



UNIVERSIDAD TÉCNICA DE MACHALA

Maestría en Software

Asignatura:

Ciencia de Datos e Inteligencia de
Negocios

Tema:

Tarea Práctica 3. Instalación y Configuración del Clúster de
Big Data.

Docente: Ing. Bertha Mazón, Mg. Inf.

Estudiantes:

ESTEBAN FABRICIO GONZABAY JIMENEZ

JIMMY FERNANDO CASTILLO CRESPIÓN

CESAR DAVID SANTILLAN VILLOTA

JORGE LUIS MIRANDA GALLEGOS

2021-2022

Consultar plataformas y herramientas de Big data.

Nombre	Logo	Link	Características	Descripción
Apache Hadoop		https://azure.microsoft.com/es-mx/services/hdinsight/#security	<ul style="list-style-type: none">• Tiene que ver con Big Data.• Hace posible a las organizaciones obtener conocimiento rápidamente a partir de cantidades masivas de datos, estructurados y no estructurados.• Gestiona datos.• Da accesibilidad a datos.• Cumple con la gobernabilidad e integración de datos.	<p>Apache Hadoop es un framework de código abierto que permite el almacenamiento distribuido y el procesamiento de grandes conjuntos de datos en base a un hardware comercial.</p> <p>El propósito principal del framework es almacenar grandes cantidades de datos y permitir consultas sobre dichos datos, que se ofrecerán con un bajo tiempo de respuesta. Esto se consigue mediante la ejecución distribuida de código en múltiples nodos (máquinas), cada uno de los cuales se encarga de procesar una parte del trabajo a realizar.</p>

Elasticsearch	 elasticsearch	https://www.elastic.co/es/	<ul style="list-style-type: none"> • Es una base de datos NoSQL orientada a documentos JSON, al estilo de MongoDB. Por lo cual no necesita que se definan esquemas a la hora de insertar los datos. • Nos permite indexar volúmenes de datos, para poder consultarlos posteriormente. • Es una base de datos distribuida que escala de manera dinámica de forma horizontal, por lo que a mayor demanda podemos ir creciendo nodos. • Acceso por API, permite acceder a los datos en tiempo real. 	<p>Elasticsearch es un motor de analítica y análisis distribuido, gratuito y abierto para todos los tipos de datos, incluidos textuales, numéricos, geoespaciales, estructurados y no estructurados.</p> <p>Elasticsearch está desarrollado a partir de Apache Lucene y fue presentado por primera vez en 2010 por Elasticsearch N.V. (ahora conocido como Elastic). Conocido por sus API REST simples, naturaleza distribuida, velocidad y escalabilidad, Elasticsearch es el componente principal del Elastic Stack, un conjunto de herramientas gratuitas y abiertas para la ingesta, el enriquecimiento, el almacenamiento, el análisis y la visualización de datos. Comúnmente denominado ELK Stack (por Elasticsearch, Logstash y Kibana), el Elastic Stack</p>

				ahora incluye una gran colección de agentes ligeros conocidos como Beats para enviar los datos a Elasticsearch.
Apache Storm	 The logo for Apache Storm features a stylized blue and orange lightning bolt icon followed by the text "APACHE STORM" in a bold, sans-serif font.	https://storm.apache.org/	<ul style="list-style-type: none"> • Administración de servicio al cliente en tiempo real • Monetización de datos • Cuadros de mando operativos o análisis de seguridad cibernética • Detección de amenazas. 	Apache Storm es un sistema utilizado para procesar datos en tiempo real. Es de fuente abierta y gratuita. Este gran sistema facilita el procesamiento de flujos ilimitados de datos. Además, su uso resulta muy simple, y puede ser utilizado con cualquier lenguaje de programación.
MongoDB	 The MongoDB logo consists of a green leaf-like icon followed by the word "mongoDB" in a lowercase, sans-serif font.	https://www.mongodb.com/es	<ul style="list-style-type: none"> • Consultas ad hoc • indexación • replicación • balanceo de carga • almacenamiento de archivos • ejecución de JavaScript del lado del servidor 	MongoDB (del inglés humongous, "enorme") es un sistema de base de datos NoSQL orientado a documentos de código abierto y escrito en C++, que en lugar de guardar los datos en tablas lo hace en estructuras de datos BSON (similar a JSON) con un esquema dinámico. Al ser un proyecto de código abierto, sus binarios están disponibles para los

				sistemas operativos Windows, GNU/Linux, OS X y Solaris y es usado en múltiples proyectos o implementaciones en empresas como MTV Network, Craigslist, BCI o Foursquare.
Apache Spark	 The logo for Apache Spark features the word "APACHE" in a small, sans-serif font above a large, stylized orange star. Below the star, the word "Spark" is written in a bold, lowercase, sans-serif font. A small trademark symbol (TM) is located at the bottom right of the "k".	http://spark.apache.org	<p>Está integrado con Apache Hadoop.</p> <p>Permite el procesamiento en tiempo real.</p> <p>Trabaja en memoria y en Disco.</p> <p>Proporciona API para Java, Scala, Python y R.</p>	<p>Apache Spark es un framework de computación en clúster open-source.</p> <p>Spark proporciona una interfaz para la programación de clusters completos con Paralelismo de Datos implícito y tolerancia a fallos.</p> <p>Se pueden programar aplicaciones usando diferentes lenguajes como Java, Scala, Python o R. pudiendo ser, según el programa, hasta 100 veces más rápido en memoria o 10 veces más en disco que Hadoop MapReduce.</p>

Python	 PYTHON	www.python.org	Simplicidad Múltiple compatibilidad Aprendizaje rápido. Herramientas de soporte. Visualización sin límites	<p>Python es un lenguaje avanzado de programación con la ventaja de ser relativamente fácil de usar para usuarios que no estén familiarizados con la informática de manera profesional, pero que necesitan trabajar con análisis de datos (estadistas, biólogos, físicos, lingüistas...).</p> <p>Es una herramienta para Big Data muy eficiente, en parte debido a la gran comunidad existente, por lo que Python dispone de muchas librerías ya hechas por otros usuarios.</p> <p>Sin embargo, tiene en su contra que no es un lenguaje muy rápido en su ejecución, por lo que suele ser empleado para tareas de integración o tareas donde no haya cálculos pesados.</p>
---------------	--	--	--	--

Apache Cassandra		https://cassandra.apache.org/download/	<ol style="list-style-type: none"> 1. Arquitectura escalable. 2. Diseño activo de principio a fin. 3. Rendimiento a escala lineal. 4. Disponibilidad continua. 5. Detección de fallos y recuperación transparente. 6. Modelo de datos flexible y dinámico. 7. Protección de datos sólida. 8. Consistencia de los datos sintonizable. 9. Replicación de datos multi-centro. 10. Compresión de datos. 11. CQL (Lenguaje de Consulta Cassandra). 	<p>Apache Cassandra es una base de datos NoSQL (Not Only SQL), un tipo de bases de datos que, por la cantidad de datos o por la evolución de los datos que hemos tenido en los últimos años, ha sido necesario crear.</p> <p>Estas nuevas bases de datos son más flexibles a la hora de recoger datos desestructurados o semiestructurados que las bases de datos tradicionales o las bases de datos relacionales.</p>
------------------	---	---	--	--

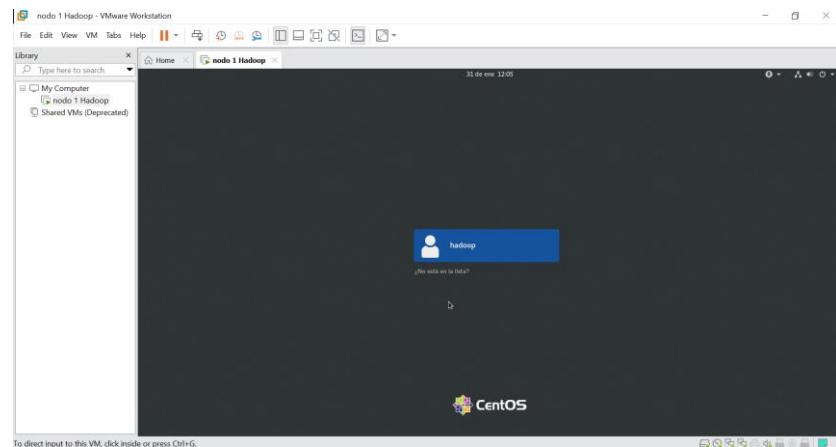
Lenguaje R		<p>https://cran.r-project.org/bin/windows/base/</p> <p>https://www.r-project.org/</p>	<p>Posibilidad de crear gráficos, basado en LaTeX.</p> <p>Gran cantidad de herramientas estadísticas.</p> <p>Posibilidad de crear tus propias funciones, además de objetos al ser su programación POO (orientada a objetos).</p> <p>Integración con distintas bases de datos.</p> <p>Puede tener un uso matemático, como sustitución a MATLAB.</p>	<p>R es un entorno de software libre (licencia GNU GLP) y lenguaje de programación interpretado, es decir, ejecuta las instrucciones directamente, sin una previa compilación del programa a instrucciones en lenguaje máquina. El término entorno, en R, se refiere a un sistema totalmente planificado y coherente, en lugar de una acumulación de herramientas específicas e inflexibles, como suele ser el caso en otros softwares de análisis de datos.</p>
Apache Drill		https://drill.apache.org/download/	<p>Alto rendimiento y baja latencia en ejecución de consultas.</p> <p>No necesita metadatos centralizados.</p> <p>Ejecución dinámica de consultas sin necesidad de</p>	<p>Apache Drill es un motor de consultas SQL para ser utilizado en entornos Big Data (Hadoop-HDFS, HBase, Hive, MongoDB...) con una baja latencia, sin necesidad de tener definido un esquema de datos previo y permitiendo combinar en una única consulta, datos</p>

		<p>que haya definido un esquema de datos.</p> <p>Lenguaje ANSI SQL (2003)</p> <p>Soporte para realizar consultas sobre datos anidados en varios niveles.</p> <p>Acceso simultaneo a múltiples fuentes de datos.</p> <p>Mediante conectores, Drill puede consumir datos de fuentes tales como sistema de ficheros local, sistema de ficheros distribuido.</p> <p>Integración con Apache Hive.</p> <p>Integración con herramientas BI</p> <p>Definición de funciones de usuario (UDFs)</p> <p>Escalabilidad</p>	de distintas fuentes y Bases de datos (relacionales o no).
--	--	---	--

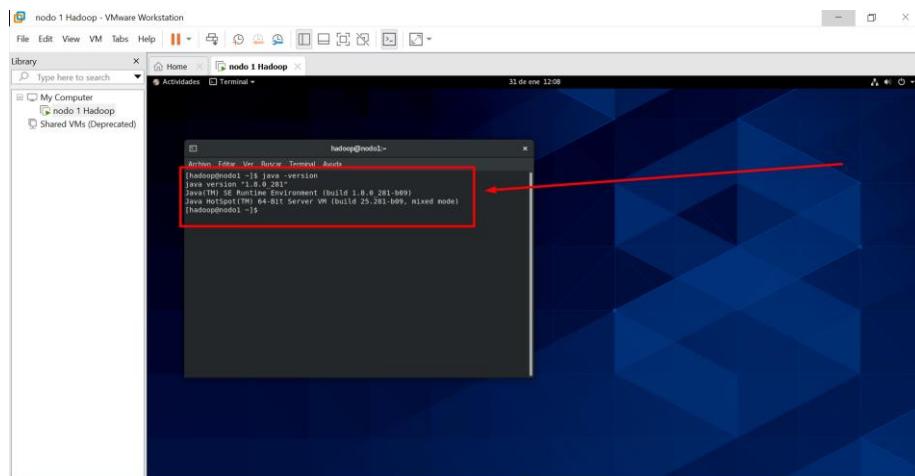
Apache Oozie	 The logo consists of the word "OOZIE" in a blue sans-serif font. Each letter is enclosed in a blue diamond shape, except for the "O" which is in a blue square. The letters are arranged in a staggered, overlapping manner.	https://oozie.apache.org/	<p>Oozie no soporta ciclos o bucles en la definición de los workflows, las definiciones workflow solo pueden ser de tipo DAG..</p> <p>La definición de una workflow para Oozie está basado en XML y ésta es llamado HPDL (Hadoop Process Definition Language).</p> <p>Los requerimiento mínimos para definir un Oozie workflow XML son nombre, punto de inicio y punto de finalización.</p> <p>Los nodos del flujo de control proporcionan un camino a la ruta de ejecución</p> <p>El nodo de inicio especifica un punto de inicio para Oozie Wrkflow</p> <p>El nodo final especifica un punto de finalización para Oozie Wrkflow</p>	<p>Apache Oozie es una herramienta para operaciones de Hadoop que permite a los administradores del clúster diseñar transformaciones de datos complejas a partir de múltiples tareas de componentes. De este modo se obtiene un mayor control sobre los trabajos y es más fácil repetirlos a intervalos predeterminados. En esencia, Oozie ayuda a los administradores a sacar más partido de Hadoop.</p>

		<p>Para un Oozie Workflow MapReduce, nosotros necesitamos definir una acción y dentro de la acción especificaciones con los parámetros del MapReduce.</p> <p>Dentro de la acciones, nosotros también podemos proporcionar los tags <ok> y <error> directamente para la siguiente acción encado de tener una ejecución satisfactoria o con error.</p>	
--	--	--	--

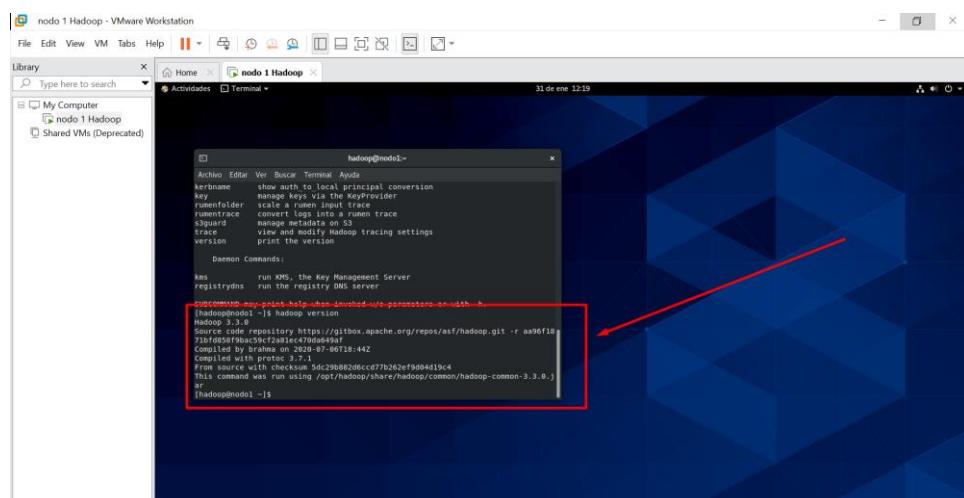
Instalación de Centos 8



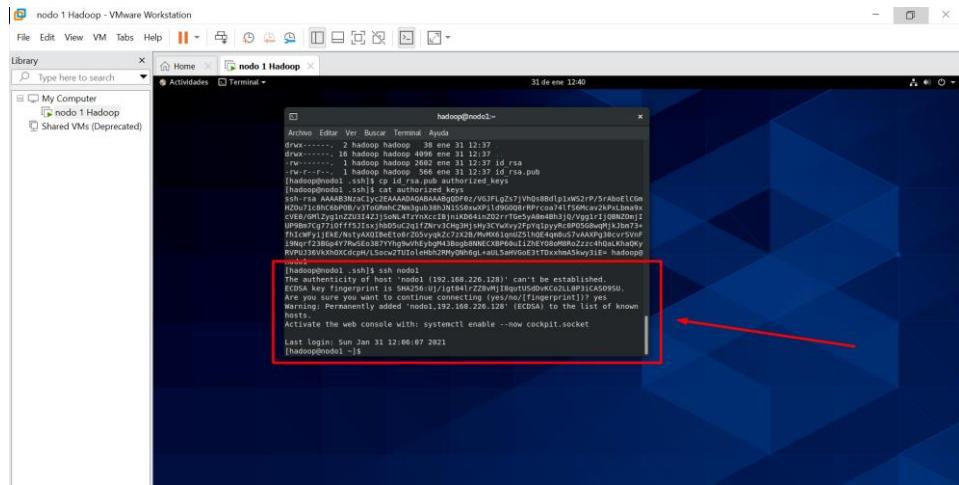
Instalación de JDK Java 8 en el nodo 1 (Master)



Variables de entorno de Hadoop en el nodo 1 (Master)



Servicio SSH en el nodo 1 (Master)



Montaje de un clúster Pseudo-distribuido de un solo nodo

Configuración del archivo core-site.xml en el nodo 1 (Master)

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://node1:9000</value>
  </property>
</configuration>
```

Configuración del archivo hdfs-site.xml en el nodo 1 (Master)

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/datos/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/datos/datanode</value>
  </property>
</configuration>
```

Inicio del servicio HDFS

```
[hadoop@nodo1 ~]$ start-dfs.sh
Starting namenodes on [nodo1]
Starting datanodes
localhost: datanode is running as process 3262. Stop it first.
Starting secondary namenodes [nodo1]
[hadoop@nodo1 ~]$ jps
4598 SecondaryNameNode
4281 NameNode
4794 Jps
3262 DataNode
```

Overview 'nodo1:9000' (active)

Started:	Sun Jan 31 14:02:35 -0500 2021
Version:	3.3.0, raa96f1071bfdb58fbac59cf2a81ec470da649af
Compiled:	Mon Jul 06 13:44:00 -0500 2020 by brahma from branch-3.3.0
Cluster ID:	CID-e5057007-01fa-4fd1-839d-4ee5128e2f60
Block Pool ID:	BP-444150833-192.168.226.128-1612119736470

Summary

Security is off.
Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).
Heap Memory used 55.7 MB of 237 MB Heap Memory. Max Heap Memory is 825 MB.
Non Heap Memory used 45.61 MB of 47.27 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	0 B
Configured Remote Capacity:	0 B
DFS Used:	0 B (100%)

Creación y carga de directorios y ficheros HDFS

```
[hadoop@nodo1 ~]$ start-dfs.sh
Starting namenodes on [nodo1]
Starting datanodes
Starting secondary namenodes [nodo1]
[hadoop@nodo1 ~]$ echo HolaMundo >> ~/Descargas/holaMundo.txt
[hadoop@nodo1 ~]$ cat ~/Descargas/holaMundo.txt
HolaMundo
[hadoop@nodo1 ~]$ hdfs dfs -mkdir /prueba
[hadoop@nodo1 ~]$ hdfs dfs -put ~/Descargas/holaMundo.txt /prueba
[hadoop@nodo1 ~]$ hdfs dfs -ls /prueba
Found 1 items
-rw-r--r-- 1 hadoop supergroup          10 2021-01-31 14:16 /prueba/holaMundo.txt
[hadoop@nodo1 ~]$ 
```

Browsing HDFS

Search: /prueba

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hadoop	supergroup	10 B	Jan 31 14:16	1	128 MB	holaMundo.txt

Showing 1 to 1 of 1 entries

Hadoop, 2020.

Servicio Yarn en el nodo 1 (Master)

The screenshot shows a Firefox browser window within a VMware Workstation interface. The URL is <http://nodo1:8042/node>. The page displays Hadoop NodeManager information for the master node. Key details include:

- Total Vmem allocated for Containers: 16.80 GB
- Total PMem allocated for Container: 8 GB
- Total VCores allocated: 8
- Resource types: memory-mb (unit=Mi), vcores
- NodeHealthStatus: true
- LastNodeHealthTime: Sun Jan 31 14:34:18 PET 2021
- NodeHealthReport: Sun Jan 31 14:34:15 PET 2021
- NodeManager started on: 3.3.0 from aa9d1871bf0858f9bac59cf2a81ec470da649af by brahma source checksum e0a276649f889c15d0e8f08ecc0c10 on 2020-07-08T11:58Z
- Hadoop Version: 3.3.0 from aa9d1871bf0858f9bac59cf2a81ec470da649af by brahma source checksum 5dc29b802d6cc77b262ef9d04d19c4 on 2020-07-08T18:44Z

Mapred-site.xml en el nodo 1 (Master)

```
[hadoop@nodo1 hadoop]$ hdfs dfs -cat /salida_texto2/part-r-00000 ~/Descargas/conteo_de_palabras.txt
HolaMundo      1
cluster        1
de             1
el             1
es             1
este           1
fichero        1
hola           1
mundo          1
para           1
texto          1
un             1
cat: `/home/hadoop/Desktop/conteo_de_palabras.txt': No such file or directory
[hadoop@nodo1 hadoop]$
```

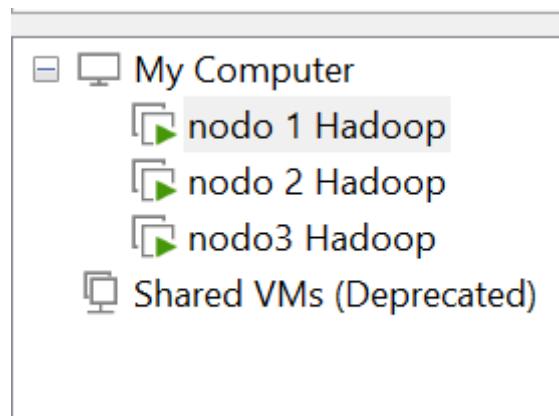
The screenshot shows a Firefox browser window within a VMware Workstation interface. The URL is http://nodo1:8088/cluster/app/application_1612121659417_0001. The page displays Application Overview for application_1612121659417_0001. Key details include:

- User: hadoop
- Name: word count
- Application Type: MAPREDUCE
- Application Tags:
- Application Priority: 0 (Higher Integer value indicates higher priority)
- YarnApplicationState: FINISHED
- Queue: default
- FinalStatus Reported by AM: SUCCEEDED
- Launched: dom ene 31 14:59:54 -0500 2021
- Finished: dom ene 31 14:59:55 -0500 2021
- Elapsed: 28sec
- Tracking URL: History
- Log Aggregation Status: DISABLED
- Application Timeout (Remaining Time): Unlimited
- Diagnostics:
- Unmanaged Application: false
- Application Node Label expression: <Not set>
- AM container Node Label expression: <DEFAULT_PARTITION>

Application Metrics

Total Resource Preempted:	<memory:0, vCores:0>
Total Number of Non-AM Containers Preempted:	0
Total Number of AM Containers Preempted:	0
Resource Preempted from Current Attempt:	<memory:0, vCores:0>
Number of Non-AM Containers Preempted from Current Attempt:	0
Aggregate Resource Allocation:	92767 MB-seconds, 54 vcore-seconds
Aggregate Preempted Resource Allocation:	0 MB-seconds, 0 vcore-seconds

Instalación de Nodos de Datos (DataNode) Virtualizado en VMware Workstation



Configuración del servicio SSH en Nodo1 (Master)

```
[hadoop@nodo1 .ssh]$ ssh nodo2
Activate the web console with: systemctl enable --now cockpit.socket

Last login: Sun Jan 31 18:12:04 2021 from 192.168.20.57
[hadoop@nodo2 ~]$ ssh nodo3
Activate the web console with: systemctl enable --now cockpit.socket

Last login: Sun Jan 31 18:14:52 2021 from 192.168.20.57
[hadoop@nodo3 ~]$
```

Información de los Datanodes

A screenshot of a Firefox browser window titled 'nodo 1 Hadoop - VMware Workstation'. The address bar shows 'nodo1:9870/dfshealth.html#tab-datanode'. The page displays 'Namenode Information' and 'Datanode information'. A histogram titled 'Datanode usage histogram' shows disk usage percentages for three nodes. Below it, a table titled 'In operation' lists three DataNodes with their details. The table includes columns for Node, Http Address, Last contact, Last Block Report, Used, Non DFS Used, Capacity, Blocks, Block pool used, and Version. Both nodes have an 'http://nodo1:9864' address and are listed as 'In operation'.

Node	Http Address	Last contact	Last Block Report	Used	Non DFS Used	Capacity	Blocks	Block pool used	Version
nodo2:9866 (192.168.20.62:9866)	http://nodo2:9864	2s	1m	4 KB	7.2 GB	63.84 GB	0	4 KB (0%)	3.3.0
nodo3:9866 (192.168.20.63:9866)	http://nodo3:9864	2s	1m	4 KB	7.2 GB	63.84 GB	0	4 KB (0%)	3.3.0

The screenshot shows the Hadoop Node Manager interface. At the top, there's a navigation bar with tabs for 'Home', 'nodo 1 Hadoop', 'nodo 2 Hadoop', and 'nodo3 Hadoop'. Below the navigation is a search bar and a sidebar with sections for 'Cluster Metrics', 'Cluster Nodes Metrics', 'Scheduler Metrics', and 'Tools'. The main content area is titled 'Nodes of the cluster' and contains several tables and charts. One table shows 'Active Nodes' with 2 entries, both labeled '/default-rack' and 'RUNNING'. Another table shows 'Decommissioning Nodes' with 0 entries.

MapReduce - Crear directorios y subir ficheros

```
[hadoop@nodo1 namenode]$ mapred --daemon start historyserver
[hadoop@nodo1 namenode]$ cd
[hadoop@nodo1 ~]$ hdfs dfs -mkdir /prueba
[hadoop@nodo1 ~]$ echo Prueba de subida al cluster >> ~/Documentos/texto_prueba.txt
[hadoop@nodo1 ~]$ cat ~/Documentos/texto_prueba.txt
Prueba de subida al cluster
[hadoop@nodo1 ~]$ hdfs dfs -put ~/Documentos/texto_prueba.txt /prueba
[hadoop@nodo1 ~]$ hdfs dfs -ls /prueba
Found 1 items
-rw-r--r-- 2 hadoop supergroup 28 2021-01-31 20:12 /prueba/texto_prueba.txt
[hadoop@nodo1 ~]$
```

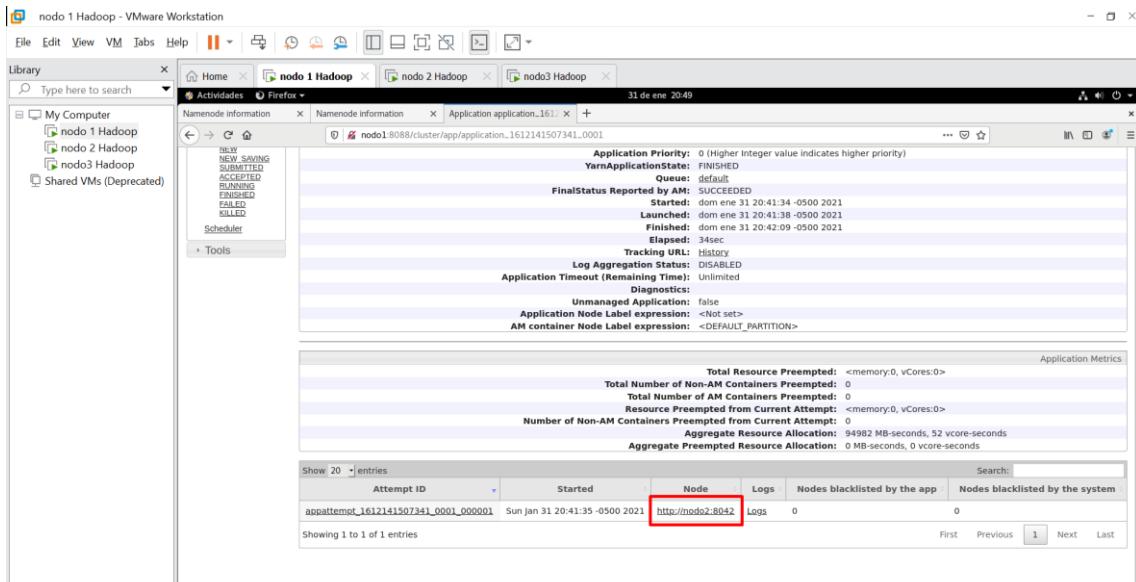
The screenshot shows the HDFS Browser interface. At the top, there's a navigation bar with tabs for 'Home', 'nodo 1 Hadoop', 'nodo 2 Hadoop', and 'nodo3 Hadoop'. Below the navigation is a search bar and a sidebar with sections for 'Hadoop', 'Overview', 'Datanodes', 'Datanode Volume Failures', 'Snapshot', 'Startup Progress', and 'Utilities'. The main content area is titled 'Browse Directory' and shows a list of files in the '/prueba' directory. The list includes 'Prueba2' (a folder), 'prueba' (a file), 'texto_prueba.txt' (a file), and 'tmp' (a folder). The table has columns for Permission, Owner, Group, Size, Last Modified, Replication, Block Size, and Name.

Lanzar un programa MapReduce en el clúster

```
[hadoop@nodo1 namenode]$ hadoop jar /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.0.jar wordcount /prueba /salida_texto1
```

```
[hadoop@nodo1 namenode]$ hdfs dfs -ls /salida_texto1
Found 2 items
-rw-r--r--  2 hadoop supergroup      0 2021-01-31 20:42 /salida_texto1/_SUCCESS
-rw-r--r--  2 hadoop supergroup    38 2021-01-31 20:42 /salida_texto1/part-r-00000
```

```
[hadoop@nodo1 namenode]$ cat ~/Documentos/conteo_de_palabras.txt  
Prueba 1  
al 1  
cluster 1  
de 1  
subida 1  
[hadoop@nodo1 namenode]$ █
```



Instalación de Apache Spark

```
[hadoop@nodo1 ~]$ spark-shell
2021-01-31 21:25:02,282 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://nodo1:4040
Spark context available as 'sc' (master = local[*], app id = local-1612146312708).
Spark session available as 'spark'.
Welcome to
```

version 3.0.1

```
Using Scala version 2.12.10 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_281)
Type in expressions to have them evaluated.
Type :help for more information.
```

scala>

```

scala> val fichero=sc.textFile("file:///opt/hadoop/spark/README.md")
fichero: org.apache.spark.rdd.RDD[String] = file:///opt/hadoop/spark/README.md MapPartitionsRDD[1] at textFile at <console>:24

scala> fichero.count()
res0: Long = 108

scala>

```

Probar Spark en el clúster

```
[hadoop@nodo1 ~]$ spark-submit --class org.apache.spark.examples.SparkPi --master yarn --deploy-mode cluster --name "apli1" /opt/hadoop/spark/examples/jars /spark-examples_2.12-3.0.1.jar 5
```

The terminal window shows the command being run:

```

[hadoop@nodo1 ~]$ spark-submit --class org.apache.spark.examples.SparkPi --master yarn --deploy-mode cluster --name "apli1" /opt/hadoop/spark/examples/jars /spark-examples_2.12-3.0.1.jar 5

```

Output from the application:

```

hadoop@nodo1:~/.ssh
hadoop@nodo1:~/.ssh
hadoop@nodo1:~/.ssh

2021-01-31 21:37:34,270 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:35,274 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:36,277 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:37,281 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:38,285 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:39,289 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:40,306 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:41,314 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:42,321 INFO yarn.Client: Application report for application_1612146704901_0001 (state: RUNNING)
2021-01-31 21:37:43,324 INFO yarn.Client: Application report for application_1612146704901_0001 (state: FINISHED)
2021-01-31 21:37:43,325 INFO yarn.Client:
    client token: N/A
    diagnostics: N/A
    ApplicationMaster host: nodo3
    ApplicationMaster RPC port: 37263
    queue: default
    start time: 1612147029864
    final status: SUCCEEDED
    tracking URL: http://nodo1:8088/proxy/application_1612146704901_0001/
    user: hadoop
2021-01-31 21:37:43,368 INFO util.ShutdownHookManager: Shutdown hook called
2021-01-31 21:37:43,371 INFO util.ShutdownHookManager: Deleting directory /tmp/spark-807ae084-973e-4ca3-8182-7d7fb9344832
2021-01-31 21:37:43,402 INFO util.ShutdownHookManager: Deleting directory /tmp/spark-e663fc75-07c3-4c5b-af97-24090f5f4b20
hadoop@nodo1 ~]$ 

```

The VMware Workstation interface shows three Hadoop nodes:

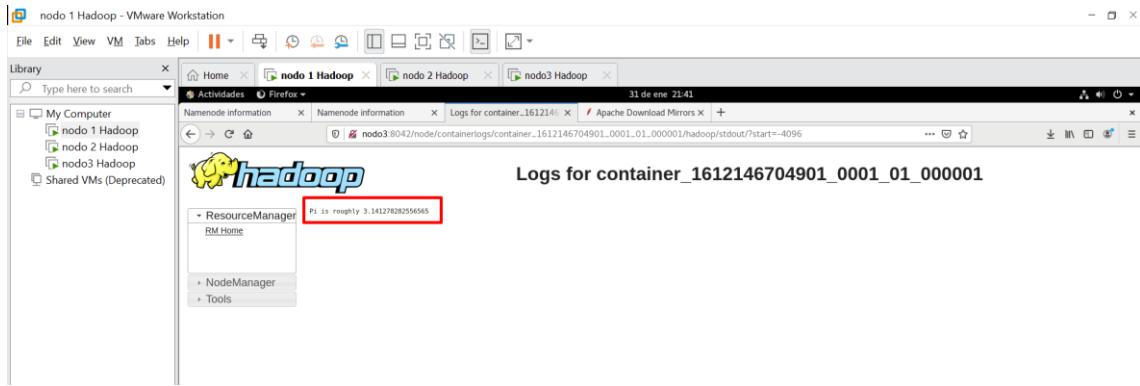
- nodo 1 Hadoop
- nodo 2 Hadoop
- nodo3 Hadoop

The Firefox browser displays the Hadoop cluster metrics page:

All Applications

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	Final
application_1612146704901_0001	hadoop	apli1	SPARK		default	0	Sun Jan 31 21:37:09 -0500 2021	Sun Jan 31 21:37:11 -0500 2021	Sun Jan 31 21:37:42 -0500 2021	FINISHED	SUCC

A red arrow points from the highlighted row in the table to the corresponding entry in the log output above.



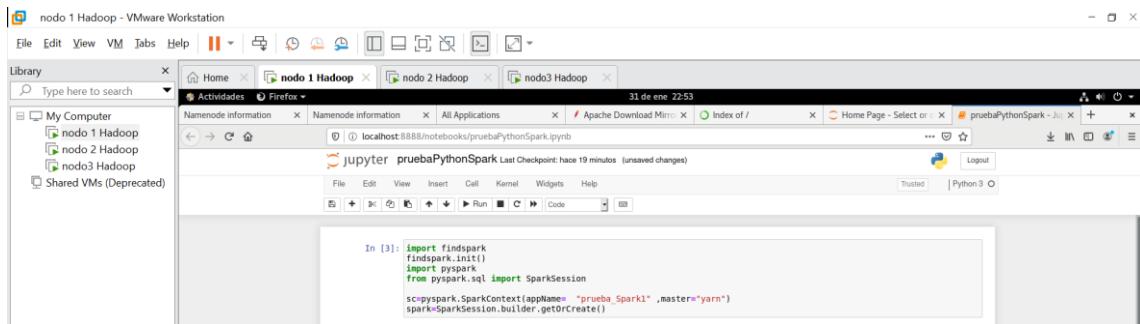
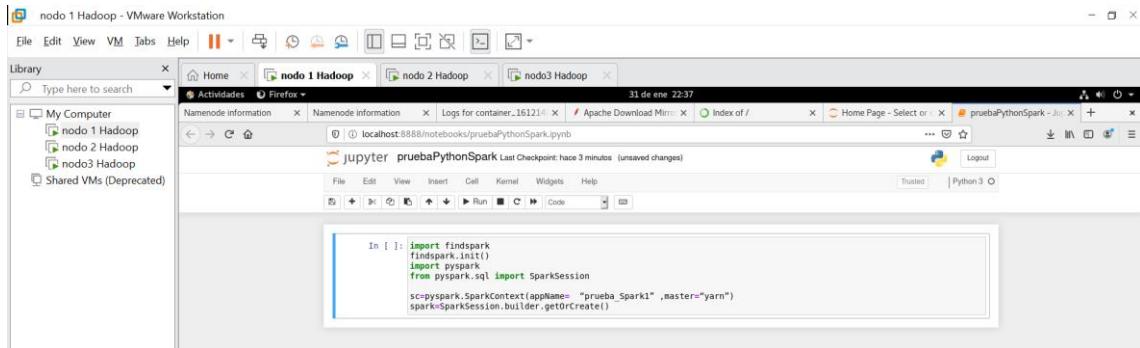
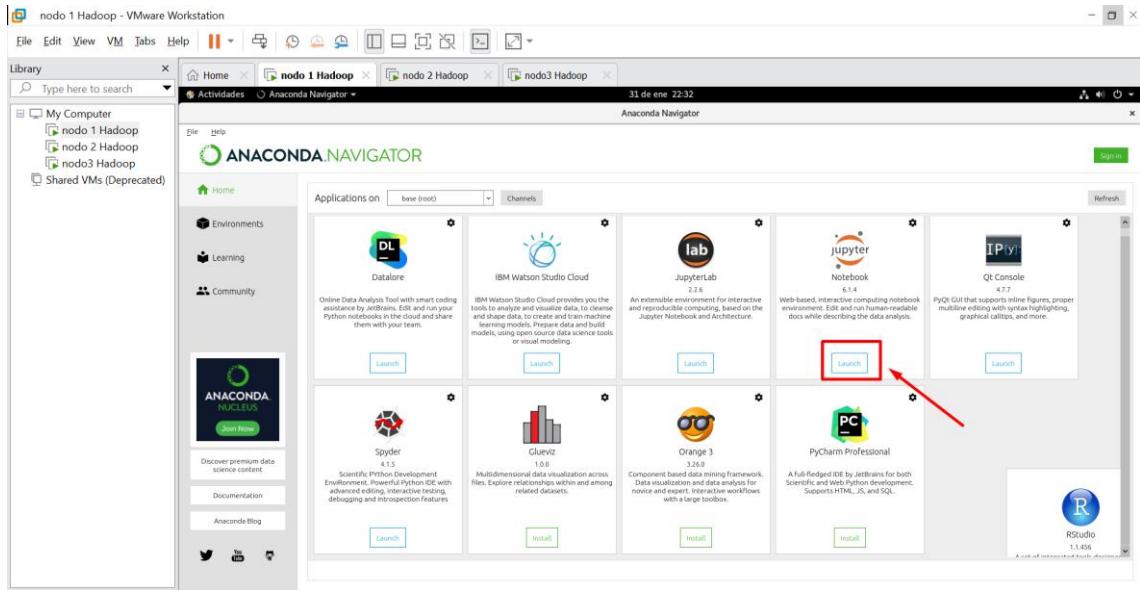
Instalación de Anaconda en el nodo Master

```
If you'd prefer that conda's base environment not be activated on startup,  
set the auto_activate_base parameter to false:  
  
conda config --set auto_activate_base false  
  
Thank you for installing Anaconda3!  
=====  
Working with Python and Jupyter notebooks is a breeze with PyCharm Pro,  
designed to be used with Anaconda. Download now and have the best data  
tools at your fingertips.  
  
PyCharm Pro for Anaconda is available at: https://www.anaconda.com/pycharm  
[hadoop@nodo1 Descargas]$
```

Conexión Jupyter Notebook con PySpark

```
[hadoop@nodo1 Descargas]$ source ~/.bashrc  
(base) [hadoop@nodo1 Descargas]$  
(base) [hadoop@nodo1 Descargas]$ cd  
(base) [hadoop@nodo1 ~]$ pip install findspark  
Collecting findspark  
  Downloading findspark-1.4.2-py2.py3-none-any.whl (4.2 kB)  
Installing collected packages: findspark  
Successfully installed findspark-1.4.2  
(base) [hadoop@nodo1 ~]$ █
```

Iniciar jupyter



ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Ru Co
application_1612146704901_0002	hadoop	prueba_Spark1	SPARK		default	0	Sun Jan 31 22:52:29 -0500 2021	Sun Jan 31 22:52:30 -0500 2021	N/A	RUNNING	UNDEFINED	2
application_1612146704901_0002	hadoop	aprili	SPARK		default	0	Sun Jan 31 22:47:45 -0500 2021	Sun Jan 31 22:47:46 -0500 2021	Sun Jan 31 22:48:19 -0500 2021	FINISHED	SUCCEEDED	N/A
application_1612146704901_0001	hadoop	aprili	SPARK		default	0	Sun Jan 31 21:37:09 -0500 2021	Sun Jan 31 21:37:11 -0500 2021	Sun Jan 31 21:37:42 -0500 2021	FINISHED	SUCCEEDED	N/A
application_1612146704901_0001	hadoop	WordCountPY	SPARK		default	0	Sun Jan 31 23:00:43 -0500 2021	Sun Jan 31 23:00:43 -0500 2021	Sun Jan 31 23:01:21 -0500 2021	FINISHED	SUCCEEDED	N/A

Conteo de palabras con Python

```
(base) [hadoop@nodo1 Documentos]$ spark-submit --master yarn --deploy-mode cluster --name "WordCountPY" WordCountPython.py /prueba/texto_prueba.txt /salida_spark_conteo_de_palabras1
```

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Ru Co
application_1612146704901_0004	hadoop	WordCountPY	SPARK		default	0	Sun Jan 31 23:00:43 -0500 2021	Sun Jan 31 23:00:43 -0500 2021	Sun Jan 31 23:01:21 -0500 2021	FINISHED	SUCCEEDED	N/A
application_1612146704901_0003	hadoop	prueba_Spark1	SPARK		default	0	Sun Jan 31 22:47:45 -0500 2021	Sun Jan 31 22:47:46 -0500 2021	Sun Jan 31 22:48:19 -0500 2021	FINISHED	SUCCEEDED	N/A
application_1612146704901_0002	hadoop	aprili	SPARK		default	0	Sun Jan 31 22:47:45 -0500 2021	Sun Jan 31 22:47:46 -0500 2021	Sun Jan 31 22:48:19 -0500 2021	FINISHED	SUCCEEDED	N/A
application_1612146704901_0001	hadoop	aprili	SPARK		default	0	Sun Jan 31 21:37:09 -0500 2021	Sun Jan 31 21:37:11 -0500 2021	Sun Jan 31 21:37:42 -0500 2021	FINISHED	SUCCEEDED	N/A

Browse Directory

/salida_spark_conteo_de_palabras1										<input type="button" value="Go!"/>	<input type="button" value=""/>	<input type="button" value=""/>	<input type="button" value=""/>
Show 25 entries										Search:			
	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name					
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	0 B	Jan 31 23:01	2	128 MB	_SUCCESS	<input type="button" value=""/>				
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	24 B	Jan 31 23:01	2	128 MB	part-00000	<input type="button" value=""/>				
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	39 B	Jan 31 23:01	2	128 MB	part-00001	<input type="button" value=""/>				

Showing 1 to 3 of 3 entries

Previous 1 Next

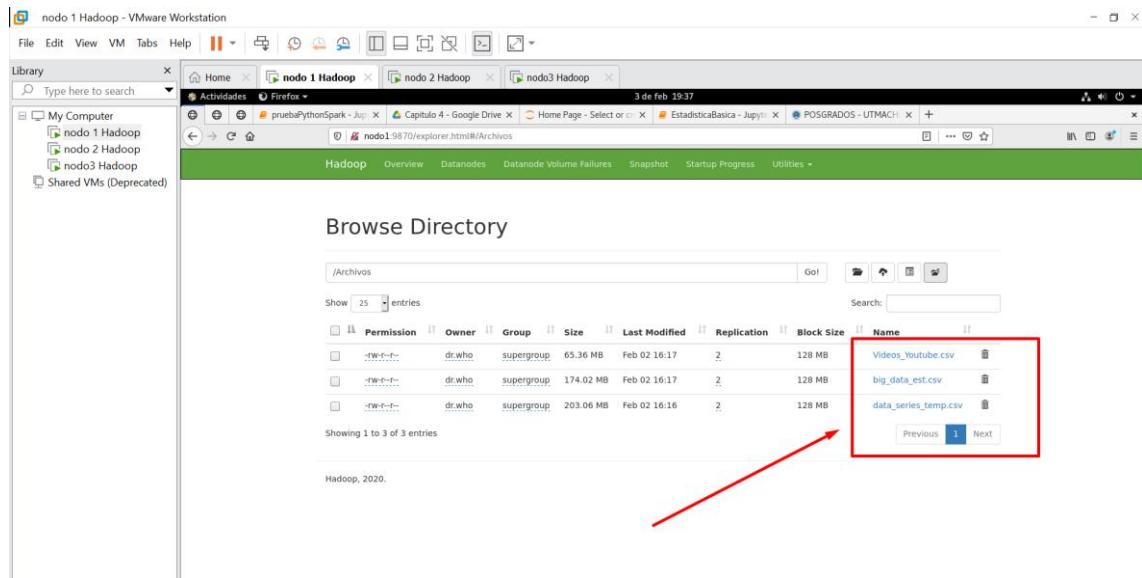
Hadoop, 2020.

Abrir una sesión pyspark con el clúster

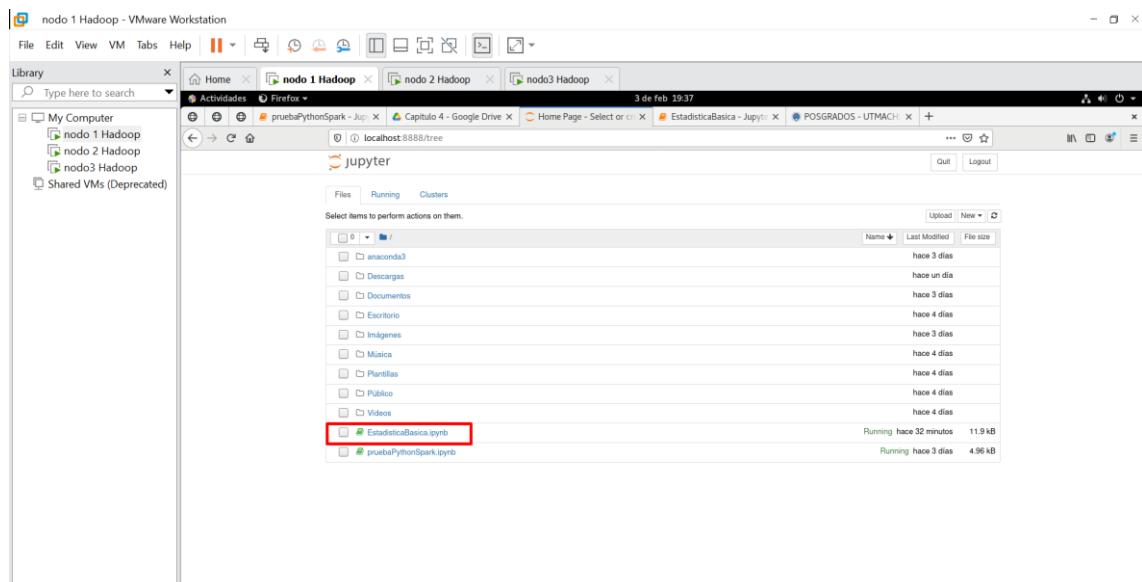
Practica 03-02-2021

Estadísticas Básicas

Carga de la data .csv



Carga del archivo jupyter



Ejecutamos el archivo EstadisticaBasica

nodo 1 Hadoop - VMware Workstation

File Edit View VM Tabs Help

Library Type here to search

- My Computer
 - nodo 1 Hadoop
 - nodo 2 Hadoop
 - nodo3 Hadoop
- Shared VMS (Deprecated)

Actividades Firefox localhost:8888/notebooks/EstadisticaBasica.ipynb

jupyter EstadisticaBasica Last Checkpoint: ayer a las 16:01 (unresolved changes)

In [2]:

```
from pyspark.sql import SparkSession
#dar un nombre al proceso y usar yarn como master
sc=pyspark.SparkContext(appName="Estadistica_Basica",master='yarn')
spark=SparkSession.builder.getOrCreate()
```

In [3]:

```
#Importar datos desde hdfs
ruta = "Archivos/data_series_temp.csv"
file_type = "csv"
infer_schema = "true"
first_row_is_header = "true"
delimiter = ","
Data = spark.read.format(file_type) \
    .option("inferSchema", infer_schema) \
    .option("header", first_row_is_header) \
    .option("sep", delimiter) \
    .load(ruta)
```

In [3]: #Visualizar los datos

```
Datos.show()
```

18]			53 9948	18 01/01/2018 0:05:48	124 01/01/20	
18]		0:19:04	48 3759	16 01/01/2018 0:09:17	60 01/01/20	
18]		0:32:31	63 10783	16 01/01/2018 0:09:39	60 01/01/20	
18]		0:32:41	27 10826	16 01/01/2018 0:10:01	60 01/01/20	
18]		0:32:39	28 10643	18 01/01/2018 0:10:02	450 01/01/20	
18]		0:20:07				

nodo 1 Hadoop - VMware Workstation

File Edit View VM Tabs Help

Library Type here to search

- My Computer
 - nodo 1 Hadoop
 - nodo 2 Hadoop
 - nodo3 Hadoop
- Shared VMS (Deprecated)

Actividades Firefox localhost:8888/notebooks/EstadisticaBasica.ipynb

jupyter EstadisticaBasica Last Checkpoint: ayer a las 16:01 (autosaved)

In [4]:

```
#Crear una lista de las columnas de tipo numérico
COLUMNAS_NUMERICAS=['Edad_Usuario','Bici','Ciclo_Estacion_Retiro','Ciclo_Estacion_Arrivo']
from pyspark.sql import functions as F
#Convertir a decimal las columnas seleccionadas
Datos = Datos.select(*[F.col(c).cast('Double').alias(c) for c in COLUMNAS_NUMERICAS])
#Ver el tipo de Datos
Datos.dtypes
```

Out[4]:

```
[('Edad_Usuario', 'double'),
 ('Bici', 'double'),
 ('Ciclo_Estacion_Retiro', 'double'),
 ('Ciclo_Estacion_Arrivo', 'double')]
```

In [5]:

```
#Descripción de una columna
Datos.select('EDAD_USUARIO').describe().show()
```

	summary	EDAD_USUARIO			
	count	3672218			
	mean	34.73662320701094			
	stddev	10.2739152954597441			

The screenshot shows a VMware Workstation window with a Firefox browser running a Jupyter Notebook. The browser has several tabs open, including 'nodo 1 Hadoop', 'nodo 2 Hadoop', 'nodo3 Hadoop', and the Jupyter interface. The Jupyter interface shows the following code and output:

```
In [5]: #Libreria para realizar funciones en Pyspark  
import pyspark.sql.functions as f  
  
In [6]: #Media de una columna  
Datos.select(f.mean('EDAD_USUARIO').alias('EDAD_PROMEDIO')).show()  
  
In [8]: Datos.select(f.avg('EDAD_USUARIO').alias('EDAD_PROMEDIO')).show()  
  
In [9]: #Valor minimo
```

To direct input to this VM, move the mouse pointer inside or press Ctrl+G.

To direct input to this VM, move the mouse pointer inside or press **Ctrl+G**.

```
In [7]: Datos.select(f.avg('EDAD_USUARIO').alias('EDAD_PROMEDIO')).show()
+-----+
| EDAD PROMEDIO |
+-----+
| 34.73662320701004 |
+-----+


In [9]: #Valor minimo
Datos.select(f.min('EDAD_USUARIO').alias('EDAD_MINIMO')).show()
+-----+
| EDAD MINIMO |
+-----+
| 16.0 |
+-----+


In [10]: #Valor maximo
Datos.select(f.max('EDAD_USUARIO').alias('EDAD MAXIMO')).show()
+-----+
| EDAD MAXIMO |
+-----+
| 118.0 |
+-----+


In [11]: Datos.select(f.round(f.mean(Datos['EDAD_USUARIO']),0).alias('EDAD_PROMEDIO_REDONDEADO')).show()
+-----+
| EDAD PROMEDIO REDONDEADO |
+-----+
| 35.0 |
+-----+
```

Práctica con la data Global SuperStore

The screenshot shows a VMware Workstation window titled "nodo 1 Hadoop - VMware Workstation". Inside, a Firefox browser is open to a "Browse Directory" page for "/Archivos". The table lists four files:

Name	Size	Last Modified	Replication	Block Size
Global Superstore1.csv	11.46 MB	Feb 03 20:02	2	128 MB
Videos_YouTube.csv	65.36 MB	Feb 02 16:17	2	128 MB
big_data_est.csv	174.02 MB	Feb 02 16:17	2	128 MB
data_series_temp.csv	203.06 MB	Feb 02 16:16	2	128 MB

A red arrow points from the "Global Superstore1.csv" file entry.

The screenshot shows a VMware Workstation window titled "nodo 1 Hadoop - VMware Workstation". Inside, a Firefox browser is open to a "jupyter" interface at "localhost:8888/tree". The sidebar shows a tree structure with several folders and files. A red arrow points from the "Visualización.ipynb" file entry in the list.

Editamos el archivo de visualizaciones.ipynb

The screenshot shows a Firefox browser window with multiple tabs open, indicating a Jupyter Notebook session on a Hadoop cluster. The main content area displays Python code for reading an Excel file using PySpark.

```
In [1]: import findspark
findspark.init()
import pyspark
from pyspark.sql import functions
from pyspark.sql import SparkSession
#ejecutar spark en el cluster
sc=spark.SparkContext(appName='Visualizacion',master='yarn')

#ejecutar spark en el contexto local
#sc=spark.SparkContext(appName='Visualizacion',master='local[2]')
spark=SparkSession.builder.getOrCreate()

Cargar datos de Excel

In [2]: #from pyspark.sql import SparkSession
#Import pandas

ruta = "/Archivos/Global_Superstore1.csv"
file_type = "csv"
infer_schema = "false"
first_row_is_header = "true"
delimiter = ","
Datos = spark.read.format(file_type) \
.option("inferSchema", infer_schema) \
.option("header", first_row_is_header) \
.option("sep", delimiter) \
.load(ruta)

df= Datos
df.show()
```

The screenshot shows a VMware Workstation window with a Firefox browser running. The browser has several tabs open, including 'nodo 1 Hadoop', 'nodo 2 Hadoop', 'nodo3 Hadoop', 'Actividades', 'Firefox', 'pruebaPythonSpark', 'ScriptsPythonBigData', 'Global Superstore...', 'Home Page - Select...', 'Visualización - Jupyter', 'EstadísticaBásica...', and 'POSGRADOS - UTM...'. The main content area of the browser is a Jupyter notebook titled 'Cargar datos de Excel'. The code in the notebook reads an Excel file named 'Global Superstore1.csv' and displays its contents as a DataFrame:

```
In [2]: #from pyspark.sql import SparkSession  
#import pandas  
  
ruta = "/Archivos/Global_Superstore1.csv"  
file_type = "csv"  
infer_schema = "false"  
first_col_is_header = "true"  
delimiter = ";"  
  
Datos = spark.read.format(file_type) \  
    .option("inferSchema", infer_schema) \  
    .option("header", first_col_is_header) \  
    .option("sep", delimiter) \  
    .load(ruta)  
  
df = Datos  
  
df.show()
```

The output of the code shows the following data:

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City
1	CA-2012-124891	31/7/2012	31/7/2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City
2	NE-2013-108241	10/8/2013	10/8/2013	Two Day	US-18451	East TEC-AC-1000033	Technology	Accessories
3	TR-2013-770878	7/2/2013	7/2/2013	Second Class	933-JR-16210	Justin Ritter	Corporate	Wollongon
4	IN-2013-770878	7/2/2013	7/2/2013	Second Class	933-JR-16210	Justin Ritter	Corporate	Wollongon
5	GR-2013-10000350	10/10/2013	10/10/2013	Two Day	OCEANIA-FUR-CH-10000350	Oceania Furniture	Corporate	Chairs
6	NO-2013-3709393	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
7	MX-2013-124891	17/10/2013	17/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
8	AU-2013-3709393	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
9	DE-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
10	GB-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
11	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
12	PT-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
13	IT-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
14	NL-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
15	SE-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
16	FI-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
17	DK-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
18	GR-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
19	IE-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
20	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
21	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
22	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
23	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
24	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
25	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
26	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
27	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
28	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
29	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
30	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
31	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
32	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
33	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
34	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
35	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
36	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
37	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
38	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
39	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
40	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
41	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
42	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
43	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
44	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
45	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
46	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
47	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
48	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
49	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
50	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
51	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
52	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
53	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
54	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
55	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
56	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
57	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
58	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
59	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
60	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
61	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
62	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
63	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
64	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
65	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
66	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
67	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
68	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
69	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
70	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
71	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
72	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
73	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
74	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
75	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
76	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
77	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
78	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
79	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
80	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
81	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
82	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
83	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
84	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
85	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
86	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
87	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
88	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
89	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
90	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
91	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
92	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
93	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
94	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
95	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
96	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
97	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
98	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
99	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
100	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
101	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
102	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
103	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
104	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
105	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
106	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
107	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
108	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
109	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
110	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
111	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
112	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
113	ES-2013-10000364	10/10/2013	10/10/2013	Two Day	925-CR-12730	Craig Heiter	Consumer	Brisbane
114	ES-2013-10000364	10/10/2013	10/10/2013</td					

