

NSI 06

```
1 0 0 0 0 1 0 0 0 1 0 1 1 1 1 1
0 0 0 1 1 0 0 0 1 1 1 1 0 0
1 1 0 1 1 0 1 0 0 1 1 0 0 1 0
1 1 0 1 1 1 0 1 0 1 1 0 1 0 1
0 1 1 0 1 1 0 1 1 0 0 0 0 0 1 1
0 1 0 0 1 1 1 1 1 0 1 1 1 1 0
1 0 0 1 1 0 0 1 1 0 0 1 0 0 1 0
1 0 0 1 0 1 0 1 0 1 0 1 1 1 0 1 0 0
1 1 1 1 1 1 1 0 0 1 0 0 1 0 0 0
0 1 1 1 1 0 0 1 1 1 0 0 1 1 1 1
0 1 1 1 1 1 1 1 0 0 1 0 0 1 0 1
0 0 1 1 1 1 1 1 0 0 1 0 1 1 1 1
1 1 1 1 0 1 1 1 1 1 1 0 0 0 1 0
0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0
0 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0
1 0 1 0 0 0 0 0 1 0 1 1 1 1 0 1
```

REPRÉSENTATION D'UN TEXTE

Connaître différentes méthodes d'encodage des caractères

Activité 1: À la découverte de la table « à ski »...avec M. Padchance !

Regarde la vidéo de M. Padchance #1

<https://drive.google.com/open?id=1XTVIn4e6TUtj5SXtkOh3jma9Vhm6NRSv>

Partie I

Un élève malin propose à ses camarades d'appuyer sur la touche Alt et tout en la laissant appuyée de taper le code 065

- 1) Dans un endroit où l'on peut écrire du texte (barre de commande, barre d'adresse, traitement de texte) testez le conseil de l'élève malin. Que se produit-il ?
- 2) Utiliser un moteur de recherche et en tapant dans la barre de recherche le nom de la touche, le code et ce que vous avez obtenu, dire à quoi correspond 065 pour le caractère obtenu.
- 3) Proposer alors un moyen d'écrire les autres lettres manquantes du clavier.
- 4) On sait comment se passer des lettres manquantes du moment qu'elles sont en majuscule, que faire pour écrire ces mêmes lettres en minuscule ?

Partie II

Tester maintenant dans la console de Python les instructions suivantes :

```
>>> ord('A')           >>> chr(65)
>>> ord('a')           >>> chr(90)
>>> ord('B')           >>> chr(123)
>>> ord(67)            >>> chr('A')
>>> ord(a)              >>> chr(A)
```

Partie III Le code de César.

- 1) A l'aide d'un moteur de recherche découvrez et lisez l'histoire derrière « Le chiffre de César »

On veut créer en python un programme permettant de traduire un texte en utilisant le système du chiffre de César. Vu que c'est un peu complexe, on va commencer par se limiter à des lettres majuscules.

- 2) Si le chiffre de César est 7 en quoi sont transformés les lettres A, D et N ?
- 3) Comment obtenir ces réponses avec une seule ligne de code (on utilisera consécutivement les deux nouvelles fonctions découvertes dans la partie 2)
- 4) En utilisant la commande de la question précédente, toujours avec 7 comme chiffre de César qu'obtient-on pour les lettres W, X et Z ?

- 5) En admettant un tel résultat créer un programme simple qui après vous avoir demandé votre chiffre de César va afficher la version codée d'un texte que vous aurez mis en mémoire, par exemple : «PADCHANCE PLEURE». Pour vérifier votre programme sachez qu'un tel texte décalé de 5 donnera : « UFIHMFESHJ%UQJZWJ »

On veut maintenant faire quelque chose de propre (on ne veut pas obtenir autre chose que des majuscules et des espaces)

- 6) En supposant que la lettre immédiatement après Z soit un « A » en quoi sont transformés W, X et Z ?
- 7) Créer la fonction `texteVnombres(txt)` qui pour un texte « txt » renvoie la liste des codes des lettres du texte
- 8) Créer la fonction `nombresVtexte(liste)` qui pour une liste de codes renvoie le texte correspondant.
- 9) On veut créer la fonction `decale(nombre,chiffreCesar)` qui pour tout « nombre » lui associe sa valeur une fois qu'on l'a décalé de la clé de codage « chiffreCesar », attention celle-ci peut être positive comme négative (quand on a envie de décoder le message rendu illisible). Pour cela il faudra faire attention au cas où le nouveau code n'est plus entre 65 et 90 à cause du fait que le code initial correspondait à un espace ou que le décalage nous fasse sortir de l'intervalle. On pourra proposer trois conditions dans le code, une si le code vaut 32 (on gardera 32), une si on est sous 65 et une après 90 (et à chaque fois corriger de manière adéquate la valeur obtenue)
- 10) A l'aide de ces trois fonctions créer le programme attendu pour convertir un texte avec le chiffre de César.

(PADCHANCE PLEURE devient avec un décalage de 5 : UFIHMFESHJ UQJZWJ)

- 11) Trouve le moyen de généraliser ton programme à un texte contenant aussi des minuscules (non accentuées)
 - a. Version simple on peut tout convertir en majuscule avant de coder (avec la fonction `upper`)
 - b. Version plus compliquée : on respecte le choix minuscule et majuscules de l'utilisateur.

- 12) Pourquoi avoir précisé non accentuées ?

- 13) Regardez la vidéo de monsieur padchance :

https://drive.google.com/open?id=1nBU1rP3tbQBSP2zG5_8rbsT4zft4cig1 puis essayer de répondre à la question posée.

Activité 2: Le ski, ce n'est qu'en hiver !

- À l'aide de la table ASCII, code cette phrase « Un âne éternue » en langage binaire et décimal, que remarques-tu ? Conclue sur l'utilisation de la table ASCII pour coder du texte en français ou en allemand.
- Au final, combien de caractères peut-on encoder grâce à la table ASCII ? En déduire le nombre de bits utilisés par cette méthode d'encodage.
- D'autres encodages sont progressivement apparus pour répondre à l'expansion d'internet à travers le monde et donc aux échanges de documents entre différents pays et donc différents caractères. On trouve par exemple, l'encodage ISO 8859-1 ainsi que l'ISO 8859-15 mais également l'UTF-8, chacun apportant des évolutions.

Sur Notepad++ , ouvre le document *phrase*

Change l'encodage comme indiqué dans le tableau ci dessous avec l'aide du document « **Ressource03 – Notepad ++** » et réponds aux questions posées :

| ENCODE en | Question |
|---|--|
| ANSI | Indique si avec cet encodage le « é » s'affiche correctement Y a t il d'autres caractères erronés ? |
| 8859-15 | Quels sont les changements ? |
| Convertir en UTF8 | Quels sont les changements ? |
| Ouvre http://xinhuanet.com (Ctrl U pour le code source) | Quel encodage est choisi pour afficher les caractères de cette page ? |
| Ouvre http://russian.news.cn (Ctrl U pour le code source) | Quel encodage est choisi pour afficher les caractères de cette page ? Que peux-tu en conclure ? |

4. Ces encodages diffèrent de l'ASCII en ajoutant (au moins) un bit supplémentaire. Combien le fait d'avoir 8 bits amène-t-il de nouvelles possibilités ?

5. Quelques précisions sur l'UTF-8

L'encodage UTF-8 utilise 1, 2, 3 ou 4 octets en respectant certaines règles :

- Un texte en ASCII de base (appelé aussi US-ASCII) est codé de manière identique en UTF-8. On utilise un octet commençant par 0.

| Caractère | Point de code (hexadécimal) | Valeur scalaire | | Codage UTF-8 |
|-----------|-----------------------------|-----------------|---------|--------------|
| | | décimal | binaire | binaire |
| A | U+0041 | 65 | 1000001 | 01000001 |

- Les octets ne sont pas remplis entièrement.
- Les bits de poids fort du premier octet forment une suite de 1 indiquant le nombre d'octets utilisés pour coder le caractère. Les octets suivants commencent tous par 10.
- Dans la norme ISO 8859-1 le « é » est codé 1110 1001, en UTF-8 on le codera **11000011 10101001**. On pourra remarquer deux choses : le codage ISO s'inscrit dans le codage UTF-8, on a comblé le premier octet avec des zéros (en italique).

Définition du nombre d'octets utilisés

| Représentation binaire UTF-8 | Signification |
|-------------------------------------|------------------------------|
| 0xxxxxxx | 1 octet codant 1 à 7 bits |
| 110xxxxx 10xxxxxx | 2 octets codant 8 à 11 bits |
| 1110xxxx 10xxxxxx 10xxxxxx | 3 octets codant 12 à 16 bits |
| 11110xxx 10xxxxxx 10xxxxxx 10xxxxxx | 4 octets codant 17 à 21 bits |

Activité 3 : Les impacts de l'encodage....

1. Ouvre python, puis trouve astucieusement le code qui va te permettre de découvrir quel est son encodage par défaut. Ecris-le au verso de ta feuille.

2. Vérifie en saisissant :

```
import sys
print(sys.getdefaultencoding())
```

3. Dans les deux extraits de code source de pages web ci-dessous, identifie, en les encerclant, dans quel encodage chacune s'affiche, et grâce à quel balise/attribut.

```
2 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
3
4 <html xmlns="http://www.w3.org/1999/xhtml">
5
6 <head>
7 <meta http-equiv="Content-Type" content="text/html; utf-8" />
8 <title>Home</title>
9 <link rel="stylesheet" type="text/css" href="styl.css" />
```

```
<!DOCTYPE html>
<html>
  <head>
    <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-15" />
    <title>iso-8859-15 Encoded Page</title>
  </head>
```

Regarde la vidéo: M. Padchance #4

<https://drive.google.com/open?id=1OuwJqFu3vf2LruGfNanqrvQEs6S-7lu8>

4. Es-tu surpris du résultat ? Explique ce qui s'est passé

5. Soit deux listes d'éléments :

Liste1 = ['a', 'Z', 'Zoo', '4', 'fin'] et

Liste2 = ['34', 'motlongenminuscule', 'MOTTRESPETIT', '?', '0', 'motenminuscule']

En t'aidant du document ressource approprié, **écris tes pronostics** sur le résultat du tri de chacune de ces listes.

Puis à l'aide d'un script Python, trie les éléments de ces listes, comme M. Padchance, et vérifie tes hypothèses. *Indice : utilise la méthode sort() de la manière suivante : liste.sort()*