

DNC Skills Assessment

Fundraising Analyst

Along with your upcoming interviews, this assignment will help us assess your technical proficiency in working with datasets as well as other relevant skill sets.

Our overall goal is to develop a broad (but shallow!) picture of your approach from start to finish. With that in mind, it's ok if elements of this assignment feel incomplete by the time you turn it in; there will be plenty of time to discuss your responses and approach in your interview.

Instructions

- Time Limit:** We estimate this assignment will require 3-5 hours to complete. You may work on your own time. If you are still working after 6 hours in total, please stop and leave a short note of what you would do with more time.
- Deadline:** Return your materials 48 hours after they are shared with you.
- File Formats:** Please submit written responses as a Google Doc, Google Sheet or PDF, and any code as a secret GitHub gist or zipped directory.
- Restrictions:** You may access any reference materials you find helpful. Do not confer with others.
- Annotations:** In your submission, please affirm that this is your own original work and note approximately how long you spent on the assignment in total.
- Questions:** Please feel free to reach out with any questions or if something is unclear. There are no “trick” questions and we’re happy to provide you with additional information as needed!

Our Expectations

This assignment gives us a chance to see how you think about a type of problem we tackle every day. It is **not** a test of your speed, grit, or ability to deliver ‘perfect’ work on a timeline.

While we are excited to see your work, we respect that you have a life outside of your open job applications and will have that in mind as we evaluate your responses.

Resources

Datasets

Three tables will be referenced in this assignment:

[donations_base_table](#)

This table is a sample of contribution-level data from our source-of-truth platform for donations to the DNC. It contains some basic information about the donations (date, amount) and join keys to our online payment processor. All donations in this sample data originated from a single email campaign (one or multiple emails sent on the same day, from the same sender, containing the same general message).

[processor_metadata](#)

A payment processor handles our online contributions before they are imported and reconciled into our primary contribution platform (see [donations_base_table](#) above). The payment processor includes important metadata about each contribution, including the specific online campaign from which the donation originated as well as flags for recurring donations.

[campaign_metadata](#)

For online campaigns like email, we have additional upstream key performance indicators from our email service provider. This dataset contains more descriptive information on the email campaign, as well as other important email metrics.

Data Dictionary

We include a [data dictionary](#) at the end of this document.

BigQuery Set-Up

If you choose to use SQL for your responses, you may use a SQL environment and syntax of your choice. We use BigQuery and [Standard SQL](#) in our day-to-day so if you'd like to emulate our environment, we've included [set-up instructions](#) at the end of this document.

Note

This step is purely optional. While we use BigQuery and Standard SQL in our workflows and this might give you a glimpse of our day-to-day, the following is not a requirement. **You will neither be penalized nor rewarded for not/using BigQuery or Standard SQL.**

Prompts

Section 1

You can answer the questions in this section with SQL queries or by using functionality in a Google Sheet or using another programming language like Python or R. Primarily, our team works in SQL but if you prefer to showcase another method, please feel free to use it. If you use Google Sheet or Excel, please clearly show the work you did to arrive at your responses including any relevant formulas or functions.

The Email Director on the Mobilization team has some questions for you about how one of their emails performed.

1. According to the *donations_base_table*, how many total donations, unique donors and revenue did the email from (fictitious) Senator Anne King generate by day? (All the donations in the *donations_base_table* are from that single campaign).
2. What was the overall open rate of this email campaign (including all the variants)? This is defined as the percent of total sent emails that were opened. Round to the nearest hundredth.
3. What was the conversion rate for new monthly recurring orders? The conversion rate is defined as the percentage of unique orders that were set up to be recurring monthly orders. Express this as a percentage (string type with suffixed “%”) rounded to the nearest integer.
4. What is the count of unique orders that are still recurring in May? For these “active” orders, what is the total revenue they have generated to date? Please respond with two values.

While answering these questions above:

5. Write a few sentences about at least two assumptions you had to make about the data sources. Would you have any follow-up questions or flags for the Email team and/or other Analytics team members?

Section 2

In this next section, **you should only complete one part.** We encourage you to pick the option based on your strongest skill set or preference. You can choose between:

- Option A: Analytics Infrastructure (SQL-based)
- Option B: Reporting and Visualizations
- Option C: Experiment Evaluation using Statistical Methods

Option A

The Sustaining Donor Manager on the Mobilization team has some questions about monthly recurring revenue generated from emails. In order to answer this question on a more regular basis, we need to create some infrastructure to support these types of requests. **For this exercise, you'll need to provide a SQL statement to generate a new table.**

1. Create a table of unique sustaining *orders* generated from this email. This table will be unique by their `order_number`. Multiple donations can be associated with a single `order_number`. If an order re-occurs for 6 months, there should be 6 donations associated with that order.
 - a. You can add any metadata that you think might be interesting about the sustaining monthly order, but you must add columns for:
 - i. **`is_active`**: a boolean flag of whether or not that gift is active - for simplicity we will define a sustaining gift as active if it was seen in May 2021
 - ii. **`count_recurring_gifts`**: how many gifts have been given to date in that recurring series
 - iii. **`first_gift_amount`**: amount of first gift in recurring series
 - iv. **`last_gift_amount`**: amount of last gift in recurring series
 - v. **`originating_email_tag`**: name of `email_tag` where the recurring gift started from

vi. **donor_id:** Id associated with first donation in the recurring order.

Note: the same recurring order may have different donor_ids due to duplication issues.

2. Add a column for **remaining_lifetime_value**. Assume on average a donor makes 12 payments on a recurring gift (excluding the first). If the donor is still active, calculate the remaining payments we might expect to receive from this recurring order.
3. Write at least two test queries to validate the data in the table you created or any assumptions you made. Indicate what you would expect the query to return if the data were valid.
4. What kind of documentation would you need to provide to other team members on the Fundraising Analytics team about your new derived table? There is no need to create the documentation but please describe what you would create to support adoption and clarity.

Option B

After answering some simple questions about the email, our Email Director now would like to have a daily refreshing report for her to check that summarizes email performance.

For questions 1-2, respond in a short (1 page) [product requirements document](#) for a potential dashboard.

1. What are 3-4 questions you would ask the Email Director before creating your report or dashboard?
2. You're partnering with another Fundraising Analyst to pull the data for your report. Please detail 2-3 requirements (for example: columns, format, delivery, timing) you might communicate to ensure you and your stakeholders have the data you need.
3. Based on your product requirements document and (dummy!) stakeholder responses, create a dashboard (1 page) including 2-3 visualizations that you think would be important to include for the Email Director who wants to understand email performance.

Option C

Our Email Director also included a test in this email to determine the best subject line for this email. They give you the context that:

- The team tested three subject lines on a subset of subscribers before sending the email to our full list.
- Based on eyeballing the results of the testing data, the team chose one of the subject lines and sent the email to the rest of the full email list.

You are helping the Email team with a more comprehensive framework for conducting experiments and evaluating testing results. Answer the following with a few sentences each. **Your solution to this question should showcase statistical methods, including an understanding of variation associated with experimentation.**

1. Based on the data, which of the three subject lines went to the full list?
2. What additional 2-3 questions would you have for the Email Director about the test they ran? You should make sure to ask questions that will address assumptions in your analysis.
3. How many emails received each version of the email, defined by the subject lines? How much revenue did each version of the subject line generate?
4. What primary metric(s) would you use to evaluate the 'best subject line' and why? Are there any other metrics that would be valuable to monitor that are not provided in this dataset?
5. How would you come to a conclusion about the subject test? Please give specific numbers, results of any statistical tests, and/or accompanying visualizations to support your answer.
6. Do you think the email team chose the right subject line to send to the full list? Would you have chosen a different one based on your analysis? How would you explain and best communicate your results to stakeholders not familiar with statistics?

Data Dictionary

table_name	Table Description
donations_base_table	Contribution-level data from source-of-truth base table
processor_metadata	Additional metadata for contribution from payment processor that handles online contributions
campaign_metadata	Email campaign data from email service provider

donations_base_table

table_name	column_name	is_nullable	data_type	description
1 donations_base_table	contribution_id	NO	STRING	Contribution Identification Number
2 donations_base_table	donor_id	NO	STRING	Donor Identification Number
3 donations_base_table	gift_date	NO	DATE	Contribution Date (YYYY-MM-DD)
4 donations_base_table	gift_amount	NO	INT64	Contribution Amount
5 donations_base_table	online_order_id	YES	STRING	Online Order-Level Identification Number
6 donations_base_table	online_reference_line_item_id	YES	STRING	Online Item-Level Identification Number

processor_metadata

table_name	column_name	is_nullable	data_type	description
1 processor_metadata	line_item_id	NO	STRING	Contribution Identification Number
2 processor_metadata	gift_date	NO	DATE	Contribution Date (YYYY-MM-DD)
3 processor_metadata	gift_amount	NO	INT64	Contribution Amount
4 processor_metadata	order_number	NO	STRING	Order-Level Identification Number

5	processor_metadata	recurring_period	NO	STRING	Recurring cadence of contribution
6	processor_metadata	recurring_sequence	NO	STRING	Sequence number of contribution in recurring series
7	processor_metadata	ref_code	YES	STRING	Campaign-level source code
8	processor_metadata	ref_code_variant	YES	STRING	Variant-level campaign source code

campaign_metadata

	table_name	column_name	is_nullable	data_type	description
1	campaign_metadata	variant_name	NO	STRING	Email Campaign Variant ID
2	campaign_metadata	email_tag	NO	STRING	Email Campaign ID
3	campaign_metadata	email_send_date	NO	DATE	Email Send Date (YYYY-MM-DD)
4	campaign_metadata	email_from	NO	STRING	Email Sender
5	campaign_metadata	email_subject	NO	STRING	Email Subject Line
6	campaign_metadata	total_sent	YES	INT64	Number of recipients associated with email
7	campaign_metadata	opens	YES	INT64	Number of opens associated with email
8	campaign_metadata	clicks	YES	INT64	Number of clicks associated with email
9	campaign_metadata	unsubs	YES	INT64	Number of unsubscribers associated with email

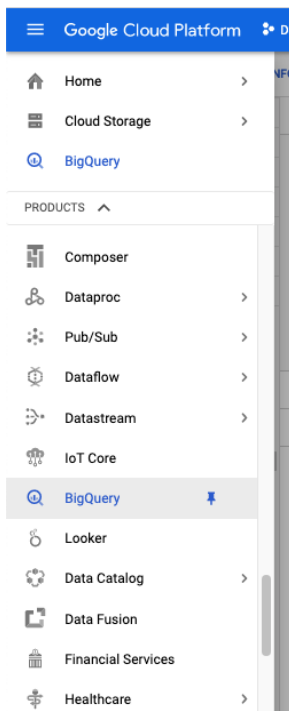
BigQuery Set Up

Note

This step is purely optional. While we use BigQuery and Standard SQL in our workflows and this might give you a glimpse of our day-to-day, the following is not a requirement. **You will neither be penalized or rewarded for not/using BigQuery or Standard SQL.**

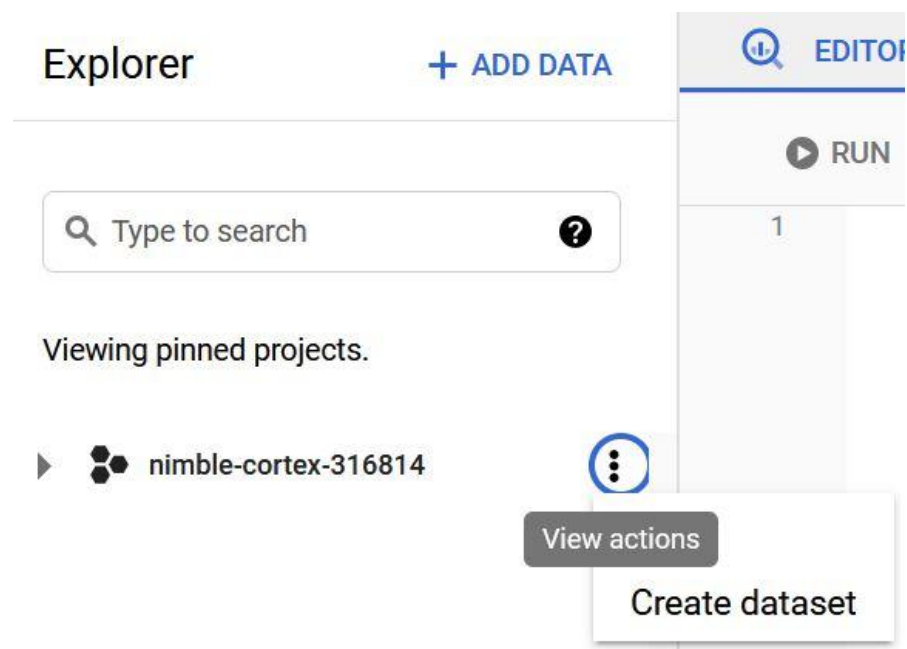
Sign Up for Google Cloud Platform Free Trial

You can sign up for a free tier of BigQuery [here](#) by clicking “Get Started for Free”. While you need to provide a credit card to complete the transaction, you will *not* be signed up automatically after your trial expires.



Open BigQuery

After logging into Google Cloud Platform, navigate to BigQuery on the left-hand menu.



Create a Dataset

First, set up a dataset by clicking the three option dots on the right of your sample project generated by GCP.

Create a Table

Load your CSVs in as tables by clicking the [+] sign to the right of your newly created dataset.

EDITOR

DNC_AS...

COM

nimble-cortex-316814:dnc_assessment

Description

None

Labels

None

Dataset info

Dataset ID	nimble-cortex-316814:dnc_assessment
Created	Jun 14, 2021, 10:48:09 AM
Default table expiration	Never
Last modified	Jun 14, 2021, 10:48:09 AM
Data location	US

Source

Create table from: Upload Select file: ? donations_base_table - Sheet1.csv Browse File format: CSV

Destination

☒ Search for a project ☐ Enter a project name

Project name My First Project Dataset name dnc_assessment Table type Native table

Table name donations_base_table

Schema

Auto detect ☐ Schema and input parameters

☐ Edit as text

Name	Type	Mode	
contribution_id	STRING	NULLABLE	×
donor_id	STRING	NULLABLE	×
gift_date	STRING	NULLABLE	×

Create table Cancel

Upload CSV

Create a table from an upload, setting the file format as a CSV.

Give it the same name as the file provided.

Input the column names explicitly. You can use the data dictionary as a guide.

Input Column Names and Types

Schema

☐ Edit as text

1 : STRING

Field names are required

For ease of set-up, you can copy and paste the below JSON arrays to set the field names and types for each table.

donations_base_table

```
{"name": "contribution_id", "type": "STRING"}, {"name": "donor_id", "type": "STRING"}, {"name": "gift_date", "type": "DATE"}, {"name": "gift_amount", "type": "INT64"}, {"name": "online_order_id", "type": "STRING"}, {"name": "online_reference_line_item_id", "type": "STRING"}
```

processor_metadata

```
{"name": "line_item_id", "type": "STRING"}, {"name": "gift_date", "type": "DATE"}, {"name": "gift_amount", "type": "INT64"}, {"name": "order_number", "type": "STRING"}, {"name": "recurring_period", "type": "STRING"}, {"name": "recurring_sequence", "type": "INT64"}, {"name": "ref_code", "type": "STRING"}, {"name": "ref_code_variant", "type": "STRING"}
```

campaign_metadata

```
{"name": "variant_name", "type": "STRING"}, {"name": "email_tag", "type": "STRING"}, {"name": "email_send_date", "type": "DATE"}, {"name": "email_from", "type": "STRING"}, {"name": "email_subject", "type": "STRING"}, {"name": "total_sent", "type": "INT64"}, {"name": "opens", "type": "INT64"}, {"name": "clicks", "type": "INT64"}, {"name": "unsubs", "type": "INT64"}
```

Set Advanced Options


You can generally follow the default settings for Advanced Options though we suggest explicitly skipping the header row.

Click “Create Table” to create your table!


Advanced options

Write preference:

Write if empty

Number of errors allowed: 

0

Unknown values: 

☐ Ignore unknown values

Field delimiter: 


Comma

Header rows to skip: 

1

Quoted newlines 

☐ Allow quoted newlines

Jagged rows 

☐ Allow jagged rows

Encryption

Data is encrypted automatically. Select an encryption key management solution.

*UNSAVE... 2

RUN

MORE

SCHEDULE

SAVE

CANCEL

1SELECT * FROM `nimble-cortex-316814.dnc_assessment.donations_base_table`

Query results

SAVE RESULTSEXPLORE DATA

Query complete (0.3 sec elapsed, 114.3 KB processed)

Job information

Results

JSON

Execution details

Row	contribution_id	donor_id	gift_date	gift_amount	online_order_id	online_reference_line_item_id
1	55373724	266299656	2021-06-03	2	PP261519211400	392168272
2	55137412	266299656	2021-05-01	2	PP261519211400	387487317
3	54872897	266299656	2021-04-01	2	PP261519211400	383217927
4	55138402	253827020	2021-05-01	8	PP261519199642	387513427

<

Rows per page:1001 - 100 of 1826

Fi

Test Table

Test your newly created table with a simple `SELECT *` statement. If all the columns were loaded in correctly, you're good to go!