Great snippet! Let’s walk through exactly what happens on **each iteration** of your for loop, showing the check, the action taken, what gets printed, and how seen and duplicate change.

Your code:

files=['data.csv','report.pdf', 'fg.txt','fg.txt','sf.csv']

seen=[]

duplicate=[]

for file in files:

if file in seen:

duplicate.append(file)

print('duplicate:',duplicate)

else:

seen.append(file)

print('seen files:',seen)

**Initial state (before the loop)**

* seen = []
* duplicate = []

**Iteration 1**

* **file** = 'data.csv'
* **Check**: 'data.csv' in seen → False
* **Action**: Append to seen
* **Prints**: seen files: ['data.csv']
* **After**:
  + seen = ['data.csv']
  + duplicate = []

**Iteration 2**

* **file** = 'report.pdf'
* **Check**: 'report.pdf' in seen → False
* **Action**: Append to seen
* **Prints**: seen files: ['data.csv', 'report.pdf']
* **After**:
  + seen = ['data.csv', 'report.pdf']
  + duplicate = []

**Iteration 3**

* **file** = 'fg.txt'
* **Check**: 'fg.txt' in seen → False
* **Action**: Append to seen
* **Prints**: seen files: ['data.csv', 'report.pdf', 'fg.txt']
* **After**:
  + seen = ['data.csv', 'report.pdf', 'fg.txt']
  + duplicate = []

**Iteration 4**

* **file** = 'fg.txt' (appears again)
* **Check**: 'fg.txt' in seen → True
* **Action**: Append to duplicate
* **Prints**: duplicate: ['fg.txt']
* **After**:
  + seen = ['data.csv', 'report.pdf', 'fg.txt'] (unchanged)
  + duplicate = ['fg.txt']

**Iteration 5**

* **file** = 'sf.csv'
* **Check**: 'sf.csv' in seen → False
* **Action**: Append to seen
* **Prints**: seen files: ['data.csv', 'report.pdf', 'fg.txt', 'sf.csv']
* **After**:
  + seen = ['data.csv', 'report.pdf', 'fg.txt', 'sf.csv']
  + duplicate = ['fg.txt']

**What you’ll see printed (in order)**

seen files: ['data.csv']

seen files: ['data.csv', 'report.pdf']

seen files: ['data.csv', 'report.pdf', 'fg.txt']

duplicate: ['fg.txt']

seen files: ['data.csv', 'report.pdf', 'fg.txt', 'sf.csv']

**Final state (after the loop finishes)**

* seen = ['data.csv', 'report.pdf', 'fg.txt', 'sf.csv']
* duplicate = ['fg.txt']

**Notes & tips**

* Your logic tracks **repeats** by name. If a file shows up 3 times, duplicate would end up like ['fg.txt', 'fg.txt'] (once for each extra occurrence).
* For large lists, if file in seen on a Python **list** is O(n). If you later care about speed, make seen a **set** for O(1) lookups (while keeping a separate list if you need order).