**HALL TICKET NUMBER: 2403A51365**

**NAME: RAYAPUDI HARITHA**

**BATCH: 24BTCAICSB14**

**AssignmentNumber:17.1**

**Task 1 – Customer Data Cleaning  
Task:  
Use AI to generate a Python script for cleaning a customer dataset.  
Instructions:  
• Handle missing values in columns (age, income, city).  
• Remove duplicate records.  
• Standardize text values (e.g., city names in lowercase).  
• Encode categorical variables (gender, city).  
Expected Output:  
• A cleaned Pandas DataFrame ready for analysis.**

**A screenshot of a computer program

AI-generated content may be incorrect.**

**A close-up of a computer error

AI-generated content may be incorrect.**

**A close up of text

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer error

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer code

AI-generated content may be incorrect.**

**A screen shot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

Task 2 – Sales Data Preprocessing  
Task:  
Preprocess sales transaction data using AI assistance.  
Instructions:  
• Convert date columns into datetime format.  
• Create new features (month, year, day of week).  
• Normalize sales amount column.  
• Detect and handle outliers in transaction amounts.  
Expected Output:  
• A transformed dataset with time-based features and normalized  
values

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A close-up of a text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Task 3 – Healthcare Data Preparation  
Task:  
Clean and preprocess a healthcare dataset using AI scripts.  
Instructions:  
• Handle missing values in blood\_pressure, cholesterol.  
• Encode categorical features like disease\_status (Yes/No).  
• Scale numerical features (min-max or standard scaling).  
• Split dataset into training and testing sets.  
Expected Output:  
• A preprocessed dataset ready for machine learning models

A computer screen with text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A computer code with black and white text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Task 4 – Real-Time Application: E-commerce Product Data

Cleaning

Scenario:

An e-commerce company has product catalog data with

inconsistencies.

Requirements:

• Standardize product names and categories.

• Fill missing prices with median values.

• Remove duplicate entries.

• Normalize product ratings (0–5 scale).

• Generate a summary report of data quality improvements (before

vs after cleaning).

Expected Output:

• A cleaned dataset with improved consistency and usability for

analytics.

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a phone

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.