

Ejercicio 1

1. Crea un rdd llamado nameRDD a partir de la siguiente lista de nombres:
 - Namelist = ["Hugo","Erick","Biel","Antonio","Manuel","Francisco",
"Eduardo","Daniel","Juan","Lucía","María","Martina","Sofía","Emma","Julia",
"Daniela","Carla","Alma","Olivia","Vega","Lola","Valentina"]
2. Crea un rdd llamado anotherRDD a partir de nameRDD cuya salida sea:
 - Nombre+ "2nd" -> ["Hugo2nd","Erick2nd","Biel2nd"...."Valentina2nd"]
3. Crea un pair rdd llamado pairRDD a partir de nameRDD usando un map cuya salida sea:
 - (Nombre, 20) -> [("Hugo",20),("Erick",20),("Biel",20)....("Valentina",20)]

Importaciones y creacion se SparkSession y SparkContext

```
In [4]: import findspark
findspark.init()

import pandas as pd
import pyspark
from pyspark.sql import SparkSession

spark = SparkSession.builder.getOrCreate()
sc = spark.sparkContext
```

Punto #1

Crea un rdd llamado nameRDD a partir de la siguiente lista de nombres:

- Namelist = ["Hugo","Erick","Biel","Antonio","Manuel","Francisco",
"Eduardo","Daniel","Juan","Lucía","María","Martina","Sofía","Emma","Julia",
"Daniela","Carla","Alma","Olivia","Vega","Lola","Valentina"]

```
In [10]: Namelist = ["Hugo","Erick","Biel","Antonio","Manuel","Francisco", "Eduardo","Daniel",
nameRDD = sc.parallelize(Namelist)
```

Punto #2

Crea un rdd llamado anotherRDD a partir de nameRDD cuya salida sea:

- Nombre+ "2nd" -> ["Hugo2nd","Erick2nd","Biel2nd"...."Valentina2nd"]

```
In [7]: anotherRDD = nameRDD.map(lambda name: name + "2nd")
anotherRDD.collect()
```

```
Out[7]: ['Hugo2nd',
        'Erick2nd',
        'Biel2nd',
        'Antonio2nd',
        'Manuel2nd',
        'Francisco2nd',
        'Eduardo2nd',
        'Daniel2nd',
        'Juan2nd',
        'Lucía2nd',
        'María2nd',
        'Martina2nd',
        'Sofía2nd',
        'Emma2nd',
        'Julia2nd',
        'Daniela2nd',
        'Carla2nd',
        'Alma2nd',
        'Olivia2nd',
        'Vega2nd',
        'Lola2nd',
        'Valentina2nd']
```

Punto # 3

Crea un pair rdd llamado pairRDD a partir de nameRDD usando un map cuya salida sea:

- (Nombre, 20) -> [("Hugo",20),("Erick",20),("Biel",20),....("Valentina",20)]

```
In [9]: pairRDD = nameRDD.map(lambda name: (name, 20))
pairRDD.collect()
```

```
Out[9]: [('Hugo', 20),
        ('Erick', 20),
        ('Biel', 20),
        ('Antonio', 20),
        ('Manuel', 20),
        ('Francisco', 20),
        ('Eduardo', 20),
        ('Daniel', 20),
        ('Juan', 20),
        ('Lucía', 20),
        ('María', 20),
        ('Martina', 20),
        ('Sofía', 20),
        ('Emma', 20),
        ('Julia', 20),
        ('Daniela', 20),
        ('Carla', 20),
        ('Alma', 20),
        ('Olivia', 20),
        ('Vega', 20),
        ('Lola', 20),
        ('Valentina', 20)]
```