

# **Progetto**

# **Introduzione alla Data Analysis**

# **Data Driven, cosa significa?**

**Con il termine data-driven, letteralmente “guidato dai dati”, si intende la costruzione di processi decisionali sempre di più basati sulle informazioni estraibili dai dati.**

## **I vantaggi "DATA DRIVEN" per un azienda**

- **Rafforzamento del vantaggio competitivo e crescita del fatturato**
- **Migliora l'acquisizione, la fiducia e la fidelizzazione dei clienti**
- **Rivoluziona le pratiche dei clienti**
- **Riduce i costi**



# Quali sono i ruoli utili?

**Data Analyst:** è colui che esplora, analizza e interpreta i dati, con l'obiettivo di estrapolare informazioni utili al processo decisionale.

**Data Engineer:** gestisce le fasi di raccolta, processamento e integrazioni dei dati. Rende i dati disponibili per le analisi nel giusto formato.

## Cosa fa il Data Engineer?

- Sviluppa, costruisce, testa e mantiene architetture di pipeline di data
- Allineare l'architettura dei dati ai requisiti aziendali
- Identificare modi per migliorare l'affidabilità, l'efficienza e la qualità dei dati
- Preparare i dati per la modellazione predittiva e prescrittiva
- Garantire la conformità con la governance dei dati e le politiche di sicurezza

## Cosa fa il Data Analyst?

- Comprende le tipologie di ricerche da svolgere
- Si assicura che le informazioni che ha non siano distorte
- Organizza i dati e crea dei cluster
- Legge i cluster con metodi statistici per ricavare info utili per il business aziendale
- Comunica i risultati tramite documenti scritti o software dedicati

# Aziende esempio di data driven e best practice

The logo for Booking.com, featuring the word "Booking" in white and ".com" in blue, set against a dark blue rectangular background.

**Utilizzato maggiormente dal mercato europeo per prenotazioni nelle strutture ricettive o case vacanze.**



**Utilizzato dal mercato prevalentemente americano, anch'esso "Online Travel Agent", con esso è possibile prenotare dal biglietto aereo, a una struttura ricettiva oppure una crociera.**

- **Formatta i set di dati in anticipo per l'utilizzo dei Big Data.**
- **Crea una fonte di analisi della verità.**
- **Bilancia i tuoi dati e ottimizza le partizioni.**
- **Checkpoint (salvataggio) qualunque set di dati verrà riutilizzato o avrà più dipendenze logiche a valle.**
- **Quando usano Spark, si usa Scala per tutto ciò che assomiglia a un ETL.**
- **Prendere in considerazione l'utilizzo o l'estensione delle classi Transformer in Spark PipelineML per distribuire rapidamente i tuoi ETL.**

# **In che modo si traggono vantaggi?**

**Si possono trarre vantaggi facendo pubblicità con strumenti di marketing automazione, così da segmentare le fasce di clienti e aumentare il livello di conversazioni anche in target meno conosciuti, per un incremento finale di vendite.**

## **Veicoli elettrici**



**Introdurre bus elettrici a scopo di ridurre l'impatto ambientale.**

## **Ridurre l'uso della carta**



**Alcune strutture propongono l'uso di app per leggere quotidiani nazionali e non, a scopo di ridurre l'uso e il circolo di carta.**

## **Mobilità ciclistica**



**Molte strutture ricettive e non, offrono la possibilità di noleggiare biciclette ed e-bike a scopo di effettuare spostamenti in modo sostenibile.**

# **Il viaggio dei dati**

## **Fase 1**

**Generare e memorizzare i  
dati allo stato grezzo**

## **Fase 2**

**ETL (estrai,trasforma,carica)**

**Preparare i dati, creare un ambiente di staging e  
trasformare i dati, pronti per l'analisi**

## **Fase 4**

**Creazione di dashboard o widgets**

## **Fase 3**

**Una volta archiviati i dati, i data engineer  
possono estrarre dal data warehouse o  
dal data lake per creare tabelle. Creano  
relazioni con i dati e le modellano.**

## **Fase 5**

**Incorporamento delle analisi nei  
prodotti e nei servizi clienti**