Overview of the Console Game Market
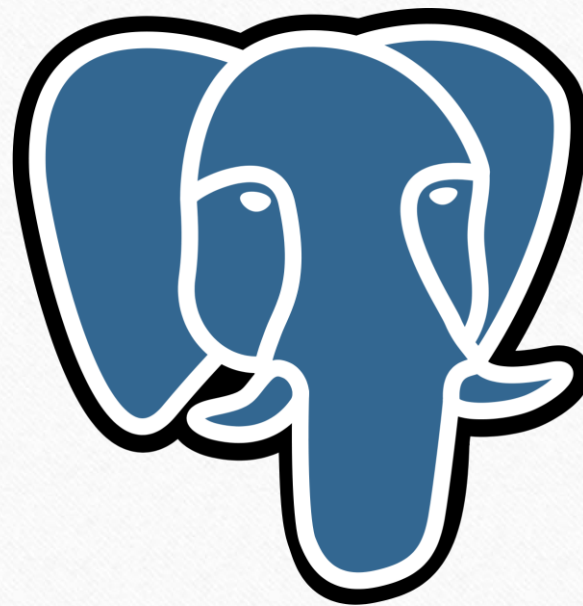
Based on data from 1985 to 2016

Ken Croker

Made With Tableau and PostgreSQL

Outline

Visualizing the Console Game Market
>   Sales Overview
>   Xbox vs PlayStation
>   Ratings Breakdown
>   How to Improve Sales
>   Comparing Publishers

Behind the Scenes: The SQL that Made this Possible
>   Creating a Database in PostgreSQL
>   Finding Insights with Queries

Visualizing the Console Game Market

With Tableau Dashboards

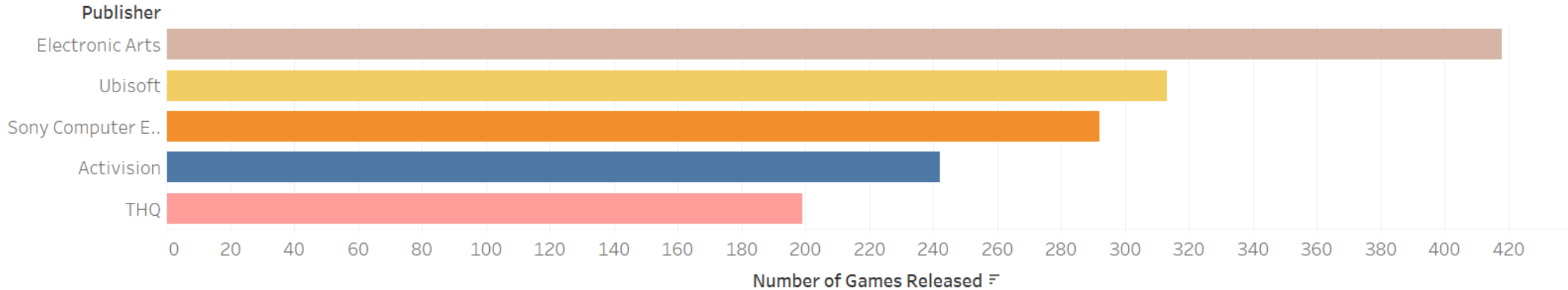# Sales Overview

Publisher
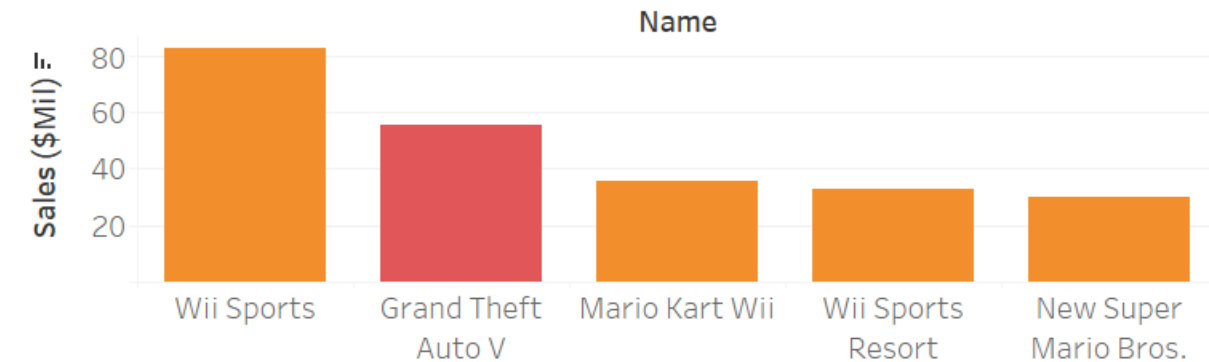Nintendo
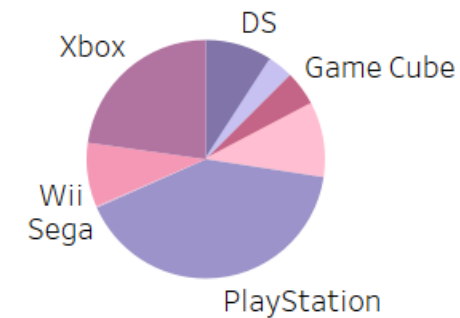Take-Two Interactive

Year of Release
1985 to 2016

Genre
All

Who Published the Most Games?

Publisher

Publisher	Number of Games Released
Electronic Arts	~418
Ubisoft	~313
Sony Computer E..	~291
Activision	~242
THQ	~198

Number of Games Released ⯆

Top Selling Games

Name

Sales ($Mil) ⯆

- Wii Sports — ~84
- Grand Theft Auto V — ~56
- Mario Kart Wii — ~36
- Wii Sports Resort — ~33
- New Super Mario Bros. — ~30

For Which Platforms were Games Developed?

DS, Game Cube, PlayStation, Sega, Wii, Xbox

Xbox vs Playstation

Number of Games Released

Platform
PlayStation
Xbox

Top Playstation Games

Name

Grand Theft Auto V
Grand Theft Auto: San Andreas
Grand Theft Auto: Vice City
Gran Turismo 3: A-Spec
Call of Duty: Black Ops II

Publisher
Activision
Sony Comput..
Take-Two Int..

Global Sales ($ Million)

Top Xbox Games

Name

Kinect Adventures!
Grand Theft Auto V
Call of Duty: Modern Warfare 3
Call of Duty: Black Ops
Call of Duty: Black Ops II

Publisher
Activision
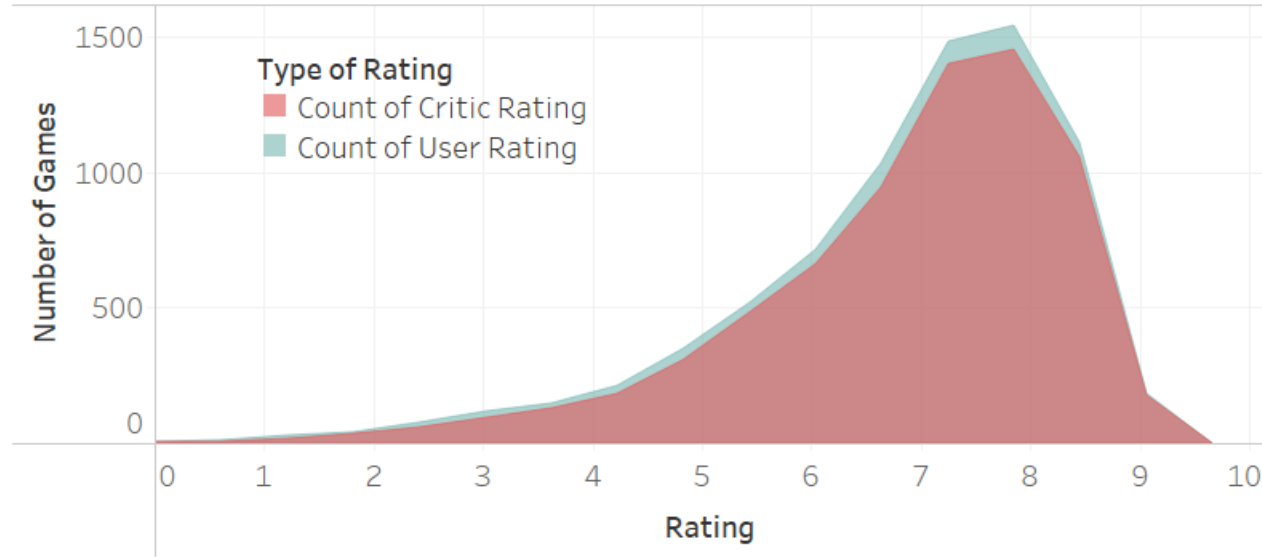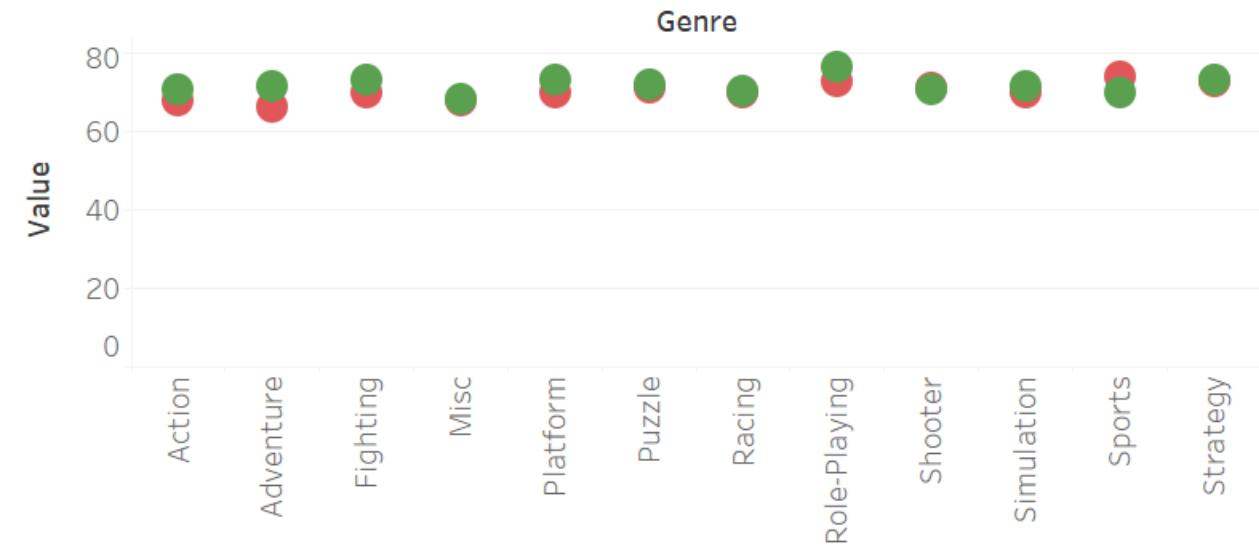Microsoft Ga..
Take-Two Int..

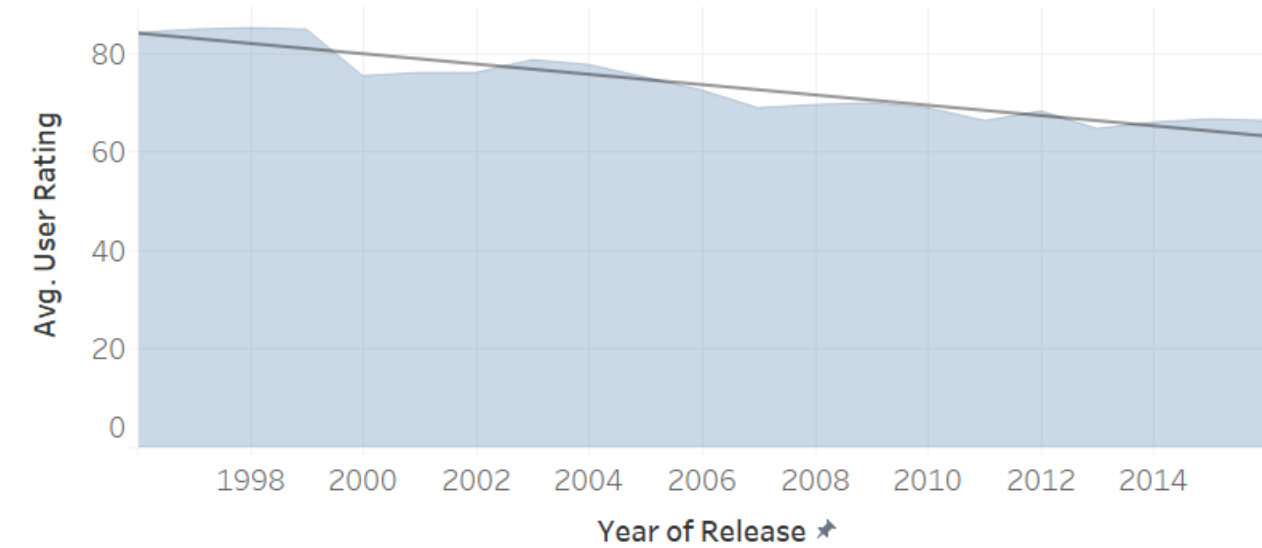Global Sales ($ Million)
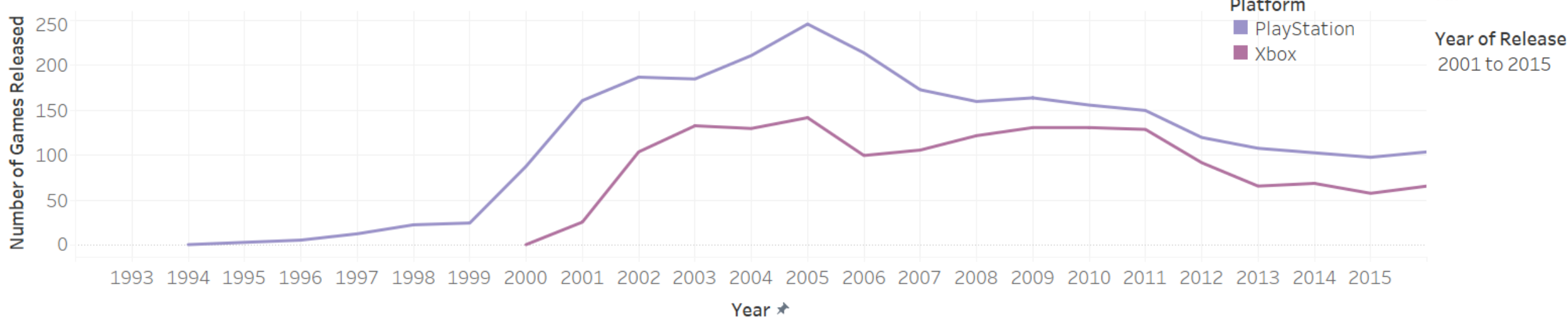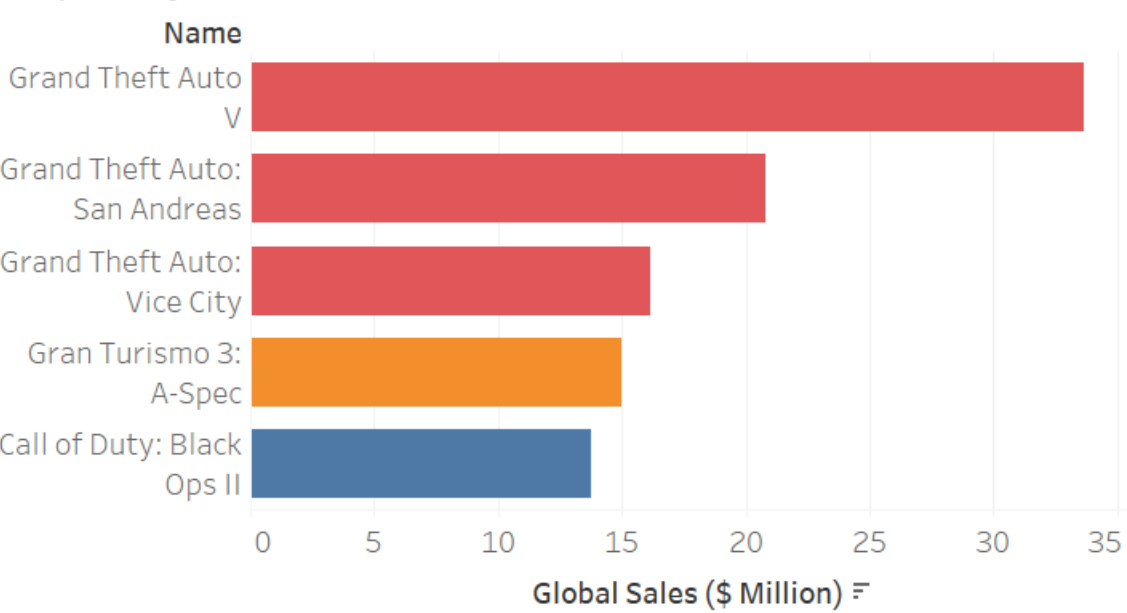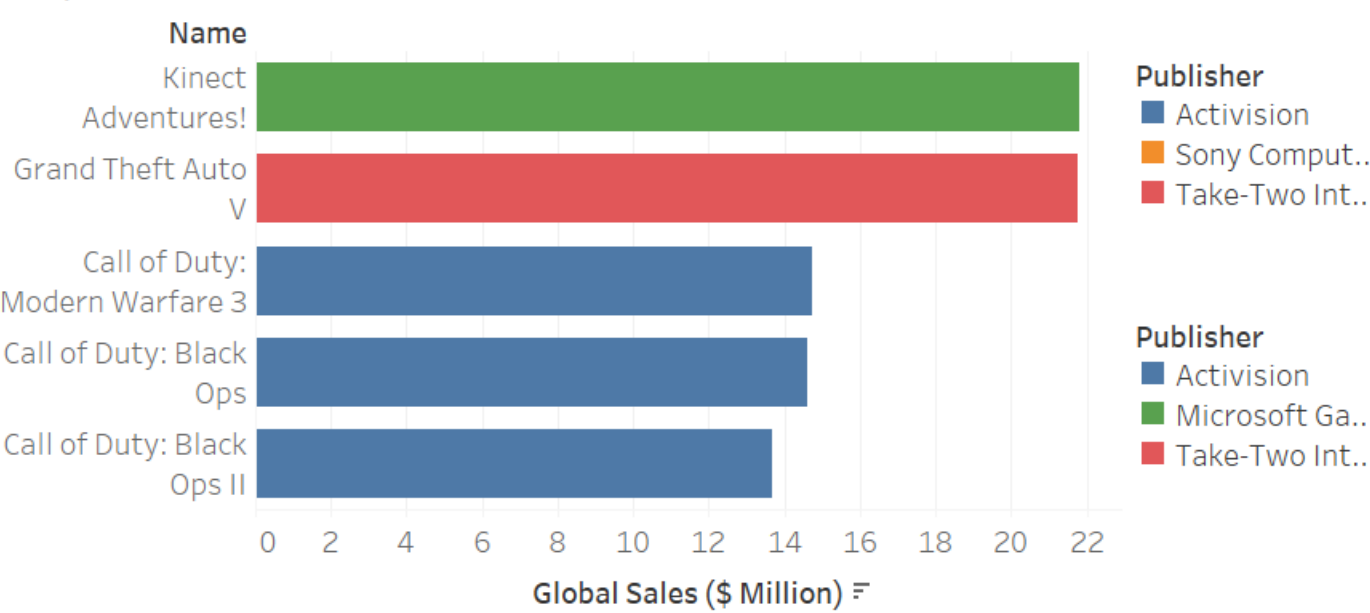
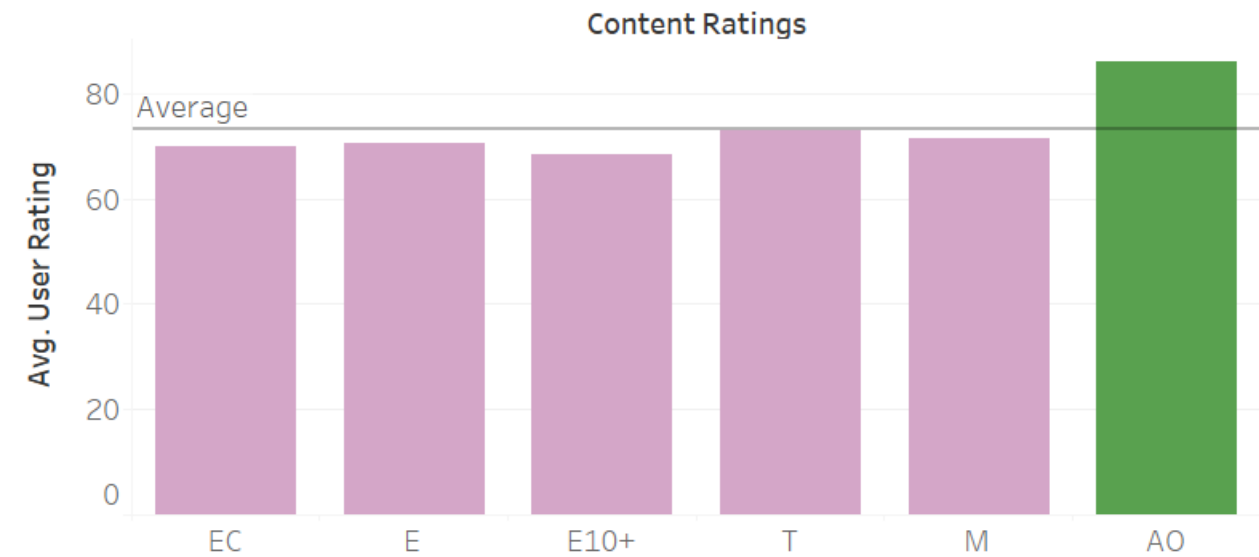# Ratings Breakdown

Ratings Histogram

Ratings Consistent Across Genres
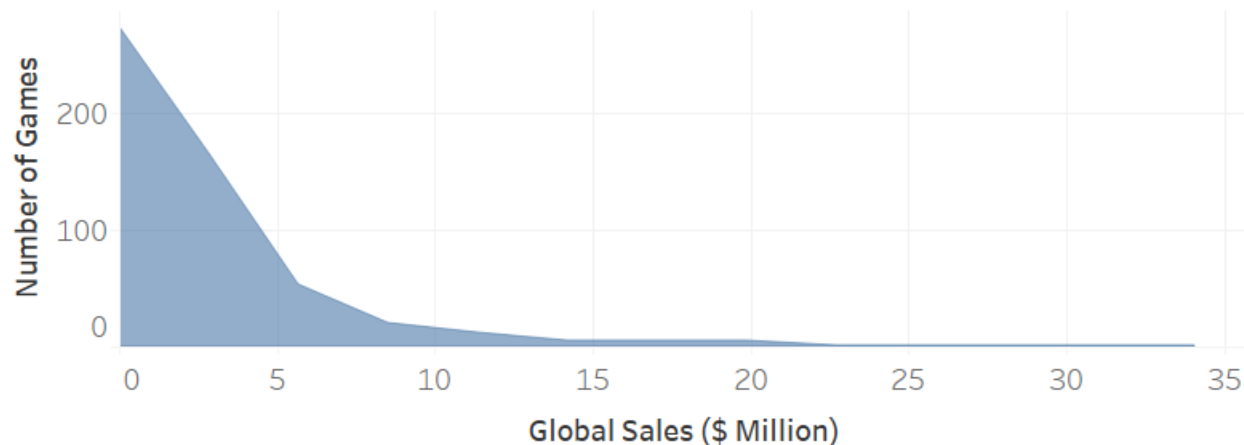

User Ratings Declined Over Time

Adult Only Games Received Higher Ratings

How to Improve Sales

Understand the Uneven Distribution of Sales

Most games generated less than $6 million in sales. A select few games made a LOT more than that.

Focus on North American and European Markets

Region
- Avg. North American S..
- Avg. European Sales
- Avg. Japanese Sales
- Avg. Other Sales
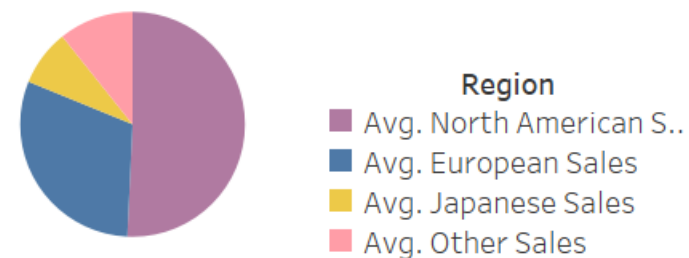
This chart shows the percentage of global sales that came from each region. Most of a video game's sales come from North America and Europe.

Go for Mature and Adult Only Content
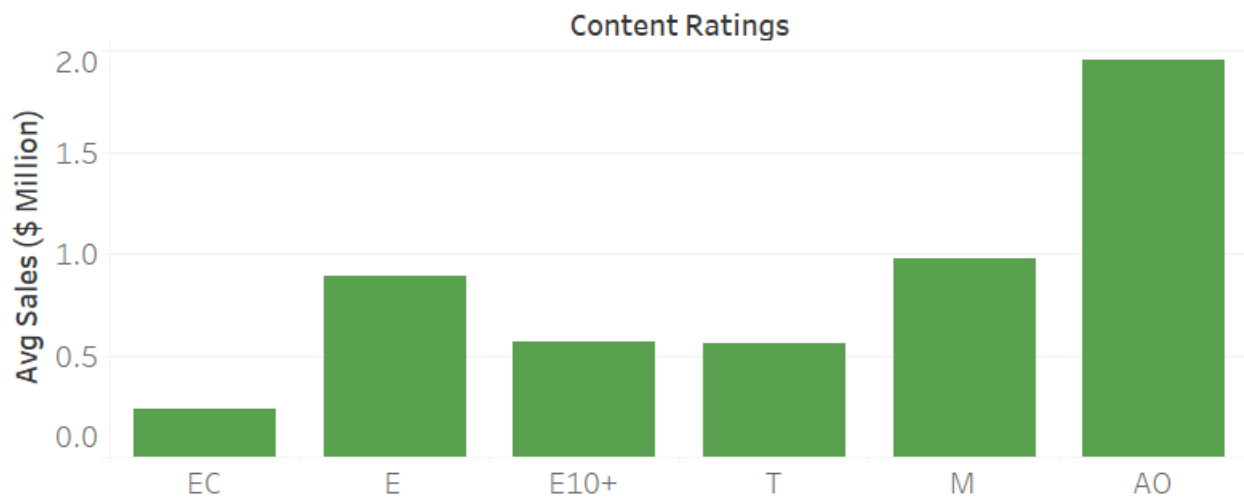
Content Ratings

Games with Mature and Adult Only content tend to generate higher average sales.

Don't Worry Too Much About Ratings
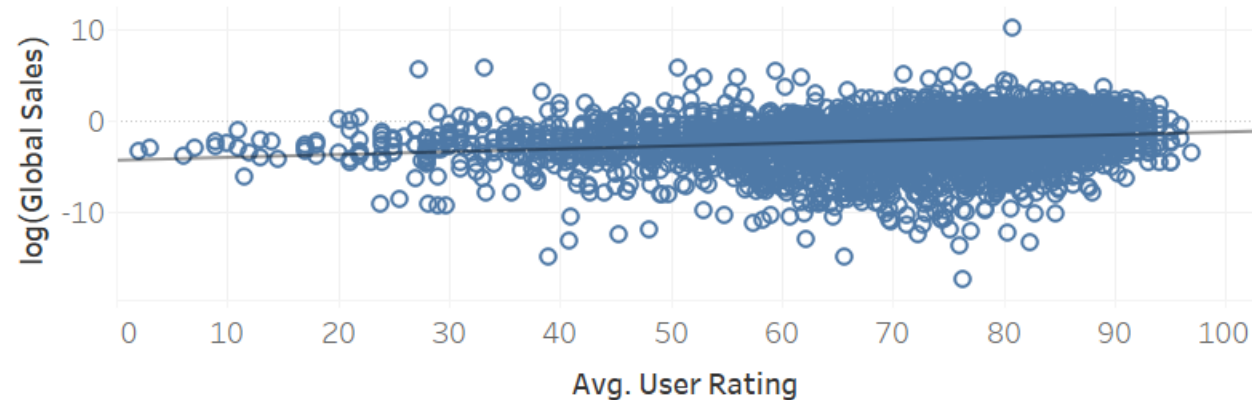
Less than 4% of the variation in (the natural log of) a game's sales can be explained by variations in its user ratings. This suggests that user ratings are not *too* correlated with a game's sales.

Comparing Publishers

A publisher can also increase sales by taking note of which publishers regularly succeed and which publishers struggle.

Conducting more research on the most successful publishers would allow us to discover how their teams operate and how they meet sales targets.

Publishers with Lowest Average Sales

Publisher

Paradox Interact..	$0.04M
JoWood Producti..	$0.05M
Idea Factory	$0.06M
Kalypso Media	$0.07M
DTP Entertainme..	$0.09M

0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4

Average Sales per Game ($ Million)
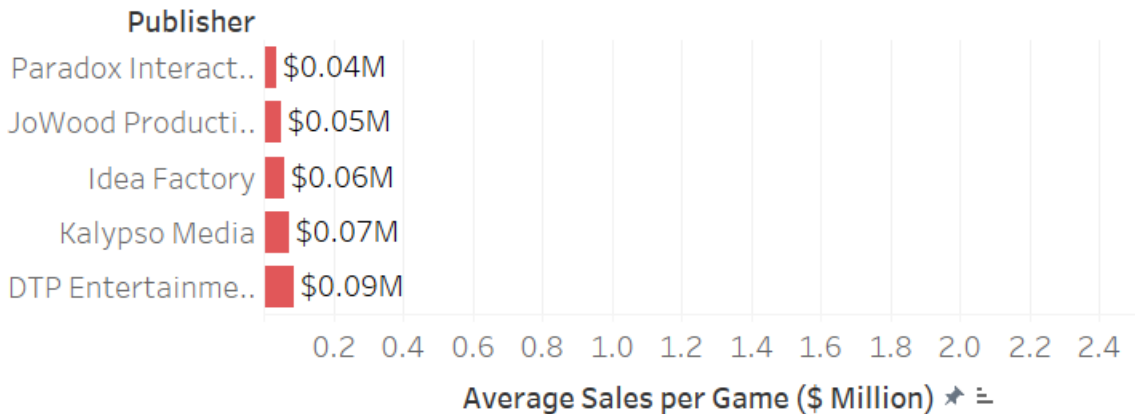
Limited to publishers who have released at least 10 unique titles in a designated time period.

# of Publishers
5

Year of Release
1988 to 2016

Publishers with Highest Average Sales

Publisher

Nintendo	
GT Interactive	
SquareSoft	
RedOctane	
Hello Games	

0.0   0.5   1.0   1.5   2.0   2.5   3.0

Average Sales per Game ($ Million)
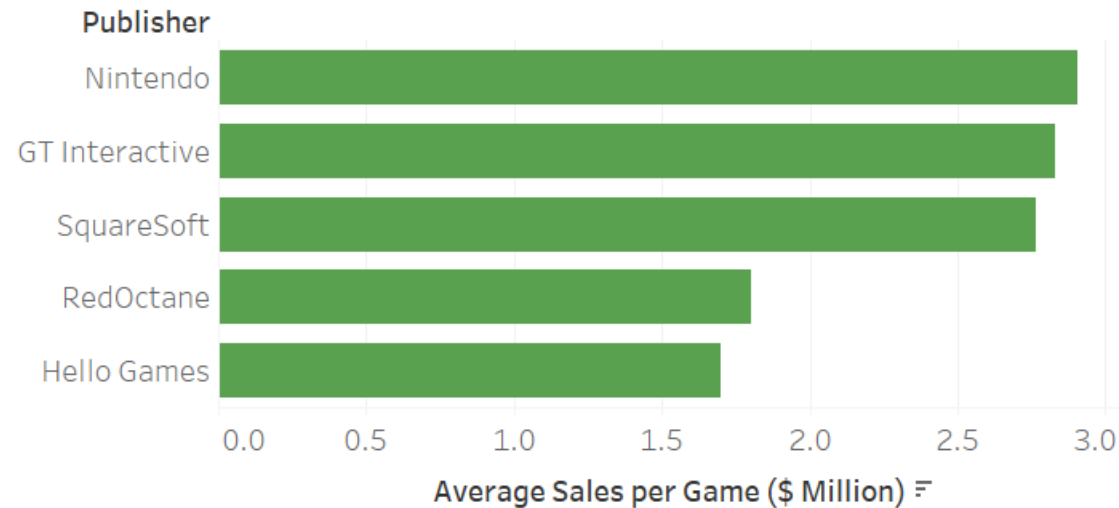
Publishers with Highest Average Ratings

Publisher

GT Interactive	
SquareSoft	
RedOctane	
Nintendo	
Hello Games	

0   10   20   30   40   50   60   70   80   90 100

Average User Rating
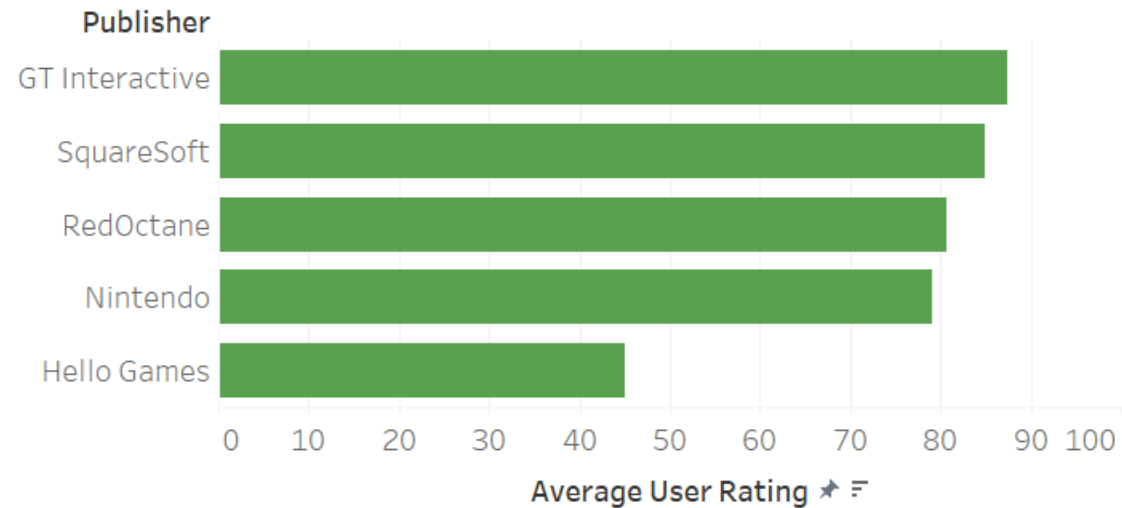
Behind the Scenes

The SQL that made these dashboards possible

Creating a Database

Before I started creating any charts or visualizations in Tableau, I first created a database using PostgreSQL.


```
ken@dellxps:~$ sudo -u postgres psql
psql (12.8 (Ubuntu 12.8-0ubuntu0.20.04.1))
Type "help" for help.

postgres=# CREATE DATABASE gaming;
CREATE DATABASE
postgres=# \c gaming
You are now connected to database "gaming" as user "postgre
gaming=#
```

Creating a Table

After setting up the database, I created a table called 'game_sales' for the videogame sales data.

I paid close attention to ensure that each column took the correct data type (text, int, float, etc).

```
ken@dell
CREATE TABLE game_sales (
        name TEXT,
        platform VARCHAR(10),
        year_released INT,
        genre TEXT,
        publisher TEXT,
        na_sales FLOAT,
        eu_sales FLOAT,
        jp_sales FLOAT,
        other_sales FLOAT,
        global_sales FLOAT,
        critic_score FLOAT,
        critic_count INT,
        user_score FLOAT,
        user_count INT,
        developer TEXT,
        rating TEXT
);
```

Importing the Data

After setting up the game_sales table, I copied the contents of the csv file into it.

Here's the syntax I used.

Distinct Game Titles

I first wanted to know how many unique titles were in the dataset, so I ran a query to COUNT how many DISTINCT values were in the name column.

The value returned is 11553, meaning there are 11553 unique console game titles in the data table.

```
ken@dellxps:~$ sudo -u postgres psql
psql (12.8 (Ubuntu 12.8-0ubuntu0.20.04.1))
Type "help" for help.

postgres=# \c gaming
You are now connected to database "gaming" as user "postgre
gaming=# SELECT COUNT(DISTINCT(name)) FROM game_sales;
 count
-------
 11553
(1 row)

gaming=#
```

Distinct Publishers

I then ran a similar query for the publishers column.

The value returned is 582, meaning there are 582 unique console game publishers in the data table.

```
                                                    ken@dellxps: ~
ken@dellxps:~$ sudo -u postgres psql
psql (12.8 (Ubuntu 12.8-0ubuntu0.20.04.1))
Type "help" for help.

postgres=# \c gaming
You are now connected to database "gaming" as user "postgre
gaming=# SELECT COUNT(DISTINCT(name)) FROM game_sales;
 count
-------
 11553
(1 row)

gaming=# SELECT COUNT(DISTINCT(publisher)) FROM game_sales;
 count
-------
   582
(1 row)

gaming=#
```

Game Titles From Nintendo

Next, I wanted to see all the game titles Nintendo released, so I ran a SELECT statement to return all values in the name column WHERE Nintendo was the publisher.

The screenshot at the bottom shows the first part of the query's output.

```
--Get all titles of games published by Nintendo
SELECT name FROM game_sales
WHERE developer = "Nintendo";
```

```
-----------------------------------------------
---------------------------------
Baseball
Mario Bros.
Mahjong
Donkey Kong
Donkey Kong Jr.
Popeye
Duck Hunt
Excitebike
Golf
Tennis
Pinball
```

PS3 Games in 2011

I wanted to restrict the table to just the top 100 PS3 games released in 2011, measured by global sales. I only wanted each game's name and global sales.

The screenshot at the bottom displays the first part of the output.

```sql
--Get top 100 PS3 games released in 2011 as measured  by global sales
SELECT DISTINCT(name), global_sales FROM game_sales
WHERE year_released = 2011 AND platform = 'PS3'
ORDER BY global_sales DESC
LIMIT 100;
```

```
                     name                    | global_sales
---------------------------------------------+--------------
 Call of Duty: Modern Warfare 3              |        13.32
 Battlefield 3                               |         7.17
 Uncharted 3: Drake's Deception             |         6.74
 FIFA 12                                     |         6.65
 The Elder Scrolls V: Skyrim                 |         6.41
 Batman: Arkham City                         |         5.48
 Assassin's Creed: Revelations              |         4.22
 LittleBigPlanet 2                           |         3.39
 L.A. Noire                                  |         3.17
 Mortal Kombat                               |         2.84
 Killzone 3                                  |         2.77
 Final Fantasy XIII-2                        |         2.63
 Saints Row: The Third                       |         2.47
```

EA's Dominance in 2011

I wondered how many of these games were published by Electronic Arts, so I ran a nested query to find out. The inner query returns the output shown in the last slide. The outer query selects the entries in the output with a publisher value of 'Electronic Arts'.

The number 56 is returned, meaning Electronic Arts made 56 of the 100 top selling PS3 games released in 2011.

```
SELECT COUNT(*) FROM game_sales
WHERE name IN (
        SELECT name FROM game_sales
        WHERE year_released = 2011 AND platform = 'PS3'
        ORDER BY global_sales DESC
        LIMIT 100)
AND publisher LIKE '%Electronic Arts%' OR publisher LIKE '%EA%';
```

```
gaming(#  LIMIT 100)
gaming-# AND publisher LIKE '%Elect
 count
-------
    56
(1 row)

gaming=#
```

Top Selling Games Released in 2016

Here's a query that returns the entries of the top 100 selling games released in 2016.

```
--Find the top selling 100 games of 2016.
SELECT DISTINCT(name), publisher, global_sales, platform FROM game_sales
WHERE year_released = 2016
ORDER BY global_sales DESC
LIMIT 100;
```

```
                          name                          |                 publisher
          | global_sales | platform
----------------------------------------------------------+----------------------------
---------+--------------+----------
FIFA 17                                                  | Electronic Arts
          |         7.59 | PS4
Pokemon Sun/Moon                                         | Nintendo
          |         7.14 | 3DS
Uncharted 4: A Thief's End                               | Sony Computer Entertainment
          |         5.38 | PS4
Call of Duty: Infinite Warfare                           | Activision
          |         4.46 | PS4
Battlefield 1                                            | Electronic Arts
          |         4.08 | PS4
Tom Clancy's The Division                                | Ubisoft
          |          3.8 | PS4
FIFA 17                                                  | Electronic Arts
          |         2.65 | XOne
Call of Duty: Infinite Warfare                           | Activision
          |         2.42 | XOne
Far Cry: Primal                                          | Ubisoft
          |         2.26 | PS4
```

Multiple Platforms Not Required

I wondered how many platforms the top selling games were released on, so I decided to run yet another query.

The values in the platform count column are varied. Some of the top selling games were released on 5 platforms, some were released on just one. Releasing a game on several platforms is NOT a requirement for getting sales.

```sql
--List all games, ordering them by descending global sales
--How much did each game sell worldwide?
--How many platforms were each of these games released on?
SELECT name, SUM(global_sales) s, COUNT(DISTINCT(platform)) FROM game_sales
GROUP BY name
ORDER BY s DESC;
```

```
                              |        s        | count
------------------------------------------------------+-------------------+------
------------------------------------------------------+-------------------+------
Wii Sports
                              |           82.53 |     1
Grand Theft Auto V
                              |           56.57 |     5
Super Mario Bros.
                              |           45.31 |     2
Tetris
                              |           35.84 |     2
Mario Kart Wii
                              |           35.52 |     1
Wii Sports Resort
                              |           32.77 |     1
Pokemon Red/Pokemon Blue
                              |           31.37 |     1
Call of Duty: Black Ops
                              |           30.82 |     5
Call of Duty: Modern Warfare 3
                              |           30.59 |     4
New Super Mario Bros.
```

Thank you for your time!

I learned a lot about the console game market through this project, and I hope you learned something interesting too!