

SUBMISSION 1:

Create an RDD with tuples where there is a key and a value. But in contrast to the example above the key is removed from the value portion of the key-value tuple. Submit the code and a print out of the first tuple.

I went about this a little different:

I defined a function to create tuples from RDD:

```
>>> def tuple_maker(item): return item[0], item[1:]
```

Then I applied it to first entry to test:

```
>>> test = narcoticsCrimeRecords.first()
```

```
>>> first_tuple = tuple_maker(test)
```

```
>>> first_tuple[0]  
u'10184515'
```

```
>>> first_tuple[1]  
[u'HY372204', u'08/06/2015 11:55:00 PM', u'033XX W DIVERSEY AVE', u'2027',  
u'NARCOTICS', u'POSS: CRACK', u'STREET', u'true', u'false', u'1412', u'014', u'35',  
u'22', u'18', u'1153440', u'1918377', u'2015', u'08/13/2015 12:57:42 PM',  
u'41.931870591', u'-87.711546895', u'\"(41.931870591', u' -87.711546895)\""]
```

SUBMISSION 2:

Submit the code for executing the above query as a Spark SQL python call. Also submit the number of rows in the result.

Spark SQL python call:

```
>>> results = sqlContext.sql('SELECT count(*) from Web_Session_Log where  
REFERERURL = "http://www.ebay.com"')
```

```
>>> results.show()
```

Number of results:

```
+----+
|_c0|
+----+
|3943|
+--+-+
```

Screenshot of additional query:

```
w205_files -- w205@ip-172-31-89-15:~ -- ssh -- 145x33

[3943]

>>> results = sqlContext.sql('SELECT * FROM Web_Session_Log')
>>> results.show()

+-----+-----+-----+-----+-----+
| DATETIME | USERID | SESSIONID | PRODUCTID | REFERERURL |
+-----+-----+-----+-----+-----+
| date | userid | sessionid | productid | refererurl |
+-----+-----+-----+-----+-----+
| 2008-01-31 15:54:25 | RequestVerifica... | ;+.ASPXAUTH=C31HD... | /product/YJ29I0CVQ | http://www.abc.com |
| 2005-12-08 02:36:30 | RequestVerifica... | ;+.ASPXAUTH=H7HTS... | /product/MVI9HHP8A | http://www.ebay.com |
| 2015-06-07 23:27:58 | RequestVerifica... | ;+.ASPXAUTH=58SZL... | /search/P5XK03AC9 | http://www.abc.com |
| 2009-03-12 03:16:27 | RequestVerifica... | ;+.ASPXAUTH=VBWZJ... | /product/A13025WBT | http://www.shophe... |
| 2014-07-23 08:36:03 | RequestVerifica... | ;+.ASPXAUTH=VXBLE... | /search/SPI9XD6LZ | http://www.facebo... |
| 2002-12-30 08:42:09 | RequestVerifica... | ;+.ASPXAUTH=YABJB... | /product/WS80XFW2 | http://www.xyz.com |
| 2004-11-03 20:29:10 | RequestVerifica... | ;+.ASPXAUTH=2F90N... | /product/01201IBUN | http://www.homesh... |
| 2012-01-26 12:39:57 | RequestVerifica... | ;+.ASPXAUTH=SEWRR... | /product/OA3QGXFIJ | http://www.xyz.com |
| 2008-04-30 02:01:34 | RequestVerifica... | ;+.ASPXAUTH=60B10... | /search/K1IRBE1DU | http://www.abc.com |
| 2003-08-23 09:44:43 | RequestVerifica... | ;+.ASPXAUTH=1NRGS... | /product/ANGKDKMKM | http://www.shophe... |
| 2008-04-09 01:24:24 | RequestVerifica... | ;+.ASPXAUTH=2Y8NA... | /product/LC94NBS9A | http://www.facebo... |
| 2000-08-07 06:45:19 | RequestVerifica... | ;+.ASPXAUTH=KS9LL... | /search/HDKWJ5ORV | http://www.facebo... |
| 2013-10-09 05:22:31 | RequestVerifica... | ;+.ASPXAUTH=UA1WH... | /search/5LPS3BTJI | http://www.facebo... |
| 2006-07-31 08:12:44 | RequestVerifica... | ;+.ASPXAUTH=GVW0... | /search/8W80TIDQP | http://www.xyz.com |
| 2014-07-27 13:23:18 | RequestVerifica... | ;+.ASPXAUTH=0Y5S5... | /search/D5S8HFH9D | http://www.facebo... |
| 2001-01-10 18:23:03 | RequestVerifica... | ;+.ASPXAUTH=VM0YI... | /product/I8VLXARNO | http://www.xyz.com |
| 2011-09-24 21:28:13 | RequestVerifica... | ;+.ASPXAUTH=SDVEA... | /search/S44PIHRYX | http://www.shophe... |
| 2008-09-19 02:52:53 | RequestVerifica... | ;+.ASPXAUTH=7NEBV... | /search/CX28DBZYW | http://www.shophe... |
| 2006-03-01 20:10:27 | RequestVerifica... | ;+.ASPXAUTH=K58W1... | /product/GGBEXER8K | http://www.amazon... |
+-----+-----+-----+-----+-----+

only showing top 20 rows

>>> |
```