

Outliers, missing pieces, and wrong numbers, oh my!

Data Quality Worksheet:

Name: _____ Date: _____

Think about it!

Imagine you're the coach of a school swim team. You're tracking your team's lap times to see how they're improving.

- Look at the data table below that shows the times (in seconds) for the 400-meter freestyle event for your team.
- Answer the questions that follow to explore how data quality can affect your analysis.

Swimmer	Time (seconds)
Alex	285
Beatrice	320
Carlos	305
Danielle	298
Ethan	422
Fiona	312

Questions:

1. **Outlier Alert!** Look at Ethan's time. Do you think it might be an outlier? Why or why not?
2. **Impact of Outliers:** If Ethan's time is an outlier, how might it affect the average lap time for the team? Calculate the average lap time with and without Ethan's time.
 - Average lap time with ALL data: _____ seconds
 - Average lap time WITHOUT Ethan's time: _____ seconds
3. **Missing the Mark:** Imagine Beatrice's time is missing because her watch malfunctioned. How would a missing time affect your analysis of the team's performance?
4. **Wrong Turns:** What if Carlos accidentally entered his time in minutes instead of seconds (so his listed time of 305 should actually be 305 seconds)? How would this mistake affect the data?

5. **Data Detective!** How can you improve the data quality for your swim team's lap times? Think about ways to avoid outliers, missing values, and incorrect data entry.

Bonus Challenge!

Think of another example from your own life where data quality might be important. Describe the situation and how you could ensure good data collection.