

# Master di I livello: ANALISI DEI DATI PER LA BUSINESS INTELLIGENCE E DATA SCIENCE, edizione 2021/2022

Zakaria Vottero Vintrella

Titolo: Primi Strumenti di Data Science: Il caso Fabbri 1905

## **ABSTRACT**

Questa tesi riporta i risultati del mio tirocinio presso Fabbri 1905, un'azienda che opera nel mercato alimentare in tutto il mondo.

Il progetto ideato e sviluppato durante lo stage, e presentato nel documento di tesi, va a riportare l'introduzione dei primi strumenti di data mining all'interno di questa impresa. L'elaborato è suddiviso in 4 parti principali: la prima parte in cui vi è la storia dell'azienda e la sua struttura, la seconda parte composta dalla descrizione dei dati ottenuti della Fabbri 1905. Successivamente, più nel dettaglio, si apre la sezione dedicata all'analisi avanzata del dato, con le sue relative problematiche, con la descrizione dei lavori pratici e metodologici. Dalla raccolta dati, data cleaning, al preprocessing del dato. Nella terza parte vengono svolte le prime analisi di data mining per la Fabbri 1905, parlerò di due tecniche utilizzate: la Market Basket Analysis e la Cluster Analysis riportando il funzionamento di quest'ultime, oltre che dal punto di vista teorico, anche approfondendo il codice utilizzato. Infine, l'ultima parte riporta i risultati ottenuti da queste tecniche e le relative conclusioni.

Gli strumenti specifici utilizzati per condurre questa analisi sono stati Microsoft SQL Server per l'estrapolazione dei dati e Python per effettuare tutte le analisi.