Files and the Filesystem

File Paths

```
/Users/christina/Documents/my_project/data.csv
/home/christina/my_project/data.csv
C:\Users\christina\Documents\my_project\data.csv
```

are examples of _____

```
../my_project/data/data.csv
data.csv
data/data.csv
```

are examples of _____

Special Patterns

In file paths,

`..` means _____

`~` means _____

Additional Concepts

The working directory is _____

_____

Working with Files

To retrieve information from a file, _____ from it. To put information into a file, _____ to it.

CSV (comma separated values) are one type of plain text file. Plain text files have no _____ such as bold text, colors, or fonts.

R and Python expect that when data is stored in CSV (or tab-delimited) files, each observation is a _____ and each variable is a _____. Rows and columns may or may not have names. Data is stored in a rectangle: each row has the same number of columns, and each column has the same number of rows.

Data Types

Common data types include:

- Boolean
- Integer
- Numeric
- Character

Boolean variables can be either _____ or _____. When converted to an integer, _____ becomes 0 and _____ becomes 1.

Most languages also have special types such as `NULL` or `None` that indicate no value. These special types are different from missing value indicators (e.g. `NA`).

Character data is also known as _____ or _____ data.

Strings

Tabs, spaces, and new line characters are examples of _____ characters.

`\n` is a _____

`\t` is a _____

A string without any characters in it (length 0) is called an _____ string.

Strings are sorted in alphabetical order. Lower case letters are different from upper case letters. The order of upper and lower case letters depends on the program.

Strings must be surrounded by _____. In R and Python, single or double _____ can be used, but they must match. Pick one style and be consistent where possible!

`"sub"` is a _____ of `"substring"`.

Concatenating strings means to _____ the strings together.

Variables

Variables let us refer to a value with a name. We can use the same name, but change the value.

`<-` in R, and `=` in Python, are _____ operators. The name of the variable goes on the _____ side, and the value goes on the _____. Everything on the right hand side is evaluated first before the value is assigned to the variable.

In R and Python, a variable with name `age_list` is the [**same/not the same**] is a variable with name `Age_list`.

If you run this code:

```
x = 3
x + 2
x
```

the value of x at the end will be _____.

If you run this code:

```
x = 3
y = x
x = x + 1
y = y + 2
y
```

the value of y at the end will be _____.

Lists, Vectors, Arrays

_____ or _____ hold multiple values of the same type. _____ hold multiple values, possibly of different types.

Elements are stored in order, and elements can be referenced by their _____. The first element has _____ 0 or 1 depending on the language. The _____ of a list, vector, or array is the number of elements in it. An empty list has a _____ of 0.

You can _____ an item to the beginning of a list or vector or _____ an item to the end.

Sometimes, lists can be _____ inside other lists.

In R and Python, you can take a slice of a list (or R vector) using the list indices:

`my_list[a:b]`

Example:

`my_list[3:6]`

In Python, `a` is the index of the _____ value, `b` is the index of the _____ value EXCLUSIVE (meaning it's not included).

In R, the first number is the index of the _____ value, and the second number is the index of the _____ value INCLUSIVE.

Assigning Values

To change the value of an element in a list, assign a new value to it:

```
my_list = [7,6,5,4]
my_list[2] = 3
my_list
```

`my_list` now contains _____.

If instead you assign a new value as:

```
my_list = [7,6,5,4]
my_list = [1,2,3]
my_list
```

`my_list` now contains _____.

Conditions

[**True or False**] When using variables with boolean values in a conditional statement, you should explicitly compare them to True or False to determine their value.

The operator to test for equality is _____.

Is the following [**True or False**]

```
(TRUE and FALSE) or (not FALSE and TRUE)
```

Flow Control

If statements determine what to do based on a condition that evaluates to [**a single/multiple**] True or False value(s).

	A	B	C	D
1	X			
2				
3		>	X	
4				X

Figure 1:

Where will the > in cell B3 of Figure 1 above end up if you execute the following statements? The point of the > indicates the direction that is "forward" facing.

```
rotate left
if space ahead of you is occupied
  move one space to your left
  rotate right
else
  move one space forward
move one space forward
if you are in column B
  move one space to your left
else if you are facing up
  move one space forward
else
  move one space backwards
```

For Loops

Loops are used to _____ the same code for _____ values.

The following code will print _____ numbers.

```
x = [1, 4, 3, 6, 7, 2]
for i in x
    if i < 5
        print i
```

Functions

The values you send to a function are called _____, while the variables that are defined in a function definition are called _____.

The _____ of non-keyword arguments must match the _____ of parameters in the function definition.

In Python, _____ arguments cannot come before _____ arguments. In R, it's more complicated!

You can [**always/never**] use the name of all of the parameters when calling a function.

It [**is/is not**] good practice to specify arguments in a function call in the order they appear in the function definition, regardless of whether you're using named/keyword parameters or not.

Paramenters without default values in a function definition are [**required/optional**].

It's [**OK/not OK**] to have variables in your script with the same name as function parameters.

The output of a function is called the _____ value.

Packages/Libraries

Packages/libraries/modules need to be _____ before using them. Some are built-in, while others need to be _____ first.

It's a [**good/bad**] idea to use packages written by other people.