Mid-term 2 project

Please interpret all your results and describe every variable you use in your analysis

1.) Agresti 4.14

Data: [http://users.stat.ufl.edu/~aa/cda/data/Infection.dat (Links to an external site.)](http://users.stat.ufl.edu/~aa/cda/data/Infection.dat)

or [Infection.dat](https://uth.instructure.com/courses/63014/files/4153132?wrap=1) [Download Infection.dat](https://uth.instructure.com/courses/63014/files/4153132/download?download_frd=1)

2) Agresti 4.17

3) **Second edition** Agresti 5.3 (To answer this question, you only need to use the information provided in Table 5.9) The pdf with the question is provided below:

[Agresti\_CDA\_EX5d3.pdf](https://uth.instructure.com/courses/63014/files/4153133?wrap=1) [Download Agresti\_CDA\_EX5d3.pdf](https://uth.instructure.com/courses/63014/files/4153133/download?download_frd=1)

4) Hosmer & Lemeshow Chapter 4: Question 5 (burn data supplied below).

**Instead of 4, you can do (it is up to you)**

**4 prime)**

(a) With the burn data fit the model on Table 4.18 (H&L): death~age+tbsa+race+inh\_inj

(b) Fit the spline model  on Table 4.21

death~ns(age,knots=c(19,44.37),Boundary.knots=c(1.1,78.87))+tbsa+race+inh\_inj

Note that depending on your R or STATA function the values for the spline coefficients on table 4.21 may not be the same (that is OK) as yours. But the coefficients for the other predictors should be the same.

5) Perform a full evaluation of model in 4 or 4prime assessment for the Burn Study data. This means - choose the appropriate GOF test and apply it, conduct ONE diagnostic test, and conduct an AUC analysis that we learned.

6) Create the plot of x=AGE vs logit(pi) for a person with

tbsa=13.55; race=White; gender=Female; inh\_inj=NO;

Note that you may not need gender if gender is not in your model (e.g. the model in 4b). Hint: the shape of your figure should be similar to Figure 4.8 in H&L

Does the curve look linear?

7) Hosmer & Lemeshow Chapter 4: Question 6a. Also, please apply an appropriate goodness of fit test to the final model.

The headers for the myopia variables names are as follows (code for R):

names(myopia) = c("ID","STUDYYEAR","MYOPIC","AGE","GENDER","SPHEQ","AL","ACD","LT","VCD",  
"SPORTHR","READHR","COMPHR","STUDYHR","TVHR","DIOPTERHR","MOMMY","DADMY")

[Data burn1000.csv](https://uth.instructure.com/courses/63014/files/4153134/download?wrap=1) [Download Data burn1000.csv](https://uth.instructure.com/courses/63014/files/4153134/download?download_frd=1)

[Myopia for HW 6.xlsx](https://uth.instructure.com/courses/63014/files/4153143/download?wrap=1) [Download Myopia for HW 6.xlsx](https://uth.instructure.com/courses/63014/files/4153143/download?download_frd=1)