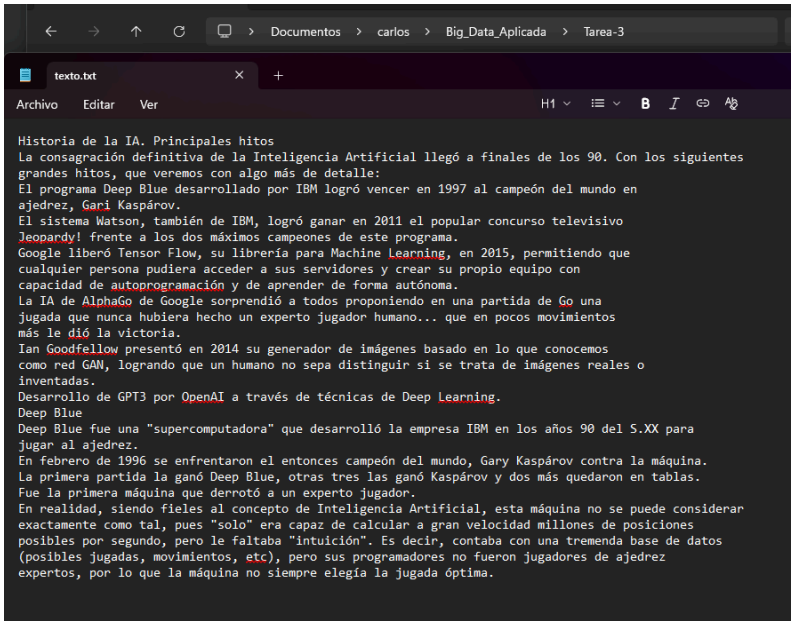
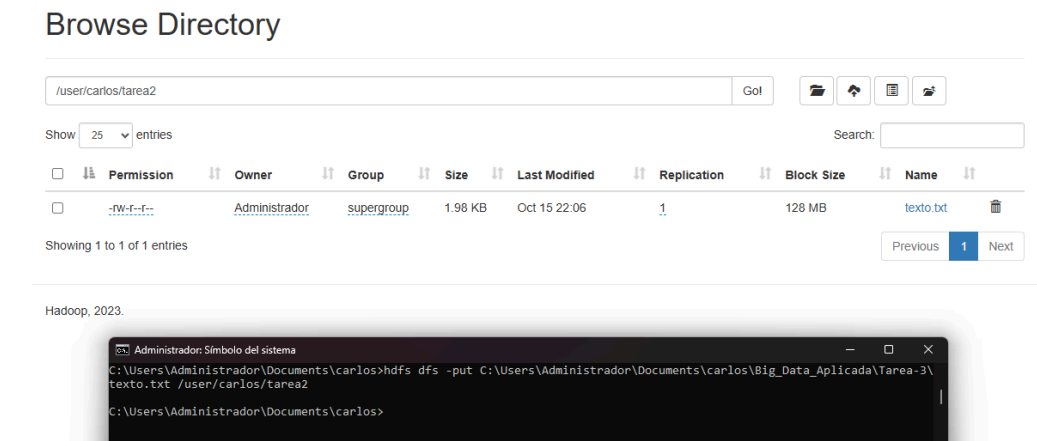


Big Data Aplicada, Tarea 3

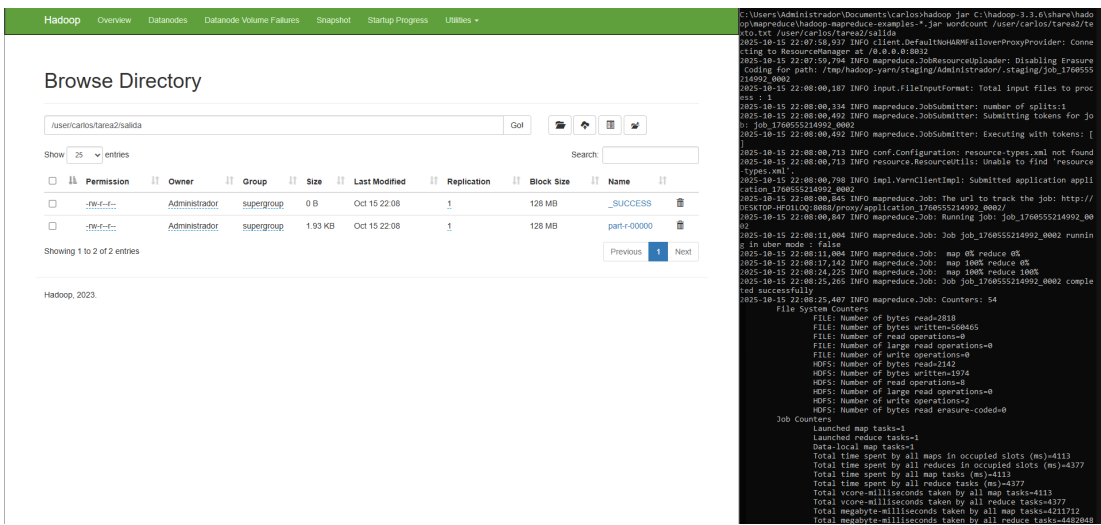
1. Crea un archivo de texto con al menos 10 líneas y varias palabras repetidas.



2. Sube el archivo a HDFS en una carpeta nueva, por ejemplo /user/tuusuario/wordcount_tarea.



3. Ejecuta el ejemplo WordCount sobre ese archivo, guardando la salida en una carpeta llamada salida dentro de la misma ruta.



4. Muestra el resultado del WordCount (cuántas veces aparece cada palabra).

```
C:\Users\Administrador\Documents\carlos>hdfs dfs -cat /user/carlos/tarea2/salida/part-r-00000
"Intuici|n". 1
"solo" 1
"supercomputadora" 1
(posibles 1
1996 1
1997 1
2011 1
2014 1
2015, 1
90 1
90. 1
AlphaGo 1
Artificial 1
Artificial, 1
Blue 3
Blue, 1
Con 1
Deep 5
Desarrollo 1
El 2
En 2
Es 1
Flow, 1
Fue 1
GAN, 1
GPT3 1
Gari 1
Gary 1
Go 1
Goodfellow 1
Google 2
Historia 1
IA 1
IA, 1
IBM 2
IBM, 1
Ian 1
Inteligencia 2
Jeopardy! 1
Kasp|irov 2
Kasp|irov. 1
La 3
Learning, 1
Learning. 1
Machine 1
OpenAI 1
Principales 1
S.XX 1
Tensor 1
Watson, 1
a 7
acceder 1
ajedrez 1
ajedrez, 1
atedrez, 1
```

```
al 3
algo 1
aprender 1
autoprogramaci|n 1
aut|noma. 1
a|os 1
basado 1
base 1
calcular 1
campeones 1
campe|n 2
capacidad 1
capaz 1
como 2
con 3
concepto 1
concurso 1
conocemos 1
consagraci|n 1
considerar 1
contaba 1
contra 1
crear 1
cualquier 1
datos 1
de 23
decir, 1
definitiva 1
del 3
derrot| 1
desarrollado 1
desarroll| 1
detalle: 1
distinguir 1
di| 1
dos 2
el 2
eleg|ja 1
empresa 1
en 10
enfrentaron 1
entonces 1
equipo 1
era 1
esta 1
este 1
etc), 1
exactamente 1
experto 2
expertos, 1
faltaba 1
febrero 1
fieles 1
finales 1
forma 1
frente 1
fue 1
fueron 1
```

Big Data Aplicada, Tarea 3

```
ganar 1
gan| 2
generador 1
gran 1
grandes 1
hecho 1
hitos 1
hitos, 1
hubiera 1
humano 1
humano... 1
imágenes 2
inventadas. 1
jugada 2
jugadas, 1
jugador 1
jugador. 1
jugadores 1
jugar 1
la 9
las 1
le 2
liber| 1
libren|a 1
liég| 1
lo 2
logrando 1
logr| 2
los 4
millones 1
movimientos 1
movimientos, 1
mundo 1
mundo, 1
miquina 3
miquina. 1
mis 3
míximos 1
no 4
nunca 1
o 1
otras 1
para 2
partida 2
permitiendo 1
pero 2
persona 1
pocos 1
popular 1
por 4
posibles 1
posiciones 1
present| 1
primera 2

programa 1
programa. 1
programadores 1
propio 1
proponiendo 1
pudiera 1
puede 1
pues 1
que 9
quedaron 1
reales 1
realidad, 1
red 1
se 3
segundo, 1
sepa 1
servidores 1
si 1
siempre 1
siendo 1
siguientes 1
sistema 1
sorprendi| 1
su 3
sus 2
tablas. 1
tal, 1
tambi|ón 1
televisivo 1
todos 1
trata 1
trav|s 1
tremenda 1
tres 1
técnicas 1
un 3
una 4
velocidad 1
vencer 1
veremos 1
victoria. 1
y 3
|ptima. 1

C:\Users\Administrador\Documents\carlos>
```

5. Responde: ¿Qué palabra aparece más veces? ¿Cuántas veces aparece la palabra "y" (si está en tu texto)?

La palabra más repetida es “de”, un total de 23 veces. La palabra “y” aparece 3 veces.