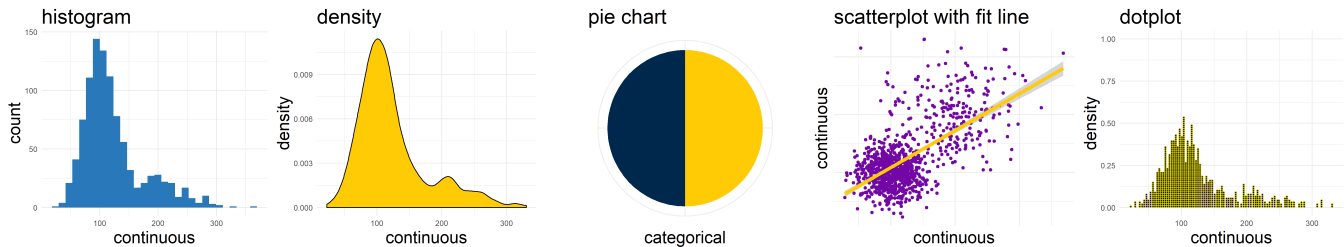


How to Choose a Chart

A Statistically Motivated Guide

Andy Grogan-Kaylor

October 08, 2019



How to Choose a Chart

Choosing the right chart to represent your data can be a daunting process. I believe that a *starting point* for this thinking is some basic statistical thinking about the *type* of variables that you have. At the broadest level, variables may be conceptualized as *categorical* variables, or *continuous* variables.

- *categorical variables* represent unordered categories like *gender*, or *religious affiliation*.
- *continuous variables* represent a continuous scale like a *mental health scale*, or a *measure of neighborhood quality*.

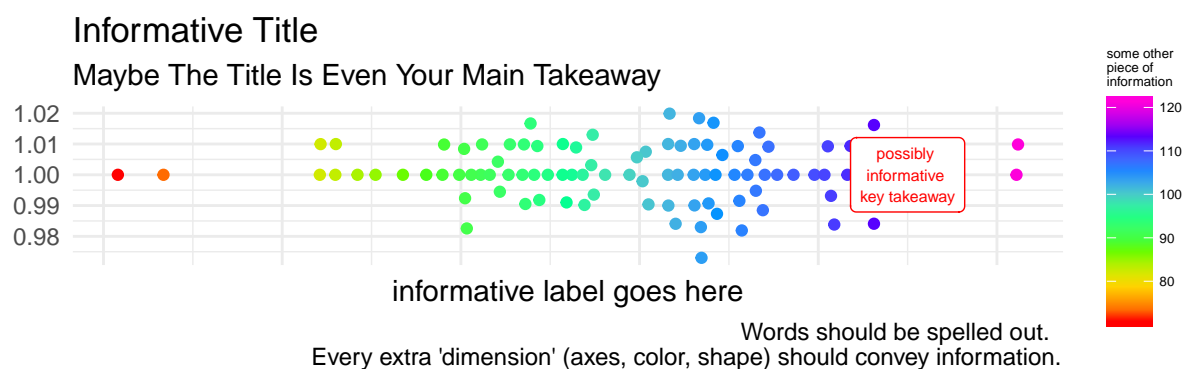
Once we have discerned the type of variable that have, there are two followup questions we may ask before deciding upon a chart strategy:

- Is our graph about **one thing at a time**?
 - How much of x is there?
 - What is the distribution of x ?
- Is our graph about **two things at a time**?
 - What is the relationship of x and y ?
 - How are x and y associated?

A Few Notes

A Note About Graph Labels

Graphs should have clear titles and labels.



A Note About Software

The principles of graphing discussed in this document transcend any particular software package, and could be implemented in many different software packages, such as SPSS, SAS, Stata, or R.

The graphs in these particular examples use ggplot2, a graphing library in R. ggplot2 graph syntax can be formidably complex, with a somewhat steep learning curve. More information about ggplot can be found [here](#).

A Note About Graph Colors

This document uses colors based upon official University of Michigan colors. Using colors that match the design scheme of your organization may be helpful.

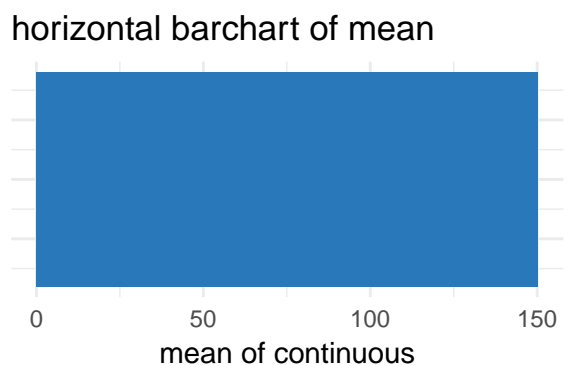
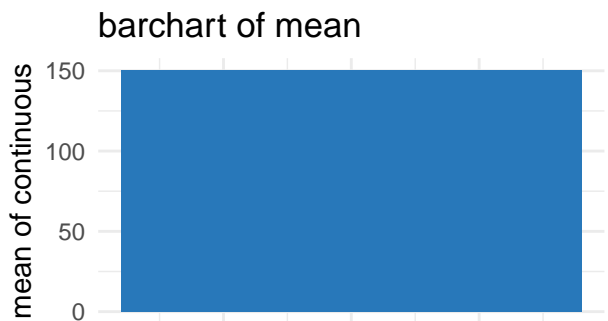
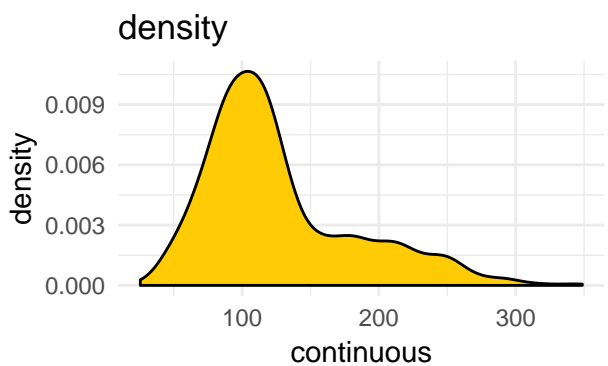
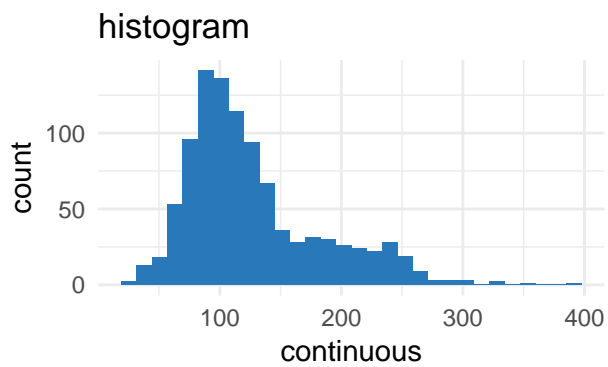
A Simulated Data File of Continuous and Categorical Data

A few randomly selected observations...

	x	y	z	u	v	w	s	q
108	126.4	99.79	100.7	Group B	Group B	Group A	Group 2	146.4
206	173.6	180.4	108.1	Group B	Group B	Group B	Group 3	203.6
956	268.7	164.9	100.9	Group B	Group B	Group B	Group 2	288.7
643	94.19	106.9	112.3	Group A	Group A	Group A	Group 3	124.2
681	143.5	48.01	103.8	Group A	Group A	Group A	Group 2	163.5
730	63.1	99.77	90.77	Group B	Group A	Group A	Group 2	83.1
989	237.4	95.51	99.55	Group B	Group B	Group B	Group 4	277.4
744	174.9	262.6	100.9	Group B	Group A	Group B	Group 2	194.9
125	111.7	93.62	136.1	Group B	Group B	Group A	Group 2	131.7
553	95.56	87.46	121.2	Group A	Group A	Group A	Group 1	105.6

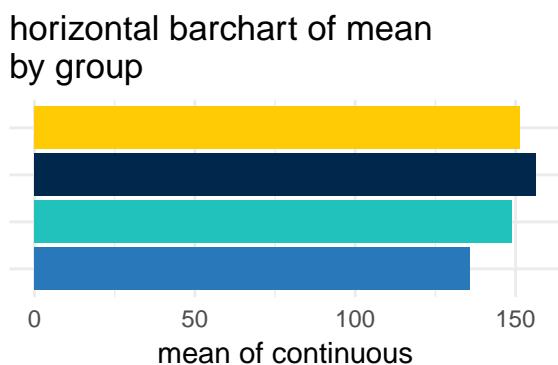
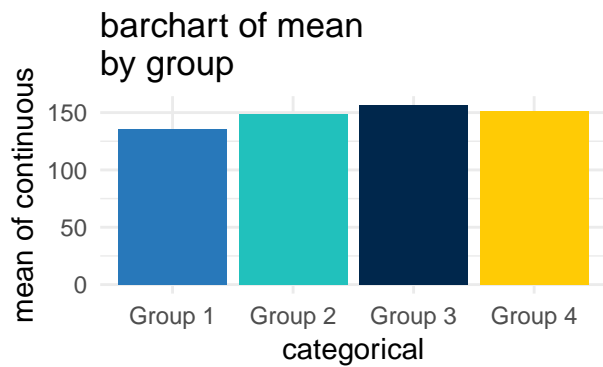
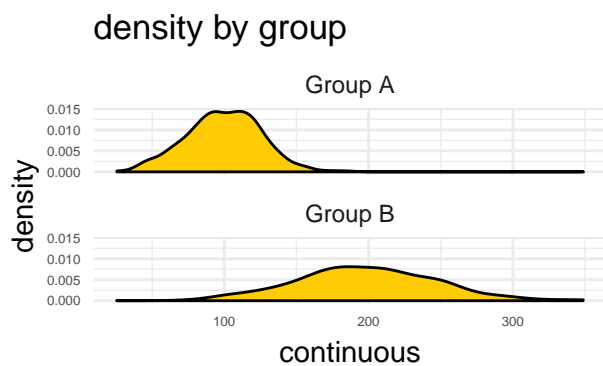
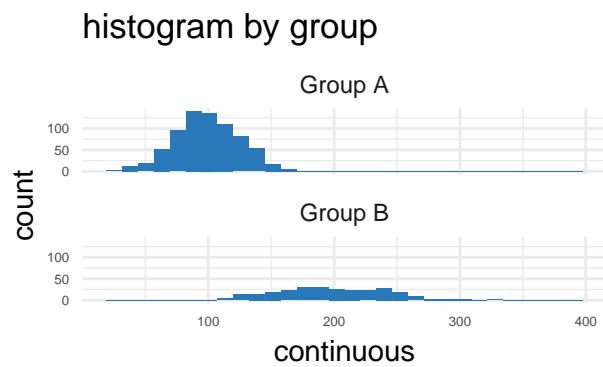
One Thing At A Time

Continuous

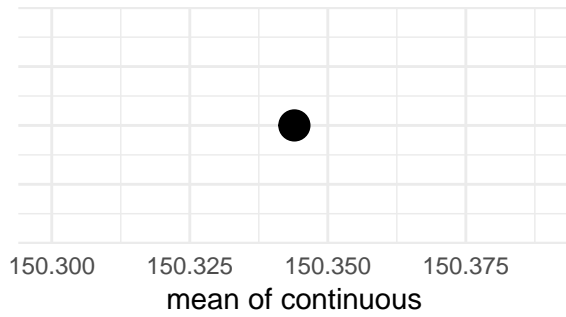


Two Things At A Time

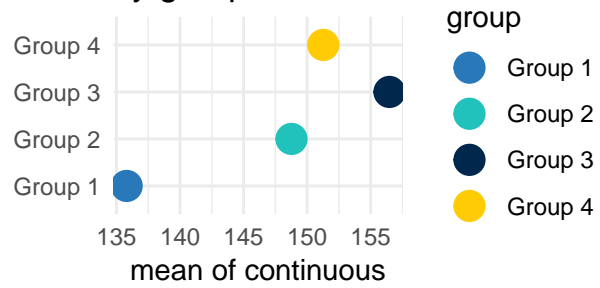
Continuous By Categorical



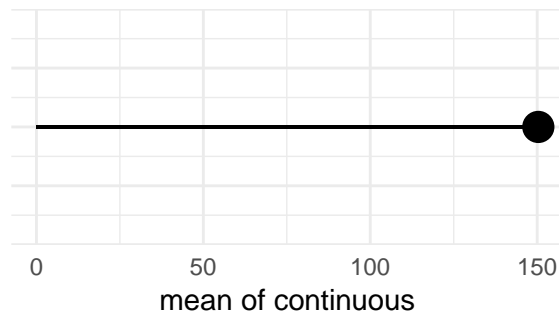
horizontal dotchart of mean



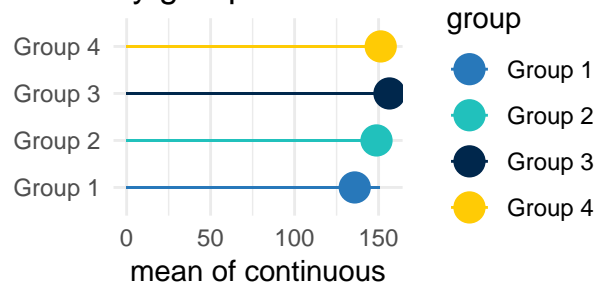
horizontal dotchart of mean by group



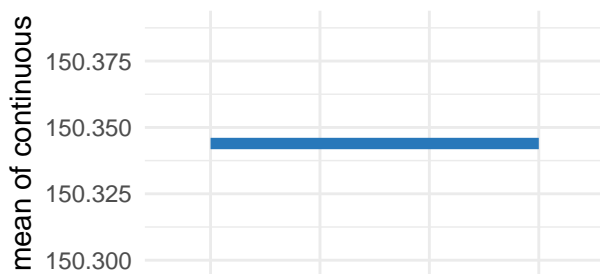
horizontal lollipop chart of mean



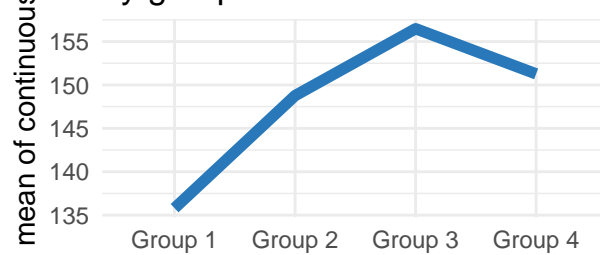
horizontal lollipop chart of mean by group



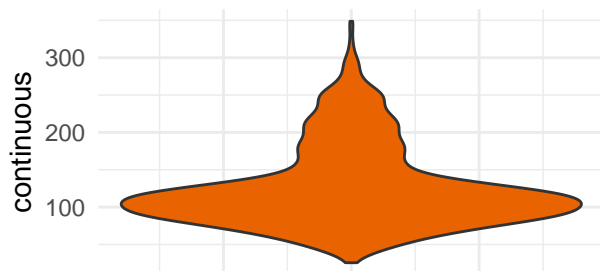
linechart of mean



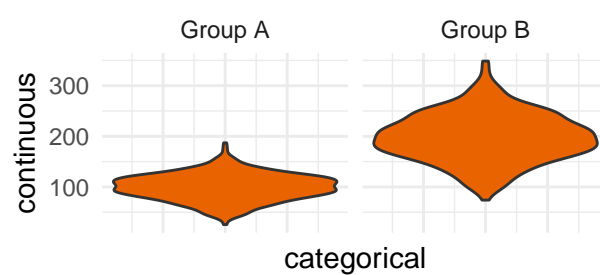
linechart of mean by group

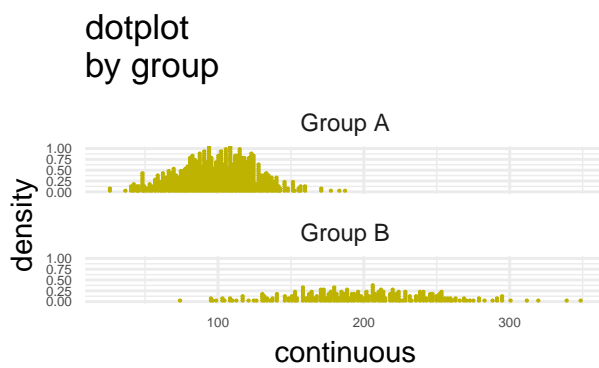
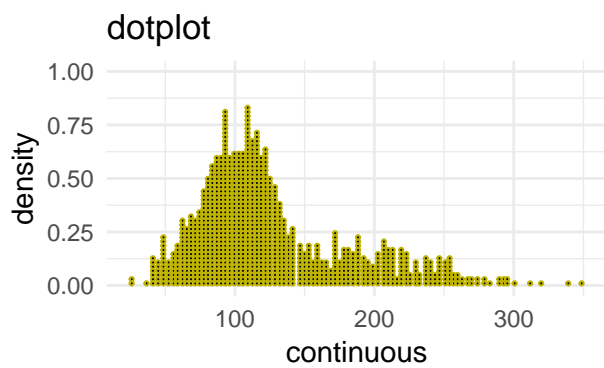
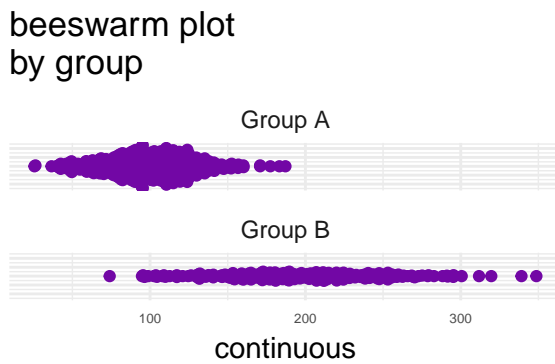
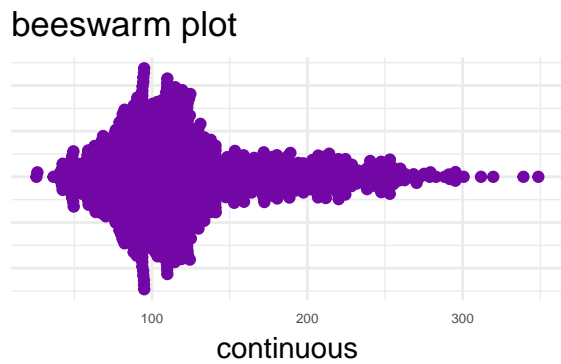
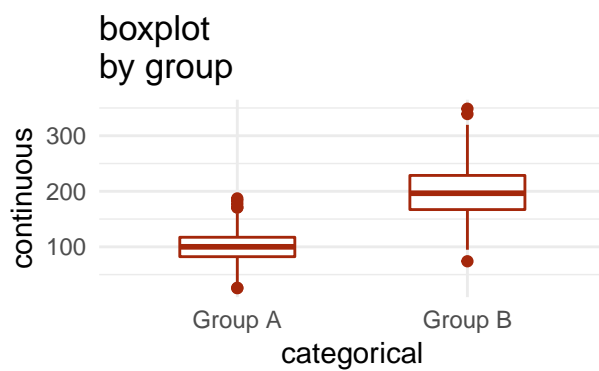
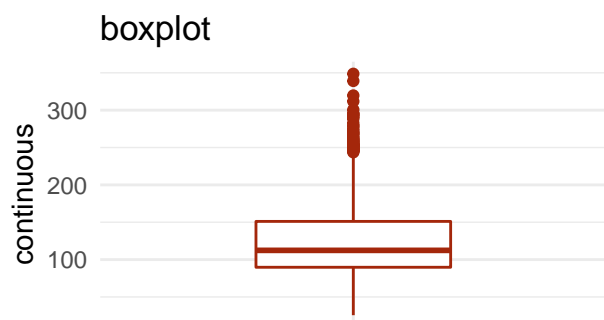


violin plot



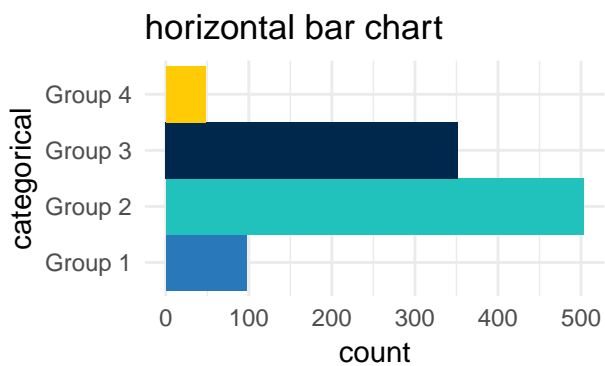
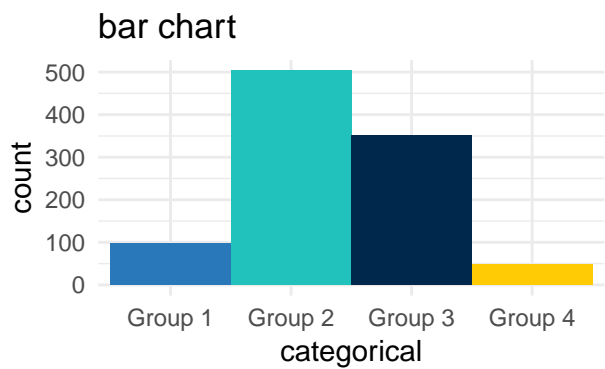
violin plot by group





One Thing At A Time

Categorical



pie chart



categorical

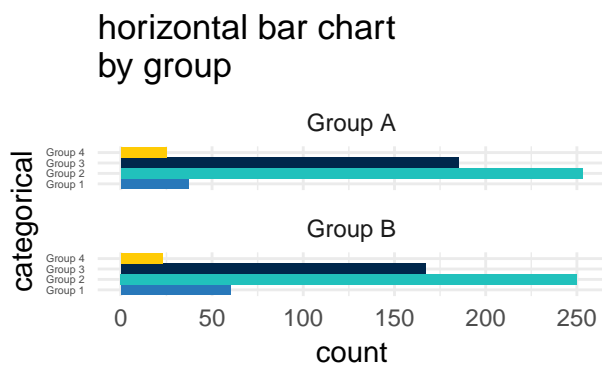
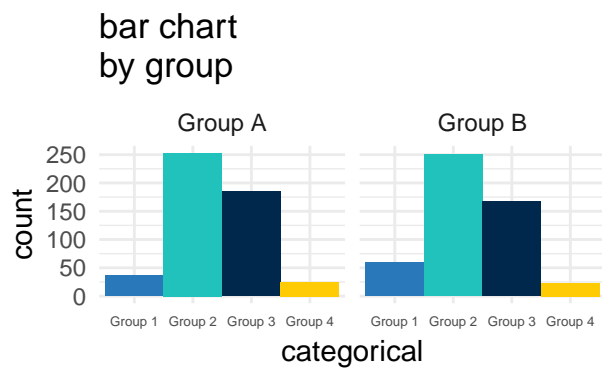
doughnut chart



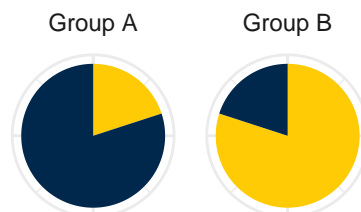
categorical

Two Things At A Time

Categorical By Categorical

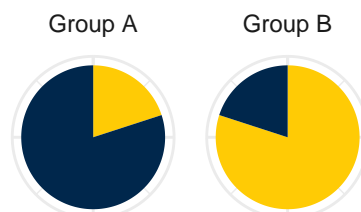


pie chart by group



categorical

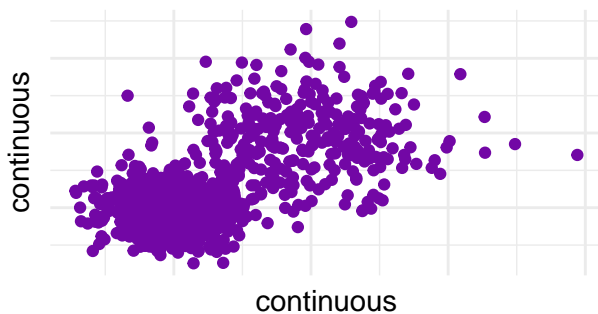
doughnut chart by group



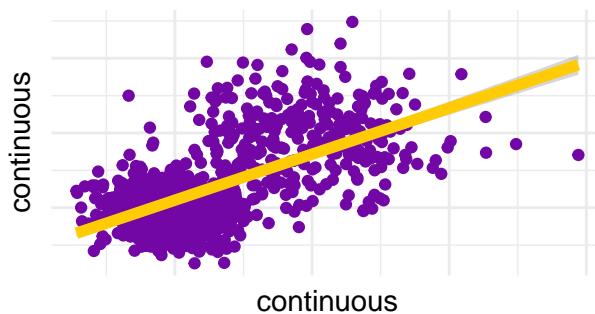
categorical

Continuous by Continuous

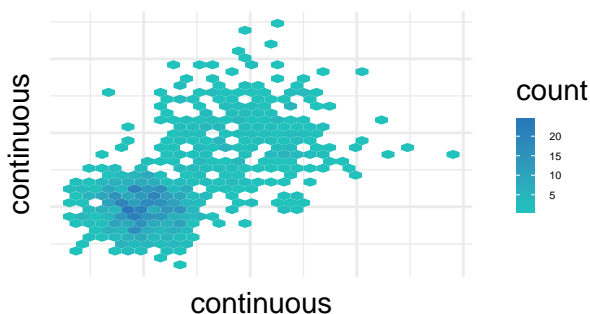
scatterplot



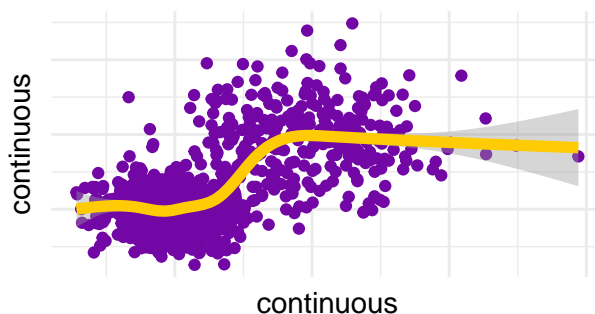
scatterplot with fit line



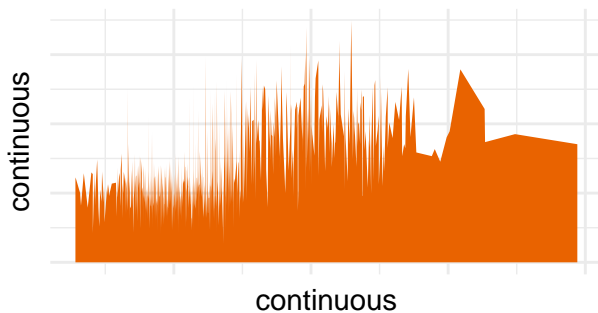
hexagon plot



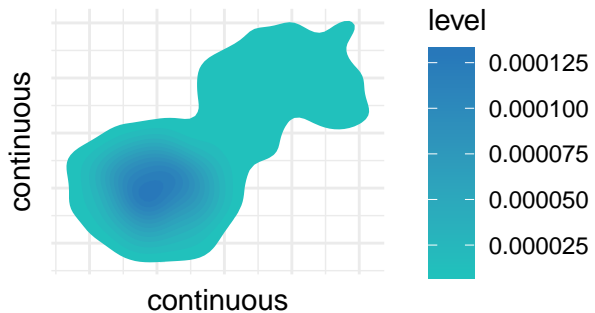
scatterplot with smoother



area plot



contour plot



Graphics made with the ggplot2 graphing library created by Hadley Wickham.

Available online at <https://agrogan1.github.io/>

How to Choose a Chart by Andrew Grogan-Kaylor is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. You are welcome to download and use this handout in your own classes, or work, as long as the handout remains properly attributed.

Last updated: October 08 2019 at 15:41