

PLAYER 1



HIGHSCORE 2500

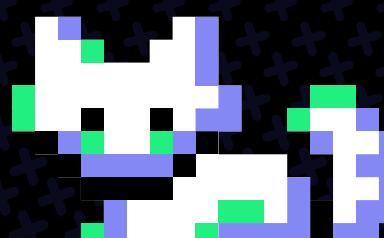


PLAYER 2



GAME RECOMMENDATIONS FROM MACHINE LEARNING

BY: NICOLE HAMMACK AND ADRIANA GALINDO





01



07



12



IN A  
DATABASE  
OF 18,906  
GAMES, 15%  
BELONG TO  
THE ROLE-  
PLAYING  
GENRE.

→ MORE PEOPLE REVIEWED ACTION GAMES BUT  
MORE NEGATIVELY

→ TOP GENRES BY MEDIAN POSITIVITY RATIO  
IS QUIZ/TRIVIA AND VISUAL NOVEL

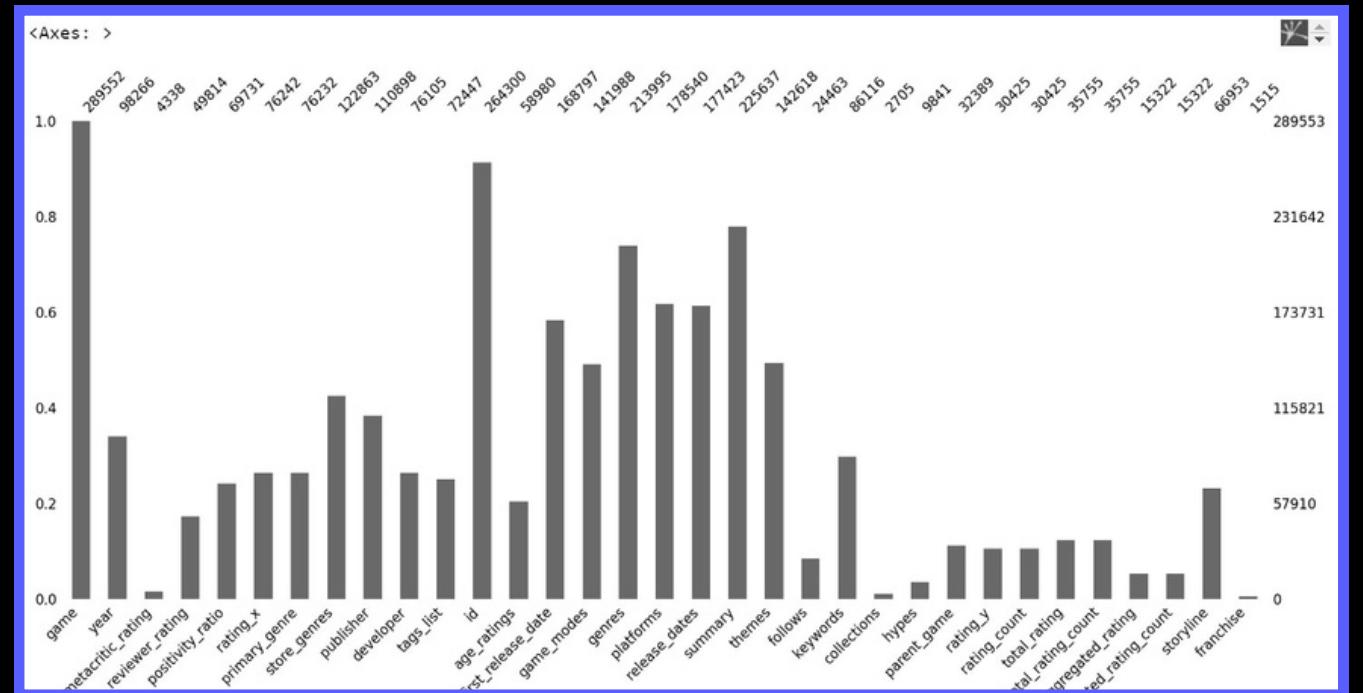
→ [TABLEAU PUBLIC DASHBOARD](#)





# STEPS

Msno.bar



One-hot encoded

```
# List of columns to be one-hot encoded
columns_to_encode = ['developer', 'game_modes', 'genres', 'themes']

# Use pd.get_dummies to perform one-hot encoding
df_encoded = pd.get_dummies(df_dropped, columns=columns_to_encode)

# Display the first few rows of the encoded DataFrame
df_encoded
```

Pulled data from  
Twitch API

Pulled data from  
Zagggle

Pulled data from  
Reddit/Github

Brought into jupyter notebook to clean

Dropped games with over 15 null values

```
#dropping rows that have over 15 columns with NaN values
df_filtered = dropped_df.dropna(thresh=dropped_df.shape[1] - 15)
df_filtered
```

Fuzzywuzzy similar names/typos  
within developers

```
developer_names = df_dropped['developer'].unique()

for i in range(len(developer_names)):
    for j in range(i + 1, len(developer_names)):
        similarity = fuzz.ratio(developer_names[i], developer_names[j])
        if similarity > 90: # Adjust the threshold as needed
            print(f'Potential typo or similar name: {developer_names[i]} vs {developer_names[j]}')
```

Potential typo or similar name: Ripknot Systems vs Ripknot System  
 Potential typo or similar name: Space Cat Studios vs Space Boat Studios  
 Potential typo or similar name: Rain Games vs Radin Games  
 Potential typo or similar name: Wanba Studio vs Wan Studio  
 Potential typo or similar name: Seen Games vs Sensen Games  
 Potential typo or similar name: Altar Games vs Altair Game  
 Potential typo or similar name: Altar Games vs Alter Games  
 Potential typo or similar name: Rava Games vs Ragiva Games  
 Potential typo or similar name: EightyEightGames vs EightyEight Games

Created new df with cleaned and encoded data

used standard scaler

determined optimal k = 8 via elbow method

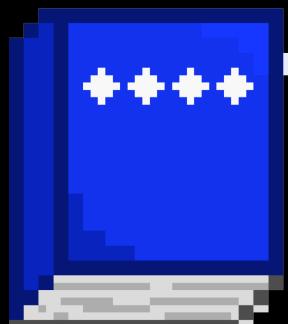
created function to find similar games within each cluster

used pickle to load/read the saved model into the flask app  
(predictor function uses the loaded model in flask)

used html and bootstrap to build website

used css for styling

user inputs game, rawg api pulled in output for game rec info



# STEPS

The screenshot shows the PIXIPAL website interface. At the top right is the logo "PIXIPAL" with the tagline "FIND YOUR NEXT ADVENTURE". Below the logo is a cartoon illustration of a young girl with brown hair and a headband, sitting at a desk with a computer monitor displaying a colorful map. On the desk are various gaming peripherals like a controller and a keyboard. The main area features a search bar with the placeholder "Game title" and two buttons: "Search" and "clear". To the right of the search bar are two game cards. The first card is for "Opus Magnum" (Year: 2017-10-19), available on Steam, GOG, and itch.io, with genres Indie, Simulation, and Puzzle. The second card is for "Mon Bazou download free[Preinstalled]" (Year: 2022-05-07), available on itch.io, with genres Simulation.

MENU

➡ 01

♦ 07

★ 12



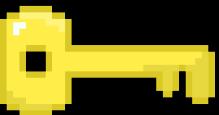
# WHAT WE WOULD HAVE DONE DIFFERENTLY



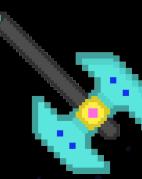
SUPERVISED  
MACHINE LEARNING/  
RANDOMFOREST



WEBSITE SHOW WHAT  
USER INPUT IS



ALL 250K+ GAMES  
HAVE COMPLETE  
DATA



PUBLISH WEBSITE  
AND GO LIVE

# RESOURCE PAGE



## TWITCH

[HTTPS://API-DOCS.IGDB.COM/#GETTING-STARTED](https://api-docs.igdb.com/#getting-started)

## ZAGGLE METACRITIC

[HTTPS://WWW.KAGGLE.COM/CODE/THOMASMANN11111/STEAM-GAMES-DATA-EXPLORATION-PREDICTION/INPUT](https://www.kaggle.com/code/thomasmann11111/steam-games-data-exploration-prediction/input)

## ZAGGLE STEAM

[HTTPS://WWW.KAGGLE.COM/CODE/HMYLEDA/STEAM-GAME-ANALYSIS/INPUT](https://www.kaggle.com/code/hmyleda/steam-game-analysis/input)

## GITHUB

[HTTPS://GITHUB.COM/THATDATASCIENCEGUY/THATDATASCIENCEGUYBLOG/TREE/MASTER/VIDEOGAMEDATA?SOURCE=POST\\_PAGE----D900C3470C79-----](https://github.com/thatdatascienceguy/thatdatascienceguyblog/tree/master/videogamedata?source=Post_Page----D900C3470C79-----)

## REDDIT

[HTTPS://WWW.REDDIT.COM/R/GAMEDEV/COMMENTS/XOQS4Z/WE\\_GATHERED\\_DATA\\_ABOUT\\_5400\\_0\\_GAMES\\_IN\\_STEAM\\_AND/?RDT=50469](https://www.reddit.com/r/gamedev/comments/xoqs4z/we_gathered_data_about_5400_0_games_in_steam_and/?rdt=50469)

# RESOURCE PAGE

Use these design resources  
in your Canva Presentation.  
Happy designing!

Don't forget to delete or hide  
this page before presenting.

