

# **Data Analytics with Cognos Project**

## **PROJECT TITLE:**

PRODUCT SALES ANALYSIS

INNOVATION PHASE

## **TEAM MEMBERS:**

MUTHU RATHNASHRI C	2021115068
NITHYA SREE K	2021115071
MUTHAMIZH VAANJINATHAN M	2021115067
NAREN KARTHICK THIRUVENGADAM	2021115069
PRASAAD S	2021115307

## **OVERVIEW**

Innovation in product sales analysis involves a series of essential steps to optimize performance and gain valuable insights. First, data collection and aggregation is paramount, as it forms the foundation of analysis. This includes gathering data on sales transactions, customer demographics, and market trends. The next step is data cleaning and preprocessing, ensuring accuracy and consistency. Once the data is ready, advanced analytics techniques, such as data mining and machine learning, can be employed to identify patterns, correlations, and outliers. Visualization tools are then used to present these findings in a digestible format. To foster innovation, predictive and prescriptive analytics can be applied, allowing for forecasting and actionable recommendations. Regular monitoring and feedback loops complete the process, enabling continuous improvement and adaptation to changing market dynamics.

## **Innovation Steps**

### **Advanced Data Collection:**

Harness the power of web scraping tools and APIs to access real-time data from a wide array of global sources, ensuring a constant influx of up-to-date information.

Cultivate partnerships and collaborations with esteemed research institutions to bolster your data resources, resulting in enriched and comprehensive datasets that drive insightful analysis.

### **Enhanced Data Preprocessing:**

Integrate cutting-edge AI-driven algorithms to streamline the handling of missing values, outliers, and inconsistencies, facilitating a more efficient data preparation process.

Employ dynamic standardization techniques that remain adaptable to the ever-evolving landscape of data formats, ensuring your data is consistently well-prepared for analysis.

**Advanced Analytics Techniques:**

Exploration of advanced analytics techniques such as predictive modelling , machine learning, and artificial intelligence to identify patterns and trends in the data that may not be immediately obvious through traditional analysis.

**Segmentation and Personalization:**

It involves Implementation of innovative customer segmentation strategies using techniques like clustering and predictive analytics. Creating personalized product recommendations and marketing strategies for different customer segments based on their preferences and behaviours.

**Social Media and Sentiment Analysis:**

Incorporation of social media data and sentiment analysis to understand customer sentiment, opinions, and trends related to the products. This can provide valuable insights into what customers are saying about the products on social platforms.

**Geo-spatial Analysis:**

Utilization of geo-spatial data to identify regional variations in sales patterns.

Understand which products perform well in specific geographic areas and tailor marketing strategies accordingly.

**Market Basket Analysis:**

To Implement market basket analysis to discover associations between products frequently purchased together. This can help in optimizing product placement, cross-selling, and upselling strategies.

**Time Series Forecasting:**

Application of time series forecasting techniques to predict future sales trends accurately. This can aid in inventory management and ensuring that the right products are in stock when demand is expected to rise.

**A/B Testing and Experimentation:**

Conduct A/B testing and experimentation to test different marketing strategies, pricing models, and product placements. Use innovative designs for experiments to optimize your strategies continually.

**Text and Image Analysis:**

To Analyze product reviews and images to extract insights about customer satisfaction and preferences. Natural language processing and image recognition technologies can be valuable in this context.

### **Data Visualization Innovations:**

Experiment with cutting-edge data visualization techniques, such as 3D visualizations, interactive data stories, and immersive VR/AR visualizations to make data more engaging and accessible.

### **Collaborative Filtering and Recommendation Systems:**

Building recommendation systems that leverage collaborative filtering to suggest products to customers based on the behavior and preferences of similar customers.

### **Blockchain for Supply Chain Transparency:**

To explore the use of blockchain technology to enhance supply chain transparency and traceability, which can improve inventory management and build trust with customers.

## **Technology Utilized**

### **IBM Cognos Analytics:**

IBM Cognos Analytics is the core technology used in the project. It is a business intelligence tool that allows organizations to explore and analyze data, create interactive dashboards, and generate reports. Cognos Analytics provides a user-friendly interface for creating visualizations, conducting ad-hoc analysis, and sharing insights with stakeholders. It supports various data sources and integrates well with different databases and file formats.

### **Structured Query Language (SQL):**

SQL is a fundamental technology for data analysis in any database-driven project. It is used for querying and manipulating data in relational databases. SQL enables users to extract specific data subsets, perform calculations, and aggregate data for analysis. In the context of the project, SQL is likely used to retrieve and preprocess data from the sales database before importing it into IBM Cognos Analytics for further analysis.

**Data Preprocessing Tools (e.g., Python, Pandas):**

Data preprocessing tools like Python programming language and Pandas library are commonly used for cleaning, transforming, and preparing raw data for analysis. Data preprocessing is crucial for ensuring data quality and reliability in the analysis phase.

Role in the Project: Python and Pandas (or similar tools) are likely used to clean and preprocess the collected sales data, making it suitable for analysis in IBM Cognos. Data preprocessing ensures accurate insights and actionable results.

**Dataset Hosting and Community:**

Kaggle is a platform for data science and machine learning competitions, hosting datasets, and providing a collaborative environment for data enthusiasts and professionals. It offers a wide range of datasets that can be used for analysis and experimentation.

Role in the Project: Kaggle is the source of the provided sales dataset. The platform serves as a repository for datasets, and it also fosters a community where data scientists and analysts can collaborate, discuss techniques, and share insights. In this project, Kaggle is used for accessing the sales data necessary for analysis.