



Project Theme: Identifying Product Bundles from Sales Data

Domain: Data Science, Machine Learning and Artificial Intelligence

Company Information:

Bicree is a 3PL AI tech enabled logistic aggregator platform, which provides customers with quality automated shipping solutions. Its services offer opportunities to ship anywhere in India. Bicree's AI-powered dashboard is highly compatible with most e-commerce platforms and marketplaces. It is also integrable with custom websites and offers rest APIs.

Bicree's leading and specialised shipping partners ensure a safe and fast order delivery across varying volumes, type or size. It provides shipping across multiple segments ranging from apparels to electronics. Their platform helps in monitoring orders easily and increase profitability. It allows customers to analyse, connect, shipping assistance, auto NDR, COD Verification and Geo Analysis.

Website Link: <http://bicree.com/>

Overview:

Through this project, learners will understand the product bundles and SKUs methodology in the supermarket. Then they will apply the data collection approach to figure out data using pandas and other open-source libraries. Further, they will analyse the data to understand customer behaviour. And prepare a Data Analytics report/Dashboard using Python.

Outcome of the Project

Learners are required to analyse the given data to understand SKUs, Sales and buying behaviour of the customers. This is essential to build a final dashboard for the same.

Data Dictionary



SKUs information- Product Category, Product Name, Product Price, Product Weight, Product Length, total skus, product company name etc, other product information, related product, price, and required sales information- Total sales, product sales, SKUs sales, cost of sales, now need customer information- purchaser name, address, income, etc.

Project Task 1: Data Cleaning

Background: Data cleaning, or data wrangling or data remediation, or data munging, refers to a variety of processes designed to transform raw data into more readily used formats. It is the method of cleaning, arranging, and enriching raw information into the desired format for better decision-making in less time. The exact methods differ from project to project depending on the data you're leveraging and the goal you're trying to achieve.

When you collect raw data sets, they can be untidy and complex. Data wrangling unifies and sorts the data so it is easy to access and translate into actionable insights. Through data wrangling, data sets can be transformed into usable and functional formats, with any bad data corrected or removed. Those who collect the data, or other non-technical stakeholders within the company, can then more quickly and easily understand the data and make better decisions based on it.

Task Explained:

As a person who is responsible for the client relations with Bicree, you are required to collect the data from all the different sources using various data collection approaches and perform Data Structuring and Data Cleaning on the data of SKUs, sales and purchases and inventory.

Resources:

Dataset:

<https://docs.google.com/spreadsheets/d/1Sq5UuhCSqiqP14QHBrIYgBYEELYwJZihxQ5d1dnVRKM/edit#gid=1386482999>

[Resource 1](#)



[Resource 2](#)

Task Submission

Summarize your work in a document that is easy to understand. Submit your document which is either in .docx or .pdf format. To submit this task you can go to the Project Task section of your dashboard and find Company Task 1 listed there. Click on choose file and upload.

Deadline

The assignment should be submitted by **Monday, 30th May 2022.**