

AI and data science jobs

Week 5, Lecture 9

<https://live.ds306.org>

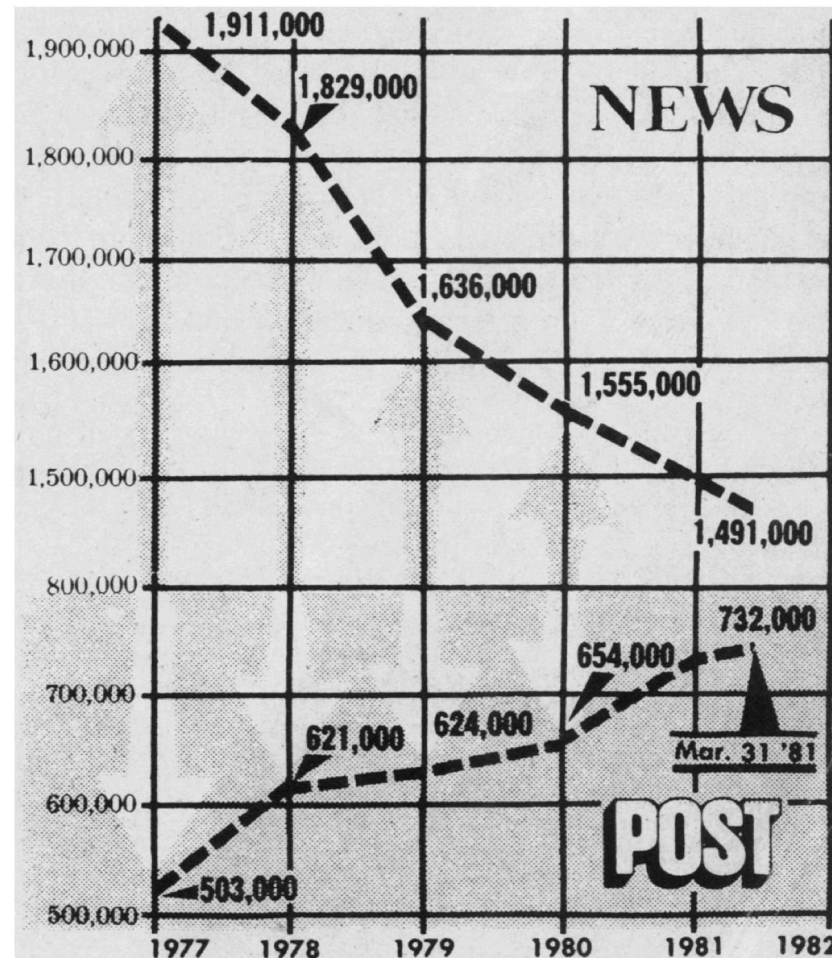
September 29, 2026

Bad plots

How much taller?



News vs. Post



What about this plot is misleading? How would you fix it?

Fix the plot

newspapers contains year, paper, and circulation.

```
newspapers |> head()  
  
# Replace this preview with your plot.
```

- Plot circulation against year, with a line for each newspaper and a shared, evenly spaced vertical scale.

On a common scale

Divide by 1,000,000 to express circulation in millions.

```
newspapers |>
  ggplot(aes(x = year, y = circulation / 1000000, color = paper)) +
  geom_line() +
  geom_point() +

  labs(x = "Year", y = "Circulation (millions)", color = "Newspaper")
```

What are recent trends in employment?

How has employment changed in the recent past for different age groups?

The data

- ADP processes payroll for employers. Stanford and ADP researchers studied employment by occupation and age.
- We'll use monthly employment summaries for software developers, from September 2021 through August 2026.
- (The publicly available data do not include a separate breakdown for data scientists.)

Source: [Canaries Dashboard](#).

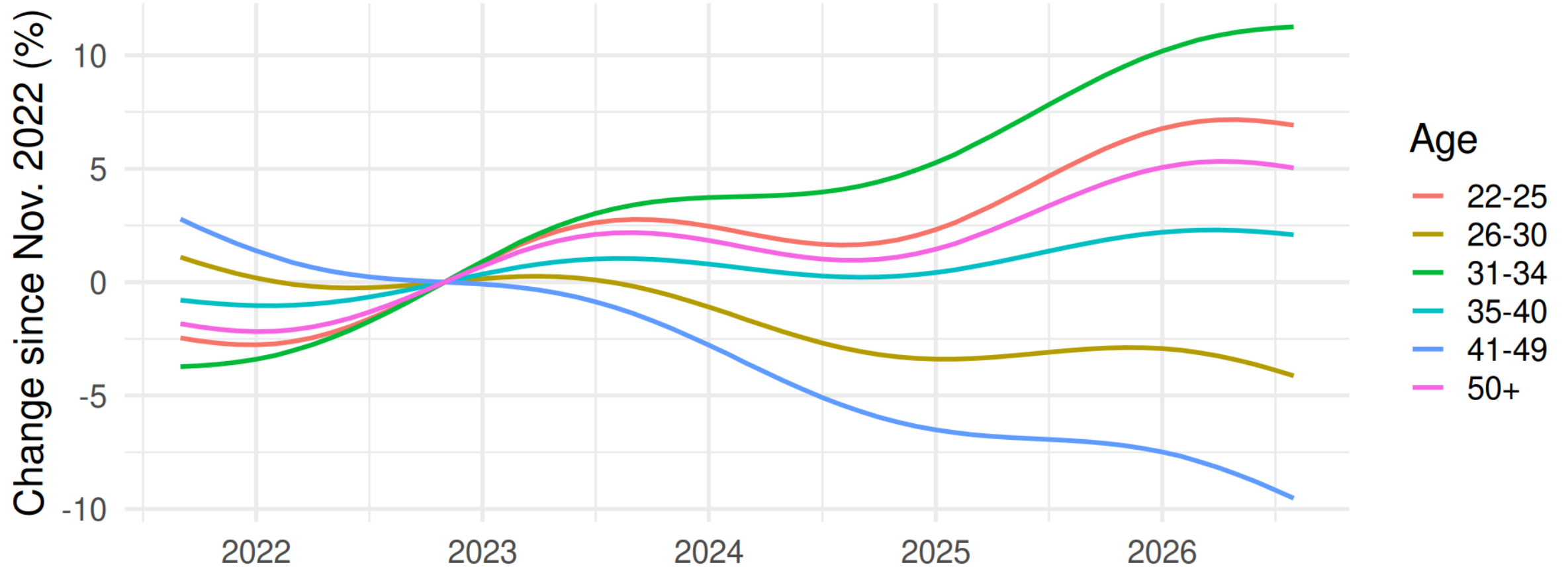
The dataset

```
software |> head()
```

- Each row is a month; each age column contains an employment **index**.
- An index of 100 is that group's November 2022 employment level.
- An index of 80 means employment is 20% below that level.

Objective

Follow employment over time, with one line for each age group.



How do we build this plot?

What columns would a table need to draw our target plot, and how could we get them from software?

What table would make that plot easy?

We need one column for age and one for employment:

```
developer_long |>
  filter(month == "2022-11-01") |>
  head(3)
```

Each row describes one month and age group.

Which parts of the original table became age and employment?

Plot employment by age

Use `pivot_longer()` to gather the six age columns, ``22-25`` : ``50+``.
(Backticks allow dashes and plus signs in column names.)

```
software  
  
# Build your pipeline one step at a time.
```

- Reshape the six age columns into age and employment, keeping month.
- Subtract 100 from employment to create `percent_change`.
- Plot percentage change over time with a separate line

All six age groups

```
developer_changes |>
  ggplot(aes(x = month, y = percent_change, color = age)) +
  geom_line() +
  # geom_vline() draws a vertical line; as.Date() converts text to a date.
  geom_vline(xintercept = as.Date("2022-11-30"),
             linetype = "dashed", color = "#0072B2") +

  # annotate() adds a label at the given position.
  annotate("text", x = as.Date("2022-11-30"), y = Inf,
          label = "ChatGPT released\nNov. 30, 2022",
          hjust = -0.1, vjust = 1.2, size = 4, color = "#0072B2") +

  labs(x = NULL, y = "Change since Nov. 2022 (%)", color = "Age",
       title = "Software-developer employment")
```

Interpreting the results

Does the age-group plot prove that ChatGPT caused changes in employment for young software developers? Why or why not?

Find the peak

When did employment among 22–25-year-old software developers peak? Does it coincide with ChatGPT's release?

```
software
```

```
# Build your pipeline one step at a time.
```

- Keep month and the employment index for ages 22–25. Name the index column employment.
- Sort the rows from highest to lowest employment.
- Keep the first row to show the peak

When was the peak?

```
software |>  
  select(month, employment = `22-25`) |>  
  arrange(desc(employment)) |>  
  head(1)
```

ChatGPT became publicly available in November 2022.

Did employment peak before or after ChatGPT's release? What does this tell us about the idea that ChatGPT started the decline?

Is this happening to other young adults?

Let's look beyond software developers.

Would you expect employment to have fallen more for young adults with a bachelor's degree or for those without one? Explain your prediction.

A second source of data

- Ozimek and Goldschlag analyzed the Current Population Survey, a survey of U.S. households.
- We'll look at ages 22–25, split by education, and ages 26–65.
- The **employment rate** is the percentage of people in a group who have a job.

Source: [EIG, March 5, 2026](#).

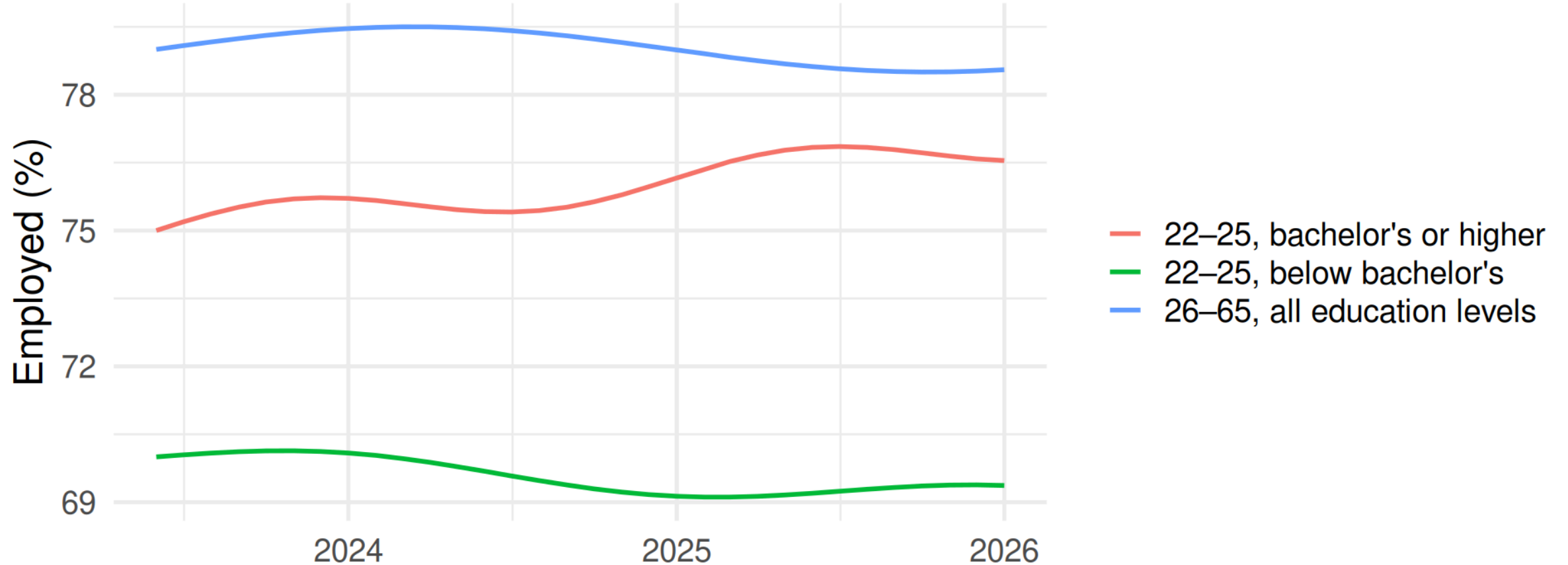
Inspect the table

```
youth |> head()
```

- Each row is a month, group, and population definition.
- employment_rate is a 12-month moving average.
- population is either "Everyone" or "Out of school".

Objective

Produce a plot that looks like:



Plot the three groups

Use youth to find out what happened to employment in each group.

```
youth
```

```
# Build your pipeline one step at a time.
```

- Keep observations from June 2023 onward.
- Keep the population labeled “Everyone”, retaining all three groups.
- Plot employment rate against month, with a separate line for each group.

What happened in each group?

```
youth |>  
  
  filter(month >= "2023-06-01", population == "Everyone") |>  
  
  ggplot(aes(x = month, y = employment_rate, color = group)) +  
  geom_line() +  
  
  labs(x = NULL, y = "Employed (%)", color = NULL)
```

Revisit your prediction

How does the graph fit your prediction? What does an explanation of the decline now need to account for?

Is COVID to blame?

One possibility is that employers expanded during the pandemic and then pulled back. What would that look like in the data?

Data on tech employment

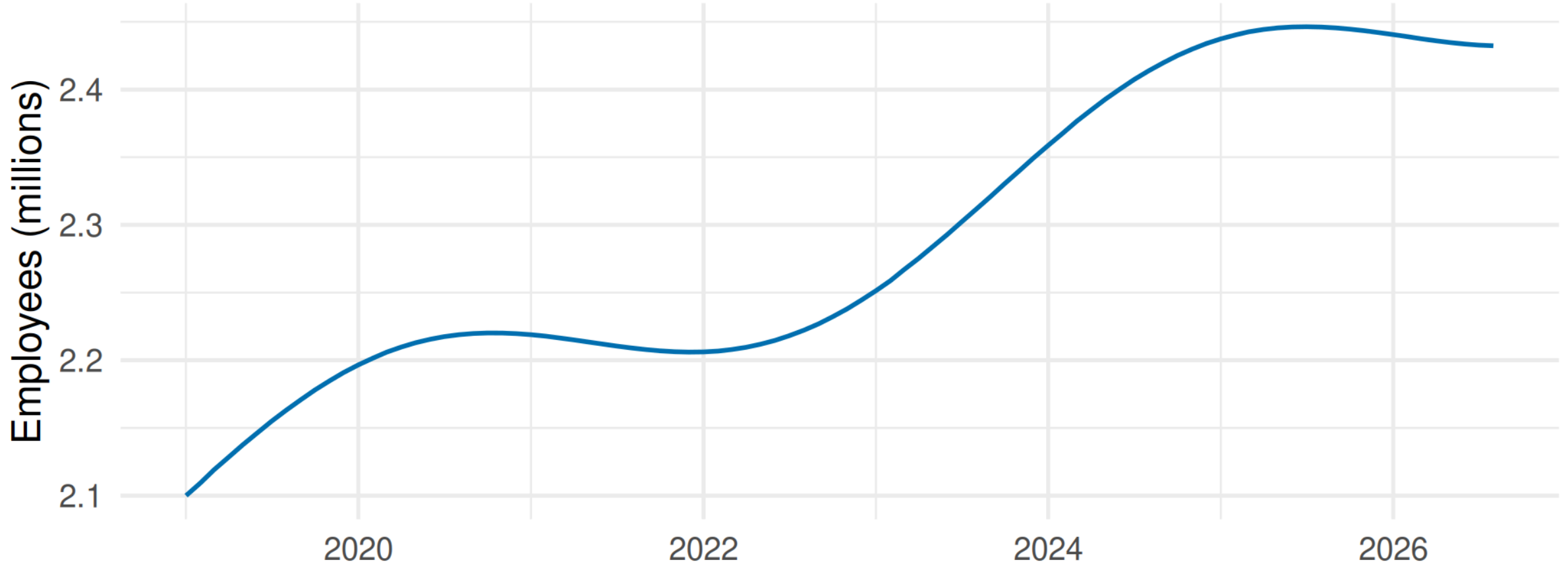
- labor_market contains monthly data from the Bureau of Labor Statistics.
- tech_jobs counts employees in computer systems design and related services, in thousands.
- This industry includes workers of all ages and occupations.

```
labor_market |>  
  select(month, tech_jobs) |>  
  head(4)
```

Source: [BLS](#), via [FRED](#).

Objective

How did employment change before, during, and after the pandemic?



Build the employment plot

Use `labor_market` to investigate whether the industry's expansion reversed.

```
labor_market  
  
# Build your pipeline one step at a time.
```

- Keep observations from January 2019 onward.
- Divide the employment values by 1,000 to express them in millions.
- Draw a line plot of employment in millions against month.

How much of the expansion reversed?

How well does your plot support the idea that employers are reversing a pandemic-era expansion? Is employment back to its pre-pandemic level?

Has hiring slowed too?

The same `labor_market` table contains monthly U.S. hiring and turnover rates.

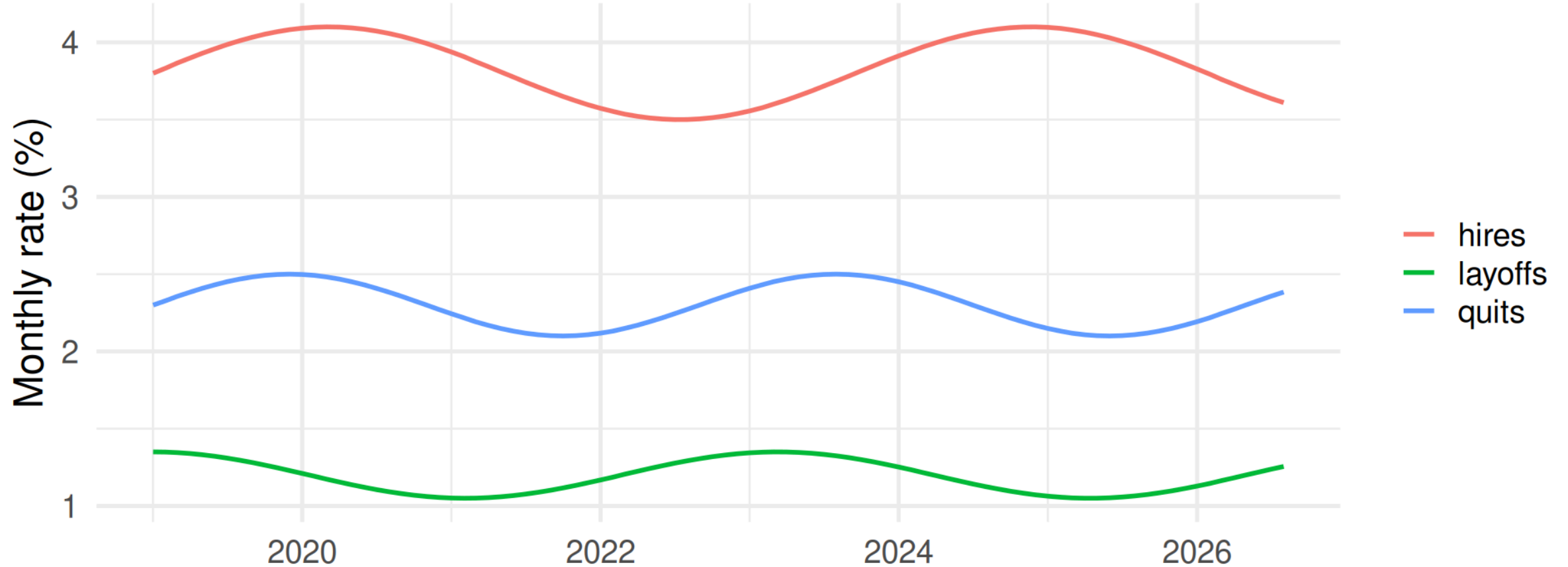
- `hires`, `quits`, and `layoffs` count monthly events per 100 employees.
- Quits are voluntary departures; layoffs includes discharges.
- These series cover nonfarm employment throughout the economy.

```
labor_market |>
  select(month, hires, quits, layoffs) |>
  head(4)
```

Source: [BLS Job Openings and Labor Turnover Survey](#).

Objective

Follow hiring, quits, and layoffs on the same plot.



Build the hiring plot

Use `labor_market` to see how hiring and departures changed.

```
labor_market  
  
# Build your pipeline one step at a time.
```

- Keep observations from January 2019 onward.
- Keep month, hires, quits, and layoffs.
- Gather the three rate columns into measure and rate, keeping month.
- Draw a line plot of rate against month, with a separate line

What does this mean for a first job?

Look at hiring, quits, and layoffs after 2022. How could these patterns make it hard to find a first job even if most people who have jobs keep them?

What would you tell a friend?

A friend says, “The job market is getting worse for young graduates because of AI.”

Use one result we found today to respond. What have we learned, and what would we still need to find out about AI’s role?

What did we learn?

- Employment fell among young software developers, while several older groups grew.
- In the household survey, employment rates fell among young adults with and without degrees.
- Tech-industry employment grew rapidly in 2021–22, then partly reversed that expansion.
- These patterns help describe the problem. We still need evidence about what caused it.

Sources and data

- Canaries Dashboard and methodology
- Researchers' hiring analysis
- EIG article and replication workbook
- Software-developer data and EIG employment rates
- BLS tech employment, hiring, quits, and layoffs; classroom CSV
- Data documentation
- Opening graphics from F25 Lecture 8; [newspaper figures](#) transcribed from the graphic.

ADP: through August 2026. EIG: through January 2026.

Week 5, Lecture 9: AI and data science jobs

Closed September 29, 2026 at 3:40 PM

16 class questions · 958 anonymous responses · 0 student questions

The following slides preserve what students submitted during class. No names or login information are included.

What about this figure is misleading? How would you fix it?

68 anonymous responses

Inconsistent scaling of figures

57

States that the people in the figure are scaled inconsistently or exaggerated relative to actual heights or each other.

Other responses

7

Missing axis or labels

4

Points out absence of a y-axis or axis labels and that the plot isn't drawn to scale, recommending clear axes/labels.

Student responses

1. 5 ft 2 looks much shorter than 5 ft 11. 4 people who are all 5 ft 2 all look different in the plot
2. All 5 foot two people have differnt heights on the plot, there is not scaling to all the heights it just shows an upward trend of heights of the women on the plot
3. All the people who are 5ft 2 are different heights in the picture. Additionally, none of the listed heights match the pictures which makes it seem like 5ft 7 is a much higher difference to 5ft 2
4. Because it just said the height no other measurement. I don't think it is good.
5. bro this is not to scale at all why is the woman on the right like 2x the size of nicki minaj if shes only 9 inches taller
6. Even though the shortes tand tallest person are only 9 inches apart they look like 2x as tall\, which is not proportional.
7. Even though there isn't a big numeric difference between the numbers, the heights are exaggerated to show a huge difference in heights, misleading the viewer of this plot

Student responses

1. For one, height is a quantitative value but isn't displayed on a scale. The pictures of the individuals are also arranged in a way that suggests they are to scale with each other, when they are clearly not.
2. Height not proportionately scaled,
3. I feel like the scaling is completely off. Like some people are both 5'2" and they are like several inches shorter in this diagram, so this really does not have the proper scaling that it should have. I would fix it by making a proper scale.
4. It doesn't have a set point of scale, it's not scaled or proportioned properly.
5. It is misleading because it is not proportional. You would think and picture the height differences to be much greater than they actually are, and there are not axes
6. it is misleading because the drawn figures are not scaled to the numbers. The labeled heights run only from about 1.54 m to 1.82 m, a small spread, but the images make some people look nearly twice as tall as others.
7. It is weird, the proportion is so bad. The difference in height is exaggerated

Student responses

1. It makes it seems like all the 5.2ft women are different heights, while the graphic scales up the women's heights terribly.
2. it shows them as so much shorter than in real life with not consistent picture
3. It's misleading because the heights listed are not comparable to the actual height differences that can be seen visually
4. lady is not half the height of other lady, heights are correlated but poorly represented by their real-world counterparts
5. Not accurately portraying heights with multiple numerical heights visually showing differences.
6. Not an even scale. Ariana and the shortest person are both 5'2 yet the figure makes them appear like they have a huge height distance. Comparatively, the jump from 5'2 to 5'7 is small.
7. Some of them look way larger although they're about the same size
8. The bodies were not stretched by the actual ratio of the height. We need to calculate the ratio and regraph it.
9. the bottom lady looks like half the height of the tallest lady even though there is only a 9 inch difference

Student responses

1. the difference between heights is greatly exaggerated in the graphic compared to the actual difference. To fix the figure adjust the sizes so they are more comparable to the real life differences.
2. the disparity between the shortest and the tallest vocalist on the boxplot makes it seem like there is a much bigger height difference than there actually is. visually inaccurate
3. The figure greatly exaggerates the height differences. Someone 1.82 m tall is only about 18% taller than someone 1.54 m tall, but the tallest figure appears roughly twice as tall. Even people labeled 1.55 m are drawn at different heights.
4. The figure is misleading because the people are scaled much more differently than their actual heights, so I would fix it by using the same scale for everyone.
5. The figure makes it seem as though this is their true size relatively even though they all start at about 5 feet and dramatize the difference after this.
6. The first three girls are all 5'2, but their apparent height differences are massive. the scale is not consistent

Student responses

1. the first three women are all 5'2 but they appear to be different heights. The entire graph starts at 5' for some reason? It should include the first 5 ft especially if they're using picture of the women rather than bars.
2. The four women on the left are the same height 5'2 but don't look the same height in the picture. Some have heels on and some don't. No axis or titles
3. The graphic is misleading because the scale makes it look like the height differences between all the women are greater than they actually are. For example, all of the first few women are 5'2, but the graphic makes it look like they differ in height.
4. The height difference looks bigger than it actually is since it's not to scale
5. The height of these people are not proportional to the actual numerical values
6. The people aren't scaled correctly.
7. The people in the graph are not on the same scale as each other. The first four women are the same height, but are shown to be different heights.
8. The proportions are off, so the 5.2 girl looks twice as short
9. The proportions do not actually represent the differences in values.

Student responses

1. The ratio and demonstration of 1 inch difference in height is not the same. I would make the 1 inch difference between all the women to be the same so that it is visually more accurate
2. The ratio of the difference in height to the absolute height of each individual is inaccurate. This also makes it seem as though body parts scale with height 1:1.
3. The scale between different heights isn't consistent. Some of the 5.2 ft heights are different heights despite both being 5.2
4. The scale is not accurate.
5. The scale is super off, not all the figures are standing on a level baseline (some wearing heels)
6. The scale is terrible, and makes the 5ft2 person look super short compared to the 5ft11 person.
7. The scale of each image is not representative of the true difference in heights between the women shown
8. The scaling and proportions are completely off in this plot. Between different heights, there is no meaningful reason why there is such variation between how the women are presented compared to the height of those before and after them.

Student responses

1. The scaling of the plot is very misleading and the heights seem uncomparable in a common sense
2. the sizes arent proportional
3. These women aren't shown in the same proportion. Noticeably, the backside of the "tallest" woman is larger than the head of the fourth woman (which looks like ariana grande). I would adjust the sizing so all the women are displayed in the same size, then order their heights.
4. They are not shown to scale so even though the first one is less than a foot shorter than the last one she is shown as about half the size
5. They aren't on the same scale. 5'11" is only 9 inches taller than 5'2", but it looks like it's double the height on this plot.
6. This figure is misleading because the proportions for the figure's height differences do not properly scale with the sizes mentioned below.
7. This figure is not scaling. The shortest girl has 1.54m while the tallest one is 1.82m. The difference between is less than a half meter, but it looks like the shortest girl is only a half the height of the tallest from the figure.

Student responses

1. This is a bad plot because it seems like the 5'11 women is double the height of the smallest one so the plot is not to scale
2. This is a bad plot because there is no scale. Four of the artists are 5 foot 2 and yet there is a large difference between them, that doesnt really make any sense.
3. This plot has scaled down the size of the people that are shorter instead of just their height
4. those are not accurate scaling of the size of the people

Student responses

1. abstract
2. Higher bars are also wider, which artificially heightens the difference in size
3. It's not consistent growth, you skip all of the heights between 5'2 and 5'7.
4. people are getting bigger both horizontal and vertical
5. The plot is going based off of the meter which is why the heights still look different despite some of these women being around the same height. the difference is too much to actually show the real heights of these people. it should just be based off feet and inches since that is a more general unit
6. the scale makes niki minaj look half the height of iggy azalia, even though they have a 9 in height difference, (the scale is off, the y axis doesn't start at zero)
7. The women are plotted like these are their actual relative heights (which is not true). If we plotted this as like a line plot then this y-axis would probably start at like five feet or something like that, we could fix it by making it start at 0.

Student responses

1. Plot doesn't show any y axis, so I don't know exactly what the scale is. Also, how are these people different from one another?
2. there is no y-axis or y-axis label. it also is not "drawn" to scale.
3. There's no title and no y axis to properly compare. also proportion is way off.
4. Uses graphical information that does not fully represent the actual height of the female vocalists. ALso no x, y axis and titles.

What about this plot is misleading? How would you fix it?

72 anonymous responses

Misleading y-axis scale

50

Notes that the y-axis jumps large values or uses an inconsistent/unnatural scale, exaggerating or hiding differences; suggests using a continuous/regular y-axis scale.

Other responses

9

Missing labels or legend

8

Points out absence of labels for lines or axes, recommending adding clear labels and a legend so viewers know what each line represents.

Unclear what plot shows

5

Expresses uncertainty about the plot's meaning or purpose without specific suggestions; indicates the plot lacks clarity.

Student responses

1. Also the proportion problem, we don't have the instinct sense of the distances between different point.
2. Axes are not labeled. Have no clue what is being measures. Y axis is not proportionally going up with a jump from 800,000 to 1,500,000.
3. Confusing what the two lines represent
4. i can't even tell what this plot is trying to say. the y-axis scale is uneven as well
5. It doesn't start at zero and the y scale isn't consistent. There is a jump between 800,000 and 1.5 million.
6. it is not clear what each axis is for, there are no labels
7. It starts at 500,000 rather than 0, leading to believe that the initial intercept is quite low.
8. It starts at 500,000.
9. It suddenly jumps between 800,000 and 1,500,000 on the Y axis.

Student responses

1. It's two different plots with the same y axis, which is really confusing.
2. left scale incorrectly jumps between numbers at an inconsistent rate. Goes from 800,000 to 1.5mil at some point when previous jumps were only by 100,000.
3. Same thing where the axis starts not at zero. Fixed in the same way, by starting the y-axis at zero.
4. The distance between 1500000 and 800000 is mistakenly represented as the same as the other vertical distances.
5. The distances between the numbers on the y axis are not even and jump in size from 100,000, to 700,000
6. The gap between 800,000 and 1.5 million is the same gap as 600,000 and 700,000, so it massively undershoots the gap between post and news size.
7. The graph doesn't start at 0 on the y axis, so the difference looks bigger than it actually is
8. The interval between 800000 and 1500000 is different from other intervals. We can either plot them on 2 graphs or present the actual intervals between these 2 measurements.

Student responses

1. The jump in the y axis from 800,000 to 1.5 million manipulates the graph to suggest that Post readership is far closer to NEWS than in reality. A consistent y-axis would clear this up.
2. The numbers don't correspond to the y-axis. The scale is also uneven
3. the plot does not use consistent changes in measurement for the y axis, jumping from 800,000 to 1,500,000
4. The plot is misleading because the y-axis skips from 800,000 to 1,500,000, which makes the two lines look closer together, so I would use a consistent scale on the y-axis.
5. The scale changes in the middle of the Y-axis and the squares don't change size
6. the scale goes from 1.5 milliion to 800k
7. The scale is not consistent; it starts in increments of 100,000 on the y axis and then jumps from 800,000 to 1,500,000
8. The unit of the height of each block is inconsistent. There is a surge from 800,000 to 1,500,000 within a block.
9. the vertical axis of the graph skips from 800K to 1.5Million, there is no standard scale system (specific increments) I would start by have a vertical axis with consistent increments

Student responses

1. The vertical scale has grids per 100,000 units except close to the middle, where it suddenly jumps from 800000 to 1500000 to make the gap between the two measurement look much closer.
2. The vertical scale is misleading because equal numerical differences are not represented by equal vertical distances.
3. The vertical scale skips from 800,000 to 1,500,000 in the same space used elsewhere for an increase of 100,000, without marking a break.
4. The Y axis is not evenly spaced
5. the y axis is very inconstant, it skips 900,000 to 1,400,000. I think the post ran the plot to make it look like they were catching up to the news
6. The y axis units skips a whole bunch. ALso unclear title/axis
7. the y units are in increments of 100,000 and then for one spot it just jumps up 700,000
8. The y-axis does not have consistent scaling
9. The y-axis does not start from 0, and the differences between the marked values are not always the same (for example, the graph goes from 700,000 to 800,000 then all the way to 1,500,000), which makes it look like the two lines are closer than they actually are.

Student responses

1. the y-axis doesn't start at 0 and its very confusing you can't really see whats going on.
2. the y-axis is broken and uneven. It jumps from the 500,000 to 800,000 band straight to the 1,500,000 to 1,900,000 band, and the gridlines are not evenly spaced,
3. The y-axis starts at 500,000 instead of 0. It makes the difference between the values at year = 1977 much larger than compared to absolute 0.
4. There are 2 lines converging, not sure what each is referring to
5. There is a huge gap in the graph between 800,000 and 1,500,000, making the graph look more normal.
6. There is a HUGE jump in the y-axis from 1,500,000 to 800,000. The rest of the boxes properly scale for every 100,000, but that one jumps unproportionally. To make this plot accurate, I'd advise changing the benchmarks so they are all consistent or adding in many boxes to fill in the mentioned gap.
7. there is a massive gap between the two trendlines that the graph does not show so it appears to be closer than it actually is
8. There is break on the y axis, making it seem like the two lines are way closer than they actually are

Student responses

1. There's a huge jump on the y-axis between 800,000 and 1,500,000 implying a much closer relationship between the two lines when in fact, the top line is still vastly outnumbering the bottom one, ran by the POST probably to exaggerate their growth, could adjust the y-axis to actually have even steps
2. This is a misleading plot because of the scale that is used on the y-axis. It does not go up in even increments but instead jumps from 800,000 to 1,500,000. This then misleads us to think the difference in 1981 is not as large as it actually was. You would need to resize the y-axis so that it goes up in equal increments.
3. This plot is not explicitly saying which data is from news vs post and has unclear labeling
4. uneven spacing on the y-axis can distort the data's appearance
5. Vertical scale is not right
6. while post has been increasing it does not really highlights that and just highlights that it is so much lower than news
7. You don't know what each of the two different lines mean, there is no labeling of the axes and no title either

Student responses

1. halfway through on the y axis, it just jumps 700k
2. The trend is actually higher, but it seems means the news one is worse than the post one.
3. The two lines are the same, one can be solid and keep one dashed, it also skips a lot of numbers on the y-axis.
4. there shouldnt be 2 points in the same column for something like this
5. This doesn't graph a function; not sure what it's actually trying to communicate without a key. Perhaps it's two different trends that use the same y-axis but are not sufficiently differentiated?
6. This plot is misleading because it has 2 lines and doesn't explain what the difference between them. The post ran it.
7. two lines on the same plot
8. w
9. we dont know what the y axis represents, and the y axis also does not start at 0, so the decline for the top line might not be as significant as it actually is compared to the incline for the bottom line since the bottom line starts at 500,000

Student responses

1. Hard to determine a trend, don't know which line represents what. What is describing the decreasing line and what is described the increasing line trend.
2. I would add labels and maybe split it in half
3. The axes are not labelled. We know one line is news and one is post, but we don't know exactly what is being measured here. Also the y-axis has an uneven scale, with the gap between 800,000 and 1,500,000 being the same size as 100,000
4. The lines aren't labeled. Additionally, the numbers presented aren't accurate to what the line graph is showing.
5. There are no units listed, the scale is extremely large, and the two lines have the same dashed pattern
6. There is no label on the y axis I have no idea what the trends are and no explanation on why there are two lines
7. What is the y axis title? I don't know what these quantities represent. Also, why do x values have two y values associated with them? What are these two lines that are plotted?
8. what is x row and y row? what is news?

Student responses

1. axis scaling
2. I have no idea what the heck this grpah is talking about and some numbers are missing pointers so idk which point they're talking about
3. I think the misleading part might be the paths? like it looks like they might converege but there is not enough information to tell me what is converging and also whether these two lines are relevant enought to converge
4. it is hard to differentiate news and post, it looks like just one line that moves forward and backwards.
5. The scales are misleading and we can't really tell what each trend means.

- Plot circulation against year, with a line for each newspaper and a shared, evenly spaced vertical scale.

63 anonymous responses

Other responses

63

Student responses

1. `ggplot(newspapers, aes(x = year, y = circulation, colour = paper)) + geom_line() + scale_y_continuous(labels = scales::label_comma())`
2. `library(ggplot2) newspapers |> ggplot(aes(x = year, y = circulation, color = paper, group = paper)) + geom_line(linewidth = 1) + geom_point() + scale_y_continuous(limits = c(0, NA), labels = scales::label_comma()) + labs(x = "Year", y = "Circulation", color = "Newspaper") + theme_minimal() # Replace this preview with your plot.`
3. `newspapers`
4. `newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.`
5. `newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.`
6. `newspapers |> ggplot(aes(x = year, y = circulation, color = paper, group = paper)) + geom_line() + labs(x = "Year", y = "Circulation")`
7. `newspapers |> ggplot(aes(x = year, y = circulation, color = paper)) + geom_line()`

Student responses

1. newspapers |> ggplot(aes(x = year, y = circulation, color = paper)) + geom_line() + labs(x = "Year", y = "Circulation", color = "Newspaper") # Replace this preview with your plot.
2. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
3. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
4. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
5. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
6. newspapers |> ggplot() + geom_line(aes(x=year, y = circulation, color = paper)) # Replace this preview with your plot.
7. newspapers |> ggplot()+ geom_line(aes(x=year, y=circulation, color = paper)) # Replace this preview with your plot.

Student responses

1. `newspapers |> ggplot(aes(x = year, y = circulation / 1000000, color = paper )) + geom_line() + geom_point() + labs( x = "Year", y = "Circulation (millions)", color = "Newspaper" )` # Replace this preview with your plot.
2. `newspapers |> ggplot(aes(x = year, y = circulation, color = paper, group = paper)) + geom_line() + scale_y_continuous() + labs(x = "Year", y = "Circulation", color = "Newspaper")`
3. `newspapers |> ggplot() + geom_line(aes(x= year, y = circulation, color = paper))` # Replace this preview with your plot.
4. `newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper))` # Replace this preview with your plot.
5. `newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper))` # Replace this preview with your plot.
6. `newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)` # Replace this preview with your plot.

Student responses

1. newspapers |> ggplot(data, aes(x = year, y = circulation, color = paper)) + geom_line() # Replace this preview with your plot.
2. newspapers |> ggplot + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
3. newspapers |> ggplot() + geom_line(aes(x=year, y= circulation, color = paper)) # Replace this preview with your plot.
4. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
5. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper))
6. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
7. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
8. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.

Student responses

1. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper))
2. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
3. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper))
4. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
5. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
6. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
7. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
8. newspapers |> ggplot() + geom_line(aes(x = year, y=circulation, color=paper))
9. newspapers |> ggplot() + geom_line(aes(x=year, y = circulation, color = paper)) # Replace this preview with your plot.

Student responses

1. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
2. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
3. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
4. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
5. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
6. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
7. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.

Student responses

1. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
2. newspapers |> ggplot() + geom_line(aes(x=year, y = circulation, color = paper)) # Replace this preview with your plot.
3. newspapers |> ggplot() + geom_line(aes(x=year,y=circulation,color=paper)) # Replace this preview with your plot.
4. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
5. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
6. newspapers |> ggplot() + geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.
7. newspapers |> ggplot() + geom_line(aes(x = year, y= circulation, color = paper)) # Replace this preview with your plot.
8. newspapers |> ggplot() + geom_line(aes(x=year, y=circulation, color =paper)) # Replace this preview with your plot.
9. newspapers |> ggplot() + geom_line(aes(x=year, y=circulation, color=paper)) # Replace this preview with your plot.

Student responses

1. `newspapers |> ggplot()+ geom_line(aes(x = year, y = circulation, color = paper)) # Replace this preview with your plot.`
2. `newspapers |> ggplot()+ geom_line(aes(x=year, y=circulation, color=paper)) # Replace this preview with your plot.`
3. `newspapers |> ggplot(newspapers, aes(x = year, y = circulation, color = paper)) + geom_line()`
4. `newspapers |> head() # Replace this preview with your plot.`
5. `newspapers |> head() # Replace this preview with your plot.`
6. `newspapers |> head() # Replace this preview with your plot.`
7. `newspapers |> head() # Replace this preview with your plot. newspapers |> ggplot() + geom_line(aes(x=year, y = circulation, color = paper))`
8. `newspapers |> head() ggplot(newspapers, aes(x = year, y = circulation, group = paper, color = paper)) + geom_line() # Replace this preview with your plot.`
9. `newspapers |> head() ggplot(newspapers, aes(x = year, y = circulation, color = paper, group = paper)) + geom_line() + labs(x = "Year", y = "Circulation", color = "Newspaper") # Replace this preview with your plot.`

Student responses

1. `newspapers |> head()` `ggplot(newspapers, aes(x = year, y = circulation, color = paper)) + geom_line() + scale_y_continuous( breaks = seq(0, max(newspapers$circulation, na.rm = TRUE), by = 100000) )` # Replace this preview with your plot.

What columns would a table need to draw our target plot, and how could we get them from `software`?

74 anonymous responses

Compute change from baseline

24

Response proposing a derived column representing change since a baseline (e.g., since Nov 2022) for each age group and month.

Age groups and employment values

17

Responses saying the table needs age (or age groups) and employment measures as columns.

Plot each row/column as a line

8

Response suggesting plotting directly from rows or drawing one line per column without specifying specific needed columns.

Group by age and plot daily rates

5

Responses recommending grouping by age and plotting the employment rate at a daily frequency.

Response themes for question 4

Index levels per group for each month

5

Response saying the table needs index-level values for every group for each month (multi-index / group+month values).

Indexes as columns to draw plot

5

Response stating that index values should be columns to enable drawing the target plot.

Percentage by age and year/month

5

Responses identifying percentage, age groups, and time (years or months) as required columns.

Reshape to long format before plotting

5

Responses recommending pivoting or using long (tidy) format so age groups become rows/values before graphing.

Student responses

1. A table need month (which we already have), age group and the index compared with Nov 2022. For age group, we would use `pivot_longer()` function from software to form a new column based on the column names of software. For index, we need to mutate a third group containing percentage of the change in each period compared with Nov 2022.
2. A table would need the columns age range, the date, and the employment level relative to nov. 01 of 2022. We'd need to subtract 100 from the change in employment and create lines for each age group.
3. A table would need to have a change column, which we could get by calculating the difference in employment numbers for each month. It also requires a month and age column.
4. A table would need to have reference values for employment rates for a given age bracket in 2022. Then, a new column would need to be created so that the current year's difference from 2022's employment rate for that age bracket is stored. These difference values are then plotted against the year.
5. Change in employment in reference to 2022, subtracting 100 from each of the employment thingys

Student responses

1. change since 2022 for each age period for different time
2. difference from nov, can calculate a new column
3. salary now, salary before, we can get them by computing the change
4. set the x axis to be time and then mutate and use a col to be y axis to be the difference between that month and nov
5. They would need all the columns of the age groups mutated to show the percent change since Nov 2022
6. To draw the target plot, the table would need a column calculating the difference between that month's index and the reference from November of 2022.
7. We need the % change from the base year, so we need to make a new column by taking the index - 100 to see the difference. We already have the dates.
8. We would need a column showing the difference between our chosen year and the year we are comparing it too. Then we would need a column for percentage.
9. we would need all the age groups and plot those by subtracting by a 100 and then on the x-axis we would just put the prgression of the years using the months.

Student responses

1. We would need to calculate the percent difference from the values from november 2022 and then plot it.
2. We would need to create a column with the difference between each year and the year we want to compare against, perhaps through mutating a new column to get the required amount compared to the current year subtracted from the 2022-11-01, as well as reshaping the data longer
3. We would need to create a new column that can calculate what the difference is from 2022. We could do this by creating a new column and doing a subtraction of november results and that month. We would need to do this for every age group too.
4. Year (or date), age group, and change in employment. We can group by age and time, summarize employment, and calculate the change from Nov. 2022.
5. year, age, and change since november we could filter and mutate to create the variables we want
6. You need a column that has the difference in employment rates for each month and age range.
7. you would need a column of change from the starting date to each reading for each age group and then you could plot them

Student responses

1. You would need a different column for each age group and then the values would need to be the change in employment since 2022.
2. You would need each age group column for employment relative to 2022 and minus them by 100 then multiply by 100 to get percent change. Make them new columns and use them to plot
3. You would need the change since November 2022 for each age group for each month (so current number minus 100)

Student responses

1. age, employment, month, percent change. percent change can be found from employment data
2. Employment over the years and the base rate from 2022
3. It would need all of them, from month to all the age columns
4. want month, age, and employment, so long format
5. we need different months and each age range. We can actually use different programs to create them
6. We need to categorize by age and condense the months together to avoid having too many data points.
7. we would need the age groups, years, and the change amount, using filter and group by.
8. we would need the employment index
9. We would need to get age employment and year

Student responses

1. we would use data from columns that are the age groups and then plot the percentage of those age groups that are working linearly
2. we would want each of the age groups to draw the target plot and we can access those through software, we can pivot longer to shorten the data based on specific age groups
3. we would want to filter to the group we want, then average the amount for each month of each age group
4. We'd need an age range enum/group column, year, and number employed
5. we'd need columns for the different age groups, we can get them by grouping the different age groups.
6. Would need age and unemployment rates for thos ages throughout the years
7. You need columns for age and the period of time that you are looking at, as well as the percentage of unployment
8. you need month, age group, and employment. The wide table has one column per age band, so pivot_longer the six age columns into two, an age column holding the old names and an employment column holding the values, keeping month.

Student responses

1. 1. Year group by 2. Age group by 3. lag(salary) 4. ggplot() + geom_line()
2. Group by year and occupation and then get the mean of salaries throughout the years.
3. months?
4. The plot would have to selectively plot each column using the x axis for age group, then have each column be a singular line.
5. we would need to half our columns show difference in employee index for each month, grouped by age over all the months. We can do this easily by plotting all values for each age group
6. we would probably just need to create line plots this is pretty understandable to me and just make it for each line
7. We would probably need to make each age group their own individual table with the months, the percentages, and the change in percentages from month to month in order to draw our target plot.
8. you would need to draw one line for each column, you could get it by just plotting each row

Student responses

1. i would group by each age group column and plot the employment rate value for each month row, and reshape by pivot_longer to make the data more clear.
2. It needs the age group columns
3. The table need the employment rate in a specific tiemline to draw the overall changes of the employment rate.
4. We would need to filter the plot to get monthly averages for each age range, and then plot those averages as a line graph
5. You would have to group by age and then calcluate the difference between the past to determine the amount of change that is going on. You could get them from software by calculating monthly averages. You would need month age and employmne tcolumns with pivot longer.

Student responses

1. Create a new column with the change in % from the previous date
2. it would need the index levels per group for every month
3. no idea, this is unclear
4. We would need age group and employment status and we could get them by filtering the table by age group and employment status
5. We would need the dates as columns, and the ages as rows.

Student responses

1. a income level and data
2. need to create a helper column that shows the difference between dates. we need to use month, age, employment, and pivot longer to collapse the ages into a row
3. probably split each year into its own row. Use lag to create a column that takes the previous year's change and adds the change from that year to have a total change column. then plot year x total change
4. we would need the indexes as column to draw our target plot
5. year, software, and a numerical value from value benplotted

Student responses

1. the dataset would need date/time, age group, and percent change we would have to rehspe the data so taht all the age groups fall under one single column
2. we could do a row for each age group, then have a column for each year, where the value is the percent change for that age group.
3. we would need percentage, age groups, and years
4. year/time, age groups, employment %
5. You would need columns for age, percent change by month, and total statistics for both month and age group of jobs held

Student responses

1. group by age range. Then use ggplot to create a plot.
2. I think we would group by different age groups, but before we do that we would probably want to do a pivot longer so that the columns become rows in the data and then graph it after that
3. If software stores age groups in separate columns, use `pivot_longer()` to create age_group and employment columns. Then use `group_by(age_group)` and `mutate()` to calculate change_pct.
4. We would need columns for month, age group, and employment index, and we could reshape software from wide to long format so all the age columns become one age-group column.
5. We would need the counts of employment for all the years on the x-axis, separated by age groups, and we would also need a function that calculates the percentage change of employment.

- Reshape the six age columns into age and employment, keeping month. - Subtract 100 from employment to create percent_change. - Plot percentage change over time with a separate line for each age group.

77 anonymous responses

Other responses

77

Student responses

1. # Build your pipeline one step at a time. `developer_long <- software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate( percent_change = employment - 100 ) developer_long |> ggplot(aes( x = month, y = percent_change, color = age )) + geom_line() + labs( x = "Month", y = "Percent change in employment", color = "Age group" )`
2. `developer_changes <- software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate( percent_change = employment - 100 ) developer_changes |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `developer_changes <- software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate( percent_change = employment - 100 ) developer_changes |> ggplot( aes( x = month, y = percent_change, color = age ) ) + geom_line() + labs( x = NULL, y = "Change since Nov. 2022 (%)", color = "Age" ) # Build your pipeline one step at a time.`
2. `developer_long <- software |> pivot_longer( `22-25`:`50+`, names_to = "age", values_to = "pct_employed" ) |> mutate( percent_change = pct_employed - 100 ) ggplot(developer_long, aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`
3. `employment_long <- software %>% pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) %>% mutate(percent_change = employment - 100) ggplot(employment_long, aes(x = month, y = percent_change, color = age)) + geom_line()`
4. `software` # Build your pipeline one step at a time.
5. `software` # Build your pipeline one step at a time.

Student responses

1. software # Build your pipeline one step at a time.
2. software # Build your pipeline one step at a time. employment_long <- employment %>% pivot_longer(cols = `22-25`:`50+`, names_to = "age", values_to = "employment") %>% mutate(percent_change = employment - 100) ggplot(employment_long, aes(x = month, y = percent_change, color = age)) + geom_line()
3. software employment_long <- software |> pivot_longer(cols = `22-25`:`50+`, names_to = "age", values_to = "employment") |> mutate(percent_change = employment - 100) ggplot(employment_long, aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.
4. software employment_long <- employment %>% pivot_longer(cols = `22-25`:`50+`, names_to = "age", values_to = "employment") %>% mutate(percent_change = employment - 100) ggplot(employment_long, aes(x = month, y = percent_change, color = age)) + geom_line() + labs(x = "Month", y = "Percentage change", color = "Age group") # Build your pipeline one step at a time.

Student responses

1. `software library(tidyverse) software |> pivot_longer(cols = -month, names_to = "age", values_to = "employment") |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line(linewidth = 1) + labs(x = NULL, y = "Change since Nov. 2022 (%)", color = "Age") + theme_minimal() # Build your pipeline one step at a time.`
2. `software pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ" |> mutate( rel_employ = pct_employ - 100 )`
3. `software software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`
4. `software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer( '22-25':'50+', names_to = 'age', values_to = 'pct_employ' ) |> mutate(rel_employ = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer( `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line()`
3. `software |> pivot_longer( `22-25`:`50+`, names_to = "age", values_to = "pct_employ" ) |> mutate(percent_change = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line()`

Student responses

1. `software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line()`
2. `software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() + labs(x = NULL, y = "Change since Nov. 2022 (%)", color = "Age") # Build your pipeline one step at a time.`
3. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "employment" ) |> mutate( percent_change = employment - 100 ) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age))`
2. `software |> pivot_longer( cols = c(`22-25`: `50+`), names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age )) # Build your pipeline one step at a time.`
3. `software |> pivot_longer( `22-25`: `50+`, names_to = "age", values_to = "pct_employ" ) |> mutate( rel_employ = pct_employ - 100 ) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer( cols = `22-25`: `50+`, names_to = "age", values_to = "percent_employ") |> mutate( percent_change = percent_employ - 100) |> ggplot() + geom_line(aes( x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate( relemploy = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = relemploy, color = age))`
2. `software |> pivot_longer('22-25': '50+', #the range of the cols names_to = "age", values_to = "pct_employ") |> mutate(perc_change = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = perc_change, color = age))`
3. `software |> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = percent_change - 100) |> ggplot() + geom_line(aes(x=month, y=percent_change, color=age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ" ) |> mutate( pct_change = pct_employ - 100 ) |> ggplot() + geom_line(aes(x = month, y = pct_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y= rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employed") |> mutate(pct_change = pct_employed - 100) |> ggplot() + geom_line(aes(x=month, y=pct_change, color = age))`
3. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "percent_employ") |> mutate(percentchange = percent_employ - 100) |> ggplot(aes(x = month, y = percentchange, color = age))+ geom_line() # Build your pipeline one step at a time.`
4. `software |> pivot_longer(cols = `22-25`:`50+`, names_to = "age", values_to = "percentage") |> mutate(percent_change = percentage - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer( `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percentchange = employment - 100) |> ggplot(aes(x = month, y = percentchange, color = age)) + geom_line()`
3. `software |> pivot_longer( '22-25' : '50+', names_to = 'age', values_to = 'pct_employ', ) |> mutate( rel_employ = pct_employ - 100 ) |> ggplot() + geom_line( aes( x = month, y = rel_employ, color = age ) ) # Build your pipeline one step at a time.`
4. `software |> pivot_longer( '22-25':'50+', names_to = "age", values_to = "pct_employ" ) |> mutate(percent_change = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer( `22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer( `22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x= month, y = percent_change, color = age)) software # Build your pipeline one step at a time.`
3. `software |> pivot_longer( `22-25`: `50+`, names_to="age", values_to="pct_employ") |> mutate(percent_change=pct_employ-100) |> ggplot() +
geom_line(aes(x=month, y=percent_change, color=age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer( cols = `22-25`: `50+`, names_to = "age_group", values_to = "employment") |> mutate(percent_change = employment - 100) |>
ggplot() + geom_line(aes(x = month, y = percent_change, color = age_group)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer('22-25' : '50+', names_to = "age", values_to = 'pct_employed') |> mutate(percent_change = pct_employed - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_rank())) # Build your pipeline one step at a time.`
2. `software |> pivot_longer('22-25' : '50+', names_to = 'age', values_to = 'percent') |> mutate(percent_change = percent - 100) |> ggplot() +
geom_line(aes(x=month, y=percent_change, color=age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer('22-25' : '50+', names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100) |> ggplot() +
geom_line(aes(x=month, y=rel_employ, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "pct_employ") |> mutate( rel_employ = pct_employ - 100 ) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "pct_employ") |> mutate( percent_change = pct_employ - 100 ) |> ggplot() + geom_line(aes(x=month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "pct_employ") |> mutate( rel_employ = pct_employ - 100) |> ggplot(aes(x = month, y = rel_employ, color = age)) + geom_line() # Build your pipeline one step at a time.`
4. `software |> pivot_longer('22-25':'50+', names_to = "age", values_to = "pct_employ", )|> mutate(percent_change = pct_employ - 100)|> ggplot()+ geom_line(aes(x = month, y = percent_change)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot(aes(x=year, y=percent_change)) + geom_line() # Build your pipeline one step at a time.`
2. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate( rel_employ = pct_employ - 100) |> print() # Build your pipeline one step at a time.`
3. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer('22-25': '50+', names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100.0) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ")|>mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x=month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer(`22-25`: `50`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geomline(aes(x=month,y=rel_employ,color=age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer(`22-25`: `50+`, names_to = "age" values_to = "pct_employ" )|>`
4. `software |> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percentchange = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percentchange, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "employment") |> mutate(percent_change = employment - 100) |> ggplot() +
geom_line(aes(x=month,y=percent_change, color = age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate (percent_change = pct_employ-100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
4. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age))`
2. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ-100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employed")|> mutate(percent_change = pct_employed-100)|> ggplot()+
geom_line(aes(x=month, y=percent_change, color=age))`
4. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(percent_change = pct_employ - 100) |> ggplot() +
geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct_employ")|> mutate(percent_change = pct_employ - 100) |> ggplot() + geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`
3. `software |> pivot_longer(`22-25`:`50+`, names_to = "age", values_to = "pct-employ") |> mutate( rel_employ = pct_employ-100)|> ggplot() + geom_line(aes(x=month, y=pct_change, color=age)) # Build your pipeline one step at a time.`
4. `software_long <- software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate( percent_change = employment - 100 ) software_long |> ggplot(aes(x = month, y = percent_change, color = age)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `software_long <- software |> pivot_longer(-month, names_to = "age", values_to = "employment") |> mutate(percent_change = employment - 100)`
`ggplot(software_long, aes(x = month, y = percent_change, colour = age)) + geom_line() + geom_hline(yintercept = 0, linetype = "dashed")`
2. `software_long <- software |> pivot_longer( cols = `22-25`:`50+`, names_to = "age", values_to = "employment" ) |> mutate(percent_change = employment - 100) |> ggplot+ geom_line(aes(x = month, y = percent_change, color = age)) # Build your pipeline one step at a time.`

Student responses

1. `software|> pivot_longer( cols = `22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate(rel_employ = pct_employ - 100)|> ggplot(aes(x = month, y = rel_employ, color = age)) + geom_line(linewidth = 1) + geom_text(data = ~ dplyr::filter(., month == max(month)), aes(label = age), hjust = -0.1, show.legend = FALSE) + scale_y_continuous(labels = scales::label_number(suffix = "%")) + labs(title = "Change in employment over time by age group", x = "Month", y = "Percent change from baseline", color = "Age group") + theme_minimal() + theme(legend.position = "none") # Build your pipeline one step at a time.`
2. `software|> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "percent_change" ) |> mutate( rel_employ = percent_change - 100)|> ggplot() + geom_line(aes(x = month, y = percent_change, color= age)) # Build your pipeline one step at a time.`

Student responses

1. `software|> pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ") |> mutate( rel_employ = pct_employ - 100 ) |> ggplot() + geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`
2. `software|>pivot_longer(`22-25`: `50+`, names_to = "age", values_to = "pct_employ")|> mutate(rel_employ = pct_employ - 100)|> ggplot()+geom_line(aes(x = month, y = rel_employ, color = age)) # Build your pipeline one step at a time.`

Does the age-group plot prove that ChatGPT caused changes in employment for young software developers? Why or why not?

79 anonymous responses

Emphasizes correlation-not-causation

22

States that the plot shows correlation but explicitly denies that this proves causation without further analysis or evidence.

Affirms causation from plot

20

Asserts the plot demonstrates ChatGPT caused employment changes, citing the observed shift in the data as sufficient evidence.

Highlights alternative explanations

18

Rejects a causal conclusion based on the plot alone, pointing to possible confounding variables or other explanations.

Calls for statistical or methodological analysis

5

Says the plot alone is insufficient and recommends formal statistical tests or more rigorous methodology to assess causation or significance.

Response themes for question 6

Mixed: correlation implies possible role

4

Acknowledges correlation and suggests ChatGPT may play a role but stops short of claiming definitive causation and notes more evidence is needed.

Notes observed decline without causal claim

4

Describes the pattern in the plot (young software developers being hired less) without asserting ChatGPT caused it.

Qualified attribution with uncertainty

3

Leans toward ChatGPT playing a role or being a likely cause but explicitly notes uncertainty or caveats about proving causation from the plot alone.

Says ChatGPT is a contributing factor

3

Acknowledges other factors may exist but asserts ChatGPT is at least one contributing reason for the decline.

Student responses

1. Although we can see that young software developers stopped working significantly when chatgpt was created but that still won't be enough evidence to say that chatgpt caused changes in employment
2. I don't think that you can prove a causal relationship solely based on the plot
3. It does not prove this, merely implies there is correlation. This is not sufficient enough to say whether or not Chat GPT is the primary factor behind lower hiring rates for entry-level software developers
4. It doesn't prove but it shows a strong correlation, as previously all the age group lines were fairly close but there was a sharp deviation after the release of chatgpt
5. It might not be causal, but there is a correlation.
6. No because correlation is not causation. For us to prove causation we need to account for potential other factors.
7. No because there just seems to be a correlation. This does not prove causality
8. No it shows correlation not causation
9. no, correlation does not equal causation

Student responses

1. No, correlation does not prove causation
2. No, it does show some correlation but that doesn't mean causation. Personally, I do think that LLMs have reduced entry level job openings, so I am inclined to think there is causation.
3. No, it doesn't necessarily prove that it caused it because they are correlated, but we haven't found proof of causation
4. No, it shows a possible relationship but we cannot prove causation.
5. No, the graph shows an correlation over time, but it does not establish causation.
6. No, the plot only shows that employment changed around the time ChatGPT was released, but it does not prove that ChatGPT caused those changes.
7. No, this shows there is correlation not causation. I think its a intereting thing to look at though, it could be a factor impacting it.
8. No. The plot shows an association over time, not proof of causation.
9. not necesarily proved. Causation != ccorrelation. However, I do feel as if the release of chatgpt did have a significant effect on this

Student responses

1. the observation does not prove causation, only shows correlation.
2. The plot signals a correlation, but it does not signal a causation because we haven't conducted an experiment proving it was the cause.
3. This does not prove that ChatGPT caused changes, although it does show a decrease in employment we cannot prove that it is due to ChatGPT or not.
4. You cannot call it direct causation, there may be some correlation but there are also other factors like years of experience at play.

Student responses

1. I do believe that it proves this because it shows a decrease after the initial release of chatgpt
2. I think ChatGPT did cause a change in the job market for young software developers. We see a drastic increase in the change in employment after the release of ChatGPT, and the numbers differentiate larger by age group.
3. It suggests but doesn't prove that the release of chatgpt has decreased the employment level for young software devs. This could be due to its occupying simple, entry-level tasks.
4. Maybe yes because all the lines intersected right when Chat gpt was released, helping show a correlation between chatgpt and the employment rate for young software developers
5. The age-group plot does prove that ChatGPT caused changes in employment for young software developers because after ChatGPT was released, the employment rate for the two youngest age groups decreased instead of the upward trend seen with other age groups.
6. Yes
7. yes because the plot shifted
8. Yes it looks like as soon as chat cpt was introduced, the amount of young software engineers dropped significantly

Student responses

1. Yes it shows that chat gpt plays a role in the employment for young swe because right when chat gpt released is whe the trend for that age group was going down
2. yes the lines split a lot after chatgpt
3. yes, because the lines for the younger age groups start sloping downwards severely after chatgpt's release. it is correlation though, we cannot prove just from this causation.
4. Yes, because we saw how the percent change decreased for younger age groups probably because the jobs that these younger people were looking for were getting replaced by chatgpt. older people would have probably been employed by then so introducing chat gpt might have just made their job easier and more efficient, helping their employment rate, however there are many other employment trends outside of chatgpt alone
5. Yes, each individual age group went in different directions after ChatGPT was introduced.
6. yes, huge changes younger people like me are gg
7. Yes, it appears to, but without a test, we don't know

Student responses

1. Yes, it shows more entry level jobs are gone because AI has taken that over. But more skilled and upper level jobs are still valuable. That is personal belief but causation vs correlation
2. Yes, the number of people has decreased
3. Yes.
4. Yes. after chatgpt emerged the employment rate of 22-25 dropped
5. Yes. It negatively affects them, as shown by the steep decrease in the graph.

Student responses

1. Although the graph shows a negative trend after chatgpt was released, we do not have sufficient evidence to suggest that as there might be other reasons that affect the salary such as population growth.
2. I do not think it necessarily proves that ChatGPT caused changes in employment for young software developers. While I think it may have an impact, I think it is important to see how many people were no longer pursuing software development roles and instead transitioned to other fields like data science, analytics, or intelligence instead. It also may leave out hiring and layoff trends from big tech as well.
3. It doesn't prove it. We do nothing to prove correlation and we don't look for anything that can confound the current relationship we see. Other things could've happened around the same time that affected the job market.
4. No because we only looked at one factor out of a variety of possible causes although it seems that there is a strong relationship between the two

Student responses

1. No it doesn't. It shows a clear trend, but there could be other factors related to chatGPT or just the time it was released that could cause the changes in employment
2. no there might be other confounding variables
3. No, because it is not stated whether or not this data is controlled for other variables. We cannot rule out a causational relationship until we determine that other variables have not had an affect on younger software developers' employment.
4. No, because we haven't thought of other confounding variables that could have affected this change in employment.
5. No, even though thats when ChatGPT was released, there could have been other reasons for the decline
6. No, it doesn't because there are other factors involved in employment such as layoffs as a result of over hiring in the pandemic era or economic causes like war.
7. No, this plot doesn't look at all the variables behind what was happening. THere could be something else that could be the reason for these trend.s

Student responses

1. no. It is observational time-series. ChatGPT's release lines up in time, but timing is not cause. Many things hit at once, the end of the pandemic hiring boom, interest-rate hikes, and broad tech layoffs in 2022 and 2023. The plot shows an association, not a cause.
2. Not necessarily. It doesn't indicate any changes in economic changes that could've affected these numbers. We could set it relative to other sectors to make a fairer comparison.
3. The age-group plot does not definitively prove that ChatGPT CAUSED changes in employment for young software developers, it merely proves that there is a correlation between the emergence of ChatGPT and falling employment rates for young software developers because there might be other factors in the job market that we might not be aware of that could provide some extra context in this relationship (confounding factors).
4. This does not directly prove that ChatGPT caused changes in employment, as many other variables could exist that we are not considering. But it does appear although it had a major impact.

Student responses

1. we cannot say the "cause", because we also have different factors affecting this relationship
2. We dont have a good reference for the before chatgpt era so we wont be able to prove that there is a change in employment because of chatgpt. We could also claim that the rise of AI is the cause and not specifically chat.
3. While the plot may signify changes in employment, we are not sure if ChatGPT was the cause of it or how statistically different the changes were between age groups.

Student responses

1. No, because there isn't enough evidence to show that there has. If we ran a t-test, we would be able to make a conclusion.
2. no, need statis tific test
3. No, the plot shows a correlation that younger software developers employment is trending down, but it's not really proof without some kind of statisitcal test
4. No, we need a more statistically rigorous approach to answer that.
5. Visually it allows the viewer to see a correlation, but we can perform a t-test to see if the differnece is statistically significant before and after the introduction of OpenAI

Student responses

1. I don't think it can prove. The plot only shows a change in 2022 for employment for young software developers, and GPT is launched in 2022, no direct evidence of these two things causing each other
2. I think so. Because younger engineers are educated with AI environment so it has become part of their skills to use AI. But old workers might still rely on themselves the most.
3. No, the plot only shows an association between ChatGPT release and changes in employment.
4. the graph shows a correlation, however, a correlation doesn't always indicate a causation. there could be confounding variables that you just cannot eliminate without an experiment

Student responses

1. Although it doesn't prove anything, it does show an interesting trend that we can look into.
2. Does not prove but does show a type of correlation. After the emergence of AI it appears that younger people applying to the job market have faced a decrease in development so there might be some type of correlation but this is just one angle and would potentially need to see why both are affected
3. it shows that young sd are getting hired less after chat
4. No, it doesn't seem to have changed the trajectory of where the line was going

Student responses

1. From this plot yes, but we need more analysis of external variables.
2. No; this is simply a graphic without strong evidence. For example, it could be the case that the population declined for people between 22 and 30. Without proper statistical methods or controls, we can't see correlation or see causation.
3. The graph itself doesn't prove causation, but it's a useful plot that can provide insight into part of the explanation as to why software developers have seen this decrease in recent years

Student responses

1. Im not sure
2. no because we cant find causation by looking at the plot alone
3. The age_group plot shows that one contributing factor to a decrease in employment for young software developers has been the introduction of ChatGPT.
There may be other factors that haven't been considered.

- Keep month and the employment index for ages 22–25. Name the index column employment. - Sort the rows from highest to lowest employment. - Keep the first row to show the peak month and its employment index.

73 anonymous responses

Other responses

73

Student responses

1. `peak_month <- software %>% select(month, employment = `22-25`) %>% arrange(desc(employment)) %>% slice(1) peak_month`
2. `software # Build your pipeline one step at a time.`
3. `software software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
4. `software developer_changes |> filter(age == "22-25") |> transmute(month, employment = percent_change) |> arrange(desc(employment)) |> slice(1)`
5. `software peak <- jobs |> select(month, employment = index_22_25) |> arrange(desc(employment)) |> slice(1) # Build your pipeline one step at a time.`
6. `software peak_month <- software_jobs %>% select(month, employment = `22-25`) %>% arrange(desc(employment)) %>% slice(1) peak_month # Build your pipeline one step at a time.`
7. `software software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice(1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment='22-25') |> arrange(employment) |> slice_head(n = 1) top_n(software, 1) # Build your pipeline one step at a time.`
2. `software |> select(month, `22-25 employment index`) |> rename(employment = `22-25 employment index`) |> arrange(desc(employment)) |> slice(1) # Build your pipeline one step at a time.`
3. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) |> slice(1) # Build your pipeline one step at a time.`
4. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> head(1) # Build your pipeline one step at a time.`
5. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1)`
6. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice(1)`
7. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice(1)`
8. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice(1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment='22-25') |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
2. `software |> select( month, employment = `22-25` ) |> arrange(desc(employment)) |> head(1) # Build your pipeline one step at a time.`
3. `software |> select(month, "22-25") |> arrange(desc(`22-25`)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
4. `software |> select(month, `22-25`) |> arrange(desc(`22-25`)) |> slice_head(n=1) # Build your pipeline one step at a time.`
5. `software |> select(month, `22-25`) |> mutate(employment = `22-25`) |> select(month, employment) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
6. `software |> select(month, employment = '22-25') |> arrange(desc(`22-25`)) |> slice_head(n = 1)`
7. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
8. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
2. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1)`
3. `software |> select(month, employment = `22-25`)|> arrange(desc(employment))|> slice_head(n=1) # Build your pipeline one step at a time.`
4. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) top_n(software, 1, '22-25')`
5. `software |> select(month, employment == '22-25') |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
6. `software |> select(month, employment= `22-25`) |> arrange(desc(employment)) |> head(n = 1) # Build your pipeline one step at a time.`
7. `software |> select(month, employment=`22-25`) |> arrange(desc(employment)) |> head(1) # Build your pipeline one step at a time.`
8. `software |> select(month, employment=`22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment = `22-25`) |> arrange(desc(`employment`)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
2. `software |> head() employment_peak <- software |> select(month, `22-25`) |> rename(employment = `22-25`) # Build your pipeline one step at a time.`
3. `software |> mutate(employment = `22-25`) |> select(month, employment)|> arrange(desc(employment))|> head(1) # Build your pipeline one step at a time.`
4. `software |> select(month, '22-25', employment) |> arrange(desc('22-25')) |> slice_head(n=1) # Build your pipeline one step at a time.`
5. `software |> select(month, '22-25') |> arrange(desc('22-25')) |> slice_head(n = 1) # Build your pipeline one step at a time.`
6. `software |> select(month, '22-25') |> arrange(desc('22-25')) |> slice_head(n = 1) # Build your pipeline one step at a time.`
7. `software |> select(month, '22-25') |> arrange(desc('22-25')) |> slice_head(n = 1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, '22-25') |> arrange(desc('22-25')) |> slice_head(n = 1) top_n(software, 1, '22-25') # Build your pipeline one step at a time.`
2. `software |> select(month, '22-25')|> arrange(desc('22-25'))|> slice_head(n=1) # Build your pipeline one step at a time.`
3. `software |> select(month, `22-25`) arrange(desc(`22-25`)) |> slice_head(n = 1)`
4. `software |> select(month, `22-25`) |> arrange(desc(`22-25`)) |> slice_head(n=1) top_n(software, 1, `22-25`) # Build your pipeline one step at a time.`
5. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
6. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
7. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
8. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
2. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head (n = 1) # Build your pipeline one step at a time.`
3. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
4. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
5. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
6. `software |> select(month, employment = `22-25`) |> arrange(desc(employment))|> slice_head(n=1) # Build your pipeline one step at a time.`
7. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
2. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) top_n(software, 1, `22-25`) # Build your pipeline one step at a time.`
3. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
4. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
5. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> head(n=1) # Build your pipeline one step at a time.`
6. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice(head(1)) # Build your pipeline one step at a time.`
7. `software |> select(month, employment = '22-25') |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`

Student responses

1. `software |> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) # Build your pipeline one step at a time.`
2. `software |> select(month, employment='22-25') |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
3. `software |> select(month, employment='22-25') |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
4. `software |> select(month, employment='22-25') |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
5. `software |> select(month, employment=`22-25`) |> arrange(desc(employment)) |> slice_head(n = 1) top_n(software, 1, `22-25`) # Build your pipeline one step at a time.`
6. `software |> select(month, employment=`22-25`) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
7. `software |> select(month,employment = `22-25) |> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
8. `software|> select(month, employment = `22-25`) |> arrange(desc(employment)) |> slice_head(n=1)`

Student responses

1. `software|> select(month, employment = `22-25`)|> arrange(desc(employment)) |> slice_head(n=1) # Build your pipeline one step at a time.`
2. `software|> select(month, employment=`22-25`)|> arrange(desc(employment))|> slice_head(n=1) # Build your pipeline one step at a time.`
3. `software|>select(month, employment = `22-25`)|> arrange(desc(employment))|> slice_head(n=1) # Build your pipeline one step at a time.`
4. `software|>select(month, employment = `22-25`)|> arrange(desc(employment))|> slice_head(n = 1) # Build your pipeline one step at a time.`
5. `top_n(software, 1, '22-25')`

Did employment peak before or after ChatGPT's release? What does this tell us about the idea that ChatGPT started the decline?

0 anonymous responses

No responses were submitted.

Would you expect employment to have fallen more for young adults with a bachelor's degree or for those without one? Explain your prediction.

78 anonymous responses

Predicts larger drop for degree holders

49

Respondents who expect greater employment losses among young adults with a bachelor's degree, indicating belief that degree-holders were more affected.

Predicts larger drop for non-degree holders

16

Respondents who expect employment to decline more among young adults without a bachelor's degree, implying less job resilience for those without higher education.

Unclear or unintelligible response

5

Responses that are too brief, nonresponsive, or garbled to determine which group the respondent predicted.

Both or ambiguous

4

Responses indicating both groups decline or giving an unclear/ambivalent prediction without specifying which fell more.

Response themes for question 9

Other responses

4

Student responses

1. Employment for adults with people with bachelors degrees. I think that this AI introduction was more impactful for people with bachelors degrees because AI allows people without access to education to be educated
2. Employment for those with a bachelor's degree to have fallen more, since AI has disrupted a lot of jobs that require a degree.
3. For those with since ai seems to be easily used in any field
4. I am going to say for people with a bachelors degree have fallen because I think it is disrupting jobs where people that are educated need to keep up with what AI can do
5. I expect employment to fall more for those with a bachelor's degree because AI could lower the threshold of education?
6. i expect employment to have fallen for people with degrees to have fallen because specialization as in masters is much more valuable. Also blue collar jobs can not be replaced by AI
7. I predict employment drops in college graduates as non college graduates hold more trades jobs that cannot be replaced by AI

Student responses

1. I predict it would be higher for those with bachelors degree as they tend to work more skilled or specialized jobs that AI particularly replaces as compared to those without degrees on average.
2. I think it dropped more for people with bachelor's degree because I think the rise of AI is affecting more white collar jobs than blue collar jobs
3. I think it would have fallen more for people with a bachelor's degree because it seems like more people are getting into trades, compared to other jobs.
4. I think people with bachelors degrees because can't really replace manual labor with AI. AI also does many repeated tasks that humans used to do
5. I think that blue collar jobs are likely to be safer from AI as it has mostly impacted white collar operational gigs so far.
6. I think that employment for people with a bachelor's degree are more effected than those without one but only slightly. Certain jobs like retail or restaurant work will always need people and have arguably more opportunities. Those with a bachelor's degree are probably only looking for relevant jobs more so too

Student responses

1. I think that employment went down for those with a bachelors degree because major companies have had major layoffs for jobs that require a degree. In contrast, jobs that don't require a degree always seem to need more people
2. I think that employment will also decrease for young adults with a bachelor's degree, although not as steep of a decline as the software developer graph was for ages 22-25
3. I think those with a bachelor's degree would be more impacted because of the thinking ability of ai causing their jobs to be taken over while labor jobs are safe.
4. I would definately expect more for those with. Since if ChatGPT is the true reason then that takes away entry level corportate jobs not jobs like waiters and waitresses
5. I would expect a decline for people with degrees because their jobs can be automated by AI while people who work in fast food are still needed.
6. i would expect bachelor's degree un-employment to be higher, i do think that llms are more impactful to white collar jobs

Student responses

1. I would expect employment for those with a bachelor's degree to have fallen more than for those without. This is because bachelor's degrees have become so common in the labor market that they do not help distinguish candidates as much as they used to. Also, many jobs that require these kinds of degrees now have many portions of work automated by AI, which makes employers reluctant to hire.
2. I would expect employment for those with a bachelors degree to have fallen more, as AI can more easily replace those kinds of jobs versus more blue-collar work and the trades.
3. I would expect employment to decrease more for people with a bachlors degree. I feel like AI affect more white color jobs than blue color jobs, so the need for white color jobs have slowly decreased
4. I would expect employment to fall farther for those with a bachelor's degree, as academic work is easier to automate with AI.
5. I would expect employment to fall more for those with bachelors degrees because I feel like AI has more of an impact on those jobs

Student responses

1. I would expect employment to have fallen because chat gpt can do a lot of entry level jobs beyond software engineering.
2. I would expect employment to have fallen more for young adults with a bachelor's degree because they are more likely to work in jobs that could be affected by AI.
3. I would expect employment to have fallen for those with a bachelor's degree because of the development of AI, but in considering all industries, this would have less of an impact than software development.
4. I would expect employment to have fallen more for people with a bachelor's degree since it is more likely to effect white collar sectors.
5. I would expect employment to have fallen more for those with a bachelor's degree. Young adults who don't go to college likely go into trades or service work, which haven't been affected by AI as much. AI mainly affects the white-collar jobs bachelors students get.
6. I would expect employment to have fallen more for young adults with a bachelor's degree because many degree-based entry-level jobs are more exposed to AI and automation.

Student responses

1. I would expect employment to have fallen to a greater extent for young individuals with college degrees, since AI can do analysis and decision making surrounding intellectual fields rather than physical work.
2. I would expect it has fallen more for people with a bachelors degree, as people without one are in fields that are less likely to be impacted by AI
3. I would expect that employment would fall more for people with bachelor's degrees because the rise of AI affects jobs not requiring bachelor's degrees less.
4. I would not because I am going to look for a job soon. Though I really expect AI can focus on reduce poverty and wars and protection more instead of taking jobs away.
5. I would predict a larger decline among young adults with a bachelor's degree because many AI-exposed professional and technical jobs require a college degree.
6. I would say those with a bachelors degrees should be fine but at tthe same time I don't have a good explanatiopn.
7. In a post AI-era, employment changes are more drastic among those with bachelor's degrees than those without due to the cognitive offloading AI allows.

Student responses

1. More because AI is currently more suited towards white collar jobs, which are jobs mostly occupied by people with bachelors degrees
2. more for those with one because it is relatively more saturated than before and ai is disrupting white collar jobs more
3. Those with a bachelor's, as AI tends to take over white-collar work
4. those with because service jobs cant be done by ai
5. with a bachelor's degree would possibly suffer more because jobs without bachelor's degree probably can't be replaced by AI so they would remain the same, but jobs with bachelor's degree is more affected by AI.
6. with a bachelors degree because ai impacts their potential careers more
7. With BS degree fallen more
8. with, because AI could do more jobs
9. With. Those without a degree wouldn't be affected because their jobs wouldn't even require that much technology.

Student responses

1. yes I expect that it might have influenced other young adults too but we have to be sure
2. Yes, I think AI is disrupting jobs for those with a degree than without one.
3. you would expect employment to have fallen more for YA with a bach degree

Student responses

1. fallen. Bachelors are more common recently, so it wouldn't mean distinction as in the past
2. for those without one, recruiters tend to gravitate towards people with a degree
3. For those without one, since this is relative to the percentages before.
4. I think I would because those without a bachelor's would have blue-collar/trade jobs and I think there is always an abundance of those jobs while there isn't an abundance of jobs that require a bachelor's
5. I would expect employment to fall more for those without a degree because the degree serves as the only differentiator between people while hiring
6. I would expect employment to fall more for those without one because the bachelor's degree has value that would make it so changes don't have as much of an effect
7. I would expect employment to fall more for young adults with a bachelor's degree because the jobs that adults without a BA are targeting are probably more hands on and jobs that can not be replaced by AI

Student responses

1. I would expect it to have fallen more for those without a bachelors degree as many jobs require candidates to have at least a bachelors degree to even be considered
2. I would expect more for those without a bachelors degree
3. I would expect the employment rate to fall more among young adults without a bachelor's degree cause their works are more and more substitutable by AI.
4. it would probably drop more for people without bachelors degrees because the jobs that don't require a degree can more easily be replaced by ai.
5. People wiithout bachelor degree
6. Probably without one, because a lot of jobs now require more experience more than ever.
7. without
8. Without. It wouldnt make sense for people without an education to be employed more than people with a college level education
9. yes because companies tend to hire those with a bachelor's degree

Student responses

1. I think it depends on the type of bachelors degree
2. I would explain it to fhjave fallen mor cus of layouts
3. I'd expect employment to have fallen for young people with a bachelor's degree, and employment to have fallen at a greater degree for those without a bachelor's degree.
4. yes
5. Yes

Student responses

1. A reasonable one either way. If AI hits knowledge work you might expect degree holders to fall more, if it is a general young-worker slowdown you might expect the non-degree group to fall more.
2. I expect that overtime as more and more people get college degrees, the job market will be less competitive for those without bachelor's degrees.
3. i think both populations employment rate drops
4. impact to more white collar jobs

Student responses

1. I actually think that employment is poor for most young adults right now. The only industry that seems to continue to increase in need is blue collar. The economy is bad. Ass even.
2. I expect the salary b/w the 2 groups will shrink, as the human work are more important when AI is released.
3. I would expect employment for those with a bachelor's degree has dropped less than those without
4. Yes, because most young adults tend to achieve the master degree after they gain the bachelor degree

**- Keep observations from June 2023 onward. -
Keep the population labeled "Everyone",
retaining all three groups. - Plot employment rate
against month, with a separate line for each
group.**

79 anonymous responses

Other responses

79

Student responses

1. `library(tidyverse) youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> count(group) |> ggplot(aes(x = month, y = employment_rate, color = group, group = group)) + geom_line(linewidth = 1) + labs(x = "Month", y = "Employment rate", color = "Group") + theme_minimal() # Build your pipeline one step at a time.`
2. `plot_data <- youth %>% filter(month >= as.Date("2023-06-01"), population == "Everyone") ggplot(plot_data, aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`
3. `youth # Build your pipeline one step at a time.`
4. `youth youth |> filter(month >= "2023-06-01" & population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`
5. `youth youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter( month >= "2023-06-01", population == "Everyone" ) |> ggplot( aes( x = month, y = employment_rate, color = group ) ) + geom_line() + labs( x = NULL, y = "Employed (%)", color = NULL ) # Build your pipeline one step at a time.`
2. `youth |> filter( month >= as.Date("2023-06-01"), population == "Everyone" ) |> ggplot(aes( x = month, y = employment_rate, color = group )) + geom_line()`
3. `youth |> filter( month >= as.Date("2023-06-01"), population == "Everyone" | group == "26-65, all education levels" ) |> ggplot( aes(x = month, y = employment_rate, color = group) + geom_line() + labs( x = NULL, y = "Employed (%)", color = NULL ) ) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter( month >= as.Date("2023-06-01"), population == "Everyone", group %in% c("AI exposure group 1", "AI exposure group 2", "AI exposure group 3")) |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line()`
5. `youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line()`

Student responses

1. `youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() + labs(x = NULL, y = "Employment rate (%)", color = "Group", title = "Employment rates for youth groups since June 2023")`
2. `youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() #`
Build your pipeline one step at a time.
3. `youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() + labs( x = "Month", y = "Employment rate", color = "Group" )`
4. `youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() + labs(x = NULL, y = "Employed (%)", color = NULL) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= as.Date("2023-06-01")) |> filter(population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line()` # Build your pipeline one step at a time.
2. `youth |> filter(population == "Everyone", group %in% unique(group)) |> filter(month >= "2023-06-01") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line()`
3. `youth |> filter(population == "Everyone") youth |> filter(month >= as.Date("2023-06-01"), population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, colour = group)) + geom_line() + labs(y = "Employment rate (% , 12-month moving average)")`
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone")` # Build your pipeline one step at a time.
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line( aes(x = month, y = employment_rate, color = group))` # Build your pipeline one step at a time.

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month, y=employment_rate, color = group)) # Build your pipeline one step at a time.`
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`
3. `youth |> filter(month >= "2023-06-01")|> filter(population == "Everyone") |> ggplot(aes(x=month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= "2023-06-01", population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate), color = group) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month,y=employment_rate, color=group))`
2. `youth |> filter (month >= '2023-06-01',population == 'Everyone') #filter (population == 'Everyone') # Build your pipeline one step at a time.`
3. `youth |> filter(month >= '2023-06-01') |> filter(population == 'Everyone') |> ggplot() + geom_line(aes(x=month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= '2023-06-01') |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month,y=employment_rate, color = group)) # Build your pipeline one step at a time.`
5. `youth |> filter(month >= "2023-06-01", population == "Everyone") |> filter(population == "Everyone")|> ggplot()+geom_line(aes(x=month, y=employment_rate, color = group)) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= "2023-06-01") filter(population == "Everyone") |> ggplot() + geom_line( aes(x = month, y = employment_rate, color = group) ) # Build your pipeline one step at a time.`
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month, y= employment_rate, color = group)) #` Build your pipeline one step at a time.
5. `youth |> filter(month >= "2023-06-01") |> filter (population == "Everyone") #` Build your pipeline one step at a time.

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") ggplot() + geom_line( aes(x = month, y = employment_rate, color = group) ) # Build your pipeline one step at a time.`
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone")|> ggplot()+ geom_line(aes(x = month, y = employment_rate, color = group)) # Build`
your pipeline one step at a time.

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == `Everyone`) |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month,y=employment_rate,color = group)) # Build`
your pipeline one step at a time.
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line( aes(x = month, y = employment_rate, color = group) ) #`
Build your pipeline one step at a time.
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group))`
5. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") # Build your pipeline one step at a time.`
6. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line( aes(x = month, y = employment_rate, color = group)) #`
Build your pipeline one step at a time.

Student responses

1. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month, y=employment_rate, color = group)) # Build your pipeline one step at a time.`
2. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
3. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
4. `youth |> filter(month >= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month, y=employment_rate, color=group)) # Build your pipeline one step at a time.`
5. `youth |> filter(month >= as.Date("2023-06-01")) |> filter(population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `youth |> filter(month >= as.Date("2023-06-01")) |> filter(population == "Everyone") |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line()`
2. `youth |> filter(month>= "2023-06-01") |> filter(population == "Everyone") |> ggplot() + geom_line(aes(x=month,y=employment_rate,color=group)) # Build your pipeline one step at a time.`
3. `youth |> filter(month>= "2023-06-01")|> filter(population == "Everyone")|> ggplot()+ geom_line(aes(x=month, y = employment_rate, color = group) ) # Build your pipeline one step at a time.`
4. `youth_filtering <- youth |> filter( month >= as.Date("2023-06-01"), population == "Everyone" ) youth_filtering |> ggplot(aes(x = month, y = employment_rate, color = group)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `youth_recent <- youth |> filter(month >= as.Date("2023-06-01")) |> filter(population == "Everyone") library(ggplot2) ggplot(youth_recent, aes(x = month, y = employment_rate, color = group)) + geom_line(linewidth = 1.1) + labs(title = "Employment Rate by Group (June 2023–Present)", x = "Month", y = "Employment Rate (%)", color = "Group") + theme_minimal()`

Student responses

1. youth.youth.filter(month >= June) # what i should do? What should i use # You should make a line plot with 3 lines, one for each group, using only recent data and only the "Everyone" population. # What to do: # 1. Filter rows to: # - month >= June 2023 # - population == "Everyone" # 2. Keep the 3 groups mentioned in the slide # 3. Use youth to compare employment across those groups # 4. Plot: # - x-axis: month # - y-axis: employment_rate # - color/group: group # - geom_line() # In R, you'll likely use dplyr + ggplot2. # Example structure: # youth |> # filter(month >= as.Date("2023-06-01"), population == "Everyone") |> # ggplot(aes(x = month, y = employment_rate, color = group)) + # geom_line() # If your group variable has a different name, replace group with the actual column name. If you want, I can help write the exact code from your dataset's column names. # Build your pipeline one step at a time.

Student responses

1. `youth|> filter(month >= "2023-06-01")|> filter(population == 'Everyone')|> ggplot()+ geom_line( aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
2. `youth|> filter(month>="2023-06-01")|> filter(population=="Everyone")|> ggplot()+ geom_line(aes(x=month,y=employment_rate, color= group)) # Build your pipeline one step at a time.`
3. `youth|>filter( month >= "2023-06-01", population == "Everyone" ) |> ggplot(aes( x = month, y = employment_rate, color = group )) + geom_line() + labs( x = "Month", y = "Employment rate", color = "Group" ) # Build your pipeline one step at a time.`
4. `youth|>filter(month >= "2023-06-01", population == "Everyone")|> ggplot()+ geom_line(aes(x = month, y = employment_rate, color = group)) # Build your pipeline one step at a time.`
5. `youth|>filter(month>="2023-06-01") # Build your pipeline one step at a time.`

How does the graph fit your prediction? What does an explanation of the decline now need to account for?

69 anonymous responses

Both education groups declining

21

Students note that employment is decreasing for both those with and without bachelor's degrees.

Overall youth employment falling

15

Students say the graph shows a decline in employment for young people overall, not matching the prediction for only one subgroup.

Economic or broad causes suggested

11

Students propose that factors beyond AI, such as the overall economy, could explain decreases affecting both groups.

Those without degrees hit hardest

7

Students assert that people without college degrees are most affected by the decline (greater impact on non-degree holders).

Response themes for question 11

Higher education still gives advantage**6**

Students observe that people with bachelor's degrees have higher employment rates despite declines and are less impacted.

General sense the rate dropped**4**

Student offers a brief, non-specific observation that the employment line decreased for the age group.

Unemployment causing the decline**3**

Student attributes the decline to more people becoming unemployed.

Other responses**2**

Student responses

1. Both groups had a decrease in employment rate, so maybe its not just AI but the economy as a whole
2. For people with bachelors, the emplyment rate is decreasing, but it is also decreasing (even more so maybe) for people without bachelors degree.
3. Is only somewhat fits my prediction. It seems young people even below bachelors are seeing a reduction in employment alongside those with a bachelors, thus it may just be an economic issue.
4. it does not fit my prediction, as those both groups show a similar rate of decreasing employment levels.
5. It does not fully fit my prediction, since we see similar downturns for both groups. We may need to somehow control for the state of the economy, which obviously affects employment
6. it matches my prediction
7. It sort of fit my prediction. I thought that there would be a decrease in young adult with bachelor's degrees, but not super steep, and that was what we observed in the plot
8. Maybe a slight yes, the graph does appear to be going downward a bit for the bachelors degree people but it is nothing drastic like I thought previously

Student responses

1. My prediction was wrong because the plot that we made shows that the employment rate was higher for young adults with a BA or higher but still decreased slightly, while the young adults without a BA or higher had a constant employment rate. Both groups employment rate declined, and the decline needs to account for COVID, and other confounding variables
2. No. Both Bachlors and those without a Bachlor degree are overall dropping by 1%. However, this decreasing trend only happens in young people but not others. We can account for age instead
3. q
4. The graph actually shows that both those with and those without bachelor's degrees faced a decline in employment with the introduction of ChatGPT.
5. The graph does fit the prediction as people with no bachelors still have a higher employment rate and seems liek the dip is the same.
6. The graph does not display the pronounced drop in employment in the bachelor degree group, which should occur if it is more affected by ai release.

Student responses

1. The graph does not generally support the idea that bachelor's people are more exposed to AI. We see a relative decline for both those with and without a degree. Just because more bachelor's people are generally employed, it does not mean that they are more/less affected that is a completely separate issue.
2. The graph does not support my prediction of there not being a trend at all. Both have declined, but those without bachelor's degree have been declining more steadily in my opinion.
3. The graph doesn't really fit my prediction. It doesn't look like there is a steep drop in either bachelors or non bachelors, as there is a small decline in both
4. The graph fits my prediction that the group with BAs decreased since AI, however it goes against my prediction for those without BAs as I predicted they would grow when in fact their employment rate fell as well. How does the explanation account for this fact, idk
5. The plot shows small decreases in both education groups.

Student responses

1. The trend in employment rate for those with a bachelor's or higher in the 22-25 crowd matched my prediction, but the employment for those without a bachelor's did not. An explanation of the decline now also needs to account for past events before AI became prevalent.
2. This plot does not because the slight decrease seen is also seen with those with below bachelors, not just bachelors or higher. So it could be affecting all youth.

Student responses

1. Both job trends downwards but all education levels is relatively stable. I do not have reasoning
2. by just looking, it looks like bachelors and no bachelors decreased by about the same amount. i cant really tell which one has a bigger slope
3. It agrees with my prediction. The decline accounts for the appearance of AI tools.
4. It doesn't fit my prediction, it shows the same (very slightly downward) trend for everyone in the age group, so an explanation has to account for the fact that there is no difference by education level
5. It fits because it decreased a little
6. My prediction was incorrect as no matter if a young person person has a bachelor's degree the employment rate seems to be going down. I think the explanation of the decline should account for the proportion of people in each age group and the availability of entry level jobs.
7. no, in real world, the all education is higher than 22-25
8. The both 22-25 groups shows a decreasing trends, the whol age group was not affected much.

Student responses

1. The graph doesn't fit my prediction because the employment rates for all young people are dropping regardless of education level. The graph needs to account for age.
2. The graph doesn't fit my prediction, because it seems like young people as a whole regardless of bachelor's or not are doing down in employment percentage
3. The graph partly fits my prediction. Because employment declined for young adults with a bachelor's degree, but it also declined for those without one. An explanation needs to account for why employment fell for young adults in both education groups.
4. The plot shows that for our age group, no matter your education level there is a decrease in employment rates. AI should not have an effect on most jobs that do not require a degree so we can conclude there is a confounding variable influencing our data.
5. There is a slightly decline in the 22-25 group and 26-35 group.
6. This graph shows a downward trend for all people aged 22-25, which does not support my hypothesis that college-educated people would be more affected
7. yeah it fits my prediction, idk what it now needs to account for

Student responses

1. An explanation of the decline needs to account for things that aren't AI, that disproportionately affect lower-education people.
2. Interestingly, not only did white-collar jobs decrease, but also non-white collar.
3. it does show that hiring for bachelors went down
4. No this graph does not fit my prediction. An explanation of the decline needs to account for
5. No, the plot suggests that bachelors degree holders are still more employed than people without. We need to account for majors and if someone majored in an AI effected role like CS, DataSci.
6. The graph fits my prediction because bacheolors dropped more. An explanation of the decline needs to account for other fators other thn gpt
7. The graph fits my prediction that all education levels would remain constant, however, the decline is shown in those without a Bachelors, while I predicted those with Bachelors would have a decline in employment.
8. The plot fit my guess. Is the covid effect the employ rate?

Student responses

1. the sharpest decline was shown in people who have bachelors or higher, so we should filter them by education level when examining the impact of AI on jobs.
2. Yes the graph matched my expectations in the job market's response to AI. The decline would need to account for factors including pursuit of graduate study, population changes, and more.
3. Yes. This shows that those with a degree have a greater decline with the presence of AI versus those who don't. And the overall employment level stays constant.

Student responses

1. does not fully fit the prediction because employment fell more for young adults below a bachelor's degree.
2. It agrees that with my prior conclusion of people without bachelors are more impacted by AI
3. It looks more positive than my prediction but the trend for people without a bachelor's degree is not promising.
4. It matches my prediction that people with no college education would be hit the most with unemployment.
5. Like i predicted, people without bachelor degree had lower employment rate and experienced much drop after the advent of ai
6. No, people without Bacehlor's are seemingly more affected by the AI boom as their drop in employment has been more stark compared to those with bacehlor's
7. yes, low degree has been impacted

Student responses

1. fit my prediction of degree vs employment
2. it fits exactly
3. No it does not fit my prediction, as although people with higher education have more employment, it still seems notably impacted by AI
4. The plot kind of fits, it shows a decline in employment of Bachelors degrees but we need to scale it , since the overall employment rate for lower edu is lower, and compare the differences to scale.
5. Yeah, definitely consistent with it. Having a bachelors degree increases chances of employment accortding to just this data.
6. Yes it supports my prediction. Those without degree are employed less since 2022, but this declining trending is lighter than while those with degrees. Maybe it's becasue the covid

Student responses

1. It fit my prediction, but an explanation now needs to consider education level
2. it makes sense that the line went down for that age group. the 26-65 is not impacted, while the other two age groups went down
3. It seems that the decline of employment percentage of people with BAs or higher is declining at a higher rate than other professions, this is probably because BAs have a higher salary on average so employers want to pay less of them.
4. The graph is following Simpson's paradox, where even though the overall employment rate of the graph is rising, the employment percentage of the individual groups of bachelors or higher or below bachelors is both following, so the graph needs to account for this, however, strictly speaking, the graph does fit my prediction that unemployment dropped slower for bachelors candidates than for non bachelors candidates

Student responses

1. both becoming unemployed at similar rates so ai does not really favor one over other
2. No, it did not fit my prediction. The employment rate stayed about the same for the two groups.
3. The graph does not fit my prediction, because the 22-25 below bachelors had a lower employment rate. The explanation of the decline now needs to account for the population of people in 22-25, because maybe more people aged out of that group while less people aged into the group.

- Keep observations from January 2019 onward.
 - Divide the employment values by 1,000 to express them in millions.
 - Draw a line plot of employment in millions against month.
-

76 anonymous responses

Other responses

76

Student responses

1. `employment_plot <- labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs(x = "Month", y = "Employment (millions)", title = "Employment over time") # Build your pipeline one step at a time.`
2. `labor_market` # Build your pipeline one step at a time.
3. `labor_market labor_market |> filter(month >= "2019-01-01") |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line()` # Build your pipeline one step at a time.
4. `labor_market filtered_data <- labor_market |> filter(month >= as.Date("2019-01-01")) mutated_data <- filtered_data |> mutate(employment_millions = employment / 1000) ggplot(mutated_data, aes(x = month, y = employment_millions)) + geom_line()` # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment = employment / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()` # Build your pipeline one step at a time.
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment = employment / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()` # Build your pipeline one step at a time.
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) + labs(y="Tech Jobs (Millions)", title = "Tech Hiring Across COVID")` # Build your pipeline one step at a time.
4. `labor_market |> filter(month > "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs))` # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= '2019-01-01') |> mutate(million_employment_val = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech hiring across covid") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(employment = employment / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) +labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs)) + geom_line() + labs( title = "Tech employment since January 2019", x = "Month", y = "Employment (millions)" ) + theme_minimal() # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs)) + geom_line() + labs( y = "tech jobs (millions)", title = "tech hiring across covid" )`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs))+ labs(y= "tech jobs in millions", title = "tech hiring across covid")`
4. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment = tech_jobs / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line() + labs(x = NULL, y = "Employees (millions)") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment = tech_jobs / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()`
Build your pipeline one step at a time.
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs( x = "Month", y = "Employment (millions)", title = "Employment from January 2019 onward" )`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs( x = "Month", y = "Employment (millions)", title = "Employment since January 2019" )` # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs(x = "Month", y = "Employment (millions)", title = "Employment since January 2019") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment_millions = employment / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs(x = NULL, y = "employment (millions)", title = "software developer employment since 2019")`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs)) + geom_line()`
4. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs)) + geom_line() + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(tech_jobs_m = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs_m)) + geom_line() + labs(y = "Tech jobs (millions)")`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000)|> ggplot()+geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate( tech_jobs = tech_jobs / 1000 ) |> ggplot() + geom_line(aes(x = month , y = tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate( tech_jobs = tech_jobs / 1000 ) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (millions)", title = "Tech Hiring across Covid") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(x = "Month", y = "Tech Jobs (Millions)", title = "Tech Hiring Before and After COVID") # Build your pipeline one step at a time.`
5. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) + labs(y=tech_jobs(millions_, title = "tech hiring across covid) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs) ) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs, labs(y="Tech Jobs (Millions)", title = "Tech Hiring across COVID")`
5. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs_millions = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs_millions)) + geom_line() + labs( x = NULL, y = "Employment (millions)" ) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01")|> mutate(jobs = tech_jobs/1000) |> ggplot(aes(x=month, y = jobs)) + geom_line() + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01")|> mutate(tech_jobs = tech_jobs / 1000)|> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)") # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate( tech_jobs_millions = tech_jobs / 1000 ) |> ggplot( aes( x = month, y = tech_jobs_millions ) ) + geom_line() + labs( x = NULL, y = "Employees (millions)" ) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month>= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01")` # Build your pipeline one step at a time.
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across Covid")` # Build your pipeline one step at a time.
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID")` # Build your pipeline one step at a time.
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID")` # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y = tech_jobs))|> labs(y = "Tech Jobs(Millions)", title= "Tech Hiring across COVID") # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs", title = "Tech Hiring across Covid") # Build your pipeline one step at a time.`
4. `labor_market |> filter(month > "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y= tech_jobs)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= '2019-01-01') |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs (y = "Tech Jobs (Millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= '2019-01-01') |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x=month,y=tech_jobs)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs,)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) #labs(y = "Tech Jobs (Millions)", # title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs( y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID" ) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot(aes(x = month, y = tech_jobs)) + geom_line() + labs( x = "Month", y = "Tech Employment (millions)", title = "Tech Employment since January 2019" ) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y = "Tech Jobs (Millions)", title = "Tech Hiring across COVID") # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate (tech_jobs = tech_jobs /1000) |> ggplot() +geom_line(aes(x = month, y=tech_jobs)) + labs(y= "tech jobs") # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(employment = tech_jobs / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()` # Build your pipeline one step at a time.
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = (tech_jobs / 1000)) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) + labs(y= "Tech Jobs (Millions)", title = "Tech Hiring across COVID")` # Build your pipeline one step at a time.
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs))` # Build your pipeline one step at a time.
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) + labs(y= "Tech Jobs(Mil)", title = "Tech hiring across COVID")` # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month,y=tech_jobs)) + labs(y="Tech Jobs (Millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs), + labs(y = "Tech Jobs (Millions)", x = "Month", title = "Tech Hiring across COVID")) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs/1000)|> ggplot() + geom_line(aes (x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
5. `labor_market |> filter(month >= "2019-01-01") |> mutate( tech_jobs = tech_jobs / 1000)|> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs = tech_jobs / 1000) |> ggplot() + geom_line(aes(x = month, y = tech_jobs)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> mutate(tech_jobs=tech_jobs/1000) |> ggplot() + geom_line(aes(x=month, y=tech_jobs)) # Build your pipeline one step at a time.`
4. `labor_market|> filter(month >= "2019-01-01")|> mutate(tech_jobs = tech_jobs/ 1000)|> ggplot() + geom_line(aes(x=month, y = tech_jobs)) # Build your pipeline one step at a time.`
5. `labor_market|> filter(month>="2019-01-01") mutate(tech_jobs=tech_jobs/1000)|> ggplot()+geom_line(aes(x=month,y=tech_jobs)) labs(y="Tech_jobs", title="Tech hiring across COVID") # Build your pipeline one step at a time.`

Student responses

1. `labor_market|> filter(month >= "2019-01-01") |> mutate(employment_millions = tech_jobs / 1000) |> ggplot(aes(x = month, y = employment_millions)) + geom_line() + labs( x = "Month", y = "Tech employment (millions)" ) # Build your pipeline one step at a time.`
2. `labor_market|>filter(month >= "2019-01-01")|> mutate(tech_jobs = tech_jobs/1000)|> ggplot()+geom_line(aes(x = month, y = tech_jobs))+ labs(y = "Tech Jobs (Millions)", title = "Tech Hiring Across COVID") # Build your pipeline one step at a time.`
3. `labor_plot<-labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(emp_million = tech_jobs / 1000) labor_plot |> ggplot(aes(x = month, y = emp_million)) + geom_line() # Build your pipeline one step at a time.`
4. `library(tidyverse) labor_market |> filter(month >= "2019-01-01") |> mutate(employment = employment / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line()`

Student responses

1. `library(tidyverse) labor_market |> filter(month >= as.Date("2019-01-01")) |> mutate(employment = employment / 1000) |> ggplot(aes(x = month, y = employment)) + geom_line(linewidth = 1) + labs(x = "Month", y = "Employment (millions)") + theme_minimal() # Build your pipeline one step at a time.`

How well does your plot support the idea that employers are reversing a pandemic-era expansion? Is employment back to its pre-pandemic level?

0 anonymous responses

No responses were submitted.

- Keep observations from January 2019 onward.
 - Keep month, hires, quits, and layoffs.
 - Gather the three rate columns into measure and rate, keeping month.
 - Draw a line plot of rate against month, with a separate line for each measure.
-

77 anonymous responses

Other responses

77

Student responses

1. labor_market # Build your pipeline one step at a time.
2. labor_market # Build your pipeline one step at a time. labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line()
3. labor_market labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.
4. labor_market labor_maket |> filter(month >= as.Date("2019-01-01"))|> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits,layoffs))|> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
2. `labor_market %>% filter(month >= as.Date("2019-01-01")) %>% select(month, hires, quits, layoffs) %>% pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) %>% ggplot(aes(x = month, y = rate, color = measure)) + geom_line()`
3. `labor_market |> filter(month > "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = hires:layoffs, names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= '2019-01-01') |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = 'measure', vals_to = 'rate') |> # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line()`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = -month, names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() + labs(x = "Month", y = "Rate", color = "Measure") # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line()`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() + labs(x = NULL, y = "Monthly rate (%)", color = NULL) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, colour = measure)) + geom_line()`
2. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() + labs(x = NULL, y = "Rate (%)", color = "measure", title = "hiring and departures in the software-developer labor market since 2019")`

Student responses

1. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = hires:layoffs, names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() + labs(x = NULL, y = "Monthly rate (%)", color = NULL) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month > "2019-01-01") |> select(month,hires,quits,layoffs) |> pivot_longer(-month, names_to = "measure",values_to = "rate")|> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( month, names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = -month, names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = -c(month), names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month, y=rate, color=measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y= rate, color= measure)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs)|> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot(aes(x=month, y = rate, color = measure))+ geom_line() # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to="measure", values_to="rate")|> ggplot() + geom_line( aes(x = month, y = rate, color = measure))`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= '2019-01-01') |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month,y=rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( -month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure))`
4. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate" ) |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(`hires`,`layoffs`, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure )) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month,y=rate,color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> # reshape everything thats not month col ggplot() + geom_line(aes(x=month, y=rate, color=measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month, y=rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> gg_plot() + geom_line (aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rates") |> ggplot() + geom_line(aes(x = month, y = rates, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to='measure', values_to='rate') |> ggplot() + geom_line(aes(x=month, y=rate, color=measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month) |> ggplot() + geom_line(aes(x = month, y = value, color = name)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01") |> select(month,hires,quits,layoffs) |> pivot_longer(~month, names_to = "measure", values_to = "rate")) ggplot()+ geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01") |> select(month,hires,quits,layoffs) |> pivot_longer(-month,names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month,y=rate,color=measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x=month, y=rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to="measure", values_to="rate") |> ggplot() + geom_line(aes(x=month, y=rate, color=measure)) # Build your pipeline one step at a time.`
4. `labor_market |> filter(month >= "2019-01-01")|> select(month, hires, quits, layoffs) pivot_longer( id = month, names_to = "measure", values_to = "rate" )|> ggplot()+geom_line(aes(x=month,y=rate,color=measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month >= "2019-01-01")|> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot() + geom_line(aes(x = month, y = rate,color = measure)) # Build your pipeline one step at a time.`
2. `labor_market |> filter(month >= "2019-01-01")|> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
4. `labor_market |> filter(month>="2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to="measure", values_to"rate") ) |> # Build your pipeline one step at a time.`

Student responses

1. `labor_market |> filter(month>="2019-01-01")|> select(month, hires, quits, layoffs) |> pivot_longer(-month, names_to = "measure", values_to = "rate") |> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`
2. `labor_market|> filter(month >= "2019-01-01")|> select (month, hires, quits, layoffs)|> pivot_longer(cols = -month, names_to = "measure", values_to = "rate")|> ggplot(aes(x = month, y = rate, color = measure)) + geom_line() + labs(title = "Hiring, quits, and layoffs since January 2019", x = "Month", y = "Rate", color = "Measure") # Build your pipeline one step at a time.`
3. `labor_market|> filter(month >= "2019-01-01")|> select(month, hires, quits, layoffs)|> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot() + geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `labor_market|> filter(month >= "2019-01-01")|> select(month,hires,quits, layoffs)|> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot()+ geom_line(aes(x=month, y = rate, color=measure)) # Build your pipeline one step at a time.`
2. `labor_market|> filter(month >= "2019-01-01")|> select(month,hires,quits,layoffs)|> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot()+ geom_line(aes(x=month, y=rate, color = measure)) # Build your pipeline one step at a time.`
3. `labor_market|>filter(month >= "2019-01-01")|> select(month, hires, quits, layoffs)|> pivot_longer(-month, names_to = "measure", values_to = "rate")|> ggplot()+geom_line(aes(x = month, y = rate, color = measure)) # Build your pipeline one step at a time.`

Student responses

1. `library(dplyr) library(tidyr) library(ggplot2) hiring_plot_data <- labor_market %>% filter(month >= as.Date("2019-01-01")) %>% select(month, hires, quits, layoffs) %>% pivot_longer(cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate") ggplot(hiring_plot_data, aes(x = month, y = rate, color = measure)) + geom_line() # Build your pipeline one step at a time.`
2. `library(tidyverse) labor_market |> filter(month >= "2019-01-01") |> select(month, hires, quits, layoffs) |> pivot_longer( cols = c(hires, quits, layoffs), names_to = "measure", values_to = "rate" ) |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line()`

Student responses

1. `library(tidyverse) labor_market |> filter(month >= as.Date("2019-01-01")) |> select(month, hires, quits, layoffs) |> pivot_longer(cols = c(hires, quits, layoffs),
names_to = "measure", values_to = "rate") |> ggplot(aes(x = month, y = rate, color = measure)) + geom_line(linewidth = 1) + labs(x = "Month", y = "Rate", color =
"Measure") + theme_minimal() # Build your pipeline one step at a time.`

Look at hiring, quits, and layoffs after 2022. How could these patterns make it hard to find a first job even if most people who have jobs keep them?

73 anonymous responses

Reduced hiring limits openings

33

Explains that slower hiring directly creates fewer available positions for job seekers, making it harder to get a first job.

Simple statement of difficulty

9

Concise assertion that getting a job will be harder without further detail.

Low quits reduce turnover

8

Notes that fewer people quitting means fewer vacancies created from employees leaving, so fewer opportunities for new entrants.

Mentions high layoffs (misaligned)

6

Claims that high layoffs reduce available new jobs or links layoffs to fewer openings, a view that conflicts with the provided pattern of low layoffs.

Response themes for question 15

Other responses**6**

Responses the model could not place reliably.

Attributes effect partly to AI/other factors**4**

Suggests that AI or other broader factors, alongside hiring changes, contribute to making it harder for first-time job seekers.

Vague or unclear reason**4**

Responses that are too brief or vague to identify a specific mechanism explaining why finding a first job is harder.

Companies not expanding reduces openings**3**

Argues that firms are less inclined to grow or expand post-COVID, so they create fewer new positions even if current employees keep their jobs.

Student responses

1. After 2022, hiring and quits both declined, while layoffs stayed relatively low. This suggests that employers were not necessarily firing large numbers of workers, but they were hiring fewer new workers. As a result, people who already had jobs could keep them, while new graduates and first-time job seekers faced fewer opportunities to enter the labor market.
2. After 2022, hiring and quits decline while layoffs remain low. Employers hire fewer people, and fewer workers leave jobs that could create replacement openings. Existing workers can therefore keep their jobs while first-time job seekers face fewer openings and more competition. Job stability for current employees does not necessarily mean good opportunities for newcomers.
3. Because all three values are stable and people are not getting laid off or quitting, there will be less opening for new jobs.
4. Because hiring has gone down, there are less jobs for new graduates.
5. Because hiring rates spiked after COVID, people keep there jobs and therefore there are less jobs open to new people in the workforce

Student responses

1. Because most jobs aren't hiring at a rate as large as they did, which means that getting a job is a little more difficult for recent college grads
2. Because the rate at which companies are hiring people is decreasing, so getting hired for your first job is still difficult, even if people who were already employed prior to 2022 were able to keep their jobs
3. companies are hiring less and less people are quitting or being laid off so there are less jobs available for new members of the workforce.
4. Getting a job will be harder because there are less hirings. It is probably a wide combination of everyone trying to budget cut and AI and other things.
5. hire and quit fallen
6. hires are pretty correlated with quits
7. hiring decreased, but quits/layoffs stayed low, so companies aren't bringing in new employees
8. Hiring has decreased even when quitting and layoffs have decreased
9. If most people who have jobs keep them, then there will inherently be less hiring.

Student responses

1. Its about as hard to get a job as before covid. Since there are less hires and less quits that means people are staying in jobs and companies dont have an incentive to hire unexperienced workers
2. People are hiring less so it is harder to get a job. AI might have an effect but the real reason is that there were a lot of people hired and companies dont necessarily need to hire any more people so this has lead to being less hires per year
3. Since hiring has slowed while layoffs stayed relatively low, there are fewer open spots for new workers, which can make finding a first job harder.
4. since hiring is going down, people wont be as likely to get their first job. less people are also quitting so there are fewer job openings
5. The amount of hires has decreased, and a relatively low amount of people are quit or laid off, so it might be difficult to get hired for a first job.
6. The graph depicts that hires have been on a steady gradual decrease since 2022, making it more difficult to find a first job since quits and layoffs are low
7. THe hires are dropping which makes it harder to find a first job

Student responses

1. The hiring rate and quit rate have declined after 2022, while layoffs have stayed the same. People keep their jobs, so less hires happen, making it harder for new workers to break into their career.
2. The hiring rate and quitting rate keep decreasing, indicating that fewer companies are hiring and more positions are stable.
3. The number of hires have decreased from 2022, making it difficult for new workers to find employment.
4. The observed pattern that hiring has been declining since COVID could make it hard to find a first job, even if most people who have jobs keep them.
5. The overall hiring rate after 2022 has steadily declined which could make finding a first job difficult
6. The percentage of hires has decreased since 2022, making it harder to find a first job out of college.
7. The rate of hiring has gone down drastically, meaning it's hard to get hired in the first place as many employers might not be looking to take on new employees at all.
8. There could be less openings
9. There is a downward trend in quits and hires so new grads have hard time breaking in.

Student responses

1. There was so much more hiring then than now
2. These patterns make it hard to find a first job because there aren't enough jobs for everyone looking for a job
3. Well if most people who have jobs keep them, then obviously it will be harder to find a first job. Getting that entry level position will be difficult.

Student responses

1. Because people keep jobs, there are less vacancies. Also, the post-covid era sees a sharp decrease in available jobs.
2. Because the job market is much more strained
3. Big tech was laying off people which resulted in displacing talented engineers taking jobs from new grads.
4. getting a job is gonna be harder
5. Hiring is reaching all time lows and follows a cyclic pattern, so if you're off pattern, the hiring is not that easy for new grads, even if quits are going down, reducing new openings, that won't help the new grads enter the job market
6. it's a tough world. they aren't hiring as much.
7. People with more experince are applyign to entry level jobs
8. since no one is willing to hire the young people, they dont have the experience to get a job no matter what
9. Yes, it could, fewer hires and quits can make it harder to get a first job even if most people keep their jobs. With less quits there are less new jobs

Student responses

1. after 2022 hires fell and quits fell, while layoffs stayed low. Low layoffs means people who have jobs keep them, but low hires plus low quits means very few openings are being created or vacated. A new graduate needs an opening, and in a low-hire, low-fire market there are almost none, so the market can be frozen for first-time seekers while incumbents are fine.
2. Because when people quit or get laid off, they're likely going to try and find a new job if they aren't retiring, meaning the new workforce is competing against an older, more experienced group for work.
3. Hiring and quit both decrease after 2022, so there are fewer new jobs and fewer positions from people leaving. This can make it harder for young people to find there first job.
4. In a "low hire low fire" market, there are not many openings for college graduates to fill because people are not leaving their jobs that much (voluntarily or involuntarily).
5. It might be harder because jobs are more static since there are less layoffs and quits so less jobs open up for new members of the workforce

Student responses

1. People who were fired in covid might be looking for jobs still now and because they have experience its harder for younger people with no experience to get the same jobs. Also, this graph sort of implies that the decrease isnt all because of AI/Covid, but instead a general decline in the economy.
2. yes, i think if people quit their jobs or like layoffs, it should be easier for me to get in
3. You have to account for the layoffs and quits. Those people already have experience, so the hires have to outweigh both of those combined

Student responses

1. amount of layoffs exceeded number of hires and there weren't positions to be filled
2. hiring is decreasing while people are quitting at a slower rates, so there are less jobs available for young people to be hired for
3. More layoffs mean there are less new jobs available. number of hires has also been declining so firms are just hiring less in general. which would make it difficult for people entering the labor market to get new jobs.
4. The combinations of layoffs and quits after covid has made it harder to find a job after 2021
5. There's not as much opportunity for employers to hire, people aren't quitting or getting laid off so there's less open roles.
6. This tells you that it will be challenging to get a job out of college in 2026 because there has been a slowdown in hiring due to overhiring during covid and a general downturn in certain aspects of the economy. More people are keeping jobs so there is not as much natural turnover because people are concerned about how they will get their next job

Student responses

1. Companies go through spurts of lots of hiring and then lots of layoffs, so it is hard to find the exact time when they are hiring a ton.
2. Hiring and quits have fallen since 2022, while layoffs have stayed relatively low.
3. Makes it hard because no one wants to quit as the quitting rate decreased to nearly 0, and layoffs have decreased as well just like hires
4. The hire and quit rate both went down, so harder to get hired for new jobs, and those already having a job doesn't give their positions.
5. There aren't many changes, meaning that it will be hard to make changes to the current job market for new jobs to open
6. When we are just looking at new jobs (open positions), there are fewer hires while there are more layoffs. Essentially, companies are trying to reduce the size of their labor force rather than replace.

Student responses

1. It seems like there are some trends with AI, but it has had a relatively moderate and toned down impact on the job market than one might originally think
2. It's more difficult in a post-covid, post-ai world to get hired as a fresh grad, rather than maintaining a job
3. Looking at the hiring, quits, and layoffs after 2022 it is hard to translate the pattern for layoffs and quits because of covid or ai
4. maybe AI and other facotrs like how lower hirjn gna dquitting affect the people who are lookign for job wven when most poepel keep theior jobs

Student responses

1. because there are not enough jobs cereates more layoffs thennhring
2. fewer hires and quits reduce job openings would make your answer clearer and stronger.
3. gradually decline
4. There aere na sata

Student responses

1. It looks like layoffs are slowly rising and hires are slowly falling. This could make it a little bit harder to get a job for young people but it may not necessarily be because of AI
2. There might not be new job openings, and the hiring rate is relatively flat
3. These patterns make it more difficult because companies are less motivated to expand than they were during covid

Use one result we found today to respond. What have we learned, and what would we still need to find out about AI's role?

0 anonymous responses

No responses were submitted.