



EDA

**Video Games Sales**

**By:  
William Baaklini**

LET'S PLAY

LET'S PLAY

LET'S PLAY

LET'S PLAY

LET'S PLAY

# WHY?

## Evaluation Strategy

- Regression: Predict Global Sales
- Classification: Predict whether a game is a "Hit"
- Clustering (Optional): Segment games based on regional sales

To explore and model the factors that influence the success of video games by analyzing sales data across different platforms, genres, and regions from 1980 to 2024.

- What genres, platforms, and publishers tend to generate higher sales?
- Can we predict if a game will be a high-seller based on its features?
- Are there distinct clusters of games based on regional performance? (optional)



# Data Used

## Dataset

**Title:** Video Game Sales 2024

**Source:** Kaggle Dataset

**Size:** ~ 64016 entries

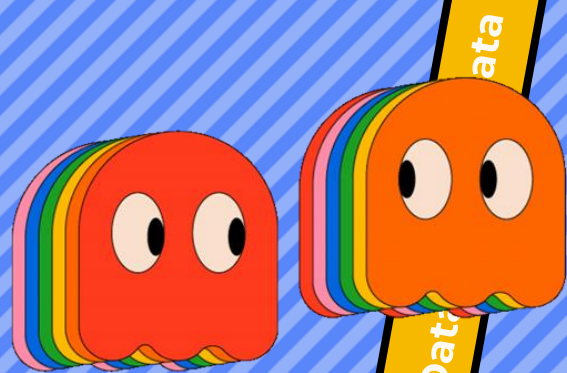
**Format:** Structured CSV

## Features

- **Name:** Title of the video game
- **Platform:** Platform released (e.g., PS4, PC, Switch)
- **Year:** Year of release
- **Genre:** Game genre (e.g., Shooter, Sports, RPG)
- **Publisher:** Game publisher
- **NA\_Sales, EU\_Sales, JP\_Sales, Other\_Sales, Global\_Sales:** Sales in millions per region

## Why?

- Longitudinal (1980–2024) — allows trend analysis
- Includes both categorical and numerical features
- Enables exploration, prediction, and segmentation





# Thanks

William Baaklini

s34859@pjawst.edu.pl

[www.williambaaklini.com](http://www.williambaaklini.com)

