

Preprocesado de la información

El fichero *windscreen2.csv* contiene datos ficticios relativos al año 2011 de una aseguradora de automóviles.

Concretamente, se trata de una tabla donde cada observación corresponde a un asegurado, donde tienes información *propia* del asegurado (edad, ciudad, tipo de coche, etc.), el coste total por siniestros relativos a las lunas de su coche (INCURR_COST), el número de siniestros de este tipo que ha sufrido (N_CLAIMS) y su exposición (EXPOSURE), es decir, la fracción del año que el sujeto ha estado asegurado.

Realiza un análisis exploratorio completo de los datos, prestando especial atención a los *missing values* y los *outliers*, y trata de establecer algún tipo de relación entre el coste anual y el resto de variables por asegurado (naturalmente, excluyendo N_CLAIMS y normalizando respecto a EXPOSURE), empleando la técnica que consideres oportuna.