

The Tripbot Restaurant

Duccio Breschi [553550], Giacomo Cerretini [543999], Camilla Poggianti [546747]

ABSTRACT

Questo progetto si colloca in un contesto economico, turistico ma anche culturale, essendo la cucina una parte essenziale della nostra cultura.

Il cibo (in particolare quello italiano) è spesso al centro dell'attenzione comune, perciò numerose sono le raccolte dati, le stesure di articoli e i siti web che guidano l'utente nella scelta del perfetto locale dove consumare un pasto.

Frequentemente tuttavia si tende a concentrarsi sul singolo ristorante senza realizzare un'analisi estesa a tutto il territorio italiano e europeo. Ad esempio è più economico mangiare al nord o al sud? Abbondano più i fast-food o le pizzerie all'estero? È a questo tipo di domande che il nostro sito si propone di rispondere.

INTRODUZIONE

La nostra cucina è molto apprezzata all'estero e viene tramandata da secoli attraverso la vita familiare, per questo è strettamente legata alla nostra cultura.

In generale gli italiani sono molto casarecci, amano mangiare a casa propria o dagli amici, ma quando decidono di consumare un pasto fuori casa molte sono le alternative valide. Come scegliere quella giusta?

Sul web molti sono i siti (fra i quali TripAdvisor, dal quale abbiamo effettuato uno scraping dei dati) con l'obiettivo di indirizzare l'utente alla scelta del ristorante adatto alle proprie esigenze, siti in cui gli utenti stessi hanno modo di confrontarsi e interagire creando recensioni e fornendo un proprio giudizio. Partendo da queste loro preziose opinioni siamo stati in grado di elaborare un quadro più generale sul cibo e i ristoranti italiani e europei, in modo tale da aiutare anche il singolo turista a scegliere la propria meta sulla base delle proprie esigenze culinarie e cercare di capire la distribuzione dei ristoranti in Italia e all'estero sulla base di alcuni aspetti da noi ritenuti fondamentali.

Molti sono anche gli articoli on-line che abbiamo consultato e che mettono in luce le peculiarità della nostra cucina, le differenze fra la cucina nordica e quella sudista e quanto gli altri paesi tentino di emularla contaminandola con le proprie culture. Si tratta tuttavia di articoli ricchi di dati poco comprensibili e osservazioni talvolta anche contrastanti. Scarseggiano su internet grafici che con una sola immagine esprimono in modo molto chiaro un concetto che richiederebbe decine di parole per risultare esaustivo.

The Tripbot Restaurant nasce dal tentativo di unire le varie osservazioni e informazioni tratte dal web con l'obiettivo di indirizzare l'utente non a un ristorante ma in una località sulla base dei propri gusti e delle proprie esigenze.

Ci siamo infatti resi conto di come non esista un sito web in grado di fornirci una visione globale su cosa voglia dire "andare a mangiare fuori" in Italia e in Europa e quali "parametri" tener di conto. Il tutto lo abbiamo fatto ricorrendo a grafici in grado di parlare da soli, che ci hanno permesso di realizzare osservazioni tutt'altro che banali e ci hanno permesso di far luce su come sia oggi il mondo della ristorazione.

Abbiamo organizzato il sito in una serie di domande, in modo da rendere più chiara e logica la sua struttura, collegando diversi aspetti a uno di carattere più generale: sulla base di cosa possiamo decidere in che località andare al ristorante?

Questa disposizione è a nostro parere molto intuitiva, accattivante e riesce ad attrarre l'attenzione dell'utente facendolo immedesimare in ciò che si dice.

Le diverse domande che abbiamo deciso di porci sono distribuite in 3 diverse sezioni:

1. **PRIMA SEZIONE → PREZZO.** Facciamo notare ai nostri utenti quale sia la regione italiana più economica a livello della ristorazione e quale invece quella più dispendiosa. Abbiamo esteso quest'analisi anche livello europeo per valutare invece quale sia il paese più costoso a livello culinario.
2. **SECONDA SEZIONE → QUALITÀ.** Non sempre prezzo alto è sinonimo di alta qualità, per questo abbiamo valutato il rapporto "qualità-prezzo" dei ristoranti prima per ogni regione italiana e successivamente per ogni stato europeo. Sempre in questa sezione abbiamo stilato una classifica dei primi 20 migliori locali italiani, creando un podio per i primi 3 classificati seguiti dai restanti 17 fuori podio. La classifica si basa sul numero delle recensioni e sul rating del ristorante.
3. **TERZA SEZIONE → CATEGORIE.** Dopo aver individuato 6 categorie di cibo da noi ritenute rilevanti (pesce, pizza, fast-food, cibo vegetariano, cibo gluten-free e cibo asiatico) abbiamo ritenuto necessario chiederci a quale di esse i ristoranti si dedichino maggiormente, sia in Italia che in Europa. Accanto quindi a uno scopo più divulgativo, si colloca anche la nostra volontà di sfatare alcuni luoghi comuni quali ad esempio "Italia=pasta e pizza". Successivamente abbiamo effettuato un confronto fra Nord, Sud e Centro Italia, realizzando per ogni categoria presa in considerazione un grafico che mettesse in luce in quale parte d'Italia quella determinata categoria fosse prevalente e perché.

Abbiamo inoltre realizzato un'ulteriore pagina interattiva per i nostri utenti, i quali potranno effettuare le proprie ricerche personalizzate. È possibile scegliere fra 2 tipi di ricerca:

- Ricerca per nome: si seleziona la regione di interesse e il nome del locale per ottenere informazioni su questo
- Ricerca personalizzata: si seleziona la regione di interesse, la categoria di cibo alla quale si è interessati, la fascia di prezzo entro la quale si vuole rientrare e il range di punteggio che il ristorante deve avere per ottenere una lista dei locali corrispondenti a quelle caratteristiche.

STATO DELL'ARTE

Come detto in precedenza, la cucina si trova alla base della nostra cultura perciò in rete sono presenti molti siti e svariati articoli che trattano l'argomento "cibo", ma ognuno di essi lo affronta in modo diverso. C'è il sito che valuta i singoli ristoranti, l'articolo che parla dell'incremento del cibo vegetariano, l'influenza del cibo italiano su quello europeo ecc ecc...

Fra i siti più noti nell'ambito della ristorazione non possiamo non citare **TripAdvisor** (<https://www.tripadvisor.it/Restaurants>), un sito più a stampo blog giornalistico/informativo in quanto comprende recensioni sui diversi locali e prevede l'espressione di un proprio

voto. Quello che manca è una visione globale della ristorazione italiana e europea, caratteristica peculiare invece del nostro sito. *The Tripbot Restaurant* si propone infatti non di valutare il singolo ristorante ma realizzare osservazioni e riportare statistiche a livello regionale e europeo sui ristoranti in base ad alcuni parametri ritenuti da noi fondamentali (prezzo, qualità e categorie di cibo).

Un altro sito interessante che abbiamo consultato è **Il consumo di cibo in Italia in tempo reale**

(<https://www.vegolosi.it/news/ristoranti-vegetariani-e-vegani-in-italia-sempre-piu-numerosi/>), un sito che ci aggiorna in tempo reale sul consumo di cibi e bevande in Italia, calcolando di secondo in secondo la spesa dei nostri connazionali. Interessante è osservare come il consumo di alcuni prodotti aumenti molto più velocemente rispetto ad altri ma è difficile stimare a colpo d'occhio il consumo di alimenti che si trovano in una fascia intermedia. Attratti da questo tipo di studio, abbiamo raccolto dei dati prendendo in considerazione non i singoli alimenti ma più in generale delle categorie di cibo e abbiamo stimato, attraverso dei grafici, qual è la categoria prevalente in Italia e in generale in Europa.

A nostro parere abbiamo reso più user-friendly l'idea trattata in questo sito da noi consultato, traendo interessanti osservazioni su quelle che sono le maggiori richieste dei clienti nell'ambito della ristorazione europea e italiana.

Quello che ci interessava a livello italiano era anche effettuare un confronto fra la cucina nordica, quella del centro e quella sudista. Un sito che ci è risultato utile è stato **La cucina italiana** (<https://kairos1.wordpress.com/>). Qui sono stati messi in evidenza quelli che sono i prodotti tipici locali e una serie di dati che noi abbiamo ritenuto più opportuno e comprensibile visualizzare mediante dei grafici. In questo sito ogni pagina è dedicata alla cucina di una parte dell'Italia (nord, centro o sud che sia) rendendo difficile il confronto immediato fra le 3. Quello che abbiamo cercato di fare nel nostro progetto è documentarci sulle varie differenze fra le cucine del nord, del centro e del sud Italia, raccogliere dei dati e elaborare dei confronti su uno stesso grafico per ottenere una visione d'insieme.

MODELLO DEI DATI

1. Selezione ed estrazione dei dati

Tutti i dati sui quali abbiamo basato le nostre analisi e le nostre osservazioni sono stati estratti dal sito di TripAdvisor (<https://www.tripadvisor.it/Restaurants>). L'estrazione di questi è stata resa possibile grazie all'utilizzo di **Scrapy**, un framework gratuito e open source scritto in Python usato per il web crawling e le attività di scraping dei dati. Con Scrapy è possibile realizzare il proprio scraping tool e i risultati ottenuti possono essere esportati in comodi file formato JSON o CSV.

Partendo da un elenco di URL delle regioni italiane e degli stati europei (costruito manualmente), abbiamo creato un **web crawler** per realizzare un elenco di URL delle città (sia italiane che europee) e, da questo, abbiamo realizzato un elenco di URL per i ristoranti di ciascuna città (italiana ed europea). Da qui, per ogni città, abbiamo estratto i dati relativi ai suoi ristoranti. Nella cartella del progetto sono presenti i codici dello scraping.

Esempio pratico:

URL della regione ricavato manualmente:

[TripAdvisor LLC \[US\] | https://www.tripadvisor.it/Restaurants-g187893-Tuscany.html](https://www.tripadvisor.it/Restaurants-g187893-Tuscany.html)

Attraverso il web crawler abbiamo estratto le URL di ogni città, ad es. Firenze per la Toscana:

[TripAdvisor LLC \[US\] | https://www.tripadvisor.it/Restaurants-g187895-Florence_Tuscany.html](https://www.tripadvisor.it/Restaurants-g187895-Florence_Tuscany.html)

Dall'URL della città abbiamo avuto accesso alla lista di tutti i ristoranti di questa e abbiamo estratto la stringa delle categorie attive per quel ristorante

Attraverso il web crawler abbiamo estratto le URL dei ristoranti per ciascuna città, ad es. "Gustarium" per Firenze:

[TripAdvisor LLC \[US\] | https://www.tripadvisor.it/Restaurant_Review-g187895-d14028759-Reviews-Gustarium-Florence_Tuscany.html](https://www.tripadvisor.it/Restaurant_Review-g187895-d14028759-Reviews-Gustarium-Florence_Tuscany.html)

Dall'URL dei ristoranti abbiamo avuto accesso ai seguenti dati per ogni ristorante

Dettagli

5,0 251 recensioni

Aperto ora 12:00 - 21:00 Orari

CUCINE Pizza, Mediterranea, Italia... **PREZZO** €

PUNTEGGI Servizio Cucina Qualità/prezzo

LOCALITÀ Via Dei Cimatori 24/r, 50122, Firenze, Italia

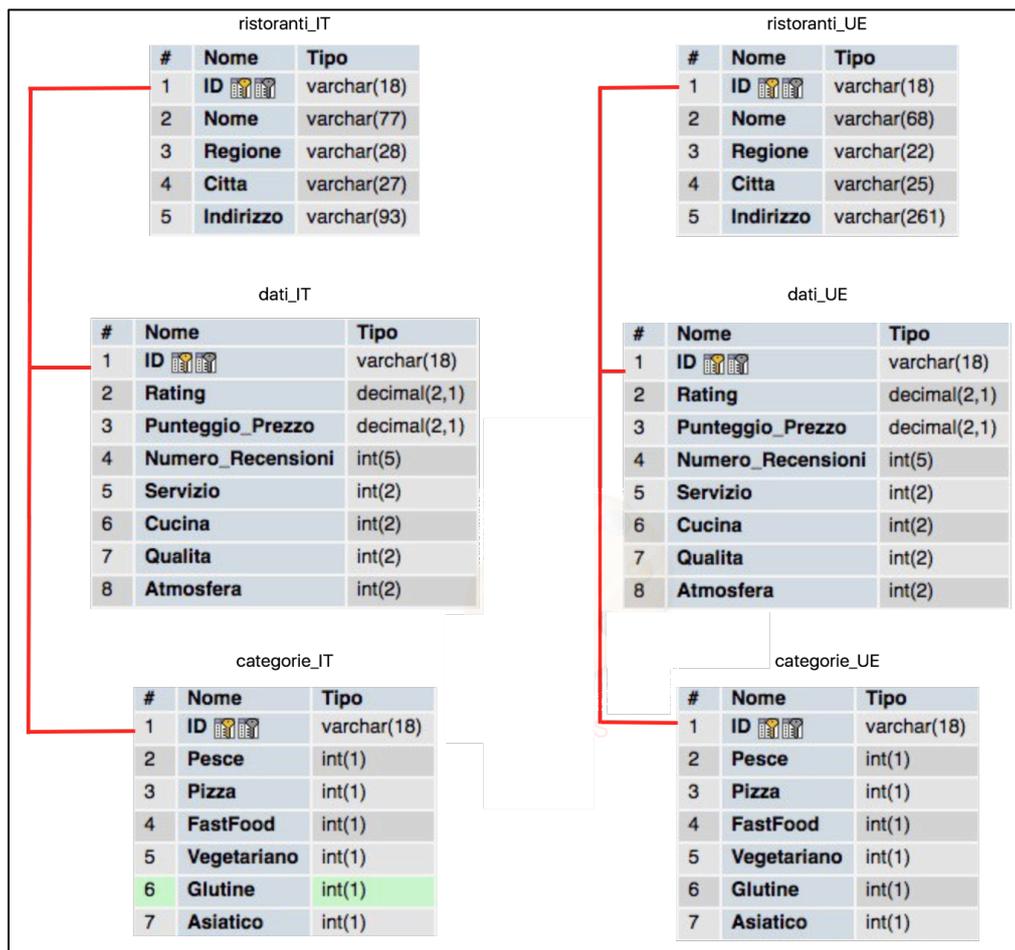
Tutti i dettagli | Migliorate questo profilo

2. Manipolazione dei dati

Una volta scaricati, i dati sono stati elaborati su Excel per la creazione di tabelle che poi sono state importate su *phpMyAdmin*. Il nostro Database contiene le seguenti 6 tabelle da noi create:

- **ristoranti_IT** e **ristoranti_UE**: contenenti rispettivamente i dati relativi a tutti i ristoranti italiani presenti su TripAdvisor e ai ristoranti europei (prendendo in considerazione le città più importanti per stato). Per ogni ristorante abbiamo specificato: l'ID (che è la primary-key), il nome del ristorante, la regione (o stato) nel quale si trova, la città e l'indirizzo specifico.
- **dati_IT** e **dati_UE**: contenenti rispettivamente una serie di punteggi riguardanti diversi aspetti per ogni ristorante italiano presente su TripAdvisor e per ogni ristorante europeo (prendendo in considerazione le città più importanti per stato). Per ogni ristorante abbiamo specificato: l'ID (che è la primary-key), il rating ovvero il voto complessivo assegnato dagli utenti TripAdvisor alla qualità del ristorante (i valori assunti da questo campo variano in un range da 1 a 5), il punteggio dato dagli utenti alla spesa nei vari ristoranti (1.0 se la spesa è bassa, 2.5 se la spesa è media e 4.0 se la spesa è alta), il numero delle recensioni per ogni ristorante, il punteggio assegnato dagli utenti al servizio, alla cucina, al rapporto qualità-prezzo e all'atmosfera (i valori assunti da questi campi variano in un range da 0 a 50).
- **categorie_IT** e **categorie_UE**: si tratta di 2 tabelle che mettono in evidenza quali sono le categorie attive per ogni ristorante. In modo particolare abbiamo tenuto di conto le seguenti 6 categorie: pesce, pizza, fast-food, cibo vegetariano, cibo senza glutine e cibo asiatico. Se un ristorante risulta attivo per una determinata categoria (ovvero possiede nel proprio menù quella pietanza), la categoria assumerà come valore 1, altrimenti 0. In più per ogni ristorante abbiamo specificato l'ID (che è la primary-key).

3. Schema del database



ANALISI DEI DATI

Per rendere le nostre osservazioni il più veritiere e attendibili possibili abbiamo normalizzato i nostri dati andando ogni volta a tenere di conto il numero totale di ristoranti per ogni regione italiana o per ogni stato europeo. In questo modo abbiamo ottenuto dei risultati che tengono conto della densità di locali per ogni regione o stato.

Analizzando i dati raccolti, siamo in grado di sostenere le seguenti tesi:

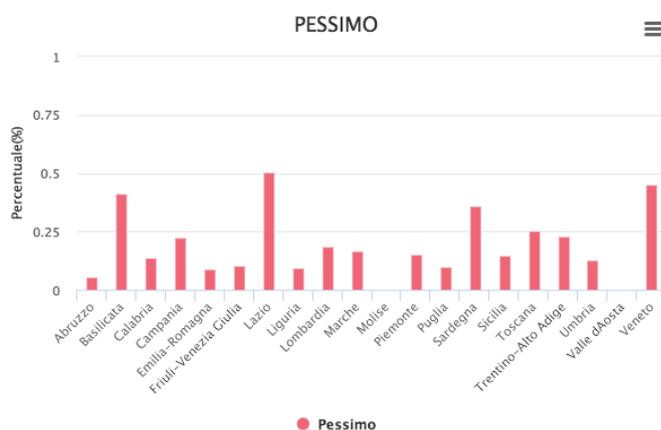
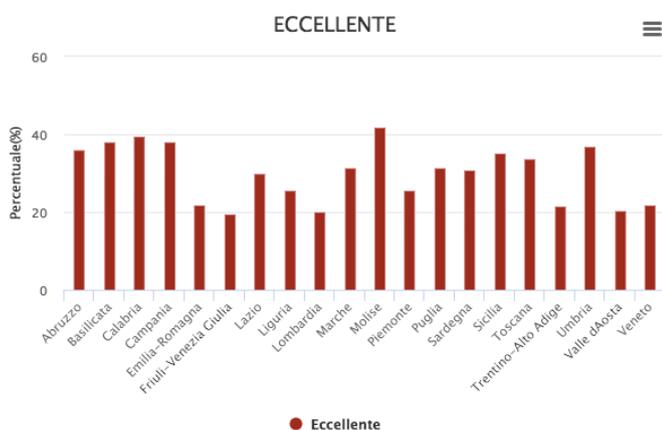
- 1) Le persone ritengono che **nella maggior parte dei ristoranti italiani la spesa sia adeguata rispetto alla quantità del consumo**, solo una piccola percentuale di ristoranti richiedono una spesa eccessiva e una modesta percentuale di locali ha un basso costo rispetto al pasto consumato. Nello specifico si deduce che la **Sicilia** sia la regione **più economica** a livello della ristorazione, mentre la **Lombardia** quella **più dispendiosa**.

Anche a livello europeo gli utenti ritengono che **nella maggior parte dei ristoranti la spesa sia adeguata** rispetto alla quantità del

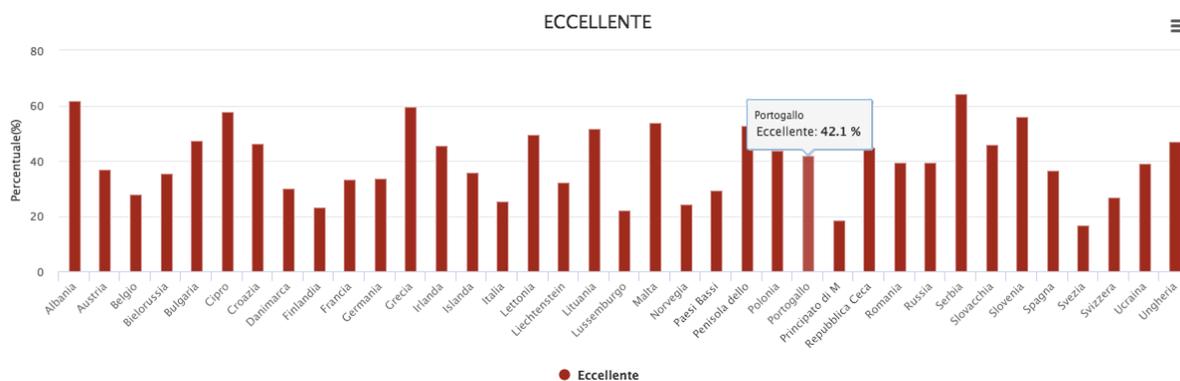
consumo ma si notano maggiori dislivelli rispetto ai risultati italiani. Si osservano diversi stati con una **più alta percentuale di ristoranti costosi**. In particolar modo spicca il **Principato di Monaco** con il maggior numero di **ristoranti costosi**, mentre la **Polonia** spicca per l'**economicità** dei suoi locali.

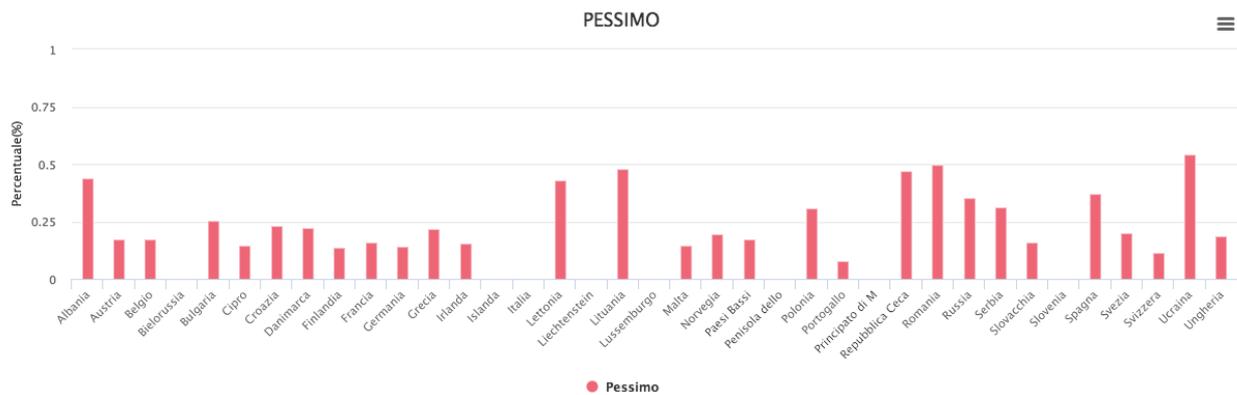


2) Non sempre prezzo alto è sinonimo di alta qualità. La **minor percentuale** dei ristoranti con un'**eccellente rapporto qualità-prezzo** si trovano in **Lazio**, mentre il **miglior rapporto qualità-prezzo** si ha in **Molise**. Infatti nessun ristorante in Molise (ma anche in Valle D'Aosta) ha ristoranti con un pessimo rapporto qualità-prezzo. Abbiamo visualizzato questi dati su questi 2 grafici: nel primo analizziamo la percentuale dei ristoranti con un eccellente rapporto qualità-prezzo (quelli a cui gli utenti TripAdvisor hanno attribuito un punteggio compreso fra 45 e 50), mentre nel secondo la percentuale dei ristoranti con un pessimo rapporto qualità-prezzo (quelli a cui gli utenti TripAdvisor hanno attribuito un punteggio compreso fra 10 e 15).



A livello europeo invece lo stato con la maggior percentuale di ristoranti aventi un **rapporto qualità-prezzo eccellente** è la **Serbia**, mentre il **Principato di Monaco** presenta una percentuale bassa di ristoranti collocabili in questa fascia, non a caso abbiamo riscontrato come si tratti dello stato europeo più caro a livello della ristorazione. Inoltre ben 8 stati italiani su 30 non hanno ristoranti con un pessimo rapporto qualità-prezzo.

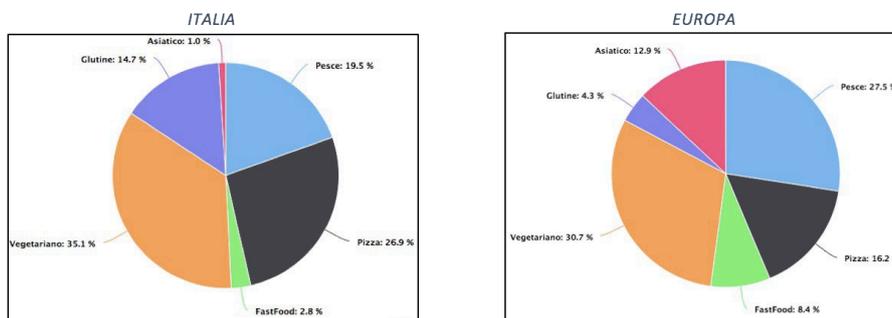




3) Realizzando una top 20 dei ristoranti italiani si deduce che il **miglior locale italiano** secondo gli utenti TripAdvisor (tenendo di conto il rating e il numero di recensioni) si trovi a **Firenze**.



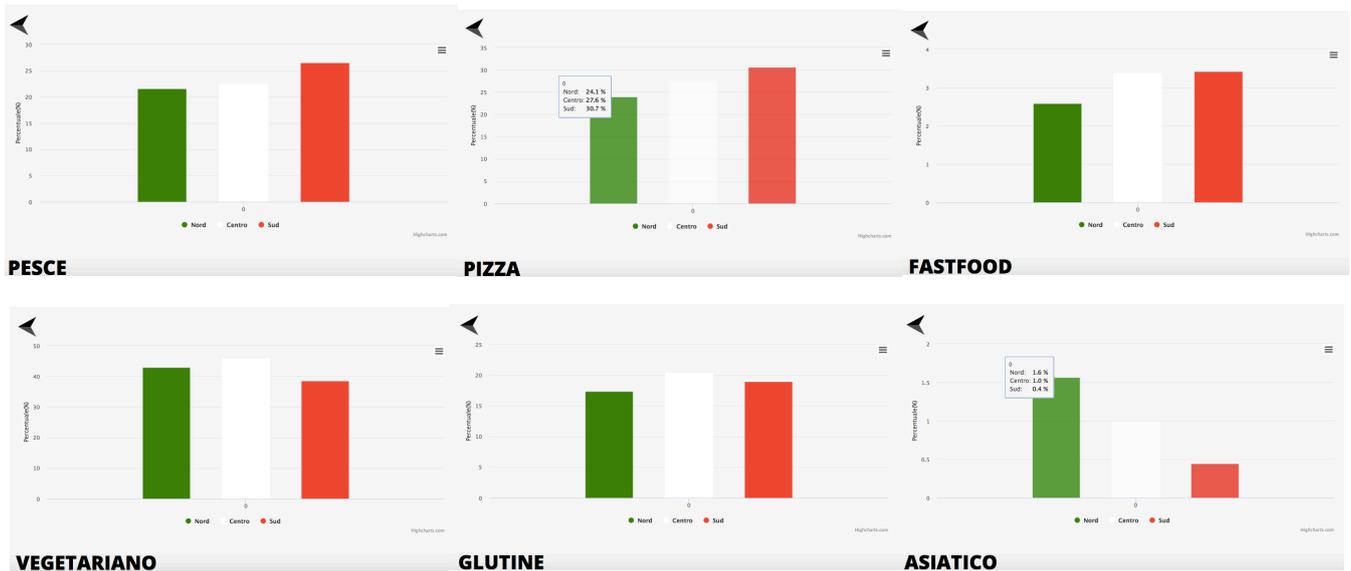
4) La **maggior parte dei ristoranti italiani e europei** si dedicano alla **cucina vegetariana**, ciò suggerisce che sta crescendo la ristorazione bio e le persone cercano il cibo sano.



5) Analizzando categoria per categoria abbiamo effettuato dei confronti fra Nord, Sud e Centro Italia arrivando a sostenere le seguenti tesi:

- La maggior parte dei **fast-food** si trovano al **Sud**, quasi ex aequo col Centro mentre la minor percentuale si trova al Nord. Probabilmente questo dipende dalla maggior economicità di questo prodotto, più accessibile da quella parte di popolazione con un reddito più basso.
- La percentuale più alta di ristoranti che servono **cibo asiatico** si registra al **Nord** e questo può essere giustificato col maggior attaccamento alla tradizione da parte del Sud e l'essere più inclini alle novità e alla contaminazione delle culture del Nord.
- Al **Centro** si ha la maggior concentrazione di ristoranti **"gluten-free"**, non a caso fonti attestano come la Toscana e la Sardegna siano particolarmente interessati alla questione.
- Il maggior numero di ristoranti che cucinano **prodotti ittici** si ha al **Sud** molto probabilmente per la richiesta dei turisti e la maggior disponibilità del pescato. Una buona percentuale però si ha anche al Nord, che è infatti una grande piattaforma di arrivo e smercio del pesce.
- La più alta densità di ristoranti **vegetariani** si registra al **Centro**, subito seguito dal Nord.

- I ristoranti che servono **pizza** sono piuttosto **ben distribuiti** in Italia, senza grandi dislivelli fra Nord, Centro e Sud, non ha caso è uno degli alimenti più tipici del nostro Paese.



CONCLUSIONI E POSSIBILI SVILUPPI

The Tripbot Restaurant consente di ottenere una visione d'insieme sulla ristorazione italiana ed europea, soffermandosi su parametri fondamentali come prezzo, qualità e categoria di cibo. Utilizzando i grafici abbiamo permesso in modo più chiaro all'utente di confermare o falsificare alcuni luoghi comuni relativi ai temi trattati.

Attraverso questo sito le persone potranno informarsi su alcuni aspetti riguardanti la ristorazione italiana ed europea, stabilendo delle mete da visitare anche in funzione delle osservazioni che sono riusciti a trarre. Il nostro sito è quindi una guida alla ristorazione in Italia ma anche più generalmente in Europa.

Possibili sviluppi di *The Tripbot Restaurant* potrebbero essere:

- **Estendere l'analisi a livello mondiale**, un approfondimento sicuramente interessante in quanto metterebbe a confronto delle culture molto diverse fra loro.
- **Condurre uno studio sull'asse temporale** e osservare come negli ultimi decenni è cambiata la ristorazione mondiale elaborando anche possibili congetture per il futuro. Un campo sul quale varrebbe la pena insistere è quello riguardante la ristorazione vegetariana e vegana che, come abbiamo avuto modo di osservare, oggi ha assunto una posizione rilevante e probabilmente aumenterà in futuro.
- **Aumentare i dati estratti da TripAdvisor**, ad esempio prendere in considerazione le coordinate geografiche dei locali per realizzare una mappa che mostri la concentrazione dei ristoranti nelle varie città. Da qui sarebbe anche interessante valutare la variazione dei prezzi e dei voti in base alla densità dei ristoranti, facendo luce sul fenomeno della competizione fra attività.
- **Realizzare una versione del sito multilingue** per permettere la comprensione anche a utenti di altri paesi.