Lecture 8: Line Charts

March 23, 2026

Course Administration

1. Tutorial 7 quiz
2. Next week: scatter plots
3. Next next week (10): guest lecture + workshop
   - Alyssa Fowers, former Wash Po
   - Please ask questions!
   - Workshop instructions online under Lecture 6
   - Post your work by Sunday April 6 at 3 pm

Course Administration

1. Tutorial 7 quiz
2. Next week: scatter plots
3. Next next week (10): guest lecture + workshop
   - Alyssa Fowers, former Wash Po
   - Please ask questions!
   - Workshop instructions online under Lecture 6
   - Post your work by Sunday April 6 at 3 pm
4. Ask for help!
   - Use Piazza
5. Presentations and final brief
   - formal instructions next week
6. Review function question from last quiz
7. Anything else?

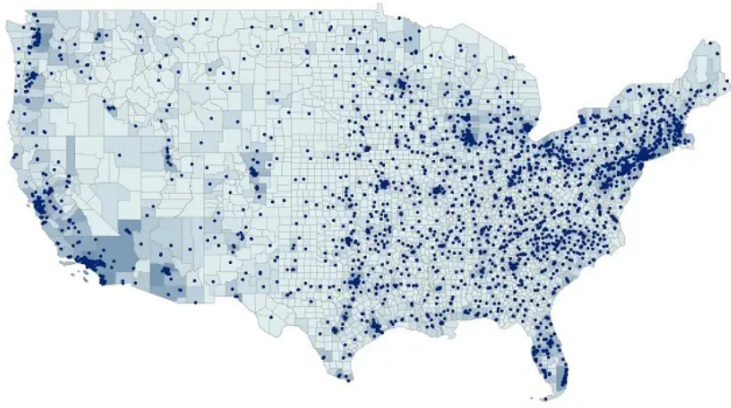
Announcement from Career People

Next Week

Post this week by Wednesday noon. Look for a line chart.

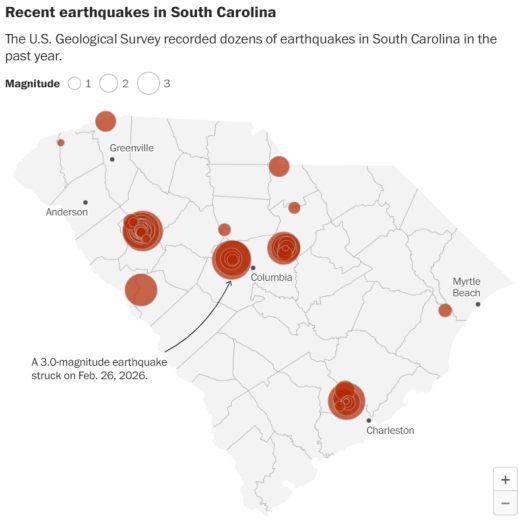
Finder	Commenter
Thomas	Elizabeth
Elizabeth	George

Thomas's Example, Comments by Elizabeth



" Points represent all non-for profit colleges and universities. Shading represents population density. The progression from light to dark represents the progression from low density to high density." •Beth Akers and Adela Soliz, "Mapping the market for higher education," *Brookings Research Report*, Sept. 24, 2015. [link].

Elizabeth's Example, Comments by George



Dennis, Brady, "Why hundreds of earthquakes have rattled South Carolina in recent years," *Washington Post*, March 15, 2026. [link]

Today

1. Line chart overview
2. Moderating lines: inflation and normalization
3. Few on stories
4. Annotations, the power thereof
5. Line charts in R

Line Charts

Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units

Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units
- Values on the vertical axis
  - usually start at zero

Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units
- Values on the vertical axis
  - usually start at zero
- Should you put dots for points?

Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units
- Values on the vertical axis
  - usually start at zero
- Should you put dots for points?
  - Con: Noisy, may add little info
  - Pro: When data are sparse, readers assume full line is data

Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units
- Values on the vertical axis
  - usually start at zero
- Should you put dots for points?
  - Con: Noisy, may add little info
  - Pro: When data are sparse, readers assume full line is data
- Slope has meaning:

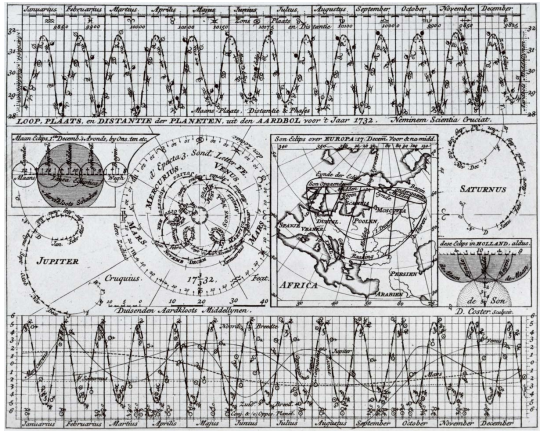
Line Charts

- Have time on the horizontal axis
  - **Always** have consistent time units
- Values on the vertical axis
  - usually start at zero
- Should you put dots for points?
  - Con: Noisy, may add little info
  - Pro: When data are sparse, readers assume full line is data
- Slope has meaning: rate of change
- More than a few lines is too much

Look at the Values on the Horizontal Axis

Line Chart, c. 1732

Nicolaas Kruik (1678-1754) "land surveyor, cartographer, astronomer and weatherman" who "liked to measure things"



Thanks to Wikipedia.

How to Call Things out in a Line Chart

How to Call Things out in a Line Chart

Think back to preattentive processing

How to Call Things out in a Line Chart

Think back to preattentive processing

- color
- size
- timing

How to Call Things out in a Line Chart

Think back to preattentive processing

- color
- size
- timing

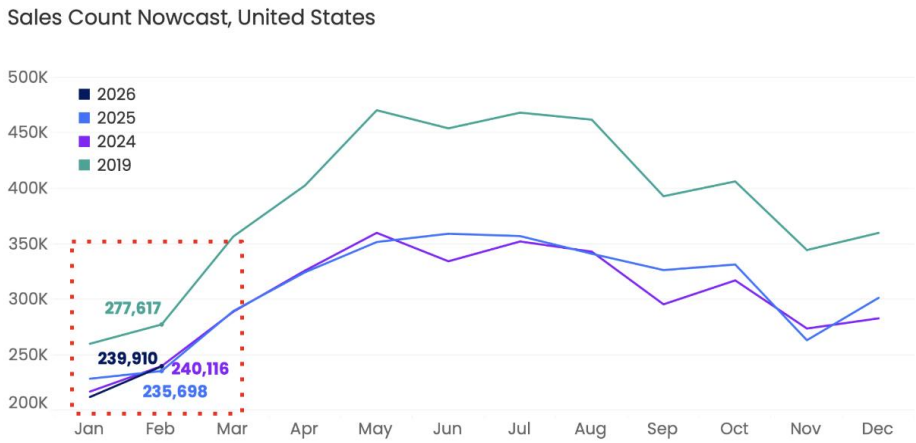
Two examples with this: when labeling fails, and an example from my work.

The National Housing Supply Landscape

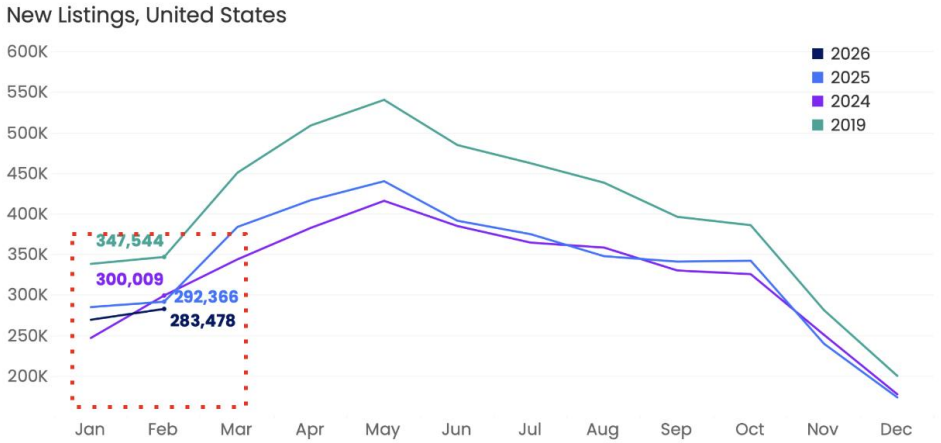
Mischa Fisher
Chief Economist, Zillow

Where is the housing market *today?*

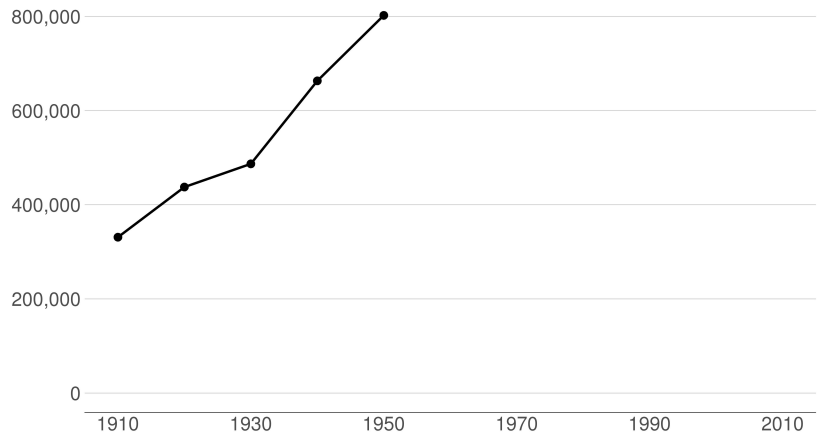
Buyers are showing signs of life



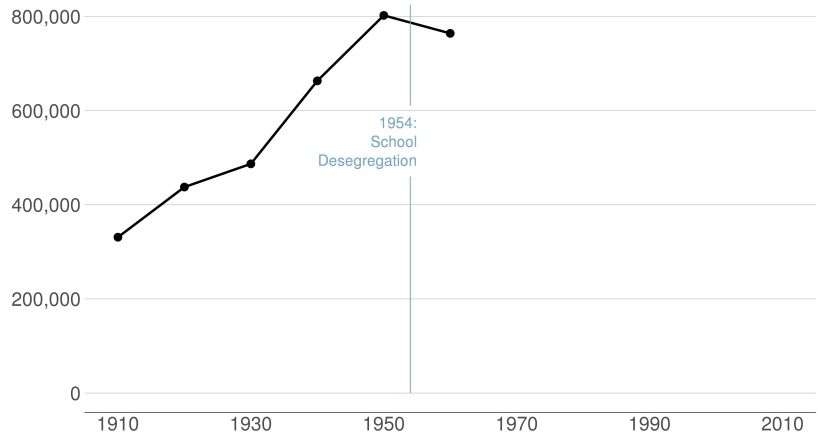
Sellers are slower to come back



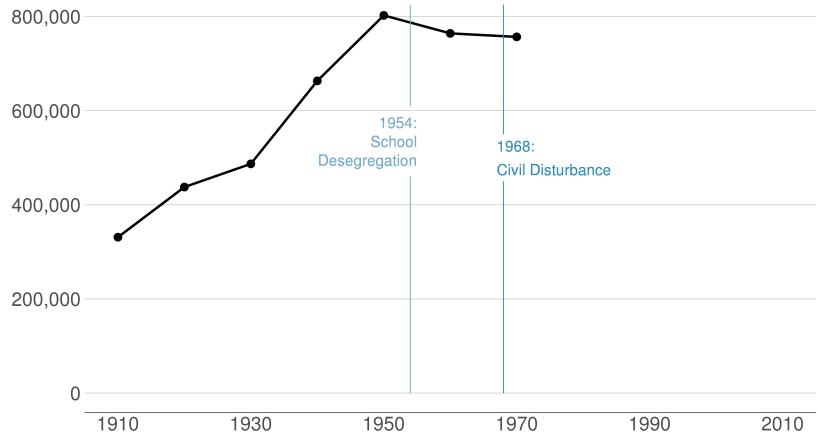
DC Gains Population Through 1950



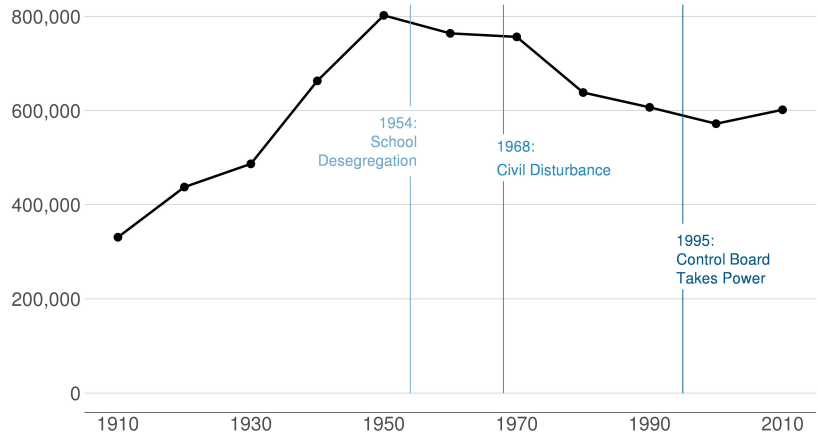
Population Loses Start with Desegregation



Continue After Civil Disturbance



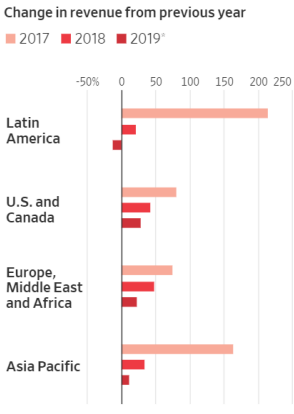
Population Turns Up After 2000



Something That Should be a Line Chart

Slower Ride

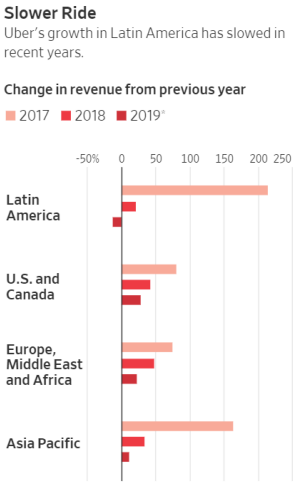
Uber's growth in Latin America has slowed in recent years.



*first nine months

Source: Uber's SEC filings

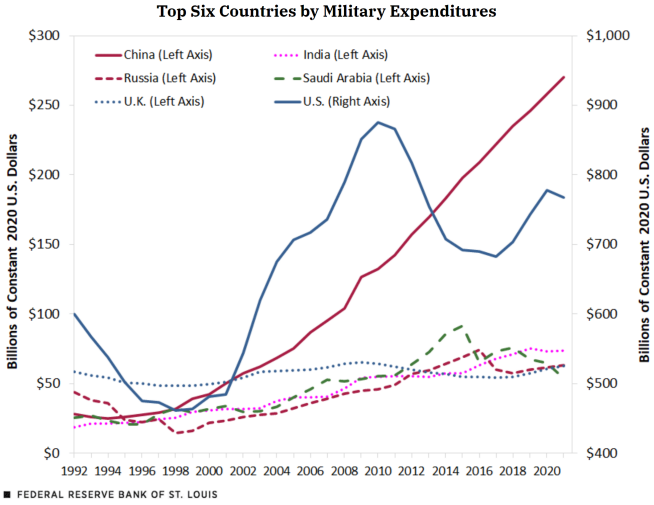
Something That Should be a Line Chart



- We use lines to show change over time
- Lines make pace of change obvious
- These bars have to point out years
- Vertical alignment of lines would show that they are the same year

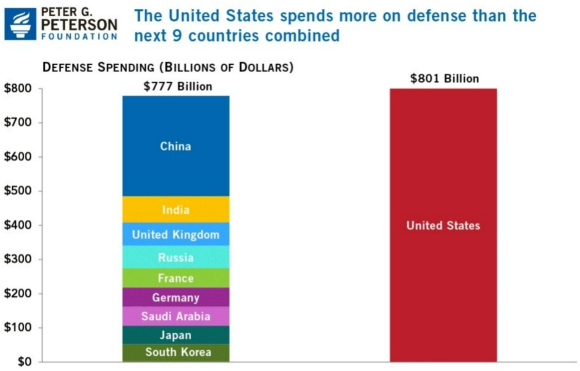
From this very interesting *WSJ* article about Soft Bank's funding of Uber and its competitors

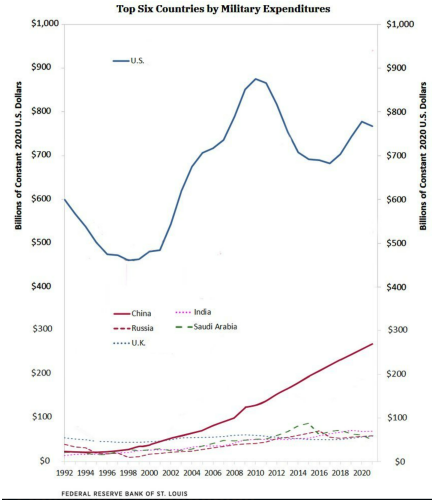
Look Carefully! What is Misleading?



With thanks to William Berry for bringing this to my attention. "Military Expenditures: How Do the Top-Spending Nations Compare?", January 3, 2023. [link]

Two Revisions





See the Twitter thread [here].

Moderating Lines: Inflation and Normalizing

Inflation

- A general increase in prices
- Sometimes we want to show dollar figures adjusting for inflation
- For example
  - 20 years ago your house was worth $100,000
  - Now it is worth $200,000
  - Does this increase exceed inflation?

Inflation

- A general increase in prices
- Sometimes we want to show dollar figures adjusting for inflation
- For example
  - 20 years ago your house was worth \$100,000
  - Now it is worth \$200,000
  - Does this increase exceed inflation?
- $\rightarrow$ adjust for inflation

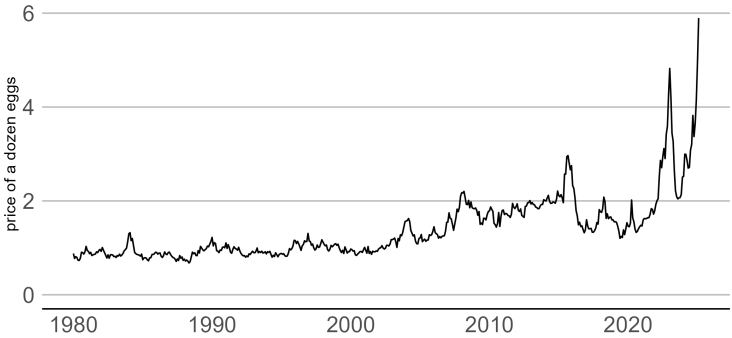
How to Adjust for Inflation?

Get data

- I recommend this page
- Use “all urban consumers”
- Download “text files” and load into R
- **Read instructions**
- Average monthly info

series_id	year	period	value	footnote_codes
CUSR0000SA0	1947	M01	21.48	
CUSR0000SA0	1947	M02	21.62	
CUSR0000SA0	1947	M03	22.00	
CUSR0000SA0	1947	M04	22.00	
CUSR0000SA0	1947	M05	21.95	
CUSR0000SA0	1947	M06	22.08	
CUSR0000SA0	1947	M07	22.23	
CUSR0000SA0	1947	M08	22.40	
CUSR0000SA0	1947	M09	22.84	
CUSR0000SA0	1947	M10	22.91	
CUSR0000SA0	1947	M11	23.06	
CUSR0000SA0	1947	M12	23.41	
CUSR0000SA0	1948	M01	23.68	
CUSR0000SA0	1948	M02	23.67	
CUSR0000SA0	1948	M03	23.50	
CUSR0000SA0	1948	M04	23.82	
CUSR0000SA0	1948	M05	24.01	
CUSR0000SA0	1948	M06	24.15	
CUSR0000SA0	1948	M07	24.40	

Nominal Price of Eggs



- to know if eggs increase more
- than overall prices
- we must adjust for inflation

Adjust to February 2025 Prices

	series_id	year	month	value
	<char>	<int>	<num>	<num>
1:	CUSR0000SA0	1980	1	78.000
2:	CUSR0000SA0	1980	2	79.000
3:	CUSR0000SA0	1980	3	80.100
4:	CUSR0000SA0	2025	1	319.086
5:	CUSR0000SA0	2025	2	319.775

- make things in 2025/02 terms
- requires scaling to 2025/02 value

Adjust to February 2025 Prices

	series_id	year	month	value
	<char>	<int>	<num>	<num>
1:	CUSR0000SA0	1980	1	78.000
2:	CUSR0000SA0	1980	2	79.000
3:	CUSR0000SA0	1980	3	80.100
4:	CUSR0000SA0	2025	1	319.086
5:	CUSR0000SA0	2025	2	319.775

- make things in 2025/02 terms
- requires scaling to 2025/02 value
- for each month, divide 2025/02 value by value

Adjust to February 2025 Prices

	series_id	year	month	value	base202502	relative
	<char>	<int>	<num>	<num>	<num>	<num>
1:	CUSR0000SA0	1980	1	78.000	319.775	4.099679
2:	CUSR0000SA0	1980	2	79.000	319.775	4.047785
3:	CUSR0000SA0	1980	3	80.100	319.775	3.992197
4:	CUSR0000SA0	2025	1	319.086	319.775	1.002159
5:	CUSR0000SA0	2025	2	319.775	319.775	1.000000

- `relative` tells us
- value relative to 2025/02
- 2025/02 must be 1

Adjust to February 2025 Prices

	series_id	year	month	value	base202502	relative
	<char>	<int>	<num>	<num>	<num>	<num>
1:	CUSR0000SA0	1980	1	78.000	319.775	4.099679
2:	CUSR0000SA0	1980	2	79.000	319.775	4.047785
3:	CUSR0000SA0	1980	3	80.100	319.775	3.992197
4:	CUSR0000SA0	2025	1	319.086	319.775	1.002159
5:	CUSR0000SA0	2025	2	319.775	319.775	1.000000

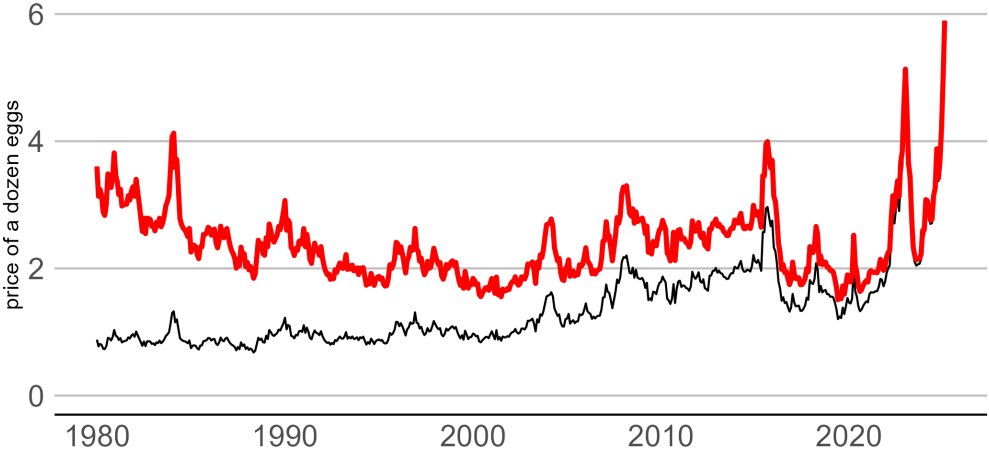
- `relative` tells us
- value relative to 2025/02
- 2025/02 must be 1
- merge with egg prices
- multiple `relative` by egg price

Adjusted Series

	series_id	year	month	value	base202502	relative	egg.price	eggs.adj
	<char>	<int>	<num>	<num>	<num>	<num>	<num>	<num>
1:	CUSR0000SA0	1980	1	78.000	319.775	4.099679	0.879	3.603618
2:	CUSR0000SA0	1980	2	79.000	319.775	4.047785	0.774	3.132985
3:	CUSR0000SA0	1980	3	80.100	319.775	3.992197	0.812	3.241664
4:	CUSR0000SA0	2025	1	319.086	319.775	1.002159	4.953	4.963695
5:	CUSR0000SA0	2025	2	319.775	319.775	1.000000	5.897	5.897000

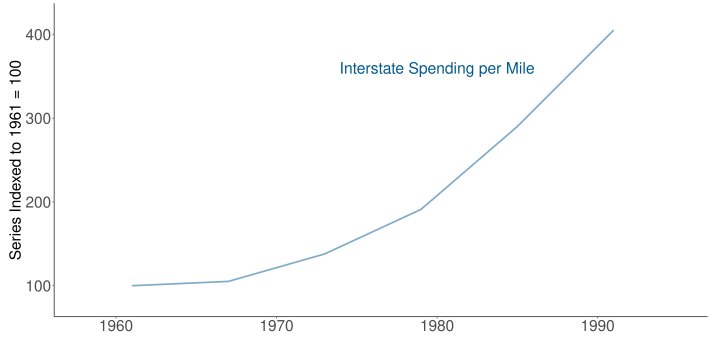
- Check!
- Are older prices higher?

Nominal and Real Egg Prices



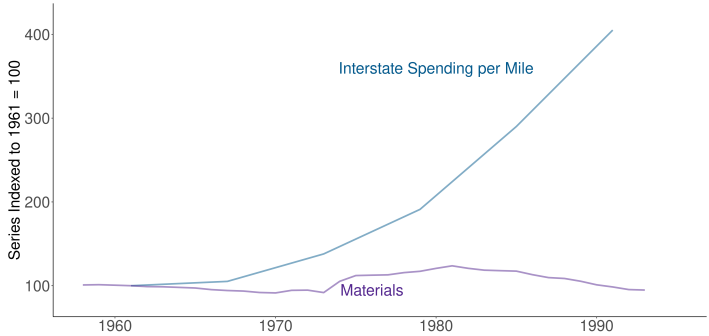
Normalizing: To Compare Series Over Time **Without** Two Axes

- Compare cost of highway mile
- To cost of highway inputs per unit



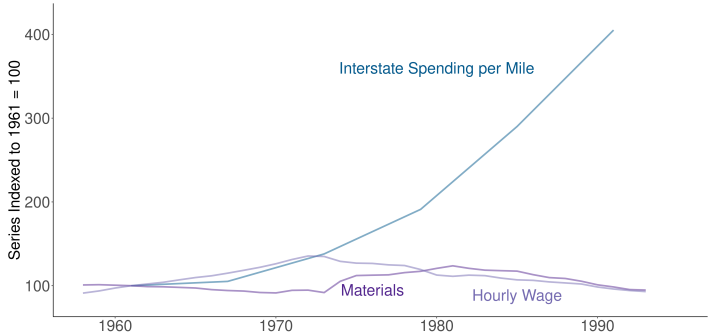
Normalizing: To Compare Series Over Time **Without** Two Axes

- Compare cost of highway mile
- To cost of highway inputs per unit



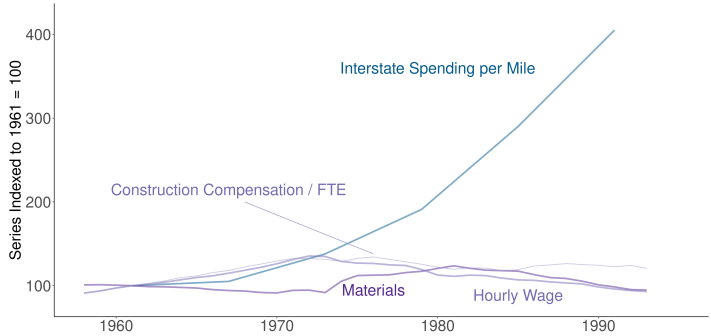
Normalizing: To Compare Series Over Time **Without** Two Axes

- Compare cost of highway mile
- To cost of highway inputs per unit



Normalizing: To Compare Series Over Time **Without** Two Axes

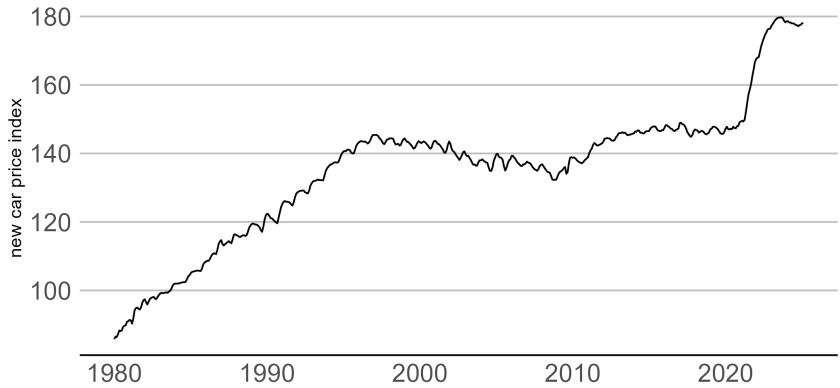
- Compare cost of highway mile
- To cost of highway inputs per unit



Which Changes Price More Since 1980: Eggs or Cars?

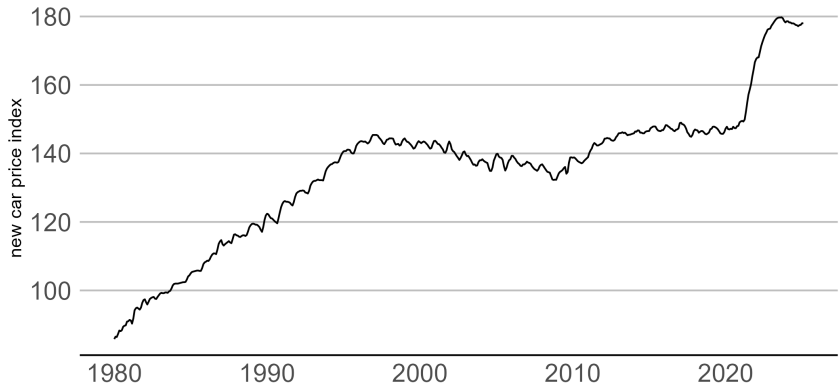
Which Changes Price More Since 1980: Eggs or Cars?

Cars price index



Which Changes Price More Since 1980: Eggs or Cars?

Cars price index



How can we show eggs and cars on the same graph?

Make Everything Relative to a Specific Date

- I choose Jan 1994
- NAFTA implementation date
- Merge egg and car prices
- Normalize each to 1 in Jan 1994

	observation_date	egg.price	car.price.index
	<IDat>	<num>	<num>
1:	1980-01-01	0.879	85.800
2:	1980-02-01	0.774	86.500
3:	1994-01-01	0.917	136.100
4:	1994-02-01	0.903	136.500
5:	2025-01-01	4.953	178.008
6:	2025-02-01	5.897	178.038

Make Everything Relative to a Specific Date

- I choose Jan 1994
- NAFTA implementation date
- Merge egg and car prices
- Normalize each to 1 in Jan 1994

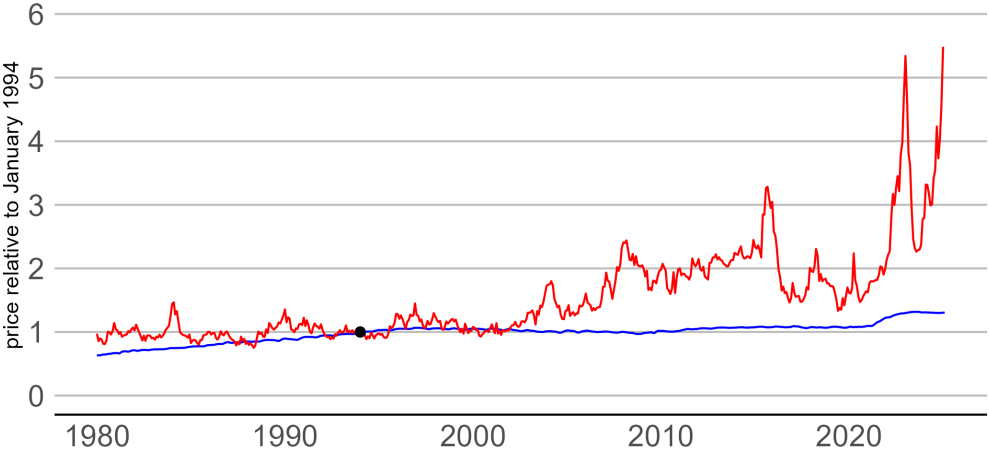
	observation_date	egg.price	egg.base	car.price.index	car.base
	<IDat>	<num>	<num>	<num>	<num>
1:	1980-01-01	0.879	0.903	85.800	136.5
2:	1980-02-01	0.774	0.903	86.500	136.5
3:	1994-01-01	0.917	0.903	136.100	136.5
4:	1994-02-01	0.903	0.903	136.500	136.5
5:	2025-01-01	4.953	0.903	178.008	136.5
6:	2025-02-01	5.897	0.903	178.038	136.5

Make Everything Relative to a Specific Date

- I choose Jan 1994
- NAFTA implementation date
- Merge egg and car prices
- Normalize each to 1 in Jan 1994

	observation_date	egg.price	egg.base	egg.94	car.price.index	car.base	car.94
	<IDat>	<num>	<num>	<num>	<num>	<num>	<num>
1:	1980-01-01	0.879	0.903	0.9734219	85.800	136.5	0.6285714
2:	1980-02-01	0.774	0.903	0.8571429	86.500	136.5	0.6336996
3:	1994-01-01	0.917	0.903	1.0155039	136.100	136.5	0.9970696
4:	1994-02-01	0.903	0.903	1.0000000	136.500	136.5	1.0000000
5:	2025-01-01	4.953	0.903	5.4850498	178.008	136.5	1.3040879
6:	2025-02-01	5.897	0.903	6.5304540	178.038	136.5	1.3043077

Inflation, Eggs, and Cars



Few on Stories

Chap 13: Telling Compelling Stories with Numbers

- Answer to "Is it a good chart?" depends on the story you're trying to tell
- The graphic can tell you about the story
- But the story can also lead you to the graphic
- Make sure you know the point that the graphic should make

Few's Components of a Compelling Story

- **Simple**
- Seamless
- Informative
- True
- **Contextual**
- Familiar
- Concrete
- Personal
- Emotional
- Actionable
- **Sequential**

Simple

- Always present the simplest possible version of your analysis first
- Summary statistics preferred to regression coefficients

Contextual

- Very important for magnitudes with which people are not familiar
- Helps us answer “so what” question
- Regression tables should have dependent variable means
- Visuals can put in context
  - dates
  - comparative categories
  - baseline mean
  - standard deviation

Contextual

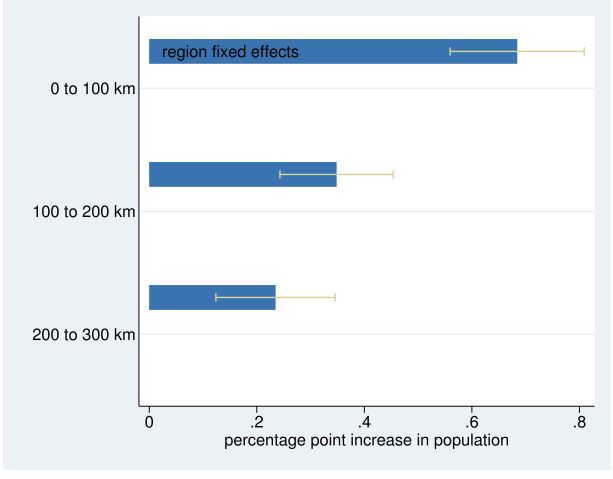
- Very important for magnitudes with which people are not familiar
- Helps us answer "so what" question
- Regression tables should have dependent variable means
- Visuals can put in context
  - dates
  - comparative categories
  - baseline mean
  - standard deviation

What does this mean for your policy brief?

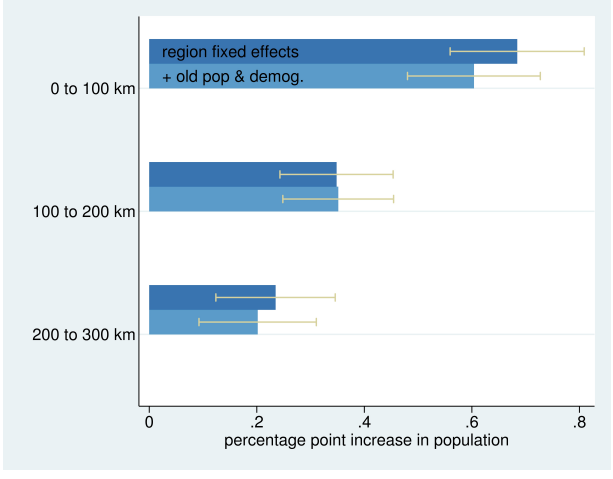
Sequential

- It is possible to present relatively complex graphics
- With proper groundwork
- Can be easer in a presentation than in a paper
- Paper/screen visuals need to be sequential differently
  - dance on screen vs dance in person

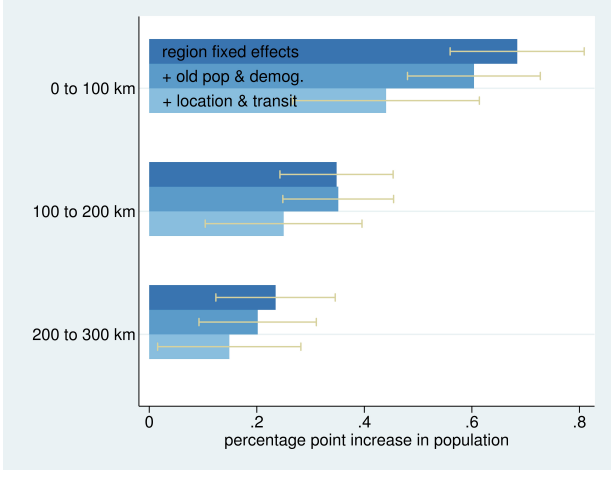
Bars with Error Bars, Building



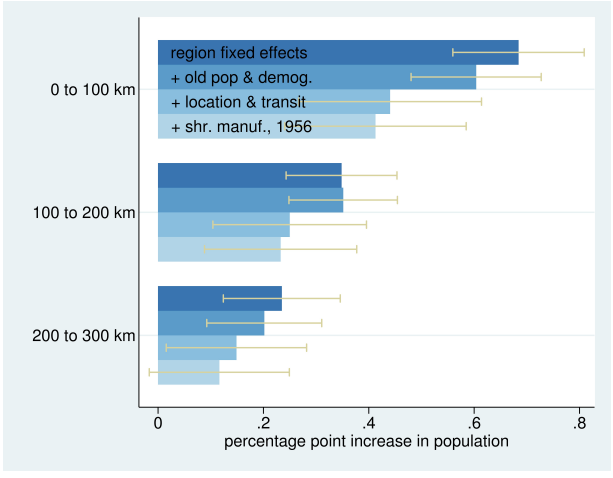
Bars with Error Bars, Building



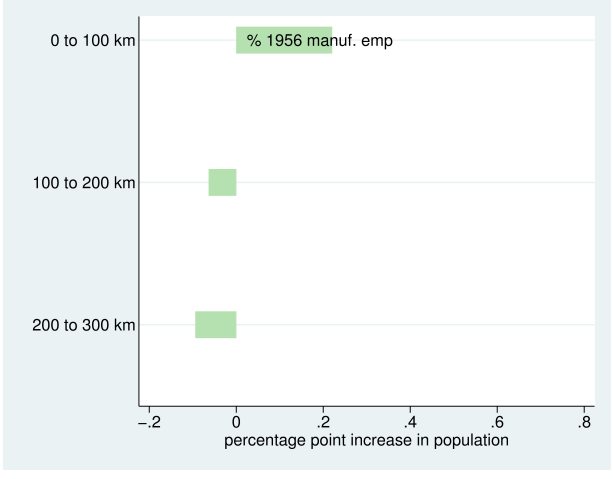
Bars with Error Bars, Building



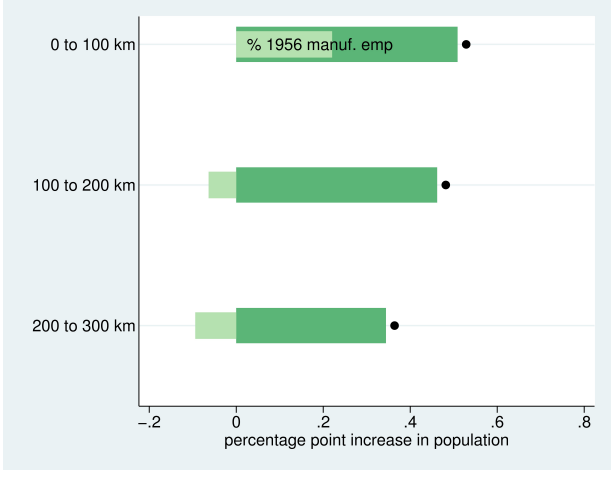
Bars with Error Bars, Building



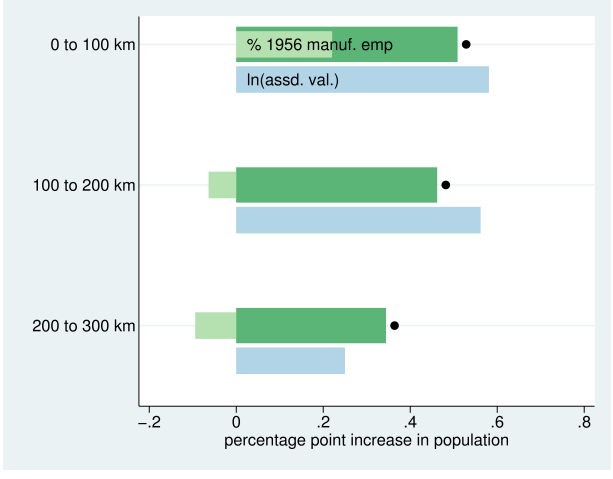
Interaction Effects



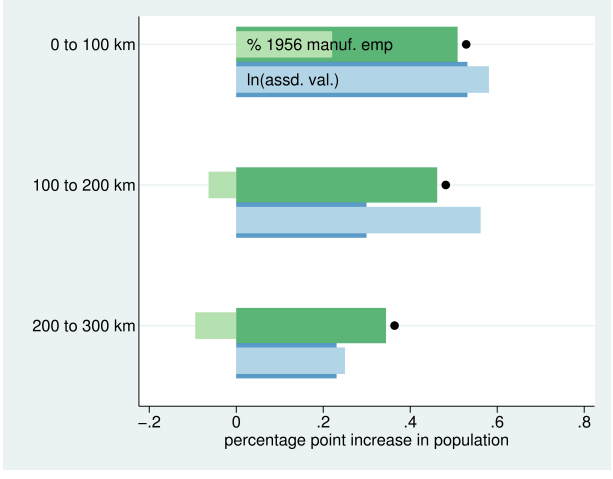
Interaction Effects



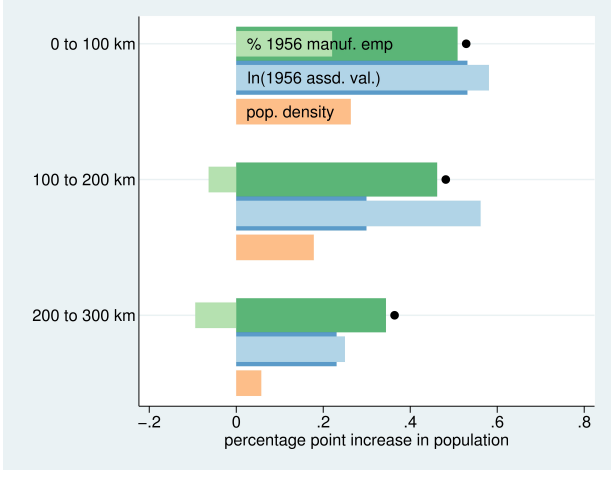
Interaction Effects



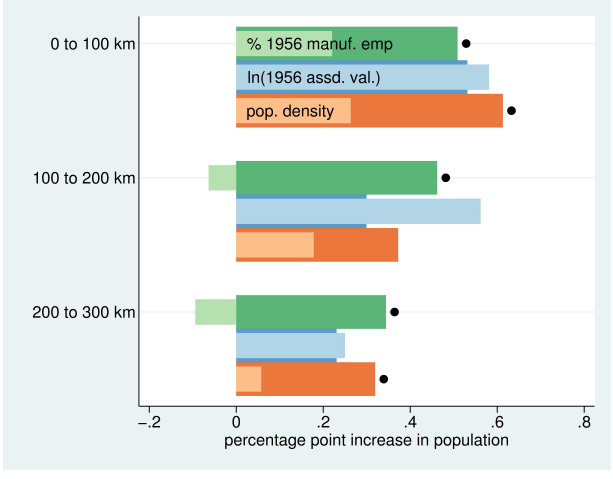
Interaction Effects



Interaction Effects

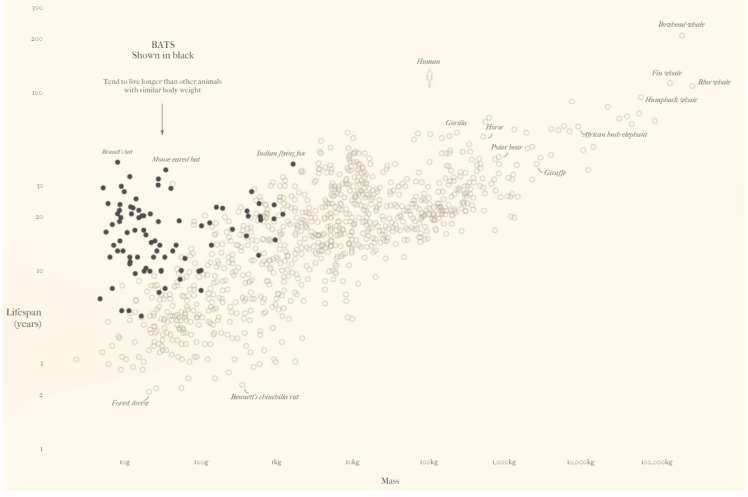


Interaction Effects



Annotations

A Great Use of Annotations



From Reuters

Today in R

Recap: Lines

- Lines are a great way to show change over time
- When using dollars, adjust for inflation unless you have a good reason not to
- Use normalization to compare more than one series when values are different
- Use on-graph annotations to point out differences

Next Lecture

- Next week: Scatters
- Next next week
  - Guest speaker
  - In-class workshop
  - We are likely to end at 5:20
- Post workshop material on time!
- I will be around to answer questions