

# NSI 1ère - Tables en csv - Introduction

QK

January 16, 2019

## 2. Rechercher et nettoyer dans une table

### Présentation

Le fichier *Winners.csv* contient les données des différents lauréats des prix nobel de 2001 à 2014.

Elles sont triées par année et prix (par exemple : 1901 - **Chemestry**) et présentent de nombreuses colonnes (*Winner, known for, field, gender, etc., Country represented for Nobel Prize*)

- Objectifs Simplifier la table
- Resultat voulu : un fichier *nobels.csv* avec comme colonnes :  
année, prix, nom, genre, pays 1901, chimie, “Jacobus Henricus van ’t Hoff”, Male, Netherlands

### Correction

```
import csv

csv_input = "Winners.csv" # fichier csv délimité par ","
listecsv = [] # stockera la liste d'entree
listeclean = [] # stockera la liste finale
csv_output = "nobels.csv" # fichier csv délimité par ","

# on lit toutes les lignes
with open(csv_input, newline='') as csvfile:
    winners = csv.reader(csvfile, delimiter=",")
    for row in winners:
        listecsv.append(row)
```

```

# les années et types de prix ne sont pas decoupés, nous devons le faire
# exemple de decoupage avec split
# print(listecsv[:2])
mots = "1901 - Chemistry".split(" ")
print(mots)

# on cree la premiere ligne
listeclean.append(["annee", "prix", "nom", "genre", "pays"])

# on ajoute chaque ligne
for row in listecsv[1:]:
    row1 = row[1].split(" ")
    listeclean.append( [row1[0], row1[2], row[0], row[4], row[12]] )

# on regarde la premiere ligne et une ligne choisie aléatoirement
print(listeclean[1:2])

import random
n = random.randint(1,len(listeclean))
print(listeclean[n : n+1])

with open(csv_output, 'w', newline='') as csvfile:
    csvwriter = csv.writer(csvfile, delimiter=',',
        quoting=csv.QUOTE_MINIMAL)
    for row in listeclean:
        csvwriter.writerow(row)

```