Of course. I've removed the **Acknowledgments section** and all mentions of the **Kaggle competition** as you requested.

Here is the updated and final version of the README.md file for your Walmart Sales Analysis project.

**Walmart Sales Data Analysis**

**About**

This project explores Walmart Sales data to understand top-performing branches, popular products, sales trends, and customer behavior. The goal is to analyze how sales strategies can be improved and optimized for the business.

**Purposes Of The Project**

The major aim of this project is to gain insight into the sales data of Walmart to understand the different factors that affect sales across various branches.

**About The Data**

This dataset contains sales transactions from three different branches of Walmart, located in Mandalay, Yangon, and Naypyitaw. The data consists of 17 columns and 1000 rows.

| Column | Description | Data Type |
| --- | --- | --- |
| invoice\_id | Invoice of the sales made | VARCHAR(30) |
| branch | Branch at which sales were made | VARCHAR(5) |
| city | The location of the branch | VARCHAR(30) |
| customer\_type | The type of the customer | VARCHAR(30) |
| gender | Gender of the customer making purchase | VARCHAR(10) |
| product\_line | Product line of the product sold | VARCHAR(100) |
| unit\_price | The price of each product | DECIMAL(10, 2) |
| quantity | The amount of the product sold | INT |
| VAT | The amount of tax on the purchase | FLOAT(6, 4) |
| total | The total cost of the purchase | DECIMAL(10, 2) |
| date | The date on which the purchase was made | DATE |
| time | The time at which the purchase was made | TIMESTAMP |
| payment\_method | The total amount paid | DECIMAL(10, 2) |
| cogs | Cost Of Goods sold | DECIMAL(10, 2) |
| gross\_margin\_percentage | Gross margin percentage | FLOAT(11, 9) |
| gross\_income | Gross Income | DECIMAL(10, 2) |
| rating | Rating | FLOAT(2, 1) |

**Approach Used**

1. **Data Wrangling:** This is the first step where inspection of data is done to ensure NULL values and missing values are detected and handled.
2. **Feature Engineering:** This step involves generating new columns from existing ones—such as time\_of\_day, day\_name, and month\_name—to provide deeper insights.
3. **Exploratory Data Analysis (EDA):** EDA is performed to answer the listed questions and achieve the project's aims.

**Business Questions To Answer**

**Generic Questions**

* How many unique cities does the data have?
* In which city is each branch?

**Product Analysis**

* How many unique product lines does the data have?
* What is the most common payment method?
* What is the most selling product line?
* What is the total revenue by month?
* What month had the largest COGS?
* What product line had the largest revenue?
* What is the city with the largest revenue?
* What product line had the largest VAT?
* Which branch sold more products than the average product sold?
* What is the most common product line by gender?
* What is the average rating of each product line?

**Sales Analysis**

* Number of sales made in each time of the day per weekday.
* Which of the customer types brings the most revenue?
* Which city has the largest tax percent/VAT?
* Which customer type pays the most in VAT?

**Customer Analysis**

* How many unique customer types does the data have?
* How many unique payment methods does the data have?
* What is the most common customer type?
* Which customer type buys the most?
* What is the gender of most of the customers?
* What is the gender distribution per branch?
* Which time of the day do customers give the most ratings?
* Which day of the week has the best average ratings?
* Which day of the week has the best average ratings per branch?

**Author**

* **Parveen Jalwal**
* **LinkedIn**: <https://www.linkedin.com/in/parveen-jalwal-201a2a302>