

Visualization and Analysis of Formula One Racing Results with InfoZoom — the Demo

Michael Spenke and Christian Beilken

FIT – Fraunhofer Institute for Applied Information Technology

Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Michael.Spenke@fit.fraunhofer.de

Abstract: We show the application of the data analysis tool InfoZoom to a database of Formula One racing results. The user can interactively explore different tabular visualisations of the data. Initially, the complete table is shown. Queries are performed by zooming on selected values displayed in the table. In this way, the user can obtain a feeling of the data, detect interesting knowledge, and gain a deep understanding of the data set.

Keywords: Data visualization, interactive data exploration, user-interfaces for databases, data mining

1 Concepts of InfoZoom

InfoZoom displays database relations in tables with attributes as rows and objects as columns. In our example each column corresponds to the participation of a driver in a certain race. The table has about 16,000 columns. The attributes include the name of the driver, his team, the date and location of the race, and the starting and final position. InfoZoom compresses this table by reducing the column width until all the 16,000 columns fit on the screen (Figure 1). The column width is then about 0.05 pixels.

Special techniques are used to make such highly compressed tables readable. The most important is that neighbouring cells with identical values are combined into one larger cell. The width of each cell indicates the number of subsequent objects with this value. If a cell is too small to display a numeric value, a short horizontal line still indicates its relative height. If the table is sorted by an attribute, its value distribution can be seen at a glance. Correlations can

be visualized by subsequent sorts of different attributes and by animated zooms into interesting areas of the table. In the *Overview Mode*, the values of each attribute are sorted independently, so that the value distributions of all attributes can be seen simultaneously (Figure 2). Like the formula-cells in a spreadsheet program, derived summary attributes (like *sum*, *count*, *average*, *maximum*, etc.) can be defined which are automatically updated by InfoZoom whenever necessary.

2 The Demo

We show that questions like the following can be answered with a few mouse clicks:

- When was the first race in Monaco?
- When did the winner lap all opponents?
- How were the final standings in 2002?
- Who had the most victories / pole positions?
- Which winner had the worst start position?

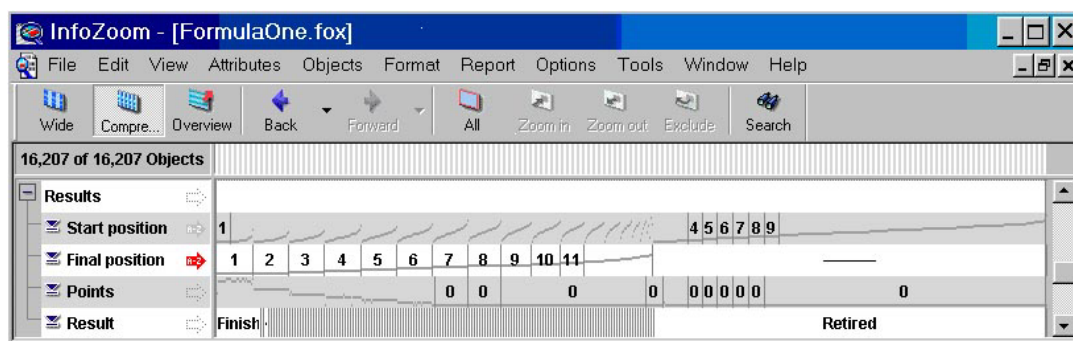


Figure 1: Some rows of the compressed table with all Formula One racing results

For example, in Figure 2 we have zoomed on the races of Michael Schumacher with two mouse clicks. We can make the following observations:

- Michael Schumacher started his career in 1991 but had comparatively few races in that year.
- He drove for *Benetton*, *Ferrari* and also one race for *Jordan*. (*Jordan* can only be seen by positioning the mouse over the small cell on the right.)
- He participated in 179 races.
- He made 945 points in his career.
- We can see that 1 and 3 are Schumacher's most frequent starting positions, and positions worse than 9 are extremely seldom.
- Schumacher won about each third race (64 times) and retired in about each fourth race. Only a few times was he lapped.

In Figure 1 we can observe the following:

- Winners usually started from a good position.
- Winners did not always get the same points.
- Nearly half of the drivers retired in each race.
- Retired drivers do not always get 0 points.
- Even the second winner did not always *finish* the race, i.e. he was lapped in some cases.

3 Related Work

The *TableLens* (Rao, 1994) is the only approach we know which also uses the basic idea of compressing

database tables until they completely fit on the screen. While InfoZoom displays each record in a column, in *TableLens* each row contains a record. Therefore, the *TableLens* cannot use the technique of combining neighbouring cells with identical values, which is vital to make textual values readable. Also, a counterpart to InfoZoom's *Overview Mode* is missing. The graphical representation of numeric values, however, is quite similar.

References

Rao, R. and Card, S. K., The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (Boston, MA, Apr 24–28, 1994)*, pp. 318–322.

Spenske, M.; Beilken, Chr., InfoZoom - Analysing Formula One racing results with an interactive data mining and visualization tool In: *Data mining II / Ebecken, N.[Editor]*, (2000), S. 455 – 464.

Spenske, M., Visualization and interactive analysis of blood parameters with InfoZoom In: *Artificial Intelligence in Medicine, Bd. 22, Nr. 2, (2001)*, S. 159 – 172.

<http://www.humanIT.de> – The InfoZoom home page. A free test version of InfoZoom can be obtained.

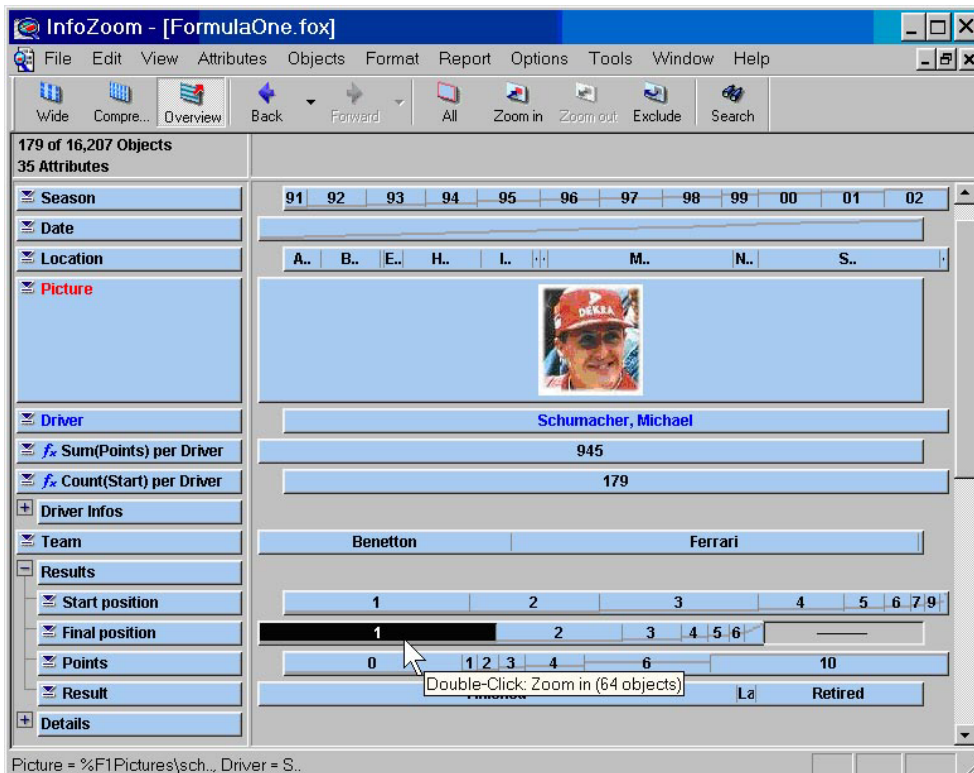


Figure 2: Overview of all races of Michael Schumacher