Data Analyst - Analytics Challenge

Cedar is growing rapidly and we are looking for a Data Analyst that can quickly transform data into useful insights and present them to other teams and clients (both technical and non-technical). We hope this challenge gives you a good idea of the kind of data that you would be working with and the kind of problems that we are trying to solve on the Data Science team. We appreciate you taking the time to do this challenge!

## About Cedar

Everyday, providers (hospitals, medical centers …) send us invoices that are ready to be sent to patients. Patients may already have previous invoices outstanding with an amount due. We are charged with contacting the patient or the guarantor (via paper mail, text, or email) to invite them to check out their bill on our website. Patients can use our website app to see the details of their bills and their explanation of benefits (insurance payments and adjustments), select a payment plan (when available), or pay their bills (via credit card, Apple Pay or a mailed-in check). They can also chat with their healthcare providers if they have billing questions through our website.

Cedar has the ability to personalize and tailor the patient experience. For example, we can vary the methods we use and the frequency with which we contact patients depending on patient preferences. We can also decide to offer payment plans or discounts based on balance sizes or characteristics of the patient’s bill. The ultimate goal is to maximize the rate at which patients take actions to resolve their balances – either through payment or by letting the healthcare provider know of any changes that need to be made to the bill.

## The Dataset

We provided you with a sample dataset of 25,000 patients that have an outstanding balance as of February 1, 2019. The dataset contains information about their invoices (amount due after insurance, days since the visit, etc.), their demographics (age, number of previous visits, etc.), and some metrics about past and current engagements (i.e. number of chat messages for past invoices). We have also provided you with data from the Census Bureau (median household income and home value by zip code) for the patient. Finally, you will also find a column indicating if the patient has made any payments in the 120 days following Feb 1st 2019 (whether partial or full).

Descriptions of the data are available in the file “columns\_description.csv”.

## The Challenge

We are interested in the following:

* What are the key factors that affect patients’ propensity to pay their bills?
* If you could run experiments to test the impact of a change on patients’ propensity to pay their bills, what would you be interested in testing first? Why? (e.g. test if sending a paper statement or an email first to the patient will impact results)

You can use any tools you want to do your analysis (Tableau, Jupyter notebooks...). The output should be a presentation of a few slides with your findings. You can assume that the presentation will be for our Product team (who knows our platform but is always looking for ways to improve it). Please send us your presentation and the code that you wrote to do your queries if any.

If you have questions, please email [data\_science@cedar.com](mailto:data_science@cedar.com), we will try to get back to you as fast as we can.

The format is open ended. You can spend as much time as you want on it but we advise not spending more than 5-7 hours.