My Data Science Master's Degree Program

GPA: 4.0
Courses Remaining: 0

By: Ryan Miller

Courseload

Spring 2021	Summer 2021	Fall 2021	Spring 2022	Summer 2022
Database Systems	Data Mining	Advanced Big Data Analytics	Modeling in Regression	Machine Learning
Java Programming		Database Administration	Advanced Statistical Modeling	Big Data: Managerial Perspective
Statistical Modeling			Digital Media Analytics	

Duration
1/18/21 - 8/16/22

Technologies and Languages

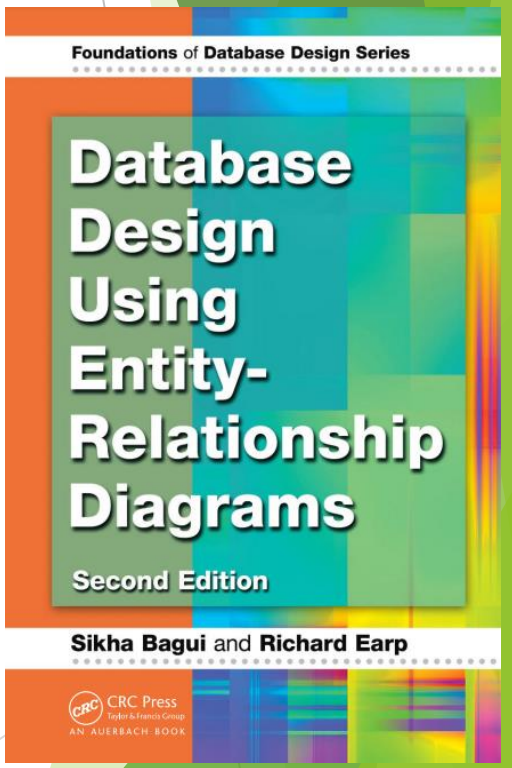
Classes Taken Thus Far

Java Programming

- Introduction to algorithms and object-oriented programming.
- Topics include Variables, Branches, Loops, Arrays, Methods, Objects and Classes, Memory Management, Input/Output, Inheritance, Abstract Classes and Interfaces, Recursion, Exceptions, Generics, Collections, Searching and Sorting Algorithms.
- There were 6 projects as well.

Database Systems

- Introduction to database systems and database management system architectures.
- Various database models are discussed with emphasis on the relational model and relational database design.
- Case applications using fourth-generation languages, such as SQL are included.



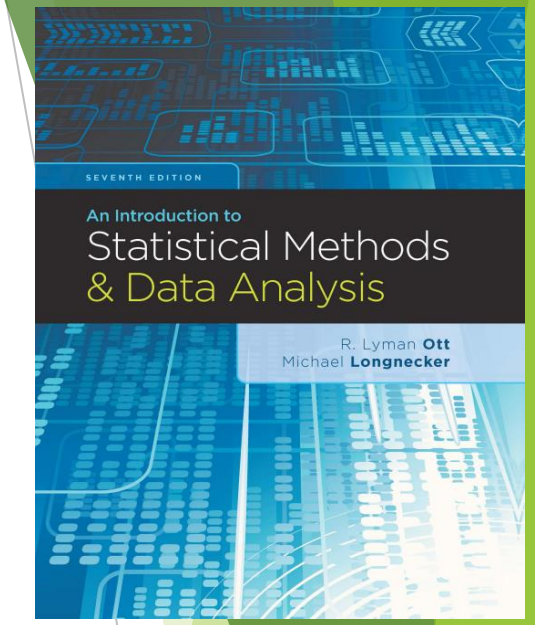
Classes Taken Thus Far (cont.)

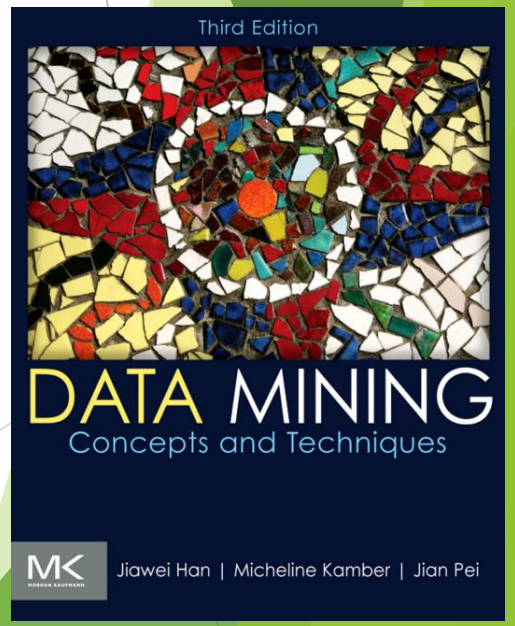
Statistical Modeling

- Examination of statistics and data analysis beyond an introductory course.
- Topics covered included data visualization, point, and interval estimation, hypothesis testing of means, variances, and proportions, and linear and logistic regressions.
- Emphasis was placed on conducting reproducible research.

Data Mining

- Introduction to database systems and database management system architectures.
- Various database models are discussed with emphasis on relational database design, OLAP architectures, and Data Warehousing.
- A course-long group project was conducted in which my group ran Association Rule, Decision Tree, Classification, and Clustering algorithms on census data and interpreted our results.





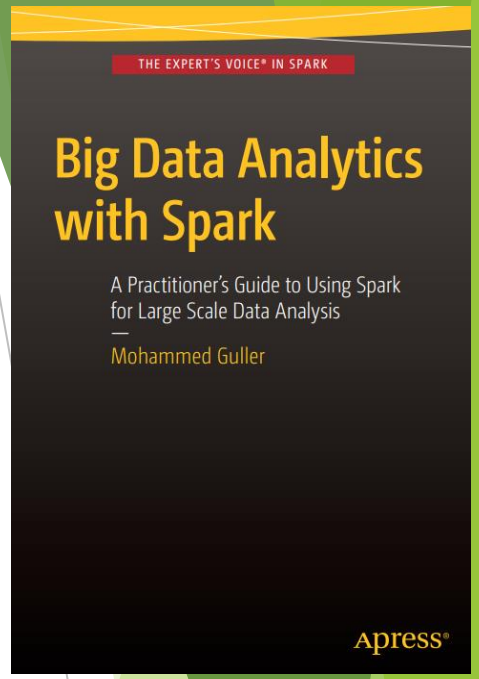
Classes Taken Thus Far (cont.)

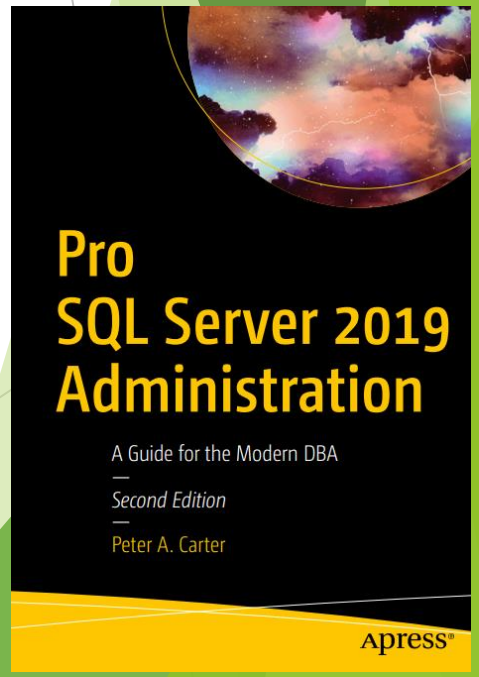
Advanced Big Data Analytics

- In this course, I studied advanced methods to handle and analyze very large data sets in Hadoop's Big Data environment.
- I worked with the Spark architecture in a MapReduce framework and learned to apply machine learning algorithms in Spark.
- Formulated a machine learning pipeline in SparkML using PCA and Random Forest Classification on network traffic identifying attacks.

Database Administration

- Database administration skills covering installation, configuration and tuning a database, administering servers and server groups, managing and optimizing schemas, tables, indexes, and views, creating logins, configuring permissions, assigning roles and performing other essential security tasks, backup and recovery strategies, automation and maintenance.





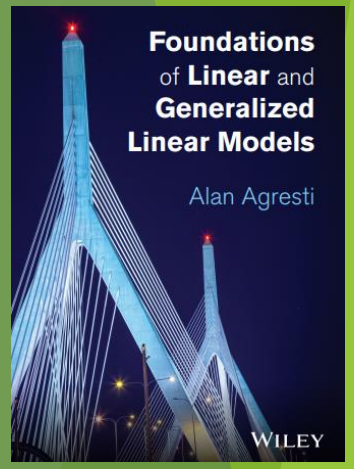
Classes Taken Thus Far (cont.)

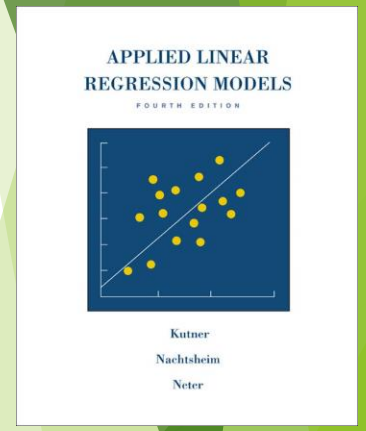
Advanced Statistical Modeling

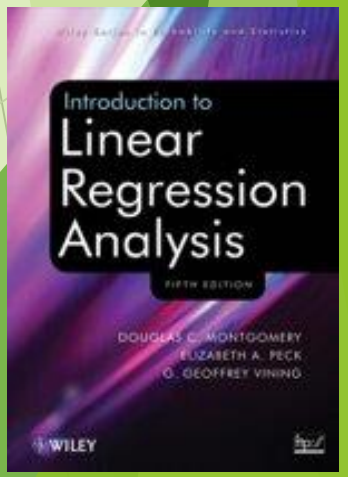
- Generalized Linear Analysis, Nonlinear Regression Analysis, Spatial Cluster Analysis, Exponential Dispersion Models, Long-Short Term Memory Networks, K-means Clustering, Bayesian Networks, Generalized Additive Models, Quantile Regression

Modeling in Regression

- Nonlinear Regression, Influence Diagnostics, Eigensystem Analysis Of X'X Matrix, Logistic Regression, Ridge Regression, Robust Regression, and Generalized Linear Models







Classes Taken Thus Far (cont.)

Machine Learning

- Support Vector Machines, Generative And Discriminative Learning, Parametric And Nonparametric Learning, Similarity Learning, Bayesian Learning, Decision tree Learning with Sci-kit Learn
- Artificial Neural Networks (ANN), Deep Learning (DL), and Convolutional Neural Networks (CNN) with Tensorflow
- Reinforcement Learning, Q-learning, Markov Decision Process (MDP), Deep Reinforcement Learning (DRL), Inverse Reinforcement Learning (IRL)
- Kernel Methods, Bias/Variance Tradeoffs, Soft Margins, and Applications of Machine Learning

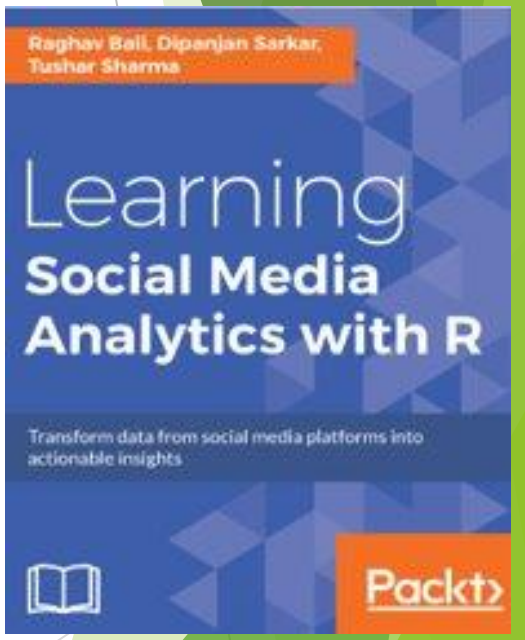
Big Data Mining: A Managerial Perspective

- We identified descriptive, predictive, and prescriptive analytics and their applications in businesses with legal, privacy, and ethical concerns.
- Sensitivity analysis and linear programming methods were utilized to capture optimized channel & resource distributions in manufacturing and logistics.
- We developed network models and optimized weighted scheduling criteria with path-finding algorithms.
- In addition, we cost-optimized economic decision trees with backwards induction methods.

Classes Taken Thus Far (cont.)

Digital Media Analytics

- Utilized the R programming language, insights, real-time analytics, graph network traversal and machine learning methods.
- Created methods for API ingestion of streaming text data from YouTube and Twitter.
- Improved my ability to implement and prove the validity of marketing campaigns while engaging a global or regional audience.
- Built a recommendation engine with natural language processing and unstructured data.



Projects

- Tensorflow/Keras Convolutional Neural Network image classification
- Spark/Scala Random Forest Classification of network traffic attacks
- Parallel Processing & Regularized Regression/Classification on car prices
- OLS Regression & case study on diamond prices
- Social media sentiment analysis comparison on COVID-19, the US FDA, and retirement planning
- Hierarchical clustering & association rule mining of US census data
- SQL Server Database creation, securing, encrypting, permissioning, & automation