

R Data Analysis Course – Week 5

Skews and Transformations



Transforming Data

Ronnie Steinitz, Ph.D.
Dian Fossey Gorilla Fund

African Marsh Harrier, Mpenge River, 2025

Resources to reinforce your understanding

The Google Drive folder includes:

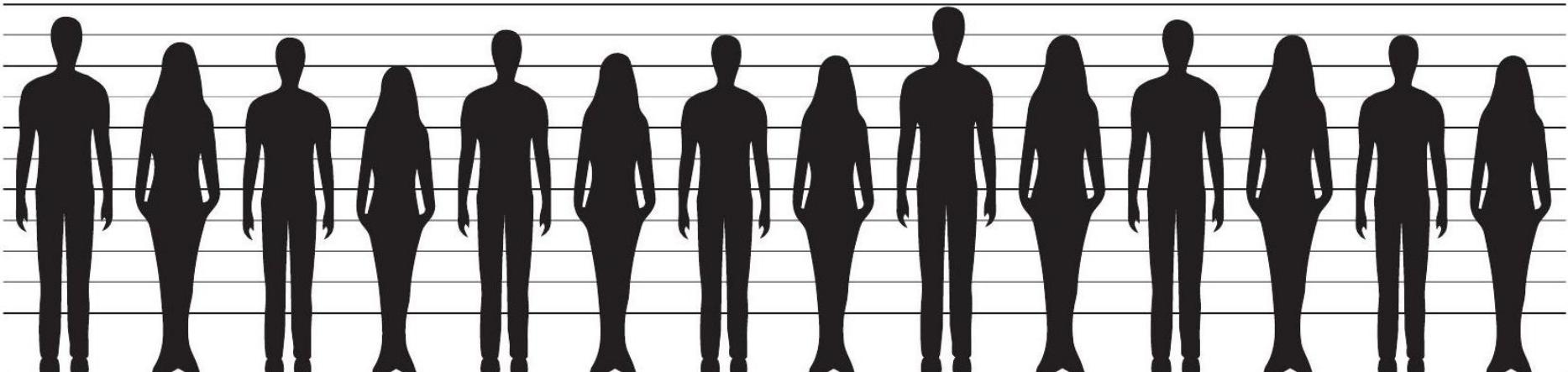
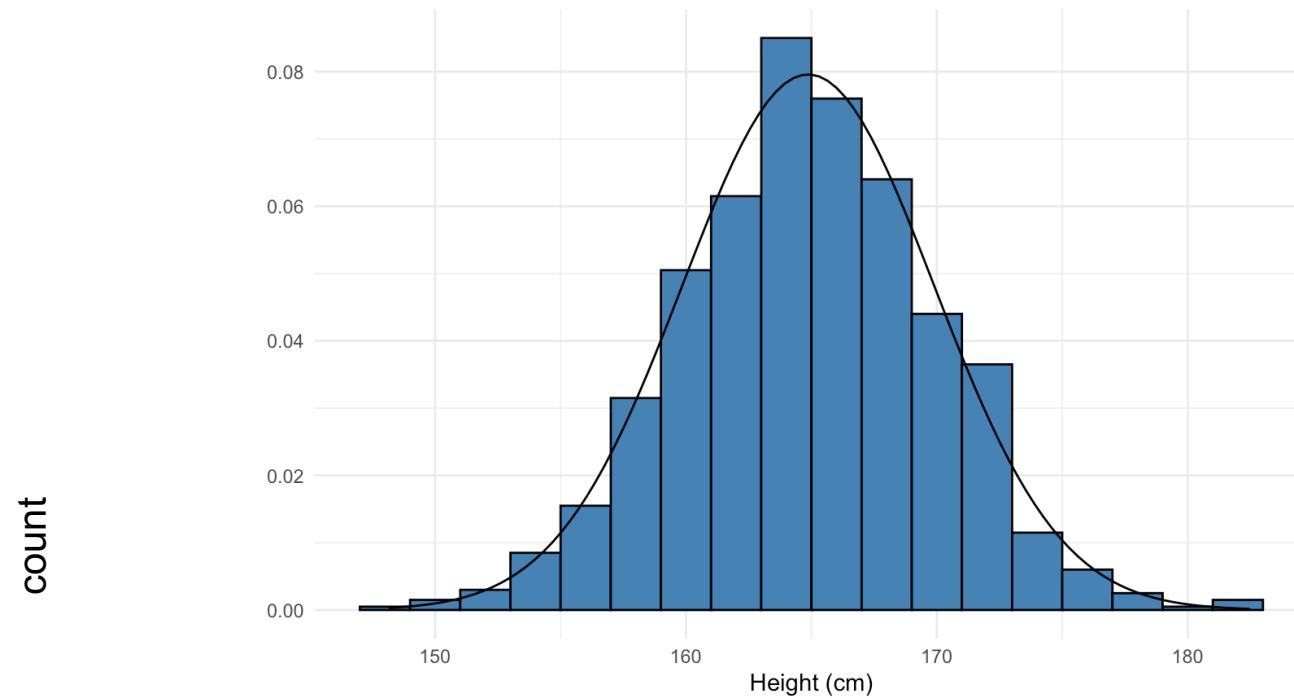
- Every week: a **complete .Rmd file** with detailed in-line explanations and example code from the lecture
- General: an updated “**Functions Learned So Far**” reference sheet

Every week: I offer help sessions: extra support with your code, project setup, or any course concepts

Distribution of Data

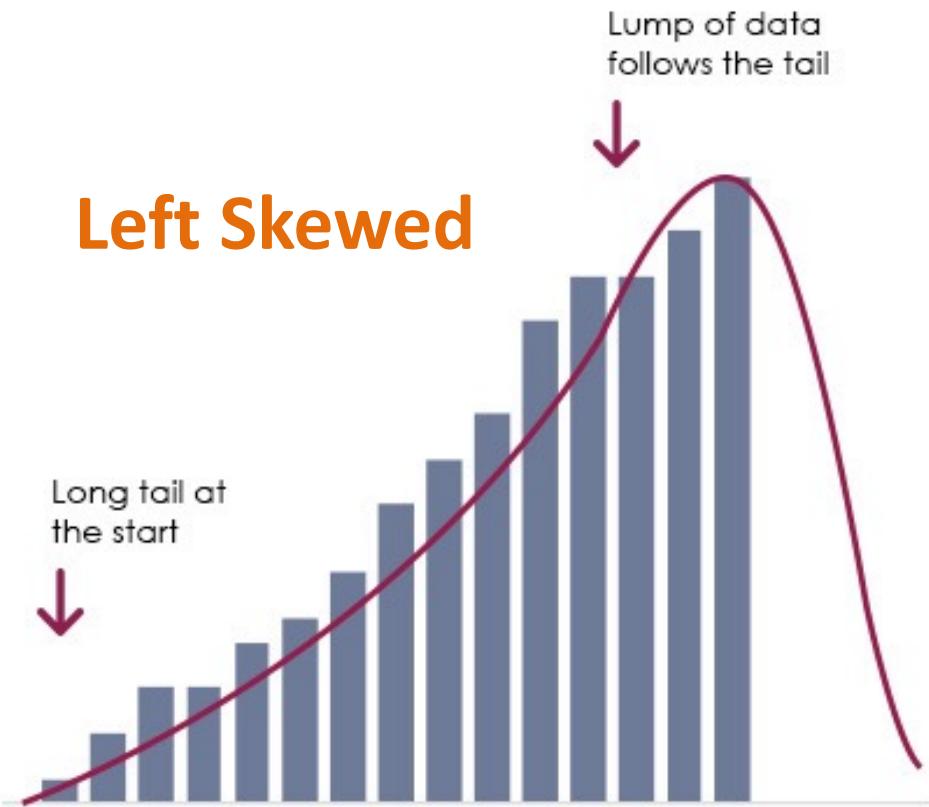
Histogram of adult height and normal curve

N = 1000, mean = 164.87, variance = 25.13



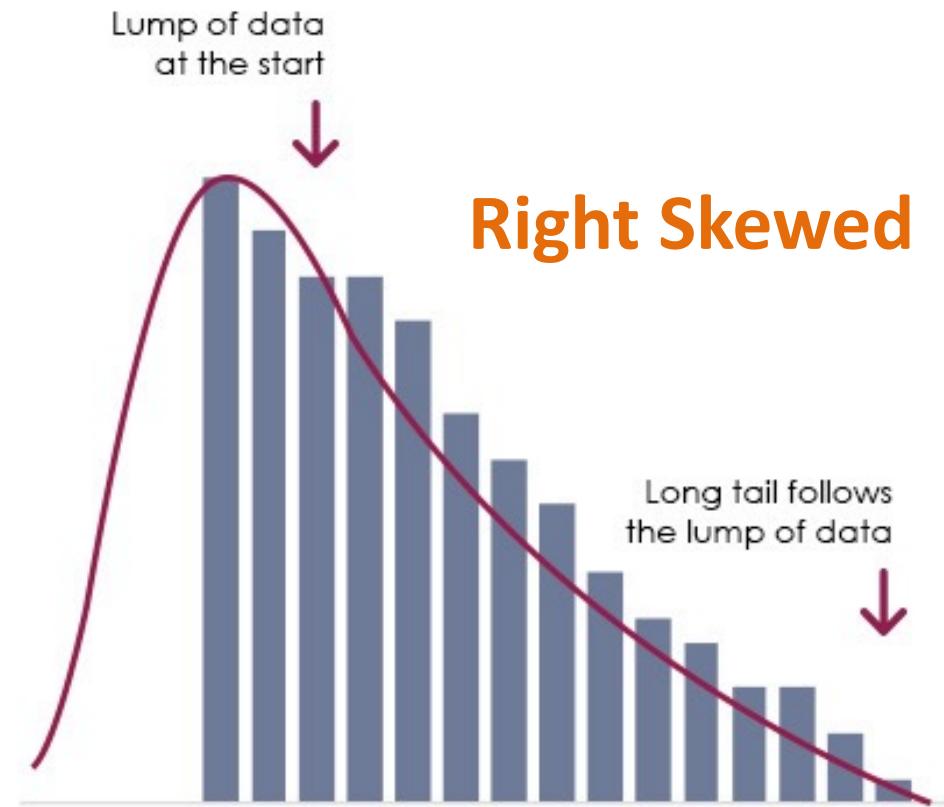
Skewed Data

Left Skewed



Negative skew
(left-skewed distribution)

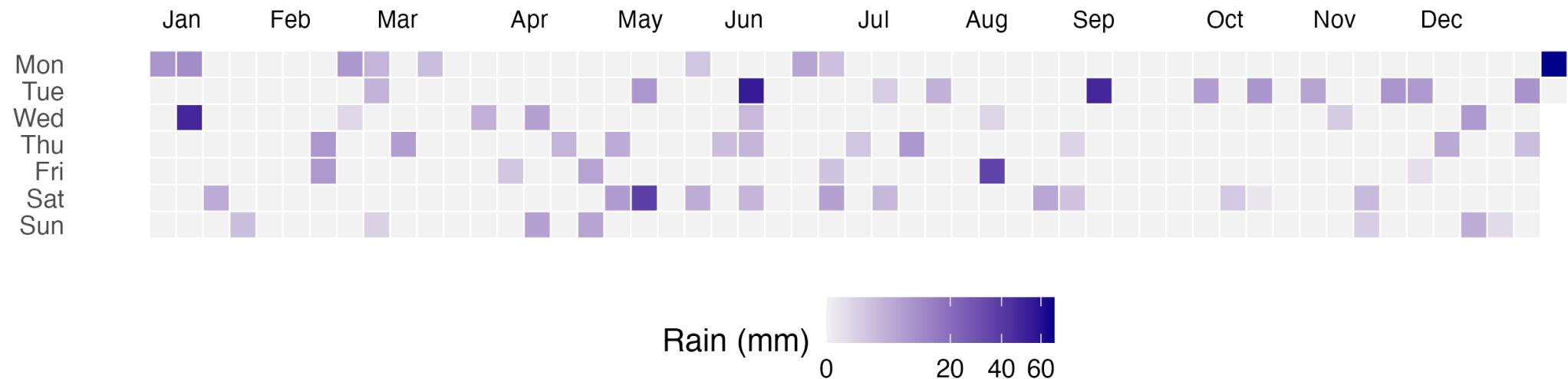
Right Skewed



Positive skew
(right-skewed distribution)

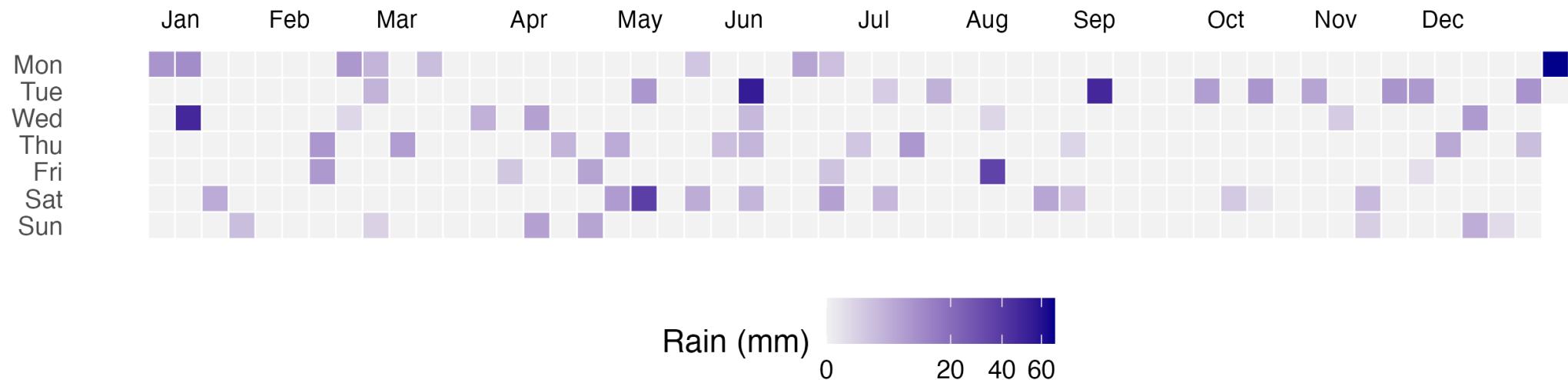
Daily Rainfall Calendar – Semi-Dry Environment

Most days are dry; occasional storms bring high rainfall

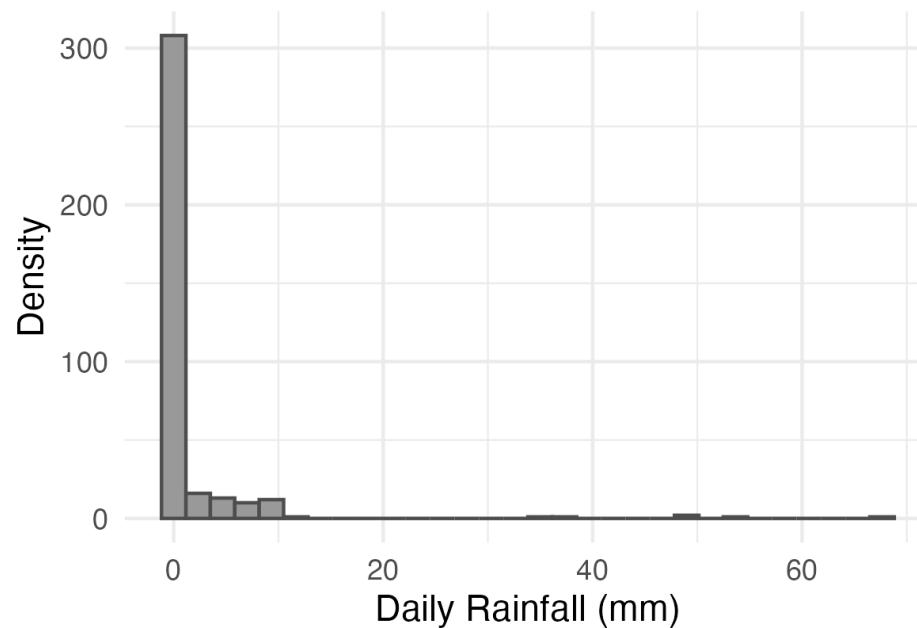


Daily Rainfall Calendar – Semi-Dry Environment

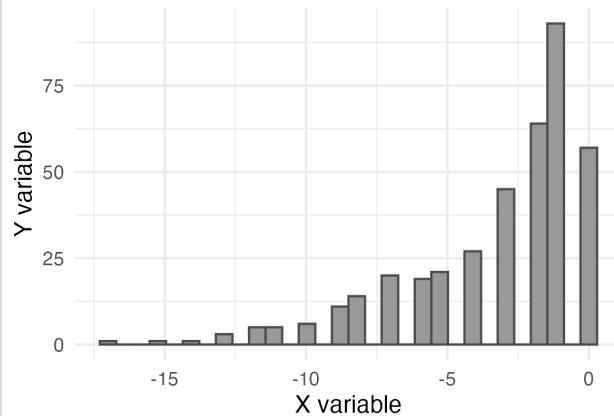
Most days are dry; occasional storms bring high rainfall



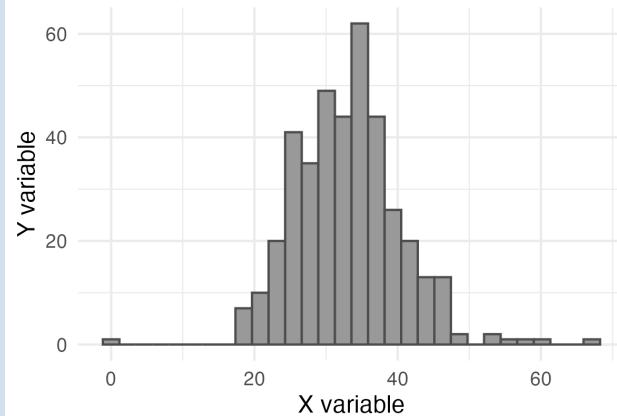
Daily Rainfall in Semi-Dry Environment



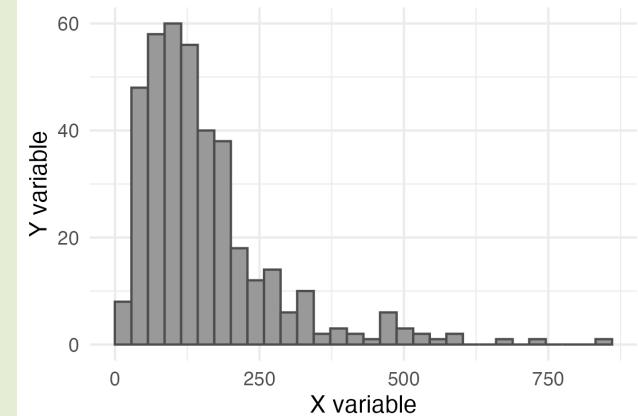
Histogram - Left-skewed



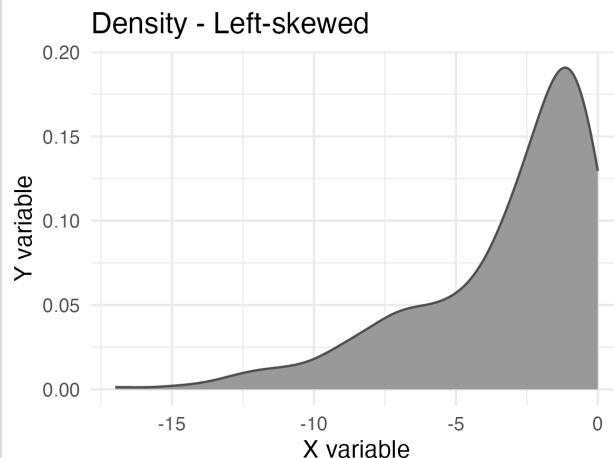
Histogram - Normal Distribution



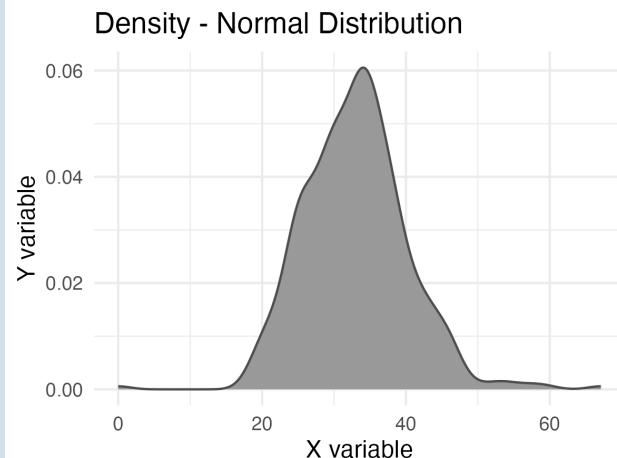
Histogram - Right-skewed



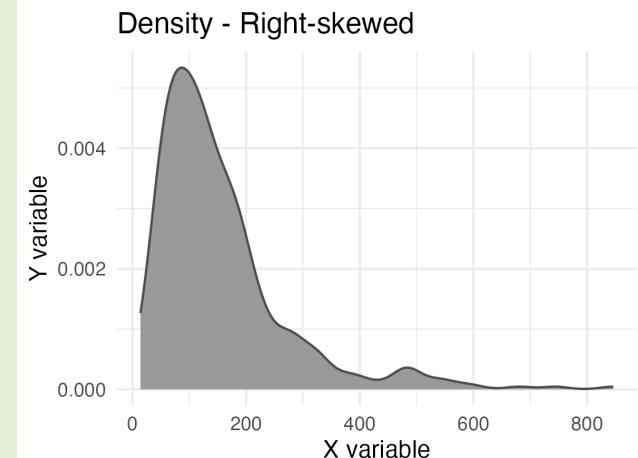
Density - Left-skewed



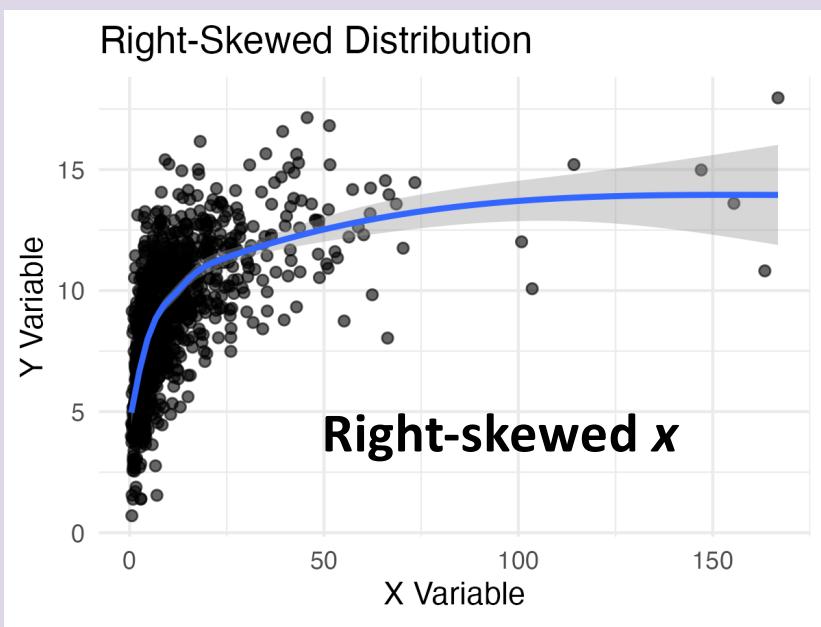
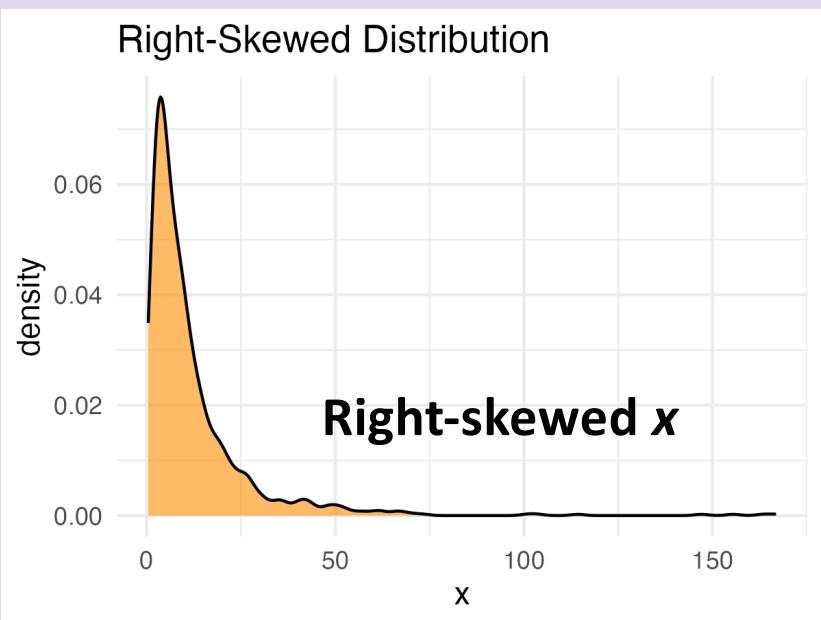
Density - Normal Distribution

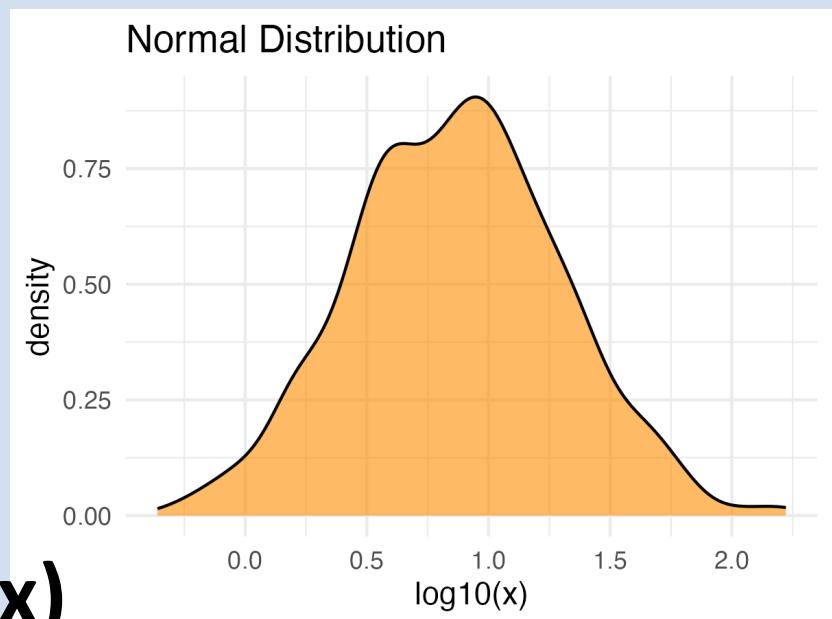
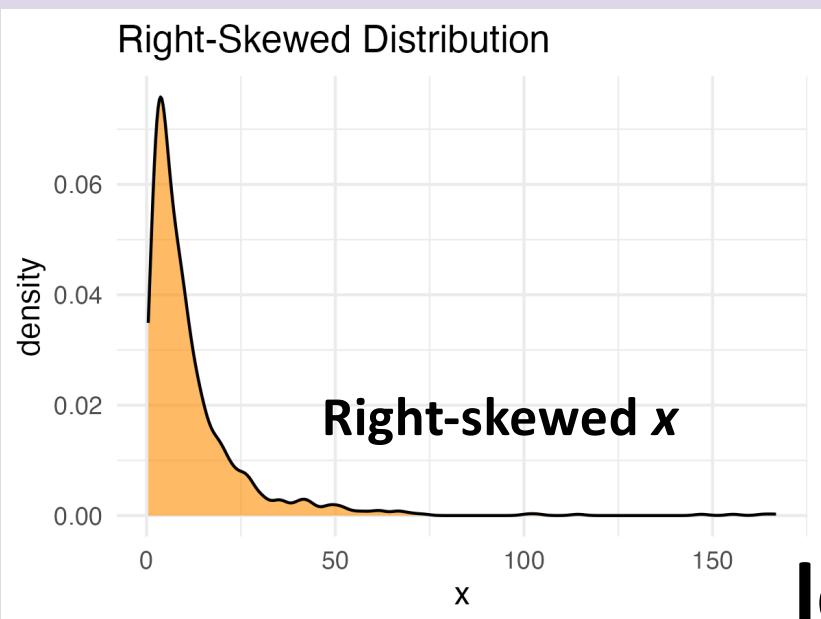


Density - Right-skewed

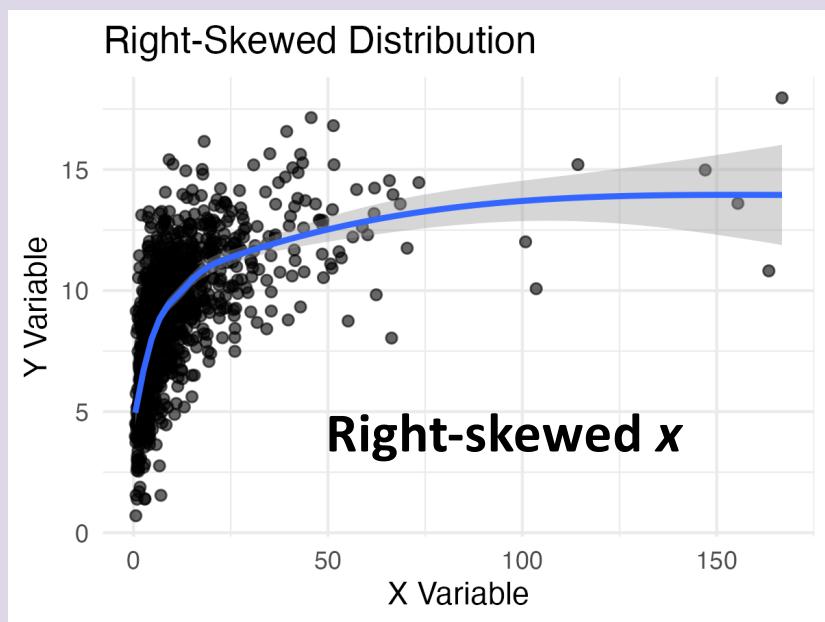


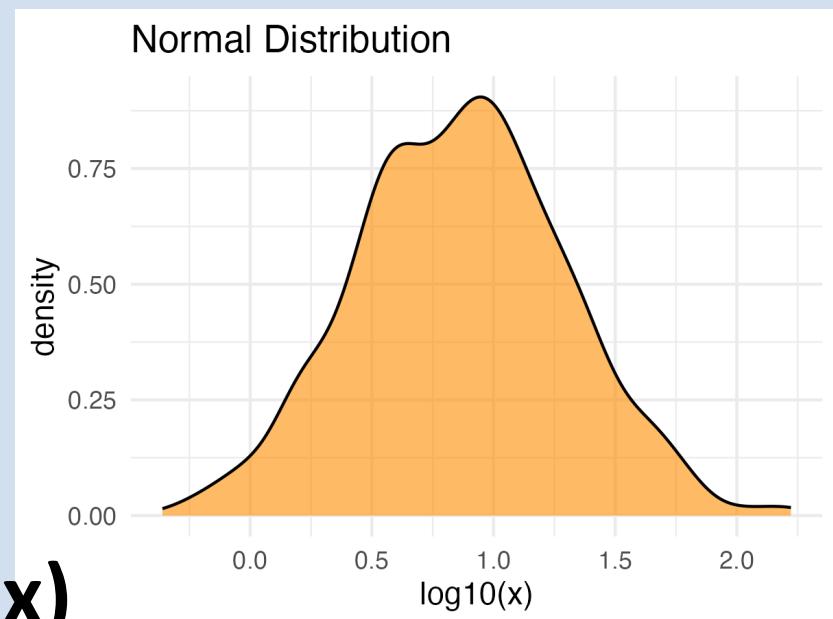
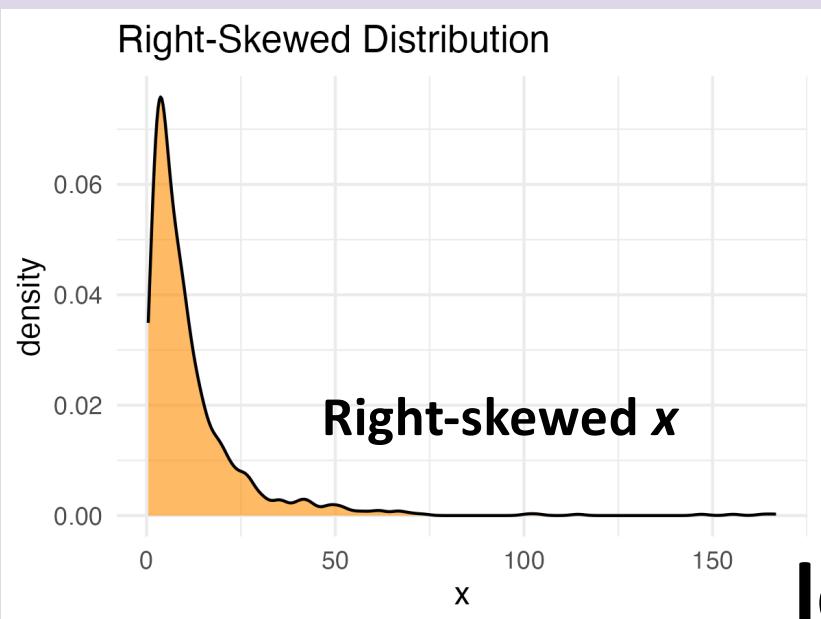
x (raw value)	log10(x)	Formula
0.1	-1.000	$10^{\text{(-1.000)}} = 0.1$
0.3	-0.523	$10^{\text{(-0.523)}} = 0.3$
0.8	-0.097	$10^{\text{(-0.097)}} = 0.8$
1	0.000	$10^{\text{(0.000)}} = 1$
2.5	0.398	$10^{\text{(0.398)}} = 2.5$
4	0.602	$10^{\text{(0.602)}} = 4$
9.5	0.978	$10^{\text{(0.978)}} = 9.5$
12	1.079	$10^{\text{(1.079)}} = 12$
25	1.398	$10^{\text{(1.398)}} = 25$
48	1.681	$10^{\text{(1.68)}} = 48$
120	2.079	$10^{\text{(2.079)}} = 120$
260	2.415	$10^{\text{(2.415)}} = 260$
480	2.681	$10^{\text{(2.681)}} = 480$
620	2.792	$10^{\text{(2.792)}} = 620$
950	2.978	$10^{\text{(2.978)}} = 950$
1000	3.000	$10^{\text{(3.000)}} = 1000$



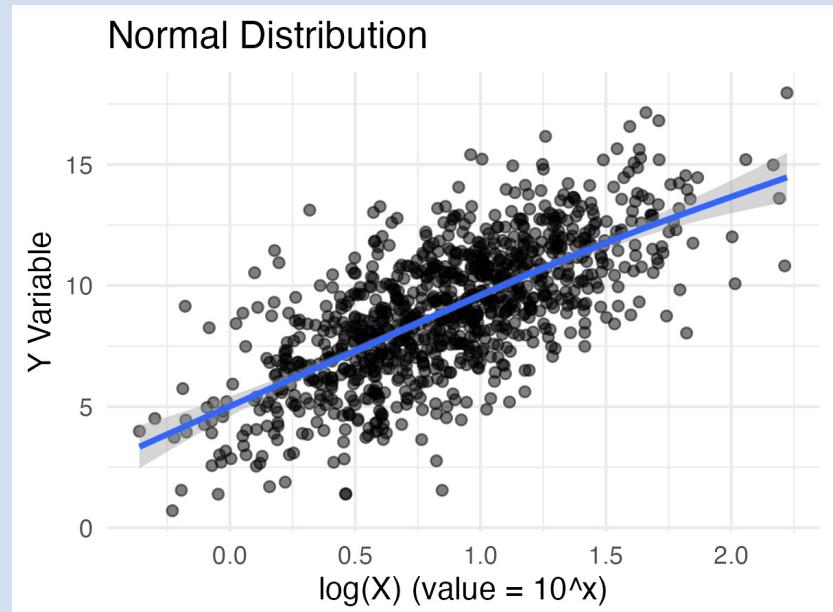
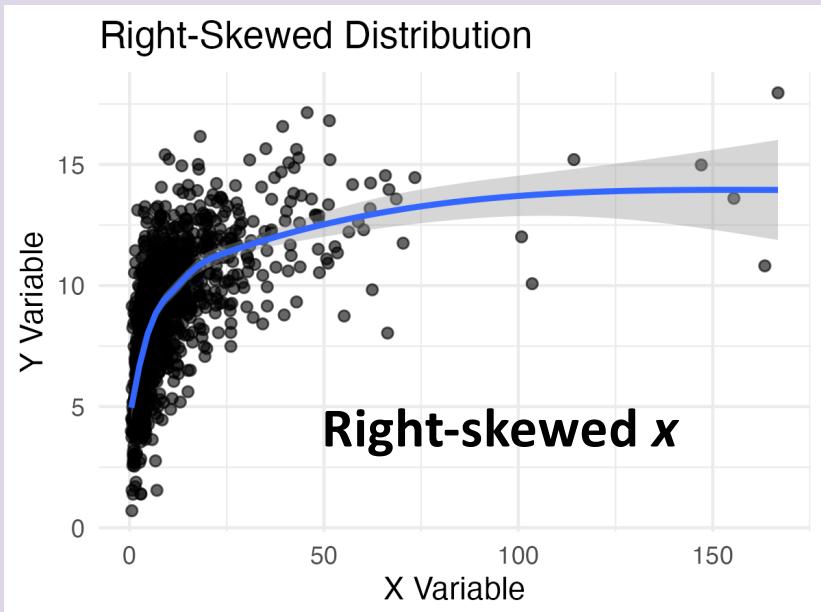


$\log_{10}(x)$





$\log_{10}(x)$



Skills Learning

code-along lecture



Mount Sabyinyo 2025

Skills Application

Laboratory exercise

Week 5 - Skills Application

Download Skills Application Instructions from [Google Drive folder > Week 5](#)

If you are working on your own dataset:

- Keep working on your *Lastname_Firstname_Data.Rmd*

If you are working on sample dataset:

- create a Markdown file just for this week.
- Name it: *Lastname_Firstname_Week5.Rmd* in `/**Week 5**`

In your new Markdown file: Use code blocks (Ctrl+Alt+i) to load packages, import dataset into the environment, and save data as an object (e.g., `*data <-* `). Then follow **instructions**.