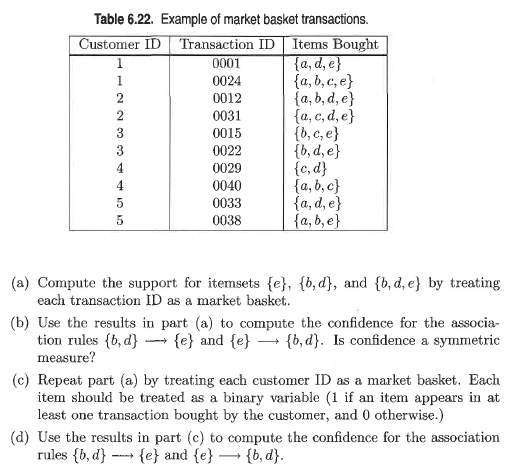
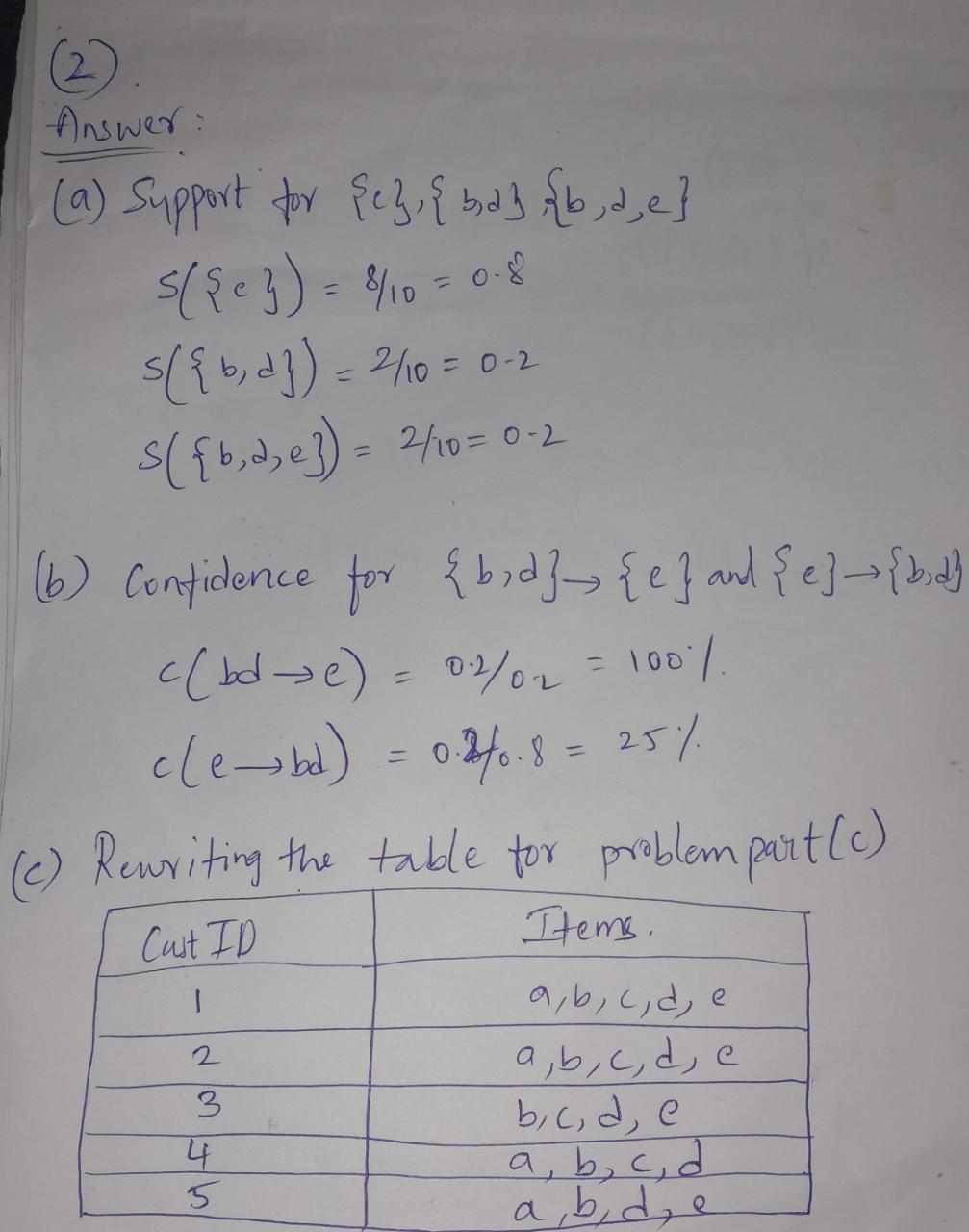
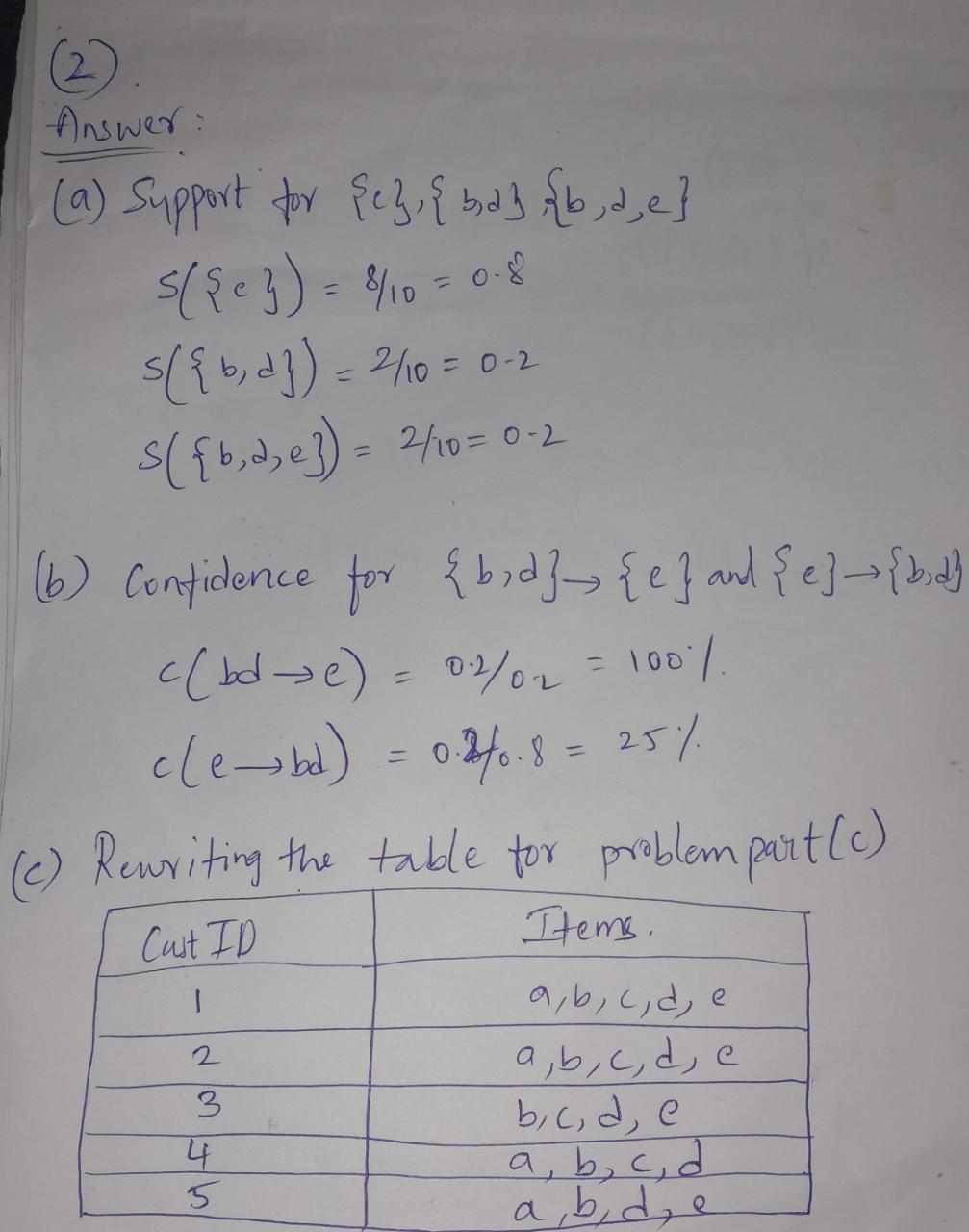
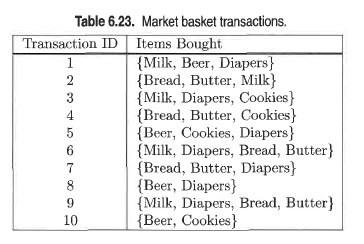
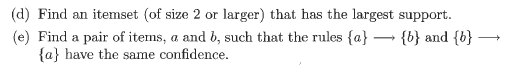
Data Mining Assignment 3

1. Read Chapter 6 (only sections 6.1 and 6.7).  
     
   **2) Do Chapter 6 textbook problem #2 (parts a,b,c,d only) on page 404.**





**3) Do Chapter 6 textbook problem #6 (parts d,e only) on page 406.**



**(d) Answer:** {Butter, Butter}

**(e) Answer:** (Beer, Cookies) or (Bread, Butter).

**4) Using the data at** [**www.stats202.com/more\_stats202\_logs.txt**](http://www.stats202.com/more_stats202_logs.txt) **and treating each row as a "market basket" compute the support and confidence for the rule ip=65.57.245.11 → "Mozilla/5.0 (X11; U; Linux i686 (x86\_64); en-US; rv:1.8.1.3) Gecko/20070309 Firefox/2.0.0.3".**

**State what the support and confidence values mean in plain English in this context.**

**Answer:**

According to the data, support(65.57.245.11) = 5021

Support = 1619/14803 = 0.10936 = 10.9%

Confidence = 1619/5021 = 0.322 = 32.2%

**Support:** it is the number of times items appear in the data.

**Confidence**: It can be interpreted as likelihood of occurring both IP address and link together.