

# Capitolo 1: BI come strumento di supporto alle decisioni

## *1.1: Business Intelligence*

Nel mondo moderno il valore dell'informazione sale incessantemente in quanto è necessario per analizzare l'operato dell'azienda, monitorare eventuali progressi verso obiettivi previsti (vendite previste, costi stimati ecc.) ed organizzare attività aziendali in maniera efficiente. Tuttavia, il reperimento dell'informazione è un processo non banale e laborioso quando si devono trattare moli di dati di dimensioni enormi quali possono essere i dati di vendita annuali di un'azienda o i costi di produzione relativi ad uno specifico prodotto. Inoltre, le aziende necessitano di tali informazioni in tempi brevi, spesso organizzate in modo personalizzato e di conservare i dati in modo centralizzato e non volatile.

Con *Business Intelligence* si intende l'insieme degli strumenti e dei procedimenti per selezionare, aggregare, correggere e trasformare i dati grezzi in conoscenza utile per supportare processi decisionali. Qui di seguito sono riportate alcune delle varie tecniche di Business Intelligence quali:

- Data Mining
- Reporting
- Analisi Statistica
- Analisi Descrittiva
- Query
- Metriche e Benchmarking delle prestazioni
- Ecc...

## 1.2: Data Warehousing

Il Data Warehousing è *una collezione di metodi, tecnologie e strumenti di ausilio al knowledge worker (dirigente, amministratore, gestore, analista) per condurre analisi dei dati finalizzate all'attuazione dei processi decisionali e al miglioramento del patrimonio informativo.* [1]

È un sistema di data management di supporto alla BI che permette di centralizzare e consolidare dati da diverse origini, mantenere grandi quantità di dati storici e su cui possono essere eseguite query e analisi. Il processo di Data Warehousing offre i seguenti vantaggi:

- Accessibilità a utenti con scarse conoscenze informatiche
- Integrazione dei dati sulla base di un modello standard dell'impresa
- Flessibilità di interrogazione per ottenere il massimo dalle informazioni presenti
- Sintesi per analisi mirate ed efficaci
- Rappresentazione multidimensionali per offrire all'utente una visione intuitiva ed efficacemente manipolabile delle informazioni
- Correttezza e completezza dei dati integrati

Il fulcro di questo processo è il Data Warehouse, *una collezione di dati di supporto per il processo decisionale che presenta le seguenti caratteristiche: orientato ai soggetti, integrato e consistente, rappresentativa dell'evoluzione temporale, non volatile.* [2]

### 1.2.1: Orientato ai soggetti

A differenza dei database relazionali, progettati con il focus sulle applicazioni che andranno ad utilizzarli, il Data Warehouse è progettato in vista degli utenti finali che andranno ad usufruirne.

[1]: Slide 9 Corso Business Intelligence, S.Rizzi

[2]: Slide 12 Corso Business Intelligence, S.Rizzi

### *1.2.2: Integrato e consistente*

Appoggiandosi a più fonti di dati eterogenee provenienti da basi di dati relazionali, sistemi informativi esterni o addirittura documenti non strutturati, è necessario unificare e assicurarsi che questi dati siano uniformi e renderli tali qualora non lo fossero.

Questo è il ruolo degli strumenti di ETL (Extraction, Transformation and Loading).



Figura 1: funzioni degli strumenti di ETL

### *1.2.3: Rappresentativa dell'evoluzione temporale*

A differenza dei DB relazionali, in un DW il tempo è parte delle chiavi e i dati contenuti non possono essere aggiornati o sovrascritti.

### *1.2.4: Non volatile*

In un DB relazionale i dati sono soggetti ad operazioni di Insert, Update e Delete, il che rende i dati non persistenti. In un DW i dati vengono caricati una volta e non verranno più modificati o rimossi. Il problema si sposta dalla gestione delle transazioni al query-throughput.

### *1.3: Architetture Data Warehouse*

Affinché un'architettura di Data Warehouse funzioni, sono necessari dei requisiti:

- Separazione: l'elaborazione analitica e transizionale devono essere mantenute il più possibile separate
- Scalabilità: l'architettura deve poter essere facilmente ridimensionata a fronte della crescita nel tempo dei volumi di dati e del numero di utenti da gestire
- Estendibilità: deve essere possibile adottare nuove tecnologie e applicazioni senza riprogettare integralmente il sistema
- Sicurezza: controllo sugli accessi
- Amministrabilità: è necessario limitare la complessità dell'attività amministrativa

A fronte di questi requisiti ci sono varie implementazioni, differenziate dal numero di livelli di cui si compongono. Qui di seguito è descritta l'architettura a due livelli.

### 1.3.1 Architettura a due livelli

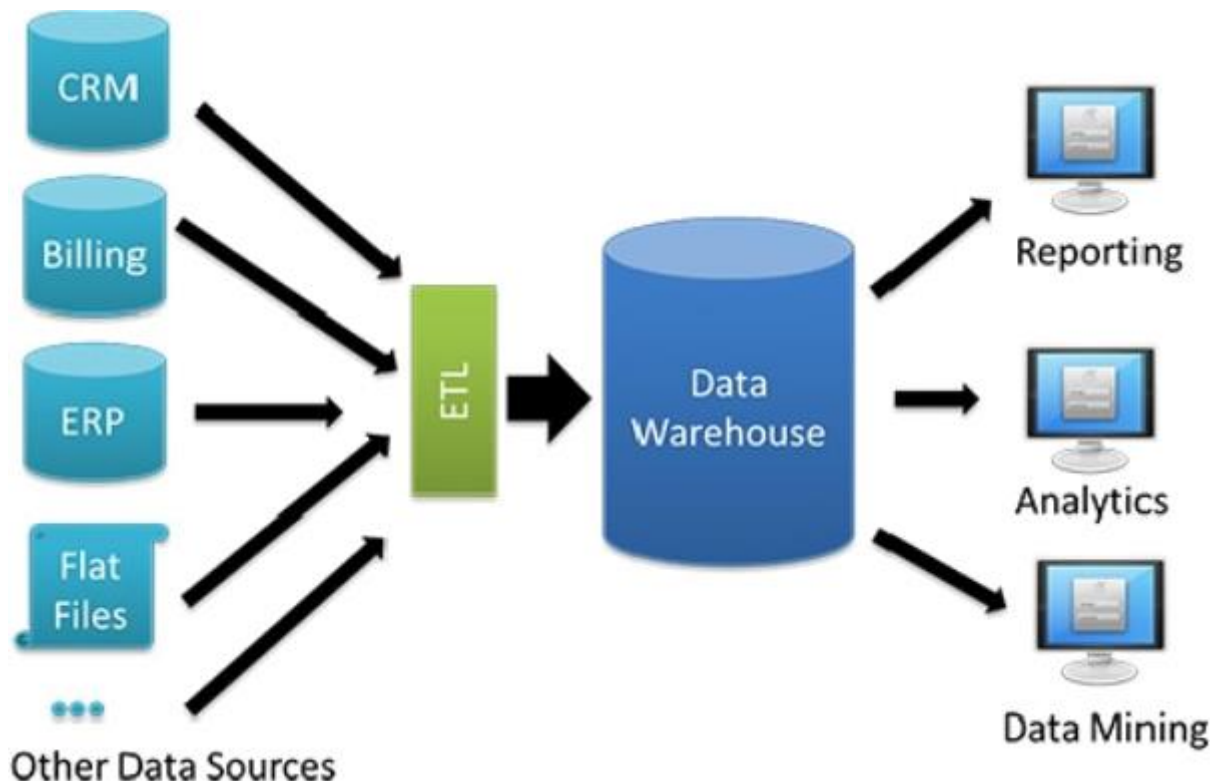


Figura 2: Architettura a due livelli

I due livelli che compongono questa architettura sono il livello sorgenti, composto da tutte le fonti di dati, e il livello Data Warehouse, dove il DW è composto da ulteriori “sotto DW” chiamati Data Mart, ossia *sottoinsiemi o aggregazioni di parte dei dati presenti nel DW primario*. Essi possono essere dipendenti, se usati per suddividere un DW che sarebbe altrimenti troppo grande per essere gestito agilmente e migliorare le prestazioni, o indipendenti, se alimentato direttamente da una fonte dati e sfruttato per facilitare la fase progettuale rendendo più difficile lo schema di accesso ai dati.

L'utilizzo di un'architettura a due livelli comporta numerosi vantaggi quali:

- Informazioni di buona qualità continuamente disponibili a livello di DW anche quando è temporaneamente precluso l'accesso alle sorgenti.
- Le interrogazioni analitiche effettuate sul DW non interferiscono con la gestione delle transazioni a livello operativo.
- L'organizzazione logica del DW è basata su un modello multidimensionale mentre le sorgenti offrono modelli relazionali o semi-strutturati.
- Discordanza temporale e di granularità tra sistemi OLTP, che trattano dati correnti e al massimo livello di dettaglio, e sistemi OLAP che operano su dati storici e di sintesi.
- A livello del warehouse è possibile impiegare tecniche specifiche per ottimizzare le prestazioni per applicazioni di analisi e reportistica

## Capitolo 2: Descrizione Software utilizzati

In questo capitolo si tratteranno i software che sono stati impiegati per l'elaborazione del progetto di tesi.

Segue un elenco dei vari software per poi approfondire ognuno di essi nello specifico:

- *SQL Server Management Studio (SSMS)*: è un ambiente integrato per la gestione di qualsiasi infrastruttura SQL. SSMS integra un'ampia gamma di strumenti grafici con numerosi editor di script avanzati per offrire accesso a SQL Server per gli sviluppatori e gli amministratori di database qualsiasi sia il livello di competenza.
- *SQL Analysis Services (SSAS)*: è un motore dati analitici usato nel supporto decisionale e nell'analisi aziendale. Fornisce funzionalità del modello di dati semantico di livello aziendale per le applicazioni di business intelligence (BI), di analisi dei dati e di Reporting, ad esempio Power BI, Excel, Reporting Services e altri strumenti di visualizzazione dei dati.
- *Visual Studio Analysis Services projects*: è un'estensione di Visual Studio che fornisce la possibilità di progettare e costruire modelli tabulari e multidimensionali dispiegati poi in SQL Server Analysis Services, Power BI o Azure Analysis Services.
- *Visual Studio Reporting Services*: è un'estensione di Visual Studio che fornisce la possibilità di progettare e creare (anche tramite wizard) report professionali.
- *SQL Server Reporting Services (SSRS)*: offre un set di servizi e strumenti locali per creare, distribuire e gestire report impaginati e per dispositivi mobili anche a partire da modelli SSAS.

Tutti i software impiegati sono proprietari Microsoft ed è stata utilizzata la distribuzione 2019 per Visual Studio e la 2017 per i servizi relativi alla suite SQL.

## *2.1 SQL Server Management Studio*



*Figura 3: logo di Microsoft SQL Server Management Studio*

Microsoft SQL Server Management Studio è uno degli ambienti di sviluppo e gestione per tipo di infrastruttura SQL più usati. Ciò è dovuto alla sua facilità d'uso per utenti meno esperti e dalla flessibilità e potenza se usato da mani esperte.

La peculiarità di SSMS è la possibilità di mantenere aperte più connessioni e navigare liberamente tra di esse mentre agisce in tempo reale su oggetti e tabelle.

L'ambiente include, oltre a tool grafici, editor testuali per scrivere script in vari linguaggi quali SQL, DAX, DMX, MDX ecc.

Siccome SSMS è stato pensato per essere affiancato a Visual Studio, un progetto sviluppato in tale ambiente gode di numerose estensioni e funzioni integrate per permettere agli sviluppatori di lavorare in maniera agile ed efficace.

## 2.2 SQL Analysis Services



*Figura 4: logo di Microsoft SQL Analysis Services*

Microsoft SQL Server Analysis Services è uno strumento del pacchetto Business Intelligence di SQL Server insieme a Reporting Services ed Integration Services e la sua funzione è offrire supporto per il data mining e OLAP (online analytical processing).

SSAS aggrega informazioni disseminati in più DataBase o tabelle in modelli tabulari o multidimensionali e supporta il linguaggio XML come DDL (data definition language) e MDX, LINQ, SQL (una parte), DMX e DAX come DML (data Manipulation Language).

## 2.3 Visual Studio Analysis Services projects

Estensione di Visual Studio che permette di creare un modello tabulare o multidimensionale a partire dai dati contenuti in DataBase. È inoltre possibile ritoccare, selezionare, ordinare e aggregare i dati prima di inserirli nel modello.

Dopodiché è possibile creare relazioni tra vari campi, gerarchie, misure sulle varie tabelle e dividere i dati in diverse partizioni o creare gruppi di utenti che potranno accedere a specifiche informazioni del modello.

Infine, il modello può essere distribuito su un server di SQL Analysis Services.



## *2.4 Visual Studio Reporting Services*

Estensione di Visual Studio che permette di creare dei report impaginati a partire da modelli tabulari o multidimensionali reperiti da un server di SQL Analysis Services.

Reporting Services consente di generare report anche tramite wizard e di gestire tutti gli aspetti di un report quali parametri (visibili o no all'utente finale), set di dati estratti dai modelli SSAS usando query scritte in SQL, DMX, DAX o MDX ed elementi grafici quali indicatori, grafici di ogni tipo.

Ultimato il report è possibile pubblicarlo su SQL Reporting Services o un'altra piattaforma di visualizzazione report.

## *2.5 SQL Server Reporting Services*



*Figura 5: logo di Microsoft SQL Server Reporting Services*

Microsoft SQL Server Reporting Services permette sia di visualizzare progetti di report creati esternamente sia di crearli. È possibile creare e gestire separatamente le origini dati, i set di dati e le pagine di report con impostazioni di sicurezza, piani di aggiornamento cache e simulazioni di esecuzione con calcolo dei tempi.

SSRS si presenta come un sito web molto intuitivo e strutturato come un filesystem dove l'utente seleziona la cartella contenente il report desiderato e visionarlo (è compito degli admin nascondere le parti non pensate per essere viste dagli utenti).