




 SOUTENANCE À PLANIFIER

Participez à un concours sur la Smart City

50%

MISSION COURS ÉVALUATION

 60 heures

Mis à jour le mardi 17 mars 2020

Vous êtes fraîchement établi en tant qu'expert indépendant spécialisé en intelligence artificielle. Vous participez régulièrement à des concours pour vous **faire la main sur de nouveaux sujets**.

Durant vos recherches, vous avez découvert l'ONG "**Data is for Good**", qui propose des challenges de Data Science en ligne sur des thématiques ayant trait au bien commun. Des associations et collectivités publiques sponsorisent ces challenges.



Start Collaborating in our Real-World Challenges



10+
Challenges



400+
Collaborators



60+
Countries

Upcoming Challenges

Connecting Land Conflicts with Policies

Degraded land is affecting billions of people. Help to build a solution together with the World Resources Institute to facilitate restoration efforts.

Challenge starts: 13 September



Vous avez décidé de participer à un challenge proposé par la ville de Paris ! Voici le cahier des charges du challenge, que vous avez trouvé sur le site :

Inscrivez-vous au challenge Data sponsorisé par la ville de Paris et **aidez Paris à devenir une smart-city !**

Dans ce challenge, ouvert à tous, vous allez réaliser une analyse exploratoire avec un jeu de données portant sur les arbres de la ville de Paris, dans le cadre du programme "Végéталisons la ville".

Vos résultats contribueront à une optimisation des tournées pour l'entretien des arbres de la ville. Eh oui, moins de tournées égale moins de trajets, et plus d'arbres entretenus.

Vous aurez ainsi un impact réel sur le futur de la ville de Paris !

Données :

Téléchargez le jeu de données des arbres de la ville de Paris à [cette adresse](#). Vous pouvez aussi le consulter dans son contexte sur opendata.paris.fr.

Règles du challenge :

- Si ce n'est pas fait, installez votre environnement de développement sur votre ordinateur, et créez un environnement virtuel dédié à ce challenge.
- Les données doivent être explorées à l'aide de Python et de ses librairies.
- Vous soumettrez votre analyse sous forme de présentation, contenant les informations suivantes :
 - Présentation générale du jeu de données
 - Démarche méthodologique d'analyse de données
 - Synthèse de l'analyse de données
- Le second livrable prendra la forme d'un Notebook Jupyter. Le Notebook sera documenté pour expliciter les différents traitements, calculs ou graphiques que vous effectuez en utilisant les fonctionnalités d'édition de texte de Jupyter. Vos explications doivent permettre à un public non technique de comprendre les différentes étapes de votre analyse et votre synthèse.

Merci pour votre inscription, et bon courage pour ce challenge ! Que le meilleur gagne !

Dans ce projet, vous devez réaliser une analyse exploratoire en utilisant Python. Attention, cette tâche peut prendre beaucoup de temps ! Soyez vigilant sur le temps passé sur ce projet, en gardant un oeil sur la durée estimée du projet, et les compétences évaluées.

Livrables

- Un **Jupyter Notebook** comportant votre exploration du jeu de données.
- Un **support de présentation de type Power Point** (aussi utilisé pour la soutenance du projet).
 - Ce document comportera trois parties : la présentation générale du jeu de données, la démarche méthodologique d'analyse de données, la synthèse de votre analyse de données.

Pour faciliter votre passage au jury, déposez sur la plateforme, dans un dossier nommé "*P2_nom_prenom*", tous les livrables du projet. Chaque livrable doit être nommé avec le numéro du projet et selon l'ordre dans lequel il apparaît, par exemple "*P2_01_notebook*", "*P2_02_présentation*", et ainsi de suite.


Soutenance

La soutenance se déroulera en visioconférence et durera 30 minutes.

Les 20 premières minutes sont consacrées à la présentation de votre analyse, que vous souhaitez soumettre pour le challenge en ligne.

Suivront ensuite 10 minutes de questions.

Référentiel d'évaluation

 Compétences	 Critères d'évaluation
Utiliser des bibliothèques python	L'analyse de données est complète si : <input type="checkbox"/> l'environnement de développement est installé et fonctionnel sur

pour réaliser
une analyse de
données
exploratoires

le poste de travail (Python et Jupyter)

- ❑ un environnement virtuel a été créé pour assurer l'isolement du projet et la gestion des dépendances
- ❑ les librairies python spécialisées ont été importées dans le Jupyter Notebook
- ❑ le jeu de données a été décrit brièvement (nombre de lignes, nombre de colonnes et nombre de valeurs manquantes) en ayant chargé le fichier plat dans un dataframe.

L'analyse de données est **pertinente** si :

- ❑ des indicateurs statistiques basiques (moyenne et écart-type) ont été calculés pour les différentes colonnes
- ❑ les ordres de grandeur des grandeurs statistiques des différentes colonnes ont été comparés

L'analyse de données est **présentable** si :

- ❑ les fonctionnalités d'édition de cellule Markdown du Jupyter Notebook sont utilisées dans au moins trois cellules pour commenter

l'analyse et la mettre en forme (titres, mise en forme, alternance de cellule d'exécution de code python et de cellule de texte explicatif)

- ❑ les titres des trois parties (la présentation générale du jeu de données, la démarche méthodologique d'analyse de données, la synthèse de votre analyse de données) sont visuellement en évidence dans le Jupyter Notebook (par exemple en gras et police de caractère plus grande que dans le reste du document)

Effectuer une
analyse
statistique
univariée

L'analyse statistique univariée est **complète** si :

- ❑ les moyennes, médianes et quantiles des distributions sont calculées
- ❑ au moins une représentation graphique d'une distribution statistique a été tracée

L'analyse statistique univariée est **pertinente** si :

- ❑ les éventuelles valeurs aberrantes ont été identifiées (avec une définition des valeurs aberrantes basée sur un multiple des quantiles de la distribution).

L'analyse statistique univariée est **présentable** si :

☐ les graphiques sont lisibles (titres alignés, légende présente, noms des abscisses et des ordonnées précisés)

Compétences évaluées



Effectuer une analyse statistique univariée



Utiliser des librairies python pour réaliser une analyse de données exploratoire

OPENCLASSROOMS



ENTREPRISES



CONTACT



EN PLUS



 Français 



