



Sofaer Global MBA

1238.3272.01

Data Driven Business Thinking

Module 3 – 2021

Course Section Details

| Day | Hour |
|------------|-------------|
| 03/03/2021 | 12:45-15:30 |
| 10/03/2021 | 12:45-15:30 |
| 17/03/2021 | 12:45-15:30 |
| 24/03/2021 | 12:45-15:30 |
| 07/04/2021 | 12:45-15:30 |
| 21/04/2021 | 12:45-15:30 |
| 28/04/2021 | 12:45-15:30 |

Lecturer: Dr. Shachar Reichman, sr@tauex.tau.ac.il. Office hours: By appointment
TA: Tal Shoshani, talshoshani@mail.tau.ac.il

Course Description

In the digital era, commercial, consumer and social processes are carried out on a broad scope in a digital environment. Operational and managerial systems, along with Internet websites and cellular devices enable routine collection of data concerning all activities performed in this environment. The ability to combine information concerning the consumer allow companies to build a complete picture of the consumption journey, from understanding consumer preferences, through analyzing business activities to the consumption decision making.

In spite of the richness of data collected, the ability to leverage advanced data analysis is not optimally utilized by companies to make decisions and achieve a significant business contribution.

The goal of the course is to develop data driven business thinking through an understanding of the tools and technologies that allow firms to address the business questions and challenges in the digital era. The course deals with data driven business solutions and focuses on data mining methods and their applications using real world cases. Students will learn data analytics methodologies with a big data analytics tool – Microsoft Azure ML Studio. During the course, we will review examples of data challenges from various domains and discuss the uses of data analytics to support managerial decisions.

Course Objectives



Sofaer Global MBA

Upon completion of the course, the student will:

- Understanding Data Driven Decision process and their managerial applications
- Learning the data analytics process
- Familiarity with online data sources and business data sources
- Learning predictive analytics models
- Understanding how to apply predictive analytics to business challenges.

Course Outline

| # | Topics | Assignments |
|---|---|---|
| 1 | Data Driven Decision (DDD) Making | |
| 2 | Data – Challenges and Opportunities <ul style="list-style-type: none"> a. Online data sources b. Data handling and cleaning c. Managing multiple data sources | DDD based on online data |
| 3 | Descriptive Analytics <ul style="list-style-type: none"> a. Business intelligence b. Data visualization | |
| 4 | Predictive Analytics <ul style="list-style-type: none"> a. Introduction b. Fitting a model to data (revisiting linear and logistic regressions) c. Classification d. Trees and random forest e. Clustering | Predictive analytics 1 – Linear Regression |
| 5 | Business Value of Predictive Analytics Results <ul style="list-style-type: none"> a. Goals, measures, and model evaluation b. Over-fitting and its avoidance | Predictive analytics 2 - Classification |
| 6 | Data Science Applications: <ul style="list-style-type: none"> a. Recommendation systems b. Network analysis | |
| | Ethics, Privacy and Security Challenges | |



Sofaer Global MBA

| | | |
|---|--|---------------|
| 7 | | Final project |
|---|--|---------------|

* note the coverage of these topics may change slightly.



Sofaer Global MBA

Course Site (Moodle)

The course Moodle site will be the primary tool used to communicate messages and material to students. It is, therefore recommended to periodically check the course site in general, periodically, before each lesson, at end of the course as well.

Course slides will be available on the course site.

Assessment and Grade Distribution

| # | Assignment type | Weight | Group size |
|---|--|--------|--------------|
| 1 | DDD - group assignment and presentation in class | 10% | 2-3 students |
| 2 | Predictive Analytics 1 – Linear Regression | 10% | 1-2 students |
| 3 | Predictive Analytics 2 - Classification | 10% | 1-2 students |
| 4 | Final project (group assignment) | 70% | 2-3 students |

Course Assignments

Assignment notes: You will complete assignments in self-selected groups of 2-3 individuals. You will submit each assignment by the beginning of the class on each assignment's due date. The general instructions for the assignments appear at the end of this document. Detailed instructions will be posted on the course site (Moodle).

Should a student become unable to complete an assignment or course requirement, s/he must notify the TA of the course in advance via email

Implications for Business Ethics

Issues in business ethics will be discussed on the 7th meeting. Specifically, we will discuss ethical issues of collecting and using online data, mobile data, and location data for business purposes. We will discuss privacy concerns of different data types and how data analytics may influence consumers and companies.

Grading Policy

As of the 2008/9 academic year the Faculty has implemented a grading policy for all graduate level courses. This policy applies to all graduate courses in the Faculty, and will be reflected in the final course grade. Accordingly, the final average of the class for this course (which is a core course) will fall between 83-87. Additional information regarding this policy



Sofaer Global MBA

can be found on the Faculty website.

Evaluation of the Course by Student

Following completion of the course students will participate in a teaching survey in order to evaluate the instructor and the course for the benefit of the students and the university.

Requires Reading

- The primary course material is the course book "**Data Science for Business**" (Provost, F., Fawcett, T., O'Reilly Publisher, 2013).

Recommended Reading

- "**Data Mining: Practical Machine Learning Tools and Techniques (Second Edition)**" (Ian H. Witten., Eibe Frank., Morgan Kaufmann Publishers)

Course Assignments:

Assignment 1 – Data Driven Decision Making

You were asked to consult a firm that is interested in opening a restaurant in Europe.

The main questions are:

- Restaurant type
 - Cuisine
 - Pricing
- Location – where to open the restaurant

Provide a data-driven recommendation, using an analysis multiple online data sources and as analysis of the different options.

Each group will present their analysis and recommendation in class (~3 minutes presentation). Submission will include the presentation slide deck.



Sofaer Global MBA

Assignment 2 - Making Your First Predictions Using Linear Regression

In the Azure ML Studio platform perform the following (the dataset will be posted on the course site on Moodle):

1. Perform data preprocessing – describe the steps you performed and the summary of the changes to the dataset (i.e. number of rows, columns, missing values and etc.).
2. Perform descriptive analytics – describe the key measures of important variables of the data, including single variable and multi variables metrics.
3. Build a regression model to predict the dependent variable based on all the variables
4. Evaluate the prediction model.
5. Estimate 3 different linear regression prediction models and compare their outputs

For all questions, the answers should include the analysis performed (including screenshots of the solution on the AzureML platform), a summary of the results, and visualizations.

Assignment 3 - Classification Models

(the dataset for the assignment will be posted on the course site on Moodle)

1. Perform data preprocessing and descriptive analytics (similar to the tasks performed in assignment 2)
2. Build classification models to predict the “Class” variable:
 - a. Perform variables selection.
 - b. Decision tree
 - i. Describe the generated tree. Could you explain the logic?
 - ii. Calculate the information gain of the first two branches of the tree.
 - c. Classification using logistic regression
3. Estimate and compare the results of the different models – which one would you recommend?

For all questions, the answers should include the analysis performed (including screenshots of the solution on the AzureML platform), a summary of the results, and visualizations.