TOYOTA

ITCS453 DATA WAREHOUSING AND DATA MINING

Noppanan Pinklao	6388043
Thanat Plongcharoen	6388044
Tanutkorn Pramualmongkon	6388152
Qiyue Shi	6388171

OVERVIEW OF THE BUSINESS CASE - GENERAL INFO

TOYOTA

Toyota Motor Corporation

- **Type:** Foreign-owned
- **Industry:** Vehicles and Components
- **Founded & Headquarter:** Toyota City, Japan (1937)
- **Founder:** Kiichiro Toyoda

Toyota Loom Company
BY **sakichi Toyoda** (1906)

Automobile **Department** was
established in Automatic Loom

First car came out in 1930
BY **Kiichiro**

Toyota Motor Corporation was
formally established **independently** in 1937

TOYOTA

Toyota's Business

- **Automotive**
  - Focus on Japan and North America
- **Financial Services**
  - Car sales financing credit card services
- **Others**
  - Invest in manufacturers develop new energy technology

EVALUATE THE PERFORMANCE

- **Quality**
- **Speed and dependability**
- **Suppliers**
- **Wide range and cost**

KEY BUSINESS PROCESSES, SCOPE AND REQUIREMENTS

TOYOTA

Goal: To analyze the information of car sales in Primorsky Krai, and in Kamchatka Krai region of Russia.

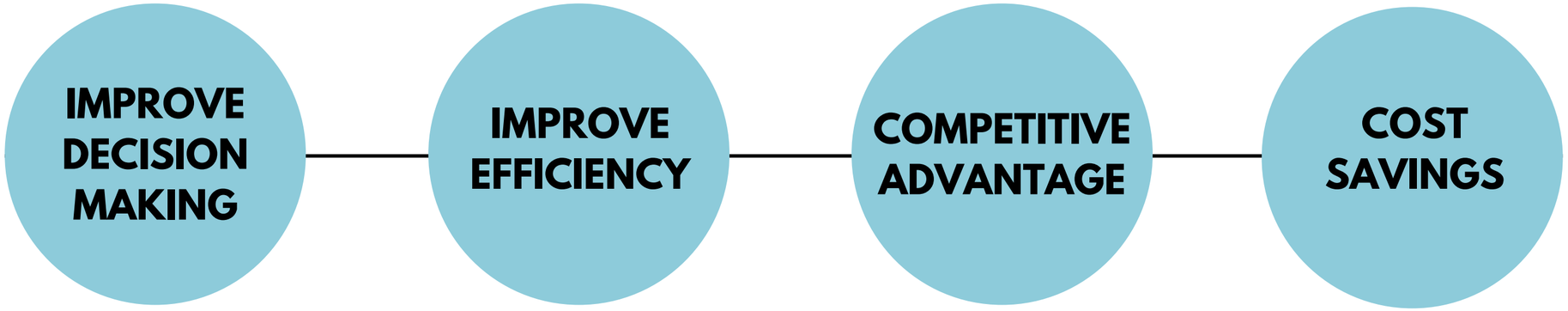
Scope:

- Descriptive Analytics
  - car models sell
  - correlation analysis between horsepower and mileage
- Predictive Analysis
  - predict the possible future outcome

Requirement:

- Identify
  - goals and End-users
- For Data
  - Sufficiently large data sets and data integrity
  - Data model
  - Data governance

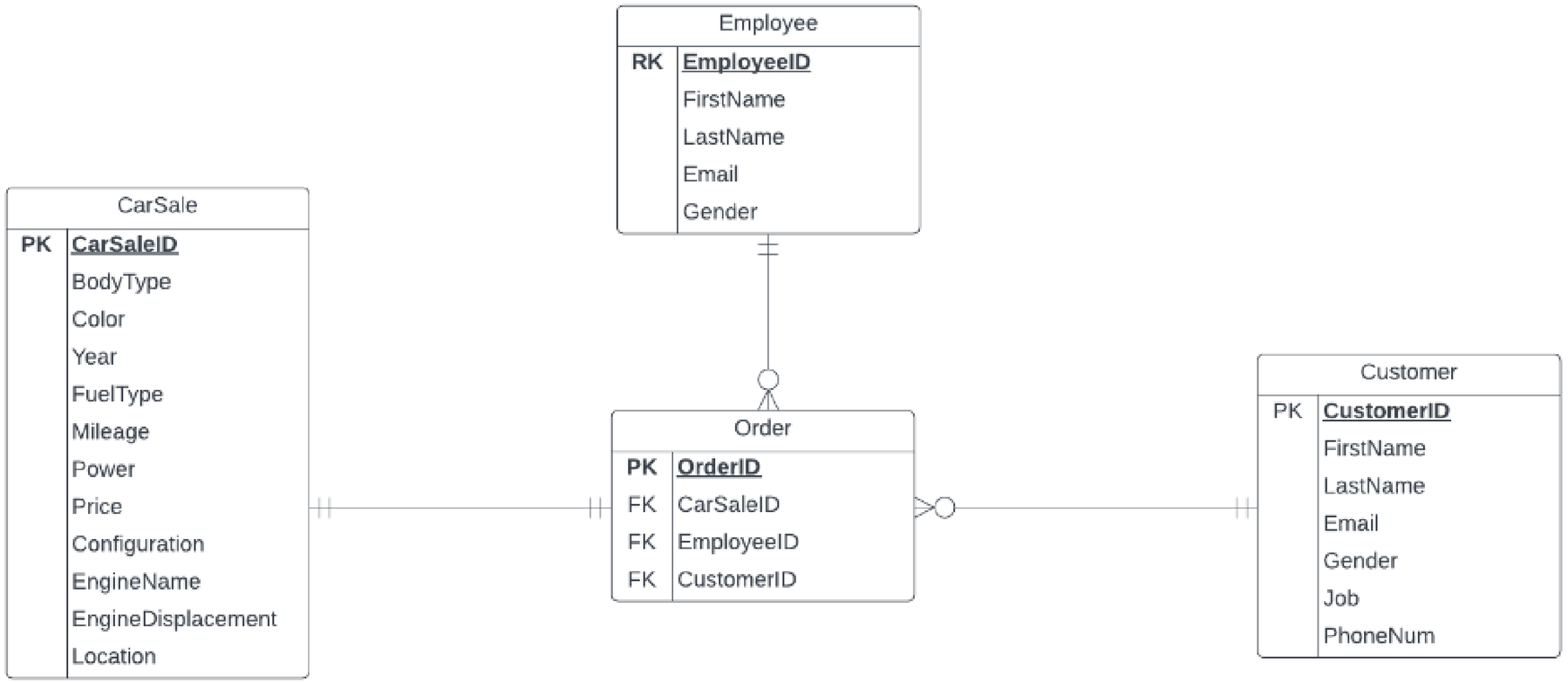
WHY DATA WAREHOUSING AND BI MATTER TO COMPANIES



Excellent data warehousing and BI provide **high-precision** decision making, help comprehend market demand, and **boost** the organization's competitiveness and customer experience.

DATA SOURCE

ERD



DATA SOURCE

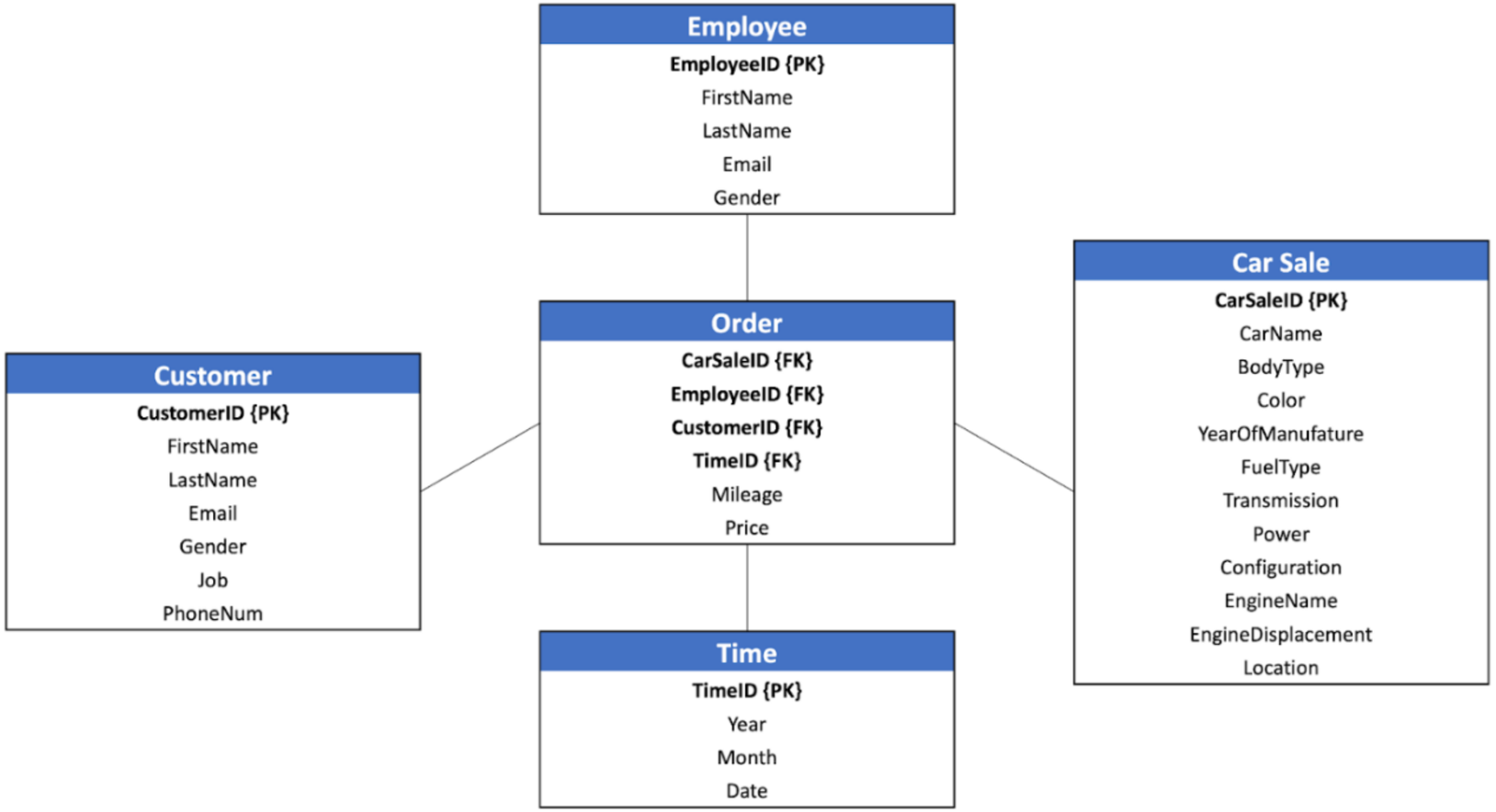
OLTP

Data dictionaries for operational database

Table Name	Field Name	Data Type	Key	FK Referenced Table	Description
CarSale	CarSaleID	int	PK		Car's unique ID
	CarName	varchar(100)			Car's name
	BodyType	varchar(100)			Body type of the car
	Color	varchar(100)			Color of the car
	YearOfManufacture	int			Manufacture year in A.D. format
	FuelType	varchar(100)			Type of fuel used, e.g. gasoline, diesel
	Transmission	varchar(10)			Transmission used
	Power	int			Horse power of the car
	Configuration	varchar(100)			The configuration of the car, e.g., 25AT, 35ATGRSport
	EngineName	varchar(100)			Name of the engine used in the car
	EngineDisplacement	varchar(10)			The measure of the cylinder volume swept by all of the pistons of a piston engine
	Location	varchar(100)			Location of the sale
Customer	CustomerID	int	PK		Customer's unique ID
	FirstName	varchar(100)			Customer's firstname
	LastName	varchar(100)			Customer's lastname
	Gender	varchar(10)			Gender of the customer
	Job	varchar(100)			Customer's occupation
	PhoneNum	int			Customer's phone number
Employee	EmployeeID	int	PK		Employee's unique ID
	FirstName	varchar(100)			Employee's firstname
	LastName	varchar(100)			Employee's lastname
	Email	varchar(100)			Email address of the employee
	Gender	varchar(10)			Gender of the employee
Order(Fact Table)	CarSaleID	int	PK, FK	CarSaleID [CarSale]	Unique ID of each specific car sale
	EmployeeID	int	PK, FK	EmployeeID [Employee]	Employee who is in charge of this specific car sale
	CustomerID	int	PK, FK	CustomerID [Customer]	Customer who boght the car
	Mileage	int			Number of miles that this car traveled
	Price	int			Price of this specific car sale

DATA WAREHOUSE DESIGN

STAR SCHEMA

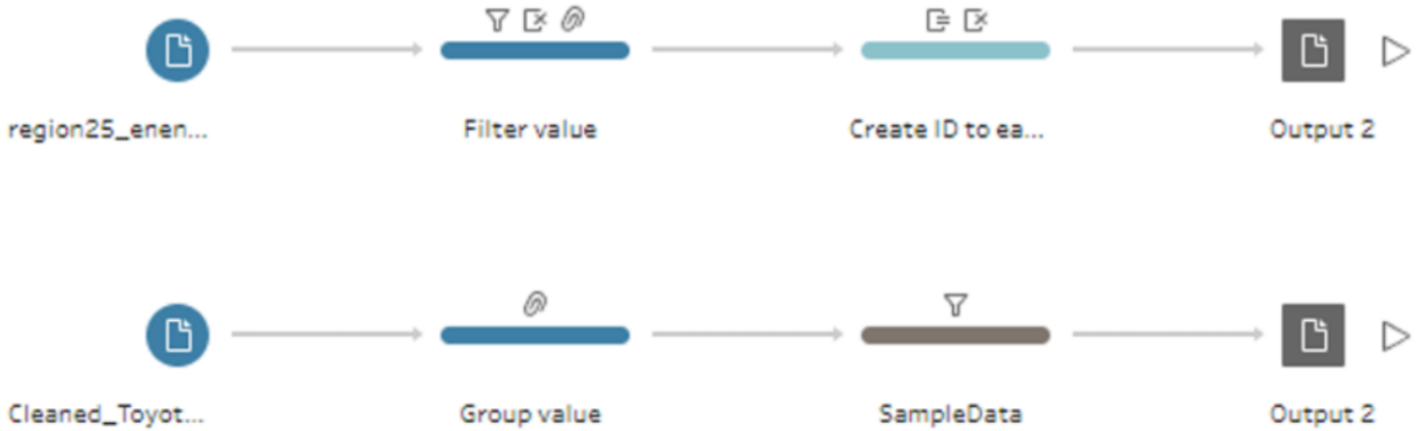


INTEGRATE DATA

ETL PROCESS

Connect to the data source

1



2

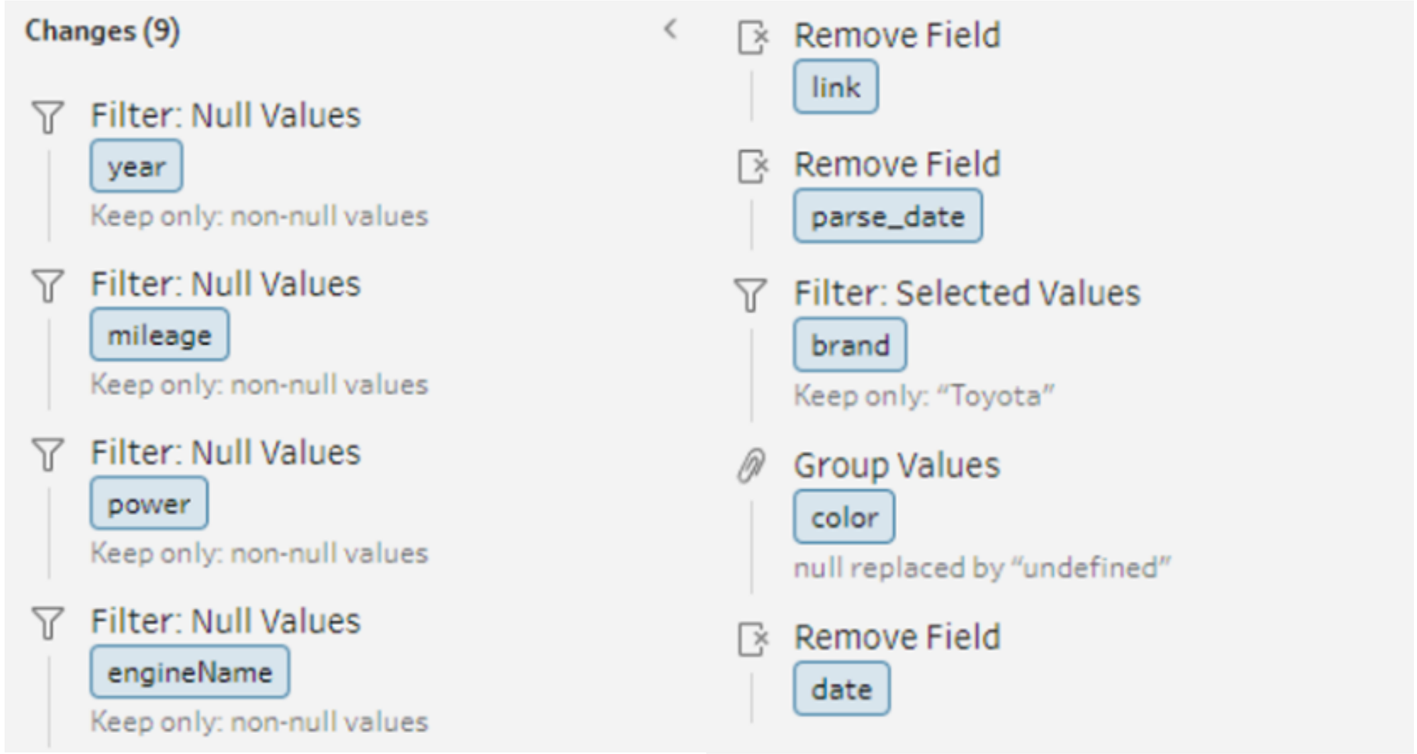
Remove column that we do not want and filter null value out

INTEGRATE DATA

ETL PROCESS

Remove column that we do not want and filter null value out

2



3

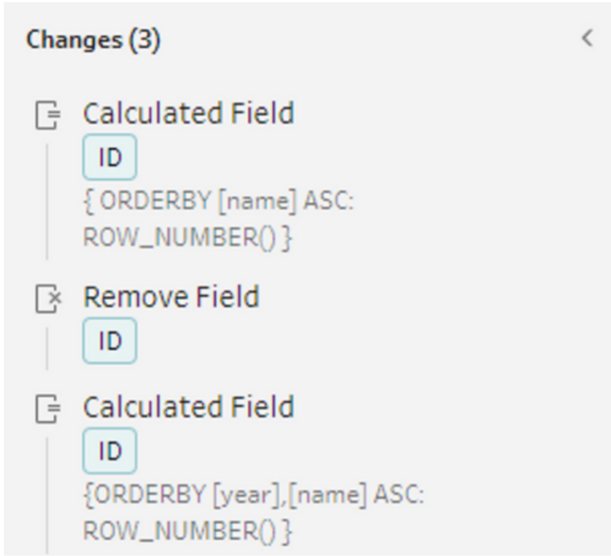
Assigning ID to each row

INTEGRATE DATA

ETL PROCESS

3

Assigning ID to each row and Grouping data



4

5

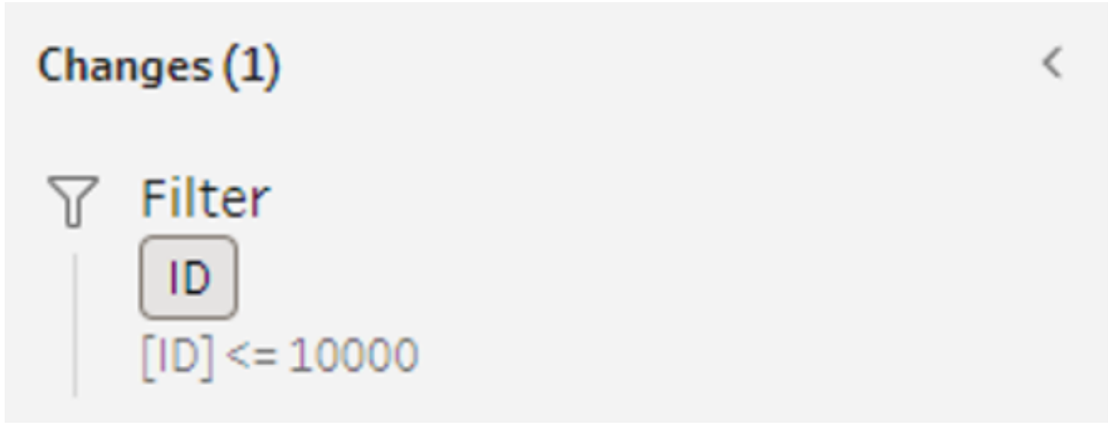
Sample Data

INTEGRATE DATA

ETL PROCESS

Sample
10000 data

5



6

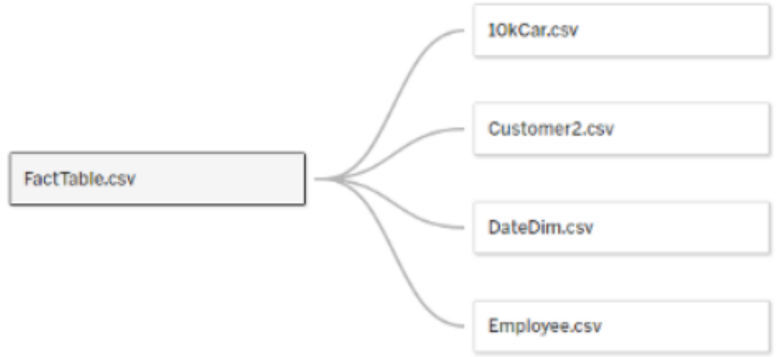
Connect each dimension to the fact table

INTEGRATE DATA

ETL PROCESS

Connect each dimension to the fact table

6



FactTable.csv — 6 fields 10000 rows

Name: FactTable.csv

Fields

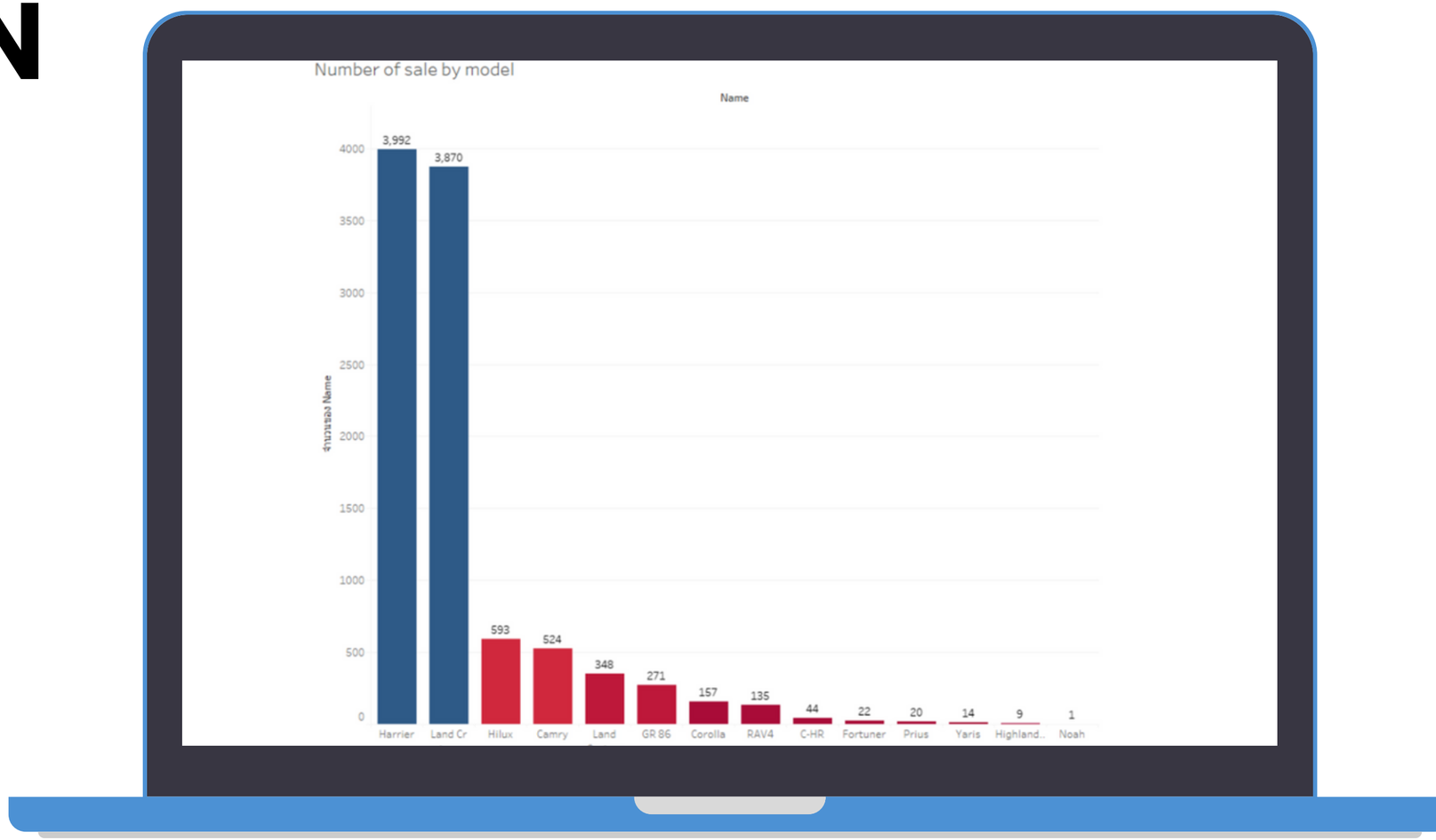
Type	Field Name	Physical Table	Remote Field N...
#	Car Sale ID	FactTable.csv	CarSaleID
#	Employee ID	FactTable.csv	EmployeeID
#	Customer ID	FactTable.csv	CustomerID
🗓	Time ID	FactTable.csv	TimeID
#	Mileage	FactTable.csv	Mileage
#	Price	FactTable.csv	Price

	FactTable.csv Car Sale ID	FactTable.csv Employee ID	FactTable.csv Customer ID	FactTable.csv Time ID	FactTable.csv Mileage	FactTable.csv Price
1	8556769	11122	20/5/2565	1,000	2,800,000	
2	5930274	71919	1/5/2565	10,000	3,860,000	
3	4891519	46400	2/1/2565	2,000	4,500,000	
4	5945578	28848	12/10/2565	10,000	3,860,000	
5	2906369	57469	15/4/2565	10,000	3,860,000	
6	2467951	94819	4/6/2565	20,000	3,200,000	
7	1845903	744	7/9/2565	20,000	3,200,000	
8	4389138	90709	24/1/2565	20,000	3,200,000	
9	6498799	28294	3/3/2565	20,000	3,200,000	
10	2927277	66643	7/7/2565	20,000	3,200,000	
11	8786823	8059	14/4/2565	20,000	3,200,000	

TOYOTA

ANALYSIS AND VISUALIZATION

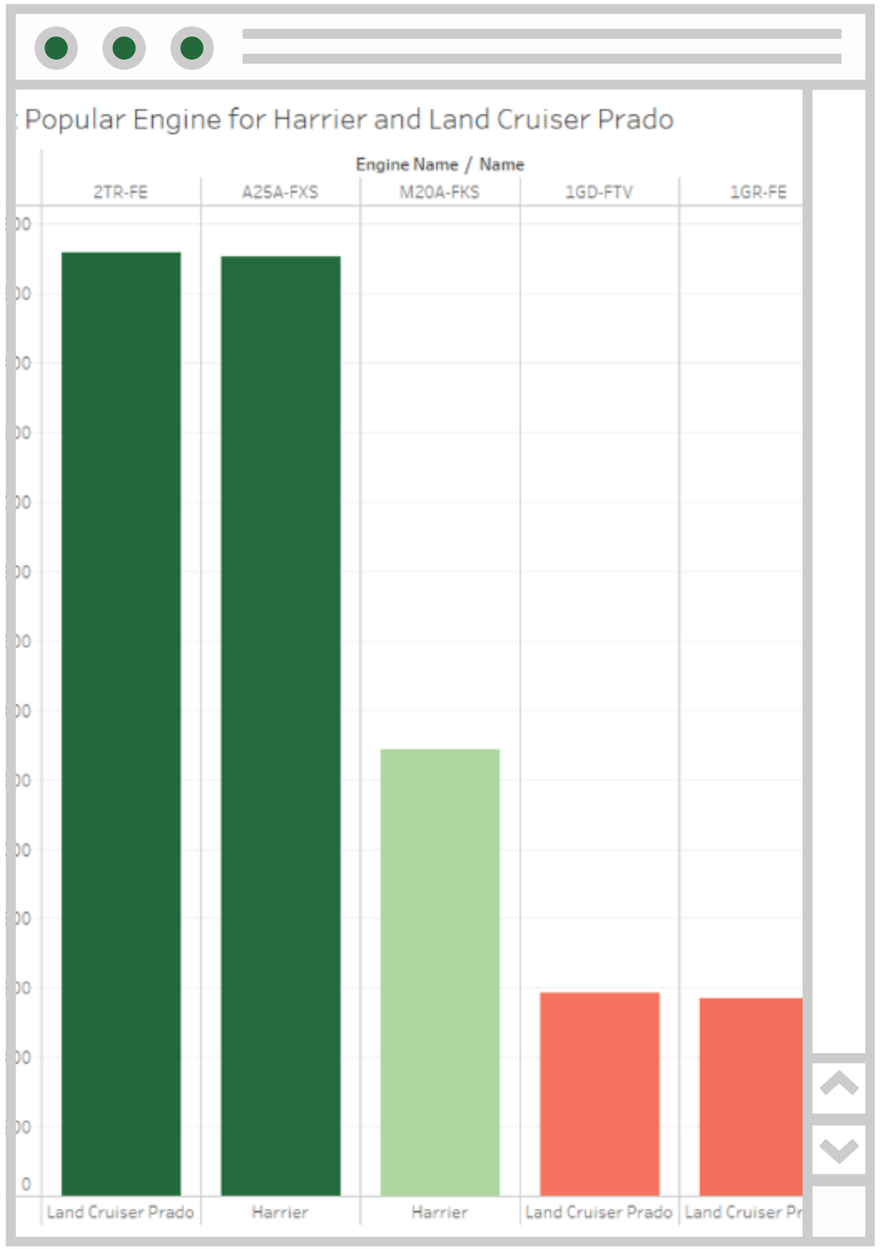
The best-selling car



TOYOTA

ANALYSIS AND VISUALIZATION

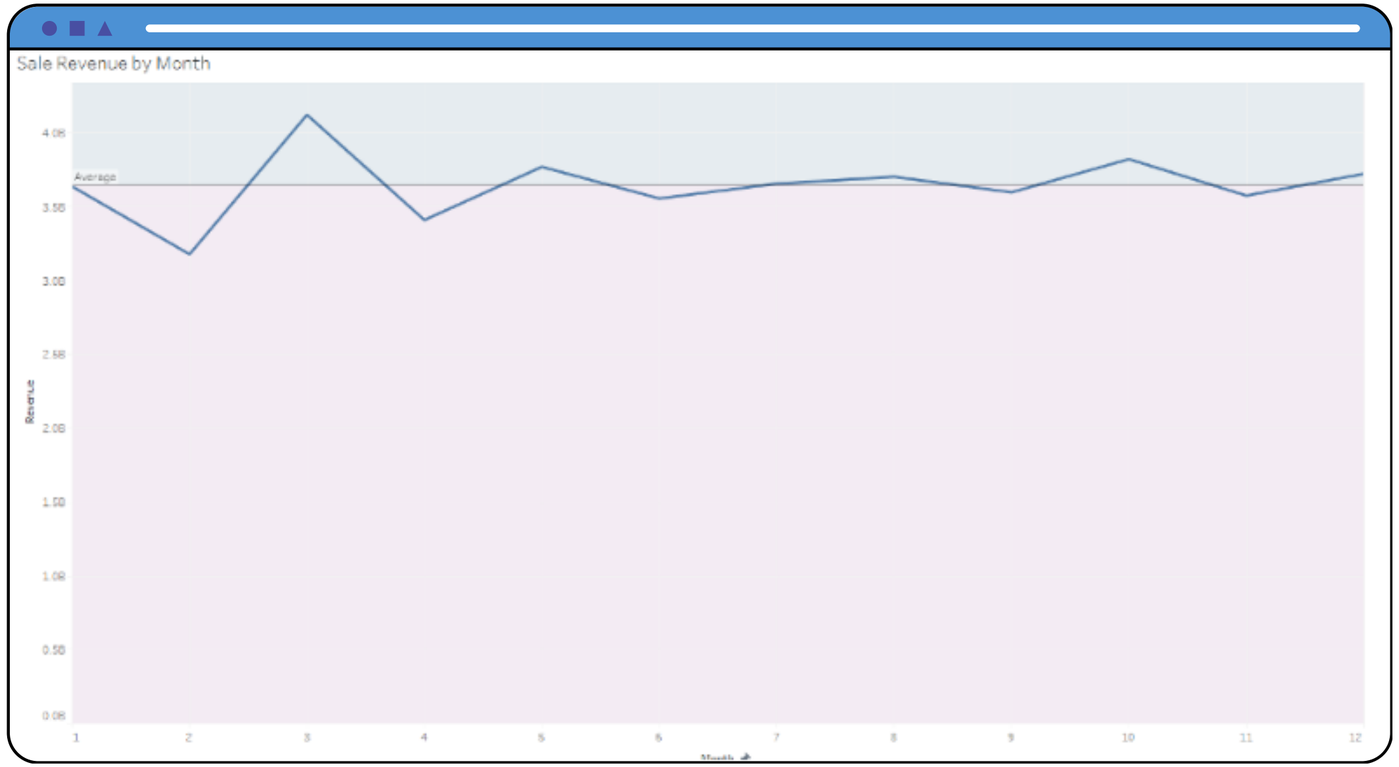
The most popular engine for Harrier and Land Cruiser Prado



TOYOTA

ANALYSIS AND VISUALIZATION

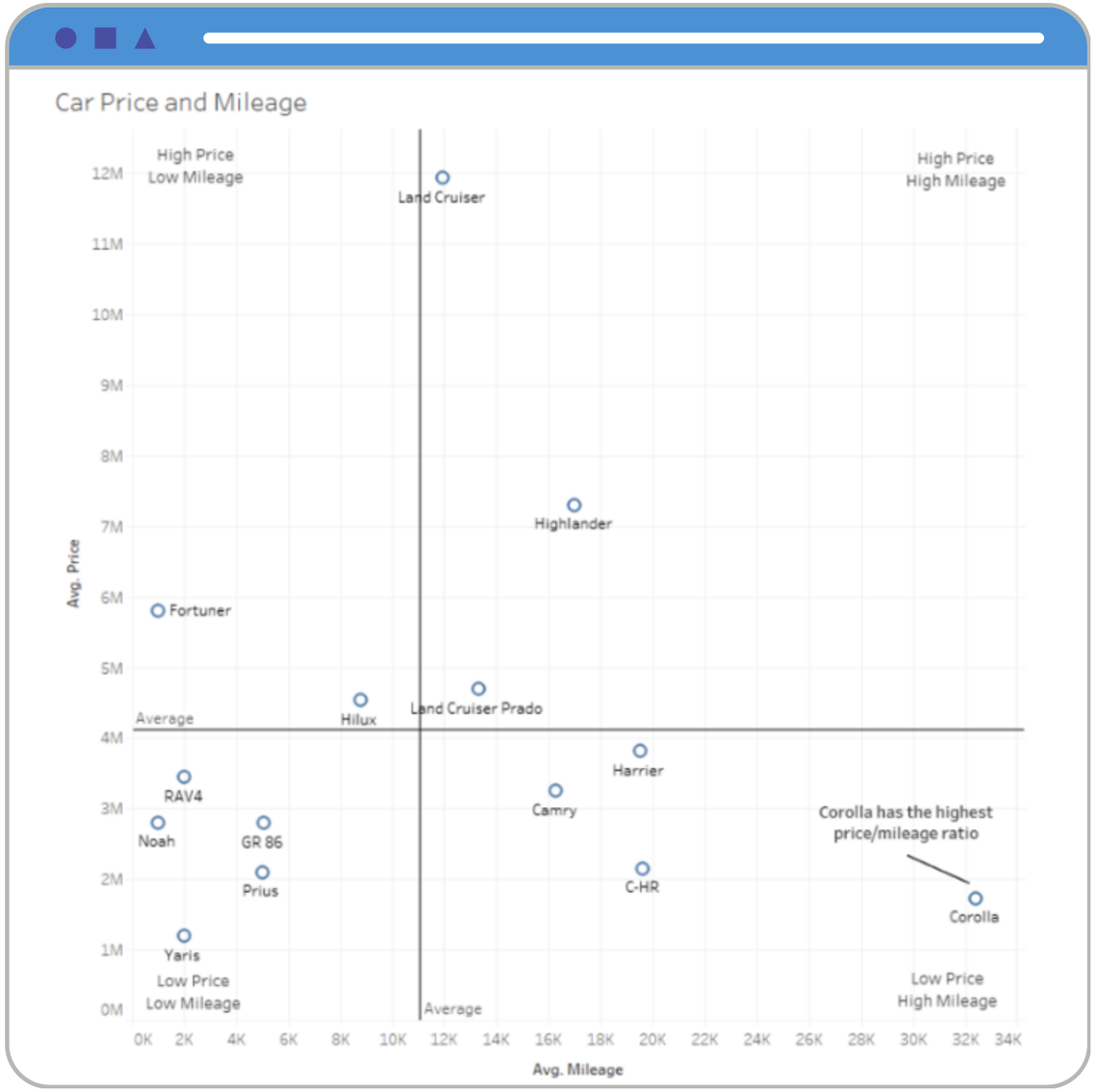
Which month had the highest number of sales



TOYOTA

ANALYSIS AND VISUALIZATION

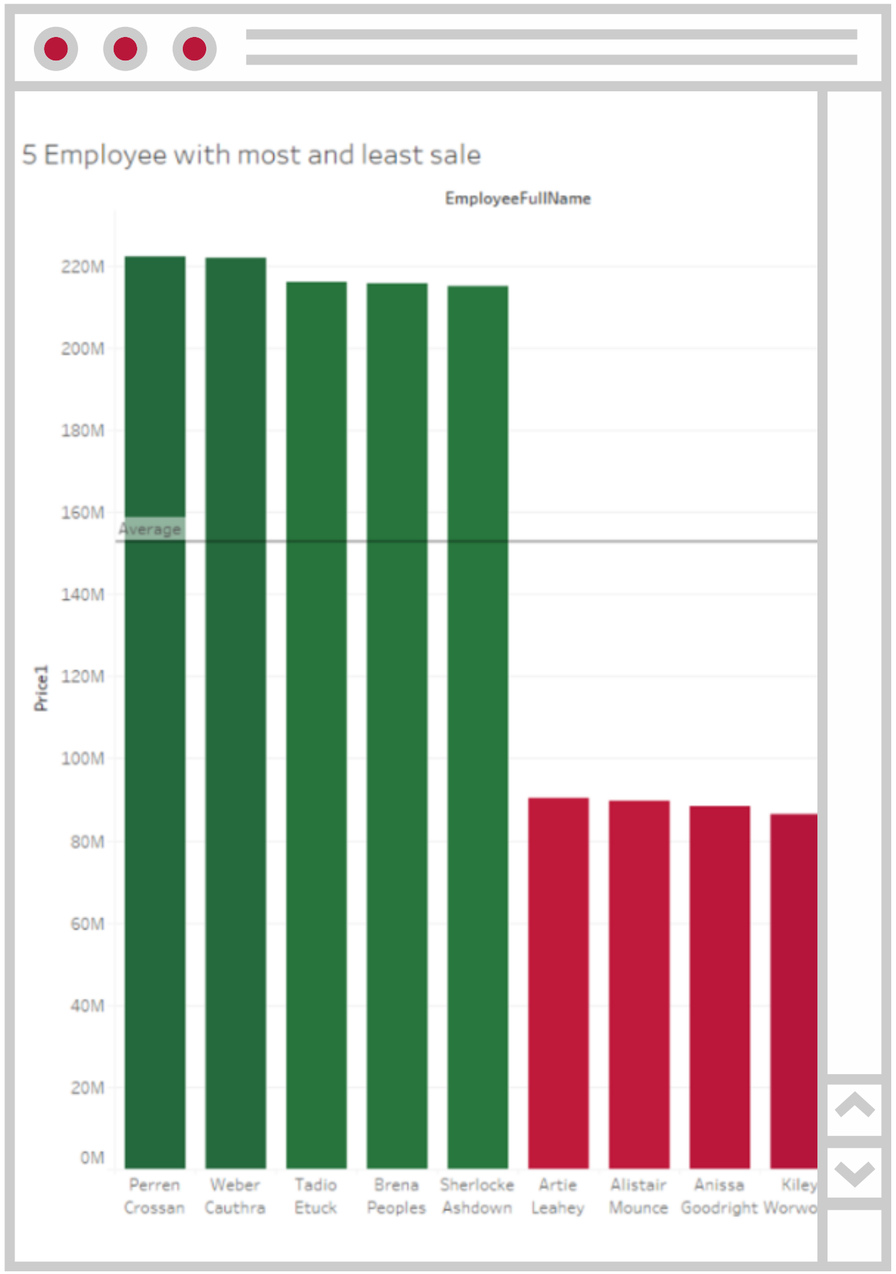
Make a tougher car with a cheaper price. This visualization shows correlation between price and mileage



TOYOTA

ANALYSIS AND VISUALIZATION

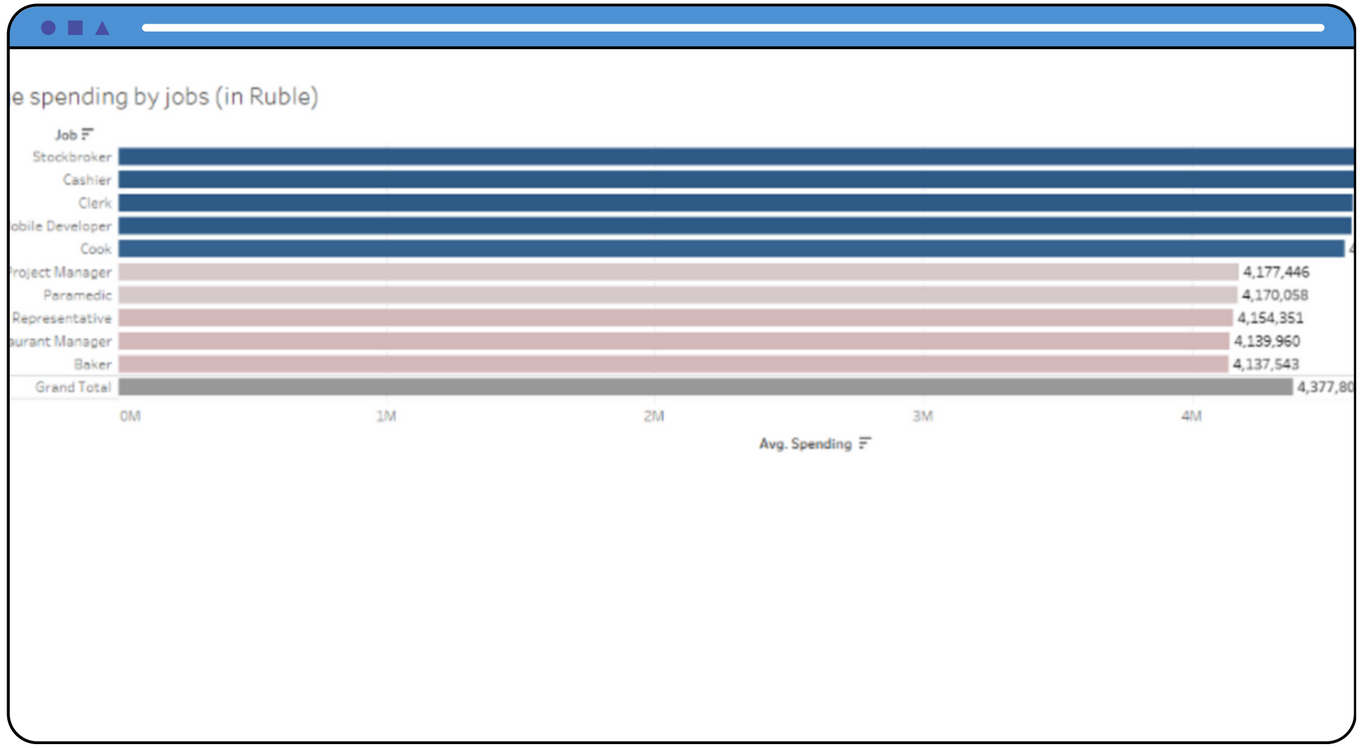
Measure employees' performance



TOYOTA

ANALYSIS AND VISUALIZATION

Advertise better on their phone. This visualization shows the average spending of each customer, grouped by jobs



DISCUSSION AND CONCLUSION

IMPROVE DECISION MAKING

Help them analyze data they have and make decisions based on their information.

IMPROVE EFFICIENCY

Process their data and use them to reduce manual effort, eliminate errors and prioritize their employees to focus on their higher-value tasks.

COMPETITIVE ADVANTAGE

Gain insights into market trends, consumer behavior, and competitor activity .

COST SAVING

Improve its operations and cut expenses by identifying inefficiencies and potential areas for improvement.

THANK YOU