Chapter 3 talks about language analysis with mathematics methods. It first introduces three major models to treat regularities in languages. Next, it uses data and examples to show the distribution of different kinds of English words. The most frequently used English words are article words and then verbs while nouns are least frequently used. Next, it introduces information theory, where mathematical formulas and methods are utilized to acquire the relationship between different distributions. However, these formulas and methods still have some disadvantages. Moreover, it introduces that using matrices and vector spaces to represent a huge number of measurements is an efficient way. In this part, PCL, a method of statistical analysis is explained from aspects of both advantages and disadvantages. Finally, a theoretical framework is illustrated, which has three phases: canonicization, determination of the event set and statistical inference.

Chapter 5 talks about attributional analysis. There are some important factors: an understanding of the requirements of the final answer and the amount as well as the type of training material available. The second one is significantly mentioned here. First, for unsupervised analysis, scatter plot is often used. Also, PCA and vector spaces mentioned before are useful. Furthermore, MDS, which could visualize the difference and distance, is also a great method. Last, cluster analysis is kind of similar to MDS speaking of the difference-distance transition. It also use a cluster to replace close pair of events. Compared with unsupervised analysis, supervised analysis requires the documents to be categorized before analysis. First method of this is to use simple statistics, which uses a lot of what I’m now learning in MATH 216. Both advantages and disadvantages of it is discussed. Besides, LDA, which is similar to PCA to some extent, uses categorized data to infer differences. In addition, distance based methods and general machine learning techniques are also useful. Last, SVMs, which are a relatively new classification method, have been applied to a huge variety of problems.