|  |
| --- |
| **BITÁCORA TÉCNICA DEL CURSO** |
| Administración de datos y su almacenamiento |
|  |
| Memoria técnica de todo aquello de lo que se hizo uso en el aspecto aplicativo del curso; sea por una tarea, sea por un ejercicio o algo que expresamente se debió de conocer. |
|  |
| **Juan José Lair Martínez** |
|  |
|  |

ÍNDICE

[INTRODUCCIÓN 3](#_Toc440894947)

[INSTALACIÓN DE RSTUDIO y Lenguaje R 4](#_Toc440894948)

[INSCRIPCIÓN A GITHUB 11](#_Toc440894949)

[INSTALACIÓN GIT 13](#_Toc440894950)

[INSTALACIÓN VIRTUALBOX 5.0.8. 16](#_Toc440894951)

[INSTALACIÓN CENTOS 7 22](#_Toc440894952)

[INSTALACIÓN HORTONWORKS DATA PLATFORM SANDBOX V.2.3 26](#_Toc440894953)

[ANEXO 29](#_Toc440894954)

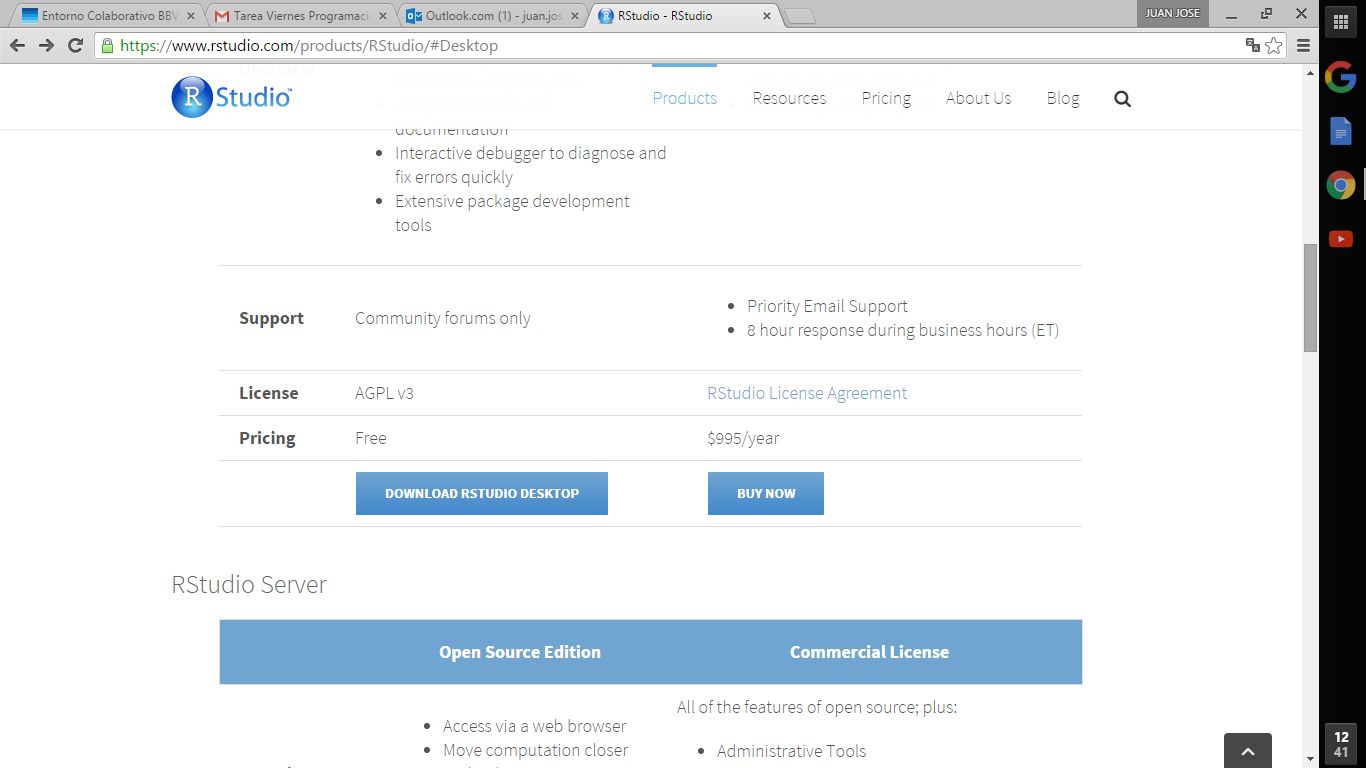
# INTRODUCCIÓN

A lo largo del curso de “*Almacenamiento de datos y su administración*”, se ha trabajado con varias herramientas para las distintas tareas elaboradas. Por ello, se ha generado esta bitácora que contiene las diferentes instalaciones de todos aquellos programas utilizados para este fin.

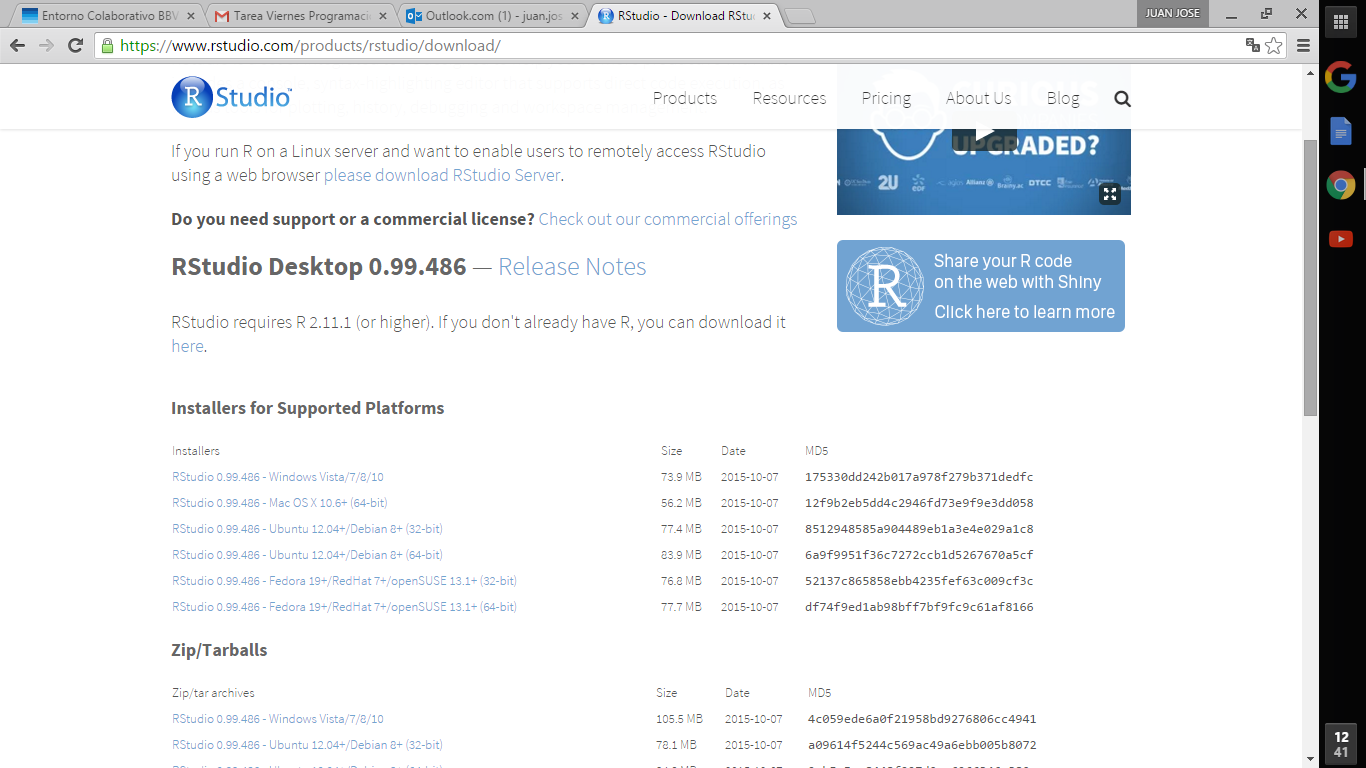
Cada una de ellas, viene desglosada en este material para dejar la evidencia necesaria, en caso de ser requerida, para posteriores instalaciones o usos de los mismos programas; ya que, en ocasiones, por el mismo tiempo de no utilizarse, se olvidan algunos detalles, incidencias o sitios de ayuda o descarga y es mejor conservar los pasos para poder llevar a cabo una instalación y, en caso de problemas o alguna incidencia o detalle en la misma instalación, documentarla para tener un soporte para su posterior uso y/o solución.

# INSTALACIÓN DE RSTUDIO y lenguaje r

1. Desde la página de **RStudio**<https://www.rstudio.com/products/rstudio/#Desktop> se procedió a descargar la versión para escritorio y se elije la versión para el sistema operativo, en nuestro caso, la versión para **Windows.**

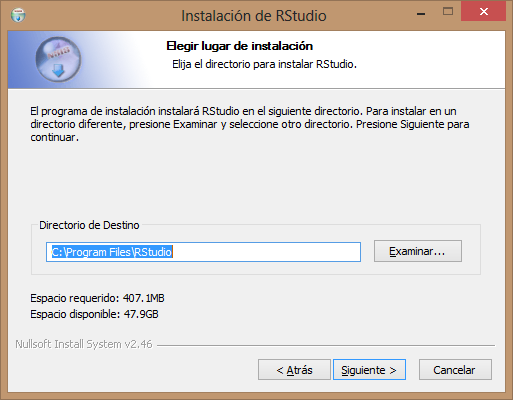
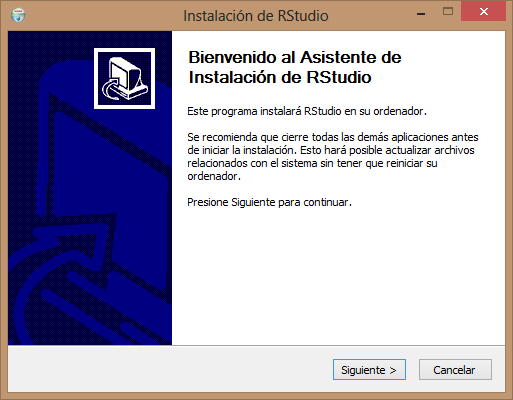


Descarga RStudio



Se elije el Sistema Operativo y se guarda el archivo

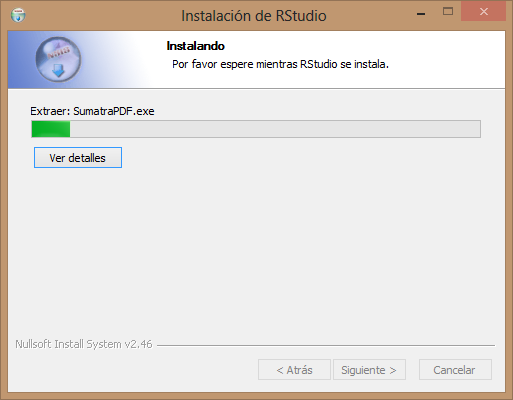
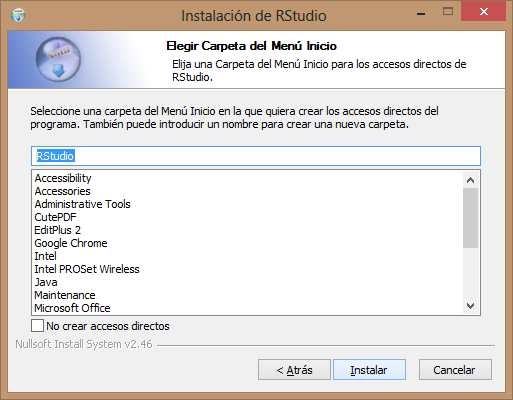
1. Una vez descargado el archivo y guardado en la carpeta elegida, se procede a ejecutar el archivo **RStudio-0.99.486.exe** y se elije el lugar de la instalación.



Se elije el lugar de la instalación

Se ejecuta el archivo descargado RStudio-0.99.486.exe

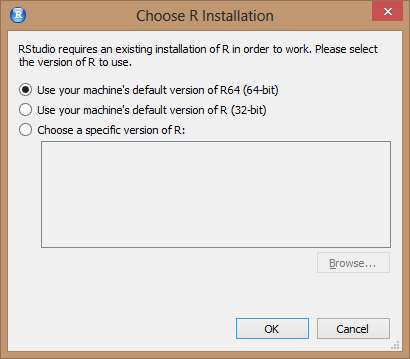
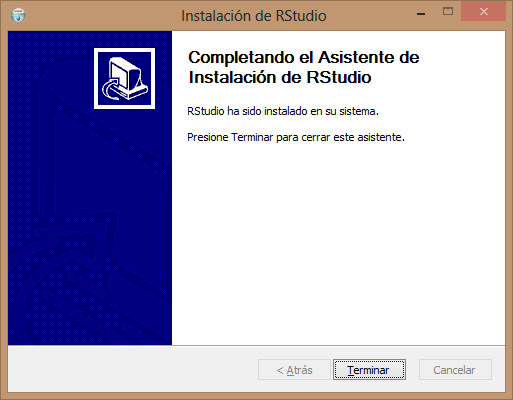
1. Se sigue con el progreso de la instalación eligiendo la carpeta del menú de inicio y se procede a instalar.



Progreso de la instalación

Se elije la carpeta del Menú de Inicio

1. Terminada la instalación, se indica que se requiere, previamente, la instalación del **lenguaje R**, por lo que es necesario descargarlo para continuar con la instalación de **RStudio**.



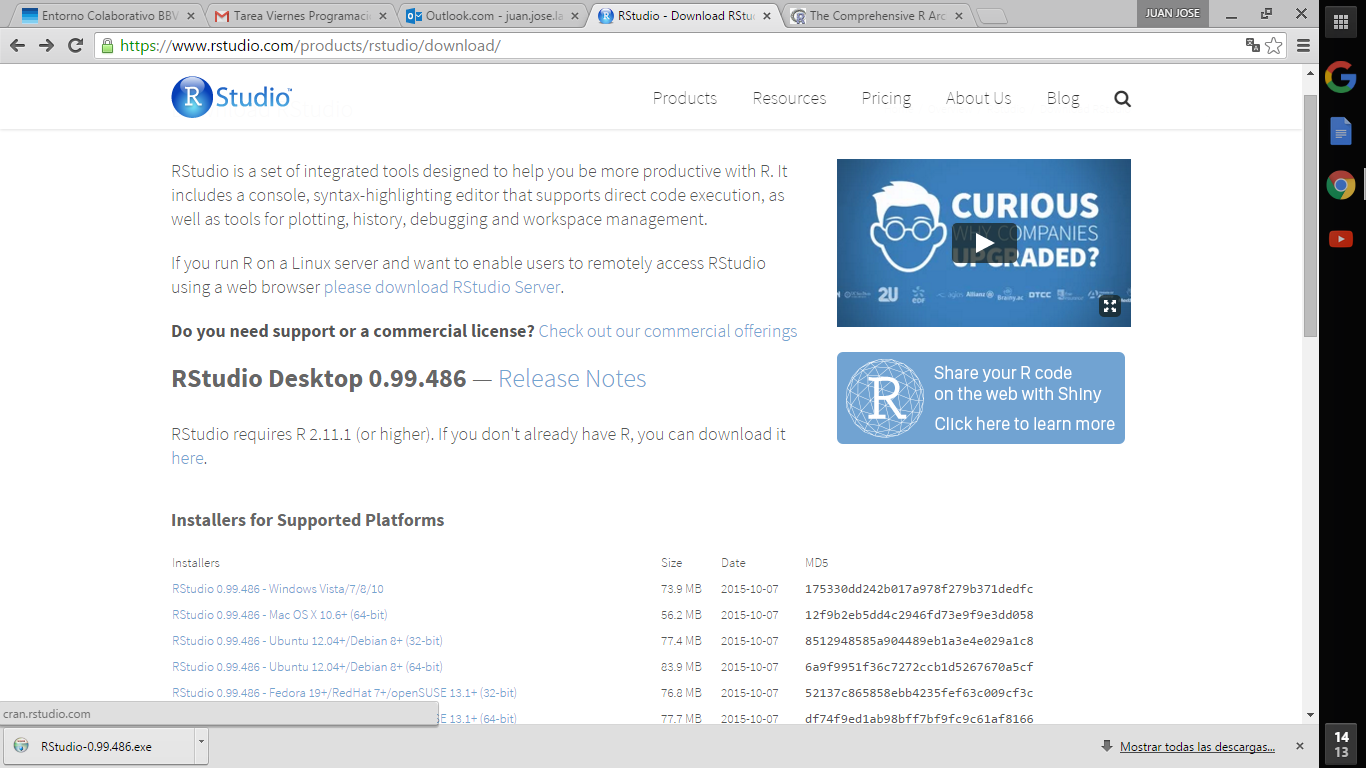
Se requiere una instalación existente de R y se selecciona para la versión específica de la PC

Se elije Terminar para completar la instalación



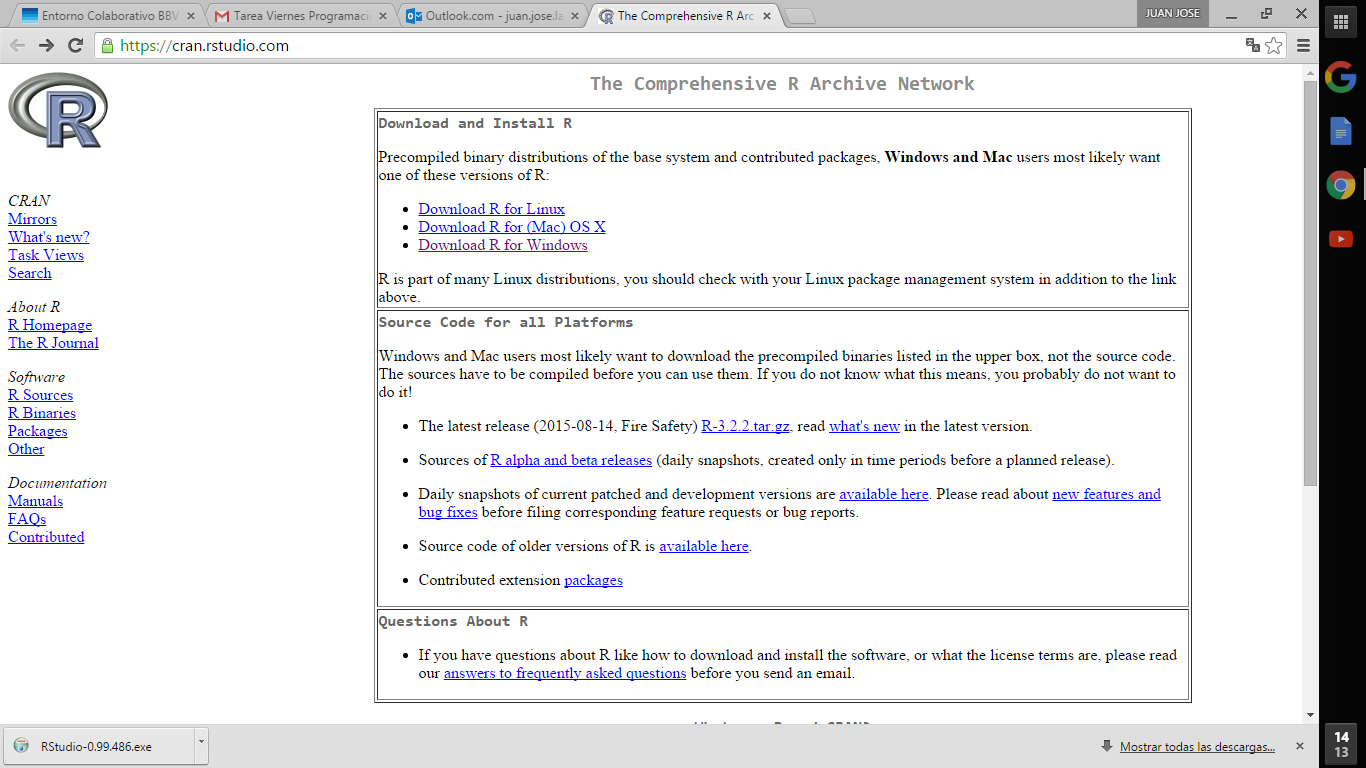
No se cuenta con una instalación previa, por lo que se procede a descargarla

1. **RStudio** requiere el *lenguaje R 2.11.1* o posterior, por lo que se procede al sitio de la descarga.

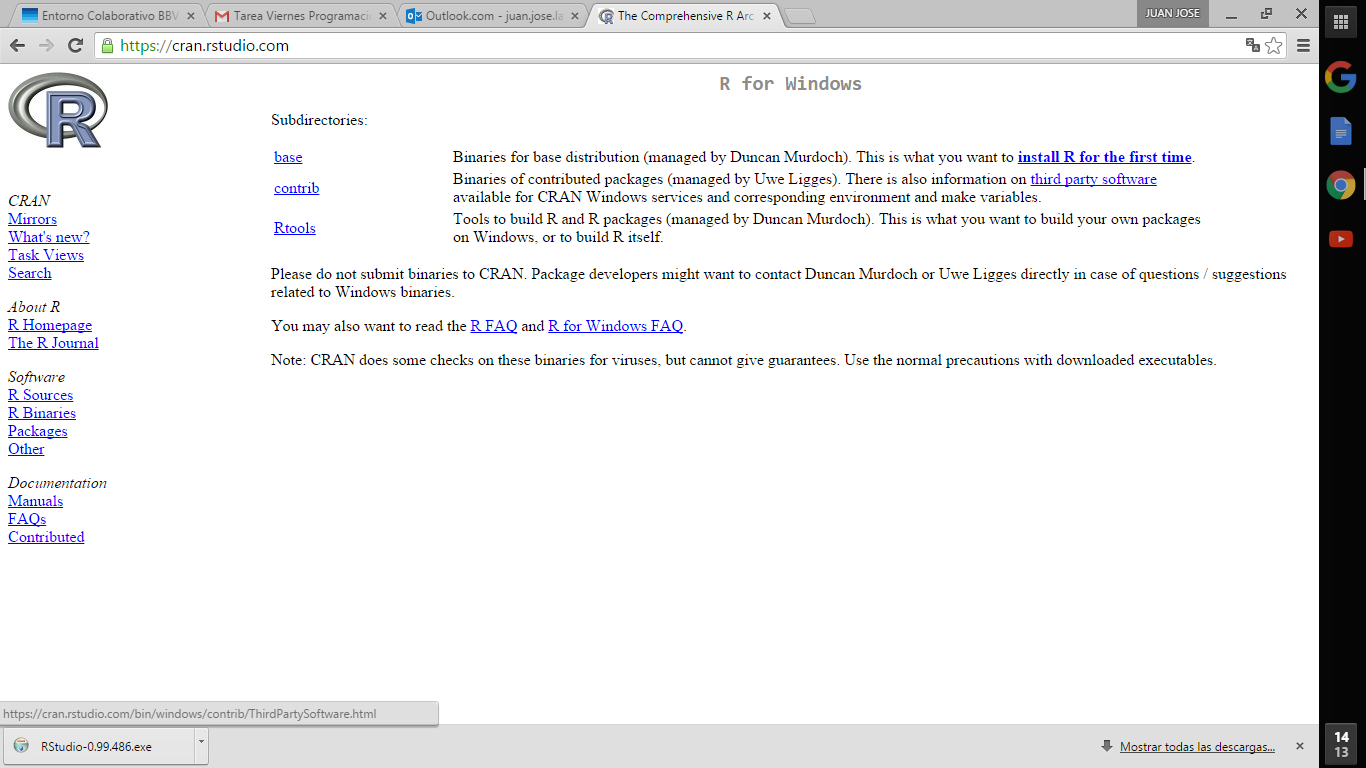


Se necesita previamente instalar R 2.11.1; por lo que se descarga e instalar

1. En la página <https://cran.rstudio.com/> se elije la opción del sistema operativo, para nuestro caso **Windows**; por lo que se procede a su descarga y se escoge *“install R forthefirst time”.*

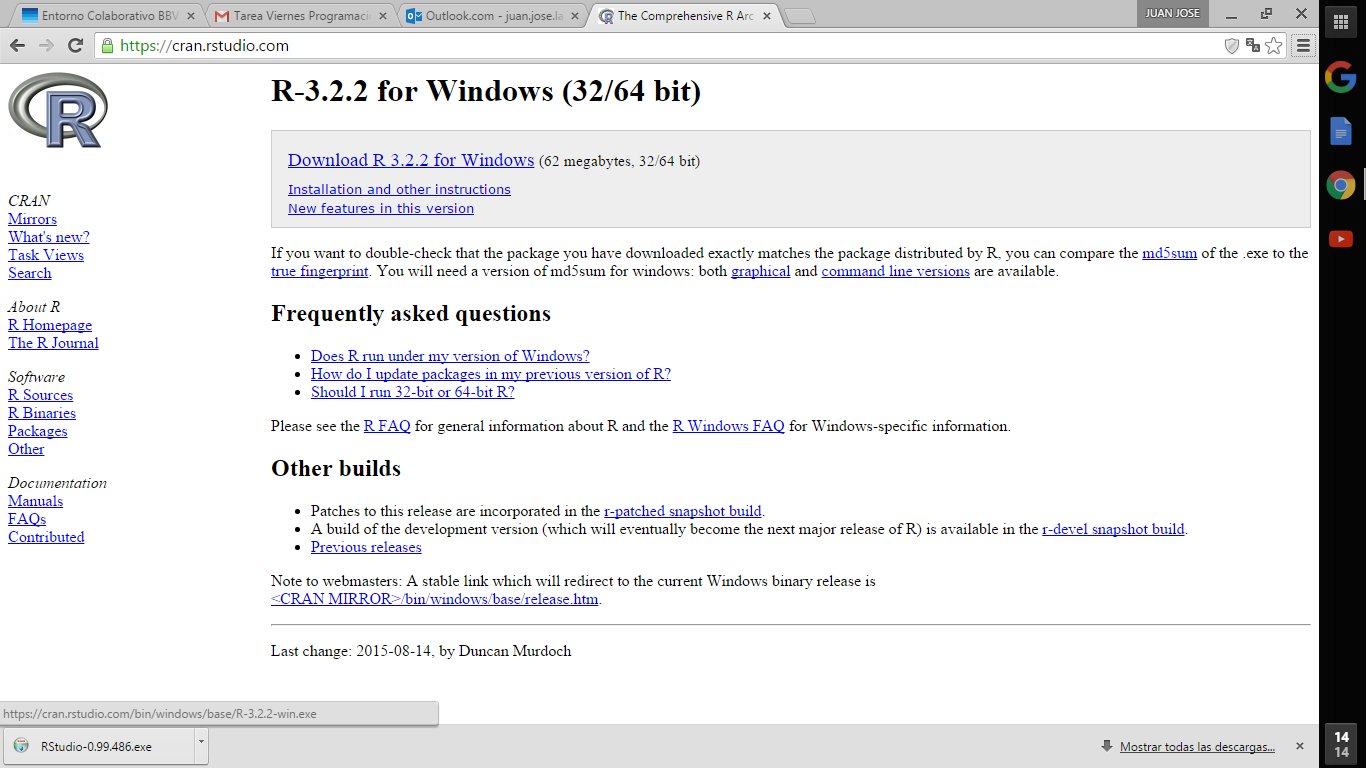


Se descarga la versión para el SO que se tenga



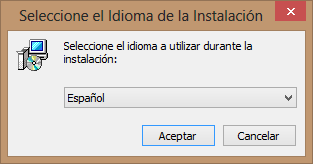
Se procede a su instalación por primera vez

1. Se procede a descargar la última versión que, al momento de la realización de éste documento, era la versión de **R-3.2.2. para Windows 32/64 bits**.

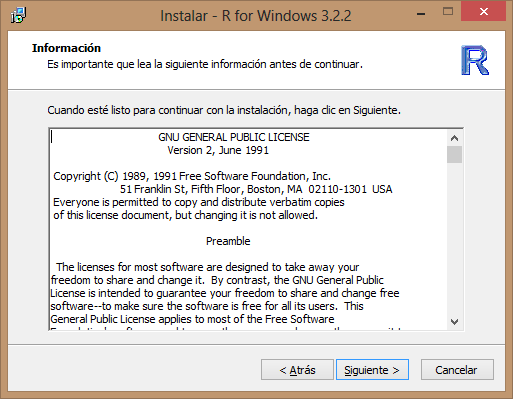
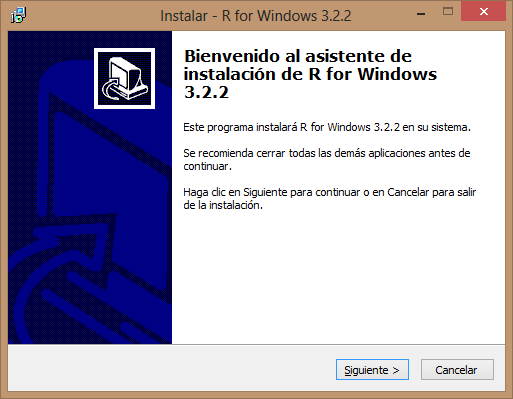


Se procede a descargar la versión de R para el SO que se tiene

1. Descargado el archivo y almacenado, se procede a su instalación seleccionando el lenguaje (para nuestro caso, *español*) y se continúa aceptando la licencia y su ruta.



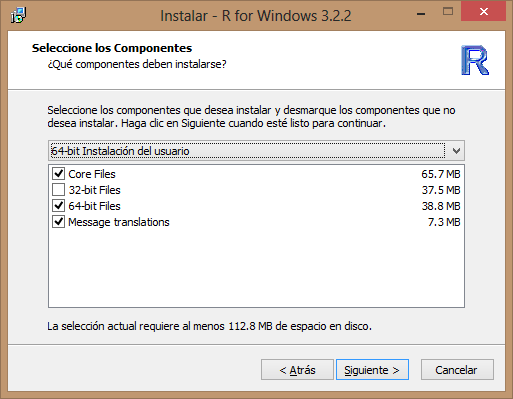
Una vez descargado el programa R-3.2.2-win, se procede a su ejecución. Se elije el lenguaje a utilizar durante la instalación



Se acepta la licencia para instalar

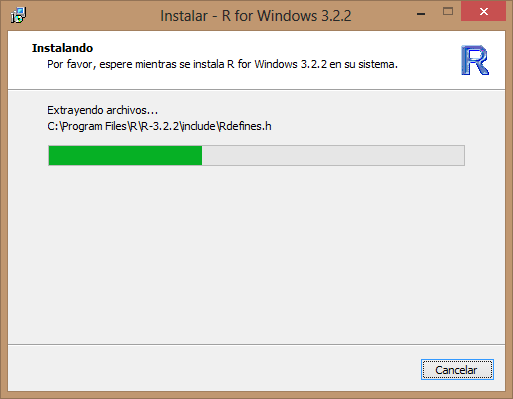
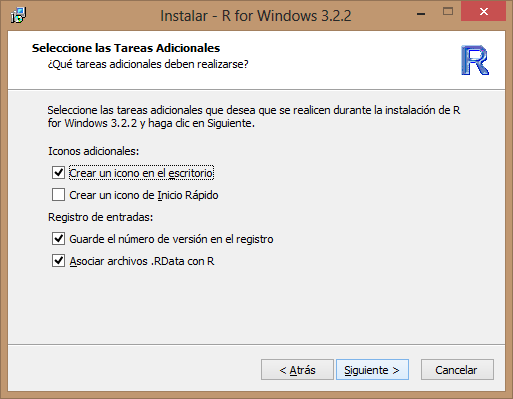
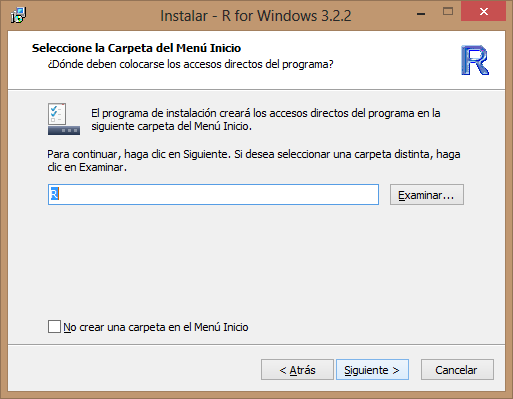
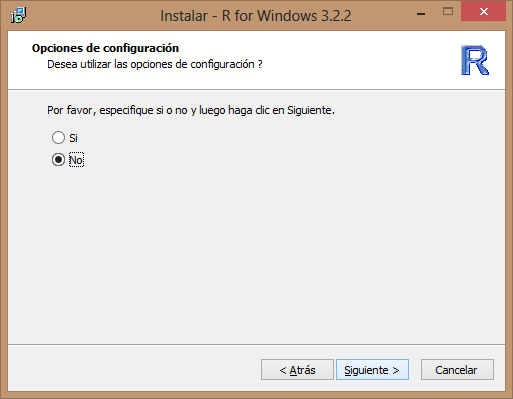
Se procede a su instalación

1. Elegida la ruta, se elijen también los componentes para la versión de nuestro sistema operativo, en nuestro caso **Windows a 64 bits** y las opciones por *default*.



Se elije la ruta en dónde se instalará

Se elijen los componentes a instalar



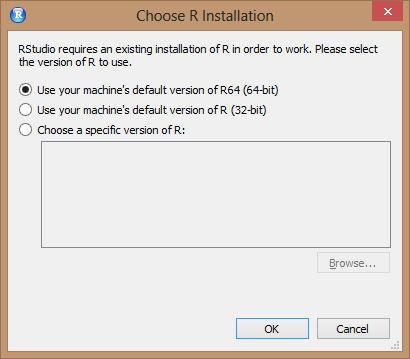
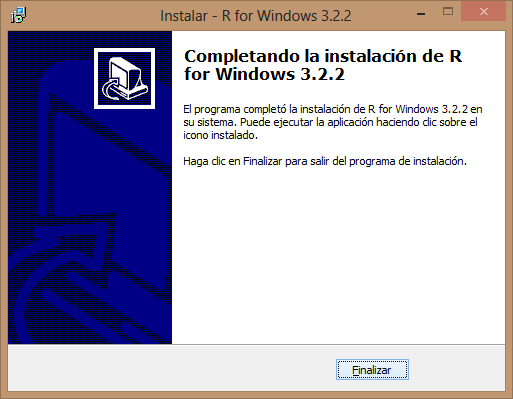
Se finaliza la instalación del lenguaje R

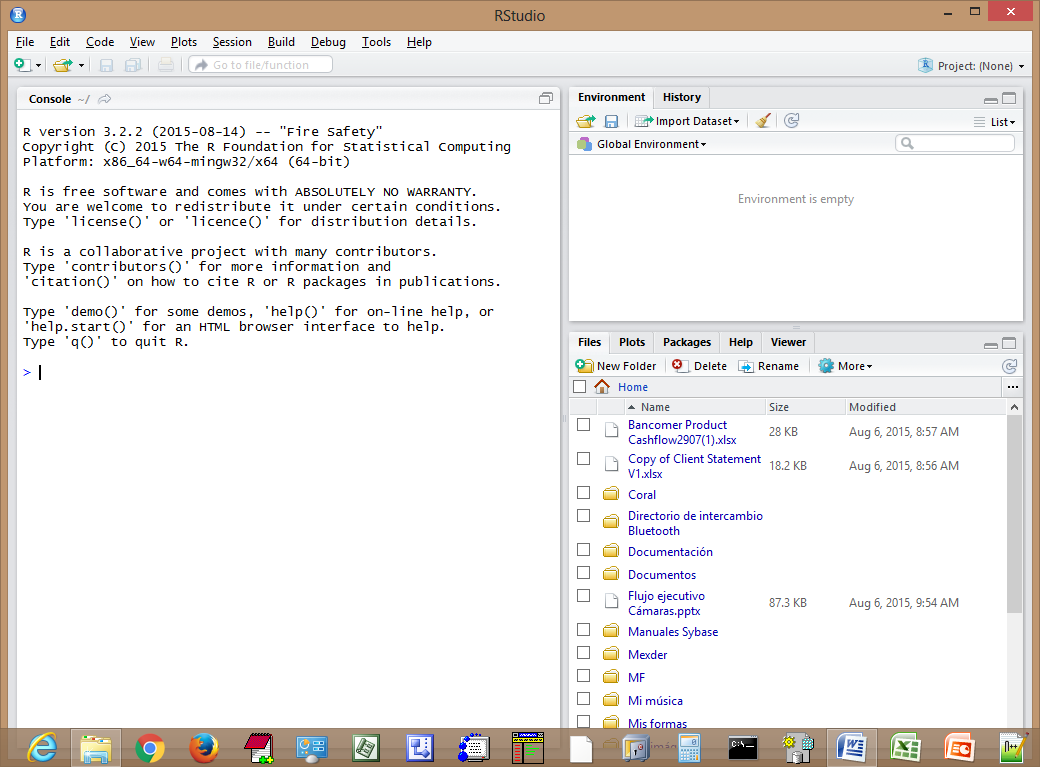
Se selecciona la carpeta del menú de inicio

Se elije la opción por *default* que tiene el programa de instalación

Finalizado todos los pasos, se muestra la consola de RStudio

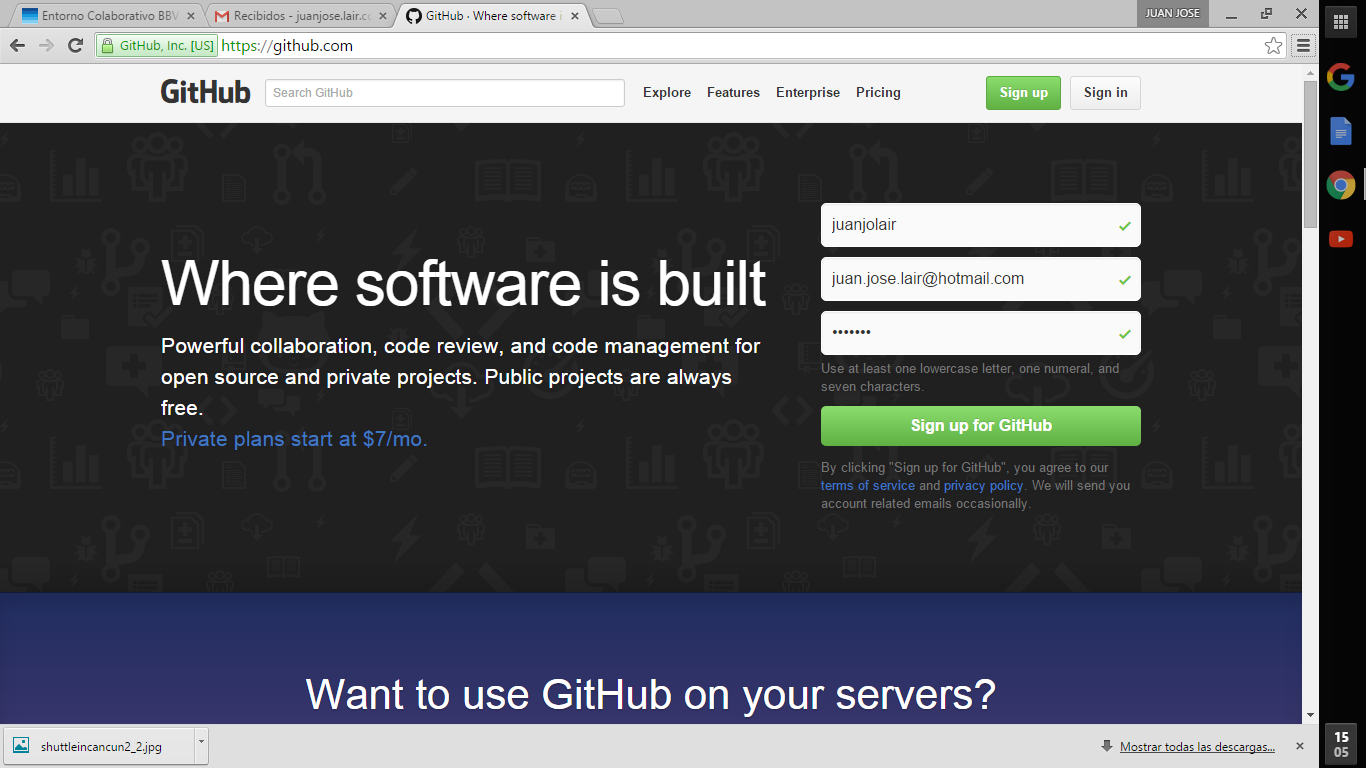
1. Completada la instalación del **lenguaje R**, se continúa con la instalación de **RStudio** eligiendo la versión necesaria, nuestro caso, a *64 bits* y se termina con la instalación.



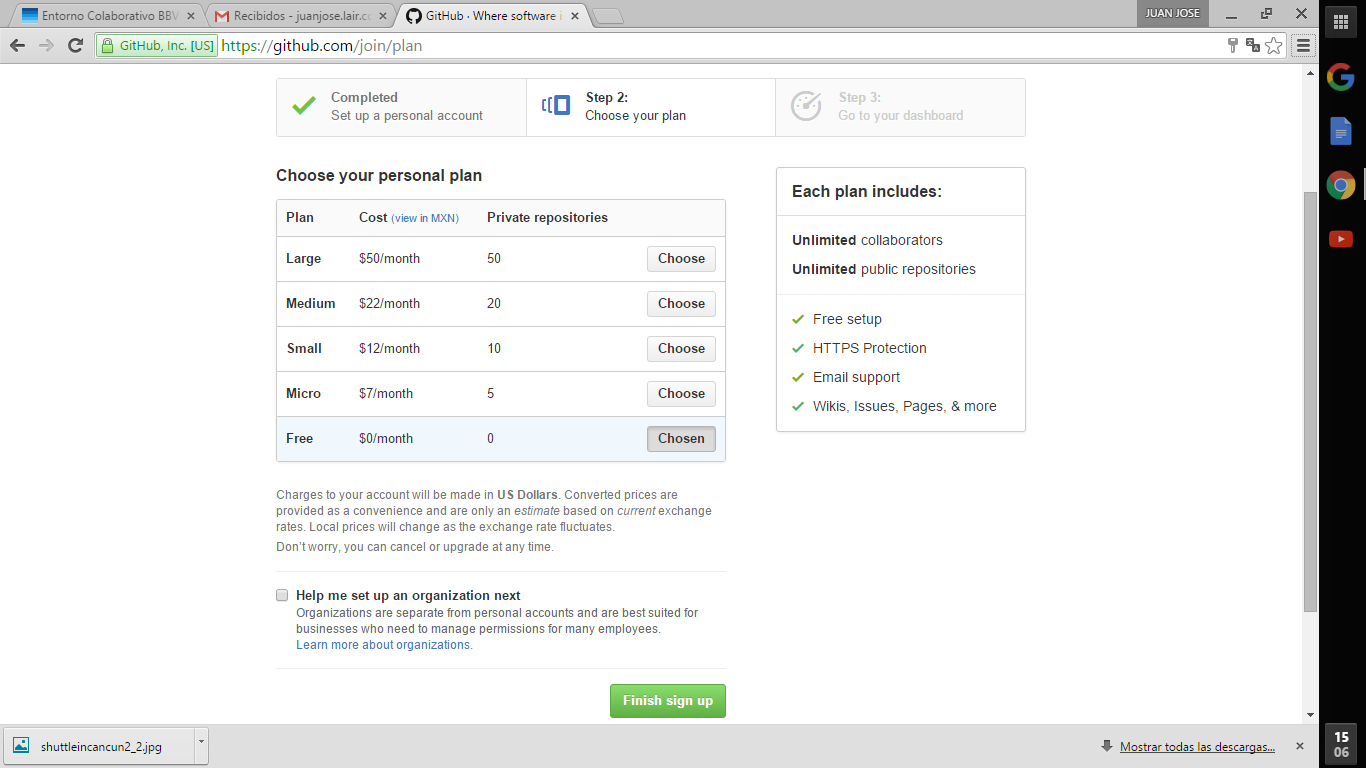


# inscripción a github

1. Desde la página de **GITHUB**<https://github.com/> se procede a crear el usuario proporcionando un *username*, correo electrónico y generando un *password*. Posteriormente se elige el plan personal de pago, para nuestro caso se escogió el plan gratuito.

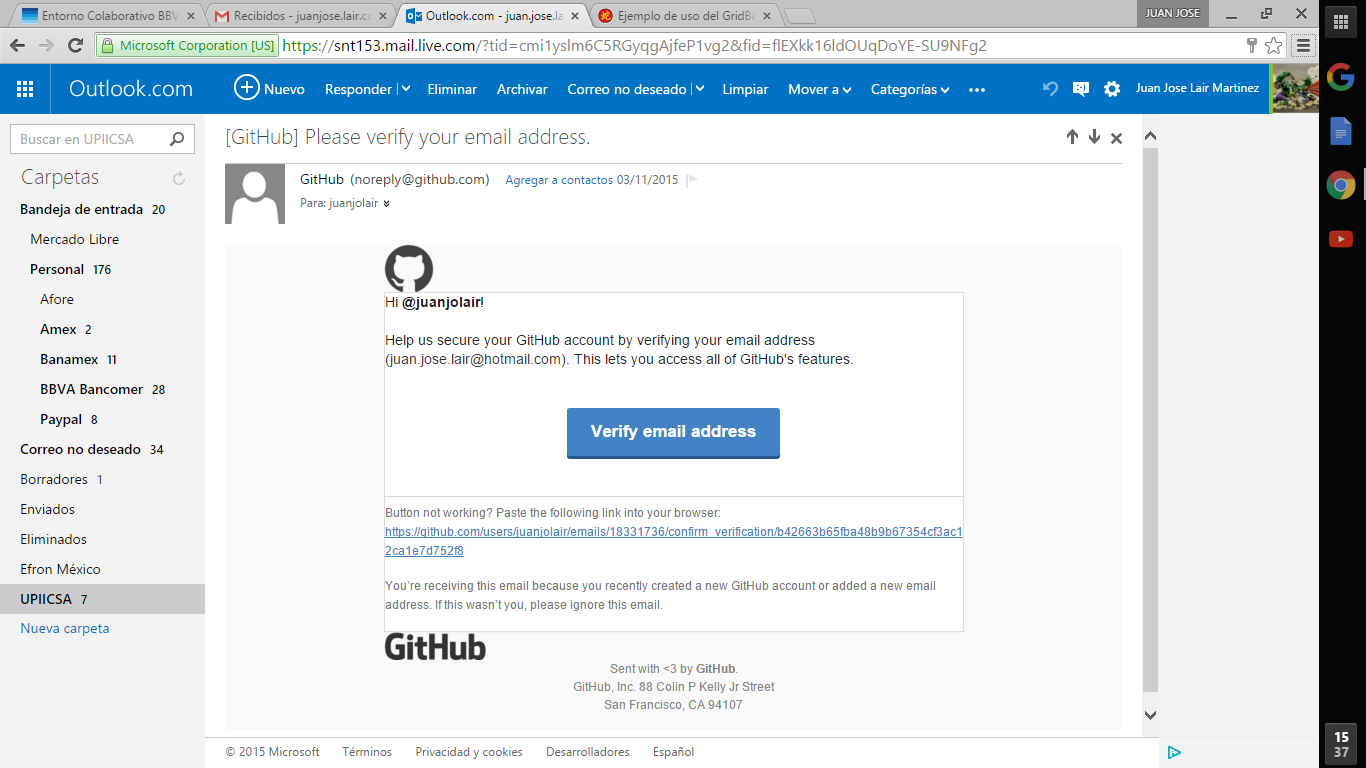


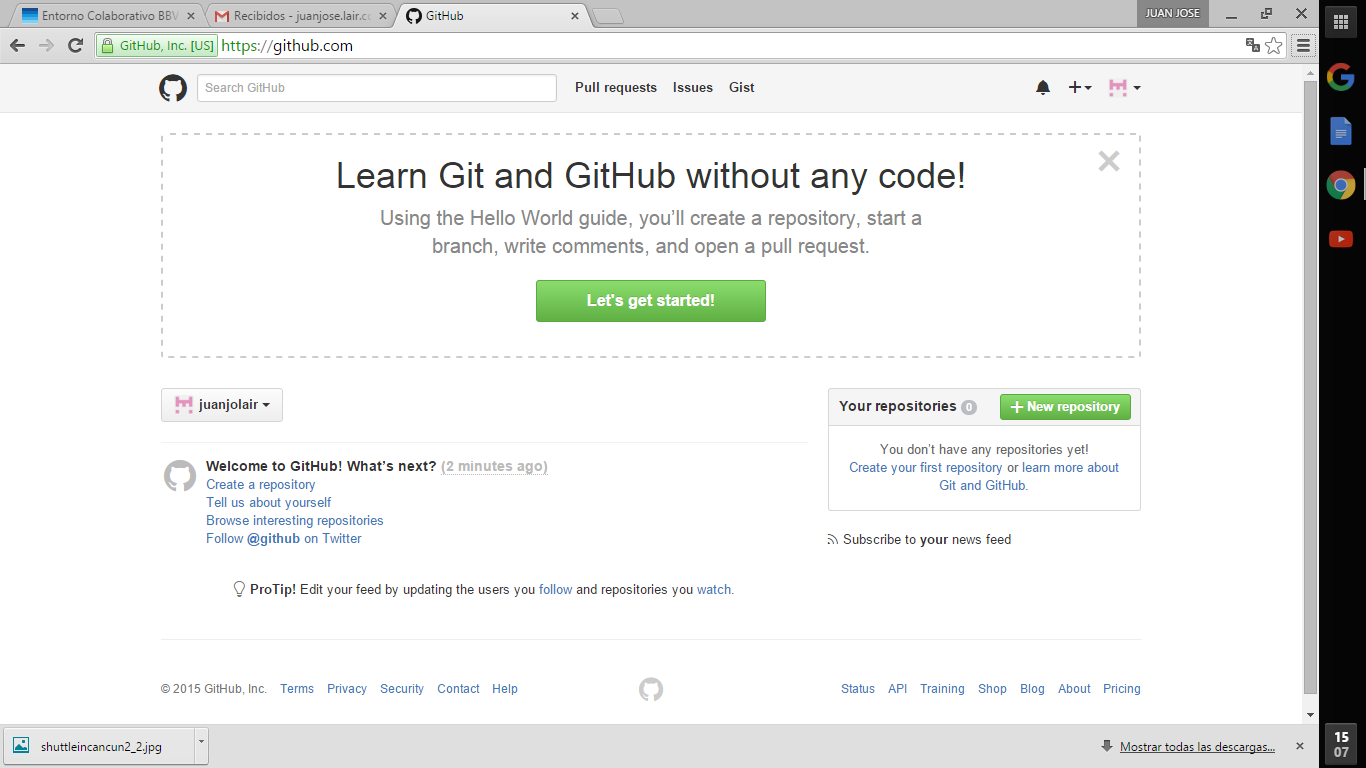
Para firmarse, primeramente se debe de realizar la creación del login



Posteriormente, se elige plan de pago o gratuito

1. Por último, se verifica el correo electrónico proporcionado en la creación del usuario yendo al repositorio del correo y se encuentra listo para utilizarse.





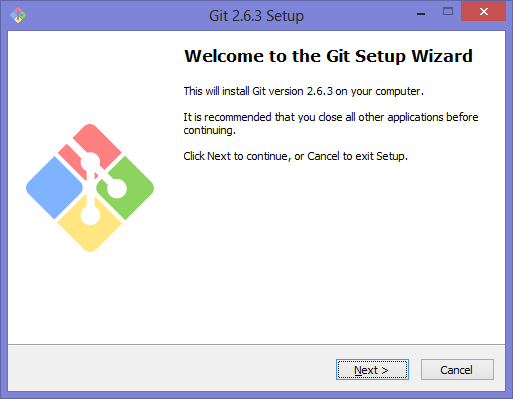
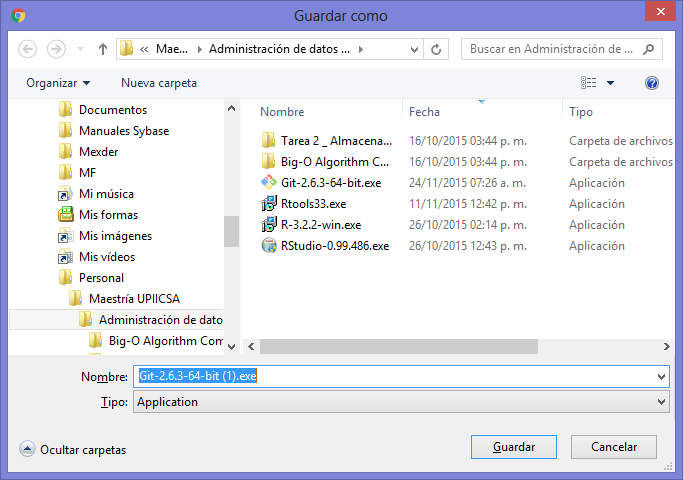
Se verifica el email ingresado y se encuentra listo para utilizarse

# INSTALACIÓN git

1. Desde la página de **GIT**<https://git-scm.com/> se procede a descargar la versión para *Windows* y guardarla en la carpeta de preferencia. Una vez descargada, se procede a la ejecución del archivo almacenado.

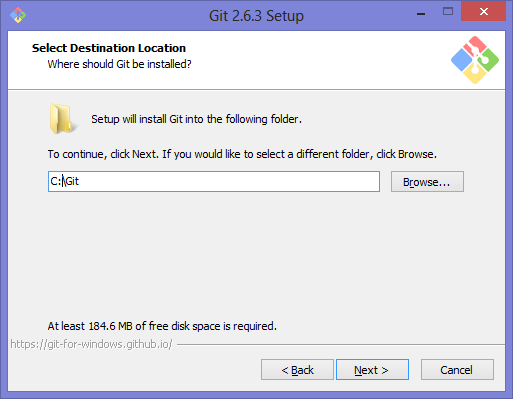


Se descarga la versión para Windows 64

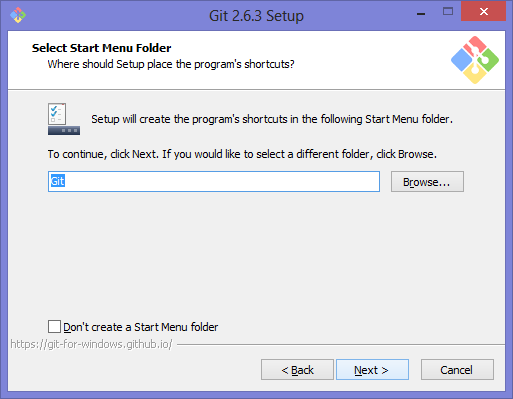
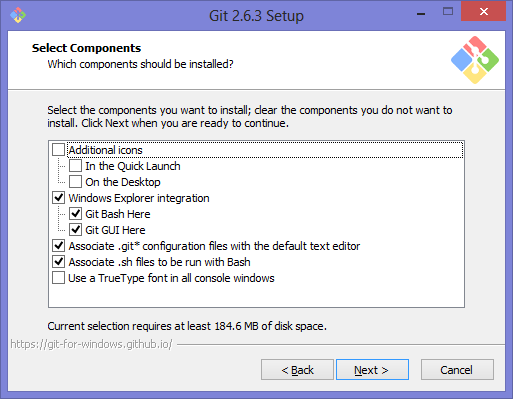


Se guarda y se procede a ejecutar

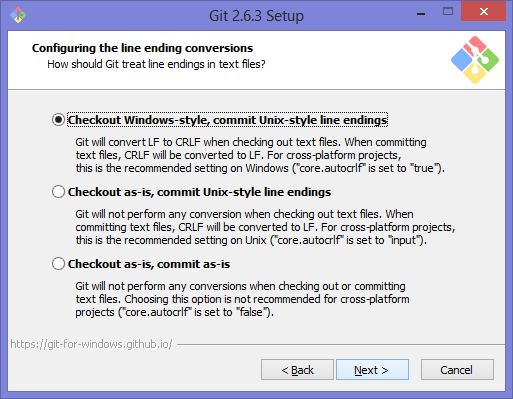
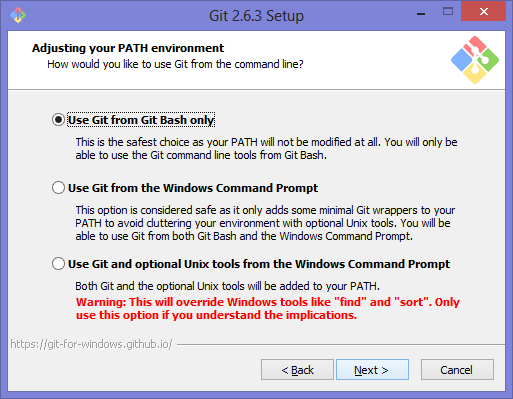
1. Se continua con la aceptación de la licencia, carpeta destino de la ruta en donde se instalará el programa y demás instrucciones tecleando “*next*”.



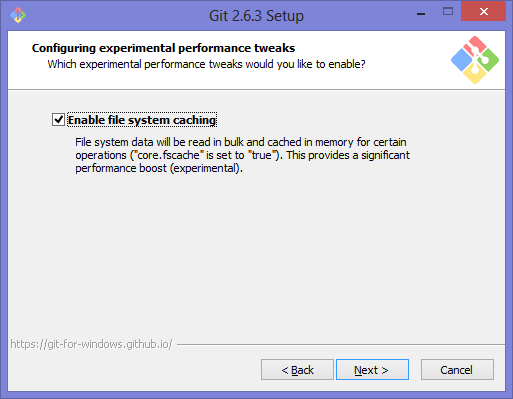
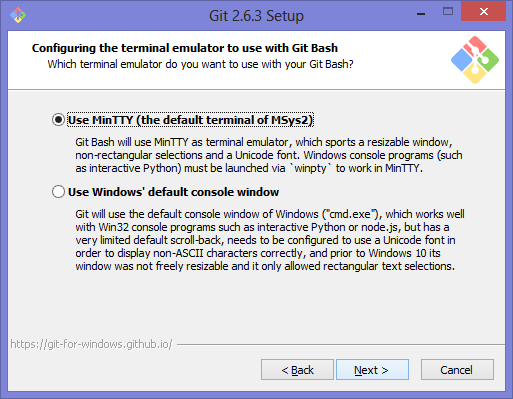
Se especifica la ruta en donde se instalará el programa



Se procede a todas las ventanas teclear “Next”

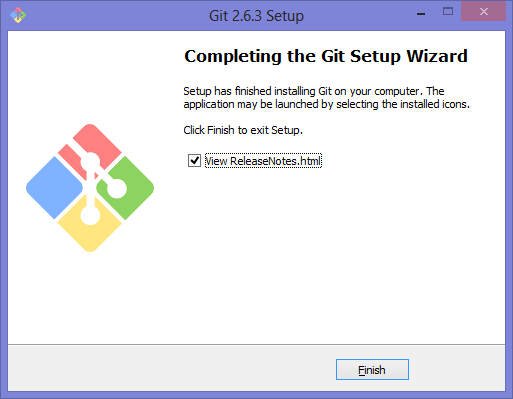
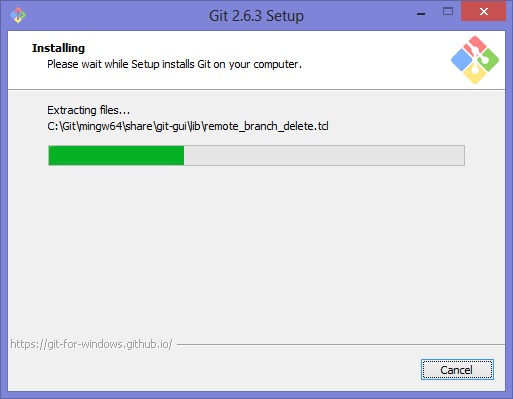


1. Se procede con la configuración de la terminal del emulador “por defecto”; así como la habilitación de la opción “*Enable file systemcaching*” para un mejor desempeño.

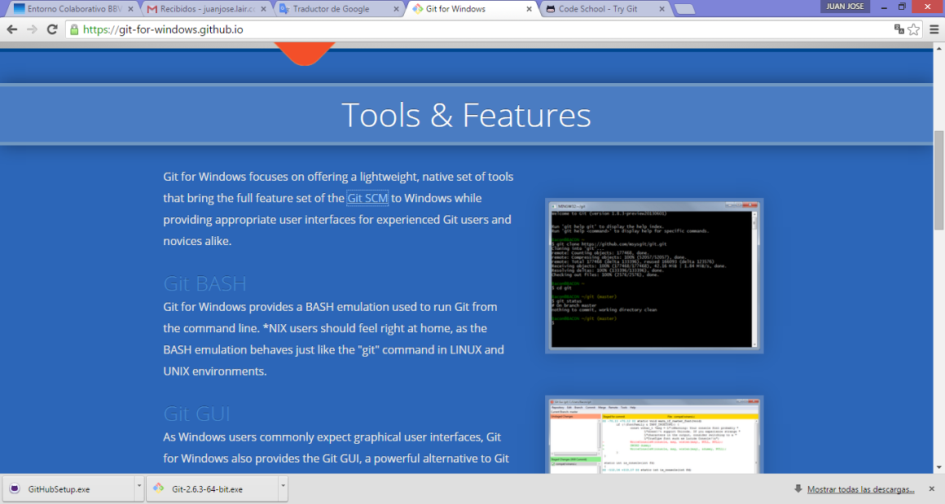


Se habilita la opción

1. Por último, se completa la instalación y se encuentra listo para utilizarse.



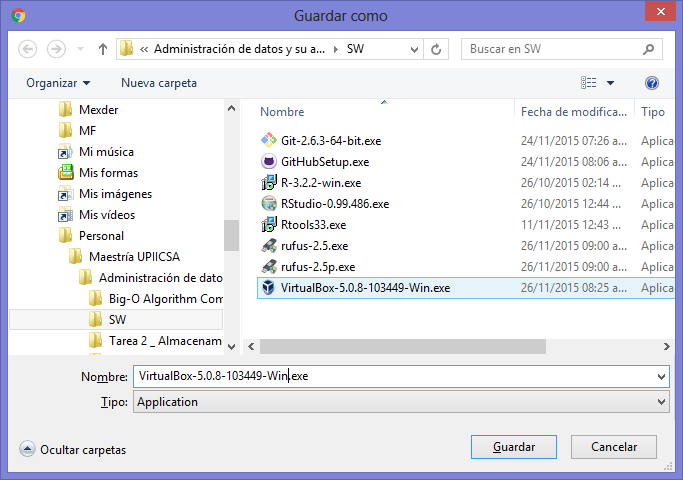
Fin de la instalación



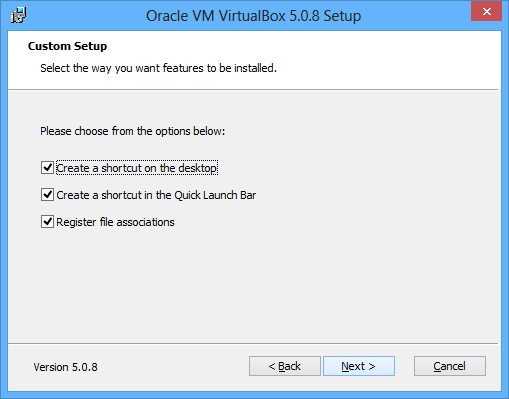
# INSTALACIÓN VIRTUALBOX 5.0.8.

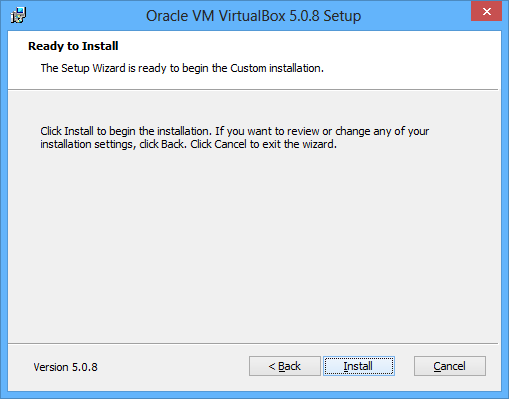
1. Desde la página de **VIRTUALBOX**<http://download.virtualbox.org/virtualbox/5.0.8/> se procede a descargar la versión para Windows y guardarla en la carpeta de preferencia. Una vez descargada, se procede a la ejecución del archivo almacenado.

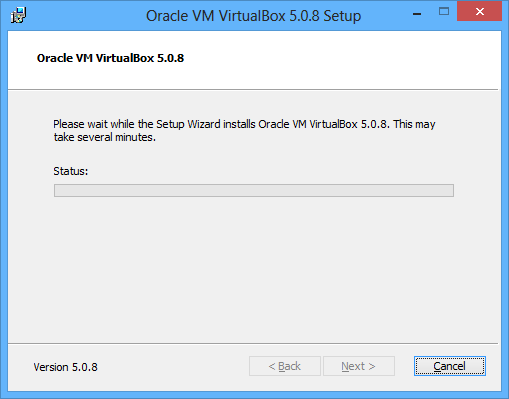




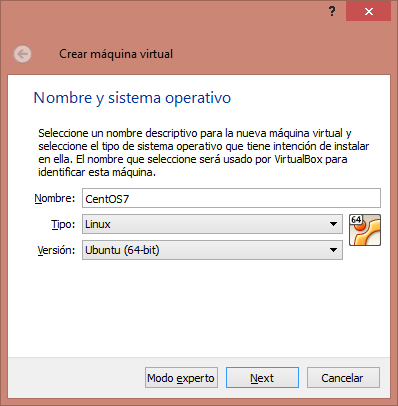
1. Se procede a elegir la ruta de instalación y se procedió a aceptar las advertenciasy otros avisos presentados hasta obtener la confirmación de instalación con éxito.



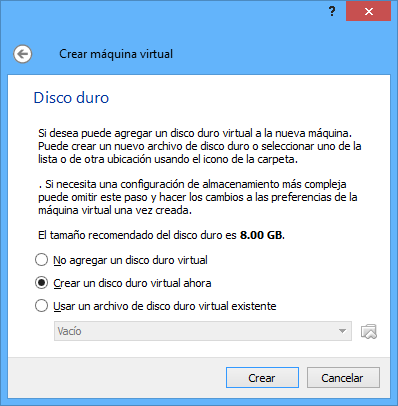




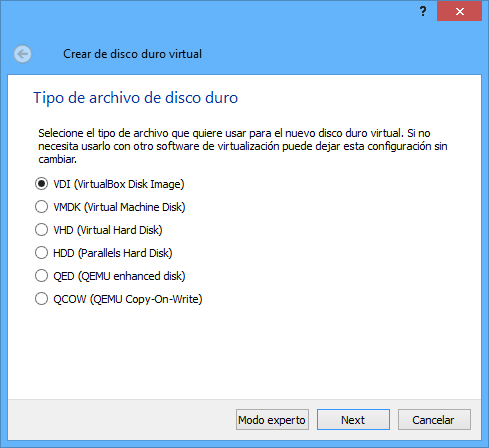
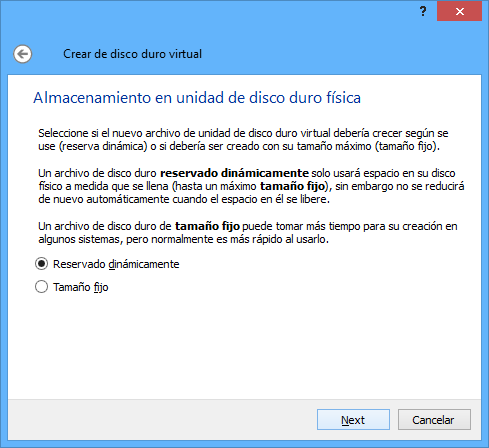
1. Una vez instalado el**VIRTUALBOX** se observa una ventana de bienvenida como la mostrada en la parte de abajo, del lado izquierdo, en donde no se muestra ninguna máquina virtual; por lo que se procede a crear una. En la barra de herramientas principal se da*click* en **NUEVA** y se proporciona un nombre y el tipo de sistema operativo. Para nuestro caso se pretende crear una máquina virtual de **LINUX**y la versión con **Ubuntu (64-bit)**.

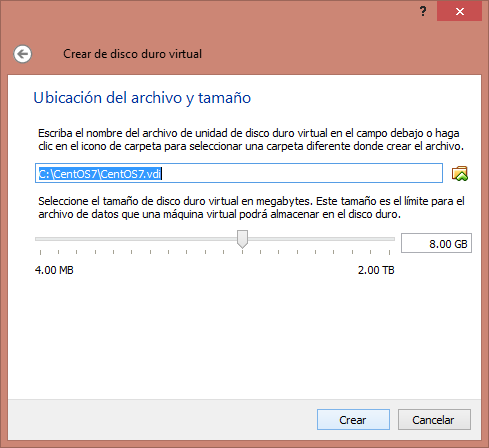
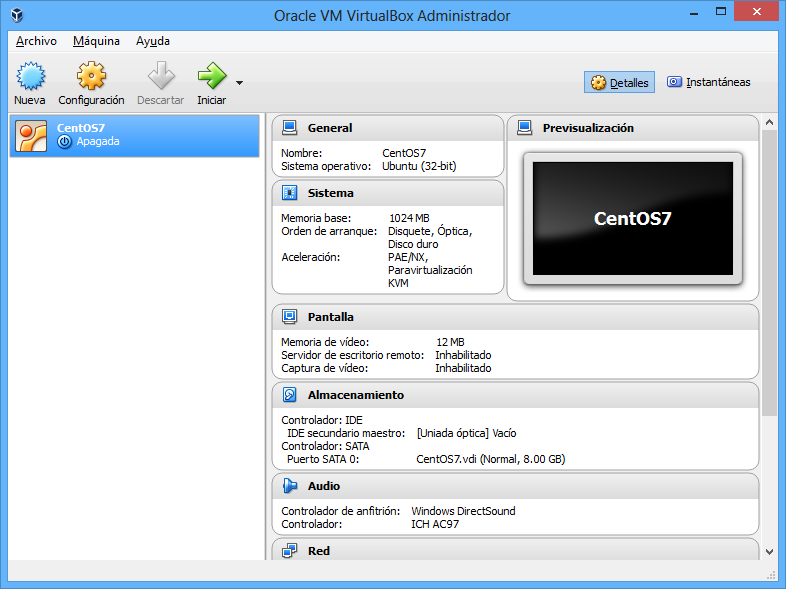
1. Posteriormente, se define el tamaño de la memoria, en nuestro caso, 1024MB y crear un disco duro virtual.

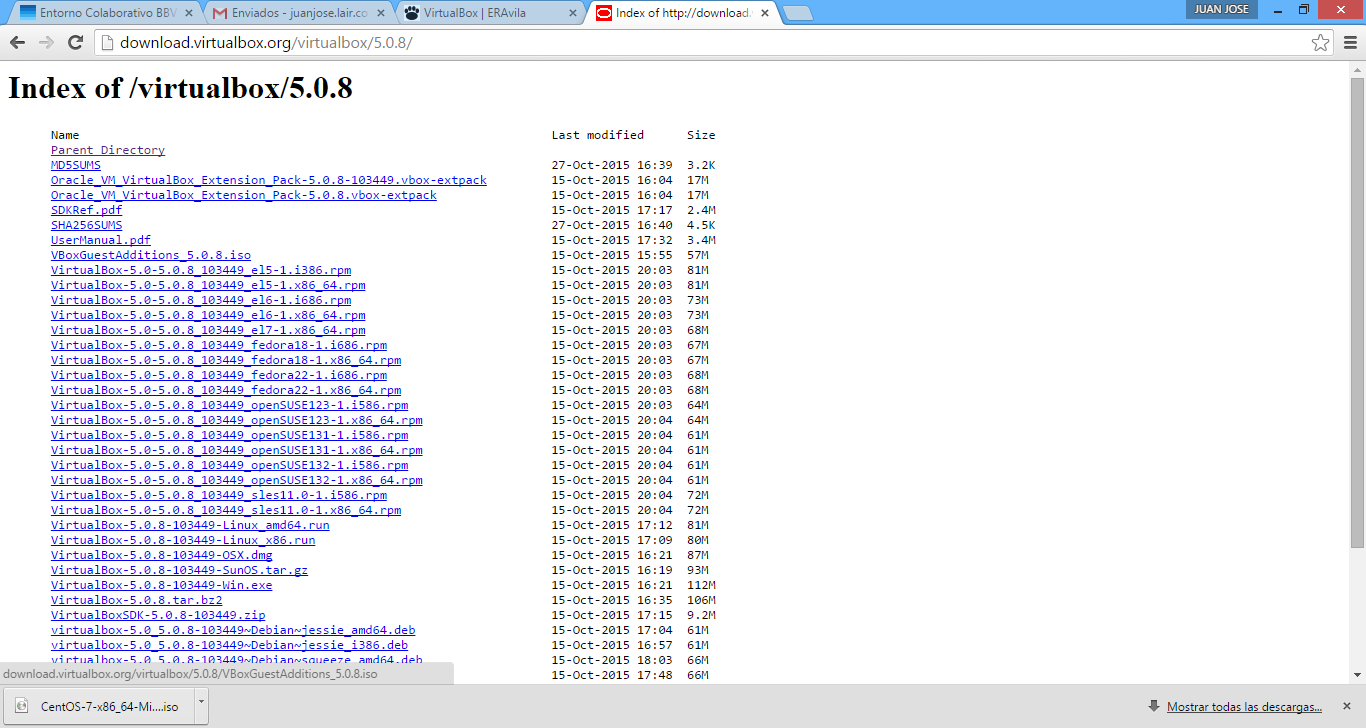
1. Y, a menos de que se busque compartir la unidad de disco duro que se cree, la opción por defecto es la usual. Además, un disco duro de tamaño fijo puede traer algunas ventajas de desempeño; pero, a menos de que se trate de un sistema para producción, la opción de crecimiento dinámico, trae mayores ventajas en el uso del espacio de disco y la administración de los archivos de la máquina virtual. Para nuestro caso, se eligió la opción de crear el disco de manera fija con un tamaño de 8GB.

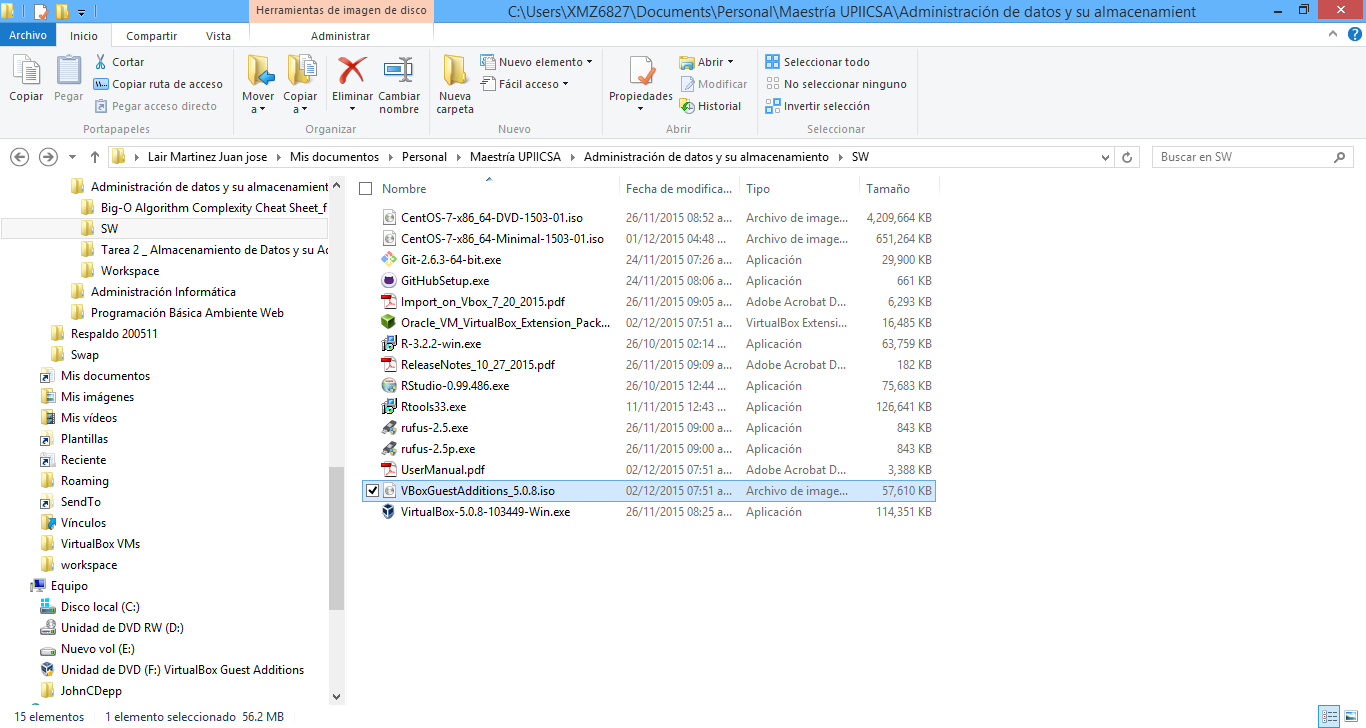
 

1. Dando la ubicación física del disco duro, por último se observa la nueva máquina virtual que deberá de aparecer en el administrador de **VIRTUALBOX** de forma “Apagada”.

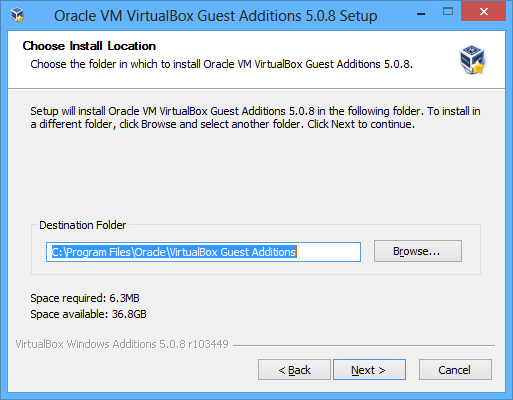
 

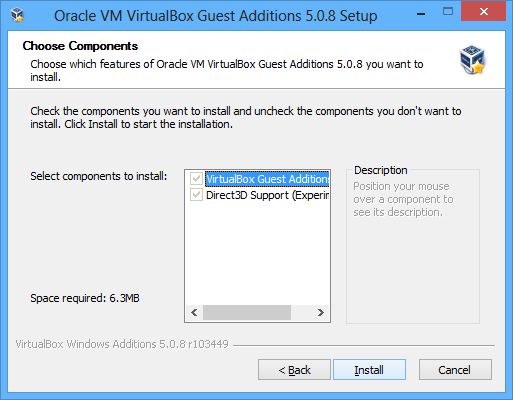
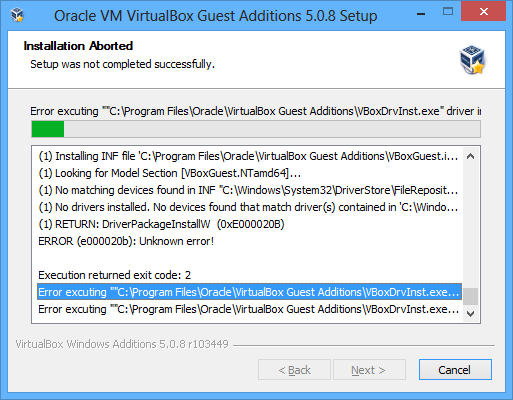
1. Creada nuestra máquina virtual, procedemos a instalar las adiciones del huésped (*VirtualBoxGuestAdditions*). Estas adiciones son difíciles de ubicar, aunque varias páginas de **Oracle** y **virtualbox.org** recomiendan instalarlas; pero se tiene problemas con la instalación como se puede ver en las siguientes imágenes.

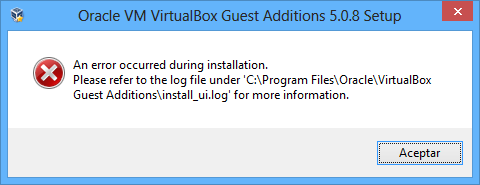




1. Como se puede observar, una vez descargado el programa para instalar y siguiendo con las recomendaciones que nos da el propio instalador, una vez pasado por la venta “*ChooseComponents*”, el instalador aborta al encontrar un error al hacer referencia a un log dentro de la ruta citada en el mesaje de la última ventana de esta página.



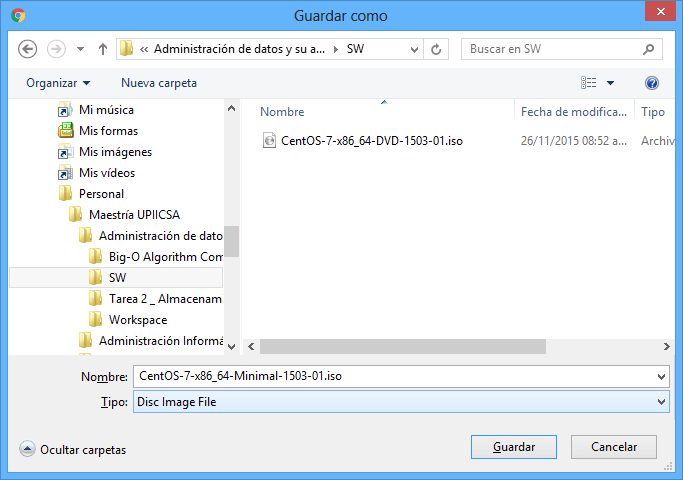
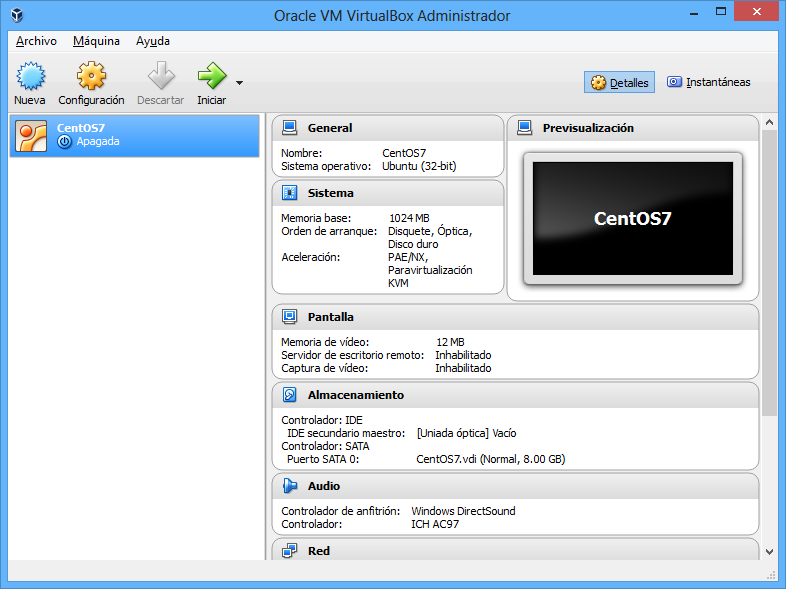
# INSTALACIÓN centos 7

1. Para la instalación de **CentOS**, se procede a su descarga desde la página [www.centos.org/download/](http://www.centos.org/download/)la versión mínima “*Minimal ISO*”. En nuestro caso, se descargó desde la dirección actual del país <http://centos.blazar.mx/7/isos/x86_64/CentOS-7-x86_64-Minimal-1511.iso>

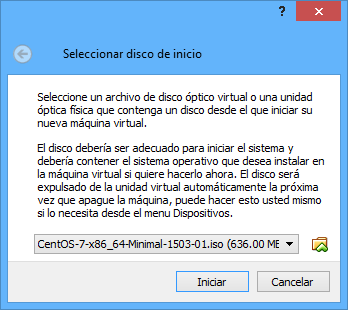
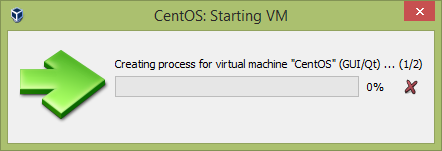


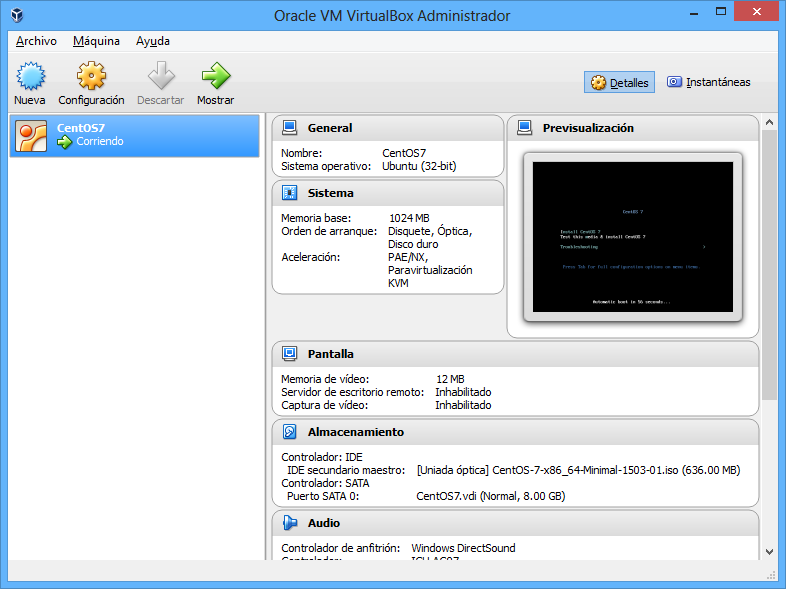


1. Se almacena en la ruta local designada y el archivo descargado y se procede a asignarlo a la máquina virtual generada anteriormente.

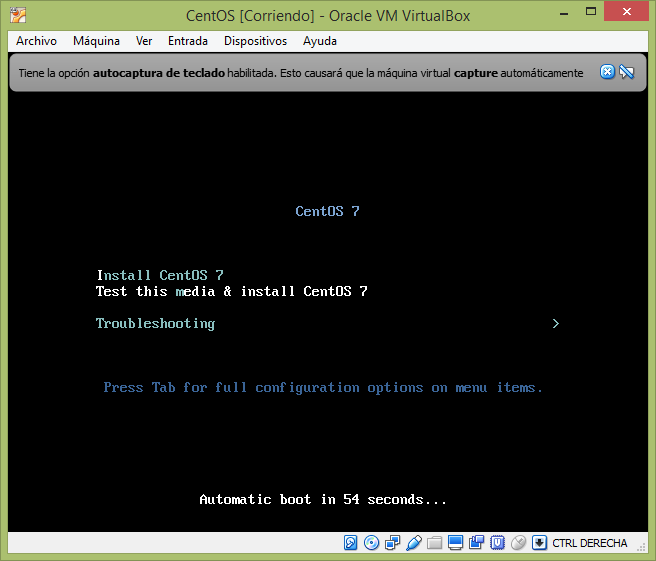
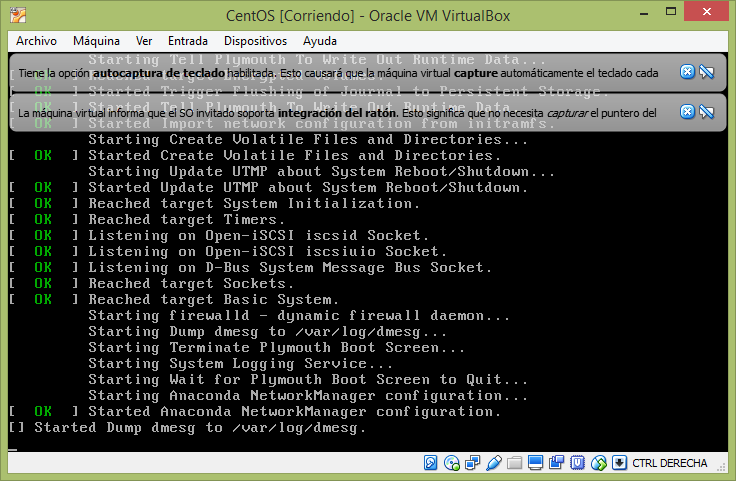
 

1. Se selecciona el archivo descargado para ser asignado a la máquina virtual y se procede a su creación. Terminado el proceso, se encuentra lista para su inicio.

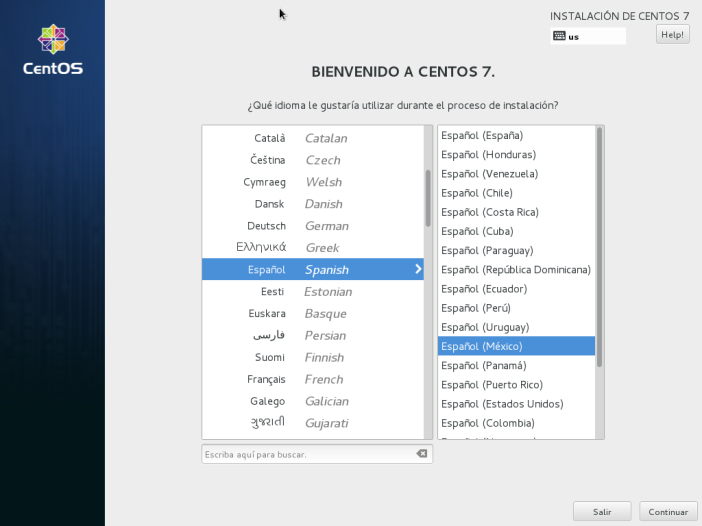
 



1. Seleccionando del menú de herramientas en la parte de INICIAR e instalar **CentOS 7** con la opción “*Install CentOS7*” y se le da *<enter>*.

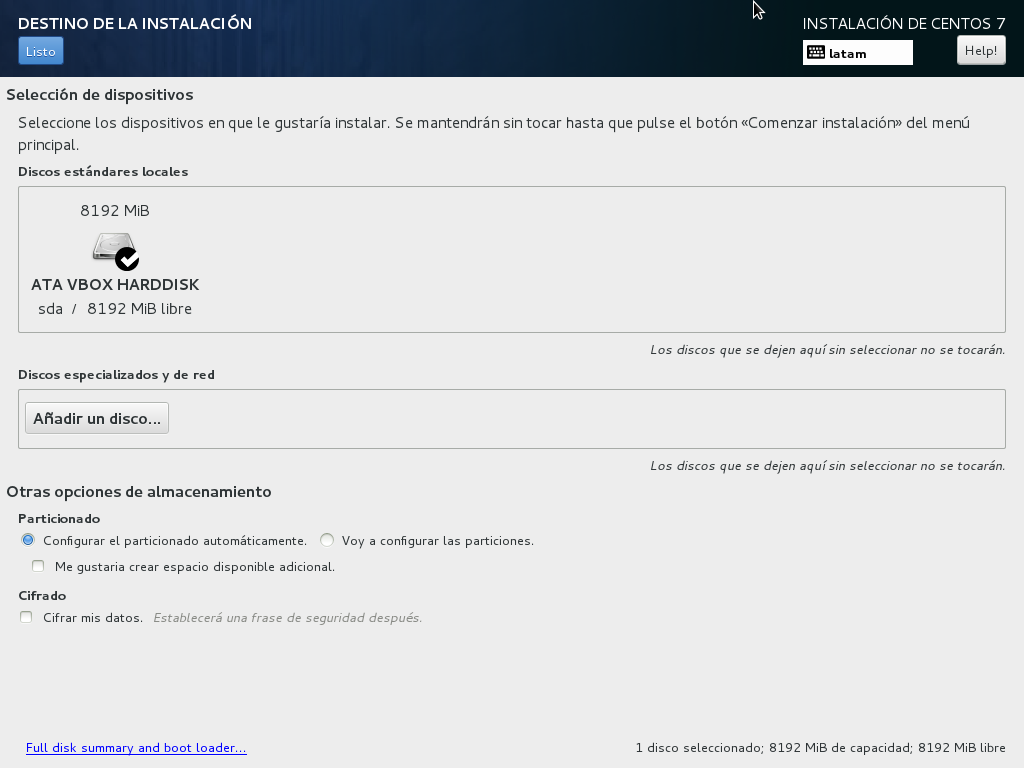
 

1. El siguiente paso es seleccionar el idioma, para nuestro caso se seleccionó “Español (México)” y se continúa con el proceso de instalación de la máquina virtual.

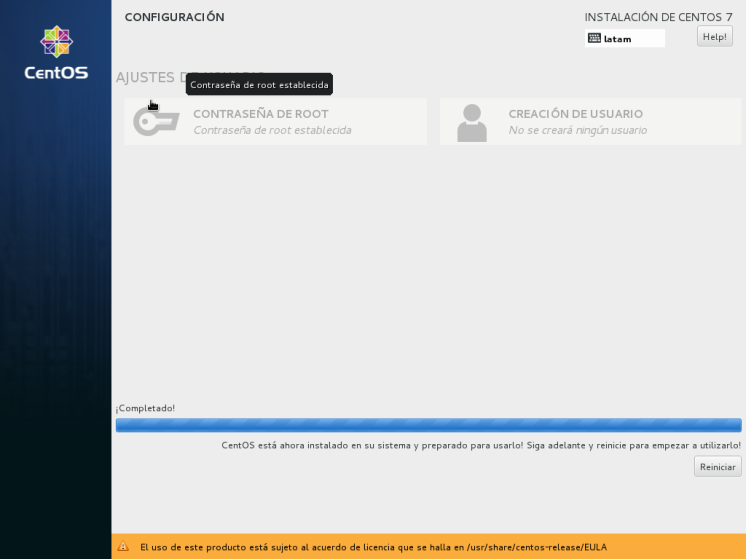




1. Se selecciona la partición automática, *Destino de la instalación*y a *Comenzar con la instalación* y la contraseña del usuario *root*.

1. Por último, se termina con la instalación y se *Reinicia* la máquina virtual y está lista para utilizarse.

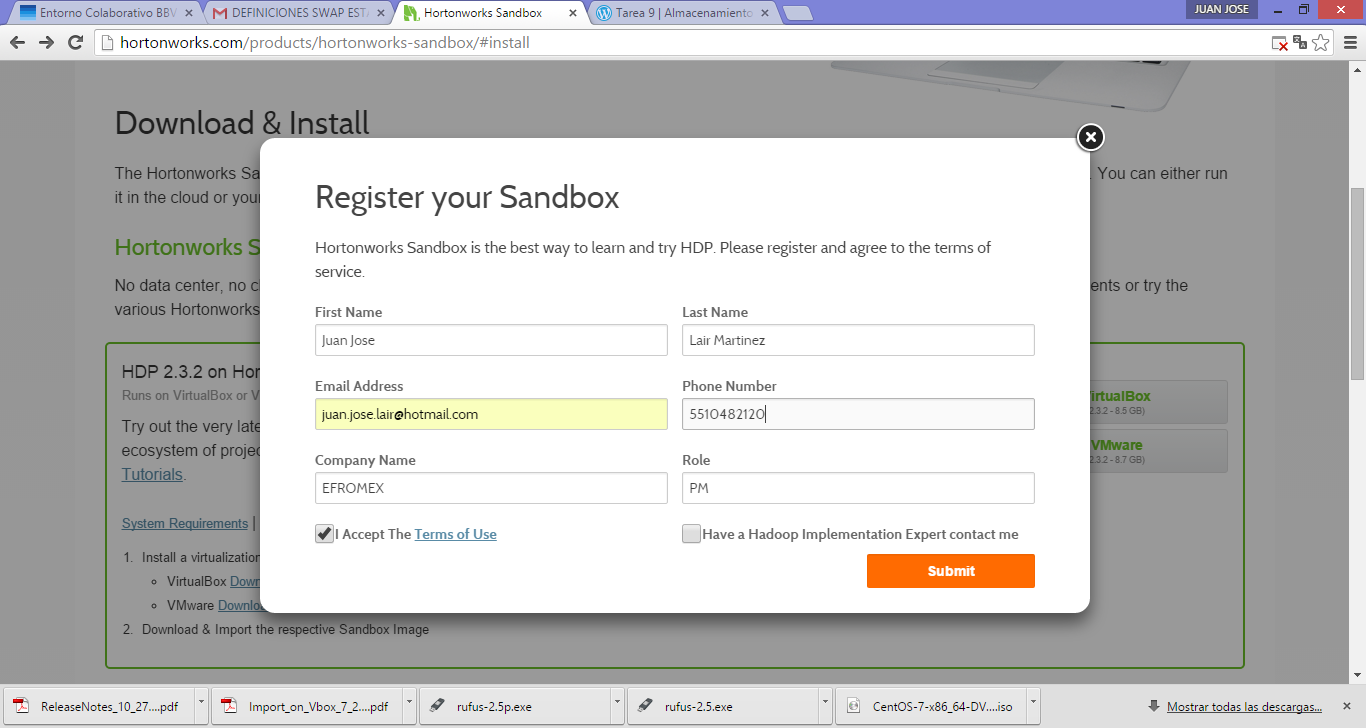




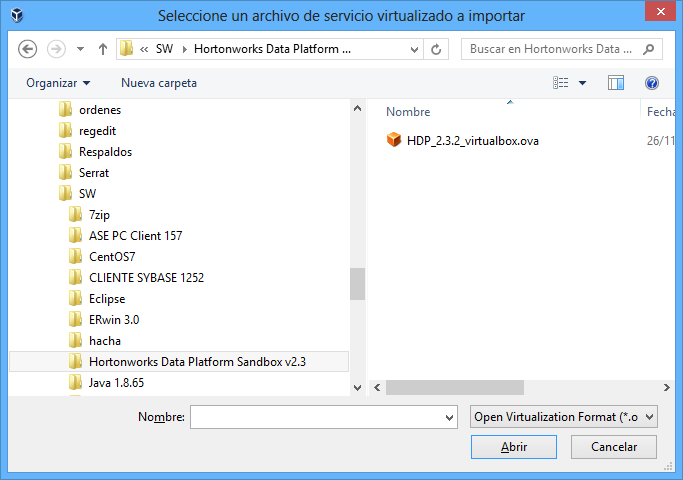
# INSTALACIÓN HORTONWORKS DATA PLATFORM SANDBOX V.2.3

1. Se procede a descargar de la página <http://hortonworks.com/products/hortonworks-sandbox/#install> el producto **HortonworksSandbox**. Primero, para ello, es necesario contar con una cuenta; por lo que se procede a la realización del registro.

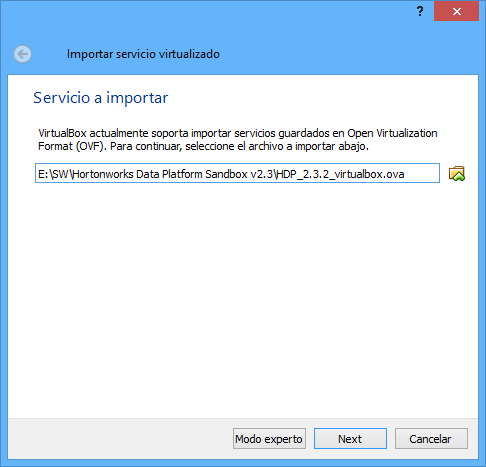
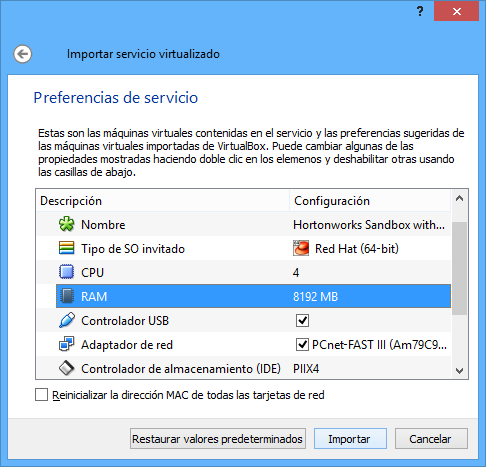


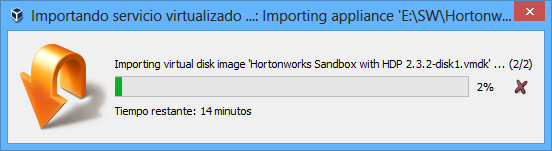


1. Almacenado el archivo en una carpeta local elegida (que por cierto, ocupa mucho espacio y se tarda mucho en descargar), se procede a *importar servicio virtualizado*; es decir, el archivo descargado.

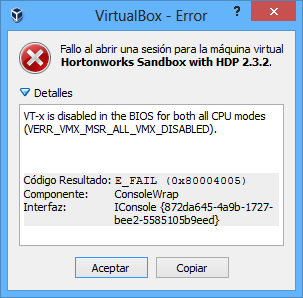
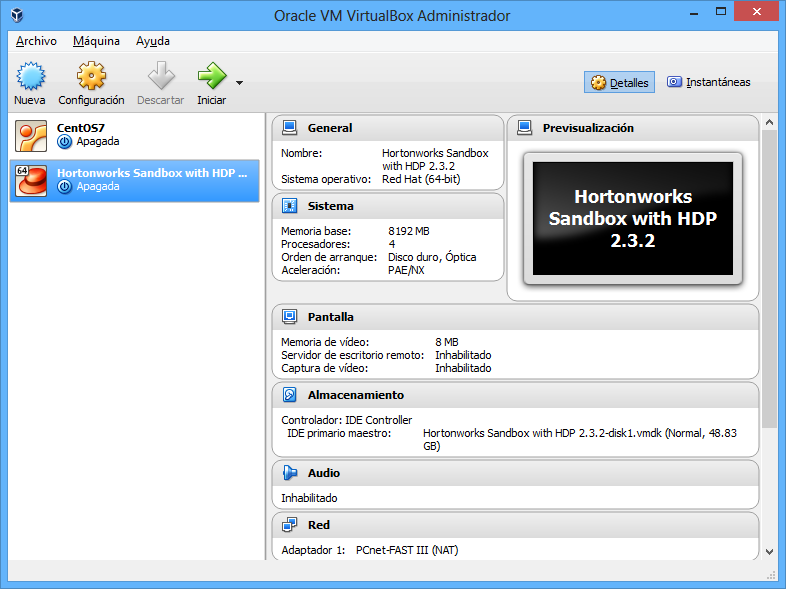
 

1. Elegido el archivo descargado en el *Servicio a importar*, se eligen las *Preferencias del servicio* y procede a importar el servicio.



1. Al momento de iniciar la máquina virtual marca el error anexo y no se pudo inicializar.



# ANEXO

**Lenguaje R.** Lenguaje y entorno de programación para análisis estadístico y gráfico. Se trata de un proyecto de *software* libre resultado de la implementación GNU (*"GNU's Not Unix"*) del premiado lenguaje S.

**RStudio.** Es un entorno de desarrollo integrado ([*IDE*](https://es.wikipedia.org/wiki/Ambiente_de_desarrollo_integrado)) para [R (lenguaje de programación)](https://es.wikipedia.org/wiki/R_(lenguaje_de_programaci%C3%B3n)) . Incluye una consola, editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo. **RStudio** está disponible para *Windows*, *Mac* y *Linux* o para navegadores conectados a **RStudio Server** **o RStudio Server Pro** (*Debian* / *Ubuntu*, *RedHat* / *CentOS*, y *SUSE Linux*).

**Git.** Es un software de [control de versiones](https://es.wikipedia.org/wiki/Control_de_versiones) diseñado por [*Linus Torvalds*](https://es.wikipedia.org/wiki/Linus_Torvalds), pensando en la eficiencia y la confiabilidad del mantenimiento de versiones de aplicaciones cuando éstas tienen un gran número de archivos de [código fuente](https://es.wikipedia.org/wiki/C%C3%B3digo_fuente).

**GitHub.** Es una plataforma de desarrollo colaborativo para alojar proyectos utilizando el sistema de [control de versiones](https://es.wikipedia.org/wiki/Control_de_versiones) [Git](https://es.wikipedia.org/wiki/Git). Utiliza el [*framework*](https://es.wikipedia.org/wiki/Framework) [**Ruby on Rails**](https://es.wikipedia.org/wiki/Ruby_on_Rails) por *GitHub, Inc.* (anteriormente conocida como *Logical Awesome*). Desde enero de [2010](https://es.wikipedia.org/wiki/2010), **GitHub** opera bajo el nombre de **GitHub, Inc**. El código se almacena de forma [pública](https://es.wikipedia.org/wiki/C%C3%B3digo_abierto), aunque también se puede hacer de forma privada, creando una cuenta de pago.

**VirtualBox.** Es un [software](https://es.wikipedia.org/wiki/Software) de [virtualización](https://es.wikipedia.org/wiki/Virtualizaci%C3%B3n) para arquitecturas [*x86*](https://es.wikipedia.org/wiki/X86)*/*[*amd64*](https://es.wikipedia.org/wiki/Amd64), creado originalmente por la empresa alemana *innotek GmbH*. Actualmente es desarrollado por [*Oracle Corporation*](https://es.wikipedia.org/wiki/Oracle_Corporation)como parte de su familia de productos de [virtualización](https://es.wikipedia.org/wiki/Virtualizaci%C3%B3n). Por medio de esta aplicación es posible instalar [sistemas operativos](https://es.wikipedia.org/wiki/Sistema_operativo) adicionales, conocidos como «sistemas invitados», dentro de otro sistema operativo «anfitrión», cada uno con su propio ambiente virtual.

**CentOS.** (*Community ENTerprise Operating System*) es una [bifurcación](https://es.wikipedia.org/wiki/Bifurcaci%C3%B3n_(desarrollo_de_software)) a nivel binario de la distribución *Linux*[*Red Hat Enterprise Linux*](https://es.wikipedia.org/wiki/Red_Hat_Enterprise_Linux)*RHEL*, compilado por voluntarios a partir del [código fuente](https://es.wikipedia.org/wiki/C%C3%B3digo_fuente) publicado por [*Red Hat*](https://es.wikipedia.org/wiki/Red_Hat). Es un sistema operativo de código abierto, basado en la distribución *Red Hat Enterprise Linux*, operándose de manera similar, y cuyo objetivo es ofrecer al usuario un software de "clase empresarial" gratuito. Se define como robusto, estable y fácil de instalar y utilizar.

**Horton Sandbox.** *Sandbox* es un entorno personal portátil *Hadoop* que viene con una docena de tutoriales interactivos. Incluye muchos de los más emocionantes acontecimientos de la última distribución *Hortonworks Data Platform*, envasados ​​en un entorno virtual que se puede poner en marcha en 15 minutos.