Intro to data exploration and visualization

SDS 323
James Scott (UT-Austin)

Types of features in a data set

Categorical: the answer to a multiple choice question:

- Chevy/Honda/Tesla
- ice cream/cake/pie

Ordinal: categorical, where the answers have an ordering but not a magnitude

- Poor, Moderate, Good, Great
- Private, Corporal, Lieutenant, Colonel, General

Numerical: numbers, whether integer or real-valued

- Beware the "faux numerical" ordinal scale

The basics of data exploration

- Boxplots
- Scatter plots
- Line graphs
- Faceting

- Tables
- Grouping/piping/summarizing
- Bar plots
- Histograms/density plots

Scatter plots

Here are several rows of a data frame about cars. Every row is a car. Every column is a feature describing the car.

```
Error in mosaic::sample(mpg, 6) : object 'mpg' not found
```