These are the main hive queries used. This is not an exhaustive list, but represents some of the important ones we used.

TABLES CREATION:

```
CREATE EXTERNAL TABLE IF NOT EXISTS LICENSES(ID
STRING,LICENSE_ID STRING, LEGAL_NAME STRING,ADDRESS STRING,
CITY STRING, ZIP_CODE INT,WARD STRING,PRECINCT
STRING,POLICE_DISTRICT STRING ,LICENSE_CODE STRING,BUSINESS
ACTIVITY ID STRING, LICENSE_TERM_START_YEAR
INT,LICENSE_TERM_END_YEAR INT,LATITUDE DECIMAL(10,6),LONGITUDE
DECIMAL(10,6))
COMMENT 'DATA ABOUT BUSINESS LICENSES'
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION '/USER/AG7982_NYU_EDU/PROJECTHIVE';


CREATE EXTERNAL TABLE IF NOT EXISTS DEMOGRAPHICS(CA
INT,CA_NAME STRING, PERCENT_OF_HOUSING_CROWDED
DECIMAL(6,6),PERCENT_HOUSEHOLD_POVERTY DECIMAL(6,6),
PERCENT_16_UNEMPLOYED DECIMAL(6,6), PERCENT_25_HSDIPLOMA
DECIMAL(6,6), PERCENT_AGED_18_64 DECIMAL(6,6),INCOME
DECIMAL(6,6) ,HARDSHIP_INDEX DECIMAL(6,6))
COMMENT 'DATA ABOUT DEMOGRAPHICS'
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION '/USER/AG7982_NYU_EDU/PROJECTHIVE/DEMOGRAPHICS';


CREATE EXTERNAL TABLE CRIMES DATA TEST (ID STRING, PRIMARY
IVPE STRING, DESCRIPTION STRING,
LOCATION DESCRIPTION STRING, ARREST STRING, DOMESTIC
STRING,BEAT STRING,DISTRICT STRING,
WARD STRING, COMMUNITY AREA INT, YEAR INT, LATITUDE DECIMAL (10,
6), LONGITUDE DECIMAL (10,6))
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION '/USER/AG7982_NYU_EDU/PROJECTHIVE/CRIMES';
```

CREATING TABLE TO GET COUNT OF CRIME INCIDENTS BY AREA:

    CREATE TABLE CRIME_COUNT_BY_CA AS SELECT COMMUNITY_AREA,
    COUNT(*) COUNT FROM CRIMES_DATA GROUP BY COMMUNITY_AREA;

CREATING TABLE FOR COMMUNITY AREAS WITH TOP 30 CRIME COUNTS:

    CREATE TABLE TOP_20_COUNT_DEMO_CRIME SELECT * FROM
    CRIME_COUNT_BY_CA WHERE COMMUNITY_AREA IS NOT NULL ORDER
    BY COUNT DESC LIMIT 30;

UNION BETWEEN TOP 30 AREAS BY CRIME AND TOP 30 AREAS BY HARDSHIP INDEX:

    SELECT CA_NAME FROM TOP_30_COUNT_DEMO_CRIME WHERE EXISTS
    (SELECT CA_NAME FROM TOP_30_COUNT_HARDSHIPINDEX WHERE
    TOP_30_COUNT_HARDSHIPINDEX.CA_NAME =
    TOP_30_COUNT_DEMO_CRIME.CA_NAME);

THE FOLLOWING QUERIES WERE USED FOR EACH UNIQUE ENTRY IN PRIMARY_TYPE:

    CREATE TABLE TRAFFICKING AS SELECT COMMUNITY_AREA,
    PRIMARY_TYPE FROM CRIME_TEST_08_12 WHERE
    PRIMARY_TYPE="HUMAN TRAFFICKING";

    CREATE TABLE TRAFFICKING_BY_CA AS SELECT COMMUNITY_AREA,
    COUNT(*) COUNT FROM TRAFFICKING GROUP BY COMMUNITY_AREA;

    CREATE TABLE DEMO_TRAFFICKING_COMBINED AS SELECT * FROM
    DEMOGRAPHICS_DATA JOIN TRAFFICKING_BY_CA ON
    COMMUNITY_AREA=CA;

    CREATE TABLE CORR_TRAFFICKING AS
    SELECT CORR(COUNT, PERCENT_OF_HOUSING_CROWDED)
    CORR_COUNT_PERCENT_OF_HOUSING_CROWDED, CORR(COUNT,
    PERCENT_HOUSEHOLD_POVERTY)

```
CORR_COUNT_PERCENT_HOUSEHOLD_POVERTY, CORR(COUNT,
PERCENT_16_UNEMPLOYED)
CORR_COUNT_PERCENT_16_UNEMPLOYED, CORR(COUNT,
PERCENT_25_HSDIPLOMA) CORR_COUNT_PERCENT_25_HSDIPLOMA ,
CORR(COUNT,PERCENT_AGED_18_64)
CORR_COUNT_PERCENT_AGED_18_64 , CORR(COUNT,INCOME )
CORR_COUNT_INCOME , CORR(COUNT, HARDSHIP_INDEX)
CORR_COUNT_HARDSHIP_INDEX
FROM DEMO_TRAFFICKING_COMBINED;

SELECT * FROM CORR_TRAFFICKING;
```