

Searching and Matching Patterns in Files



Maaike van Putten

Trainer & Software Developer

@brightboost | www.brightboost.nl

Searching patterns in files



Overview



Searching and matching patterns in files

Demos:

- Search a file for a word
- Find file names in a file
- Get all the (likely) names in a file
- Search a file for a pattern



Demo



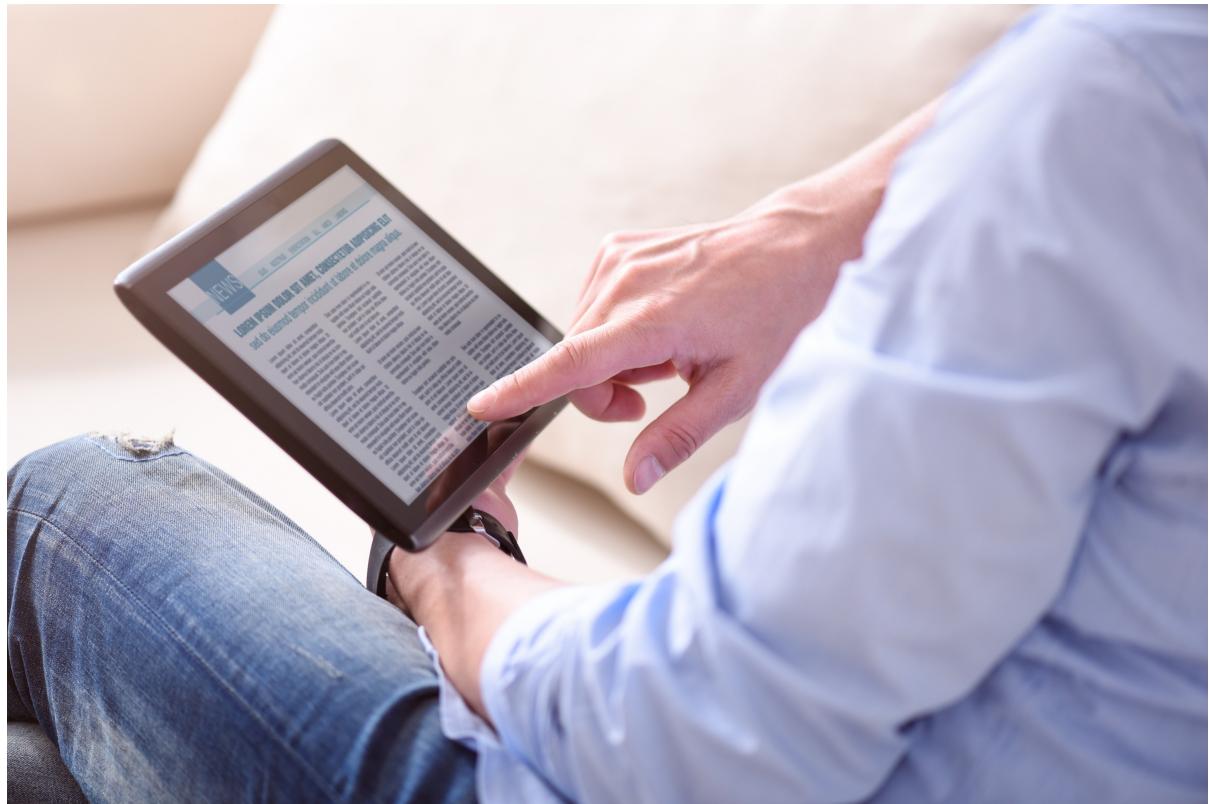
Search a file to find the first match

Steps:

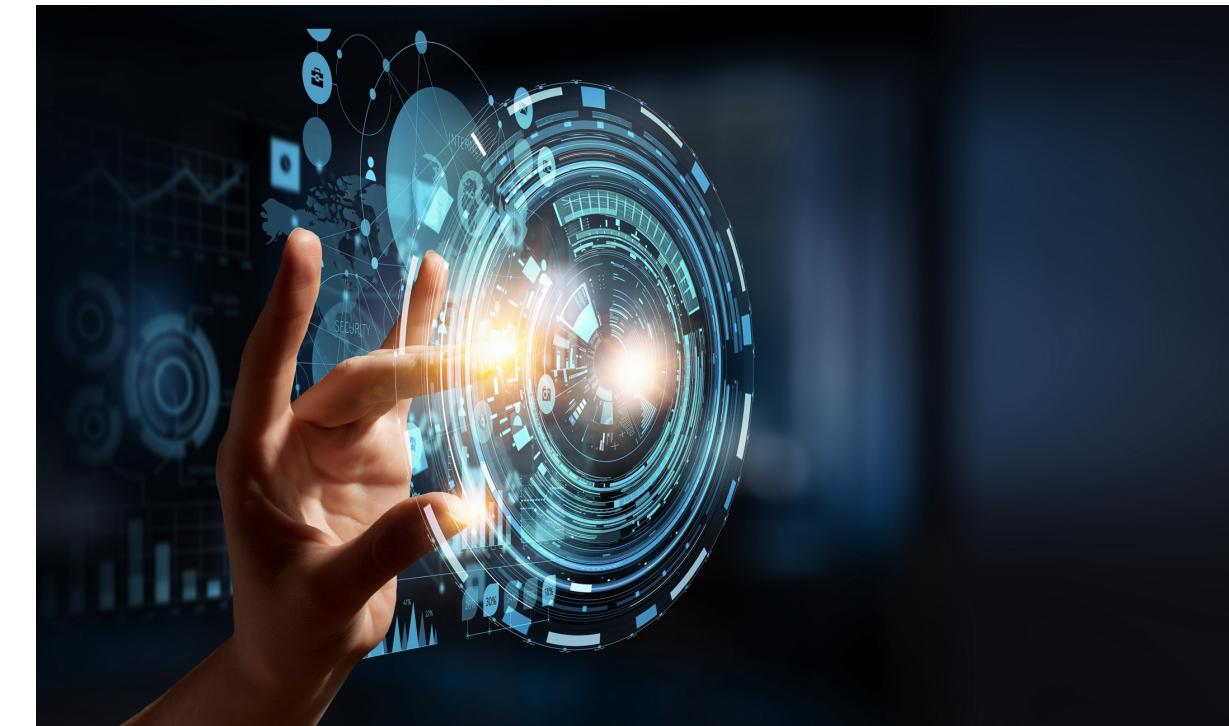
- Open a file
- Use the Python search function
- Print the outcome



Reading Files



Line by line
Processing the contents of the file
line by line as we go.



All at once
Loading the entire file in the
memory before we start.



Demo

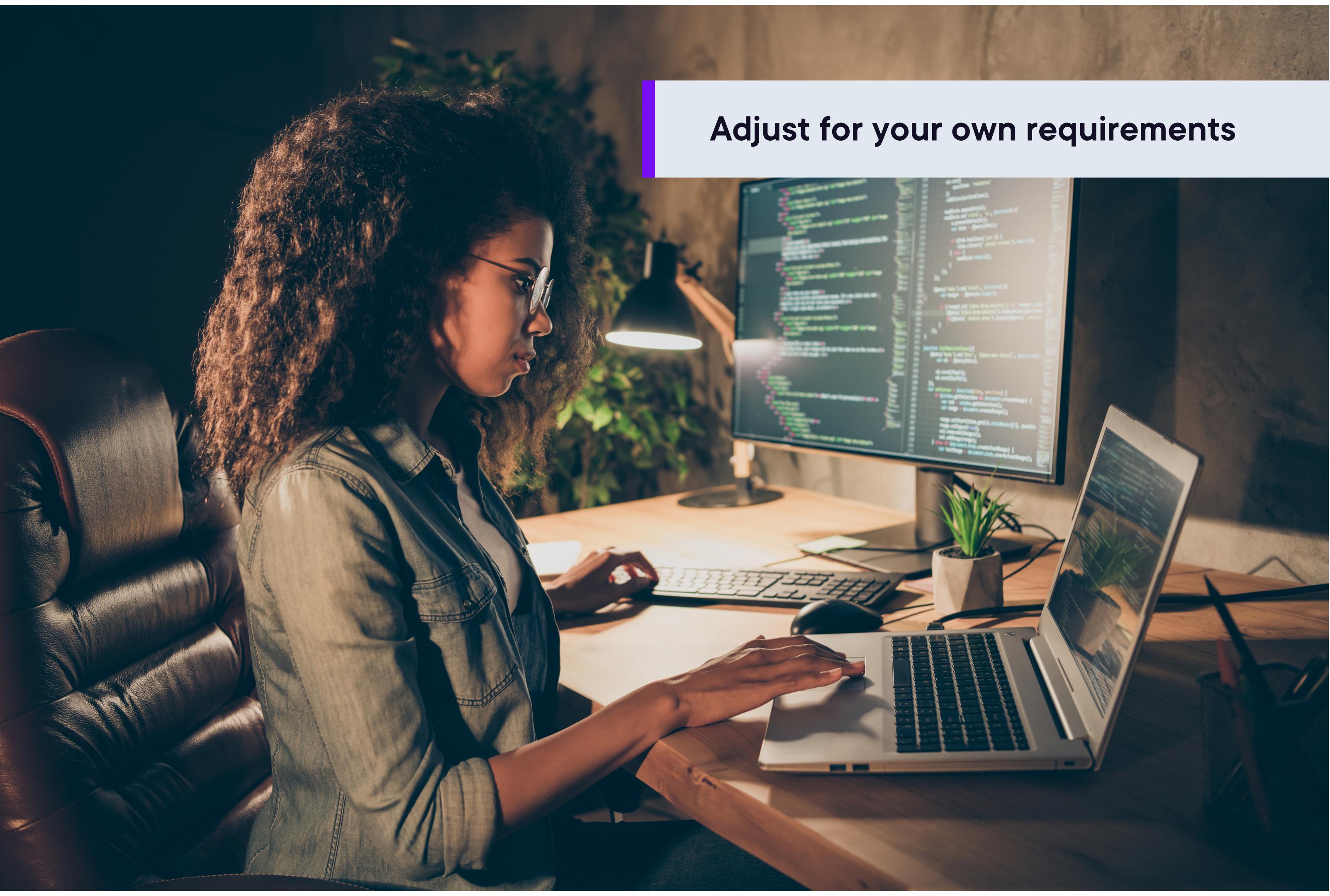


Search a file for file names

Steps:

- Create common extensions list
- Read a file
- Check for a regex pattern made using the extensions list
- Find all the matches with the `findall` function





Adjust for your own requirements





Search for specific patterns



Demo



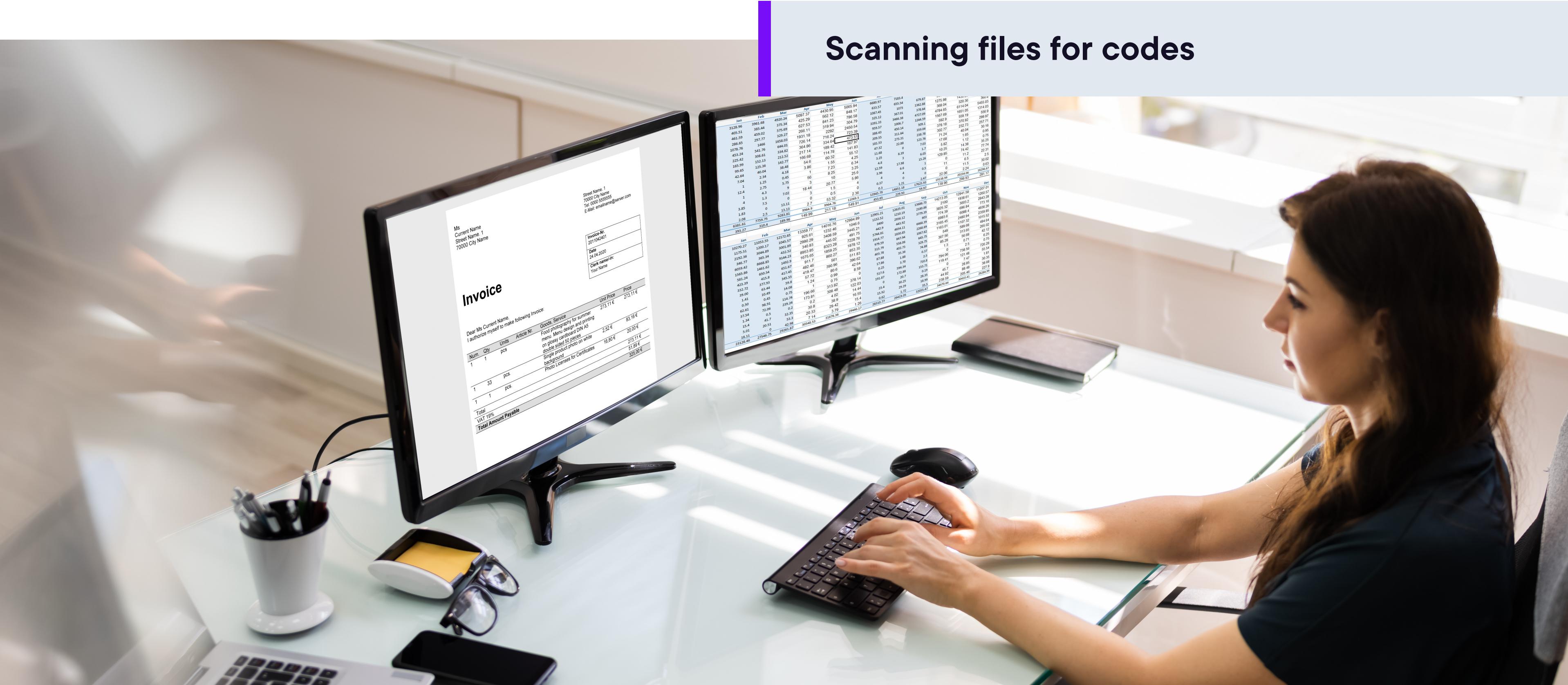
Searching and matching patterns in files

Steps:

- Define regex patterns
- Search for them in the file
- Display the result using the match objects



Scanning files for codes



Demo



Find custom product codes in a file

Steps

- Read a file
- Create the pattern
- Use the `findall` function
- Display the matches



Extract Pattern from a File

```
import re

with open("sample.txt", "r") as f:
    text = f.read()

pattern = r"\b[A-Z]{2}\d{2}-[A-Z]{2}\d{2}\b|\b\d{3}-\d{3}\b"

matches = re.findall(pattern, text)
print("All product codes:", matches)
```



Extract Pattern from a File

```
import re

with open("sample.txt", "r") as f:
    text = f.read()

pattern = r"\b[A-Z]{2}\d{2}-[A-Z]{2}\d{2}\b|\b\d{3}-\d{3}\b"

matches = re.findall(pattern, text)
print("All product codes:", matches)
```



Extract Pattern from a File

```
import re

with open("sample.txt", "r") as f:
    text = f.read()

pattern = r"\b[A-Z]{2}\d{2}-[A-Z]{2}\d{2}\b | \b\d{3}-\d{3}\b"

matches = re.findall(pattern, text)
print("All product codes:", matches)
```



Demo



Find custom product codes in multiple files

Steps

- Wrap the product code finding piece in a function and adjust it to write to a file
- Create a list of files
- Loop over the list of files and call the function

