**Data Analyst Test Task**

**Introduction**

Welcome to the Data Analyst Test Task. Our objective with this task is to assess your process when doing data analysis. We want to understand your skills and how you approach data challenges. We are aware that there are multiple ways to tackle each different problem and, in this sense, there is no single right answer. What matters the most is your though process, the choices you make and how you derive and present your results.

**What we expect**

Simply put, the expected outcome of this task is a data analysis report.  
Some ground rules and guidelines on this:

* We expect a professionally written report. Be succinct, clear and provide actionable insights. We want to understand how you present and communicate your findings. Make sure the story you���re telling is easy to understand and to act upon. Imagine the CEO will read it!
* Send all code, the report and all support files in a zip file. We want to understand not only how you work but also be able to review the code and the process behind the analysis. A clean, reproducible and well defined data analysis project is a huge plus.

**Quick Game Analytics Primer**

The analysis is about a free to play mobile game A/B test. In the context of mobile games there are some concepts and variables that are very important to us game analysts and to the industry as a whole. Here���s a small selection of things you should know, that are relevant when dealing with new users:

* A **new user** is someone who never played the game before. The date when a user first started playing is called the **Acquisition Date**. We can identify it by finding the date of the first ever recorded activity of the user.
* The **Retention Rate at Day N** is the percentage of users who played the game N days after their Acquisition Date. If for example 1000 new users started playing on May 1st, and of those 400 also played on March 3rd, then we say the Retention Rate at Day 2 is 40% for May 1st.
* The **Cumulative ARPU at Day N** measures how much revenue comes from each user N days after he or she started playing the game. This is a measure of revenue normalised by the size of the cohort of new users. It���s calculated by taking all of the revenue generated by users between their Acquisition Date and N days after and dividing it by the number of new users on that same Acquisition Date.  
  ARPU stands for Average Revenue Per User.
* The **Cumulative Conversion Rate at Day N** is the number of users who made a purchase in the game at some point between their Acquisition Date and N days after, divided by the size of the cohort of new users. Typical values for N are 1, 7, 15, 30.
* An **in-app purchase** is a single purchase a user does inside the game, using real world money. This is the main source of our revenue.

**Task!**

**Introduction**

We���ve conducted an A/B test in one of our games aimed only at new users. This game contains a shop where users can buy several different products that make their game experience better. This menu is where they can access the majority of in-app purchases.  
Currently the shop only appears to the users after they played at least 2 matches. This choice aims to reduce the distractions during the user���s first games, but also to make sure the user sees the shop only after engaging a bit with the game, thus increasing the chances that he or she will buy one of the products.  
We believe however that we can tune this parameter to better avoid users leaving the game and improve the chances they���ll make a purchase. Our hypothesis is that by delaying the appearance of the shop our Retention and Conversion Rates will increase, ultimately increasing our average revenue per user. Conversely we believe that showing it earlier has a negative effect on those metrics.

To test the hypotheses, during a 15 day period, all of the new users where divided into three fairly equally sized test groups:

* Users in the *control\_group* had access to the shop after playing 2 matches.
* Users in the *test\_group\_a* had access to the shop after playing 4 matches.
* Users in the *test\_group\_b* always had access to the shop, regardless of the number of matches played.

**Data**

You can find the datasets for this task in this link: <https://drive.google.com/file/d/1w8y5MFFJY6Gw7vKdsETEuXfWkAzXAapw/view?usp=sharing>

Inside the zip file you���ll find three sets of data.  
The file *data\_daily\_activity.csv* contains the daily activity of our users. Each row represents a day the user played the game, i.e., there is one row per each day the user was active.  
This contains all the user���s records from the moment they started playing the game to the last day of the A/B test.  
This dataset has approximately 236K rows and has the following structure:

* **userId** The unique user ID of each user active.
* **date** The calendar date when the user was active.
* **countryCode** The two letter country code from where the user played the game.
* **platform** The mobile OS used by the user to play the game.
* **abTestGroup** The group to which the user was assigned to.

The data in the file *data\_daily\_matches.csv* registers the number of matches played by each user, on each day where that the user played a match.  
This dataset has the following structure:

* **userId** The unique user ID.
* **date** The calendar date when the user played a match.
* **matches** the number of matches played on the given date.

Finally you���ll find the file *data\_in\_app\_purchases.csv* which registers all the in-app purchases made during the period in analysis.  
This dataset has the following structure:

* **userId** The unique user ID.
* **date** The calendar date when the the purchase was completed.
* **product** An identifier for the product being bought.
* **cost** The cost of the product in question in US dollars, i.e., how much the user spent on this particular purchase.

**Question**

We need to make a decision regarding the in-game shop and when to show it to our new users, so we need your help. Please report your findings and recommendations with statistical evidence, and help us visualise the results.  
Our objective is to keep users playing our game, make the shop attractive with the end goal of improving our average revenue per user.  
Having said this, the question is simple: should we maintain the shop as it is or should we adopt a new strategy?