

TheAnalyticsTeam

# **Sprocket Central Pty Ltd**

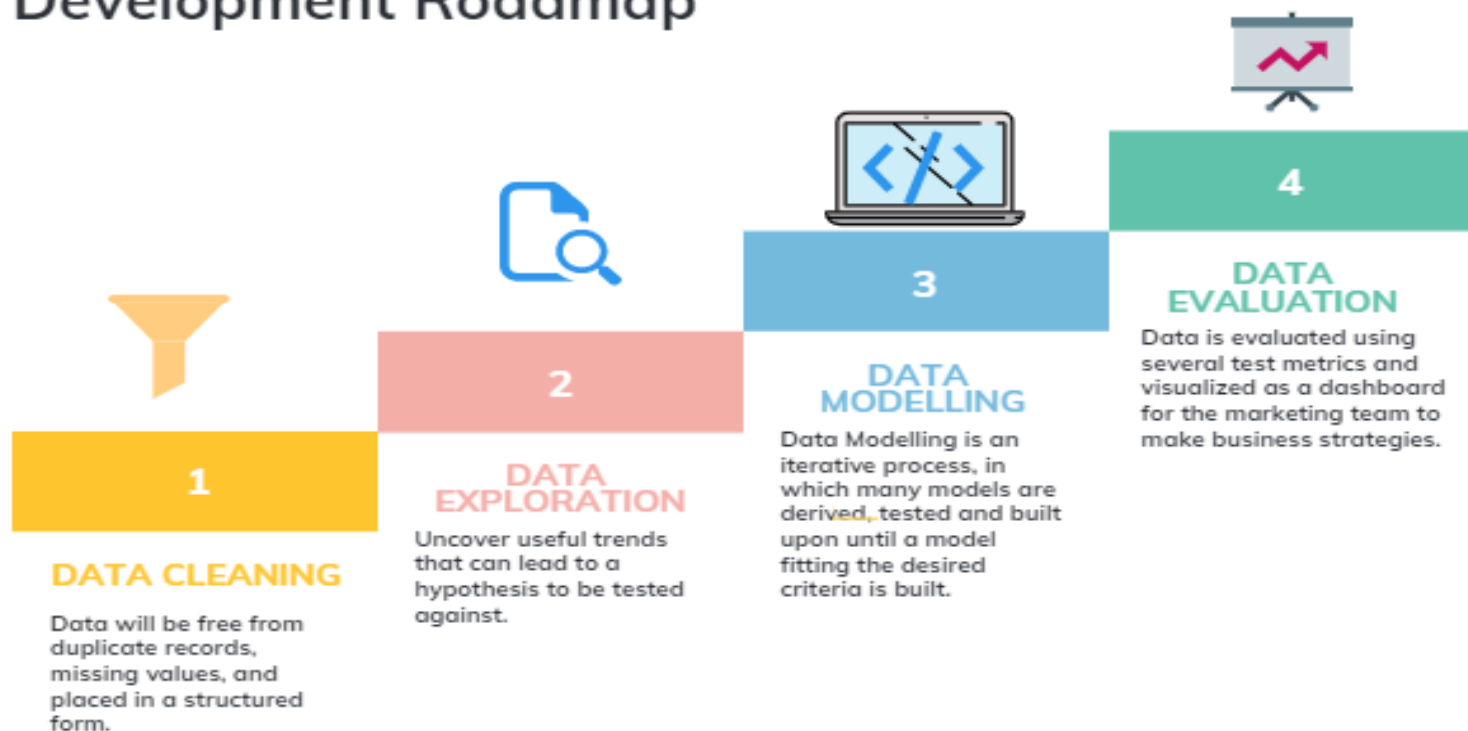
Data analytics approach

[Data Analytics Team] - [Engagement Manager], [Senior Consultant], [Abdelrahman Akmal]

# Agenda

1. Introduction
2. Data Exploration
3. Model Development
4. Interpretation

## SPRocket Central Pty LTD. Development Roadmap



# Data Cleaning

It is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.



**Removing  
Duplicates**

**Fix  
structural  
errors**

**Filter  
unwanted  
outliers**

**Handle  
missing data**

**Validate &  
QA**

# Data Exploration

**Exploratory Data Analysis is a data analytics process to understand the data in depth and learn the different data characteristics, often with visual means.**

```
graph LR; A[Univariate Analysis] --> B[Bivariate Analysis<br/>1. Numeric – Numeric<br/>2. Numeric – Categorical<br/>3. Categorical - Categorical]; B --> C[Feature Engineering]
```

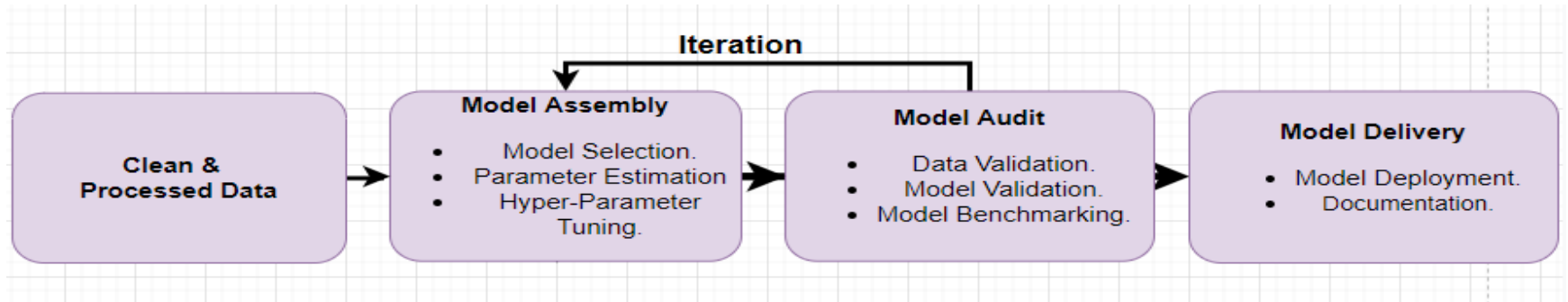
**Univariate Analysis**

**Bivariate Analysis**  
1. Numeric – Numeric  
2. Numeric – Categorical  
3. Categorical - Categorical

**Feature Engineering**

# Modeling

It is an iterative process, in which many models are derived, tested and built upon until a model fitting the desired criteria is built.

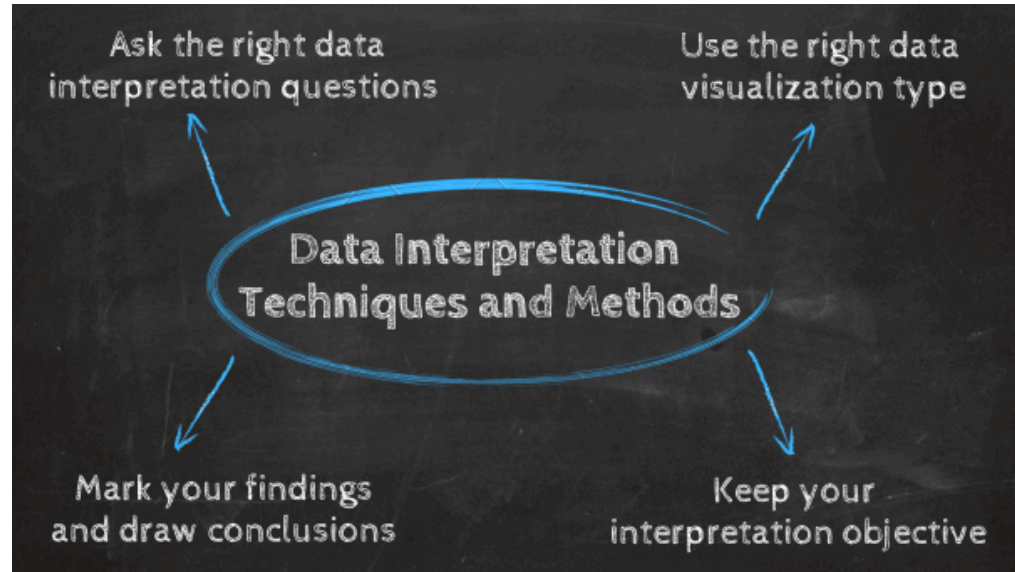


- The first step is to form the hypothesis about our defined problem.
- We use a statistical models to test that hypothesis, for every model we tune its parameter to get the best performances.
- Evaluate the performance of the models using the suitable metrics like ROC Curve, precision, recall, F1 Score, and accuracy.
- Finally, we select the model with the best performance deploy it and make

# Interpretation

**Model interpretability is to understand and explain the steps and decisions a machine learning model takes when making predictions**

- We will be reporting the findings in form of visualizations which will demonstrate customers demographics, targeted customers, and top selling goods to help to make decision-making process to be more reliable .
- We will also interpret our predictive models in a very simple model that will its understanding process easy regardless the background of recipients.



**Thank You !**