# **Project Proposal**

## **Project Title**

Analyzing Worldwide App Store Top Apps Revenue (Sep 25, 2024 – Sep 24, 2025)

## **Purpose**

Provide actionable insights from App Store top apps revenue data to directly support the **Business Development (BD) team**. The goal is to identify high-revenue publishers, breakout apps with strong growth, and category trends that can inform partnership targeting, potential cooperation opportunities, and long-term BD strategy.

## **Outcomes**

* A structured data cleaning and processing pipeline in Power BI or Python/Pandas.
* An initial exploratory data analysis (EDA) to identify revenue distribution, concentration, and category splits.
* Visualization concepts (charts, dashboards) for storytelling around top apps and publishers.
* A framework for ongoing reporting (e.g., monthly/quarterly updates).
* **Prioritization for BD:** Highlight insights that matter most – identifying top revenue partners, spotting fast-growth apps, and mapping category opportunities.

## **Dataset**

* **Source:** Sensor Tower, *App Store Top Apps Revenue (Sep 25, 2024 – Sep 24, 2025, Worldwide)*
* **Format:** CSV (UTF-16, tab-delimited)
* **Scope:** Top apps by worldwide revenue, including app details, publisher details, category, sub-genre, downloads, revenue, DAU, RPD, ARPDAU, and country fields.
* **Status:** Raw file, not yet cleaned or processed.

## **Initial Analysis Plan**

### **1. Data Cleaning**

* Import CSV with correct encoding (UTF-16, tab delimiter).
* Trim and clean text fields (App Name, Publisher Name, Category, Sub-genre, Country).
* Standardize column data types (dates, numeric revenue/downloads, percentages).
* Handle missing/null values.
* Create a unified Publisher Display field.
* Keep only relevant columns for analysis.

### **2. Exploratory Data Analysis (EDA)**

* Inspect distribution of revenue across apps and publishers.
* Identify top apps and publishers by absolute revenue.
* Explore revenue concentration (Top 5 / Top 10 share).
* Examine PoP Growth % to detect breakout apps.
* Category and sub-genre revenue splits.
* **Game Sub-genre:** Analyze revenue contribution and growth by sub-genre to identify which gameplay archetypes are trending.
* **Game Product Model:** Segment revenue by monetization model (e.g., Free-to-Play, Premium) to see which models dominate.
* **Publisher Country:** Map top publishers to their headquarters country to highlight regional strengths and potential BD outreach geographies.
* **Release Timing:** Compare each app’s *Earliest Release Date* with its revenue in the selected time series to see how older vs. newer titles perform, and whether longevity or recency drives success.

### **3. Analysis**

* Calculate KPIs: total revenue, publisher concentration (HHI), Top 5/Top 10 shares.
* Rank apps and publishers by revenue.
* Identify breakout apps (high revenue + high PoP growth).
* Compare revenue patterns across **Game Sub-genres** and **Product Models**.
* Link publisher revenue with **Publisher Country** to reveal which regions dominate.
* Examine correlations between **Release Date** and revenue performance.

### **4. Visualization**

* **Bar charts:** Top 10 apps and Top 10 publishers by revenue.
* **Treemap/Pie:** Category and sub-genre revenue shares.
* **Stacked bar/column:** Revenue by Game Product Model.
* **Map visual:** Publisher revenue aggregated by country.
* **Scatter plot:** Release Year vs. Revenue (bubble size = Downloads).
* **KPI cards:** Total revenue, Top 5/10 share, HHI.

### **5. Data Storytelling**

* Present a **diverse view of the market**: who leads by revenue, who is growing fastest, which sub-genres and monetization models are trending.
* Highlight **regional strengths** by showing which publisher countries dominate, and where cooperation opportunities exist.
* Show the balance between **evergreen titles vs. newer releases**, emphasizing how game age relates to success.
* Build a narrative for BD: from **top revenue partners** (secure cooperation), to **fast-growth apps** (early engagement), to **category opportunities** (expand pipeline).