Session 1:
Levels of Corporate Decision

Anàlisi de Dades i Explotació de la Informació

Grau d'Enginyeria Informatica.

Information System track

Prof. Lídia Montero

lidia.montero@upc.edu

FIB

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Departament d'Estadística i Investigació Operativa

Levels of corporate decision

A company is a business organization aiming at creating value

Level	Decisions
Operational level	The invoice isn't accurate Restock inventory The mean daily production is erratic fluctuating The response answer of a web query is expelling potential costumers Offer credit to costumers according the rules Launch a promotion
Middle management	Design a commercial campaign Develop a departmental budget Rengineer the manufacturing process Design a new corporate web
Seniormanagement	Approve capital budget. Enter in a new (or exit) market. Decide long term goals

Decision taking in a firm: All levels of management need to take decisions

Levels of corporate decision

senior management

middle management
business areas

production and service workers
business processes

- All levels of management need to take decisions

- Many times in a “context of uncertainty”

Major Business Areas

FUNCTIONAL BUSINESS AREA	PURPOSE
Sales and Marketing	Selling the organization's products and services
Manufacturing, production and delivering	Producing and delivering products and services
Finance and accounting	Managing the organization's financial assets and maintaining the organization's financial records
Human ressources	Attracting, developing and maintaining the organization's labor force, maintaining employee records

Business Processes

- Each Functional Area generates its own business processes: assembling a product, identifying costumers, paying creditors, hiring employees, …
- All business processes generates a path of data of the process. Transactions processing systems provide the tool to gather all theses elementary activities and transactions of the organization (sales, invoices, cash deposits, payroll, inventory, shipping, produced items, complaints, web visits, e-selling …) Daily routine Transactional Systems needed at Operational Level
- Transactional Processes are major producers of information for the organization

Decision Support Systems to monitor sectorial areas of business (weekly, monthly and yearly reports), helping the process decision making and predict future performance

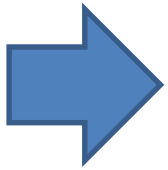
Decision Support Systems

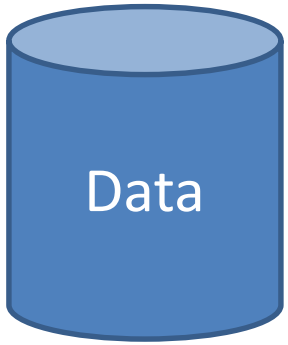
Level	decision	Data form	Tools and statistical level
Operational level	Structu-red. Short term	Files, BBDD	Data coming from TPS. Control charts Routinely hypothesis tests
Middle Manage-ment		BBDD, DW (OLAP)	Decision Support Systems: To monitor business processes. Excel spreadsheet (weekly, monthly and yearly reports) Sophisticated Data Mining tools: profiling, association of events, clustering, classification, prediction, forecasting, …. Web Mining, Text mining To discover niches in the market. To better customize the product or service. To avoid attrition. Business Intelligence systems
Senior Ma-nagement	Unstruc-tured. Long term	BBDD, DW (OLAP), external	Executive Information Systems Balanced scoreboard: Joint monitoring of Financial, Processes, Costumers and Human capital. Internal and external info. Tangible and intangible key indicators.

From Business Process to Data

Business processes are concrete workflows of material, information and knowledge – sets of activities. Each business process generates its own application (or part of it).

The output of the application is stored in DDBB (or files).





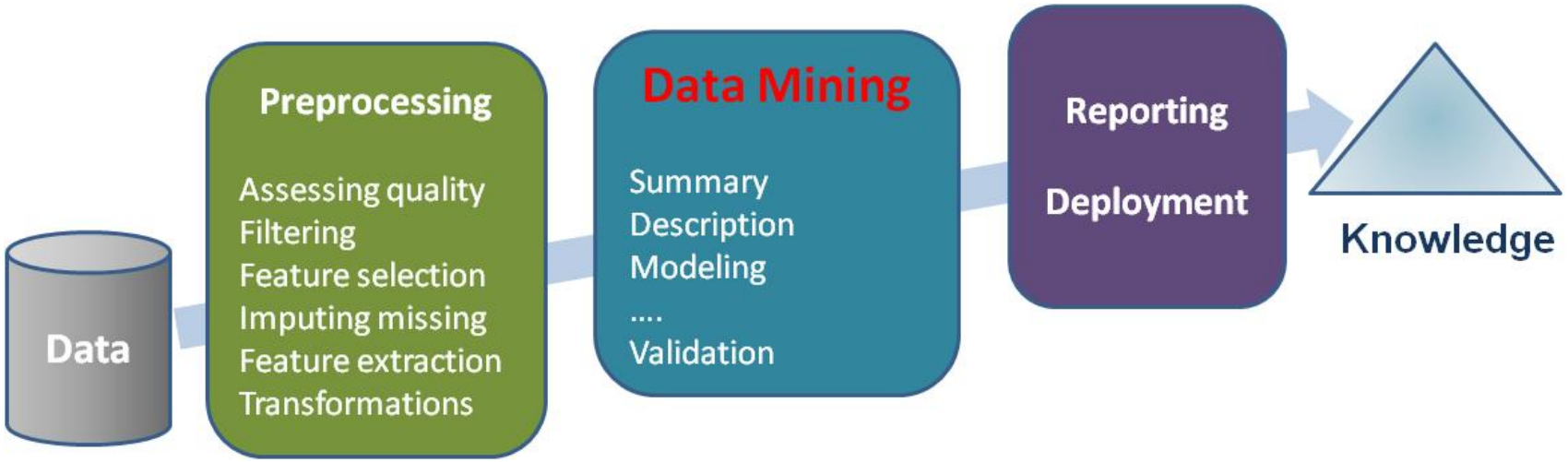
- All levels of management need to take decisions
- Many times in a “context of uncertainty”
- But supported by “experimental data”

The value of Data

Paradigm of the information era:

Data is the new driving force of businesses and governments.

Data is a key value for organizations



Program

Unit		Weeks
1	Levels of corporate decision.	1
2	Quality of data.	1
3	Multivariate Analysis: Profiling , Principal component analysis, Correspondence analysis, Multiple correspondence analysis, Clustering	6
4	Statistical Modeling. Prediction	5
5	Quality Control	2

Planning of the course

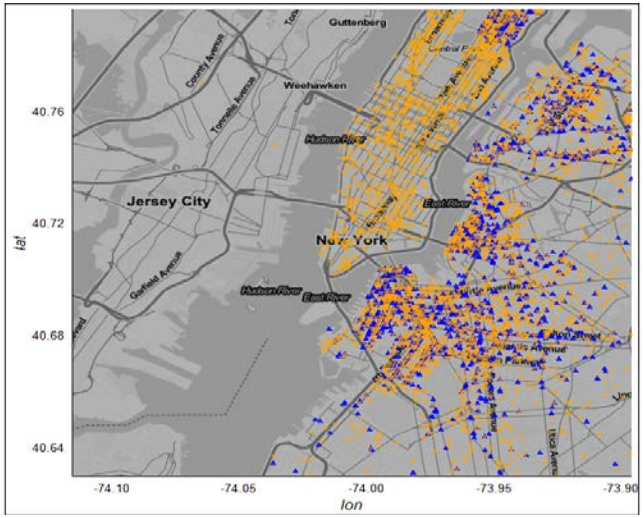
ADEI SCHEDULE. Course 2020-21. Term 1			
		Wednesday 17-19h	Friday 15-17h
Week		Laboratory - Online	Theory -A6202
1	Presentation of Subject	16-sep	18-sep
	Levels of Corporate Decision		
2	Data Quality	23-sep	25-sep
	Statistical Modeling: Imputation		
3	Profiling	30-sep	02-oct
4	Multivariate Analysis : PCA	07-oct	09-oct
5	Multivariate Analysis : PCA	14-oct	16-oct
6	Multivariate Analysis: PCA/Clustering	21-oct	23-oct
7	Multivariate Analysis: MCA	28-oct	30-oct
8	Multivariate Analysis: MCA/Clustering	04-nov	10-nov
			QUIZ 1
9	Statistical Modeling: Numeric Target and Covariates	11-nov	13-nov
10	Statistical Modeling: Numeric Target and Covariates	18-nov	20-nov
11	Statistical Modeling: Numeric Target - Diagnostics	25-nov	27-nov
12	Statistical Modeling: Numeric Target /	02-dic	04-dic
	Factors and Covariates as Explanatory variables		
13	Statistical Modeling - Binary Target	09-dic	11-dic
	Factors and Covariates as Explanatory variables		
14	Statistical Modeling - Binary Target - Diagnostics	16-dic	18-dic
			QUIZ 2
15			
		Final Exam	7 Gen 2021
		11:30 - 14:30	
	Teaching		
	Lídia Montero Teaching		
	Lídia Montero Teaching		
	Assessment:		
	2 Quiz (Part 1 i 2)	30%	
	5 Deliverables: 3 Part 1 (ACP, ACM) I 2 Part 2 (MR, BR)	30%	
	1 Final Presentation (Oral). Conclussions in English	40%	
	Assignments in groups of 2 - Quiz individually		

Reports on Data description and Statistical Modeling:

Case study: NY Cabs data – Green Taxi Trip records

- **The course project is concerned with Multivariant Data Analysis and model building for response variables: Y- Total Amount (Numeric Target) and binary factor Y.bin-'TipIsGiven' (Binary Target) for trips in the green Taxi Trip Records.**
- **A random sample containing 5000 registers of green taxi records has to be retained by each group.**

Practical Deliverables	Deadline
3 Reports on Data cleaning, feature selection and profiling, multivariant analysis and numeric and binary targets modeling (limited to 50 pages each)	Over the course
Report on Multivariant Analysis and Statistical Modeling- Presentation of Executive Summary of the Case Study	End of Course



Carry on the Case Study

Case study: groups two students

The slide presentation will be presented orally, with an executive summary and/or conclusions presented in English.

The presentation will be done the same final exam day

Evaluation

The evaluation of the course integrates the three phases of learning process: knowledge, skills and competencies.

•**The knowledge is assessed by two quizzes**, *in the middle and last week of the course. If you fail this exam, students may have a final resit. (score T).*

•**The skills assessed from the delivery of several deliverables (3) related to the course**. *Each of the blocks involve a practice that students will perform by groups of 2 (Score L) and post in the selected Atenea tasks. A Final Deliverable should be also posted.*

•**The case study will be evaluated based on the oral presentation (score P).** *In the presentation of case study that generic skills will be assessed. In any case, the presentation of the case study is compulsory.*

The final grade will obtained weighing the three scores: Final Mark = 0.4P + 0.3L + 0.3T.

Generic skills will be assessed on the scale: Fail, Pass, Good and Very good (D,C,B and A).

Generic competences

- English

  Evaluated from the

  – Presentation of the executive summary (5 min.)
  – Slides of presentation

- Reasoning

  Evaluated from the answers to the questions raised by the oral presentation of the Case Study.

To assess the competence on English, it will be required to have written in English the report on the Case Study, moreover at the beginning of the presentation, the student must do an outline of the work in English as well. Regarding the reasoning competence, it will be assessed from the answers given to the presentation of the Case Study.

Software

- The software to be used during the course will be R and RStudio.
- Each block will use its specific packages and functions.
- ***cran.r-project.org/***
- https://www.r-project.org/nosvn/conferences/useR-2013/Tutorials/Kuhn.html
- **A Complete Tutorial to learn Data Science in R from Scratch**
- https://www.analyticsvidhya.com/blog/2016/02/complete-tutorial-learn-data-science-scratch/

Recommended books

✓Francois Husson, Sebastien Le, Jérôme Pagès (2011) **Exploratory Multivariate Analysis by Example Using R.** Chapman & Hall/CRC Computer Science & Data Analysis

✓Fox, J. ***Applied Regression Analysis and Generalized Linear Models***. Sage Publications, Edition 2015.

✓Fox and Weisberg ***An R Companion to Applied Regression***. Sage Publications, Edition 2011.

✓Wickham, H. ***ggplot2: Elegant Graphics for Data Analysis***. Springer New York, 2009.

✓John, Peter W.M. , Statistical Methods in Engineering and Quality Assurance , Wiley-Interscience , 1990 , ISBN:0471829862. http://ebookee.org/Statistical-Methods-in-Engineering-and-Quality-Assurance_225786.html

✓Maindonald, J and Braun, John , Data Analysis and Graphics Using R , Cambridge University Press , 2007 , ISBN:9780521861168. http://cran.r-project.org/doc/contrib/usingR.pdf

✓Aluja Banet, Tomas y Morineau, Alain , Aprender de los Datos: El Análisis de Componentes Principales , EUB , 1999 , ISBN:84-8312-022-4.