



Data Analyst Evaluation



▼ Type

Evaluations



Deadline

7 Days From Date of Assignment



Property

Empty

Hello!

Thank you for applying, and welcome to the second round of evaluation.

```
SELECT "WELCOME" from data_analyst_evaluation_table;
```

● Please read this document very carefully. As a data analyst - you are in charge of putting data to life - by bringing out insights. And we can't emphasize enough how important reading is.

Your goal is to crunch the data and present it - in the most lucid way to the business team. Please label the columns, test your queries etc, before submitting the same.

Deadline - Within 7 days of the date when it was assigned

Grace period - 1 day


Once you submit - we will usually complete the evaluation in 7 working days. We go through every single evaluation in detail. If you are shortlisted, then we will reach out to you, if not, then feel free to re-apply with us - after 6 months. We are unable to provide feedback/solution to keep the integrity of the process.

Why us?

- Async Work Mode
- Flat organisation
- Super Cool Team
- (Super) Fast Machines
- Interesting Problems
- Work life balance
- A lot of autonomy
- Unlimited Remote
- Engineers - Of the engineers, for the engineers, by the engineers

Index:

1. Problem Set 1 - Working with Regex
2. Problem Set 2 - SQL
3. Problem Set 3 - Working with BigQuery
4. Submission Process

 Please carefully read the instructions and document the notebook/code well - as we will be running that in order to evaluate the assignment.

 As with anything, please read this entire document before attempting the assessment.

Why read the next section

Before you proceed with the evaluation, which may take hours of your time, it's very important that you read about what we do and what we are looking for - so that it aligns with your goals.

A Brief Preface

1. What we are looking for?

An ideal candidate should be extremely good with data analysis. He/She should be able to write efficient queries.

2. About this role 🏢

We hire people with broad sets of technical skills. You'd find that you are handling many of the projects by your own. It's tough with a very steep learning curve.

This role requires data gathering, extensive writing 🖋️ - for example - explaining the results etc.

Document - where you would find the data sources -

<https://docs.google.com/spreadsheets/d/1VIfunqcshcNug7cJEzzg4cGyhN7FC9TUEcypfPTRMc/edit#gid=1218415869>

Problem Set I - Regex

1. Write a query to extract the keywords from the url column. This query should be run on a Google Sheet. (This should be done via SQL and not via by any formulas.) Also, please note that it keywords - with multiple words - should be separated by space.

```
https://docs.google.com/spreadsheets/d/1VIfunqcshcNug7cJEzzg4cGyhN7FC9TUEcusp=sharing
```

Resource - How to run Queries on Google Sheets - [Link](#)

Problem Set 2 - SQL

Suppose there's a log of security related incidences for about 60 days along with instance_id. Now, there are cases when the incident is reported but the incident_id is not.

- Your goal is to find out - how many incidents have **incident_id** on a given day. How many percentage of incidents have a non-null id?


- The time is in UTC, so convert that into IST.
- What time of the day does the incident_id percentage is lower?

Bonus Points - Use Google DataStudio - to help visualize the same.

Problem Set 3 - Working with BigQuery

Suppose you get a raw data of your website visitors and you are asked:

1. Identify how many users came via a certain referrer url.
2. Also, convert the **visitStartTime** property to IST.

 **Hint:** Group the number of rows grouped by hit→ referral property. (If the hit→referral is not present, please consider that as a null and group this too).

hits										
Row	hitNumber	time	hour	minute	isSecure	isInteraction	isEntrance	isExit	referrer	pag
1	1	0	3	39	null	true	true	true	http://www.google.com/	✓

Submission Process

Fill this submission form - <https://forms.gle/xZFo6ApfGJWdCTuu9> We only consider the submissions through Google Form.

(Before you fill up the form, please read the below points)

All these problems can be solved into a single google sheet - call them solution_1, solution_2, solution_3.

Important note for the third problem - using any cell - eg: A2, A3, and just add your query there.

If you are submitting the DataStudio link - then please put that in a separate sheet

Submission Template (It's not necessary to follow it as it is - it's just given to give you a broad theme of how you can submit the assignment)

```
https://docs.google.com/spreadsheets/d/1siRBL0Ipi0x4uF21JhaAEt26_G8-xP-  
AF62fvwhu_0/edit#gid=2044387380
```

Note: Please make sure you give the **viewer access** to anyone with the link for the **Google Sheet**. (Refer Screenshot)

General access



Anyone with the link

Anyone on the internet with the link can view

Viewer