

## ***Développement des programmes avec SparkSql***

**Exercice:** moyenne de nombre d'amis (friends) par âge

**Exercice:** fréquences de mots avec dataFrames

### **Fonctions SQL**

- from pyspark.sql import functions as func
- func.explode() – comme flatmap, éclater les colonnes en lignes
- func.split()
- func.lower()
- Passer les colonnes en paramètres:
  - func.split(inputDF.value, \\w+)
  - filter(wordsDF.word, != "")
  - C'est possible d'utiliser func.col("culmnName") to refer to a column

**Exercice:** MIN-MAX Temperature **avec dataFrames**