

**ALMACENAMIENTO DE DATOS Y SU ADMINISTRACIÓN**



**MAESTRÍA EN CIENCIAS DE LA INFORMÁTICA**

**NOMBRE: JUAN JOSÉ LAIR MARTÍNEZ**

**BOLETA: B150848**

TAREA 5: Práctica 1

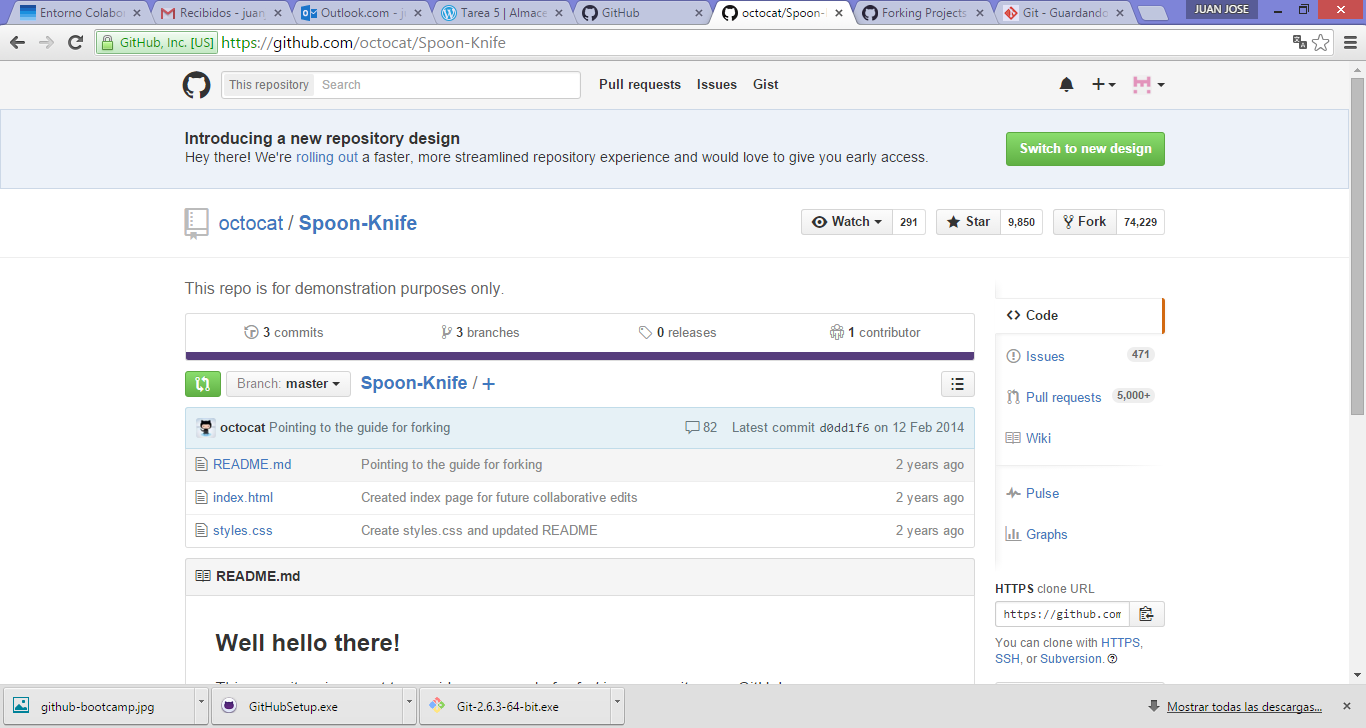
ACTIVIDADES

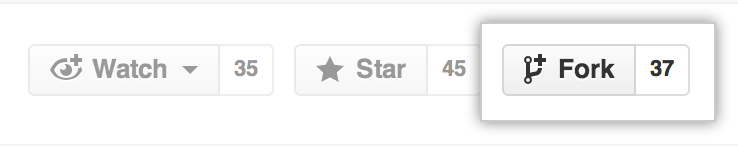
**PRIMERA PARTE**

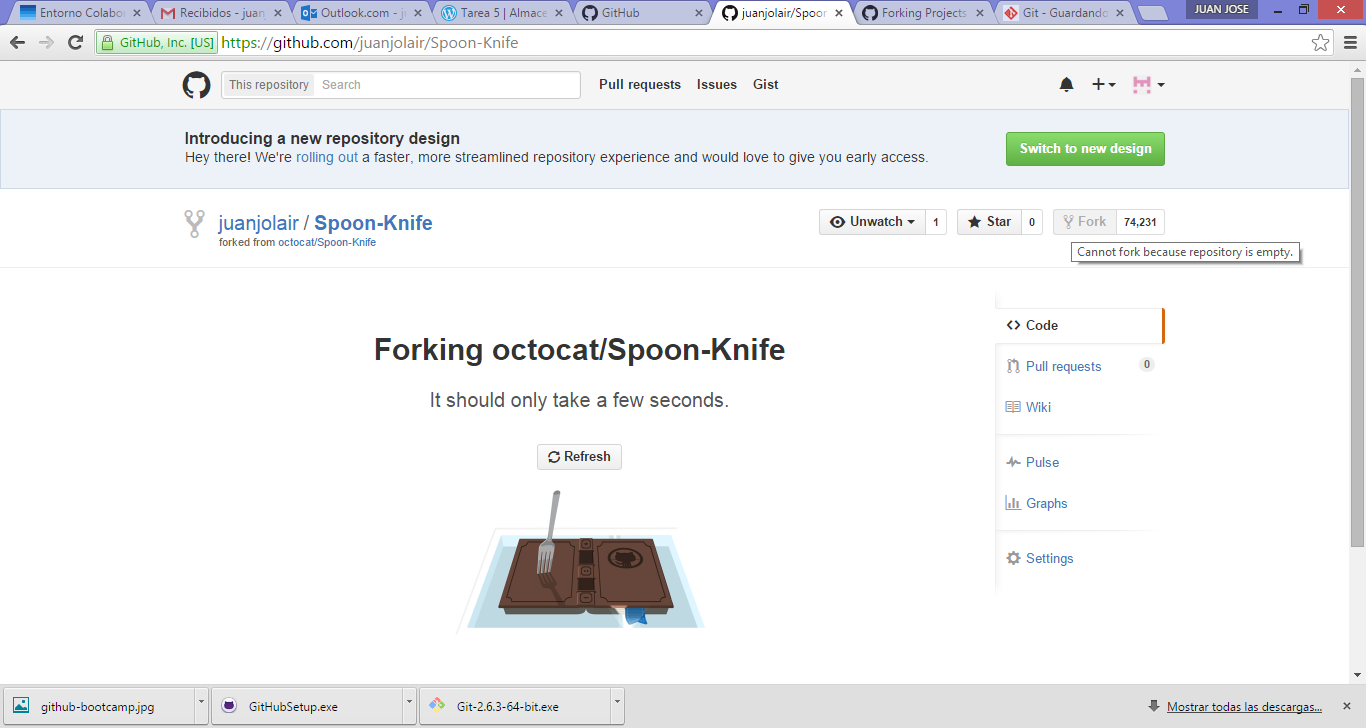
1. Siga el ejemplo sugerido en la página del “**GitHub Bootcamp**” sobre creación de repositorios (*click* en la imagen para ir al artículo) y proceda con el “*fork*” del repositorio **Spoon-Knife** que en éste se indica.

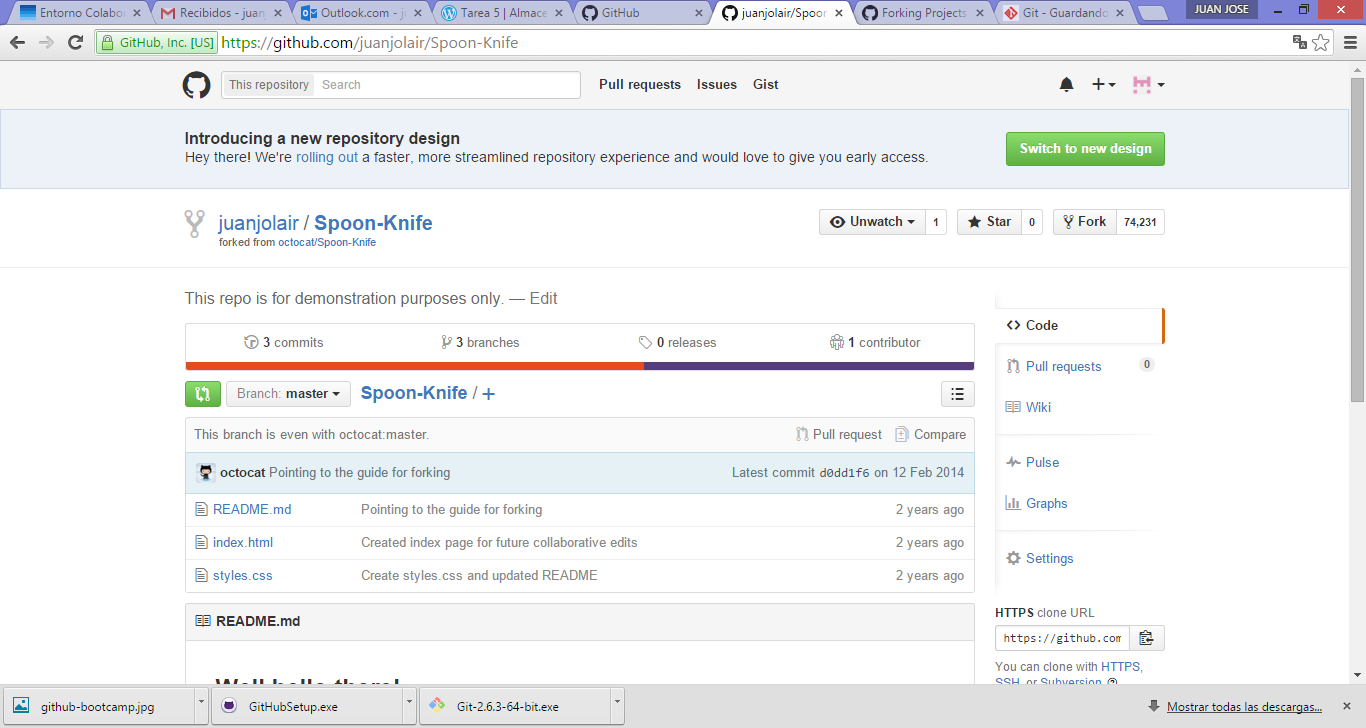


1. Tome evidencia de la creación del repositorio bifurcado ya creado en su cuenta.



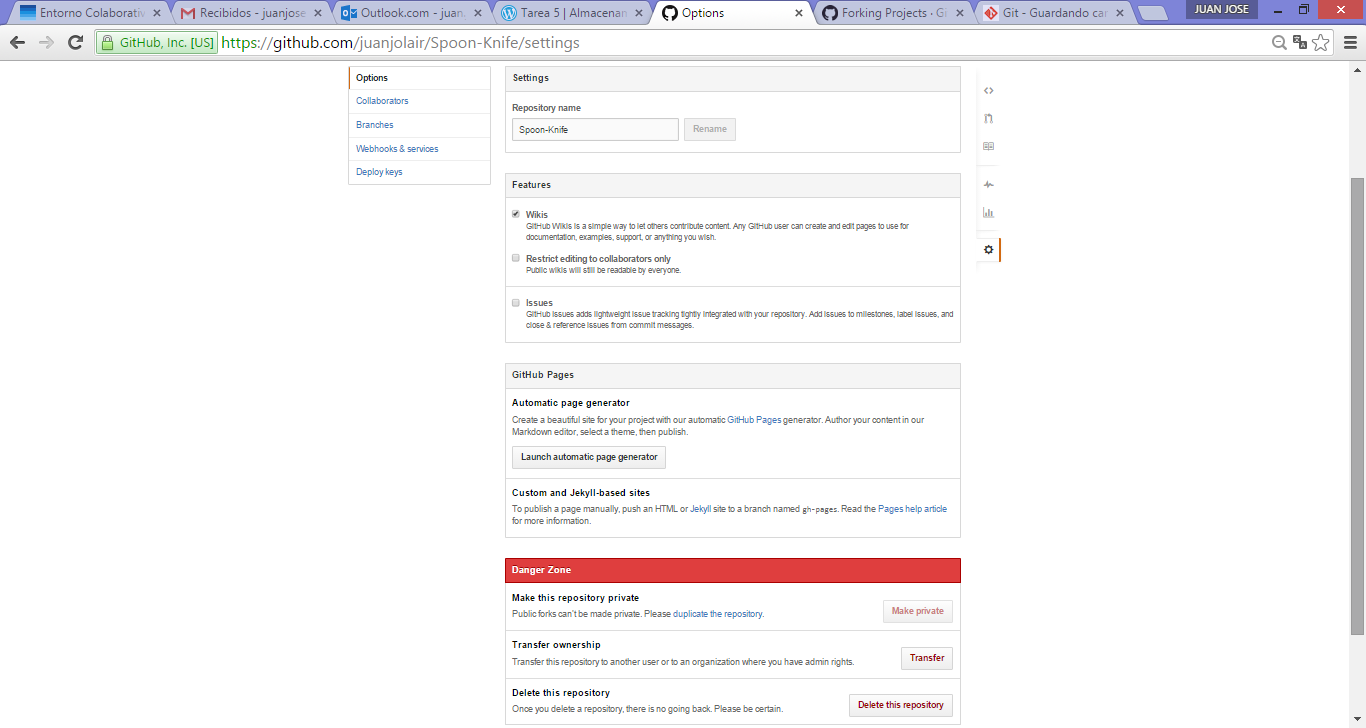


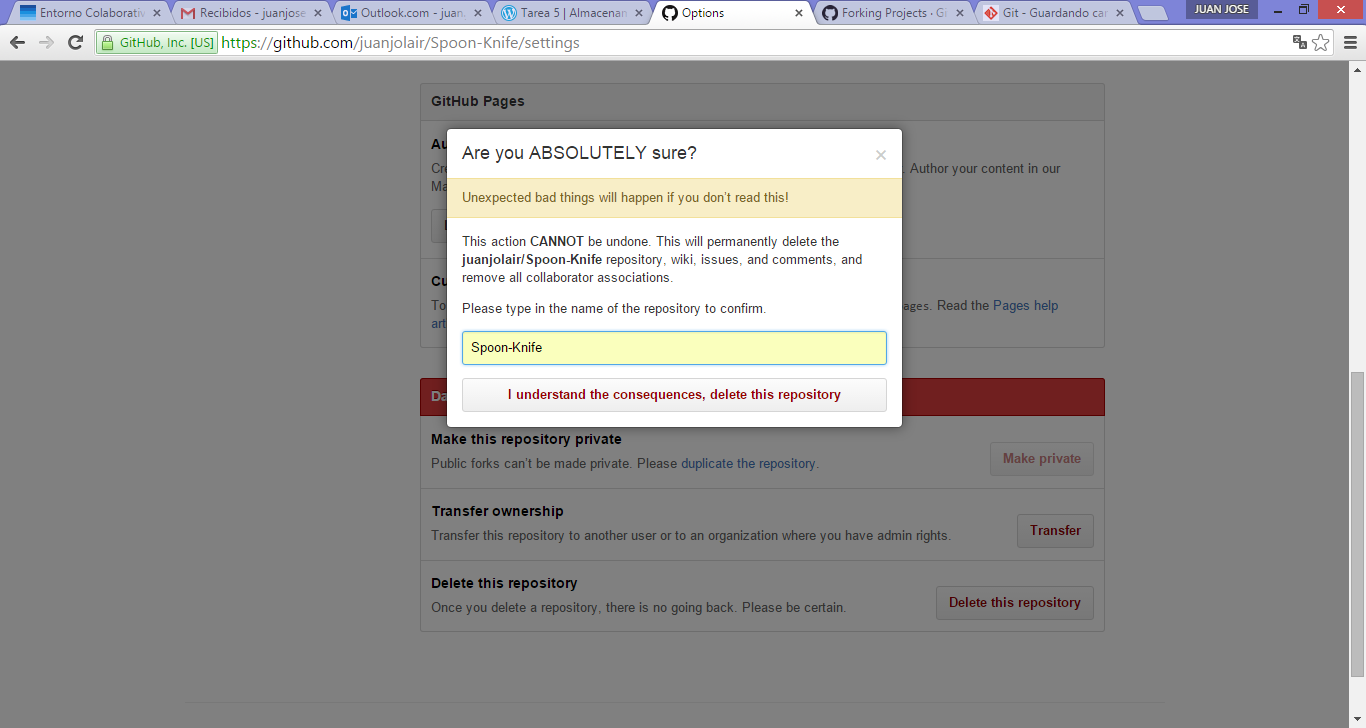




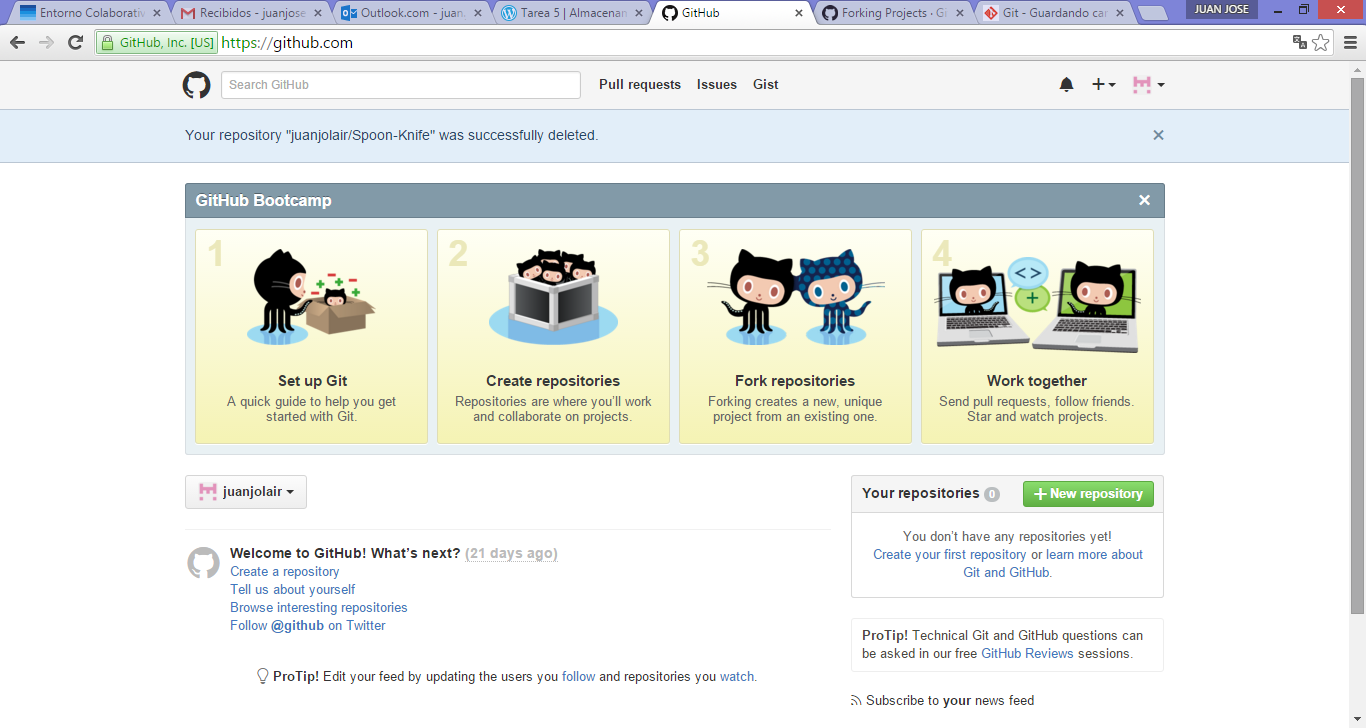
1. Investigue como eliminarlo y proceda con su borrado.

Dentro de la opción “Settings” se encuentra un apartado llamado “Danger Zone” en donde podemos ubicar la opción para eliminar el repositorio en el apartado “Delete this repository”.



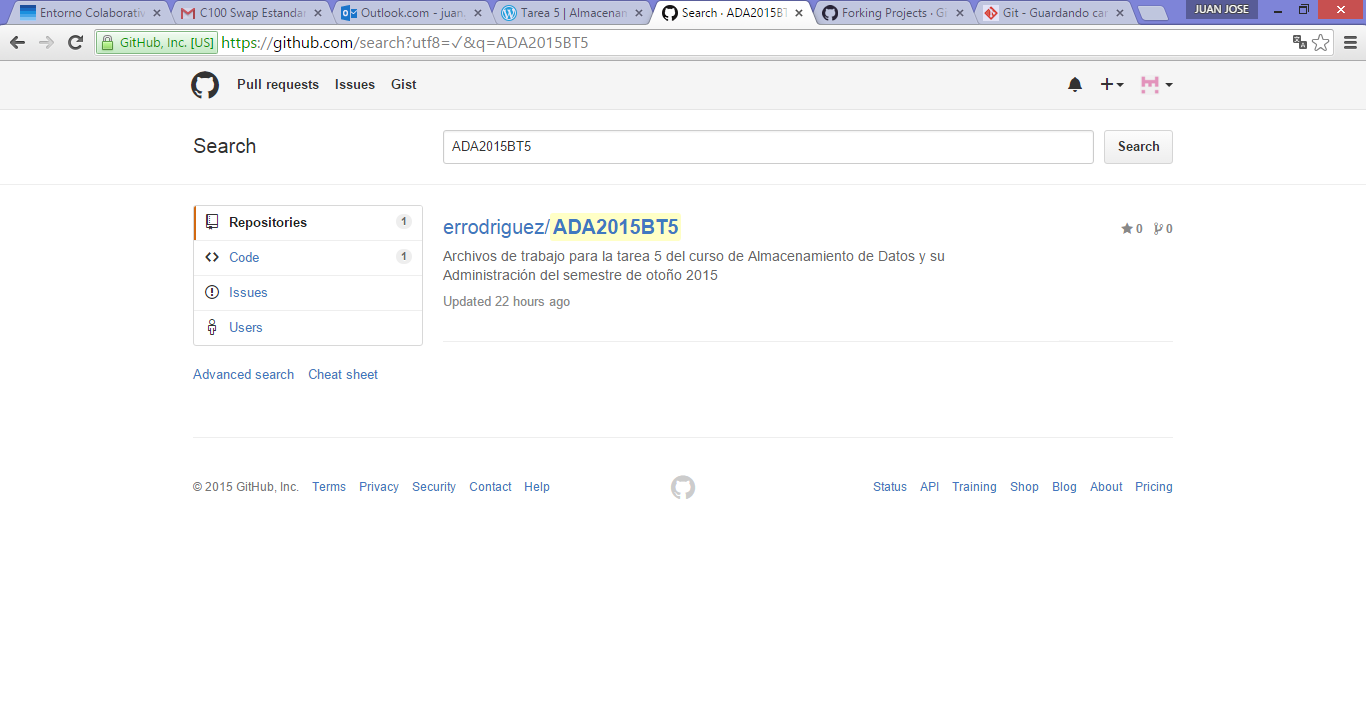


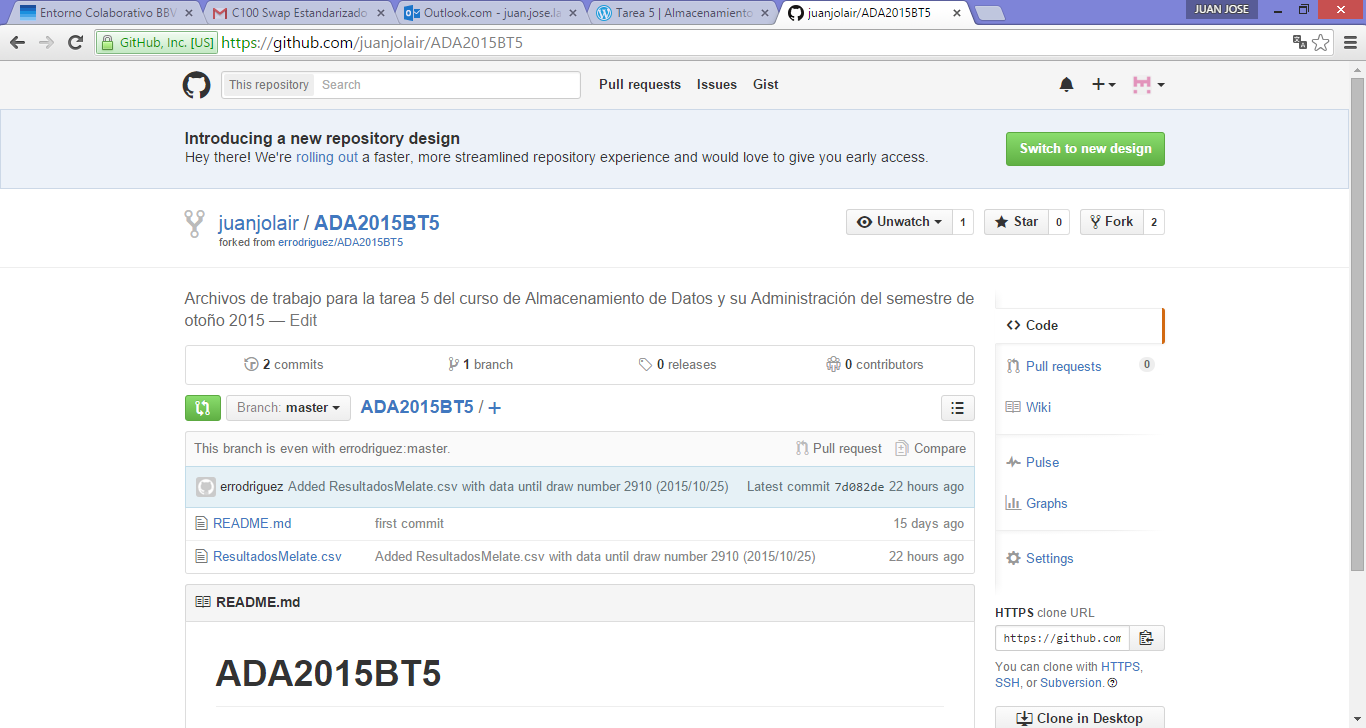
Se especifica la carpeta que se desea borrar, en este caso, Spoon-Knife y, con ello, se elimina el repositorio regresando a la pantalla inicial y muestra el mensaje: “Your repository "juanjolair/Spoon-Knife" was successfully deleted.”.



**SEGUNDA PARTE**

1. Busque el repositorio ADA2015BT5 de la cuenta errodriguez y bifúrquelo. Esto en el *browser*, los siguientes pasos en una ventana de **MSDOS** o *Shell*  **Unix**, según sea el caso.





1. Si Ud. ya ha configurado localmente a git en su equipo pase al siguiente punto, en otro caso continúe leyendo. Proceda con la configuración general, de primera vez ([first- time setup](https://git-scm.com/book/en/v2/Getting-Started-First-Time-Git-Setup)). Muestre comandos y resultados.

Se configura tu email

Se configura tu identidad



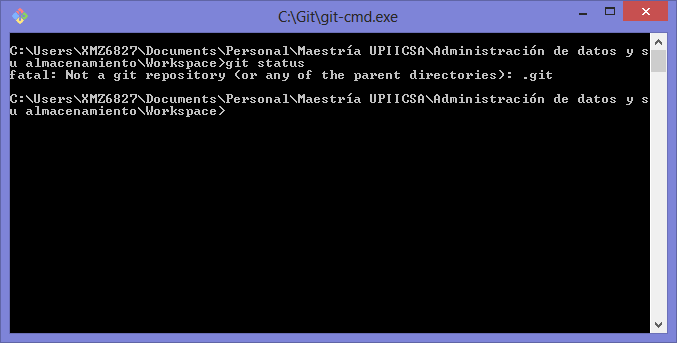
Se configura tu editor

1. Liste la configuración global con que termina este paso.

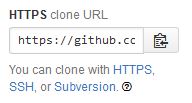


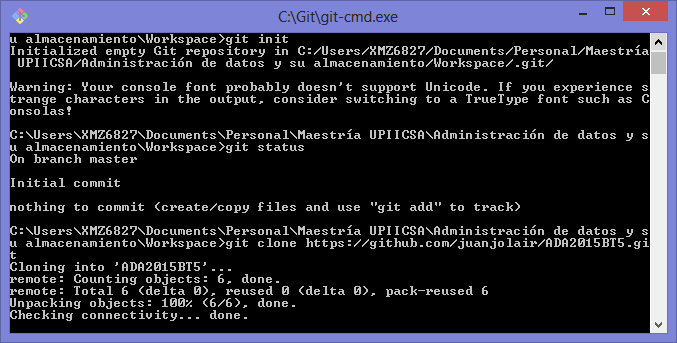
Se comprueba tu configuración

1. Ubique una ruta en su equipo donde desea trabajar. Cree un directorio de ser necesario. Proporcione evidencia de que este directorio no es un depósito git.



1. En la ruta seleccionada proceda con la clonación de repositorio bifurcado en el primer paso de esta sección. El URL a usar debe aparecer en la página del repositorio, buscar:



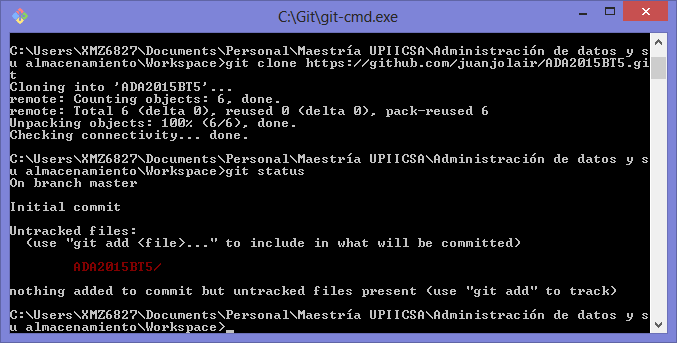


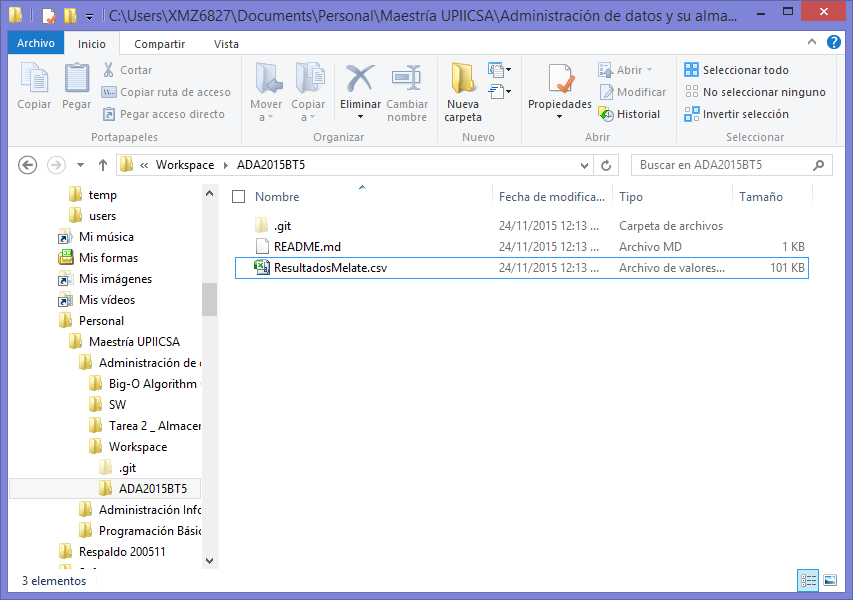
1. Investigue al respecto de la diferencia entre clonar y bifurcar un repositorio. Indique si son o no parte de la funcionalidad de git o si alguna es propia de GitHub.

Cuando hacemos *fork* de un repositorio, se crea un nuevo repositorio dentro de nuestra cuenta de github con una URL diferente (<https://github.com/juanjolair/ADA2015BT5.git>). Posteriormente se debe de realizar una clonación (*git clone*) de esa copia sobre la que se va a trabajar y las modificaciones se hacen sobre nuestra copia de nuestro repositorio. En otro caso, si realizamos una clonación directa sobre la URL del repositorio, las modificaciones se haría directamente a ese repositorio, en este caso, a la URL que es: <https://github.com/errodriguez/ADA2015BT5.git> si se tuvieran permisos.



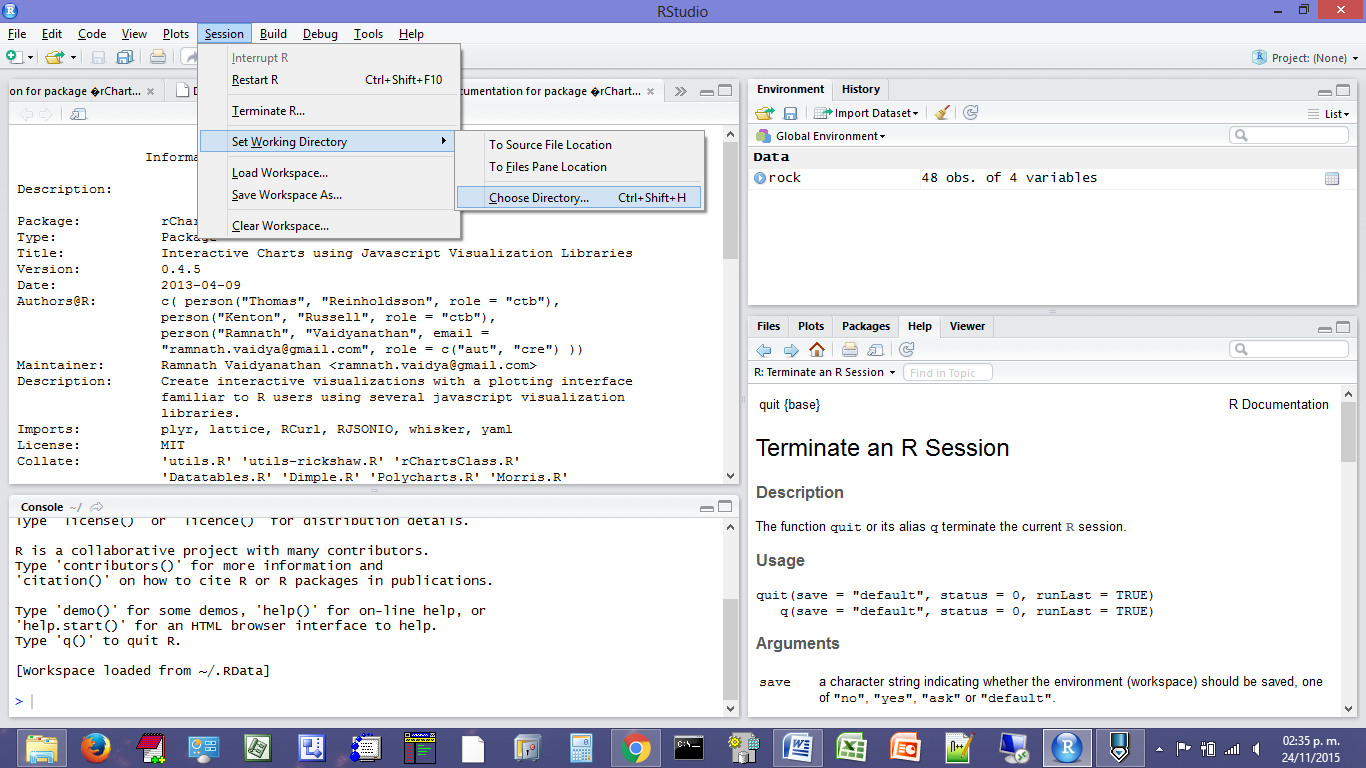
1. El clonado deberá crear un directorio nuevo con el nombre del repositorio. Proporcione evidencia de que éste es un almacén git.

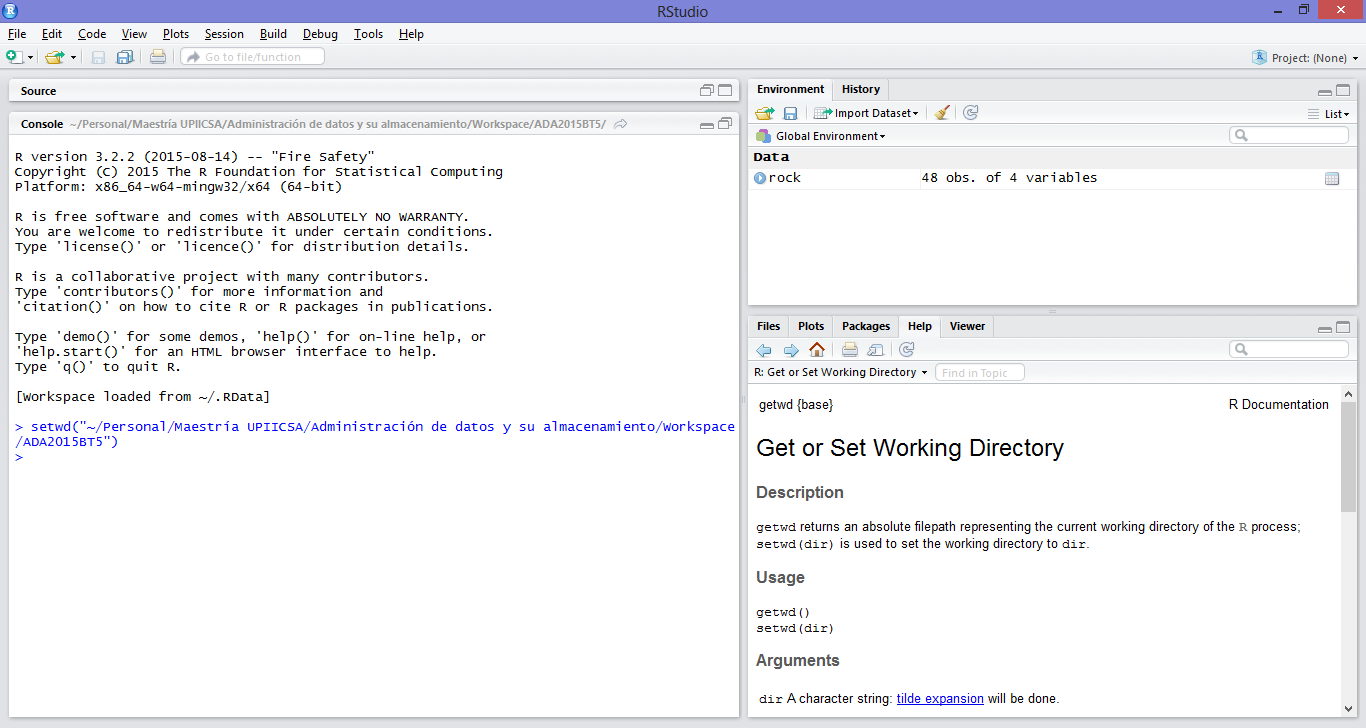




**TERCERA PARTE (Usando RStudio)**

1. Fije como directorio de trabajo el directorio creado en la segunda parte de este trabajo. Tome evidencia.

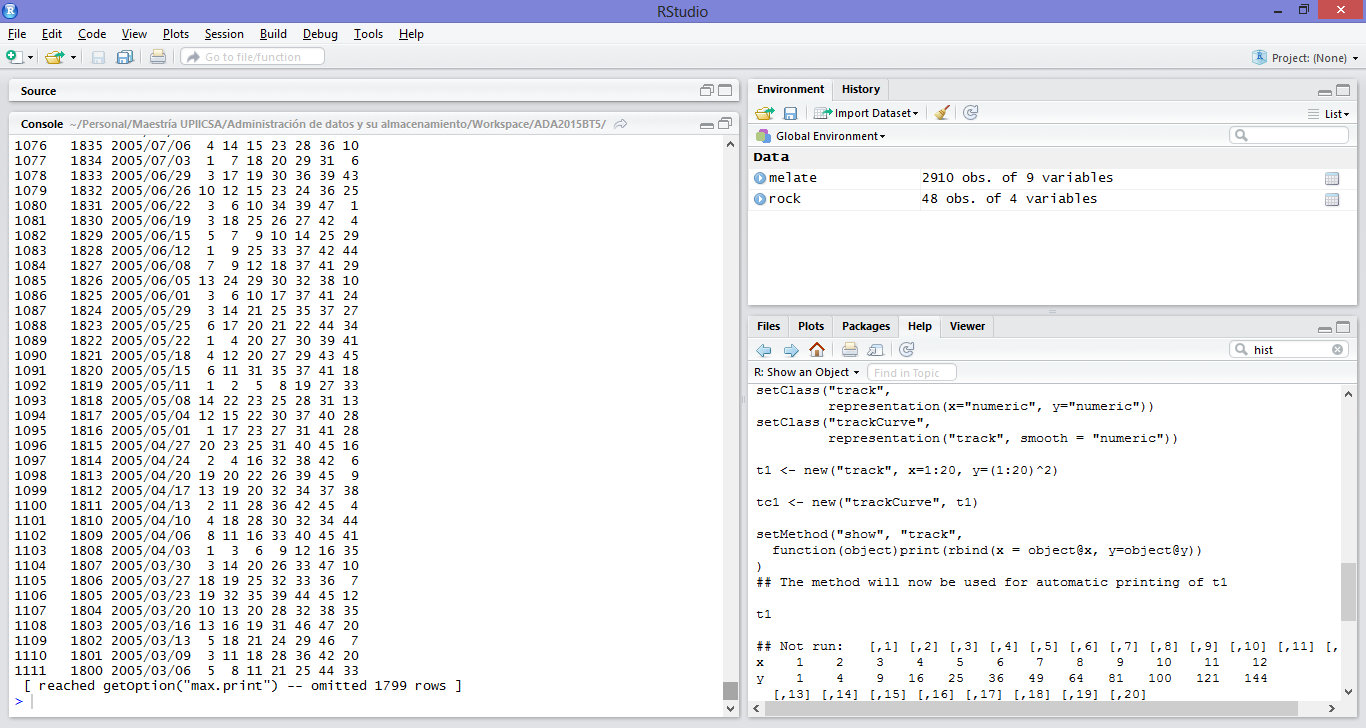




1. Pruebe primero la ventana de la consola de **RStudio** las instrucciones necesarias para las siguientes acciones:

* Lea el archivo CSV dentro del repositorio git (tip: una forma es mediante la función read.csv)

> read.csv("ResultadosMelate.csv")



* Haga una primera exploración del archivo mostrando:
  + Encabezados

> melate <- read.csv("ResultadosMelate.csv")

> colnames(melate)

[1] "SORTEO" "FECHA" "N1" "N2" "N3" "N4" "N5" "N6" "A"

* + Estructura del archivo y algunos registros

> nrow(melate)

[1] 2910

> ncol(melate)

[1] 9

> read.csv("ResultadosMelate.csv", skip = 2895)

X16 X1984.12.02 X2 X11 X13 X22 X30 X36 X24

1 15 1984/11/25 6 9 15 23 32 37 7

2 14 1984/11/18 4 8 25 32 36 38 27

3 13 1984/11/11 6 9 23 24 28 36 26

4 12 1984/11/04 3 14 24 32 36 39 28

5 11 1984/10/28 3 5 14 15 26 34 28

6 10 1984/10/21 2 3 24 25 29 34 7

7 9 1984/10/14 4 6 10 15 32 36 33

8 8 1984/10/07 4 12 17 19 20 30 25

9 7 1984/09/30 12 14 21 29 32 39 26

10 6 1984/09/23 12 23 29 31 32 36 1

11 5 1984/09/16 4 10 13 18 33 39 11

12 4 1984/09/09 8 12 15 18 20 34 37

13 3 1984/09/02 16 18 19 21 29 32 27

14 2 1984/08/26 5 6 24 30 34 38 32

15 1 1984/08/19 1 4 7 11 19 30 36

* + Un resumen (summary)

> summary(melate)

SORTEO FECHA N1 N2 N3

Min. : 1.0 1997/01/26: 2 Min. : 1.000 Min. : 2.00 Min. : 3.00

1st Qu.: 728.2 1997/10/22: 2 1st Qu.: 3.000 1st Qu.: 8.00 1st Qu.:14.00

Median :1455.5 1984/08/19: 1 Median : 5.000 Median :13.00 Median :20.00

Mean :1455.5 1984/08/26: 1 Mean : 6.801 Mean :13.58 Mean :20.68

3rd Qu.:2182.8 1984/09/02: 1 3rd Qu.: 9.000 3rd Qu.:18.00 3rd Qu.:26.00

Max. :2910.0 1984/09/09: 1 Max. :35.000 Max. :44.00 Max. :52.00

(Other) :2902

N4 N5 N6 A

Min. : 4.00 Min. :10.00 Min. :14.00 Min. : 1.00

1st Qu.:22.00 1st Qu.:29.00 1st Qu.:37.00 1st Qu.:12.00

Median :27.50 Median :35.00 Median :41.00 Median :24.00

Mean :27.72 Mean :34.67 Mean :41.57 Mean :24.42

3rd Qu.:34.00 3rd Qu.:40.00 3rd Qu.:47.00 3rd Qu.:36.00

Max. :54.00 Max. :55.00 Max. :56.00 Max. :56.00

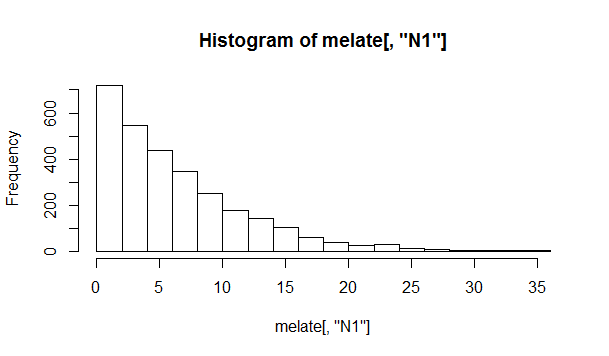
* + Dimensiones del archivo

> dim(melate)

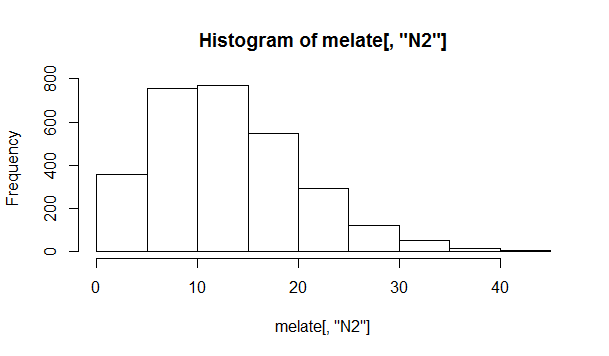
[1] 2910 9

* + El archivo cuenta con varias columnas. Seis de ellas son de particular importancia. Haga un histograma por cada una.

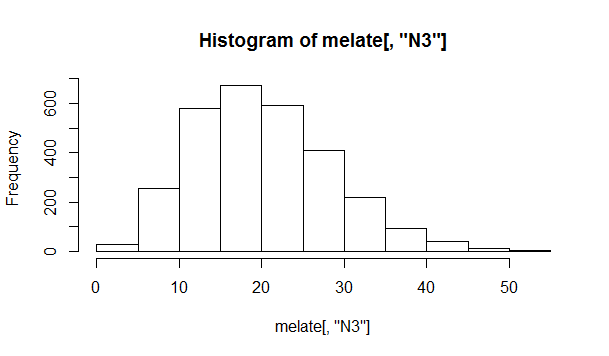
> hist(melate[,"N1"])



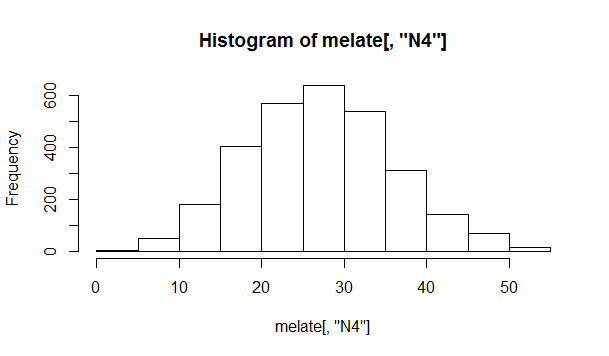
> hist(melate[,"N2"])



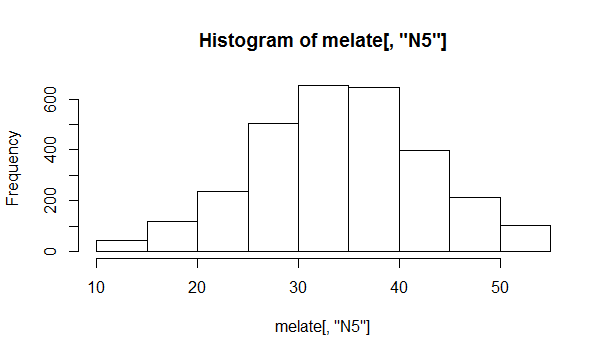
> hist(melate[,"N3"])



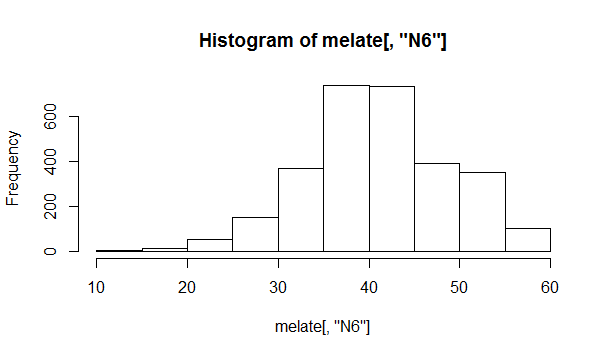
> hist(melate[,"N4"])



> hist(melate[,"N5"])



> hist(melate[,"N6"])



* + Opcional: cualquier otra acción sencilla exploratoria del contenido del archivo es bienvenida.

> min(melate[,"N1"])

[1] 1

> max(melate[,"N1"])

[1] 35

> min(melate[,"N2"])

[1] 2

> max(melate[,"N2"])

[1] 44

> min(melate[,"N3"])

[1] 3

> max(melate[,"N3"])

[1] 52

> min(melate[,"N4"])

[1] 4

> max(melate[,"N4"])

[1] 54

> min(melate[,"N5"])

[1] 10

> max(melate[,"N5"])

[1] 55

> min(melate[,"N6"])

[1] 14

> max(melate[,"N6"])

[1] 56

> min(melate[,"SORTEO"])

[1] 1

> max(melate[,"SORTEO"])

[1] 2910

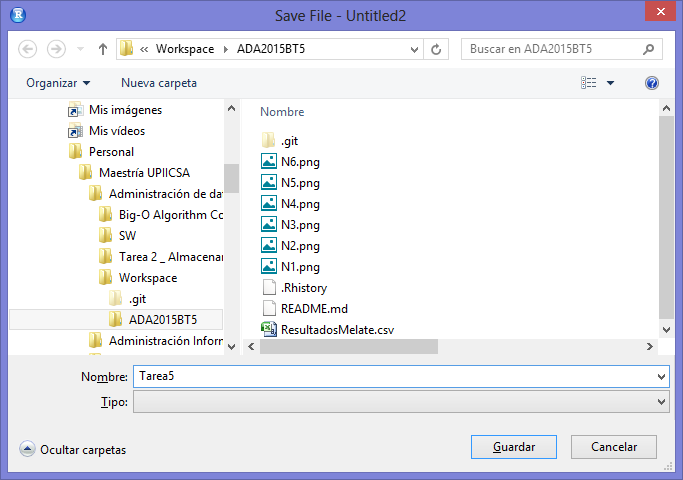
> min(melate[,"A"])

[1] 1

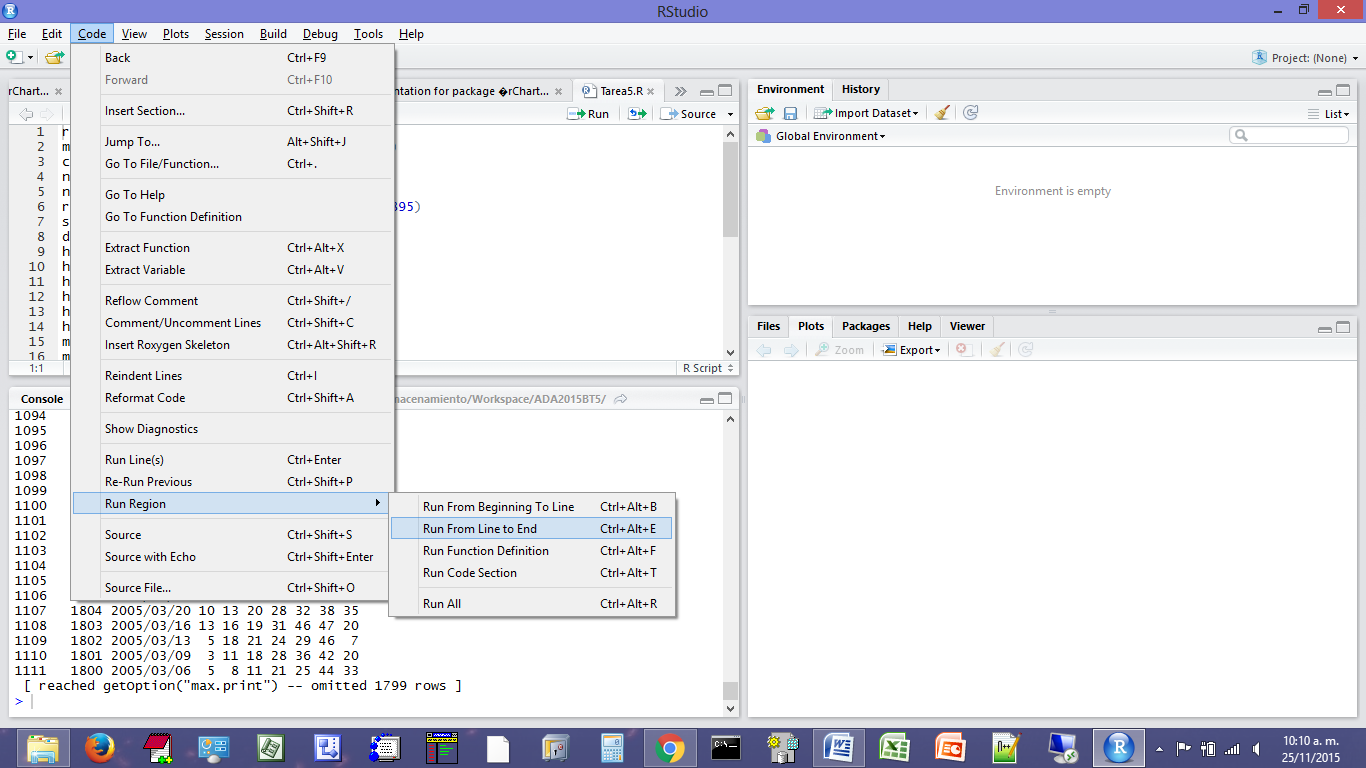
> max(melate[,"A"])

[1] 56

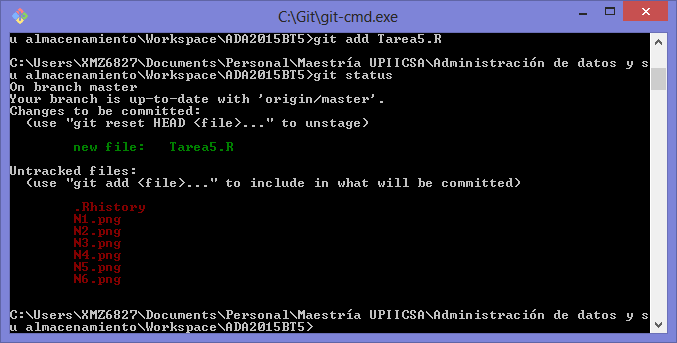
1. Desde el menú de **RStudio**, cree un nuevo *script* **R** (**File** -> **New File** -> **R Script**) en el que colocará las instrucciones manualmente probadas.



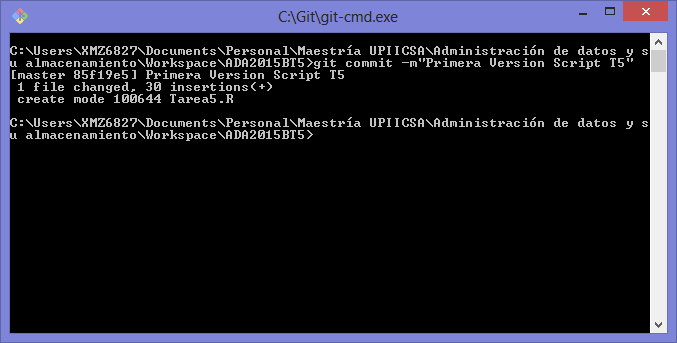
1. Ejecute el archivo para validar su funcionamiento.



1. Cuando considere que su script está terminado agréguelo  a la lista de monitoreo de git y sincronícelo con su repositorio. Tome evidencia del proceso.

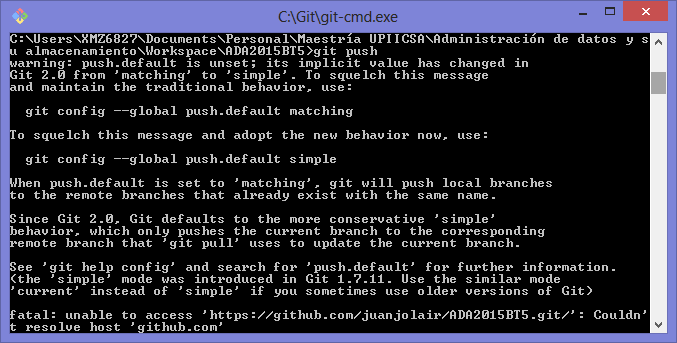


git add Tarea5.R

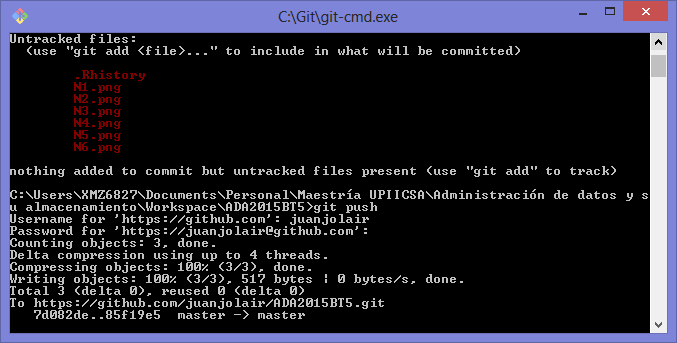


git commit –m”Primera Version Script T5”

Al momento de utilizar el comando “git push” nos marcó el error de que era necesario previamente configurar el valor de “simple” por “matching” con el comando git config –global push.default matching

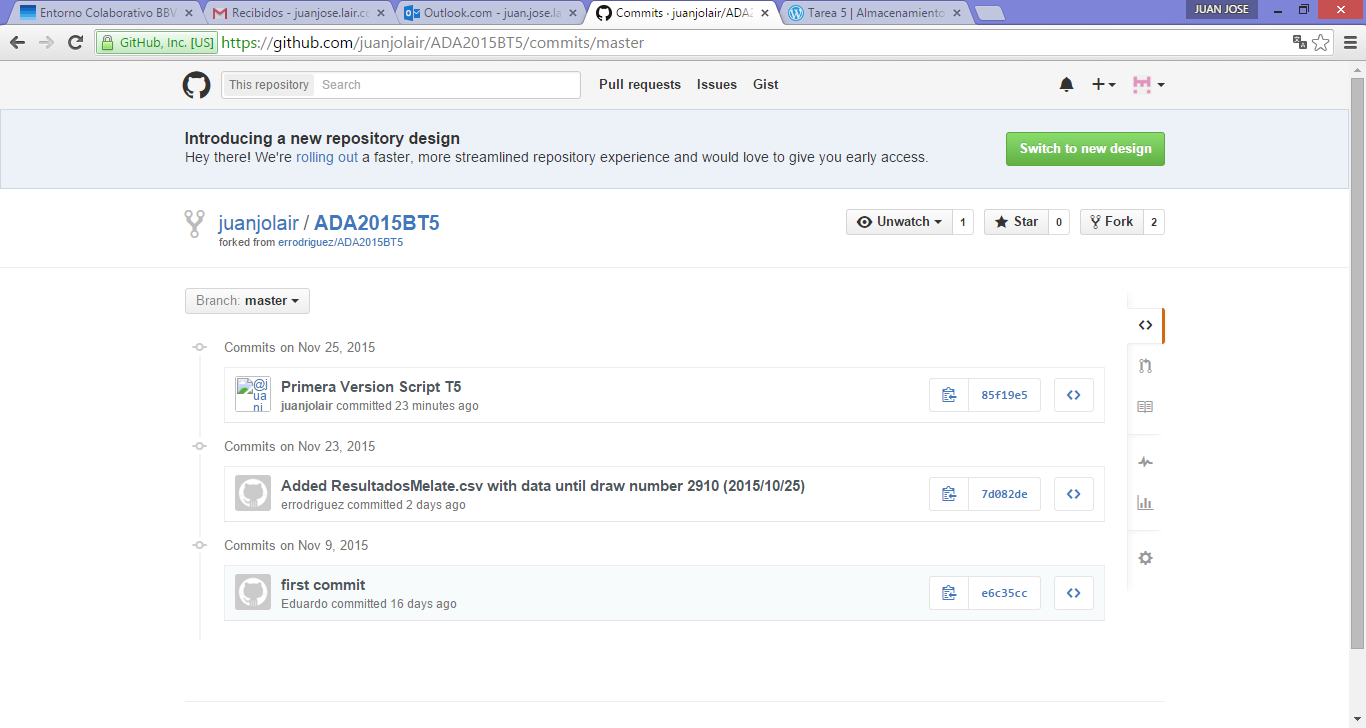


git config –global push.default matching

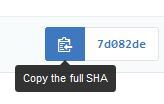


git push

1. De *click* en su contador de *commits*



1. y *de la página resultante copie el*hash*SHA del último*commit*de su*script*y documéntelo en su trabajo*



85f19e5f62539a36641b3f6c75e1987960720a03

<https://github.com/juanjolair/ADA2015BT5/blob/85f19e5f62539a36641b3f6c75e1987960720a03/Tarea5.R>

1. Exprese sus observaciones, opiniones o conclusiones sobre el archivo explorado y documente su *script* en el documento a presentar.

El archivo es un conjunto de datos que representa a los sorteos realizados por Melate los días domingos y miércoles (desde su funcionamiento) desde el 19 de agosto de 1984 hasta el 25 de octubre del presente año. Dicta las combinaciones ganadoras de cada sorteo, desde el número 1 hasta el 2910. Cada número de las 6 (N1, N2, N3, N4, N5, N6) posibilidades que existen más el adicional (A).

Dentro de cada uno de los histogramas, podemos ver la frecuencia que se da por cada número del 1-56 en cada esfera (las cuales han ido aumentando de 40 hasta 56) y observamos que para la primera esfera el rango mayor de posibilidad se encuentra en los primeros números decreciendo hasta el 35, para la segunda esfera observamos que el rango mayor se encuentra en el 10. Para la tercera esfera se encuentra en el rango del 15 al 20; así como para la cuarta esfera la media aritmética es entre el 27 y 28. Con respecto a la quinta esfera, podemos observar que los valores entre 30 al 40 son los más altos dentro del histograma y; por último, para la sexta esfera dentro del 35 al 45, siendo el 41 su media.

**Script**

read.csv("ResultadosMelate.csv")

melate <- read.csv("ResultadosMelate.csv")

colnames(melate)

nrow(melate)

ncol(melate)

read.csv("ResultadosMelate.csv", skip = 2895)

summary(melate)

dim(melate)

hist(melate[,"N1"])

hist(melate[,"N2"])

hist(melate[,"N3"])

hist(melate[,"N4"])

hist(melate[,"N5"])

hist(melate[,"N6"])

min(melate[,"N1"])

max(melate[,"N1"])

min(melate[,"N2"])

max(melate[,"N2"])

min(melate[,"N3"])

max(melate[,"N3"])

min(melate[,"N4"])

max(melate[,"N4"])

min(melate[,"N5"])

max(melate[,"N5"])

min(melate[,"N6"])

max(melate[,"N6"])

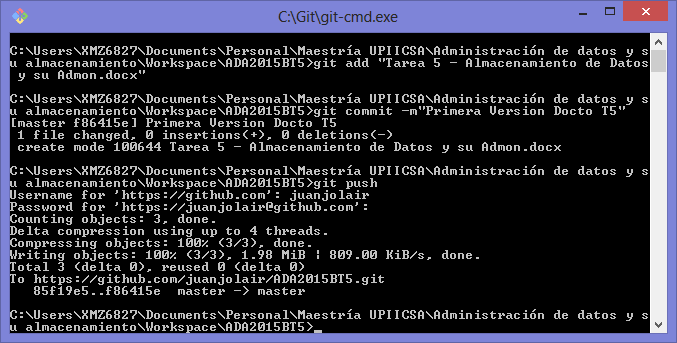
min(melate[,"SORTEO"])

max(melate[,"SORTEO"])

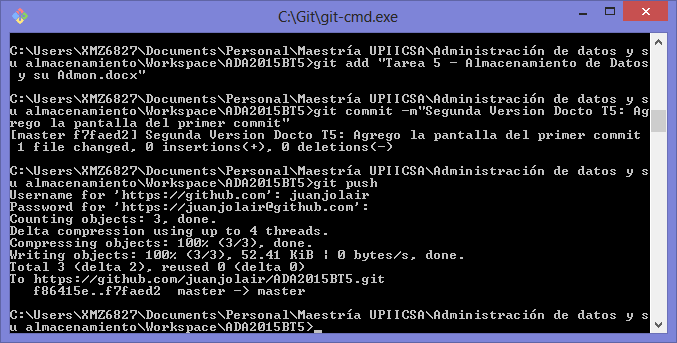
min(melate[,"A"])

max(melate[,"A"])

1. Sincronice  el documento que presentará con el repositorio. Indique que es el 1er *commit* en su texto de sincronización.

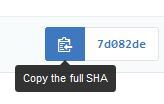


1. Agregue el **URL** de su repositorio al documento como una referencia bibliográfica y sincronice de nuevo (2do *commit*). Indique las razones en su mensaje de sincronización.



BIBLIOGRAFÍA

URL: 2º commit



f7faed2e5d619188b906b9c334a606f367802bdd

<https://github.com/juanjolair/ADA2015BT5/commit/f7faed2e5d619188b906b9c334a606f367802bdd#diff-106321ba2ef5a3a41d441658cd4f30df>