

# Analysis of Placer.ai Dataset

## Introduction

The objective of this analysis was to evaluate the potential value of the Placer.ai dataset in enhancing our investment due diligence process. By leveraging customer segmentation and predictive modeling, we aimed to uncover actionable insights that could inform strategic decision-making and optimize marketing efforts.

## Methodology

### Data Preprocessing

```
# Import useful libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
# Load data to pandas
reader = spark.read.option("header", "true")
root_path = "/Volumes/dev/placer/imports"
metrics_df = reader.csv(f"{root_path}/metrics").toPandas()
metadata_df = reader.csv(f"{root_path}/metadata").toPandas()
```

## Exploratory Analysis

- Initial Data Inspection:

Conducted an initial exploration to understand the structure and content of the data. The data includes various sales metrics and metadata essential for deeper analysis.

- Descriptive Statistics:

Computed summary statistics to get an overview of the central tendency and dispersion of the data. Used visualizations such as histograms and box plots to identify data distributions and potential outliers.