Just how good is The Last of Us Part 2?

An analysis of the critically acclaimed sequel

1 | Project Overview

1.1 | Introduction

The successor to one of the most critically acclaimed games of all time, The Last of Us, faced the weighty challenge of meeting exceptionally high expectations. As my all-time favorite game, The Last of Us Part 2 intrigued me deeply, prompting a question thats surely been in the minds of many: Just how good is it? Nearing four years since its release, my goal is to unravel the layers of its greatness.

I've been inspired to answer this question for a long time, and when it was revealed that the game had sold over 10 million copies in 2022 [1], I was sure I had the answer. But what does 10 million copies even mean? Are sales the only way to measure a games success? Can I show through other metrics, such as game reviews, game comparisons, and game profits that The Last of Us Part 2 is a great sequel?

I wanted to focus mainly on The Last of Us Part 2, and "successful" games while drawing as many comparisons as possible with other games that share similarities with TLoU2. Things like divisiveness, and the success of other sequels. I am aware that a user by the name of Lázaro ran a sentiment analysis on the game [2] and though I plan on using NLP to do sentiment analysis much like he did, I aim to explore the question at a greater length, beyond that of user/critic reviews. There are currently no known notebooks or data analysis on the success of The Last of Us Part 2.

1.2 | Project background

Soon after its release, The Last of Us Part 2 faced a lot of criticism, especially with quite a number of negative reviews. Some folks who didn't like the game went to extremes, demanding refunds and even sending death threats and offensive slurs to the developers and actors [3]. I remember at that time, users were intentionally using bots or just leaving a 0/10 rating with no 'review' to bring down the game's rating. This fuss continued at the end of 2020 during the Game Awards when the public was clearly upset that The Last of Us Part 2 snagged 'Game of the Year'.

Even with all the negative reviews, there were still lots of positive ones. Back in late June 2020, if you checked Metacritic, it was almost a 50-50 split between 0/10 and 10/10 reviews. The game was divisive, with some singing its praises and others swearing off Naughty Dog games altogether. At that time, it felt like a close call, but over time, I've wondered how many people really liked the game. Crossing the 10 million units sold mark is a big deal for AAA games, and TLoU2 notched up a ton of awards, making it one of the most awarded games ever.

So, does the game truly deserve all the praise, or is the hate it got justified?

1.3 | Aims and Objectives

With this project, I would like to do the following:

Aims

1. Assess the overall reception and impact of The Last of Us Part 2 in the gaming community.
2. Examine the factors contributing to the game's divisive nature, considering both positive and negative aspects.
3. Investigate the validity and significance of sales figures (10 million copies sold) as a measure of the game's success.
4. Analyze the influence of user reviews, both positive and negative, on the perception and reputation of the game.
5. Explore the extent to which awards and accolades contribute to the assessment of a game's quality and success.

Objectives

1. Conduct a comprehensive review analysis, examining sentiments and key themes in both positive and negative reviews.
2. Utilize Natural Language Processing (NLP) to perform sentiment analysis, extending beyond user and critic reviews, with a focus on in-depth exploration.
3. Compare The Last of Us Part 2 with similar games in terms of sales, critical reception, and player engagement.
4. Evaluate the impact of controversies surrounding the game, including the backlash and negative sentiments, on its long-term success.
5. Scrutinize the correlation between sales figures and game quality, considering the potential influence of marketing and public perception.
6. Investigate the role of awards in shaping the narrative around a game, examining how accolades contribute to its overall reception.

1.4 | Data

1.4.1 Data requirements

In the pursuit of a comprehensive analysis of The Last of Us Part 2, a deliberate selection of diverse datasets was made, aiming to encapsulate various dimensions that contribute to defining the game's success or failure. This inclusive approach spans both qualitative and quantitative datasets, encompassing aspects such as user reviews and sales figures.

To ensure a thorough exploration of the game's performance, it is imperative that the data encompasses a wide spectrum of the video game industry. This involves incorporating datasets that provide insights into different facets of gaming, allowing for a nuanced understanding of the factors influencing The Last of Us Part 2's standing within the industry.

Given the exclusivity of The Last of Us Part 2 to the PlayStation platform, a judicious consideration of this aspect is vital. While acknowledging its exclusivity, the analysis should not be confined solely to comparisons with other PlayStation-exclusive titles. Instead, the scope should extend to include a diverse range of games, both exclusive and non-exclusive to PlayStation, offering a comprehensive view of the game's performance against a broader industry backdrop.

The data requirements for this analysis involve a rich and varied collection of datasets, encompassing qualitative and quantitative metrics, and spanning the breadth of the video game industry. The exclusivity of The Last of Us Part 2 should be acknowledged, but the comparisons should transcend platform exclusivity to capture a holistic perspective on the game's success or failure.

1.4.2 Choice of data

The dataset compiled for this analysis comprises information that is readily available in the public domain, having been previously gathered for various notebooks and data science projects. It was a deliberate choice to utilize data that is widely used by others in their analyses of video games, ensuring a level of comparability and relevance within the broader context of gaming analytics.

A key objective in selecting the data was to strike a balance between opinion-based insights and factual information. Recognizing that much of the discourse surrounding the game revolves around subjective viewpoints, the dataset aims to capture both the qualitative aspects of discussions and the objective facts. This approach is essential to discern whether the game's perceived success aligns with the empirical evidence.

To maintain impartiality and objectivity in the analysis, a conscious effort was made to source data without inherent bias toward any particular stance in the argument. The goal

was to foster a comprehensive understanding of the game's performance by integrating diverse perspectives and verifiable facts, ensuring a robust and nuanced evaluation.

1.4.3 Limitations of data

Obtaining up-to-date information on video game sales and metrics presents challenges due to the selective nature of data disclosure by game companies. Typically, companies choose when and which data to reveal, often showcasing information that portrays the game in a favorable light. Sales figures deemed successful are often presented within broad brackets, such as "sold over", without specifying precise numbers. Additionally, the analysis focuses exclusively on reviews posted to Metacritic, excluding other review sites.

While this exploration falls under the realm of data science, it's crucial to acknowledge that comparing exact figures may be insufficient. The primary objective of utilizing data in this context is to identify trends that could provide insights into the factors contributing to success or failure. Rather than emphasizing precise numbers, the project aims to position The Last of Us Part 2 on a scale relative to other successful games, gauging its standing within the broader landscape of gaming achievements.

1.4.4 Ethical considerations

In the process of conducting this analysis, it is crucial to address the ethical implications of utilizing scraped data from Metacritic. It should be noted that Metacritic explicitly states in its terms of service that scraping or automated access to its content is prohibited. Despite this policy, the decision was made to use scraped data for the purposes of this analysis.

The ethical considerations stem from the potential violation of Metacritic's terms of service, raising concerns about data ownership and the right to control access to their platform. While efforts were made to respect ethical standards throughout the analysis, it is essential to acknowledge the inherent tension between data accessibility and the terms set by content providers.

Researchers and analysts should exercise caution when considering the use of scraped data, always prioritizing compliance with platform policies and terms of service. The decision to use scraped data from Metacritic was made with the understanding that it may pose ethical concerns, and users are encouraged to evaluate the ethical implications in the broader context of data usage and intellectual property rights.

It is important to emphasize that I do not endorse the practice of actively scraping Metacritic, and I have personally refrained from engaging in such activities. The inclusion of a dataset containing reviews serves as a smaller component within a broader project concept. I am open to promptly excluding this dataset from my research should any user concerns arise due to my use of a scapped dataset.

All presented data was sourced without infringing upon any intellectual properties or encountering paywalls. Regarding user reviews, I have intentionally omitted user names and

the reviews' content to safeguard their identity. While the data utilized in this project includes only chosen usernames and lacks personal information, it poses no risk of leading to personally identifiable distinctions or causing harm to anyone.

2 | Methodology and Analysis

2.1 | Sentiment Analysis on Reviews

Importing modules

Below I import and download all the relevant modules needed to do my analysis. I use textblob for polarity and subjectivity scores.

In [2]:
```
import pandas as pd
import matplotlib.pyplot as plt
import string
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize, sent_tokenize
from nltk.sentiment import SentimentIntensityAnalyzer

%pip install textblob
from textblob import TextBlob
%pip install wordcloud
from wordcloud import WordCloud

nltk.download('stopwords')
nltk.download('punkt')
nltk.download('vader_lexicon')
```

```
Requirement already satisfied: textblob in /Users/kyle/anaconda3/lib/python3.1
1/site-packages (0.17.1)
Requirement already satisfied: nltk>=3.1 in /Users/kyle/anaconda3/lib/python3.
11/site-packages (from textblob) (3.8.1)
Requirement already satisfied: click in /Users/kyle/anaconda3/lib/python3.11/s
ite-packages (from nltk>=3.1->textblob) (8.0.4)
Requirement already satisfied: joblib in /Users/kyle/anaconda3/lib/python3.11/
site-packages (from nltk>=3.1->textblob) (1.2.0)
Requirement already satisfied: regex>=2021.8.3 in /Users/kyle/anaconda3/lib/py
thon3.11/site-packages (from nltk>=3.1->textblob) (2022.7.9)
Requirement already satisfied: tqdm in /Users/kyle/anaconda3/lib/python3.11/si
te-packages (from nltk>=3.1->textblob) (4.65.0)
Note: you may need to restart the kernel to use updated packages.
Requirement already satisfied: wordcloud in /Users/kyle/anaconda3/lib/python3.
11/site-packages (1.9.3)
Requirement already satisfied: numpy>=1.6.1 in /Users/kyle/anaconda3/lib/pytho
n3.11/site-packages (from wordcloud) (1.24.3)
Requirement already satisfied: pillow in /Users/kyle/anaconda3/lib/python3.11/
site-packages (from wordcloud) (10.0.1)
Requirement already satisfied: matplotlib in /Users/kyle/anaconda3/lib/python
3.11/site-packages (from wordcloud) (3.7.2)
Requirement already satisfied: contourpy>=1.0.1 in /Users/kyle/anaconda3/lib/p
ython3.11/site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cycler>=0.10 in /Users/kyle/anaconda3/lib/pytho
n3.11/site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in /Users/kyle/anaconda3/lib/
python3.11/site-packages (from matplotlib->wordcloud) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /Users/kyle/anaconda3/lib/
python3.11/site-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in /Users/kyle/anaconda3/lib/py
thon3.11/site-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in /Users/kyle/anaconda3/
lib/python3.11/site-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in /Users/kyle/anaconda3/l
ib/python3.11/site-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: six>=1.5 in /Users/kyle/anaconda3/lib/python3.1
1/site-packages (from python-dateutil>=2.7->matplotlib->wordcloud) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
zsh:1: command not found: pip
zsh:1: command not found: pyppeteer-install
[nltk_data] Downloading package stopwords to /Users/kyle/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /Users/kyle/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     /Users/kyle/nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
```

Out[2]:
```
True
```

Helper functions

Below I define functions I need to process and handle the data.

In the data cleaning process, I identified several unnecessary columns that wouldn't contribute to my sentiment analysis. Focusing specifically on the reviews, I excluded

columns such as the number of views, which wouldn't enhance my exploration. I also ensured that all reviews were in English and excluded entries without a review.

For tokenization, I eliminated punctuation and numbers, as sentiment analysis doesn't rely on them. Additionally, I removed stop words from the reviews to analyze the most impactful words in determining sentiment.

In [158…

```
def getReviewsDataframe(filepath):
    dfStart = pd.read_csv(filepath)

    #Only keep the records that were written in English
    dfEnglish = dfStart[dfStart['language'] == 'English']
    dfEnglish.reset_index(drop=True, inplace=True)

    #Remove unnessary columns, keep data that'll be used in analysis
    dfSimplified = dfEnglish.drop(['type_review', 'views', 'votes', 'language']

    # Remove any null values in the 'review' column if they exist
    dfSimplified = dfSimplified.dropna(subset=['review'])

    return dfSimplified

def getAllScores():
    scores = df['score'].value_counts()
    scoresDf = pd.DataFrame({'count': scores.values}, index=scores.index)
    scoresDf = scoresDf.sort_index()

    return scoresDf

def tokenizeReview(text):
    tokens = word_tokenize(text)

    #Convert the token to all lowercase
    lowercaseTokens = [word.lower() for word in tokens]

    #Remove punctuation and any numbers
    cleanedTokens = [t for t in lowercaseTokens if t.isalpha()]

    #Remove stop words
    stopWords = stopwords.words('english')
    stopwordsRemovedTokens = [word for word in cleanedTokens if word.lower() no

    return stopwordsRemovedTokens

def getSentimentAnalysis():
    tokenizedReviews = []
    sia = SentimentIntensityAnalyzer()

    for index, row in df.iterrows():
        tokenizedReviews.append(tokenizeReview(row['review']))

    compoundScores = [sia.polarity_scores(' '.join(tokens))['compound'] for tok

    return compoundScores

def getPolarity():
    tokenizedReviews = []
    polarityScores = []
```

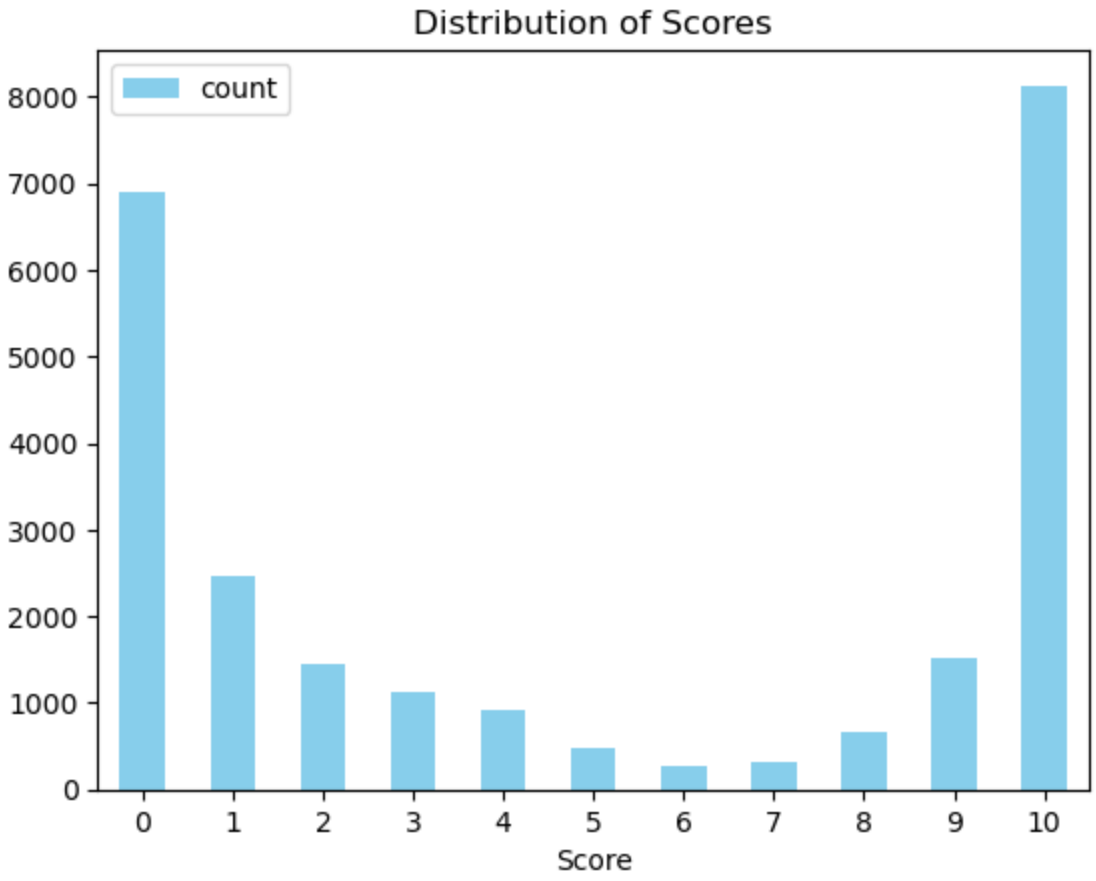
```
    for index, row in df.iterrows():
        tokenizedReviews.append(tokenizeReview(row['review']))

        blob = TextBlob(row['review'])
        polarityScores.append(blob.sentiment.polarity)

    return polarityScores

def getSubjectivity():
    tokenizedReviews = []
    subjectivityScores = []

    for index, row in df.iterrows():
        tokenizedReviews.append(tokenizeReview(row['review']))

        blob = TextBlob(row['review'])
        subjectivityScores.append(blob.sentiment.subjectivity)

    return subjectivityScores
```

Extracting the general distribution of scores

Displayed below is a bar graph illustrating the distribution of user scores, ranging from 0 to 10, where 0 indicates an extremely negative rating and 10 represents an exceptionally positive one.

In [160…

```
#Taken from Lázaro's own study
sentimentCSVFilePath = 'tlou2_reviews.csv'

#Build the dataframe
df = getReviewsDataframe(sentimentCSVFilePath)

#Get the scores
scores = getAllScores()

#Plot a graph of all scores
scores.plot(kind='bar', color='skyblue', rot=0)
plt.xlabel('Score')
plt.title('Distribution of Scores')
plt.show()
```



Superficially, this highlights the divisive nature of the game, with pronounced opinions prevalent in the majority of reviews.

Percentage of extreme scores

The following code calculates the total percentage of 0 and 10 scores

In [13]:
```
#Calculate the percentage of 0 and 10 scores
totalScoreCount = scores['count'].sum()

percentageDivisiveScores = (scores.loc[0, 'count'] + scores.loc[10, 'count']) /

print(f"Reviews with scores of 0 and 10 take {percentageDivisiveScores:.2f}% o
```

```
Reviews with scores of 0 and 10 take 62.05% of the total scores.
```

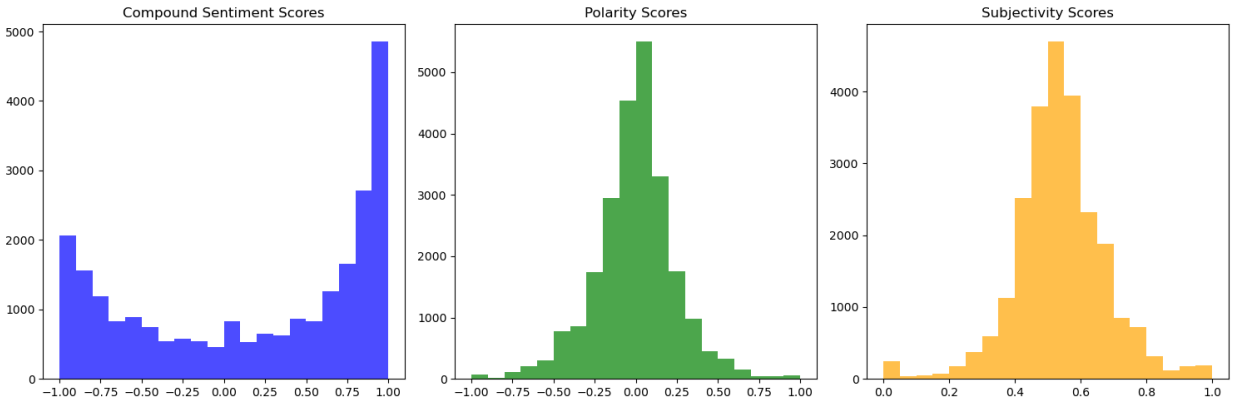
With a significant majority of 62% falling into the extremes of scores 0 or 10, there is a clear tendency toward polarized opinions. On average, users seem to either regard it as the best game ever or the worst game ever.

Extracting sentiment, polarity, and subjectivity

Presented below are three graphs providing insights into the overall sentiment, polarity, and subjectivity of all English reviews.

In [187…

```
compoundScores = getSentimentAnalysis()
polarityScores = getPolarity()
subjectivityScores = getSubjectivity()

fig, axs = plt.subplots(1, 3, figsize=(15, 5))

# Plot histograms
axs[0].hist(compoundScores, bins=20, color='blue', alpha=0.7)
axs[0].set_title('Compound Sentiment Scores')

axs[1].hist(polarityScores, bins=20, color='green', alpha=0.7)
axs[1].set_title('Polarity Scores')

axs[2].hist(subjectivityScores, bins=20, color='orange', alpha=0.7)
axs[2].set_title('Subjectivity Scores')

# Adjust layout
plt.tight_layout()
plt.show()
```



The sentiment analyses reveal intriguing patterns. The overall compound sentiment tends to be more positive than negative, supported by a slight rightward tendency in the polarity scores. Given that the analyzed content consists of personal reviews, the high subjectivity scores suggest a skew toward opinions rather than objective information.

However, upon closer inspection of reviews with somewhat positive sentiment scores, a noteworthy trend emerges. Many individuals expressing negative sentiments adopt a polite and euphemistic language, leading to an inflation of their sentiment scores towards the positive spectrum. Consequently, while these charts offer broad insights into the data, a granular analysis may be impacted by the inherent nature of sentiment analysis, where scores can be influenced by the nuanced expression of opinions.

With that being said, it appears that a larger number of users expressed favorable opinions about the game compared to those with negative sentiments. Additionally, the overall sentiment leans more toward positivity than negativity.

Finding what users liked and disliked

But what exactly did people like or dislike about the game?

I'm intrigued by the specific reasons behind both the admiration and criticism of the game. To delve into the sentiments of those who enjoyed it and those who loathed it, I'll create separate word clouds for each category to identify common themes.

Below, I've crafted two functions to preprocess strings for use in a word cloud—one designed for negative reviews and the other for positive reviews.

In [15]:
```
def getNegativeStringOfReviews():
    #Only take the reviews that scored lower than 5
    filteredDf = df[df['score'] < 5]

    combinedReviews = ''
    for index, row in filteredDf.iterrows():
        if not pd.isnull(row['review']):
            combinedReviews += row['review']

    return combinedReviews

def getPositiveStringOfReviews():
    #Only take the reviews that scored higher than 5
    filteredDf = df[df['score'] > 5]

    combinedReviews = ''
    for index, row in filteredDf.iterrows():
        if not pd.isnull(row['review']):
            combinedReviews += row['review']

    return combinedReviews
```

Below I generate two word clouds for both positive and negative reviews

In [205…
```
negativeWordcloud = WordCloud(width=400, height=400, background_color='white')
positiveWordcloud = WordCloud(width=400, height=400, background_color='white')

# Plot the word clouds side by side
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.imshow(negativeWordcloud, interpolation='bilinear')
plt.axis("off")
plt.title('Negative reviews')

plt.subplot(1, 2, 2)
plt.imshow(positiveWordcloud, interpolation='bilinear')
plt.axis("off")
plt.title('Positive reviews')

plt.show()
```

The word cloud analysis reveals some intriguing patterns. Notably, terms such as "character," "game," "story," and "Ellie" are prevalent in both positive and negative reviews. This suggests a divisive narrative, where users express both admiration and disdain for the game based on similar aspects.

A striking observation is the increased prominence of the name "Joel" in negative reviews compared to positive ones. This aligns with the context that Joel was a beloved character in the first game, "The Last of Us Part 1." In the sequel, "The Last of Us Part 2," the character Abby intentionally kills Joel in a graphic and brutal way.

It's noteworthy that Abby's name appears more frequently in positive reviews. This may indicate a nuanced perspective among users who appreciated Abby's role or story arc. Interestingly, the word cloud suggests that users who disliked the game often attribute their dissatisfaction to Joel's death rather than specifically Abby being the perpetrator. Adding context, the game's narrative centers around Joel's death, serving as a pivotal event shaping around 80% of the storyline.

It's important to highlight that Lázaro conducted a comprehensive analysis on The Last of Us Part 2 reviews, offering a more intricate examination. If you find my analysis intriguing, I recommend exploring Lázaro's work, which delves into a broader spectrum of user sentiments and provides in-depth data insights. My analysis, while more straightforward, aligns closely with Lázaro's conclusions. For a comprehensive understanding of user perspectives on the game and detailed data exploration, referring to Lázaro's notebook is highly recommended.

2.2 | Analysis on Sales Data

In framing this analysis, it's worth noting that obtaining precise sales figures for video games can be challenging, as companies often release data selectively, emphasizing success. Acquiring up-to-date data is also a hurdle, and many datasets lack sales figures for The Last of Us Part 2. With that being said what we are trying to analyze here is what roughly shows a successful game in terms of sales, and more importantly what do sales figures look like for games (and sequels) that are widely considered "successes". The data should reveal trends associated with success. My primary focus is on identifying patterns rather than fixating on precise numbers, which are subject to constant change.

Helper functions

Here are the functions utilized to preprocess the data for this analysis.

To streamline the dataset and focus on the key information, unnecessary columns from the CSV file were eliminated. The refinement process involved retaining only the data relevant to game sales figures, specifically those games that achieved a sales threshold of at least 1 million copies. This targeted approach aims to declutter the data, enabling a more precise exploration of the desired insights.

In [155…

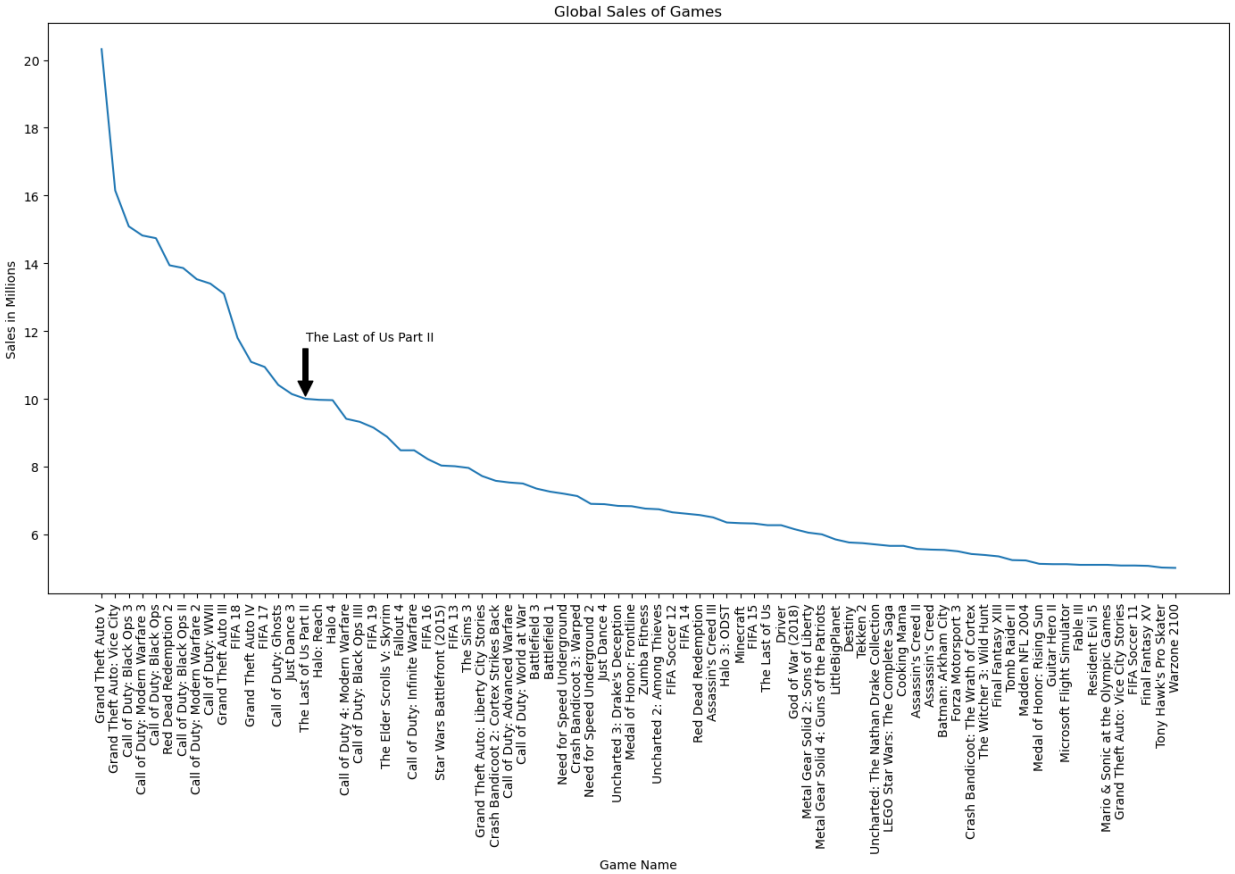
```
def getGameSalesDataframe(filepath):
    dfStart = pd.read_csv(filepath)

    #Set rank to be the index, as it is in the csv
    dfStart.set_index('Rank', inplace=True)

    #Remove unnessary columns, keep data that'll be used in analysis
    dfSimplified = dfStart.drop(['basename', 'Genre', 'ESRB_Rating', 'Publishe
                                 'VGChartz_Score', 'Year', 'Last_Update', 'url
                                 'Vgchartzscore', 'img_url'], axis=1)

    #Keep games that have any value in the any of 'sales' columns
    salesColumns = ['Global_Sales', 'NA_Sales', 'PAL_Sales', 'JP_Sales', 'Othe
    dfFiltered = dfSimplified.dropna(subset=salesColumns, how="all")

    #Only include games that have sold at least 1 million copies
    dfFiltered = dfFiltered[dfFiltered[salesColumns].apply(lambda x: (x > 1.0)

    #Fix the 'rank' column to start from 0, and be an index instead
    dfFiltered.reset_index(drop=True, inplace=True)

    return dfFiltered

def getGameSalesAbove(dataframe, value):
    df = dataframe
    df = df[df['Global_Sales'] > value]
    return df

def addTLOU2ToDataframe(dataframe):
    record = {'Name': 'The Last of Us Part II', 'Platform': 'PS4', 'Critic_Sco

    dataframe.loc[len(dataframe)] = record
```

Place 'The Last of Us Part 2' amongst other game sales

Below I create a dataframe to plot a graph of where the game stands amongst other games

In [161…

```
gameSalesCSVFilePath = 'game_sales_data.csv'

gameSalesDf = getGameSalesDataframe(gameSalesCSVFilePath)
gameSalesDf = getGameSalesAbove(gameSalesDf, 5.0)
addTLOU2ToDataframe(gameSalesDf)

#Remove duplicate games, and sort in ascending order
gameSalesDf['Global_Sales'] = pd.to_numeric(gameSalesDf['Global_Sales'], errors
gameSalesDf = gameSalesDf.loc[gameSalesDf.groupby('Name')['Global_Sales'].idxma
gameSalesDf = gameSalesDf.sort_values(by='Global_Sales', ascending=False)

#Plot the graph
plt.figure(figsize=(14, 10))
plt.plot(gameSalesDf['Name'], gameSalesDf['Global_Sales'], label='Global Sales

highlightIndex = gameSalesDf.index[gameSalesDf['Name'] == 'The Last of Us Part
plt.annotate("The Last of Us Part II",
             xy=(15, gameSalesDf.loc[highlightIndex, 'Global_Sales']),
             xytext=(15, 12),
             arrowprops=dict(facecolor='black', shrink=0.05),
             horizontalalignment='left',
             verticalalignment='top')

plt.title('Global Sales of Games')
plt.xlabel('Game Name')
plt.ylabel('Sales in Millions')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()
```



This graph specifically illustrates the sales data for The Last of Us on the PS3, indicating approximately 6.27 million units sold. Despite the data reflecting The Last of Us sales on PS3, it's crucial to consider the broader context. For instance, The Last of Us, including its remaster for PS4, has reportedly sold over 17 million units according to Statista [1]. This highlights the limitation of the graph in representing the comprehensive sales performance of the game across all platforms.

Comparatively, the highly successful Red Dead Redemption 2, a 'Game of the Year' title and a single-platform sequel, achieved sales of 13.94 million. This success positions The Last of Us Part II closely behind Red Dead Redemption 2 in terms of sales figures.

In a historical perspective, the first Red Dead Redemption, exclusive to PS3, sold 6.57 million copies. The sequel, Red Dead Redemption 2, surpassed this figure by selling an additional 7.37 million units. While precise figures for The Last of Us Part II are not available, it has surpassed The Last of Us (on PS3 only) by more than 3.8 million units. This achievement not only signifies growth but also positions The Last of Us Part II as a success, with sales comparable to over half of what Red Dead Redemption 2 accomplished. Dismissing the success of TLoU2 based on these comparisons would be unwarranted.

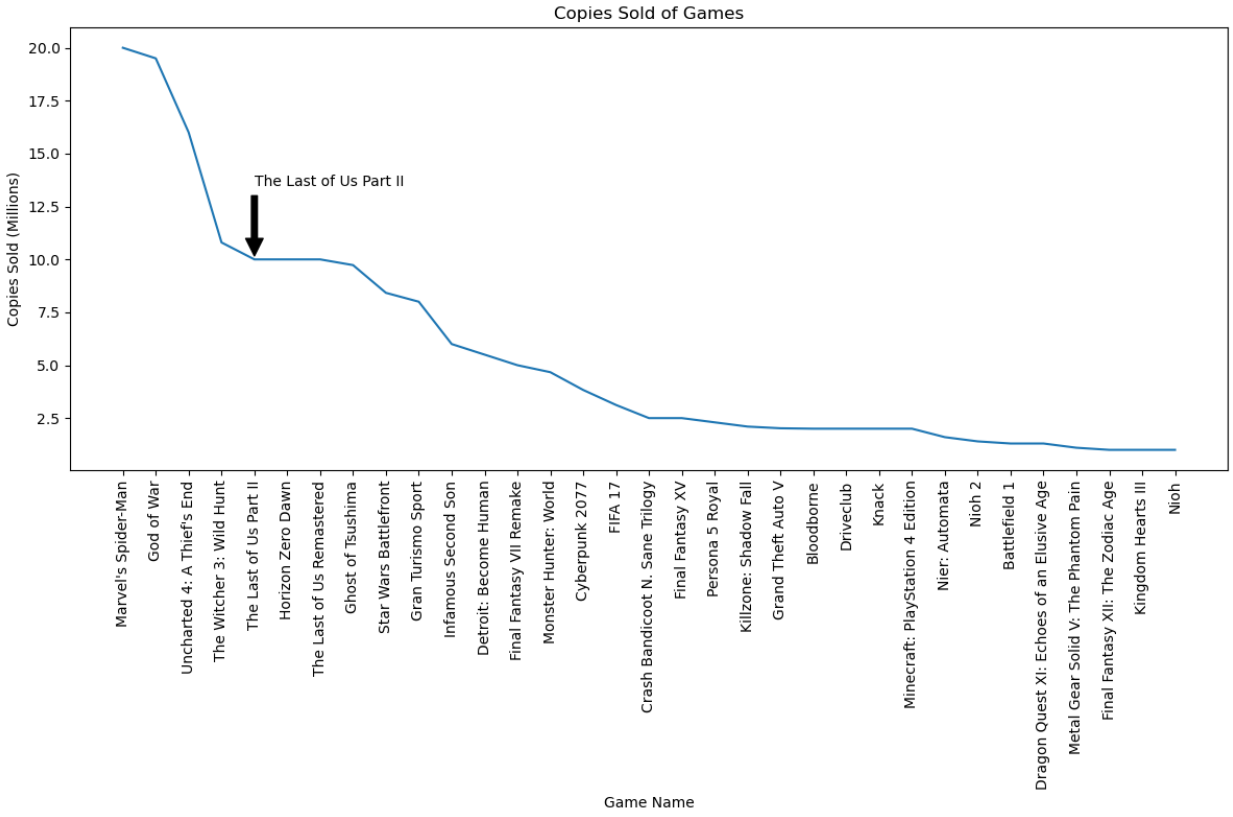
What is note worthy is that The Last of Us Part 2 hasn't reached the same amount of sales as Red Dead Redemption 2 despite them being of very similar in nature. It's definitely a successful game, but basing the data off the game Red Dead Redemption 2... it should've done better, a lot better.

Only PlayStation

Below is a graph of a list of the best selling games on PlayStation 4

In [154…

```
ps4df = pd.read_csv('best_selling_PS4_games.csv')

#Convert copies sold to floats, remove the word "million"
ps4df['Copies sold'] = ps4df['Copies sold'].replace('[\$,]', '', regex=True).r

#Plot the graph
plt.figure(figsize=(12, 8))
plt.plot(ps4df['Game'], ps4df['Copies sold'])

highlightIndex = ps4df.index[ps4df['Game'] == 'The Last of Us Part II'].tolist

plt.annotate("The Last of Us Part II",
             xy=(4, ps4df.loc[highlightIndex, 'Copies sold']),
             xytext=(4, 14),
             arrowprops=dict(facecolor='black', shrink=0.05),
             horizontalalignment='left',
             verticalalignment='top')

plt.title('Copies Sold of Games')
plt.xlabel('Game Name')
plt.ylabel('Copies Sold (Millions)')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()
```



Once again, it's worth noting that The Last of Us Part II faces an uneven comparison, given its late release in the PlayStation 4 lifecycle. The preceding games had a substantial head

start of at least two additional years of sales. Despite this inherent disadvantage, the data highlights the game's remarkable performance, particularly in achieving sales comparable to The Last of Us Remastered.

Both the presented graphs strongly suggest that The Last of Us Part II can be deemed a success. Although its sales figures lack precision, the game stands out among the best-selling titles on the platform. However, it falls short of surpassing the success of the original The Last of Us, a widely acclaimed masterpiece among fans. A high bar to reach sure but based on sales alone it should've sold a lot more than it did; it should be one of the highest selling games and should've at least sold near 14 million like Red Dead Redemption 2. Despite the game being a success, it wasn't as successful as it could've been.

2.3 | Analysis on Budgets and Profit

Continuing the comparison with Red Dead Redemption 2, let's delve into specific figures related to these two games.

The Last of Us Part II generated over `$250 million` in revenue within its first three days of release, selling over four million copies worldwide, marking it as the fastest-selling PlayStation 4 exclusive during its release weekend [4].

In contrast, Red Dead Redemption 2 achieved the largest opening weekend in entertainment history, amassing over `$725 million` in revenue within three days and shipping over 17 million copies in two weeks [5].

It's evident that Red Dead Redemption 2 enjoyed tremendous success, partly owing to its availability on multiple platforms. This hints at The Last of Us Part II having the potential to be approximately three times more successful.

Another noteworthy PlayStation exclusive, God of War, generated over `$400 million` in revenue within three months of release [6]. God of War is widely considered a highly profitable game. Intriguingly, The Last of Us Part II achieved `$250 million` in just three days, surpassing God of War's three-month revenue, while the latter took three months to reach `$400 million`.

The budget for The Last of Us Part II was `$220 million` [7], and it became profitable in its release weekend alone.

With over 10 million copies sold (estimated revenue of `$50` per game), The Last of Us Part II would have made around `$500 million`, resulting in a profit of `$250 million`.

God of War, selling roughly 23 million copies with an estimated production cost of around `$291 million` [8], would have made approximately `$700 million` in profits.

While these numbers are estimates, it's crucial to note that in 2019, God of War officially confirmed making `$500 million` in revenue [8], not the roughly estimated `$700 million` I calculated. Nevertheless, this perspective places The Last of Us Part II in the context of other highly successful games.

The data indicates that despite having a similar budget, The Last of Us Part II didn't yield as much profit as other games listed here. For instance, the sequel to God of War has sold over 15 million copies [9] with a similar budget of around `$200 million`. This comparison underscores that, in terms of budgets and profits, The Last of Us Part II, while good, doesn't rank among the best.

I conclude that this discrepancy can be attributed to the divisive nature of the game. Its initial six months were marked by extremely negative reviews and reception, likely hindering some of its potential growth from the outset. The leak of Joel's death, a significant factor cited by users for disliking the game, prior to its release, probably also impacted its sales negatively.

2.4 | Analysis on Game Awards and Reception

As mentioned earlier, the initial reception of the game undoubtedly influenced its sales, suggesting that it had the potential to achieve higher numbers.

Despite its sales figures, The Last of Us Part II held the title of the most awarded video game for an extended period, securing 322 awards [10]. Only recently was it surpassed by Elden Ring with 323 awards. Notably, God of War ranks 4th, and Red Dead Redemption 2 is in 9th place.

Considering the critical acclaim, The Last of Us Part II outshone some games that surpassed it in sales and profits. This fact solidifies the argument that user reception significantly impacted its sales performance compared to the most successful games. Extrapolating from its critical acclaim, the game should have surpassed the success of many titles discussed in this report.

It's evident that sales and critical reception don't always align, indicating a distinction between the two aspects.

3 | Conclusion

The data presented in this report unequivocally establishes The Last of Us Part 2 as a successful game. Every metric explored here reinforces this assertion. The primary goal of this exploratory report was to identify aspects that users may have disliked about the game. The conclusion drawn is that the strong negative response upon its release significantly hampered its potential for higher sales and profits. Additionally, a glimpse into the game's

acclaimed success reveals that, for those who truly appreciate it, The Last of Us Part 2 undeniably stands among the greatest games ever created.

Addressing the initial question posed – "Just how good is The Last of Us Part 2?" – I contend that, at the very least, the game stands alongside some of the finest games ever made. The analysis indicates that the game had the potential to be among the very best. While it may not claim the title of 'the' very best, it certainly merits inclusion in the conversation.

What does that mean for 'The Last of Us Part 3'? All I can say, is the data supports the making of it.

4 | References

4.1 | Cited Sources

[1] Statista, Lifetime unit sales of The Last of Us Part II worldwide as of June 2022, Statista. [Online]. Available: https://www.statista.com/statistics/1135726/the-last-of-us-part-ii-units-sold/
[2] Lázaro, The Last of Us II: a sentiment analysis, Kaggle. [Online]. Available: https://www.kaggle.com/code/lazaro97/the-last-of-us-ii-a-sentiment-analysis
[3] Dave Trumbore, 'The Last of Us: Part II': One Year Later and the Worst Takes Still Remain, Collider. [Online]. Available: https://collider.com/last-of-us-2-controversy-explained/
[4] Wikipedia, The Last of Us Part II, Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/The_Last_of_Us_Part_II#Sales
[5] Wikipedia, Red Dead Redemption 2, Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Red_Dead_Redemption_2#Sales
[6] Wikipedia, God of War (franchise), Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/God_of_War_(franchise)#Critical_reception
[7] Wikipedia, List of most expensive video games to develop, Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/List_of_most_expensive_video_games_to_develop
[8] Marie L., How much did God of War 2018 cost to make?, Rankiing. [Online]. Available: https://rankiing.net/how-much-did-god-of-war-2018-cost-to-make/
[9] Statista, Lifetime unit sales generated by God of War: Ragnarök worldwide as of November 2023, Statista. [Online]. Available: https://www.statista.com/statistics/1364679/god-of-war-ragnaroek-lifetime-unit-sales/#:~:text=As%20of%20November%202023%2C%20the,selling%20first%2Dparty%20title
[10] Reddit, The Last of Us Part II has been dethroned by Elden Ring as the most awarded video game of all time, Reddit. [Online]. Available: https://www.reddit.com/r/thelastofus/comments/1025m1g/the_last_of_us_part_ii_has_been_det

4.2 | Datasets

https://www.kaggle.com/datasets/lazaro97/the-last-of-us-reviews

https://www.kaggle.com/datasets/holmjason2/videogamedata

https://www.kaggle.com/datasets/drahulsingh/list-of-best-selling-ps4-games