

## How big is your subset?

You have the following `loans` DataFrame which contains loan and credit score data for consumers, and some metadata such as their first and last names. You want to find both complete and incomplete duplicates using `.duplicated()`.

first_name	last_name	credit_score	has_loan
Justin	Saddlemeyer	600	1
Hadrien	Lacroix	450	0

Choose the **correct** usage of `.duplicated()` below:

☒ Answer the question

50XP

### Possible Answers

- ☐

`loans.duplicated()`  
Because the default method returns both complete and incomplete duplicates.

press 1
- ☐

`loans.duplicated(subset = 'first_name')`  
Because constraining the duplicate rows to the first name lets me find incomplete duplicates as well.

press 2
- ☒

`loans.duplicated(subset = ['first_name', 'last_name'], keep = False)`  
Because subsetting on consumer metadata and not discarding any duplicate returns all duplicated rows.

press 3
- ☐

`loans.duplicated(subset = ['first_name', 'last_name'], keep = 'first')`  
Because this drops all duplicates.

press 4