# Video Game Recommender System

Jawad Toufaili
Sebastian Salazar
Shivangi Soni
Vivek Saahil

# About the Videogame Industry

- Development, marketing and monetization of videogames

- Console big three companies: Nintendo (Switch), Sony (PlayStation), and Microsoft (Xbox)

- Industry annual revenue (2018): US$ 134.9 billion

- Different developer formats: First, second-, and third-party developers

- Big push to sell games in digital format

- Indie developers (third-party) sales problem
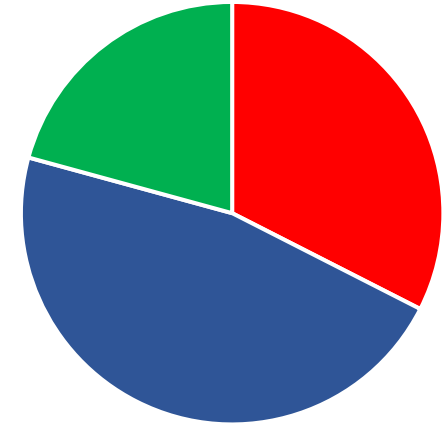


■ Nintendo   ■ Playstation   ■ Xbox

# Problem Statement: GameStop Corp.

## GameStop's Losing Big to Digital
GameStop is confronting increasing pressure from digital gaming companies.

**Now that Google and Nintendo offer digital video games, GameStop could have the same fate as Blockbuster**

GameStop: its all fun and games until someone gets disintermediated



NYSE: GME – Stock Price
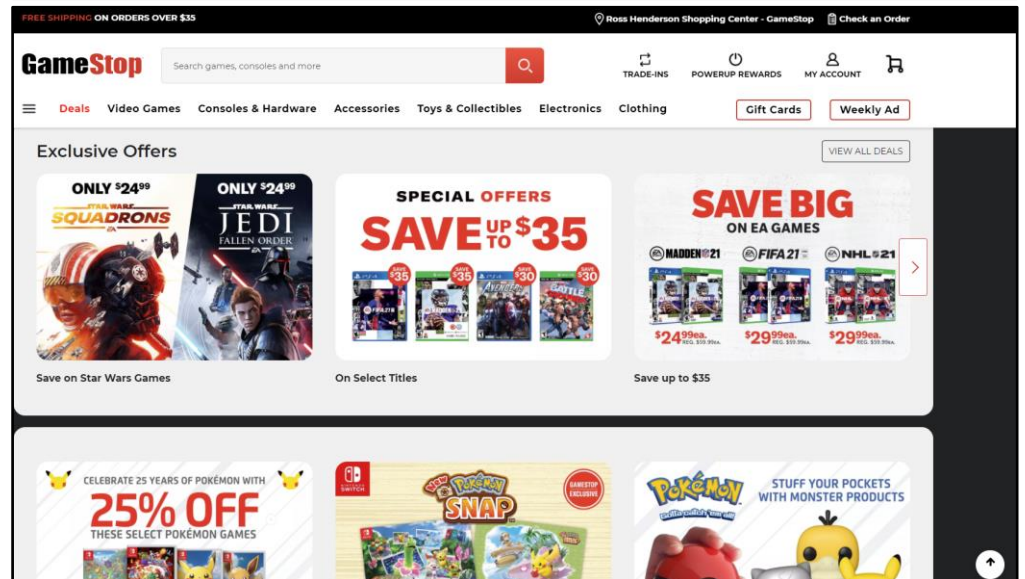
GameStop Appoints Chief Technology Officer

Announces Two Additional Executive Hires to Support Transformation

GameStop Announces Additional Board Refreshment to Accelerate Transformation

Aligns with Ryan Cohen of RC Ventures on the Immediate Appointment of Three New Directors with Significant E-Commerce and Technology Experience

Confirms the Board's Commitment to Supporting GameStop's Pursuit of Growth and Market Leadership

UPDATE – GameStop Invites Customers to Explore New Redesigned Website

# Analyzing Potential Opportunities

# Analyzing Potential Opportunities

```
                    OLS Regression Results
================================================================
Dep. Variable:        total_sales_USDMM   R-squared:              0.810
Model:                          OLS       Adj. R-squared:         0.788
Method:              Least Squares        F-statistic:            36.10
                                          Prob (F-statistic):     8.57e-29
                                          Log-Likelihood:         -643.53
No. Observations:               105       AIC:                    1311.
Df Residuals:                    93       BIC:                    1343.
Df Model:                        11
Covariance Type:          nonrobust
================================================================
                                       coef    std err      t      P>|t|
----------------------------------------------------------------
const                               52.0706    23.259    2.239    0.028
genre_racing                       257.9498    93.262    2.766    0.007
genre_shooter                      151.7964    64.627    2.349    0.021
genre_simulation                  -124.1876    60.091   -2.067    0.042
rating_E                            69.3619    31.850    2.178    0.032
rating_M                          -111.3395    40.928   -2.720    0.008
meta_overview_Generally favorable reviews  -219.5786  24.005  -9.147  0.000
meta_overview_Mixed or average reviews     -248.3889  27.331  -9.088  0.000
meta_overview_Universal acclaim             520.0381  39.984  13.006  0.000
user_overview_Generally favorable reviews   185.8357  35.362   5.255  0.000
user_overview_Generally unfavorable reviews 226.9570  65.181   3.482  0.001
user_overview_Mixed or average reviews      196.3866  37.782   5.198  0.000
user_overview_Universal acclaim            -557.1087 104.380  -5.337  0.000
Developer_type_First                        197.9016  35.616   5.557  0.000
================================================================
```

* Regression was run over the games that sales or shipping information could be extracted

## Insights

1. Generally unfavorable games can still have high sales. This can be an indicator that popular games might have sales for franchise influence or developer popularity

2. Universally acclaim games do not necessarily have larger sales, emphasizing the fact that game quality is not a defining feature of commercial success

3. Games developed by first party companies sell more, which may be related to marketing, popularity and customer top of mind

# Analyzing Potential Opportunities

- Low selling games include the same attributes on user reviews as high selling games. This shows that sell performance is not directly related with the videogame content, but other innate commercial attributes impact them



More Popular (Greater than total 25 favorable or mixed reviews)



Less Popular (Less than or equal to total of 25 favorable or mixed reviews)
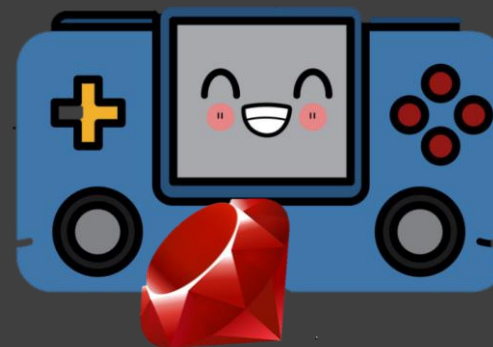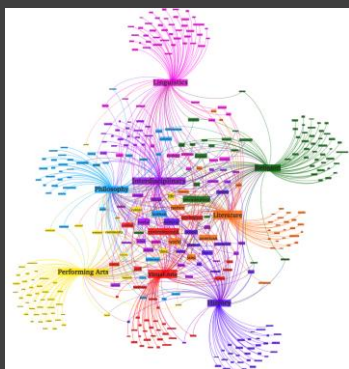
# Solution: Creating a Competitive Advantage

- Quality is not always correlated to high sales

- Marketing and popularity drive sales in the industry

- Indie developers look for more exposure for their games as great games can be easily missed

- **Opportunity: GameStop can become the go-to market to discover and acquire top-indie physical games.** (unexploited)

-  Small developers have less negotiation power:
  - Higher margin on sales for GameStop
  -  More competitive prices through discounts

- Potential exclusive benefits for shopping indies at GameStop

# User/Critic-Review driven Recommender System

Scrape videogame info and reviews* from Metacritic (critics and users)

Used TF-IDF to vectorize all text and NMF to reduce dimensionality

Determine recommendations using clustering (with cosine similarity)

Recommend games based on user and critic reviews and "hidden gem" games

Increase sales through best deals and brand popularity
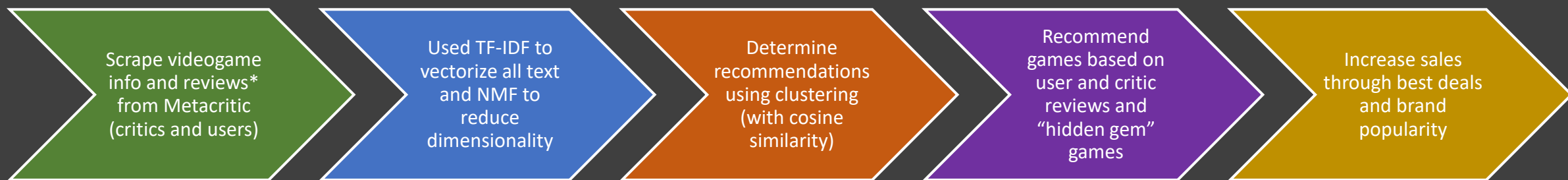
*The beta version of this solution will be done on Nintendo Switch as it is the console that pushes more indie sales*

# Building the Solution: Recommender Inner-Workings

## Data Pre-processing

- Lowercasing Reviews
- Removing Stop words, Punctuation and Game Titles
- Aggregating all reviews per game (User/Critic separately)
- Keeping Unique Game Titles
- Lemmatization
- Removing words with frequency < 50
- Removing non-essential words

## Defining Game Topics

- Topic Modelling
- TF/IDF Vectorizer – for vectorizing reviews
- Techniques used for dimension reduction – NMF, LSA, LDA
- Best topics obtained through NMF at n = 25 topics

## Identifying User's Matching Games

- Recommender system using K-means Clustering algorithm with cosine similarity on the vectorized comments

## Recommending Hidden Gems

- An additional layer of constraints, which only recommends games with less than 25 total number of user reviews (either favorable/mixed/average reviews)
- Less than 17 total number of critic reviews(either favorable/mixed/average reviews)

# GameStop Game Recommender Demo



```
IPython console
Console 1/A

****** Welcome to GameStop RecoBot ******

**** The Ultimate Game Recommender ****

*******************************************

Hello there! What's your favourite game?
Try 'explore', if not sure

User:
```

# Conclusions

- Recommendations varied greatly depending on user input
  - For example, cross genre recommendations
- Different recommendations generated from users' reviews vs. critics' reviews
  - Relatively more accurate recommendations generated from critics' review as compared to users' reviews
  - Potentially due to the terminology and jargon used in critics' reviews

**Future Improvements:**
- Including other features such as genre, ratings, sentiment scores of comments to get more accurate recommendations
- Incorporating other consoles to get more variety in dataset
- Building a recommender engine that puts more weight on the important attributes in the comments
- Finding more data overlay sources that could help with determining hidden gems
- Building an interactive recommendation engine using Dash