# Study Plan

**Introduction to My Background**

I am currently a fourth year Ph.D. candidate in the Department of Statistics at Purdue University, and expected to get my Ph.D. degree in August 2016.

Realizing the importance of data analysis, as well as encouraged by my dream, I decided to go to USA for my advanced study in 2010. I originally started as a Master student and then transferred to Ph.D. program within the same department at the end of my second academic year. Over the past almost three years, I maintained a 3.81 GPA, which is very hard in a highly competitive Ph.D. program. By the end of my second year, even though I was still a Master student, I have finished the required 8 courses for the Ph.D. program and passed all three Ph.D. Qualifying Examination.

**Why do I want to pursue the Ph.D. degree in Statistics?**

Four years of undergraduate studies have introduced me to the industry of mathematics. I did perfectly well in my major, information and calculation science, and cultivated comparatively math and programming capabilities, which helped me lay a solid foundation for future graduate studies in Statistics.

During my first two years' studying in the Department of Statistics, I have taken some basic courses in statistics, and participated in many interesting statistical seminars. Especially during the last semester, STAT695T, taught by Prof. William Cleveland, profoundly stimulated my interests in the area of data visualization statistics. The more I learn, the more eager I feel the willing to become an expert in data visualization area. I really enjoy the act of discovering something new and also the feeling of truly understanding it. I have also come to realize that doing research as a Ph.D. student requires far more than just completing the course work; it also requires motivation, innovation, commitment, and perseverance. I need further education and training to make myself think sophisticatedly, work creatively, and study independently during my research life in the future.

All above experiences and observation lead me to continue my study and pursue Ph.D. degree in large and complex data visualization area. I felt passionate about challenge that there is still a lot of space to improve the visualization performance. Thus, encouraged by advisor, Prof. William Cleveland who is the pioneer of data visualization, I decided to continue my graduate study in the field of Statistics for a Ph.D. degree.

**Detailed Study Plan**

The normal length of the study is about 60 to 72 months. Ph.D. candidates must complete at least 90 semester hours of coursework. By the end of this semester, Fall 2013, I have finished 79 semester hours. For the detailed course information please check the academic transcript attached. For the rest of time, I will register at least 11 hours for STAT699 research Ph.D. Thesis for the research credits.

At the same time, I am going to take research assistantship from Prof. William S. Cleveland, and perform responsibilities as a RA required by professor and department. For the last year of my Ph.D. program, I will conduct research and working on the dissertation which is focusing on data visualization of spatial time series data with the guidance from my advisor and other professors, and finish writing dissertation and thesis. Finally, I will publish my paper on top tie conference and pass Ph.D. final examination (oral defense of the dissertation).

**My Research Interest**

My research mainly focuses on how to apply the divide and recombine method to the data visualization for large spatial time series data. Divide and recombine is a new method which is trying to speed up the data analysis process by dividing the data first and then recombining the results under the concept of parallel mathematically computing. And spatial time series data is the data which is time series data, like temperature and precipitation over time with spatial factor, like elevation, longitude, and altitude. I am currently working with Prof. William S. Cleveland (wsc@purdue.edu) who is mainly working on large and complex data analysis. Prof. Cleveland and I both believe that spatial time series data is another important area for large data visualization and analysis, and we are planning to break the traditional analysis method for spatial time series data, and create a new mathematical analysis and visualization method under the divide and recombine mechanism.

Currently I am financially supported by my advisor, Prof. William S. Cleveland. The department head of statistics is Prof. Rebecca W. Doerge (doerge@purdue.edu)