R Data Visualization

Wednesday, 9 April 2025
Presented by Satoshi Koiso

Table of contents

Table of contents

R vs. Excel (Google spreadsheet) vs. STATA

Software	Price	Pros	Cons
R	Free	• Versatile • Better graphical tools • Comprehensive documentation • Strong community of users • Large selection of libraries	• Steeper learning curve • Longer scripts compared to STATA • Most work will require installation of libraries • Updates to R and libraries often affect running older scripts
Excel	Free for online or US$99.99/year	• Easiest to use	• Cannot handle large data files • Prone to human error • Hard to conduct **probabilistic sensitivity analysis**
Google spreadsheet	Free		
STATA	US$870+/year (for a single-user)	• Easier to use than R (regression) • More out-of-the-box statistical tools	• Limited data visualization capabilities compared to R • Can only import one data set at a time

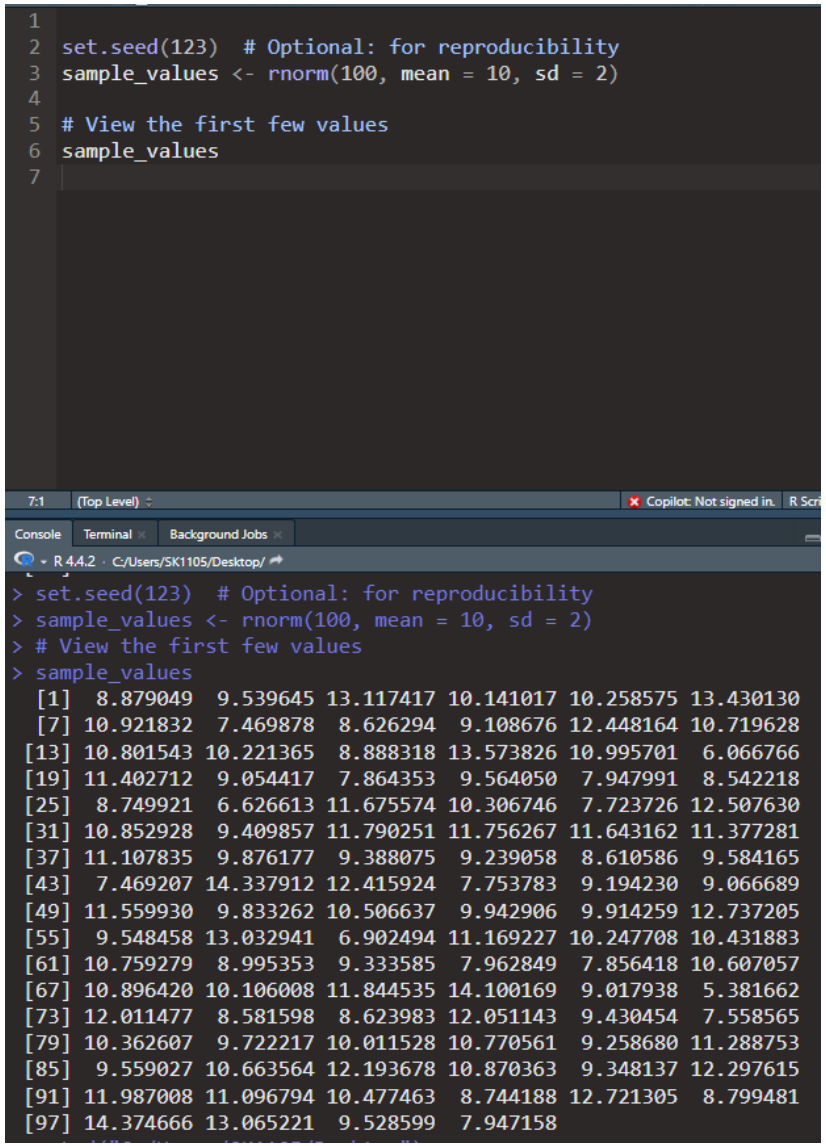
Probabilistic Sensitivity Analysis

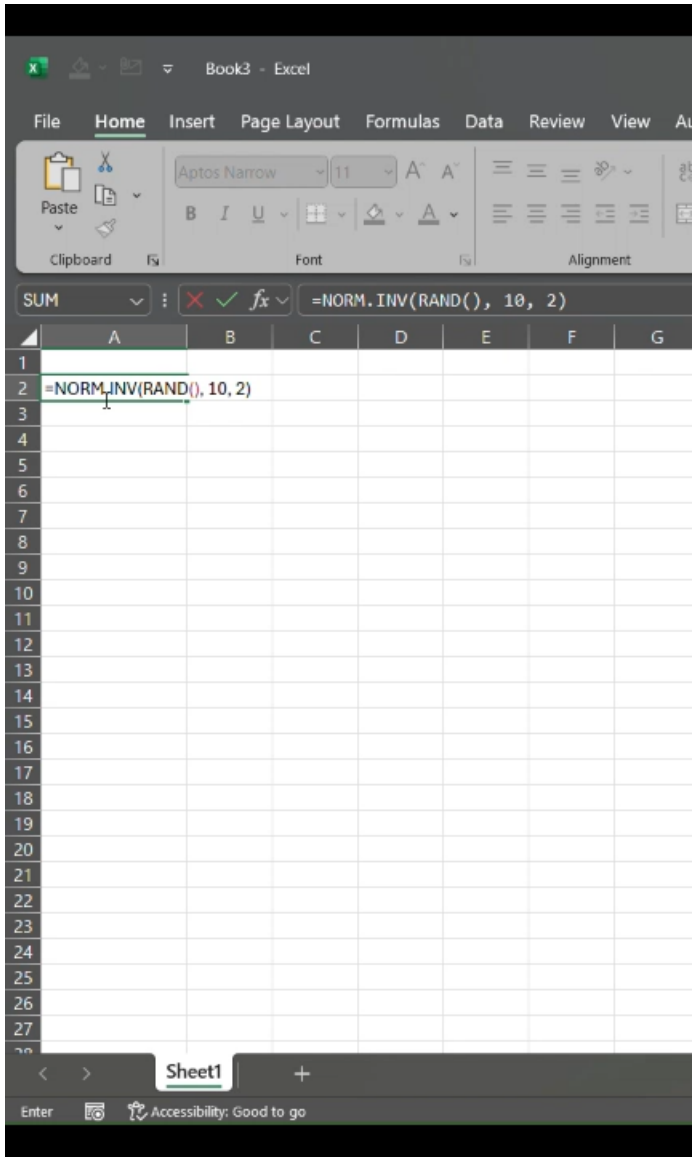
- In PSA, we sample numbers from a distribution to account for uncertainty.

- Example: Sampling 100 numbers from a normal distribution with a mean of 10 and SD of 2.

- R
  - Can reproduce the same result by setting a seed.
  - Just three lines of code.

- Excel/Google spreadsheet
  - Results are unstable.
  - Much more steps.

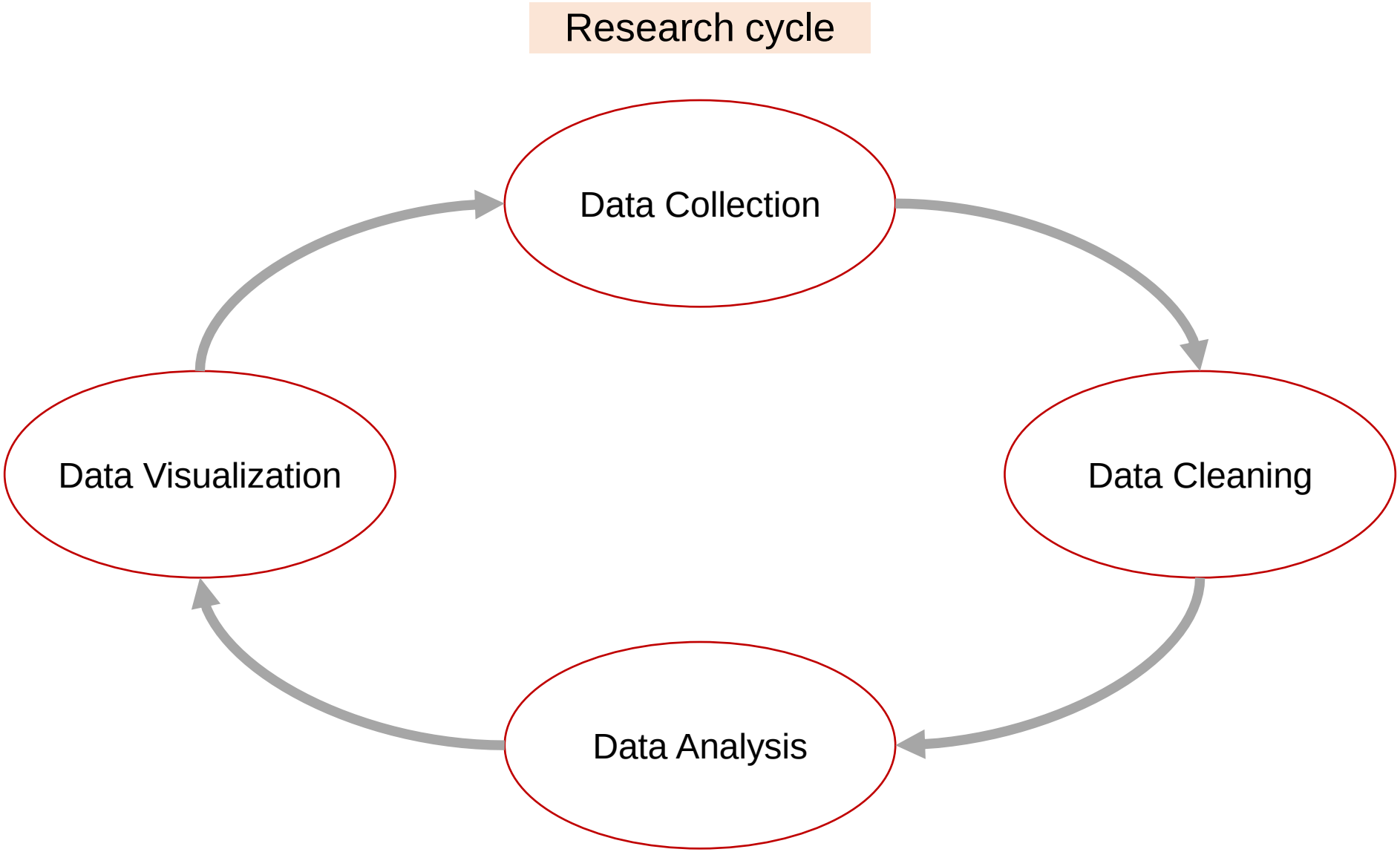




What should R be used for?

- Data analysis and figure creation for one specific project
- Sensitivity analyses involving many runs
- Integrating data with online
- Geospatial analyses

Scope of this session



Scope of this session

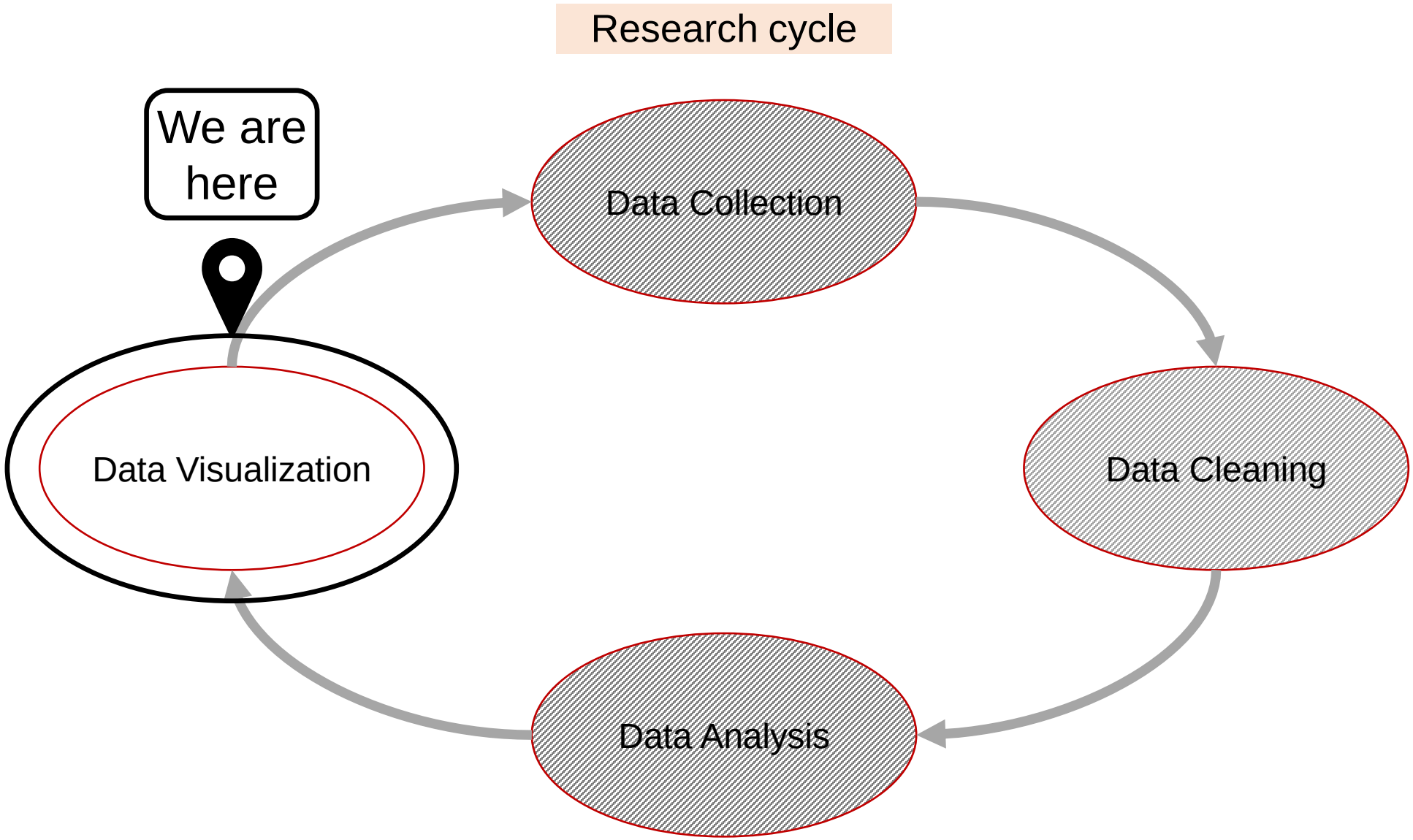


Table of contents

Histogram

Use case:

- Show distributions of data.
- The height of a bar represents frequency of occurrence of values within each bin.

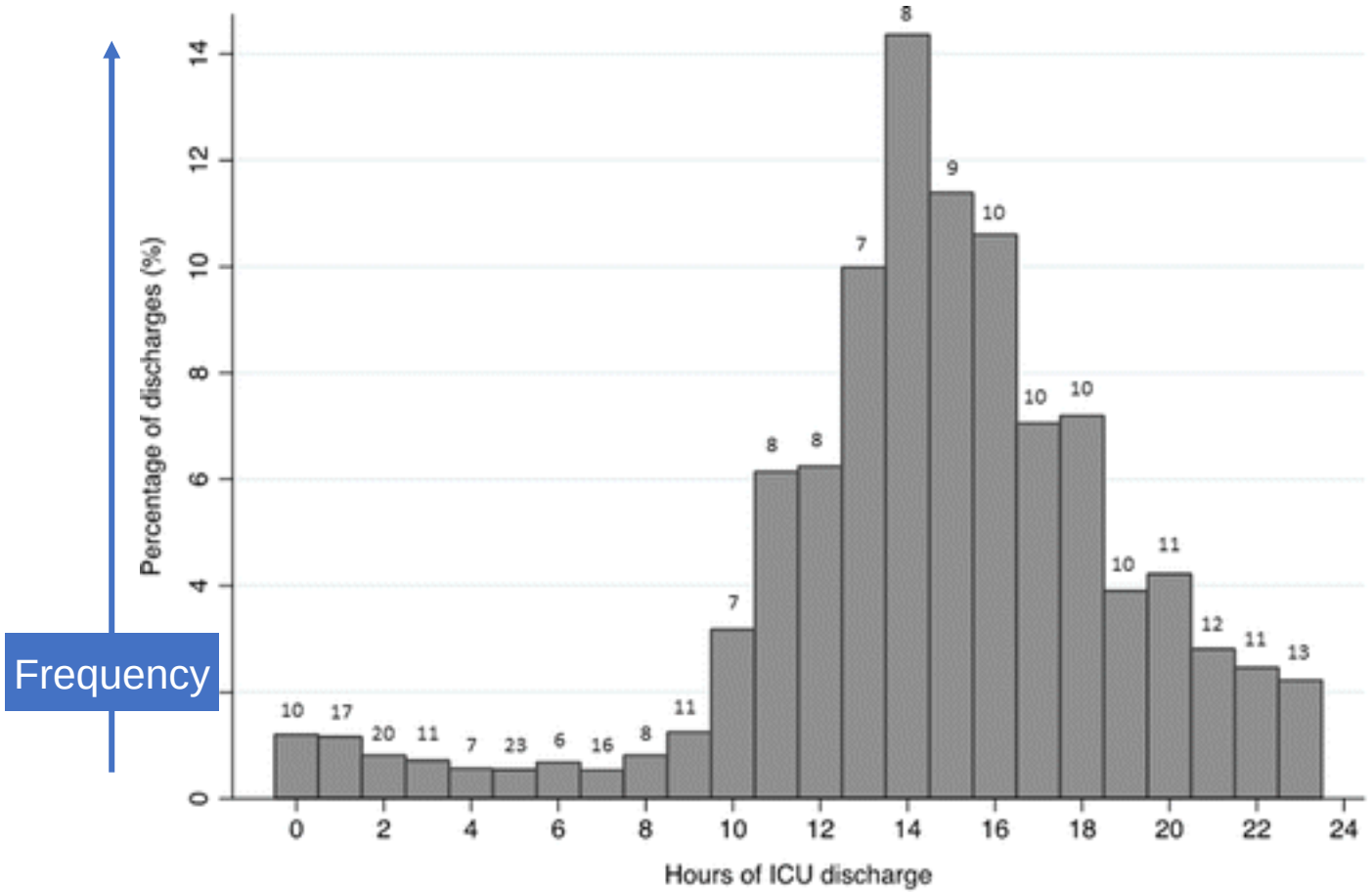


Figure1. Histogram showing the distribution of hours of ICU discharge for the study population as well as the unadjusted mortality rates distributed according to hour of discharge
From Azevedo, L.C., de Souza, I.A., Zygun, D.A. et al. Association Between Nighttime Discharge from the Intensive Care Unit and Hospital Mortality: A Multi-Center Retrospective Cohort Study. BMC Health Serv Res 15, 378 (2015). https://doi.org/10.1186/s12913-015-1044-4

Bar graph

Use case:

- Compare data across different groups.

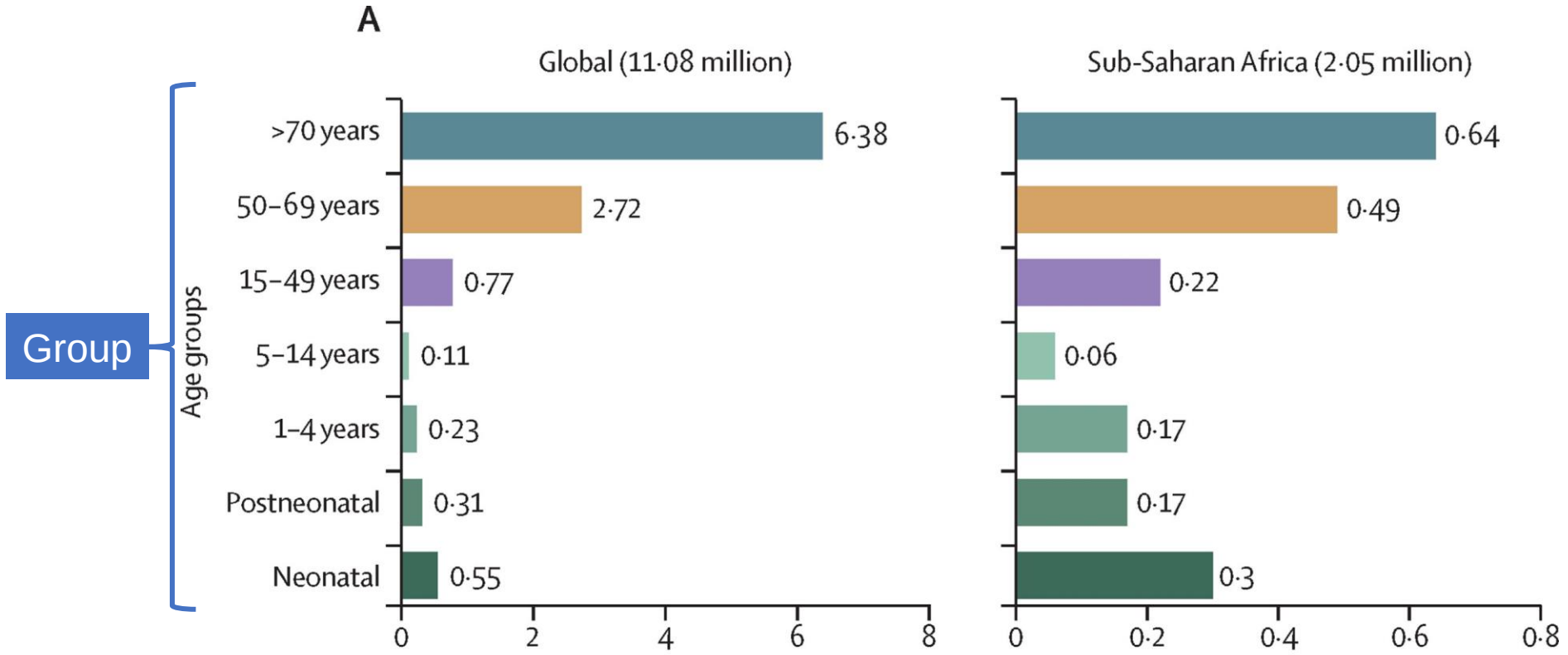


Figure 8. Cumulative deaths averted (in millions) in two different scenarios, by age, global versus sub-Saharan Africa, 2025–2050 (A) Gram-negative drug scenario.
From GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. Lancet. 2024;404(10459):1199-1226. doi:10.1016/S0140-6736(24)01867-1

Stacked bar graph

Use case:

- Compare data across subgroups in different groups.
- e.g. Across a group of strategies, showing which subgroup of costs occupies how much of the total costs.

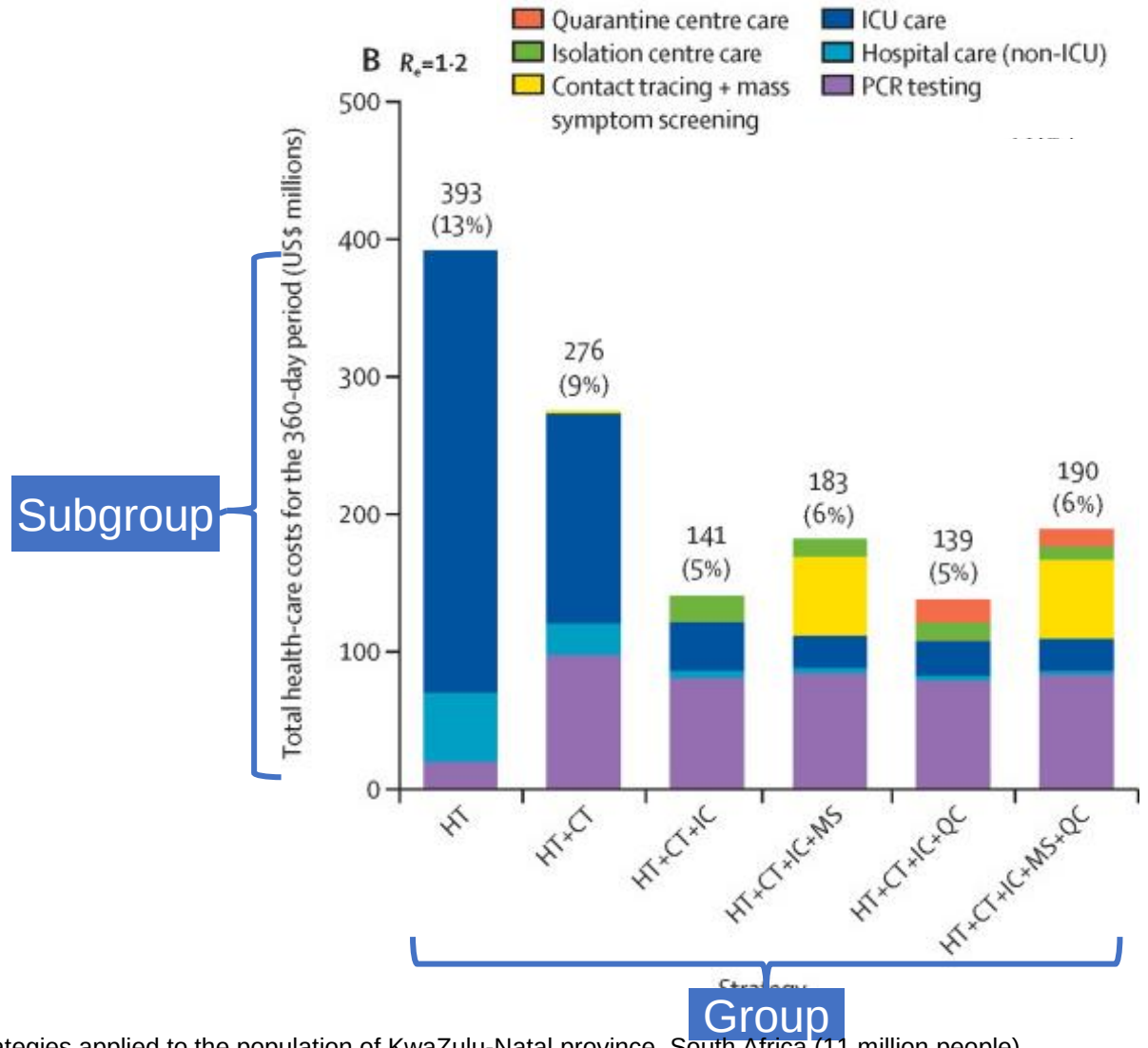


Figure 2. Budget impact analysis of contributors to health-care costs of COVID-19 intervention strategies applied to the population of KwaZulu-Natal province, South Africa (11 million people)
From Reddy KP, Shebl FM, Foote JHA, et al. Cost-effectiveness of public health strategies for COVID-19 epidemic control in South Africa: a microsimulation modelling study. Lancet Glob Health. 2021;9(2):e120-e129. doi:10.1016/S2214-109X(20)30452-6

Scatter plot

Use case:

- Explore a relationship between two variables.
- Visualize trends in data and shape of data.

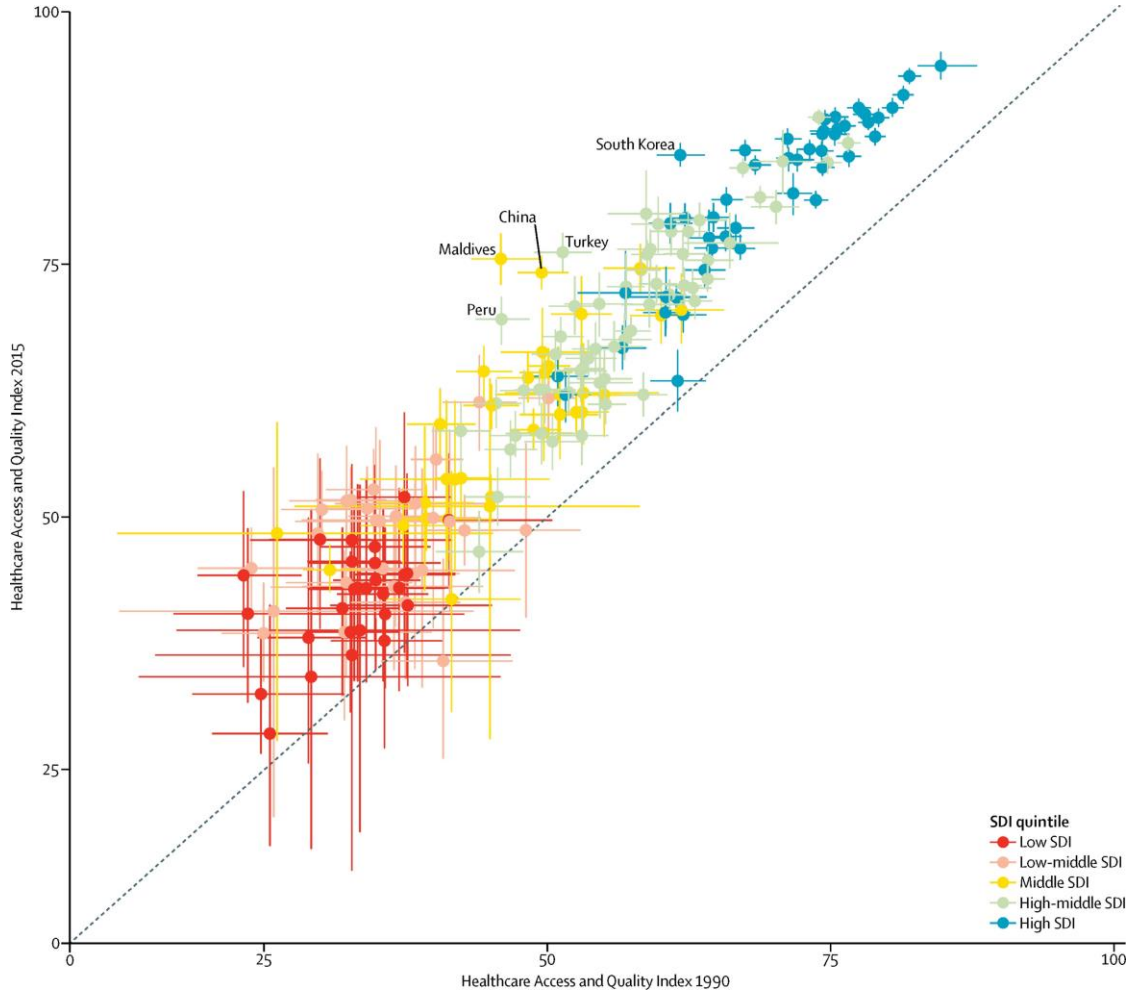


Figure 3 Comparison of 1990 and 2015 HAQ Index estimates, with uncertainty, by country or territory
From Healthcare Access and Quality Index based on mortality from causes amenable to personal health care in 195 countries and territories, 1990–2015: a novel analysis from the Global Burden of Disease Study 2015 Barber, Ryan M et al. The Lancet, Volume 390, Issue 10091, 231 - 266

Cost-effectiveness scatter plot

Use case:

- Show how parameter uncertainty impacts cost-effectiveness.
- Each dot shows the result from each recalculation in a probabilistic sensitivity analysis.

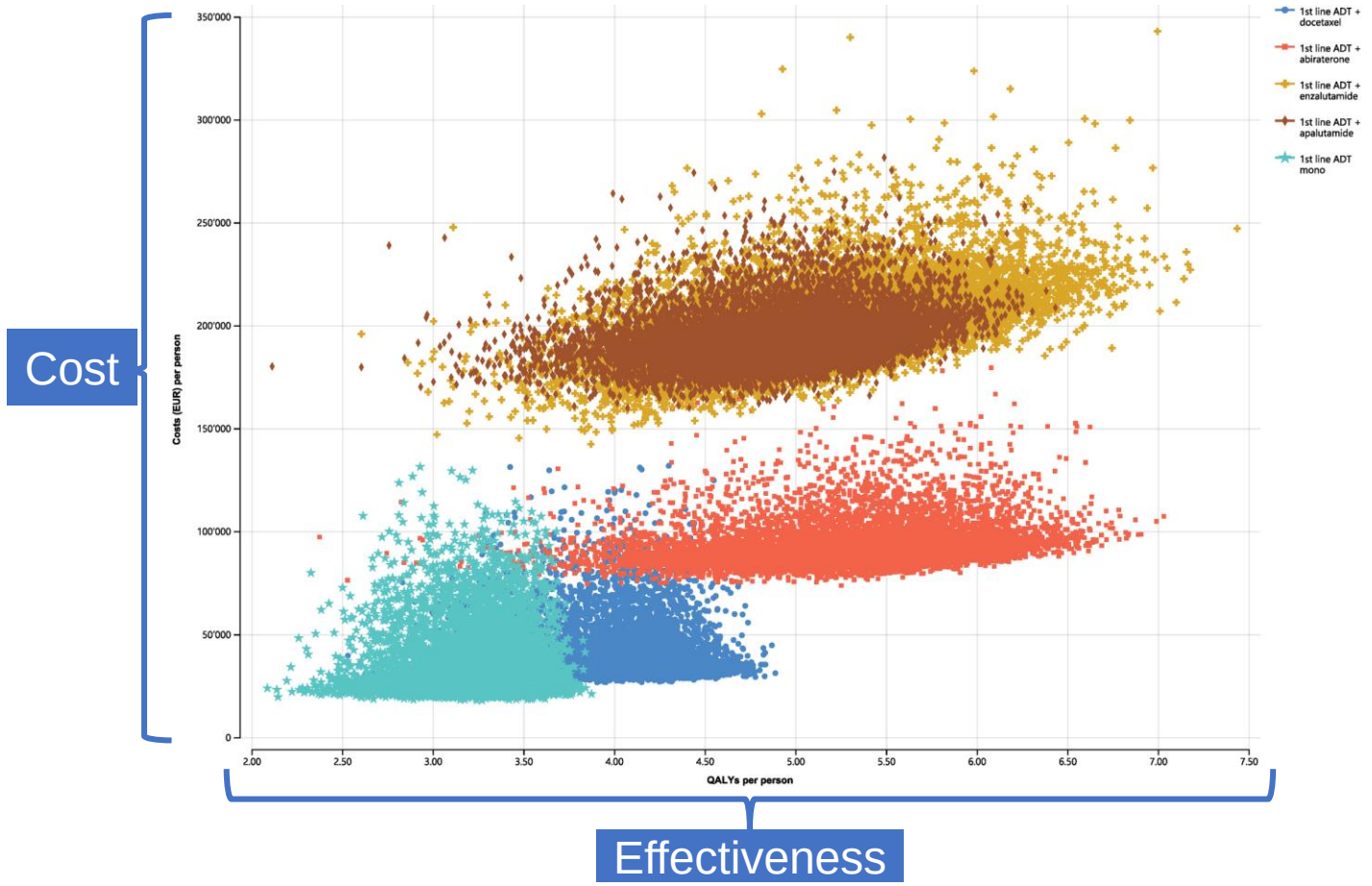
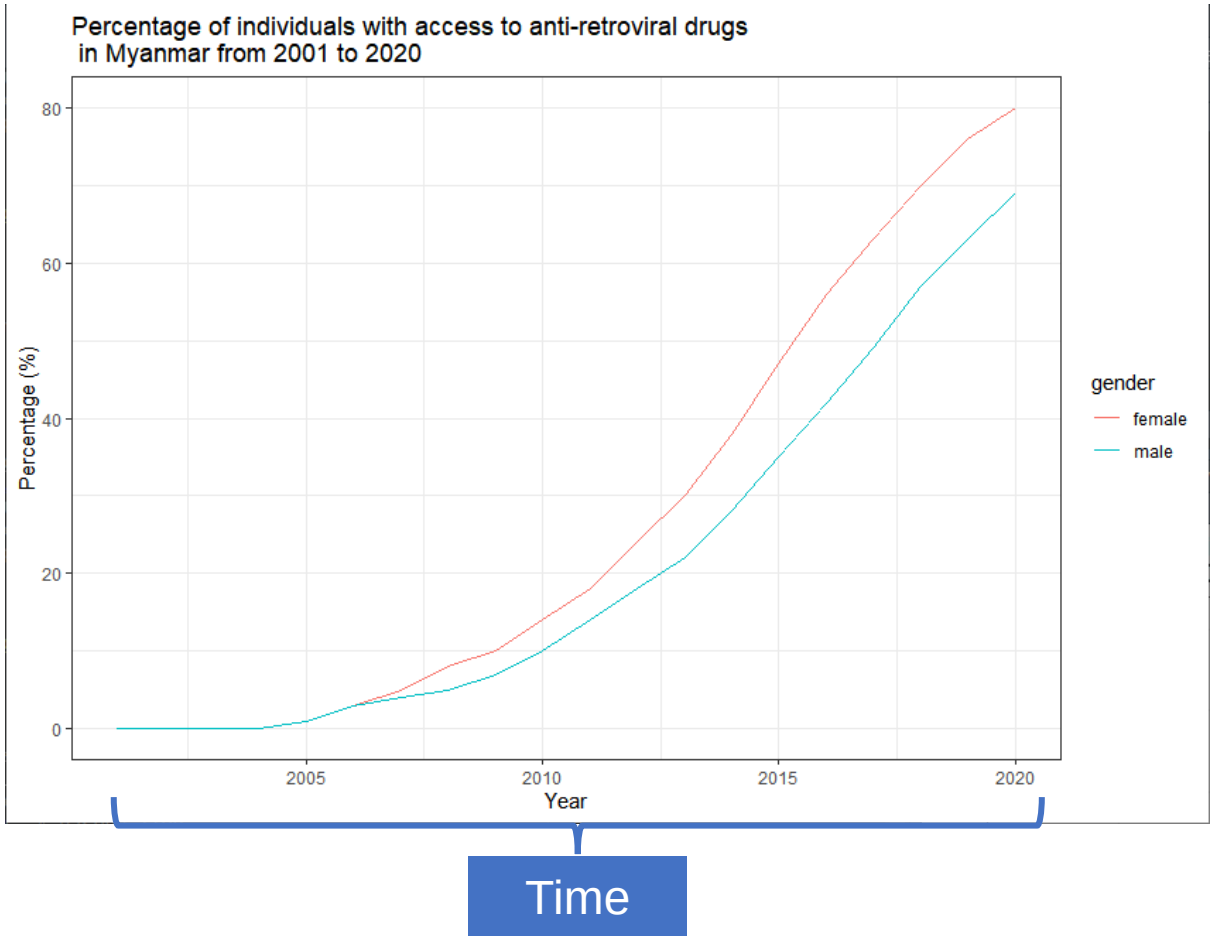


Fig 3. Cost-effectiveness scatter plot based on probabilistic sensitivity analysis.
From Barbier MC, Tomonaga Y, Menges D, Yebyo HG, Haile SR, Puhan MA, et al. (2022) Survival modelling and cost-effectiveness analysis of treatments for newly diagnosed metastatic hormone-sensitive prostate cancer. PLoS ONE 17(11): e0277282. https://doi.org/10.1371/journal.pone.0277282

Line graph

Use case:

- Visualize values of variables with a gradual change of a continuous variable (often time).



Survival curve

Use case:

- Plot model output of proportion of people alive over time.
- Compare several intervention arms.

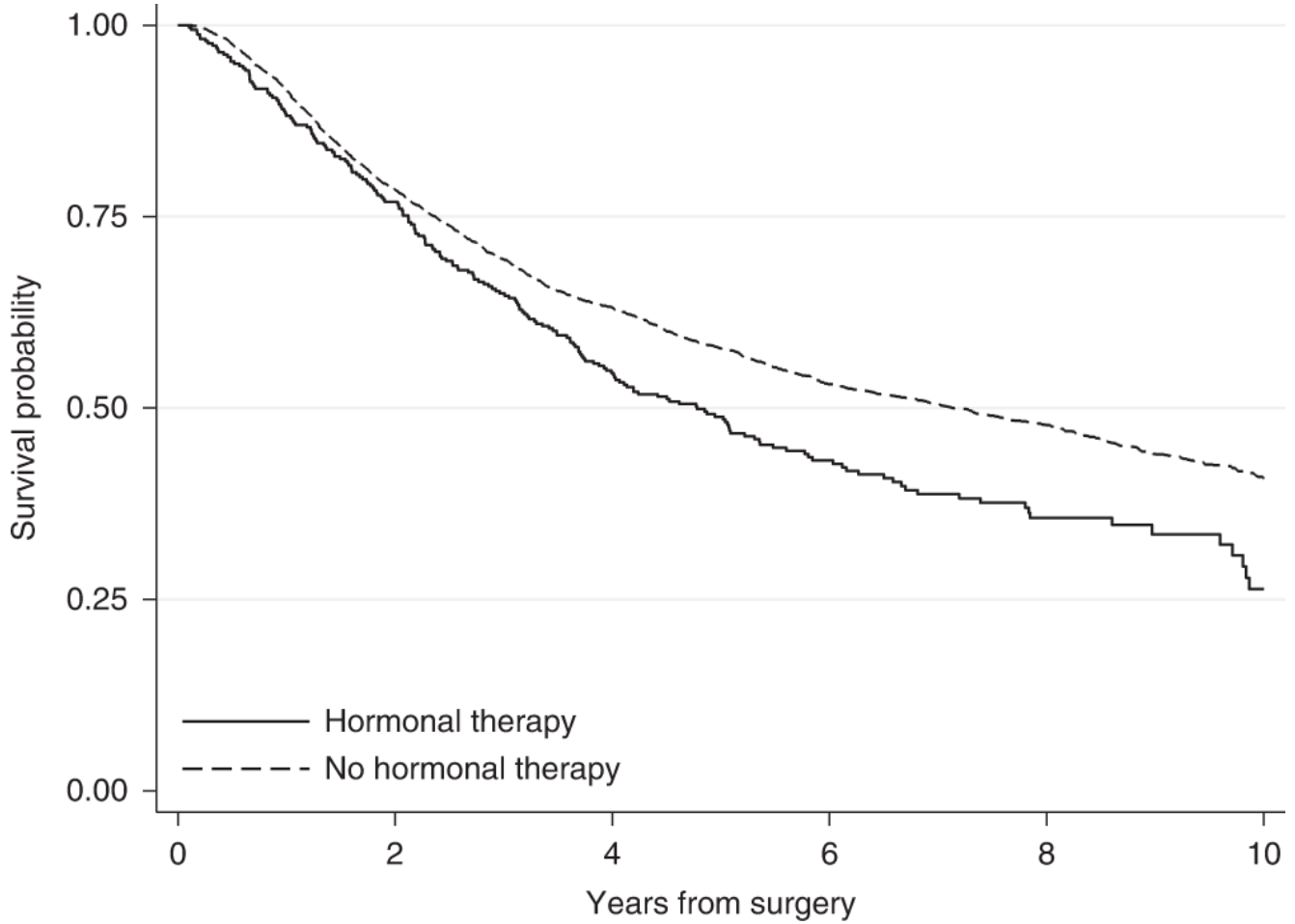


Fig. 1: Kaplan–Meier survival curves.
From Syriopoulou, E., Wästerlid, T., Lambert, P.C. et al. Standardised survival probabilities: a useful and informative tool for reporting regression models for survival data. Br J Cancer 127, 1808–1815 (2022). https://doi.org/10.1038/s41416-022-01949-6

Cost-Effectiveness frontier

Use case:

- Compare competing strategies' cost-effectiveness.
- Line connects successive points representing cost-effective strategies at different values of the cost-effectiveness threshold.
- Points not on the frontier are not considered cost-effective at any threshold.

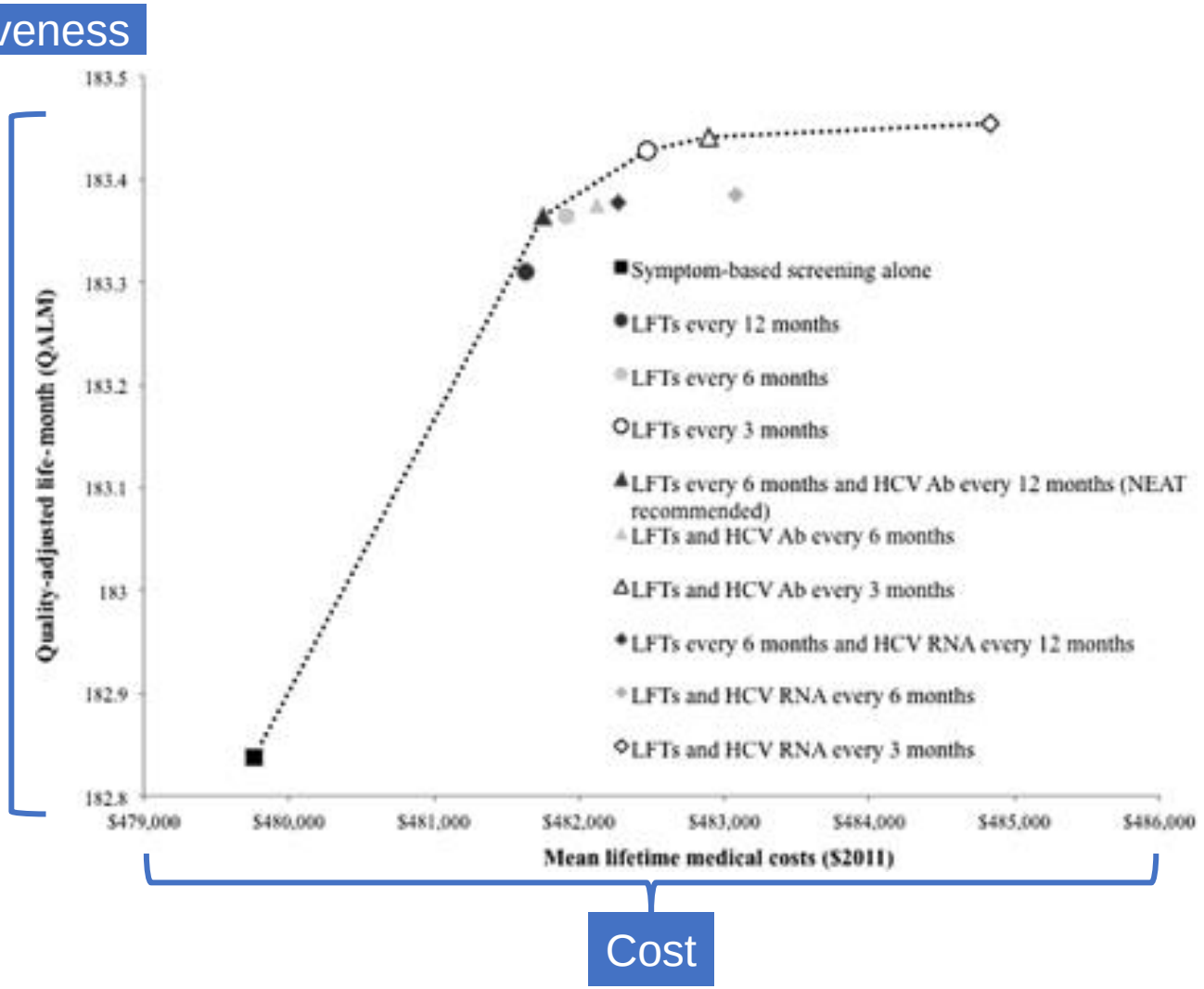


Figure 1.Efficiency frontier of a cost-effectiveness analysis of screening for acute hepatitis C virus (HCV) infection in human immunodeficiency virus (HIV)–infected men who have sex with men.
From Benjamin P. Linas, Angela Y. Wong, Bruce R. Schackman, Arthur Y. Kim, Kenneth A. Freedberg, Cost-effective Screening for Acute Hepatitis C Virus Infection in HIV-Infected Men Who Have Sex With Men, Clinical Infectious Diseases, Volume 55, Issue 2, 15 July 2012, Pages 279–290, https://doi.org/10.1093/cid/cis382

How to interpret Cost-Effectiveness frontier

Use case:

- The slope between two points represents the incremental cost-effectiveness ratio (ICER).
- In the right graph, ICERs are
  - A: $43,700/QALY
  - B: $129,700/QALY
- Depending on the decision-maker's preference, the optimal strategies would be different.

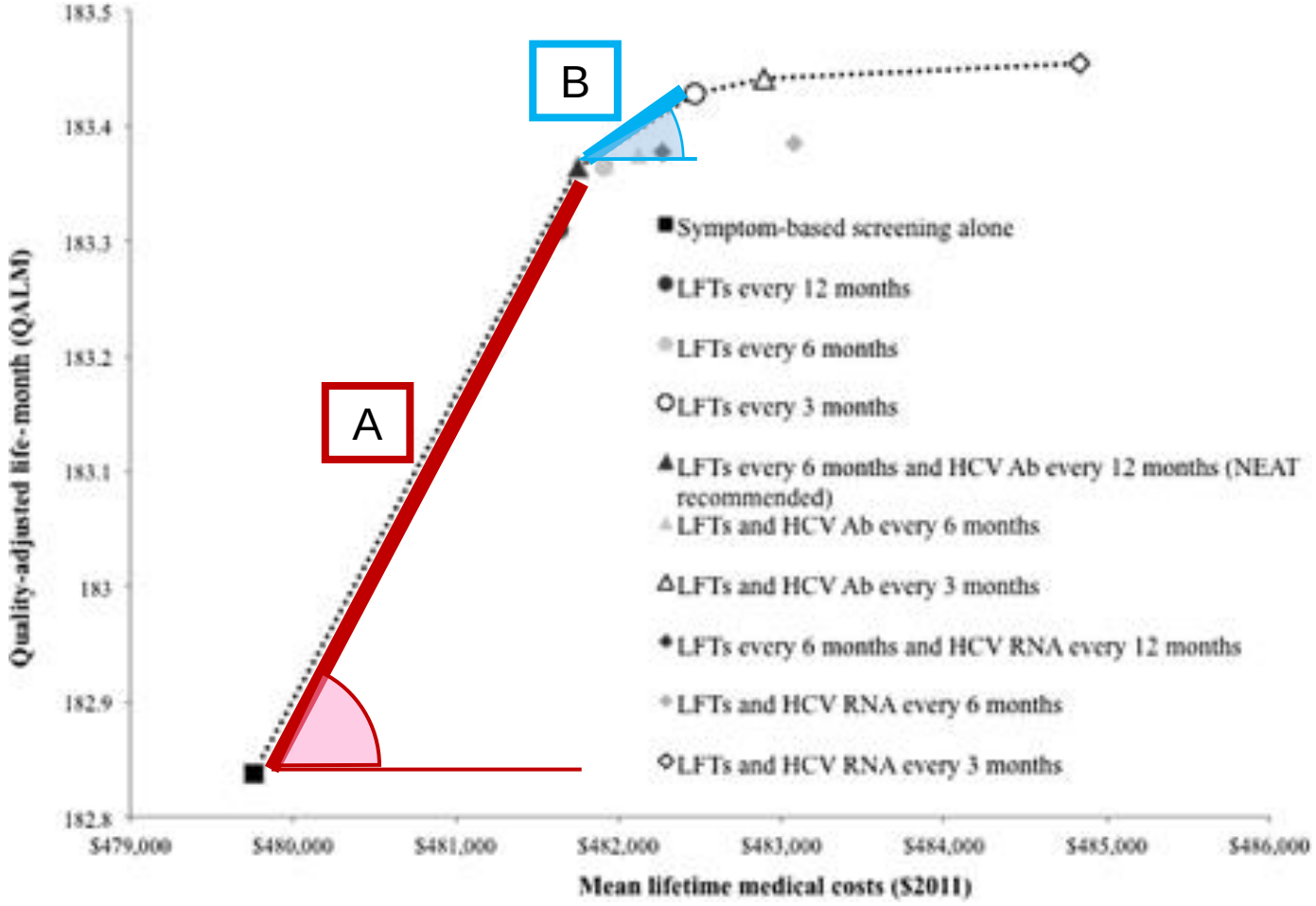


Figure 1.Efficiency frontier of a cost-effectiveness analysis of screening for acute hepatitis C virus (HCV) infection in human immunodeficiency virus (HIV)–infected men who have sex with men.
From Benjamin P. Linas, Angela Y. Wong, Bruce R. Schackman, Arthur Y. Kim, Kenneth A. Freedberg, Cost-effective Screening for Acute Hepatitis C Virus Infection in HIV-Infected Men Who Have Sex With Men, Clinical Infectious Diseases, Volume 55, Issue 2, 15 July 2012, Pages 279–290, https://doi.org/10.1093/cid/cis382

Cost-effectiveness acceptability curve

Use case:

- Shows the percentage of calculations that favor each strategy over a Willingness-to-pay range.
- As WTP is higher, more effective strategies are more likely to be chosen.

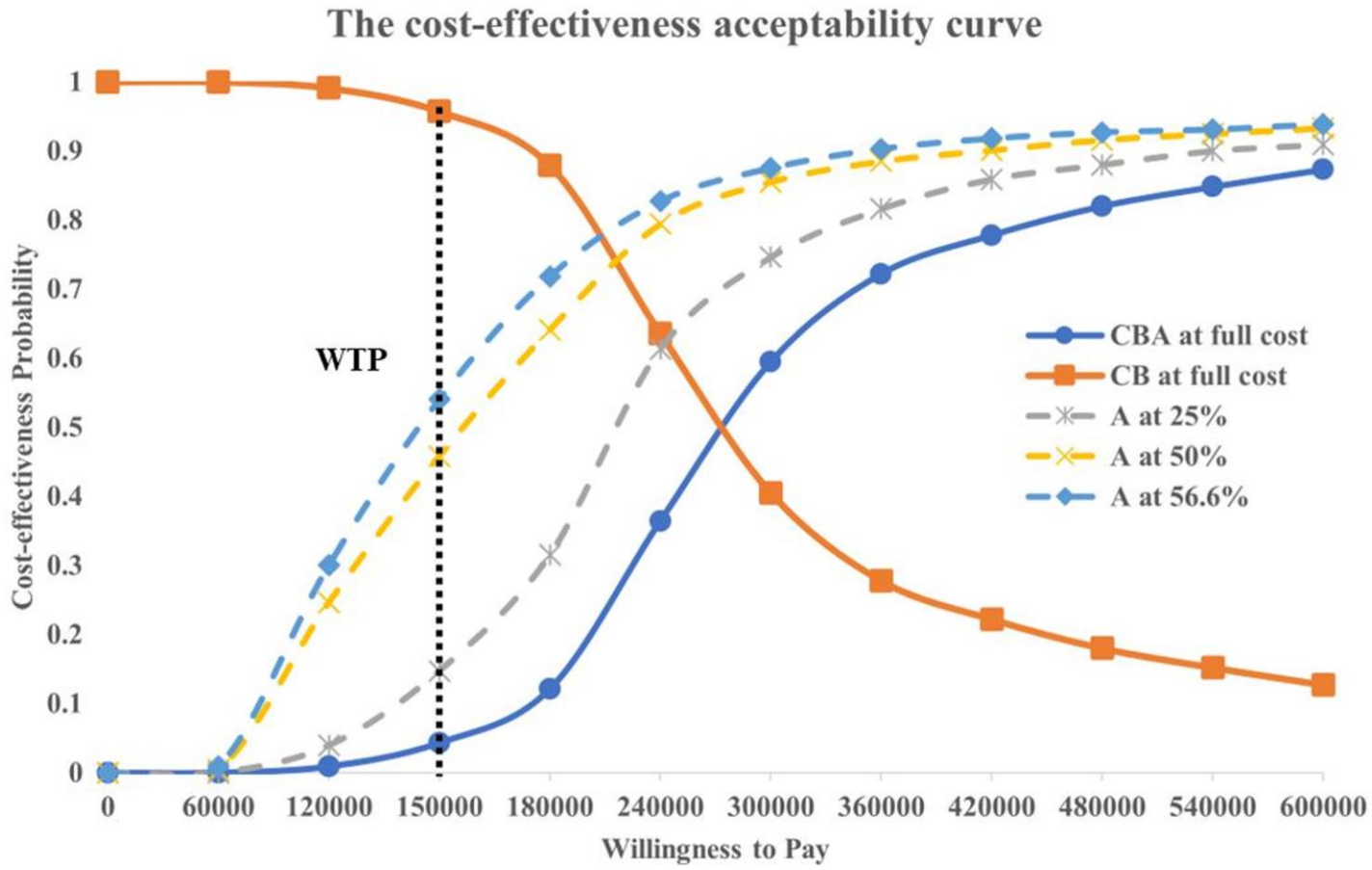


Figure 3. The cost-effectiveness acceptability curve of different atezolizumab cost.
From Lei, Jianying, MSc, et al. “First-Line Treatment With Atezolizumab Plus Bevacizumab and Chemotherapy for US Patients With Metastatic, Persistent, or Recurrent Cervical Cancer: A Cost-Effectiveness Analysis.” Value in Health, vol. 27, no. 11, 2024, pp. 1528–34, https://doi.org/10.1016/j.jval.2024.07.013.

What is (deterministic) sensitivity analysis?

- SA evaluates the effects of input uncertainty on the conclusion.

Input parameters	Values
HIV Testing cost	$5
HIV Prevalence	1%
HIV Testing sensitivity	80%
PreP cost	$20
ART cost	$50

Conclusion:
Testing every three years is cost-effective than testing every year.

If the input values are not certain, how would this conclusion be different?

One-way SA:

- Changing a single parameter value one by one.
- E.g.
  1) Testing cost: $1
  2) Testing cost: $10
  3) Prevalence: 0.1%
  4) Prevalence: 10% …

Two-way (Multi-way) SA:

- Changing two parameter values together.
- E.g.
  1) Testing cost: $1 & Prevalence: 0.1%
  2) Testing cost: $1 & Prevalence: 10%
  3) Testing cost: $10 & Prevalence: 0.1%
  4) Testing cost: $10 & Prevalence: 10%

  ….

Tornado graph (used in one-way sensitivity analyses)

Use case:

- Show the effect on a specific model output of changing inputs to the maximum or minimum value in a range.

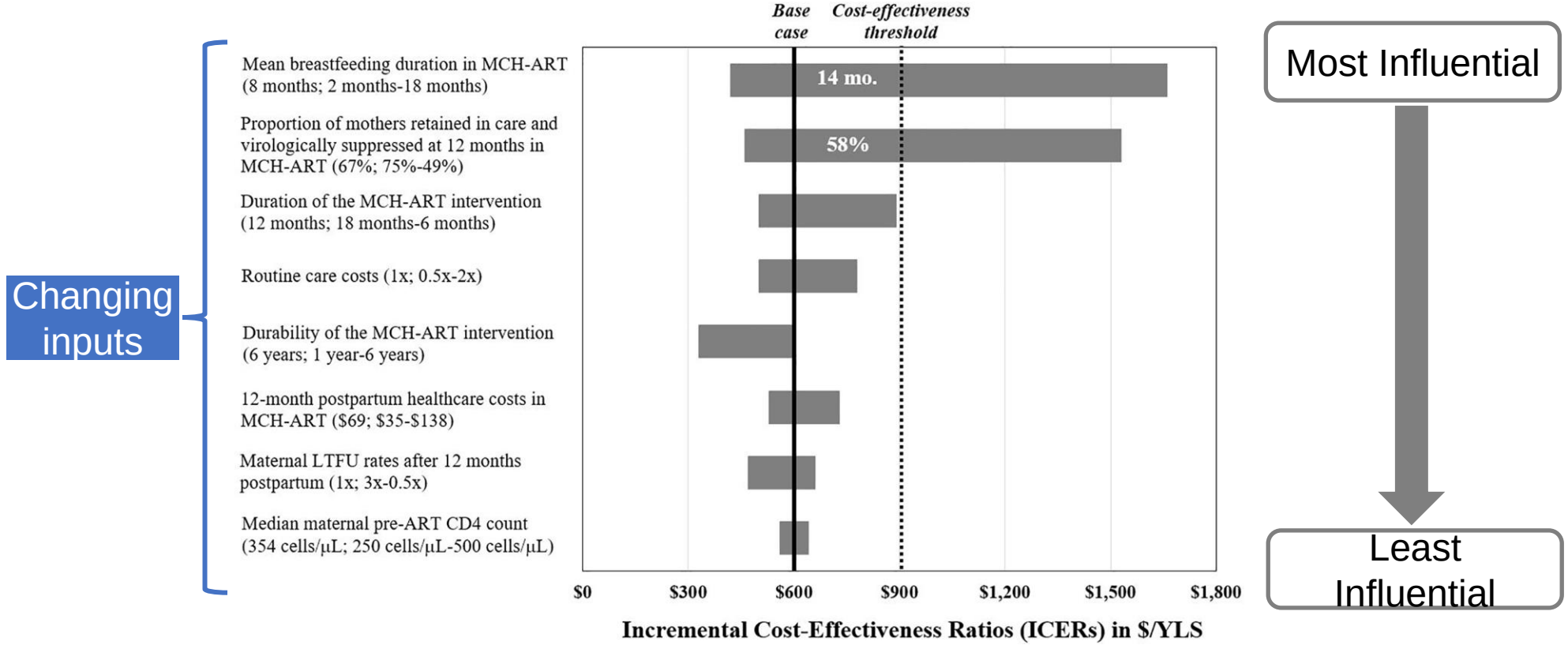


Fig 2. Tornado diagram of one-way sensitivity analyses.
From Dugdale CM, Phillips TK, Myer L, Hyle EP, Brittain K, Freedberg KA, et al. (2019) Cost-effectiveness of integrating postpartum antiretroviral therapy and infant care into maternal & child health services in South Africa. PLoS ONE 14(11): e0225104. https://doi.org/10.1371/journal.pone.0225104

Heatmap (used in two-way sensitivity analyses)

Use case:

- Show which decision to choose as a function of two input parameters in two-way sensitivity analysis.
- The color of the cell shows the optimal decision conditioning on the selected values of the two input parameters.
- Typically compare the efficiency and cost of different strategies.

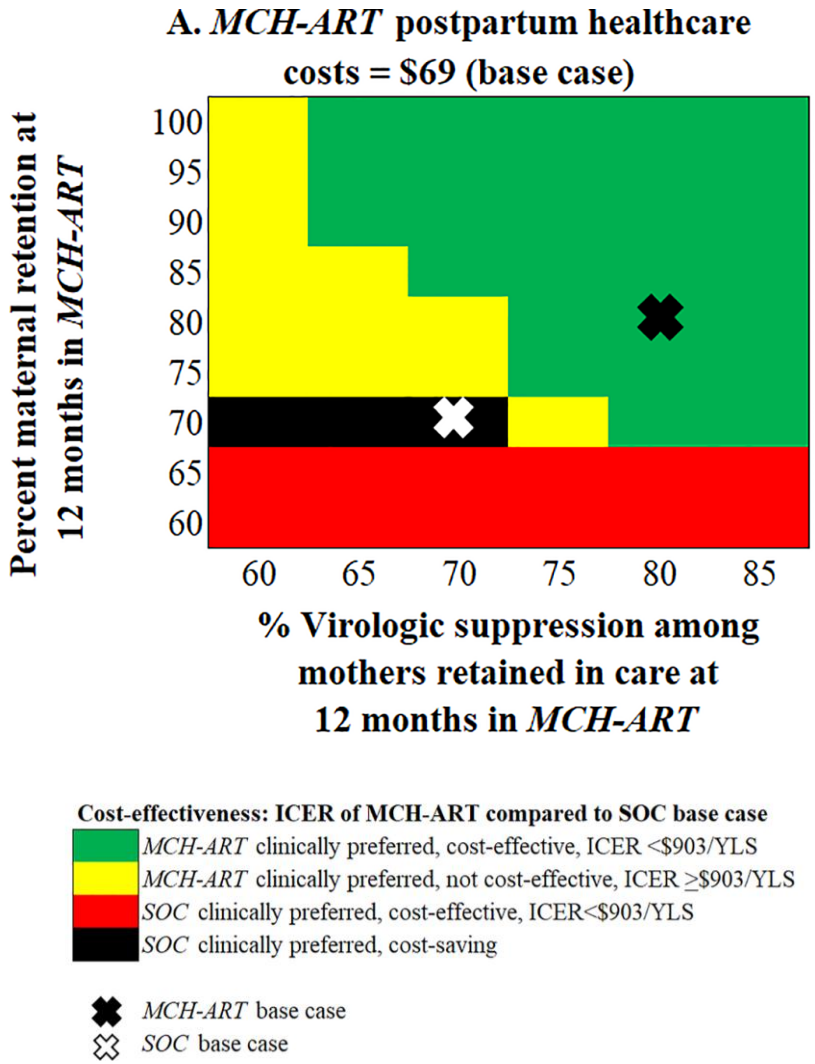


Fig 3. Multi-way sensitivity analyses.
From Dugdale CM, Phillips TK, Myer L, Hyle EP, Brittain K, Freedberg KA, et al. (2019) Cost-effectiveness of integrating postpartum antiretroviral therapy and infant care into maternal & child health services in South Africa. PLoS ONE 14(11): e0225104. https://doi.org/10.1371/journal.pone.0225104

Forest plot

Use case:

- Compare results in meta-analyses and systematic reviews.
- Display results (often odds/hazard ratios and relative risk) with the lower and upper bounds of each study.

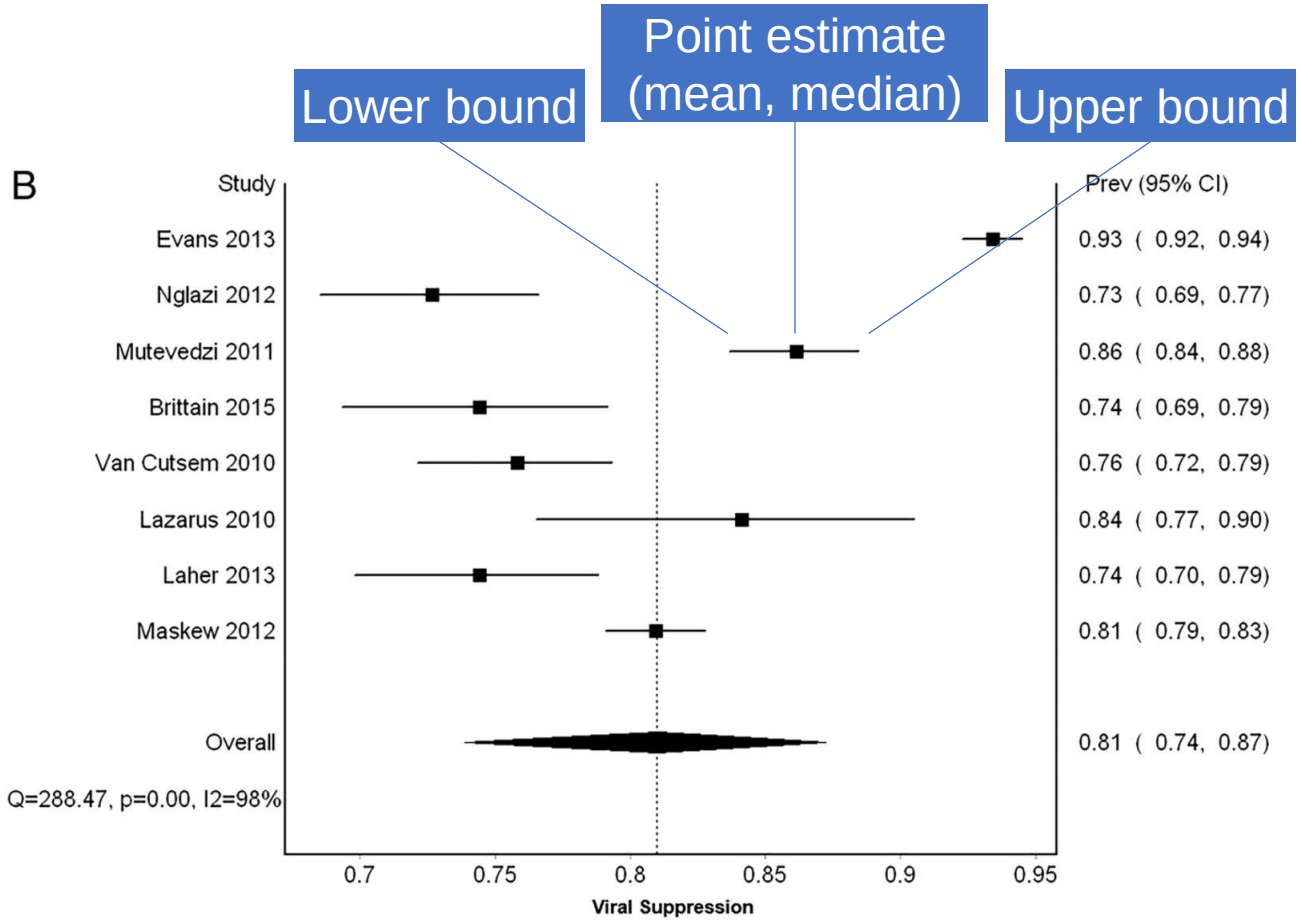
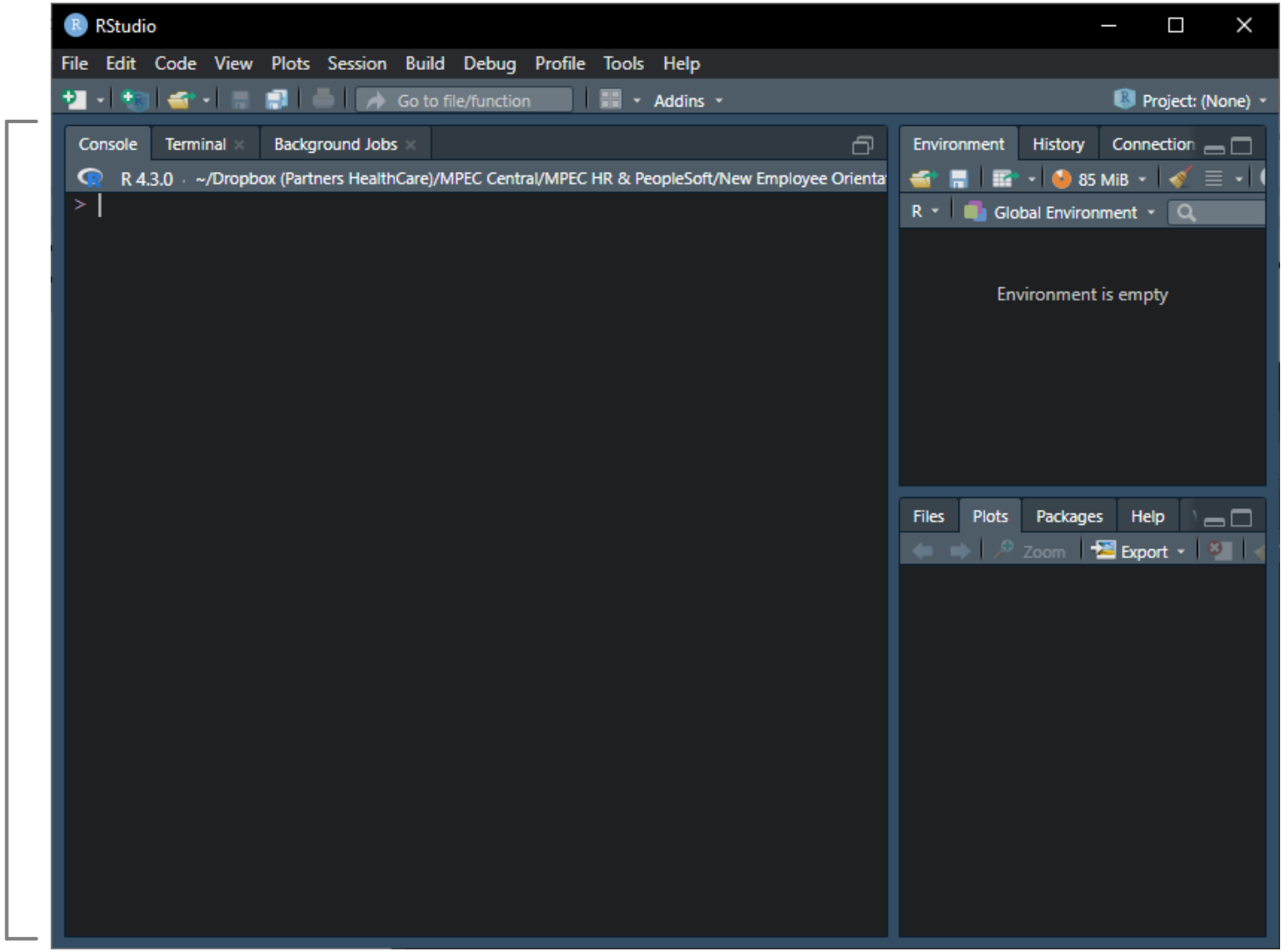


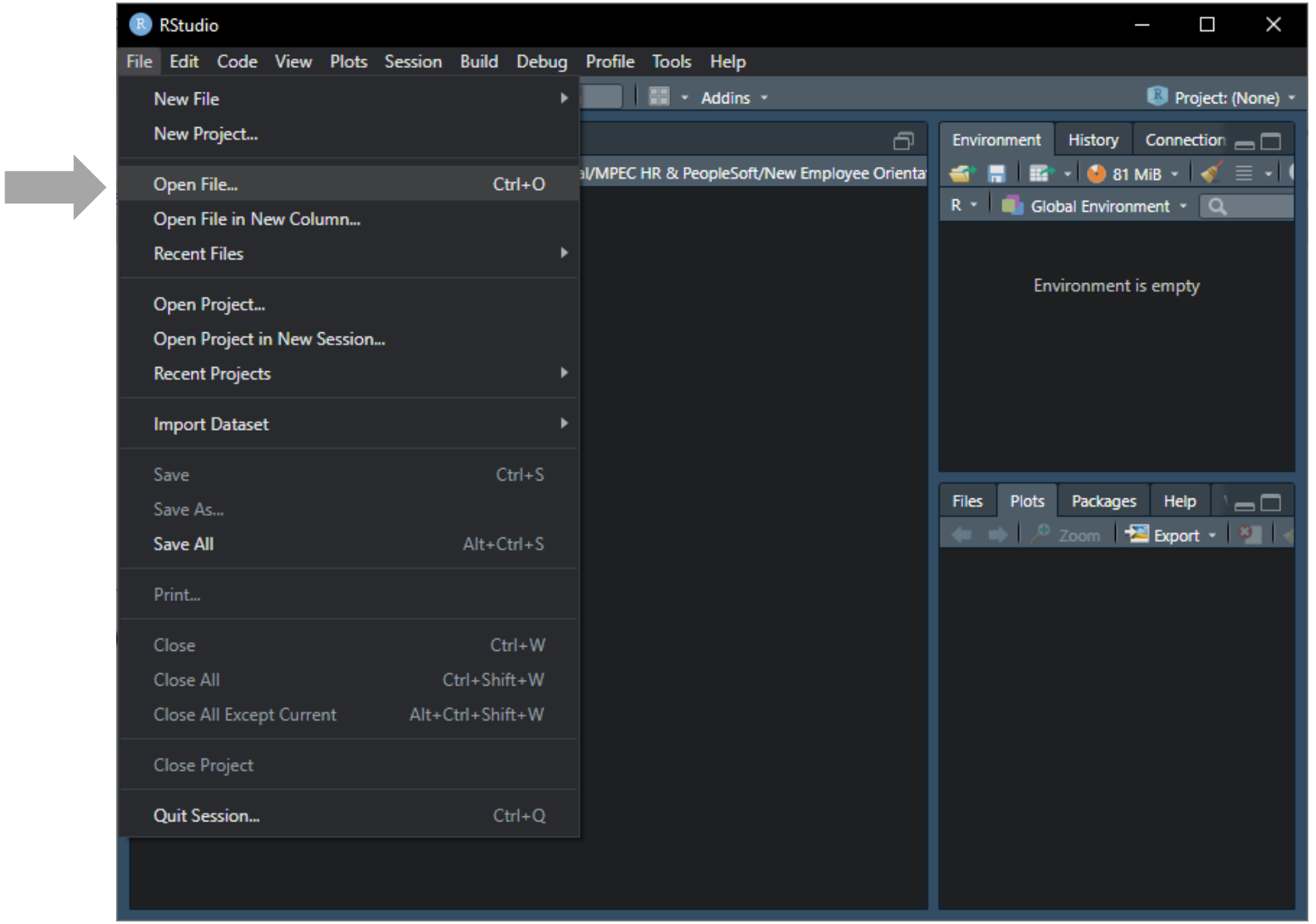
Figure 3 (B) Meta-analysis viral suppression: Forest plot of the proportion of virally suppressed HIV-infected adolescents and young adults in South Africa.
From Brian C Zanoni, Moherndran Archary, Sarah Buchan, Ingrid T Katz, Jessica E Haberer - Systematic review and meta-analysis of the adolescent HIV continuum of care in South Africa: the Cresting Wave: BMJ Global Health 2016;1:e000004.

Table of contents

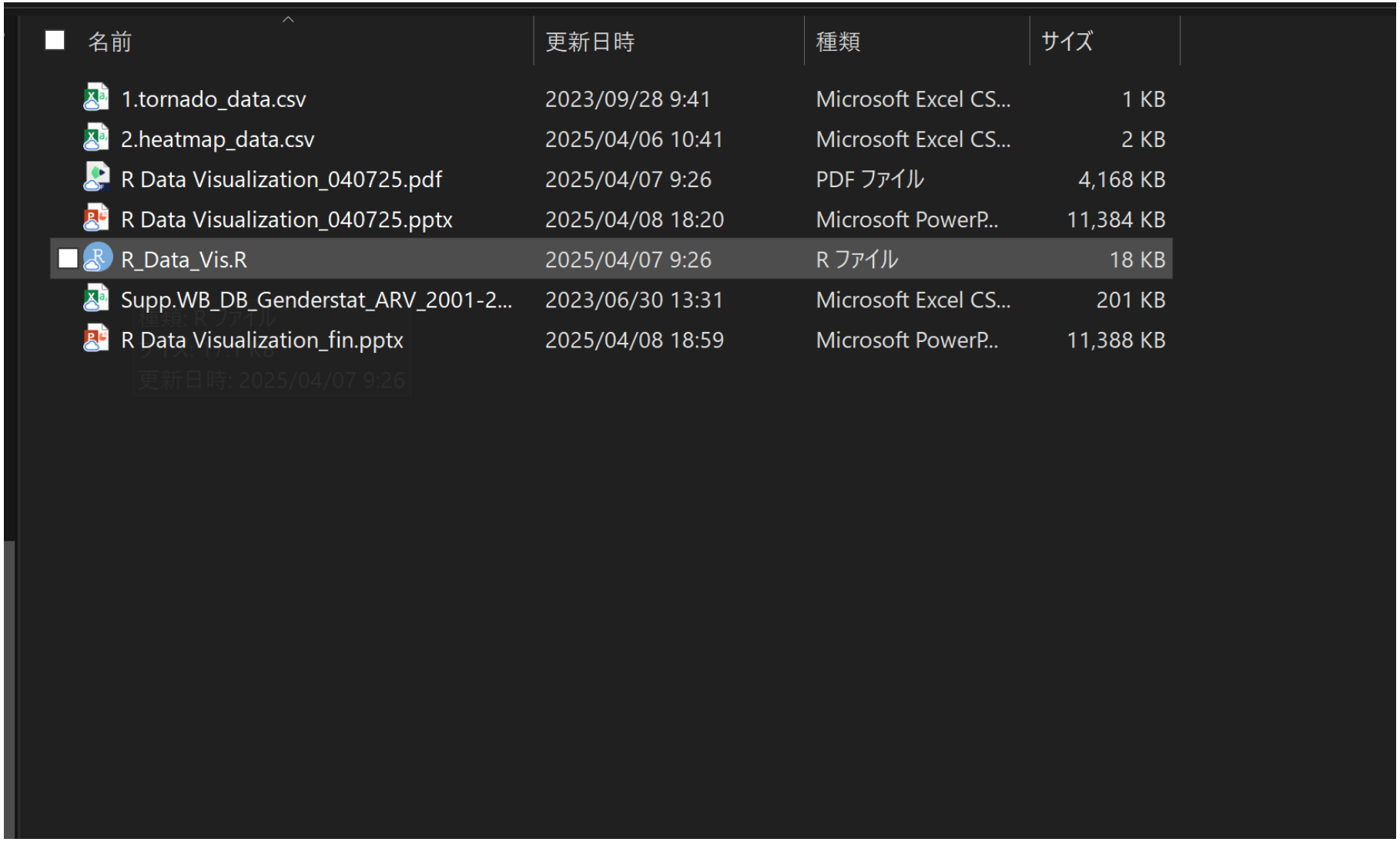
Rstudio: Integrated Development Environment



Opening Script Step 1



Opening Script Step 2



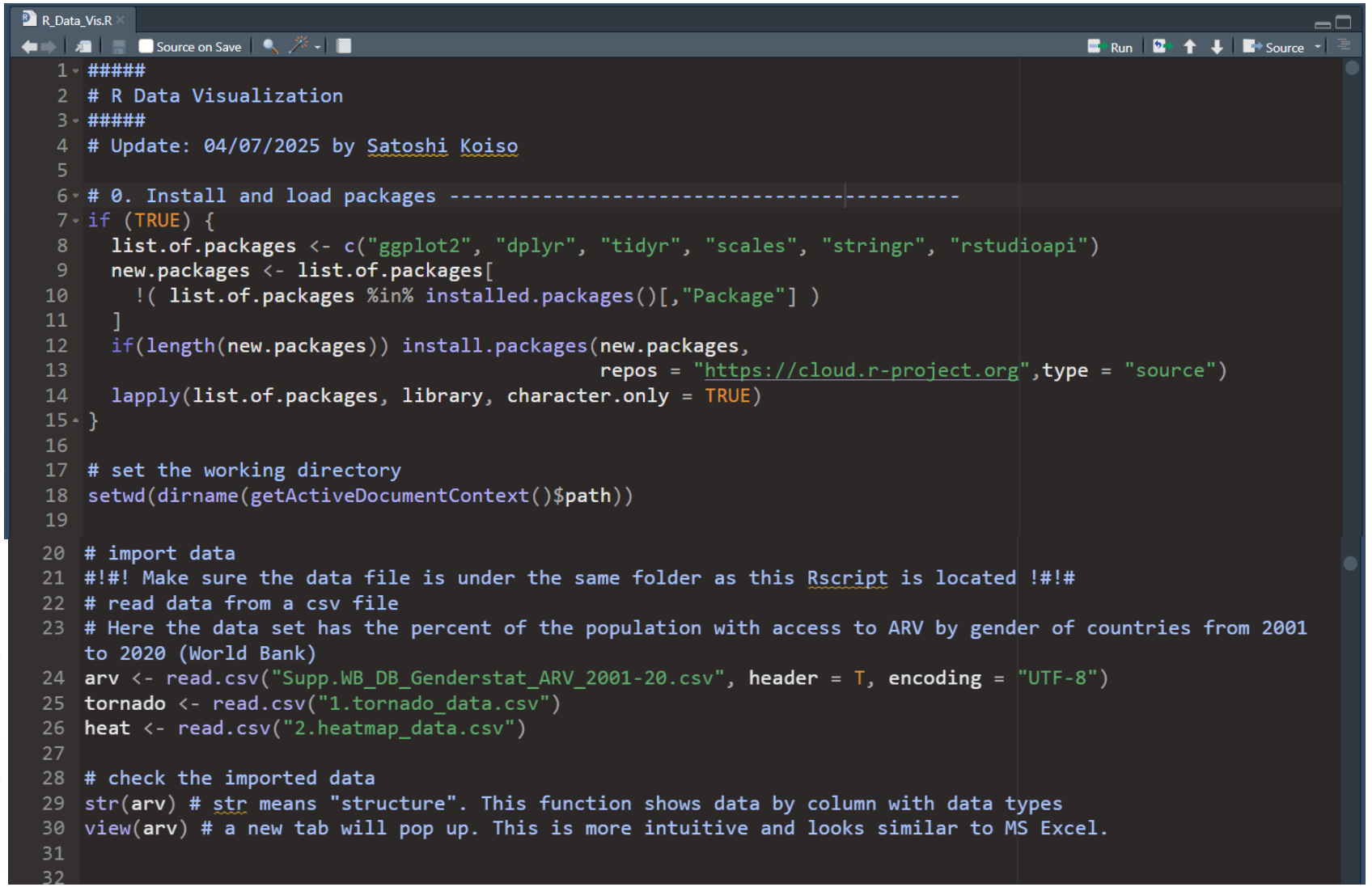
Script components

Load packages (packages include useful R functions)

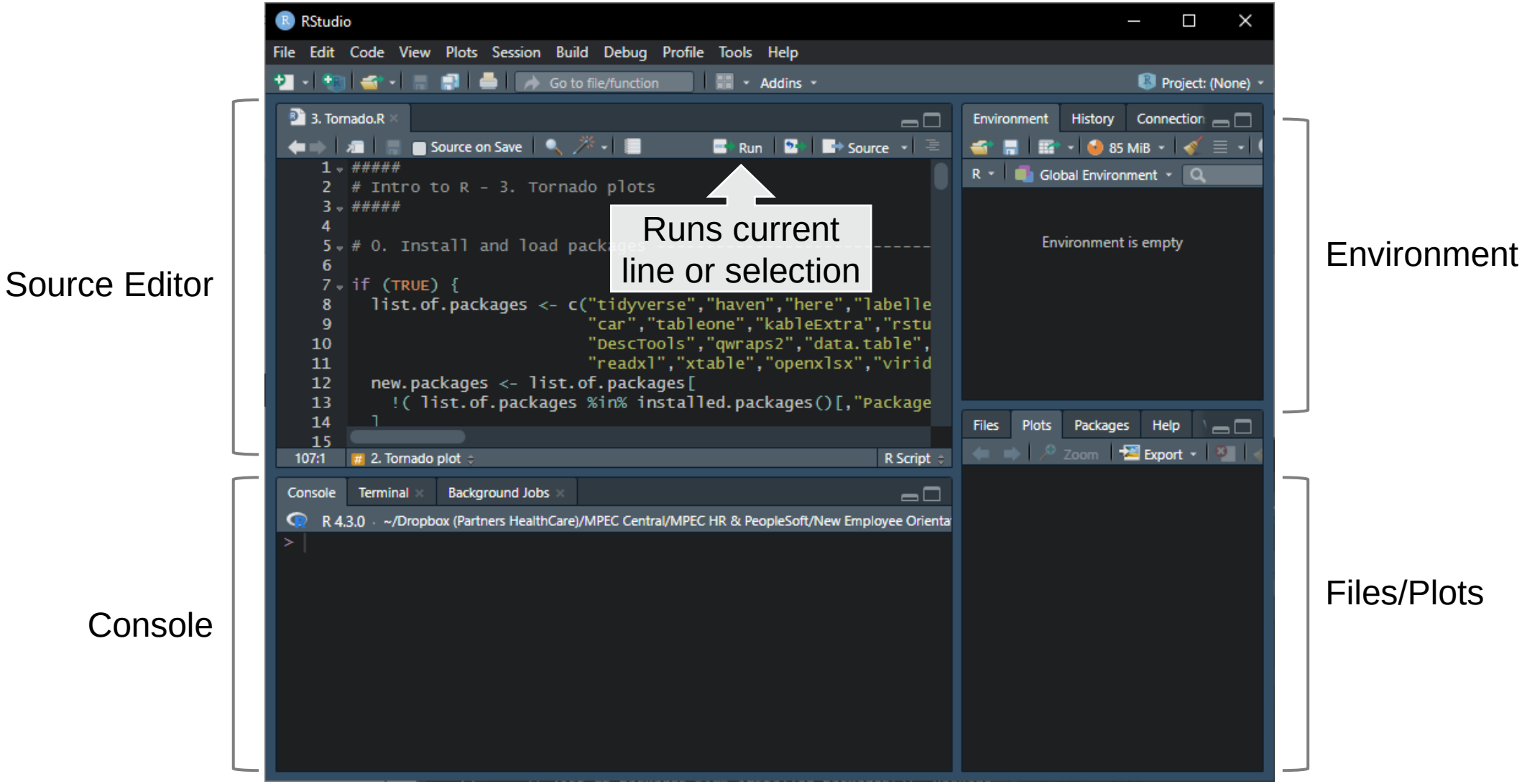
Set the directory (folder where the R script file is located)

Read in CSV file

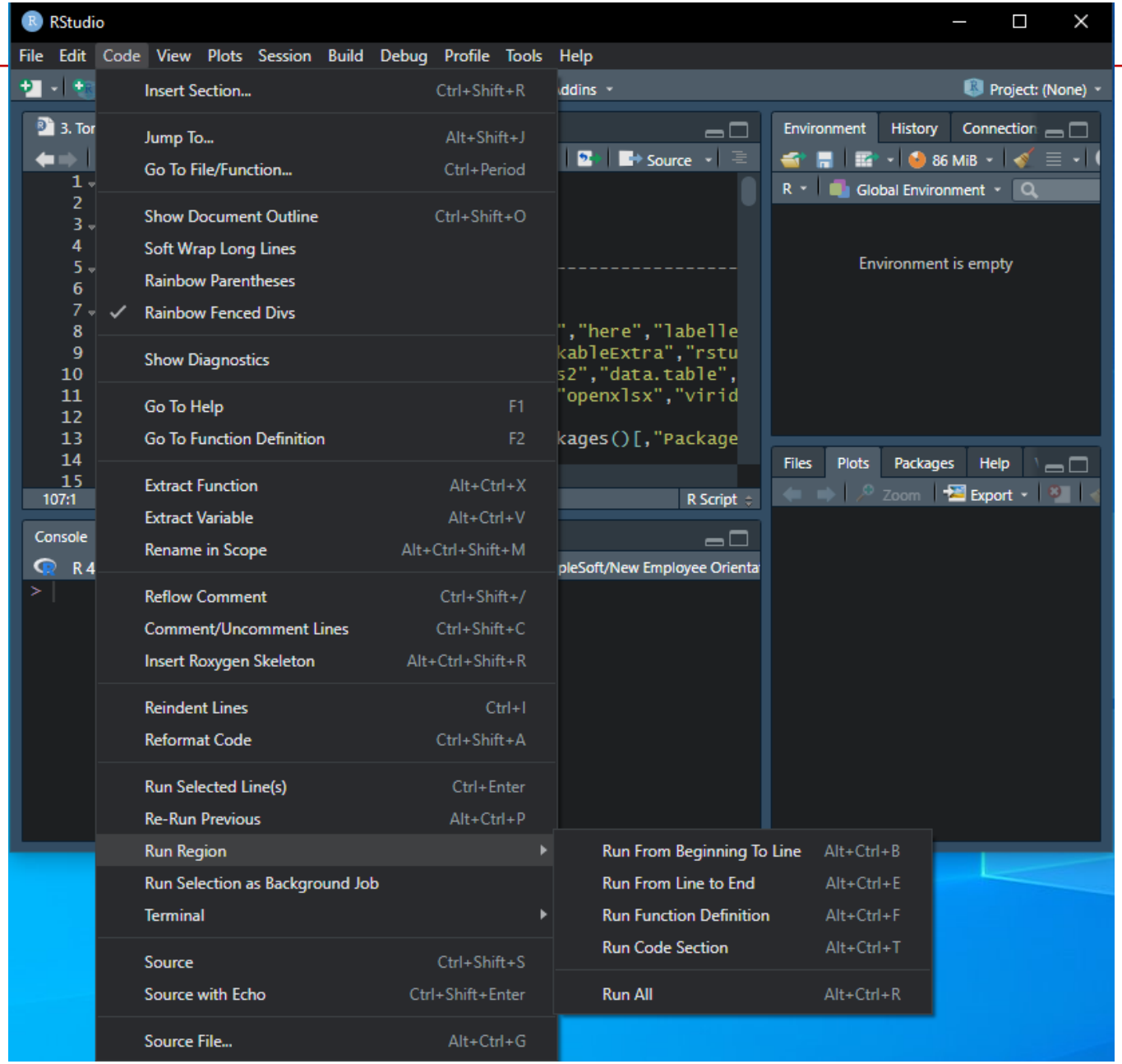
Check data



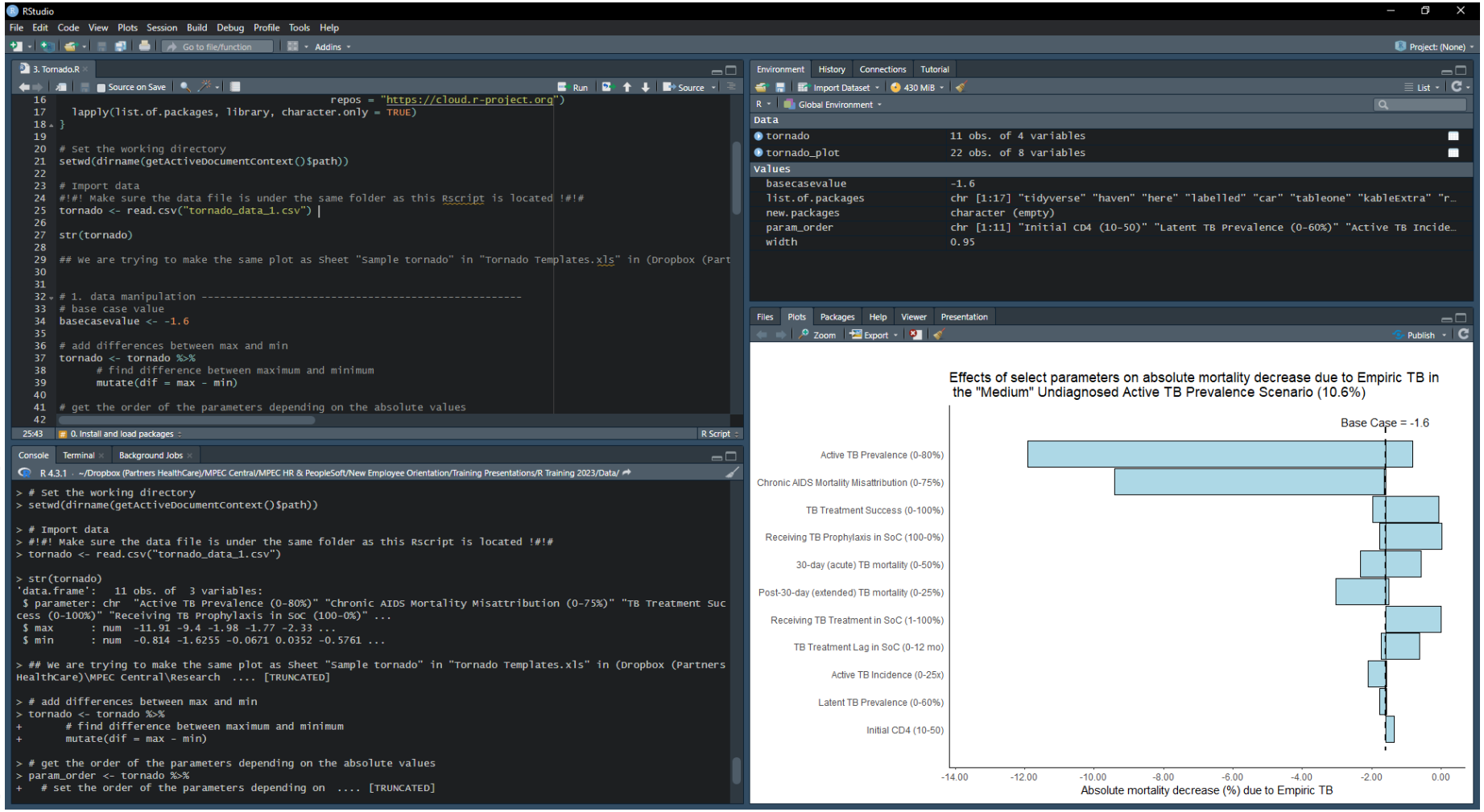
Running script



Run options


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Insert Section... Ctrl+Shift+R
Jump To... Alt+Shift+J
Go To File/Function... Ctrl+Period
Show Document Outline Ctrl+Shift+O
Soft Wrap Long Lines
Rainbow Parentheses
Rainbow Fenced Divs
Show Diagnostics
Go To Help F1
Go To Function Definition F2
Extract Function Alt+Ctrl+X
Extract Variable Alt+Ctrl+V
Rename in Scope Alt+Ctrl+Shift+M
Reflow Comment Ctrl+Shift+/
Comment/Uncomment Lines Ctrl+Shift+C
Insert Roxygen Skeleton Alt+Ctrl+Shift+R
Reindent Lines Ctrl+I
Reformat Code Ctrl+Shift+A
Run Selected Line(s) Ctrl+Enter
Re-Run Previous Alt+Ctrl+P
Run Region
Run Selection as Background Job
Terminal
Source Ctrl+Shift+S
Source with Echo Ctrl+Shift+Enter
Source File... Alt+Ctrl+G
Run From Beginning To Line Alt+Ctrl+B
Run From Line to End Alt+Ctrl+E
Run Function Definition Alt+Ctrl+F
Run Code Section Alt+Ctrl+T
Run All Alt+Ctrl+R
Project: (None)
Environment History Connection
Global Environment
Environment is empty
Files Plots Packages Help
Console

Output from Run All

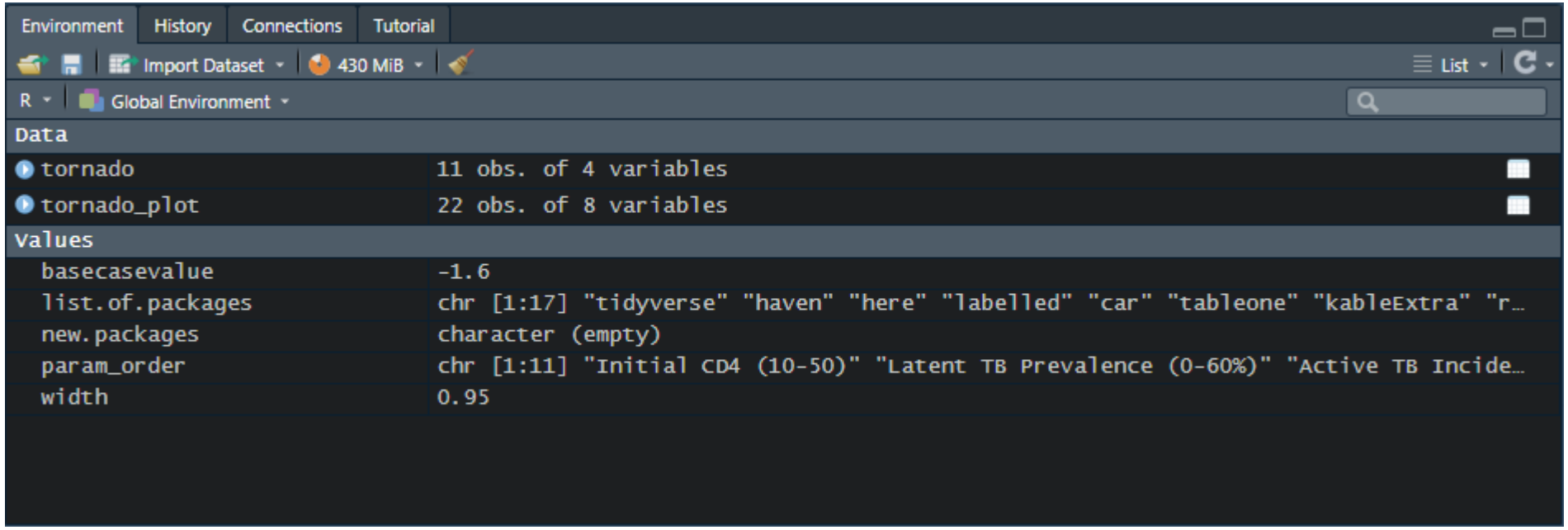


Console

Environment

Files/Plots

Environment pane

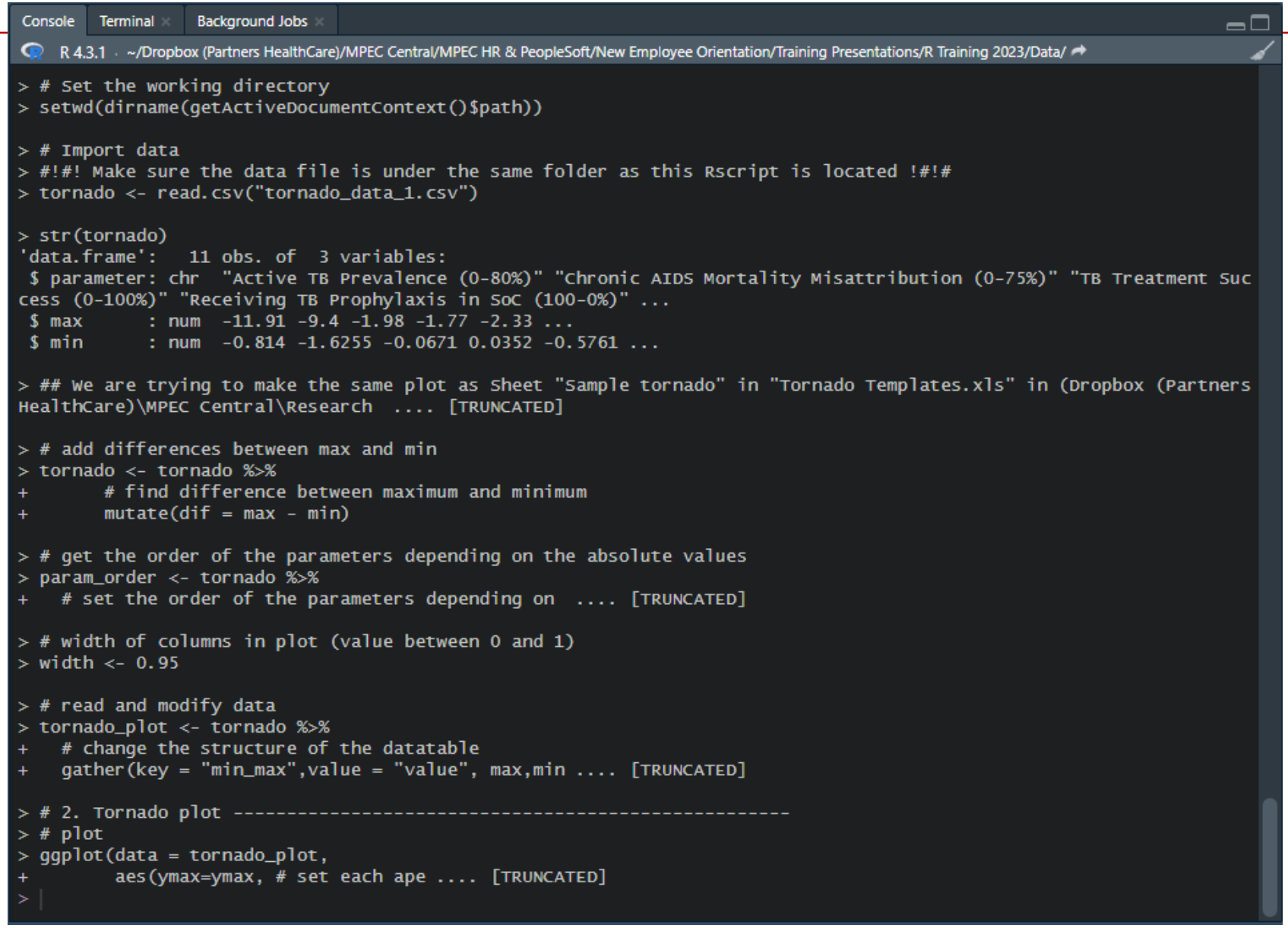


Tables of data

Variables, vectors, lists, functions, etc.

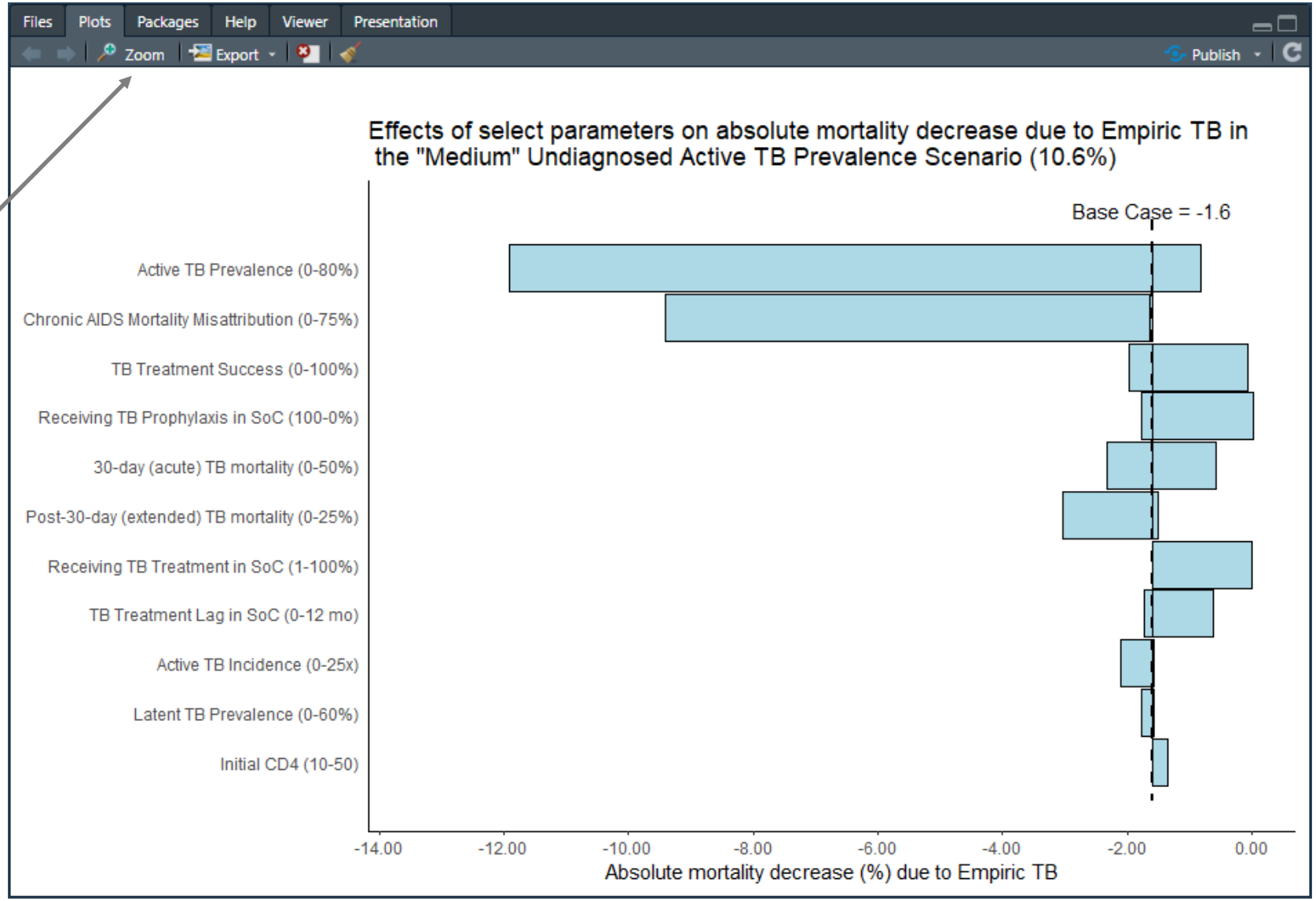
Console pane

- Displays commands that have been run or are being run
- Can directly type and run commands in the console

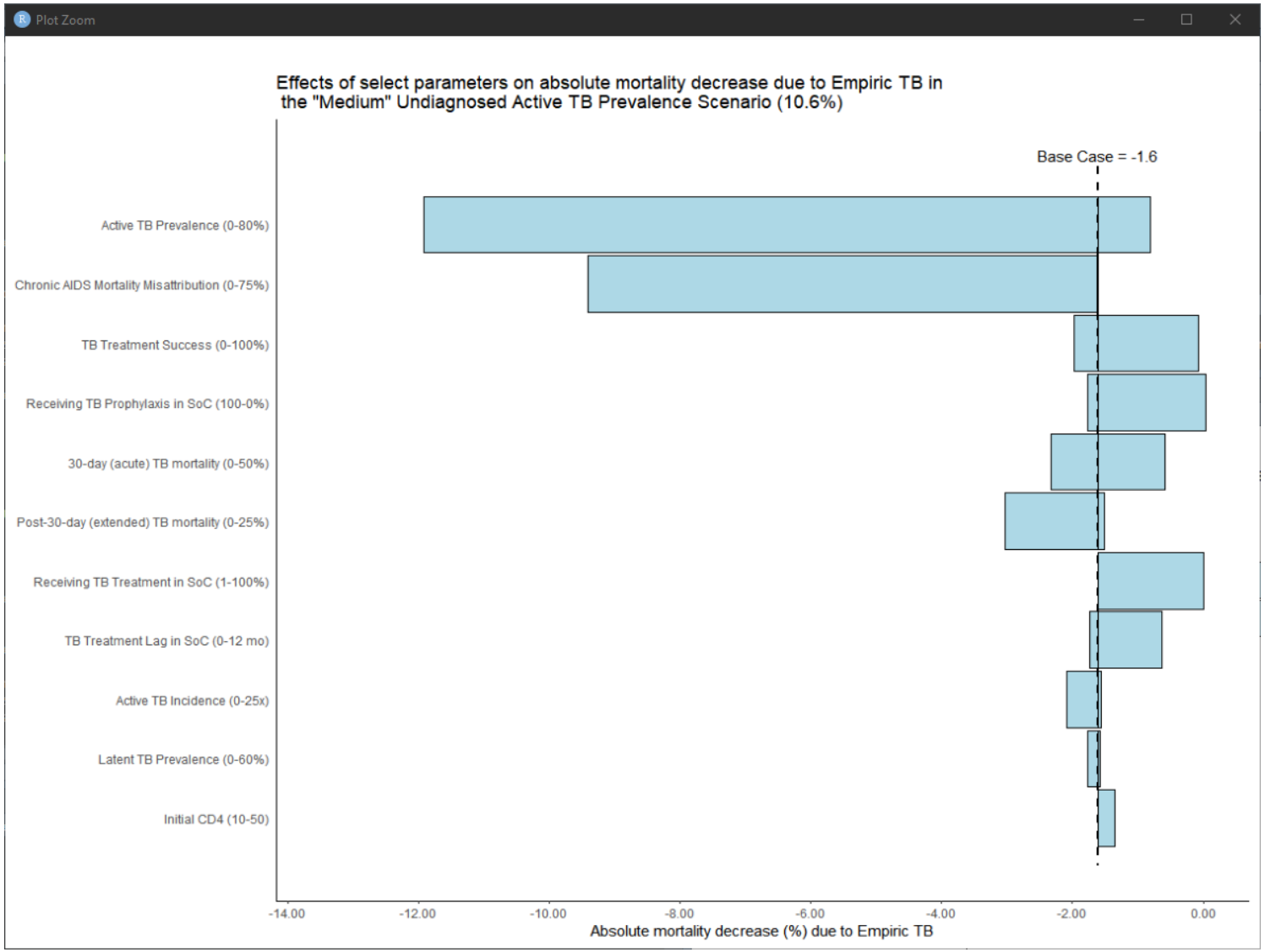


Files/Plots pane

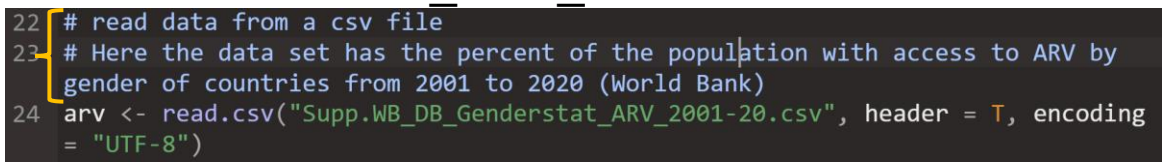
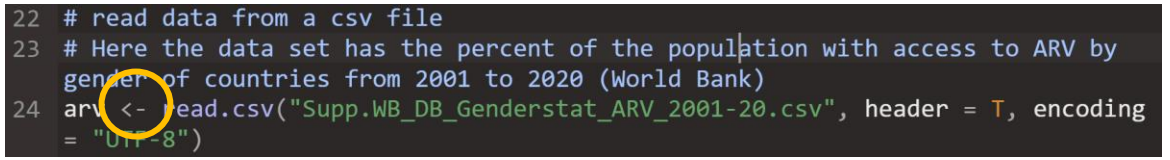
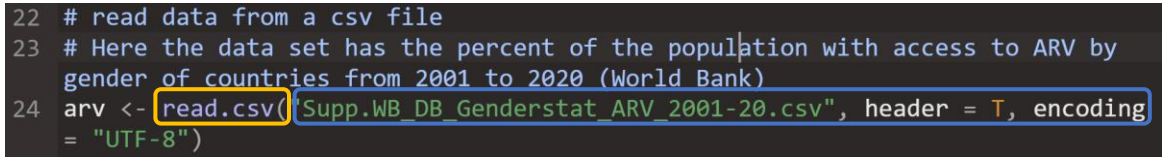
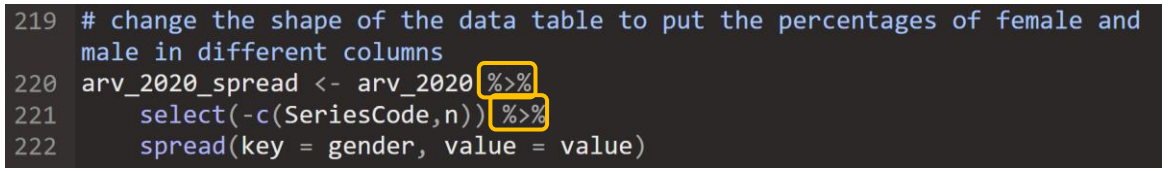
Opens figure in a separate window



Plot Zoom


Plot Zoom
Effects of select parameters on absolute mortality decrease due to Empiric TB in the "Medium" Undiagnosed Active TB Prevalence Scenario (10.6%)
Base Case = -1.6
Active TB Prevalence (0-80%)
Chronic AIDS Mortality Misattribution (0-75%)
TB Treatment Success (0-100%)
Receiving TB Prophylaxis in SoC (100-0%)
30-day (acute) TB mortality (0-50%)
Post-30-day (extended) TB mortality (0-25%)
Receiving TB Treatment in SoC (1-100%)
TB Treatment Lag in SoC (0-12 mo)
Active TB Incidence (0-25x)
Latent TB Prevalence (0-60%)
Initial CD4 (10-50)
-14.00
-12.00
-10.00
-8.00
-6.00
-4.00
-2.00
0.00
Absolute mortality decrease (%) due to Empiric TB

Basic Operation in R

Command	Meaning	Example
Line starting with #	Comments that don't influence functions or operations.	Line 22 and 23in "R_Data_Vis.R" 
a <- b	Assign b to a. b can be the results of functions and calculations. e.g., a <- 1 + 1, then typing "a" will give 2.	Line 24 in "R_Data_Vis.R" 
function(arguments)	The **function** works by following specified **arguments**	Line 24 read.csv in "R_Data_Vis.R" 
%>% or **|>**	To link one function to another.	Line 220 in "R_Data_Vis.R" 

Frequently used functions

Function	Purpose
read.csv()	To read data from csv. file
read_excel()	To read data from MS Excel file
str()	To check the data and type of data of each column
View()	To check the whole data in a new tab, which looks similar to an Excel format
ggplot()	To make plots. Only ggplot function doesn't work. - Know what column of data should be placed on x- and y-axes. - Know what additional functions you need to create the type of graph you want to make and specify additional functions with "**+**" following ggplot() (Not "**%>%**"!)

Google whenever you come across unknown functions!

Table of contents

0. Install and load packages (and data!)

- 0. Install and load packages (and data!)
  - **Don’t change anything until Line 18 “setwd(dirname(….))”**
  - Import data from csv. data on percentage of people with access to ARV 2001-2020 (World Bank), and toy data for tornado and heatmap: read.csv()
  - Check the imported data
    - `str(arv)`

```
'data.frame':   2440 obs. of  9 variables:
 $ n           : int  1 2 9 10 11 12 25 26 29 30 ...
 $ SeriesName  : chr  "Access to anti-retroviral drugs" "Access to anti-retroviral dr
anti-retroviral drugs" ...
 $ gender      : chr  "female" "male" "female" "male" ...
 $ unit        : chr  "%" "%" "%" "%" ...
 $ SeriesCode  : chr  "SH.HIV.ARTC.FE.ZS" "SH.HIV.ARTC.MA.ZS" "SH.HIV.ARTC.FE.ZS" "SH
 $ CountryName: chr  "Afghanistan" "Afghanistan" "Argentina" "Argentina" ...
 $ CountryCode: chr  "AFG" "AFG" "ARG" "ARG" ...
 $ year        : int  2001 2001 2001 2001 2001 2001 2001 2001 2001 2001 ...
 $ value       : int  0 0 20 23 54 50 0 0 0 0 ...
```

	A	B	C	D	E	F	G	H	I
1	n	SeriesName	gender	unit	SeriesCode	CountryName	CountryCode	year	value
2	1	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Afghanistan	AFG	2001	0
3	2	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Afghanistan	AFG	2001	0
4	9	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Argentina	ARG	2001	20
5	10	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Argentina	ARG	2001	23
6	11	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Australia	AUS	2001	54
7	12	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Australia	AUS	2001	50
8	25	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Benin	BEN	2001	0
9	26	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Benin	BEN	2001	0
10	29	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Bolivia	BOL	2001	0
11	30	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Bolivia	BOL	2001	0
12	31	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Botswana	BWA	2001	1
13	32	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Botswana	BWA	2001	1
14	35	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Burkina Faso	BFA	2001	0
15	36	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Burkina Faso	BFA	2001	0
16	39	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cabo Verde	CPV	2001	0
17	40	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cabo Verde	CPV	2001	0
18	49	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Chile	CHL	2001	12
19	50	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Chile	CHL	2001	12
20	51	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Colombia	COL	2001	0
21	52	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Colombia	COL	2001	0
22	61	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cote d'Ivoire	CIV	2001	0
23	62	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cote d'Ivoire	CIV	2001	0
24	65	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cuba	CUB	2001	4
25	66	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cuba	CUB	2001	3
26	69	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Denmark	DNK	2001	47
27	70	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Denmark	DNK	2001	46

0. Install and load packages (and data!)

- `View(arv)`

Untitled2* | travel_modeling.Rmd | 1. Histo_Bar_Line_Scatter.R* | arv | Untitled1*

Filter

	n	SeriesName	gender	unit	SeriesCode	CountryName	CountryCode	year	value
1	1	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Afghanistan	AFG	2001	0
2	2	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Afghanistan	AFG	2001	0
3	9	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Argentina	ARG	2001	20
4	10	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Argentina	ARG	2001	23
5	11	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Australia	AUS	2001	54
6	12	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Australia	AUS	2001	50
7	25	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Benin	BEN	2001	0
8	26	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Benin	BEN	2001	0
9	29	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Bolivia	BOL	2001	0
10	30	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Bolivia	BOL	2001	0
11	31	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Botswana	BWA	2001	1
12	32	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Botswana	BWA	2001	1

	A	B	C	D	E	F	G	H	I
1	n	SeriesName	gender	unit	SeriesCode	CountryName	CountryCode	year	value
2	1	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Afghanistan	AFG	2001	0
3	2	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Afghanistan	AFG	2001	0
4	9	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Argentina	ARG	2001	20
5	10	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Argentina	ARG	2001	23
6	11	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Australia	AUS	2001	54
7	12	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Australia	AUS	2001	50
8	25	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Benin	BEN	2001	0
9	26	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Benin	BEN	2001	0
10	29	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Bolivia	BOL	2001	0
11	30	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Bolivia	BOL	2001	0
12	31	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Botswana	BWA	2001	1
13	32	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Botswana	BWA	2001	1
14	35	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Burkina Faso	BFA	2001	0
15	36	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Burkina Faso	BFA	2001	0
16	39	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cabo Verde	CPV	2001	0
17	40	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cabo Verde	CPV	2001	0
18	49	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Chile	CHL	2001	12
19	50	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Chile	CHL	2001	12
20	51	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Colombia	COL	2001	0
21	52	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Colombia	COL	2001	0
22	61	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cote d'Ivoire	CIV	2001	0
23	62	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cote d'Ivoire	CIV	2001	0
24	65	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Cuba	CUB	2001	4
25	66	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Cuba	CUB	2001	3
26	69	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Denmark	DNK	2001	47
27	70	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Denmark	DNK	2001	46

1. Histogram

- 1. Histogram
  - Aim: To visualize the distribution of countries by percentage of females who have access to ARV in the world in 2020
  - Subset data with “female” in “gender” column and “2020” in “year” column

```
37 # select data of female in 2020
38 arv_fe_2020 <- arv[arv$gender=="female" & arv$year == 2020,]
```

  - (Check data)

Some ideas for playing around:
- How about “male” in “gender”?
- How about other years?

1. Histogram

- 1. Histogram
  - Plot the data: here **ggplot()** comes!
    - Specify the data (“arv_fe_2020”) and which variable is used on the x-axis (“value”).
    - Specify the type of graph: **geom_histogram()**
  - Save the plot in a png. format

```
# plot
ggplot(data=arv_fe_2020,  # specify the data you want to use
       aes(x=value)) +  # specify the column that you want to show in th
-axis
  # specify the type of your graph
  geom_histogram(
    fill = "grey",  # which color is used to fill the bins
    color = "black",  # which color is used to draw the borders of bins
    bins = 10,       # number of bins
  ) +
  # add title
  ggtitle("Distribution of the percentage of female who have access to a
-retroviral drugs \n in the world in 2020") +  #"\n" means change a line
  # add x-axis label
  xlab("Percentage (%)") +
  # add y-axis label
  ylab("Number of countries") +
  # customize the background theme
  theme_bw()
```

```
# save the figure in png
ggsave("hist.png", width = 8, height = 6)
```

Some ideas for playing around:

- How about other colors inside bins?
- How about other colors for the borders of the bins?
- How about different numbers of bins?
- How about different titles and labels in the x-/y-axes?

Table of contents

2. Bar chart

- 2. Bar chart
  - Aim: To visualize the percentage of females who have access to ARV in 2020 by country
  - Use the same data as the histogram (“arv_fe_2020”)
  - Plot the data: **ggplot()**!
    - Specify the data (“arv_fe_2020”) and which variable is used on the x-axis (“CountryName”) and the y-axis (“value”).
    - Specify the type of graph: **geom_bar()**

```
94  # plot
95  ggplot(data=arv_fe_2020, # specify data
96         aes(x=CountryName, # specify the column that you want to show in the x-axis
97             y=value)) +    # specify the column that you want to show in the y-axis
98    # specify the type of your graph
```

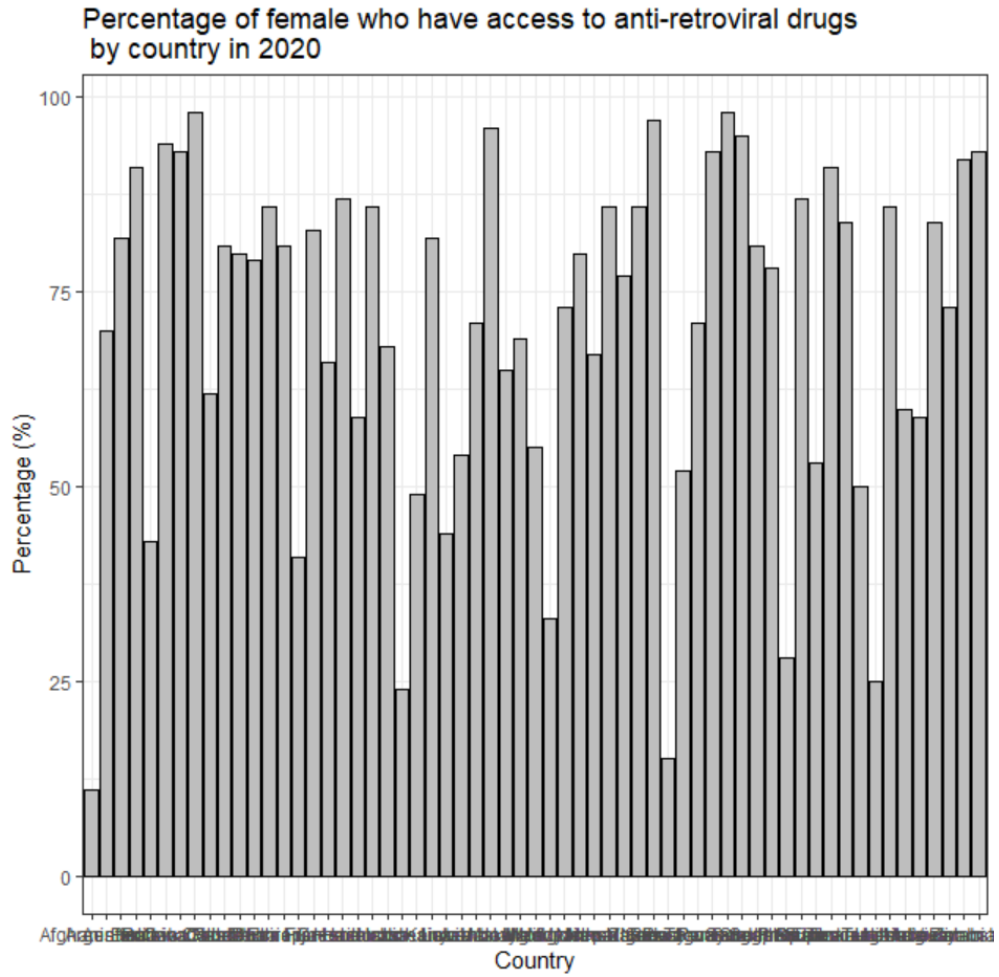
2. Bar chart

- 2. Bar chart
  - Specify the type of graph: **geom_bar()**
    - **Don’t forget “ stat = “identity” ”!**
    - Without it, the plot will show the counts of the variable of the x-axis.

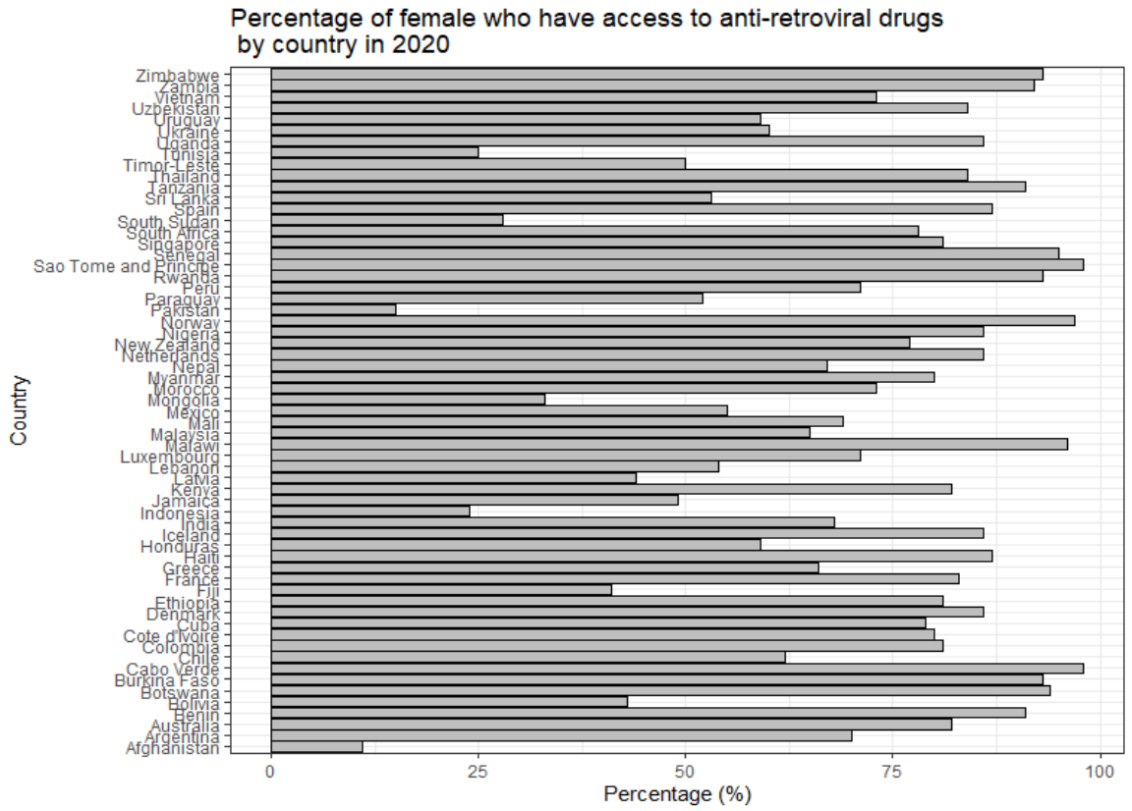
```
  # specify the type of your graph
  geom_bar(
    fill = "grey",  # which color is used to fill the bins
    color = "black",  # which color is used to draw the borders of bins
    stat = "identity" # when you specify x and y in geom_bar, don't forget this argument
  ) +
  # add title
  ggtitle("Percentage of female who have access to anti-retroviral drugs \n by country in 2020") +
  # add x-axis label
  xlab("Country") +
  # add y-axis label
  ylab("Percentage (%)") +
  # customize the background theme
  theme_bw()
```

2. Bar chart

- 2. Bar chart



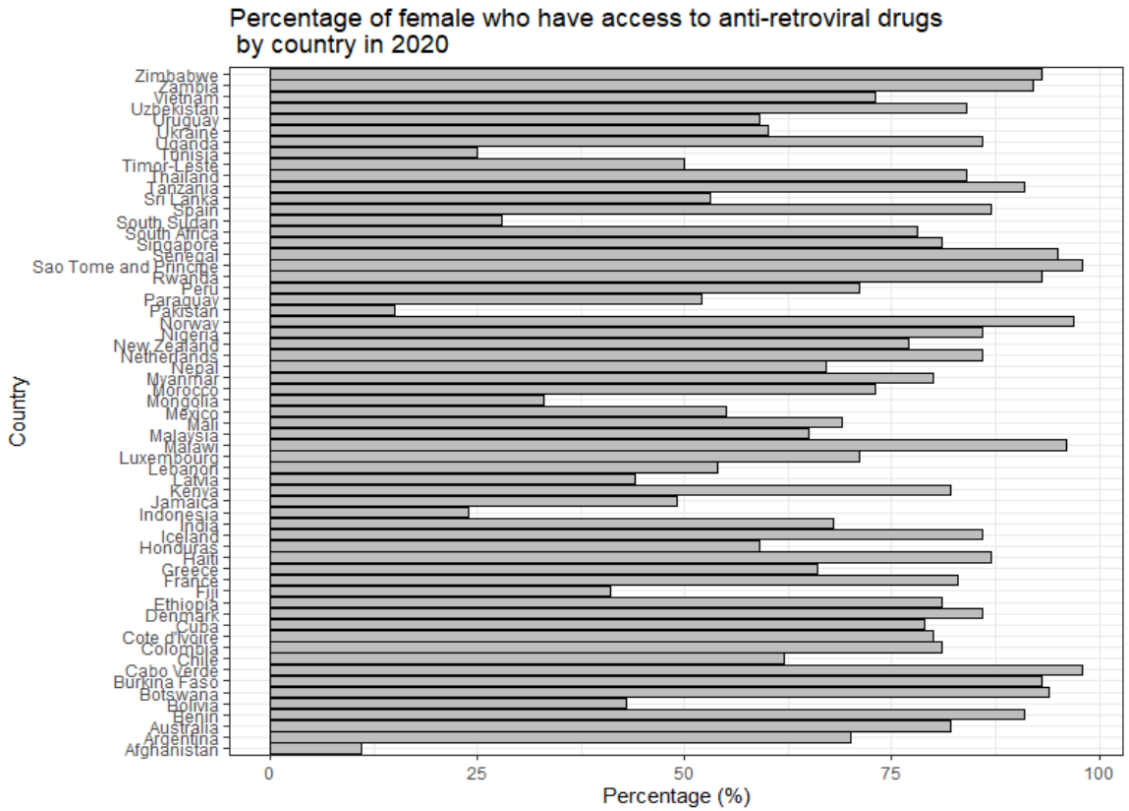

```
133    coord_flip() # to switch x and y
```



2. Bar chart

- 2. Bar chart

```
136 ggplot(data=arv_fe_2020, # specify data
137        aes(x=reorder(CountryName, value), # x-axis, rec
138            y=value)) +    # specify the column that you
```



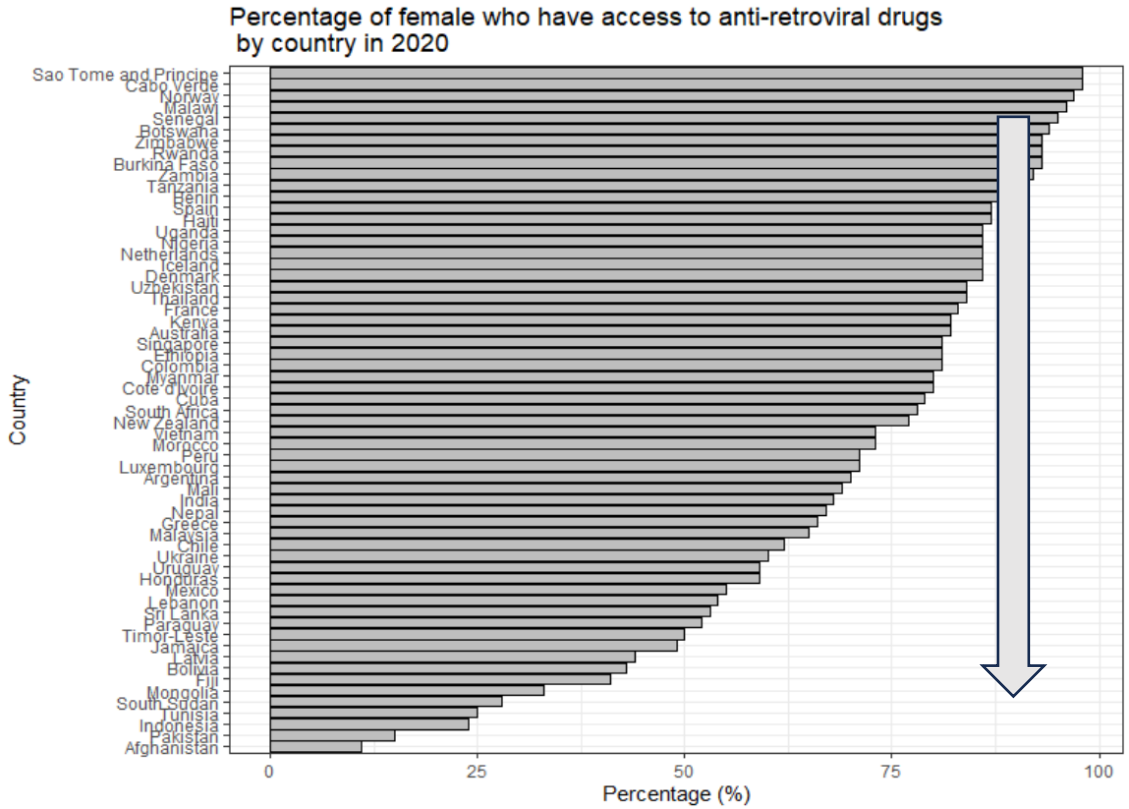


Table of contents

2.1. Stacked bar chart

- 2.1. Stacked bar chart
  - Aim: To visualize the percentage of females and males who have access to ARV in 2020 by country
  - Subset data with “2020” in “year” column

```
160 # select data of female & male in 2020
161 arv_2020 <- arv[arv$year == 2020,]
```

  - (Check data)

Some ideas for playing around:
- How about other years?

2.1. Stacked bar chart

- 2.1. Stacked bar chart
  - Plot the data: **ggplot()**!
    - Specify the data (“arv_fe_2020”) and which variable is used on the x-axis (“CountryName” with reorder()), the y-axis (“value”), and **the color to fill the bars (“gender”)**
    - Specify the type of graph: **geom_bar()**
      - Don’t forget “ **position= “stack” ”**!

```
166 # plot
167 ggplot(data=arv_2020, # specify the data
168        aes(x=reorder(CountryName, value), # x-axis, re
169            y=value,    # specify the column that you w
170            fill = gender)) +  # specify the subgroup
171   # specify the type of your graph
172   geom_bar(
173     stat = "identity", # when you specify x and y in g
174     position = "stack" # when you make a stacked barch
175   ) +
176   # add title
177   ggtitle("Percentage of individuals with access to ar
    2020") +
178   # add x-axis label
179   xlab("Country") +
180   # add y-axis label
181   ylab("Percentage (%)") +
182   # customize the background theme
183   theme_bw() +
184   coord_flip() # to switch x and y
```

2.1. Stacked bar chart

- 2.1. Stacked bar chart

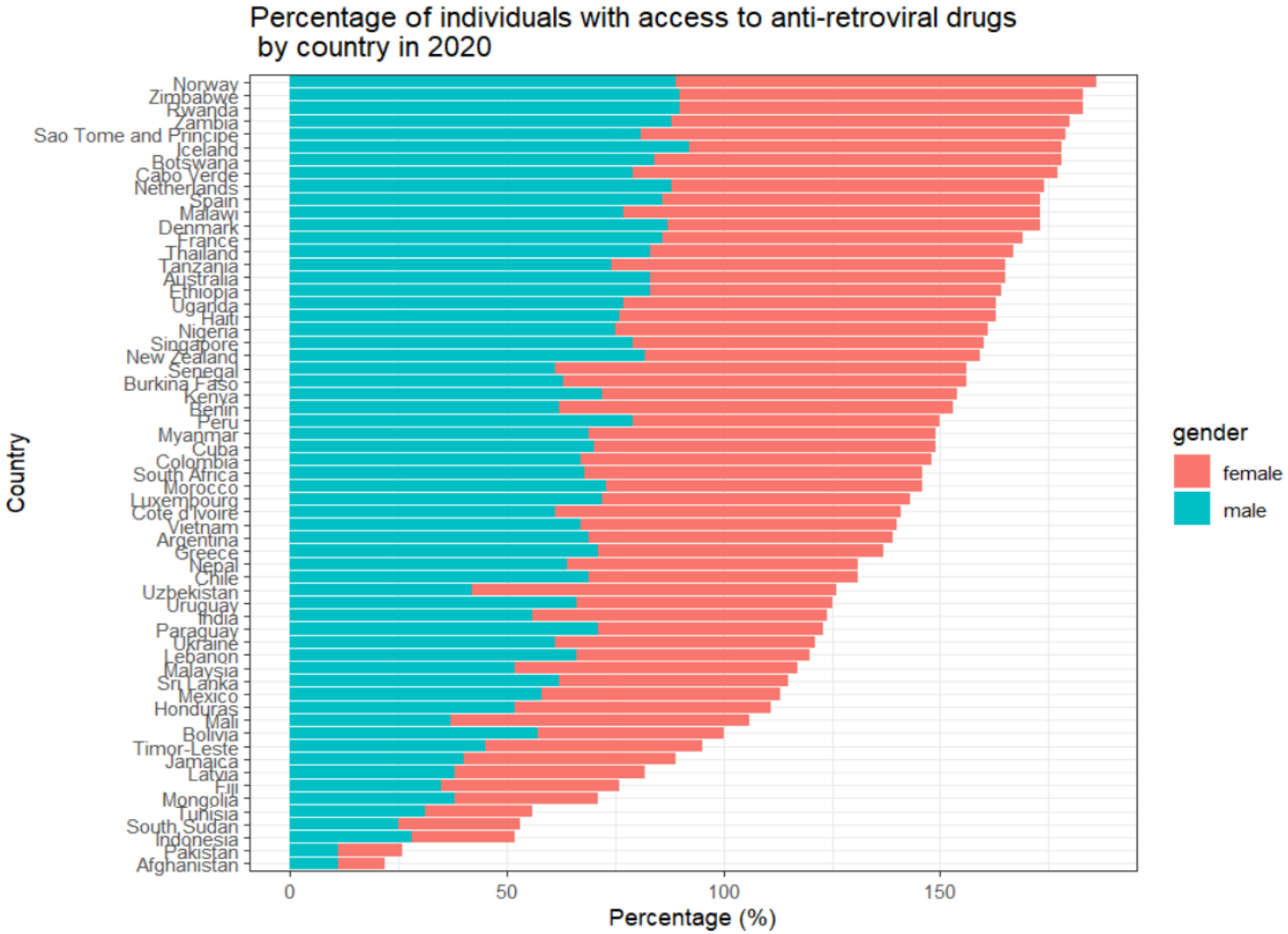


Table of contents

3. Line chart

- 3. Line chart
  - Aim: To visualize the trend of the percentages of females and males who have access to ARV in Myanmar from 2001 to 2020
  - Subset data with “Myanmar” in “CountryName” column

```
191 # select data of Myanmar from 2001 to 2020
192 arv_myanmar <- arv[arv$CountryName=="Myanmar",]
```

  - (Check data)

Some ideas for playing around:
- How about other countries?

3. Line chart

- 3. Line chart
  - Plot the data: **ggplot()**!
    - Specify the data (“arv_myanmar”) and which variable is used on the x-axis (“year”), the y-axis (“value”), how to **group the lines (“gender”)**, and the **color of the lines (“gender”)**
    - Specify the type of graph: **geom_line()**

```
197 # plot
198 ggplot(data=arv_myanmar, # specify data
199        aes(x=year, # specify the column that y
200            y=value,  # # specify the column th
201            group = gender,    # different line
202            color = gender)) +   # different co
203   # specify the type of your graph
204   geom_line() +
205   # add title
206   ggtitle("Percentage of individuals with acce
    from 2001 to 2020") +
207   # add x-axis label
208   xlab("Year") +
209   # add y-axis label
210   ylab("Percentage (%)") +
211   # customize the background theme
212   theme_bw()
```

3. Line chart

- 3. Line chart

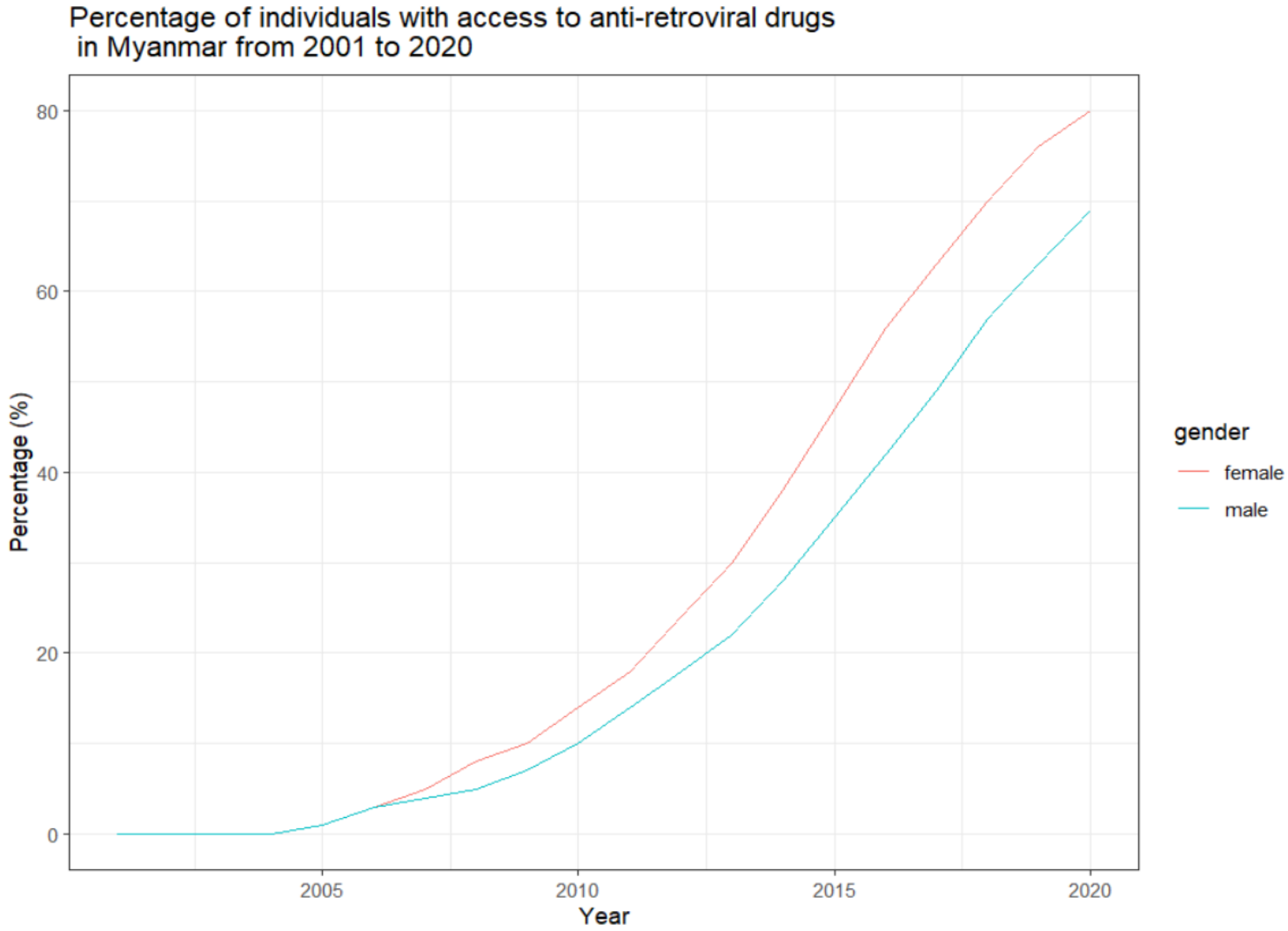


Table of contents

4. Scatter plot

- 4. Scatter plot
  - Aim: To visualize the relationship between the percentages of females and males who have access to ARV in 2020

  - Change the data structure of “arv_2020”

	n	SeriesName	gender	unit	SeriesCode	CountryName	CountryCode	year	value
19	4827	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Afghanistan	AFG	2020	1
20	4828	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Afghanistan	AFG	2020	1
21	4835	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Argentina	ARG	2020	7
22	4836	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Argentina	ARG	2020	6
23	4837	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Australia	AUS	2020	8
24	4838	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Australia	AUS	2020	8
25	4851	Access to anti-retroviral drugs	female	%	SH.HIV.ARTC.FE.ZS	Benin	BEN	2020	9
26	4852	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Benin	BEN	2020	6
27	4855	Access to anti-retroviral drugs	male	%	SH.HIV.ARTC.MA.ZS	Bolivia	BOL	2020	5

4. Scatter plot

- 4. Scatter plot
  - Change the data structure of “arv_2020” to “arv_2020_spread” with different columns with each gender

```
219 # change the shape of the data table to put the p
220 arv_2020_spread <- arv_2020 %>%
221     select(-c(SeriesCode,n)) %>%
222     spread(key = gender, value = value)
```

	SeriesName	unit	CountryName	CountryCode	year	female	male
1	Access to anti-retroviral drugs	%	Afghanistan	AFG	2020	11	11
2	Access to anti-retroviral drugs	%	Argentina	ARG	2020	70	69
3	Access to anti-retroviral drugs	%	Australia	AUS	2020	82	83
4	Access to anti-retroviral drugs	%	Benin	BEN	2020	91	62
5	Access to anti-retroviral drugs	%	Bolivia	BOL	2020	43	57
6	Access to anti-retroviral drugs	%	Botswana	BWA	2020	94	84
7	Access to anti-retroviral drugs	%	Burkina Faso	BFA	2020	93	63

4. Scatter plot

- 4. Scatter plot
  - Plot the data: **ggplot()**!
    - Specify the data (“arv_2020_spread”) and which variable is used on the x-axis (“female”), and the y-axis (“male”)
    - Specify the type of graph: **geom_point()**

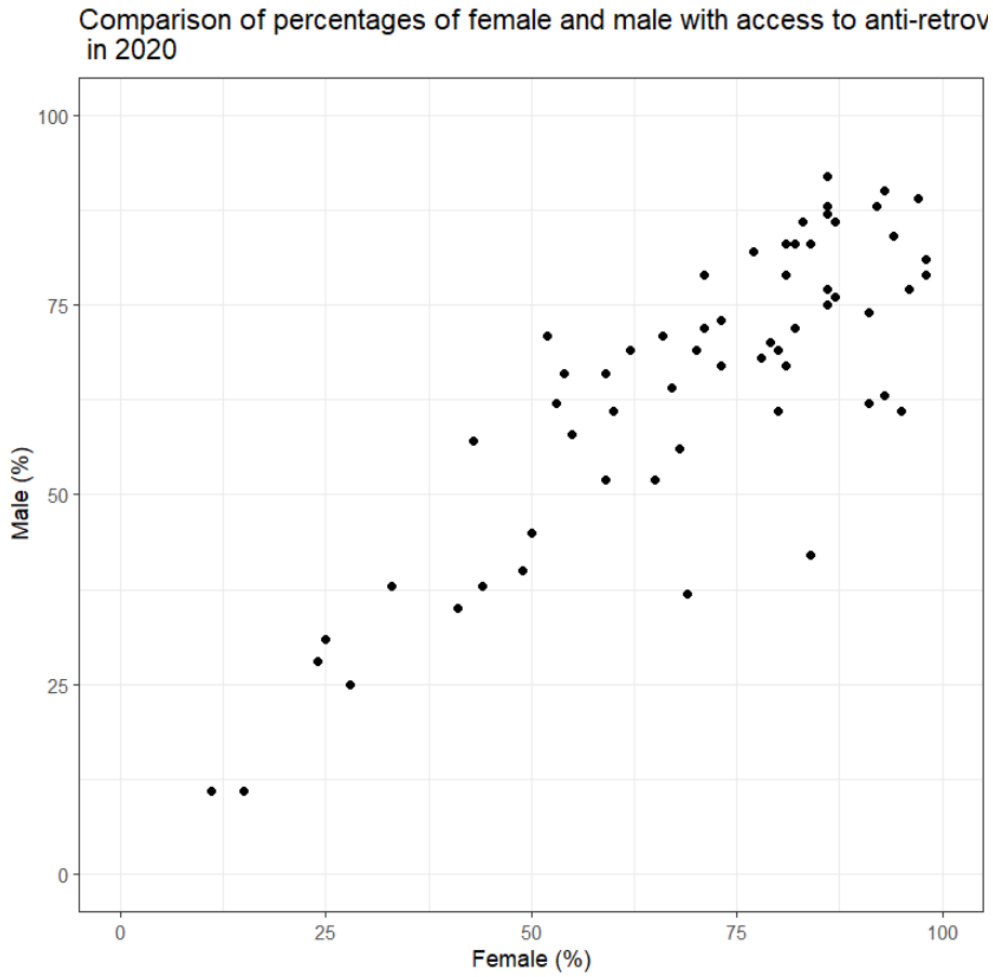
```
227 # plot
228 ggplot(data=arv_2020_spread, # specify data
229        aes(x=female, # specify the column tha
230            y=male)) +   # specify the column
231   # specify the type of your graph
232   geom_point() +
233   # make sure x-axis extends to 100%
234   xlim(0,100) +
235   # make sure y-axis extends to 100%
236   ylim(0,100) +
237   # add title
238   ggtitle("Comparison of percentages of femal
    2020") +
239   # add x-axis label
240   xlab("Female (%)") +
241   # add y-axis label
242   ylab("Male (%)") +
243   # customize the background theme
244   theme_bw()
```

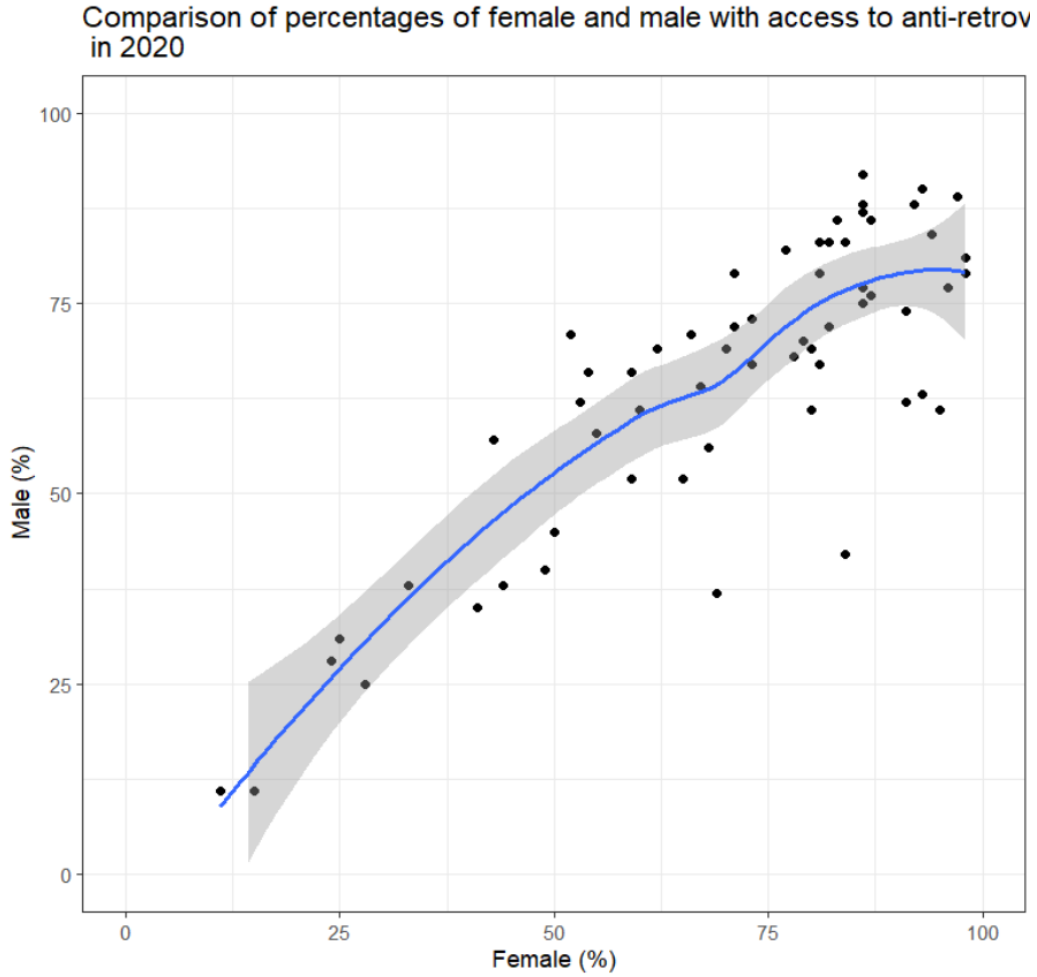
Some ideas for playing around:
- How about flipping the x-axis and y-axis?

4. Scatter plot

- 4. Scatter plot

```
252   # add the linear trend and the confidence interval
253   geom_smooth() +
```

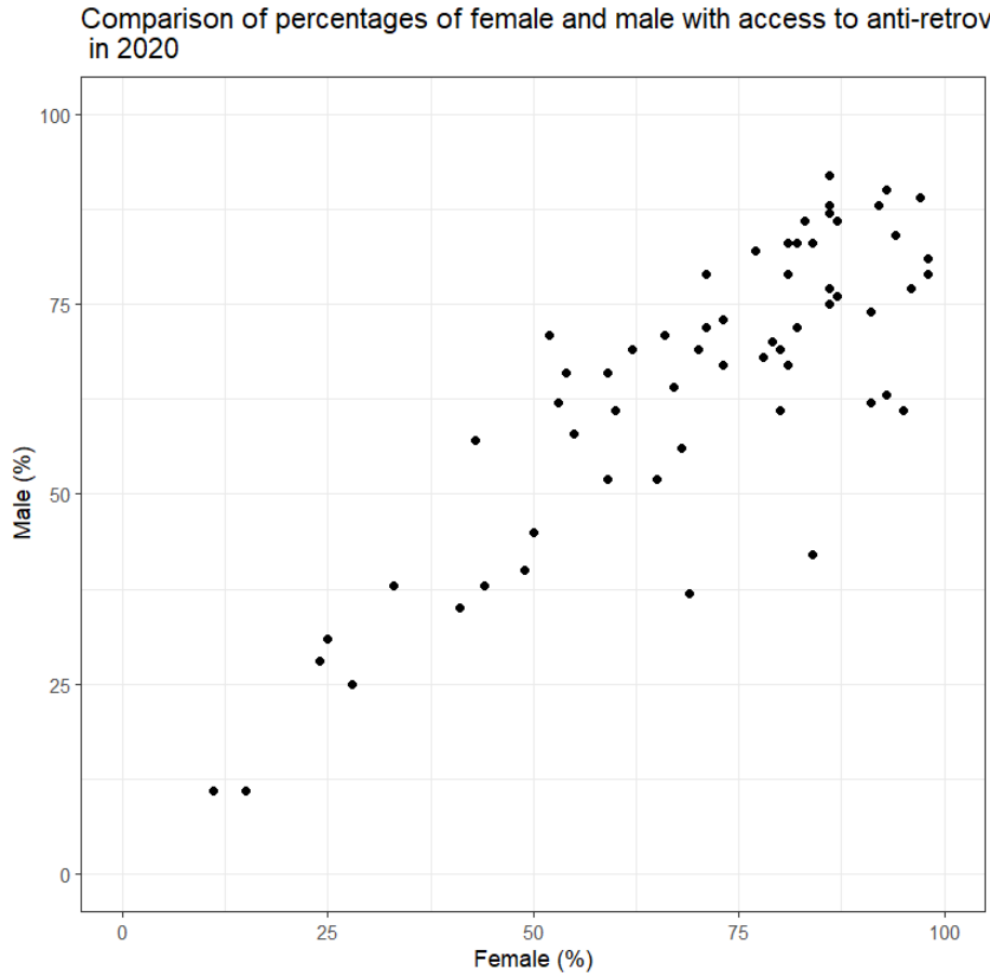




4. Scatter plot

- 4. Scatter plot

```
273   # add country names
274   geom_text(
275     label = arv_2020_spread$CountryName, # specify texts
276     nudge_x = 1, nudge_y = 3,   # shift the texts along x and y axis
277     check_overlap = TRUE    # avoid overlap
278   ) +
```



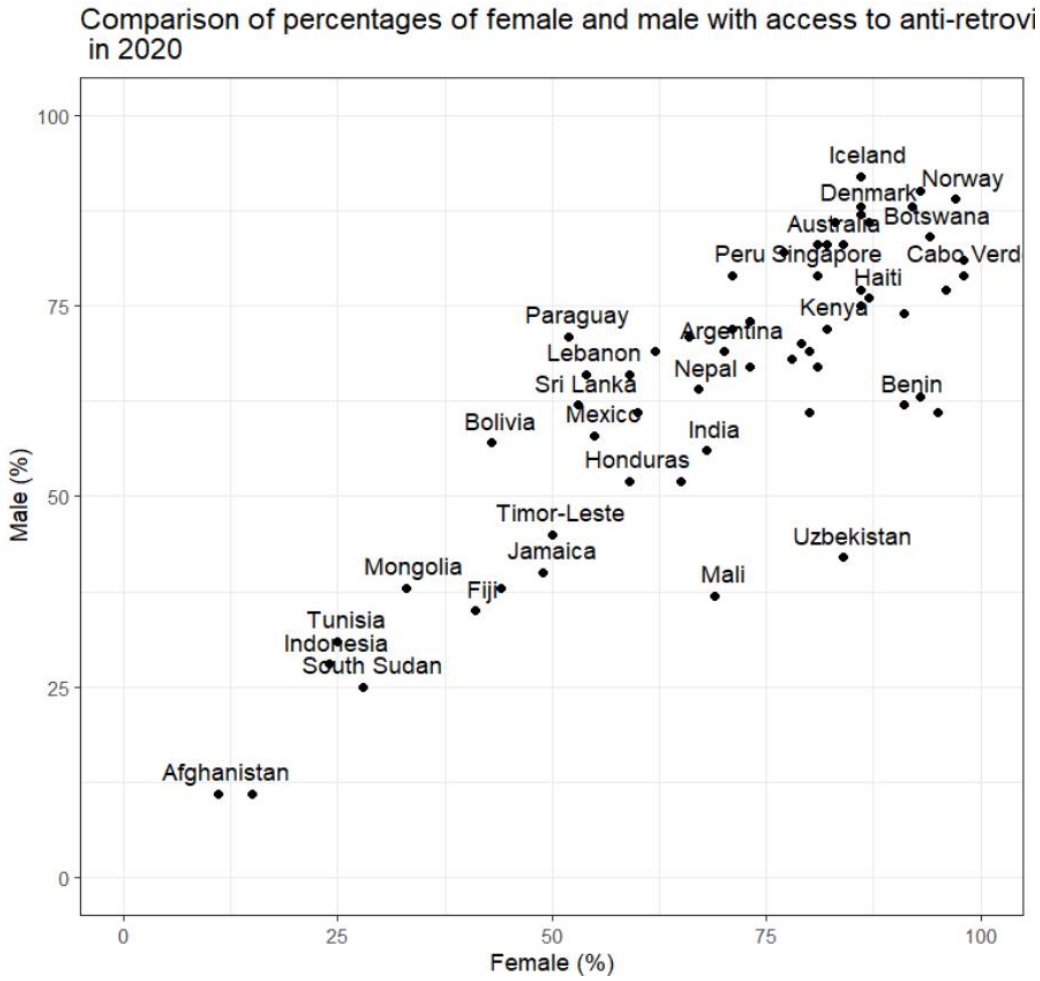
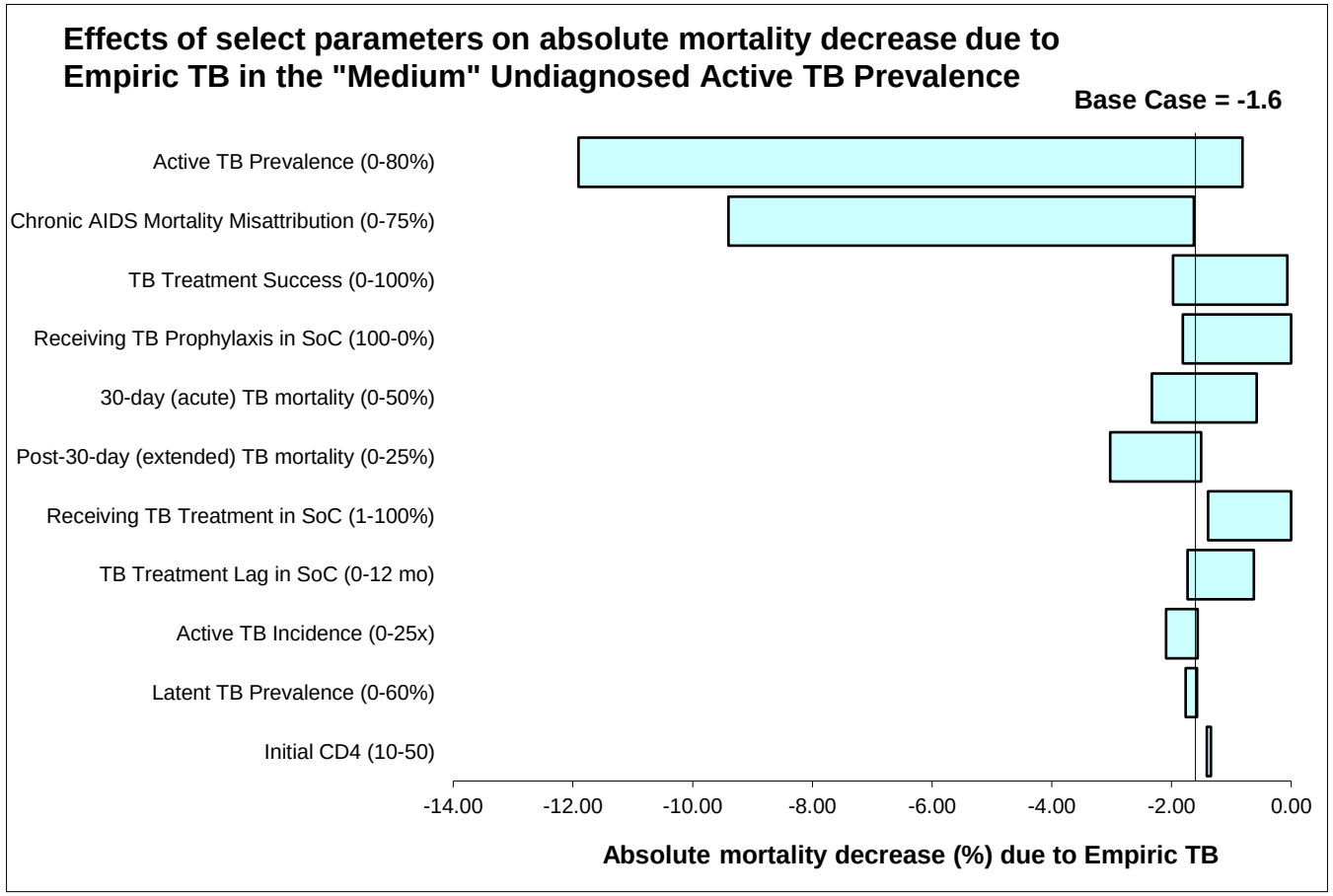


Table of contents

5. Tornado plot

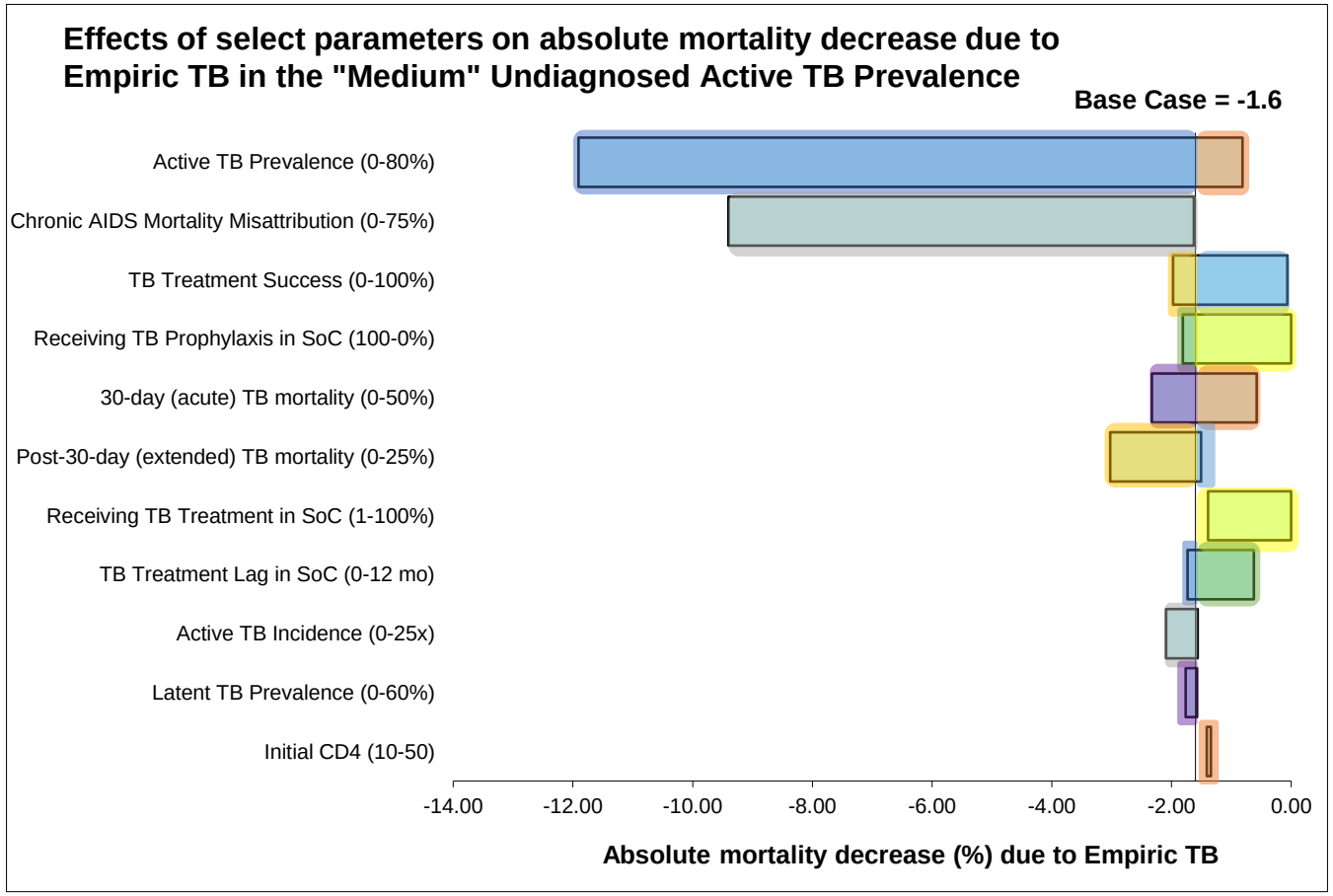
- Aim: To visualize the effect ranges on outputs when changing one input parameter



- Looks like a bar chart, but…

5. Tornado plot

- Aim: To visualize the effect ranges on outputs when changing one input parameter



- Looks like a bar chart, but…
- What we're doing is **to put rectangles**!

- Game plan:
  - Import result data of min and max
  - Find differences (will be side lengths)
  - Identify the positions of each apex of rectangles

5. Tornado plot

- Needed data:
  - Input parameters that varied in your one-way sensitivity analysis
  - Maximum output values by the input parameter
  - Minimum output values by the input parameter
  - (Base case output value)

	A	B	C
1	parameter	max	min
2	Active TB Prevalence (0-80%)	-11.9092	-0.814
3	Chronic AIDS Mortality Misattribution (0-75%)	-9.4012	-1.6255
4	TB Treatment Success (0-100%)	-1.9762	-0.0671
5	Receiving TB Prophylaxis in SoC (100-0%)	-1.7739	0.0352
6	30-day (acute) TB mortality (0-50%)	-2.3272	-0.5761
7	Post-30-day (extended) TB mortality (0-25%)	-3.0225	-1.504
8	Receiving TB Treatment in SoC (1-100%)	-1.3884	0
9	TB Treatment Lag in SoC (0-12 mo)	-1.7323	-0.6258
10	Active TB Incidence (0-25x)	-2.092	-1.5625
11	Latent TB Prevalence (0-60%)	-1.7641	-1.5715
12	Initial CD4 (10-50)	-1.4125	-1.3429

5. Tornado plot

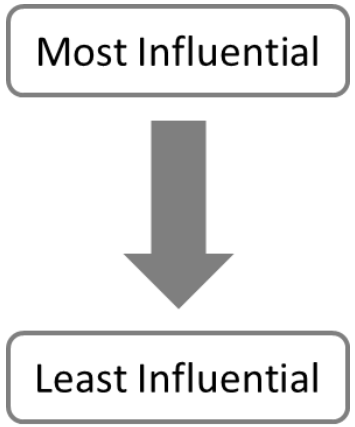
- Calculate the differences between the maximum and minimum values

```
302 # add differences between max and min
303 tornado <- tornado %>%
304       # find difference between maximum and minimum
305       mutate(dif = max - min)
```

	parameter	max	min	dif
1	Active TB Prevalence (0-80%)	-11.9092	-0.8140	-11.0952
2	Chronic AIDS Mortality Misattribution (0-75%)	-9.4012	-1.6255	-7.7757
3	TB Treatment Success (0-100%)	-1.9762	-0.0671	-1.9091
4	Receiving TB Prophylaxis in SoC (100-0%)	-1.7739	0.0352	-1.8091

- Identify the order of parameters depending on the absolute values of differences

```
307 # get the order of the parameters depending on the absolute values
308 param_order <- tornado %>%
309   # set the order of the parameters depending on the absolute values
310   arrange(desc(dif)) %>%
311   mutate(parameter = factor(x = parameter, levels = parameter)) %>%
312   select(parameter) %>%
313   # change the data type from list to vector
314   unlist() %>%
315   levels()
```



5. Tornado plot

- Identify the positions of each apex of each rectangle

```
322 # read and modify data
323 tornado_plot <- tornado %>%
324   # change the structure of the datatable
325   gather(key = "min_max",value = "value",max,min) %>%
```

Separate one parameter into two boxes

parameter	max	min
Active TB Prevalence (0-80%)	-11.9092	-0.814
Chronic AIDS Mortality Misattribution (0-75%)	-9.4012	-1.6255
TB Treatment Success (0-100%)	-1.9762	-0.0671
Receiving TB Prophylaxis in SoC (100-0%)	-1.7739	0.0352
30-day (acute) TB mortality (0-50%)	-2.3272	-0.5761
Post-30-day (extended) TB mortality (0-25%)	-3.0225	-1.504
Receiving TB Treatment in SoC (1-100%)	-1.3884	0
TB Treatment Lag in SoC (0-12 mo)	-1.7323	-0.6258
Active TB Incidence (0-25x)	-2.092	-1.5625
Latent TB Prevalence (0-60%)	-1.7641	-1.5715
Initial CD4 (10-50)	-1.4125	-1.3429

	parameter	dif	min_max	value	ymin	ymax	xmin	xmax
1	Active TB Prevalence (0-80%)	-11.0952	max	-11.9092	-11.9092	-1.6000	10.525	11.475
2	Chronic AIDS Mortality Misattribution (0-75%)	-7.7757	max	-9.4012	-9.4012	-1.6000	9.525	10.475
3	TB Treatment Success (0-100%)	-1.9091	max	-1.9762	-1.9762	-1.6000	8.525	9.475
4	Receiving TB Prophylaxis in SoC (100-0%)	-1.8091	max	-1.7739	-1.7739	-1.6000	7.525	8.475
5	30-day (acute) TB mortality (0-50%)	-1.7511	max	-2.3272	-2.3272	-1.6000	6.525	7.475
6	Post-30-day (extended) TB mortality (0-25%)	-1.5185	max	-3.0225	-3.0225	-1.6000	5.525	6.475
7	Receiving TB Treatment in SoC (1-100%)	-1.3884	max	-1.3884	-1.6000	-1.3884	4.525	5.475
8	TB Treatment Lag in SoC (0-12 mo)	-1.1065	max	-1.7323	-1.7323	-1.6000	3.525	4.475
9	Active TB Incidence (0-25x)	-0.5295	max	-2.0920	-2.0920	-1.6000	2.525	3.475
10	Latent TB Prevalence (0-60%)	-0.1926	max	-1.7641	-1.7641	-1.6000	1.525	2.475
11	Initial CD4 (10-50)	-0.0696	max	-1.4125	-1.6000	-1.4125	0.525	1.475
12	Active TB Prevalence (0-80%)	-11.0952	min	-0.8140	-1.6000	-0.8140	10.525	11.475
13	Chronic AIDS Mortality Misattribution (0-75%)	-7.7757	min	-1.6255	-1.6255	-1.6000	9.525	10.475

5. Tornado plot

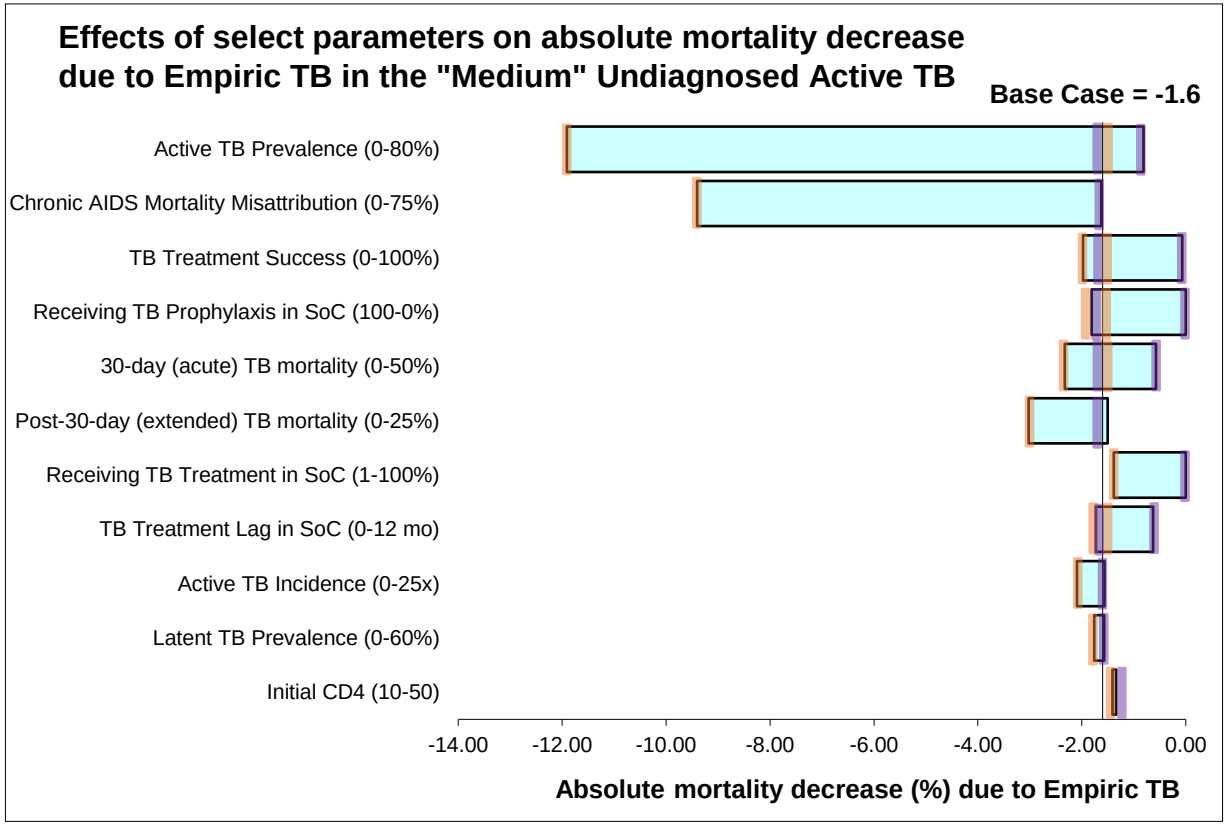
- Identify the positions of each apex of each rectangle

```
326   # calculate the positions of apexes of rectangles
327   mutate(parameter=factor(parameter,levels = param_order),
328          ymin=pmin(value, basecasevalue),
329          ymax=pmax(value, basecasevalue),
330          xmin=as.numeric(parameter)-width/2,
331          xmax=as.numeric(parameter)+width/2)
```

Keep parameter order

Get position of rectangle vertical sides depending on base case values.

Get position of rectangle horizontal sides.



	parameter	dif	min_max	value	ymin	ymax	xmin	xmax
1	Active TB Prevalence (0-80%)	-11.0952	max	-11.9092	-11.9092	-1.6000	10.525	11.475
2	Chronic AIDS Mortality Misattribution (0-75%)	-7.7757	max	-9.4012	-9.4012	-1.6000	9.525	10.475
3	TB Treatment Success (0-100%)	-1.9091	max	-1.9762	-1.9762	-1.6000	8.525	9.475
4	Receiving TB Prophylaxis in SoC (100-0%)	-1.8091	max	-1.7739	-1.7739	-1.6000	7.525	8.475
5	30-day (acute) TB mortality (0-50%)	-1.7511	max	-2.3272	-2.3272	-1.6000	6.525	7.475
6	Post-30-day (extended) TB mortality (0-25%)	-1.5185	max	-3.0225	-3.0225	-1.6000	5.525	6.475
7	Receiving TB Treatment in SoC (1-100%)	-1.3884	max	-1.3884	-1.6000	-1.3884	4.525	5.475
8	TB Treatment Lag in SoC (0-12 mo)	-1.1065	max	-1.7323	-1.7323	-1.6000	3.525	4.475
9	Active TB Incidence (0-25x)	-0.5295	max	-2.0920	-2.0920	-1.6000	2.525	3.475
10	Latent TB Prevalence (0-60%)	-0.1926	max	-1.7641	-1.7641	-1.6000	1.525	2.475
11	Initial CD4 (10-50)	-0.0696	max	-1.4125	-1.6000	-1.4125	0.525	1.475
12	Active TB Prevalence (0-80%)	-11.0952	min	-0.8140	-1.6000	-0.8140	10.525	11.475
13	Chronic AIDS Mortality Misattribution (0-75%)	-7.7757	min	-1.6255	-1.6255	-1.6000	9.525	10.475

5. Tornado plot

- Plot the data: **ggplot()**!
  - Specify the data (“tornado_plot”) and which variable is used for the maximum and minimum values on the x-axis and the y-axis (“xmax”, “xmin”, “ymax”, and “ymin”)
  - Specify the type of graph: **geom_rect()**
  - Specify the color to fill the rectangles and the borders of the rectangles: **scale_fill_manual()** and **scale_color_manual()**

```
337 # plot
338 ggplot(data = tornado_plot,
339        aes(ymax=ymax, # set each apex
340            ymin=ymin,
341            xmax=xmax,
342            xmin=xmin)) +
343   # specify the type of your graph and fill and border color
344   geom_rect(aes(fill="", color="")) +
345   scale_fill_manual(values = "lightblue") + # color inside the rect
346   scale_color_manual(values = "black") + # color of the rect borders
```

5. Tornado plot

- Add a vertical line to show the base case result with an annotation:

geom_segment()

```
366   # add vertical line to show the baseline result
367   geom_segment(aes(x=max(xmax) + 0.5,
368                    xend=0.2,
369                    y=basecasevalue,
370                    yend=basecasevalue),
371                linetype = 2,  # dashed line
372                linewidth = 1) +
373   # add annotation of the vertical line
374   annotate("text", # set annotation type
               x = max(tornado_plot$xmax) + 0.7, y = basecasevalue, #
               label = paste0("Base Case = ", basecasevalue)) +
```

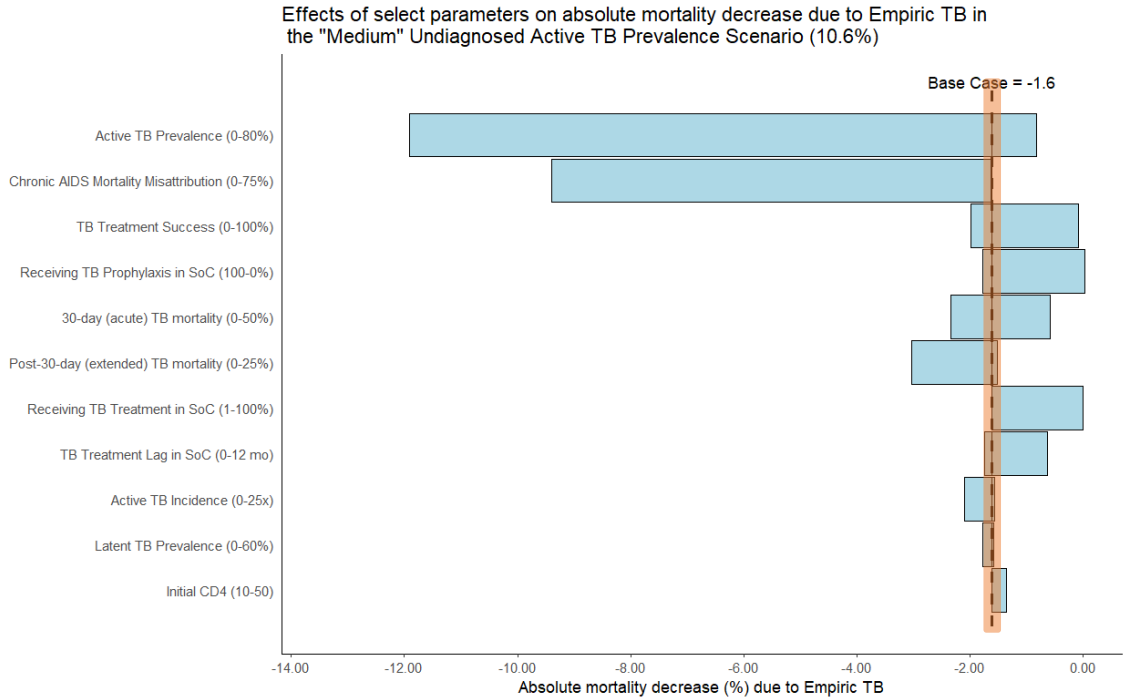
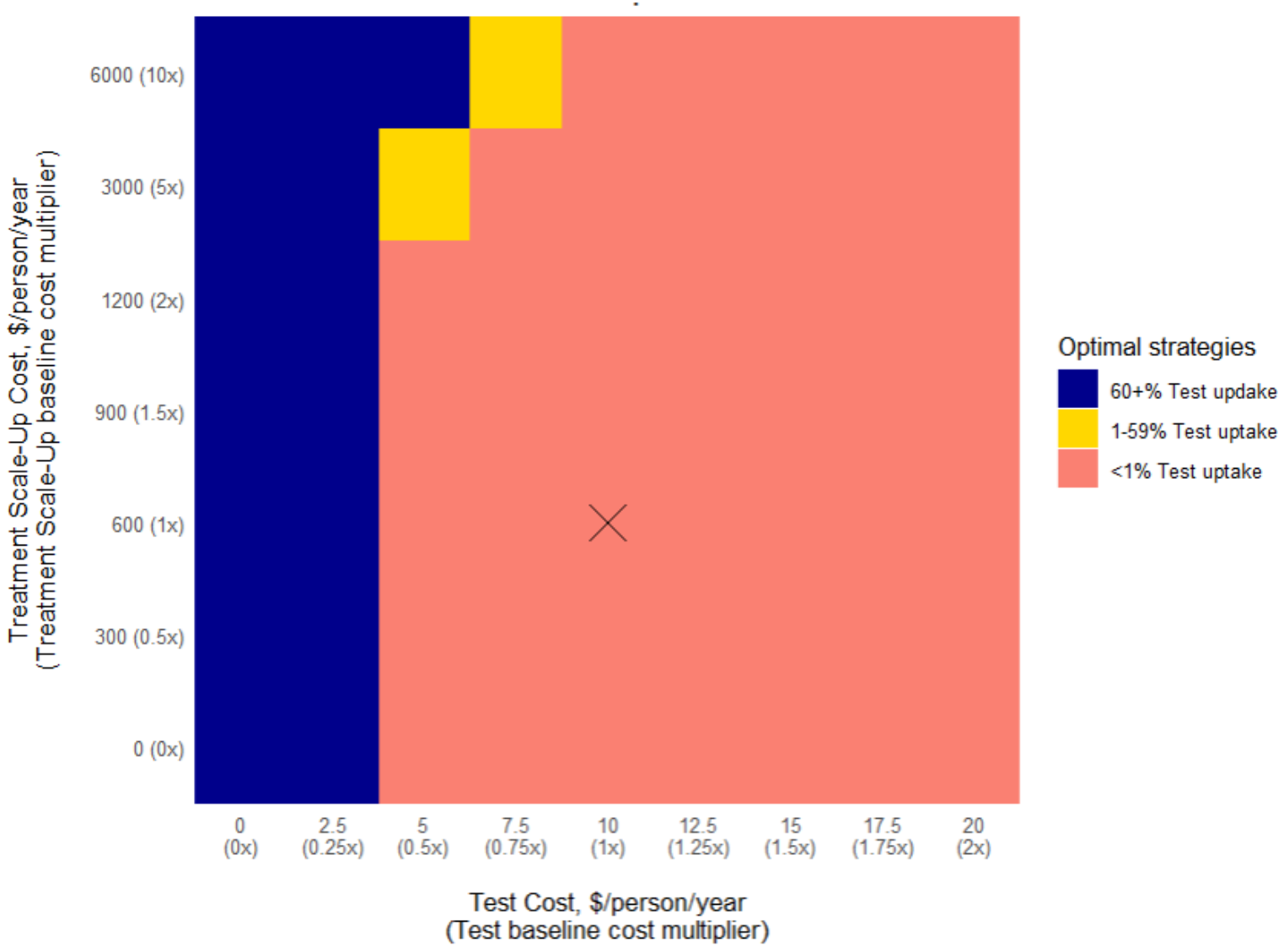


Table of contents

6. Heatmap

- Aim: To visualize the most cost-effective strategies depending on variation of two conditions

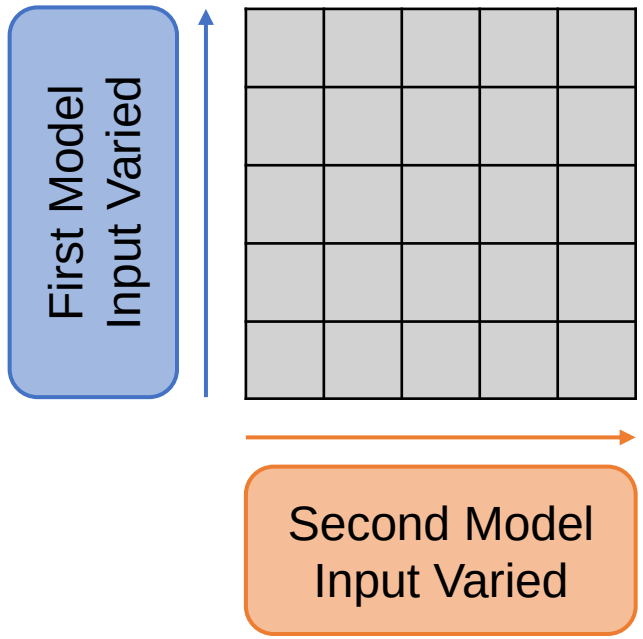


6. Heatmap

- Needed data:
  - Optimal strategies identified by each permutation of the variations of two conditions
  - (Base case conditions and result)

```
26  heat <- read.csv("2.heatmap_data.csv")
```

	Tx_cost_times	Test_cost_times	Optimal_strategy
1	0	0	60+% Test updake
2	0	0.25	60+% Test updake
3	0	0.5	<1% Test uptake
4	0	0.75	<1% Test uptake
5	0	1	<1% Test uptake
6	0	1.25	<1% Test uptake
7	0	1.5	<1% Test uptake
8	0	1.75	<1% Test uptake
9	0	2	<1% Test uptake
10	0.5	0	60+% Test updake



6. Heatmap

- Modifying data for plotting (cont'):
  - Use factor levels to plot Test and Treatment scale up cost in ascending order
  - Create x and y axis labels

```
393 # create "levels" for the x- and y-axis to show the result in ascending order
394 heat <- heat %>%
395   mutate(Tx_cost_times = factor(Tx_cost_times, levels = unique(heat$Tx_cost_times)),
396          Test_cost_times = factor(Test_cost_times, levels = unique(heat$Test_cost_times)))
397
398 # create x axis labels
399 heat$x_labels = paste0(
400   as.numeric(as.character(heat$Test_cost_times))*Test_cost_base,
401   "\n(",heat$Test_cost_times, "x)")
402
403 # create y axis labels
404 heat$y_labels = paste0(
405   as.numeric(as.character(heat$Tx_cost_times))*Tx_cost_base,
406   " (",heat$Tx_cost_times, "x)")
```

6. Heatmap

	Tx_cost_times	Test_cost_times	Optimal_strategy	x_labels	y_labels
1	0	0	60+% Test updake	0 (0x)	0 (0x)
2	0	0.25	60+% Test updake	2.5 (0.25x)	0 (0x)
3	0	0.5	<1% Test uptake	5 (0.5x)	0 (0x)
4	0	0.75	<1% Test uptake	7.5 (0.75x)	0 (0x)
5	0	1	<1% Test uptake	10 (1x)	0 (0x)
6	0	1.25	<1% Test uptake	12.5 (1.25x)	0 (0x)
7	0	1.5	<1% Test uptake	15 (1.5x)	0 (0x)
8	0	1.75	<1% Test uptake	17.5 (1.75x)	0 (0x)
9	0	2	<1% Test uptake	20 (2x)	0 (0x)
10	0.5	0	60+% Test updake	0 (0x)	300 (0.5x)
11	0.5	0.25	60+% Test updake	2.5 (0.25x)	300 (0.5x)
12	0.5	0.5	<1% Test uptake	5 (0.5x)	300 (0.5x)
13	0.5	0.75	<1% Test uptake	7.5 (0.75x)	300 (0.5x)
14	0.5	1	<1% Test uptake	10 (1x)	300 (0.5x)
15	0.5	1.25	<1% Test uptake	12.5 (1.25x)	300 (0.5x)
16	0.5	1.5	<1% Test uptake	15 (1.5x)	300 (0.5x)

6. Heatmap

Specify the data (“heat”) and which variable is used on the x and y axis.

Create heatmap

Draw “X” as base case

Plot title and axes labels

Specify color for each optimal strategy

Format figure

```
# plot
ggplot(data = heat,  # specify the data
       aes(x=Test_cost_times, # specify the x-axis
           y=Tx_cost_times, # specify the y-axis
           fill = Optimal_strategy)) + # specify the color showing the optimal strategy
  geom_tile() +
  # add "X" at the base case
  geom_point(
    data = heat[(heat$Tx_cost_times == 1) & (heat$Test_cost_times == 1),],
    aes(x = Test_cost_times, y = Tx_cost_times), fill = "NA",
    color = "black", size = 7, shape = 4
  ) +
  theme_classic() +
  # add plot title
  ggtitle("Test cost and Treatment scale-up cost") +
  # x axis label
  xlab("Test Cost, $/person/year\n(Test baseline cost multiplier)") +
  # y axis label
  ylab("Treatment Scale-Up Cost, $/person/year\n(Treatment Scale-Up baseline cost multiplier)") +
  # set strategy colors, reverse legend order, name legend "Optimal strategies"
  scale_fill_manual(values = c("salmon","gold","darkblue"),
                    guide = guide_legend(reverse = TRUE),
                    name = "Optimal strategies") +
  # adjust distance between x axis labels and plot, add x axis tick labels
  scale_x_discrete(expand = c(0,0),
                   labels = unique(heat$x_labels)) +
  # adjust distance between y axis labels and plot, add y axis tick labels
  scale_y_discrete(expand = c(0,0),
                   labels = unique(heat$y_labels)) +
  # remove axis ticks and lines, add margin between axes and axes titles
  theme(axis.ticks = element_blank(),
        axis.line = element_blank(),
        axis.title.y = element_text(margin = unit(c(0,5,0,0), "mm")),
        axis.title.x = element_text(margin = unit(c(5,0,0,0), "mm")))
```

Resources

- Data Visualization Cheatsheet
- Cheatsheet Updates
- R Color Cheatsheet
- ggplot2 Documentation
- Coloring for Colorblindness (davidmathlogic.com)

Thank you!

Satoshi Koiso

Email: ash.satoshi.koiso@gmail.com

LinkedIn: https://www.linkedin.com/in/satoshi-koiso/