

## I) *Latin1 vs Utf8*

1) On veut écrire **sur deux lignes** les 3 mots ci-dessous :

Le Père

Noël

a) Peut-on coder ces caractères en Ascii ? en Iso8859-1 ? en Utf8 ?

b) A l'aide uniquement de la table Iso8859-1 du cours, encoder les caractères de ces 2 lignes en nombres hexadécimaux (sans omettre de coder le saut de ligne).

c) Avec le convertisseur en ligne <https://www.rapidtables.com/convert/number/hex-to-ascii.html>, choisir comme encodage « ISO-8859-1 » et saisir les nombres hexadécimaux trouvés ci-dessus. Le texte affiché correspond-il à ce que l'on souhaite ?

d) Ouvrez l'application « Geany » de votre chromebook, écrivez « Le Père Noël » sur deux lignes, enregistrez ce texte dans un fichier et réouvrez ce fichier avec l'éditeur hexadécimal : <https://hexed.it>. Comparer les nombres hexadécimaux affichés par hexedit avec ceux trouvés tout à l'heure. L'application Geany a utilisé quel encodage ? Comment a été encodé le saut de ligne ?

2) Le premier mot de la Bible est « **בראשית** ». En hébreu, cela veut dire « au commencement » et cela s'écrit de la droite vers la gauche avec les lettres beth, resh, aleph, shin, yud et tav.

a) Peut-on coder ces caractères en ascii ? en iso8859-1 ? en utf8 ?

b) Chercher sur <https://utf8-chartable.de> par quels nombres hexadécimaux coder ces caractères en utf-8 ?

c) Recopier ces nombres dans le convertisseur en ligne rapidtables ci-dessus. En choisissant le bon encodage, est-ce que « **בראשית** » est affiché correctement ?

## II) *Point de code Unicode vs encodage Utf8*

1) En tapant dans la console python : `help(chr)` puis `help(ord)`, déterminer ce que font les fonctions chr() et ord(). En déduire les points de code Unicode des trois caractères suivants : @ ù £

2) Écrire un script Python qui affiche les 50 premières lignes de 16 caractères de la table Unicode. Pourquoi y a-t-il des « trous » dans certaines des lignes de 16 caractères ?

3) En grec, le mot « eau » se dit « **νερό** ». A l'aide de la console python, déterminer les points de code Unicode de ces 4 lettres grecques **en binaire**.

4) En vous appuyant sur votre cours, en déduire les encodages de ces lettres en UTF8.

5) Recopier ces nombres binaires dans le convertisseur en ligne rapidtables ci-dessus. En choisissant le bon encodage, est-ce que « **νερό** » est affiché correctement ?