**Class 17 Text Processing**

**Key Ideas**

* There are string methods that perform case conversions
* There are string methods that deal with spacing: removing spaces, and justifying strings
* There are string methods that join strings and split strings
* There are string methods that find occurrences of substrings within strings, and find/replace
* There are “is” string methods that ask questions and return True/False
* There is a **string** module that contains useful string constants, such as a string of all lowercase letters, a string of all upper case letters, a string of all letters, and a string of all digits

**Built-ins**

**Methods**

* **capitalize**: capitalizes only the first word in a string
* **title**: capitalizes all words in a string
* **swapcase**: converts upper to lower and lower to upper
* **strip**: deletes both leading and trailing blanks
* **lstrip**: deletes only leading blanks
* **rstrip**: deletes only trailing blanks (on the right)
* **zfill**: specifies a field width, and pads with leading 0’s
* **center**: centers a string within a specified width
* **ljust**: left justifies a string within a specified width
* **rjust**: right justifies a string within a specified width
* **join**: concatenates all of the strings in an iterable; called with a string which is the separator
* **split**: splits a string into substrings, using a space as the default delimiter (although another delimiter can be specified) and returns a list of the substrings
* **rsplit**: just like split, except it splits from the right (which only matters if maxsplit is specified)
* **splitlines**: split lines in a string, using \n as the delimiter; returns a list
* **partition**: splits a string at the first occurrence of a specified separator and returns 3 things in a tuple: the left part, the separator, and the right part
* **rpartition**: same as partition but splits at the last occurrence of the specified separator
* **rindex**: returns the index of the last occurrence of the beginning of a substring within a string
* **find**: returns the index of the first occurrence of a substring within a string but if the substring is not found, returns -1
* **rfind**: returns the index of the last occurrence of a substring within a string but if the substring is not found, returns -1
* **replace**: finds occurrences of a substring within a string and replaces them with another string (by default, all, but a count can be passed)
* **isalpha**: returns True if all characters in a string are letters of the alphabet
* **isalnum**: returns True if all characters are letters or numbers
* **isdigit**: returns True if all characters in a string are digits

**String Module Constants**

* + **ascii\_lowercase**: a string containing all of the lower case letters of the alphabet
  + **ascii\_uppercase**: a string containing all of the upper case letters of the alphabet
  + **ascii\_letters**: a string containing all letters of the alphabet
  + **digits**: a string containing all of the digits

**Assessment Question**

Which of the following are true? Check all that apply.

The title method concatenates a ‘!’ to the end of a string.

**The rstrip method deletes only trailing blanks from a string**

All string methods that manipulate strings change the value of the string variable

**There are methods that left-justify and right-justify strings**