**Title: Preparing Statistician to be Successful Data Scientist**

**Description:** With recent big data revolution, enterprises ranging from FORTUNE 500 to startups across the US are hungry for data scientists to bring valuable business insight from all the data collected. Statisticians are great data scientist candidates, but there are relatively few data scientists with statistics education background. The course aims to prepare statisticians to be successful data scientists. To fill the gap between statisticians and data scientists, we will walk through the needed data science skills such as data access and manipulation (i.e., extract-transform-load) in production environments.

Data science is a combination of science and art with data as the foundation. We will also cover the art part including data science project flow, general pitfalls in data science projects, and soft skills to communicate with business stakeholders effectively. It will be hands-on. We will use the Databricks community edition cloud platform and R-Studio to illustrate programming, big data platform (such as Spark) and standard machine learning algorithms.

**Instructors**: Dr. Ming Li (Amazon) and Dr. Hui Lin (DowDuPont)

Dr. Ming Li is currently a Senior Data Scientist at Amazon and an Adjunct Faculty of Department of Marketing and Business Analytics at Texas A&M University-Commerce. He was the Chair of Quality & Productivity Section of ASA for 2017. He was a Data Scientist at Walmart and a Statistical Leader at General Electric Global Research Center. He obtained his Ph.D. in Statistics from Iowa State University in 2010. With in-depth statistics background and a few years’ experience in data science, he has trained and mentored numerous junior data scientist with a different background such as statistician, programmer, software developer, database administrator and business analyst. He is also an Instructor of Amazons internal Machine Learning University and was one of the key founding members of Walmart’s Analytics Rotational Program which bridges the skill gaps between new hires and productive data scientists.

Dr. Hui Lin is currently Data Scientist at DowDuPont. She is the leader in the company at applying advanced data science to enhance Marketing and Sales Effectiveness. She has been providing statistical leadership for a broad range of predictive analytics and market research analysis since 2013. She is a co-founder of Central Iowa R User Group, blogger of scientistcafe.com and 2018 Program Chair of ASA Statistics in Marketing Section. She enjoys making analytics accessible to a broad audience and teaches tutorials and workshops for practitioners on data science. She holds MS and Ph.D. in statistics from Iowa State University, BS in mathematical statistics from Beijing Normal University.

