

# Projet guidé: Analyse de data

Sondage Thanksgiving

# Introduction au dataset

- Utiliser la fonction `pandas.read_csv()` pour lire le fichier `thanksgiving.csv`.
  - Spécifier dans les paramètres de la fonction l'argument `encoding='latin-1'` car ce dataset n'est pas encodé normalement
  - Assigner le dataframe à la variable `data`
- Afficher les premières lignes de data.
- Afficher le nom des colonnes avec l'attribut `columns`.

# Filtrer des données

- Utiliser la méthode `Series.value_counts()` pour afficher le décompte du nombre de réponses pour chaque catégorie de la colonne 'Do you celebrate Thanksgiving?'.
- Filtrer et garder toutes les lignes du dataframe `data` pour lesquelles la réponse à la question 'Do you celebrate Thanksgiving?' est 'Yes'.
- Assigner ce nouveau dataframe à `data` et afficher le.

# Exploration des repas de thanksgiving

- Utiliser la méthode `Series.value_counts()` pour afficher combien de fois chaque résultat apparaît pour la question et colonne 'What is typically the main dish at your Thanksgiving dinner?'.
- Afficher la colonne 'Do you typically have gravy?' pour les lignes du dataframe `data` pour lesquelles la colonne 'What is typically the main dish at your Thanksgiving dinner?' vaut 'Tofurkey' pour la dinde de tofu.
  - Créer un filtre qui sélectionne seulement les lignes du dataframe `data` où 'What is typically the main dish at your Thanksgiving dinner?' vaut 'Tofurkey'
  - Assigner le résultat à la variable `tofurkey`
  - Afficher la colonne 'Do you typically have gravy?' de ce nouveau dataframe `tofurkey`

# Tendances des desserts pour Thanksgiving

- Créer un objet Series indiquant avec des booléens les valeurs de la colonne 'Which type of pie is typically served at your Thanksgiving dinner? Please select all that apply. - Apple' qui sont nulles. Assigner le résultat à la variable `apple_isnull`.
- Créer un objet Series indiquant avec des booléens les valeurs de la colonne 'Which type of pie is typically served at your Thanksgiving dinner? Please select all that apply. - Pumpkin' qui sont nulles. Assigner le résultat à la variable `pumpkin_isnull`.
- Créer un objet Series indiquant avec des booléens les valeurs de la colonne 'Which type of pie is typically served at your Thanksgiving dinner? Please select all that apply. - Pecan' qui sont nulles. Assigner le résultat à la variable `pecan_isnull`.
- Combiner ces 3 objets Series avec l'opérateur `&` et assigner le résultat à la variable `pies`.
- Afficher les valeurs uniques et combien de fois elles apparaissent dans la colonne de `pies`.