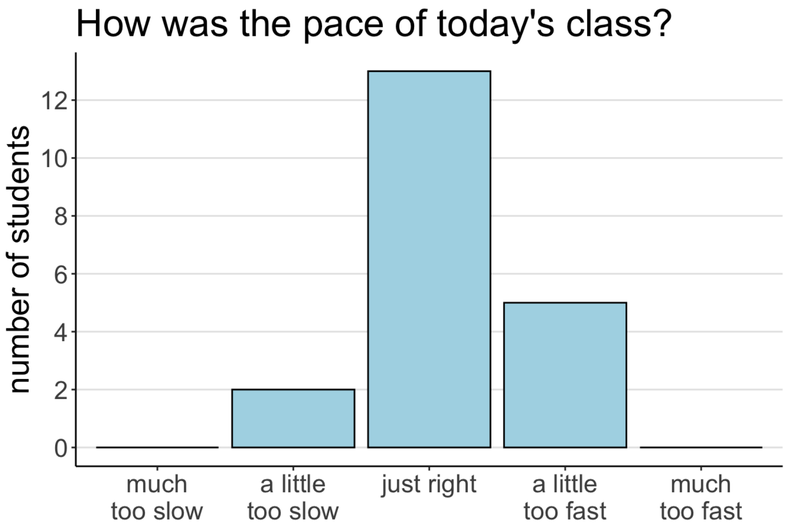
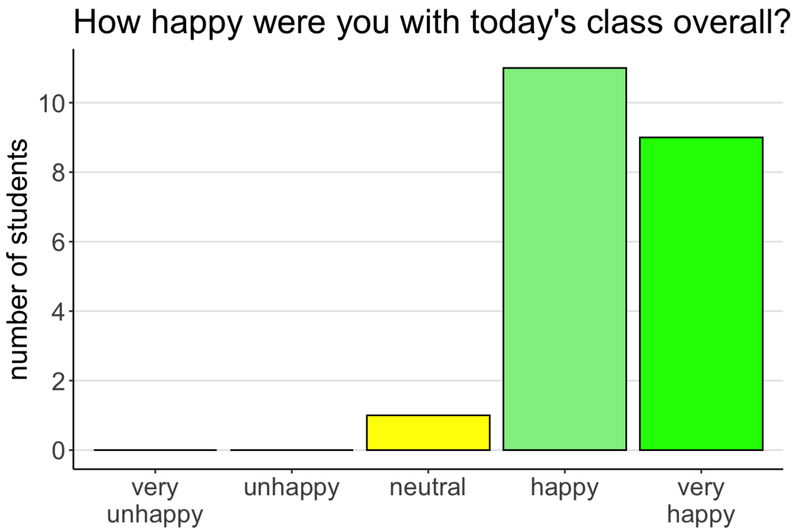
Data wrangling 2

1/15/2025

Your feedback

Your feedback





If the TAs can reposition themselves around the room so they can better see us holding blue notes up for help that would be great. Thank you!

I liked that there were more instructions before the practice

Data wrangling time ...

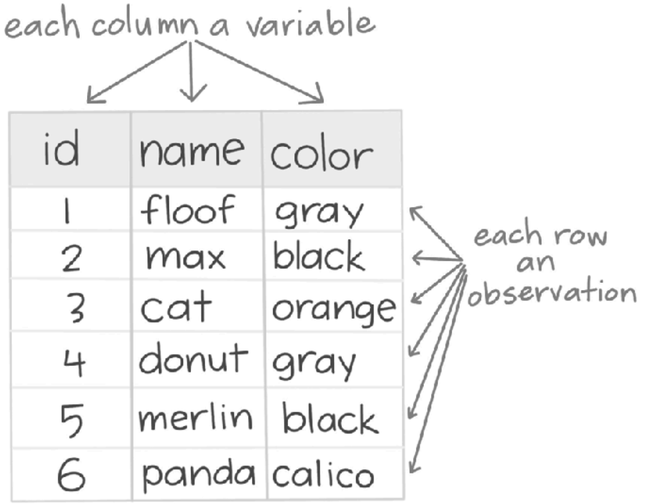
Tidy data

"TIDY DATA is a standard way of mapping the meaning of a dataset to its structure."

–HADLEY WICKHAM

In tidy data:

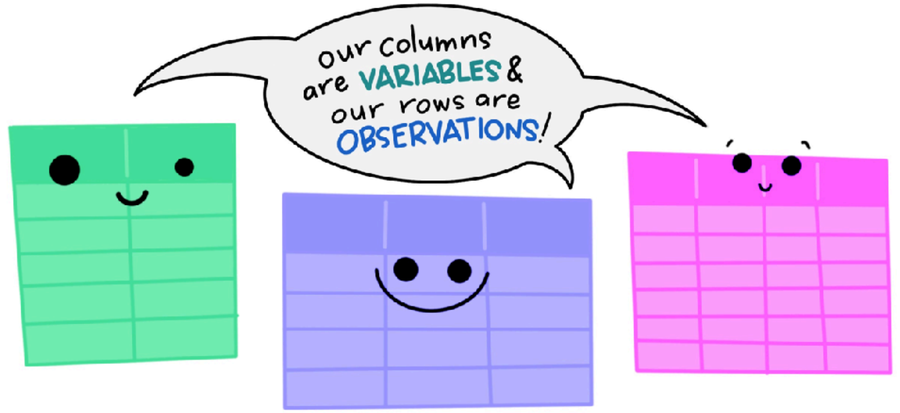
- each variable forms a column
- each observation forms a row
- each cell is a single measurement



Wickham, H. (2014). Tidy Data. Journal of Statistical Software 59 (10). DOI: 10.18637/jss.v059.i10

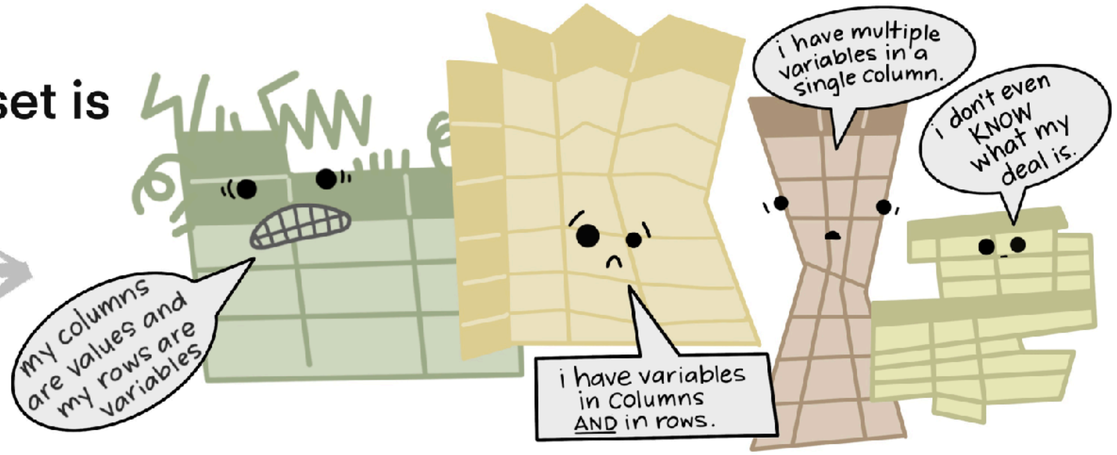
Tidy data

The standard structure of tidy data means that "tidy datasets are all alike..."



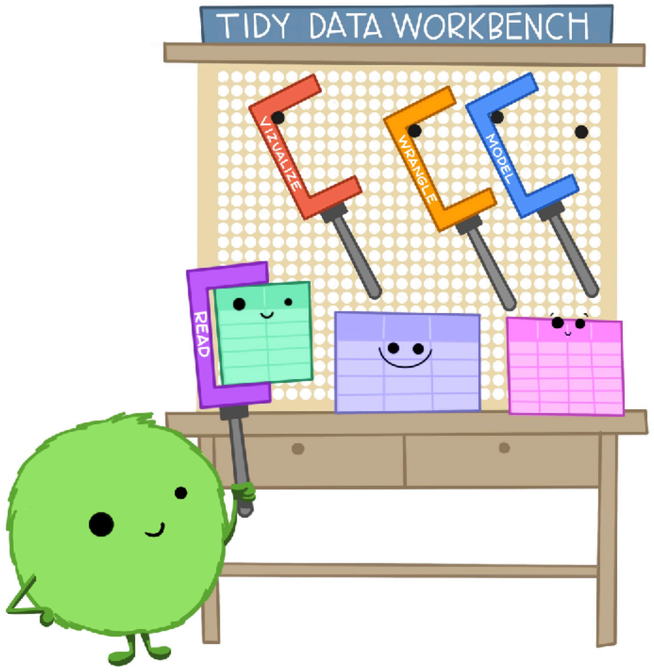
"...but every messy dataset is messy in its own way."

–HADLEY WICKHAM

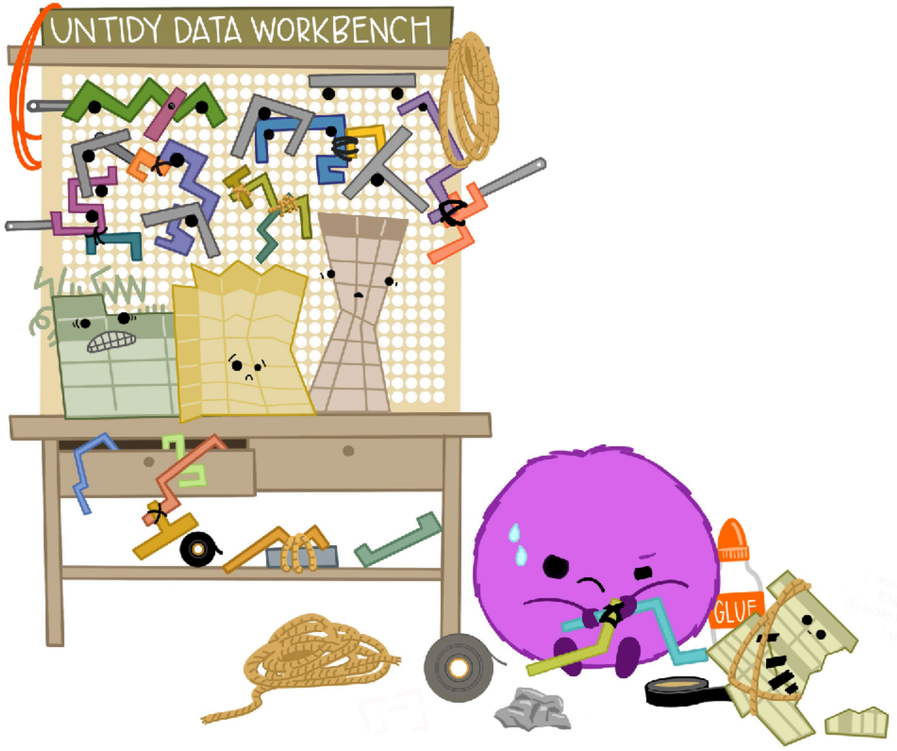


Tidy data

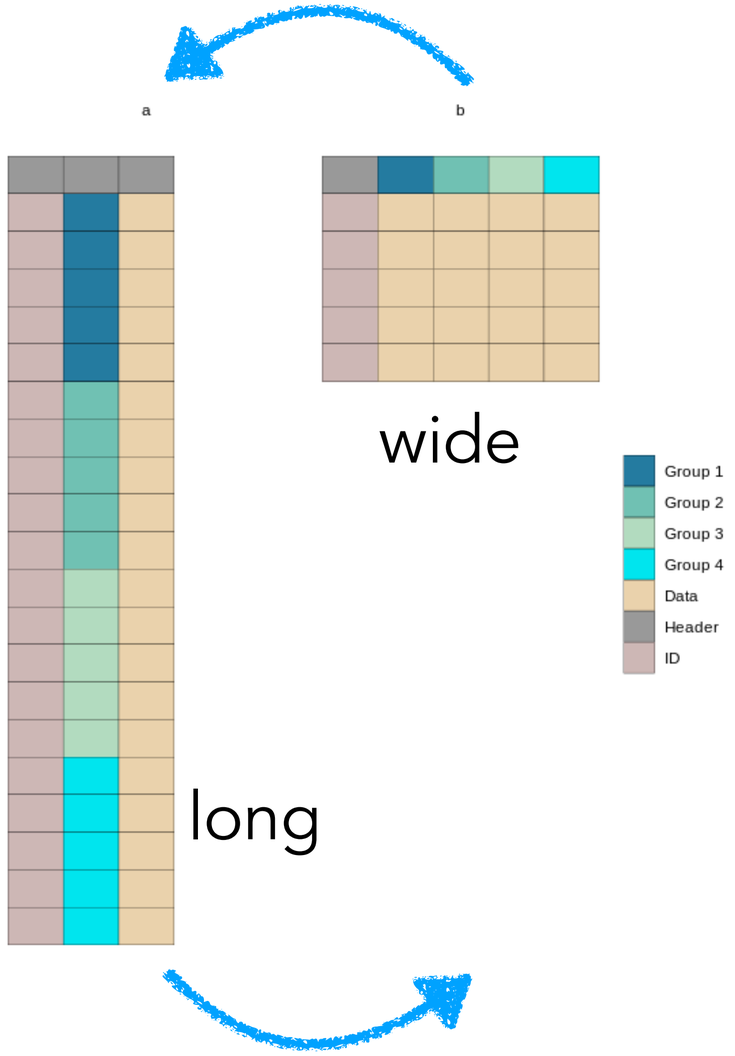
When working with tidy data, we can use the same tools in similar ways for different datasets...



...but working with untidy data often means reinventing the wheel with one-time approaches that are hard to iterate or reuse.

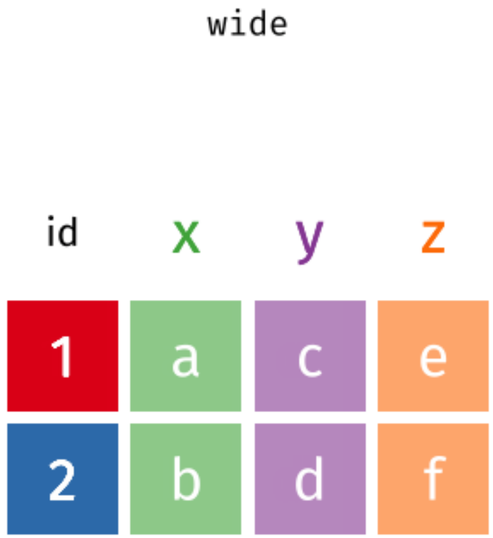


pivot_longer()



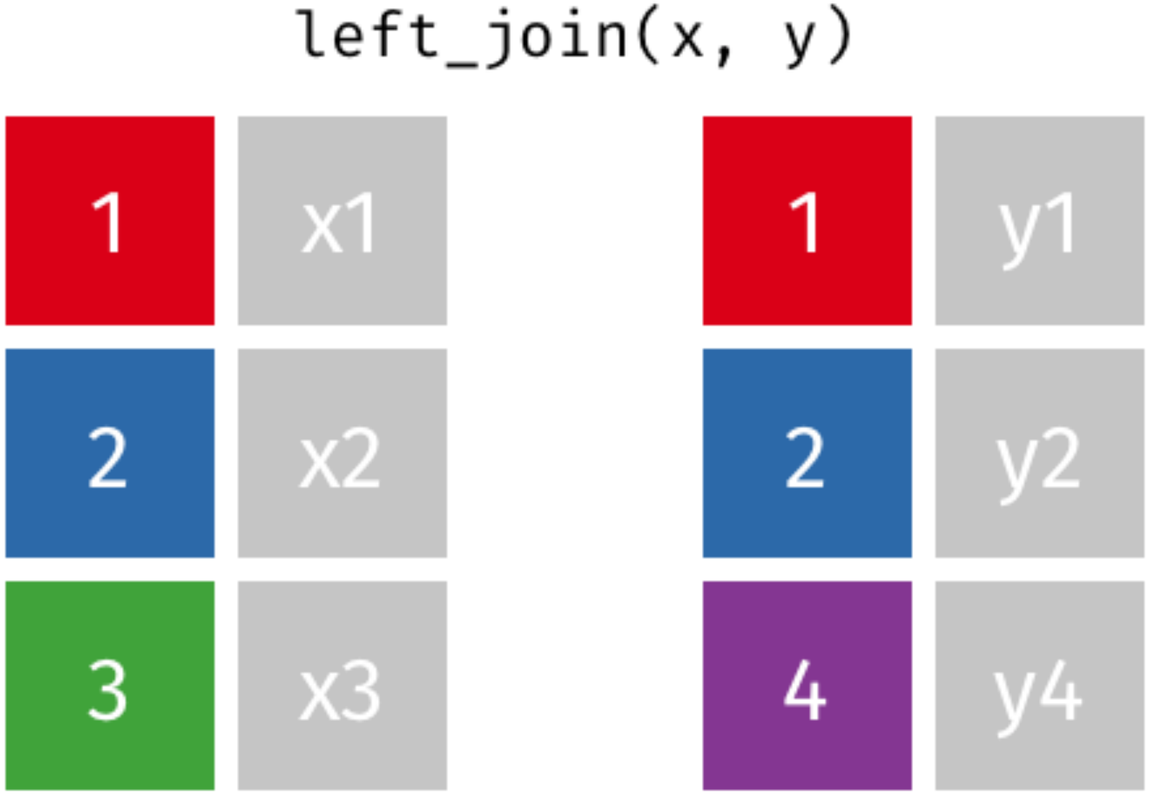
pivot_wider()

pivot_longer() & pivot_wider()



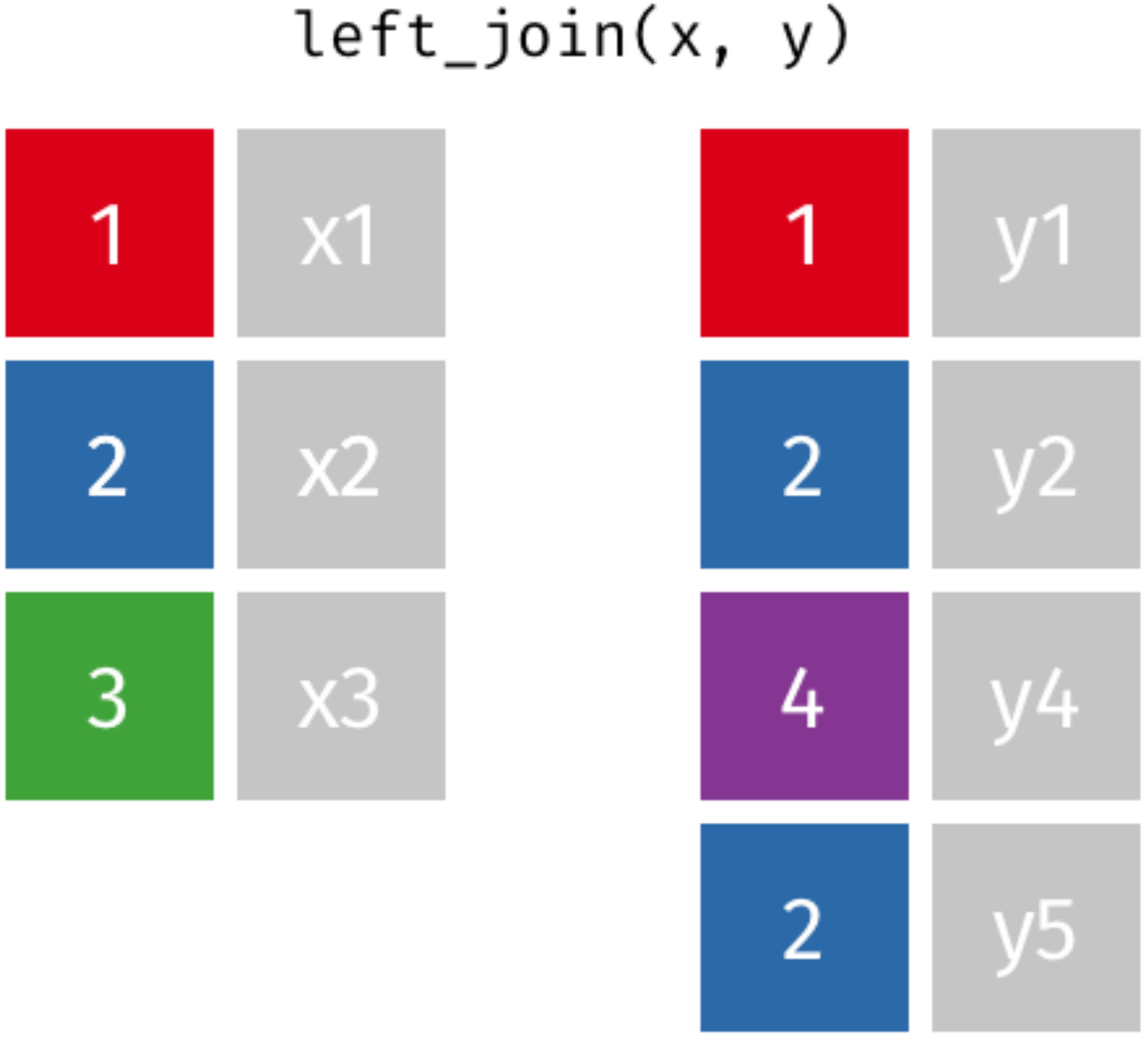
https://github.com/gadenbuie/tidyexplain

left_join()



https://github.com/gadenbuie/tidyexplain

left_join()



https://github.com/gadenbuie/tidyexplain

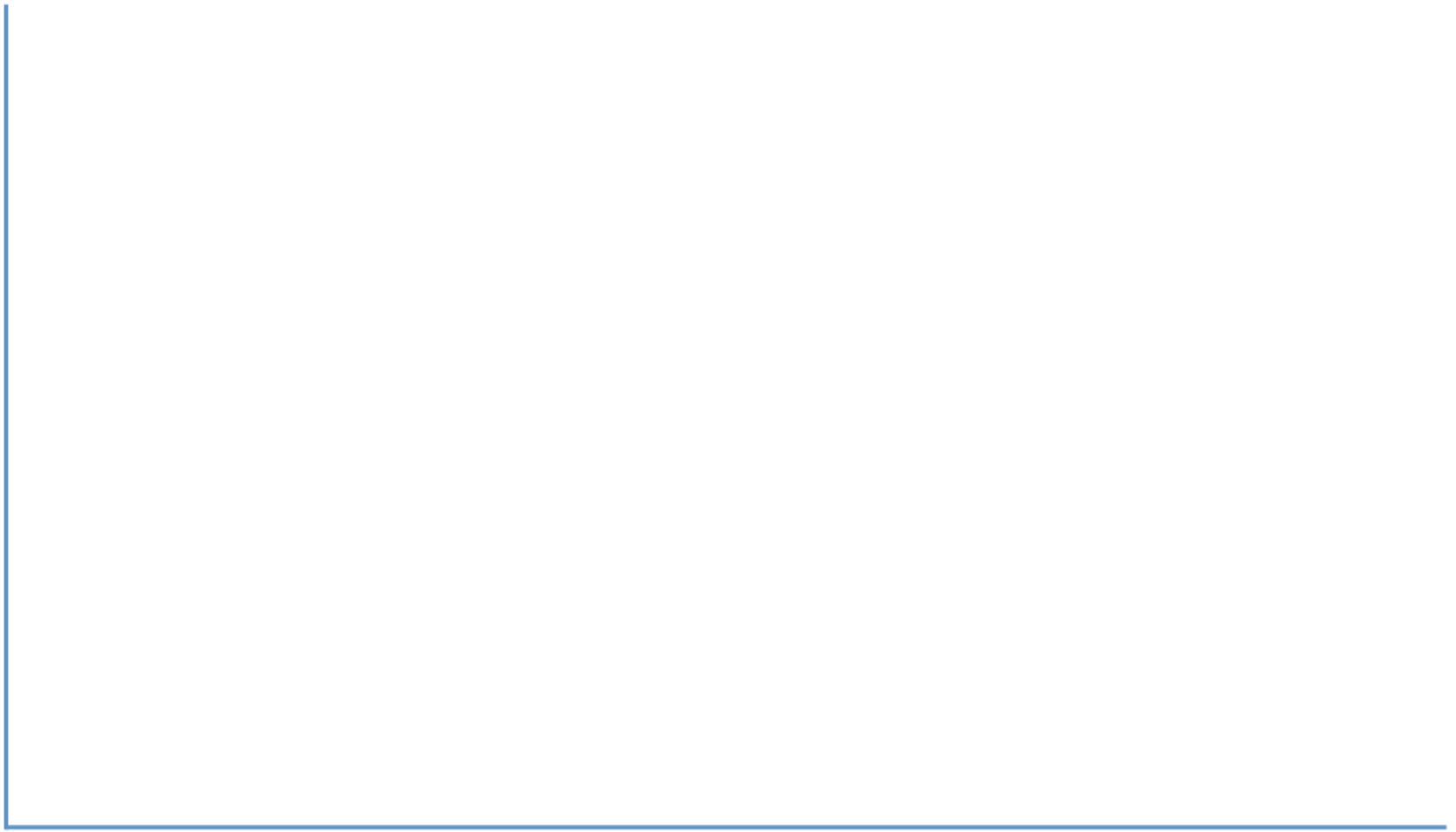
Timer

blue

pink

Feedback

How was the pace of today's class?



How happy were you with today's class overall?

What did you like about today's class? What could be improved next time?

Thank you!