Preslav Angelov, Huy Le, Evin Tolentino

Professor Ben Kim

BUAN 5510

8 July 2019

Group Assignment 1

1. Provide the description of your database, including the overview, sizes of the files, number of tables, attributes of each table, number of rows for each table, relationships among the tables, and others.

Our database contains property value and tax information for Pierce County in Washington State. This information is collected, maintained, and published by the Assessor-Treasurer's Office. Our database contains 9 table, Appraisal Account, Improvement, Improvement Builtas, Improvement Detail, Land Attribute, Sale, Seg Merge, Tax Account, and Tax Description.

The Appraisal Account table is 5.8 MB in size, has 24 columns (attributes), has 331342 rows, has a one-to-one relationship with tax account, a one-to-many relationship with Improvement, and a one-to-many relationship with Land Attribute.

The Improvement table is 3.7 MB in size, has 25 columns (attributes), has 687895 rows, has a one-to-many relationship with Appraisal Account, a one-to-many relationship with Improvement Builtas, and a one-to-many relationship with Improvement Detail.

The Improvement Builtas table is 4.1 MB in size, has 26 columns (attributes), has 693344 rows, and has a one-to-many relationship with Improvement.

The Improvement Details table is 9.1 MB in size, has 5 columns (attributes), has 2286446 rows, and has a one-to-many relationship with Improvement.

The Land Attribute table is 1.9 MB in size, has 3 columns (attributes), has 553243 rows, has a one-to-many relationship with Appraisal Account.

The Sale table is 15.0 MB in size, has 12 columns (attributes), has 524229 rows, and does not have any established relationship with other tables in the database because the information contained in the Sale table is historical data.

The Seg Merge table is 0.6 MB in size, has 6 columns (attributes), has 170439 rows, and does not have any established relationship with other tables in the database because the information contained in the Seg Merge table is historical data.

The Tax Account table is 8.2 MB in size, has 28 columns (attributes), has 343639 rows, has a one-to-one relationship with appraisal account, and a one-to-many relationship with tax description.

The Tax Description table is 11.9 MB in size, has 3 columns (attributes), has 1193564 rows, and has a one-to-many relationship with tax account.

1. Provide the problem statement. Include why and how answering these problems or questions can be interesting and relevant. Please remember you may want to change or make the problems more specific later after completing the literature review. As said in class, we will take an iterative approach for our analysis.

As our group came together to draft our problem statement, we initially thought of who our audience would be as this would affect how we go about and present our analysis of our database. Would our analysis assist real estate agencies as their agents work with clients considering selling or buying their homes, or would it our analysis assist the Assessor-Treasurer's Office so those living in Pierce County or looking into living in Pierce County can access our analysis of Pierce County’s property value and tax information? Further, we explored any additional information that would could incorporate into our database, such as Pierce County demographic, crime, and school information. We also pointed out whether we wanted to approach our problem statement as a prediction or classification problem. Finally, our group landed on the problem statement of what factors impact Pierce County property value and taxes? Moreover, our group in interested in narrowing our broad problem statement to primarily focus on how either Pierce County crime information or Pierce County school information impacts property value and taxes and vice versa. Our decision to use either Pierce County crime information or Pierce County school information will be based on further research and review of the data available and how we can integrate that information into our database.

Reference/Appendix

1. Identify the sources of the datasets. If you downloaded the data from the websites, provide the URLs. Include the URLs in the "References" in your paper. (Appendix)

Main Database Link: <https://www.co.pierce.wa.us/736/Data-Downloads>

Additional Data Links:

<https://gisdata-piercecowa.opendata.arcgis.com/datasets/crime-data/data> (crime)

<https://gisdata-piercecowa.opendata.arcgis.com/datasets/schools/data> (school)

<https://gisdata-piercecowa.opendata.arcgis.com/datasets/address-points/data> (address points)

1. Upload the data to your SQL Server database. When you upload the datasets, make sure you review the raw data and choose the appropriate data types and sizes. (Appendix)
2. While doing this step, create a data dictionary for each attribute, including the name, description, valid domain values, data type, length, and example values. Many students use Excel to build a data dictionary. You can find a good example from the Freddie Mac site (http://www.freddiemac.com/fmac-resources/research/pdf/user\_guide.pdf (Links to an external site.). In addition, please add a column showing whether the attribute is nominal, ordinal, or numerical. Also, add a column showing the null ratio. Your data dictionary will be part of the Appendix of your paper. (Appendix)