

Project Checklist

- *Clarify your Motivation.*

- ☐ What is your goal for this project?
- ☐ What end use cases should your data be prepped for ?
- ☐ Is it for an analytics table, app back end, source of truth table?
- ☐ Decide on ETL or ELT?
- ☐ Tools you are looking to use in the project.

Skills to Showcase: Understanding business needs, business processes and requirements gathering.

- *Gather, explore and assess the data.*

- ☐ Identify and gather data with at least two different data formats (csv, json & xlsx).
- ☐ Gather data with at least one million rows.

- *Model the data*

- ☐ Plan out the Conceptual Model
 - ☐ Organise the business rules and define entities.
 - ☐ Map out the business processes
 - ☐ Build the Entity Relationship (ER) diagram.
 - ☐ Determine the schema (eg: snowflake or star)
 - ☐ Decide on Fact and Dimension tables.
- ☐ Refine using the logical Model
 - ☐ Define the implementation steps for the data model
 - ☐ Add attributes and define integrity constraints.
 - ☐ Build logical model, irrespective of the database type
- ☐ Create the Physical Model

Skills to Showcase: Defining data model based on a problem statement.

- *Build data pipelines*

- ☐ Modularize physical model into functions
- ☐ Build your Extract, Transform and Load(ETL) pipelines.
- ☐ Prepare data dictionaries
- ☐ Data Quality checks
 - ☐ Integrity constraints on the dataset
 - ☐ Unit tests to make sure that the data is meeting the business requirements.
 - ☐ Count checks to ensure completeness.

Skills to showcase : Knowledge on software engineer skills.

- *Deployment*

- ☐ Build it out to a dashboard and/or
- ☐ Data Visualisation