

Using the space provided, please address the following items:

1. Provide a summary of the different descriptive statistics you looked at and WHY.
 - The dataset date range from April 2005 to Jan 2022. The timespan of data is important as it enables the tracking of improvement over time, which gives key insight. Besides, we can make sure that we are not using outdated data.
 - Descriptive statistics such as mean, median, minimum, maximum and mode for each column is calculated. It allows to check the quality of data by detecting potential outliers, collection or encoding errors.
2. Submit 2-3 key points you may have discovered about the data, e.g. new relationships? Aha's! Did you come up with additional ideas for other things to review?
 - Discovered that cities with higher number of reviews are major cities
 - Most of the coffee company with reviews operating from morning till afternoon.
 - Seems like there is relationship between review count and average rating. The higher the review count, the higher the average review rating.
3. Did you prove or disprove any of your initial hypotheses? If so, which one and what do you plan to do next?
 - Data shows urban city having higher reviews than rural city, which prove the initial hypothesis. Further, need to prove that is higher reviews count brings higher rating.
 - Data shows coffee shop with higher reviews generally operate from morning to afternoon, this somehow proved the hypothesis that people love to buy coffee at that period. Further, need to prove the period brings higher rating.
4. What additional questions are you seeking to answer?
 - I would like to know what text the people are commenting for high rating coffee company. Based on the repeated words, we might conclude that is the key feature for a coffee company.