



# **Gaming Industry: Sales Drivers for Games**

**Cime Ajouz  
James Janovsky  
Rashmi Francis  
Victor Fonseca**

**December/2019**

# What makes a game sell well?

## Genre, target market, licensed platforms

What type of game to  
release, where and for  
which systems?

## Ratings

Correlate historical  
ratings to historical sales

## Sales prediction

Prepare a database that  
can help in future  
predictive endeavors



# Project walkthrough

## DATA COLLECTION

- Dataset "Video Games Sales" (Kaggle, 2016), based on a scrape of VGChartz.com
- RAWG API, scraped for this project
- Key: game name

## DATASET COMPILATION

- Levenshtein Distance Ratio, for strings comparison

Ex.:

- Street Fighter 2
- Street Fighter II
- Street Fighter 2 – The World Warrior

## EXPLORATORY ANALYSIS

- Analyzing sales in function of some drivers
- Correlation analysis ratings x sales

## > TOOLS WE USED:

- BeautifulSoup
- Requests
- Json
- Pandas
- Numpy
- Matplotlib
- Seaborn



# Data dictionary

Source	Attribute name	Description
Video Game Dataset (Kaggle)	Name	Game name
	Platform	Platform (console) for which the game was released
	Year	Year of release
	Genre	Genre of the game
	Publisher	Publisher of the game
	NA_Sales	Sales in North America (in millions of units)
	EU_Sales	Sales in Europe (in millions of units)
	JP_Sales	Sales in Japan (in millions of units)
	Other_Sales	Sales in the rest of the world (in millions of units)
	Global_Sales	Total sales (worldwide)
RAWG API	API_release_date	Release date (used to confirm the game matching)
	API_playtime	Game estimated playtime
	API_rating	Game rating for RAWG users (1-5)
	API_n_platforms	Number of platforms for which the game was released
	API_search_name	Game name found on the API ((used to confirm the game matching)



# CHART 1 - Yearly Sales by Market

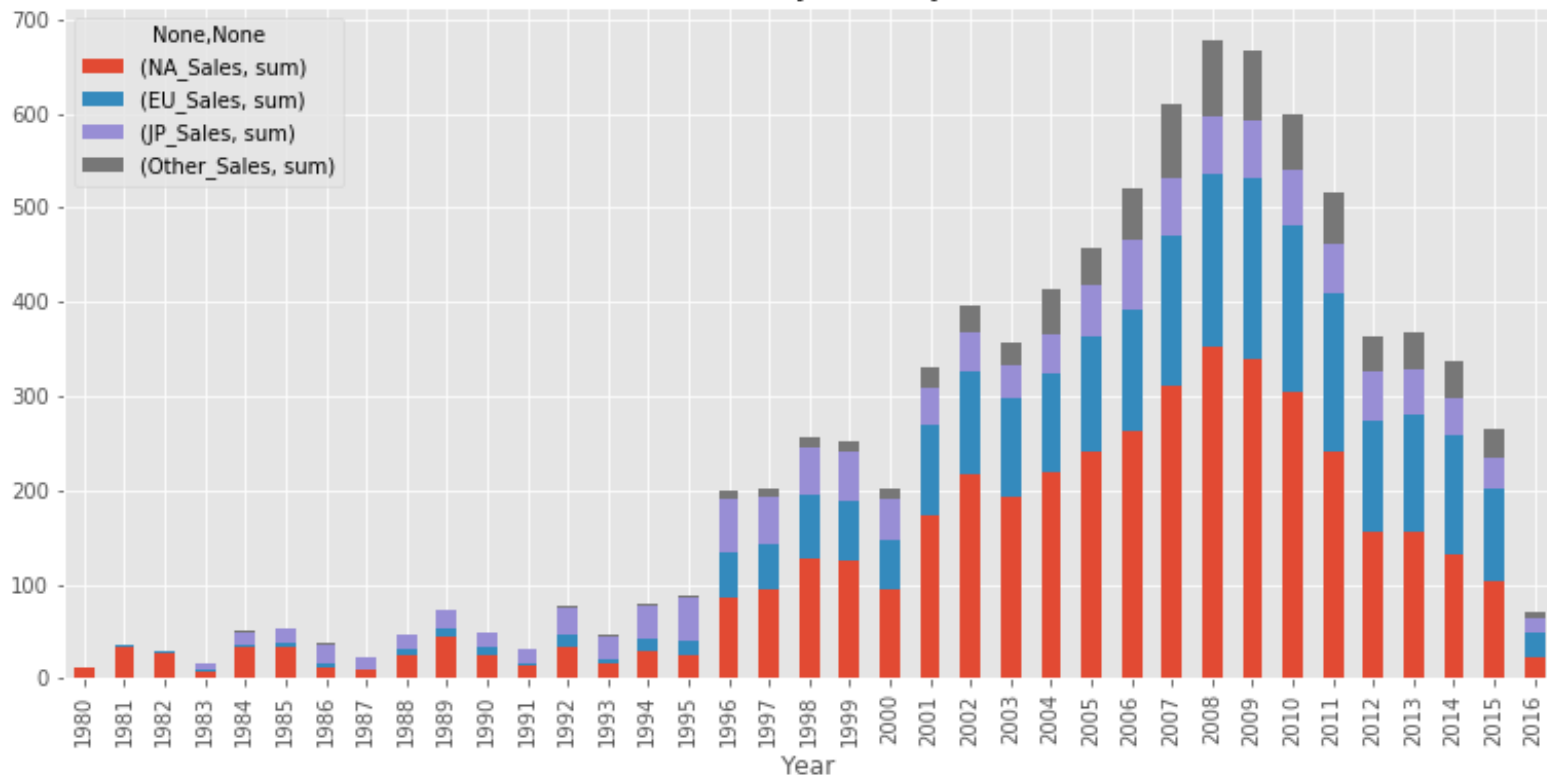
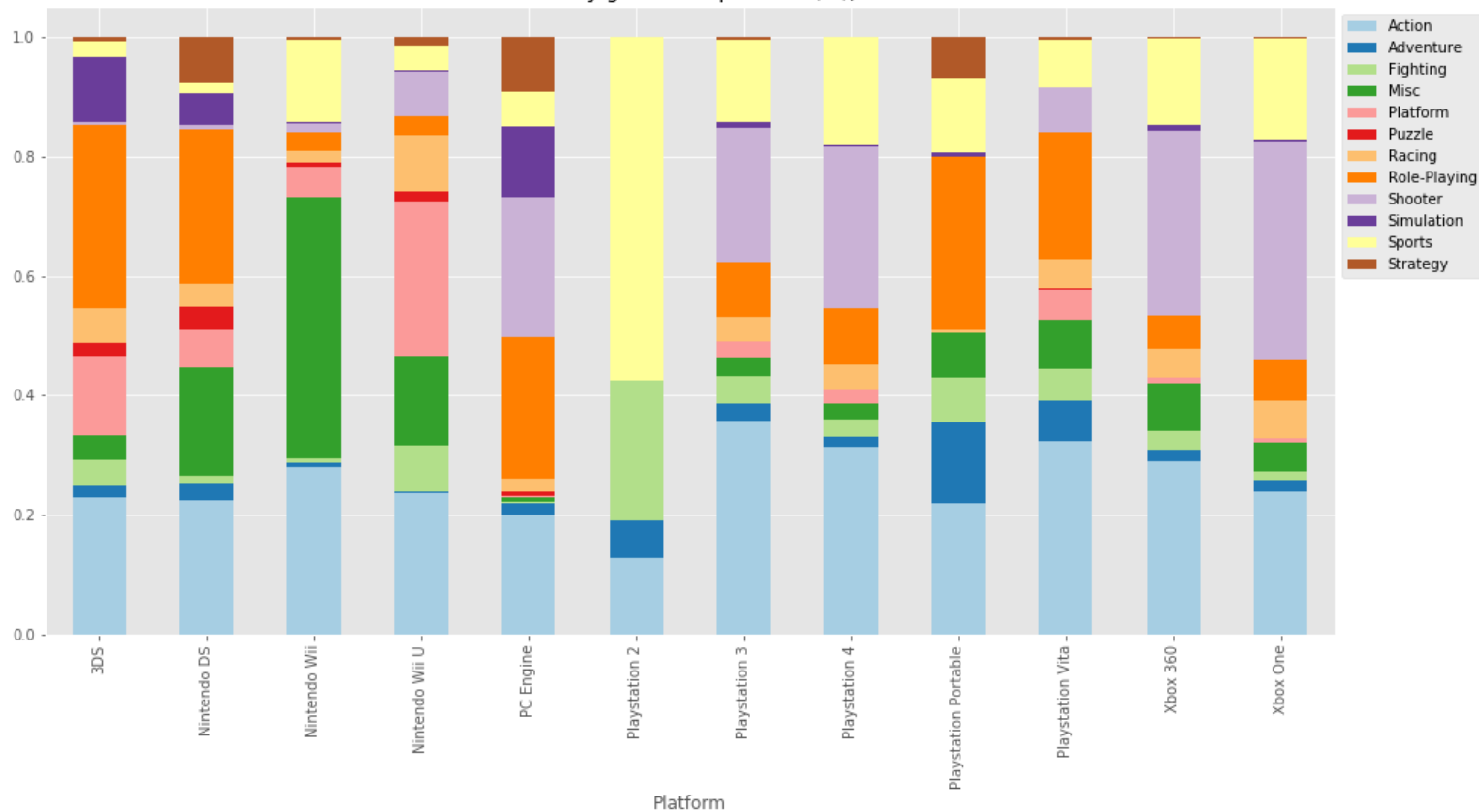


CHART 2 - Games sales by genre and platform (%), from 2010 to 2016





**Thank  
you**