BIG DATA & BUSINESS ANALYTICS & INTELLIGENCE

BACHELOR IN MANAGEMENT INFORMATION SYSTEMS
Professor: ALVARO JOSÉ MÉNDEZ LÓPEZ
E-mail: ajmendez@faculty.ie.edu

Academic year: 18-19
Degree course: THIRD
Semester: 1º
Category: COMPULSORY
Number of credits: 6.0
Language: English

PREREQUISITES
Some previous knowledge of Descriptive and Inferential Statistics is beneficial for better understanding of subject.

SUBJECT DESCRIPTION
The subject is a broad introduction to several areas within Analytics subject. Business Intelligence will be studied to understand what has happened in the past and what’s happening currently. To accomplish this task, a leader and very well known BI tool, Tableau, will be used.

The core of the subject is Business Analytics, an area that is used to predict what will happen in the future. Also, it will be covered from a practical point of view, giving the alumni a tool that will serve to understand and practice with topics seen.

OBJECTIVES AND SKILLS
The main objective of this course is to learn and exercise “multivariate thinking”. The students will be able to understand and use the standard collection of multivariate statistical methods operated in behavioral sciences and market research.

To achieve this goal, it is necessary that we first learn how to analyze past data, by building dashboards that allow to summarize Key Performance Indicators.

The amount of mathematics and statistical basis will be strongly reduced in this course expending the majority of the time in:

- Emphasizing the applications of multivariate techniques in a every day's practice in real business and industry
- Intensively practicing how to run analysis and properly interpret and present the results obtained for helping decision making processes

More specifically, the learning expected outcomes of the course are:
- To have a broad view of how a dashboard is built and what are the needs in terms of information that these require.
- To have a broad panorama of the different contexts in which multivariate analysis becomes a valuable resource, from pure business context to more research-oriented activities.
- To clearly understand the real usefulness of these techniques; for that, the teachers will illustrate theory with real interdisciplinary examples and exercises coming from own experience as data analysis consultants.
- To learn the motivation behind every different technique, when to apply each one and how to come to meaningful conclusions.
- To learn how to fruitfully conduct multivariate analysis using a popular statistical platform
- To learn how to present the results of a data analysis session as they would be presented in a professional context empowering students’ abilities to see and communicate statistical ideas effectively through written and oral presentations.

**METHODOLOGY**

- Flipped - classroom method: For each lecture, a script will be uploaded into the campus before the class. I strongly recommend students to read these documents before each class and even do a first try on the guided examples contained. The goal is to be well prepared when that session takes place.
- Lectures: The lectures are intended to present the basic material, solve the questions and doubts that have arisen when reading the script and practice with the platforms.
- Hands on: The student will also take an active role in the course solving specific exercises in groups, using the different techniques and delivering an executive document and a technical annex with the results.

<table>
<thead>
<tr>
<th>Teaching methodology</th>
<th>Weighting</th>
<th>Estimated time a student should dedicate to prepare for and participate in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lectures</td>
<td>40.0 %</td>
<td>60 hours</td>
</tr>
<tr>
<td>Discussions</td>
<td>10.0 %</td>
<td>15 hours</td>
</tr>
<tr>
<td>Exercises</td>
<td>0.0 %</td>
<td>0 hours</td>
</tr>
<tr>
<td>Group work</td>
<td>30.0 %</td>
<td>45 hours</td>
</tr>
<tr>
<td>Other individual studying</td>
<td>20.0 %</td>
<td>30 hours</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100.0 %</td>
<td>150 hours</td>
</tr>
</tbody>
</table>
PROGRAM

SESSIONS 1 - 2

INTRODUCTION TO COURSE AND MAIN ANALYTIC CONCEPTS
- Introduction to the course (Syllabus, Objectives, Tools, Rules, Grading, Topics)
- What is Big Data? How Data Analytics is changing. Business Cases
- Business Intelligence vs. Business Analytics
- Main techniques in Business Analytics
- Brief introduction to Artificial Intelligence
- Review of platforms to be used

SESSION 3 (FACE TO FACE)

INTRODUCTION TO BUSINESS INTELLIGENCE
- What is Business Intelligence and main benefits
- Needed Components in Business Intelligence
- Relevance of data preparation and data quality
- Leader Vendors
- Introduction to software used: Tableau

SESSIONS 4 - 8

CREATION OF PROFESSIONAL DASHBOARDS WITH TABLEAU
- Initial Visualizations with Tableau: Charts
- Creation of Maps, Dashboards and Histories
- Explore basic data connections with Tableau
- Creation of advanced charts
- Use of advanced data connections and data merging tools
- Additional features of Tableau: Parameters, Filters, Calculations ..
- Advanced Dashboards

SESSION 9 (FACE TO FACE)

INTRODUCTION TO BUSINESS ANALYTICS
- Supervised vs. Unsupervised techniques
- CRISP-DM model
- Main algorithms to be covered
- Introduction to Business Analytics platform used
- Data Manipulation & Feature Engineering

SESSIONS 10 - 11
Going a little farther than the classical descriptive statistical analysis: EXPLORING SIMPLE RELATIONSHIPS BETWEEN TWO VARIABLES

- Basic Ideas about bivariate analysis
- A simple guide to explore bivariate relationships. The rule of the 3 stages: graphs, then numbers then (maybe) statistical inference
- How to measure relations between categorical variables. Using graphs to explore categorical relationship. 2 way Crosstabs. Statistics: Chi-square test. The risk of Crosstab analysis (Simpson paradox)
- How to measure relations between continuous and categorical variables Using graphs to explore categorical – metrical relationship. Explore procedure. Statistics: Independent and paired samples T-test. ANOVA
- How to measure relations between continuous variables. Using graphs to explore relationship: scatterplots. What can we learn from simple linear correlation?

SESSIONS 12 - 14

How to group customers in marketing valuable segments. NON guided segmentation analysis: CLUSTER ANALYSIS

- CLUSTER ANALYSIS Basics:
  - Definition: what a Cluster Analysis is Understanding
  - Why / when Cluster Analysis is useful
  - How to evaluate a CLUSTER solution
- Understanding the basic technical background of a Cluster Analysis Hierarchical Clustering:
  - When to use hierarchical clustering Technical background
  - Procedure setting and interpretation
  - More about distance and agglomeration procedures
  - Practical exercises
- K-means and Two-Step Clustering:When to use either of both:
  - A brief note about both algorithms
  - Practical exercises

SESSIONS 15 - 17

PRINCIPAL COMPONENTS ANALYSIS

- Why PCA is so interesting: uses of Factor/PCA analysis.
- Dimensionality reduction and measurement of unobserved “factors”
- Understanding PCA analysis within a
  - Setting a PCA analysis
  - Reading a PCA results
  - Basic concepts (communality, explained variance, load matrix, ....)
  - How to better interpret the outcomes: Rotations
- Additional Practice

SESSIONS 18 - 20
CLASSIFICATION TECHNIQUES: How to build-up predictive modeling for categorical variables:
TREES Analysis
- Advantages of Trees analysis for business oriented analysis
- A brief introduction to Trees machinery: How do they work (CHAID and CRT explained)
- Using the platform for growing a TREE
  - Basic Settings
  - Interpreting a TREE output (What to look for. Key concepts: Gains, Response, Index,...)
  - Interpreting a TREE prediction (how to handle classification outputs)

SESSIONS 21 - 23
ASSOCIATION ANALYSIS: Events that occur together
- Introduction to Association. What’s the benefit behind. Basket Market analysis.
- Define Rule and Measure Formats: Antecedent, Consequent, Support, Rule Support & Confidence.
- Identify association Algorithm. Examine the APRIORI algorithm to generate rules
- Advanced Terms & concepts to improve the use of association algorithms: Lift, Confidence Difference and Confidence Ratio.
- Practical Examples of an Association Model

SESSIONS 24 - 26
How to really explore casual relationships? MULTIVARIATE REGRESSION ANALYSIS
- Introductory notes and basic definitions:
  - What is a regression analysis?
  - How a linear regression looks like?
  - What does it mean to estimate coefficients / parameters?
- How to avoid common regression mistakes
- LEARNING BY DOING: how to run (and recognize) a good regression analysis?

SESSIONS 27 - 28
INTRODUCTION TO ARTIFICIAL INTELLIGENCE
- What has been AI evolution from beginning?
- Main current areas: Chat bots, Image Recognition, Robotics, Artificial vision, language recognition
- Business examples for every area
- Main research fields

SESSION 29
REVIEW SESSION
Practical session in which we will test and review main concepts and practices seen along the course
BIBLIOGRAPHY
The program is divided in sections, grouping 3-4 classes for every technique. For each of them, ad-hoc documents will be uploaded and they will include:

- Information about the business concept related to each technique
- A theoretical explanation about the statistic technique
- Guided examples on how to perform each analysis and how to interpret it
- Optional additional exercises to further practice

Apart from the above, some very well-known textbooks and manuals might also be of interest for students, such as:

- *Data Science for Business*, Foster provost & Tom Fawcett. O’REILLY
- *Applied Predictive Modeling*, Max Kuhn & Kjell Johnson. Springer

The above manuals are not compulsory references.

EVALUATION CRITERIA

---

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Percentage</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation</td>
<td>10 %</td>
<td>See guide below</td>
</tr>
<tr>
<td>Individual Work</td>
<td>20 %</td>
<td>See criteria below</td>
</tr>
<tr>
<td>Workgroups</td>
<td>30 %</td>
<td>See criteria below</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40 %</td>
<td>See criteria below</td>
</tr>
</tbody>
</table>

CLASS PARTICIPATION
The ideal class performance will be characterized by the person who:

- Plays an active role in discussions, common activities and dialogue, providing value to the other fellows, etc
- Is fully prepared (in advance) at every class session.
- KEY RULE: Class participation is not measured by the amount of contributions that alumni do, but the quality of them

INDIVIDUAL WORK
Intermediate brief on line tests will be given to students at the end of each program section. Each test will score 0-100. Every student will be allowed to drop the worst score of these tests before computing the average. Some practice work will also be required.

WORKGROUPS
Every student should be able to work, in groups, on a set of selected analytical exercises prepared by the professor for each one of the major subjects. Every group should prepare a simple Executive Summary with no more than 3 pages and, in a different style, a technical annex supporting conclusions. Every single exercise will be evaluated with a numerical mark according to the following ideal criteria for the maximum grade:
- **Ideal Executive Summary**: Clear, concise, major points emphasized, well organized, clear recommendations, strong conclusion or call for action. Points easy to follow.
- **Ideal Technical Annex**: Displayed an excellent understanding of the technique. Showed excellent mastery of content, application and implications. Excellent research depth.

**FINAL EXERCISE**

The final exercise will include both a theoretical and a practical part. It would be evaluated with a numerical mark from 0 to 100.

**PROFESSOR BIO**

Professor: **ALVARO JOSÉ MÉNDEZ LÓPEZ**

E-mail: ajmendez@faculty.ie.edu

Professor: **ALVARO JOSÉ MÉNDEZ LÓPEZ**

+25 years experience in Analytics (Coca-Cola, SPSS & IBM)

+ 25 years experience as BI, Statistics, Econometrics & Machine Learning Professor
   (in both private and public institutions)

Founder and CEO of Advanced Analytics company
Co-founder of Demoscopic company

Doctor Candidate in Econometrics and Quantitative Analysis, UAM, Madrid, Spain.

Microsoft Certified Solution Developer

**OTHER INFORMATION**

Personal e-mail: alvaroj@gmail.com

**CODE OF CONDUCT IN CLASS**

1. **Be on time**: Students arriving more than 5 minutes late will be marked as “Absent”.
   Only students that notify in advance in writing that they will be late for a specific session may be granted an exception (at the discretion of the professor).

2. **If applicable, bring your name card and strictly follow the seating chart.** It helps faculty members and fellow students learn your names.

3. **Do not leave the room during the lecture**: Students are not allowed to leave the room during lectures. If a student leaves the room during lectures, he/she will not be allowed to re-enter and, therefore, will be marked as “Absent”.
   Only students that notify that they have a special reason to leave the session early will be granted an exception (at the discretion of the professor).

4. **Do not engage in side conversation**. As a sign of respect toward the person presenting the lecture (the teacher as well as fellow students), side conversations are not allowed. If you have a question, raise your hand and ask it. If you do not want to ask it during the lecture, feel free to approach your teacher after class.
If a student is disrupting the flow of the lecture, he/she will be asked to leave the classroom and, consequently, will be marked as “Absent”.

5. **Use your laptop for course-related purposes only.** The use of laptops during lectures must be authorized by the professor. The use of Social Media or accessing any type of content not related to the lecture is penalized. The student will be asked to leave the room and, consequently, will be marked as “Absent”.

6. **No cellular phones:** IE University implements a “Phone-free Classroom” policy and, therefore, the use of phones, tablets, etc. is forbidden inside the classroom. Failing to abide by this rule entails expulsion from the room and will be counted as one absence.

7. **Escalation policy: 1/3/5.** Items 4, 5, and 6 above entail expulsion from the classroom and the consequent marking of the student as “Absent.” IE University implements an “escalation policy”: The first time a student is asked to leave the room for disciplinary reasons (as per items 4, 5, and 6 above), the student will incur one absence, the second time it will count as three absences, and from the third time onward, any expulsion from the classroom due to disciplinary issues will entail 5 absences.