

f.Introduction to Advanced Business Analytics

Information are at the base.

- Descriptive Analytics: aims to understand what happened in the past
- Diagnostic Analytics:
- Predictive Analytics: tries to predict what will happen
- Prescriptive Analytics:

Data and information aren't the same thing: data is raw and unprocessed facts, it is not useful in this state, information is processed and organized data.

The cycle between data and information is refined by each passage.

Business intelligence is based on insight (capacity of having a deep understanding of someone/something). Now we can move to performance manager and then reach the advanced analytics part.

Technological knowledge, used to inform the client to which technology is the best for him. It is needed but is half of the equation, you have to have soft-skill, business knowledge.

A business intelligence platform gives ways to see data, it focuses on Descriptive and Diagnostic analytics. Its goal is to empower organization and give them an advance.

Business intelligence is a set of methodologies, processes, architectures and technologies that transform raw data into meaningful and useful information used to enable more effective strategic tactical decisions.

Data visualization is fundamental for the understanding of the data and for enabling the meaning in them. CRV is the metric that tries to show the lifetime of a client and indicate what happens if we lose them.

A business intelligence architecture is composed of:

- A source as ERP, databases, external data, manual inputs, etc...
- Staging Layer: layer storing the raw data information
- Information model: we have to organize the data, create context/meaning to the data
- Data Visualization: let you see information about data.
- Self service: the users can build their own reports/can use the infrastructure to create new data

Where to spend time if you want to work in data analytics ecosystem

Hard Skill:

- Math and statistics: basic math/statistics cover the 80% of problems.
- ETL/ELT: example is Azure DataFactory
- Coding: python, SQL

- Data visualization and storytelling: DataBricks, Excel
Soft -Skill
- Critical thinking
- Business acumen
- Problem solving
- Collaboration

From data to value

Data Architectures

Describes how data is managed from collection through utilization

First pattern is ETL(Extract Transform Load).

Second pattern ELT(Extract Load Transform) we load the data in a data lake and then transform them.

Project methodologies

Define each initiative as a project. A project is composed by a set of task and you have to define a methodology/framework which define the design of the project).

- Waterfall project: the scope is defined by a blueprint that act as a reference to the stakeholder and contains all information related to the project(Everything outside is considered a change request). It is not uncommon that the customer is unsatisfied. The team is composed by a vertical team, there is a project manager who coordinate all the people.
- Agile: it is defined an initial scope, it gives a framework to incorporate the user feedback in the development of the project. The team is self organized and don't require a vertical project manager. Three main actors: developers(create a plan for the sprint backlog, adapt their plan each day toward the sprint goal), scrum planner(coaching the team members in self-managements), product owner(communicate the product goal, clearly communicate the product backlog items)

These two approaches can be combined in a hybrid approach and we pick some drivers. We want the low technology uncertainty of the waterfall approach and the capability of adapt to feedback.

You have to develop the art of taking requirements to have the maximum clearness of the requirement and restrain of the project. You have to understand the business needs and intended outcomes of an analytics project.

Garbage IN Garbage OUT

It is better to have a okay model but to have good data in it. Data quality refers to the overall reliability, accuracy and completeness of data. In a company there could be specific roles related to data quality, backend have to implements some type of data quality rule(technical) but

also from the business part of the company there has to be data quality.

Data Observability:

Data Modelling

It is the process of creating a structured representation of data to understand its relationships, constraints and organization. It involves defining entities, databases,

Data dictionary: guide/references to all the data in the data model

A semantic model is a Logical Layer that contains the relations, transformation and calculations between data sources needed to create reports and dashboards. The direction of the arrows indicate the direction of the relation

Dimensions: describes the entities in the model, attribute to slice the fact

Facts: stores information regarding an event or an observation

We can change how data is organized with Normalization(Different tables for the various facts) and Denormalization(One huge table with all the attributes). We use Normalization when we talk about normal day works as it reduces redundancy and simplifies updates instead we use denormalization in data lake where we want to avoid relationships and want fast query. We can use denormalization to a certain degree also in normalization as it can help to reduce complexity.

Data Visualization

Interactive exploration and graphical representation of data.