

CEE 412/CET 522 Transportation Data Management and Visualization

Assignment 4 | Winter 2020

Please answer the following seven questions based on Exercise III Part 2.

Please **list necessary R/SQL code** and **the results/outputs of the models/parameters** when you answer the following questions.

1. **What variables are significant at $\alpha = 0.05$? [5 points]**
2. **What is the dispersion parameter and AIC? [5 points]**
3. **Based on the AIC, which model would you choose? Explain why. [5 points]**
4. **Is the new NB model better or worse than the original NB model (in terms of AIC)? [5 points]**
5. **Show the top 15 road segments in terms of ARP. [5 points]**
6. **Would a road segment which is predicted (by NB model) to have a very high accident count be likely to be identified as a high safety treatment priority using this method? Why or why not? [5 points]**

(Note that the answer to this question is somewhat open ended, full credit will be given as long as you demonstrate a reasonable understanding of how to interpret the ARP in hotspot identification.)
7. **Show the top 3 road segments for each route in terms of ARP based on the EB results you saved in your personal database. [5 points]**

(Please list necessary SQL query and queried results.)