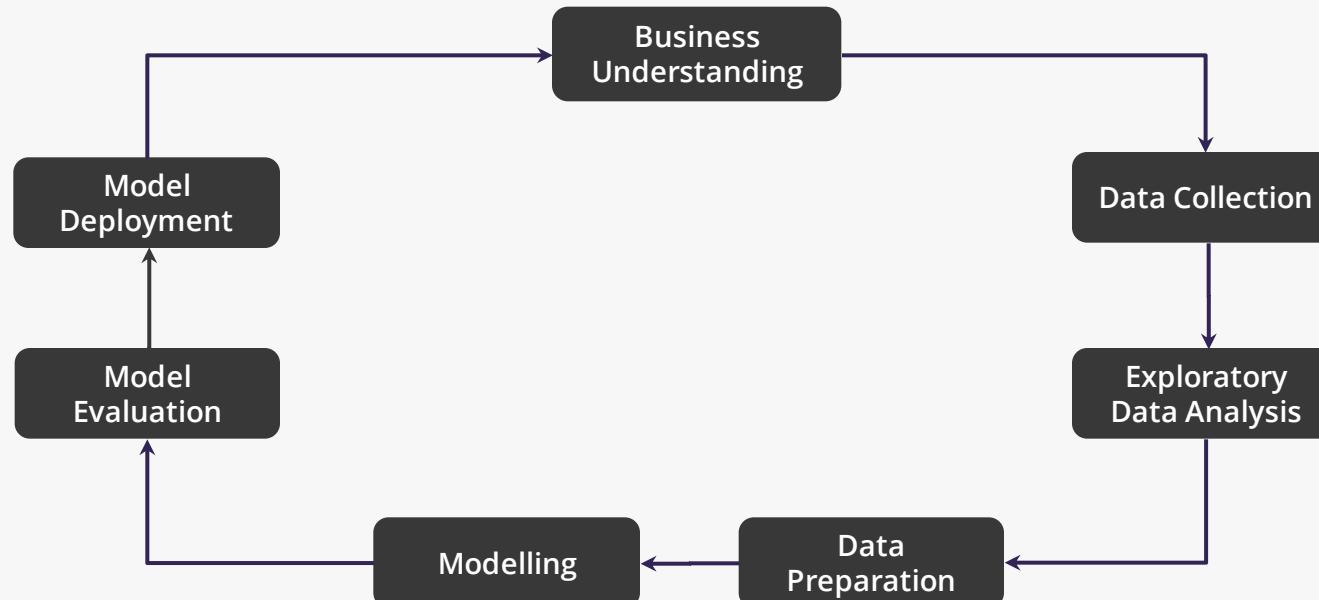


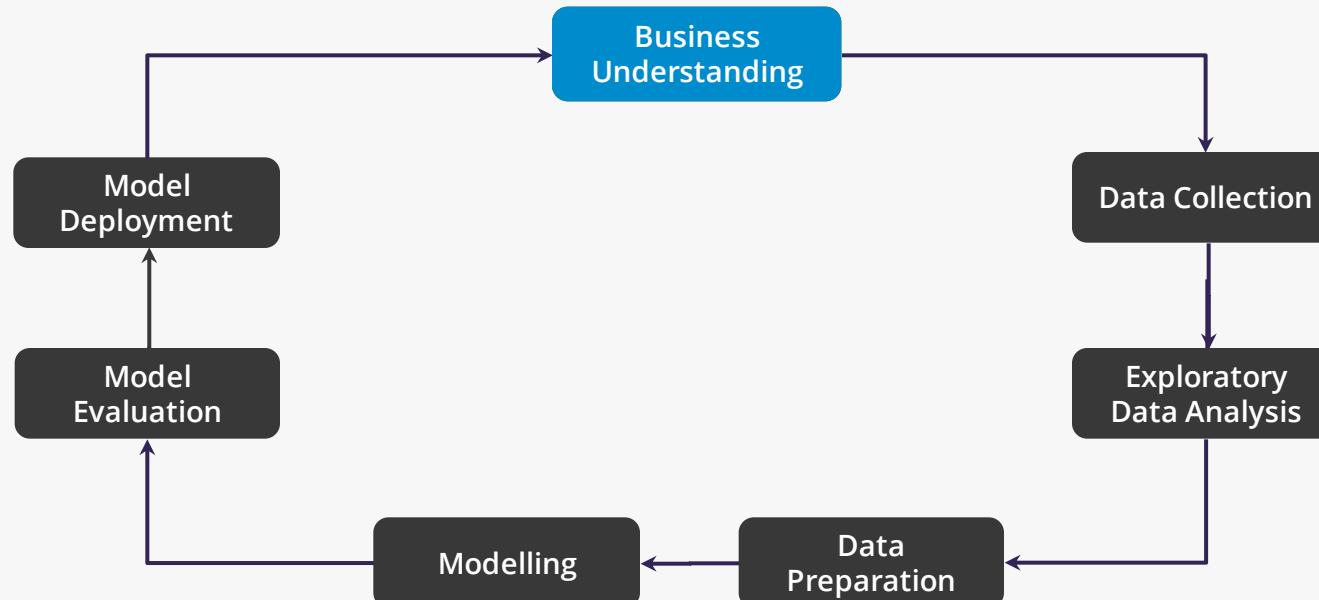


Machine Learning Workflow

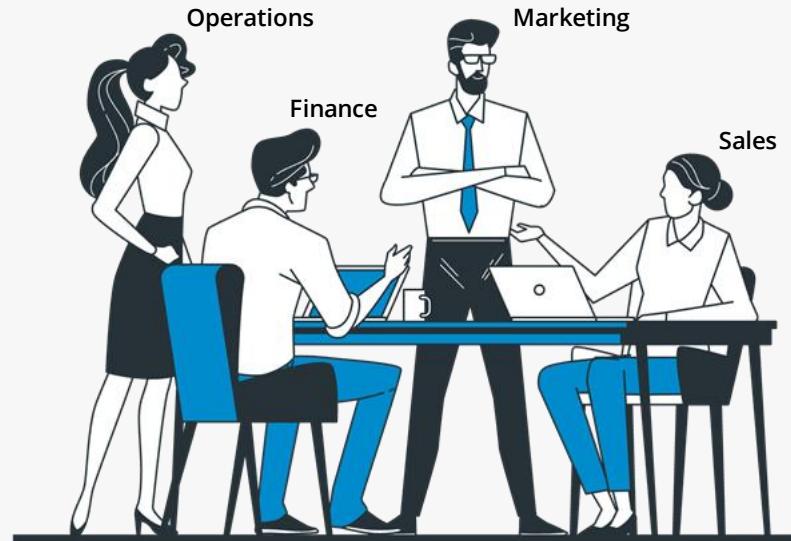
Machine Learning Workflow



Machine Learning Workflow

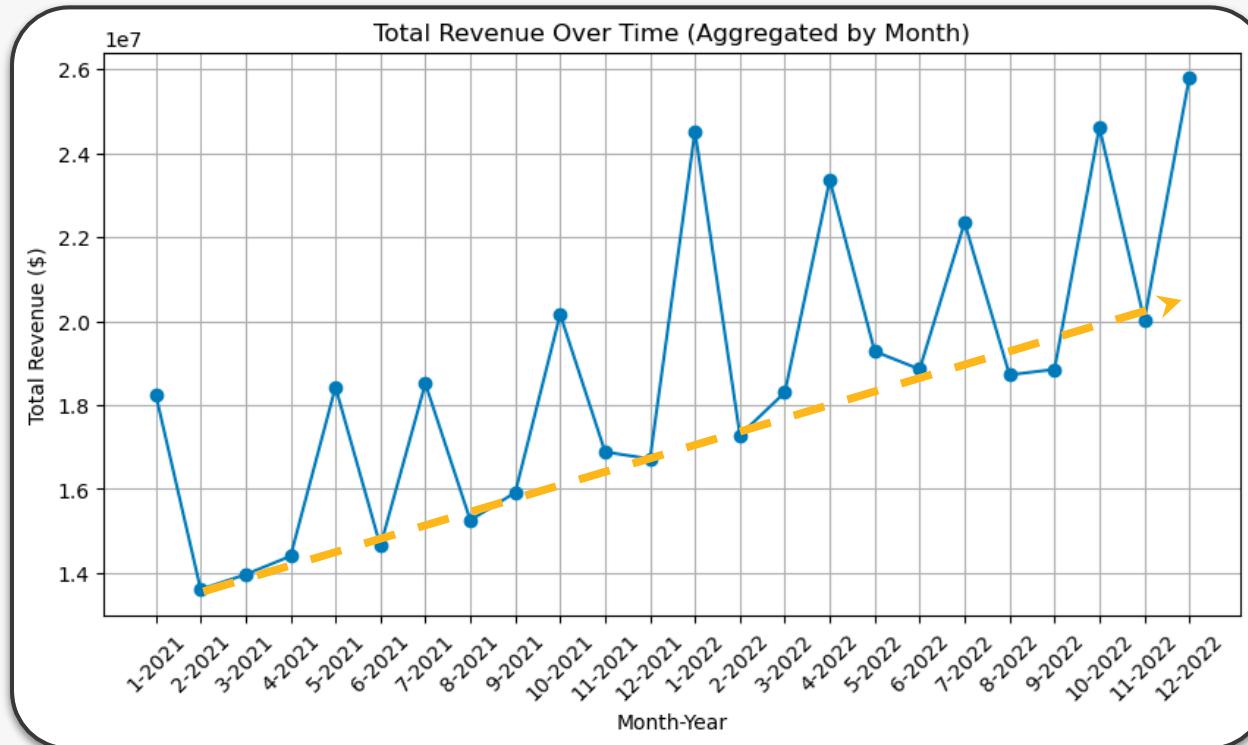


Business Understanding

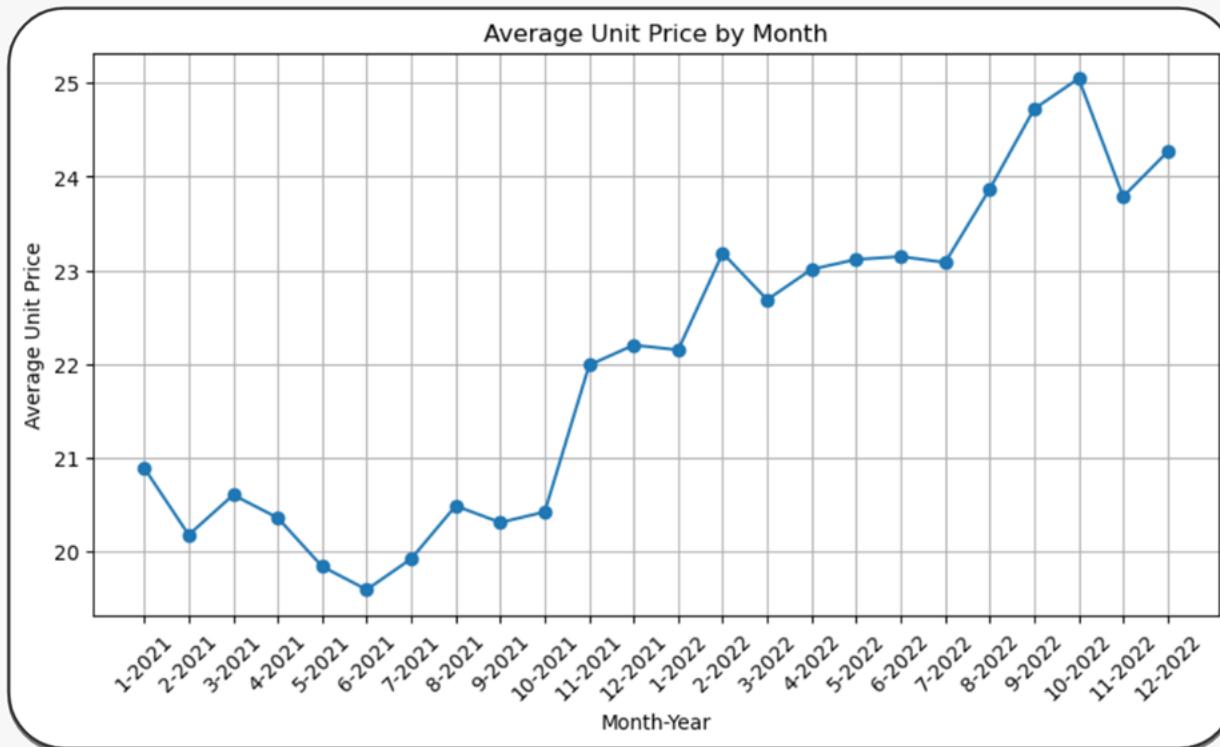


Stakeholder Meeting

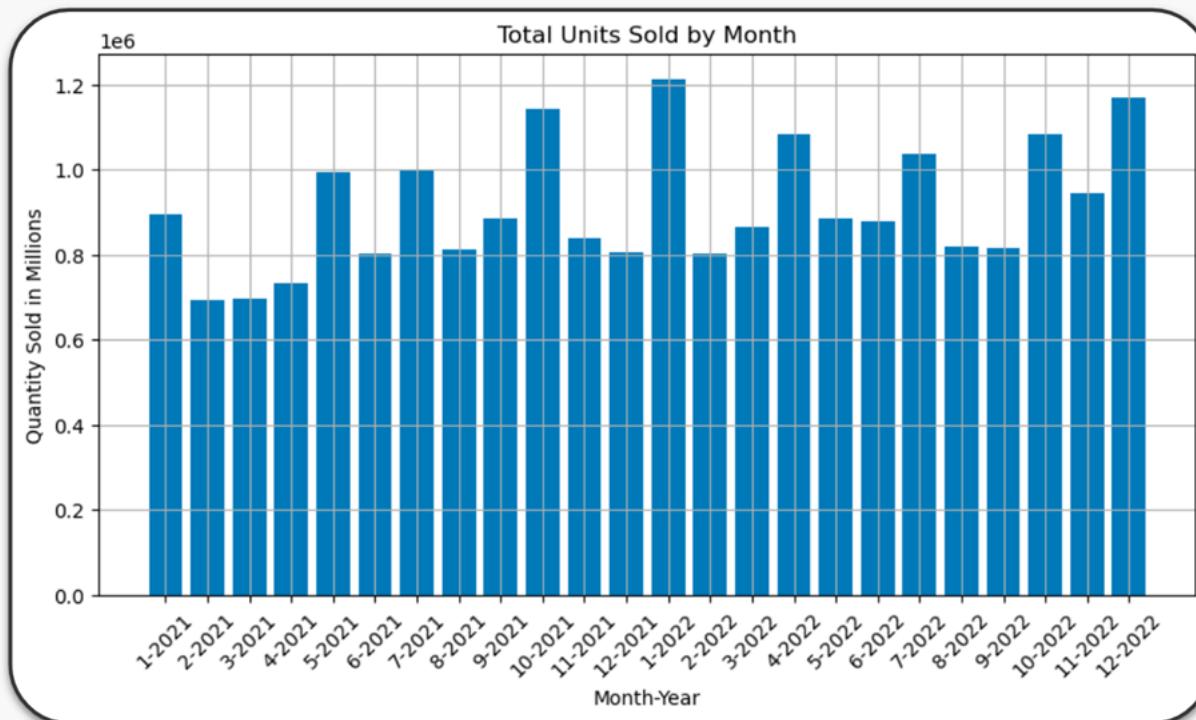
Business Understanding



Business Understanding



Business Understanding



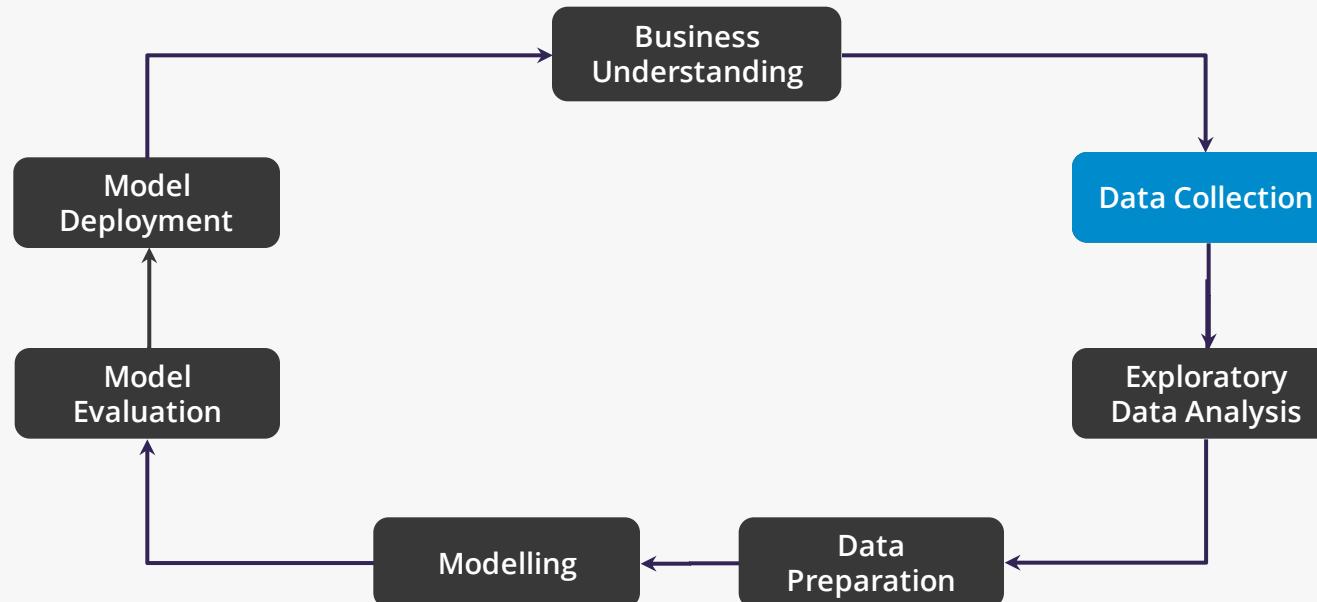
Business Understanding



Problem Statements

1. Identify the factors to predict units sold for products
2. Identify which products will sell more than 1000 units
3. Create a new-age bundled online marketing strategy

Machine Learning Workflow



Data Collection



Stakeholder Meeting

Understand the factors and map required data to solve the problem

Data Collection

POS_Data

Online_Data

Offline_Data

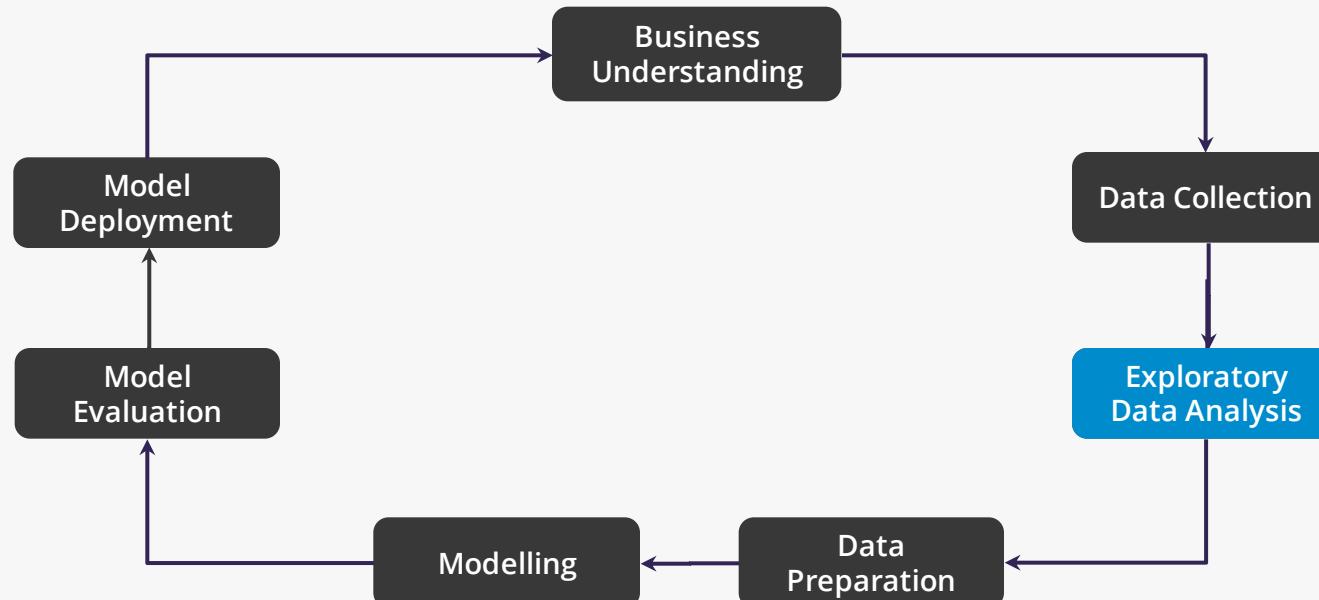
Product_Attribute_Data

VPC_Data

Search_Rank_Data

Synergix Solutions has collected data in 6 different tables.
Let's take a look at what each of these tables contain.

Machine Learning Workflow



Exploratory Data Analysis



In EDA, data is **analyzed** and **visualized** using various techniques by the user to obtain relevant insights.

Exploratory Data Analysis

1

Generate relevant insights for understanding



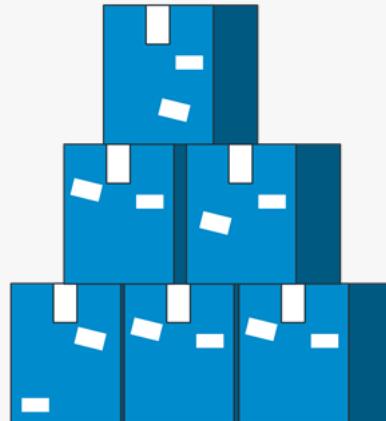
Generate insights on:

- A) What has been happening?
- B) What is working?
- C) What needs improvement

Exploratory Data Analysis

2

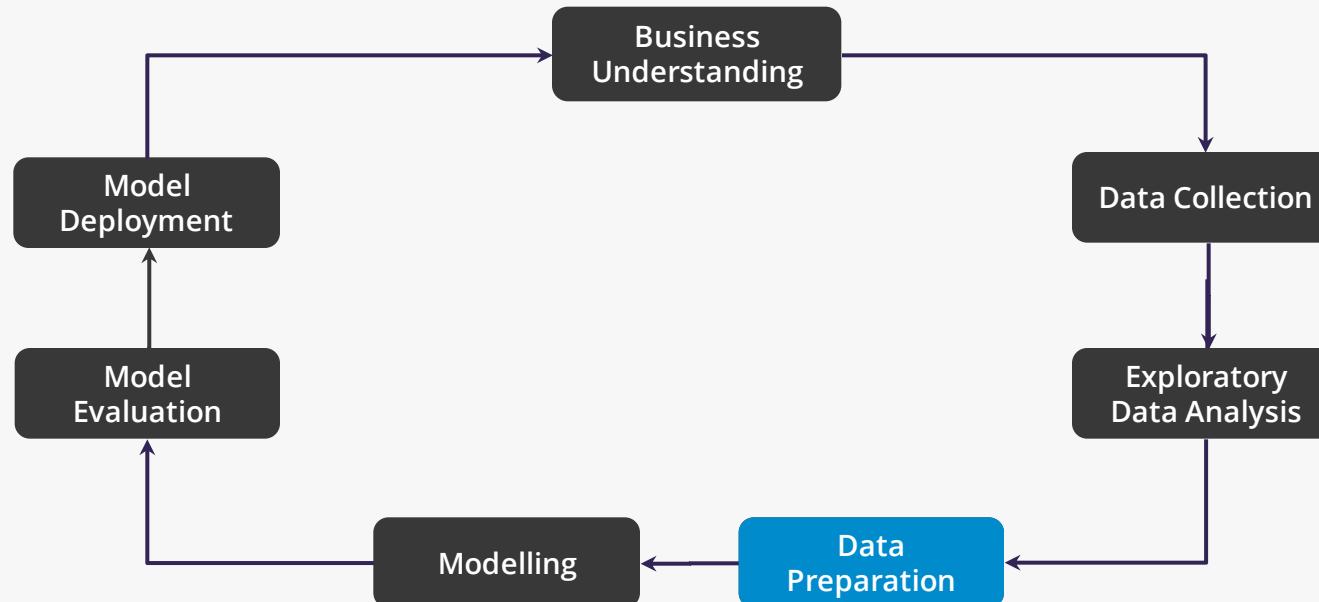
Identify important features for problem solving



In our case:

To find which factors impact the Units_sold.

Machine Learning Workflow

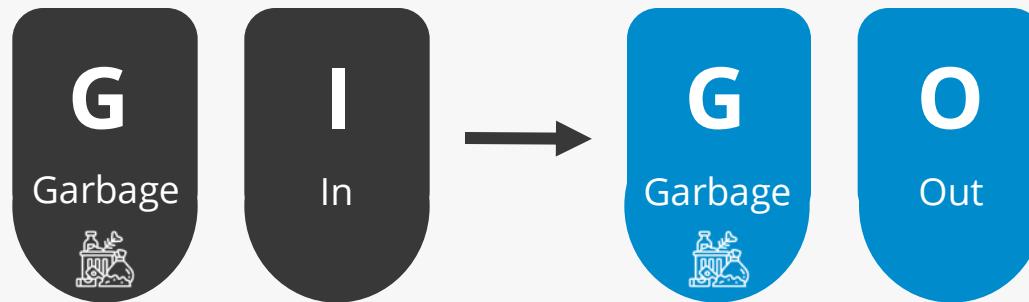


Data Preparation



This step ensures that the data is accurate, consistent and ready for modelling

Data Preparation



Data Preparation

- 1 Data Division
- 2 Data Cleaning
- 3 Data Restructuring
- 4 Data Transformation
- 5 Feature Engineering

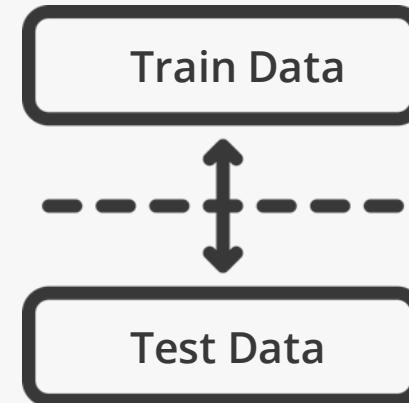
Data Preparation

1

Data Division



This process improves the reliability of the model



Data Preparation

2

Data Cleaning



Removing invalid values



Imputing missing values



Treating outliers



Example: Imputing missing values for Page_traffic in POS_Data with mean of overall traffic

Data Preparation

3

Data Transformation



Converting data from one format to another



Example: Converting the segment column in pos_data into integers

Data Preparation

4

Data Restructuring



Combining data from multiple sources to easily analyze and work with it

Table A		
Category	Sub-Category	Value
Product A	Category 1	100
Product A	Category 2	200
Product B	Category 1	150
Product B	Category 2	300
Product C	Category 1	200
Product C	Category 2	400
Product D	Category 1	250
Product D	Category 2	500

Example: Combine multiple tables for Synergix Solutions to analyze it easily.

Data Preparation

5

Feature Engineering



Create new features by combining existing ones

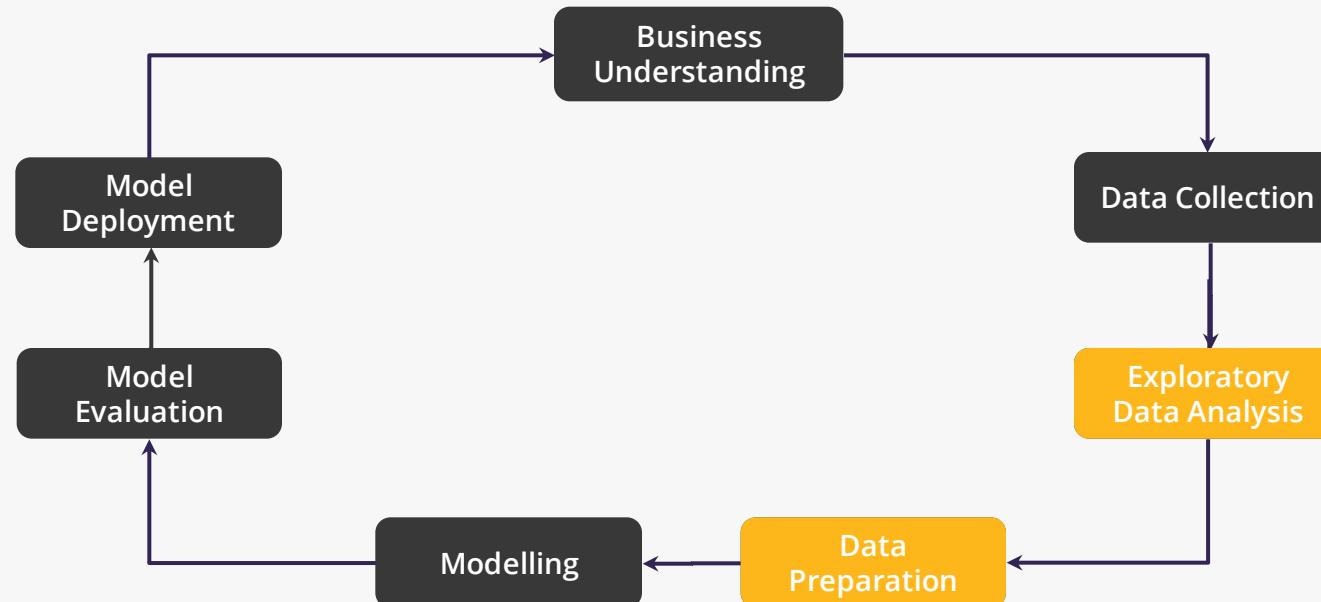


Requires domain knowledge

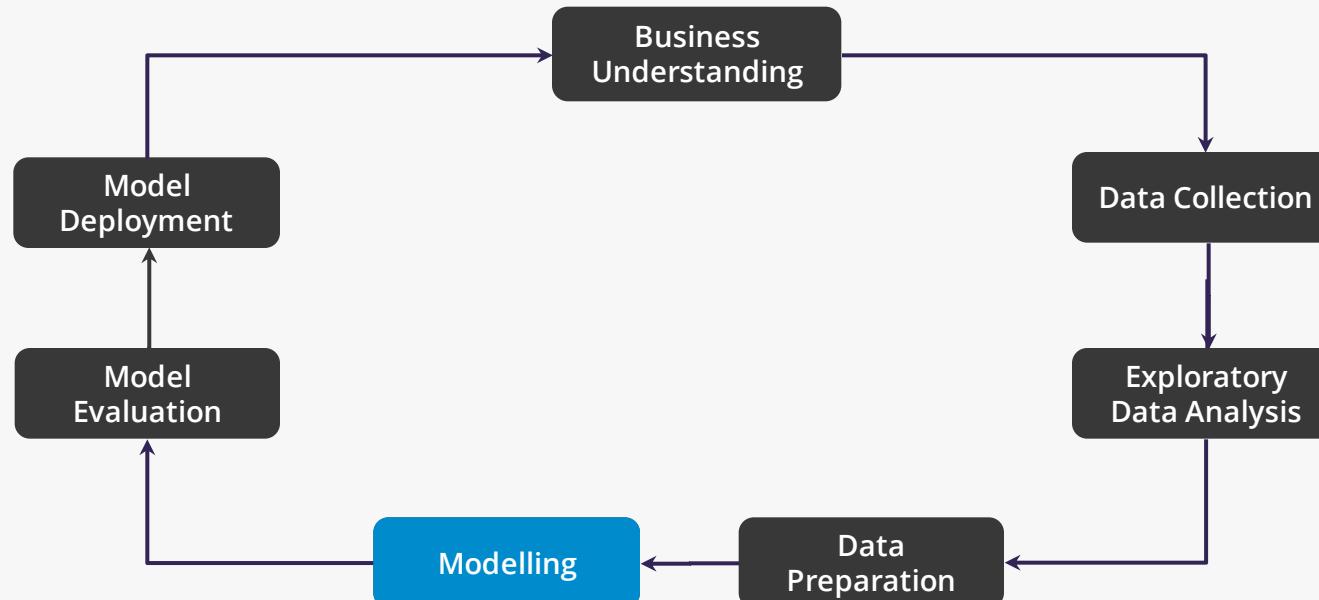


Example: Create a new feature called click through rate

Machine Learning Workflow



Machine Learning Workflow



Modelling



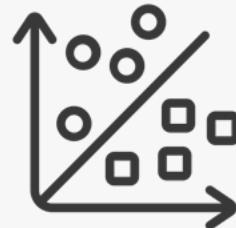
Try different ML algorithms and train the model using the prepared dataset

Modelling

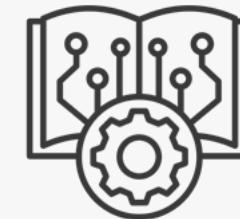
3 Problem Statements



Regression Model

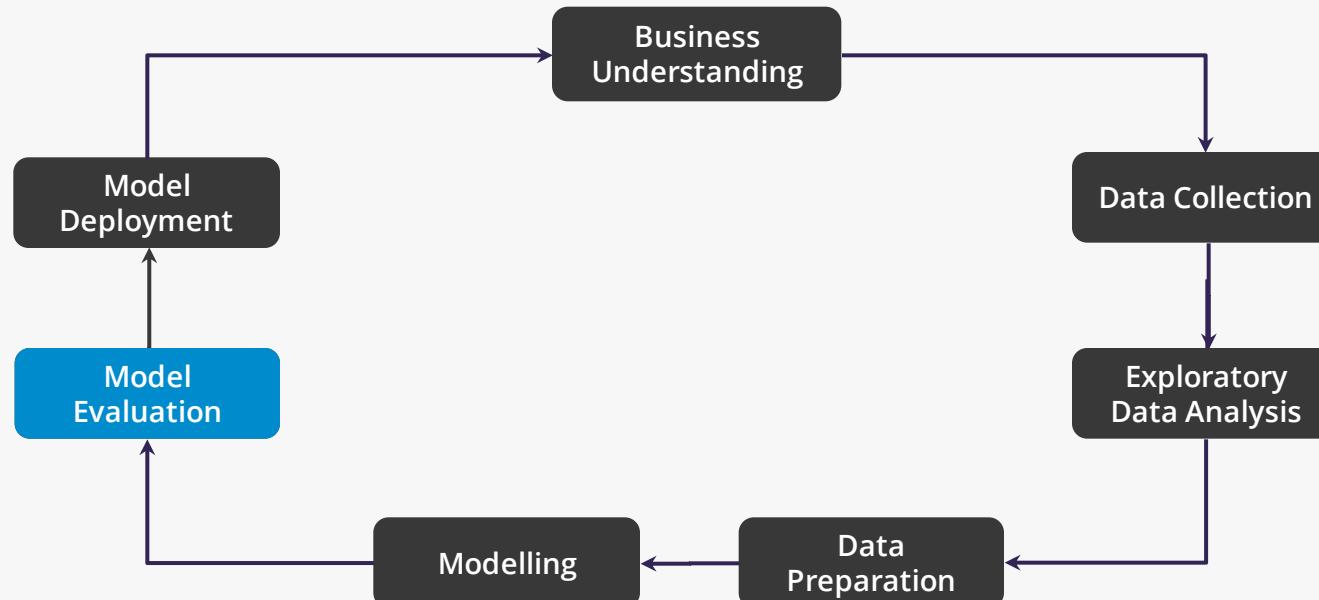


Classification Model

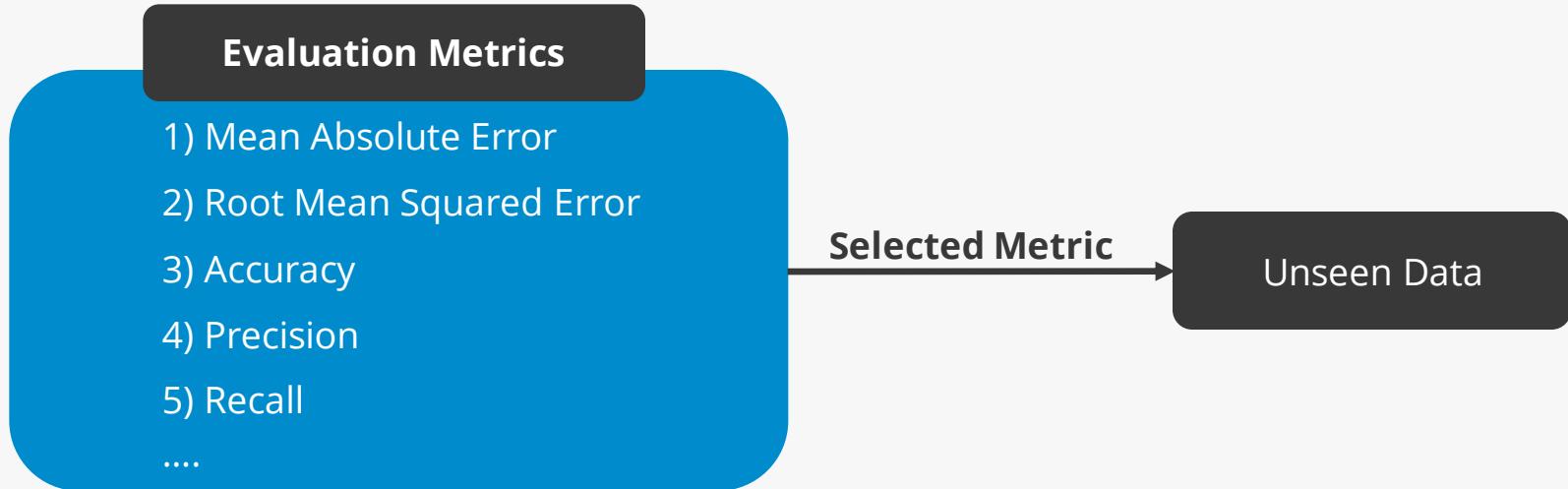


Unsupervised ML Model

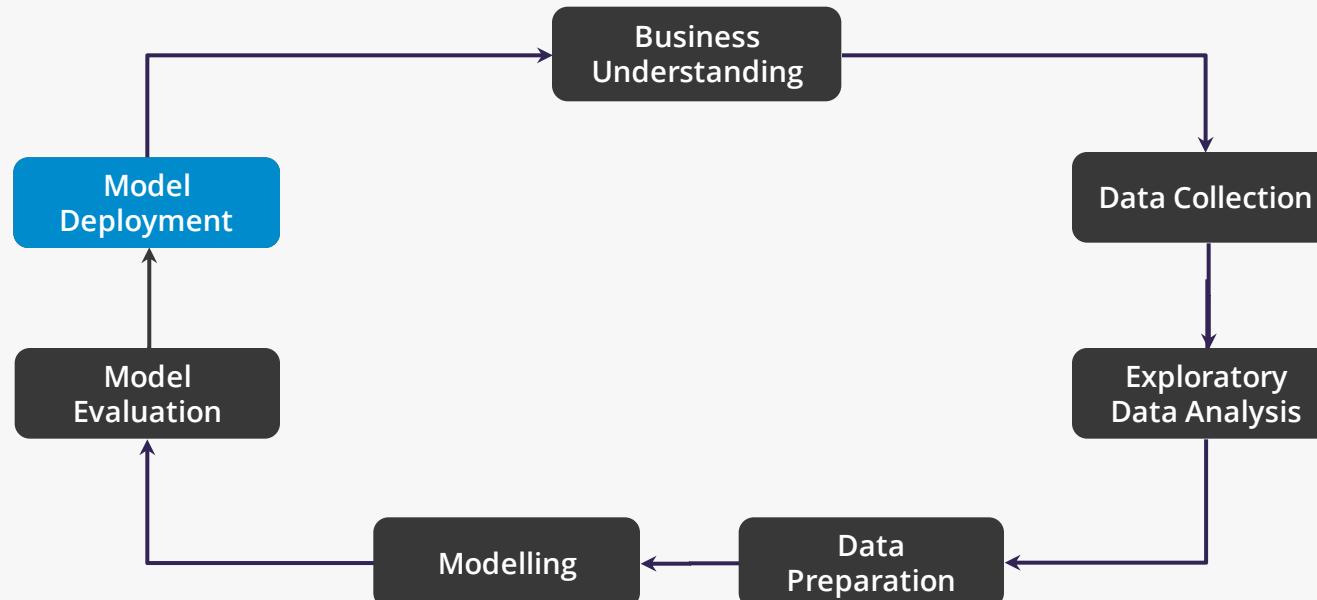
Modelling



Machine Learning Workflow



Machine Learning Workflow

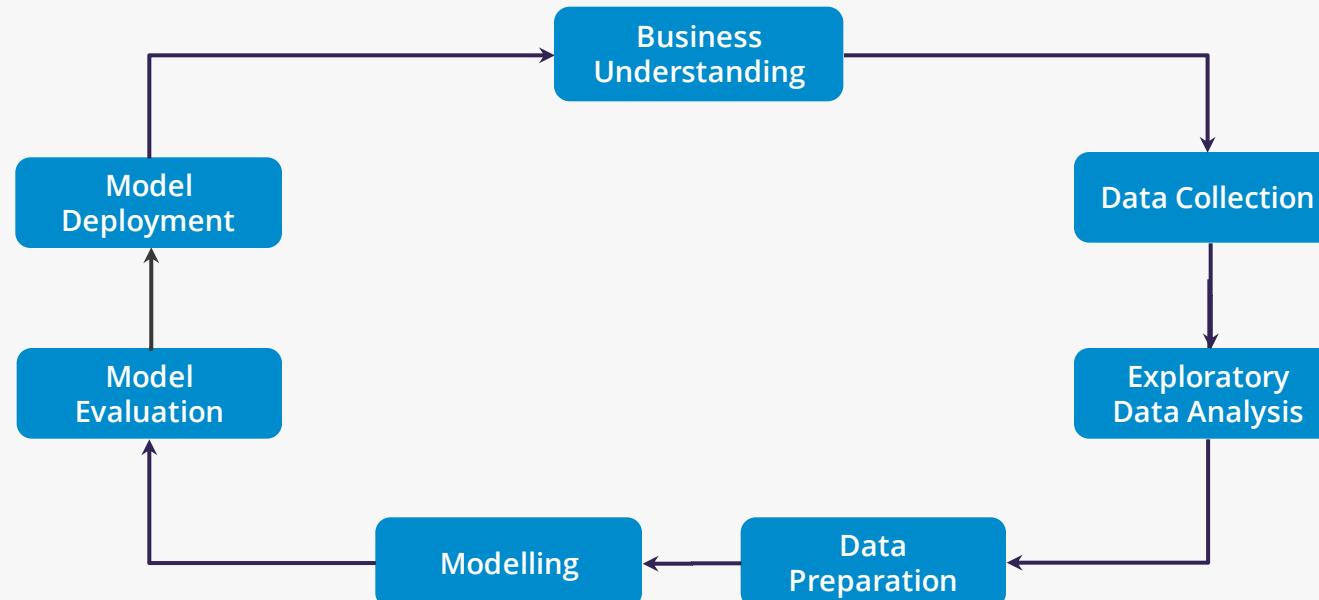


Model Deployment



In this step the trained and evaluated model is deployed or integrated into the eco-system.

Machine Learning Workflow



Machine Learning Workflow

