

Politics, money, and happiness

Week 9, Lecture 16

<https://live.ds306.org>

October 29, 2026

Picking up from Tuesday

- We investigated what young adults think about AI, financial security, and public policy.
- We combined responses from several files into one table.
- Today we'll finish the policy questions and investigate a new dataset.

Today's questions

1. Where do young independents fall politically?
2. Does money buy happiness?

We'll use grouped summaries, reshaping, correlation, and regression.

Tuesday's dataset

Each row gives one response category for one policy and group.

policies

Which policy divides the parties most?

We want a table of support percentages and the gap between the parties, ordered by the size of the gap:

```
# A tibble: 4 × 5
  item                Democrat Republican   gap size
  <chr>                <dbl>      <dbl> <dbl> <dbl>
1 Increased deportations  18.5        67.2 -48.7  48.7
2 Guarantee legal abortion rights 76.5        29.4  47.1  47.1
3 School sports teams by sex at birth 34.2        79.5 -45.3  45.3
4 Trump's tariffs        9.31        46.8 -37.5  37.5
```

gap is Democrat minus Republican. size is the absolute value of that gap.

A new way to build data pipelines

- Each block contains one step. Connect the circles to build your pipeline.
- Check what each step does to the table as you go.
- Click **Run** to inspect the result in your workspace.
- Use **Check progress** for feedback on your steps, or **Submit plot** for a plot question.
- Use **Help** if you need help with your pipeline.

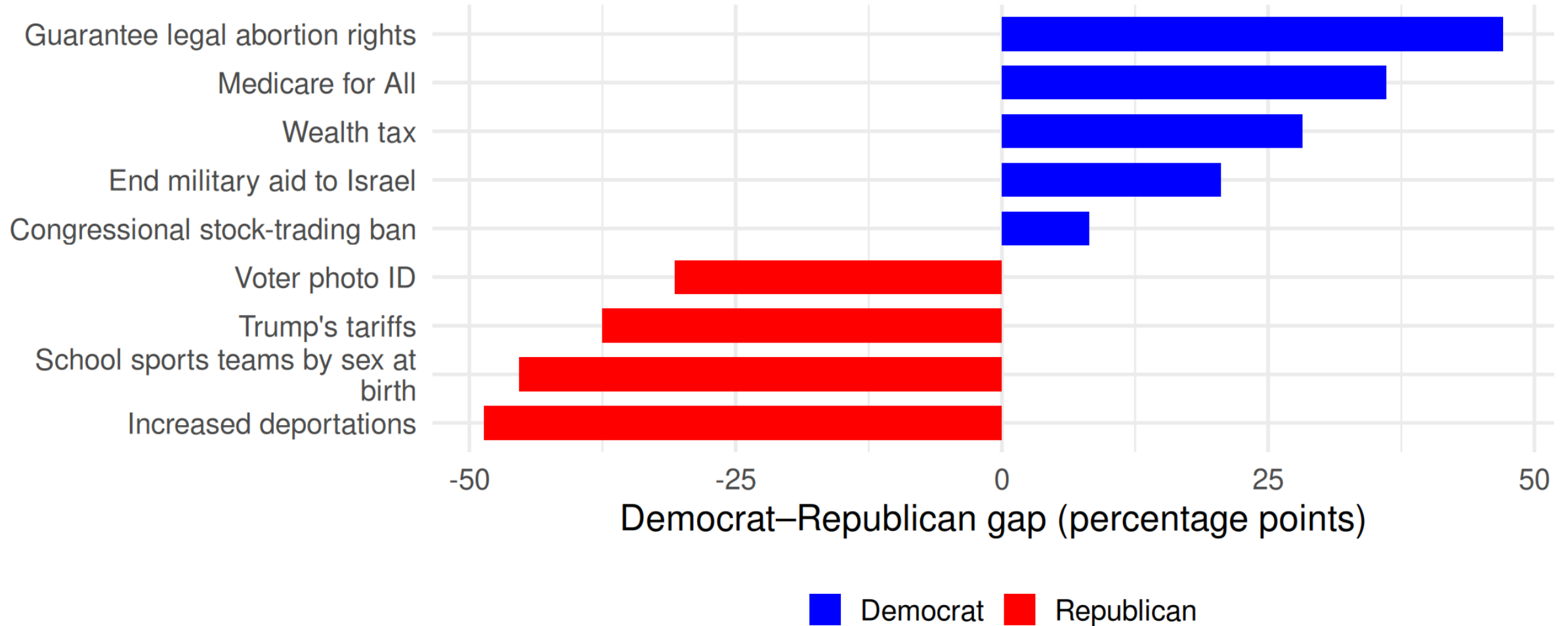
Which policy divides the parties most?

Start with `party_support`, which has one row per policy and party. Build the table above to find the policies with the largest gaps.

```
party_support
```

- Put each party's support in its own column, with one row per policy.
- Calculate gap as Democrat minus Republican, and size as the absolute value of gap.
- Put the policies with the largest size

Which way do the gaps lean?



Build the party-gap plot

Start from `policy_wide`. Use the column block with size on X and policy on Y.

```
policy_wide
```

- Calculate gap as Democrat minus Republican, and size as the absolute value of gap.
- Create leaning: Democrat when gap is positive, Republican otherwise.
- Order the policy labels with `item =`

Which party shares independents' policy preferences?

Do policies that are popular among independents also tend to be popular among Democrats, Republicans, both parties, or neither?

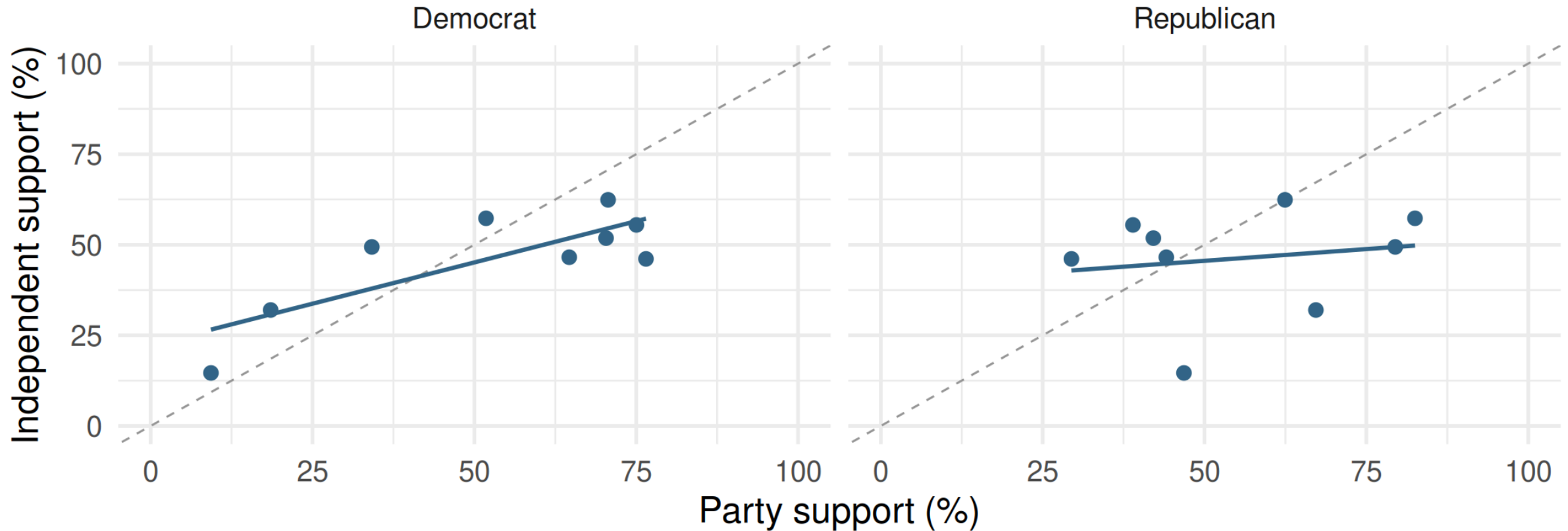
Plot independent vs. parties

Add a scatterplot with **Democrat** on X and **Independent** on Y. Then try **Republican** on X.

```
policy_wide
```

Submit one of your scatterplots. Which party shows the stronger relationship between its support percentages and independents' support percentages?

Independents and party support



Correlation with Democrats: **0.80**. With Republicans: **0.17**.

Find policies closer to Republicans

Measure each party's distance from independent support.

```
policy_wide
```

- Calculate distance_D and distance_R as the absolute differences from Independent support.
- Keep the policies where distance_R is smaller than distance_D.
- Display each policy

A surprising result

Independents are closer to Republicans on **five of the nine policies**.

```
# A tibble: 5 × 4
```

item	Democrat	Independent	Republican
<chr>	<dbl>	<dbl>	<dbl>
1 Congressional stock-trading ban	70.7	62.4	62.4
2 End military aid to Israel	64.7	46.6	44.1
3 Guarantee legal abortion rights	76.5	46.1	29.4
4 Medicare for All	75.0	55.5	38.9
5 Wealth tax	70.3	51.8	42.1

How can this happen when the correlation with Democrats was **0.80**?

How can both results be true?

- Policies with higher Democratic support generally also have higher independent support. This produces the strong correlation.
- Independents often support those policies less strongly. For legal abortion rights, support is **77% among Democrats, 46% among independents, and 29% among Republicans**. Independents are closer to Republicans on this policy.
- Even if independent support were exactly 20 points below Democratic support on every policy, their correlation would be **1**.

Correlation can be high even when the groups' support percentages are far apart.

Q2: Does money buy happiness?

On Tuesday, only about one in four young adults (one in three in this class!) expected to be financially better off than their parents.

How closely is wealth related to how people feel about their lives?

Which country is happiest?



Which of these countries do you think has the highest average life satisfaction?

Two sources of data

- The **World Happiness Report** asks people to rate their lives on a ladder from 0 (worst possible life) to 10 (best possible life).
- The **World Bank** measures GDP per person, adjusted for purchasing power, in constant 2021 international dollars.
- We have two tables to combine: life ratings from 2023–2025 and GDP from 2024.

[World Happiness Report](#) · [World Bank GDP per capita](#)

Two tables to combine

Some rows from `life_ratings`:

```
# A tibble: 4 × 2
  country      happiness
  <chr>         <dbl>
1 Finland      7.76
2 Costa Rica   7.44
3 United States 6.82
4 Singapore    6.58
```

And from `gdp`:

```
# A tibble: 4 × 3
  country      gdp_per_capita gdp_year
  <chr>         <dbl>    <dbl>
1 Costa Rica   27325.    2024
2 Finland     55901.    2024
3 Singapore   134549.    2024
4 United States 75698.    2024
```

Income and happiness: join and plot

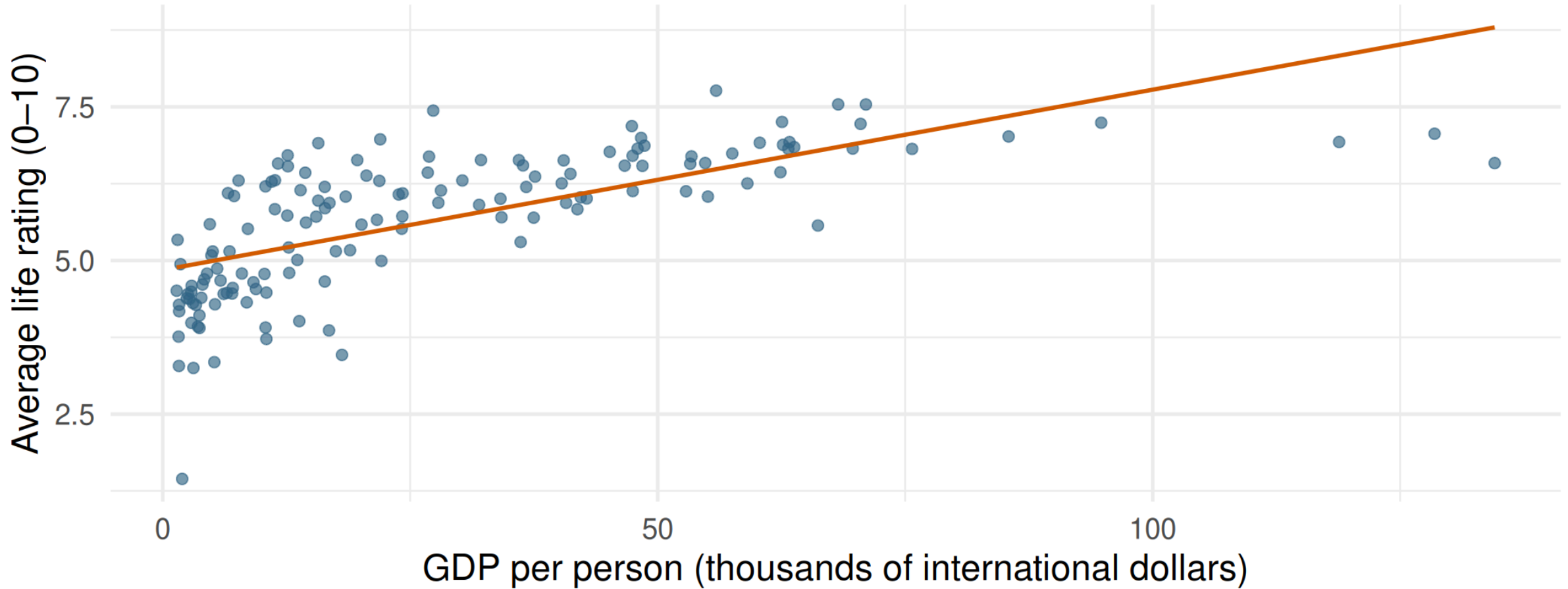
Join the two sources and plot the joint distribution of income and happiness.

```
life_ratings
```

Join `life_ratings` to `gdp` by country, keeping countries that appear in both tables.

Make a scatterplot with `gdp_per_capita` on X and `happiness` on Y. Each point should represent one country.

What do you notice?



Fit a line

Use `income_data`, where `income_thousands` is GDP per person in thousands. Add an **lm** block with the formula `happiness ~ income_thousands`.

```
income_data
```

- Fit a simple regression of happiness on `income_thousands` and inspect the model summary.

Where does the line go wrong?

Look at the scatterplot and fitted line. Where does the line tend to predict too much happiness? Where does it predict too little?

Plot residuals against fitted values

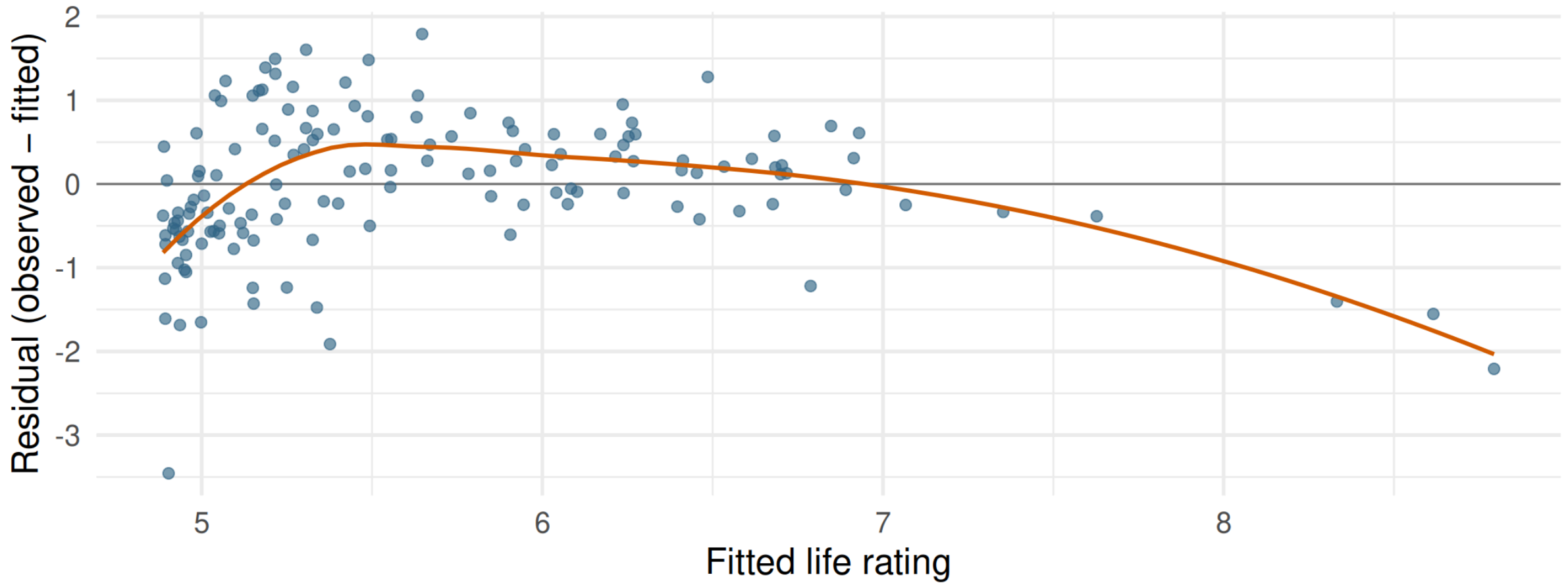
Use `income_data` and the fitted model `income_fit`. Add a column named `fitted` using `predict(income_fit)`.

Calculate `resid` as `happiness - fitted`.

```
income_data
```

Make a scatterplot with `fitted` on X and `resid` on Y. Submit your plot. Do you see a pattern?

What do the residuals show?



Is another \$10,000 equally valuable everywhere?

Would an extra \$10,000 of GDP per person have the same relationship with happiness in a country with GDP per person of \$5,000 as in one with \$80,000? What feature of our plot supports your answer?

Measure income in doublings

- Going from \$10,000 to \$20,000 doubles income. So does going from \$40,000 to \$80,000.
- We can use `log2(gdp_per_capita / 10000)` to count doublings from \$10,000.
- The values 0, 1, 2, and 3 correspond to \$10,000, \$20,000, \$40,000, and \$80,000.

The `income_data` table already contains this column, named `income_doublings`.

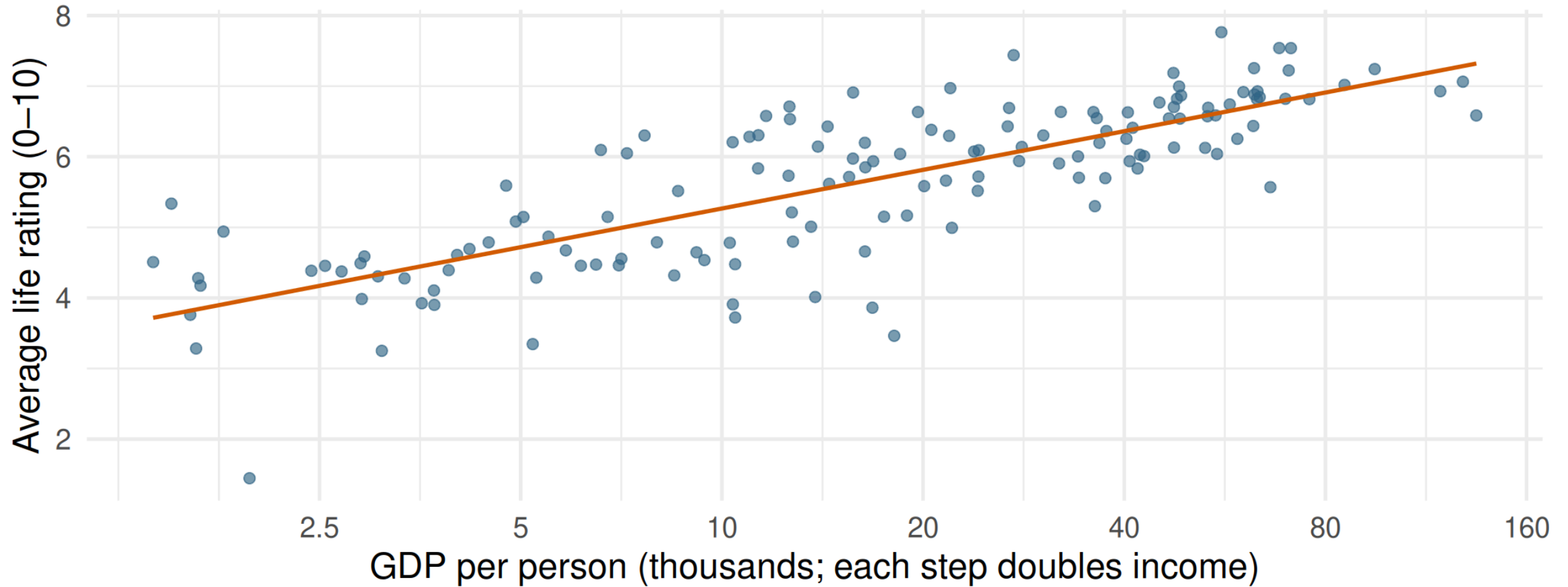
Fit the new line

Start with `income_data`. Add an `lm` block with `happiness ~ income_doublings`.

```
income_data
```

- Fit happiness on `income_doublings`. Inspect the slope and R-squared in the model summary.

Does this look more linear?



Interpret the slope

(Intercept)	income_doublings
5.2670116	0.5476111

According to this line, how much higher is the predicted life rating when income per person doubles? Explain using the fitted coefficient.

How much did the fit improve?

`model_errors` contains each country's residual from both models. Which model has the smaller average absolute residual?

```
model_errors
```

- Calculate `absolute_resid` as `abs(resid)`.
- Group by model.
- Calculate `mean_absolute_residual` as the mean of `absolute_resid` for each model.

Which countries stand out?

- Some countries report much higher life satisfaction than the line predicts.
- We can find them using residuals: observed happiness minus fitted happiness.
- A positive residual means the country's life rating is above its fitted value.

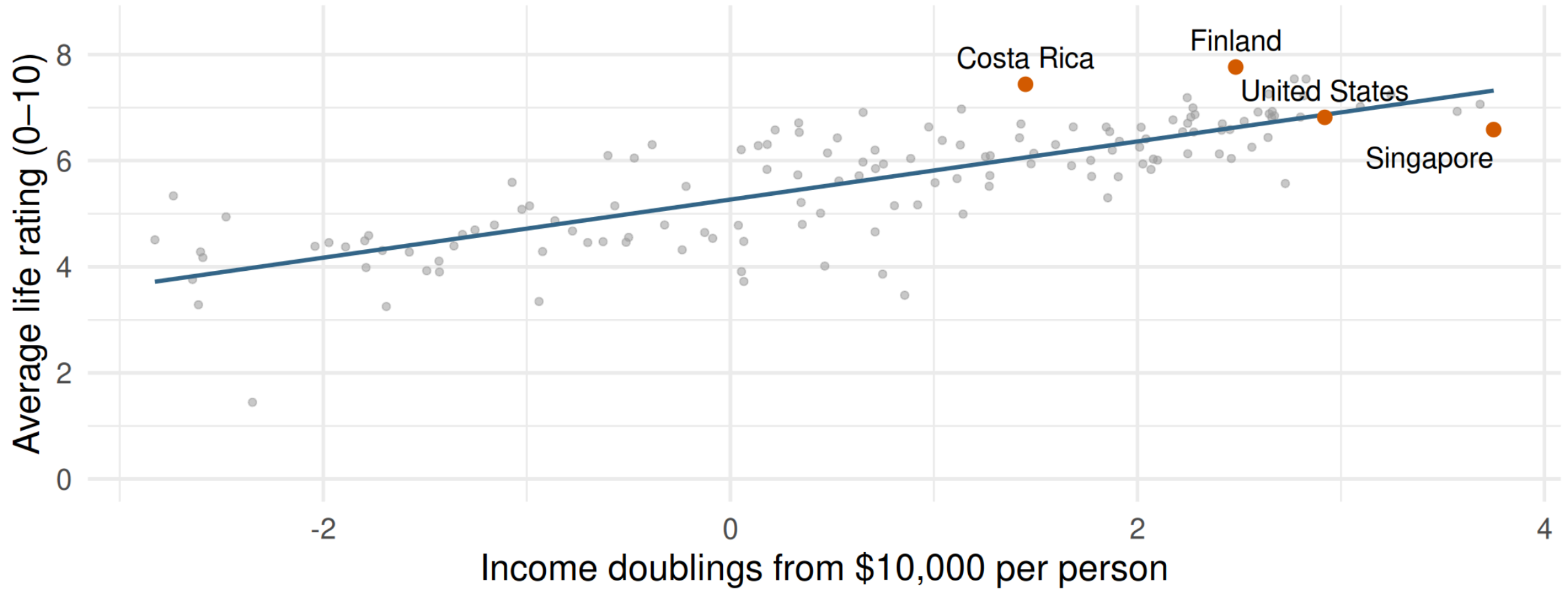
Find the positive surprises

`happiness_residuals` contains a `resid` column from our fitted model. Find the five countries furthest above the line.

```
happiness_residuals
```

- Put the countries with the largest residuals first.
- Keep the first five countries.
- Display country, happiness, gdp_per_capita, and resid.

Money leaves plenty unexplained



What would you investigate next?

Choose a country whose life rating surprised you. What additional data would you want to explain its position? Describe one analysis you would try.

Two investigations

- For the policy questions, we summarized support, reshaped the table, and checked how independents aligned with the parties.
- For happiness, we fitted a regression, investigated a pattern in its errors, and tried a transformed predictor.
- In both cases, inspecting the result helped us decide what to do next.

Week 9, Lecture 16: Politics, money, and happiness

Closed October 29, 2026 at 3:47 PM

16 class questions · 1087 anonymous responses · 0 student questions

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

CLASS RECORD · QUESTION 1

- Put each party's support in its own column, with one row per policy. - Calculate gap as Democrat minus Republican, and size as the absolute value of gap. - Put the policies with the largest size first. - Display item, Democrat, Republican, gap, and size.

73 anonymous responses

1. Other responses73

Student responses

1. `largest_gaps <- party_support %>% pivot_wider( names_from = party, values_from = support ) %>% mutate( gap = Democrat - Republican, size = abs(gap) ) %>% arrange(desc(size)) %>% select(item = policy, Democrat, Republican, gap, size)`
2. `party_support party_support |> tidyr::pivot_wider( names_from = party, values_from = support ) |> dplyr::mutate( gap = Democrat - Republican, size = abs(gap) ) |> dplyr::arrange(dplyr::desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support %>% pivot_wider(names_from = group, values_from = support) %>% mutate( gap = Democrat - Republican, size = abs(gap) ) %>% arrange(desc(size)) %>% select(item, Democrat, Republican, gap, size)`
4. `party_support |> dplyr::filter(group %in% c("Democrat", "Republican")) |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> filter(group != "Independent") |> pivot_wider( names_from = group, values_from = support ) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
2. `party_support |> pivot_wider( names_from = group, values_from = support ) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
3. `party_support |> pivot_wider( names_from = group, values_from = support ) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size) |> slice(1:5)`
4. `party_support |> pivot_wider( names_from = party, values_from = support ) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`

Student responses

1. `party_support |> pivot_wider(names_from = group, values_from = support) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
2. `party_support |> pivot_wider(names_from = group, values_from = support) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
3. `party_support |> pivot_wider(names_from = group, values_from = support) |> mutate(gap = Democrat - Republican, size = abs(gap)) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider( names_from = group, values_from = support ) |> dplyr::mutate( gap = Democrat - Republican, size = abs(gap) ) |> dplyr::arrange(dplyr::desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`

Student responses

1. `party_support |> tidyr::pivot_wider( names_from = party, values_from = support ) |> dplyr::mutate( gap = Democrat - Republican, size = abs(gap) ) |> dplyr::arrange(dplyr::desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(dplyr::desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = (Democrat - Republican), size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = (Democrat - Republican), size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap))item) |> dplyr::arrange(desc(`Wealth tax`))`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::filter(item %in% c("Increased deportations", "Guarantee legal abortion rights", "School sports teams by sex at birth", "Trump's tariffs")) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party support |> tidyr::pivot wider(names from = group, values from =`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(Gap = Democrat - Republican, Size = abs(Gap)) |> dplyr::arrange(desc(Size)) |> dplyr::select(item, Democrat, Republican, Gap, Size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size=abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size))`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(Democrat - Republican)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat,`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::select(item, Democrat, Republican, gap, size) |> dplyr::arrange(desc(size))`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::select(item, Democrat, Republican, gap, size) |> dplyr::arrange(desc(size))`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::arrange(desc(abs(gap)))`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::arrange(desc(gap)) |> dplyr::select(item, Democrat, Republican, gap)`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::mutate(size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::mutate(size = abs(gap)) |> dplyr::select(item, Democrat, Republican, gap, size) |> dplyr::arrange(desc(size))`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat-Republican, size = abs(gap)) |> dplyr::arrange(desc(gap)) |> dplyr::select(item, Democrat, Republican, gap, size)`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat-Republican, size = abs(gap)) |>`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat-Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::select(item, Democrat, Republican) |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::mutate(size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate()`
4. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap,`

Student responses

1. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
2. `party_support |> tidyr::pivot_wider(names_from = group, values_from = support) |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::arrange(desc(size)) |> dplyr::select(item, Democrat, Republican, gap, size)`
3. `party_support |> tidyr::pivot_wider(names_from = item, values_from = support)`
4. `party_support |> tidyr::pivot_wider(names_from = item, values_from = support) |> dplyr::mutate()`
5. `party_support |> tidyr::pivot_wider(names_from = item, values_from = support) |> dplyr::mutate() |> dplyr::arrange(group) |> dplyr::select(group`

Student responses

1. `party_support |> pivot_wider(id_cols=item, names_from=group, values_from=support) |> mutate(gap = Democrat - Republican, size = abs(gap)) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
2. `party_support |> pivot_wider(names_from = group, values_from = support) |> mutate(gap = Democrat - Republican, size = abs(gap)) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`
3. `party_supportparty_support |> filter(group %in% c("Democrat", "Republican")) |> pivot_wider( names_from = group, values_from = support ) |> mutate( gap = Democrat - Republican, size = abs(gap) ) |> arrange(desc(size)) |> select(item, Democrat, Republican, gap, size)`

- Calculate gap as Democrat minus Republican, and size as the absolute value of gap. - Create leaning: Democrat when gap is positive, Republican otherwise. - Order the policy labels with `item = reorder(item, size)`, so that the largest gap appears at the top. - Draw horizontal bars with size on X, policy on Y, and fill determined by leaning.

71 anonymous responses

1. Other responses71

Student responses

1. `ggplot2::ggplot( policy_wide |> dplyr::mutate( gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size) ), ggplot2::aes(x = size, y = item, fill = leaning) ) + ggplot2::geom_col()`
2. `ggplot2::ggplot( policy_wide |> dplyr::mutate( gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ), ggplot2::aes( x = size, y = item, fill = leaning ) ) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item,size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican" ), item_ordered = reorder(item, gap)), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican" ), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "blue", Republican = "red"), name = NULL)`
4. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican,`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item)) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(Gap = Democrat - Republican, Leaning = ifelse(Gap > 0, "Democrat", "Republican"), item = reorder(item, Gap)), ggplot2::aes(x = Gap, y = item, fill = factor(Leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, gap)), ggplot2::aes(x = gap, y = item)) + ggplot2::geom_col()`
4. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item_ordered = reorder(item, gap)), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item_reordered = reorder(item, size)), ggplot2::aes(x = gap, y = item_reordered, fill = factor(leaning))) + ggplot2::geom_col()`
4. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item =`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap>0, 'Democrat', 'Republican'), item_ordered = reorder(item, gap)), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = 'blue', Republican = 'red'), name = NULL)`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(Gap = Democrat - Republican, Size = abs(Gap), Leaning = if_else((Gap > 0), "Democrat", "Republican"), item = reorder(item, Size)) |> dplyr::arrange(desc(Size)), ggplot2::aes(x = Size, y = item, fill = factor(Leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "blue", Republican = "red"), name = NULL)`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) +`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "Blue", Republican = "Red"), name = NULL)`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item_ordered = reorder(item, gap)), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item,gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "Blue", Republican = "Red"), name = NULL)`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item_ordered = reorder(item, size)), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::mutate(leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "blue", Republican = "red"), name = NULL)`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican" ), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
4. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"),`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, Democrat, Republican), item = reorder(item, gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col() + ggplot2::scale_fill_manual(values = c(Democrat = "blue", Republican = "red"), name = NULL)`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap >0, "Democrat", "Republican"), item = reorder(item,gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap>0, "Democrat", "Republican"), item_ordered = reorder(item, gap) ), ggplot2::aes(x = gap, y = item_ordered, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::mutate(leaning = ifelse(gap > 0, "Democrat", "Republican"), policy = reorder(item, size)), ggplot2::aes(x = size, y = policy, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::mutate(leaning=ifelse((gap>=0), "Democrat", "Republican")) |> dplyr::mutate(item=reorder(item, size)) |> dplyr::arrange(desc(gap)), ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::mutate(leaning = if_else(gap > 0 , "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)), ggplot2::aes(x = size, y = item, fill = factor(leaning))) + ggplot2::geom_col()`

Student responses

1. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat - Republican)
|> dplyr::mutate(size = abs(gap)) |> dplyr::mutate(leaning = ifelse(gap > 0,
"Democrat", "Republican"), item = reorder(item, size)), ggplot2::aes(x = size, y
= item, fill = factor(leaning))) + ggplot2::geom_col()`
2. `ggplot2::ggplot(policy_wide |> dplyr::mutate(gap = Democrat-Republican,
leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap),
item_ordered = reorder(item, size)), ggplot2::aes(x = gap, y = item_ordered,
fill = factor(leaning))) + ggplot2::geom_col()`
3. `ggplot2::ggplot(policy_wide |> dplyr::select(item, Democrat, Republican) |>
dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap>0,
"Democrat", "Republican"), size = abs(gap), item = reorder(item,gap)) |>
dplyr::mutate(leaning = ifelse(gap>0, "Democrat", "Republican")),
ggplot2::aes(x = gap, y = item, fill = factor(leaning))) + ggplot2::geom_col() +
ggplot2::scale_fill_manual(values = c(Democrat = "blue", Republican = "red"),`

Student responses

1. `library(tidyverse) policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
2. `policy_wide`
3. `policy_wide library(dplyr) library(ggplot2) policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
4. `policy_wide policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col() + scale_y_discrete(limits = rev)`

Student responses

1. `policy_wide %>% mutate(gap = Democrat - Republican, size = abs(gap),
 leaning = if_else(gap > 0, "Democrat", "Republican")) %>% mutate(item =
 reorder(item, size)) %>% ggplot(aes(x = size, y = item, fill = leaning)) +
 geom_col() + labs(x = "Size (Absolute Gap)", y = "Policy Item", fill = "Leaning")
 + theme_minimal()`
2. `policy_wide %>% mutate(gap = Democrat - Republican, size = abs(gap),
 leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item,
 size)) %>% ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
3. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap),
 leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item,
 size)) |> ggplot2::ggplot(ggplot2::aes(x = size, y = item, fill = leaning)) +
 ggplot2::geom_col()`
4. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning =
 ifelse(gap > 0, "Democrat", "Republican"))`

Student responses

1. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"))`
2. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::mutate(leaning = ifelse(gap > 0, Democrat, Republican)) |> dplyr::mutate(item = reorder(item, size))`
3. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"))`
4. `policy_wide |> dplyr::mutate(gap = Democrat - Republican) size = abs(gap) leaning = ifelse(gap>0, "Democrat", "Republican")`
5. `policy_wide |> dplyr::mutate(gap = Democrat - Republican) |> dplyr::mutate(leaning = if_else(gap > 0, "Democrat", "Republican"))`
6. `policy_wide |> dplyr::mutate(gap=Democrat-Republican, leaning=ifelse(gap >0, "Democrat","Republican"))`

Student responses

1. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
2. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
3. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
4. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`

Student responses

1. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col() + labs( x = "Size of Support Gap", y = "Policy", fill = "Leaning" )`
2. `policy_wide |> mutate( gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col() + labs( x = "Party gap (percentage points)", y = "Policy", fill = "Leaning" )`
3. `policy_wide |> mutate(gap = Democrat - Republican, size = abs(gap), leaning = ifelse(gap > 0, "Democrat", "Republican"), item = reorder(item, size)) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`

Student responses

1. `policy_wide |> dplyr::mutate(gap = Democrat - Republican, size = abs(gap)) |> dplyr::mutate(leaning = if_else(gap > 0, "Democrat", "Republican")) |> dplyr::mutate(item = reorder(item, size)) |> ggplot2::ggplot() + ggplot2::geom_col(aes(x = size, y = item, fill = leaning))`
2. `policy_wide |> mutate(gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size)) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
3. `policy_wide |> mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)) |> ggplot(aes(x=size, y=item, fill = leaning)) + geom_col()`
4. `policy_wide |> mutate(gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size)) |> ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`

Student responses

1. `policy_wide %>% mutate( gap = Democrat - Republican, size = abs(gap), leaning = if_else(gap > 0, "Democrat", "Republican"), item = reorder(item, size) ) %>% ggplot(aes(x = size, y = item, fill = leaning)) + geom_col()`
2. `x <- policy_wide |> mutate(gap = Democrat - Republican, leaning = ifelse(gap > 0, "Democrat", "Republican"), size = abs(gap), item = reorder(item, size)) |> arrange(desc(gap)) ggplot(x, aes(x = item, y = size, fill = leaning)) + geom_col() + coord_flip() + labs(x = "Policy", y = "Size of Gap", fill = "Leaning")`

Do policies that are popular among independents also tend to be popular among Democrats, Republicans, both parties, or neither?

67 anonymous responses

1. Other responses67

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1. 1

2. 1

3. 1

4. 1

5. 1

6. 1

7. 1

8. 1

9. 1

Student responses

1. 1

2. 1

3. 2

4. 2

5. 2

6. 2

7. 2

8. 2

9. 2

Student responses

1.2

2.2

3.2

4.2

5.2

6.2

7.2

8.2

9.2

Student responses

1.2

2.2

3.2

4.2

5.2

6.2

7.2

8.2

9.2

Student responses

1.2

2.2

3.2

4.2

5.2

6.2

7.2

8.2

9.2

Student responses

1.2

2.2

3.2

4.2

5.2

6.2

7.2

8.2

9.2

Student responses

1. 2

2. 3

3. 3

4. 3

Submit one of your scatterplots. Which party shows the stronger relationship between its support percentages and independents' support percentages?

60 anonymous responses

1. Other responses 60

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

1.

2.

3.

4.

5.

6.

7.

8. democrat

9. democrat is a stronger plot

Student responses

1. Democrat seems to show more of a liner trend.
2. Democrat shows a stronger relationship.
3. democrats have strong relatiobship
4. democrats show stronger relationships to independents
5. Democrats show the stronger relationship. In the Democrat plot the points line up tightly along an upward trend (correlation 0.80), so policies with high Democratic support also have high independent support.
6. Democrats show the stronger relationship. In the Democrat plot, points rise together in a clear positive trend: policies Democrats support more, Independents also tend to support more. In the Republican plot, points are scattered with no clear line, so knowing Republican support tells you little about Independent support.
7. Democrats show the stronger relationship. In this plot, as Democratic

Student responses

1. It seems like correlation is more clear between democrat and ind than republican
2. Republican vs independent
3. seems to be a positive correlation between democrats and independents
4. Seems to be closer?
5. Shows Republican vs Independent for specific topics
6. This plot shows a possible relationship between Democrats and Independents

CLASS RECORD · QUESTION 5

- Calculate distance_D and distance_R as the absolute differences from Independent support. - Keep the policies where distance_R is smaller than distance_D. - Display each policy and its Democrat, Independent, and Republican support percentages.

72 anonymous responses

1. Other responses72

Student responses

1. `filtered_policies <- policy_wide %>% # 2. Calculate the absolute distance from Independent support for both parties mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) %>% # 3. Keep only policies where distance_R is smaller than distance_D filter(distance_R < distance_D) %>% # 4. Display each policy alongside Democrat, Independent, and Republican percentages select(policy, Democrat, Independent, Republican)`
2. `policy_wide`
3. `policy_wide library(dplyr) policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
4. `policy_wide policies %>% mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) %>%`

Student responses

1. `policy_wide policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
2. `policy_wide %>% mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) %>% filter(distance_R < distance_D) %>% select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`

Student responses

1. `policy_wide |> dplyr::mutate( distance_D = abs(Democrat-Independent),distance_R = abs(Republican-Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(differenceD=abs(Democrat-Independent), differenceR=abs(Republican-Independent)) |> dplyr::filter(differenceR<differenceD)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican, distance_D, distance_R)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance D)`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R <`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_D > distance_R) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican, distance_D, distance_R)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican) ) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter((distance_R < distance_D)) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R <`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican, distance_D, distance_R)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican, distance D, distance R)`

Student responses

1. `policy_wide |> dplyr::mutate(distance_d = abs(Independent - Democrat), distance_r = abs(Independent - Republican)) |> dplyr::filter(distance_r < distance_d) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R <`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R <`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |>`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_d = abs(Independent - Democrat), distance_r = abs(Independent - Republican)) |> dplyr::filter(distance_r < distance_d) |> dplyr::select(item, Democrat, Independent, Republican)`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
4. `policy_wide |> dplyr::mutate(distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
5. `policy_wide |> dplyr::mutate(distance_D = Independent - Democrat, distance_R = Independent - Republican) |> dplyr::filter(distance_R <`

Student responses

1. `policy_wide |> dplyr::mutate(distance_D= abs(Independent - Democrat), distance_R = abs(Independent - Republican) ) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
2. `policy_wide |> dplyr::mutate(distance_R = abs(Republican - Independent), distance_D = abs(Democrat - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Democrat, Independent, Republican)`
3. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
4. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
5. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R <`

Student responses

1. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
2. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
3. `policy_wide |> mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican) 2 * pt(-abs(-2.12), df = 33)`
4. `policy_wide |> mutate( distance_D = abs(Independent - Democrat), distance_R = abs(Independent - Republican) ) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
5. `policy_wide |> mutate( distance_D = abs(Independent - Democrat)`

Student responses

1. `policy_wide |> mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
2. `policy_wide |> mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
3. `policy_wide |> dplyr::mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> dplyr::filter(distance_R < distance_D) |> dplyr::select(item, Independent, Democrat, Republican)`
4. `policy_wide |> mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> filter(distance_R < distance_D) |> select(item, Democrat, Independent, Republican)`
5. `policy_wide |> mutate(distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent)) |> filter(distance_R <`

Student responses

1. `policy_wide %>% mutate( distance_D = abs(Democrat - Independent), distance_R = abs(Republican - Independent) ) %>% filter(distance_R < distance_D) %>% select(item, Democrat, Independent, Republican)`

Which of these countries do you think has the highest average life satisfaction?

71 anonymous responses

1. Other responses 71

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.0

7.0

8.0

9.0

Student responses

1.0

2.0

3.0

4.0

5.0

6.1

7.2

8.2

9.3

Student responses

1.3

2.3

3.3

4.3

5.3

6.3

7.3

8.3

Join ``life_ratings`` to ``gdp`` by ``country``, keeping countries that appear in both tables. Make a scatterplot with ``gdp_per_capita`` on X and ``happiness`` on Y. Each point should represent one country. Submit your plot and describe the shape of the relationship.

65 anonymous responses

1. Other responses65

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

1.

2.

3.

4.

5.

6.

7.

8. as GDP per capita increases, its a logarithmic increase in happiness, initial big increase, then tapers out

9. blue

Student responses

1. happiness rises steeply as GDP per capita goes from low to moderate levels, then flattens out among richer countries. That's a positive relationship that is curved (concave, roughly logarithmic) rather than linear, with a fair amount of scatter around the trend
2. Happiness tends to increase with GDP per capita, but the relationship levels off. At low incomes, happiness rises steeply as GDP increases, while among richer countries happiness is mostly between about 6 and 7.5 and does not keep rising much with extra income. So the shape looks curved rather than a straight line.
3. It looks log, like it spikes up and then flatlines
4. logarithmic
5. Logarithmic relationship/shape of the plot
6. low income countries have a lower happiness rating but once a country

Student responses

1. positive but curved (concave), with diminishing returns. Happiness climbs steeply as GDP rises across poor and middle-income countries, then flattens out among the richest, the points level off rather than continuing up at the same rate.
2. positive correlation
3. positive relationship
4. positive shape, curved?
5. Positive, nonlinear (concave) relationship. Among lower-income countries (under about \$20k), happiness rises steeply, from roughly 3–5 up to about 6. Above about \$25k, the curve flattens: happiness stays around 6–7.5 even as GDP per capita grows past \$100k, showing diminishing returns. Poorer countries also vary much more in happiness than richer ones. A log scale on the x-axis would likely make it look more linear.

Student responses

1. There is a positive relationship between GDP per capita and happiness. Countries with higher GDP per capita generally have higher happiness scores. The relationship appears nonlinear, with happiness increasing more slowly at higher income levels.
2. There is a positive relationship between GDP per capita and happiness. Countries with higher GDP per capita tend to have higher happiness scores, but the relationship is not perfectly linear.
3. There is an initial upwards trend between gdp and happiness, but relationship doesn't appear to be linear and plateaus.
4. This is generally a positive relationship, having a higher gdp per capita generally coorelates with a higher happiness rating.

- Fit a simple regression of happiness on income_thousands and inspect the model summary.

66 anonymous responses

1. Other responses66

Student responses

1. `happiness_model <- lm(happiness ~ income_thousands, data = income_data)`
`summary(happiness_model)`
2. `happiness_regression <- stats::lm(happiness ~ income_thousands, data = income_data)`
`summary(happiness_regression)`
3. `income_data`
4. `income_data income_data |> lm(happiness ~ income_thousands, data = _) |>`
`summary()`
5. `income_data model <- lm(happiness ~ income_thousands, data = income_data)`
`summary(model)`
6. `income_data model <- lm(happiness ~ income_thousands, data = income_data)`
`summary(model)`
7. `income_data lm(formula = happiness ~ income_thousands, data = income_data)`

Student responses

1. `income_data |> lm(formula = happiness ~ income_thousands) |> summary()`
2. `income_data |> lm(happiness ~ income_thousands, data = _)`
3. `income_data |> lm(happiness ~ income_thousands, data = _)`
4. `income_data |> lm(happiness ~ income_thousands, data = _) |> summary()`
5. `income_data |> lm(happiness ~ income_thousands, data = _) |> summary()`
6. `income_data |> mutate(income_thousands = gdp_per_capita / 1000) |>
 lm(happiness ~ income_thousands, data = _) |> summary()`
7. `income_data |> stats::lm(happiness = income_thousands, data = _)`
8. `income_data |> stats::lm(happiness ~ gdp_per_capita, data = _)`
9. `income_data |> stats::lm(happiness ~ gdp_per_capita*income_thousands, data
 = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
2. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
3. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
4. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
5. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
6. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
7. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
8. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
9. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
2. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
3. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
4. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
5. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
6. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
7. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
8. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
9. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
2. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
3. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
4. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
5. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
6. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
7. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
8. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
9. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
2. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
3. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
4. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
5. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
6. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
7. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
8. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
9. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_thousands, data = _)`
2. `income_data |> stats::lm(happiness ~ income_thousands, data = _) |> summary()`
3. `income_data |> stats::lm(happiness~income_thousands, data = income_data)`
4. `income_data |> stats::lm(income_thousands ~ happiness, data = _)`
5. `income_model <- lm(happiness ~ income_thousands, data = income_data)
summary(income_model)`
6. `income_model <- lm(happiness ~ income_thousands, data = income_data)
summary(income_model)`
7. `income_model <- lm(happiness ~ income_thousands, data = income_data)
summary(income_model)`
8. `lm(happiness ~ income_thousands, data = income_data) |> summary()`

Student responses

1. `model <- lm(happiness ~ income_thousands, data = income_data);
summary(model)`
2. `model_raw <- lm(happiness ~ income_thousands, data = income_data)
summary(model_raw)`
3. `x <- lm(happiness ~ income_thousands, data = income_data) summary(x)`

Look at the scatterplot and fitted line. Where does the line tend to predict too much happiness? Where does it predict too little?

71 anonymous responses

1. Overpredicts for highs, under for lows23

Respondent says the line predicts too much happiness at high income values and too little at low incomes.

2. Prediction error relative to zero point11

Respondent refers to accuracy above/below zero or a baseline value rather than income-based misprediction patterns.

3. Overpredicts for middle, under for extremes8

Respondent believes the line overestimates happiness for middle-income countries and underestimates it for both very low and very high incomes, implying a curved relationship.

4. Inaccurate mainly at tails6

Respondent notes the model is inaccurate primarily at the extreme low or high ends of income (tails) without specifying direction.

Response themes for question 9

1. Overpredicts at low, under at high⁶

Respondent specifically says the line predicts too much happiness at low incomes and too little at high incomes.

2. Unclear or single-word response⁶

Response too brief or unclear to assign to a substantive misprediction pattern.

3. Under for lower, over for higher incomes (swapped)⁶

Respondent states the line underpredicts at lower incomes and overpredicts at higher incomes (opposite of some others).

4. Underpredicts for highs, over for lows⁵

Respondent says the line predicts too little happiness for high-income observations and too much for low-income ones.

Student responses

1. At high amounts of gdp per capita, the model overpredicts happiness. At low amounts it underpredicts
2. it over predicates happiness for wealthy countries and under predicates poorer countries
3. It predicts slightly less happiness below 75 GDP per person and slightly too much above 75 GDP per person.
4. It predicts too much happiness for high income, too little for low income
5. it predicts too much happiness for the lowest incomes and too little for the highest
6. It predicts too much happiness when the income is high, and too little happiness when the income is low
7. Its predicts too little happiness for the lowest income countries and too much happiness for the highest income countries.

Student responses

1. overpredicts how happy you are later in the scatterplot (richer) and underpredicts early on (less rich)
2. Tends to predict too much happiness when there is a high gdp per person. It predicts too low when the gdp per person is lower
3. the higher gdp, the more the line tend to predict too much happiness
4. The line overestimates happiness at the highest incomes and underestimates it at the mid to low incomes.
5. The line predicts too much happiness toward higher GDPs, overestimating the happiness in that region.
6. The line predicts too much happiness when income increases towards the end of the right side of the scatterplot and the model predicts too little happiness when income is smaller, or the left side of the scatterplot.
7. The line tend to predict too mush for very high GDP population and little for

Student responses

1. The regression line overpredicts happiness at a high gdp per capita. It predicts too little happiness for a lower gdp per capita.
2. too much happiness at 150 thousand and too little at 0
3. too much at high gdp, too little at low gdp
4. too much on the highs, too little on the lows
5. when the income is very high, the line tends to predict too much happiness

Student responses

1. Around a gpd per capita of 25000 has lots of data with a postiive trend, but data after does not show a strong positive trend as before
2. becomes less accurate as gdp increases
3. Fitness is too low
4. For low income(gdp), it predicts lower, and for high income countries, it predicts higher
5. it over predicts at the high end of gdp and under predicts in the lower end and middle.
6. It predicts too high when income is higher and too low when income is lower so towards the ends
7. Predicts too much above 5.0, too little below 5.0
8. The fitted line overpredicts happiness at higher GDPs and underpredicts at lower GDPs around 25 000

Student responses

1. The line predicts too much happiness when the points are too below the line
2. The poorest countries (near \$0–5k). Many sit at happiness 3–4.5, below where the line starts. The richest countries (above about \$80k). The line keeps climbing, but their actual happiness levels off around 6.5–7. Lower-middle and middle-income countries (about \$10k–50k). Happiness has already jumped to about 6–7, faster than the line rises.

Student responses

1. over predicts happiness in the richest and poorest countries, under predicts in countries around 25-50 thousand GDP per capita
2. overpredicted at lower and higher gdp. predicts too little in the middle
3. The line predicts too much happiness at lower incomes and higher incomes, but it predicts too little in middle incomes because its a logarithmic curve
4. the line tends to predict too much happiness for countries with middle incomes and too little happiness for countries with very low or very high incomes, maybe c curve
5. The line tends to predict too much happiness for very low- and very high-income countries, and too little happiness for countries in the middle, because the relationship is curved rather than perfectly linear.
6. The line tends to underestimate happiness for some middle-income countries and overestimate happiness for some high-income countries. The

Student responses

1. at the tails
2. I think it is least accurate as you get more money as the real line flatlines, so when you have a lot of money the happiness goes down, but if you don't have that much money, smaller increases and decreases are pretty constant.
3. Predicts too much happiness for very rich countries and too little for rather mid to poorer countries
4. the ends
5. The line predicts too much happiness for the richest countries, where most points sit below the line at high GDP per person. It predicts too little in the middle, roughly \$10,000 to \$30,000 per person, where many countries sit above the line. That pattern is what you would expect if the real relationship flattens out instead of staying straight.
6. When having high income happier, too little income little happiness

Student responses

1. It predicts too much happiness at low gdp, and too little at high gdp
2. It predicts too little happiness for middle ground GDP and too much happiness for lower GDP
3. predicts too much at the beginning and too little at the end
4. the line predicts too much happiness at lower gdp and less at higher gdp
5. underpredict low and vice versa
6. We overpredict happiness at very low GDP, and under predict at middling levels of GDP

Student responses

1. at the end of the line
2. income
3. people with too much wealthy are not happier
4. The line tends to predict too much happiness for very high-income countries and too little happiness for some middle-income countries.
5. too much happiness at low GDP
6. When gdp is small.

Student responses

1. Between 10 and 50 thousand, the model tends to underpredict happiness. Upwards of 50 thousand it tends to more frequently overpredict happiness
2. It predicts too little towards a large x
3. The line predicts too little at the beginning of the scatterplot, while it predicts too much towards the end with the richer countries.
4. the line predicts too much for richer, and it predicts too little around 30 gdp per capita
5. under predict for lower incomes and over predict for higher incomes
6. Underpredicts happiness in the first 25 thousands. And overpredicts in the top GDP countries.

Student responses

1. The graph when its towards the end overpredicts happiness because at one point there isn't enough data points to justify our prediction
2. The line predicts too little happiness for low- and middle-income countries. The line predicts too much happiness for the richest countries.
3. The linear model tends to overpredict happiness for countries with very low GDP per capita and for countries with very high GDP per capita.
4. We tend to overpredict ahppaines for poor people and underpredict happiness for rich people
5. We tend to overpredict happiness at lower x values and underpredict at higher x values.

Make a scatterplot with `fitted` on X and `resid` on Y. Submit your plot. Do you see a pattern?

64 anonymous responses

1. Other responses 64

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Student responses

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
9. As fitted increases, residual decreases.

Student responses

1. es, there's a clear arch (upside-down U) instead of random scatter around 0. At the lowest fitted values (about 4.9), residuals are mostly negative, down to about -3.5. Around fitted 5.2–5.6 they're mostly positive (up to about +1.8). At the highest fitted values (above 7), they're all negative (about -0.4 to -2.2). So the line overpredicts happiness for the poorest and richest countries and underpredicts it for middle-income ones. The spread is also widest at low fitted values. Both patterns me
2. It is inversely proportional. Resid vs fitted
3. Resid tend to be larger when happiness level increases
4. The residuals vs fitted values plot shows some relationship between the two
- which implies that the true model is not linear
5. there is a curved pattern
6. There is a parabolic shape

Student responses

1. Yes, I see a pattern. The residuals are mostly positive in the middle, around fitted values of about 5.2 to 5.6, and mostly negative at low and especially at high fitted values. That curved shape suggests the straight line is missing the flattening relationship, rather than the residuals just scattering randomly around zero.
2. yes, there's a clear pattern. The residuals aren't a flat, formless band around zero, they curve. They start low, rise to positive in the middle range of fitted values, then bend back down and go negative at the high end.

Would an extra \$10,000 of GDP per person have the same relationship with happiness in a country with GDP per person of \$5,000 as in one with \$80,000? What feature of our plot supports your answer?

75 anonymous responses

1. Effect differs: diminishing returns43

Students state that an extra \$10,000 has a larger positive association with happiness in poorer countries than in richer ones, citing the plot's curved pattern that flattens at higher incomes.

2. Says effect differs without detail10

Student indicates the effect is not the same but gives no supporting reasoning or plot feature.

3. Says effect is the same5

Student asserts the relationship would be the same across low- and high-income countries.

4. Notes fitted line decreases at high income4

Response themes for question 11

1. References utility function or concave regression⁴

Student invokes economic utility shapes or concave regression (e.g., square-root utility, concave fitted line) to explain diminishing marginal benefit of income.

2. States poorer person would be happier with increase⁴

Student directly asserts that the person in the poorer country would experience more happiness from the extra \$10,000, without referencing plot shape or fitted line.

3. Mentions parabolic/peak pattern³

Student says the plot shows a peak or parabola so increases at very high incomes might decrease happiness, implying a non-monotonic relationship.

4. Other responses²

Student responses

1. _..._ :...: \ . * :...: :...: :...: :...: \ :...: ' ! \ :...: ' * * ^ ^ / \ ^ / \ / _ _ \ / \ / _ _ \ | [] | / _ _ \ | [] |
| | | [] | | _ _ | | _ _ | | _ _ | ~~~~~
2. An extra 10000 GDP would more greatly benefit people in country with lower GDP as the model is fairly accurate
3. I think an extra 10,000 would have a higher impact on happiness in a country with 5,000 GDP per person because happiness sees a drastic increase toward the beginning of the plot
4. If you already have more money, the additional income would not increase happiness as much as in a lower GDP country.
5. no
6. No Diminishing returns / concave curve: The relationship between GDP per person and happiness is logarithmic/concave rather than linear. The slope is very steep

Student responses

1. No the extra money would have a different impact on the two countries because that added money can be more valuable based on the current income each country already makes
2. No you dont got as happy as the person making 5,000. the feature thatll support this is that the curve flattens at the end
3. No, an extra \$10,000 does not have the same relationship with happiness in both countries because in lower income countries an extra \$10,000 represents a massive percentage increase in resources.
4. No, an extra \$10,000 would probably have a bigger relationship with happiness in a lower-income country than in a higher-income country, because the plot curves and starts to flatten as GDP per person gets higher.
5. No, as GDP increases, there are diminishing returns. This is explained because our graph is not linear

Student responses

1. No, because according to the theory of diminishing marginal utility, if the person with \$80,000 gets an additional 10,000 dollars, they are not as happy as the person who started with 5000 dollars and received an additional 10,000
2. no, because the relationship is not linear, as demonstrated by the better fit line being nonlinear. 10k of GDP would have a significantly larger effect at 5k than 80k.
3. No, because the scatterplot rises steeply at low GDP and flattens at high GDP, an extra \$10,000 matters much more to a country at \$5,000 than to one at \$80,000.
4. No, I don't think the effect is the same. An extra \$10,000 would likely have a bigger relationship with happiness in a poorer country than in a richer country. The scatterplot shows that happiness increases quickly at lower GDP levels, but the increase becomes smaller at higher GDP levels.

Student responses

1. No, I don't think the effect would be the same. An extra \$10,000 would likely make a bigger difference in happiness for a country with \$5,000 GDP per person than one with \$80,000. The scatterplot shows a curved relationship that starts to flatten at higher incomes. This suggests that happiness increases with income, but the increase gets smaller as countries become wealthier.
2. No, increasing the amount of money at higher GDP has diminishing returns because the relationship is not linear. There is a horizontal asymptote of happiness.
3. No, it is like diminishing marginal returns in econ, because if you have less you are more happy
4. no, it would give more happiness to ppl with less gdp, as we can see by the higher residual there
5. No, it would not be the same. An extra \$10,000 would matter much more at

Student responses

1. No, the 10,000s has a different impact depending on the local economy. The poor fit of the linear model supports my answer, we should use a non-linear model instead
2. no, the changing slope for different x supports this answer, it's not exactly linear from the plot
3. No, the curve in the plot
4. No, the extra 10,000 would be much more impactful for the country with a gdp of 5,000
5. No, the extra GDP would be more significant for individuals with 5000 average GDP. The plots whows that the model is not linear
6. No, the plot is nonlinear meaning the marginal utility is different depending on how much money you have
7. No, the regression line is concave down which indicates diminishing marginal

Student responses

1. No! This does not mean the same to all. The benefit of the additional good is not as high for the person that already has a higher GDP per person. We see this by the model showing it constantly thinking higher income countries will be happier than they actually are and lower income countries with less happiness than observed
2. No. An additional \$10,000 in GDP per capita appears to be associated with a larger increase in happiness for lower-income countries than for higher-income countries.
3. No. An extra \$10,000 buys much more happiness at \$5,000 than at \$80,000. The feature that supports this is the curvature of the scatterplot.
4. No. An extra \$10,000 of GDP per person is likely to have a larger positive relationship with happiness in a poorer country than in a richer country. The scatterplot shows a curved pattern that becomes flatter at higher income levels, suggesting diminishing returns to income.

Student responses

1. No. In the scatterplot, going from about \$5k to \$15k is linked to happiness rising from roughly 4.5 to about 6, a gain of about 1.5 points. Going from \$80k to \$90k shows almost no change, since happiness is already flat around 7. The feature that shows this is the curve's shape: it's steep at low incomes and flattens at high incomes (diminishing returns). The residual plot backs this up, because its arch shows a single straight-line slope can't describe both ends. This suggests happiness relates
2. No. We see this in the negative residual in high income countries. The same increase in income appears to contribute less to their happiness than for low income countries.
3. The 15k will have a larger increase in happiness than the 90k. Our line shouldn't be linear because as GDP increase the rate at which happiness is increasing decreases.
4. The extent of the person who earn \$80,000 will be less happy than the one

Student responses

1. utility functions are usually square root of consumption so increase in already high consumptions have a much smaller effect than the same jump when your consumption is already low

Student responses

1. I think no? Not really sure what the question is saying. I think the plot shows a inverse relationship that the lower negative residual tends to be vad. Residual is happiness - fitted, so being overpredicted makes you sadded
2. it would have a different relationship because of scale
3. no
4. No
5. no it would not have the same relationship
6. No it wouldny be the same, that amount of money is different for each of teh situations in that 10k is much more a help to teh 5k gdp than the 80k gpd population
7. no, it would affect the person of 5000 gdp more than the 80000, since it is a greater relative increase
8. No. There is no linear relationship because income would matter less in a

QUESTION 11 · Says effect differs without detail · 2/2

Student responses

1. Probably, not, since different countries would have different standard of living.

QUESTION 11 · Says effect is the same · 1/1

Student responses

1. it would negatively impact
2. No, GDP is viewed differently in different countries depending on their utility.
3. No, since there would still be the added affects of living in that country versus the other. Independent
4. yes
5. yes

Student responses

1. According to our graph, happiness may increase but then decrease as the GDP goes higher.
2. It would not have the same relationship because our fitted line indicates that as GDP increases happiness decreases.
3. It would not have the same relationship because the slope would be positive for the first group but could be negative for the second group.
4. No. Because the fitted line increases first, then decreases.

Student responses

1. 1
2. Its different the fittingline support this
3. no the wealthy diff is too big the guy with 5000 would only feel pain living around with other rich guy
4. No, if you have 5000, the plot is good at showing the relationship with happiness, but not with higher incomes

QUESTION 11 · States poorer person would be happier with increase · 1/1

Student responses

1. It would not have the same relationship because the plot shows that happiness is more associated with income at lower income levels.
2. No, because there is a bigger increase among the person with less money
3. No, the person in the poorer country would give more happiness, the model is not linear
4. The person with 5,000 would be happier with the 10,000. In our plot we see bigger residuals at lower incomes

Student responses

1. Absolutely not. Adding \$10,000 to a country with a lower GDP benefits way more according to our non-linear model than a higher GDP country, which could arguably become less happy as a result based on our model!
2. No it would not, there is a clear parabolic peak of happiness at a certain GDP, so at 5,000 increasing money will increase happiness but at 80,000 increasing money will actually decrease happiness
3. No, having a different GDP per person will end up having different relationship with happiness based on the residual plot as it shows a trend of going down after a specific income

- Fit happiness on income_doublings. Inspect the slope and R-squared in the model summary.

64 anonymous responses

1. Other responses64

Student responses

1. `#income_data |> lm(formula = happiness ~ income_doublings, data = income_data)`
2. `check_income <- stats::lm(happiness ~ income_doublings, data = income_data) summary(check_income)`
3. `income_data income_fit2 <- lm( happiness ~ income_doublings, data = income_data ) summary(income_fit2)`
4. `income_data model <- lm(happiness ~ income_doublings, data = income_data) summary(model)`
5. `income_data %>% lm(happiness ~ income_doublings, data = .) %>% summary()`
6. `income_data |> lm(formula = happiness ~ income_doublings) |> summary()`
7. `income_data |> lm(happiness ~ income_doublings, data = _) |> summary()`
8. `income_data |> lm(happiness ~ income_doublings, data = _) |> summary()`

Student responses

1. `income_data |> lm(happiness ~ income_doublings, data = _) |> summary()`
2. `income_data |> lm(happiness ~ income_doublings, data = df)` To inspect the slope and R-squared: `income_data |> lm(happiness ~ income_doublings, data = df) |> summary()`
3. `income_data |> lm(happiness ~ income_doublings, data = income_data) |> summary()`
4. `income_data |> lm(happiness ~ income_doublings) |> summary()`
5. `income_data |> stats::lm(happiness ~ income_doublings , data = _)`
6. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
7. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
8. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
9. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
2. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
3. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
4. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
5. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
6. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
7. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
8. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
9. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
2. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
3. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
4. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
5. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
6. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
7. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
8. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
9. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
2. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
3. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
4. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
5. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
6. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
7. `income_data |> stats::lm(happiness ~ income_doublings, data = _)`
8. `income_data |> stats::lm(happiness ~ income_doublings, data = _) |>
summary()`
9. `income_data |> stats::lm(happiness ~ income_doublings, data = ...)`

Student responses

1. `income_data |> stats::lm(happiness ~ income_doublings, data = income_data, data = _)`
2. `income_data |> stats::lm(happiness ~ income_doublings, data = income_data)`
3. `income_data |> stats::lm(happiness ~ income_doublings, data = income_data)`
4. `income_data |> stats::lm(happiness ~ income_doublings, data = income_data)`
5. `income_data |> stats::lm(happiness ~ income_doublings, dataset = income_data, data = _)`
6. `income_data |> stats::lm(happiness ~ income_doublings)`
7. `income_data |> stats::lm(happiness~income_doublings, data = _)`
8. `income_fit <- lm( happiness ~ income_doublings, data = income_data )`
`summary(income_fit)`
9. `income_fit <- lm(happiness ~ income_doublings, data = income_data)`

Student responses

1. `income_fit_log <- lm( happiness ~ income_doublings, data = income_data )`
`summary(income_fit_log)`
2. `income_fit_log <- lm( happiness ~ income_doublings, data = income_data )`
`summary(income_fit_log)`
3. `income_fit2 <- lm( happiness ~ income_doublings, data = income_data )`
`summary(income_fit2)`
4. `lm( happiness ~ income_doublings, data = income_data )`
5. `lm(happiness ~ income_doublings, data = income_data)`
6. `lm(happiness~income_doublings, data = income_data)`
7. `model <- lm(happiness ~ income_doublings, data = income_data) coef(model)`
`["income_thousands"] summary(model)$r.squared`
8. `model <- lm(happiness ~ income_doublings, data = income_data);`
`summary(model)`

Student responses

1. `summary(lm(happiness ~ income_doublings, data = income_data))`

According to this line, how much higher is the predicted life rating when income per person doubles? Explain using the fitted coefficient.

67 anonymous responses

1. About 0.55 with more precision13

Respondents state the coefficient value more precisely (around 0.5476 or 0.548) and connect it to the increase per doubling.

2. Around 0.54 increase12

Respondents interpret the fitted coefficient as implying about a 0.54 increase in predicted life rating when income doubles.

3. Around 0.55 increase9

Respondent interprets the fitted coefficient as implying about a 0.55 increase in predicted life rating when income doubles.

4. Precise 0.5476 statement8

Explicit single-sentence statements asserting the predicted increase equals 0.5476 when income doubles.

Response themes for question 13

1. About 1 or ~5 (overestimates)7

Responses that give rounded or larger estimates (around 1 or about 5), indicating overestimation relative to the fitted coefficient.

2. Other responses7

Responses the model could not place reliably.

3. Interprets as 54 increase (scale confusion)6

Respondent gives 54, indicating a likely confusion about decimal placement or coefficient scaling rather than the small fractional increase.

4. Negative five (sign confusion)5

Respondent states a negative five, indicating confusion about the sign or magnitude of the coefficient.

Student responses

1. 0.5476
2. 0.5476111
3. 0.5476111 higher every time the gdp doubles
4. 0.5476111, which means the rating is about 0.54 higher when income doubles
5. fitted coefficient for income_doublings is 0.5476, meaning each doubling of income is associated with a 0.55-point increase in predicted happiness
6. It increases by an average of 0.548 points.
7. Since the coefficient on income_doublings is 0.548, each doubling of income per person raises the predicted life rating by about 0.55 points
8. The predicted life rate is 0.548 higher when income per person doubles.
9. The predicted life rating is 0.55 points higher (or approximately 0.548 points

Student responses

1. When GDP per capita doubles, the predicted life satisfaction rating increases by approximately 0.548 points.
2. When income doubles, predicted life rating increases by 0.548
3. When income per person doubles, the predicted life rating increases by approximately 0.548 points. This is because the fitted coefficient for income_doublings is 0.5476111, meaning each additional doubling of income is associated with a 0.548-point increase in predicted happiness.
4. When income per person doubles, the predicted life satisfaction rating increases by about 0.55 points, because the slope for income_doublings is 0.5476.

QUESTION 13 · Around 0.54 increase · 1/2

Student responses

1. .54
2. .54
3. .54 higher
4. 0.54
5. 0.54
6. 0.54
7. It is higher by .54 points when income doubles.
8. Life rating is .54 higher when income doubles.
9. The predicted life rating increases by 0.54

QUESTION 13 · Around 0.54 increase · 2/2

Student responses

1. When income doubles, predicted life is higher by 0.54 years.
2. When income doubles, the predicted life rating is 0.54 points higher
3. When income per person doubles, the predicted life rating increases by 0.54

Student responses

1. .5
2. .5 points more
3. .55
4. 0.5 points higher.
5. 0.55
6. Each time income doubles, life ratings increases by about 0.55 points
7. income per person doubles, the model predicts life rating to be about 0.55 points higher
8. it goes up by 0.55 every time income doubles per person explained by out model
9. the predicted life rating increases by 0.55 when income doubles.

Student responses

1. 0.5476 values of happiness
2. 0.5476111 happiness units (slope of the fit)
3. Increases by 0.547 every time income doubles.
4. It would be .547, that's the coefficient for income_doublings
5. The life rating is predicted to be 0.5476 higher when the income per person doubles.
6. The predicted life rating is higher by 0.5476.
7. when income per person doubles the predicted life rating increases by about .547
8. When income per person doubles, the predicted life rating goes up by 0.5476 when compared to the average life rating

Student responses

1. ~5
2. 0.54. it is this much higher because of the coefficient
3. 1.8
4. $5.26 + 0.54$
5. about 0.54 pts, meaning a double in income leads to 1/2 pt increase in happiness
6. Basically you increase by half a life point
7. When income per person doubles, predicted life rating increases by about 5.27

Student responses

1. 0.548
2. $0.55 * \text{income_doublings} + 5.27$, every 0.55 point rise in life rating would clarify the prediction better.
3. About .54 increase
4. About 0.55 points higher on the 0 to 10 scale. The fitted line is $\text{happiness} = 5.267 + 0.5476 \times \text{income_doublings}$. The predictor `income_doublings` counts how many times income has doubled from \$10,000, so increasing it by 1 means income per person doubles. The slope is the change in predicted happiness per one-unit increase in the predictor, so each doubling of income raises the predicted life rating by the slope, 0.5476, roughly half a point.
5. about half as high
6. less than 1 point
7. When the income per person doubles, the predicted life rating increases by

Student responses

1. .5
2. 0.548
3. much higher
4. The predicted life rating increases by about 0.55 points when income per person doubles, because the fitted coefficient for `income\_doublings` is 0.5476.
5. The slope coefficient for income_doublings is 0.5476111. Because the predictor variable income_doublings tracks the number of times income doubles. A 1-unit increase in income_doublings represents a doubling of GDP per person
6. When GDP per person doubles, the predicted life rating increases by about 0.55 points. This is because the slope coefficient for income_doublings is 0.5476.

Student responses

1. -5

2. 0.547

3. 0.5476

4. it increases by 0.547, or half a point, via the coef

5. Predicted life rating (or happiness?) is a half-year longer when income doubles. ^ ♥ /|、 ㄥ - ´) (` - 7 / - , | | ? - ? \ (_ , J J じ し _ ,) ノ

- Calculate absolute_resid as abs(resid). - Group by model. - Calculate mean_absolute_residual as the mean of absolute_resid for each model.

67 anonymous responses

1. Other responses67

Student responses

1. `library(dplyr) model_errors %>% mutate(absolute_resid = abs(resid)) %>%
group_by(model) %>% summarize(mean_absolute_residual =
mean(absolute_resid))`
2. `model_errors library(dplyr) model_errors |> mutate(absolute_resid =
abs(resid)) |> group_by(model) |> summarise(mean_absolute_residual =
mean(absolute_resid)) |> arrange(mean_absolute_residual)`
3. `model_errors model_errors %>% mutate(absolute_resid = abs(resid)) %>%
group_by(model) %>% summarize(mean_absolute_residual =
mean(absolute_resid))`
4. `model_errors model_errors |> mutate(absolute_resid = abs(resid)) |>
group_by(model) |> summarize(mean_absolute_residual =
mean(absolute_resid))`
5. `model_errors model_errors |> mutate(absolute_resid = abs(resid)) |>
group_by(model) |> summarize(mean_absolute_residual =`

Student responses

1. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::mutate(mean_absolute_residual =
 mean(abs_resid))`
2. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::mutate(mean_absolute_residual =
 mean(abs_resid))`
3. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid, na.rm = TRUE))`
4. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid, na.rm = TRUE))`
5. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =`

Student responses

1. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid))`
2. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid))`
3. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid))`
4. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =
 mean(abs_resid))`
5. `model_errors |> dplyr::mutate(abs_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_abs_resid =`

Student responses

1. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::mutate(mean_absolute_resid = mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::mutate(mean_absolute_resid = mean(absolute_resid))`
3. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::mutate(mean_absolute_residual = mean(absolute_resid))`
4. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::mutate(mean_absolute_residual = mean(absolute_resid))`
5. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::mutate(mean_absolute_residual =`

Student responses

1. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_resid =
 mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
3. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
4. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
5. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =`

Student responses

1. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
3. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
4. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
5. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =`

Student responses

1. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
3. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
4. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
5. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =`

Student responses

1. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
3. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
4. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_residual =
 mean(absolute_resid))`
5. `model_errors |> dplyr::mutate(absolute_resid=abs(resid)) |>
 dplyr::group_by(model) |> dplyr::summarize(mean_absolute_resid =`

Student responses

1. `model_errors |> mutate( absolute_resid = abs(resid) ) |> group_by(model) |> summarize( mean_absolute_residual = mean(absolute_resid) )`
2. `model_errors |> mutate( absolute_resid = abs(resid) ) |> group_by(model) |> summarize( mean_absolute_residual = mean(absolute_resid) )`
3. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarise( mean_absolute_residual = mean(absolute_resid) )`
4. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarise(mean_absolute_residual = mean(absolute_resid))`
5. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarize(mean_absolute_residual = mean(absolute_resid))`
6. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarize(mean_absolute_residual = mean(absolute_resid))`
7. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |>`

Student responses

1. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarize(mean_absolute_residual = mean(absolute_resid))`
2. `model_errors |> dplyr::mutate(absolute_resid = abs(resid)) |> dplyr::group_by(model) |> dplyr::summarise( mean_absolute_residual = mean(absolute_resid) )`
3. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarize(mean_absolute_residual = mean(absolute_resid))`
4. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarise(mean_absolute_residual = mean(absolute_resid))`
5. `model_errors |> mutate(absolute_resid = abs(resid)) |> group_by(model) |> summarize(mean_absolute_residual = mean(absolute_resid))`
6. `model_errors["absolute_resid"] = model_errors["resid"].abs() # 2 & 3. Group by model and calculate the mean of absolute_resid summary = (`

CLASS RECORD · QUESTION 15

- **Put the countries with the largest residuals first.**
 - **Keep the first five countries. - Display country, happiness, gdp_per_capita, and resid.**
-

64 anonymous responses

1. Other responses64

Student responses

1. happiness_residuals
2. happiness_residuals happiness_residuals %>% arrange(desc(resid)) %>% slice_head(n = 5) %>% select(country, happiness, gdp_per_capita, resid)
3. happiness_residuals happiness_residuals |> arrange(desc(resid)) |> slice_head(n = 5) |> select(country, happiness, gdp_per_capita, resid)
4. happiness_residuals library(dplyr) happiness_residuals |> arrange(desc(resid)) |> slice(1:5) |> select(country, happiness, gdp_per_capita, resid)
5. happiness_residuals %>% arrange(desc(resid)) %>% slice_head(n = 5) %>% select(country, happiness, gdp_per_capita, resid)
6. happiness_residuals %>% arrange(desc(resid)) %>% slice_head(n = 5) %>% select(country, happiness, gdp_per_capita, resid)
7. happiness_residuals |> arrange(desc(resid)) |> slice_head(n = 5) |>

Student responses

1. `happiness_residuals |> arrange(desc(resid)) |> slice_head(n = 5) |> select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> arrange(desc(resid)) |> slice_head(n = 5) |> select(country, happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> arrange(desc(resid)) |> slice(1:5) |> select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> arrange(desc(resid)) |> slice(1:5) |> select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> arrange(desc(resid)) |> slice(1:5) |> select(country, happiness, gdp_per_capita, resid)`
6. `happiness_residuals |> arrange(desc(resid)) |> slice(1:5) |> select(country, happiness, gdp_per_capita, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid))`

Student responses

1. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5)`
2. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5)`
4. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>
dplyr::select(country, happiness, gdp_per_capita, gdp_year, life_years, rank,
life_lower, life_upper, iso3, wb_country, income_thousands,
income_doublings, resid)`
5. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>
dplyr::select(country, happiness, gdp_per_capita, gdp_year, resid)`
6. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>
dplyr::select(country, happiness, gdp_per_capita, gdp_year, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>
dplyr::select(country, happiness, gdp_per_capita, resid)`

Student responses

1. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
6. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>`

Student responses

1. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
6. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>`

Student responses

1. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
6. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |>`

Student responses

1. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> dplyr::arrange(desc(resid)) |> head(5) |> select(country, happiness, gdp_per_capita, resid)`
6. `happiness_residuals |> dplyr::arrange(desc(resid)) |> head(5) |> select(country, happiness, gdp_per_capita, resid)`
7. `happiness_residuals |> dplyr::arrange(desc(resid)) |> head(5)|>`

Student responses

1. `happiness_residuals |> dplyr::arrange(dplyr::desc(resid)) |> dplyr::slice(1:5)
|> dplyr::select(country, happiness, gdp_per_capita, resid)`
2. `happiness_residuals |> dplyr::mutate(abs_resid = abs(resid)) |>
dplyr::arrange(desc(resid)) |> dplyr::slice(1:5) |> dplyr::select(country,
happiness, gdp_per_capita, resid)`
3. `happiness_residuals |> dplyr::arrange(desc(resid)) |> head(n = 5) |>
dplyr::select(country, happiness, gdp_per_capita, resid)`
4. `happiness_residuals |> arrange(desc(resid)) |> slice_head(n = 5) |>
select(country, happiness, gdp_per_capita, resid)`
5. `happiness_residuals |> arrange(desc(resid)) |> slice_head(n=5) |>
select(country, happiness, gdp_per_capita, resid)`
6. `top_five_above = happiness_residuals.sort_values(by="resid",
ascending=False).head(5)[["country", "happiness", "gdp_per_capita", "resid"]]`

Choose a country whose life rating surprised you. What additional data would you want to explain its position? Describe one analysis you would try.

70 anonymous responses

1. Named country plus demographic note14

Student names Costa Rica and mentions population size as a relevant factor to investigate.

2. Individual income and happiness link10

Student suggests analyzing individual-level earnings alongside personal happiness scores to explain life ratings.

3. Named country: China10

Student names China as the surprising country with no additional details or analyses offered.

4. Named country: Singapore9

Student simply names Singapore as the surprising country without further data or analysis suggestions.

Response themes for question 16

1. Requests data on work, housing, health, poverty⁸

Students ask for specific socioeconomic and health variables—hours worked, housing, healthcare access, life expectancy, poverty rates—to explain life ratings.

2. Suggests crime or healthcare data⁷

Student proposes examining crime rates or measures of healthcare quality to explain a country's life rating.

3. Costa Rica with desire for experiments or demographics⁶

Student names Costa Rica and requests controlled studies or mentions demographic/education/employment variables as relevant.

4. Unclear or no content⁶

Responses that are empty, punctuation-only, or too vague to infer a viewpoint or suggested data/analysis.

Student responses

1. costa rica
2. costa rica and population size
3. Costa Rica surprised me. I think we would need something like climate to explain more of this data.
4. Costa Rica surprises me; I'd like to check cost of living to see if it anti-correlates
5. Costa Rica. I'd want to look at the cost of living and the cost of healthcare, because income isn't the only factor that can explain the ratings. I would construct another linear regression model.
6. I am surprised by Costa Rica, as it is a very poor country. I would be interested in looking at its quality of life or other similar stats.
7. I think Costa Rica, I can't imagine this, since it's located in Latin America
8. I want to look at the disparity between the wealthy and those who aren't

Student responses

1. I would want to look at how peaceful that country is maybe. Not really shocked by any of them.
2. I'm surprised by how high costa rica is because it is such a small island, and I would think there would be a lower quality of life
3. Mozambique, I'd want to look at other characteristics included in the table like number of pets.
4. Other data like crime and housing rates, as well as cost of living would be important. Costa Rica was by far the biggest surprise out of the countres for me
5. Qualityb of living could be another data piece we could use, analyzing those residuals based on another factor

Student responses

1. Costa Rica Surprised me, I would want to see data about government spending/ public works, I would analyze approval of government spending
2. Costa Rica surprised me. I would run a more detailed linear relationship on jsut Costa Rica
3. Costa Rica surprised me. I would want to examine life expectancy, political freedom.
4. finland I wanna try it's population policy
5. I did not know that Costa Rica would have such a high life rating.
6. I think the US surprised me I would want to look into what its like for different incomes
7. I was surprised by Costa Rica. It may be interesting to examine climate data, and rank the countries by a set temperature or precipitation range.
8. I would look into the rights that citizens have in each country to see if more

Student responses

1. Singapore is quite surprising due to its higher GDP with lower happiness. We can do a multiregression test, adding more variables social support freedom health / life expectancy corruption perception work-life balance inequality

Student responses

1. china
2. Costa Rica surprised me because its life rating is 7.44 even though GDP per person is only about \$27,000, well above what the line predicts. I would want data on social support, healthy life expectancy, and perceived freedom or corruption, which the World Happiness Report also measures. I would fit a multiple regression adding social support to income doublings and check whether Costa Rica's residual gets smaller and whether the model's R-squared improves.
3. Costa Rica surprised me because its life satisfaction is very high even though its GDP per person is much lower than many other highly ranked countries. I would look at factors like social support, health, life expectancy, and work-life balance, and compare those measures with countries that have similar GDP levels to see what might explain Costa Rica's higher happiness.
4. finland. I would also like to see average days of sunlight

Student responses

1. I would investigate Costa Rica. I would collect data on social support, healthcare, and life expectancy. Then, I would use multiple linear regression to examine whether these factors help explain Costa Rica's higher-than-predicted happiness.
2. living cost, competition pressure
3. Mozambique, id want to know why it has the highest residual. maybe data on the level of healthcare or education and run a linear regression on it and the happiness
4. osta Rica surprised me: its life rating sits well above what its GDP per person predicts (a large positive residual), so income alone doesn't explain why it's so happy. Additional data I'd want: measures that capture quality of life beyond raw output. Social support and trust (do people have someone to count on), healthy life expectancy, freedom to make life choices, inequality (a Gini coefficient, since average GDP hides how income is spread), and generosity

Student responses

1. United States supresied me the most beause it was right on the line and in the middle. Some more data that would be nie is other facotrs that play a role in deciding the happiness becasue our variable was already simply defined.

Student responses

1. I was surprised by Costa Rica as you normally only hear about the Nordic countrys being very happy and not much in the west as much. I would want to see what the quality of life actually looks like to see how many hours people work a week, what their healthcare system is like, and the amount of income inequality they experience.
2. I would investigate categorical variables such as which factors are causing the most unhappiness like cost of living, healthcare, politics, etc. and then perform and independence test to see whether these variables are independent or not
3. other things like health, freedom of speech, etc
4. Singapore surprised me the most. I would look into cost of living and CPI in that country to investigate people's buying power, maybe marriage/divorce rates, and level of education. I would do this by plotting points on a boxplot to seperate effects.

Student responses

1. Singapore. I really like its strict laws to protect overall safety, but maybe the area is too limited and the climate is too hot, and people too busy so the happiness level is not as high as I expected, I think we can add a working load analysis
2. Singapore. Maybe working time per week. (High gdp could due to long working time, thus making happiness less). I will fit a new model using all variables
3. singapore, it was lower than expected; one analysis would be anova

Student responses

1. Costa Rica surprised me because it has a high happiness score even though its GDP per person is much lower than many wealthy countries. I would want data on social support, healthcare, and work-life balance to understand why. I would try a regression analysis that includes these factors along with GDP per person to see whether they help explain Costa Rica's high happiness rating.
2. Costa Rica. Because its happiness rating is surprisingly high compared to its GDP per person. I would like to look at data on social support, healthcare, and work-life balance. I would try adding these factors to the regression model to see if they help explain why Costa Rica has such a high happiness rating.
3. Could investigate population statistics, employment, health, etc.
4. I think that Costa Rica was quite surprising. I think I'd want to use life quality factors like crime rate and social pressure to investigate them.

Student responses

1. it surprised me that kosovo had such a large residual, I would want to look at life expectancy or cost of living, and regress happiness on income, gdp, cost of living, and life expectancy to see if they explain more of the variance
2. other current events happening within the country

Student responses

1. data on crime and safety in that country
2. I want to analyze the correlation between education level or intensity with happiness.
3. I would collect data on healthcare. While the GDP of a country is telling of its healthcare, it's not always the most accurate. In some cases, countries could be investing more into healthcare than expected of their GDP, so the people might be happier despite having less income because the health system will still take care of them.
4. I would look at crime rates or healthcare quality, using similar analysis like scatterplots and regression on a country with a high residual in our current analysis, like costa rica.
5. I would want to investigate data related to housing and healthcare.
6. Maybe try the meat demand or "Big mac" price score

Student responses

1. Costa Rica, I would investigate healthcare and life expectancy and make another linear model with more predictors
2. costa rica, i would want more controlled experiments man idk
3. Employment and Education Data, average amount of kids , healthcare, ertc. I am interested in costa rica. I would test on scatterplots and lms for associations between these and happiness.
4. I was mostly surprised with Costa Rica, as its economy I would say is more dependent on things such as tourism. I would thus want to filter out certain economic areas.
5. I'm surprised by how high Costa Rica's life rating is given its current GDP, I would want to learn more about its healthcare system specifically to help explain its position, I would try to quantify where in the world its healthcare system lies

Student responses

1. A country like Mozambique surprised me, since I expected countries affected by colonialism to have lower happiness scores. I am curious if we added the factor of colonized by European powers, what would our plot look like.
2. Afghanistan. I would survey threat to life, feeling of safety, living condition, you might have to separate based on gender
3. Costa Rica. I would also like to investigate its geographical position, i.e. it is located in Latin America. I will try to compare countries in different regions (Asia/Europe/Africa/Australia/North America/Latin America/South America).
4. Health
5. I think more layered data would be helpful, like some indirect factor could reveal more things
6. Political satisfaction could be a useful way to analyze life rating. If most