Raccolta e analisi dei requisiti per un progetto di analisi dati

Progettazione di

Dovevuoidormire.it

Neli Staykova & Giuseppe Schillaci

Epicode Day1 Exe 1

## Descrizione del caso di studio

La piattaforma online Dovevuoidormire.it permette agli utenti di cercare delle strutture ricettive come appartamenti, ville o hotel in Italia. Gli utenti possono prenotare un alloggio sul territorio italiano in specifiche date scelte dall’utente.

Nella base dati sono presenti 10.000 utenti e 1000 strutture registrati. Dovevuoidormire.it è una piattaforma che permette l’incontro tra utenti e proprietari di strutture ricettive.

Dovevuoidormire.it suggerisce ai proprietari delle strutture come aumentare i profitti, consiglia agli utenti i periodi e le strutture migliori dove alloggiare ed effettua con i dati in suo possesso delle analisi sull’andamento delle prenotazioni e quali saranno le città piu’ scelte.

## Analisi e raccolta dei requisiti

**Requisiti per gli Utenti**

Gli utenti che prenotano devono essere maggiorenni

Gli utenti devono essere registrati

**Requisiti per le Strutture**

Una struttura può ospitare massimo massimo 4 ospiti

Le strutture possono essere di tre tipologie Hotel, Villa o Appartamento

Le strutture devono essere registrate in un apposito registro delle Agenzie delle Entrate e al comune.

**Requisiti per le Prenotazioni**

Utenti possono effettuare solo una prenotazione alla volta

**Team necessario**

Data Engineer

Progetta e crea la base di dati tramite

Data Analyst

Analizza i dati tramite interrogazioni della base di dati e crea dei report

Data scientist

Crea un modello predittivo per prevedere l’andamento futuro delle prenotazioni.

## 

## Strutturazione dei requisiti in gruppi di frasi omogenee

Utenti

Per gli utenti (10.000), identificati da un codice numerico univoco, il nome, il cognome, l’età, il genere, la città, la nazione e lo stato civile.

Strutture

Per le strutture (1000), identificate da un codice numerico univoco, il nome della struttura, il tipo della struttura, la città, il prezzo e la foto della struttura.

Prenotazioni

Le prenotazioni sono identificate da un codice numerico univoco, una data di inizio prenotazione, una data di fine prenotazione, l’importo della prenotazione e il numero degli ospiti.

## Glossario dei termini

| Termine | Descrizione | Sinonimi | Collegamenti |
| --- | --- | --- | --- |
| Utente | Persona che effettua una prenotazione | Ospite | Strutture, Prenotazioni |
| Struttura | Luogo dove le persone scelgono di alloggiare | Alloggio | Utenti, Prenotazioni |
| Prenotazione | Pernottamento dell utente nella struttura | Soggiorno | Utenti, Struttura |

## 

## 

## 

## Tipologia di dati

Utenti

| id\_Utente | numerico | strutturati |
| --- | --- | --- |
| Nome | testo | strutturati |
| Cognome | testo | strutturati |
| Età | numerico | strutturati |
| Genere | boolean | strutturati |
| Stato\_Civile | sposato/single | strutturati |
| Città | testo | strutturati |
| Nazione | testo | strutturati |

Strutture

| id\_Struttura | numerico | strutturati |
| --- | --- | --- |
| Nome\_Struttura | testo | strutturati |
| Tipo\_Struttura | categorico | strutturati |
| Città | testo | strutturati |
| Prezzo | numerico | strutturati |
| Foto\_Struttura | non strutturato | non strutturati |

Prenotazioni

| ID\_Prenotazione | numerico | strutturati |
| --- | --- | --- |
| id\_struttura\_prenotata | numerico | strutturati |
| id\_cliente | numerico | strutturati |
| Data\_Inizio | data | strutturati |
| Data\_Fine | data | strutturati |
| Importo | numerico | strutturati |
| Numero\_Ospiti | numerico | strutturati |

## Supporto alle decisioni aziendali

Analisi di business

* Trend
  + Analisi dei periodi storici
* Peak seasons
  + Periodo preferito dagli utenti
* Peak location
  + Quali Città hanno avuto più prenotazioni
  + Quali tipi di strutture sono state maggiormente prenotate
* Target customers
  + Quali tipologie di utente ha effettuato più prenotazioni
* Campagne Marketing/Advertising
  + Consigliare le strutture di quale città promuovere con campagne di advertising. Sapere l’effetto delle campagne marketing effettuate prima e dopo la data delle campagne

Analisi predittiva

* Trend: Previsione nei prossimi 3 mesi
* Peak seasons
  + Periodo preferito dagli utenti
* Peak location
  + Quali Città saranno le preferite dagli utenti nei prossimi 3 mesi
* Target customers
  + Quali tipologie di utente effettuerà più prenotazioni nei prossimi 3 mesi
* Campagne Marketing/Advertising
  + Consigliare le strutture di quale città promuovere con campagne di advertising
  + Sapere l’effetto delle campagne marketing effettuate prima e dopo la data delle campagne

## Storia/Infografica

1. Infografica sul cliente tipo per le città più grandi.
2. Infografica su tipo di prenotazione maggiormente effettuata
   1. Venezia uomo sposato, sopra i 40, per san valentino;
   2. Quanti uomini sposati hanno prenotato a san valentino a venezia;
   3. Milano, uomini e donne under 35,nel periodo della fashion week.
3. Infografica su città con più prenotazioni (immagine sotto)

Di seguito un’Infografica creata sui dati delle locations piu’ prenotate. L’immagine mostra il paese italia con dei cerchi di grandezza che varia al numero di prenotazioni. Maggiore è la dimensione del cerchio maggiore è il numero delle prenotazioni nella città.



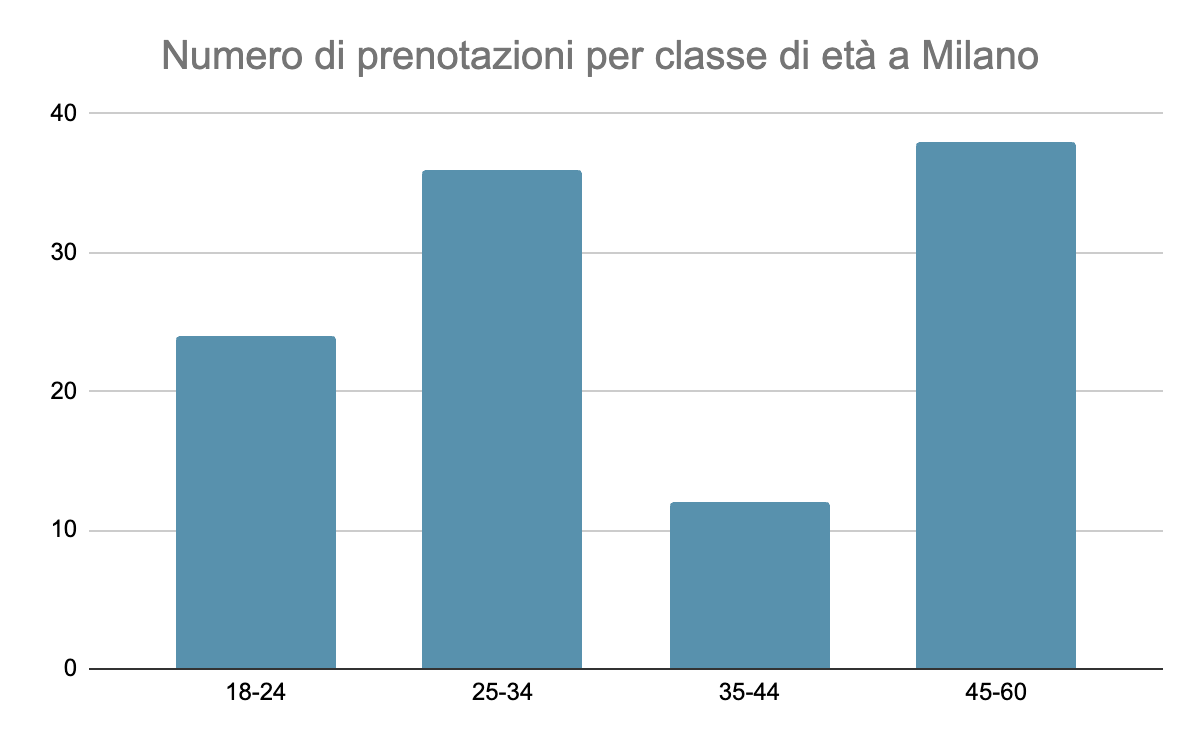
## Outcome

* Numero di prenotazioni in un anno a Milano per classi di età (18-25, 25-34, 35-44, 45-60) per individuare il target group del Marketing team.
* Vogliamo sapere la Media, Varianza, Mediana del prezzo delle strutture a Milano.
* Individuare il cliente tipo per le città più grandi
* Quante prenotazioni ci sono per Natale a roma
* Che età hanno prenotato maggiormente a roma durante il Natale.
* Quante prenotazioni sono state effettuate da uomini sposati a Milano.
* Che tipologia di struttura è maggiormente scelta da persone di età minore di 30 anni.
* Seleziona i nomi e i cognomi delle persone che hanno fatto almeno una prenotazione nel mese di Ottobre.

## Caso d’uso

Interrogazione della base di dati delle tabelle utenti e prenotazioni. In particolare ID\_prenotazione, ID\_utente, raggruppati per età (18-25, 25-34, 35-44, 45-60). Il risultato sarà un grafico di tipo istogramma da cui possiamo identificare che il maggior numero di utenti appartiene alla fascia [45-60], seguita dalla fascia [25-34].

I dati sono da moltiplicare per 1000 i dati sono riferiti all’ultimo anno 2021.



## Possibili implementazioni future

1. Creare un algoritmo per consigliare agli utenti la migliore città per le prossime vacanze in base al loro budget.
2. Vendere i dati tramite API a pagamento a altri servizi per lo studio dei dati nelle città
3. Rendere liberi i dati ad enti governativi o di ricerca per fare ricerche sull’andamento del numero di persone nelle città in giorni molto importanti come eventi nazionali e internazionali.