

Computer Vision with Python

Nennouche Mohamed

Rekrouk Maroua

Séance 2

Sommaire

- Présentation des objectifs de la séance
- Qu'est ce que l'analyse des données ?
- Qu'est ce que la visualisation des données ?
- Intérêt pour un data scientist ?
- Environnement de travail
- Introduction aux outils d'analyse et de visualisation de données
 - NumPy
 - Matplotlib
 - Pandas
 - Seaborn
 - Projet analyse des données : Titanic
 - Projet visualisation des données : Iris
- Projet à proposer

Objectif de la séance

1. Faire les premiers pas dans la data science
2. Avoir une idée sur les outils d'analyse et de visualisation des données avec Python
3. Comprends le flot de travail pour aborder un problème d'analyse de données
4. Aborder de vraies données

Analyse et visualisation des données

Analyse des données

Méthodes mathématiques (essentiellement géométriques) permettant d'extraire les relations et les informations à partir des données :

- Corrélation
- Analyse en composantes principales
- Analyse factorielle
- ect...

Visualisation des données

ensemble de méthodes permettant de résumer de manière graphique de données statistiques qualitatives et surtout quantitatives afin de montrer les liens entre des ensembles de ces données :

- Graphique à barres
- Histogramme
- Graphique à secteurs
- Distribution en plan ou en espace
- ect...

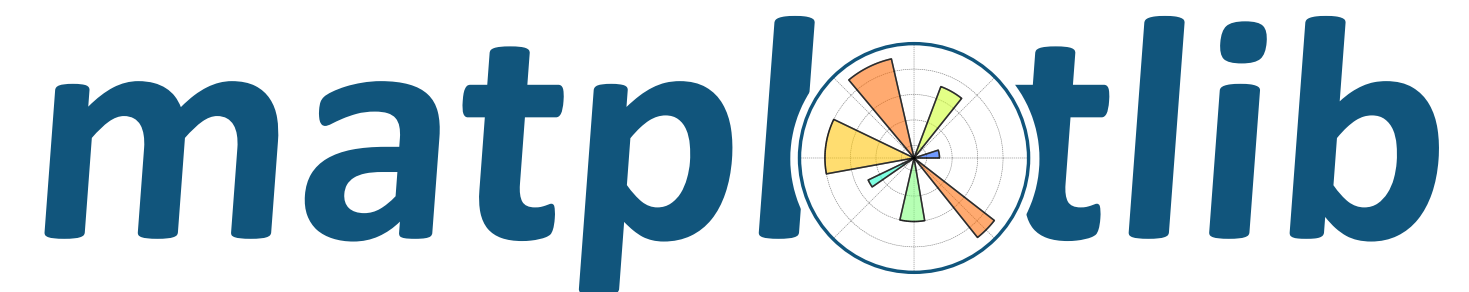
Quelques exemples

Intérêt pour un data scientist

L'analyse et la visualisation des données est utile pour un data scientist pour :

- La compréhension des données qu'il a
- La compréhension des relations entre les différentes caractéristiques des données étudiées
- Analyse de la qualité des données étudiées.

Librairies utilisées en Python



Analyse et visualisation des données avec Python

Prochaine séance

On verra la prochaine séance les bases du Machine Learning en étudiant les modèles les plus utilisés ainsi que son implémentation

- Les bases du Machine Learning
- Théorie derrière les principaux modèles de Machine Learning : KNN, SVM, Random Forest, Naïve Bayes
- Flot de travail pour traiter un problème de Machine Learning
- Evaluation des performances atteintes

Merci et à la prochaine