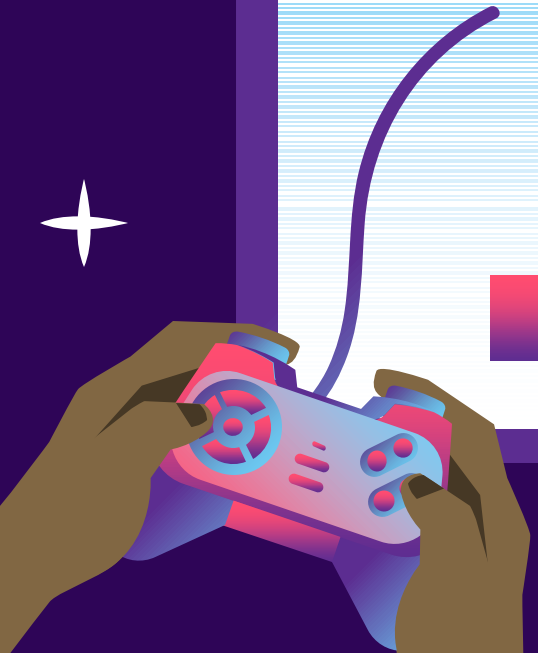
Professional Video Game Reviews

Milestone #1: Data Preparation (PRI)



Gonçalo de Abreu Matias up202108703@up.pt
Henrique Filipe Pereira da Silva Caridade up202108817@up.pt
Pedro da Cunha Teixeira e Faro Beirão up202108718@up.pt
Pedro Miguel Marques da Silva Pinto up202108826@up.pt

Project Overview

01 Datasets

02 Pipeline

03 Domain Conceptual Model

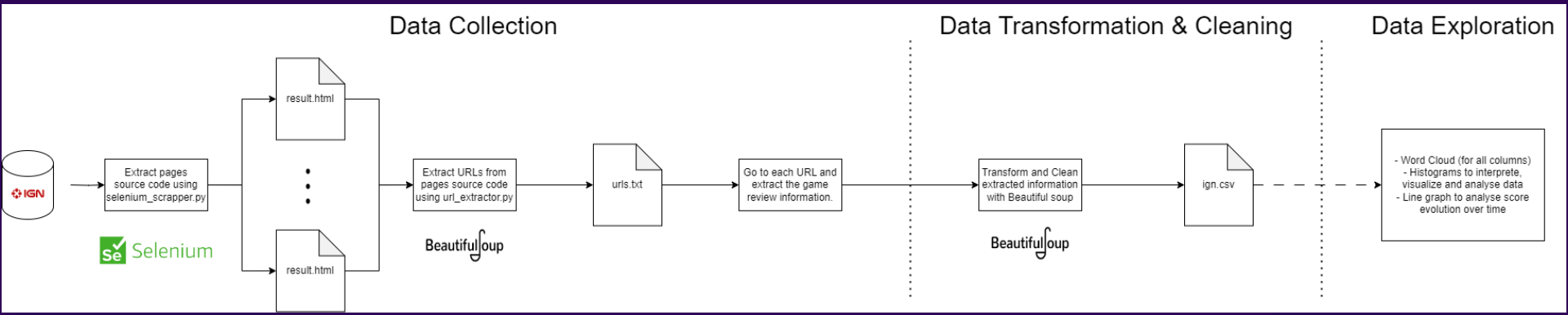
04 Data Characterization

05 Search Scenarios

01 Datasets

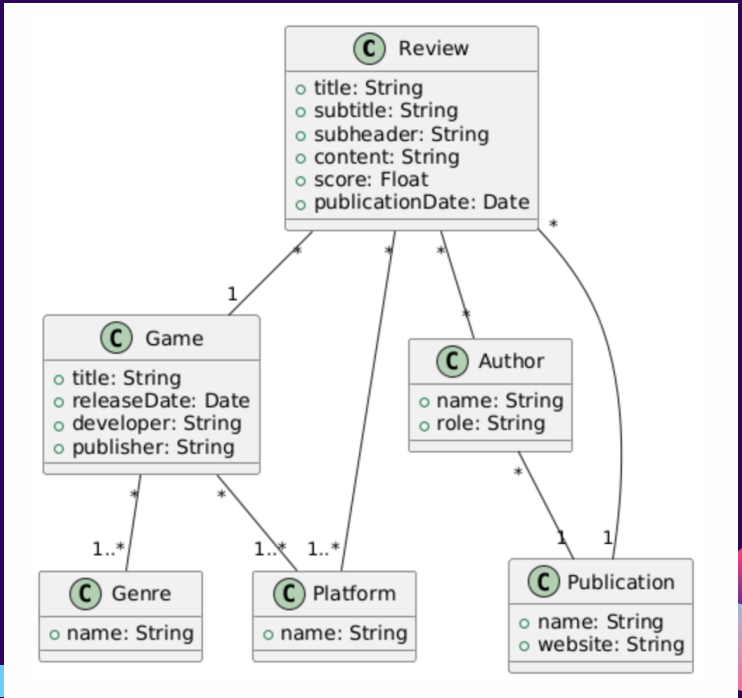
- Number of rows: 6937
- Number of columns: 5 ['Title', 'Subtitle', 'Subheader', 'Content', 'Score']
- ign.csv
- From *www.ign.com/reviews/games*

02 Pipeline


Data Collection
Data Transformation & Cleaning
Data Exploration
IGN
Extract pages source code using selenium_scrapper.py
Selenium
result.html
result.html
Extract URLs from pages source code using url_extractor.py
Beautiful Soup
urls.txt
Go to each URL and extract the game review information.
Transform and Clean extracted information with Beautiful soup
Beautiful Soup
ign.csv
- Word Cloud (for all columns)
- Histograms to interprete, visualize and analyse data
- Line graph to analyse score evolution over time

03

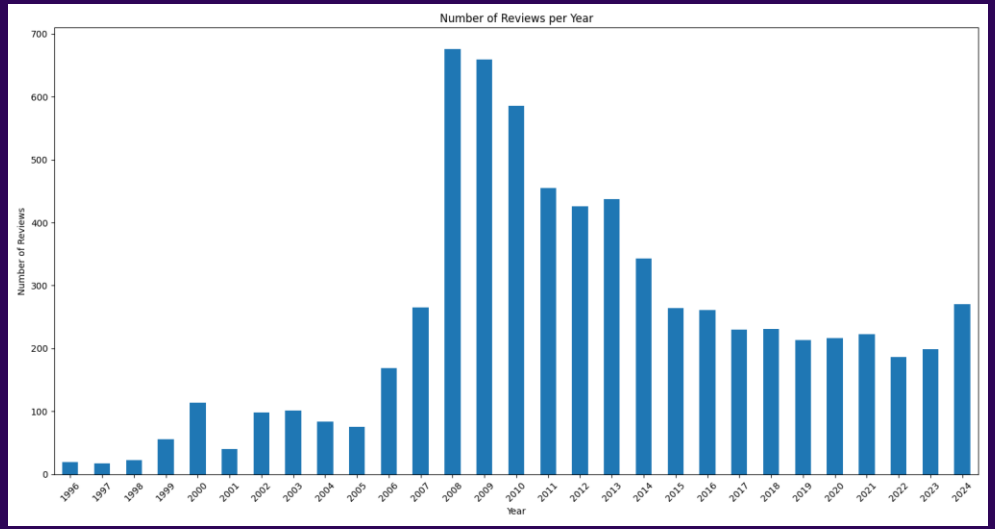
Domain Conceptual Model



04

Data Characterization

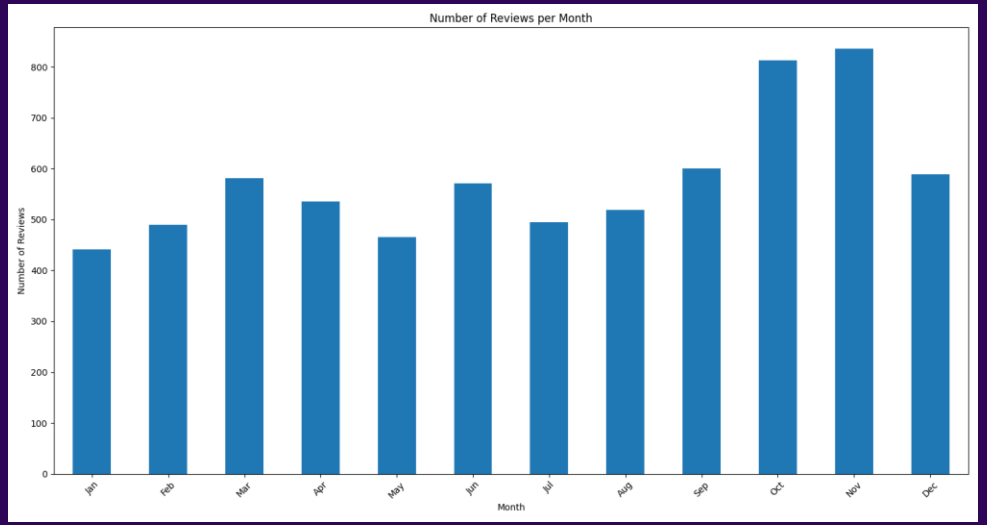
Number of Reviews per Year



04

Data Characterization

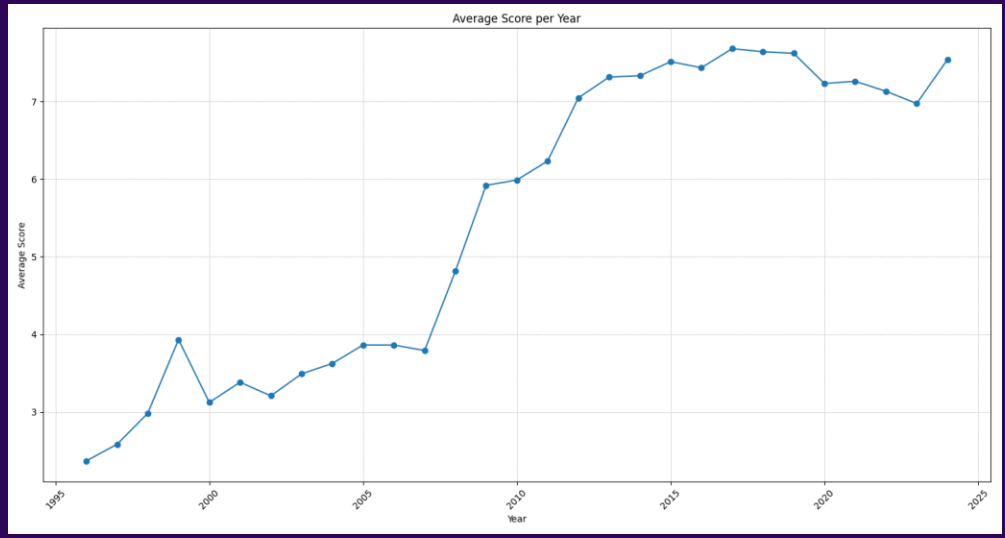
Number of Reviews per Month



04

Data Characterization

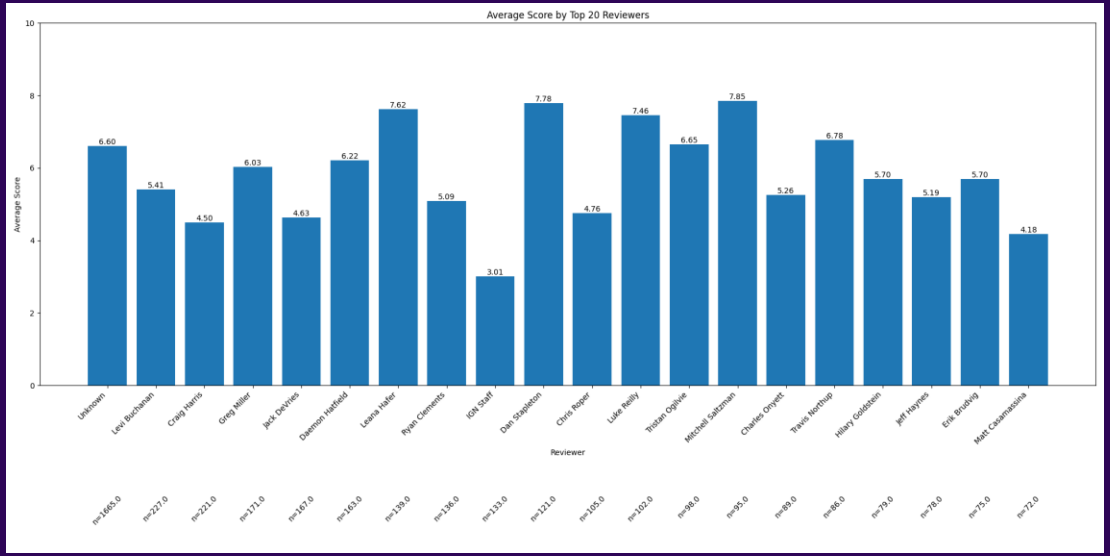
Average Score per Year



04

Data Characterization

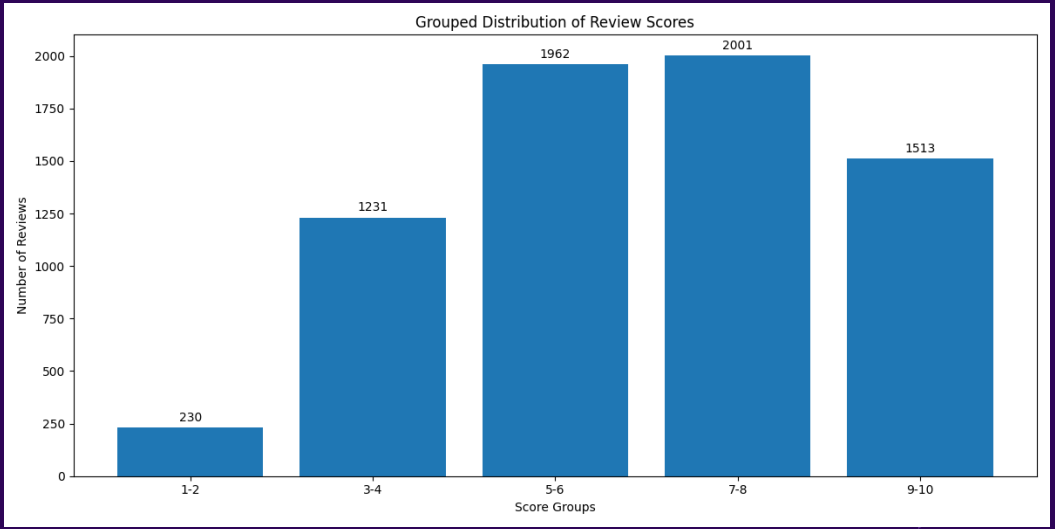
Average Score by Top 20 Reviewers



04

Data Characterization

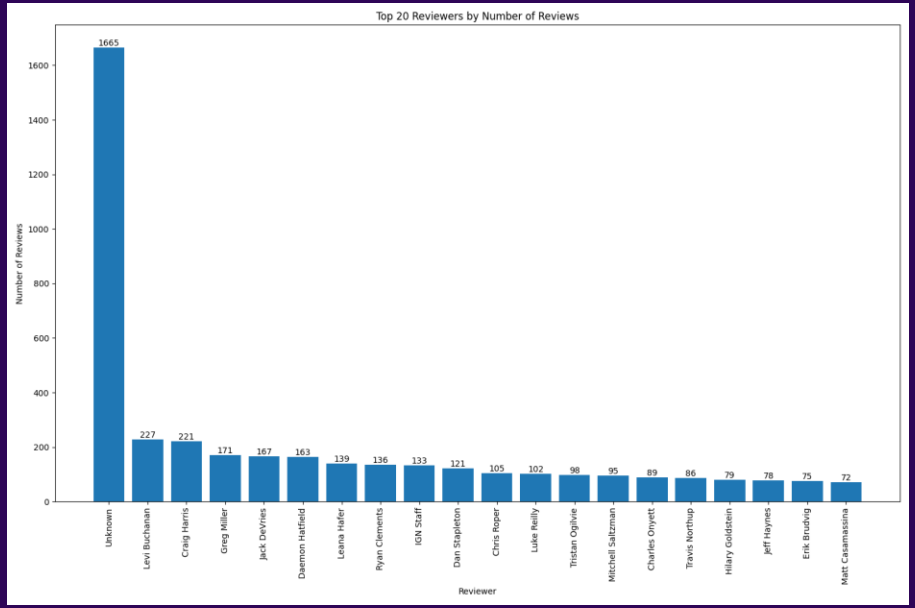
Grouped Distribution of Review Scores



04

Data Characterization

Top 20 Reviewers by Number of Reviews



04

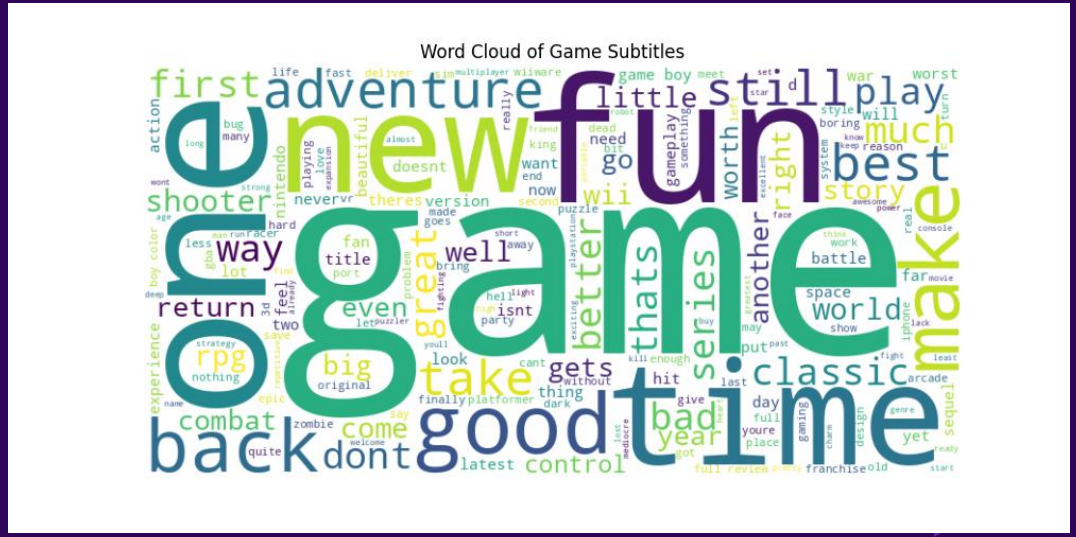
Data Characterization

Word Cloud of Game Titles

04

Data Characterization

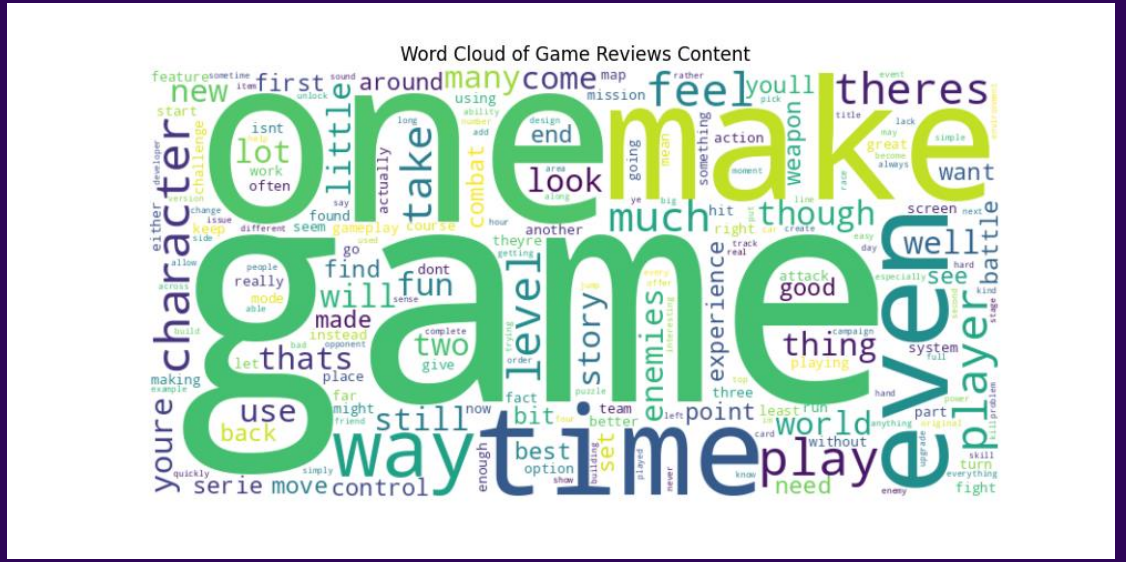
Word Cloud of Game Subtitles



04

Data Characterization

Word Cloud of Game Reviews Content



05

Search Scenarios

You can enter a subtitle here if you need it

Search Scenarios

Titles

Search game reviews by title

Authors

Finding reviews by author

Scores

Finding reviews based on game score

Subtitles

Exploring reviews based on subtitles

Dates

Finding reviews based on their release date

Genres

Interest in reviews based on game genre

Conclusion

All objectives were accomplished for this milestone

Thanks!

Do you have any questions?