Lecture 4: Histograms

February 9, 2026

Course Administration

1. Tutorial 3 quiz

Course Administration

1. Tutorial 3 quiz
2. Will return Tutorial 2 quiz at end of class
3. Policy brief proposal comments posted – alert me if you don't see them
4. Reminder: Fully composed chart due class after holiday **Feb. 23**
   - if there is something you want to do, but can't figure out how
   - write it in words accompanying the graph
   - I want to be sure you know what to do
   - we can work on how to do it
   - look at the grading sheet for complete graph elements

Course Administration

1. Tutorial 3 quiz
2. Will return Tutorial 2 quiz at end of class
3. Policy brief proposal comments posted – alert me if you don't see them
4. Reminder: Fully composed chart due class after holiday **Feb. 23**
   - if there is something you want to do, but can't figure out how
   - write it in words accompanying the graph
   - I want to be sure you know what to do
   - we can work on how to do it
   - look at the grading sheet for complete graph elements
5. Anything else lingering?

General Policy Brief Proposal Feedback

Good work and interesting topics.

General Policy Brief Proposal Feedback

Good work and interesting topics.

Successful proposals

- have a point to make and a story to tell
- clearly set out the ≥ 2 data sources you're using
- explain how you're planning on aggregating data
- give a sense of having some thoughts about the graphics you'd like to do or the points you'd like to make
- for this class, aggregation does not mean merging together – it means going from one unit of observation to another

General Policy Brief Proposal Feedback

Good work and interesting topics.

Successful proposals

- have a point to make and a story to tell
- clearly set out the ≥ 2 data sources you're using
- explain how you're planning on aggregating data
- give a sense of having some thoughts about the graphics you'd like to do or the points you'd like to make
- for this class, aggregation does not mean merging together – it means going from one unit of observation to another

Do not

- report regression coefficients

General Policy Brief Proposal Feedback

Good work and interesting topics.

Successful proposals

- have a point to make and a story to tell
- clearly set out the ≥ 2 data sources you're using
- explain how you're planning on aggregating data
- give a sense of having some thoughts about the graphics you'd like to do or the points you'd like to make
- for this class, aggregation does not mean merging together – it means going from one unit of observation to another

Do not

- report regression coefficients

In the best final work, graphics drive narrative

Looking Forward to the Final Product

- Final product needs 8 to 10 graphics
- Graphics with some basic descriptives often set the stage
- May be helpful to think about summary statistics before correlations
- With new data, good practice for you to match published summary stats
- As relevant, consider adding in decennial census/acs data to add demographics
  - Use nhgis.org for geographic census data
- Come see me about data sources

Looking Forward to the Final Product

- Final product needs 8 to 10 graphics
- Graphics with some basic descriptives often set the stage
- May be helpful to think about summary statistics before correlations
- With new data, good practice for you to match published summary stats
- As relevant, consider adding in decennial census/acs data to add demographics
  - Use nhgis.org for geographic census data
- Come see me about data sources
- Expect to have problems

Looking Forward to the Final Product

- Final product needs 8 to 10 graphics
- Graphics with some basic descriptives often set the stage
- May be helpful to think about summary statistics before correlations
- With new data, good practice for you to match published summary stats
- As relevant, consider adding in decennial census/acs data to add demographics
  - Use nhgis.org for geographic census data
- Come see me about data sources
- Expect to have problems
- Next deadline: Lecture 5, one fully composed chart

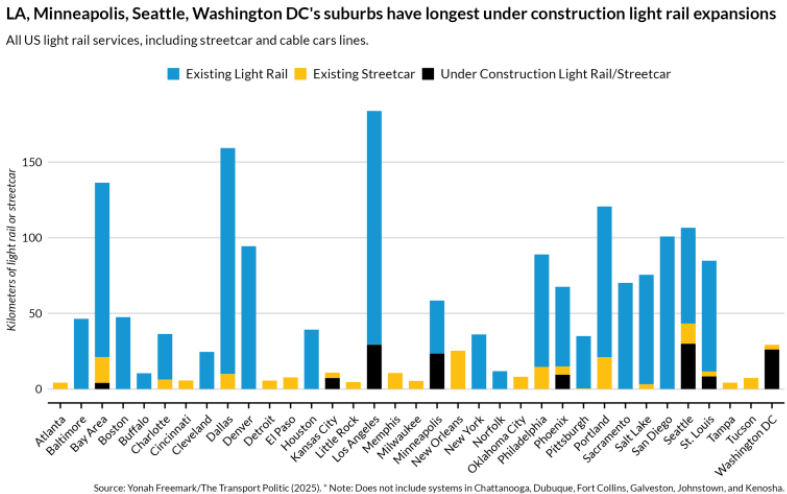
Next Week's Good Bad and Ugly

Find a histogram. Post by Wednesday noon. Post the link on the google sheet.

Finder	Commenter
Kibin	Chase
Tosha	Otsile

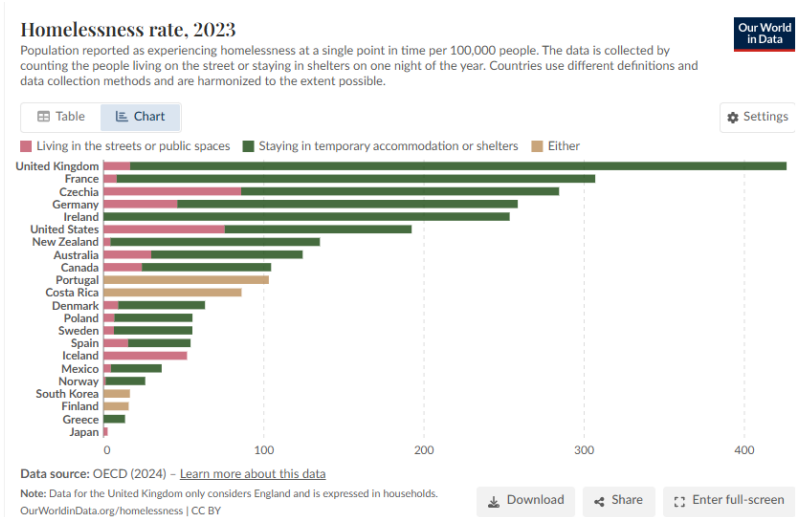
Next week is a holiday.

Sawyer on Chase's Graph



From Yonah Freemark's blog, *The Transport Public*

Kibin on Samantha's Graph



From Our World in Data, link.

How is AI Helpful in Coding?

Using AI Helpfully

- Which AI(s) are you using?
- What problems has it solved for you?
- What problems did it not solve, or solved wrong?
- What do you think it should be able to do for you but has not?

Histograms

Thinking About Histograms

1. What is a histogram and why use one?
2. Most basic histogram
3. Histogram history
4. Histogram variations

1. What are histograms and why should you use them?

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

- mean
- median
- mode
- variance
- percentiles

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

- mean
- median
- mode
- variance
- percentiles

Two ways to think about distributions.

Ex.: Number of commuters by jurisdiction in DMV.

1. levels

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

- mean
- median
- mode
- variance
- percentiles

Two ways to think about distributions.

Ex.: Number of commuters by jurisdiction in DMV.

1. levels
   - number of commuters by jurisdiction

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

- mean
- median
- mode
- variance
- percentiles

Two ways to think about distributions.

Ex.: Number of commuters by jurisdiction in DMV.

1. levels
   - number of commuters by jurisdiction
2. shares

Histograms Show the Distribution of **One** Variable

What are non-graphical ways of describing the distribution of a continuous variable?

- mean
- median
- mode
- variance
- percentiles

Two ways to think about distributions.

Ex.: Number of commuters by jurisdiction in DMV.

1. levels
   - number of commuters by jurisdiction
2. shares
   - share of commuters by jurisdiction

When Do You Want to Show a Histogram?

- For when you need to communicate beyond the mean
- Or show relative prevalence

When Do You Want to Show a Histogram?

- For when you need to communicate beyond the mean
- Or show relative prevalence
- Examples
  - How many people at each part of the income distribution?
  - How much of an outlier is Caitlin Clark?
  - How much variation is there in prices for airline tickets?
  - My work: distribution of condo prices vs other residential prices

When Do You Want to Show a Histogram?

- For when you need to communicate beyond the mean
- Or show relative prevalence
- Examples
  - How many people at each part of the income distribution?
  - How much of an outlier is Caitlin Clark?
  - How much variation is there in prices for airline tickets?
  - My work: distribution of condo prices vs other residential prices
- You give some examples!

When Do You Want to Show a Histogram?

- For when you need to communicate beyond the mean
- Or show relative prevalence
- Examples
  - How many people at each part of the income distribution?
  - How much of an outlier is Caitlin Clark?
  - How much variation is there in prices for airline tickets?
  - My work: distribution of condo prices vs other residential prices
- You give some examples!

But beware! Histograms are harder to explain than bars.

2. Most Basic Histogram

Notionally, to Create a Histogram

- Take a variable
- Make bins by value
- Count the number of observations in each bin
- Plot bars with that number

Notionally, to Create a Histogram

Imaginary Income Data

Person	Income
A	4
B	11
C	12
D	3
E	0

Notionally, to Create a Histogram

Imaginary Income Data

Person	Income
A	4
B	11
C	12
D	3
E	0

Decide on bins

Notionally, to Create a Histogram

Imaginary Income Data

Person	Income
A	4
B	11
C	12
D	3
E	0

Decide on bins

Adding a Bin

Person	Income	Bin
A	4	1-5
B	11	11-15
C	12	11-15
D	3	1-5
E	0	0

Notionally, to Create a Histogram

Imaginary Income Data

Person	Income
A	4
B	11
C	12
D	3
E	0

Decide on bins

Adding a Bin

Person	Income	Bin
A	4	1-5
B	11	11-15
C	12	11-15
D	3	1-5
E	0	0

Binned Dataset

Bin	No.
0	1
1-5	2
6-10	0
11-15	2

Notionally, to Create a Histogram

Imaginary Income Data

Person	Income
A	4
B	11
C	12
D	3
E	0

Decide on bins

Adding a Bin

Person	Income	Bin
A	4	1-5
B	11	11-15
C	12	11-15
D	3	1-5
E	0	0

Binned Dataset

Bin	No.
0	1
1-5	2
6-10	0
11-15	2

Graph this one!

Key Features of Histograms

- Looks like a bar chart
- But! unlike a bar chart, histogram bars touch, to indicate continuity
- Give me more examples of when a histogram would be useful

3. Histogram History

The Histogram Inventor

Karl Pearson (1857-1936) as a young man

The Histogram Inventor

Karl Pearson (1857-1936) as a young man

A big thinker

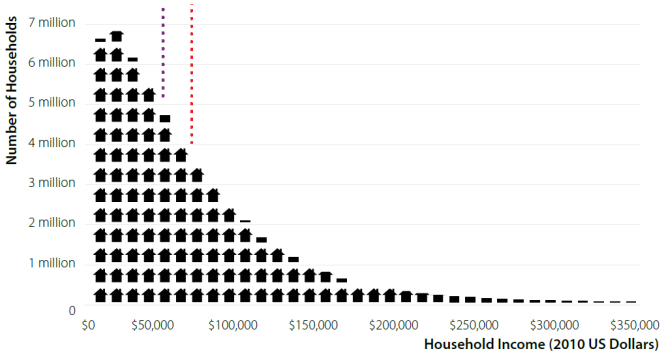
- father of mathematical statistics
- publishes first histogram, 1895
- fervent eugenicist
- early suffragist
- turned down knighthood due to socialist beliefs

4. Histogram Variations

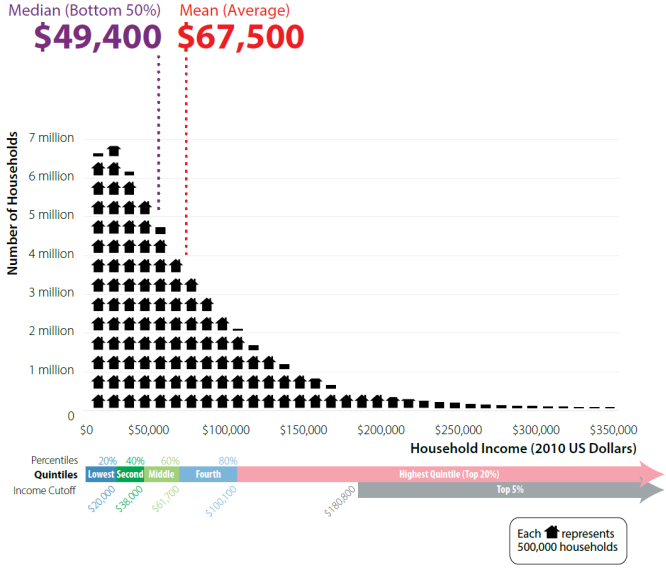
Histogram Variations

1. The classic: touching bars
2. The modified classic: bars with colors by category
3. Make bars into lines: frequency polygon
4. Smooth further: density distribution

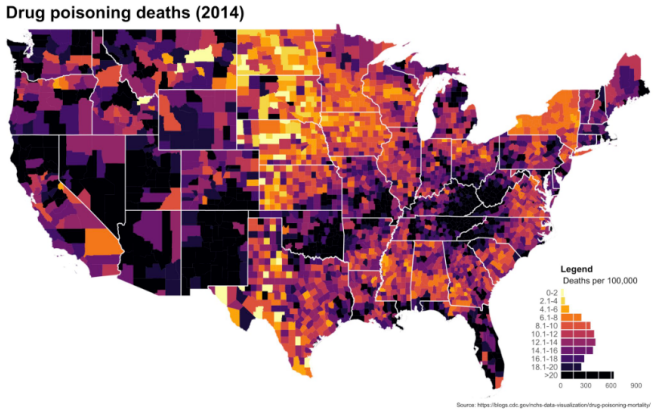
1. The Classic: Mulbrandon's Income Histogram



1. The Classic: Mulbrandon's Income Histogram

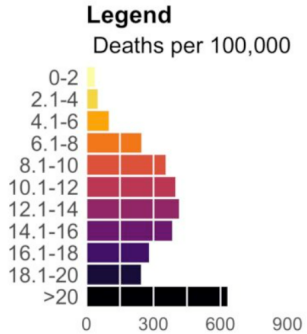


As a Map Legend



From https://mathewkiang.com/2017/01/16/using-histogram-legend-choropleths/

Legend, Now Visible



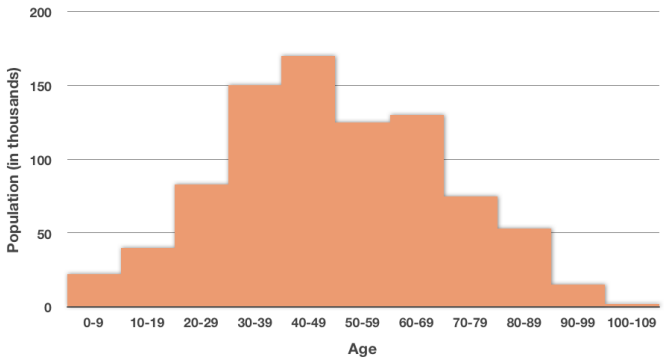
With code on how to do this! From Matthew Kiang's website

2. The Modified Classic: Colored in Bars

Missing a good example – you'll do this in the tutorial

3. Make Bars into Lines: Frequency Polygon

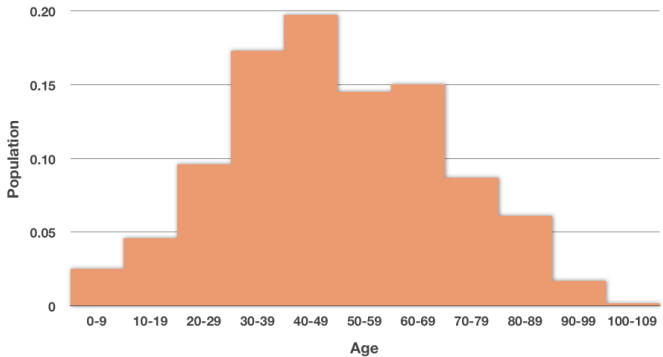
Basic Histogram: Age Distribution



Thank you Krista King math

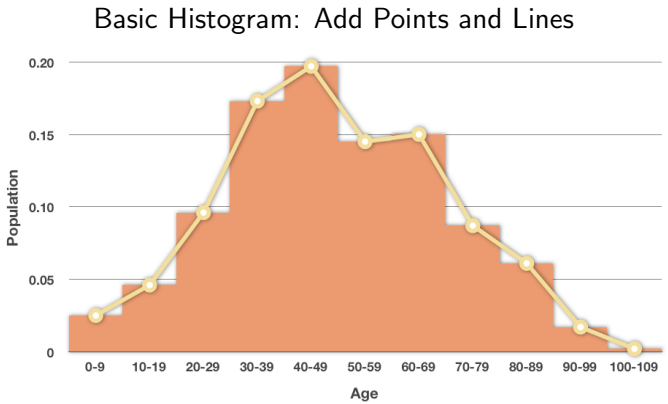
3. Make Bars into Lines: Frequency Polygon

Basic Histogram: Age Distribution – The Same?



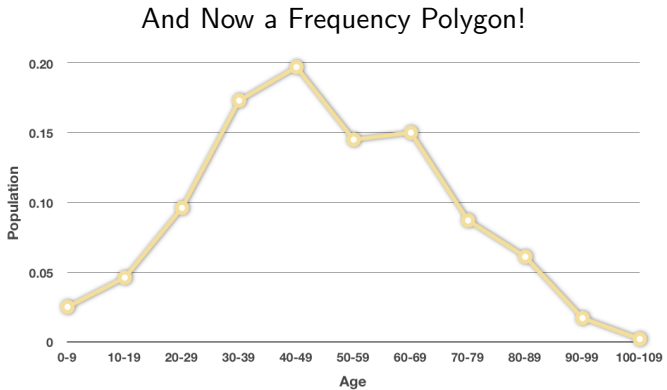
Thank you Krista King math

3. Make Bars into Lines: Frequency Polygon



Thank you Krista King math

3. Make Bars into Lines: Frequency Polygon

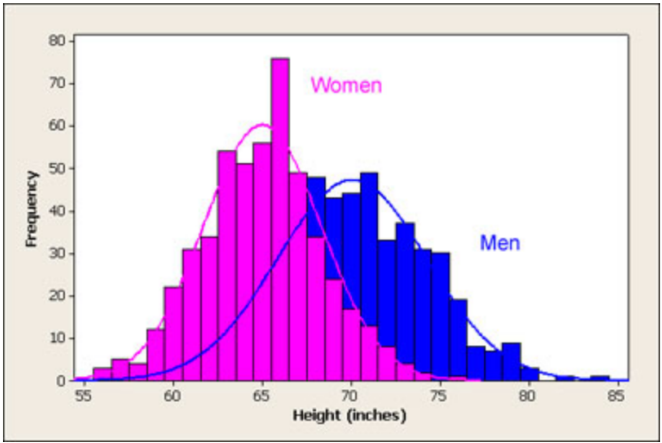


Thank you Krista King math

4. Density Curves: Smoothed Histograms

- Imagine many very thin bars
- This yields a curve
- Sometimes it is more helpful to draw the curve

Height: Note the Curves

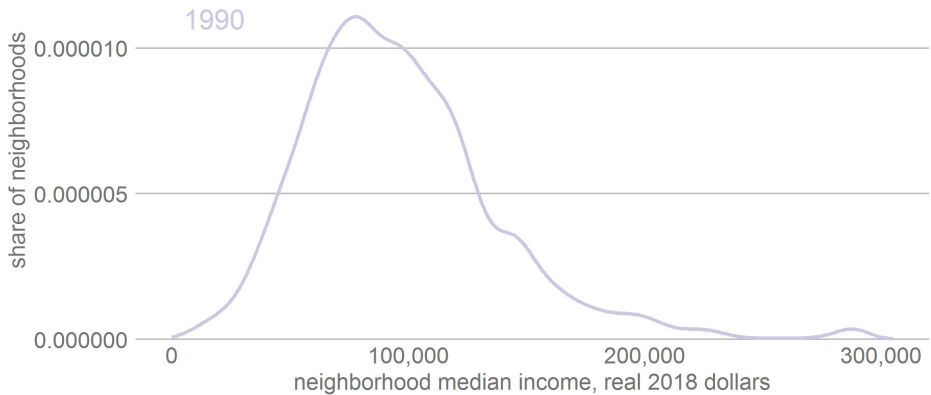


What can you learn about men's versus women's height?

From http://www.usablestats.com/lessons/normal

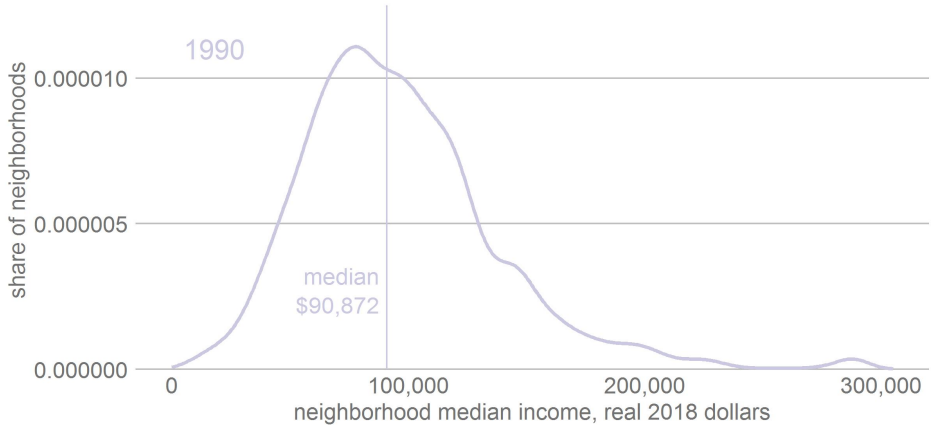
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



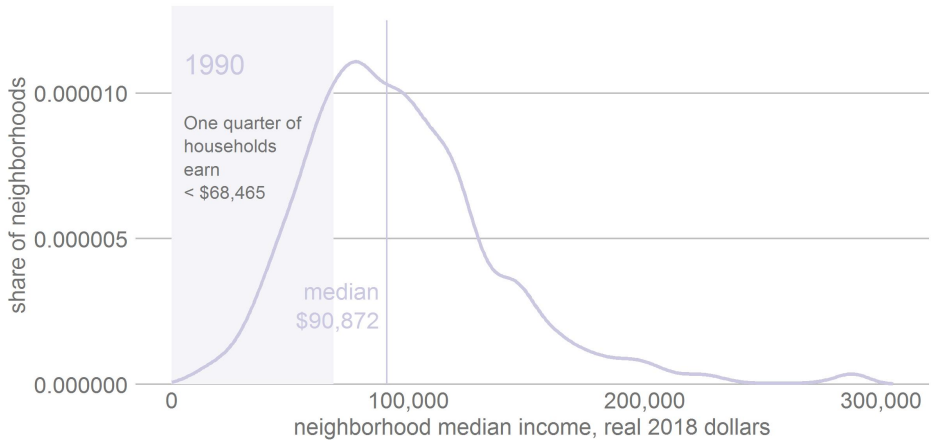
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



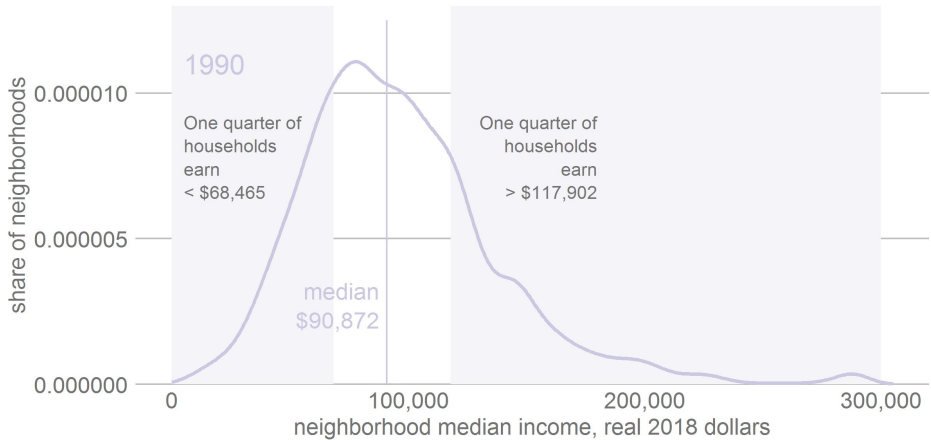
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



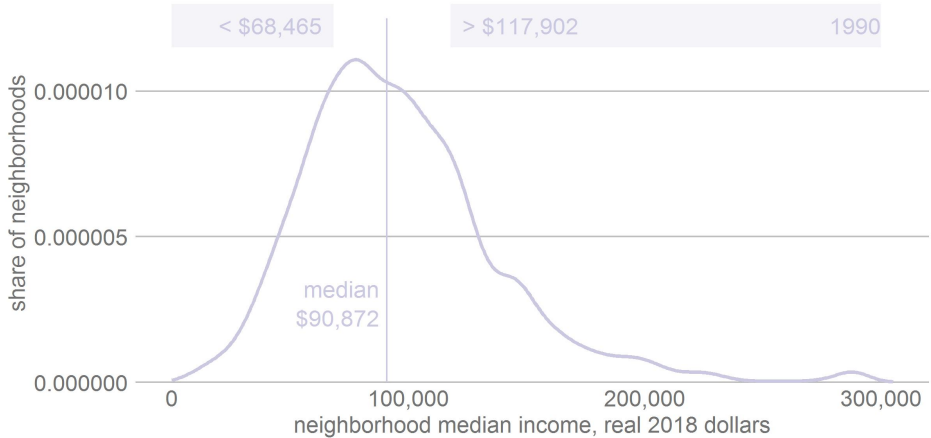
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



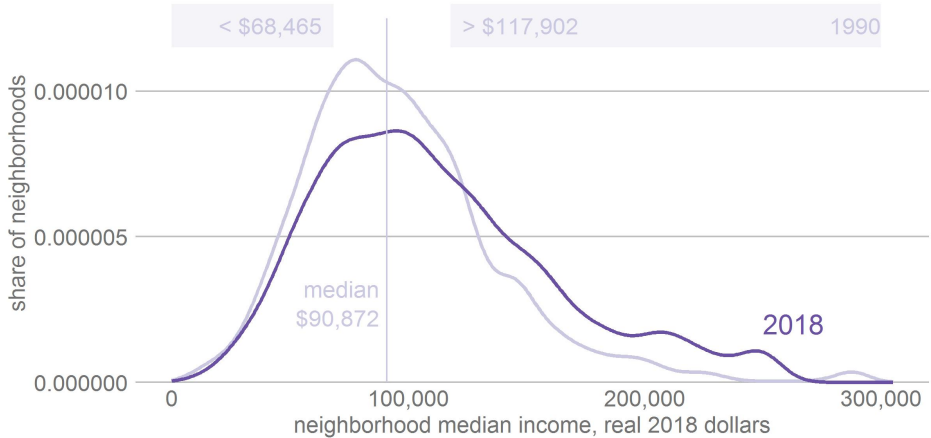
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



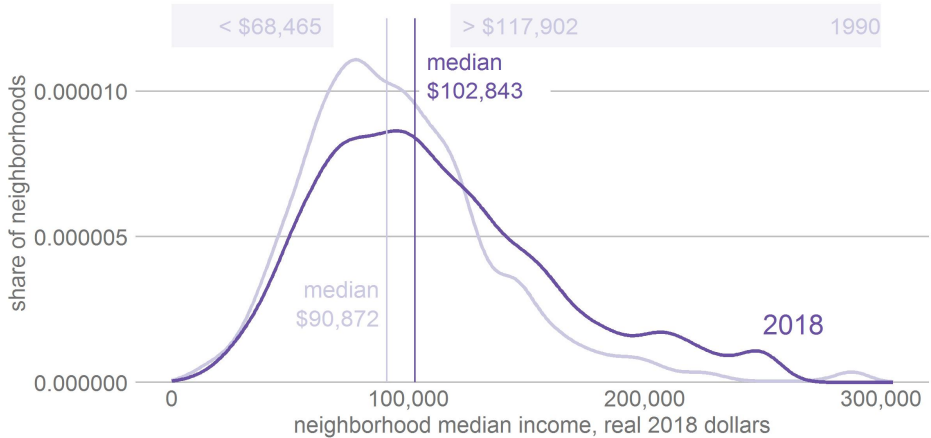
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



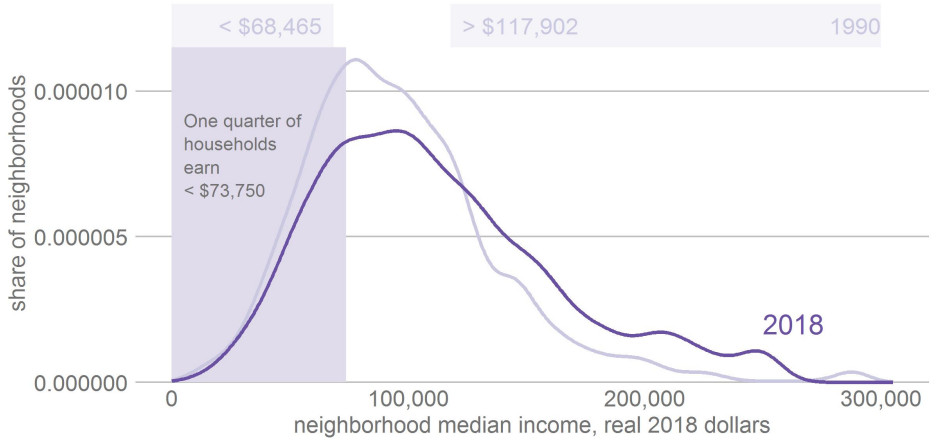
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



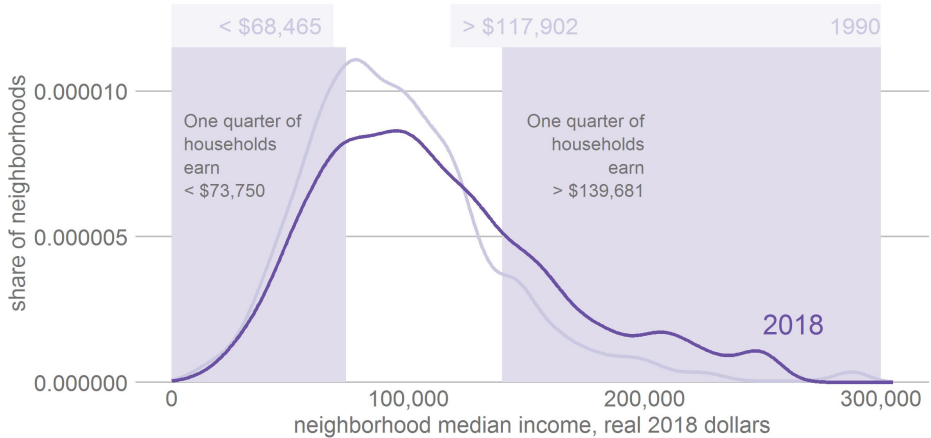
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



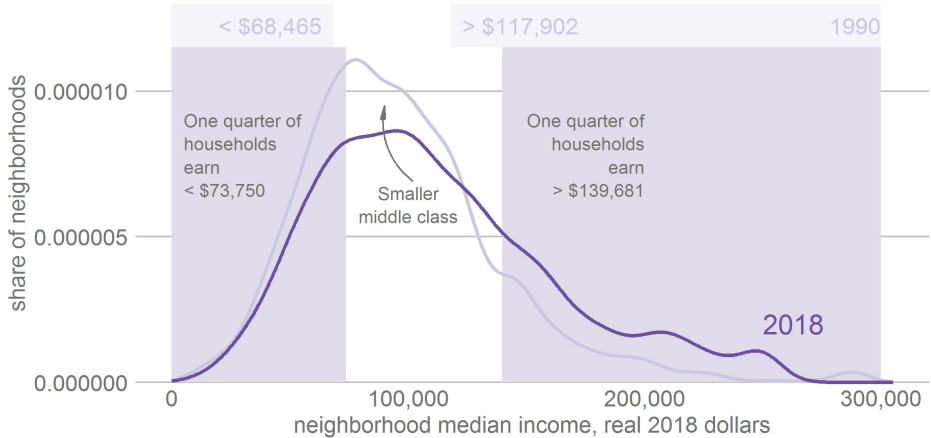
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out



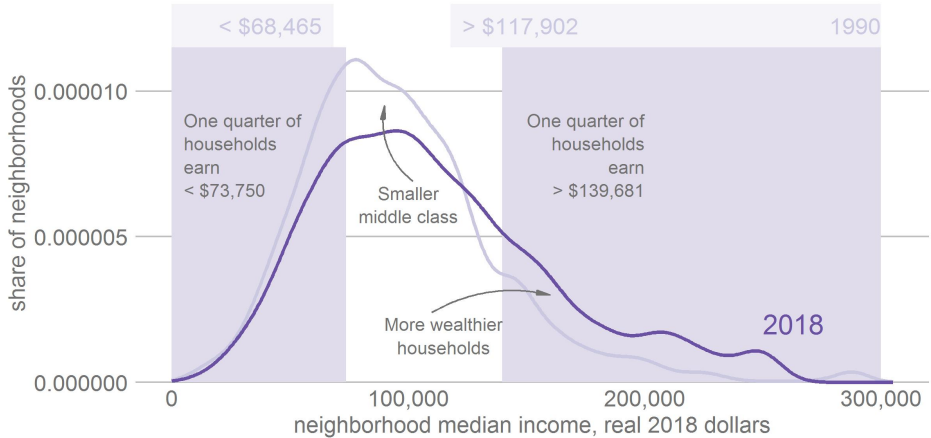
INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out

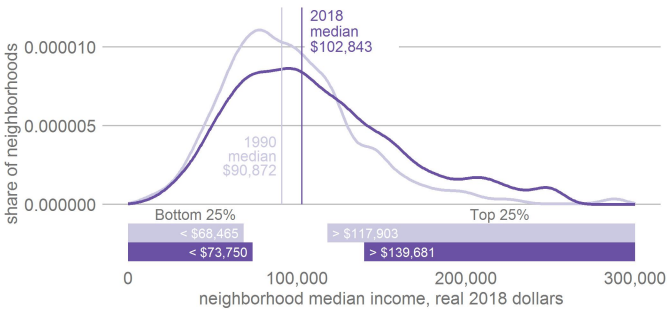


INCOME

Incomes Have Grown, But the Middle Class Has Hollowed Out

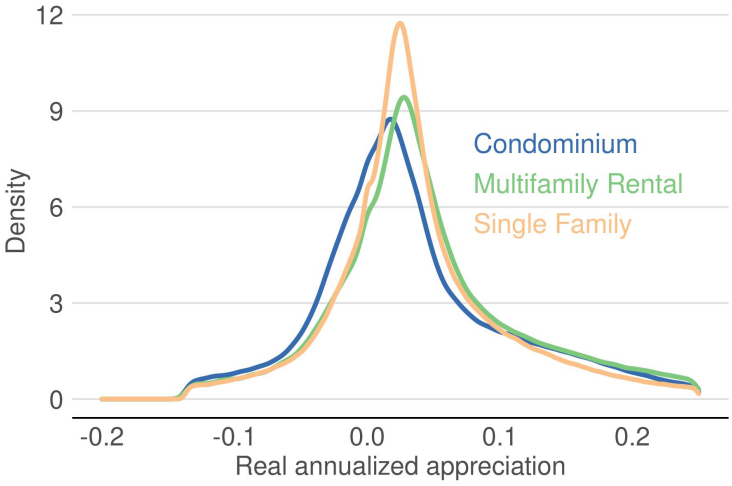


Income Distribution in the DC Metro Area Over Time



- print version
- was never satisfied with y axis
- light purple probably too light
- goal was to show 25th and 75th percentiles
- and change therein

Comparison of Price Appreciation at Distribution Tails



Goal: compare distribution of price appreciation by type

R Histograms

Next Class: Maps

- Happy Presidents Day next week
- See you week after next week
- Be ready for quiz on Tutorial 4
- Turn in fully composed chart assignment via link in assignment
- Manson, *Mapping and Society*, Chapter 7