

DSC540-T303-Data-Preparation-Week1-2

December 4, 2024

```
[ ]: # Weeks 1 & 2 Exercises
```

```
[85]: # 1. Install the latest version of Anaconda.  
# Print the version of Anaconda.  
import sys  
print(sys.version)  
print(sys.version_info)
```

3.12.4 | packaged by Anaconda, Inc. | (main, Jun 18 2024, 15:03:56) [MSC v.1929
64 bit (AMD64)]
sys.version_info(major=3, minor=12, micro=4, releaselevel='final', serial=0)

```
[1]: ## 2. Create a Jupyter notebook where you create a list, iterate over the list  
# and sort your results,  
## generate random numbers, add to the list, and then print your results.
```

```
[3]: # Answer:  
# step 1: Creates a list of numbers  
numbers = [5, 3, 8, 1, 9]  
print("Original List:", numbers)  
  
# step 2: Iterate over the list and print each number  
for number in numbers:  
    print(number)  
  
# step3: Sort a list  
sorted_numbers = sorted(numbers)  
print("Sorted List:", sorted_numbers)
```

Original List: [5, 3, 8, 1, 9]

5

3

8

1

9

Sorted List: [1, 3, 5, 8, 9]

```
[5]: # step5: Generate 5 random integers between 10 and 50
import random
new_random_numbers = [random.randint(10, 50) for _ in range(5)]
print("Generated Random Numbers:", new_random_numbers)
```

Generated Random Numbers: [50, 39, 30, 46, 30]

```
[7]: # step6: Add it to the previous list and print
sorted_numbers.extend(new_random_numbers)
print(sorted_numbers)
```

[1, 3, 5, 8, 9, 50, 39, 30, 46, 30]

```
[11]: ## 3. Create a line chart with Matplotlib and the following data file.
## a. Data file: world-population.xlsm

# Import required libraries.
# Matplotlib is a comprehensive library for creating static, animated, and
    ↳ interactive visualizations in Python
# pandas is a fast, powerful, flexible and easy to use open source data
    ↳ analysis and manipulation tool,
# built on top of the Python programming language

import matplotlib.pyplot as plt
import pandas as pd

# Step 1: Read data from Excel file
file_path = "world-population.xlsm"
data = pd.read_excel(file_path)

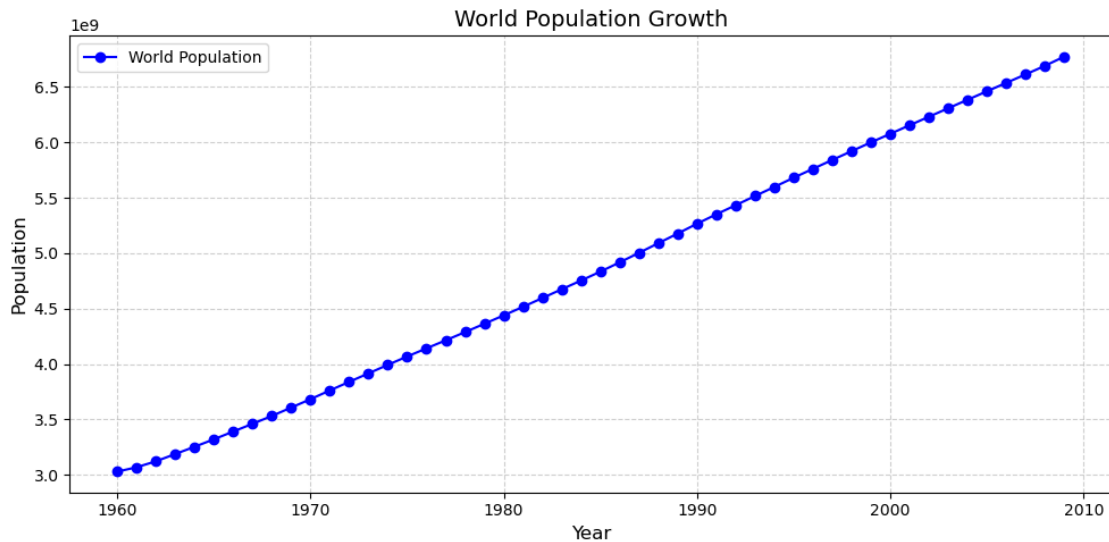
# Step 2: Extract columns (columns are 'Year' and 'Population')
years = data["Year"]
population = data["Population"]

# Step 3: Create a line chart
plt.figure(figsize=(10, 5)) # Set figure size
plt.plot(years, population, marker="o", linestyle="--", color="blue",
    ↳ label="World Population")

# Step 4: Customize the chart
plt.title("World Population Growth", fontsize=14)
plt.xlabel("Year", fontsize=12)
plt.ylabel("Population", fontsize=12)
plt.grid(True, linestyle="--", alpha=0.6)
plt.legend()
plt.tight_layout() # Adjust layout to prevent clipping

# Step 5: Display the chart
```

```
plt.show()
```



```
[ ]: ## 4a. The Data Wrangling Workshop: Activity 1.01 page 23
# Handling list
# Generate a list of random numbers and then generate another list from the
# first one,
# which only contain the numbers that are divisible by 3. Repeat the experiment
# 10 times, see the output each time.
# Calculate the average difference between the length of the two lists.
```

```
[17]: #import random library to generate the random numbers
import random
```

```
[19]: # Generate 100 random number
random_number_list = [random.randint(0, 100) for _ in range(100)]
```

```
[21]: # Print the output
print("Generated Random Numbers:", random_number_list)
```

```
Generated Random Numbers: [16, 84, 3, 1, 32, 13, 89, 96, 4, 44, 5, 85, 43, 46,
11, 33, 68, 34, 51, 63, 81, 46, 61, 35, 16, 94, 61, 72, 86, 82, 95, 26, 45, 93,
33, 27, 7, 65, 8, 5, 20, 16, 97, 50, 79, 91, 26, 21, 75, 95, 12, 19, 8, 22, 6,
1, 90, 85, 91, 87, 46, 47, 88, 62, 98, 64, 87, 27, 20, 44, 55, 55, 13, 61, 15,
39, 53, 12, 61, 0, 10, 9, 17, 15, 45, 13, 96, 56, 28, 26, 31, 79, 22, 57, 92, 4,
60, 4, 28, 1]
```

```
[ ]: # Generate another list from the first one, which only contain the numbers that
# are divisible by 3.
```

```
[23]: list_with_divisible_by_3 = [a for a in random_number_list if a % 3 == 0]
list_with_divisible_by_3
```

```
[23]: [84,
3,
96,
33,
51,
63,
81,
72,
45,
93,
33,
27,
21,
75,
12,
6,
90,
87,
87,
27,
15,
39,
12,
0,
9,
15,
45,
96,
57,
60]
```

```
[25]: # Find the difference in length of random list and the list of numbers which
      ↪ are divisible by 3
length_of_random_list = len(random_number_list)
length_of_3_divisible_list = len(list_with_divisible_by_3)
difference = length_of_random_list - length_of_3_divisible_list
difference
```

```
[25]: 70
```

```
[29]: # Repeat the experiment 10 times, see the output each time.
      # define constant for number of experiments.
      NUMBER_OF_EXPERIMENTS = 10
```

```

# Declare empty list to store the difference for each experiments
difference_list = []
for i in range(0, NUMBER_OF_EXPERIMENTS):
    random_number_list = [random.randint(0, 100) for x in range(0, 100)]
    list_with_divisible_by_3 = [a for a in random_number_list if a % 3 == 0]

    length_of_random_list = len(random_number_list)
    length_of_3_divisible_list = len(list_with_divisible_by_3)
    difference = length_of_random_list - length_of_3_divisible_list
    difference_list.append(difference)

# Print the data from the list
difference_list

```

[29]: [63, 68, 58, 60, 74, 66, 66, 64, 73, 64]

```

[31]: # Calculate the average difference
avg_diff = sum(difference_list) / float(len(difference_list))
avg_diff

```

[31]: 65.6

```

[29]: ## 4b. The Data Wrangling Workshop: Activity 1.02 page 43
# Analyzing a multiline string and generating the unique word count
# Get multiple text and ave it in a Python variable
# Get rid of all ne line in it uing string methods.
# Get all the unique words and occurences from the string
# Repeat the steps to find all unique words and occurrences, without
↳ considering case sensitivity.

```

```

[33]: # Create a multiline_text variable from copying the text from the first chapter
↳ of Pride & Prejudice
# https://gutenberg.org/files/1342/1342-h/1342-h.htm#Chapter_I

multiline_text= """IT is a truth universally acknowledged, that a single man in
↳ possession of a good fortune must be in want of a wife.

However little known the feelings or views of such a man may be on his first
↳ entering a neighbourhood, this truth is so well fixed in the minds of the
↳ surrounding families, that he is considered as the rightful property of some
↳ one or other of their daughters.

"My dear Mr. Bennet," said his lady to him one day, "have you heard that
↳ Netherfield Park is let at last?{2}"

Mr. Bennet replied that he had not.

```

"But it is," returned she; "for Mrs. Long has just been here, and she told me
↳all about it."

Mr. Bennet made no answer.

"Do not you want to know who has taken it?" cried his wife, impatiently.

"You want to tell me, and I have no objection to hearing it."
This was invitation enough.

"Why, my dear, you must know, Mrs. Long says that Netherfield is taken by a
↳young man of large fortune from the north of England; that he came down on
↳Monday in a chaise and four to see the place, and was so much delighted with
↳it that he agreed with Mr. Morris immediately; that he is to take possession
↳before Michaelmas, and some of his servants are to be in the house by the
↳end of next week.{3}"

"What is his name?"

"Bingley."

"Is he married or single?"

"Oh, single, my dear, to be sure! A single man of large fortune; four or five
↳thousand a year. What a fine thing for our girls!"

"How so? how can it affect them?"

"My dear Mr. Bennet," replied his wife, "how can you be so tiresome? You must
↳know that I am thinking of his marrying one of them."

"Is that his design in settling here?"

"Design? Nonsense, how can you talk so! But it is very likely that he may fall
↳in love with one of them, and therefore you must visit him as soon as he
↳comes."

"I see no occasion for that. You and the girls may go-or you may send them by
↳themselves, which perhaps will be still better; for as you are as handsome
↳as any of them, Mr. Bingley might like you the best of the party."

"My dear, you flatter me. I certainly have had my share of beauty, but I do not
↳pretend to be anything extraordinary now. When a woman has five grown-up
↳daughters, she ought to give over thinking of her own beauty."

"In such cases, a woman has not often much beauty to think of."

"But, my dear, you must indeed go and see Mr. Bingley when he comes into the neighbourhood."

"It is more than I engage for, I assure you."

"But consider your daughters. Only think what an establishment it would be for one of them. Sir William and Lady Lucas are determined to go, merely on that account; for in general, you know, they visit no new comers. Indeed you must go, for it will be impossible for us to visit him, if you do not."

"You are over scrupulous, surely. I dare say Mr. Bingley will be very glad to see you; and I will send a few lines by you to assure him of my hearty consent to his marrying whichever he chooses of the girls-though I must throw in a good word for my little Lizzy."

"I desire you will do no such thing. Lizzy is not a bit better than the others: and I am sure she is not half so handsome as Jane, nor half so good-humoured as Lydia. But you are always giving her the preference."

"They have none of them much to recommend them," replied he: "they are all silly and ignorant like other girls; but Lizzy has something more of quickness than her sisters."

"Mr. Bennet, how can you abuse your own children in such a way? You take delight in vexing me. You have no compassion on my poor nerves."

"You mistake me, my dear. I have a high respect for your nerves. They are my old friends. I have heard you mention them with consideration these twenty years at least."

"Ah, you do not know what I suffer."

"But I hope you will get over it, and live to see many young men of four thousand a year come into the neighbourhood."

"It will be no use to us, if twenty such should come, since you will not visit them."

"Depend upon it, my dear, that when there are twenty, I will visit them all."

```
Mr. Bennet was so odd a mixture of quick parts, sarcastic humour, reserve, and
↳caprice, that the experience of three-and-twenty years had been insufficient
↳to{5} make his wife understand his character. Her mind was less difficult to
↳develope. She was a woman of mean understanding, little information, and
↳uncertain temper. When she was discontented, she fancied herself nervous.
↳The business of her life was to get her daughters married: its solace was
↳visiting and news. I hope Mr. Bingley will like it."
```

```
[35]: # Find the type of the string
type(multiline_text)
```

```
[35]: str
```

```
[37]: # Find the length of the string
len(multiline_text)
```

```
[37]: 4517
```

```
[39]: # Remove new line charecter
multiline_text = multiline_text.replace('\n', "")
multiline_text
```

```
[39]: 'IT is a truth universally acknowledged, that a single man in possession of a
good fortune must be in want of a wife.However little known the feelings or
views of such a man may be on his first entering a neighbourhood, this truth is
so well fixed in the minds of the surrounding families, that he is considered as
the rightful property of some one or other of their daughters."My dear Mr.
Bennet," said his lady to him one day, "have you heard that Netherfield Park is
let at last?{2}"Mr. Bennet replied that he had not."But it is," returned she;
"for Mrs. Long has just been here, and she told me all about it."Mr. Bennet made
no answer."Do not you want to know who has taken it?" cried his wife,
impatiently."You want to tell me, and I have no objection to hearing it."This
was invitation enough."Why, my dear, you must know, Mrs. Long says that
Netherfield is taken by a young man of large fortune from the north of England;
that he came down on Monday in a chaise and four to see the place, and was so
much delighted with it that he agreed with Mr. Morris immediately; that he is to
take possession before Michaelmas, and some of his servants are to be in the
house by the end of next week.{3}"What is his name?"Bingley."Is he married or
single?"Oh, single, my dear, to be sure! A single man of large fortune; four or
five thousand a year. What a fine thing for our girls!"How so? how can it
affect them?"My dear Mr. Bennet," replied his wife, "how can you be so
tiresome? You must know that I am thinking of his marrying one of them."Is that
his design in settling here?"Design? Nonsense, how can you talk so! But it is
very likely that he may fall in love with one of them, and therefore you must
visit him as soon as he comes."I see no occasion for that. You and the girls
may go-or you may send them by themselves, which perhaps will be still better;
for as you are as handsome as any of them, Mr. Bingley might like you the best
```


of the party."My dear, you flatter me. I certainly have had my share of beauty, but I do not pretend to be anything extraordinary now. When a woman has five grown-up daughters, she ought to give over thinking of her own beauty."In such cases, a woman has not often much beauty to think of."But, my dear, you must indeed go and see Mr. Bingley when he comes into the neighbourhood."It is more than I engage for, I assure you."But consider your daughters. Only think what an establishment it would be for one of them. Sir William and Lady Lucas are determined to go, merely on that account; for in general, you know, they visit no new{4} comers. Indeed you must go, for it will be impossible for us to visit him, if you do not."You are over scrupulous, surely. I dare say Mr. Bingley will be very glad to see you; and I will send a few lines by you to assure him of my hearty consent to his marrying whichever he chooses of the girls-though I must throw in a good word for my little Lizzy."I desire you will do no such thing. Lizzy is not a bit better than the others: and I am sure she is not half so handsome as Jane, nor half so good-humoured as Lydia. But you are always giving her the preference."They have none of them much to recommend them," replied he: "they are all silly and ignorant like other girls; but Lizzy has something more of quickness than her sisters."Mr. Bennet, how can you abuse your own children in such a way? You take delight in vexing me. You have no compassion on my poor nerves."You mistake me, my dear. I have a high respect for your nerves. They are my old friends. I have heard you mention them with consideration these twenty years at least."Ah, you do not know what I suffer."But I hope you will get over it, and live to see many young men of four thousand a year come into the neighbourhood."It will be no use to us, if twenty such should come, since you will not visit them."Depend upon it, my dear, that when there are twenty, I will visit them all."Mr. Bennet was so odd a mixture of quick parts, sarcastic humour, reserve, and caprice, that the experience of three-and-twenty years had been insufficient to{5} make his wife understand his character. Her mind was less difficult to develope. She was a woman of mean understanding, little information, and uncertain temper. When she was discontented, she fancied herself nervous. The business of her life was to get her daughters married: its solace was visiting and news. I hope Mr. Bingley will like it.'

```
[41]: # remove special chars, punctuation etc.
cleaned_multiline_text = ""
for char in multiline_text:
    if char == " ":
        cleaned_multiline_text += char
    elif char.isalnum(): # using the isalnum() method of strings.
        cleaned_multiline_text += char
    else:
        cleaned_multiline_text += " "
```

```
[43]: cleaned_multiline_text
```

[43]: 'IT is a truth universally acknowledged that a single man in possession of a good fortune must be in want of a wife However little known the feelings or views of such a man may be on his first entering a neighbourhood this truth is so well fixed in the minds of the surrounding families that he is considered as the rightful property of some one or other of their daughters My dear Mr Bennet said his lady to him one day have you heard that Netherfield Park is let at last 2 Mr Bennet replied that he had not But it is returned she for Mrs Long has just been here and she told me all about it Mr Bennet made no answer Do not you want to know who has taken it cried his wife impatiently You want to tell me and I have no objection to hearing it This was invitation enough Why my dear you must know Mrs Long says that Netherfield is taken by a young man of large fortune from the north of England that he came down on Monday in a chaise and four to see the place and was so much delighted with it that he agreed with Mr Morris immediately that he is to take possession before Michaelmas and some of his servants are to be in the house by the end of next week 3 What is his name Bingley Is he married or single Oh single my dear to be sure A single man of large fortune four or five thousand a year What a fine thing for our girls How so how can it affect them My dear Mr Bennet replied his wife how can you be so tiresome You must know that I am thinking of his marrying one of them Is that his design in settling here Design Nonsense how can you talk so But it is very likely that he may fall in love with one of them and therefore you must visit him as soon as he comes I see no occasion for that You and the girls may go or you may send them by themselves which perhaps will be still better for as you are as handsome as any of them Mr Bingley might like you the best of the party My dear you flatter me I certainly have had my share of beauty but I do not pretend to be anything extraordinary now When a woman has five grown up daughters she ought to give over thinking of her own beauty In such cases a woman has not often much beauty to think of But my dear you must indeed go and see Mr Bingley when he comes into the neighbourhood It is more than I engage for I assure you But consider your daughters Only think what an establishment it would be for one of them Sir William and Lady Lucas are determined to go merely on that account for in general you know they visit no new 4 comers Indeed you must go for it will be impossible for us to visit him if you do not You are over scrupulous surely I dare say Mr Bingley will be very glad to see you and I will send a few lines by you to assure him of my hearty consent to his marrying whichever he chooses of the girls though I must throw in a good word for my little Lizzy I desire you will do no such thing Lizzy is not a bit better than the others and I am sure she is not half so handsome as Jane nor half so good humoured as Lydia But you are always giving her the preference They have none of them much to recommend them replied he they are all silly and ignorant like other girls but Lizzy has something more of quickness than her sisters Mr Bennet how can you abuse your own children in such a way You take delight in vexing me You have no compassion on my poor nerves You mistake me my dear I have a high respect for your nerves They are my old friends I have heard you mention them with consideration these twenty years at least Ah you do not know what I suffer

But I hope you will get over it and live to see many young men of four thousand a year come into the neighbourhood It will be no use to us if twenty such should come since you will not visit them Depend upon it my dear that when there are twenty I will visit them all Mr Bennet was so odd a mixture of quick parts sarcastic humour reserve and caprice that the experience of three and twenty years had been insufficient to 5 make his wife understand his character Her mind was less difficult to develope She was a woman of mean understanding little information and uncertain temper When she was discontented she fancied herself nervous The business of her life was to get her daughters married its solace was visiting and news I hope Mr Bingley will like it '

```
[45]: # Generate a list of all the word from the cleaned string using the split_
      ↪command
list_of_words = cleaned_multiline_text.split()
list_of_words
```

```
[45]: ['IT',
      'is',
      'a',
      'truth',
      'universally',
      'acknowledged',
      'that',
      'a',
      'single',
      'man',
      'in',
      'possession',
      'of',
      'a',
      'good',
      'fortune',
      'must',
      'be',
      'in',
      'want',
      'of',
      'a',
      'wife',
      'However',
      'little',
      'known',
      'the',
      'feelings',
      'or',
      'views',
```

'of',
'such',
'a',
'man',
'may',
'be',
'on',
'his',
'first',
'entering',
'a',
'neighbourhood',
'this',
'truth',
'is',
'so',
'well',
'fixed',
'in',
'the',
'minds',
'of',
'the',
'surrounding',
'families',
'that',
'he',
'is',
'considered',
'as',
'the',
'rightful',
'property',
'of',
'some',
'one',
'or',
'other',
'of',
'their',
'daughters',
'My',
'dear',
'Mr',
'Bennet',
'said',
'his',

'lady',
'to',
'him',
'one',
'day',
'have',
'you',
'heard',
'that',
'Netherfield',
'Park',
'is',
'let',
'at',
'last',
'2',
'Mr',
'Bennet',
'replied',
'that',
'he',
'had',
'not',
'But',
'it',
'is',
'returned',
'she',
'for',
'Mrs',
'Long',
'has',
'just',
'been',
'here',
'and',
'she',
'told',
'me',
'all',
'about',
'it',
'Mr',
'Bennet',
'made',
'no',
'answer',

'Do',
'not',
'you',
'want',
'to',
'know',
'who',
'has',
'taken',
'it',
'cried',
'his',
'wife',
'impatiently',
'You',
'want',
'to',
'tell',
'me',
'and',
'I',
'have',
'no',
'objection',
'to',
'hearing',
'it',
'This',
'was',
'invitation',
'enough',
'Why',
'my',
'dear',
'you',
'must',
'know',
'Mrs',
'Long',
'says',
'that',
'Netherfield',
'is',
'taken',
'by',
'a',
'young',

'man',
'of',
'large',
'fortune',
'from',
'the',
'north',
'of',
'England',
'that',
'he',
'came',
'down',
'on',
'Monday',
'in',
'a',
'chaise',
'and',
'four',
'to',
'see',
'the',
'place',
'and',
'was',
'so',
'much',
'delighted',
'with',
'it',
'that',
'he',
'agreed',
'with',
'Mr',
'Morris',
'immediately',
'that',
'he',
'is',
'to',
'take',
'possession',
'before',
'Michaelmas',
'and',

'some',
'of',
'his',
'servants',
'are',
'to',
'be',
'in',
'the',
'house',
'by',
'the',
'end',
'of',
'next',
'week',
'3',
'What',
'is',
'his',
'name',
'Bingley',
'Is',
'he',
'married',
'or',
'single',
'Oh',
'single',
'my',
'dear',
'to',
'be',
'sure',
'A',
'single',
'man',
'of',
'large',
'fortune',
'four',
'or',
'five',
'thousand',
'a',
'year',
'What',

'a',
'fine',
'thing',
'for',
'our',
'girls',
'How',
'so',
'how',
'can',
'it',
'affect',
'them',
'My',
'dear',
'Mr',
'Bennet',
'replied',
'his',
'wife',
'how',
'can',
'you',
'be',
'so',
'tiresome',
'You',
'must',
'know',
'that',
'I',
'am',
'thinking',
'of',
'his',
'marrying',
'one',
'of',
'them',
'Is',
'that',
'his',
'design',
'in',
'settling',
'here',
'Design',

'Nonsense',
'how',
'can',
'you',
'talk',
'so',
'But',
'it',
'is',
'very',
'likely',
'that',
'he',
'may',
'fall',
'in',
'love',
'with',
'one',
'of',
'them',
'and',
'therefore',
'you',
'must',
'visit',
'him',
'as',
'soon',
'as',
'he',
'comes',
'I',
'see',
'no',
'occasion',
'for',
'that',
'You',
'and',
'the',
'girls',
'may',
'go',
'or',
'you',
'may',

'send',
'them',
'by',
'themselves',
'which',
'perhaps',
'will',
'be',
'still',
'better',
'for',
'as',
'you',
'are',
'as',
'handsome',
'as',
'any',
'of',
'them',
'Mr',
'Bingley',
'might',
'like',
'you',
'the',
'best',
'of',
'the',
'party',
'My',
'dear',
'you',
'flatter',
'me',
'I',
'certainly',
'have',
'had',
'my',
'share',
'of',
'beauty',
'but',
'I',
'do',
'not',

'pretend',
'to',
'be',
'anything',
'extraordinary',
'now',
'When',
'a',
'woman',
'has',
'five',
'grown',
'up',
'daughters',
'she',
'ought',
'to',
'give',
'over',
'thinking',
'of',
'her',
'own',
'beauty',
'In',
'such',
'cases',
'a',
'woman',
'has',
'not',
'often',
'much',
'beauty',
'to',
'think',
'of',
'But',
'my',
'dear',
'you',
'must',
'indeed',
'go',
'and',
'see',
'Mr',

'Bingley',
'when',
'he',
'comes',
'into',
'the',
'neighbourhood',
'It',
'is',
'more',
'than',
'I',
'engage',
'for',
'I',
'assure',
'you',
'But',
'consider',
'your',
'daughters',
'Only',
'think',
'what',
'an',
'establishment',
'it',
'would',
'be',
'for',
'one',
'of',
'them',
'Sir',
'William',
'and',
'Lady',
'Lucas',
'are',
'determined',
'to',
'go',
'merely',
'on',
'that',
'account',
'for',

'in',
'general',
'you',
'know',
'they',
'visit',
'no',
'new',
'4',
'comers',
'Indeed',
'you',
'must',
'go',
'for',
'it',
'will',
'be',
'impossible',
'for',
'us',
'to',
'visit',
'him',
'if',
'you',
'do',
'not',
'You',
'are',
'over',
'scrupulous',
'surely',
'I',
'dare',
'say',
'Mr',
'Bingley',
'will',
'be',
'very',
'glad',
'to',
'see',
'you',
'and',
'I',

'will',
'send',
'a',
'few',
'lines',
'by',
'you',
'to',
'assure',
'him',
'of',
'my',
'hearty',
'consent',
'to',
'his',
'marrying',
'whichever',
'he',
'chooses',
'of',
'the',
'girls',
'though',
'I',
'must',
'throw',
'in',
'a',
'good',
'word',
'for',
'my',
'little',
'Lizzy',
'I',
'desire',
'you',
'will',
'do',
'no',
'such',
'thing',
'Lizzy',
'is',
'not',
'a',

'bit',
'better',
'than',
'the',
'others',
'and',
'I',
'am',
'sure',
'she',
'is',
'not',
'half',
'so',
'handsome',
'as',
'Jane',
'nor',
'half',
'so',
'good',
'humoured',
'as',
'Lydia',
'But',
'you',
'are',
'always',
'giving',
'her',
'the',
'preference',
'They',
'have',
'none',
'of',
'them',
'much',
'to',
'recommend',
'them',
'replied',
'he',
'they',
'are',
'all',
'silly',

'and',
'ignorant',
'like',
'other',
'girls',
'but',
'Lizzy',
'has',
'something',
'more',
'of',
'quickness',
'than',
'her',
'sisters',
'Mr',
'Bennet',
'how',
'can',
'you',
'abuse',
'your',
'own',
'children',
'in',
'such',
'a',
'way',
'You',
'take',
'delight',
'in',
'vexing',
'me',
'You',
'have',
'no',
'compassion',
'on',
'my',
'poor',
'nerves',
'You',
'mistake',
'me',
'my',
'dear',

'I',
'have',
'a',
'high',
'respect',
'for',
'your',
'nerves',
'They',
'are',
'my',
'old',
'friends',
'I',
'have',
'heard',
'you',
'mention',
'them',
'with',
'consideration',
'these',
'twenty',
'years',
'at',
'least',
'Ah',
'you',
'do',
'not',
'know',
'what',
'I',
'suffer',
'But',
'I',
'hope',
'you',
'will',
'get',
'over',
'it',
'and',
'live',
'to',
'see',
'many',

'young',
'men',
'of',
'four',
'thousand',
'a',
'year',
'come',
'into',
'the',
'neighbourhood',
'It',
'will',
'be',
'no',
'use',
'to',
'us',
'if',
'twenty',
'such',
'should',
'come',
'since',
'you',
'will',
'not',
'visit',
'them',
'Depend',
'upon',
'it',
'my',
'dear',
'that',
'when',
'there',
'are',
'twenty',
'I',
'will',
'visit',
'them',
'all',
'Mr',
'Bennet',
'was',

'so',
'odd',
'a',
'mixture',
'of',
'quick',
'parts',
'sarcastic',
'humour',
'reserve',
'and',
'caprice',
'that',
'the',
'experience',
'of',
'three',
'and',
'twenty',
'years',
'had',
'been',
'insufficient',
'to',
'5',
'make',
'his',
'wife',
'understand',
'his',
'character',
'Her',
'mind',
'was',
'less',
'difficult',
'to',
'develope',
'She',
'was',
'a',
'woman',
'of',
'mean',
'understanding',
'little',
'information',

```
'and',
'uncertain',
'temper',
'When',
'she',
'was',
'discontented',
'she',
'fancied',
'herself',
'nervous',
'The',
'business',
'of',
'her',
'life',
'was',
'to',
'get',
'her',
'daughters',
'married',
'its',
'solace',
'was',
'visiting',
'and',
'news',
'I',
'hope',
'Mr',
'Bingley',
'will',
'like',
'it']
```

```
[47]: # Find the number of words
len(list_of_words)
```

```
[47]: 864
```

```
[53]: # Create a list from the list with unique words
unique_words_as_dict = dict.fromkeys(list_of_words)
len(list(unique_words_as_dict.keys()))
```

```
[53]: 346
```

```
[55]: for word in list_of_words:
        if unique_words_as_dict[word] is None:
            unique_words_as_dict[word] = 1
        else:
            unique_words_as_dict[word] += 1
unique_words_as_dict
```

```
[55]: {'IT': 1,
      'is': 12,
      'a': 20,
      'truth': 2,
      'universally': 1,
      'acknowledged': 1,
      'that': 15,
      'single': 4,
      'man': 4,
      'in': 11,
      'possession': 2,
      'of': 29,
      'good': 3,
      'fortune': 3,
      'must': 7,
      'be': 11,
      'want': 3,
      'wife': 4,
      'However': 1,
      'little': 3,
      'known': 1,
      'the': 17,
      'feelings': 1,
      'or': 5,
      'views': 1,
      'such': 5,
      'may': 4,
      'on': 4,
      'his': 11,
      'first': 1,
      'entering': 1,
      'neighbourhood': 3,
      'this': 1,
      'so': 8,
      'well': 1,
      'fixed': 1,
      'minds': 1,
      'surrounding': 1,
      'families': 1,
      'he': 11,
```

'considered': 1,
'as': 8,
'rightful': 1,
'property': 1,
'some': 2,
'one': 5,
'other': 2,
'their': 1,
'daughters': 4,
'My': 3,
'dear': 8,
'Mr': 11,
'Bennet': 6,
'said': 1,
'lady': 1,
'to': 22,
'him': 4,
'day': 1,
'have': 7,
'you': 24,
'heard': 2,
'Netherfield': 2,
'Park': 1,
'let': 1,
'at': 2,
'last': 1,
'2': 1,
'replied': 3,
'had': 3,
'not': 9,
'But': 6,
'it': 12,
'returned': 1,
'she': 6,
'for': 11,
'Mrs': 2,
'Long': 2,
'has': 5,
'just': 1,
'been': 2,
'here': 2,
'and': 17,
'told': 1,
'me': 5,
'all': 3,
'about': 1,
'made': 1,

'no': 7,
'answer': 1,
'Do': 1,
'know': 5,
'who': 1,
'taken': 2,
'cried': 1,
'impatiently': 1,
'You': 7,
'tell': 1,
'I': 18,
'objection': 1,
'hearing': 1,
'This': 1,
'was': 8,
'invitation': 1,
'enough': 1,
'Why': 1,
'my': 10,
'says': 1,
'by': 4,
'young': 2,
'large': 2,
'from': 1,
'north': 1,
'England': 1,
'came': 1,
'down': 1,
'Monday': 1,
'chaise': 1,
'four': 3,
'see': 5,
'place': 1,
'much': 3,
'delighted': 1,
'with': 4,
'agreed': 1,
'Morris': 1,
'immediately': 1,
'take': 2,
'before': 1,
'Michaelmas': 1,
'servants': 1,
'are': 8,
'house': 1,
'end': 1,
'next': 1,

'week': 1,
'3': 1,
'What': 2,
'name': 1,
'Bingley': 5,
'Is': 2,
'married': 2,
'Oh': 1,
'sure': 2,
'A': 1,
'five': 2,
'thousand': 2,
'year': 2,
'fine': 1,
'thing': 2,
'our': 1,
'girls': 4,
'How': 1,
'how': 4,
'can': 4,
'affect': 1,
'them': 11,
'tiresome': 1,
'am': 2,
'thinking': 2,
'marrying': 2,
'design': 1,
'settling': 1,
'Design': 1,
'Nonsense': 1,
'talk': 1,
'very': 2,
'likely': 1,
'fall': 1,
'love': 1,
'therefore': 1,
'visit': 5,
'soon': 1,
'comes': 2,
'occasion': 1,
'go': 4,
'send': 2,
'themselves': 1,
'which': 1,
'perhaps': 1,
'will': 10,
'still': 1,

'better': 2,
'handsome': 2,
'any': 1,
'might': 1,
'like': 3,
'best': 1,
'party': 1,
'flatter': 1,
'certainly': 1,
'share': 1,
'beauty': 3,
'but': 2,
'do': 4,
'pretend': 1,
'anything': 1,
'extraordinary': 1,
'now': 1,
'When': 2,
'woman': 3,
'grown': 1,
'up': 1,
'ought': 1,
'give': 1,
'over': 3,
'her': 5,
'own': 2,
'In': 1,
'cases': 1,
'often': 1,
'think': 2,
'indeed': 1,
'when': 2,
'into': 2,
'It': 2,
'more': 2,
'than': 3,
'engage': 1,
'assure': 2,
'consider': 1,
'your': 3,
'Only': 1,
'what': 2,
'an': 1,
'establishment': 1,
'would': 1,
'Sir': 1,
'William': 1,

'Lady': 1,
'Lucas': 1,
'determined': 1,
'merely': 1,
'account': 1,
'general': 1,
'they': 2,
'new': 1,
'4': 1,
'comers': 1,
'Indeed': 1,
'impossible': 1,
'us': 2,
'if': 2,
'scrupulous': 1,
'surely': 1,
'dare': 1,
'say': 1,
'glad': 1,
'few': 1,
'lines': 1,
'hearty': 1,
'consent': 1,
'whichever': 1,
'chooses': 1,
'though': 1,
'throw': 1,
'word': 1,
'Lizzy': 3,
'desire': 1,
'bit': 1,
'others': 1,
'half': 2,
'Jane': 1,
'nor': 1,
'humoured': 1,
'Lydia': 1,
'always': 1,
'giving': 1,
'preference': 1,
'They': 2,
'none': 1,
'recommend': 1,
'silly': 1,
'ignorant': 1,
'something': 1,
'quickness': 1,

'sisters': 1,
'abuse': 1,
'children': 1,
'way': 1,
'delight': 1,
'vexing': 1,
'compassion': 1,
'poor': 1,
'nerves': 2,
'mistake': 1,
'high': 1,
'respect': 1,
'old': 1,
'friends': 1,
'mention': 1,
'consideration': 1,
'these': 1,
'twenty': 4,
'years': 2,
'least': 1,
'Ah': 1,
'suffer': 1,
'hope': 2,
'get': 2,
'live': 1,
'many': 1,
'men': 1,
'come': 2,
'use': 1,
'should': 1,
'since': 1,
'Depend': 1,
'upon': 1,
'there': 1,
'odd': 1,
'mixture': 1,
'quick': 1,
'parts': 1,
'sarcastic': 1,
'humour': 1,
'reserve': 1,
'caprice': 1,
'experience': 1,
'three': 1,
'insufficient': 1,
'5': 1,
'make': 1,

```

'understand': 1,
'character': 1,
'Her': 1,
'mind': 1,
'less': 1,
'difficult': 1,
'develope': 1,
'She': 1,
'mean': 1,
'understanding': 1,
'information': 1,
'uncertain': 1,
'temper': 1,
'discontented': 1,
'fancied': 1,
'herself': 1,
'nervous': 1,
'The': 1,
'business': 1,
'life': 1,
'its': 1,
'solace': 1,
'visiting': 1,
'news': 1}

```

```

[57]: top_words = sorted(unique_words_as_dict.items(), key=lambda key_val_tuple:
↳key_val_tuple[1], reverse=True)
top_words[:25]

```

```

[57]: [('of', 29),
('you', 24),
('to', 22),
('a', 20),
('I', 18),
('the', 17),
('and', 17),
('that', 15),
('is', 12),
('it', 12),
('in', 11),
('be', 11),
('his', 11),
('he', 11),
('Mr', 11),
('for', 11),
('them', 11),
('my', 10),

```

```

('will', 10),
('not', 9),
('so', 8),
('as', 8),
('dear', 8),
('was', 8),
('are', 8)]

```

```

[61]: ## 4c. The Data Wrangling Workshop: Activity 2.01 page 70
# Permutation, Iterator, Lambda and List
#Lookup the definition of permutation and dropwhile from itertools
# itertools.permutations - Generates all possible orderings of a given
    ↳iterable (e.g., string, list, etc.).
# itertools.dropwhile - Drops elements from the input iterable as long as the
    ↳condition specified in the predicate is True.
# Once the predicate is False, it yields the remaining elements.

# import required libraries
from itertools import permutations, dropwhile

# make use of '?' to describe the libraries
permutations?
dropwhile?

```

Init signature: `dropwhile(predicate, iterable, /)`

Docstring:

Drop items from the iterable while `predicate(item)` is true.

Afterwards, return every element until the iterable is exhausted.

Type: `type`

Subclasses:

```

[63]: # write an expression to generate all the possible 3 digit number using 1, 2
    ↳and 3
permutations(range(3))

```

[63]: `<itertools.permutations at 0x1548ffee840>`

```

[65]: # Loop over and print iterator expression generated before. Use isinstance
    ↳method to verify element are tuples
for number_tuple in permutations(range(3)):
    print(number_tuple)
    assert isinstance(number_tuple, tuple)

```

```

(0, 1, 2)
(0, 2, 1)
(1, 0, 2)
(1, 2, 0)

```

```
(2, 0, 1)
(2, 1, 0)
```

```
[67]: # Drop any leading zero from tuple and print it.
for number_tuple in permutations(range(3)):
    print(list(dropwhile(lambda x: x <= 0, number_tuple)))
```

```
[1, 2]
[2, 1]
[1, 0, 2]
[1, 2, 0]
[2, 0, 1]
[2, 1, 0]
```

```
[69]: # Return the whole number from the list

import math

# create a function which take the values from each element of the tuple and
↳ return into hole number
def convert_to_number(number_stack):
    final_number = 0
    for i in range(0, len(number_stack)):
        final_number += (number_stack.pop() * (math.pow(10, i)))
    return final_number

for number_tuple in permutations(range(3)):
    number_stack = list(dropwhile(lambda x: x <= 0, number_tuple))
    print(convert_to_number(number_stack))
```

```
12.0
21.0
102.0
120.0
201.0
210.0
```

```
[33]: ## 4d. The Data Wrangling Workshop: Activity 2.02 page 81
```

```
[ ]: # Designing your own csv parser
```

```
[71]: # import zip_longest from itertools
from itertools import zip_longest
```

```
[79]: # create a function 'return_dict_from_csv_line' which will take 2 inputs -> 1)
↳ header values 2) line
def return_dict_from_csv_line(header, line):
```

```

    # Zip them. The zip_longest function is part of the itertools module in
    ↪Python.
    # It takes multiple iterables as input and returns an iterator that
    ↪aggregates elements from each iterable.
    # The key difference between zip_longest and the standard zip function is
    ↪how they handle iterables of different lengths.
    zipped_line = zip_longest(header, line, fillvalue=None)
    # Use dict comprehension to generate the final dict
    ret_dict = {kv[0]: kv[1] for kv in zipped_line}
    return ret_dict

```

```

[ ]: # Open the file in r mode in inside a with block. read the first line and use
    ↪tring method to generate a list of all the column name
    # fill up the key value variable

```

```

[75]: with open("sales_record.csv", "r") as fd:
    first_line = fd.readline()
    header = first_line.replace("\n", "").split(",")
    for i, line in enumerate(fd):
        # Here we loop over the first 10 lines in order to not to make the
        ↪output too big
        line = line.replace("\n", "").split(",")
        d = return_dict_from_csv_line(header, line)
        print(d)
        if i > 10:
            break

```

```

{'Region': 'Central America and the Caribbean', 'Country': 'Antigua and Barbuda',
 'Item Type': 'Baby Food', 'Sales Channel': 'Online', 'Order Priority': 'M',
 'Order Date': '12/20/2013', 'Order ID': '957081544', 'Ship Date': '1/11/2014',
 'Units Sold': '552', 'Unit Price': '255.28', 'Unit Cost': '159.42', 'Total
Revenue': '140914.56', 'Total Cost': '87999.84', 'Total Profit': '52914.72'}
{'Region': 'Central America and the Caribbean', 'Country': 'Panama', 'Item
Type': 'Snacks', 'Sales Channel': 'Offline', 'Order Priority': 'C', 'Order
Date': '7/5/2010', 'Order ID': '301644504', 'Ship Date': '7/26/2010', 'Units
Sold': '2167', 'Unit Price': '152.58', 'Unit Cost': '97.44', 'Total Revenue':
'330640.86', 'Total Cost': '211152.48', 'Total Profit': '119488.38'}
{'Region': 'Europe', 'Country': 'Czech Republic', 'Item Type': 'Beverages',
 'Sales Channel': 'Offline', 'Order Priority': 'C', 'Order Date': '9/12/2011',
 'Order ID': '478051030', 'Ship Date': '9/29/2011', 'Units Sold': '4778', 'Unit
Price': '47.45', 'Unit Cost': '31.79', 'Total Revenue': '226716.10', 'Total
Cost': '151892.62', 'Total Profit': '74823.48'}
{'Region': 'Asia', 'Country': 'North Korea', 'Item Type': 'Cereal', 'Sales
Channel': 'Offline', 'Order Priority': 'L', 'Order Date': '5/13/2010', 'Order
ID': '892599952', 'Ship Date': '6/15/2010', 'Units Sold': '9016', 'Unit Price':
'205.70', 'Unit Cost': '117.11', 'Total Revenue': '1854591.20', 'Total Cost':
'1055863.76', 'Total Profit': '798727.44'}

```



```
{'Region': 'Asia', 'Country': 'Sri Lanka', 'Item Type': 'Snacks', 'Sales
Channel': 'Offline', 'Order Priority': 'C', 'Order Date': '7/20/2015', 'Order
ID': '571902596', 'Ship Date': '7/27/2015', 'Units Sold': '7542', 'Unit Price':
'152.58', 'Unit Cost': '97.44', 'Total Revenue': '1150758.36', 'Total Cost':
'734892.48', 'Total Profit': '415865.88'}
{'Region': 'Middle East and North Africa', 'Country': 'Morocco', 'Item Type':
'Personal Care', 'Sales Channel': 'Offline', 'Order Priority': 'L', 'Order
Date': '11/8/2010', 'Order ID': '412882792', 'Ship Date': '11/22/2010', 'Units
Sold': '48', 'Unit Price': '81.73', 'Unit Cost': '56.67', 'Total Revenue':
'3923.04', 'Total Cost': '2720.16', 'Total Profit': '1202.88'}
{'Region': 'Australia and Oceania', 'Country': 'Federated States of Micronesia',
'Item Type': 'Clothes', 'Sales Channel': 'Offline', 'Order Priority': 'H',
'Order Date': '3/28/2011', 'Order ID': '932776868', 'Ship Date': '5/10/2011',
'Units Sold': '8258', 'Unit Price': '109.28', 'Unit Cost': '35.84', 'Total
Revenue': '902434.24', 'Total Cost': '295966.72', 'Total Profit': '606467.52'}
{'Region': 'Europe', 'Country': 'Bosnia and Herzegovina', 'Item Type':
'Clothes', 'Sales Channel': 'Online', 'Order Priority': 'M', 'Order Date':
'10/14/2013', 'Order ID': '919133651', 'Ship Date': '11/4/2013', 'Units Sold':
'927', 'Unit Price': '109.28', 'Unit Cost': '35.84', 'Total Revenue':
'101302.56', 'Total Cost': '33223.68', 'Total Profit': '68078.88'}
{'Region': 'Middle East and North Africa', 'Country': 'Afghanistan', 'Item
Type': 'Clothes', 'Sales Channel': 'Offline', 'Order Priority': 'M', 'Order
Date': '8/27/2016', 'Order ID': '579814469', 'Ship Date': '10/5/2016', 'Units
Sold': '8841', 'Unit Price': '109.28', 'Unit Cost': '35.84', 'Total Revenue':
'966144.48', 'Total Cost': '316861.44', 'Total Profit': '649283.04'}
{'Region': 'Sub-Saharan Africa', 'Country': 'Ethiopia', 'Item Type': 'Baby
Food', 'Sales Channel': 'Online', 'Order Priority': 'M', 'Order Date':
'4/13/2015', 'Order ID': '192993152', 'Ship Date': '5/7/2015', 'Units Sold':
'9817', 'Unit Price': '255.28', 'Unit Cost': '159.42', 'Total Revenue':
'2506083.76', 'Total Cost': '1565026.14', 'Total Profit': '941057.62'}
{'Region': 'Middle East and North Africa', 'Country': 'Turkey', 'Item Type':
'Office Supplies', 'Sales Channel': 'Offline', 'Order Priority': 'C', 'Order
Date': '9/25/2013', 'Order ID': '557156026', 'Ship Date': '10/15/2013', 'Units
Sold': '3704', 'Unit Price': '651.21', 'Unit Cost': '524.96', 'Total Revenue':
'2412081.84', 'Total Cost': '1944451.84', 'Total Profit': '467630.00'}
{'Region': 'Middle East and North Africa', 'Country': 'Oman', 'Item Type':
'Cosmetics', 'Sales Channel': 'Online', 'Order Priority': 'M', 'Order Date':
'5/12/2013', 'Order ID': '741101920', 'Ship Date': '5/17/2013', 'Units Sold':
'7382', 'Unit Price': '437.20', 'Unit Cost': '263.33', 'Total Revenue':
'3227410.40', 'Total Cost': '1943902.06', 'Total Profit': '1283508.34'}
```