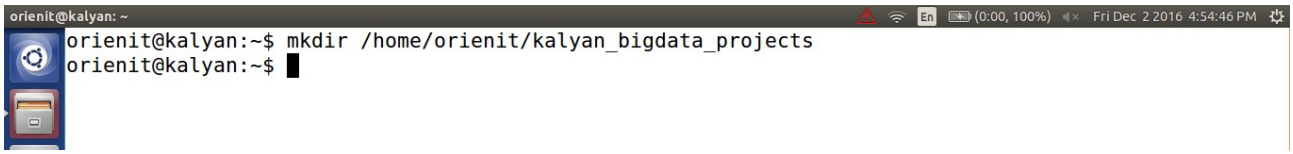
# Kalyan Big Data Real Time Projects

How to generate **large amount of sample data** with simple techniques for **Big Data Projects.** Follow the below commands to generate **large amount of sample data**.

i) Create '**kalyan_bigdata_projects**' folder in **user home** (i.e **/home/orienit**)

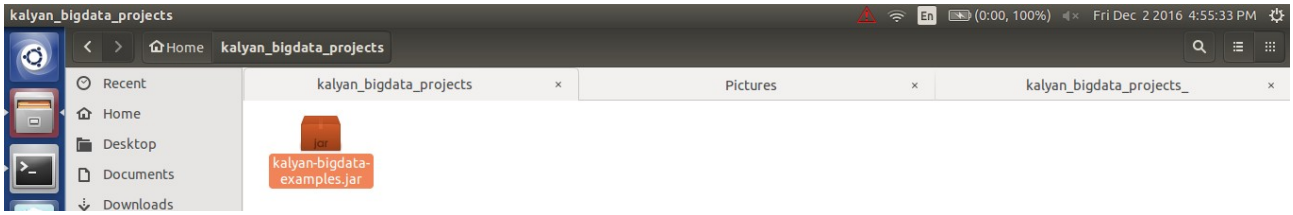
**Command:** *mkdir /home/orienit/kalyan_bigdata_projects*




ii) Download '**kalyan-bigdata-examples.jar**' jar file from this link.

(https://github.com/kalyanhadooptraining/kalyan-bigdata-realtime-projects/blob/master/kalyan/kalyan-bigdata-examples.jar)

iii) Copy '**kalyan-bigdata-examples.jar**' jar file into '**/home/orienit/kalyan_bigdata_projects**' folder




## We are going to learn below Use Cases

**Use Case 1:** Generating Sample **Server Logs** with simple command

**Use Case 2:** Generating Sample **Users** in **JSON** format with simple command

**Use Case 3:** Generating Sample **Users** in **CSV** format with simple command

**Use Case 4:** Generating Sample **Users** in **TSV** format with simple command

**Use Case 5:** Generating Sample **Users** in **DELIMITED** format with simple command

**Use Case 6:** Generating Sample **Product Log** in **JSON** format with simple command

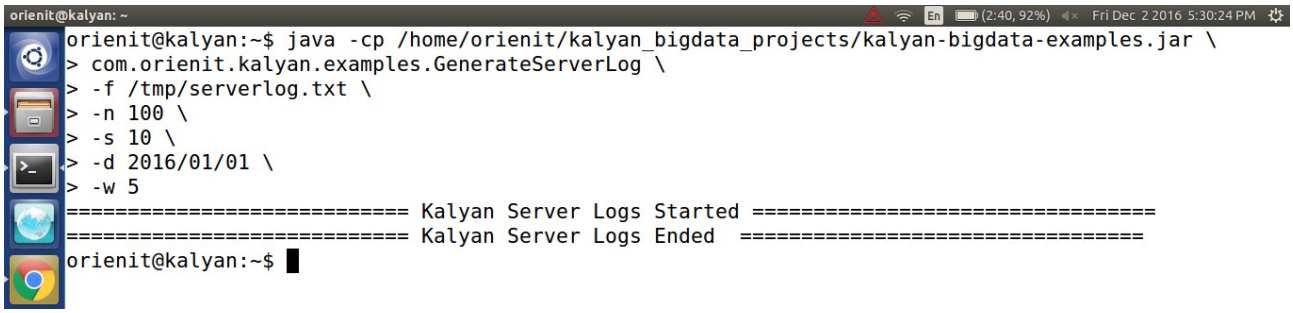
**Use Case 7:** Generating Sample **Product Log** in **CSV** format with simple command

**Use Case 8:** Generating Sample **Product Log** in **TSV** format with simple command

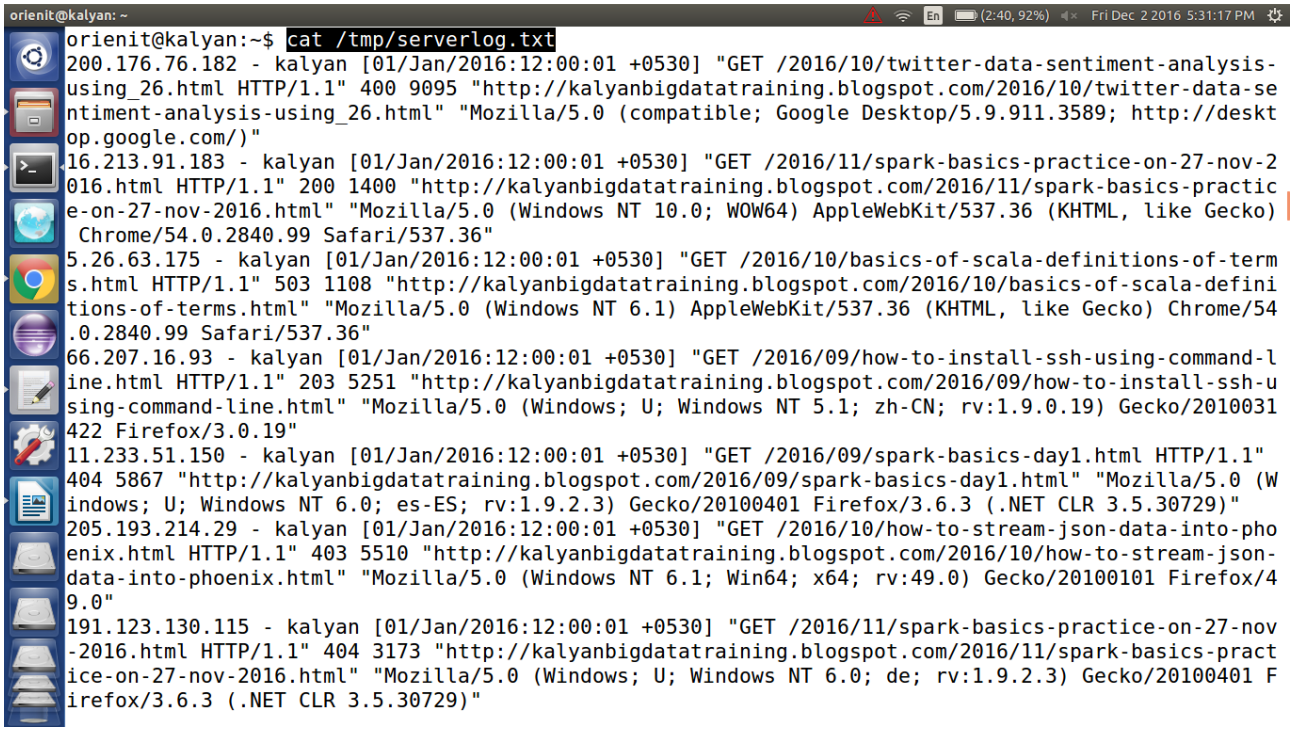
**Use Case 9:** Generating Sample **Product Log** in **DELIMITED** format with simple command

# Kalyan Big Data Projects – Use Case 1
# Generating Sample Server Logs with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateServerLog \
-f /tmp/serverlog.txt \
-n 100 \
-s 10 \
-d 2016/01/01 \
-w 5
```

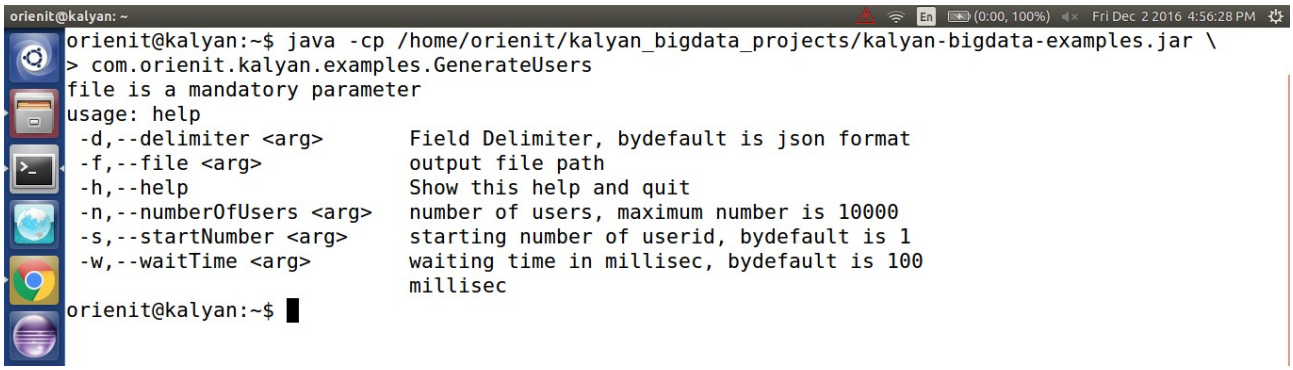



Read **SERVER LOG** data



# Kalyan Big Data Projects – Use Case
# Generating Sample Users with simple command

java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateUsers

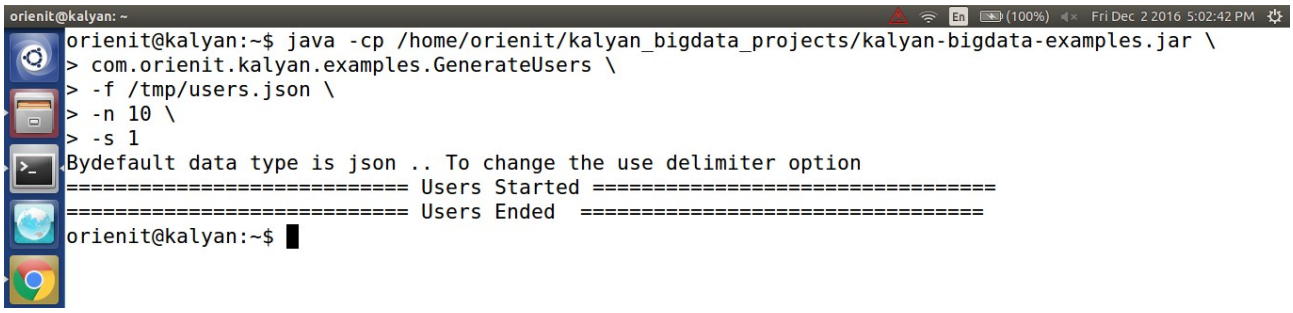



We can pass different aguments for above command

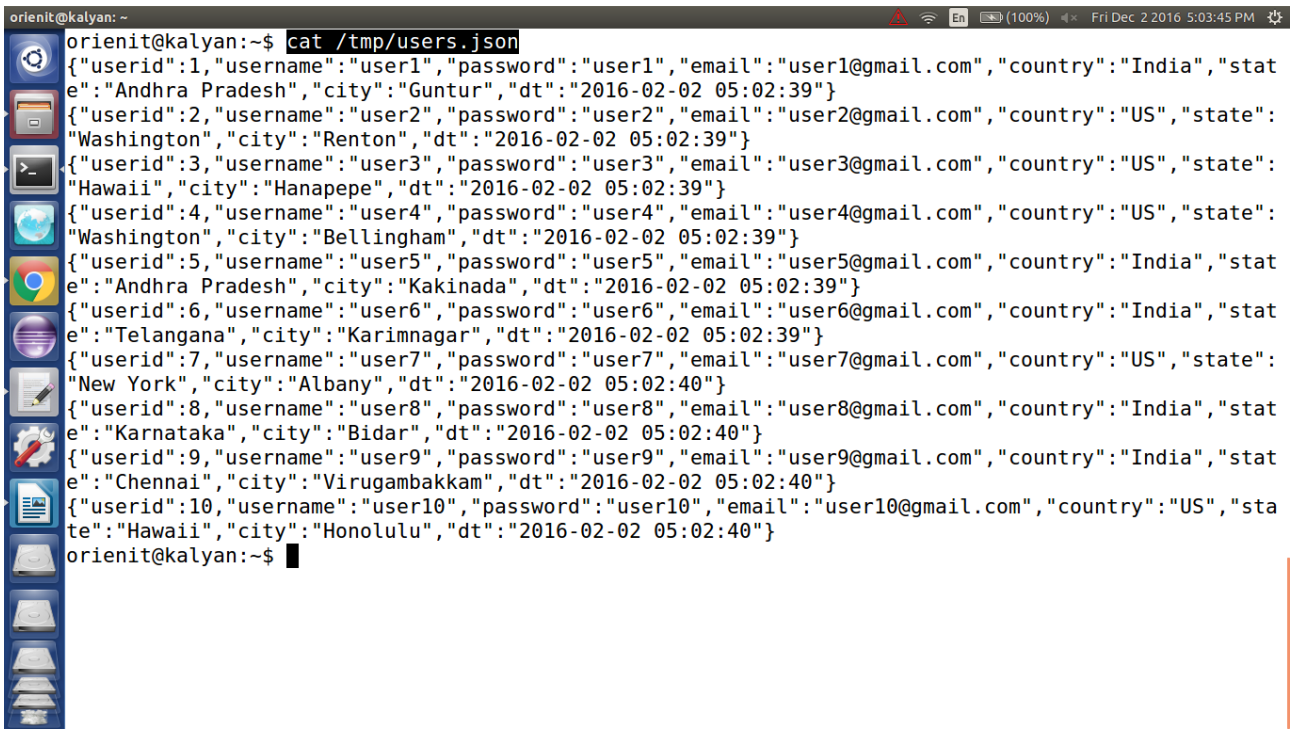
-d => field delimiter like (tab, comma, semicolon, etc )
-f => output file path
-n => number of users, maximum number is 10000
-s => starting number of user id, bydefault is 1
-w => waiting time in milli sec, bydefault is 100 millisec

# Kalyan Big Data Projects – Use Case 2
# Generating Sample Users in JSON format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateUsers \
-f /tmp/users.json \
-n 10 \
-s 1
```

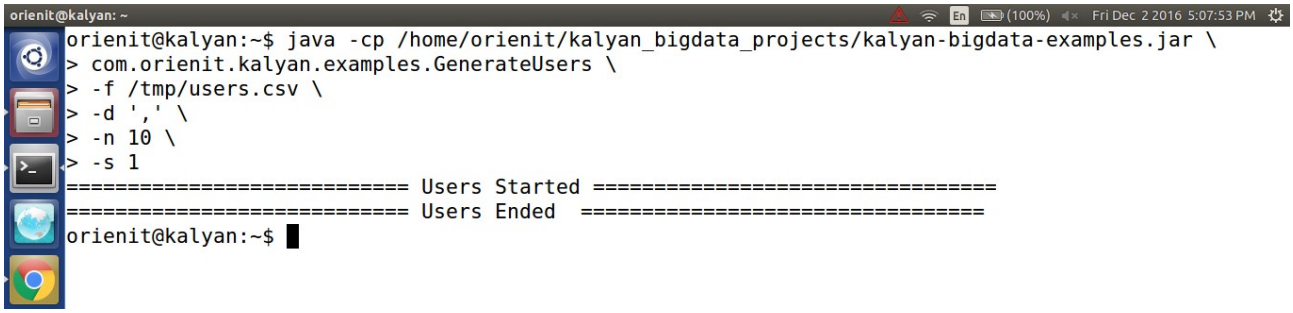



Read **JSON** Data

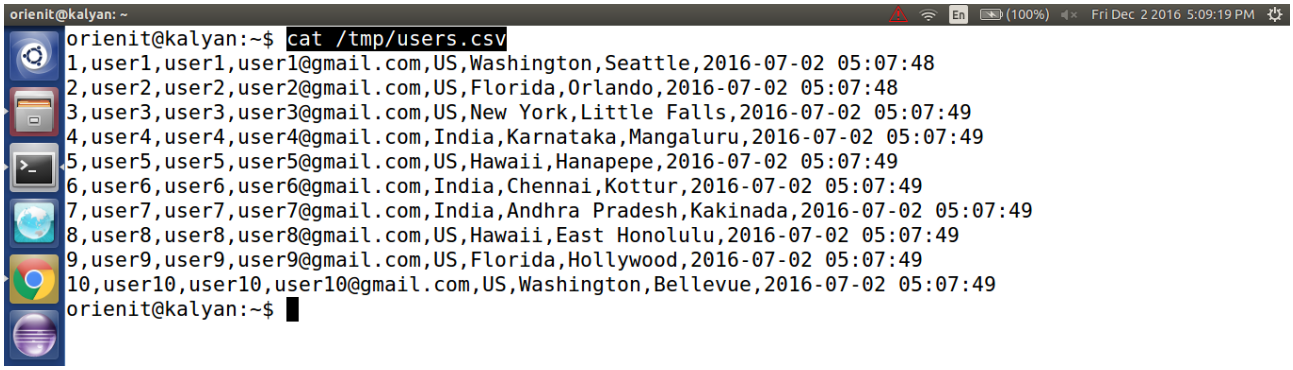


# Kalyan Big Data Projects – Use Case 3
# Generating Sample Users in CSV format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateUsers \
-f /tmp/users.csv \
-d ',' \
-n 10 \
-s 1
```

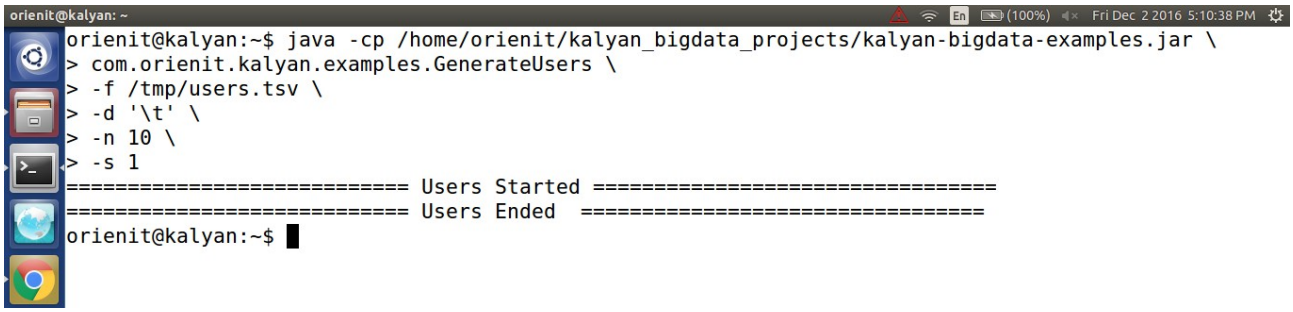



Read **CSV** data

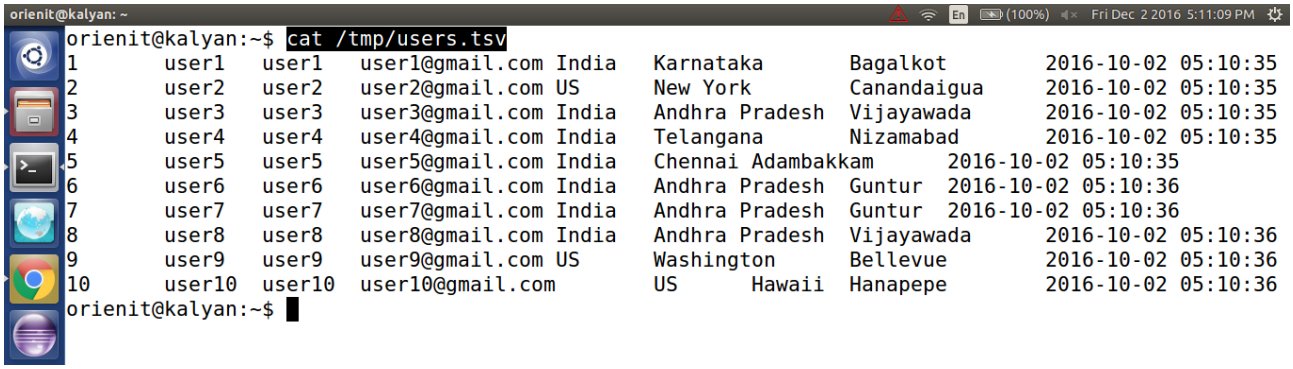


# Kalyan Big Data Projects – Use Case 4
# Generating Sample Users in TSV format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateUsers \
-f /tmp/users.tsv \
-d '\t' \
-n 10 \
-s 1
```

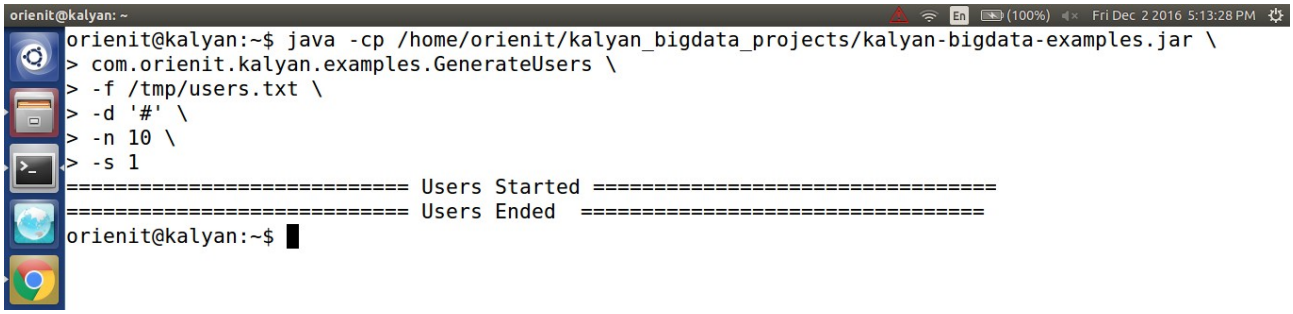



Read **TSV** data

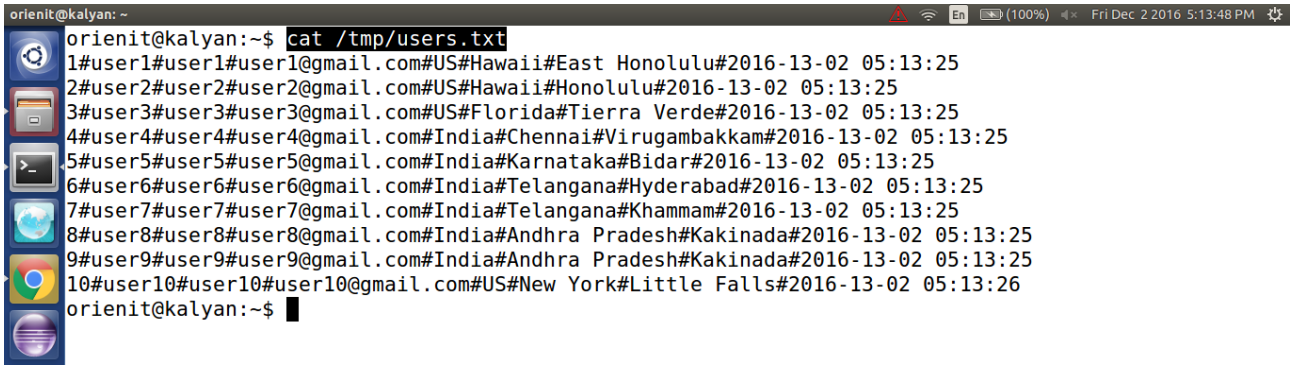


# Kalyan Big Data Projects – Use Case 5
# Generating Sample Users in DELIMITER format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateUsers \
-f /tmp/users.txt \
-d '#' \
-n 10 \
-s 1
```

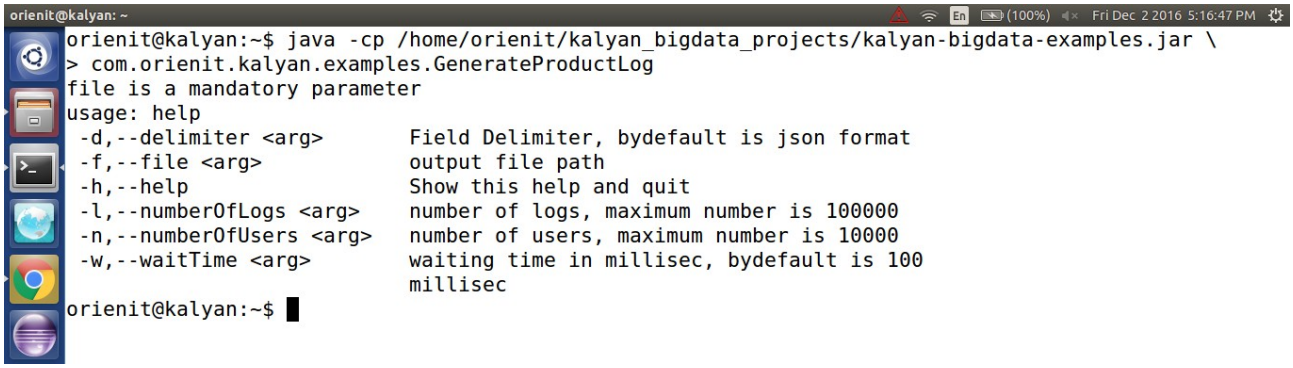



Read Any **DELIMITED** Data



# Kalyan Big Data Projects – Use Case
# Generating Sample Product Log with simple command

java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateProductLog

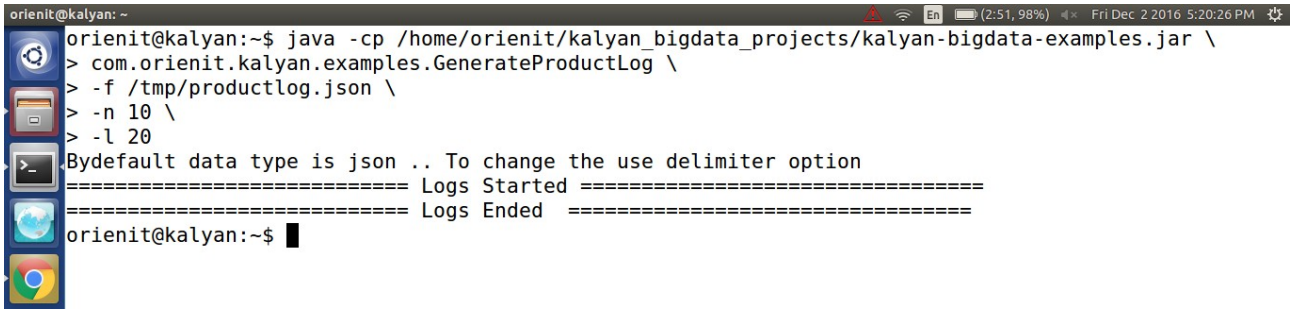



We can pass different aguments for above command

-d => field delimiter like (tab, comma, semicolon, etc )
-f => output file path
-l => number of logs, maximum number is 100000
-n => number of users, maximum number is 10000
-w => waiting time in milli sec, bydefault is 100 millisec

# Kalyan Big Data Projects – Use Case 6
# Generating Sample Product Log in JSON format with simple command

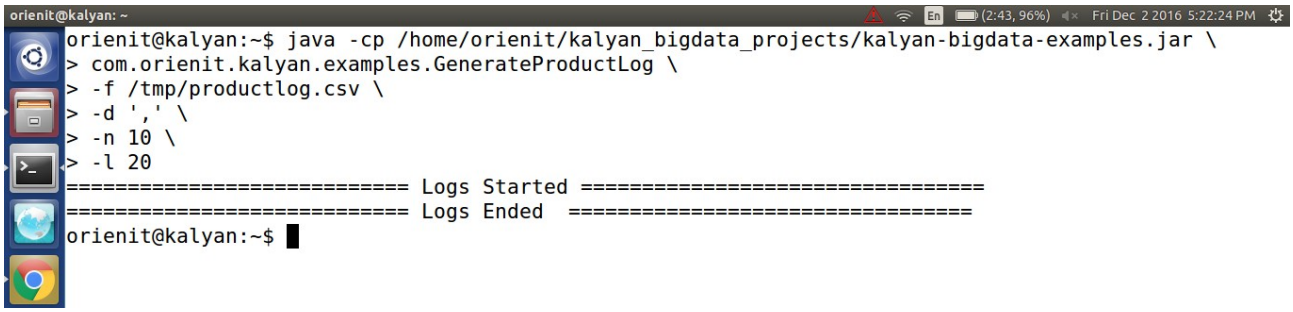
```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateProductLog \
-f /tmp/productlog.json \
-n 10 \
-l 20
```



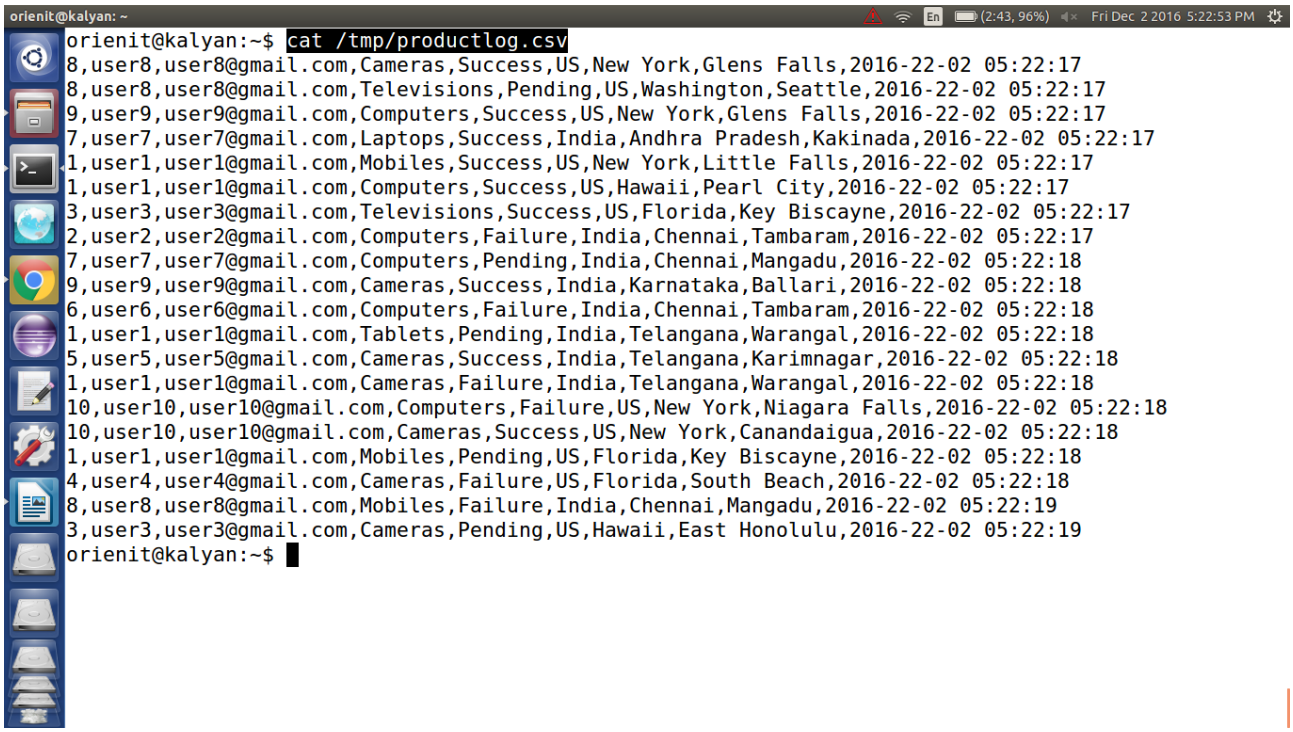

Read **JSON** data

# Kalyan Big Data Projects – Use Case 7
# Generating Sample Product Log in CSV format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateProductLog \
-f /tmp/productlog.csv \
-d ',' \
-n 10 \
-l 20
```

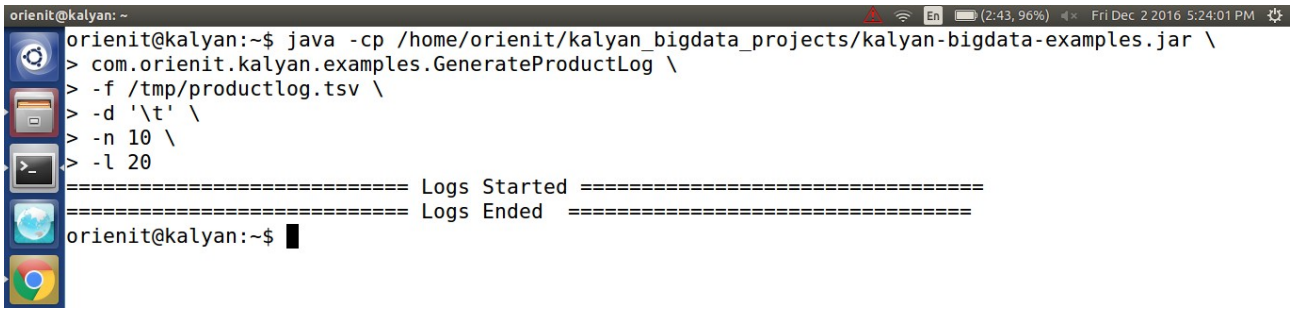



Read **CSV** data

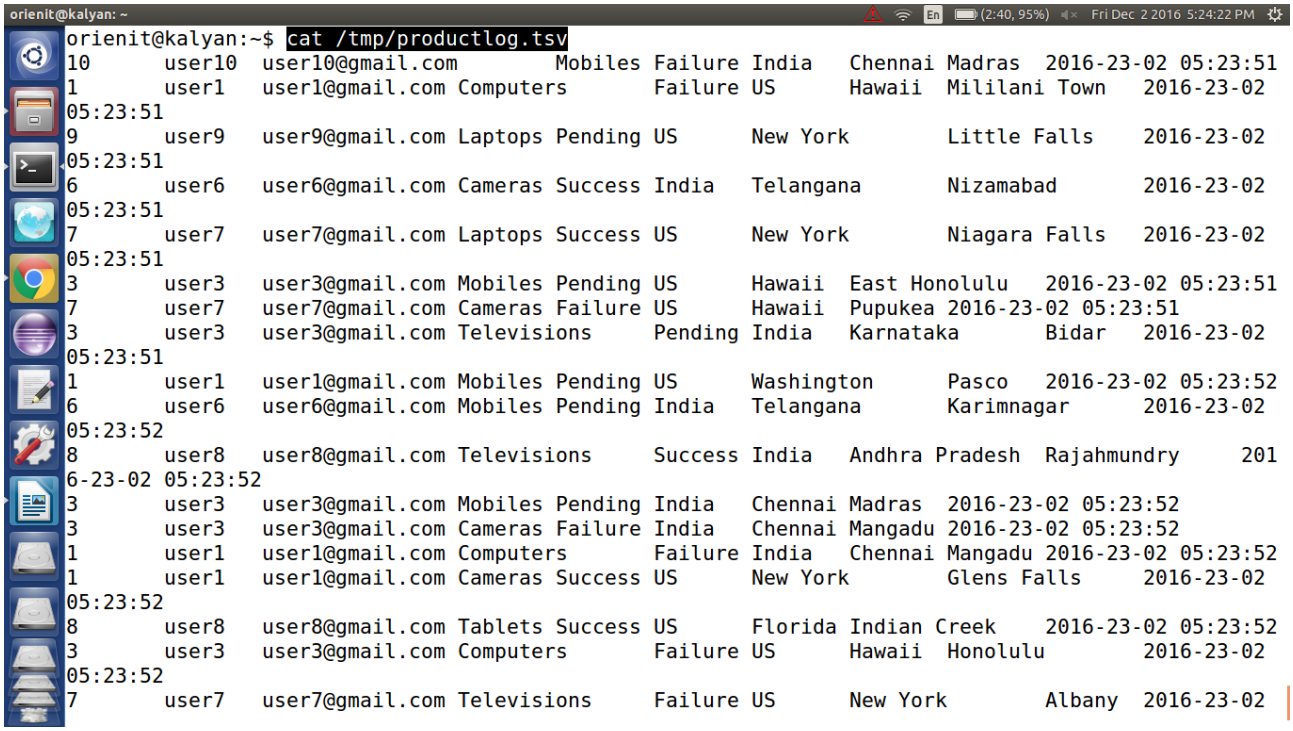


# Kalyan Big Data Projects – Use Case 6
# Generating Sample Product Log in TSV format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateProductLog \
-f /tmp/productlog.tsv \
-d '\t' \
-n 10 \
-l 20
```

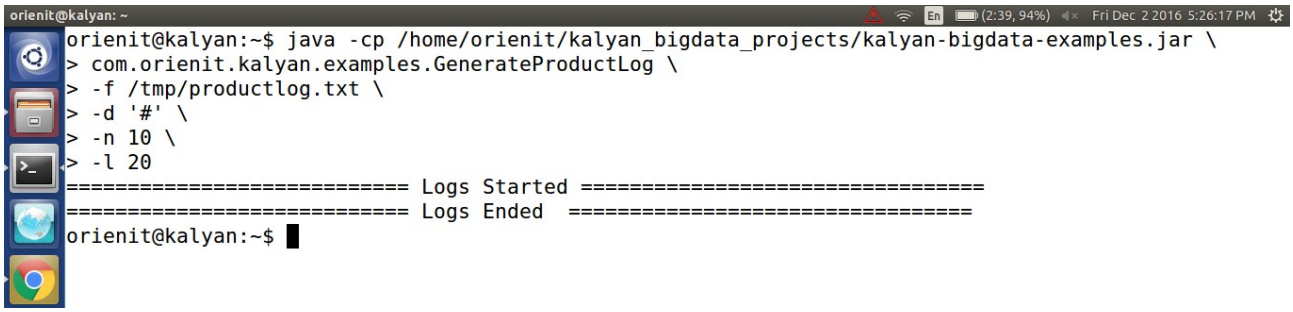



Read **TSV** data



# Kalyan Big Data Projects – Use Case 9
# Generating Sample Product Log in DELIMITED format with simple command

```
java -cp /home/orienit/kalyan_bigdata_projects/kalyan-bigdata-examples.jar \
com.orienit.kalyan.examples.GenerateProductLog \
-f /tmp/productlog.txt \
-d '#' \
-n 10 \
-l 20
```




Read Any **DELIMITED** data