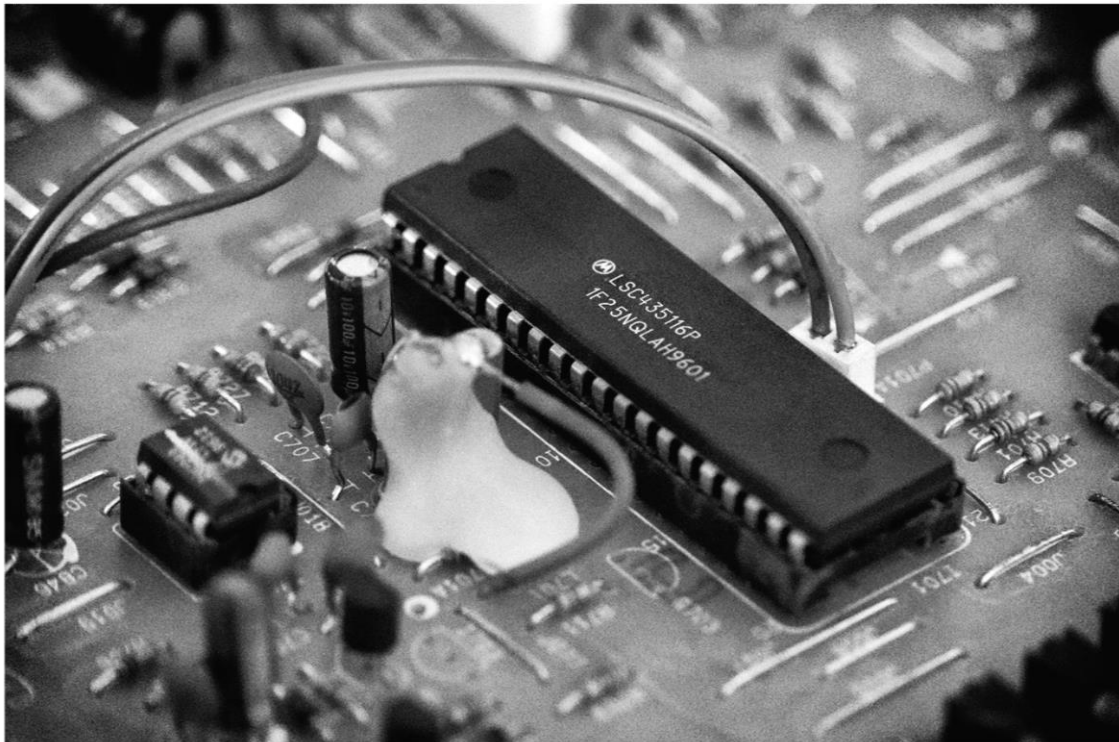


# ACTIVIDAD 02 - HADOOP



UNIVERSIDAD EUROPEA Grandes Volúmenes de Datos

## Contenido

.....	0
Contenido .....	1
1. Obtén una imagen de Hadoop para Docker: Puedes buscar en Docker Hub una imagen de Hadoop que alguien más haya creado.....	2
a. Iniciar y verificar el contenedor de Hadoop .....	2
2. Practica Hadoop .....	3
a. Crear un nuevo directorio en HDFS.....	3
b. Copiar los archivos al contenedor .....	3
c. Listar los archivos en el directorio HDFS creado .....	3
d. Leer el contenido de los archivos (Sustituye localfile.txt son los archivos descargados):.	4
3. Operación map-reduce.....	4



## 2. Practica Hadoop

### a. Crear un nuevo directorio en HDFS

Para poder crear un nuevo directorio en Hadoop simplemente nos conectaremos al contenedor correspondiente y ejecutaremos el comando mkdir seguido de la carpeta a crear:

```
C:\Users\Javie\Desktop\Actividad 02
λ docker exec -it 8e74a7e3953b hdfs dfs -mkdir /test
```

### b. Copiar los archivos al contenedor

Para copiar los archivos que se han descargado para la práctica deberemos de usar el comando put dentro de Hadoop, podemos ver un ejemplo en la siguiente captura:

```
C:\Users\Javie\Desktop\Actividad 02
λ docker exec -it 8e74a7e3953b hdfs dfs -put /datos /test
2024-01-20 11:19:11,298 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:11,796 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:11,874 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:12,299 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:12,321 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:12,653 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:13,197 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:14,349 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:15,366 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:15,396 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:16,003 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:16,134 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:16,491 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:16,684 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:17,060 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:17,222 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:17,694 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:17,827 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:18,392 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:18,577 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:18,603 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2024-01-20 11:19:19,244 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
```

Cabe destacar que los datos fueron descargados previamente y guardados en la carpeta /datos.

### c. Listar los archivos en el directorio HDFS creado

Una vez todos los datos hayan sido copiados deberemos de listar la carpeta contenedora de los mismos para verificar su correcto copiado, podemos verlo en la siguiente captura:

```
C:\Users\Javie\Desktop\Actividad 02
λ docker exec -it 8e74a7e3953b hdfs dfs -ls /test/datos
Found 13 items
-rw-r--r-- 3 root supergroup 8330 2024-01-20 11:19 /test/datos/games.csv
-rw-r--r-- 3 root supergroup 103939 2024-01-20 11:19 /test/datos/players.csv
-rw-r--r-- 3 root supergroup 3546466 2024-01-20 11:19 /test/datos/plays.csv
-rw-r--r-- 3 root supergroup 535391 2024-01-20 11:19 /test/datos/tackles.csv
-rw-r--r-- 3 root supergroup 185611101 2024-01-20 11:19 /test/datos/tracking_week_1.csv
-rw-r--r-- 3 root supergroup 180879231 2024-01-20 11:19 /test/datos/tracking_week_2.csv
-rw-r--r-- 3 root supergroup 186867093 2024-01-20 11:19 /test/datos/tracking_week_3.csv
-rw-r--r-- 3 root supergroup 186949967 2024-01-20 11:19 /test/datos/tracking_week_4.csv
-rw-r--r-- 3 root supergroup 194650924 2024-01-20 11:19 /test/datos/tracking_week_5.csv
-rw-r--r-- 3 root supergroup 164721387 2024-01-20 11:19 /test/datos/tracking_week_6.csv
-rw-r--r-- 3 root supergroup 170426365 2024-01-20 11:19 /test/datos/tracking_week_7.csv
-rw-r--r-- 3 root supergroup 185526698 2024-01-20 11:19 /test/datos/tracking_week_8.csv
-rw-r--r-- 3 root supergroup 151477618 2024-01-20 11:19 /test/datos/tracking_week_9.csv
```



d. Leer el contenido de los archivos (Sustituye localfile.txt son los archivos descargados):

Por último, podremos leer los datos utilizando el comando cat, a continuación se muestra una captura con el contenido de uno de los csv:

```
cmd
EBACK, 6, 85338993, 0.992694322, 0.007388978, -0.44E-05, 1.980613195, -0.853761717, NA, NA, NA, NA
2022103081, 111, 44901, D'Onta Foreman, (14:46) D.Foreman left tackle to CAR 27 for 7 yards (H.Walker), .1, 1, 10, CAR, ATL, CAR, 20, 14:46, 0, 0, NA, NA, 7, N, 30, SINGLEBACK, 8,
7, 0, -414432824, 0.664287094, 0.335712906, -0.021885043, 0.021885043, 0.87253832, 0.320073067, NA, NA, NA, NA
2022103010, 674, 42837, Malcolm Brown, (1:49) M.Brown left guard to SF 2 for no gain (O.Burks; F.Warner), .1, 1, 2, LA, SF, SF, 2, 1:49, 0, 0, NA, NA, 0, 0, N, 108, SINGLEBACK, 8,
0.617183270, 0.672842555, 0.327157445, -0.008896835, 0.008896835, 6.169178677, -0.34703473, NA, NA, NA, NA
2022102309, 2837, 43334, Derrick Henry, (12:27) D.Henry left guard to 50 for 5 yards (B.Danego), .4, 1, 10, TEN, IND, TEN, 45, 12:27, 16, 7, NA, NA, 5, 5, N, 65, SINGLEBACK, 8, 0.3
19702834, 0.900857959, 0.099142041, 0.003179185, -0.003179185, 2.449594241, 0.027104782, NA, NA, NA, NA
2022101610, 3205, 47857, Devin Singletary, (15:00) (Shotgun) D.Singletary up the middle to KC 46 for -2 yards (D.Harris), .4, 3, 1, BUF, KC, KC, 44, 15:00, 17, 17, NA, NA, -2
7, N, 54, SHOTGUN, 6, 0.668349981, 0.368281423, 0.631798577, 0.10371827, -0.10371827, 3.108311815, -2.934996895, NA, NA, NA, NA
2022100911, 703, 44017, James Conner, (4:19) (No Huddle, Shotgun) K.Murray pass short left to J.Conner to ARZ 48 for 8 yards (T.Edwards; K.White) [F.Cox], .1, 3, 8
ARI, PHI, ARI, 40, 4:19, 0, 7, C, 1, NA, 8, N, 50, SHOTGUN, 6, 0.978497028, 0.254828572, 0.745171428, 0.054996356, -0.054996356, 1.800985672, NA, NA, NA, NA
2022091100, 2353, 39975, Cordarrelle Patterson, (8:09) (No Huddle, Shotgun) C.Patterson right end to ATL 46 for 1 yard (D.Davis), .3, 1, 10, ATL, NO, ATL, 45, 8:09, 16, 1
0, NA, NA, 1, 1, N, 55, PISTOL, 7, 0.427988201, 0.732696291, 0.267303709, -0.005706199, 0.005706199, 2.759341507, -0.473847568, NA, NA, NA, NA
2022101610, 120, 47857, Devin Singletary, (12:59) J.Allen pass short right to D.Singletary to KC 43 for 2 yards (J.Matson), .1, 2, 1, BUF, KC, KC, 45, 12:59, 0, 0, C, 2, NA, 2
2, N, 55, I_FORM, 6, 0.408131571, 0.389211208, 0.630788762, -0.000732027, 0.000732027, 3.549283146, -0.273190378, NA, NA, NA, NA
2022092510, 55, 39983, Zach Ertz, (15:00) K.Murray pass short right to Z.Ertz to ARZ 31 for 6 yards (N.Scott), .1, 1, 10, ARI, LA, ARI, 25, 15:00, 0, 0, C, 1, NA, 6, N, 85, SING
LEBACK, 7, 0.497866362, 0.422090501, 0.577909499, 0.00217206, -0.00217206, 1.298698674, 0.027665524, NA, NA, NA, NA
2022092508, 1650, 50439, DeVonta Smith, (4:17) (Shotgun) J.Hurts pass short right to D.Smith pushed ob at WAS 11 for 31 yards (B.McCain), .2, 3, 5, PHI, WAS, WAS, 42, 4:1
7, 0, 10, C, 2, NA, 21, 31, N, 52, EMPTY, 5, 0.954746555, 0.114033366, 0.885964634, -0.036627208, 0.036627208, 2.319217055, 2.638453011, NA, NA, NA, NA
2022110608, 1621, 39987, Geno Smith, (4:40) (Shotgun) G.Smith left tackle to SEA 29 for 5 yards (B.Baker), .2, 2, 20, SEA, ARI, SEA, 24, 4:40, 7, 10, NA, NA, 5, 5, N, 34, EMPTY, 5
0.693718553, 0.457725476, 0.542274524, 0.017328877, -0.017328877, -0.052599534, -0.343133176, NA, NA, NA, NA
2022091111, 770, 41282, Davante Adams, (2:18) (Shotgun) D.Carr pass short right to D.Adams pushed ob at LV 7 for 5 yards (A.Samuel), .1, 1, 10, LV, LAC, LV, 2, 0:28, 3, 3, 6
5, NA, 5, 5, 108, SHOTGUN, 6, 0.680043876, 0.592175107, 0.407824899, -0.002112253, -0.002112253, -0.093150512, -0.221519772, NA, NA, NA, NA
2022091108, 3295, 38538, Ryan Tannehill, (1:18) R.Tannehill right tackle to NYG 29 for -2 yards (J.Ellis), .4, 1, 10, TEN, NYG, NYG, 20, 1:18, 20, 21, NA, NA, -2, -2, N, 37, JUMBO
0, 0.083381347, 0.720931950, 0.279066802, -0.02668134, 0.02668134, 2.386558984, 0, NA, NA, NA, NA
2022101604, 3219, 41258, Teddy Bridgewater, (10:50) (Shotgun) T.Bridgewater scrambles right end to MIN 4 for 10 yards (Z.Smith), .4, 2, 6, MIN, MIN, MIN, 14, 10:56, 3, 16, R
NA, NA, 10, 10, N, 24, EMPTY, 5, 0.715717826, 0.099444012, 0.900551183, 0.011829672, -0.011829672, 4.593822069, 1.330651956, NA, NA, NA, NA
2022091804, 378, 44816, Leonard Fournette, (9:02) L.Fournette right guard to TB 25 for no gain (S.Tuttle; C.Jordan), .1, 1, 10, TB, NO, TB, 25, 9:02, 3, 0, NA, NA, 0, 0, N, 35, I
_FORM, 7, 0.518475413, 0.49560002, 0.58439978, 0.021048598, -0.021048598, -0.558141164, NA, NA, NA, NA
2022100909, 3674, 43251, Carson Wentz, (1:34) (No Huddle, Shotgun) C.Wentz scrambles up the middle to TEN 39 for 8 yards (T.Mitchell; K.Byard), .4, 2, 5, WAS, TEN, TE
N, 47, 1:34, 17, 21, R, NA, NA, 8, 8, N, 57, SHOTGUN, 5, 0.945711255, 0.36861375, 0.693988525, 0.032121075, -0.032121075, 2.580679753, 0.498856777, NA, NA, NA, NA
2022100901, 3120, 54528, James Cook, (14:11) B.Hart reported in as eligible (J.Cook up the middle for 24 yards (J.Hard; K.Thibodeaux), .4, 1, 10, BUF, PIT, PIT, 24, 14:11, 31, 3, 3
NA, NA, 24, 24, N, 86, SINGLEBACK, 8, 0.674119397, 0.998898715, 0.001181285, 0.000245086, -0.000245086, 4.184517902, 2.815482098, NA, NA, NA, NA
2022092512, 740, 44995, Aaron Jones, (4:33) (Shotgun) A.Jones right end to TB 14 for -2 yards (C.Davis, L.David), .1, 1, 10, GB, TB, TB, 12, 4:33, 3, 0, NA, NA, -2, -2, N, 22,
SHOTGUN, 7, 0.358594179, 0.519621961, 0.480378039, 0.02100473, -0.02100473, 3.149817327, -0.862474349, NA, NA, NA, NA
2022100901, 524, 52453, Najee Harris, (6:30) (Shotgun) N.Harris left tackle to BUF 11 for -1 yards (T.Johnson, T.Dodson), .1, 2, 7, PIT, BUF, BUF, 10, 6:30, 7, 0, NA, NA, -1
1, -1, N, 100, SHOTGUN, 3, 0.708302811, 0.903337468, 0.096662532, 0.00425415, -0.00425415, 4.809419064, -0.88493249, NA, NA, NA, NA
2022091805, 3046, 46071, Saquon Barkley, (13:43) (No Huddle) S.Barkley right end pushed ob at CAR 38 for no gain (Y.Gross-Matos; C.Henderson), .4, 2, 7, NYG, CAR, CAR, 3
8, 13:43, 13, 13, NA, NA, 0, 0, N, 72, SINGLEBACK, 6, 0.867895365, 0.565924726, 0.434075274, -0.030270286, 0.030270286, 2.964200744, -0.875009551, NA, NA, NA, NA
2022102302, 1215, 54042, Caleb Hunkley, (9:38) (Shotgun) C.Hunkley right tackle to CIN 28 for 2 yards (V.Bell; G.Park), .2, 1, 10, ATL, CIN, CIN, 27, 9:38, 21, 0, NA, NA, 2
2, N, 37, PISTOL, 6, 0.704550087, 0.947739481, 0.952269519, 0.000588939, -0.000588939, 4.137341118, -0.237006564, NA, NA, NA, NA
2022092511, 3337, 54506, Kenneth Walker, (11:46) (Shotgun) G.Smith pass short middle to K.Walker to SEA 45 for 5 yards (R.Evans), .4, 3, 20, SEA, ATL, SEA, 40, 11:46, 23, 2
7, C, 5, NA, 5, N, 50, SHOTGUN, 4, 0.962818268, 0.268827739, 0.713172261, -0.043963686, 0.043963686, -0.240246499, -0.947411295, NA, NA, NA, NA
2022100909, 441, 54597, Romeo Doubs, (8:00) (Shotgun) A.Rodgers pass short right to R.Doubs to 08 27 for -1 yards (J.Ward, K.Thibodeaux), .1, 3, 10, GB, NYG, GB, 28, 8:0
00, 3, 0, C, -3, NA, 1, -1, N, 82, SHOTGUN, 5, 0.970363915, 0.856418841, 0.143581159, -0.021287875, 0.021287875, 0.013747229, -1.641053633, NA, NA, NA, NA
2022091110, 2925, 39983, Zach Ertz, (3:56) (Shotgun) K.Murray pass short middle to Z.Ertz to ARZ 33 for 8 yards (N.Bolton), .3, 1, 10, ARI, KC, ARI, 25, 3:56, 7, 37, C, 5, NA,
8, 8, N, 85, EMPTY, 5, 0.747570097, 0.001760755, 0.998239245, -3.71E-06, 3.71E-06, 1.279283918, 0.425718037, NA, NA, NA, NA
2022100912, 2238, 42358, Melvin Gordon, (9:14) M.Gordon left tackle to 50 for 1 yard (D.Deablo), .3, 1, 10, DEN, LV, DEN, 49, 9:14, 19, 16, NA, NA, 1, 1, N, 59, WILDCAT, 6, 0.51152
5691, 0.570692504, 0.420837436, 0.01820916, -0.01820916, 2.870688786, -0.352056304, NA, NA, NA, NA
2022092507, 310, 43334, Derrick Henry, (8:54) (Shotgun) R.Tannehill pass short middle to D.Henry to LV 5 for 6 yards (J.Brown, D.Deablo), .1, 1, 10, TEN, LV, LV, 11, 8:
54, 0, 0, C, 3, NA, 6, 6, N, 99, SHOTGUN, 6, 0.565141678, 0.665483802, 0.334516998, -9.79E-05, 9.79E-05, 5.14574742, 0.291840269, NA, NA, NA, NA
2022091105, 315, 33084, Matt Ryan, (9:10) (Shotgun) M.Ryan scrambles right end to H0T 46 for 2 yards (K.Gruiger-Hill), .1, 3, 1, IND, HOU, HOU, 48, 9:28, 0, 0, R, NA, NA, 2, 2, N
58, SHOTGUN, 7, 0.63081531, 0.22538253, 0.777451747, 0.00674157, 0.00674157, 2.82484379, 0.336768883, NA, NA, NA, NA
2022100201, 288, 39975, Cordarrelle Patterson, (10:10) C.Patterson left end to ATL 10 for 6 yards (A.Wright), .1, 1, 10, ATL, CLE, ATL, 4, 10:10, 0, 0, NA, NA, 6, 6, N, 14, I_FOR
M, 8, 0.188302457, 0.553130433, 0.446869567, 0.000985105, -0.000985105, 0.006170481, 0.031114305, NA, NA, NA, NA
2022092508, 2194, 41292, Jarvis Landry, (13:41) J.Winston pass short left to J.Landry to NO 35 for 13 yards (K.Woods), .3, 1, 10, NO, CAR, NO, 22, 13:41, 13, 0, C, 13, NA, 13, 1
3, N, 22, SINGLEBACK, 7, 0.524089297, 0.821099510, 0.160894481, -0.021405812, 0.021405812, 3.389707026, 0.540565064, NA, NA, NA, NA
2022110605, 3822, 53000, J.J. Taylor, (1:10) J.Taylor right tackle to IND 47 for 2 yards (S.Gilmore, K.Paye), .4, 3, 7, NE, IND, IND, 49, 1:18, 26, 3, NA, NA, 2, 2, N, 59, SING
LEBACK, 8, 0.618373501, 0.998757417, 0.001242583, -4.71E-05, 4.71E-05, 1.66698603, -1.082212784, NA, NA, NA, NA
2022103001, 3650, 44901, D'Onta Foreman, (2:14) (Shotgun) D.Foreman right end to CAR 26 for 1 yard (A.Ebiketie), .4, 1, 10, CAR, ATL, CAR, 25, 2:14, 31, 28, NA, NA, 1, 1, N, 35,
SHOTGUN, 6, 0.824573310, 0.69604093, 0.30393107, 0.024965978, -0.024965978, 1.46306466, 0.477651574, NA, NA, NA, NA
2022101300, 2082, 54563, Brian Robinson, (14:19) (Shotgun) B.Robinson left end to WAS 36 for 7 yards (R.Smith), .3, 2, 6, WAS, CHI, WAS, 29, 14:19, 0, 3, NA, NA, 7, 7, N, 39, SHO
TGUN, 6, 0.61963564, 0.33897665, 0.666102335, -0.021749627, 0.021749627, 1.13147496, 0.945554113, NA, NA, NA, NA
2022110607, 2806, 54563, Brian Robinson, (15:00) B.Robinson right end to MIN 6 for 5 yards (J.Hicks), .4, 1, 10, WAS, MIN, MIN, 11, 15:00, 10, 7, NA, NA, 5, 5, N, 21, SINGLEBACK,
7, 0.369950851, 0.72168906, 0.227831014, -0.00812417, -0.00812417, 4.063586025, 0.44453927, NA, NA, NA, NA
2022102307, 1224, 43358, Melvin Gordon, (14:18) (Shotgun) M.Gordon left tackle to NYJ 49 for 1 yard (Q.Williams), .2, 2, 8, DEN, NYJ, NYJ, 50, 14:18, 6, 7, NA, NA, 1, 1, N, 60, SH
OTGUN, 6, 0.721504867, 0.45803922, 0.541196078, -0.036081951, 0.036081951, 1.939442847, -0.81722467, NA, NA, NA, NA
2022100213, 811, 40011, Travis Kelce, (1:27) P.Mahomes pass short left to T.Kelce to KC 33 for 7 yards (W.Gholston), .1, 2, 2, KC, TB, KC, 26, 1:27, 3, 14, C, 5, NA, 7, N, 84, S
INGLEBACK, 7, 0.366100824, 0.167661197, 0.832338803, 0.001441572, -0.001441572, 1.018607682, 0.354830387, NA, NA, NA, NA
2022092504, 2690, 52725, Ben Ellsford, (8:31) K.Cousins pass short right to B.Ellsford to MIN 35 for 12 yards (A.Grunariye), .3, 1, 10, MIN, DET, MIN, 29, 8:31, 14, 17, C, 12
NA, NA, 12, 12, N, 87, I_FORM, 8, 0.467717916, 0.569607757, 0.430392243, 0.024004817, -0.024004817, 1.105008532, 0.715667434, NA, NA, NA, NA
2022100910, 105, 44959, George Kittle, (13:43) J.Garoppolo pass deep left to G.Kittle pushed ob at SF 49 for 20 yards (J.Burris), .1, 3, 6, SF, CAR, SF, 29, 13:
43, 0, 0, C, 17, NA, 20, 20, N, 81, SHOTGUN, 7, 0.99047025, 0.395102009, 0.604897991, -0.068615705, 0.068615705, 0.172975704, 2.69119564, NA, NA, NA, NA
2022101608, 3371, 44860, Joe Mixon, (9:35) (Shotgun) J.Mixon left end to CIN 47 for 6 yards (C.Harris, D.Davis), .4, 1, 10, CIN, NO, CIN, 41, 9:35, 26, 21, NA, NA, 6, 6, N, 69
SHOTGUN, 6, 0.639438927, 0.66705386, 0.33294614, 0.005164288, 0.005164288, 2.115356216, 0.203819486, NA, NA, NA, NA
2022100207, 2777, 52449, Jonathan Taylor, (2:02) (Shotgun) J.Taylor up the middle to TEN 36 for -2 yards (D.Long, J.Schobert), .3, 1, 10, IND, TEN, TEN, 34, 2:02, 17, 24
NA, NA, -2, -2, N, 44, SINGLEBACK, 7, 0.518695235, 0.410611343, 0.589388657, -0.046648037, 0.046648037, 3.946232241, -0.976839406, NA, NA, NA, NA
2022102308, 905, 54572, Damion Pierce, (12:51) C.Hack reported in as eligible (D.Pierce up the middle to LV 14 for 3 yards (D.Perryman; D.Deablo), .2, 3, 1, HOU, LV, L
V, 17, 12:51, 3, 0, NA, NA, 3, 3, N, 27, I_FORM, 8, 0.11525352, 0.726639198, -0.273360082, -0.015027672, 0.015027672, 6.685974378, 0.25157968, NA, NA, NA, NA
cmd.exe
```

### 3. Operación map-reduce

La operación map-reduce que se ha utilizado ha sido para contar el número de pesos que se repiten en el archivo players.csv. En primer lugar, se obtenido el contenido del fichero csv, una vez se ha almacenado en una variable se divide en líneas haciendo un Split mediante el salto de carro (\n). Una vez tengamos divididas las líneas simplemente se ejecuta un bucle foreach que recorre cada una de las mismas. Por cada línea se realiza otro Split usando como separador el carácter “,”. Obtenemos la altura y la almacenamos para ser contada.

```

public class CustomMapper extends MapReduceBase implements Mapper {
    private final static IntWritable one = new IntWritable(value:1);
    private String[] lines;

    @Override
    public void map(Object key, Object value, OutputCollector output, Reporter reporter) throws IOException {
        String content = value.toString();
        lines = content.split(regex:"\\n");

        for (String line : lines) {
            int weight = Integer.valueOf(line.split(regex:",")[2]);
            output.collect(weight, one);
        }
    }
}

```

La segunda parte consta simplemente de obtener los valores anteriores e ir sumando para obtener el número de veces que se repite. Con esto conseguimos obtener el número de personas que tienen el mismo peso.

```

public class CustomReducer extends MapReduceBase implements Reducer<Text, IntWritable, Text, IntWritable> {

    @Override
    public void configure(JobConf job) {
        // TODO Auto-generated method stub
        throw new UnsupportedOperationException(message:"Unimplemented method 'configure'");
    }

    @Override
    public void close() throws IOException {
        // TODO Auto-generated method stub
        throw new UnsupportedOperationException(message:"Unimplemented method 'close'");
    }

    @Override
    public void reduce(Text key, Iterator<IntWritable> values, OutputCollector<Text, IntWritable> output,
        Reporter reporter) throws IOException {
        int sum = 0;

        while (values.hasNext()) {
            sum += values.next().get();
        }

        output.collect(key, new IntWritable(sum));
    }
}

```