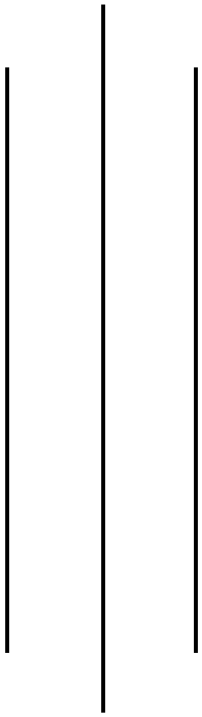
Assignment – I

Data Cleaning Using String Functions for Variables and Data Structures



Submitted To

Samriddha Pathak

Digital Pathshala Instructor

Submitted By

Rajendra Chimala

Submitted At

2025-July-5

Contents

1. Remove Space 4
2. Remove Leading Characters 4
3. Remove Trailing Characters 4
4. Capitalize a Sentence 5
5. Title Case a Name 5
6. Clean List of Names 5
7. Remove Custom Characters 6
8. Capitalize All Names in List 6
9. Clean Dictionary Values 7
10. Title Case Sentences in List 7
11. Clean and Title Case 7
12. Clean List of Emails 8
13. Remove Leading Numbers 8
14. Clean Nested List 8
15. Capitalize After Cleaning 9
16. Clean Dictionary Keys 9
17. Clean and Deduplicate the Names 9
18. Remove Multiple Character 10
19. Conditional Cleaning in List 10
20. Clean and Group By First Letter 10
21. Clean Set of String 11
22. Complex Nested Cleaning 11
23. Custom Title Function 12
24. Clean and Format Emails 12

25. **Multi Step Cleaning**.....12
26. **In-Place Cleaning**.....13
27. **Clean and Count Unique Word**.....13
28. **Clean Dictionary Sentences**.....13
29. **Selective Character Removal**.....14
30. **Batch Clean and Sort**.....14

1. Remove Space

Assignment: Remove leading and trailing spaces from a string using strip().

```
DP-Python > Assignment-1.py > ...
1   text = "    Rajendra Chimala      "
2   cleanData = text.strip()
3
4   print(f"Data Before Cleaning :  {text}")
5   print(f"Data After Cleaning :  {cleanData}")
6
```

Click here to run my code on browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data Before Cleaning :     Rajendra Chimala      
Data After Cleaning :  Rajendra Chimala
```

2. Remove Leading Characters

Assignment: Remove all leading asterisks from a string using lstrip().

```
DP-Python > Assignment-1.py > ...
1   text = "*****###Rajendra Chimala"
2   cleanData = text.lstrip("*#")
3
4   print(f"Data Before Cleaning :  {text}")
5   print(f"Data After Cleaning :  {cleanData}")
6
7
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data Before Cleaning :  *****###Rajendra Chimala
Data After Cleaning :  Rajendra Chimala
```

3. Remove Trailing Characters

Assignment: Remove all trailing exclamation marks from a string using rstrip().

```
DP-Python > Assignment-1.py > ...
1   text  = "Rajenrda Chimala**$$!!"
2   cleanData = text.rstrip("**$$!!")
3
4   print(f"Data Before Cleaning :  {text}")
5   print(f"Data After Cleaning :  {cleanData}")
```

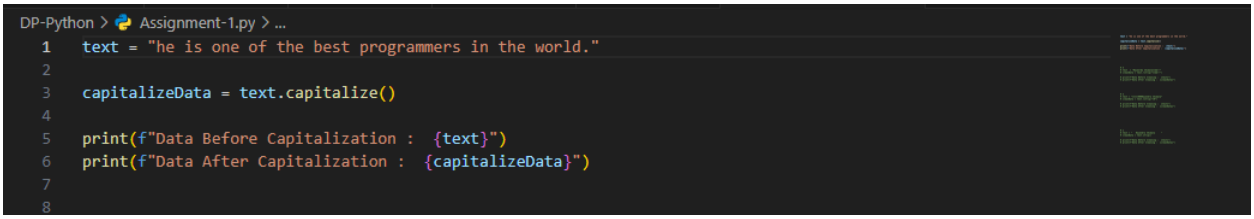
Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data Before Cleaning :  Rajenrda Chimala**$$!!
Data After Cleaning :  Rajenrda Chimala
```

Click here to run my code on browser

4. Capitalize a Sentence

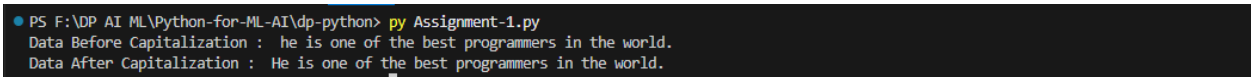
Assignment: Capitalize only the first letter of a lowercase sentence using capitalize().



```
DP-Python > Assignment-1.py > ...
1  text = "he is one of the best programmers in the world."
2
3  capitalizeData = text.capitalize()
4
5  print(f"Data Before Capitalization :  {text}")
6  print(f"Data After Capitalization :  {capitalizeData}")
7
8
```

Click here to run my code on browser

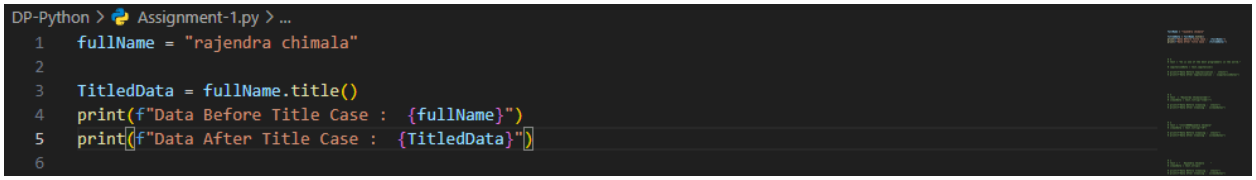
Output:



```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data Before Capitalization :  he is one of the best programmers in the world.
Data After Capitalization :  He is one of the best programmers in the world.
```

5. Title Case a Name

Assignment: Convert a name to title case using title().



```
DP-Python > Assignment-1.py > ...
1  fullName = "rajendra chimala"
2
3  TitledData = fullName.title()
4  print(f"Data Before Title Case :  {fullName}")
5  print(f"Data After Title Case :  {TitledData}")
6
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data Before Title Case :  rajendra chimala
Data After Title Case :  Rajendra Chimala
```

6. Clean List of Names

Assignment: Remove leading and trailing spaces from each name in a list

```
DP-Python > Assignment-1.py > ...
1  fruitNames = [" Apple        ", "    Banana    ", " Mango", "  Pomegranet   ", "  Pineapple   "]
2
3  cleanedNames = [fName.strip() for fName in fruitNames]
4
5  print("Cleaned Fruit Names:")
6  for name in cleanedNames:
7      print(name)
8
```

Click here to run my code on browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Fruit Names:
Apple
Banana
Mango
Pomegranet
Pineapple
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

7. Remove Custom Characters

Assignment: Remove # and $ from both ends of a string using strip().

```
DP-Python > Assignment-1.py > ...
1   text = " *$!#  Rajendra Chimala    **$$!! "
2
3   cleanText = text.strip("*$!# ")
4
5   print(f"Data After Cleaning :  {cleanText}")
6
```

Click here to run my code on browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Cleaning :  Rajendra Chimala
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

8. Capitalize All Names in List

Assignment: Capitalize each name in a list using capitalize().

```
DP-Python > Assignment-1.py > ...
1   rawNames = ["rajendra","binod","suresh","karthik","ram","sita","gita","sanjay"]
2
3   capitalizeData = [name.capitalize() for name in rawNames]
4   print("Names After Capitalization:")
5   for name in capitalizeData:
6       print(name)
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Names After Capitalization:
Rajendra
Binod
Suresh
Karthik
Ram
Sita
Gita
Sanjay
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

9. Clean Dictionary Values

Assignment: Remove trailing spaces from all values in a dictionary.

```
DP-Python > Assignment-1.py > ...
1   data = {"name": "Rajendra    ", "city": "Dhanghadhi   ", "country": "Nepal   ","age": "25   "}
2   cleanedData = {k: v.rstrip() for k, v in data.items()}
3   print("Data After Cleaning:")
4   for key, value in cleanedData.items():
5       print(f"{key}: {value}")
```

Click here to run my code on browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Cleaning:
name: Rajendra
city: Dhanghadhi
country: Nepal
age: 25
```

10. Title Case Sentences in List

Assignment: Convert each sentence in a list to title case.

```
DP-Python > Assignment-1.py > ...
1   myList = ["hello world", "good morning", "python programming", "data science"]
2   result = [ml.title() for ml in myList]
3   print("Data After Title Case:")
4   for item in result:
5       print(item)
```

Click here to run my code on browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Title Case:
Hello World
Good Morning
Python Programming
Data Science
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

11. Clean and Title Case

Assignment: Remove spaces and convert to title case

```
DP-Python > Assignment-1.py > ...
    You, 2 minutes ago | 1 author (You)
1   text = " WellCome to dIGITAl PathShala "
2   formatedData = text.strip().title()
3
4   print("Formated Data : ",formatedData)
5
```

Click here to run my code in browser

Output

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Formated Data :  Wellcome To Digital Pathshala
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

12. Clean List of Emails

Assignment: Remove leading/trailing spaces from each email in a list.

```
DP-Python > Assignment-1.py > ...
    You, 9 seconds ago | 1 author (You)
1   raw_emails = [" ram@gmail.com", "sita@yahoo.com ","          rajendra@xmail.com"]
2   cleanedEmails = [e.strip() for e in raw_emails]
3   print(cleanedEmails)
4
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
['ram@gmail.com', 'sita@yahoo.com', 'rajendra@xmail.com']
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

13. Remove Leading Numbers

Assignment: Remove all leading digits from a string using lstrip().

```
DP-Python > Assignment-1.py > ...
    ...
1   text = "12345Welcome to Digital Pathshala"
2   cleaned = text.lstrip("1234")
3   print("Data After Cleaning:", cleaned)
4
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Cleaning: Welcome to Digital Pathshala
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

14. Clean Nested List

Assignment: Remove spaces from each string in a nested list

```
DP-Python > Assignment-1.py > ...
    You, 6 seconds ago | 1 author (You)
1   nestedList = [["  Rajendra Chimala  ", "  Binod Kumar  "], ["  Suresh Yadav  ", "  Karthik Sharma  "], ["  Ram
    Prasad  ", "  Sita Devi  "], ["  Gita Kumari  ", "  Sanjay Singh  "]]
2   cleaned = [[item.strip() for item in sublist] for sublist in nestedList]
3
4   print("Cleaned Nested List:", cleaned)
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Nested List: [['Rajendra Chimala', 'Binod Kumar'], ['Suresh Yadav', 'Karthik Sharma'], ['Ram Prasad', 'Sita Devi'], ['Gita Kumari', 'Sanja
y Singh']]
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

15. Capitalize After Cleaning

Assignment: Clean a string and capitalize the first letter.

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Cleaning and Capitalize :  Good morning!
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

16. Clean Dictionary Keys

Assignment: Remove trailing underscores from all dictionary keys.

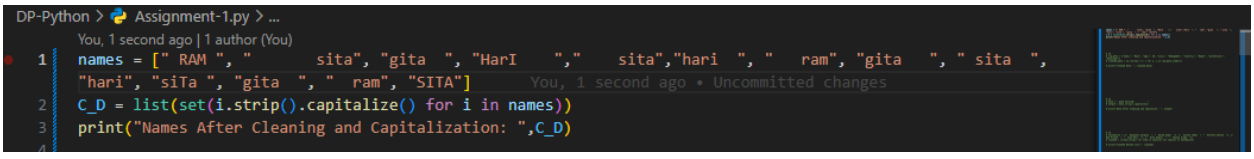
Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Data:  {'name': 'Raju', 'age': 20, 'city': 'Dhangadhi', 'country': 'Nepal', 'profession': 'Developer'}
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

17. Clean and Deduplicate the Names

Assignment: Clean, capitalize, and deduplicate names in a list.



Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Names After Cleaning and Capitalization:  ['Hari', 'Ram', 'Gita', 'Sita']
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

18. Remove Multiple Character

Assignment: Remove *, -, and spaces from both ends of a string.

```
DP-Python > Assignment-1.py > ...
    You, 1 second ago | 1 author (You)
1   raw_data = "$$$$@@@@&!!!Rajendra*___*---*^^^"
2   cleaned_data = raw_data.strip("*-$!_-^@&")      You, 1 second ago • Uncommitted changes
3   print("Cleaned Data: ", cleaned_data)
4
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Data:  Rajendra
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

19. Conditional Cleaning in List

Assignment: Remove leading # only if present in each string in a list.

```
DP-Python > Assignment-1.py > ...
    You, 7 seconds ago | 1 author (You)
1
2   my_list = ["#Development", "#AI", "ML","#DSA","NLP"]
3   C_D = [x.lstrip("#") if x.startswith("#") else x for x in my_list]
4
5   print("Cleaned Data: ", C_D)
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Data:  ['Development', 'AI', 'ML', 'DSA', 'NLP']
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

20. Clean and Group By First Letter

Assignment: Clean and group product names by their first letter (case-insensitive).

```python
products = [" apple", "-Banana", "apricot", "banana ","mango","momo","------orange"]
cleaned_data = [p.strip(" -").capitalize() for p in products]
grouped_data = {}
for p in cleaned_data:
 key = p[0].upper()
 grouped_data.setdefault(key, []).append(p)

print(grouped_data)
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
{'A': ['Apple', 'Apricot'], 'B': ['Banana', 'Banana'], 'M': ['Mango', 'Momo'], 'O': ['Orange']}
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

21. Clean Set of String

Assignment: Clean all strings in a set of special-character-laden strings

```python
raw_data = {"***ram***", "  sita    ", " Gita@@@@@@@@@@@@@@@ ","333333333hari-----"}
cleaned_data = {rd.strip("*@3- ").capitalize() for rd in raw_data}
print("Cleaned Data: ", cleaned_data)
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Data:  {'Gita', 'Sita', 'Ram', 'Hari'}
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

22. Complex Nested Cleaning

Assignment: Clean and title-case all strings in dictionary values (lists).

```python
raw_data = {"animals": [" rat", "rabbit ","  tiger"], "plants": ["palm ", " apple","neem"]}
cleaned_data = {i: [j.strip().title() for j in x] for i, x in raw_data.items()}

print("Cleaned Data: ", cleaned_data)
```

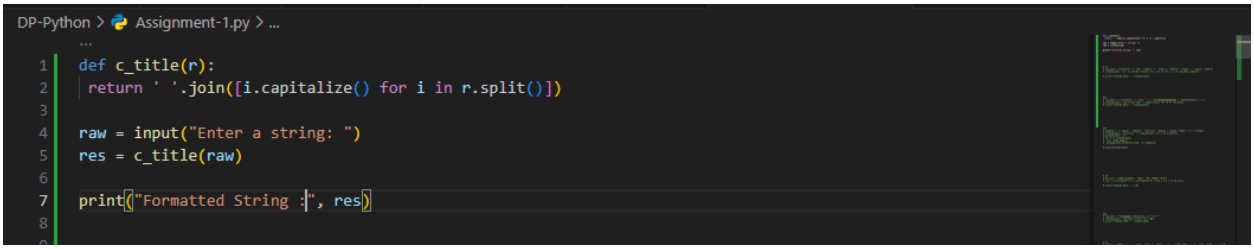
Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Cleaned Data:  {'animals': ['Rat', 'Rabbit', 'Tiger'], 'plants': ['Palm', 'Apple', 'Neem']}
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

23.Custom Title Function

Assignment: Implement your own title() using only capitalize() and loops.



```
DP-Python > Assignment-1.py > ...
...
1  def c_title(r):
2      return ' '.join([i.capitalize() for i in r.split()])
3
4  raw = input("Enter a string: ")
5  res = c_title(raw)
6
7  print("Formatted String :", res)
8
```

Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Enter a string: hello everyone i am rajendra chimala
Formatted String:  Hello Everyone I Am Rajendra Chimala
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

24. Clean and Format Emails

Assignment: Clean emails (strip spaces, lowercase all, except capitalize first letter before @).

```
DP-Python > Assignment-1.py > ...
You, 1 second ago | 1 author (You)
1  raw_emails = [" Ram@Example.com", "sitA@EXAMPLE.COM "]
2  cleaned_emails = [i.strip().split("@")[0].capitalize() + "@" + i.strip().split("@")[1].lower() for i in
   raw_emails]
3
4  print("Cleaned Emails: ", cleaned_emails)
```

Click here to run my code on browser

25. Multi Step Cleaning

Assignment: Remove leading numbers, trailing punctuation, and title-case the string.

```
You, 1 second ago | 1 author (You)
1  raw_data = "345235good morning!$234          "
2  cleaned_data = raw_data.strip("0123456789").rstrip("0123456789!$ ").title()
3
4  print("Data After Cleaning and Title Case: ", cleaned_data)
5
```

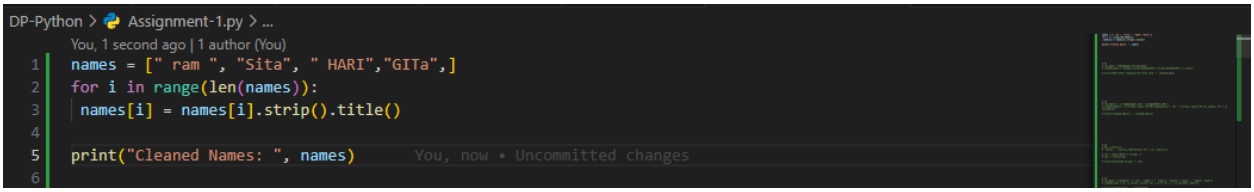
Click here to run my code on browser

Output:

```
PS F:\DP AI ML\Python-for-ML-AI\dp-python> py Assignment-1.py
Data After Cleaning and Title Case:  Good Morning
PS F:\DP AI ML\Python-for-ML-AI\dp-python>
```

26. In-Place Cleaning

Assignment: Clean and title-case a list of strings in place (no new list).

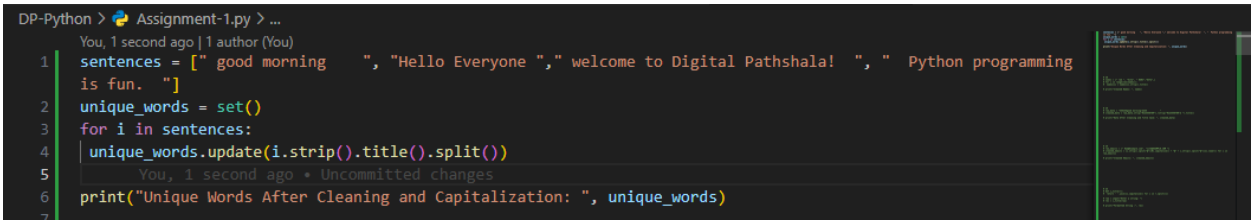


Click here to run my code in browser

Output:

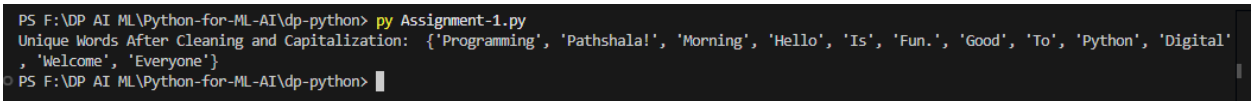
27. Clean and Count Unique Word

Assignment: Clean sentences, split into words, count unique words.



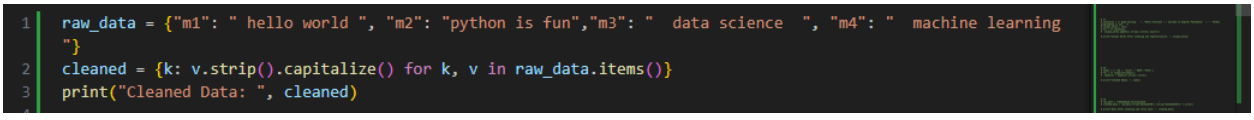
Click here to run my code on browser

Output:



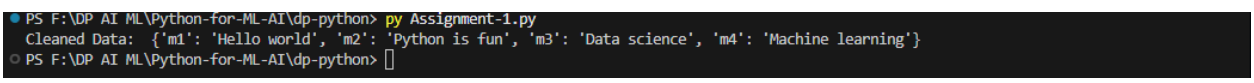
28. Clean Dictionary Sentences

Assignment: Clean and capitalize only the first word of each dictionary value.



Click here to run my code on browser

Output:



29. Selective Character Removal

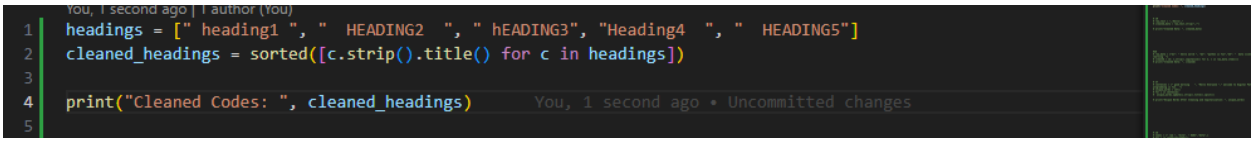
Assignment: Remove only leading/trailing underscores and dashes.

Click here to run my code on browser

Output:

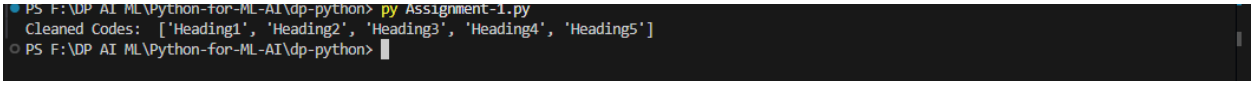
30. Batch Clean and Sort

Assignment: Clean and title-case product codes, then sort.



Click here to run my code on browser

Output:



---END---