

Travaux pratiques python

Préparation de l'environnement :

- Installer Anaconda Anaconda3-2024.02-1-Windows-x86_64.exe
- Utiliser Jupyter d'Anaconda pour faire les exercices.

Remarque :

Pour chaque notebook/exercice, expliquer via des commentaires les lignes de codes.

À la fin de chaque exercice :

- téléchargez le fichier jupyter et
- exporter fichier jupyter en fichier pdf ou html
- puis télécharger le sur OneDrive

Exercice 1 : diagramme à barres

Créer un graphique en barres pour comparer les pourboires entre Hommes et Femmes en utilisant barplot de seaborn.

1. Importer les bibliothèques de visualisation et de chargement des données (seaborn et pyplot de matplotlib)
2. Charger le jeu de données Tips depuis Seaborn.
3. Le jeu de données tips est de quel type ? justifier la réponse avec un code python ?
4. Afficher les cinq premières lignes de le jeu de données tips
5. Initialiser une figure avec une taille de 10x6 pouces.
6. Utiliser Seaborn pour tracer un graphique en barres de tips avec le Genre sur l'axe des X et pourboire sur l'axe des Y, et choisir une palette de couleurs 'Set2'
7. Définir le titre du graphique,
8. Définir le titre de l'axe des x,
9. Définir le titre de l'axe des y,
10. Afficher le graphique
11. Interpréter le graphique, qu'est-ce que vous voyez ?

Exercice 2 : diagramme de dispersion

1. Importer seaborn et pyplot de matplotlib
2. Charger le jeu de données tips depuis seaborn
3. Le jeu de données tips est de quel type ? justifier cela avec un code python ?
4. Afficher les cinq premières lignes de le jeu de données tips
5. Créer une figure avec une figsize de (4,10) pouces
6. Créer le graphique de nuage de point en spécifiant le jeu de données, la variable x et y
7. Définir le titre de graphique,
8. Définir le titre de l'axe des x,
9. Définir de l'axe des y
10. Forcer l'utilisation d'une grille sur la figure
11. Afficher le graphique
12. Interpréter le résultat ?
13. Changer les dimensions de la figure de (4,10) à (4,4)
14. Exécuter de nouveau le code, qu'est ce que vous remarquez ? est ce que le changement de dimensions de la figure peut rendre le graphique plus/moins lisible ?

Exercice 3 : Histogramme

1. Importez **Matplotlib** et **Seaborn** pour charger et visualiser les données.
2. Charger le jeu de données tips
3. Affichez les premières lignes des données pour vérifier la structure de le jeu de données tips
4. Créez un histogramme pour visualiser la distribution des pourboires tips
5. Spécifiez le nombre de *bins* (intervalle de regroupement des données) à utiliser pour l'histogramme.
6. Ajoutez un titre à votre histogramme décrivant ce qui est visualisé.
1. Ajoutez des titres pour l'axe des abscisses et l'axe des ordonnées
2. Ajoutez une grille pour améliorer la lisibilité du graphique.
3. afficher votre histogramme
4. Interpréter l'histogramme
5. Varier les valeurs de bins et exécuter de nouveau le code.
6. Est ce qu'il y a une Valeur de bins qui assure une lecture plus adéquate de l'histogramme.

Exercice 4 : boîtes à moustaches

1. Importer les bibliothèques pour la visualisation et le chargement des données :

2. Charger le jeu de données mpg de Seaborn
3. Le jeu de données mpg est de quel type ? justifier cela avec un code python ?
4. Afficher les cinq premières lignes de le jeu de données mpg
5. Initialiser une figure avec une taille de (10,6) pouces
6. Créer un box plot pour observer la consommation de carburant mpg en fonction de nombres de cylindres
7. Définir le titre de graphique,
8. Définir le titre de l'axe des x,
9. Définir de l'axe des y
10. Interpréter les résultats ? quelles sont les voitures les plus efficaces et les moins efficaces en consommation de carburant ?
11. Quelle relation entre la variable mpg et l'efficacité de voitures en termes de consommation de carburant?
12. Quelle relation entre le nombre de cylindres et l'efficacité de voitures en termes de consommation de carburant?

Exercice 5 : graphique à bulles

En se basant sur le jeu de données *Gapminder*, nous utiliserons le graphique à bulles interactif pour analyser la relation entre trois variables à savoir :

- PIB par habitant (**X**) : Représentant la richesse économique par personne.
- Espérance de vie (**Y**) : Indiquant la santé et le développement des pays.
- Population (**taille de la bulle**) : Montrant la taille relative de la population de chaque pays.

Pour faire cela, écrit le code répondant aux questions suivantes :

1. Importer la bibliothèque Plotly Express
2. Charger le jeu de données Gapminder depuis Plotly Express
3. Afficher les cinq dernières lignes du jeu de données
4. Le jeu de données Gapminder est de quel type ? justifier la réponse avec un code python ?
5. Afficher la structure de données du jeu de données

6. Afficher un résumé statistique du jeu de données
7. Filtrer le jeu de données pour avoir uniquement les données de 2007
8. Affecter les données de 2007 à une variable df2007
9. créer le graphique à bulles en tenant compte des paramètres suivants :

- df2007 : Sélectionner uniquement les données de l'année 2007 pour une analyse précise.
- x="gdpPercap" : L'axe des X représente le PIB par habitant.
- y="lifeExp" : L'axe des Y représente l'espérance de vie.
- size="pop" : La taille des bulles représente la population des pays.
- color="continent" : Les bulles sont colorées en fonction des continents pour une distinction géographique.
- hover_name="country" : Le nom du pays s'affiche lors du survol de la bulle pour des informations supplémentaires.
- log_x=True : L'échelle logarithmique de l'axe des X permet de mieux visualiser les différences de PIB par habitant.
- size_max=60 : Limite la taille maximale des bulles pour une meilleure lisibilité.

10. afficher le graphique à bulles

Exercice 6 : TAF Histogramme

1. Concevoir un histogramme de deux jeu de données représentant respectivement :
 - des notes d'une classe excellente
 - et des notes d'une autre classe Moyenne.
2. Choisissez une Valeur de bins adequate.
3. Comparer les deux graphiques et dire ce que vous voyez