# LAB ASSESMENT 17.4

TASK 1: Use AI to generate a Python script for cleaning an employee dataset.

PROMPT: please create a sample employee dataset having employee id ,name, age,dept,payscale,qualification,joining date with some missing values in python

Sample Employee Dataset:

A screenshot of a computer program

AI-generated content may be incorrect.PROMPT 2: please clean the data set ie fill the missing values

A screenshot of a computer

AI-generated content may be incorrect.

TASK 2: Use AI to generate a script for preprocessing a sales transaction dataset.

A screenshot of a computer

AI-generated content may be incorrect.PROMPT 1: generate a sales transaction dataset using python

Prompt 2 : from sales dataset convert the transactions dates to proper date time format,create a new column Month-Year from the transaction date,remove rows with negative,zero,invalid transactions, and Normalize the "transaction\_amount" column using Min-Max scaling.

A screenshot of a computer

AI-generated content may be incorrect.

TASK 3: Use AI to generate a script for cleaning healthcare patient records**.**

PROMPT 1: generate a python based data set on healthcare patient records

A screenshot of a computer

AI-generated content may be incorrect.

PROMPT 2: IN THE dataset fill missing values in numeric columns,standardize the units(cm to m),Correct inconsistent categorical labels ,and drop irrelevant columns regarding the datasetA screenshot of a computer

AI-generated content may be incorrect.

TASK 4: Use AI to write a script to preprocess a social media text dataset.

PROMPT 1: generate me a social media text dataset

A screen shot of a computer

AI-generated content may be incorrect.

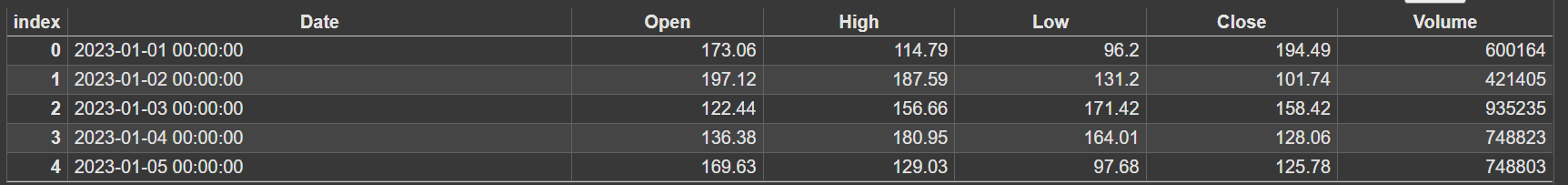
PROMPT 2: from the dataset,remove special characters,urls,emojis from text,and convert them into lowercase and Tokenize and remove stopwords. and lemmatize the words

A screen shot of a computer

AI-generated content may be incorrect.

TASK 5: Use AI to create a preprocessing script for a financial dataset.

PROMPT 1: generate a **Financial Dataset Feature Engineering**

****

PROMPT2: from the data set handle the missing values in stock price and volume, create a new feature of moving average (7-day, 30-day)• Normalize continuous variables using StandardScaler • Encode categorical columns (sector, company\_name).

A screenshot of a computer program

AI-generated content may be incorrect.