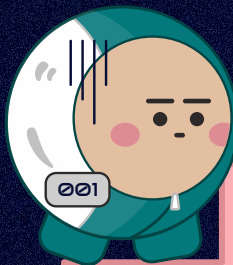
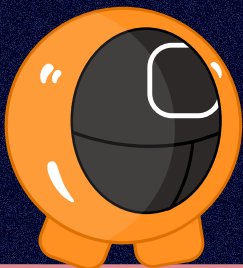


THE GAME INDUSTRY

By: Andrew Mukurazita, Davy Xu,
Gengchen Zhang, Jenny W.



THE GAMES

001

Introduction

- Dataset
- Purpose of the dataset
- Access

002

Data Overview

- Contents of dataset
- Unit of Measures

003

Scope of Data

- General direction
- Kaggle

004

Figures

- Graphs presentation
- Analytical details

005

Summary

- Potential Pros & Cons
- Future application

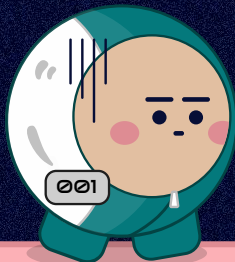
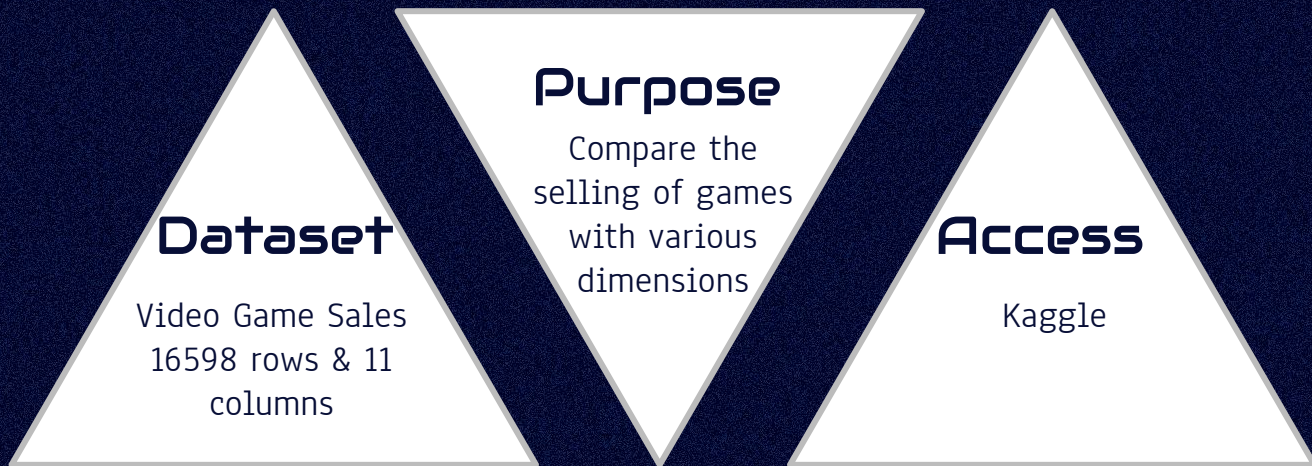


01

Introduction

Dataset, purpose and access

THE PLAYERS





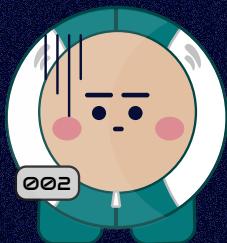
Data Overview

Contents of Dataset
& Unit of Measures

WHO IS WHO

```
Console Terminal x Background Jobs x
R 4.2.1 · C:/Users/26292/Desktop/
> glimpse(game)
Rows: 16,598
Columns: 11
$ Rank      <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, ...
$ Name      <chr> "Wii Sports", "Super Mario Bros.", "Mario Kart ...
$ Platform  <chr> "Wii", "NES", "Wii", "Wii", "GB", "GB", "DS", "...
$ Year      <chr> "2006", "1985", "2008", "2009", "1996", "1989",...
$ Genre     <chr> "Sports", "Platform", "Racing", "Sports", "Role...
$ Publisher <chr> "Nintendo", "Nintendo", "Nintendo", "Nintendo",...
$ NA_sales  <dbl> 41.49, 29.08, 15.85, 15.75, 11.27, 23.20, 11.38...
$ EU_sales  <dbl> 29.02, 3.58, 12.88, 11.01, 8.89, 2.26, 9.23, 9...
$ JP_sales  <dbl> 3.77, 6.81, 3.79, 3.28, 10.22, 4.22, 6.50, 2.93...
$ Other_sales <dbl> 8.46, 0.77, 3.31, 2.96, 1.00, 0.58, 2.90, 2.85,...
$ Global_sales <dbl> 82.74, 40.24, 35.82, 33.00, 31.37, 30.26, 30.01...
> #16598 16
```


WHO IS WHO



Contents of Dataset

Rank: Ranking of overall sales
Names: Name of the game
Platform: platform of the game release
Year: Year of the game's release 1980-2018
Genre: Type of the game
Publisher: Publisher of the game
NA, EU, JP, Other, Global sales: Sales in different regions



Unit of Measures

All Sales data have millions as the unit of measures



Scope of Dataset

Scope of the Dataset

- The dataset we chose includes video games with sales greater than 100,000 copies
- The dataset is the original data scrapped from <http://www.vgchartz.com/gamedb/> by GregorUT and tuvtran in github
- Reason we chose this dataset: reasonable for consulting-based analysis



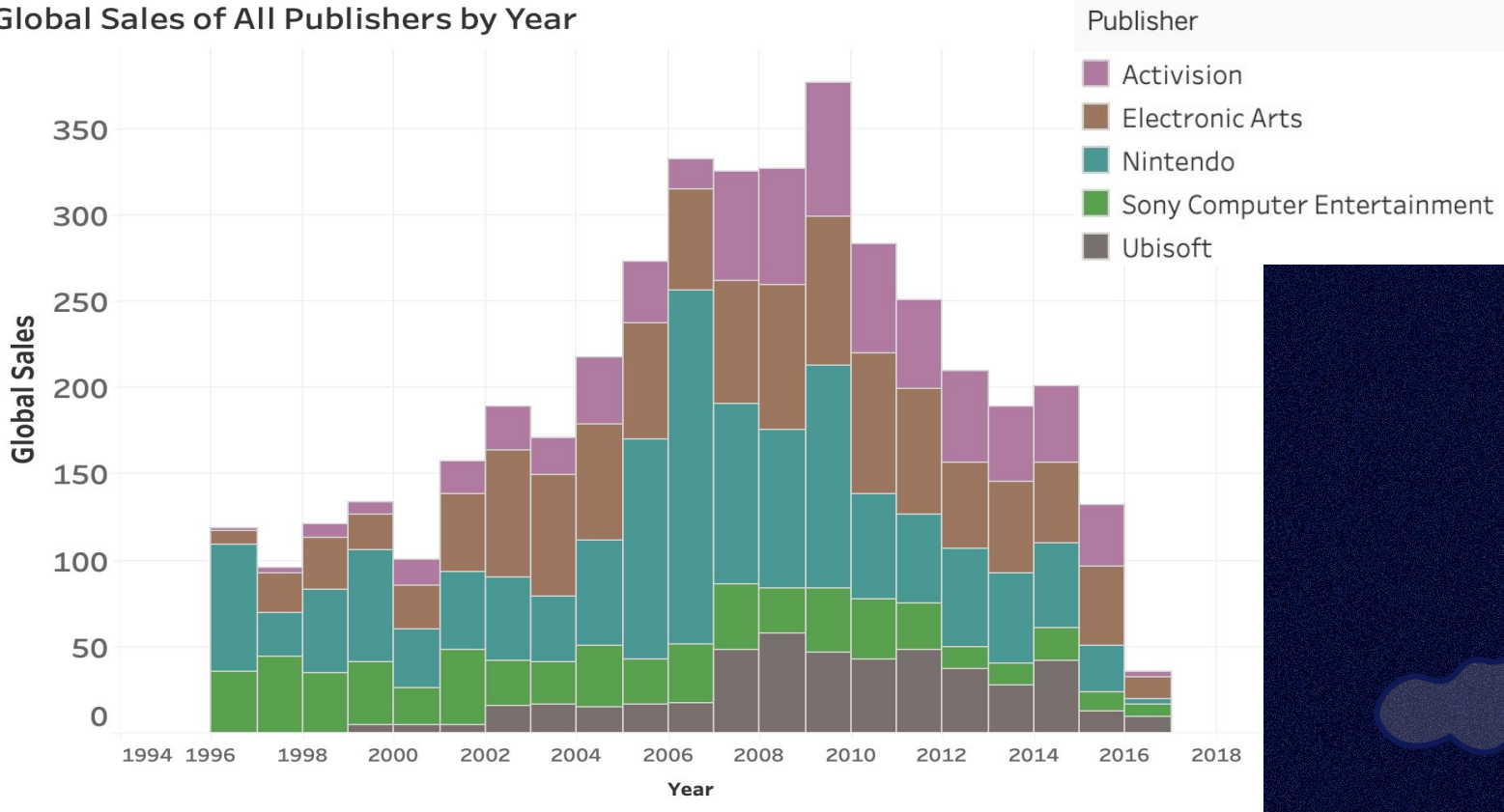
04

Graphs

Graphs & Analysis

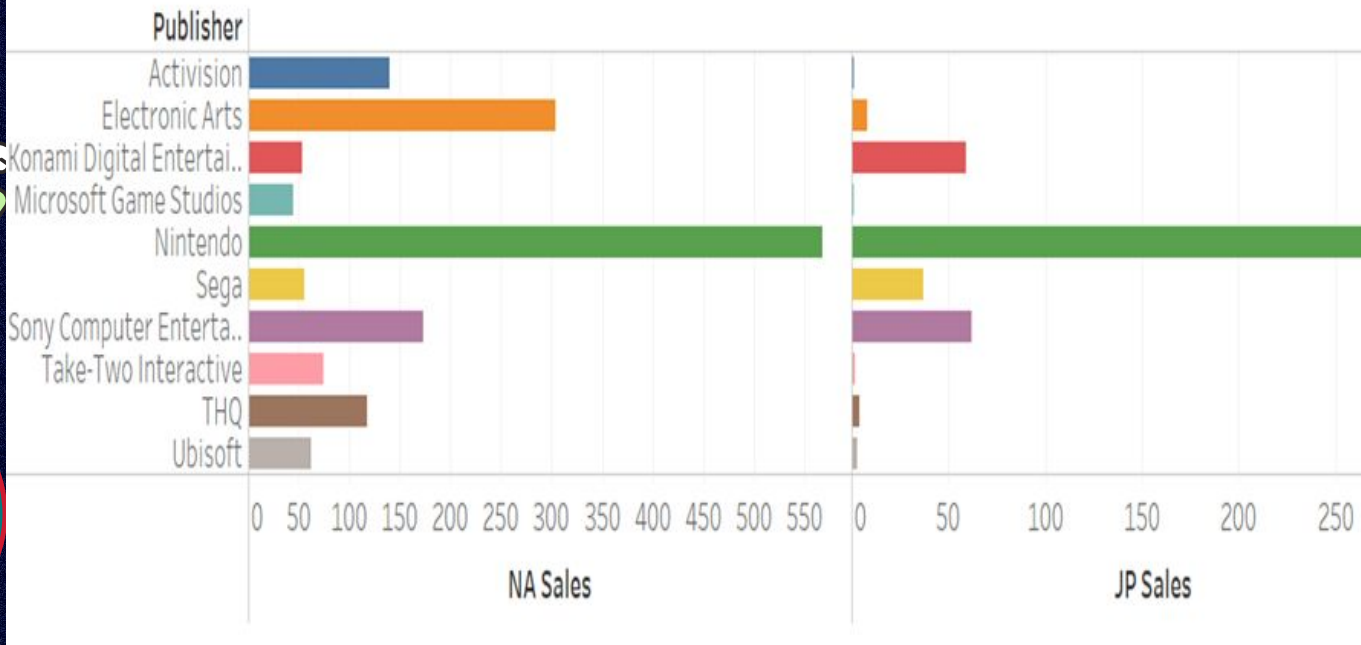
Best selling publishers over time

Global Sales of All Publishers by Year

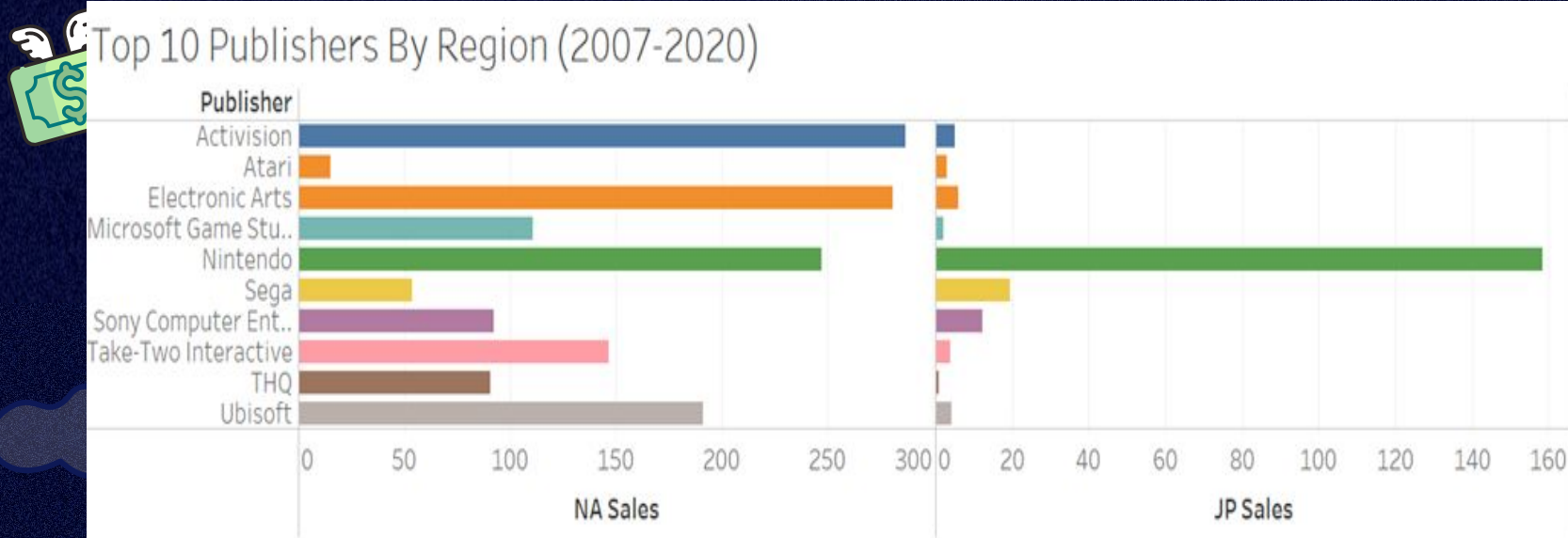


Best Selling Publishers Pre-2007 Era

Top 10 Publishers By Region (1980-2006)



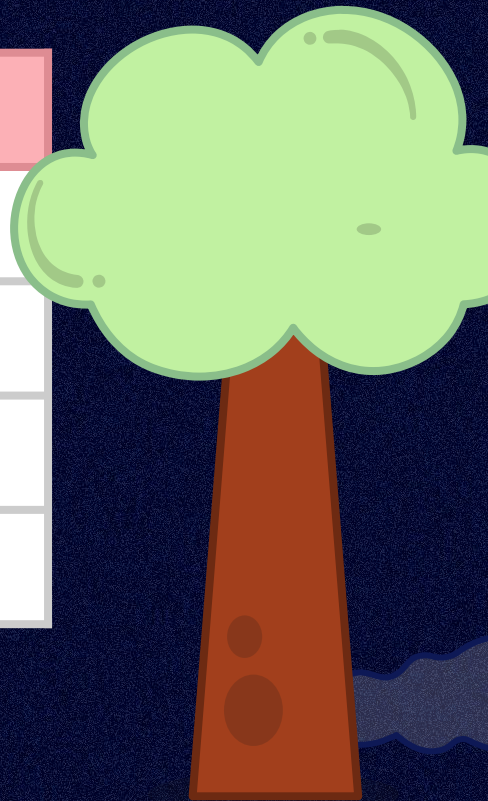
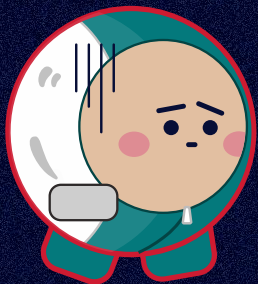
Best Selling Publishers Post 2007 Era



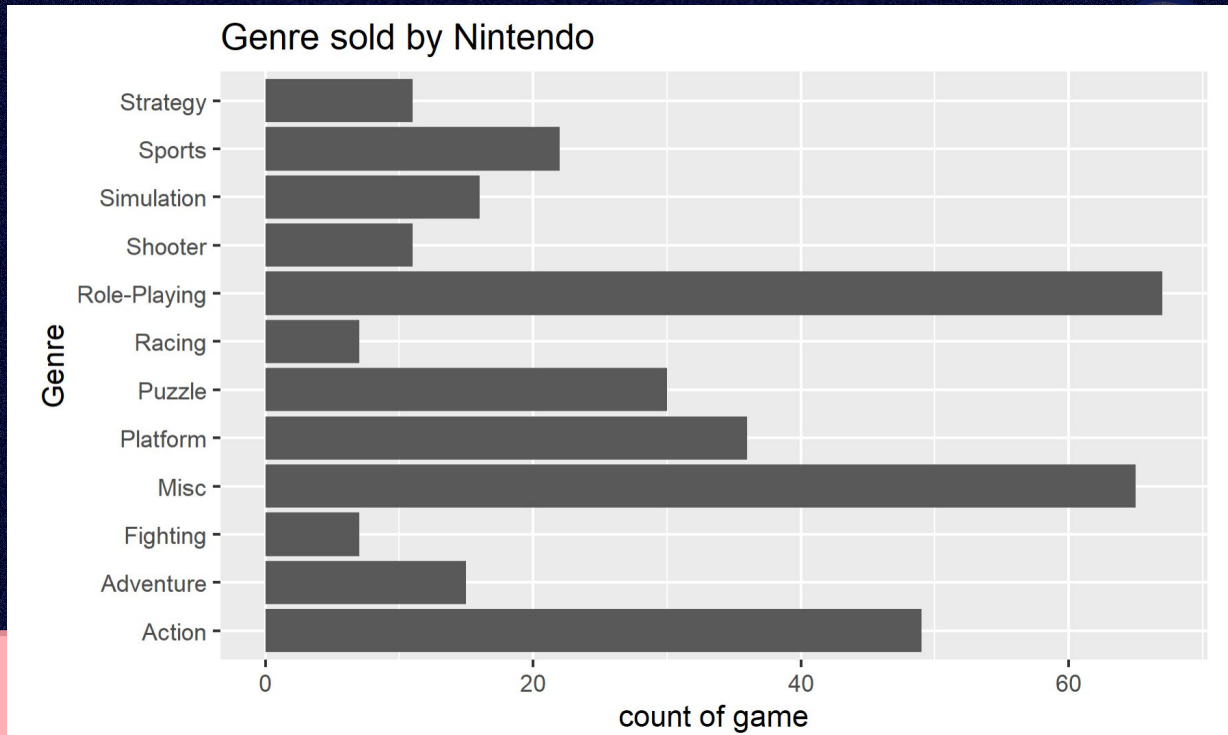
Sum of NA Sales and sum of JP Sales for each Publisher. Color shows details about Publisher. The data is filtered on Year, which ranges from 2007 to 2020. The view is filtered on sum of NA Sales and Publisher. The sum of NA Sales filter includes everything. The Publisher filter keeps 10 of 579 members.

Who's the best? (Company)

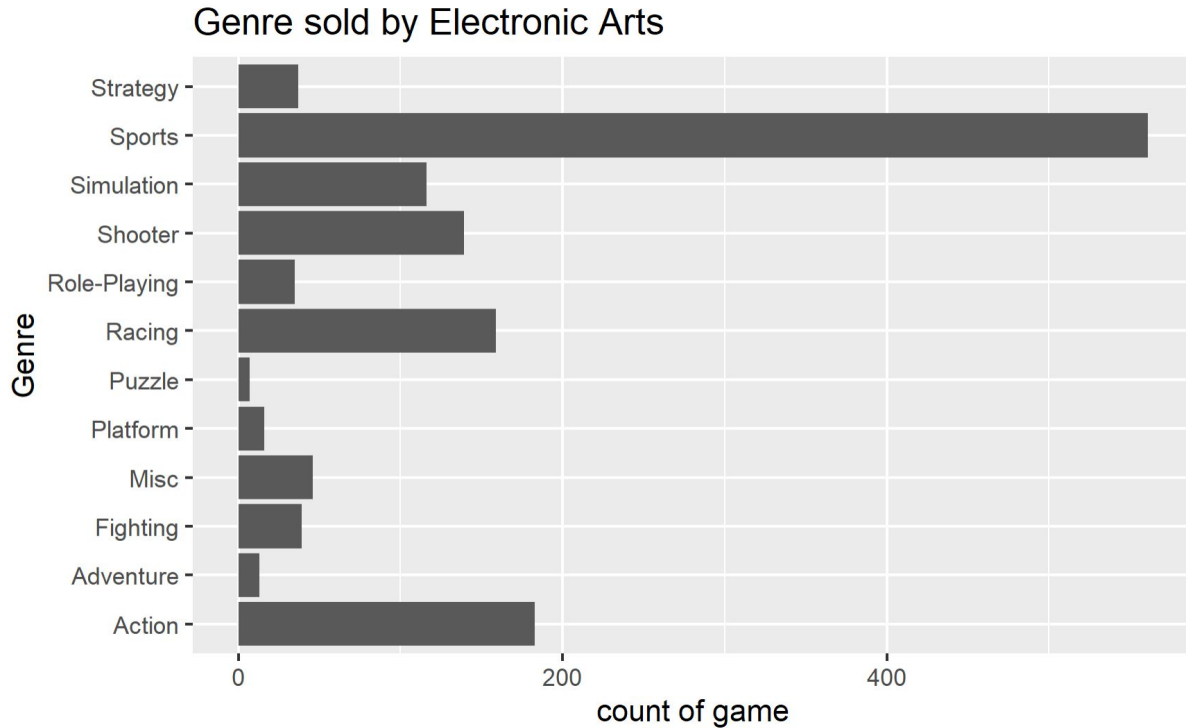
Region	Publisher
EU	Nintendo, EA, Activision
NA	Nintendo, EA, Activision
JP	Nintendo
Other	EA, Nintendo, Activision



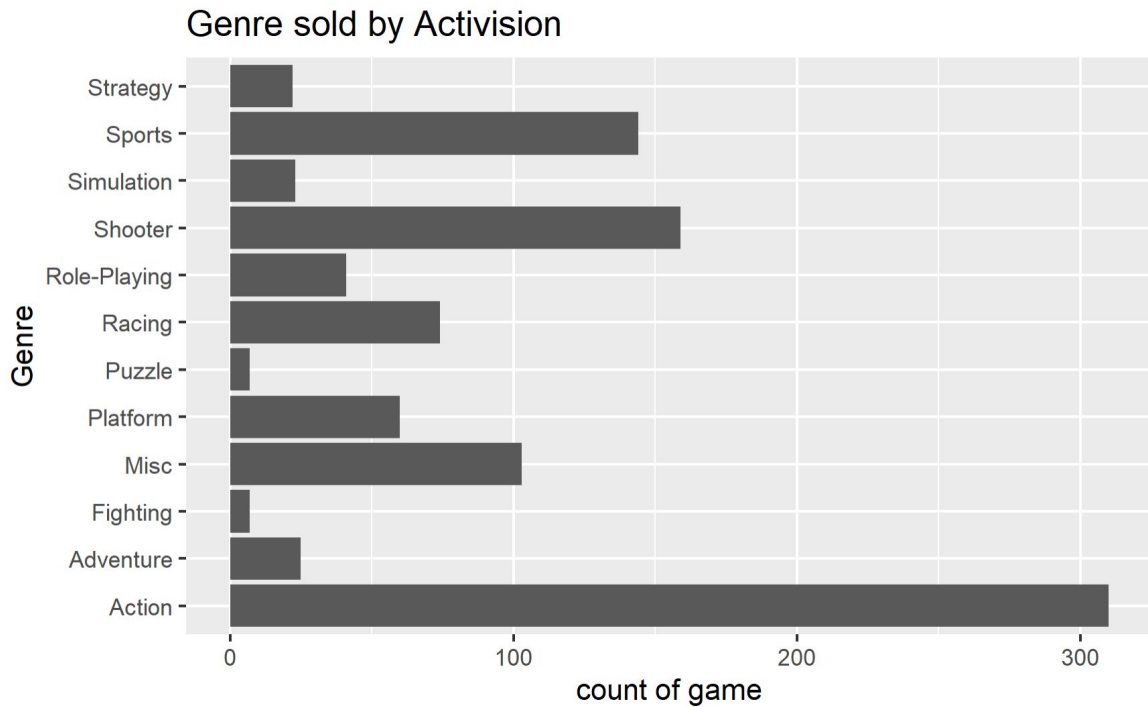
Genre sold by Nintendo



Genre sold by Electronic Arts

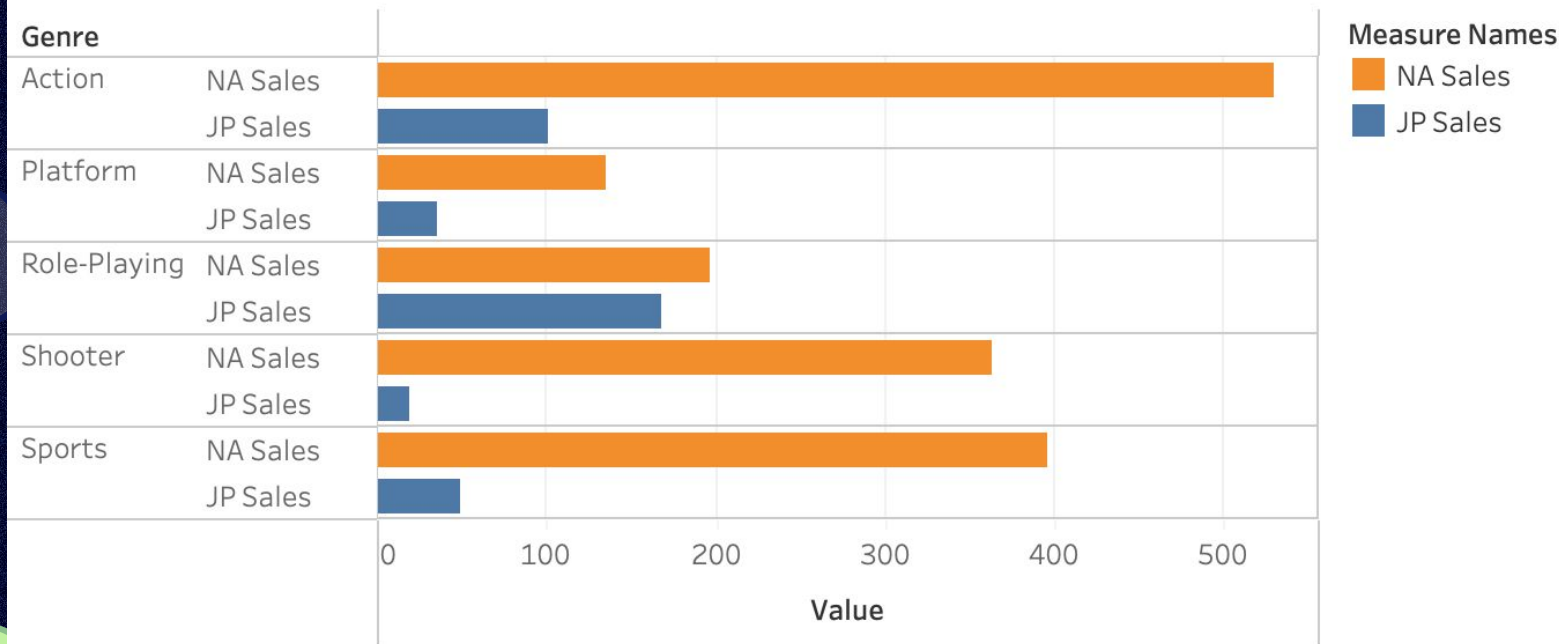


Genre sold by Activision



Regional sales group by Genre

<Genre Preference>



NA Sales and JP Sales for each Genre. Color shows details about NA Sales and JP Sales. The data is filtered on Year, which ranges from 2006 to 2018. The view is filtered on Genre, which keeps Action, Platform, Role-Playing, Shooter and Sports.



Summary

You can enter a subtitle here if you
need it

Future Application



Indie game
developer



Pop culture
analysis

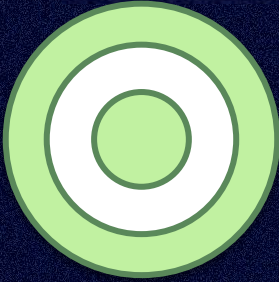


Demographic
analysis



Investors

Pros & Cons



- Clustered
- Data is clean
- Provides an overview for the industry
- Collected from official sources



- N/A in years
- Effect by time
- The life span of platforms are different
- This dataset is a scrapped result from the original sources

Thanks!

