Revenue Analysis & Forecast

Data Source: [Spotify Revenue, Expenses, and Premium Users via Kaggle](#)

Project Lead: Kimberly Mizrahi, 2024

1. Objective

The primary objective of this project is to develop a machine learning model to predict Spotify's future revenue based on historical data. The goal is to utilize historical data, including revenue, costs, user statistics, and various key performance indicators (KPIs) to inform strategic decisions related to pricing, cost management, and marketing investments.

---

2. Scope

- **Data Analysis**: This project will include a comprehensive descriptive analysis to understand trends in the data, such as revenue growth, costs, user base expansion, and advertising performance.
- **Model Development**: A predictive model (time series, regression, or machine learning) will be created to forecast Spotify's revenue using multiple variables.
- **Feature Engineering**: Analysis of key features, including premium revenue, cost of revenue, monthly active users (MAUs), and advertising revenue, to build robust prediction models.
- **Evaluation Metrics**: Metrics such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE) will be used to assess model performance.
- **Model Deployment**: The final model will be prepared for deployment with the potential for dynamic updates using new data.

---

3. Stakeholders

- **Primary Stakeholders**: C-Suite Executives, Sales, Marketing, and Product Management Teams at Spotify.
- **Secondary Stakeholders**: Data Analysts, Financial Planners and Business Intelligence Units.

4. Data Overview

The dataset, spanning from March 2017 to March 2023, showcases changes in revenue streams and costs over time, allowing for the analysis of profitability, user engagement, and expense trends across both premium and advertising models.

Variables	Description
Date	The date of each data entry.
Year, Month, Quarter	Temporal information, useful for time series analysis and trends.
Cost of Revenue	The cost associated with generating revenue (e.g., operational costs, royalties).
Gross Profit	Total revenue minus the cost of revenue.
Premium Revenue	Revenue generated from premium subscriptions.
Ad Revenue	Revenue generated from advertising.
Monthly Active Users	Total, Premium, and Ad-Supported Monthly Active Users (MAUs).
Premium ARPU	Average revenue per user (ARPU) for premium subscribers.
Sales, R&D, Admin Costs	Costs incurred in sales and marketing, research and development, and general and administrative functions.

5. Guiding Questions for Analysis

1. **Revenue Trends**: How has total revenue evolved over time, and what are the key drivers (premium revenue vs ad revenue)?
2. **Cost Analysis**: How do the costs of revenue (both total and premium) compare to revenue trends?
3. **Profitability**: What is the gross profit margin for both premium and ad-supported revenue streams, and how has it evolved?
4. **User Engagement**: What are the trends in Monthly Active Users (MAUs), and how does user growth correlate with revenue changes?
5. **Customer Segmentation**: How has Premium Average Revenue Per User (ARPU) changed over time, and are there noticeable trends in user monetization?
6. **Advertising Performance**: How has advertising revenue performed compared to premium revenue, and are advertising costs proportional to revenue gains?
7. **Cost Efficiency**: Are sales and marketing efforts yielding higher revenues, and what is the return on investment?
8. **Seasonality or Cyclicality**: Are there specific periods where revenue or costs fluctuate significantly?
9. **Premium vs. Ad Model**: Which business model (premium vs ad-based) is more profitable over time, and how should the business balance these two streams?
10. **Operational Costs**: How do general administrative costs evolve over time, and are they aligned with revenue growth?

6. Descriptive Analysis (Preliminary Findings)

From a preliminary examination of the dataset:

- **Revenue Trends**: Spotify's total revenue shows a clear upward trajectory from 2017 to 2022. There are significant jumps at the end of each fiscal year (Q4), indicating seasonal or promotional influences.
- **Premium vs. Ad Revenue**: Premium revenue forms a substantial portion of total revenue. However, ad revenue has shown a steady increase, indicating growing profitability from advertising.
- **Cost Structure**: Spotify's costs (e.g., Cost of Revenue, Sales & Marketing) increase proportionally with revenue, but efforts in cost management can be seen (e.g., lower cost per user over time).
- **User Base**: Monthly Active Users (MAUs) have grown significantly, with a larger percentage of users contributing to premium services. Ad-supported users have also increased, contributing to ad revenue growth.
- **Average Revenue per User (ARPU)**: ARPU is relatively stable but shows gradual increases, especially in premium users.

---

7. Deliverables

- **Predictive Model**: A machine learning model to forecast future revenue.
- **Scenario Analysis**: Forecast revenue under different potential scenarios, such as varying user growth rates or changes in ad revenue, to help guide strategic planning.
- **Visualizations**: Clear and actionable visualizations of revenue trends, costs, and user metrics.
- **Insights**: Data-driven insights on Spotify's revenue growth, cost structure, and market performance.
- **Report**: A comprehensive report, including recommendations based on predictive analysis and insights from the data.

---

8. Tools & Technologies

- **Data Cleaning & Processing**: Excel, Python (Pandas, Numpy)
- **Modeling**: Scikit-learn, Time Series Forecasting Tools (e.g., ARIMA)
- **Visualization**: Tableau, Matplotlib, Seaborn
- **Deployment**: Jupyter Notebooks, GitHub, PowerPoint for presentations