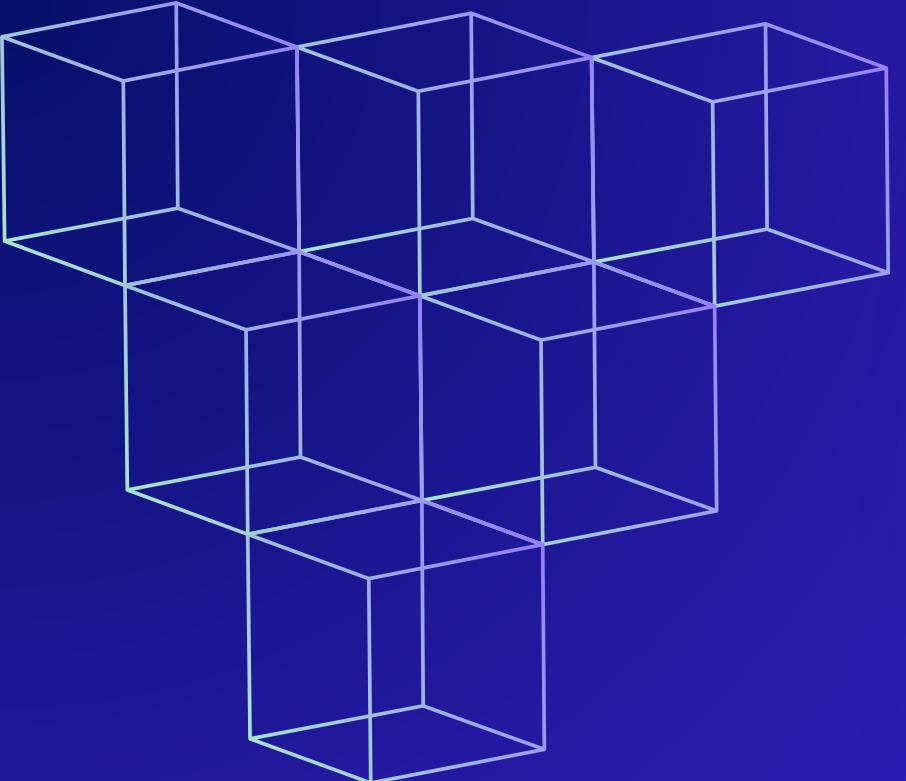


# STEAM GAMES

# PROJECT

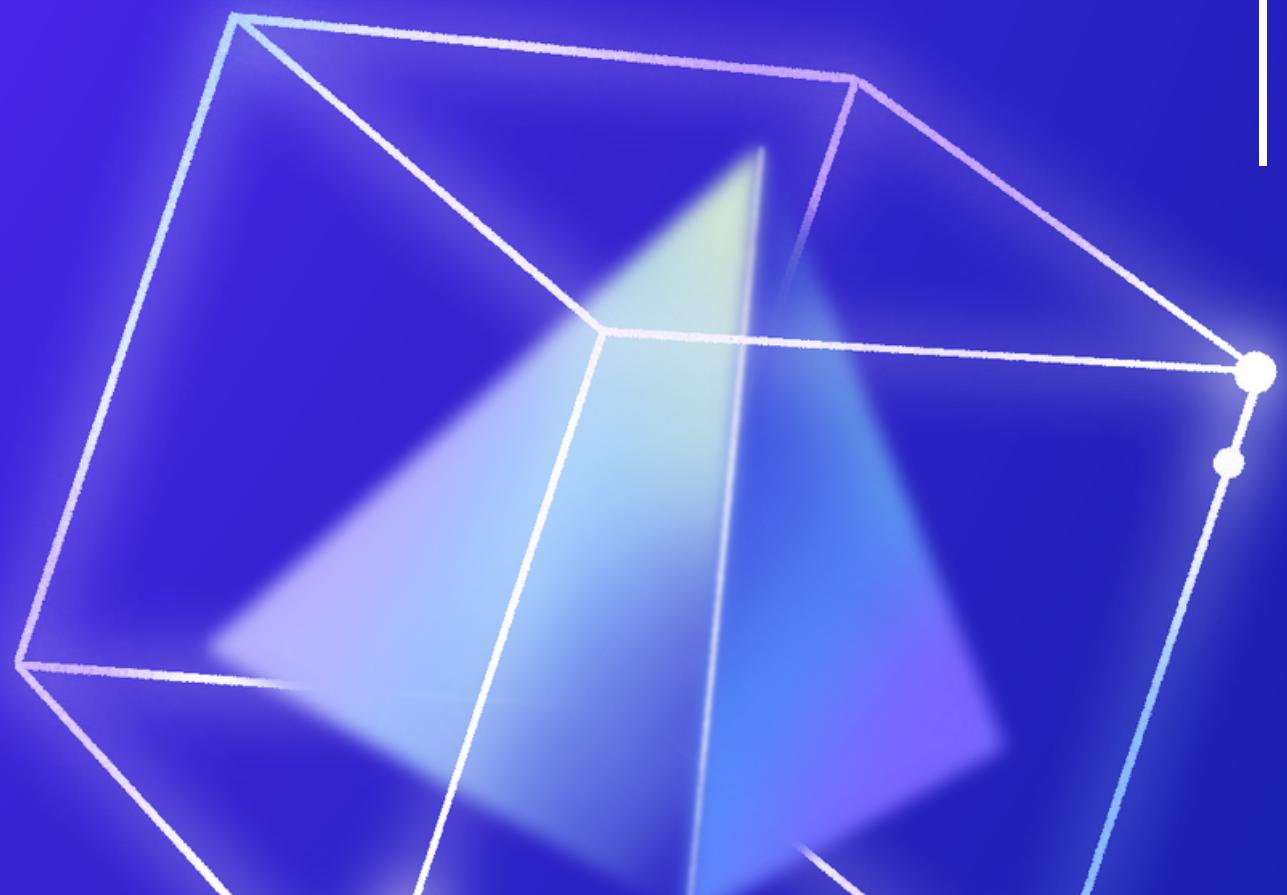
Polina Bykova





# TABLE OF CONTENTS

• Introduction	01
• Why games?	02
• Methodology	03
• Results and Achievements	04
• Plots	05
• Conclusion	09
• Small quiz	



# INTRODUCTION

The goal of our research today is not just to solve the data puzzle, but to understand what numbers will tell us about player preferences, industry trends, and how this data can become a valuable tool for developers and investors.



# WHY GAMES?

**Games provide an opportunity to relax and have fun. They allow people to escape from their daily routine, immerse themselves in other worlds and experience new experiences.**

**Playing games promotes interaction and communication with other players. It can be a way for people to communicate, especially those who may have difficulty communicating in real life.**

**Some games develop cognitive skills, logic, strategic thinking and problem solving. It can be a form of learning hidden behind entertainment.**



## METHODOLOGY

The [pandas library](#) was used during the data analysis phase of the project. Starting with the data collection phase, [CSV files](#) were imported from [Kaggle](#). Using the pandas functionality, data preprocessing was performed, including removing duplicates and transforming data if necessary.

The [Matplotlib library](#) was used to visualize the analysis results and create informative graphs. Using Matplotlib, various types of graphs were constructed to visually present the results of the data analysis, making them more understandable and accessible for interpretation.

# RESULTS AND ACHIEVEMENTS

## **As a results:**

- Created a time series plot to visualize how the number of game releases has changed.
- Explored the distribution of games across different platforms (Windows, Mac, Linux, Steam Deck). This provided insights into platform preferences among developers and users.
- Examined the distribution of game ratings. Built a bar chart to understand the spread of ratings in the dataset.

```
[1]: # Importing all necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import os
from collections import Counter
from datetime import datetime
pd.set_option('display.float_format', lambda x: '%.3f' % x)
```

The dataset consists of three main entities:

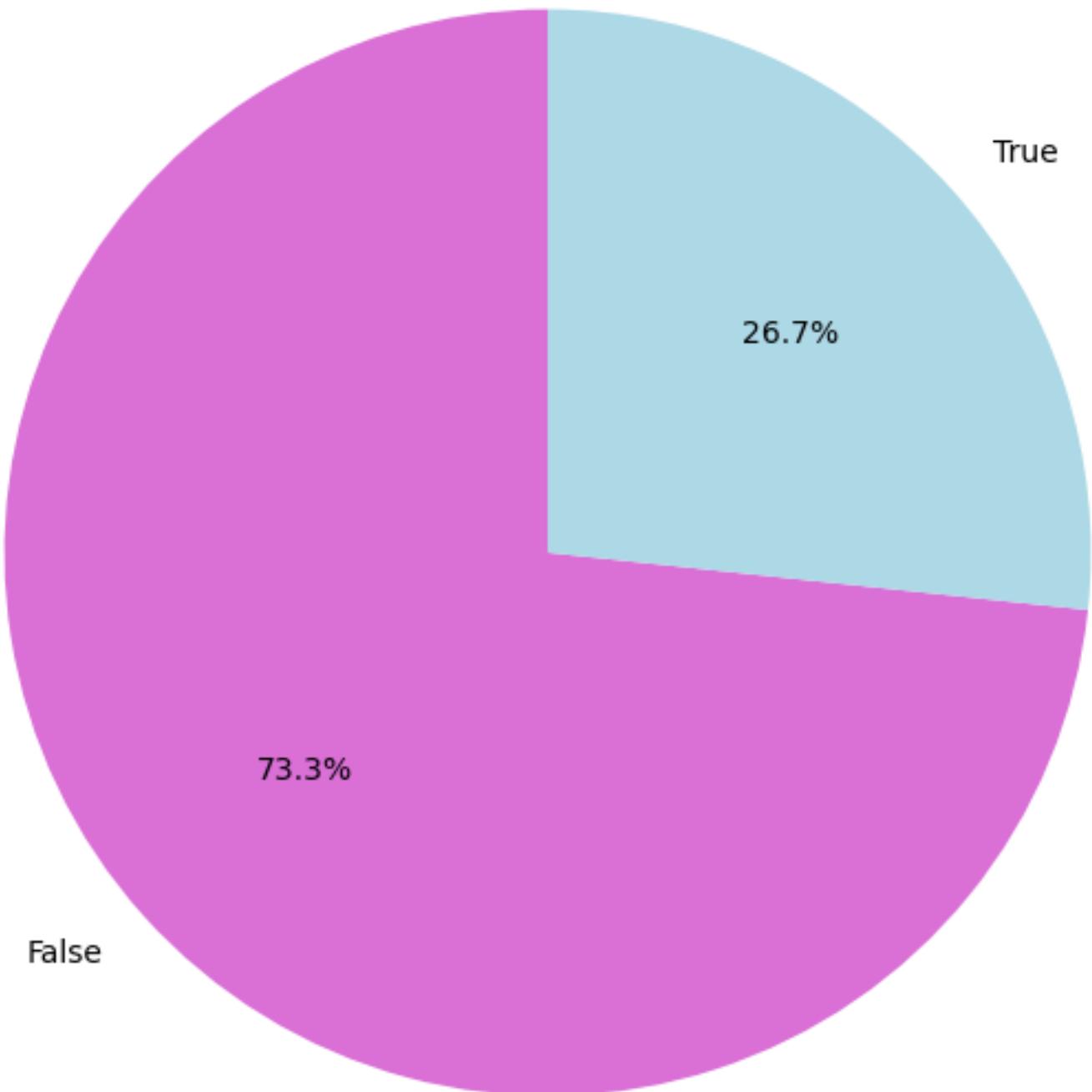
- [games.csv](#) - a table of games (or add-ons) information on ratings, pricing in US dollars \$, release date, etc. A piece of extra non-tabular details on games, such as descriptions and tags, is in a [metadata file](#);
- [users.csv](#) - a table of user profiles' public information: the number of purchased products and reviews published;
- [recommendations.csv](#) - a table of user reviews: whether the user recommends a product. The table represents a many-many relation between a game entity and a user entity.

```
[2]: # Loading the data
games_data = pd.read_csv('./games.csv')
recommendations_data = pd.read_csv('./recommendations.csv')
users_data = pd.read_csv('./users.csv')
```

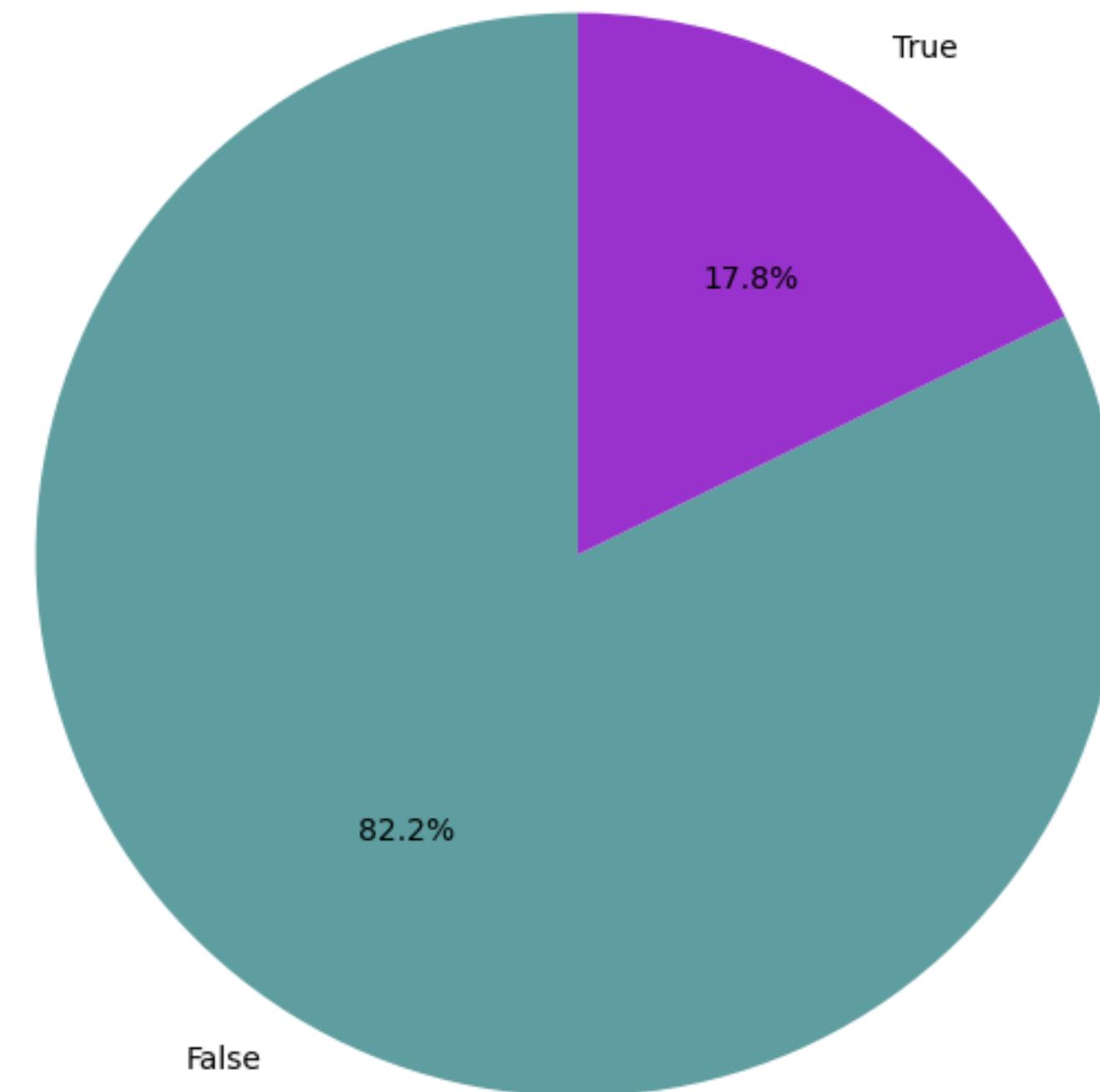
```
[3]: # Looking at our data
games_data.head(5)
```

	app_id	title	date_release	win	mac	linux	rating	positive_ratio	user_reviews	price_final	price_original	discount	steam_deck
0	13500	Prince of Persia: Warrior Within™	2008-11-21	True	False	False	Very Positive	84	2199	9.990	9.990	0.000	True
1	22364	BRINK: Agents of Change	2011-08-03	True	False	False	Positive	85	21	2.990	2.990	0.000	True
2	113020	Monaco: What's Yours Is Mine	2013-04-24	True	True	True	Very Positive	92	3722	14.990	14.990	0.000	True
3	226560	Escape Dead Island	2014-11-18	True	False	False	Mixed	61	873	14.990	14.990	0.000	True
4	249050	Dungeon of the ENDLESS™	2014-10-27	True	True	False	Very Positive	88	8784	11.990	11.990	0.000	True

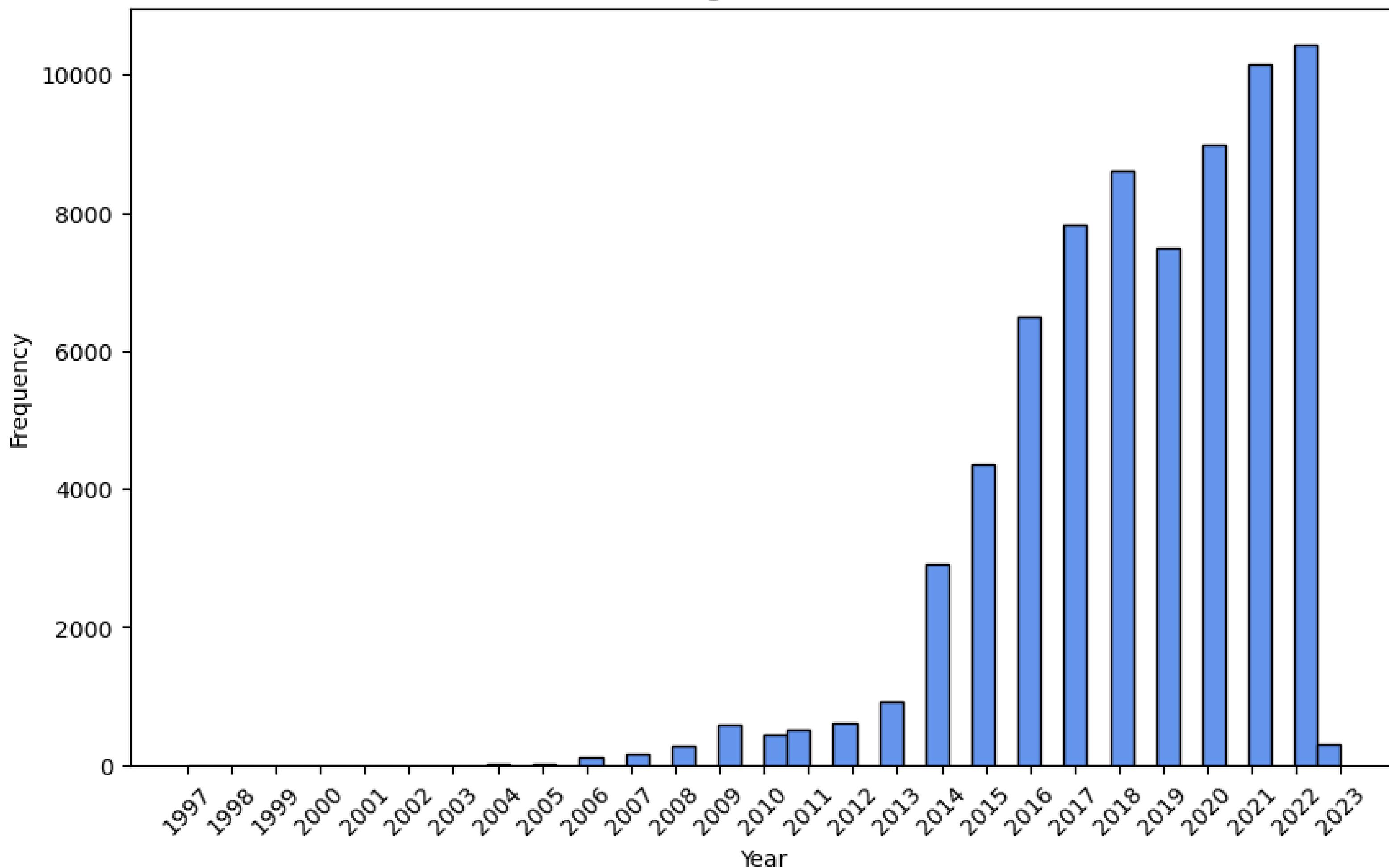
Game is available on Mac OS



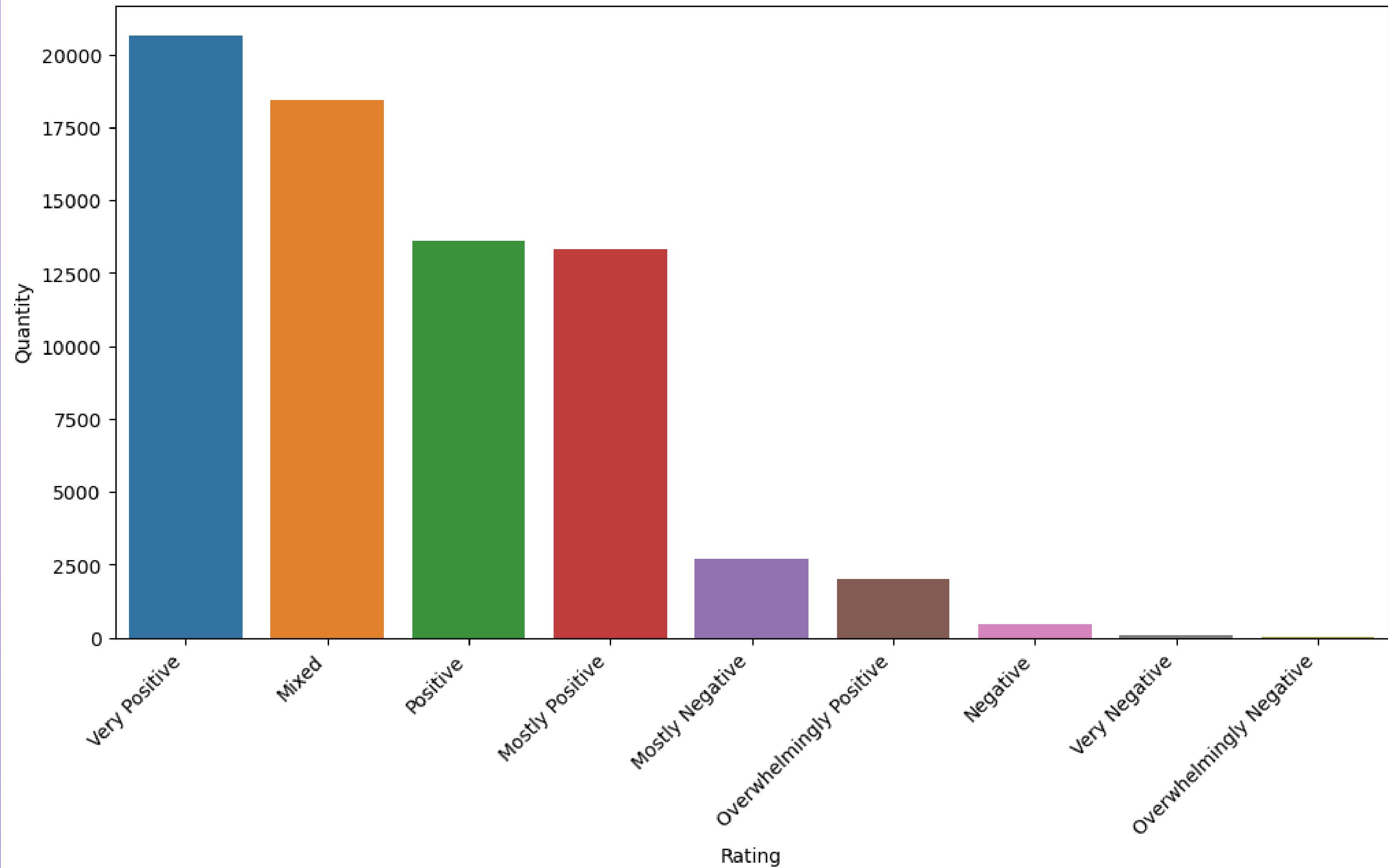
Game is available on Linux



### Histogram of Years



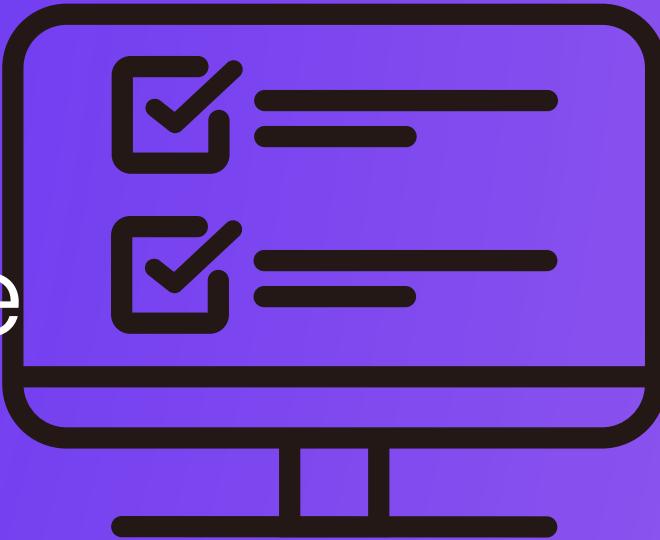
Distribution of ratings



# CONCLUSION

To sum it up, analyzing data from Steam helps game folks make smart choices. By understanding what players like, how the price behaves, and how games perform, they can make better decisions about making and selling games.

It's like having a roadmap for success in the gaming world.



THANK YOU !

