DATA2001

Data Science Case Study Analysis

Name: Benjamin Israel Student ID: 45372995

Case Study: Walmart

Shape, arrow

Description automatically generated

*Below are some example questions and tips to structure your review of the case studies. DO not feel restricted or confined to these questions only in your case study analysis*

What data science problem is being addressed in the case study

Walmart is a large business in a competitive industry and they therefore require near immediate analysis of product price, availability and demand.

What data was collected/created in the case study, and what were the characteristics of the data (size, types, etc)

200 streams of internal and external data including 40 – petabyte data base of sales transactions of the previous weeks.

How was the data stored, curated, and prepared for analytics

Hadoop framework was used to create a ‘data democracy’ where all units of the business had access to the data. Spark, Cassandra, R and SAS where used in analytical applications.

What analytics were performed

Error catching in real time ie. Products not being listed.

Foresight: predicting demand for a product.

Insight: answering general business concerns.

What data stories were created, why and how did they respond to the original problem

Absence of sales for one product in the data stream where analysed as being a error by the shop. The shop was contacted and the error rectified.