Team Data sci-analyst

Metodologia di lavoro

I dataset (DS) fornitici riportano dati sul settore merceologi “scarpe”, sia da uomo che da donna:

n.3 dataset riportanti dati di scarpe da donna

n.1 dataset riportanti dati di scarpe da donna

Tool utilizzati: notebook Jupyter, principalmente librerie Pandas, Numpy, Matplot, Sys, Math, Json.

Un primo scopo dell’analisi dei dati è quello di trovare delle correlazioni tra gli stessi che ci potessero dare delle indicazioni interessanti sullo spazio di mercato in cui fare la nostra proposta di lancio di un nuovo brand di scarpe, in ottica di estensione di brand.

Abbiamo fatto una prima analisi della tipologia di dati presenti nei dataset. Da questa prima generica analisi, abbiamo estrapolato le seguenti informazioni: i dataset riportano i dati di modelli di scarpe distribuiti attraverso il canale online, nel mercato statunitense (la maggior parte dei modelli riporta la valuta USD).  
Dall’analisi del contenute delle colonne, si è deciso di procedere lavorando solo con i dati di nostro interesse, estrapolando quindi solo le colonne che riportano i dati che abbiamo ritenuto utili per le nostre analisi, eliminando tutte quelle colonne che riportavano informazioni verosimilmente usate interamente dai diversi distributori, per la catalogazione dei diversi articoli; abbiamo altresì eliminato tutte le colonne che riportavano dati in quantità poco rilevante rispetto alla totalità degli items del DS.

La colonna “recensioni” , anche se non presente in tutti i DS, ci è sembrata, in questa prima fase, quella più interessante per estrarre una prima indicazione sulla categoria di scarpe che ha destato più interesse tra quelle presenti, a prescindere che il giudizio sull’articolo sia negativo o positivo. Non avendo, infatti, dati di vendita, abbiamo ritenuto che la presenza delle recensioni ci desse intanto un’indicazioni di quali articoli sono stati venduti e quindi di maggiore interesse.  
Sia che il giudizio sia negativo che positivo, abbiamo quindi individuato uno spazio di azione; nel primo caso, possiamo agire sul miglioramento dell’articolo, nel secondo caso, possiamo inserirci in diretta concorrenza.

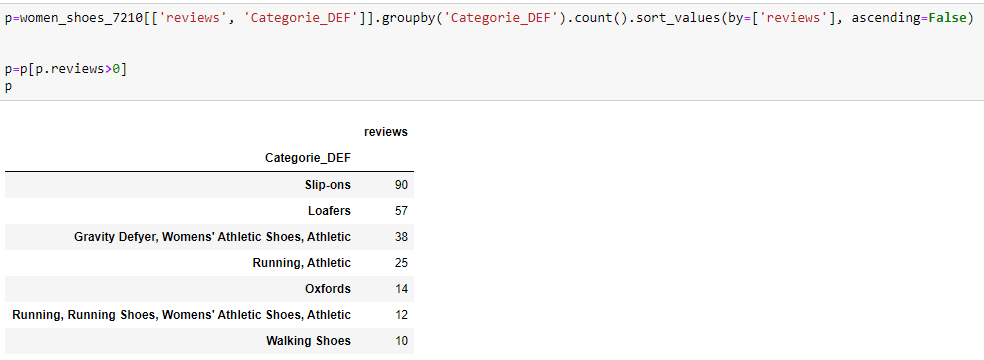
Da tale analisi sulle recensioni, si è quindi concluso che tra le categorie più recensite ci sono le scarpe casual/sneakers. Questa indicazione è stata riportata quindi al tema mkt.

Ogni Item del DS riportava poi sia un prezzo minimo che massimo (probabilmente legati a promozioni in corso sui colori e/o taglie); si è quindi ritenuto necessario estrarre una colonna prezzo medio.

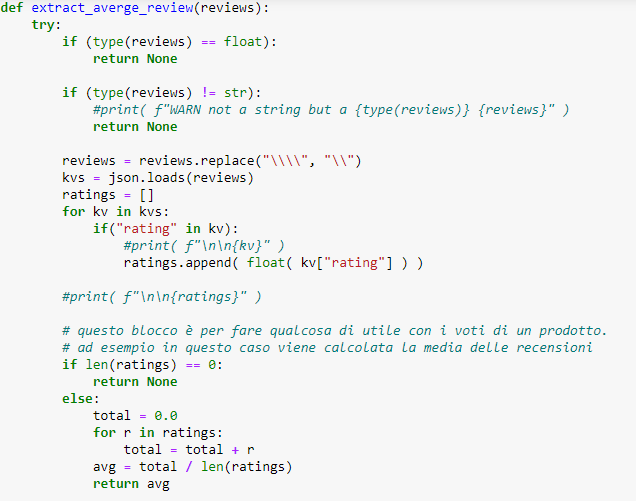


Per confermare la scelta della categoria casual/sportiva, abbiamo analizzato il numero di di recensioni presenti per ogni categoria di nostro interesse.

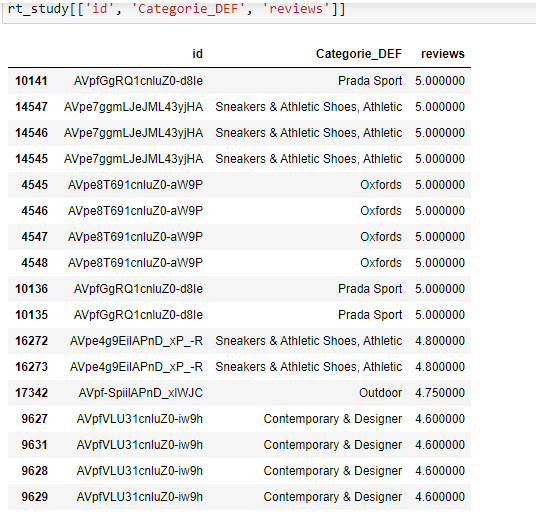




Il rating delle recensioni sono all’interno di un Jason, abbiamo quindi prima estratto il valore rating da tutte le recensioni e poi abbiamo fatto una media, in quanto molti degli articoli recensiti hanno più di una recensione.



Di seguito vediamo che le categorie di nostro interesse hanno tutte delle buone recensioni



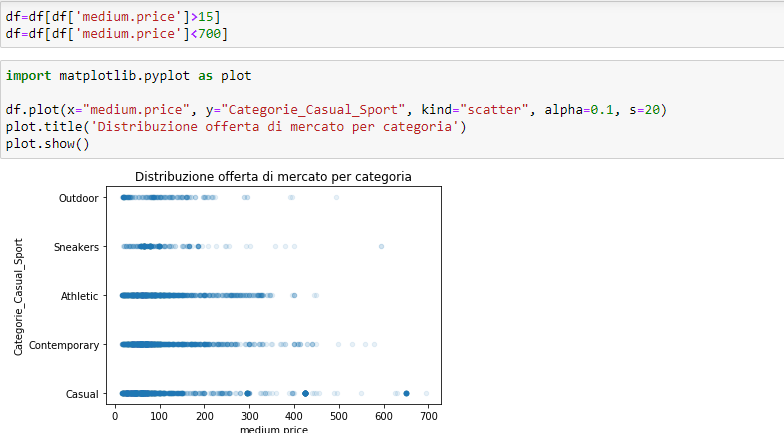
ANALISI DEI PREZZI

Abbiamo fatto una prima analisi dei prezzi in base alle categorie

Abbiamo raggruppato le categorie di nostro interesse in macro categorie, in modo che i dati fossero più leggibili

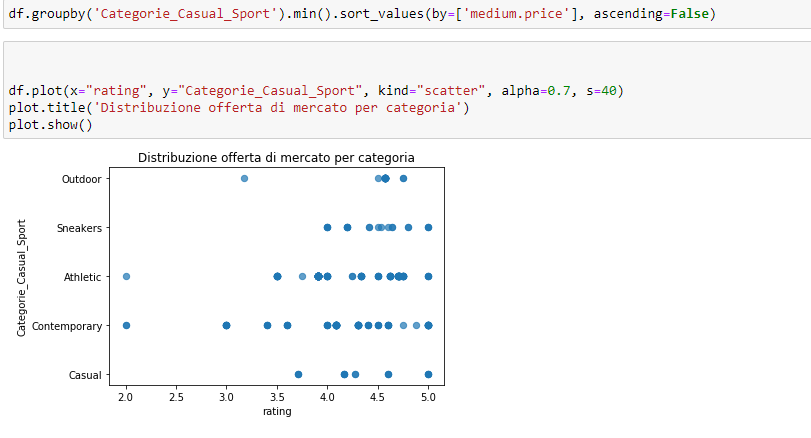


Per minimizzare il rischio di inserire outliers all’interno della distribuzione dei prezzi per categoria, abbiamo deciso di eliminare items con prezzi troppo alti o troppo bassi.



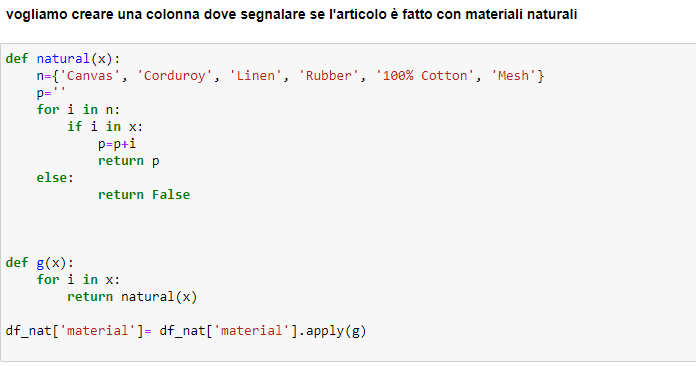
**Proposta di interpretazione dati: il mercato delle sneakers offre un'interessante opportunità nella fascia media alta, che non è saturo come quello di fascia bassa, notiamo in particolare come potremmo agilmente "inserirci" con un prezzo sui 150 euro/dollari in quanto andiamo ad intercettare una vera e propria carenza di offerta.**

**Abbiamo poi fatto una distribuzione dei prezzi in base al rating delle recensioni**

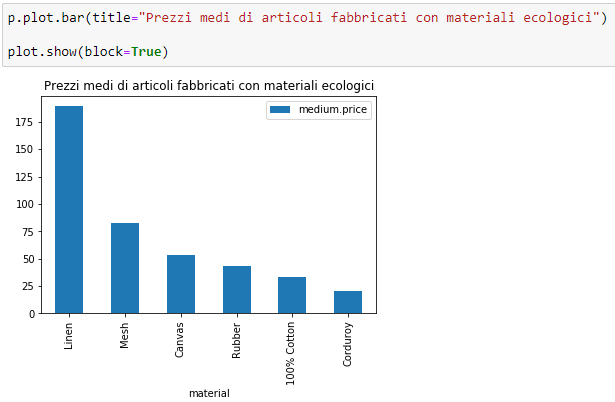


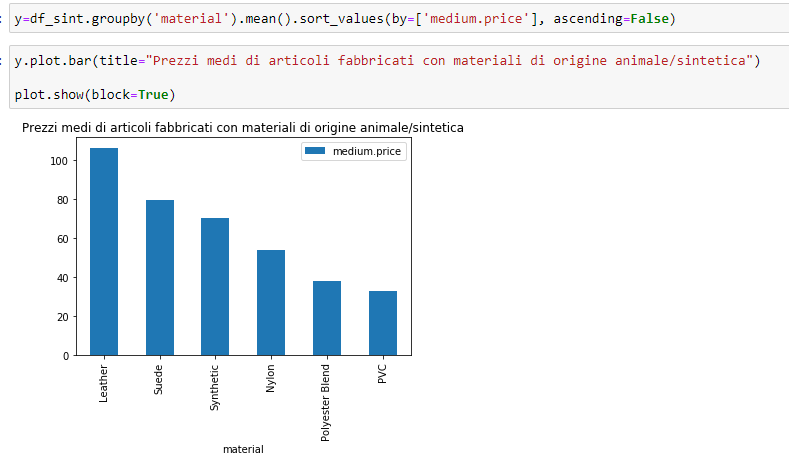
Notiamo come la categoria sneackers/casual sia ben posizionata con punteggi di rating piuttosto alti

Abbiamo poi analizzato la distribuzione dei prezzi in base ai materiali in tutte le categorie a dimostrazione del fatto che i materiali da noi categorizzati come eco siano competitivi con altri materiali non\_eco

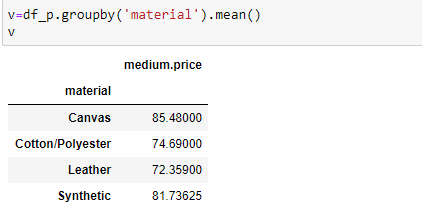








Infine abbiamo analizzato ricavato la distribuzione dei prezzi in base ai materiali, relativi solo alle categorie di nostro interesse casual/sneackers





Un'analisi più da vicino sul tipo di scarpa "sneakers" svela che la nostra ambizione nel voler proporre un articolo di costo medio, competitivo ma di qualità, è giustificato dall'attuale offerta di mercato. Il canvas offre un elevato rapporto qualità prezzo rispetto materiali sintetici e misti, e si propone come alternativa "vegana" ai prodotti fabbricati con derivati animali, con un incremento di prezzo rispetto a ques'ultimi che rappresentano la qualità a il valore aggiunto che siamo convinti i nostri consumatori saranno felici di pagare.

