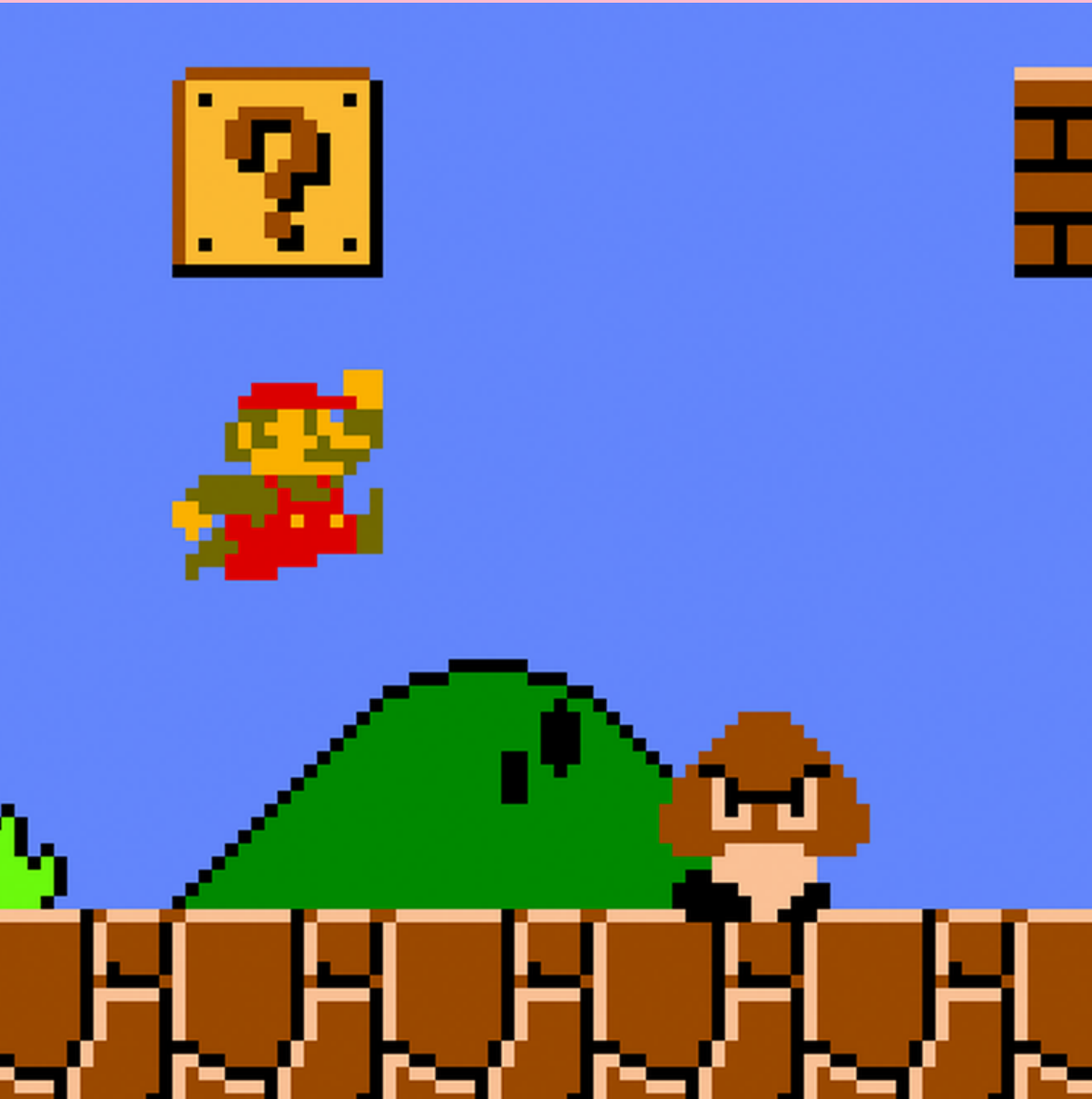


# GOOD GAME; WELL PLAYED?

*The Makings of a Successful Videogame*

Joanie Ipili  
Patricia Matias  
Alyssa Paglumotan  
Mei Villanueva





# IT ALL STARTED WITH THIS GUY

Mickey Mouse who?

Mario revolutionized the 2D and 3D gaming platform

And sold over 500 million copies worldwide!

# SINCE THEN, VIDEOGAMES TOOK ON POP CULTURE

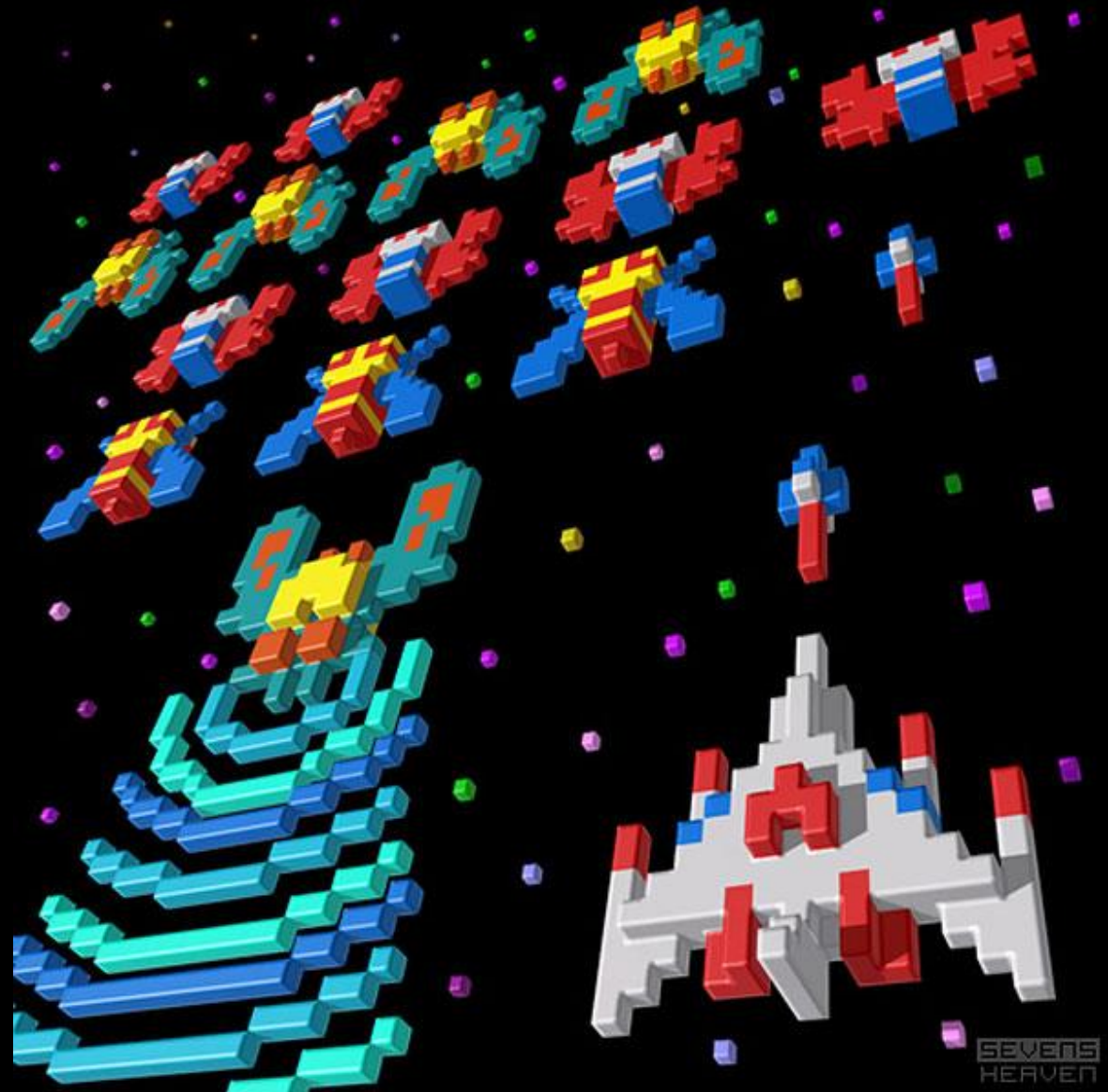
And gave rise to millions of games in different genres, replicated in multiple platforms



# IT'S THE BEST TIME TO CREATE GAMES

Gaming industry is one of the fastest rising markets in the world with **134.9B** in annual total revenue reported in 2018 worldwide, expected to generate **152.1B** by the end of 2019.

*GamesIndustry.biz*

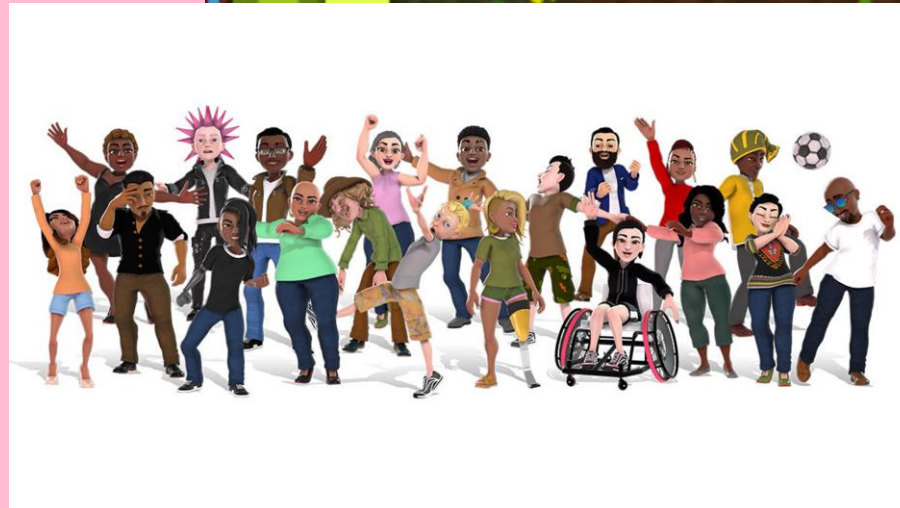


SEVENS  
HEAVEN



# ONLINE IS CHANGING THE GAME

Scale is massive  
Interactions are enormous  
Activity is trackable



# GAMES LIVE AND DIE

With the sheer number of games being produced, it's no surprise that some make it big, while others don't.



# NOT ALL GAMES ARE CREATED EQUAL

For a game publisher, it's quite a challenge.

*What makes a game successful? Which type of game  
has the best market potential?  
What should I optimize for?*





1,000,000 GAMES

1,000,000,000,000 PLAYTIME

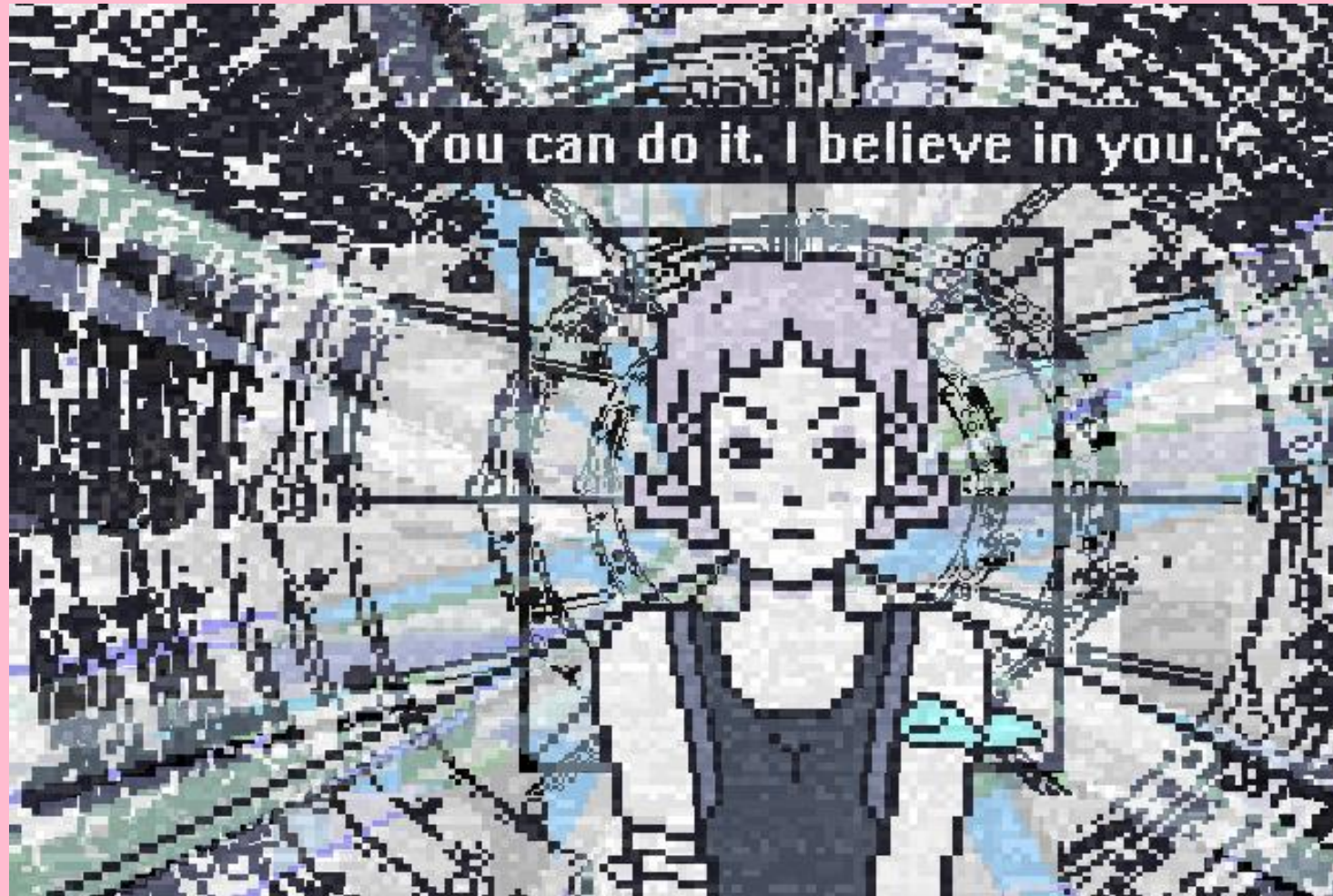
*DATA DATA DATA*





# SO HERE WE ARE

We want to try to describe and predict the performance of games across different genres, with the end goal of identifying the core factors critical to product success.



*With the help of data ;)*

# WHAT'S IN IT?



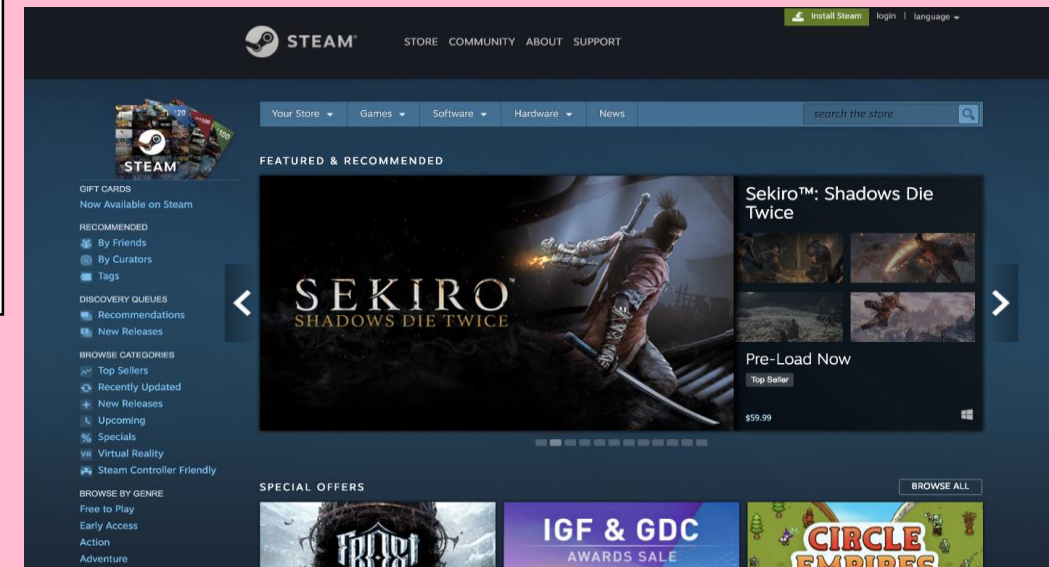
1. Game app data – Price, Genre, Multiplayer, Release Date, Developer, Publisher

2. Ratings

3. Achievement Percentage

4. So much more we can't extract :-)

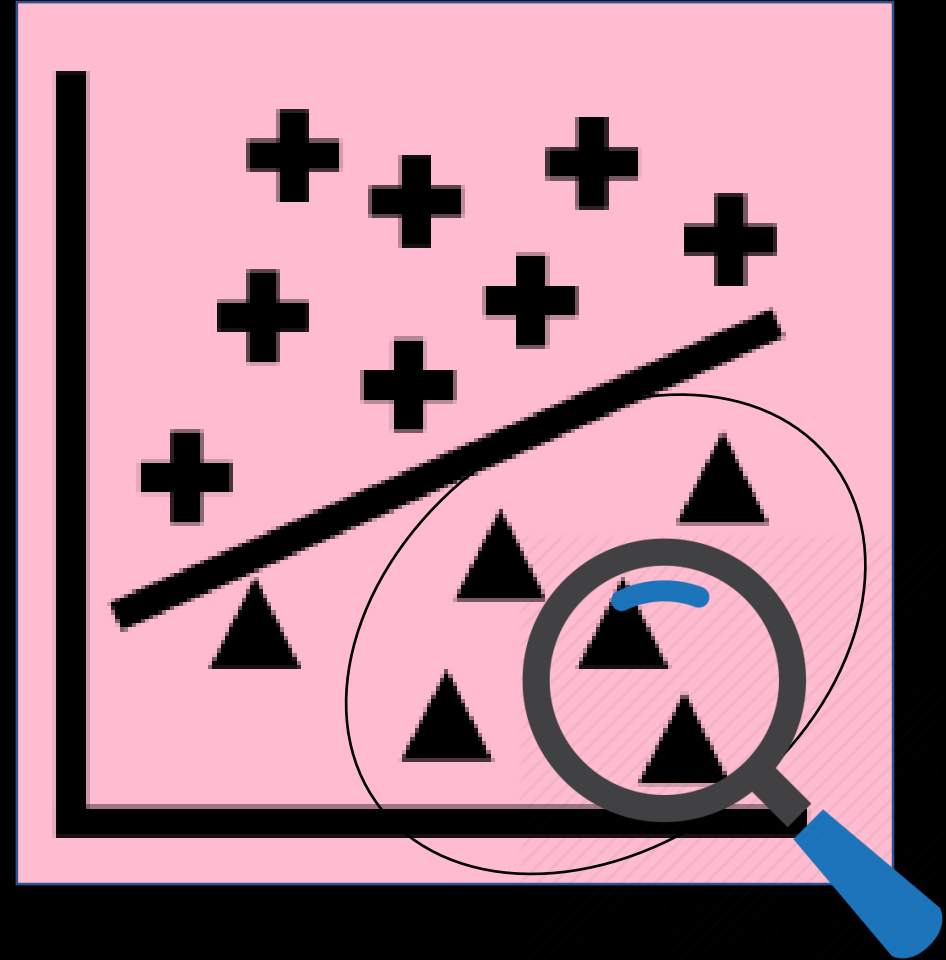
## DATA FROM STEAM



	Title	Type	Price	Free_or_paid	High_priced	Release_Date	Release_Day	Release_Month	Release_Year	Release_Spring	...
App_ID											
220	Half-Life 2	game	9.99	1	0	11/18/2004	18	11	2004	0	...
240	Counter-Strike: Source	game	19.99	1	1	11/01/2004	1	11	2004	0	...
300	Day of Defeat: Source	game	9.99	1	0	07/12/2010	12	7	2010	0	...
380	Half-Life 2: Episode One	game	7.99	1	0	08/01/2006	1	8	2006	0	...

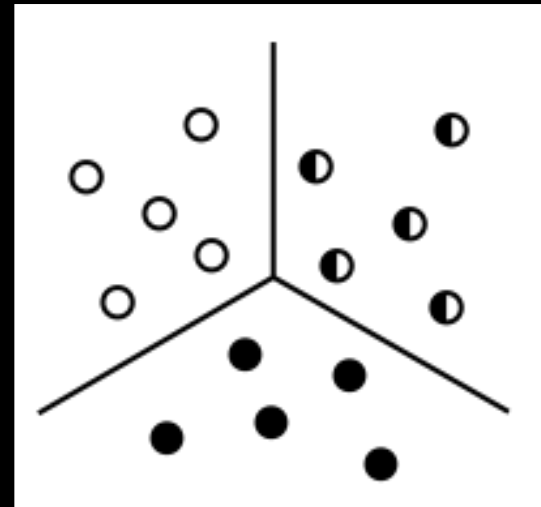
# SO, HOW?

1. Cluster games based on market performance with key indicators a) ratings and b) achievement percentage
2. Identify factors affecting market performance
3. Describe performance clusters based on identified factors



# WHAT WE DID

1. 180gb extracted, struggled to open, cleaned, and feature-engineered using EmEditor, SQL, Excel, OpenRefine
2. Ran K-means algorithm to segment the games based on performance
3. Ran a data analysis to describe the similarities and differences of the identified performance clusters
4. Analyzed!





# MATRIX OF GAMING EXCELLENCE

## Performance Factors

1. Satisfaction, depicted by *Ratings*
2. Usage and Engagement, depicted by *Achievement Percentage*



# EARLY FINDINGS

**0 = Easy wins:**

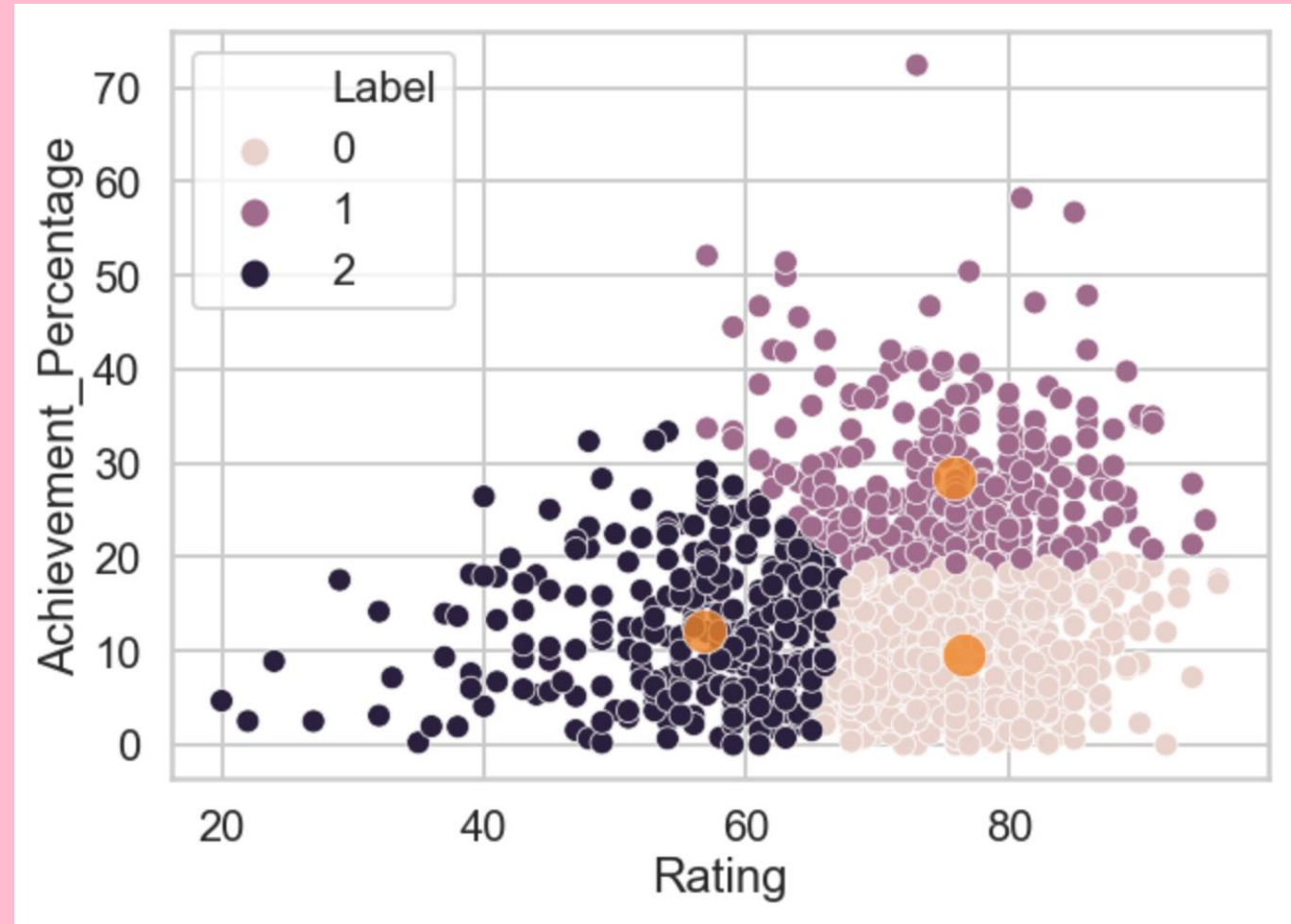
Low Completion, High Rating

**1 = Legends:**

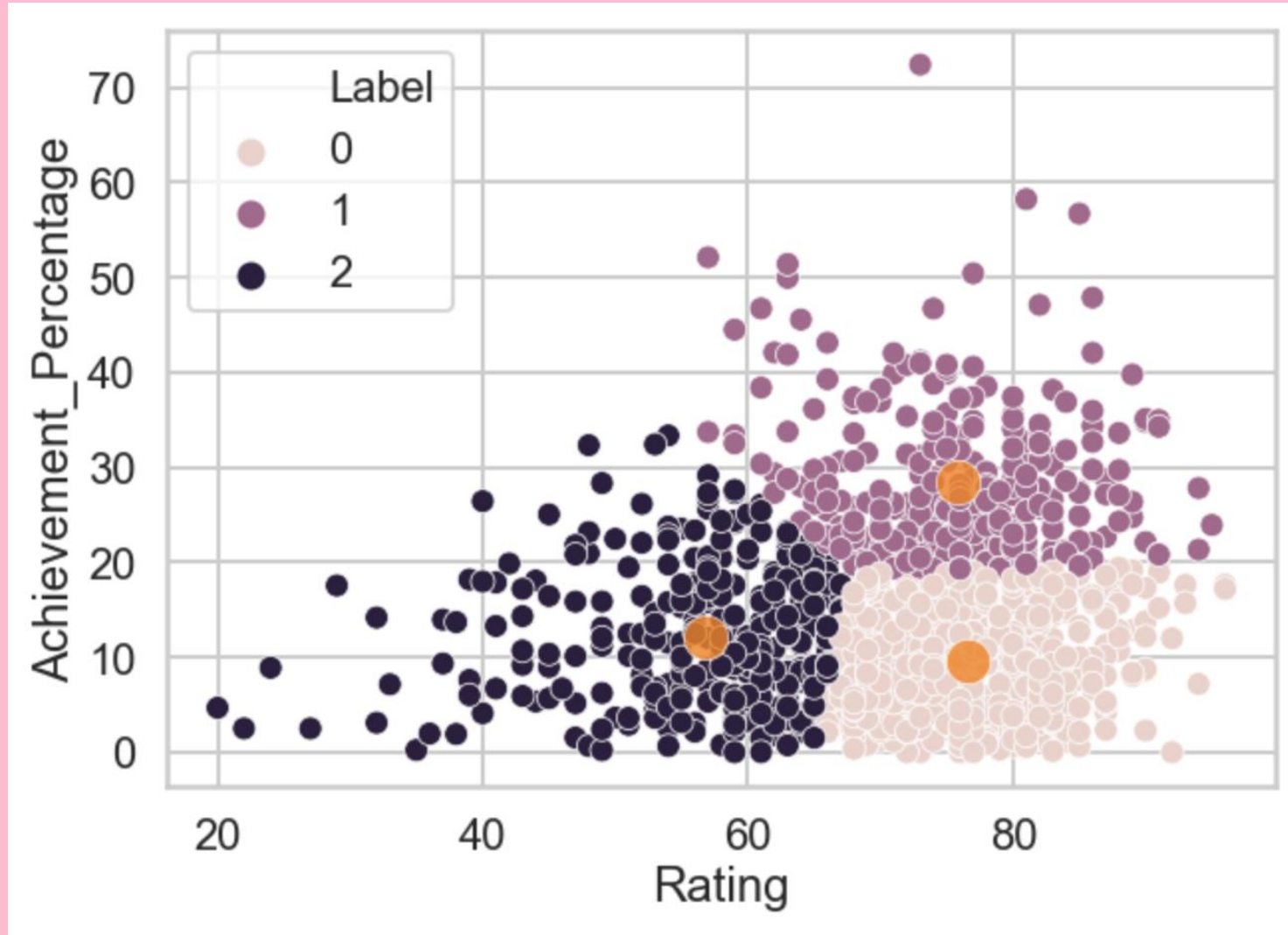
High Completion, High Rating

**2 = Steady Sheep:**

Mid Completion, Low Rating



# EARLY THOUGHTS



**2 = Steady Sheep:**  
Try not to invest too  
much resources in  
this category.

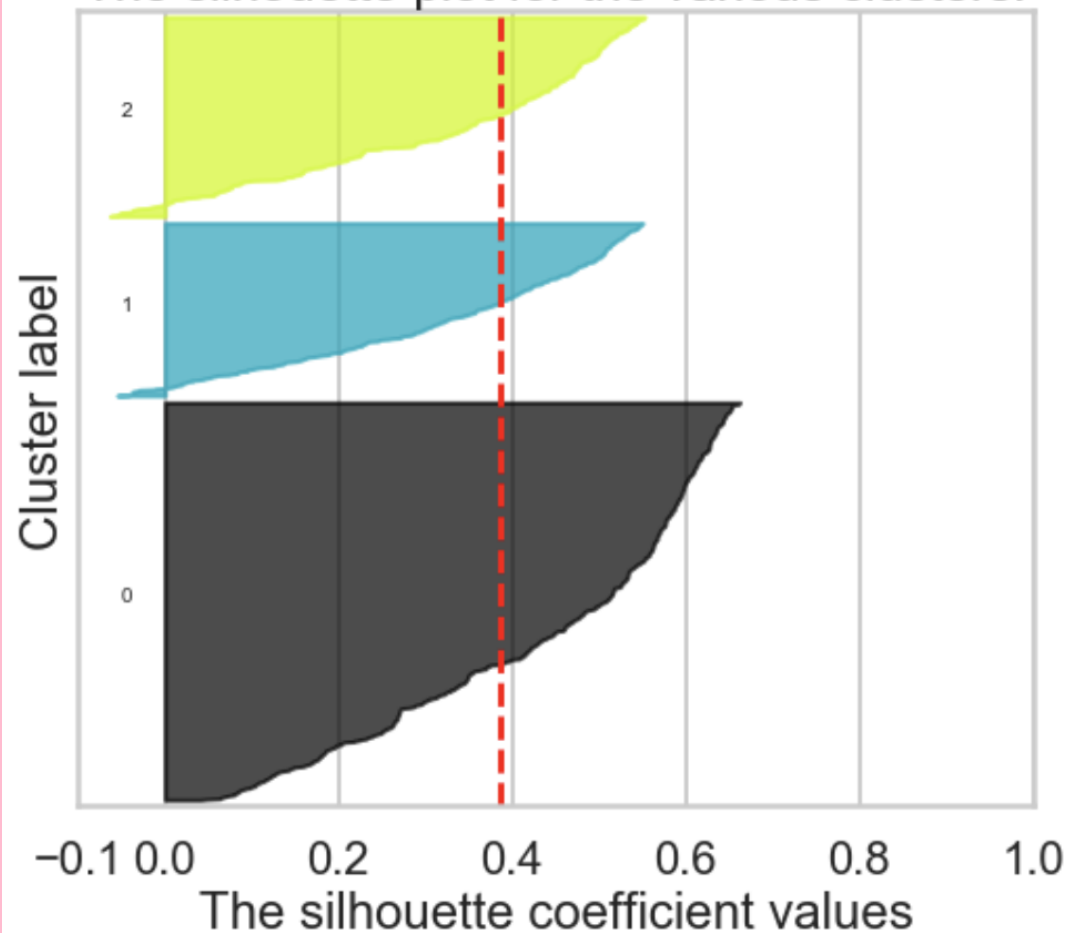
**1 = Legends:**  
Aim for these!

**0 = Easy wins:**  
Minimum investment;  
maximum returns.

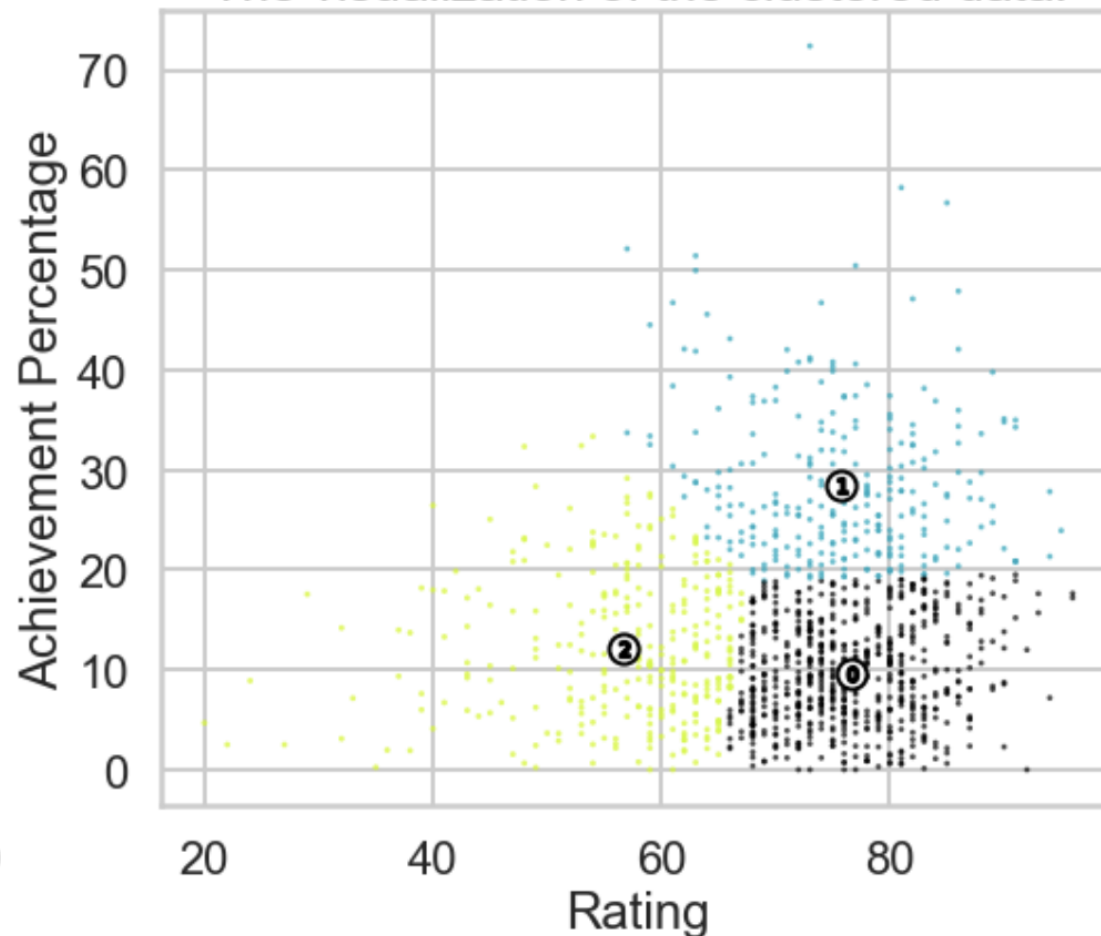
# JUST TO CHECK

Silhouette analysis for KMeans clustering on sample data with  $n\_clusters = 3$

The silhouette plot for the various clusters.



The visualization of the clustered data.





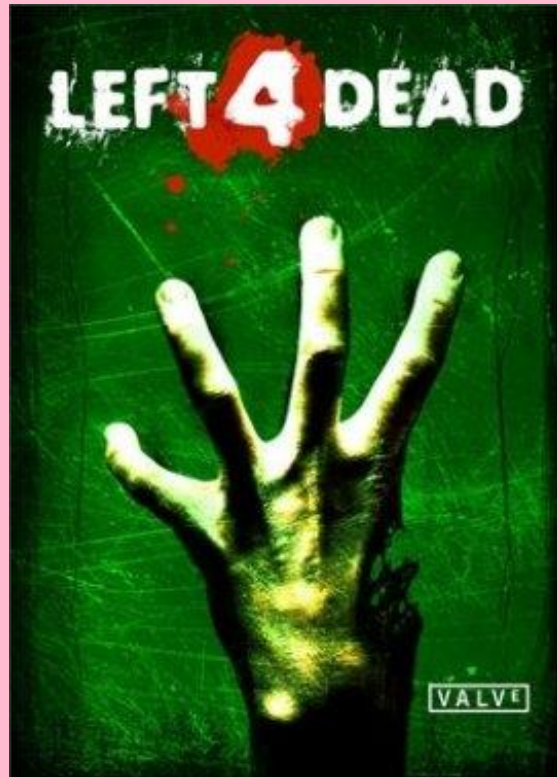
## A pixel art screenshot from the video game Super Mario Bros. The scene is set in a level with a brown brick floor. On the left, Mario, wearing his iconic red cap and overalls, stands on a platform. He is looking towards the right. In the center-right, a Goomba enemy, a small grey creature with a white belly and a single eye, is walking towards Mario. The background consists of a large, stylized mountain with a blue and white color scheme, set against a bright blue sky with white clouds. The overall aesthetic is that of a classic 8-bit video game.

## *What makes up a Legend?*

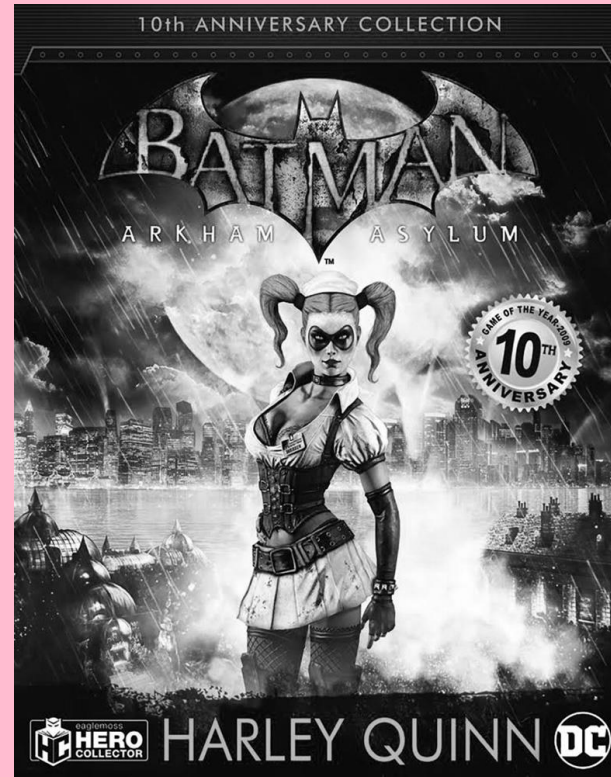
- Specific genres? Free or paid? Timing? Level of interaction?

## How does pricing affect performance?

# ARE LEGENDS MULTIPLAYER?



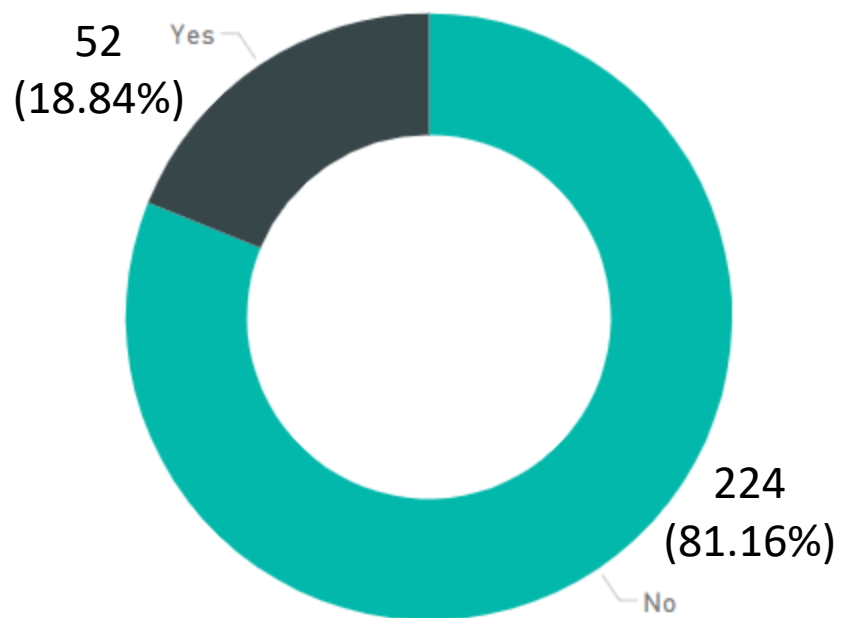
A. For sure, the more the merrier!



B. Nah, single player wins!

# ARE LEGENDS MULTIPLAYER?

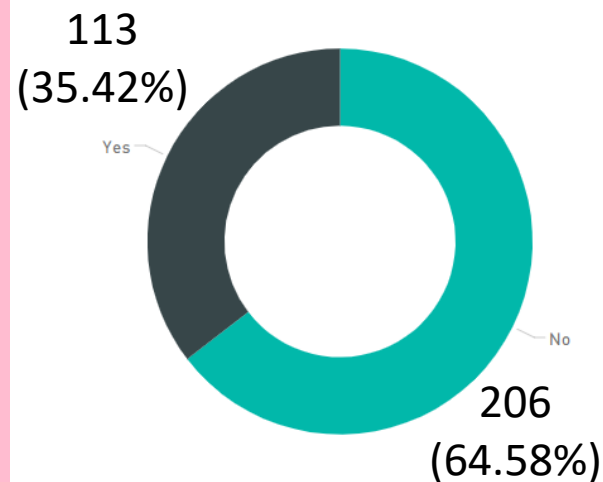
Count of Multiplayer by Multiplayer



Legends

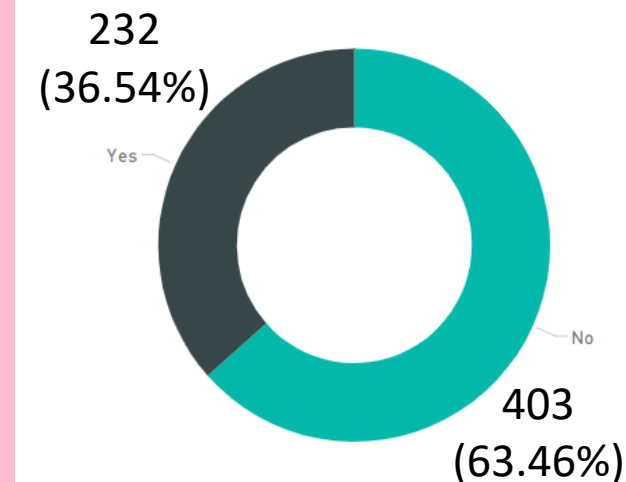
Legends has a much larger percentage of *single player games* than Steady Sheep and Easy Wins.

Count of Multiplayer by Multiplayer



Steady Sheep

Count of Multiplayer by Multiplayer



Easy Wins

# WHICH GENRE DOMINATES?



A. Indie



B. Racing



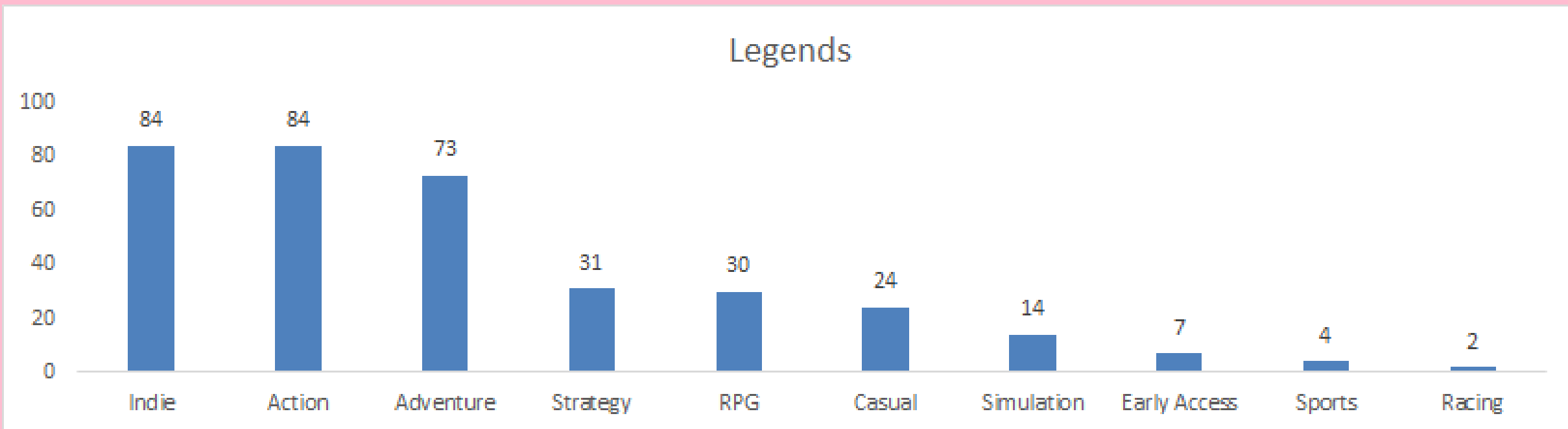
C. Action



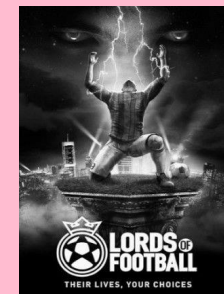
D. Sports



# LEGENDARY GENRES

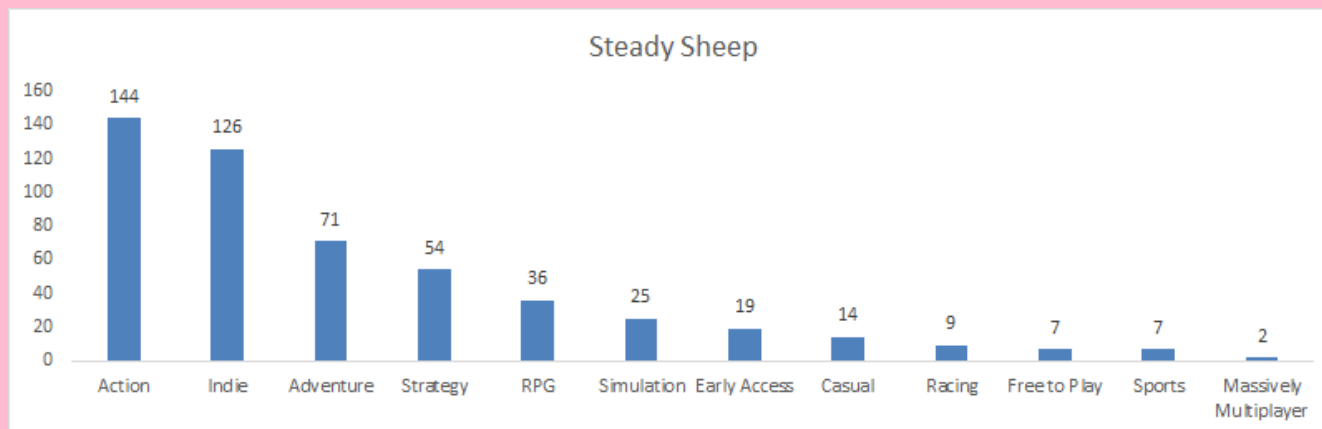
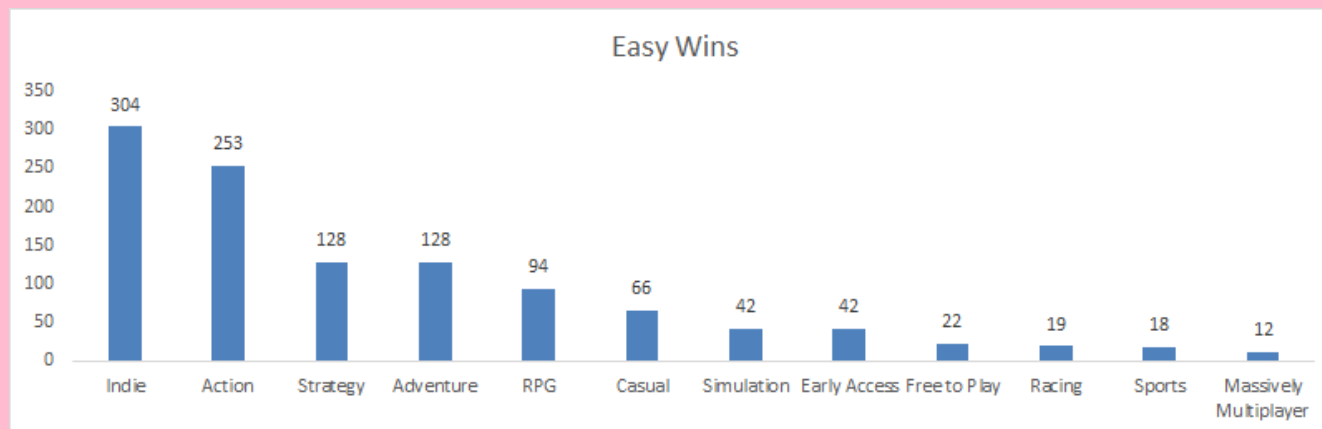
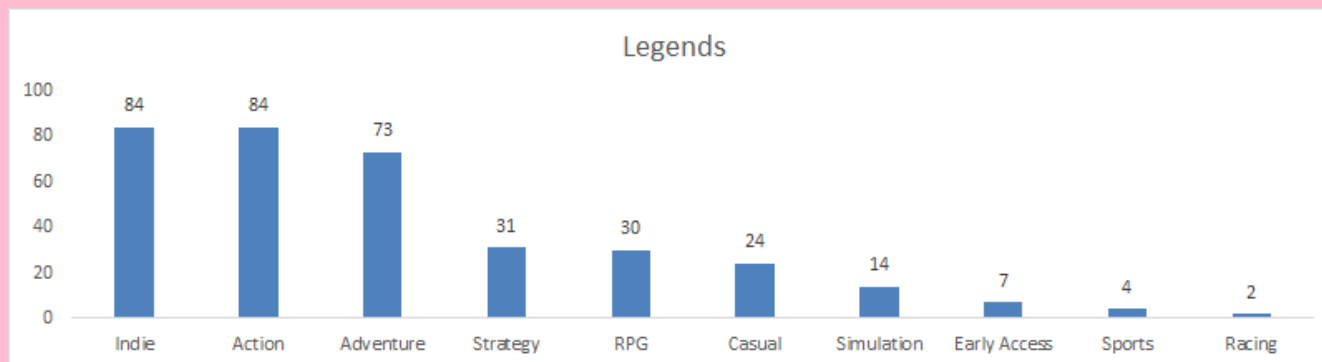


- Indie, Action, and Adventure rule!
- Sports and Racing don't perform as well as we've expected.



# LEGENDS CAN STAND ALONE

- Legendary games don't need social pull. *Massively multiplayer games* are absent in Legends.



# GOOD NEWS!

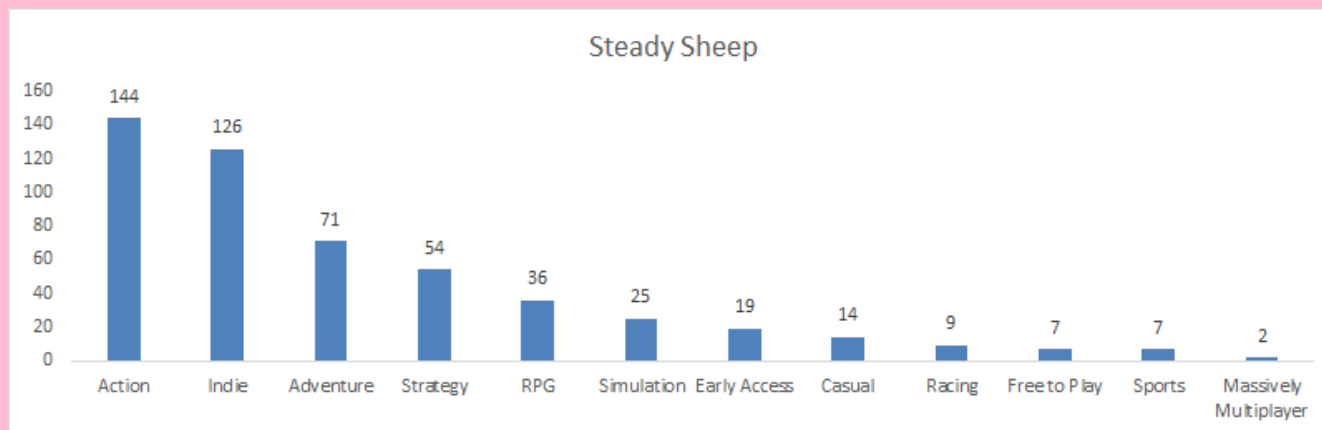
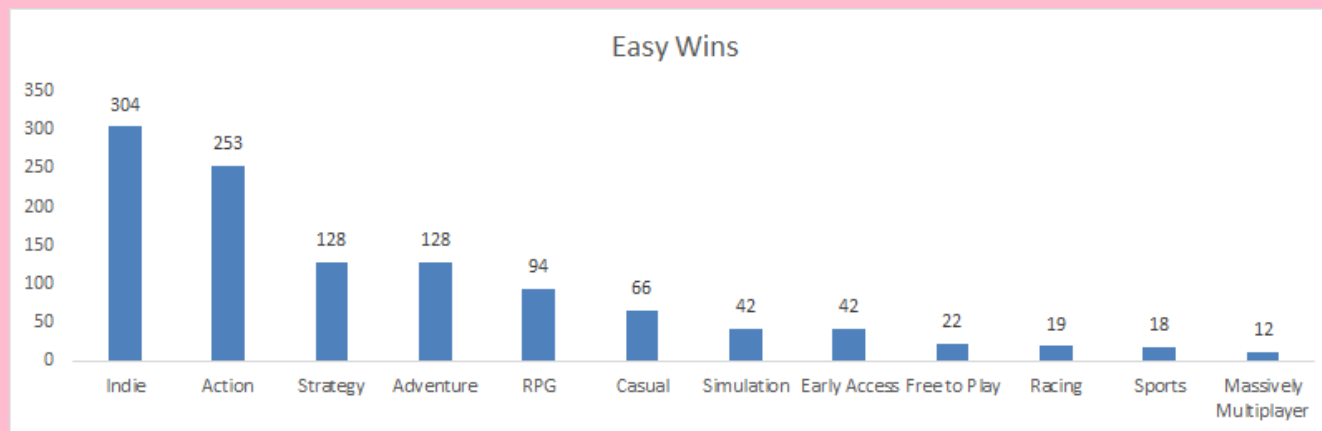
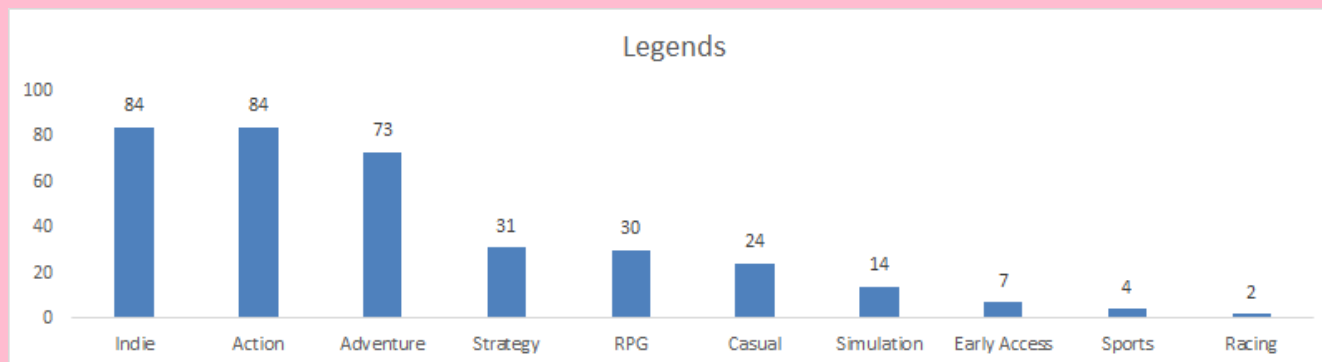
You don't need \$\$\$ from a large publisher to make it to the top of the gaming industry.

*Indie games*, created by small independent developers, prove to be successful among Legends and Easy Wins.

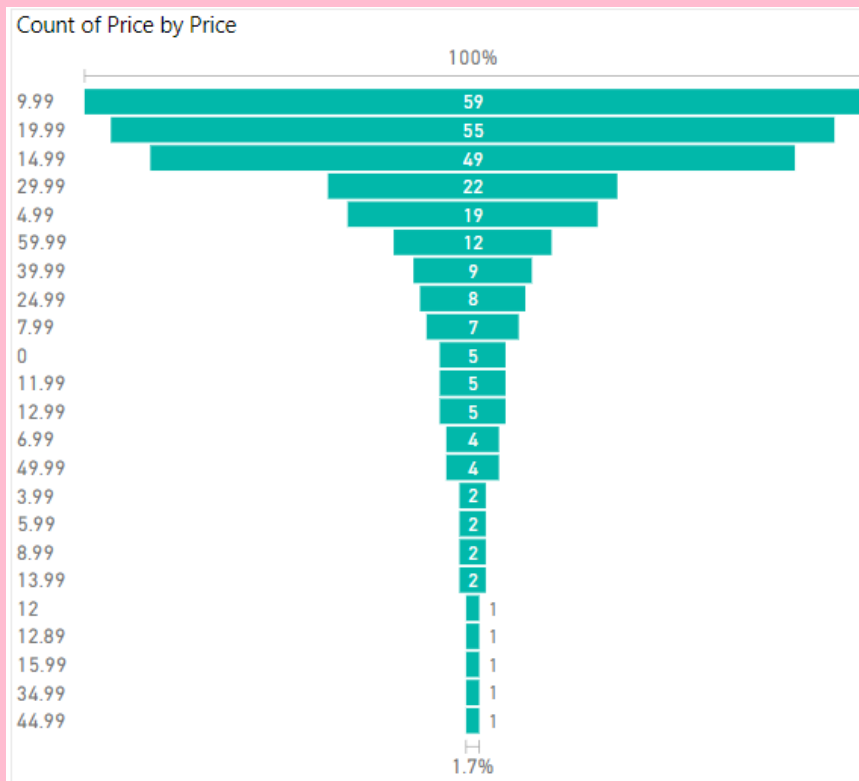


# LEGENDS COME AT A COST!

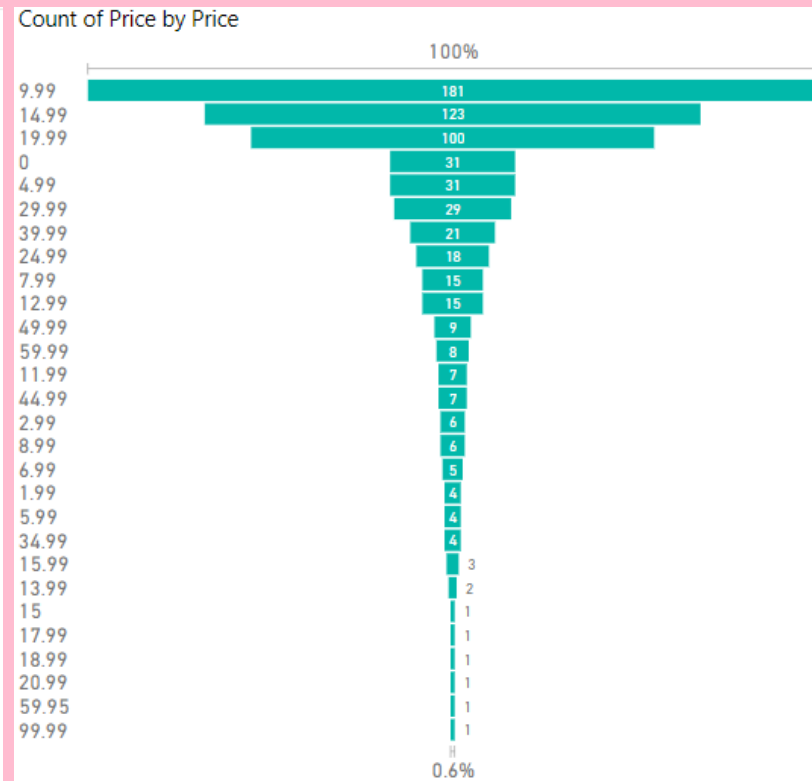
- There are no "Free to Play" games in Legends, unlike other clusters.



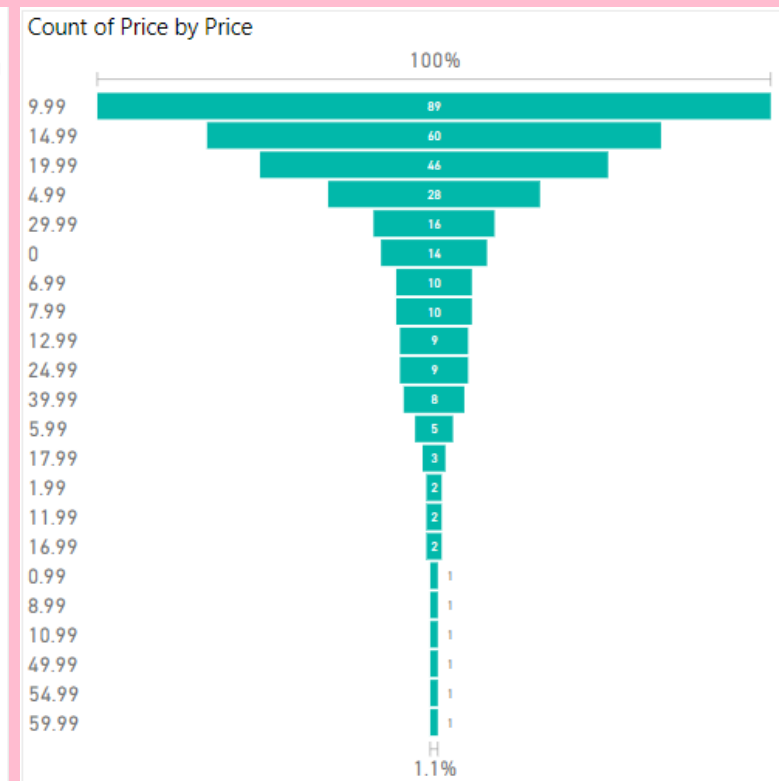




Legends



Easy Wins



Steady Sheep

# DON'T PANIC! KEY TAKEAWAYS:

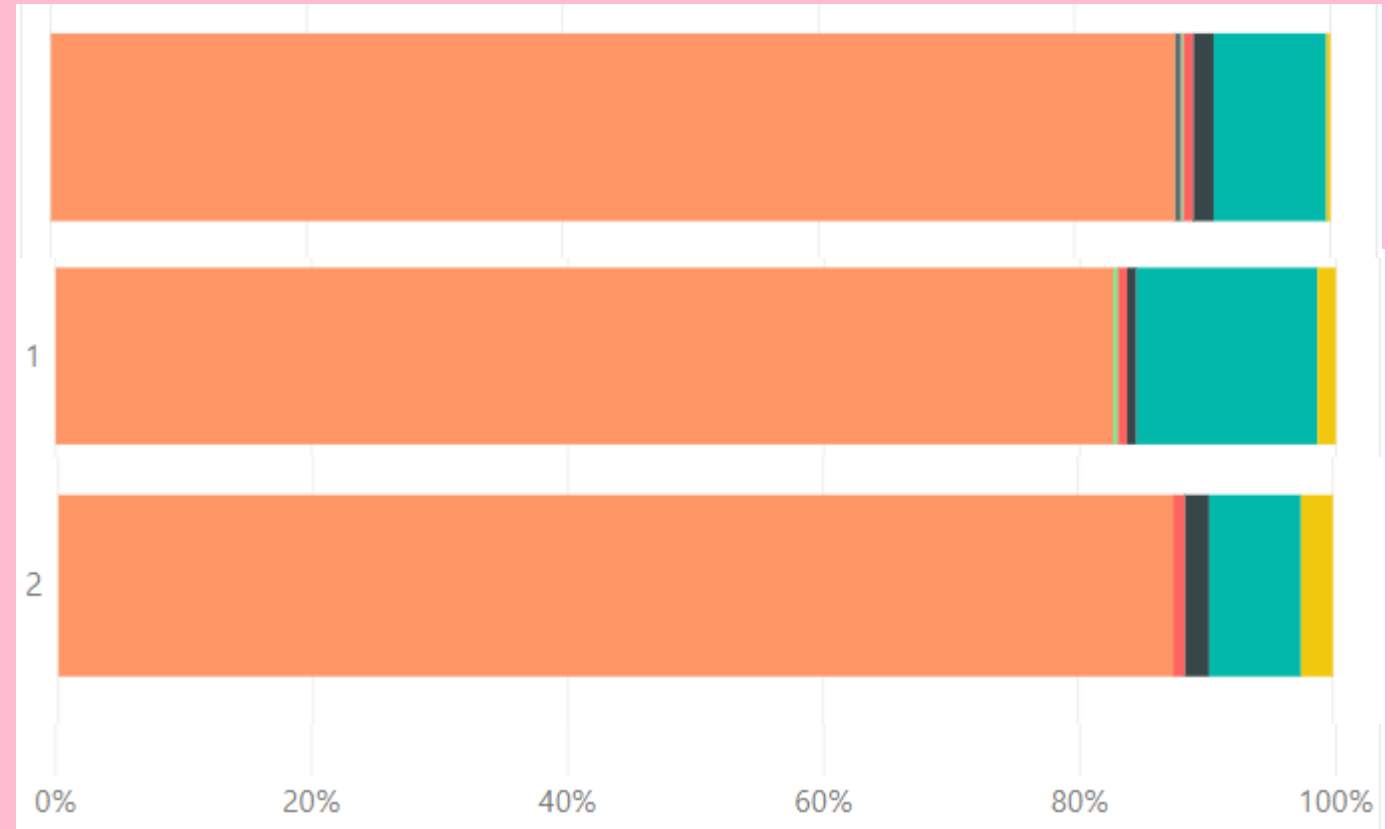
- Legends are \$\$\$ generators. They don't come free, and they are pretty high-priced (Has the highest count for USD 29.99 and 59.99)
- Easy wins are cheap compared to others!

# LET'S TALK AGE

- Good news for you, young ones! Most games across all clusters generally don't have an age requirement.
- Legends, though, need **mature gamers** more than other clusters! You need to be 17 and up to play 17% of the legendary games.

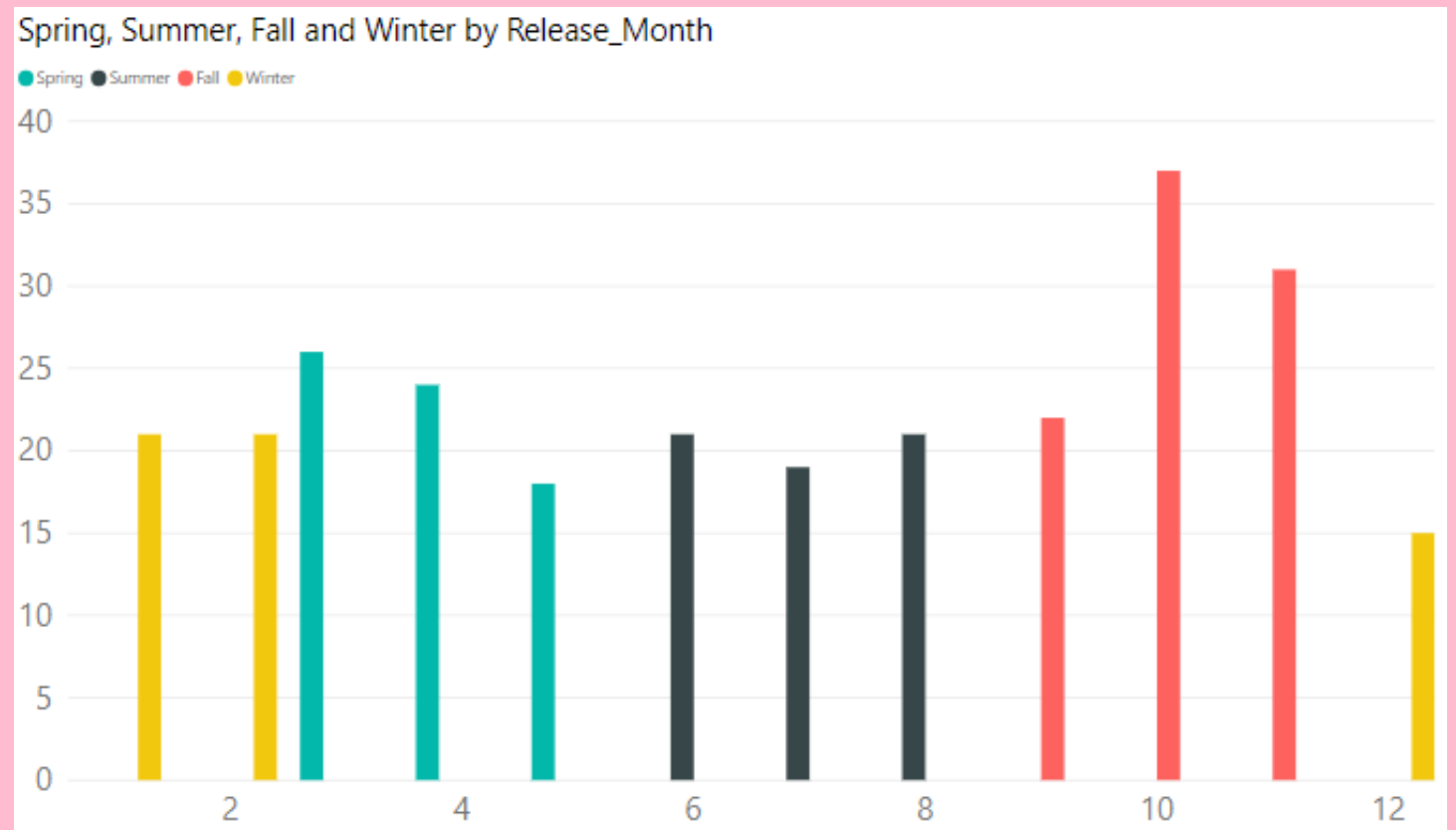
Count of Required\_Age by Cluster\_Label and Required\_Age

Required\_Age 0 12 13 16 17 18



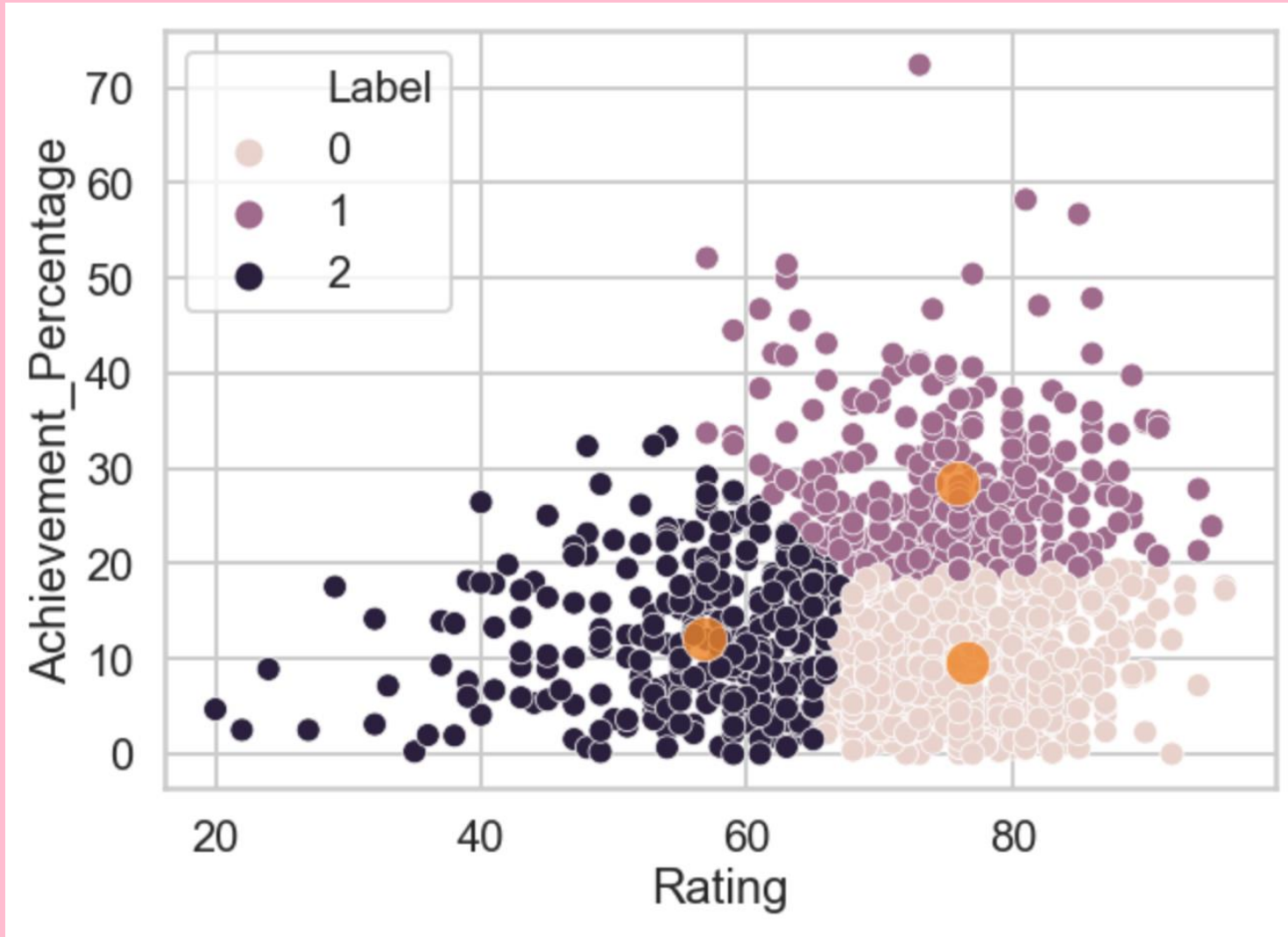
# CERRY ON TOP: TIMING MATTERS

- Legends get released right around Fall season!
- December isn't as good a launch time as we think!



Legends

# CHOOSING WHICH TO OPTIMIZE FOR



**1 = Legends:  
Aim for these!**

**2= Steady Sheep:  
Try not to invest too  
much resources in  
this category.**

**0 = Easy wins:  
Minimum  
investment;  
maximum returns.**



# LEGENDS

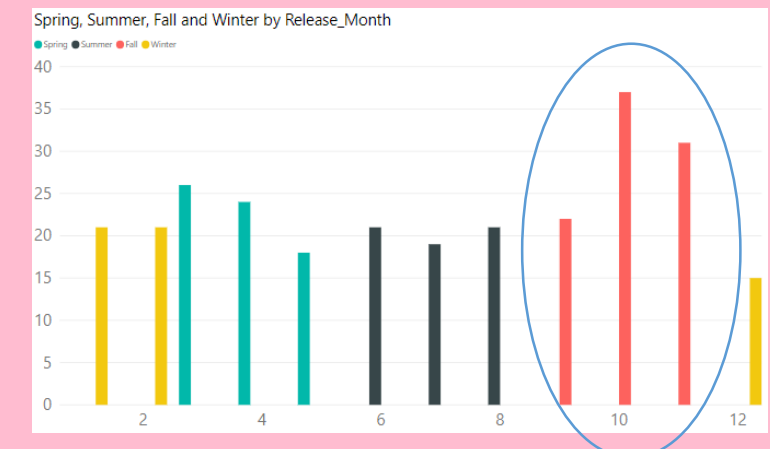
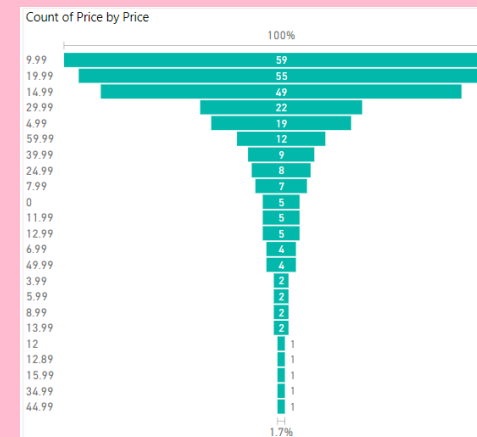
*If you want to optimize for quality, brand reputation, and huge release sales*

# EASY WINS

*If you want to optimize for a quick time-to-market, less production spend, and high user acquisition, in exchange of a longer sales cycle / slower ROI*

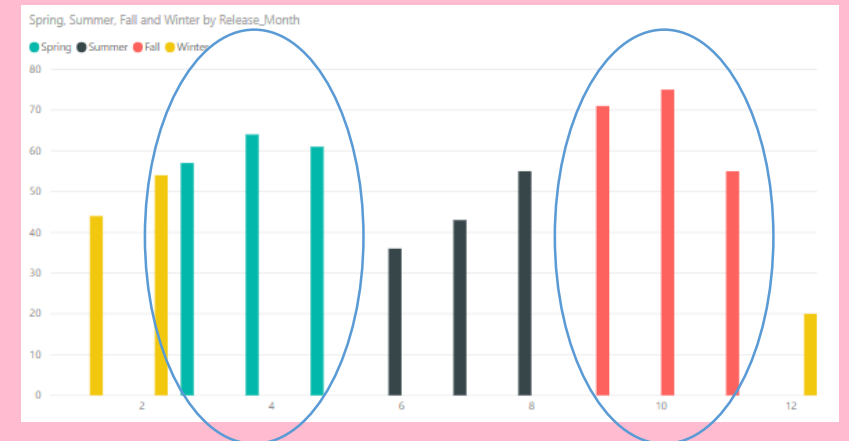
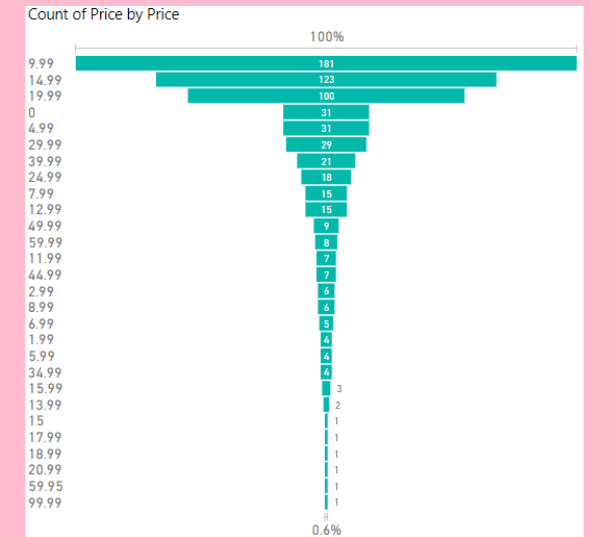
# SO YOU WANT TO CREATE A LEGEND?

1. Go for Action and Adventure games
2. No need to invest on multiplayer set-ups; focus on creating a great single player experience
3. You may charge high! Aim for 19.99 – 59.99 USD
4. Release around Fall



# SO YOU WANT TO CREATE AN EASY WIN?

1. Action, Strategy, and Adventure are your best bets
2. Single player will do okay, but you might want to consider Multiplayer to increase user volume
3. Price it reasonably (0 to 9.99) and focus on user volume, maybe consider charging for in-app purchases
4. Release around Spring and Fall



# YOU GOT THIS!

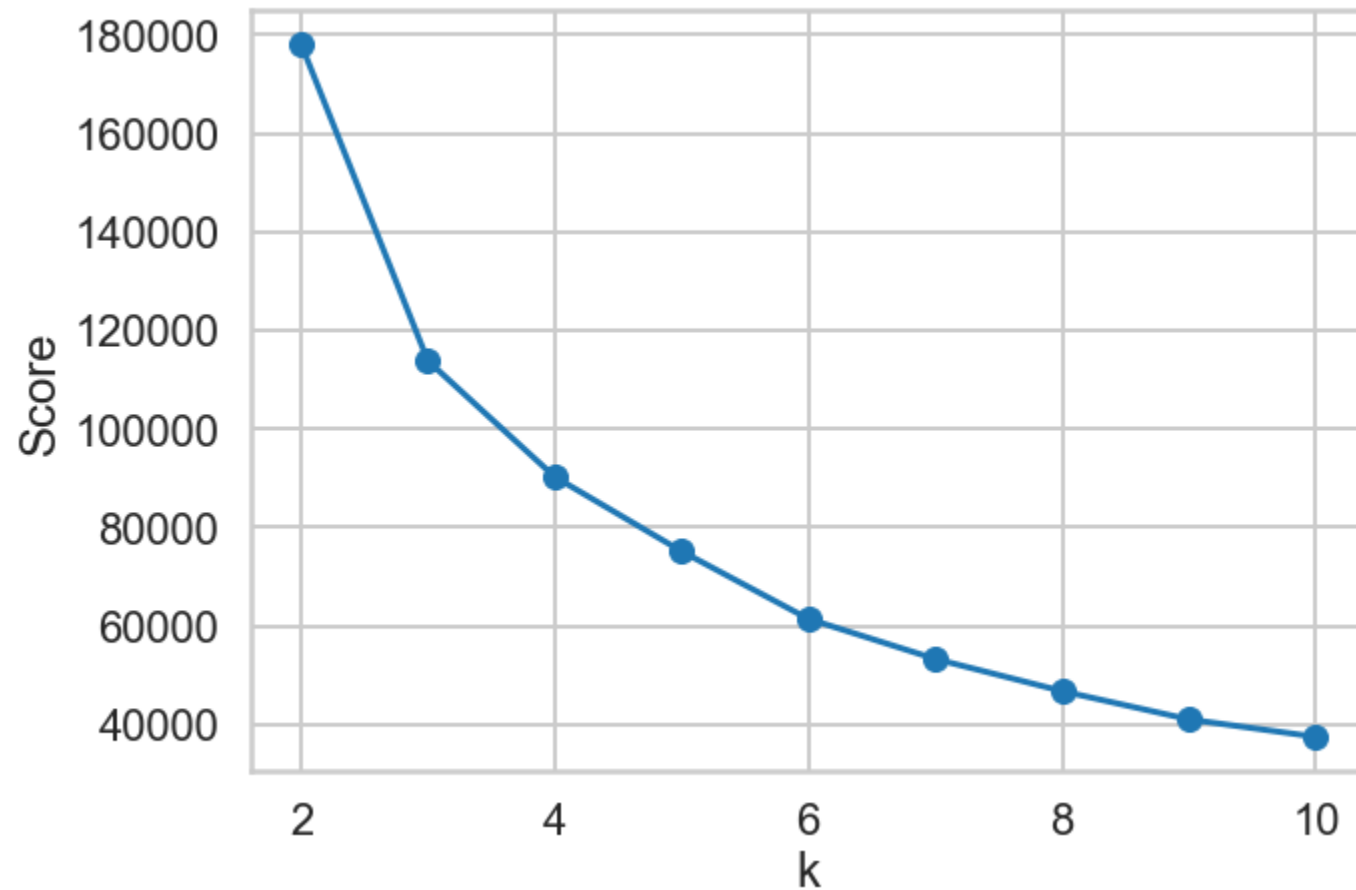
It's the best time for indie developers - armed with data, there's nothing you can't do!





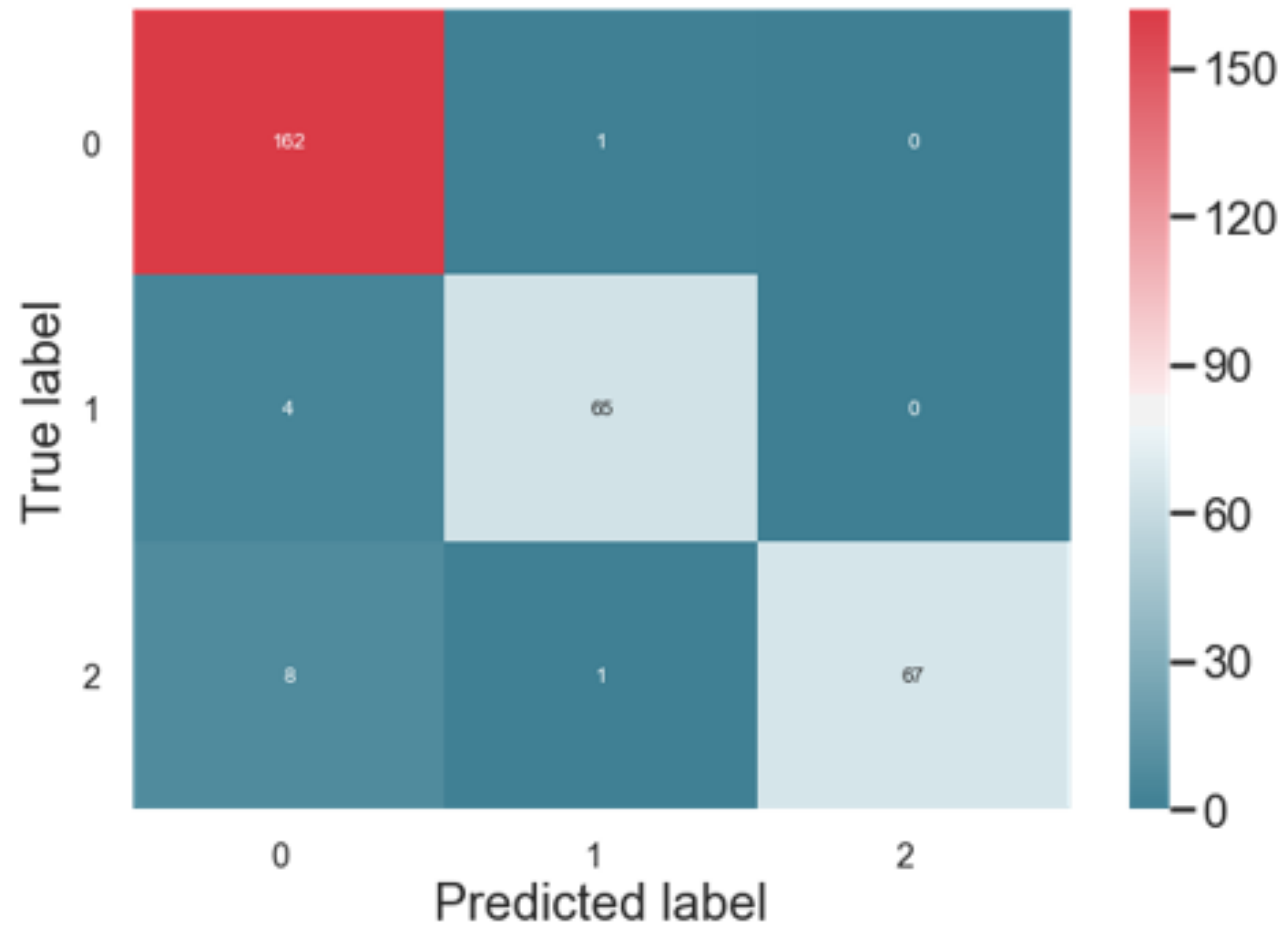


# KMeans



2	178097.87219222126
3	113793.45212042582
4	90229.58637406639
5	75119.61395990041
6	61386.04041812572
7	53252.581704052005
8	46715.885247064885
9	40975.7952072396
10	37433.55712766705

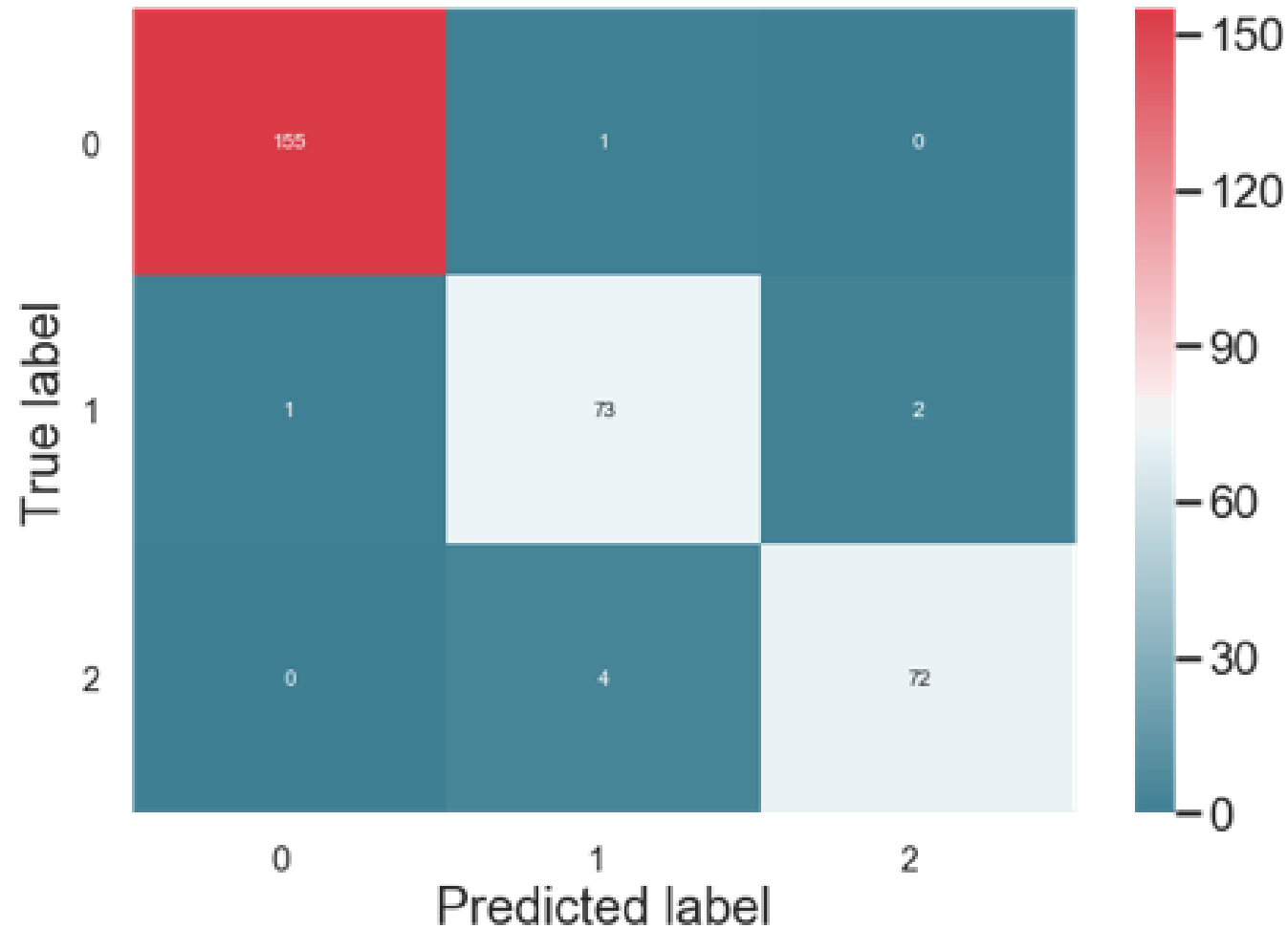
# Naive Bayes



	precision	recall	f1-score	support
0	0.93	0.99	0.96	163
1	0.97	0.94	0.96	69
2	1.00	0.88	0.94	76
micro avg	0.95	0.95	0.95	308
macro avg	0.97	0.94	0.95	308
weighted avg	0.96	0.95	0.95	308

('Accuracy score:', 0.9545454545454546)

# Decision Tree (GINI)

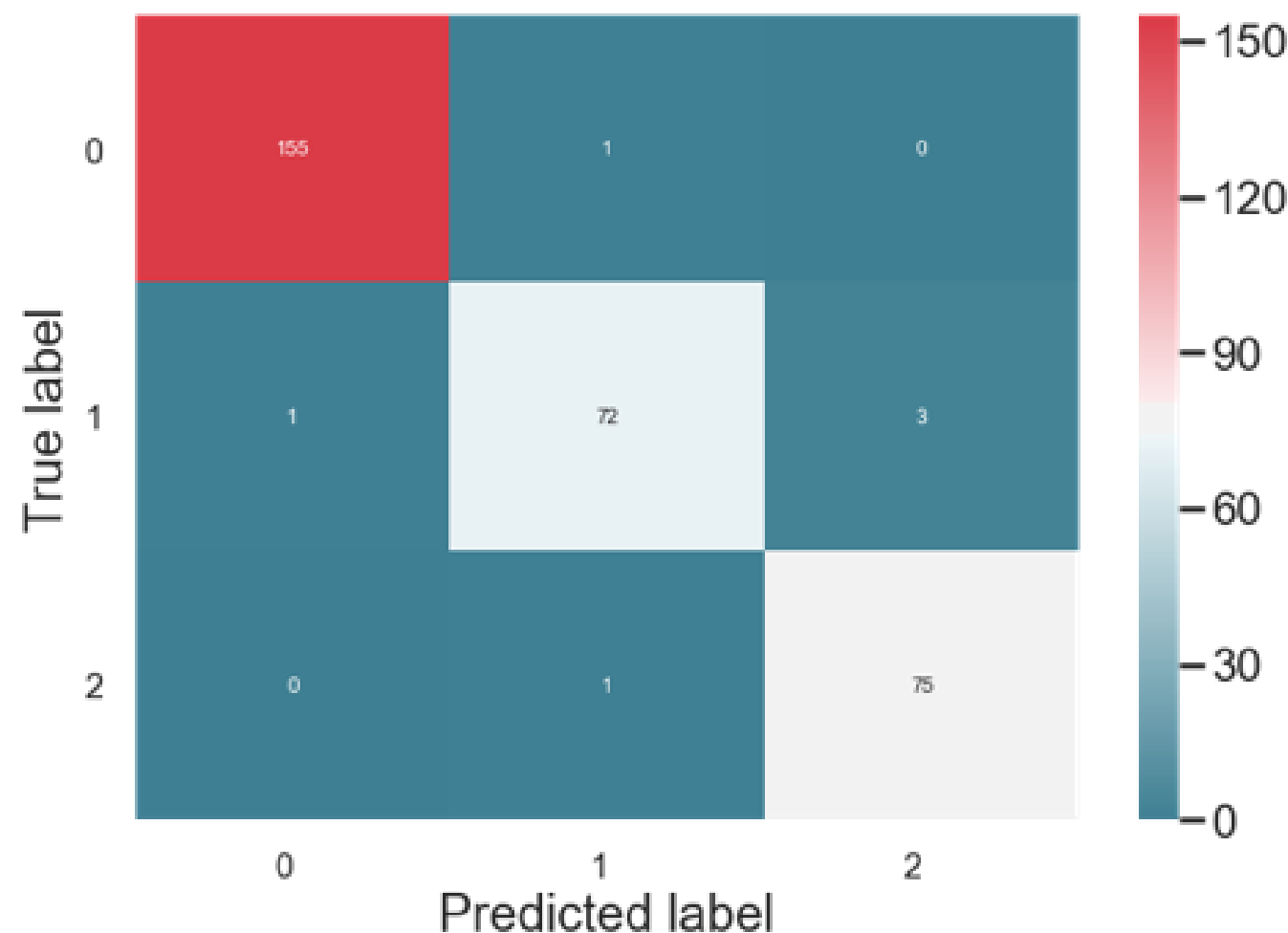


Accuracy score: 0.97

	precision	recall	f1-score	support
0	0.99	0.99	0.99	156
1	0.94	0.96	0.95	76
2	0.97	0.95	0.96	76
micro avg	0.97	0.97	0.97	308
macro avg	0.97	0.97	0.97	308
weighted avg	0.97	0.97	0.97	308



# DECISION TREE (ENTROPY)



Accuracy score: 0.98

	precision	recall	f1-score	support
0	0.99	0.99	0.99	156
1	0.97	0.95	0.96	76
2	0.96	0.99	0.97	76
micro avg	0.98	0.98	0.98	308
macro avg	0.98	0.98	0.98	308
weighted avg	0.98	0.98	0.98	308