



# Bar Graphs in ggplot2

By Zaira, Leo, Riya



# Intro

- Great way to visualize numeric values on the y axis, separated into categories by the x axis
- Basic plots are easy to read and highlight differences in data in a way that is easy for our visual processing centers to understand
- Not always the best options for continuous variables
- Heights of bars can represent *counts* of data or *variables*
  - defined by stat = “identity” or “bin”
  - *Labels within bars help!*

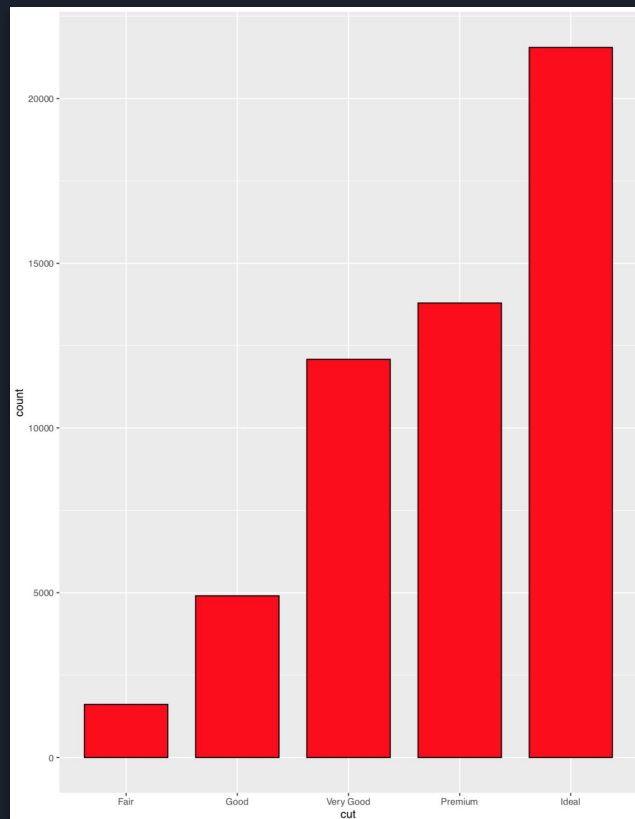


# Dataset

	carat	cut	color	clarity	depth	table	price	x	y	z
1	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
2	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
3	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31
4	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63
5	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75
6	0.24	Very Good	J	VVS2	62.8	57.0	336	3.94	3.96	2.48
7	0.24	Very Good	I	VVS1	62.3	57.0	336	3.95	3.98	2.47
8	0.26	Very Good	H	SI1	61.9	55.0	337	4.07	4.11	2.53
9	0.22	Fair	E	VS2	65.1	61.0	337	3.87	3.78	2.49
10	0.23	Very Good	H	VS1	59.4	61.0	338	4.00	4.05	2.39
11	0.30	Good	J	SI1	64.0	55.0	339	4.25	4.28	2.73
12	0.23	Ideal	J	VS1	62.8	56.0	340	3.93	3.90	2.46
13	0.22	Premium	F	SI1	60.4	61.0	342	3.88	3.84	2.33
14	0.31	Ideal	J	SI2	62.2	54.0	344	4.35	4.37	2.71

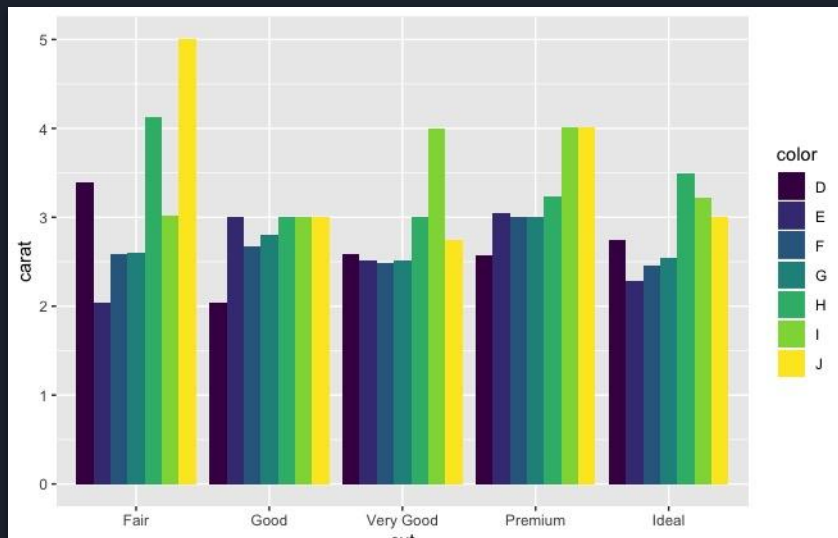
# Bar Plot (Basic) and Code

```
3 library(ggplot2)
4
5 data("diamonds")
6 force(diamonds)
7 # Here, we call our dataset, diamonds.
8
9 basic_graph = ggplot(diamonds, aes(x=cut))
10 # Now we have a variable that represents our foundation that we'll build the graph on top of.
11
12 basic_graph + geom_bar()
13 # Here is just a basic graph with no aesthetics.
14 # This graph is telling us the amount of each cut of the diamonds observed.
15 # See how we used our variable, 'basic_graph'
16 # This way, we don't have to retype "ggplot(diamonds, aes(x=cut))" each time.
17
18 basic_graph + geom_bar(fill="red", color="black", width=0.75)
19 # Here we have some basic aesthetics:
20 # 'fill' defines the color of the bars;
21 # 'color' defines the color of the outline of the bars;
22 # 'width' defines the width of the bars.
```



# Grouping Bar Plot and Code

```
24 graph_2 = ggplot(diamonds, aes(x=cut, y=carat, fill=color))
25 graph_2 + geom_bar(stat="identity", position="dodge")
26 # Here is an even more complicated graph that does some stuff we haven't done yet.
27 # This graph shows the carat of each diamond in the data set, by cut quality.
28 # Furthermore, it shows the different colors by cut quality, and the carats of those different colors.
29 # We allocate diamond color to each specific cut quality by using 'fill=color' in conjuncton with 'position="dodge"'
30 # This is a premium example of the perfect use case for a Bar Plot:
31 # We have two sets of categorical data: Color and Cut, and one set of continuous data: Carat.
32
```

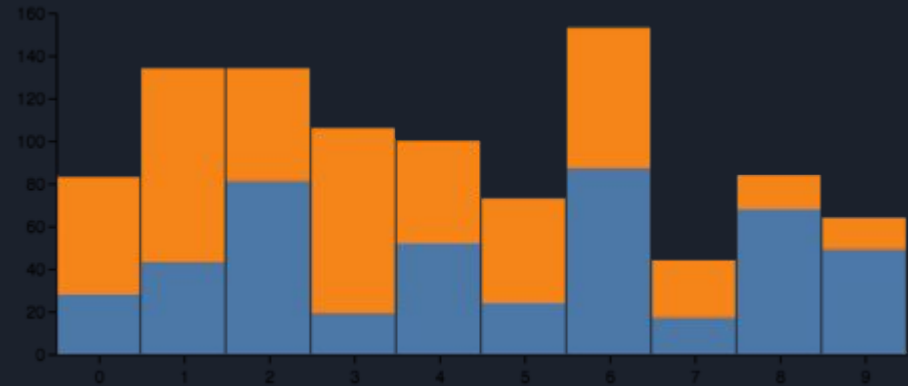


# Stacked Bar Graph

- Stacks bars on top of each other vertically
- Useful for displaying proportions within the same data set

→ (This is also called a **100% stacked bar graph**)

- Can manipulate the stacking order depending on what you want to highlight!





# Sources

R Documentation database:

[https://www.rdocumentation.org/packages/ggplot2/versions/1.0.1/topics/geom\\_bar](https://www.rdocumentation.org/packages/ggplot2/versions/1.0.1/topics/geom_bar)

Chang Textbook Ch. 3