



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Gestionale

A.a. 2024/2025

Sessione di Laurea ottobre 2025

Analisi della Digital Voice-of- Customer mediante Topic Modeling, detrending e Carte di Controllo per variabili

Relatori:

Prof. Federico Barravecchia

Prof. Luca Mastrogiacono

Candidato:

Fabio Cricelli

INDICE

INTRODUZIONE	6
CAPITOLO 1 – CONCETTI TEORICI	10
1.1 La Customer Satisfaction e la Customer Loyalty	10
1.2 La Voice of Customer (VoC)	16
1.2.1 Gli User-Generated Contents (UGC)	19
1.3 Il Quality Tracking	21
1.3.1 Il Topic Modeling	21
1.3.2 Il modello Latent Dirichlet Allocation (LDA)	25
1.3.3 L’algoritmo Structural Topic Model (STM)	26
1.4 Le Carte di Controllo	30
1.4.1 Carta di controllo $\bar{X}$ -S	33
1.4.2 Applicazione delle Carte di Controllo per l’Analisi della digital VoC	35
CAPITOLO 2 – METODOLOGIA	38
2.1 Preparazione del Database	39
2.2 Pre-Processing	41
2.3 Selezione del numero di Topic	43
2.4 Etichettatura dei Topic	45
2.5 Validazione dei dati	48
CAPITOLO 3 – CASO DI STUDIO	54
3.1 Preparazione del Database e Pre-processing	54
3.2 Selezione del numero di Topic	57
3.3 Applicazione dell’algoritmo Structural Topic Model (STM)	59
3.4 Etichettatura dei Topic	61
3.5 Validazione del modello	64
3.6 Analisi dei risultati	66
3.6.1 MRP	66
3.6.2 MTP	69
3.6.3 IMTP	70
3.7 Applicazione delle Carte di Controllo	73
3.7.1 Detrending e calcolo dei residui	73

3.7.2 Applicazione della carta $\bar{X}$ -S	77
3.7.3 Analisi critica della metodologia	93
CAPITOLO 4 – CONCLUSIONI	99
APPENDICE I	103
APPENDICE II	106
APPENDICE III	118
APPENDICE IV	122
BIBLIOGRAFIA	126

INDICE DELLE FIGURE

Figura 1: Modello di McDougall e Levesque (2000), che descrive come la Customer satisfaction sia raggiungibile tramite un insieme di più fattori.	155
Figura 2: Gerarchia di classificazione dei Topic Model (Kherwa e Bansal, 2018). Essi possono essere suddivisi in due macrocategorie (probabilisti e non probabilistici). A partire da questa distinzione, esistono numerosi modelli che presentano caratteristiche diverse rispetto agli altri.	222
Figura 3: Funzionamento di un algoritmo di Topic Modeling (Barravecchia et al., 2022). A seguito dell'input considerato (ad esempio un database di VoC), l'algoritmo restituisce come output la Topical Content e la Topical Prevalence.	244
Figura 4: Esempio del funzionamento del modello LDA (Blei, 2012)	255
Figura 5: Rappresentazione del metodo STM (Roberts et al., 2016). Il Modello è un insieme di tre sotto modelli (Topical Prevalence, Topical Content e Core Language).....	288
Figura 6: Esempio di una carta di controllo (Montgomery, 2009). In azzurro è raffigurato il processo, mentre le tre linee orizzontali rappresentano la linea centrale (CL), il limite superiore (UCL) e il limite inferiore (LCL) della carta utilizzata.	311
Figura 7: Metodologia sviluppata per il Topic Modeling (Mastrogiacomo et al., 2021). Essa è rappresentata da una serie di sette step consequenziali, partendo dall'estrazione del database e finendo con l'analisi dei risultati. Una volta terminati tutti gli step, è possibile proseguire con l'applicazione delle Carte di controllo.	388
Figura 8: Esempio di una rappresentazione grafica delle metriche "Held-out likelihood", "Residuals", "Semantic Coherence" e "Lower Bound". Per il presente lavoro di tesi verrà sempre considerata la metrica Held-out likelihood.	444
Figura 9: Differenze principali tra le metriche automatiche e le metriche supervisionate per la validazione dei dati (Barravecchia et al., 2021).....	488
Figura 10: Principali step da effettuare per la validazione dei dati (Barravecchia et al., 2021). Dopo l'assegnazione umana e automatica dei topic di un campione di recensioni, viene fatto un confronto tramite il calcolo di alcune metriche.	49
Figura 11: Metriche per la validazione dei dati (Costa et al., 2007; Franceschini et al., 2019; Maria Navin e Pankaja, 2016; Zaki e McColl-Kennedy, 2020; Barravecchia et al., 2021). Tutte le formule delle metriche descritte vengono costruite partendo dai valori veri positivi, falsi positivi, veri negativi e falsi negativi.	522
Figura 12: Valori soglia delle metriche per la validazione dei dati (Barravecchia et al., 2021).	533
Figura 13: Script lanciato per individuare il valore K.	577
Figura 14: Grafici che illustrano i valori per le metriche Held-Out Likelihood, Semantic Coherence, Residuals e Lower Bound. Considerando la prima metrica, viene scelto un valore $k=12$	588
Figura 15: Rappresentazioni grafiche della metrica MRP per i dodici topic analizzati.	688
Figura 16: Grafico IMTP per il Topic T4 Costi del servizio, dove è possibile osservare una tendenza positiva.	722
Figura 17: Grafico IMTP per il Topic T7 Affidabilità ristorante, dove è possibile osservare una tendenza negativa.....	722
Figura 18: Grafico dei residui per il Topic 1 Affidabilità dell'app, dopo aver effettuato il detrending.	755

Figura 19: Confronto tra il grafico iniziale (a) e il residuo (b) dopo il detrending, per il Topic 4 Costi del servizio.	766
Figura 20: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 1 Affidabilità dell'app	800
Figura 21: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 2 Accesso all'account	811
Figura 22: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 3 Disponibilità dei prodotti	822
Figura 23: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 4 Costi del servizio.....	833
Figura 24: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 5 Qualità del servizio	844
Figura 25: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 6 Gestione del rimborso	855
Figura 26: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 7 Affidabilità ristorante	866
Figura 27: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 8 Personalizzazione dell'ordine	877
Figura 28: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 9 Puntualità della consegna	888
Figura 29: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 10 Affidabilità del servizio clienti.....	89
Figura 30: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 11 Consegna all'indirizzo voluto e localizzazione	900
Figura 31: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 12 Completezza dell'ordine.....	911
Figura 32: Confronto tra la Carta $\bar{X}$ applicata all'IMTP iniziale (a) e quella applicata al residuo (b) per il Topic 4 Costi del servizio.	944
Figura 33: Calcolo della Carta $\bar{X}$ per un IMTP senza detrending (a) e uno con detrending (b). Il grafico (b) mostra una diminuzione dei punti di fuori controllo.	955
Figura 34: Calcolo della Carta $\bar{X}$ per un IMTP senza detrending (a) e uno con detrending (b). Il grafico (b) mostra un aumento dei punti di fuori controllo.....	966

INDICE DELLE TABELLE

Tabella 1: Strumenti utilizzati per l'individuazione dei bisogni espliciti del cliente. (Freeman e Radziwill, 2018)	188
Tabella 2: Strumenti utilizzati per l'individuazione dei bisogni taciti del cliente. (Freeman e Radziwill, 2018)	188
Tabella 3: Distribuzione numerica delle recensioni all'interno del database iniziale, dopo le fasi di Pre-processing.	555
Tabella 4: Esempio di distribuzione della Topical Prevalence su tre recensioni e dodici topic.	59
Tabella 5: Risultati ottenuti nella fase di labelling delle determinanti di qualità. La Tabella presenta il nome scelto per l'argomento e una breve descrizione che lo caratterizza.	633
Tabella 6: Calcolo indicatori di validazione. In verde sono evidenziate le metriche che rientrano nei valori soglia; viceversa, in rosso quelle che non ci rientrano.	644
Tabella 7: Calcolo del valore MRP per la company Glovo.	677
Tabella 8: Calcolo del valore MTP per la company Glovo. La somma di tutti i valori, come indicato anche nella formula (3.3), equivale a 1.	700
Tabella 9: Calcolo dei residui per le prime dieci osservazioni del Topic 1 Affidabilità dell'app. Partendo dal valore iniziale y_t , è stato sottratto il valore $\hat{y}_t$ utilizzando le formule (3.5) e (3.6).	744
Tabella 10: Confronto tra l'equazione della retta iniziale e quella della retta ottenuta dopo il detrending.	766
Tabella 11: Valori di $\bar{x}$ (grande media del processo) e $\bar{s}$ (scarto tipo medio) per ogni topic.	788
Tabella 12: Numero di punti fuori controllo per ogni Topic, per la Carta $\bar{X}$	922

INTRODUZIONE

Nel panorama economico odierno, caratterizzato da un'intensa competizione e da clienti sempre più esigenti, la capacità di un'impresa di comprendere e soddisfare i propri consumatori rappresenta un fattore determinante per il successo e la sostenibilità della stessa nel lungo periodo. La qualità percepita di un prodotto o servizio è diventata un elemento strategico fondamentale, non solo per differenziare l'offerta sul mercato, ma soprattutto per costruire relazioni durature con la clientela; pertanto, comprendere i fattori che influenzano questa percezione è essenziale per indirizzare le nuove scelte aziendali e attuare interventi mirati al miglioramento continuo.

Tradizionalmente, le aziende hanno adottato strumenti di monitoraggio della qualità basati principalmente su interazioni dirette col cliente, come questionari, interviste o sondaggi periodici. Sebbene queste metodologie forniscano informazioni preziose, risultano spesso costose, limitate nel tempo e, a volte, invasive per l'utente. L'evoluzione tecnologica e la diffusione delle piattaforme digitali hanno tuttavia trasformato radicalmente il rapporto tra aziende e consumatori, generando notevoli quantità di dati provenienti dai contenuti che gli utenti online condividono spontaneamente. Recensioni, commenti e valutazioni pubblicate su app e altri aggregatori, note nell'insieme come Digital Voice-of-Customer (VoC), costituiscono quindi una risorsa di conoscenza immediata e a basso costo.

Questa nuova prospettiva consente alle imprese di raccogliere feedback costanti e aggiornati in tempo reale, sfruttando i cosiddetti User Generated Contents (UGC). L'analisi automatizzata di questi dati permette di individuare temi e tendenze ricorrenti, offrendo una comprensione più approfondita delle esperienze dei

consumatori e delle criticità percepite. In particolare, il settore del food delivery, trainato dalla crescente diffusione di applicazioni e piattaforme digitali, rappresenta un contesto ideale per sperimentare approcci innovativi di monitoraggio della qualità del servizio: la rapidità dei processi, l'intensità delle interazioni e la quantità di recensioni disponibili rendono necessarie metodologie analitiche robuste e flessibili.

In questo contesto si inserisce la presente tesi, che si propone di esplorare l'applicazione di strumenti di Controllo Statistico del Processo alla Digital VoC. In particolare, la ricerca si concentra sull'utilizzo delle carte di controllo $\bar{X} - S$, strumenti solitamente usati nei processi industriali per monitorare media e variabilità, con l'obiettivo di valutarne l'efficacia nel rilevare variazioni significative nei temi estratti tramite algoritmi di Topic Modeling. Il settore analizzato nel caso di studio è il food delivery, con particolare attenzione ai processi che presentano diversi trend nel tempo.

Gli obiettivi specifici della ricerca sono:

- Indagare se le Carte $\bar{X} - S$ siano efficaci nel rilevare variazioni nella percezione della qualità del servizio, a partire dall'analisi della Digital VoC tramite l'applicazione di algoritmi di Topic Modeling;
- Valutare l'impatto dell'attività di eliminazione dei trend (denominata detrending) sulla capacità della carta di distinguere anomalie reali da semplici effetti dovuti all'andamento temporale;
- Confrontare i risultati ottenuti con e senza detrending, evidenziando vantaggi e limiti dei due approcci;
- Definire raccomandazioni pratiche su quando e come utilizzare i suddetti strumenti in contesti aziendali.

Il gap di ricerca affrontato è evidente: la letteratura offre numerosi studi sull'uso delle carte di controllo in ambito manifatturiero e su applicazioni del Topic Modeling alla Digital VoC, ma pochi studi collegano questi due ambiti, valutando l'utilizzo del controllo statistico del processo a dati qualitativi e dinamici provenienti dalle recensioni online.

La scelta di focalizzarsi sulle carte $\bar{X} - S$ rispetto ad altre tipologie di carte, risponde alla necessità di testare uno strumento ampiamente diffuso e di immediata interpretabilità, per valutare il comportamento anche in situazioni più complesse, come ad esempio la presenza di tendenze. Tale approccio contribuisce ad arricchire la letteratura sul quality tracking nei servizi digitali, aprendo nuove prospettive di ricerca.

Il contributo pratico per le aziende è altrettanto rilevante: i risultati ottenuti mostrano come il monitoraggio della Digital VoC tramite carte di controllo possa supportare i decision maker nell'identificazione tempestiva di segnali di peggioramento o miglioramento della percezione del servizio. In particolare, il detrending permette di evitare falsi allarmi in presenza di tendenze, mentre l'osservazione parallela dei dati originali consente di mantenere il focus anche sull'evoluzione complessiva del processo. Le imprese del settore del food delivery, e più in generale quelle operanti nei servizi digitali, possono quindi adottare questo approccio per rafforzare le proprie strategie di ascolto e miglioramento continuo.

Il primo capitolo del lavoro comprende un'analisi della letteratura, andando a illustrare lo stato dell'arte riguardante argomenti come Customer satisfaction, Digital Voice-of-customer, Topic Modeling, Quality tracking e applicazione delle carte di controllo agli User Generated Contents.

A seguire, il secondo capitolo descrive, sempre a livello teorico, la metodologia adottata per raggiungere l'obiettivo della ricerca: in particolare, il framework proposto da Barravecchia et al. (2021) inizia con la raccolta delle recensioni che andranno a costituire il dataset di partenza. Su tali dati viene applicata un'attività di pre-processing, realizzata attraverso specifici algoritmi implementati nel software R, con lo scopo di migliorare la qualità e renderli adeguati alle analisi successive. L'algoritmo di Topic Modeling, successivamente, avrà il compito di individuare gli argomenti maggiormente discussi all'interno del corpus testuale, lasciando all'analista il compito di assegnare denominazioni esplicative e interpretabili. L'ultimo step è dedicato alla validazione del modello, condotta tramite il calcolo di apposite metriche, indispensabili per verificare l'affidabilità dei risultati e stabilire le possibilità di proseguire con l'applicazione delle carte di controllo.

Il terzo capitolo è dedicato invece all'applicazione pratica al caso di studio. Inizia con la raccolta sul web delle recensioni, tramite un'attività chiamata web scraping, per poi seguire la metodologia descritta precedentemente, fino ad arrivare alla fase di validazione complessiva del modello. A questo punto, vengono calcolati, tramite differenti indicatori, le metriche che andranno a formare gli andamenti temporali su cui verranno applicati gli strumenti di monitoraggio. Viene effettuata una valutazione sui trend che caratterizzano i processi di ogni topic analizzato, per poi applicare le Carte di Controllo $\bar{X} - S$. Il capitolo termina con un'analisi critica dei risultati, fornendo una serie di confronti e commenti.

Il quarto capitolo infine sintetizza le conclusioni, evidenziando contributi, limiti, implicazioni pratiche e prospettive per sviluppi futuri.

CAPITOLO 1 – CONCETTI TEORICI

1.1 La Customer Satisfaction e la Customer Loyalty

In un contesto di crescita economica estremamente competitivo, la sopravvivenza e lo sviluppo delle imprese dipendono da una conoscenza precisa dei consumatori. Diventa quindi di fondamentale importanza per ogni responsabile del marketing e gestione della qualità comprenderli a fondo, in quanto la loro soddisfazione rappresenta il miglior indicatore di come l'organizzazione apparirà in futuro. Una parte essenziale del successo di qualsiasi azienda è la soddisfazione del cliente, che si basa interamente su una solida relazione con esso (Musafir, 2017); in assenza di consumatori, infatti, un'impresa non esisterebbe (Chattopadhyay, 2019).

La soddisfazione, o insoddisfazione, è solitamente vista come una funzione delle prestazioni, in relazione alle aspettative o ad altri criteri di confronto. Dalle numerose definizioni fornite dalla letteratura, emerge un denominatore comune: una valutazione post-adozione riguardante una scelta di acquisto specifica (Westbrook e Oliver, 1991). La decisione di acquisto può essere valutata dopo aver effettuato la scelta, ma comunque prima dell'acquisto reale del prodotto o servizio, mentre la soddisfazione potrebbe anche manifestarsi prima della decisione o anche senza un acquisto o una scelta effettiva (Giese e Cote, 2002). Quando un cliente è soddisfatto, si verifica un impatto positivo sul valore dell'azienda, influenzandone gli utili, in quanto tenderà a spendere maggiormente e a raccomandare l'impresa ad altri consumatori (Musafir, 2017); analogamente, quando la customer satisfaction inizia a diminuire, cala anche il rendimento dell'organizzazione (O'Sullivan e Mccallig, 2012).

Il livello di soddisfazione (o insoddisfazione) del cliente, deriva dall'entità della differenza tra le aspettative e la performance percepita, che può risultare positiva, nulla o negativa (Guido, 2015). L'aspettativa è influenzata dal livello di conoscenza rispetto a un determinato insieme di performance di un prodotto o servizio; in termini generali, può essere descritta come "il risultato anticipato di una situazione probabilistica" (Reber, 1985, p. 257).

Esistono in letteratura una serie di indicatori che hanno come scopo la misurazione della soddisfazione del cliente; uno tra questi è il Customer Satisfaction Index (CSI), il quale considera a sua volta i seguenti cinque aspetti (Ilieska, 2013):

- **Aspettative dei clienti:** rappresentano il livello di anticipazione riguardo la qualità di un prodotto o servizio fornito da un'impresa. Esse possono essere influenzate sia da esperienze dirette di consumo passate, sia da informazioni non esperienziali come il passaparola o la pubblicità.
- **Qualità percepita:** viene valutata in termini di personalizzazione (quanto un servizio soddisfa le esigenze individuali del compratore) e affidabilità (presenza o meno di problemi post adozione).
- **Valore percepito:** è una misura della qualità che tiene conto del prezzo pagato, andando a considerare quindi il prezzo complessivo rispetto alla qualità e la qualità complessiva rispetto al prezzo. Il prezzo percepito solitamente è cruciale nella decisione di acquisto iniziale.
- **Reclami dei clienti:** vengono misurati come la percentuale di rispondenti che hanno riscontrato problemi con un prodotto o servizio entro un certo periodo di tempo dall'acquisto, segnalando un reclamo all'azienda. La soddisfazione e i reclami hanno un rapporto inversamente proporzionale.

- **Fedeltà dei clienti:** considera la probabilità del cliente di effettuare nuovamente un acquisto dallo stesso fornitore in futuro. È una componente critica in quanto rappresenta un importante indicatore di redditività aziendale.

La soddisfazione influisce positivamente sulla fedeltà.

Le aspettative dei consumatori non sempre corrispondono necessariamente a ciò di cui hanno bisogno o a ciò che desiderano; quindi, l'insoddisfazione potrebbe emergere a prescindere dal livello della performance. Nonostante spesso siano allineate ad altri criteri, – ad esempio, un bisogno, un desiderio o una performance ideale – le aspettative sono sviluppate separatamente dal consumatore in base alle circostanze (Wirtz e Mattila, 2001). L'individuo, a volte, potrebbe inoltre ritenersi insoddisfatto di determinati aspetti delle prestazioni di un prodotto o servizio, di cui non era a conoscenza fino a prima di acquistarlo (Guido, 2015).

La prestazione del prodotto o servizio costituisce il punto di partenza dal quale viene misurata la soddisfazione del consumatore, portandolo ad ottenere determinati risultati che, tuttavia, sono soggettivi e derivanti da una serie di processi cognitivi. (Oliver, 1997). La velocità di navigazione in una determinata app o sito internet può rappresentare un esempio di come essa possa essere misurata oggettivamente, tramite determinate metriche, ma anche percepita in base al concetto di velocità che possiede ogni individuo a livello soggettivo. Secondo Anderson (1973), esiste un limite superato il quale il consumatore non tollera più la distanza tra la performance percepita e l'aspettativa riposta. Quando tale limite viene oltrepassato, il prodotto o servizio ottiene una valutazione negativa.

Identificare i requisiti specifici del cliente per ottenere la sua soddisfazione è fondamentale per le imprese; la raccolta delle stesse informazioni in momenti diversi

può infatti essere di grossa importanza per identificare tendenze e schemi che si sviluppino con l'evoluzione e i cambiamenti dell'organizzazione, perché ciò che funziona in un determinato periodo potrebbe non essere efficace in un altro (Chattopadhyay, 2019). Tramite la raccolta e il confronto temporale delle informazioni, è possibile adattare e modificare i servizi e i prodotti in modo da soddisfare i requisiti e migliorare le pratiche aziendali (Chattopadhyay, 2019).

Secondo Nobar e Rostamzadeh (2018), un forte senso di lealtà del consumatore contribuisce alla creazione di un legame di appartenenza e una vicinanza psicologica con un marchio, identificandolo come familiare. Le ricerche dimostrano pertanto che la customer loyalty ha un significativo effetto positivo sulla forza del marchio; infatti, i comportamenti e gli atteggiamenti di lealtà incidono sul valore del brand (Šerić e Gil-Saura, 2019).

La relazione tra la customer satisfaction e la customer loyalty risulta essere il collegamento più debole nel modello della catena di servizio-profitto (Heskett et al., 1997). La soddisfazione dei clienti, quindi, non garantisce necessariamente che essi rimarranno fedeli a un'azienda (Jones e Sasser, 1995). Secondo Oliver (1999) la soddisfazione non si trasforma direttamente in fedeltà, ma ne rappresenta un precursore da coltivare. Esistono infatti numerosi clienti che, anche se soddisfatti, cambiano comunque fornitore (Reichheld, 1995).

Il modello di McDougall e Levesque (2000), illustrato in Figura 1, descrive la soddisfazione del cliente nel contesto dei servizi, sostenendo che la qualità percepita e il valore del servizio influenzino la soddisfazione del consumatore, la quale a sua volta incide sulle intenzioni future. La qualità percepita del servizio si articola in due dimensioni: il servizio di base (Core) e la modalità di erogazione

(Relazionale). In aggiunta a queste dimensioni, esistono aspetti tangibili, di empatia, reattività, affidabilità e garanzia che i clienti tengono in considerazione per giudicare complessivamente un servizio (Parasuraman et al., 1985). Le determinanti sono in tutto dieci, e possono essere riassunte come segue:

- **Accesso:** facilità di contatto e avvicinamento all'impresa, ottenibile ad esempio avendo degli orari di apertura più comodi, essendo raggiungibili telefonicamente e avendo tempi di attesa brevi.
- **Affidabilità:** costanza nelle prestazioni e nel fornire il servizio correttamente, mantenendo le promesse nei riguardi del cliente. Alcuni esempi sono la precisione nella fatturazione e nelle registrazioni.
- **Competenza:** possesso delle conoscenze sul servizio e sulla sua vendita; questo obiettivo è raggiungibile preparando in modo adeguato tutto il personale.
- **Comunicazione:** conoscenza del cliente e dei metodi di comunicazione appropriati con esso, mantenendo ad esempio la chiarezza sulla descrizione del servizio e del suo prezzo.
- **Conoscenza del cliente:** intercettazione delle esigenze del cliente in qualsiasi momento della fase di acquisto o adozione del servizio.
- **Cortesìa:** gentilezza e rispetto da parte del personale di contatto.
- **Credibilità:** interesse sincero nei confronti delle esigenze del consumatore, trasmettendo onestà e fiducia.
- **Elementi tangibili:** supporto materiale al servizio. Sono incluse tutte le strumentazioni e le attrezzature utilizzate per permettere una corretta fornitura.

- **Sicurezza:** libertà dal rischio, dal pericolo e da ogni eventuale dubbio che potrebbe avere il consumatore, in merito ad esempio ai dati sensibili che fornisce.
- **Reattività:** tempestività del servizio e prontezza del personale in tutte le fasi della vendita.

La fidelizzazione del cliente, che si traduce in transazioni ripetute, risulta essere l'obiettivo principale degli erogatori di servizi. Identificando l'influenza di tali fattori, i fornitori possono concentrarsi maggiormente sulla soddisfazione del cliente, portando come risultato un incremento delle vendite.



Figura 1: Modello di McDougall e Levesque (2000), che descrive come la Customer satisfaction sia raggiungibile tramite un insieme di più fattori.

1.2 La Voice of Customer (VoC)

Con l'evolversi dell'ambiente competitivo del mercato, la concezione di creazione del valore si è trasformata dall'essere esclusivamente prodotta dalle aziende all'essere co-creata insieme da imprese e clienti (Vargo e Lusch, 2017). I suggerimenti e le opinioni del consumatore rispecchiano la volontà di trasformazione del mercato e rappresentano, pertanto, una fonte per aiutare le organizzazioni a creare valore e a intercettare i bisogni dei clienti, che sono sempre in evoluzione, con lo scopo di migliorare il loro benessere (Groth, 2005, Yang et al., 2021).

La "Voice of Customer" (VoC) è un termine impiegato nel mondo del business per descrivere il processo di raccolta requisiti dei consumatori; si tratta di una tecnica che genera un alto quantitativo di desideri e necessità dei clienti, organizzati gerarchicamente e prioritizzati tenendo in considerazione il grado di soddisfazione e le alternative attuali (Griffin e Hauser, 1991).

Inizialmente, la voce è stata presentata come una reazione del consumatore alla sua insoddisfazione nei confronti del prodotto o servizio specifico, cercando di ottenere dall'impresa un rimedio alla situazione (Redress Seeking) (Singh, 1988). Tuttavia, secondo Hirschman (1970), questa definizione si discosta in parte dalla vera nozione di "voce", in quanto essa non consiste solo nel contestare, ma anche nell'esprimere opinioni che apportano valori positivi all'azienda.

Per identificare e organizzare i bisogni, è necessario innanzitutto capire qual è la differenza tra essi. È possibile distinguere tre tipi di bisogni (Freeman e Radziwill, 2018):

- **Bisogni espliciti:** sono quelli direttamente espressi dal cliente.

- **Bisogni impliciti:** non possono essere espressi, ma possono essere determinati tramite dei mezzi indiretti.
- **Bisogni “silenziosi”:** possono essere identificati solamente esaminando il contesto, le relazioni e i bisogni espressi dalle tendenze sociali e societarie.

La grande maggioranza dei programmi che le organizzazioni adottano per ascoltare la voce del cliente, incorpora più input provenienti da strumenti differenti (Freeman e Radziwill, 2018). I metodi utilizzati per individuare i bisogni impliciti provengono principalmente da diversi campi di studi, come ad esempio la filosofia, l’analisi dei dati e la psicologia; alcuni di questi prevedono l’osservazione diretta del cliente oppure la scelta di una preferenza tra due opzioni (Test A/B). Le Tabelle 1 e 2 illustrano invece alcuni dei metodi più utilizzati per raccogliere i bisogni dichiarati espressamente dai consumatori e i bisogni silenziosi (Freeman e Radziwill, 2018):

TECNICA	DESCRIZIONE
Osservazioni dei clienti / “Lens Model”	Osservare direttamente i clienti mentre utilizzano il prodotto o servizio e parlano ad alta voce con un addetto dell’azienda.
Sondaggi	I sondaggi consistono in una serie di domande predefinite, per raccogliere facilmente dati quantificabili per l’impresa. Esistono vari metodi per condurre un sondaggio: di persona, al telefono, tramite sito web.
Benchmarking	È molto utile per le organizzazioni per comprendere come le altre imprese riescono a soddisfare i bisogni dei loro clienti, individuando particolari aree dove poter apportare dei miglioramenti.
Gemba Visits	Molto spesso, i clienti insoddisfatti non esprimono una lamentela, pertanto l’osservazione diretta sul posto di lavoro potrebbe rivelarsi la soluzione migliore per intercettare i motivi di insoddisfazione.

Focus group	I focus group consentono di trascorrere del tempo con un ristretto panel di clienti per intercettare informazioni specifiche sui loro bisogni. Solitamente, vengono svolti di persona.
Brainstorming	Si tratta di uno strumento molto utile per raccogliere idee e suggerimenti in modo meno strutturato dei sondaggi.
Analisi dei Social media, delle Chat e del sito Web	Tramite queste analisi, è possibile acquisire delle informazioni dai feedback espressi sui social media, dalle conversazioni con i rappresentanti del servizio e dal traffico web verso sezioni specifiche del sito aziendale.
Moduli di feedback	Vengono inviati subito dopo l'interazione del cliente con l'impresa che fornisce il prodotto o servizio.
Metodo Delphi	Questo metodo consiste nel presentare più cicli di questionari a degli esperti del settore, fino ad arrivare ad una risposta comune.

Tabella 1: Strumenti utilizzati per l'individuazione dei bisogni espliciti del cliente. (Freeman e Radziwill, 2018)

TECNICA	DESCRIZIONE
Tabella della Voce del Cliente (VoCT)	Esamina i bisogni vaghi e confusi dei consumatori per trasformarli in bisogni quantificabili e raggiungibili.
Topic Modeling	Questo strumento è utile per raccogliere informazioni dai testi non strutturati (ad esempio recensioni online, post sui social media), identificando in essi dei temi o priorità.
Analisi del sentimento	I contenuti dei testi non strutturati vengono confrontati con lessici stabiliti per determinare se il sentimento complessivo è positivo o negativo.

Tabella 2: Strumenti utilizzati per l'individuazione dei bisogni taciti del cliente. (Freeman e Radziwill, 2018)

1.2.1 Gli User-Generated Contents (UGC)

Il concetto di “User-Generated Contents” (UGC) ha guadagnato rilevanza all’inizio dagli anni 2000, con la diffusione dei personal computer e con l’introduzione di interfacce user-friendly (motori di ricerca, condivisione di video e social network), dove gli utenti hanno acquisito improvvisamente degli strumenti per produrre dei contenuti (Bolin, 2018). Gli UGC possono essere definiti quindi come dei “contenuti media (digitali) creati da individui al di fuori delle istituzioni professionali, spesso senza compenso e resi disponibili al pubblico” (Daubs, 2019, p. 1825).

Secondo Vickery e Wunsch-Vincent (2010), i fattori che spiegano la rapida crescita degli UGC sono quattro:

- **Tecnologici**, come ad esempio l’accesso alla banda larga e il miglioramento degli hardware e dei software.
- **Sociali**, come l’accettazione della cultura della condivisione dei contenuti e l’ascesa dei nativi digitali della nuova generazione.
- **Economici**, rappresentati dalle barriere di costo più basse per accedere ai servizi e ai dispositivi.
- **Istituzionali**, ovvero l’adozione di licenze per la creazione libera di contenuti.

Ogni tipo di contenuto può ricadere sotto la tipologia di UGC, purché soddisfatti tre criteri. Innanzitutto, il contenuto deve essere stato pubblicato, altrimenti non può essere classificato come UGC. In secondo luogo, il contenuto deve essere originale, ossia deve essere generato da un autore con una componente di creatività. Infine, deve essere prodotto volontariamente dall’utente, senza alcuna ricompensa, non coinvolgendo alcun tipo di pagamento o profitto (Naeem e Okafor, 2019).

Uno dei temi più importanti per poter analizzare i contenuti generati dagli utenti è la scelta delle piattaforme che veicolano l'UGC, come blog, social media, aggregatori di recensioni e siti di e-commerce (Santos, 2021). La forma più diffusa di UGC (o "digital VoC") è rappresentata dalle recensioni rilasciate dai clienti; tali contenuti possono essere considerati come dei feedback su prodotti o servizi ed essere suddivisi in UGC di tipo "strutturato", come le valutazioni numeriche, e/o di tipo "non strutturato", come i commenti testuali (Zhang et al., 2016). L'analisi di questi dati pubblicati sul web può essere uno strumento in più per le organizzazioni per raccogliere, a basso costo, la VoC e intercettare le esigenze e le aspettative dei consumatori. È inoltre interesse delle imprese avere delle valutazioni positive in merito al prodotto o al servizio offerto, in modo tale che esse possano portare in futuro ad acquisire nuovi clienti (Tirunillai e Tellis, 2012). Alcuni studi hanno evidenziato, infatti, che gli UGC hanno una certa influenza nella decisione di acquisto dei clienti, influenzata in particolare da fattori come la quantità e la qualità delle recensioni, la persuasività percepita e la credibilità della fonte e della piattaforma. (Park et al., 2007, Lee et al., 2008, Zhang et al., 2014).

1.3 Il Quality Tracking

Le principali tecniche adottate dalle imprese per monitorare la qualità nel tempo si basano sulla raccolta della VoC, tramite focus group, interviste e questionari. Questi metodi, tuttavia, presentano alcuni limiti come l'eccessivo tempo richiesto e il ristretto bacino di utenti raggiungibile (Barravecchia et al., 2023). Gli approcci pratici attualmente considerati sono i seguenti tre (Bandaru et al., 2015):

- **Interviste strutturate e questionari:** tramite questo approccio è possibile conoscere la percezione del consumatore in un determinato istante.
- **Valutazioni post-vendita:** permettono di ottenere la valutazione del cliente dopo l'adozione di un certo prodotto o servizio.
- **Monitoraggio continuo:** questa tecnica è utile per raccogliere le informazioni di una certa determinante di qualità di un prodotto o servizio nel tempo in modo continuo.

Con la digitalizzazione delle imprese, sono nate nuove opportunità per gestire la qualità di un servizio o di un prodotto, sfruttando i dati generati direttamente dai consumatori stessi; i testi presenti sul web sottoforma di digital VoC possono infatti essere raccolti e interpretati da una serie di strumenti innovativi di data mining (Sony et al., 2020, Mastrogiacomo et al. 2021).

1.3.1 Il Topic Modeling

I dati non strutturati, presenti ad esempio nei testi del set di recensioni analizzato in questa tesi e facenti parte della digital VoC, per essere interpretati hanno spesso bisogno di strumenti di text mining, tra i quali è possibile annoverare i Topic Models;

grazie a questi algoritmi, è possibile estrarre delle determinanti latenti della qualità di un prodotto o un servizio, le quali possono influenzare negativamente o positivamente la percezione del consumatore (Barravecchia et al., 2020, Mastrogiacomo et al., 2021).

Il Topic Modeling è una tecnica che include una serie di algoritmi di machine-learning progettati per rivelare, scoprire e annotare la struttura dei temi presenti in una raccolta di documenti (Blei, 2012). Esistono numerose tecniche di topic modeling che hanno lo scopo di analizzare le statistiche e ricavare il significato delle parole di un determinato testo. La Figura 2 illustra come tali tecniche possano essere suddivise in due grandi macrocategorie, ovvero quelle basate su un modello probabilistico e quelle basate su un modello non probabilistico (algebrico) (Kherwa e Bansal, 2018):

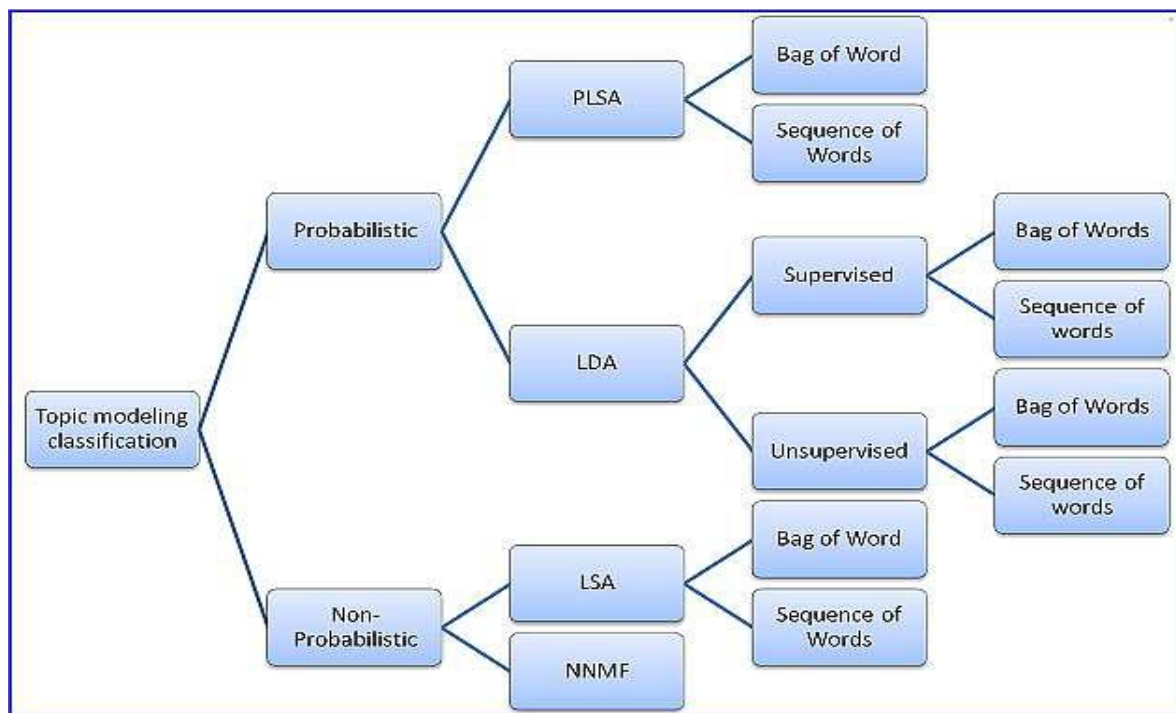


Figura 2: Gerarchia di classificazione dei Topic Model (Kherwa e Bansal, 2018). Essi possono essere suddivisi in due macrocategorie (probabilisti e non probabilistici). A partire da questa distinzione, esistono numerosi modelli che presentano caratteristiche diverse rispetto agli altri.

L'approccio non probabilistico si basa su tecniche algebriche di fattorizzazione delle matrici e nasce nei primi anni '90 con il concetto di Latent Semantic Analysis (LSA) e Non-Negative Matrix Factorization (NNMF). Entrambi lavorano su un approccio "Bag of Words" (BoW), dove ciò che è rilevante è il solo conteggio dei termini in un documento, trascurando completamente l'ordine degli stessi.

Il modello algebrico è stato in seguito superato da approcci di tipo probabilistico, i quali aggiungono appunto il concetto di probabilità; tra i principali modelli, è possibile nominare il Probabilistic Latent Semantic Analysis (PLSA) e il Latent Dirichlet Allocation (LDA); quest'ultima tecnica è stata studiata con approcci di tipo non supervisionato e supervisionato, in merito all'apprendimento del modello. L'ultimo livello di classificazione riguarda la sequenza delle parole analizzate durante il processo di ricerca degli argomenti: oltre al già citato "Bag of Words", il quale è ancora oggi l'approccio dominante, esiste la modalità "Sequence of Words", che oltre a conteggiare i termini in un documento, ne analizza anche il relativo ordine.

Utilizzando quindi le suddette tecniche, è possibile individuare un insieme di argomenti che caratterizzano un determinato corpus testuale, associando, per ogni argomento, delle parole chiave e restituendo una distribuzione degli stessi all'interno del singolo record che viene dato come input (Blei et al., 2003). La Figura 3 offre una rappresentazione grafica del funzionamento generale di un algoritmo di topic modeling.

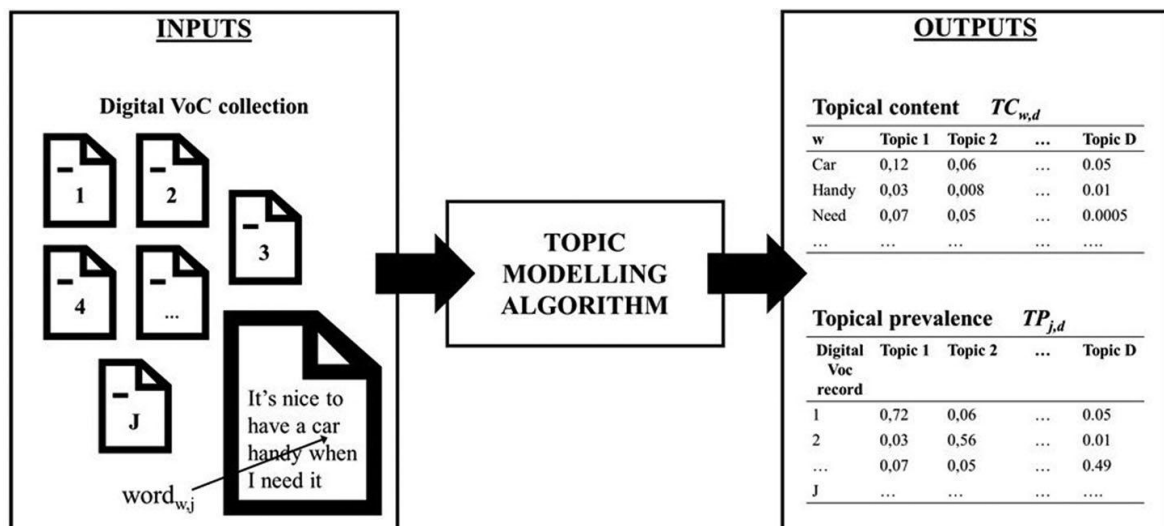


Figura 3: Funzionamento di un algoritmo di Topic Modeling (Barravecchia et al., 2022). A seguito dell'input considerato (ad esempio un database di VoC), l'algoritmo restituisce come output la Topical Content e la Topical Prevalence.

È possibile identificare i seguenti parametri:

- **w**: parola chiave del vocabolario relativo alla digital VoC.
- **d**: topic identificato dall'algoritmo.
- **j**: numero del record analizzato.

Gli output restituiti dall'algoritmo sono la *Topical Content* ($TC_{w,d}$) e la *Topical Prevalence* ($TP_{j,d}$); la prima rappresenta il peso associato a ciascuna w-esima parola chiave all'interno del d-esimo argomento analizzato, mentre la seconda è una distribuzione di probabilità multinomiale che mostra la proporzione del d-esimo topic nel j-esimo record dato come input (Barravecchia et al., 2022).

1.3.2 Il modello Latent Dirichlet Allocation (LDA)

Il Latent Dirichlet Allocation è considerato il più semplice tra gli algoritmi di Topic Modeling e si basa sul fatto che un documento può presentare al suo interno una molteplicità di argomenti differenti. La figura 4 rappresenta il funzionamento semplificato del modello LDA nell'esperimento condotto da Blei (2012).

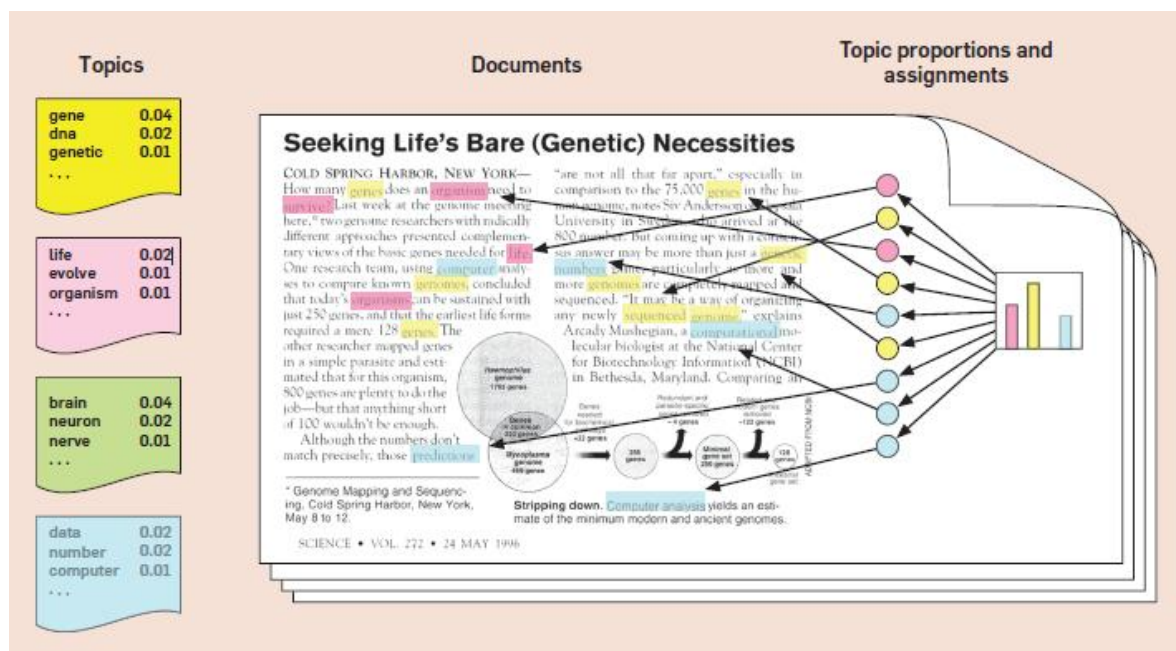


Figura 4: Esempio del funzionamento del modello LDA (Blei, 2012)

All'estrema sinistra della figura sono illustrati i topic: in questo esperimento sono stati assunti quattro argomenti, distinti per colore giallo, rosa, verde e azzurro. I principali limiti del modello si basano sul fatto che esso non fornisce il numero esatto di argomenti, ma è un parametro che deve essere impostato in fase iniziale; in aggiunta, per ogni argomento, l'algoritmo non restituisce un titolo rappresentativo, che deve essere anch'esso impostato dall'uomo. Sulla destra, invece, è presente

un istogramma che indica la probabilità associata per ciascuno dei quattro topic presenti nel documento.

L'algoritmo scelto per le analisi condotte in questa tesi è lo "Structural Topic Model" (STM) (Roberts et al., 2014), che si basa sul modello LDA e verrà approfondito in sezione 1.3.3.

1.3.3 L'algoritmo Structural Topic Model (STM)

Un Topic Model come, ad esempio, quello basato sull'algoritmo LDA, non tiene conto della dimensione temporale, considerando l'intero corpus dei dati come se fossero stati generati tutti nello stesso momento. Poiché è nato l'interesse nel comprendere come determinati argomenti nascono e si evolvono nel tempo, sono emersi altri algoritmi come appunto lo Structural Topic Model. Sia LDA che STM si basano sul presupposto che ogni argomento sia definito da un vocabolario fisso all'interno del corpus stesso, rendendo gli argomenti uguali per ciascun documento, indipendentemente dai metadati che lo rappresentano (luogo, rating, data) (Villamor et al., 2023).

Il modello può essere considerato una combinazione e un'estensione di altri tre algoritmi, che sono il Correlated Topic Model (CTM), lo Sparse Additive Generative (SAGE) e il Dirichlet-Multinomial Regression (DMR). La tecnica STM consente l'introduzione delle covariate all'interno delle distribuzioni che legano il documento all'argomento e l'argomento alla singola parola; il risultato è un documento che è composto da più argomenti analizzabili e correlati tra loro. (Roberts et al., 2013; Roberts et al., 2014).

Per poter spiegare nel dettaglio l'algoritmo, è necessario introdurre prima le seguenti notazioni (Roberts et al., 2016):

- **d** : numero del documento indicato, appartenente all'insieme $\{1, \dots, D\}$.
- **n** : numero di parole all'interno del documento d , con $n \in \{1, \dots, N_d\}$.
- **v** : numero di vocaboli unici all'interno del vocabolario formato dalle parole $w_{d,n}$ all'interno del documento d , con $v \in \{1, \dots, V\}$.
- **k** : numero del topic di riferimento, appartenente a $\{1, \dots, K\}$, dove K è il totale dei topic che viene definito a priori.
- **X** : matrice della prevalenza tematica, la cui dimensione è $D \times P$. Le righe della matrice sono indicate con x_d e contengono i documenti (D), mentre le colonne presentano le covariate che influenzano la topical prevalence (P).
- **Y** : matrice del contenuto tematico, con dimensione $D \times A$. Le righe sono indicate con y_d e contengono anche qui i documenti (D), mentre nelle colonne sono presenti le covariate che influenzano la topical content (A).
- **m_v** : distribuzione multinomiale del termine v nel vocabolario.

Lo structural topic model può essere a questo punto analizzato e suddiviso in tre sotto-modelli, rappresentati nella Figura 5 (Roberts et al., 2016):

1. Modello della Topical Prevalence;
2. Modello della Topical Content;
3. Modello della Core Language.

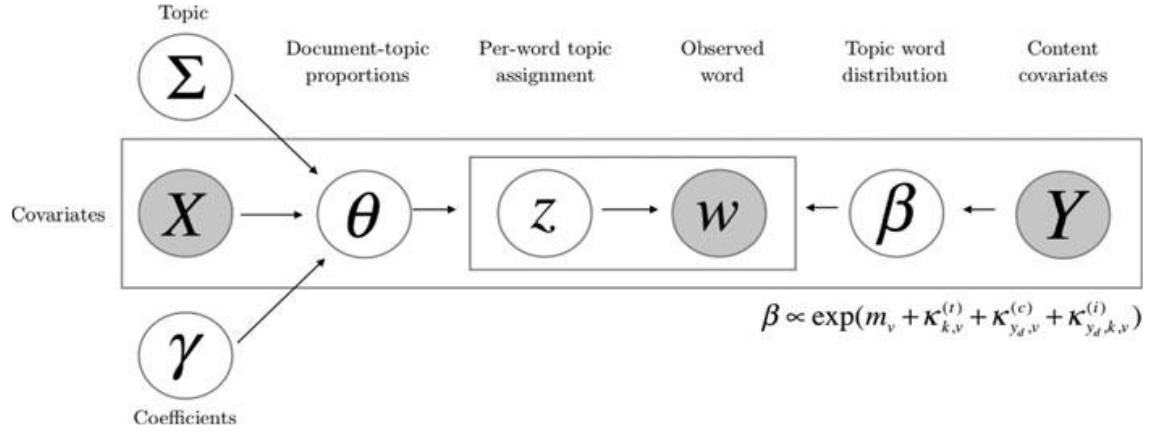


Figura 5: Rappresentazione del metodo STM (Roberts et al., 2016). Il Modello è un insieme di tre sotto modelli (Topical Prevalence, Topical Content e Core Language).

L'assunzione di base sul modello della topical prevalence per un determinato documento ϑ_d è che essa segua una distribuzione normale-logistica, rappresentata dalla seguente formula, in funzione del documento X_d , del topic k e della relazione tra covariata e topic γ :

$$\vartheta^d | X_{dY}, k \sim \text{LogisticNormal}(\mu = X_{dY}, k) \quad (1.1)$$

Si può definire a questo punto una nuova matrice ϑ data dal risultato di $D \times K$ e che presenta sulle righe i documenti del database, sulle colonne i K topic e all'interno la prevalenza tematica $\vartheta_{d,k}$ del d -esimo documento in relazione all'argomento di riferimento (Roberts et al., 2019).

Il secondo modello è rappresentato da una relazione che include al suo interno le deviazioni del topic, della covariata e della relazione covariata-topic:

$$\beta_{k,v} \propto \exp(m_v + k_{v^k} + k_{v^{y^k}} + k_{v^{y,k}}) \quad (1.2)$$

La matrice di riferimento (β^T), ha una dimensione $V \times K$; essa ha sulle righe i singoli termini presenti nel vocabolario, sulle colonne i topic e all'interno la topical content $\beta_{k,v}$ che relaziona il vocabolo all'argomento considerato (Roberts et al., 2019).

L'ultimo modello combina i risultati delle precedenti due matrici, assegnando un topic a ogni documento d , in base alla parola n . Per ogni topic scelto, inoltre, viene calcolata la probabilità di estrazione delle parole all'interno di quel documento ($w_{d,n}$).

Le formule di riferimento sono le seguenti (Roberts et al., 2019):

$$z_{d,n} | \vartheta_d \sim \text{Multinomial}(\vartheta_d) \quad (1.3)$$

$$w_{d,n} | z_{d,n}, \beta_{d,k=z_{d,n}} \sim \text{Multinomial}(\beta_{d,k=z_{d,n}}) \quad (1.4)$$

1.4 Le Carte di Controllo

Le Carte di Controllo sono degli strumenti utilizzati al fine di monitorare in modo continuo una determinata caratteristica di un processo o una determinante di qualità di un prodotto o servizio (Montgomery, 2009). In generale, si tratta di un diagramma che presenta sull'asse delle ascisse i campioni da considerare, mentre sull'asse delle ordinate sono presenti i valori della determinante analizzata. Ogni carta di controllo presenta inoltre tre linee di riferimento fondamentali: la Linea Centrale di Controllo (CL), che rappresenta il valore atteso della variabile analizzata, e i Limiti Superiore e Inferiore di Controllo, abbreviati rispettivamente con UCL e LCL e calcolati a partire dalla variabilità intrinseca del processo. Questi ultimi definiscono l'intervallo entro cui ci si aspetta che i dati si distribuiscano quando il processo è sotto controllo statistico; il superamento dei limiti o l'emergere di andamenti anomali rispetto alla linea centrale, costituisce un segnale di allarme che richiede ulteriori indagini (Franceschini et al., 2020).

È possibile distinguere due categorie principali di Carte di Controllo, che sono quelle “per attributi” e quelle “per variabili” (Montgomery, 2009). Le Carte per variabili vengono principalmente utilizzate quando i valori sono quantificabili tramite unità continue, come ad esempio la massa, la lunghezza e il tempo; le Carte per attributi, invece, vengono impiegate nei casi in cui i dati del campione sono rappresentabili solamente tramite degli stati, in numero finito, come per esempio lo stato di conformità (difettoso o non difettoso) (Franceschini et al., 2020; Montgomery, 2009). La corretta scelta della Carta di Controllo dipende inoltre dalla numerosità del campione che viene selezionato, la quale influisce anche sul valore dei parametri utilizzati per la costruzione della CL e dei limiti (Montgomery, 2009).

In determinate situazioni, tuttavia, l'utilizzo delle carte tradizionali può non essere ottimale. Questo accade, ad esempio, quando la media del processo non risulta stabile nel tempo o quando si desidera rilevare variazioni di piccola entità. In questi casi si può ricorrere ad altre tipologie di carte come quelle a media mobile, progettate per essere più sensibili ai cambiamenti gradualmente e capaci di cogliere tempestivamente deviazioni rispetto al comportamento atteso.

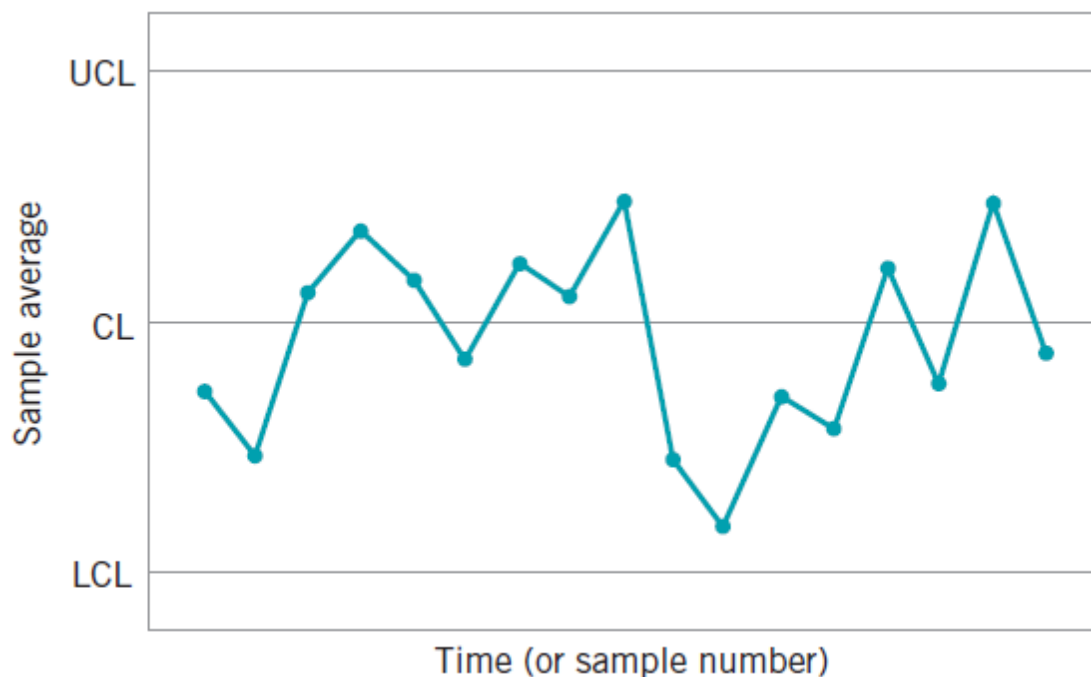


Figura 6: Esempio di una carta di controllo (Montgomery, 2009). In azzurro è raffigurato il processo, mentre le tre linee orizzontali rappresentano la linea centrale (CL), il limite superiore (UCL) e il limite inferiore (LCL) della carta utilizzata.

La Figura 6 mostra un esempio di una caratteristica di qualità analizzata tramite una carta di controllo. In questo caso, il processo risulta per intero all'interno della zona compresa tra i due limiti, pertanto può essere considerato come in controllo. Alternativamente, se il processo non fosse stato in controllo, sarebbero state necessarie alcune azioni correttive. Infine, esistono particolari casi in cui il processo rientra all'interno dei due limiti, ma i valori analizzati non presentano un andamento

casuale, rendendo necessario quindi l'intervento correttivo, trattandosi di un fuori controllo.

Le formule per costruire la linea centrale e i limiti superiore e inferiore di una carta di controllo sono le seguenti:

$$CL = \mu_w \quad (1.5)$$

$$UCL = \mu_w + L\sigma_w \quad (1.6)$$

$$LCL = \mu_w - L\sigma_w \quad (1.7)$$

dove:

- μ_w : media campionaria della statistica w .
- σ_w : deviazione standard campionaria della statistica w .
- L : distanza tra la linea centrale e il limite di riferimento, solitamente impostata a $L=3$, associando una probabilità di errore del primo tipo $\alpha=0,0027$, che indica la probabilità di rilevare un falso allarme, ovvero i casi in cui viene evidenziato un fuori controllo mentre invece in realtà non è così. Maggiore è il valore di L , minore sarà il valore di α ; al contrario, aumenterà il rischio di secondo tipo β , ovvero la probabilità di non rilevare un fuori controllo.

La fase di “impianto” di una Carta di Controllo prevede infine la scelta della frequenza di campionamento, derivante da valutazioni non oggettive relative alla carta e al campione selezionati.

Nella fase di “controllo” si studia il comportamento dei campioni, andando alla ricerca dell'anomalia. È doveroso sottolineare che, se il processo risulta in controllo, non vuol dire che necessariamente il prodotto o servizio analizzato sia conforme alle specifiche. Oltre alla presenza di punti al di fuori dei limiti superiori e inferiori,

possono essere utilizzati anche altri criteri per l'identificazione dei fuori controllo; tali criteri prendono il nome di Western Electric Rules (Western Electric, 1956; Montgomery, 2009).

1.4.1 Carta di controllo $\bar{X}$ -S

La Carta di controllo $\bar{X}$ -S appartiene alla categoria delle carte di controllo per variabili, ovvero dove il valore da monitorare può essere espresso tramite unità quantitative continue, come ad esempio tempo, massa e lunghezza.

La suddetta carta risulta di grande utilità nei casi in cui la numerosità n del campione analizzato è variabile e maggiore di 10 (Montgomery, 2009).

La fase di impianto inizia con il calcolo per ogni campione della media $\bar{x}_i$ e della deviazione standard s_i , tramite le seguenti due formule:

$$\bar{x}_i = \frac{\sum_{j=1}^n x_{j,i}}{n} \quad (1.8)$$

$$s_i = \sqrt{\frac{\sum_{j=1}^n (x_{j,i} - \bar{x}_i)^2}{n-1}} \quad (1.9)$$

dove:

i = indice che conta il numero dei campioni;

n = numerosità del campione.

Una volta ottenuti i due parametri, il passaggio successivo consiste nel calcolare la grande media del processo $\bar{\bar{x}}$ e lo scarto tipo medio $\bar{s}$:

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} \quad (1.10)$$

$$\bar{s} = \frac{\sum_{i=1}^k s_i}{k} \quad (1.11)$$

dove in questo caso k equivale al numero totale di campioni.

Per il calcolo dei limiti di controllo è necessario inoltre conoscere il valore di alcuni parametri, tutti dipendenti dalla numerosità del campione n (tendenzialmente quando è maggiore di 25), ovvero c_4 , A_3 , B_3 e B_4 :

$$c_4 \cong \frac{4(n-1)}{4n-3} \quad (1.12)$$

$$A_3 = \frac{3}{c_4 \sqrt{n}} \quad (1.13)$$

$$B_3 = 1 - \frac{3}{c_4 \sqrt{2(n-1)}} \quad (1.14)$$

$$B_4 = 1 + \frac{3}{c_4 \sqrt{2(n-1)}} \quad (1.15)$$

Ricavati i valori dei suddetti parametri, è possibile a questo punto procedere con la costruzione delle carte di controllo (Montgomery, 2009).

Per la Carta $\bar{X}$ si avrà:

$$\begin{cases} UCL_{\bar{X}} = \bar{\bar{x}} + A_3 \bar{s} \\ CL_{\bar{X}} = \bar{\bar{x}} \\ LCL_{\bar{X}} = \bar{\bar{x}} - A_3 \bar{s} \end{cases} \quad (1.16)$$

mentre per la carta S:

$$\begin{cases} UCL_S = B_4 \bar{s} \\ CL_S = \bar{s} \\ LCL_S = B_3 \bar{s} \end{cases} \quad (1.17)$$

1.4.2 Applicazione delle Carte di Controllo per l'Analisi della digital VoC

L'analisi della digital Voice of Customer permette alle imprese di superare alcune delle problematiche riscontrate nelle tecniche tradizionali di raccolta delle informazioni. Studiando la digital VoC risulta molto più semplice e affidabile, infatti, intercettare i giudizi dei clienti (Sweeney et al., 2008).

Se un certo argomento viene discusso, è molto probabile che faccia parte degli aspetti critici che influiscono sulla qualità percepita del prodotto o servizio di riferimento. Gli algoritmi di Topic Modeling sono in grado di individuare quali sono questi argomenti discussi e come la percezione della qualità varia nel tempo; la variabilità della Topical Prevalence (TP), quindi del risultato restituito dall'algoritmo, può essere studiata utilizzando come strumento le carte di controllo (Barravecchia et al., 2022; Barravecchia et al., 2020; Mastrogiacomo et al., 2021).

Gli studi riguardanti l'utilizzo del controllo statistico dei processi sugli UGC sono ancora pochi in letteratura. Tra i primi studiosi è possibile citare Lo (2008), il quale utilizzò delle carte di controllo per attributi per monitorare la proporzione dei reclami dei consumatori su un sito web, ottenuta tramite l'algoritmo Support Vector Machine (SVM). Ashton ed Evangelopoulos, nei loro studi del 2012, 2014 e 2015, hanno utilizzato l'algoritmo LSA per individuare l'evoluzione di diversi argomenti nel tempo,

proponendo per il loro monitoraggio delle particolari carte di controllo per la media mobile esponenziale. Nel 2019, Liang e Wang hanno studiato una metodologia che prende in considerazione anche la distribuzione del “sentiment” del consumatore, facendo una distinzione tra pareri positivi e negativi.

Per poter implementare correttamente una Carta di Controllo, è necessario considerare i seguenti quattro aspetti (Barravecchia et al., 2020; Mastrogiacomo et al., 2021):

- **Il tipo di Carta:** è possibile, ad esempio, utilizzare più tipi di variabili, come quelle continue, dove si considera il livello di discussione di una determinante, o discrete, dove si analizza la determinante più discussa all'interno di un record.
- **Il periodo di campionamento:** la frequenza con la quale i record vengono raggruppati in un campione è molto importante, in quanto influisce sui risultati.
- **La gestione delle determinanti che presentano dei trend:** è possibile avere delle determinanti che presentano un andamento nel tempo stazionario, in aumento o in diminuzione; queste particolari tendenze potrebbero aver bisogno di carte apposite per il loro monitoraggio (Barravecchia et al., 2020; Mastrogiacomo et al., 2021).
- **La gestione di nuove determinanti di qualità:** la scomparsa e la nascita di alcune determinanti di qualità necessitano di uno studio nel dettaglio per poter selezionare la giusta Carta di Controllo.

Con l'applicazione delle Carte di Controllo alle determinanti di qualità negli UGC, è possibile rilevare delle anomalie, ovvero la variabilità sistematica (dovuta a

distorsioni nel processo) e la variabilità naturale (dovuta all'effetto cumulativo di piccole cause non controllabili). L'individuazione di un fuori controllo potrebbe indicare una variazione del gradimento del cliente in relazione a un particolare aspetto, offrendo degli importanti spunti alle organizzazioni per poter analizzare il dato e procedere con l'opportuna azione correttiva (Barravecchia et al., 2022).

CAPITOLO 2 – METODOLOGIA

La metodologia adottata per intercettare le determinanti di qualità è un processo lineare composto da vari step, che verranno descritti nei paragrafi di questo capitolo. Una sintesi del percorso è illustrata in Figura 7 (Mastrogiacomo et al., 2021): da una fase di raccolta della Digital VoC, si arriva a una definizione chiara e validata delle determinanti di qualità che caratterizzano il dataset; l'ultimo step, riguardante l'analisi dei risultati ottenuti, verrà introdotto nel capitolo successivo, dedicato alla concreta applicazione della suddetta metodologia al caso di studio in esame.

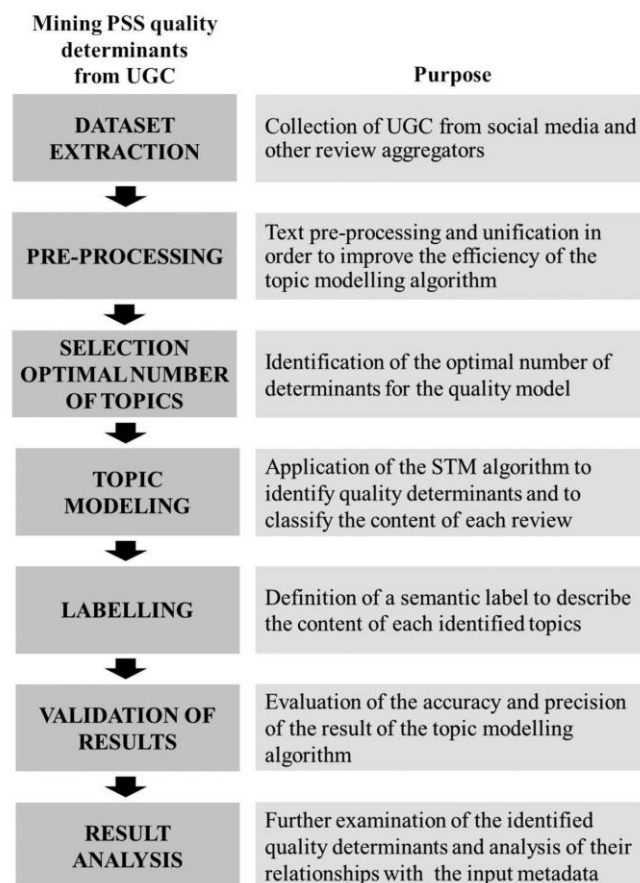


Figura 7: Metodologia sviluppata per il Topic Modeling (Mastrogiacomo et al., 2021). Essa è rappresentata da una serie di sette step consequenziali, partendo dall'estrazione del database e finendo con l'analisi dei risultati. Una volta terminati tutti gli step, è possibile proseguire con l'applicazione delle Carte di controllo.

2.1 Preparazione del Database

Il processo proposto inizia con l'estrazione e la preparazione del database, ovvero le attività che consistono nel raccogliere una grande quantità di UGC da social media e da aggregatori di recensioni, come ad esempio TripAdvisor, Trustpilot e Yelp. L'insieme di queste attività prende il nome di "web scraping" e grazie a questa tecnica è possibile estrapolare da un sito web una serie di informazioni su una recensione, come il titolo, il nome dell'impresa recensita, l'autore, la data, il punteggio assegnato e la nazionalità, oltre ovviamente al corpo della recensione stessa (Mastrogiacomo et al, 2021). Esistono attualmente molteplici applicazioni, librerie e servizi software che si occupano di web scraping, tra cui BeautifulSoup (Richardson, 2004), Scrapy (Scrapy developers, 2008), Octoparse (Octopus Data Inc., 2016), Diffbot (Diffbot Technologies Corp., 2008), ParseHub (ParseHub Inc., 2015), Apify (Apify Technologies s.r.o., 2015) e Content Grabber (Sequentum International, 2011).

Nonostante il web scraping sia una tecnica molto potente ed efficace per ottenere un grande insieme di informazioni, essa potrebbe essere considerata come controversa, sollevando questioni legali relative al diritto d'autore e ai termini di servizio; per questo motivo, alcune società tra cui Google e Meta stanno limitando la possibilità di raccogliere i dati all'interno delle loro piattaforme (Zhao, 2017; Zaki e McColl-Kennedy, 2020).

È di fondamentale importanza, per un risultato di qualità, ottenere un dataset più vario e completo possibile, cercando di raccogliere informazioni diversificate tra loro (Mastrogiacomo et al, 2021). Questo obiettivo è raggiungibile, per esempio,

ricercando i dati tra più siti web, includendo recensioni che appartengono a uno spettro temporale abbastanza ampio da permettere un'analisi ottimale.

2.2 Pre-Processing

Per aumentare l'efficienza dell'algoritmo di topic modeling utilizzato, è necessario pre-elaborare e unificare i record contenuti all'interno del database selezionato, con lo scopo di ridurre il più possibile il rumore delle recensioni (Meyer et al., 2008; Tirunillai e Tellis, 2014; Guo et al., 2017; Salur e Aydin, 2020).

Le operazioni di pre-processamento più comuni sono le seguenti:

- Conversione del testo in minuscolo in modo da eliminare l'ambiguità con le lettere maiuscole.
- Eliminazione di numeri e punteggiatura, in quanto essi non forniscono informazioni utili per l'analisi.
- Rimozione di congiunzioni e preposizioni, che prendono il nome di "Stopwords", come "and", "the", "when", "which", "where", "is", "at", "on".
- Rimozione delle parole che presentano più di 15 caratteri o meno di 2 caratteri. Parole molto lunghe, oltre ad essere rare, sono spesso derivate da errori di scraping, concatenazioni di più termini, sequenze non significative o URL, mentre parole di uno o due caratteri non hanno un particolare valore semantico. Rimuovere queste tipologie di parole, pertanto, aiuta a ridurre il rumore e migliorare la qualità dei dati.
- Rimozione delle parole che all'interno del dataset si presentano con una bassa frequenza, in quanto potrebbero creare una distorsione nell'applicazione dell'algoritmo. Queste parole non portano un significativo contributo al modello in quanto sono spesso derivate da errori di digitazione, nomi propri molto rari o abbreviazioni; mantenerle, quindi, farebbe aumentare inutilmente la dimensione del vocabolario generato dal software.

- Applicazione del processo di “stemming”, che consiste nel ridurre il numero di parole andando a rimuovere le desinenze morfologiche di alcune parole inglesi (Jivani, 2011). Ad esempio, le parole “likes”, “liked”, “likely”, e “liking” vengono trasformate nella parola “like”. Anche in questo caso, come nel punto precedente, l'intento principale è quello di non aumentare la dimensionalità del vocabolario, in modo da ottenere una maggiore efficienza.
- Eliminazione delle parole che generalmente non fanno parte al contesto e all'argomento scelto, come “another”, “mean”, “review”, “made”, “problem”, “done”, “did”. La presenza di parole comuni come quelle appena citate, non aggiunge alcun valore informativo durante l'analisi dei topic.
- Identificazione e accorpamento in n-grammi delle parole più comuni che solitamente risultano combinate tra loro; alle parole “customer service” per esempio verrà eliminato lo spazio, facendole diventare un'unica parola “customerservice” (Durrani et al., 2015). Questa operazione viene effettuata in quanto alcuni concetti nei testi non si esprimono tramite parole singole, ma con combinazioni di termini, come quelli appena citati. L'accorpamento in n-grammi quindi permette di migliorare la rappresentazione del testo e conservare il contesto semantico.

2.3 Selezione del numero di Topic

Prima dell'applicazione dell'algoritmo STM, una fase molto importante consiste nella scelta del parametro K , che rappresenta il numero di topic, dell'intero corpus di dati, da voler analizzare tramite il modello. Il metodo maggiormente utilizzato per determinare il numero ottimale di topic è stato proposto da Wallach, Mimno e McCallum nel 2009 e si basa su un approccio iterativo: una volta selezionato il parametro K , viene applicato l'algoritmo valutandone la performance; se la performance non viene ritenuta soddisfacente, si procede con l'iterazione del processo, andando a selezionare altri numeri di topic da analizzare fin quando si ritiene necessario (Wallach et al., 2009).

Per valutare la performance del modello e ottenere il numero ottimale di K , esistono varie metriche che possono essere considerate, tra cui l'Held-out likelihood, che è l'indice di riferimento utilizzato per le analisi di questa tesi. L'Held-out likelihood è una metrica che fornisce informazioni relative al modello e a quanto esso sia adatto a spiegare la variabilità complessiva dell'intero set di dati. Il modello viene creato andando ad analizzare circa il 90% degli UGC disponibili, restituendo la probabilità che il modello stesso possa essere coerente con il restante 10% non utilizzato (Scott e Baldrige, 2013; Roberts et al., 2014).

Oltre all'Held-out likelihood esistono anche altre metriche, come la Semantic Coherence, la Lower Bound e la Residuals (Roberts et al., 2014); la Figura 8 offre una panoramica delle quattro metriche citate, tramite un output ottenibile usando il software RStudio.

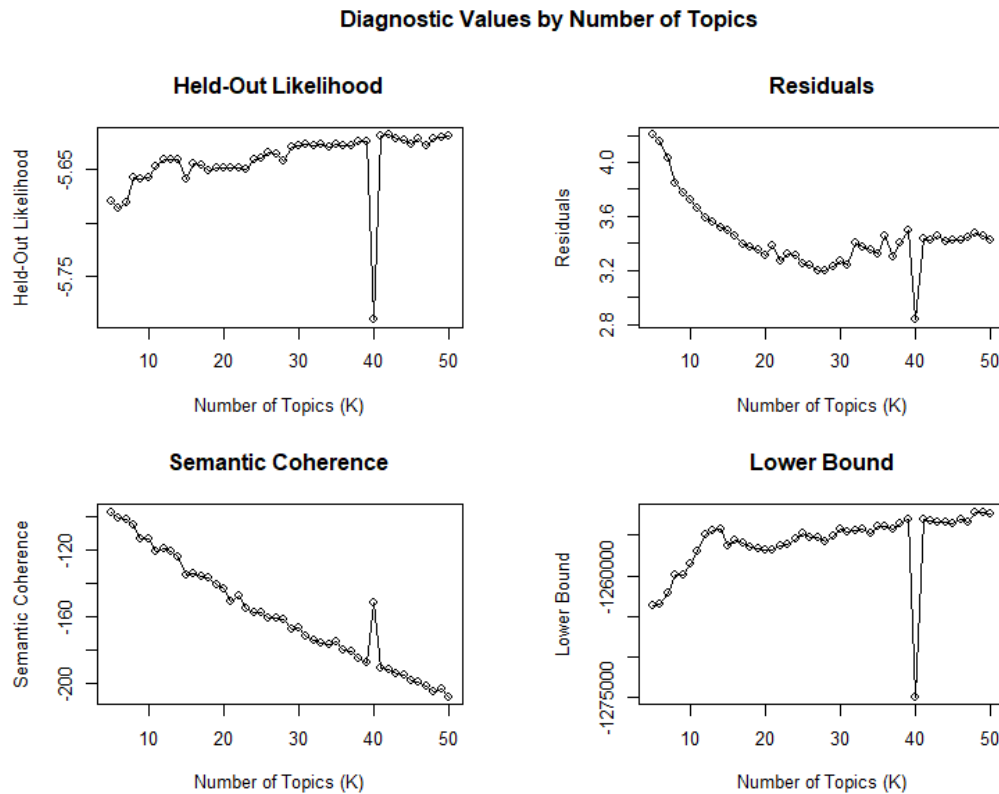


Figura 8: Esempio di una rappresentazione grafica delle metriche "Held-out likelihood", "Residuals", "Semantic Coherence" e "Lower Bound". Per il presente lavoro di tesi verrà sempre considerata la metrica Held-out likelihood.

Il numero ottimale di topic, andando a considerare l'Held-out likelihood, si ottiene calcolando la metrica per diversi valori di k e osservando l'andamento sul relativo grafico. Generalmente si sceglie il valore di k corrispondente al massimo della curva, in quanto rappresenta la migliore capacità predittiva. Alternativamente, se dopo un certo punto l'andamento della metrica diventa marginale e la curva assume un andamento stazionario, si preferisce selezionare il valore dal quale parte tale comportamento. Nell'esempio in Figura 8, è presente un andamento stazionario che parte da un valore $k=12$, che quindi potrebbe essere preso come riferimento per una prima iterazione per valutare la bontà del modello.

2.4 Etichettatura dei Topic

Una volta selezionato il numero K che si vuole analizzare, è possibile lanciare l'algoritmo STM indicato nella sezione 1.3.3; esso restituisce, per ogni topic, una lista delle parole più significative che rappresentano il dataset. A questo punto si procede con la fase di etichettatura dei topic, anche chiamata *labelling*. Il modello, infatti, restituisce per ogni argomento analizzato un insieme di parole con una determinata probabilità di presenza all'interno del corpus di UGC, ma non assegna automaticamente un nome. Il labelling consiste quindi nell'interpretare queste parole e sintetizzarle in un'etichetta semantica coerente, in modo da rendere il topic di facile comprensione e utilizzabile per le fasi successive della ricerca. Ad esempio, un topic caratterizzato da alcuni termini come *delivery*, *shipping* e *delay*, potrebbe essere etichettato come *tempi di consegna*.

La fase di Labelling può essere svolta in modo automatico dall'algoritmo, oppure in modo manuale dall'uomo. Nel primo caso, l'algoritmo propone un'etichetta basata sulle parole con maggiore Topical Prevalence o su metriche di coerenza semantica; si tratta di un approccio rapido ma talvolta produce risultati poco interpretabili. Nel caso manuale, l'analista esamina le parole di ciascun topic e assegna un'etichetta in base al loro significato complessivo. Si tratta di una fase abbastanza critica del processo, in quanto un'etichetta, se non assegnata con precisione, potrebbe falsare i dati nella successiva fase di validazione; si ritiene, pertanto, che il metodo di assegnazione migliore sia ancora quello manuale, seppur il più dispendioso (Blei, 2012).

Esistono vari criteri che permettono all'algoritmo di ottenere una lista delle parole più rappresentative all'interno dei vari topic, tra i quali (Roberts et al., 2019):

- **Highest probability:** è un criterio che si basa sul calcolo delle parole con le più alta probabilità di comparire all'interno dell'argomento.
- **FREX:** le parole conteggiate per questo criterio sono quelle più frequenti, ma anche esclusive.
- **Lift:** il punteggio è calcolato considerando la distribuzione delle parole all'interno dell'argomento, dividendola per la distribuzione empirica della probabilità del conteggio delle parole.
- **Score:** vengono considerate le parole con il punteggio più alto assegnato da un algoritmo LDA.

Ad esempio, considerando un topic relativo ad alcune recensioni sui ristoranti, l'algoritmo potrebbe generare parole come *food*, *pizza*, *service*, *waiter*, *price*, *experience*.

- Con il criterio Highest probability verrebbero riportate le parole più frequenti, come ad esempio *food*, *service*, *price*.
- Con FREX, che privilegia l'esclusività, emergerebbero termini più distintivi come *pizza* e *waiter*.
- Con Lift si avrebbero parole molto specifiche per ogni topic e poco diffuse altrove, come *waiter* ed *experience*.
- Con Score, infine, l'algoritmo produrrebbe parole rilevanti come *pizza*, *service* ed *experience*.

Per ottenere un'etichettatura più precisa e coerente possibile con le parole restituite dal metodo, è possibile considerare una combinazione tra i criteri appena elencati; in aggiunta, l'algoritmo è in grado di restituire, per ogni argomento, un numero

selezionato di recensioni più significative, in modo da avere una migliore panoramica sul topic degli UGC (Roberts et al, 2019; Mastrogiacomo et al., 2021).

Un esempio, potrebbero essere le seguenti tre recensioni:

- **Recensione 1:** The *food* was amazing and the *pizza* was one of the best I've ever had. The *service* was quick and friendly, and the *waiter* even recommended a great wine. Totally worth the *price*!
- **Recensione 2:** I had a wonderful dining *experience*. The *waiter* was very attentive and made sure everything was perfect. The *pizza* came out hot and delicious, and overall the service was excellent compared to the low *price*.
- **Recensione 3:** The *food* quality was average, but the *service* was really disappointing. The *waiter* seemed distracted, and for the *price* we paid, the *experience* could have been much better. Only the *pizza* saved the evening.

2.5 Validazione dei dati

Durante la fase di validazione dei dati, si stabilisce la bontà dei risultati ottenuti tramite il modello, ma nonostante la sua importanza, in letteratura non sono state proposte molte metodologie pratiche. Esistono una serie di metriche automatiche, fornite da algoritmi, che permettono di misurare e testare le performance dei topic, ma esse non tengono in considerazione la semantica e il significato degli argomenti analizzati, con il rischio di effettuare possibili errori nei calcoli. Dall'altro lato, alcuni studiosi hanno proposto alcune metodologie di tipo supervisionato, che richiedono la presenza umana; i risultati, in questo senso, saranno più precisi, a scapito del tempo e delle risorse impiegate che saranno maggiori rispetto agli algoritmi automatici (Barravecchia et al., 2021; Chang et al., 2009).

La Figura 9 mostra una panoramica delle differenze che è possibile riscontrare tra i criteri automatici e i criteri supervisionati.

	Automatic criteria	Supervised criteria
Required time	Low	High
Fully automated evaluation process	Yes	No
Semantic evaluation	No	Yes
Need for human input	No	Yes
Topic intelligibility evaluation	No	Yes
Suitable for setting model parameters	Yes	No
Sample size	Complete database	Sample of documents

Figura 9: Differenze principali tra le metriche automatiche e le metriche supervisionate per la validazione dei dati (Barravecchia et al., 2021)

Il metodo supervisionato di validazione dei risultati proposto da Barravecchia et al. (2021) si basa su quattro step, rappresentati in Figura 10:

1. Estrazione di un campione casuale e assegnazione umana dei topic;

2. Assegnazione automatica dei topic;
3. Confronto tra i risultati ottenuti nei precedenti due punti;
4. Calcolo delle metriche.

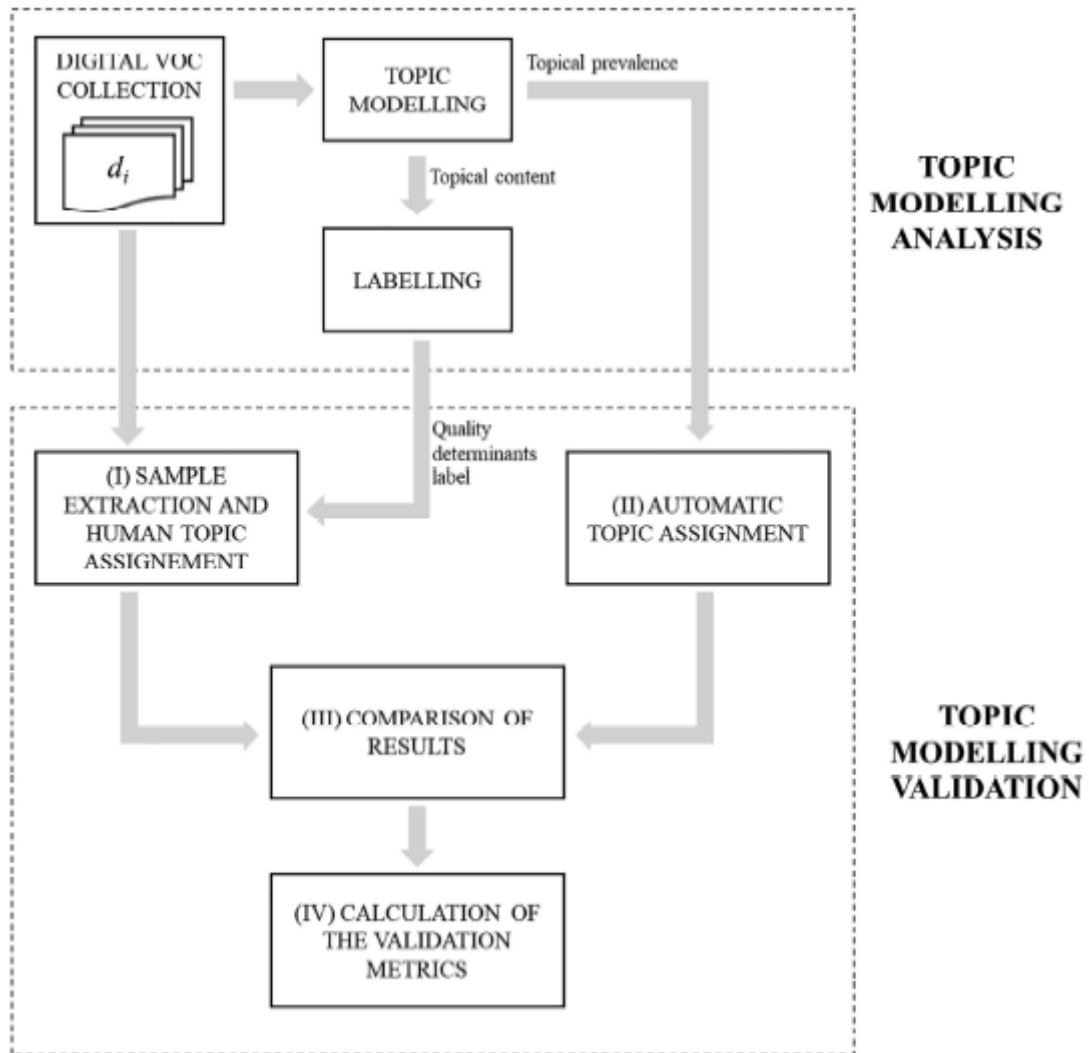


Figura 10: Principali step da effettuare per la validazione dei dati (Barravecchia et al., 2021). Dopo l'assegnazione umana e automatica dei topic di un campione di recensioni, viene fatto un confronto tramite il calcolo di alcune metriche.

Dopo aver estratto in modo casuale dal database un certo numero di documenti (ad esempio $n=100$), si procede con l'assegnazione umana dei topic a cui appartengono tali contenuti. Successivamente, si passa all'assegnazione automatica degli

argomenti, partendo dalla topical prevalence ottenuta. Esistono tre approcci differenti per questo step:

- **Highest probability:** per ogni contenuto viene assegnato il topic che presenta il valore maggiore di topical prevalence.
- **Soglia statica:** sono assegnati tutti i topic che superano una certa soglia statica definita a priori.
- **Soglia dinamica:** in questo caso la soglia non è più fissa, ma varia in base alla recensione selezionata. Il calcolo della soglia si basa sulle seguenti formule:

$$DT_i = Q3_i + (1,5 \times IQR_i) \quad (2.5)$$

$$IQR = Q3 - Q1 \quad (2.6)$$

dove Q3 rappresenta il terzo quartile, Q1 il primo quartile, IQR il range interquartile e DT il valore della soglia dinamica. Anche qui, tutti i valori che superano la soglia dinamica vengono considerati.

Il confronto tra le assegnazioni ottenute con il metodo automatico e il metodo umano, viene effettuato tramite il calcolo di quattro indicatori (Costa et al., 2007):

- **True Positive (tp):** l'assegnazione umana e quella automatica coincidono, riconoscendo lo stesso topic per il contenuto analizzato.

Esempio con la recensione "The waiter was kind and the service was excellent.":

- Assegnazione umana: stabilisce che si parla di servizio.
- Assegnazione automatica: riconosce anch'essa che l'argomento è il servizio clienti.

- **True Negative (tn):** sia l'assegnazione umana che quella automatica non riconoscono un topic.

Esempio con la recensione "The pizza was delicious and the crust was perfect.":

- Assegnazione umana: stabilisce che *non* si parla di servizio.
- Assegnazione automatica: riconosce anch'essa che l'argomento *non* è il servizio clienti.

- **False Positive (fp):** l'assegnazione automatica riconosce un topic che quella umana non ha riconosciuto (Errore del I tipo).

Esempio con la recensione "The pizza was too high compared to other restaurants."

- Assegnazione umana: stabilisce che si parla di prezzo
- Assegnazione automatica: stabilisce che si tratta di servizio clienti.

- **False Negative (fn):** l'assegnazione umana riconosce un topic che l'assegnazione automatica non ha riconosciuto (Errore del II tipo).

Esempio con la recensione "The waiter ignored our table and the service was terrible."

- Assegnazione umana: stabilisce che si parla di servizio clienti.
- Assegnazione automatica: stabilisce che non si tratta di servizio clienti.

L'ultimo step, dopo aver effettuato il confronto, consiste nel calcolo di una serie di indicatori, le cui formule sono rappresentate in Figura 11 (Costa et al., 2007; Franceschini et al., 2019; Maria Navin e Pankaja, 2016; Zaki e McColl-Kennedy, 2020):

		Human topic assignment (true condition)			
		T_i existence	T_i non-existence	<i>Accuracy</i> $\frac{\sum_{i=1}^n tp_i + \sum_{i=1}^n tn_i}{\sum_{i=1}^n tp_i + \sum_{i=1}^n tn_i + \sum_{i=1}^n fp_i + \sum_{i=1}^n fn_i}$	
Automatic topic assignment	T_i existence	True Positive (tp) Correct inference	False Positive (fp) Type I error	<i>Precision</i> $\frac{\sum_{i=1}^n tp_i}{\sum_{i=1}^n tp_i + \sum_{i=1}^n fp_i}$	<i>False discovery rate</i> $\frac{\sum_{i=1}^n fp_i}{\sum_{i=1}^n tp_i + \sum_{i=1}^n fp_i}$
	T_i non-existence	False Negative (fn) Type II error	True Negative (tn) Correct inference	<i>False omission rate</i> $\frac{\sum_{i=1}^n fn_i}{\sum_{i=1}^n fn_i + \sum_{i=1}^n tn_i}$	<i>Negative predictive value</i> $\frac{\sum_{i=1}^n tn_i}{\sum_{i=1}^n fn_i + \sum_{i=1}^n tn_i}$
		<i>Recall</i> $\frac{\sum_{i=1}^n tp_i}{\sum_{i=1}^n tp_i + \sum_{i=1}^n fn_i}$	<i>Fall-out</i> $\frac{\sum_{i=1}^n fp_i}{\sum_{i=1}^n fp_i + \sum_{i=1}^n tn_i}$	<i>F₁ Score</i> $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$	
		<i>Miss rate</i> $\frac{\sum_{i=1}^n fn_i}{\sum_{i=1}^n tp_i + \sum_{i=1}^n fn_i}$	<i>Specificity</i> $\frac{\sum_{i=1}^n tn_i}{\sum_{i=1}^n fp_i + \sum_{i=1}^n tn_i}$		

Figura 11: Metriche per la validazione dei dati (Costa et al., 2007; Franceschini et al., 2019; Maria Navin e Pankaja, 2016; Zaki e McColl-Kennedy, 2020; Barravecchia et al., 2021). Tutte le formule delle metriche descritte vengono costruite partendo dai valori veri positivi, falsi positivi, veri negativi e falsi negativi.

Ognuna di queste metriche presenta inoltre dei valori target all'interno dei quali esse devono rientrare per considerare come valido l'insieme dei dati analizzati. Quando invece i valori si discostano abbondantemente dai valori target, è possibile procedere tramite le seguenti azioni correttive, elencate dalla meno impattante alla più impattante in termini di struttura del modello (Barravecchia et al., 2021):

1. Verifica della rappresentatività delle etichette assegnate ai vari topic;
2. Verifica della correttezza dei parametri di input (numero topic);

3. Verifica della correttezza nella scelta dei metadati, se considerati nel modello;
4. Verifica dell'omogeneità e della pertinenza del database.

I valori soglia appena nominati sono rappresentati in Figura 12:

Indicators	Range	Direction	Target values
Accuracy	[0;1]	High is good	>0.95
Recall	[0;1]	High is good	>0.70
Precision	[0;1]	High is good	>0.70
F ₁ score	[0;1]	High is good	>0.70
Miss-rate	[0;1]	Low is good	<0.20
Fall-out	[0;1]	Low is good	<0.05
Specificity	[0;1]	High is good	>0.90
False omission rate	[0;1]	Low is good	<0.05
False discovery rate	[0;1]	Low is good	<0.05
Negative predictive value	[0;1]	High is good	>0.90

Figura 12: Valori soglia delle metriche per la validazione dei dati (Barravecchia et al., 2021).

CAPITOLO 3 – CASO DI STUDIO

3.1 Preparazione del Database e Pre-processing

Per la prima fase del lavoro impostato sul presente caso di studio, è stato scelto “Octoparse” come strumento di web scraping al fine di ottenere un database più vario possibile (Octopus Data Inc., 2016). Tramite il suddetto software è stato possibile analizzare gli UCG di quattro dei principali competitor del mercato del food delivery: Glovo, Just Eat, Uber Eats e Deliveroo. Poiché queste aziende sono presenti nel digitale principalmente sottoforma di applicazioni per telefonia mobile, l'aggregatore di recensioni selezionato per formare il database è stato il Play Store di Google, in modo da intercettare tutte le recensioni riguardanti la parte tecnica e informatica, oltre a quella dei servizi dei ristoranti, che non dipendono dalle app. L'altro sito web inizialmente scelto come fonte per la ricerca delle recensioni è Trustpilot; tuttavia, questo secondo aggregatore di UCG non è stato selezionato in quanto l'applicativo Octoparse non forniva tutti i dati necessari al proseguimento delle analisi.

I dati estrapolati per il database comprendono:

- Nome della compagnia: viene identificata l'azienda a cui si riferisce la recensione. È un'informazione molto utile per raggruppare i dati.
- Fonte: indica la piattaforma da cui la recensione è stata prelevata (Play Store Google).
- Nome dell'utente recensore: rappresenta l'autore del contenuto.

- Data della pubblicazione: riporta giorno mese e anno in cui la recensione è stata pubblicata. È fondamentale per le analisi temporali e per la valutazione dei trend.
- Rating (sottoforma di valore numerico da 1 a 5): è la valutazione numerica attribuita al servizio dall'utente, dove 1 rappresenta la minima soddisfazione e 5 la massima. È un indicatore immediato della percezione di qualità.
- Corpo della recensione: è il testo libero scritto dall'autore. È l'elemento chiave per l'impostazione delle successive fasi di preprocessing e analisi.

In tutto sono state raccolte 14.599 recensioni, in uno spettro temporale che va dal settembre del 2018 al maggio del 2023. Una volta terminata la raccolta, si è proceduto a effettuare la prima fase di Pre-processing. Sul software "R Studio" è stato caricato l'intero set di UCG e un secondo file denominato "Stopwords", contenente al suo interno una lista di 1255 parole che, tramite l'applicativo, sono state eliminate nell'intero corpus del primo documento. Successivamente, sempre tramite software, sono state eliminate le parole che appaiono nel database meno di 15 volte.

Come risultato, si è ottenuta l'eliminazione di 21 recensioni, per un totale di 14.578, distribuite come in Tabella 3, a fronte delle 14.599 iniziali.

Azienda	2018	2019	2020	2021	2022	2023	Totale
Deliveroo	99	495	706	713	764	1020	3797
Glovo	55	507	703	1034	849	883	4031
Uber Eats	0	0	0	0	117	3123	3240
Just Eat	142	376	487	675	651	1179	3510
Totale	296	1378	1896	2422	2381	6205	14578

Tabella 3: Distribuzione numerica delle recensioni all'interno del database iniziale, dopo le fasi di Pre-processing.

Questo nuovo database è stato quindi preso come punto di partenza per l'applicazione dell'algoritmo STM; l'intero script lanciato sul R Studio, che comprende le operazioni svolte in questo paragrafo e nei successivi, è presente e consultabile in Appendice I.

3.2 Selezione del numero di Topic

La fase successiva al Pre-processing consiste nel selezionare, anche tramite varie iterazioni se necessario, il valore K corrispondente al numero di Topic da voler analizzare con l'algoritmo STM. Per la scelta dei giusti parametri, è stato lanciato su RStudio il seguente script in Figura 13:

```
c=(5:50) #c vettore con numeri da 5 a 10
K<-c
storage <- searchK(out$documents, out$vocab, K, data = meta)
plot(storage)
write.csv(unlist(storage$results),file='optimizationresults.csv')
```

Figura 13: Script lanciato per individuare il valore K.

Il risultato è stato la creazione automatica di un file .csv con all'interno i valori delle varie metriche per ogni K compreso tra 5 e 50. La Figura 14 mostra di seguito i grafici ottenuti.

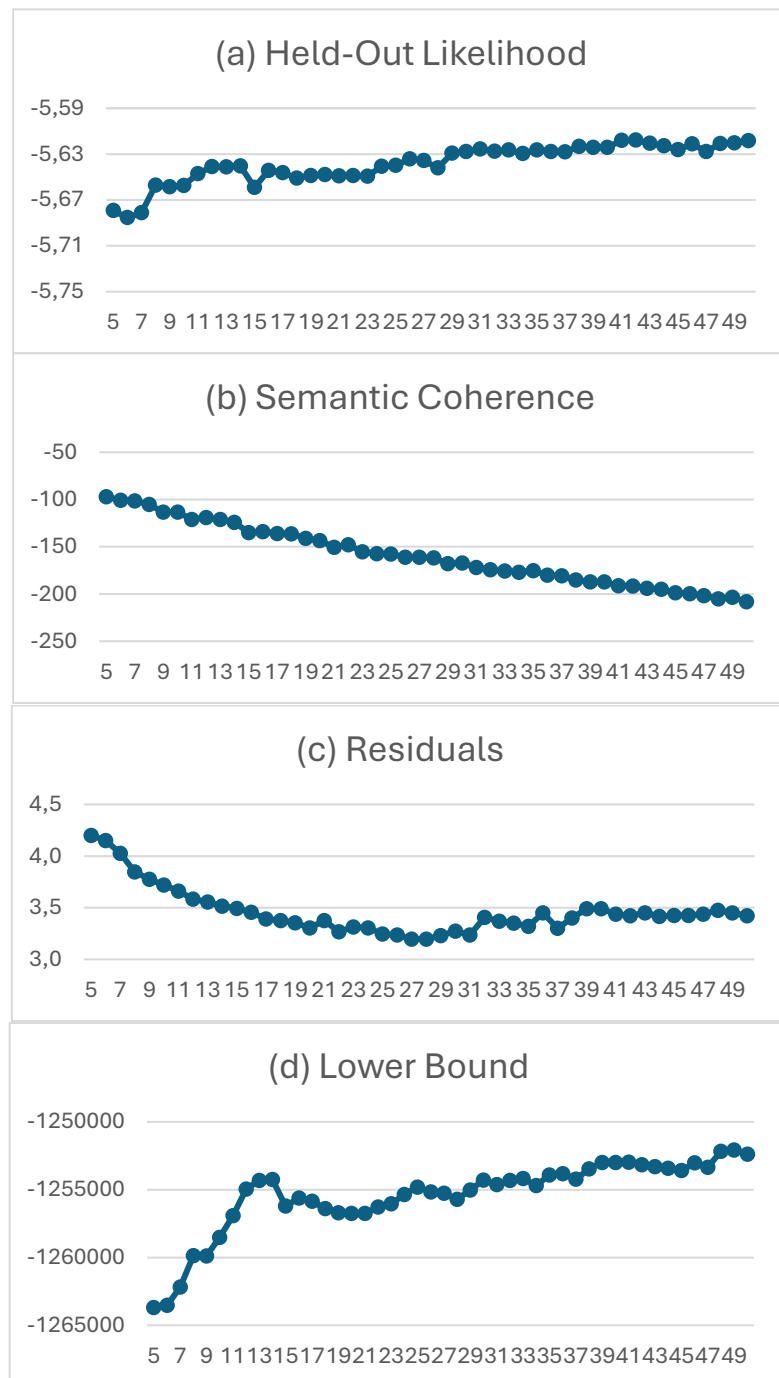


Figura 14: Grafici che illustrano i valori per le metriche Held-Out Likelihood, Semantic Coherence, Residuals e Lower Bound. Considerando la prima metrica, viene scelto un valore $k=12$.

Soffermendosi sulla metrica Held-Out Likelihood è possibile notare un andamento pressoché stazionario a partire da $K=12$; proprio questo valore, infatti, è stato selezionato come parametro di input per il lancio dell'algoritmo STM.

3.3 Applicazione dell'algoritmo Structural Topic Model (STM)

A partire dal numero scelto di topic da analizzare, ovvero 12, è stato lanciato l'algoritmo STM che ha dato come risultato sei file .csv. I primi cinque di questi contengono, per ogni recensione del database, i parametri raccolti tramite web scraping (nome dell'azienda, fonte, data della recensione, rating e corpo della recensione). L'ultimo file, invece, è un documento che associa ad ogni recensione un valore della Topical Prevalence (TP), per ogni argomento analizzato. A titolo di esempio, in Tabella 4 è indicata la distribuzione della TP di tre recensioni (R1, R2 e R3), considerando dodici topic.

	R1	R2	R3
TP Topic 1	0,10	0,05	0,15
TP Topic 2	0,05	0,10	0,05
TP Topic 3	0,20	0,05	0,05
TP Topic 4	0,15	0,05	0,10
TP Topic 5	0,05	0,15	0,05
TP Topic 6	0,05	0,20	0,05
TP Topic 7	0,10	0,05	0,20
TP Topic 8	0,05	0,10	0,10
TP Topic 9	0,05	0,05	0,05
TP Topic 10	0,10	0,05	0,05
TP Topic 11	0,05	0,10	0,05
TP Topic 12	0,05	0,05	0,10
Somma TP	1,00	1,00	1,00

Tabella 4: Esempio di distribuzione della Topical Prevalence su tre recensioni e dodici topic.

Come è possibile notare dalla Tabella 4 di esempio, per ogni recensione la somma delle Topical Prevalence costituisce l'unità.

Tutti i file .csv che il software RStudio ha restituito come output, sono stati aggregati in un unico e completo database; nello specifico, il documento risultante presentava 14.578 righe, ognuna per ogni recensione, e le seguenti colonne: Fonte, Nome azienda, Data della recensione, Rating della recensione, Corpo della recensione e dodici colonne rappresentanti la Topical prevalence per ogni topic.

3.4 Etichettatura dei Topic

Come specificato nel precedente capitolo, la fase di labelling è una delle più critiche dell'intero processo, in quanto un'etichettatura non corretta potrebbe portare a un'errata validazione dei dati.

Per il presente caso, è stato impostato un codice su RStudio in modo da ottenere, per ognuno dei 12 topic, una lista delle 10 parole più utilizzate secondo le quattro metriche Highest Prob, FREX, Lift e Score. Successivamente, sono state ottenute, per ogni topic, le 5 recensioni più significative. A titolo di esempio, vengono riportati di seguito i risultati riguardanti solamente il Topic 1; i dati relativi a tutti gli argomenti per intero sono presenti invece in Appendice II.

Topic 1 Top Words:

- **Highest Prob:** app, work, use, updat, open, order, restaur, fix, notif, uninstal
- **FREX:** open, notif, load, screen, updat, crash, stuck, user, button, function
- **Lift:** user-, anim, load, refresh, brows, crash, malfunc, aggress, intuit, open
- **Score:** user-, app, work, notif, open, updat, bug, screen, load, search

Topic 1:

1. Just installed the app, opened it and was greeted with a message saying that an update was needed to continue using the app. Odd, but I clicked the update link....which takes me back to the Google Play Just Eat page with options to 'Uninstall' or 'Open' the app. I click the 'Open' link....and round and round we go. Good for nothing at the moment.
2. Use MenuLog and Foodora instead. Deliberately bombards me with notifications via the app. But... you can't turn off via the app; you have to login via the website. Go to website and login, it then defaults back to the app... Please. I'll just uninstall rather than deal with this stupid loop. I turned off notifications twice and they keep coming back; every updates turn back on I guess because I never enable them; with any app. Annoying marketing making things more difficult than they need to

- be. Just allow the notifications to be turned off permanently.
3. Great service, bad app. Million of bugs. Keyboard doesn't open for search, back button gesture doesn't work, repeatedly opens some screens, closes the app after opening from notification, also has unremovable notifications
 4. It was great at the beginning but now i am very disappointed. When i open the app, most of the times (9 out of 10), it shows only the loading screen. Only when i close and open several times the app i can pass the loading screen. I tried to uninstall and install the app several times but this doesn't fix the problem. I thought that only i have this problem but several of my friends have this problem
 5. Awful, horrible app design. Basics like the back button don't work, you can get stuck in infinite loops in the referral screen with no way to exit it but to kill the app, bugs galore and terrible CX. Worst app I've ever used.

Sulla base delle parole e delle recensioni più significative per le determinanti di qualità analizzate, si è proceduto all'etichettatura di ognuna di esse, ottenendo il risultato in Tabella 5, dove è presente inoltre una breve descrizione dell'argomento:

Nome del topic	Descrizione
T1: Affidabilità dell'app	Navigabilità e intuitività dell'applicazione, presenza di eventuali bug e velocità di caricamento delle varie pagine.
T2: Accesso all'account	Accesso all'applicazione tramite varie tipologie di account (Google, Facebook, Apple, numero di telefono, e-mail). Sono incluse anche le varie modalità di pagamento previste dall'applicazione (ApplePay, GooglePay, Paypal, Satispay, carta di debito, buoni pasto).
T3: Disponibilità dei prodotti	Varietà di prodotti disponibili all'interno dei menu dei singoli ristoranti.

T4: Costi del servizio	Costi aggiuntivi e tasse da pagare per ogni ordine effettuato. Sono inclusi i costi per gli abbonamenti premium.
T5: Qualità del servizio	Qualità generale offerta dall'azienda, comprendente sia l'esperienza online che quella offline.
T6: Gestione del rimborso	Sono intese le modalità e i tempi con cui vengono concessi ed emessi i rimborsi (rimborso con buono spendibile solamente tramite la stessa app, oppure rimborso tramite la modalità di pagamento indicata in fase di ordine). Si includono anche le motivazioni per le quali un rimborso viene emesso o meno e l'ammontare dell'eventuale rimborso.
T7: Affidabilità ristorante	Qualità del cibo offerto dal ristorante e completezza del menu.
T8: Personalizzazione dell'ordine	Possibilità per l'utente di modificare i vari piatti dell'ordine aggiungendo o sottraendo ingredienti. Si intende inoltre la possibilità di inserire eventuali note per il ristorante, indicando ad esempio allergie/intolleranze o informazioni aggiuntive.
T9: Puntualità della consegna	Corrispondenza tra il tempo previsto di consegna indicato inizialmente dall'applicazione e quello effettivo.
T10: Affidabilità del servizio clienti	Velocità e completezza di risposta da parte del servizio clienti.
T11: Consegna all'indirizzo voluto e localizzazione	Possibilità di indicare, oltre all'indirizzo di consegna, anche l'effettiva posizione sulla mappa.
T12: Completezza dell'ordine	Corrispondenza tra quanto ordinato e quanto ricevuto, ad esempio se sono presenti tutte le pietanze indicate nell'ordine, comprensive di eventuali modifiche o richieste.

Tabella 5: Risultati ottenuti nella fase di labelling delle determinanti di qualità. La Tabella presenta il nome scelto per l'argomento e una breve descrizione che lo caratterizza.

3.5 Validazione del modello

Per verificare che il modello generato dall'algoritmo di Topic Modeling rappresenti in modo corretto il contenuto del database analizzato, è necessario passare da una fase di validazione dei dati, come indicato in sezione 2.5.

Dopo aver terminato la fase di etichettatura delle determinanti di qualità, è stato estratto in ordine casuale un set composto da cento recensioni appartenenti al file iniziale; per ognuna di esse sono stati assegnati, manualmente, zero, uno o più topic tra quelli precedentemente etichettati.

Le stesse cento recensioni sono state analizzate anche a livello automatico, sfruttando la distribuzione multinomiale di probabilità (topical prevalence) che l'algoritmo ha assegnato per ogni record. Per identificare i topic più rappresentativi è stato utilizzato l'approccio della soglia dinamica descritta nel paragrafo 2.6, ottenendo i seguenti risultati in Tabella 6:

		Assegnazione manuale dei topic (condizione vera)			
		Esistenza del topic	Non esistenza del topic	Accuracy 0,95	
Assegnazione automatica dei topic	Esistenza del topic	Veri positivi (tp) 102	Falsi positivi (fp) 18	Precision 0,85	False discovery rate 0,15
	Non esistenza del topic	Falsi negativi (fn) 41	Veri negativi (tn) 1039	False omission rate 0,04	Negative predictive value 0,96
		Recall 0,71	Fall-out 0,02	F1 Score 0,78	
		Miss-rate 0,29	Specificity 0,98		

Tabella 6: Calcolo indicatori di validazione. In verde sono evidenziate le metriche che rientrano nei valori soglia; viceversa, in rosso quelle che non ci rientrano.

Dal calcolo degli indicatori di validazione, presi dalla Figura 11 e presenti in Tabella 6, confrontati con i valori soglia in Figura 12, è possibile osservare che essi rientrano all'interno del range corretto a eccezione degli indicatori Miss-rate e False discovery rate, rispettivamente a causa di un alto numero di falsi negativi e un alto numero di falsi positivi. Queste anomalie possono essere dovute alla natura stessa dei dati testuali, caratterizzati da una forte variabilità linguistica e sovrapposizione tra i temi estratti, che rendono più complessa una classificazione netta degli argomenti; inoltre, l'eterogeneità delle recensioni, spesso brevi e informali, potrebbe aver accentuato la difficoltà del modello nel distinguere correttamente alcuni casi limite. Nonostante tali valori, il set di dati è stato ritenuto idoneo a proseguire con le analisi, poiché gli altri indicatori rientrano ampiamente nelle soglie di validazione e l'impatto dei falsi positivi e falsi negativi non compromette in maniera significativa la rappresentatività complessiva dei risultati. L'obiettivo principale della ricerca non è infatti la perfetta corrispondenza di ogni singola osservazione, ma l'individuazione delle tendenze generali e dei temi più rilevanti nella Voice of Customer; in quest'ottica, la qualità dei dati è stata ritenuta sufficiente per garantire l'affidabilità delle fasi successive.

3.6 Analisi dei risultati

3.6.1 MRP

Per l'analisi dei risultati e la successiva applicazione delle carte di controllo si è scelto di proseguire considerando solo il player col maggior numero di recensioni rimanenti nel dataset, ovvero Glovo con 4.029 recensioni. Le suddette analisi iniziano con il calcolo dell'indice MRP (Mean Rating Proportion), che indica quanto è discusso un topic all'interno delle recensioni, considerando uno specifico livello di rating, e rappresentato dalla seguente formula (Barravecchia et al., 2023).

$$MRP_{t,k} = \frac{\sum_{i \in R_k} TP_{i,t}}{|R_k|} \quad (3.1)$$

dove:

- **t** = topic;
- **k** = livello di rating;
- **R_k** = sottoinsieme di recensioni associate al livello di rating k;
- **$TP_{i,t}$** = topical prevalence associata al topic t nella recensione i;
- **$|R_k|$** = cardinalità R_k .

Sono state raggruppate, all'interno dell'intero database, le recensioni appartenenti a uno stesso valore di valutazione, ottenendo 2.585 record per il rating 1, 395 per il rating 2, 272 per il rating 3, 323 per il rating 4 e 454 per il rating 5.

Da questi numeri, utilizzando la formula (3.1), per ogni livello di rating e per ogni topic trattato è stato calcolato il valore corrispondente, illustrato in Tabella 7 e nei grafici presenti in Figura 15.

	MRP1	MRP2	MRP3	MRP4	MRP5
T1: Affidabilità dell'app	0,145	0,189	0,179	0,152	0,121
T2: Accesso all'account	0,132	0,107	0,112	0,084	0,050
T3: Disponibilità dei prodotti	0,056	0,060	0,068	0,067	0,060
T4: Costi del servizio	0,086	0,080	0,072	0,066	0,070
T5: Qualità del servizio	0,075	0,066	0,068	0,080	0,111
T6: Gestione del rimborso	0,071	0,048	0,043	0,046	0,047
T7: Affidabilità ristorante	0,131	0,108	0,096	0,093	0,098
T8: Personalizzazione dell'ordine	0,047	0,063	0,070	0,100	0,080
T9: Puntualità della consegna	0,071	0,084	0,095	0,110	0,171
T10: Affidabilità del servizio clienti	0,078	0,067	0,057	0,055	0,062
T11: Consegna all'indirizzo voluto e localizzazione	0,061	0,083	0,097	0,101	0,084
T12: Completezza dell'ordine	0,047	0,044	0,043	0,046	0,047

Tabella 7: Calcolo del valore MRP per la company Glovo.

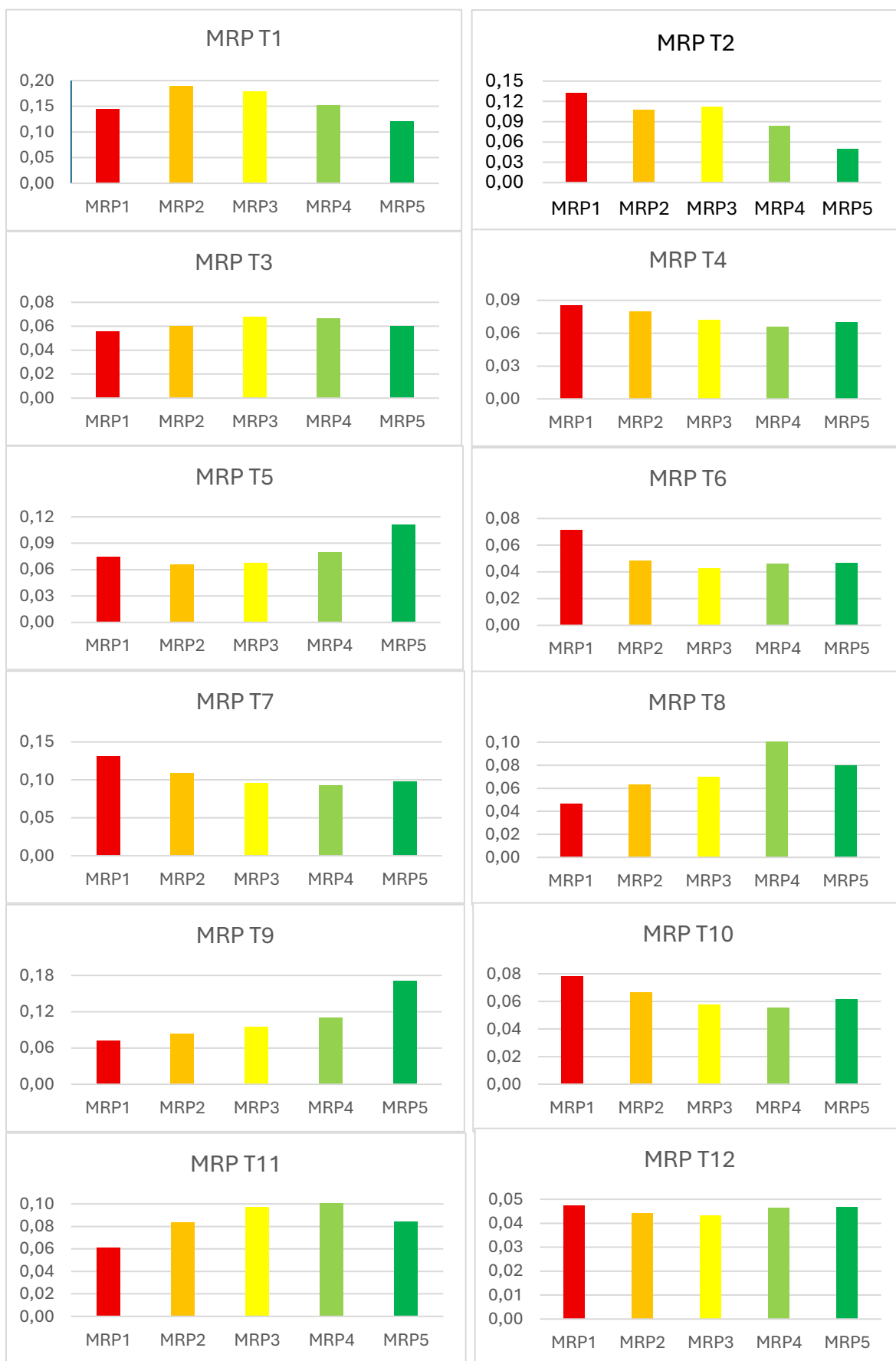


Figura 15: Rappresentazioni grafiche della metrica MRP per i dodici topic analizzati.

Osservando i risultati ottenuti, è possibile affermare che la distribuzione dei punteggi delle recensioni è molto variabile a seconda dell'argomento analizzato. Ad esempio, il *Topic 9 Puntualità della consegna* presenta una tendenza verso il valore MRP 5 (punteggio massimo); ciò significa che tendenzialmente i clienti sono soddisfatti riguardo i tempi di consegna. Al contrario, il *Topic 6 Gestione del rimborso* presenta una tendenza contraria, verso il punteggio minimo, lasciando intendere che il cliente medio non è soddisfatto della procedura di rimborso per i pagamenti effettuati.

3.6.2 MTP

L'analisi è proseguita con il calcolo dei valori MTP (Mean Topic Proportion), i quali indicano in media quanto un topic è discusso all'interno del dataset (Barravecchia et al., 2023). Le relative formule sono indicate di seguito:

$$MTP_t = \frac{\sum_{i=1}^N TP_{i,t}}{N} \quad (3.2)$$

$$\sum_{t=1}^T MTP_t = 1 \quad (3.3)$$

dove N è il numero dei documenti analizzati e $TP_{i,t}$ è la topical prevalence associata al topic t nella recensione i . Sono state effettuate, utilizzando le formule (3.2) e (3.3), le sommatorie della TP per ogni determinante di qualità, che divise per il numero totale di recensioni (4.029) permettono di ottenere i risultati illustrati in Tabella 8:

	MTP
T1: Affidabilità dell'app	0,149
T2: Accesso all'account	0,115
T3: Disponibilità dei prodotti	0,058
T4: Costi del servizio	0,081
T5: Qualità del servizio	0,078
T6: Gestione del rimborso	0,062
T7: Affidabilità ristorante	0,120
T8: Personalizzazione dell'ordine	0,058
T9: Puntualità della consegna	0,088
T10: Affidabilità del servizio clienti	0,072
T11: Consegna all'indirizzo voluto e localizzazione	0,072
T12: Completezza dell'ordine	0,047
SOMMA MTP	1,000

***Tabella 8:** Calcolo del valore MTP per la company Glovo. La somma di tutti i valori, come indicato anche nella formula (3.3), equivale a 1.*

3.6.3 IMTP

È a questo punto necessario introdurre anche il concetto di intervallo di tempo, in questo caso considerato mensile, dal settembre del 2018 al maggio del 2023. La scelta dell'intervallo mensile risponde all'esigenza di adottare una granularità sufficientemente dettagliata da cogliere le principali variazioni nella percezione dei clienti, evitando allo stesso tempo l'eccessiva frammentazione che deriverebbe da intervalli più brevi (ad esempio settimanali) e la perdita di informazioni che

caratterizzerebbe aggregazioni troppo ampie, come quelle annuali. Per analizzare quindi come il livello di discussione di un topic varia nel tempo, bisogna ottenere il valore IMTP (Interval Mean Topic Prevalence), tramite la formula (3.4):

$$IMTP_{t,n} = \frac{\sum_j^{R_n} TP_{j,t}}{|R_n|} \quad (3.4)$$

dove:

- **t** = topic;
- **n** = periodo considerato;
- **R_n** = sottoinsieme delle recensioni associate al periodo n;
- **TP_{j,t}** = topical prevalence associata al topic t nella recensione j;
- **|R_n|** = cardinalità R_n.

Eseguiti i calcoli relativi al numero di recensioni e alla somma dei TP per ogni topic, considerando lo specifico intervallo di tempo, è possibile eseguire la formula (3.4) sopra descritta, ottenendo delle visualizzazioni grafiche che indicano l'andamento dell'entità di discussione della determinante di qualità nel tempo.

Plottando i valori all'interno di grafici con la data sulle ascisse e la Topical Prevalence sulle ordinate, si ricavano i risultati presenti in Appendice III per ognuno dei 12 topic analizzati.

A titolo di esempio, vengono di seguito riportati in Figura 16 e 17 i grafici relativi al *Topic 4 Costi del servizio* e *Topic 7 Affidabilità del ristorante*, i quali presentano

rispettivamente una tendenza positiva e una negativa, come si può notare dalla linea tratteggiata e dall'espressione della retta indicata in rosso.

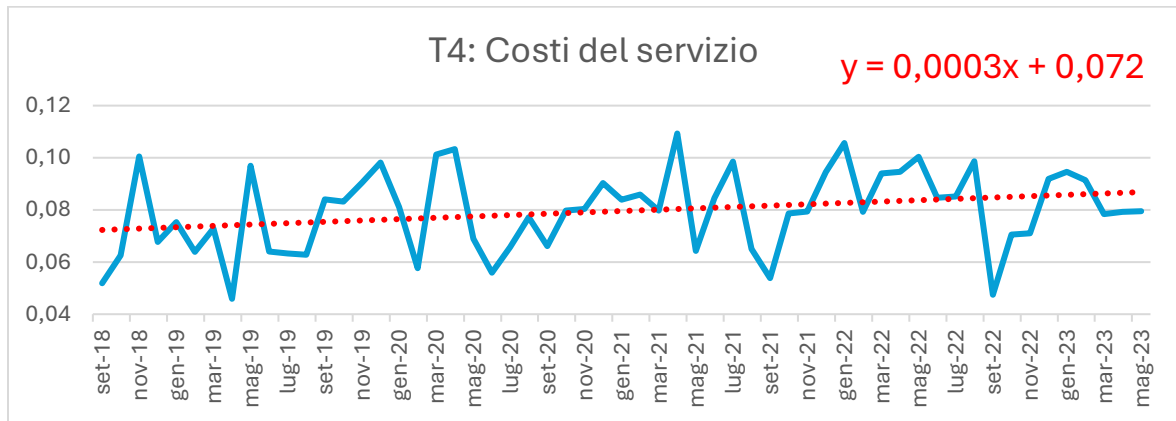


Figura 16: Grafico IMTP per il Topic T4 Costi del servizio, dove è possibile osservare una tendenza positiva.

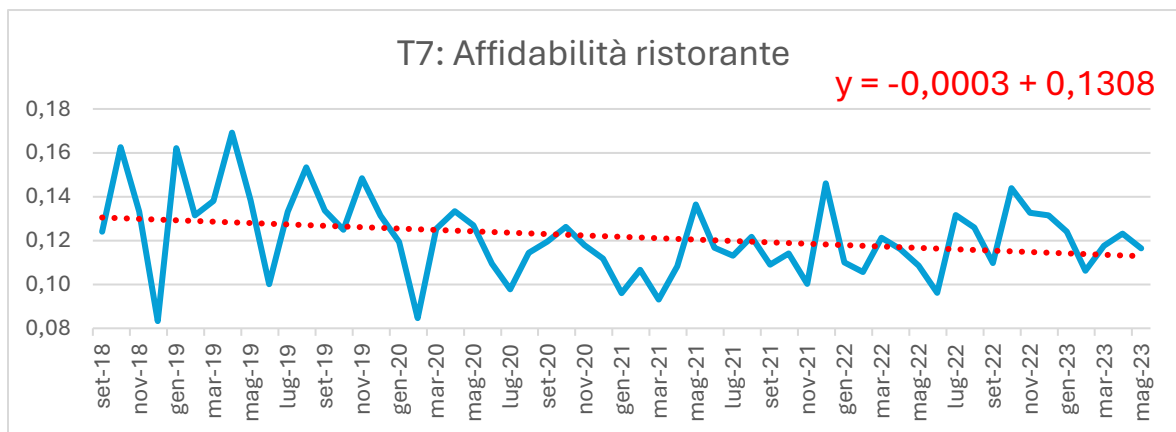


Figura 17: Grafico IMTP per il Topic T7 Affidabilità ristorante, dove è possibile osservare una tendenza negativa.

3.7 Applicazione delle Carte di Controllo

A questo punto del lavoro, è possibile procedere con l'applicazione della carta di controllo. Per il presente caso di studio è stata scelta la carta $\bar{X} - S$, ma prima di iniziare con il calcolo dei vari limiti, è necessario fare alcune riflessioni, descritte nel successivo paragrafo, sui risultati ricavati dai grafici dell'indice IMTP.

3.7.1 Detrending e calcolo dei residui

Osservando i grafici ottenuti nelle Figure 16-17 e in generale riportati in Appendice III è possibile osservare tramite la retta rossa tratteggiata che tutti i Topic analizzati presentano una tendenza, che sia essa positiva o negativa.

La presenza di una certa tendenza potrebbe alterare i risultati ottenuti con l'applicazione della Carta di controllo utilizzata, in quanto essa ha come caratteristica una Linea Centrale fissa, che non considera eventuali trend.

Si è proceduto quindi a adottare la cosiddetta tecnica di Detrending, che consiste come dice la parola stessa nell'eliminazione di una specifica tendenza, ottenendo come risultato dei residui che possono essere utilizzati per i successivi calcoli.

La formula utilizzata per il calcolo dei residui è la (3.5) e consiste in una sottrazione dove, per ogni intervallo di tempo t considerato, il primo valore corrisponde all'indice IMTP già calcolato, mentre il secondo si ottiene con il calcolo della retta di tendenza rappresentata dalla formula (3.6):

$$y_t - \hat{y}_t \quad (3.5)$$

$$\hat{y}_t = mx + q \quad (3.6)$$

dove m è il coefficiente angolare e q la variabile indipendente. Poiché le osservazioni sono identificate da date mensili (ad esempio set-18, ott-18, nov-18 e così via), per semplicità operativa si è scelto di sostituire le date con degli indici numerici progressivi. Alla prima osservazione, quindi, è stato assegnato il valore $x=1$, alla successiva $x=2$, fino ad arrivare all'ultima osservazione (mag-23 con $x=57$).

In Tabella 9 viene riportato, come esempio, il calcolo dei residui per il Topic 1 per i primi 10 mesi, considerando la retta di tendenza $y = 0,000175 - 0,144568$.

Mese di osservazione	y (valore IMTP iniziale)	Formula $y_t - \hat{y}_t$	Residuo
Settembre 2018	0,137535	$0,137535 - (0,000175 \times 1 - 0,144568)$	-0,007208
Ottobre 2018	0,189714	$0,189714 - (0,000175 \times 2 - 0,144568)$	0,044796
Novembre 2018	0,127116	$0,127116 - (0,000175 \times 3 - 0,144568)$	-0,017977
Dicembre 2018	0,149478	$0,149478 - (0,000175 \times 4 - 0,144568)$	0,004210
Gennaio 2019	0,123356	$0,123356 - (0,000175 \times 5 - 0,144568)$	-0,022088
Febbraio 2019	0,140078	$0,140078 - (0,000175 \times 6 - 0,144568)$	-0,005542
Marzo 2019	0,122985	$0,122985 - (0,000175 \times 7 - 0,144568)$	-0,022809
Aprile 2019	0,158514	$0,158514 - (0,000175 \times 8 - 0,144568)$	0,012544
Maggio 2019	0,138922	$0,138922 - (0,000175 \times 9 - 0,144568)$	-0,007223
Giugno 2019	0,138062	$0,138062 - (0,000175 \times 10 - 0,144568)$	-0,008259

Tabella 9: Calcolo dei residui per le prime dieci osservazioni del Topic 1 Affidabilità dell'app. Partendo dal valore iniziale y_t , è stato sottratto il valore $\hat{y}_t$ utilizzando le formule (3.5) e (3.6).

Dai calcoli ottenuti considerando l'intero intervallo di osservazioni, è possibile ricavare il grafico dei residui indicato in Figura 18.

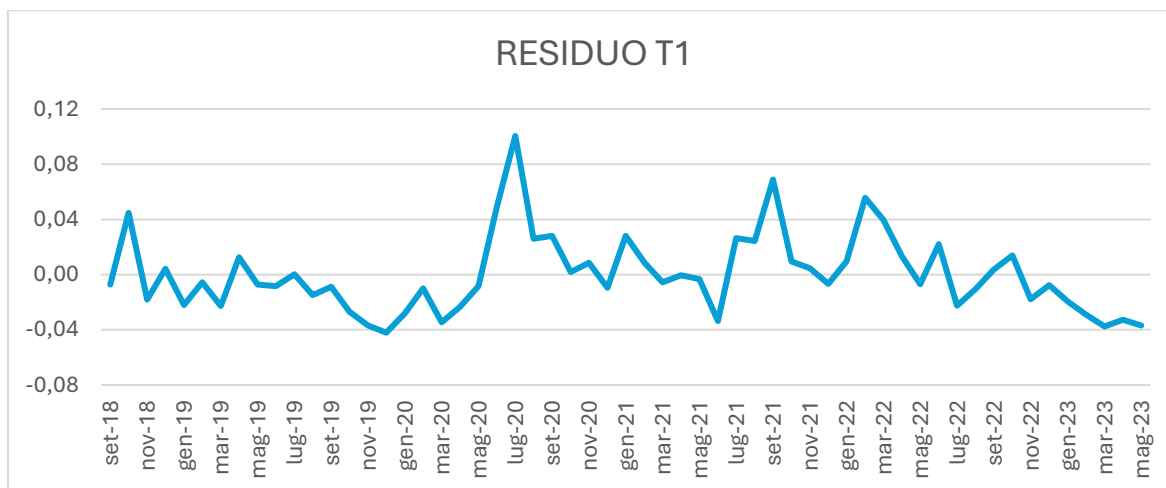


Figura 18: Grafico dei residui per il Topic 1 Affidabilità dell'app, dopo aver effettuato il detrending.

Prima di procedere con l'applicazione delle Carte di controllo, è necessario fare un'ulteriore osservazione in merito alla tecnica di detrending appena descritta. Ogni grafico ottenuto plottando i valori IMTP dei topic presenta una certa tendenza, che per semplicità nel presente caso di studio è sempre stata considerata come lineare. Ciò significa che, con altre regressioni (ad esempio polinomiale, logaritmica, esponenziale ecc.), alcuni residui avrebbero potuto ottenere un risultato più preciso in fase di eliminazione del trend.

In Tabella 10 di seguito è possibile osservare un confronto tra la retta di pendenza iniziale e quella ottenuta dopo il detrending, mentre i grafici sono consultabili in Appendice IV:

	Retta iniziale	Retta post detrending
T1: Affidabilità dell'app	$y = 2E-04x + 0,1446$	$y = 2E-19x - 6E-17$
T2: Accesso all'account	$y = 1E-04x + 0,1134$	$y = -1E-19x + 1E-17$
T3: Disponibilità dei prodotti	$y = 2E-04x + 0,0534$	$y = -1E-19x - 5E-19$
T4: Costi del servizio	$y = 3E-04x + 0,0720$	$y = 5E-20x + 5E-18$
T5: Qualità del servizio	$y = -2E-04x + 0,0845$	$y = 7E-20x - 2E-17$

T6: Gestione del rimborso	$y = 1E-04x + 0,0575$	$y = -1E-19x - 1E-17$
T7: Affidabilità ristorante	$y = -3E-04x + 0,1308$	$y = 6E-20x - 3E-17$
T8: Personalizzazione dell'ordine	$y = -1E-04x + 0,0618$	$y = -1E-20x - 9E-18$
T9: Puntualità della consegna	$y = 2E-05x + 0,0846$	$y = 6E-20x - 1E-17$
T10: Affidabilità del servizio clienti	$y = -3E-04x + 0,0814$	$y = -9E-21x - 1E-18$
T11: Consegna all'indirizzo voluto e localizzazione	$y = -1E-05x + 0,0719$	$y = 2E-20x - 1E-18$
T12: Completezza dell'ordine	$y = 1E-04x + 0,0442$	$y = -3E-20x + 5E-18$

Tabella 10: Confronto tra l'equazione della retta iniziale e quella della retta ottenuta dopo il detrending.

A titolo di esempio, per comprendere al meglio la Tabella 10, viene proposto in Figura 19 un confronto tra il grafico iniziale e il residuo calcolato per il *Topic 4 Costi del servizio*; osservando l'immagine, infatti, è possibile notare come si siano registrati dei cambiamenti sia sui valori della retta, sia su quelli delle ordinate:

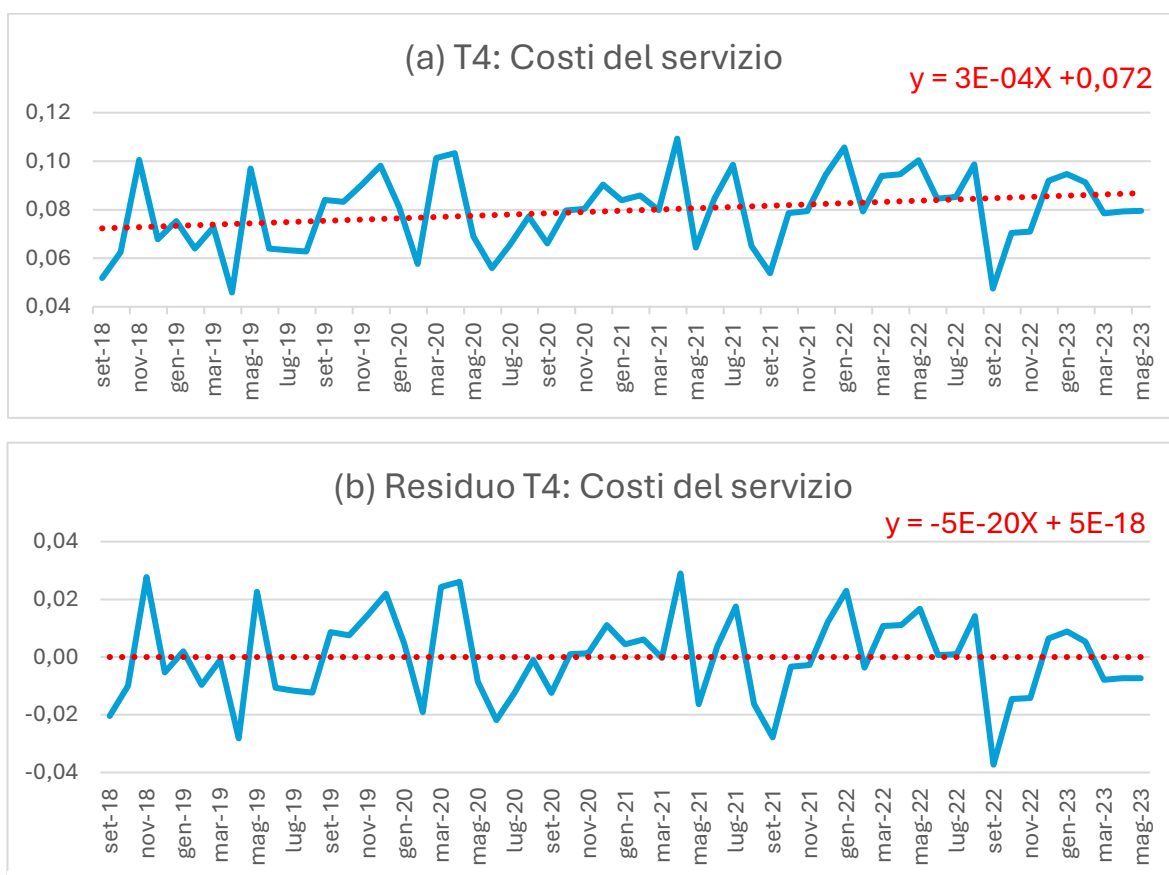


Figura 19: Confronto tra il grafico iniziale (a) e il residuo (b) dopo il detrending, per il Topic 4 Costi del servizio.

In particolare, come si può notare sempre in Figura 19b, nel grafico dei residui la retta in rosso presenta un coefficiente angolare prossimo allo zero. Tuttavia, è doveroso considerare che tutti i coefficienti angolari presenti in Tabella 10 sono già in partenza molto bassi, quindi in fase di detrending non è possibile aspettarsi un cambiamento significativo del processo. Dal momento che, in ogni caso, non è possibile ottenere una linea di tendenza ottimale (coefficiente nullo) su cui strutturare i calcoli, i residui ricavati hanno mantenuto un certo trend non nullo, seppur minore rispetto al caso iniziale.

Nei paragrafi 3.7.2 e 3.7.3, dedicati al calcolo dei limiti per la carta $\bar{X} - S$ ai processi detrendizzati e all'analisi critica della metodologia, verranno effettuate delle valutazioni sull'effettiva efficacia di questo strumento nei casi in cui siano presenti trend non troppo accentuati, come in questo caso.

3.7.2 Applicazione della carta $\bar{X}-S$

A partire dai valori dei residui ottenuti nei dodici topic analizzati, si è iniziato a calcolare la grande media del processo $\bar{\bar{x}}$. I valori IMTP (e quindi anche dei residui), calcolati per ogni osservazione temporale ovvero mensilmente, possono essere considerati come una media degli IMTP di tutti i record appartenenti alla stessa mensilità di riferimento. Ad esempio, nel settembre 2018 sono state rilevate nove recensioni; pertanto, il valore IMTP di 0,137535 (e del residuo di -0,007208) è stato ottenuto come una media dei valori per i nove record appena citati.

Per il calcolo della grande media, pertanto, è stato sufficiente per ogni topic calcolare la media dei residui, considerando tutte le 57 mensilità analizzate. Avendo

effettuato a livello preliminare il detrending dei processi, la grande media per tutti i topic è risultata approssimabile allo zero.

In merito alla deviazione standard s_i è necessario fare un passo indietro e tornare al database di partenza ottenuto al paragrafo (3.6): per ottenere il valore per ogni mese della deviazione è necessario avere i dati della Topical Prevalence non ancora raggruppati per data. A questo punto, impostando una tabella pivot su Excel con calcolo automatico della deviazione standard, è stato possibile ricavare il valore mensile s_i , e di conseguenza anche il valore dello scarto tipo medio $\bar{s}$, necessario per ottenere i limiti e le linee centrali delle carte.

In Tabella 11 di seguito, per ogni topic vengono riportati i valori $\bar{x}$ e $\bar{s}$:

	$\bar{x}$	$\bar{s}$
T1: Affidabilità dell'app	≈ 0	0,121598
T2: Accesso all'account	≈ 0	0,150913
T3: Disponibilità dei prodotti	≈ 0	0,059718
T4: Costi del servizio	≈ 0	0,087306
T5: Qualità del servizio	≈ 0	0,053049
T6: Gestione del rimborso	≈ 0	0,061912
T7: Affidabilità ristorante	≈ 0	0,104859
T8: Personalizzazione dell'ordine	≈ 0	0,068004
T9: Puntualità della consegna	≈ 0	0,073229
T10: Affidabilità del servizio clienti	≈ 0	0,068447
T11: Consegna all'indirizzo voluto e localizzazione	≈ 0	0,072198
T12: Completezza dell'ordine	≈ 0	0,035487

Tabella 11: Valori di $\bar{x}$ (grande media del processo) e $\bar{s}$ (scarto tipo medio) per ogni topic.

Per il calcolo dei limiti della carta $\bar{X}$, l'ultimo parametro richiesto è A_3 , il quale è stato possibile ricavare tramite tavole consultabili online, per numerosità del campione fino a $n=25$. Per numerosità maggiori di 25, sono state applicate le formule (1.12) e (1.13). Aggiungendo ai grafici dei residui la linea centrale e i limiti, tramite le formule (1.16) e per ognuno dei 12 topic analizzati, si ottengono i grafici delle carte $\bar{X}$.

In merito alla costruzione della carta S, invece, per la linea centrale è stato considerato lo scarto tipo medio $\bar{s}$ presente in Tabella 11; per i limiti inferiori e superiori invece sono state utilizzate le formule (1.12), (1.14), (1.15) e (1.17).

Nelle pagine seguenti (Figure 20-31) verranno illustrati, per ognuno degli argomenti analizzati, i grafici relativi a:

- a) IMTP iniziale, in verde;
- b) Applicazione della carta $\bar{X}$, in celeste con i limiti in rosso;
- c) Applicazione della carta S, in giallo con i limiti in rosso.

- **Topic 1 Affidabilità dell'app:**

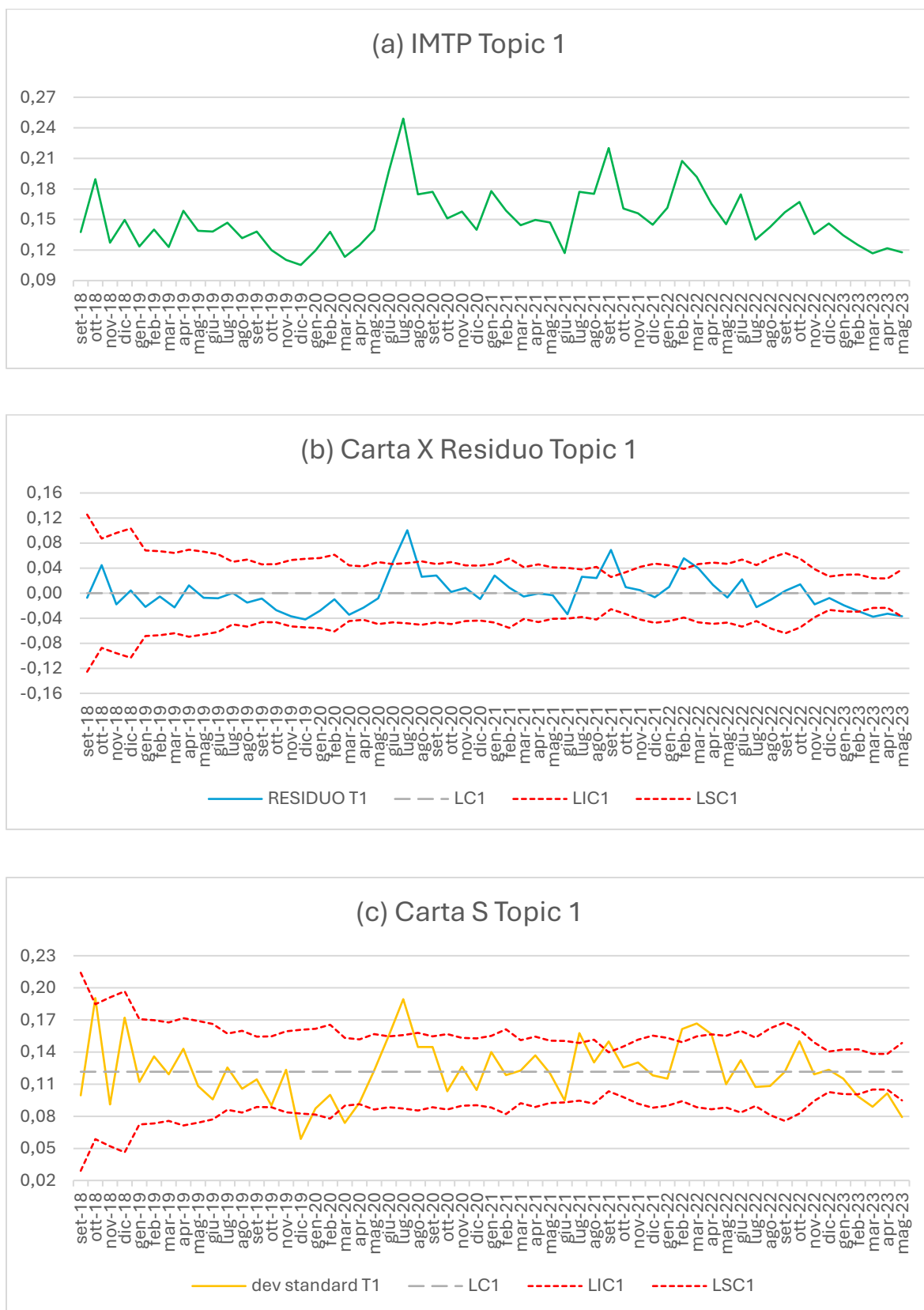


Figura 20: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 1 Affidabilità dell'app

- **Topic 2 Accesso all'account:**

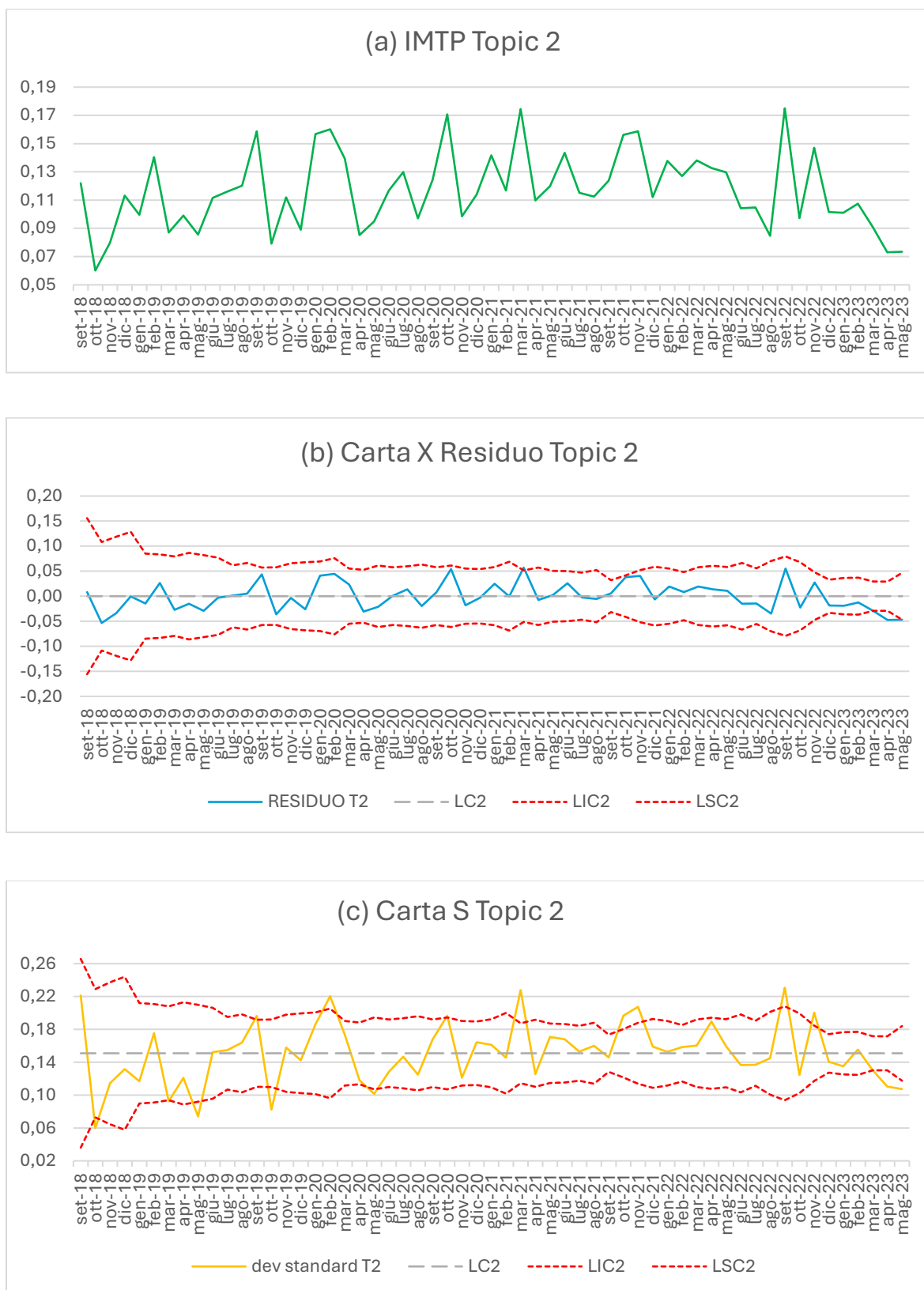


Figura 21: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 2 Accesso all'account

- **Topic 3 Disponibilità dei prodotti:**

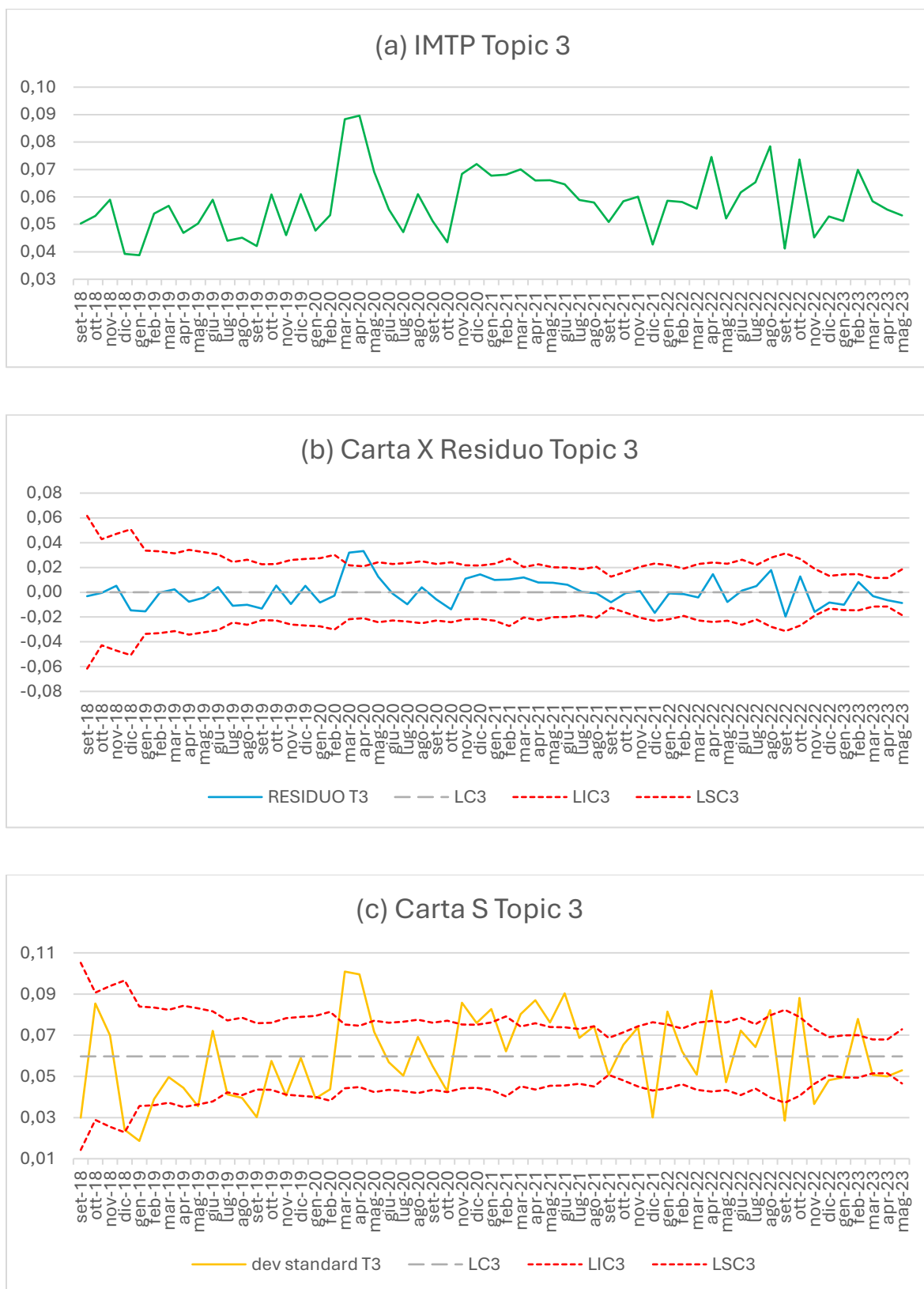


Figura 22: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 3 Disponibilità dei prodotti

- **Topic 4 Costi del servizio:**

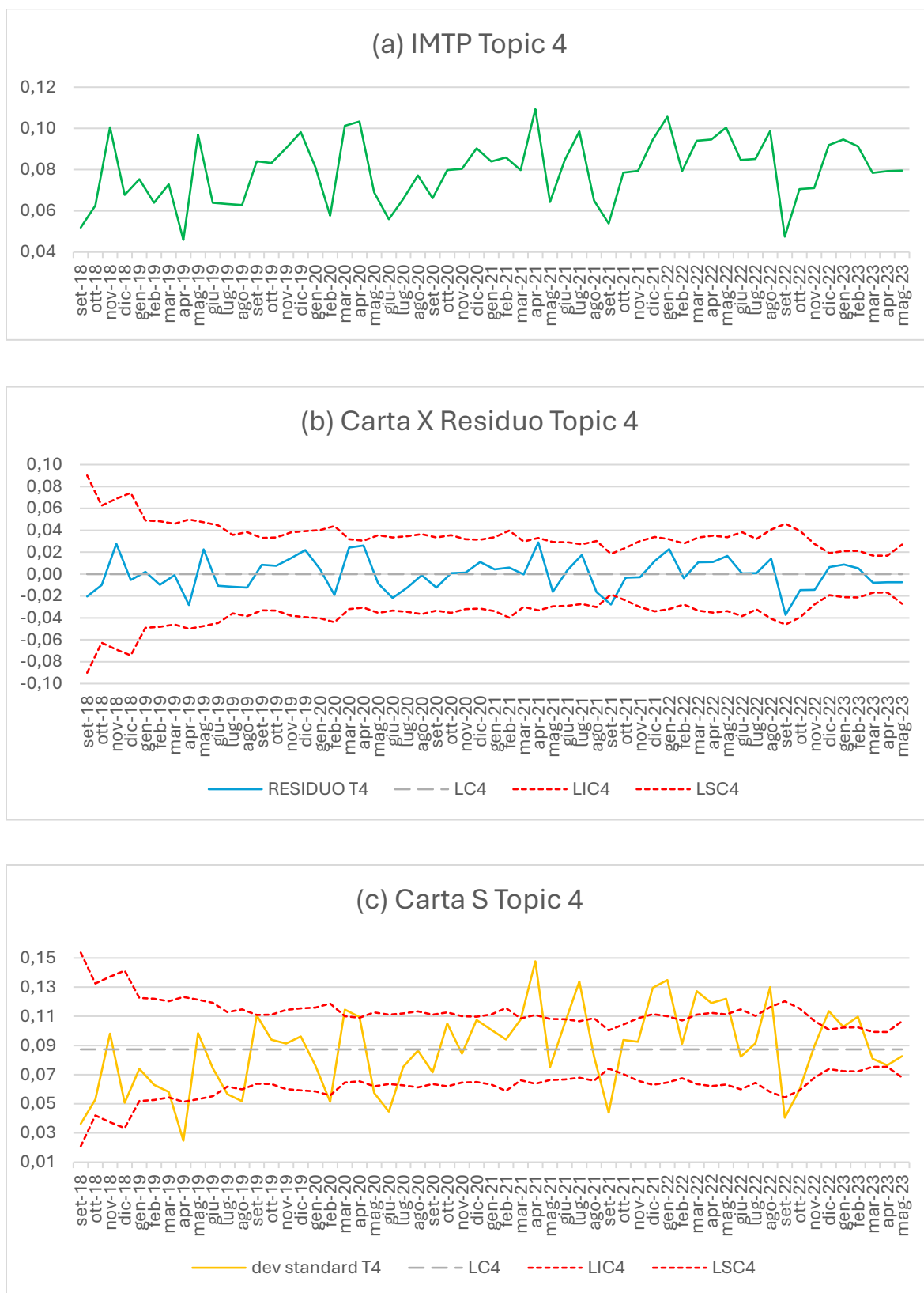


Figura 23: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 4 Costi del servizio

- **Topic 5 Qualità del servizio:**

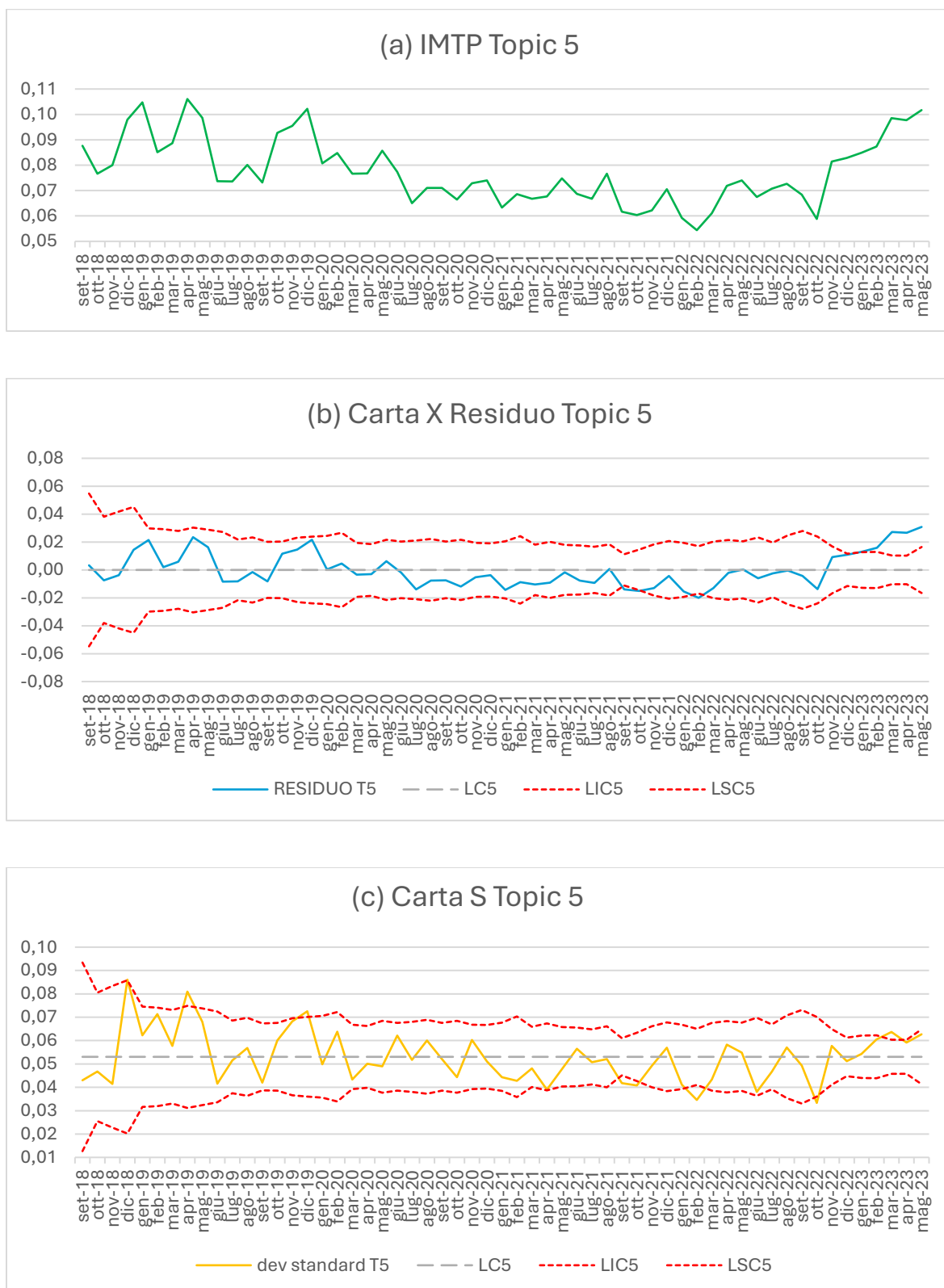


Figura 24: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 5 Qualità del servizio

- **Topic 6 Gestione del rimborso:**

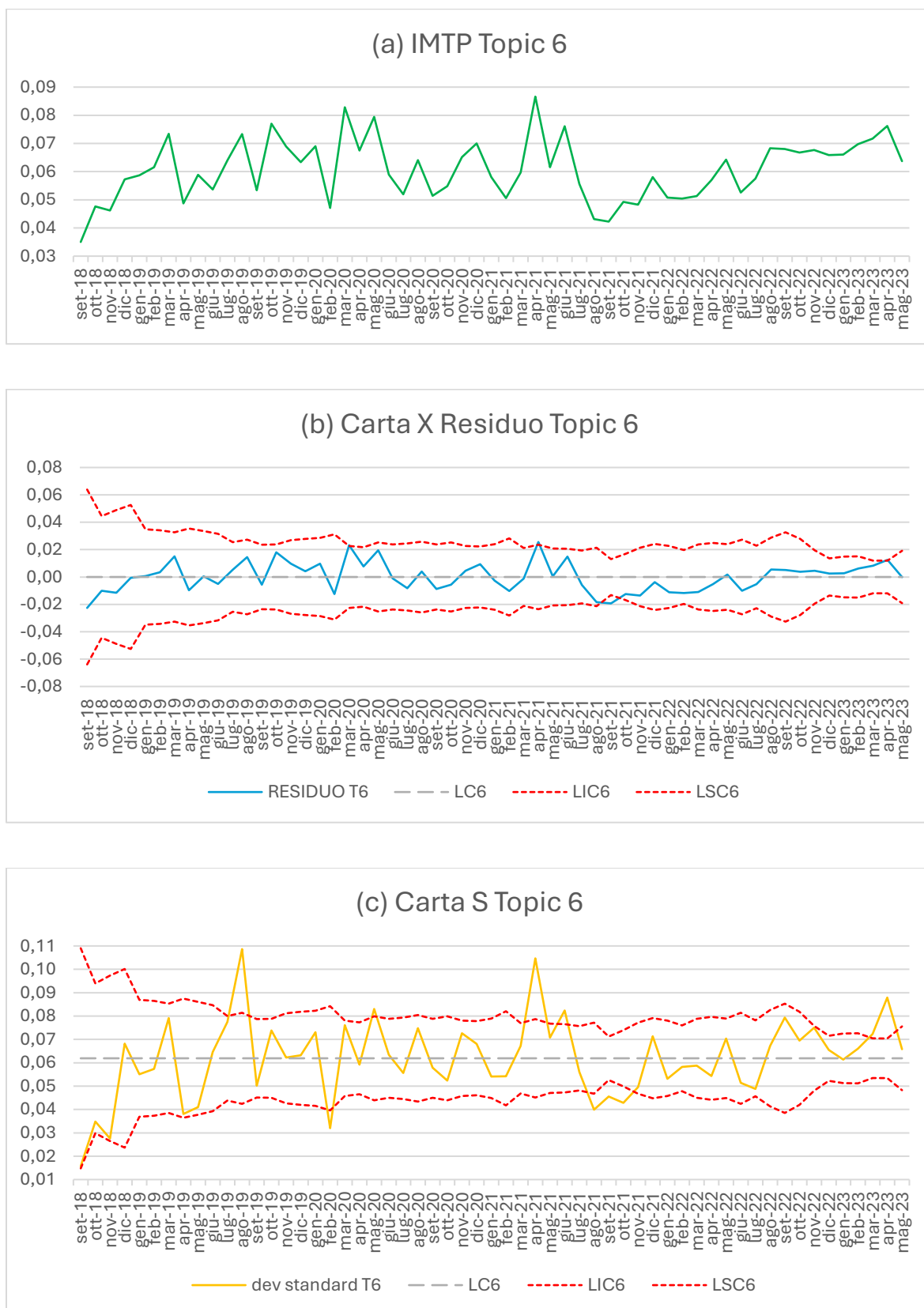


Figura 25: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 6 Gestione del rimborso

- **Topic 7 Affidabilità ristorante:**

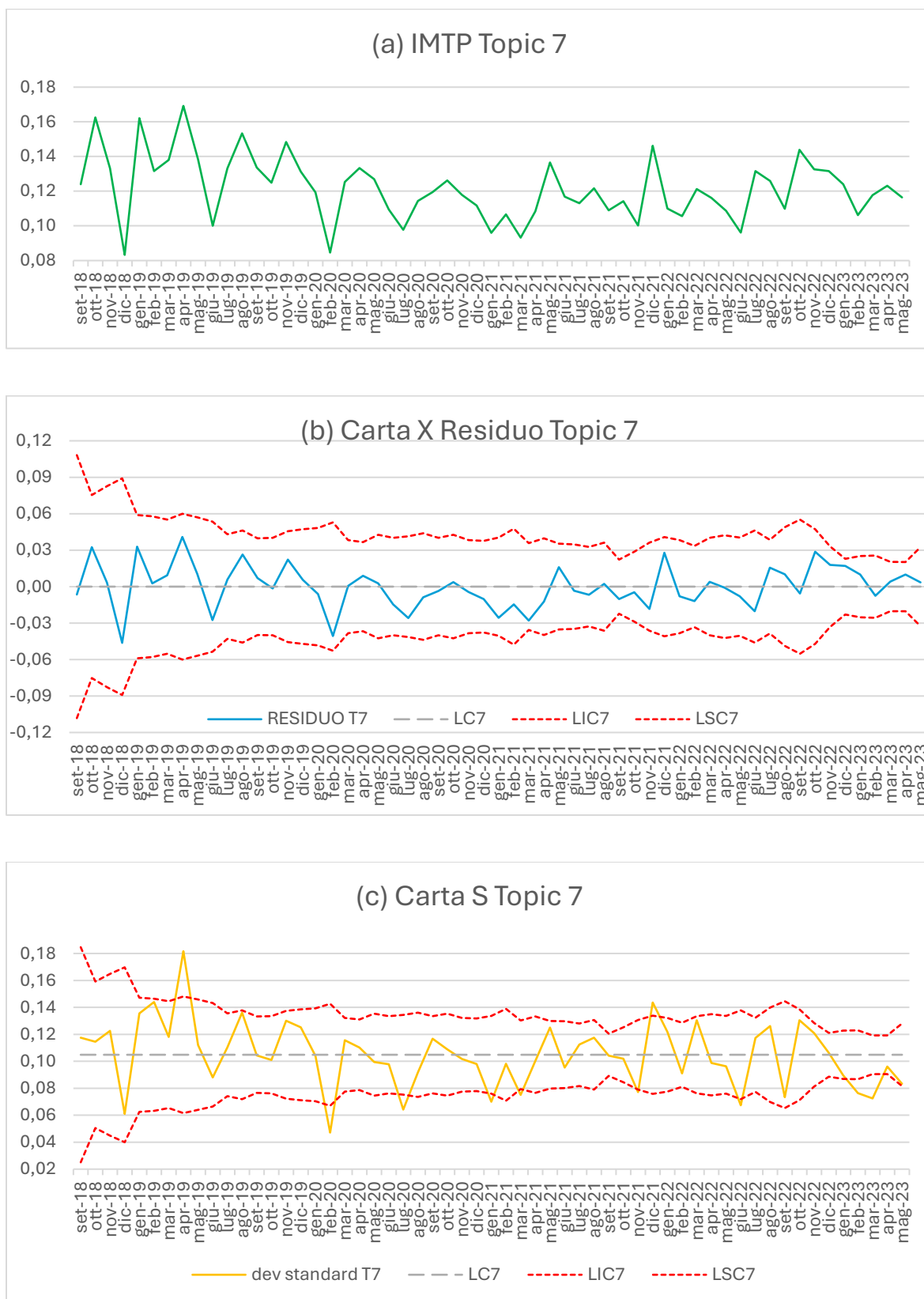


Figura 26: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 7 Affidabilità ristorante

- **Topic 8 Personalizzazione dell'ordine:**

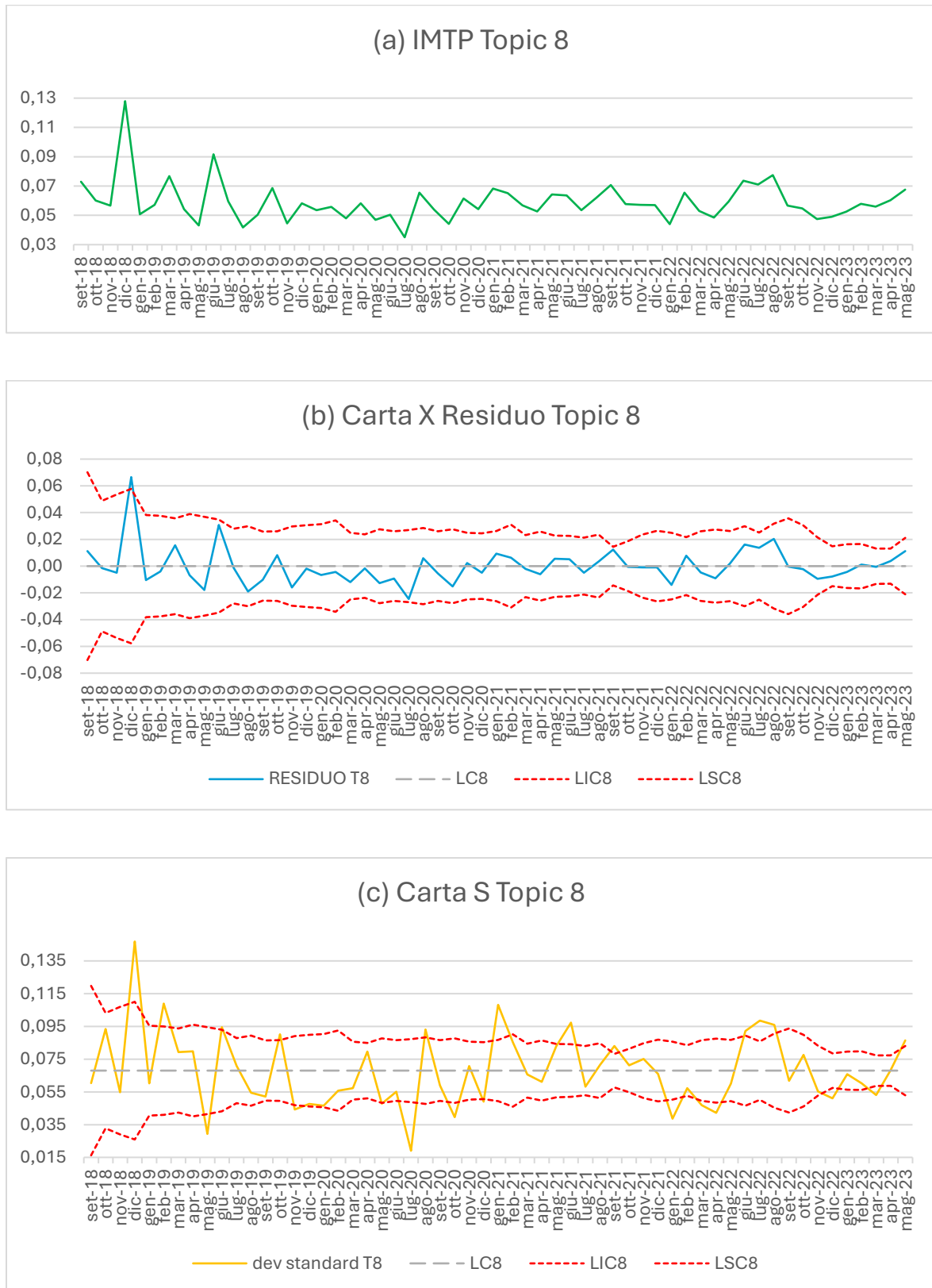


Figura 27: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 8 Personalizzazione dell'ordine

- **Topic 9 Puntualità della consegna:**

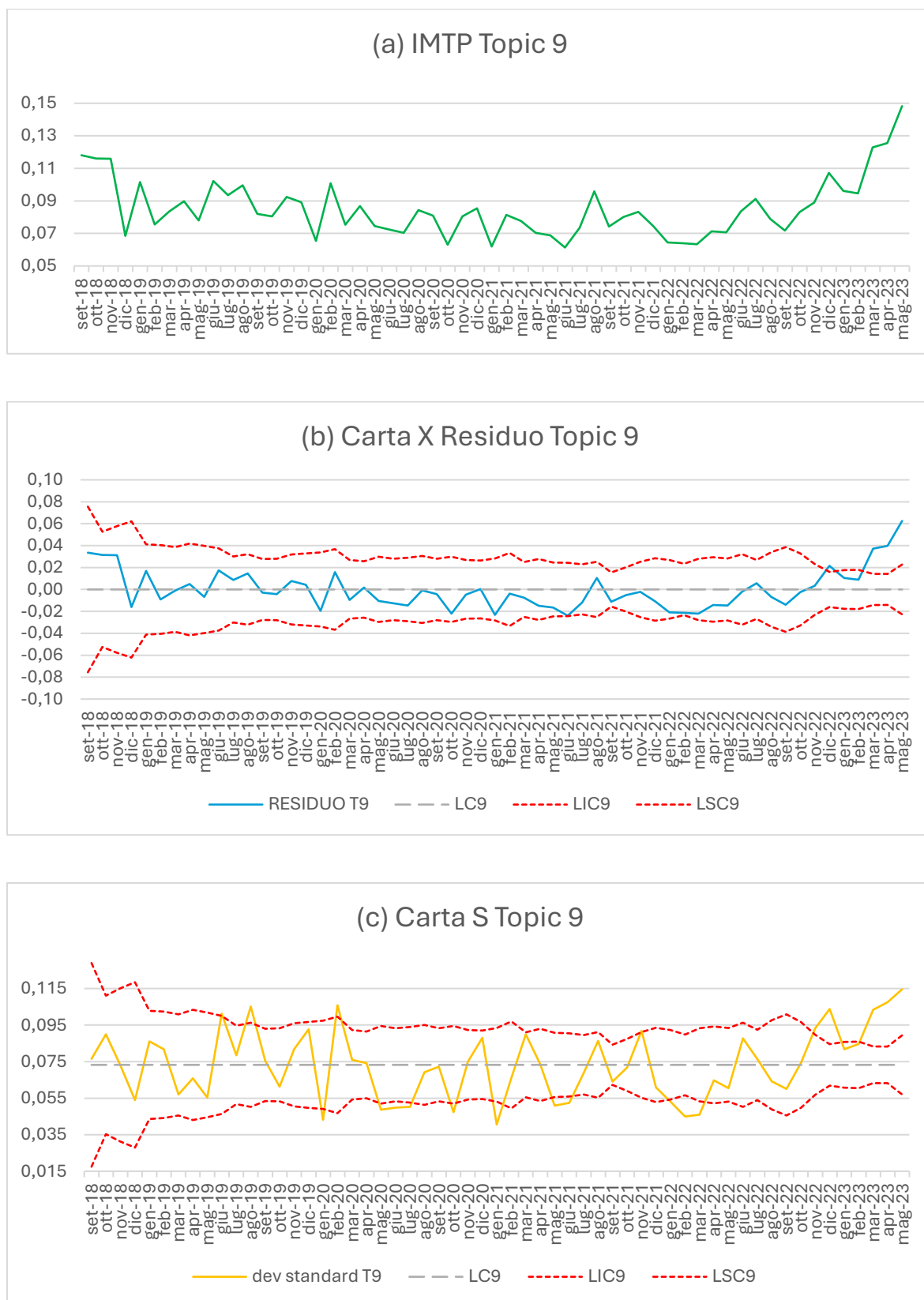


Figura 28: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 9 Puntualità della consegna

- **Topic 10 Affidabilità del servizio clienti:**

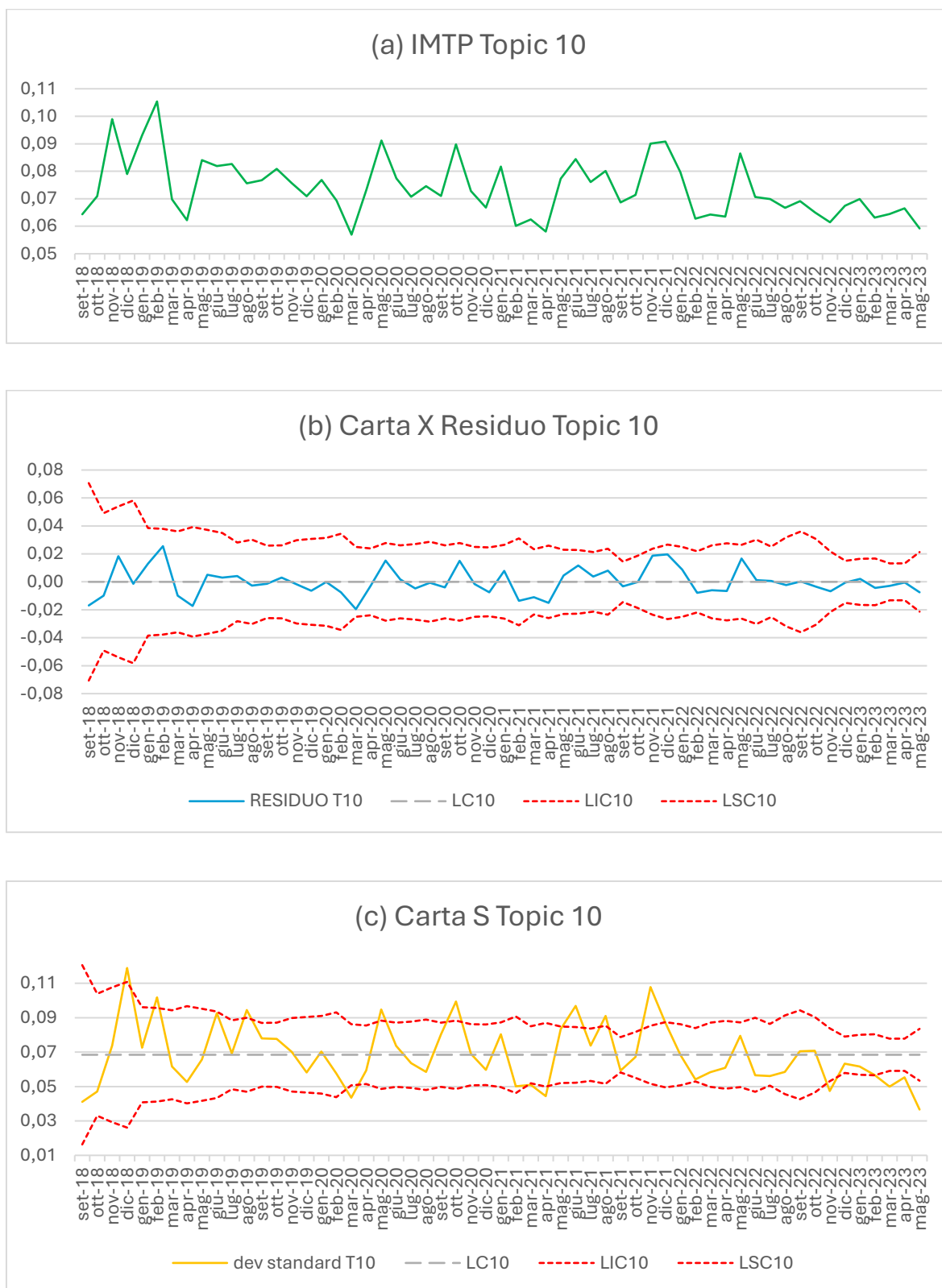


Figura 29: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 10 Affidabilità del servizio clienti

- **Topic 11 Consegna all'indirizzo voluto e localizzazione:**

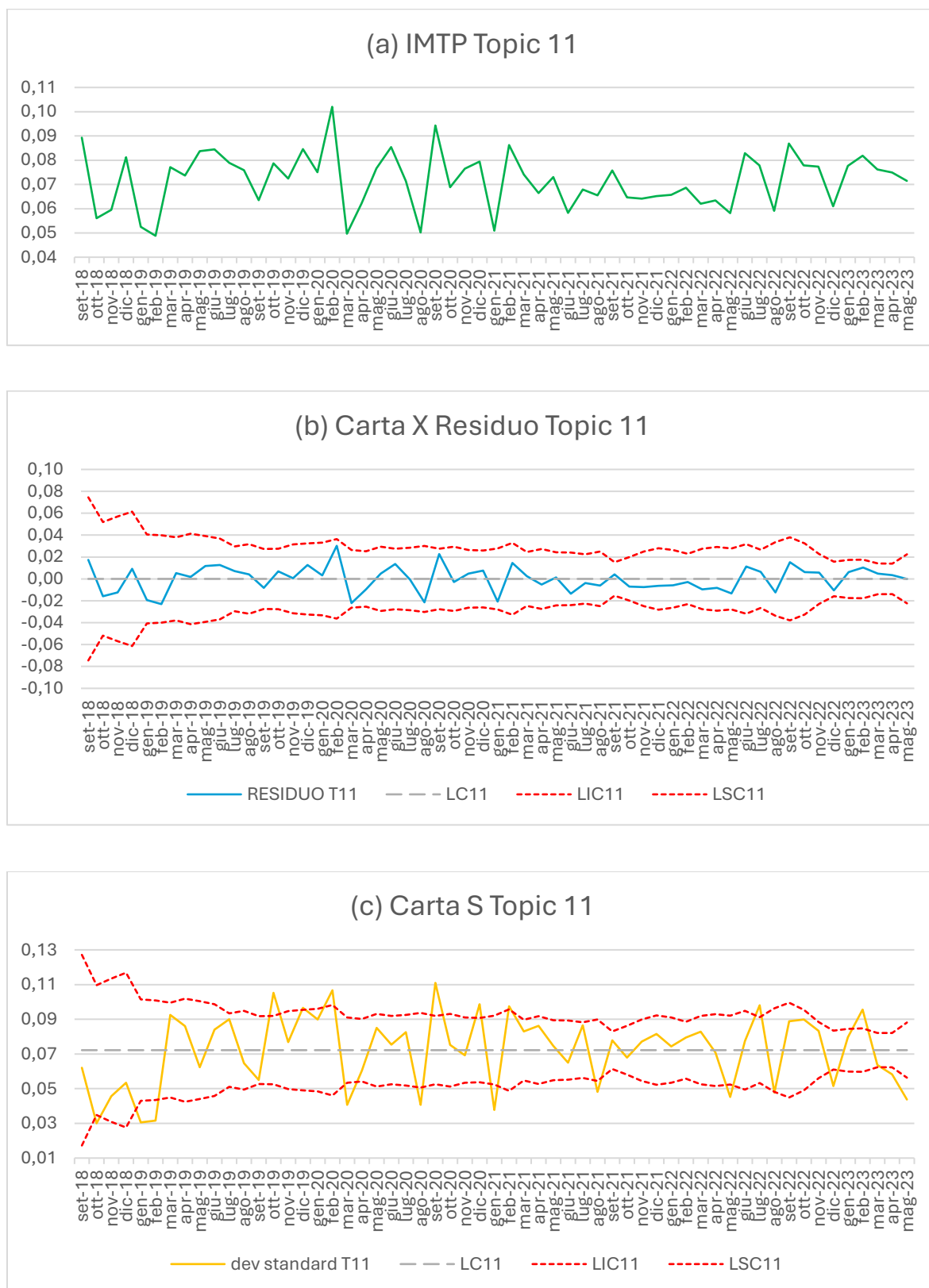


Figura 30: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 11 Consegna all'indirizzo voluto e localizzazione

- **Topic 12 Completezza dell'ordine:**

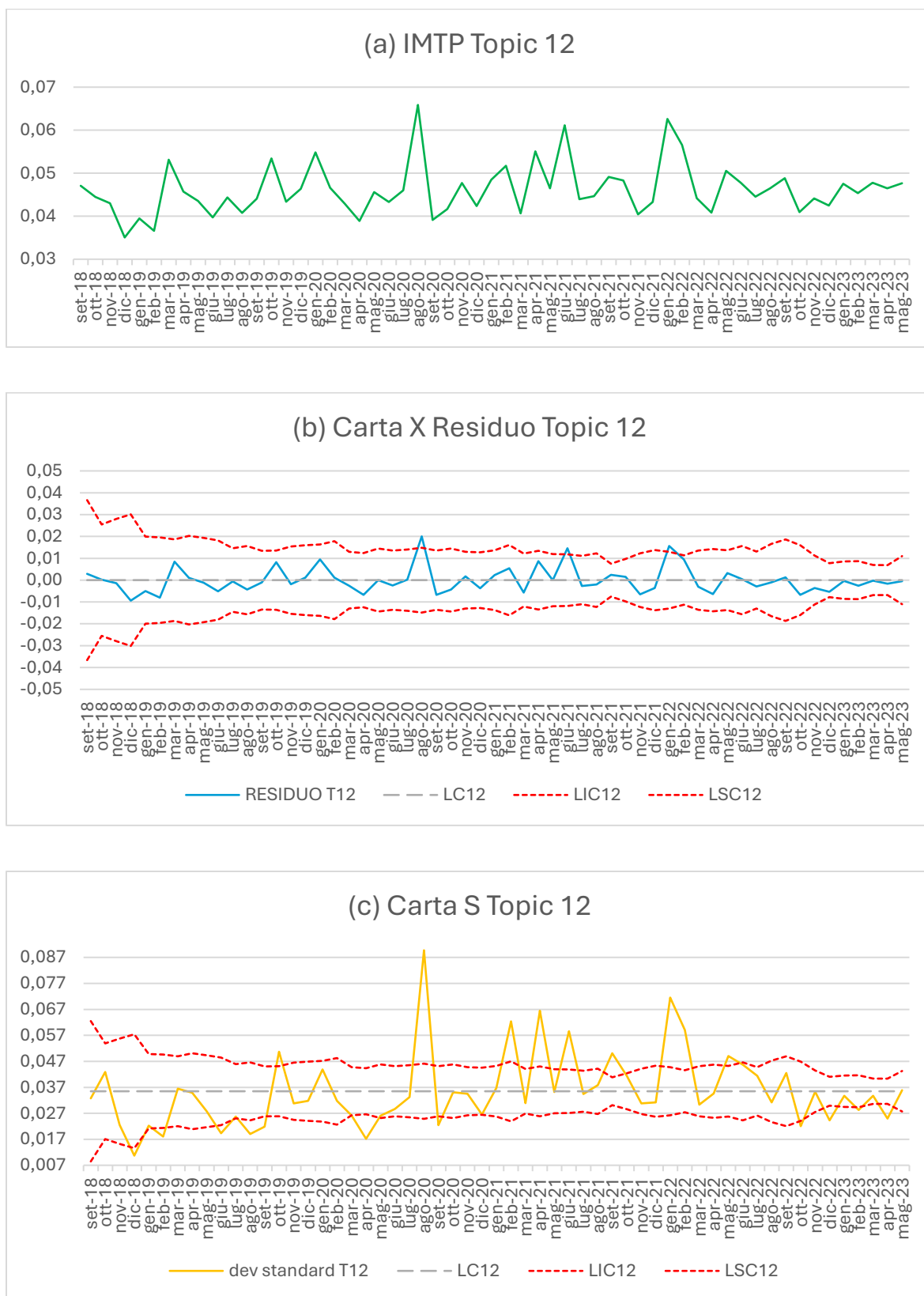


Figura 31: Grafico iniziale IMTP (a), Carta $\bar{X}$ (b) e Carta S (c) per il Topic 12 Completezza dell'ordine

Osservando i processi evidenziati dalla Figura 20 alla Figura 31, riguardanti l'applicazione della carta $\bar{X}$ ai 12 topic, è possibile notare come l'andamento della Topical Prevalence si possa considerare in controllo solamente in 3 casi, ovvero per il *Topic 7 Affidabilità del ristorante* (Fig. 26b), il *Topic 10 Affidabilità del servizio clienti* (Fig. 29b) e il *Topic 11 Consegna all'indirizzo voluto e localizzazione* (Fig. 30b); ciò significa che la prevalenza tematica dei suddetti tre argomenti è rimasta all'interno dei limiti per tutto l'andamento temporale considerato.

Il discorso opposto vale per i restanti nove argomenti, dove sono presenti diversi punti di fuori controllo all'interno del processo. Nello specifico, la Tabella 12 riporta per ogni topic il numero esatto di punti di fuori controllo.

	Numero di punti fuori controllo
T1: Affidabilità dell'app	6
T2: Accesso all'account	3
T3: Disponibilità dei prodotti	2
T4: Costi del servizio	1
T5: Qualità del servizio	8
T6: Gestione del rimborso	4
T7: Affidabilità ristorante	0
T8: Personalizzazione dell'ordine	1
T9: Puntualità della consegna	4
T10: Affidabilità del servizio clienti	0
T11: Consegna all'indirizzo voluto e localizzazione	0
T12: Completezza dell'ordine	3

Tabella 12: Numero di punti fuori controllo per ogni Topic, per la Carta $\bar{X}$.

In merito ai processi riguardanti l'applicazione della Carta S , si può constatare come la variabilità risulti molto alta; è possibile infatti notare, anche ad occhio nudo, come nessuno dei dodici argomenti risulti in controllo, con numerosi punti che vanno al di fuori dei limiti inferiori e superiori delle carte.

3.7.3 Analisi critica della metodologia

Per la valutazione complessiva della Carta $\bar{X} - S$ e della metodologia proposta, è necessario effettuare un'analisi sui seguenti fattori:

Fattori esterni: Osservando specialmente le Carte $\bar{X}$, è possibile notare come la maggior parte dei punti che vanno al di fuori dei limiti appartengano al periodo compreso tra gli anni 2020 e 2021, coincidente con la pandemia globale da Covid-19. In questo periodo, infatti, a causa delle numerose restrizioni e lockdown il mercato del food delivery potrebbe aver registrato un importante incremento del bacino di utenti registrati e dei conseguenti ordini effettuati, con maggiore attenzione ad alcuni argomenti critici come l'igiene e la modalità di consegna. Per questo motivo, un repentino incremento dell'utilizzo delle app di food delivery potrebbe essere una parziale giustificazione dei punti di fuori controllo.

Fattori di calcolo e acquisizione dei dati: come specificato nel paragrafo 3.7.1, prima di procedere con il calcolo della Carta $\bar{X}$ si è notato che, una volta ottenuti i grafici dell'andamento temporale dell'IMTP, essi presentavano delle tendenze. Per questo motivo, è stata effettuata un'attività di detrending, che ha permesso di eliminare, per quanto possibile, i trend esistenti. Avendo approssimato ogni tendenza come una linea retta, alcuni residui potrebbero aver mantenuto qualche anomalia, notevole abbastanza da permettere la conseguente presenza di alterazioni sui calcoli.

Tuttavia, come è possibile notare anche in Figura 19 e in Tabella 10 (sezione 3.7.1), le pendenze appartenenti ai grafici IMTP iniziali non si possono in realtà considerare sufficientemente accentuate da permettere, tramite detrending, un cambiamento importante sui risultati ottenuti. A sostegno di ciò, viene mostrato in Figura 32 una comparazione tra la Carta $\bar{X}$ applicata all'IMTP iniziale e quella applicata al grafico dei residui, per il *Topic 4 Costi del servizio*.

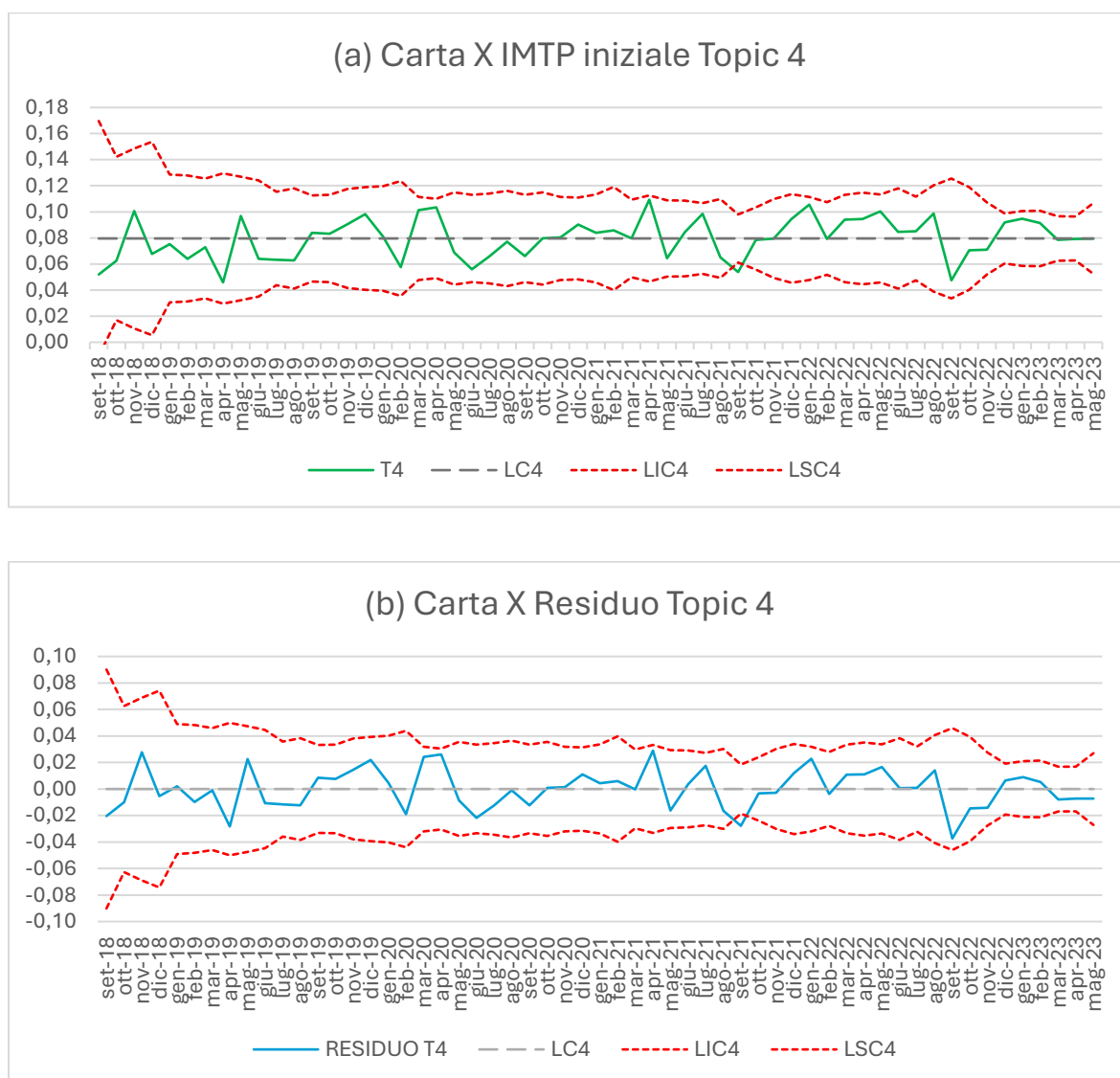


Figura 32: Confronto tra la Carta $\bar{X}$ applicata all'IMTP iniziale (a) e quella applicata al residuo (b) per il Topic 4 Costi del servizio.

Osservando la Figura 32 si può quindi notare che, a parte per i valori delle ordinate, il profilo del processo rimane perlopiù simile, con la presenza sempre di un punto di fuori controllo in corrispondenza di settembre 2021.

È doveroso inoltre sottolineare che effettuare un processo di detrending non porta necessariamente all'ottenimento di un numero minore di punti di fuori controllo. A conferma di quanto sostenuto, la Figura 33 mostra un esempio in cui dopo la fase di detrending si ottiene effettivamente una riduzione dei punti di fuori controllo, mentre la Figura 34 mostra l'effetto contrario, anche con un'approssimazione polinomiale della linea di tendenza, quindi più precisa di quella lineare.

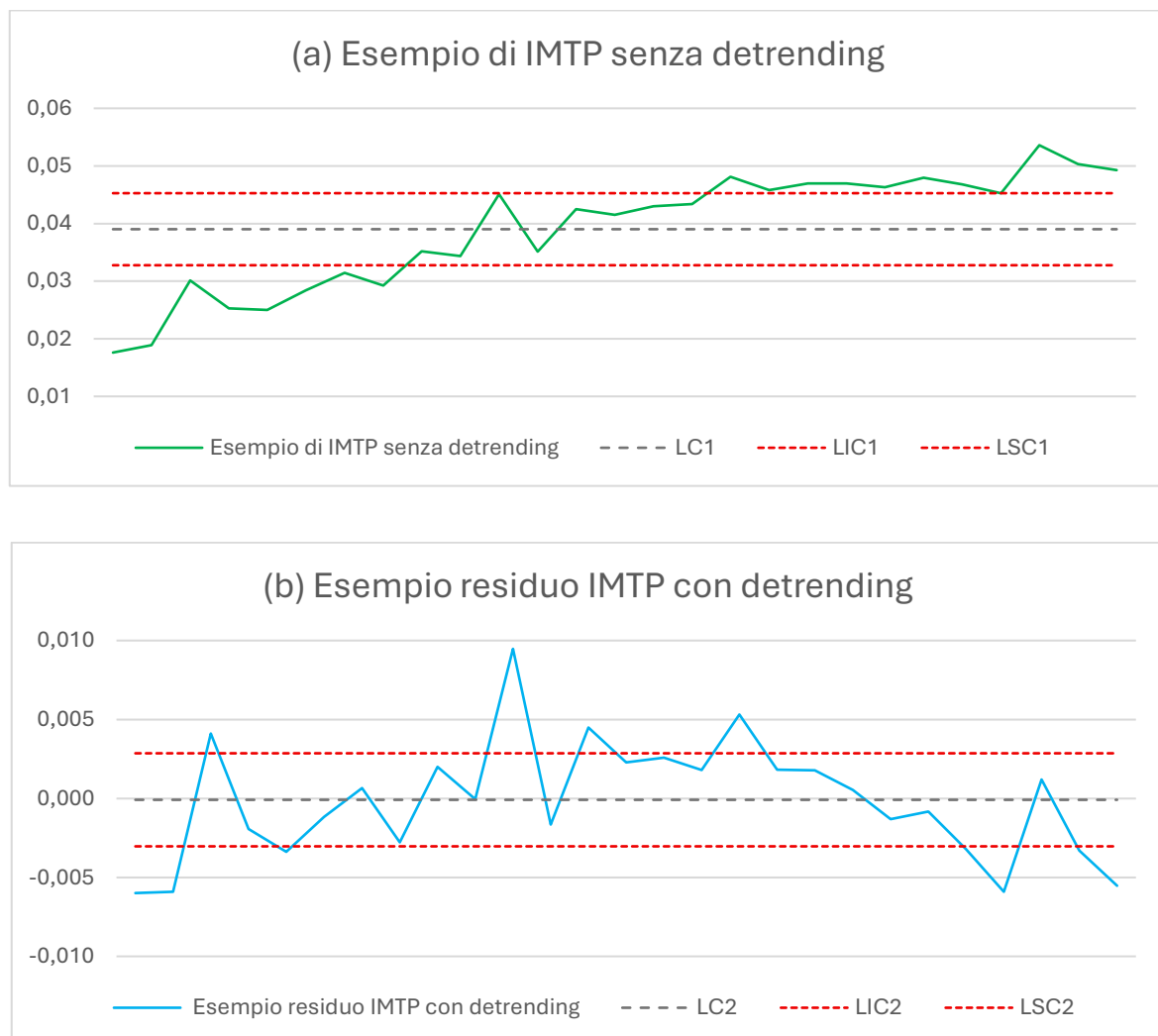


Figura 33: Calcolo della Carta $\bar{X}$ per un IMTP senza detrending (a) e uno con detrending (b). Il grafico (b) mostra una diminuzione dei punti di fuori controllo.

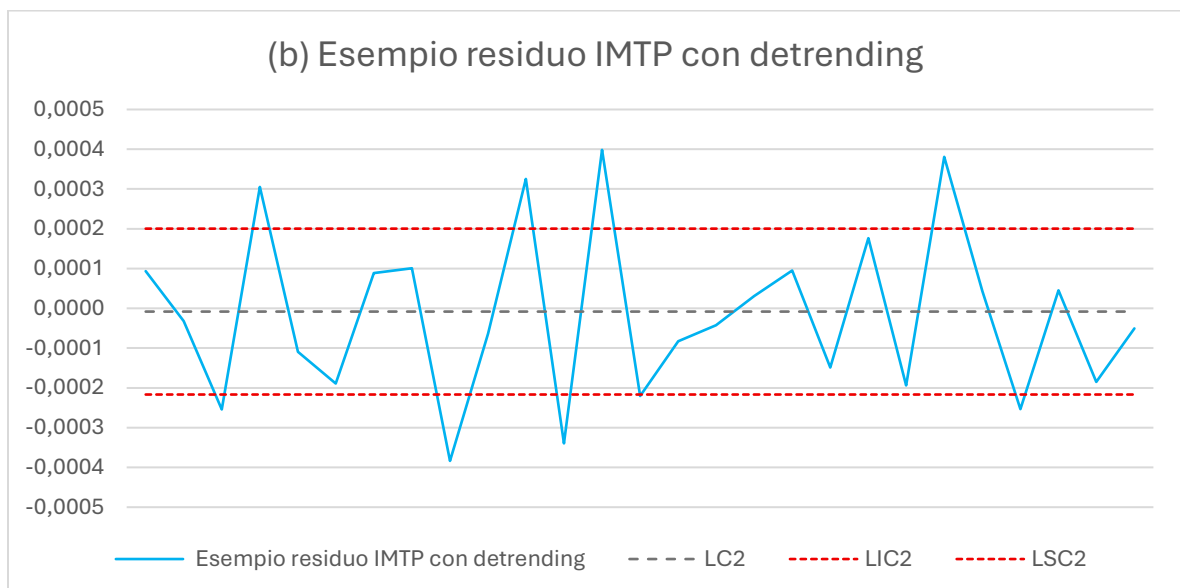
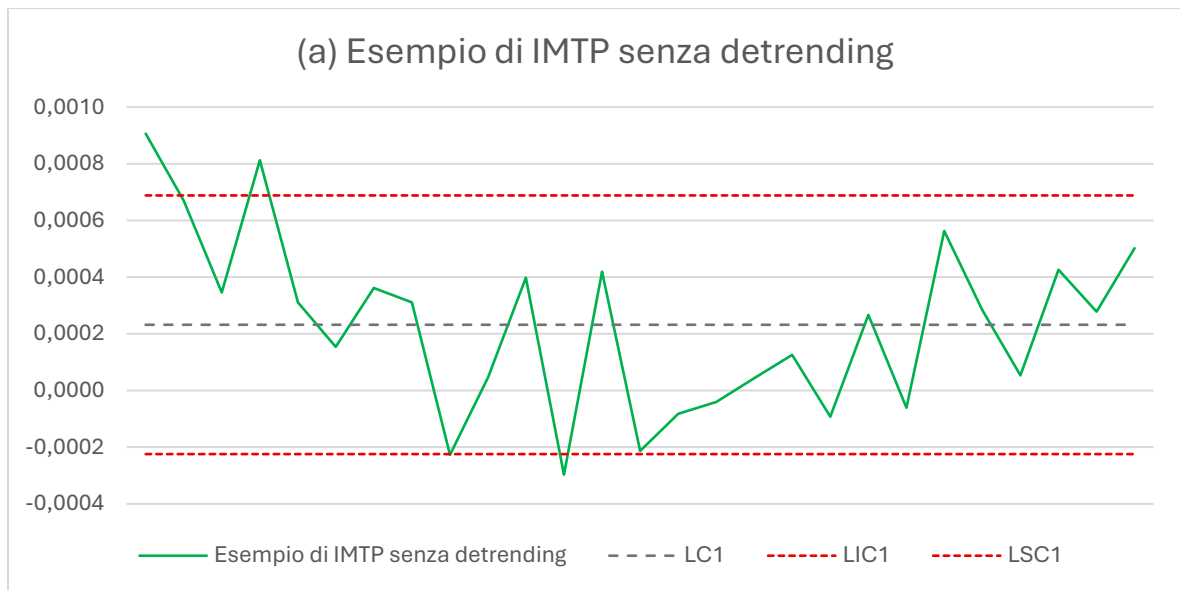


Figura 34: Calcolo della Carta $\bar{X}$ per un IMTP senza detrending (a) e uno con detrending (b). Il grafico (b) mostra un aumento dei punti di fuori controllo.

In Figura 33, il numero di punti fuori controllo è diminuito del 42% (da 19 a 11), mentre in Figura 34 è aumentato del 125% (da 4 a 9); il caso in cui invece il numero di punti fuori controllo è rimasto lo stesso, è quello indicato nel confronto in Figura 32. Ciò significa che l'attività di detrending è utile per minimizzare le anomalie potenzialmente causate dalla presenza di progressioni, ma non vuol dire che

necessariamente il processo, una volta applicate le carte, risulti più in controllo rispetto a prima.

Un ulteriore aspetto da considerare è quello legato all'ampiezza dei limiti superiori e inferiori, che dipendono dalla numerosità del campione analizzato: per campioni abbastanza ampi, come quelli considerati nel presente caso di studio, i limiti di controllo diventano più stringenti, andando ad aumentare la probabilità che un valore appartenente al processo sia fuori controllo; al contrario, per numerosità basse, i limiti risultano più ampi, permettendo la presenza di più punti all'interno di essi.

Per quanto riguarda la Carta S , la modalità di acquisizione dei dati della deviazione standard è stata differente, in quanto si è partiti dal database iniziale delle Topical Prevalence, e non dai residui ottenuti dopo il detrending; questa operazione potrebbe essere stata una delle cause della presenza dei numerosi punti fuori controllo.

In generale, la metodologia proposta basata sull'utilizzo di algoritmi di Topic Modeling STM risulta di fondamentale importanza nel raccogliere ed evidenziare gli argomenti più discussi all'interno degli UGC che costituiscono la Digital Voice-of-Customer. L'applicazione delle carte di controllo con e senza detrending, ha dimostrato come questa fase risulti utile maggiormente nei casi in cui le tendenze sono molto accentuate: eliminando i trend, infatti, vengono annullate anche le anomalie dovute ad essi, favorendo l'analisi del processo anche con uno strumento a media fissa e non pendente, come la Carta $\bar{X} - S$.

Per le procedure di calcolo effettuate nel presente lavoro si può affermare che la Carta $\bar{X}$ riesce a intercettare con sufficiente precisione anomalie macroscopiche

(come, ad esempio, gli effetti del periodo pandemico), ma la componente S , in particolare a causa dell'alta variabilità dei processi analizzati per i dodici topic, risulta molto sensibile e di difficile interpretazione. Tuttavia, soprattutto per la sua semplicità di applicazione anche in contesti dove sono presenti delle tendenze, come in questo caso dove è stato effettuato il detrending, essa potrebbe essere sicuramente presa in considerazione in un contesto aziendale al fine di monitorare l'andamento della Topical Prevalence. Dal punto di vista operativo, i segnali forniti dalle carte possono essere interpretati dalle imprese come piccoli allarmi; la presenza di punti fuori controllo infatti indica che la percezione dei clienti rispetto a uno specifico tema sta subendo variazioni inattese. Tutto ciò può guidare i responsabili verso approfondimenti mirati e interventi tempestivi. Ad esempio, un picco fuori controllo sul topic relativo alla puntualità delle consegne, potrebbe suggerire la necessità di indagare i tempi medi di evasione degli ordini o l'affidabilità della logistica. A tal proposito, è molto importante associare al monitoraggio anche il collegamento con relativi KPI aziendali, come ad esempio il tasso di reclami sui ritardi, la percentuale di ordini incompleti e il tempo medio di rimborso, andando ad agire quando ritenuto necessario.

CAPITOLO 4 – CONCLUSIONI

Il presente lavoro di tesi aveva come obiettivo principale valutare l'efficacia della carta di controllo $\bar{X} - S$ nell'analisi della Digital Voice of Customer (VoC), relativamente al settore dei servizi del food delivery, con la finalità di comprendere se tale strumento statistico potesse costituire un valido supporto per le aziende operanti in questo ambito nel monitoraggio della qualità percepita e nel prendere decisioni strategiche tempestive. La domanda di ricerca di partenza può essere così riassunta: *le Carte $\bar{X} - S$ con detrending sono applicabili ed efficaci per analizzare i risultati ottenuti dopo l'applicazione di algoritmi di Topic Modeling sulla Digital Voice-of-Customer?*

La ricerca è iniziata con l'estrazione, tramite dei software, del database contenente l'insieme di recensioni pubblicate dagli utenti, utilizzate come fonte di dati per le successive analisi. È stato utilizzato un algoritmo di Topic Modeling al fine di intercettare gli argomenti più discussi all'interno dell'intero set di dati. Dopo aver superato la fase di validazione, sono state effettuate delle analisi sulle recensioni appartenenti limitatamente all'azienda Glovo, con la valutazione delle varie tendenze grafiche, l'applicazione del detrending e l'utilizzo della Carta $\bar{X} - S$ per il monitoraggio.

L'analisi condotta ha permesso di dare una risposta articolata. I risultati ottenuti hanno evidenziato che la carta di controllo $\bar{X} - S$, pur essendo uno strumento consolidato nel campo del controllo statistico in processi industriali e produttivi, non si è dimostrata pienamente adeguata al monitoraggio di dati testuali e qualitativi caratterizzati da alta variabilità, come quelli derivanti dagli User Generated Contents (UGC). Tuttavia, seppur non con effetti diretti sui risultati del presente lavoro, è stato

dimostrato come in generale l'attività di detrending, grazie a una lettura più stabile delle dinamiche temporali, permetta di migliorare la qualità del monitoraggio nei casi di applicazione della Carta $\bar{X} - S$, andando a diminuire notevolmente la possibilità di rilevamento di punti fuori controllo dovuti all'effetto delle pendenze. In altre parole, il detrending rappresenta un passaggio molto utile quando l'andamento dei processi mostra una chiara tendenza, che sia essa lineare o no, poiché permette di distinguere variazioni reali da semplici effetti del trend. Viceversa, in presenza di tendenze molto contenute, i benefici del detrending risultano marginali e può essere utile mantenere una visione anche sull'IMTP grezzo, per non perdere informazioni legate all'evoluzione del fenomeno. Questo esito, sebbene circoscritto a questo caso di studio in particolare, rappresenta comunque un contributo scientifico interessante sia per i ricercatori che per le aziende, evidenziando infatti i limiti di applicabilità degli strumenti tradizionali di controllo statistico dei processi, nei casi in cui vengono applicati a dati non convenzionali e altamente dinamici, come quello delle piattaforme digitali.

Il contributo originale del presente lavoro di tesi risiede nell'aver collegato, in modo sperimentale, un approccio tradizionale di controllo statistico a un dominio innovativo come quello della Digital VoC analizzata con algoritmi di Topic Modeling, limitatamente ai casi in cui i processi mostrano delle tendenze che vengono eliminate tramite una fase di detrending. In letteratura sono pochi i lavori che esplorano tale legame, e i risultati ottenuti evidenziano limiti e potenzialità che aprono la strada a ulteriori sviluppi; in particolare, lo studio ha dimostrato che strumenti classici, pur non trasferibili in maniera immediata, possono essere adattati e valutati criticamente per comprendere meglio i dati generati dagli utenti nelle piattaforme digitali.

Occorre comunque sottolineare alcuni limiti della metodologia proposta. In primo luogo, la base dati utilizzata, seppur significativa, è stata circoscritta a recensioni pubblicate su un solo aggregatore (Play Store Google), con una conseguente possibile riduzione della rappresentatività dei risultati. Questo può aver introdotto un bias legato al profilo degli utenti Android, con un impatto amplificato su alcune specifiche dinamiche della community (es. sul *Topic 1 Affidabilità dell'app*), influenzando così gli IMTP osservati e riducendo la possibilità di generalizzare i risultati all'intera popolazione di clienti del servizio.

Un secondo limite è rappresentato dalla modalità di acquisizione dei dati: se da un lato è stato possibile effettuare il detrending per la costruzione della Carta $\bar{X}$ con risultati soddisfacenti, per la Carta S non è stato possibile operare allo stesso modo, ottenendo dei risultati distorti e poco informativi, evidenziando la sensibilità dello strumento. Su questo punto è doveroso sottolineare che, a causa delle leggere tendenze riscontrate nei grafici IMTP, i risultati applicati alla Carta $\bar{X}$ non si sono discostati molto dal caso originale. Questa considerazione lascia ipotizzare che anche nel caso di applicazione della Carta S al processo detrendizzato, i risultati sarebbero stati simili a quelli ottenuti sui processi iniziali, i quali si possono considerare a questo punto caratterizzati da un'alta variabilità, che giustifica i punti di fuori controllo.

Infine, l'applicazione di una singola tipologia di carta di controllo ha impedito un confronto diretto con altri approcci statistici o di machine learning nel contesto dell'analisi dei dati di natura testuale.

Alla luce di queste considerazioni, i futuri sviluppi della ricerca potrebbero concentrarsi sui seguenti aspetti:

- L'esplorazione di strumenti alternativi di controllo statistico meglio equipaggiati per gestire i trend, come le carte a media mobile, andando a confrontare i risultati ottenuti con esse rispetto a quelli ricavati nel presente lavoro con l'applicazione dello strumento $\bar{X} - S$ dopo una fase di detrending;
- L'adozione di approcci ibridi che combinino metodi statistici tradizionali con algoritmi di machine learning, capaci di adattarsi meglio alla complessità dei dati digitali;
- L'ampliamento del dataset, includendo più piattaforme e/o periodi temporali più lunghi (raggruppandoli potenzialmente in trimestri o semestri per rafforzare la significatività dei risultati e consentire un calcolo più stabile della variabilità, nel caso di carte come la $\bar{X} - S$).

In sintesi, questa tesi ha dimostrato che non tutti gli strumenti di controllo statistico classici risultano immediatamente trasferibili alla Digital Voice of Customer, e che la valutazione critica di tali strumenti è essenziale per sviluppare metodologie alternative. Ad esempio, il loro utilizzo insieme a una fase di detrending può offrire indicazioni utili quando i processi mostrano forti tendenze, mentre in altri casi è consigliabile mantenere anche una lettura dei dati grezzi. Sebbene la carta $\bar{X} - S$ non si sia dimostrata lo strumento più efficace in questo contesto, la ricerca compiuta rappresenta un passo importante verso la comprensione dei limiti e della potenzialità degli approcci statistici nel monitoraggio della qualità percepita dai clienti. Le aziende possono quindi trarre vantaggio dai risultati di questo approccio, interpretando i segnali delle carte non come decisioni definitive, ma come alert da integrare con altri KPI operativi, per supportare scelte strategiche e migliorare la qualità del servizio offerto.

APPENDICE I

```
#-----  
#INPUT DATI  
#-----  
data <- read.csv2("Database_food.csv") #CSV RECENSIONI  
rimuovere <- read.csv2("Stopwords.csv")  
  
#in alternativa all'utilizzo della funzione read.csv si può creare la lista di stopwords  
direttamente all'interno dello script  
  
# -----  
# PREPARE AND PRE-PROCESS DATA  
# -----  
#PRE-PROCESSING  
  
processed <- textProcessor(data$Review, metadata = data, customstopwords =  
rimuovere$Stopwords, verbose=TRUE) #impostare la variabile corretta per le custom  
stopwords  
  
out <- prepDocuments(processed$documents, processed$vocab, processed$meta,  
lower.thresh=15, verbose=TRUE) #ELIMINA PAROLE CHE APPAIONO MENO DI 15 VOLTE  
  
docs <- out$Review  
vocab <- out$vocab  
meta <- out$meta  
  
  
#CREAZIONE FILE CON LE REVIEWS RIMANENTI. Viene salvato un file per ogni colonna del  
database in ingresso.  
  
write.csv (meta$Fonte,file='Fonte_Rimanenti.csv')  
write.csv (meta$Company,file='Company_Rimanenti.csv')  
write.csv (meta$Data,file='Data_Rimanenti.csv')  
write.csv (meta$Rating,file='Rating_Rimanenti.csv')  
write.csv (meta$Review,file='Review_Rimanenti.csv')  
  
#-----  
# ESTIMATE THE STRUCTURAL TOPIC MODEL
```



```
# -----
```

```
# VALUTAZIONE NUMERO DI TOPIC IDEALE
```

```
# 4 DIAGNOSTIC DIMENSIONS (held-out likelihood, Residuals, Semantic Coherence, Lower Bound)
```

```
# la funzione serchK applica iterativamente STM variando il numero di topic secondo l'insieme c
```

```
# I risultati sono salvati nel file optimizationresults.csv
```

```
c= (5:50) #c vettore con numeri da 5 a 10
```

```
K<-c
```

```
storage <- searchK(out$documents, out$vocab, K, data = meta)
```

```
plot(storage)
```

```
write.csv(unlist(storage$results),file='optimizationresults.csv')
```

```
#APPLICAZIONE STM
```

```
k=12 #impostare il numero di topic da estrarre
```

```
poliblogPrevFit <- stm(documents=out$documents, vocab=out$vocab,K=k ,max.em.its=75, data=out$meta, init.type="Spectral")
```

```
write.csv (poliblogPrevFit$theta,file='matrice_review_topics.csv')
```

```
#per ottenere il file finale da analizzare unire con excel i file delle review rimanenti e la matrice review/topics
```

```
#PLOT RISULTATI TOP TOPICS/PERCENTUALE TOPICS NEL TEXT CORPUS
```

```
plot(poliblogPrevFit,type="summary")
```

```
#-----
```

```
# LABELLING
```

```
# -----
```

```
#STAMPA TOP 10 KEYWORDS PER I 12 TOPICS UTILIZZANDO 4 RANKINGS (Highest probability, FREX, Score, Lift)
```

```
labelTopics(poliblogPrevFit, c(1:12), n=10)
```

```
#STAMPA d DOCUMENTI PIU' RAPPRESENTATIVI PER IL TOPIC t
```

```
t=12 #impostare il numero del topic di cui si vogliono visualizzare i documenti più significativi
```

```
d=5 #impostare numero documenti da visualizzare
```

```
thoughts <- findThoughts(poliblogPrevFit, texts=out$meta$Review, n=d, topics=t)
```

```
thoughts
```

APPENDICE II

Topic 1 Top Words:

- **Highest Prob:** app, work, use, updat, open, order, restaur, fix, notif, uninstal
- **FREX:** open, notif, load, screen, updat, crash, stuck, user, button, function
- **Lift:** user-, anim, load, refresh, brows, crash, malfunct, aggress, intuit, open
- **Score:** user-, app, work, notif, open, updat, bug, screen, load, search

Topic 2 Top Words:

- **Highest Prob:** card, phone, payment, number, code, account, use, app, promo, chang
- **FREX:** code, number, payment, card, phone, method, promo, verifi, log, countri
- **Lift:** laggi, sms, verif, verifi, digit, spain, invalid, password, spanish, code
- **Score:** card, laggi, number, payment, code, phone, account, promo, log, password

Topic 3 Top Words:

- **Highest Prob:** item, miss, order, someth, product, avail, shop, store, use, applic
- **FREX:** product, shop, groceri, miss, supermarket, someth, item, unavail, lot, stock
- **Lift:** comfort, groceri, shop, product, substitut, suffici, supermarket, stock, weve, particular
- **Score:** item, miss, comfort, product, someth, shop, groceri, avail, choic, store

Topic 4 Top Words:

- **Highest Prob:** charg, deliveri, pay, fee, price, cost, amount, order, plus, expans
- **FREX:** price, fee, charg, expans, plus, subscript, cost, cheaper, higher, minimum
- **Lift:** hike, inflat, surcharg, tax, price, fee, minimum, higher, weather, cheaper
- **Score:** fee, charg, price, hike, pay, deliveri, expans, cost, amount, plus

Topic 5 Top Words:

- **Highest Prob:** servic, use, deliveri, issu, absolut, care, deliv, provid, help, late

- **FREX:** servic, complaint, absolut, love, rude, communic, compens, shock, team, unprofession
- **Lift:** speed, complaint, satisfactori, unprofession, love, deal, profession, shock, dread, favor
- **Score:** servic, speed, deliveri, issu, absolut, care, compens, rude, late, team

Topic 6 Top Words:

- **Highest Prob:** refund, money, credit, receiv, order, refus, account, paid, ask, request
- **FREX:** money, refus, refund, request, paid, credit, transact, proof, taken, polici
- **Lift:** atroci, photograph, disput, proof, evid, owe, prove, upload, scammer, rob
- **Score:** refund, atroci, money, credit, refus, account, receiv, paid, bank, request

Topic 7 Top Words:

- **Highest Prob:** order, cancel, hour, wait, restaur, call, deliv, minut, place, close
- **FREX:** cancel, wait, hour, call, minut, prepar, assign, close, took, min
- **Lift:** son, assign, cancel, wait, hour, prepar, fulfil, reorder, call, row
- **Score:** cancel, hour, order, wait, son, call, minut, restaur, min, deliv

Topic 8 Top Words:

- **Highest Prob:** restaur, option, meal, add, order, thing, menu, save, choos, note
- **FREX:** meal, navig, easier, sauc, thing, effici, note, box, save, allergi
- **Lift:** extra, salad, chees, sauc, beauti, ingredi, customis, burger, dessert, contain
- **Score:** extra, effici, afford, meal, add, restaur, menu, option, save, thing

Topic 9 Top Words:

- **Highest Prob:** time, deliveri, cold, take, arriv, deliv, courier, late, quick, order
- **FREX:** estim, courier, quick, hot, time, consist, track, awesom, cold, warm
- **Lift:** safe, warm, bicycl, tasti, awesom, bike, earli, estim, predict, delici
- **Score:** safe, time, deliveri, cold, estim, quick, hot, arriv, courier, fast

Topic 10 Top Words:

- **Highest Prob:** support, help, contact, chat, use, app, offer, restaur, system, voucher
- **FREX:** review, voucher, coupon, question, bot, support, answer, oper, chat, standard
- **Lift:** eas, oper, review, hygien, tech, bot, holiday, takeout, human, question
- **Score:** eas, support, chat, contact, help, voucher, review, answer, coupon, offer

Topic 11 Top Words:

- **Highest Prob:** driver, address, find, deliv, locat, tip, option, deliveri, door, app
- **FREX:** address, door, hous, map, street, outsid, driver, area, locat, gps
- **Lift:** train, pin, street, gps, road, build, hous, knock, map, outsid
- **Score:** train, driver, address, tip, locat, map, door, find, hous, translat

Topic 12 Top Words:

- **Highest Prob:** order, app, use, place, tri, done, get, differ, delet, that
- **FREX:** youv, done, guess, tri, whether, delet, come, that, differ, let
- **Lift:** popup, accid, usual, youv, dismiss, true, everywher, whether, oop, done
- **Score:** popup, order, app, use, done, tri, youv, delet, differ, place

Topic 1:

1. Just installed the app, opened it and was greeted with a message saying that an update was needed to continue using the app. Odd, but I clicked the update link....which takes me back to the Google Play Just Eat page with options to 'Uninstall' or 'Open' the app. I click the 'Open' link....and round and round we go. Good for nothing at the moment.
2. Use MenuLog and Foodora instead. Deliberately bombards me with notifications via the app. But... you can't turn off via the app; you have to login via the website. Go to website and login, it then defaults back to the app... Please. I'll just uninstall rather than deal with this stupid loop. I turned off notifications twice and they keep coming back; every updates turn back on I guess because I never enable them; with any app. Annoying marketing making things more difficult than they need to be. Just allow the notifications to be turned off permanently.

3. Great service, bad app. Million of bugs. Keyboard doesn't open for search, back button gesture doesn't work, repeatedly opens some screens, closes the app after opening from notification, also has unremovable notifications
4. It was great at the beginning but now i am very disappointed. When i open the app, most of the times (9 out of 10), it shows only the loading screen. Only when i close and open several times the app i can pass the loading screen. I tried to uninstall and install the app several times but this doesn't fix the problem. I thought that only i have this problem but several of my friends have this problem
5. Awful, horrible app design. Basics like the back button don't work, you can get stuck in infinite loops in the referral screen with no way to exit it but to kill the app, bugs galore and terrible CX. Worst app I've ever used.

Topic 2:

1. Can't log in, I have used the app before then I remove it and now I can't enter after installing it again. The issue is that I have changed my phone number, so after I log in with Facebook or with mail and password I get to a page that ask for a phone code, that they are sending to my old phone. In the web page, after you reset your password you are not login so I can't update my phone. The weird thing is that I can change my password but not my phone number.
2. Lost my phone and got a new number. There is NO way to use my account. It keeps sending me SMS codes when I try to log in the web but absolutely ZERO option to change number or verify account. It's a shame cuz I was using this almost daily.
3. Was not able to order food because Glovo did not want to verify any of mine phone numbers. Said "this number can not be verified" but the phone numbers are valid. When tried to fill it in settings said that I have reached the maximum number of attempts. Now I should stay hungry. Do not recommed
4. I cant enter on my account, just nothing happening when Im trying to log in. I wrote to support but they are telling me that account working but I cant still access to it. Creating new account don't solve the issue, i cant enter my phone number, getting error. pls fix it
5. They hacked my Facebook account. I chose to sign up with my FB account and they reset my password + now i can no longer log in. I checked and apparently now theres a new email associated with my FB account that is not mine.

I've been trying to recover it for days... Don't trust this app

Topic 3:

1. It's really easy to use. Very easy to find past orders, items I liked, so I can order again. Lots of choices of restaurants. It's been such a comfort during lockdowns because it's so simple.
2. Awful experience. I order groceries from the supermarket Glovo and they delivered expired products... Some had mold. Don't buy here, you will regret it
3. Items always missing, no matter who you order from the drivers never check, reporting the missing items is a nightmare as if you order a meal deal and part of it's missing you can only say All of it's missing which is not true. It's just rubbish. worse but if you spend ages selecting all the meal deal options only for them not to deliver! Pointless.
4. I added food in cart and when I tried to order it said the store or products are not available. I tried it a few times with the same result. Time wasted. No way to submit a bug besides this form.
5. Every time I order groceries through Deliveroo, something seems to be missing, the last two times I ordered I had over £10 worth of items missing per order, including a bottle of wine. Deliveroo customer service says that there is no option to give me a refund for any of the missing items in either of the orders, despite the fact that I never received the items. I never have this issue with any other company, only Deliveroo.

Topic 4:

1. FEES FEES FEES..... Not bad except for the delivery fees added. I tip the delivery person PLUS delivery fees PLUS they ask for restaurant tip...it never ends... Then if you don't want the delivery fees...YOU HAVE TO PAY FOR THAT TOO ! FEES FEES FEES.... They offer 2 for 1 but with the delivery fees it ruins the deal. The fact a delivery often stops elsewhere while on the way to you unless you pay more ruins another part of the experience hence the 3 stars.
2. If I have to pay taxes and FEES then why can't I know what these fees are for and how much they are a 9 dollar meal shouldn't 17 dollars even with a 3 dollar tip it's like having to pay 75 cents for every dollar you spend, and taxes ain't that high, so what are these fees ??? and why is it hidden if I'm forced to pay it anyway.
3. Made an order and it said that the order cost x with a discount applied. But I was taxed full price without the discount. It doesn't bother me that I paid the full

- price but don't trick people into thinking they will have a discount but charge them the full price.
4. If I would describe my experience with Glovo as "corporate greed" it wouldn't do it justice. Let me explain. I started using Glovo right after launch here in Romania. The strategy they applied seems to mimic the Romanian business mentality perfectly, which is "get rich fast, nevermind what comes tomorrow". At first they had RON 2.99 transport fee. A very competitive fee compared to the competition, Uber Eats. As months went by, they started to add 1 RON approximately each month. Fair enough I said, that was a starting fee, as they made their way onto the market. Then, they started to surpass the competition with unjustified transport fees for some establishments which were actually closer to me than others. Then, they introduced a "bad weather" fee, a fee which is activated by the app and thus sometimes only triggers when it's a cloudy summer day, no rain or anything. Then, as if that was not enough, they introduced a minimum order fee, just like those delivery services or the regular restaurants which do not sport a transportation fee at all, only a minimum order fee to justify long distances. But Glovo has this fee for all restaurants. Because why not? And then, when I finally thought I saw the greed threshold at its peak, they introduced the wrapping fee, now you need to pay for the wrapping. And you are forced to choose it, no way around it, and then sometimes the app gets bugged and charges you 1 wrapping for each Chinese dumpling and so for 10 dumplings you pay for example 10 RON and then for the wrapping you pay 20 RON. I guess the corporate meetings at Glovo are only about brainstorming a new fee. I am surprised they are still on the market and I will never ever order from them again. Such businesses deserve bankruptcy.
 5. Only app that adds a surcharge on top of a minimum price. E.g. The minimum price is 30 (local currency) and the surcharge rate is applied below 45. Other apps not only offer free delivery but don't have this surcharge garbage. What you offer does not justify the extra greed.

Topic 5:

1. Horrible, horrible service. Multiple times my order was not delivered to me (but labeled as delivered) and customer service does nothing to help. They obviously do not screen drivers, drivers do not care and just unprofessional. Was issued a refund once after pleading with the customer service. They do not care, they do not have a professional protocol, they do not establish good

- communication between the customer and the company. would not use it again in a million years.
2. Just stick to just eat or uber. Over 3 hours to deliver food, lie about the reason for the delay, deliver cold food and then offer a completely unacceptable level of compensation. Won't be using again. Customer service is horrible. No point talking to the reps as they can't do anything. They do promise to update their processes which is absolutely pointless to the problem you are trying to solve at that point. I'm uninstalling.
 3. I had a truly terrible experience with a delivery service in France recently, and I feel compelled to share my thoughts in a review. Unfortunately, my food delivery was stolen by the delivery person, which was a major disappointment. Even worse, when I tried to reach out to customer service for a refund, they provided no assistance whatsoever and even refused to provide me with the name of the delivery person involved. I found this to be terrible.
 4. Both myself and my partner have orders more than 2 hours or just not turn up late recently. No offer of compensation was offered from just eat for their poor service. When you receive bad service you shouldn't have to ask for a company to make it right it should offered as recognition that their service was not up to scratch on that occasion. As a result I have uninstalled the app.
 5. I ordered from this app, the delivery service was appalling. My food came an hour late, cold and was completely wrong, none of the food was right, none! The delivery lady was rude, didn't say a thing so I suspect she knew it was the wrong order and still delivered it without apology. Awful. I'm not using this delivery service even again, absolutely fuming.

Topic 6:

1. THEY ARE THIEVES! they are scamming people, they used my bank account as a way to make money transaction for them, I tried to contact them about the issue and they said they are not in charge about this and that I should consult my bank agency, and it obviously says in my bank chart GLOVO, which is their mistake not the bank. never deal with them again, worst experience!
2. Service is generally good however don't hold your breath when it comes to refunds. The restaurant sent us the completely wrong food. I reported it immediately to Justeat and send them photographs to prove it. Refund was approved however it is now 2 weeks later and still have not recieved my refund.
3. This app is pure scam,I ordered and was debited twice still didn't get order and yet no refund was made since

March 16th was told it takes 3-10 working days to get a refund but till date non and have sent screenshot of my bank transactions history yet no refund was told to go and be getting bank statement.. useless app refund my money

4. It is too good to be true. Every order will clutter your bank statements with 3 line items (you send them money then they still charge and usually refund your 'deposit'?). Eventually your bank statement is huge and that's when they don't always refund your deposit and make random withdrawals from your account. Getting a refund takes weeks of sending the same screenshots over and over again and repeating yourself for hours. If they do eventually refund you it will be a random lesser amount.
5. Refused to refund me despite me sending photographic evidence of the receipt that I recieved the wrong order

Topic 7:

1. Very lousy . Most vendors don't accept orders .. waited long time then cancel order your orders or hair kept you waiting for them to accept when you have no idea how long u need to wait , or you will realize your order being cancelled after expecting the order for 1 hour or more
2. I made an order, restaurant accepted it, waited for one hour then called restaurant, they said they had other orders with higher prices so they will not attend to my order of pizza for 11 Euros, and they didnt even bother to cancel the order, they kept me waiting. I got in touch with deliveroo via chat, all they had to do is call the restaurant, then tell me sorry, and cancel my order! Not the best first experience
3. Placed my order, kept checking the status of my order, it went through the process of preparing food for 30 minutes, even the rider was at restaurant for pickup, but the order got cancelled by the restaurant. Went on live chat, as something like this happening on other delivery apps like Ubereats, Ubereats would compensate with a discount code immediately. However, deliveroo live chat said it's normal for restaurants to cancel 30 min into the order. Don't accept my order if it can't be delivered
4. Very bad experience with this app, cancelled the order just after 3 mins of placing it but they still delivered after 2 hours of cancellation. The worst part is there live chat thing has very lazy people, the person who I was chatting with kept on lingering the matter for 25mins and didn't cancelled the order which was already cancelled 2 hours earlier.
5. Worthless and theives I ordered a 27 pound order from this company and the restaurant never delivered and

deliveroo did nothing to help me bare in mind I have mental disabilities the restaurant marked it as delivered while I was waiting for my food all night and they called me a liar and deliveroo listened to the restaurant who clearly didn't deliver my order I called the restaurant 3 times to cancel and they didn't want yo cancel it Never depend on deliveroo Get uber one way more reliable

Topic 8:

1. Some restaurants with Uber only give you part of their menu. If you order from Skip you get the entire menu but for some reason Uber limits the menu. For example, Chuck's roadhouse has brochuto bread on the menu through skip. You can't order it through Uber. Also subway doesn't show all of the topping choices through Uber. If you know that they are you can ask in the notes and hope that the person filling the order reads it but why don't they put all the choices on the menu?
2. Terrible app. How you can order a subway if there is no option to add ingredient? How can you remove an ingredient that you don't like? writing it as a dietary requirement is just stupid. And many times you write that you don't want an specific ingredient and the app just don't include this information and if you have problems with allergy like me, you will constantly get food that you can't eat. UBER EATS MUCH BETTER IN THIS CASE. SAVE TIME AND LET YOU PERSONALIZE YOUR ORDER IN A SIMPLE WAY
3. Super great!! And easy to navigate. I just wish the menu options gave more add on options like single burger patties or side of beef kind of thing. otherwise, super fabulous thank you :)
4. Not easy to amend an ongoing order, can't do it from basket , have go back through restaurant list, find supplier, go to their menu again to add items... Can't search by required item, only restaurant names Eg. If I want a restaurant that does 'Coke' I have to trawl through every menu... Restaurants that ask you to write your multiple choice in the description, but don't list the choices.
5. Working as it should, the only issue is that I get indecisive. If there was description of what some dishes are with unfamiliar names it would make it easier trying new things. And pictures. I wish all menus came with pictures.

Topic 9:

1. I am having one of the worst experiences with this establishment. while they used to be good, lately every order has been arriving later than the promised time. They claim that the later delivery time ensures that the

- food arrives warm and fresh, but each time the food has been cold and inedible. Initially, they provide an early time estimate upon placing the order, and we plan accordingly. However, lately they have been consistently delaying the delivery, disrupting our plans.
2. They are almost always late (between 20 and 40 min late, systematically). The problem is that the initial estimated time of arrival is not to be trusted. Couldn't they change this issue easily by being more honest or accurate in their initial estimation?
 3. deloveroo is good in theory.. they promise delivery times that they consistently fail to meet. The whole tracking of order is not accurate as it feels like the order is getting ready quickly and then the delivery takes ages but when you ask the delivery people why it was late they say they were waiting the food. obviously there is no customer service to check, track and complain. This would have been an amazing patience tracking application, tests your patience and keeps time accurately.
 4. The app is really really good. A lot of restaurants, a lot of deals, simple to use and fast. I can't give 5 stars because they give really bad estimates for when will the delivery arrive. It's a shame honestly, because they consistently give unrealistic times (like having the food delivered in 15-20m) while it ends up taking 30-50m most of the time. I mean, if you give an estimate time, I think the priority has to be to let the clients know when, rather than pressure the restaurants with it.
 5. Lately the delivery times have been taking longer. It will show an expected delivery time that will repeatedly get extended. Often past the latest arrival.

Topic 10:

1. My previous 2 reviews were bad. This is still vary bad. Example : you search for a know restaurant like McDonald's. Result:not found. You dig a lot and find that the rest is closed at that time. what if you just provide a feedback like: Your rest is closed at this time. I got coupons in my email on 12 Nov. Tried to order today, 14 Nov. Coupons are not found. Tried to contact support from app. No support, except for the FAQ way beyond any standard of customer service
2. One of the worst app designs I've seen in a while! Can't contact the support easily. When you finally reach support they are everything but helpful. 1/5 would not recommend to anyone! would rather use the alternative such as wolt!
3. I have an issue in the app and it says to contact customer support. But there is no contact information for customer support! I can't get this issue resolved meaning it's

impossible to use the app now. I hope Glovo could contact me to resolve this but until then I'm extremely disappointed.

4. App says restaurants are closed when they aren't, and you can't actually contact anyone when it comes to customer support. Just a bunch of bots to further frustrate you.
5. Contacting support shouldn't be this hard. They purposely make it as hard as they can to actually talk to a human, redirecting you to countless useless FAQ pages.

Topic 11:

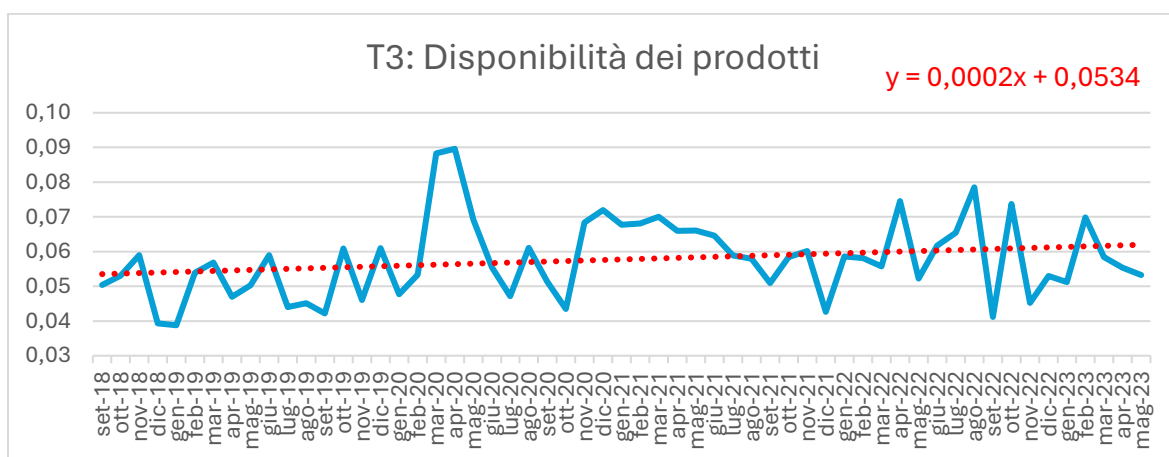
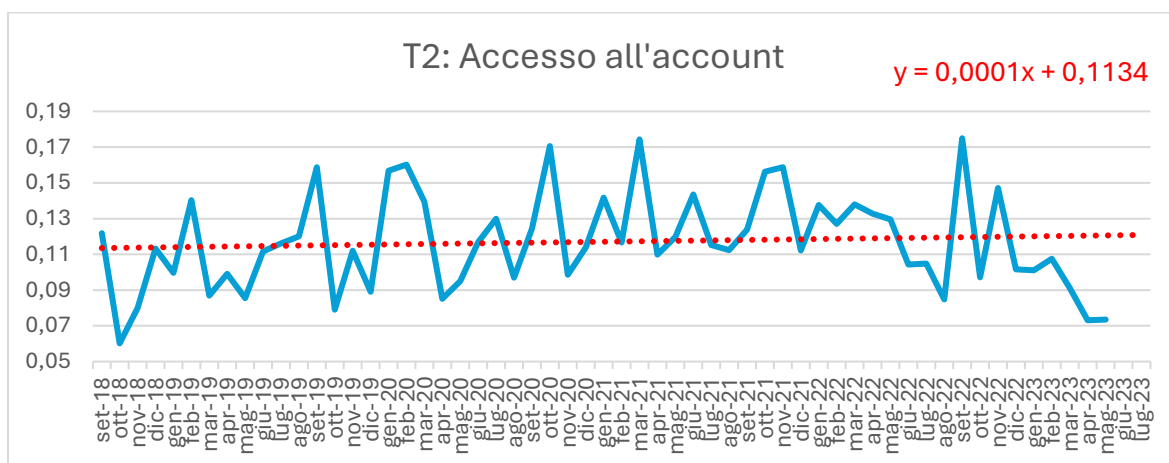
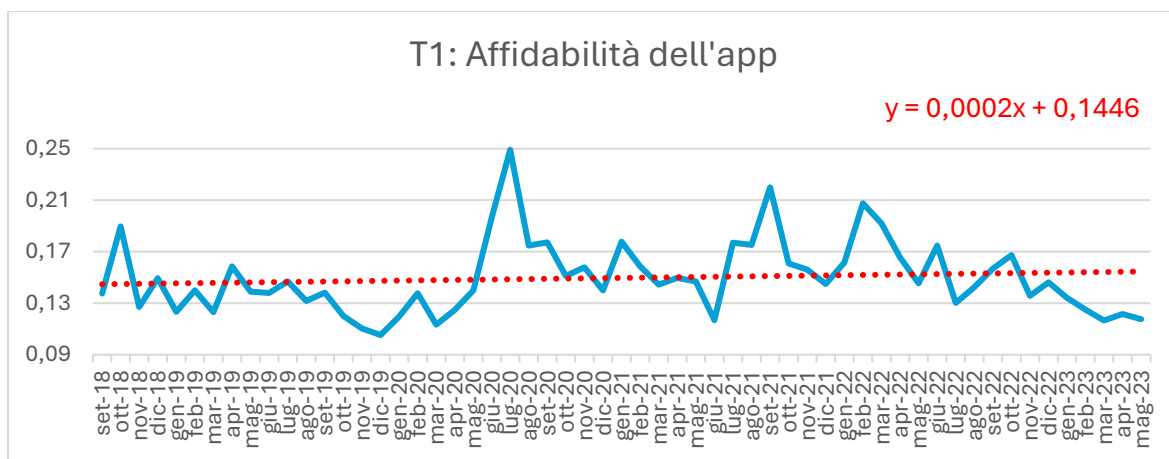
1. very poor service. the delivery driver was looking for my house. i gave clear instructions to where my house is. it was very easy to find. but my driver was 2 miles away looking for my address. and i couldn't call the driver because with deliveroo u can only call once they are near u. i couldn't text her. because there is no option for text. and also there is no option for rating. my 2nd driver did the same thing. he marked delivered without even coming close to this address.
2. For the most part this app is great. However when it comes to the map, the pin to locate my address is always incorrectly placed on the map, so the drivers always get confused. Just some feedback that should be flagged.
3. Delivered my food to the wrong address. wouldn't allow me to change the address even though the driver hadn't picked up the food yet and even though the wrong address was only five minutes drive from the correct address.
4. For the most part my experience has been okay with Uber Eats. My only critique would be that the drivers delivering the product/food need to follow the customers instructions. In the past I've left several instructions and notes regarding my delivery, for the driver to do the total opposite. There are several drivers I wished I could've taking my tips back from because they didn't deserve it by not following my instructions. FYI it's very important to pay attention to details.
5. The postcode for my road is wrong and the drivers keep turning up at a different address. Lived in this house nearly 10 years and it has never changed. The pin pointer doesn't seem to recognise my house either. The only delivery service that has ever had a problem...

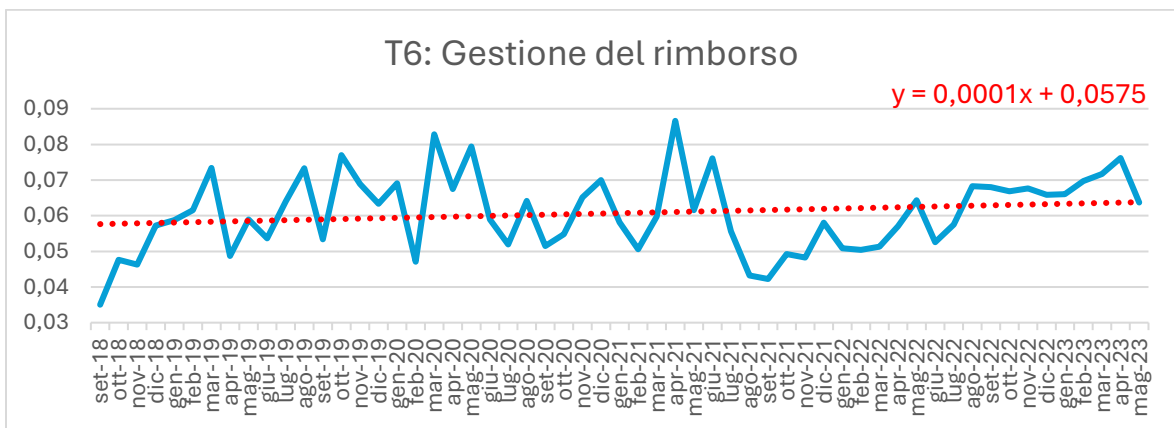
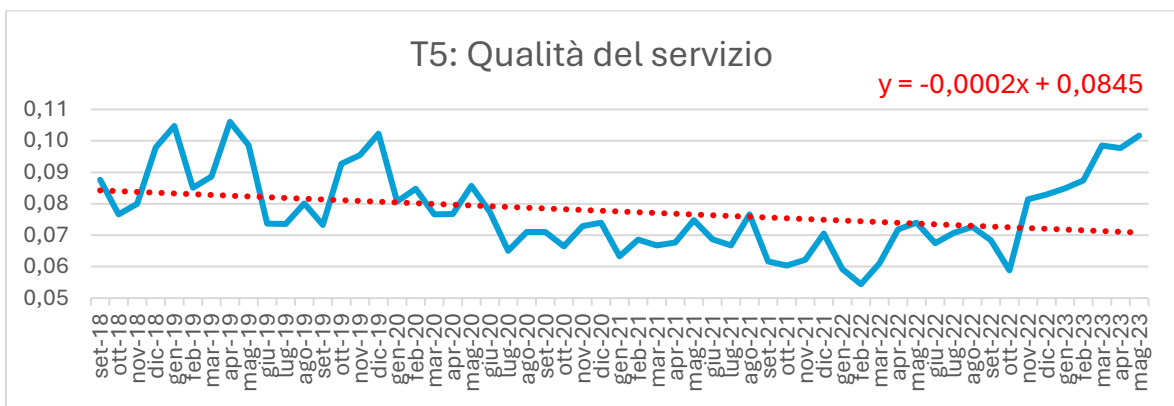
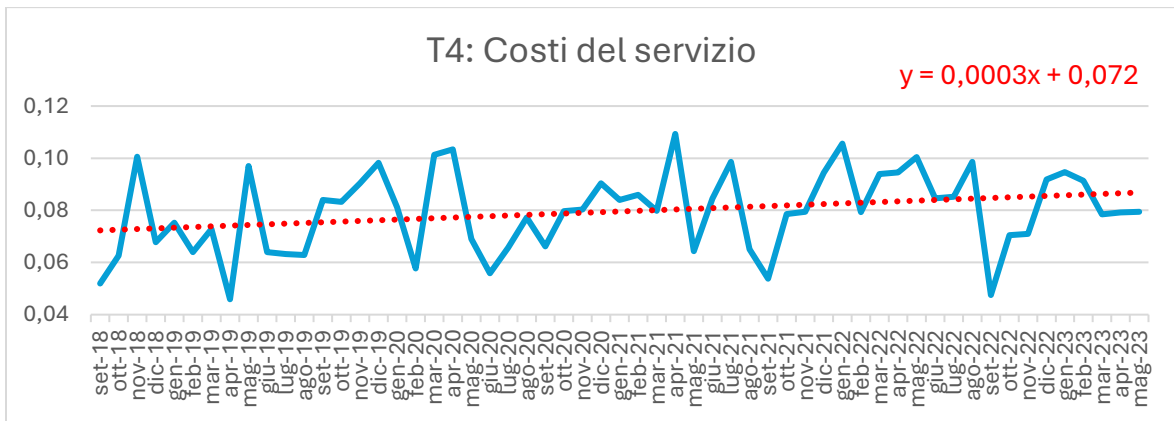
Topic 12:

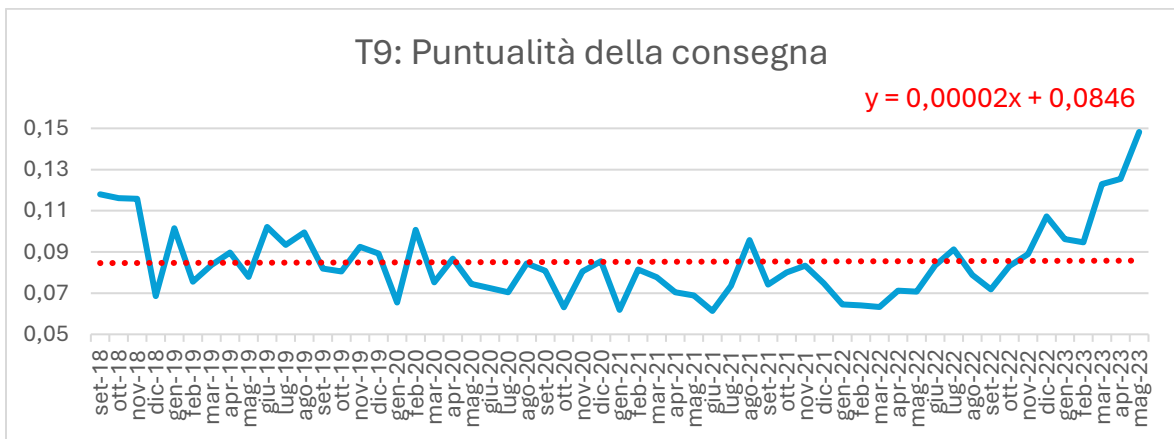
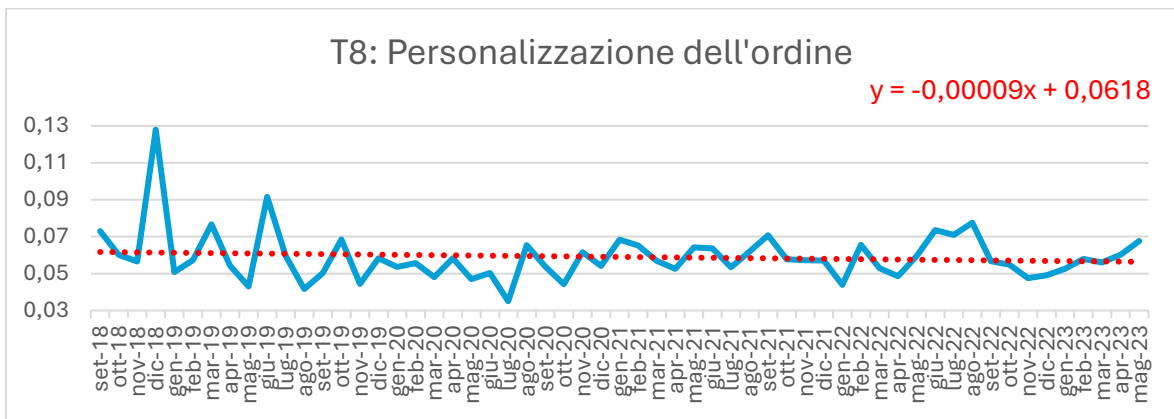
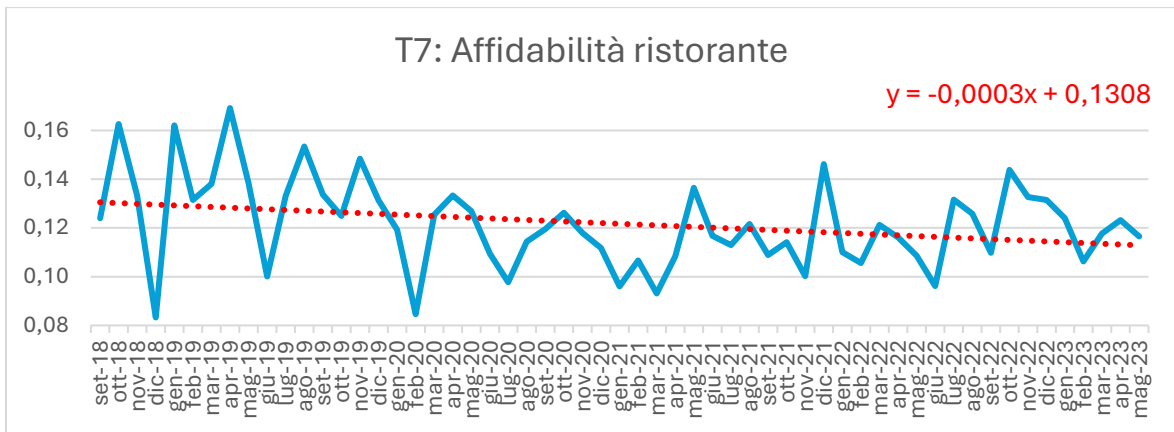
1. Easy to understand how to use it but the order together system is total garbage. Very unclear whether you've actually joined a group order or not when it bounces you back and forth between the app and your internet app version of it. Multiple times when I've joined a group

- order and it turns out that once I order, I've ordered alone instead causing arguments within the group.
2. Very Poor, kindly do not use. I used the app to order food twice. The first time the order which came was not what I had ordered for , they had to engage me in physical exercise of going back to the store and coming back with the correct order, what was the use of using the app? The second time the items in my order weren't fully delivered I had to call and go back again for my goods which were left in the store, what is the use of downloading such an inconvenient app?
 3. Every single time I order something from anywhere on the damn app, my order is wrong or I end up with the wrong order, and nothing changes whether you rate the place or complain
 4. Its one thing restaurants getting orders wrong, but when every order is wrong and the only thing you can do about it is NOTHING except get app credit... getting a family meal is a no go. Deleting this app and i advise everyone else to go direct with takaway companies. Deliveroo you need to make changes. I hope you go bust.
 5. Terrible app. I was enjoying it until they decided to go to a different location give out my food to a random person then say that I will be charged and there's nothing they can do about it. I deleted it and I advice against using this app.

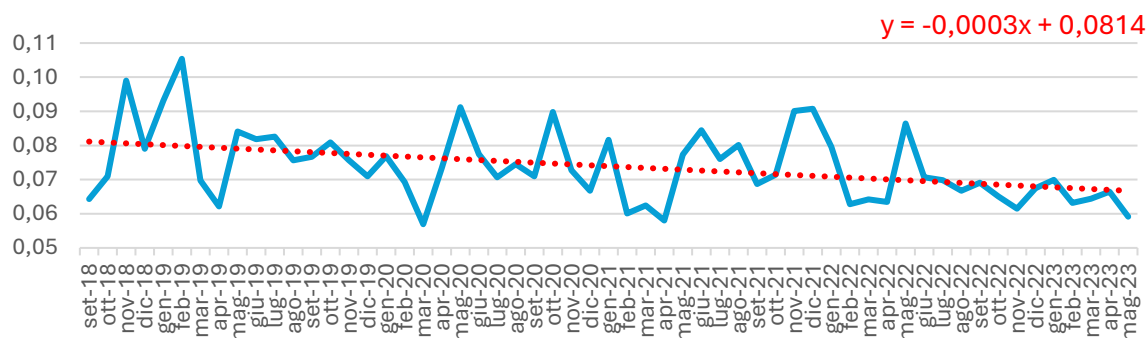
APPENDICE III



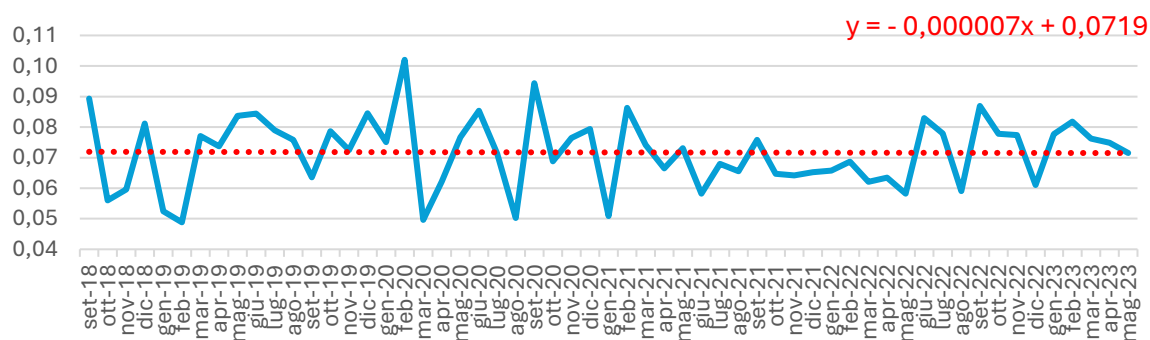




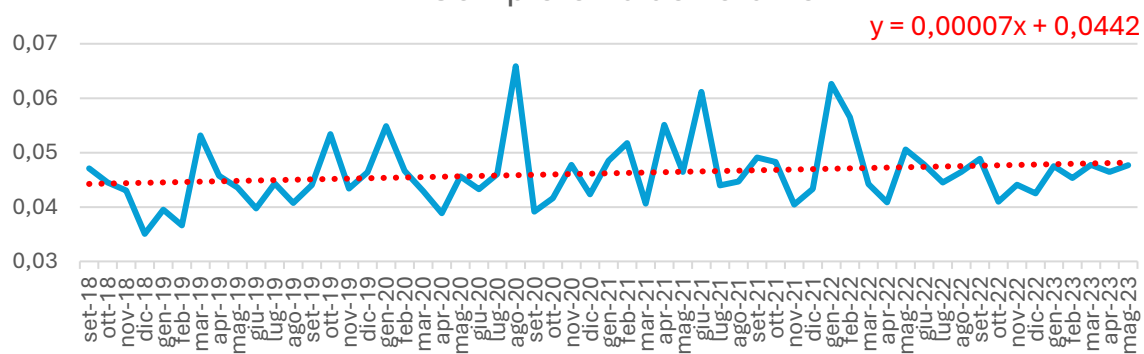
T10: Affidabilità del servizio clienti



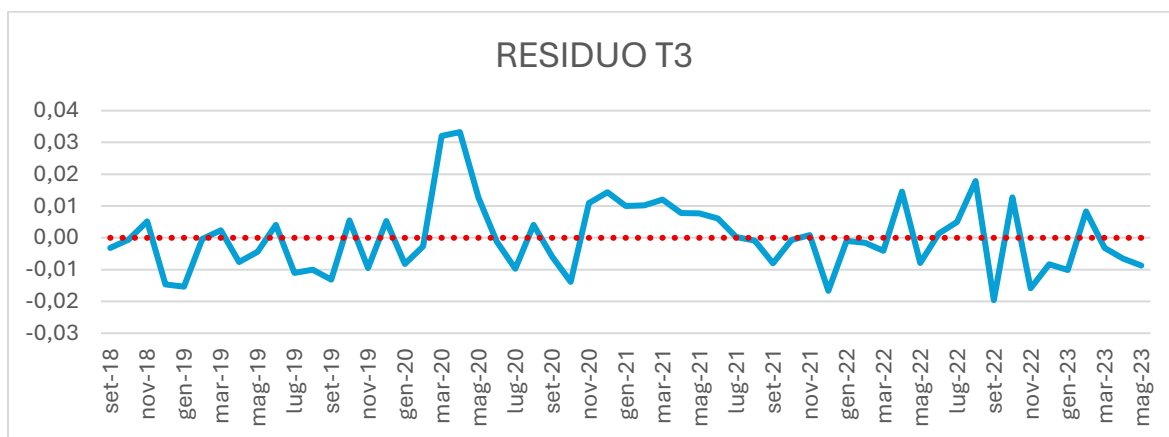
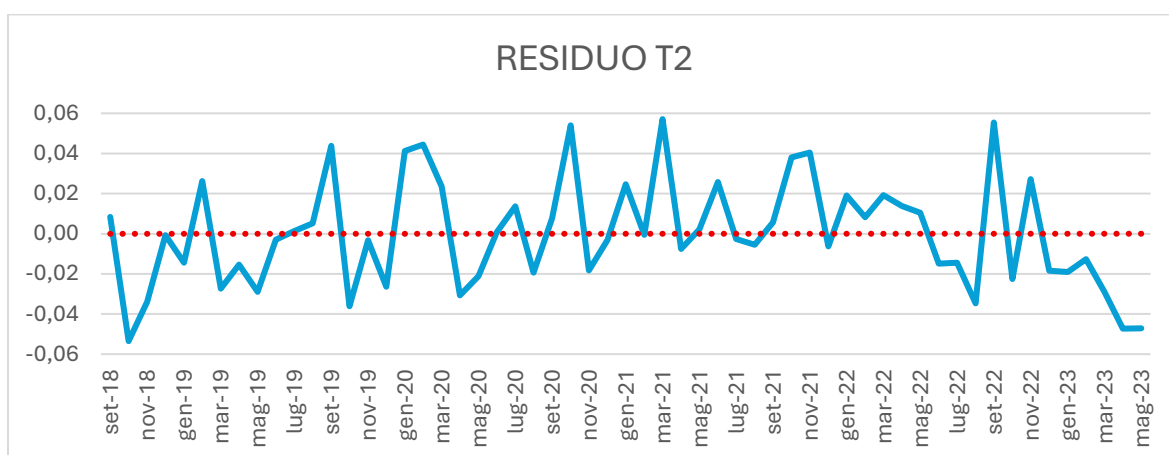
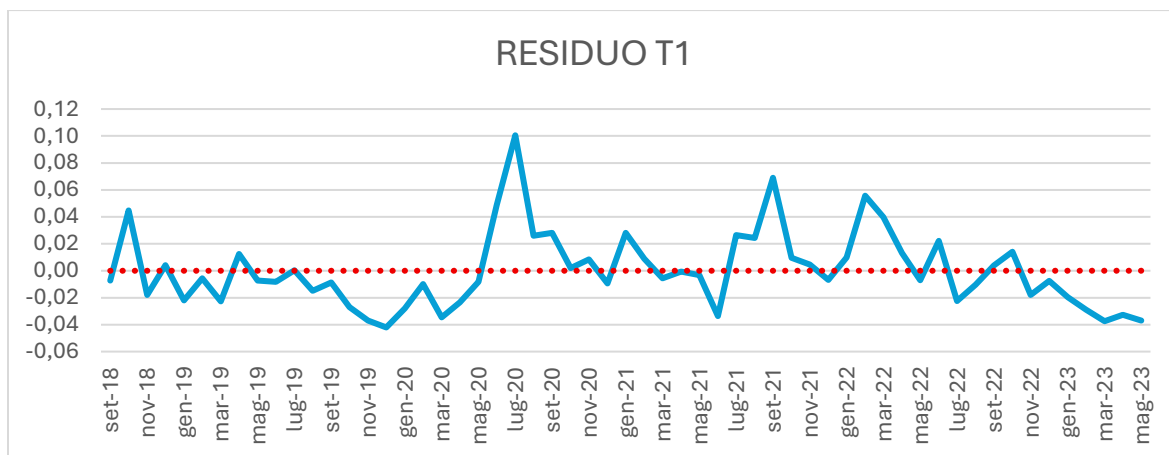
T11: Consegna all'indirizzo voluto e localizzazione

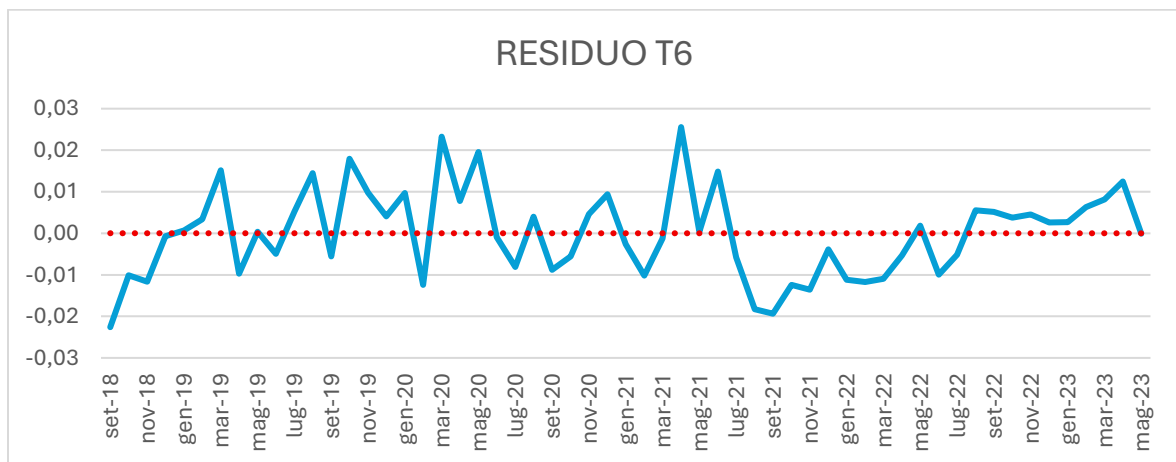
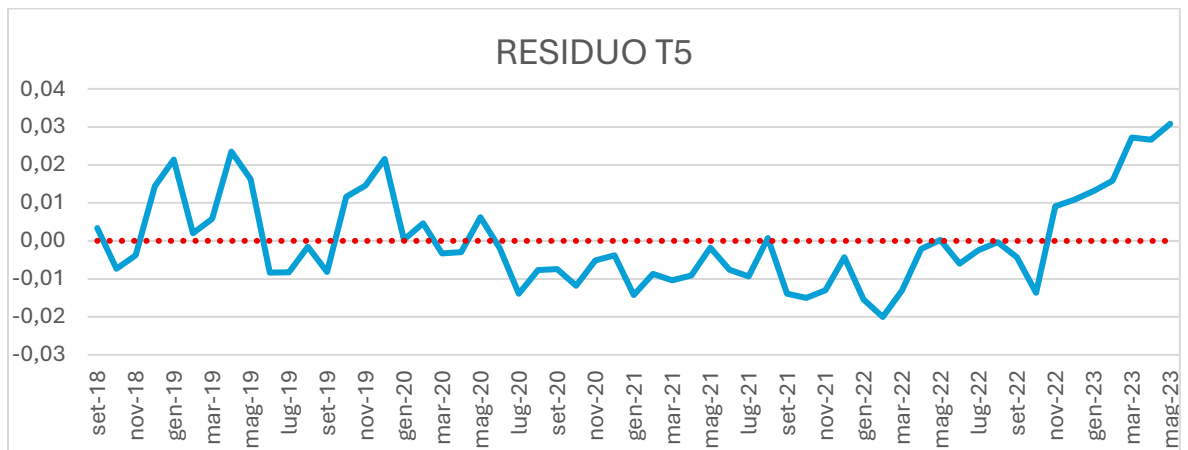
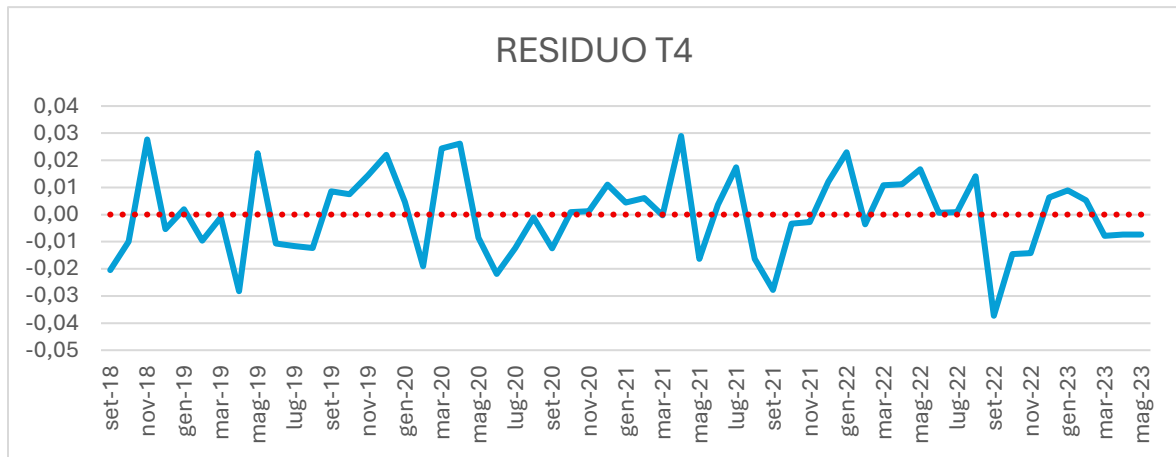


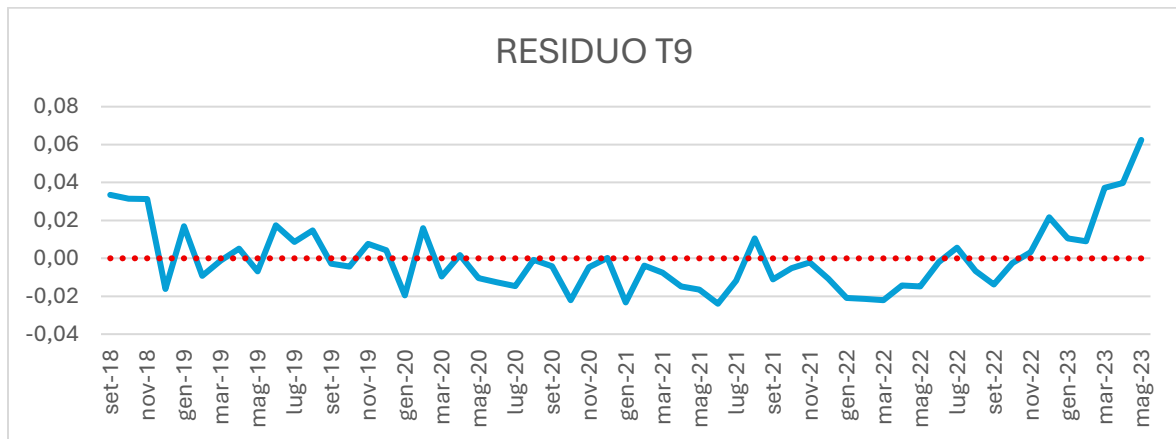
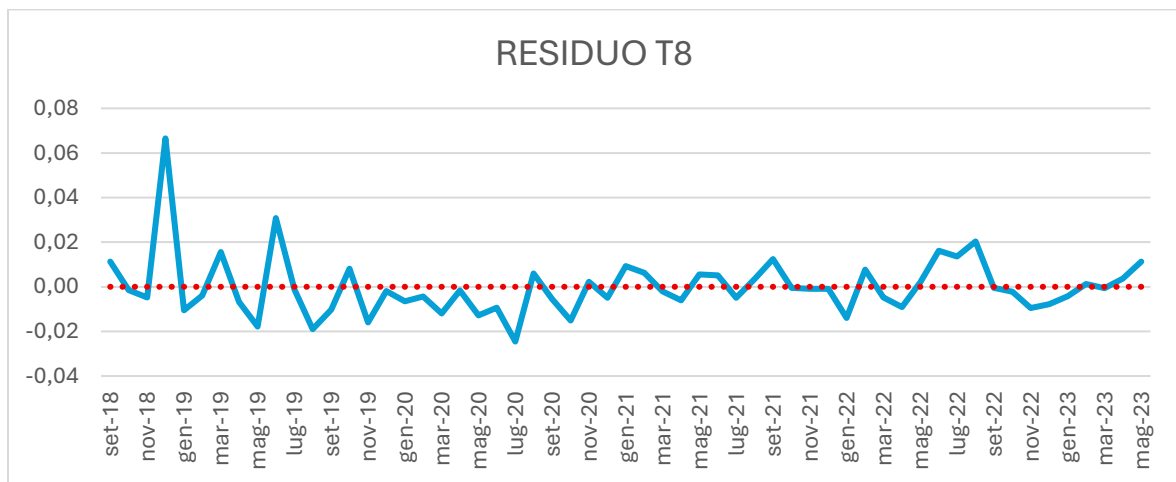
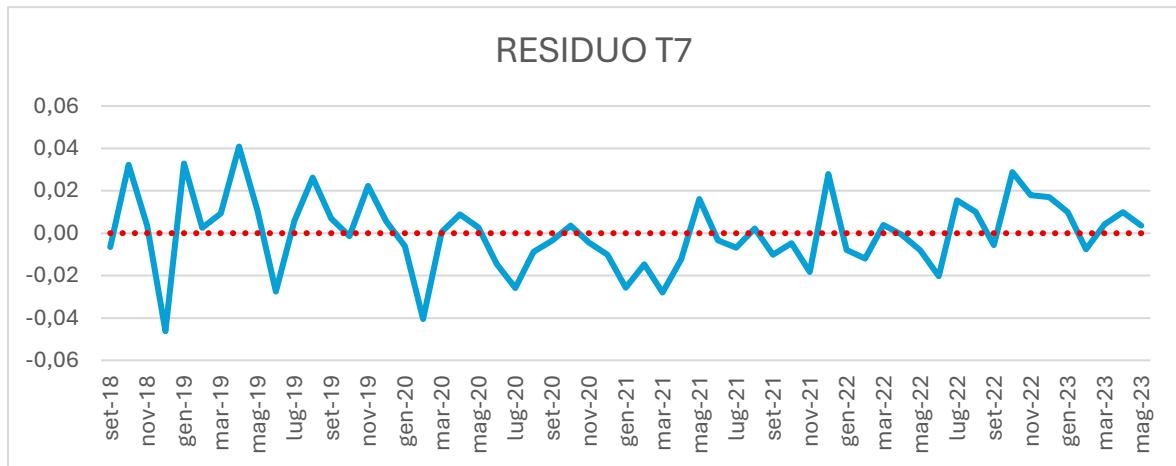
T12: Completezza dell'ordine

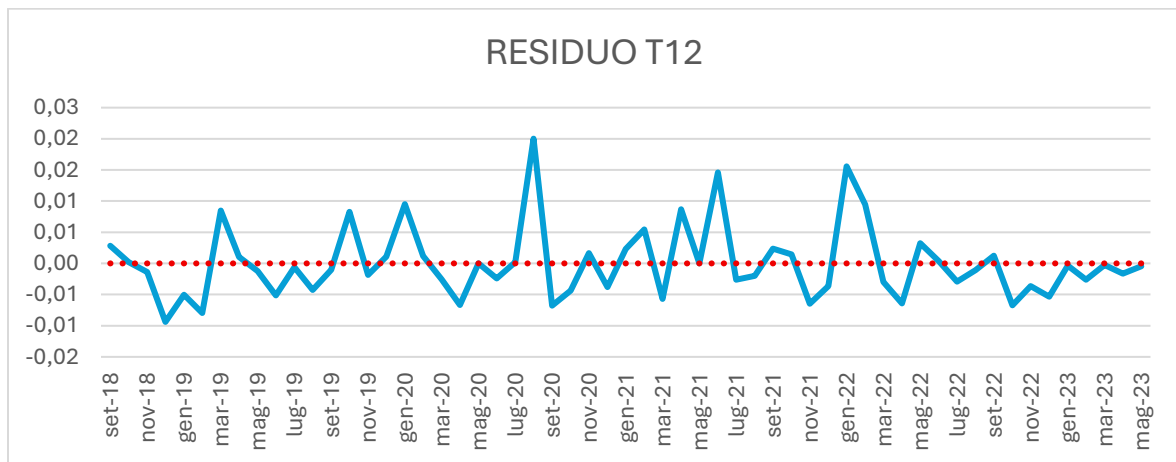
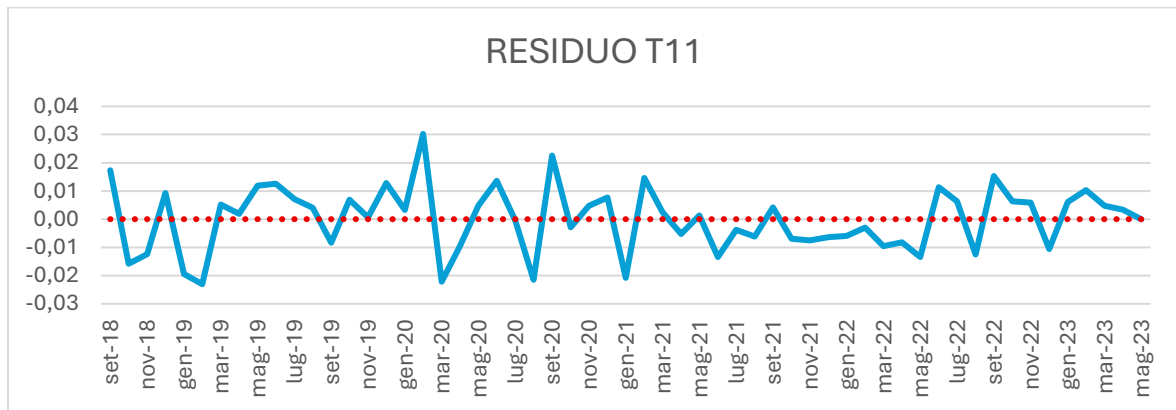
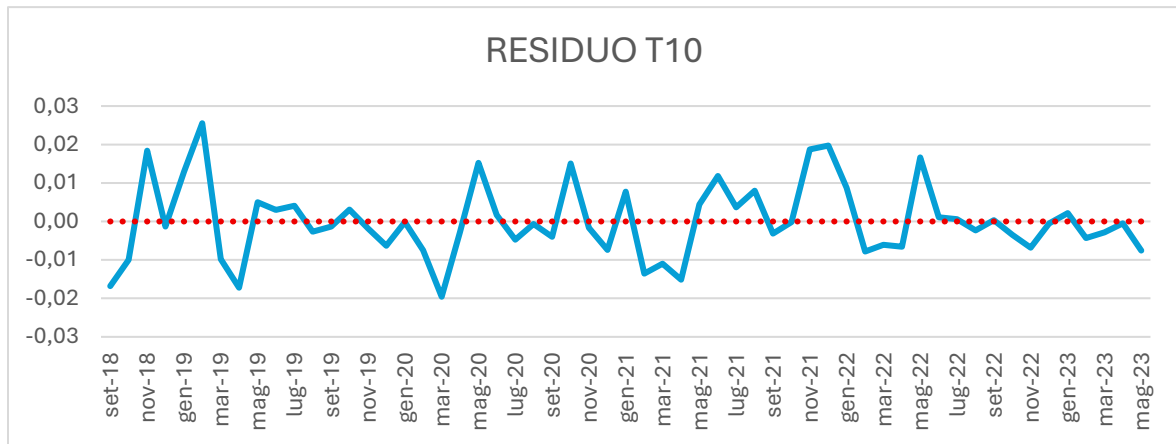


APPENDICE IV









BIBLIOGRAFIA

Anderson, R.E. (1973). *Consumer dissatisfaction: the effect of disconfirmed expectancy on perceived product performance*. Journal of Marketing Research, 10(11), pp. 38–44.

Apify Technologies s.r.o. (2015). Apify. <https://apify.com/>. Ultimo accesso: 11/09/2025.

Ashton, T., Evangelopoulos, N. (2012). *Control charts for customer comments: a case study and a research agenda*. Proceeding of the Southwest Decision Sciences Institute, pp. 661-669.

Ashton, T., Evangelopoulos, N., Prybutok, V. (2014). *Extending monitoring methods to textual data: a research agenda*. Quality & Quantity, 48(4), pp. 2277-2294.

Ashton, T., Evangelopoulos, N., Prybutok, V.R. (2015). *Quantitative quality control from qualitative data: control charts with latent semantic analysis*. Quality & Quantity, 49(3), pp. 1081-1099.

Bandaru, S., Gaur, A., Deb, K., Khare, V., Chougule, R., Bandyopadhyay, P. (2015). *Development, analysis and applications of a quantitative methodology for assessing customer satisfaction using evolutionary optimization*. Applied Soft Computing, 30(1), pp. 265-278.

Barravecchia, F., Mastrogiacomo, L., Franceschini, F. (2020). *Categorizing quality determinants in mining user-generated contents*. Sustainability, 12(23), 9944.

- Barravecchia, F., Mastrogiacomo, L., Franceschini, F. (2021). *Digital voice-of-customer processing by topic modelling algorithms: insights to validate empirical results*. International Journal of Quality & Reliability Management, pp. 1-18.
- Barravecchia, F., Mastrogiacomo, L., Franceschini, F. (2023). *Product quality tracking based on digital Voice-of-Customers*. Total Quality Management & Business Excellence.
- Barravecchia, F., Mastrogiacomo, L., Tavani, L., Franceschini, F. (2022). *Statistical Process Control techniques to monitor quality determinants in digital Voice-of-Customer*. Proceedings of the 5th ICQEM Conference.
- Blei, D.M. (2012). *Probabilistic topic models*. Communication of the ACM, 55(4), pp. 77-84.
- Blei, D.M., Ng, A.Y., Jordan, M.I. (2003). *Latent Dirichlet allocation*. Journal of Machine Learning Research, pp. 993-1022.
- Bolin, G. (2021). *User-Generated Content (UGC): Understanding the Activity of Media Use in the Age of Digital Reproduction*. Digital Roots, pp. 267-279.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J.L., Blei, D.M. (2009). *Reading tea leaves: how humans interpret topic models*. Advances in Neural Information Processing Systems, pp- 288-296.
- Chattopadhyay, P. (2019). *A Study on the Impact of Service Quality on Customer Satisfaction and Customer Loyalty with Reference to Service Marketing Context: Theoretical Approach*. Iconic research and engineering journals, 3(1), pp. 89-96.

Costa, E., Lorena, A., Carvalho, A., Freitas, A. (2007). *A review of performance evaluation measures for hierarchical classifiers*. Evaluation Methods for Machine Learning II: Papers from the AAAI-2007 Workshop, pp. 1-6.

Daubs, M. (2019). *User-Generated Content*. SAGE International Encyclopedia of Mass Media and Society, 5(1), pp. 1825-1827.

Diffbot Technologies Corp. (2008). Diffbot. <https://www.diffbot.com/>. Ultimo accesso: 11/09/2025.

Durrani, N., Schmid, H., Fraser, A., Koenh, P., Schütze, H. (2015). *The operation sequence model-combining N-gram-based and phrase-based statistical machine translation*. Computational Linguistics, 41(2), pp. 185-214.

Franceschini, F., Galetto, M., Maisano, D. (2019). *Designing Performance Measurement Systems*. Management for Professionals, Springer Nature, Cham.

Franceschini, F., Galetto, M., Maisano, D.A., Mastrogiacomo, L. (2020). *Ingegneria della Qualità (Quarta edizione)*. C.L.U.T. Editrice.

Freeman, G., Radziwill, N. (2018). *Voice of the Customer (VoC): A Review of Techniques to Reveal and Prioritize Requirements for Quality*. Journal of Quality Management Systems, Applied Engineering, and Technology Management, 2018(3), pp. 1-29.

Giese, J.L., Cote, J.A. (2002). *Defining Consumer Satisfaction*. Academy of Marketing Science Review, 2000(1), pp. 1-24.

Griffin, A., Hauser, J. (1991). *The Voice of the Customer*. Marketing Science, 12(1), pp. 1-6.

- Growth, M. (2005). *Customers as good soldiers: examining citizenship behaviors in internet service deliveries*. Journal of Management, 31(1), pp. 7-27.
- Guido, G. (2015). *Customer Satisfaction*. Wiley Encyclopedia of Management, 9(117), pp. 1-3.
- Guo, Y., Barnes, S.J., Jia, Q. (2017). *Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent Dirichlet allocation*. Tourism Management, 59, pp. 467-483.
- Heskett, J.L., Sasser, W.E.J., Schlesinger, L.A. (1997). *The service profit chain*. New York.
- Hirschman, A.O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press, 25.
- Ilieska, K. (2013). *Customer Satisfaction index – as a Base for Strategic Marketing Management*. TEM Journal, 2(4), pp. 327-331.
- Jivani, A.G. (2011). *A comparative study of stemming algorithms*. International Journal of Computer Applications in Technology, 2(6), pp. 1930-1938.
- Jones, T.O., Sasser, W.E. (1995). *Why satisfied customers defect*. Harvard Business Review, 73(6), pp. 88-99.
- Kherwa, P., Bansal, P. (2018). *Topic Modeling: A Comprehensive Review*. ICST Transactions on Scalable Information Systems, 7(24), pp. 1-16.
- Lee, J., Park, D.H., Han, I. (2008). *The effect of negative online consumer reviews on product attitude: An information processing view*. Electronic Commerce Research and Applications, 7(3), pp. 341–352.

- Liang, Q., Wang, K. (2019). *Monitoring of user-generated reviews via a sequential reverse joint sentiment-topic model*. Quality and Reliability Engineering International, 35, pp. 1180-1199.
- Lo, S. (2008). *Web service quality control based on text mining using support vector machine*. Expert Systems with Applications, 34(1), pp. 603-610.
- Maria Navin, J.R., Pankaja, R. (2016). *Performance analysis of text classification algorithms using confusion matrix*. International Journal of Engineering and Technical Research (IJETR), pp. 869-2321.
- Mastrogioacomo, L., Barravecchia, F., Franceschini, F., Marimon, F. (2021). *Mining quality determinants of Product-Service Systems from unstructured User-Generated Contents: The case of car-sharing*. Quality Engineering, 33(3), pp. 425–442.
- McDougall, G.H.G., Levesque, T. (2000). *Customer Satisfaction with Services: Putting Perceived Value into the Equation*. Journal of Services Marketing, 15(4), pp. 392-410.
- Meyer, D., Hornik, K., Feinerer, I. (2008). *Text mining infrastructure in R*. Journal of Statistical Software, 25(5), pp. 1-54.
- Montgomery, D.C. (2009). *Introduction to Statistical Quality Control*. 6th ed. John Wiley & Sons Inc.
- Musafir, S. (2017). *Consumer satisfaction and consumer delight*. International Journal of Advanced Research in Management and Social Sciences, 6(5), pp. 328-339.

- Naeem, M., Okafor, S. (2019). *User-Generated Content and Consumer Brand Engagement*. Leveraging Computer-Mediated Marketing Environments, Cap. 9.
- Nobar, H.B.K., Rostamzadeh, R. (2018). *The Impact of Customer Satisfaction, Customer Experience, and Customer Loyalty on Brand Power: Empirical Evidence from Hotel Industry*. Journal of Business Economics and Management, 19(2), pp. 417-430.
- O'Sullivan, D., McCallig, J. (2012). *Customer Satisfaction, Earnings and Firm Value*. European Journal of Marketing, 46, pp. 827-843.
- Octopus Data Inc. (2016). Octoparse. <https://www.octoparse.com/>. Ultimo accesso: 11/09/2025.
- Oliver, R.L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. The McGraw-Hill Companies.
- Oliver, R. L. (1999). *Whence consumer loyalty?*. Journal of Marketing, 63(Special Issue), pp. 33-44.
- Parasuraman, A.P., Zeithaml, V.A., Berry, L.L. (1985). *A Conceptual Model of Service Quality and its Implications for Future Research*. Journal of Marketing, 49(1), pp. 41-50.
- Park, D.H., Lee, J., Han, I. (2007). *The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement*. International Journal of Electronic Commerce, 11(4), pp. 125–148.
- ParseHub Inc. (2015). ParseHub. <https://www.parsehub.com/>. Ultimo accesso: 11/09/2025.
- Reber, A.S. (1985). *The Penguin Dictionary of Psychology*. Penguin Group, p. 257.

Reichheld, F.F. (1995). *The loyalty effect: the satisfaction trap*. Cap. 5, Boston, Bain & Company.

Roberts, M.E., Stewart, B.M., Airoidi, E.M. (2016). *A Model of Text for Experimentation in the Social Sciences*. Journal of the American Statistical Association, 111(515), pp. 988-1003.

Roberts, M., Stewart, B., Tingle, D., Airoidi, E. (2013). *The Structural Topic Model and Applied Social Science*. ICONIP, pp. 1-4.

Roberts, M.E., Stewart, B.M., Tingley, D. (2019). *STM: R package for structural topic models*. Journal of Statistical Software, 91(2), pp. 1-40.

Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., Rand, D.G. (2014). *Structural Topic Models for Open-Ended Survey Responses*. American Journal of Political Science, 58(4), pp. 1064-1082.

Richardson, L. (2004). Beautiful Soup. Crummy, <https://www.crummy.com/software/BeautifulSoup/>. Ultimo accesso: 11/09/2025.

Salur, M.U., Aydin, I. (2020). *A Novel Hybrid Deep Learning Model for Sentiment Classification*. IEEE Access, 8, pp. 58080-58093.

Santos, M.L.B. (2021). *The “so-called” UGC: an updated definition of user-generated content in the age of social media*. Online Information Review, 46(1), pp. 95-113.

Scott, J., Baldridge, J. (2013). *A recursive estimate for the predictive likelihood in a topic model*. Journal of Machine Learning Research, 31, pp. 527-535.

Scrapy developers (2008). Scrapy. <https://www.scrapy.org/>. Ultimo accesso: 11/09/2025.

Šerić, M., Gil-Saura, I. (2019). *Understanding Brand Equity in Hotel Firms. What Is the role of Brand Loyalty and Satisfaction?*. International Journal of Contemporary Hospitality Management, 31(9), pp. 3526-3546.

Sequentum International. (2011). Content Grabber. <https://www.contentgrabber.com/>. Ultimo accesso: 11/09/2025.

Singh, J. (1988). *Consumer complaint intentions and behavior: definitional and taxonomical issues*. Journal of Marketing, 52(1), pp. 93-107.

Sony, M., Antony, J., Douglas, J. A. (2020). *Essential ingredients for the implementation of Quality 4.0*. The TQM Journal, 32(4), pp. 779–793.

Sweeney, J.C., Soutar, G.N, Mazzarol, T. (2008). *Factors influencing word of mouth effectiveness: receiver perspectives*. European Journal of Marketing, 42(3/4), pp. 344-364.

Tirunillai, S., Tellis, G.J. (2012). *Does chatter really matter? Dynamics of user-generated content and stock performance*. Marketing Science, 31(2), pp. 198-215.

Tirunillai, S., Tellis, G.J. (2014). *Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation*. Journal of Marketing Research, 51(4), pp. 463-479.

Vargo, S.L., Lusch, R.F. (2017). *Service-dominant logic 2025*. International Journal of Research in Marketing, 34(1), pp. 46-67.

Vickery, G., Wunsch-Vincent, S. (2010). *Partecipative Web: User-created Content, Web 2.0. Wikis and social networking*, OECD Paris, Cap. 3.

- Villamor, M., Kirsch, D., Prieto-Nañez, F. (2023). *The promise of machine-learning-driven text analysis techniques for historical research: topic modeling and word embedding*. Management & Organizational History, 18(1), pp.1-16.
- Wallach, H.M., Mimno, D.M., McCallum, A. (2009). *Rethinking LDA: Why priors matter*. Advances in Neural Information Processing Systems, 23, pp. 1973-1981.
- Westbrook, R.A., Oliver, R.L. (1991). *The dimensionality of consumption emotion patterns and consumer satisfaction*. Journal of Consumer Research, 18(1), pp. 84–91.
- Western Electric (1956). *Statistical Quality Control Handbook*. Western Electric Corporation, Indianapolis, IN.
- Wirtz, J., Mattila, A. (2001). *Exploring the role of alternative perceived performance measures and needs congruency in the consumer satisfaction process*. Journal of Consumer Psychology, 11(3), pp. 181–192.
- Yang, B., Yu, H., Yu, Y., Liu, M. (2021). *Community experience promotes customer voice: co-creation value perspective*. Marketing Intelligence & Planning, 39(6), pp. 825-841.
- Zaki, M., McColl-Kennedy, J.R. (2020). *Text Mining Analysis Roadmap (TMAR) for Service*, Journal of Services Marketing, 34(1), pp. 30-47.
- Zhang, K.Z.K., Zhao, S.J., Cheung C.M.K, Lee, M.K.O. (2014). *Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model*. Decision Support Systems, 67(1), pp. 78–89.

Zhang, X., Yu, Y., Li, H., Lin, Z. (2016). *Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews*. Online Information Review, 40(1), pp. 119-145.

Zhao, B. (2017). *Web scraping*, Encyclopedia of Big Data, Springer International Publishing, pp. 1-3.