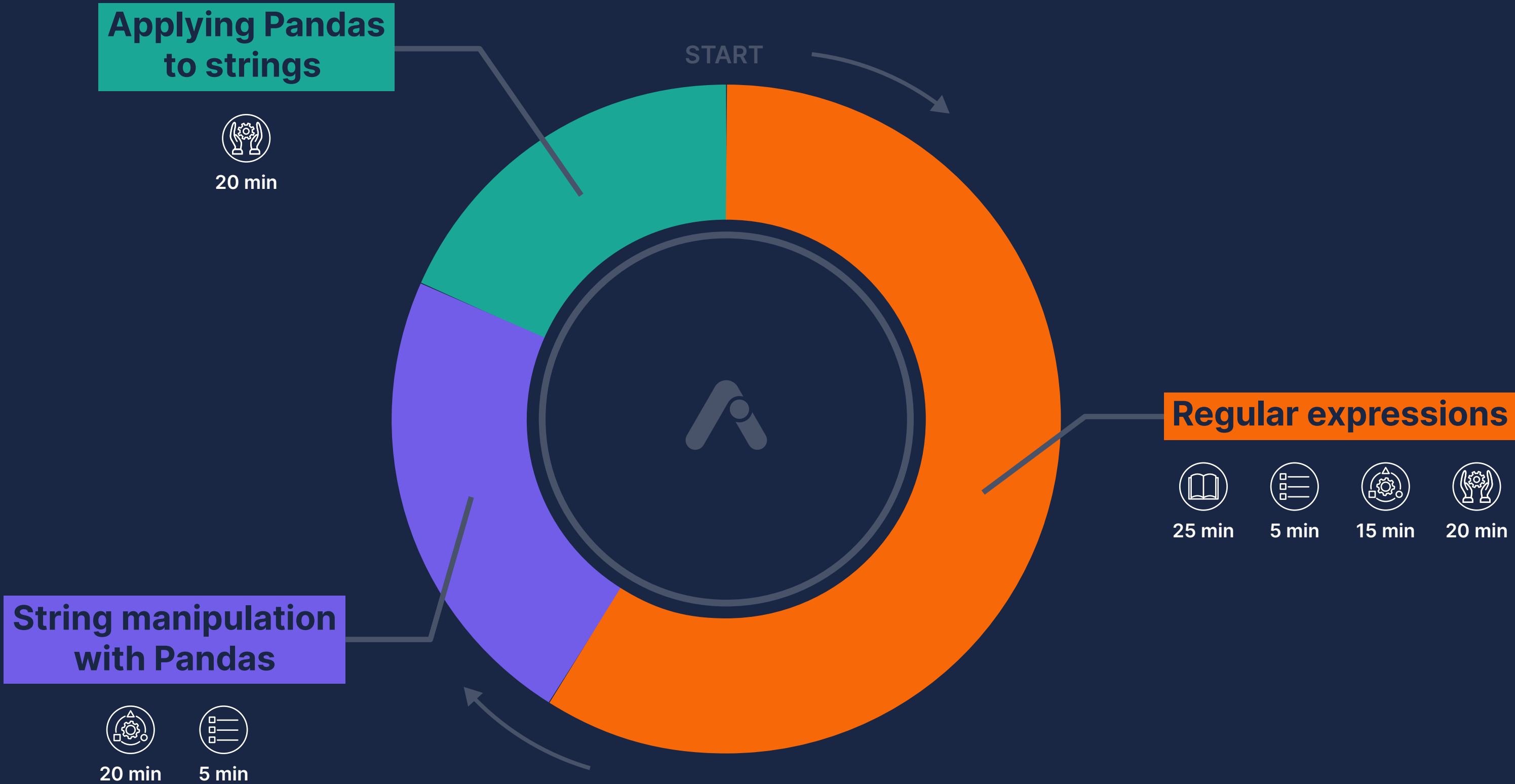# **Advanced string manipulation** – Lesson overview

Regular expressions (regex) in Python are essential tools for anyone working with **text data**, enabling efficient **searching, editing**, and **data manipulation**. Mastering regex opens up a world of possibilities for data scientists and engineers, allowing for **precise pattern matching and text processing in large datasets**.

In this lesson, we'll explore the basics of Python's regex library, focusing on practical applications such as **extracting specific information** and using **compiled regex objects** for performance. We'll gain a solid understanding of **regex patterns, methods**, and their use in **real-world data processing challenges**, enhancing our data manipulation skills in Python, especially when combined with **Pandas** for complex data tasks.

Learning objectives

- Understand the basics of Python's regex library, including the use of regex objects, character classes, and quantifiers for string pattern matching.

- Utilise various regex methods for different pattern matching and string manipulation tasks.

- Interpret and employ regex flags and basic pattern components effectively in Python.

- Create and analyse complex regex patterns to efficiently extract and manipulate data.

- Integrate regex with Pandas for advanced string operations within Python data structures.

Slide deck

Examples

Knowledge questions

Exercise

Applying Pandas to strings

20 min

START

Regular expressions

25 min   5 min   15 min   20 min

String manipulation with Pandas

20 min   5 min