**High Impact Skills Development Program** **AI & Data Science**

**Name: Irshad Ali**

**Roll No: DSAI-GIL-203**

**Gmail:** [**irshadalirahi001@gmail.com**](mailto:irshadalirahi001@gmail.com)

**Title: Data Visualization Project.**

**Profile link: https://public.tableau.com/app/profile/irshad.ali5123/viz/ProjectDatavis2/Dashboard1?publish=yes.**

**1. Introduction:**

The dataset I selected from VGChartz 2024 contains comprehensive information on video game sales, covering variables such as game titles, platforms, regions, genres, and release years. The objective of my Exploratory Data Analysis (EDA) was to uncover insights related to sales performance across regions, genres, and time periods. Through visualizations such as regional sales breakdowns, genre-based performance, and historical sales trends, I aim to support future AI/ML tasks, like predicting game sales or classifying popular genres. These insights will help in understanding patterns that can improve recommendation systems, forecast trends, and guide strategic decision-making.

**2. Visualization Process:**

For each visualization, I carefully selected chart types that best represented the data. The process included:

* **Total Sales by Region**: I used a bar chart, as it provides a clear comparison between regions and displays overall contribution to global sales. This was useful to highlight regional dominance.
* **Average Sales by Genre**: I chose a bar chart to show the distribution of sales across different genres. Which game genres are the most popular based on sales.
* **Top 10 Games by Total Sales**: A horizontal bar chart was chosen, given its readability for ranked data. This made it easy to compare sales between the highest-grossing games.
* **Sales Trends Over Time**: I used a line chart to track trends across years, revealing peaks and troughs in sales that aligned with a pie chart to show the proportional breakdown of sales by genre. This was effective in conveying major industry events.
* **Distribution of Total Sales by Genre**: The box plot shows the distribution of total sales for different game genres. It highlights the variability and median sales within each genre, with some genres showing a wider range of success.

**3. Decision-Making Justification:**

In designing these visualizations, I focused on clarity, aesthetics, and usability:

* **Clarity**: I ensured each chart was simple and focused, with appropriate axis labels, legends, and titles.
* **Aesthetics**: I maintained a consistent design throughout the dashboard, using clean, minimalistic layouts that avoided clutter.
* **Interactivity**: In dashboards, allowing users to interact with filters (e.g., by genre or region) enhances the analytical experience, offering flexibility in exploring the data at multiple levels.

**4. Challenges and Solutions:**

One of the main challenges I faced was handling missing or incomplete data, especially for older games where sales data was sparse across regions. To address this, I implemented imputation techniques for missing data and excluded games with incomplete datasets in critical analyses to maintain accuracy.

**5. Conclusion:**

The key insights from my EDA include identifying the most profitable regions and genres, as well as significant sales trends over time. For example, North America and Japan emerged as dominant regions, while genres like action and sports consistently topped the charts in terms of sales. These findings will be instrumental in informing future AI/ML tasks. For instance, genre and region-specific insights could help build models that predict game sales more accurately, while temporal sales patterns could enhance models forecasting the success of upcoming releases.