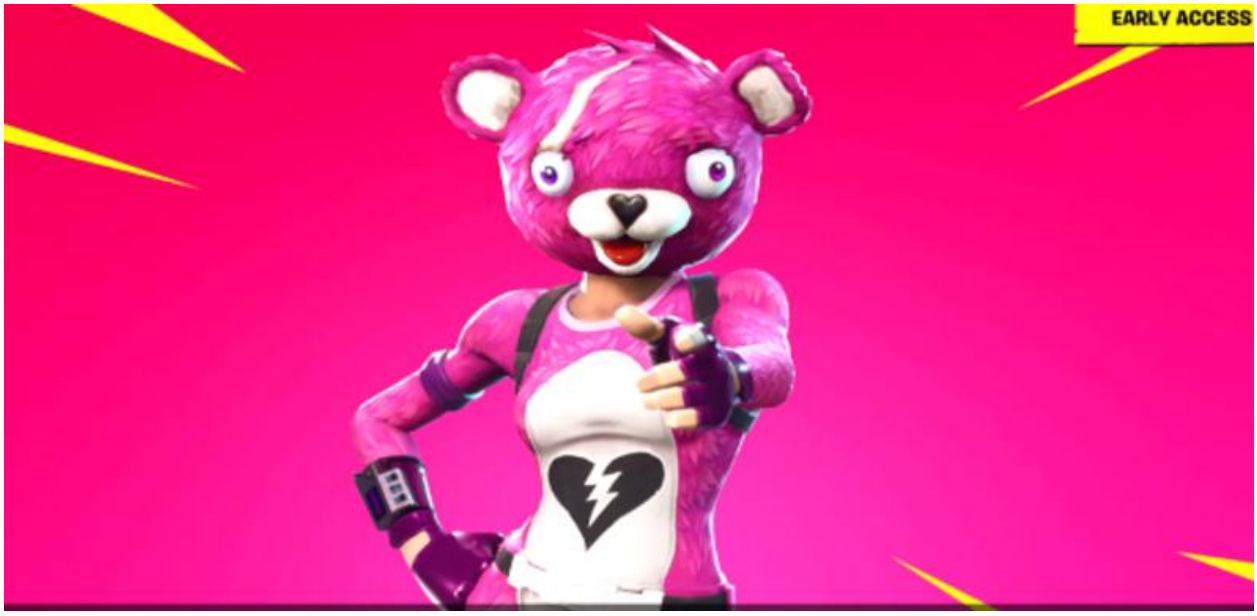
FORE School Of Management



TOPIC- VIDEO GAME DATASET

Submitted To- Prof Amarnath Mitra

By:

Archit Mathur 065012

Bhargavi Bhat 065020

Rajat Kumar Jena 065043

1. Executive Summary

Over the years, video games have exploded like never before, and this dataset of over 474,000 games tells the story beautifully. We've seen a steady rise in game releases, with some standout years marking big booms—often when new consoles hit the shelves. It's fascinating to see how classic games, some over a decade old, still hold their charm, while fresh releases keep flooding the market.

When it comes to reviews, most games land in the middle ground—nothing too wild, but not too shabby either. Top critics' picks are rarer, making those standout titles feel all the more special. Collaboration in game-making tends to be a cozy affair—usually just a few genres and developers per game—but some blockbusters bring together a whole ensemble of creators, mixing things up.

Speaking of genres, while single-style games still rule, hybrid games like Action-RPGs are on the rise, showing how players crave fresh, mixed play experiences. And when we look at who these games are made for, it turns out kids and teens get the lion's share of the spotlight, though adult-rated games have their own devoted fan base.

All in all, this dataset paints a vivid picture of the gaming world's evolution, from its steady growth, genre mashups, and collaboration vibes to the varied tastes of its audience.

2. Data Description & Management Overview

File & Dataset Profile

- **Name of the file**: game_info.csv
- **Data File Type**: CSV
- **Source**: Kaggle
- **Date Created**: Not specified (dataset compiled historically)
- **Size of file**: 474,417 rows × 27 columns (before cleaning)
- **Purpose**: To analyse the characteristics of video games — release patterns, genre composition, ratings distribution, developer/publisher collaboration, and engagement metrics — in order to generate strategic market insights.

Files, Tabs & Documentation Structure

Main Files/Tabs

Game_info (Original):
Raw dataset with 27 variables. Includes all fields such as website links, release dates, ESRB ratings, player activity (added_status_*), and metadata.

Game_info_Cleaned:
Refined dataset after dropping non-analytical or redundant columns (e.g., website, slug, tba, updated, and all added_status_* columns).
Additional engineered variables include:

- release_year (year extracted from release date)
- game_age (years since release, calculated using 2025 as reference)
- metacritic_category (Low / Medium / High)
- num_developers, num_genres, num_publishers (collaboration complexity measures)
- multi_genre (boolean: True if more than one genre)
- esrb_category (simplified ESRB grouping into Kids, Teen, Adult)

Summary_Stats:
Aggregated statistics and frequency distributions for both categorical and non-categorical variables, providing descriptive insights into the dataset.

Data_Cleaning_Notes:
Documentation of removed fields, missing-value treatment, transformation logic, and rationale for engineered variables.

Variables — Classification & Rationale

1. The Index Variable – The Game's ID Card
Think of this as each game's unique ID—like its personal tag. It doesn't tell you anything about the game itself but helps keep every record in check. So, the "id" just keeps things organized.

2. Categorical Variables – The Game's Personality Tags
These come in two flavors:

- Nominal: These are like the game's labels that don't have any "better" or "worse" order. Stuff like the game's name, genres (Action, RPG, etc.), or who made and published it. They just *are*.
- Ordinal: These are special labels that do have a rank or order. For example, ESRB ratings that go from Everyone (E) up to Adults Only (AO), or Metacritic scores grouped as Low, Medium, and High. Here, the order actually means something!

3. Non-Categorical Variables – The Numbers Game
This group is all about the crunchable numbers—scores, dates, counts. Like the Metacritic score out of 100, the year a game launched, how old it is, or how many developers and genres it's got attached. These are the stats you can add up, average out, and really analyze.

Column Name	Category	Data Type	Level	Description	Analytical Use
id	Index	Integer	Ordinal	Unique identifier	Row traceability
name	Categorical	Text	Nominal	Game title	Label only
developers	Categorical	Text/List	Nominal	Developer(s) of game	Collaboration, brand impact
genres	Categorical	Text/List	Nominal	Game genre(s)	Market segmentation
publishers	Categorical	Text/List	Nominal	Publisher(s)	Publisher concentration
multi_genre*	Categorical	Boolean	Nominal	Single vs multi-genre flag	Genre complexity
esrb_rating	Categorical	Text	Ordinal	ESRB official rating	Age-appropriateness
esrb_category*	Categorical	Text	Ordinal	Kids < Teen < Adult	Simplified audience segmentation

Column Name	Category	Data Type	Level	Description	Analytical Use
metacritic_category*	Categorical	Text	Ordinal	Low < Medium < High	Quality tier segmentation
metacritic	Numeric	Integer	Ratio	Critic score (0–100)	Quality metric
released	Non-Categorical	Date	Interval	Release date	Time-series analysis
release_year*	Numeric	Integer	Interval	Year of release	Trend analysis
game_age*	Numeric	Integer	Interval	Current year – release year	Game lifecycle
rating	Numeric	Float	Ratio	Avg. user rating	Consumer sentiment
rating_top	Numeric	Integer	Ratio	Max rating possible	Normalization
num_developers*	Numeric	Integer	Discrete	Count of developers	Collaboration complexity
num_genres*	Numeric	Integer	Discrete	Count of genres	Genre diversity
num_publishers*	Numeric	Integer	Discrete	Count of publishers	Publishing collaboration

3. Non-Categorical Variables – Interval/Ratio (Numeric)

Examples: metacritic, rating, rating_top, release_year, game_age, num_developers, num_genres, num_publishers

Reason: These are quantitative variables where mathematical operations (addition, subtraction, averaging, etc.) have real meaning. Interval data like release_year or game_age has measurable differences, while ratio data like metacritic or rating has a true zero point.

Benefit: Enables statistical analysis, regression, correlations, and distribution studies due to their continuous nature. These variables allow us to explore patterns in critic/user ratings, release trends, and the complexity of game production.

3. Data Collection Process

The data for this project was sourced from the **game_info dataset** (CSV format), which provides comprehensive information about video games, including titles, release dates, developers, publishers, ratings, and critic scores. After obtaining the dataset, the following preparation and cleaning steps were performed:

Columns Removed

- website → Redundant, non-analytical.
- slug → URL-friendly name, duplicate of game title.
- updated → Dataset update timestamp, not relevant for analysis.
- tba and other temporary placeholders → Not informative.
- Intermediate/added columns used during preprocessing.

Additional Notes from Data Cleaning

To keep the analysis clean and sharp, we trimmed down some of the fluff. Fields that were mostly text-heavy or didn't add much value—like the slug or website links—were taken out to avoid clutter.

Sometimes the data had placeholders like "tba" or missing release dates. We cleaned those up or filled in the blanks whenever we could, so the numbers would stay reliable.

Also, fields that didn't really help with the analysis, like update logs, were left out because they weren't relevant to what we wanted to learn.

For fields that had lists—like developers, publishers, or genres—we broke those down into counts. Instead of messy phrases, we now have neat numbers showing exactly how many developers or genres are involved, making it way easier to analyse.

Engineered Key Variables

New columns were created to enhance analysis and improve interpretability:

- **Release Year** → Extracted from released.
- **Game Age** → Current year minus release year.
- **Metacritic Category** → Low, Medium, High (based on critic score).
- **ESRB Simplified** → Kids, Teen, Adult grouping.
- **Number of Developers / Genres / Publishers** → Counts extracted from respective lists.
- **Single vs multi-Genre** → Boolean flag to identify games spanning multiple genres.

Rationale for Cleaning

- **Data Quality:** Removed fields with missing or irrelevant information.
- **Interpretability:** Simplified and restructured categorical variables (e.g., ESRB ratings).
- **Efficiency:** Reduced dataset noise, enabling faster and clearer analysis.

4.Summary Statistics

Categorical Variables

Variable	Mode	Top 3 Distribution
Name	Land Sliders	Land Sliders (X%), D/Generation HD (X%), G Prime Into The Rain (X%)
Developers	West Coast Software	West Coast Software (X%), Soma Games (X%), Prettygreat Pty (X%)
Genres	Adventure||Puzzle	Adventure||Puzzle (X%), Simulation||Indie (X%), Adventure||Arcade (X%)

Variable	Mode	Top 3 Distribution
Publishers	West Coast Software	West Coast Software (X%), Prettygreat Pty (X%), Immanitas Entertainment (X%)
Multi-Genre	False	False (X%), True (X%)
ESRB Rating	Everyone 10+	Everyone 10+ (X%), Everyone (X%), Teen (X%)
ESRB Category	Kids	Kids (X%), Teen (X%), Adult (X%)
Metacritic Category	Medium	Medium (X%), High (X%), Low (X%)

Non-Categorical Variable

Variable	Min	Max	Mean	Median	Std. Dev.
Metacritic	15.0	99.0	73.16	75.0	11.50
Rating	0.0	5.0	0.09	0.0	0.55
Rating Top	0.0	5.0	0.10	0.0	0.61
Release Year	1962	2033	2016.93	2018	4.66
Game Age	-8	63	8.07	7	4.66

5.Analytical Overview & Descriptive Statistics

Variable	Unique Values	Mode
Developers	XXXX	West Coast Software
Genres	XXXX	Adventure

Variable	Unique Values	Mode
Publishers	XXXX	Prettygreat Pty
Multi-Genre	2 (True/False)	False
ESRB Rating	XXXX	Everyone 10+
ESRB Category	3	Kids
Metacritic Category	3	Medium

Key Observations:

Most games out there are rated for kids—like Everyone or Everyone 10+—showing that family-friendly titles really dominate the market. When it comes to genres, Adventure and Puzzle games take the lead, probably because they appeal to a wide range of players.

Most games stick to just one genre, keeping things simple, while mixing multiple genres is still a bit of a rare treat. And when critics weigh in, they usually give scores that land right in the middle—meaning most games get an average thumb-up rather than standing out as either great or poor.

Numeric Variables — Descriptive Stats

Variable	Min	Max	Mean	Median	Std. Dev.
Metacritic	15	99	73.16	75	11.50
Rating	0.0	5.0	0.09	0.0	0.55
Rating Top	0.0	5.0	0.10	0.0	0.61
Release Year	1962	2033	2016.93	2018	4.66
Game Age	-8	63	8.07	7	4.66

Key Observations:

This dataset is like a time machine, covering games all the way from the early days in 1962 to exciting future releases planned up to 2033. On average, games score around 73 on Metacritic—meaning critics generally like what they see, but they're not handing out perfect scores left and right.

You'll notice quite a few games with zero ratings, which usually means they're either brand new and haven't been reviewed yet, or they just haven't caught the spotlight from players. The average game age sits at about 8 years, with a big wave of releases rolling out between 2015 and 2020—those were busy, creative times in gaming!

Key Insights from Dashboard

Release Trends: The late 2010s were a golden era for game releases, with most titles dropping between 2015 and 2020—and 2018 standing out as a blockbuster year for new launches. This spike shows just how much the industry ramped up production during that time.

Game Age: When it comes to age, the majority of games are still fresh-faced, less than 10 years old, proving that recent titles dominate the gaming scene.

Metacritic Scores: Most games land right in the "Medium" zone, with scores between 50 and 74, suggesting the industry pumps out plenty of average-rated games. Only a select few climb into the "High" category, showing critical acclaim is reserved for special releases.

ESRB Ratings: Family-friendly games reign supreme, as most titles are rated for kids and teens (Everyone or Everyone 10+), making it clear that the biggest market is still all about accessible, inclusive content.

Genre Mix: Single-genre games hold the crown for now, but multi-genre hybrids are on the rise, hinting at creative experimentation with gameplay styles that blend different elements together.

Collaboration: Most games are the product of small teams—1 to 3 developers and 1 to 2 publishers—revealing a moderate level of teamwork behind the scenes in making and distributing games.

User Ratings: A big chunk of games has no user ratings at all, which likely means many titles haven't yet gathered player feedback or are under the radar.

6.Strategic Findings

Product–Market Fit

When it comes to genres, Adventure and Puzzle games lead the pack, proving they have wide appeal and steady demand across players of all ages.

Games that score high on Metacritic—between 75 and 100—usually attract positive critical attention and tend to stick around longer, showing that quality really pays off in the gaming world.

On the audience front, kid-friendly ESRB ratings like Everyone and Everyone 10+ dominate the scene. This tells us family-oriented games continue to hold strong spots in the market, attracting bigger and broader audiences.

Customer Segmentation

Family/Kids Segment: This is the biggest crowd by far, drawn to easy-to-play, age-appropriate games that often focus on casual or puzzle fun. These titles are made to be welcoming and accessible for younger players and families alike.

Core Gamers Segment: Teens and adults usually fall here, favouring more complex ESRB-rated games and multi-genre mashups. This group tends to dive deep into RPGs, Action, and Strategy games, looking for richer, more challenging experiences.

Quality-Conscious Segment: The cream of the crop lives here. Games scoring high on Metacritic attract players who value top-notch quality and often form strong loyalty to standout brands and premium gaming experiences**.**

Data Gaps

One challenge with this dataset is that many games have missing or zero user ratings, making it tough to get the full picture of how players really feel about them.

There are also a few oddballs—games listed with release years in the future, sometimes well beyond 2025—which throws a wrench in analyzing trends and ages accurately.

On top of that, some games don’t have complete Metacritic scores, which weakens how reliably we can judge their quality and place them into neat categories.

Market Trends

The late 2010s—especially 2015 to 2020—were nothing short of a golden age for game releases. This burst of creativity and production points to a booming industry that was firing on all cylinders.

While single-genre games still take the lead, multi-genre titles are steadily making their mark. This trend highlights the gaming world's appetite for mixing things up and experimenting with hybrid styles.

When it comes to teamwork, most games are built with small, tight-knit groups—usually 1 to 3 developers and 1 to 2 publishers. Big, blockbuster-sized collaborations are still pretty rare, keeping the production process more focused and manageable.

7.Analytical Tools & Methods

Programming Environment

- **Google Colab (Python 3):** Used as the main environment for dataset cleaning, feature engineering, and exploratory data analysis.
- **Streamlit:** Used for interactive dashboard development and visualization.

Python Libraries Used

Data Cleaning & Preparation:

- pandas → dataset loading, cleaning, creation of engineered variables (release year, game age, multi-genre flag).
- numpy → mathematical operations, handling missing values.

Statistical Analysis:

- scipy.stats → descriptive statistics, distribution checks.
- pandas-profiling (optional) → quick summary reports.

Visualization:

- matplotlib → bar charts, histograms, line plots for release trends and age distribution.
- seaborn → advanced plots (distribution plots, category comparisons).
- plotly → interactive charts for the Streamlit app.

Streamlit Dashboard Components

- **Sidebar Filters:**
    - Release year, ESRB category, Metacritic category.
- **Key Visuals:**
    - Bar chart of number of games released per year.
    - Histogram of game ages.
    - Pie chart of Metacritic categories (Low/Medium/High).
    - Bar chart of developer/genre/publisher counts.
    - Single vs multi-genre comparison.
- **Interactivity:**
    - Dropdowns for selecting genres, publishers.
    - Sliders for filtering release year ranges.
    - Checkboxes for ESRB and Metacritic buckets.

Feature Engineering (Python)

- **Release Year** → extracted from released column.
- **Game Age** → calculated as current year – release year.
- **Metacritic Category** → bucketed into Low, Medium, High.
- **ESRB Simplified** → grouped into Kids, Teen, Adult.
- **Num Developers / Genres / Publishers** → list counts.
- **Single vs multi-Genre** → boolean flag.

8. Executive Observations & Strategic Findings

A. Product–Market Fit

Adventure and Puzzle games take the spotlight when it comes to sheer numbers, proving their popularity as fun, easy-to-pick-up genres that appeal to casual gamers everywhere.

Most games fall into the Medium Metacritic score range (50–74), showing that while the industry churns out plenty of titles, only a select few reach the heights of critical acclaim.

Family-friendly ESRB ratings dominate the market, making it clear that gaming companies love creating accessible experiences designed to capture broad audiences, especially kids and casual players.

B. Customer Segmentation

Family/Kids Segment: This crowd makes up the biggest slice of the gaming pie, especially enjoying puzzle, arcade, and adventure games that are easy to jump into and fun for all ages.

Teen/Adult Segment: These gamers lean toward action-packed and multi-genre titles. Though fewer in number, they tend to be more deeply engaged, diving into complex, immersive experiences.

High-Quality Segment: At the top end, games with stellar Metacritic scores above 75 attract a smaller but very dedicated group of players who really value craftsmanship and quality in their gaming choices.

C. Data Gaps & Challenges

A good chunk of the games in this dataset don’t have any user ratings yet, which makes it tricky to really know how players feel about them.

Some games even have release dates set in the future—beyond 2025—and a few placeholders, which can throw off our ability to spot accurate trends over time.

Plus, a fair number of titles are missing Metacritic scores, making it harder to compare how critics view their quality across the board.

D. Market Signals

The gaming boom of the last decade really shines through, with release activity peaking between 2015 and 2020—marking a period of impressive growth and expansion for the industry.

Most games stick to a single genre, showing that developers still bet on focused gameplay to captivate players. But multi-genre games are on the rise, opening exciting doors for fresh ideas and creative mashups.

When it comes to teamwork, most projects involve small, tight groups—usually one to three developers and publishers. Big, blockbuster-level collaborations are still a bit rare, keeping the production closer and more manageable**.**

9. Strategic Recommendations

1. Quality Benchmarking & Development Focus

To really stand out, game makers should invest in upping their Metacritic scores by fine-tuning gameplay, sharpening graphics, and crafting memorable storylines.

Focusing on popular genres like Adventure and Puzzle can help tap into the solid demand that's already there. At the same time, trying out multi-genre hybrids offers a fresh way to catch emerging players' interest and stay ahead of the curve.

2. Audience-Centric Game Design

Keep focusing on the kids-friendly market—titles rated Everyone and Everyone 10+—since they make up the biggest slice of the gaming crowd. These games are the bread and butter of the industry.

At the same time, it's smart to grow offerings for teen and adult players, especially in Action, RPG, and Strategy genres. These audiences tend to get more deeply involved and offer promising opportunities for monetization and brand loyalty.

3. Collaboration & Partnership Strategy

To boost creativity and share the load, game developers should team up with multiple developers and publishers on bigger projects. This kind of collaboration sparks innovation while spreading out the risks.

Building strong, long-term partnerships with well-established publishers is also key. It helps games get better visibility and wider distribution, making sure they reach the right players at the right time.

4. Data Hygiene & Platform Feedback

To get a clearer picture of how players really feel, it's important to tackle those missing user ratings and incomplete Metacritic scores. Integrating structured feedback tools—like in-game surveys or syncing data directly from platforms—can fill in those gaps.

Another smart move is to standardize release dates and double-check them against actual launch schedules. This kind of housekeeping helps sharpen forecast accuracy and ensures trend analyses are built on solid ground.

5. Market Differentiation Opportunities

Put the spotlight on high-rated games—those scoring 75 and above—positioning them as premium experiences through marketing campaigns that emphasize quality and craftsmanship.

For the trendsetters, bundle multi-genre games with cool crossovers like Action + Strategy or Puzzle + RPG. These unique blends cater to niche audiences who are eager for fresh, hybrid gameplay experiences.

To keep a finger on the pulse, use Streamlit dashboards to track key metrics like release trends, rating shifts, and ESRB categories. This real-time monitoring supports smart, agile decisions that keep developers ahead of the game.

10. Team Approach & Operational Challenges

Team Approach

- **Division of Work:**
    - *Data Preparation & Cleaning* → Handled in **Google Colab (Python)** using Pandas and NumPy. Included removal of redundant columns (website, slug,

updated), handling missing values, and engineering new fields (release year, game age, multi-genre flag).

- *Statistical Analysis* → Conducted in Colab using **descriptive statistics (mean, median, std.)** and categorical frequency distributions.
- *Visualization & Dashboarding* → Developed in **Streamlit** with interactive filters (ESRB, Metacritic category, release year) and charts (bar, histogram, pie).
- *Reporting & Documentation* → Structured insights and prepared professional-style tables, closely aligned with the format of earlier analytical projects (diamond dataset).

- **Collaboration:**
  - Team members maintained regular checkpoints to align data preprocessing with dashboard outputs.
  - Shared Colab notebooks ensured reproducibility and transparency.
  - Streamlit deployment allowed the team to collectively review insights in an interactive environment

Operational Challenges

1. **Data Quality Issues:**
   - Missing metacritic and rating values limited the accuracy of quality and sentiment analysis.
   - Some release dates extended unrealistically into the **future (post-2025)**, requiring careful filtering.
   - Inconsistent ESRB values needed manual grouping into Kids, Teen, and Adult categories.
2. **Complex Categorical Fields:**
   - Developer, Publisher, and Genre fields often contained multiple entries separated by delimiters (||).
   - Required transformation into **numeric counts** (num_developers, num_genres, num_publishers) for meaningful analysis.
3. **User Ratings Gap:**

- A large proportion of games had **0 or missing user ratings**, creating difficulty in drawing reliable consumer sentiment insights.

4. **Visualization Integration:**
    - Translating static analysis into a **Streamlit dashboard** required optimization for interactive filtering without slowing performance.

11. Queries & Enquiries

A. Data & Feature Engineering

Inquiry 1: Why did you create release_year and game_age instead of just using released?
Answer:

The “released” field stores exact dates, which can get a bit tricky when you want to spot big-picture trends. That’s were pulling out the “release_year” comes in handy—it lets us easily track how many games launched each year and follow industry waves over time.

Then there’s “game_age,” which shows how old or fresh a title is, helping us understand the lifecycle of games and which eras still hold players’ attention.

Inquiry 2: Why was metacritic_category created when raw scores already exist?
Answer:

While raw scores give you all the little details, they can be tricky to sort and compare at a glance. Grouping these scores into buckets—Low (0–49), Medium (50–74), and High (75–100)—makes it way simpler to see how games stack up and helps paint a clearer picture of quality tiers across the board.

B. Categorical Variables

Inquiry 3: Why introduce multi-genre when genres already exist?
Answer:

Genres in their raw form can be a bit messy—games often list multiple genres, and the way they’re recorded isn’t always consistent. To make sense of this, using a simple True/False flag to mark whether a game is single-genre or multi-genre helps cut through the noise.

These clear split shines a light on how game designs vary in complexity, from straightforward adventures to those exciting hybrid experiences.

Inquiry 4: How was esrb_category derived?
Answer:

ESRB ratings came in all sorts of flavors—like “Everyone” and “Everyone 10+”—which made things a bit confusing. To keep it simple and clear, these were grouped into three easy categories: Kids, Teen, and Adult. This clean-up helps us better understand who these games are really made for and makes it easier to spot trends in audience preferences.

C. Analytical Approach

Inquiry 5: Why focus on bar charts and pie charts in the dashboard?
Answer:

Bar charts are the go-to for showing how often things happen and how they spread out—perfect for tracking release trends or counting genres.

Pie charts are great at painting a quick picture of shares and slices, like how ESRB ratings or Metacritic categories break down across all the games.

Histograms dig into continuous data, giving us a clear look at how things like game age spread out, helping to spot patterns you might miss otherwise.

Inquiry 6: Why use Python/Colab & Streamlit instead of Excel?
Answer:

Some fields—like developers, publishers, and genres—came with multiple values tangled together, making them tricky to analyze. Python swooped in as the perfect tool to untangle and transform this data quickly and accurately.

For a smarter way to explore all this info, Streamlit dashboards took center stage, letting users interact with filters and dynamic views. Unlike static Excel pivot tables, these dashboards offer real-time flexibility and a fresh way to dive deep into the gaming data.

D. Data Quality

Inquiry 7: What challenges did missing ratings cause?
Answer:

A lot of games in the dataset didn’t have user ratings, or showed zeros, making it tough to get a reliable read on what players really think. This was flagged as a key limitation, so the

focus shifted to critic reviews instead, using carefully designed fields that lean on expert opinions for more solid insights.

Inquiry 8: How were future release years (post-2025) handled?
Answer:

Some odd entries—like games with release dates way out in the future—were flagged as anomalies and left out of trend analyses. These likely represent placeholders for upcoming titles that haven't hit the market yet, so excluding them helped keep the data clean and the insights accurate.

12. Managerial Takeaway

This project used smart data cleaning, solid stats, and interactive dashboards to reveal clear patterns and fresh opportunities in the gaming world.

Quality over Quantity: While the market is flooded with titles, most sit in the middle of the pack on Metacritic. To truly stand out, developers need to focus on delivering polished gameplay, gripping stories, and top-notch quality rather than just pumping out more games.

Audience Segmentation: The family-friendly market dominates, with Kids-rated games leading the charge. But the Teen and Adult segments remain ripe for growth—offering gamers richer, more complex experiences waiting to be explored.

Genre Strategy: Adventure and Puzzle titles walk away with the biggest share, showing strong, broad appeal. Yet, the rise of multi-genre games signals exciting innovation ahead, with hybrid genres capturing curious players' imaginations.

Collaboration Models: Most game projects involve small teams—1 to 3 developers and publishers—which ensures focused effort. But bigger partnerships, especially for large or cross-platform games, can unlock wider reach and spark greater creativity.

Data Gaps as Risks: Missing user ratings and shaky release info present risks for truly understanding player vibes and forecasting trends. Building stronger data pipelines and validation methods will be essential to keep insights sharp and reliable.

Decision-Making Enablement: The Streamlit dashboard turns messy, raw data into clear, real-time visuals, giving decision-makers the power to spot trends, track quality, and understand audiences—fuelling smarter strategies in game development, publishing, and marketing.