

## ANALYTICS ENGINEER CHALLENGE DATA ENABLEMENT TEAM

**Hello!** Thank you for your interest in Loadsmart. We appreciate your interest in our company and are excited that you have decided to move on to this step! We hope you find it challenging and fun, and we're looking forward to connecting with you to discuss the results!

The challenge will allow us to learn more about your technical skills. We will use your results along with other dimensions when considering your suitability for the position. This is not the only factor we will be looking at regarding your candidacy. If, for any reason, you are not able to complete the entire challenge, whatever progress you make will be taken into consideration with the broader picture. Please make as much progress as you can and we'll go through Q&A in the technical interview.

You will have up to seven days to complete the challenge and send it back to us. **Please be sure you've confirmed the due date with the representative who sent you the challenge.**

Once you feel your challenge is complete, send your github repository link via email to the representative who sent it to you, containing the README.md with the steps necessary to test your proposed solution. Please make the subject of the email "YOUR FULL NAME - Analytics Engineer Challenge".

Good luck, and have fun!

\*\*\*\*\*

We have attached a data set for you to work with. **Please look for the csv file attached to the email. Here's what we ask you to do:**

1. **Dimensional modeling, SQL and dbt skills:**

- a. **Ingest the data and build a dimensional data model (Start schema):** use **dbt(data build tool)** and a supported database to show your data model. This will make sure that data analysis will be available through SQL;

2. **Python skills:**

- a. **Python functions:** Please create one or more of the following python functions on a Jupyter Notebook::
  - i. **Create a Python function to split the lane column:** create a function that will receive a lane value and split it into 4 new columns: pickup\_clty, pickup\_state, delivery\_city and delivery\_state;
  - ii. **Create a Python function to send a csv file via email:** create a function to send a csv file to an email. The Python function should receive a csv file path, the subject of the email and an email body;
  - iii. **Create a Python function to send a csv file via sFTP:** create a function to send a csv file to an sFTP. The Python function should receive a csv file path and the destination file path;
- b. **Export a csv file using a Python script:** please use the same Python Jupyter Notebook created above to write a Python script that will read your dimensional model and create an export to a csv file. This file should have the list of loadsmart\_ids that were delivered in the last month available on the raw csv data that we sent to you. Below is the list of columns that needs to be on the exported csv file:

- i. loadsmart\_id
- ii. shipper\_name
- iii. delivery\_date
- iv. pickup\_city
- v. pickup\_state
- vi. delivery\_city
- vii. delivery\_state
- viii. book\_price
- ix. Carrier\_name

If you have the ability of creating reports please also try to accomplish this last requirement:

3. **Data visualization:**

- a. **Create a report using the modeled data:** provide a visual analysis so you can make a proof of concept of how your data model works, and how we can use it. You can use Power BI Desktop or Superset.

**Here's what you'll need to deliver in the github repository:**

- 1. The dbt project and scripts used for **creating** the dimensional data model on the database, with any specific instructions needed for us to reproduce it;
- 2. The Python Jupyter Notebook on a specific folder on your github repository that has the dbt models;
- 3. The README.md file with the steps necessary to test your proposed solution;
- 4. If you were able to create the report, please send to us the Power BI semantic model and report or the Superset report you created.

The table headers may be a little confusing since you don't have a complete understanding of our language in the logistics business, so please do your best to review and make assumptions. Unfortunately, we will not be able to answer questions about the challenge while you are doing it, but we encourage you to write them down so we can discuss during the review portion of the interview.

We hope you enjoy working on this project!

Thank you,

Loadsmart Team.