DSC2008 Business Analytics—Data and Decisions

Semester II, 2016/2017

"Most companies today have plenty of data. However, creating intelligence and gleaning real insights from this data is what continues to elude organizations."—*Competing on Analytics: The New Science of Winning.*

Business decisions are often made under uncertainty. In the modern business environment, technological advances facilitate the collection of huge amounts of data which can potentially improve the decision-making process. Successful businesses make use of Business Analytics and Business Intelligence, which are fundamentally based on quantitative statistical methods and optimization procedures, to identify patterns and trends in their data which eventually lead to insightful projections and realistic predictions.

This module relates to DSC1007 Business Analytics — Models and Decisions (Business Analytics I). It introduces students to the fundamental concepts of statistical inference such as parameter estimation and hypothesis testing, as well as to statistical tools useful in business analytics, such as regression analysis, time series analysis and cluster analysis. This continues the theme of delivering hands-on experience in modules focusing on analytics and operations.

This module was co-designed, and is co-taught, by the Department of Decision Sciences in the NUS Business School and the Department of Statistics and Applied Probability in the Faculty of Science, to draw upon the relevant expertise from the two units.

The module provides all BBA students with a common statistical grounding for Business Analytics, upon which specialization may be built depending on each student's chosen major.

In keeping with the principles of Rigor and Relevance of Business Analytics I, students are expected to acquire the following knowledge and abilities.

Rigor

Building on the foundations of probability from Business Analytics — Models and Decisions, this module covers fundamental concepts underpinning the following business analytic tools:

1. Data summarization: pivot table;
2. Statistical inference: sampling distributions, confidence intervals and hypothesis testing;
3. Regression analysis: linear and multiple regression, regression diagnostics and model building;
4. Time series analysis: smoothing, regression-based models, ARIMA models and forecasting;
5. Clustering and market segmentation: K-means method and Hierarchical cluster analysis.

Emphasis will be made on how, what and why certain tools are useful and, and what their ramifications would be when used in practice. Hands-on assignments and a Project will practise students on model building.

Relevance

Module content makes use of examples that are based on current events and timely business topics.

Adopting the Plan/Do/Report problem-solving approach, worked examples show students how to clearly define the business decision to be made and *plan* which method to use, *do* the business analysis with data-based numerical evidence and with the help of illuminating graphical displays, and finally *report* their findings and recommendations to the decision maker. This approach is reinforced using case-study projects involving real data in which students investigate a business-related question or make a business decision. This module also makes active use of the Excel spreadsheets and the SAS software.

Students will be equipped with the ability to "tell a story" and provide insights based on (big) data given to them in their future workplaces (e.g. during their internships).

Prerequisites

None.

Module Outline

- Describing data (Week 1-2)
  - Data type
  - Data statistics
    - Description
    - Sampling
  - Data distributions
    - Standard distributions
    - Sampling distributions
      - Confidence Interval
      - Central Limit Theorem

- Inferring from Data (Week 2-4)
  - Testing hypothesis
  - Comparing distributions
    - Histogram
    - Boxplot
  - Comparing averages
    - 1 sample
    - 2 samples
    - Multiple samples
      - Analysis of Variance
  - Relating variables (Week 4)
    - Correlation

- Predicting beyond Data (Week 4-6)
  - Simple regression
  - Multiple regression
    - Variable selection

- Transformations
- Different Intercepts & Slopes
- Outliers
- Missing values
- Logistic regression

Time Series Analysis (Week 7-10)

- Multiple-regression-based time series models
- Smoothing methods
  - Simple moving average
  - Weighted moving average
  - Single exponential smoothing
  - Holt-Winters exponential smoothing
- Autoregressive Integrated Moving Average models

Cluster Analysis (Week 11-12)

- Distance measure
- Hierarchical cluster analysis
- K-means method

Assessments

Two Individual Tutorial Assignments: 20% (10% each)
Group Project (report and presentation): 20%
Class Participation: 20%
Final Examination: 40%