

Devoirs à faire à la maison

Visualisation des données de transport à l'aide de Python

Module : **Recherche et analyse dans le domaine des transports**

Pratique 5 : **Visualisation des données dans la recherche sur les transports**

Date limite : Une semaine à compter de la date de publication du devoir.

Contents

Objectif	2
Résultats d'apprentissage	2
Tâches assignées	2
Critères d'évaluation	5
Directives pour la soumission	5

Objectif

Développer des compétences pratiques dans l'utilisation de Python (Google Colab) pour créer, interpréter et communiquer des visualisations de données sur les transports qui soutiennent des décisions de planification fondées sur des preuves.

Résultats d'apprentissage

À l'issue de ce travail, les élèves seront capables de :

- Charger et inspecter les données réelles sur la fréquentation des transports dans Python.
- Créer des graphiques linéaires, des histogrammes, des nuages de points et des cartes thermiques à l'aide de Colab.
- Interpréter les principaux modèles, tendances et relations dans les données.
- Communiquer des informations claires et étayées par des données pour la planification des transports.

Tâches assignées

Ensemble de données: `advanced_transport_ridership_practical.xlsx`

Partie 1: Exploration de l'ensemble de données

- Chargez l'ensemble de données dans Google Colab.
- Explorez les colonnes et les types de données.
- Vérifiez s'il y a des valeurs manquantes ou des incohérences.
- Rédigez un bref aperçu de l'ensemble de données (3 à 5 phrases).

Indice

Réfléchissez à la signification de chaque colonne et à l'exhaustivité de vos données. Posez-vous les questions suivantes: que représente chaque colonne? Y a-t-il des valeurs manquantes? Utilisez le code suivant:

```
import pandas as pd
df = pd.read_excel('advanced_transport_ridership_practical.xlsx')
df.info()
df.describe()
df.head()
```

Partie 2: Graphique linéaire

- Créez un graphique linéaire illustrant l'évolution du nombre d'usagers des bus, trains et métros au fil du temps.
- Ajoutez des titres, des étiquettes d'axe et une légende.
- Enregistrez le graphique et incluez-le dans votre rapport.

Indice

Posez-vous la question suivante: comment le nombre d'usagers évolue-t-il au fil du temps ? Recherchez les pics, les baisses ou les tendances saisonnières.

```
import matplotlib.pyplot as plt
df.plot(x='Date', y=['Bus_Ridership', 'Train_Ridership', '
    Metro_Ridership'])
plt.title('Ridership Over Time')
plt.xlabel('Date')
plt.ylabel('Number of Passengers')
plt.legend()
plt.show()
```

Partie 3: Histogramme

- Créez un histogramme pour une colonne de fréquentation (par exemple, Bus).
- Personnalisez le nombre de classes si nécessaire.

Indice

Réfléchissez: quelle est la fourchette la plus fréquente en termes de fréquentation ? Y a-t-il des valeurs extrêmes (valeurs aberrantes)?

```
plt.hist(df['Bus_Ridership'], bins=20, color='skyblue', edgecolor='black
    ')
plt.title('Bus Ridership Distribution')
plt.xlabel('Passengers')
plt.ylabel('Frequency')
plt.show()
```

Partie 4: Diagramme de dispersion

- Créez un nuage de points comparant la fréquentation des bus et des trains.
- Ajoutez des titres et des étiquettes d'axes.

Indice

Question: Existe-t-il une corrélation entre la fréquentation des bus et celle des trains? Une fréquentation plus élevée des bus signifie-t-elle une utilisation plus importante des trains?

```
plt.scatter(df['Bus_Ridership'], df['Train_Ridership'], alpha=0.6)
plt.title('Bus vs Train Ridership')
plt.xlabel('Bus Ridership')
plt.ylabel('Train Ridership')
plt.show()
```

Partie 5: Interprétation et réflexion

- Rédigez 5 à 7 phrases résumant vos conclusions.
- Réfléchissez au graphique qui vous a été le plus utile et expliquez pourquoi.

Indice

Réfléchissez aux tendances qui se sont démarquées et à la manière dont ces informations pourraient aider un planificateur ou un décideur politique.

Partie 6: Carte thermique

- Créer une carte thermique montrant les modèles d'utilisation par station ou type de jour.

Indice

Demandez : Quels sont les moments ou les stations les plus fréquentés? Utilisez ce code:

```
import seaborn as sns
pivot = df.pivot_table(values='Bus_Ridership', index='DayType', columns='Station')
sns.heatmap(pivot, annot=True, cmap='YlGnBu')
plt.title('Bus Ridership Heatmap')
plt.show()
```

Critères d'évaluation

Composant	Poids
Exploration des ensembles de données	15%
Graphique linéaire	20%
Histogramme	20%
Diagramme de dispersion	20%
Interprétation	15%
Carte thermique	10%

Directives pour la soumission

- Soumettez un cahier Colab (.ipynb) ou un fichier PDF exporté contenant tous les codes et résultats.
- Incluez vos réponses écrites et vos graphiques de manière claire sous chaque section.
- Nommez le fichier `VotreNom_Pratique5_Visualisation.pdf`.
- Téléchargez-le sur la plateforme ou envoyez-le par e-mail selon les instructions de votre professeur.