Real World Datasets

.csv file is popular in read world

Splitting queries to multiple lines in Jupyter

Use backslash "\" to split the query into multiple lines:

single/double quotes to handle name with space

```
%sql SELECT Id, `Name of Dog`, \
     FROM dogs \
     WHERE `Name of Dog`='Huggy'
```

→ *in Python, we can use backslash to split query in multiple line for improve readability*

Or use %%sql in the first row of the cell in the notebook:

`%%sql` → *to each line redundant*

```
SELECT Id, `Name of Dog`,
     FROM dogs
     WHERE `Name of Dog`='Huggy'
```

Using quotes in Jupyter notebooks

used pd

```
data = pandas.read_sql(query_statement,
connection_variable)
```

First assign queries to variables:

Double for column's name

single for SQL query

```
query_statement = 'SELECT "Name of Dog" FROM dogs'
```

Use a backslash \ as the escape character in cases where the query contains single quotes:

```
query_statement = 'SELECT * FROM dogs
     WHERE "Name of Dog"=\'Huggy\' '
```

Getting Table and Column Details

Getting a list of tables in the database

DB2
SYSCAT.TABLES

SQLite3
sqlite_master

MySQL
SHOW TABLES

Getting a list of tables in the database

Query system catalog to get a list of tables & their properties:

SQLite3
SELECT name FROM sqlite_master WHERE type="table"

MySQL
SHOW TABLES

Now retrieve the the list of columns in SCHOOLS table and their column type (datatype) and length.

```
58]:  # type in your query to retrieve all column names in the SCHOOLS table along with their datatypes and length
      %sql SELECT name,type,length(type) FROM PRAGMA_TABLE_INFO('Chicago');
       * sqlite:///RealWorldData.db
      Done.
```

Get all names columns and details

name	type	length(type)
School_ID	INTEGER	7
NAME_OF_SCHOOL	TEXT	4
Elementary, Middle, or High School	TEXT	4
Street_Address	TEXT	4
City	TEXT	4
State	TEXT	4
ZIP_Code	INTEGER	7
Phone_Number	TEXT	4