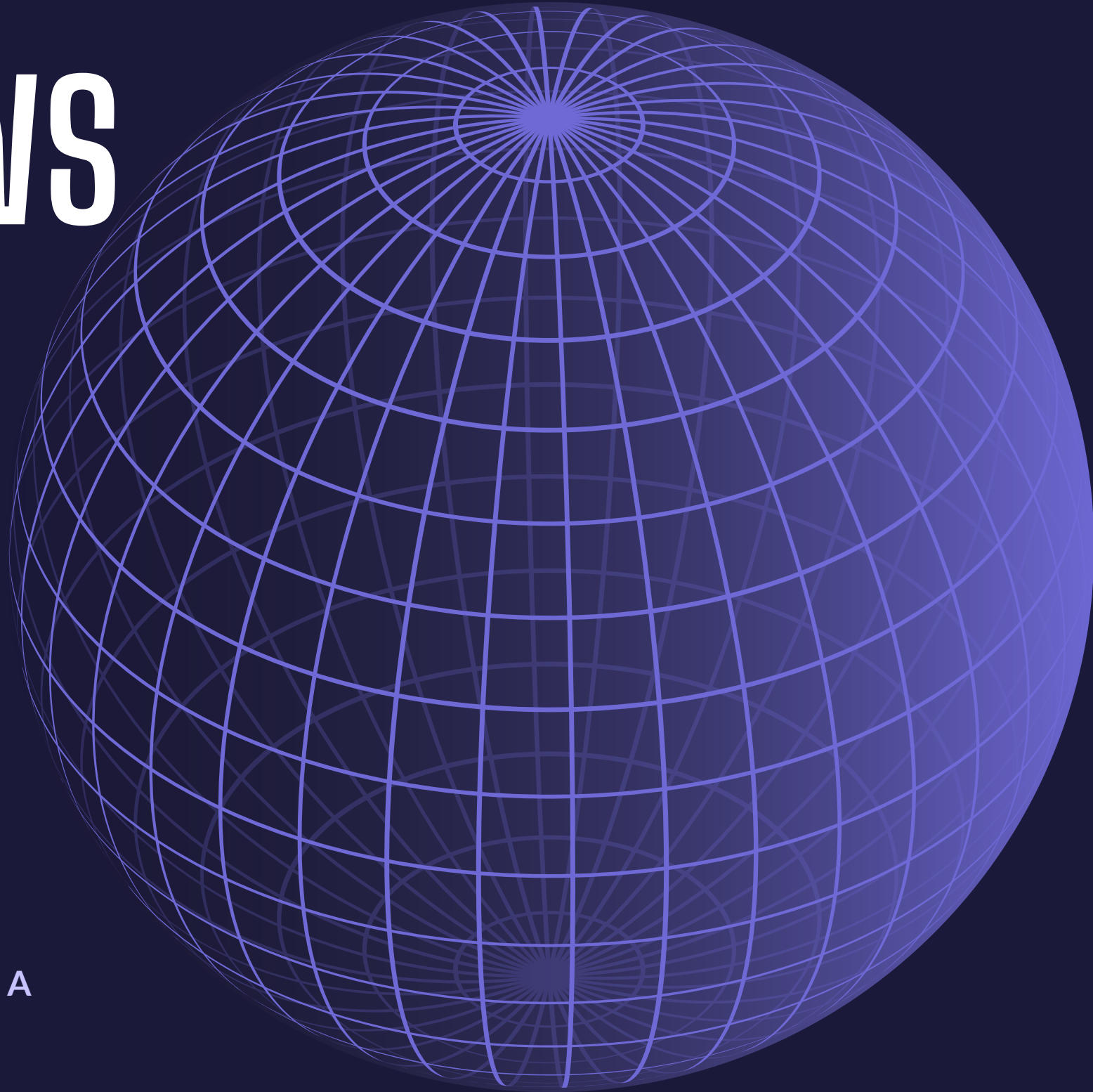


PLAYER REVIEWS FOR DEVELOPMENT

TEAM 7 : LUCIEN WILLIAMS, RHEA DESAI,
CHAITALI DESHMUKH, VANESSA SUNG, TINA
ZHAO





WHY THEY MATTER

for developers

- *Player reviews reveal:*
 - What players *value*
 - Where they feel *frustrated*
 - Which game characteristics consistently lead to *high satisfaction*
- *This can provide early evidence of product-market fit*
 - focus on certain features
 - price points
 - genres



DATA INSIGHTS

Our stem dataset includes:



- 51.5M English reviews
 - Data includes *sentiment, playtime...*
- 111K Steam games
 - Data includes *price, genre, tags...*

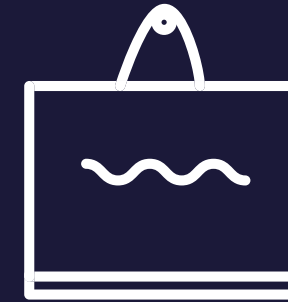
1) KAGGLE: STEAM USER REVIEWS GET LIST API - 100 MILLION+ STEAM REVIEWS 43GB ([HTTPS://WWW.KAGGLE.COM/DATASETS/KIERANPOC/STEAM-REVIEWS](https://www.kaggle.com/datasets/kieranpoc/steam-reviews))

2) KAGGLE: STEAM API & STEAM SPY - STEAM GAMES DATASET 1GB ([HTTPS://WWW.KAGGLE.COM/DATASETS/FRONKONGAMES/STEAM-GAMES-DATASET](https://www.kaggle.com/datasets/fronkongames/steam-games-dataset))

UNDERSTANDING THE DATA DEEPER

OVERALL SENTIMENT TRENDS

- 85% of all English reviews are positive
 - Steam players tend to give favorable ratings
 - So negative sentiment stands out more
- Negative reviews usually relate to something structural
 - Monetization
 - Bugs
 - Gameplay systems



GENRE EFFECTS

- Creative tools (tags related to animation, photo editing, modeling...) had the highest positive review rates

UNDERSTANDING THE DATA DEEPER

CTD.

PRICING INSIGHTS

- **High-priced games (over \$60):**
 - Were not more positively received compared to lower-priced games
 - Lead to increased expectations
 - Must be justified by good content and execution



POPULARITY VS QUALITY GAP

- **High Concurrent Users (CCU) do not equate to high positive review sentiment**
 - Developers should not target sales and high user count as the only metrics for success
 - Quality can thrive in niches that are invisible through CCU and sales





PLAYER SENTIMENT OVER TIME

METHOD

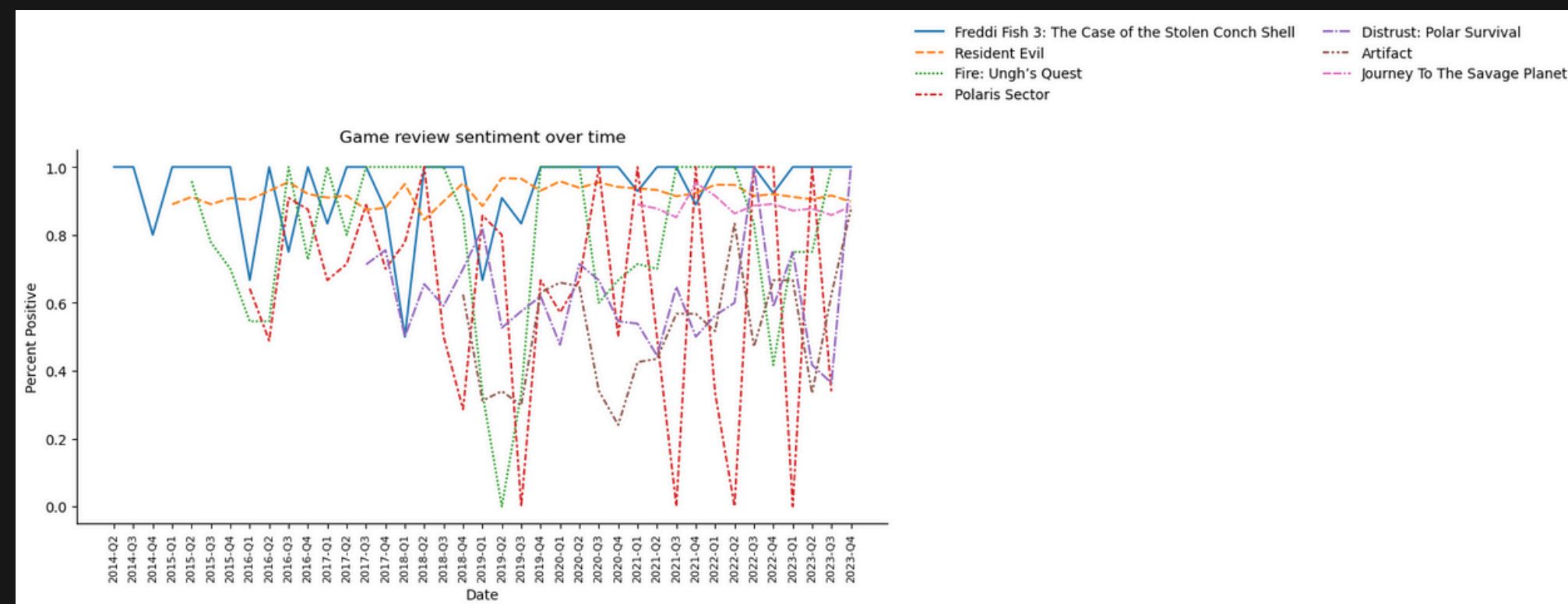
We took a small sample of games that had been released at least 1 year prior to the dataset collection date and aggregated reviews by quarter

RESULT

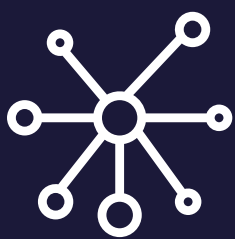
Games in our sample showed volatility and no concrete trend given the data we had

TAKEAWAY

Developers can use similar time-based analysis to track review sentiment and overlay it with key events, e.g. major updates, game news, etc. By including key events, time-based analysis would show audience reaction to these changes



MACHINE LEARNING



K-MEANS CLUSTERING

CLUSTER 0

Widely played, expensive mainstream games with very high positive review rates and massive review volume, reflecting broad popularity.

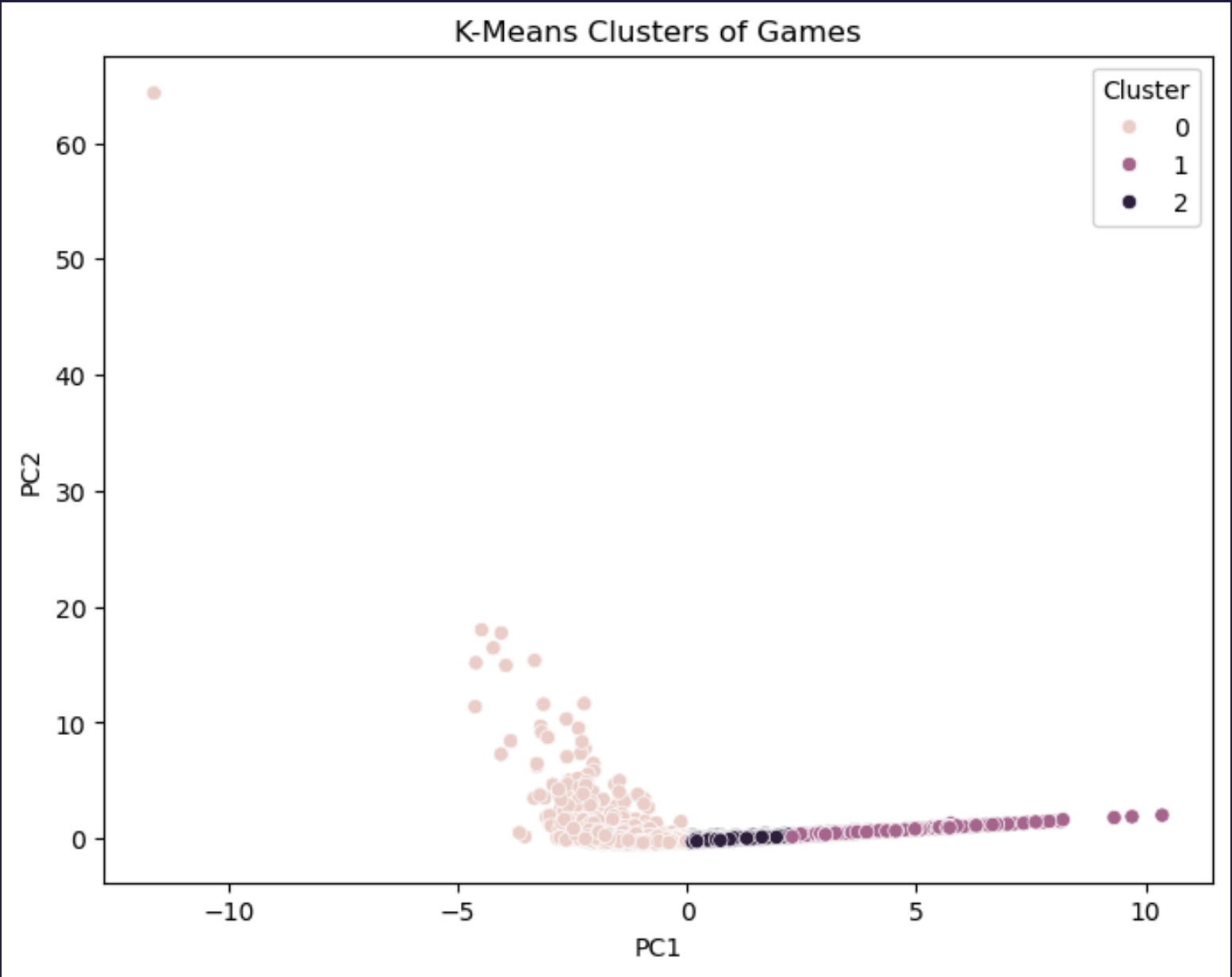
CLUSTER 1

Low-priced, lower-rated niche titles that attract high-engagement “power users” who leave mixed feedback.

CLUSTER 2

These are moderately priced, moderately rated mid-tier games with steady but not breakout engagement across all metrics.

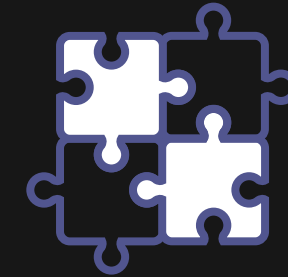
| Cluster | Number of Games | % of Positive reviews | Avg. Number of Reviews | Avg. Price |
|---------|-----------------|-----------------------|------------------------|------------|
| 0 | 5577 | 84.08% | 7968 | \$15.63 |
| 1 | 704 | 50.51% | 89 | \$2.94 |
| 2 | 2818 | 67.71% | 451 | \$4.99 |





REVIEWS FOR DEVELOPMENT

Integrating Insights with Development



1

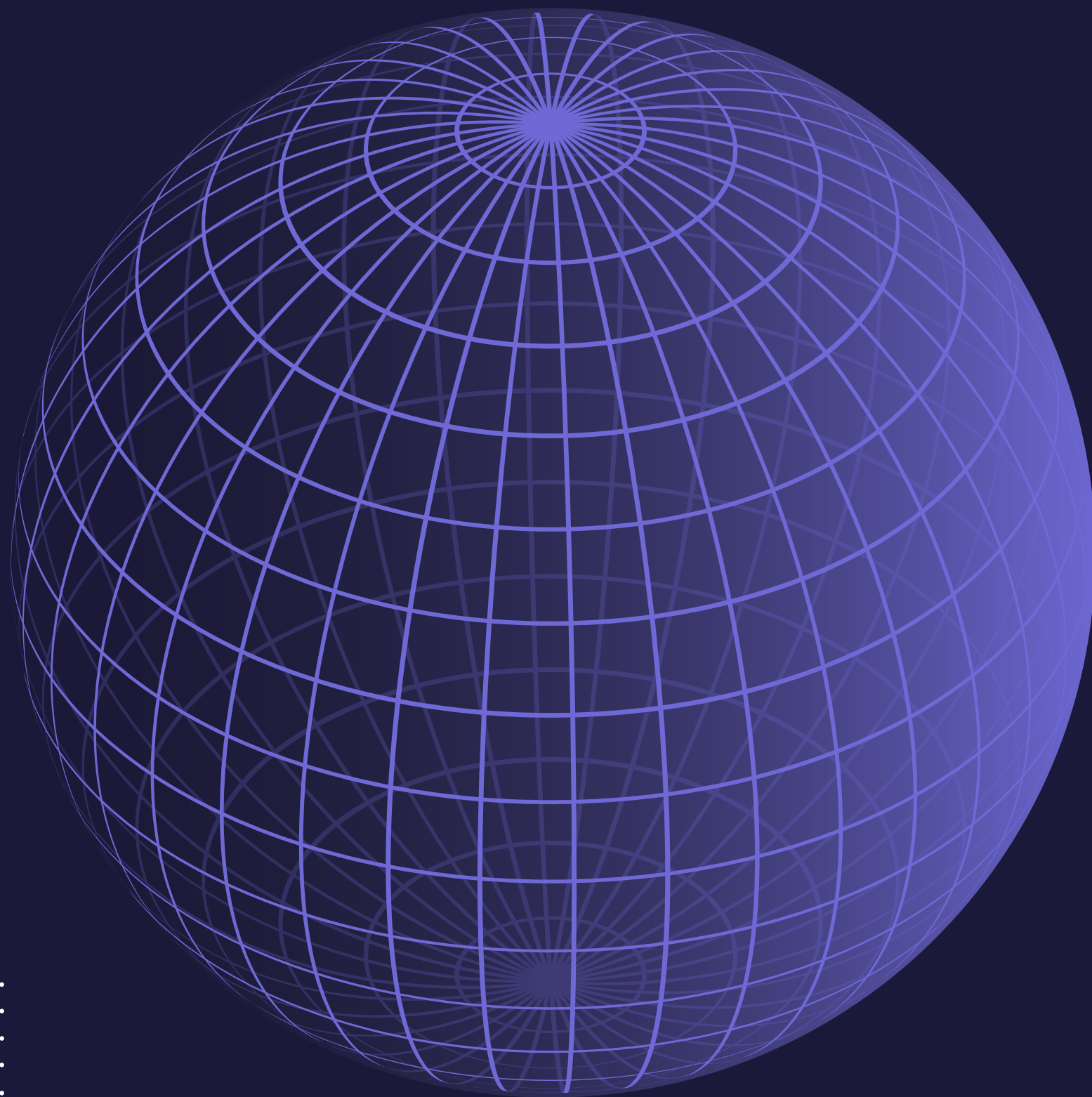
Success is **not purely financial**, nor is it purely driven by **popularity**

2

Expectations surrounding the chosen **game genre(s)** and the **quality of the content** matter more than other factors (e.g. price)

3

Listening to **feedback and tracking sentiment over time** in response to game changes allow **developers** to deliver on expectations



THANK YOU!

