

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

---

SCUOLA DI INGEGNERIA E ARCHITETTURA

DIPARTIMENTO

INGEGNERIA INDUSTRIALE

TESI DI LAUREA

in

Business Intelligence

***La Business Intelligence per il monitoraggio delle vendite:  
il caso Ducati Motor Holding***

CANDIDATO

Lucchetta Linda

RELATORE

Chiar.mo Prof.  
Grandi Fabio

CORRELATORI

Rizzi Stefano  
Golfieri Francesco

Anno Accademico 2015/2016

Sessione II



# Sommario

INTRODUZIONE .....	1
CAPITOLO 1: INTRODUZIONE ALLA BUSINESS INTELLIGENCE.....	3
Inefficienze del sistema di reporting tradizionale .....	3
Verso la Business Intelligence .....	4
Definizione di Business Intelligence.....	5
Storia della Business Intelligence.....	6
La svolta del 1958 .....	7
Sviluppi ed evoluzioni fino al 1980 .....	8
Dal 1980 e il 1990 .....	8
Business Intelligence 1.0.....	9
Business Intelligence 2.0.....	9
Modern Day.....	10
Obiettivi della BI.....	11
Caratteristiche della BI.....	12
La struttura della BI.....	14
Extraction, Transformation and Loading (ETL).....	14
Data Warehouse.....	16
La Piramide Informativa della BI.....	22
Rappresentazione Multidimensionale .....	23
Reportistica.....	24
OLAP.....	26
Analisi What-If.....	28
Data Mining.....	29
Sintesi di un Ambiente di Data Warehousing .....	31
Top Business Intelligence Companies .....	32
VANTAGGI DELLA BI.....	35
Ruoli nella Business Intelligence e composizione dei team.....	35
Ruoli .....	36
Business Analytics .....	40

TERMINOLOGIA .....	41
CAPITOLO 2: BUSINESS INTELLIGENCE E BIG DATA NEL MERCATO DELL'AUTOMOTIVE.....	
Definizione di BIG DATA .....	45
BIG DATA all'interno delle aziende .....	47
Le 9 sorgenti dei BIG DATA .....	47
Alcuni tool di support per la gestione dei BIG DATA.....	48
Differenze con la Business Intelligence .....	48
Tipi di analisi effettuabili sui Big Data .....	49
Ultimi anni.....	51
BIG DATA e Business Intelligence nel mercato dell'automotive .....	52
Key Performance Indicator (KPI) .....	54
Importanza delle funzionalità della BI per un'azienda nell'automotive .....	56
Aumentare la visibilità sui dati e permettere il data-sharing .....	57
Creare utenti finali information-empowered .....	57
Migliorare l'efficienza di pianificazione grazie a sistemi real-time .....	58
Come i BIG DATA impattano le vendite nel settore dell'automotive? .....	59
L'impatto dei dati sulla customer retention.....	60
Come i dati possono migliorare la customer retention? .....	60
CAPITOLO 3: DUCATI MOTOR HOLDING .....	
Storia .....	62
Origini.....	62
Dopoguerra .....	63
Il controllo statale .....	64
Gli anni Cagiva.....	64
Anni recenti .....	64
2012 .....	65
2016 .....	65
Ducati oggi .....	65
Ducati Group .....	65
Vendite Ducati.....	67
Vendite Ducati Motor Holding per categoria di moto.....	67

Le Vendite delle Top 100 Moto.....	68
CAPITOLO 4: PROGETTO BUSINESS ANALYTICS.....	70
Dimensionamento dati ducati.....	71
Fonti dati Ducati.....	72
Conseguenze introduzione tool di Business Analytics.....	72
Architettura logica desiderata:.....	73
Selezione dei vendor .....	73
Parametri.....	73
Vendor valutati e prodotti proposti.....	74
Conclusioni.....	75
Fase di Analisi.....	77
Aree aziendali coinvolte .....	77
Assessment .....	78
Sviluppi successivi .....	79
Realizzazione roadmap di progetto .....	80
Fase 1 .....	80
Fase 2.....	86
Roadmap di progetto.....	90
Sales.....	90
Dimensioni di analisi: prodotto e mercato (1).....	91
Sell-in e sell-out: il mondo sales all'interno di Ducati (2-3).....	92
Claim.....	94
DCS- Dealer Communication System .....	94
CAPITOLO 5: REPORTING SALES La reportistica per il monitoraggio delle vendite e la sua applicazione in Ducati Motor Holding.....	97
Area sales .....	97
Overview reporting: .....	99
Come il reporting può aiutare un'impresa a vendere di più.....	100
Sei consigli per realizzare migliori dashboards per il monitoraggio delle vendite .....	103
Cinque grafici che ogni responsabile delle vendite dovrebbe avere a disposizione .....	105
The bullet.....	105
KPI map.....	106

I trend stagionali .....	108
Sales vs Profit scatter plot.....	109
The sparklines chart.....	110
Quattro grafici per l'analisi dei dati raccolti dai social media. ....	111
Slope chart for growth of followers/reach .....	111
Mappa dei click sui link regionali / globali .....	112
Grafico di dispersione (scatter plot) .....	113
Click-through rate in un grafico boxplot .....	113
Reportistica App WDW .....	114
Reportistica in Ducati prima e dopo la BI.....	117
Data quality assurance .....	118
Prima della BI.....	122
Le novità della Business Intelligence .....	125
Front end.....	131
Bibliografia.....	138

# INTRODUZIONE

“La Business Intelligence per il monitoraggio delle vendite: il caso Ducati Motor Holding”.

L’obiettivo di questa tesi è quello di illustrare cos’è la Business Intelligence e di mostrare i cambiamenti verificatisi in Ducati Motor Holding, in seguito alla sua adozione, in termini di realizzazione di report e dashboard per il monitoraggio delle vendite.

L’elaborato inizia con una panoramica generale sulla storia e gli utilizzi della Business Intelligence nella quale vengono toccati i principali fondamenti teorici: Data Warehouse, data mining, analisi what-if, rappresentazione multidimensionale dei dati, costruzione del team di BI eccetera.

Ho deciso poi di proseguire mediante un focus sui Big Data e di convogliare l’attenzione sul loro utilizzo e utilità nel settore dell’automotive (inteso nella sua accezione più generica e cioè non solo come mercato delle auto, ma anche delle moto), portando in questo modo ad un naturale collegamento con la realtà Ducati.

Si apre così una breve overview sull’azienda descrivendone la storia, la struttura commerciale attraverso la quale vengono gestite le vendite e la gamma dei prodotti.

Dal quarto capitolo si entra nel vivo dell’argomento: la Business Intelligence in Ducati. Si inizia descrivendo le fasi che hanno fino ad ora caratterizzato il progetto di Business Analytics di cui faccio parte in prima persona per poi concentrarsi, a livello prima teorico e poi pratico, sul reporting sales e cioè sulla reportistica basata, per l’appunto, sul monitoraggio delle vendite.





# **CAPITOLO 1:**

## **INTRODUZIONE ALLA BUSINESS INTELLIGENCE**

Una parte importante del patrimonio di conoscenza di un'azienda è da sempre contenuta in forma strutturata nelle basi dati delle applicazioni gestionali. Conseguentemente, prima dell'introduzione dei sistemi di Data Warehousing, le informazioni venivano rese disponibili sotto forma di report statici, onerosi da produrre e spesso obsoleti una volta arrivati nelle mani degli utenti.

Di conseguenza le organizzazioni si basavano più sull'intuito dei manager che su dati quantitativi per valutare l'andamento dell'impresa e prendere decisioni importanti.

### **Inefficienze del sistema di reporting tradizionale**

L'era digitale (chiamata anche digital age, information age o computer age) è definita come il periodo di tempo dal 1970 caratterizzato dall'introduzione dei personal computer e delle tecnologie connesse alla trasmissione libera e veloce dei dati. A partire da questo periodo, le aziende il cui business era basato su informazioni digitalizzate, hanno acquistato sempre più controllo sulle proprie dinamiche interne grazie a tecniche di analisi dei dati sempre più evolute aumentando in questo modo il loro vantaggio competitivo. Parallelamente, la digital age ha fatto emergere con sempre maggiore evidenza alcuni limiti del sistema di reporting tradizionale:

- **Stabilità:** ormai ogni settore economico è attraversato da forti turbolenze: innovazioni tecnologiche, nuovi player, cambiamenti degli scenari politici difficilmente rappresentabile attraverso un sistema di reporting statico;
- **Incapacità di sintetizzare i fattori critici di successo** come la qualità, il livello di servizio, il grado di innovazione, il time to market, l'immagine, l'apporto delle Risorse Umane alla competitività

aziendale, un carattere eccessivamente contabile delle informazioni che pone un focus quasi esclusivo sui risultati economico finanziari;

- ***Enfasi eccessiva sul breve periodo*** che non permette di accrescere la capacità dell'azienda di analisi interpretativa e quindi di apprendimento sui fenomeni attuali che possono influenzare le performance future;
- ***Gestione settoriale***: si presta eccessiva attenzione alle singole unità organizzative trascurando in questo modo le numerose interazioni che invece legano tra loro le unità stesse. Ci si concentra più su una visione settoriale che su una visione organizzativa d'insieme;
- ***Sistema "a consuntivo" e solo per poche persone*** che comporta valutazioni incomplete, tardive e ormai inutili.

## **Verso la Business Intelligence**

Il reporting direzionale è quell'insieme di rendiconti, tabelle e grafici opportunamente strutturati per area di responsabilità, che mettono a confronto i dati consuntivati con quelli programmati, al fine di:

- Evidenziare il rispetto degli obiettivi programmati dalla direzione e dal management;
- Identificare le eventuali cause di scostamento;
- Favorire l'adozione di azioni correttive e di conseguenza valutare le prestazioni dei gruppi di lavoro.

Ma tale sistema di reporting è un'utopia o si può realizzare? E cosa serve per realizzarlo?

Per consentire alle aziende di sfruttare il proprio patrimonio informativo, in vista di decisioni tattiche e strategiche, sono stati messi a punto numerosi strumenti denominati di Business Intelligence (BI).

Grazie alla BI è possibile raccogliere tutte le informazioni utili, gestirle e trasformarle in informazioni a supporto decisionale.

La Business Intelligence è un sistema di modelli, metodi, processi, persone e strumenti che rendono possibile la raccolta regolare ed organizzata del patrimonio dati generato da un'azienda. Inoltre attraverso elaborazioni, analisi o aggregazioni, ne permettono la trasformazione in informazioni, la loro conservazione, reperibilità e presentazione in una forma semplice, flessibile ed efficace, tale da costituire un supporto alle decisioni strategiche, tattiche ed operative.

Il sistema di Business Intelligence comporta dunque:

- La raccolta dei dati del patrimonio dell'azienda;
- La loro pulizia, validazione e integrazione;
- La successiva elaborazione, aggregazione e analisi;
- L'utilizzo fondamentale di questa mole di informazioni nei processi decisionali.

## **Definizione di Business Intelligence**

Il termine “Business Intelligence” nasce nel 1958 ad opera di Hans Peter Luhn, ricercatore e inventore tedesco che all'epoca lavorava per IBM.

La BI rappresenta lo strumento chiave dell'evoluzione verso una gestione sempre più efficace e strategica delle informazioni. Le condizioni di contesto in cui operano le organizzazioni, le risorse umane e tecnologiche a disposizione e la maggiore o minore apertura all'innovazione hanno determinato, infatti, un processo evolutivo del sistema informativo aziendale articolato in vari stadi.

Il primo di essi è costituito dalla creazione di una base dati (“database”), con l'obiettivo di raccogliere la gran mole di dati con cui le organizzazioni interagiscono quotidianamente.

Il passo successivo è, poi, quello di rendere i dati archiviati a disposizione del processo decisionale aziendale; a tal fine è fondamentale costruire un “Data Warehouse”, una sorta di grande magazzino integrato di dati aziendali

riorganizzati in modo funzionale alla loro condivisione e alla riutilizzabilità nei vari processi operativi.

Il Data Warehouse può essere suddiviso in vari sottoinsiemi, detti “data mart”, per permettere analisi puntuali di specifici settori dell’azienda o di variabili del business.

Organizzati i dati in modo strutturato, è possibile poi, attraverso gli strumenti di Business Intelligence, analizzare la performance aziendale attuale, prevedere quella futura e presentare i risultati di queste analisi alla direzione, che li utilizzerà poi per maturare le decisioni chiave della vita dell’azienda.

La massima evoluzione del sistema informativo è rappresentata, infine, dalla possibilità di accedere alle funzionalità sopra descritte attraverso internet, intranet ed altri canali elettronici, massimizzando, quindi, la condivisione e la riutilizzabilità delle informazioni da parte di tutti gli utenti.

Inizialmente la BI era utilizzata soprattutto da analisti e da altri professionisti dell’IT che avevano il compito di realizzare analisi e di produrre report per gli utenti di business che li richiedevano. In seguito, anche business executives hanno iniziato ad utilizzare software di BI, grazie anche allo sviluppo di tool di self-service BI e di data discovery.

La Business Intelligence combina un vasto set di applicazioni di analisi e software di visualizzazione dei dati al fine di realizzare grafici, tabelle o altri report informativi, e anche tool per realizzare dashboard e scorecards di performance che permettono di visualizzare dati sulla base di diverse metriche di business e i KPI aziendali esistenti e desiderati.

## **Storia della Business Intelligence**

Le applicazioni di Business Intelligence del giorno d’oggi sono soluzioni software che consentono di effettuare analisi dei dati per misurare e migliorare la qualità e il servizio dei processi di business. Questa è tuttavia

la situazione attuale ma ci sono voluti diversi anni per arrivare alle soluzioni di BI conosciute oggi.

L'opera *Cyclopaedia of Commercial and Business Anecdotes* di Richard Millar Devens del 1865 contiene il primo utilizzo, di cui si ha conoscenza, del termine "Business Intelligence" usato per descrivere il modo in cui il banchiere Sir Henry Furnese ottenne il successo: egli era a conoscenza delle questioni politiche, delle instabilità e del mercato prima dei suoi competitors. Nonostante Furnese finì per utilizzare questo vantaggio conoscitivo per duplici fini e diventò famoso per essere un finanziatore corrotto (strozzino), la sua idea di raccogliere informazioni sulle condizioni del business seminò i suoi frutti ed ebbe un notevole seguito.

Dal punto di vista del business, l'esigenza di utilizzare i dati per fare analisi emerse fin dai primi anni '70, ovvero fin dalla prima fase di informatizzazione dei processi aziendali "core": contabilità, fatturazione, acquisti, vendite, logistica, ecc.

Uno dei principali motivi di insuccesso delle soluzioni BI dagli anni '70 agli anni '90 fu che l'IT non era pronta ad accogliere e soddisfare i requisiti della BI: mancavano sia gli strumenti che le competenze specifiche. Lo stesso mondo accademico in quegli anni considerava ancora i sistemi di BI come "accessori" o come soluzioni di "serie B" rispetto ai sistemi di processo detti anche sistemi transazionali.

In questo stato delle cose, i professionisti IT, di fronte alle richieste del business, procedevano "empiricamente" e nella migliore delle ipotesi affrontavano e tentavano di risolvere i requisiti BI come facevano con i sistemi transazionali. Risultato: insuccesso assicurato.

## **La svolta del 1958**

Fu con la pubblicazione nel 1958 dell'articolo considerato la pietra miliare della BI scritto da Hans Peter Luhn, che al tempo lavorava in IBM, che il potenziale di questa tecnologia venne riconosciuto.

Questo articolo, intitolato “*A Business Intelligence System*”, descriveva un sistema automatico sviluppato per diffondere informazioni nelle varie aree di un’organizzazione sia essa industriale, scientifica e di governo.

Alla vigilia della rinascita dopo la seconda guerra mondiale, questi settori avevano bisogno di un modo per organizzare e semplificare la mole di dati scientifici e tecnologici in rapida crescita.

Il lavoro di Luhn fece di più che introdurre ed espandere le possibilità di un concetto; egli oggi è conosciuto come il padre della Business Intelligence.

## **Sviluppi ed evoluzioni fino al 1980**

Con l’avvento dei computer nel mondo del business, le aziende ebbero finalmente un’alternativa agli archivi cartacei per la memorizzazione e archiviazione dei dati.

L’invenzione di IBM degli hard disk nel 1956 rivoluzionò il sistema di archiviazione dei dati. Furono creati strumenti tecnologici per la memorizzazione di moli di dati crescenti come floppy disk, laser disk e altri.

Questo diede il via alla creazione del primo data base management system chiamato decision support system (DSS).

Dal 1970 alcuni vendor di software BI iniziarono a proporre tool che rendevano possibile l’accesso e l’organizzazione di questi dati. Ma questa tecnologia era nuova e ai primi sviluppi, e soprattutto era difficile da utilizzare.

## **Dal 1980 e il 1990**

La concorrenza tra i vendor di software di BI portarono a numerosi passi in avanti fino ad arrivare ai primi veri e propri Data Warehouse che andarono a migliorare il flusso di dati dai sistemi operazionali ai sistemi di supporto alle decisioni.

Il Data Warehousing ridusse drasticamente il tempo necessario per l'accesso alle informazioni richieste in quanto dati che tradizionalmente erano memorizzati in diversi database ora iniziano ad essere tutti collocati in un'unica postazione.

Di fianco a questo primo sviluppo sorsero altri elementi che oggi stanno alla base della BI come per esempio tool di ETL (Extract, Transform, and Load) e software OLAP (Online Analytical Processing).

Negli anni successivi, questa fase di sviluppo venne battezzata come Business Intelligence 1.0.

## **Business Intelligence 1.0**

A mano a mano che la Business Intelligence prese piede, tra la fine del 1990 e gli inizi del 2000, dozzine di nuovi vendor entrarono nel mercato.

Durante questo periodo, due erano le funzioni principali della BI: produrre report, e organizzare e visualizzare dati in maniera presentabile al board delle aziende. Continuavano tuttavia a persistere due problemi principali legati soprattutto alla tecnologia: la complessità e il tempo. Troppi progetti venivano ancora reindirizzati all'area IT a dimostrazione del fatto che gli utenti non erano ancora in grado di eseguire i task relativi alla BI in maniera autonoma. Inoltre, poiché i dati erano raggruppati in "silos", era necessario molto tempo per formalizzare e realizzare report. Solo esperti tecnici erano in grado di utilizzare software di analisi avanzata dei dati.

In seguito questi tool iniziarono ad evolvere e a orientarsi sempre più verso users non esperti e senza particolari esperienze tecniche, ma questa ulteriore evoluzione avvenne lentamente.

## **Business Intelligence 2.0**

L'alba del ventunesimo secolo segnò un punto di svolta grazie a sviluppi tecnologici che portarono notevoli miglioramenti in termini di complessità e velocità di elaborazione dei dati.

La BI 2.0 riguardò molte differenti tecnologie come ad esempio processi in real-time, che includevano informazioni provenienti da eventi così come apparivano nel Data Warehouse, permettendo alle aziende di prendere decisioni basandosi sulle informazioni più recenti disponibili. Altre tecnologie che iniziarono ad essere usate includevano accessi self-service per utenti non esperti, questo comportò che gli impiegati potevano ora completare progetti senza dover necessariamente essere supportati o in contatto con l'area IT.

Presto la BI non fu più semplicemente una utility aggiuntiva, o un mero vantaggio, ma stava diventando una necessità delle aziende per rimanere competitive e persino per rimanere a galla in un mondo nuovo, interamente guidato dai dati.

## **Modern Day**

Tool specifici, l'espansione di opzioni self-service e l'aumento delle visualizzazioni, sono tre tra i più importanti tratti della successiva frontiera dell'evoluzione della BI.

Oggi i tool della BI sono spesso disegnati seguendo una ben precisa industry. Conosciuta come "software verticalization", questa crescita di tool specifici per le industry ha contribuito significativamente all'aumento delle adozioni della Business Intelligence.

I tool di visualizzazione iniziarono ad evolvere fino ad includere sempre più gli utenti finali. Sempre più piattaforme potenziarono gli utenti grazie ad accessi self-service: essi potevano esplorare e utilizzare i dati per conto loro senza la necessità di alcun tipo di training.

I Big Data sono un argomento interessante per molte aziende, le quali negli ultimi anni hanno investito più di 15 miliardi di dollari, finanziando lo sviluppo di software per la gestione e l'analisi dei dati. Questo è accaduto perché le economie più forti sono molto motivate all'analisi di enormi quantità di dati: basti pensare che ci sono oltre 4,6 miliardi di smartphone



attivi e circa 2 miliardi di persone hanno accesso a internet. Il volume dei dati in circolazione si è evoluto nel corso degli anni:

- nel 1986 i dati erano 281 PetaBytes;
- nel 1993 i dati erano 471 PetaBytes;
- nel 2000 i dati erano 2,2 Exabyte;
- nel 2007 i dati erano 65 Exabyte;
- nel 2014 si prevede uno scambio di oltre 650 Exabyte

La quantità di Big Data e l'ampio uso di dati non strutturati non permette l'utilizzo dei tradizionali RDBMS (data base management system relazionali), che non rendono possibile archiviazione e velocità di analisi. Nell'ambito della Business Analytics nascono nuovi modelli di rappresentazione in grado di gestire tale mole di dati con elaborazioni in parallelo dei database. Architetture di elaborazione distribuita di grandi insiemi di dati sono offerte da MapReduce di Google, e dalla controparte open source Apache Hadoop.

## **Obiettivi della BI**

Gli obiettivi della BI sono:

- Passare da opinioni a fatti;
- Aumentare la qualità dell'informazione;
- Dare informazioni significative;
- Diffondere informazioni;
- Condividere informazioni.

Scegliere in un'azienda le soluzioni idonee di DW/BI richiede una visione strategica d'insieme considerando che si hanno a disposizione diverse classi di prodotti (di tipo build or buy), per diverse tipologie di utenti (utenti finali, manager di linea, direzione), in diverse aree aziendali, su diverse architetture tecnologiche.



Figura 1 Caratteristiche principali di un sistema di DW

## Caratteristiche della BI

Un sistema di BI deve rispondere a requisiti di funzionalità e progettazione che vanno ben oltre quelli propri di un normale ambiente di reportistica facente parte di un'applicazione gestionale. In particolare, un sistema di BI deve possedere le seguenti caratteristiche:

- **Facilità d'uso:** presentare i dati in un formato che sia facile da leggere e da interpretare, dove sia possibile navigare sui dati

seguendo dei percorsi di analisi facilmente comprensibili dall'utente finale;

- **Velocità:** possibilità di trattare grandi volumi di dati con tempi di risposta quasi istantanei grazie all'uso di tecniche di modellazione, memorizzazione e indicizzazione dei dati orientate all'analisi piuttosto che all'aggiornamento dei dati;
- **Integrazione:** integrare tra loro dati provenienti da fonti differenti, sia interne che esterne all'azienda. Il processo di integrazione deve essere affidabile e testato, in modo che gli utenti possano fare affidamento sui dati presenti nel DW. Se i dati provenienti dai sistemi operazionali non sono puliti ed affidabili, prima di essere inseriti nel DW devono passare attraverso un processo di pulizia e certificazione;
- **Storicizzazione:** mantenere la storia dei cambiamenti subiti da certi attributi selezionati, per permettere analisi storiche contestualizzate;
- **Identificazione di trend ed anomalie:** gli strumenti devono facilitare l'identificazione di trend nei dati, ad esempio confrontando periodi e prodotti diversi. Queste operazioni sono possibili solo con l'utilizzo di strumenti interattivi che permettano di effettuare operazioni di drilldown/roll-up (visualizzazione di un dato a diversi livelli di dettaglio) e di slice & dice (cambiamento delle dimensioni di analisi sui due assi);
- **Subject orientation:** presentare i dati in modo da ottenere la visione completa di un processo aziendale (supply chain, vendite, qualità...), attraversando i confini delle singole aree dei sistemi gestionali;
- **Simulazione scenari:** in certi casi (applicazioni di budgeting, forecasting and planning) deve essere possibile impostare degli scenari e confrontarli poi con i valori reali ("actual");
- **Indipendenza dal reparto IT:** gli strumenti di analisi e reportistica devono dare la possibilità agli utenti finali di crearsi da soli i report di cui hanno bisogno;

- *Adattabilità nel tempo*, intesa come la capacità di affrontare e sfruttare le inevitabili evoluzioni della realtà aziendale, dei sistemi operazionali e delle esigenze di analisi;
- *Sicurezza*: deve essere possibile controllare al tempo stesso in maniera stretta e flessibile l'accesso ai dati, che in molti casi includono informazioni altamente riservate.

## **La struttura della BI**

### **Extraction, Transformation and Loading (ETL)**

Il ruolo degli strumenti di ETL è quello di alimentare una sorgente dati singola, dettagliata, esauriente e di alta qualità che possa a sua volta alimentare il Data Warehouse. Le operazioni da essi svolte vengono spesso indicate con il termine *riconciliazione* che, durante il processo di alimentazione del Data Warehouse (DW) avviene in due occasioni: quando il DW viene popolato per la prima volta e periodicamente quando viene aggiornato. La riconciliazione consiste di quattro distinti processi detti rispettivamente:

1. Estrazione (extraction o capture);
2. Pulitura (cleaning o scrubbing);
3. Trasformazione (transformation);
4. Caricamento (loading).

In linea generale il confine tra pulitura e trasformazione è abbastanza nebuloso quindi per semplicità si assume che l'operazione di pulitura sia essenzialmente mirata alla correzione dei valori dei dati, mentre la trasformazione si occupa più propriamente del loro formato.

#### ***Estrazione***

Durante questa fase i dati rilevanti vengono estratti dalle sorgenti e questa operazione può essere di tipo:

- **Statico:** viene effettuata quando il DW deve essere popolato per la prima volta e consiste concettualmente in una fotografia dei dati operazionali;
- **Incrementale:** viene usata per l'aggiornamento periodico del DW, e cattura solamente i cambiamenti avvenuti nelle sorgenti dall'ultima estrazione. L'idea alla base è quella di utilizzare i cambiamenti registrati a livello dei dati per aggiornare il DW. I benefici derivabili sono: volume molto piccolo dei dati coinvolti di volta in volta nell'operazione rispetto all'estrazione statica; la maggior parte dei dati nel DW restano invariati e vengono analizzati solo i datai che hanno subito modifiche. Vengono usate tecniche CDC (Change Data Capture) che permettono di monitorare le sorgenti dati con l'obiettivo di individuare i cambiamenti avvenuti a livello dei dati. Queste tecniche sono particolarmente importanti per la Wata Warehouse maintenance grazie alla propagazione dei cambiamenti rilevati a livello della sorgente.

### ***Pulitura***

È la fase che si occupa di migliorare la qualità dei dati andando ad eliminare dati “sporchi” dovuti a duplicazioni, inconsistenze, dati mancanti, valori errati etc.

Le principali funzionalità di pulitura dei dati riscontrabili negli strumenti ETL sono la correzione e l'omogeneizzazione, che utilizzano dizionari appositi per correggere errori di scrittura e riconoscere sinonimi, e la pulitura basata su regole, che applica regole proprie del dominio per stabilire le corrette corrispondenze tra i valori.

### ***Trasformazione***

È la fase centrale del processo di riconciliazione e ha l'obiettivo di convertire i dati dal formato operativo sorgente a quello del DW. Tra le funzionalità di questo livello per l'alimentazione del livello dei dati riconciliati si hanno:

- conversione e normalizzazione, che operano sia a livello di formato di memorizzazione sia a livello di unità di misura al fine di uniformare i dati;
- matching, che stabilisce corrispondenze tra campi equivalenti in sorgenti diverse;
- selezione, che riduce, se necessario, il numero di campi e record rispetto alle sorgenti.

Nella fase di alimentazione del DW si hanno invece due sostanziali differenze:

- la normalizzazione viene sostituita dalla denormalizzazione;
- si introduce l'aggregazione che realizza le opportune sintesi dei dati.

### ***Caricamento***

In questa fase avviene il caricamento dei dati sul DW attraverso due modalità alternative:

- ***refresh***: i dati vengono riscritti integralmente sostituendo completamente quelli precedenti. In generale questa tecnica viene utilizzata solo durante la fase iniziale di popolamento del DW;
- ***update***: vengono aggiunti al DW solo i cambiamenti avvenuti sui dati senza sovrascrivere ad ogni iterazione tutti i dati. Questa tecnica viene utilizzata, in abbinamento all'estrazione incrementale per l'aggiornamento periodico del DW.

### **Data Warehouse**

Il Data Warehouse è uno strumento (o un insieme di strumenti) software avente lo scopo di memorizzare tutti i dati di interesse aziendale, sia che siano provenienti dai sistemi gestionali sia da fonti esterne, in una forma ottimizzata per effettuare interrogazioni e non scritture tradizionali (che rimangono appannaggio dei normali database operazionali "OLTP").

Per questo sono state concepite strutture di dati alternative a quelle dei database operazionali; mentre questi si basano sui concetti e sulle regole relazionali (Entity-Relationship), i Data Warehouse sono generalmente basati sul modello dimensionale o Star Schema, ottimizzato per rispondere velocemente a interrogazioni di vario tipo.

L'attività di Data Warehousing, cioè di costruzione e gestione di un Data Warehouse, comprende varie fasi:

- individuazione dei dati di partenza;
- conversione, estrazione e pulitura dei dati;
- utilizzo di un DBMS per gestire il Data Warehouse;
- Impiego di strumenti di Business Intelligence per accedervi.

### ***Caratteristiche del Data Warehouse***

- Il Data Warehouse deve consentire l'accesso ai dati aziendali garantendo bassi tempi di attesa fra l'interrogazione dei dati e l'output di risultati;
- I dati esposti devono essere consistenti: un particolare valore deve essere sempre lo stesso indipendentemente dal momento dell'interrogazione o dalla modalità di interrogazione dei dati;
- I dati esposti possono essere ricombinati e separati rispetto agli oggetti di calcolo definiti dal business;
- Nel concetto di Data Warehouse rientrano anche gli strumenti software necessari a facilitare l'interrogazione dei dati e a garantire una presentazione chiara della sintesi delle analisi;
- Il DW deve contenere dati coerenti e "ripuliti": capita che fonti dati importanti per il DW trascinino errori di diverso genere che possono avere un impatto devastante sulla qualità dei dati stessi. Parte fondamentale del processo di realizzazione di un sistema di DW è rappresentata proprio da attività di recupero, trasformazione e pulitura dei dati;

- La qualità del DW può mettere in luce problematiche nascoste all'interno dei processi aziendali, tipicamente l'assenza di dati relativamente ad una particolare attività dovuta all'opzionalità di compilazione dei dati può evidenziare, proprio grazie al DW, la necessità di verificare a livello di processi aziendali l'opportunità di definirla come obbligatoria;
- Il Data Warehouse è un sistema in sola lettura, gli utenti non eseguono azioni di aggiunta, modifica o eliminazione dei dati.

### ***Data Base vs Data Warehouse***

Il Data Warehousing è quindi:

- orientato ai soggetti di interesse;
- integrato e consistente;
- rappresentativo dell'evoluzione temporale;
- non volatile.

La differenza principale che intercorre tra un DB e un DW è nel concetto di denormalizzazione dei dati. Tale ridondanza permette di ottimizzare le caratteristiche per cui il DW viene progettato.

	<b>Data base</b>	<b>Data warehouse</b>
<b>Definizione</b>	Ogni collezione di dati organizzati con le funzioni di: memorizzazione, accessibilità e recupero dei dati.	Un tipo di data base che integra copie di dati transazionali provenienti da sistemi sorgenti differenti e li memorizza in un'unica sorgente per renderli disponibili per analisi future.
<b>Tipologia</b>	Ci sono differenti tipologie di database, ma per la maggior parte dei casi ci si riferisce ad un database applicativo OLTP. Altri tipi di database includono XML, file CSV, flat text e persino fogli Excel.	Un data warehouse è un database OLAP in cui i dati sono rappresentati secondo una struttura multidimensionale. Nota: non tutte le OLAP sono create allo stesso modo e differiscono in base alla modellazione dei dati.
<b>Similarità</b>	Sia i sistemi OLAP che OLTP memorizzano e gestiscono dati in forma di tabelle, colonne, indici, chiavi.	



<b>Utilizzo</b>	Tipicamente limitato ad una singola applicazione: un'applicazione = un data base. OLTP permette la realizzazione di processi transazionali real-time.	Può contenere e memorizzare dati per ogni tipo e numero di applicazioni: un data warehouse = infinite applicazioni e infiniti database. OLAP permette di mettere a disposizione dell'organizzazione una sola fonte dati come guida per l'analisi e per il decision making.
<b>Service Level Agreement (SLA)</b>	I database OLTP tipicamente hanno uptime del 99,99%. Il database è direttamente collegato all'applicazione di front end. Il dato è disponibile in real time al fine di essere accessibile immediatamente ("real-and-now") in caso di bisogno.	Con i database OLAP, i SLA sono molto più flessibili perché si mettono in preventivo occasionali downtime per il caricamento dei dati. Il database OLAP è separato dalle applicazioni di front end, cosa che gli permette di essere modulare. I dati sono aggiornati ogni volta che è necessario. Questi database sono utili per il decision making e per analisi dei trend storici.
<b>Ottimizzazione</b>	Ottimizzato per realizzare operazioni di lettura e scrittura da parte di un'unica transazione. Performare query di analisi molto lunghe su questo tipo di database non è una buona norma, perché questo impatta sulle performance del sistema. Una query analitica potrebbe richiedere diversi minuti durante i quali gli utenti vengono messi in stand-by e non possono operare.	Ottimizzati per operazioni di lettura/analisi di grandi set di dati. Un Data Warehouse è pensato per query di analisi molto lunghe e va ad eliminare i limiti di performance che tali analisi potrebbero comportare su un sistema transazionale.
<b>Organizzazione dei dati</b>	In un database OLTP tabelle molto complesse sono unite tra loro (mediante join) e i dati sono normalizzati (in questo modo i dati non vengono mai duplicati).	In un database OLAP, i dati sono organizzati in maniera tale da facilitare la reportistica e l'analisi. Il dato è denormalizzato per aumentare la velocità di risposta per le query analitiche e la facilità di utilizzo da parte degli utenti.

<b>Reporting e analisi</b>	A causa del numero di tabelle collegate, realizzare query di analisi è molto difficile. Esse generalmente richiedono l'esperienza di uno sviluppatore o di un database administrator familiare con quella applicazione. Il reporting è tipicamente limitato a più bisogni statici. È utile, ma non permette analisi approfondite.	Con meno tabelle in join diventa molto più semplice realizzare delle query analitiche. Questo significa che utenti non molto esperti possono soddisfare i propri bisogni senza dover ricorrere necessariamente all'aiuto dell'IT. Le possibilità di reporting e di analisi sono infinite. Un data warehouse permette diversi tipi di analisi: <ul style="list-style-type: none"> <li>• descrittiva (cosa sta accadendo);</li> <li>• diagnostica (perché ciò sta accadendo);</li> <li>• predittiva (cosa accadrà);</li> <li>• prescrittiva (cosa fare a proposito).</li> </ul>
<b>Utenti</b>	Migliaia.	Centinaia.
<b>Integrazione dei dati</b>	Per applicazione.	Per soggetto.
<b>Accesso</b>	A centinaia di record, in lettura e scrittura.	A milioni di record, per lo più in lettura.
<b>Scopo</b>	Dipende dall'applicazione.	Supporto alle decisioni.
<b>Dati</b>	Elementari, sia numerici sia alfanumerici.	Di sintesi, prevalentemente numerici.
<b>Copertura temporale</b>	Solo dati correnti.	Dati correnti e storici.
<b>Aggiornamenti</b>	Continui.	Periodici.
<b>Sviluppo</b>	A cascata.	Iterativo.

### ***Data Mart***

Un data mart (DM) è un raccogliatore di dati specializzato su un particolare segmento/ area aziendale. Contiene un'immagine di una porzione i dati e permette di formulare strategie sulla base degli andamenti passati. Viene collocato a valle di un DW ed è alimentato a partire da esso, di cui costituisce, in pratica, un estratto.

Detto in termini più tecnici, un data mart è un sottoinsieme logico o fisico di un DW di maggiori dimensioni. La differenza fondamentale consiste nel fatto che la creazione del Data Warehouse avviene prima in maniera generalizzata per poi subire modifiche per potersi adattare a esigenze specifiche, mentre il data mart viene generalmente creato per venire incontro ad una esigenza già determinata.

La necessità di creare un sistema separato per il DM rispetto al DW può riassumersi nelle seguenti motivazioni:

- necessità di utilizzare un diverso schema;
- migliorare la performance separando il computer dedicato;
- garantire una maggiore sicurezza dovendo autorizzare l'accesso ad un insieme minore di dati.

In generale è possibile affermare che più un approccio è articolato più offre vantaggi in termini di flessibilità e di ROI, ma nello stesso tempo fa aumentare il rischio di perdere il controllo sull'architettura.

## La Piramide Informativa della BI

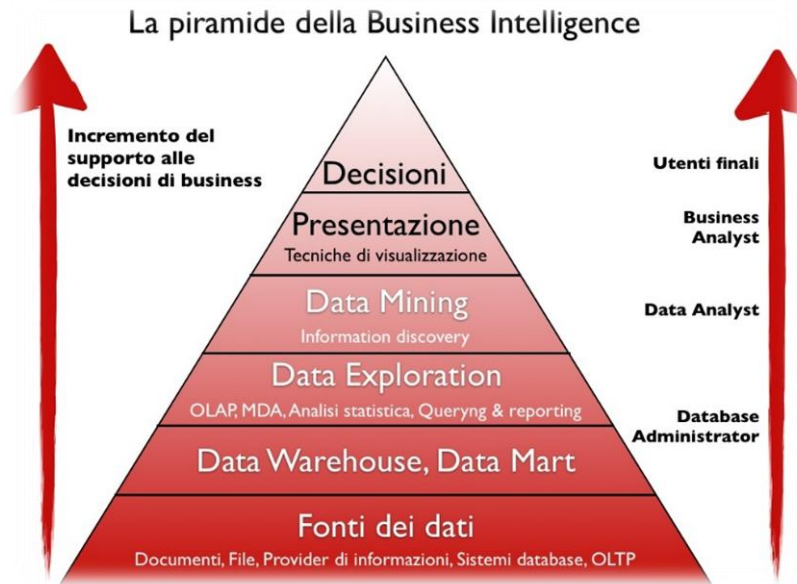


Figura 2 Rappresentazione piramidale della BI

Quello della Business Intelligence è un sistema che coinvolge ogni livello dell'infrastruttura tecnologica dell'azienda, poiché tutte le informazioni devono essere distillate per rendere più efficienti e precisi i processi decisionali.

La struttura della BI può quindi essere pensata come una piramide, alla base della quale si colloca il processo di raccolta dei dati attraverso le infrastrutture ad esso collegate. Salendo ai piani superiori si incontrano gli strumenti analitici basati sulla statistica e le attività di data mining. Ovviamente le implementazioni più raffinate di un processo di Business Intelligence richiedono grande potenza di calcolo e sono costose, ma possono ricavare previsioni attendibili anche a medio termine dell'andamento del business.

Una volta implementata in azienda una struttura di BI, bisogna poi saper usare gli strumenti a disposizione e, per comprendere appieno tecniche e applicazioni, sono necessarie competenze trasversali. Bisogna prima di tutto sapersi approcciare ai dati da una prospettiva business, il che può non essere facile per chi parte da una formazione troppo tecnica.

## Rappresentazione Multidimensionale

Il modello multidimensionale prende le mosse dalla constatazione che gli oggetti che influenzano il processo decisionale sono fatti del mondo aziendale. Le occorrenze di un fatto corrispondono a eventi accaduti. Per ciascun fatto, interessano in particolare un insieme di misure che descrivono quantitativamente gli eventi.

Gli eventi che accadono in un'azienda sono troppi per poterli analizzare singolarmente e quindi si è soliti collocarli in uno spazio *n-dimensionale* i cui assi (chiamati dimensioni di analisi) definiscono diverse prospettive per la loro identificazione.

È proprio questo concetto di dimensione che ha dato origine alla metafora del *cubo* per la rappresentazione dei dati multidimensionali. Secondo questa metafora, gli eventi corrispondono a celle di un cubo i cui spigoli rappresentano le dimensioni di analisi. Ogni cella del cubo contiene un valore per ciascuna misura.

La modellazione dati all'interno di un DW riguarda quindi tre concetti fondamentali propri del management:

1. fatto: un concetto sul quale centrare l'analisi. Modella un evento che accade nell'azienda;
2. misura: una proprietà atomica di un fatto da analizzare. Ne descrive un aspetto quantitativo;
3. dimensione: descrive una prospettiva lungo la quale effettuare l'analisi.

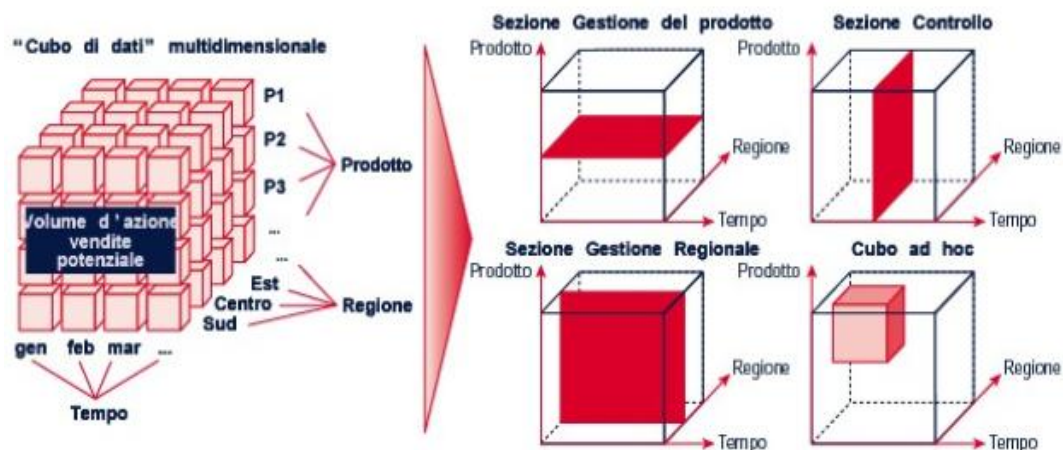


Figura 3 Rappresentazione multidimensionale dei dati

I due principali approcci all'interrogazione di un DW da parte degli utenti finali sono: la reportistica e l'OLAP.

## Reportistica

I sistemi di reportistica vengono sviluppati in ambiti complessi che hanno previsto una soluzione di Data Warehouse. Una delle finalità di un processo di DW è proprio quella di strutturare un contesto informativo hardware-software capace di rispondere alle esigenze dello scenario organizzativo nel senso più ampio.

Col crescere dei dati accumulati a disposizione delle organizzazioni, i vantaggi di un'elaborazione centralizzata dei documenti si rivelano nei tempi di esecuzione dei singoli documenti di reportistica: la particolare configurazione hardware delle postazioni su cui vengono - a livello fisico - ospitate le risorse del sistema permette l'ottimizzazione delle richieste al sistema e ne diminuisce il carico di attività rispetto alla situazione in cui singoli utenti ricercano informazioni sul sistema individualmente.

Un documento, una volta elaborato e generato, viene validato dalle strutture preposte e viene distribuito (ed aggiornato con cadenza periodica) agli appartenenti all'organizzazione che ne diventano i fruitori.

### ***Le fasi dello sviluppo di un sistema di reportistica***

Un processo di sviluppo di un sistema di reportistica è genericamente composto dalle seguenti fasi che possono essere ampliate o ridotte in conseguenza dei particolari ambienti di sviluppo, dei differenti contesti macroeconomici di attività dell'organizzazione:

- identificazione delle esigenze informative e di visualizzazione;
- identificazione del contesto informativo e delle fonti;
- identificazione della configurazione del sistema hardware/software;
- fase di integrazione hardware/software delle risorse informative;
- preparazione del report;
- validazione del report;
- fase di collaudo del sistema;
- fase di esercizio del sistema di reportistica.

Queste fasi non sono da intendersi necessariamente come consecutive in quanto alcune possono anche svolgersi in concomitanza.

### ***Il report***

Il documento prodotto viene chiamato *report* e si presenta come una combinazione di tabelle e grafici che presentano le misure di rilievo per i vari fenomeni analizzati, disaggregate e destrutturate secondo le esigenze. Tali misure costituiscono una base comune per le analisi successive.

### ***Gli strumenti informatici per i sistemi di reportistica***

I sistemi di reportistica hanno favorito la nascita ed il proliferare di una categoria di prodotti per la cosiddetta 'Office Automation'. Tali prodotti, per lo sviluppo dei quali si sono fatte concorrenza le maggiori software-house mondiali, si sono nel corso degli anni evoluti ed hanno allargato il loro spettro di standardizzazione fino a divenire strumenti integrati con quelli propri della Business Intelligence e convergendo verso vere e proprie piattaforme di 'Desktop Automation'.

## **OLAP**

Il termine OLAP, acronimo dell'espressione On-Line Analytical Processing, designa un insieme di tecniche software per l'analisi interattiva e veloce di grandi quantità di dati, che è possibile esaminare in modalità piuttosto complesse.

### ***Funzionalità***

La creazione di un database OLAP consiste nell'effettuare una fotografia delle informazioni in un determinato momento e trasformare queste singole informazioni in dati multidimensionali.

Eseguendo successivamente delle interrogazioni sui dati, in tal modo strutturati, è possibile ottenere risposte in tempi decisamente ridotti rispetto alle stesse operazioni effettuate su altre tipologie di database.

Una struttura OLAP creata per questo scopo è chiamata *cubo* multidimensionale. Ci sono diversi modi per creare un cubo, ma il più conosciuto è quello che utilizza uno schema *a stella*; al centro è posta la tabella maggiore dei *facts* che elenca i principali elementi su cui sarà costruita l'interrogazione, e collegate ad essa vi sono varie tabelle delle *dimensioni* che specificano come saranno aggregati i dati.

Un sistema OLAP permette di:

- studiare una grande quantità di dati;
- analizzare i dati da prospettive diverse;
- supportare i processi decisionali.

### ***Tipi di sistemi OLAP***

Partendo dai concetti di base appena descritti, si può precisare che esistono tre tipologie di sistemi OLAP: multidimensionale (MOLAP: *Multidimensional OLAP*), relazionale (ROLAP: *Relational OLAP*) e ibrido (HOLAP: *Hybrid OLAP*).



- MOLAP è la tipologia più usata e ci si riferisce ad essa comunemente con il termine OLAP. Sfrutta un database di riepilogo avente un motore specifico per l'analisi multidimensionale e crea le *dimensioni* con un misto di dettaglio ed aggregazioni. Risulta la scelta migliore per quantità di dati ridotte, perché è rapido nel calcolare aggregazioni e restituire risultati; tuttavia crea enormi quantità di dati intermedi.
- ROLAP lavora direttamente con database relazionali. I dati e le tabelle delle dimensioni sono memorizzati come tabelle relazionali e nuove tabelle sono create per memorizzare le informazioni di aggregazione. È considerato più scalabile e presenta requisiti di archiviazione e memoria minori; tuttavia, è lento nella creazione delle tabelle e nel generare il rapporto circa le interrogazioni.
- HOLAP utilizza tabelle relazionali per memorizzare i dati e le tabelle multidimensionali per le aggregazioni "speculative". Come dice il nome, questo sistema è un ibrido, poiché viene creato più velocemente di un sistema ROLAP ed è al tempo stesso più scalabile di un MOLAP.

### ***Caratteristiche di un OLAP***

Le funzioni di base di uno strumento OLAP sono:

- pivoting: è l'operazione di rotazione delle dimensioni di analisi;
- slicing: è l'operazione di *estrazione* di un subset di informazioni dall'aggregato che si sta analizzando. L'operazione di slicing viene eseguita fissando uno specifico valore per una delle dimensioni del "cubo", estraendo quindi una "fetta" e ottenendo un nuovo cubo con una dimensione in meno rispetto a quello di partenza;
- dicing: è l'operazione di *estrazione* di un subset di informazioni dall'aggregato che si sta analizzando. L'operazione di dicing viene eseguita quando l'analisi viene focalizzata su un sottoinsieme del "cubo" avente particolare interesse per l'analista;
- drill-down: è l'operazione di "esplosione" del dato nelle sue determinanti. L'operazione di drill-down può essere eseguita

seguendo due diversi percorsi: la *gerarchia* costruita sulla dimensione di analisi, oppure la *relazione matematica* che lega un dato calcolato alle sue determinanti;

- *drill-across*: è l'operazione mediante la quale si naviga attraverso uno stesso livello nell'ambito di una gerarchia. Il passaggio dalla famiglia di prodotti alla lista dei prodotti è un'operazione di drill-down, il passaggio da una famiglia a un'altra famiglia è un'operazione di drill-across;
- *drill-through*: concettualmente simile al drill-down, è l'operazione mediante la quale si passa da un livello aggregato al livello di dettaglio appartenente alla base dati normalizzata.
- *roll-up*: operazione inversa al drill-down.

## **Analisi What-If**

Con questo termine si indicano tecniche di analisi previsionale che permettono di valutare il comportamento di un sistema reale assumendo un particolare insieme di condizioni iniziali. Questo tipo di analisi supera uno dei limiti fondamentali di reportistica e OLAP, ossia quello di registrare solo il passato e di non consentire di analizzare scenari futuri. Per raggiungere l'obiettivo, il progettista deve essere in grado di disegnare un modello che riproduca con buona approssimazione il comportamento del sistema in esame; ovviamente maggiore sarà l'aderenza del modello alla realtà, maggiore sarà l'accuratezza delle stime ottenute.

È possibile classificare le tecniche di analisi what-if in base all'approccio utilizzato nell'elaborare il modello:

- *tecniche deduttive*: cercano di ricavare il modello di un sistema basandosi sul comportamento che questo ha esibito durante un certo intervallo temporale. Sono anche dette tecniche estensionali perché si basano appunto su un insieme di eventi accaduti nel passato;

- tecniche deduttive: sono mirate all'identificazione e alla caratterizzazione dei rapporti di tipo causa- effetto tra i componenti del sistema.

## **Data Mining**

L'ultimo livello comune a tutte le architetture di Data Warehousing è quello dell'analisi, il *data mining*. Questo è un processo per estrarre conoscenza da banche dati di grandi dimensioni tramite l'applicazione di algoritmi che individuano le associazioni nascoste tra le informazioni e le rendono visibili. Tecniche di data mining sono state usate per anni, prima dell'avvento del DW, in applicazioni specialistiche non commerciali. Tramite il Data Warehousing, il data mining viene trasportato dal mondo dell'analisi scientifica a quello dell'analisi commerciale, aprendo la via a numerose applicazioni utili al supporto decisionale permettendo:

- ricerche di mercato;
- studio dell'efficacia del marketing;
- studio della segmentazione del mercato per identificare profili cliente significativi;
- analisi delle abitudini di acquisto;
- pianificazione aziendale;
- altro.

In presenza di moli di dati molto elevate, nonostante l'utilizzo di strumenti di analisi sofisticati (per esempio OLAP) l'utente non è sempre in grado di individuare tutti i pattern significativi in essi presenti. Il data mining raccoglie un insieme di tecnologie e metodologie con l'obiettivo di aiutare l'utente alla ricerca di questi pattern.

Due sono gli elementi fondamentali del processo di estrazione di conoscenza (*knowledge discovery*) di cui il data mining costituisce il fulcro:

- dati: insieme di descrizioni di eventi derivate da una base di dati, da un DW, da un data stream;
- pattern: rappresentazione sintetica e ricca di semantica di un insieme di dati.

Oggi il data mining ha una duplice valenza:

- estrazione, con tecniche analitiche all'avanguardia, di formazione implicita, nascosta, da dati già strutturati, per renderla disponibile e direttamente utilizzabile;
- esplorazione ed analisi, eseguite in modo automatico o semiautomatico su grandi quantità di dati allo scopo di scoprire pattern (schemi) significativi.

In entrambi i casi i concetti di informazione e di significato sono legati strettamente al dominio applicativo in cui si esegue data mining, in altre parole un dato può essere interessante o trascurabile a seconda del tipo di applicazione in cui si vuole operare.

## Sintesi di un Ambiente di Data Warehousing

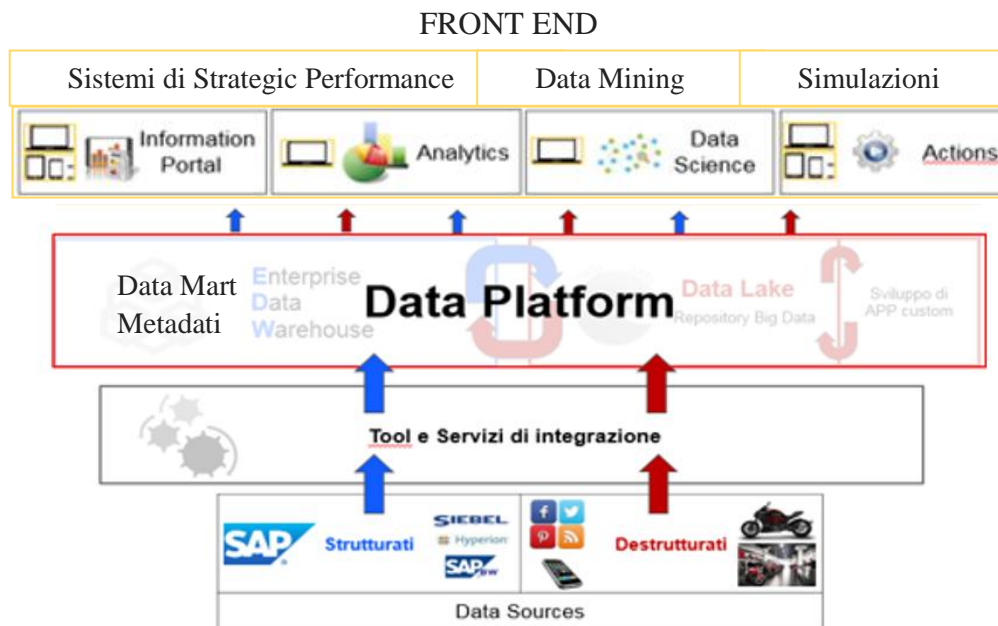


Figura 4 Schema dell'ambiente di DW

Un DW è un sistema di Business Intelligence basato su una base dati relazionale. Si tratta di una base dati separata rispetto al sistema gestionale, strutturata per contenere una replica parziale del sistema informativo aziendale in un formato che renda ottimale l'analisi delle informazioni.

Viene adottato per rispondere a esigenze quali:

- trasformare i dati in informazioni;
- fornire informazioni giuste al momento giusto e alle persone giuste;
- integrazioni tra fonti dati eterogenee;
- elevata profondità storica e indipendenza dai sistemi sorgente;
- pulizia dei dati e certificazione dei dati a tutti i livelli;
- gestione per eccezioni dei processi aziendali.

Dal punto di vista tecnico/architetturale la definizione di un ambiente di Data Warehouse può essere così semplificata:

- i dati sorgente, oggetto dell'analisi, vengono forniti di sistemi transazionali e possono essere di varia natura (sistemi ERP, fogli Excel ecc.);
- tali dati sono caricati nel Data Warehouse attraverso procedure di ETL;
- nel Data Warehouse i dati sono presenti al massimo dettaglio e storicizzati. Tuttavia spesso raggiungono dimensioni comunque poco compatibili con l'interrogazione diretta attraverso tool di analisi: vengono così generati dal DW aggregati tematici di dati (data mart) sui quali verranno scatenate le query di analisi;
- un notevole numero di tool permettono di effettuare analisi di ogni tipo e di migliorare ulteriormente le performance delle interrogazioni, i server OLAP (database multi-dimensionali) possono pre-calcolare le possibili combinazioni analitiche e memorizzarle fisicamente in un file binario (con esplosione delle dimensioni), tool di analisi appositi possono interrogare i dati sfruttando puntatori che indicano la posizione del dato pre-calcolato anziché attendere i tempi di elaborazione del motore database. I server OLAP possono memorizzare i dati con diverse modalità utilizzando strutture relazionali (ROLAP) o multidimensionali (MOLAP), per database smisurati (centinaia di giga) spesso si utilizza una modalità ibrida (HOLAP).

## **Top Business Intelligence Companies**

1. Microsoft: Microsoft Business Intelligence portfolio include MS Office- Excel, SQL Server, SharePoint e Power BI.
  - MS Office- Excel dà agli utenti la possibilità di scoprire, analizzare e visualizzare in completa autonomia i dati grazie ai potenti servizi offerti;
  - SharePoint permette la collaborazione e lo scambio di dati e report in un ambiente sicuro e organizzato;

- Power BI è un software basato sul cloud che offre servizi ad utenti non esperti;
  - SQL Server: Microsoft SQL Server è un DBMS relazionale (*Relational Database Management System* RDBMS); nelle prime versioni era utilizzato per basi dati medio-piccole, ma a partire dalla versione 2000 è stato utilizzato anche per la gestione di basi dati di grandi dimensioni. SQL Server offre prestazioni in-memory integrate nel database per transazioni e query, fornisce analisi approfondite più veloci sui dati grazie a strumenti di analisi e rende disponibili soluzioni Big Data a livello Enterprise.
2. Tableau: include una famiglia di prodotti di BI per la visualizzazione interattiva dei dati quali Desktop, Server, Public e Online.
    - Tableau Desktop è un tool drag-and-drop per l'esplorazione e la visualizzazione dei dati;
    - Tableau Online è una versione "host" di Tableau Server.
  3. Oracle: è il provider di Oracle Business Intelligence Enterprise Edition che è una piattaforma di BI che offre un ampio set di possibilità tra cui dashboard interattive, query ad hoc, report finanziari, scorecards amministrative e strategiche, etc;
  4. IBM: i prodotti di Business Intelligence di IBM, come Cognos BI e Cognos Insight), sono pensati per integrarsi a vicenda e con soluzioni di terzi;
  5. SAS: offre sistemi integrati, robusti e flessibili per analisi statistiche e predittive. Il portfolio include Enterprise BI Server, Visual Analytics and Office Analytics;
  6. SAP: è il provider di numerose soluzioni di BI. Il portfolio include SAP BI suite, Business Objects BI Suite, Lumira e SAP Crystal Reports.

Di seguito si riporta il quadrante di Gartner<sup>1</sup> nel quale sono rappresentati e classificati i principali vendor di software di BI.




---

<sup>1</sup> Gartner è un'importante azienda nel mondo IT che si occupa di analisi e ricerche di mercato/prodotto e advising. L'obiettivo di questa azienda è quello di supportare le decisioni strategiche dei propri clienti aiutandoli con consulenze e report in grado di fornire un punto di vista "super partes" sullo stato generale di un mercato, di un'azienda o dei suoi prodotti. Tutti i vendor fanno la fila per poter essere inclusi nei report di Gartner e, per molti clienti, i risultati di queste ricerche sono fondamentali per definire le strategie aziendali riguardo l'IT.



## **Vantaggi della BI**

La BI permette:

- ottimizzazione dei processi mediante la condivisione di informazioni con clienti, fornitori e partner;
- controllo del mercato:
  - monitoraggio del mercato;
  - individuazione del trend e opportunità di sviluppo del business.
- semplicità:
  - uso degli strumenti senza necessariamente avere conoscenze tecniche di business.
- flessibilità:
  - sperimentazione di modelli e soluzioni diverse;
  - analisi dinamiche.
- condivisione informazioni:
  - le aziende offrono ai dipendenti un accesso facile ed intuitivo alle informazioni strategiche;
  - condivisione delle responsabilità per la crescita del business.
- decision making “più consapevole” e veloce.
- ottimizzazione dei processi interni di business.

## **Ruoli nella Business Intelligence e composizione del team**

*Nei progetti di Business Intelligence di qualsiasi dimensione, la composizione del team è un elemento cruciale per il successo. Qual è l'organizzazione più efficiente per un team funzionale di BI? Come i ruoli individuali e le varie discipline possono contribuire al raggiungimento di un risultato di successo?*

Sebbene tutte le organizzazioni siano tra loro differenti, i ruoli all'interno del team tipicamente dipendono dalla maturità delle organizzazioni, dalla grandezza dei team stessi, dai set di skills disponibili e dalla cultura organizzativa. I ruoli non sono necessariamente correlati agli individui; una persona potrebbe ricoprire più di un ruolo, o i ruoli potrebbero essere suddivisi tra più individui.

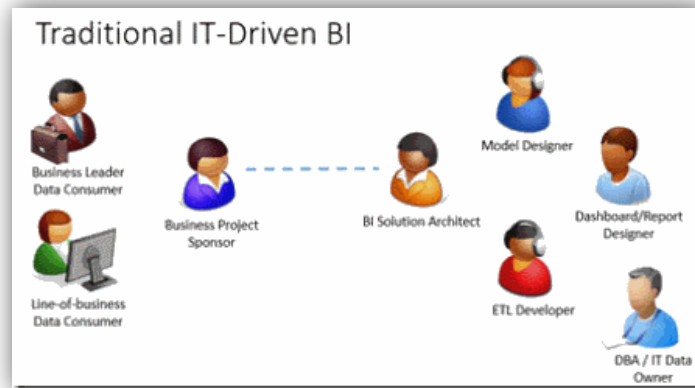
Team di BI di successo sono generalmente poco numerosi e lavorano in cicli brevi; in questo modo riescono a consegnare risultati a piccoli passi in modo che i business users e gli stakeholders possano fornire feedback senza dover aspettare lo sviluppo di lunghi cicli di lavoro.

In generale, un progetto di BI è diretto principalmente dal gruppo dell'IT o dal business, ma elementi indispensabili al team sono anche individui provenienti dall'area su cui in quel momento si focalizza il progetto.

## **Ruoli**

- **Business Sponsorship**

Avere un forte business sponsor dà al progetto obiettivi e direzioni. Il diagramma sottostante mostra un tradizionale progetto di BI guidato dall'IT. Da notare la relazione diretta tra il Business Project Sponsor, responsabile della definizione e dell'approvazione delle necessità del business e il BI Solution Architect, che è responsabile della gestione dei corrispondenti requisiti tecnici, di condurre il team di progetto e di prendere decisioni riguardo la soluzione architettonica che verrà adottata.



Altre capacità tecniche specifiche dell'IT varieranno in base alla dimensione e alla portata del progetto. In progetti più estesi alcuni ruoli che potrebbero essere considerati sono ad esempio *Data Warehouse Architect* e *Data Mart Developer*;

- **ETL Developer**

Sviluppa i pacchetti e i database usati per caricare e trasformare i dati dai sistemi sorgente ai data mart. Lo sviluppo dell'ETL è spesso un lavoro sottovalutato che invece si rivela più arduo del previsto. Il miglioramento della qualità dei dati è uno dei compiti dell'ETL che tuttavia può essere molto time-consuming. Investire al fine di migliorare in anticipo la qualità dei dati può essere molto utile in quanto porterebbe alla riduzione di tempi e di costi;

- **DBA/IT DATA Owner**

Il DBA gioca un ruolo fondamentale e dovrebbe essere coinvolto in molti aspetti all'interno dei progetti di BI. Oltre a gestire i server e i database utilizzati, garantisce l'accesso ai sistemi sorgente, crea e assegna agli account e aiuta a risolvere i problemi di comunicazione tra sviluppatori, utenti e responsabili della fase di test. Il DBA dovrebbe avere tempo sufficiente per supportare efficacemente il team e partecipare alle riunioni;

- **Model Designer**

Elabora e sviluppa modelli semantici usando tecnologie come SQL Server Analysis Service o Power Pivot. Il Model Designer può inoltre disegnare lo schema relazionale dei data mart anche se a volte questo compito potrebbe venire assegnato a un data mart schema designer dedicato. Il livello di impegno di un data mart designer dipende dall'estensione dei progetti e dalla maturità dell'organizzazione e anche dall'esistenza o meno in azienda di un data warehouse o di un master data hub. Il model designer deve inoltre stabilire le gerarchie e le relazioni tra i dati e gli attributi usati nel reporting. I calcoli previsti usano linguaggi specifici come DAX e DMX per definire misure e KPI;

- **Dashboard/Report Designer**

Le visualizzazioni dei dati tipicamente rientrano in classi pre-definite di dashboard e report realizzati per essere consegnati e utilizzati dagli utenti o per il self-reporting. Entrambe le soluzioni richiedono lavori antecedenti per permettere agli utenti di accedere e navigare in maniera efficace ed efficiente i dati di business. La realizzazione di report convenzionali richiede l'utilizzo di tool come Reporting Services per costruire pixel-perfect report (con questo termine si indica la possibilità offerta agli utenti di poter manipolare la grandezza e la posizione degli elementi del report, ma anche di regolare le dimensioni della pagina da stampare con un elevato grado di precisione) con caratteristiche che permettano di filtrare i dati e altre operazioni come drill-down dei dati e drill-through tra i report;

- **BI Solution Architect**

È colui che conduce il progetto e che mantiene allineati gli sviluppi progressivi dei lavori con il piano originario e ingaggia i membri del team e dà loro la giusta occupazione. I progetti di BI spesso richiedono di risolvere problemi complessi e profondi ma consistono anche in opportunità per fare pratica nel trovare nuove soluzioni creative e interessanti. Questo dinamismo tuttavia potrebbe portare i

progetti a diffondersi in maniera caotica in ogni direzione e quindi è di fondamentale importanza che ci sia un team leader in grado di convogliare i membri del team verso obiettivi prioritari e in grado di individuare i giusti trade-off decisionali. In grandi progetti, il Solution Architect, potrebbe essere un ruolo singolare con compito di staffing, di gestione delle negoziazioni e meeting management. In progetti minori, potrebbe essere semplicemente un membro senior del team;

- **Data Analyst**

Il profilo del tipico business user è cambiato negli ultimi anni. Il ruolo del Data Analyst o Data Scientist definisce questo user non più come un semplice consumatore di dati, ma come un abile realizzatore di analisi dei dati;

- **Data Steward**

Questo ruolo è critico nelle soluzioni di business a lungo termine. Il DBA controlla la sicurezza e la manutenzione del database mentre il Data Stewart assicura che una lista di record principali sia disponibile per users e analisti e spesso funge da coordinatore tra sorgenti applicative, sviluppatori e il DBA. Alcune aziende utilizzano software di Mater Data Management per assistere questo ruolo.



## Business Analytics

Spesso la BI è utilizzata in maniera intercambiabile con la Business Analytics (BA); in altri casi la Business Analytics è considerata come un'analisi dei dati più approfondita e avanzata e comprendente anche la BI.

<b>BI vs BA</b>	<b>Business Intelligence</b>	<b>Business Analytics</b>
<b>Risponde alle domande</b>	Perché è successo? Quando? Chi? Quante volte?	Perché è successo? Succederà ancora? Cosa succede se cambiamo "x"? Cosa ci possono dire in più i dati rispetto a quello che abbiamo sempre chiesto?
<b>Include</b>	Reporting (KPI e altre misure) Query ad hoc OLAP (cubi, drilling, slice & dice) Dashboard e scorecards BI operational / real time Monitoraggio automatico e segnalazione errori	Analisi statistiche/ quantitative Data mining Analisi predittiva e modeling Analisi dei Big Data Analisi testuale

Esempi di utilizzo della BA includono:

- esplorazione dei dati per trovare nuovi pattern e relazioni (data mining);
- analisi delle motivazioni del verificarsi di un determinato risultato (analisi statistiche, analisi quantitative);
- effettuare test per verificare diverse ipotesi a livello decisionale (A/B testing, multivariate testing);
- prevedere risultati futuri (predictive modeling, predictive analysis).

A causa della crescente popolarità della Business Analytics, sempre più spesso i vendor di applicazioni di BI includono tra i loro prodotti con funzionalità di BA.

## Terminologia

- ***Analisi what-if:*** tecnica di analisi previsionale che permette di simulare il comportamento di un sistema reale assumendo un particolare insieme di condizioni iniziali;
- ***Budgeting:*** è il bilancio di previsione e rientra tra gli strumenti fondamentali di programmazione e controllo dell'azienda;
- ***Business Intelligence:*** disciplina che consente ai decisori aziendali di capire, attraverso soluzioni software, i fattori chiave del business e conseguentemente di prendere le migliori decisioni;
- ***Business Process Management:*** è l'insieme di attività necessarie per definire, ottimizzare, monitorare e integrare i processi aziendali;
- ***BPM (Business Performance management):*** insieme di attività atte a misurare le prestazioni aziendali incoraggiando l'efficacia dei processi e l'uso efficiente delle risorse umane, materiali ed economiche;
- ***CPM:*** Corporate Performance Management, è l'insieme delle attività per definire le strategie aziendali e assicurarne il corretto sviluppo;

- **CRM:** Customer Relationship Management, è l'insieme delle attività di gestione dei rapporti con la clientela;
- **Cubo:** metafora adottata per indicare la rappresentazione dei dati nel modello multidimensionale. Ogni evento accaduto viene visto come una cella di un cubo i cui spigoli rappresentano le dimensioni di analisi;
- **Data Warehousing:** metodi, tecnologie e strumenti di ausilio al "lavoratore della conoscenza" per condurre analisi dei dati finalizzate all'attuazione di processi decisionali e al miglioramento del patrimonio informativo aziendale;
- **Data Mining:** estrazione di informazioni utili da grandi moli di dati con tecniche di tipo statistico;
- **Dashboard:** forniscono una rappresentazione per immagini delle prestazioni dell'intera organizzazione. E' una raccolta di analisi e report di KPI che offre agli utenti una singola visualizzazione dei dati con l'obiettivo di aiutarli a monitorare le informazioni associate a un lavoro, a un progetto o ad un obiettivo;
- **Data mart:** un sottoinsieme o un'aggregazione dei dati presenti nel datawarehouse, contenente l'insieme delle informazioni rilevanti per una particolare area di business, una particolare divisione dell'azienda, una particolare categoria di soggetti;
- **Database operativa:** contiene i dati generati da operazioni principalmente di carattere amministrativo, svolte all'interno dei processi gestionali;
- **Data warehouse:** una collezione di dati di supporto per il processo decisionale orientata ai soggetti di interesse, integrata e consistente, rappresentativa dell'evoluzione temporale e non volatile;
- **DFM (dimensional fact model):** modello concettuale per la progettazione di data warehouse;
- **Drill-across:** operatore OLAP che stabilisce un collegamento tra due o più cubi correlati al fine di compararne i dati, per esempio calcolando espressioni che coinvolgono misure prese dai due cubi;



- **Drill- down:** operatore OLAP che diminuisce l'aggregazione dei dati in un cubo introducendo un ulteriore livello di dettaglio;
- **Drill-through:** operatore OLAP che permette il passaggio dei dati aggregati multidimensionali di un cubo ai dati operazionali presenti nelle sorgenti o nel livello riconciliato;
- **ETL (Extraction, transformation, loading):** componente fondamentale del Data Warehouse con l'obiettivo di estrarre, trasformare e caricare i dati provenienti da sorgenti diverse su un Data Warehouse;
- **Entity/ Relationship:** diffuso modello concettuale per la progettazione di basi di dati operazionali;
- **Estrazione:** fase del processo di Data Warehousing durante la quale i dati rilevanti vengono estratti dalle sorgenti;
- **Join:** date due relazioni e un predicato booleano che coinvolge attributi di entrambe, l'operazione di join collega tutte le coppie di tuple prese dalle due relazioni per le quali il predicato è soddisfatto;
- **Knowledge Management:** condivisione della conoscenza;
- **Modello multidimensionale:** modello alla base dei sistemi di Data Warehousing, vede i dati come punti in uno spazio le cui dimensioni corrispondono ad altrettante possibili dimensioni di analisi; ciascun punto, rappresentativo di un evento accaduto nell'azienda, viene descritto tramite un insieme di misure di interesse per il processo decisionale;
- **Modello relazionale:** modello logico per la memorizzazione dei dati sotto forma di relazioni o tabelle;
- **MOLAP:** implementazione del modello multidimensionale su un modello logico ad hoc nel quale i dati multidimensionale possono essere direttamente rappresentati sotto forma di matrici ad accesso posizionale;
- **OLAP (on-line analytical processing):** metodologia di analisi dei dati tipica dei sistemi di data warehouse;
- **OLTP (on- line transactional processing):** elaborazione interattiva dei dati basata su transazioni;

- **Planning:** processo di pianificazione;
- **Pattern:** nel DFM, insieme di attributi delle gerarchie che definiscono il livello di aggregazione per un'interrogazione. Nel data mining, modello significativo estratto dai dati grezzi;
- **Pulitura:** fase del processo di Data Warehousing durante la quale viene migliorata la qualità dei dati;
- **Report:** sintesi chiara e schematica dei dati aziendali;
- **ROLAP:** implementazione del modello multidimensionale sul modello logico relazionale, effettuata tipicamente attraverso uno schema a stella;
- **Roll-up:** operatore OLAP che aggrega un insieme di eventi in un cubo diminuendone il livello di dettaglio;
- **Scorecards:** consentono la misurazione delle performance aziendali sotto diversi profili: finanziario, cliente, interno e innovazione. E' una raccolta di informazioni in linea con gli obiettivi strategici dell'organizzazione, che consente ai dipendenti di sapere in che modo i loro obiettivi e le loro attività si correlano agli obiettivi aziendali. Mostra velocemente come si sta andando a fronte degli obiettivi. Comunica la strategia e gli obiettivi;
- **Slicing:** operatore OLAP che riduce la dimensionalità di un cubo fissando un valore per una delle sue dimensioni.

## **CAPITOLO 2:**

# **BUSINESS INTELLIGENCE E BIG DATA NEL MERCATO DELL'AUTOMOTIVE**

### **Definizione di BIG DATA**

Big data è il termine usato per descrivere una raccolta di dati così estesa in termini di volume, velocità e varietà da richiedere tecnologie e metodi analitici specifici per l'estrazione di informazioni a valore.

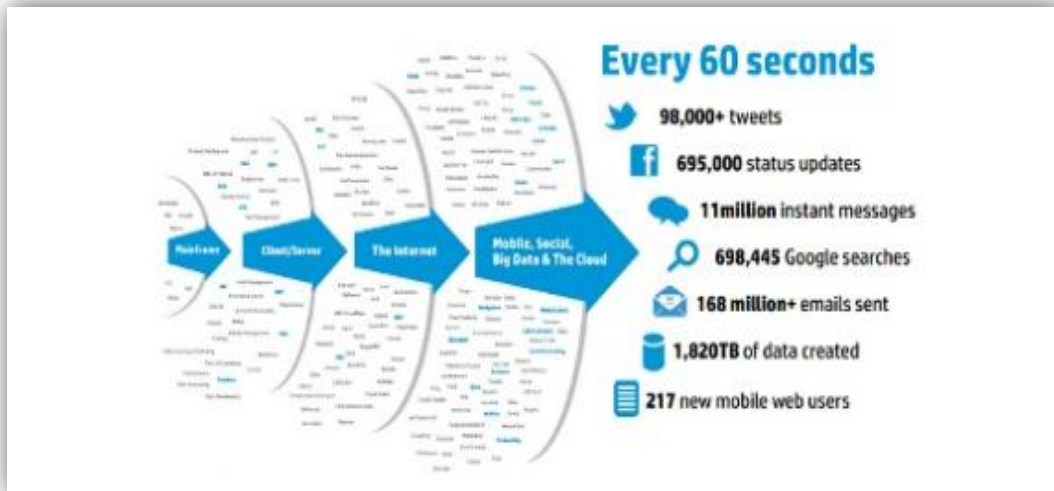
Il progressivo aumento della dimensione dei dataset è legato alla necessità di effettuare analisi su un unico insieme di dati, con l'obiettivo di estrarre informazioni aggiuntive rispetto a quelle che si potrebbero ottenere analizzandone piccoli set tra loro separati.



*Figura 1 Crescita dei BIG DATA*

Big data rappresenta anche l'interrelazione di dati provenienti potenzialmente da fonti eterogenee, quindi non soltanto i dati strutturati, come i database, ma anche non strutturati, come immagini, email, dati GPS, informazioni prese dai social network.

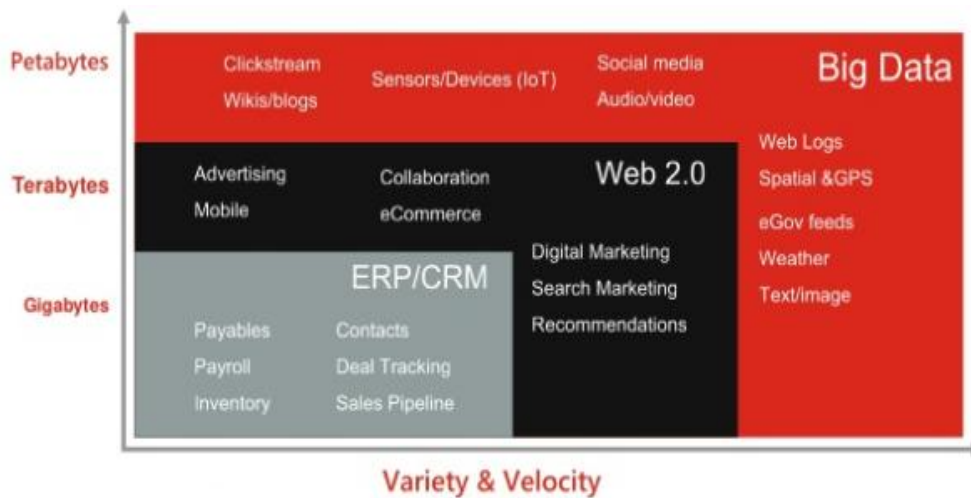
Con i big data la mole dei dati è dell'ordine degli Zettabyte, ovvero miliardi di Terabyte. Si richiede quindi una potenza di calcolo parallelo e massivo con strumenti dedicati eseguiti su decine, centinaia o anche migliaia di server.



La complessità dei big data è soprattutto legata alle così dette 4 'V':

- **Volume:** immense quantità di dati che devono essere processati. Rappresenta la dimensione effettiva del dataset; l'ampio volume di dati che è possibile raccogliere oggi potrebbe apparentemente rappresentare un problema; in realtà la questione risulta facilmente risolvibile con l'adozione di cloud e della virtualizzazione che aiutano nella gestione del grosso volume di dati disponibili, semplificando i processi di raccolta, immagazzinamento e accesso ai dati;
- **Velocità:** si riferisce alla velocità di generazione dei dati. Si fa riferimento ad un processamento dei dati ad alta frequenza o real time;
- **Varietà:** riferita alle varie tipologie di dati strutturati e non;
- **Variabilità:** "messiness", questa caratteristica si riferisce alla possibilità di inconsistenza dei dati.

## BIG DATA all'interno delle aziende



## Le 9 sorgenti dei BIG DATA

- **Archivi:** archivi di documenti scannerizzati, statements, moduli di assicurazioni, corrispondenze dei clienti, rapporti e cartelle mediche, file stampati che contengono sistemi originali di record tra le organizzazioni e i loro clienti;
- **Documenti:** XLS, PDF, CSV, email, Word, PPT, HTML, XML, JSON, etc.;
- **Media:** immagini, video, audio, Flash, podcast, live streams, etc.;
- **Data storage:** SQL, NoSQL, Hadoop, file system, etc.;
- **Business Apps:** project management, marketing automation, produttività, CRM, gestione dei contenuti del sistema ERP, HR, talent management, gestione del magazzino, intranets, Goggle Docs;
- **Public web:** government, competitors, traffic, censo, servizi sanitari, finanza pubblica, World Bank, Wikipedia, etc.;
- **Social media:** Twitter, LinkedIn, Facebook, Blog, Google+, Instagram, Pintarest, etc.;
- **Machine log data:** server data, mobile location, mobile app usage, application logs etc.;
- **Sensor data:** medical devices, smart electric, car sensors, satelliti, video games, processori collocati su veicoli etc.

## Alcuni tool di supporto per la gestione dei BIG DATA

- Splunk;
- Google analytics (usato soprattutto per analisi del traffico Web);
- Mixpanel (per user-event data);
- Geckoboard (per aggregating data sources).

## Differenze con la Business Intelligence

La crescente maturità del concetto di Big Data mette in evidenza le differenze con la Business Intelligence, in materia di dati e del loro utilizzo:

- Business Intelligence utilizza la statistica descrittiva<sup>2</sup> con dati ad alta densità di informazione per misurare cose, rilevare tendenze, ecc., cioè utilizza dataset limitati, dati puliti e modelli semplici;
- Big Data utilizzano invece la statistica inferenziale<sup>3</sup> e concetti di identificazione di sistemi non lineari, per dedurre leggi (regressioni, relazioni non lineari, ed effetti causali) da grandi insiemi di dati e per rivelare i rapporti, le dipendenze, ed effettuare previsioni di risultati e comportamenti, cioè utilizza dataset eterogenei (non correlati tra loro), dati raw e modelli predittivi complessi.

---

<sup>2</sup> La **statistica descrittiva** è la branca della statistica che studia i criteri di rilevazione, classificazione, sintesi e rappresentazione dei dati appresi dallo studio di una popolazione o di una parte di essa (detta campione).

I risultati ottenuti nell'ambito della statistica descrittiva si possono definire certi, a meno di errori di misurazione dovuti al caso, che sono in media pari a zero. Da questo punto di vista si differenzia dalla statistica inferenziale, alla quale sono associati, inoltre, errori di valutazione.

<sup>3</sup> L'**inferenza statistica** (o **statistica inferenziale**) è il procedimento per cui si inducono le caratteristiche di una popolazione dall'osservazione di una parte di essa (detta "campione"), selezionata solitamente mediante un esperimento casuale (aleatorio).

## Tipi di analisi effettuabili sui Big Data

Sui dati possono essere effettuati diversi tipi di analisi che, in ordine di difficoltà e complessità sono:

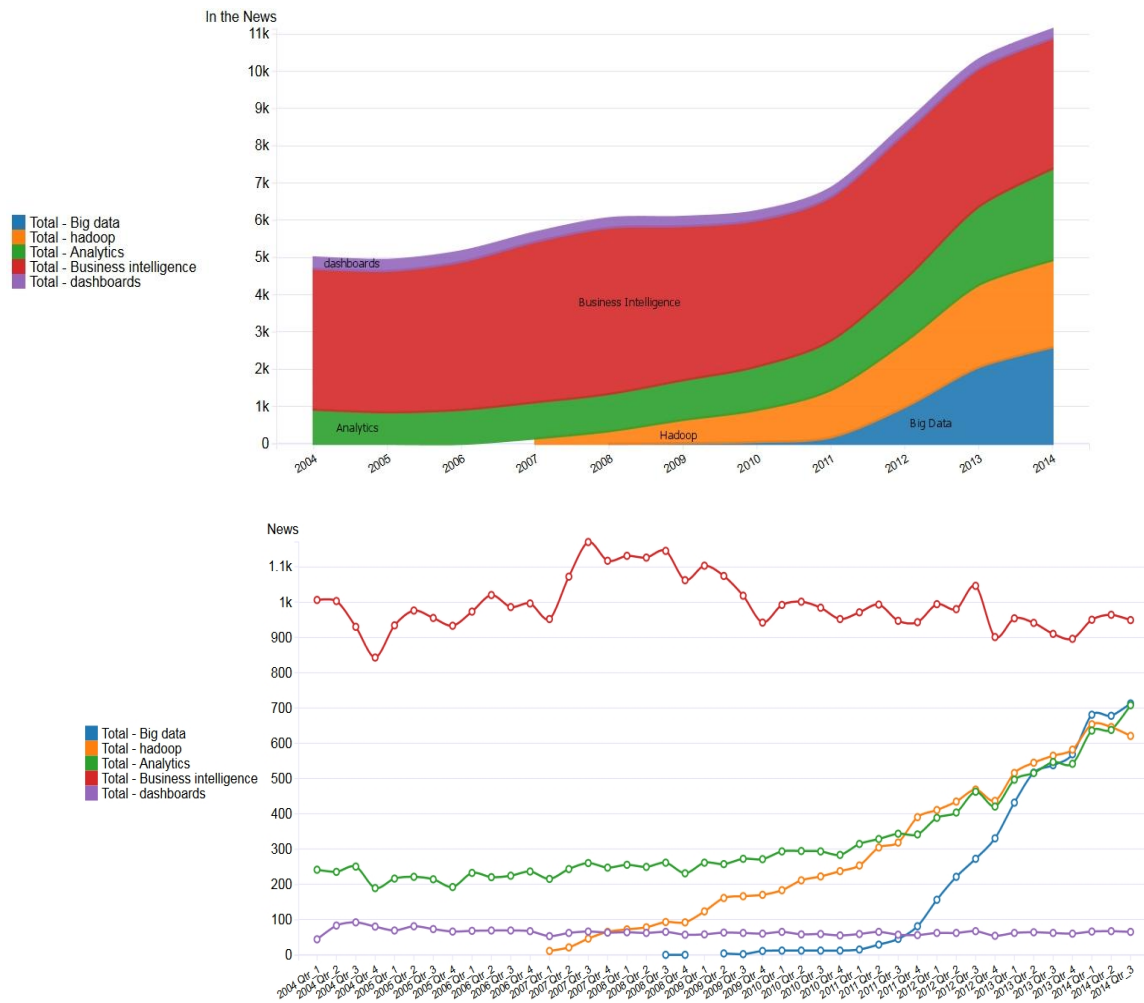
- ✓ Analisi descrittive:
  - l'obiettivo è la descrizione di un set di dati;
  - questa è la prima tipologia di analisi dei dati che viene effettuata;
  - viene comunemente applicata a dati censiti;
  - descrizione ed interpretazione dei dati sono step differenti;
  - le descrizioni non possono normalmente essere generalizzate senza addizionali modelli statistici.
  
- ✓ Analisi esplorative:
  - l'obiettivo è quello di trovare relazioni di cui al momento non si sa nulla;
  - i modelli esplorativi sono ottimi per scoprire nuove connessioni;
  - sono inoltre utili per definire studi futuri;
  - le singole analisi esplorative non sono usate da sole per generalizzare o predire connessioni.
  
- ✓ Analisi inferenziale:
  - usano un campione di dati relativamente piccolo per dire qualcosa su una popolazione più grande;
  - l'inferenza è comunemente l'obiettivo dei modelli statistici;
  - l'inferenza riguarda sia la stima della misura oggetto di analisi sia l'incertezza della stima effettuata;
  - l'inferenza dipende sia dalla popolazione che dallo schema campione.

- ✓ Analisi predittiva:
  - l'obiettivo di questa tipologia di analisi è quello di permettere l'utilizzo di dati relativi ad un oggetto per predire valori e dati relativi ad un altro oggetto;
  - se X predice Y non significa che X causa Y;
  - predizioni accurate dipendono soprattutto dalla misurazione dei giusti valori.
  
- ✓ Analisi causale:
  - l'obiettivo è quello di scoprire cosa accade ad una variabile quando viene apportato un cambiamento su un'altra variabile;
  - generalmente studi randomized sono necessari per individuare le cause;
  - ci sono approcci che permettono di inserire questa causalità anche in studi non-randomized;
  - i modelli casuali sono in genere i "gold standard" per le analisi dei dati.
  
- ✓ Analisi meccanicistica:
  - l'obiettivo è quello di capire esattamente quali cambiamenti di variabili inducono cambiamenti in altre variabili concentrandosi su singoli obiettivi di analisi;
  - molto difficili da dedurre se non in situazioni molto semplici;
  - quest' analisi viene generalmente modellata da un set deterministico di equazioni;
  - se le equazioni sono note ma i parametri non lo sono, questi possono essere dedotti da analisi dei dati.

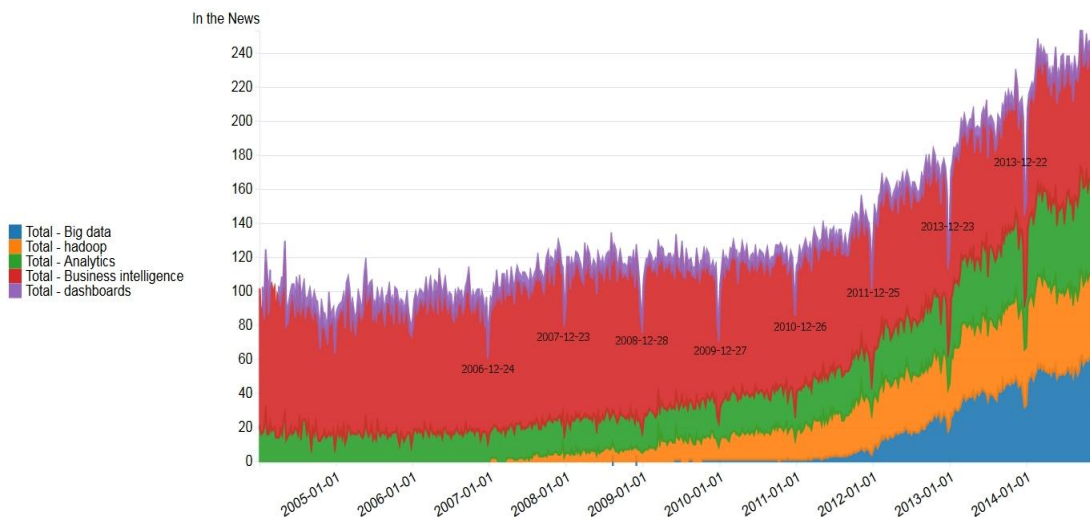


## Ultimi anni

Come è cambiato, nel corso degli anni, l'interesse per i Big Data, Hadoop<sup>4</sup>, la Business Intelligence, la Business Analytics e l'utilizzo di dashboards?



<sup>4</sup> Hadoop supporta le analitiche di business gestendo i dati non strutturati che arrivano sempre più numerosi dal mondo dei social media e dalla Internet of Things. Il framework, infatti, permette alle applicazioni di lavorare velocemente con migliaia di nodi e petabyte di dati, attraverso alcune tecniche di ottimizzazione dello storage.



Evidentemente le nuove e le vecchie tecnologie di gestione e analisi dei dati (tool di Business Intelligence, dashboards, Big Data ecc.) sono in continua espansione e il loro processo di sviluppo non presenta segnali di possibili futuri rallentamenti.

## **BIG DATA e Business Intelligence nel mercato dell'automotive**

La Business Intelligence aiuta le aziende del mercato dell'automotive (inteso nella sua accezione più generica: mercato di auto e moto) a spostarsi sempre più velocemente all'interno del mercato per rispondere alle nuove esigenze dei clienti e ad avere crescente visibilità finanziaria sulle performance del business aumentando così l'efficienza degli utenti.

Data l'intensa natura competitiva dei mercati, l'utilizzo efficiente delle informazioni è di vitale importanza per il successo. L'intensa competizione e la crisi economica hanno aumentato la pressione sul margine di profitto e hanno ristretto la "window of opportunity" per l'introduzione di nuovi prodotti. All'interno di questo scenario le industrie dell'automotive continuano a rispondere globalizzando le loro supply chain alla ricerca di costi sempre più contenuti. Queste lunghe e complesse supply chain presentano un numero sostanziale di cambiamenti finanziari e di business, inclusi:

- design-centre-to-dealer time-to-market performance;
- domanda di prodotti sempre più innovativi per clienti sempre più sofisticati;
- giacenze in eccesso e un incorretto mix di materiali in giacenza nei canali di distribuzione;
- lead-time lunghi e rigidi per i percorsi critici;
- garanzie e servizi post vendita.

La BI, oltre a dare significato ai dati, permette di generare soluzioni mirate per vincere le sfide generate nel local business e porta ad ulteriori benefici, come l'accelerazione del time to market per nuovi prodotti al fine di stimolare con continuità la domanda.

Una compagnia di successo nel settore dell'automotive, richiede aggiornamenti real-time delle informazioni che riguardano clienti, fornitori e competitors, delle decisioni collegate alle vendite, alla gestione della qualità, allo sviluppo di nuovi prodotti e alla realizzazione di piani di produzione.

Utilizzare e condividere gli strumenti e le potenzialità della BI è importante per ottenere gli outcomes di business desiderati, specifici miglioramenti sia dei prodotti che della qualità e per ottenere tutte le informazioni necessarie dai e per i clienti.

Le soluzioni offerte dalla BI permettono alle industrie dell'automotive di ottimizzare gli outcomes di business attraverso:

- l'accelerazione dell'innovazione dei prodotti;
- il miglioramento della qualità e dei tempi per le decisioni di business;
- l'espansione dell'utilizzo delle informazioni e il miglioramento della collaborazione attraverso tutti i livelli della supply chain globale;
- il miglioramento della visibilità della supply chain e la minimizzazione dei disturbi alla produzione;

- la gestione dei rischi e il rispetto delle norme;
- la riduzione dei costi di garanzia e l'aiuto per migliorare la qualità dei prodotti.

La BI permette inoltre di migliorare le performance del business; aiuta le aziende dell'automotive a raggiungere importanti miglioramenti in termini di business performance, inclusi:

- riduzione degli sprechi e dei costi;
- guida all'innovazione dei prodotti;
- miglioramento della customer satisfaction;
- adattamento al cambiamento;
- utilizzo dei dati per identificare miglioramenti di business e di processo;
- aumento del knowledge management.

La BI è molto ricercata dalla maggior parte delle organizzazioni. I business leader vogliono ottenere metriche corrette in real-time, collegate direttamente ai processi di business, ai sistemi di dati e che siano conosciute dalle persone. La conoscenza derivata può essere continuamente ottimizzata grazie alla possibilità di ottenere sempre nuove informazioni a valore dalle grandi moli di dati a loro disposizione. La visibilità e la conoscenza che ne derivano aiutano tutti gli aspetti del business apportando soprattutto un miglioramento delle operazioni di produzione evidenziandone gli sprechi e le ragioni delle perdite di produzione o aumentando nel complesso l'efficienza della supply chain globale.

## **Key Performance Indicator (KPI)**

Le aziende nel settore dell'automotive sono coinvolte nelle fasi di design, sviluppo, produzione, marketing e vendita di moto-veicoli. La Business Intelligence crea report di analisi che partono da KPI (Key Performance Indicator, cioè indicatori di performance) che sono direttamente collegati con gli obiettivi del business. Questi KPI possono essere associati, in linea

generale, ad ogni obiettivo di performance, inclusi: l'aumento delle vendite data l'ottimizzazione del livello delle giacenze, la creazione di una lean supply chain, la riduzione del livello delle scorte e dei ricambi, minori tempi per la realizzazione di nuovi prodotti e un aumento della collaborazione con i fornitori. I KPI più utilizzati nel settore dell'automotive sono:

- Efficiency of SCM (supply chain management) operations:
  - inventory turns;
  - total costs;
  - use of technology;
  - document issues;
  - supply chain disruptions;
  - supply/demand imbalances;
  - operating environment;
  - outsourcing and off-shoring production;
  - manufacturing cycle time.
  
- Productivity and competitiveness:
  - production equal to demand (supply vs demand curve);
  - velocity and flexibility;
  - on-time delivery.
  
- Costs:
  - logistics wage;
  - inventory carrying costs;
  - outsourced logistics cost as a percentage of GDP (gross domestic product);
  - cost per unit;
  - opportunity costs;
  - shrinkage;
  - insurance and taxes;
  - total obsolescence;
  - distribution costs (warehousing).

- Monitor business operation:
  - percentage of logistics outsourcing;
  - eliminate/reduction of waste;
  - project time- frame.
  
- Effectiveness of business control:
  - fill rate;
  - forecasting;
  - gross margin;
  - total operating revenue;
  - cost of goods sold;
  - hub (logistic network);
  - just-in-time;
  - lean logistics;
  - radiofrequency identification (RFID);
  - lead-time.
  
- Inventory management:
  - inventory turns ratio;
  - raw materials availability;
  - finished goods in hand;
  - total costs.

## **Importanza delle funzionalità della BI per un'azienda dell'automotive**

La BI è un'importante svolta tecnologica in ambito di memorizzazione e gestione dei dati che utilizza e sfrutta tecnologie avanzate. È innanzitutto un potente frame work analitico per ricavare informazioni utili e per realizzare una supply chain più flessibile e dinamica.

Le seguenti funzionalità della BI possono essere fondamentali per aziende operanti nel mercato dell'automotive:

- Aumentare la visibilità dei e sui dati e permettere un migliore scambio dei dati (data sharing);
- Creare utenti finali information-empowered, cioè mette a loro disposizione molte più informazioni rispetto al passato per permettere loro di prendere decisioni più consapevoli e sicure;
- Migliorare l'efficienza pianificativa grazie a sistemi real time.

### **Aumentare la visibilità sui dati e permettere il data-sharing**

Le grandi aziende del mercato dell'automotive spesso hanno più sistemi ERP (enterprise resource planning). La BI aiuta queste imprese ad ottenere informazioni utili osservando i dati a disposizione da differenti prospettive poiché presenta capacità di reporting superiori rispetto all'ERP grazie all'utilizzo di metodologie di analisi avanzate incentrate sull'obiettivo di permettere ai manager di prendere decisioni migliori e in tempi più rapidi. I report creati con sistemi di Business Intelligence differiscono da quelli dell'ERP perché utilizzano dati provenienti da differenti fonti e aree funzionali quali: domanda, cliente, supply chain e fornitori. Al contrario i report dell'ERP tendono a focalizzarsi su una funzione specifica. L'approccio della BI aumenta la collaborazione e distribuisce le responsabilità.

### **Creare utenti finali information-empowered**

La forza lavoro presente nel settore dell'automotive è sempre più inserita nel modello "knowledge worker". Le decisioni più critiche non saranno più prese da un gruppo d'élite ristretto di manager, ma, al contrario, anche gli utenti qualificati potranno prendere decisioni importanti e attività quotidiane e semplici potranno essere automatizzate. Questo tipo di decisioni, naturalmente, non potranno essere prese arbitrariamente; i lavoratori hanno bisogno di informazioni recenti, significative e consistenti in linea cioè con quanto visto dagli altri membri dell'azienda e con quanto rilevato da analisi predittive precedentemente effettuate.

La Business Intelligence permette di formare information-empowered users dando loro i tool di cui hanno bisogno per raggiungere in maniera efficace ed efficiente i loro obiettivi di performance. Per esempio, grazie alla BI un sales manager può ricevere report giornalieri, mensili o su base annuale grazie ai quali può approcciarsi meglio al raggiungimento degli obiettivi di business (quali l'incremento delle vendite) e agli shareholders. Il reporting è strutturato in modo tale da supportare una collaborazione produttiva sia all'interno che all'esterno dell'organizzazione, e include dashboard riassuntive, drill down dei dati, indicatori e analisi root-cause (tutto questo con l'obiettivo di aiutare i dipendenti a raggiungere i loro KPI).

La disponibilità real-time dei dati nel formato richiesto per favorire il decision-making diventa una sfida per gli utenti operazionali. La BI offre tool innovativi e una struttura agile per generare il tipo di analisi e informazioni di cui i manager hanno bisogno.

### **Migliorare l'efficienza di pianificazione grazie a sistemi real-time**

Migliorare l'efficienza di pianificazione è uno dei benefici più importanti e velocemente ottenibili grazie alla BI. Un'azienda nel settore dell'automotive realizza molte pianificazioni basate su previsioni di mercato, come per esempio vendite, operations e acquisti. Tuttavia, cambiamenti di mercato possono rendere questi piani inefficaci o obsoleti ben prima che i manager possano realizzare di trovarsi di fronte a problemi più o meno gravi come una situazione di fuori stock, un problema del controllo della qualità, o una promozione pubblicitaria per un prodotto non disponibile.

La BI chiude il gap tra pianificazione ed esecuzione offrendo l'utilizzo di dati real-time per la pianificazione e per il processo di revisione della pianificazione.

Non appena un piano è completo, la BI inizia a determinare se il trend attuale sta procedendo con quanto pianificato. Se questo non sta avvenendo, i manager vengono immediatamente informati della discrepanza, non alla



fine del mese o del quadrimestre, in modo tale da poter aggiustare i piani velocemente.

La BI, inoltre, migliora velocità e qualità della preparazione e della revisione dei piani. Se un automaker sta utilizzando un motore di pianificazione avanzato, la BI aiuta ad ottenere il massimo da quest'ultimo migliorando i dati di input e il reporting. Se l'azienda non sta usando un motore di pianificazione avanzato, la BI aiuta a migliorare il ciclo di pianificazione velocizzando e consolidando l'accuratezza dei dati actual delle vendite, della produzione, della logistica, provenienti da varie risorse e sistemi.

## **Come i BIG DATA impattano le vendite nel settore dell'automotive?**

I Big Data comprendono una gran quantità di dati strutturati e semi-strutturati da utilizzarsi per ottenere informazioni a valore per l'azienda. Il loro impatto nel settore dell'automotive è enorme, da veicoli connessi a gestioni più intelligenti dei concessionari, delle lavorazioni e persino delle linee produttive.

Ma la grande maggioranza delle aziende non hanno ancora compreso il ruolo che i Big Data possono giocare nel rendere gli obiettivi di business sempre più vicini, piuttosto che sogni irraggiungibili.

I Big Data possono essere utili in ogni step della supply chain, dalla pianificazione della produzione e dal monitoraggio degli errori fino al marketing, alle vendite e alla customer loyalty.

Tutti sanno che i clienti di ogni industria hanno sempre nuove esigenze e richiedono sempre maggiore attenzione sia dall'industria che dai suoi fornitori. Quando i clienti stanno comprando una nuova macchina o stanno cercando le pastiglie dei freni, l'azienda deve essere in grado di rispondere rapidamente, con le giuste informazioni e sempre con il giusto prezzo.

In breve, la nascita dei Big Data significa che in ogni momento è possibile avere il controllo pieno su ciò che avviene in azienda. Non c'è più la necessità di aspettare l'IT per avere i report sul business o aggiornamenti di stato su un particolare ordine, in quanto tutto questo è possibile in un secondo al semplice click del mouse o direttamente sullo smartphone.

I Big Data hanno spianato la strada per soluzioni che processano ogni singolo pacchetto di dati di business e lo mettono a disposizione in un formato chiaro e semplice da interpretare e comprendere: dashboard, report e grafici che possono fornire tutte le informazioni di cui si ha bisogno.

## **L'impatto dei dati sulla customer retention**

Mentre i dati continuano a regnare nel settore dell'automotive, sempre più spesso ci si rifiuta di vedere come essi potrebbero beneficiare ai clienti. Quando sono utilizzati in modo corretto, i Big Data possono avere un impatto immenso sulla customer loyalty e retention.

Nella loro forma più semplice, i Big Data possono prendere la forma del sistema di CRM dell'azienda. Molte aziende dell'automotive hanno già introdotto software che possono offrire questo tipo di informazioni in pochi secondi.

## **Come i dati possono migliorare la customer retention?**

Con una completa comprensione di chi è il cliente, cosa egli tende a comprare, cosa al contrario non acquista e così via, è possibile raggiungere un livello eccellente di customer care che assicura la fedeltà del cliente stesso. Usando i Big Data, le aziende dell'automotive possono mantenere un'utile e positiva relazione con i propri clienti.

I Big Data offrono inoltre la potenzialità di identificare gli imminenti bisogni dei clienti. È possibile per esempio utilizzare il trend degli acquisti e altri dati dei clienti per predire cosa i clienti potrebbero acquistare in futuro e quale linea di prodotto potrebbe essere la migliore per i loro bisogni.

Anche il marketing può essere uno dei più grandi beneficiari dei Big Data. Questa area aziendale tiene costantemente traccia di questi dati per monitorare e analizzare le condizioni di mercato, le performance delle campagne pubblicitarie e per guidare il ROI. Con la loro analisi, il marketing può estrarre precise informazioni utili per realizzare campagne personalizzate per i propri clienti.

## CAPITOLO 3: DUCATI MOTOR HOLDING

Ducati Motor Holding S.p.A. è una casa motociclistica italiana con sede a Borgo Panigale, Bologna.

### Storia

#### Origini



*Figura 1 Radio Ducati anni '50*

L'azienda nacque nel 1926 per volontà dell'ingegnere Antonio Cavalieri Ducati con il nome di *Società Scientifica Radio Brevetti Ducati*, specializzata nella ricerca e produzione di tecnologie per le comunicazioni radio. Ben presto, grazie ai figli di Antonio Ducati (morto un anno dopo la fondazione), l'azienda cominciò ad affermarsi, per poi spaziare in svariati campi industriali. I figli Adriano, Bruno e Marcello Cavalieri Ducati iniziarono la loro attività con la produzione di un condensatore denominato "Manens", nello scantinato di un edificio situato nel centro di Bologna, in Via Collegio di Spagna. Tra il 1930 e il 1934 la produzione venne ampliata e spostata presso la villa di proprietà della famiglia Ducati.

Nel 1935 venne realizzato lo stabilimento dove hanno attualmente sede la Ducati Motor Holding Spa e la Ducati Energia Spa. La produzione venne ampliata con la realizzazione delle prime apparecchiature radiofoniche, antenne radio, i primi sistemi di comunicazione interfonica, proiettori cinematografici, addizionatrici e rasoi elettrici.

Durante il secondo conflitto mondiale, la Ducati fu obbligata, come tante altre aziende italiane, a convertire la produzione da uso civile a uso militare. In seguito all'armistizio dell'8 settembre 1943, la fabbrica fu occupata dalle truppe tedesche; successivamente venne bombardata e distrutta il 12 ottobre 1944.

## **Dopoguerra**

Nel 1946 nacque il reparto motociclistico, come branca dell'azienda, allo scopo di produrre su licenza il *Cucciolo*, un motore monocilindrico di 48 cc con cambio a due velocità da applicare come propulsore ausiliario ad una normale bicicletta, progettato dall'avvocato Aldo Farinelli.

Nel 1954 avvenne la scissione dell'azienda in *Ducati Meccanica* e *Ducati Elettrotecnica*: la Ducati Meccanica seguì la realizzazione di motoveicoli, mentre la Ducati Elettrotecnica continuò la strada inizialmente percorsa dalla famiglia Ducati. Sempre nel 1954 venne assunto in Ducati Meccanica Fabio Taglioni, il progettista romagnolo che sviluppò per Ducati, tra il 1954 e il 1984, oltre mille progetti di moto e motori, ma soprattutto le tecnologie utilizzate ancora oggi, il sistema desmodromico<sup>5</sup>, il motore bicilindrico e il telaio a traliccio.

---

<sup>5</sup> Il nome "*desmodromico*" deriva dalle parole greche "desmos" = controllata e "dromos" = corsa.

Già dal nome, quindi, si intuisce che si tratta di un sistema di distribuzione nel quale la corsa della valvola è controllata sia durante la fase di apertura che in quella di chiusura. Proprio in quest'ultima fase si racchiude tutta la differenza con il "classico" sistema di distribuzione, nel quale infatti a richiamare le valvole ci sono delle molle, che hanno dei tempi di reazione noti, che anche se rapidissimi non possono essere ignorati. Il sistema desmodromico, cioè a corsa controllata, dispone di due camme per ogni valvola: una che si occupa dell'apertura, come in qualsiasi motore, e un'altra che si occupa della chiusura (al posto delle molle che in questo caso sono assenti) e possono essere sia SOHC che DOHC, ovvero sia monoalbero che bi-albero. Il sistema che abitualmente si abbrevia in "desmo" è spesso usato nei motori che devono avere prestazioni alte, anzi altissime, proprio perché quando il numero di giri del motore diventa particolarmente elevato si creano circostanze critiche, infatti il tempo di reazione delle molle per la chiusura delle valvole limita la possibilità di cercare di raggiungere regimi sempre più elevati.

## **Il controllo statale**

A partire dal 1975 la Ducati passò sotto il controllo dello Stato Italiano: in questo periodo l'azienda si concentrò sulla produzione di motori diesel per conto della controllante ponendo in secondo piano le moto.

Il risultato fu il tracollo delle vendite, passate da 7000 moto vendute nel 1981 a meno di 2000 nel 1984.

Nel frattempo, nei primi anni ottanta, la varesina Cagiva aveva stretto con la casa bolognese un accordo di fornitura dei motori Pantah per poter creare le proprie moto di grossa cilindrata: col tempo i rapporti si strinsero e nel 1985 i fratelli Claudio e Gianfranco Castiglioni acquistarono la Ducati, con l'idea di spostare la produzione delle moto a Varese, per lasciare a Borgo Panigale l'assemblaggio dei motori "Pantah".

## **Gli anni Cagiva**

L'azienda varesina mantenne il controllo fino al 1996, anno in cui il Texas Pacific Group acquistò il 51% delle azioni. Questo decennio fu caratterizzato da forti investimenti nelle competizioni, con i successi nelle gare Superbike a trainare le vendite dell'azienda.

## **Anni recenti**

Il Texas Pacific Group acquistò il rimanente 49% della Ducati nel 1998; l'anno successivo l'azienda mutò denominazione in **Ducati Motor Holding SPA**. Nello stesso anno fu organizzata la prima edizione del World Ducati Week (WDW): l'evento dedicato ai tifosi Ducatisti. In quell'occasione fu inaugurato il nuovo *Museo Ducati*, che raccoglie i veicoli da competizione dal Cucciolo fino alle moderne Desmosedici.

Nel 2006 il marchio Ducati è ritornato in mani italiane con l'acquisto da parte di Investindustrial Holdings.

## 2012

In questo anno la casa automobilistica di Ingolstadt, Audi, controllata dal gruppo Volkswagen, ha raggiunto l'accordo per acquistare Ducati dal fondo Investindustrial per un prezzo di circa 860 milioni di euro.

Le sisters italiane ora appartenenti al *gruppo Volkswagen - Audi* sono:

- Volkswagen Group Italia (Verona);
- Lamborghini (Sant'Agata Bolognese);
- Ducati (Borgo Panigale);
- Italdesign Giugiaro (Moncalieri, Torino).

## 2016

Il 2016 non è un anno come gli altri: coincide con il novantesimo anniversario dell'azienda, celebrazioni incoronate dall'organizzazione del WDW (World Ducati Week).

## Ducati oggi

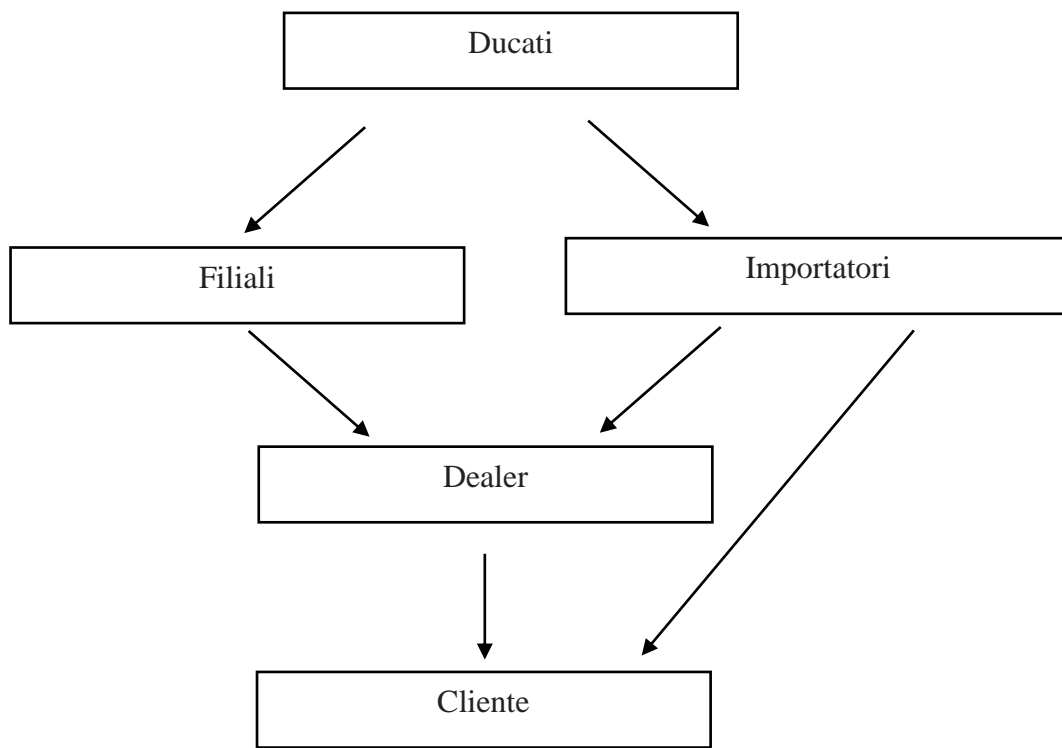
### Ducati Group

All'interno del gruppo Ducati due sono gli stabilimenti produttivi: DMH, Ducati Motor Holding con sede a Borgo Panigale, e DMT, Ducati Motor Thailand. Situazione a parte è quella che riguarda invece la sede brasiliana che acquista pacchetti contenenti componenti di moto che vengono poi assemblate da una società esterna (DAFRA).



Nella sede a Borgo Panigale le moto vengono quasi del tutto completate fatta eccezione per alcune componenti come specchietti, serbatoi ecc. che sono montate solo successivamente in un magazzino ad Anzola dell'Emilia.

Qui le moto vengono ultimate, “vestite” e in seguito imballate e preparate per la spedizione a filiali e importatori.



*Figura 2 Schema stilizzato della gerarchia delle vendite*

I dealer (o concessionari) appartenenti alla rete commerciale di Ducati sono 809 e sono presenti in moltissimi paesi del mondo. Negli stati in cui i dealer sono presenti in maggior numero, sono state costituite le filiali, una sorta di “ufficio centrale” a cui i dealer fanno capo.

Le filiali sono 12: Ducati North America (DNA), Brasile, Svizzera, Germania, Ducati Western Europe (DWE), India, Giappone, Ducati North Europe (DNE), Messico, Canada, UK, Ducati Motor Thailand (DMT).

Una volta spedite, le moto arrivano direttamente ai dealer e non alle filiali poiché queste ultime non sono dotate di magazzini propri. Fanno eccezione DNA, Giappone e India che invece hanno depositi locali. Nei paesi in cui non ci sono filiali come ad esempio in Spagna, Austria, Australia eccetera, la vendita ai dealer avviene attraverso l’intermediazione di importatori (o importers), considerati normali clienti nel flusso logistico Ducati. Questi



ultimi organizzano in modo autonomo le proprie vendite alla sottorete commerciale o direttamente al cliente finale.

Per i contatti con i dealer si utilizza il DCS (Dealer Communication System), cioè un portale web tramite il quale un concessionario può effettuare, modificare e vedere a che punto è la preparazione dei propri ordini siano essi relativi a moto o ad accessori (termine generico usato per indicare le componenti delle moto es: marmitte, specchietti ecc.).

## Vendite Ducati

### Vendite Ducati Motor Holding per categoria di moto

Il grafico sottostante mostra le vendite di moto divise per famiglia, confrontando gli andamenti degli anni 2014 vs 2015 (con focus sui mesi di novembre e dicembre) e 2015 vs 2016 (confrontando i valori relativi al mese di gennaio).

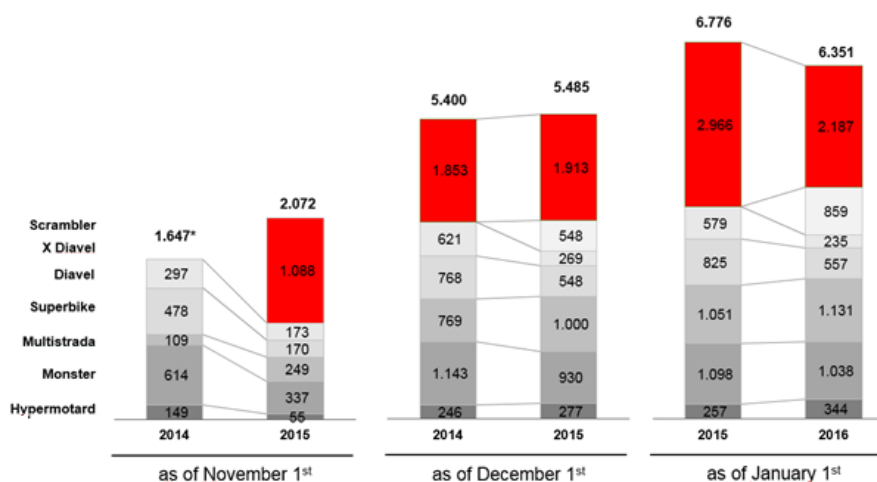


Figura 3 Customer orders portfolio by family

Come si può vedere con estrema chiarezza, l'introduzione sul mercato nel 2015 della moto **Scrambler** (le cui vendite sono rappresentate nell'istogramma attraverso il segmento rosso) si è rivelata una mossa vincente per l'azienda, permettendole di incrementare notevolmente le proprie vendite.



## Le Vendite delle Top 100 Moto

Italia - Mercato Moto: Top 100 giugno 2015				
Rank	Marca	Modello	Segmento	Gen-Giu 2015
1	BMW	R 1200 GS	Enduro	2.039
2	YAMAHA	MT-09 TRACER	Naked	1.831
3	DUCATI	SCRAMBLER	Naked	1.639
4	HONDA	NC 700/750 X	Enduro	1.530
5	BMW	R 1200 GS ADVENTURE	Enduro	1.353
6	YAMAHA	MT-07	Naked	1.352
7	DUCATI	MULTISTRADA 1200	Turismo	976
8	YAMAHA	MT-09	Naked	767
9	BMW	R 1200 R	Naked	607
10	KAWASAKI	Z 800	Naked	599
11	KAWASAKI	VERSYS 650	Enduro	551
12	TRIUMPH	TIGER 800	Enduro	532
13	HONDA	CROSSRUNNER	Turismo	511
14	TRIUMPH	BONNEVILLE	Naked	498
15	HONDA	CB 650 F	Naked	473
16	HARLEY-D.	XL883N IRON 883	Custom	441
17	DUCATI	MONSTER 821	Naked	416
18	BMW	R 1200 RT	Turismo	405
19	MOTO GUZZI	V7 STONE	Naked	398
20	SUZUKI	DL 1000 V-STROM	Enduro	396
21	DUCATI	HYPERMOTARD/SP/STRADA	Supermotard	395

Figura 4 Sintesi delle vendite di giugno 2015 ( [www.motociclismo.it](http://www.motociclismo.it) )

Quando nacque, il modello Scrambler era principalmente destinato al mercato americano. Subì continue modifiche fino al 1968, quando la moto iniziò ad essere commercializzata in Italia e uscirono i veri Scrambler con i motori a "carter largo", prima nella versione 250 e 350 e poi, nel 1969, nella 450.

La produzione venne interrotta nel 1975 fino al ritorno del 2015. Sin dagli anni Novanta, infatti, si parlava in fabbrica di una nuova Scrambler, un modello che vent'anni prima aveva avuto un buon successo in Italia come negli USA. Allora non se ne fece nulla, l'idea rimase tale, accantonata in favore di altri modelli, tutti finalizzati alla sportività e alla tecnologia. Un'accelerazione al progetto Scrambler si è finalmente avuta lo scorso anno, quando si è arrivati alla consapevolezza che il mercato stava cambiando profondamente, con nuovi segmenti in fase di espansione.

Con questo modello DMH è entrata in un segmento che interessa tuttora un'ampia fascia di utenze, senza distinzioni di età perché il gradimento è "trasversale": piace soprattutto ai giovani, ma pure ai motociclisti di lunga esperienza e a qualche nostalgico delle Scrambler che uscirono di scena nel 1975.

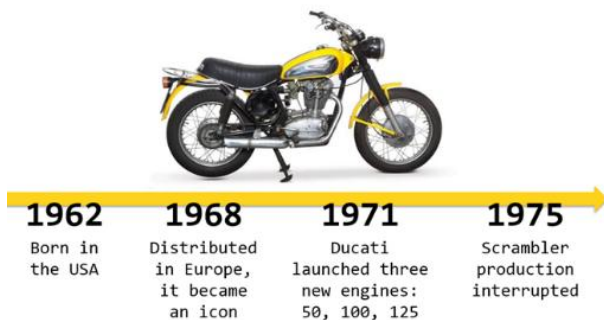


Figura 5  
La storia di Scrambler in sintesi

Figura 6  
I principali modelli Scrambler



# CAPITOLO 4: PROGETTO BUSINESS ANALYTICS

Attualmente Ducati Motor Holding (DMH) utilizza differenti basi di dati e software per la memorizzazione e la gestione dei dati utilizzati per le attività di gestione e controllo.

L'obiettivo del progetto è la creazione di una Data Platform comune a tutta l'azienda.

Con il termine "Data Platform" s' intende una piattaforma IT a supporto dei processi di business per quel che riguarda l'estrazione, l'integrazione tra fonti diverse, l'analisi avanzata e la presentazione grafica dei dati.

L'idea che sta alla base del progetto prevede il classico approccio di un qualsiasi progetto di Business Intelligence/Analytics: migliorare la gestione dei dati al fine di rendere più chiare e immediatamente individuabili informazioni indispensabili all'azienda.



In Ducati, come in ogni azienda modernamente strutturata, quotidianamente si lavora con dati. Questi possono essere:

- Strutturati: conservati in database e organizzati secondo schemi e tabelle rigide;  
Questa è la tipologia di dati più indicata per i *modelli di gestione relazionale* delle informazioni;
- Non strutturati: conservati senza alcuno schema. In questo caso, i sistemi di gestione dei dati utilizzabili sono quelli basati sul modello dell'information retrieval. L'**information retrieval (IR)** (lett:

reperimento delle informazioni) è l'insieme delle tecniche utilizzate per gestire la rappresentazione, la memorizzazione, l'organizzazione e l'accesso ad oggetti contenenti informazioni quali documenti, pagine web, cataloghi online e oggetti multimediali;

- **“Big Data”**: termine usato per descrivere una raccolta di dati così estesa in termini di volume, velocità e varietà da richiedere tecnologie e metodi analitici specifici per l'estrazione di valore.

## **Dimensionamento dati Ducati**

- Moto vendute in un anno;
- fatturato annuo, suddiviso in base ai prodotti Ducati (Apparel, Accessori etc.);
- movimenti di magazzino;
- ordini di acquisto;
- numero e qualità dei fornitori;
- 201 mila followers su Twitter;
- 900 mila Client e prospect su Siebel;
- 3 milioni di like su Facebook;
- 1,4 milioni di MUV (Monthly Unique Visitors) su Ducati.com;
- migliaia di threads su forum, blog, gruppi di discussione online.

I dati enterprise (presenti su diverse piattaforme quali, tra le principali, SAP, Siebel, Hyperion) sono, quindi, circa 3 TB.

## Fonti dati Ducati

Dati Tradizionali		Big Data
Fonti applicative	Fonti non applicative	
Dati strutturati provenienti da applicazioni già implementate in azienda.  Es: <ul style="list-style-type: none"> <li>• fatturato consuntivo;</li> <li>• movimenti di magazzino;</li> <li>• acquistato;</li> <li>• dati customer base;</li> <li>• budget e PR;</li> <li>• livello di servizio.</li> </ul>	Dati provenienti da fonti diverse dalle applicazioni interne. Possono essere più o meno strutturati e complessi.  Es: <ul style="list-style-type: none"> <li>• dati di mercato esterni;</li> <li>• analisi utenti "custom".</li> </ul>	Dati non strutturati che provengono da fonti dati diverse, interne ed esterne all'azienda.  Es: <ul style="list-style-type: none"> <li>• listening web (social media, forum etc.);</li> <li>• Internet of Things;</li> <li>• dati da sito web e app Ducati.</li> </ul>

## Conseguenze all' introduzione di tool di Business

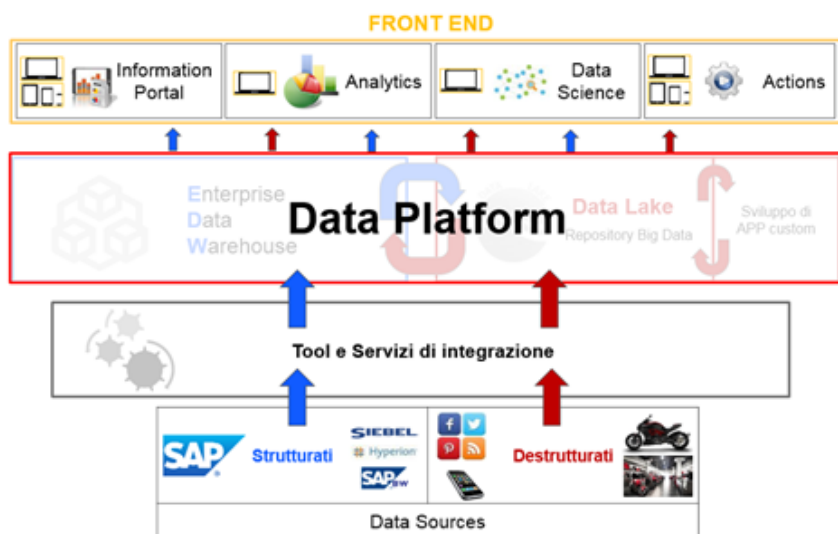
### Analytics

AS IS	TO BE
Report manuali aggiornati periodicamente.	Report automatici sempre online.
Attività business: 90% estrazione e pulizia, 10% analisi.	Attività business: 10% controllo integrità, 90% analisi.
Dettagli non navigabili a partire dalla sintesi.	Drill down dal KPI univoco fino al minimo dettaglio.
EXCEL / Power Point.	Cubi di analisi / Dashboard.
Analisi ad hoc = richiesta all'IT.	Analisi ad hoc = pochi click in autonomia.

Tale progetto permetterà di passare dai dati a informazioni di valore:

- velocemente;
- integrando dati Enterprise con quelli Web/Social;
- garantendo user friendliness;
- su passato, presente e futuro;
- attraverso dati certificati;
- in modo condiviso, ma profilato;
- via report standard o analisi libera;
- su PC, smartphone o tablet.

## Architettura logica desiderata:



L'architettura desiderata è quella classica di un sistema di Business Intelligence: datawarehouse per il salvataggio e la gestione dei dati, strumenti di ETL per la pulizia e il caricamento dei dati e infine strumenti di front end che permettano una visualizzazione chiara, immediata ed esplicativa dei dati.

## Selezione dei vendor

La prima fase di questo progetto di sviluppo consiste nella scelta dei software a supporto della futura piattaforma.

## Parametri

La selezione dei vendor è avvenuta tramite una valutazione dei prodotti da loro offerti inerenti alle diverse aree della piattaforma di Business Intelligence. I prodotti sono stati suddivisi secondo le seguenti categorie:

- front end;
- Data Warehouse;
- ETL;
- Big Data.

I parametri utilizzati per confrontare economicamente le diverse offerte sono: data storage, user number e il relativo valore economico per ETL, EDW, Data Lake, Front end Enterprise Reporting, Front end Self Service e Front end Data Science (Predictive) su un orizzonte temporale di tre anni.

Si è valutata inoltre la possibilità di suddividere gli utenti in tre diverse categorie:

1. *Power Users*: utenti che possono creare report e dashboard in modalità self-service e condividerli in autonomia;
2. *Interactive Users*: consultano report e dashboard con la possibilità di navigare il dato ed effettuare modifiche personali;
3. *Consumers*: consultano report e dashboard senza modificarne i contenuti.

## **Vendor valutati e prodotti proposti**

In una prima fase di scouting Ducati ha interpellato numerosi vendor di prodotti e servizi legati alla Business Analytics. Tra questi:

- alcuni (come Microsoft o SAP) si propongono sul mercato come fornitori end-to-end, in grado di soddisfare in un'unica soluzione verticale tutte le esigenze aziendali in termini di estrazione, modellazione e fruizione dei dati;
- altri (come Tableau o Pivotal) si propongono sul mercato come specialisti settoriali, esperti di un ambito specifico e fornitori di una soluzione “best in class” nel proprio ambito.

Una lista non esaustiva dei vendor interpellati nell'ambito dell'indagine di mercato comprende:

- SAP: fornitore del sistema ERP di DMH;
- SAS: tra le piattaforme best in class per l'analisi statistica avanzata;
- Pivotal: tra le piattaforme best in class per l'integrazione tra i dati destrutturati e quelli relazionali;



- Microsoft: tra le piattaforme best in class per la Business Intelligence, anche date le forti esternalità ottenibili in termini di integrazione con la suite Office;
- Microstrategy: tra le piattaforme best in class per la modellazione del metadato aziendale.

## Conclusioni

La fase di analisi è proseguita incentrandosi solo su due vendor selezionati (Microsoft e SAP) valutando le loro proposte sia in termini di Cloud che On Premise.

Per la scelta finale si è anche tenuto conto del quadrante di Gartner:



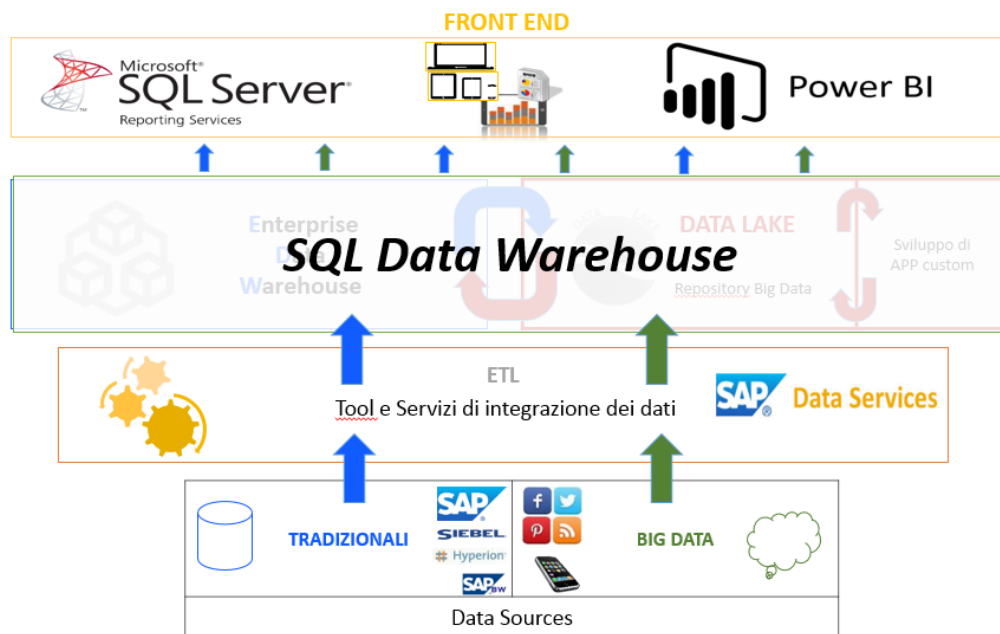
In grigio sono riportate le posizioni occupate dalle aziende l'anno scorso, in blu le nuove collocazioni individuate da Gartner. Da questa analisi è facilmente visibile che Microsoft è stata l'unica azienda che è riuscita a mantenere la propria posizione nel quadrante "leader" mentre al contrario SAP e SAS sono state trasferite nella classe dei "visionari" segnalando in questo modo un calo nelle abilità di esecuzione dei loro prodotti.

La conclusione di questa seconda fase di analisi dei vendor si è conclusa con una certa propensione nei confronti di Microsoft.

Alla fine, in seguito a successive analisi e colloqui con i vendor, si è optato per una soluzione “mista”:

- come ETL si è deciso di adottare SAP Data Services, che verrà installato su una macchina virtuale;
- come Enterprise Data Warehouse è stato scelto Microsoft SQL Server Parallel Data Warehouse;
- come strumento di front end (reportistica e predictive analysis) si è optato per l’utilizzo di Power BI e Microsoft SQL Server reporting services (che verrà installato su una seconda macchina virtuale);
- il tutto avverrà in cloud appoggiandosi al cloud service Azure di Microsoft.

Questo è quindi lo schema finale dell’infrastruttura software scelta:



## Fase di Analisi

Il progetto è iniziato con la fase di analisi il 15 Novembre 2015 con la realizzazione del primo assessment.

Con il termine *assessment* si intendono riunioni di gruppo alle quali partecipano alcuni membri dell'area aziendale di volta in volta coinvolta con l'obiettivo di indagare sulla situazione aziendale attuale a livello di reportistica, raccolta e gestione dei dati.

## Aree aziendali coinvolte

Codice Area	Nome Area
CDG	Controlling
CEO	CEO
CFO	Direzione AFC
CRED_MAN	Credit Management
CRM	CRM
DAFRA	Produzione Brasile DAFRA
DIR_PUR	Direct Purchases
DMT	Ducati Motor Thailand
DWE	Ducati West Europe
GSPM_HSE	General Serv. Plant Mant. And HSE
HR_CR	HR and Corporate Responsibility
IND_PUR	Indirect Purchases
LOG	Logistics
MHT	Process Engineering
MKT	Marketing
MRP	MRP
ND	Network Development
PRC	PR and Communication
PROD_MKT	Product Marketing
QUA	Quality
SAL_IMP	Sales Importers
SALES	Sales
SALES_ST_PL	Sales Strategies and Planning
SERV	Service
TAX_ACC	Tax & Accounts
TRE	Treasury
VEH_ENG	Vehicle and Engine Assembly

Fino a marzo 2016 sono stati realizzati circa 27 incontri tra self assessment e interviste e 4 Big Data Workshop aventi l'obiettivo di sensibilizzare le aree sull'utilizzo e l'analisi dei dati destrutturati utili per migliorare la Brand Reputation, la comunicazione, i processi di vendita sul territorio, i servizi post-vendita e la pianificazione delle vendite e dei flussi logistici.

In preparazione all'intervista di assessment era stato precedentemente richiesto ad ogni area di raccogliere le attuali esigenze informative compilando in maniera guidata appositi documenti Excel.

La riga seguente è un esempio di compilazione							
Vendite mensili	Comparazione del venduto dell'anno in corso con quello dell'anno precedente	Mario Rossi	Sales Bikes	Analisi Venduto	4	Venduto del Mese, Ordinato	Sell in %
Titolo report (Per ogni report censito è necessario allegare il file Excel \ PDF identificato con il suo nome)	Descrizione report	Owner del report	Ente di riferimento (selezione)	Principale processo supportato dal report	Importanza del report (1 bassa - 4 alta)	Misure principali del report	Eventuali KPI calcolati
<b>Informazioni report esistenti</b>							
<b>Informazioni report desiderati</b>							

Figura 1 Esempio di documento da compilare dalle singole aree aziendali prima degli assessment

## Assessment

Gli obiettivi degli assessment sono stati:

1. individuare e analizzare le esigenze informative attuali e desiderate;
2. individuare le modalità di misurazione delle performance;
3. individuare le inefficienze in termini di FTE<sup>6</sup> dedicati al reporting, duplicazione di attività dovute ad incomprensioni o a carenze informative sugli attuali sistemi;
4. individuare le criticità sui sistemi alimentanti (dati non disponibili, dati non affidabili, dati in ritardo);
5. valutazione di effort e benefici relativi alla realizzazione di ogni report.

Gli output degli assessment, opportunamente rielaborati dovranno mettere i process owner nelle condizioni di prendere le seguenti decisioni:

- roadmap di progetto;

<sup>6</sup> FTE: Full Time Equivalent. È l'unità di misura dell'impegno lavorativo necessario a realizzare un'attività. Un'attività che impiega 1 FTE impiega costantemente 8 ore di lavoro al giorno.

- interventi sui sistemi alimentanti;
- scelta tecnologica.

Gli obiettivi dell'azienda relativi alla realizzazione del progetto sono:

- automazione delle attività non a valore aggiunto;
- valorizzazione del patrimonio informativo disponibile;
- soddisfazione del business e aumento della potenzialità analitica;
- riduzione dei costi per efficienza e semplificazione.

## **Sviluppi successivi**

Dopo la realizzazione degli assessment, partendo da base consulenziale, ho personalmente lavorato alla realizzazione di un data base contenente tutti i dati raccolti e ho successivamente sviluppato un sistema per classificare le richieste degli utenti. L'obiettivo di tale classificazione è di individuare un ranking da utilizzarsi poi come linea guida per gli sviluppi successivi del progetto.

Sono stati schedulati più di 489 report<sup>7</sup> (attualmente realizzati in azienda o desiderati dalle varie aree), corredati da informazioni aggiuntive relative come, per esempio, il dipartimento in cui il report è utilizzato o richiesto, la fonte dati utilizzata, la dimensione di analisi, l'ammontare di tempo necessario per redigere tali report e la frequenza con cui deve essere presentato, le difficoltà incontrate dai diretti interessati in fase di realizzazione, il rischio di commettere errori durante la stesura, il numero di destinatari a cui viene consegnato il report e la sua importanza.

A questo punto abbiamo ritenuto necessario, ai fini del progetto, associare tali report ad uno o più interventi, intesi come insiemi di attività del progetto di Business Intelligence volti a rendere fruibili informazioni che ora non sono direttamente ricavabili dai dati aziendali e quindi ad incrementare il

---

<sup>7</sup> È un documento generalmente costituito da visualizzazioni tabellari e grafiche.

livello di affidabilità del sistema di reportistica attualmente adottato. Abbiamo in questo modo ottenuto un database contenente 218 interventi.

Lo step successivo ha riguardato la necessità di realizzare un ranking degli interventi per individuare una roadmap di progetto, cioè una linea guida, costituita dalle date di inizio e di rilascio dei vari interventi fino alla realizzazione completa di tutti i flussi dati previsti.

## **Realizzazione roadmap di progetto**

### **Fase 1**

Per la definizione della roadmap di progetto abbiamo deciso di definire e calcolare impact ed effort per intervento da realizzare.

*Impact* = “beneficio”, impatto positivo rilevabile dalle singole aree aziendali in seguito al miglioramento/ automazione dei report in esse compilati o in seguito alla realizzazione di report al giorno d’oggi non presenti in Ducati.

*Effort* = costi in termini di tempo e denaro per la realizzazione dell’intervento calcolati tenendo conto di:

- volume di dati considerati dall’intervento;
- complessità della fonte dati;
- complessità delle logiche di calcolo;
- numero di dipartimenti coinvolti.

### **Step 1**

In un primo momento mi sono concentrata sui singoli report per i quali ho calcolato gli impact considerando le seguenti variabili:

- importanza del report;
- difficoltà di realizzazione;
- rischio presenza errori;

- FTE necessari per la stesura del report (calcolati come “tempo speso ore” diviso la frequenza di aggiornamento del report stesso);
- n
- u

REPORT_COD	ESISTENTE_O_DESIDE	FTE_ANNO	IMPORTANZA	DIFFICOLTA	RISCHIO_PRESENZA_ERRORI	NUM_DESTINATARI
CDG-001	Esistente	0,125	4	2	2	2
CDG-009	Desiderato	0	3	1	1	30
HR_CR-124	Desiderata	0	1	1	1	1
IND_PUR-005	Desiderato	0	1	1	1	1
QUA-021	Esistente	0,075	1	1	1	1
QUA-022	Esistente	0,075	4	1	1	1
QUA-023	Esistente	0,05	4	1	1	1

Figura 2 Porzione del database nella quale sono riportati i valori di impact di alcuni report realizzati in azienda.

I valori delle prime tre voci ci sono stati forniti direttamente dalle aree aziendali durante gli assessment e hanno un valore numerico compreso tra 1 e 4. Parallelamente, per poter utilizzare la medesima scala numerica, ho ritenuto opportuno clusterizzare i valori relativi al numero di destinatari e agli FTE per associare poi a questi un valore da 1 a 4.

Attraverso poi la scrittura di query in linguaggio SQL è stato possibile visualizzare e calcolare l’impatto totale di ogni report dato dalla sommatoria dei valori sopra citati diviso il numero di interventi per report (il ragionamento che è stato fatto alla base è che: il beneficio di un report è effettivamente ottenuto dall’azienda solo quando tutti gli interventi ad esso correlati vengono compiuti. Premesso ciò l’impatto del report deve essere diviso per il numero di interventi che permettono di ottenere il 100% del beneficio).

Di seguito riporto la query per il calcolo dell’impact per singolo report:

```
“SELECT dipartimento_cod, t.report_cod, Max(T.Importanza) AS
punteggio_importanza, Max(t.rischio_presenza_errori) AS
punteggio_rischio_presenza_errori, Max(t.difficolta) AS
punteggio_difficolta_realizzazione, Max(C.punteggio) AS
punteggio_destinatari, Max(D.punteggio) AS Punteggio_FTE, Count
(T.report_cod) AS num_interventi,
```

(Max(importanza)+Max(rischio\_presenza\_errori)+Max(difficolta)+Max(C.punteggio)+Max(D.punteggio))/num\_interventi AS beneficio

FROM ana\_report as t, ass\_report\_intervento as b,  
v\_ass\_punteggio\_destinatari as c, v\_ass\_punteggio\_fte as d

WHERE (((t.report\_cod)=[b].[report\_cod] and  
(t.report\_cod)=[c].[report\_cod] and  
(t.report\_cod)=[d].[report\_cod]))

GROUP BY t.report\_cod, dipartimento\_cod”

REPORT_COD	BENEFICIO_DA_REPORT
CRED_MAN-001	3,20
CRED_MAN-002	3,67
CRED_MAN-003	2,20
CRED_MAN-004	4,00
CRED_MAN-005	2,00
CRED_MAN-006	1,33
CRED_MAN-007	1,33
CRM_ND-008	4,50
CRM_ND-013	1,75
DWE-001	1,43
DWE-002	1,67
ND-001	1,67
CFO-001	1,78
CFO-002	0,81

Figura 3 Tabella riportante l'impact totale per singolo report

A questo punto, è saltato immediatamente all'occhio che, a fronte di una valutazione dell'impacto totale registrato a livello di singola area, vi era un forte sbilanciamento a causa del fatto che alcuni dipartimenti hanno richiesto molti più report rispetto ad altri. Ad esempio l'area HR, con più di 120 report, avrebbe avuto, a livello di impact totale un punteggio molto più alto rispetto al Credit Management al quale sono assegnati solo 7 report. Seguendo quindi il ragionamento fatto in precedenza, nella roadmap di progetto sarebbero stati inseriti come primi interventi da effettuare quelle che coinvolgevano l'area HR.

Per normalizzare i risultati ho allora deciso di raggruppare i report calcolando il loro impact per dipartimento, clusterizzare i risultati così ottenuti ed assegnar loro un punteggio da 0 a 11 (impact totale per



dipartimento basso = punteggio alto e, viceversa, punteggio impact totale alto = punteggio basso).

Questo punteggio è stato poi sommato all'impact per singolo report calcolato in precedenza, ottenendo il valore finale di IMPACT PER REPORT.

REPORT_COD	BENEFICIO_DA_REPORT	PUNTEGGIO_DA_DIPARTIMENTO	TOTALE_BENEFICIO_REPORT
DWE-001	1,43	11,00	12,43
TRE-005	2,00	9,00	11,00
GSPM_HSE-003	2,20	7,00	9,20
SAL_IMP-004	3,50	7,00	10,50
CDG-001	4,00	2,00	6,00
MKT-008	1,33	1,00	2,33
HR_CR-012	13,00	0,00	13,00
HR_CR-015	10,00	0,00	10,00
QUA-035	2,25	0,00	2,25
VEH_ENG-001	1,33	0,00	1,33

## Step 2

A questo punto ho trasferito l'analisi a livello di impatto per intervento. Ho calcolato l'impatto totale per intervento raggruppando i report interessati dall'intervento stesso e moltiplicando i risultati ottenuti per un peso strategico assegnato agli interventi in base alle richieste emerse durante le interviste con CEO e CFO.

INTERVENTO_COD	BENEFICIO_NON_PESATO	BENEFICIO_PESATO
Acquistato / Infocord	247,01	494,02
Ageing Scaduto	42,07	42,07
Allocazioni Bikes a Dealer	39,25	78,50
Analisi di clima	39,00	39,00
Assegnatari auto	4,00	4,00
Assenze DAFRA	8,42	8,42
Assunzioni	47,00	47,00
ATP (Disponibilità a magazzino Ducati)	32,25	128,99
Audit garanzia	9,00	54,00
Avanzamento progetti	5,00	5,00
Azioni Corporate Responsibility	4,00	4,00

### Step 3

Al passo successivo mi sono concentrata sugli effort.

Come detto in precedenza le variabili analizzate sono:

- volume dati a cui è stato dato un punteggio da 1 a 5;
- complessità della fonte dati;
- complessità delle logiche di calcolo;
- numero di dipartimenti coinvolti, valori che sono stati prima clusterizzati e poi a questi è stato assegnato un punteggio da 1 a 4.

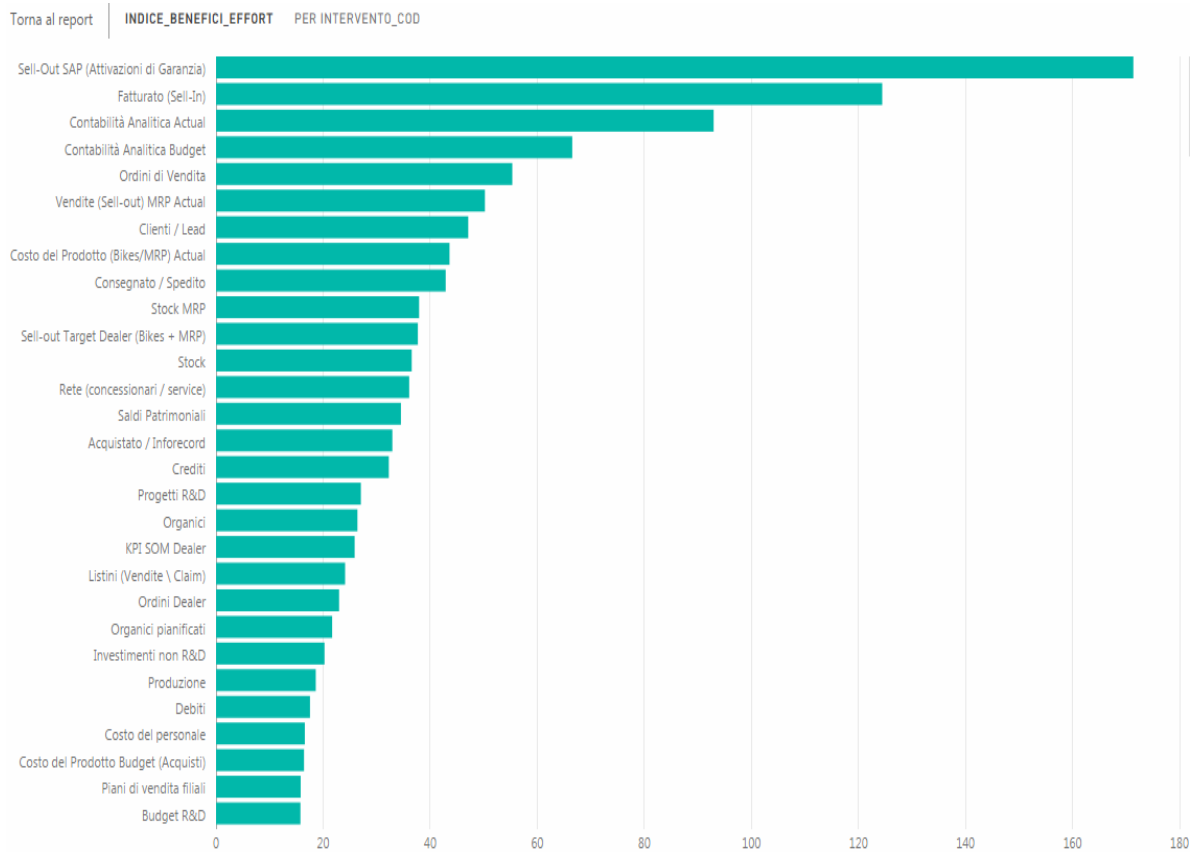
In questo caso i punteggi **non** sono stati assegnati seguendo la medesima scala in quanto abbiamo ritenuto alcune variabili più costose di altre in termini di effort.

INTERVENTO_COD	VOLUME_DATI	COMPLESSITA_LOGICHE_DI_CALCOLO	COMPLESSITA_FONTE_DATI
Budget Acquisti Ricambi	2	2	2
Forecast Acquisti	1	2	2
Gare D'appalto	1	1	6
Prezzi materie prime	1	2	8
Azioni Corporate Responsibility	1	1	10
Bollettini Service	2	1	8
Cessioni di credito	2	6	4
Contatti Work with us	1	2	8
Contratti cartacei	2	1	10
Contratti QA	1	4	4

Alla fine siamo quindi arrivati ad avere valori di impact ed effort per ogni intervento.

INTERVENTO_COD	BENEFICIO_NON_PESATO	BENEFICIO_PESATO	EFFORT
Acquistato / Inforecord	247,01	494,02	15,00
Ageing Scaduto	42,07	42,07	12,00
Allocazioni Bikes a Dealer	39,25	78,50	12,00
Analisi di clima	39,00	39,00	11,00
Assegnatari auto	4,00	4,00	11,00
Assenze DAFRA	8,42	8,42	11,00
Assunzioni	47,00	47,00	11,00
ATP (Disponibilità a magazzino Ducati)	32,25	128,99	12,00
Audit garanzia	9,00	54,00	8,00

A questo punto utilizzando strumenti quali Power BI e Tableau ho rappresentato graficamente i risultati ottenuti ordinando per valori decrescenti dell'indice Impact / Effort.



## Fase 2

Questa prima classificazione degli interventi ci è stata utile per la realizzazione dello step successivo: l'individuazione di fasi (date dall'unione di interventi) e macro-fasi (nate dal raggruppamento di più fasi) e quindi della roadmap di progetto.

### *Fasi*

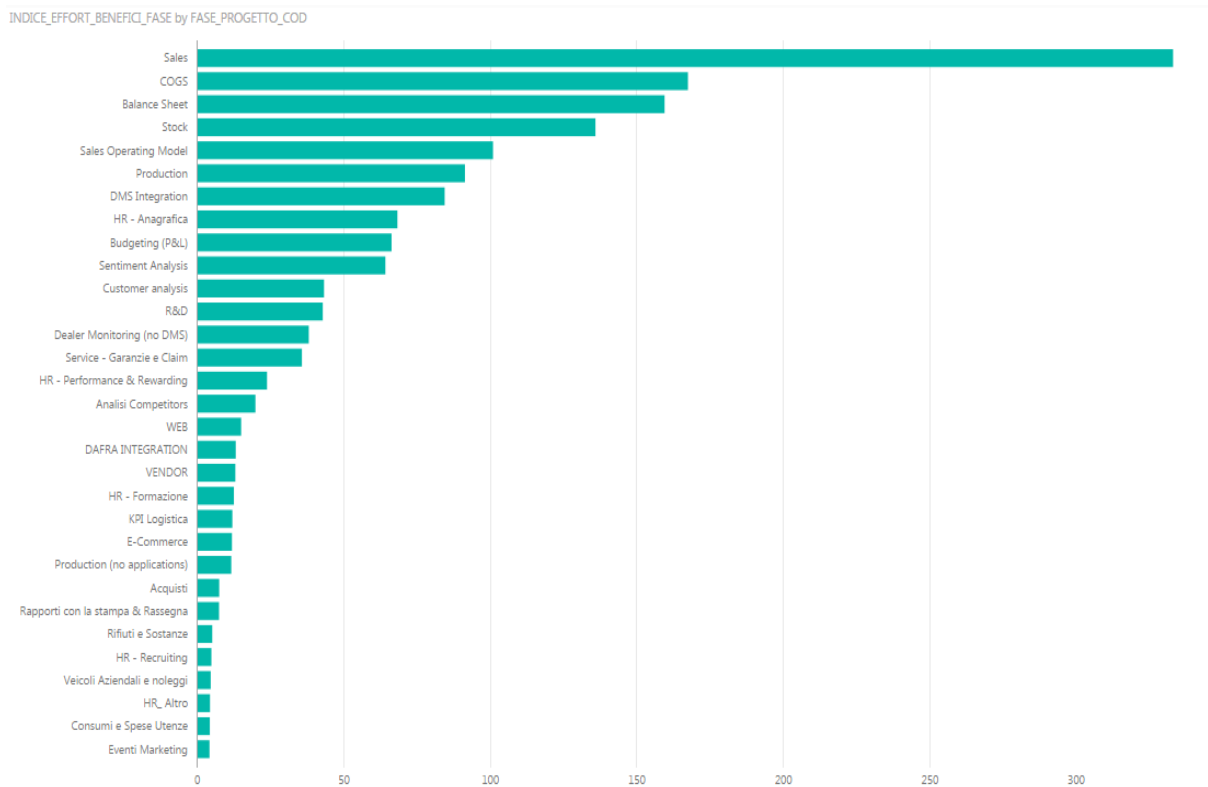


Figura 3 Grafico ranking fasi sulla base dell'indice benefici/effort

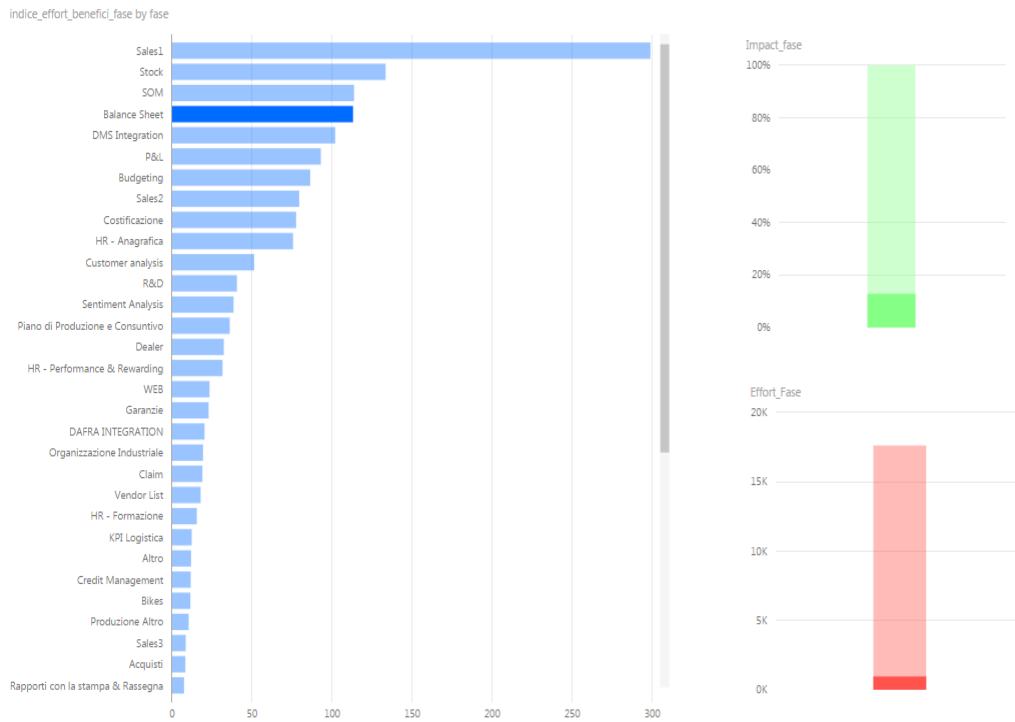


Figura 4 Dashboard riassuntiva con ranking fase e % impact ed effort

## Macro fasi

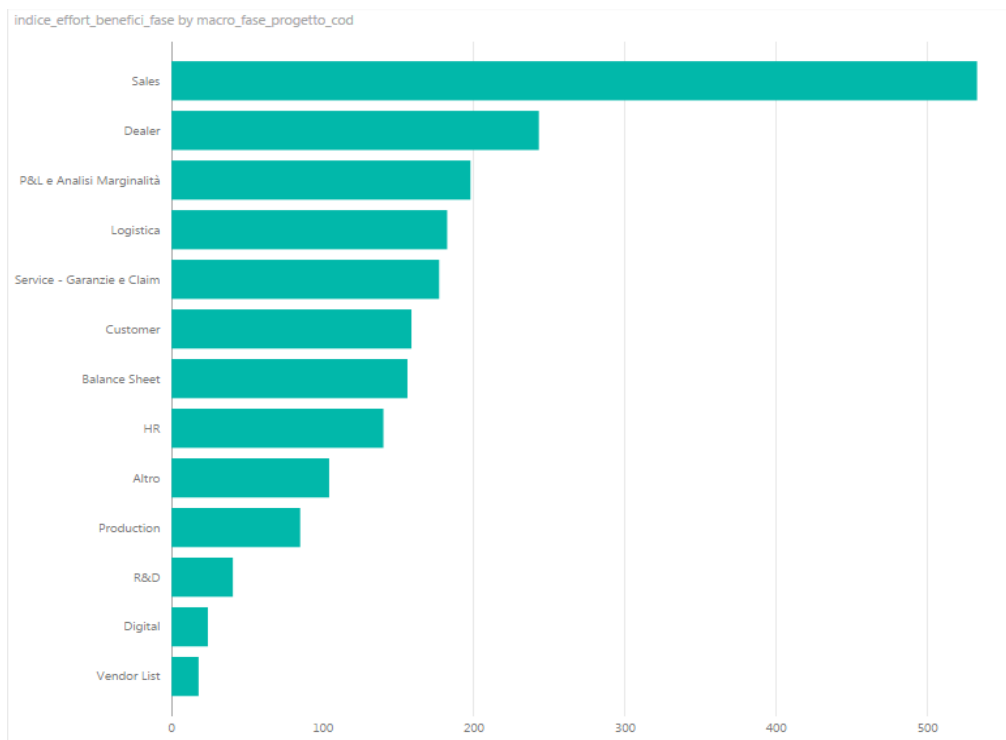


Figura 5 Grafico della classificazione delle macrofasi in base all'indice impact/effort

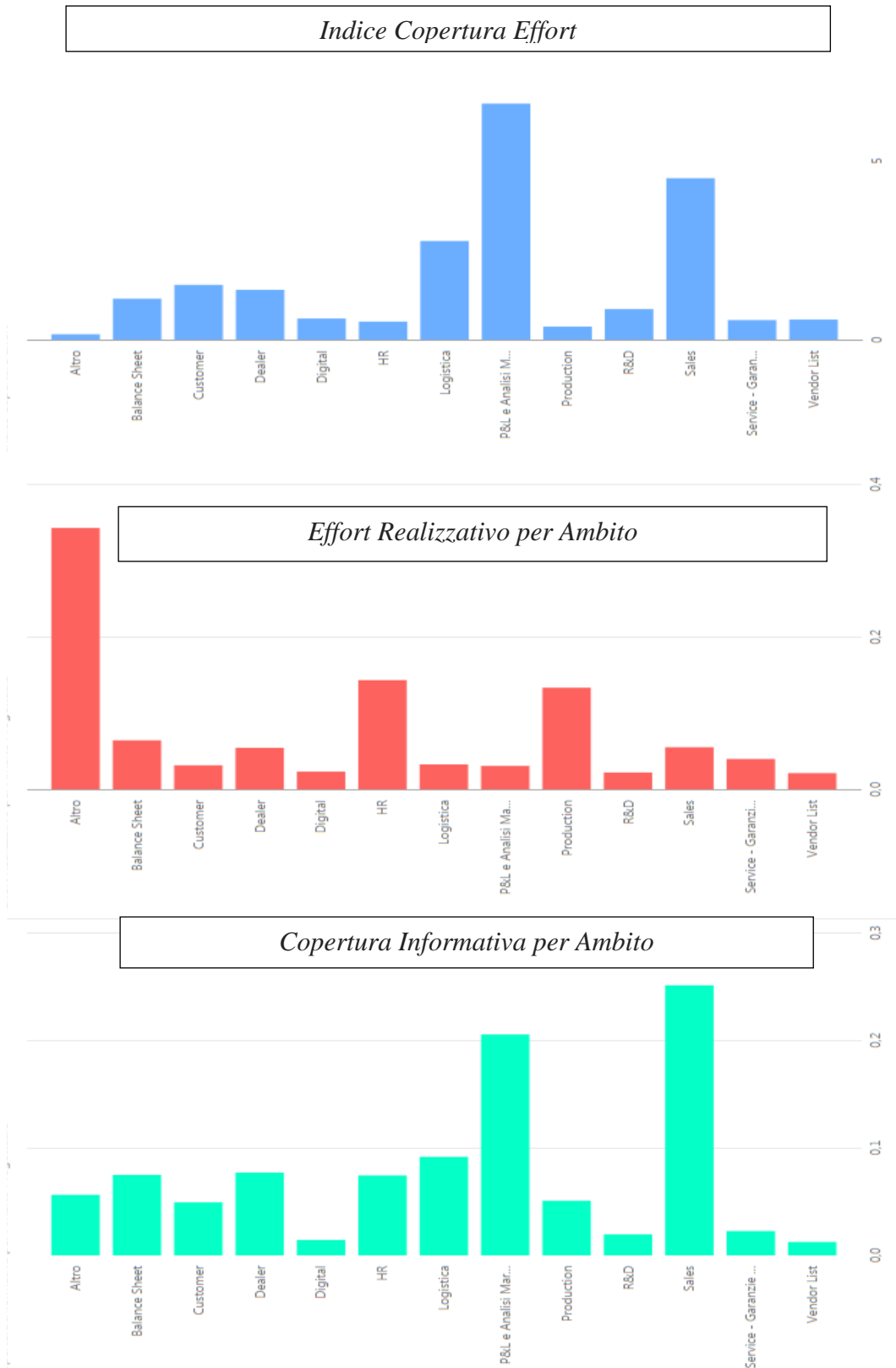


Figura 6 Grafici con focus su Macro aree e relativi valori di Impact ed Effort

## Dashboard riassuntiva

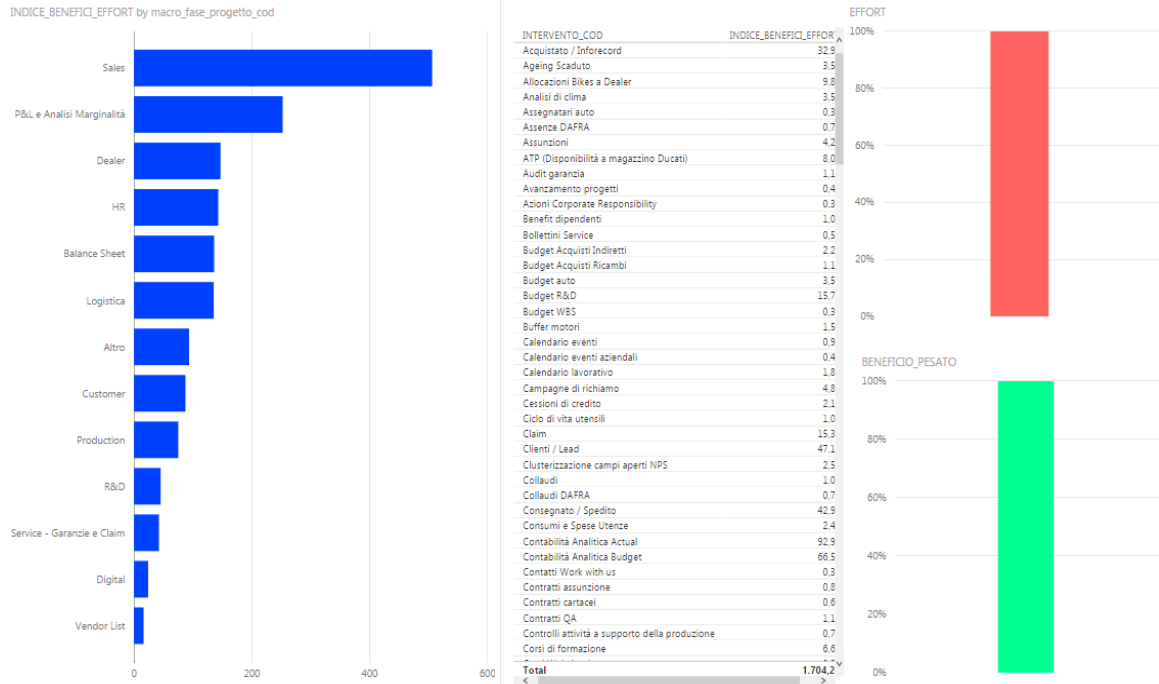


Figura 7 schermata raffigurante i dati generali contenuti nella dashboard

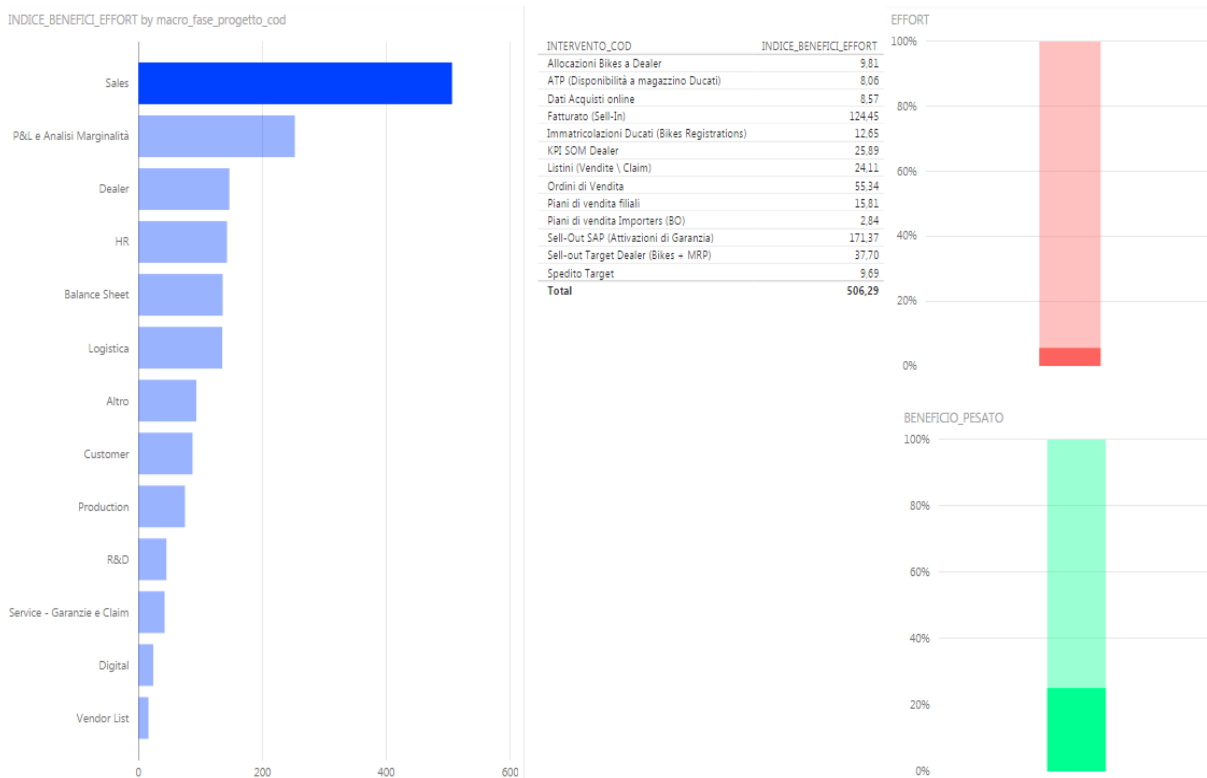


Figura 8 Schermata raffigurante le funzionalità della dashboard con focus sulla macro area di intervento "sales"

Attraverso questa dashboard realizzata con Power BI è possibile selezionare alcune macrofasi e vedere a quali fasi e interventi corrispondono e quindi la percentuale di Impact ed Effort che permettono di raggiungere in seguito alla loro realizzazione.

## **Roadmap di progetto**

Da quanto emerso nelle fasi di analisi precedenti i primi interventi da effettuare sono quelli relativi alle macrofasi “Sales” e “P&L Analisi Marginalità” in quanto sono quelle che presentano i migliori valori dell’indice Impact/Effort e così via.

Tenendo conto di questi risultati e del core business aziendale si è deciso di iniziare concentrandosi sugli interventi relativi a sell-in e sell-out sia Bikes che CAS (Commercial After Sales, inteso come accessori, apparel etc) sia in termini di valore che di quantità.

## **Sales**

Il passo successivo del progetto di Business Analytics consiste nella preparazione della piattaforma hardware e software desiderata e nell’organizzazione di tavoli tematici nei quali discutere di concetti relativi al prodotto dell’azienda Ducati, analizzato a diversi livelli di dettaglio, e al mercato. L’obiettivo è quello di allineare le diverse aree aziendali creando una base d’azione e un linguaggio comune per una migliore integrazione e comprensione dei dati.

Si è deciso di procedere percorrendo tre differenti vie parallele:

1. estrapolare dalla reportistica attualmente realizzata in azienda le diverse declinazioni d’ analisi a cui sono sottoposte le dimensioni “*prodotto*” e “*geografia del mercato*”;
2. realizzare incontri con le principali aree aziendali coinvolte dagli interventi sopracitati (in particolare sales e sales planning);
3. raccogliere informazioni utili da altri progetti connessi all’area sales per ottenere una visione completa sulle dinamiche dell’area.



## **Dimensioni di analisi: prodotto e mercato (1)**

Come detto in precedenza, analizzando la reportistica attualmente realizzata in azienda, è stato rilevato che le dimensioni di prodotto e mercato vengono utilizzate e declinate in modi diversi a seconda delle aree aziendali analizzate e dei report realizzati. Per unificare l'utilizzo di tali dimensioni di analisi si è pensato di organizzare dei tavoli tematici nei quali le persone interessate vengono letteralmente messe attorno ad un tavolo con l'obiettivo di trovare una linea d'azione comune.

Analizzando la reportistica attuale ho rilevato che alcune delle declinazioni della dimensione prodotto sono:

- linea di produzione;
- modello/versione;
- standard nazione;
- telaio;
- segmento;
- famiglia;
- colore;
- pacchetti;
- cilindrata;
- codice materiale.

Per quel che riguarda la geografia del mercato è emerso che l'attenzione è invece posta su:

- gerarchia geografica;
- importers;
- plant;
- dealer;
- filiali.

Alcune delle incongruenze rilevate riguardano ad esempio i nomi utilizzati per indicare la sede di Borgo Panigale (DMH/Italy), in alcuni report gli importatori Spagna e Portogallo sono considerati insieme in altri, invece, separatamente. Problema simile è quello rilevato per l'importatore APAC che in certi casi comprende anche la Cina in altri no. Relativamente al prodotto è stato riscontrato che in alcuni report i dati relativi alle moto erano rappresentati seguendo una gerarchia 'segmento-famiglia-gruppo-modello', in altri alcuni di questi livelli non erano presenti con conseguente asimmetria nei valori numerici in essi contenuti.

### **Sell-in e sell-out: il mondo sales all'interno di Ducati (2-3)**

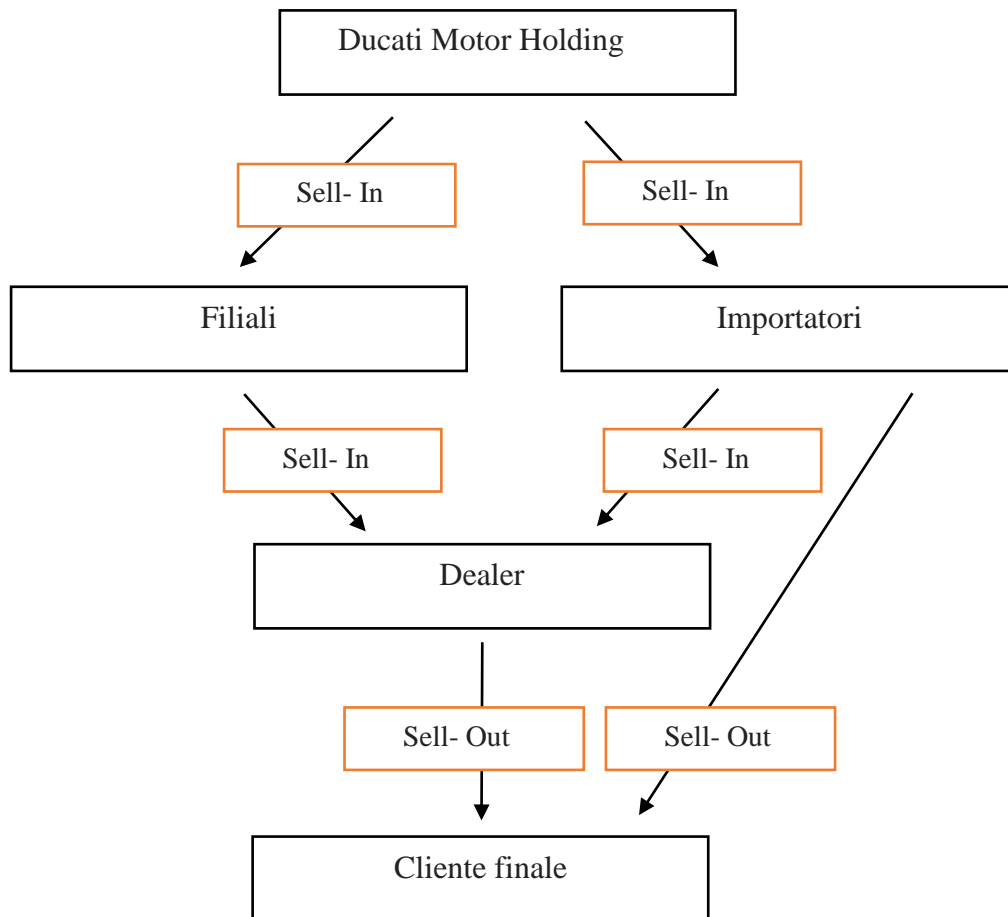
Le vendite di Ducati riguardano diversi prodotti che vanno al di là del core business delle moto.



*Figura 9 Prodotti offerti da Ducati*

Concentrandosi esclusivamente sul mondo bikes, si può parlare di sell-in e di sell-out:

- Sell-in: con questo termine si intende la vendita di moto tra gli impianti produttivi (Ducati Motor Thailand e Ducati Motor Holding) e importers o dealer, quindi si tratta di flussi inter- company;
- Sell-out: si intende la vendita di moto ai clienti finali.



Una moto viene considerata effettivamente venduta, e quindi registrata come sell-out, se ha associata un'immatricolazione e un'attivazione di garanzia.

Le attivazioni di garanzie vengono effettuate direttamente dai dealer tramite DCS (Dealer Communication System)<sup>8</sup> ogni volta che ad una moto viene

---

<sup>8</sup> DCS (Dealer Communication System): portale web tramite il quale un concessionario (dealer) può effettuare e modificare gli ordini di moto, accessori eccetera e vedere il work in progress dell'ordine stesso.

associato un cliente che ne ha fatto richiesta. Ci sono diversi tipi di garanzie che possono essere attivate:

- Garanzia standard: è il tipo di garanzia contrattuale con durata di 24 mesi;
- Ever Red (presente solo per alcuni paesi): è una garanzia, di durata pari a 12 o 24 mesi a seconda della volontà del cliente, che va ad aggiungersi alla garanzia contrattuale iniziale.

## Claim

Sul DCS avviene anche la gestione delle claim. Quando un cliente ha un problema con la moto, contatta il dealer il quale avvia una procedura di validazione della claim (visibile non solo su DCS, ma anche su SAP). Una volta ricevuta l'autorizzazione a procedere, il dealer effettua la riparazione o la sostituzione.

## DCS- Dealer Communication System

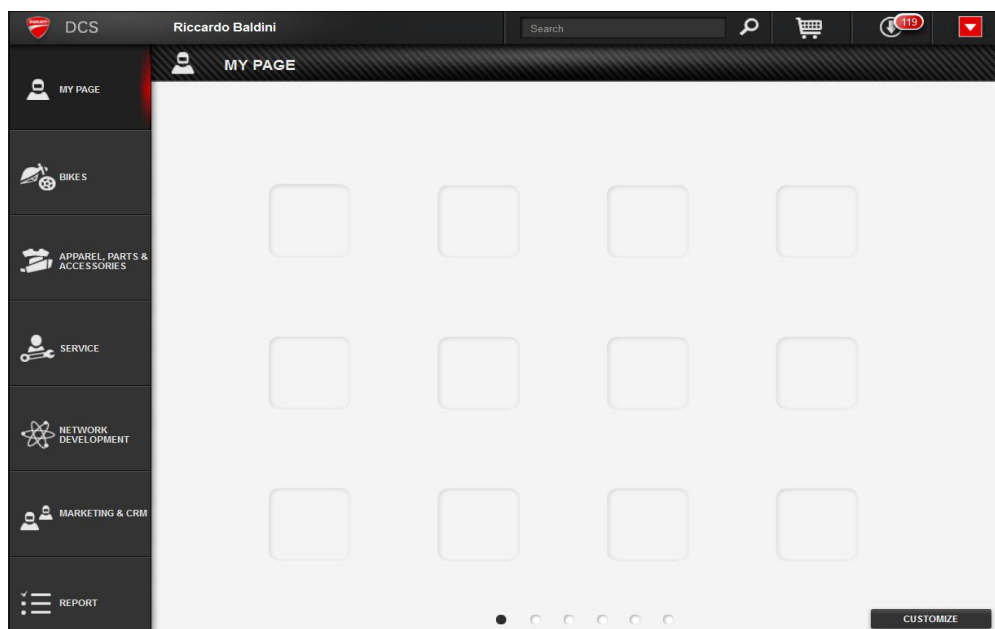


Figura 10 Schermata iniziale del DCS

Il cuore della gestione dei rapporti con i dealer e dei dati relativi alle vendite bikes e accessori è il DCS. Questo portale web permette ai dealer di effettuare e/o modificare i propri ordini, di registrare attivazioni di garanzia e immatricolazioni del veicolo e di vedere lo stato di avanzamento del proprio ordine.

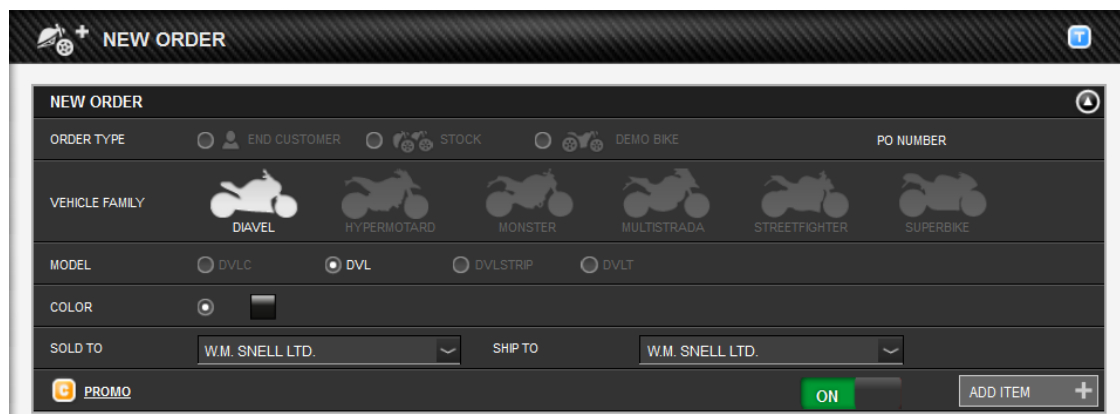


Figura 11 Schermata per effettuare un nuovo ordine da DCS

Attualmente in azienda è in fase di realizzazione un progetto che ha l'obiettivo di migliorare quest' ultima funzionalità del sito al fine di rendere più affidabile l'ATP (available to promise) da intendersi in generale come miglioramento della precisione della data di prevista consegna che viene comunicata ai dealer. Inoltre il perfezionamento del DCS permette alla sede DMH di avere maggiore visibilità sulle moto in stock presso i propri dealer e le moto effettivamente vendute al cliente finale (al momento in azienda, una volta che le moto sono consegnate ai dealer, si perde visibilità sul prodotto).

Attualmente il DCS mostra solo tre diversi step attraverso i quali passa l'ordine del dealer: aperto, in preparazione e in transito (il carico è stato spedito e quindi ora è in carico allo spedizioniere). Si prevede di "aggiungere" due ulteriori stati.

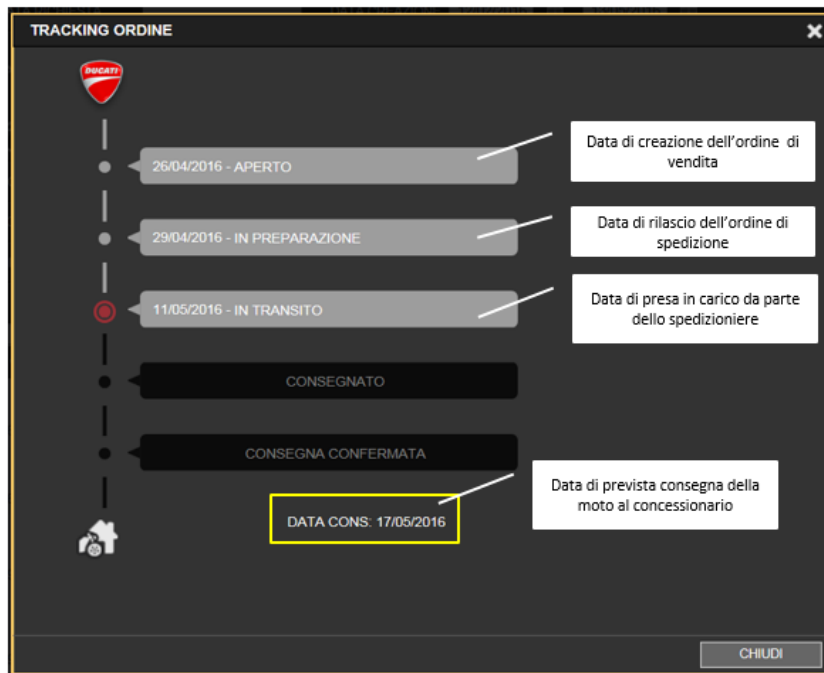


Figura 12 Screenshot rappresentante le fasi visibili ora dai dealer in relazione al visibilità dello stato di avanzamento dell'ordine

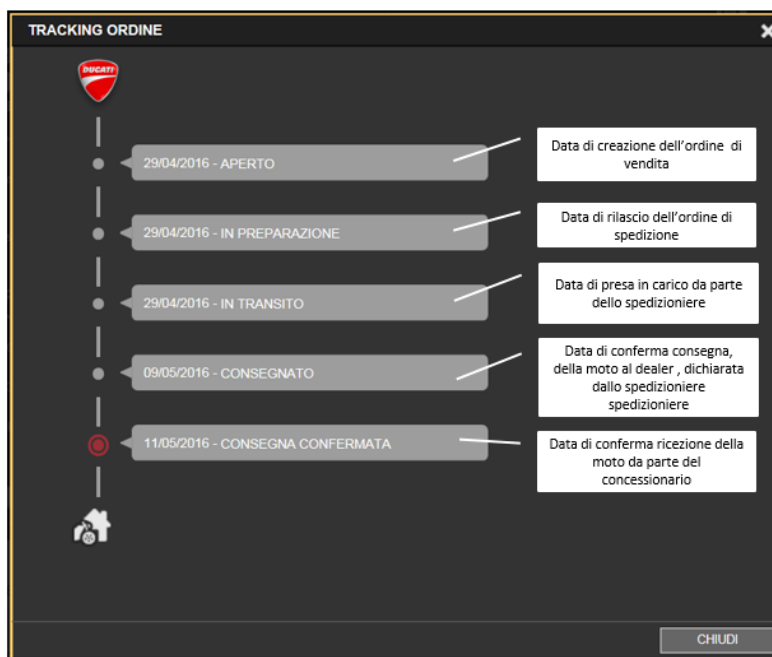


Figura 13 Screenshot DCS con stati aggiuntivi

## **CAPITOLO 5: REPORTING SALES**

### **La reportistica per il monitoraggio delle vendite e la sua applicazione in Ducati Motor Holding**

Parallelamente allo svolgimento delle attività di cui si è parlato sopra (analisi reportistica attuale, tavoli tematici etc.), il progetto, che ha come obiettivo finale l'introduzione della Business Intelligence in Ducati Motor Holding, è proseguito con la fase di estrazione dei dati dalle diverse fonti utilizzate in azienda. A tale operazione hanno fatto seguito: pulizia, trasformazione e caricamento dei dati stessi prima all'interno della staging area e poi del datawarehouse con successiva realizzazione dei primi report a cui io stessa ho lavorato.

Prima di affrontare il cambiamento avvenuto in Ducati a livello di reportistica grazie all'introduzione della BI, propongo una panoramica teorica sull'area vendite all'interno di un'azienda e su quali sono i grafici che non dovrebbero mai essere dimenticati al momento dell'analisi dei relativi dati.

#### **Area sales**

Una delle principali aree aziendali è sicuramente la funzione **sales** che si occupa della gestione delle vendite. In generale nelle aziende si possono individuare differenti funzionalità svolte all'interno di quest'area tra cui attività decisionali, operative e di controllo.

All'interno dell'area sales in Ducati si possono rilevare differenti attività a ciascuna delle quali è associata una specifica sotto area:

- Product marketing: analizza i dati relativi alle vendite provenienti dalle diverse filiali e le confronta con quelli dei principali competitors a diversi livelli di dettaglio quali: cilindrata, modello (Panigale 1299, Monster 1200, Scrambler 800 Icon etc.), famiglie di prodotti (Naked, Sport, Touring etc.) e altri;

- Sales: suddivisa a sua volta in *Area Direzione Vendite* che si occupa della gestione delle filiali e *APAC e importatori*, che gestisce e controlla i rapporti con gli importatori APAC, India e Thailandia. Alcuni dei principali KPI utilizzati in questa area sono:
  - sell-in: quantità di moto spostate con flussi intercompany;
  - sell-out: quantità di moto vendute al cliente finale;
  - stock stimato dalla rete commerciale: dato, in linea generale, dalla differenza tra i due precedenti valori. Nello stock non devono essere considerate quelle bikes che, seppur presenti presso i dealer, appartengono alla categoria courtesy e demo;
  - Customer order: numero di clienti che hanno firmato un contratto per l'acquisto di una moto, ma non l'hanno ancora ritirata. Sono le moto presenti nello stock dei dealer a cui è associato un end-customer, cioè un cliente finale;
  - Dealer order: numero di ordini che i dealer inseriscono a sistema tramite DCS.
- Sales strategies and planning: si occupa della previsione delle vendite, realizzando principalmente analisi statistiche;
- Sales Importers: si occupa della gestione dei rapporti con gli importatori. Alcuni dei KPI utilizzati sono:
  - Avanzamento fatturato, sia per le bikes che per le diverse categorie di business del commercial after sales (accessori, apparel e ricambi);
  - Status degli ordini presenti a sistema, cioè back-order o ordini aperti;
  - Analisi posizione finanziaria degli importers nei confronti di Ducati.



## Overview reporting:

Importance of BI Trends from "Not important at all" (0) to "Very important" (10)



Figura 1 Ranking delle differenti attività che possono essere svolte sui dati nelle diverse realtà aziendali

Senza dati e informazione non è possibile prendere decisioni ragionate e mirate, ma, parallelamente, dati e informazioni sono inutili se non si è in grado di leggerli ed interpretarli e la tabella riportata sopra mostra proprio l'importanza della data discovery e della data visualization in tutte le realtà aziendali.

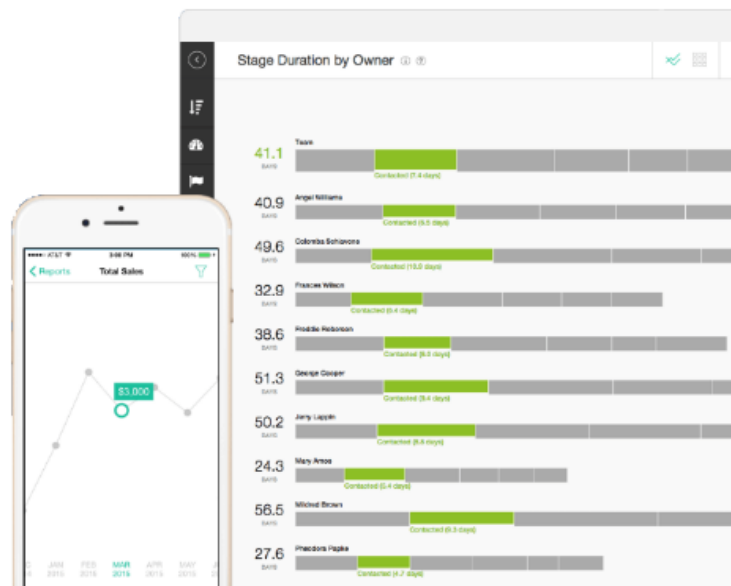


Figura 2 Esempio di dashboard per l'analisi dei dati

L'analisi puntuale del processo di vendita è condizione necessaria per la costruzione del successo di un'azienda in quanto permette ai manager di effettuare analisi rilevanti al fine di aumentare i profitti aziendali, ma anche di individuare eventuali criticità. A queste si aggiunge la possibilità di prevedere i profitti futuri e l'andamento delle performance.

## Come il reporting può aiutare un'impresa a vendere di più

1. Aiuta ad individuare i colli di bottiglia presenti nel processo di vendita dei prodotti;
2. Permette di scendere ad un livello di dettaglio sempre maggiore in termini di analisi dei dati e di effettuare poi il roll-up degli stessi, permette cioè di tornare ad un livello di dettaglio minore;
3. Permette una visualizzazione grafica dei dati rendendoli in questo modo più facilmente analizzabili garantendo un'individuazione pressoché immediata di anomalie;
4. L'analisi dei Big Data permette di comparare i dati attuali con quelli storici relativi, per esempio, alle vendite effettuate nello stesso intervallo temporale in anni diversi (es: studio di stagionalità);

5. Permette di identificare quale ambito del marketing sta raggiungendo i risultati migliori (in seguito a questi risultati si potrebbe ad esempio pensare di effettuare ulteriori investimenti in quel campo oppure di concentrare l'attenzione su altri strumenti o rami del marketing per raggiungere una clientela differente eccetera).

Gli strumenti di reportistica (visual reporting, dashboard, mobile reporting etc.) permettono di concentrare la propria analisi su alcuni specifici aspetti del business come per esempio:

- le ragioni che stanno alla base dei dati e dei risultati aziendali;
- i motivi per cui un team riesce a superare un target di vendita mentre altri team no;
- quando è necessario realizzare un radicale cambiamento del processo delle vendite perché quello esistente è ormai obsoleto e poco efficiente;
- quando è necessario realizzare maggiori investimenti in termini di marketing e sponsorizzazione del marchio;
- come un determinato team sta performando;
- comparare i risultati e identificare opportunità di miglioramento.

La creazione di “cruscotti gestionali” (così vengono chiamati insieme di report, grafici e tabelle che contengono le informazioni riassuntive e i dati di maggiore interesse raccolti nell' intervallo di tempo analizzato), basati su piattaforme di Business Intelligence, rende agevolmente consultabili tutte le informazioni di cui il Top Management ha bisogno e permette di ottenere nuove prospettive di analisi oltre che a indicazioni sul raggiungimento di obiettivi prefissati e trend futuri.

I report, i grafici, le immagini e gli indicatori di performance, se correttamente individuati e combinati, consentono di monitorare in modo interattivo e intuitivo i principali obiettivi e parametri di business.

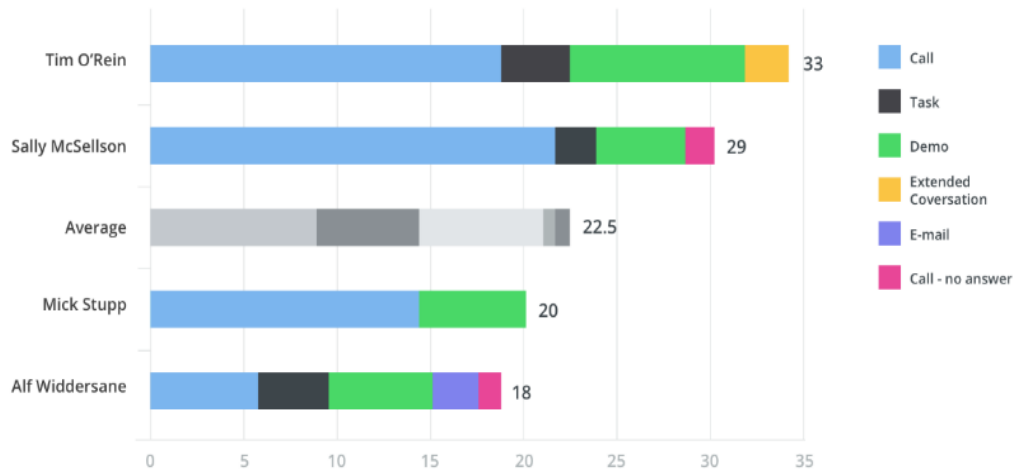


Figura 3 Esempio di grafico per l'analisi dei dati

### How many deals were won?

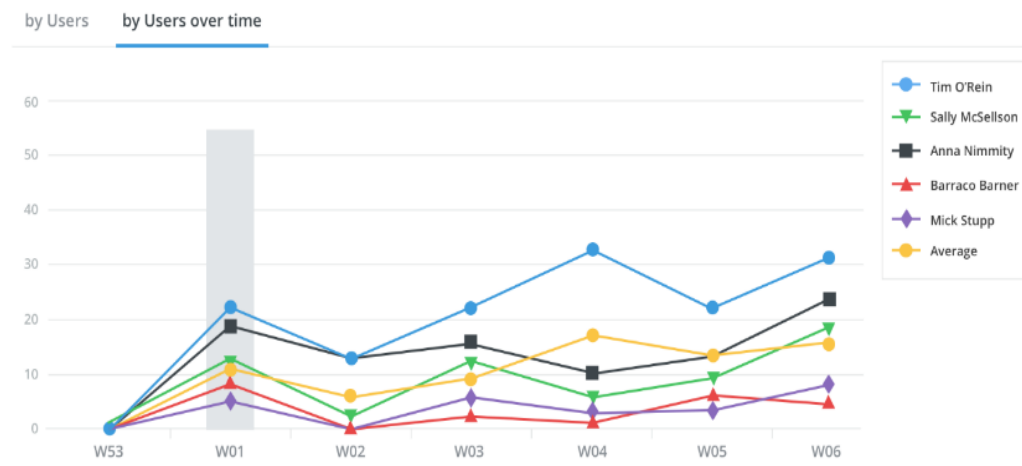


Figura 4 Secondo esempio di grafico

## Sei consigli per realizzare migliori dashboard per il monitoraggio delle vendite

### 1. Conoscere i propri dati e la loro collocazione.

Conoscere i propri dati è particolarmente importante per considerare opportunità differenti. Analizzare i dati e saperli osservare da prospettive diverse favorisce l'elaborazione di molteplici strategie e piani d'azioni. Importante è anche quello che viene chiamato "data lineage" definito come il ciclo di vita dei dati che include la loro origine e come si sono modificati nel tempo. Il data lineage descrive cosa è accaduto ai dati dopo il verificarsi di determinati eventi o processi, aiuta ad avere visibilità analitica, a rintracciare e a rimediare agli errori compiuti. Può quindi essere visto come una sorta di film nel quale vengono riportati tutti i cambiamenti avvenuti nei dati in relazione a processi, decisioni, eventi eccetera;

### 2. Ottimizzare e organizzare i propri dati in maniera tale da avere una sola fonte di verità.

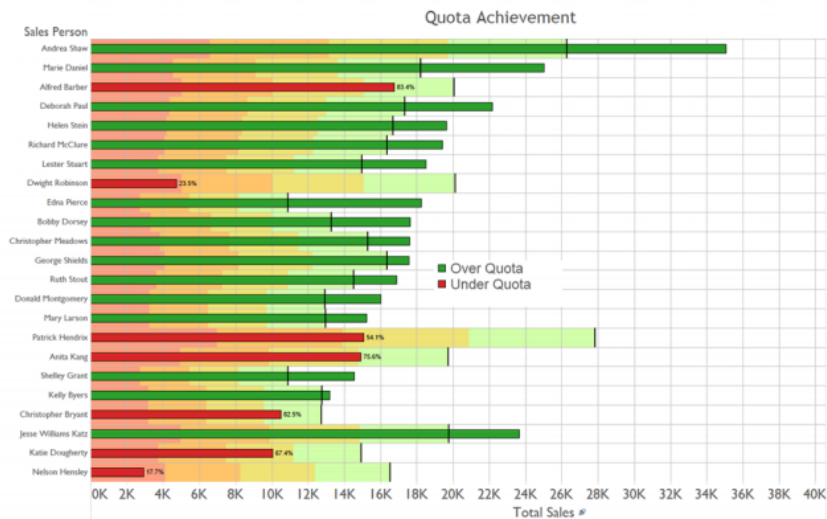
I dati delle vendite da soli, senza informazioni geografiche o senza i dati del marketing hanno poco valore e sono utili per supportare solo un range limitato di decisioni. Utilizzare molteplici fonti dati per un'unica dashboard è fondamentale per estrarre il massimo potere informativo dai dati aziendali. La combinazione dei dati, conosciuta anche come data mash-up, è la capacità di unire e connettere tra loro dati integrati provenienti da sorgenti differenti. Questo permette di avere una visione unificata dei propri prospetti, delle vendite, dei propri utenti e obiettivi. *Misurazione delle performance*: capire l'andamento delle vendite è un elemento chiave per comprendere quali tipi di aggiustamenti è necessario applicare al proprio business. Le classiche analisi comparative, come per esempio l'andamento delle vendite negli anni, sono semplici da realizzare; ma è necessario guardare anche alla storia del proprio business, sia in maniera cumulata sia facendo riferimento alle singole porzioni del business. Unire i dati delle

vendite a quelli del prodotto permette di capire quali sono i trend e le nuove opportunità presenti nel mercato;

3. Scegliere metriche significative: le metriche utilizzate nelle dashboard devono essere rilevanti per individuare quanto si è prossimi al raggiungimento degli obiettivi prefissati. Al fine di rimanere focalizzati sugli obiettivi, bisogna prestare molta attenzione nel selezionare le metriche corrette. Per questo bisogna fare riferimento ai KPI aziendali e a come sono misurati. Bisogna poi allineare le azioni ad ogni indicatore e usare la dashboard risultante per guidare le azioni dell'impresa (Data Driven);
4. Realizzare visualizzazioni interattive: le dashboard sono importanti perché permettono a tutti di accedere e visualizzare gli stessi dati. Dashboard interattive permettono di realizzare differenti analisi, applicare filtri per focalizzarsi solo su determinati valori, realizzare calcoli veloci, effettuare drill-down dei dati, tutto questo con delle semplici selezioni effettuate direttamente sulla dashboard stessa. Grazie a questa interattività le montagne di dati che si hanno a disposizione vengono trasformati in piani realizzabili e in veloci processi di decision making;
5. Utilizzare Current and Live Data: i dati storici possono essere usati per osservare trend e andamenti che hanno caratterizzato la storia di un'azienda negli anni, ma la corretta linea del tempo che deve essere usata nelle dashboard è quella che permette di avere dati aggiornati e di effettuare analisi real time;
6. Permettere agli utenti di effettuare analisi in maniera indipendente (Self-Reliance): il tutto non deve dipendere dall'IT ed è importante fornire alle varie aree aziendali e ai vari uffici la possibilità di realizzare dashboard per tener traccia dei propri prospetti, obiettivi e KPI.

## Cinque grafici che ogni responsabile delle vendite dovrebbe avere a disposizione

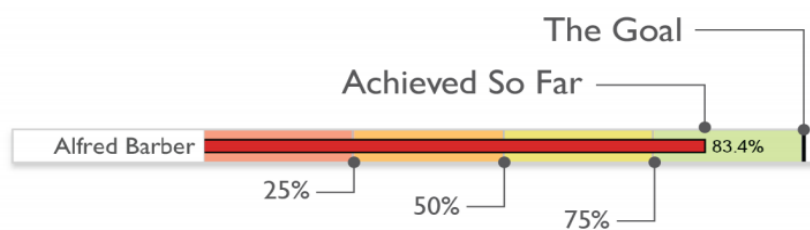
### The bullet



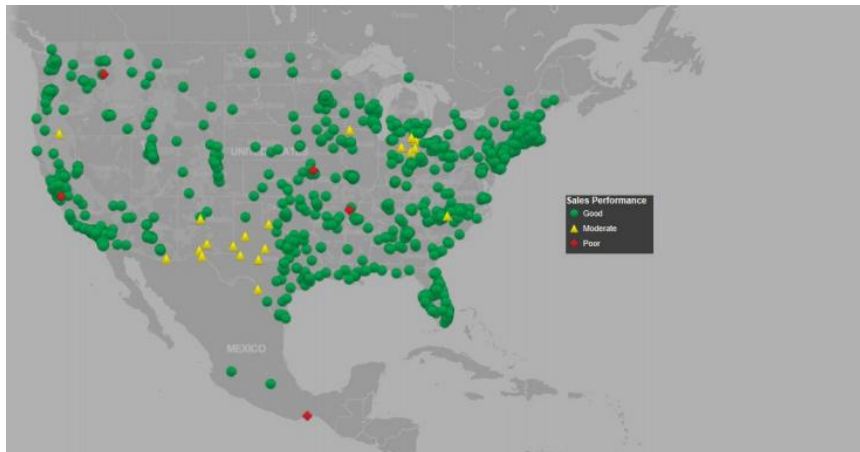
Tutti i bullet graphs riguardano la misurazione dei progressi verso un target prestabilito. I più analizzati sono in merito a:

- numero di offerte;
- quota delle vendite;
- performance individuali.

Il colore della barra mostra immediatamente se l'obiettivo sta per essere raggiunto o se si è ancora piuttosto lontani, mentre la lunghezza indica la posizione del work in progress rispetto all'obiettivo. I diversi colori utilizzati permettono di individuare la posizione attuale rispetto a benchmark di riferimento.



## KPI map



Attraverso la mappa dei KPI è possibile vedere immediatamente dove dovrebbero o non dovrebbero essere spese risorse e quindi effettuare ulteriori investimenti.

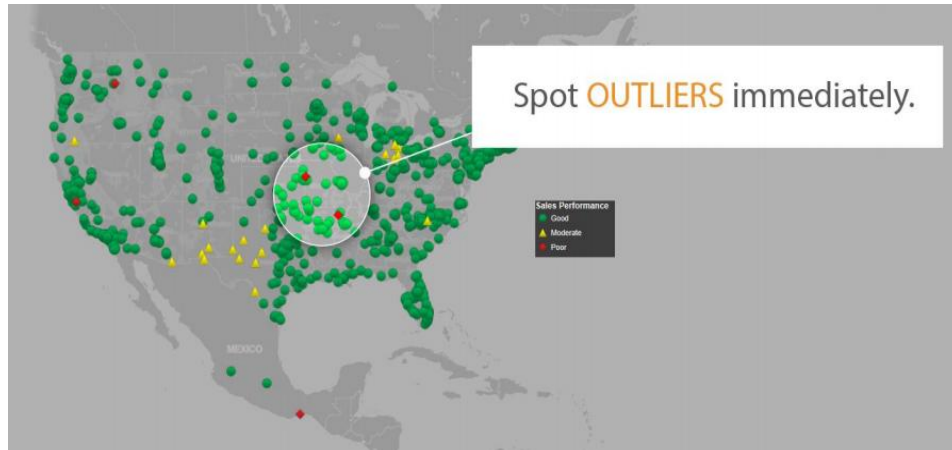
Attraverso questo tipo di grafico è possibile ricavare informazioni riguardo:

- anomalie territoriali;
- difficoltà locali;
- performance per area geografica;
- performance per regione.

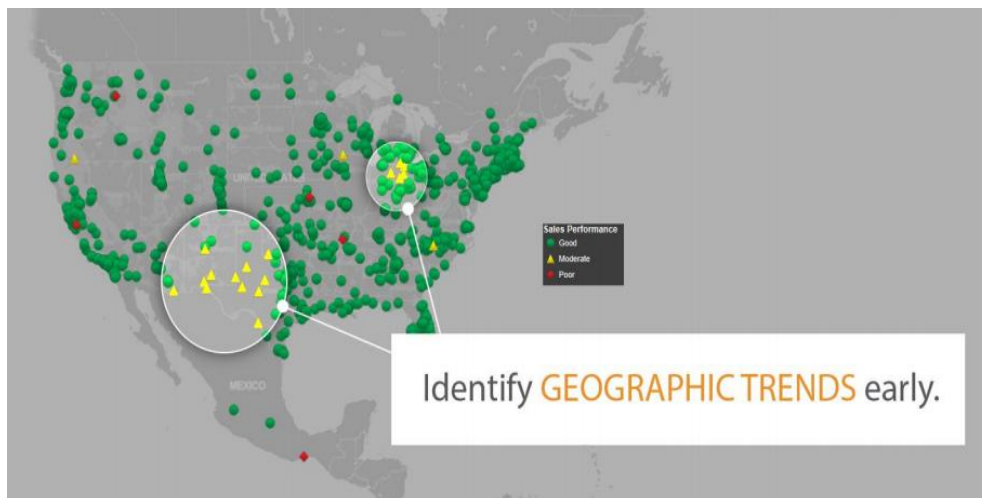


Attraverso questo tipo di rappresentazione sono immediatamente individuabili:

- Gli outliers<sup>9</sup>;



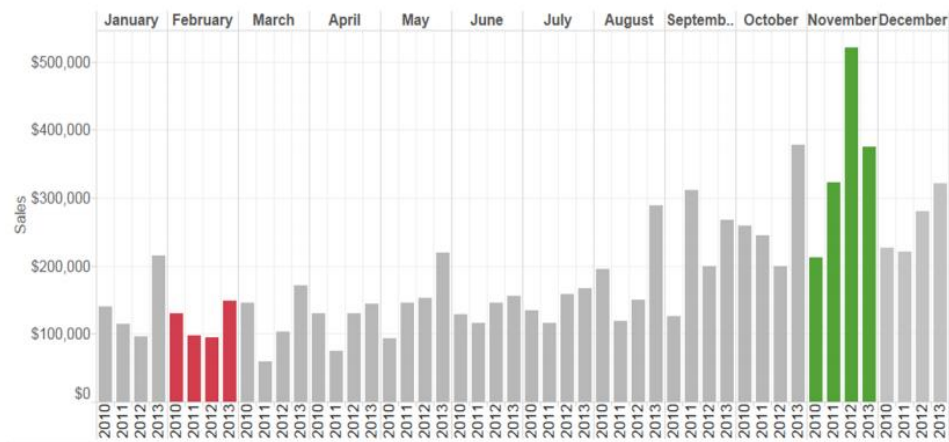
- I trend geografici;



---

<sup>9</sup> Outlier: è un termine utilizzato in statistica per definire, in un insieme di osservazioni, un valore anomalo e aberrante; un valore quindi chiaramente distante dalle altre osservazioni disponibili. Solitamente questi dati vengono esclusi dalle analisi statistiche perché possono portare a risultati fuorvianti.

## I trend stagionali

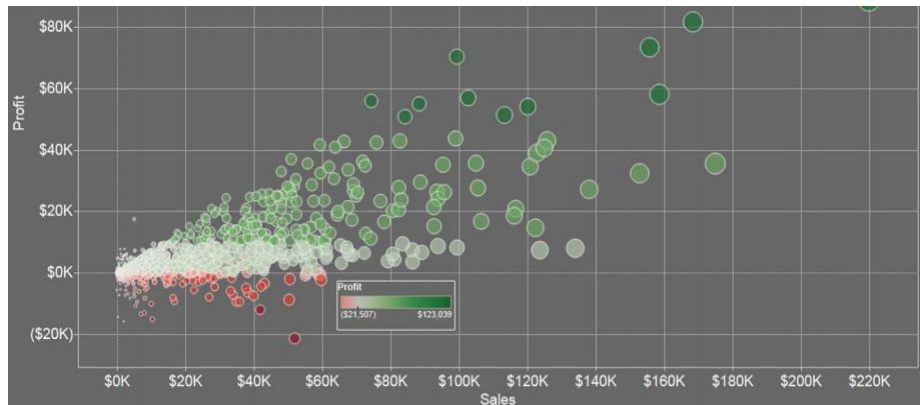


Questo report mi permette di individuare i trend dovuti alla stagionalità e offre, di conseguenza, la possibilità di effettuare previsioni e piani.

Avendo i risultati mensili di diversi anni disposti a calendario uno vicino all'altro, è possibile vedere facilmente i trend ciclici ed individuare i picchi minimi e massimi annuali delle vendite.



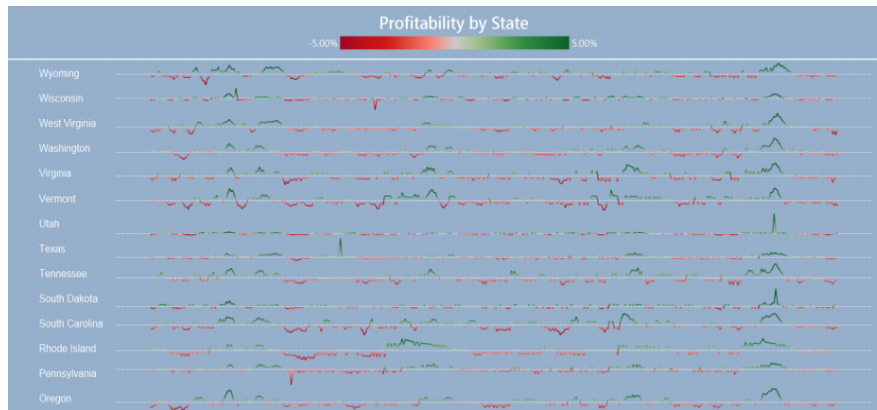
## Sales vs Profit scatter plot



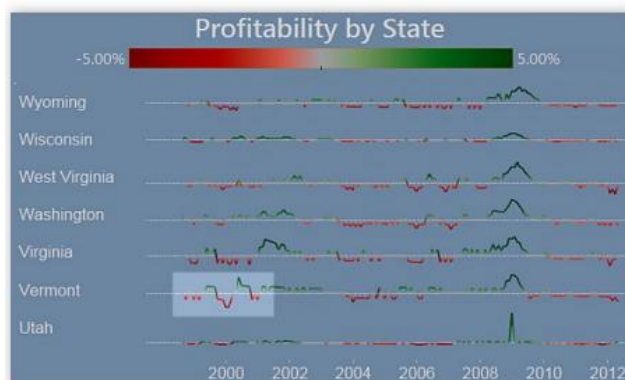
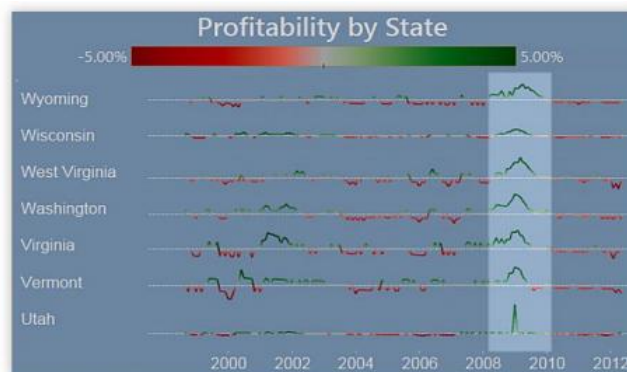
Registrazione molte vendite non significa per forza avere elevati profitti. Questo grafico permette non solo di valutare la frequenza di un determinato evento (nel caso delle vendite va a valutare il loro volume in termini di quantità) ma anche la sua intensità (riprendendo l'esempio precedente si valuta il valore economico riscosso dalle vendite effettuate).



## The sparklines chart



L'importante non è conoscere ogni dettaglio, ma essere in grado di individuare rapidamente picchi, positivi o negativi che siano, ed essere in grado di intervenire rapidamente. Il grafico sparkline aiuta a individuare tali anomalie e a capire che qualcosa di strano si sta verificando.

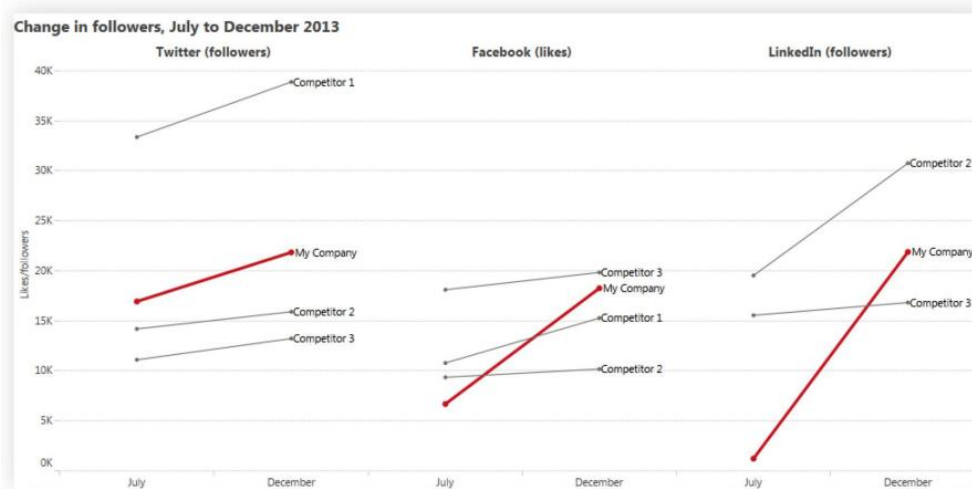


## Quattro grafici per l'analisi dei dati raccolti dai social media

Oggi la tecnologia regna sovrana nella vita di tutti i giorni e strumenti quali internet e i social media (Facebook, Twitter, Instagram) sono ormai diventati fondamentali per un'impresa per pubblicizzare il proprio brand, lanciare anteprime sui nuovi prodotti che stanno per essere immessi sul mercato, raccogliere i pareri dei pionieri, cioè dei clienti che per primi hanno acquistato e utilizzato il nuovo prodotto proposto sul mercato e molto altro.

Il mondo del web è quindi fortemente legato alle vendite di un'impresa e per questo motivo riporto un breve excursus sui grafici più importanti relativi alla raccolta e all'analisi dei dati provenienti dai social media.

### Slope chart for growth of followers/reach

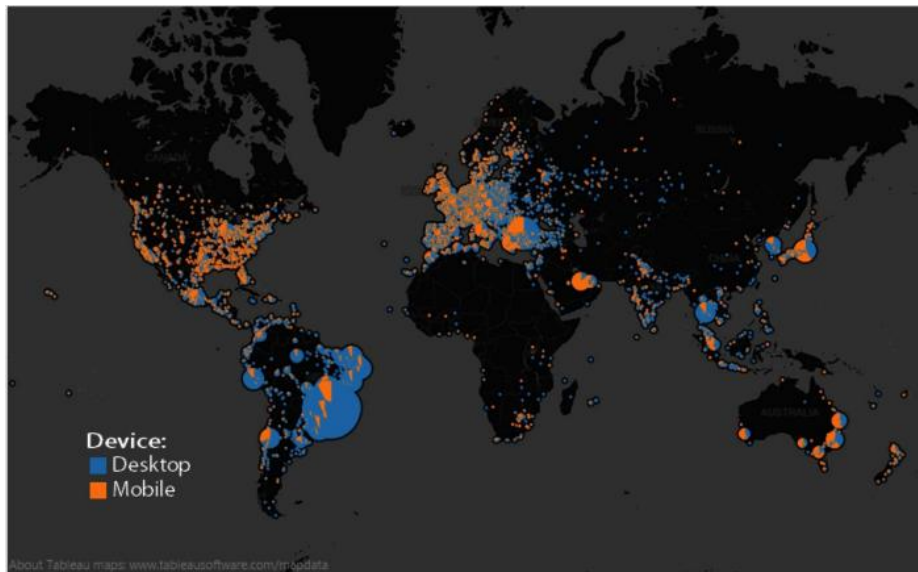


Questo grafico mostra i cambiamenti intercorsi nell'intervallo di tempo preso in analisi, permettendo comparazioni molto semplici fra il tasso di partenza e quello di arrivo (i diversi colori utilizzati permettono di mettere in evidenza una determinata azienda o area aziendale nel caso in cui il focus di analisi sia interno).

I dati che possono essere rappresentati sono:

- tasso di crescita dei follower;
- citazioni e menzioni;
- engagement totale raggiunto.

### **Mappa dei click sui link regionali / globali**

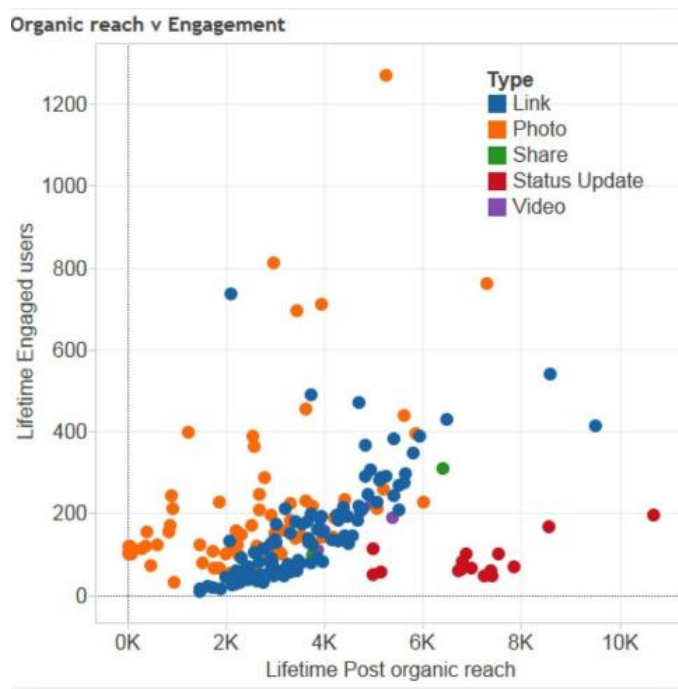


Questa mappa permette di visualizzare il numero di click sui link relativi all'azienda posta sotto analisi e permette di vedere dove si concentra la maggior parte dell'audience dell'azienda e quali differenze regionali concorrono.

I dati visualizzabili sono:

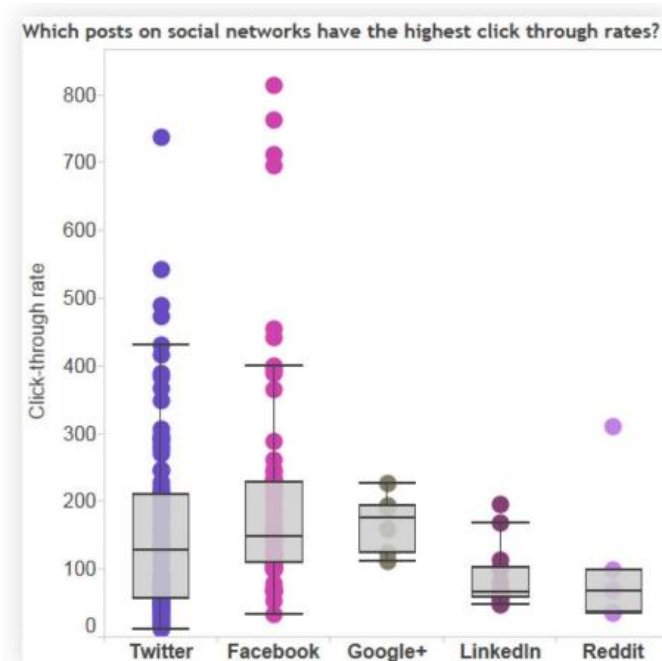
- click sui link all'interno di social network;
- re-tweet e post di Facebook con annessa posizione geografica (geo-tagged posts);
- localizzazione dei followers.

## Grafico di dispersione (scatter plot)



Questo grafico permette di comparare due o più misure per trovare relazioni e correlazioni. Rende più facile la possibilità di individuare outliers e clusters di dati.

## Click-through rate in un grafico boxplot



Questo grafico mostra il tasso di visualizzazioni e condivisioni di link avvenuti su diversi canali durante un orizzonte temporale di 6 mesi. I rettangoli grigi evidenziano la media e il quartile di ogni social network. Questa visualizzazione permette di confrontare distribuzioni di dati attraverso diverse categorie.

I dati visualizzabili sono:

- tasso di visualizzazione e condivisioni;
- engagement raggiunto per categoria di post su Facebook;
- follower of followers, cioè il livello di diffusione raggiunto dal post attraverso “passa parola” confrontandolo con il valore relativo ai propri competitors.

## **Reportistica App WDW**

A tal proposito riporto alcune immagini relative ad una dashboard navigabile che io stessa ho realizzato durante il mio tirocinio in Ducati. Tale dashboard riporta dati raccolti dall'applicazione creata appositamente per l'evento WDW (Word Ducati Week), evento biennale dedicato a tutti i ducati e ai fan del brand della durata di 3 giorni che ha avuto luogo a luglio 2016 e che ha visto il susseguirsi di concerti, spettacoli, eventi e molto altro.

Attraverso il collegamento dell'applicazione direttamente al sito Google Analytics e alla successiva connessione tra quest'ultimo e il software utilizzato per la reportistica, nel nostro caso Power BI, è stato possibile realizzare i seguenti report:

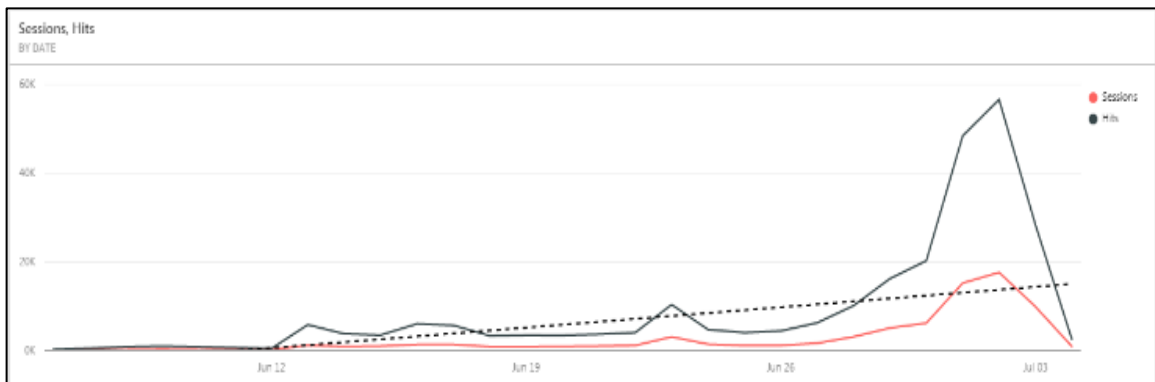
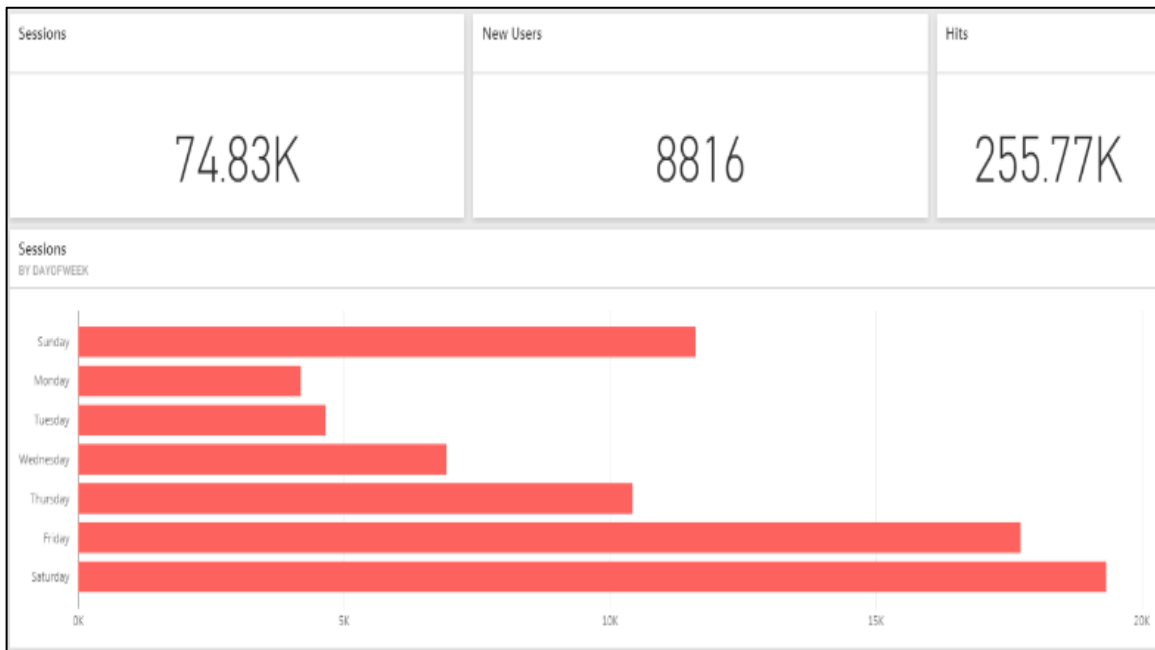


# WDW2016

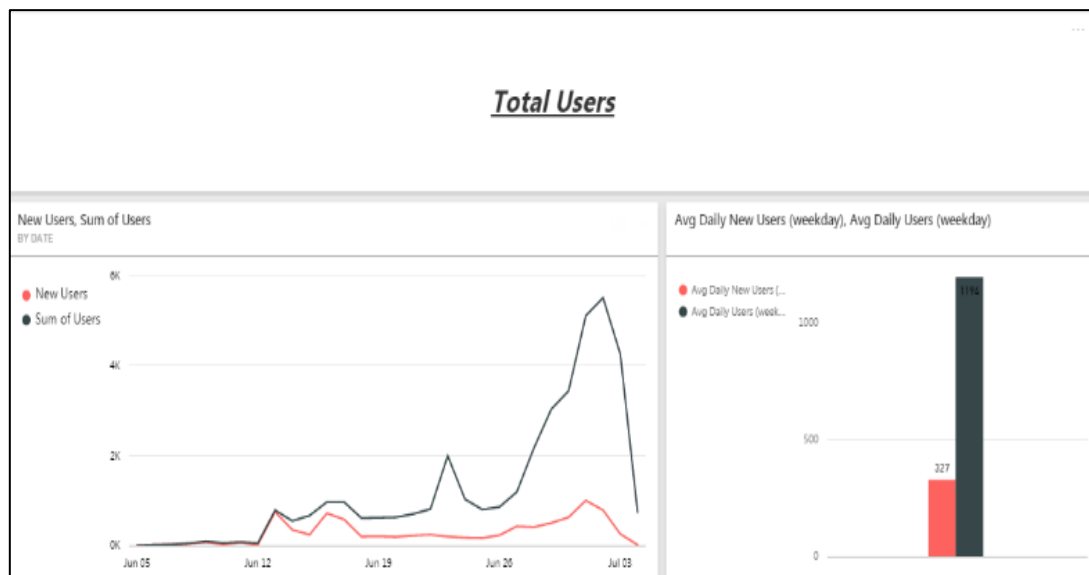
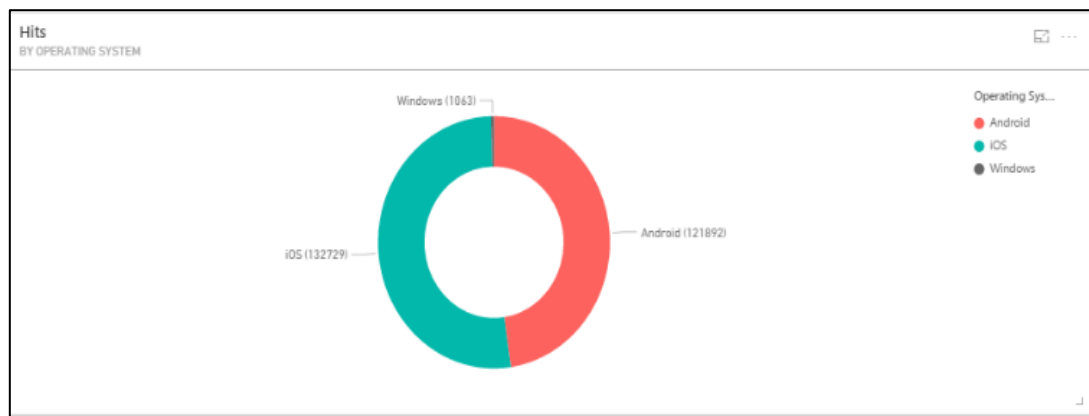
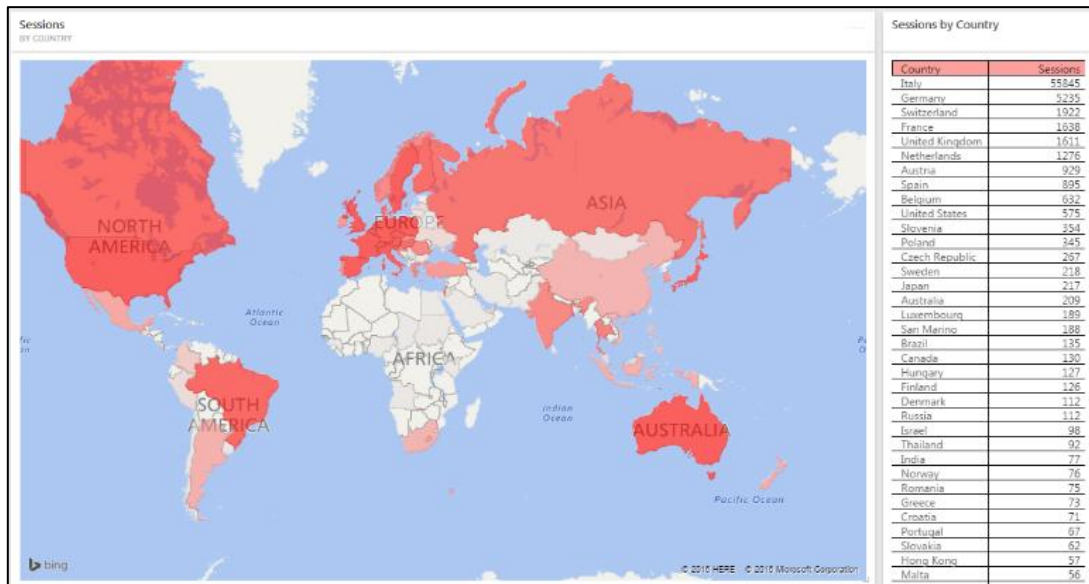
## WORLDUCATIWEEK

Dati App WDW [1-3 Luglio 2016]

Site Traffic



System Usage



Le immagini riportate contengono dati raccolti fino a qualche giorno prima l'inizio dell'evento e mostrano dati relativi al numero di visualizzazioni, al numero di persone che hanno effettuato il download dell'applicazione, i paesi in cui le visualizzazioni sono state maggiori, il tipo di sistema operativo degli apparecchi digitali utilizzati per visualizzare l'app (di questa applicazione sono infatti state realizzate due diverse versioni, una per dispositivi Android e l'altra per I-Phone) e molti altri.

I dati contenuti possono essere aggiornati con un semplice click sul tasto "refresh" direttamente da Power BI.

## **Reportistica in Ducati prima e dopo la BI**

L'obiettivo di questo paragrafo conclusivo è quello di mostrare la differenza tra la reportistica realizzata in Ducati prima e dopo l'adozione della Business Intelligence.

Il confronto, basato sulla realizzazione di questi report, si concentrerà specialmente su:

- metodologia utilizzata per la realizzazione della reportistica prima e dopo l'adozione della BI (in termini di fonti dati, software utilizzati, elaborazione dei dati, operazioni di data quality e data assurance);
- tipologia di report realizzati, relative possibilità di analisi e software utilizzati.

Si riporteranno tre esempi pratici di report che hanno come elementi di analisi: la stabilità degli ordini dell'azienda, l'aging dello stock presente presso i dealer e gli ordini che vengono confermati dai vettori che consegnano la merce (si precisa che i dati presenti nei grafici non sono reali e sono stati appositamente modificati per motivi di privacy aziendale).

### **1. Order stability:**

- **Obiettivo:** Analizzare la stabilità degli ordini in termini di cancellazione e modifiche;
- **Scope:** Filiali europee;

- **Frequenza di misurazione KPI:** mensile;
- **Orizzonte temporale:** 6 mesi nel passato e 2 nel futuro;
- **Fonti dati:** SAP.

## 2. Dealer Inventory Aging:

- **Obiettivo:** Misurazione della percentuale di moto in giacenza presso il dealer rispetto al totale dell'inventario facendo una distinzione tra quelle presenti da meno di 30 giorni, da 30 a 60, da 60 a 90 e da più di 90 giorni;
- **Scope:** Filiali europee, filiali overseas (DNA, Giappone);
- **Frequenza di misurazione KPI:** ogni 15 giorni;
- **Orizzonte temporale:** su tutto l'orizzonte temporale possibile;
- **Fonte Base Dati:** SAP.

## 3. Pod Tracking:

- **Obiettivo:** Misurare il livello di uso, da parte dei provider logistici, della nuova app POD (go live in Italia: 07/03/2016; go live nel resto d'Europa: 01/04/2016);
- **Scope:** Filiali europee, filiali overseas (DNA, Giappone);
- **Frequenza di misurazione KPI:** Settimanale;
- **Orizzonte Temporale:** ultime 8 settimane;
- **Fonte Dati:** SAP.

## **Data quality assurance**

Per ognuno dei 3 report descritti precedentemente devono essere effettuate elaborazioni manuali sui dati ai fini di garantirne la significatività. Descrivendo i passaggi necessari è possibile comprendere al meglio l'efficienza ottenuta automatizzando il processo grazie alla BI:

### Order stability

Per la realizzazione del report è necessario estrarre i dati attraverso l'esecuzione di due query SAP differenti, una contenente una lista degli ordini ricevuti da Ducati nel periodo di tempo considerato e l'altra riportante tutte le modifiche subite nel tempo dai vari ordini (esempio: ordini aperti e

poi cancellati, oppure modifiche relative alla data di consegna). Si ottengono quindi due fogli Excel nei quali dovranno essere eliminati dei dati e aggiunte delle colonne:

- ✓ aggiungere le colonne “Stocking Orders” e “Standard Orders”: denominazione data agli ordini in base alle utenze che li hanno creati;
- ✓ aggiungere la colonna “Ordini cancellati” riportante i valori dei soli ordini con codice dello “stato dell’ordine” uguale a ZD e Z0 (che nello specifico stanno ad indicare ordini annullati e cancellati);
- ✓ aggiungere la colonna “Ordini Modificati” riportante solo i valori degli ordini che hanno subito una modifica della data di consegna prevista al dealer;
- ✓ eliminare dalla lista degli ordini tutti quelli che hanno più posizioni e tutti gli ordini aventi quantità maggiore di uno;
- ✓ eliminare tutti questi ordini segnalati come modificati per i quali il cambio di data è avvenuto nello stesso mese (per esempio un ordine previsto per il 4 aprile viene posticipato al 24, in quanto è una modifica non considerata rilevante dal punto di vista delle conseguenze sulla produzione) e tutti gli ordini per i quali il delta, misurato in giorni, tra la nuova e la vecchia data è superiore ai 300 giorni in quanto si tratta di errori successivamente corretti.

Il report finale quindi riporta tutti gli ordini con data di prevista consegna presente nell’orizzonte temporale di analisi e che hanno in seguito subito almeno una modifica di data rilevante o una cancellazione.

### **Dealer Inventory Aging**

Per la realizzazione del report è necessario estrarre i dati attraverso l’esecuzione di una query SAP andando precedentemente ad inserire i codici identificativi dei dealer Ducati ancora attivi. Questa lista mi è giunta in

formato Excel. Si ottiene quindi la lista degli ordini inviati a questi concessionari. Le modifiche da effettuare al file così ottenuto sono:

- ✓ cancellare tutte le righe per le quali non esiste una fattura eccetto per il dealer di Ducati Roma;
- ✓ aggiungere la colonna “Descrizione motivo ordine”: indica il tipo di ordine (End-Customer, Stock, Demo) attraverso la seguente formula:

```

=+SE(O(AA2="zef";AA2="zec");"END
CUSTOMER";SE(T2 <> "";"DEMO";"STOCK"))
ifthenelse (DEALERINVENTORY02."Motivo
ordine"='ZEC' OR
DEALERINVENTORY02."Motivo
ordine"='ZEF','END CUSTOMER,
ifthenelse ( DEALERINVENTORY02."Demo
Order" = ", 'STOCK','DEMO')));

```

- ✓ aggiungere la colonna “Lead Time Trasporto”: indicante il tempo che intercorre dall’uscita merci al momento in cui arriva presso il dealer. Tali tempi sono contenuti all’interno del file “LT trasporto” proveniente dalla logistica. La compilazione della tabella può essere fatta con un cercavert ( o vlookup) sulla colonna del paese di destinazione delle moto:

```

= +VLOOKUP(Inventory! D2;'LT trasp'! A: H; 8; 0);

```

- ✓ aggiungere la colonna “Consegnata presso il dealer” ottenuta dalla data di spedizione più il lead time di trasporto;
- ✓ aggiungere la colonna “Aging” che va a calcolare l’anzianità della moto presente nell’inventario del dealer calcolato facendo la differenza tra la data odierna e la data di consegna presso il dealer;

- ✓ inserire la colonna “Cluster”: riporta una clusterizzazione dell’aging suddividendo lo stock in: <30giorni, 30-60 giorni, 60-90 giorni, >90giorni:

= +IF(AE2 <= 30;" <= 30gg";IF(AND(AE2 > 30;AE2 = 60);"30 – 60gg";IF(AND(AE2 > 60;AE2 <= 90);"60 – 90gg";" > 90gg"))).

### Pod Tracking

Anche in questo caso i dati necessari vengono estratti direttamente da SAP. Si ottiene un elenco di ordini che posso aver associati diversi stati indicanti: ordine aperto, ordine chiuso, ordine confermato dal vettore e ordine confermato dal dealer. Le modifiche da apportare al file sono:

- ✓ eliminare tutte le consegne che non hanno un’uscita merci;
- ✓ eliminare all’interno di una stessa consegna i “codici stato” che si ripetono più volte; a tal proposito si consiglia di procedere nel modo seguente:

- aggiungere la colonna “Chiave” come prima colonna in cui inserire la formula:

= B2&J2;

- ordinare per chiave (colonna A);
- inserire nella prima colonna disponibile (colonna R) la formula seguente:

= SE(E(CONTA.SE(A:A,A2) > 1;A2 = A3); Elimina;"") ;

- eliminare le righe in cui nella colonna R, appena creata, è valorizzato con “Elimina”;
- cancellare la colonna ”Chiave” (colonna A);
- ordinare per “Consegna” (colonna A) e “Contatore” (colonna B);

- ✓ aggiungere la colonna “POD Tracking:” riporta l’informazione di quale POD (proof of delivery) è stata fatta per una specifica consegna. In particolar modo l’informazione può essere:

- “**POD Vettore**” nel caso in cui esista la conferma da parte del vettore logistico;
- “**POD Dealer**” se è stata effettuata la conferma da parte del dealer;
- “**In transito**” se non c’è stata alcuna conferma (POD) né da parte del vettore logistico né da parte del dealer;

= +IF(I2 <> 1; ""; IF(OR(COUNTIF(A:A; A2) = 1; AND(COUNTIF(A:A; A2) = 3; I3 = 3)); "In Transito"; IF(OR(AND(COUNTIF(A:A; A2) = 2; I3 = 900); COUNTIF(A:A; A2) = 3; AND(COUNTIF(A:A; A2) = 4; I5 = 900); COUNTIF(A:A; A2) = 5); "POD Vettore"; "POD Dealer"))) ;

- ✓ aggiungere la colonna ”Week”: numero della settimana della consegna prevista:

= WEEKNUM(H2)

- ✓ aggiungere la colonna “Country”: indicazione relativa al gruppo clienti a cui è destinata la consegna. Tali informazioni sono contenute nel file “Itinerari” fornito dalla logistica:

= VLOOKUP(C:C; Itinerari! A:H; 3; 0) .

## Prima della BI

Come detto, prima della BI il lavoro di analisi dei dati e realizzazione dei report era basato su estrazioni di dati direttamente su SAP ed eventuale loro unione con altri file Excel (come per esempio nei report relativi al sell-in



bikes in cui i dati contenuti in SAP vengono uniti con dati contenuti in file Excel provenienti direttamente dalle filiali).

I file Excel, come tutti sanno, si presentano in forma tabellare e offre possibilità di rielaborazione da parte dall'utente.

Una volta raccolto il materiale necessario per la realizzazione dei report, i vari file venivano modificati in maniera manuale attraverso l'aggiunta di colonne calcolate (le formule più usate comprendono gli operatori classici di Excel quali "if" o "vlookup" per ottenere un unico foglio contenente tutti i dati utili. Su questa versione definitiva veniva poi costruito un grafico pivot.

Equipment	Material	Client	Sold to Dealer Name	Descr Gr CI	Descr Gr CI 2	Dealer Invoice	Dealer Invoice	Area	Mani	Area Manager	Name	AM Est.	Man	Dat	Elke	Prod.	Super Model	Segment	Model	Power	Model Y	Family	Standard	ation	Color	
2	D2000267	E-49	BOB CYCLE MOTORCYCLE CO, LTD	USA	CNA	800003	AREA CNA S30	S30		800003	AREA CNA S30	530				25.02.2015	SCR-III	NAVED SMALL	Icon	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Red	I	
3	D2000267	E-49	CYCLE DOOL INC	USA	CNA	800009	AREA CNA S25	S25		800009	AREA CNA S25	525	02.06.2015	12.04.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	
4	D2000268	E-49	MOTO SPORTS	USA	CNA	800001	AREA CNA S26	S26		800001	AREA CNA S26	526	16.08.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
5	D2000268	E-49	POWERSPORTS OF KENT WA	USA	CNA	800001	AREA CNA S26	S26		800001	AREA CNA S26	526	12.06.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
6	D2000276	E-49	HEDSON VALLEY MOTORCYCLE USA	USA	CNA	800003	AREA CNA S21	S21		800003	AREA CNA S21	521	10.07.2015	SCR-III	NAVED SMALL	Icon	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Yellow	I	I	I	I	
7	D2000268	E-49	BEARH STATE MOTORCYCLES USA	USA	CNA	800009	AREA CNA S25	S25		800009	AREA CNA S25	525	17.05.2016	16.07.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I
8	D2000268	E-49	MOTIES MOTORSPORTS, LLC	USA	CNA	800002	AREA CNA S27	S27		800002	AREA CNA S27	527	22.07.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
9	D2000268	E-49	TOURING SPORT	USA	CNA	800007	AREA CNA S22	S22		800007	AREA CNA S22	522	05.08.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
10	D2000268	E-49	MOTOSPORTS (ASB) S.A. (A-C) Mexico	CNA	CNA	800001	AREA CNA S21	S21		800001	AREA CNA S21	521	07.01.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
11	D2000276	E-49	AGS MOTORCYCLE PARTS, INC.	USA	CNA	800009	AREA CNA S23	S23		800009	AREA CNA S23	523	17.05.2016	14.08.2015	SCR-III	NAVED SMALL	Icon	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Yellow	I	I	I	
12	D2000268	E-49	COMMONWEALTH MOTORCYCL USA	USA	CNA	800002	AREA CNA S27	S27		800002	AREA CNA S27	527	21.03.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
13	D2000268	E-49	FRONTLINE EUROSPORTS LLC	USA	CNA	800002	AREA CNA S27	S27		800002	AREA CNA S27	527	27.01.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
14	D2000268	E-49	VALLEY MOTORSPORTS, INC.	USA	CNA	800003	AREA CNA S21	S21		800003	AREA CNA S21	521	12.12.2015	SCR-III	FULL THRO'	NAVED SMALL	Full Throttle	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Deep Blue	I	I	I	
15	D2000268	E-49	MOTO ALSTIN LP	USA	CNA	800009	AREA CNA S24	S24		800009	AREA CNA S24	524	22.12.2015	16.11.2015	SCR-III	FULL THRO'	NAVED SMALL	Full Throttle	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Deep Blue	I	I	
16	D2000268	E-49	Bob Weaver Motorsports and Mar	USA	CNA	800003	AREA CNA S21	S21		800003	AREA CNA S21	521	17.10.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
17	D2000268	E-49	SPC ENTERPRISES, INC	USA	CNA	800009	AREA CNA S24	S24		800009	AREA CNA S24	524	26.10.2015	SCR-III	FULL THRO'	NAVED SMALL	Full Throttle	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Deep Blue	I	I	I	
18	D2000268	E-49	DUKATI ATTILLI, LLC	USA	CNA	800009	AREA CNA S25	S25		800009	AREA CNA S25	525	27.01.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
19	D2000268	E-49	DUKATI MOTOR BEACH, INC.	USA	CNA	800009	AREA CNA S23	S23		800009	AREA CNA S23	523	26.10.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
20	D2000268	E-49	VANDOVEN MOTORRAD LTD	Canada	CNA	800003	AREA CNA S30	S30		800003	AREA CNA S30	530	03.12.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
21	D2000276	E-49	SO CAL TRIUMPH, INC	USA	CNA	800009	AREA CNA S23	S23		800009	AREA CNA S23	523	10.12.2015	SCR-III	NAVED SMALL	Icon	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Yellow	I	I	I		
22	D2000268	E-49	MOTO EUROPA, LLC	USA	CNA	800009	AREA CNA S24	S24		800009	AREA CNA S24	524	21.12.2015	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
23	D2000268	E-49	VIP ENTERPRISES OF OHLAND USA	USA	CNA	800007	AREA CNA S22	S22		800007	AREA CNA S22	522	04.10.2015	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
24	D2000268	E-49	CALIFORNIA SPEED SPORTS, INC	USA	CNA	800003	AREA CNA S23	S23		800003	AREA CNA S23	523	18.01.2016	14.01.2016	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I
25	D2000276	E-49	MOTALA, INC.	USA	CNA	800009	AREA CNA S23	S23		800009	AREA CNA S23	523	04.01.2016	SCR-III	NAVED SMALL	Icon	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Yellow	I	I	I		
26	D2000268	E-49	THE MOTORCYCLE SHOP, INC.	USA	CNA	800009	AREA CNA S25	S25		800009	AREA CNA S25	525	22.02.2016	08.01.2016	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I
27	D2000268	E-49	SO CAL TRIUMPH, INC.	USA	CNA	800009	AREA CNA S23	S23		800009	AREA CNA S23	523	16.01.2016	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
28	D2000268	E-49	THE MOTORCYCLE SHOP, INC.	USA	CNA	800003	AREA CNA S25	S25		800003	AREA CNA S25	525	16.01.2016	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
29	D2000268	E-49	RETAIL MOTORCYCLE VENTURE USA	USA	CNA	800009	AREA CNA S24	S24		800009	AREA CNA S24	524	09.01.2016	SCR-III	FULL THRO'	NAVED SMALL	Full Throttle	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Deep Blue	I	I	I	
30	D2000268	E-49	CYCLER, INC.	USA	CNA	800002	AREA CNA S27	S27		800002	AREA CNA S27	527	12.01.2016	SCR-III	FULL THRO'	NAVED SMALL	Full Throttle	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Deep Blue	I	I	I	
31	D2000268	E-49	WISSELL MOTORSPORTS	USA	CNA	800002	AREA CNA S27	S27		800002	AREA CNA S27	527	10.01.2016	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Orange	Su	I	I	
32	D2000268	E-49	Bob Weaver Motorsports and Mar	USA	CNA	800003	AREA CNA S21	S21		800003	AREA CNA S21	521	22.01.2016	SCR-III	CLASSIC	NAVED SMALL	Classic	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	
33	D2000268	E-49	MG Maxoral LLC	USA	CNA	800003	AREA CNA S25	S25		800003	AREA CNA S25	525	20.01.2016	SCR-III	UPBEANED	NAVED SMALL	Urban Enduro	MOTO STANDARD	16	SCRAMBLER	VERBODEN CALIFORNIA	Vit	Green	I	I	

Figura 5 Esempio di come si presentano i dati contenuti in file Excel

Quella sopra riportata è un'immagine relativa alla configurazione di un classico foglio Excel nel quale sono già state apportate delle modifiche manuali e aggiunte colonne calcolate. Le formule si presentano come poco leggibili e manutenibili, in una forma di questo tipo:

```

=+SE(AR2="Demo"; "DEMO";
SE(AU2<=30;"<=30gg";
SE(E(AU2>30;AU2<=60);"30-60gg";
SE(E(AU2>60;AU2<=90);"60-90gg";">90gg"))))

```

Evidentemente non sono formule particolarmente complesse, il vero problema relativo alla costruzione di queste basi dati è il rischio di commettere errori al momento della selezione dei campi corretti da inserire nelle formule. Inoltre, poiché questi fogli si basano su molteplici interventi manuali, ogni volta che si presentava la necessità di aggiornare il report, non era sufficiente un semplice refresh dei dati, ma tutto il lavoro svolto andava rifatto. Questo procedimento risultava ovviamente molto time consuming.

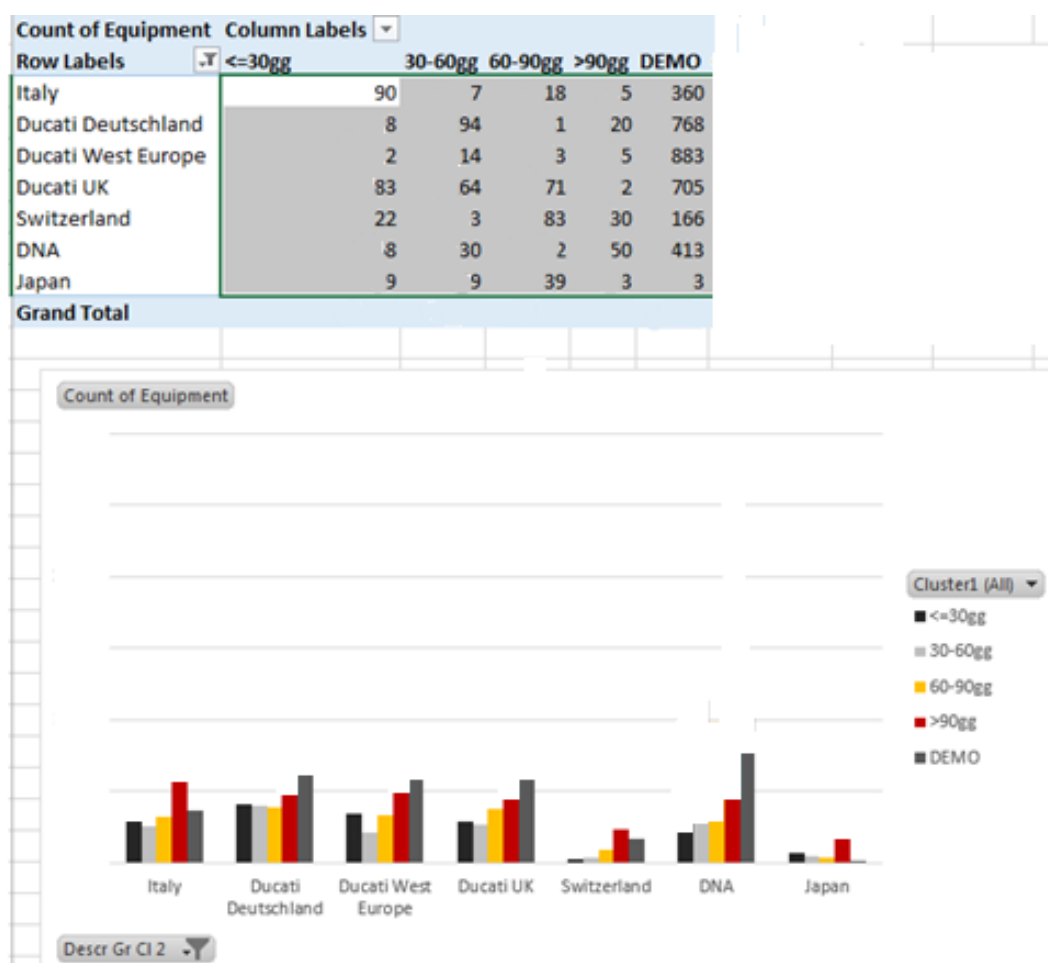


Figura 6 Esempio di tabella e grafico pivot costruito direttamente su un file Excel (\* i dati riportati non sono reali e sono stati appositamente alterati)

Altro limite dei report realizzati prima dell'introduzione della Business Intelligence, oltre alla loro staticità, è la mancanza di navigabilità. Una volta realizzato un report i dati visualizzabili erano quelli inseriti nella tabella pivot sottostante al grafico, non si poteva scendere ad un livello di dettaglio

maggiore se non realizzando un altro grafico o modificando quello già preparato.

Vantaggi di Excel	Svantaggi
Strumento già diffuso e conosciuto da tutti all'interno dell'azienda.	Staticità.
Strumento utilizzato anche dalle filiali e usato per invio di dati non presenti su SAP.	Non navigabilità.
----	Realizzazione report molto time consuming.
----	Elevato rischio di errori per la presenza di molte operazioni manuali.

## Le novità della Business Intelligence

Con l'introduzione della BI in azienda si è automatizzato l'intero flusso dati, dal momento del reperimento di tutti i file necessari fino alla visualizzazione del report finale.

Il software utilizzato per tale automazione è SAP Data Services.

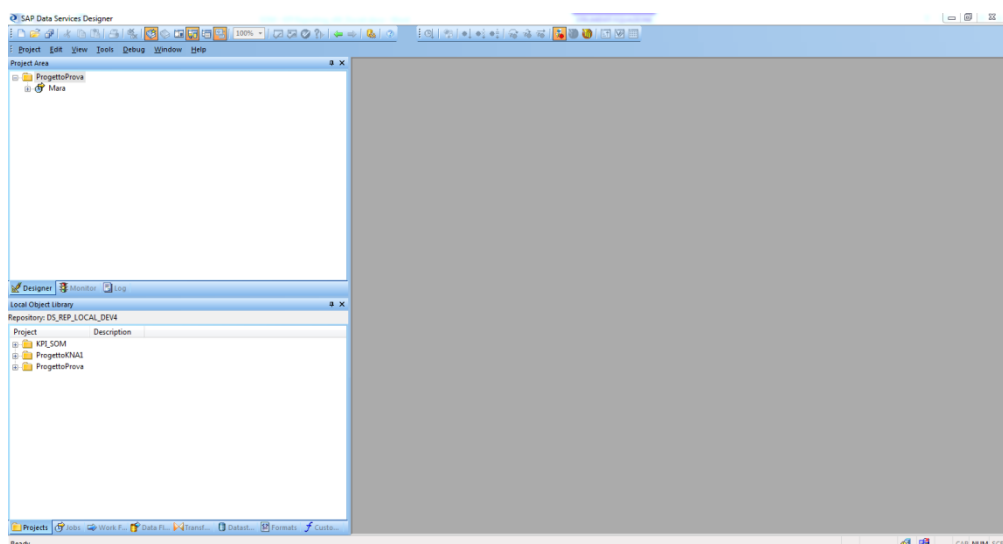


Figura 7 Schermata iniziale di SAP Data Services

Su questo tool è possibile accedere e scaricare direttamente le tabelle di interesse contenute nel database applicativo sottostante l'ERP SAP, inserire file Excel e impostare query per aggiungere e modificare i dati.

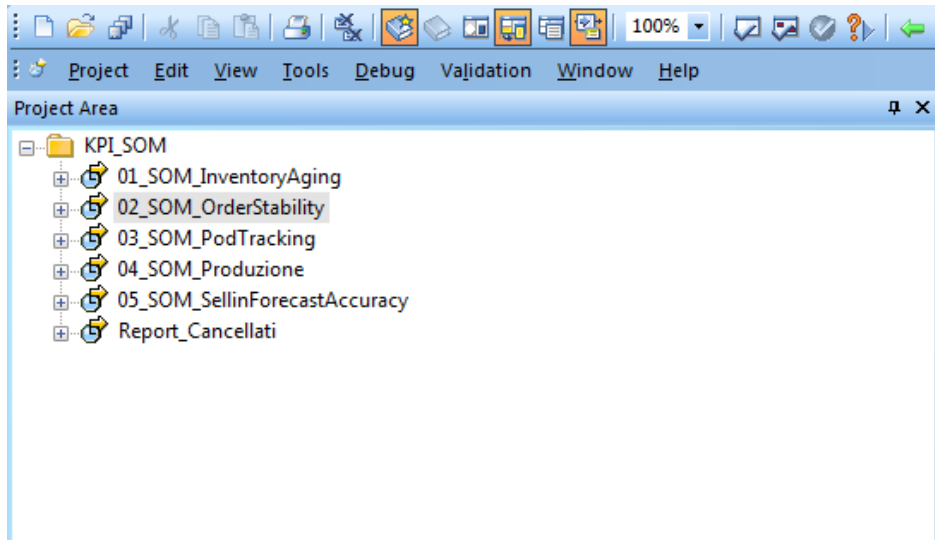


Figura 8 Elenco Job KPI

All'interno di SAP Data Services è possibile creare un job per ogni KPI che si vuole realizzare. All'interno di ogni job sono contenuti diversi Data Flow nei quali è possibile impostare tutte le operazioni che in precedenza venivano effettuate manualmente su Excel.

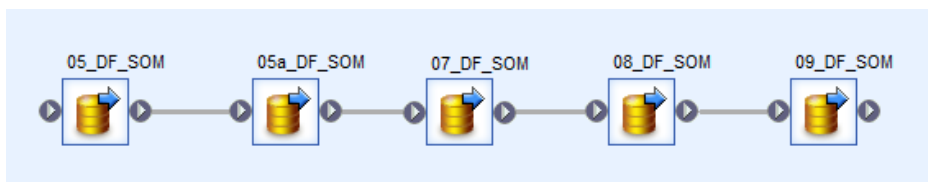


Figura 9 Flusso di Data Flow per la realizzazione del KPI Order Stability

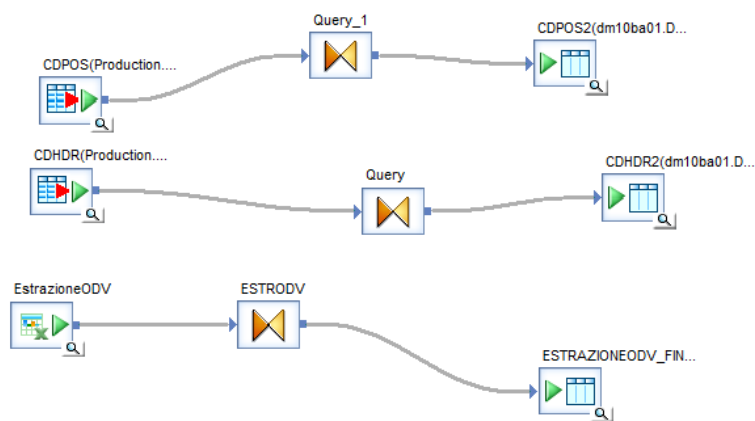


Figura 10 Operazioni effettuate del Data Flow 5 della figura 17

Nella figura precedente (figura 10) è possibile vedere alcune delle potenzialità di SAP Data Services come per esempio l'estrazione di tabelle direttamente da SAP (in questo caso si parla delle tabelle CDPOS e CDHDR), il caricamento diretto di file Excel e l'impostazione di query di modifica o di selezione dei dati.

Infatti, dato che queste tabelle hanno un elevato numero di righe, per impedire di appesantire eccessivamente il server su cui vengono caricati i dati in SAP DS, è possibile inserire delle query nelle quali specificare solo alcune delle colonne da selezionare e impostare filtri, come si vede nella figura sottostante.

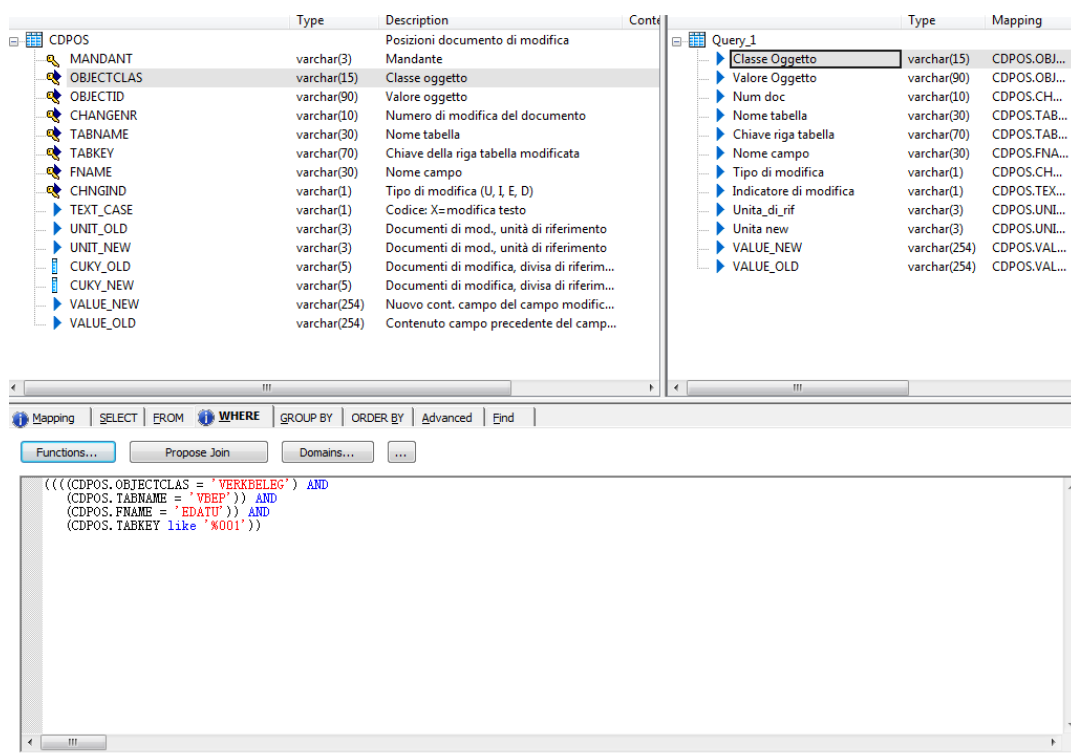


Figura 11 Schermata riportante alcuni filtri impostati all'interno di una query in SAP DS

È possibile inoltre effettuare join tra tabelle e aggiungere delle colonne calcolate, esattamente come in precedenza veniva fatto sui file Excel. Il vantaggio ora è dato dal fatto che, al momento in cui si verifica la necessità di aggiornare i report, è sufficiente rilanciare i job di SAP DS senza ulteriori interventi manuali. Il lavoro di data quality assurance quindi, viene ora impostato solamente una volta con un notevole risparmio in termini di tempo.

The screenshot displays a database query editor with two schema views: 'Schema In: CDHDR2' and 'Schema Out: Query'. The 'CDHDR2' schema lists fields like OBJECTCLAS, OBJECTID, CHANGENR, USERNAME, and UDATE. The 'Query' schema lists fields like NOME, DATA, OGGETTO, Val# ogg#, N# doc#, Nome tabella, Chiave tabella, Nome campo, UNITA\_DI\_RIF, Unità new, Nuovo Valore, Val# prec#, and Indicatore di modifica.

The SQL editor shows the following query structure:

```

SELECT FROM WHERE GROUP BY ORDER BY Advanced Find
Input schema(s) From Join rank Cache
CDHDR2 [x] 0 Automatic
CDPOS2 [x] 0 Automatic

Join pairs:
Left Join Type Right Join Condition
CDPOS2 Inner join CDHDR2 ... CDPOS2."Classe oggetto" = CDHDR2.OBJECTCLAS and
CDPOS2."Valore oggetto" = CDHDR2.OBJECTID and
CDPOS2."Num doc" = CDHDR2.CHANGENR

FROM clause:
SET('u_on_lexi' = CDPOS2."Classe oggetto" = CDHDR2.OBJECTCLAS and
CDPOS2."Valore oggetto" = CDHDR2.OBJECTID and
CDPOS2."Num doc" = CDHDR2.CHANGENR')

```

Figura 12 Esempio di join tra due tabelle

The screenshot displays a database query editor with two schema views: 'Schema In: DEALERINVENTORY02' and 'Schema Out: Query Descrizione\_Motivo\_Ordine'. The 'DEALERINVENTORY02' schema lists fields like CONCESSIONARIO, Sold to Dealer Name, Dealer Invoice doc#, Dealer Invoice date, Area Manager, Area Manager Name, Data di inizio cont# garanzia, Bike Prod# Date, Super Model, SEGMENT, Model Year, Standard nation, COLOR, Owner Country, Owner Region, INTERCHANGE, Demo Order, and DCS End-Customer. The 'Query Descrizione\_Motivo\_Ordine' schema lists fields like Descrizione\_Motivo\_Ordine, EQUIPMENT, MATERIALE, VENDITORE, CONCESSIONARIO, Sold to Dealer Name, Dealer Invoice doc#, Dealer Invoice date, Area Manager, Area Manager Name, Data di inizio cont# garanzia, Bike Prod# Date, Super Model, SEGMENT, Model Year, Standard nation, and COLOR.

The SQL editor shows the following query structure:

```

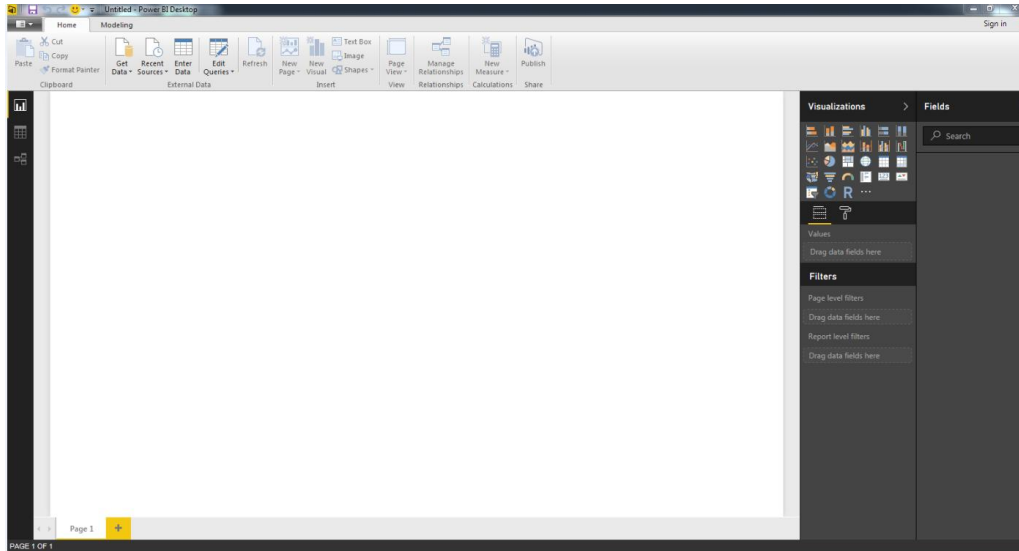
Mapping SELECT FROM WHERE GROUP BY ORDER BY Advanced Find
for: "Query Descrizione_Motivo_Ordine".Descrizione_Motivo_Ordine
Functions... Apply Schema Remapping...

ifthenelse (DEALERINVENTORY02."Motivo ordine"='ZBC' OR DEALERINVENTORY02."Motivo ordine"='ZEF', 'END CUSTOMER',
ifthenelse (DEALERINVENTORY02."Demo Order" = 'X', 'DEMO', 'STOCK'))

```

Figura 13 Esempio di colonna calcolata in SAD DS

Una volta completato il flusso dati è possibile accedere al software di front end e caricare direttamente la tabella ottenuta/ elaborata in SAD Data Services. Come detto in precedenza il programma scelto per la reportistica è Power BI, che offre differenti soluzioni in termini di visualizzazione dati (istogrammi, mappe, torte, grafici di dispersione eccetera).



*Figura 14 Schermata iniziale di Power BI*



## Front end

### Dealer Inventory Aging

I report realizzati prima della BI si basavano sulle offerte presenti in merito su Excel. In seguito riporto i grafici così realizzati:

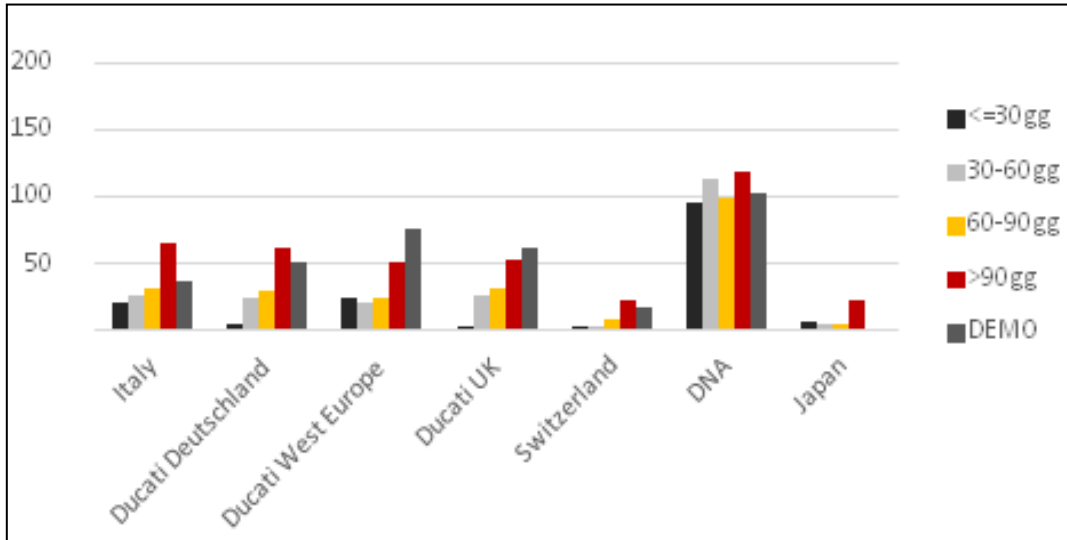


Figura 15 Report Dealer Inventory Aging Excel  
(\* i dati riportati non sono reali e sono stati appositamente alterati)

Di seguito invece le dashboard navigabili realizzabili tramite Power BI:

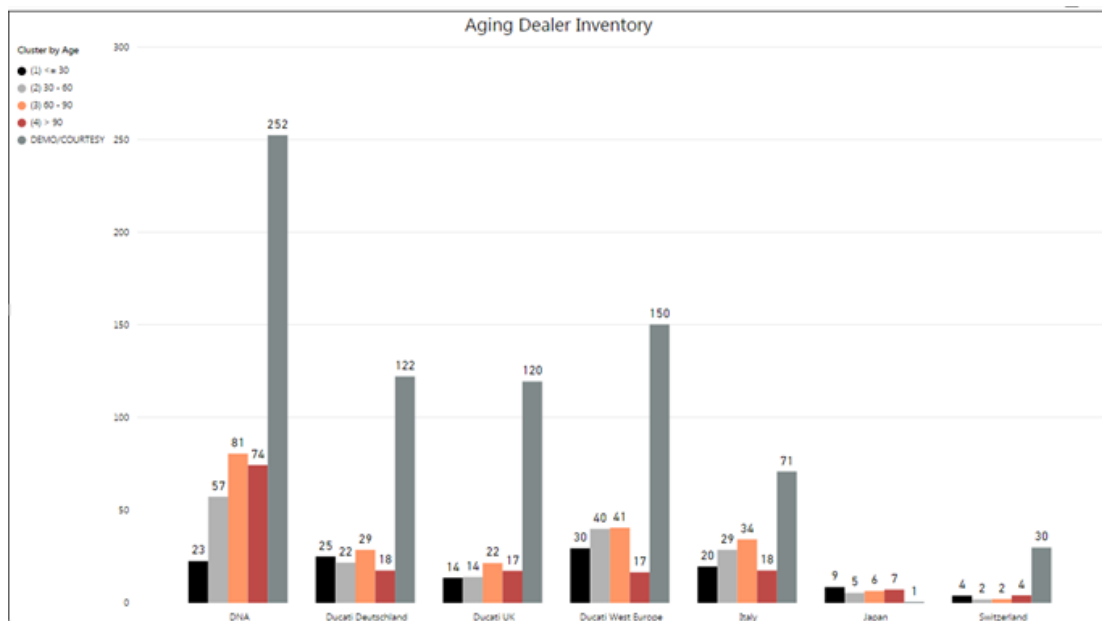


Figura 16 Report Dealer Inventory Aging Power BI  
(\* i dati riportati non sono reali e sono stati appositamente alterati)

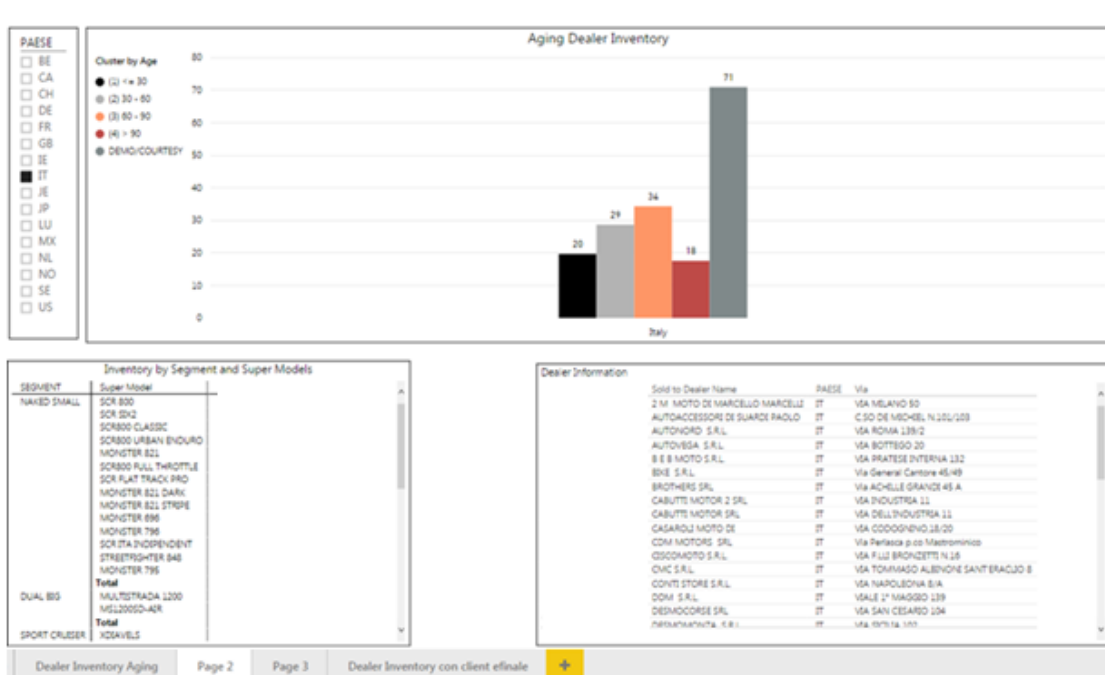


Figura 17 Esempio di dashboard navigabile  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

In figura 17 è riportata una dashboard navigabile contenente un focus sulle moto contenute negli stock dei dealer italiani. Sulla sinistra ho inserito un filtro attraverso il quale è possibile selezionare il paese che si intende porre sotto analisi (tenendo premuto il tasto *Ctrl* è possibile effettuare una selezione multipla); seguono un istogramma simile al grafico totale riportato sopra, ma con solo i dati relativi all'Italia e due tabelle, una contenente i modelli di moto costituenti lo stock dei dealer del paese in esame e l'altra l'elenco dei dealer che presentano tali giacenze.

Semplicemente selezionando un paese diverso, tutte le visualizzazioni contenute nella dashboard si aggiornano automaticamente. È inoltre possibile selezionare una sola colonna del grafico ed entrare ancora più nel dettaglio dei dati. In figura 18 ho impostato il filtro sulla Svizzera e selezionato la colonna dell'istogramma con aging dello stock tra 60 e 90 giorni. In questo modo nelle tabelle presenti nella parte inferiore del report sono visibili i dati relativi ai dealer che contengono quelle 2 moto e i modelli di quest'ultime.

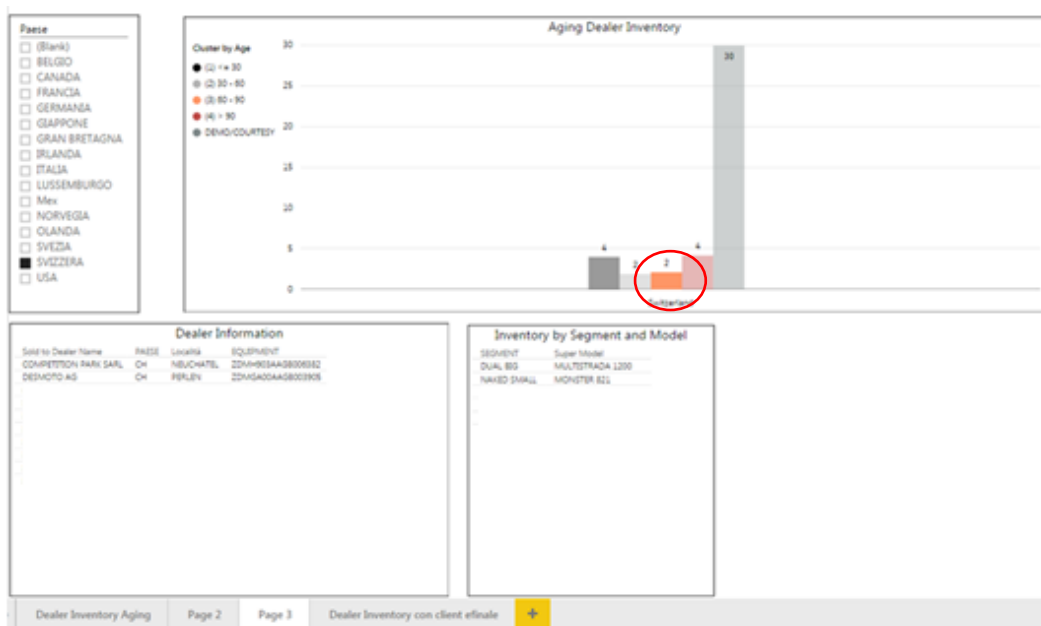


Figura 18 Dashboard con focus su Svizzera  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

### Pod Tracking

Nell'istogramma sono riportate le cumulate degli ordini spediti suddivisi per settimana dell'anno mentre le spezzate rappresentano la percentuale degli ordini confermati tramite l'applicazione POD sul totale. Ogni colore è rappresentativo dell'andamento di una diversa filiale.

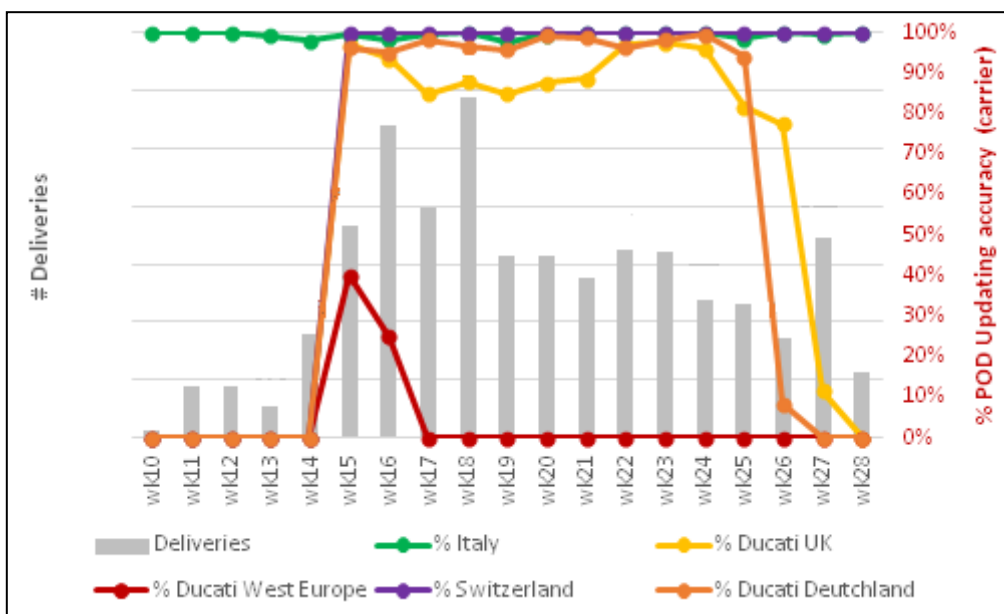


Figura 18 Report POD Tracking Excel  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

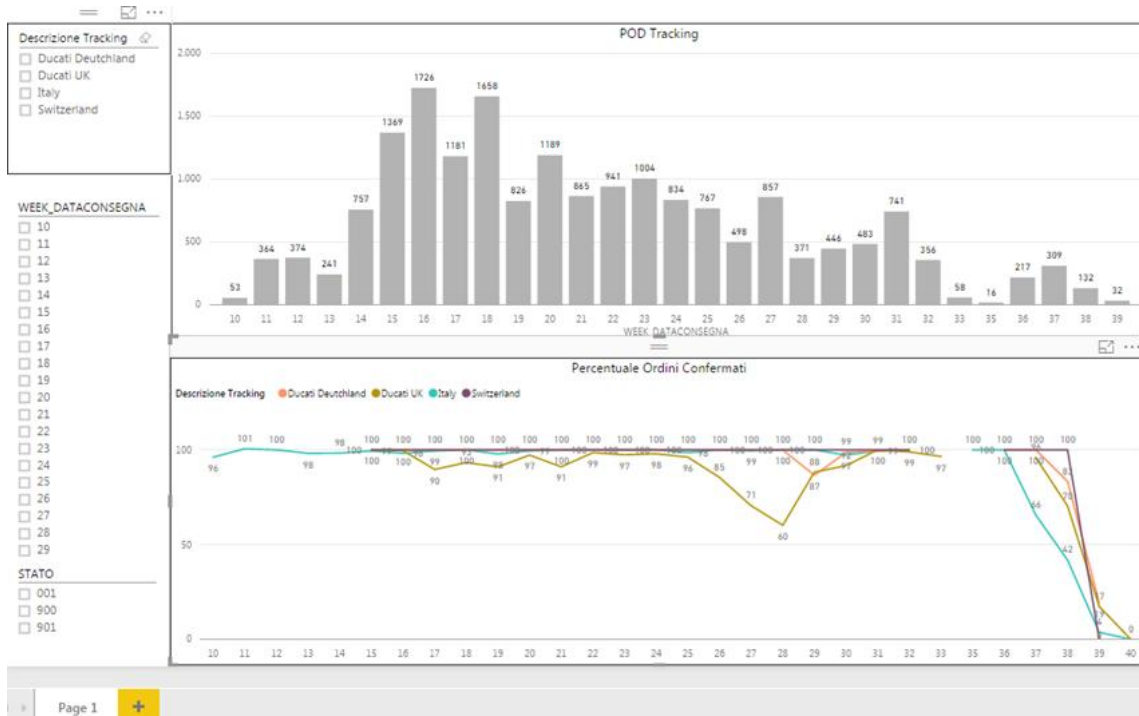


Figura 19 Report POD Tracking Power BI  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

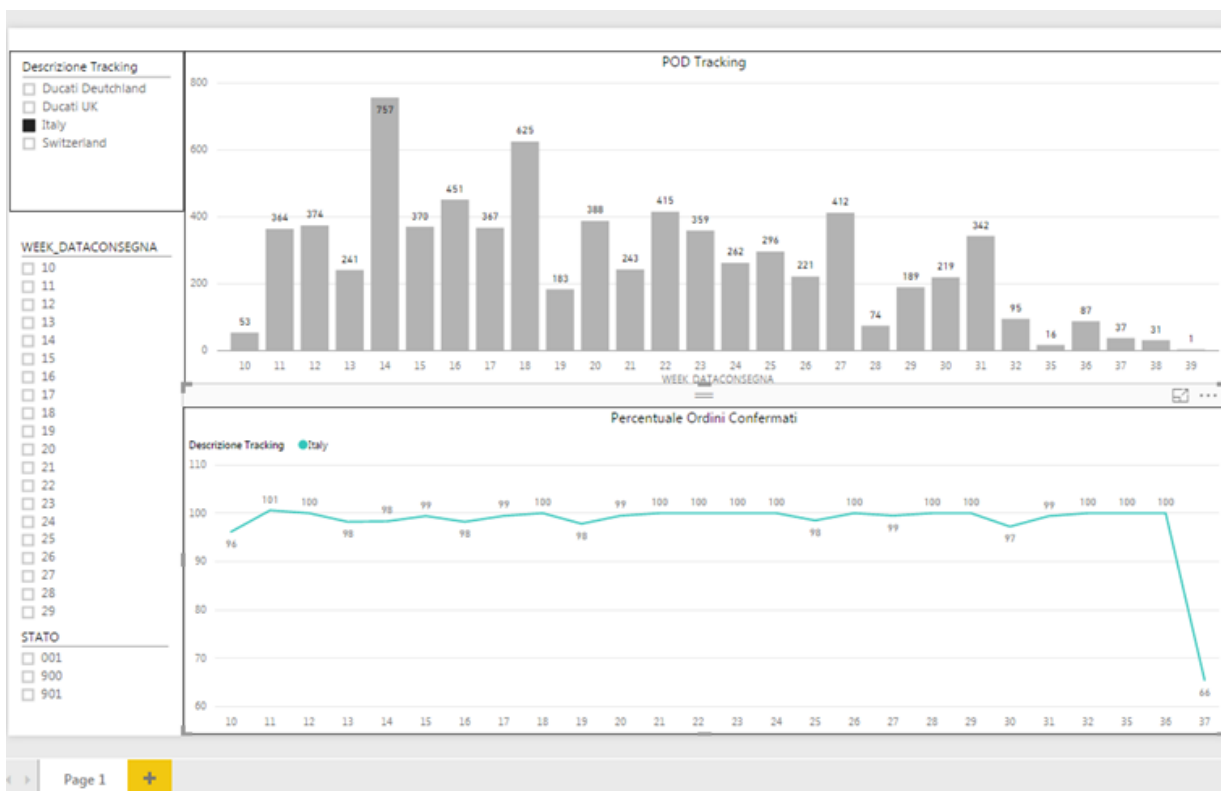


Figura 20 Report POD Tracking con focus sull' Italia  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

## Order Stability

Nell'istogramma è riportata la cumulata mensile degli ordini facendo una distinzione tra quota parte di standard orders e quella di stocking orders. Le due spezzate rappresentano invece le percentuali degli ordini che erano originariamente previsti per quel mese, ma hanno subito in seguito una cancellazione o una modifica.

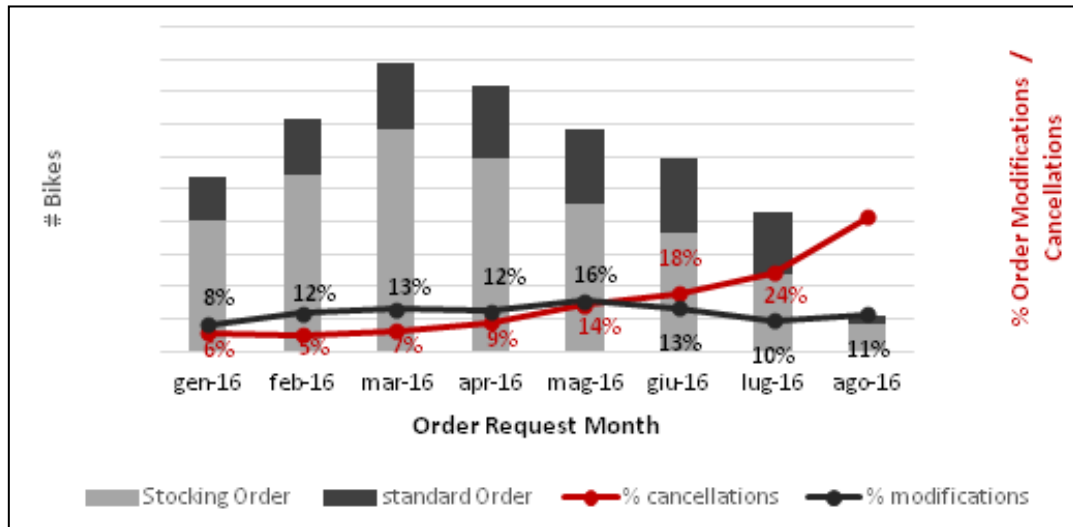


Figura 21 Report Order Stability Excel  
(\* i dati riportati non sono reali e sono stati appositamente alterati)

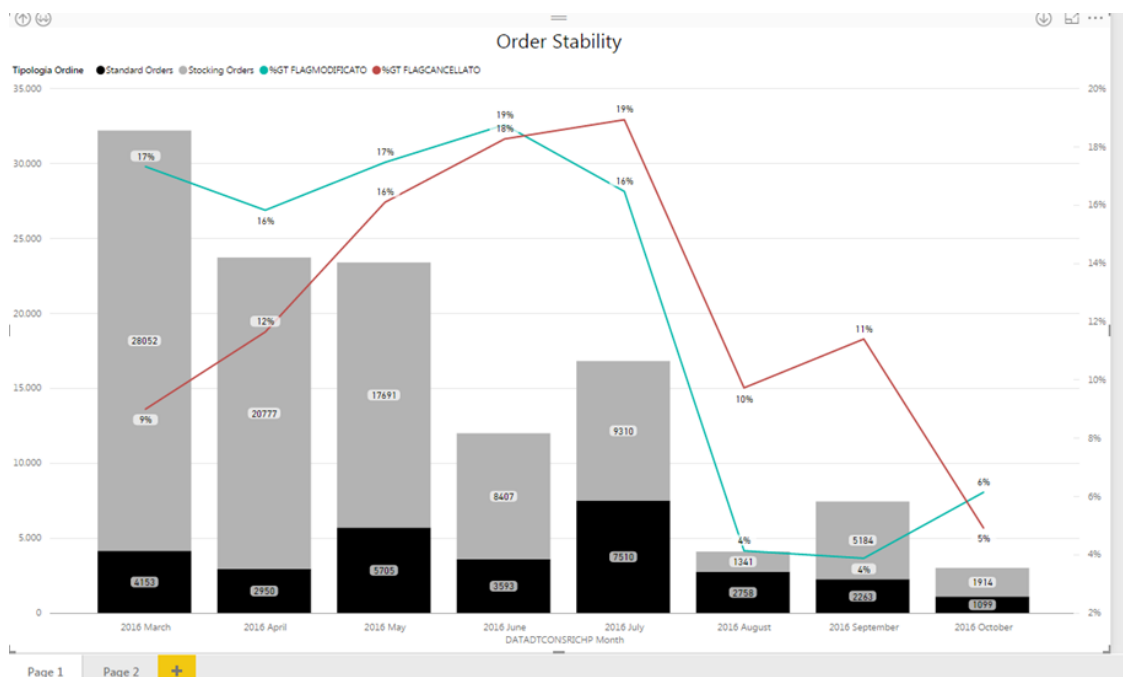


Figura 22 Report Order Stability Power BI  
(\* i dati riportati non sono reali e sono stati appositamente alterati)

In seguito a richieste aggiuntive giunte dall'area sales ho messo a punto due ulteriori dashboard con focus sugli ordini cancellati. Oltre agli ordini totali cancellati è possibile vedere quale country ne ha registrato il maggior numero, quali modelli di bikes sono stati coinvolti e quanto tempo prima della data di prevista consegna è avvenuta tale cancellazione, per valutare l'effetto provocato sulla produzione.

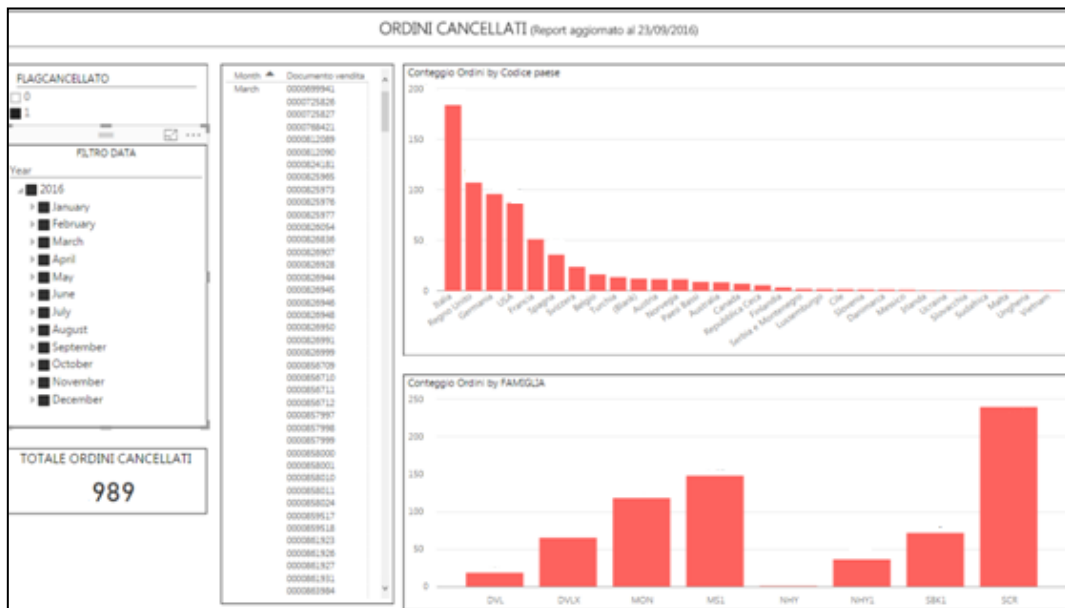


Figura 23 Ordini cancellati per country e modello  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

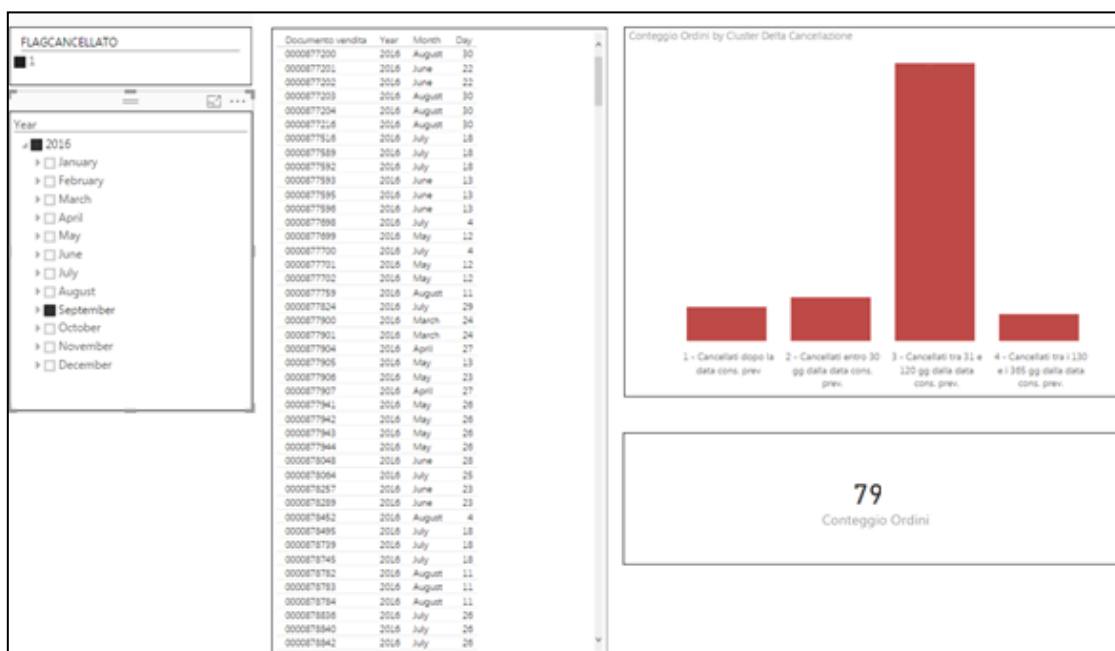


Figura 24 Report ordini cancellati con preavviso sulla data di consegna  
 (\* i dati riportati non sono reali e sono stati appositamente alterati)

## *Conclusioni*

<i><b>PRE-BI</b></i>	<i><b>POST-BI</b></i>
Elevato rischio presenza errori.	Rischio presenza errori basso.
Elaborazioni manuali da ripetere ad ogni aggiornamento report.	Elaborazioni sui dati automatizzate ed impostate solo una volta valida per tutti i successivi aggiornamenti.
Report statici.	Dashboard navigabili contenenti molteplici dati.
Attività di reporting molto time consuming.	Notevole risparmio di tempo.
---	Possibilità di visualizzare le dashboard su tablet e cellulari.
---	Dashboard sempre disponibili online.

Il progetto di Business Intelligence / Analytics in Ducati è appena stato avviato e questi sono i primi risultati operativi utili per il miglioramento dei processi interni.

I prossimi passi di progetto prevedono di:

- approfondire maggiormente l'analisi coinvolgendo i key user dell'area sales nella definizione di KPI aggiuntivi e di visualizzazioni ulteriori, anche al fine di renderli autonomi nello sviluppo di nuove dashboard;
- replicare l'approccio analitico implementato nelle altre aree dell'azienda seguendo la roadmap definita attraverso l'indice impact \ effort trattato al capitolo 4.

## Bibliografia

- [1] M. Golfarelli e S. Rizzi, Data Warehouse, teoria e pratica della progettazione.
- [2] «Search Data Management - Business Intelligence,» [Online]. Available: <http://searchdatamanagement.techtarget.com/definition/business-intelligence>.
- [3] «History of BI,» [Online]. Available: <https://www.betterbuys.com/bi/history-of-business-intelligence/>.
- [4] «Search Business Analytics,» [Online]. Available: <http://searchbusinessanalytics.techtarget.com/definition/business-analytics-BA>.
- [5] «Search Data Management,» [Online]. Available: <http://searchdatamanagement.techtarget.com/definition/business-intelligence>.
- [6] «Gartner,» [Online]. Available: <https://www.gartner.com/>.
- [7] «Sistemi di reportistica,» [Online]. Available: [https://it.wikipedia.org/wiki/Sistemi\\_di\\_reportistica](https://it.wikipedia.org/wiki/Sistemi_di_reportistica).
- [8] «La business intelligence per lo sviluppo di politiche competitive,» [Online]. Available: <http://www.slideshare.net/SimcoConsulting/la-business-intelligence-per-lo-sviluppo-di-politiche-competitive>.
- [9] «B.I. Strategy BUSINESS INTELLIGENCE Introduzione al Data Warehouse,» [Online]. Available: <http://slideplayer.it/slide/939088/>.
- [10] [Online]. Available: [https://it.wikipedia.org/wiki/Statistica\\_descrittiva](https://it.wikipedia.org/wiki/Statistica_descrittiva).



- [11] [Online]. Available: [https://it.wikipedia.org/wiki/Big\\_data](https://it.wikipedia.org/wiki/Big_data).
- [12] [Online]. Available: [http://www.tuttosport.com/news/motori/due-ruote/mercato/2015/07/03-2161962/ducati\\_mercato\\_successo\\_scrambler\\_e\\_non\\_solo/?cookieAccept](http://www.tuttosport.com/news/motori/due-ruote/mercato/2015/07/03-2161962/ducati_mercato_successo_scrambler_e_non_solo/?cookieAccept).
- [13] [Online]. Available: <http://www.gazzetta.it/Passione-Motori/Moto/14-01-2016/ducati-2015-vendite-record-scrambler-regina-140261702594.shtml>.
- [14] [Online]. Available: [http://www.slideshare.net/sfamiliar/working-with-big-data-jan-2016-part-1/8-CONTEXT\\_WHATS\\_BIG\\_DATATHE\\_FOUR](http://www.slideshare.net/sfamiliar/working-with-big-data-jan-2016-part-1/8-CONTEXT_WHATS_BIG_DATATHE_FOUR) .
- [15] [Online]. Available: [http://www.slideshare.net/businessintelligence/business-intelligence-bi-for-automotive?qid=b4654a7a-bd49-420e-aa84-bada2812c61f&v=&b=&from\\_search=8](http://www.slideshare.net/businessintelligence/business-intelligence-bi-for-automotive?qid=b4654a7a-bd49-420e-aa84-bada2812c61f&v=&b=&from_search=8) .
- [16] [Online]. Available: <http://www.slideshare.net/sales-i/how-is-big-data-impacting-sales-in-the-automotive-industry>.
- [17] [Online]. Available: <http://social.technet.microsoft.com/wiki/contents/articles/25319.business-intelligence-roles-and-team-composition.aspx>.
- [18] [Online]. Available: <http://www.primitivelogic.com/2015/04/21/what-is-pixel-perfect-reporting-why-is-it-important/>.
- [19] [Online]. Available: <http://heavyrider.corriere.it/2014/10/22/ducati-scrambler-successo-annunciatoimparando-anche-dagli-errori-del-passato/>.
- [20] [Online]. Available: <http://searchcio.techtarget.com/definition/Information-Age>.