## CUNY MSDA - IS607 : DATA ACQUISITION AND MANAGEMENT

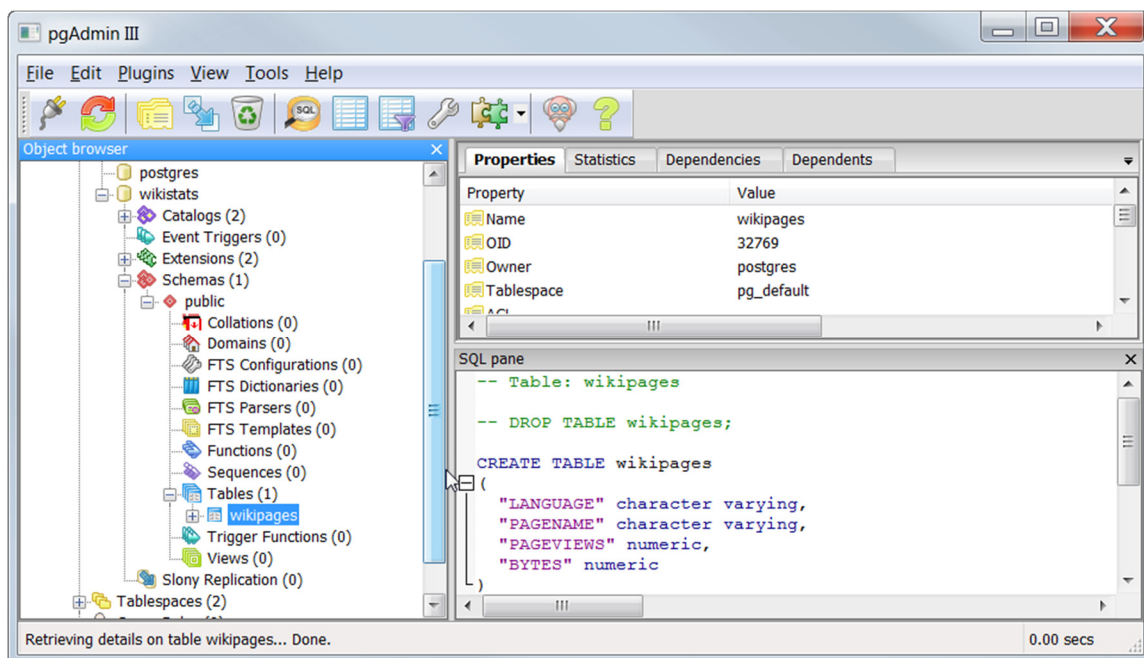## Week 7 Project

## A. Leopold HILLAH

**##### Creation of  a new database #####**

Creation of **wikistats** database.

```sql
CREATE DATABASE wikistats
  WITH ENCODING='UTF8'
       OWNER=postgres
       TEMPLATE=postgres
       LC_COLLATE='English_United Kingdom.1252'
       LC_CTYPE='English_United Kingdom.1252'
       CONNECTION LIMIT=-1
       TABLESPACE=pg_default;

COMMENT ON DATABASE wikistats
  IS 'Wikipedia Traffic Statistics';
```

```sql
CREATE TABLE wikipages
(
  "LANGUAGE" character varying,
  "PAGENAME" character varying,
  "PAGEVIEWS" numeric,
  "BYTES" numeric
)
WITH (
  OIDS=FALSE
);
ALTER TABLE wikipages
  OWNER TO postgres;
COMMENT ON TABLE wikipages
  IS 'Pageview statistics';
```

**##### Loading Data File #####**

The relevant SQL statement is:

**COPY WIKIPAGES FROM 'F:\R-Stats\20141001140000.txt' (DELIMITER(' '));**

**##### The five Wikipedia pages that were visited most often #####**

The relevant SQL statement is:

```
SELECT "PAGENAME", SUM("PAGEVIEWS") FROM WIKIPAGES
GROUP BY "PAGENAME"
ORDER BY 2 DESC
LIMIT 5;
```

**Results:**

| PAGENAME | sum |
|---|---|
| Database | 257 |
| Data:image/png base64,iVBORw0KGgoAAAANSUhEUgAAAAEAAAAu | 220 |
| Data | 203 |
| Data_mining | 196 |
| Data_Protection_Act_1998 | 174 |