

Analisi delle Vendite Decathlon Milano

Progetto di Probabilità e Statistica per l'Informatica

Si vogliono analizzare le vendite di un Decathlon di Milano. In particolare valutiamo l'ipotesi di ampliare l'area assistenza e riparazioni, in quanto il nostro superiore crede che, oltre a migliorare il servizio alla clientela, influisce positivamente sulle vendite delle componenti.

Prefazione

Abbiamo raccolto dati per le 52 settimane dell'anno solare 2022. Si assume che i dati rilevati siano rappresentativi e non influenzati da eventi esterni. In questo modo, i dati possono essere utilizzati come gruppo di controllo valido per comprendere se l'area assistenza e riparazioni ha influito positivamente sulle vendite dei componenti.

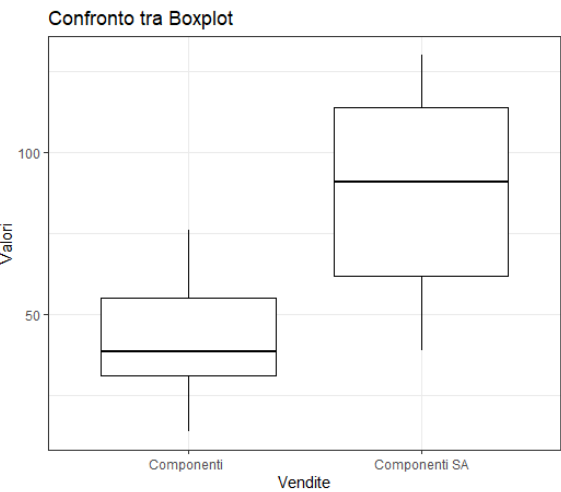
Si suppone che altri fattori esterni, che potrebbero aver influito sulle vendite durante il periodo di osservazione, non abbiano avuto effetti significativi rispetto all'area assistenza o riparazioni, o che tali effetti siano trascurabili.

Analisi dei Dati

Si ritiene opportuno effettuare prima una descrizione dei dati tramite indici di posizione e dispersione della statistica descrittiva. Osserviamo, in particolare, i quartili delle due colonne.

Indice	Vendite	
	Componenti	Componenti SA
Media Campionaria	42,48	87,98
Primo Quartile	31	61,5
Mediana Campionaria	38,5	91
Terzo Quartile	55	114,5
Varianza Campionaria	214,73	719,78
Deviazione Standard Campionaria	14,65	26,83

Si creano dei box plot per confrontare graficamente le vendite dei componenti con e senza servizio di assistenza.



Possiamo osservare come i quartili delle vendite dei componenti con servizio assistenza sono generalmente maggiori rispetto a quelli della vendita dei componenti senza tale servizio. Ciò suggerisce che l'area assistenza potrebbe influenzare positivamente sulle vendite.

Test di Ipotesi

Si vuole verificare se c'è sufficiente evidenza statistica per confermare quanto emerge dalla rappresentazione grafica dei dati. Si formula l'ipotesi che il servizio assistenza non influisca positivamente sulla vendita dei componenti. Si ottiene l'ipotesi nulla ed alternativa come segue:

$$H_0: \mu_{C_SA} \leq \mu_C \qquad H_1: \mu_{C_SA} > \mu_C$$

Verifica Ipotesi

Si può utilizzare un test t-student per confrontare le medie delle due distribuzioni in quanto avendo un numero di dati grande si può assumere la distribuzione normale degli stessi. Si esegue il test a livello di significatività 5%.

```
> t.test(valori$Componenti_SA, valori$Componenti, paired = T, alternative = "greater",
conf.level=0.95)

Paired t-test

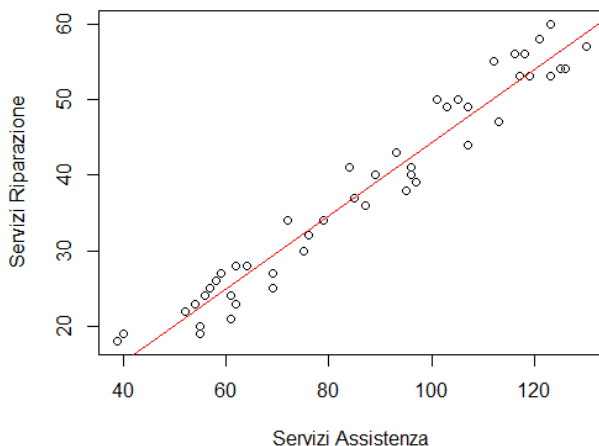
data:  valori$Componenti_SA and valori$Componenti
t = 21.048, df = 51, p-value < 2.2e-16
alternative hypothesis: true mean difference is greater than 0
95 percent confidence interval:
 41.87849      Inf
sample estimates:
mean difference
      45.5
```

Poiché il valore p è di molto inferiore al livello di significatività α ($p\text{-value} < 0.0001$), possiamo respingere l'ipotesi nulla e dunque accettiamo l'ipotesi alternativa, dimostrando empiricamente che la vendita con servizio assistenza influisce positivamente sulla vendita dei componenti.

Regressione Lineare

Vogliamo adesso considerare l'ipotesi che all'aumentare dei servizi di riparazione effettuati, aumentano anche i servizi di assistenza, in quanto è ragionevole ipotizzare che le componenti da sostituire vengano prese da quest'ultima area. Per farlo, effettuiamo un test di regressione lineare.

Regressione lineare



Graficamente si evince come al crescere dei servizi di riparazione e manutenzione, crescono anche i servizi di assistenza.

Possiamo verificare anche tale correlazione con il coefficiente di correlazione lineare $r = 0.975$ il quale indica correlazione positiva significativa tra i due dati.

Conclusioni

In conclusione, i risultati dell'analisi delle vendite indicano che l'area assistenza ha un effetto positivo sulle vendite dei componenti e che la vendita dei componenti tramite servizio assistenza ha correlazione con il numero di riparazioni.

I box plot mostrano che le vendite dei componenti con servizio assistenza presentano una mediana più elevata rispetto a quelle senza servizio assistenza. Inoltre, il test di ipotesi conferma che la differenza tra le due distribuzioni è significativa.

Il test di regressione lineare dimostra la correlazione positiva significativa tra i servizi di assistenza e i servizi di manutenzione.

Si consiglia dunque un ampliamento dell'area servizio assistenza e riparazioni.

Tuttavia, è importante notare che questa conclusione si basa esclusivamente sui dati forniti e sull'analisi specifica condotta