PRACTICE CASE FOR SOCIAL MEDIA ANALYSIS WITH SPARK

Import Library dari Python API

```
      __              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /__ / .__/\_,_/_/ /_/\_\   version 2.4.0.cloudera2
      /_/

Using Python version 2.7.5 (default, Aug  7 2019 00:51:29)
SparkSession available as 'spark'.
```

Load Data (.csv)

```
>>> df = spark.read.csv('/user/cloudera/clean_tweet.csv', header=True)
>>> df.show()
+--------------------+------+
|                text|target|
+--------------------+------+
|awww that s a bum...|     0|
|is upset that he ...|     0|
|i dived many time...|     0|
|my whole body fee...|     0|
|no it s not behav...|     0|
|   not the whole crew|     0|
|          need a hug|     0|
|hey long time no ...|     0|
|k nope they didn ...|     0|
|        que me muera|     0|
|spring break in p...|     0|
|i just re pierced...|     0|
|i couldn t bear t...|     0|
|it it counts idk ...|     0|
|i would ve been t...|     0|
|i wish i got to w...|     0|
|hollis death scen...|     0|
|  about to file taxes|     0|
|ahh ive always wa...|     0|
|oh dear were you ...|     0|
+--------------------+------+
only showing top 20 rows
```