# LAB ASSIGNMENT-17

Task-1:

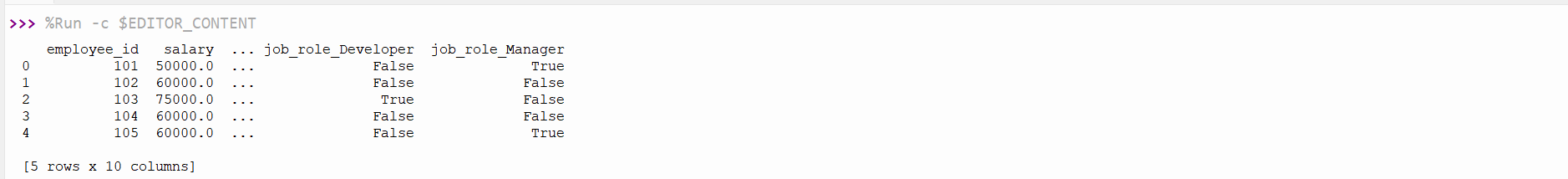
Prompt:

Create a Python script to clean an employee dataset: fill missing values (salary, department, joining\_date), convert joining\_date to datetime, standardize department names (e.g., 'HR', 'hr', 'Human Resources' → 'HR'), and encode categorical variables (department, job\_role).

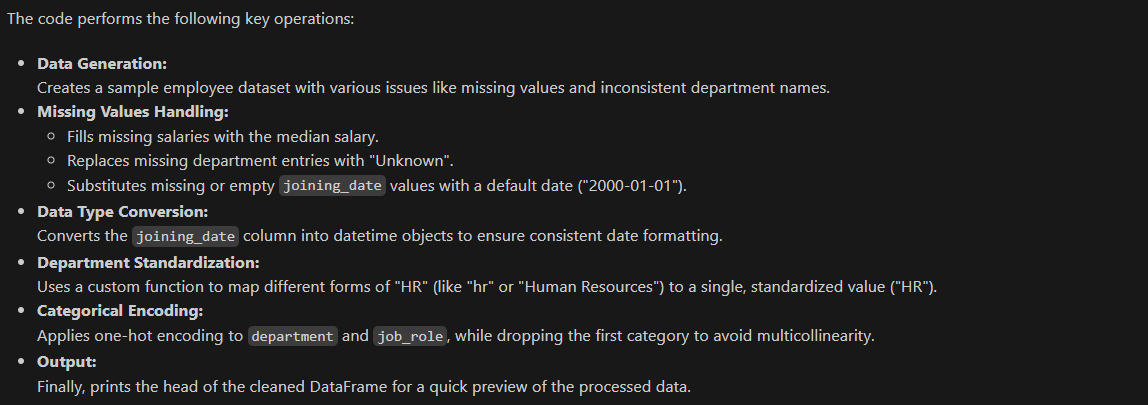
Code and Output:







Explanation:



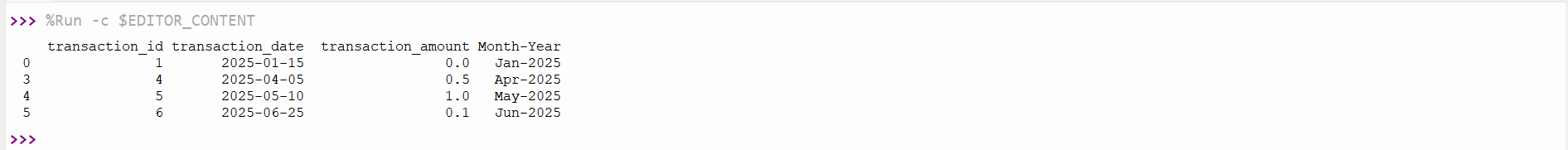
Task 2:

Prompt:

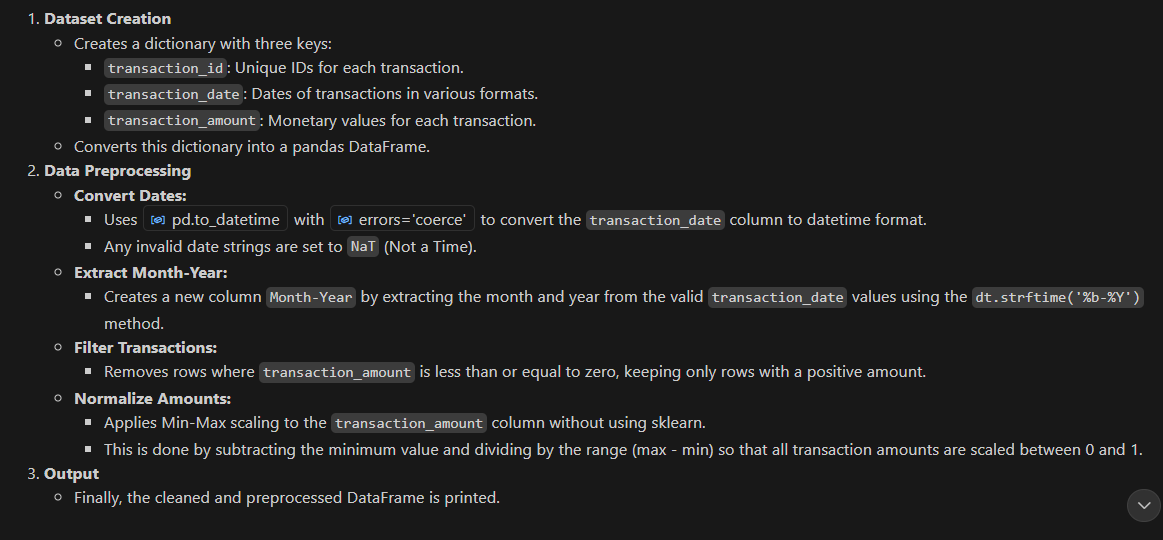
Generate a Python script to preprocess a sales dataset: convert transaction\_date to datetime, add Month-Year column, remove rows with non-positive transaction\_amount, and apply Min-Max scaling to transaction\_amount.

Code and Output:





Explanation:



**Task 3**:

Prompt:

Write a Python script to clean healthcare patient data: fill missing numeric values with column mean, convert height from cm to meters, standardize gender labels (e.g., 'M', 'Male', 'male' → 'Male'), and drop the patient\_id column.

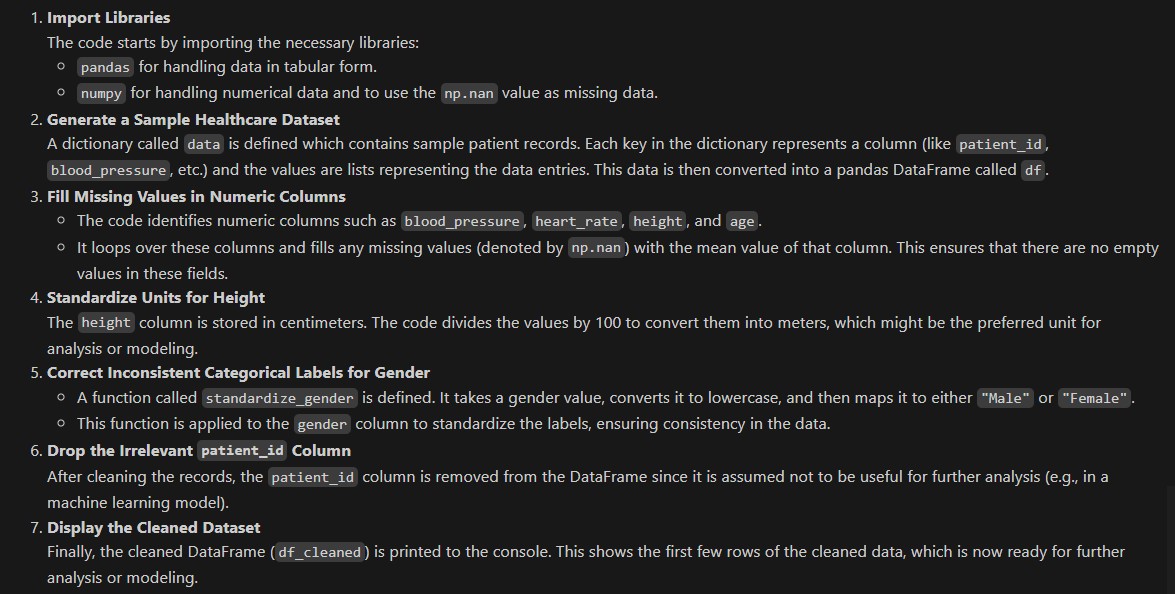
Code and Output:







Explanation:

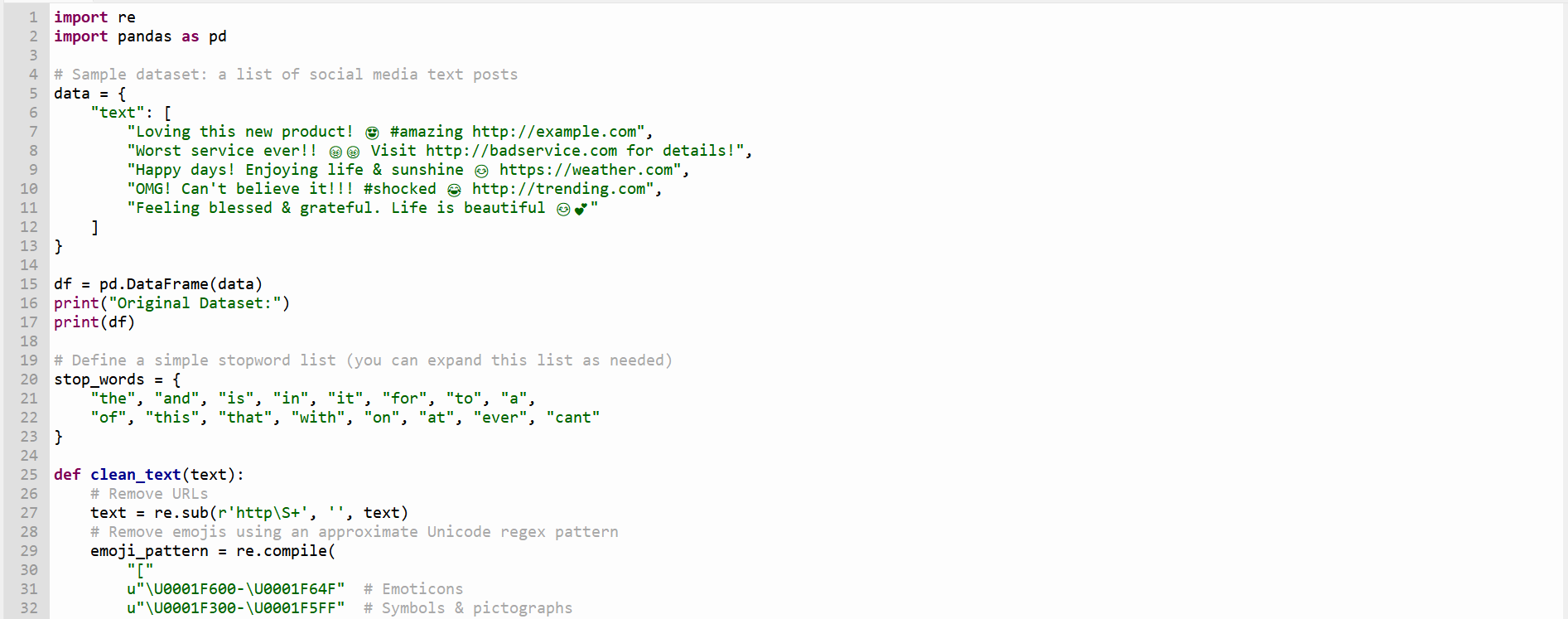


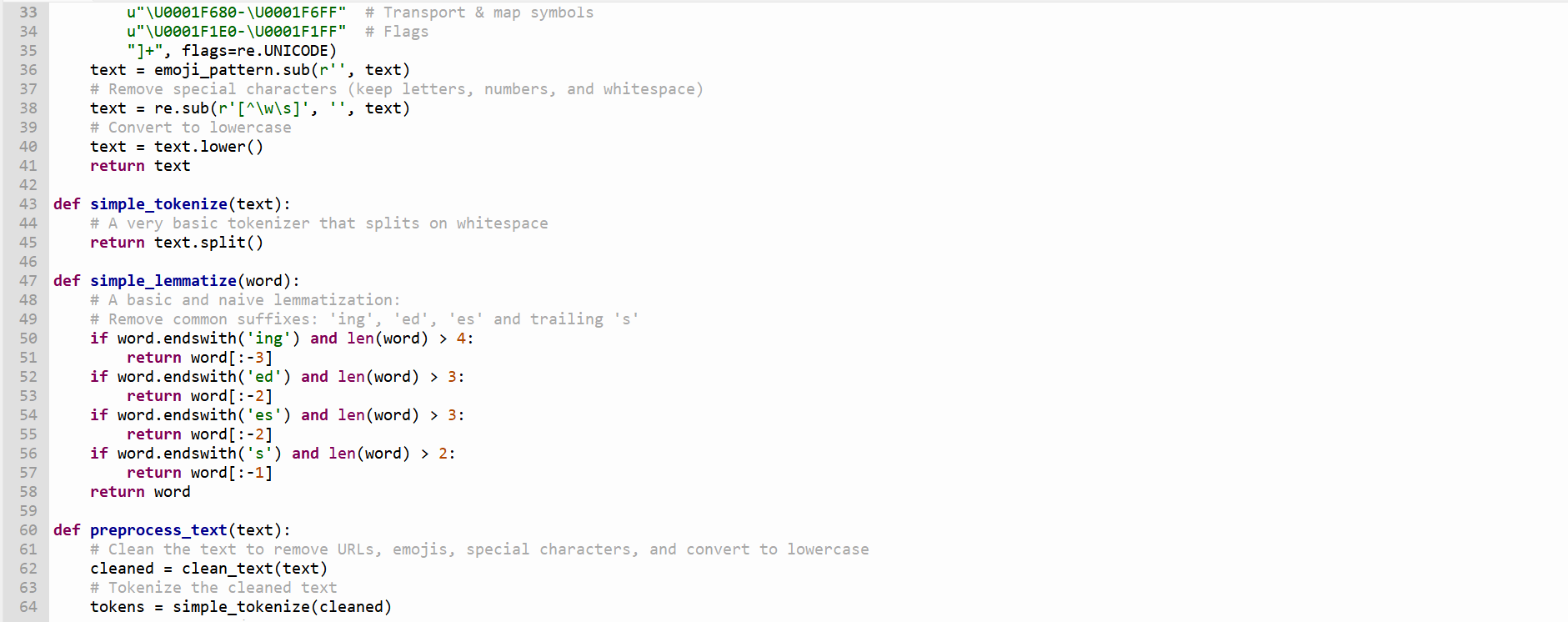
Task 4:

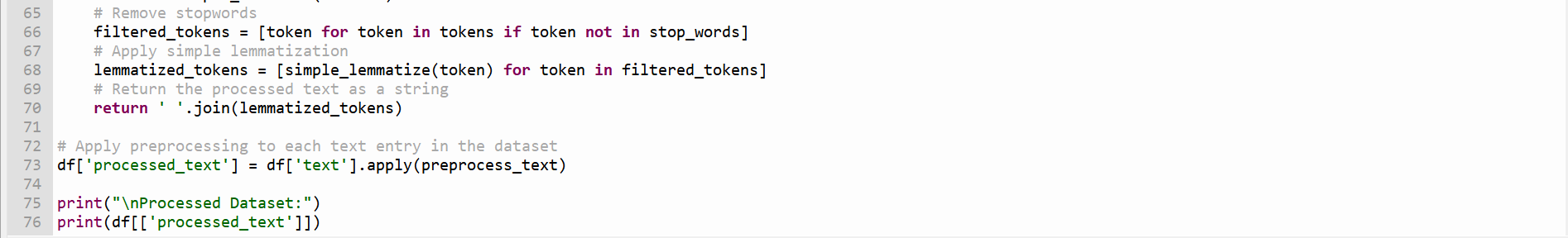
Prompt:

Write a Python script to clean social media text data: remove special characters, URLs, and emojis; convert text to lowercase; tokenize and remove stopwords; and apply lemmatization to prepare for sentiment analysis.

Code & Output:

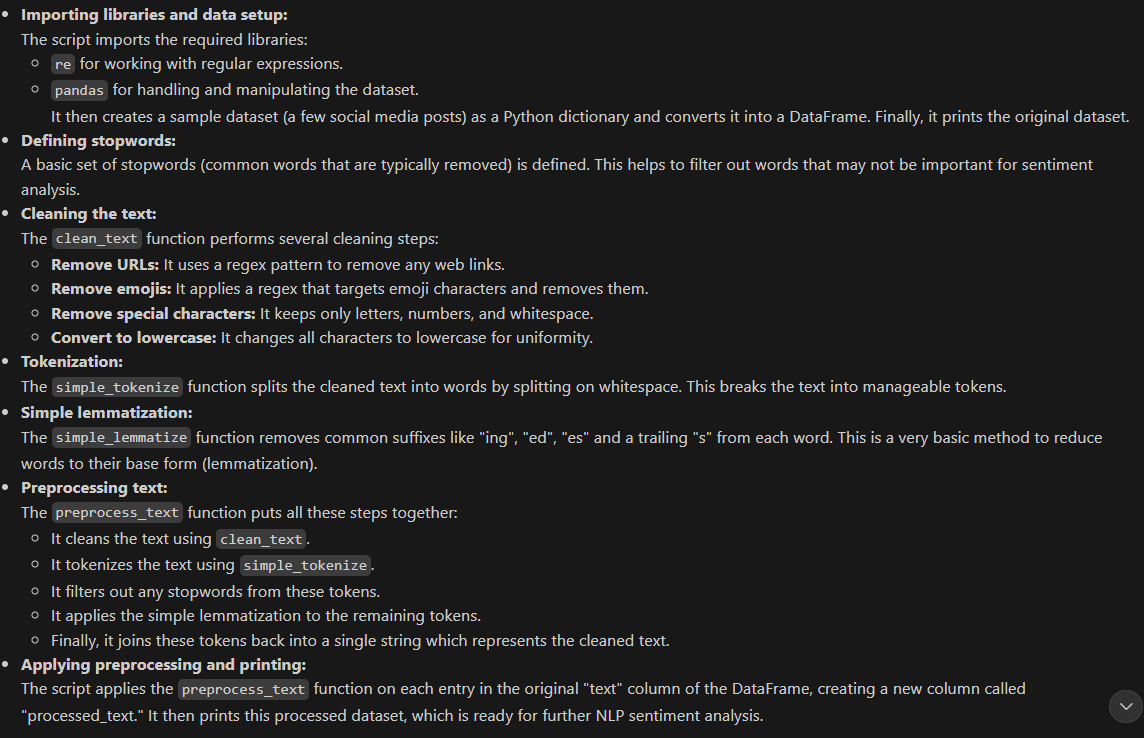








Explanation:

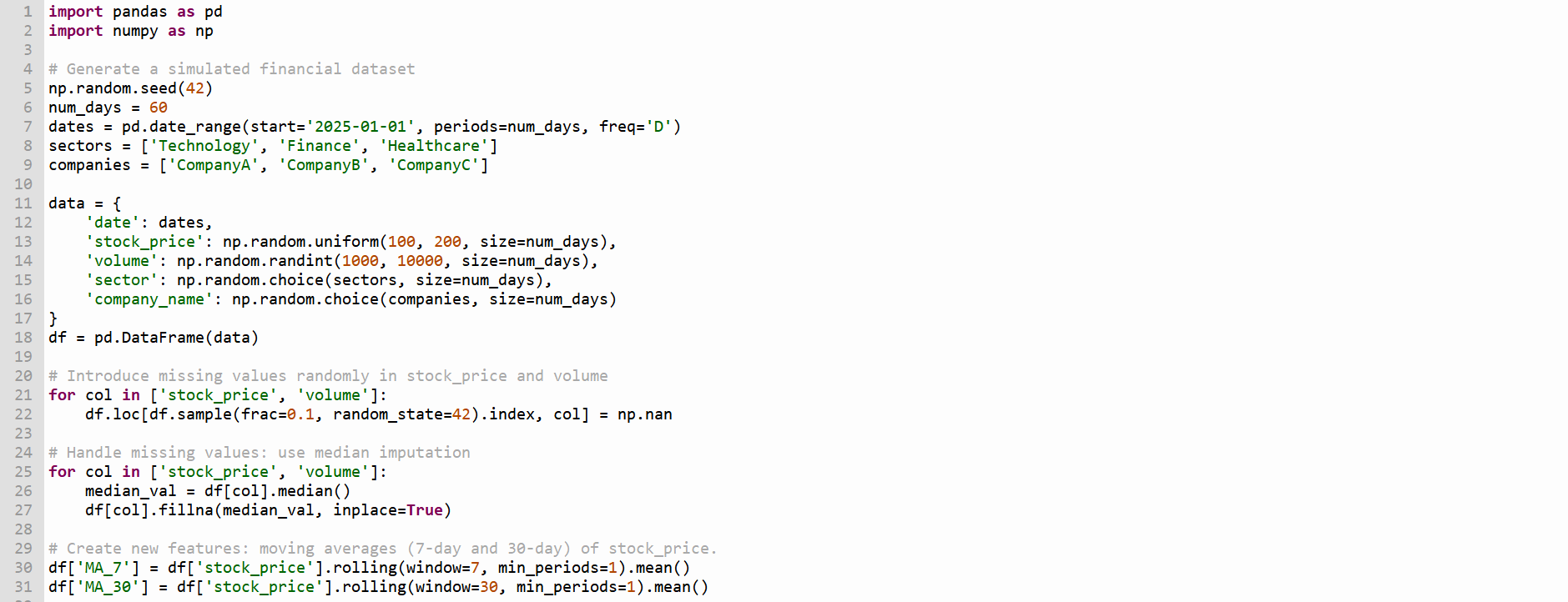


Task 5:

Prompt:

Write a Python script to preprocess a financial dataset: fill missing values in stock price and volume, create 7-day and 30-day moving averages, normalize continuous columns with StandardScaler, and encode categorical columns like sector and company\_name.

Code & Output:







Explanation:

