

Raggruppamento Articoli di Vendita - Panoramica del Progetto

Obiettivo del Progetto

- **Scopo:** Trovare e classificare gli articoli che sono stabili e prevedibili nella vendita.
 - **Metodo:** Utilizzare dati storici per guidare decisioni più informate sulla gestione dello stock.
-

Tecniche utilizzate

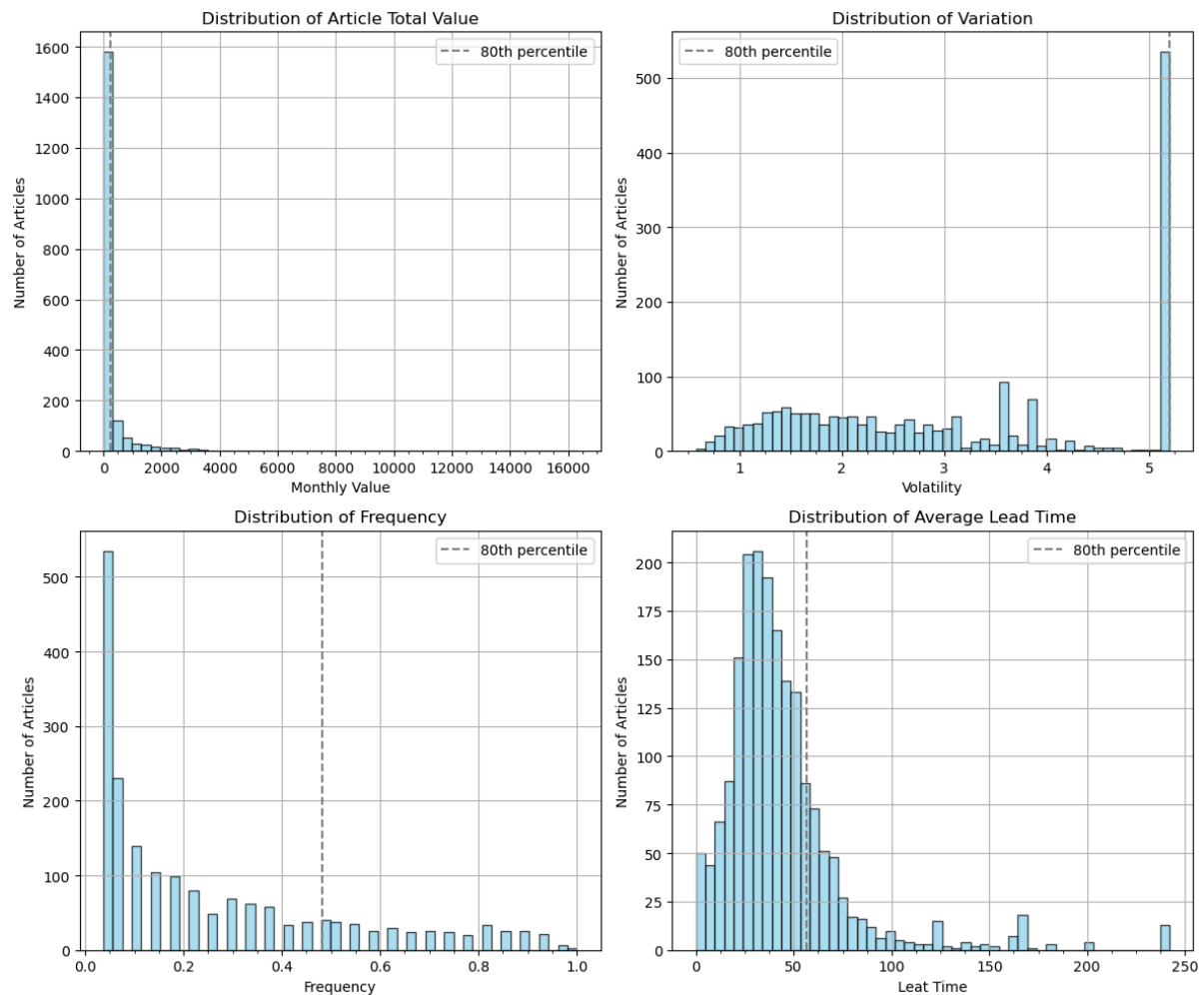
- **PCA:** Analisi delle componenti principali (riduzione della dimensionalità).
 - **Clustering:** Selezione di oggetti con caratteristiche omogenee a partire da un insieme di dati.
-

Passaggi Principali

Preparazione dei Dati

- Il dataset include:
 - Codici articolo
 - Quantità vendute mensili
 - Costo unitario
 - Lead time (tempo di consegna)
- La selezione dei dati segue queste logiche:
 - Limite dei dati tra 2023-01 fino a 2025-02 incl.
 - Limite delle vendite complete (con date di consegna)
 - che hanno un prezzo di vendita
 - i prezzi di vendita sono stati selezionati in base a un criterio di data (più recenti)
- Sono stati calcolati alcuni indicatori chiave per ogni articolo:
 - **Fatturato medio mensile** (quanto vale economicamente ciascun articolo)
 - **Stabilità delle vendite** (quanto sono regolari mese per mese)
 - **Frequenza di vendita** (in quanti mesi l'articolo è stato venduto)
 - **Costo medio unitario** (il valore di vendita della singola unità)
 - **Lead Time medio** (quanto tempo intercorre tra ordine e vendita)

Istogrammi delle principali variabili



L'immagine mostra la distribuzione di tutti gli articoli combinati in raggruppamenti secondo le variabili principali. Si notano in particolare:

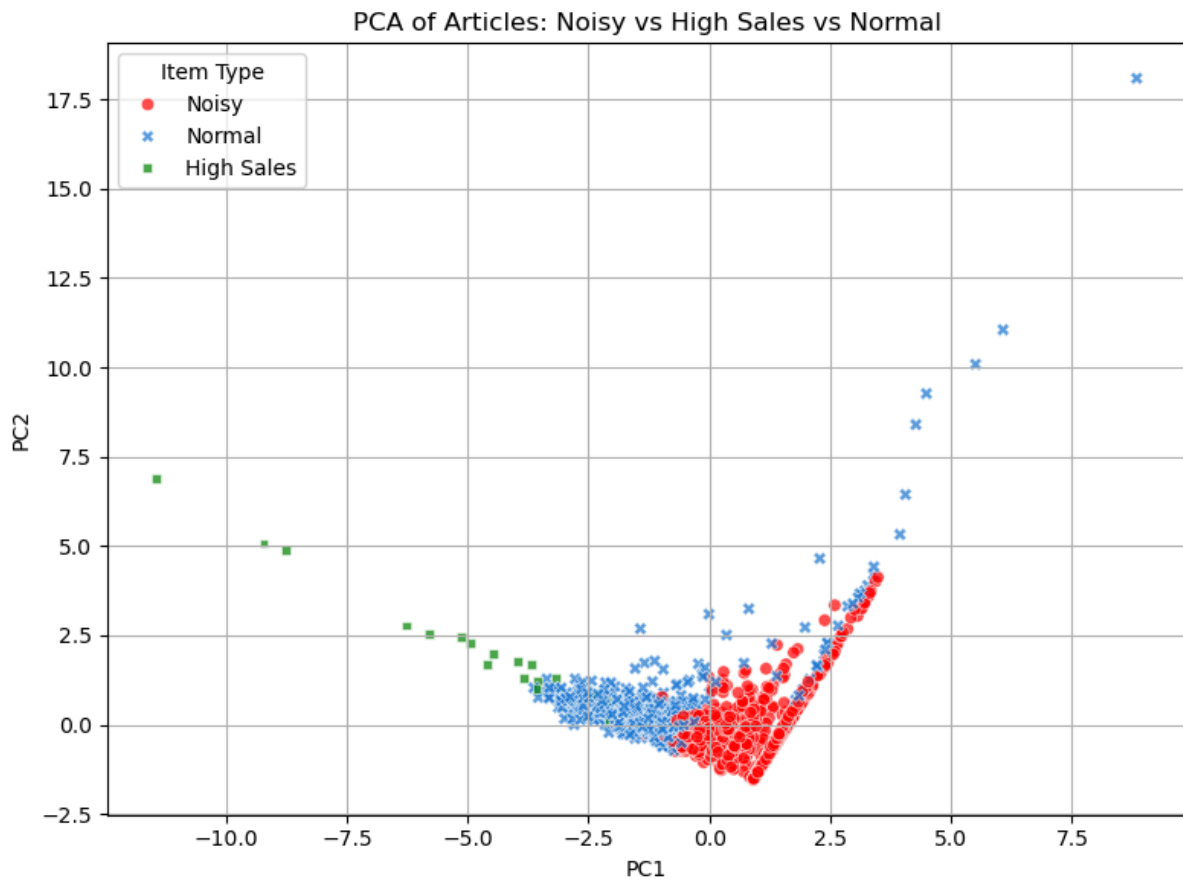
- **Fatturato:** notare la lunga "coda" a destra.
- **Volatilità:** notare il peso del "estremamente volatile" a destra.

Identificazione di Outliers

- Sono stati evidenziati articoli particolari:
 - **Articoli con vendite molto elevate** (top 1%)
 - **Articoli con costo unitario molto alto** (top 2%)
 - **Articoli "rumorosi"** (vendite instabili e poco frequenti)

Questi articoli verranno trattati con attenzione a parte.

Proiezione 2D dei principali componenti

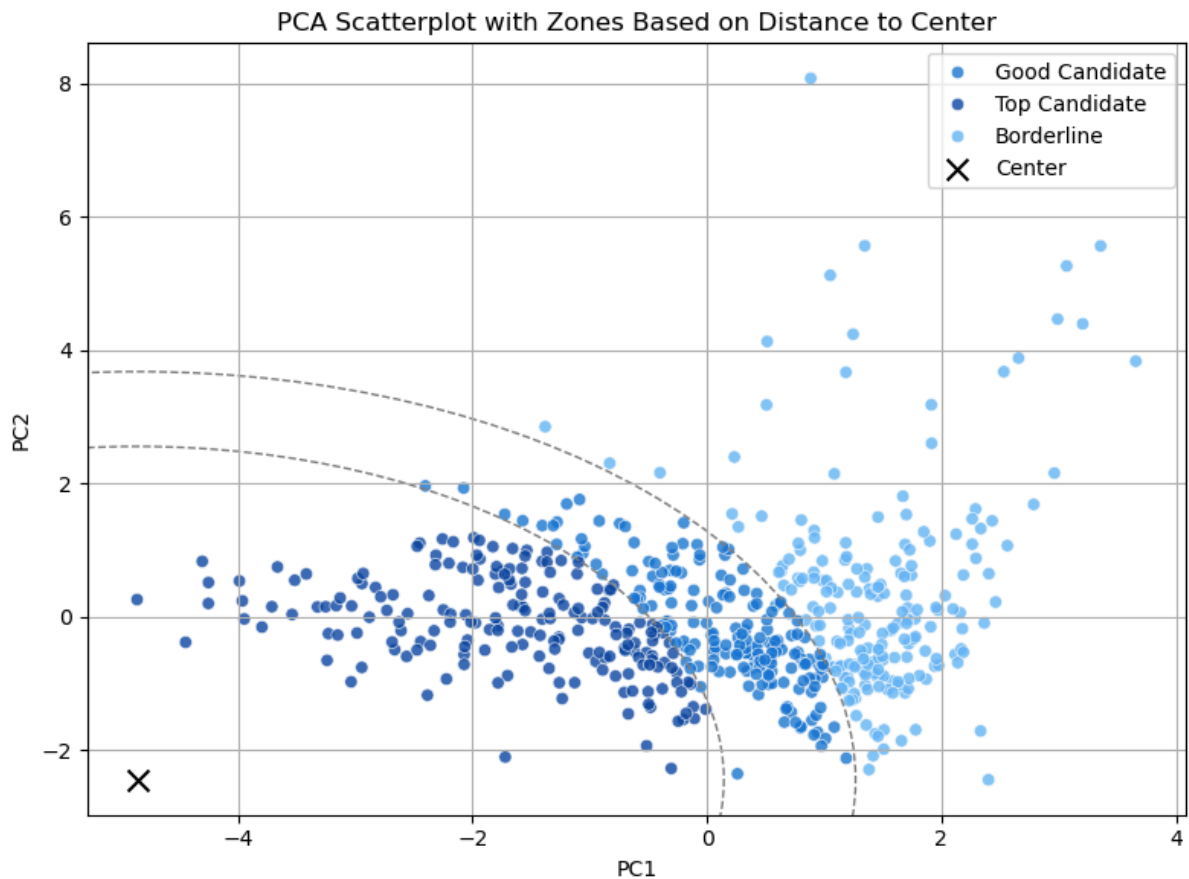


L'immagine distribuisce sul piano cartesiano gli articoli a seconda di etichette: outliers, articoli da tenere e articoli con valori di fatturato estremi. Gli outliers rappresentano circa 2/3 del totale; mentre gli altri due gruppi rappresentano 1/3.

Raggruppamento e Prioritizzazione Visiva

Step 1: focus su articoli ordinari (non outliers)

- Considerando le diverse variabili insieme, abbiamo rappresentato graficamente gli articoli per identificare quelli più vicini al "profilo ideale":
 - Vendite regolari
 - Costi contenuti
 - Alta frequenza di vendita
- Sono stati tracciati **cerchi** per delimitare:
 - **Top Candidate**
 - **Good Candidate**
 - **Borderline**

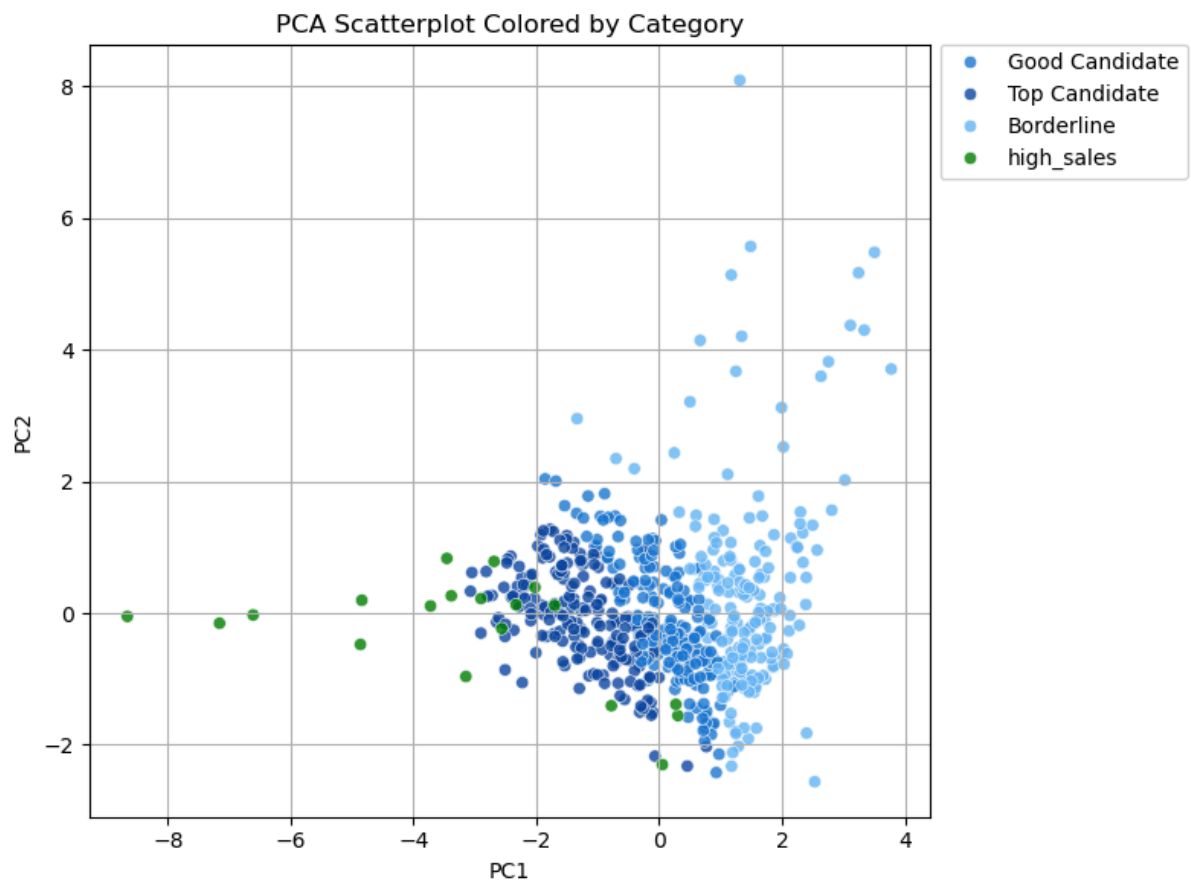


Risultato Finale

Step 2: inclusione di articoli con vendite estreme.

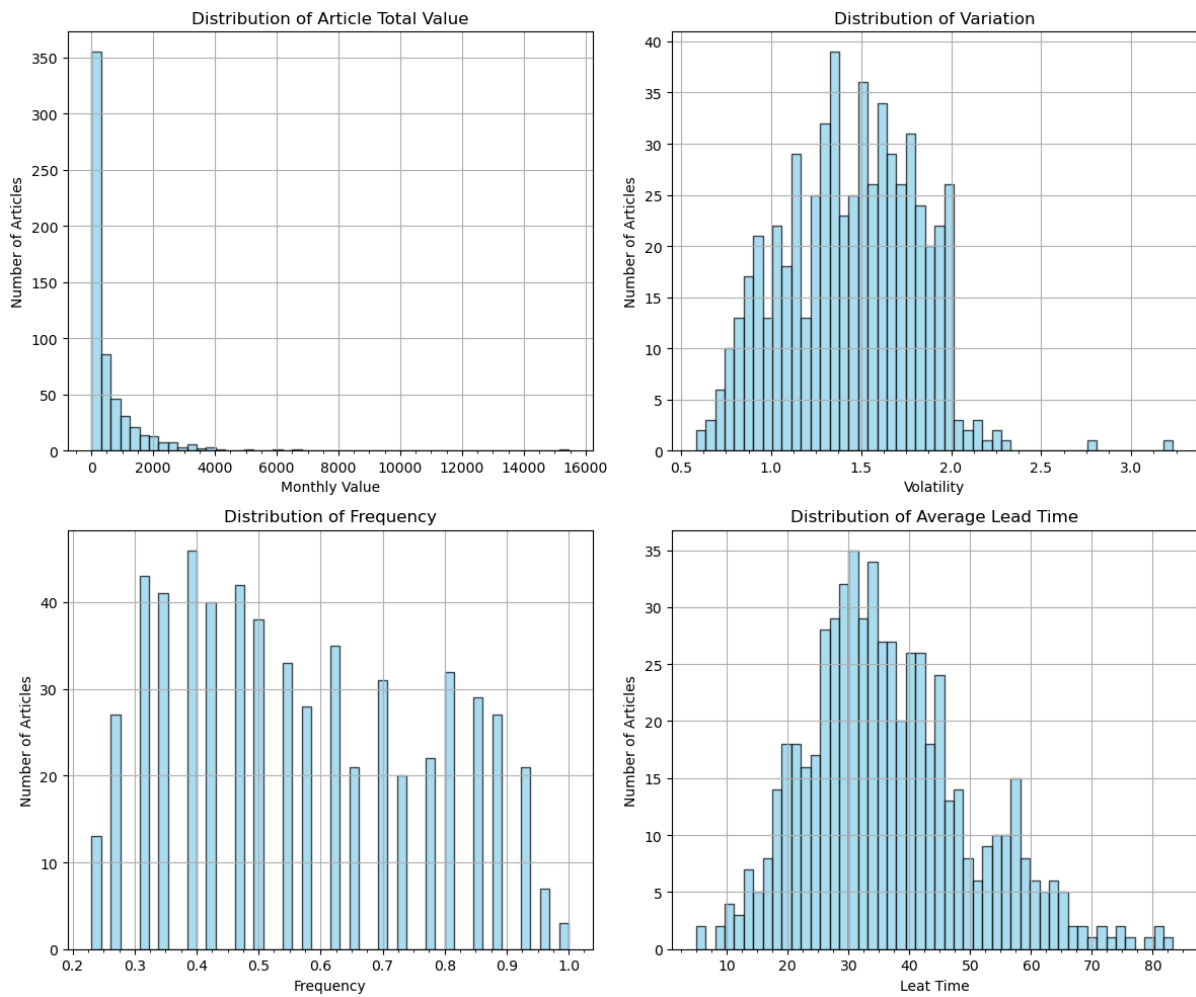
- Una lista pulita di articoli, etichettati come:
 - "Top Candidate" - Migliori candidati per essere mantenuti a stock
 - "Good Candidate" - Da considerare
 - "Borderline" - Più rischiosi
 - Eventuale flag "High Sales" o "Noisy" per articoli particolari

Proiezione in 2D dei candidati da mantenere a stock



L'immagine plotta e raggruppa a seconda delle variabili considerate le 4 'zone'. Una revisione manuale degli articoli e delle quantità può informare la produzione sugli articoli e i volumi da tenere a magazzino.

Istogrammi delle principali variabili con focus sugli articoli ordinari



L'immagine mostra come le curve di distribuzione si normalizzino, infatti abbiamo eliminato gli outliers.



Lista degli Articoli e dei Parametri

	category	item_id	zone_type	zone	avg_monthly_sales	avg_unit_cost	avg_lead_time	avg_total_value	cv_sales	freq_ratio
0	Carrelli a sfere (30)	4271	Standard	Top Candidate	12.654	6.06	39.523	76.682	0.938	0.846
1	Carrelli su rotaia T (40)	622.492	Standard	Top Candidate	6.269	108.17	42.180	678.143	0.734	0.885
2	Winch - Accessori (150)	2121	Standard	Top Candidate	22.000	37.14	28.574	817.080	1.111	0.808
3	Stopper (120)	507.111	Standard	Top Candidate	13.154	53.03	25.969	697.548	1.052	0.846
4	Carrelli su rotaia T (40)	623.402	Standard	Top Candidate	6.000	87.57	34.167	525.420	0.938	0.846
5	Carrelli su rotaia T (40)	623.462	Standard	Top Candidate	5.423	73.50	32.228	398.596	1.634	0.731
6	Rotaie (100)	4510-2000	Standard	Top Candidate	2.577	80.09	29.882	206.386	1.322	0.654
7	Winch - Accessori (150)	2022	Standard	Top Candidate	3.462	38.39	33.694	132.888	1.075	0.692
8	Stopper (120)	507.112	Standard	Top Candidate	18.923	59.38	29.526	1123.652	0.895	0.923
9	Winch - Accessori (150)	2021	Standard	Top Candidate	45.500	34.66	37.217	1577.030	0.933	0.846

Lo screenshot mostra le prime 10 righe della tabella contenente:

- Codice articolo
- Rating del Candidato da tenere a magazzino: ('top', 'good', 'borderline')
- Parametri mensili:
 - 'Volume di vendita',
 - 'Costo unitario',
 - 'Lead Time',
 - 'Fatturato',
 - 'Volatilità',
 - 'Frequenza'
- Categoria articolo

Specchietto Riassuntivo dei Cluster

	zone_type	zone	count
0	Outlier	Noisy Item	1213
1	Outlier	Costly Item	37
2	Outlier	Bullet Item	19
3	Standard	Borderline	197
4	Standard	Top Candidate	192
5	Standard	Good Candidate	191

Lo screenshot mostra il conteggio articoli per Tipo e Rating

Valore per il Business

- Concentrarsi su investimenti di stock più sicuri e prevedibili.
- Ridurre il rischio di mantenere articoli costosi o poco richiesti.
- Fornire raccomandazioni chiare e basate sui dati per la gestione dell'inventario.
- Ulteriori ricerche
 - visualizzare i dati di performance mensili (KPI)

Conclusione

Questo approccio combina analisi quantitativa, buon senso pratico e visualizzazione dei dati per supportare **decisioni di business più intelligenti**.