1. The goal of EDA is to first look at the data structure and organization to see whether there are inputs that have to be replaced or streamlined or cleaned. Next, the data is evaluated for content in order to identify relevant information and variables of interest. Finally, key features are identified among categories or variable that are relevant to address the goals of the study. New features could be introduced and engineered.
2. A. In a costumer product review on a 1-5 scale, some reviews might be submitted with either no number selected or multiple numbers selected. These rows would have to be omitted from the dataset. I would also need to know whether this data is anonymous and how well the survey represents the product. The two key components are how reliable the answers are and how valid the answers to score the product.

B. To investigate the customer reviews, I would clean the data of non- or multi-input rows and next choose the most relevant survey questions that 1) a person would answer honestly 2) are not biased to any particular sample of the population, if such biased exists in the sample group 3) focus on questions that are not ambiguous, confusing, or subject to interpretation. I might also group the survey information that might evaluate broader categories of a data feature. This might be accomplished by adding a new column that assigns groups of survey information into particular variables.

C. Data could be categorized as how easy to e-commerce site is to operate. Another feature could evaluate the overall sentiment on the price of the product. One could also see if customer answers seem to follow a certain pattern of answers. These patterns could identify relationships between certain customer behaviors and preferences. The data might also immediately identify elements of the e-commerce company that are complete not working well.