

FE512 Final Project Data Import 1

```
In [61]: %load_ext sql
%sql mysql+pymysql://root:@fe512_mysql/fe512db
#%sql CREATE DATABASE hint
%sql USE hint;
%sql SELECT DATABASE();
%sql SHOW TABLES;
```

The sql extension is already loaded. To reload it, use:

```
%reload_ext sql
* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.
* mysql+pymysql://root:***@fe512_mysql/fe512db
19 rows affected.
```

Out[61]:

Tables_in_hint
Cancer_Cat_Value_Label
Race_Cat2_Value_Label
cancer_info
prostate_analysis
section_a
section_b
section_c
section_d
section_e
section_f
section_g
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

Initial process step1:Show datasheet already exist in FE512DB

Create Table in Database_Section_A

```
In [3]: %%sql
DROP TABLE IF EXISTS section_a;
CREATE TABLE IF NOT EXISTS section_a (
    PersonID VARCHAR(20),
    SeekHealthInfo INT,
    WhereSeekHealthInfo INT,
    WhoLookingFor INT,
    LotOfEffort INT,
    Frustrated INT,
    ConcernedQuality INT,
    TooHardUnderstand INT,
    ConfidentGetHealthInf INT,
    TrustDoctor INT,
    TrustFamily INT,
    TrustNewsMag INT,
    TrustRadio INT,
    TrustInternet INT,
    TrustTelevision INT,
    TrustGov INT,
    TrustCharities INT,
    TrustReligiousOrgs INT,
    StrongNeedHealthInfo INT,
    StrongNeedHealthInfo_OS VARCHAR(255),
    SeekCancerInfo INT,
    InternetCancerInfoSelf INT
);

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_a'")
  result = self._query(query)

Out[3]: []
```

In [4]: `%sql SHOW TABLES;`

* mysql+pymysql://root:***@fe512_mysql/fe512db
9 rows affected.

Out[4]:

Tables_in_hint
section_a
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [7]:

```
%%sql
LOAD DATA
  INFILE '/home/data/section_a.csv'
  INTO TABLE section_a
  FIELDS
    TERMINATED BY ','
  LINES
    TERMINATED BY '\r\n'
    STARTING BY ''
    IGNORE 1 LINES;
```

* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.

Out[7]: []

In [8]: `%sql SELECT * FROM section_a LIMIT 1;`

* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.

Out[8]:

PersonID	SeekHealthInfo	WhereSeekHealthInfo	WhoLookingFor	LotOfEffort	Frustrate
60000001-02	1	7	1	4	4

%%sql

```
LOAD DATA INFILE '/home/data/section_a.csv' INTO TABLE section_a FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY '"' ESCAPED BY '\\' LINES TERMINATED BY '\n' IGNORE 1 LINES;
```

Create Table in Database_Section_B

```
In [9]: %%sql
DROP TABLE IF EXISTS section_b;
CREATE TABLE IF NOT EXISTS section_b (
    PersonID VARCHAR(20),
    UseInternet INT,
    Internet_DialUp INT,
    Internet_BroadBnd INT,
    Internet_Cell INT,
    Internet_WiFi INT,
    WhereUseInternet_Home INT,
    WhereUseInternet_Work INT,
    WhereUseInternet_School INT,
    WhereUseInternet_PublicPlace INT,
    WhereUseInternet_MobileDevice INT,
    WhereUseInternet_GamingDevice INT,
    Electronic_SelfHealthInfo INT,
    Electronic_HealthInfoSE INT,
    Electronic_BuyMedicine INT,
    Electronic_HCPSearch INT,
    Electronic_TalkDoctor INT,
    Electronic_MadeAppts INT,
    Electronic_TrackedHealthCosts INT,
    Electronic_CompletedForms INT,
    Electronic_TestResults INT,
    HaveDevice_Tablet INT,
    HaveDevice_SmartPh INT,
    HaveDevice_CellPh INT,
    TabletHealthWellnessApps INT,
    Tablet_AchieveGoal INT,
    Tablet_MakeDecision INT,
    Tablet_DiscussionsHCP INT,
    OtherDevTrackHealth INT,
    SharedHealthDeviceInfo INT,
    IntRsn_VisitedSocNet INT,
    IntRsn_SharedSocNet INT,
    IntRsn_WroteBlog INT,
    IntRsn_SupportGroup INT,
    IntRsn_YouTube INT,
    TextFromDoctor INT
);
```

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.
```

```
/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_b'")
    result = self._query(query)
```

```
Out[9]: []
```

In [10]: `%sql SHOW TABLES;`

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
10 rows affected.
```

Out[10]:

Tables_in_hint
section_a
section_b
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [11]: `%%sql`

```
LOAD DATA
  INFILE '/home/data/section_b.csv'
  INTO TABLE section_b
  FIELDS
    TERMINATED BY ','
  OPTIONALLY ENCLOSED BY '"'
  ESCAPED BY '\\'
  LINES
    TERMINATED BY '\n'
  IGNORE 1 LINES;
```

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.
```

Out[11]: []

In [13]: `%sql SELECT * FROM section_b LIMIT 1`

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.
```

Out[13]:

PersonID	UseInternet	Internet_DialUp	Internet_BroadBnd	Internet_Cell	Internet_WiFi
60000001-02	1	2	2	2	1

Create Table in Database_Section_C

```
In [14]: %%sql
DROP TABLE IF EXISTS section_c;
CREATE TABLE IF NOT EXISTS section_c (
    PersonID VARCHAR(20),
    RegularProvider INT,
    MostRecentCheckup2 INT,
    FreqGoProvider INT,
    ChanceAskQuestions INT,
    FeelingsAddressed INT,
    InvolvedDecisions INT,
    UnderstoodNextSteps INT,
    ExplainedClearly INT,
    SpentEnoughTime INT,
    HelpUncertainty INT,
    QualityCare INT,
    ProbCare_BringTest INT,
    ProbCare_WaitLong INT,
    ProbCare_RedoTest INT,
    ProbCare_ProvideHist INT,
    HealthIns_InsuranceEmp INT,
    HealthIns_InsurancePriv INT,
    HealthIns_Medicare INT,
    HealthIns_Medicaid INT,
    HealthIns_Tricare INT,
    HealthIns_VA INT,
    HealthIns_IHS INT,
    HealthIns_Other INT,
    HealthIns_Other_OS VARCHAR(255)
);

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_c'")
  result = self._query(query)

Out[14]: []
```

In [35]: `%sql SHOW TABLES;`

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
13 rows affected.
```

Out[35]:

Tables_in_hint
section_a
section_b
section_c
section_d
section_e
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

Create Table in Database_Section_D

In [26]: %%sql

```

DROP TABLE IF EXISTS section_d; #db name
CREATE TABLE IF NOT EXISTS section_d(
    PersonID VARCHAR(20),
    ProviderMaintainEMR2 INT,
    OfferedAccessHCP2 INT,
    OfferedAccessInsurer2 INT,
    HCPEncourageOnlineRec INT,
    AccessOnlineRecord INT,
    NotAccessed_SpeakDirectly INT,
    NotAccessed_NoInternet INT,
    NotAccessed_NoNeed INT,
    NotAccessed_ConcernedPrivacy INT,
    NotAccessed_NoRecord INT,
    NotAccessed_Other INT,
    NotAccessed_Other_OS VARCHAR(255),
    RecordsOnline_Labs INT,
    RecordsOnline_Meds INT,
    RecordsOnline_HealthProbs INT,
    RecordsOnline_Allergies INT,
    RecordsOnline_VisitSummary INT,
    RecordsOnline_ClinNotes INT,
    RecordsOnline_Immunizations INT,
    RecordsOnline_MakeAppt INT,
    RecordsOnline_RefillMeds INT,
    RecordsOnline_Paperwork INT,
    RecordsOnline_RequestCorrection INT,
    RecordsOnline_MessageHCP INT,
    RecordsOnline_ViewResults INT,
    RecordsOnline_MonitorHealth INT,
    RecordsOnline_DownloadHealth INT,
    RecordsOnline_AddHealthInfo INT,
    RecordsOnline_MakeDecision INT,
    ESent_AnotherHCP INT,
    ESent_Family INT,
    ESent_HealthApp INT,
    UnderstandOnlineMedRec INT,
    UsefulOnlineMedRec INT,
    ConfidentInfoSafe INT,
    WithheldInfoPrivacy INT,
    ElectInfoSafe INT,
    AccessFamilyMedRec INT,
    AccessedFamRec_TheirPwd INT,
    AccessedFamRec_MyPwd INT
);

```

```

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

```

Out[26]: []

In [27]: **%sql** SHOW TABLES;

* mysql+pymysql://root:***@fe512_mysql/fe512db
12 rows affected.

Out[27]:

Tables_in_hint
section_a
section_b
section_c
section_d
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [28]: **%%sql**

```
LOAD DATA
  INFILE '/home/data/section_d.csv'
  INTO TABLE section_d
  FIELDS
    TERMINATED BY ','
    OPTIONALLY ENCLOSED BY '"'
    ESCAPED BY '\\'
  LINES
    TERMINATED BY '\n'
  IGNORE 1 LINES;
```

* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.

Out[28]: []

In [29]: **%sql** SELECT * FROM section_d LIMIT 1;

* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.

Out[29]:

PersonID	ProviderMaintainEMR2	OfferedAccessHCP2	OfferedAccessInsurer2	HCPEnc
60000001-02	1	1	3	1

```
In [ ]: %%sql
LOAD DATA
  INFILE '/home/data/section_d.csv'
  INTO TABLE section_d
  FIELDS
    TERMINATED BY ','
  LINES
    TERMINATED BY '\r\n'
    STARTING BY ''
  IGNORE 1 LINES;
```

Create Table in Database_Section_E

```
In [30]: %%sql
DROP TABLE IF EXISTS section_e;
CREATE TABLE IF NOT EXISTS section_e (
  PersonID VARCHAR(20),
  Caregiving_Child INT,
  Caregiving_Spouse INT,
  Caregiving_Parent INT,
  Caregiving_Family INT,
  Caregiving_Friend INT,
  Caregiving_No INT,
  CaregivingWho_Cat INT,
  Caregiving_Cancer INT,
  Caregiving_Alzheimers INT,
  Caregiving_OrthoMusc INT,
  Caregiving_MentalHealth INT,
  Caregiving_ChronicCond INT,
  Caregiving_NeuroDev INT,
  Caregiving_AcuteCond INT,
  Caregiving_Aging INT,
  Caregiving_NotSure INT,
  Caregiving_Other INT,
  Caregiving_Other_OS VARCHAR(255),
  CaregivingCond_Cat INT,
  Caregiving_HoursPerWeek INT
);

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_e'")
  result = self._query(query)
```

Out[30]: []

In [31]: **%sql** SHOW TABLES;

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
13 rows affected.
```

Out[31]:

Tables_in_hint
section_a
section_b
section_c
section_d
section_e
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [32]: **%%sql**

```
LOAD DATA
  INFILE '/home/data/section_e.csv'
  INTO TABLE section_e
  FIELDS
    TERMINATED BY ','
    OPTIONALLY ENCLOSED BY '"'
    ESCAPED BY '\\'
  LINES
    TERMINATED BY '\n'
  IGNORE 1 LINES;
```

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.
```

Out[32]: []

In [33]: `%sql SELECT * FROM section_e LIMIT 1;`

* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.

Out[33]:

PersonID	Caregiving_Child	Caregiving_Spouse	Caregiving_Parent	Caregiving_Family
60000001-02	2	2	2	2

Create Table in Database_Section_F

In [36]: `%%sql`

```
DROP TABLE IF EXISTS section_f;
CREATE TABLE IF NOT EXISTS section_f (
    PersonID VARCHAR(20),
    HeardDNATest INT,
    GeneticTestUse_DetermineRisk INT,
    GeneticTestUse_DetermineTx INT,
    GeneticTestUse_DetermineMed INT,
    GeneticTestUse_DeterminePass INT,
    GeneticTestUse_Cat INT,
    HadTest_Paternity INT,
    HadTest_Ancestry INT,
    HadTest_DNAFing INT,
    HadTest_CFCarrier INT,
    HadTest_BRCA INT,
    HadTest_Lynch INT,
    HadTest_None INT,
    HadTest_NotSure INT,
    HadTest_Other INT,
    HadTest_Other_OS VARCHAR(255),
    HadTest_Cat INT
);
```

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_f'")
result = self._query(query)

Out[36]: []

In [37]: `%sql SHOW TABLES;`

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
14 rows affected.
```

Out[37]:

Tables_in_hint
section_a
section_b
section_c
section_d
section_e
section_f
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [38]: `%%sql`

```
LOAD DATA
  INFILE '/home/data/section_f.csv'
  INTO TABLE section_f
  FIELDS
    TERMINATED BY ','
    OPTIONALLY ENCLOSED BY '"'
    ESCAPED BY '\\'
  LINES
    TERMINATED BY '\n'
  IGNORE 1 LINES;
```

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.
```

Out[38]: []

In [39]: `%sql SELECT * FROM section_f LIMIT 1;`

* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.

Out[39]:

PersonID	HeardDNATest	GeneticTestUse_DetermineRisk	GeneticTestUse_DetermineTx
60000001-02	1	1	2

Create Table section_g

In [40]: `%%sql
DROP TABLE IF EXISTS section_g;
CREATE TABLE IF NOT EXISTS section_g (
 PersonID VARCHAR(20),
 GeneralHealth INT,
 OwnAbilityTakeCareHealth INT,
 MedConditions_Diabetes INT,
 MedConditions_HighBP INT,
 MedConditions_HeartCondition INT,
 MedConditions_LungDisease INT,
 MedConditions_Arthritis INT,
 MedConditions_Depression INT,
 Height_Feet INT,
 Height_Inches INT,
 Weight INT,
 LittleInterest INT,
 Hopeless INT,
 Nervous INT,
 Worrying INT,
 EmotionalSupport INT,
 TalkHealthFriends INT,
 HelpDailyChores INT,
 Deaf INT
);`

* mysql+pymysql://root:***@fe512_mysql/fe512db
0 rows affected.
0 rows affected.

/opt/conda/lib/python3.6/site-packages/pymysql/cursors.py:170: Warning:
(1051, "Unknown table 'hint.section_g'")
result = self._query(query)

Out[40]: []

In [41]: `%sql SHOW TABLES;`

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
15 rows affected.
```

Out[41]:

Tables_in_hint
section_a
section_b
section_c
section_d
section_e
section_f
section_g
section_h
section_i
section_j
section_k
section_l
section_m
section_n
section_o

In [42]: `%%sql`

```
LOAD DATA
  INFILE '/home/data/section_g.csv'
  INTO TABLE section_g
  FIELDS
    TERMINATED BY ','
    OPTIONALLY ENCLOSED BY '"'
    ESCAPED BY '\\'
  LINES
    TERMINATED BY '\n'
  IGNORE 1 LINES;
```

```
* mysql+pymysql://root:***@fe512_mysql/fe512db
3285 rows affected.
```

Out[42]: []

In [43]:

%sql SELECT * from section_g LIMIT 1;

* mysql+pymysql://root:***@fe512_mysql/fe512db
1 rows affected.

Out[43]:

PersonID	GeneralHealth	OwnAbilityTakeCareHealth	MedConditions_Diabetes	MedCon
60000001-02	2	2	2	1