

Reconstituting the Datagrams Dataset

- Split files are more easily transferred. In the destination, reconstruct the original file:

1. download the .zip files
2. test the integrity of .zip files:

```
date; zip -T datag-p1.zip; date  
zip -T datag-p2.zip  
zip -T datag-p3.zip  
(No errors should occur)
```



Validación de integridad de datos

3. Decompress files:

```
date; unzip datag-p1.zip; unzip datag-p2.zip; unzip datag-p3.zip; date
```

Descomprimir (extraer)
cada uno de los archivos
<comando1> ; <comando2>;

4. Reconstruct the original:

```
cat datag-p1 datag-p2 datag-p3 > datagrams.zip
```

Unir las tres partes de
archivos en uno solo.

Ejecución secuencial

5. Test integrity and decompress the original file:

```
date; zip -T datagrams.zip; date  
date; unzip datagrams.zip; date
```

Validar la integridad del nuevo archivo (union de las tres parte

Descomprimir (extraer) el contenido del nuevo archiv

Comandos

Navegación

cd – Cambio de directorio
pwd – Directorio actual
Ls – Directorio actual

Gestión de Archivos

mkdir: Crear directorio.
mv: Mover o renombrar archivos.
cp: Copiar archivos.
rm: Eliminar archivos y directorios.

Compresión y Descompresión

zip: Comprimir archivos.
unzip: Descomprimir archivos.

Espacio

du : Uso de disco
df : espacio libre en disco

wc : contador de palabras

date; comando; date;

Descargar archivos

wget -O "ARCHIVO" "RUTA_WEB"
--show-progress

Pruebas (Test -T) de Integridad
zip -T "archive.zip"

Descomprimir/Comprimir

date; zip "files"; date;
Date; unzip "files"; date;

```
CARLOS@LAPTOP-D81000AR /cygdrive/f/CarlosJaramillo/ICESI/datos-mcd
$ date; ls -la | grep .zip; date
Sat May  4 08:20:16 -05 2024
drwxrwxr-x+ 1 CARLOS Ninguno          0 May  4 08:17 2. unzip
-rw-rw-r--+ 1 CARLOS Ninguno 6101848686 May 20 2020 cloudera-quickstart-vm-5.4.2-0-virtualbox-GTM-configured.zip
-rw-rw-r--+ 1 CARLOS Ninguno 4831833831 Nov 11 08:21 datag-211MRows.zip
-rw-rw-r--+ 1 CARLOS Ninguno 18604990632 Nov 11 21:48 datagrams-nonstd.zip
-rw-rw-r--+ 1 CARLOS Ninguno 18604990632 May  3 08:10 datagrams.zip
-rwxrwxr-x+ 1 CARLOS Ninguno 11524877 Jul 11 2020 datagrams500k.zip
-rw-rw-r--+ 1 CARLOS Ninguno 238694237 Apr 23 22:40 virtual.zip
Sat May  4 08:20:16 -05 2024
```



1_wget.txt



2_zip.txt



3_unzip.txt



4_cat.txt



5_last.txt

Uso de disco

`df -h`

`du -h`

`wc -l archivo.txt`

COMANDO BASE

`head -n # archivo.txt`

`tail -n # archivo.txt`

`wc -l archivo.txt`

`awk '{print NF}' archivo.txt | sort -nu | tail -n 1`