Joyce Woznica  
IST 652, Week 7

Activity 7.3: Unicode and Python

Code:

﻿# Week 7

# Joyce Woznica

﻿# Unicode

15

0xFF

0x7F

'\u0394'

"\U00000394"

"\N{GREEK CAPITAL LETTER DELTA}"

b'\xFFab'

type(b'ab')

type(b'\xFFab')

len(b'\xFFab')

u=chr(40960) + 'abcd' + chr(1972)

type(u)

u

print(u)

u.encode('ascii','ignore')

u.encode('ascii','replace')

u.encode('ascii', 'xmlcharrefreplace')

u.encode('ascii', 'backslashreplace')

Output:

﻿﻿15

Out[86]: 15

0xFF

Out[87]: 255

0x7F

Out[88]: 127

'\u0394'

Out[89]: 'Δ'

"\U00000394"

Out[90]: 'Δ'

"\N{GREEK CAPITAL LETTER DELTA}"

Out[91]: 'Δ'

b'\xFFab'

Out[92]: b'\xffab'

type(b'ab')

Out[93]: bytes

type(b'\xFFab')

Out[94]: bytes

len(b'\xFFab')

Out[95]: 3

u=chr(40960) + 'abcd' + chr(1972)

type(u)

Out[97]: str

u

Out[98]: 'ꀀabcd\u07b4'

print(u)

ꀀabcd޴

u.encode('ascii','ignore')

Out[100]: b'abcd'

u.encode('ascii','replace')

Out[101]: b'?abcd?'

u.encode('ascii', 'xmlcharrefreplace')

Out[102]: b'&#40960;abcd&#1972;'

u.encode('ascii', 'backslashreplace')

Out[103]: b'\\ua000abcd\\u07b4'