ACTSC 632 – Assignment 1 – due on June 26

For this project you will use data from the "Motorcycle" data set available in the `insuranceData` package in which it is named `dataOhlsson`. The main goal of this project is to compute multipliers for the premium based on the different rating factors and corresponding levels deemed important in the model.

The current tariff is given in Table 1 at the end of the project description. It is based on four rating factors and was established some time ago.

1. Download the data set `dataOhlsson` and briefly discuss the data found in there (e.g., how many rating factors, what are the levels for each, how is the exposure determined, etc.). If there are any problems with the data, explain how you dealt with them.

2. Use only the rating factors contained in the current tariff. Compute the exposure (in policy years), the claim frequency and the average claim severity for each rating factor (i.e., for each possible value of each rating factor). You should group the numerical rating factors into the same intervals as in Table 1.

3. Use a relative Poisson glm to determine relativities for the claim frequency, using the current rating factors. Provide a 95% confidence interval for each relativity. Comment on the overall fit of this model to the data.

4. Use a Gamma glm (with log link function) to determine relativities for the severity, still using the current rating factors. Provide a 95% confidence interval for each relativity. Comment on the overall fit of this model to the data.

5. Assess whether rating factors for the policyholder's age and sex would have a significant impact. Include an interaction term. You should group the age factor into intervals of 0-30 and 30-100.

6. Now combine your multiplier estimates for the frequency and severity data from part 3 and part 4 to get multipliers (and associated 95% confidence intervals) for the premium overall and then propose a new tariff based on your analysis. Compare the results to the old tariff.

7. Comment on any further analysis that should be considered before deciding on a final tariff.

Your report should be formatted using R Markdown. It should be professionally organized, with a brief introduction. You are writing mainly for your colleagues in the pricing unit.

You can collaborate with one other student on the assignment. Please indicate both names on your final report. Any online resources used should be properly cited.

Please submit your report in pdf format to the LEARN dropbox as well as all code used.

Table 1: Current tariff rating factors and relativities

Rating Factor	Class	Description	Relativity
Zone	1		7.768
	2		4.227
	3		1.336
	4		1.000
	5		1.734
	6		1.402
	7		1.402
Motorcycle class	1		0.625
	2		0.769
	3		1.000
	4		1.406
	5		1.875
	6		4.062
	7		6.873
Motorcycle Age	1	0-1 years	2.000
	2	2-4 years	1.200
	3	≥ 5 years	1.000
Bonus Class	1	1-2	1.250
	2	3-4	1.125
	3	5-7	1.000

In addition, you will find the following useful: an easy-to-copy version of the values listed in Table 1 is: 7.768, 4.227, 1.336, 1.000, 1.734, 1.402, 1.402, 0.625, 0.769, 1.000, 1.406, 1.875, 4.062, 6.873, 2.000, 1.200, 1.000, 1.250, 1.125, 1.000.