

## 4 Cleaning data for analysis

0%

Here, you'll dive into some of the grittier aspects of data cleaning. You'll learn about string manipulation and pattern matching to deal with unstructured data, and then explore techniques to deal with missing or duplicate data. You'll also learn the valuable skill of programmatically checking your data for consistency, which will give you confidence that your code is running correctly and that the results of your analysis are reliable!

▶ Data types	50 xp
</> Converting data types	100 xp
</> Working with numeric data	100 xp
▶ Using regular expressions to clean strings	50 xp
</> String parsing with regular expressions	100 xp
</> Extracting numerical values from strings	100 xp
</> Pattern matching	100 xp
▶ Using functions to clean data	50 xp
</> Custom functions to clean data	100 xp
</> Lambda functions	100 xp
▶ Duplicate and missing data	50 xp
</> Dropping duplicate data	100 xp
</> Filling missing data	100 xp
▶ Testing with asserts	50 xp
</> Testing your data with asserts	100 xp

HIDE CHAPTER DETAILS

Continue Chapter