

## Case Study

### Snowflake, dbt & Power BI or Tableau



#### The task

In this case study we would like you to build an ELT process, bringing together data sources to enable further analysis. The result of this process will be to enable further analysis to understand whether there is a correlation between weather conditions and the number of bikes hired in New York City from Citibike.

Create an analytics warehouse, creating appropriate tables including a table with metrics for:

- Number of Trips
- Average Duration
- Distance Travelled (hint - haversine function)
- etc.

Integrate the Citibike and weather data to enable investigation of the correlation between the weather and the number of bikes based on the weather conditions - Rain, Snow, Clouds or Clear?

The outcome of this case study should be an ELT process in dbt and Snowflake, and if you have time visualisations in Power BI or Tableau (whichever you prefer).

It's up to you on how you would like to present your results, but something along the lines of tables created and populated in Snowflake, your dbt project, and a dashboard to highlight a story, if you can find one.



#### What we are looking for

We will be looking for the following minimum outputs;

- Show how you handle a slightly vague request like this. Feel free to ask me any questions but in general we are the customer, so we don't know what we really want and need to be guided.
- Gain a working knowledge and/or show your skills in Snowflake, dbt and Power BI or Tableau.
- Show you understand how to create a database, schemas and tables or views in Snowflake in a clear and logical way.

- Show you understand what an ELT process is, what are the important components, and how to create one in a clear and logical way.
- Present to me what is built, who you saw as the end user, the approach you took, how you built the database and ELT process, and possibly dashboards, and the thought processes you went through in delivering the outcome.
- When you present, imagine you are talking to us as the client, explain the benefits of using the transformations you've built?
- Is there a correlation between the number of bikes and weather?



## The Data

The Citybike data is available in CSV format here:

<https://www.citibikenyc.com/system-data>, but you can also load it directly from the S3 bucket via an external stage, here:  
s3://snowflake-workshop-lab/citibike-trips

The New York City weather data is in JSON format and can be loaded directly from the S3 bucket via an external stage here:  
s3://snowflake-workshop-lab/weather-nyc



## Software

You can download full trial versions of the software below:

- Snowflake: 30 day trial [here](#)
- dbt: [here](#)
- Power BI - [here](#)
- Tableau: 14 day trial [here](#)

You can also access training material in the links below:

- Snowflake: [Zero to Snowflake](#); [Snowflake Training Resources](#)
- dbt: [dbt Cloud Fundamentals](#)
- Power BI: [Power BI Learning Paths](#)
- Tableau: [Free Tableau Training](#)



**Good luck!**

We hope you enjoy using some potentially new tools, and the opportunity to showcase your abilities.

Any questions give me a shout - [sarah@firn.nz](mailto:sarah@firn.nz)

Your recruiter will be in touch with a time to present.