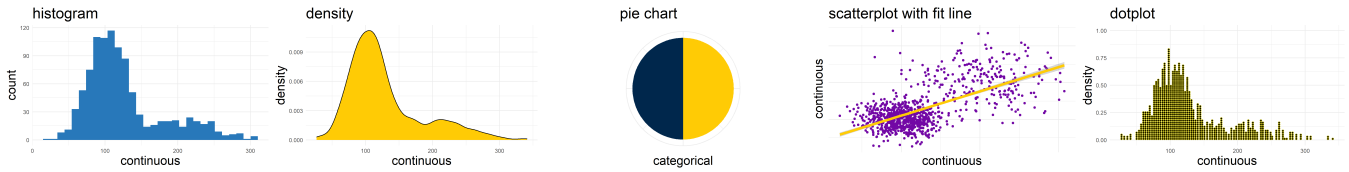
How to Choose a Chart

A Statistically Motivated Guide

Andy Grogan-Kaylor

March 10, 2021



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*.

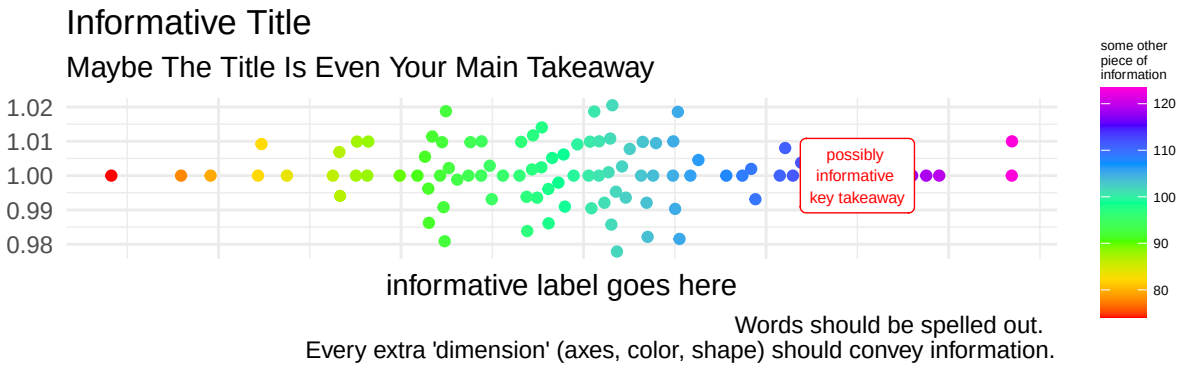
Once we have discerned the type of variable that have, there are two followup questions we may ask before deciding upon a chart strategy:

- 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.



Words should be spelled out.
Every extra 'dimension' (axes, color, shape) should convey information.

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
976	74.33	106.3	95.23	Group B	Group B	Group A	Group 2	94.33
215	218	230.7	71.15	Group B	Group B	Group B	Group 4	258
989	106	105.3	106.6	Group B	Group B	Group A	Group 3	136
325	110.9	113.9	97.18	Group A	Group B	Group A	Group 3	140.9
189	251	149.5	112	Group B	Group B	Group B	Group 2	271
416	80.44	119.3	107.3	Group A	Group A	Group A	Group 3	110.4
117	193.3	163.5	94.79	Group B	Group B	Group B	Group 2	213.3
509	230.9	201.1	92.95	Group A	Group A	Group B	Group 2	250.9

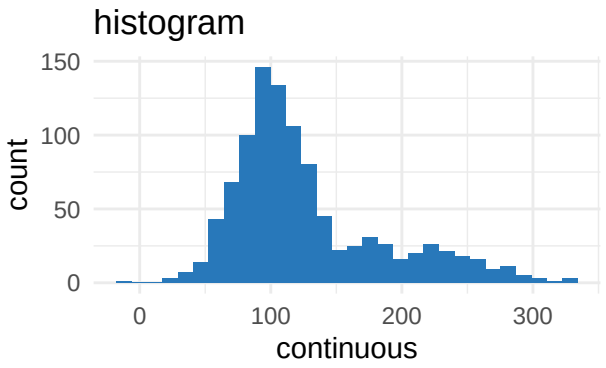
	x	y	z	u	v	w	s	q
301	76.98	209.3	80.38	Group A	Group B	Group B	Group 1	86.98
743	218.9	214.6	110.5	Group B	Group A	Group B	Group 3	248.9

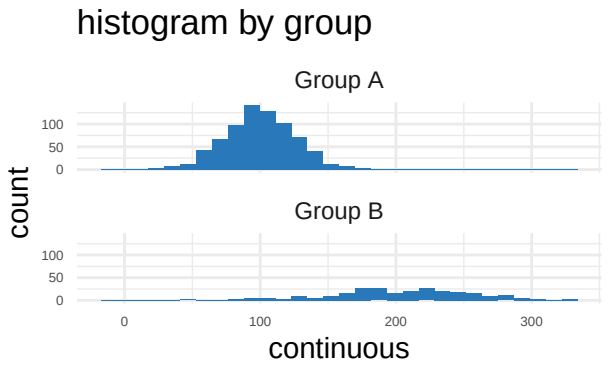
One Thing At A Time

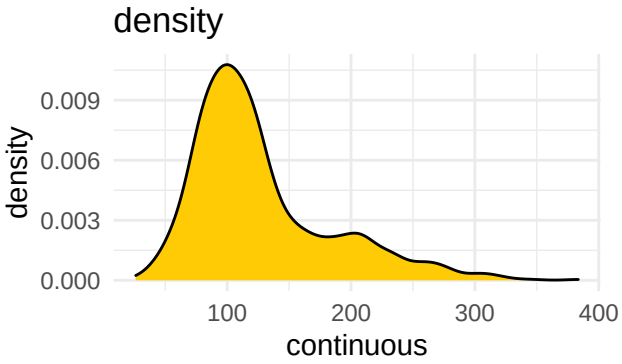
Two Things At A Time

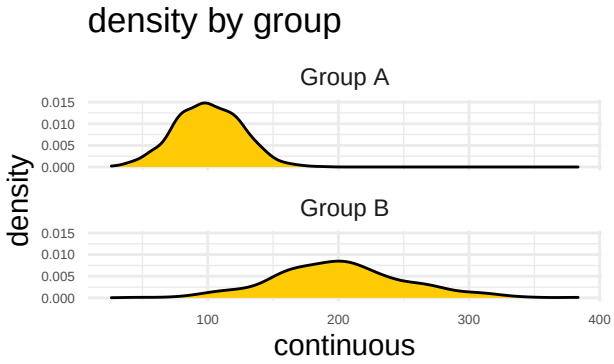
Continuous

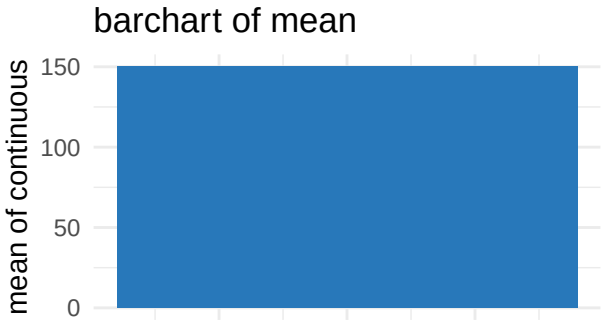
Continuous By Categorical

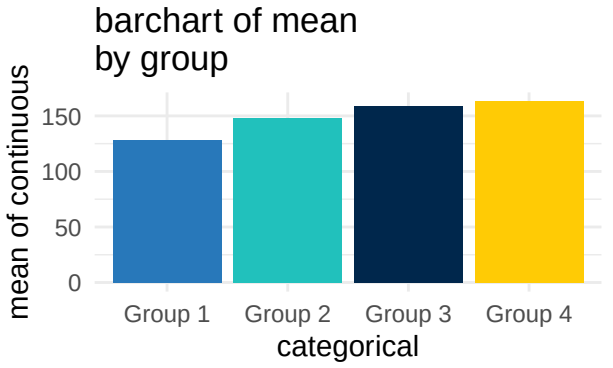




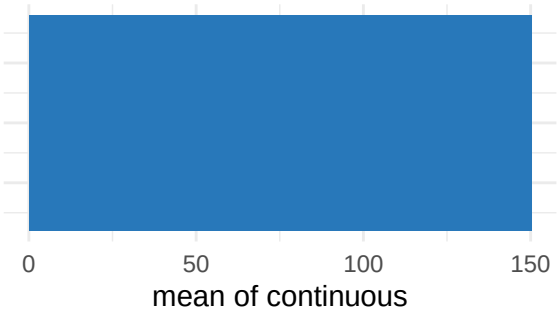




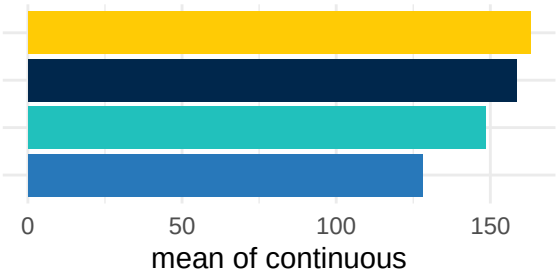




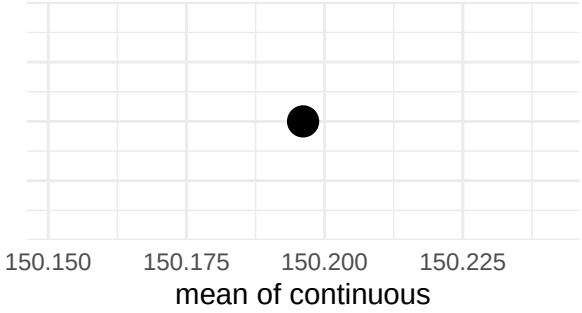
horizontal barchart of mean



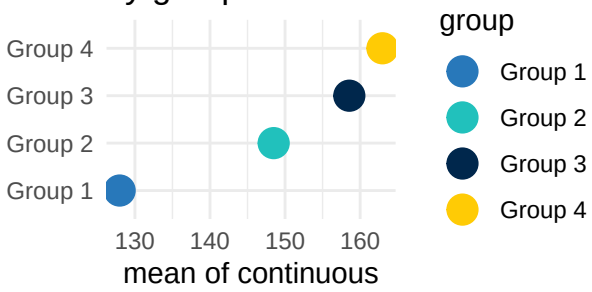
horizontal barchart of mean by group



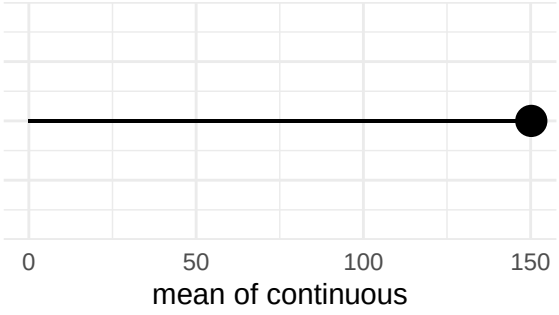
horizontal dotchart of mean



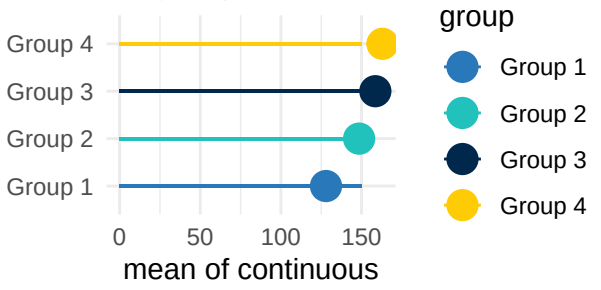
horizontal dotchart of mean by group



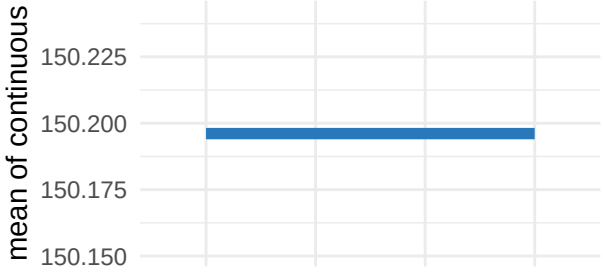
horizontal lollipop chart of mean



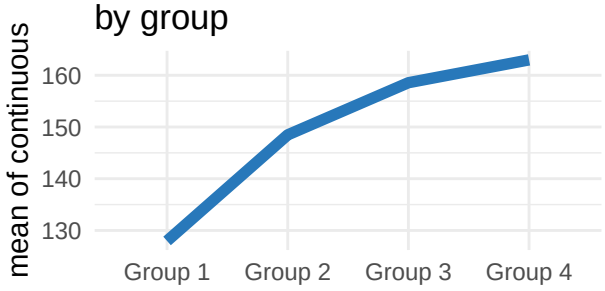
horizontal lollipop chart of mean by group

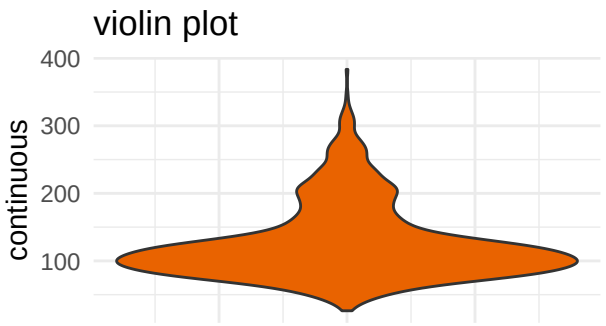


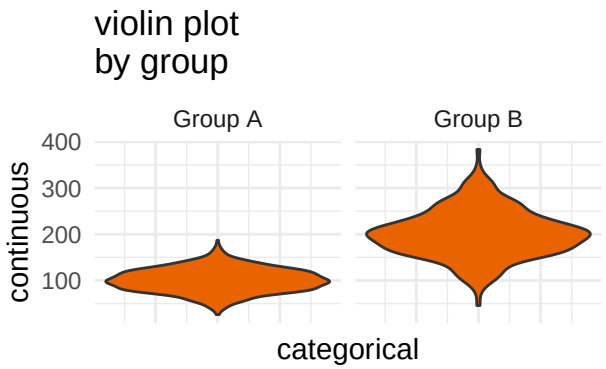
linechart of mean

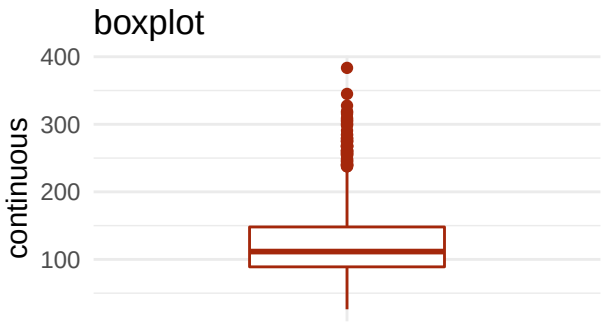


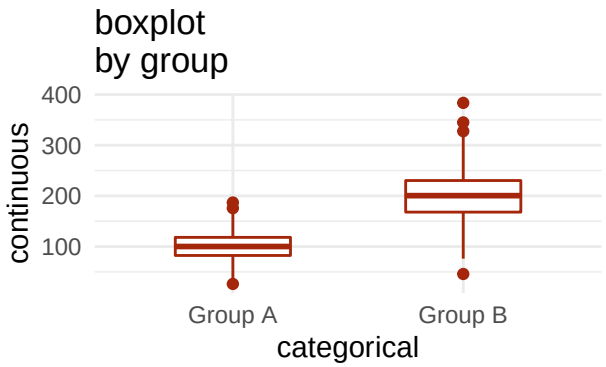
linechart of mean by group

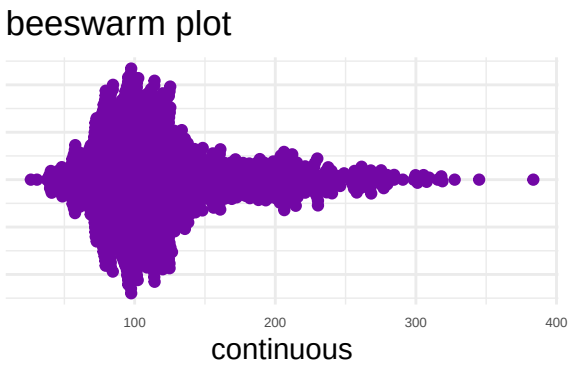


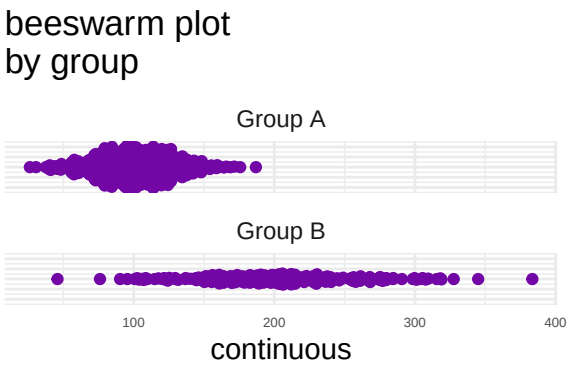

violin plot
400
300
200
100
continuous

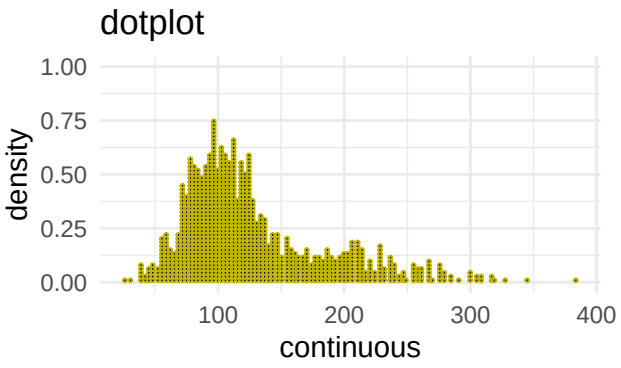

violin plot
by group
Group A
Group B
400
300
200
100
continuous
categorical

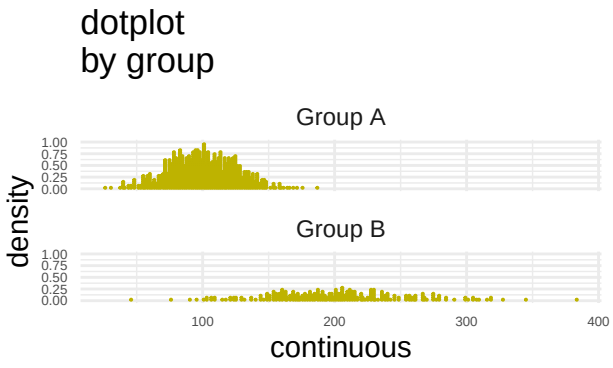

boxplot
400
300
200
100
continuous


boxplot
by group
400
300
200
100
continuous
Group A
Group B
categorical


beeswarm plot
100
200
300
400
continuous


beeswarm plot
by group
Group A
Group B
100
200
300
400
continuous


dotplot
1.00
0.75
0.50
0.25
0.00
density
100
200
300
400
continuous


dotplot
by group
Group A
Group B
1.00
0.75
0.50
0.25
0.00
density
100
200
300
400
continuous

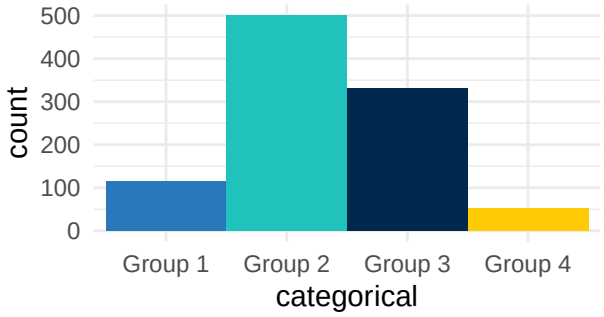
One Thing At A Time

Two Things At A Time

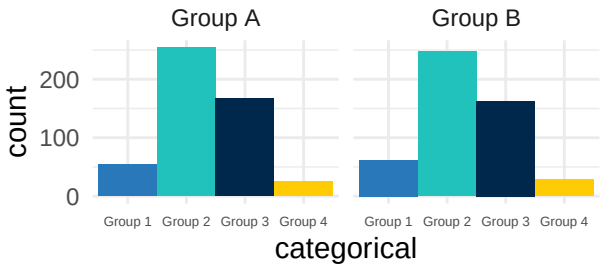
Categorical

Categorical By Categorical

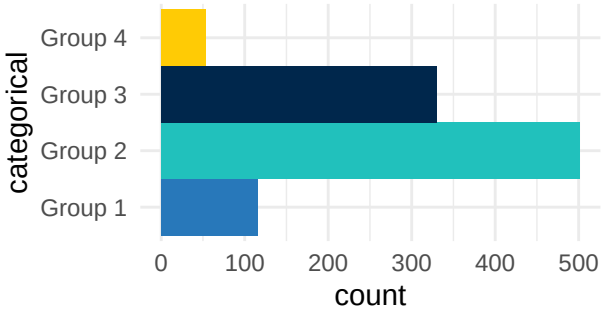
bar chart



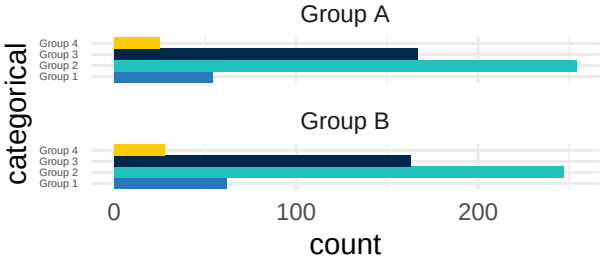
bar chart
by group



horizontal bar chart



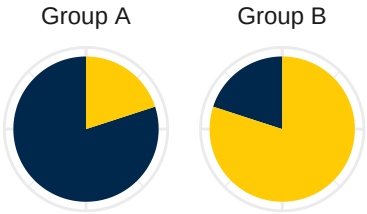
horizontal bar chart
by group



pie chart

categorical

pie chart
by group

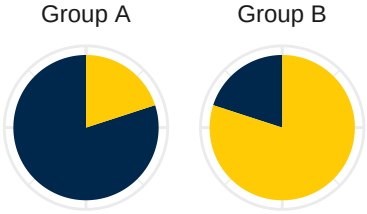


categorical

doughnut chart

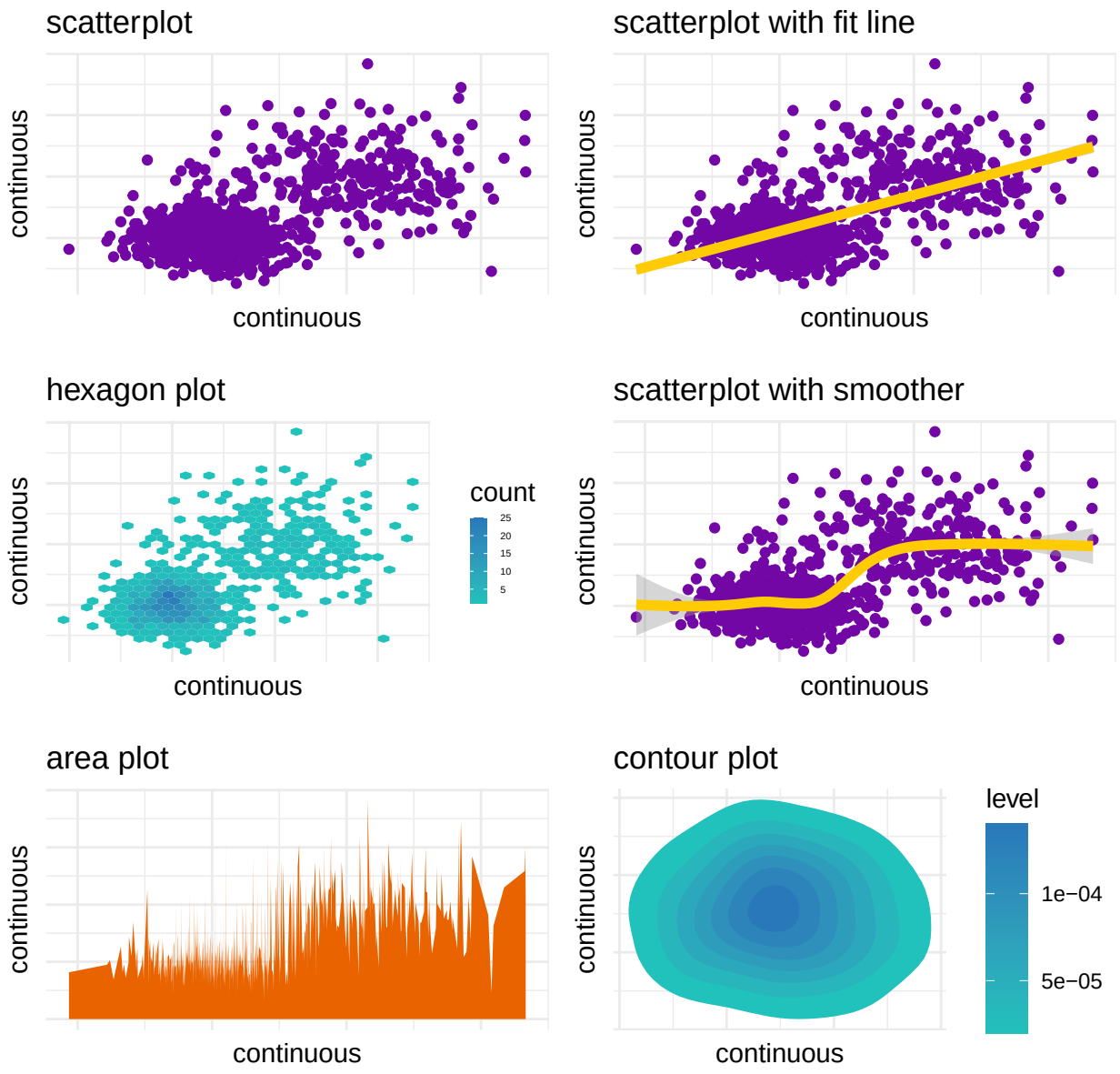
categorical

doughnut chart
by group



categorical

Continuous by Continuous



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

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

Last updated: `March 10 2021` at `16:10`