DATA ENGINEER
CASE SCENARIO

Ver 1.0

Prepared for

Nawatech

By

PT. Nawa Darsana Teknologi

Nawatech

Company Details

PT. Nawa Darsana Teknologi
(Nawatech)
Gedung Office 8, Lantai 18 Unit A, SCBD Lot. 28
Jl Jend Sudirman Kav. 52 - 53
Senayan - Kebayoran Baru
Jakarta Selatan 12190
Telp. 021- 29552754

The information contained in this document represents the current view of Nawatech, on the issues discussed as of the date of publication. Because Nawatech must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Nawatech, and Nawatech cannot guarantee the accuracy of any information presented after the date of publication.

This reviewer's guide is for informational purposes only. NAWATECH MAKES NO WARRANTIES, EXPRESS OR IMPLIED, AS TO THE INFORMATION IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Nawatech.

Nawatech may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Nawatech, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Table of Contents

1 Case 1

From the **motorcycle.csv** file there are 8 columns of summarized dummy data transactions in motorcycle sale and distribution in Indonesia:

Sale_id	ID of individual rows / transactions
Purchase_date	The date of transactions
Motorcycle_name	The name of motorcycle
Motorcycle_group	The group of motorcycle
Dealer_origin	Origin city of dealer of transactions
Price	Motorcycle pricing
Qty	The quantity of unit
Total	Total payment of transactions per unit

Task 1

Analyze the data, is there any anomalies in the data? How do you analyze it?

Task 2

If there are any anomalies in the data, remove them. Then, create Power BI Dashboard that represents:

1. Comparison between each motorcycles type
2. Comparison between each motorcycle group
3. Trend of sales each motorcycles type
4. Trend of sales each motorcycle group
5. Create your own reports based on the data, what insights that you can build?

2 Case 2

Create a high level architecture flow/diagram of ETL/ Pipeline Data Ingestion consists of:

Data sources: Database (SQL Server or Oracle or MySQL or else..), Files (CSV, XLSX, JSON, AVRO, ORC, PARQUET, or ELSE)

Data warehousing: SQL Server or Cloudera or Big Query or else..

Data visualization: Power BI, Tableau, Looker Studio, or else...

You can create it within on-premises environments or cloud environments or even hybrid environments (both on-premises and cloud In a single flow/diagram). Please give your explanation about the flow/diagram.