Data Analytics Tasks

Note:

Participants must complete at least
3 tasks for the 2-week internship and
4 tasks for the 1-month internship.

Level 1

Task 1: Sales Performance Dashboard Using Excel

Description:

- Dataset (Recommended): Superstore Sales Dataset (Kaggle)
- Import a simple sales dataset into Excel
- Clean and organize the data using formatting and formulas
- Use pivot tables to summarize key metrics: total revenue, units sold, monthly trends
- Create clean, readable charts to visualize overall performance

Tools & Libraries:

Excel

Covered Topics

Excel basics | Pivot tables & charts | Business metrics

Bonus:

Add slicers to filter by region or product category
Calculate YoY or MoM change

Task 2: Exploratory Data Analysis on Titanic Dataset

Description:

- Dataset (Recommended): Titanic: Machine Learning from Disaster (Kaggle)
- Explore the classic Titanic dataset using Python
- Perform data cleaning: handle missing values, convert types
- Generate summary statistics and group-based insights (e.g., survival by gender/class)
- Visualize key patterns and correlations

Tools & Libraries:

Python Pandas Seaborn or Matplotlib

Covered Topics

EDA | Data types & missing data | Data visualization

Bonus:

Visualize survival rates using bar plots and heatmaps

Data Analytics Tasks

Level 2

Task 3: Customer Segmentation Using RFM Analysis

Description:

- Dataset (Recommended): Online Retail Dataset (UCI)
- Use a sales dataset to analyze customer behavior based on:
  - Recency: How recently a customer made a purchase
  - Frequency: How often they made purchases
  - Monetary: How much money they spent
- Assign scores to each customer using these three metrics
- Group similar customers based on their scores
- Suggest simple marketing ideas for each group (e.g. discounts for loyal buyers, re-engagement for inactive ones)

Tools & Libraries:

Python | Pandas | Seaborn (or Excel)

Covered Topics

Feature engineering | Segmentation logic | Customer analytics

Bonus:

Visualize RFM segments with heatmaps or bar charts

Task 4: Data Cleaning and Insight Generation from Survey Data

Description:

- Dataset (Recommended): Kaggle Data Science Survey 2017-2021 (Kaggle)
- Work with real-world survey data containing missing values, duplicates, and inconsistent formatting
- Clean the dataset and apply label encoding or mapping for categorical variables
- Extract meaningful insights about respondent (a person who answered the survey) behavior or preferences

Tools & Libraries:

Python or Excel

Covered Topics

Data cleaning | Categorical handling | Reporting

Bonus:

Create a summary dashboard or chart showing top 5 insights

Data Analytics Tasks

Level 2

Task 5: SQL-Based Analysis of Product Sales

Description:

- Dataset (Recommended): Chinook Database (Kaggle)
- Use SQL queries to analyze product sales from a relational database
- Answer key business questions: top-selling products, revenue per region, monthly performance
- Write JOINs to combine product and sales tables

Tools & Libraries:

SQL (SQLite, PostgreSQL, MySQL or BigQuery)

Covered Topics

SQL queries | JOINs & aggregations | Business logic from raw data

Bonus:

Use a window function (e.g., ROW_NUMBER or RANK)

Data Analytics Tasks

Level 3

Task 6: Web Scraping and Analysis of Job Postings

Description:

- Scrape job listings from a real website (like Indeed or a static HTML page)
- Extract fields such as job title, location, company, skills, and posted date
- Clean and analyze the data to find most in-demand roles or skills

Tools & Libraries:

Python | Pandas | BeautifulSoup or Selenium

Covered Topics

Web scraping | Text cleaning | Skill frequency analysis

Bonus:

Visualize most frequent skills by city

Task 7: Time Series Breakdown of Retail Sales

Description:

- Dataset (Recommended): Walmart Sales Forecasting (Kaggle)
- Analyze a retail dataset with monthly sales
- Plot trends, moving averages, and seasonal patterns
- Break down revenue by product and region over time

Tools & Libraries:

Python (Pandas, Matplotlib) or Excel

Covered Topics

Time series analysis | Trend & seasonality | Visualization over time

Bonus:

Add simple forecasting (rolling mean or exponential smoothing)

Data Analytics Tasks

Level 3

Task 8: Power BI Dashboard for Retail Insights

Description:

- Dataset (Recommended): Superstore Sales Dataset (Kaggle)
- Use Power BI to create an interactive dashboard from a cleaned retail dataset
- Include KPIs such as total revenue, top products, sales by region
- Add filters, slicers, and custom visuals

Tools & Libraries:

Power BI Desktop

Covered Topics

Data modeling | Interactive visualization | BI dashboarding

Bonus:

Publish the dashboard to the web

Data Analytics Tasks

Industry Level

Task 9: Business Insights & Executive Report for E-Commerce Dataset

Description:

- Dataset (Recommended): Brazilian E-Commerce Public Dataset by Olist (Kaggle)
- Conduct a full-cycle analysis on an e-commerce dataset
- Clean and preprocess sales, product, and customer data
- Perform EDA and segmentation, visualize patterns, and generate actionable business insights
- Present your findings in a professional report or interactive dashboard

Tools & Libraries:

Python, Excel, Power BI or Tableau (you can use any of them)

Covered Topics

End-to-end analytics workflow | Business storytelling | Decision-oriented reporting

Bonus:

Include a slide or one-pager executive summary with your final submission