Core Areas of Competency for Graduates of M.S in Business Analytics Program

The following are the six key areas of competency for that plan to cover through the M.S. in Business Analytics Program

1. Data Mining/ Machine Learning
2. Programming and Software Tools
3. Applied Probability and Statistics
4. Databases and Data Processing
5. Quantitative Algorithms
6. Business Acumen

The rest of this document describes the topics covered related to each area.

1. **Data Mining/ Machine Learning**
2. **Principles of Machine Learning and Data Modelling** 
   * Data Structures and Types of Variables
   * Supervised vs. Unsupervised Machine Learning Modelling
   * Data Preparation Techniques
   * Feature Engineering
   * Evaluation of Machine Learning Models
   * Optimizing Machine Learning Models
   * Ensemble Learning
   * Common Mistakes in Modelling
3. **Regression Modelling** 
   * Concepts and Definitions
   * Performance Metrics
   * Linear Regression
   * Generalized Linear Models (GLM)
4. **Classification Modelling** 
   * Concepts and Definitions
   * Performance Metrics
   * Logistic Regression
   * k-Nearest Neighbor (k-NN)
   * Naïve Bayes
   * Decision Trees (applied to Regression as well)
   * Random Forrest (applied to Regression as well)
   * Gradient Boosted Machines (applied to Regression as well)
   * Support Vector Machines (applied to Regression as well)
   * Neural Networks (applied to Regression as well)
5. **Segmentation and Clustering**
   * Concepts and Definitions
   * Performance Metrics
   * k-means Clustering
   * Hierarchical Clustering
6. **Recommendation Systems** 
   * Concepts and Definitions
   * Performance Metrics
   * Apriori algorithm for association mining
7. **Time Series Analysis** 
   * Concepts and Definitions
   * Performance Metrics
   * Stationarity, causality and invertibility
   * Autoregressive Integrated Moving Average (ARIMA)
8. **Graph Analytics** 
   * Concepts and Definitions
   * Centrality and Connectivity Measures
   * Application to Social Network Analysis
9. **Text Analytics** 
   * Concepts and Definitions
   * Feature Extraction
   * Topic Modelling
   * Sentiments Analysis
10. **Programming and Software Tools**
11. **Data Mining, Machine Learning and Quantitative Programming: R**

Implementation of the following Data Mining/ Machine Learning methods:

* + - Linear Regression
    - Generalized Linear Models
    - Logistic Regression
    - Decision Trees
    - Random Forrest
    - Gradient Boosted Machines
    - Support Vector Machines
    - Neural Networks

Implementation of the following Quantitative methods:

* + - Linear Programming
    - Integer Programming
    - Goal Programming
    - Simulated Annealing
    - Network Models
    - Genetic Algorithms/ Programing

1. **Open Source Data Mining Tool: WEKA**

Implementation of the following Data Mining/ Machine Learning methods:

* + - k-Nearest Neighbor (k-NN)
    - Naïve Bayes
    - Decision Trees
    - Bagged/Boosted Trees
    - Association Mining

1. **Data Preparation General Purpose Programing: R**
   * + Calculating Various Statistics and Math Calculations
     + Calculating Probability Values
     + Data Input/ Export
     + Data Cleansing
     + Data Wrangling and Data Subsetting
     + Feature Engineering
     + Applying summarization and Aggregate functions
2. **Database: SQL**
   * + Principals of Database Design
     + Using SQL to Create, Update and Delete Tables
     + Using SQL to Select a subset of Data
     + Using SQL to Join Tables
     + Using SQL to perform various Aggregate Functions
3. **Visualization: R/ Tableau**

* Using R ‘ggplot’ for explanatory analysis and communicating the insights
* Using R ‘Shiny’ for interactive visualization and dash boarding
* Using Tableau for explanatory analysis and communicating the insights

1. **Software Repository and Development Platforms: Github/Git**

* Creating a new repository
* Fork and Push changes to a repository
* Clone a public project
* Send a pull request/ Merge changes from a pull request

1. **Big Data and High performance Computing: Spark, Hadoop, AWS, Azure, MLlib, R**

* Spark and Big Data Ecosystem
* Using Spark's MLlib for Machine Learning
* Scale up Spark jobs using Amazon Web Services
* Using R in Azure Machine Learning Studio
* Parallel computing using R

1. **Applied Probability and Statistics**

* Statistics: Measures of Central Tendencies
* Statistics: Measures of Dispersion
* Statistics: Measures of Skewness
* Statistics: Measures of Dependence
* Statistics: Statistical Significance
* Statistics: A/B Testing
* Probability: Probability Distributing Functions
* Probability: Normal Distribution
* Probability: Uncertainty and Confidence Intervals
* Probability: Conditional Probabilities
* Probability: Bayesian Probability
* Probability: Information Entropy

1. **Databases and Data Processing**
2. **Relational Databases**

* Concepts and Definitions
* Entity-Relationship Diagrams
* Structured Query Language (SQL)
* Normalization, Transaction management and Concurrency Control
* SQL as an Analytical Tool
* Intro to NoSQL Databases and Applications

1. **Big Data Platforms**

* Big Data Paradigms (e.g. MapReduce)
* Big Data Platforms (e.g. Hadoop)
* Big Data Extraction/Integration
* Big Graph Processing
* Big Data Stream Techniques and Algorithms

1. **Quantitative Algorithms**

* Linear Programming
* Duality in Linear Programming
* Integer Programming
* Goal Programming
* Simulated Annealing
* Network Models
* Genetic Algorithms/ Programing

1. **Business Acumen**

* Practical Case Studies Based on Real-World Data from Different Industries
* Formulation of Business Problems to Solve Using Analytics Group Projects
* Data Collection and Communication of Findings
* Operationalizing Analytical Models in Practice
* Common Mistakes in Analytical Modeling

**Appendix I: List of Core Courses for the MS in MSBA Courses**

**Fall Offerings**

* + MIS 64036:   Business Analytics
  + MIS 64082: Database Management and Database Analytics
  + CS 63015: Data Mining Techniques
  + MIS 64018: Quantitative Management Modeling

**Spring Offerings**

* + MIS 64037: Advanced Data Mining and Predictive Analytics
  + CS 63016: Big Data Analytics
  + MIS 64038: Analytics in Practice

**Summer Offerings**

* + MIS 64098:   Capstone Project in Business Analytics