1. The following are the names and functions of string object types in Python 3.X

* **<stirng>.isdecimal() ->** Returns True if all characters in a string are decimal.
* **<string>.isalnum() ->** Returns True if all characters in the string are AlphaNumeric.
* **<string>.istitle() ->** Returns True if first character in a string is in Uppercase.
* **<string>.partition(<sub\_string>) ->** Splits string at first occurance of sub string and returns a tuple of 3 elements.
* **<string>.rpartition(<sub\_string>) ->** Splits string at last occurance of sub string and returns a tuple of 3 elements.
* **<string>.isidentifier() ->** Returns True if give string is a valid identifier name.
* **len(<string>) ->** Returns the length of the given string.
* **<string>.index(<sub\_string>) ->** Returns the lowest index of substring if substring is found in the string.
* **<string>.rindex(<sub\_string>) ->** Returns the highest index of substring if substring is found in the string.
* **max(<string>) ->** Returns the highest Alphabetical Character in the string as per ASCII.
* **min(<string>) ->** Returns the lowest Alphabetical Character in the string as per ASCII.
* **<string>.splitlines() ->** Returns a list of lines in the string.
* **<string>.capitalize() ->** Returns the string with first character capitalized.
* **<string>.upper() ->** Returns the string with all characters in uppercase.
* **<string>.lower() ->** Returns the string with all characters in lowercase
* **<string>.casefold() ->** Returns the string in lowercase which can be used for caseless comparisions.
* **<string>.expandtabs(no\_of\_spaces) ->** Replaces tabs in a string with specified no of spaces default is 8
* **<string>.find(<sub\_string>) ->** Returns lowest index of substring if substring is found in the string else returns -1.
* **<string>.rfind(<sub\_string>) ->** Returns highest index of substring if substring is found in the string else returns -1.
* **<string>.count(<char>) ->** Returns the no of occurances of the char in the given string.
* **<string>.split(<sep>) ->** Returns list of words seperated by given sep else seperated by whitespace.
* **<string>.rsplit(<sep>) ->** Returns list of words seperated by given sep else seperated by whitespace scanning from end.
* **<string>.lstrip() ->** Returns a copy of where leading whitespaces are removed.
* **<string>.rstrip() ->** Returns a copy of where trailed whitespaces are removed.
* **<string>.strip() ->** Returns a copy of where both leading and trailing whitespaces are removed.
* **<string>.swapcase() ->** Swaps lowercase characters with uppercase and vice versa.
* **<sep>.join(<list>) ->** Concatenates a list or tuple of words with intervening occuernces of sep.
* **<string>.translate(<mapping\_table>) ->** translates the characters using table.
* **<string>.maketrans(<dict>) ->** Creating a mapping translation tbale usable for **<string>.translate(<mapping\_table>)**
* **<string>.replace(<char\_1>,<char\_2>) ->** Replace all occurances of char\_1 with char\_2 in string.
* **<string>.encode() ->** Encodes string into any encoding supported by python. Default encoding is **UTF-8**.
* **<string>.ljust(<no\_of\_spaces>) ->** Left-justify in a field of given width.
* **<string>.rjust(<no\_of\_spaces>) ->** Right-justify in a field of given width.
* **<string>.center(<no\_of\_spaces>) ->** Center-justify in a field of given width.
* **<stirng>.zfill(<length>) ->** Zfill adds zeros to the begining of string until the specified length is reached.

2. In Python3 default format of strings is Unicode Whereas in Pyton2 we need to explicitly mention Unicode value using u.

3. In Python 3.x **unidecode()** method from unidecode library can be used to put non-ASCII Unicode Characters in a string.

Example:

**from** unidecode **import** unidecode

print(unidecode(u'ko\u017eu\u0161\u010dek'))

print(unidecode("आप नीचे अपनी भाषा और इनपुट उपकरण चुनें और लिखना आरंभ करें"))

print(unidecode("谢谢你"))

print(unidecode("ありがとう。"))

print(unidecode("улыбаться Владимир Путин"))

4. The major difference between these two is that a text file contains textual information in the form of alphabets, digits and special characters or symbols. On the other hand, a binary file contains bytes or a compiled version of a text file.

When a file is opened in **text mode**, reading its data automatically decodes its content (as per the platform default or as per provided encoding), and returns it as a **str**; writing operation takes a **str**, and automatically encodes it before transferring to the file. Text mode files also support universal end-of-line translation, and encoding specification arguments.

When a file is opened in **binary mode** by adding a **b** to the mode string argument in the open() call, reading its data does not decode it in any way, and simply returns its content raw and unchanged, as a bytes object; writing takes a bytes object and transfers it to the file unchanged. Binary-mode files also accept a bytearray object for the content to be written to the file.

5. Use of **encode()** and **decode()** method can be used to you interpret a Unicode text file containing text encoded in a different encoding than your platform's default, by default encoding parameter is **UTF-8.**

6. Use str.encode() and file.write() to make a Unicode text file in a particular encoding format, default encoding format is UTF-18.

* Call str.encode(encoding) with encoding set to utf8 to encode str.
* Call open(file, mode) to open a file with mode set to wb . wb writes to files in binary mode & preserves UTF-8format.
* Call file.write(data) to write data to the file.

Example:

unicode\_text = u"улыбаться Владимир Путин"

encoded\_unicode\_text = unicode\_text.encode("utf-8")

print(encoded\_unicode\_text,end="\n\n")

w\_file = open("w\_file.txt","wb")

w\_file.write(encoded\_unicode\_text)

w\_file.close()

r\_file = open("w\_file.txt","rb")

data = r\_file.read()

print(data)

7. **Unicode** represents most written languages in the world. **ASCII** has its equivalent in Unicode. The difference between ASCII and Unicode is that ASCII represents lowercase letters (a-z), uppercase letters (A-Z), digits (0–9) and symbols such as punctuation marks while Unicode represents letters of English, Arabic, Greek etc. mathematical symbols, historical scripts, emoji covering a wide range of characters than ASCII.

8. Python 3 stores strings as Unicode by default whereas Python 2 requires you to mark a string with a **u** if you want to store it as Unicode. Unicode strings are more versatile than ASCII strings, which are the Python 3.X default, as they can store letters from foreign languages as well as emoji and the standard Roman letters and numerals.