Project 1 Supplementary Explanation

[File Composition]

If you download "part1.zip" file, the part1 folder is as follow:

ebay_data
bulk-loading.pdf
project-1-instructions.pdf
skeleton_parser.py

And to submit Part 1 of the project, gather the following files in a single submission directory named "part1_{your_student_id}" ex) part1_20161234:

design.pdf
{your_parser_with_student_id}.py ex)20161234.py which is based on skeleton_parser.py
runParser.sh
create.sql
load.txt
query1~7.sql

The files marked in yellow are the files you need to create. And I'll give you a hint about files marked in green. (refer to below).

Based on ebay_data, how to make a schema is your freedom. So, it doesn't matter how many tables you make or what you name it. But, you need to explain the information about your schema in "design.pdf" file.

[Project Summary]

1. Make a parser which converts the JSON files to .dat files (.dat is the data file extension)
   In data file, the attributes should be separated by delimiter '|'

ex>

```
"zysset"|"NE"|"USA"|574
"zyzest"|"ST. JOHN'S,NF"|"Canada"|781
"zzclassiccollectibles"|"San Diego, CA"|"USA"|703
"zzphillipsa"|"CHARLOTTE, NC"|"USA"|36
```

(It's just sample. This is my "user.dat" file. You don't have to do this. Whatever field you make it's up to you.)

2. The process of putting information of data file into sqlite3 database.
3. Solve 7 problems in the document(query1~7.sql)

(The following steps are required to perform the above tasks.)

1.

./runParser.sh (screenshot)

```
#RunParser
rm -rf *.dat;
python my_parser.py ./ebay_data/items-*.json

sort user.dat | uniq > user_table.dat
sort bids.dat | uniq > bids_table.dat
sort category.dat | uniq > category_table.dat
sort item.dat | uniq > item_table.dat
```

2.

sqlite3 test.db < create.sql

sqlite3 test.db < load.txt (screenshot)

```
.separator |
.import item_table.dat Item
UPDATE Item SET Buy_Price = null WHERE Buy_Price = 'NULL';
UPDATE Item SET Description = null WHERE Description = 'NULL';
.import user_table.dat User
UPDATE User SET Country = null WHERE Country = 'NULL';
UPDATE User SET Location = null WHERE Location = 'NULL';
.import bids_table.dat Bids
.import category_table.dat Category
```

(Now, you are ready to create your query. (query1~7.sql))

(Note: All screenshots are based on my schema, so you need to modify it to apply your system.)

[Development Environment]

Use Python3 Not Python2

(Please re-download skeleton_parser.py on Facebook or KLMS.)

Case 1. (Window)

- Install bash shell (Search Google for the keyword “Window bash shell install”)
- Use bash

Case 2. (Mac OS)

- Use terminal (You don’t have to install sqlite3. It’s already installed.)