Harry Potter

RESEARCH METHODS FOR BUSINESS ADMINISTRATION

STUDENTS:

BĂLUNĂ ȘTEFAN-DANIEL
DIDU IULIA-ALEXANDRA
POPA ANA-MARIA

Harry Potter

MOVIE REVIEWS ANALYSIS

1. INTRODUCTION

"Harry Potter and the Philosopher's Stone" was a massive success both critically and commercially, marking the beginning of the highly popular film franchise based on J.K. Rowling's book series. The film was a blockbuster at the box office, grossing over $974 million worldwide. It became the highest-grossing film of 2001 and one of the highest-grossing films of all time at the point of its release.

It had a profound impact on popular culture, introducing a new generation of fans to the wizarding world and becoming a cultural phenomenon that extended beyond the realm of literature. The film received several awards and nominations. Although it didn't win any Academy Awards, it was nominated for three Oscars in the categories of Best Art Direction, Best Costume Design, and Best Original Score.

The success of the first film laid the foundation for the entire Harry Potter film series. Warner Bros. continued to adapt each of the subsequent novels into movies, resulting in a total of eight films that collectively became one of the highest-grossing franchises in cinematic history.

The story follows a young orphan named Harry Potter, played by Daniel Radcliffe, who discovers on his eleventh birthday that he is a wizard. He also learns that he is famous in the wizarding world for surviving an attack by the dark wizard Lord Voldemort as an infant, which left him with a lightning-shaped scar on his forehead.

Harry is invited to attend Hogwarts School of Witchcraft and Wizardry, where he makes friends with Ron Weasley (Rupert Grint) and Hermione Granger (Emma Watson). Together, they uncover the mystery surrounding the Philosopher's Stone, a magical object that grants immortality. They discover that someone is attempting to steal it, leading to a dangerous adventure involving magical creatures, puzzles, and the dark forces that Harry thought he had escaped.

Throughout the film, the trio learns about friendship, courage, and the magical world they are now a part of. The movie introduces viewers to various iconic elements of the Harry Potter universe, including Quidditch, magical creatures like Fluffy the three-headed dog, and the wise and kind headmaster Albus Dumbledore, played by Richard Harris.

"Harry Potter and the Philosopher's Stone" sets the stage for the subsequent films in the series, establishing the characters, the magical world, and the ongoing battle between Harry and the forces of darkness led by Voldemort.

2. AIM OF THE ANALYSIS

Given the fact that the Harry Potter movies were a worldwide success and their impact lasts even nowadays, the aim of our analysis was to find out how the audience felt about the first movie of the Harry Potter franchise, to check their reviews and to find a pattern based on age, gender and country with regards to their favourite characters, rating and, last but not least, to find the most used words in all of the ratings, generating a wordcloud.

3. METHODS

In this section of the project, we will first outline the procedures by describing the code and its intended function. We will next go over the outcomes and all the necessary data to support our findings. Python was the processing language of choice for data analysis, where data was sorted, processed, and presented in order to illustrate and explain the results.

We started by importing the necessary tools and addons in order to make our program work.

```
!pip install pandas

import pandas as pd

raw_csv_data = pd.read_csv("/Users/iuliadidu/Desktop/harry_potter_reviews.csv")

# Import the necessary Python libraries

import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.ticker as mtick

from PIL import Image
!pip install wordcloud

from wordcloud import WordCloud
```

Then, we read the CSV file, where all the necessary data was stored and we displayed basic information about the data frame.

```
# Read the data from the csv file. Save it in a DataFrame.
df = pd.read_csv("/Users/iuliadidu/Desktop/harry_potter_reviews.csv")
df.head()

# Display basic information about the DataFrame

df.info()

# There are 491 rows and 8 columns
# Fortunatelly, there are no missing values
```
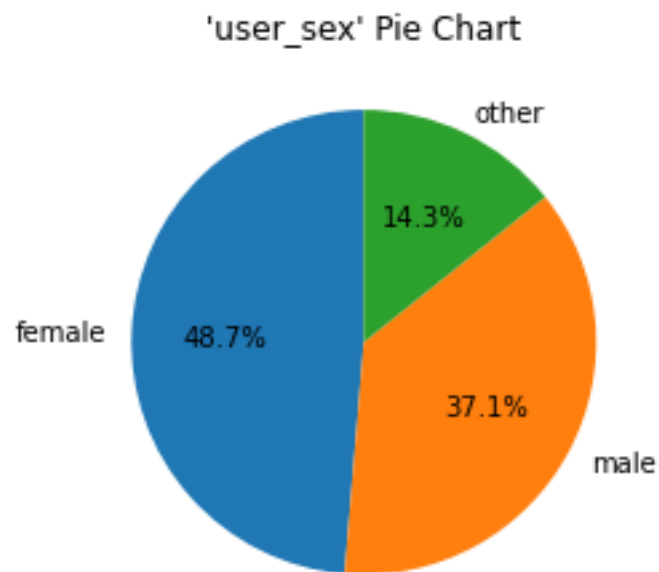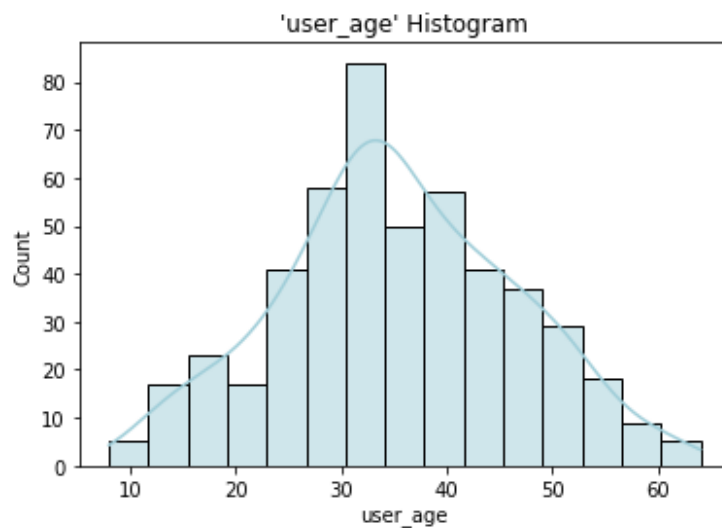
User sex pie chart:

```python
user_sex_counts = df['user_sex'].value_counts(normalize=True)

plt.pie(
    x=user_sex_counts.values,
    labels=user_sex_counts.index,
    autopct='%1.1f%%',
    startangle=90,
)
_ = plt.title("'user_sex' Pie Chart")
```


User age histogram:

```python
ax = sns.histplot(df['user_age'], color='#A0CED9', kde=True)
_ = ax.set(title="'user_age' Histogram")
```

User country count plot:

```python
# Sort the list of unique countries by number of occurrences
sorted_countries = df['user_country'].value_counts().index

# Count Plot of the different countries in data
ax1 = sns.countplot(
    x=df['user_country'],
    color='#A0CED9',
    order=sorted_countries
)

# Create a second y-axis for percentages (%)
ax2 = ax1.twinx()

# Compute the pct of ocurrences of each country
pct_series = (df['user_country'].value_counts() / len(df)) * 100

# (Dummy) Plot to show percentages on the second y-axis
sns.lineplot(
    x=pct_series.index,
    y=pct_series.values,
    color='white',
    marker='o',
    ax=ax2
)
# Note that the actual lineplot's shape and form does not matter

# Send the dummy white percent plot to the back
ax2.set_zorder(1)
ax1.set_zorder(2)

# Rotate the Xlabels (45º, diagonal)
ax1.set_xticklabels(
    labels=sorted_countries,
    rotation=45,
    ha='right',
    rotation_mode='anchor',
)
ax1.set_ylabel('Count')
ax2.set_ylabel('Percentage', color='black')

# Set the second y-axis to percentage
ax2.yaxis.set_major_formatter(mtick.PercentFormatter())

# Add the title
_ = ax1.set_title("'user_country' Count Plot")
```
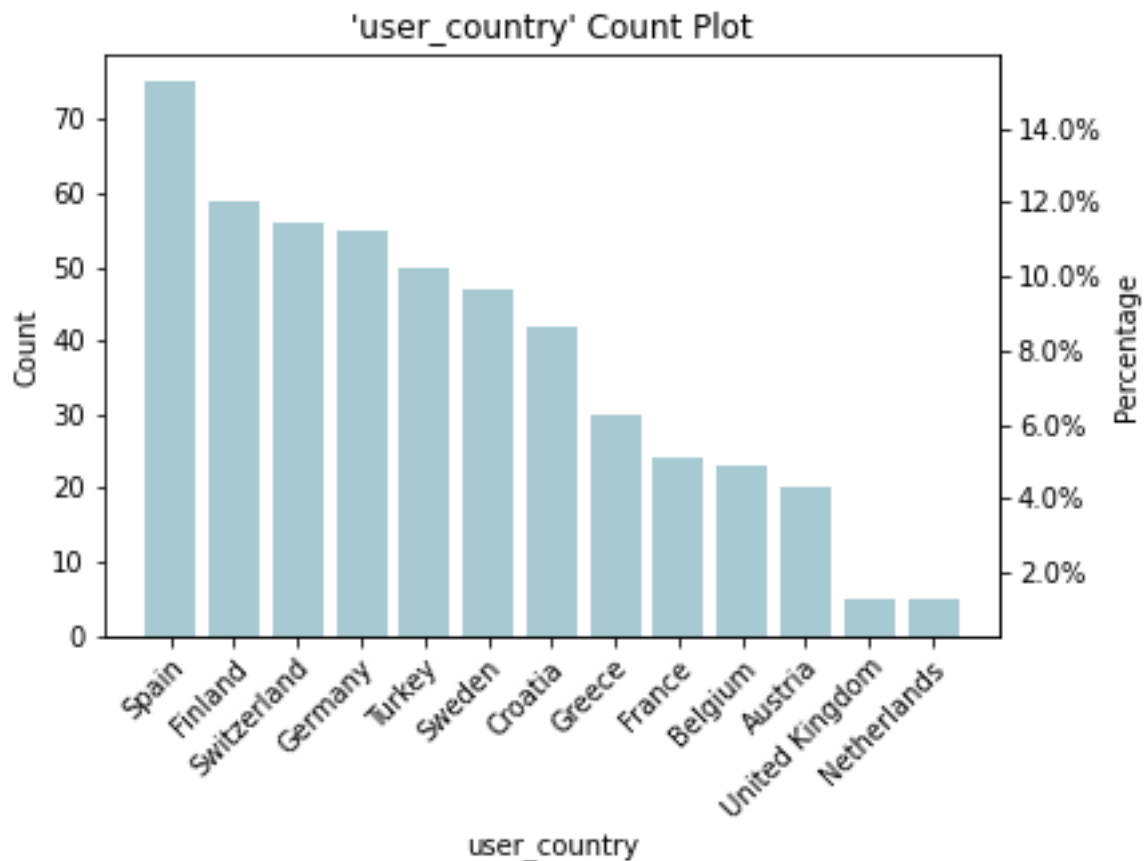
'user_country' Count Plot

Rating count plot:

```
# Sort the list of ratings ascending
rating_labels = sorted(df['rating'].unique())

# Count Plot of the different ratings in data
ax1 = sns.countplot(
    x=df['rating'],
    color='#A0CED9',
    order=rating_labels
)

# Create a second y-axis for percentages (%)
ax2 = ax1.twinx()

# Compute the pct of ocurrences of each rating
pct_series = (df['rating'].value_counts() / len(df)) * 100

# (Dummy) Plot to show percentages on the second y-axis
sns.lineplot(
    x=pct_series.index,
    y=pct_series.values,
    color='white',
    marker='o',
    ax=ax2
)
# Note that the actual lineplot's shape and form does not matter

# Send the dummy white percent plot to the back
ax2.set_zorder(1)
ax1.set_zorder(2)

# Rotate the Xlabels (45º, diagonal)
ax1.set_xticklabels(
    labels=rating_labels,
    rotation=45,
    ha='right',
    rotation_mode='anchor',
)
ax1.set_ylabel('Count')
ax2.set_ylabel('Percentage', color='black')

# Set the second y-axis to percentage
ax2.yaxis.set_major_formatter(mtick.PercentFormatter())

# Add the title
_ = ax1.set_title("'rating' Count Plot")
```
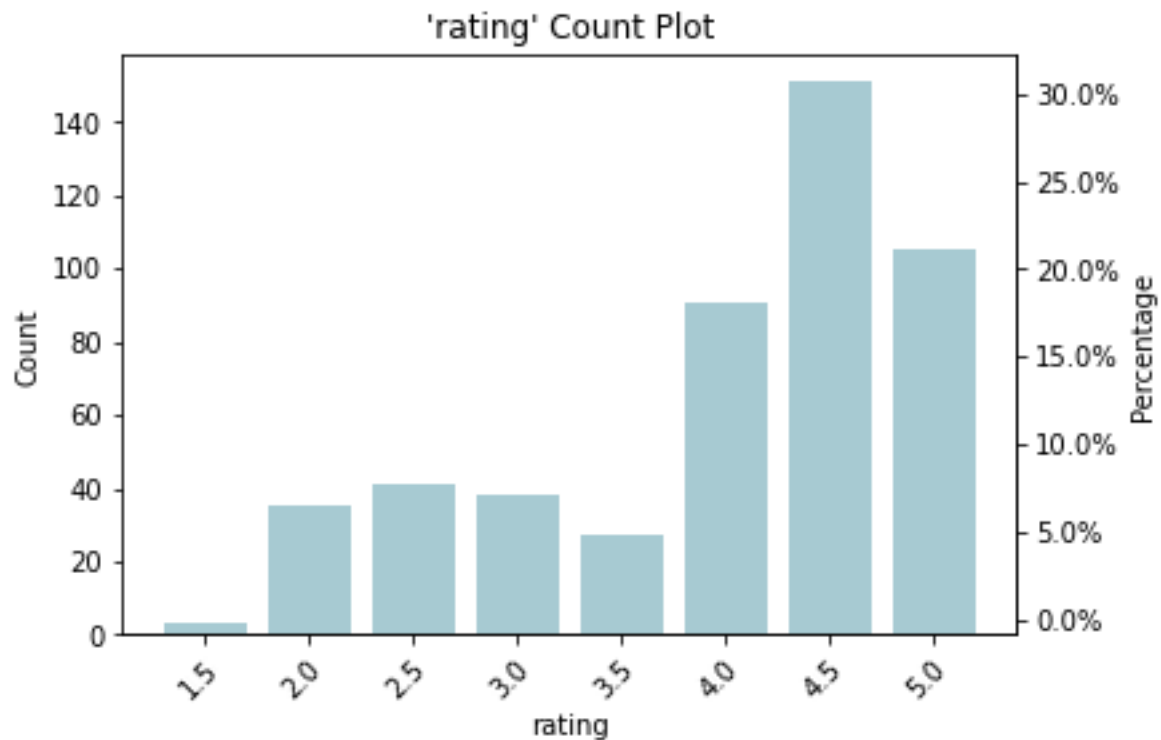
'rating' Count Plot

Next, we generated a word cloud which includes the most used words in the ratings:

```python
# Let's create a wordcloud based on all the comments (reviews)
text = ' '.join(review for review in df['comment'])


# Create and generate a word cloud image:
wordcloud = WordCloud(
    background_color="white",
).generate(text)

# Display the generated image:
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
```

And, last but not least, we generated a favourite character count plot:

```python
# Sort the list of characters by the number of occurrences
sorted_fav_chars = df['favourite_character'].value_counts().index

# (Sorted) Count Plot of the different characters in data
ax1 = sns.countplot(
    x=df['favourite_character'],
    color='#A0CED9',
    order = sorted_fav_chars
)

# Create a second y-axis for percentages (%)
ax2 = ax1.twinx()

# Compute the pct of ocurrences of each character
pct_series = (df['favourite_character'].value_counts() / len(df)) * 100

# (Dummy) Plot to show percentages on the second y-axis
sns.lineplot(
    x=pct_series.index,
    y=pct_series.values,
    color='white',
    marker='o',
    ax=ax2
)
# Note that the actual lineplot's shape and form does not matter

# Send the dummy white percent plot to the back
ax2.set_zorder(1)
ax1.set_zorder(2)

# Rotate the Xlabels (45º, diagonal)
ax1.set_xticklabels(
    labels=sorted_fav_chars,
    rotation=45,
    ha='right',
    rotation_mode='anchor',
)
ax1.set_ylabel('Count')
ax2.set_ylabel('Percentage', color='black')

# Set the second y-axis to percentage
ax2.yaxis.set_major_formatter(mtick.PercentFormatter())

# Add the title
_ = ax1.set_title("'favourite_character' Count Plot")
```
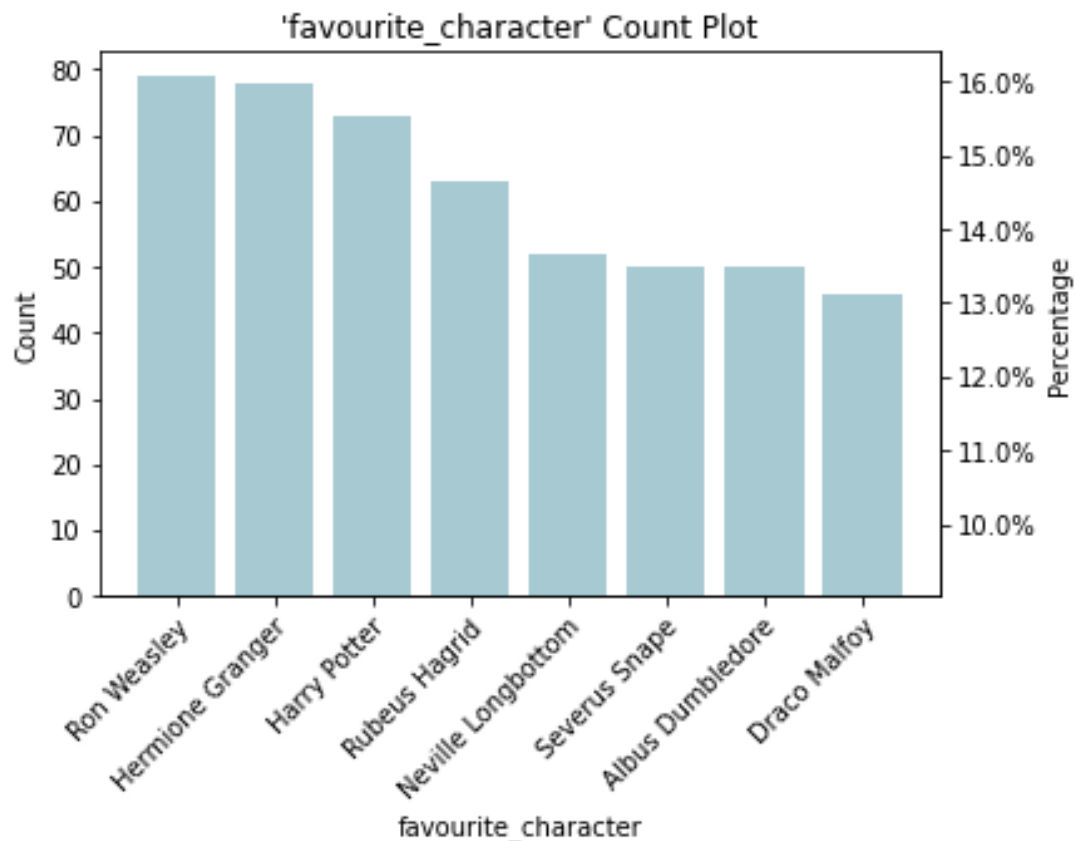

'favourite_character' Count Plot

4. DISCUSSION AND FINDINGS

In order to sum up, our main findings are the following:

a. the females were more opinionated and eager to share their reviews with the rest of the world. They represent 48.7% of the audience.

b. the ages of the users are between 10 and 60, most of the users which expressed their opinions being between 30 and 35 years old. This shows that Harry Potter movies are magical for each age, keeping not only the children engaged, but also adults.

c. Most reviews are coming from Spain, followed by Finland and Switzerland and the least present are the reviews from UK and The Netherlands.

d. The most used words in the wordcloud were "film", "Draco Malfoy", "Neville Longbottom", "Severus Snape" and "magical world".

e. With regards to the favourite characters, there is no surprise that the most loved characters by the audience are the magic trio: Ron Weasley, Harmonie Granger and Harry Potter. The most disliked one is Draco Malfoy, followed by Albus Dumbledore.

5. CONCLUSION

In conclusion, the impact of the Harry Potter series on readers of all ages and, later on, of the Harry Potter franchise, especially the first movie which we chose to analyse, has been profound and far-reaching since the release of the first book, "Harry Potter and the Philosopher's Stone" (titled "Harry Potter and the Sorcerer's Stone" in the United States), in 1997. The findings of our analysis show that each and every one of us needs a little magic in their life, a place to escape from the reality and from the every day life, which allows us to feel like a child again.