**Analisi e raccolta dei requisiti**

**Argomento**:   
L’obiettivo principale è quello di ingaggiare il maggior numero di utenti utilizzatori dell’app “Adidas run” partendo dall’analisi delle vendite regionali, dati demografici residenti regionali e dati sull’utilizzo dell’app per singolo utente.

**Descrizione caso studio:**

L’ufficio vendite Italia di “Adidas run” richiede un’analisi sulle vendite di gennaio 2021 incrociate con gli utenti attivi dell’app “Adidas run” nello stesso periodo al fine di valutare una campagna di marketing che miri ad incentivare l’acquisto del nuovo paio di scarpe da corsa. Il dataset con i dati demografici della popolazione residente per regione ci permetterà di dare una misura dei clienti attivi nelle regioni prescelte e valutare della attività anche negli store fisici delle maggiori città. Per lo studio utilizzeremo anche i dati generati dal comportamento dell’utente nell’app “Adidas run”.

**Open Data:** i dati in nostro possesso sono dati strutturati in formato cvs, xml e json.

**Fonte dei requisiti**:

* Demografia\_regionale\_Italia\_2021 **(XML) –** Fonte ISTAT -

<http://dati.istat.it/viewhtml.aspx?il=blank&vh=0000&vf=0&vcq=1100&graph=0&view-metadata=1&lang=it&QueryId=18460&metadata=DCIS_POPRES1>

Licenza: Creative Commons (CC BY 3.0 IT)

* User\_app **(JSON)** – Fonte “Adidas Run” –

Licenza: dati proprietari dell’azienda e coperti da copyright

* Venduto\_gen21 **(CSV)** – Fonte “Adidas Run” –

Licenza: dati proprietari dell’azienda e coperti da copyright

Classificazione qualità Data: 3 stelle per tutti I dataset

**Struttura dataset**

L’attributo chiave: “regione”, “anno” e “mese” che nel dataset csv esprime le vendite (quantità e valore) di “Adidas run” per regione, nel dataset xml i dati demografici espressi in regione, e nel terzo dataset vengono riportati i dati per ogni singolo utente iscritto l’app “adidas run”. Il tutto circoscritto al mese di gennaio 2021.

Il dataset “Demografia\_regionale\_Italia\_2021” dispone dei seguenti dati: regione, nome regione, popolazione residente nella stessa regione anno 2021, divisa per genere (uomo, donna). I dati precedenti sono disponibili per tutte le regioni italiane.

Il dataset “User\_app” dispone dei seguenti dati: codice identificativo utente, nome, cognome, genere, età, indirizzo, città, via, n civico, km percorsi nel mese di gennaio 2021, regione dove ha praticato l’attività motoria, accesso all’app nel mese di gennaio 2021 e date degli accessi. Il campione si compone di 10 utenti delle regioni più rappresentate.

Il dataset “Venduto\_gen21” dispone dei seguenti dati: anno, mese, regione dove sono avvenute le vendite, quantità di scarpe venduta nello stesso periodo, vendite di scarpe a valore per lo stesso periodo.

**Analisi e raccolta dei requisiti**

Di seguito le statistiche utilizzate con il dataset “User\_app”: media, max, min e frequenza accessi degli utenti in app, divisi per regioni, genere (uomo, donna); media, max, min dei km percorsi dagli utenti divisi per regioni, genere (uomo, donna), età. Media, min, max del numero di accessi per età, genere, regione.

Di seguito le statistiche utilizzate con il dataset “Vendite\_gen21”: media, max, min delle vendite a quantità per regione; media, max, min delle vendite a valore per regione.

Il dataset “Demografia\_regionale\_Italia\_2021”: invece è stato utilizzato da confronto con le statistiche dei due dataset precedenti, utilizzando i dati della popolazione divisa per genere, e regione. Pesando ogni genere sul totale della popolazione.

**Strutturazione dei requisiti**

**Frase di carattere generale:**

L’obiettivo principale è quello di ingaggiare il maggior numero di utenti dell’app “Adidas run”. L’ufficio vendite Italia di “Adidas run” richiede un’analisi sulle vendite di gennaio 2021 incrociate con gli utenti attivi dell’app “Adidas run” nello stesso periodo al fine di valutare una campagna di marketing che miri ad incentivare l’acquisto del nuovo paio di scarpe da corsa.

**Outcome**

L’obiettivo dell’analisi è studiare il comportamento dell’utente in app per fidelizzarlo e indurlo ad acquistare le nuove scarpe da running, con una particolare focus legato al genere dell’utente. Si vuole produrre una campagna commerciale completamente su app in modo da poter inserire in app il modello di scarpe utilizzate che possano generare informazioni future per la progettazione di nuove scarpe. Grazie ai dati divisi per regione si potrà organizzare degli eventi di diversa natura e grandezza in funzione della numerosità di utenti residenti nella stessa regione o che praticano l’attività fisica nella regione di riferimento.

**Figure interessate per l’analisi:**

* Data Engineer: ha raccolto i dati sulle vendite e sugli utenti dell’app.
* Data Analyst: ha elaborato i dati in funzione delle richieste dell’ufficio vendite che punta a conoscere meglio gli i comportamenti degli utenti.
* Data scientist: dall’analisi comportamentale degli utenti, struttura un modello predittivo sulle nuove funzionalità dell’app.

**Glossario**

|  |  |  |  |
| --- | --- | --- | --- |
| **Termine** | **Descrizione** | **Sinonimi** | **Collegamenti** |
| Dati demografici | Dati che indicano composizione quantitativa di una popolazione per una determinata area | Indicatore di popolazione | clienti |
| Fidelizzazione | Insieme di azioni di marketing da parte di un’azienda per aumentare i punti di contatto con il cliente | ingaggio | Clienti |
| Clienti | Persona che acquista le scarpe da running | acquirente | Fidelizzazione, dati demografici, app |
| App | Software creato dall’azienda per far accedere i clienti a tutti i servizi proprietari | Applicativo web | clienti |