Sales Analysis 2019 - Project Documentation (PACE Framework)

# 1. Planning

## Business Objective

Strategic Business Goal:

Driving smarter decision-making and accelerate revenue growth by transforming transactional sales data into meaningful business insight

Anaytical Business Goals:

* Identifying best-selling product by revenue and volume
* Reveal Sales trends by time and geography
* Enable data driven actions in marketing and inventory planning

## Problem Statement

* Which month has the best sales? Why?
* Which city has the best sales? Why?
* Which product should be increase the marketing activity?
* Identify trendline revenue?
* Identify golden time to make advertisement?

## Scope of Analysis

* Sales data from 2019 January to 2019 December
* List of variables:

+ Product

+ Quantity Order

+ Price each

+ Sales

+ Order date, Hour, Month

+ City

* Key analyses performed include:

+ Trendline sales by month

+ Sales by city

+ Sales by product

+ Sales by hour per day

# 2. Analyzing

## Dataset Overview

* Dataset Original: Kaggle
* Number of Objectives: 371,901
* Key attribute:

|  |  |  |  |
| --- | --- | --- | --- |
| Attribute | Data Type | Meaning | Analytics Role |
| Order ID | String | Unique identifier for each transaction | Used to distinguish individual sales; not analyzed deeply |
| Product | String | Name of the product sold | Analyze best-selling and highest-revenue products |
| Quantity Ordered | Integer | Number of units purchased | Analyze total units sold; identify high-volume products |
| Price Each | Float | Unit price of the product | Used to calculate total sales revenue |
| Order Date | Datetime | Date and time when the order was placed | Analyze sales trends over time (month, day, hour) |
| Purchase Address | String | Delivery address of the order | Extract city information for location-based analysis |
| Sales (add more) | Float (calcualated) | Total revenue per transaction | Analyze total revenue across dimensions |
| Month (add more) | Integer (extracted) | Month of the order | Analyze seasonal trends and monthly sales performance |
| Hours (add more) | Integer (extracted) | Hour of the order | Analyze hourly purchasing behavior |
| City | String (extracted) | City of delivery | Analyze geographical distribution of sales |

## Data Preprocessing Steps

* Merging 12 – month dataset into 1 file “annualSales2019.csv”
* Extracting “Month” column
* Get rid of ‘NaN’ and ‘ERROR’ value

## Exploratory Data Analysis (EDA)

Key findings from the EDA process include:

* December recorded the highest total sales volume
* San Francisco was the city with the largest sales contribution
* Macbook Pro Laptop generated the most revenue among all products, but with few quantity
* Most sales occurred between 18:00 to 20:00, indicating peak shopping hours

## Key Analyses Conducted: Analyzing Sales trend based on

* Month
* City
* Product
* Hour

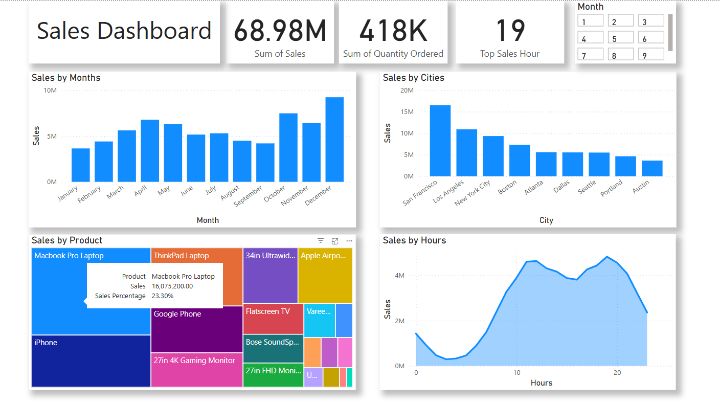
# 3. Constructing (Building a Visualization)

## Dashboard Description

The dashboard was built to provide a comprehensive and intuitive overview of sales performance throughhout 2019.

It consolidates key business metrics and visualization that allow users to quickly identify trends across time, location, product lines and purchasing hours.

This interactive dashboard serves as the foundation for extracting actionable business insights.



The dashboard consist of:

* KPI Cards: Total Sales, Total Orders and Greatest Sales Hour
* Column Chart: Sales Trends across Months & City
* Treemap: Sales Contribution by Product
* Linechart: Sales Performance by Hours
* Slicer: Filter all the chart and KPI Cards base on Month

## Why these visuals were chosen

|  |  |
| --- | --- |
| Visual | Reason |
| KPI Cards | Summarize key metrics at a glance, allowing users to grasp the overall performance within the first few seconds. |
| Line Chart (Sales by Month) | Clearly illustrates the sales trend over time, making seasonality and fluctuations easy to observe. |
| Tree map (Sales by Product) | Enables a proportional comparison of sales contributions across different products. |
| Line Chart (Sales by Hours) | Analyzes customer purchasing behavior by hour, helping to identify peak sales periods. |
| Column Chart (Sales by City) | Displays sales distribution across cities in a straightforward, easy-to-compare format. |

# 4. Executing

## Key Insights & Recommendation

(Ghi 3-5 điểm rút ra quan trọng nhất.

(Dựa trên insight, đề xuất những hành động cần thiết.)

1. December is the highest revenue time

Sales peaks in December, indicating strong seasonal demand likely driven by holiday shopping behaviour

* Launch targeted holiday campaigns and promotions starting from late November to maximize December sales potential.

1. San Francisco dominates in sales revenue.

Among all cities, San Francisco contributed the largest share to total sales, suggesting a strong market presence in this area.

* Prioritize marketing efforts and inventory allocation in San Francisco to strengthen market leadership and boost profitability.

1. MacBook Pro Laptop generates the most revenue.

This product alone accounts for the highest proportion of overall sales, highlighting the importance of high-end electronics.

* Focus on bundling offers, upselling accessories, and extended warranty packages for MacBook Pro customers to further increase revenue.

1. Evening hours (18:00–20:00) see the highest purchasing activity.

Sales spike significantly during the evening, indicating a potential optimization window for promotions or ad targeting.

* Schedule digital ads, email promotions, and flash sales during peak evening hours to capitalize on consumer shopping behavior.

1. Sales distribution across months reveals steady seasonal fluctuations.

Besides December’s peak, other months like April and May also show strong performance, implying multiple shopping seasons throughout the year.

* Develop mid-year sales events (e.g., "Summer Promotions" in April/May) to stabilize revenue flow outside the holiday season.

## Potential Business Impact

Implementing the insights and recommended actions could lead to significant improvements in business performance.  
Specifically:

* **Maximizing peak season revenue b**y strategically focusing on December and mid-year peaks, the business can capture a larger share of seasonal demand.
* **Optimizing marketing and inventory efforts** by concentrating on high-performing cities like San Francisco ensures more efficient allocation of resources.
* **Boosting high-value product sales** by leveraging the strong performance of products like the MacBook Pro Laptop can significantly enhance average order value and profitability.
* **Enhancing customer engagement:** Timing promotional activities to align with peak shopping hours increases the likelihood of conversion and customer satisfaction.
* **Sustaining steady revenue flow:** Understanding and planning around monthly sales trends allows for better cash flow management and reduces sales volatility throughout the year.

Overall, these insights empower the business to make data-driven decisions that strengthen competitive advantage, improve operational efficiency, and drive sustainable long-term growth.