

# Python File IO

- Download the SSA data for baby names in US for years 1880-2104 from <http://www.ssa.gov/oact/babynames/limits.html> (choose the national data set)
- Unzip the file onto your hard drive;
- Open a sample txt file and review the contents

Read the following Python program. What is its function?

```
import os
path = '/Users/serge/projects/SchoolNova/IT102/python/babynames/names'
data = {} #Dictionary
searchName=input("Search for name: ")
for dirEntry in os.listdir(path):
    dirEntryPath = os.path.join(path, dirEntry)
    if os.path.isfile(dirEntryPath) and dirEntryPath.endswith('.txt'):
        with open(dirEntryPath, 'r') as namesFile:
            data[dirEntry[3:7]] = namesFile.read()
for (y, babynames) in data.items():
    allNames = babynames.split("\n")
    for line in allNames:
        nameData = line.split(",")
        if nameData[0] == searchName:
            print(nameData[2] + " children named " + nameData[0] + " in " + y)
```

## Class and Home Exercise

1. Can you change the program such that it sums up the number of children named with the input name since 1880?
2. Can you change the program such that it displays the year(s) when the input name was most popular?
3. Can you change the program such that it shows name similar to the input name? For example, if the input is “Maria”, the program will show “Mariam”, “Marianne” etc.
4. Use Python language reference (<https://docs.python.org/3/reference/index.html>), your favorite search engine and StackOverflow.com to find solutions.