```
# Start
# Import spaCy (Python's natural language processing library).
# Import pandas to read in the csv file and to create a data frame of the dataset.
# Import spacytextblob from the TextBlob library for polarity and sentiment analysis.
# Import WordCloud, matplotlib.pyplot and defaultdict for analysing the effectiveness of identifying the positive and
# negative words for the sentiment analysis.
# Load the simple english language model to enable natural language processing tasks.
# Read in the amazon product reviews file and call the data frame 'products'.
# Creating a function to preprocess the reviews from the 'reviews.text' column for lemmatisation.
# Create a function to get a polarity score of the reviews.
# Create a function to get polarity and subjectivity scores of the reviews.
# Create a function to get the sentiment in words based off the polarity scores of the reviews.
# Select only the relevant columns for analysis.
# Check for any duplicate reviews and drop rows with duplicate reviews.
# Remove any rows with null values for a complete dataset as long as it does not affect the integrity and representativeness
# of the dataset largely and assign this to a new DataFrame called 'clean_data'.
# Ensure that all values are easily understood under each column and adjust otherwise.
# Apply the preprocess function to the 'reviews.text' column to preprocess the text and store the preprocessed reviews
# in a new column called 'processed_reviews'.
# Apply the function to get the polarity scores to the preprocessed reviews and store in a new column.
# Apply the sentiment score function to get the tuples of polarity and subjectivity on the preprocessed reviews and store
# in a new column.
# Apply the sentiment function on the processed reviews to identify whether the sentiment of the reviews are
# positive/negative/neutral and store in a new column.
# In order to compare the effectiveness of these, compare the sentiments in words against the doRecommend column
# and the reviews.title column to easily identify if the sentiments identified are along the same lines as the
# reviews.title is a summary of what the review represents.
# To further assess the effectiveness of this analysis, conduct the sentiment analysis on the 'reviews.title' column
# after preprocessing to check whether the sentiment in words matches our own intuitive understanding of the sentiment.
# Select two reviews that have been identified to have the same sentiment and check their similarity scores.
# Select two reviews at random and check their similarity scores and assess whether it is in alignment
# to the polarity scores.
# Initialise positive and negative words dictionaries to store positive and negative words identified from the
# processed reviews.
# Create a function to enable the above.
# Use WordCloud and matplotlib.pyplot to create a visualisation of the positive and negative words identified.
# End
```