

Projet (4h)

Plan de la séance :

1. (1h) Mise en contexte et discussions ;
2. (2h) Préparation du projet en équipes ;
3. (1h) Présentation de solutions et discussions.

Contexte :

C'est un fait bien connu que les « Millenials¹ » aiment les toasts à l'avocat. C'est aussi un fait bien connu que tous les « Millenials » vivent dans le sous-sol de leurs parents.

De toute évidence, ils n'achètent pas de maison parce qu'ils achètent trop de toasts à l'avocat ! Mais peut-être qu'il y a de l'espoir, si un « Millenial » pouvait trouver une ville avec des avocats pas chers, il pourrait vivre le rêve américain « Millenial » et avoir sa propre maison.

Données (vous ne disposez que de ce résumé):

- **avocado_survey.scv** – 18.2k lignes de données obtenues de 2000 à 2021 sur :
 1. **Date** - la date de l'observation ;
 2. **Ville** – le nom de la ville ;
 3. **Shop** – le nom du vendeur ;
 4. **Type** – le type du fruit : conventionnel ou biologique ;
 5. **TotalVolume** - nombre total d'avocats vendus par ce vendeur à la date de l'observation ;
 6. **AveragePrice** - le prix moyen d'un seul avocat en dollars.
- **house_prices.csv** – 1460 lignes de données obtenues en 2021 sur :
 1. **Date** - la date de l'observation ;
 2. **Ville** - la ville ;
 3. **Surface** - surface en m² ;
 4. **Constr_year** - l'année de la construction ;
 5. **Neighbourhood** – le type du quartier : très aisé, aisé, peu aisé ;
 6. **Price** - le prix de la maison, en dollars.

1 Adultes nés entre la fin des années 1980 et le milieu des années 1990.

Objectifs :

1. **(Analyse de données)** Trouvez la ville où, en 2021, un « Millenial » peut :
 - 1) Avoir son toast à l'avocat le moins cher ;
 - 2) Acheter la maison la moins cher avec un crédit intégral ;
 - 3) 1) et 2) ensemble.
2. **(Machine Learning)** On suppose que le revenu d'un Millenial lui permet juste d'acheter un avocat par jour au prix moyen dans son lieu de résidence. Prévoyez le changement de prix de l'avocat selon l'historique des prix, et calculez combien d'années il devrait se priver de toasts à l'avocat payer pour sa maison s'il l'a achetée en 2021 entièrement à crédit.

Faites une présentation qui contient votre plan d'action pour résoudre ces problèmes-là en utilisant des méthodes étudiées pendant nos séances. Votre réponse peut se structurer autour de réponses détaillées aux questions suivantes :

1. Comment pouvez-vous trouver la ville "optimale" demandée dans l'objectif 1 en vous basant sur les derniers prix en 2021 des bases de données ?
 - Quel est le problème principal posé en termes d'Analyse de données ?
 - Comment allez-vous préparer les données ?
 - Quels types de caractéristiques de données allez-vous utiliser ?
2. Pour l'objectif 2, comment pouvez-vous prédire le prix de l'avocat dans la ville choisie pour les prochaines années (à partir de 2022) ?
 - Quel est le problème principal posé en termes de Machine Learning ?
 - Quels types de caractéristiques de données allez-vous utiliser ?
 - Quelle famille de méthodes pouvez-vous appliquer ? Quelle méthode ?