Almedia's Data Science Case

Almedia is looking for data scientists, analysts, and engineers to start building its data team.

This case aims toward providing a similar context as to what the company is living, while allowing the candidate to demonstrate his knowledge and skills in any of those fields.

You will be presented with a database, simulating real data, as well as context over the company's business model and operations.

It is your challenge to understand the data, identify business opportunities, and to derive actionable insights as well as data driven solutions to optimize the workflow and generate value to the company.

You are free to present your results in any format, but keep in mind that those will display not only your competencies and areas of expertise, but also your seniority.

Once completed, you may submit your solution to the HR member with whom you are in contact with. In normal circumstances, we should respond within a week.

Be creative, have fun, and good luck!

The Company

Almedia's main product is [Freecash](#) (please do check it out).

Freecash allows users to earn real money by playing games, participating in surveys and promotions.

On signup, the user can explore an offerwall, pick and choose an offer of an app they might like, start playing & earning.

Each offer has customized tasks, which when achieved, reward the user with a certain amount of coins, which in turn, can be used to cash out $$ (1000c ⇸ 1$).

Companies advertise with us, because we are able to provide quality customers at a competitive rate compared to traditional user acquisition channels.

In short, our business model is:

1. We spend money to acquire new users.
2. Some of these users engage in our partner's offers. We get paid for those users.
3. Users may achieve certain goals, in which case they get rewarded by us.

While some goals are easily achievable (such as installing an app), others might require the user to dedicate a considerable amount of time, or to make in-app purchases. Both of which are metrics our partners try to optimize for when considering which acquisition channel to invest in.

Departments & Processes

The company is on a hyper growth stage, with most of its 100 employees having joined in the previous months.

Departments, teams, and processes are constantly being revised as we try to improve and optimize our operation.

While there are no specific constraints as to what data can work on, the core teams and processes that drive profitability are:

1. User Acquisition
2. Offer Management
3. Anti Fraud

Those teams have some data dependencies (KPIs, analysis, monitoring), but for the purpose of this exercise, feel free to consider they do not.

The Data

The data shared is a subset of a replica of our production database.
Right now all our analysis and work takes place on that replica.

While we are not sharing all of its content, and anonymizing part of it, the provided data should suffice for a well versed candidate to explore different hypotheses, propose, and build solutions.

Entities shared are:

signups: User creation.
user_offers: Lists user's engaged offers.
offers: Offers that were displayed to users.
chargebacks: Fraud related events.
purchases: User in app purchases.

Their attributes should be (somewhat) self-explanatory.

It's important to notice that the data is not perfect, nor is it meant to be.
We are working to develop a well documented and structured data warehouse, but in the meantime, we are looking for collaborators who are capable of discerning between noise and valid information, as well as to make reasonable assumptions about the system and its outputs.

Objective

Generate value by driving process change & creation through data driven solutions.

These may be analysis, models, data pipelines, dashboards, reports, and/or whatever comes to mind - as long as there are valid hypotheses and clear deliverables.

We do not expect you to tackle all of the above, work on what you are passionate about and don't lose track of **the objective**.

While we do not impose time constraints, a good solution usually takes the better part of a weekend. Spend as little or as much time as you believe is necessary.

Don't hesitate to ask questions, but do try to figure out things by yourself by making reasonable assumptions. We don't always have someone available to help, hence being able to make assumptions is part of the job.

While we use mostly python, notebooks, and SQL, you are free to consider any technology on the proposed solutions.

Simplicity and awareness is preferred over the implementation of complex techniques which you do not master, or that show diminishing returns.

Depending on the proposed solutions, you are free to consider the business or technical stakeholders. Present your results accordingly.

If your solution satisfies our minimum criterias, you should be invited to a technical/business presentation. Solutions that do not satisfy our criteria, but that show potential, will receive in depth feedback.

Have fun & good luck.