connotation words when compared to the other car types. If we had more information about the state, city, manufacturing unit for each car/brand, etc., we could have extracted useful information that could identify the places/units that elicited the greatest number of complaints.

OTHER APPLICATIONS OF TEXT MINING

Unstructured text is very common and, in fact, may represent the majority of information available to a particular research or data mining project. The selection of tools or techniques available with STATISTICA, along with the Text Mining module, can help organizations to solve a variety of problems. A few to mention are the following:

1. Extracting information reflecting customers/employees/public—opinions, needs, and interest (e.g., visualizing semantic spaces using 2D, 3D plots);
2. Filtering unwanted documents/emails (using stop list, include lists, etc.);
3. Predicting customer satisfaction levels (e.g., negative connotations);
4. Clustering similar words/documents. (e.g., reviews, research papers, survey data, etc.);
5. Classifying or organizing documents (e.g., electronic documents about general information can be classified into different subgroups);
6. Predicting/routing new documents, etc. (The rules for clustering or classifying or predicting can be used to score new documents.)

CONCLUSION

This simple tutorial is intended to help you understand how the STATISTICA Text Miner module, along with numerous STATISTICA Data Miner tools and techniques, can be used for finding solutions to problems that require knowledge of language and computing technology. More importantly, extraction of useful insights or information from unstructured data could be used as input for decision-making purposes.