Data Engineer Technical Test

Clevertouch Consulting

Wednesday, 19 November 2025

V2.1

clevertouch
love martech

1 Introduction

A global B2B marketing company who utilises multiple marketing channels (Web, Email, etc…) has approached you to help their marketing team gain better visibility and insight into how their Leads are interacting with their campaigns

Today their data is isolated and fragmented. Campaign data is produced by their Data team, while Activity Logs come with various platforms, and additional third-party behavioural data is also available via an API.

The issue they face is that while generating this data isn't a challenge, none of these data sets share a common data schema or structure.

For them to assess whether they'd like to work with you longer term, they have provided you with sample datasets from all sources and have assigned you a task to complete. The data sets are;

- **Campaign Data (Parquet)**
  An export provided by the data team containing Campaign ID, Names, Dates & Owners.

- **Activity Data (CSV)**
  Provided by the marketing team, and has relational data structures to Campaign Data

- **Behavioural Data (JSON)**
  Data provided by third-parties via the REST API

2 The Task

The client is asking you to complete two parts of a task, with each part being possible to independently review and assess.

Part One: Ingestion & Transformation
They're expecting a more standardised data model, so are asking you to ingest the three data sets and build a standard structure so that the data may be more easily understood, related and reported on.

- Load each data set
- Produce a standard data structure for all three data sets
- Demonstrate your thinking and decisions for the chosen approach

Part Two: Reporting & Analysis
With the new standard structure they're asking you to produce a simple Excel report that provides greater insight into their data.

Your reporting structure, and content should at a minimum showcase:

- Number of activities per campaign
- Most commonly engaged activity type per campaign
- Number of activities per person

You may also add any additional insights and data points you believe may be of value.

clevertouch
love martech

3 Your Artefact

Please upload your artefact to a public git repository, or private if you'd rather and provide us access. Please do this upon starting your work, and provide us with access so we may also review your commit history as part of the review.

We will assess your artefact **one working day** before your interview, and so while you're free to continue to work on it, and we *may* take those improvements into consideration, 24 hours before your interview is the point where we will clone and begin to review your progress.

It must be – where appropriate – written in Python, and please ensure that you have the correct project configuration to allow us to easily implement any packages you have installed, run your project and we'll need a supporting README with any instructions or guidance you believe will be helpful.

4 Considerations

Please do keep in mind while working on your project, that there are a few rules and considerations we'd like you to be aware of.

- The use of AI to complete your artefact is strictly forbidden. You may utilise AI to facilitate better ideation, or conceptual thinking, but any code, logic or deliverables must be completed solely by yourself

- After the completion of your artefact, you'll be invited into our offices for a formal interview, and at this time we will be asking you to walk us through your code, and our Engineers will have a chance to ask you questions

- You may reach out to the Clevertouch team for clarification on the task, but we will be unable to provide any technical support or guidance

- You are welcome to add any further concepts, elements or processes that you believe would add value to the task. Do keep in mind the minimum delivery

- You're free to spend as much time as you'd like on this project, but overall we're expecting the deliverable to take approximately 4 hours