Data Dictionary:

ERP_project.xlsx is a database, which contains 10 different datasets, the following is a basic description of all the datasets and their dimensionalities.

1. Dataset: MaintenanceData

Description: Contains historical records of maintenance activities, including equipment details, actions taken, and sensor readings.
Dimensionality: 500 rows, 14 columns

Column Name	Description	Data Type
MaintenanceID	Unique identifier for each maintenance record	Categorical
Date	Date when the maintenance occurred	Date
EquipmentID	Unique identifier for the equipment	Categorical
EquipmentType	Type of equipment (e.g., Conveyor, Pump)	Categorical
TechnicianID	Unique identifier for the technician who performed maintenance	Categorical
MaintenanceType	Type of maintenance (e.g., Preventive, Corrective)	Categorical
IssueDescription	Description of the issues encountered during maintenance	Text
ActionsTaken	Description of the actions performed during maintenance	Text
DowntimeHours	Number of hours the equipment was down for maintenance	Numerical
MaintenanceCost	Cost incurred during the maintenance	Numerical
NextScheduledMaintenance	Date when the next maintenance is scheduled	Date
SensorReading_Temperature	Temperature reading from the equipment during maintenance	Numerical
SensorReading_Vibration	Vibration level reading from the equipment	Numerical
Comments	Additional comments or observations	Text

2. Dataset: InventoryData

Description: Contains details about the inventory items, including stock levels, reorder points, and supplier information.

Dimensionality: 200 rows, 14 columns

Column Name	Description	Data Type
ItemID	Unique identifier for each inventory item	Categorical
ItemName	Name or description of the item	Text
Category	Category/type of item (e.g., Spare Part, Consumable)	Categorical
QuantityOnHand	Current stock level of the item	Numerical
ReorderLevel	Stock level at which reordering should occur	Numerical
ReorderQuantity	Quantity of the item to reorder	Numerical
LeadTimeDays	Number of days between ordering and receiving the item	Numerical
SupplierID	Unique identifier for the supplier	Categorical
UnitCost	Cost per unit of the item	Numerical
UsageRateMonthly	Average monthly usage of the item	Numerical
LastOrderDate	Date when the item was last ordered	Date
NextExpectedDelivery	Date when the next delivery is expected	Date
StorageLocation	Location where the item is stored (e.g., Warehouse A)	Categorical
Comments	Additional comments or notes	Text

3. Dataset: OperationalData

Description: Tracks data related to operational performance, including production volumes, downtime, and energy consumption

Dimensionality: 1,000 rows, 14 columns

Column Name	Description	Data Type
OperationID	Unique identifier for each operational record	Categorical
Date	Date of the operation	Date
Shift	Operational shift (e.g., Morning, Evening)	Categorical
EquipmentID	Unique identifier for the equipment used	Categorical
ProductionVolume	Quantity of products produced during the operation	Numerical
OperationalHours	Number of hours the equipment was operational	Numerical
DowntimeHours	Number of hours the equipment was down during the operation	Numerical
EnergyConsumption_kWh	Energy consumption (in kWh) during the operation	Numerical
DefectRate	Percentage of defective products produced	Numerical
LaborHours	Total labor hours for the operation	Numerical
EquipmentEfficiency	Efficiency of the equipment during operation (as a percentage)	Numerical
Temperature	Temperature reading during the operation	Numerical
Pressure	Pressure reading (if applicable)	Numerical
Comments	Additional notes or comments about the operation	Text

4. Dataset: UserActivityData

Description: Logs of user activities, including login times, actions performed, and IP addresses.
Dimensionality: 300 rows, 10 columns

Column Name	Description	Data Type
ActivityID	Unique identifier for each user activity	Categorical
Timestamp	Date and time of the activity	Date
UserID	Unique identifier for the user	Categorical
UserName	Username of the user who performed the activity	Categorical
UserRole	Role of the user (Admin, Technician, etc.)	Categorical
ActionType	Type of action performed (e.g., Login, DataEdit, Logout)	Categorical
ActionDetails	Detailed description of the action	Text
IPAddress	IP address of the user	Categorical
SuccessFlag	Whether the action was successful (Yes, No)	Boolean
Comments	Additional comments or notes	Text

5. Dataset: MaintenanceLogs

Description: Unstructured text data related to maintenance logs, which is used for NLP analysis.
Dimensionality: 200 rows, 5 columns

Column Name	Description	Data Type
LogID	Unique identifier for each maintenance log	Categorical
Date	Date of the maintenance log	Date
EquipmentID	Unique identifier for the equipment	Categorical
TechnicianID	Unique identifier for the technician	Categorical
LogText	Unstructured text description of the maintenance log	Text

6. Dataset: InventoryTransactions

Description: Tracks inventory transactions, including stock received, issued, or adjusted.
Dimensionality: 1,000 rows, 9 columns

Column Name	Description	Data Type
TransactionID	Unique identifier for each inventory transaction	Categorical
Date	Date of the transaction	Date
ItemID	Unique identifier for the inventory item	Categorical
TransactionType	Type of transaction (Received, Issued, Adjusted)	Categorical
Quantity	Quantity involved in the transaction	Numerical
BalanceAfterTransaction	Stock level after the transaction	Numerical

Reference	Reference number (e.g., Purchase Order, Work Order)	Categorical
Location	Location of the transaction (e.g., Warehouse)	Categorical
Comments	Additional comments or notes	Text

7. Dataset: EquipmentDetails

Description: Metadata about the equipment, including installation and maintenance schedules.
Dimensionality: 50 rows, 12 columns

Column Name	Description	Data Type
EquipmentID	Unique identifier for each piece of equipment	Categorical
EquipmentName	Name of the equipment	Categorical
EquipmentType	Type/category of equipment (e.g., Pump, Conveyor)	Categorical
InstallationDate	Date when the equipment was installed	Date
Manufacturer	Name of the manufacturer	Categorical
ModelNumber	Model number of the equipment	Categorical
SerialNumber	Serial number of the equipment	Categorical
Location	Physical location of the equipment	Categorical
WarrantyExpiryDate	Date when the equipment's warranty expires	Date
LastMaintenanceDate	Date of the last maintenance performed	Date
NextMaintenanceDue	Date when the next maintenance is due	Date
Comments	Additional comments or observations	Text

8. Dataset: SupplierData

Description: Details about suppliers, including contact information and preferred items.
Dimensionality: 20 rows, 13 columns

Column Name	Description	Data Type
SupplierID	Unique identifier for each supplier	Categorical
SupplierName	Name of the supplier	Categorical
ContactPerson	Contact person for the supplier	Categorical
PhoneNumber	Supplier's contact phone number	Categorical
Email	Supplier's email address	Categorical
Address	Supplier's physical address	Categorical
City	City of the supplier	Categorical
State	State of the supplier	Categorical

PostalCode	Postal code of the supplier	Categorical
Country	Country of the supplier	Categorical
PreferredItems	Items that the supplier typically provides	Text
PaymentTerms	Payment terms agreed upon with the supplier	Text
Comments	Additional comments or notes	Text

9. Dataset: UserData

Description: Stores user credentials and roles for the system's authentication mechanism.
Dimensionality: 15 rows, 12 columns

Column Name	Description	Data Type
UserID	Unique identifier for each user	Categorical
UserName	Username for logging in	Categorical
PasswordHash	Password in plain text (for now)	Categorical
FirstName	First name of the user	Categorical
LastName	Last name of the user	Categorical
Email	Email address of the user	Categorical
PhoneNumber	User's phone number	Categorical
UserRole	Role assigned to the user (Admin, Manager, Technician)	Categorical
DateCreated	Date when the user account was created	Date
LastLoginDate	Date and time of the last login	Date
ActiveStatus	Whether the account is active (Active, Inactive)	Boolean
Comments	Additional notes or comments	Text

10. Dataset: ShiftData

Description: Provides information about operational shifts and supervisors.
Dimensionality: 3 rows, 6 columns

Column Name	Description	Data Type
ShiftID	Unique identifier for each shift	Categorical
ShiftName	Name of the shift (e.g., Morning, Night)	Categorical
StartTime	Time when the shift starts	Time
EndTime	Time when the shift ends	Time
SupervisorID	UserID of the supervisor for the shift	Categorical
Comments	Additional comments or notes	Text