ANALYSIS OF YOUTUBE ANALYTICS DATASET

GROUP 4

Howard Muchaki - 668712

Patrick Keegan - 668964

UNITED STATES INTERNATIONAL UNIVERSITY

CLASS DSA 3050A

[**1. Business case 3**](#_30ucqc4tn404)

[1.1 Problem statement 3](#_vkl1vuflqu6)

[1.2 Objectives 4](#_gyrgjx5byyae)

[**This study aims to: 4**](#_puj312hpfr7)

[**1. Identify factors contributing to churn and develop retention strategies. 4**](#_puj312hpfr7)

[**2. Analyse engagement and revenue trends to inform content scheduling. 4**](#_puj312hpfr7)

[**3. Maximise ROI through optimised ad revenue and premium subscription growth. 4**](#_puj312hpfr7)

[**4. Develop actionable recommendations based on seasonal, temporal, and behavioural insights. 4**](#_puj312hpfr7)

[**5. Establish a roadmap for sustainable channel growth and community building. 4**](#_gvyoq9opxn25)

[**2. Data architecture 4**](#_cdmi4hoq67sy)

[**3. Key insights 6**](#_2zfy8wttbxq3)

[**4. ROI Projections 7**](#_zaoojfxdnrk5)

[**5. Recommendation 7**](#_2oidvee76bhm)

[**6. Conclusion 8**](#_4d5cvt23jtt7)

# 1. Business case

A popular YouTuber approached our team seeking guidance on effectively managing their channel. Their primary goals were to minimise subscriber loss, enhance viewer retention, and maximise their return on investment (ROI) through both ad revenue and premium subscriptions. To achieve these objectives, we discussed strategies such as optimising content for engagement, analysing viewer demographics to tailor uploads, and employing effective marketing techniques to promote exclusive content for premium members. By focusing on fostering a loyal community and delivering consistent value, the YouTuber aims to build a sustainable and profitable channel while creating meaningful connections with their audience.

## 1.1 Problem statement

YouTube channels experience fluctuations in subscriber counts based on various operational strategies and content management approaches. Understanding the factors that contribute to unsubscriptions is crucial for content creators, as it impacts both audience retention and revenue generation. This process involves analysing specific elements such as video quality, consistency of uploads, viewer interaction, and the relevance of content to the target audience. Additionally, it's important to evaluate the effects of unsubscriptions on overall revenue streams, including ad revenue, sponsorship opportunities, and merchandise sales. A comprehensive understanding of viewer engagement metrics—such as watch time, likes, comments, and shares—can provide valuable insights into subscriber behaviour and preferences, ultimately guiding creators in optimising their channels to foster loyalty and growth.

## 1.2 Objectives

# This study aims to:

# Identify factors contributing to churn and develop retention strategies.

# Analyse engagement and revenue trends to inform content scheduling.

# Maximise ROI through optimised ad revenue and premium subscription growth.

# Develop actionable recommendations based on seasonal, temporal, and behavioural insights.

# Establish a roadmap for sustainable channel growth and community building.

# 2. Data architecture

The provided data is organised in a star schema, a widely used data architecture for data warehousing and business intelligence. This structure is ideal for creating dashboards and reports because it simplifies the querying process and improves performance. The architecture consists of a central fact table and multiple dimension tables.

* **Fact Table:** “**fact\_video\_performance.csv”** serves as the central fact table. It contains the core metrics (facts) of interest for video performance, such as Views, Likes, Shares, Watch Time, and Estimated Revenue. Each row in this table corresponds to a single video's performance metrics.
* **Dimension Tables:** These tables provide context to the facts. The provided data includes the following dimension tables:
  + **dim\_channel.csv**: A very small table that provides details about the YouTube channel itself, such as Channel\_ID, Channel\_Name, Country, and Subscribers\_Total.
  + **dim\_date.csv**: A table that breaks down dates into various components like Day, Month, Quarter, and Year. This is crucial for time-series analysis.
  + **dim\_video.csv**: This table contains descriptive attributes for each video, such as Video Duration, Video Publish Time, Month, Year, and Video Thumbnail CTR (%).

The tables have been constructed after performing data mining using the Apriori Algorithm, which further utilises association rules(in-built function) that extract the frequent itemsets in a dataset. This gives us all the impact each variable has on one another in our dataset, including the interpretation of what makes a YouTube Channel or Video good for the content creators and finding the target audience to then be aware of and reduce some of the concerns our client had. From viewership to subscriber retention.

Further, we will look at the star schema and find what has been of help to us

# 3. Key insights

The day of the week that exhibits the highest levels of engagement, measured by likes, shares, and comments, is Tuesday. Analysis indicates that two seasons, specifically spring and winter, experience minimal churn, characterised by a churn probability of less than 0.3. Conversely, the remaining two seasons demonstrate a moderate churn probability, ranging from greater than 0.5 to not exceeding 0.8. Channel growth experienced an upward trajectory over the first three years, reaching its zenith in 2018, while the retention rate remained stable from the onset of the COVID-19 pandemic until 2024.

Factors influencing churn include the publishing season, with summer and fall showing the lowest retention rates. Additionally, the day of the week plays a significant role, as both Tuesday and Sunday are associated with heightened interaction with content. This increased engagement subsequently leads to a high revenue per 1,000 views, with Tuesday generating revenue of $8.04 and YouTube ad revenue amounting to $1,150. Furthermore, August is identified as the month with the least churn, attributed to the majority of school seasons being in recess during this period.

In terms of revenue, Tuesdays and Fridays record the highest average revenue, with the Fall season generating the most revenue overall. From our analysis, the YouTuber can maximise earnings by prioritising posts on Tuesdays and Fridays and focusing content around the Fall and Summer seasons, while using the Winter season for strategic planning or upskilling.

Additionally, ad impressions peak within the first three days of publishing. To capitalise on this, the YouTuber should consider investing in sponsored video visibility during this period and employ strategies such as teasers and countdowns before publishing. These actions can help maximise views and, consequently, ad impressions, further driving revenue growth.

# 4. ROI Projections

Using the provided data and recommended strategies:

* Posting on high-engagement days (Tuesdays & Fridays): Expected 15–20% uplift in RPM, increasing from $8.04 to approximately $9.25 per 1,000 views.
* Targeting Fall and Summer seasons: Potential 10–15% increase in seasonal revenue by aligning content with peak audience activity.
* Churn reduction (Summer/Fall from 0.5–0.8 to ≤0.4): Estimated 8–12% increase in annual subscriber retention, compounding long-term growth.
* Investing in early video promotion (first 72 hours): Projected 25–30% increase in ad impressions.
* Leveraging low churn months (eg, August): Intensified posting during school holidays could add 5–7% to monthly revenue, driven by higher viewer availability.

# 5. Recommendation

1. **Schedule strategically**: Focus posting on **Tuesdays and Fridays** to capitalise on peak engagement and RPM.
2. **Seasonal alignment**: Publish revenue-focused content in **Fall and Summer**; use **Winter** for strategic planning and skill-building.
3. **Retention-focused initiatives**: Introduce loyalty perks (exclusive content, Q&A sessions) during high-churn periods to stabilise retention.
4. **Boost early visibility**: Run targeted paid promotions and teasers within the first 72 hours of publishing to maximise impressions.
5. **Optimise August activity**: Increase uploads when churn is lowest (August) to drive cumulative audience growth.
6. **Data-driven refinements**: Regularly track watch time, comments, and shares to fine-tune upload frequency and content themes.

# 6. Conclusion

By strategically aligning content schedules with engagement data, addressing churn through loyalty-focused initiatives, and optimising ad monetisation during peak impression windows, the YouTuber stands to increase ROI by 20–25% within the next content cycle. Combining these actions with targeted seasonal strategies and early visibility promotion will not only improve immediate revenue performance but also establish a long-term foundation for sustainable growth, stronger community loyalty, and higher premium subscription uptake.