Les Mains à la pâte

Se familiariser avec Python

- **Nouveau** : aujourd'hui, **on avance sans théorie juste en tâtonnant et en découvrant TOUT par la pratique (essais/erreurs)**

Les Mains à la pâte

Analyse du fichier arbres.csv : programmer en python

Comment récupérerait-on le contenu du fichier Arbres.csv dans R?

```
read_csv("G:/Arbres.csv", sep=";", dec=",")
```

Analyse du fichier arbres.csv : programmer en python

Comment récupérerait-on le contenu du fichier Arbres.csv dans R?

```
read_csv("G:/Arbres.csv", sep=";", dec=",")
```

Comment récupérer le contenu du fichier Arbres.csv avec Pandas?

```
pd.read_csv("G:/Arbres.csv", sep=";", decimal=",", encoding='latin1')
```

Pandas est une bibliothèque écrite pour le langage Python permettant la manipulation et l'analyse des données (dataframes).

Analyse du fichier arbres.csv : programmer en python

Comment récupérer le contenu du fichier arbres.csv avec Pandas?

```
pd.read_csv("C:/Users/marc/Desktop/DS5/Arbres.csv", sep=";", decimal=".", encoding='latin1')
```

Pandas est une bibliothèque écrite pour le langage Python permettant la manipulation et l'analyse des données (dataframes).

On en sait assez pour commencer...

Analyse du fichier arbres.csv : programmer en python

Comment récupérer le contenu du fichier `arbres.csv` avec Pandas?

```
arbres=pd.read_csv("C:/Users/marc/Desktop/DS5/Arbres.csv",sep=";",  
decimal="," ,encoding='latin1')
```

Création du dataframe `arbres`
qui contient les données du fichier `Arbres.csv`

```
print(arbres.head(3))
```

```
print(arbres.describe())
```

Quels sont vos commentaires et autres remarques ?

Analyse du fichier arbres.csv : programme python

Création du dataframe `arbres` qui contient les données de `Arbres.csv`

```
arbres=pd.read_csv("C:/Users/marc/Desktop/DS5/Arbres.csv",sep=";",  
decimal="," ,encoding='latin1')
```

Premier jet de statistique descriptive

```
print(arbres.head(3))
```

```
arbres.describe()
```

```
print(arbres.groupby('essence').describe())
```

```
print(arbres.groupby('essence').hauteur.describe())
```

```
print(arbres.boxplot(column="hauteur",by="essence"))
```

Quels sont vos commentaires et autres remarques ?

Analyse du fichier arbres.csv : programme python

```
arbres=pd.read_csv("C:/Users/marc/Desktop/DS5/Arbres.csv",sep=";",  
decimal="," ,encoding='latin1')
```

Du côté des graphes (statistique descriptive, suite)

Les graphes les plus logiques

```
print(arbres.boxplot(column="hauteur",by="essence"))
```

```
histo3 = arbres.hauteur.hist()
```

```
histo2 = arbres.hauteur.hist(by=arbres.essence)
```

```
arbres.boxplot(column="hauteur",by="etat")
```

```
arbres.plot.scatter("hauteur","valeur")
```

Choisissez votre style!