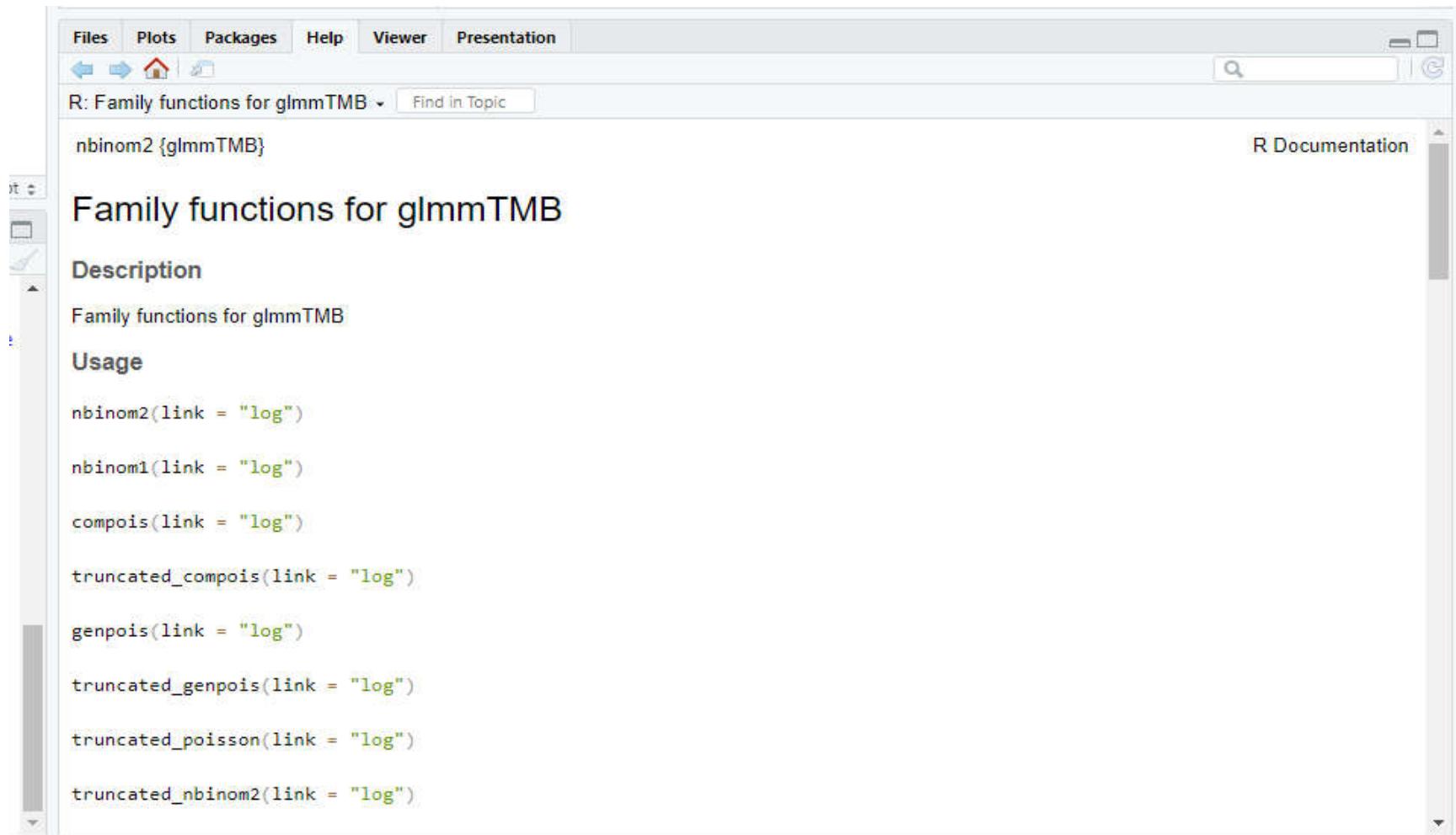
Output pane



Same to update already installed packages

4

Output pane



The **Help tab** provides access to R documentation for all functions

The letters "QR" are displayed in a large, bold, blue font. The letter "Q" is white with a gray outline, and the letter "R" is solid blue. They are positioned within a thick, light gray circle.

Intermission