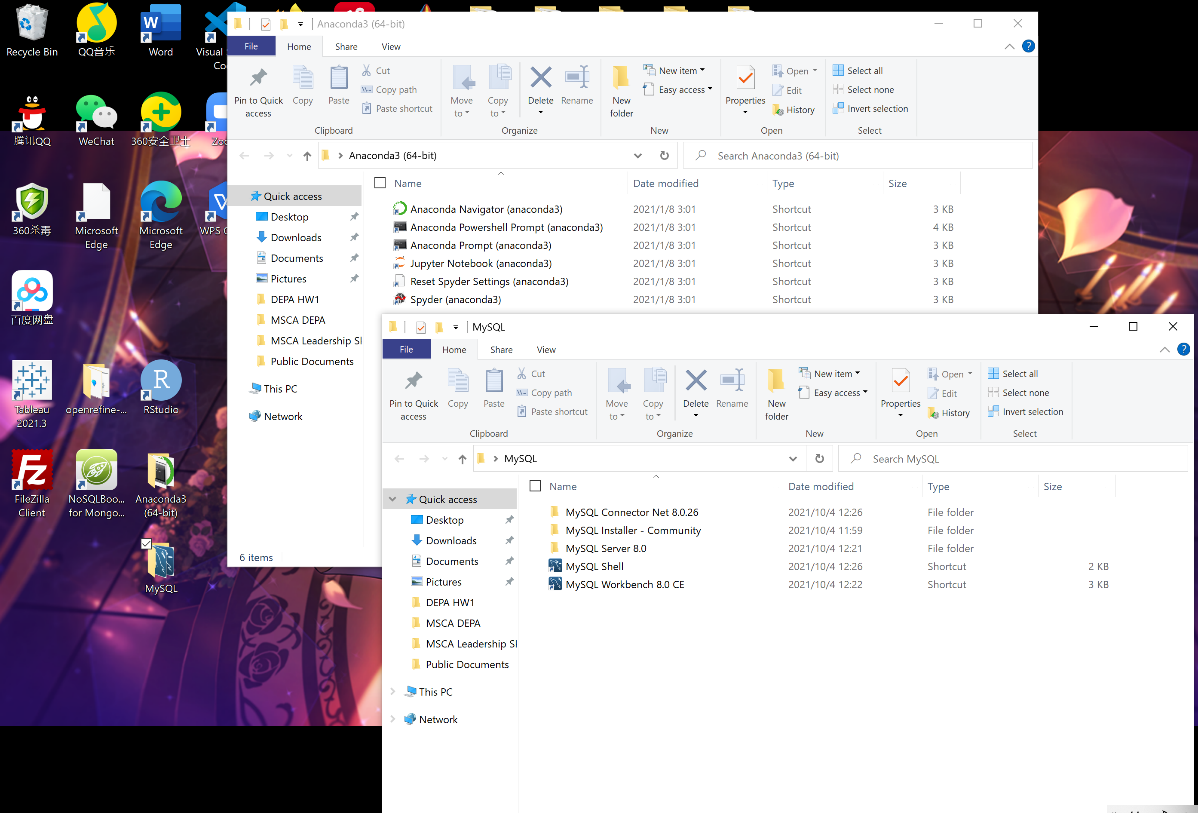
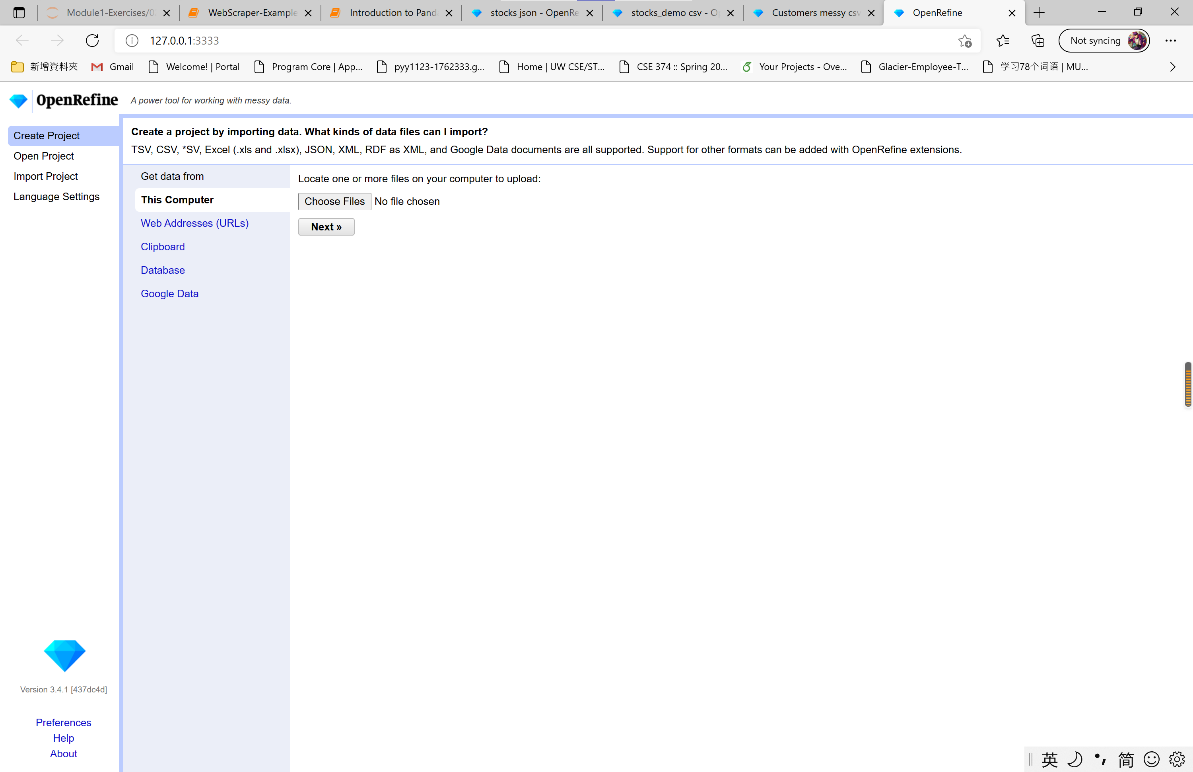
**Assignment 1:**

**Part A:**

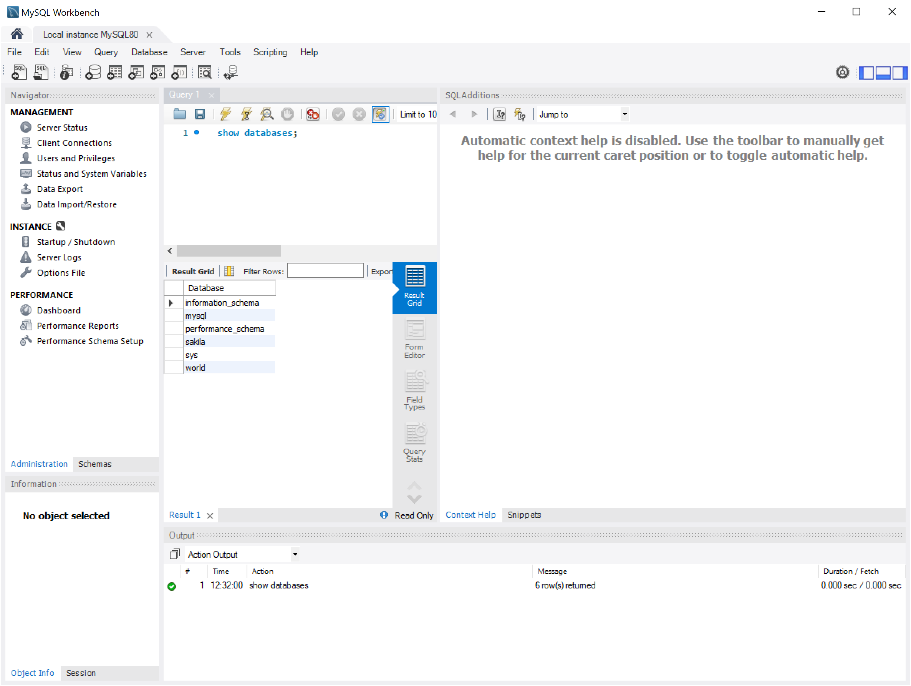
1. **Installation**

****

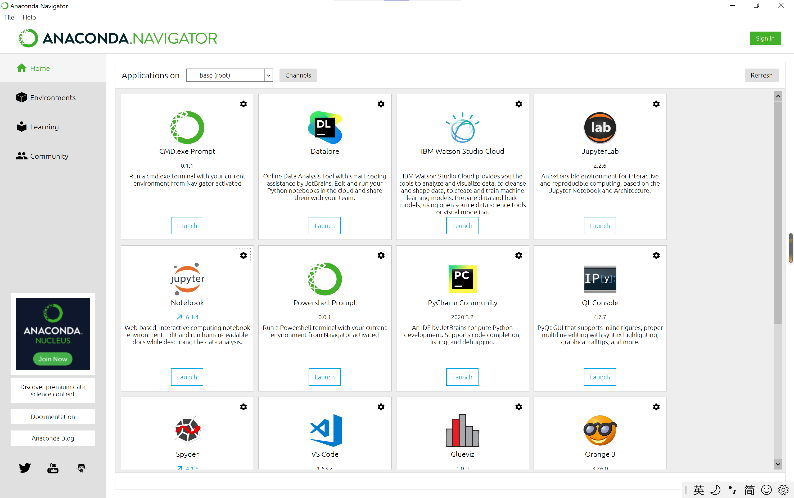
* 1. **OpenRefine**

****

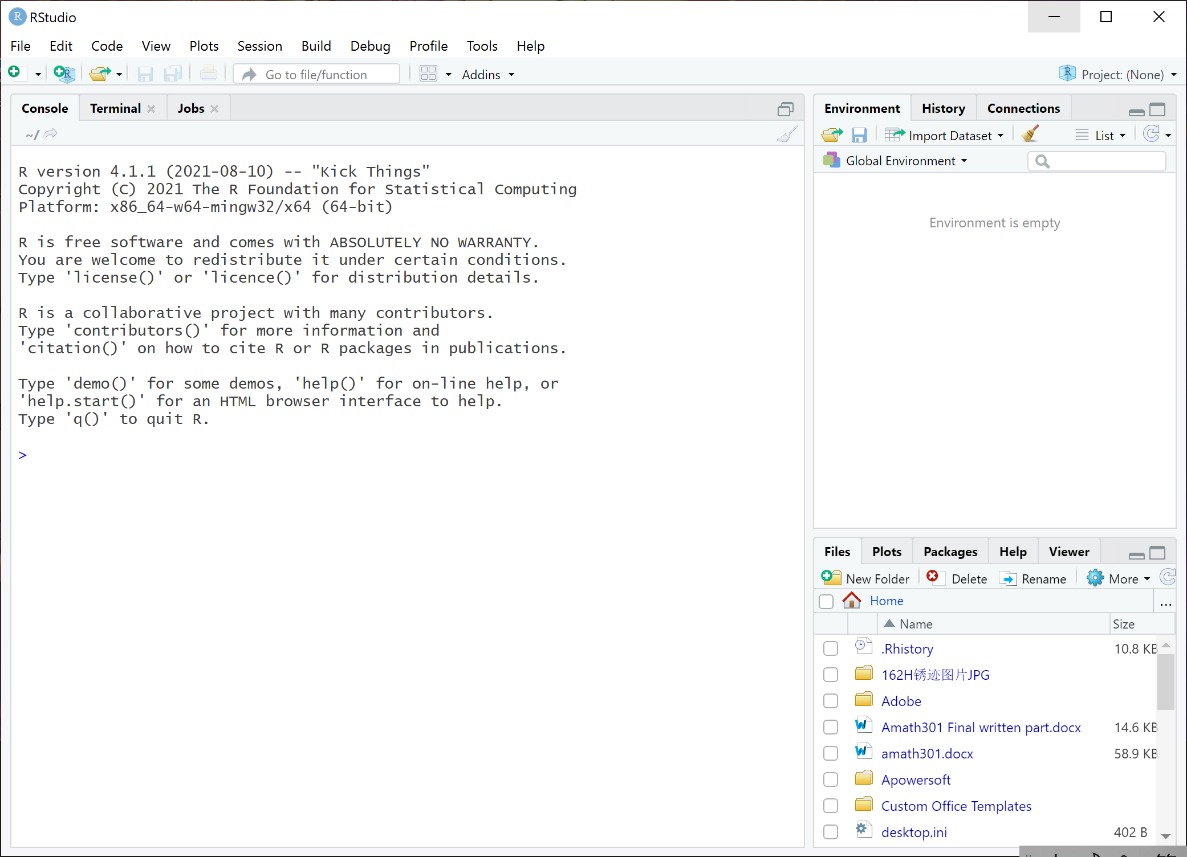
* 1. **MySql**

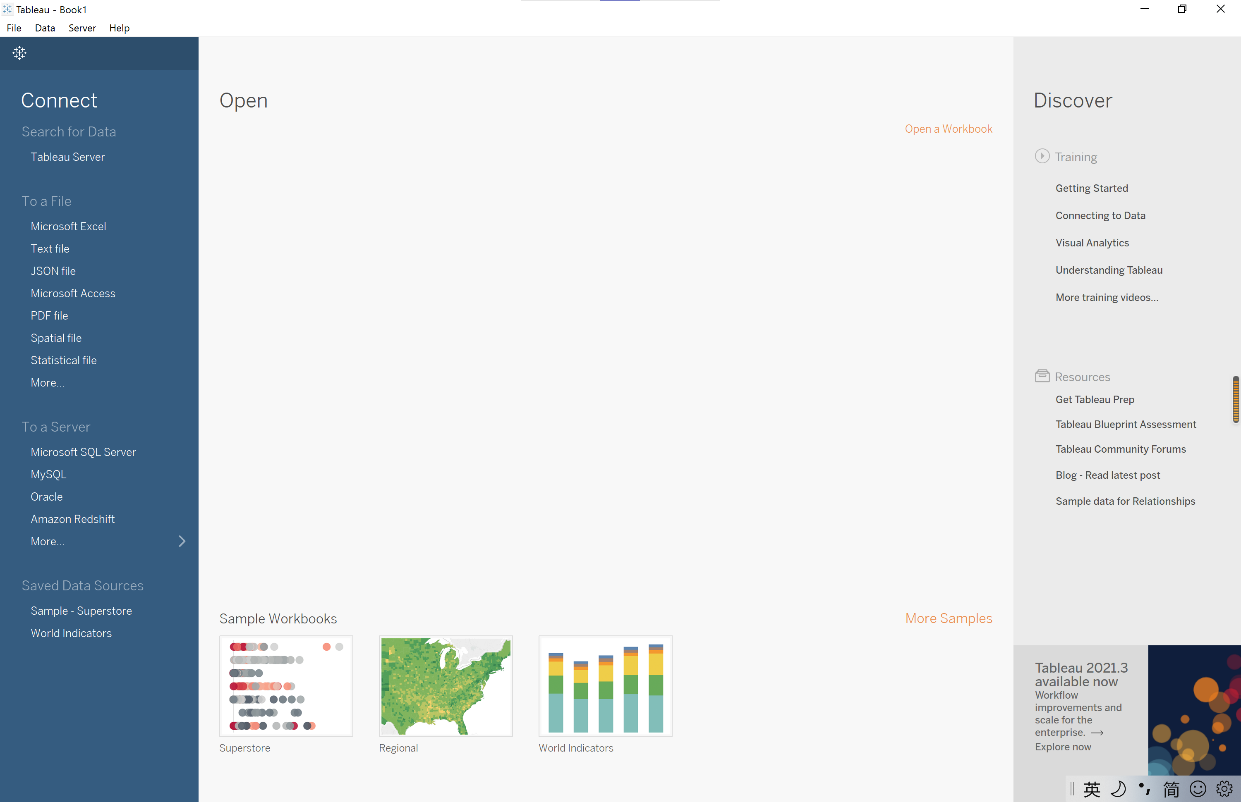
****

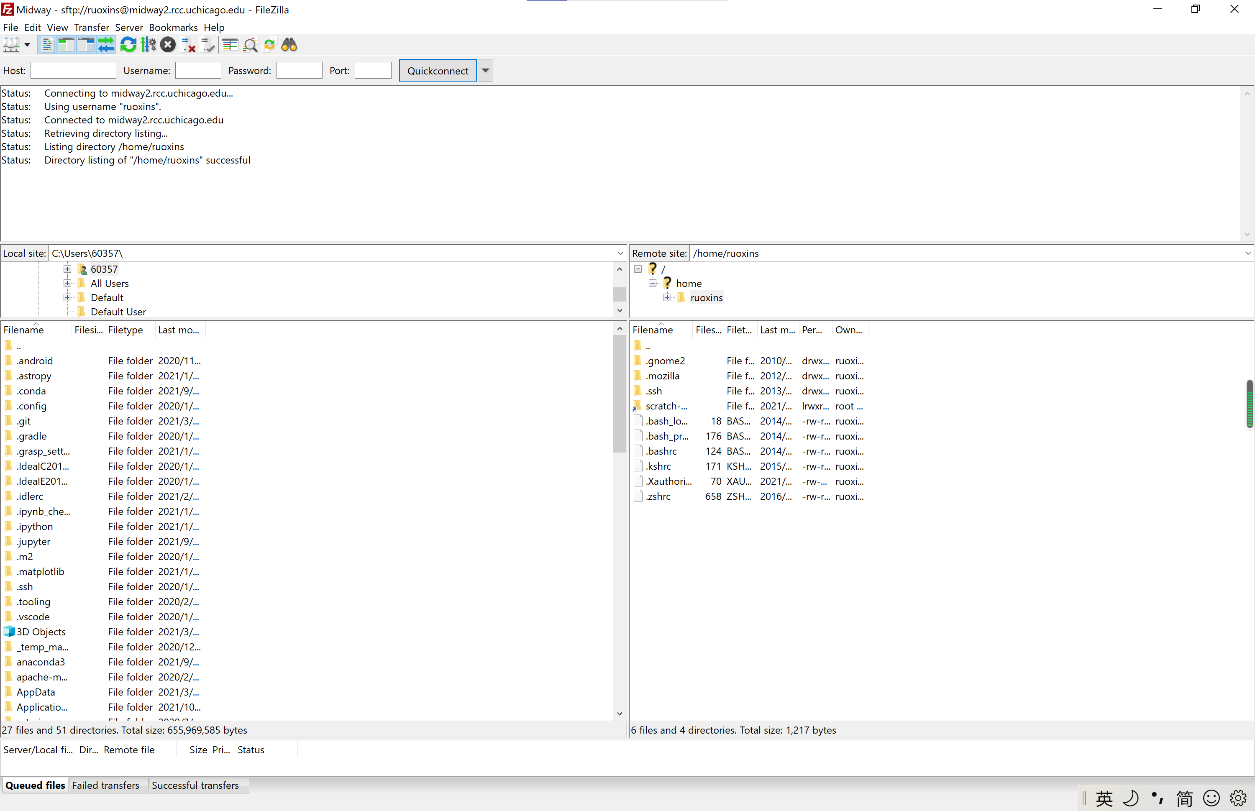
* 1. **Anaconda**

****

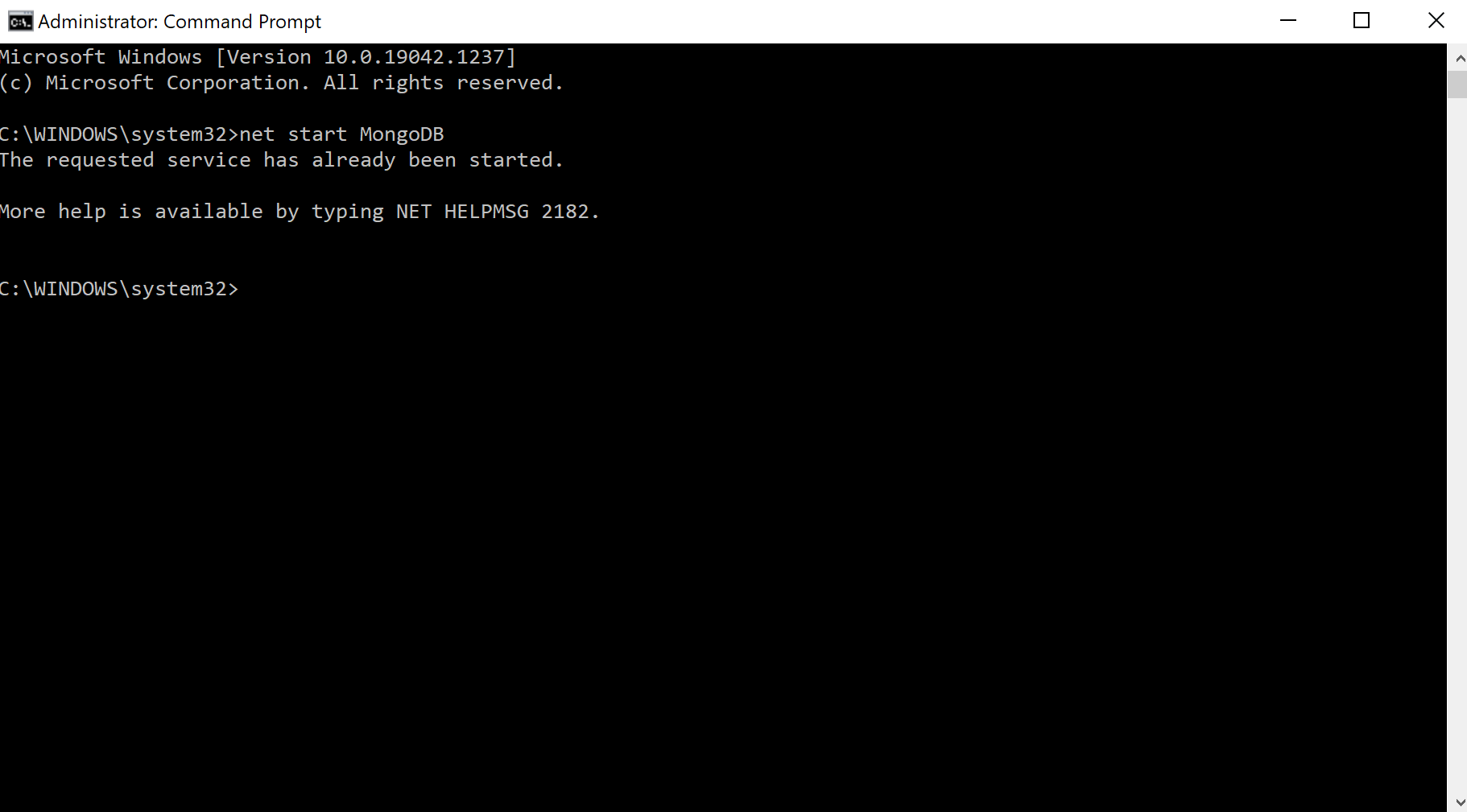
* 1. **R-studio**

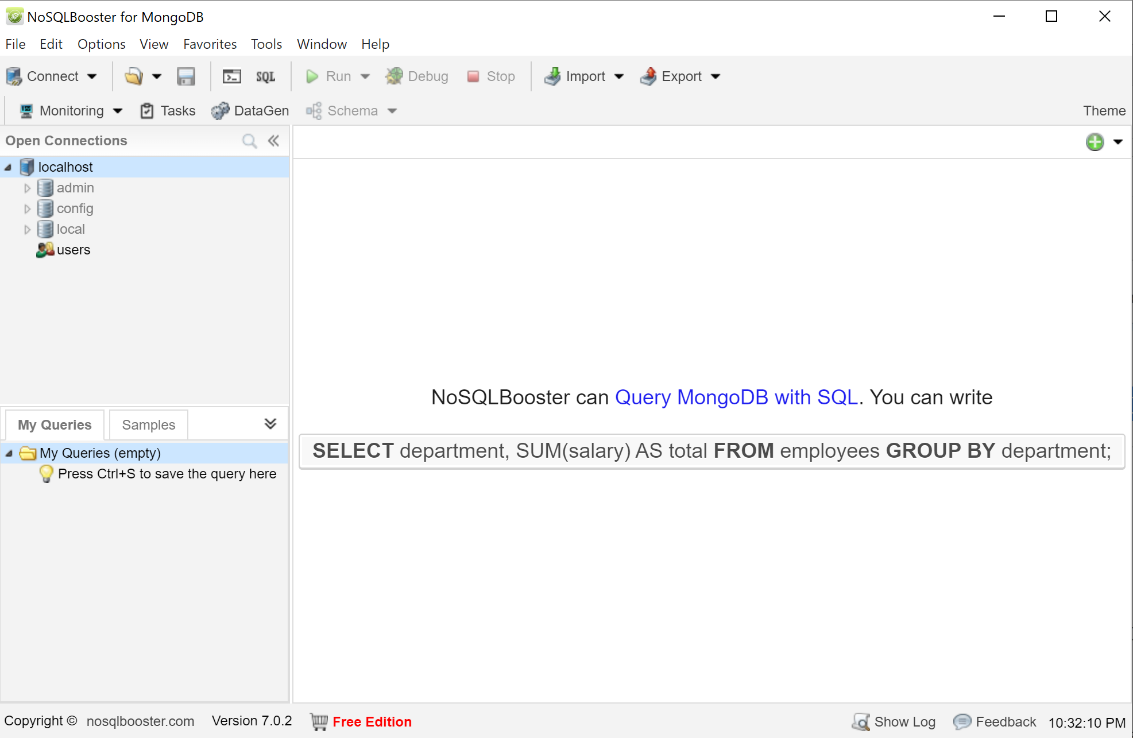
****

* 1. **Tableau**
  2. **FileZilla**

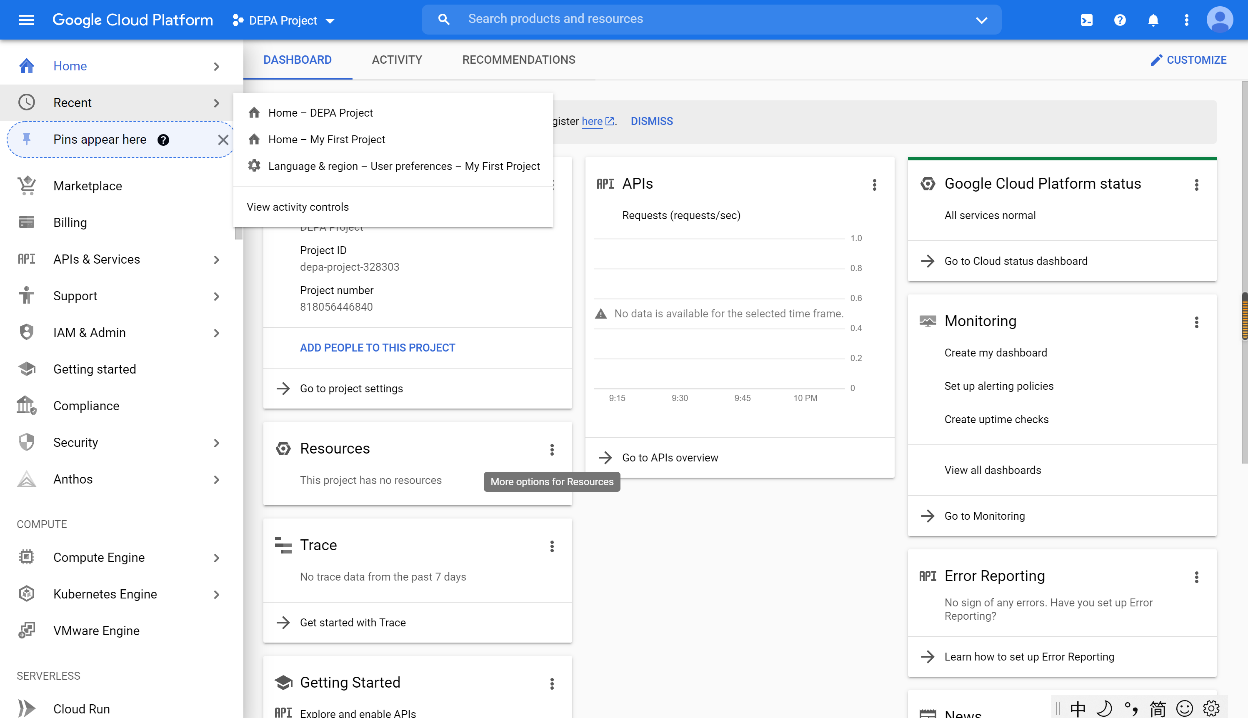
****

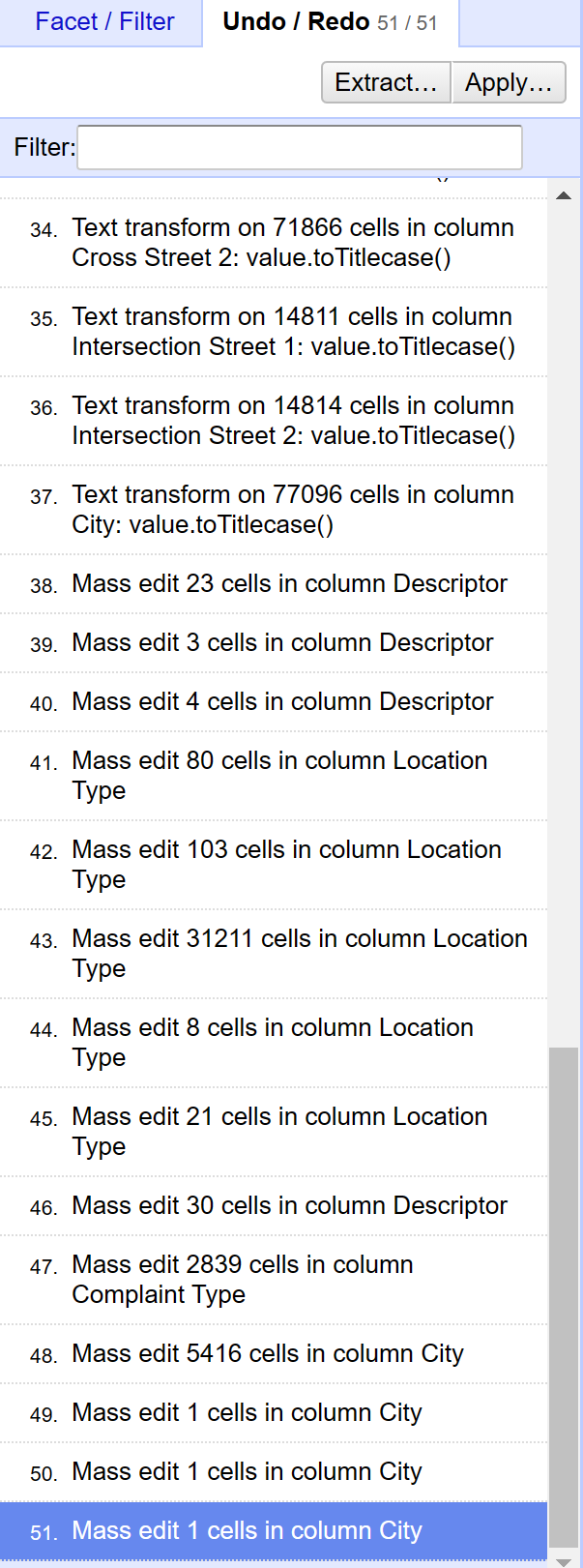
* 1. **MongoDB**

****

****

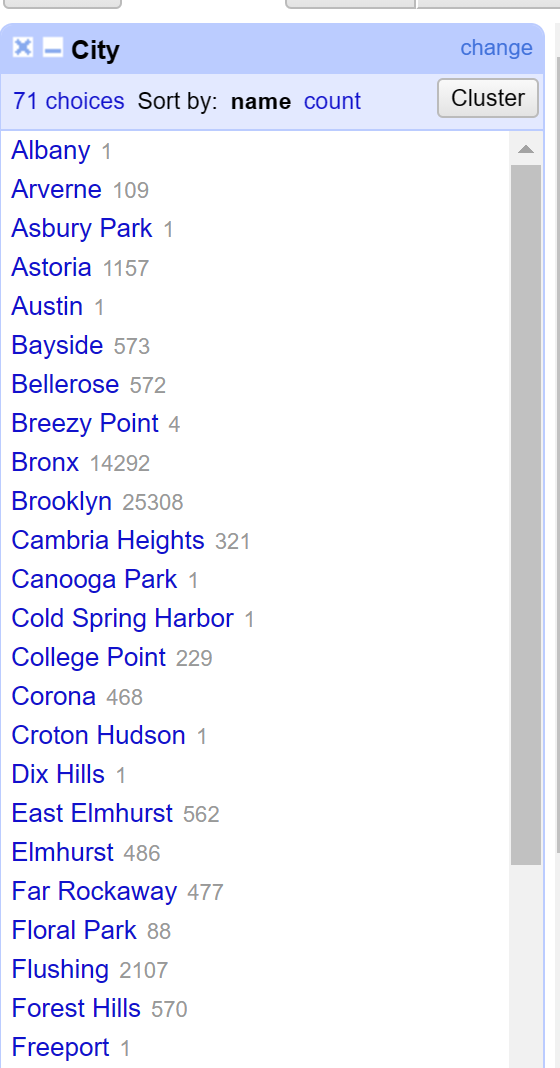
* 1. **GCPS**

****

1. ****

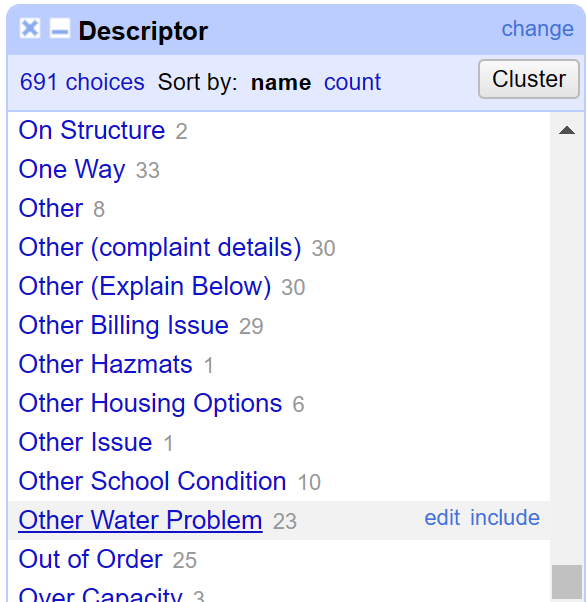
**Part d – “City” cleanup**

* 1. Convert City to title case, then Cluster and Merge the column

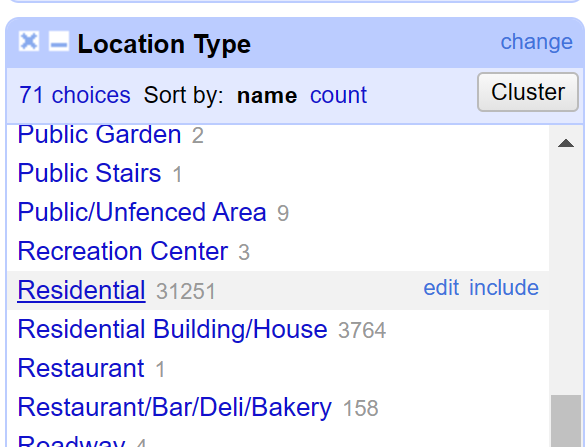
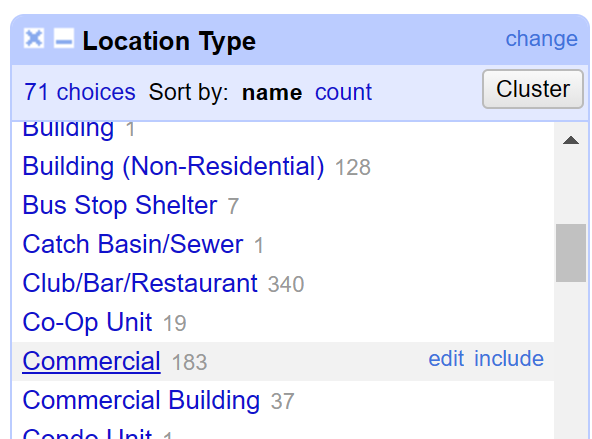
****

**Part e – “Descriptor” cleanup**

* 1. 1. "Other Water problem(WZZ)", "Other Water problem(QZZ)" as "Other Water Problem"
  2. 2. "Commercial 421 A/B Exemptions" as "Commercial Exemption"
  3. 3. "Commercial Exemption" "Commercial Other Exemption" as "Commercial Exemption"

****

****

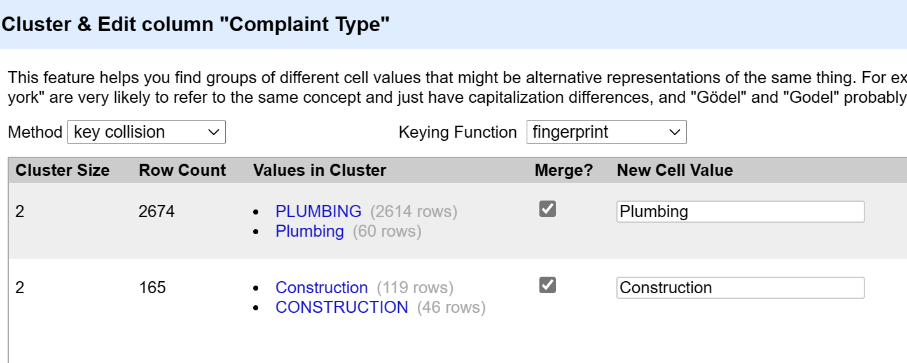
****

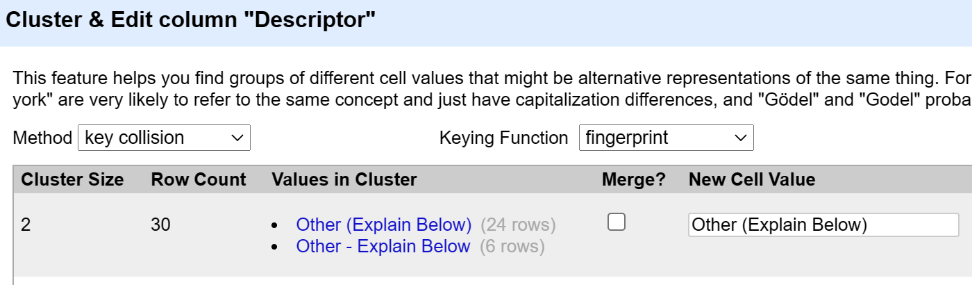
**Part f – “Location Type” cleanup**

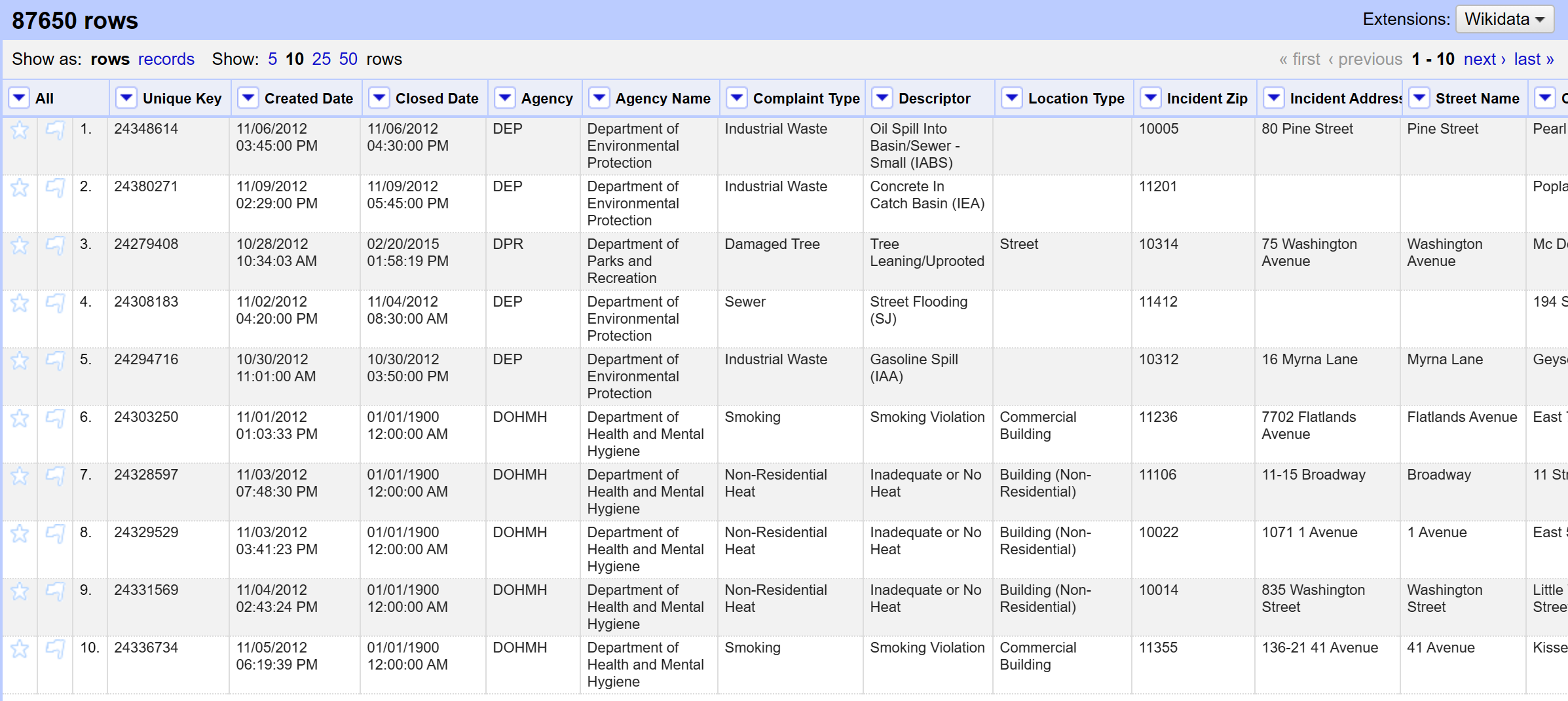
* 1. 1. "Comercial", "Commercial", "Store/Commercial" as "Commercial"
  2. 2. "RESIDENTIAL BUILDING", "Residential Building", "Residence" as "Residential"
  3. 3. "Street/Sidewalk", "Street and Sidewalk" as "Street/Sidewalk"

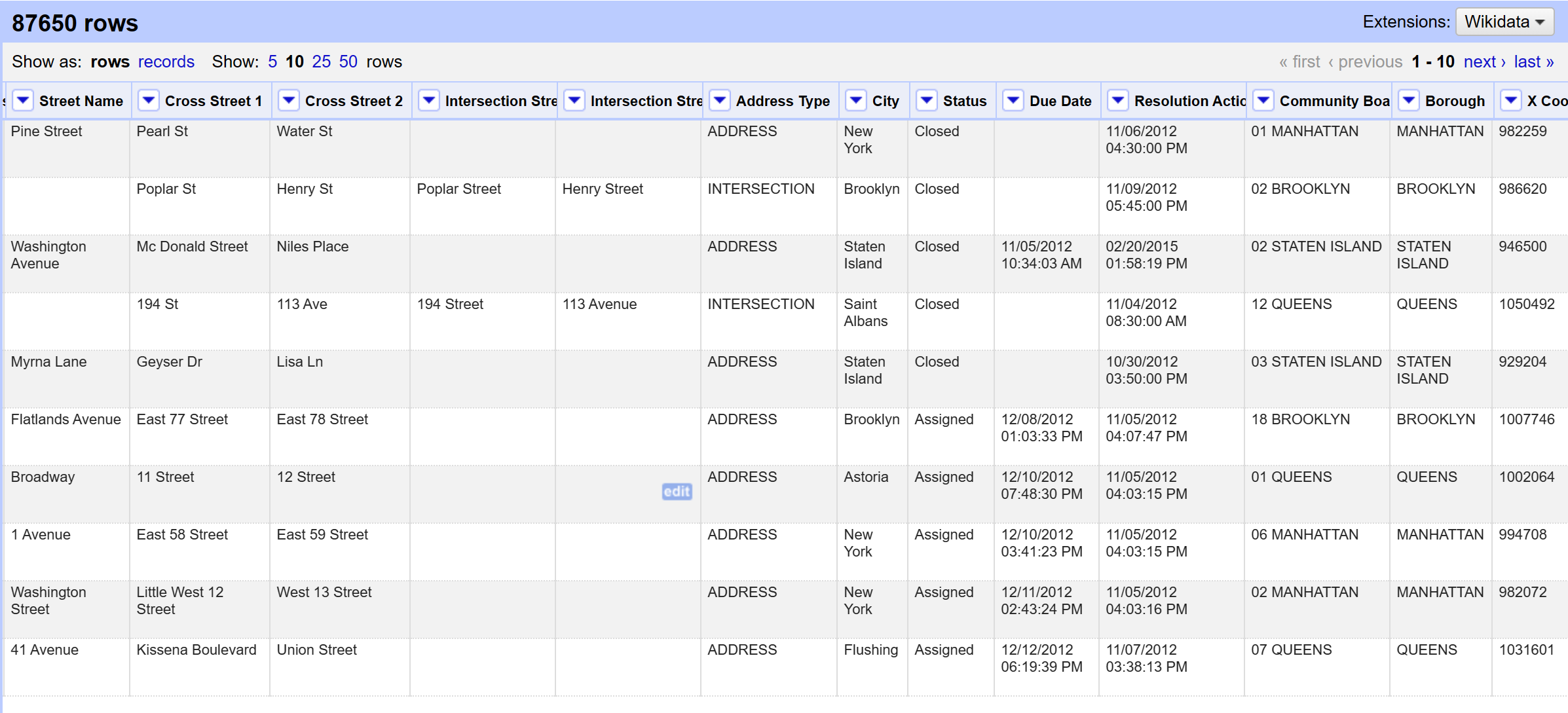
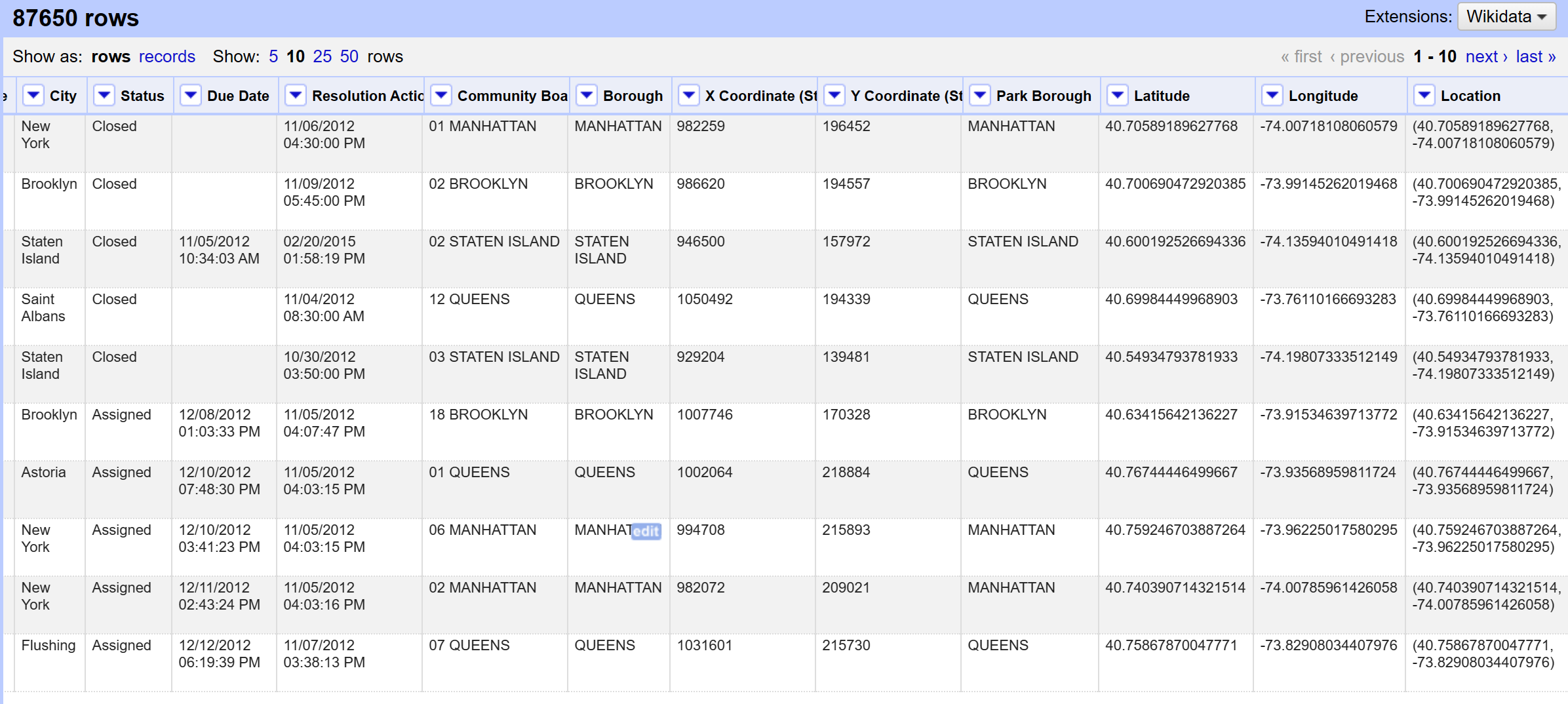
****

**Part g – 2 extra cleanup**

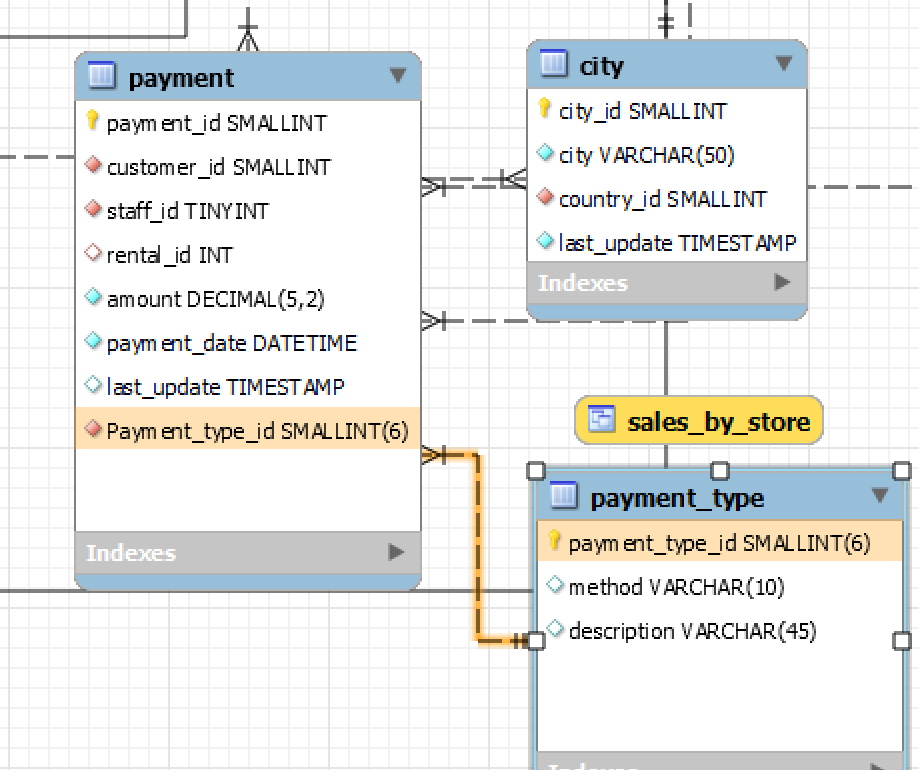
****

****

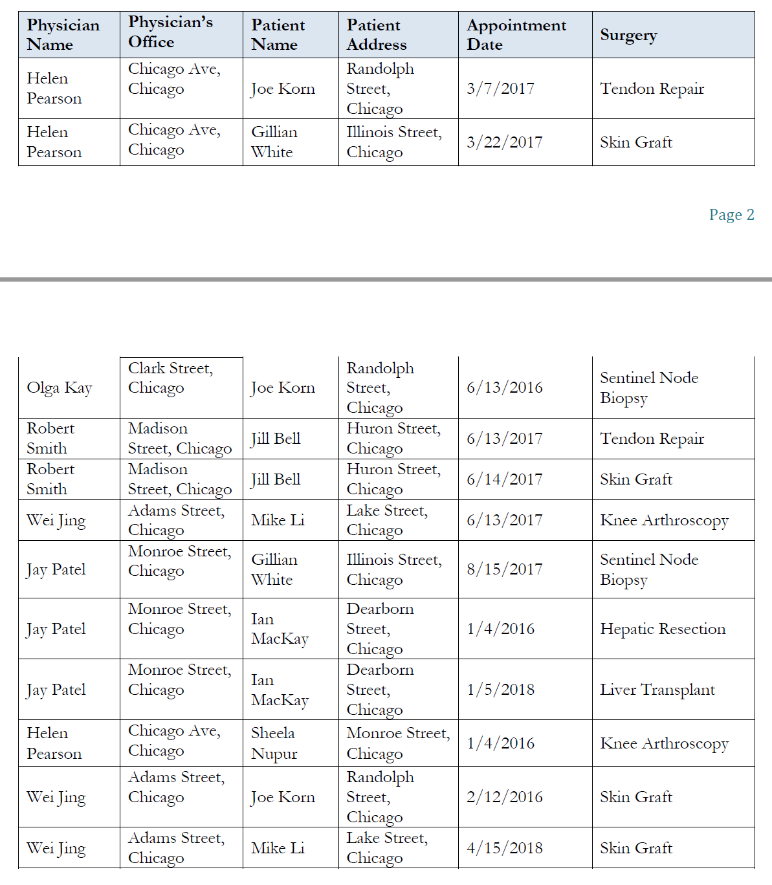
****

**** ****

**Part B:**

1. ****

|  |  |  |  |
| --- | --- | --- | --- |
| **Field**  **(Attributes)** | **Primary Key**  **(Y/N)** | **Foreign Key**  **(Y/N)** | **Related Tables(s)**  **(only enter this for foreign key fields) & Type of relationship between tables** |
| payment\_id | Y | N | / |
| customer\_id | N | Y | Related tables: customer; Relationship: 1 to n  1 customer to many payments |
| staff\_id | N | Y | Related tables: staff; Relationship: 1 to n  1 staff to many payments |
| rental\_id | N | Y | Related tables: rental; Relationship: 1 to n  1 rental to many payments |
| Payment\_type\_id | N | Y | Related tables: payment\_type; Relationship: 1 to n  1 payment type to many payments |
| amount | N | N | / |
| payment\_date | N | N | / |
| last\_update | N | N | / |

1. **Normalization**
2. **Provide examples of insertion, deletion, and modification anomalies.**

**insertion anomaly:**

**Cannot insert new physician name without knowing other needed information of the physician listed in the table. We cannot simply add a physician without knowing the patient and surgery information.**

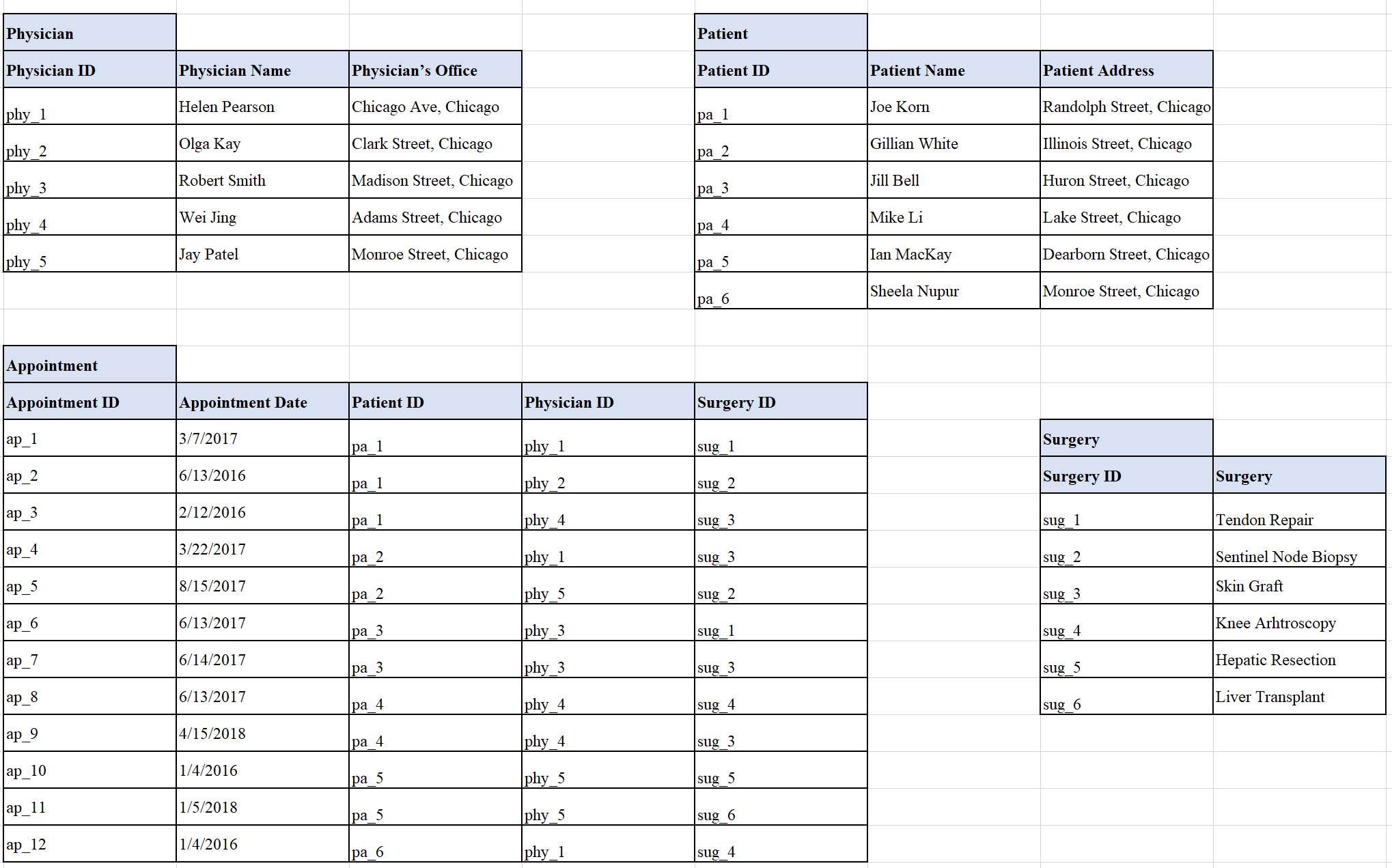
**deletion anomaly:**

**We cannot delete the Patient “Joe Korn” since we would also lose all the information for the Physician “Olga Kay”.**

**modification anomaly:**

**To change the address of Patient “Jill Bell”, 2 records have to be modified.**

1. **Normalize this table to 3NF and list any assumptions.**

**Assumptions: Since the table only shows a very small part of the data, we don’t have many information. We could assume that for each appointment of surgery, there are only 1 patient and 1 physician.**

**Normalized into 4 tables: Physician, Patient, Appointment and Surgery.**

**Each physician, patient, appointment and surgery have their unique id.**