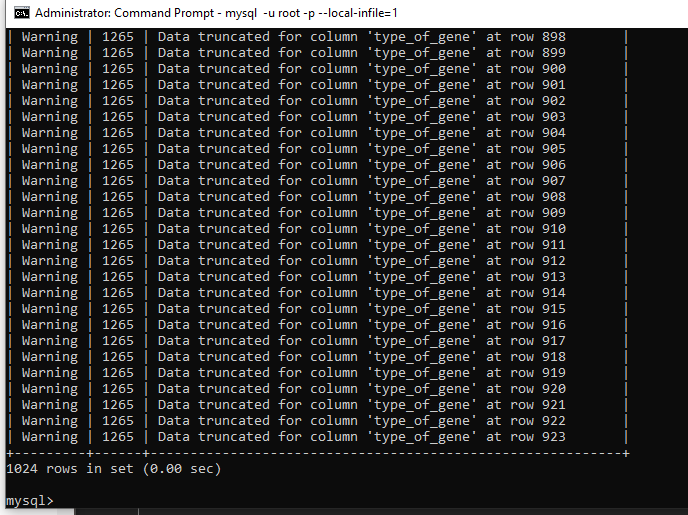
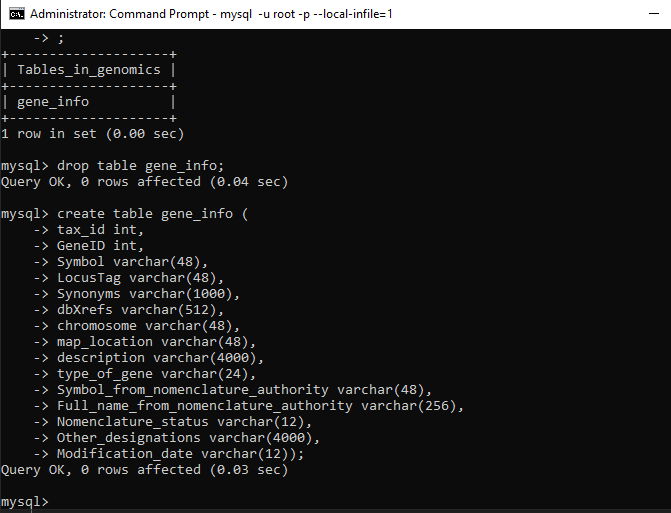
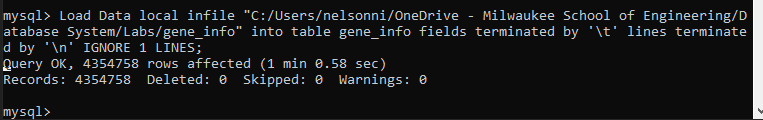
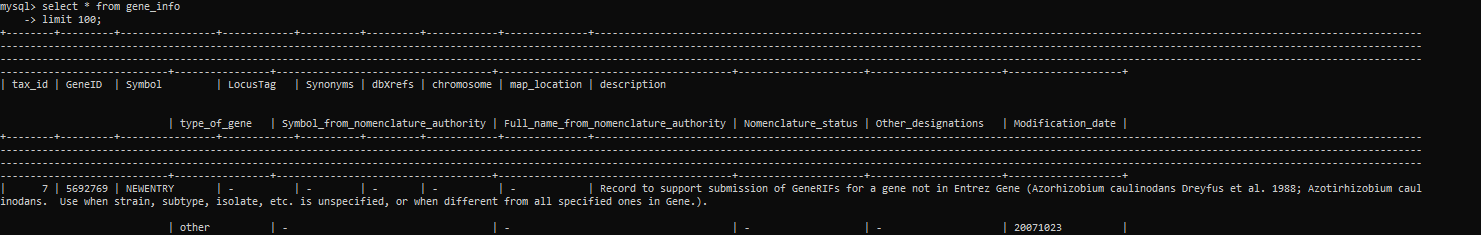
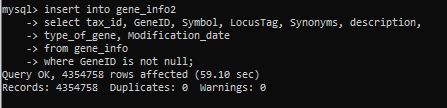
Initial Observations of data in excel

* There are 1000 entries
* There seem to be 4 columns that are empty
* There are very few column headers, making the data hard to read
* All but 2 rows seem to be alphanumeric identifiers

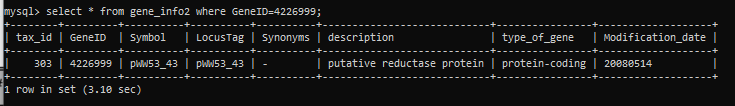
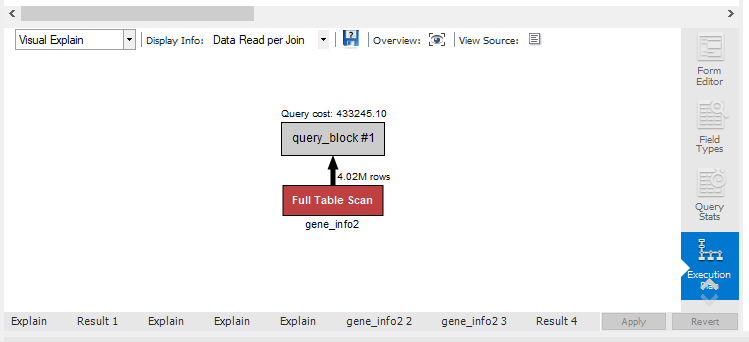
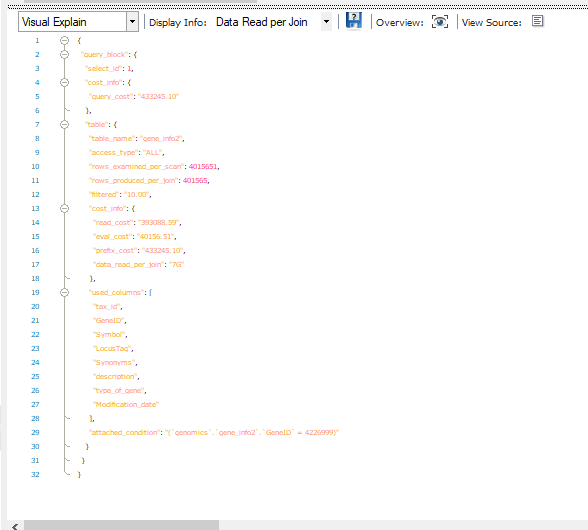
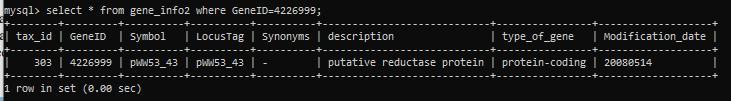
Create a database table

* The SQl seems to be a formatter for geneinfo\_1000, where each entry separated as a common acts as a header for each column of geneinfo\_1000. Also it appears that each “header” has the number of entries it expects to have in its column.

Load Data local Infile "C:/Users/nelsonni/OneDrive - Milwaukee School of Engineering/Database System/Labs/gene\_info" into table genomics.gene\_info fields terminated by '\t' lines terminated by '\n' IGNORE 1 LINES;

* 
* 
* 
* 
* Creating work table
* 
* 

Primary keys and indexes

* 
* 
* 
* Updating primary key
* 
* 
* Create index
* 
* Import gene2pubmed
* 