

UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II



Corso di Laurea Magistrale in Ingegneria Informatica

DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELLE TECNOLOGIE
DELL'INFORMAZIONE

ELABORATO SULLA BUSINESS INTELLIGENCE IN

US Restaurant Ranking Analysis

Professore:

Vincenzo Moscato

Candidati:

Benfenati Domenico M63001165

Banchini Federico M63001162

Fusco Camilla M63001180

Vitrano Arianna M63001171

ANNO ACCADEMICO 2021/2022

Primo Semestre

Indice

1	Introduzione	2
1.1	Business Understanding	2
1.2	Data Understanding	3
1.3	Data Preparing	3
1.3.1	Sorgente dei Dati	3
1.3.2	Formato dei dati	3
1.3.3	Modello dei dati	4
2	Analisi	5
2.1	Dashboard 1	5
2.2	Dashboard 2	6
2.3	Dashboard 3	7
2.4	Dashboard 4	8
2.5	Dashboard 5	9

Capitolo 1

Introduzione

Il seguente elaborato si focalizza sulla trattazione dei dati "**Restaurant Ranking 2020**", tramite il software *PowerBI*, con il quale è stata eseguita l'analisi dei dati stessi. Nel primo capitolo viene quindi analizzato il dataset, per capirne le varie features e per aumentare il grado di efficienza dell'analisi stessa. Nel capitolo successivo vengono poi inserite e descritte le varie dashboard che sono state create per l'analisi grafica dei dati di test, e ne viene fornita una descrizione nel dettaglio delle sue componenti.

1.1 Business Understanding

Tale elaborato ha l'obiettivo di analizzare un insieme di dati relativi ai ristoranti e alle catene di ristorazione negli Stati Uniti. I dati sulle ristorazioni provengono dal sito *Restaurant Business Online*, che si occupa di raccogliere e stilare classifiche circa le catene di ristorazione e i ristoranti indipendenti presenti sul territorio americano. L'analisi è stata condotta tramite Power BI, noto servizio di Business Analytics prodotto da Microsoft, utilizzato ampiamente in ambito aziendale per applicazioni di Business Intelligence. L'obiettivo dell'analisi è quello di mostrare le informazioni delle vendite e dei punti vendita dei vari ristoranti, attraverso un'interfaccia interattiva, semplice ed intuitiva. In questo modo si fornisce all'utente una visione chiara della situazione di ogni ristorante e di ogni catena, a seconda dei dati scelti. Al fine di fornire sia una vista panoramica che una vista più dettagliata della situazione, sono state progettate alcune dashboard che analizzano i dati da punti di vista diversi e a livelli di dettaglio diversi. In questo modo si è riusciti ad andare incontro alle varie esigenze che caratterizzano tipologie di utenti differenti.

1.2 Data Understanding

Il dataset è composto da tre tabelle, **Top250**, **Independent100** e **Future50**, sviluppate come segue:

- **Top250** racchiude i dati delle migliori 250 catene di ristoranti americani. Le informazioni rilevanti sono maturate utilizzando una combinazione di archiviazioni finanziarie di società pubbliche, indagini dirette degli operatori, documenti di divulgazione del franchising e algoritmi di valutazione proprietari, sulla base dei risultati dell'anno solare 2019. I dati riguardano la località, le vendite sostenute da quella catena, il numero di punti vendita del ristorante, l'incremento annuo percentuale in termini di vendite e di numero di punti vendita, dall'anno 2019 all'anno 2020, e la categoria di ristorante a cui quella catena fa riferimento;
- **Independent100** contiene i dati dei ristoranti indipendenti con i maggiori incassi. Solo i ristoranti con non più di cinque sedi sono considerati "indipendenti" ai fini di questa lista (anche se è possibile che un ristorante che condivide il nome con una catena ma è gestito separatamente si qualifichi come indipendente). Le classifiche si basano sulle vendite lorde 2019 di cibo e bevande;
- **Future50** è una classifica dei ristoranti con più rapida crescita, che hanno vendite annuali tra i 20 e i 50 milioni di dollari. Le classifiche si basano sul cambiamento percentuale nelle vendite a livello di sistema dal 2018 al 2019; le catene devono anche aver aperto nuove unità nel 2019 per qualificarsi. I conteggi delle unità sono al 31 dicembre 2019.

1.3 Data Preparing

1.3.1 Sorgente dei Dati

I dati sono stati reperiti dal sito www.kaggle.com in maniera statica, dal momento che tale dataset non è soggetto a cambiamenti frequenti. Sono state quindi importate le tabelle all'interno dello strumento di analisi per farne una prima elaborazione.

1.3.2 Formato dei dati

Per garantire la corretta interpretazione dei dati da parte di Power BI è stato necessario definire opportunamente il formato dei dati. Ad esempio:

- in tutte le tabelle è stato necessario ridefinire il nome delle colonne, adattando i nomi alla lingua italiana al fine di rendere la visualizzazione più consona;

- per poter interpretare correttamente i valori di località e stato relativi al posizionamento geografico di stati e città, è stato necessario separare alcune colonne che includevano entrambi in un solo valore, andando a creare due colonne tramite il separatore ',';

1.3.3 Modello dei dati

La modellazione dei dati è avvenuta in maniera indipendente tra le tre tabelle, questo perchè non è rilevante analizzare oltre le tabelle a se, anche eventuali relazioni che è possibile evidenziare tra esse. Si va quindi a definire l'insieme delle tabelle, ognuna delle quali rappresenta un dataset con i relativi attributi. Di seguito il modello dei dati rappresentato in *PowerBI*.

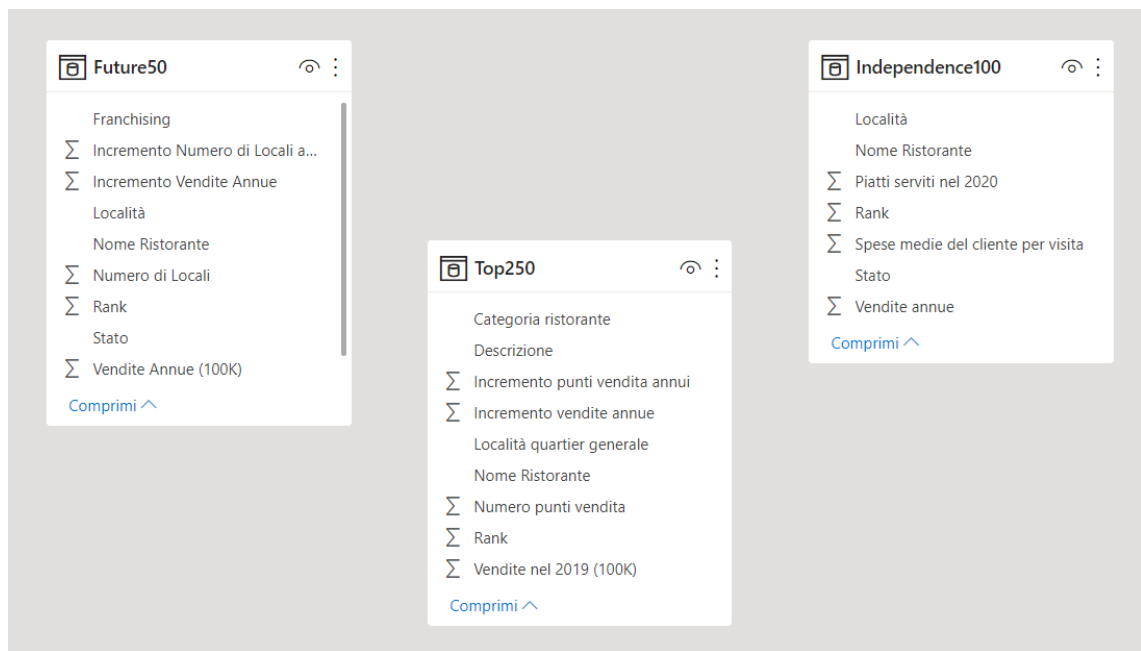


Figura 1.1: Tabelle relative ai dati

Capitolo 2

Analisi

Di seguito vengono elencate le varie dashboard create usando il tool di business intelligence. Ogni schermata ha in comune un *navigator* tra le dashboard, che permette la navigazione tra esse.

2.1 Dashboard 1

La prima dashboard, raffigurata di seguito, è relativa ad un riepilogo generale di alcuni indicatori spiegati in seguito.

Dashboard di Riepilogo Generale

Future 50				
50	1735	1689	168,49	
Ristoranti nella Future 50	Totale numero di locali	Totale vendite annue nel 2019	Totale incremento vendite annue	
Top 250				
250	212519	310685	73,46	
Ristoranti nella Top 250	Totale numero di locali	Totale vendite annue nel 2019	Totale incremento vendite annue	
Independence 100				
100	1918343432		31716666	
Ristoranti nella Independence 100	Totale vendite annue		Totale piatti serviti nel 2020	
Riepilogo generale	Piatti serviti per Città	Vendite e incrementi annui per categoria	Numero di locali e vendite annue	Numero di locali per rank

Figura 2.1: Vista generale sui dati delle tabelle

Andando ad analizzare tutti gli elementi della dashboard, si identificano le seguenti componenti:

- **FUTURE 50**: scheda di riepilogo multipla che mostra il numero totale di ristoranti all'interno della classifica, il totale dei locali, di vendite annue e di incrementi subiti per tutti i ristoranti in classifica;
- **TOP 250**: scheda di riepilogo multipla all'interno della quale troviamo il numero totale di ristoranti tra tutte le top 250 catene di ristoranti, il totale dei punti vendita delle catene, di vendite annue e di incrementi subiti per tutte le catene presenti in classifica;
- **INDEPENDENCE 100**: scheda di riepilogo contenente il numero di ristoranti indipendenti presenti in classifica, il totale delle vendite annue di tali ristoranti e il numero totale di piatti che i ristoranti hanno servito nell'anno 2020.

2.2 Dashboard 2

La seconda dashboard, raffigurata di seguito, è utile per analizzare il numero di piatti serviti in ogni stato e in ogni ristorante. La seconda dashboard è così composta:

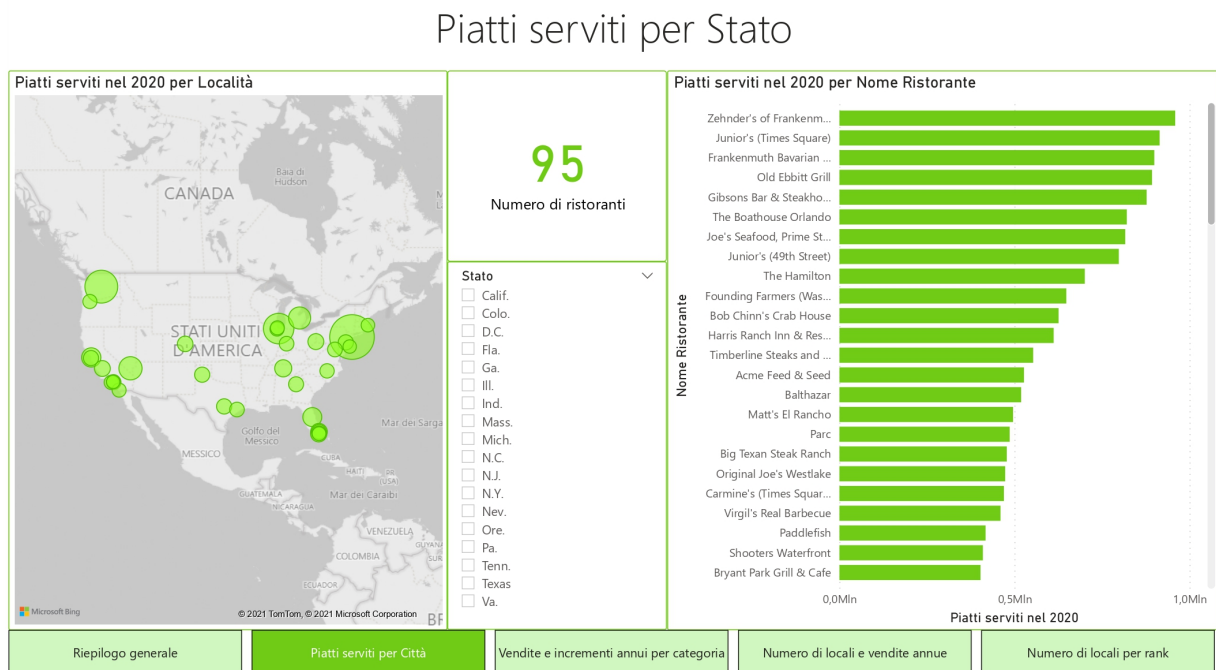


Figura 2.2: Analisi sui piatti venduti per ogni stato e per ogni ristorante

- **Piatti serviti nel 2020 per località**: Mappa a bolle, che indica gli stati dove sono presenti i ristoranti della classifica *Independence 100*. La dimensione delle bolle è data dal numero di piatti distribuiti in quella località.

- **Numero di ristoranti:** scheda di riepilogo che mostra dinamicamente il numero di ristoranti situati nella zona o nello stato selezionato;
- **PIATTI SERVITI NEL 2020 PER NOME RISTORANTE:** istogramma a barre orizzontali che indica in maniera visuale quale ristorante ha servito il maggior numero di piatti.

La dashboard è corredata infine di un filtro che permette di selezionare lo stato americano su cui focalizzare l'analisi.

2.3 Dashboard 3

La terza dashboard è relativa alle vendite e ai loro incrementi, che sono stati registrati nel 2019 da parte delle catene di ristoranti nella top 250.

Dashboard Vendite e incrementi per Categoria

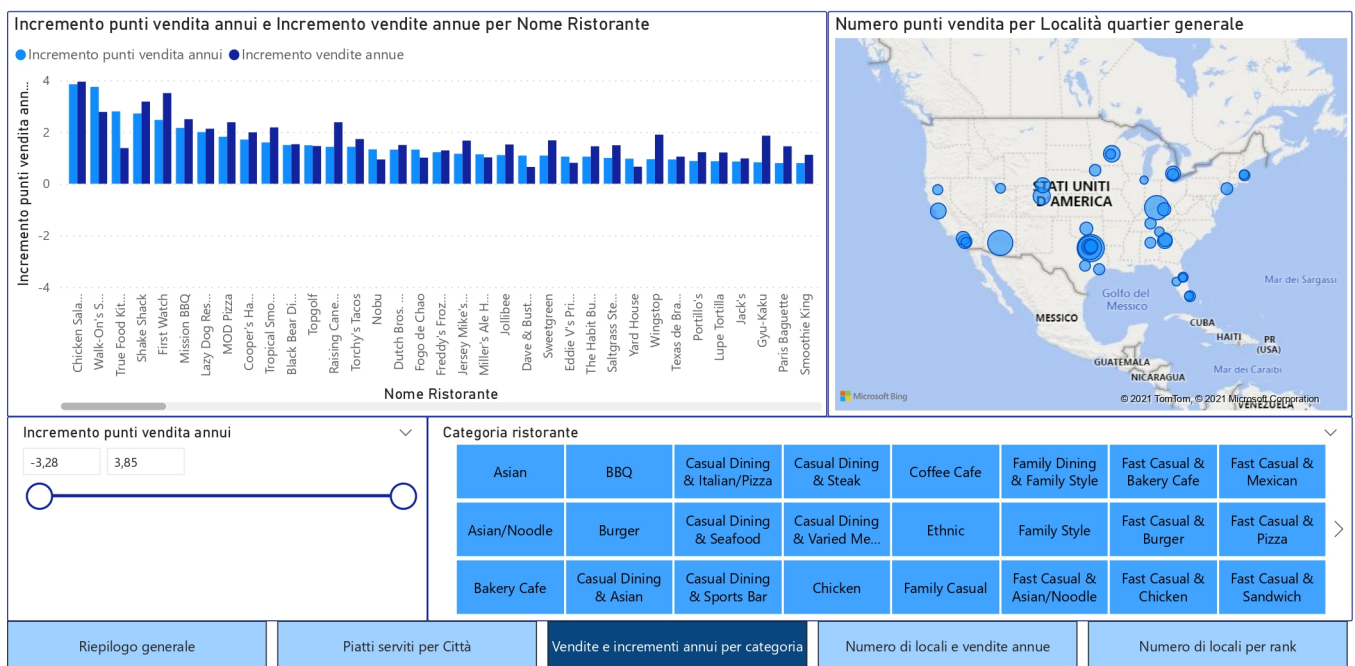


Figura 2.3: Analisi sulle vendite della top 250 ristoranti

Evidenziando gli elementi di tale dashboard si ha:

- **Incremento di vendite e locali per nome ristorante:** istogramma multiplo a barre orizzontali che indica per ogni catena di ristoranti, l'incremento di vendite e di locali che esso ha subito dal 2019 al 202;
- **Numero di punti vendita per località di quartier generale:** mappa a bolle che mostra l'ubicazione dei vari quartier generali delle varie compagnie di ristoranti, dove la

dimensione della bolla indica il numero di punti vendita della catena stessa. Il dataset contiene numerosi locali per cui il quartier generale non viene indicato; ciò viene segnalato con NA all'interno della colonna del quartier generale. Tali catene non è possibile visualizzarle tramite mappa, dal momento che non si dispone della geolocalizzazione delle stesse.

Nella sezione inferiore della vista, sono presenti due tipologie di filtri: il primo permette di selezionare un range tramite uno slider, che evidenzia il valore massimo e minimo di incremento di locali che ha subito la catena, per filtrare i dati all'interno degli altri elementi, mentre il secondo permette di filtrare l'analisi per una data categoria di catena, che sia essa un fast food, una catena di locali familiari, locali etnici, ecc. tramite una pulsantiera.

2.4 Dashboard 4

La quarta vista si occupa di analizzare la tabella dei ristoranti presenti nella Future 50, al fine di visualizzarne vendite e numero di locali per località.

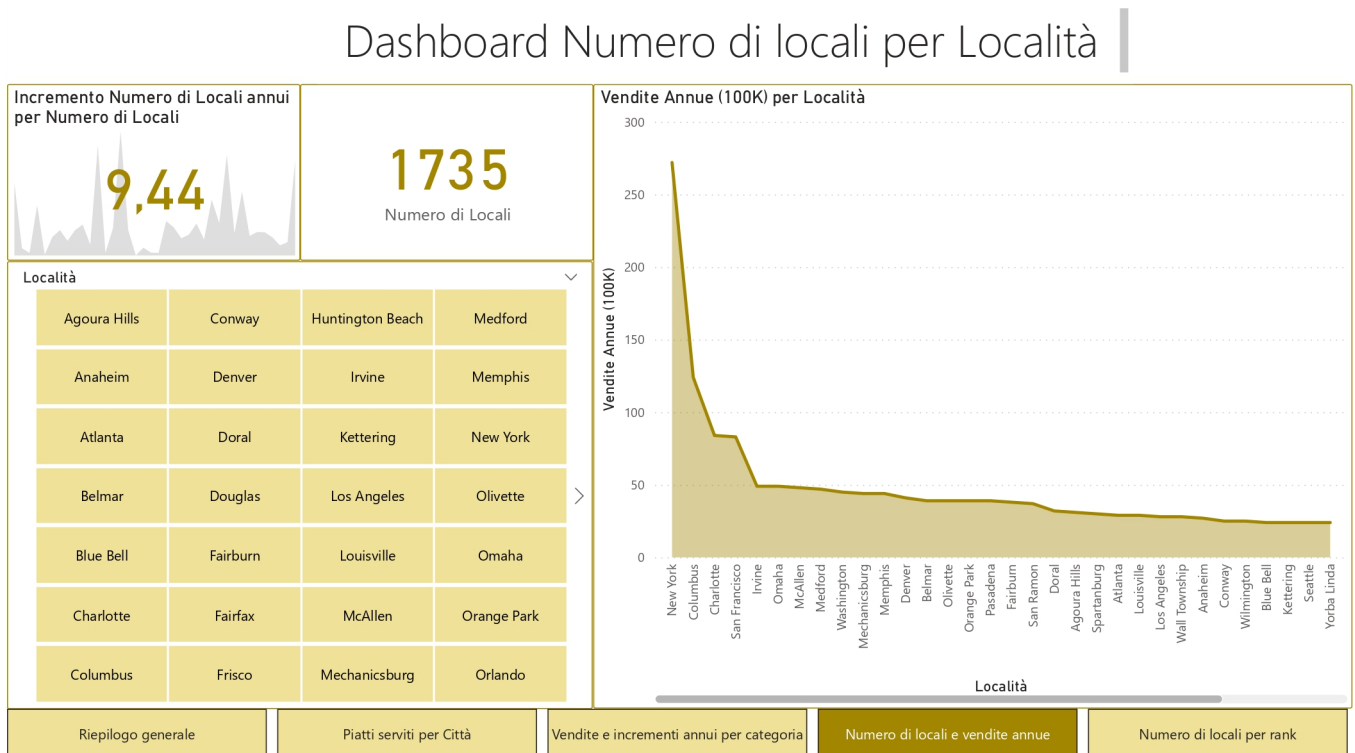


Figura 2.4: Vista su locali e vendite della Future 50 ristoranti

Tale dashboard è composta dai seguenti elementi:

- **Incremento locali per numero di locali:** indicatore di performance per stabilire l'incremento di locali avvenuto nella località selezionata, se non selezionata, l'incremento è mediato su tutte le località;
- **Vendite annue per località:** grafico a linee per rappresentare l'andamento delle vendite annue per località;

Nella dashboard è incluso un filtro a pulsanti per la selezione della località da analizzare nel dettaglio.

2.5 Dashboard 5

La quinta dashboard permette di visualizzare dati relativi alle vendite di ristoranti indipendenti analizzati per posizione in classifica.

Dashboard Numero di locali per Rank

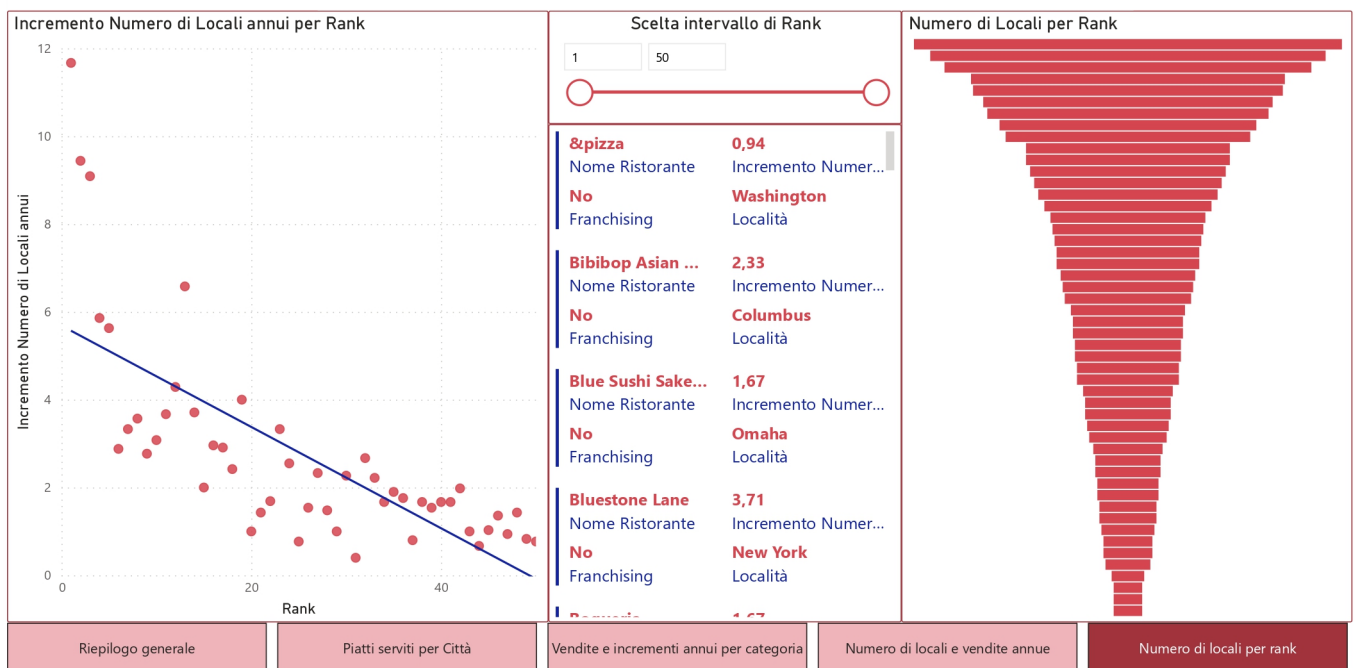


Figura 2.5: Analisi di locali indipendenti per ranking

Analizzando la dashboard, si identificano le seguenti componenti:

- **Incremento locali per rank:** grafico a punti che rappresenta l'incremento di locali che uno specifico ristorante ha subito, rappresentati sulla posizione del ristorante stesso. Tale grafico presenta una trend-line che mostra come l'andamento delle aperture di nuovi locali sia correlato al rank che quel ristorante ha all'interno della classifica: infatti, un

ristorante di rank più alto ha un incremento di numero di locali annui maggiore rispetto ad un ristorante che si trova più giù in classifica;

- **Numero di locali per rank:** grafico ad imbuto che ordina i ristoranti della Future 50, rappresentati dal proprio rank in classifica, in base al numero di punti vendita che essi presentano. Tale grafico permette di stabilire se c'è una relazione tra il numero di locali effettivo e la posizione in classifica del ristorante stesso;

La dashboard si completa di uno slider, che permette di selezionare l'intervallo di rank da analizzare, e di una scheda multipla, che raggruppa le informazioni su incremento di vendite, località e franchising dei ristoranti soggetti all'analisi.