**Job Assessment - Data Scientist**

|  |
| --- |
| **PURPOSE:** |

**Congratulations on making it to the assessment phase!** This next step is required of all candidates**, and has 3 purposes.**

● The assessment helps the hiring manager further determine if you should be brought in for an in person interview with the whole team.

● The assessment helps you gain a deeper understanding of the day to day of the role, and of HubSpot’s business. This should help you further determine if this role is the right next step for you, in your career.

● The assessment will help you, and the hiring team, have more meaningful conversations in your next interview(s).

|  |
| --- |
| **ASSESSMENT:** |

As a business, we are trying to understand who would be our target audience for one of our products -- CRM. Please provide insight around the following two questions:

First, we would like to understand who our current customer base is and how different they are from non-customers.

Second, ideally, we would like to build a model so that we can use it to identify our future customers.

You can use any method you like to respond to the above, but please offer a brief summary of why you chose the method you did. Also please feel free to note questions you had while going through the process. Datasets to use for this assessment and definitions of fields are below. Please do not spend more than three hours on this assessment (cumulatively).

Attached are three dataset and here are their descriptions:

**customers.csv.** This file includes a sample list of our customers.

**noncustomers.csv.** This file includes a list of companies that are not our current customers.

**usage\_actions.csv.** This file includes behavioral data for our users. Please note, HubSpot is a freemium-model business, meaning one doesn't have to be paid customer to be able to use HubSpot. Usage may take place before actually buying the product or even after stopping being a customer.

**CLOSEDATE:** date when they became a customer

**MRR:** acronym for "monthly recurring revenue", a monthly payment amount customers make

**ALEXA\_RANK:** a score given by Alexa considering many aspects of a business, such as traffic, performance etc. For example, Google has a score of 1, Facebook has a score of 4 etc.

**EMPLOYEE\_RANGE:** min and max number of employee size of the company

**INDUSTRY:** industry of the company

id: unique identifier, you may also treat it as if their portal number within which they have taken the actions listed below

**WHEN\_TIMESTAMP:** date when usage activity happened

**ACTIONS\_CRM\_CONTACTS:** number of actions users have taken in their portal on Contacts property.

**ACTIONS\_CRM\_COMPANIES:** number of actions users have taken in their portal on Companies property.

**ACTIONS\_CRM\_DEALS:** number of actions users have taken in their portal on Deals property.

**ACTIONS\_EMAIL:** number of actions users have taken in their portal on Email property.

**USERS\_CRM\_CONTACTS:** number of users that have used Contacts property.

**USERS\_CRM\_COMPANIES:** number of users that have used Companies property.

**USERS\_CRM\_DEALS:** number of users that have used Deals property.

**USERS\_EMAIL:** number of users that have used Email property.

|  |
| --- |
| **TIMELINE:** |

Please complete assessment within 3 business days. If you have a conflict and need more time, please reach out to the recruiter to coordinate. We look forward to reviewing!