

Data Analytics e dintorni

Tommaso Rossi

25.9.2018

Descriptive data analytics

- Trasformare in informazione un groviglio di dati... «dando loro un'occhiata» osservandoli non tali e quali uno dopo l'altro, la nostra mente non è molto abile a catturare informazioni in questo modo, ma attraverso una rappresentazione grafica (alternativamente riassumendoli in pochi valori che possano essere guardati e interpretati direttamente)
- Il descriptive data analytics consente di trasformare i dati in evidenze e numeri permettendo così:
 - di imparare da ciò che è accaduto nel passato, ovvero da ciò che accaduto un secondo, un'ora, un anno fa, ...
 - di prendere decisioni in modo consapevole
- La maggior parte delle analisi che vengono condotte per compiere scelte di tipo operativo, tattico e strategico in azienda ricadono nella categoria del descriptive data analytics

Descriptive data analytics

- Le principali tecniche descrittive che possono essere efficacemente impiegate per mettere in luce caratteristiche prima nascoste e insite nel dato sono:
 - analisi univariate (media, mediana, range, varianza, deviazione standard)

Postazione	Assemblaggio 4
Calciobalilla	Tempo [min]
1	4
2	3,7
3	3,8
4	3,2
5	3,7
6	3,8
7	3,2
8	2,5
9	1,8
MEDIA	3,30
DEV STD	0,73
IRE	0,22

Questi tipi di analisi possono essere condotte con strumenti molto semplici come, per esempio, Excel grazie a funzioni come: `media()`, `dev.st()`

Misura di posizione

Misura di dispersione

Descriptive data analytics

- Le principali tecniche descrittive che possono essere efficacemente impiegate per mettere in luce caratteristiche prima nascoste e insite nel dato sono:
 - analisi bivariate (covarianza, correlazione)

Codice	Vendite 2013	Vendite 2014	Vendite 2015	k	Rk
6883	65	44	56	1	0,376975
	40	73	70	2	0,128205
	105	85	77	3	-0,09674
	92	69	124	4	-0,208
	136	102	117	5	-0,42236
	96	113	110	6	-0,52068
	85	113	108	7	-0,54924
	60	34	20	8	-0,18597
	79	75	63	9	0,016336
	83	83	70	10	0,20532
	39	49	42	11	0,461621
	21	21	41	12	0,727915

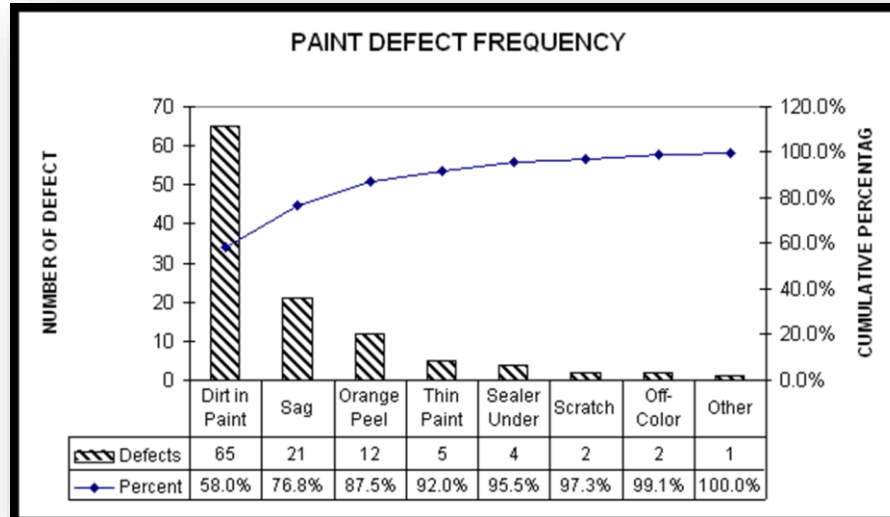
Questi tipi di analisi possono essere condotte con strumenti molto semplici come, per esempio, Excel grazie a funzioni come: `correlazione()`, `covarianza.c()`

Indici di correlazione a 'k' mesi

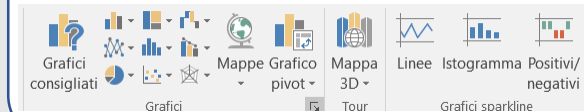
L'indice di correlazione è massimo per $k=12$. Questo significa che la domanda del codice è caratterizzata da una stagionalità di passo 12

Descriptive data analytics

- Le principali tecniche descrittive che possono essere efficacemente impiegate per mettere in luce caratteristiche prima nascoste e insite nel dato sono:
 - grafici (a dispersione, istogrammi, box-plot, ...)

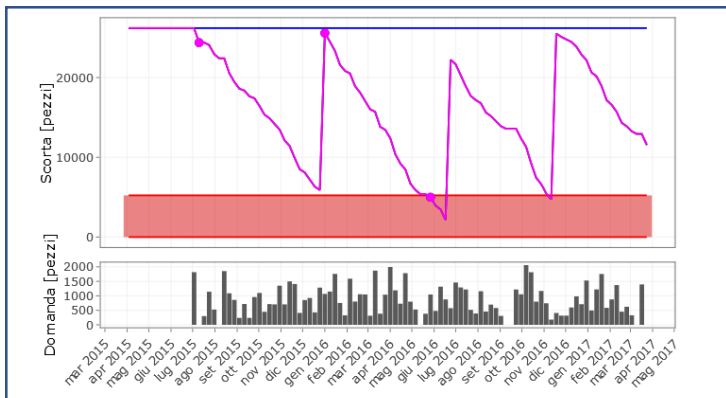


Questi tipi di analisi possono essere condotte con strumenti molto semplici come, per esempio, Excel grazie a



Descriptive data analytics – esempi

Obiettivo: visualizzare l'andamento della domanda dei codici venduti da un'azienda distributrice di prodotti ottici e, quindi, visualizzare l'andamento della loro giacenza e il valore assunto da alcuni indicatori di sintesi. Il fine ultimo era quello di verificare quali modalità di gestione dei materiali potessero essere adottate dall'azienda (modelli pull vs modelli push)



SERVER & STORAGE



Area spedizioni							
Regione	Provincia	SIGLA	1000	2000	3000	4000	
Giulia	Sapri	SP	15,2	14,9	14,2	13,78	
Piemonte	Alessandria	AL	8,7	8,2	7,7	7,45	
Marche	Ancona	AN	10,5	10	9,5	8,5	
Valle d'Aosta	Aosta	AO	11,5	11	10,5	10	
Toscana	Arezzo	AR	9,7	9,2	8,7	8,2	
Marche	Ascoli Piceno	AP	10,5	10	9,5	8,5	
Diemonte	AV	BT	8,7	8,2	7,7	7,45	
Campania	Avellino	AV	11,4	11,2	10,7	10,2	
CITTA DESTINAZIONE			PR.	NAZIONE	NR. IMB.	Q.TA. NETTA	Q.TA. LORDA
Puglia	TREBASELEONE	PD	ITALIA	1	185,00	201,00	0
Veneto	FROSSASCO	TO	ITALIA	2	400,00	432,00	0
Campania	GIUSEGGANA	TV	ITALIA	8	1.760,00	1.872,00	0
Lombardia	SUSEGGANA	TV	ITALIA	8	1.840,00	1.952,00	0
Piemonte	NERVESSA DELLA BATTAGLIA	TV	ITALIA	1	230,00	234,00	0
Emilia Romagna	NERVESSA DELLA BATTAGLIA	TV	ITALIA	1	230,00	244,00	0
Lombardia	CAPRIANO DEL COLLE	BS	ITALIA	4	880,00	936,00	0
	CAPRIANO DEL COLLE	BS	ITALIA	4	920,00	976,00	0
	RIVALTA SCRIVIA	AL	ITALIA	4	800,00	856,00	0
	BERGAMO	BG	ITALIA	2	400,00	437,00	0

ARTICOLO	MEDIA	STDEV	IRE	Gestione
528	4,76	4,88452	1,02616	Previsione
526	0,15	0,568195	3,787964	Lampy
525	0,16	0,605807	3,786291	Lampy
524	0,24	0,717132	2,98805	Lampy
522	5,97	4,718182	0,790315	Previsione
521	9,13	6,978853	0,764387	Previsione
520	0,95	1,34807	1,419021	Previsione
518	14,23	9,048901	0,635903	Scorta
516	0,85	1,3038	1,533882	Previsione
514	5,21	4,760277	0,913681	Previsione

Descriptive data analytics – esempi

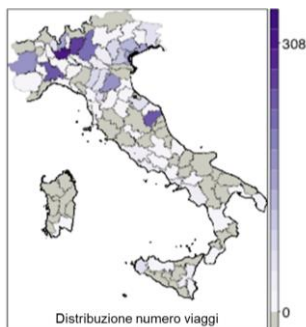
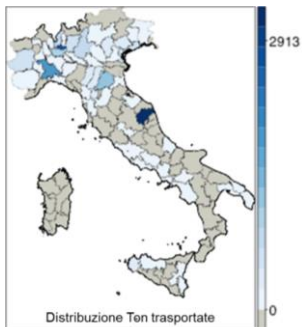
Obiettivo: evidenziare disallineamenti tra quantità spedite verso una determinata provincia/un determinato NUTS e numero di spedizioni effettuate verso la medesima provincia/il medesimo NUTS. Il fine ultimo era quello di razionalizzare la logistica di un'azienda operante nel settore delle resine individuando le situazioni inefficienti (basse quantità spedite vs elevato numero di spedizioni)



SERVER & STORAGE



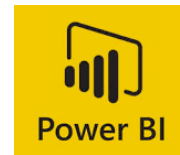
Area spedizioni							
Regione	Provincia	SIGLA		1000	2000	3000	4000
Giulia	Sagreda	MS		15,2	14,9	14,2	13,75
Piemonte	Alessandria	AL		8,7	8,2	7,7	7,45
Marche	Ancona	AN		10,5	10	9,5	8,5
Valle d'Aosta	Aosta	AO		13,5	11	10,5	10
Toscana	Arezzo	AR		9,7	9,2	8,7	8,2
Marche	Ascoli Piceno	AP		10,5	10	9,5	8,5
Diemonte	AV	BT		8,7	8,2	7,7	7,45
Campania	Avellino	AV		11,4	1,3	10,7	8,2
CITTA DESTINAZIONE	PR.	NAZIONE		NR. IMB.	Q.TA NETTA	Q.TA LORDA	
Puglia_TREBASSELEONE	PD	ITALIA		1	185,00	201,00	
Veneto_FROSSASCO	TO	ITALIA		2	400,00	432,00	
Campania_GUSEGANA	TV	ITALIA		8	1.760,00	1.872,00	
Lombardia_SUSEGANA	TV	ITALIA		8	1.840,00	1.952,00	
Piemonte_MERVESSA DELLA BATTAGLIA	TV	ITALIA		1	220,00	234,00	
Emilia_Romagna_MERVESSA DELLA BATTAGLIA	TV	ITALIA		1	230,00	244,00	
Lombardia_CAPRIANO DEL COLLE	BS	ITALIA		4	880,00	936,00	
Capriano del Colle	BS	ITALIA		4	920,00	976,00	
Rivalta Scrivia	AL	ITALIA		4	800,00	856,00	
Bergamo	BG	ITALIA		2	400,00	437,00	



Altri software:



Power Map



...

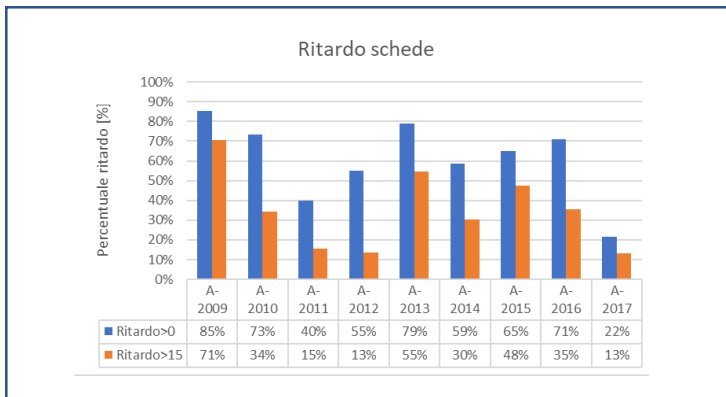
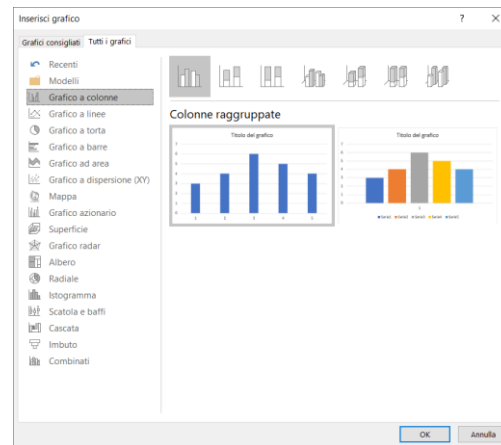


Descriptive data analytics – esempi

Obiettivo: valutare l'andamento del numero di commesse in ritardo per un'azienda del settore delle macchine per la lavorazione di materie plastiche e gomma. Il fine ultimo era quello di verificare se l'impressione degli operatori, ovvero una situazione di ritardo cronica, fosse vera o meno



MVCODICE	MVDESART	MVATCCO	MVOTAMOV	MVVCODCOM	MVDATEVA	MVCODA
S.13053	1225_5/E13.LD	01001	1		01/07/2015	
E.13059	1425/E13.LD	01001	1		20/03/2015	
054.56992	PRO.MULT.CENT. E13-LD D.05092/A P.1	01025	2	S.13053	08/05/2015	OFFICINA
054.54551	PRO.MULT.EST E13-LD 350x37 D.05092/P.6	01025	4	S.13053	08/05/2015	OFFICINA
054.54552	PRO.MULT.LARD. E13.1224x54 D.10035/A P.1	01025	1	S.13053	08/05/2015	OFFICINA
054.54724	PRO.MULT.LARD. E13.1224x54 D.12024/F P.2	01025	1	S.13053	08/05/2015	OFFICINA
053.53179	PROT. CAVI 1 F. 22x340 D.05008/P.1	01025	1	S.13053	08/05/2015	OFFICINA
056.68260	COIBENT. E13-LD 397x212 D.13012/3 P.1	01025	1	S.13053	08/05/2015	OFFICINA
056.68261	COIBENT. E13-LD 397x212 D.13012/3 P.2	01025	2	S.13053	08/05/2015	OFFICINA
056.68436	COIBENT. E13-LD 609x777 D.13012/3 P.3	01025	1	S.13053	08/05/2015	OFFICINA
056.68437	COIBENT. E13-LD 609x777 D.13012/3 P.4	01025	1	S.13053	08/05/2015	OFFICINA
056.68438	COIBENT. E13-LD 609x305 D.13012/3 P.5	01025	1	S.13053	08/05/2015	OFFICINA
056.68439	COIBENT. E13-LD 609x305 D.13012/3 P.6	01025	1	S.13053	08/05/2015	OFFICINA
056.67796	COIB. POST. E13-LD 394x396 D.10035/5 P.8	01025	1	S.13053	08/05/2015	OFFICINA
056.67785	COIB. POST. E13-LD 394x396 D.10035/5 P.7	01025	1	S.13053	08/05/2015	OFFICINA
056.67911	COIB. LAT. E13-LD 456x396 D.11015/6 P.1	01025	1	S.13053	08/05/2015	OFFICINA
056.67912	COIB. LAT. E13-LD 456x396 D.11015/6 P.2	01025	1	S.13053	08/05/2015	OFFICINA
052.52728	CANAL. CAVI T-C 1060x52x52 D.10035/P.1	01025	1	S.13053	08/05/2015	OFFICINA
052.52656	COPERCHIO CANAL. 100x25x42 D.08001/8 P.2	01025	4	S.13053	08/05/2015	OFFICINA
052.52657	PIASTR. C-CANALINA 52x50 D.08001/8 P.3	01025	4	S.13053	08/05/2015	OFFICINA



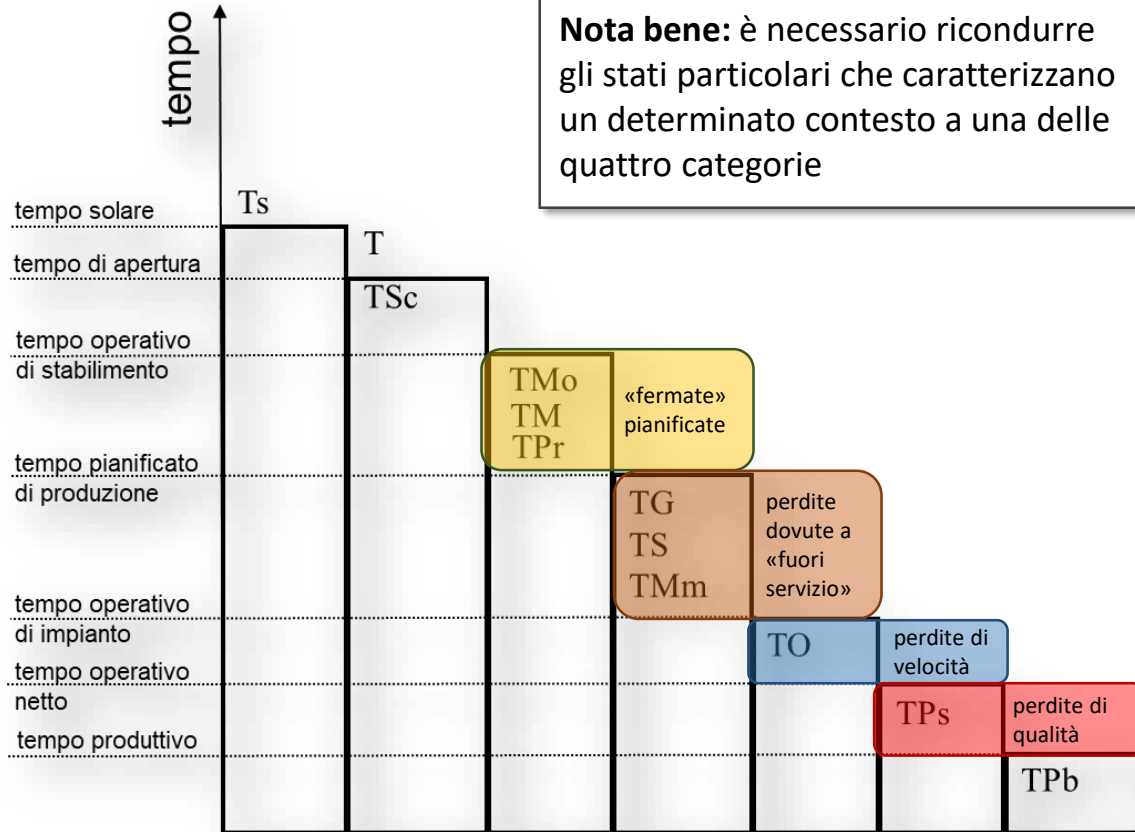
Descriptive data analytics – applicazioni

- In sintesi, le principali applicazioni in ambito operations riguardano:
 - calcolo delle misure di prestazione di efficienza (utilizzo e rendimento degli impianti, OEE, utilizzo e rendimento della manodopera, resa dei materiali);
 - calcolo delle misure di prestazione di efficacia di sistemi produttivi operanti su commessa (ritardo pianificato e non pianificato delle diverse fasi del processo di soddisfazione dell'ordine cliente, lead time delle fasi);
 - calcolo delle misure di prestazione di efficacia di sistemi produttivi operanti per il magazzino (incidenza del potenziale stock-out, persistenza del potenziale stock-out, incidenza delle righe non evase in prima consegna, ...);
 - media e deviazione standard dei tempi ciclo per tipologia di item/risorsa;
 - analisi ABC riguardanti giacenze, consumi, difetti, fermi macchina, ...;
 - analisi di densità riguardanti quantità vendute, numero di spedizioni, missioni di logistica interna, ...;
 - visualizzazione grafica di costi di produzione, tariffe di trasporto, ...;
 - matrici origine-destinazione;
 - correlazione dei dati di domanda per identificare le componenti della domanda stessa.

Descriptive data analytics – OEE

- Bisogna partire dagli stati in cui un impianto può trovarsi:
 - T, tempo solare di apertura dello stabilimento in cui l'impianto è inserito (tempo potenzialmente produttivo)
 - TPb, tempo che, in T, l'impianto ha trascorso nello stato «produzione buona»
 - TPs, tempo che, in T, l'impianto ha trascorso nello stato «produzione di scarto»
 - TS, tempo che, in T, l'impianto ha trascorso nello stato «set-up»
 - TPr, tempo che, in T, l'impianto ha trascorso nello stato «produzione per prove»
 - TG, tempo che, in T, l'impianto ha trascorso nello stato «fermo per guasti»
 - TM, tempo che, in T, l'impianto ha trascorso nello stato «fermo per manutenzione programmata»
 - TMo, tempo che, in T, l'impianto ha trascorso nello stato «fermo per mancanza ordini»
 - TMM, tempo che, in T, l'impianto ha trascorso nello stato «fermo per mancanza materiali»
 - TSc, tempo che, in T, l'impianto ha trascorso nello stato «fermo per scioperi»
 - TO, tempo che, in T, l'impianto ha trascorso nello stato «micro-fermo organizzativo»

Descriptive data analytics – OEE



Descriptive data analytics – OEE

- Il calcolo dell'OEE:

$$\begin{aligned}
 OEE &= \frac{\text{DISPONIBILITA' (A)} \\ &\quad \text{Operating Time}}{\text{Planned production Time}} \times \frac{\text{PERFORMANCE (P)} \\ &\quad \text{Net Operating Time}}{\text{Operating Time}} \times \\ &\quad \times \frac{\text{Fully Productive Time}}{\text{Net Operating Time}} = \\ &\quad \text{QUALITA' (Q)} \\ &= \frac{\text{Fully Productive Time}}{\text{Planned production Time}}
 \end{aligned}$$

Descriptive data analytics – OEE

- Il calcolo dell'OEE:

$$\begin{aligned}
 A(OEE) &= \frac{\text{Operating Time}}{\text{Planned Production Time}} = \\
 &= \frac{T - TSc - TMo - TMm - TG - TM - TPr - TS}{T - TSc - TMo - TM - TPr}
 \end{aligned}$$

Descriptive data analytics – OEE

- Il calcolo dell'OEE:

$$\begin{aligned}
 P(OEE) &= \frac{\text{Net Operating Time}}{\text{Operating Time}} = \\
 &= \frac{\sum_i (\overline{TPb}_i + \overline{TPs}_i)}{T - TSc - TMo - TMm - TG - TM - TPr - TS}
 \end{aligned}$$

Descriptive data analytics – OEE

- Il calcolo dell'OEE:

$$Q(OEE) = \frac{\text{Fully Productive Time}}{\text{Net Operating Time}} =$$
$$= \frac{\sum_i \overline{TPb_i}}{\sum_i \left(\overline{TPb_i} + \overline{TPs_i} \right)}$$

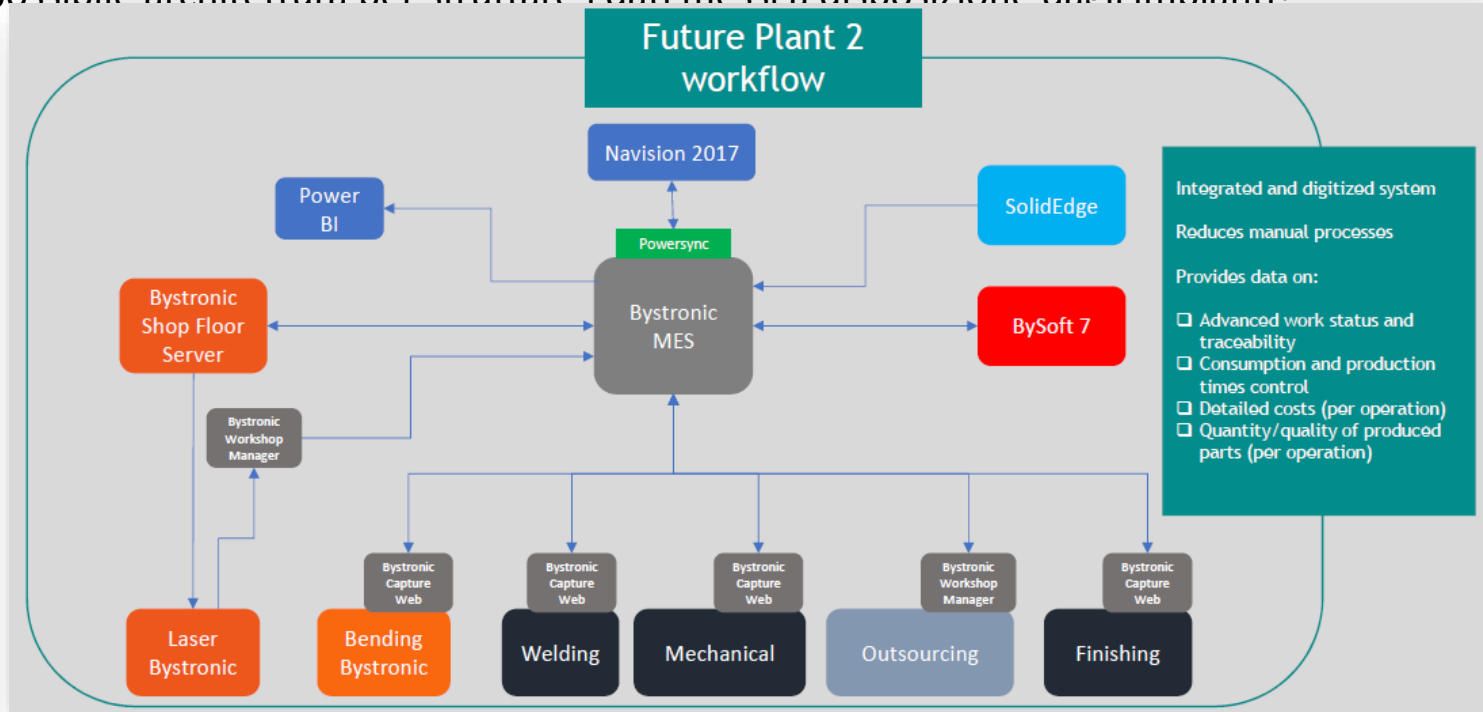
Descriptive data analytics – OEE

Pilastro	Caratterizzazione	Costruttori	Clieni
Sistemi di produzione avanzati	<p>attenzione: è importante definire congiuntamente con l'azienda fornitrice dell'impianto i dati che l'impianto stesso deve mettere a disposizione</p>		●
Manifattura additiva			●
Realtà aumentata			●
Simulazione			●
Di processo		○	●
Integraz	<p>tutto sommato ragionevole:</p> <ul style="list-style-type: none"> • impossibilità di fornire una visione integrata <ul style="list-style-type: none"> • rischio di confrontare «pere con mele» • necessità di «remotare» il dato 		●
Industri			●
Cloud			●
			○
Dati a disposizione del cliente mediante protocolli standard			●
Descriptive data analytics (OEE)		●	●

Legenda - livello di importanza: alto ●, medio-alto ●, medio ●, basso ●, nullo ○

Descriptive data analytics – OEE

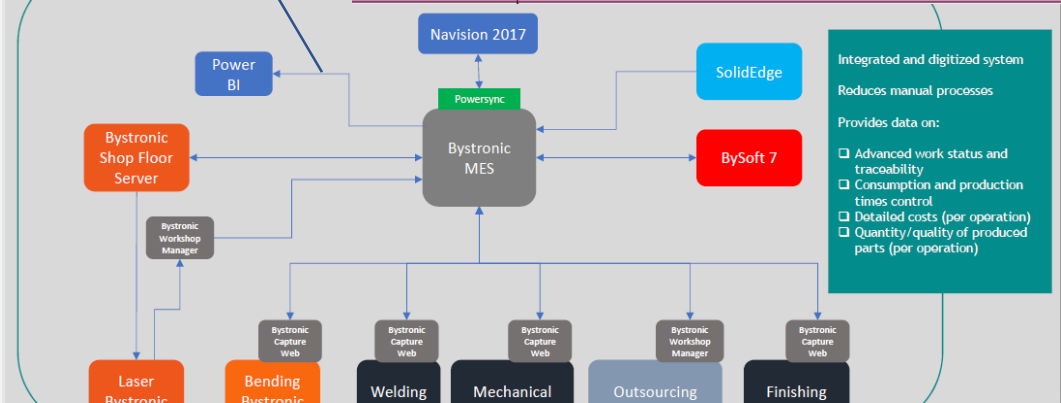
- Quale possibile architettura per sfruttare i dati messi a disposizione dagli impianti?



Descriptive data analytics – OEE

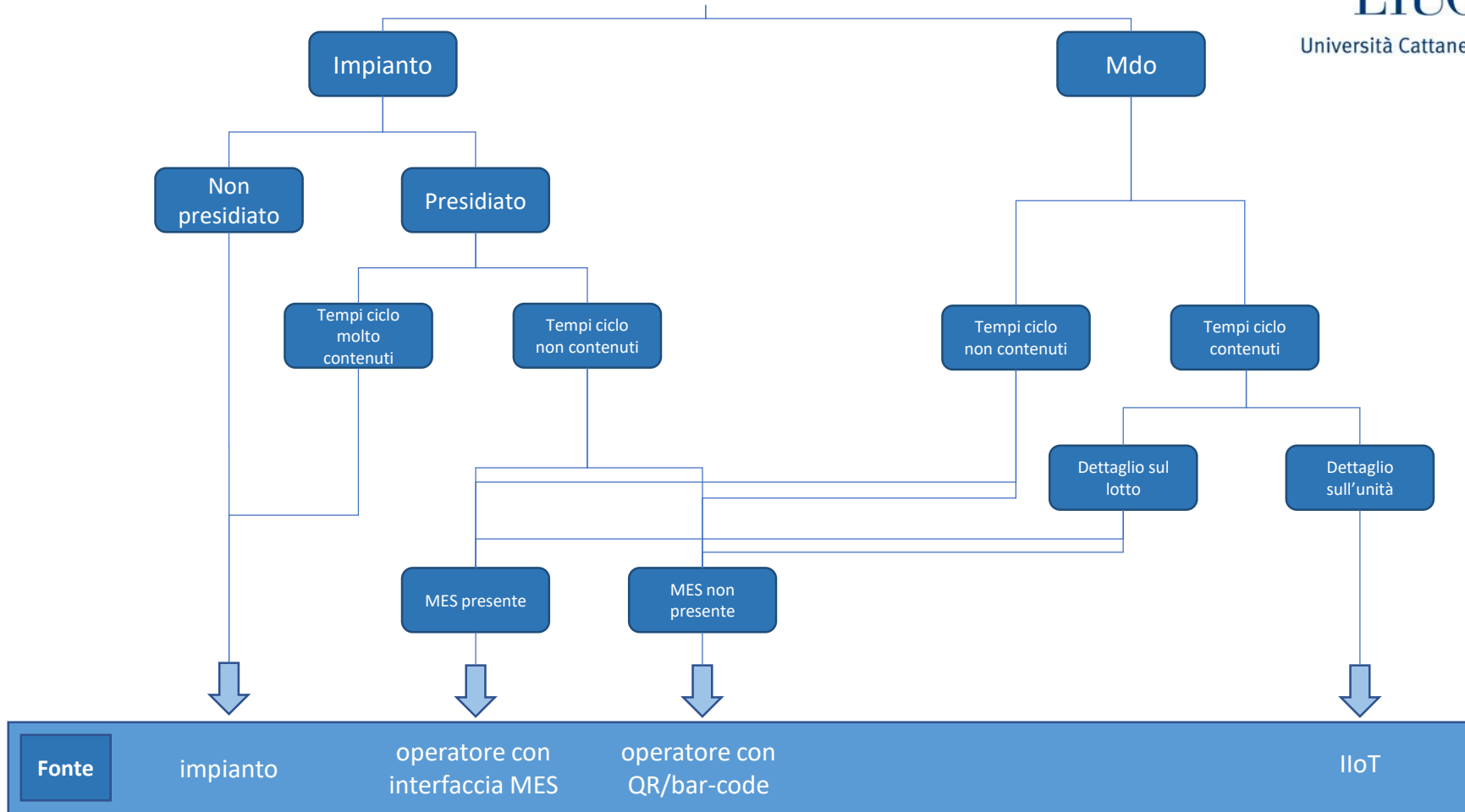
vanno definiti il formato dei dati che si vogliono ottenere dal MES e il metodo di scambio (es. web services, file csv, ecc.)

Macro-caratteristica	Caratteristica	Descrizione	Popolarità	Must have ● Should have ●
collegamento macchine	schede I/O	rilevazione segnali elettrici che coincidono con gli stati della macchina	0%	●
	PLC	creazione di interfacce con i PLC delle macchine al fine di importare, controllare e monitorare alcuni dati di produzione	0%	●



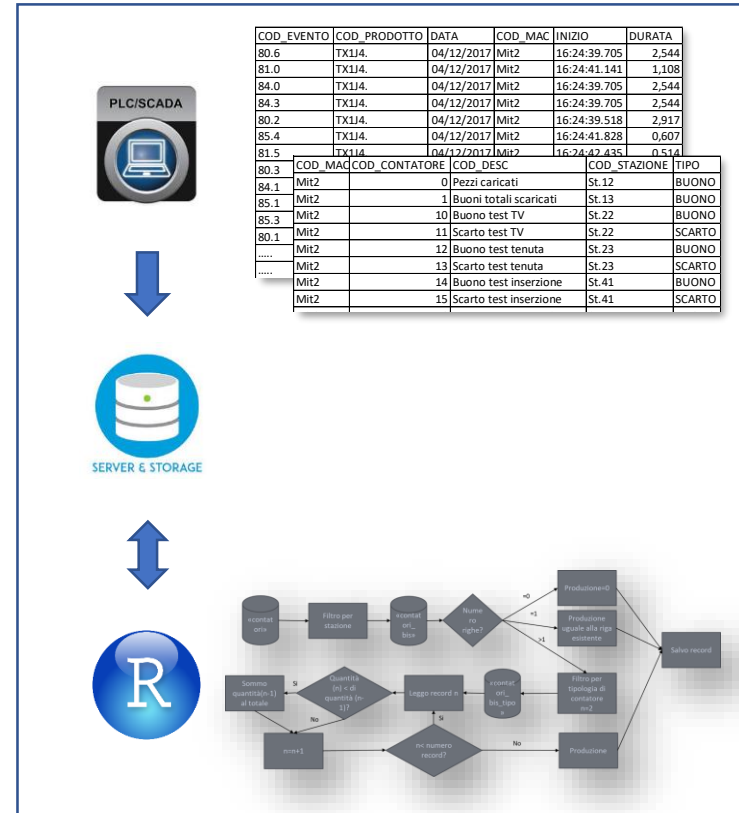
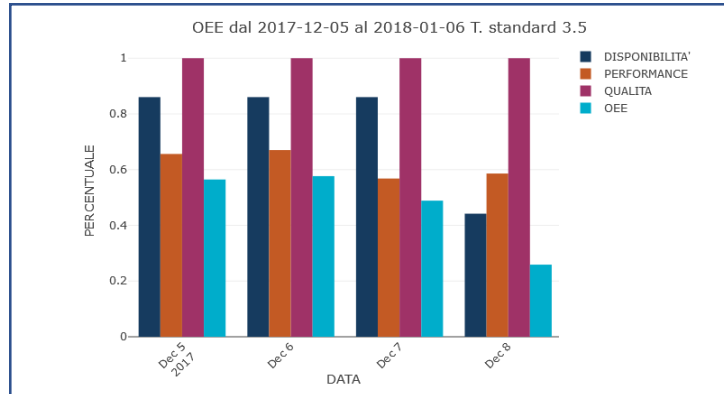
Caratteristica	Descrizione	Popolarità	Must have ● Should have ●
elaborazione KPI	capacità del sistema di effettuare in modo automatico del data analytics e restituire all'utente gli specifici KPI	8%	●

Elemento da monitorare



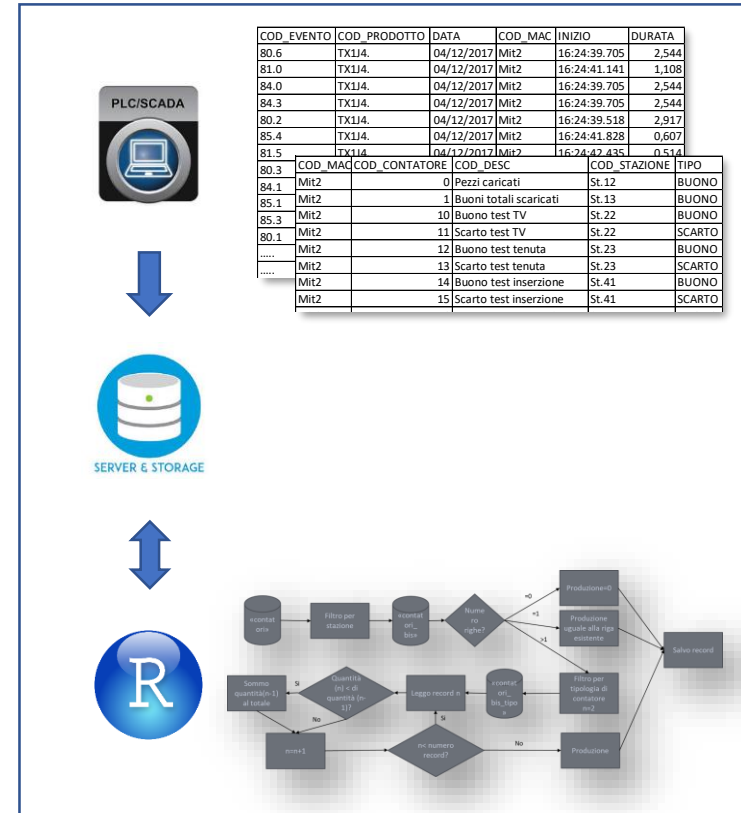
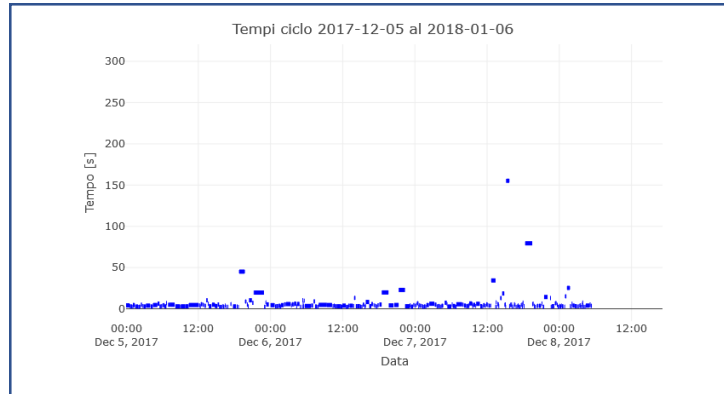
Descriptive data analytics – OEE (esempio)

Obiettivo: individuare mediante una misura di sintesi l'efficienza delle diverse stazioni che compongono una tavola rotante per l'assemblaggio automatico di componentistica per il settore del «bianco». Il fine ultimo era quello di identificare le stazioni a più bassa efficienza in modo tale da guidare azioni di miglioramento che, incrementando la loro efficienza, portassero ad aumentare l'efficienza dell'intero impianto



Descriptive data analytics – esempi

Obiettivo: calcolare i tempi ciclo delle diverse stazioni che compongono una tavola rotante per l'assemblaggio automatico di componentistica per il settore del «bianco». Il fine ultimo era quello di confrontare tali tempi ciclo per individuare le stazioni collo di bottiglia in modo tale da guidare azioni di miglioramento che, riducendo il loro tempo ciclo, portassero ad aumentare il throughput dell'intero impianto



Descriptive data analytics – OEE (esempio)

OBIETTIVO: definizione dell'obiettivo che si vuole raggiungere

Ex. Operations

METODOLOGIA: individuazione di quale tecnica/tecnologia adottare per raggiungere l'obiettivo

Ex. Operations

Ex. Analista

DATI: individuazione del minimo numero di dati da raccogliere

Ex. Operations

Ex. Macchina

Ex. Analista

RACCOLTA: selezione della metodologia attraverso la quale è possibile ottenere i dati individuati in precedenza

Ex. Macchina

DECODIFICA: associazione ad ogni dato un valore intellegibile dagli operatori

Ex. Operations

Ex. Macchina

Ex. Analista

ANALISI: sviluppo di strumenti per analisi l'automatica e in tempo reale dei dati raccolti ed evidenziazione dei risultati

Ex. Operations

Ex. Analista

Software per descriptive data analytics

CARATTERISTICHE PER LA SELEZIONE

- Necessità dichiarate dai futuri utilizzatori

REQUISITI



- Dipendente dal numero di features richieste dagli utilizzatori

COSTO



- La provenienza dei dati da analizzare:
 - Excel file
 - Database
 - ..

DATA SOURCE



- Aggiornamento istantaneo e automatico
- Collegamento con cloud data source

INTEGRAZIONE



- Numero di utilizzatori
- Numero di fattori
- Crescita con la società

SCALABILITA'

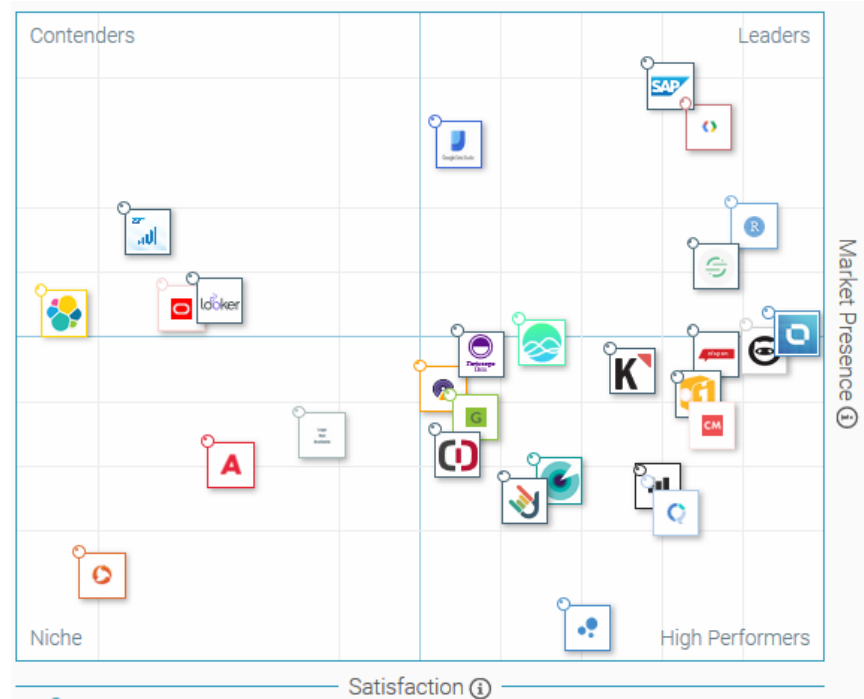


- Dinamicità dei dati e delle analisi che si vogliono svolgere
- Profondità delle analisi che si vogliono condurre

COMPLESSITA'

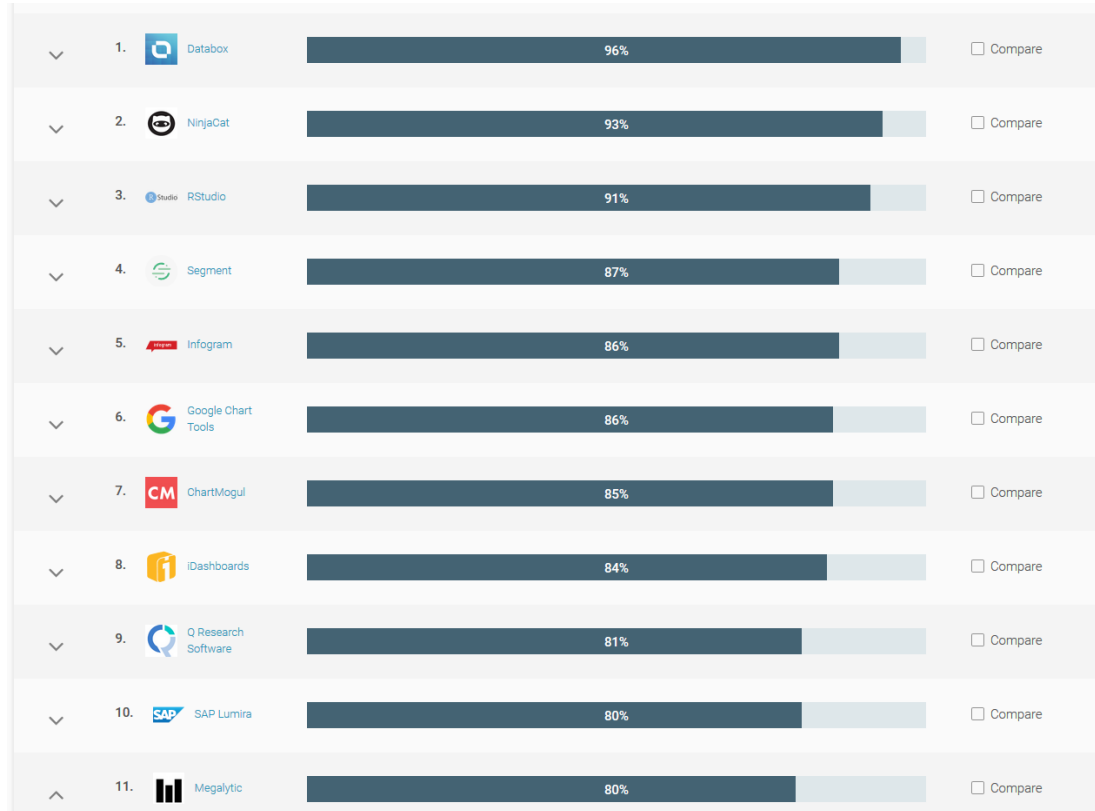


Software per descriptive data analytics

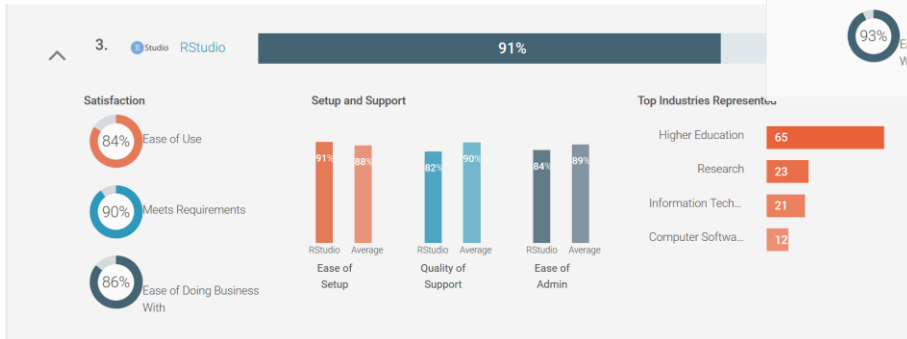
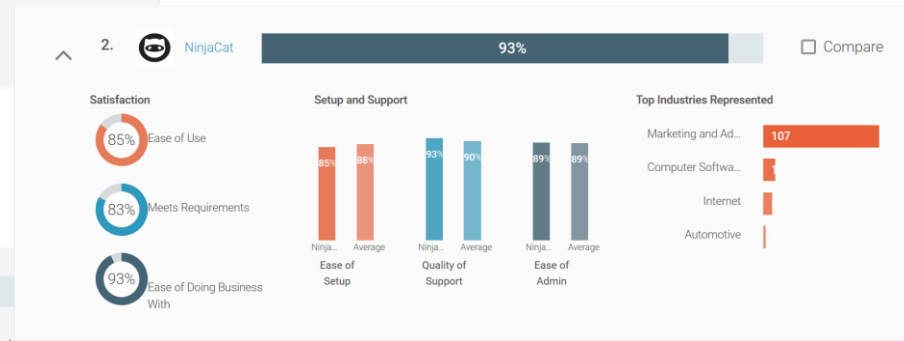
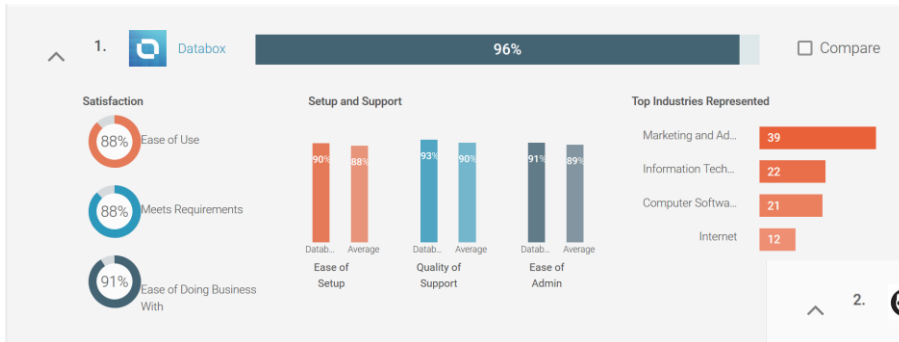


- Market Presence: calcolato attraverso oltre 10 differenti «social source» che forniscono indicazioni su market share, dimensione dei vendor e impatto sociale
- Satisfaction: calcolato utilizzando i giudizi dati da reali utilizzatori





















Software per descriptive data analytics



Software per descriptive data analytics



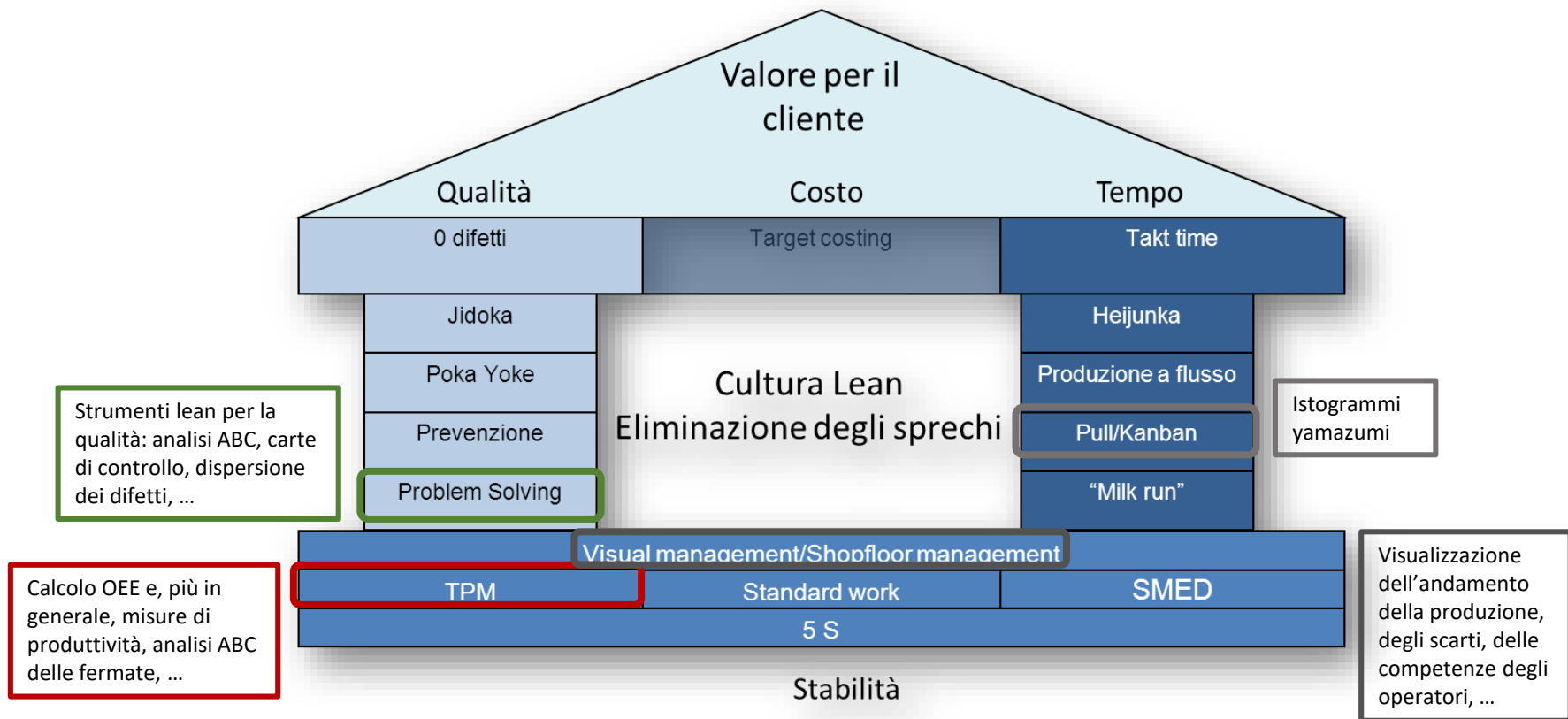
Software per descriptive data analytics

1.	 Databox		8.7	<input type="checkbox"/> Compare
2.	 Segment		8.7	<input type="checkbox"/> Compare
3.	 NinjaCat		8.6	<input type="checkbox"/> Compare
4.	 ChartMogul		8.5	<input type="checkbox"/> Compare
5.	 Megalytic		8.3	<input type="checkbox"/> Compare
6.	 Google Chart Tools		8.3	<input type="checkbox"/> Compare
7.	 RStudio		8.2	<input type="checkbox"/> Compare
8.	 Geckoboard		8.2	<input type="checkbox"/> Compare
9.	 iDashboards		8.1	<input type="checkbox"/> Compare
10.	 ClicData		8.0	<input type="checkbox"/> Compare

Software per descriptive data analytics



Descriptive data analytics & lean



Predictive data analytics

- Sulla base di dati storici (per esempio, esiti di esperimenti) formulare un modello che consenta di predire il comportamento di un sistema

Data mining vs Predictive data analytics

- Il data mining è la disciplina che sfrutta tecniche statistiche e di machine learning e che è finalizzata a estrarre nuova conoscenza dalla base di dati, conoscenza rappresentata da:
 - pattern (regolarità)
 - predittori
- Il predictive data analytics è l'insieme delle tecniche di data mining utili alla predizione di eventi o di comportamenti sulla base di dati storici, utili a estrarre predizioni accurate da una base di dati
- In virtù di quanto sopra riportato, da qui in poi si farà riferimento al termine più generale di data mining

Data mining

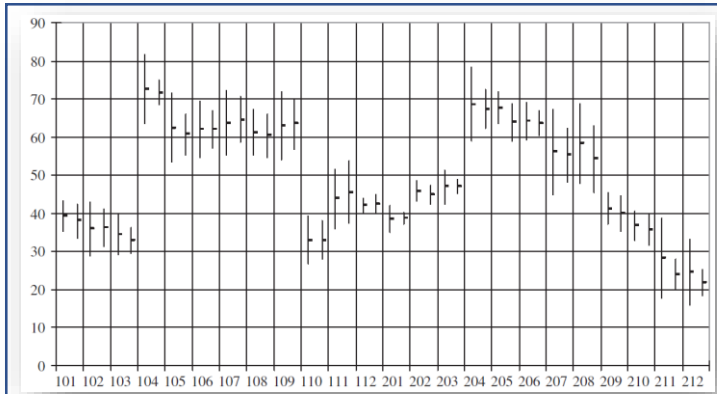
- Il data mining, a seconda delle tecniche utilizzate, permette di effettuare due tipi di analisi:
 - Interpretazione: cioè l'identificazione di regolarità o pattern nei dati e la rappresentazione degli stessi come attraverso regole e criteri che gli utilizzatori possono comprendere
 - Predizione: cioè anticipare il valore di una variabile casuale nel futuro o stimare la probabilità che si verifichi un evento
- Più in dettaglio, il data mining consente di affrontare i seguenti problemi:
 - Clustering: dividere i dati in gruppi omogenei
 - Regressione: creare funzioni che emulano l'andamento dei dati
 - Classificazione: associare ai dati una classe tra quelle disponibili
 - Regole associative: trovare relazioni tra gruppi di dati

Data mining

- Le fasi del data mining sono le seguenti:
 - Scelta della tecnica: a seconda dell'obiettivo che ci si è posti si sceglie il tipo di tecnica da utilizzare (tecniche di interpretazione vs tecniche di predizione)
 - Scelta del modello: spesso non si sa a priori quale modello dà risultati più precisi, quindi nella pratica se ne provano diversi
 - Inizializzazione, simulazione e selezione della parametrizzazione del modello: si divide il set di dati, una parte si usa per «allenare» il modello, cioè fargli imparare le caratteristiche del fenomeno, un'altra parte si usa per testare l'accuratezza del modello. Si sceglie quindi il migliore (non sempre questa fase risulta necessaria)
 - Utilizzo del modello: si utilizza il modello per calcolare le caratteristiche che si vuole studiare su dati inediti

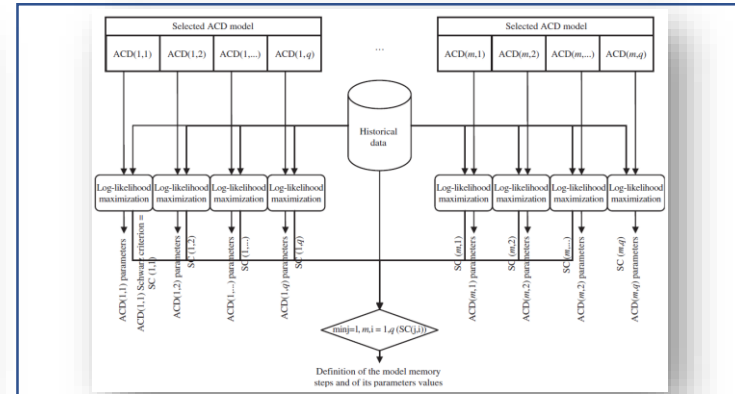
Data mining – esempi

Obiettivo: prevedere l'occorrenza dei minor stoppages (rotture delle fibre di vetro) nei reparti di fibratura/bobinatura di un'azienda produttrice di rocche di fibre di vetro. Il fine ultimo era quello di popolare con tali previsioni un modello di simulazione per dimensionare il numero di operatori necessari in fibratura per ridurre il più possibile le perdite di vetro fuso



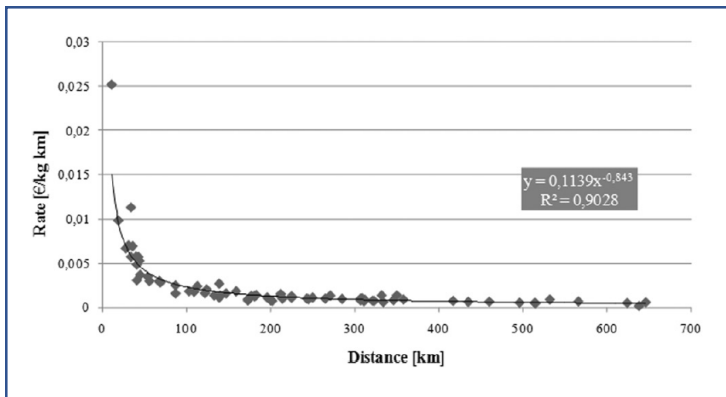
Variable of interest	Duration
Frequency	High/Very high
Intervals	Irregular
Series	Auto-correlated
Typifying phenomenon	Clustering

$$\psi_i = \omega + \sum_{j=1}^m \alpha_j x_{i-j} + \sum_{j=1}^q \beta_j \psi_{i-j}.$$



Data mining – esempi

Obiettivo: stimare la tariffa di trasporto [€/kg*km] all'interno di una determinata regione europea al variare dell'origine dei flussi di spedizione. I dati di partenza erano rappresentati: (i) dalle tariffe per fasce di peso applicate dal trasportatore tra l'origine attuale e le diverse aree in cui la regione europea considerata era suddivisa, (ii) le quantità trasportate dall'origine attuale e le diverse aree.

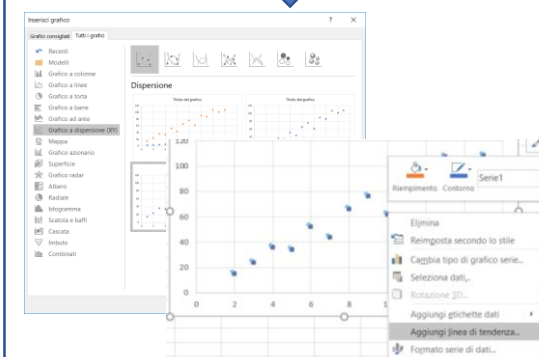


Tariffs (€/kg) Destinations (zip code)	Distance (km)	Weight ranges (kg)										
		≤20	20-50	50-100	100-200	200-300	300-500	500-1000	1000-2000	2000-3000	3000-5000	> 5000
Wien (10)	28	1.50	1.04	0.80	0.69	0.61	0.56	0.34	0.19	0.15	0.11	0.10
Wien (11)	31	1.50	1.04	0.80	0.69	0.61	0.56	0.34	0.19	0.15	0.11	0.10
Wien (12)	34	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (20)	56	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (21)	44	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (22)	54	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (23)	43	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (24)	40	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (25)	11	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (26)	19	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (27)	36	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11
Stockerau (28)	41	1.75	1.21	0.93	0.80	0.71	0.66	0.39	0.25	0.19	0.14	0.11

↓ Media pesata sulle quantità trasportate

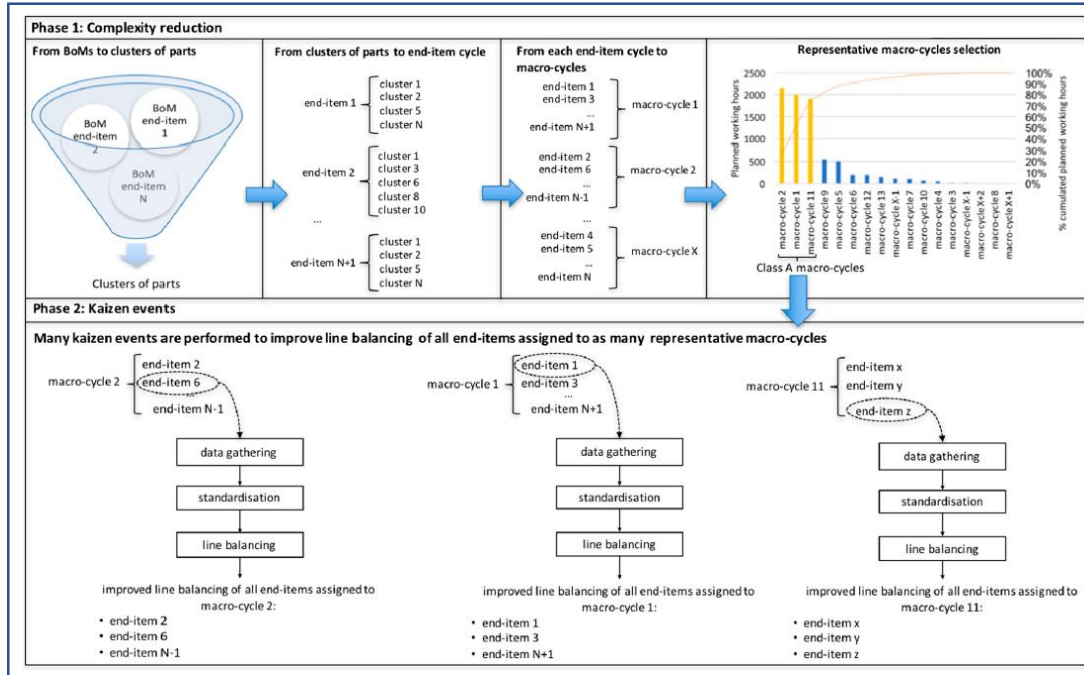
Delivery point code	10	11	12	20	21	22	23	24	25	26	27	28
Weighted	0.5547	0.6446	0.5402	0.5020	0.6903	0.5630	0.7274	0.6826	0.4721	0.5514	0.7386	0.3783
€/kg rate												
Distance (km)	28	31	34	56	44	54	43	40	11	19	36	41
€/kg km rate	0.0198	0.0208	0.0159	0.0090	0.0157	0.0104	0.0169	0.0171	0.0429	0.0290	0.0205	0.0092

↓



Data mining – esempi

Obiettivo: individuare tra più di 300 codici da confezionare su linee semi-automatiche quelli sui quali condurre attività kaizen.



Featuring
 Microsoft®
Visual Basic®
 Technology



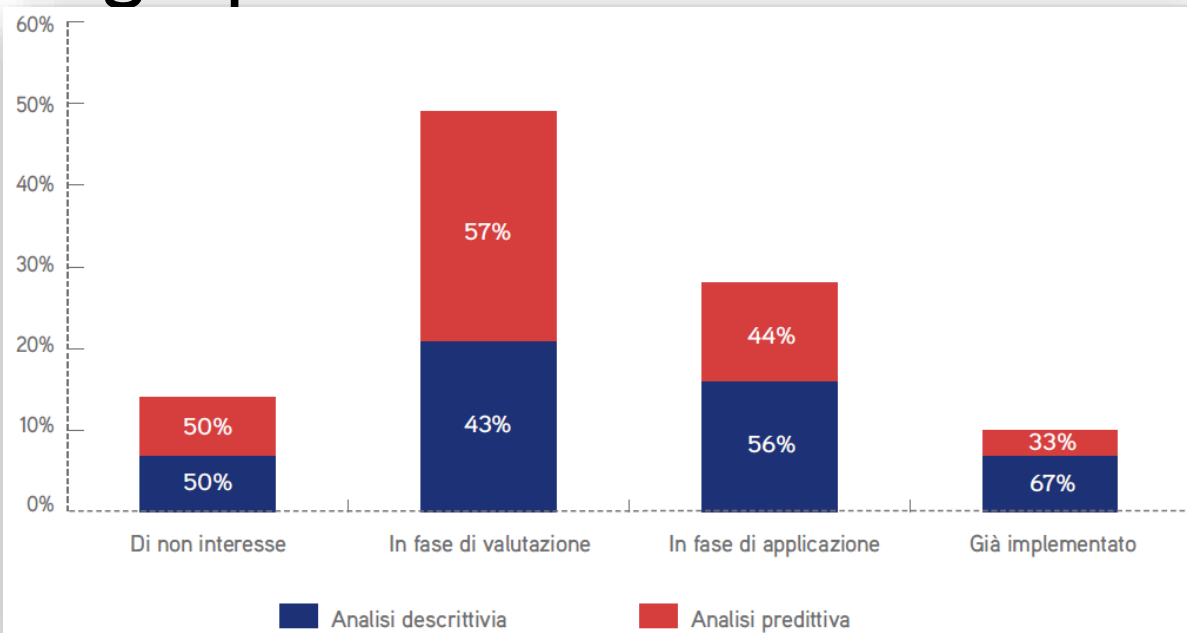
Data mining – applicazioni

- In sintesi, le principali applicazioni in ambito operations riguardano:
 - previsione della domanda (metodi causali e tecniche estrapolative);
 - rappresentazione dell'occorrenza di minor stoppages (metodi causali e tecniche estrapolative, machine learning);
 - stima dei costi di trasporto al variare del punto di origine (regressioni);
 - identificazione delle cause effettive di un determinato fenomeno tra quelle potenziali (tecniche statistiche e regole associative);
 - group technology (tecniche e algoritmi di clustering);
 - manutenzione su condizione e manutenzione predittiva (regole associative, metodi causali e tecniche estrapolative, machine learning);
 - ...

Data mining – predictive maintenance



Data mining – predictive maintenance



standard		
Predictive data analytics per garantire adattività	<input checked="" type="radio"/>	<input type="radio"/>
Predictive data analytics a supporto della manutenzione	<input type="radio"/>	<input checked="" type="radio"/>

Legenda - livello di importanza: alto ●, medio-alto ●, medio ●, basso ●, nullo ○

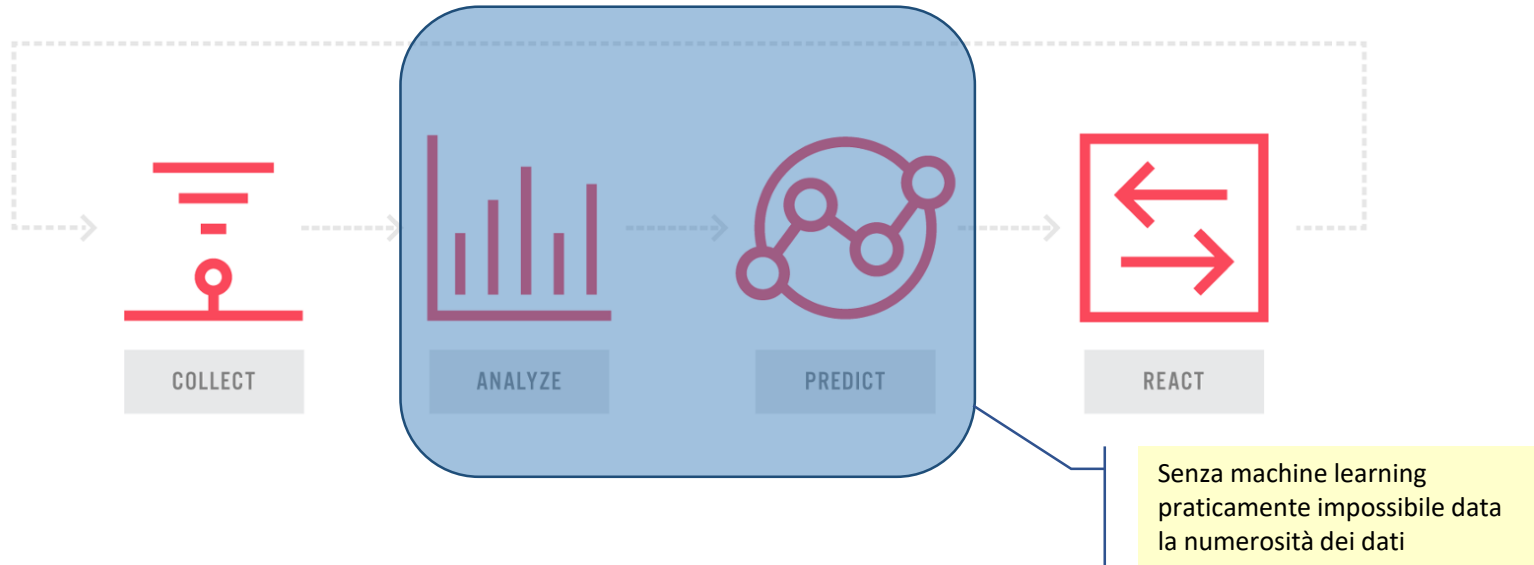
Data mining – predictive maintenance

- Per mettere in piedi un sistema di predictive maintenance è necessario implementare quello che si chiama monitoraggio delle condizioni
- In particolare, le principali tecniche di monitoraggio delle condizioni riguardano:
 - vibrazioni
 - termografia
 - lubrificanti
 - assorbimento energia
 - aspetti visivi
 - ...

I tipi di condizioni da monitorare dipendono dalle tipologie di componenti dell'impianto per il quale si vuole implementare il sistema di predictive maintenance. È opportuno impostare prima una **FMECA**

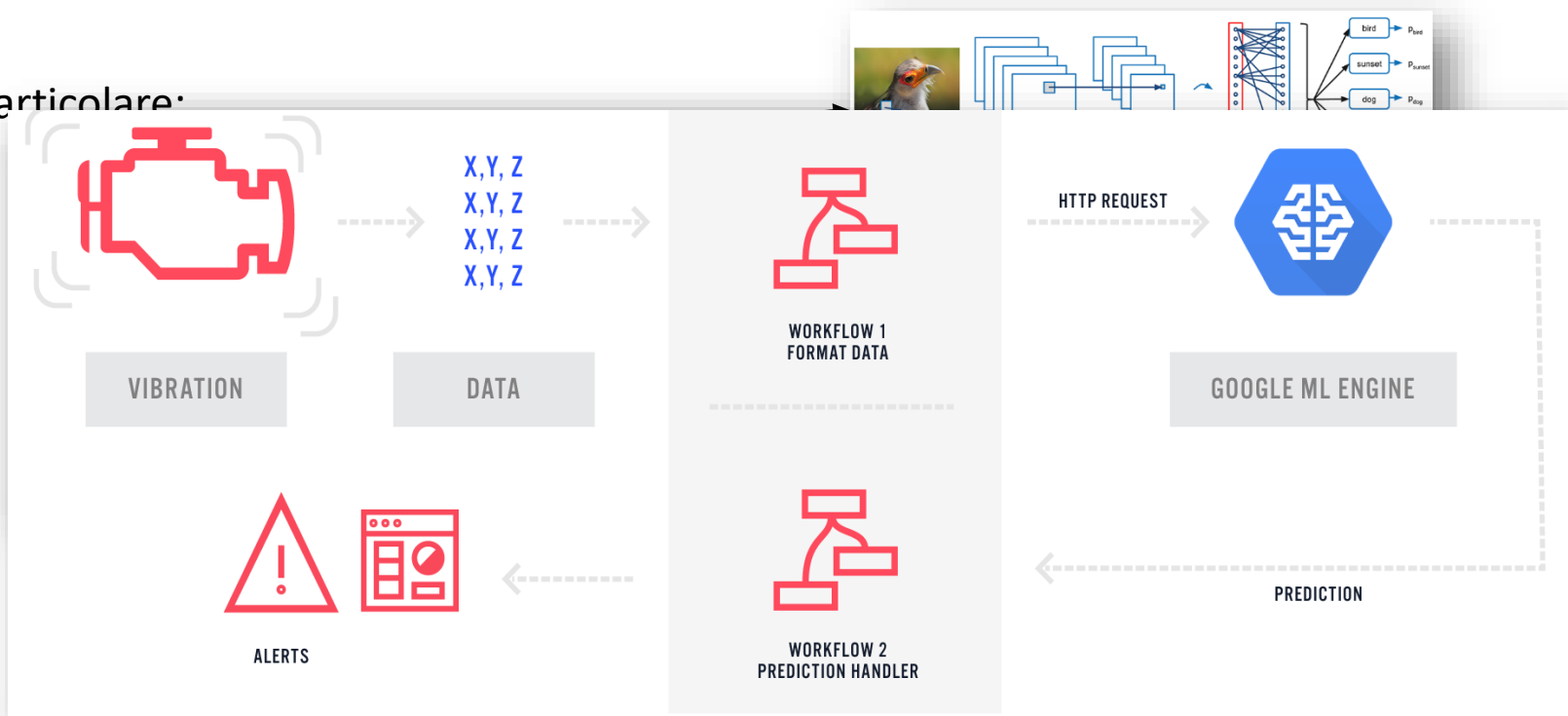
Data mining – predictive maintenance

- Una volta che il monitoraggio delle condizioni è attivo si riceveranno milioni di dati e da questi si dovrà capire quale fattore influenza realmente lo stato dell'impianto

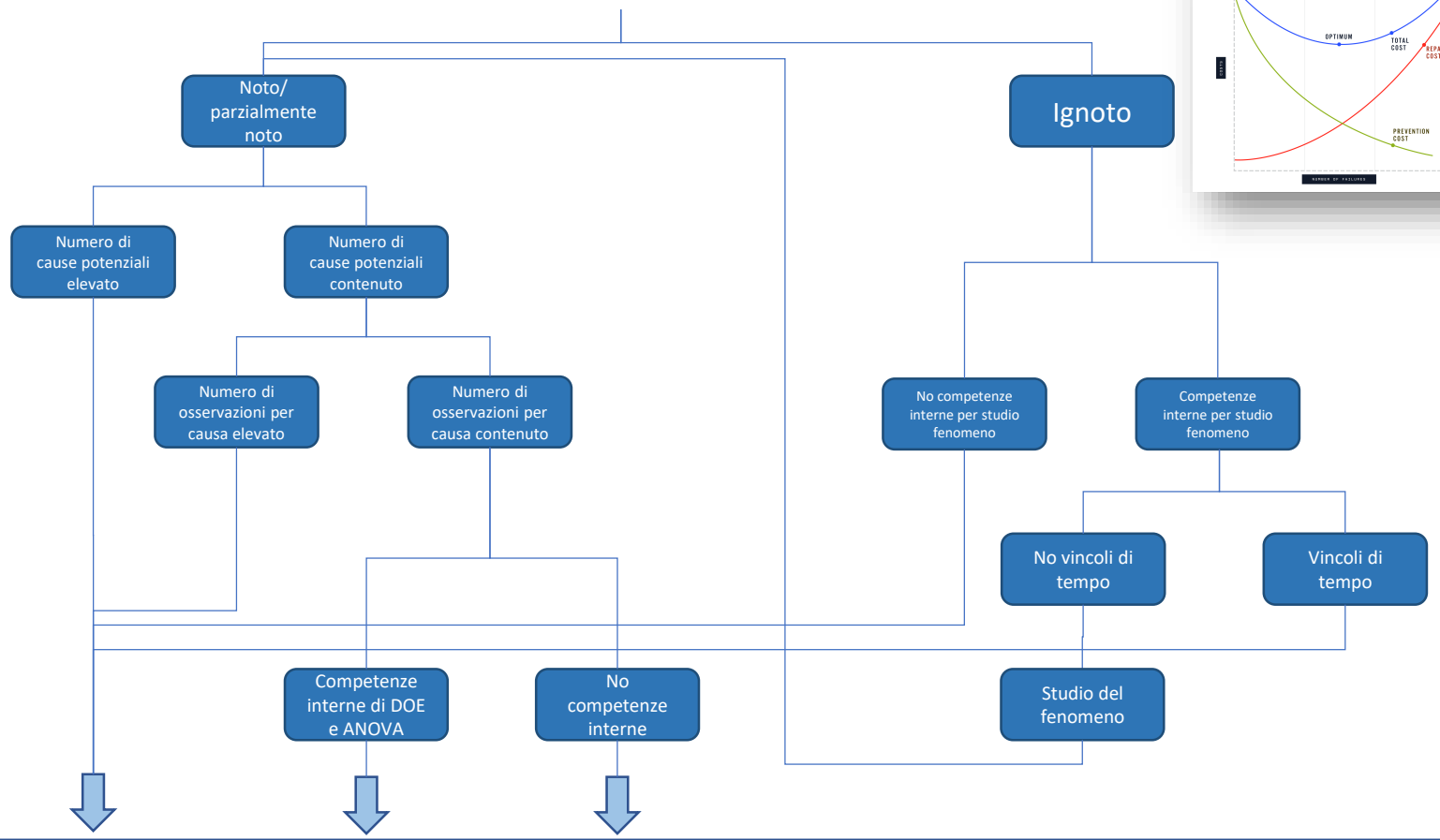
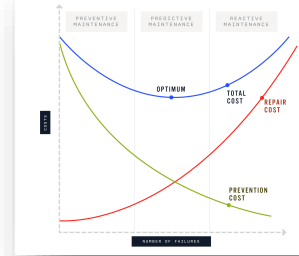


Data mining – predictive maintenance

- In particolare:



Fenomeno



Fonte	machine learning (terza parte)	DOE/ANOVA	DOE/ANOVA (terza parte)
--------------	--------------------------------	-----------	-------------------------

Data mining – predictive maintenance



- Responsabile del miglioramento sia termini di performance che di costi
- Autorità a intervenire nei diversi livelli
- Deve avere conoscenze sulle macchine e sulla loro progettazione
- Deve essere in grado di applicare le «best practice» sia in ambito di manutenzione che di operations
- Avere conoscenza delle tecniche di manutenzione predittiva

Data mining – predictive maintenance

SUPER STAR

- Selezionare tra il personale presente in aziende chi ha una profonda conoscenza in uno o più degli ambiti sopra indicati
- Verificare le effettive conoscenze delle persone selezionate
- Poiché poche persone avranno conoscenze avanzate sulla manutenzione predittiva è necessario una formazione di tutto il team selezionato
- Per svolgere le analisi è consigliato utilizzare una root-cause analisi

HIRING

- Assumere un ingegnere esperto in ambito affidabilità
- Solitamente coloro che possiedono queste capacità ingegneri specializzati che richiedono un alto salario
- Ci sono pochi ingegneri veramente esperti in questo campo
- Solitamente preferiscono un rapporto consulenziale rispetto a un rapporto duraturo

Data mining – previsione (esemp

Obiettivo: stimare l'andamento della domanda dei diversi codici venduti da un'azienda distributrice di prodotti ottici dando anche un range di domanda minima e domanda massima e rappresentando il tutto graficamente. Il fine ultimo era quello di ridurre sia i costi di mantenimento a scorta, sia gli stockout



SERVER & STORAGE

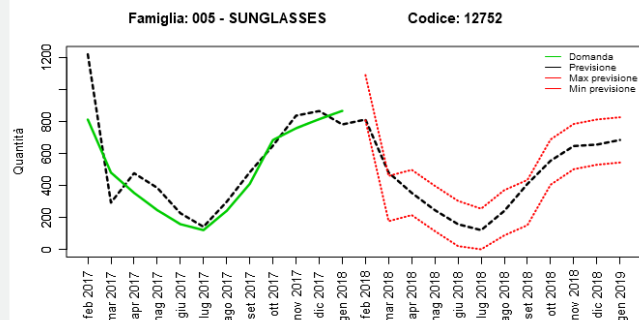


ce	Descrizio	Gruppo	Sotto_gru	Flag	Vendite_2	Vendite_2	Vendite_2	Vendite_2	Vendite_2
10	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	20	2	19	26	17
20	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	4	9	3	8	16
21	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	4	9	0	6	2
22	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	2	15	2	9	6
23	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	2	9	12	22	6
40	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	6	14	7	12	6
41	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	6	14	5	10	16
42	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	7	8	10	13	7
43	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	6	8	9	14	3
44	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	7	3	4	4	24
46	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	7	9	7	4	3
47	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	8	10	5	4	3
48	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	9	10	6	4	4
49	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	7	13	6	3	5
54	CERNIERA	A01 - PAR	A00 - CER	0 - ATTIVO	0	0	0	0	0

Famiglia: 005 - SUNGLASSES

Codice: 12752

Rownames	feb 2017	mar 2017	apr 2017	mag 2017	giu 2017	lug 2017	set 2017	ott 2017	nov 2017	dic 2017	gen 2018	feb 2018	mar 2018	apr 2018	mag 2018	giu 2018	lug 2018	ago 2018	set 2018	ott 2018	nov 2018	dic 2018	gen 2019	
Mn. previsione	0	0	0	0	0	0	0	0	0	0	0	807	176	214	114	20	0	87	151	404	501	529	543	
Previsione	1220	292	476	385	226	141	297	482	648	836	865	781	812	480	353	245	157	120	240	409	554	646	655	684
Max. previsione	0	0	0	0	0	0	0	0	0	0	0	1090	459	497	397	303	254	370	434	687	784	812	826	
Domanda	812	480	353	245	157	120	240	409	683	757	814	866	193	0	0	0	0	0	0	0	0	0	0	0
Domanda master	0	0	0	0	0	0	40	0	818	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0
Domanda spot	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errore	-408	188	-123	-140	-69	-21	-57	-73	35	-79	-51	85	0	0	0	0	0	0	0	0	0	0	0	0



Data mining – previsione (esempio)

OBIETTIVO: definizione dell'obiettivo che si vuole raggiungere

Ex. Acquisti

METODOLOGIA: individuazione di quale tecnica/tecnologia adottare per raggiungere l'obiettivo

Ex. Acquisti

Ex. Analista

DATI: individuazione del minimo numero di dati da raccogliere

Ex. Acquisti

Ex. IT

Ex. Analista

RACCOLTA: selezione della metodologia attraverso la quale è possibile ottenere i dati individuati in precedenza

Ex. IT

DECODIFICA: associazione ad ogni dato un valore intellegibile dagli operatori

Ex. Acquisti

Ex. IT

Ex. Analista

ANALISI: sviluppo di strumenti per analisi l'automatica e in tempo reale dei dati raccolti ed evidenziazione dei risultati

Ex. Acquisti

Ex. Analista

Software per data mining

CARATTERISTICHE PER LA SELEZIONE

- La previsione del futuro deve essere accettata e integrata con informazioni

UTILIZZABILITA'



- Verificare che il costo di implementazione non sia maggiore dei benefici

COSTO



- Capacità di mostrare la previsione in modo conciso

SINTETICITA'



- Aggiornamento istantaneo e automatico
- Collegamento con cloud data source

INTEGRAZIONE



- Numero di utilizzatori
- Numero di fattori
- Crescita con la società

SCALABILITA'

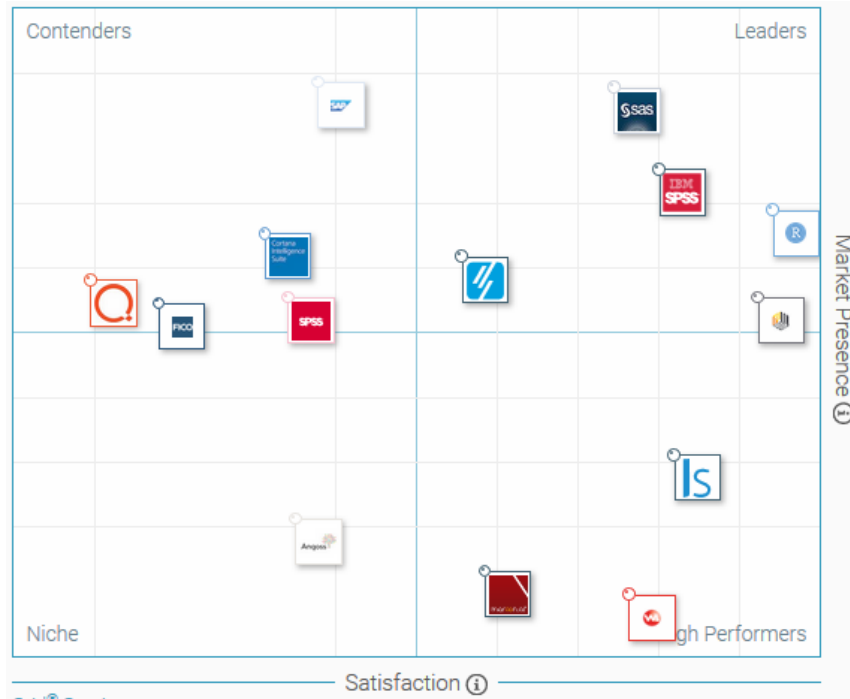


- Verificare la necessità di sviluppare algoritmi o metodologie personalizzate

APERTURA

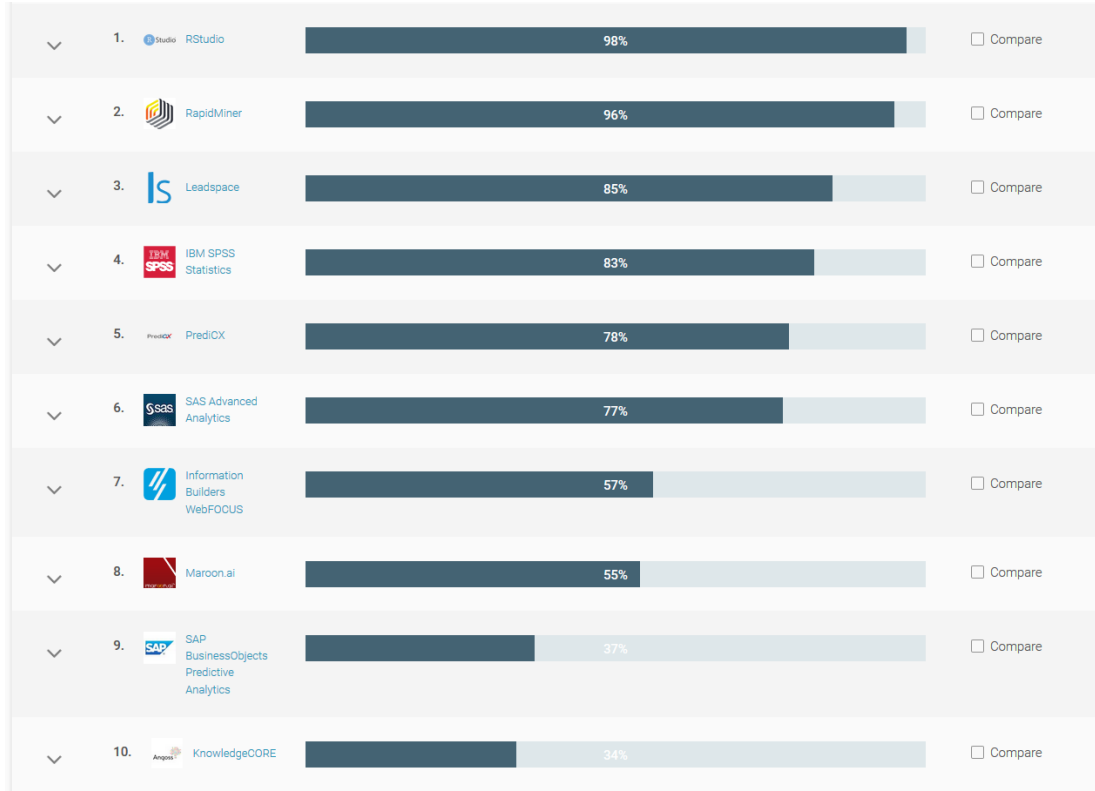


Software per data mining

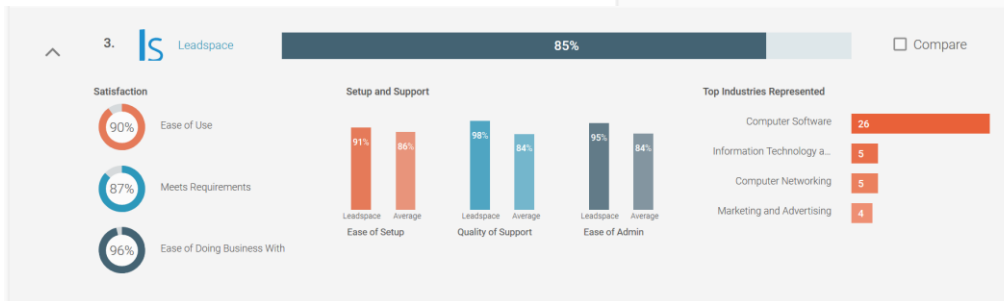
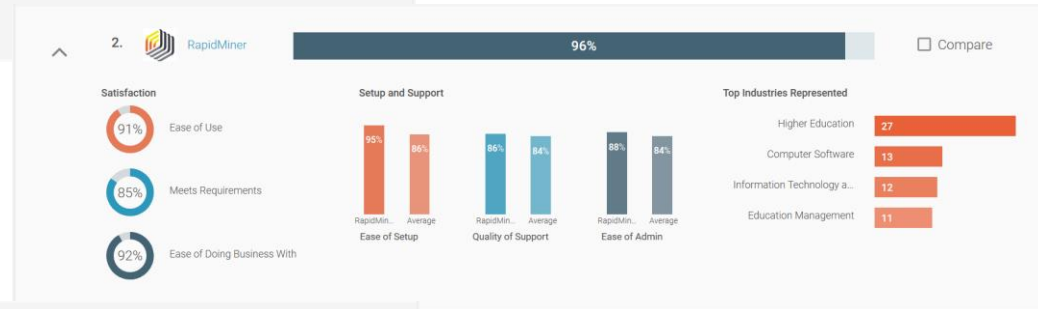
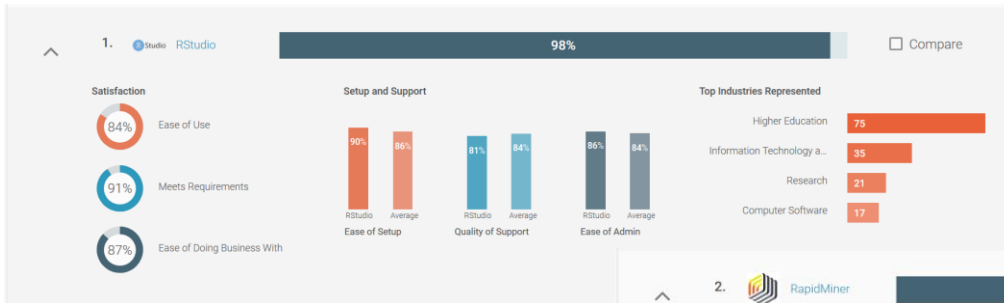


- Market Presence: calcolato attraverso oltre 10 differenti «social source» che forniscono indicazioni su market share, dimensione dei vendor e impatto sociale
- Satisfaction: calcolato utilizzando i giudizi dati da reali utilizzatori

















Software per data mining



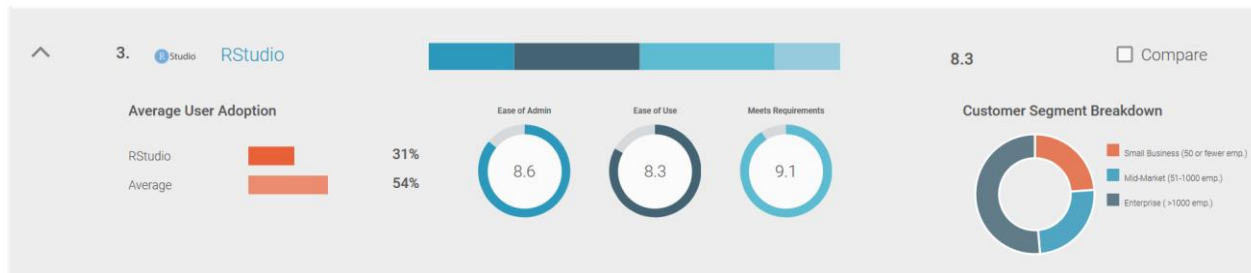
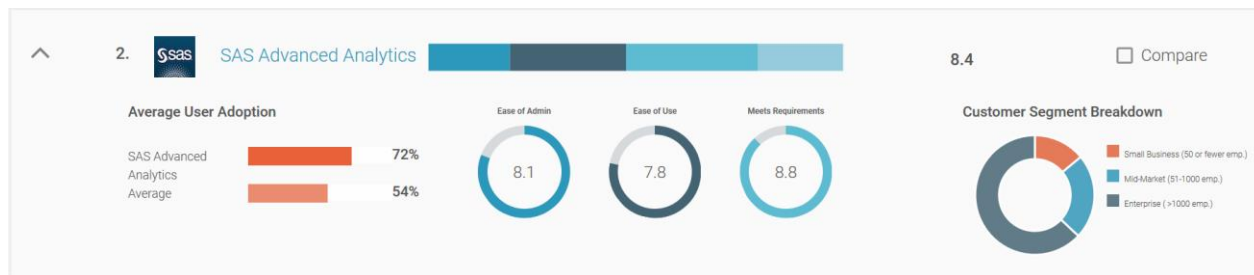
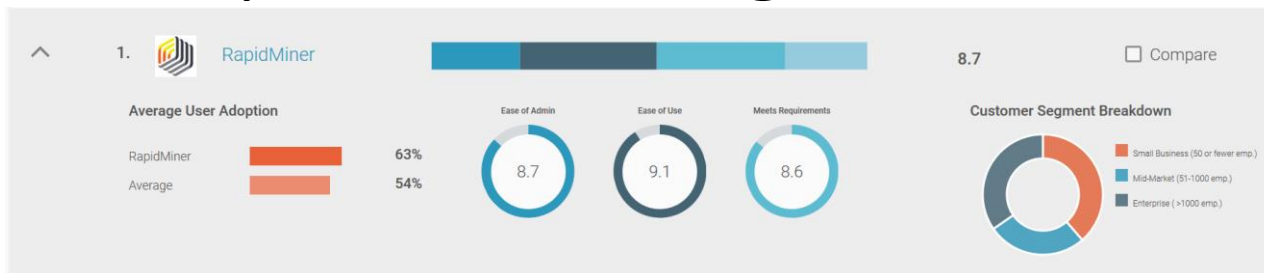
Software per data mining



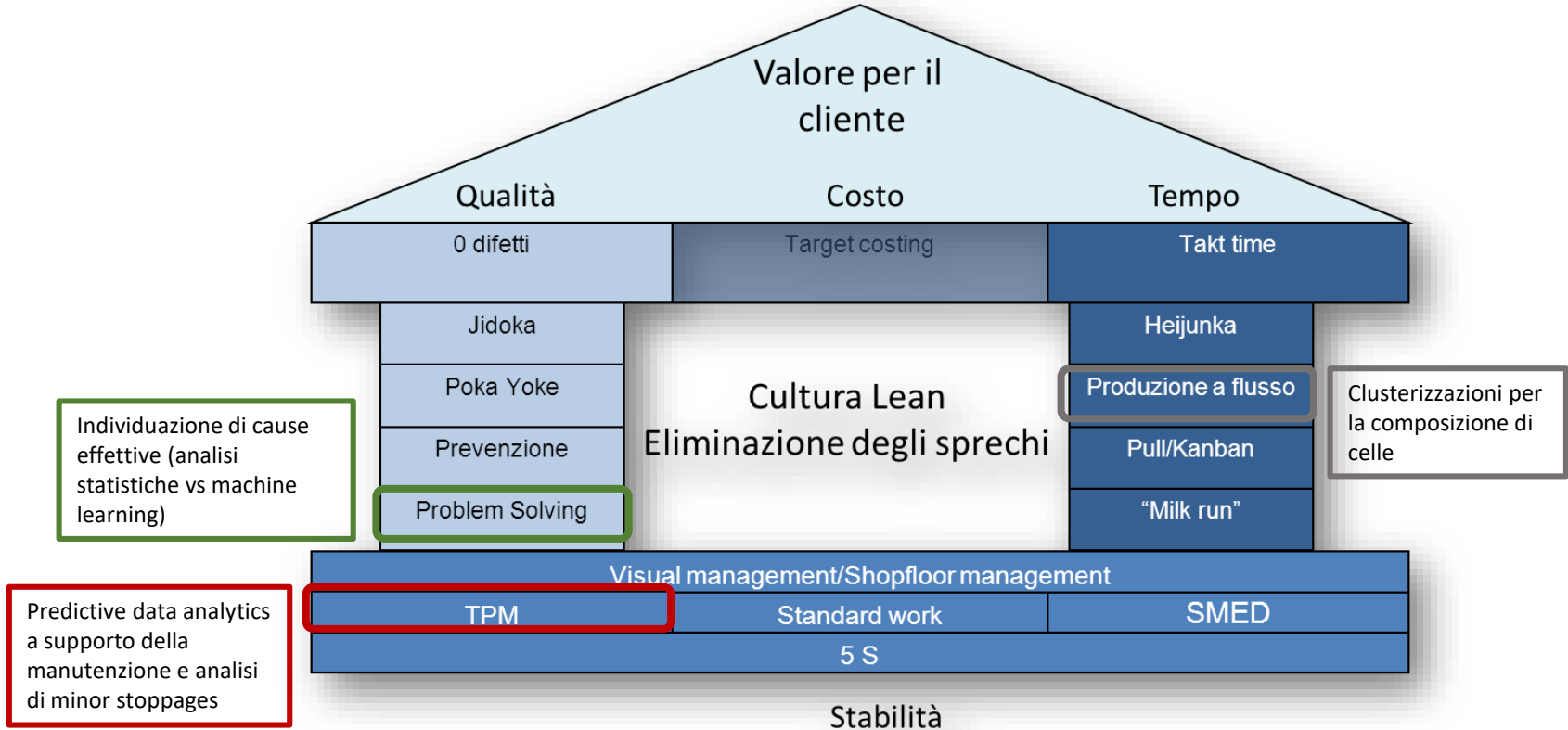
Software per data mining

1.	 RapidMiner		8.7	<input type="checkbox"/> Compare
2.	 SAS Advanced Analytics		8.4	<input type="checkbox"/> Compare
3.	 RStudio		8.3	<input type="checkbox"/> Compare
4.	 LeadSpace		8.2	<input type="checkbox"/> Compare
5.	 IBM SPSS Statistics		8.1	<input type="checkbox"/> Compare
6.	 Maroon.ai		7.7	<input type="checkbox"/> Compare
7.	 SAP BusinessObjects Predictive Analytics		7.6	<input type="checkbox"/> Compare
8.	 Information Builders WebFOCUS		7.6	<input type="checkbox"/> Compare

Software per data mining



Data mining & lean



Prescriptive data analytics

- Il prescriptive data analytics è l'insieme di tecniche e di strumenti che consentono di prendere, con riferimento a un determinato contesto, le decisioni che ottimizzano gli obiettivi d'interesse
- Sono tecniche e strumenti che permettono di risolvere:
 - Problemi di ottimizzazione vincolata
 - Problemi di ottimizzazione combinatoria
- Si parla di «prescriptive» data analytics perché arriva a prescrivere decisioni che dovrebbero essere prese per raggiungere l'obiettivo
- Le tecniche di prescriptive data analytics sono più complesse rispetto a quelle di descriptive e di predictive data analytics e, quindi, sono meno applicate (complessità, decisione delegata allo strumento che applica la tecnica, necessità di forte personalizzazione)

Prescriptive data analytics

Problemi di ottimizzazione vincolata

- Vengono modellizzati tramite formalismi matematici che rappresentano
 - le decisioni come variabili numeriche (interi, reali, binarie), chiamate variabili decisionali, che sono le leve su cui il decisore può agire con l'obiettivo di trovarne il valore che corrisponde alla scelta migliore
 - le relazioni tra le decisioni come vincoli tra le variabili (equazioni lineari o non lineari) che sono le limitazioni che restringono il campo di esistenza delle variabili ossia il range entro cui sono ammesse le soluzioni. Possono essere :
 - \leq (evidenzia un limite superiore),
 - \geq (evidenzia un limite inferiore),
 - $=$ (evidenzia una relazione fissata tra le variabili)
 - la qualità delle decisioni come una funzione sulle variabili, chiamata funzione obiettivo, che deve essere massimizzata o minimizzata
 - i parametri/coefficienti: sia le relazioni di vincolo (dis/equazioni), sia la funzione obiettivo sono formate anche da parametri e da coefficienti d'impiego che, a differenza delle variabili decisionali, hanno valori fissati assunti con certezza
- I modelli così creati, chiamati anche «programmi matematici» o «programmi a vincoli» vengono risolti da un software, chiamato solver (risolutore), che prescrive i valori che le variabili decisionali devono assumere

Prescriptive data analytics

Problemi di ottimizzazione vincolata

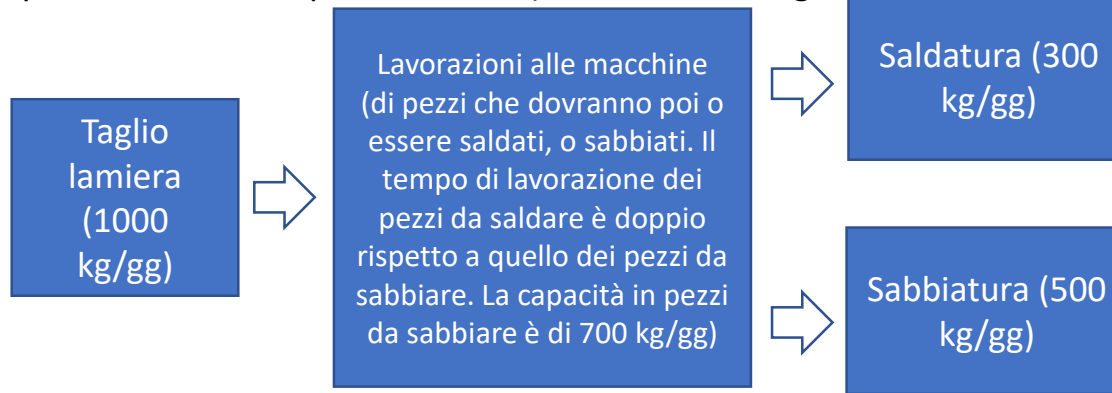
- La modellazione di tali problemi avviene generalmente secondo le seguenti fasi:
 1. Definizione delle variabili decisionali (che cosa bisogna decidere?). Devono essere esplicitate in modo preciso, sia come descrizione che come unità di misura. Per esempio x_1 =numero di pezzi dell'articolo 1 prodotti mensilmente [pz/mese]
 2. Formulazione della funzione obiettivo (che cosa si deve massimizzare o minimizzare?). Definizione di una forma funzionale in termini di combinazione lineare o non lineare delle variabili decisionali
 3. Formulazione delle relazioni di vincolo (cosa limita il valore delle variabili decisionali?). Definizione delle disequazioni o delle equazioni di vincolo identificando i parametri o i coefficienti di impiego per ciascuna variabile decisionale

Prescriptive data analytics

Problemi di ottimizzazione vincolata

- Esempio

- 2 tipologie di prodotti: saldati e piani sabbiati ($mc_{\text{Saldati}}=1 \text{ €/kg}$, $mc_{\text{Sabbiati}}=0.7 \text{ €/kg}$)



- Qual è il mix di Saldati e Sabbiati che massimizza il profitto dell'azienda?

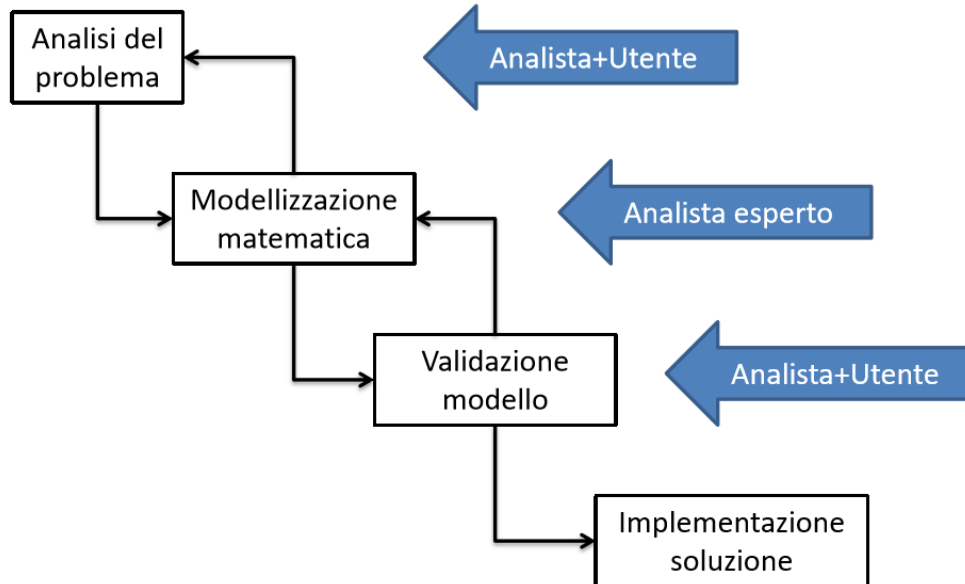
Prescriptive data analytics

Problemi di ottimizzazione combinatoria

- La soluzione di tali problemi è data da una stringa di valori (una combinazione). Un tipico esempio di problema di ottimizzazione combinatoria è il problema del commesso viaggiatore (visitare 5 città A-B-C-D-E nell'ordine che minimizza il tempo di viaggio)
- La risoluzione avviene tramite tecniche che cercano/creano le combinazioni più buone. Il numero di combinazioni generalmente è molto elevato e, di conseguenza, non è possibile esplorarle tutte. L'efficienza dell'algoritmo di risoluzione sta nella capacità di trovare la soluzione facendo il minor numero di tentativi tentativi

Prescriptive data analytics


- Tipicamente uno studio di prescriptive data analytics segue questo schema:



Prescriptive data analytics – esempi

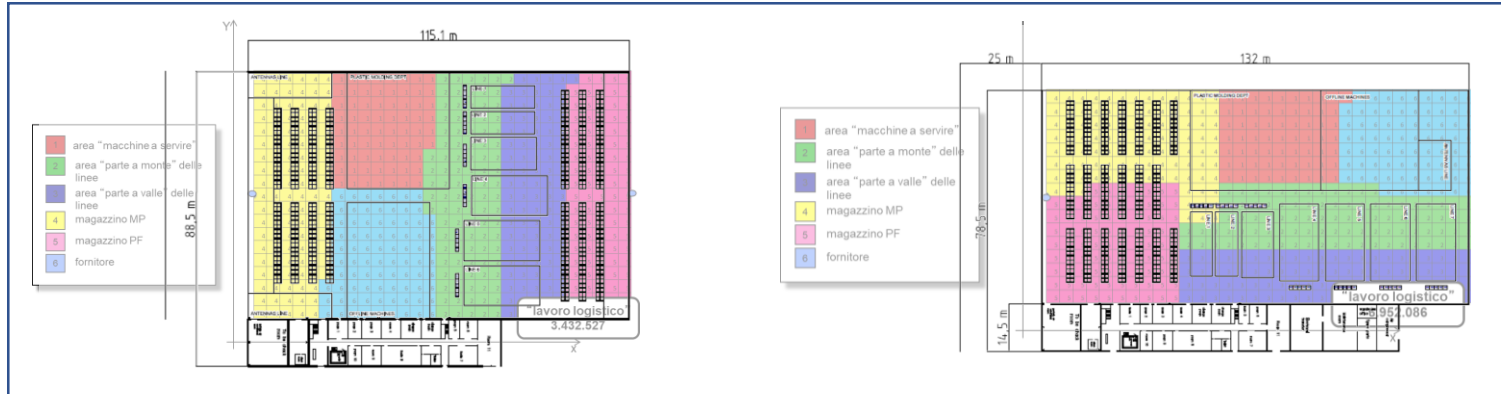
Obiettivo: definire il layout ottimale per la produzione di avvisatori acustici per automobili. I dati di input sono: i) le aree e le loro dimensioni presenti nel plant, ii) i componenti e i flussi questi all'interno delle aree, iii) una grandezza caratteristica come il peso in kg. Grazie a questi dati è possibile definire quale sia il layout con il minor lavoro logistico

$$\min \sum Q_i * d_i$$



Line	Codice	AMBSGX	HRB	LLHX	Totale
Magazzino MP - "parte a monte" delle linee					1452332,966
	A	48.760,37	23.287,30	16.741,75	
	B	504.712,57	365.651,47	173.292,15	
	D	-	-	-	
	G	1.088,75	1.485,63	-	
	J	-	445,69	-	
	N	188.820,05	-	63.618,37	
	L	-	-	-	
	R	7.776,77	3.714,08	-	
	P	-	57.739,98	-	
	F	-	26.969,99	-	
Magazzino MP - "parte a valle" delle linee					18.958,00
	G	195.974,87	222.844,98	91.657,10	
	U	-	-	25.900,37	
Magazzino MP - Fornitore	K	235.091,85	112.276,73	60.718,36	428086,930
Magazzino componenti elettronici - "parte a monte" delle linee	E	-	-	25.252,83	42934,0767
Fornitore - "parte a monte" delle linee	K	235.091,85	112.276,73	60.718,36	428086,930
"Macchine a servizio" gruppi oscillanti - "parte a monte" delle linee	I	356.697,25	239.855,48	108.888,35	705441,0773
		7.118,20	112.053,88	39.491,30	268663,4709
		16.194,03	30.418,34	-	76612,37139
		6.065,88	2.896,98	-	8962,86766
		19.782,63	1.215.210,62	699.096,32	3.674.089,77

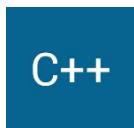
Featuring Microsoft Visual Basic Technology



Prescriptive data analytics – esempi

Obiettivo: definire la sequenza con la quale andare a «sfridare» i fogli di lamiera tagliati al laser in modo da far sì che si riduca il tempo perché una commessa possa avanzare dal taglio alla piega (ovvero, perché tutti i componenti tagliati della commessa siano stati sfridati). Il fine ultimo era duplice: ridurre il tempo di attraversamento delle commesse e ridurre lo spazio utilizzato nell'area sfrido necessario per «parcheggiare» i pallet di commessa

Featuring
 Microsoft®
Visual Basic®
 Technology

...

```

Microsoft Visual Basic, Applications Edition - Algoritmo_20171122.xlsm - [Modulo1 (codice)]
File Modifica Visualizza Inserisci Formato Debug Esegui Strumenti Aggiunte Finestra 2
Riga 6, Col 25

(generale)
Sub Algoritmo_sequenziamento_fogli ()
' Algoritmo_sequenziamento_fogli Macro

Dim Foglio1 As Worksheet
Dim Foglio2 As Worksheet

Dim i As Integer
Dim j As Integer
Dim riga As Integer
Dim colonna As Integer
Dim num_righe As Integer
Dim num_colonne As Integer
Dim contatore As Integer
Dim avanzamento As Integer
Dim benchmark As Integer

Dim array_bom() As Integer
Dim Matrice() As Integer
Dim Matrice_scelta() As Integer
Dim Matrice_pezzi_rimasti() As Integer

Set Foglio1 = Sheets("Foglio1")
Set Foglio2 = Sheets("Foglio2")

i = 2
j = 2
riga = 1
colonna = 1
benchmark = 10000
contatore_verticale_verifica = 0

'conteggio del numero di commesse
While (Foglio1.Cells(i, 1) <> "")
    i = i + 1
Wend

num_righe = i - 1
i = 1

'conteggio del numero di fogli
  
```

Prescriptive data analytics – esempi

Obiettivo: sviluppare una procedura MRP a capacità finita e sua applicazione a un'azienda che produce su commessa attuatori di valvole. Il fine ultimo era quello di consentire all'azienda di compiere una pianificazione a ritroso dei propri reparti produttivi rispettando i vincoli rappresentati dalle loro capacità produttive

$$\min \sum_f \sum_i \sum_t X_{f,i,t} \cdot FC_t,$$

$$\text{CumOrder}_{f,i,t} \geq \text{CumNReqWLot}_{f,i,t} \quad \forall f, \forall i, \forall t$$

$$\sum_i X_{f,i,t} \cdot \text{lot}_{f,i,t} \cdot \text{ct}_{f,i,r} \leq \text{pc}_{f,r,t} \quad \forall r, \forall t$$



...

$$\text{GReq}_{f,i,t} = \left(\sum_{j \neq i} X_{f,j,t} \cdot \text{lot}_{f,j} \cdot u_{f,i,j} \right) + D_{f,i,t} \quad \forall f, \forall i, \forall t,$$

$$A_{f,i,t} = \max \left(0, A_{f,i,t-1} - \text{GReq}_{f,i,t-1} \right) \quad \forall f, \forall i, \forall t,$$

$$\text{NReq}_{f,i,t} = \max \left(0, \text{GReq}_{f,i,t} - A_{f,i,t} \right) \quad \forall f, \forall i, \forall t,$$

$$\text{NReqW}_{f,i,t} = \text{NReq}_{f,i,t} \cdot (1 + w_{f,i}) \quad \forall f, \forall i, \forall t,$$

$$\text{NReqWLot}_{f,i,t} = \frac{\text{NReqW}_{f,i,t}}{\text{lot}_{f,i}} \quad \forall f, \forall i, \forall t,$$

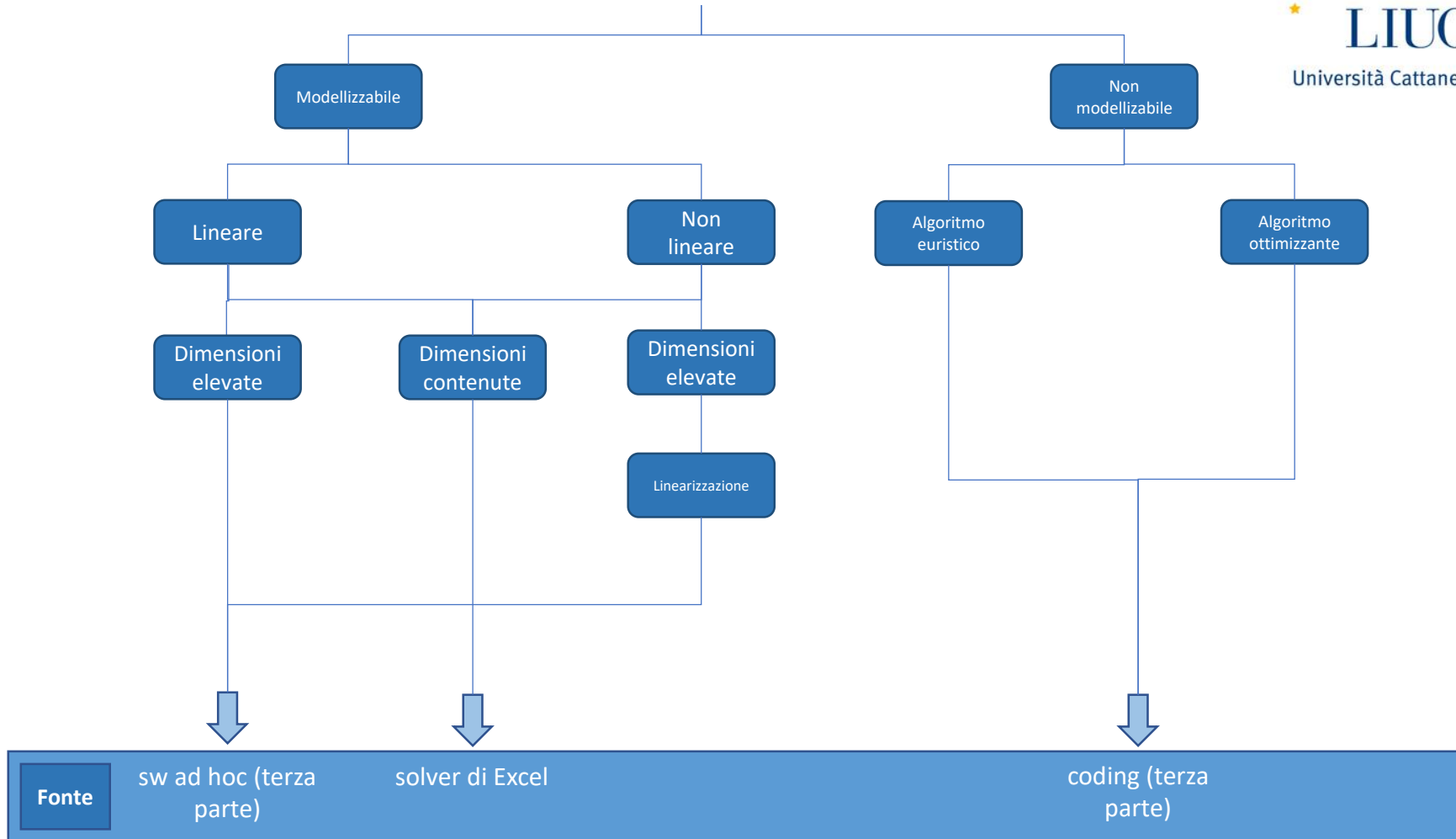
$$\text{CumNReqWLot}_{f,i,t} = \sum_{k=1}^t \text{NReqWLot}_{f,i,k} \quad \forall f, \forall i, \forall t,$$

$$\text{CumOrder}_{f,i,t} = \sum_{k=1}^t X_{f,i,k} \quad \forall f, \forall i, \forall t.$$

Prescriptive data analytics – applicazioni

- Le principali applicazioni in ambito operations riguardano:
 - supply chain design (LP, MILP, NLP);
 - progettazione/riprogettazione di lay-out (MILP);
 - pianificazione aggregata (LP, MILP);
 - gestione dei materiali a capacità finita (MILP, NLP);
 - schedulazione della produzione (algoritmi ottimizzanti, algoritmi euristici);
 - nesting, sorting, ... (algoritmi ottimizzanti, algoritmi euristici);
 - bilanciamento linee di assemblaggio (MILP);
 - ...

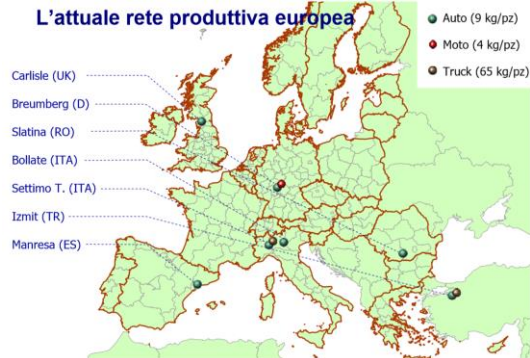
Problema decisionale



Prescriptive data analytics – SCD (esempio)

Obiettivo: identificare il numero e la localizzazione ottimale dei Regional Distribution Warehouse RDW da cui servire tutto il mercato europeo, in modo tale da minimizzare l'overall distribution cost (somma dei costi di trasporto inbound, di distribuzione, dei costi di warehousing e handling) nel rispetto del livello di servizio desiderato

L'attuale rete produttiva europea



La base dati di progetto



- Quantità ricevuta da ciascun RDW dagli stabilimenti del gruppo, in termini quantità (kg) e spedizioni, suddiviso per le 3 business unit



- Quantità consegnata da ciascun RDW ai clienti raggruppati per ZIP Code (2 digit), in termini quantità (kg o pezzi), spedizioni, suddiviso per le 3 business unit (solo per IT, F, D, ES)



- Per RDW sono riportate le distanze stradali verso ciascun punto di consegna (o centroide) e da ciascun Plant o porto di sbarco



- Per ciascun operatore logistico attuale tariffe di trasporto secondario (articolate nelle fasce di peso e classi di distanza), handling & housing e i costi di trasporto primario (a viaggio)



- Per ciascun ZIP Code è riportato il numero dei punti di consegna

La base dati elaborata



- Per ciascun paese, sono riportate le quantità consegnate dal RDW di competenza ai clienti raggruppati per ZIP Code, ripartite per fasce di peso, incrociate con le distanze km e con il numero di clienti per area.
- Sono stati analizzati i "profil" delle spedizioni per classi di peso



- Per ciascuno dei 1322 centroidi in cui viene suddivisa secondo il criterio NUTS3, sono riportate le **quantità consegnate**, le spedizioni ricevute, il costo di distribuzione, il numero di clienti, le coordinate LAT-LON e lo ZIP Code

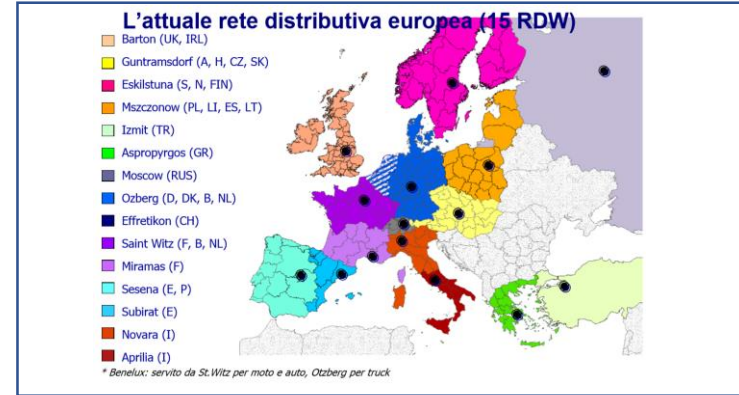
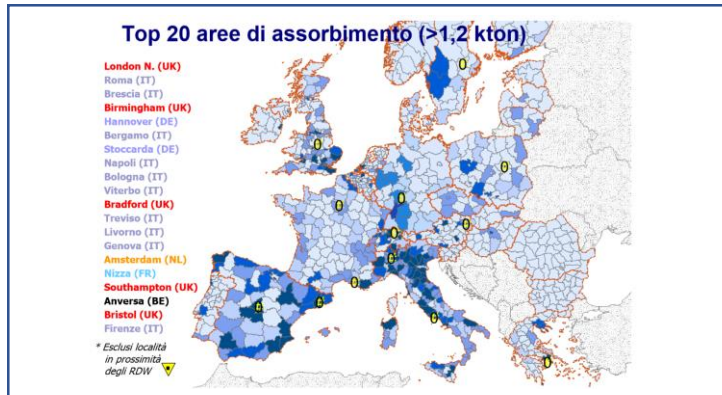
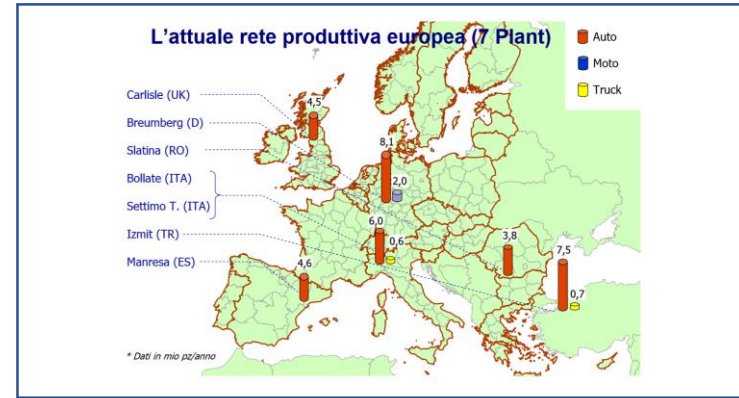
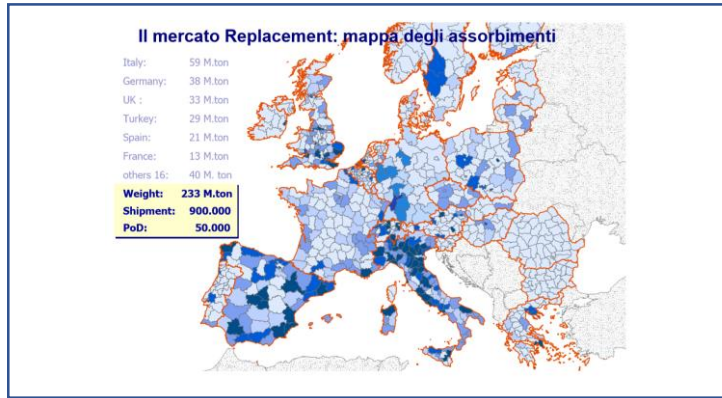


- Per ciascun paese, sulla base delle attuali tariffe di trasporto secondario, è stato **modellizzato il costo unitario** (€/kg_km) che meglio interpretasse l'attuale contratto in essere con i 3PL



- E' un **quadro di sintesi** in cui per ciascun paese, sono state riportate le info salienti, i dati di flusso, i costi unitari e sono stati stimati alcuni KPI di verifica e i costi annui di distribuzione

Prescriptive data analytics – SCD (esempio)



Prescriptive data analytics – SCD (esempio)

$$\min \left(\sum_{h=1}^H \sum_{j=1}^J cs_{hj} \cdot d_{hj} \cdot k_{hj} \cdot D_j + \sum_{h=1}^H \sum_{j=1}^J cw_h \cdot k_{hj} \cdot D_j + \sum_{h=1}^H \sum_{p=1}^P cp_{ph} \cdot \sum_{j=1}^J p_{pj} \cdot k_{hj} \cdot D_j \right)$$

$$\sum_{h=1}^H k_{h,j} = 1 \quad \forall j$$

$$k_{h,j} \leq I_{h,j} \cdot k_h \quad \forall h, j$$

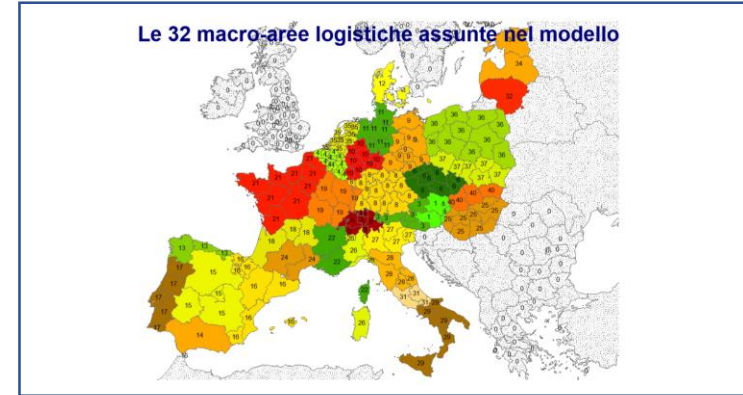
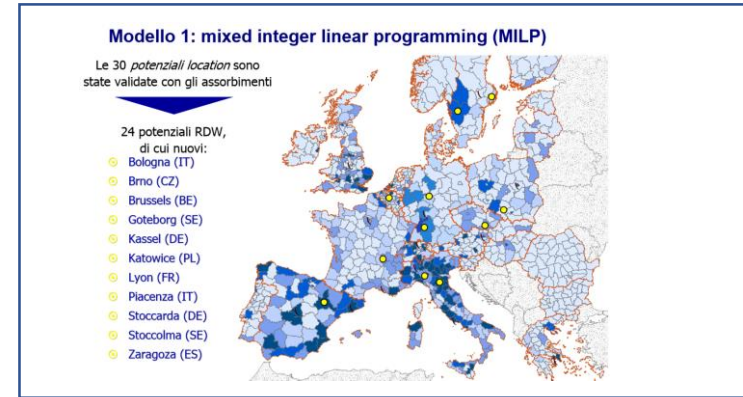
$$k_{h,j} = b_{hj} \quad \forall h, j$$

$$k_h = b_{hj} \quad \forall h$$

$$\sum_{j=1}^J k_{h,j} \cdot \frac{D_j}{S_j - ITR_h} > 4000 \quad \forall h.$$

A	B	C	D	E	F	G
1	SUTS_3	REFURON	X_COORD	Y_COORD	ZP_COORD	ADP_i
2	AT111	MITTELBERGSLAND	11172,7814200000	-279765,4174000000		
3	AT112	NORDRHEINLAND	13163,2612000000	-237514,2210000000	Essen	
4	AT113	SUEBBERGSLAND	9482,8017100000	-11301,8417800000		
5	AT121	MIDTVESTLJANDESVUENEN	6,34621701100	-246234,0624000000		
6	AT122	MIDR-OSTERREICH-REDE	6761,1518070000	-24461,6248170000		
7	AT123	SASKT FVELTEN	4841,6111070000	-20489,7501800000	S. Pölten	
8	AT124	VALDVEITEL	2177,2217170000	-43607,2784000000		
9	AT125	VESTVEITEL	9549,2510800000	-45140,8317180000		
10	AT126	WESTENLANDS-REITEL	10047,2024900000	-481798,2120000000	Stokholm	
11	AT127	WESTENLJANDESVUENEN	10014,1718000000	-246791,4415000000		

ADP _i	NUTS3 cluster	ADP _i demand (kg)	ADP _i geographical coordinates	
			X coordinate	Y coordinate
ADP ₁	AT112	95,009	131,116	-237,154
ADP ₂	AT123	465,655	48,041	-204,049
ADP ₃	AT126	297,091	100,847	-183,598
ADP ₄	AT130	503,437	102,108	-198,224
ADP ₅	AT121	403,415	-75,938	-375,822
ADP ₆	AT122	764,405	29,379	-319,175
ADP ₇	AT132	1,201,131	-63,222	-194,536
ADP ₈	AT125	450,261	-133,736	-248,635
ADP ₉	AT132	503,743	-276,214	-304,616



Prescriptive data analytics – SCD (esempio)

Dato un RDW potenziale (RDWh: ah=1) si calcolano due curve isocrone, luogo dei punti raggiungibili entro 24h e 48h dal RDW. Questi si traducono in vincoli di servizio attraverso variabili “booleane”.



Quadro riassuntivo della copertura per le 32 macro-aree

	Macroaree																																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40				
Güntramsdorf																																												
Effretikon																																												
Ottberg																																												
Sesana																																												
Subrata																																												
St. Witz																																												
Miramare																																												
Novara																																												
Aprilia																																												
Mazzorow																																												
Brussels																																												
Birno																																												
Kassel																																												
Stuttgard																																												
Zaragoza																																												
Lyon																																												
Piacenza																																												
Bologna																																												
Katowice																																												



Software per prescriptive data analytics

CARATTERISTICHE PER LA SELEZIONE

- Effort richiesto per l'implementazione del modello di programmazione

SVILUPPO



- Verificare che il costo di implementazione non sia maggiore dei benefici

COSTO



- Tipologia di dato che viene modellizzato

TIPO DATI



- Tempo richiesto per convergere alla soluzione ottima

VELOCITA'



- Numero di variabili da considerare
- Numero di vincoli da considerare

SCALABILITA'

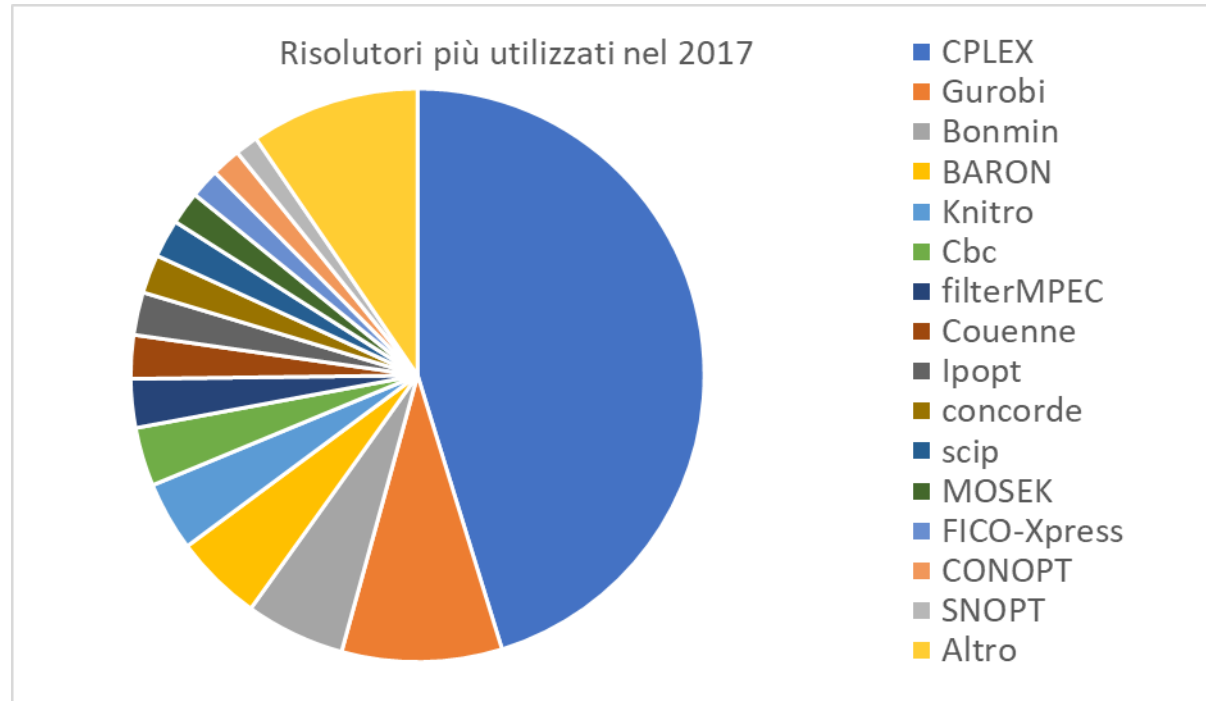


- Tipologia di modello matematico che viene applicato per la risoluzione

ALGORITMO

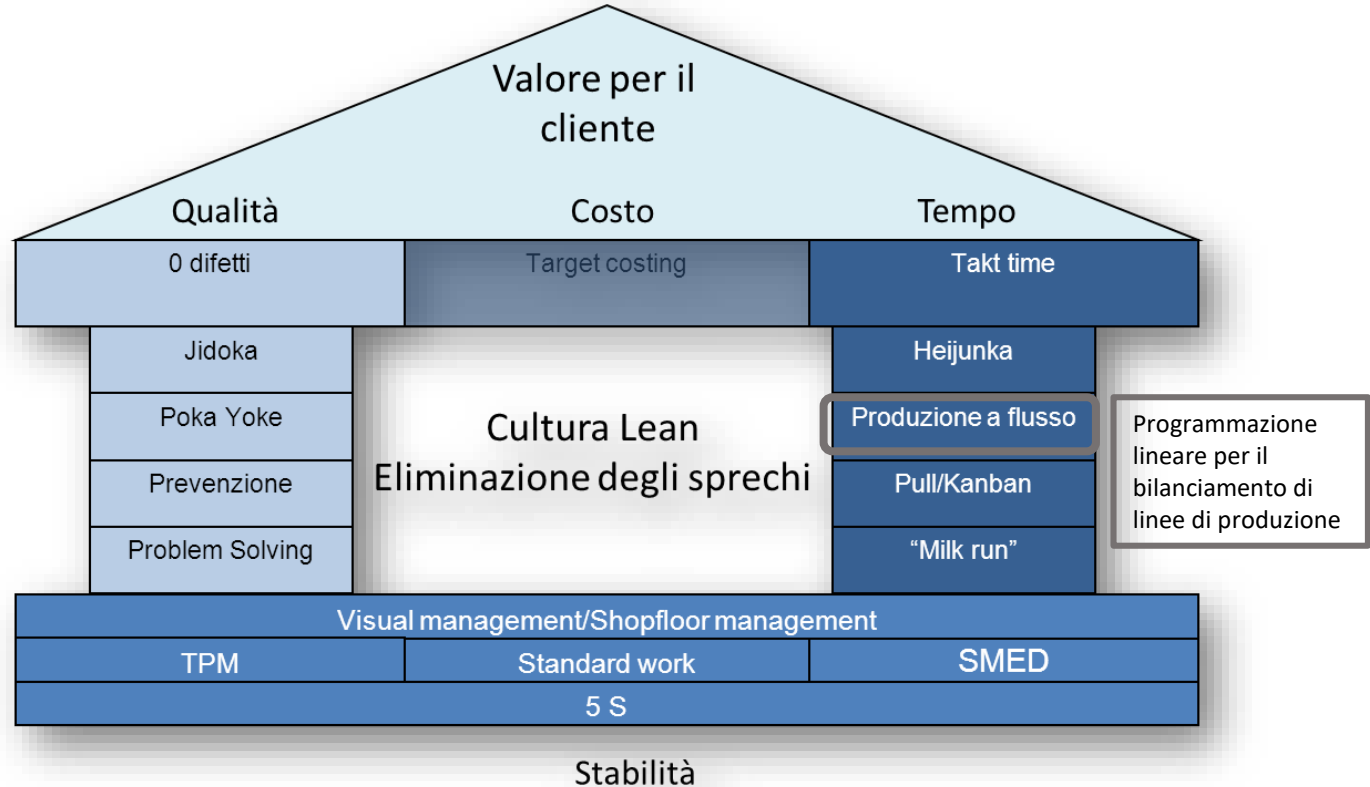


Software per prescriptive data analytics



Problemi di prescriptive analytics risolti attraverso il server NEOS nel 2017

Prescriptive data analytics & lean



Non solo prescriptive data analytics...

- Quando il sistema per il quale si vogliono dare prescrizioni non risponde a leggi fisiche/matematiche note, ovvero quando non è possibile scrivere un insieme di disequazioni/equazioni di vincolo e una funzione obiettivo...
- ...quando le performance del sistema dipendono fortemente da aspetti dinamici (es. concorrenza su una o più risorse)...
- ...quando i valori assumibili da grandezze caratteristiche del sistema non sono certi ma aleatori...
- ...tecniche di prescriptive data analytics o non possono essere impiegate o possono esserlo a prezzo di grandi sforzi sia modellistici, sia computazionali

Simulazione «a eventi discreti»

- Riproduce nel tempo l'evoluzione di un sistema non continuo (come è un sistema logistico-produttivo) consentendo di conoscere in anticipo il comportamento del sistema medesimo
- Generando una storia artificiale del sistema, la simulazione consente di valutare ex ante le prestazioni che il sistema avrà in esercizio e di condurre analisi di sensitività
- Obiettivi progettuali:
 - dimensionamento parco macchine
 - dimensionamento sistema logistico
 - analisi impiantistiche
 - capacity planning
 - analisi colli di bottiglia
 - verifica lay-out
- Obiettivi gestionali:
 - verifica dell'influenza di guasti e problemi di alimentazione materiali
 - ottimizzazione politiche gestione materiali
 - ottimizzazione politiche gestione manodopera
 - analisi prestazioni sistema al variare del mix e dei volumi

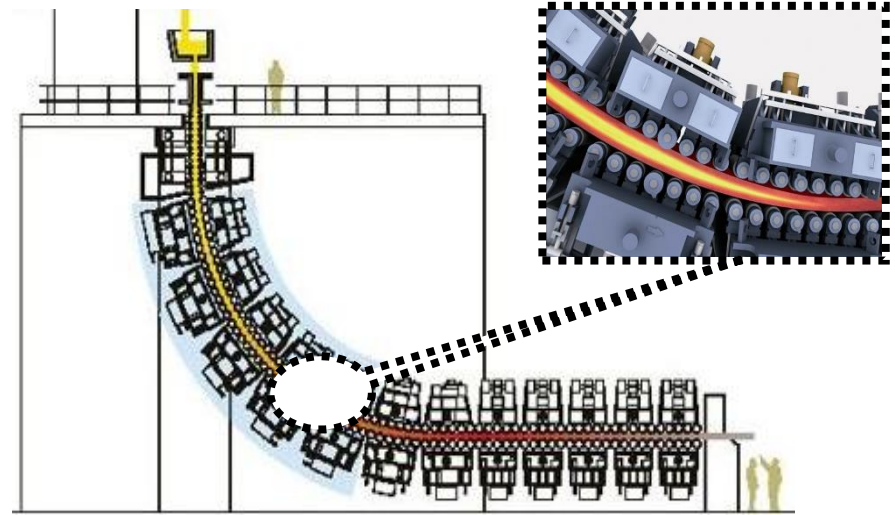
Simulazione «a eventi discreti»

Quali sono le ragioni che spingono a sperimentare su un modello di simulazione anziché sul sistema reale?

- Costo, è meno costoso sperimentare su un modello che sul sistema reale;
- tempo, la sperimentazione su un modello è più veloce rispetto a quella sul sistema reale;
- ripetibilità, su un modello le condizioni al contorno possono essere fissate;
- il sistema reale non esiste ancora.

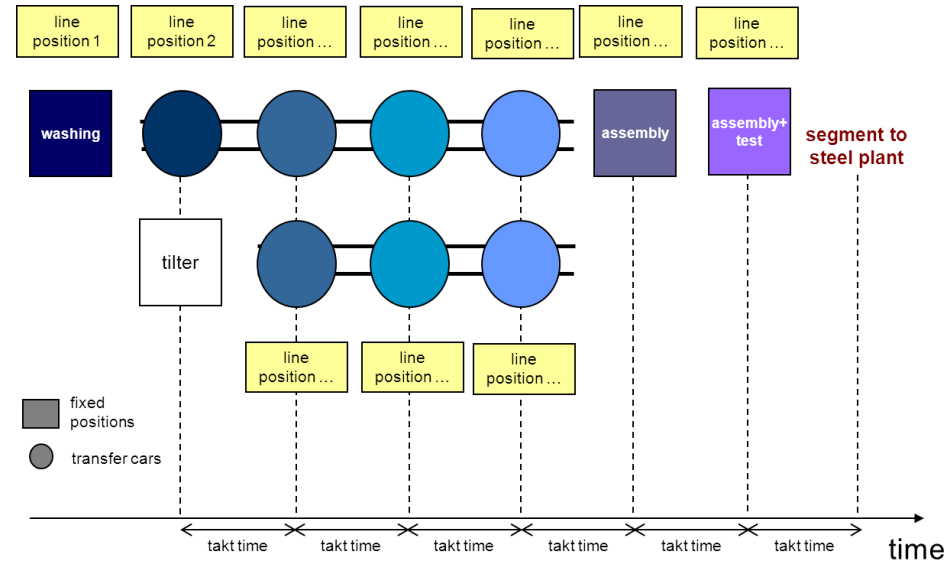
Simulazione «a eventi discreti» – esempio

- Il contesto è quello di un'acciaiera che aveva la necessità di riorganizzare la manutenzione dei segmenti delle sue colate continue
- I segmenti (70 ton l'uno) erano mantenuti a posto fisso con più squadre di manutentori che si alternavano nei diversi turni (il repair shop di un'acciaiera lavora su 3 turni come l'acciaiera stessa)
- I segmenti arrivavano a batch al repair shop e tornavano a batch alle colate dopo aver ricevuto un contenuto di lavoro di circa 600 ore-uomo

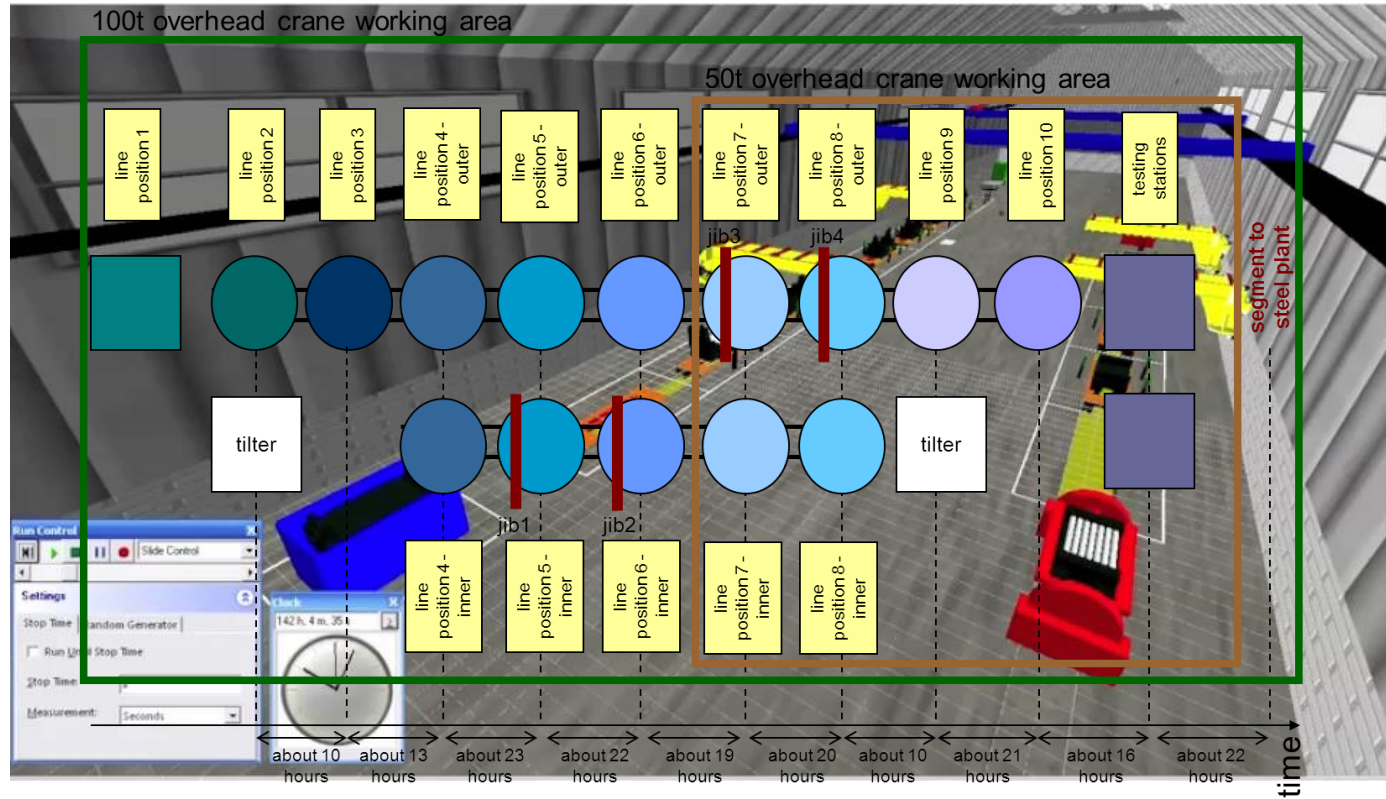


Simulazione «a eventi discreti» – esempio

- Riorganizzare la manutenzione dei segmenti in linea (la linea dovrebbe essere doppia per parallelizzare le attività di manutenzione sull'outer e sull'inner)
- I segmenti verrebbero mossi su carri oleodinamici che scorrono su rotaie e asservite alle postazioni della linea dovrebbero esserci jib crane e tilter
- L'obiettivo dell'acciaieria è che ogni giorno 1 segmento torni dal repair shop alla colata
- **Ma quanti carri, quante jib crane, quanti tilter sono necessari per tale obiettivo?**



Simulazione «a eventi discreti» – esempio



Le tendenze in ambito «simulazione»

Step 1: Formulate problem
State model objective precisely

Step 2: Specify independent and dependent variables
Define independent variables
Define dependent variables

Step 3: Develop and validate conceptual model
Specify assumptions, algorithms, and model components

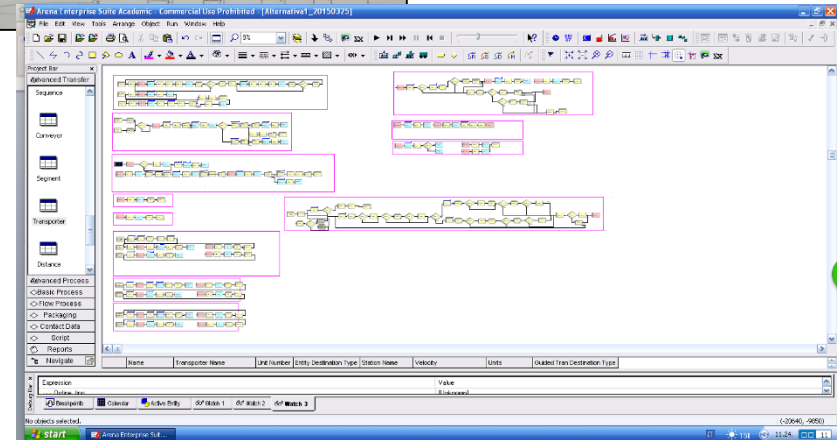
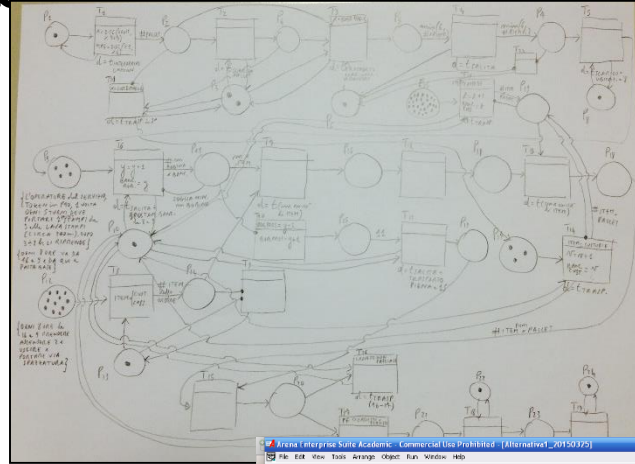
Step 4: Collect data
Define data requirements

Step 5: Develop and verify computer-based model
Develop a detailed flowchart
Choose programming environment

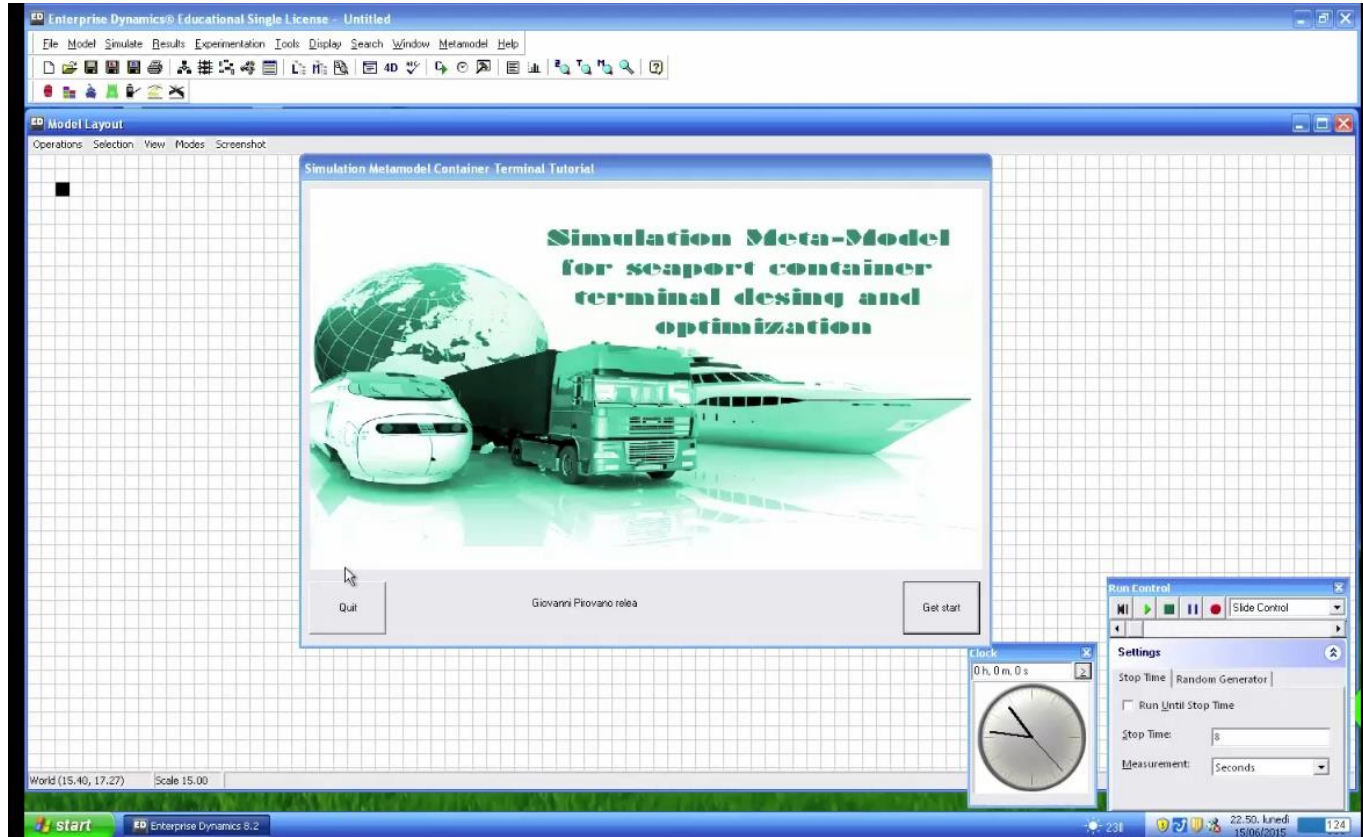
Step 6: Validate the model
Perform a structured walk-through
Check for reasonableness of results
Perform sensitivity analysis

Step 7: Perform simulations
Specify sample size
Specify run length and warm-up period
Perform simulation runs

Step 8: Analyze and document results
Establish appropriate statistical techniques
Document results



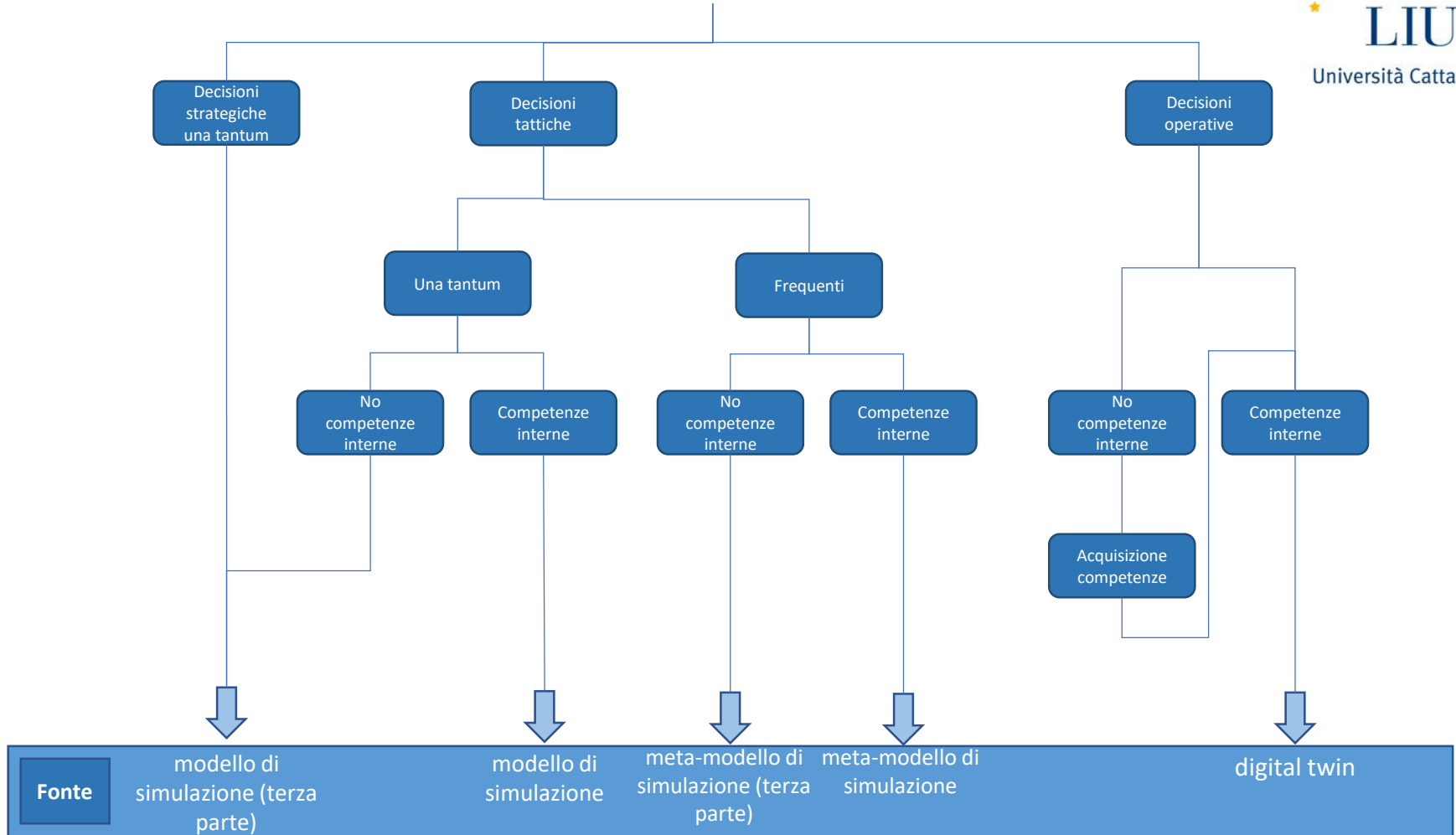
Le tendenze in ambito «simulazione»

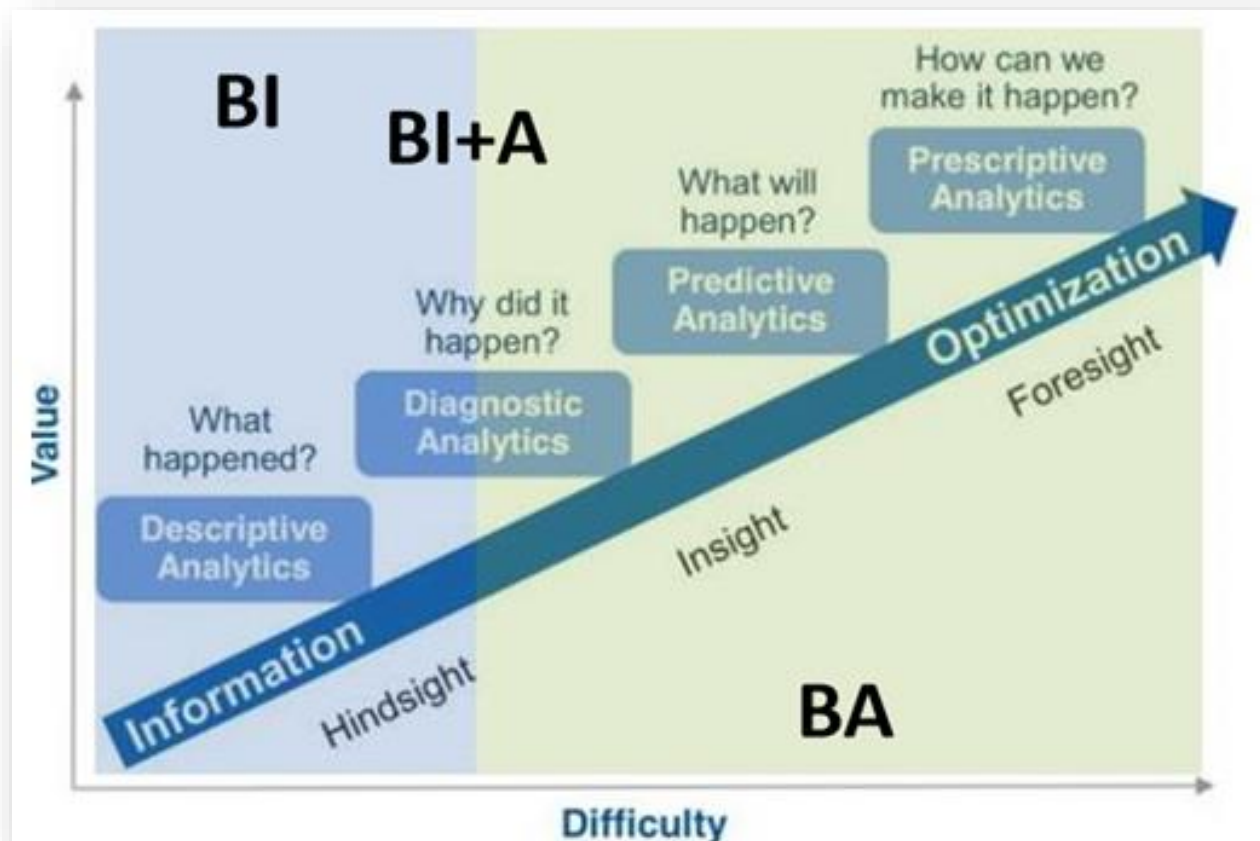



Le tendenze in ambito «simulazione»

- Modelli di simulazione sincronizzati con i sistemi fisici in quanto alimentati con dati provenienti dal campo in tempo quasi reale.
- In questo modo, è possibile trarre vantaggio dalla simulazione anche nell'ambito delle decisioni di tipo operativo (con impatto nel breve-brevissimo periodo).
- Il digital twin del sistema logistico-produttivo alimentato con dati di campo in tempo quasi reale consente al decisore di verificare il comportamento del sistema fisico nell'immediato futuro a fronte di differenti alternative in termini, per esempio, di sequenza con la quale mettere in lavorazione determinati ordini di produzione su una particolare risorsa o di piano di manutenzione programmata all'interno di un certo reparto.
- In questo modo è possibile una nuova modalità di pianificazione della produzione, adatta a contesti affetti da alta variabilità e che richiedono tempi brevi di adattamento.

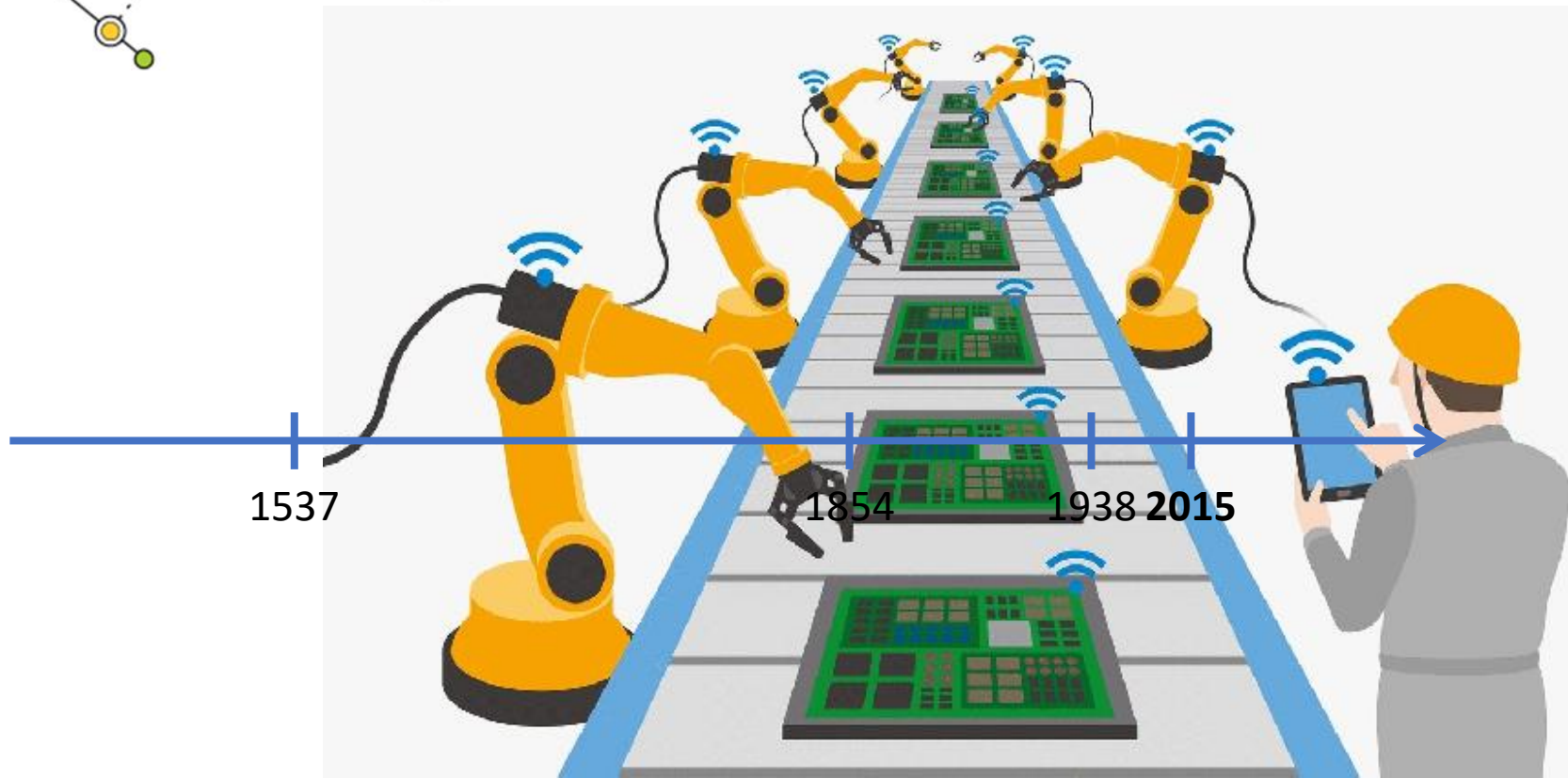
Sistema con aleatorietà e dinamicità

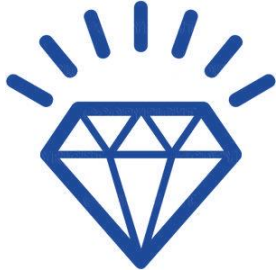






internet of things





Data preparation

- Le tecniche analitiche necessitano che il dato sia «affidabile», cioè non contenga anomalie: valori assenti, di formato errata, valori ridondanti, o che la dimensione dell'insieme del dato sia consona allo strumento utilizzato
- I passi di data preparation sono i seguenti:
 - Data validation
 - Data transformation
 - Data reduction

Data preparation

- Il data validation è il passo che viene eseguito per gestire:
 - Incompletezze, valori assenti (es. per via di manipolazioni precedenti o mancati inserimenti)
 - Rumore, valori errati o anomali (outliers) (es. per via di malfunzionamenti nei dispositivi di rilevamento o trasmissione)
 - Inconsistenze, valori che hanno rappresentazioni o unità di misura diverse (es. per via della provenienza da fonti che utilizzano standard diversi). Si gestiscono tramite trasformazioni del dato (data transformation)

Data preparation

- Le incompletezze possono essere gestite mediante:
 - Eliminazione: si elimina il record che contiene l'elemento vuoto
 - Inspection: si ispezionano gli elementi assenti e si popolano con valori verosimili
 - Identification: si popola l'elemento con un valore convenzionale (es. -1 assente)
 - Substitution: si popola l'elemento con un valore che è «concorde» con altri valori (di record adiacenti per dati su serie temporali o di un gruppo sufficientemente grande in altri casi), per esempio si può usare la media dei valori vicini
- Il rumore può essere gestito identificando record che hanno outliers e utilizzando eliminazione e substitution
- Le inconsistenze possono essere gestite mediante:
 - Standardization: si traduce il valore nell'unità di misura voluta
 - Feature extraction: si elabora il dato secondo la logica che lo trasforma nella forma voluta, quindi può essere una funzione matematica

Data reduction

- Il data reduction è il passo che serve a riportare il campione di dati (es. il numero di record) a una dimensione gestibile dagli strumenti che si utilizzano. Il data reduction considera tre criteri:
 - Efficienza: ogni algoritmo (per approssimazione: programma) impiega un determinato numero di passi per concludere la propria esecuzione. Il tempo di computazione (di calcolo) dipende dal numero di passi (e dal tempo di clock del CPU utilizzato). Il numero effettivo di passi dipende dalla dimensione dei dati in input al programma. Quindi maggiore è la dimensione del dato, maggiore è il tempo di computazione.
 - Accuratezza: diminuire la dimensione dei dati in input può abbassare il tempo di computazione, ma può anche diminuire l'accuratezza del risultato, quindi dell'intera analisi. Perciò si cerca di ridurre la dimensione del dato tanto quanto è possibile, ma senza inficiare la precisione del risultato.
 - Semplicità: spesso i fenomeni che si studiano dipendono da tantissime variabili, difficili da interpretare nel loro complesso. Potrebbe essere quindi utile ignorarne alcune (diminuendo l'accuratezza dell'analisi) a favore di una semplicità di lettura e analisi dei risultati

Data reduction

- Il data reduction può avvenire per:
 - Sampling: si campiona un sottoinsieme di record, casualmente o seguendo specifiche regole (probabilistiche)
 - Feature selection: si eliminano «le colonne» che non sono necessarie, cioè che non contengono quantità che incidono sull'analisi (quindi che non sono correlate con le quantità che si vogliono misurare o stimare)