

# **GUVI - Project**

# **Walmart Sales Analysis**

# **Pipeline**

11 November 2024

Soumya Sankar Panda

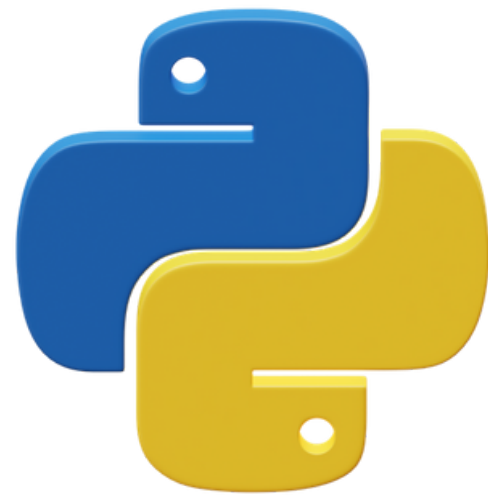
# Project Overview

This project demonstrates a data pipeline for analyzing Walmart sales and customer data using PySpark. Through the pipeline, we answer key business questions such as :

1. total sales by state
2. most purchased products and monthly sales trends.

These are some of the queries ..

# Technologies Used



+



+



# Outcomes

1. Using PySpark Do the Data Analysis
2. The analysis is achieved by performing SQL-like transformations

# Project Flow

