

STI3 IPD

April 17, 2021

1 Devoir 1 Python for Data Science 2021

1.1 EXERCICE 1

Question 1. Commencer par importer pandas et numpy en utilisant la convention:

```
[ ]: # Ecrire le code ici
```

Question 2. Maintenant charger le fichier `input/automobile_data.csv` dans le **DataFrame** `cars_df` et afficher les 10 première lignes:

```
[ ]: # Ecrire le code ici
```

Question 3. Afficher les 8 dernières du **DataFrame**.

```
[ ]: # Ecrire le code ici
```

Question 4. Donner le nombre de lignes et de colonnes de ce Dataframe.

```
[ ]: # Ecrire le code ici
```

Question 5. Récupérez le nombre d'entrées non nulles par colonne. La colonne `price` contient-elle des valeurs nulles?

```
[ ]: # A faire: Afficher les informations du dataframe en utilisant .info()
```

Question 6. Sélectionnez maintenant toutes les lignes avec une valeur dans la colonne 'price', et stockons-les dans un nouveau DataFrame appelé `cars_with_price_df`.

Quel est le **shape** de ce nouveau DataFrame? Confirmez votre réponse en utilisant le résultat précédent.

```
[ ]: # Ecrire le code ici
```

Question 7. Combien de `constructeurs(make)` distincts existe-il dans le jeu de données?

```
[ ]: # Ecrire le code ici
```

Question 8. Affichez tous les constructeurs qui ont publié plus de 10 modèles. Enregistrez ces résultats dans une nouvelle variable `big_makers`.

```
[ ]: # Ecrire le code ici
```

Question 9. Récupérez les statistiques suivantes: - la voiture avec la puissance(horsepower) maximale - la voiture avec la longueur(length) minimale - la moyenne de la largeur(width) des voitures

```
[ ]: # Ecrire le code ici
```

1.2 EXERCICE 2:

Ecrivez une fonction appelée **isprime** qui prend un entier positif N et détermine s'il est premier ou non. Renvoie le N s'il est premier et rien si ce n'est pas le cas.

Ensuite, en utilisant une **list comprehension** et la fonction **isprime**, créez une liste **myprimes** qui contient tous les nombres premiers inférieurs à 100.

```
[ ]: # Ecrire le code ici
```

1.3 EXERCICE 3 : Vrai ou Faux

- Python est un langage compilé
- les listes sont modifiable contrairement aux tuples
- Data Science produit de la connaissance à partir des données
- DataFrame peut-être vu comme une table sur excel

```
[ ]:
```