# CEN445 - INTRODUCTION TO DATA VISUALIZATION PROJECT REPORT

**Project Title:** *Global Video Game Sales Exploratory Dashboard* **Team Members:** *2021555027 Cenk Göl, 202255560 Muhammed Hilmi  Salkutlu, 2021555004 İlhan Altunbaş*

**1. Dataset and Context**   We selected the "vgsales.csv " dataset, which contains records of video games with sales greater than 10,000 copies. The dataset includes 16,599 rows and 11 columns, featuring dimensions such as Platform, Year, Genre, Publisher, and numerical sales data for North America (NA), Europe (EU), Japan (JP), and Global markets.

- **Goal:**  To analyze the gaming industry's evolution, understand regional market differences, and identify dominant genres and publishers through interactive visualizations.
- **Preprocessing:**  We cleaned the data by removing rows with missing Year and Publisher values to ensure accuracy. We also converted the Year column to integers for better time-series analysis.

**2. Visualization Techniques & Contributions**   We developed an interactive dashboard using Python, Streamlit, and Plotly. The project features 9 distinct visualizations, with interactivity such as zooming, filtering, and hovering.

*Cenk: Focused on statistical distributions and hierarchy.*

- **Interactive Scatter Plot:** Analyzed the correlation between NA and EU sales using marginal histograms and trendlines.
- **Box Plot :** Visualized sales distributions across genres. Implemented a "Deep Zoom" feature to focus on the 0-2M range, revealing median values hidden by outliers.
- **Dynamic Treemap:** Created a hierarchical view of market share, allowing users to dynamically switch layers (e.g., Publisher → Genre → Platform).

*İlhan: Focused on flow and temporal trends.*

- **Sankey Diagram:** Visualized the multi-stage sales distribution from the Top 5 Publishers to specific Genres and Platforms.
- **Stacked Area Chart:** Illustrated the sales evolution and market lifecycles of the Top 12 platforms over the years.
- **Line Chart:** Tracked the annual global sales performance of specific publishers with interactive zooming features.

*Hilmi: Focused on comparative and ranking analysis.*

- *Parallel Coordinates:* Compared regional sales profiles (NA, EU, JP, Other) across different game genres.
- **Interactive Heatmap (Choropleth):** A geographical visualization showing the regional footprint of the top-selling game for each year. It colors countries based on which region (Americas, Europe, Japan, Other) the year's hit game sold best in, highlighting global shifts in market dominance.
- *Bar Chart:* Ranked the Top 20 best-selling games of all time.

**3. Key Insights & Findings**

- **Market Dynamics & Sales Distribution:** The correlation analysis reveals a strong positive relationship between North American and European markets, indicating similar regional gaming preferences. Furthermore, the Box Plot highlights a significant "Long Tail" effect: while the industry average is skewed by massive outliers like Wii Sports, the vast majority of games sell under 1 million copies. This demonstrates that "hit" games are rare exceptions rather than the norm in the gaming ecosystem.

- **Publisher Strategy & Lifecycle:** The Sankey Diagram demonstrates that top publishers do not distribute their efforts randomly; instead, they clearly specialize in specific genre-platform pipelines to maximize their reach. Complementing this, the temporal analysis confirms that publisher success is rarely linear. The trends show "Hit-Driven Volatility," where revenue peaks align sharply with major console launches and blockbuster releases rather than steady, continuous growth.

- **Global Shifts & Best-Sellers:** The Geographical Heatmap visualizes a historical shift in market dominance, highlighting how top-selling titles have evolved from being purely North America-centric to relying increasingly on European sales in recent years. Additionally, the Top 20 Games ranking confirms the massive scale difference between standard titles and industry giants, illustrating that high sales figures stem from both bundled casual games (e.g., Nintendo Wii era) and standalone multi-platform releases.

**4. Project Link :**   GitHub Repository: CEN445 INTRODUCTION TO DATA VISUALIZATION COURSE -  ASSIGNMENT