

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Triennale in Informatica per il management

Business Analytics

Tesi di Laurea in Basi di Dati

Relatore:
Chiar.mo Prof.
Danilo Montesi

Presentata da:
Alice Casari

Sessione I
Anno Accademico 2011/2012

Indice

1	Introduzione	5
2	Business Analytics	7
2.1	Business Analytics: la BI che guarda avanti	7
2.2	Definizione e caratteristiche della Business Analytics	8
2.3	Il modello della Business Analytics	10
2.3.1	Caso di studio	11
2.4	Il livello strategico del modello della Business Analytics	16
2.5	Il livello funzionale del modello della Business Analytics	19
2.5.1	Caso di studio: CRM	20
2.6	Il livello analitico del modello della Business Analytics	23
2.6.1	Data Mining	28
2.7	Il livello di data warehouse del modello della Business Analytics	29
2.8	Il livello delle fonti dei dati del modello della Business Analytics	31
2.9	Valutazione e priorità dei progetti di Business Analytics	33
3	Sistemi software per la Business Analytics	37
4	Business Analytics con Pentaho	39
4.1	Creare report con Pentaho	41
4.2	Creare analisi di report con Pentaho	50
4.3	Creare dashboard con Pentaho	56
4.4	Conclusioni	56
5	Ringraziamenti, bibliografia ed elenco delle figure	59

Capitolo 1

Introduzione

La Business Analytics si occupa di attività di analisi strategiche per generare vantaggi competitivi sia per le grandi imprese, sia per le piccole e medie imprese.

Introduciamo la Business Analytics partendo dal suo predecessore la Business Intelligence e spiegandone gli innegabili vantaggi apportati mettendo queste due metodologie a confronto.

La Business Analytics è rappresentata da un modello guida composto da vari livelli, ognuno con le proprie caratteristiche che è opportuno approfondire attraverso anche semplici esempi.

Concludiamo con un elenco dei software più famosi e utilizzati per mettere in pratica la Business Analytics e descrivendo in dettaglio uno di questi.

Visto il crescente interesse e utilizzo di questa nuova metodologia possiamo dire che la Business Analytics sarà ovunque e in ogni momento.

Capitolo 2

Business Analytics

In tempi di rapidi cambiamenti e di crescente complessità, un rapido apprendimento diventa importante. Per questo è necessaria una visione strategica su ciò che è necessario attuare, per ottenere un rapido apprendimento e infine, la creazione di valore.

Tutto questo diventa possibile attraverso la Business Analytics.

2.1 Business Analytics: la BI che guarda avanti

Il successo di una qualsiasi attività, specie nel business, dipende in massima parte dalla capacità di prevedere il futuro cogliendone i segni anticipatori. Questi segni esistono: molti sanno dire, a cose fatte, quali erano i segnali che annunciavano l'accaduto. Quello che fa la differenza è saperlo prima.

Lo sviluppo delle applicazioni di business intelligence sono state il primo, grande, risultato verso l'individuazione e lo studio dei segnali nascosti. La BI è un insieme di metodologie, processi, architetture e tecnologie che trasformano dati grezzi in informazioni significative e utili, attraverso l'utilizzo di Data Warehouse e Data Mining per l'estrazione e l'analisi dei dati.

Le applicazioni di Business Intelligence, se progettate bene e basate su dati aggiornati, offrono una visione approfondita del passato o anche del presente, identificando gli schemi che determinano comportamento e prestazioni delle funzioni aziendali: vendite, produzione, approvvigionamento, marketing, amministrazione, finanza e così via.

Possono quindi dirci cosa è successo e perché (il che non è poco), ma la BI ha dei limiti fondamentali in quanto non è in grado né di dare una visione di

possibili eventi futuri né di prevedere gli sviluppi futuri di un'azione. Sono punti fondamentali per chi deve decidere, si tratta di sostituire una scelta soggettiva basata sull'intuito con una basata su alternative analizzate per quanto possibile in modo oggettivo.

La BI risponde a domande come:

- Cosa è successo?
- Quante volte?
- Dove è il problema?
- Quali azioni sono necessarie?

La Business Analytics è una disciplina avanzata della BI, dove i vantaggi competitivi delle aziende derivano dal ricorso di sistemi più avanzati di studio delle informazioni. La BA risponde a domande come:

- Cosa sta succedendo ora?
- Perché sta succedendo?
- Cosa succederà dopo? (previsioni)
- Qual'è la cosa migliore che può accadere? (ottimizzazioni)

Approfondiamo quindi la Business Analytics al fine di mettere più attenzione agli elementi mancanti della BI che sono ormai fondamentali in questa era basata sulle analisi.

2.2 Definizione e caratteristiche della Business Analytics

La BA può essere definita tramite questa frase: “dare il giusto supporto decisionale alle persone giuste al momento giusto”.

Le analisi di business consistono nel migliorare il processo decisionale, i suoi processi operativi e la competitività ottenuta quando un'impresa è in possesso di fatti rilevanti e sa come usarli.

Le iniziative di BA che hanno successo sono sempre strettamente collegate con la strategia dell'organizzazione (missione, visione, e gli obiettivi) e vengono messe in atto per rafforzare la capacità dei processi di business di muoversi nella giusta direzione verso i propri obiettivi.

2.2. DEFINIZIONE E CARATTERISTICHE DELLA BUSINESS ANALYTICS⁹

La Business Analytics consente alle organizzazioni di “guardare in avanti” e ne beneficiano tutti gli aspetti della catena del valore di un’organizzazione, tra cui:

- Logistica in entrata: ricezione, stoccaggio, controllo del magazzino e pianificazione dei trasporti.
- Operazioni: tra cui manutenzione imballaggi, apparecchiature, test e tutte le attività che aggiungono valore dalla materia prima al prodotto finale.
- Logistica in uscita: le attività necessarie per ottenere i prodotti finiti sul mercato, tra cui lo stoccaggio e gestione della distribuzione.
- Marketing e vendite: attività rivolte all’acquirente per l’acquisto del prodotto, compresa la selezione dei canali, pubblicità, promozione, vendita, prezzi e ottimizzazione degli spazi.
- Servizi: attività che mantengono il valore di un prodotto, incluso il supporto ai clienti, riparazioni, installazioni, formazione, gestione dei ricambi e altro ancora.

In questo modo, la BA guida l’innovazione e migliora la velocità di risposta ai cambiamenti del mercato e dell’ambiente di un’organizzazione.

I limiti della BI vengono quindi superati permettendo di anticipare lo sviluppo degli eventi e di prevedere l’effetto che potranno avere le decisioni in futuro.

Per esempio:

- nel campo finanziario si possono prevedere le reazioni e i commenti degli analisti al prossimo annuncio dei risultati della società;
- nel campo del marketing si può prevedere quali prodotti, quali prezzi e quali sistemi di vendita permetteranno di ottenere una maggiore fedeltà alla marca;
- nel campo dei rischi manageriali si potranno prevedere e prevenire frodi, sprechi, abusi e minacce;
- e tanto altro...

Possiamo anche dire che la BA consiste in una continua esplorazione interattiva e analisi delle performance di business del passato, al fine di conoscere e di migliorare la pianificazione aziendale attuale e futura, attraverso la ricerca di eccezioni e anomalie e la ricerca di nuovi pattern ottimali.

2.3 Il modello della Business Analytics

Il modello della BA fornisce uno schema per la comprensione e la creazione di analisi di successo in qualsiasi tipo di organizzazione.

La Business Analytics è una disciplina stratificata e gerarchica, le informazioni richieste si muovono dall'alto (l'ambiente guidato dalle imprese) verso il basso (l'ambiente guidato dalla tecnica), mentre il flusso successivo di informazioni si spostano dal basso verso l'alto.

Vediamo l'immagine che rappresenta questo modello.

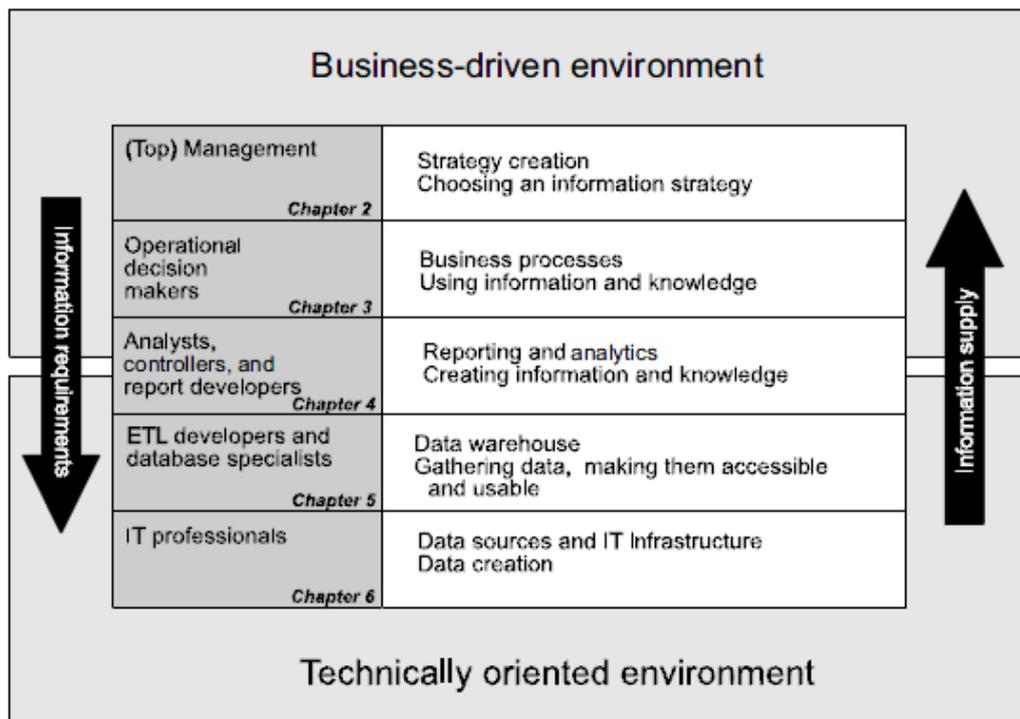


Figura 2.1: Il Modello della Business Analytics

Nel **primo strato** vengono prese le decisioni strategiche e vengono scelte le informazioni strategiche.

Il **secondo livello** necessita di informazioni e conoscenza per supportare la scelta della strategia aziendale.

Nel **terzo strato** gli analisti, coloro che controllano e gli sviluppatori di report creano informazioni e conoscenza per innovare e ottimizzare le loro attività.

Nel **penultimo livello** si trova l'ambiente operativo dove vengono lavorati i dati attraverso datawarehouse e gli specialisti rendono questi dati accessibili

agli altri.

Nell'**ultimo strato** del modello i dati primari generati da risorse vengono gestiti e sviluppati.

La strategia aziendale (visione, missione, obiettivi), di cui è responsabile il primo livello, normalmente viene tradotta in una serie di indicatori di prestazione chiamati **KPIs**, con lo scopo di misurare il grado di processo e di successo.

Il contenuto di questi indicatori dipende da quale processo di business vogliamo controllare.

L'obiettivo della BA è quello di cambiare i processi di business e le azioni, in modo che siano mirate verso il raggiungimento degli obiettivi standard.

2.3.1 Caso di studio

Introduciamo un esempio semplificato per comprendere meglio il ruolo dei KPIs per ottenere obiettivi strategici.

In questo esempio abbiamo un'impresa che produce e vende mezzi di trasporto (auto, moto etc). Di questi prodotti è già disponibile un'analisi in cui si vedono le vendite totali di ogni articolo e un'analisi dove si vede la distribuzione delle vendite nei vari territori.

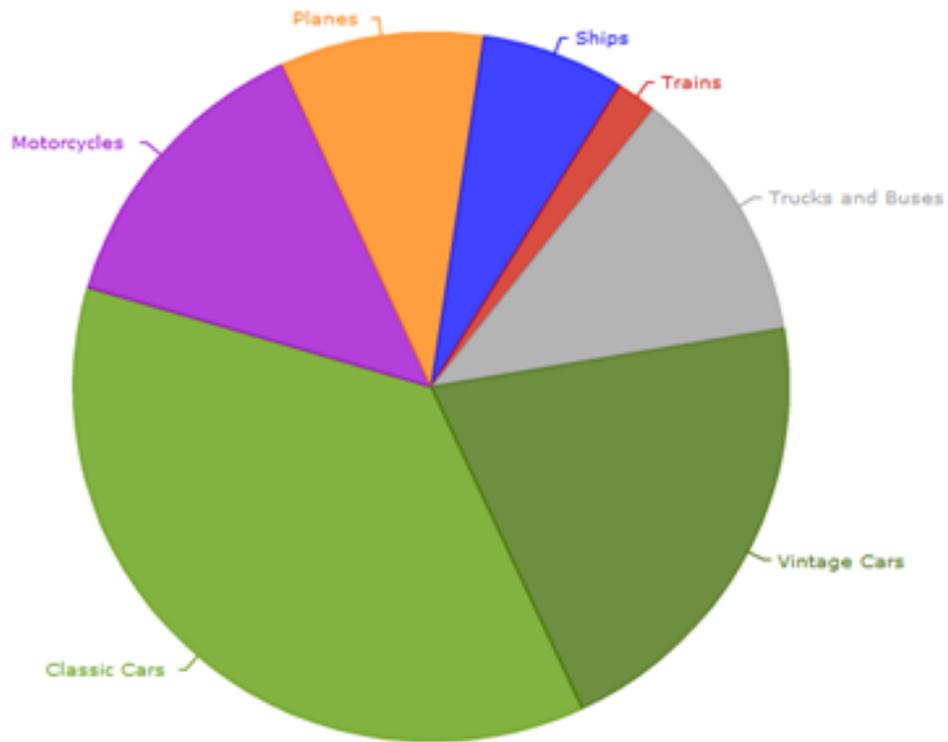


Figura 2.2: Linee di prodotto - Vendite totali

Questo grafico a torta rappresenta la suddivisione delle vendite totali dell'azienda in base al tipo di prodotto venduto.

Per esempio le vendite maggiori sono del tipo Classic Card mentre le vendite minori sono del tipo Trains.

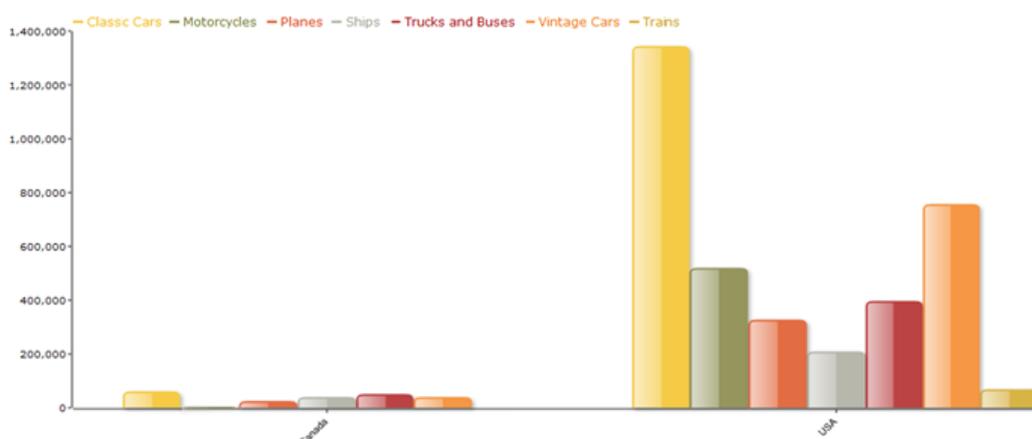


Figura 2.3: Linee di prodotto - Vendite in base al territorio

Questo grafico a barre invece mostra la suddivisione delle vendite totali dei vari prodotti in base al territorio di vendita.

Nell'asse delle ascisse sono presenti i due territori di vendita (Canada, USA), mentre nell'asse delle ordinate sono rappresentati i ricavi totali delle vendite.

Gli obiettivi di questa impresa sono di raggiungere una vendita media di 4100 prodotti, una cancellazione di ordini pari a 0, e di raggiungere una media di articoli per vendita di 38.

Allo stato attuale la vendita media di prodotti è 3800, le cancellazioni degli ordini sono 6, e la media di articoli per vendita è di 30.

La strategia di business e gli obiettivi sono quindi presentati tramite i seguenti parametri o KPI:

- **KPI 1:** Vendita media = 4100. Valore attuale = 3800.
- **KPI 2:** Cancellazioni = 0. Valore attuale = 6.
- **KPI 3:** Articoli per vendita = 38. Valore attuale = 30.

Quello che ci si aspetta è che una diminuzione totale delle cancellazioni e quindi una maggiore soddisfazione dei clienti porti automaticamente ad un aumento delle vendite e anche degli articoli di ogni vendita.

Si tratta quindi di rapportarsi con i clienti in modo più efficiente.

Bisogna cercare i clienti più fedeli e dargli maggiori servizi e maggiori offerte per evitare una loro "fuga", bisogna acquisire nuovi clienti e fidelizzarli,

bisogna cercare l'offerta giusta per ogni territorio in base alle loro caratteristiche.

La necessità è quindi quella di segmentare i clienti e segmentare il territorio. Le domande primarie per l'analisi sono:

- Chi sono i miei clienti?
- Quali sono le loro preferenze?
- Chi compra cosa e quando?
- In che territori c'è maggior vendita?
- Quali sono le preferenze dei clienti nei vari territori?

L'analista si rende conto rapidamente che egli non è in possesso di dati sufficienti circa i clienti e i territori per essere in grado di elaborare i loro profili. Questi dati potrebbero però essere immessi in un data warehouse e, successivamente, costituirebbero la base della creazione di conoscenza dei profili dei clienti per i diversi territori di vendita.

In ambiente analitico, l'analista ha ora accesso ai dati e inizia a trasformare i dati raccolti in informazioni e conoscenza.

Attraverso lo studio dei profili dei clienti si possono indagare le preferenze, la classe sociale, la cultura e tutto questo porta a offerte mirate e precise.

Lo studio dei territori può portare alla luce necessità di veicoli particolari per un territorio montuoso o di pianura, con un clima secco o piovoso, tutte caratteristiche che permettono all'impresa di avere piena conoscenza di quale prodotto è giusto ed aumentare la soddisfazione dei clienti.

Dopo sei mesi di attuazione di questa strategia in continua evoluzione con i dati prodotti dalla BA l'impresa è arrivata ad avere i seguenti valori obiettivo, mostrati dai "cruscotti" della BA.



Figura 2.4: KPI 1 : Vendita media dei prodotti

L'obiettivo KPI1 era di 4100 vendite medie per prodotto. Dopo sei mesi il valore raggiunto è quello mostrato dalla figura.



Figura 2.5: KPI 2: Cancellazione degli ordini

L'obiettivo KPI2 era di 0 cancellazioni dei prodotti da parte dei clienti. Dopo sei mesi il valore raggiunto è quello mostrato dalla figura.



Figura 2.6: KPI 3: Articoli per vendita

L'obiettivo KPI3 era di vendere 38 articoli per tipologia di prodotto. Dopo sei mesi il valore raggiunto è quello mostrato dalla figura.

Come possiamo vedere i risultati sono molto buoni, i valori non sono precisamente quelli indicati come obiettivo, ma l'azienda è sulla buona strada per raggiungere i propri obiettivi strategici generali.

Senza la BA non si sarebbe potuti arrivare a questa strategia e a queste risorse di dati e sicuramente non si sarebbe arrivati a questo risultato.

2.4 Il livello strategico del modello della Business Analytics

Il livello strategico non si focalizza solamente su come sviluppare una strategia di business, ma dimostra l'esistenza di importanti relazioni tra le strategie di business globali e le informazioni che la funzione di BA è in grado di fornire in questo contesto.

Due sono le domande fondamentali:

- Come può la funzione di BA influenzare il processo di strategia globale dell'organizzazione?
- In che modo la strategia aziendale globale influenza la funzione di BA?

La strategia aziendale ha di solito degli obiettivi specifici, che sono spesso supportati da alcune aspettative, su come ogni singolo reparto dovrebbe raggiungere questi obiettivi.

Ci sono diversi gradi di integrazione tra la funzione di BA e la strategia della

2.4. IL LIVELLO STRATEGICO DEL MODELLO DELLA BUSINESS ANALYTICS¹⁷

società e questo serve per capire il pieno potenziale di un'integrazione completa.

Mostriamo questi diversi gradi di integrazione attraverso quattro scenari.

Primo scenario Nel primo scenario non esiste alcun legame tra la funzione di BA e la strategia della società. Si tratta di aziende che hanno pochi dati per cui è impossibile utilizzare la BA come strategia, oppure si tratta di piccole imprese dove il costo di gestione di un data warehouse è più grande rispetto al valore del supporto decisionale creato.

Queste aziende devono maturare questi limiti o continuare con strategie non legate all'informazione.

Secondo scenario Nel secondo scenario le funzioni di BA vengono utilizzate per monitorare la realizzazione delle funzioni dei singoli obiettivi decisi dalla strategia. Oltre a monitorare però la funzione di Business Analytics non dà alcun contributo alla creazione della strategia.

La funzione BA è quindi unicamente impiegata per il controllo del raggiungimento degli obiettivi definiti dalla strategia.

Gli obiettivi espressi in parametri hanno però dei requisiti da rispettare, devono essere:

- **Specifici:** obiettivi, come ad esempio quanti clienti dobbiamo avere entro la fine dell'anno, a quando dovrà ammontare il nostro fatturato, etc.
- **Misurabili:** se non è misurabile, non è un obiettivo rilevante. Se non sappiamo quanti clienti abbiamo, dobbiamo trovare altri obiettivi.
- **Concordati:** l'organizzazione deve accettare gli obiettivi. Se non vengono accettati saranno ignorati o neutralizzati.
- **Realistici:** gli obiettivi devono essere realistici, se non hanno alcuna possibilità di essere raggiunti devono essere cambiati con altri obiettivi.
- **Vincolati nel tempo:** stabilire i tempi necessari e i tempi massimi per raggiungere gli obiettivi.

Terzo scenario Nel terzo scenario esiste invece un dialogo continuo che fa sì che le singole funzioni si ottimizzino sulla base delle informazioni prodotte dalla BA. La Business Analytics supporta quindi attivamente la strategia aziendale migliorandone le prestazioni.

L'impresa come punto di partenza definisce una strategia da attuare nei vari reparti che compongono l'azienda. Il coordinamento avviene attraverso l'identificazione dei cosiddetti fattori critici di successo, che sono gli elementi che sono essenziali per avere una strategia di successo.

È in questa fase che definiamo le varie KPI.

Quando la strategia è messa in moto, il progresso viene misurato su base continuativa. In generale, gli indicatori KPI sono raramente raggiunti con precisione, e si cerca di analizzarli per capirne il motivo. “Abbiamo trascurato qualche opportunità potenziale, o forse mancano alcune competenze nell'organizzazione?”

L'ottimizzazione della strategia avviene quando usiamo questo tipo di apprendimento per migliorare i nostri processi di business e garantire così che l'organizzazione mantenga la sua agilità tra i processi di strategia annuale.

Quarto scenario Nell'ultimo scenario l'informazione scaturita dalla BA è utilizzata come una risorsa strategica e può essere usata per determinare la strategia.

Queste aziende analizzano sistematicamente le opportunità e le minacce del mercato, i punti di forza e di debolezza e per questo possono ottenere un concreto vantaggio competitivo.

Questa prospettiva è prioritaria per le imprese che agiscono sulla base di innovazione di prodotto o strategie di leadership di prodotto.

Dal momento che le opportunità strategiche dipendono anche dal modo in cui un'azienda ha deciso di competere, bisogna identificare il rapporto tra le proprie discipline competitive e le informazioni potenzialmente più rilevanti. Esempi: l'innovazione di prodotto, analisi di business in grado di fornire informazioni sulla capacità dei prodotti di creare reddito durante il loro intero ciclo di vita, ottenere informazioni su quale prodotto sviluppare e per quale segmento di clientela, etc.

2.5 Il livello funzionale del modello della Business Analytics

Nel secondo livello del modello della BA si identifica come arrivare ad avere alcuni obiettivi generali per i reparti ed essere in grado di specificare i requisiti di informazione.

Quindi, viene specificato quali informazioni servono per realizzare gli obiettivi dei dipartimenti sulla base della strategia aziendale.

A questo punto è importante distinguere due tipi di informazioni diverse: **lead information** e **lag information**.

Le lead information rappresentano qualcosa che viene prima, e le lag information descrivono qualcosa che viene dopo.

Le lead information descrivono le informazioni o le conoscenze necessarie per iniziare con un nuovo processo o migliorarne uno già esistente. L'opposto è l'informazione lag, che è la continua misurazione di come il processo si sta sviluppando.

Lo scopo delle informazioni lag è, quindi, di controllare se stiamo raggiungendo i nostri obiettivi o se abbiamo bisogno di apportare alcune modifiche. In altre parole, stiamo parlando di conoscenza proattiva, informazioni da utilizzare per creare nuovi processi e conoscenza reattiva, che monitora i processi che sono già attivi e funzionanti.

Le lead information sono quindi più astratte mentre le lag information sono una segnalazione automatizzata su indicatori chiave, che indicano se quel processo è riuscito a conseguire gli obiettivi stabiliti.

Se è la prima volta che affrontiamo un problema, dobbiamo avviare un processo nuovo, perché stiamo facendo qualcosa per la prima volta. Questo significa anche che non abbiamo alcuna conoscenza storica sull'argomento e dobbiamo quindi fare una pianificazione sulla base di altre informazioni, come le indicazioni provenienti da Internet o l'esperienza in generale.

Ciò di cui stiamo parlando sono le lead information, pertanto si tratta di informazioni che abbiamo bisogno di avere, prima di cominciare il nostro processo.

Mentre iniziamo questo nuovo processo, riceviamo una grande quantità di lag information che controllano il nostro processo. Si tratta di informazioni che raccogliamo e utilizziamo nel corso del processo e possono renderlo più veloce e più facile.

Se invece non è la prima volta che affrontiamo questo problema abbiamo già della conoscenza, abbiamo, in altre parole, le lag information che ci dicono tutte le statistiche e le analisi effettuate sul processo. Sulla base delle lag information siamo in grado di generare nuove lead information, perché sappiamo quali sono i nostri nuovi obiettivi sulla base della situazione attuale. Possiamo imparare dalla nostra conoscenza interna ed ottimizzare il processo.

Abbiamo quindi discusso di due prospettive diverse.

Una prospettiva è la creazione di un processo per la prima volta, che comprende l'identificazione delle lead information e delle lag information, in modo da poter iniziare e gestire il processo dato.

La seconda prospettiva si basa su un processo già stabilito e si possono utilizzare le lag information salvate che descrivono la correlazione tra il processo e il modo in cui esso è stato influenzato. Quindi, sulla base delle informazioni lag, possiamo generare informazioni lead, e questo ciclo di apprendimento è continuo.

Quando si ottimizzano le risorse in un processo di business bisogna comunque prendere in considerazione la soddisfazione degli utenti, che è una regola fondamentale nella gestione delle prestazioni.

2.5.1 Caso di studio: CRM

Utilizziamo un modello per descrivere il rapporto tra gli obiettivi definiti a livello strategico ed i nuovi processi che necessitano informazioni successive e continue.

Sulla base della sua strategia locale, ogni reparto deve individuare i fattori critici di successo, che sono gli elementi del piano che devono avere un esito positivo.

É importante notare che se solo uno dei fattori critici di successo non riesce, l'intera strategia potrebbe fallire.

Sulla base dei fattori critici di successo, la funzione di BA dovrà fornire vari tipi di informazioni, in generale, fornisce lag information per consentire di verificare se il reparto sta arrivando ai risultati richiesti.

A volte alla BA può anche essere chiesto di fornire lead information, per esempio se stiamo attuando una campagna di vendite verso la nostra base di clientela esistente, si può essere interessati al profilo dei clienti e dei loro consumi.

Alla funzione BA è però quasi sempre richiesto di fornire lag information nel

2.5. IL LIVELLO FUNZIONALE DEL MODELLO DELLA BUSINESS ANALYTICS²¹

contesto di grandi progetti strategici, il punto è se le verrà mai richiesto di fornire lead information per creare nuovi progetti.

Ora analizziamo la creazione di nuovi processi di business attraverso quattro livelli fondamentali concentrandoci sulle attività di CRM.

Primo livello: individuare gli obiettivi Prendiamo come esempio un reparto marketing, esso ha due scopi principali:

- attrarre nuovi clienti attraverso campagne in televisione, riviste e altri media;
- mantenere i clienti esistenti.

Al fine di raggiungere buone relazioni con i clienti, molte aziende adottano programmi di dialogo, per informare i clienti di nuovi prezzi, negozi, prodotti etc.

Queste comunicazioni hanno lo scopo di estendere e rafforzare il coinvolgimento del cliente.

Esempio di obiettivi: alla fine dell'anno, la base dei clienti della società deve essere aumentata del 10% e il reddito medio per cliente deve essere aumentato del 10%.

Secondo livello: identificare una strategia operativa A questo livello siamo già a conoscenza degli obiettivi, ciò significa che la strategia deve basarsi sulla creazione di crescita della base di clienti.

I responsabili di CRM devono attuare una segmentazione dei clienti, dividendoli in categorie basate sul valore che essi generano per l'impresa.

Per esempio i clienti si possono dividere in clienti oro, argento e bronzo.

Una strategia possibile può essere quella di aumentare del 10% i clienti d'oro, così facendo si otterrebbe un aumento dei ricavi medi del 5%, il 3% è un aumento della quota di mercato e il restante 2% lo utilizzo per dedicarmi agli altri clienti.

La strategia è quindi divisa in due parti: mantenere i clienti d'oro e avviare delle attività per aumentare la base di clienti.

Terzo livello: identificare i Fattori Critici di Successo Questa fase nel processo verso una strategia di informazione si basa sul fatto che quando si implementa una strategia, si avviano un gran numero di attività e alcune di queste attività sono più critiche di altre.

Un punto importante è ridurre l'abbandono dei nostri clienti, ma sarebbe

uno spreco fare offerte a tutti i clienti per paura che se ne vadano. Bisogna individuare chi vuole andare via e il perchè, in modo da fare offerte mirate. Tutto questo si basa su strategie di CRM che utilizzano ovviamente le analisi della BA.

La stessa questione si ha quando vogliamo aumentare la nostra base di clientela, anche qui è uno spreco fare campagne per attirare nuovi clienti come spedire lettere a tutti etc., bisogna riconoscere i clienti giusti e la giusta offerta per loro al momento giusto.

Quarto livello: identificare le lead e lag information Le informazioni che si cercano in connessione con la strategia di fidelizzazione dei clienti devono rispondere alla domanda: Quali clienti mi stanno lasciando, quando e perché?

Una volta che hai le informazioni è possibile effettuare l'offerta giusta per i clienti giusti al momento giusto.

Tutte queste sono lead information, si tratta di informazioni che sono necessarie per iniziare le nuove attività imprenditoriali.

Ovviamente è importante avere continue informazioni su come le singole campagne stanno andando e se si stanno raggiungendo gli obiettivi, e queste sono lag information.

Fino ad ora abbiamo mostrato come avviare e ottimizzare i processi operativi aziendali, ora ci concentreremo su come utilizzare le informazioni come risorsa strategica.

Questo procedimento andrebbe esplorato per ogni reparto dell'impresa ma continuiamo a concentrarci sulle attività di CRM.

L'attività di gestione delle relazioni con i clienti è uno dei processi più sostenuti dalla BA. Ci sono diverse ragioni per questo, prima di tutto perché possono visibilmente aggiungere valore nel brevissimo termine e perchè molte delle informazioni possedute sono informazioni sui clienti.

In generale ci sono tre aree che interessano il CRM:

- Ottenere clienti importanti.
- Aumentare il valore dei clienti esistenti.
- Mantenere i clienti.

2.6. IL LIVELLO ANALITICO DEL MODELLO DELLA BUSINESS ANALYTICS²³

Tutto questo si può tradurre nell'ottimizzazione del **lifetime value** di un cliente, che serve per valutare il valore del cliente nel tempo.

Il lifetime value corrisponde al valore attuale netto di tutti i contributi, presenti e futuri, di un cliente nella sua vita di relazione con l'azienda.

Un punto importante è che deve essere più basso il costo per acquisire un cliente rispetto al reddito che questo porterà all'organizzazione.

Bisogna poi analizzare attraverso quale canale ottenere i clienti, con quali campagne e in quale segmento.

(Market Basket Analysis) Attraverso specifiche analisi della BA si possono per esempio studiare gli oggetti che vengono generalmente comprati insieme dai consumatori e attuare una disposizione fisica degli oggetti in modo che questi siano vicini per invogliare il cliente a comprare sempre questi oggetti insieme in modo comodo e veloce.

Per quanto riguarda invece lo studio di come mantenere i clienti, attraverso la BA è possibile monitorare i due diversi modi in cui si perdono i clienti:

- alcuni clienti sono soddisfatti dell'impresa e dei prodotti/servizi ma ricevono un'offerta migliore da altri;
- alcuni clienti sono davvero insoddisfatti e per questo vengono persi.

Esiste anche una regola chiamata regola 80/20 o Zipf's law, che dice che l'azienda genera l'80% del suo profitto dal 20% dei suoi clienti.

Una società dunque deve fare molto per mantenere questo 20% ed eseguire strategie di conservazione per questo gruppo di clienti.

Quando si tratta invece di clienti poco redditizi si dovrebbe semplicemente rinunciare a questi clienti, o almeno ridurre al minimo tutti i costi che riguardano questi.

2.6 Il livello analitico del modello della Business Analytics

Fino a questo momento si è parlato dei tipi di informazioni che un'organizzazione richiede a livello strategico e quali requisiti queste informazioni devono avere.

Ora ci concentreremo sui vari metodi analitici che generano le informazioni richieste e forniscono conoscenza.

Cominciamo con la distinzione fondamentale tra tre concetti: dati, informazioni e conoscenze.

I **dati** sono definiti come trasportatori di informazioni e raramente hanno valore per l'utente. Un dato è ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione. In informatica: elementi di informazione costituiti da simboli che debbono essere elaborati.

Le **informazioni** sono dati aggregati ad un livello in cui hanno senso per il supporto decisionale. L'informazione è un notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.

Quando la BA genera **conoscenza** significa che le informazioni sono state analizzate e interpretate.

Ora possiamo specificare tre chiari requisiti dei nostri analisti e il loro centro di competenza:

- competenze di business;
- competenze di metodo (i kit di attrezzi devono essere adeguati);
- competenze sui dati (comprensione tecnica).

Competenze di business L'analista deve comprendere il processo di business che sta sostenendo e in che modo le informazioni fornite e la conoscenza appresa può generare valore aggiunto a livello strategico.

L'analista deve essere in grado di trasmettere, all'area business, la possibilità di utilizzare le informazioni come parametro competitivo e di ottimizzare autonomamente l'informazione e la conoscenza in modo tale che all'utente venga dato il supporto decisionale migliore possibile.

In altre parole, gli analisti devono essere in grado di ampliare la loro funzione, in modo da non eseguire solo una serie di compiti, ma in modo da essere in grado di ottenere il maggior valore possibile dal volume di informazioni e conoscenze che ottengono.

Competenze di metodo L'analista dovrebbe essere in grado di visualizzare le informazioni in modo tale che l'utente riceva una panoramica di tutto il materiale in modo chiaro e immediato, inoltre, deve essere in grado di fornire più delle informazioni scaturite da un modello e partecipare direttamente all'analisi dei dati per assicurare di ottenere solo dati corretti e rilevanti.

Con questo non si chiede all'analista di spiegare matrici di covarianza e fare

2.6. IL LIVELLO ANALITICO DEL MODELLO DELLA BUSINESS ANALYTICS²⁵

calcoli, perchè abbiamo i software per questo, ma il requisito richiestogli è quello di essere in grado di trarre le giuste conclusioni dai test.

Competenze sui dati Il requisito finale che si chiede a un'analista è quello di avere una conoscenza di base su come recuperare ed elaborare i dati. Oltre a questo si richiede anche la competenza di lavorare direttamente con un data warehouse.

Si noti, tuttavia, che tali competenze personali e professionali non hanno bisogno di essere racchiuse in una sola persona, ma solo di essere rappresentate nell'organizzazione e collegate quando richiesto.

Il ruolo dell'analista nella catena del valore complessivo della BA, si estende dal raccogliere i dati nella parte tecnica dell'organizzazione, a fornire informazioni o conoscenze alla parte business dell'organizzazione.

Se un analista non utilizza una corretta metodologia analitica, perde della conoscenza potenziale con perdite di valore, vediamo quindi come evitare questo.

Ci sono alcune domande chiave per assicurare che il dialogo tra l'analista e il destinatario fornisca una panoramica sul quali informazioni sono più rilevanti, e quale metodo utilizzare.

Elenchiamo le tre domande necessarie:

1. Determinare con il titolare del processo, quali competenze di analisi quantitative, o di gestione dei dati e sviluppo di report sono richieste;
2. Determinare se le ipotesi o i dati guidati dalle analisi, rendono migliore il supporto decisionale;
3. Determinare se il metodo ha l'obiettivo di esaminare la correlazione tra una variabile dipendente e un gran numero di altre variabili, oppure se l'obiettivo è identificare diversi tipi di strutture nei dati.

Lo sviluppo di report è ormai necessario per fornire le informazioni in modo appropriato, chiaro e immediatamente comprensibile all'utente.

Esistono diversi tipi di report ognuno con caratteristiche diverse per soddisfare diverse esigenze.

Parliamo di report ad hoc, report aggiornati manualmente, report automatizzati (On Demand) e report automatizzati (Event Driven).

Report ad hoc I report ad hoc sono il tipo di consegna richiesta dal cliente se abbiamo informazioni che ci servono in connessione con, ad esempio, un caso aziendale o un sospetto o una questione critica che deve essere confermata o negata.

Possiamo per esempio avere il sospetto che il segmento del settore pubblico rifiuti alcuni nostri prodotti, e abbiamo quindi bisogno di una relazione su questo particolare problema.

Report aggiornati manualmente I report aggiornati manualmente vengono normalmente utilizzati in connessione con progetti e quindi hanno una durata limitata.

Questo valore a breve termine rende finanziariamente impraticabile mettere questi rapporti in produzione regolare.

In alternativa, questi report potrebbero arrivare in tempo, ma alcuni utenti non hanno accesso ai sistemi di reporting della società o semplicemente non riescono a comprenderli e quindi rimangono comunque inutilizzati.

Questo metodo porta l'analista ad utilizzare tutto il suo tempo per l'aggiornamento dei report standard, che esso crea per gli utenti. Due sono le conseguenze: in primo luogo, non otteniamo un ulteriore sviluppo della conoscenza che l'analista invece potrebbe creare, in secondo luogo, l'intero flusso delle informazioni in azienda si ferma, quando l'analista cambia lavoro.

Report automatizzati: On Demand Questo tipo di report viene in genere fornito in connessione con le implementazioni di data warehouse e si basa sugli utenti che hanno accesso ad una moltitudine di informazioni che vengono aggiornate su base regolare.

Con l'espressione on demand si intende che l'utente legge i report e visualizza le informazioni solo quando lo richiede.

In genere, si ha un'interfaccia utente personalizzata, controllata dal login dell'utente, che assicura che l'utente veda solo le informazioni rilevanti, e che tutte le informazioni personali (ad esempio, lo stipendio e la malattia) non siano accessibili pubblicamente per l'organizzazione.

Report automatizzati: Event Driven Questo tipo di relazione funziona come i report on demand (su richiesta), con la differenza che si ricordano dell'utente quando legge i report.

Questi report sono guidati dagli eventi (event driven) e l'evento che innesca un report può essere qualsiasi cosa, dal trascorrere di un intervallo di tempo al fatto che alcuni valori critici siano stati superati.

Il vantaggio di tale soluzione è che gli utenti vengono avvertiti quando i pro-

2.6. IL LIVELLO ANALITICO DEL MODELLO DELLA BUSINESS ANALYTICS²⁷

blemi si verificano in modo che questi possano essere risolti a breve termine, piuttosto che lasciare agli utenti la decisione di quando visualizzare i report scoprendo così dei problemi in ritardo.

Nelle sezioni precedenti, si è discusso la differenza tra lead e lag information, e si è sottolineato che le informazioni lag in genere sono distribuite attraverso i report. Ciò significa che deve essere un requisito che una strategia di informazione includa una serie di report che, attraverso la misurazione dei processi aziendali critici, sia in grado di fornire il supporto per la strategia di business scelta.

Tra le analisi possibili utilizzando la Business Analytics quella rivoluzionaria abbiamo detto essere la previsione.

La previsione è utilizzata per due cose:

1. creare proiezioni delle tendenze;
2. imparare da correlazioni storiche.

I metodi di previsione sono quindi strumenti estremamente preziosi in termini di ottimizzazione dei processi, dove vogliamo sapere, sulla base dei KPI, come possiamo ottimizzare le nostre prestazioni.

Le previsioni e tutti gli altri tipi di analisi svolti dalla BA fanno uso di tecniche di data mining. Il data mining si occupa di estrazioni non banali di informazioni implicite, precedentemente non note, e potenzialmente utili. Si occupa di esplorare e analizzare, tramite strumenti automatici o semiautomatici, grandi quantità di dati allo scopo di identificare pattern significativi. Spesso si fa data mining su di un data warehouse.

Il processo di data mining può essere suddiviso in tre fasi:

1. creare un certo numero di modelli;
2. selezionare il modello migliore;
3. utilizzare il modello selezionato.

Questi modelli vengono creati da delle attività principali di data mining che è opportuno menzionare.

2.6.1 Data Mining

Le attività di data mining sono suddivise in due tipologie: metodi predittivi (utilizzano alcune variabili per prevedere il valore futuro di altre variabili); metodi descrittivi (consistono nell'identificare pattern comprensibili che descrivano i dati).

Fanno parte dei metodi predittivi la classificazione e la regressione, mentre fanno parte dei metodi descrittivi il clustering, l'identificazione di regole associative, l'identificazione di pattern sequenziali e l'identificazione di anomalie.

Classificazione Data una collezione di record (training set), ogni record contiene un insieme di attributi, alcuni dei quali indicano la classe del record. Lo scopo è di assegnare nuovi record ad una (o più) classi il più accuratamente possibile, identificando un modello per gli attributi di classe, cioè una funzione sui valori degli altri attributi.

Un test set viene utilizzato per determinare l'accuratezza del modello. I test set e i training set sono ottenuti entrambi dai dati disponibili, che vengono suddivisi nei due tipi di insiemi.

Es: ridurre il costo di marketing individuando un insieme ristretto di consumatori che più probabilmente compreranno un nuovo modello di cellulare. Approccio: utilizzare i dati per un prodotto simile introdotto in precedenza, individuare chi ha comprato o meno il prodotto, questo attributo (Sì/No) è l'attributo di classe. Raccogliere informazioni relative a questi clienti e utilizzarle come attributi di input per "addestrare" il classificatore.

Regressione Predire il valore di una variabile continua sulla base dei valori di altre variabili assumendo l'esistenza di un modello di dipendenza (lineare o non lineare).

Esempi:

- Predire le vendite di un nuovo prodotto sulla base delle spese pubblicitarie.
- Predire la velocità del vento in funzione della temperatura, l'umidità, la pressione, e altri parametri.
- Predire l'andamento di indici di borsa.

Clustering Dato un insieme di dati, ognuno descritto da un insieme di attributi, e una misura di similarità, identificare cluster (gruppi) tali che:

2.7. IL LIVELLO DI DATA WAREHOUSE DEL MODELLO DELLA BUSINESS ANALYTICS²⁹

- I dati in un cluster siano più simili tra loro.
- I dati in due cluster diversi siano meno simili.

Es: attuare una segmentazione del mercato raccogliendo le informazioni relative ai clienti e identificare poi cluster di clienti simili.

Misurare i cluster ponendo attenzione ai pattern di acquisto di clienti nello stesso cluster o in cluster differenti.

Identificazione di regole associative Dato un insieme di record, ciascuno contenente alcuni oggetti, produrre regole di dipendenza per predire l'occorrenza di alcuni oggetti sulla base della presenza di altri.

Es: processare i dati raccolti alla cassa tramite gli scanner di codici a barre per identificare i prodotti che vengono acquistati insieme da un numero significativo di clienti.

Identificazione di pattern sequenziali Dato un insieme di oggetti, con ogni oggetto associato al tempo in cui esso si manifesta, identificare regole che predicano dipendenze sequenziali tra eventi differenti.

Es: seguire la navigazione degli utenti e cercare le regole su come questi si muovono attraverso il web.

Identificazione di anomalie Identificare deviazioni significative dal comportamento abituale.

Es: identificazione di frodi con carte di credito.

Le informazioni presentate da questi processi metodici non sono destinate ad essere successivamente interpretate e trasferite in conoscenza, perché questi risultati analitici sono indiscutibili.

2.7 Il livello di data warehouse del modello della Business Analytics

Nel livello precedente del modello abbiamo guardato i processi che trasformano i dati grezzi immagazzinati in informazioni e conoscenza, in questo livello, invece, vedremo come memorizzare i dati e i vantaggi di avere un data warehouse.

I vantaggi di avere un data warehouse sono quelli di dare all'organizzazione

una piattaforma comune di informazioni, che garantisce dati coerenti, integrati e validi in tutte le aree di business.

Ciò è essenziale se una società vuole ottenere un quadro il più completo possibile dei suoi clienti.

Un data warehouse è costituito da una parte tecnica e una parte commerciale.

La parte tecnica deve garantire che i dati dell'organizzazione vengano raccolti dai suoi sistemi di origine, e che vengano poi memorizzati, combinati, strutturati, e purificati indipendentemente dalla piattaforma del sistema di origine.

La parte commerciale di un data warehouse deve assicurare che le figure chiave desiderate e i report possano essere creati.

Ci sono molte buone ragioni per l'integrazione di dati in un data warehouse globale, elenchiamone alcune.

- Evitare isole di informazioni e processi manuali di connessione.
- Evitare il sovraccarico dei sistemi di origine con reportistica giornaliera e analisi.
- Integrare i dati provenienti da diversi sistemi sorgente.
- Aggregare i dati per le esigenze aziendali.
- Aggiungere nuovi termini, regole e dati, non presenti nei sistemi di origine.
- Tenere la documentazione dei metadati al momento della raccolta dei dati.
- Garantire la scalabilità necessaria per garantire la futura gestione di un aumento dei volumi dei dati.
- Garantire la coerenza dei dati validi in tutte le aree di business. (una sola versione della verità).

Siccome tutto si basa sui dati è importante assicurarsi che questi siano di alta qualità.

La qualità dei dati è il risultato della completezza dei dati, è non avere duplicati, è il livello di accuratezza e la coerenza in tutta l'organizzazione complessiva.

I dati provenienti dai sistemi di origine non possono essere utilizzati in modo efficace fino a che non sono stati analizzati e purificati, per questo, per garantire che la scarsa qualità dei dati provenienti da fonti esterne non distrugga o riduca la qualità dei processi interni e le applicazioni, le organizzazioni dovrebbero stabilire un firewall di qualità dei dati nel loro data warehouse.

Nella fase successiva del miglioramento della qualità dei dati, il processo inizia con lo sviluppo di dati migliori. In altre parole, ciò significa correggere gli errori, assicurare la precisione e la convalida e la standardizzazione dei dati al fine di aumentare la loro affidabilità.

2.8 Il livello delle fonti dei dati del modello della Business Analytics

I dati di origine per le analisi di business vengono creati per esempio nel sistema di gestione finanziaria e dei dati di vendita (dati degli ordini), oppure vengono creati attraverso degli ordini sul sito Web della società.

È qui che la qualità dei dati è della massima importanza, perché è qui che i dati vengono creati.

In questo livello risponderemo a molte domande fondamentali del tipo:

- Come può un'azienda raccogliere i dati di origine?
- Come possono essere suddivisi i tipi di dati?
- Come possono essere suddivise le analisi sui dati?
- Come possiamo migliorare la qualità dei dati dei sistemi di origine?

Quando usiamo il termine “sistemi di origine”, il nostro punto di partenza è un data warehouse, dove i sistemi di origine sono le fonti di dati su cui esso si basa.

Quindi i dati in un data warehouse vengono d'altrove e vengono salvati in base alle regole di business e alle esigenze di informazione della società.

Facciamo alcuni esempi di sistemi di origine:

- **Sistemi di fatturazione.** Analizzando i dati provenienti dalle fatture, possiamo effettuare segmentazioni basate sul comportamento o sul valore dei clienti.

- **Sistemi di promemoria.** Questi sistemi inviano solleciti ai clienti che non saldano le loro bollette in tempo.
- **Sistemi di CRM.** Questi sistemi contengono la storia sulle chiamate dei clienti e le conversazioni. Si tratta di informazioni chiave sui clienti che possono fornire input per l'analisi del comportamento.
- **Sistemi di informazione sui prodotti e consumi.** Queste informazioni possono dirci qualcosa sui prodotti e servizi che vengono venduti nel tempo.
- **Sistemi di informazione sui clienti.** Si tratta di informazioni di base sui nostri clienti, sui quali vogliamo raccogliere tutte le informazioni di mercato.
- **Sistemi di informazione commerciale.** Si tratta di informazioni sul numero dei dipendenti, o sulle figure contabili.
- **Sistemi di informazione sulla produzione.** Questo tipo di informazioni possono essere utilizzate per ottimizzare i processi di produzione, controllo scorte, acquisti, e così via.
- **Sistemi di informazione sui KPIs.** Questi sistemi vengono utilizzati per i processi di monitoraggio nel presente, ma possono essere utilizzati successivamente per l'ottimizzazione dei processi, in quanto rivelano le correlazioni tra le attività e le prestazioni finanziarie che ne risultano.
- **Sistemi di data mining.** Questi sistemi generano risultati come segmentazioni o nuovi modelli di vendita etc, che vengono utilizzati per creare conoscenza.

Ora che abbiamo le informazioni la questione è come si usano e soprattutto quali informazioni usare. Possiamo utilizzare due valutazioni: quanto è utile l'informazione?; quanto è accessibile l'informazione?

Nelle grandi organizzazioni spesso vediamo una duplice funzione di business analytics, vale a dire gli analisti di mercato e gli analisti di data warehouse. I due gruppi sono a volte chiamati, rispettivamente, gli analisti esterni e gli analisti interni, in base alle loro fonti di informazione.

Gli analisti esterni in genere utilizzano, per raccogliere dati, questionari e interviste per avere un contatto diretto con i clienti. Questi sono ciò che noi chiamiamo dati primari, che sono i dati raccolti per un determinato scopo.

2.9. VALUTAZIONE E PRIORITÀ DEI PROGETTI DI BUSINESS ANALYTICS 33

Uno dei problemi con questo tipo di analisi è che è costoso inviare questionari a tutta la clientela ogni trimestre, in più potremmo anche infastidire i nostri clienti con un interrogatorio costante.

Gli analisti interni utilizzano i dati presenti nel data warehouse, le loro informazioni provengono dai già citati modelli di data mining.

Quello che è importante è guardare le due fonti di informazione come integrate tra di loro, piuttosto che in competizione.

Quello che conta non è dove otteniamo le nostre informazioni, ma come si applicano.

Come accennato sopra la qualità dei dati è molto importante, le organizzazioni con dati di alta qualità hanno garantita la competitività, l'aumento dell'efficienza, il miglioramento dei servizi al cliente, e la redditività.

Le organizzazioni con scarsa qualità dei dati spendono molto tempo a lavorare con rapporti contraddittori di bilanci, che portano a decisioni sbagliate grazie a dati vecchi, incoerenti ed errati.

Al fine di migliorare la qualità dei dati in modo efficiente, dobbiamo iniziare alla fonte con la convalida, per esempio la qualità può essere migliorata in modo significativo rendendo obbligatorio compilare tutti i campi importanti nei sistemi di origine.

2.9 Valutazione e priorità dei progetti di Business Analytics

La questione di come dare la priorità ai progetti di Business Analytics porta ad altre due domande:

1. in che ordine devo realizzare le iniziative di BA?
2. quali iniziative non dovrebbero essere implementate del tutto?

Quando assegnamo ai progetti una priorità, è importante verificare se il progetto è strategico.

I progetti vengono proposti sia dall'ambiente strategico sia dall'ambiente operativo, e in generale quelli proposti dall'ambiente operativo hanno maggiore priorità rispetto agli altri.

Elenchiamo anche vari modi semplici e intuitivi per valutare i progetti.

- Un progetto con costi bassi e un rendimento elevato sarà, ovviamente, tra i preferiti.
- Un progetto con gli stessi costi, ma un rendimento minore avrà meno priorità rispetto al precedente.
- Un progetto con costi elevati e una bassa redditività deve essere respinto.
- Un progetto con costi elevati e un rendimento elevato nel lungo periodo verrà valutato in base alla sua rilevanza.

Se vogliamo calcolare il valore di un progetto si utilizza il valore attuale netto (VAN) che è spesso calcolato per i progetti a lungo termine.

Il VAN può essere calcolato così:

costi di implementazione + flusso di cassa annuale / rendimento richiesto.

Tuttavia ci sono molti casi in cui è molto difficile produrre stime sui flussi di cassa futuri per calcolare gli investimenti in termini di entrate dai progetti di BA.

In questi casi dobbiamo utilizzare le analisi costi / benefici, che dovrebbero indicare se il progetto è fattibile per l'organizzazione in relazione al suo costo e ai suoi rischi.

La parte descrittiva del rapporto costi / benefici nell'esempio dell'impresa che produce e vende mezzi di trasporto fatto in precedenza potrebbe essere la seguente.

Titolo Conoscere le preferenze dei clienti e adattare la produzione e la promozione a queste.

Stato attuale Lo stato attuale è caratterizzato da una produzione senza conoscenza dei clienti e dei luoghi di vendita. Ogni scelta è dettata da pura congettura e non si basa sulla conoscenza dei fatti.

Le conseguenze dell'assenza di conoscenza Il reparto di produzione dell'impresa non può lavorare in modo mirato ed adattare i processi alle preferenze dei clienti. La conseguenza è un rendimento del capitale ridotto rispetto al potenziale.

2.9. VALUTAZIONE E PRIORITÀ DEI PROGETTI DI BUSINESS ANALYTICS³⁵

Fattori critici di successo Dal momento che è la prima volta che un'iniziativa di BA è implementata nel reparto di produzione, la disponibilità al cambiamento è fondamentale per il successo del progetto. Un altro fattore critico di successo è riuscire a raccogliere i dati desiderati sulle caratteristiche dei clienti e sulle loro preferenze nei vari luoghi di vendita e questi dati devono essere di qualità adeguata.

Destinatari Il reparto di produzione e di vendita dell'impresa è il target dell'iniziativa di BA.

Rischi Il rischio è collegato alla raccolta di dati e alla disponibilità al cambiamento.

Se si è in grado di elaborare una stima ragionevole della creazione di valore di un progetto è preferibile, altrimenti, si dovrà fare una valutazione qualitativa del valore del progetto.

Capitolo 3

Sistemi software per la Business Analytics

Le nuove tecnologie e le nuove tecniche stanno consentendo alle organizzazioni di ogni tipo e dimensione di avere nuove intuizioni per la performance aziendale.

Oggi più che mai, le aziende si sforzano di capire e analizzare le enormi quantità di dati raccolti da diverse tecnologie e situati in diversi “magazzini”.

Sono necessari ormai software che consentano agli utenti finali di ottenere direttamente i dati che permettano di approfondire l’esplorazione e la comprensione della società, dei clienti e delle performance.

Le tecnologie di prossima generazione e gli strumenti consentono una facile integrazione delle diverse fonti di dati, essi eliminano le limitazioni e le performance tradizionali di archiviazione per quanto riguarda la quantità dei dati, il numero di utenti finali, e dei metodi di analisi.

Tutto questo deve essere permesso da software “facili da usare”.

Esistono già numerosi software in grado di praticare la Business Analytics, tra questi i più conosciuti sono:

- SAS
- Pentaho
- Cognos
- e-SoftEasy Business Analytics

Capitolo 4

Business Analytics con Pentaho

Pentaho è un semplice software con un'ottima interfaccia grafica per creare report, effettuare analisi e generare dashboard.

Tutto questo utilizzando i principi della Business Analytics e tutti i vantaggi che questa porta.

Precedentemente è stato fatto un esempio per spiegare e approfondire il ruolo dei KPIs per ottenere obiettivi strategici. Abbiamo parlato di un'impresa che produce e vende mezzi di trasporto (auto, moto etc) e abbiamo visto di questi prodotti un'analisi delle vendite totali di ogni articolo e un'analisi della distribuzione delle vendite nei vari territori.

Tutti i grafici visualizzati sono stati creati dal software Pentaho e sempre con questo software è stata fatta la definizione degli obiettivi tramite i kpis e ne abbiamo visto i risultati grafici tramite i dashboard.

Prima di procedere vediamo un grafico che mostra l'architettura del software Pentaho e come sono composte le sue varie parti.

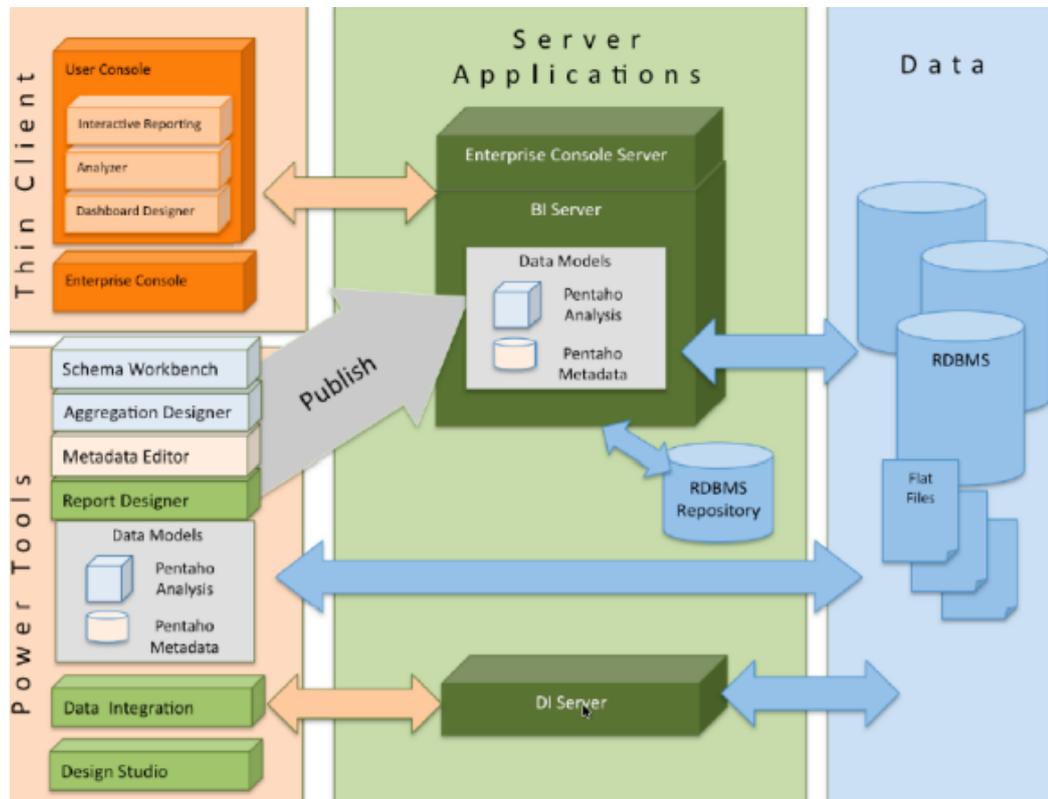


Figura 4.1: Architettura del software Pentaho

Thin Client Pentaho User Console è la principale interfaccia web-based che permette all'utente finale di visualizzare, creare e pianificare i report, report interattivi, e dashboard Analyzer.

Vediamo una breve descrizione delle sue componenti:

- **Interactive Reporting:** fornisce un reporting altamente interattivo e facile da usare, per creare semplici report ad hoc e operativi, senza dipendere da IT o dagli sviluppatori di report. Consente di accedere alle fonti di dati e metadati di Pentaho;
- **Analyzer:** fornisce un reporting analitico intuitivo e interattivo, consentendo agli utenti, non tecnici aziendali, di comprendere rapidamente le informazioni aziendali. Consente di accedere alle fonti di analisi dei dati di Pentaho e di effettuare ordinamento e filtraggio avanzato;
- **Dashboard Designer:** consente agli utenti aziendali di creare ricchi cruscotti interattivi.

Power Tools Gli strumenti di progettazione sono applicazioni desktop che permettono di creare un ricco reporting dell'ambiente.

Vediamo una breve descrizione delle sue componenti:

- **Schema Workbench:** costruisce fonti di dati di analisi (cubi ROLAP) per facilitare l'esplorazione e analisi dei dati per gli utenti business;
- **Aggregation Designer:** è un ambiente grafico usato per aumentare le prestazioni delle query di analisi di uno schema Pentaho, attraverso la creazione di tabelle aggregate;
- **Metadata Editor:** costruisce fonti di dati e metadati, è una rappresentazione del modello dei dati di un database relazionale in cui gli utenti aziendali possono creare query senza dover conoscere il linguaggio SQL;
- **Report Designer:** è un ambiente di progettazione visuale, che rende facile per gli autori di report creare rapidamente sofisticati report graficamente perfetti, per affrontare le esigenze di tipo operativo, di reporting finanziario e della produzione;
- **Data Integration:** è un intuitivo ambiente di progettazione, che fornisce una potente estrazione, trasformazione e caricamento dei dati;
- **Design Studio:** viene utilizzato per creare sequenze d'azioni. Consente la completa personalizzazione e integrazione tra tutti i componenti all'interno di Pentaho.

Server Applications Il server supporta la Pentaho User Console, in primo luogo, soddisfa gli utenti con reporting, analisi, dashboard e comprende servizi per la sicurezza, la programmazione e l'analisi dei dati.

Server Enterprise Console supporta la Enterprise Console e include i servizi per la gestione della sicurezza, programmazione, repository e la configurazione del server.

Data Intergration Server, viene utilizzato per eseguire i lavori di integrazione e trasformazione dei dati e offre servizi come la pianificazione e la gestione dei contenuti (compresa la revisione e integrazione della sicurezza).

4.1 Creare report con Pentaho

Utilizzando Pentaho è possibile creare un semplice report includendo anche grafici che si autogenerano in base ai dati inseriti.

Le caratteristiche principali di questo software per i report sono:

- Consente di accedere direttamente alle fonti di dati.
- Consente un'interazione facile e immediata con i report, progettato appositamente per tutti i tipi di utenti.
- Consente di aggiungere, spostare ed eliminare i campi direttamente all'interno del report.
- Consente la formattazione in riga, il filtraggio, l'ordinamento, il raggruppamento, le aggregazioni e i calcoli sommari.
- Si basa sul principio "Quello che vedi è quello che ottieni".



Figura 4.2: Interfaccia per la creazione di Report

Questa immagine illustrata è l'interfaccia principale per la creazione dei report, mentre le immagini successive illustrano i passaggi da eseguire per impostare le opzioni giuste sui dati per come si vogliono ottenere.

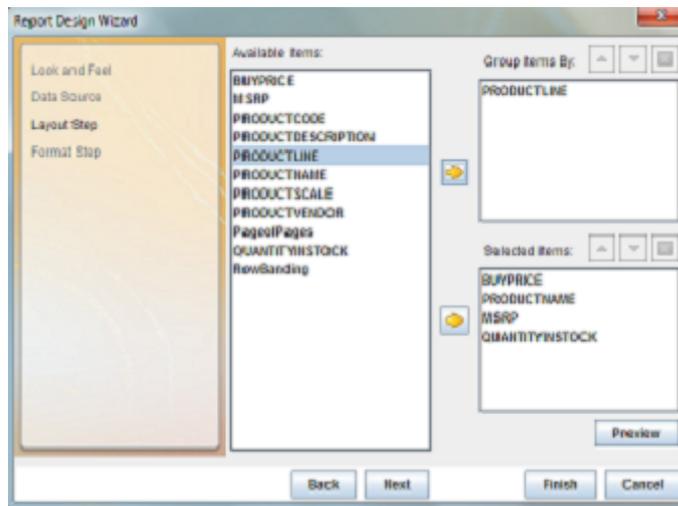


Figura 4.3: Fase di impostazione: scelta dei dati

Inizialmente scegliamo i dati che sono di nostro interesse e il dato che utilizziamo per raggrupparli.

Nel prossimo passaggio per ogni dato selezionato precedentemente dobbiamo indicare il formato adatto e eventualmente il tipo di raggruppamento del dato.

Nelle immagini seguenti vediamo le impostazioni selezionate per ogni dato scelto.

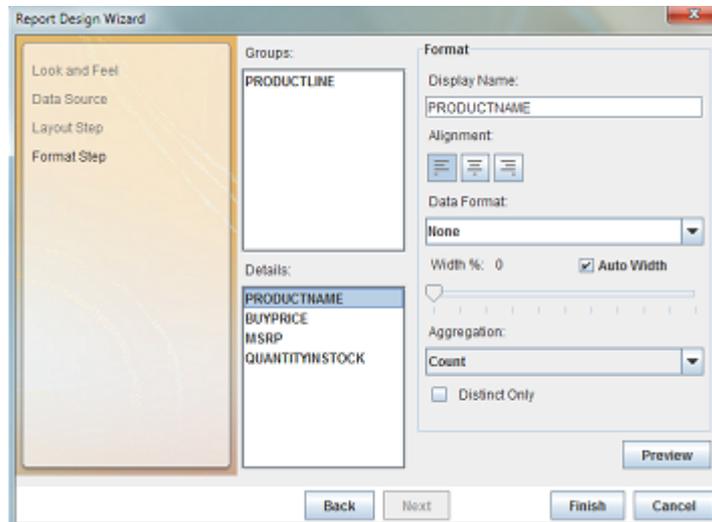


Figura 4.4: Fase di impostazione: scelta del formato del primo dato

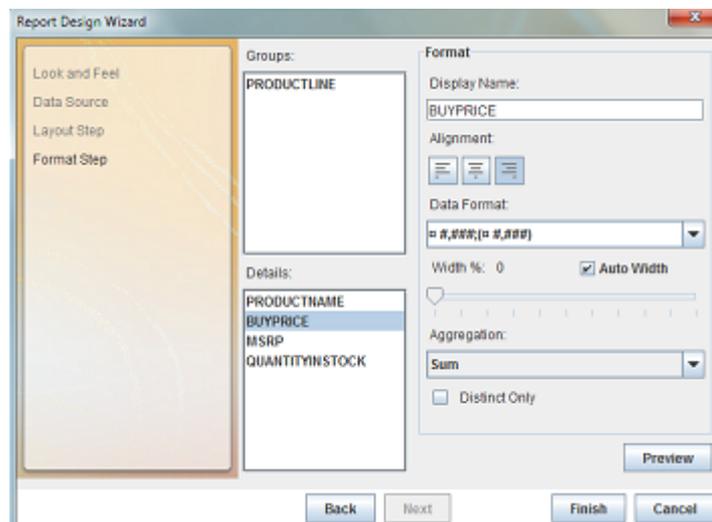


Figura 4.5: Fase di impostazione: scelta del formato del secondo dato

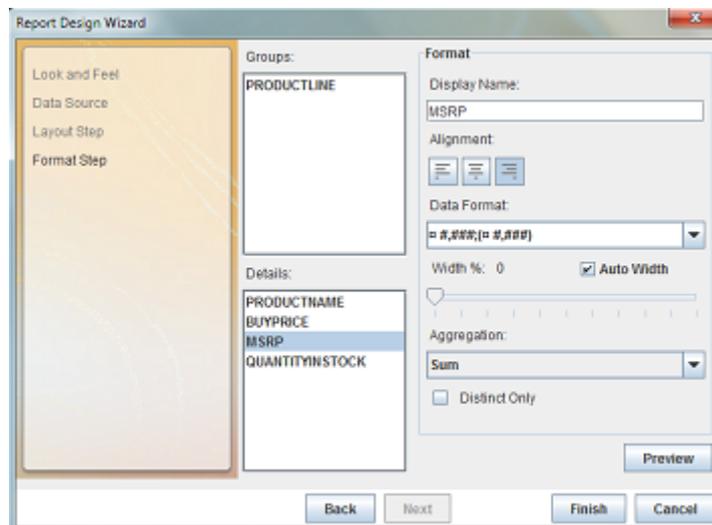


Figura 4.6: Fase di impostazione: scelta del formato del terzo dato

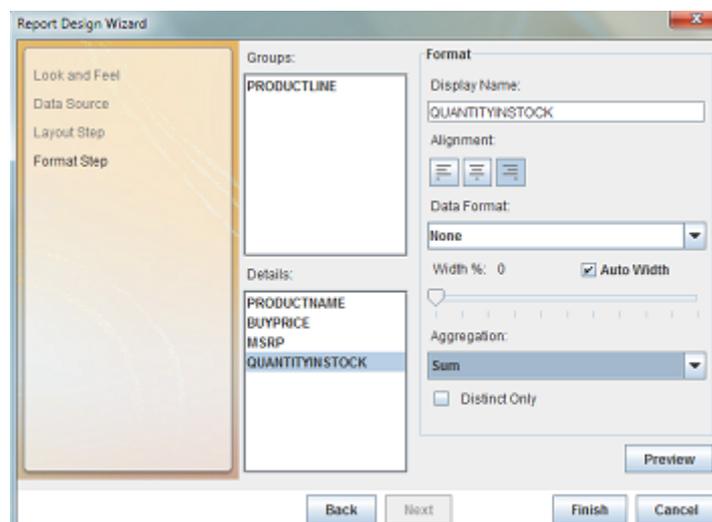


Figura 4.7: Fase di impostazione: scelta del formato del quarto dato

Dopo aver inserito tutte le impostazioni il report viene creato. Illustriamo il report.

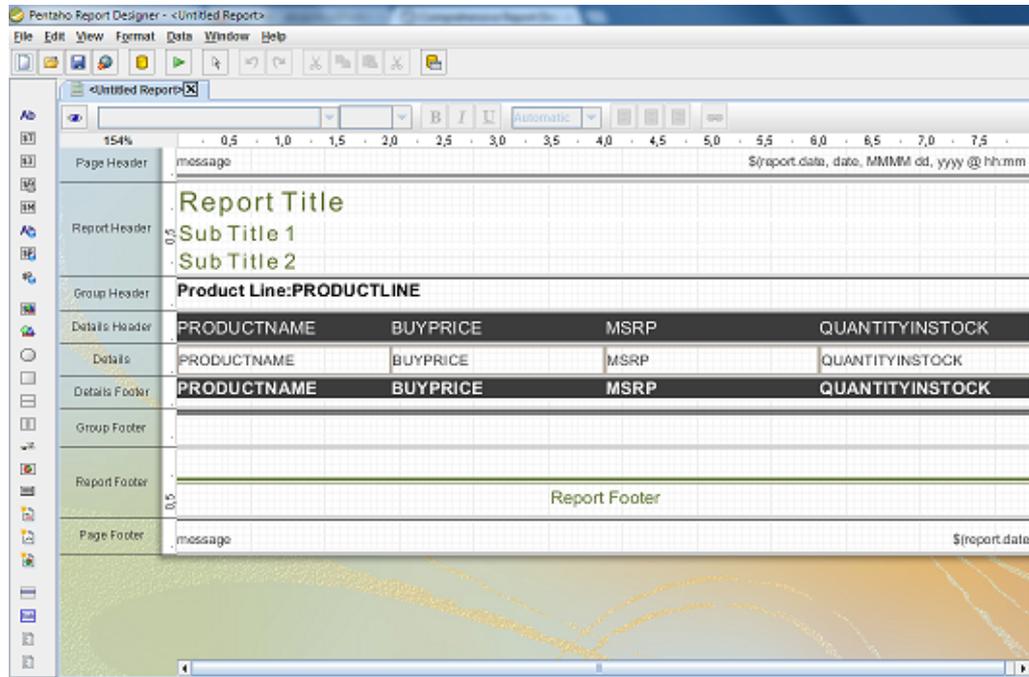
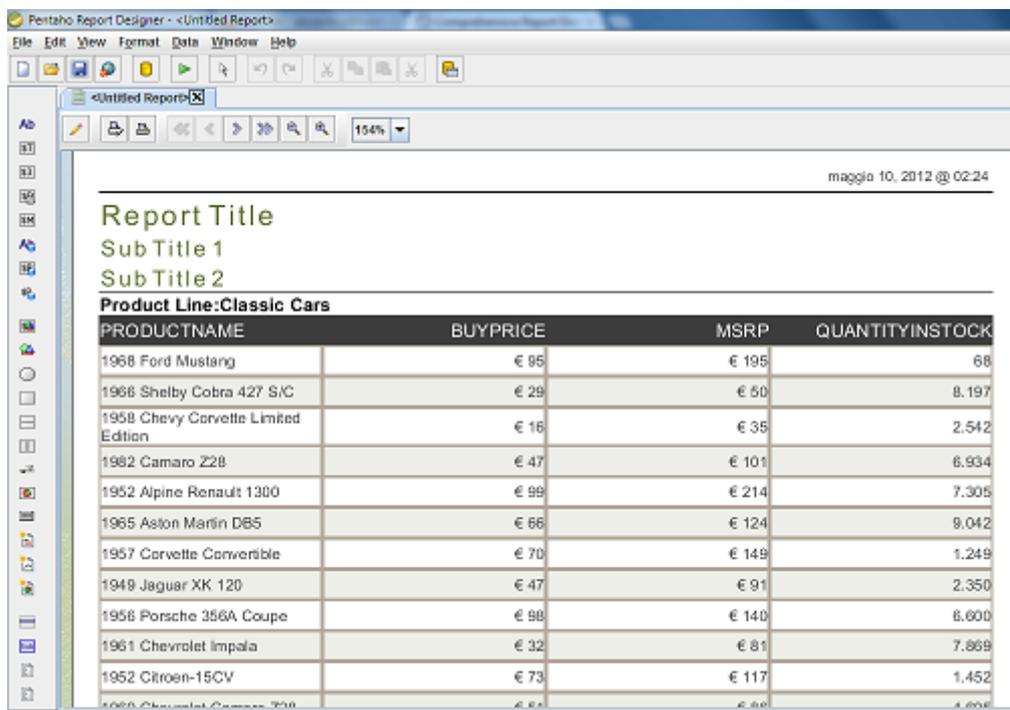


Figura 4.8: Report creato in fase di modifica

Questa immagine rappresenta il report appena creato con i dati scelti. Il report è subito in fase di modifica dove è possibile in ogni momento modificare il report, inserire titoli, immagini e grafici e queste modifiche sono immediatamente visibili.



Report Title
Sub Title 1
Sub Title 2

Product Line: Classic Cars

PRODUCTNAME	BUYPRICE	MSRP	QUANTITYINSTOCK
1968 Ford Mustang	€ 95	€ 195	68
1966 Shelby Cobra 427 S/C	€ 29	€ 50	8.197
1958 Chevy Corvette Limited Edition	€ 16	€ 35	2.542
1982 Camaro Z28	€ 47	€ 101	8.934
1952 Alpine Renault 1300	€ 99	€ 214	7.305
1965 Aston Martin DB5	€ 66	€ 124	9.042
1957 Corvette Convertible	€ 70	€ 149	1.249
1949 Jaguar XK 120	€ 47	€ 91	2.350
1956 Porsche 356A Coupe	€ 98	€ 140	8.600
1961 Chevrolet Impala	€ 32	€ 81	7.869
1952 Citroen-15CV	€ 73	€ 117	1.452
1960 Chevrolet Corvair 700	€ 64	€ 99	4.698

Figura 4.9: Report creato e visualizzato correttamente

In questa immagine invece il report non è in fase di modifica ma è salvato e visualizzato come realmente si visualizza un report.

Ricordiamo che in qualsiasi momento si può tornare alla fase di modifica per aggiungere elementi.

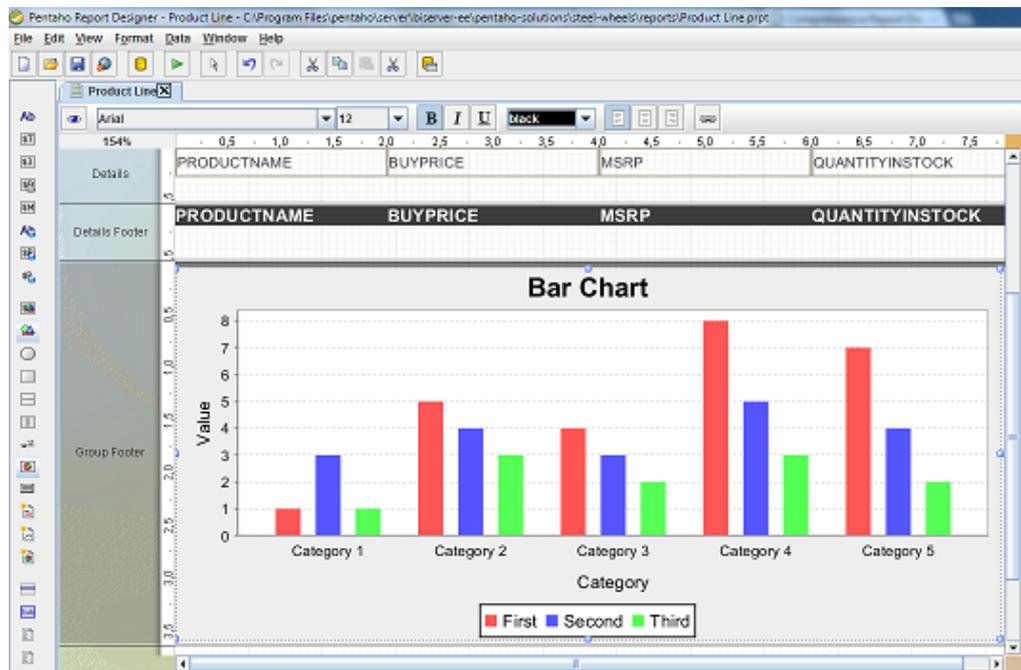


Figura 4.10: Aggiungo un grafico al report in fase modifica

Tornando alla modifica del report ora selezioniamo l'aggiunta di un grafico. Non bisogna fare altro che trascinare l'immagine del grafico nel punto desiderato e deciderne la dimensione adatta, sarà poi durante il salvataggio del report quando ritorna nella fase non modificabile che i grafici verranno creati automaticamente sulla base dei dati passati durante le impostazioni precedenti.

Come si vede da i seguenti due grafici nel report definitivo viene visualizzato il grafico generato automaticamente per ogni linea di prodotto.

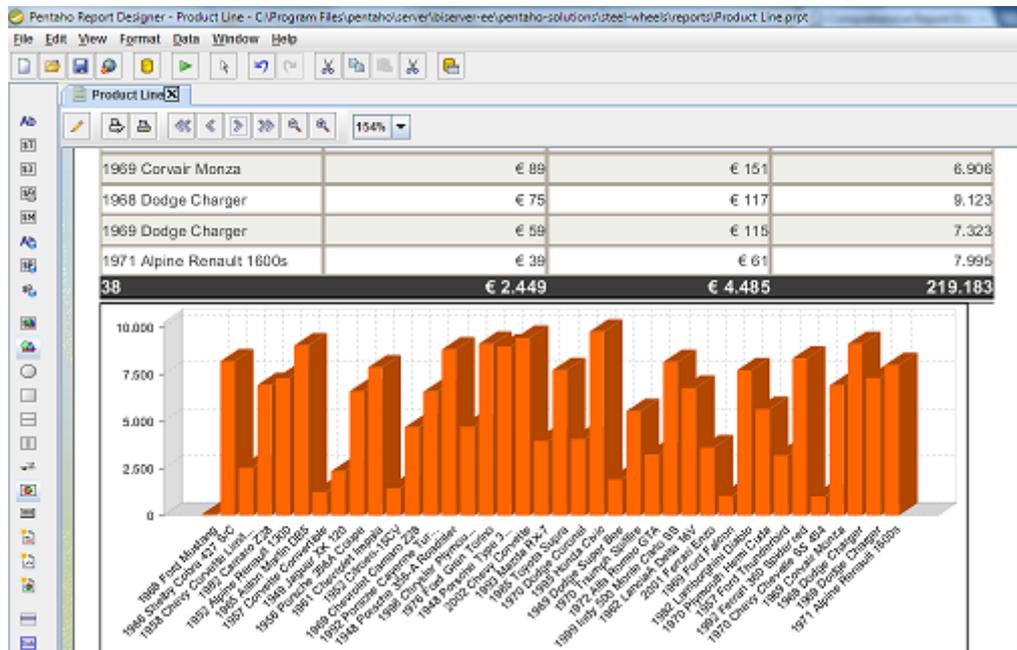


Figura 4.11: Report modificato con grafico

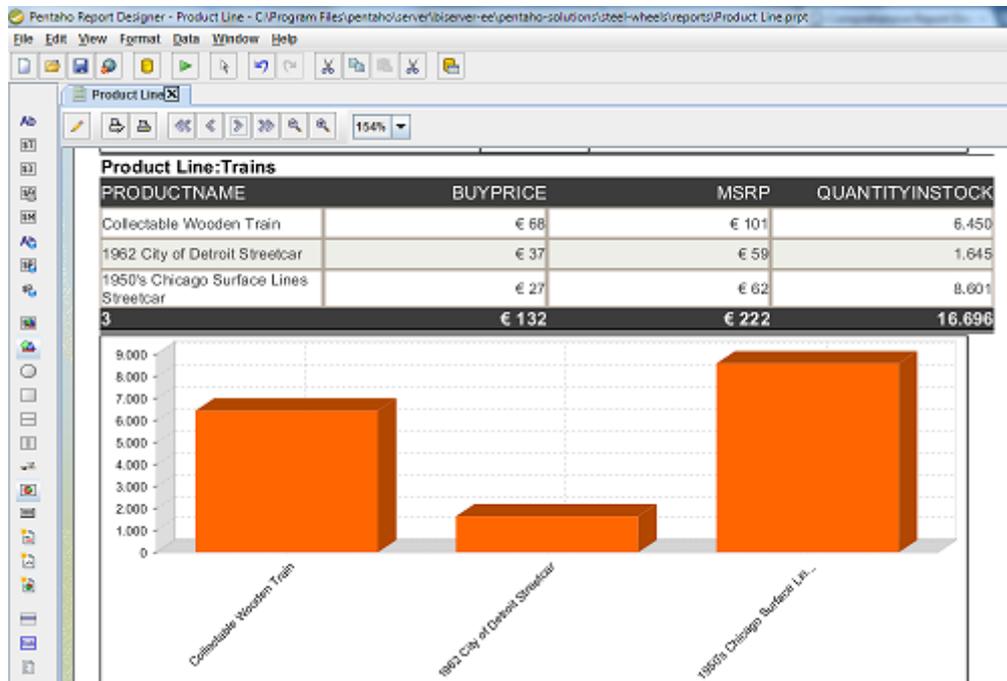


Figura 4.12: Ulteriore pagina del report

4.2 Creare analisi di report con Pentaho

Uno strumento importante di questo software è la creazione di analisi visibili sia in forma tabellare sia attraverso vari tipi di grafici.

Le caratteristiche principali di questo software per i report sono:

- Consente di accedere direttamente alle fonti di dati.
- Consente la creazione di analisi dei dati scelti dall'utente.
- Consente ordinamento e filtraggio avanzato.
- Consente la visualizzazione delle analisi attraverso vari tipi di grafici.

Vediamo come è possibile effettuare analisi sempre utilizzando i medesimi dati.

Mostriamo l'interfaccia di questo strumento.

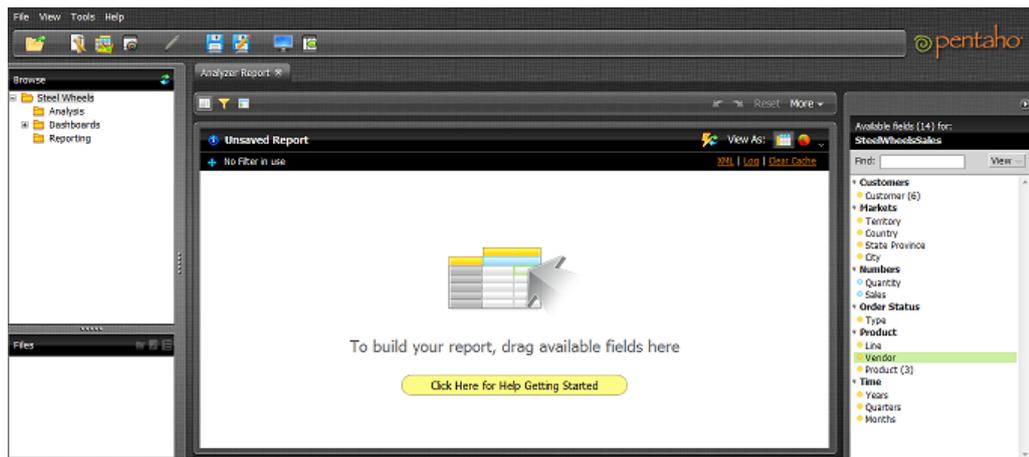


Figura 4.13: Interfaccia dello strumento per analizzare i dati

Ora vediamo meglio le varie parti di questo strumento. Analizziamo prima l'area di destra di questa immagine.

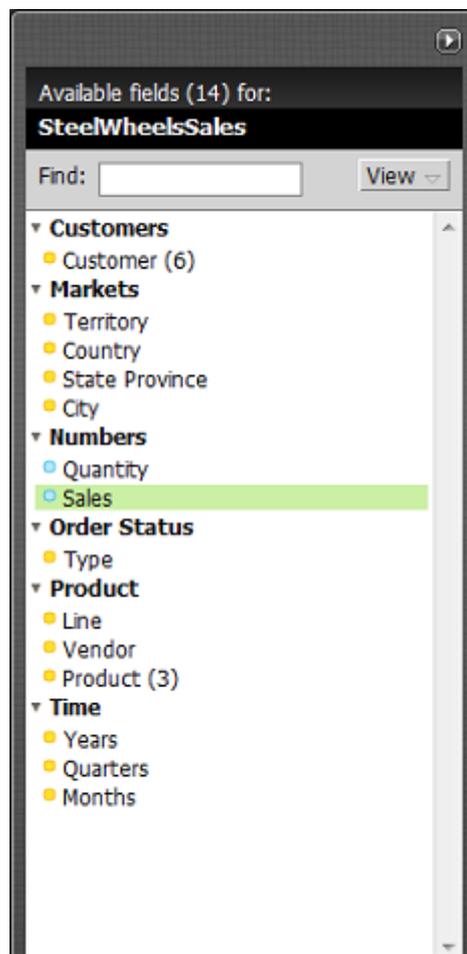


Figura 4.14: Area a destra dell'interfaccia dove vengono elencate le possibili combinazioni di dati

In questa area sono elencati i dati presenti divisi in categorie. Questi dati possono tutti essere utilizzati per creare nuove analisi, basta trascinare in nome del dato che ci interessa nell'area centrale e immediatamente si crea una tabella con questi dati. È possibile trascinare e combinare i dati a piacere per creare tutte le analisi possibili desiderate.

Vediamo ora quattro esempi dove la prima immagine è la tabella che risulta dai dati scelti, mentre la seconda immagine è il grafico relativo che si crea cliccando sul pulsante che raffigura il grafico a torta e poi scegliendo il grafico che si vuole per la visualizzazione.

Facciamo questi quattro esempi con dati diversi e grafici di tipo diverso.

Analyzer Report

Unsaved Report

No Filter in use

XML | Log | Clear Cache

Your report is ready. Rows: 7 Columns: 2

Line	Quantity	Sales
Classic Cars	35.552	4.091.420
Motorcycles	12.708	1.274.125
Planes	11.860	1.076.757
Ships	8.499	748.671
Trains	2.818	234.469
Trucks and Buses	11.001	1.154.281
Vintage Cars	22.893	2.066.226

Figura 4.15: Prima analisi

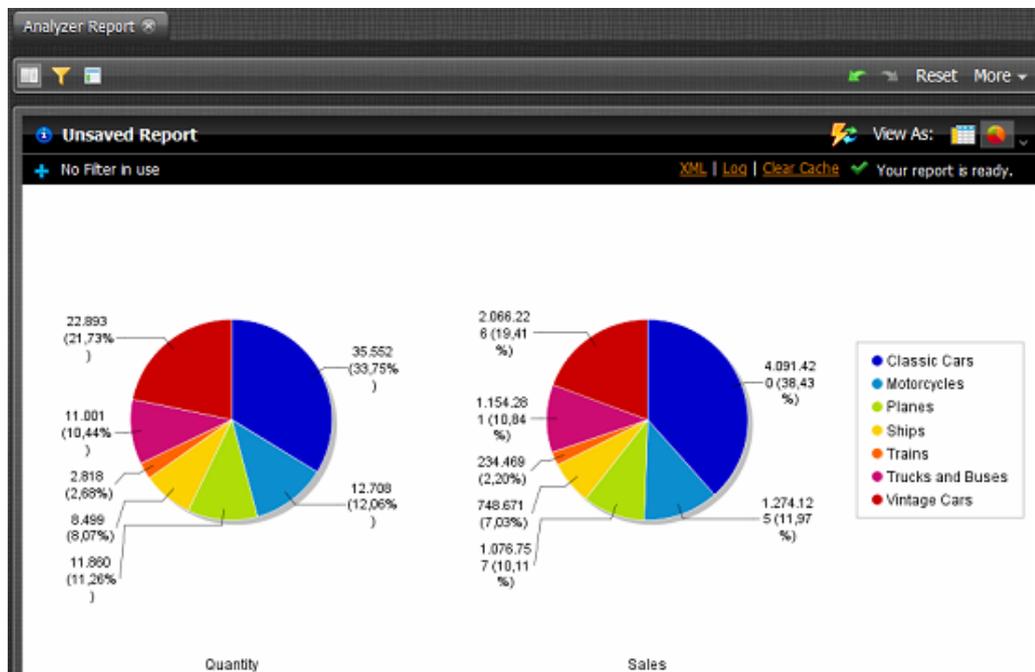


Figura 4.16: Grafico della prima analisi

The screenshot shows the 'Analyzer Report' window with an 'Unsaved Report' tab. The report displays a table with three columns: 'Line', 'Years', and 'Sales'. The data is as follows:

Line	Years	Sales
Classic Cars	2003	1.514.407
	2004	1.838.275
	2005	738.738
Motorcycles	2003	397.220
	2004	590.580
	2005	286.325
Planes	2003	347.755
	2004	528.928
	2005	200.074
Ships	2003	244.821
	2004	375.672
	2005	128.178
Trains	2003	72.802
	2004	124.750
	2005	36.917
Trucks and Buses	2003	420.430
	2004	531.976
	2005	201.875
Vintage Cars	2003	679.949
	2004	
	2005	

Figura 4.17: Seconda analisi

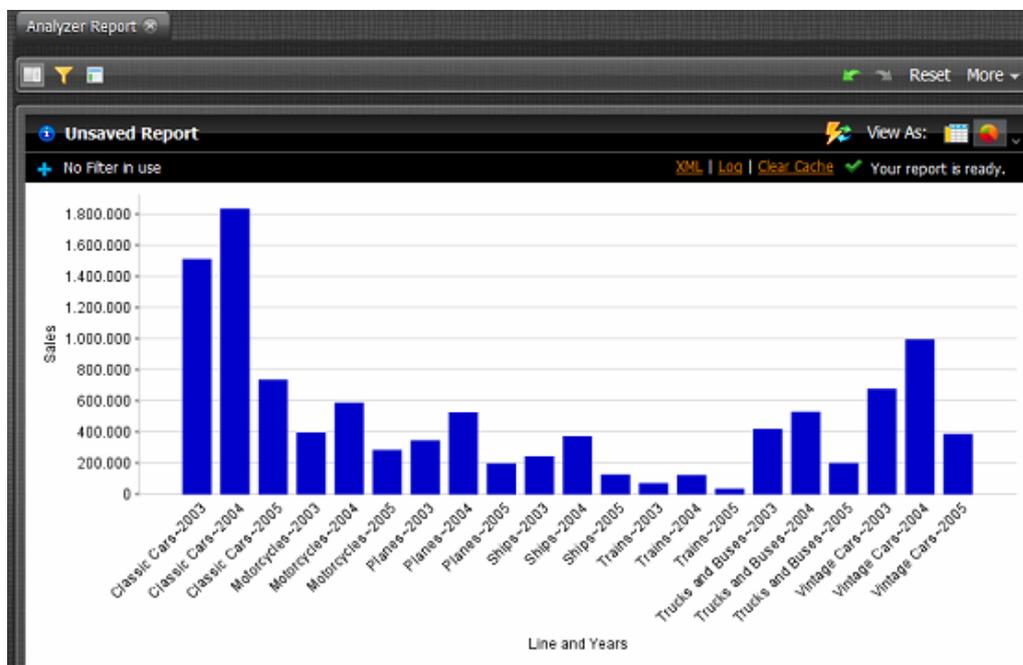


Figura 4.18: Grafico della seconda analisi

Analyzer Report

Unsaved Report

No Filter in use

XML | Log | Clear Cache

Your report is ready. Rows: 33 Columns: 1

Line	Type	Quantity
Classic Cars	Cancelled	659
	Disputed	174
	In Process	564
	On Hold	512
	Resolved	294
	Shipped	33.349
Motorcycles	Cancelled	45
	Disputed	259
	On Hold	49
	Shipped	12.355
Planes	Cancelled	600
	Disputed	63
	On Hold	418
	Resolved	442
	Shipped	10.337
Ships	Cancelled	661
	Disputed	32
	On Hold	271
	Resolved	404

Figura 4.19: Terza analisi

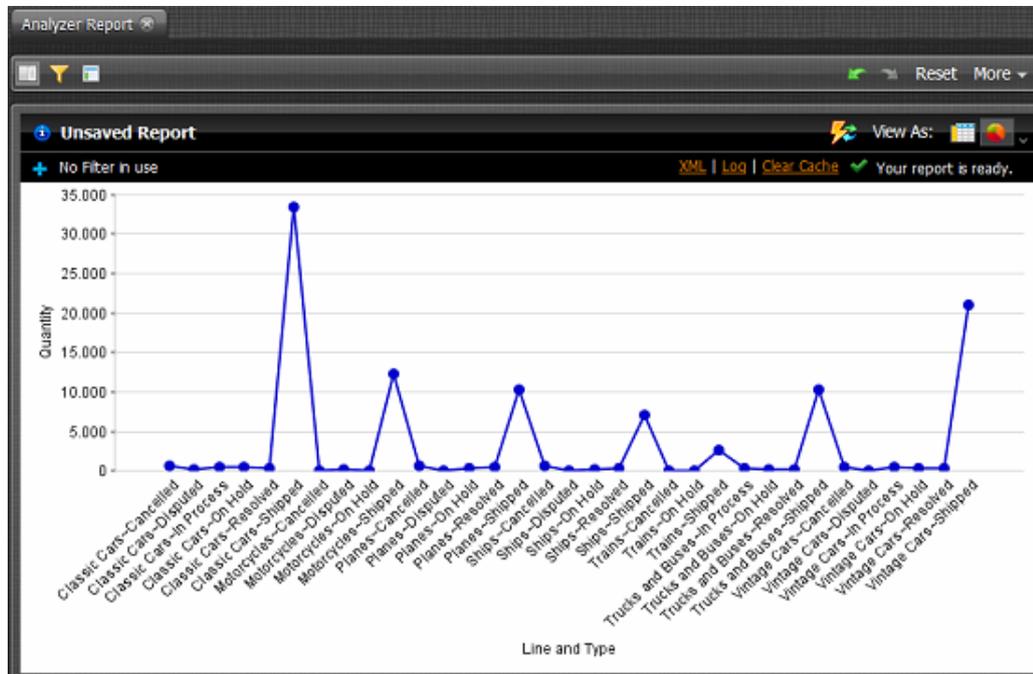


Figura 4.20: Grafico della terza analisi

Type	Years	Months	Quantity
Cancelled	2003	Oct	550
		Nov	288
	2004	May	454
Disputed	2005	Jun	1.342
		Apr	174
In Process	2005	May	423
	2004	Nov	1.490
On Hold	2005	Apr	217
	2005	May	1.156
Resolved	2003	Oct	506
		Nov	288
	2004	Nov	253
	2005	Jan	469
		Mar	650
		Jan	1.357
		Feb	1.449
		Mar	1.755
		Apr	1.993
		May	2.017

Figura 4.21: Quarta analisi

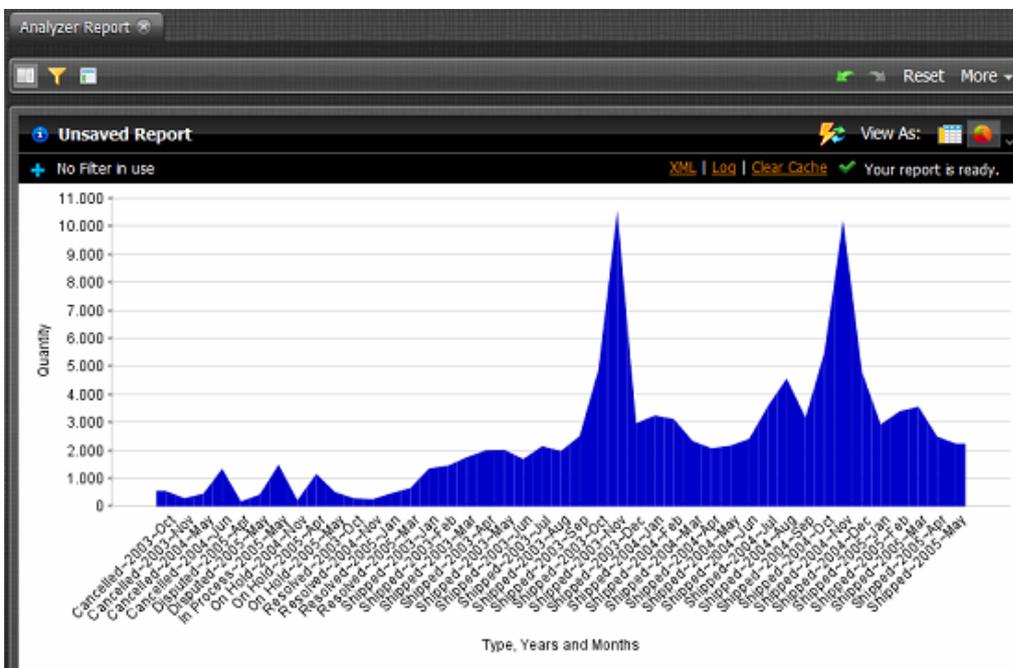


Figura 4.22: Grafico della quarta analisi

4.3 Creare dashboard con Pentaho

Un'altra funzione importante del software è la creazione dei dashboard (di cui abbiamo visto una dimostrazione precedentemente) che sono uno strumento di gestione che consente di misurare determinati valori della vostra azienda in modo rapido ed efficiente.

Invece di basarsi su rapporti cartacei obsoleti per prendere decisioni importanti, un cruscotto fornisce informazioni aggiornate e pertinenti ai dipendenti attraverso l'utilizzo di una combinazione di grafici, scorecard e componenti di report.

Questi cruscotti sono altamente visivi, sono facili da leggere e interpretare, e forniscono ai dipendenti un accesso rapido alle informazioni che aiutano a rispondere a domande critiche in merito alle prestazioni dei loro servizi specifici.

Con soli pochi click del mouse, i dipendenti riescono a sapere se un piano che stanno attuando è efficace, come si paragona alla concorrenza, ed eventuali altri parametri definiti dall'utente. Il vantaggio è ovviamente che possono accedere a questo tipo di informazioni facilmente e da qualsiasi luogo attraverso il web.

Per creare questi cruscotti interattivi basta semplicemente selezionare il layout, il tema e il contenuto che si desidera visualizzare. Questo contenuto può includere report e grafici e fornisce anche la possibilità di aggiungere controlli dinamici di filtro per filtrare l'intero contenuto del loro dashboard usando un semplice elenco di selezione.

4.4 Conclusioni

Pentaho come abbiamo potuto vedere è un ottimo software per applicare analisi di Business Analytics ed è adatto a tutti i tipi di utenti.

Oltre alle funzionalità approfondite, con questo software sono possibili anche altre operazioni interessanti presenti nel seguente elenco.

- Creare nuove connessioni a database.
- Aggiungere dati presi da una vasta tipologia di risorse.
- Aggregare dati utilizzando vari modelli.

- Creare sequenze d'azione che sono un documento XML che definisce un insieme ordinato di definizioni di azioni che insieme provvedono ad una singola attività, ma sono il più piccolo compito completo che la soluzione del motore della piattaforma è in grado di eseguire.
- Incorporare il motore di Pentaho in una nuova o già esistente applicazione Java. Questo consente ai programmatori di recuperare le informazioni da una fonte di dati, le elaborano in base ai parametri specificati, quindi generano output leggibile dall'utente.
- Valutare report e analisi.

In conclusione, la Business Analytics ottiene sempre più attenzioni dalle imprese che ne hanno ormai compreso l'immenso potere e valore e cercano di portare questo a proprio vantaggio.

Di software come abbiamo detto c'è ormai una vasta scelta, Pentaho per quello che è stato esaminato sembra un ottimo strumento facile da utilizzare e comprendere.

Ogni prodotto esistente ha i suoi punti di forza che si adattano bene a certe imprese e meno ad altre, sta ad ogni singolo utente o ad ogni singola impresa studiare le proprie necessità in base alle proprie caratteristiche e scoprire il software più adatto.

Indipendentemente dal prodotto scelto il fatto che rimane rilevante è la potenza della Business Analytics e quello che con essa si può arrivare a fare.

Capitolo 5

Ringraziamenti, bibliografia ed elenco delle figure

Colgo questa occasione per fare un ringraziamento speciale a tutte le persone che hanno lasciato un segno in questi anni.

Ringrazio per prima la persona che mi è stata più vicino in assoluto anche nei momenti per me più difficili, nel bene o nel male, tra una litigata e un'altra. La certezza di avere una persona che non ti abbandona mai e che per quanto a volte odi e a volte ami non smetterai mai di cercare.

Sei tu mamma che rimani sempre al mio fianco dopo qualsiasi cosa.

Questo è il mio regalo per te e papà, vi voglio bene.

Ringrazio di cuore i miei splendidi compagni di università che mi hanno fatto passare tre anni fantastici pieni di allegria e divertimento, tra voi ho conosciuto persone davvero speciali che porterò sempre nel cuore. Le più fantastiche amiche che si possano desiderare Valentina e Sara per fortuna che vi ho conosciute. E come non nominare i ragazzi più cool di tutta informatica per il management: Brizietto, Svay, Giga, Den, Pippo, Lippa e Ubi vi adoro ragazzi! A tutti voi e a chi non ho nominato se no non finisco più grazie davvero vi voglio bene sul serio.

Ringrazio le mie amiche di sempre, quelle che ti conoscono da una vita, quelle che non hanno bisogno di parole per capire di cosa hai bisogno, quelle che sanno ascoltarti e darti consigli nei momenti difficili e che sanno farti ridere come pochi senza bisogno di fare tanto. Mery, Sara, Bea e Eli la nostra amicizia non finirà mai pur non vedendoci spesso adesso, pur avendo mille cose da fare e vite diverse per noi ci siamo sempre state e il tempo o la lontananza questo non lo cambia.

Ringrazio le mie mitiche ballerine che condividono con me una passione da sempre che non si limita solo alla danza ma va ben oltre rendendoci quasi impossibile non vederci per troppo tempo! Volevo dirvi che anche se ci vogliono trasformare in ballerine già pensionate l'amore che noi abbiamo condiviso da una vita intera non ce lo porterà via nessuno... I nostri momenti di gloria o di pianto rimarranno nel cuore per sempre.

Ringrazio anche le ultime persone conosciute ma che in poco tempo sono diventate insostituibili facendomi passare weekend pieni di allegria e divertimento che assolutamente ci volevano dopo tanto studio e tanta fatica! Siete proprio una bella compagnia e sebbene i miei difetti saltino sempre subito alla luce spero di aver mostrato anche la parte migliore di me e se non è così con il tempo verrà fuori...

Infine (non per importanza) ringrazio veramente con tutta l'anima Matte e la sua famiglia e un'altra volta la mia cara Vale per avermi ospitata, aiutata e coccolata in un momento molto difficile, dove lontano dalla mia casa, lontano dalla mia famiglia e piena di paura per colpa del terremoto ho dovuto trovare la forza di continuare a studiare per arrivare a questo giorno! Non troverò mai parole a sufficienza per farvi capire quanto è stato importante e quanto vi sono riconoscente. Grazie, grazie e grazie.

Sono ancora tante le cose che vorrei dire, le persone che vorrei nominare e i ringraziamenti che vorrei fare, perchè tante persone hanno fatto parte della mia vita in questi tre anni, cerco di farlo nel modo più breve ma significativo possibile.

Ringrazio le persone che hanno iniziato questo percorso con me, ringrazio le persone che ho incontrato durante questo percorso, ringrazio chi è arrivato con me alla fine di questo percorso e ringrazio con amore chi con me inizierà quello nuovo.

Bibliografia

- [1] Gert H.N. Laursen, Jesper Thorlund. *Business Analytics for Managers: Taking Business Intelligence Beyond Reporting*. Wiley and SAS Business Series, 2010.
- [2] Giampiero Carli Ballola. *Business analytics: la BI che guarda avanti*. Articolo, 26/05/2010.
- [3] Agostino Lorenzi, Richelmo Giupponi. *Teoria dei sistemi operativi e delle reti. Sistema informativo aziendale*. Atlas 2009.
- [4] Mike Beller, Alan Bennett. *Next Generation Business Analytics Technology Trends*. <http://www.lightshippartners.com/Analytics.html>
- [5] Anna Brown. *Brain trust: Enabling the confident enterprise with business analytics*. <http://www.sas.com/bareport>
- [6] Pat Bates, Mike Biere, Rex Wiederanders, Alan Meyer, Bill Wong . *New Intelligence for a Smarter Planet: Driving Business Innovation with IBM Analytic Solutions*. MC Press Online, 2009.
- [7] Pang-Ning Tan, Michael Steinbach, Vipin Kumar. *Introduction to Data Mining*. Pearson International Edition, Addison-Wesley, 2005.
- [8] Mike Beller. *Next Generation Business Analytics: Technologies and techniques to develop new insights into company performance*. <http://www.lightshippartners.com/Analytics.html>
- [9] *Pentaho Business Analytics*. <http://www.pentahoevalcenter.com>

Elenco delle figure

2.1	Il Modello della Business Analytics	10
2.2	Linee di prodotto - Vendite totali	12
2.3	Linee di prodotto - Vendite in base al territorio	13
2.4	KPI 1 : Vendita media dei prodotti	15
2.5	KPI 2: Cancellazione degli ordini	15
2.6	KPI 3: Articoli per vendita	16
4.1	Architettura del software Pentaho	40
4.2	Interfaccia per la creazione di Report	42
4.3	Fase di impostazione: scelta dei dati	43
4.4	Fase di impostazione: scelta del formato del primo dato	44
4.5	Fase di impostazione: scelta del formato del secondo dato	44
4.6	Fase di impostazione: scelta del formato del terzo dato	45
4.7	Fase di impostazione: scelta del formato del quarto dato	45
4.8	Report creato in fase di modifica	46
4.9	Report creato e visualizzato correttamente	47
4.10	Aggiungo un grafico al report in fase modifica	48
4.11	Report modificato con grafico	49
4.12	Ulteriore pagina del report	49
4.13	Interfaccia dello strumento per analizzare i dati	50
4.14	Area a destra dell'interfaccia dove vengono elencate le possibili combinazioni di dati	51
4.15	Prima analisi	52
4.16	Grafico della prima analisi	52
4.17	Seconda analisi	53
4.18	Grafico della seconda analisi	53
4.19	Terza analisi	54
4.20	Grafico della terza analisi	54
4.21	Quarta analisi	55
4.22	Grafico della quarta analisi	55