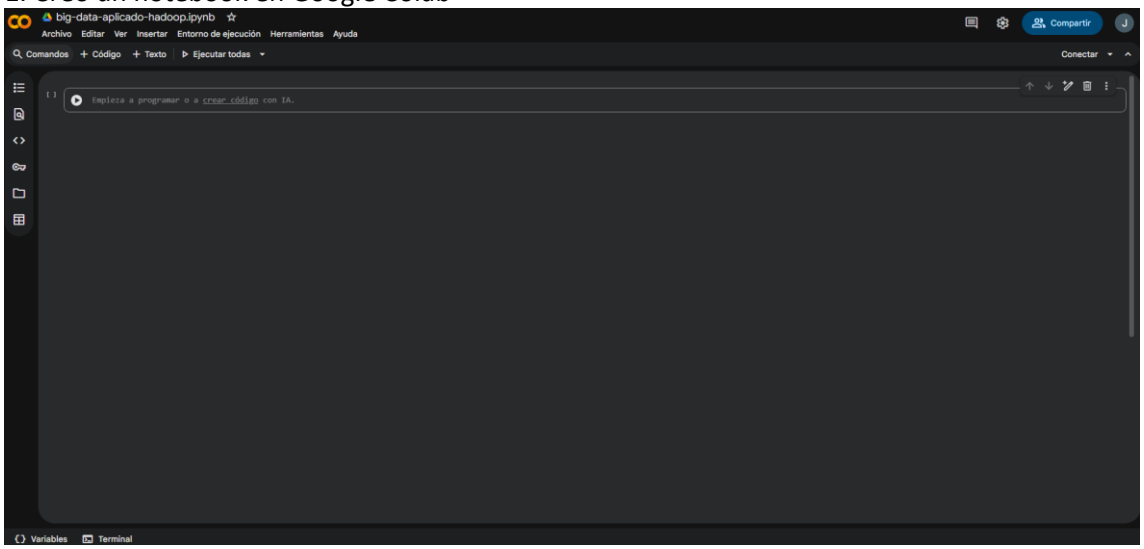
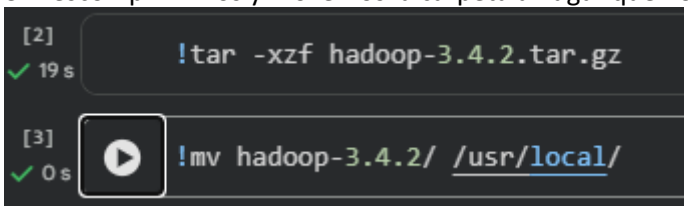
Solución tarea 2.1

1. Creo un notebook en Google Colab



2. Instalo hadoop: Hemos descargado la versión 3.4.2 que es la más reciente

3. Descomprimimos y movemos la carpeta al lugar que nos indican



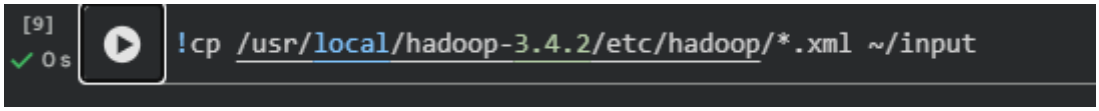
4. Configuramos el Hadoop JAVA HOME

- Buscamos la dirección de Java en la máquina Google Colab y establecemos el valor de la variable

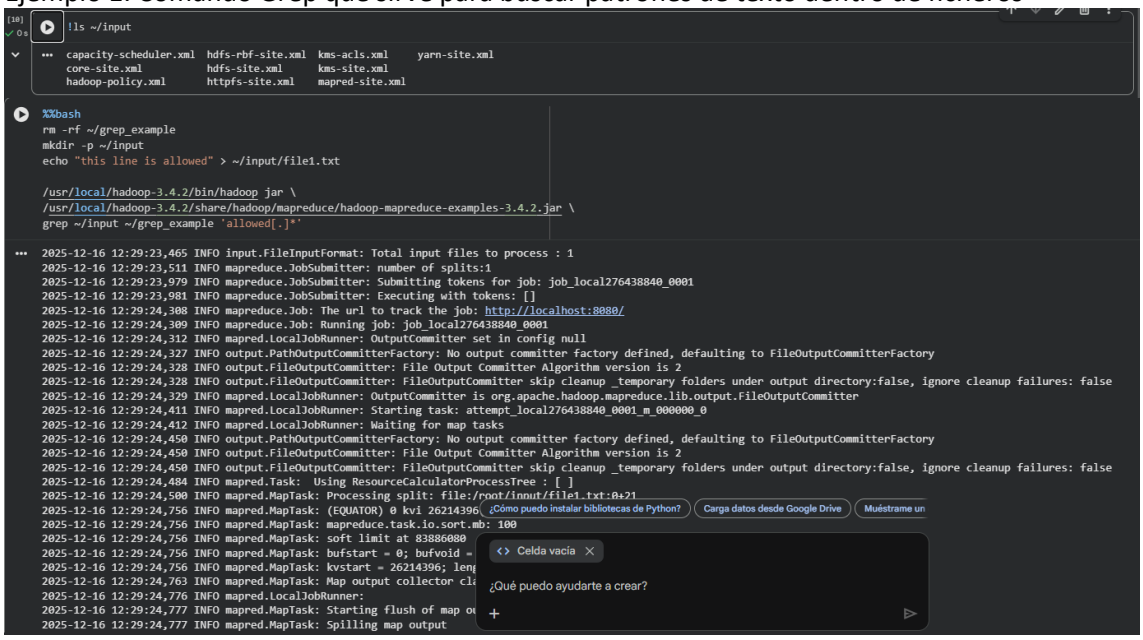
5. Ejecutamos Hadoop

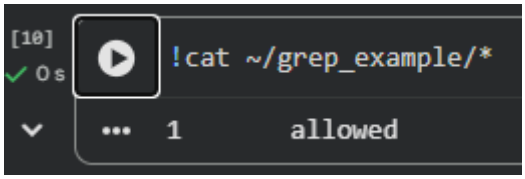
- Comandos de prueba

- Jar de ejemplos

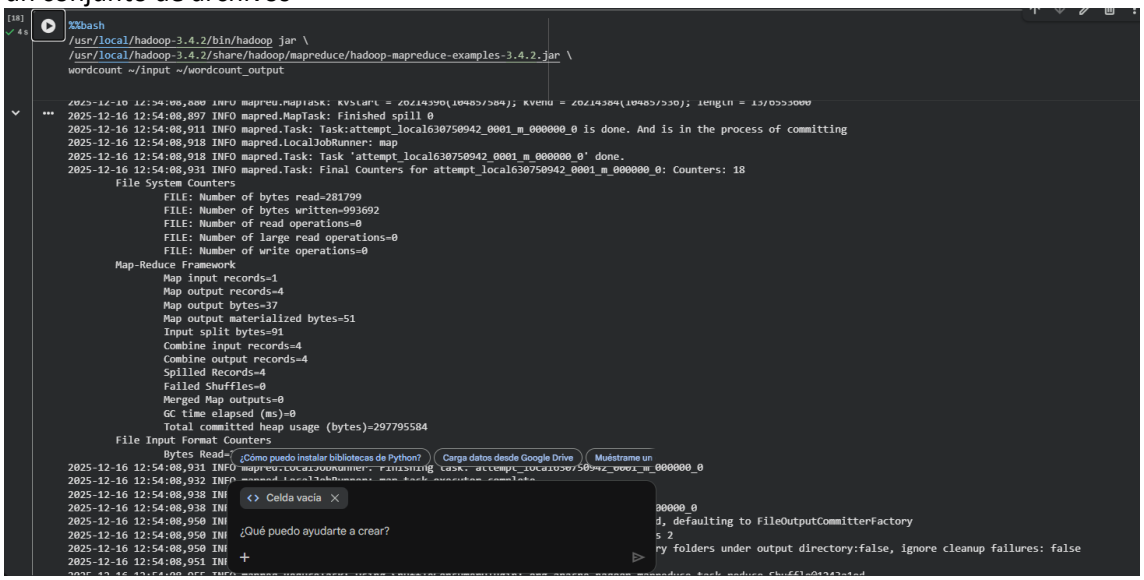


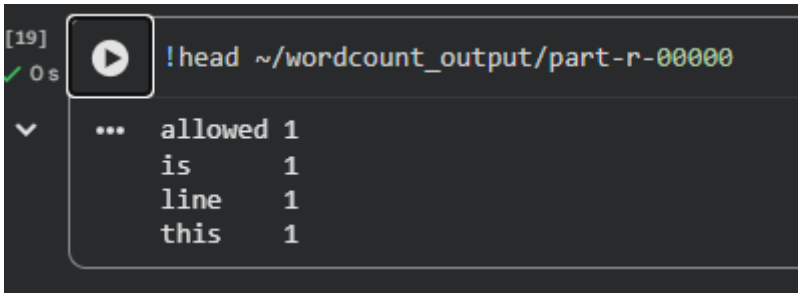
Ejemplo 1: Comando Grep que sirve para buscar patrones de texto dentro de ficheros



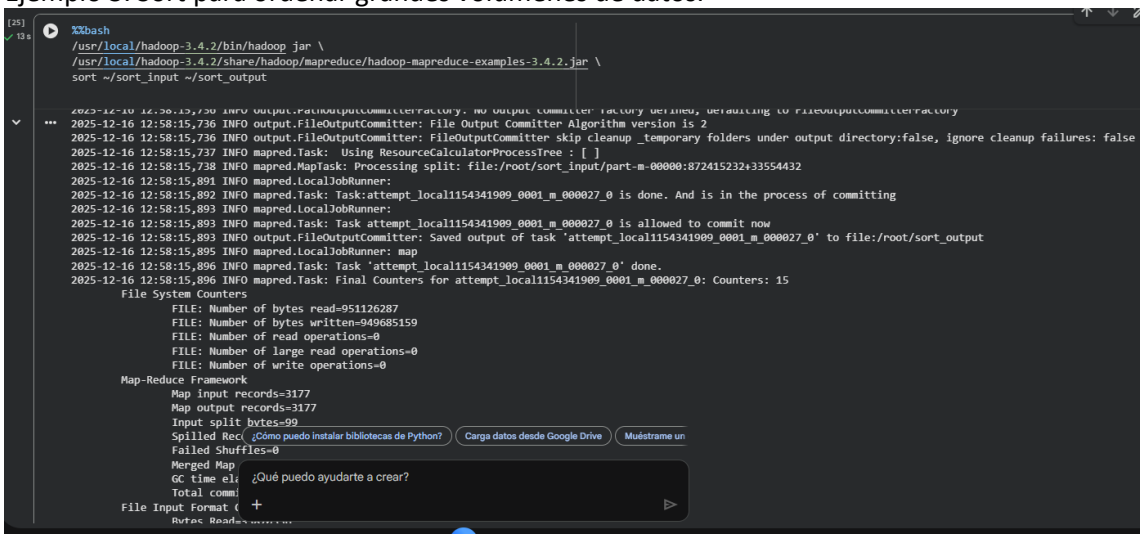


Ejemplo 2: WordCount que se utiliza para contar el número de apariciones de cada palabra en un conjunto de archivos

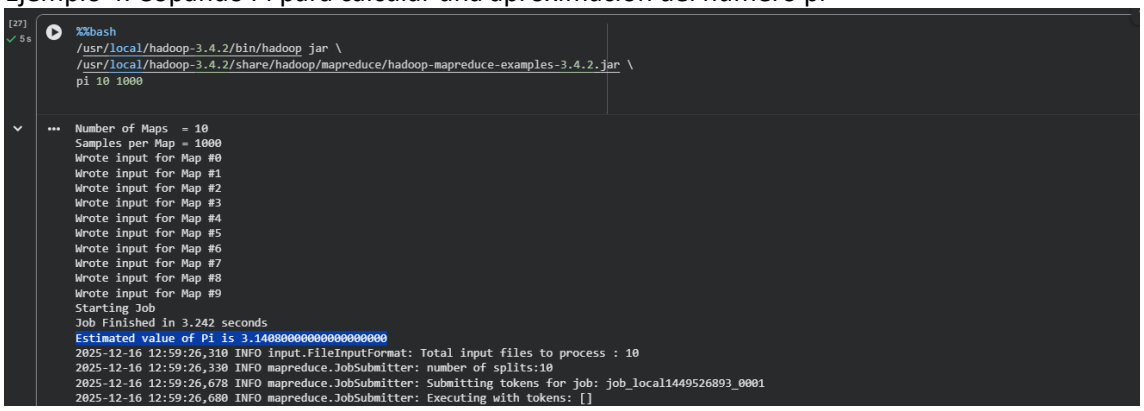




Ejemplo 3: Sort para ordenar grandes volúmenes de datos.



Ejemplo 4: Copando Pi para calcular una aproximación del número pi



```
[27]
✓ 5 s
%%bash
/usr/local/hadoop-3.4.2/bin/hadoop jar \
/usr/local/hadoop-3.4.2/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.4.2.jar \
pi 10 1000

Number of Maps  = 10
Samples per Map = 1000
Wrote input for Map #0
Wrote input for Map #1
Wrote input for Map #2
Wrote input for Map #3
Wrote input for Map #4
Wrote input for Map #5
Wrote input for Map #6
Wrote input for Map #7
Wrote input for Map #8
Wrote input for Map #9
Starting Job
Job Finished in 3.242 seconds
Estimated value of Pi is 3.14080000000000000000
2025-12-16 12:59:26,310 INFO input.FileInputFormat: Total input files to process : 10
2025-12-16 12:59:26,330 INFO mapreduce.JobSubmitter: number of splits:10
2025-12-16 12:59:26,678 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1449526893_0001
2025-12-16 12:59:26,680 INFO mapreduce.JobSubmitter: Executing with tokens: []
```