Reproduced analysis of FARS data

setwd("C:/Users/bkim7/R/fars\_project/")  
library(tidyr)  
library(dplyr)  
library(ggplot2)  
load("data/clean\_fars.RData")  
source("R/fars\_functions.R")

# Results

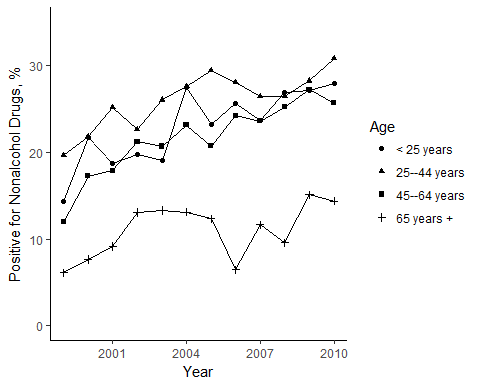
## Percentages of drivers testing positive by drug type, sex, and year group

clean\_fars %>%  
mutate(year\_cat = cut(year, breaks = c(1999, 2002, 2006, 2010),  
labels = c("1999-2002", "2003-2006",  
"2007-2010"),  
include.lowest = TRUE, right = TRUE)) %>%  
filter(!is.na(sex)) %>%  
group\_by(drug\_type, sex, year\_cat) %>%  
summarize(n\_non\_missing = sum(!is.na(positive\_for\_drug)),  
positive\_test = sum(positive\_for\_drug, na.rm = TRUE),  
perc\_positive = round(100 \* positive\_test / n\_non\_missing, 1)) %>%  
select(drug\_type, sex, year\_cat, perc\_positive) %>%  
unite(sex\_year\_cat, sex, year\_cat) %>%  
spread(sex\_year\_cat, perc\_positive) %>%  
knitr::kable(col.names = c("Drug type", "F 1999-2002",  
"F 2003-2006", "F 2007-2010",  
"M 1999-2002", "M 2003-2006",  
"M 2007-2010"))

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Drug type | F 1999-2002 | F 2003-2006 | F 2007-2010 | M 1999-2002 | M 2003-2006 | M 2007-2010 |
| Alcohol | 26.4 | 24.3 | 27.1 | 43.2 | 42.9 | 43.3 |
| Cannabinoid | 2.8 | 5.7 | 7.3 | 5.8 | 10.3 | 11.8 |
| Depressant | 3.4 | 3.8 | 4.8 | 2.0 | 2.5 | 3.2 |
| Narcotic | 4.2 | 4.9 | 7.0 | 2.2 | 3.4 | 4.0 |
| Other | 5.6 | 6.6 | 7.2 | 4.3 | 4.5 | 4.2 |
| Stimulant | 7.2 | 9.1 | 8.7 | 10.5 | 11.9 | 9.2 |

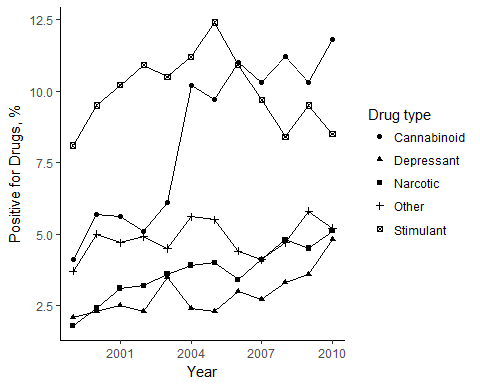
## Figure 1: Prevalence of nonalcohol drugs in fatally injured drivers by year and age group

#Total percent positive for non-alcohol drugs for each year grouped by agecat  
fars\_plot1 <- clean\_fars %>%  
 filter(as.character(drug\_type) != "Alcohol") %>%  
 filter(!is.na(agecat)) %>%  
 group\_by(year, agecat, unique\_id) %>%  
 summarize(positive\_test = any(positive\_for\_drug, na.rm = TRUE)) %>%  
 ungroup() %>%  
 group\_by(year, agecat) %>%  
 summarize(perc\_positive = mean(positive\_test) \* 100)  
  
ggplot(fars\_plot1, aes(x = year, y = perc\_positive, shape = agecat)) +  
 geom\_point() +  
 geom\_line() +  
 scale\_y\_continuous(limits = c(0, 35)) +  
 theme\_classic() +  
 labs(x = "Year", y = "Positive for Nonalcohol Drugs, %", shape = "Age")



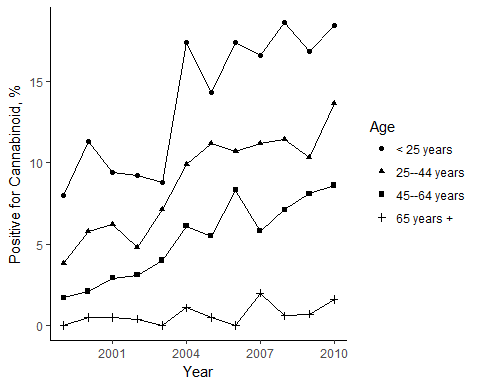
## Figure 2: Prevalence of nonalcohol drugs in fatally injured drivers by year and drug type

fars\_plot2 <- clean\_fars %>%  
 filter(as.character(drug\_type) != "Alcohol") %>%  
 filter(!is.na(agecat)) %>%  
 group\_by(year, drug\_type) %>%  
 summarize(n\_non\_missing = sum(!is.na(positive\_for\_drug)),  
 positive\_test = sum(positive\_for\_drug, na.rm = TRUE),  
 perc\_positive = round(100 \* positive\_test / n\_non\_missing, 1))  
  
ggplot(fars\_plot2, aes(x = year, y = perc\_positive, shape = drug\_type)) +  
 geom\_point() +  
 geom\_line() +  
 theme\_classic() +  
 labs(x = "Year", y = "Positive for Drugs, %", shape = "Drug type")



## Figure 3: Prevalence of cannabinoid drugs in fatally injured drivers by year and age group

fars\_plot3 <- clean\_fars %>%  
 filter(as.character(drug\_type) %in% "Cannabinoid",  
 !is.na(agecat)) %>%  
 group\_by(year, agecat) %>%  
 summarize(n\_non\_missing = sum(!is.na(positive\_for\_drug)),  
 positive\_test = sum(positive\_for\_drug, na.rm = TRUE),  
 perc\_positive = round(100 \* positive\_test / n\_non\_missing, 1))  
  
ggplot(fars\_plot3, aes(x = year, y = perc\_positive, shape = agecat)) +  
 geom\_point() +  
 geom\_line() +  
 theme\_classic() +  
 labs(x = "Year", y = "Positive for Cannabinoid, %", shape = "Age")



## Prevalence of drugs in fatally injured drivers for 1999 and 2010 by drugtype

table1 <- clean\_fars %>%  
 filter(year %in% c(1999, 2010)) %>%  
 group\_by(drug\_type, year) %>%  
 summarize(n\_non\_missing = sum(!is.na(positive\_for\_drug)),  
 positive\_test = sum(positive\_for\_drug, na.rm = TRUE),  
 percent = perc\_cis(x = positive\_test, n = n\_non\_missing)) %>%  
 select(drug\_type, year, percent) %>%  
 spread(key = year, value = percent) %>%  
 knitr::kable(col.names = c("Drug Type", "1999", "2010"))  
table1

|  |  |  |
| --- | --- | --- |
| Drug Type | 1999 | 2010 |
| Alcohol | 38.7% (36.5%, 40.9%) | 39.1% (36.7%, 41.5%) |
| Cannabinoid | 4.1% (3.2%, 5%) | 11.8% (10.2%, 13.4%) |
| Depressant | 2.1% (1.4%, 2.8%) | 4.9% (3.8%, 6%) |
| Narcotic | 1.8% (1.2%, 2.4%) | 5.1% (4%, 6.2%) |
| Other | 3.7% (2.8%, 4.6%) | 5.2% (4.1%, 6.3%) |
| Stimulant | 8.1% (6.8%, 9.4%) | 8.6% (7.2%, 10%) |

## Statistics for testing for trend in prevalence of drugs over study years by drug type using Cochran-Armitage trend test

drug\_list <- c("Alcohol", "Nonalcohol", "Narcotic", "Depressant",  
"Stimulant", "Cannabinoid", "Other")  
drug\_trend\_tests\_ca <- lapply(drug\_list, test\_trend\_ca)  
drug\_trend\_tests\_ca <- dplyr::bind\_rows(drug\_trend\_tests\_ca) %>%  
dplyr::mutate(drug = drug\_list) %>%  
dplyr::select(drug, Z, p.value)  
drug\_trend\_tests\_ca %>% knitr::kable()

|  |  |  |
| --- | --- | --- |
| drug | Z | p.value |
| Alcohol | 1.2 | 0.228 |
| Nonalcohol | 9.9 | 0.000 |
| Narcotic | 6.7 | 0.000 |
| Depressant | 4.7 | 0.000 |
| Stimulant | 0.5 | 0.604 |
| Cannabinoid | 13.6 | 0.000 |
| Other | 1.4 | 0.157 |

## Statistics for testing for trend in prevalence of drugs over study years by drug type using Wald test of logistic regression coefficient for "year"

drug\_list <- c("Alcohol", "Nonalcohol", "Narcotic", "Depressant",  
"Stimulant", "Cannabinoid", "Other")  
drug\_trend\_tests\_log\_reg <- lapply(drug\_list, test\_trend\_log\_reg)  
drug\_trend\_tests\_log\_reg <- dplyr::bind\_rows(drug\_trend\_tests\_log\_reg) %>%  
dplyr::mutate(drug = drug\_list) %>%  
dplyr::select(drug, Z, p.value)  
drug\_trend\_tests\_log\_reg %>% knitr::kable()

|  |  |  |
| --- | --- | --- |
| drug | Z | p.value |
| Alcohol | 1.2 | 0.228 |
| Nonalcohol | 11.2 | 0.000 |
| Narcotic | 6.6 | 0.000 |
| Depressant | 4.7 | 0.000 |
| Stimulant | -0.5 | 0.604 |
| Cannabinoid | 13.5 | 0.000 |
| Other | 1.4 | 0.158 |