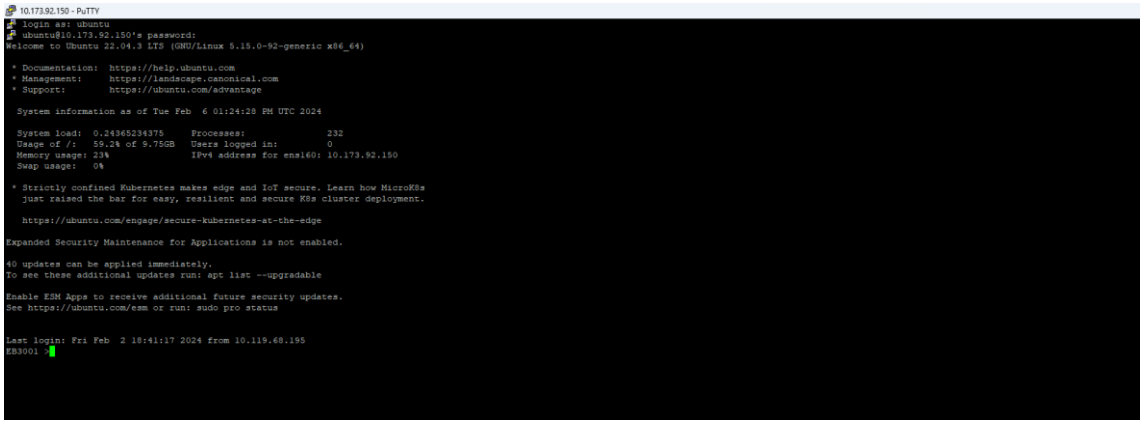
Installation Spark -

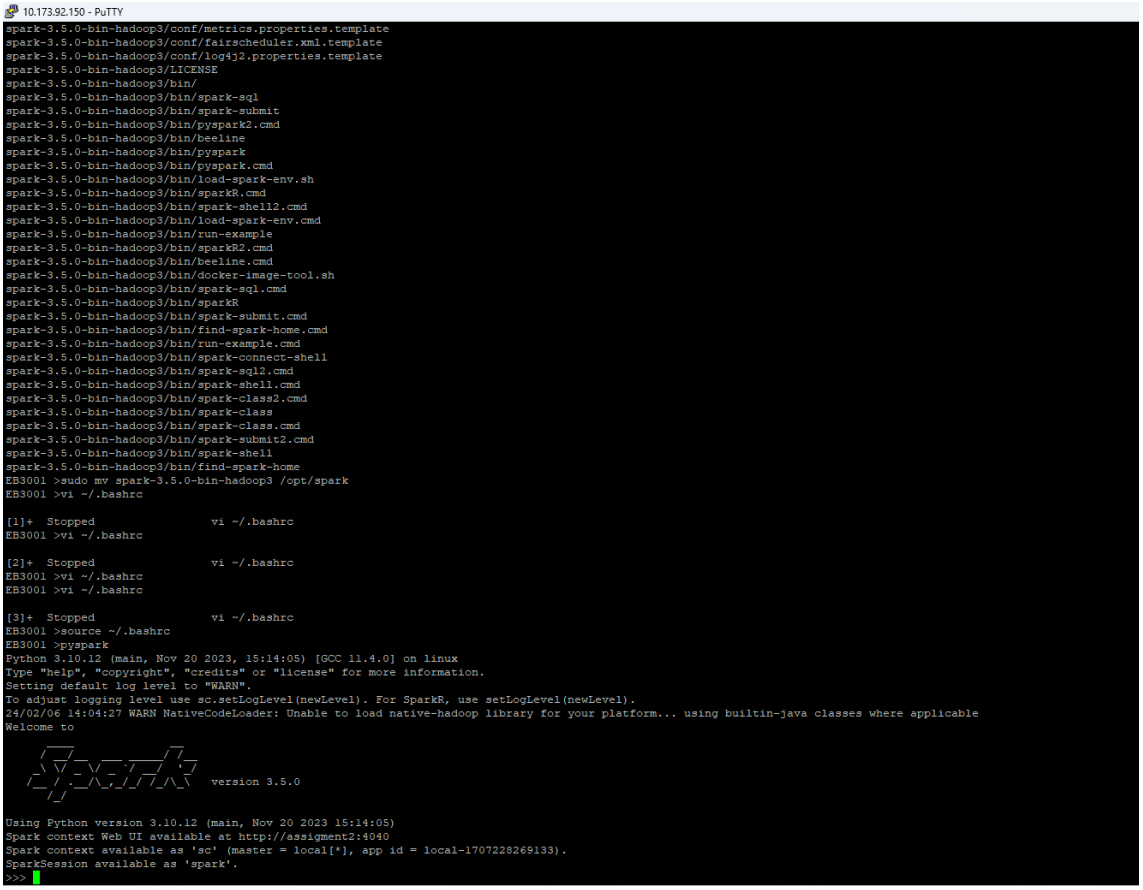


10.173.92.150 - PuTTY

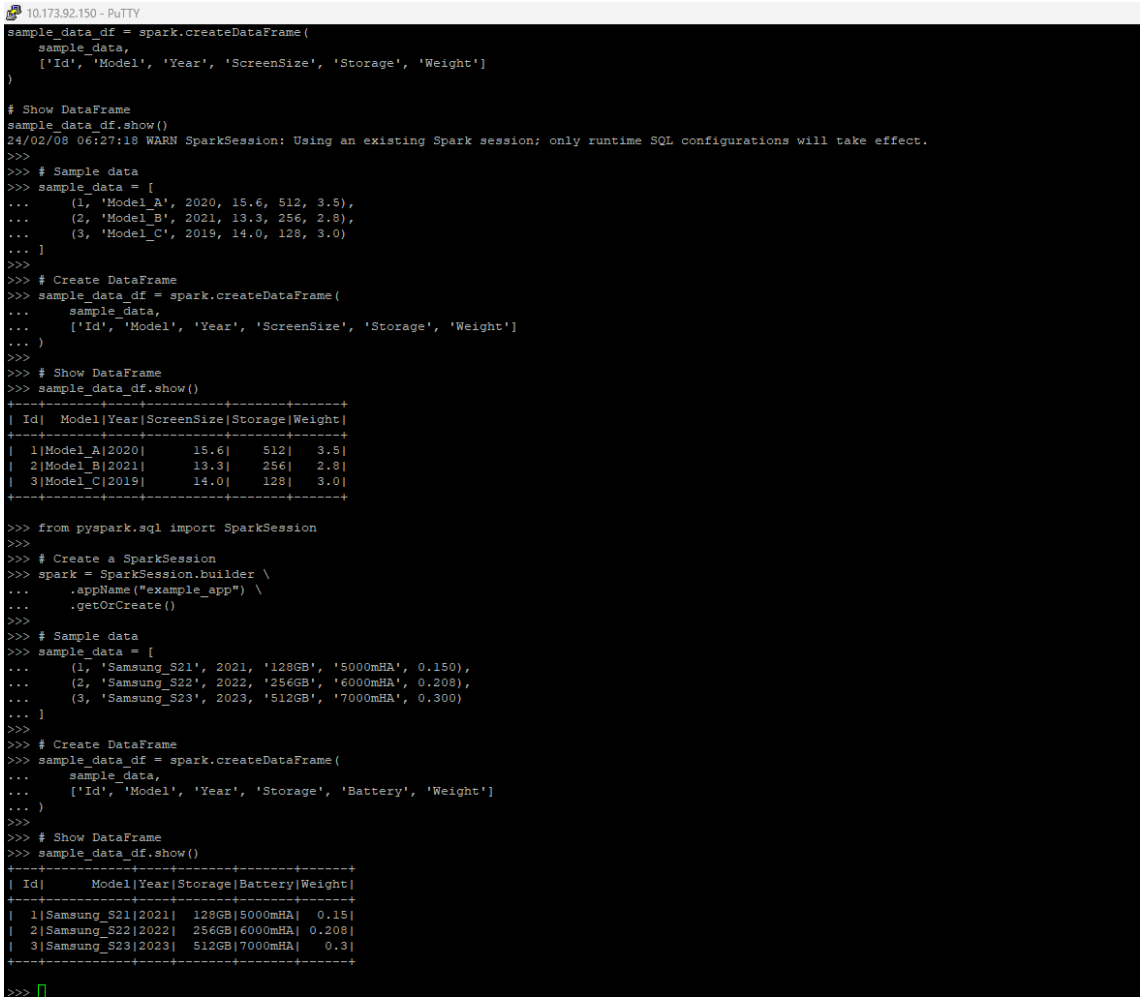
```
Unpacking libxdmcp-dev:amd64 (1:1.1.3-0ubuntu5) ...
Selecting previously unselected package xtrans-dev.
Preparing to unpack .../94-xtrans-dev_1.4.0-1_all.deb ...
Unpacking xtrans-dev (1.4.0-1) ...
Selecting previously unselected package libxcb1-dev:amd64.
Preparing to unpack .../95-libxcb1-dev_1.14-3ubuntu3_amd64.deb ...
Unpacking libxcb1-dev:amd64 (1.14-3ubuntu3) ...
Selecting previously unselected package libx11-dev:amd64.
Preparing to unpack .../96-libx11-dev_2%3a1.7.5-1ubuntu0.3_amd64.deb ...
Unpacking libx11-dev:amd64 (2:1.7.5-1ubuntu0.3) ...
Selecting previously unselected package libxt-dev:amd64.
Preparing to unpack .../97-libxt-dev_1%3a1.2.1-1_amd64.deb ...
Unpacking libxt-dev:amd64 (1:1.2.1-1) ...
Setting up libgraphite2-3:amd64 (1.3.14-1build2) ...
Setting up libxcb-dri3-0:amd64 (1.14-3ubuntu3) ...
Setting up liblcms2-2:amd64 (2.12~rc1-2build2) ...
Setting up libx11-xcb1:amd64 (2:1.7.5-1ubuntu0.3) ...
Setting up libpciaccess0:amd64 (0.16-3) ...
Setting up session-migration (0.3.6) ...
Created symlink /etc/systemd/user/graphical-session-pre.target.wants/session-migration.service → /usr/lib/systemd/user/session-migration.service.
Setting up libdrm-nouveau2:amd64 (2.4.113-2~ubuntu0.22.04.1) ...
Setting up libxcb-xfixes0:amd64 (1.14-3ubuntu3) ...
Setting up libxpm4:amd64 (1:3.5.12-1ubuntu0.22.04.2) ...
Setting up libxi6:amd64 (2:1.8-1build1) ...
Setting up java-common (0.72build2) ...
Setting up libxrender1:amd64 (1:0.9.10-1build4) ...
Setting up libdrm-radeon1:amd64 (2.4.113-2~ubuntu0.22.04.1) ...
Setting up libglvnd0:amd64 (1.4.0-1) ...
Setting up libxcb-glx0:amd64 (1.14-3ubuntu3) ...
Setting up libdrm-intel1:amd64 (2.4.113-2~ubuntu0.22.04.1) ...
Setting up libxcb-shape0:amd64 (1.14-3ubuntu3) ...
Setting up libglvnd0:amd64 (1.4.0-1) ...
Setting up libxcb-glx0:amd64 (1.14-3ubuntu3) ...
Setting up libdrm-intel1:amd64 (2.4.113-2~ubuntu0.22.04.1) ...
Setting up libxcb-shape0:amd64 (1.14-3ubuntu3) ...
Setting up x11-common (1:7.7+23ubuntu2) ...
Setting up libsensors-config (1:3.6.0-7ubuntu1) ...
Setting up libxxf86dga1:amd64 (2:1.1.5-0ubuntu3) ...
Setting up libxcb-shm0:amd64 (1.14-3ubuntu3) ...
Setting up libatspi2.0-0:amd64 (2.44.0-3) ...
Setting up libpthread-stubs0-dev:amd64 (0.4-1build2) ...
Setting up libxxf86vm1:amd64 (1:1.1.4-1build3) ...
Setting up libxcb-present0:amd64 (1.14-3ubuntu3) ...
Setting up libdconf1:amd64 (0.40.0-3) ...
Setting up libasound2-data (1.2.6.1-1ubuntu1) ...
Setting up xtrans-dev (1.4.0-1) ...
Setting up libfontenc1:amd64 (1:1.1.4-1build3) ...
Setting up libxfixes3:amd64 (1:6.0.0-1) ...
Setting up libxcb-sync1:amd64 (1.14-3ubuntu3) ...
Setting up libavahi-common-data:amd64 (0.8-5ubuntu5.2) ...
Setting up libxinerama1:amd64 (2:1.1.4-3) ...
Setting up libxv1:amd64 (2:1.0.11-1build2) ...
Setting up libxrandr2:amd64 (2:1.5.2-1build1) ...
Setting up fonts-dejavu-core (2.37-2build1) ...
Setting up libsensors5:amd64 (1:3.6.0-7ubuntu1) ...
Setting up libjpeg-turbo8:amd64 (2.1.2-0ubuntu1) ...
Setting up libglapi-mesa:amd64 (23.0.4-0ubuntu1~22.04.1) ...
Setting up libxcb-dri2-0:amd64 (1.14-3ubuntu3) ...
Setting up libgif7:amd64 (5.1.9-2build2) ...
Setting up libatk1.0-data (2.36.0-3build1) ...
Setting up fonts-dejavu-extra (2.37-2build1) ...

Progress: [ 70%] [##########################################################################################.........................]
```

10.173.92.150 - PuTTY

```
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for ca-certificates (20230311ubuntu0.22.04.1) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...

done.
done.
Setting up at-spi2-core (2.44.0-3) ...
Setting up openjdk-11-jre-headless:amd64 (11.0.21+9-0ubuntu1~22.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/java to provide /usr/bin/java (java) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jjs to provide /usr/bin/jjs (jjs) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/keytool to provide /usr/bin/keytool (keytool) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmid to provide /usr/bin/rmid (rmid) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmiregistry to provide /usr/bin/rmiregistry (rmiregistry) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/pack200 to provide /usr/bin/pack200 (pack200) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/unpack200 to provide /usr/bin/unpack200 (unpack200) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/lib/jexec to provide /usr/bin/jexec (jexec) in auto mode
Setting up openjdk-11-jre:amd64 (11.0.21+9-0ubuntu1~22.04) ...
Setting up openjdk-11-jdk-headless:amd64 (11.0.21+9-0ubuntu1~22.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jar to provide /usr/bin/jar (jar) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jarsigner to provide /usr/bin/jarsigner (jarsigner) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/javac to provide /usr/bin/javac (javac) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/javadoc to provide /usr/bin/javadoc (javadoc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/javap to provide /usr/bin/javap (javap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jcmd to provide /usr/bin/jcmd (jcmd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdb to provide /usr/bin/jdb (jdb) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeprscan to provide /usr/bin/jdeprscan (jdeprscan) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeps to provide /usr/bin/jdeps (jdeps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jfr to provide /usr/bin/jfr (jfr) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jimage to provide /usr/bin/jimage (jimage) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jinfo to provide /usr/bin/jinfo (jinfo) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jlink to provide /usr/bin/jlink (jlink) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmap to provide /usr/bin/jmap (jmap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmod to provide /usr/bin/jmod (jmod) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jps to provide /usr/bin/jps (jps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jrunscript to provide /usr/bin/jrunscript (jrunscript) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jshell to provide /usr/bin/jshell (jshell) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstack to provide /usr/bin/jstack (jstack) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstat to provide /usr/bin/jstat (jstat) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstatd to provide /usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmic to provide /usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/serialver to provide /usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jaotc to provide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up default-jre (2:1.11-72build2) ...
Setting up default-jdk-headless (2:1.11-72build2) ...
Setting up openjdk-11-jdk:amd64 (11.0.21+9-0ubuntu1~22.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Setting up default-jdk (2:1.11-72build2) ...
Scanning processes...
Scanning linux images...

Running kernel seems to be up-to-date.

No services need to be restarted.

No containers need to be restarted.

No user sessions are running outdated binaries.

No VM guests are running outdated hypervisor (qemu) binaries on this host.
EB3001 >
```


10.173.92.150 - PuTTY
spark-3.5.0-bin-hadoop3/conf/metrics.properties.template
spark-3.5.0-bin-hadoop3/conf/fairscheduler.xml.template
spark-3.5.0-bin-hadoop3/conf/log4j2.properties.template
spark-3.5.0-bin-hadoop3/LICENSE
spark-3.5.0-bin-hadoop3/bin/
spark-3.5.0-bin-hadoop3/bin/spark-sql
spark-3.5.0-bin-hadoop3/bin/spark-submit
spark-3.5.0-bin-hadoop3/bin/pyspark2.cmd
spark-3.5.0-bin-hadoop3/bin/beeline
spark-3.5.0-bin-hadoop3/bin/pyspark
spark-3.5.0-bin-hadoop3/bin/pyspark.cmd
spark-3.5.0-bin-hadoop3/bin/load-spark-env.sh
spark-3.5.0-bin-hadoop3/bin/sparkR.cmd
spark-3.5.0-bin-hadoop3/bin/spark-shell2.cmd
spark-3.5.0-bin-hadoop3/bin/load-spark-env.cmd
spark-3.5.0-bin-hadoop3/bin/run-example
spark-3.5.0-bin-hadoop3/bin/sparkR2.cmd
spark-3.5.0-bin-hadoop3/bin/beeline.cmd
spark-3.5.0-bin-hadoop3/bin/docker-image-tool.sh
spark-3.5.0-bin-hadoop3/bin/spark-sql.cmd
spark-3.5.0-bin-hadoop3/bin/sparkR
spark-3.5.0-bin-hadoop3/bin/spark-submit.cmd
spark-3.5.0-bin-hadoop3/bin/find-spark-home.cmd
spark-3.5.0-bin-hadoop3/bin/run-example.cmd
spark-3.5.0-bin-hadoop3/bin/spark-connect-shell
spark-3.5.0-bin-hadoop3/bin/spark-sql2.cmd
spark-3.5.0-bin-hadoop3/bin/spark-shell.cmd
spark-3.5.0-bin-hadoop3/bin/spark-class2.cmd
spark-3.5.0-bin-hadoop3/bin/spark-class
spark-3.5.0-bin-hadoop3/bin/spark-class.cmd
spark-3.5.0-bin-hadoop3/bin/spark-submit2.cmd
spark-3.5.0-bin-hadoop3/bin/spark-shell
spark-3.5.0-bin-hadoop3/bin/find-spark-home
EB3001 >sudo mv spark-3.5.0-bin-hadoop3 /opt/spark
EB3001 >vi ~/.bashrc
[1]+ Stopped vi ~/.bashrc
EB3001 >vi ~/.bashrc
[2]+ Stopped vi ~/.bashrc
EB3001 >vi ~/.bashrc
EB3001 >vi ~/.bashrc
[3]+ Stopped vi ~/.bashrc
EB3001 >source ~/.bashrc
EB3001 >pyspark
Python 3.10.12 (main, Nov 20 2023, 15:14:05) [GCC 11.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
24/02/06 14:04:27 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Welcome to
version 3.5.0
Using Python version 3.10.12 (main, Nov 20 2023 15:14:05)
Spark context Web UI available at http://assigment2:4040
Spark context available as 'sc' (master = local[*], app id = local-1707228269133).
SparkSession available as 'spark'.
>>>

In this step–up sample data – Samsung models



Creating DataFrames