"You may also like"

Netflix and Amazon

Recommendation systems process data from people to recommend what you might like.

- Netflix / YouTube - Shows
- Amazon - Books and more
- Last.fm - Music
- Twitter - Who to follow

$1M Netflix Prize

Unit 2 Recommendation Systems

APPLICATIONS

In this unit, we'll learn about the computing science field of **recommendation systems**, which **predict what you might enjoy** based on past experience.

PROGRAMMING LANGUAGE

In Python, we'll be learning the syntax and keywords to implement our algorithms.

ALGORITHMS

We'll learn more about **loops**, **integer** and **float** types, **calculating**, **variables**, operators and **files**, and more.

DOCUMENTATION AND TESTING

We will show how to tell if our program is any good or not!

Today's Topics

- Review **for** loops with **range**
- Review **integer** and **float** data types
- Accumulator pattern

Next week, we will work with data files and you'll learn how to process large amounts of data!

Popularity Contest

Recommendation algorithms

Basic What's the Most Popular?

One way to recommend something is simply to propose the most popular one, for example: most read news story, most visited cafe, most bought toaster.

Advanced People like you also liked...

Let's say Andrea likes *apples*, **bananas** and **cherries**.
And let's say Bob likes *durian*, **bananas**, and **cherries**.

Maybe Andrea would like *durian* (?!)
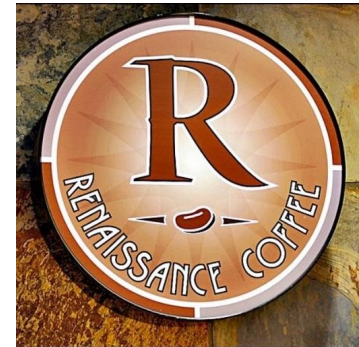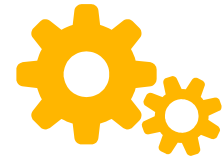And maybe Bob would like *apples*.
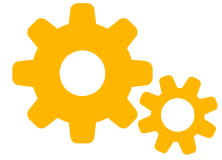
SFU coffee

First challenge: find out what the most popular cafe at SFU is.

Where do I start?

Write an interactive program to ask 5 people their favourite SFU cafe, and output the number of "likes" per cafe.

Design an Algorithm   **Write it in Python**   **Test and Deploy**   **Change the World**
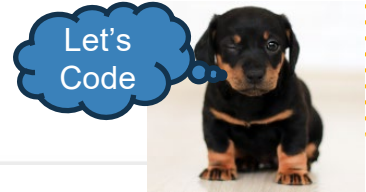
A Counting **Algorithm**

```
1   # SFU Popular Cafes
2   # Author: Angelica Lim
3   # Date: Nov. 30, 2017
4
5   # This program gets input from 5 users to find out
6   # what their favourite cafe on campus is.
7   # It prints out the number of people who like Starbucks, Renaissance
8   # or Tim Hortons
9
```

Think: How would you do this in person?

REDACTED
(for the moment)

Let's Code

Fork this: https://replit.com/@ajlsfu/HumbleTestyApplescript#main.py

ALGORITHM   9

Counting in **Python**

```
1    # SFU Popular Cafes
2    # Author: Angelica Lim
3    # Date: Nov. 30, 2017
4
5    # This program gets input from 5 users to find out
6    # what their favourite cafe on campus is.
7    # It prints out the number of people who like Starbucks, Renaissance
8    # or Tim Hortons
9
10   # Initialize tallies
11   starbucks_tally = 0
12
13   # Ask the user what their favourite cafe is
14   favourite_cafe = input("What's your favourite cafe?")
15
16   # If they say starbucks, add one to a starbucks tally
17   if favourite_cafe.lower() == "starbucks":
18     starbucks_tally = starbucks_tally + 1
19
20   print(starbucks_tally)
21   # If they say rennaissance, add one to a renaissance tally
22   # If they say tim hortons, add one to a timmy's tally
23   # Do the above 5 times
24   # In the end, print out the tallies
25   |
```

> The variable starbucks_tally is of type **integer**. We **initialize** it to 0.

> This is how you can add to yourself, or "accumulate"

http://interactivepython.org/runestone/static/thinkcspy/SimplePythonData/OperatorsandOperands.html

Counting in **Python**

```python
9
10  # Initialize tallies
11  starbucks_tally = 0
12  renaissance_tally = 0
13
14  # Ask the user what their favourite cafe is
15  favourite_cafe = input("What's your favourite cafe?")
16
17  # If they say starbucks, add one to a starbucks tally
18 ▾ if favourite_cafe.lower() == "starbucks":
19      starbucks_tally = starbucks_tally + 1
20
21  # If they say rennaissance, add one to a renaissance tally
22 ▾ elif favourite_cafe.lower() == "renaissance":
23      renaissance_tally = renaissance_tally + 1
24
25  # If they say tim hortons, add one to a timmy's tally
26  # Do the above 5 times
27  # In the end, print out the tallies
```

Please fill in this part

Counting in **Python**

```
10    # Initialize tallies
11    starbucks_tally = 0
12    renaissance_tally = 0
13    timmys_tally = 0
14
15 ▾  for i in range(5):
16        # Ask the user what their favourite cafe is
17        favourite_cafe = input("What's your favourite cafe?")
18
19        # If they say starbucks, add one to a starbucks tally
20 ▾      if favourite_cafe.lower() == "starbucks":
21            starbucks_tally = starbucks_tally + 1
22
23        # If they say rennaissance, add one to a renaissance tally
24 ▾      elif favourite_cafe.lower() == "renaissance":
25            renaissance_tally = renaissance_tally + 1
26
27        # If they say tim hortons, add one to a timmy's tally
28 ▾      elif favourite_cafe.lower() == "tim hortons":
29            timmys_tally = timmys_tally + 1
30
31    # Do the above 5 times
32    # In the end, print out the tallies
33    print(starbucks_tally)
34    print(renaissance_tally)
35    print(timmys_tally)
```

This line means: Do the following **indented block 5 times**, and set the value of **i** to **0, 1, 2, 3** and **4** each time, respectively.

String Conversion Python

```
32    # In the end, print out the tallies
33    print("Starbucks: " + starbucks_tally)
34    print(renaissance_tally)
35    print(timmys_tally)
```

```
Python 3.6.1 (default, Dec 2015, 13:05:11)
[GCC 4.8.2] on linux
>
What's your favourite cafe? starbucks
What's your favourite cafe? starbucks
What's your favourite cafe? renaissance
What's your favourite cafe? tim hortons
What's your favourite cafe? starbucks
Traceback (most recent call last):
  File "python", line 33, in <module>
TypeError: must be str, not int
>
```

```
32    # In the end, print out the tallies
33    print("Starbucks: " + str(starbucks_tally))
34    print(renaissance_tally)
35    print(timmys_tally)
```

```
Python 3.6.1 (default, Dec 2015, 13:05:11)
[GCC 4.8.2] on linux
>
What's your favourite cafe? starbucks
What's your favourite cafe? starbucks
What's your favourite cafe? renaissance
What's your favourite cafe? tim hortons
What's your favourite cafe? starbucks
Starbucks: 3
1
1
>
```
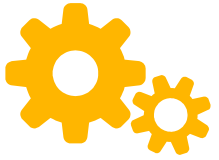
Data **types**

	Example
string	"hello"
integer or int	7
float	3.5

e.g.   favourite_food = "tempura"
      my_age = 19
      → We say that favourite_food is a **string** type and my_age is an **int**

Counting in **Python**

```python
10  # Initialize tallies
11  starbucks_tally = 0
12  renaissance_tally = 0
13  timmys_tally = 0
14
15  # Do the below 5 times
16  for i in range(5):
17      # Ask the user what their favourite cafe is
18      favourite_cafe = input("What's your favourite cafe?")
19
20      # If they say starbucks, add one to a starbucks tally
21      if favourite_cafe.lower() == "starbucks":
22          starbucks_tally = starbucks_tally + 1
23
24      # If they say rennaissance, add one to a renaissance tally
25      elif favourite_cafe.lower() == "renaissance":
26          renaissance_tally = renaissance_tally + 1
27
28      # If they say tim hortons, add one to a timmy's tally
29      elif favourite_cafe.lower() == "tim hortons":
30          timmys_tally = timmys_tally + 1
31
32  # In the end, print out the tallies
33  print("Starbucks: " + str(starbucks_tally))
34  print("Renaissance: " + str(renaissance_tally))
35  print("Tim Hortons: " + str(timmys_tally))
```

You can only concatenate strings with strings, so you need to convert your integer to a **string** type using **str()**
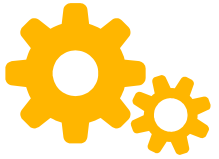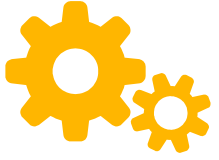
What's that i thing?

```
10  # Initialize tallies
11  starbucks_tally = 0
12  renaissance_tally = 0
13  timmys_tally = 0
14  other_tally = 0
15
16  # Do the below 5 times
17  for i in range(5):
18      # Ask the user what their favourite cafe is
19      favourite_cafe = input(str(i) + ". What's your favourite cafe?")
20
21      # If they say starbucks, add one to a starbucks tally
22      if favourite_cafe.lower() == "starbucks":
23          starbucks_tally = starbucks_tally + 1
24
25      # If they say rennaissance, add one to a renaissance tally
26      elif favourite_cafe.lower() == "renaissance":
27          renaissance_tally = renaissance_tally + 1
28
29      # If they say tims or timmys or tim horton
30      # add one to a timmy's tally
31      elif "tim" in favourite_cafe.lower():
32          timmys_tally = timmys_tally + 1
33
34      # Otherwise, add one to an "other" category
35      else:
36          other_tally = other_tally + 1
37
38  # In the end, print out the tallies
39  print("Starbucks: " + str(starbucks_tally))
40  print("Renaissance: " + str(renaissance_tally))
41  print("Tim Hortons: " + str(timmys_tally))
42  print("Other: " + str(other_tally))
```

```
0. What's your favourite cafe? starbucks
1. What's your favourite cafe? renaissance
2. What's your favourite cafe? timmy's
3. What's your favourite cafe? starbucks
4. What's your favourite cafe? renaissance
Starbucks: 2
Renaissance: 2
Tim Hortons: 1
Other: 0
>
```

What's that **i** thing?

```
15
16   # Do the below 5 times
17-  for i in range(5):
18       # Ask the user what their favourite cafe is
19       favourite_cafe = input(str(i) + ". What's your favourite cafe?")
20
```

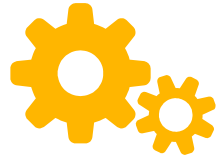**range(5)** is equivalent to **[0, 1, 2, 3, 4]**

```
0. What's your favourite cafe? starbucks
1. What's your favourite cafe? renaissance
2. What's your favourite cafe? timmy's
3. What's your favourite cafe? starbucks
4. What's your favourite cafe? renaissance
```

Remember, in computing science, we start counting from 0…. Sorry!

Update our Program

- Display the loop count in our prompt.
  - Make it start at 1…

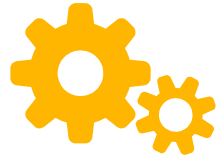- At the end, display each store's percentage of business
  - Format: 23.5%

Let's Code

Manipulating the i variable

```python
1   # SFU Popular Cafes
2   # Angelica Lim
3   # Feb. 1, 2021
4   # A survey to deduce the most popular cafe at SFU
5   # Ask 5 users what their favourite cafe on campus is
6   # from Starbucks, Renaissance, Tim Hortons.
7
8   # Initialize tallies to 0
9   starbucks_tally = 0
10  tim_hortons_tally = 0
11  renaissance_tally = 0
12
13  for i in range(5):
14      # Ask the user what their favourite cafe is
15      favourite_cafe = input(str(i+1)+". What's your favourite cafe (Starbucks,
        Tim Hortons, Renaissance)? ").lower().strip(" .!")
16
17      # If they say starbucks, add 1 to the starbucks tally
18      if favourite_cafe == "starbucks":
19          starbucks_tally += 1
20
21      elif favourite_cafe == "tim hortons":
22          tim_hortons_tally += 1
23
24      elif favourite_cafe == "renaissance":
25          renaissance_tally += 1
```

Another way to manipulate the i variable, other than using range

```
1. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
2. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
3. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
4. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
5. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
```

Calculating **percentages**

```python
# Initialize tallies
starbucks_tally = 0
tim_tally = 0
renaissance_tally = 0

# Repeat the following 5 times

for i in range(5):  # Range(5) is the same as [0, 1, 2, 3, 4]
    # Ask users what their favourite cafe is
    favourite_cafe = input("What is your favourite cafe at SFU?").lower()

    # If starbucks, add 1 to the starbucks tally
    if favourite_cafe == "starbucks":
        starbucks_tally = starbucks_tally + 1

    # Otherwise if renaissance, add 1 to renaissance tally
    elif favourite_cafe == "tim hortons":
        tim_tally = tim_tally + 1

    # Otherwise if tim hortons, add 1 to tim hortons tally
    elif favourite_cafe == "renaissance":
        renaissance_tally = renaissance_tally + 1

# Print out the number of likes per cafe
print("Starbucks: " + str( starbucks_tally/5 ) )
print("Tim Horton's: " + str( tim_tally/5 ) )
print("Renaissance: " + str( renaissance_tally/5 ) )
```
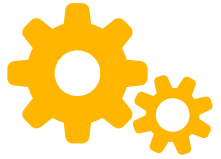
```
Python 3.6.1 (default, Dec 2015, 13:05:11)
[GCC 4.8.2] on linux
>
What is your favourite cafe at SFU? renaissance
What is your favourite cafe at SFU? starbucks
What is your favourite cafe at SFU? starbucks
What is your favourite cafe at SFU? tim hortons
What is your favourite cafe at SFU? tim hortons
Starbucks: 0.4
Tim Horton's: 0.4
Renaissance: 0.2
>
```

Dividing with a /

Calculating percentages

```python
# Initialize tallies
starbucks_tally = 0
tim_tally = 0
renaissance_tally = 0

# Repeat the following 5 times

for i in range(5):  # Range(5) is the same as [0, 1, 2, 3, 4]
    # Ask users what their favourite cafe is
    favourite_cafe = input("What is your favourite cafe at SFU?").lower()

    # If starbucks, add 1 to the starbucks tally
    if favourite_cafe == "starbucks":
        starbucks_tally = starbucks_tally + 1

    # Otherwise if renaissance, add 1 to renaissance tally
    elif favourite_cafe == "tim hortons":
        tim_tally = tim_tally + 1

    # Otherwise if tim hortons, add 1 to tim hortons tally
    elif favourite_cafe == "renaissance":
        renaissance_tally = renaissance_tally + 1

# Print out the number of likes per cafe
print("Starbucks: " + str( starbucks_tally/5*100 ) + "%" )
print("Tim Horton's: " + str( tim_tally/5*100 ) + "%" )
print("Renaissance: " + str( renaissance_tally/5*100 ) + "%" )
```

```
Python 3.6.1 (default, Dec 2015, 13:05:11)
[GCC 4.8.2] on linux
>
What is your favourite cafe at SFU? tim hortons
What is your favourite cafe at SFU? starbucks
What is your favourite cafe at SFU? starbucks
What is your favourite cafe at SFU? renaissance
What is your favourite cafe at SFU? tim hortons
Starbucks: 40.0%
Tim Horton's: 40.0%
Renaissance: 20.0%
>
```

Multiplying with *

Using string formatting
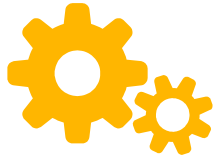
- `str(my_percentage)` converts to a string, but no control of format

```
1  # Unformatted output
2  score = 98.52353
3  print("Score: " + str(score) + "%")
4  print("Score: {}%".format(score))
5  print(f"Score: {score}%")
```

```
Score: 98.52353%
Score: 98.52353%
Score: 98.52353%
```

- Use `.format()` or f-strings (formatted string literals)

```
1  # Formatted output
2  score = 98.52353
3  print("Score: {:.2f}%".format(score))
4  print(f"Score: {score:.2f}%")
```

```
Score: 98.52%
Score: 98.52%
```

This says, print out a float with everything before the decimal and 2 places after it.

Let's Code

Using string formatting

```python
1   # SFU Popular Cafes
2   # Angelica Lim
3   # Feb. 1, 2021
4   # A survey to deduce the most popular cafe at SFU
5   # Ask 5 users what their favourite cafe on campus is
6   # from Starbucks, Renaissance, Tim Hortons.
7
8   # Initialize tallies to 0
9   starbucks_tally = 0
10  tim_hortons_tally = 0
11  renaissance_tally = 0
12
13  for i in range(5):
14      # Ask the user what their favourite cafe is
15      favourite_cafe = input(str(i+1)+". What's your favourite cafe (Starbucks, Tim
        Hortons, Renaissance)? ").lower().strip(" .!")
16
17      # If they say starbucks, add 1 to the starbucks tally
18      if favourite_cafe == "starbucks":
19          starbucks_tally += 1
20
21      elif favourite_cafe == "tim hortons":
22          tim_hortons_tally += 1
23
24      elif favourite_cafe == "renaissance":
25          renaissance_tally += 1
26
27  # Prints out the percentage of people who like
28  # Starbucks
29  print("Starbucks: {:.2f}%".format(starbucks_tally/5*100))
30  print("Tim Hortons: {:.2f}%".format(tim_hortons_tally/5*100))
31  print("Renaissance: {:.2f}%".format(renaissance_tally/5*100))
```

```
1. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
2. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
3. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
4. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
5. What's your favourite cafe (Starbucks, Tim Hortons, Renaissance)? starbucks
Starbucks: 100.00%
Tim Hortons: 0.00%
Renaissance: 0.00%
>
```

This says, print out a float with everything before the decimal and 2 places after it.
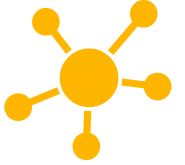
Let's **review** some concepts

How do you translate "do the following things 5 times" into Python?

How would you initialize a variable called "tally" to 0?

How would you add 1 to a variable "tally" that was already initialized to 0?

How do you convert an integer into a string?

What does the following Python code print as the first line?

```
for i in range(10,0,-1):
    print("{:.2f} dollars".format(i*1/10))
```

"You may also like"

Question 1

What is the data type of the variable **mystery** and **wonder** in this code?

```
mystery = 0
wonder = 0.
```

A. Mystery is an **int**, wonder is an **int**.

B. Mystery is an **int**, wonder is a **float**.

C. Mystery is a **float**, wonder is a **float**.

D. Mystery is an **int**, wonder is a **string**

Question 2

What is another way to write this line of code?

```
for i in [0, 1, 2]:
```

A. `i =0, i=1, i=2:`

B. `for i in range(2):`

C. `for I in ["0" to "2"]:`

D. `for i in range(3):`

Question **3**

Will this code run? If not, how can you fix it?

```
score = 10
print("Your score is " + score)
```

A. `score = 10.0`
B. `print("Your score is " + "score")`
C. `print("Your score is " + str(score))`
D. `print("Your score is " + int(score))`

Question 4

What is another way to write this code?

```
score = 10.553
print("Your score is {:.1f}".format(score))
```

A. `print(f"Your score is {score:.1f}")`

B. `print("Your score is " + score)`

C. `print("Your score is {:.1f}", score)`

D. `print("Your score is 10.55"`

Let's **review** some concepts

What do you need to do to numeric input before performing calculations on it?

What are two data types for numbers? What's the difference?

What function could you use to get the number of elements in a list?

What will this code output?

```
print( 3 / 3)
```

What is wrong with this code fragment?

```
score += 1
```