R HW #4 questions about the article

Brian Kim

1. The unit of observation for the Brady and Li study is the prevalence of alcohol or drugs in toxicological testing done on drivers killed within one-hour of a motor vehicle crash. The FARS dataset that provides the information at this observation level is the PERSON data file.
2. The study inclusion criteria were drivers who died within one-hour of a crash between January 1, 1999 and December 31, 2010. Only six states were included: California, Hawaii, Illinois, New Hampshire, Rhode Island, and West Virginia. Drivers with any missing data were excluded. The variable names that will be used to create filter statements in R and the variable values to keep are as follows:

|  |  |
| --- | --- |
| **Variable** | **Value** |
| PER\_TYP | 1 |
| INJ\_SEV | 4 |
| ALC\_RES | 0-94 |
| DRUG\_RES1 | 100-295, 300-395, 400-495, 500-595, 600-695, 700-795, 800-895, 900-995, 996. |
| DRUG\_RES2 | 100-295, 300-395, 400-495, 500-595, 600-695, 700-795, 800-895, 900-995, 996. |
| DRUG\_RES3 | 100-295, 300-395, 400-495, 500-595, 600-695, 700-795, 800-895, 900-995, 996. |
| STATE | 6, 15, 17, 33, 44, or 54 |
| LAG\_HRS | 0-1 |

1. Results are stratified by age category, year, year category, alcohol level, non-alcohol drug category, and sex.

|  |  |
| --- | --- |
| **Variable** | **Value** |
| AGE | This variable value is given as an integer of the person’s age at the time of the crash in years. Data until 2008 categorizes the value 97 as people 97 years or older and the value 99 as unknown. Data for 2009 has values of 0-120 as the person’s age in years and 999 as unknown. Data for 2010 is the same as 2009 except the value 998 is not reported. |
| SEX | Sex is categorical. All years have values of 1 as male, 2 as female, and 9 as unknown. Data for 2010 has the value 8 as not reported. |
| DEATH\_YR | Data from 1998-later displays the year of the death and values of 9999 are unknown. |
| ALC\_RES | This variable is an integer and gives the blood alcohol content (BAC) test result in g/dL. The values 0-93 are the actual value of the test result. Value 94 represents BAC 0.94 or greater. |
| DRUGRES1, DRUGRES2, DRUGRES3 | This variable has values that are categorical. 1993-later data has values of 100-295 as narcotic, 300-395 as depressant, 400-495 as stimulant, 500-595 as hallucinogen, 600-695 as cannabinoid, 700-795 as PCP, 800-895 as anabolic steroid, 900-995 as inhalant, 996 as other drugs. |
|  |  |
|  |  |