



Profundiza más

Recurso de Profundización

Ejemplos en Python de la implementación de MapReduce.

Ejemplo 1.

Contador de palabras: este *script* cuenta la frecuencia de cada palabra en un conjunto de texto.

```
from mrjob.job import MRJob

class WordCount(MRJob):
    def mapper(self, _, line):
        for word in line.split():
            yield word.lower(), 1

    def reducer(self, word, counts):
        yield word, sum(counts)

if __name__ == "__main__":
    WordCount.run()
```

Explicación

- **Mapper:** divide el texto en palabras y emite cada una con el valor 1.
- **Reducer:** suma las ocurrencias de cada palabra.

Ejemplo de entrada: hola mundo hola data science

Ejemplo de salida:

```
"hola" 2
"mundo" 1
"data" 1
"science" 1
```

Ejemplo 2.

Promedio de calificaciones por estudiante: este *script* calcula el promedio de calificaciones por estudiante.



Profundiza más

```
from mrjob.job import MRJob

class AverageGrade(MRJob):
    def mapper(self, _, line):
        student, grade = line.split(',')
        yield student, float(grade)

    def reducer(self, student, grades):
        grades_list = list(grades)
        yield student, sum(grades_list) / len(grades_list)

if __name__ == "__main__":
    AverageGrade.run()
```

Explicación

- **Mapper:** divide la línea en estudiante, calificación y emite el estudiante con su calificación.
- **Reducer:** calcula el promedio de calificaciones por estudiante.

Ejemplo de entrada:

Ana,85 Carlos,90 Ana,95 Carlos,80

```
"Ana" 90.0
"Carlos" 85.0
```