

# Raccolta e analisi dei requisiti

## Descrizione generale

Si vuole analizzare le abitudini di ascolto e comportamenti degli utenti della piattaforma Spotify con il fine di indirizzare le campagne pubblicitarie in maniera coerente.

## Analisi e raccolta dei requisiti

### Dati:

- ID utente
- Fascia d'età: 14-18, 19-25, 26-34, 35-44, 45-60, 60+
- Genere: maschio, femmina, non specificato
- Paese:
- Genere musicale: Trap, Musica Italiana, Club, Reggaeton, Musica da studio, Musica Internazionale, Indie, etc..
- Playlist: Concentrazione, Motivazionali, Relax, Playlist anni 80, Playlist anni 90, etc.
- Podcast: Economia, Storia, Notizie, Hobby, etc.
- Tempo di ascolto medio al giorno ( in ore ): <0.5, 0.5-2, 2-4, >4
- Formato: Musica, Podcast
- Orario d'ascolto: 06:00-12:00, 12:00-14:00, 14:00-18:00, 18:00-21:00, 21:00-24:00, 24:00-06:00
- Dispositivo d'ascolto: auto, telefono, assistente vocale, pc,
- Sistema operativo: iOS, macOS, Windows, Android, Huawei, Linux
- Comportamento: molto attivo, e poco attivo.

### Requisiti

Una prima ipotesi di ricerca sarà quella di poter raggruppare gli utenti in base al loro potere d'acquisto. Utilizzando come parametri la fascia d'età, il paese, il dispositivo d'ascolto e il sistema operativo applicheremo un algoritmo di machine learning di classificazione che ci permetterà di individuare dei cluster raggruppabili nelle seguenti categorie: basso spendente, medio spendente, alto spendente. L'intervallo di analisi sarà degli ultimi 5 anni.

Una seconda ipotesi di ricerca sarà quella di poter raggruppare gli utenti in base a delle categorie di abitudini. Utilizzando come parametri fascia d'età, genere, paese, genere musicale, tempo di ascolto, formato, orario d'ascolto, dispositivo d'ascolto, sistema operativo applicheremo un algoritmo di machine learning di classificazione che ci permetterà di individuare dei cluster raggruppabili nelle seguenti categorie: Studente, lavoratore, casalingo, sportivo, pensionato, etc. L'intervallo di analisi saranno degli ultimi 5 anni.

Una terza ipotesi di ricerca sarà quella di poter raggruppare gli utenti in base a due categorie che li identificano come ads-lover e ads-averse. Utilizzando come parametri tempo medio di ascolto al giorno, il comportamento, dispositivo d'ascolto. L'intervallo di analisi sarà degli ultimi 5 anni.

## Glossario dei termini

**Utenti:** tutti coloro che hanno un'iscrizione gratuita o premium alla piattaforma Spotify.

Sinonimi: ascoltatori

Collegamenti: Azienda, ID utente,

**Id utente:** identificazione univoca dell'utente sotto forma di numero.

Sinonimi:

Collegamenti: Utente, Azienda

**Campagne pubblicitarie:** si intendono tutte le attività che aziende, società, e privati legate alla pubblicità all'interno della piattaforma Spotify.

Sinonimi: advertising, adv, ads, pubblicità

Collegamenti: Azienda, Utente

**Abitudini d'ascolto:** si intendono i pattern ricorrenti degli utenti nell'utilizzo della piattaforma Spotify quando ascoltano i contenuti.

Sinonimi:

Collegamenti: Utente

**Genere musicale:** classificazione di brani in base al loro genere musicale prevalente.

Sinonimi: categoria musicale.

Collegamenti: Utente, Playlist

**Playlist:** insieme di brani raggruppati in base a dei fattori comuni come il genere musicale o il momento d'ascolto.

Sinonimi:

Collegamenti: Utente, Genere musicale

**Podcast:** contenuto audio sotto forma di racconti, format raggruppati in categorie di argomento.

Sinonimi:

Collegamenti: Utente, Categoria podcast

**Tempo di ascolto medio al giorno:** tempo di ascolto medio al giorno per singolo utente.

Sinonimi: tempo di ascolto

Collegamenti: Utente

**Formato:** Tipologia di contenuto audio che si suddivide in musica e podcast.

Sinonimi:

Collegamenti: PODCAST

**Dispositivo d'ascolto:** Dispositivo su cui viene ascoltato il contenuto.

Sinonimi: dispositivo, dispositivo di riproduzione

Collegamenti: Utente

**Comportamento molto attivo:** Utenti che eseguono ricerche specifiche dei brani, che cambiano spesso brano e che è più propenso a saltare le campagne pubblicitarie.

Sinonimi:

Collegamenti: Utente

**Comportamento poco attivo:** Utenti che tendono ad ascoltare i brani senza delle specifiche ricerche, e utilizzando i tasti di riproduzione del brano. Sono meno propensi a saltare le campagne pubblicitarie.

Sinonimi:

Collegamenti: Utente

**Potere d'acquisto:** predisposizione dell'utente a spendere soldi.

Sinonimi:

Collegamenti: Utente

**Basso spendente:** Utente con possibilità economica limitata.

Sinonimi:

Collegamenti: Utente

**Medio spendente:** Utente con possibilità economica normale.

Sinonimi:

Collegamenti: Utente

**Alto spendente:** Utente con possibilità economica sopra la norma.

Sinonimi:

Collegamenti: Utente

**Ads-lover:** Utenti che sono più propensi a consumare interamente le pubblicità.

Sinonimi:

Collegamenti: Utente

**Ads-averse:** Utenti che sono più propensi a saltare le pubblicità.

Sinonimi:

Collegamenti: Utente

# Dettaglio dell'analisi

## Tipologia di dato

Per come abbiamo impostato il problema, i dati da noi utilizzati sono prevalentemente dati strutturati, in quanto possiamo raggrupparli in maniera organizzata all'interno di tabelle e ben definiti. Per quanto riguarda il dato "comportamento" è di tipo non strutturato all'origine che però convertiremo in dato strutturato per questa analisi, riducendolo ai due soli campi "molto attivo" e "poco attivo".

## Output

La nostra ricerca è un'analisi di business e un'analisi predittiva. Il nostro output sarà una tabella dove ogni utente identificato dal dato "Id utente" viene associato alle categorie oggetto di studio in base a: potere d'acquisto, abitudini e la sua predisposizione a consumare la pubblicità. Inoltre avremo un algoritmo ML allenato per classificare tutti gli utenti entranti nelle diverse categorie. In base alle categorie che abbiamo individuato, creeremo un servizio di vendita degli spazi pubblicitari all'interno della piattaforma Spotify dove i nostri clienti potranno veicolare la pubblicità ad un pubblico di loro interesse. Lo scopo della nostra ricerca è quindi finalizzato ad aumentare le entrate di Spotify tramite la pubblicità sulla piattaforma.

## Ruoli

Data Engineer: Creazione e mantenimento del database.

Data Scientist: utilizzo di algoritmi di machine learning per la classificazione.

Business Analyst: commissiona il progetto.

## Storie

Utilizzeremo i dati per:

1. Individuare le categorie musicali più ascoltate per fasce d'età.
2. Individuare le categorie musicali più ascoltate per paese.
3. Individuare le categorie di podcast più ascoltate per fasce d'età.
4. Individuare le categorie musicali più ascoltate per fasce d'orario d'ascolto.
5. Individuare le categorie di podcast più ascoltate per genere
6. Capire il tempo medio d'ascolto per genere