Today's Agenda

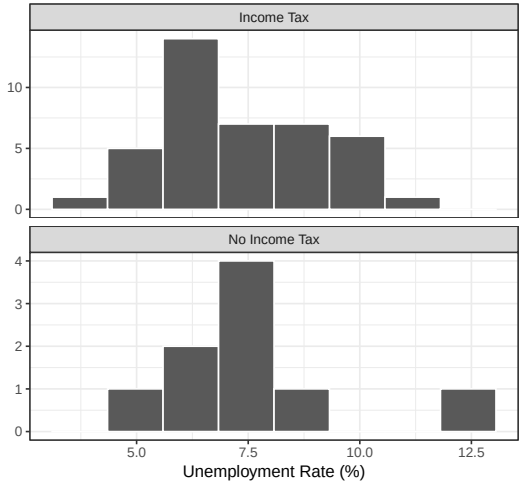
Practice building and polishing univariate, bivariate and multivariate visualizations

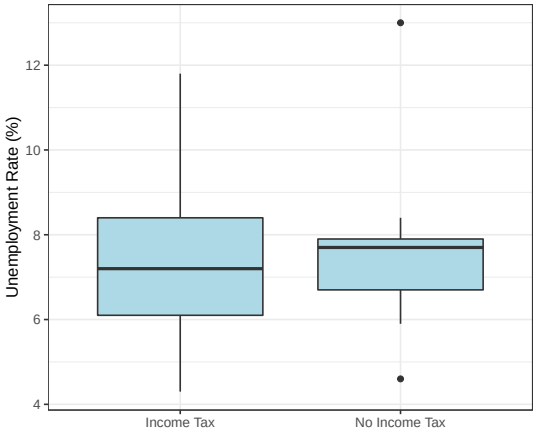
Justin Leinaweaver (Spring 2022)

Practice from Tuesday

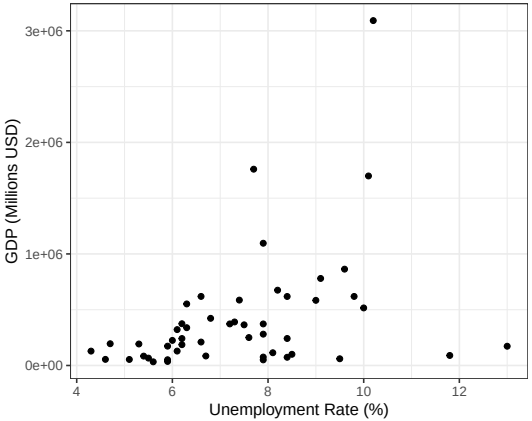
1. Make two histograms of unemployment, one for states with an income tax and one for states without.
2. Remake the above as a box plot
3. Make and analyze the following four scatter plots:
   - GDP (actual) x Unemployment
   - GDP (actual) x Bachelors' Degrees
   - GDP (rate) x Unemployment
   - GDP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

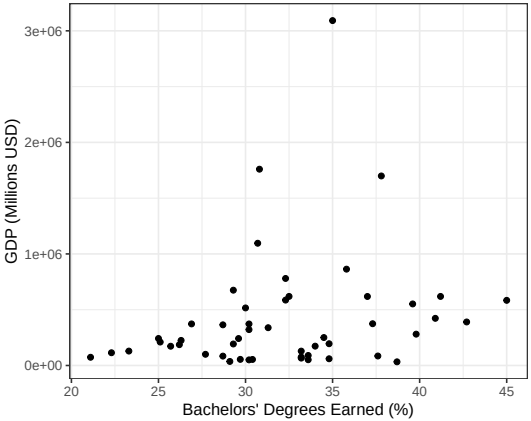
Unemployment x Income Taxes



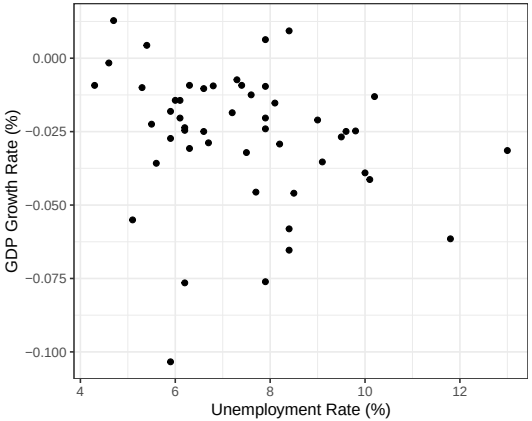


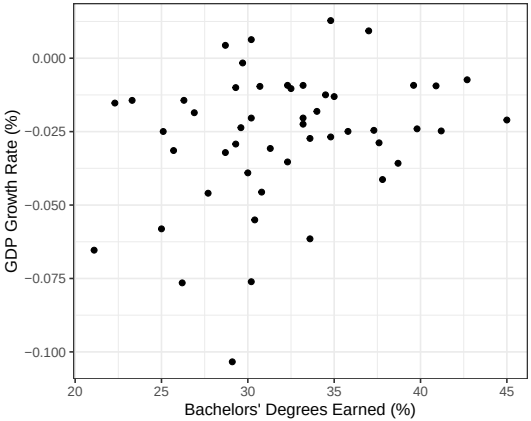
Scatter plots of GDP (actual)





Scatter plots of GDP (rate)





Knaflic, C. (2015). Choosing an Effective Visual. In *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley.

chapter two

choosing an effective visual

There are many different graphs and other types of visual displays of information, but a handful will work for the majority of your needs. When I look back over the 150+ visuals that I created for workshops and consulting projects in the past year, there were only a dozen different types of visuals that I used (Figure 2.1). These are the visuals we'll focus on in this chapter.

Professional Visualizations

1. Informative titles
2. Figure labels
3. Clean axis labels
4. Source info
5. No chart junk

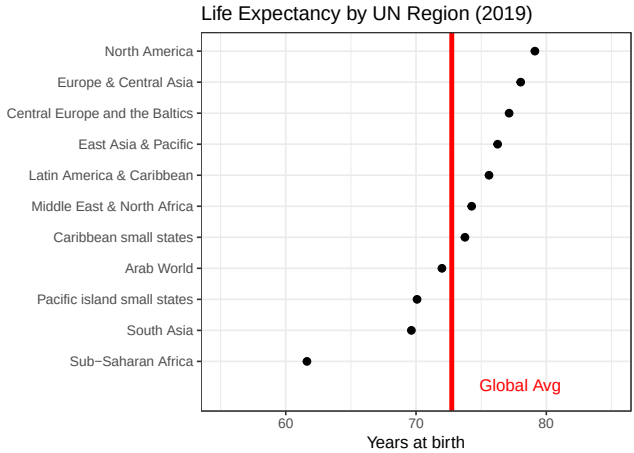


Figure 1: Global life expectancies for 2019 are taken from the World Bank's World Development Indicators (WDI) database and organized by UN region.

Work, Family, and Well-Being in the United States, 1990 (ICPSR 6666)

Version Date: Jun 10, 1996 Cite this study | Share this page

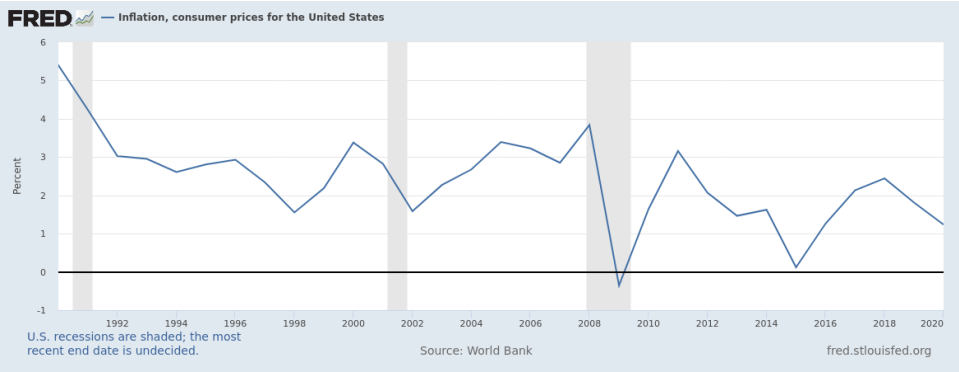
Principal Investigator(s):
Catherine E. Ross

https://doi.org/10.3886/ICPSR06666.v1

Version V1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	height	weight	male	earn	earnk	ethnicity	education	mother_education	father_education	walk	exercise	smokenow	tense	angry	age
2	74	210	1	50000	50	White	16	16	16	3	3	2	0	0	45
3	66	125	0	60000	60	White	16	16	16	6	5	1	0	0	58
4	64	126	0	30000	30	White	16	16	16	8	1	2	1	1	29
5	65	200	0	25000	25	White	17	17	NA	8	1	2	0	0	57
6	63	110	0	50000	50	Other	16	16	16	5	6	2	0	0	91
7	68	165	0	62000	62	Black	18	18	18	1	1	2	2	2	54
8	63	190	0	51000	51	White	17	17	17	3	1	2	4	4	39
9	64	125	0	9000	9	White	15	15	15	7	4	1	4	4	26
10	62	200	0	29000	29	White	12	12	12	2	2	2	0	0	49
11	73	230	1	32000	32	White	17	17	17	7	1	1	0	0	46
12	72	176	1	2000	2	Hispanic	15	15	15	8	1	2	0	0	21
13	72	265	1	35000	35	White	NA	NA	NA	1	1	2	0	0	53
14	72	160	1	27000	27	White	12	12	12	1	2	2	1	1	26
15	70	225	1	6530	6.53	White	16	16	NA	4	1	2	0	0	65
16	63	107	0	0	0	White	14	14	14	7	4	2	2	2	50

Inflation, consumer prices



- Since 1990 $\approx +102.66\%$
- **New Variable**: earnk2021 = earnk x 2.0266

Univariate Analysis

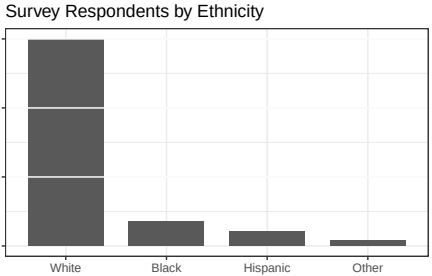
Visualize the following variables:

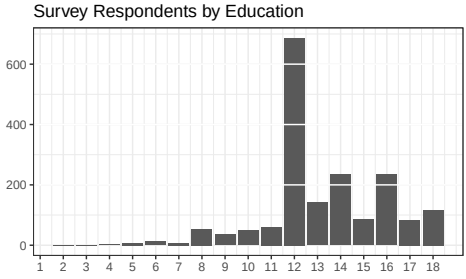
1. Ethnicity
2. Education
3. Yearly Income (earnk2021)

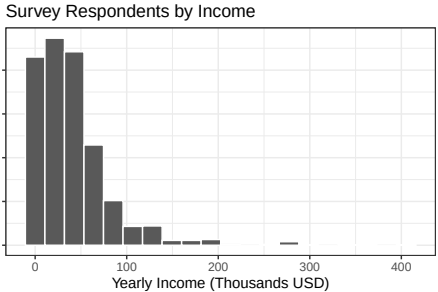
Univariate Analysis

Visualize the following variables:

1. Ethnicity - **Bar plot**
2. Education - **Bar plot**
3. Yearly Income (earnk2021) - **Histogram**


Survey Respondents by Ethnicity
1500
1000
500
0
White
Black
Hispanic
Other

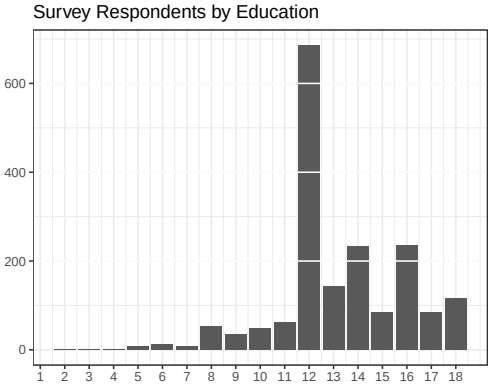

Survey Respondents by Education
600
400
200
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

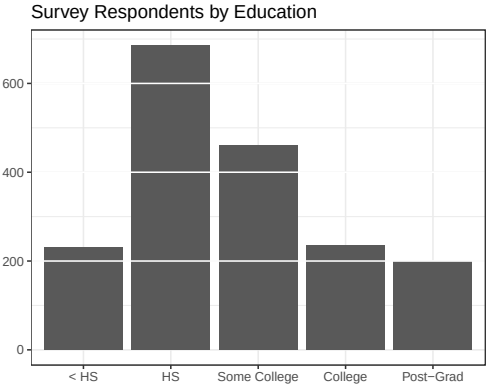

Survey Respondents by Income
400
300
200
100
0
0
100
200
300
400
Yearly Income (Thousands USD)

Univariate Analysis

Recode education and remake the visualization:

- $<$ HS (1-11)
- HS (12)
- Some college (13-15)
- College (16)
- Post-Grad (17-18)


Survey Respondents by Education
600
400
200
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

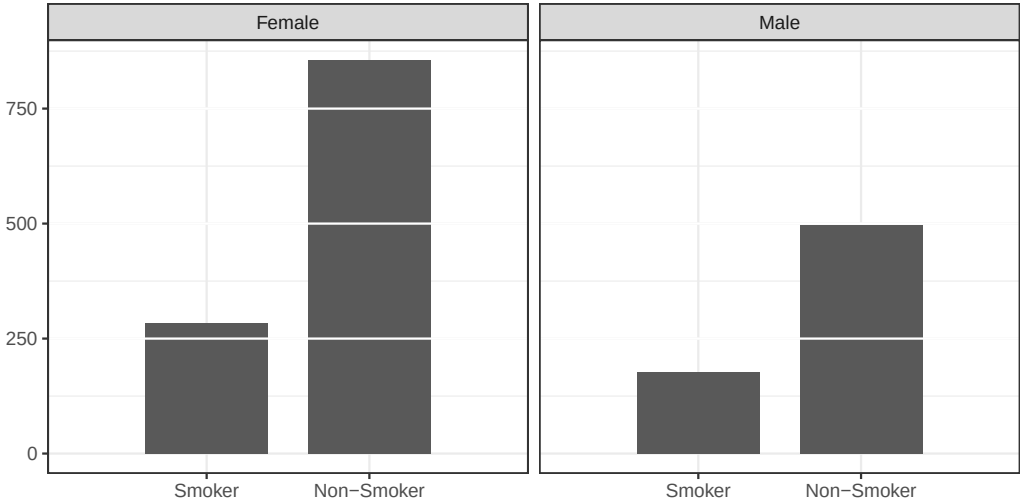

Survey Respondents by Education
600
400
200
0
< HS
HS
Some College
College
Post−Grad

Bivariate Analyses

Visualize the following relationship:

- Gender x Smoker

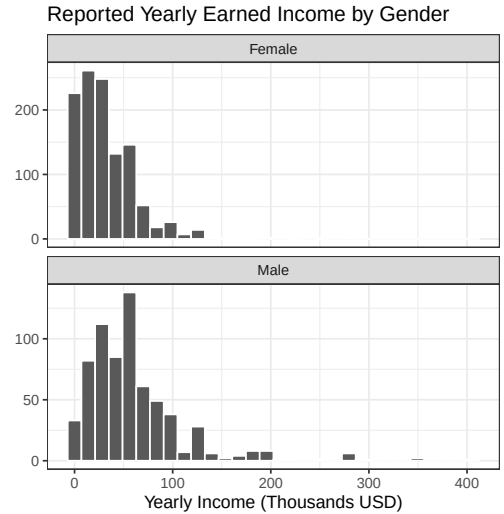
Reported Smoking Rates by Gender

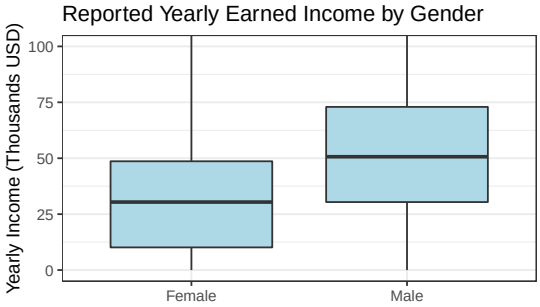


Bivariate Analyses

Visualize the following relationship:

- Income x Gender

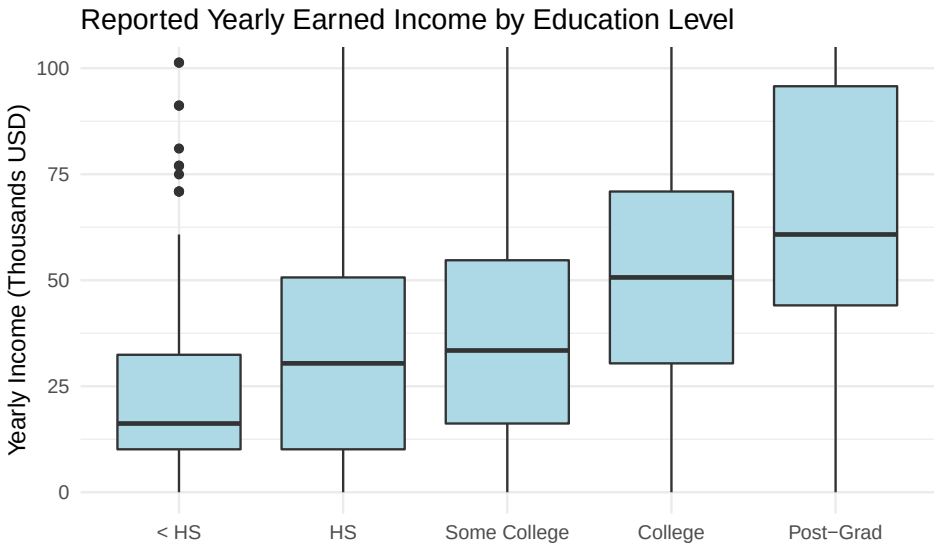

Reported Yearly Earned Income by Gender
Female
Male
200
100
0
100
50
0
0
100
200
300
400
Yearly Income (Thousands USD)


Reported Yearly Earned Income by Gender
Yearly Income (Thousands USD)
100
75
50
25
0
Female
Male

Bivariate Analyses

Visualize the following relationship:

- Income x Education (recoded)


Reported Yearly Earned Income by Education Level
Yearly Income (Thousands USD)
100
75
50
25
0
< HS
HS
Some College
College
Post−Grad

Multivariate Analyses

Visualize the following relationship:

- Income x Education (original) x Gender

Multivariate Analyses

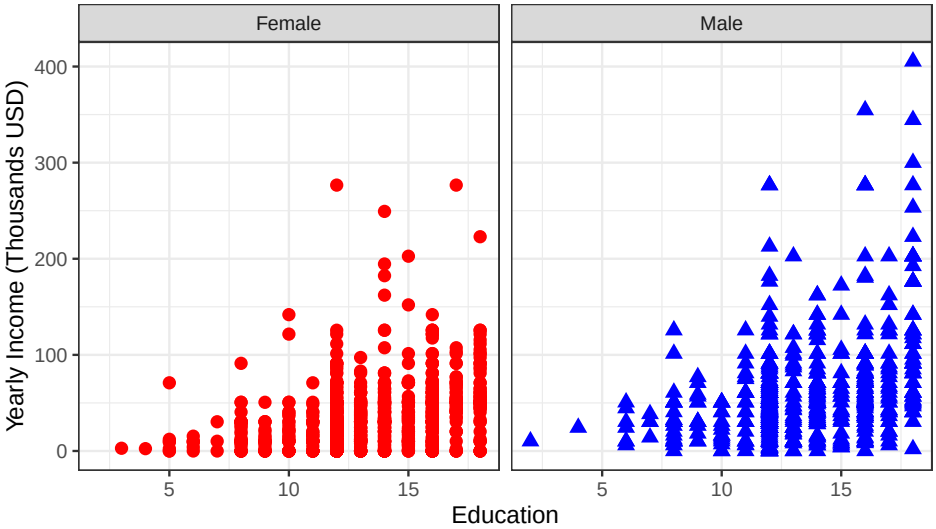
Visualize the following relationship:

- Income x Education (original) x Gender

Make two scatter plots

1. Income x Education for males
2. Income x Education for females

Does education correlate with income?



Multivariate Analyses

Visualize the following relationship:

- Height x Weight x Gender

Basic Demographics of the Survey Respondents