

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale

TESI DI LAUREA MAGISTRALE



Business Intelligence: Dashboard e applicazioni

Relatori

Prof. Luca ARDITO

Prof. Riccardo COPPOLA

Candidato

Debora MIGNOGNA

Sessione di Laurea Dicembre 2021

Indice

Elenco delle figure	IV
Abbreviazioni	VII
1 Introduzione	1
1.1 Il problema	1
1.2 Descrizione del progetto	2
1.3 Struttura della tesi	3
2 Trasformazione digitale	4
2.1 Storia della trasformazione digitale	4
2.1.1 I driver della trasformazione digitale	7
2.2 Digitalizzazione delle aziende italiane	8
2.2.1 Misurare la trasformazione digitale	8
2.2.2 I dati	9
2.2.3 Il livello di digital maturity nelle imprese italiane	12
2.2.4 Uno sguardo al futuro	13
3 Business Intelligence	14
3.1 Definizione	14
3.2 Il valore dei dati	15
3.3 Data Science	16
3.3.1 I possibili obiettivi dell'analisi dei dati	17
3.4 Knowledge Discovery from Data	18
3.5 Struttura ed architettura	19
3.5.1 I Data Warehouse	20
3.5.2 Progettare un data warehouse	22
3.5.3 L'analisi dei dati	25
4 Data visualization	26
4.1 Business Dashboard in ambito aziendale	30

4.2	Power BI	35
4.2.1	Importazione e trasformazione dei dati	37
4.2.2	Creazione, pubblicazione e condivisione delle dashboard . . .	40
5	Caso studio	43
5.1	Descrizione	43
5.1.1	Analisi AS IS	44
5.1.2	Analisi TO BE	45
5.2	Dati	47
6	Sviluppo del progetto	58
6.1	Creazione report	58
6.2	Dashboard	72
7	Conclusioni	78
	Bibliografia	81

Elenco delle figure

2.1	Le quattro rivoluzioni industriali	5
2.2	Tecnologie adottate dalle imprese italiane con almeno 10 addetti nel triennio 2016-2018	10
2.3	Tecnologie adottate dalle imprese italiane con almeno 10 addetti nel triennio 2016-2018	12
3.1	Processo di estrazione della conoscenza	19
3.2	Architetture per data warehouse [10]	22
3.3	Progettazione di un DW [10]	24
4.1	Grafici di confronto: Simple bar, Grouped bar, Floating bar, Bubble, Radar, Wordcloud [12]	27
4.2	Grafici di tendenza: Line e Area [12]	28
4.3	Grafici di proporzione: Donut, Pie, Stacked bar, Stacked area, Meter, Gauge [12]	28
4.4	Grafici di correlazione: Scatterplot, Heatmap, Parallel coordinates [12]	29
4.5	Grafici di relazione e connessione: Tree diagram [12]	29
4.6	Mappe: Chloropleth ap, Proportional symbol, Connecting lines [12]	29
4.7	Esempio di Operational dashboard	31
4.8	Esempio di Strategic dashboard	32
4.9	Esempio di Tactical dashboard	33
4.10	Gli elementi di Power BI [14]	35
4.11	Finestra di dialogo Power BI	36
4.12	Importazione dati Power BI	37
4.13	Editor di Power Query	38
4.14	Modello Power BI	39
4.15	Creazione relazioni	40
4.16	Data view	41
4.17	Dashboard	41
5.1	Microsoft Access	43

5.2	Microsoft Access: maschera ordini	45
5.3	Class Diagram	48
5.4	Tabella Ordine Cliente	49
5.5	Tabella RigA Ordine Cliente	50
5.6	Tabella Giacenza Magazzino	52
5.7	Tabella Ordine di Acquisto	53
5.8	Tabella Riga Ordine di Acquisto	54
5.9	Dimensional Fact Model: Ordine	56
5.10	Dimensional Fact Model: Acquisto	56
6.1	Esempio Tipi di Dato	59
6.2	Modello Relazionale: Ordine Cliente	59
6.3	Totale ordinato dell'anno in corso	60
6.4	Dettaglio mensile dell'ordinato dell'anno in corso	60
6.5	Filtro dei dati	61
6.6	Analisi fatturato dell'anno in corso	61
6.7	Totale ordinato YtD e Totale Rolling 2021	62
6.8	Spedito ed ordinato dell'anno in corso	62
6.9	Fatturato, ordinato e spedito del mese in corso	63
6.10	Dettaglio ordini del mese in corso	64
6.11	Backlog dell'anno corrente	65
6.12	Backlog degli anni successivi a quello in corso	65
6.13	Dettaglio ordini cliente anno in corso	66
6.14	Dettaglio ordini cliente anno in corso	67
6.15	Dettaglio clienti	67
6.16	Dettaglio clienti e relativi giorni di ritardo nella consegna degli ordini	68
6.17	Giacenza di magazzino alla data attuale	68
6.18	Delta giacenza di magazzino rispetto al mese precedente	69
6.19	Analisi ordini d'acquisto in ritardo	69
6.20	Analisi ordini d'acquisto futuri	70
6.21	Analisi TOP 10 fornitori	70
6.22	Analisi nuovi clienti	71
6.23	Analisi nuovi fornitori	71
6.24	Recupero dati Power BI Service	72
6.25	Dashboard 1	73
6.26	Dashboard 2	74
6.27	Dashboard 3	75
6.28	Dashboard 4	76
6.29	Recupero dati Power BI Service	76

Abbreviazioni

AI

artificial intelligence

IoT

Internet of Things

TD

Trasformazione digitale

AR

Realtà aumentata

VR

Realtà virtuale

BI

Business Intelligence

DB

Data Base

KDD

Knowledge Discovery from Data

SQL

Structured Query Language

DW

Data Warehouse

ETL

Extract, Transform, Load

OLAP

On-Line Analytical Processing

DAX

Data Analysis eXpressions

Capitolo 1

Introduzione

1.1 Il problema

Per avere successo nel mondo aziendale, le organizzazioni devono porre attenzione ai propri dati ed estrarne la giusta conoscenza, perchè questo sia possibile è necessario avere una piattaforma di Business Intelligence che permetta a tutti di poter osservare e comprendere i dati.

Rendendone più facile l'accesso e l'analisi, tali piattaforme, rappresentano un supporto alla strategia di BI dell'organizzazione.

Le necessità sono quelle di monitorare i cambiamenti nel tempo, fare previsione, per esempio, dei profitti e delle vendite, analizzare i risultati di progetti o campagne promosse dall'azienda, visualizzare i dati.

Per questi motivi, implementare una piattaforma di Business Intelligence è fondamentale.

Una delle funzioni più popolari della BI è la capacità di creare e personalizzare le dashboard. Si ha la possibilità di creare moduli, diagrammi e grafici che possono essere aggiornati in tempo reale.

È possibile pubblicare le dashboard su piattaforme interne distinte, per questo è fondamentale scegliere una piattaforma che offra funzionalità self-service che garantisca la possibilità di accedere ai dati, creare le proprie analisi e coadiuvare l'organizzazione nella creazione di una cultura dell'analisi.

L'analisi visiva consente di esplorare intuitivamente i dati, offrendo informazioni ovunque ci si trovi e permettendo di condividere i risultati facilmente.

La giusta piattaforma di analisi offre connessioni ai dati ottimizzate, in questo modo, è possibile raggiungere i dati più velocemente.

Il software deve consentire l'accesso e l'analisi live, deve avere la possibilità di eseguire rapidamente una query, con poco o nessuno sforzo per la codifica.

La piattaforma di BI deve permettere anche di scegliere e di distribuire le analisi

nel cloud, sul server locale o off-site in hosting.

Il sistema deve integrarsi perfettamente con la strategia dei dati esistente, senza rivoluzionare l'infrastruttura dei dati attuale.

Deve consentire di interagire con i dati e di usufruire delle informazioni, a prescindere dal livello di competenza, deve promuovere una cultura dei dati, ponendoli al centro delle decisioni aziendali.

Un giusto progetto di Business Intelligence dà all'organizzazione la capacità di dotarsi di un approccio all'analisi incentrato sull'azienda e abilitato dall'IT [1].

1.2 Descrizione del progetto

Il seguente elaborato nasce sulla base del tirocinio svolto presso Goma Elettronica SpA, azienda italiana leader nella produzione e nella distribuzione di sistemi embedded integrati per la difesa, l'industria e la ricerca.

Nata come azienda di commercio di apparecchiature elettroniche, divenne in breve tempo il distributore italiano di aziende leader in ambito internazionale, concentrandosi nel corso degli anni verso produttori attivi nel neonato mercato embedded e allo stesso tempo acquisendo know-how ed esperienza nell'ambito della mecatronica. L'ambito aziendale in cui si è svolto il tirocinio è quello dell'information technology, in particolare, nel settore IT dedicato alla gestione dei dati aziendali e del sistema informativo; è stato ricoperto il ruolo di Junior ERP Engineer.

Il tema centrale dell'elaborato è la Business Intelligence, in particolare la Data Visualization.

La data visualization è definita da Gartner come la rappresentazione grafica delle informazioni, statistiche o dei fenomeni che riguardano le aziende. Essa consente l'esplorazione visuale ed interattiva dei dati, di ogni dimensione, natura o origine attraverso l'uso dei grafici, mappe e qualsiasi altro strumento di visualizzazione.

In questo modo, è possibile raggiungere un livello più dettagliato di analisi, che spesso è accompagnato anche dalla scoperta di insights dall'alto potenziale strategico [2].

Il progetto verte sulla revisione di alcuni dei report aziendali già esistenti sulla piattaforma Microsoft Access, tali report sono stati studiati e progettati sul Software Power BI con una veste grafica diversa.

Per prima cosa è stata condotta una profonda analisi dei report per capire cosa questi volessero comunicare e come questi potessero essere di supporto al business aziendale. Il passo successivo è stato estrarre i dati originali del DB aziendale, i dati sono stati analizzati, modificati ed infine presentati in Dashboard diverse a seconda del settore aziendale a cui fanno riferimento.

Il progetto di BI pone le fondamenta su tre diversi ambiti aziendali relativi all'analisi degli Ordini Cliente, del Magazzino e degli Ordini d'Acquisto.

Il software utilizzato per la creazione di Dashboard è stato Power BI una piattaforma Microsoft unificata e scalata per la Business Intelligence che riesce a connettere i dati e visualizzarli per ottenere informazioni più dettagliate e chiare.

1.3 Struttura della tesi

La tesi è composta da sei capitoli, il primo introduce l'elaborato descrivendo il problema affrontato ed il progetto intrapreso.

Il secondo capitolo tratta il tema della trasformazione digitale mettendone in evidenza l'importanza: si parte dalla storia della TD in relazione al rapido sviluppo della tecnologia degli ultimi decenni, passando per la sua evoluzione fino all'analisi della situazione delle aziende italiane le quali cercano di riorganizzare procedure, strutture e strategie per seguire il progresso.

Il terzo capitolo esamina uno dei processi della trasformazione digitale: la Business Intelligence. In particolare ci si concentra sull'importanza dei dati, ancora sottovalutata dalle aziende; si descrive la disciplina della data science, il processo di estrazione della conoscenza (Knowledge Discovery from Data) ed infine si analizzano la struttura e l'architettura dei progetti di BI.

Successivamente si pone l'accento sulla Data Visualization, focus del progetto sviluppato, dopo aver introdotto il tema ed illustrato l'importanza degli strumenti di visualizzazione, si passa alla spiegazione del tool scelto: Power BI.

Il quinto capitolo descrive il caso studio analizzato, delineando la situazione dalla quale si parte e i desiderata finali, si pone l'accento sui dati studiati e sul passaggio dal sistema relazionale a quello decisionale.

Il capitolo che segue illustra minuziosamente lo sviluppo del progetto presentando i dettagli dei report e delle dashboard create fino alla pubblicazione e condivisione del lavoro svolto.

L'elaborato di tesi termina con un capitolo che riporta le conclusioni del lavoro ponendo l'accento sul valore aggiunto che esso apporta all'organizzazione aziendale.

Capitolo 2

Trasformazione digitale

2.1 Storia della trasformazione digitale

Negli ultimi anni la trasformazione digitale ha subito una veloce accelerazione. L'uso delle tecnologie digitali ha avuto un impatto rilevante in molti campi della nostra vita, ogni attività umana è stata interessata dai cambiamenti.

E' corretto parlare di una nuova tecnologia o si è solo di fronte al normale sviluppo tecnologico? Capire e confrontare l'importanza degli sviluppi tecnologici moderni rispetto a quelli del passato è un'ardua impresa.

Quella a cui si sta assistendo ora viene definita "la quarta rivoluzione" o Industria 4.0, cominciata intorno al 2012 ed attualmente in corso.

Con questa espressione si intende la crescente fusione tra mondo fisico, digitale e biologico.

La quarta rivoluzione sta cambiando la società come mai prima d'ora ma, tale trasformazione, si basa sulle fondamenta gettate dalle prime tre rivoluzioni industriali (Figura 2.1).

La prima nel 1784 iniziata con la nascita della macchina a vapore e la meccanizzazione della produzione ha portato ad un forte cambiamento sociale spinto dall'urbanizzazione.

La seconda rivoluzione nel 1870 con il via alla produzione di massa attraverso l'uso dell'elettricità, l'avvento del motore a scoppio e l'aumento dell'utilizzo del petrolio come nuova fonte energetica.

La terza nel 1970 con l'avvento dell'informatica, la nascita dei computer e della tecnologia digitale, da qui l'avvio dell'automazione della produzione e lo sconvolgimento dei vecchi settori e la creazione di nuovi.

La quarta rivoluzione è il processo che porterà alla produzione industriale completamente automatizzata e interconnessa [3], è una somma dei progressi in intelligenza artificiale (IA), realtà virtuale, robotica, Internet delle Cose (IoT), stampa 3D,

ingegneria genetica, computer quantistici e altre tecnologie.

È la forza collettiva che sta dietro molti prodotti e servizi che stanno rapidamente diventando indispensabili per la vita moderna.

La quarta rivoluzione industriale sta spianando la strada ai cambiamenti che avvengono nel modo in cui viviamo e sta sconvolgendo radicalmente quasi tutti i settori aziendali.

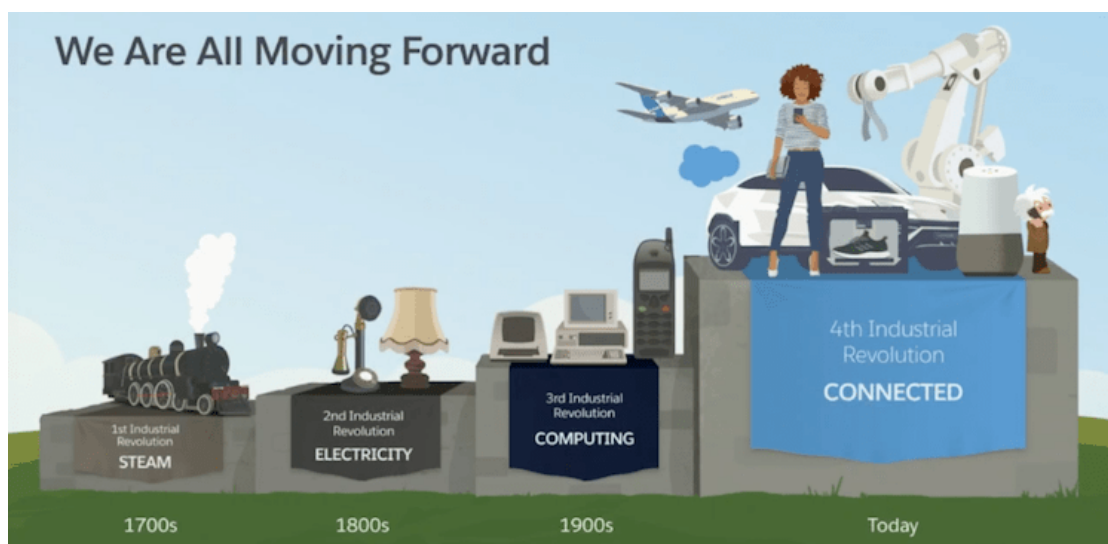


Figura 2.1: Le quattro rivoluzioni industriali

La digitalizzazione ha profondamente trasformato le imprese e la società.

Il maggiore impatto che ha avuto internet sul modo di fare business è sicuramente la possibilità di collegarsi con qualsiasi persona in qualsiasi luogo si trovi, ciò ha dato la possibilità alle imprese di crescere ed espandersi e quindi la possibilità di muoversi a livello internazionale.

Si possono identificare alcuni punti di svolta centrali per le aziende: il primo riguarda la comunicazione radicalmente modificata dall'utilizzo della posta elettronica, la quale ha garantito un veloce scambio di informazioni sia interno all'azienda che esterno.

Il secondo punto di svolta si è raggiunto con l'utilizzo dei siti web aziendali, questi consentono di creare un forte marchio e danno la possibilità ai consumatori di potersi informare sulla società e sui prodotti, forniscono anche ai clienti un mezzo per poter comunicare con l'azienda.

Altro cambiamento importante è stato quello dell' e-commerce, che ha cambiato il modo di fare acquisti degli utenti, le aziende hanno modificato il processo di vendita e ciò ha sicuramente portato loro dei vantaggi, avere un e-commerce ben strutturato e funzionante è stato fondamentale per poter affrontare il recente periodo storico

caratterizzato dalla pandemia.

Il rapido sviluppo ha consentito alle aziende di incrementare l'efficacia e l'efficienza del business ma anche l'evoluzione di nuove opportunità di crescita e sviluppo.

Dal punto di vista aziendale, la trasformazione digitale è il processo di integrazione delle tecnologie digitali in tutti gli aspetti del business, un processo che comporta cambiamenti tecnologici, culturali, operazionali e a livello di generazione di valore. La trasformazione può essere adottata in tutte le fasi aziendali, dalla comunicazione interna, alla produzione, alla vendita e alla distribuzione.

Tale fenomeno globale è caratterizzato da diversi aspetti: connessione, elaborazione dei dati ed automazione.

Calcolatori ed infrastrutture di rete consentono di connettere, attraverso il sistema digitale, tutti i processi aziendali, collegare in tempo reale i diversi ambiti dell'attività economica.

La connessione tra prodotti, processi, macchine e persone porta alla generazione di dati, i quali vengono definiti "il petrolio dell'era digitale".

La tecnologia tradizionale si fonde con quella artificiale per creare sistemi e macchine in grado di lavorare autonomamente aumentando l'efficienza, grazie ad una maggiore produttività a costi ridotti.

Il modo di fare impresa degli ultimi anni è stato trasformato dall' IoT, dai cloud, dall'evoluzione dell'analytics che rappresentano solo l'inizio di una rivoluzione tecnologica quanto culturale all'interno delle organizzazioni.

La trasformazione digitale non si limita solo ad un particolare tipo di business ma coinvolge aziende di qualsiasi dimensione, operanti nei più diversi settori; digitalizzarsi è quindi fondamentale per sopravvivere nel mercato, chi rimane indietro dovrà pagare un caro prezzo.

Tuttavia, non basta adottare soluzioni e tecnologie innovative, è fondamentale ridisegnare la struttura organizzativa fin dalle sue basi, concependo un'organizzazione nuova che comprende ogni aspetto, dal business alla cultura aziendale.

È fondamentale quindi, capire che la cultura è formata da ciò che leader e dipendenti dimostrano di valorizzare e si concretizza soprattutto per ciò che fanno, non solo per ciò che dicono.

Pertanto, una cultura organizzativa allineata ha la missione di plasmare i comportamenti, formare un'identità, generare impegno collettivo e, infine, fornire stabilità per un'azienda.

In altre parole, la cultura è il fondamento che determina un'identità tra le persone, mantenendole allineate verso un obiettivo comune.

Ecco perché è così importante parlare di cultura quando si pensa alle azioni di trasformazione digitale.

Una cultura fondata su innovazione, adattabilità e collaborazione può aiutare a potenziare, abbracciare e sfruttare la trasformazione digitale.[4]

Quindi, l'investimento in tecnologie non è una soluzione definitiva e sicura ma è qualcosa che deve essere continuamente aggiornato, gestito e curato. Si tratta però di un investimento necessario per tutte le tipologie di azienda.

2.1.1 I driver della trasformazione digitale

La trasformazione digitale è stata guidata ed accelerata da cambiamenti tecnologici e sociali.

Sicuramente, uno dei motori della TD è la customer experience, l'obiettivo è diventato quello di fornire una migliore esperienza per il cliente, coinvolgerlo in maniera efficace, allineando l'esperienza alle aspettative.

Per far ciò, le aziende devono conoscere le abitudini e le necessità dei consumatori in modo da migliorare il processo di acquisto, la comunicazione e la condivisione delle informazioni.

Come detto precedentemente, l'utente ha oggi molteplici fonti a disposizione per conoscere le caratteristiche dei diversi prodotti, questo fornisce una maggiore consapevolezza di ciò che si sta acquistando.

L'e-commerce ha radicalmente trasformato il processo di acquisto che non permette più un'interazione fisica con il cliente, le aziende devono avere la capacità di attirare i consumatori catturandone l'interesse e il gradimento attraverso siti web, pubblicità online e social.

La comunicazione e la condivisione, hanno subito un forte cambiamento, se prima i canali tradizionali quali la stampa e la televisione erano fondamentali per catturare l'attenzione degli utenti, adesso lo sono i blog, i social e i commenti di fashion blogger e influencer.

Sono proprio queste figure ad aver creato un nuovo modo di fare pubblicità incontrollabile e volubile, siamo nell'epoca in cui un gesto o un semplice Tweet da parte di persone con tanto seguito possono far precipitare in borsa azioni di aziende multinazionali.

Altro ruolo fondamentale lo ricoprono le recensioni, gli utenti a seguito di un acquisto o di un'esperienza condividono i propri pareri, le proprie opinioni evidenziando caratteristiche positive o negative; consultare le recensioni prima di procedere con l'acquisto è diventata ormai una pratica comune.

Un driver importante è l'evoluzione dei mercati, l'innovazione consente ad individui privi di esperienza ma con buone intuizioni di poter diventare leader del settore.

I confini del mercato non sono più ben definiti ma mutano rapidamente, le imprese con una grande conoscenza digitale hanno la possibilità di entrare in nuovi mercati basando il proprio vantaggio sulle soluzioni digitali.

Lo scenario fortemente dinamico richiede alle aziende velocità, dinamicità per far fronte alle minacce dei concorrenti.

Una componente chiave, sulla quale ci soffermeremo successivamente, è quella dei i

Big Data.

Le aziende raccolgono ingenti quantità di informazioni interne ed esterne all'azienda, la sfida è saper organizzare i dati per tradurli in vantaggio competitivo attraverso l'analytics e la data science.

Tra le fonti più importanti di dati si trova l'industria 4.0, che ha portato ad un'evoluzione dei processi produttivi; sensori, automazione e robotizzazione rendono i macchinari sempre più efficaci ed efficienti.

A supportare la TD sono state anche le nuove piattaforme IT, i sistemi devono assicurare stabilità, affidabilità, ed avere una forte capacità di calcolo.

Le informazioni vengono create, elaborate, diffuse e utilizzate in un processo rapido e continuo.[5]

2.2 Digitalizzazione delle aziende italiane

2.2.1 Misurare la trasformazione digitale

La progressiva digitalizzazione dei processi aziendali viene adottata come chiave di lettura per una serie di fenomeni che stanno influenzando le strategie imprenditoriali italiane come lo sviluppo di progetti di innovazione, l'emergere di nuovi modelli di business, la diffusione e l'utilizzo di nuove tecnologie e l'impatto della digitalizzazione sulla forza lavoro.

Calcolare l'avanzamento della digitalizzazione in maniera oggettiva è un lavoro a dir poco difficile, poiché, come si è visto precedentemente, questo è un fenomeno complesso e multidimensionale, non si ha ancora un'esauriente definizione statistica. Ciò rende necessario identificare diversi indicatori che, se opportunamente combinati tra loro, riescano a dare un'interpretazione unitaria.

Negli ultimi anni la trasformazione digitale è vista come un processo di evoluzione dell'organizzazione e della cultura che mira a raggiungere la digital maturity.

Deloitte ha condotto una ricerca per identificare i fattori che portano al successo della DT [6]. Dopo aver intervistato 1200 dirigenti, i risultati mostrano che per azionare il processo di maturazione digitale di un'azienda sono decisivi i seguenti sette pivot:

- Infrastruttura flessibile e sicura. L'implementazione di un'infrastruttura tecnologica che bilanci le esigenze di sicurezza e privacy in grado di supportare la domanda aziendale.
- Padronanza dei dati. L'aggregazione, l'attivazione e la monetizzazione di dati sottoutilizzati, incorporandoli in prodotti, servizi e operazioni per aumentare l'efficienza, la crescita dei ricavi e il coinvolgimento dei clienti.

- Rete aperta di talenti ed esperti digitali. E' fondamentale riorganizzare i programmi di formazione concentrandoli sulle competenze digitali, far lavorare i team attraverso modelli flessibili per accedere rapidamente alle competenze richieste e flettere la forza lavoro dell'organizzazione in base alle esigenze aziendali.
- Ecosystem engagement. Collaborare con partner esterni, tra cui organi di ricerca e sviluppo, technology hub e start-up per ottenere l'accesso a risorse come tecnologia, proprietà intellettuale o personale per aumentare la capacità dell'organizzazione di migliorare, innovare e crescere.
- Intelligent workflows. Implementare e ricalibrare continuamente i processi che sfruttano al massimo le risorse umane e tecnologiche per produrre costantemente risultati positivi e liberare risorse per azioni di alto livello.
- Unified customer experience. Fornire una user experience eccellente costruita attorno a una visione a 360 gradi del cliente, condivisa in tutta l'azienda in modo che i clienti vivano interazioni digitali e umane ben coordinate e piacevoli ma al tempo stesso efficienti in ambienti coinvolgenti.
- Business model adaptability. Ampliare la gamma di modelli di business e flussi di ricavi dell'organizzazione ottimizzando ogni offerta, adattandola alle condizioni variabili del mercato per aumentare le entrate e la redditività.

Nonostante questi punti siano fondamentali, non bastano per portare avanti un processo di TD completo, è fondamentale la presenza di fattori "soft" complementari. L'ISTAT per condurre la propria analisi ha considerato gli investimenti in tecnologie digitali di tipo infrastrutturale come per esempio connessione a Internet e acquisto di servizi cloud e ha individuato gli investimenti specializzati in Big Data, applicazioni di IoT, stampa 3D, robotica ecc.

In questo senso, l'investimento in infrastrutture digitali è stato inteso come condizione per ottimizzare i flussi informativi all'interno dell'impresa, con effetti positivi in termini di efficienza e competitività [7].

2.2.2 I dati

Le informazioni riguardanti la digitalizzazione delle imprese italiane vengono analizzate nel censimento permanente delle imprese che l'ISTAT raccoglie a cadenza triennale.

Dal censimento del periodo 2016-2018 si evince che il 77.5% [7] delle imprese con almeno 10 addetti ha investito in almeno una delle tecnologie individuate come fattore chiave della digitalizzazione (Software gestionale, fibra ottica, 4G-5G, Cloud,

cyber-security, IoT, robot, stampa 3D, simulazione, Big Data, AR - VR).

La maggior parte delle imprese utilizza un numero limitato di tecnologie, dando priorità agli investimenti infrastrutturali e lasciando eventualmente ad una fase successiva l'adozione di tecnologie applicative.

L'utilizzo di infrastrutture digitali giunge a saturazione già tra le imprese meno digitalizzate (quelle con investimenti in 4 o 5 tecnologie), solo molto più lentamente si diffondono applicazioni più complesse e con maggiore impatto sui processi aziendali: appena il 16,6% [7] delle imprese ha adottato almeno una tecnologia tra Internet delle cose, realtà aumentata o virtuale, analisi dei Big Data, automazione avanzata, simulazione e stampa 3D come si evince dal seguente grafico (Figura 2.2).

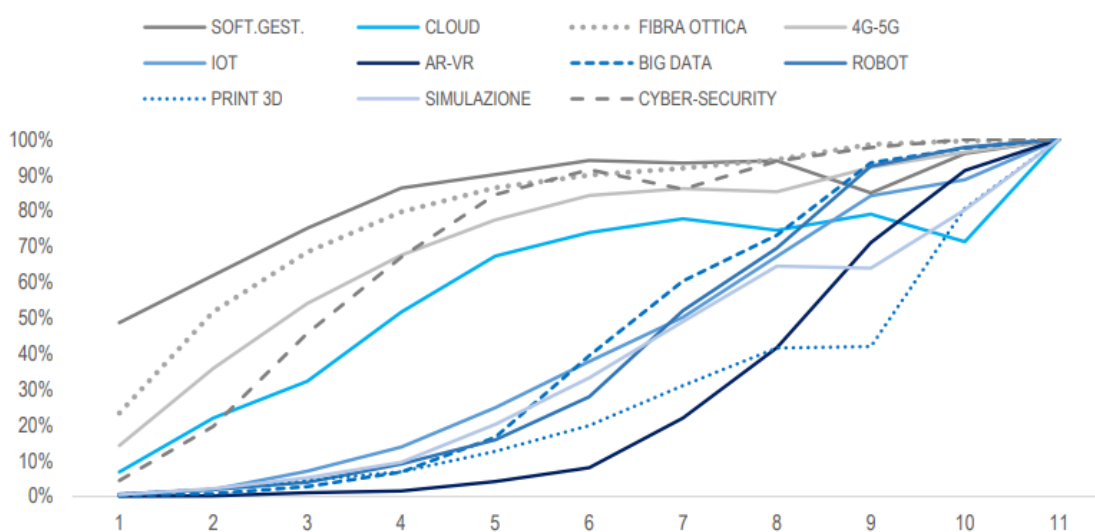


Figura 2.2: Tecnologie adottate dalle imprese italiane con almeno 10 addetti nel triennio 2016-2018

Dallo studio si evince che la digitalizzazione avvenga in due diversi stadi principali: il primo è caratterizzato dalle condizioni tecniche e culturali, il secondo dall'adozione di soluzioni applicative che aumentano efficienza e produttività.

Un fattore determinante è la dimensione dell'azienda che influisce sugli investimenti. Tra le imprese con 10-19 addetti ad effettuare un investimento è il 73,2% [7] delle imprese, tra quelle con più di 500 addetti hanno effettuato investimenti il 97,1% [7].

A livello settoriale, tra i servizi sono le telecomunicazioni ad avere una maggiore percentuale di investimento (94,2%) [7], invece nel settore manifatturiero si trova in cima la farmaceutica (94,1%) [7].

Dal rapporto annuale del 2021 pubblicato dall'ISTAT e presentato venerdì 9 Luglio 2021 a Roma, si evince che nel 2021 le tecnologie digitali sono state essenziali per la prosecuzione delle attività produttive e l'erogazione di servizi pubblici a causa dell'avvento della pandemia da Covid-19 che ha limitato fortemente gli spostamenti. Se già da tempo le tecnologie digitali erano un asset importante per l'attività delle imprese, nel 2021 rappresentano una componente strategica sia per il mantenimento della competitività attraverso l'innovazione, sia per l'evoluzione dei sistemi produttivi verso una maggiore sostenibilità.

Per questo, l'Italia ha destinato a progetti di digitalizzazione circa il 27% [8] dei 235 miliardi di risorse comprese nel proprio Programma Nazionale di Ripresa e Resilienza (222 mld) e nei fondi React-Eu (13 mld).

Ciò mira anche a colmare il divario del nostro paese rispetto all'UE in diversi ambiti di applicazione.

L'importanza raggiunta dalle ICT nell'economia può essere esemplificata dalla crescita vigorosa dell'incidenza delle professioni ICT, che ha raggiunto nel 2020 il 4,3% [8] dell'occupazione totale nell'Ue (ma solo il 3,6 per cento in Italia).

Tra le imprese con più di 10 addetti, oggi più della metà del personale usa quotidianamente computer connessi a Internet nell'attività lavorativa (il 56 % nell'Ue e il 53% in Italia) [8]; in controtendenza, nel 2020, è invece scesa la quota di imprese che ha svolto formazione ICT (pari al 15% in Italia e al 20% nell'Ue) [8], con un forte calo rispetto all'anno precedente.

Altra componente da analizzare è l'uso di servizi cloud, la crisi sanitaria ne ha spinto l'utilizzo per poter condividere dati e utilizzare applicativi da luoghi diversi; all'interno dell'Ue il loro utilizzo è cresciuto fino al 36,0% ed è cresciuto del 21,0% l'acquisto di software evoluti come per esempio i software aziendali [8].

In Italia l' aumento è stato ben maggiore, si è giunti al 59,0% per l'utilizzo di servizi cloud e al 32,0% per l'acquisto di software evoluti [8].

Questa crescita è stata probabilmente frutto degli incentivi previsti nella Legge di Bilancio del 2019 sull'industria 4.0.

Per quanto riguarda l'uso di sistemi e dispositivi interconnessi e controllo remoto l'Italia è in posizione avanzata rispetto all'Ue, è invece, in linea con la media europea per quanto riguarda l'uso dell' AI e della robotica.

Al contrario invece, gli ambiti in cui si può notare un divario negativo sono il commercio elettronico e l'uso delle tecniche di analisi di grandi volumi di dati (Big Data), ciò probabilmente a causa della mancanza di personale qualificato in grado di valorizzarli.

L'emergenza sanitaria ha dato quindi una spinta alla digitalizzazione e i benefici di ciò potremmo osservarli nel medio-lungo periodo.

2.2.3 Il livello di digital maturity nelle imprese italiane

Nell'analisi condotta sulla base del medesimo censimento relativo al triennio 2016-2018, sono stati individuati quattro profili di impresa definiti sulla base della combinazione di diverse soluzioni tecnologiche (Figura 2.3).

Il primo gruppo include le imprese che si distinguono per aver adottato almeno un software gestionale, aver affrontato un investimento in tecnologie infrastrutturali. Vengono definite imprese "asistematiche", sono caratterizzate da una potenzialità digitale ma a causa della loro dimensione o del settore a cui appartengono hanno difficoltà ad implementare un vero e proprio assetto organizzativo interamente digitalizzato.

Circa il 45% [7] del totale delle aziende compone il secondo gruppo, le imprese con tale profilo sono dette "costruttive" e mostrano una chiara spinta alla digitalizzazione.

Questo gruppo rivela l'interesse verso l'investimento in sicurezza fondamentale per proteggere i propri apparati, vengono colte le opportunità offerte dall'integrazione di diverse tecnologie come l'IoT e la sensoristica.

Il gruppo delle "sperimentatrici" individua le aziende che sperimentano diverse soluzioni informatiche, queste sono caratterizzate dagli investimenti in Big Data, simulazione, robotica ecc..

Si può notare come questo gruppo sia formato da azienda con più di 100 addetti e chiaramente una maggiore capacità finanziaria.

L'ultimo insieme, composto solo dal 3,8% [7] delle imprese, individua le aziende "mature" caratterizzate da un utilizzo integrato delle tecnologie disponibili.

Sia a livello settoriale che a livello di ripartizione geografica questi gruppi risultano omogenei.

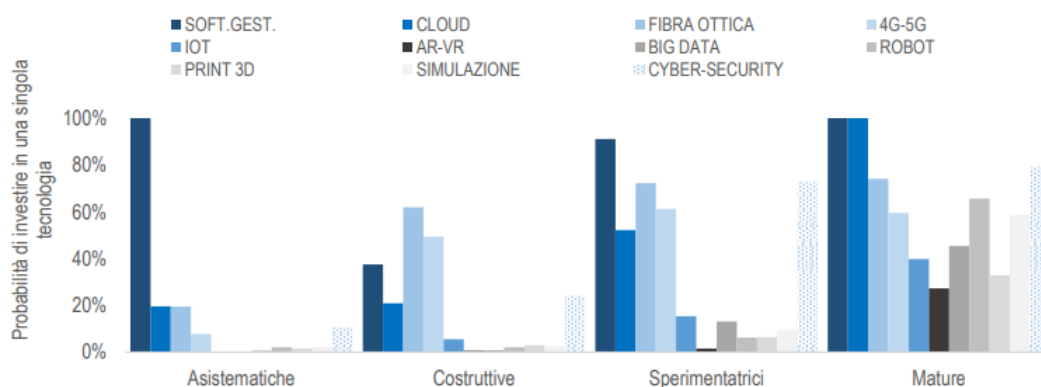


Figura 2.3: Tecnologie adottate dalle imprese italiane con almeno 10 addetti nel triennio 2016-2018

Nel 2019 il 22,5% delle aziende furono definite "asistematiche", il 35,2% "costruttive", il 17,1% rientrava nel gruppo delle "sperimentatrici" e solo il 3,0%, in calo rispetto al triennio 2016-2018 faceva parte del gruppo delle imprese "mature" [7]. Nel 2021 le aziende hanno dovuto implementare strategie difensive, di rilancio e di espansione ed in relazione ai gruppi precedentemente descritti si conferma una maggiore reattività delle imprese più strutturate, anche dal punto di vista tecnologico.

Solo il 4,1% delle imprese digitalmente mature ha dovuto affrontare un ridimensionamento delle attività, contro quote più che doppie di imprese nelle altre categorie [7].

Tra queste ultime, le imprese definite asistematiche e costruttive hanno sofferto per la crisi più delle imprese non digitalizzate.

Le imprese che hanno potuto perseguire la commercializzazione online presentano un'incidenza ridotta (14,9% in totale)[7]; queste, inoltre, non sono fortemente differenziate per capacità tecnologica, poiché tali soluzioni non richiedono grandi investimenti e competenze (pur conferendo un vantaggio nel caso delle sperimentatrici e digitalmente mature).

2.2.4 Uno sguardo al futuro

Due strategie emergono, infine, come particolarmente rilevanti per le imprese nell'immediato futuro: una ristrutturazione, in termini di riorganizzazione dei processi e degli spazi di lavoro, spesso, conseguenza diretta dell'emergenza sanitaria.

Un ulteriore sforzo di innovazione, indirizzato alla produzione di nuovi beni, all'offerta di nuovi servizi o l'adozione di nuovi processi produttivi.

La prima strategia vede le imprese sperimentatrici fortemente impegnate a consolidare gli investimenti effettuati già in periodo pre-crisi.

Sul lato dell'innovazione, invece, emerge con evidenza il vantaggio acquisito rispetto agli altri gruppi di imprese dalle unità digitalmente mature che sono, in molti casi, anche le imprese di maggiore dimensione e solidità finanziaria [8].

Capitolo 3

Business Intelligence

3.1 Definizione

Come descritto nel capitolo precedente, la Business Intelligence è parte integrante del processo di trasformazione digitale.

Le aziende italiane investono ancora poco in BI sottovalutandone l'alto potenziale. Per Business Intelligence si intendono tutti i processi e gli strumenti attraverso i quali un'azienda riesce a raccogliere dati di diversa natura per analizzarli e trarne decisioni strategiche.

Il concetto di Business Intelligence tradizionale è nato per la prima volta negli anni '60, per indicare un sistema di condivisione delle informazioni nelle organizzazioni. La BI inizia a diffondersi tra le aziende solo nei primi anni 2000 ma il suo pieno utilizzo era ancora ostacolato dai lunghi tempi di elaborazione dei dati e dalla difficoltà di utilizzo, infatti, erano pochi i tecnici esperti in grado di utilizzare software di analisi avanzata.

Con il passare degli anni gli sviluppi tecnologici segnarono un vero e proprio punto di svolta, velocità di elaborazione e complessità migliorarono notevolmente.

Pian piano la BI diventò una necessità per le aziende per continuare ad essere competitive.

Attualmente la Business Intelligence consente alle aziende di prendere decisioni con maggiore consapevolezza, mostrando dati attuali e storici all'interno del contesto aziendale.

Gli analisti possono sfruttare la BI per fornire benchmark su prestazioni e concorrenti, per consentire all'organizzazione di funzionare in modo più fluido e più efficiente.

Inoltre, gli analisti possono individuare facilmente i trend di mercato per aumentare le vendite o gli introiti.

Se usati in modo efficace, i dati giusti possono essere utili per qualsiasi attività,

dalla conformità alle assunzioni [9].

3.2 Il valore dei dati

Al giorno d'oggi, si ha la capacità di raccogliere grandi volumi di dati e storicizzarli facilmente, ciò deve essere considerato separatamente rispetto alla capacità di trasformare i dati in valore e quindi fare in modo che questi siano di supporto al processo decisionale.

La raccolta dei dati può essere automatizzata, nel contesto di estrazione della conoscenza le cose cambiano, infatti, è difficile definire un processo personalizzato di analisi dei dati in quanto bisogna conoscere il dato ed elaborarlo utilizzando diversi algoritmi iterativi.

Una delle definizioni di Big Data riporta la seguente frase:

“Data whose scale, diversity and complexity, require new architectures, techniques, algorithms and analytics to manage it and extract value and hidden knowledge from it.”

I Big Data sono di grossi volumi, di tipo diverso e con complessità diverse in base alla tipologia, per gestire tutti i dati servono nuovi paradigmi, nuove architetture, nuove tecniche e algoritmi.

Quando si parla di un'applicazione Big Data si prevede il coinvolgimento e la raccolta di grossi volumi di dati, bisogna identificarne le sorgenti, memorizzarli e successivamente è necessario analizzarli attraverso il data mining e il machine learning.

Gli attori che intervengono in questo contesto sono molteplici, le informazioni fornite in output sono diverse.

L'analista analizza i dati per estrarne della conoscenza, quest'ultima deve essere restituita agli utenti in maniera chiara e comprensibile, gli utenti a loro volta, possono accedere ai dati e passare all'analisi per rispondere alle domande che riguardano l'azienda.

Il coinvolgimento degli stakeholders nel processo di estrazione della conoscenza è molto importante.

Nel contesto aziendale le principali sorgenti di dati sono persone, funzioni, applicazioni, reti tecnologiche e procedure che hanno come obiettivo finale quello di rendere disponibili una serie di informazioni e dati nel momento e nel luogo desiderati.

Sono state individuate cinque V per caratterizzare la complessità dei dati che si vogliono gestire:

- Volume

- Velocity
- Variety
- Veracity
- Value

La prima V è quella del volume che aumenta esponenzialmente nel tempo, in pochi anni è cambiata la disponibilità di dati a livello individuale, oggi tecnologie cloud e virtualizzazione aiutano nella gestione dell'enorme volume di dati disponibili, semplificando i processi di raccolta, immagazzinamento e accesso agli stessi.

La seconda V è la velocità con cui i dati vengono generati ed elaborati, quando i dati vengono raccolti con grande velocità si deve anche essere capaci di analizzarli altrettanto velocemente altrimenti il dato nel tempo può perdere la propria attrattività, il valore viene dall'analisi del dato grazie agli algoritmi che devono però essere veloci.

Altra V è la variety quindi varietà in termini di provenienza, tipi e strutture, il formato dei dati condivisi è diverso e questo complica le analisi.

La veracity indica che i grossi dati hanno spesso problemi di qualità, se si ha un dato qualitativamente buono si otterrà un modello accurato ma se si ha un dato di qualità bassa si avrà un risultato scarso; è necessario quindi esplorare i dati per capirne la qualità.

La V più importante è quella del valore cioè la capacità di trasformare il dato per l'analisi, il risultato va tradotto in azioni che migliorano il business.

Queste caratteristiche, tuttavia, se da un lato aumentano la complessità dell'ambiente di lavoro, dall'altro, in diversi settori industriali, stanno già generando nuovi flussi di reddito, grazie allo sfruttamento intelligente dei dati stessi.

3.3 Data Science

In riferimento all'analisi dei dati è nata la disciplina della Data Science cioè la scienza che analizza i dati utilizzando diverse tecniche per ottenere il maggior beneficio.

Nel processo di Data Science si ha generazione, acquisizione, storicizzazione e analisi; all'interno dell'analisi si trova la capacità di elaborazione e combinazione di algoritmi diversi che danno la possibilità di estrarre conoscenza.

- Generazione: i dati possono essere ottenuti attraverso la storicizzazione passiva che si ha quando vengono acquisiti i dati relativi ad un evento che si è già

verificato, si ha generazione attiva quando l'utente decide di generare nuovo contenuto e metterlo a disposizione, si ha produzione automatica quando i dati vengono acquisiti automaticamente, si pensi ad esempio, ai dati rilevati dai sensori delle macchine utilizzate nei processi produttivi.

- **Acquisizione:** il dato generato è acquisito per essere analizzato, l'acquisizione può avvenire attraverso collection pull-based o push-based, tramite una trasmissione in cui i dati vengono trasferiti al data center attraverso collegamenti ad alta capacità, oppure i dati possono essere acquisiti dopo il preprocessing quindi dopo l'integrazione, la pulizia e l'eliminazione della ridondanza.
- **Storicizzazione:** il dato è memorizzato con opportuni sistemi, grazie a infrastrutture, tecnologie ed architetture di rete.
Lo storage influisce anche sull'estrazione della conoscenza.
- **Analisi:** l'analisi è la parte più complessa che richiede personalizzazione, esistono numerosi algoritmi ma questi devono essere studiati ed adattati alle diverse realtà.
Nel contesto dell'analisi dei dati è importante avere un obiettivo che spesso può essere non definito ma è necessario capire cosa potrebbe essere utile per il supporto del business, il contesto e le metodologie dipendono dagli obiettivi.

3.3.1 I possibili obiettivi dell'analisi dei dati

L'analisi dei dati è una branca molto varia, si possono individuare tre macro-obiettivi associati a tre macro-categorie di algoritmi.

L'obiettivo descrittivo è il primo livello, prevede la descrizione di una situazione analizzando dati ed eventi passati ed attuali, si analizza una collezione di dati per poter conoscere un modello, una caratterizzazione, analizzare le relazioni per comprendere il tipo di dato, questa analisi si applica, per esempio, quando si studia la clientela per caratterizzare il business di un'azienda.

L'obiettivo predittivo utilizza tecniche statistiche e apprendimento automatico per fare previsioni, simulare scenari possibili che aiutino a prevedere i risultati di scelte, azioni e decisioni prima che queste vengano applicate nella realtà.

Vengono sfruttati i dati storici per estrarre ulteriori informazioni, non si tratta di previsioni certe ma di tecniche analitiche che consentono di capire con quali probabilità possano verificarsi determinate situazioni.

L'analisi predittiva viene utilizzata, per esempio, quando si esamina la clientela per calcolare le possibilità di perdere clienti consentendo alle aziende di intervenire in tempo.

L'obiettivo prescrittivo utilizza combinazioni di matematica e statistica per poter fare ipotesi, modellazioni e simulazioni in grado di guidare il processo decisionale e

l'efficacia delle decisioni stesse.

L'analisi prescrittiva spiega il perché di determinati avvenimenti, fa uso di machine learning e intelligenza artificiale.

Questi tre macro-settori vanno poi declinati e combinati.

Per condurre le analisi sono necessari diversi algoritmi, per ognuno di essi è necessario sapere cosa riceve in input, cosa rilascia in output, quando conviene utilizzarlo, come configurare i parametri, ed infine, l'esperienza suggerisce come combinare i diversi algoritmi.

3.4 Knowledge Discovery from Data

Il processo di estrazione di conoscenza (KDD) segue diverse fasi ed è iterativo (Figura 3.1).

Il compito del data scientist è quello di identificare la combinazione e la configurazione migliore degli algoritmi da utilizzare in un determinato dominio.

La configurazione di un processo di data mining è molto complessa, è necessario conoscere gli algoritmi e capire i diversi aspetti di interesse dei dati per individuare una soluzione efficace; in ogni step si effettua un'operazione specifica ed è necessario utilizzare un approccio sperimentale.

Per definire una tecnica di trasformazione del dato è necessario testare diverse soluzioni, valutare come queste operano ed essere in grado di esaminare l'impatto sull'intero processo.

A questo proposito, è utile estrarre una sotto parte di dati cioè fare un campionamento, si prepara la pipeline migliore e l'intero processo di mining viene ripetuto su tutta la collezione di dati.

All'interno di questo processo esistono diverse fasi: si parte da un set di dati che, come detto precedentemente, il più delle volte necessita di essere suddiviso in modo tale da proseguire analizzando solo alcuni sottogruppi; la fase di selection è delicata poiché se si commettono degli errori nel campionamento le analisi successive non saranno qualitativamente corrette.

Dopo aver campionato il dataset originario si passa alla fase di preprocessing, i dati del mondo reale sono "sporchi" devono essere sottoposti ai processi di data cleaning (riduzione degli effetti del rumore, identificazione e rimozione degli outlier, risoluzione delle inconsistenze) e data integration (riconciliazione dei dati estratti dalle diverse sorgenti, integrazione dei meta dati, identificazione e risoluzione dei conflitti tra i valori dei dati, gestione delle ridondanze).

Esistono diverse tecniche ed algoritmi per effettuare la fase di preprocessing.

Successivamente i dati vengono trasformati, la data transformation permette di rappresentare i dati in uno spazio diverso, le tecniche di trasformazione assumono un ruolo centrale perché possono consentire l'identificazione di conoscenza utile su

collezioni tendenzialmente sparse.

Quando i dati sono sparsi catturare i pattern rari potrebbe essere difficile, il dato dopo la trasformazione è più fruibile ed utile.

Nel momento in cui si applica la trasformazione si fa un'analisi comparativa in un contesto sperimentale ossia, la pipeline deve essere applicata sul dataset prima e dopo la trasformazione e per essere efficace, i risultati ottenuti dopo la trasformazione devono essere migliori rispetto a quelli ottenuti prima della stessa.

Una volta che il dato è stato trasformato, è utilizzato come input per estrarre conoscenza nascosta, questa va interpretata dall'esperto di dominio al fine di poter identificare una conoscenza fruibile dagli altri stakeholders.

Per l'estrazione della conoscenza si possono utilizzare algoritmi predittivi, attività di regressione, algoritmi di clustering, tecniche di rilevamento delle anomalie ecc.. Infine, si ha l'interpretazione e la visualizzazione del modello; questo ultimo step non è banale in quanto è necessario presentare il modello in maniera chiara ed intuitiva; grafici e dashboard sono fondamentali per trasmettere agli utenti finali il risultato dell'analisi.

Ci soffermeremo sulla data visualization nel capitolo successivo.

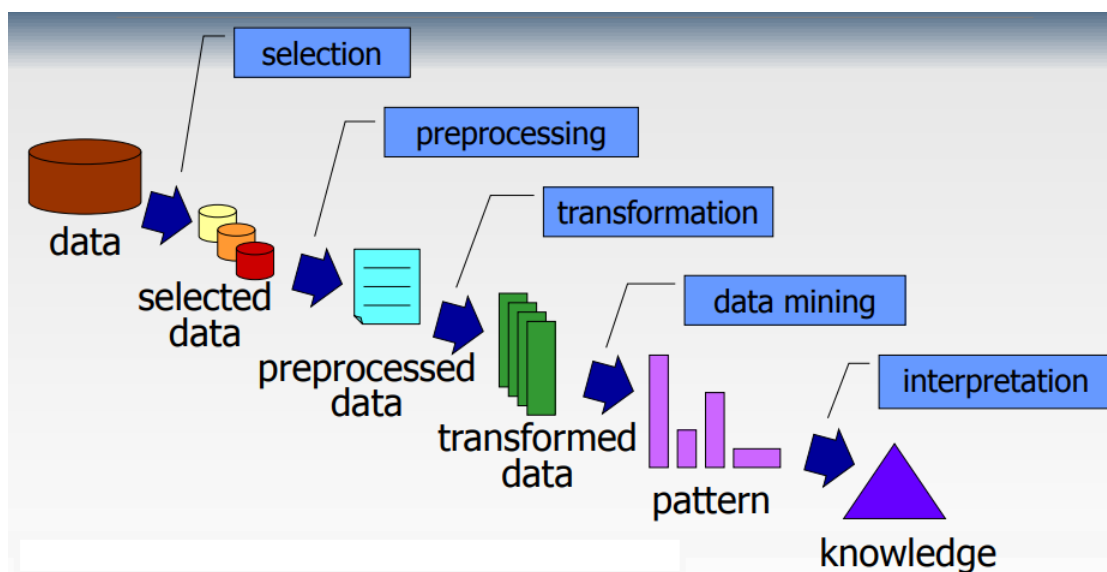


Figura 3.1: Processo di estrazione della conoscenza

3.5 Struttura ed architettura

Padroneggiare e leggere i dati in un contesto aziendale, come discusso in precedenza, è la chiave del successo.

L'attività di Business Intelligence pone le sue radici nei dati e di conseguenza,

i database ovvero gli archivi strutturati dediti alla collezione delle informazioni devono essere efficienti dal punto di vista della velocità di memorizzazione, della varietà di dati da gestire e della veridicità.

I DB si distinguono in due grandi classi, i DB relazionali e i DB non-relazionali.

I database relazionali sono basati su tabelle, ogni record è strutturato come una riga, per ogni tabella esiste uno schema predefinito e le modifiche sono consentite ma molto difficili e costose; i DB relazionali sono scalabili verticalmente ovvero, vengono ridimensionati attraverso la potenza dell'hardware; per definire e manipolare i dati si utilizza il linguaggio SQL (Structured Query Language).

I database non-relazionali detti anche NoSql sono soluzioni di archiviazione specializzate, basate su documenti e coppie chiave-valore, il cambiamento dello schema è dinamico per ogni documento; i DB non-relazionali sono adatti a dati semi-strutturati o non strutturati; sono scalabili orizzontalmente, sono ridimensionati aumentando i server in pool di risorse per ridurre il carico.

In sintesi, le caratteristiche del mondo relazionale sono: standardizzazione, maturità, funzionalità utili anche se costose computazionalmente; quello che ad oggi serve però è un mondo molto più flessibile che si adatta meglio ai database non-relazionali. Alla base di un database c'è il DataBase Management System (DBMS) definito come il "motore" del DB che è in grado di gestire le operazioni effettuate sulle informazioni e ridurre l'inconsistenza dei dati, assicurare affidabilità, persistenza, efficienza ed efficacia.

I DBMS sono fondamentali per l'ottimizzazione delle prestazioni in termini di risorse e durata.

3.5.1 I Data Warehouse

Storicizzare solo i dati non basta, è fondamentale saperli utilizzare nel modo corretto, per questo, è stato introdotto il concetto di Data Warehouse, un insieme di tecniche e metodologie che danno la possibilità di definire indicatori per supportare il processo decisionale.

Si devono distinguere i sistemi relazionali che servono ad offrire la gestione delle attività ordinarie dai sistemi decisionali che servono per supportare le decisioni, entrambi lavorano sugli stessi dati ma con granularità differenti.

Il supporto decisionale riceve i dati da diverse fonti, li integra e crea un repository comune con un dettaglio minore rispetto ai dati originali.

L'obiettivo è modellare i dati aziendali per ottenere informazioni fruibili ed indicatori che possono essere facilmente analizzati per supportare le decisioni.

I dati di un'azienda si possono aggregare ottenendo diversi livelli di dettaglio; potrebbe essere necessaria un'analisi dei dati in tempo reale, giornaliera, semestrale o annuale, il contenuto informativo varia a seconda dell'attore di business per cui è

stata pensata l'analisi.

Il DW è stato ideato per creare dashboard personalizzate, che differiscono per periodicità, attori e processo decisionale che devono supportare.

Il processo di raccolta dei dati da diverse sorgenti e la successiva organizzazione e centralizzazione in un unico repository è detto ETL (Extract, Transform, Load), tale processo richiede una progettazione attenta, in modo da evitare problemi di integrazione successivi.

La prima fase è quella di estrazione dei dati da sorgenti esterne quali database, registri di attività, report ecc., la seconda, più critica, è la trasformazione, vengono applicate una serie di regole come, per esempio, la standardizzazione e l'ordinamento per garantire l'accessibilità e la qualità dei dati e vengono effettuate conversioni di formato; l'ultima fase è quella di caricamento nel data warehouse o nei data marts (sottoinsiemi dipartimentali focalizzati su un determinato settore) che può essere eseguita tramite caricamento completo o refresh periodici.

Le architetture per DW sono tipicamente a due o più livelli, ciò garantisce la separazione dei dati in ingresso da quelli oggetto di analisi e una maggiore scalabilità. L'architettura più semplice è composta da due livelli, sorgenti e DW, le caratteristiche di questo tipo di architettura sono il disaccoppiamento delle sorgenti, la facilità di gestione delle differenti granularità temporali dei dati operazionali ed analitici e la necessità di svolgere velocemente la preparazione dei dati (ETL).

Quando, invece, il sistema è più complesso viene inserito un terzo livello di appoggio tra i due precedenti detto staging area (Figura 3.2), questa è un'area di transito che permette di separare l'elaborazione ET dal caricamento nel DW, in questo modo è possibile compiere operazioni più complesse di pulizia e trasformazione dei dati, gli svantaggi sono l'introduzione di ulteriore ridondanza e l'aumento dello spazio necessario per i dati [10].

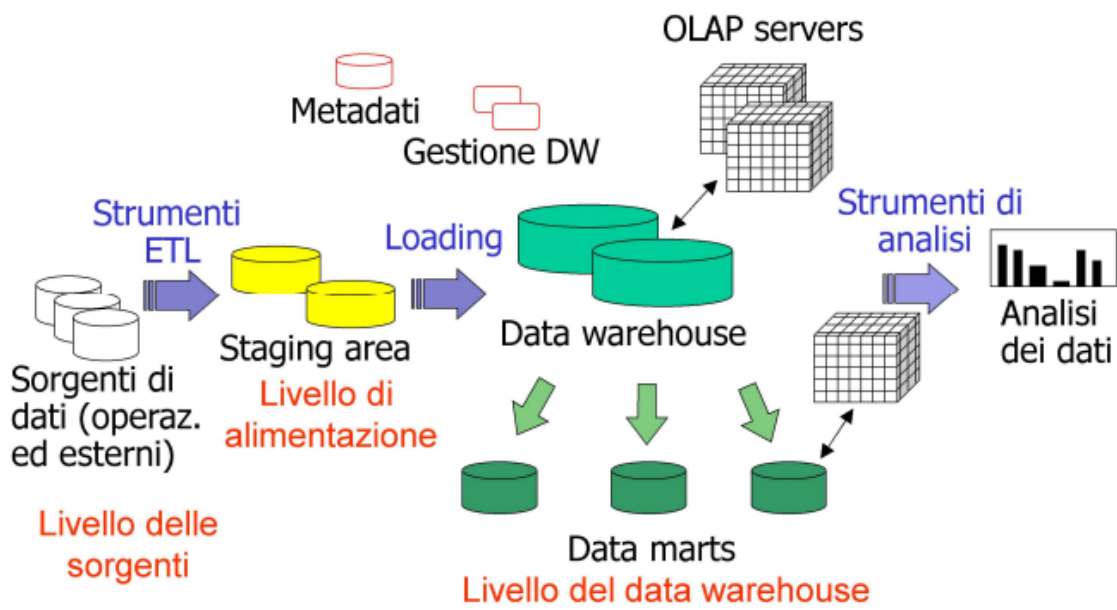


Figura 3.2: Architetture per data warehouse [10]

I dati possono essere rappresentati secondo uno schema multidimensionale o relazionale: nel primo caso i dati sono rappresentati come un ipercubo con tre o più dimensioni che indicano le misure su cui si esegue l'analisi, gli elementi vengono individuati dall'intersezione delle diverse dimensioni.

Nella rappresentazione relazionale, invece, le misure vengono memorizzate in una tabella dei fatti, le dimensioni descrivono il contesto di ogni misura nella tabella secondo un modello a stella.

3.5.2 Progettare un data warehouse

La progettazione di un DW può essere sviluppata con un approccio top-down o un approccio bottom-up.

Nel primo caso il data warehouse deve fornire una visione globale e completa dei dati aziendali, il costo della progettazione è significativo e il tempo di realizzazione è lungo.

Nel caso di un approccio di tipo bottom-up, invece, la realizzazione è incrementale, si aggiungono data mart definiti su settori aziendali specifici, il costo e il tempo di consegna sono contenuti.

Il progetto di un DW richiede diverse fasi, alcune delle quali possono essere parallelizzate, altre richiedono l'ottenimento di risultati che provengono da più attività precedenti.

Le fasi dello sviluppo sono la definizione dei requisiti, l'attuazione e la manutenzione,

l'intero processo è chiaramente individuabile in figura 3.3.

Bisogna sottolineare alcuni aspetti principali: il primo aspetto è quello legato ai dati, è fondamentale conoscere le sorgenti dei dati e i desiderata degli attori di business, si devono raccogliere le esigenze di analisi dei dati che dovranno essere soddisfatte ed i vincoli realizzativi dovuti a sistemi informativi esistenti.

Il DW deve prima essere progettato concettualmente cioè devono essere definiti gli interessi, le dimensioni e i KPI che si vogliono calcolare.

E' necessario mappare gli obiettivi con i dati disponibili andando ad evidenziare le informazioni non ancora accessibili per capire se possono essere estratte oppure bisogna integrare i dati con quelli provenienti da sorgenti diverse.

Segue un approccio metodologico, vengono descritti i requisiti applicativi per la costruzione del data mart: si definiscono gli eventi di interesse detti anche "fatti" da monitorare.

Ogni fatto rappresenta una categoria di eventi di interesse per l'azienda, gli eventi sono caratterizzati da dimensioni descrittive, intervalli di storicizzazione e misure di interesse.

Devono essere definiti i requisiti strutturali quali la periodicità dell'alimentazione, lo spazio disponibile, il tipo di architettura del sistema e la pianificazione del deployment.

Il processo continua con la progettazione concettuale che non prevede un formalismo di modellazione; una convenzione efficace è il Dimensional Fact Model che, grazie alla redazione di un modello grafico, offre una documentazione di progetto utile.

I grafici sono composti da fatti che modellano un insieme di eventi di interesse, dimensioni che descrivono le coordinate di analisi di un fatto e misure che ne descrivono una proprietà numerica.

Il modello concettuale è poi trasformato in modello logico che prevede la trasformazione delle dimensioni e dei fatti in tabelle.

Esistono diversi schemi possibili tra cui lo schema a stella, molto utilizzato, che prevede una tabella per ogni dimensione e una chiave primaria generata artificialmente, le gerarchie non vengono rappresentate esplicitamente e la rappresentazione è completamente de-normalizzata.

Nello schema a stella ogni schema di fatto viene tradotto in una tabella, la chiave primaria è data dalla combinazione delle chiavi esterne delle dimensioni e le dimensioni sono attributi della tabella.

La progettazione deve essere ottimizzata, è molto importante selezionare le strutture adatte per supportare le interrogazioni più frequenti, prevedendo, per esempio, l'uso di viste materializzate e considerare i vincoli di spazio e tempo disponibili per i dati e il loro aggiornamento.

Una volta decise le tecnologie, il progetto passa nella fase attuativa, dopo aver costruito la struttura questa può essere utilizzata dai vari attori.

3.5.3 L'analisi dei dati

Le operazioni di analisi nel DW riguardano funzioni di aggregazione sulle misure definite lungo una o più dimensioni di analisi, è utile stabilire delle finestre mobili su cui effettuare calcoli.

Le funzioni di aggregazione applicate sulle misure sono fondamentali, la finestra di analisi può essere statica o dinamica nel tempo, le operazioni di confronto sono molto utili a livello di business, per esempio potrebbe essere utile confrontare l'incasso di oggi con quello medio, la sola funzione GROUP BY potrebbe essere riduttiva.

Il dato deve essere esplorabile cioè dal dato aggregato si deve poter effettuare un'ispezione con granularità più specifica oppure si deve poter eseguire il calcolo della stessa aggregazione con granularità minore.

Per poter eseguire le interrogazioni esistono prevalentemente tre strumenti, un ambiente controllato di query, strumenti specifici di query e generazione di rapporti e strumenti di data mining.

Nell'ambiente controllato di query sono definite delle strutture prefissate normalmente parametriche, procedure specifiche di analisi e rapporti con struttura prefissata. Negli ambienti ad hoc è invece possibile definire interrogazioni OLAP di tipo arbitrario progettate al momento dall'utente, in questo caso il risultato andrà a rappresentare i rapporti che servono ai diversi attori di business.

Le analisi OLAP (online analytical process) permettono di analizzare i dati online ossia di andare ad esplorare il dato partendo da un alto livello si procede ad una fase di analisi più di dettaglio o viceversa.

Le operazioni di ricerca disponibili applicate a diverse query SQL sono [11]:

- **roll up** che consente di effettuare una riduzione di dettaglio rispetto ad una delle dimensioni presenti con il conseguente aumento di livello in una gerarchia oppure di eliminare una delle dimensioni presenti e il **drill down** che consente l'aumento di dettaglio di una delle dimensioni con la conseguente riduzione di un livello in gerarchia oppure l'aggiunta di una dimensione;
- **slice and dice** che operano una riduzione del volume dei dati da analizzare mediante uno o più predicati di uguaglianza;
- **pivot di tabelle** che riorganizzano l'orientamento della struttura multidimensionale senza variare il livello di dettaglio;
- **ordinamento**.

Capitolo 4

Data visualization

A cosa servono i dati se non si è in grado di leggerli?

Solo quando questi acquisiscono un significato, si tramutano in conoscenza.

Con l'aumento della quantità di Big Data, sempre più persone utilizzano strumenti di visualizzazione dei dati per accedere a informazioni dettagliate sul proprio computer e sui dispositivi mobili.

I dati da soli, infatti, non hanno nessun valore concreto e nessun significato.

Se non si sanno trarre deduzioni basate sui dati raccolti, allora questi sono inutili.

Ecco perché si ha bisogno di strumenti che facilitino il lavoro di interpretazione.

La data visualization è definita come l'esplorazione visuale e la rappresentazione grafica di dati di qualunque dimensione, natura e origine.

La rappresentazione visiva consente di comprendere velocemente le informazioni, identificare le relazioni e i modelli, evidenziare i trend emergenti e condividere le scoperte con gli altri.

L'arte di visualizzare i dati è piuttosto antica, l'iconografia si è sviluppata per illustrare un fenomeno con semplicità e trasmettere un messaggio che sarebbe difficile da spiegare con i soli numeri.

I grafici, come si intendono oggi, furono inventati intorno al 1700 da uno scozzese, William Playfair, esperto di economia politica, successivamente, alla fine dell'Ottocento l'iconografia si impose come strumento informativo e di studio.

Nell'era digitale alcune iconografie sono diventate interattive ed aggiornabili, la data visualization si è rapidamente evoluta in un concetto che unisce scienza e arte e che certamente modificherà il business nei prossimi anni.

La DV è nata per esplorare e analizzare i dati in maniera visuale e avere una vista di sintesi sul business, inoltre, è possibile identificare le aree di business da migliorare, comprendere i fattori chiave che influenzano determinati trend, effettuare previsioni.

Ci sono diverse tecniche di data visualization, le più frequenti sono la básica rappresentazione dei dati tramite grafici, mappe e KPI, altra tecnica prevede l'utilizzo di dashboard che permettono una visione complessiva di singole visualizzazioni

coerenti tra di loro e collegate da uno specifico argomento.

Una dashboard è un "cruscotto" che racchiude elementi grafici strutturati quali tabelle, istogrammi, diagrammi, mappe ecc., che rendono accessibili le informazioni complesse a colpo d'occhio.

I grafici che si possono inserire nelle dashboard sono generalmente suddivisi in base ai loro obiettivi o alle loro caratteristiche visive.

La prima categoria è quella dei grafici di confronto (Figura 4.1) pensati per visualizzare le differenze tra gli elementi, questi si basano sulla capacità dell'occhio umano di identificare forme più lunghe o più grandi, il posizionamento e l'affiancamento degli elementi visivi rendono il confronto molto semplice.

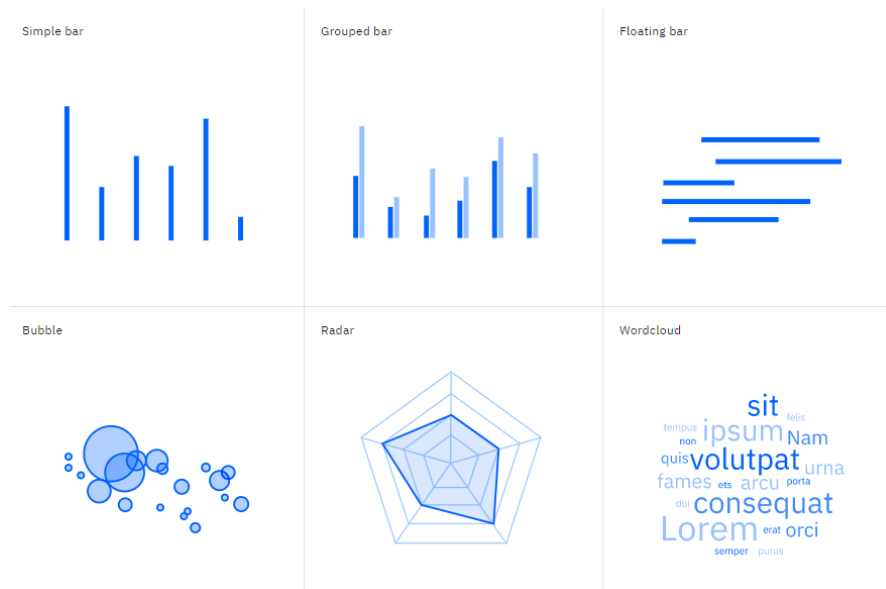


Figura 4.1: Grafici di confronto: Simple bar, Grouped bar, Floating bar, Bubble, Radar, Wordcloud [12]

I grafici di tendenza (Figura 4.2) rappresentano i dati correlati alla dimensione temporale, vengono usati per tenere traccia della trasformazione degli eventi, si basano sulla direzione degli assi per mostrare l'evoluzione dei valori.

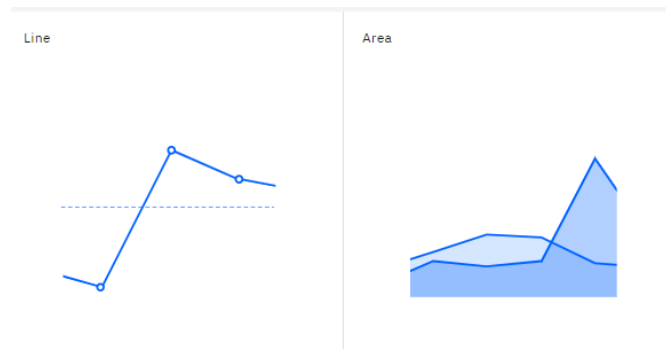


Figura 4.2: Grafici di tendenza: Line e Area [12]

Seguono grafici di proporzione (Figura 4.3) che mostrano la suddivisione di un valore tra diverse categorie o gruppi, vengono usati per rappresentare le percentuali o i valori assoluti, tali grafici sono ottimali per i dati categorizzati.

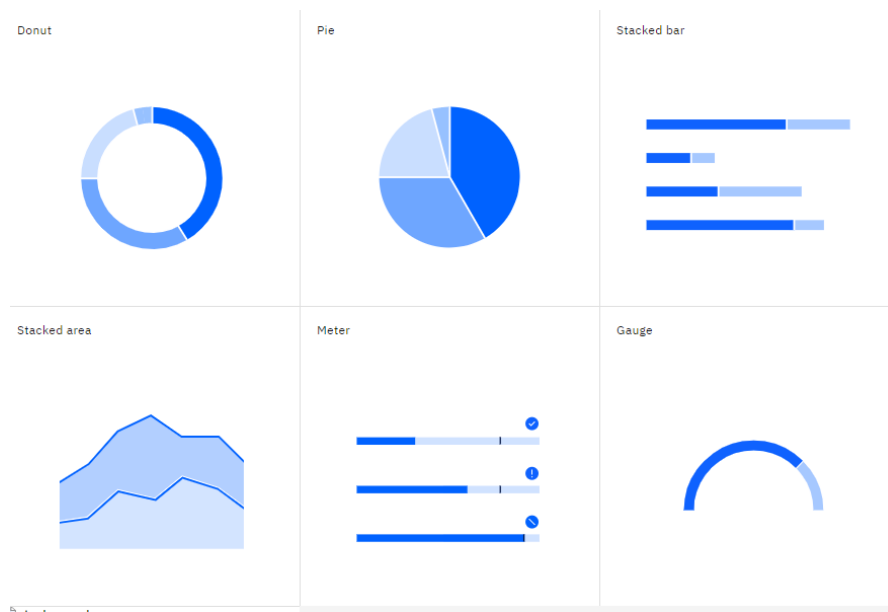


Figura 4.3: Grafici di proporzione: Donut, Pie, Stacked bar, Stacked area, Meter, Gauge [12]

I grafici di correlazione (Figura 4.4) sono adatti per evidenziare le possibili correlazioni tra due o più indicatori e come potrebbero influenzarsi a vicenda, questa tipologia di grafici ha l'obiettivo di rendere più semplice l'individuazione di comportamenti combinati.

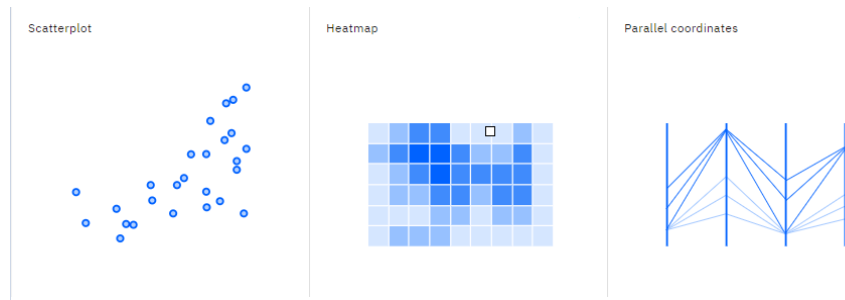


Figura 4.4: Grafici di correlazione: Scatterplot, Heatmap, Parallel coordinates [12]

Grafici di relazione (Figura 4.5) e connessione indicati per rappresentare le gerarchie.

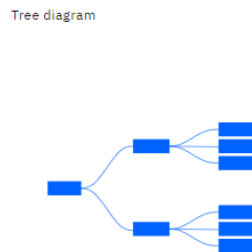


Figura 4.5: Grafici di relazione e connessione: Tree diagram [12]

Infine, si hanno le mappe (Figura 4.6), il modo più semplice e immediato per comunicare informazioni geo localizzate, queste consentono all'utente di riconoscere aree e luoghi, di comprendere il contesto geografico dell'argomento basandosi sulla posizione degli elementi.

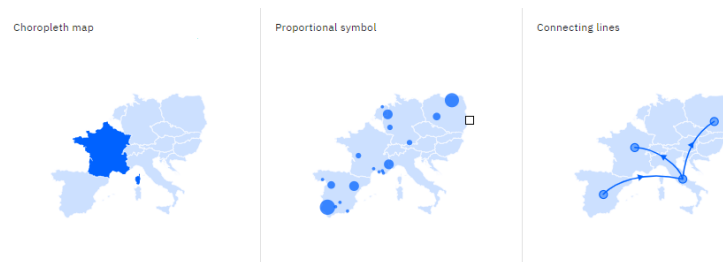


Figura 4.6: Mappe: Chloropleth ap, Proportional symbol, Connecting lines [12]

Prima di utilizzare una nuova tecnologia come la DV bisogna capire quali sono gli obiettivi e i bisogni degli stakeholders.

Intraprendere un progetto di DV significa prima di tutto comprendere i dati che si desidera visualizzare, le dimensioni e la cardinalità, determinare cosa si sta visualizzando e quale tipo di informazioni si vuole comunicare.

Conoscere gli utenti finali è decisivo per l'elaborazione dell'informazione visiva, in quanto, bisogna utilizzare una visualizzazione che trasmetta le informazioni nella maniera più appropriata e semplice.

Infatti, una delle sfide maggiori è decidere quale visualizzazione utilizzare per rappresentare al meglio le informazioni.

Per presentare i dati estratti durante il processo di analisi bisogna usare una grafica adeguata, non tutte le tipologie esprimono la stessa cosa, alcune potrebbero non adattarsi a ciò che si desidera comunicare.

Enfatizzare le informazioni principali aiutando il lettore nel processo di comprensione, evitare di riportare i dati in semplici tabelle, concentrarsi nel trarre conclusioni, evidenziare i dati positivi e negativi, comparare il presente con dati riferiti a periodi temporali precedenti.

4.1 Business Dashboard in ambito aziendale

Tutti i settori professionali possono trarre beneficio dall'avere delle grafiche più comprensibili e per ognuno di questi le necessità sono diverse.

Un template di visualizzazione dei dati aiuta la pianificazione strategica ed il monitoraggio in real-time, due obiettivi che sembrano opposti ma che si completano l'un l'altro.

Osservare i dati permette di modificare le azioni a seconda dei risultati ottenuti, dando inoltre una visione a lungo termine definita per il progetto o per l'area aziendale a cui si sta facendo riferimento.

Le dashboard sono un mezzo di analisi, confronto con previsioni e obiettivi, supportano le fasi di monitoraggio e condivisione delle strategie e delle risorse.

Dopo aver completato il lungo processo di analisi dei dati si può procedere con la visualizzazione delle informazioni attraverso l'ausilio di software.

I manager si rivolgono sempre di più ai cruscotti aziendali per gestire le organizzazioni, il business e i processi aziendali.

Fornire ai manager una dashboard ben progettata che mostri le informazioni chiave è fondamentale per monitorare gli eventi, scoprire i problemi e migliorare le prestazioni.

La disponibilità delle informazioni e la rapidità del processo decisionale sono fattori chiave per il successo di un'azienda.

Ogni dashboard è pensata per un'area e un utente specifico.

Disporre di un sistema in grado di visualizzare tali informazioni in maniera ottimale e immediata diventa di importanza fondamentale.

La progettazione dei diversi cruscotti è compito dell'esperto di BI ed è la fase finale del complesso processo di analisi dei dati precedentemente analizzato.

Esistono tre tipologie di dashbord principali:

- operational dashboard monitorano i processi operativi, forniscono informazioni dettagliate e poco riassunte, le metriche sono aggiornate quotidianamente o addirittura ogni ora o minuto, questa tipologia è più di supporto al monitoraggio che all'analisi. Sono impiegate prevalentemente da operatori di prima linea e supervisori.

Un esempio è un cruscotto operativo di marketing (Figura 4.7) che mostra il rendimento di tre campagne di marketing nelle ultime 12 settimane:

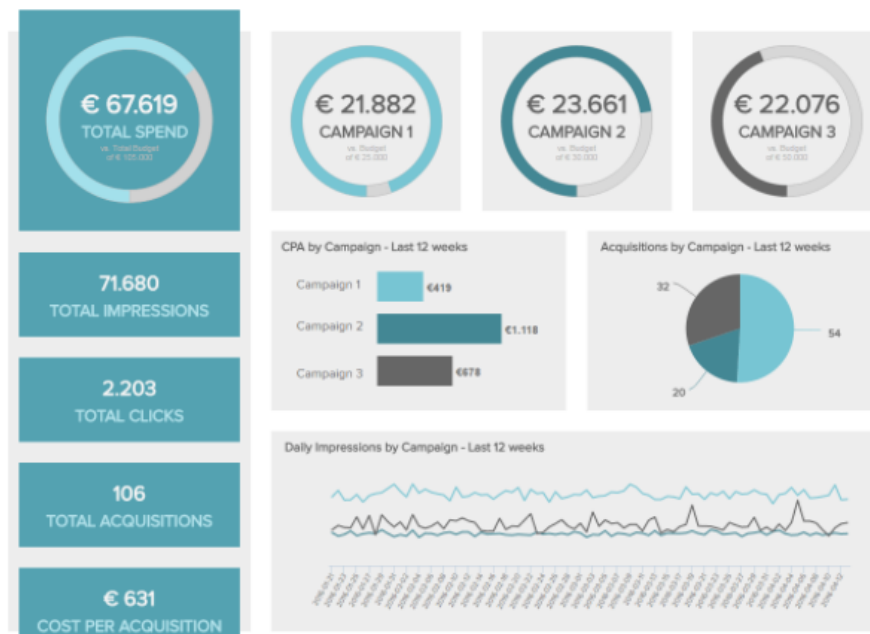


Figura 4.7: Esempio di Operational dashboard

tale dashboard fornisce importanti informazioni operative e indicatori di prestazioni chiave per il team di marketing sul costo per l'acquisizione, il numero totale di clic, le acquisizioni totali ottenute e l'importo totale speso nella campagna specifica.

Eventuali modifiche significative saranno immediatamente chiare al team di marketing. Questo tipo di visualizzazione è utile perché un reparto o un'agenzia di marketing può adattare le proprie attività operative in base a dati in tempo reale e i team non devono attendere report e analisi estesi e tradizionali presentati in un foglio di calcolo [13].

- strategic dashboard monitorano il raggiungimento di obiettivi strategici con l'obiettivo di allinearli all'organizzazione, queste sono di supporto al monitoraggio e vengono di frequente affiancate da Balanced Scorecard. Di solito sono complesse nella loro creazione, forniscono un impatto a livello aziendale e sono utilizzate principalmente dal management di livello senior. Un esempio di strategic dashboard è quella che riporta i dati di gestione aziendale (Figura 4.8):



Figura 4.8: Esempio di Strategic dashboard

La dashboard si concentra sulle entrate totali, è impostata su un intervallo di tempo specifico e include KPI significativi: costi di acquisizione del cliente, valore del cliente e obiettivo di vendita.

Questa dashboard risponde a quanto segue: quali sono i clienti con più valore, quali sono le entrate attuali rispetto allo stesso periodo dell'anno scorso?

Pur affrontando valori specifici, incorporando specifici indicatori di prestazione e utilizzando un linguaggio qualitativo e qualitativo comune, questa dashboard rappresenta il valore aziendale e la specifica linea d'azione, utilizzando metriche e analisi di confronto [13].

- tactical dashboard tracciano i processi e i progetti collegati ad una parte dell'organizzazione o ad un limitato gruppo di risorse, sono utilizzate per comparare le performance di una specifica area aziendale, per pianificare i

budget, fare previsioni o monitorare i risultati; questa tipologia viene sfruttata molto da manager e business analysts. Le dashboard tattiche aiutano a guidare gli utenti attraverso il processo decisionale. Sfruttano la natura interattiva dei dashboard fornendo agli utenti la possibilità di esplorare i dati.

L'esempio seguente (Figura 4.9) mostra una panoramica dettagliata di un progetto con tempistiche specifiche e l'efficienza delle parti coinvolte. È possibile definire rischi specifici, visualizzare i progressi complessivi e i tempi medi di esecuzione di attività specifiche. Al termine del progetto, è possibile creare un report IT completo, valutare i risultati e rendere più efficaci i progetti futuri.



Figura 4.9: Esempio di Tactical dashboard

Avendo lo strumento giusto sotto forma di dashboard delle operazioni IT, una singola schermata può fornire a un project manager tutti i dati di cui ha bisogno per analizzare tutti gli aspetti importanti del progetto.

Questo particolare elemento visivo è impostato per monitorare gli sforzi di gestione del progetto e dare allarme in caso di anomalie all'interno del processo [13].

Le dashboard oltre ad essere strumenti di Business Intelligence, come mostrato negli esempi precedenti, sono anche considerati strumenti fondamentali per il Project Management, poiché sono i principali strumenti di reporting per la gestione dei progetti, misurano le performance rispetto ai target e alle soglie; metriche e KPI vengono impiegati soprattutto nelle fasi di monitoraggio di un progetto.

Il primo punto per creare una dashboard efficiente è sicuramente confrontarsi con l'utente finale che dovrà utilizzarla, comprendere quali sono gli elementi che si vogliono individuare o monitorare.

Una volta compresa l'area aziendale che si deve analizzare bisogna identificare le

metriche e i KPI.

Le metriche sono trend o numeri che derivano da misurazioni, le caratteristiche principali di una metrica sono fornire informazioni utili, essere misurata con accuratezza, riflettere il reale stato dell'elemento analizzato, supportare la gestione ed essere uno strumento di valutazione per prendere decisioni.

Un KPI, invece, è una metrica che misura quanto efficacemente si sta svolgendo un'attività, la giusta scelta del KPI consente di identificare più velocemente le problematiche e permette di migliorare le prestazioni.

Per essere efficiente un KPI deve essere SMART cioè specifico, misurabile, attendibile, realistico e basato sul tempo.

Altro requisito imprescindibile è riuscire a trasmettere un quadro della situazione a prima vista scegliendo adeguatamente la struttura, più la dashboard risulta comprensibile più sarà facile interpretare i dati.

Infine, ogni "cruscotto" dovrà essere aggiornato in real-time, il modo migliore per farlo è quello di implementare un software che in maniera automatica raccolga i dati dai data base aziendali.

I software disponibili per sviluppare un progetto di Business Intelligence sono diversi, di seguito una breve descrizione delle tre piattaforme più famose:

- Power BI (scelto per la realizzazione del progetto) con i suoi diversi applicativi permette di creare report interattivi. Si possono connettere dati provenienti da qualsiasi fonte, pulirli ed analizzarli in maniera semplice e veloce. Con questo software si possono completamente personalizzare i report. E' uno strumento molto diffuso perché semplice ed intuitivo.
- Qlik è un software in cloud, utilizzabile gratuitamente da massimo 5 utenti, il punto di forza è la possibilità di prendere decisioni in team, accelerando i processi decisionali.
La piattaforma abbraccia varie fasi del ciclo di vita dell'analytics, in particolare, sebbene Qlik Data Integration Platform sia un'offerta autonoma, mette a disposizione potenti funzionalità di integrazione e spostamento dei dati.
- Tableau è un software gratuito offerto da Tableau Software, una delle più famose società statunitensi che si occupano della creazione di software di Business Intelligence. Con Tableau Public è possibile realizzare dashboard e report interattivi, archivarli in 10GB di spazio, connettersi con altri fruitori, condividere le realizzazioni tramite i social media e/o inserirle in un sito web o un blog.

4.2 Power BI

Nel mondo dei Big Data, la tecnologia e i tool che usano sistemi di Data Visualization sono fondamentali per analizzare le informazioni, per trarre conclusioni e per prendere decisioni.

Il software scelto per l'elaborazione del progetto di Data Visualization è Power BI, prodotto fornito da Microsoft realizzato nel 2010 ed aggiornato nel 2015.

Power BI permette di connettere, modellare e visualizzare con facilità i dati creando report personalizzati.

Grazie alla connessione delle molteplici origini di dati, con la scalabilità necessaria, permette di analizzare, condividere e promuovere le informazioni dettagliate nell'intera organizzazione, mantenendo accuratezza, coerenza e sicurezza dei dati [14]. Power BI è costituito da diversi elementi (Figura 4.10) che interagiscono tra loro, consentono agli utenti di creare, condividere e sfruttare al meglio le informazioni aziendali a livello professionale:

- Power BI Desktop un'applicazione desktop Windows,
- servizio Power BI un servizio SaaS (Software as a Service) online,
- Power BI per dispositivi Windows, iOS e Android.

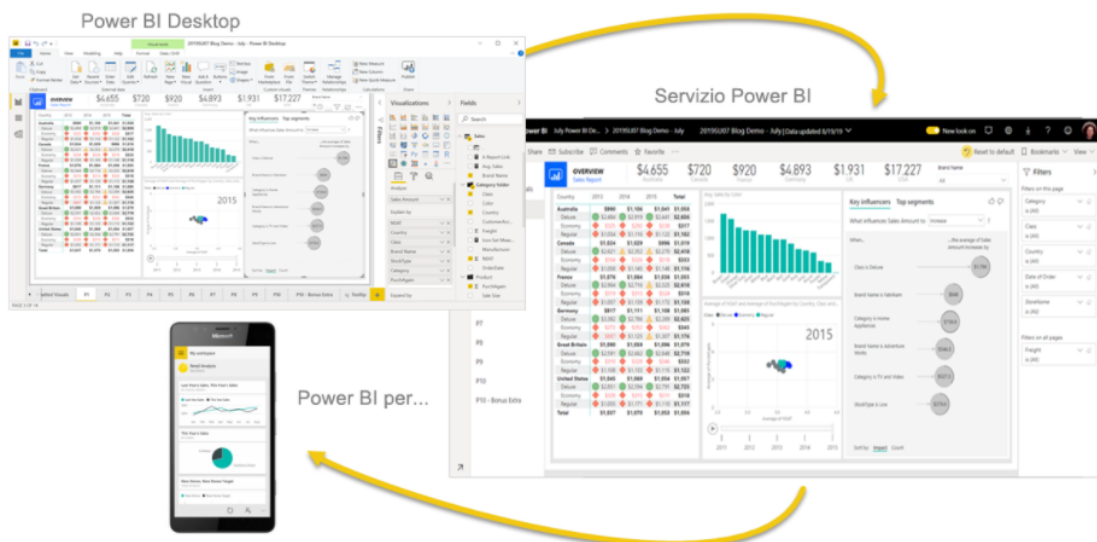


Figura 4.10: Gli elementi di Power BI [14]

Oltre a questi tre elementi sono inclusi un generatore di report per la creazione dei report stessi da condividere nel servizio Power BI e un server di report locale in cui

è possibile pubblicarli dopo averli creati in Power BI Desktop [15].

Power BI Desktop è un'applicazione gratuita, viene installata in locale e consente di connettersi ai dati.

La prima volta che si apre l'applicazione ci si troverà avanti una pagina come quella illustrata in figura 4.11, costituita da cinque aree principali [16].

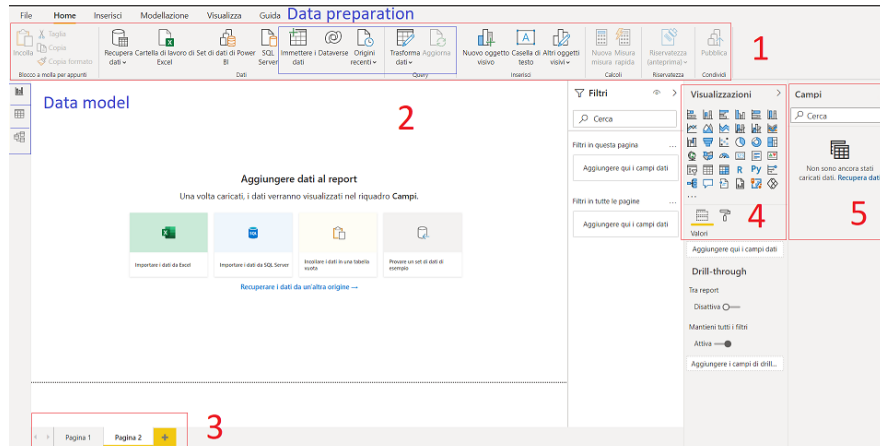


Figura 4.11: Finestra di dialogo Power BI

1. Barra multifunzione - Presenta le attività comuni associate ai report e alle visualizzazioni.
2. Visualizzazione Report o area di disegno - Posizione in cui vengono create e disposte le visualizzazioni. La visualizzazione dati consente di visualizzare tutti i dati disponibili nel report. Si tratta di un modo semplice per controllare rapidamente i tipi di dati e convalidarli. Quest'area consente di impostare visivamente la relazione tra tabelle o elementi.
3. Scheda Pagine - Situata nella parte inferiore della pagina, questa area consente di selezionare o aggiungere una pagina al report.
4. Riquadro Visualizzazioni - Area in cui è possibile modificare le visualizzazioni, personalizzare i colori o gli assi, applicare filtri, trascinare i campi e altro ancora.
5. Riquadro Campi - Area in cui gli elementi della query e i filtri possono essere trascinati nella visualizzazione Report o nell'area Filtri del riquadro Visualizzazioni.

Nella pagina iniziale si possono individuare due funzioni principali quella di data preparation in cui è possibile connettere i dati e trasformarli per creare un modello

mediante delle relazioni e quella di data model suddivisa in tre diverse aree: la prima detta di data view consente di creare oggetti visivi, la seconda di data modelling consente la visualizzazione e la trasformazione di tabelle ed attributi, l'ultima area è quella di data relationship permette di creare relazioni tra le diverse tabelle.

4.2.1 Importazione e trasformazione dei dati

I dati sono il fulcro di Power BI.

È possibile ottenere dati da una qualsiasi delle origini disponibili (Figura 4.12) nella sezione di data preparation è fondamentale fare la scelta giusta ottimizzando i tempi di aggiornamento ed evitando, quando possibile, l'importazione di dati statici.

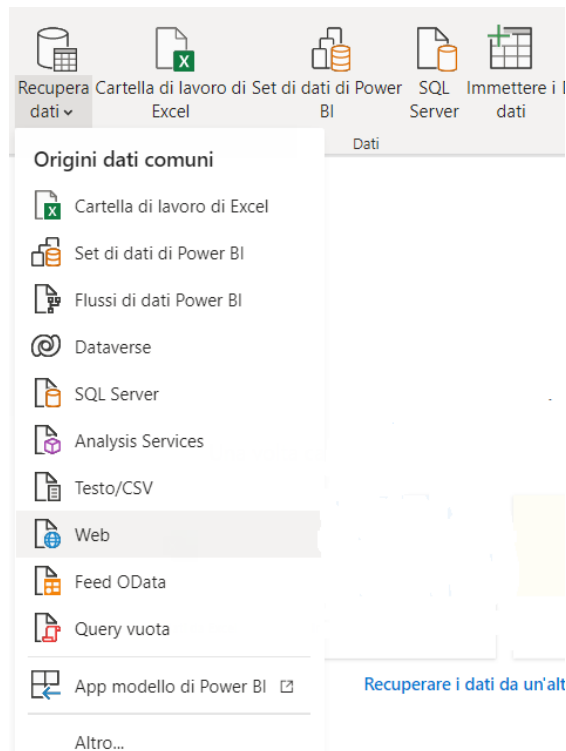


Figura 4.12: Importazione dati Power BI

Le origini dei dati sono circa 80 divise tra servizi online, file in locale o database privati.

In fase di importazione lo sviluppatore è tenuto a scegliere tra le modalità di importazione: Import mode, DirectQuery o Live Connection.

L'Import mode è l'opzione più comune, prevede l'estrazione dei dati dalla sorgente e la conseguente archiviazione direttamente nella memoria di Power BI, una memoria temporanea che prevede una metodologia di archiviazione che aumenta le prestazioni dei risultati delle query e comprime lo spazio occupato.

Il DirectQuery non prevede una vera e propria importazione dei dati ma consente l'interazione diretta tra la sorgente e le dashboard, ogni volta che i report vengono aggiornati si esegue una query nella sorgente dati primaria.

Questa tipologia però non è supportata da tutte le origini dei dati e le prestazioni delle query potrebbero essere più lente rispetto all'Import mode in quanto i tempi di attesa dipenderanno dalle prestazioni della sorgente, d'altro canto, questa modalità è preferibile quando il tasso di produzione dei dati è alto e l'aggiornamento deve avvenire in tempo reale, la sicurezza dei dati è garantita in quanto solo chi può accedere al data base originale potrà accedere alle dashboard.

Il Live Connection è supportato da poche origini di dati e, come il Direct Query, prevede una connessione diretta con la sorgente, la differenza risiede nelle prestazioni perché in quest'ultimo caso viene utilizzato il motore di archiviazione di Power BI.

Una volta completata la fase di importazione si passa alla trasformazione dei dati (Figura 4.13).

E' fondamentale che i dati siano pre-processati, filtrati e modificati.

Le funzionalità di trasformazione di Power BI sono molto intuitive ed avvengono nell' Editor di Power Query il quale converte automaticamente ogni passaggio in linguaggio 'M'.

Tutti gli step di cleaning vengono memorizzati in una pipeline che consente di tornare indietro ed avere una visione generale del processo.

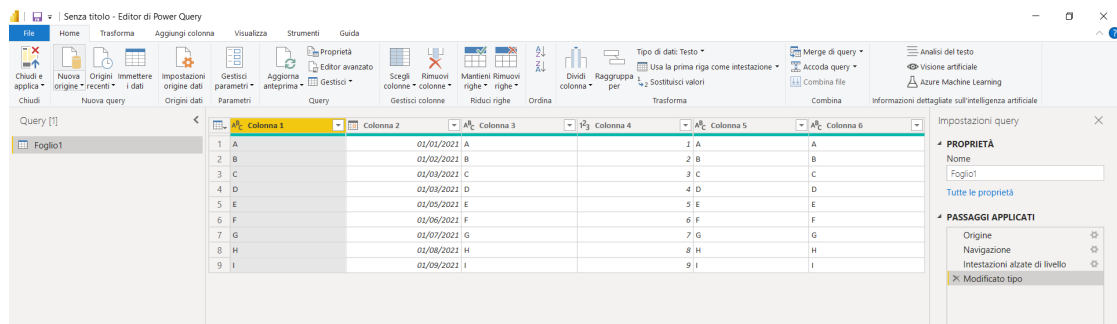


Figura 4.13: Editor di Power Query

Tra le trasformazioni più comuni si trova il cambiamento del formato delle colonne, l'aggiunta o la rimozione di colonne e il merge tra tabelle.

Una volta completate le operazioni di querying sulle tabelle, è necessario cliccare sull'icona in alto a sinistra "Close & Apply" per salvare le modifiche apportate e chiudere l'editor.

Terminata la fase di data preparation inizia quella di data modelling (Figura 4.14) che consente principalmente di stabilire delle relazioni tra le tabelle per creare un modello.

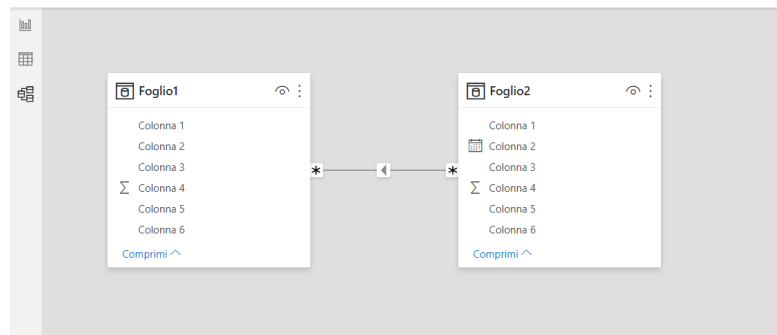


Figura 4.14: Modello Power BI

Power BI consente di visualizzare le tabelle e creare delle relazioni con una specifica direzione indicata dalla freccia ed una cardinalità (Figura 4.15).

Crea relazione



Consente di selezionare le tabelle e le colonne correlate.

Foglio1

Colonna 1	Colonna 2	Colonna 3	Colonna 4	Colonna 5	Colonna 6
A	venerdì 1 gennaio 2021	A	1	A	A
B	lunedì 1 febbraio 2021	B	2	B	B
C	lunedì 1 marzo 2021	C	3	C	C

Foglio2

Colonna 1	Colonna 2	Colonna 3	Colonna 4	Colonna 5	Colonna 6
A	venerdì 1 gennaio 2021	A	1	A	A
B	lunedì 1 febbraio 2021	B	2	B	B
C	lunedì 1 marzo 2021	C	3	C	C

Cardinalità

Uno-a-uno (1:1)

Direzione filtro incrociato

Entrambi

☒ Imposta come relazione attiva
☐ Considera integrità referenziale

OK

Annulla

Figura 4.15: Creazione relazioni

Lo sviluppatore potrebbe dover aggiungere determinate misure, colonne calcolate o tabelle calcolate, questo è possibile grazie al linguaggio DAX (Data Analysis eXpressions) il quale include molteplici funzioni di aggregazione che consentono di eseguire calcoli avanzati.

4.2.2 Creazione, pubblicazione e condivisione delle dashboard

Dopo aver elaborato i dati occorre spostarsi nella sezione di data view (Figura 4.16) nella quale è possibile progettare il report attraverso pochi passaggi:

1. Dalla sezione Campi si selezionano le tabelle e i relativi attributi che si vogliono includere nel report,
2. l'area Visualizzazioni consente di scegliere tra gli elementi visivi a disposizione (tabelle, matrici, grafici a torta o ad anello, mappe, grafici a barre, ad area o a linee ecc.), oltre ai diagrammi presenti nel programma di default, è possibile aggiungerne di nuovi da un portale nel quale gli sviluppatori sono abilitati ad aggiungere i propri.

3. l'area Formato permette di modificare le dimensioni del testo, lo stile, i titoli ecc.,
4. l'area Filtri prevede l'aggiunta di diversi filtri a livello di pagine, di drill-through e a livello di singolo report.

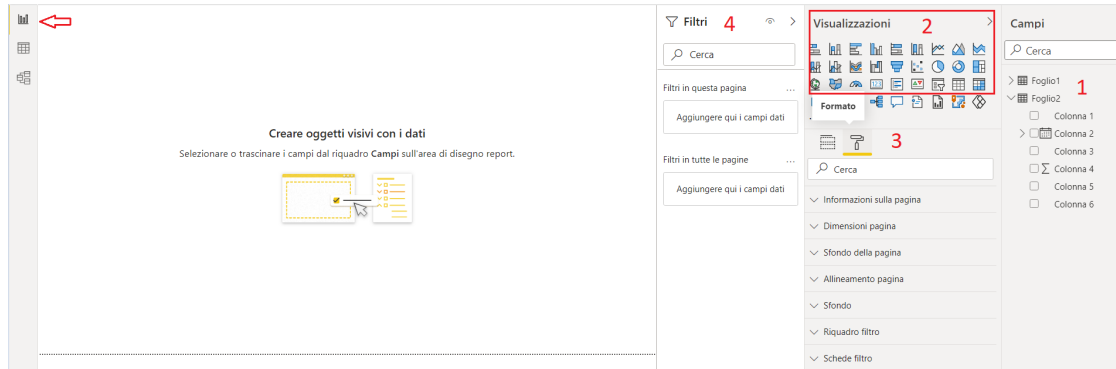


Figura 4.16: Data view

Dopo aver creato i report è possibile creare una dashboard. Una dashboard può essere intesa come una via d'accesso ai report e ai set di dati sottostanti (Figura 4.17). Selezionando una visualizzazione si accede al report usato per crearla.

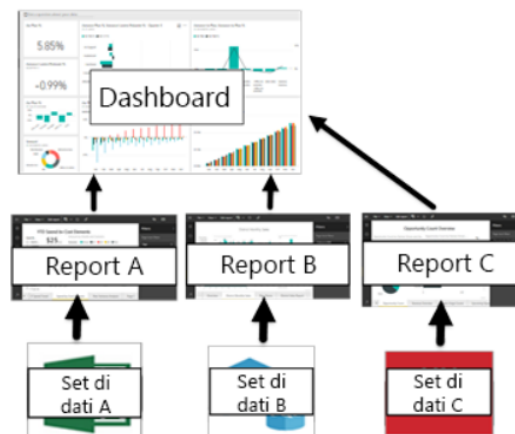


Figura 4.17: Dashboard

E' importante sottolineare quale sia la differenza tra report e dashboard: una dashboard è composta da una sola pagina, è possibile porre domande sui dati in essa contenuti, è possibile impostare degli avvisi per e-mail quando, per esempio,

vengono soddisfatte determinate condizioni e non permette di visualizzare i campi e le tabelle del set di dati sottostante.

Un report, invece, è composto da una o più pagine, l'origine dei dati è unica, permette di filtrare, evidenziare e selezionare i dati e consente di visualizzare tabelle campi e valori del set di dati [17].

Si può pubblicare un report anche partendo da Power BI Desktop e scegliendo l'area di lavoro in cui eseguire la condivisione.

Ciascuno degli elementi visivi può essere selezionato e aggiunto ad una determinata dashboard, in questa fase bisogna conoscere le priorità delle informazioni per costruire la dashboard in modo tale da far risaltare ciò che più è rilevante per l'organizzazione.

È fondamentale creare una struttura che abbia una logica di disposizione degli strumenti di visualizzazione dei dati, che devono essere facilmente navigabili.

La possibilità di controllare la modalità di condivisione del lavoro è una delle funzionalità più importanti del servizio Power BI.

È possibile creare aree di lavoro per collaborare con i colleghi, è quindi possibile aggregare i report e distribuirli come App.

Una volta completata la dashboard, si possono porre domande sui dati.

L'ultimo step è quello della condivisione, ciò può essere fatto semplicemente inserendo l'indirizzo e-mail dell'utente con cui si vuole condividere le informazioni, è possibile scegliere anche la modalità di condivisione: in modifica o solo in lettura dei dati.

Capitolo 5

Caso studio

5.1 Descrizione

L'intero percorso in azienda è stato incentrato sui dati. Il ruolo di Junior ERP Engineer ha consentito di condurre una profonda analisi sui dati aziendali a partire da come questi vengono storicizzati sino alla presentazione agli utenti.

Oltre all'ERP utilizzato per la gestione quotidiana, l'azienda utilizza una serie di applicativi esterni per creare maschere utili per il business, tra questi troviamo Microsoft Access.

Microsoft Access è un'interfaccia software per la gestione di basi di dati di tipo relazionale. Può utilizzare dati immagazzinati in formato Access/Jet, SQL Server, Oracle o qualsiasi database in formato compatibile ODBC.

L'elemento principale di Access sono le maschere ovvero uno strumento grafico utile alle interazioni dell'utente con i dati della tabella. Possono essere sviluppati progetti di varie dimensioni. Per applicazioni semplici è possibile utilizzare template già presenti nel pacchetto, per esigenze di sviluppo più evolute è disponibile il linguaggio di programmazione Microsoft Visual Basic [18].

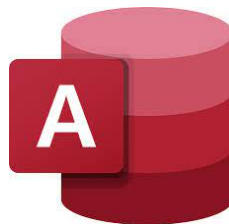


Figura 5.1: Microsoft Access

Il progetto intrapreso prevede l'analisi delle maschere Access e la successiva implementazione di nuove Dashboard in Power BI.

In particolare, è stata scelta una maschera Access, sono state analizzate tutte le informazioni che questa presenta a livello tabulare, la maschera è stata scomposta, è stata ricercata l'origine di tutti i dati presentati e, partendo dalle sorgenti, sono state progettate le dashboard su Power BI.

Oltre a presentare in maniera diversa le informazioni già esistenti, spinti dalla necessità dei diversi attori di conoscere ulteriori metriche e KPI, è stata condotta un'analisi mirata ad introdurre nuovi dati nei report.

5.1.1 Analisi AS IS

Il progetto si basa sull'analisi degli Ordini dell'azienda e sulle giacenze di magazzino. Il report già esistente prevede la presentazione dei dati in formato tabulare, questa è una grande limitazione nella presentazione dei dati.

Mostrare i dati sotto forma di tabelle è un rischio in quanto, le analisi effettuate potrebbero perdere valore perché non comprese adeguatamente da chi dovrà leggerle e prendere decisioni.

E' necessario scegliere la visualizzazione corretta per evidenziare gli aspetti più importanti dei dati, in questo modo è possibile illuminare nuove intuizioni e comunicarle in modo più persuasivo. Ciò può comportare azioni più intelligenti e risultati migliori per l'azienda.

Per essere utile l'interfaccia che mostra i dati all'utente finale deve essere personalizzabile, interattiva ed evitare la confusione.

La maschera Access attualmente disponibile per la visualizzazione dei dati non possiede le caratteristiche sopra elencate, per questo è necessaria una revisione sia grafica sia progettuale delle informazioni.

Lo strumento ad oggi utilizzato (Figura 5.2) è composto da diverse sezioni verticali:

- l'area sinistra presenta l'analisi del fatturato, in particolare, vengono presentati i dati relativi all'ordinato attuale, l'ordinato del mese in corso, il fatturato del mese e il valore degli ordini spediti ancora da fatturare. Può essere consultata la situazione dell'anno che include fatturato, backlog e totale. Segue un'analisi del Backlog con dettaglio mensile per l'anno in corso e con dettaglio annuale per gli anni successivi.
- l'area centrale presenta numerosi pulsanti che, se cliccati, provocano l'apertura di altre maschere. I dati consultabili sono il riepilogo dell'ordinato per mese, per anno, per venditore, per prodotti. In questa area è presente anche la situazione del magazzino.
- l'ultima area fornisce i dati riguardanti il fatturato per cliente e mostra una tabella dettagliata degli ordini di acquisto specificando gli ordini di acquisto scaduti cioè non arrivati in tempo e gli ordini da ricevere con dettaglio mensile.

Figura 5.2: Microsoft Access: maschera ordini

Cliccando su alcuni dati, come per esempio sul backlog del mese in corso, la maschera consente di consultare informazioni più dettagliate, effettua quindi un drill-through dei dati che vengono presentati in formato tabulare, nel caso in esempio, vengono proposti tutti i prodotti con le relative informazioni che devono essere ancora spediti.

5.1.2 Analisi TO BE

Il progetto deve consentire che tutti i dati, output dell'analisi compiuta, siano visualizzabili ed interpretabili facilmente.

Per prima cosa i dati devono essere estratti ed aggiornati in tempo reale, l'interfaccia utente deve essere personalizzabile e soprattutto, interattiva. L'utente deve poter navigare nelle finestre e visualizzare i dati partendo dal livello di sintesi aggiungendo o eliminando filtri e selettori che consentono l'esplorazione e l'analisi di alcuni dettagli, senza mai perdere la visione d'insieme; i dati complessi devono essere comprensibili e accessibili per gli utenti non tecnici.

Il risultato finale deve essere condiviso per poter promuovere la collaborazione.

Lo scopo è quello di descrivere e far comprendere l'andamento dei processi aziendali nel passato e nel presente, oltre a generare previsioni sugli eventi futuri.

L'obiettivo finale è facilitare le decisioni sulle prossime azioni da intraprendere.

A livello operativo, le dashboard misurano i processi aziendali in dettaglio, mentre a livello di management forniscono un quadro sintetico della performance globale dell'azienda.

Prima di sviluppare il progetto, è stato fondamentale intervistare gli attori per comprendere quali fossero le metriche, i KPI e le informazioni principali da visualizzare.

Al termine delle interviste, sono stati individuati tre macro argomenti diversi da affrontare: gli ordini cliente, la giacenza di magazzino e gli ordini d'acquisto.

- Ordini cliente.

Gli ordini cliente, intesi come ordini di vendita, contengono numerose informazioni fondamentali per monitorare l'andamento aziendale.

Tra le prime metriche che si desidera visualizzare troviamo l'ordinato, ovvero il volume d'affari ed il fatturato annuali e mensili. Tali valori sono indicatori economici che misurano l'ammontare complessivo dei ricavi registrati da un'impresa e derivanti dalla vendita dei beni prodotti e/o da prestazione di servizi.

In genere, il fatturato viene assunto come indicatore del giro d'affari, quindi della consistenza economica dell'impresa.

L'ordinato aziendale viene continuamente rapportato al totale degli ordini spediti ovvero la quantità, in termini economici, di ordini che sono stati già evasi, è importante tenere i due valori allineati e non ritardare le spedizioni. Negli ordini si trovano le informazioni relative alle tipologie di prodotti e servizi venduti dall'azienda, queste permettono di condurre un'analisi del business aziendale, prendere decisioni e monitorare l'allineamento alle strategie scelte dall'azienda. Per esempio, è possibile capire a quanto ammonta, in percentuale, il valore della merce prodotta internamente dall'azienda rispetto alla rivendita dei beni.

Analizzare tutte le potenzialità del proprio mercato di riferimento è fondamentale per la creazione del business plan aziendale e quindi per la programmazione delle strategie e delle attività dell'impresa.

E' possibile analizzare i diversi mercati in cui si lavora ed individuare i settori emergenti o quelli in cui il business è meno interessante; lo studio approfondito relativo alla domanda generica dei consumatori è in grado di suggerire all'azienda quali siano i possibili settori nei quali investire le proprie risorse. Inoltre, si è deciso di analizzare le performances dei Commerciali ovvero di coloro che hanno il compito di occuparsi della gestione degli ordini, nonché delle relazioni con la clientela, con la quale hanno rapporti di assistenza durante la trattativa commerciale.

A partire dagli ordini è possibile individuare i principali clienti dell'azienda, tale l'analisi è uno strumento fondamentale per valutare l'esperienza d'acquisto dei propri clienti ed elaborare strategie di marketing finalizzate alla loro fidelizzazione.

Come ultima analisi è stato ritenuto utile approfondire il portafoglio clienti in termini di nuovi clienti e provenienza geografica.

- Magazzino.

La gestione del magazzino riveste un ruolo fondamentale in azienda, garantisce

vantaggi a vari livelli: gestione della produzione, approvvigionamento, logistica e acquisti, influenzando anche i costi di produzione.

La contabilità gestionale di magazzino è uno degli aspetti che in un'industria manifatturiera deve essere costantemente monitorato.

Da un'attenta gestione del magazzino possono derivare vantaggi a livello aziendale come l'automazione di numerosi processi e la possibilità di ottenere informazioni utili per il management e per l'intera azienda, ciò può portare le aziende a diventare più efficienti e più reattive nel controllo dei cambiamenti di business.

Per questi motivi, si è deciso di tenere traccia delle giacenze di magazzino sulla base della tipologia dei prodotti e di monitorare la variazione delle giacenze rispetto al mese precedente.

- Ordini d'acquisto.

La gestione degli ordini d'acquisto è strettamente legata a quella del magazzino.

Lo scopo è gestire in modo ottimizzato gli approvvigionamenti di beni, semilavorati e servizi necessari per supportare l'attività produttiva e, più in generale, il business dell'azienda in modo efficace, riducendo al minimo i costi.

E' fondamentale tenere sotto controllo gli ordini ed in particolar modo gli arrivi, per questo ne è richiesto il monitoraggio.

Bisogna tenere traccia degli ordini che sono in ritardo per capire come questi possano impattare sulla produzione e di conseguenza sugli ordini di vendita.

A partire da tali ordini è possibile individuare i principali fornitori e la loro affidabilità in termini di ritardi nella consegna.

5.2 Dati

Le informazioni raccolte in un database aziendale sono innumerevoli, per questo, quando bisogna fare un'analisi si devono conoscere i 'desiderata' degli utenti finali. La prima operazione compiuta, dopo aver scelto la maschera da riportare su Power BI e dopo aver compreso le necessità degli utenti, è stata risalire all'origine dei dati.

Navigando la maschera Access in modalità "Visualizzazione Struttura" e cliccando su ognuna delle varie sotto-maschere che la compongono, si apre la "Finestra delle proprietà" che specifica il Formato, i Dati e l'Evento che la caratterizzano.

Dalla voce "Origine record" si può risalire alla query che consente di estrarre i dati presentati, è necessario analizzare le query fino ad intercettare la fonte originale dei dati ovvero l'archivio del database aziendale nel quale sono memorizzati i dati.

Le tabelle interessate, nel nostro caso, sono diverse perché, come detto precedentemente, la maschera oggetto del progetto presenta dati di diversa natura: ordini clienti, magazzino e ordini d'acquisto.

Per la protezione dei dati aziendali, i nomi delle tabelle e degli attributi riportati di seguito non sono quelli reali, le strutture sono state ricreate per la presentazione dei dati.

La figura 5.3 mostra il class diagram complessivo del progetto.

Il centro dell'intero progetto è la tabella PRODOTTO nella quale sono inserite le informazioni di tutti gli articoli dell'azienda (Articoli di acquisto, di vendita, semilavorati ecc.).

La parte sinistra dell'immagine riporta tutte le tabelle necessarie per la gestione degli ordini clienti: ogni ordine cliente (Tabella ORDINE-CLIENTE) è effettuato da un cliente (Tabella CLIENTE) e gestito da un Commerciale (Tabella COMMERCIALE) il quale si occupa del rapporto con il cliente; l'ordine è composto da una o più righe (Tabella RIGHE-ORDINE-CLIENTE) ed a sua volta ogni riga afferisce ad un unico prodotto di vendita (Tabella PRODOTTO).

La vista in basso nell'immagine (Tabella GIACENZA-MAGAZZINO) archivia tutte le transazioni del magazzino con lo scopo di calcolare la giacenza, ogni transazione è chiaramente collegata ad un unico prodotto (Tabella PRODOTTO).

La parte destra dell'immagine descrive la gestione degli ordini di acquisto: un ordine di acquisto è riferito ad un fornitore (Tabella FORNITORE) ed è composto da una o più righe (Tabella RIGHE-ORDINE-ACQUISTO), ogni riga corrisponde ad un determinato prodotto che si vuole acquistare (Tabella PRODOTTO).

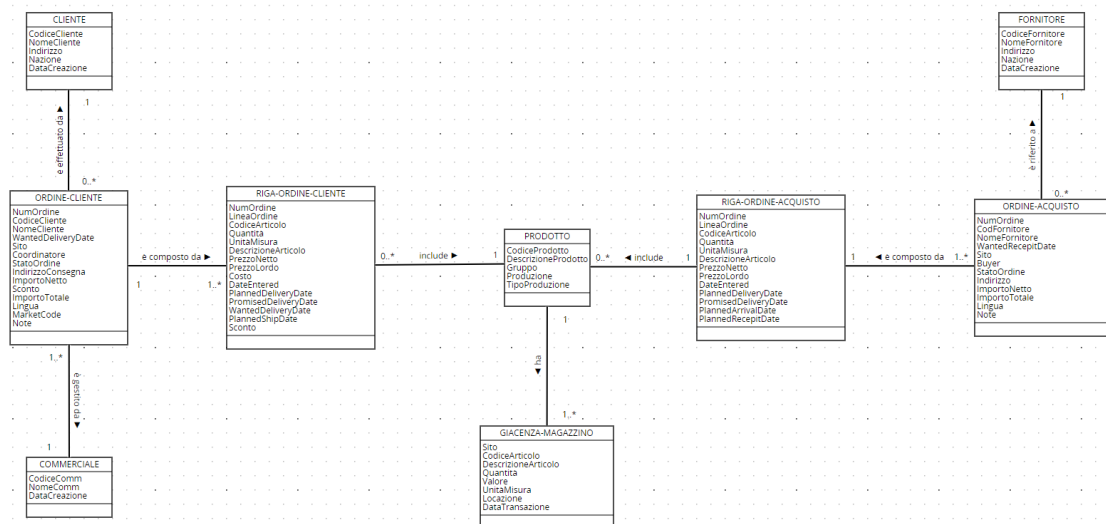


Figura 5.3: Class Diagram

Le tabelle PRODOTTO, CLIENTE, COMMERCIALE, FORNITORE sono poco dinamiche, memorizzano dati che raramente cambiano nel tempo, sono spesso modificate per il solo inserimento di nuovi record. Le rimanenti tabelle invece, sono più complesse, per questo motivo è utile analizzarle nel dettaglio. La prima tabella analizzata (Figura 5.4) è quella degli ordini cliente che chiameremo "ORDINE-CLIENTE".

ORDINE-CLIENTE
NumOrdine
CodiceCliente
NomeCliente
WantedDeliveryDate
Sito
Coordinatore
StatoOrdine
IndirizzoConsegna
ImportoNetto
Sconto
ImportoTotale
Lingua
MarketCode
Note

Figura 5.4: Tabella Ordine Cliente

Gli attributi della in questione sono diversi, analizziamo i principali:

- NumOrdine il numero identificativo assegnato all'ordine,
- CodiceCliente l'identificativo assegnato al cliente che effettua l'ordine,
- WantedDeliverydate la data in cui il cliente desidera che l'ordine venga consegnato all'indirizzo di consegna specificato,
- Coordinatore l'ID della persona solitamente responsabile del coordinamento della gestione delle offerte di vendita e degli ordini dei clienti,
- StatoOrdine lo stato dell'ordine cliente indica cosa è stato fatto e quale dovrebbe essere l'attività successiva. Sono disponibili diverse operazioni a seconda dello stato. Man mano che l'ordine si sposta nel flusso dell'ordine, lo stato cambia automaticamente. Lo stato dell'ordine può essere: Pianificato, Rilasciato, Prenotato, Prelevato, Consegnato parzialmente, Consegnato, Annullato, Fatturato/Chiuso o Bloccato.

- **ImportoNetto** l'importo netto dell'ordine, espresso nella valuta di base dell'azienda. Viene calcolato come la somma degli importi netti nella valuta di base per tutte le righe dell'ordine,
- **Sconto** lo sconto aggiuntivo totale per l'intero ordine,
- **ImportoTotale** l'importo lordo totale dell'ordine, espresso nella valuta dell'ordine. Viene calcolato come somma degli importi lordi delle righe. Questo importo include gli importi delle imposte,
- **MarketCode** indica il mercato di riferimento dell'ordine per esempio Medicaale, Difesa, Ricerca ecc..

In riferimento agli ordini abbiamo un altro archivio importante, quello delle Righe Ordine Cliente (Figura 5.5) chiamato "RIGA-ORDINE-CLIENTE", questa, riporta alcune delle informazioni principali dalla tabella degli ordini ed aggiunge dati relativi alle singole righe.

RIGA-ORDINE-CLIENTE
NumOrdine
LineaOrdine
CodiceArticolo
Quantita
UnitaMisura
DescrizioneArticolo
PrezzoNetto
PrezzoLordo
Costo
DateEntered
PlannedDeliveryDate
PromisedDeliveryDate
WantedDeliveryDate
PlannedShipDate
Sconto

Figura 5.5: Tabella RigA Ordine Cliente

Gli attributi principali sono:

- **NumOrdine** il numero identificativo assegnato all'ordine,
- **LineaOrdine** l'ID della riga dell'ordine creato automaticamente al momento della registrazione della riga dell'ordine,
- **CodiceArticolo** l'identificativo della parte di vendita che il cliente desidera acquistare,

- Quantità indica la quantità desiderata dal cliente espressa in una specifica unità di misura,
- PrezzoNetto il prezzo unitario dell'articolo recuperato da un contratto o da un listino prezzi,
- PrezzoLordo il prezzo che include tasse espresso nella valuta dell'ordine,
- Sconto lo sconto riferito alla singola riga,
- Costo il costo unitario dell'articolo,
- DateEntered la data in cui il record è stato creato,
- PlannedDeliveryDate la data in cui è prevista la consegna della merce all'indirizzo del cliente. Tale data viene comunicata al cliente in fase di stampa della conferma d'ordine,
- PromisedDeliveryDate la data in cui la merce è promessa per essere consegnata all'indirizzo del cliente. Inizialmente corrisponde alla data di consegna pianificata, questa cambierà quando la data di consegna pianificata viene aggiornata,
- WantedDeliveryDate la data in cui il cliente desidera che l'ordine venga consegnato all'indirizzo di consegna specificato. Questa data viene recuperata dall'intestazione dell'ordine del cliente,
- PlannedShipDate questa è la data della prevista partenza dalla sede del venditore ed è calcolata come il giorno successivo all'ultimo giorno di ritiro; calcolata al fine di rispettare la data di consegna promessa.

La seconda sezione riguarda la gestione del magazzino, gli archivi coinvolti sono molteplici in quanto gli articoli possono essere ricevuti tramite gli acquisti, spostati attraverso delle transazioni interne, trasferiti e ricevuti dal reparto Produzione ecc., ogni movimento è registrato in una tabella diversa. Per facilitare la gestione è stata creata una vista che raccoglie tutti i tipi di transazioni e calcola le giacenze di magazzino. Un altro aspetto importante che tale vista tiene in considerazione è la valutazione del magazzino che può essere fatta in diversi modi, quello utilizzato in questo caso è il FIFO (First In First Out). Chiameremo questa vista "GIACENZA-MAGAZZINO" (Figura 5.6), gli attributi determinanti sono:

GIACENZA-MAGAZZINO	
Sito	
CodiceArticolo	
DescrizioneArticolo	
GruppoArticolo	
Quantita	
Valore	
UnitaMisura	
Locazione	
DataTransazione	

Figura 5.6: Tabella Giacenza Magazzino

- CodiceArticolo l'identificativo dell'articolo,
- GruppoArticolo ovvero la tipologia di appartenenza dell'articolo, le diverse opzioni sono PP purchased parts , RM raw material, SA sub assembled, FG finished goods,
- Quantità dell'articolo in giacenza,
- Valore riferito all'articolo valorizzato FIFO,
- Locazione ovvero il codice del magazzino in cui l'articolo è presente,
- DataTransazione la data dell'ultima transazione importante per tracciare le variazioni del magazzino.

La terza ed ultima parte riguarda gli ordini di acquisto, anche in questo caso, come quello degli ordini cliente, abbiamo una tabella della testata degli ordini di acquisto ed una delle righe ordine di acquisto, chiameremo le tabelle rispettivamente "ORDINE-ACQUISTO" e "RIGA-ORDINE-ACQUISTO". Analizziamo gli attributi della tabella Ordine di Acquisto (Figura 5.7):

ORDINE-ACQUISTO
NumOrdine
CodFornitore
NomeFornitore
WantedRecepitDate
Sito
Buyer
StatoOrdine
Indirizzo
ImportoNetto
ImportoTotale
Lingua
Note

Figura 5.7: Tabella Ordine di Acquisto

- NumOrdine il numero identificativo assegnato all'ordine,
- CodFornitore l'identificativo assegnato al fornitore,
- WantedRecepitDate la data di entrta in giacenza desiderata,
- Buyer l'ID della persona responsabile della gestione degli acquisti,
- StatoOrdine Lo stato corrente dell'ordine di acquisto. Lo stato può essere il seguente:
 - Pianificato:** assegnato automaticamente al salvataggio dell'intestazione dell'ordine di acquisto.
 - Arrestato:** l'ordine di acquisto è temporaneamente bloccato per un'ulteriore gestione nel flusso dell'ordine.
 - Rilasciato:** l'ordine di acquisto è pronto per l'ulteriore elaborazione nel flusso dell'ordine.
 - Confermato:** il fornitore ha approvato la consegna per le date e le quantità desiderate sull'ordine di acquisto, oppure è stata inviata una nota di modifiche e l'ordine di acquisto è stato confermato con alcune differenze.
 - Arrivato:** tutte le parti ordinate o la quantità parziale sono arrivate in sede, ma devono essere (ispezionate e) ricevute in magazzino.
 - Ricevuto:** tutte le parti ordinate o la quantità parziale sono arrivate e sono state (ispezionate e) ricevute in magazzino.
 - Chiuso:** tutte le parti ordinate sono state ricevute in magazzino o le righe dell'ordine di acquisto rimanenti sono state chiuse o annullate manualmente.
 - Annullato:** l'ordine di acquisto non è più valido.

- **ImportoNetto** l'importo netto dell'ordine, espresso nella valuta di base dell'azienda. Viene calcolato come la somma degli importi netti nella valuta di base per tutte le righe dell'ordine,
- **ImportoTotale** l'importo lordo totale dell'ordine, espresso nella valuta dell'ordine. Viene calcolato come somma degli importi lordi delle righe. Questo importo include gli importi delle imposte.

L'ultima tabella analizzata è "RIGA-ORDINE-ACQUISTO" (Figura 5.8) delle righe ordine di acquisto composta principalmente dai seguenti attributi:

RIGA-ORDINE-ACQUISTO
NumOrdine
LineaOrdine
CodiceArticolo
Quantita
UnitaMisura
DescrizioneArticolo
PrezzoNetto
PrezzoLordo
DateEntered
PlannedDeliveryDate
PromisedDeliveryDate
PlannedArrivalDate
PlannedReceiptDate

Figura 5.8: Tabella Riga Ordine di Acquisto

- **NumOrdine** il numero identificativo assegnato all'ordine,
- **LineaOrdine** l'ID della riga dell'ordine creato automaticamente al momento della registrazione della riga dell'ordine,
- **CodiceArticolo** l'identificativo della parte di acquisto che si desidera acquistare,
- **Quantità** indica la quantità desiderata dall'azienda espressa in una specifica unità di misura,
- **PrezzoNetto** l'importo totale dell'ordine, espresso nella valuta di base dell'azienda. Viene calcolato come la somma degli importi netti per tutte le righe d'ordine. Nell'importo vengono considerati eventuali sconti e costi aggiuntivi derivanti da un contratto, o da qualsiasi listino prezzi, prezzo standard, gruppo di acquisto per fornitore,
- **PrezzoLordo** l'importo lordo totale dell'ordine. Viene calcolato come somma degli importi lordi di tutte le righe dell'ordine nell'ordine specifico. Questo importo include gli importi delle imposte di tutte le righe ordine.

- DateEntered la data in cui il record è stato creato,
- PlannedDeliveryDate la data in cui è prevista la consegna delle merci all'indirizzo del destinatario, come specificato nella riga dell'ordine di acquisto. La formula per calcolare la data di consegna pianificata è la seguente: $\text{Data di consegna pianificata} = (\text{Data di arrivo pianificata} - \text{Tempo di consegna del trasporto})$,
- PromisedDeliveryDate la data in cui il fornitore ha accettato o confermato di consegnare in un luogo concordato. Inizialmente corrisponde alla data di consegna desiderata,
- PlannedArrivalDate la data in cui si prevede che gli articoli arrivino al luogo di arrivo del sito. La data di arrivo pianificata viene calcolata come $\text{Data di arrivo pianificata} = (\text{Data di ricevimento pianificata} - \text{Tempo di ispezione interna} - \text{Tempo di consegna del trasporto interno})$,
- PlannedRecepitDate questa è la data in cui si prevede che gli articoli vengano ispezionati e siano pronti presso l'ubicazione di magazzino.

Oltre alle tabelle sopra descritte per lo sviluppo del progetto sono state utilizzate altre tabelle standard come, per esempio, quelle che contengono l'anagrafica dei Clienti e dei Fornitori e la tabella degli Articoli d'acquisto e di vendita.

Una volta presa conoscenza delle caratteristiche del data set, al fine che ne venissero identificate le peculiarità e le correlazioni principali, si è passati alla fase relativa alle trasformazioni dei dati a disposizione.

Trasferire archivi interi all'interno di Power BI è dispendioso in quanto i dati che ci occorrono sono molto meno dettagliati rispetto a quelli sopra descritti, è conveniente aggregare le informazioni per arrivare ai desiderata. Per raggiungere tale obiettivo, è utile creare un modello concettuale per ogni diversa area, dedurre da quest'ultimo un modello logico relazionale e progettare delle viste che consentano di avere i dati pronti per costruire facilmente i report desiderati.

Il primo Fatto creato (Figura 5.9) è quello degli ordini cliente:

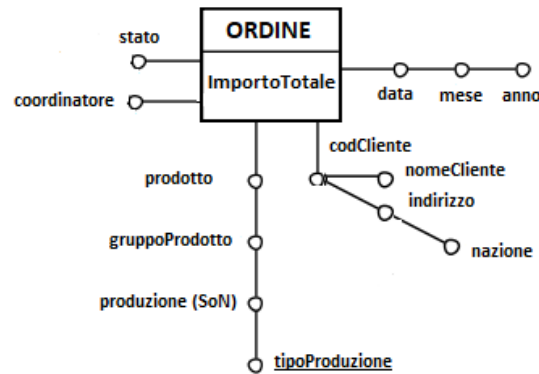


Figura 5.9: Dimensional Fact Model: Ordine

La misura individuata è l'importo totale, le gerarchie invece sono il Tempo, il Cliente, il Prodotto, lo Stato.

Dal modello concettuale si può dedurre al modello logico relazionale, le chiavi primarie delle tabelle sono sottolineate.

Tempo (Data, mese, anno)

Cliente (CodCliente, nomeCliente, indirizzo, nazione)

Prodotto (CodProdotto, gruppoProdotto, produzione, tipoProduzione)

Ordine (Data, CodCliente, CodProdotto, Stato, Coordinatore, importoTotale)

Notiamo due attributi particolari della gerarchia Prodotto, il primo è "produzione" il quale è un attributo booleano e indica se il prodotto a cui si fa riferimento è da produrre oppure no; il secondo è "tipoProduzione" il quale è un attributo opzionale.

Il secondo modello concettuale creato gestisce gli Ordini di Acquisto (Figura 5.10):

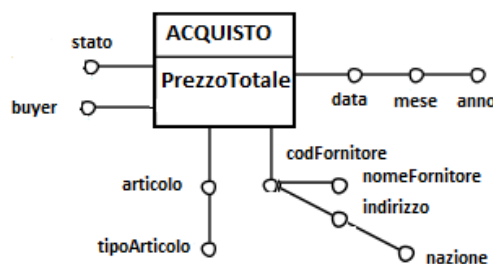


Figura 5.10: Dimensional Fact Model: Acquisto

La misura è il prezzo totale, le gerarchie sono il Tempo, il Fornitore, l'Articolo, lo Stato e il Buyer.

Dal modello concettuale si deriva il modello logico relazionale, le chiavi primarie delle tabelle sono sottolineate.

Tempo (Data, mese, anno)

Fornitore (cidFornitore, nomeFornitore, indirizzo, nazione)

Prodotto (Articolo, tipoArticolo)

Ordine (Data, CodFornitore, articolo, Stato, buyer, prezzoTotale)

Per quanto riguarda, invece, i dati delle giacenze di magazzino non è stato necessario costruire il modello concettuale ed il modello logico relazionale perché i dati erano già stati estratti correttamente attraverso la vista logica creata per ottenere le informazioni senza ridondanze.

Capitolo 6

Sviluppo del progetto

6.1 Creazione report

L'approccio utilizzato per lo sviluppo del progetto è stato quello Agile, il quale ha permesso uno sviluppo incrementale ed interattivo, caratterizzato da cicli di durata limitata e rilasci incrementali.

La metodologia Agile permette di ottimizzare e migliorare il processo di sviluppo nel tentativo di diminuire il rischio di mancati obiettivi.

Questa metodologia permette di fornire un prodotto migliore, in modo più rapido, attraverso sessioni brevi e interattive.

Agile permette di rispondere rapidamente al cambiamento, interagire con i diversi stakeholders, prioritizzare rapidamente le attività più importanti, rilevare velocemente problemi e difetti e consegnare in maniera evolutiva parti del progetto funzionanti.

Ciò ha permesso di fornire un prodotto migliore, in modo più rapido rispetto all'approccio tradizionale attraverso sessioni brevi ed interattive.

Gli obiettivi sono stati definiti nel tempo e monitorati quotidianamente, questo ha consentito una gestione trasparente e flessibile ed è stato fondamentale per il coinvolgimento di tutti gli stakeholders.

Per prima cosa, è stata rivista la connessione alle fonti di origine ed il caricamento su Power BI Desktop.

Essendo i dati in questione importati da una fonte sicura e coerente, non è stato necessario effettuare le operazioni di pulizia, correzione dei dati, rilevazione di incoerenze e lacune, al contrario, è stato necessario controllare che le informazioni fossero recepite correttamente da Power BI Desktop. (Figura 6.1)

Ovvero, al momento dell'inserimento, tutti gli attributi dovevano appartenere della tipologia corretta: le date di tipo "Data/Ora", i numeri di tipo "Numero decimale" o "Numero intero" come mostrato nella figura seguente.

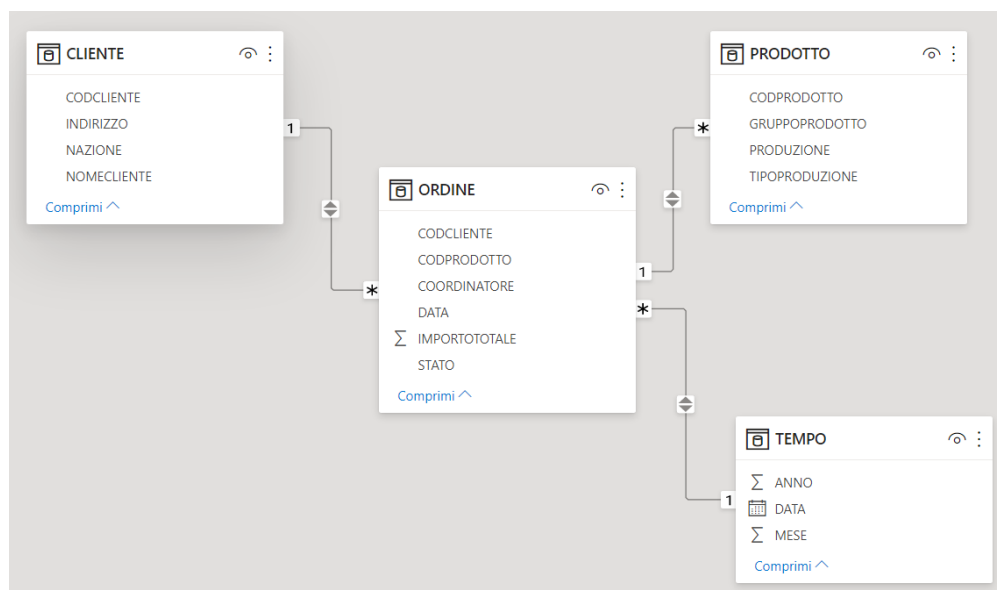
A ^B ORDINE	DATA_ORDINE	A ^B AGENTE	1.2 IMPORTO_TOT_EURO	A ^B STATO
08/01/20			0 Released	

Figura 6.1: Esempio Tipi di Dato

Prima di procedere alla progettazione dei report si è passati alla sezione di data model quindi alla creazione delle relazioni tra le diverse tabelle.

Questa fase è stata decisiva in quanto i diversi aggregati di dati risultavano strettamente dipendenti l'uno dall'altro.

Di seguito, in figura 6.2, il modello relazionale riferito agli ordini Cliente.

**Figura 6.2:** Modello Relazionale: Ordine Cliente

Terminate le fasi inerenti all'elaborazione e alla gestione dei dati, si è passati alla parte pertinente alla reportistica.

La finalità di tale step è la rappresentazione per via grafica dei dati e delle relative relazioni, al fine di generarne una conoscenza accessibile e fruibile.

Intento principale, del resto, era la realizzazione di uno strumento di supporto decisionale, che potesse conferire valore aggiunto alle scelte e all'interpretazione dei diversi contesti e ambiti coinvolti.

Con lo scopo di ottimizzare l'esperienza visiva e facilitarne la comprensione, sono state realizzate differenti pagine, ognuna inerente ad una specifica tematica e con propri e dettagliati report.

La prima sezione realizzata è stata quella relativa agli ordini cliente.

Uno dei dati principali che si desidera visualizzare è l'ordinato totale dell'anno

in corso, tale dato è assoluto, per questo si è scelto di inserirlo nel report come numero, in Power BI questa rappresentazione prende il nome di Scheda (Figura 6.3).

A tale valore è stato affiancato un grafico a linee (Figura 6.4) che presenta il dettaglio dell'ordinato per mese.

47,29Mln

ORDINATO

Figura 6.3: Totale ordinato dell'anno in corso

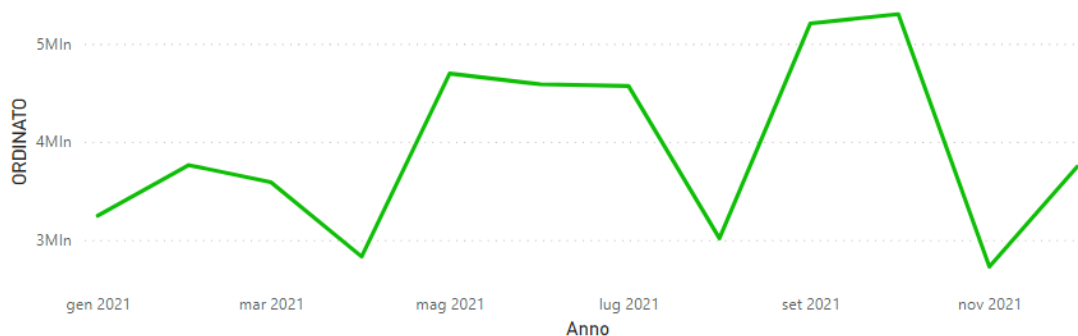
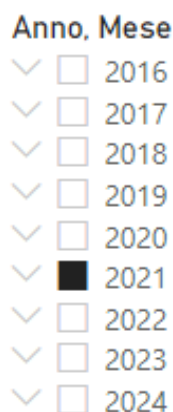


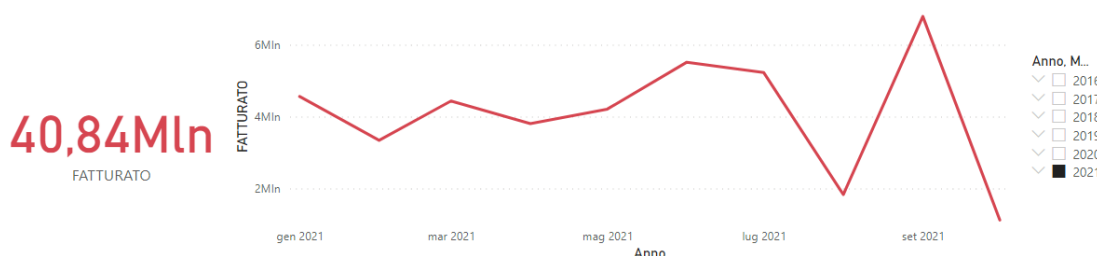
Figura 6.4: Dettaglio mensile dell'ordinato dell'anno in corso

Le precedenti rappresentazioni sono accompagnate da un filtro dei dati (Figura 6.5) che consente di selezionare un anno diverso da quello in corso, in questo modo si facilita il confronto con gli anni precedenti e si può consultare anche l'ordinato previsto per gli anni futuri.

Inoltre, possono essere confrontati gli andamenti annuali in modo da rilevare stagionalità o particolari trend.

**Figura 6.5:** Filtro dei dati

Il secondo valore, fondamentale per le analisi della solidità aziendale, è il fatturato, esattamente come per il valore precedente è stata utilizzata una scheda per mostrare il valore assoluto dell'anno in corso ed un grafico a linee per rappresentare il dettaglio mensile (Figura 6.6), il filtro consente di navigare tra i vari anni.

**Figura 6.6:** Analisi fatturato dell'anno in corso

Seguono altri due valori strettamente legati, l'ordinato Year to Date cioè il totale dell'ordinato dall'inizio dell'anno fino al giorno in cui viene visualizzato il report ed il Rolling dell'anno in corso,¹ tale valore è stato calcolato sommando l'ordinato YtD all'ordinato dell'anno precedente relativo al periodo che va dalla

¹Il rolling forecast è un dato che effettua delle previsioni sul futuro per un periodo di tempo continuo sulla base di dati storici. Diversamente dalla previsione statica che prevede i numeri sulla base di un unico periodo, come da gennaio a dicembre, il rolling forecast lascia alle spalle un mese non appena questo giunge a termine, prevedendo automaticamente il mese successivo. In questo modo, si hanno previsioni sempre volte al futuro sulla base del periodo di tempo e dei dati più recenti. Le previsioni rolling forniscono le informazioni più accurate e aggiornate, consentendo alle aziende di modificare piani o strategie al fine di ottimizzare le loro prestazioni.

data dell'estrazione alla fine dell'anno, in questo modo si ottiene una previsione. E' importante porre vicini questi valori perché se confrontati forniscono un'indicazione sul possibile andamento dell'anno (Figura 6.7).

Totale ordinato ytd	Totale Rolling 2021
40,81Mln	46,63Mln

Figura 6.7: Totale ordinato YtD e Totale Rolling 2021

Il report continua con un grafico ad aree in cui si possono notare l'ordinato e lo spedito dell'anno in corso per ogni mese (Figura 6.8).

Nei mesi in cui l'ordinato è maggiore dello spedito significa che non tutti gli ordini con Data di Spedizione prevista per quel mese sono stati evasi, al contrario, quando il valore dello spedito è maggiore dell'ordinato sono stati spediti più ordini di quelli previsti probabilmente perché sono stati inviati ordini "in ritardo" o sono stati anticipati degli ordini dei mesi successivi.

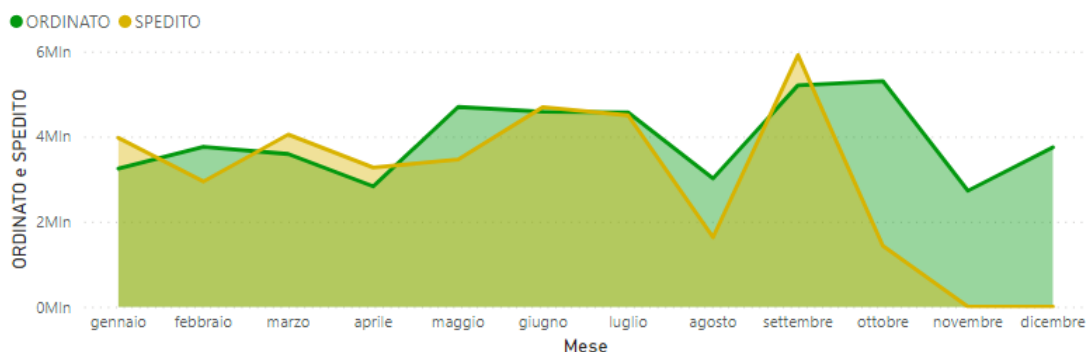


Figura 6.8: Spedito ed ordinato dell'anno in corso

Successivamente sono stati dettagliati i dati del mese in corso (Figura 6.9): in particolare possono essere visualizzati il fatturato e l'ordinato come Schede.

Un misuratore fornisce l'indicazione degli ordini ancora da spedire nel mese, il valore massimo è dato dall'ordinato mensile mentre il valore specifico rappresenta il backlog ² mensile, l'obiettivo aziendale deve essere quello di portare a zero il

²Il termine backlog è usato per indicare nel portafoglio ordini quelli che un'impresa industriale ha già ottenuto ma che deve ancora soddisfare. Nell'analisi della società, il backlog fornisce informazioni sul possibile livello del fatturato che la società potrà raggiungere.

misuratore entro la fine del mese.

Un grafico ad anello mostra la percentuale di ordini fatturati e quella degli ordini spediti ma non ancora fatturati, questi valori completano il dato del fatturato mensile, fornendo una visione più generale.

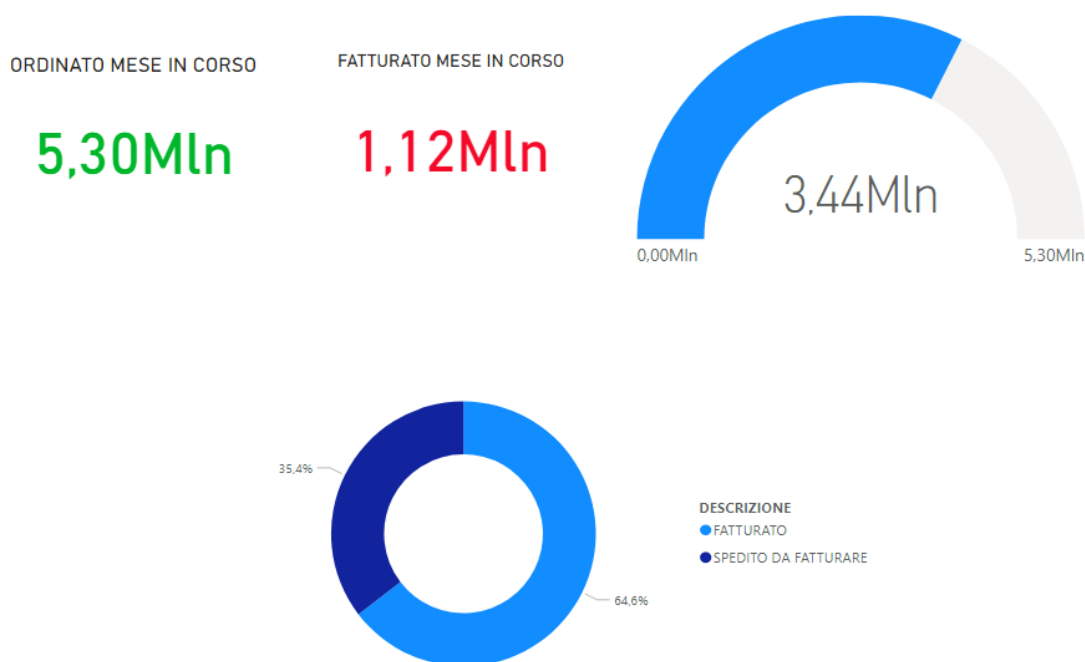


Figura 6.9: Fatturato, ordinato e spedito del mese in corso

Seguono tre grafici che forniscono dettagli ulteriori relativamente agli ordini del mese (Figura 6.10).

Il primo grafico ad anello divide gli ordini per Market code ovvero per il mercato a cui fanno riferimento, esempi sono il mercato della difesa, dei trasporti, delle telecomunicazioni ecc. tale grafico è importante per comprendere in quali settori l'azienda è inserita, quali sono i business in cui è possibile aumentare la propria quota e quali sono i settori in cui si sta perdendo.

Il primo dei due grafici successivi, invece, divide gli ordini del mese in base alla tipologia (Produzione, Acquisto, Servizio, Riparazione e Spese); il secondo effettua un drill down della tipologia Prodotto e divide i dati assegnandoli alla categoria di produzione corretta (Assemblaggio, Meccanica, Laboratorio ecc.), in questo modo è possibile monitorare il carico della produzione, relativamente agli ordini da evadere nel mese in corso.

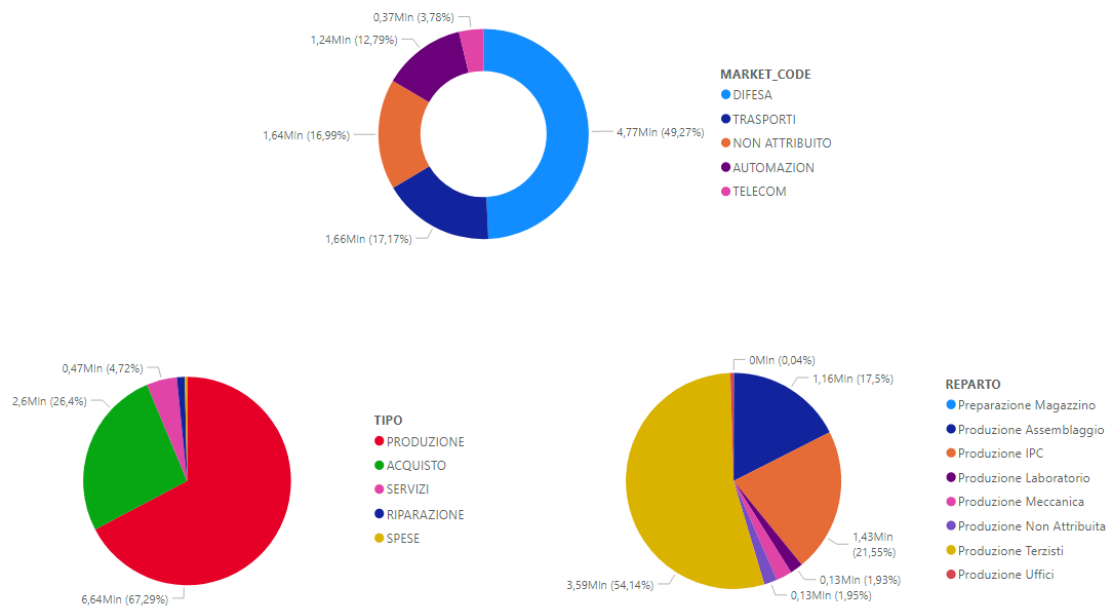


Figura 6.10: Dettaglio ordini del mese in corso

È importante notare come, quando si clicca su uno specifico Market code (o su un tipo o un reparto), gli altri grafici mutano mostrando solo i dati relativi alla parte di ordini selezionata, ciò permette di condurre analisi ancora più approfondite. Una delle metriche che si vuole monitorare è il Backlog poiché questo rappresenta gli ordini effettuati dai clienti che devono essere evasi nell'anno o negli anni successivi ma che sono già una fonte di guadagno certa.

Data la sua importanza per l'azienda, sono stati costruiti diversi grafici per rappresentarne l'andamento.

Innanzitutto, l'azienda divide il backlog dell'anno in corso rispetto a quello degli anni successivi.

È necessario fornire un dettaglio del backlog, per questo sono stati creati due differenti grafici:

- il primo (Figura 6.11) rappresenta il backlog mensile dell'anno in corso, tale grafico fornisce un'indicazione sugli ordini ancora non evasi e quelli da evadere nel breve-medio periodo. Per fare ciò è stato scelto un istogramma che riporta sull'asse delle ascisse i mesi e sull'asse delle ordinate il valore del backlog in Euro.

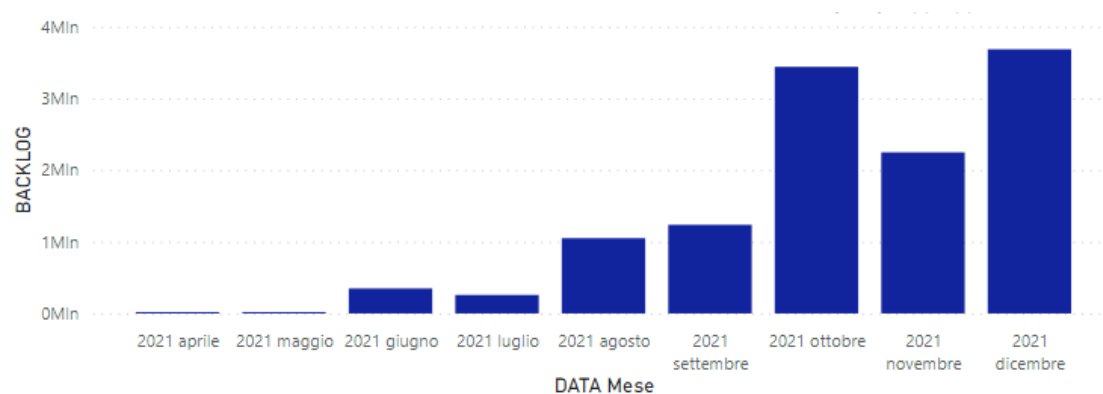


Figura 6.11: Backlog dell'anno corrente

- il secondo grafico (Figura 6.12), invece, consente la visualizzazione del backlog annuale per gli anni successivi a quello in corso, questo dato consente di avere un'idea del valore monetario degli ordini già confermati per il futuro. L'istogramma consente di individuare l'andamento del backlog, tale grafico ha gli anni sull'asse delle ascisse e il valore monetario, in Euro, del backlog sull'asse delle ordinate.

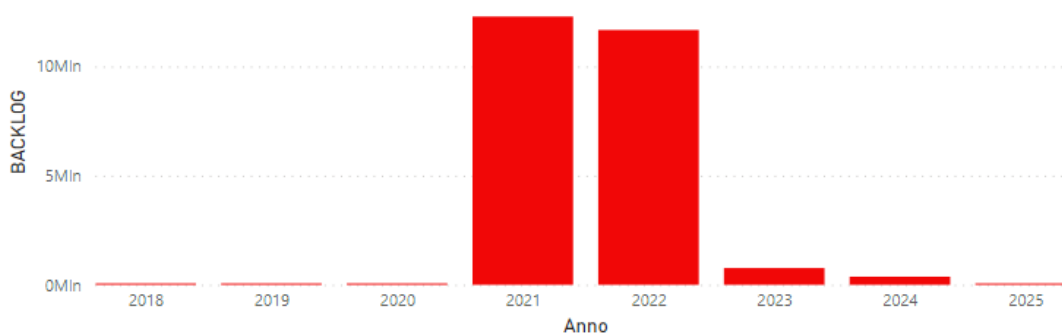


Figura 6.12: Backlog degli anni successivi a quello in corso

Le medesime informazioni realizzate per il dettaglio mensile, sono disponibili per l'intero anno in corso, ed eventualmente, modificando il filtro dati, anche per gli anni successivi o precedenti (Figura 6.13).

I dati sono presentati attraverso grafici diversi:

- il grafico ad anello indica gli ordini fatturati YtD, il backlog YtD e lo spedito ancora da fatturare.

- il grafico a linee attribuisce gli ordini ai diversi Market code, sull'asse delle ascisse si trovano le diverse tipologie di mercato e sull'asse delle ordinate l'importo totale in Euro.
- i successivo grafico a torta divide gli ordini dell'anno in base alla tipologia (Produzione, Acquisto, Servizio, Riparazione e Spese);
- l'ultimo grafico effettua un drill down della tipologia Prodotto e divide i dati assegnandoli alla categoria di produzione corretta (Assemblaggio, Meccanica, Laboratorio ecc.)

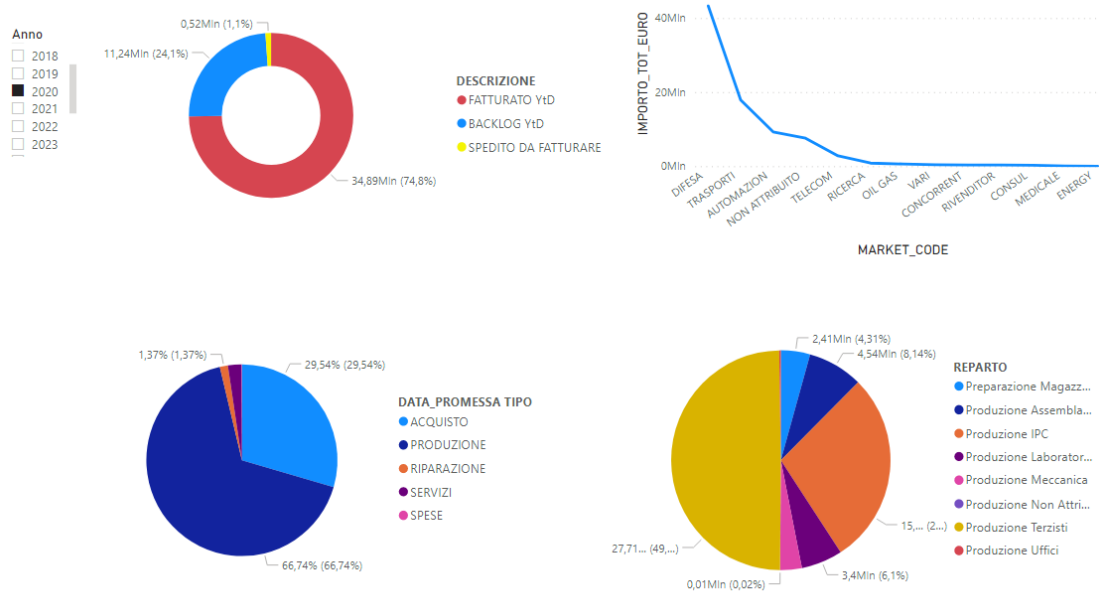


Figura 6.13: Dettaglio ordini cliente anno in corso

Altro aspetto analizzato rispetto agli ordini cliente è quello dell'andamento dei Commerciali ovvero di coloro che si occupano della stipula degli ordini con i clienti, un grafico a barre (Figura 6.14) in pila consente di individuare a quanto ammonta l'ordinato annuale o mensile di competenza; sulle ascisse è collocato l'importo totale in euro e sulle ordinate sono posti gli Agenti. Il filtro consente di scegliere un anno e/o un particolare mese.

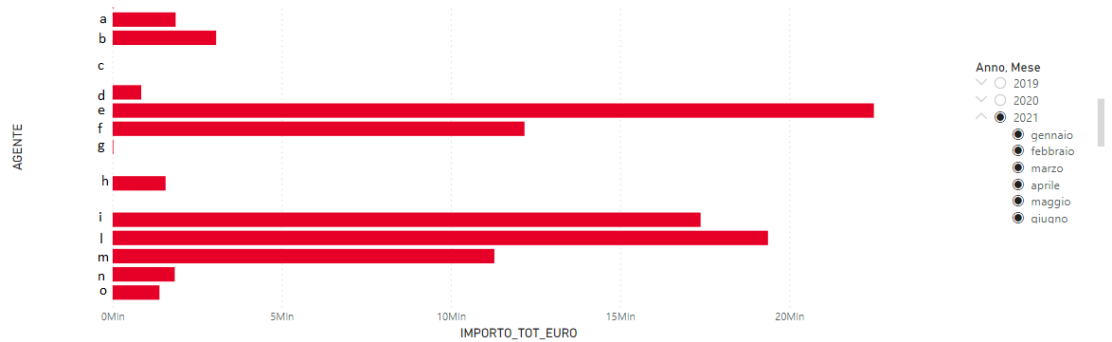


Figura 6.14: Dettaglio ordini cliente anno in corso

Per compiere un'analisi totale sugli ordini cliente è importante anche sapere quali siano i principali clienti dell'azienda, per questo è stato creato un grafico ad anello (Figura 6.15) che mostra i primi Dieci clienti per fatturato annuale.

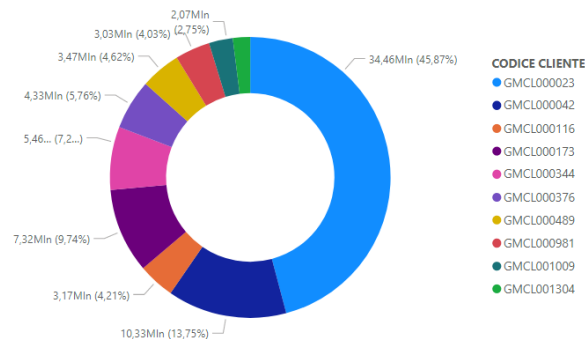


Figura 6.15: Dettaglio clienti

Selezionando uno dei clienti dal grafico è possibile visualizzare una scheda che presenta il numero di giorni di ritardo nella consegna degli ordini per quel cliente specifico considerando gli ultimi 12 mesi (Figura 6.16).

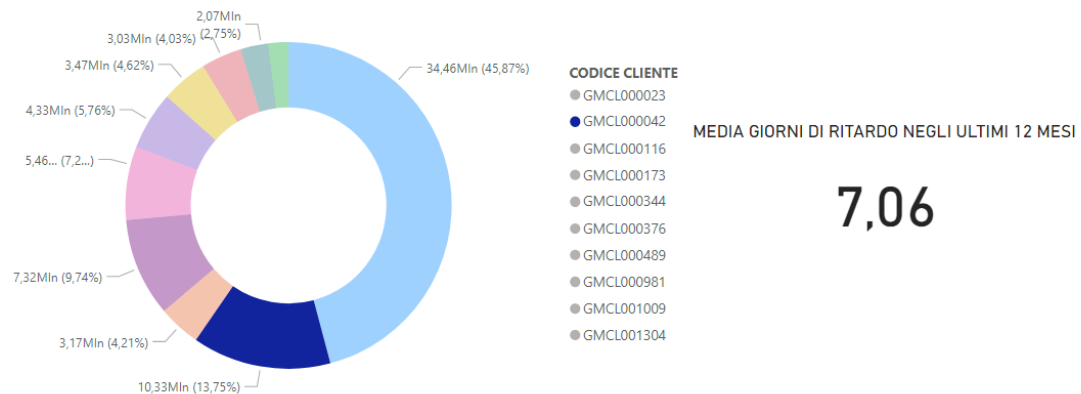


Figura 6.16: Dettaglio clienti e relativi giorni di ritardo nella consegna degli ordini

Una volta concluse le analisi sugli ordini clienti, si è passati alla realizzazione dei report riguardanti il magazzino.

Si è scelto di utilizzare due istogrammi per rappresentare le giacenze: il primo (Figura 6.17) rappresenta la giacenza del magazzino attuale, il secondo (Figura 6.18), invece, mostra i delta rispetto alla fine del mese precedente, ossia, quale è stata la variazione del magazzino dell'ultimo mese rispetto alle diverse categorie; entrambi i grafici hanno sull'asse delle ascisse gli accounting group cioè le diverse categorie di materiale (Materia prima, Prodotto finito, Sotto-assemblato e parte d'acquisto) sull'asse delle ordinate invece troviamo il valore delle giacenze in Euro.

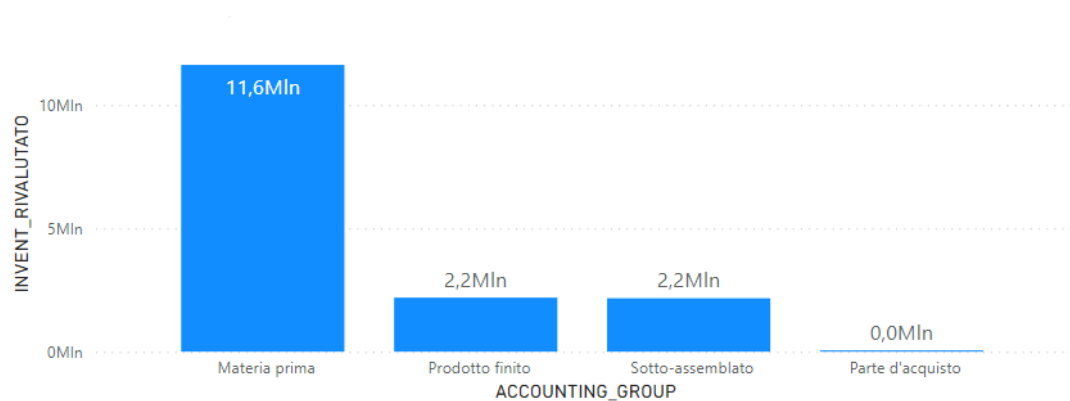


Figura 6.17: Giacenza di magazzino alla data attuale

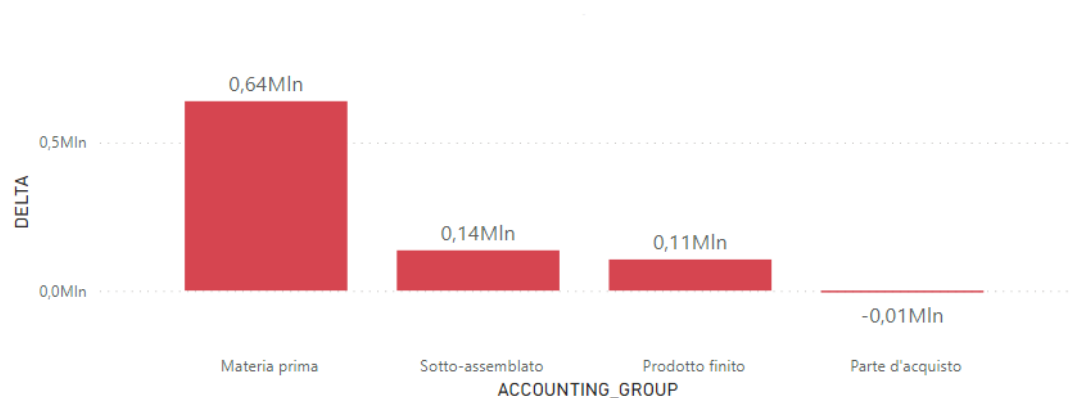


Figura 6.18: Delta giacenza di magazzino rispetto al mese precedente

La terza ed ultima sezione del report è quella relativa agli ordini d'acquisto. Per prima cosa è importante analizzare gli ordini in ritardo ovvero quegli ordini che sarebbero dovuti essere consegnati alla data attuale ma ancora non sono arrivati. Per visualizzare ciò è stato usato un istogramma (Figura 6.19) che ha sull'asse delle ascisse i mesi e gli anni e sull'asse delle ordinate l'importo totale in Euro.

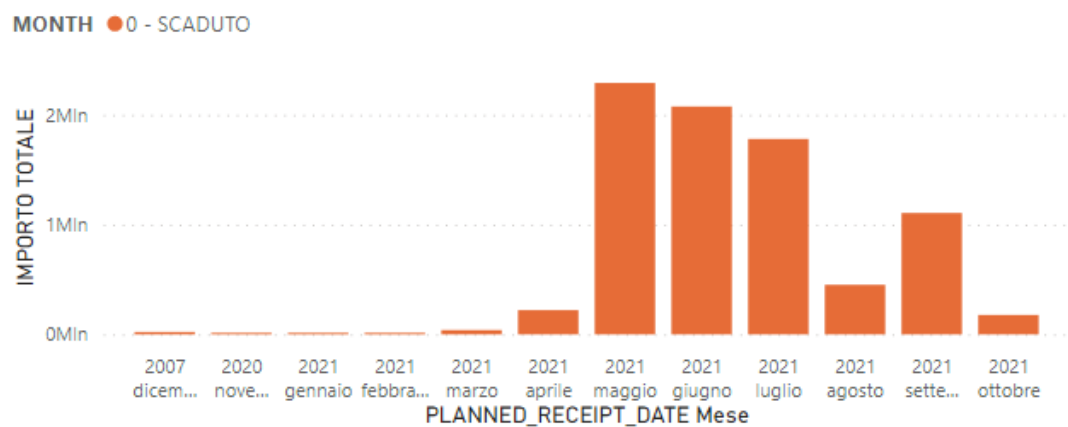


Figura 6.19: Analisi ordini d'acquisto in ritardo

Successivamente, si vogliono monitorare gli Arrivi riferiti agli ordini di acquisto, per questo è stato realizzato un grafico a linee (Figura 6.20) che ha sull'asse delle X il tempo e sull'asse delle Y l'importo totale degli ordini; un filtro consente di visualizzare il dettaglio annuale o di navigare tra gli anni.

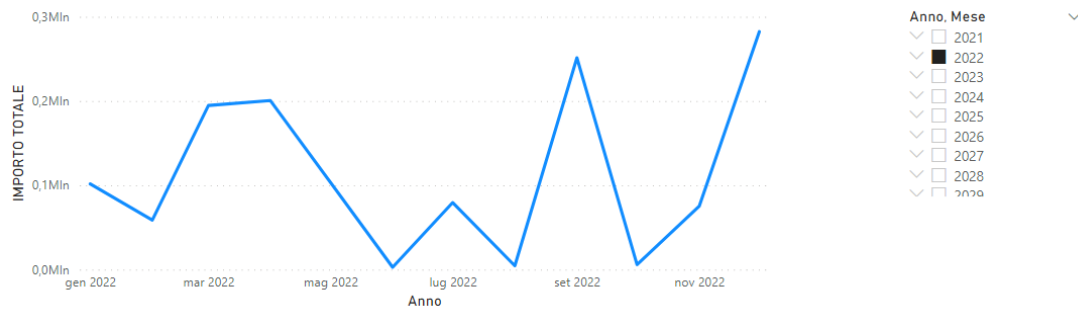


Figura 6.20: Analisi ordini d'acquisto futuri

Esattamente come fatto per i clienti, è necessario sapere quali e chi siano i principali fornitori rispetto all'ordinato, ciò è visualizzabile grazie ad un grafico ad anello (Figura 6.21) che presenta i TOP 10 fornitori.

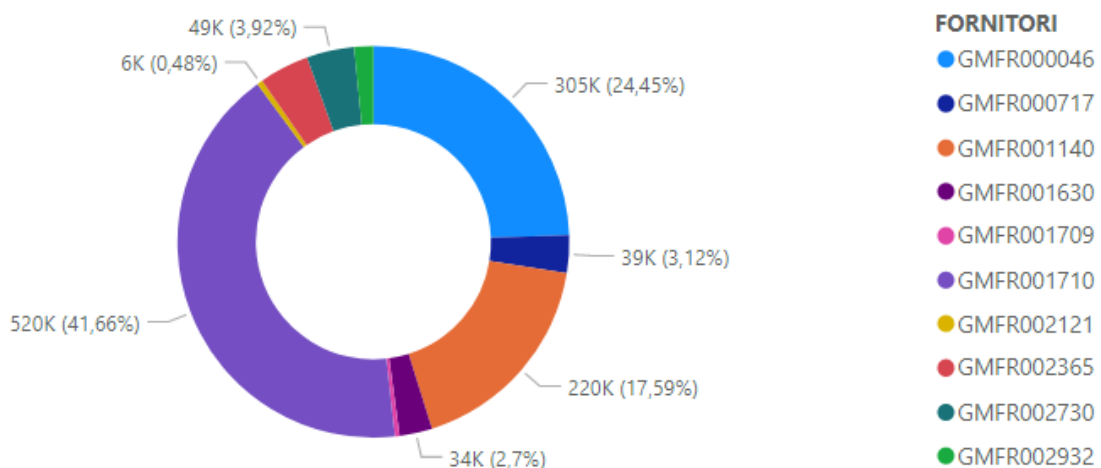


Figura 6.21: Analisi TOP 10 fornitori

In conclusione, è stata condotta un'analisi sui nuovi fornitori e sui nuovi clienti e le loro collocazioni geografiche in modo tale da poter scoprire nuovi mercati e nuove opportunità.

Due grafici a cascata (Figure 6.22 e 6.23) che riportano sull'asse delle ascisse le nazioni e sull'asse delle ordinate il numero dei nuovi clienti consentono di visualizzare tali informazioni a prima vista.

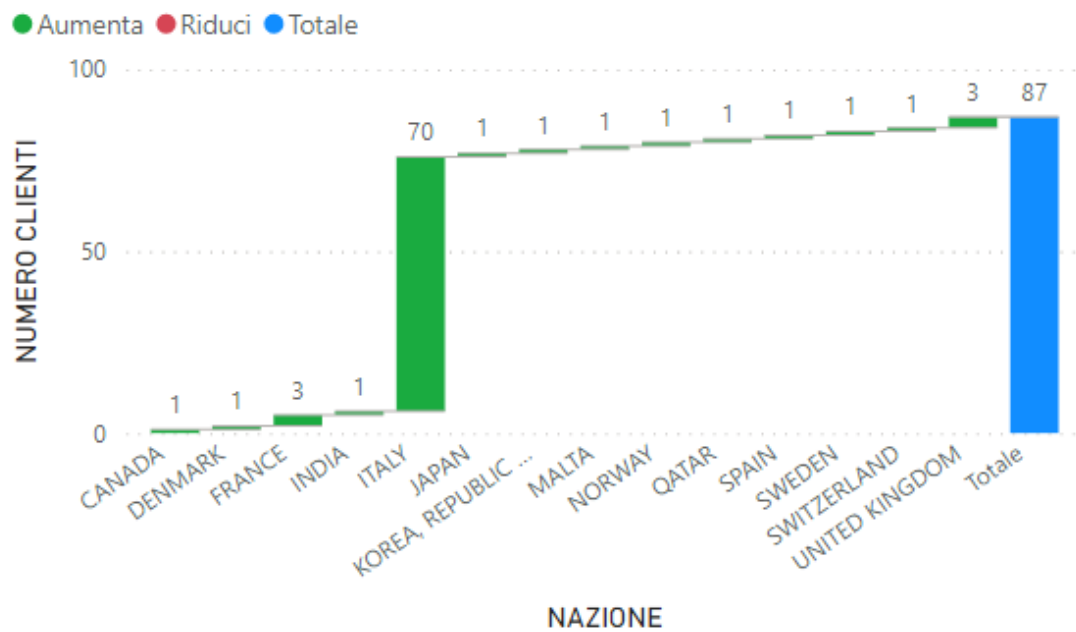


Figura 6.22: Analisi nuovi clienti

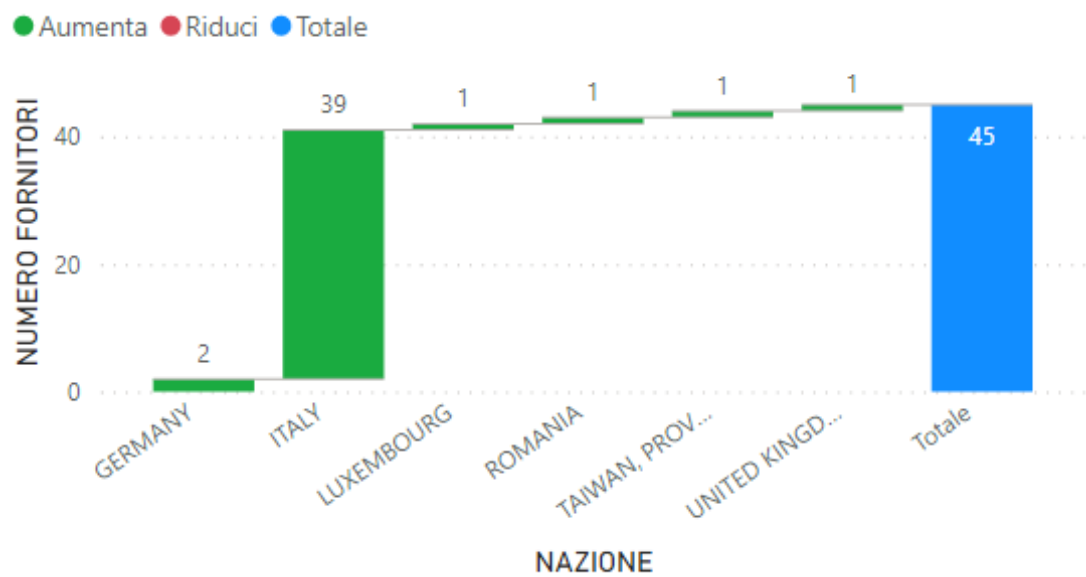


Figura 6.23: Analisi nuovi fornitori

6.2 Dashboard

Gli ultimi step del progetto sono stati la creazione delle Dashboard e la successiva condivisione.

Ciò è stato possibile mediante l'ambiente dedito di Power BI Service che ha consentito il poter caricare rispettivamente il data set ed e le relative visuals.

Da quest'ultime, si è avuto modo di creare le dashboard, da porre in condivisione con l'eventuale resto del personale.

Quest'ultimo è un servizio estremamente utile e comodo per poter meglio consentire l'allineamento e la comunicazione delle informazioni estrapolate.

Sono state create più dashboard a seconda della tematica affrontata o dell'utente di destinazione.

Dopo aver effettuato l'accesso su Power BI Service è possibile recuperare i dati dall'esterno (Figura 6.24) , nel caso specifico è stato importato direttamente il file .pbix generato in Power BI Desktop.

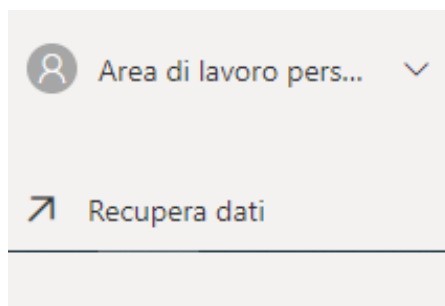


Figura 6.24: Recupero dati Power BI Service

Il file viene aperto in Visualizzazione di lettura ma è possibile passare alla Visualizzazione di modifica per effettuare delle modifiche e creare le dashboard desiderate.

Per creare le diverse strutture sono state seguite delle regole generali: le dashboard devono avere un layout logico che mostri le diverse informazioni secondo un flusso coerente.

I dettagli, le funzionalità o le immagini extra possono complicare le dashboard e distogliere l'attenzione dal messaggio che si vuole comunicare, creando confusione e riducendo il coinvolgimento o l'interesse nei dati.

Se i dati sono presentati in modo diverso, usando i colori per mettere in evidenza, cambiando tipo di grafico, assegnando un'etichetta agli indicatori e personalizzando le informazioni, gli utenti saranno invogliati a interagire con determinati dettagli. Coinvolgere i diversi utenti nella creazione delle dashboard per avere diversi punti di vista o porre domande in modo da poter approfondire l'analisi per raccontare i

dati al meglio.

Man mano che si diffonde la cultura dell'analisi nelle organizzazioni, è auspicabile avere un feedback continuo.

Per ottenere un risultato migliore è utile prevedere diverse bozze ed iterazioni [19]. Sono state create diverse dashboard, di seguito vengono presentate le principali:

- La prima dashboard presenta la situazione di Ordinato, Fatturato e Spedito dell'anno selezionato, di default quando si apre la pagina si visualizzano i dati riferiti all'anno corrente (Figura 6.25).

Le informazioni presentate sono utili alla direzione per monitorare l'andamento annuale, individuare particolari trend e comprendere se l'azienda sta perseguendo i propri obiettivi o meno.

Il grafico inerente allo spedito e all'ordinato, invece, è utile per comprendere quanti ordini (in termini di valore totale) sono stati evasi per ogni mese, i mesi con uno spedito basso rispetto all'ordinato potrebbero essere indice di una cattiva gestione o di ritardi; questo potrebbe essere uno spunto per effettuare analisi più specifiche.

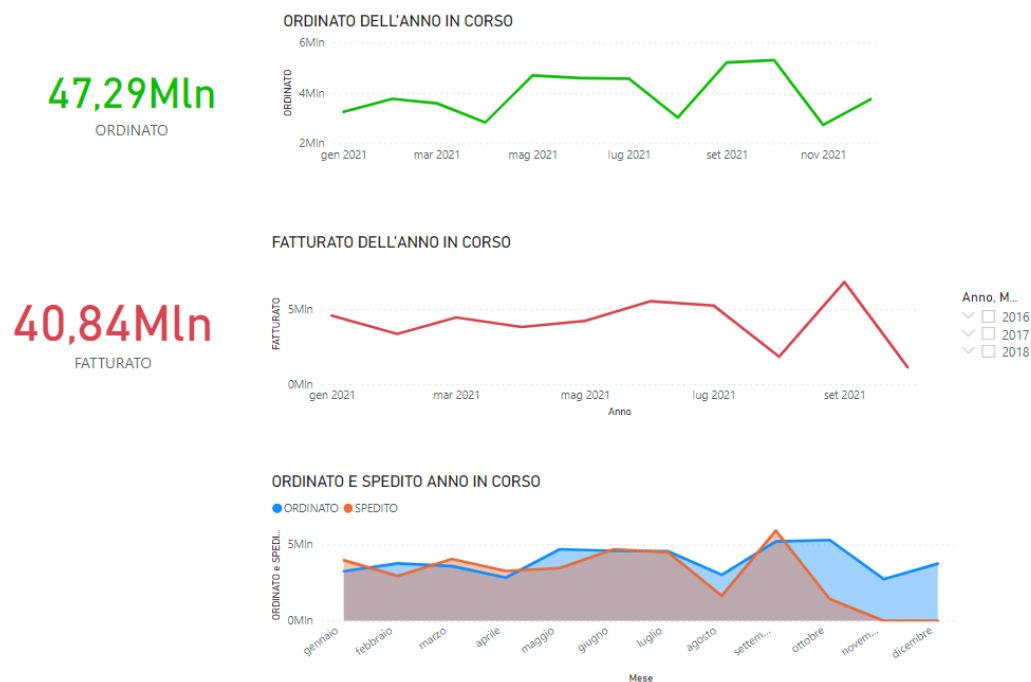


Figura 6.25: Dashboard 1

- La seconda dashboard propone i dati del mese in corso, possiamo notare i valori relativi all'Ordinato, al Fatturato, al valore degli ordini ancora da

spedire ed il dettaglio degli ordini per market_code e per tipologia (Figura 6.26).

Tali informazioni sono utili per effettuare un'analisi di alto livello, il misuratore consente di capire a quanto ammonta il valore degli ordini ancora da spedire, chiaramente, all'inizio del mese il livello sarà molto alto, l'obiettivo è quello di portare a zero il misuratore entro la fine del mese.

Gli altri grafici consentono di avere un'idea sulla diversa distribuzione degli ordini per market_code e per tipologia.

In questa dashboard è possibile notare uno dei punti di forza di Power BI, l'interattività: cliccando sui dati di un grafico, per esempio, sulla tipologia "Produzione" del grafico a torta, tutti gli altri grafici cambiano automaticamente presentando solo i valori inerenti alla categoria selezionata (il valore dell'ordinato mensile dei prodotti che rientrano nella categoria "Produzione" etc.).

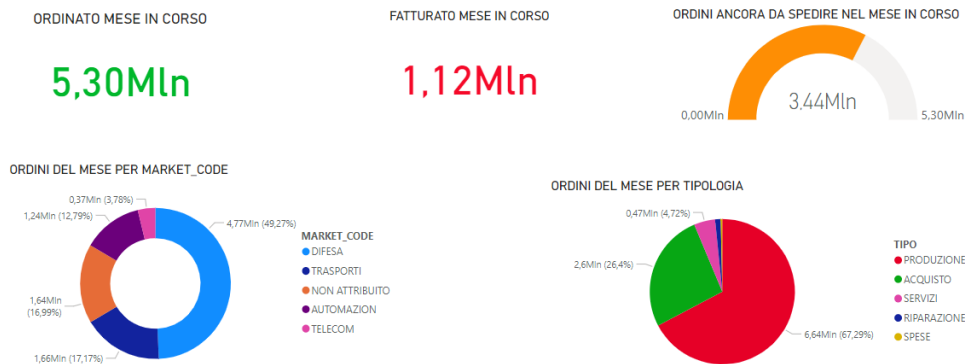


Figura 6.26: Dashboard 2

- Una dashboard importante è quella che presenta i dati relativi al magazzino, i due grafici mostrano rispettivamente la giacenza del magazzino alla data attuale e la differenza delle giacenze attuali rispetto al mese precedente (Figura 6.27).

Le giacenze sono presentate sulla base delle tipologie di materiale.

Questa visualizzazione è utile al responsabile del magazzino per una corretta gestione delle scorte.

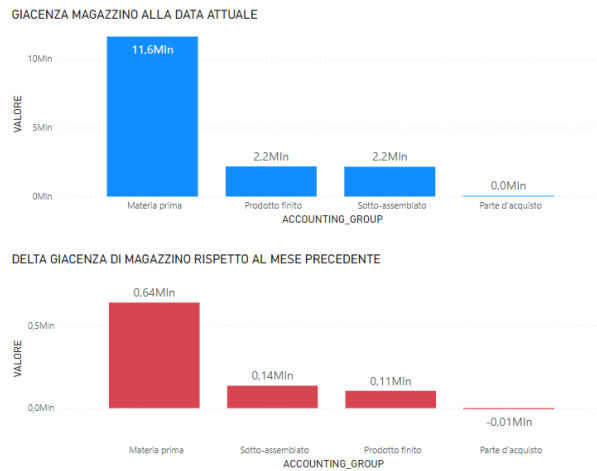


Figura 6.27: Dashboard 3

- Ultima dashboard analizzata è quella relativa agli ordini d'acquisto. I grafici mostrano un'analisi degli ordini non ancora ricevuti, degli ordini da ricevere nei mesi successivi ed il valore degli acquisti annuali per fornitore (Figura 6.28). Il filtro consente di visualizzare i dati relativi ad un altro anno. Tali rappresentazioni sono utili per l'Ufficio acquisti che deve monitorare lo stato degli ordini e per il responsabile di magazzino che dovrà gestire il ricevimento del materiale.

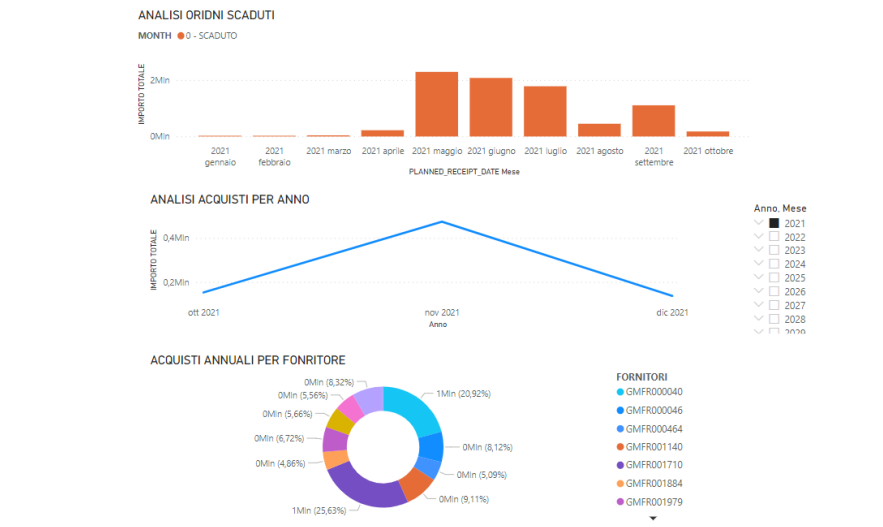


Figura 6.28: Dashboard 4

L'ultimo passaggio è la condivisione del lavoro.

In Power BI è possibile creare aree di lavoro, ovvero spazi per collaborare con i colleghi per creare e ridefinire raccolte di dashboard, report e report impaginati. Quando si crea un'area di lavoro classica, si crea un gruppo di Microsoft 365 sottostante associato.

Tutta l'amministrazione dell'area di lavoro viene gestita in Microsoft 365.

È possibile aggiungere altri colleghi a queste aree di lavoro come membri o amministratori.

Nell'area di lavoro (Figura 6.29) tutti gli utenti possono collaborare su dashboard, report e altri articoli che si intende pubblicare e rendere disponibili a un pubblico più ampio.

Tutti gli utenti che vengono aggiunti a un'area di lavoro devono possedere una licenza di Power BI Pro.

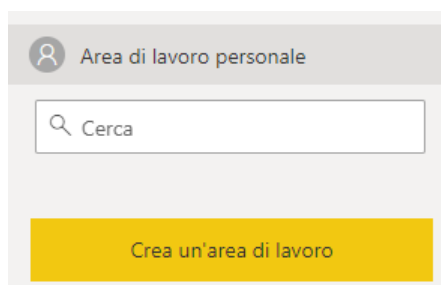


Figura 6.29: Recupero dati Power BI Service

Quando si crea un'App dall'area di lavoro si sceglie se renderla Pubblica, chiunque nell'organizzazione potrà visualizzare il contenuto dell'area di lavoro o Privata cioè solo i membri dell'area di lavoro possono visualizzarne il contenuto.

Dopo aver creato il gruppo, non è possibile modificare l'impostazione di gruppo pubblico o privato.

È anche possibile scegliere se i membri potranno modificare o avere l'accesso di sola visualizzazione.

Al termine, Power BI crea l'area di lavoro che verrà aperta e visualizzata nell'elenco delle aree di lavoro di cui si è membri.

Gli amministratori possono a questo punto selezionare gli elementi per apportare modifiche, aggiungere nuovi membri o modificarne le autorizzazioni.

Per distribuire contenuto ufficiale a un ampio gruppo di destinatari all'interno dell'organizzazione, è possibile pubblicare un'app dall'area di lavoro.

Quando il contenuto nell'area di lavoro è pronto, scegliere le dashboard e i report da pubblicare e quindi pubblicarli come App [20].

Capitolo 7

Conclusioni

Con l'avvento dei Big Data è stato creato un nuovo mondo, le aziende producono giornalmente una grande quantità di dati ed il progresso tecnologico consente di raccogliarli. Grazie alla digitalizzazione delle organizzazioni e con l'utilizzo di strumenti di analisi è possibile far sì che questi siano un fattore chiave nella competizione e un potente mezzo di supporto per le scelte strategiche aziendali orientando i decision makers nelle decisioni e nelle strategie in modo puntuale ed efficace.

La business intelligence è uno strumento ormai indispensabile all'interno di una azienda, consente di trarre il massimo beneficio dai dati.

Solo attraverso la costante misurazione delle performance è possibile ottimizzare i processi, risparmiare le risorse ed avere un quadro chiaro e dettagliato dello stato dell'organizzazione.

Uno degli strumenti che non può mancare per monitorare le performance dei processi è la Data Visualization.

Il focus della presente tesi è stato lo sviluppo di un progetto di business intelligence particolarmente incentrato sulla Data Visualization, la quale permette agli utenti che necessitano di risposte, di effettuare le analisi pur non disponendo di competenze statistiche o informatiche avanzate, interagendo con le dashboard.

Una delle difficoltà più grandi è stata la definizione del contenuto in termini di organizzazione dei dati, scelta dei KPI e chiarezza di presentazione.

Il progetto di BI ha consentito di condurre un'analisi approfondita di molti ed importanti dati aziendali, di produrre report efficienti e facilmente interpretabili. Per fare ciò è stata necessaria una rielaborazione dei dati utile ad individuare i punti di forza e di debolezza degli stessi.

Rispetto alla situazione AS IS i dati non sono più presentati in formato tabulare ma attraverso grafici ad hoc, tale tipologia di rappresentazione ha consentito di personalizzare l'interfaccia che mostra i dati all'utente rendendola interattiva ed intuitiva.

La corretta implementazione della Data Visualization ha reso i dati maggiormente comprensibili ed ha consentito di evitare la confusione.

Passaggi critici sono stati il coinvolgimento degli utenti nello sviluppo e la progettazione di un sistema di dati coerente.

I dati presenti in un sistema aziendale sono numerosi ed è necessario focalizzarsi solo sulle informazioni rilevanti per gli utenti.

E' stata posta molta attenzione nella creazione di report e dashboard, infatti, una volta caricati i dati e creato il modello è stato molto utile rappresentarli.

Le informazioni rilevanti devono essere visibili a colpo d'occhio, si è rivelato molto efficace inserire dei grafici ed adottare strumenti e colori idonei a rappresentare tutte le diverse situazioni.

I report sono stati creati in modo che fossero facilmente navigabili, l'impiego di filtri, funzioni di drill down e drillthrough hanno reso così l'interfaccia intuitiva.

L'approccio agile ha consentito una gestione trasparente e flessibile, infatti, creare dashboard è un processo continuo che si chiude solo quando si ottiene il risultato desiderato. Le organizzazioni e le persone attraversano momenti di cambiamento, per questo una dashboard efficace deve essere flessibile e adattarsi alle esigenze e agli obiettivi aziendali in continuo mutamento, ma non bisogna dimenticare che, per avere risultati concreti, è essenziale aumentare l'affidabilità e la credibilità dei dati.

Guardando al futuro, possiamo affermare che si potrà usufruire di tecniche sempre più sofisticate per l'estrapolazione della conoscenza partendo dai dati provenienti da diverse sorgenti.

Sarà possibile creare report informativi sempre più avanzati e particolareggiati, nel contempo più intuitivi e immediati grazie all'impiego di nuovi modelli e tecniche d'intelligenza artificiale.

I servizi di BI mireranno a fornire risposte in real time, minimizzando i tempi di aggiornamento, ciò sarà facilitato dal progresso delle infrastrutture tecnologiche che puntano ad aumentare la velocità di calcolo.

Bibliografia

- [1] Tableau. *Piattaforme di BI*. URL: <https://www.tableau.com/it-it/learn/articles/business-intelligence/choosing-bi-platforms> (cit. a p. 2).
- [2] ITREVIEW SMART BUSINESS VISIONS. *Data visualization: un nuovo modo di osservare i dati*. URL: <https://itreview.it/data-visualization-un-nuovo-modo-di-osservare-i-dati/> (cit. a p. 2).
- [3] Salesforce Italia. *Che cos'è la quarta rivoluzione industriale?* URL: <https://www.salesforce.com/it/blog/2019/08/che-cosa-quarta-rivoluzione-industriale.html> (cit. a p. 4).
- [4] Dan Porto. *Perché il ruolo della cultura organizzativa è essenziale per la trasformazione digitale?* URL: <https://www.blog-management.it/2021/02/24/cultura-organizzativa-trasformazione-digitale/> (cit. a p. 6).
- [5] Luca Dal Porto. «La trasformazione digitale nelle imprese: fenomeni digitali e pratiche organizzative dopo l'avvento della trasformazione». In: (2016) (cit. a p. 8).
- [6] David Schatsky Ragu Gurumurthy. *Pivoting to digital maturity*. 2019. URL: https://www2.deloitte.com/content/dam/Deloitte/it/Documents/financial-services/Pivoting-to-digital-maturity_Deloitte.pdf (cit. a p. 8).
- [7] ISTAT. *Digitalizzazione e tecnologia nelle imprese italiane*. URL: https://www.istat.it/it/files//2020/08/REPORT_DIGITALIZZAZIONE_CENSIMPRESE_PC.pdf (cit. alle pp. 9, 10, 12, 13).
- [8] ISTAT. *IL SISTEMA DELLE IMPRESE: TRA CRISI E RIPRESA*. 2021. URL: https://www.istat.it/storage/rapporto-annuale/2021/Capitolo_4.pdf (cit. alle pp. 11, 13).
- [9] Tableau. *Cos'è la business intelligence? La tua guida alla BI e al perché è importante*. URL: <https://www.tableau.com/it-it/learn/articles/business-intelligence> (cit. a p. 15).

- [10] Elena Baralis. *Data warehouse: analisi dei dati*. URL: <https://dbdmg.polito.it/wordpress/wp-content/uploads/2010/12/3-DWquery2.pdf> (cit. alle pp. 21, 22, 24).
- [11] Elena Baralis. *Data warehouse: introduzione*. URL: https://dbdmg.polito.it/wordpress/wp-content/uploads/2021/04/1-DW_intro_x2.pdf (cit. a p. 25).
- [12] IBM design language. *Charts*. URL: <https://www.ibm.com/design/language/data-visualization/charts> (cit. alle pp. 27–29).
- [13] Datapine. *Tipi di dashboard*. URL: <https://www.datapine.com/blog/strategic-operational-analytical-tactical-dashboards/#strategic-dashboard-example> (cit. alle pp. 31–33).
- [14] Microsoft. *Power BI*. URL: <https://powerbi.microsoft.com/it-it/what-is-power-bi/> (cit. a p. 35).
- [15] Microsoft. *Che cos'è Power BI*. URL: <https://docs.microsoft.com/it-it/power-bi/fundamentals/power-bi-overview> (cit. a p. 36).
- [16] Microsoft. *Modellazione e visualizzazione dei dati*. URL: <https://docs.microsoft.com/it-it/learn/modules/introduction-power-bi/2a-data-modeling-visualizations> (cit. a p. 36).
- [17] Microsoft. *Dashboard per gli utenti aziendali del servizio Power BI*. URL: <https://docs.microsoft.com/it-it/power-bi/consumer/end-user-dashboards> (cit. a p. 42).
- [18] Wikipedia. *Microsoft Access*. URL: https://it.wikipedia.org/wiki/Microsoft_Access (cit. a p. 43).
- [19] Tableau. *Creazione di dashboard*. URL: <https://www.tableau.com/it-it> (cit. a p. 73).
- [20] Microsoft. *Creare aree di lavoro in Power BI*. URL: <https://docs.microsoft.com/it-it/power-bi/collaborate-share/service-create-workspaces> (cit. a p. 77).