**2025 Rihal Codestacker Challenge**

Navigating the Digital Bazaar

**Storyline**

After what happened to the CityX, officials in the MENA region want to have more insight into e-commerce because the world of e-commerce is undergoing a transformative evolution. Traditional marketplaces blend with modern online platforms, creating a rich and complex ecosystem where data flows from every corner—from sales figures and customer reviews to web traffic and social media trends. Business leaders need clear, actionable insights to navigate this dynamic landscape and drive strategic decisions. Your mission is to embark on a comprehensive data journey, collecting and integrating diverse datasets on major e-commerce players across the region. Uncover hidden trends, generate impactful insights, and craft a creative visualization that delivers measurable business value. This would help with preventing crime in a rapidly growing industry.

**Problem Statement**

The MENA e-commerce sector is characterized by its rapid growth and data diversity. The challenge is to aggregate data from various publicly available sources—including web scraping, public datasets, and APIs—covering sales performance, customer sentiment, website analytics, and social media influence. You are tasked with cleaning, integrating, and analyzing this data using Python, SQL, Power BI, or other open-source BI tools. The final deliverable should be an interactive dashboard and custom visualization that not only reveal key insights but also empower management with the data they need to make informed decisions.

**Your Challenge**

Design and implement a data analysis project that covers the entire data journey:

* **Data Collection:** Identify and gather data from at least three distinct sources relevant to big e-commerce businesses in the MENA region.
* **Data Cleaning & Integration:** Preprocess and merge the datasets into a unified, analyzable structure.
* **Insight Generation:** Conduct exploratory data analysis to uncover trends, anomalies, and key performance indicators.
* **Visualization:** Develop a custom and creative visualization that effectively communicates your insights.
* **Dashboard Creation:** Build an interactive dashboard using Python (e.g., Plotly, Dash), SQL, Power BI, or other open-source BI tools to present your findings.
* **Journey Documentation:** Upload a clear and detailed report documenting every step of your journey—from data acquisition to final insights.

**Objectives & Challenge Levels**

**Level 1 (30 points): Data Collection & Preprocessing**

* **Task:** Identify and gather data from at least three relevant sources (e.g., public APIs, web scraping, open datasets) covering aspects like sales, customer engagement, and market trends in the MENA region.
* **Deliverable:** A unified dataset created after thorough cleaning and merging. Document your data sources and preprocessing steps.

**Level 2 (25 points): Data Analysis & Insight Discovery**

* **Task:** Perform exploratory data analysis (EDA) to understand data distributions, trends, and anomalies. Use Python or SQL to calculate key metrics (e.g., sales growth, customer sentiment, market penetration).
* **Deliverable:** A detailed analysis report highlighting at least three actionable insights that can drive business decisions.

**Level 3 (30 points): Custom Visualization & Dashboard Creation**

* **Task:** Develop a custom, creative visualization that uniquely represents your findings. Build an interactive dashboard that allows stakeholders to explore data and insights dynamically.
* **Deliverable:** A fully functional dashboard that is both visually appealing and robust, using tools such as Python (Plotly, Dash), SQL, or Power BI.

**Level 4 (15 points): Presentation & Reporting**

* **Task:** Prepare a comprehensive report that narrates your entire data journey. This should include your methodologies, challenges, solutions, and key learnings.
* **Deliverable:** Upload a clear, detailed report documenting every step—from data collection and cleaning to analysis and visualization.

**Bonus Task: Advanced Analytics and Predictive Modeling**

* **Task:** (Optional) Develop a predictive model to forecast future trends in the MENA e-commerce market. Validate your model and integrate its forecasts into your dashboard.
* **Deliverable:** An enhanced dashboard featuring advanced analytics that provide forward-looking business insights. Extra credit will be awarded for innovative approaches that further add business value.

**Submission Guidelines**

* Include all necessary scripts, dashboards (e.g., app.py for a web-based dashboard), and support files.
* Provide a detailed README with instructions on how to run your project and a comprehensive report documenting your entire data journey. Happy Analyzing and Good Luck!