



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 produced?
- what genre has the most sales?
- What are top 5 Platform have the most sales?
- What relationship between sales and years?

Tools

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

TO DO:

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