Area di interesse 1: prezzi delle case

<https://www.kaggle.com/datasets/vikrishnan/boston-house-prices>

<https://www.kaggle.com/datasets/shibumohapatra/house-price>

Informazioni che potremmo ricavare:

* Zona in cui si trova
* Dimensioni e numero stanze
* Numero dei servizi accessori (bagni)
* Numero di stanze extra (giardino, terrazzo)
* A quale piano si trova
* Presenza o meno di ascensore
* Data di costruzione
* Vista
* Orientamento
* Efficienza energetica
* Domanda per l’immobile (quanto è richiesto)

Analisi che potremmo svolgere:

* Analisi del prezzo di vendita per zona
* Analisi del prezzo di vendita per dimensioni e numero di stanze
* Analisi delle dimensioni delle case in una determinata zona
* Analisi del contesto economico e sociale in cui si inserisce l’immobile
* Presenza di ascensore in base alla data di costruzione, stato dell’immobile
* Analisi del prezzo di vendita in base alla presenza di fonti rinnovabili ed efficienza energetica

Figure professionali utili:

* Data analyst: raccogliere, pulire e presentare i dati al manager
* Data scientist: previsione dei prezzi delle case nel futuro

Area di interesse 2: password e sicurezza

<https://www.kaggle.com/datasets/morph1max/password-security-sber-dataset>

Informazioni che vorremmo ricavare:

* Lunghezza
* Numero di lettere
* Numero di numeri
* Numero di caratteri maiuscoli
* Numero di caratteri minuscoli
* Numero di caratteri speciali
* Numero di vocali
* Presenza di sillabe

Analisi che potremmo svolgere:

* Analisi sulla complessità della password
* Analisi sulla frequenza d’uso di una password
* Livello di casualità delle combinazioni scelte

Figure professionali utili:

* Data scientist: uso dell’intelligenza artificiale per testare la sicurezza delle password e
* Data analyst: capire quale sia la probabilità che si scelga una data password
* Data engineer: sviluppare metodi per proteggere i dati come associare il riconoscimento facciale
* Cyber Security Specialist: applica tecnologie, processi e misure di protezione progettate per ridurre il rischio di attacchi informatici

Area di interesse 3: previsione della cancellazione delle prenotazioni negli hotel

<https://www.kaggle.com/datasets/youssefaboelwafa/hotel-booking-cancellation-prediction>

Informazioni che vorremmo ricavare:

* Numero identificativo della prenotazione
* Prezzo della stanza
* Presenza di offerte
* Numero di persone
* Numero di adulti/bambini
* Numero di notti prenotate
* Numero di notti nel fine settimana
* Richiesta di un posto auto
* Possibilità di effettuare la cancellazione gratuita
* Servizi inclusi nel soggiorno (pasti, spa)
* Meteo delle giornate prenotate

Analisi che potremmo svolgere:

* Meteo delle giornate nel periodo in cui c’è cancellazione
* Presenza di eventi nelle giornate prenotate
* Analisi dei costi negli alberghi in cui c’è cancellazione rispetto agli altri alberghi
* Con quanto anticipo avviene la cancellazione e se c’entrano eventi particolari (scioperi)
* Che tipo di clienti effettua la cancellazione (single, coppie o famiglie)

Figure professionali utili:

Data analyst: analisi delle statistiche di cancellazione nello scorso trimestre

Data scientist: previsione del trend delle cancellazioni