

# Plotting with ggplot2

2020-07-09

# Quick Review: Types of Variables

## Categorical

A **categorical** variable is a variable with a limited number of fixed descriptions; basically a label.

- unordered
  - No natural ordering
  - Ex: sample IDs, genotypes, phenotypes
- ordered
  - Natural way to order them
  - Ex: survey responses (poor, fine, ok, very good, good), chromosomes (chr1, chr2, chr3, etc.)

## Numeric

- discrete
  - Values are indivisible (or dividing them makes no sense); aka count data.
  - Ex: counts of people, read counts
- continuous
  - Values can be divided and expressing them as a divided value, even if the divisions aren't necessary are present, is fine.
  - Ex: height, weight

# ggplot2

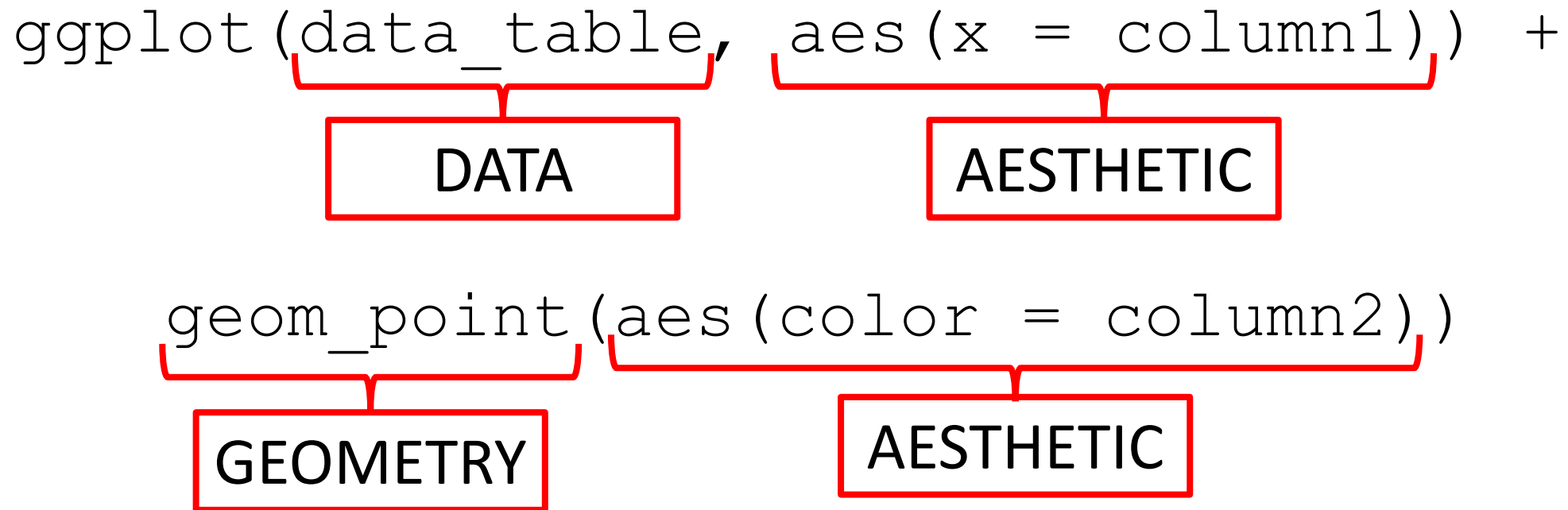


“ggplot2 is a system for declaratively creating graphics, based on [The Grammar of Graphics](#). You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.”

grammar	description
data	The table you want to visualize
geometry	What shape you want to give that visualization, ex: scatter plot, boxplot, violin plot, bar plot, histogram, density plot
aesthetic	The appearance of the geometry, ex: size, shape, color

# The philosophy of `ggplot`

Data, geometry, and aesthetics are **independent**.



DEMO WITH  
RStudio