

### **Ejercicio 1: Contar Palabras**

Objetivo: Contar la cantidad de veces que aparece cada palabra en un fichero.

Descripción:

1. Crea un archivo de texto con varias líneas de texto.
2. Escribe un programa MapReduce que cuente las palabras en el archivo.

Hola mundo, este es un ejercicio.

Las palabras en este texto son varias.

El objetivo es contar cuántas veces aparece cada palabra.

¡Espero que esto sea útil!

Programación y análisis de datos son temas interesantes.

### **Ejercicio 2: Encuentra el Máximo**

Objetivo: Encontrar el número máximo en un fichero donde cada línea es un número.

Descripción:

1. Crea un archivo con una lista de números, uno por línea.
2. Escribe un programa MapReduce para encontrar el número máximo.

12

7

45

23

89

### **Ejercicio 3: Contar las Ocurrencias de Cada Letra**

Objetivo: Contar la frecuencia de cada letra en un conjunto de documentos.

Descripción:

1. Crea un archivo de texto con varias líneas de texto.
2. Escribe un programa MapReduce que cuente las letras en el archivo, ignorando los espacios y caracteres especiales.

Hola mundo

Esto es una prueba

Contando letras

El análisis es divertido

MapReduce es una técnica poderosa

### **Ejercicio 4: Contar líneas de texto**

Objetivo: Contar cuántas líneas hay en un archivo.

Descripción:

1. Crea un archivo de texto con varias líneas.
2. Escribe un programa MapReduce que cuente el número de líneas en el archivo.

Primera línea de texto

Segunda línea de texto

Tercera línea de texto

Cuarta línea de texto

Quinta línea de texto

### **Ejercicio 5: Análisis de Sentimientos**

Objetivo: Realizar un análisis de sentimientos básico contando palabras positivas y negativas.

Descripción:

1. Crea un archivo de texto que contenga reseñas de productos o servicios.
2. Define listas de palabras positivas y negativas.
3. Escribe un programa MapReduce que cuente el número de palabras positivas y negativas en las reseñas.

Este producto es excelente y maravilloso.

No me gustó, fue una experiencia horrible.

Me siento satisfecho con el servicio.

El producto es malo, no lo recomendaría.

Fue un día fantástico, estoy muy contento.

### **Ejercicio 6: Detección de Spam en Correos Electrónicos**

Objetivo: Clasificar correos electrónicos como spam o no spam basándose en la frecuencia de palabras.

Descripción:

1. Crea un archivo con varios correos electrónicos, cada uno en una línea, donde cada línea contiene el texto del correo.
2. Define un conjunto de palabras que comúnmente aparecen en los correos electrónicos de spam.
3. Escribe un programa MapReduce que cuente cuántas palabras de spam hay en cada correo.

¡Gana dinero rápido con nuestra oferta exclusiva!

Tu cuenta ha sido seleccionada para recibir una promoción.

Hola, ¿quieres saber más sobre nuestros servicios?

Este mensaje es sobre un producto increíble que es gratis.

No te pierdas esta oferta limitada, actúa ahora.

### **Ejercicio 7: Análisis de Tiempos de Respuesta**

Objetivo: Calcular el tiempo promedio de respuesta para diferentes endpoints de una API.

Descripción:

1. Crea un archivo de texto con registros de tiempos de respuesta de solicitudes, donde cada línea contenga la URL y el tiempo (por ejemplo, /api/v1/endpoint1,200).
2. Escribe un programa MapReduce que calcule el tiempo promedio para cada endpoint.

```
/api/v1/endpoint1,200
/api/v1/endpoint2,350
/api/v1/endpoint1,150
/api/v1/endpoint3,400
/api/v1/endpoint2,300
```

### **Ejercicio 8: Análisis de Datos de Ventas**

Objetivo: Calcular las ventas totales y promedio por producto.

Descripción:

1. Crea un archivo de texto donde cada línea contenga el nombre del producto y la cantidad vendida (por ejemplo, productoA,100).
2. Escribe un programa MapReduce que calcule la cantidad total y el promedio de ventas por producto.

```
productoA,100
productoB,200
productoA,150
productoC,300
productoB,250
```

### **Ejercicio 9: Análisis de Clima**

Objetivo: Calcular la temperatura promedio por ciudad.

Descripción:

1. Crea un archivo de texto donde cada línea contenga la ciudad, la fecha y la temperatura (por ejemplo, CiudadA,2024-01-01,15).
2. Escribe un programa MapReduce que calcule la temperatura promedio por ciudad.

```
CiudadA,2024-01-01,15
CiudadB,2024-01-01,20
CiudadA,2024-01-02,17
CiudadB,2024-01-02,19
CiudadC,2024-01-01,10
```

### **Ejercicio 10: Análisis de Datos de Encuestas**

Objetivo: Calcular la media y la desviación estándar de respuestas a preguntas de encuestas.

Descripción:

1. Crea un archivo donde cada línea contenga un identificador de encuesta y la respuesta (por ejemplo, encuesta1,4).
2. Escribe un programa MapReduce que calcule la media y la desviación estándar de las respuestas por encuesta.

encuesta1,4  
encuesta1,5  
encuesta2,3  
encuesta2,4  
encuesta1,5

### **Ejercicio 11: Análisis de Ventas por Región**

Objetivo: Calcular las ventas totales y promedio por región.

Descripción:

1. Crea un archivo donde cada línea contenga la región y el monto de la venta (por ejemplo, Norte,200).
2. Escribe un programa MapReduce que calcule las ventas totales y el promedio por región.

Norte,200  
Sur,150  
Norte,300  
Este,250  
Oeste,400

### **Ejercicio 12: Análisis de Datos de Clientes**

Objetivo: Calcular el total de compras y el número de compras por cliente.

Descripción:

1. Crea un archivo donde cada línea contenga un identificador de cliente y el monto de la compra (por ejemplo, cliente1,200).
2. Escribe un programa MapReduce que calcule el total de compras y el número de compras por cliente.

cliente1,200  
cliente2,150  
cliente1,300  
cliente3,400  
cliente2,250

### **Ejercicio 13: Clasificación de Artículos por Etiquetas**

Objetivo: Contar cuántos artículos tienen cada etiqueta.

Descripción:

1. Crea un archivo donde cada línea contenga un identificador de artículo y sus etiquetas (por ejemplo, articulo1,etiqueta1,etiqueta2).
2. Escribe un programa MapReduce que cuente el número de artículos para cada etiqueta.

```
articulo1,etiqueta1,etiqueta2  
articulo2,etiqueta2,etiqueta3  
articulo3,etiqueta1  
articulo4,etiqueta3,etiqueta4  
articulo5,etiqueta1,etiqueta2
```

#### **Ejercicio 14: Análisis de Datos de Tráfico de Red**

Objetivo: Contar el número de paquetes por dirección IP.

Descripción:

1. Crea un archivo de texto donde cada línea contenga una dirección IP y el tamaño del paquete (por ejemplo, 192.168.1.1,500).
2. Escribe un programa MapReduce que cuente el número total de paquetes recibidos por cada dirección IP.

```
192.168.1.1,500  
192.168.1.2,300  
192.168.1.1,250  
192.168.1.3,400  
192.168.1.2,150
```

#### **Ejercicio 15: Análisis de Comentarios de Productos**

Objetivo: Calcular la puntuación promedio de los comentarios para cada producto.

Descripción:

1. Crea un archivo donde cada línea contenga el identificador del producto y la puntuación del comentario (por ejemplo, producto1,4).
2. Escribe un programa MapReduce que calcule la puntuación promedio de los comentarios por producto.

```
producto1,4  
producto2,5  
producto1,3  
producto3,4  
producto2,2
```

#### **Ejercicio 16: Análisis de Conexiones de Redes Sociales**

Objetivo: Contar las conexiones de cada usuario en una red social.

Descripción:

1. Crea un archivo de texto donde cada línea contenga un usuario y sus conexiones (por ejemplo, usuario1,usuario2,usuario3).
2. Escribe un programa MapReduce que cuente cuántas conexiones tiene cada usuario.

usuario1,usuario2,usuario3  
usuario3,usuario5,usuario2