

Hi there,

Congratulations, you have been successful in progressing to the next stage of the interview process! You are to complete a work sample which will help us understand your thought processes and be able to see your technical skills. Given you won't be familiar with our business model yet, the goal is to understand your ways of working and thinking.

We believe that there are a number of benefits to conducting a work sample as part of the selection process. These include the opportunity to:

- Develop your insight into hipages, so that you are more informed about whether this is somewhere that you believe that you can have an impact and will enjoy working at;
- Work with members of the team with whom you will have regular contact;
- Demonstrate your ideas, communication style and the ways in which you would approach implementation;
- Provide us with insight on your approach to business issues & opportunities;
- Discuss real ideas for moving the business forward.

Over the years, we have found this to be a highly beneficial two-way process in helping both parties determine if hipages is a 'fit' for them

Presentation:

- The presentation should take approximately 30 minutes with 10 minutes Q&A.
- Please send the completed material via [Tableau Public](#) and your choice of [GitHub](#) / [GitLab](#) to Jason (Senior Talent Acquisition Manager) before the interview. If you have questions about the brief, or if you need further information please reach out to ognenloparski@hipagesgroup.com.au

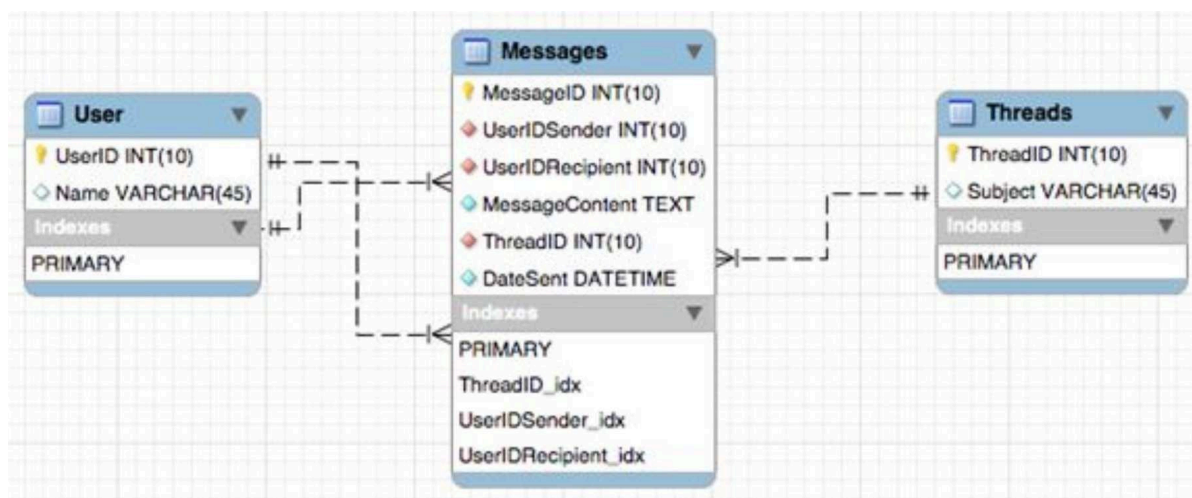
All the best!

Task 1: SQL Programming

Scenario

One of the analysts in the team is unavailable and you have been asked to quickly jump in to help analyse some data and build SQL queries that will later be used to build a BI report for the messaging platform features that was launched in the hipages app recently.

Assume the following database schema with 3 tables: Users, Threads and Messages. Any message is linked to a conversation (Identified by a ThreadID), has a sender (UserIDSender) and a recipient (UserIDRecipient) that each link to a UserID in the User table. A user can be either the sender or the recipient of a message.



Requirement

Can you come up with a SQL query that returns:

- The names and the number of messages sent by each user
- The total number of messages sent stratified by weekday
- The most recent message from each thread that has no response yet
- For the conversation with the most messages: all user data and message contents ordered chronologically so one can follow the whole conversation

Please share the solution via GitHub/GitLab

Task 2: Exploratory Data Analysis (EDA)

Scenario

The tracking has gone live for some time now, with a range of activities having been deployed with the objective of achieving the goal. The team would like for you to have a look at the jobs data again and use exploratory data analysis techniques.

Please find attached some (artificial) data which represents the type of data we capture when processing a job from a consumer. In the attached comma separated text file (jobs.csv) you'll find:

- **time_of_post** = The time the job was posted on hipages
- **latitude** = The latitude of the job location
- **longitude** = The longitude of the job location
- **category** = An integer which represents the category for which the job is assigned. A category (for example, painting, plumbing or air conditioning) allows us to assign the job to tradies who will be interested in the job.
- **number_of_tradies** = The number of tradies who we can send this job out to
- **estimated_size** = The estimated size of the job, this is proportional to the amount of revenue the tradie might get from the job
- **number_of_impressions** = The number of impressions the job gets from the tradies
- **accepted** = A Boolean which represents if a tradie accepts a job invitation (meaning that they have indicated to hipages that they would like us to provide them with the contact details of the consumer). 1 means at least one tradie was interested while 0 means no tradie was interested. This is the target variable you should focus to uncover its causal relationships with the other data.

Requirement

Analyse the data provided and present insights and recommendations. The dataset is designed to see how you approach an analytical task. You can use any approach as you see fit. Some possible scenarios that you could explore in your EDA could include:

- What data preparation steps will you implement?
- Which parameters influence if a job would be accepted?
- Can we predict using the data we have if a job would be accepted? If yes, how? If no, why not?
- Which visualisations would best communicate the findings?

Please share visualisations via Tableau Public, and use GitHub/GitLab for everything else

Task 3: Solution Design (Theory)

Scenario

A product manager at hipages has an idea that might lead to higher lead utilisation by mandating 4 leads per job. Before fully rolling out the new feature, they want to make sure that things will actually improve, and are thus asking for your help in performing an A/B test. They have heard about terms like frequentist and bayesian approach, but they don't know what that means. It is up to you to design and implement a solution, as well as educate the product manager and the wider team.

When a consumer heads to hipages they post a job against a category with some basic details like description and location, the hipages platform then matches this to the most suitable tradies and those tradies can claim a lead. A job will usually have 3 leads (3 separate tradies) who will be able to speak with the consumer and provide a quote. Our business model is based on charging tradies for the lead.

Requirement

Please design a solution for A/B testing which will include:

- all the preparatory work before performing the experiment,
- the technical details of the execution,
- as well as measurement, tracking and interpreting the results.

Feel free to assume a fictitious data structure and schema as well as event taxonomy for the purposes of this section. Please present your ideas in any way that you believe will best achieve the outcome, and share that via GitHub/GitLab.