



أكاديمية سدايا  
SDAIA Academy

## Exploratory Data Analysis

### Project Proposal



Raghad Alnasser

Nov 2021

## **Introduction**

This dataset contains a list of video games with sales greater than 100,000 copies. It contains 16,600 rows and 11 columns.

## **Dataset**

- Rank - Ranking of overall sales (int)
- Name - The games name (categorical)
- Platform - Platform of the games release (categorical)
- Year - Year of the game's release (categorical)
- Genre - Genre of the game (categorical)
- Publisher - Publisher of the game (categorical)
- NA\_Sales - Sales in North America (in millions) (float)
- EU\_Sales - Sales in Europe (in millions) (float)
- JP\_Sales - Sales in Japan (in millions) (float)
- Other\_Sales - Sales in the rest of the world (in millions) (float)
- Global\_Sales - Total worldwide sales (float)

## **Questions:**

- Which platforms are the most popular?
- How many games sales - Super Mario All-Stars- in Japan?
- How many releases in 2012?
- What is the least selling game in North America?
- What is the least selling game in 2008 in Europe?
- What is the most frequent genre of games?

## **Tools**

- Python Language.
- Python library (pandas, numpy, matplotlib).
- Jupyter notebook
- Excel

## **TO DO:**

- Explore the data and analysis.
- Answer the questions.