

# FOOD SALES PREDICTION PROJECT

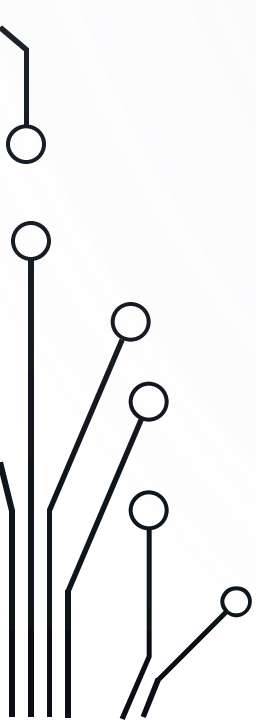
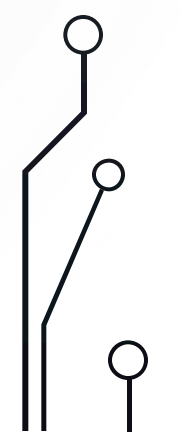
MARICEL ANDAL





# AGENDA

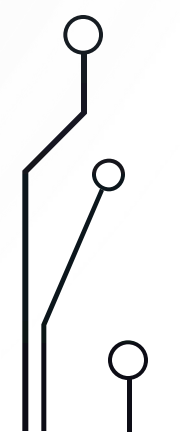
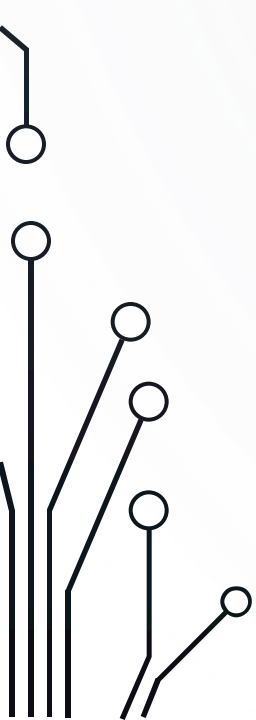


- Background
  - Problem Statement
  - Objectives
  - Dataset
  - Predictive Modeling Results
  - Recommendation
- 
- 



## BACKGROUND

As analyst, we are tasked to help the retailer understand the properties of products and outlets that play a crucial roles in increasing sales.

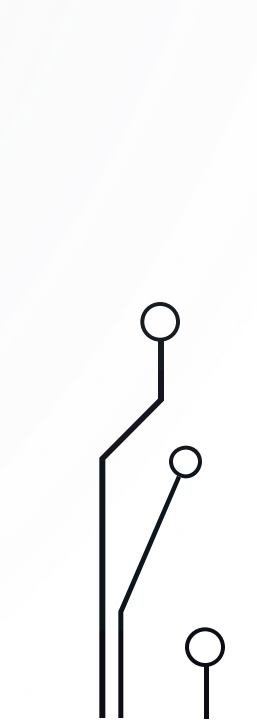
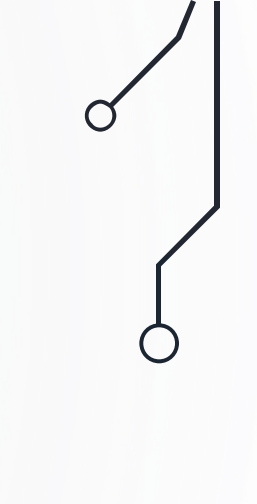
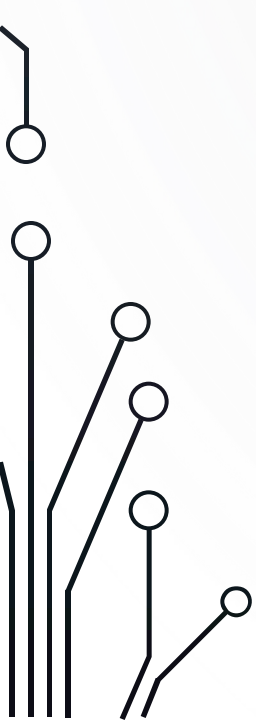




## **PROBLEM STATEMENT**

What is the current state of the business and what are the sales drivers?

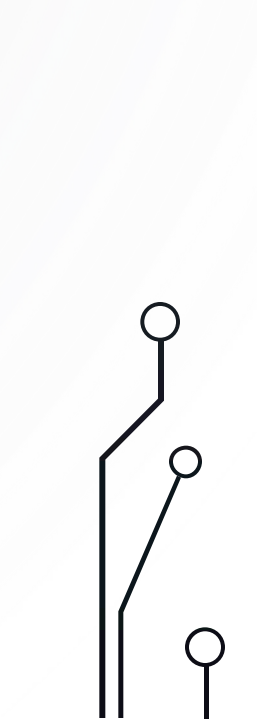
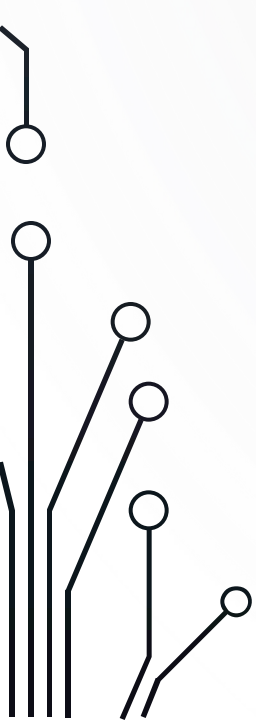
As retailers, where will we get additional business growth (sales) and by how much?





## OBJECTIVES

To identify and uncover new sales opportunities for our stakeholders (customers and retailers) and to better understand the trends in our data crucial in increasing sales.





# DATASET

8,523

Total Number of Rows  
(2023)



Source: sales\_predictions\_2023.csv\*

DATA USED	DESCRIPTION
Item_Weight	Weight of the items being sold
Item_Fat_Content	Whether low fat or regular
Item_Visibility	Visibility of the items being sold
Item_Type	Food by categories sold
Item_MRP	Maximum Retail Price (MRP) of the items being sold
Outlet_Size	Whether small, medium or high
Outlet_Location_Type	Whether tier1, tier2 or tier3
Outlet_Type	Whether grocery, supermarket1, supermarket2, supermarket3
Item_Outlet_Sales	Sales

\*Notes:

- The purpose of our model is to figure out its relationship to the randomness so we exclude 'Item\_Identifier', 'Outlet\_Identifier' because they are irrelevant columns.
- Outlet\_Establishment\_Year has datatype: datetime

# PREDICTIVE MODELING RESULTS

Metrics	BEST FIT MODEL Decision Tree		Baseline		Linear Regression	
	Training	Testing	Training	Testing	Training	Testing
RMSE	1081	1062	1719	1675	1207	1173
R2	0.60	0.60	0.00	-0.00	0.51	0.51



DECISION TREE MODEL : **GOOD FIT**

# RECOMMENDATION

- Based on the results, **DECISION TREE REGRESSION** beat the baseline and the Linear Regression models.
- We recommend to use it in our food sales prediction, results show it has **HIGHER R2** of 61% for both training and testing datasets.
- Moreover, the model has **LOWER RMSE** or root mean square error (training and testing data have RMSE of 1081 and 1062, respectively)..
- Higher r2 and lower RMSE values show that the Decision Tree model makes more accurate predictions and fits the data well. Overall, it is a **GOOD FIT** for our **FOOD SALES PREDICTION**.