AXIA COHORT 8 PROJECT BRIEF

Data Analytics Track

AUGUST 1, 2025

AXIA AFRICA

OUTLINE

- Introduction
- Microsoft Excel Project
- Microsoft Power BI Project
- Structured Query Language (SQL) Project
- Python Project

INTRODUCTION

Dear Data Enthusiasts,

You've come a long way! Over the past four months, you've explored and practiced four powerful tools – Excel, Power BI, SQL, and Python. Now, it's time to bring everything together. This document contains your final project tasks, each one carefully designed to help you apply what you've learned in real-world situations.

These projects aren't just about testing your knowledge – they're about showing you how far you've come and giving you a chance to build something you'll be proud of. Dive in, give it your best, and most importantly, enjoy the process. You've got this!

AXIA Data Analytics Team

S/N	WEEK	TOOL	SCORE
1	**Week 1**	**Microsoft Excel Project**	**25**
2	**Week 2**	**Microsoft Power BI Project**	**25**
3	**Week 3**	**SQL Project**	**25**
4	**Week 4**	**Python Project**	**25**
		TOTAL SCORE	**100**

MICROSOFT EXCEL

In this project, you'll be working with a messy dataset that simulates real-world data problems. Your job is to clean and organize the data using Microsoft Excel so it's ready for analysis. You'll deal with inconsistent entries, spacing issues, missing or incorrect values, and more. This is your chance to apply all the Excel cleaning techniques you've learned, from text functions to filtering, sorting, and logical formulas.

Key Tasks

1. Autofit Columns and Rows.
2. Identify and Remove Duplicates.
3. Trim Extra Spaces.
4. Eliminate Blank Cells.
5. Convert Data into Table.
6. Use Find and Replace to correct errors.
7. Validate data to be sure it is thoroughly clean.

NOTE: You are to create a comprehensive report that contains everything you did to come up with the solution in the details. Let the document be in **PDF** format before submitting.

Find the dirty dataset here: Dataset

MICROSOFT POWER BI

You're working as a data analyst for a company that runs digital marketing campaigns across different channels: Email, Instagram Ads, Influencer Marketing, and more. Your team wants a fully interactive dashboard to monitor campaign performance and help optimize future strategies. Your job is to build a clean, professional dashboard that clearly communicates key insights and allows for dynamic filtering across campaigns, dates, products, and marketing channels.

Tips for Creating Dashboard

1. KPIs:

Show Total Ad Spend, Impressions, Clicks, Conversions, Revenue, and Overall ROI using card visuals.

2. Visuals:

Include charts for Ad Spend by Channel, Clicks vs Impressions, Conversion Rate by Category, Revenue by Product, ROI by Product/Category, and time-based trends for Spend, Conversions, ROI, and Clicks.

3. Interactivity:

Add slicers for Campaign Date, Product Name, Product Category, and Marketing Channel.

4. DAX Measures: Create calculated fields for CTR, Conversion Rate, and ROI (double-check vs dataset ROI).

5. Design Tips:

Keep it clean and readable, use a clear title and summary, and apply consistent colors (e.g., green for high ROI, red for low).

NOTE: You are to create a comprehensive report that contains everything you did to come up with the solution in the details. Let the document be in **PDF** format before submitting.

Find the marketing dataset here: Dataset

STRUCTURED QUERY LANGUAGE (SQL)

This project is designed to demonstrate SQL skills and techniques typically used by data analysts to explore, clean, and analyze retail sales data. The project involves setting up a retail sales database, performing exploratory data analysis (EDA), and answering specific business questions through SQL queries.

Objectives

1. Set up a retail sales database: Create and populate a retail sales database with the provided sales data.

2. Data Cleaning: Identify and remove any records with missing or null values.

3. Exploratory Data Analysis (EDA): Perform basic exploratory data analysis to understand the dataset.

4. Business Analysis: Use SQL to answer specific business questions and derive insights from the sales data.

Project Questions

1. Import the data into your SQL server.

2. Check for null values and clean the data.

3. Answer the following questions:
i. Write a SQL query to retrieve all columns for sales made on '2022-11-05

ii. Write a SQL query to retrieve all transactions where the category is 'Clothing' and the quantity sold is more than 4 in the month of Nov-2022

iii. Write a SQL query to calculate the total sales (total_sale) for each category.

iv. Write a SQL query to find the average age of customers who purchased items from the 'Beauty' category.

v. Write a SQL query to find all transactions where the total_sale is greater than 1000.

vi. Write a SQL query to find the total number of transactions (transaction_id) made by each gender in each category.

vii. Write a SQL query to calculate the average sale for each month. Find out best-selling month in each year

viii. Write a SQL query to find the top 5 customers based on the highest total sales.

ix. Write a SQL query to find the number of unique customers who purchased items from each category.

x. Write a SQL query to create each shift and number of orders (Example Morning <12, Afternoon Between 12 & 17, Evening >17)

NOTE: You are to create a comprehensive report that contains everything you did to come up with the solution in the details. Let the document be in **PDF** format before submitting.

Find the retail sales dataset here: Dataset

PYTHON

This project is designed to demonstrate your Python data analysis skills using the Pandas and Matplotlib libraries. You'll be working with raw sales data – your task is to clean it, explore it, and extract useful business insights. The goal is to simulate the day-to-day responsibilities of a data analyst: transforming messy data into clear, actionable findings using Python.

Objectives

1. Import the sales dataset into a Pandas Data Frame and get familiar with its structure and content.

2. Identify and handle missing values, remove duplicates, fix formatting issues (extra spaces), and convert columns to appropriate data types.

3. Use descriptive statistics and aggregation techniques to explore trends, distributions, and group-wise summaries.

4. Answer key business questions using Python and Pandas.

Project Questions

1. What was the Most Preferred Payment Method?

2. Which one was the Most Selling Product by Quantity and by Revenue?

3. Which City had maximum revenue, and Which Manager earned maximum revenue?

4. What was the Average Revenue?

5. What was the Average Revenue of November & December?

6. What was the Standard Deviation of Revenue and Quantity?

7. What was the Variance of Revenue and Quantity?

8. Was the revenue increasing or decreasing over the time?

9. What was the Average 'Quantity Sold' & 'Average Revenue' for each product?

10. What was the total number of orders or sales made?

NOTE: You are to create a comprehensive report that contains everything you did to come up with the solution in the details. Let the document be in **PDF** format before submitting.

Find the sales dataset here: Dataset