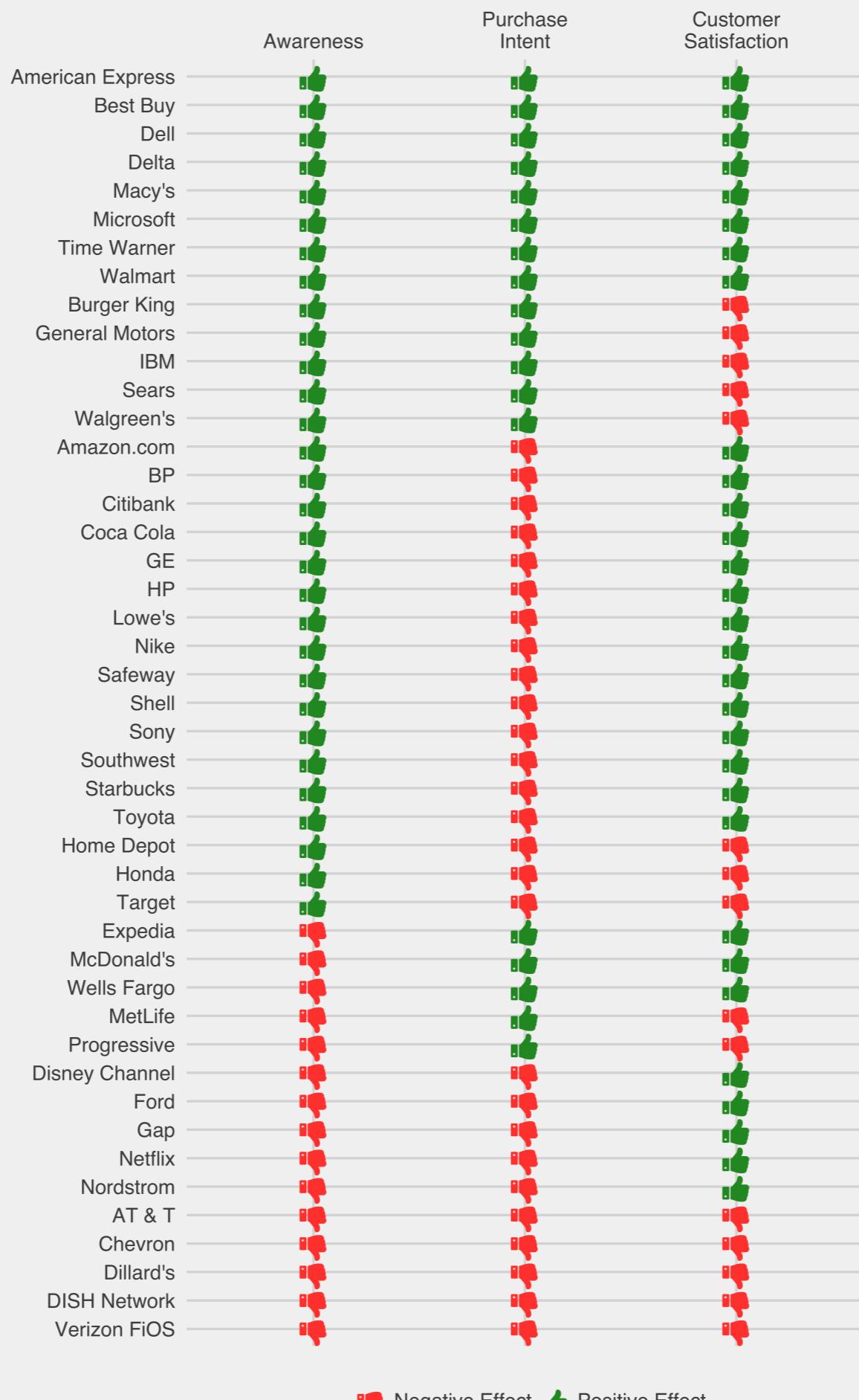


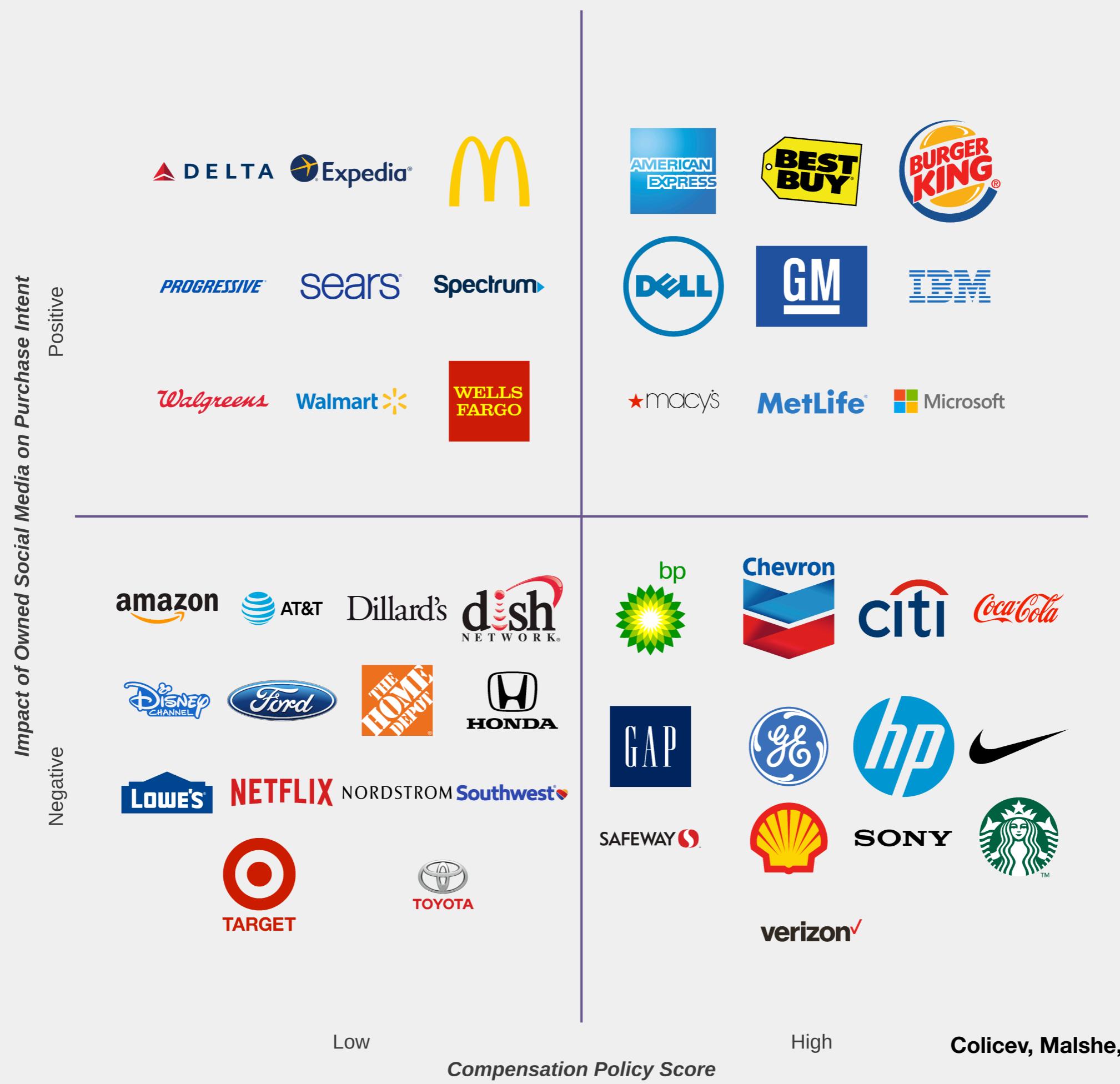
ggplot2

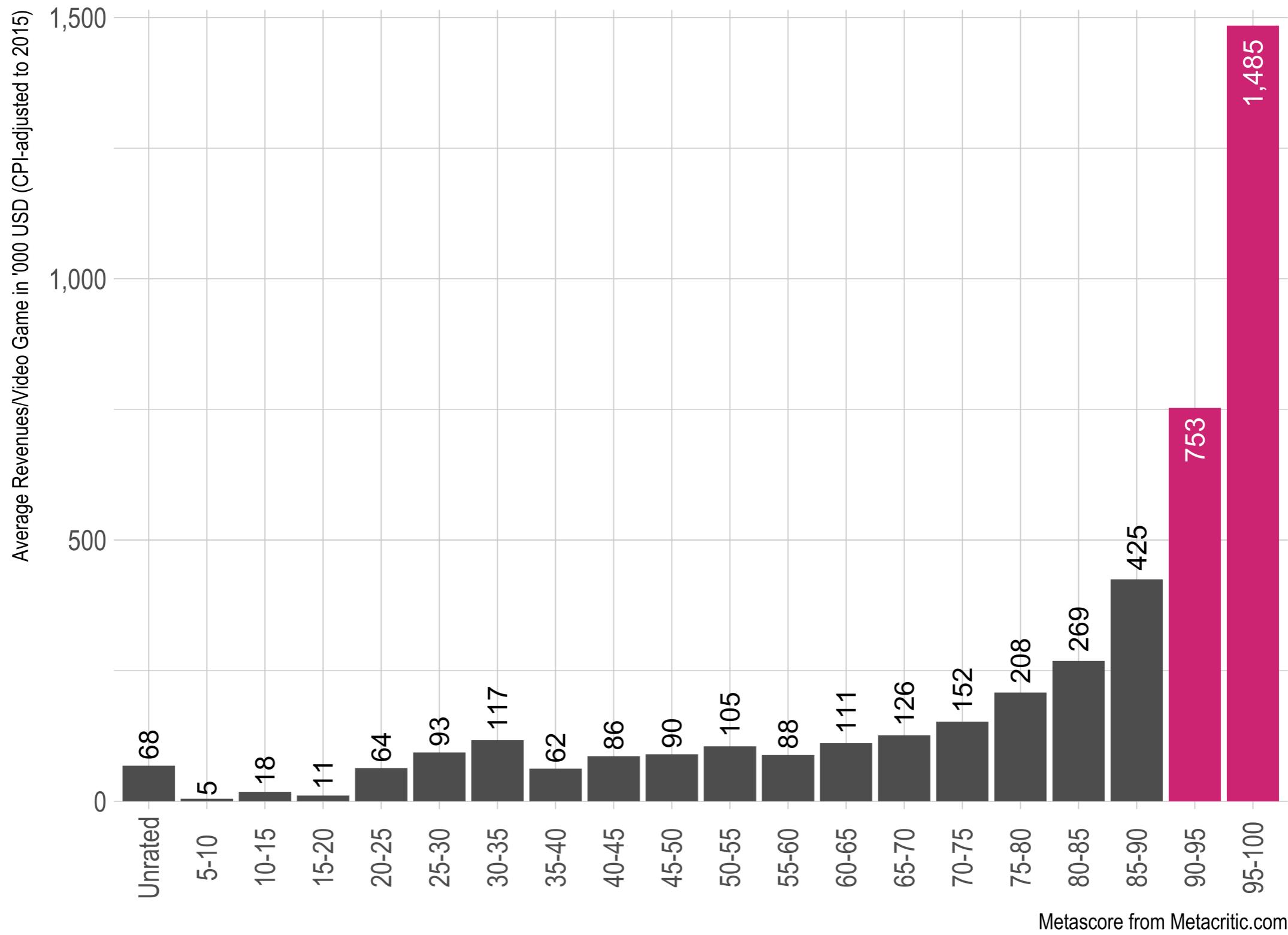
Prof. Ashwin Malshe

A Few Examples from My Work

Effectiveness of Owned Social Media

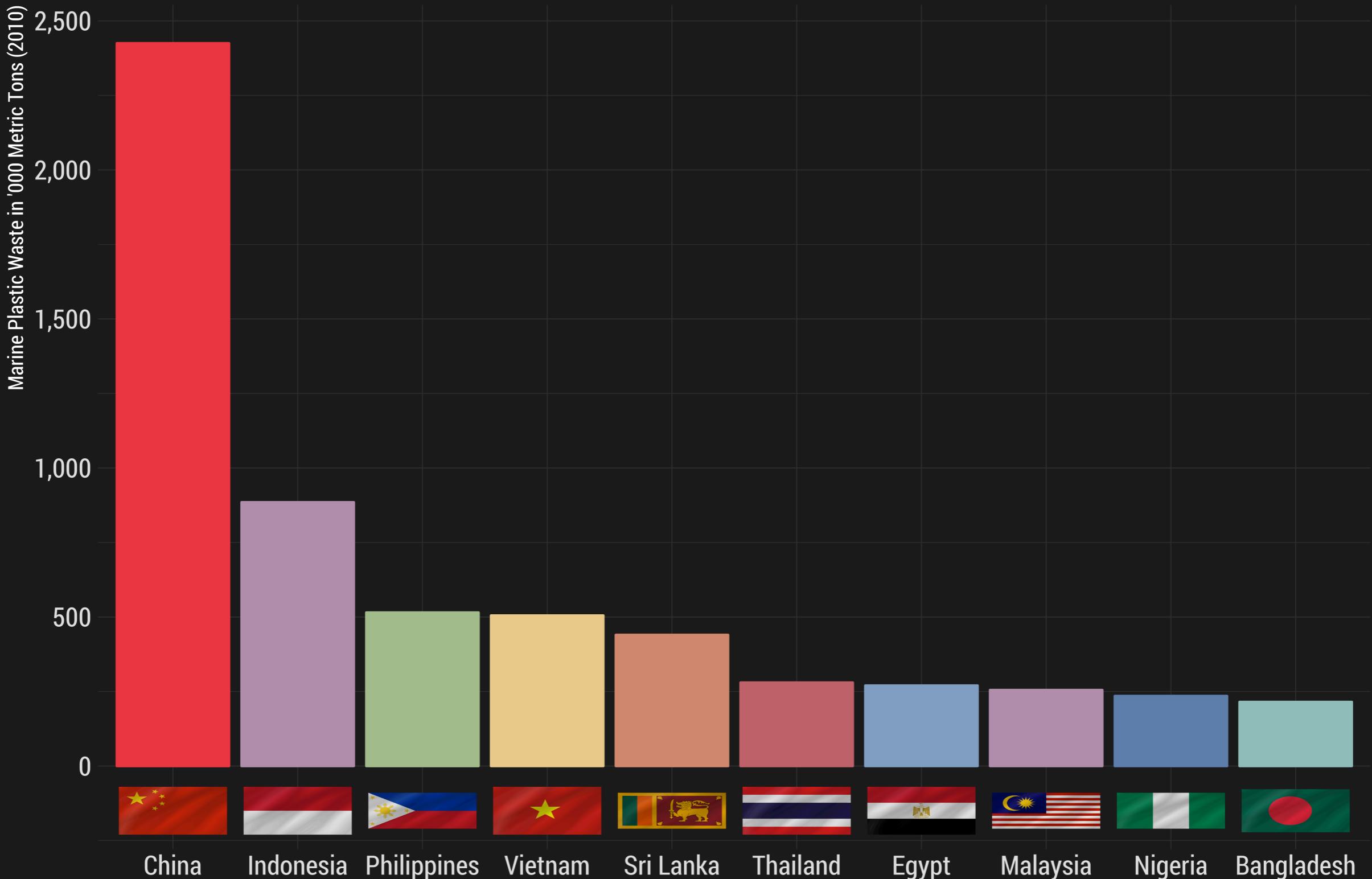






Gretz, Malshe, Bauer, and Basuroy 2019

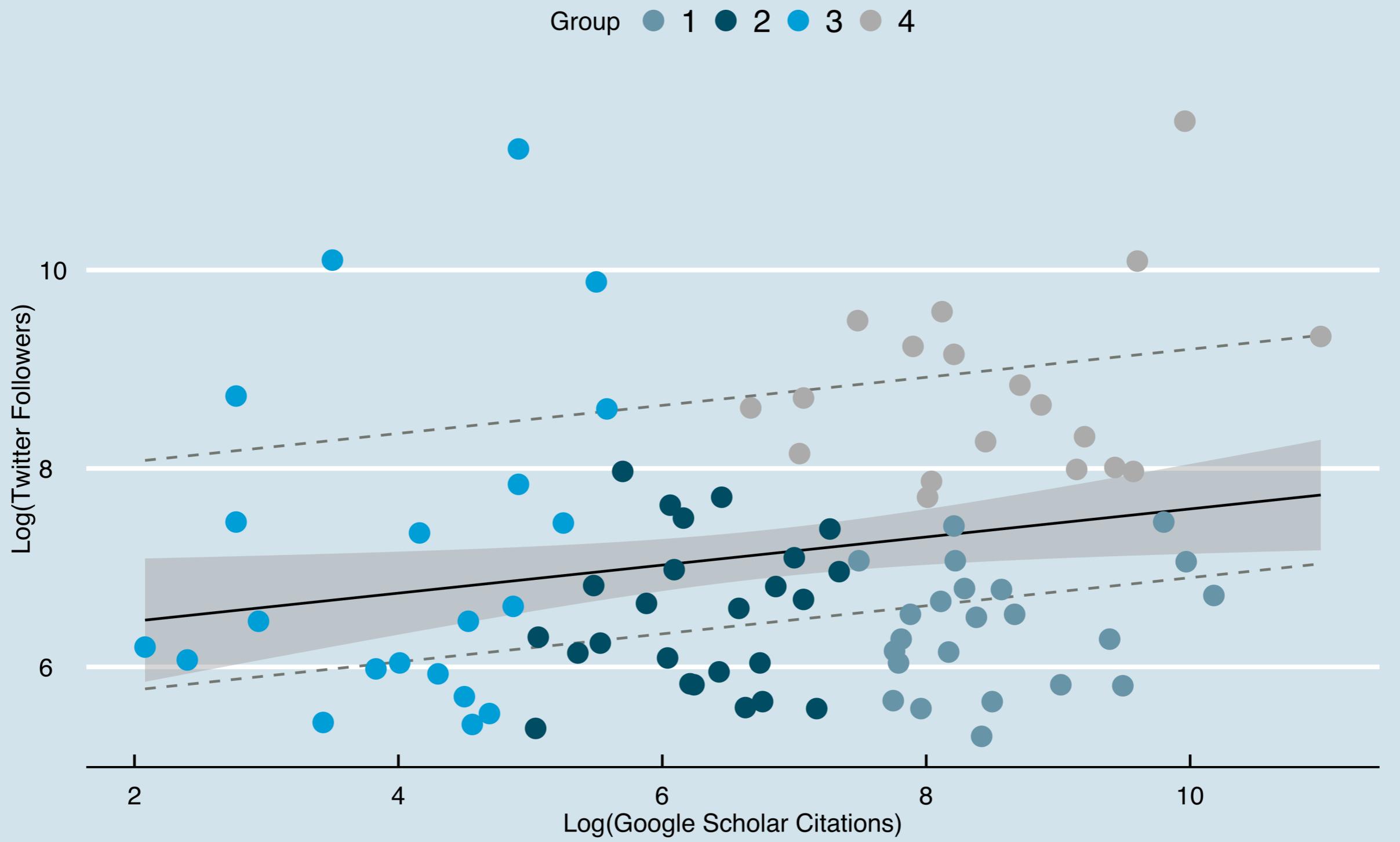
Top 10 Nations to Dump Plastic in the Oceans



© Ashwin Malshe 2019

Source: Jambeck et al. (2015), 'Plastic waste inputs from land into the ocean,' Science

Google Scholar Citations and Twitter Followers



Daily Returns in 2018

BSE returns from close of day_t to open of day_{t+1} ($r_{xy} = 0.61$)



Grammar of Graphics

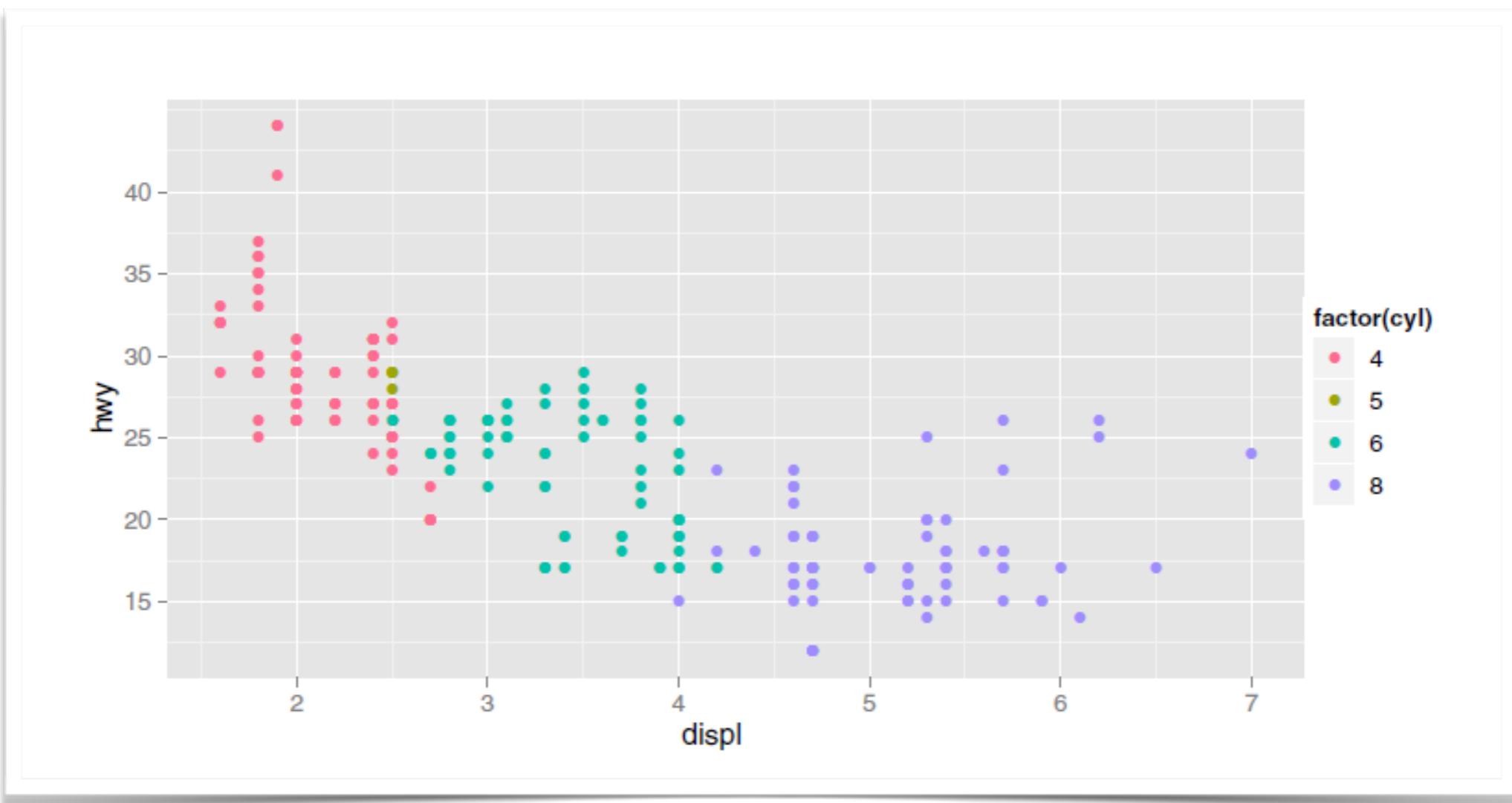
- The layered grammar of graphics tells us that a statistical graphic is a **mapping from data to aesthetic attributes** (color, shape, size) of geometric objects (points, lines, bars).
- The plot may also contain statistical transformations of the data and is drawn on a specific coordinate system.
- Faceting can be used to generate the same plot for different subsets of the dataset.
- It is the combination of these independent components that make up a graphic.

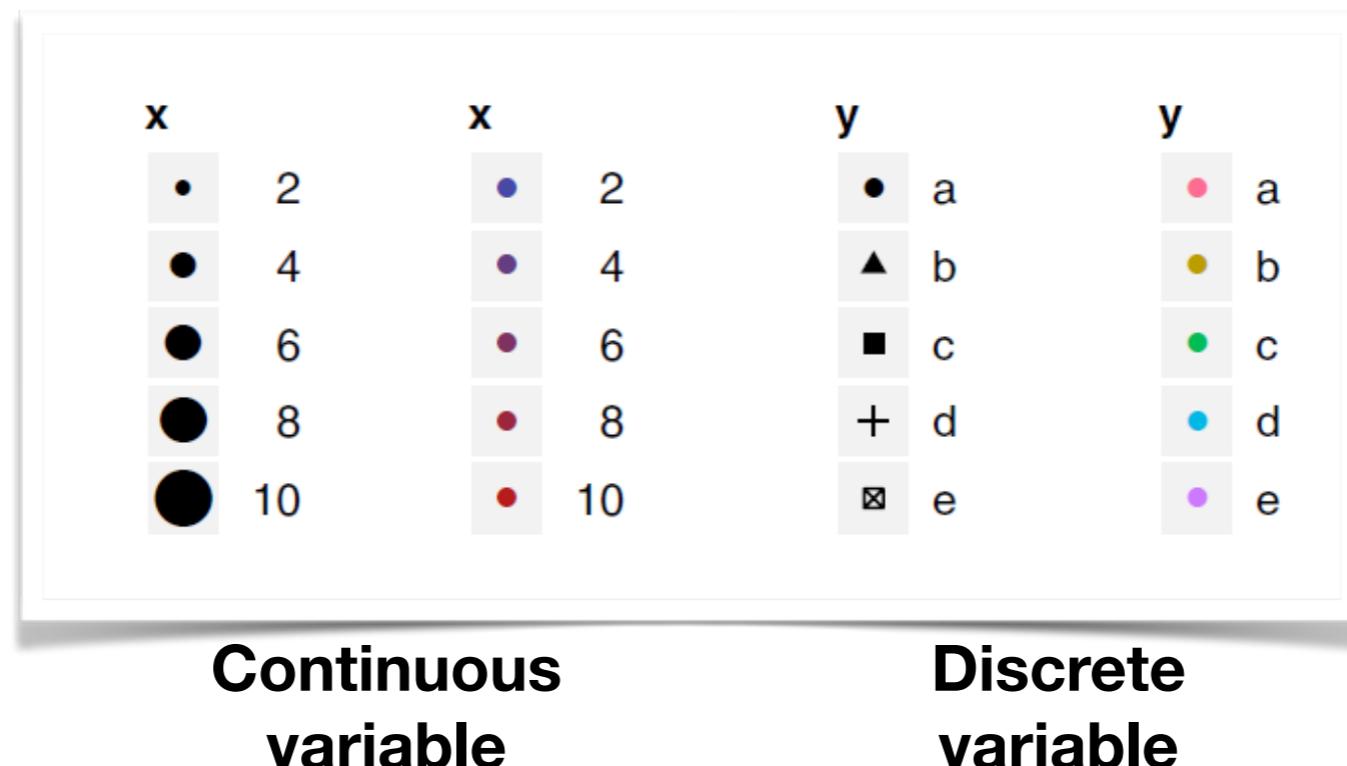
Plots Components

- Data that you want to visualize and a set of aesthetic mappings
- Layers made up of geometric elements and statistical transformation
- The scales map values in the data space to values in an aesthetic space
- A coordinate system, coord for short, describes how data coordinates are mapped to the plane of the graphic
- A faceting specification describes how to break up the data into subsets
- A theme which controls the finer points of display

A quick note on aesthetics mapping

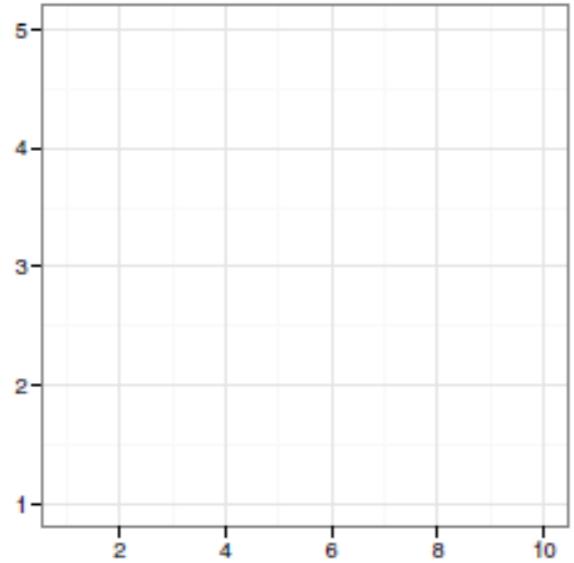
- Aesthetics are the properties that can be perceived on the graphic. Each aesthetic can be mapped to a variable, or set to a constant value.



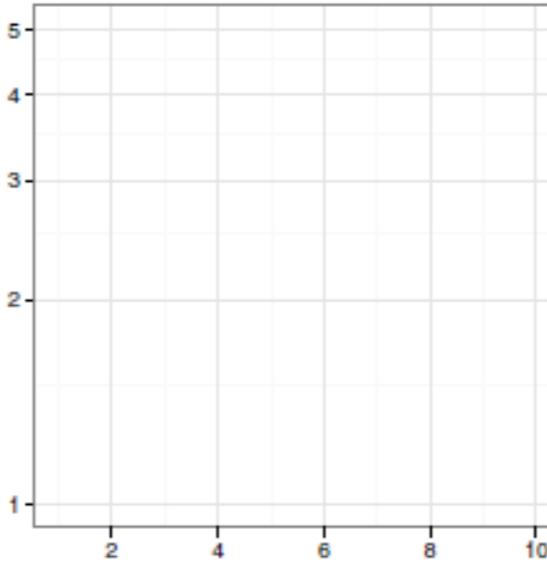


Scales

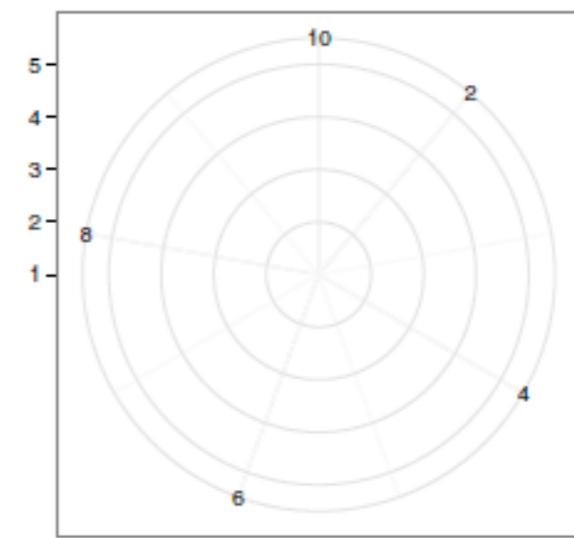
- A **scale** controls the mapping from data to aesthetic attributes, and we need a scale for every aesthetic used on a plot.
- Each scale operates across all the data in the plot, ensuring a consistent mapping from data to aesthetics.



Cartesian



Semi-log



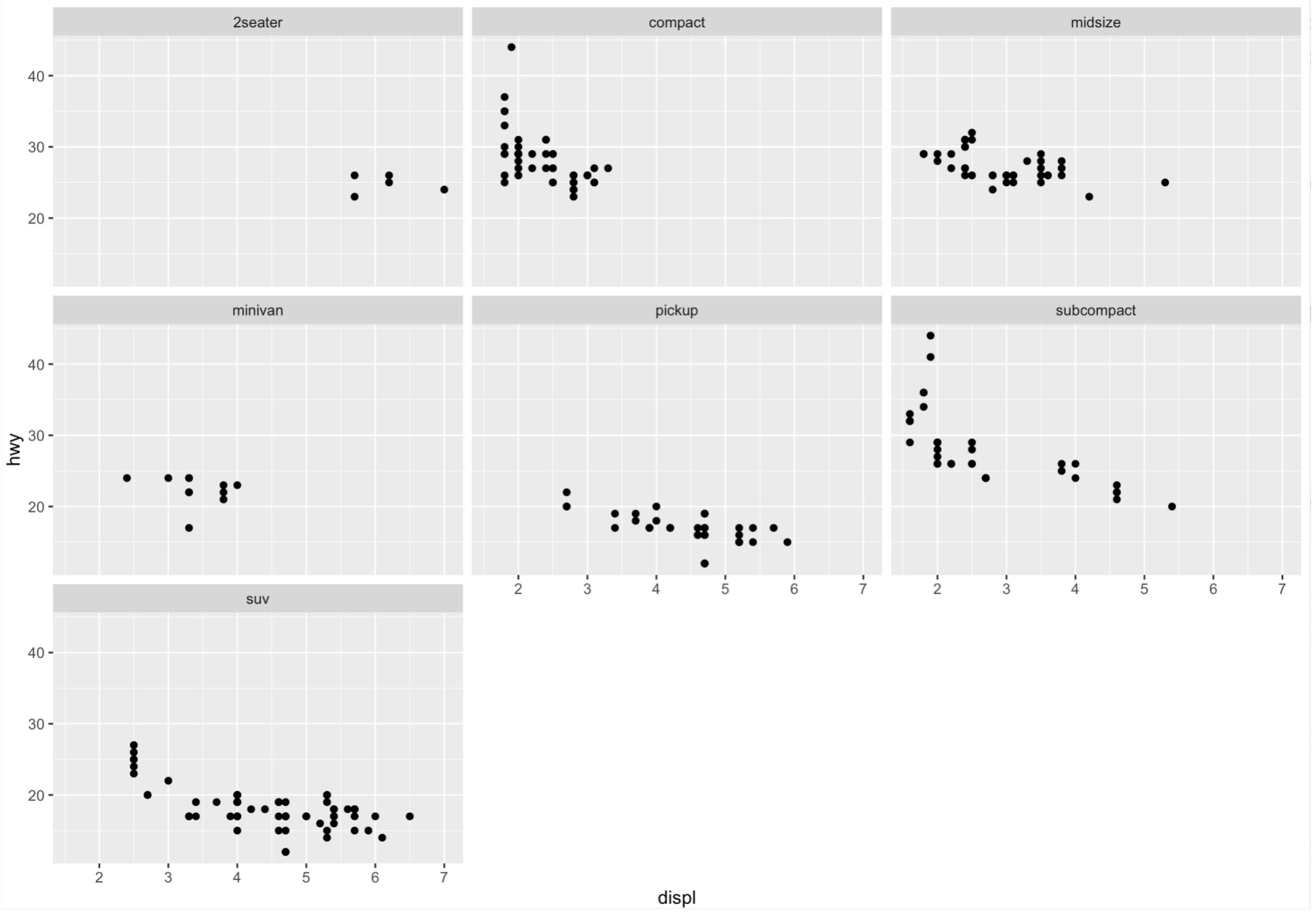
Polar

Coordinate system

- A coordinate system, or **coord** for short, maps the position of objects onto the plane of the plot
- Position is often specified by two coordinates (x, y)
- Coordinate systems control how the axes and grid lines are drawn

Faceting

- This makes it easy to create small multiples each showing a different subset of the whole dataset
- This is a powerful tool when investigating whether patterns hold across all conditions
- The faceting specification describes which variables should be used to split up the data, and whether position scales should be free or constrained



Examples of geoms in ggplot2

Name	Description
abline	Line, specified by slope and intercept
area	Area plots
bar	Bars, rectangles with bases on y-axis
blank	Blank, draws nothing
boxplot	Box-and-whisker plot
contour	Display contours of a 3d surface in 2d
crossbar	Hollow bar with middle indicated by horizontal line
density	Display a smooth density estimate
density_2d	Contours from a 2d density estimate
errorbar	Error bars
histogram	Histogram
hline	Line, horizontal
interval	Base for all interval (range) geoms
jitter	Points, jittered to reduce overplotting
line	Connect observations, in order of x value
linerange	An interval represented by a vertical line
path	Connect observations, in original order
point	Points, as for a scatterplot
pointrange	An interval represented by a vertical line, with a point in the middle
polygon	Polygon, a filled path
quantile	Add quantile lines from a quantile regression
ribbon	Ribbons, y range with continuous x values
rug	Marginal rug plots
segment	Single line segments
smooth	Add a smoothed condition mean
step	Connect observations by stairs
text	Textual annotations
tile	Tile plot as densely as possible, assuming that every tile is the same size
vline	Line, vertical

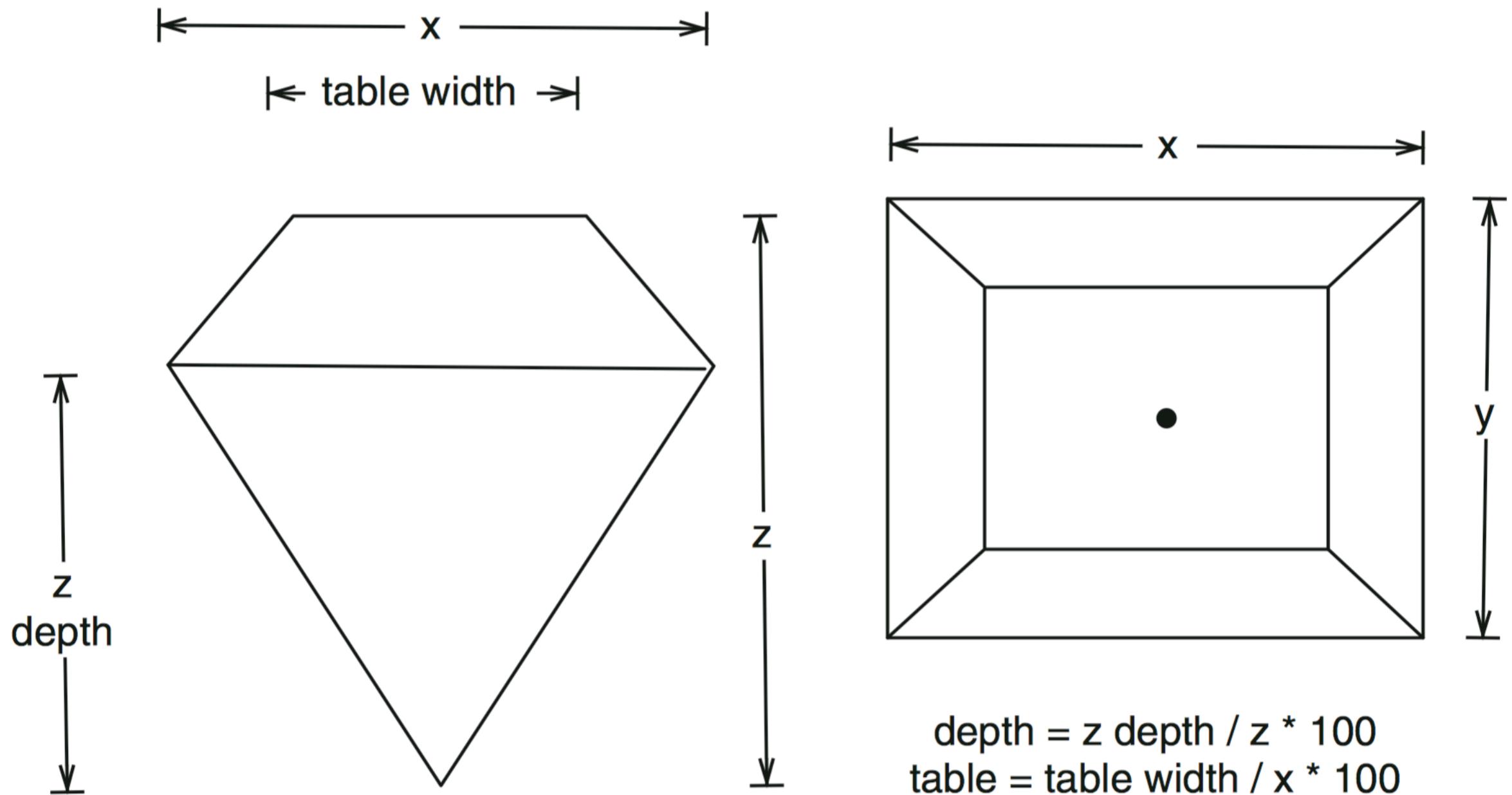


Fig. 3.1 How the variables x, y, z, table and depth are measured