

# SDA Quiz R/Statistics

James Bullard

January 18, 2008

**Please do this quiz using Sweave. This will allow for easy presentation of the results!**

In the data directory there are is a file “traffic.csv”. This file consists of traffic accidents in Buenos Aires starting in 2001 and continuing for 973 days.

1. First, take a small exploratory look at the “traffic” dataset. Try to find any associations and figure out nice ways to plot some of these relationships.
2. Stratify by month and by day of the week, see if this makes a difference. Don’t forget the helpful split function.
3. Fit a normal linear regression model to the traffic data. First, fit the model directly and have a look at the residuals. Second, fit a linear model to the log transformed values.
4. Fit a poisson regression model to the data. Compare this to the linear regression model - maybe looking at different residual plots.