

## DATA ENGINEER OR DATA ARCHITECT CHALLENGE

### BI TEAM

**Hello!** Thank you for your interest in Loadsmart and to be part of this amazing company, we 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 and experience. We will use your results along with other inputs 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 possible, and we'll go through Q&A in the technical interview.

You will have up to five 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 results via email to the representative who sent it to you. Please make the subject of the email "YOUR FULL NAME - Data Challenge".  
Good luck, and have fun!

\*\*\*\*\*

**Attached to the email, there is a data set for you to work with.**

We know this is a simple data set file. Still, we believe with it you can show your skills. Assume this is the data pipeline you have to deal with, so implement or consider having the necessary steps to deliver this data with quality, performance, and meaningful, even not possible to implement all you would like to do. Please, in some way, let us know what is in your mind.

**What we expect from you:**

1. **Build a dimensional data model** in a free/open-source database of your choice (PostgreSQL, MySQL, SQLite and so), so it can hold the data set we sent you and make data analysis available through SQL.
2. **Load the data from** the data set to the database model you've created, using an open-source ETL tool or language you want, but **processing the data using Python (Pandas) is a great plus!**

**As a bonus:**

3. **Provide a visual analysis** so you can make a proof of concept of how your data model works and how we can use it for a **self-service** data analysis. You can use **any** data visualization of **your preference**. • In case of doubt, we recommend Power BI. But that should be used only as a tiebreaker. We will assess the quality of your deliverable **only** - the choice of tool will not be a factor - and you must choose the tool you feel most comfortable with.

**We going to reproduce what you did here, so please try to make our journey easy, write a good walkthrough on how to "setup" your developed solution.**

**Here's what you'll need to deliver:**

1. The scripts used for **creating** the database's data model, with any specific instructions needed for us to reproduce it.
2. The scripts used for **loading** the data to the model, with any specific instructions needed for us to reproduce it.
3. If you choose to go through the “**Bonus**” then we need the resulting **dashboard/data set** you used to prove your solution's concept.

**General information:**

The table headers may not make too much sense due to our transportation business, so please do your best to figure out and make assumptions. Unfortunately, we will not be available to answer questions about the challenge while you are doing it, but we encourage you to write them down to discuss them during the review part of the interview.

We hope you enjoy working on this project!

Thank you,

Loadsmart Data Team.