Ethics Pledge

Consistent with the above statements, all homework exercises, tests and exams that are designated as individual assignments MUST contain the following signed statement before they can be accepted for grading.

I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination. I further pledge that I have not copied any material from a book, article, the Internet or any other source except where I have expressly cited the source.

Signature: Haodong Zhao Date: Apr 28th. 2019

Please note that assignments in this class may be submitted to

www.turnitin.com, a web-based anti-plagiarism system, for an evaluation of their originality.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Reading review**

**Extreme Visualization:**

**Squeezing a Billion Records into a Million Pixels**

visual presentations enable discovery of relationships, clusters, outliers, gaps, and other patterns. Scaling visual presentations from millions to billions of records will require collaborative research efforts in information visualization and database management to enable rapid aggregation, meaningful coordinated windows, and effective summary graphics.

The appeal of information visualization is to gain a deeper understanding of an important phenomenon that are represented in a database. Information visualizations are different from geographic and scientific visualizations since they have no inherent 2D or 3D structure but are designed to deal with multi-dimensional and more importantly multi-variate data.

An important contribution from the database community will be to develop scalable data structures and algorithms that support rapid update of visual displays for billion record databases. To meet data needs and user experience, need to overcome:

1. Database performance during exploration
2. Display performance to ensure 100msec updates
3. Visual representations that are compact and information abundant
4. Human perception of rich displays with specialized markers, aggregation icons, and density plots
5. Cognitively comprehensible interaction controls and coordinated windows

The author introduces several methods of visualization：

1. Atomic visualizations: one marker per data record
2. Aggregate visualizations: one marker per thousand data records
3. Density plot visualizations: color coded areas how users where to explore

Most operating systems users have shifted from command line interfaces to graphical user interfaces, greatly expanding the audience for computing. The narrow community of database query language users will expand greatly as effective visualization interfaces enable rapid and comprehensible access to large databases. If strong collaborations can be arranged between information visualization and database management researchers and implementers, then the use of billion record visualizations could become widespread.