# Understanding and Visualizing Data

**Course Project**

#### **Part One – Draft a Data-Collection Plan**

This part identifies a situation that requires to make a decision, identify what data will be used to inform the decision, and draft a plan to collect the data we need.

Step 1: Situation that requires a decision

|  |
| --- |
| *Sri Venkateshwara TVS is a two-wheeler sales showroom,*  *always approaches to have a long-lasting relationship with its customers.*  *Being the licensed motorcycle dealer of Rasipuram, they sell*  *motor cycle, moped and scooter categories.*  *Every vehicle will have a model number based on its year of production. The showroom has 2020-models at present, and by January'21, 2021-models will be under sales. The customers always prefer to buy new models. So, the dealer wants to make sure whether most of the 2020 model vehicles were sold by January’21. As the dealer planned for a year end sale, he wants to decide which vehicle type needs more and which vehicle type needs less discount based on the sales done in the last 4 months of 2019-2020. Discounts on prices should not disturb their investment for the next quarter.*  *The dealer needs,*   1. *Sales growth analysis report* 2. *The top 5 selling vehicle categories (Post-lockdown analysis)* 3. *Vehicles in stock list* 4. *Discount price list* |

Parameters or options in the decision

|  |
| --- |
| *Growth rate (sales growth) in order to make more investment in the next quarter is the main parameter considered in the decision* |

Key stakeholders

|  |
| --- |
| *Dealer and customers are the key stakeholders.* |

Expectation on data to help illuminate the decision

|  |
| --- |
| *I expect the dataset to have both the sales data and the price list. With the sales data, we can find the top sellers and stock details. Based on the price list, we can identify the discount range from highest to lowest.*  *The goal is to clear 2020-model vehicles to welcome 2021-models. With the above findings, the decision will be precise and supports the showroom in two-wheeler direct sale(year-end sale).*  *The decision will also help the dealer to sign up with new customers.* |

Step 2: Identifying data that helps in better understanding of the situation.

What are the key performance indicators for your situation?

|  |
| --- |
| *Orders, Sales, Total Sales (INR)(columns/variables) are the key performance indicators.* |

What defines the range of cases that are considered?

|  |
| --- |
| *Month and vehicle types defines the range of cases. In this case we have 23 models and months of Sept, Oct, Nov, Dec. Range will be around 90-100* |

What are the variables you will consider?

|  |  |
| --- | --- |
| 1. *Category* | *5. Month* |
| 1. *Sub-Category* | *6. Sales* |
| 1. *Model* | *7. In stock* |
| 1. *Orders* | *8. Total Sales (INR)* |

Categorical or quantitative variable

|  |  |
| --- | --- |
| *Categorical*  *Variables* | *Category* |
| *Sub-Category* |
| *Model* |
| *Month* |
| *Quantitative*  *Variables* | *Orders* |
| *Sales* |
| *In stock* |
| *Total Sales (INR)* |

Purpose of each variable in informing the decision

|  |  |
| --- | --- |
| *Order date, Order, Sales variables* | *To monitor, evaluate, and analyze sales performance.* |
| *Category, subcategory, model, Total Sales, in Stock variables* | *To identify top categories in revenue.*  *To identify vehicle types in stock.* |
| *Orders, Sales Variables* | *Calculated value on orders and sales based on this variable helps us to finalize the appropriate discount range at the end.* |

Step 3: Developing a data-gathering plan.

Included [

Where will the data come from?

Is it observational or experimental data?

Who will collect it?

How much data will be needed (sample size)?

How to assure that it is representative of the population?

What steps are taken to mitigate potential bias?]

|  |
| --- |
| *Data collection will take place at Sri Venkateshwara TVS sales showroom, Rasipuram. When a vehicle gets sold, the salesperson enters the details in an excel sheet. The AD will supervise, and the dealer will monitor. I have planned to collect this observational data for the months September, October, November and December. 90+ total observations have to be collected. With 25 - 30 cases from every month of this quarter. Under the vehicle categories, we have around 23 models. So, the data collection for every month should have 23 entries. Additional entries are for any special editions. The dealer wants to analyze the sales done from Sep to Dec. The data includes only the sales done in the showroom. So, the collected data is representative of the population. Educating the sales persons with the details about each vehicle. This will make the customer to increase their interest in the features available in each vehicle. To ensure accuracy in data collection, each entry should be checked twice before moving on to the next entry. This will reduce the possible bias in data collection.* |

#### **Part Two – Identify Data Summaries and Visualizations**

This part will identify the statistical summaries and visualizations believed will best help to make our decision.

Summary statistics to inform the decision

|  |
| --- |
| *Average sales(monthly) and Number of vehicles sold for years 2018-2019 and 2019-2020. Respective summary statistics are provided in Sales2018-2019 and Sales2019-2020 sheets respectively in TVS\_SalesData.xlsx* |

Statistics that are sensitive to or resistant to outliers.

|  |
| --- |
| *Statistics that are sensitive to outliers will affect our predictive outcomes, and the decision might be riskier. Statistics resistant to outliers will help the data to make the right decision. So, I am interested in statistics resistant to outliers.* |

Visualizations to inform the decision.(Sample)

|  |
| --- |
| * *Visualization - Orders vs Sales comparison (2019-2020) (Detailed visualizations are provided in the TVS\_SalesData Excel workbook.* |

#### **Part Three – Data and Your Decision**

This part will decide about whether the data we have or planned to have been going to be sufficient to make a good decision.

Part a: About the data-model-insight framework

|  |
| --- |
| *With this, I am attempting to show sales growth comparison, to identify the best sellers, and help the dealer to find the right solution to improve his sales in the last quarter of 2020.* |

KPIs for the situation we are trying to understand.

|  |
| --- |
| *Our aim is to increase the sales growth comparatively more than the previous year. For this, we need to increase the vehicle sales and to clear the 2020-models at present. KPIs are the average sales (compare years 2019 and 2020), in-stock vehicles and ways to attract more customers to buy vehicles.* |

Relationship between our variables and the KPIs?

|  |
| --- |
| *Sales, Sales (in INR) will show us the sales growth of the four months. On comparison with in-stock analysis, it will help to find new ways to attract customers to buy vehicles.* |

What are the limitations of your model?

|  |
| --- |
| *Missing or error data, scope of the data (sep-dec 2019, 2020), should include first-hand sales and not second-hand or exchange offer vehicles*. |

Do you feel your model, as defined, is “good enough” to inform your decision? Why or why not?

|  |
| --- |
| *The model clearly visualizes the comparative sales growth of the years 2018-2019 and 2019-2020. Also having a bar chart of top sellers and stock alert, it makes it easier to allot discount ranges for the required models. Monthly sales growth (category, sub-category and models) will help the dealer to best analyze the sales growth in this post-lockdown period. Based on the orders-sales comparison graph I have increased and decreased the discount percentage. So, the present model will be ‘good enough’ to inform my decision.* |

Part b: Your Project Dashboard

If you have data for your project, you may be able to put together a dashboard in Excel that is a good working model. If so, include an Excel workbook as part of your project submission.

As an alternative, you may turn in a mockup that indicates what elements you would like your dashboard to include.

In either case, your dashboard or mockup should include

* a readout of KPI’s, clearly labeled as such
* one or more visualizations
* summary statistics, as needed

Indicate how the elements of the dashboard are connected. In a mockup, this could be arrows drawn between different elements on the dashboard. For a dashboard with actual data, you might include this information as text.

|  |
| --- |
| *I have attached the visualizations (Graphs and dashboard) needed in* ***VISUALIZATIONS (with decision) sheet*** *in the attached* ***TVS\_SalesData.xlsx Excel*** *workbook.* |