

# Data Analysis Project

**Data Analyst:** Emil Perez

**Client/Sponsor:** Amazon E-Commerce Analytics Team

**Dataset Provider:** Amazon Products Sales Dataset 2023 [[Kaggle](#)]

## Purpose:

This project aims to analyze customer sentiment based on Amazon product reviews. The goal is to identify key trends, issues, and opportunities for improving customer satisfaction and product performance. This is all testing data for learning and development reasons.

## Scope / Major Project Activities:

What are the major parts of this project? List out the high-level steps, activities, or stages of the project, and give a brief description for each.

Activity	Description
Data Collection	Extract Amazon reviews for specific product categories (e.g., electronics, clothing) from the database.
Data Cleaning	Remove duplicates, handle missing values, and preprocess text for analysis.
Sentiment Analysis	Apply Natural Language Processing (NLP) techniques to classify review sentiments as positive, negative, or neutral.
Visualization	Create dashboards showing trends by product, category, and geography.
Reporting	Generate actionable insights and recommendations based on the analysis.

## This project does not include:

Specify the things that this project isn't responsible for doing (out of scope). For instance, "this project does not involve a summation of 2019 data analysis"

- Analysis of reviews outside the year 2023.
- Predictive modeling for future reviews.
- Integration of insights into the client's live systems.

---

### Deliverables:

A specific list of things that your project will deliver.

Deliverable	Description/ Details
Cleaned Dataset	A structured and cleaned dataset ready for sentiment analysis.
Sentiment Dashboard	Interactive Power BI dashboards showing review sentiment trends.
Final Report	Comprehensive report with visualizations, trends, and actionable recommendations.

### Schedule Overview / Major Milestones:

The expected schedule for the project. This can be defined by milestones (e.g. "all data is cleaned and processed"), periods of time ("Week 1 / Week 2"), or other ways based on the needs of the project.

Milestone	Expected Completion Date	Description/Details
Data Collection	Week 1 (Jan 5, 2023)	Extract all relevant reviews for the analysis.
Data Cleaning	Week 2 (Jan 12, 2023)	Ensure the dataset is ready for analysis.
Sentiment Analysis	Week 3 (Jan 19, 2023)	Conduct sentiment classification.
Dashboard Creation	Week 4 (Jan 26, 2023)	Build and validate visualizations.
Final Report Submission	Feb 2, 2023	Deliver the final report and presentation to stakeholders.

### \*Estimated date for completion:

This is my "if all goes well and I have everything I need, this is when I'll be done" date.

**February 2, 2023**