

**Repte 11b:**  
**Què fa que els jugadors es quedin?**

*Anàlisi de Retenció a Steam*



Ilias Dahchouri

Neil Pradas

Albert Guillaumet

Adrià Muro

David Morillo

## 1. Introducció: La Importància de la Retenció a Steam

Steam és la plataforma de distribució digital més gran per a jocs en PC, amb una base d'usuaris massiva i diversa. La retenció de jugadors és un dels indicadors més importants per mesurar la salut dels jocs i la plataforma en general.

Mantenir els jugadors actius és clau per a l'èxit continuu tant per als desenvolupadors de videojocs com per a Steam. La retenció de jugadors es pot mesurar en funció de diverses mètriques, com el temps de joc, la freqüència amb què els jugadors tornen a jugar i la durada total de les sessions de joc. Definir correctament aquests paràmetres és el primer pas cap a una comprensió profunda de com les característiques dels jocs afecten el comportament dels jugadors.



Una alta retenció no només beneficia els desenvolupadors, que obtenen ingressos continus mitjançant contingut descarregable (DLC), microtransaccions i expansions, sinó també a Steam, que pot atreure nous usuaris i mantenir la plataforma com una opció atractiva per als jugadors.

## 2. Definició del Problema: Què Determina la Retenció a Steam?

El projecte pretén investigar quines característiques dels jocs influeixen en la retenció de jugadors a Steam, com ara el gènere del joc, el preu, les ressenyes dels usuaris i altres elements de disseny.

La retenció de jugadors pot estar relacionada amb diversos factors, com la jugabilitat, la narrativa, els gràfics, el rendiment tècnic, les actualitzacions del joc i la possibilitat de jugar en línia o amb altres jugadors.

Les dades disponibles inclouen descripcions de jocs, puntuacions d'usuaris, temps de joc, assoliments, preus i informació sobre els desenvolupadors. Així, la qualitat d'un joc, així com la seva capacitat per mantenir l'interès dels jugadors a llarg termini, es converteixen en elements fonamentals per analitzar la retenció.

### 3. Significança del Problema: Per Què Importa Entendre la Retenció?

Comprendre què determina la retenció de jugadors és fonamental per a diversos aspectes de la indústria del videojoc:

- **Desenvolupament de jocs:** Identificar quines característiques són més efectives per mantenir els jugadors interessats permet als desenvolupadors dissenyar jocs més atractius i duradors.
- **Màrqueting i vendes:** Entendre els factors que milloren la retenció ajuda a optimitzar les campanyes publicitàries i estratègies de preus.
- **Plataforma Steam:** Millorar els algorismes de recomanació de Steam per destacar jocs amb bona retenció, oferint una millor experiència als usuaris.
- **Indústria del videojoc:** Obtenir coneixements sobre les tendències de mercat i comparar el rendiment dels jocs permet als desenvolupadors i editors fer ajustos estratègics per millorar les vendes.

### 4. Fonts de Dades Disponibles i la seva Rellevància per a l'Anàlisi de la Retenció

Per analitzar la retenció, es poden utilitzar diversos conjunts de dades que proporcionen informació valuosa sobre les característiques dels jocs, la seva popularitat i el comportament dels jugadors:

- **Steam Games Complete Dataset:** Aquest conjunt de dades conté informació detallada sobre molts jocs, com el gènere, el nom, les descripcions, les ressenyes, les dates de llançament i els preus. És útil per analitzar com les característiques d'un joc poden influir en la seva retenció.

<https://www.kaggle.com/datasets/trolukovich/steam-games-complete-dataset>

- **Steam Store 2024 Hot Picks and Reviews:** Conté informació sobre els jocs més populars el 2024, incloent preus, ressenyes i descomptes. Pot ajudar a identificar tendències actuals i factors de retenció.

<https://www.kaggle.com/datasets/kanchana1990/steam-store-2024-hot-picks-and-reviews>

- **Top 1500 Games on Steam by Revenue 09-09-2024:** Proporciona dades sobre els jocs amb més ingressos fins al setembre de 2024. Analitzar el temps de joc mitjà pot proporcionar indicis sobre el comportament dels jugadors.

<https://www.kaggle.com/datasets/alicektopcu/top-1500-games-on-steam-by-revenue-09-09-2024>

- **Steam Player Data:** Aquest conjunt de dades mostra el nombre mitjà de jugadors mensuals per als jocs més populars, oferint informació directa sobre la retenció al llarg del temps.

<https://www.kaggle.com/datasets/jackogozaly/steam-player-data>

## 5. Anàlisi DAFO

Fortaleses	Oportunitats
Dades a Kaggle: Disponibilitat de datasets públics amb variables importants per a l'anàlisi.	Millorar disseny de jocs: Els resultats poden ajudar a crear experiències més atractives.
Dades directes d'activitat: Informació precisa sobre el comportament real dels jugadors.	Recomanacions a empreses: Aplicació pràctica per a companyies sense estratègies clares de retenció.
SteamDB com a complement: Font fiable que permet validar i ampliar les dades disponibles.	

Febleses	Amenaces
Definició complexa de la retenció: La retenció pot mesurar-se de diverses maneres (dies jugats, sessions, etc.), cosa que dificulta la seva estandardització.	Qualitat de les dades limitada: Possibles biaixos o dades incompletes que afectin els resultats.
Correlació no implica causalitat: Les dades poden mostrar relacions estadístiques que no necessàriament indiquen causes reals.	Inconsistències: Diferències en com es registren o es reporten les dades.
Factors externs: Elements com esdeveniments mundials o competència poden afectar la retenció sense estar relacionats directament amb el joc.	Definició variable de retenció: Distintes empreses poden utilitzar mètriques diferents.
	Factors absents: Variables importants que no estan incloses en els datasets disponibles.

## 6. Model de Dades

Per al disseny del model de dades orientat al dashboard d'anàlisi de la retenció de jugadors a Steam, estructurarem la informació en diverses taules interconnectades.

La taula principal contindrà dades dels jugadors, incloent-hi l'identificador d'usuari, el temps total jugat, la freqüència de sessions i la data de la primera i última activitat.

Es vincularà amb una taula de jocs, que inclourà atributs com el gènere, el model de monetització, la puntuació mitjana de les ressenyes i les actualitzacions rebudes.

A més, incorporarem una taula d'esdeveniments, on es registraran accions específiques com compres de DLC, participació en esdeveniments especials i interaccions amb la comunitat. Tanmateix, dades relacionades amb el mercat de videojocs (tendències, números de ventes...) es poden incloure a la taula d'esdeveniments.



Aquest enfocament permetrà generar visualitzacions detallades sobre les **tendències del mercat** i l'impacte de diferents factors en la **retenció**, com la influència del **gènere del joc**, l'efecte de les **ofertes** o el paper de les **ressenyes** dels usuaris.

La normalització del model garantirà una millor eficiència en la consulta de dades, mentre que l'ús d'índexs i agregacions optimitzarà el rendiment en l'explotació de la informació.

Així, el dashboard proporcionarà una visió clara i accionable per comprendre millor els patrons de retenció i millorar les estratègies dels desenvolupadors.

## 7. Conclusions

- **Importància de la Retenció de Jugadors:** La retenció de jugadors és un dels indicadors més crítics per mesurar l'èxit d'un joc i la salut de la plataforma Steam en general. Com més temps els jugadors continuen jugant, més probable és que els desenvolupadors obtinguin ingressos continus mitjançant DLC, expansions i microtransaccions.
- **Factores Multitudinaris:** La retenció pot dependre de diversos factors, incloent-hi la jugabilitat, la narrativa, les opcions multijugador, les actualitzacions del joc i els preus. És essencial identificar quins d'aquests factors tenen un impacte més gran en la retenció per tal de dissenyar estratègies d'optimització.
- **Dades Clau per a l'Anàlisi:** Els conjunts de dades disponibles, com el Steam Games Complete Dataset i Steam Player Data, ofereixen una gran quantitat d'informació que pot ser útil per investigar la relació entre les característiques del joc i el comportament dels jugadors. La recopilació d'aquestes dades és fonamental per aprofundir en la comprensió dels patrons de retenció.
- **Reptes en la Mesura de la Retenció:** La retenció es pot mesurar de diverses maneres (temps de joc, freqüència de sessions, nombre de dies jugats), el que pot portar a resultats contradictoris si no es defineixen clarament les mètriques. Això pot complicar l'anàlisi i la comparació dels resultats entre diferents jocs.
- **Impacte dels Factors Externs:** A l'hora d'analitzar la retenció, cal tenir en compte factors externs com la competència, les tendències del mercat o esdeveniments globals. Aquests poden influir en el comportament dels jugadors sense que estiguin directament relacionats amb el disseny del joc.
- **Teòrica Aplicabilitat del Model de Dades:** Tot i que el model de dades proposat no s'ha implementat encara, es pot preveure que un sistema ben estructurat pot proporcionar informació valuosa sobre com les característiques d'un joc afecten la retenció. La creació de taules interconnectades permetria analitzar les dades de manera més eficient i obtenir resultats accionables per a desenvolupar millores estratègiques.
- **Aplicació dels Resultats en el Disseny de Jocs:** A mesura que s'analitzen les dades sobre retenció, els resultats podrien ajudar a orientar els desenvolupadors en la creació de jocs més atractius i duradors. La identificació dels factors clau que retenen els jugadors ajudaria a millorar el disseny, el màrqueting i les estratègies de monetització de futurs jocs.