

---

---

ID 529 Final Project

# Data Visualization Group 2

Lydia Lin, Hui Miao, Tina Geng, Majeda Alzaydan,  
Kevin Ma, Eduardo Ortiz-Panozo, Suzanne McCluskey

---

---

# Why Data visualization?

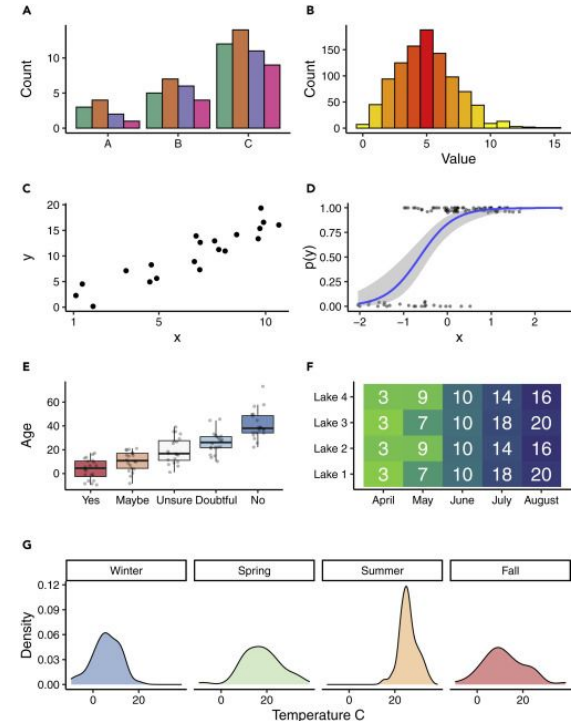
- Hans Rosling: "If your idea is worth spreading, then presentation matters."
- Data presented in a clear visual way goes a long way with driving conversations, providing context, explaining points of view, and getting people on board.
- Visualizations and storytelling complement each other:
  - Visualisation needs to tell a story to the audience.
  - Storytelling helps the viewer gain insight from the data.

# Benefits of good data visualization - Our takes

- Efficiency
- Conciseness
- Improves understanding of results
- Informative
- Greater impact than text alone

# 10 Principles of Effective Data Visualization

1. Diagram First
2. Use right software
3. Use an effective geometry and show data
4. Colors always mean something
5. Include uncertainty
6. Facet/Panel when possible
7. Understand difference of Data and Model
8. Detailed Caption
9. Infographic
10. Get an Opinion



# Special topics in data visualization

- Visualizations for audiences affected by color-blindness
- Size and resolution
- Axis scales and scaling
- Visualizations in grayscale

# Visualizations for audiences affected by color-blindness

- Red & blue – color blind safe colors (also orange & yellow)
- Green - the number one among the colors to avoid for the color blind
- The first rule: **avoid combining red and green**

Normal vision



Protanopia  
Red-Blind



Deuteranopia  
Green-Blind



Trianopia  
Blue-Blind



Normal vision



Protanopia  
Red-Blind



Deuteranopia  
Green-Blind



Trianopia  
Blue-Blind



# Size and resolution

Recommended image files formats to save graph are

- TIFF (Tagged Image File Format)
- EPS (Encapsulated PostScript)
- PDF (Portable Document Format)

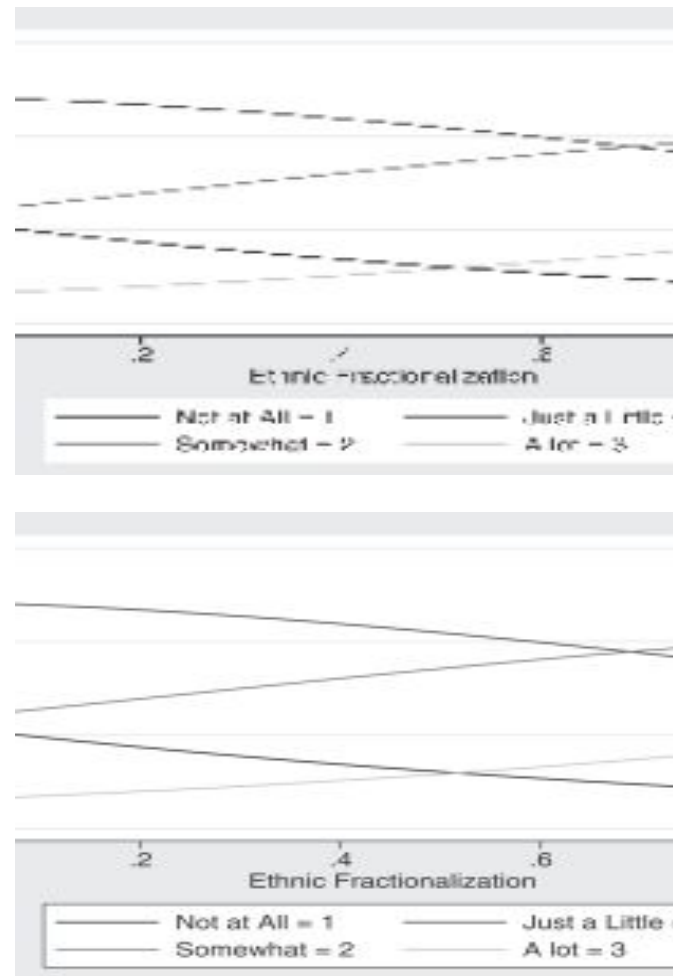
The best option is PDF as it does not required to specify

- height
- width

the benefit of pdf provide high resolution and crisp image.

example:

```
if (FALSE) {ggplot(mtcars, aes(mpg, wt)) + geom_point()
ggsave("mtcars.pdf")}
```

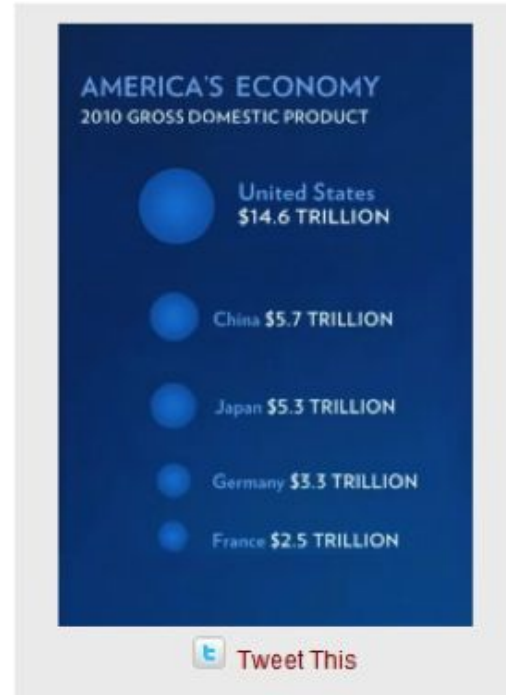
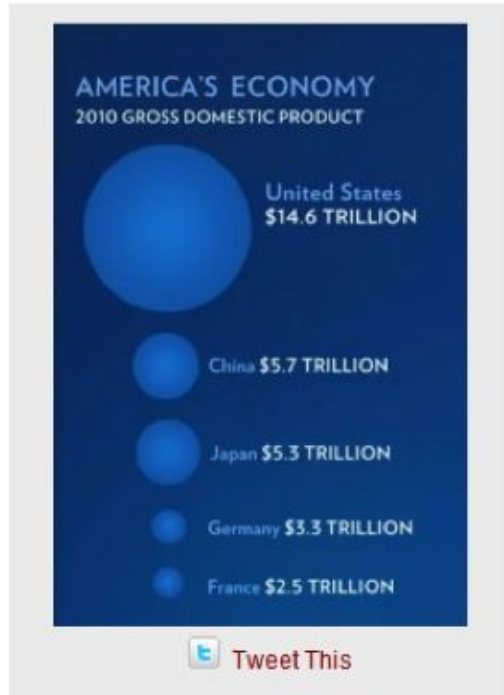


# Axis scales and scaling

- We may want to change default axis scales in plots
  - Log scale
  - Harmonizing a set of plots
- Ommiting baseline values and/or truncating scales may be misleading
- Base R: plot with `xlim()`, `ylim()` and `log=` arguments
- `ggplot2`: `+xlim()`, `+ylim()` functions,  
`+scale_y_continuous(tra`  
`s='log10')`
- `scales` package: tools to override the default break, labels, transformations and palettes



# Axis scales and scaling

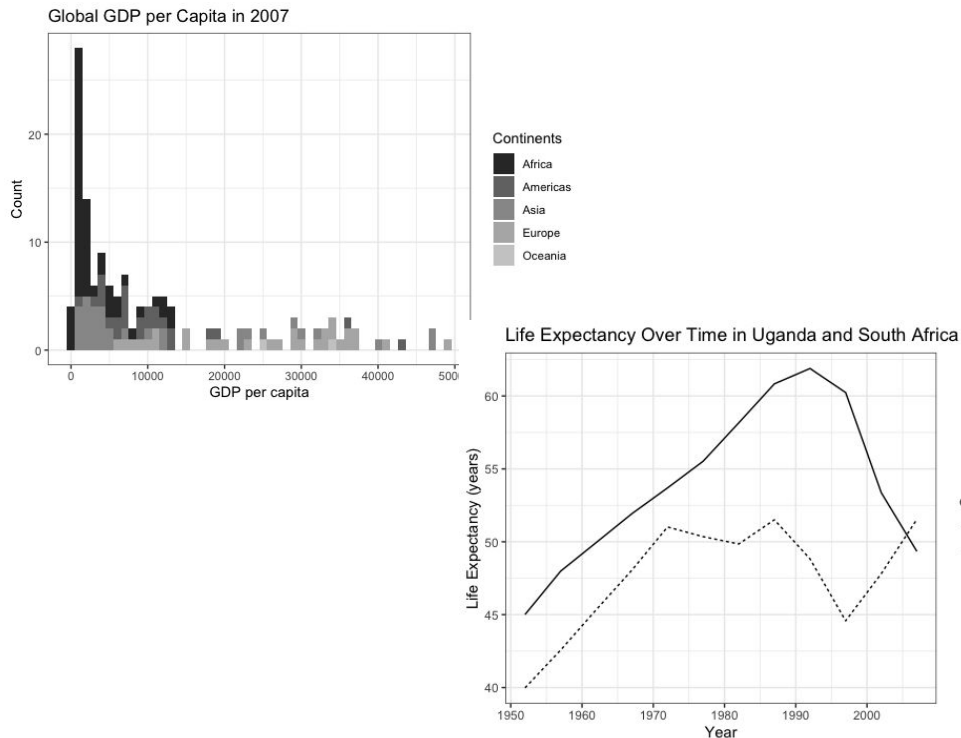


[https://www.data-to-viz.com/caveat/radius\\_or\\_area.html#](https://www.data-to-viz.com/caveat/radius_or_area.html#)

# Visualizations in the grayscale

- Many journals only print in black and white or charge \$\$\$ for color figures
- Grayscale a must if likely that images will be printed or photocopied
- R Tips in ggplot
  - Remove default grey background from ggplot figures (`theme_bw()`)
  - Use of grayscale fill for bar graphs, maps, etc (`scale_fill_grey()`)
  - Use of different line types for line graphs (`linetype`)

# Visualizations in the grayscale



- R Tips in ggplot
  - Remove default grey background from ggplot figures (`theme_bw()`)
  - Use of grayscale fill for bar graphs, maps, etc (`scale_fill_grey()`)
  - Use of different line types for line graphs (`linetype`)

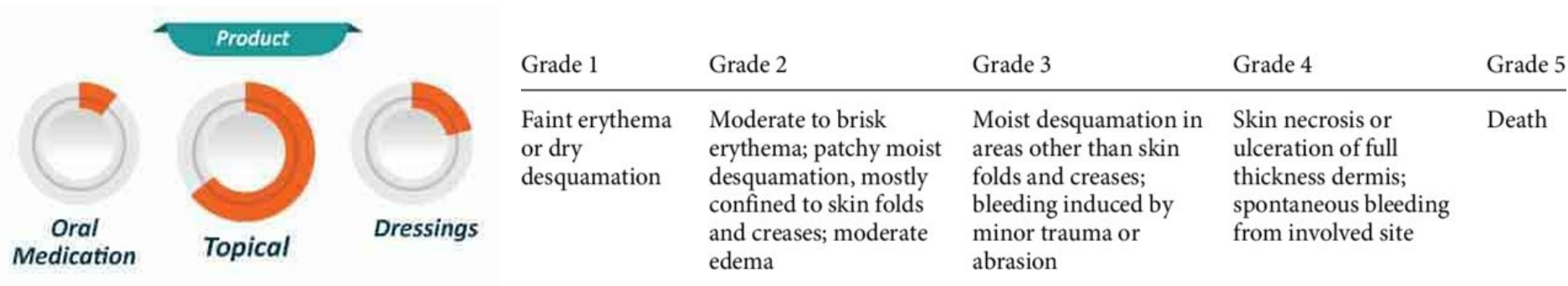
# Data visualization with R – towards the best clinical practice

The most common side effects in head and neck ca survivors are **radiation dermatitis** and oral mucositis.

How to determine the most effective way to **prevent radiation dermatitis in pts w/ head and neck cancer**?

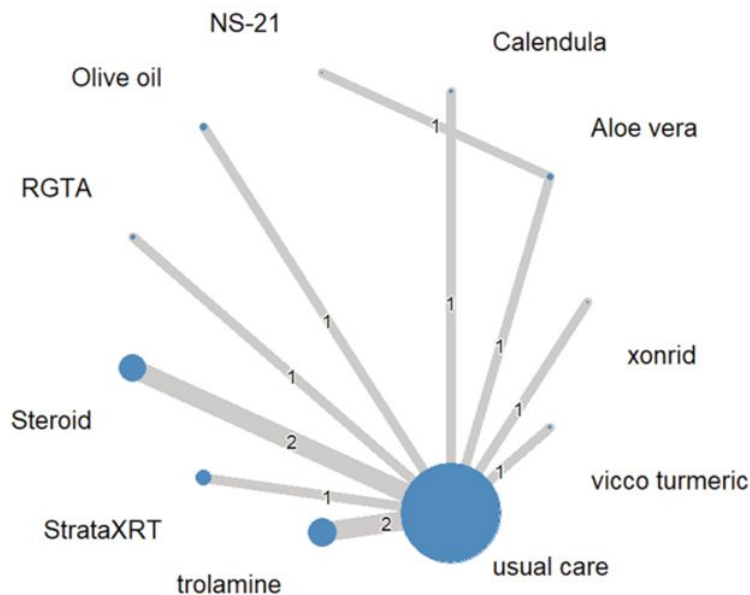
**R package netmeta** for network meta-analysis of RCTs.

**Generalized linear mixed models** were used to conduct network meta-analysis with the restricted maximum likelihood method for estimation.

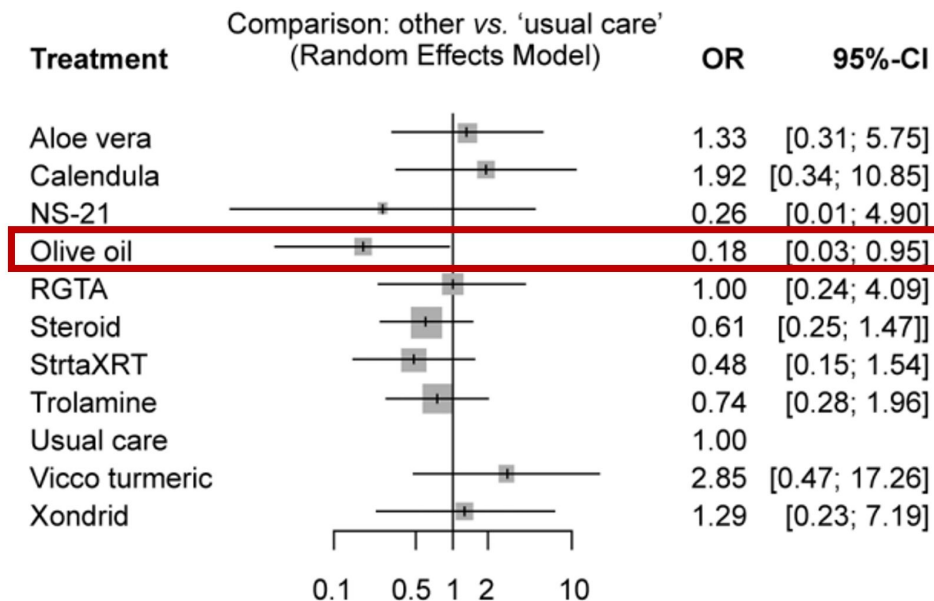


*Common Terminology Criteria for Adverse Events proposed by National Cancer Institute (NCI-CTCAE) criteria*

# Data visualization with R – towards the best clinical practice



Network plot of 12 RCTs included in the network meta-analysis



Forest plot for the network meta-analysis

*Topical use of olive oil before radiotherapy was associated with significantly reduced odds of radiation dermatitis in pts w/ HNC.*

# References

- <https://stackoverflow.com/questions/13501217/ggplot2-for-grayscale-printouts>
- <https://www.datylon.com/blog/data-visualization-for-colorblind-readers>
- [Making High-Resolution Graphics for Academic Publishing | Society for Political Methodology \(polmeth.org\)](#)



**Thank you**

