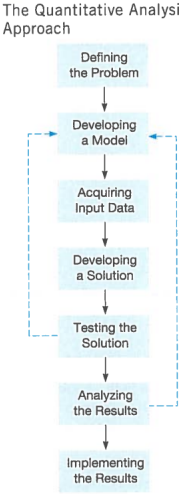
Today's Agenda

Exploring bivariate and multivariate visualizations

Justin Leinaweaver (Spring 2022)

Dataset 1: The Motivating Problem



What drives economic investment in US states?

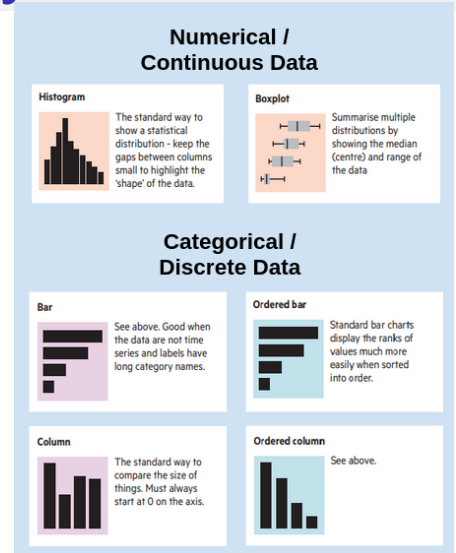
Why do some states attract greater investment by companies and individuals than others?

Last Week: Univariate Analyses

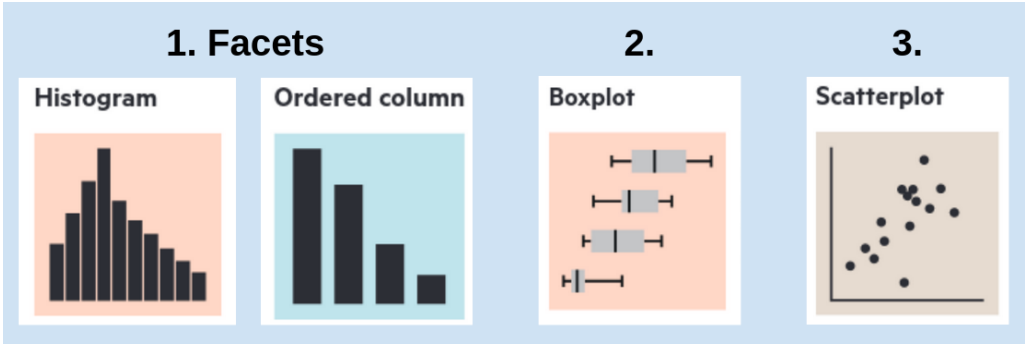
Measures of Central Tendency
- Mean
- Median

Deviations from Central Tendency
- Standard deviation

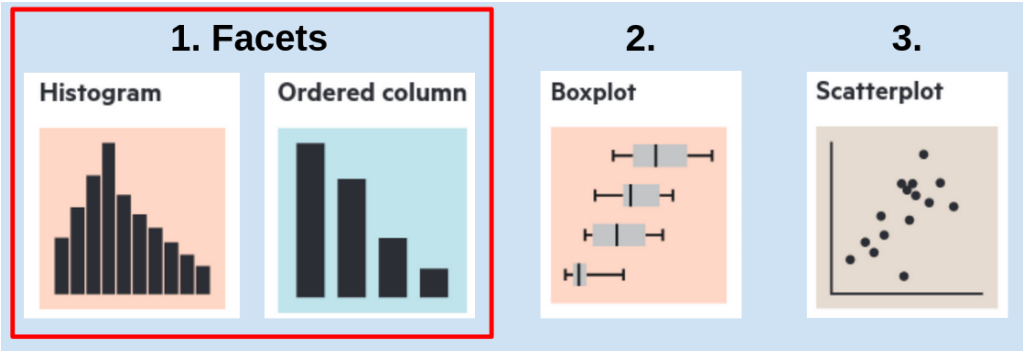
Measures of Variability
- Range
- IQR



Bivariate and Multivariate Visualizations



Bivariate and Multivariate Visualizations



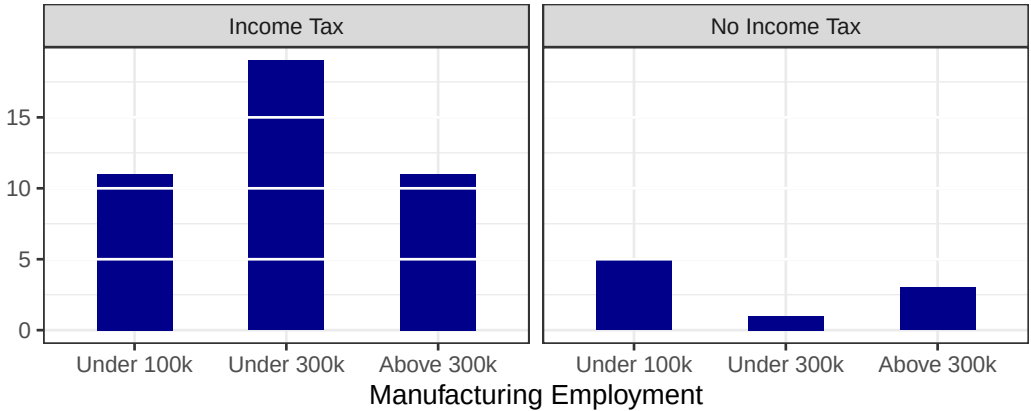
Bivariate Viz: Categorical x Categorical

1. Using Facets to Extend Univariate Visualizations

Make two bar plots of manufacturing category

1. One for states with an income tax, and
2. One for states without an income tax.

Bivariate Viz: Categorical x Categorical



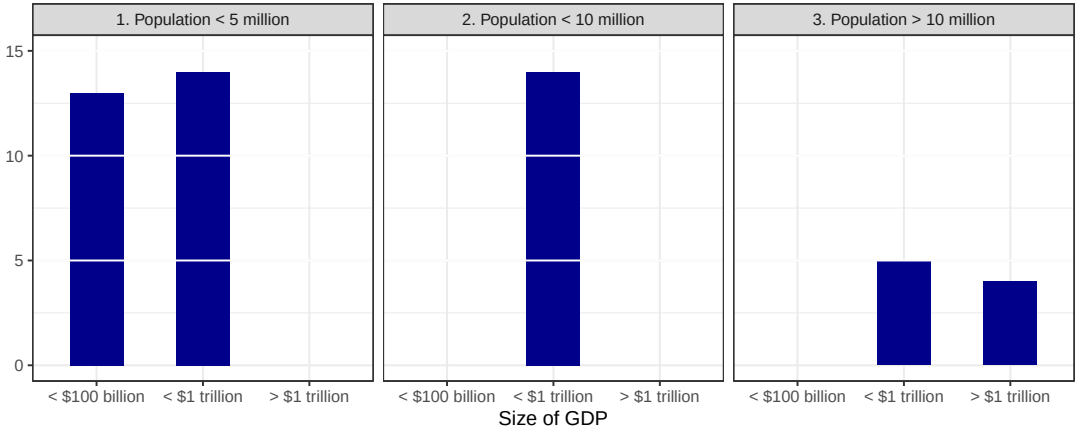
Bivariate Viz: Categorical x Categorical

1. Using Facets to Extend Univariate Visualizations

Make three bar plots of GDP category

1. pop_category = "Under 5 million"
2. pop_category = "Under 10 million"
3. pop_category = "Above 10 million"

Bivariate Viz: Categorical x Categorical

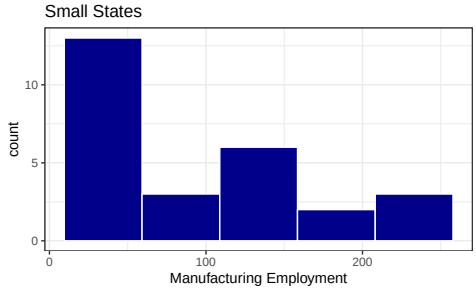


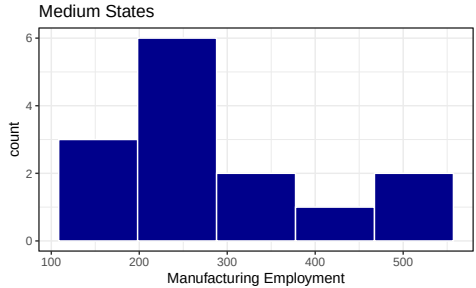
Bivariate Viz: Numerical x Categorical

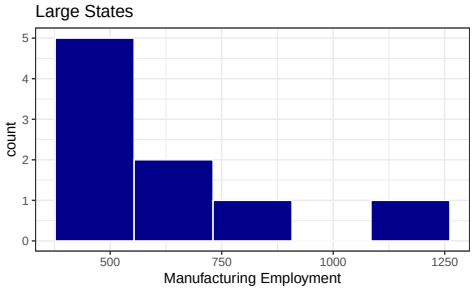
1. Using Facets to Extend Univariate Visualizations

Make three histograms of manufacturing employment (5 bins)

1. pop_category = "Under 5 million"
2. pop_category = "Under 10 million"
3. pop_category = "Above 10 million"

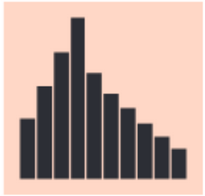

Small States
count
10
5
0
0
100
200
Manufacturing Employment


Medium States
count
6
4
2
0
100
200
300
400
500
Manufacturing Employment


Large States
count
5
4
3
2
1
0
500
750
1000
1250
Manufacturing Employment

1. Facets

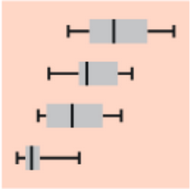
Histogram



Ordered column

2.

Boxplot



3.

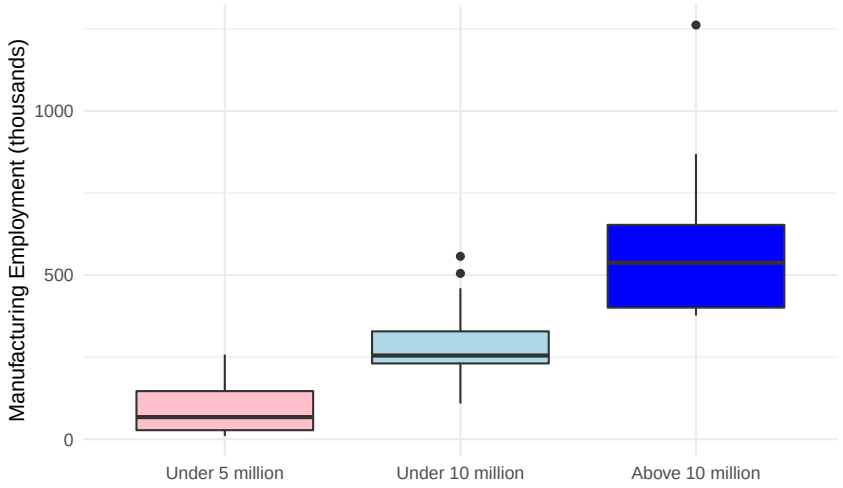
Scatterplot

Using Box Plots: Numerical x Categorical

Remake the last visualization using a box plot instead of histograms with facets.

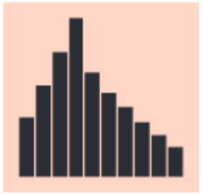
e.g. separate boxes for each population category

Bivariate Viz: Numerical x Categorical



1. Facets

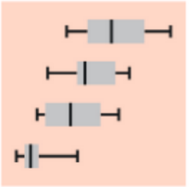
Histogram



Ordered column

2.

Boxplot



3.

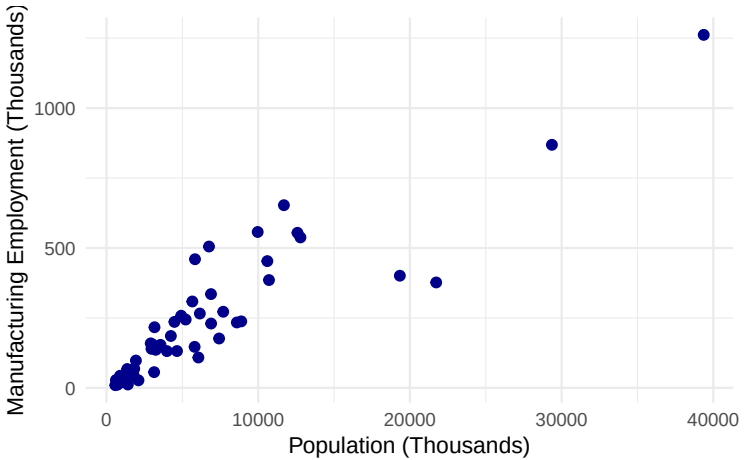
Scatterplot

Using Scatter Plots: Numerical x Numerical

Do states with bigger populations have higher levels of employment in manufacturing?

Make a scatter plot of manufacturing employment and population.

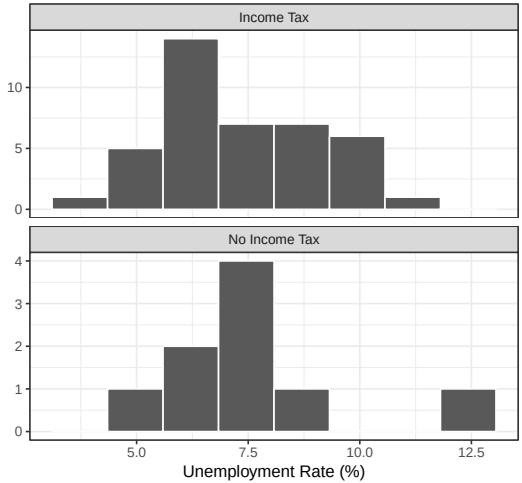
Bivariate Viz: Numerical x Numerical

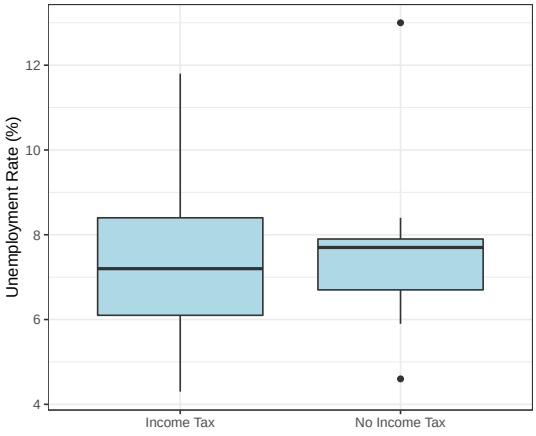


Time to Practice!

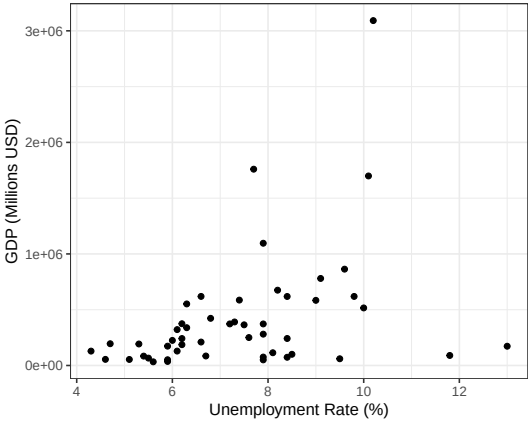
1. Make two histograms of unemployment, one for states with an income tax and one for states without.
2. Remake the above as a box plot
3. Make and analyze the following four scatter plots:
   - GDP (actual) x Unemployment
   - GDP (actual) x Bachelors' Degrees
   - GDP (rate) x Unemployment
   - GDP (rate) x Bachelors' Degrees

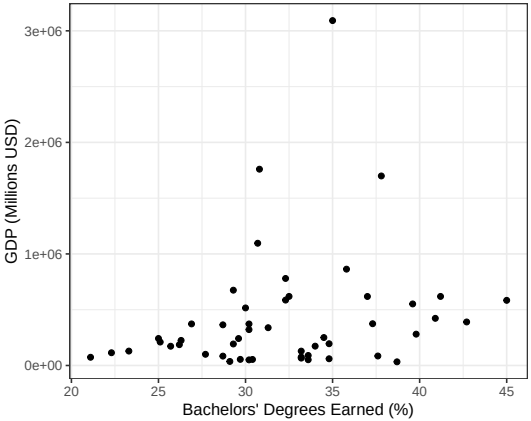
Unemployment x Income Taxes





Scatter plots of GDP (actual)





Scatter plots of GDP (rate)