Viz6 - Extra Credit

Hayden Ginman

2023-10-07

How did the impaired driving death rates change from 2012 to 2014 for each age group? Which age group and gender had the highest and lowest impaired driving death rates in both 2012 and 2014?

```
# Load Packages into coding environment
library("ggplot2")
library("dplyr")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

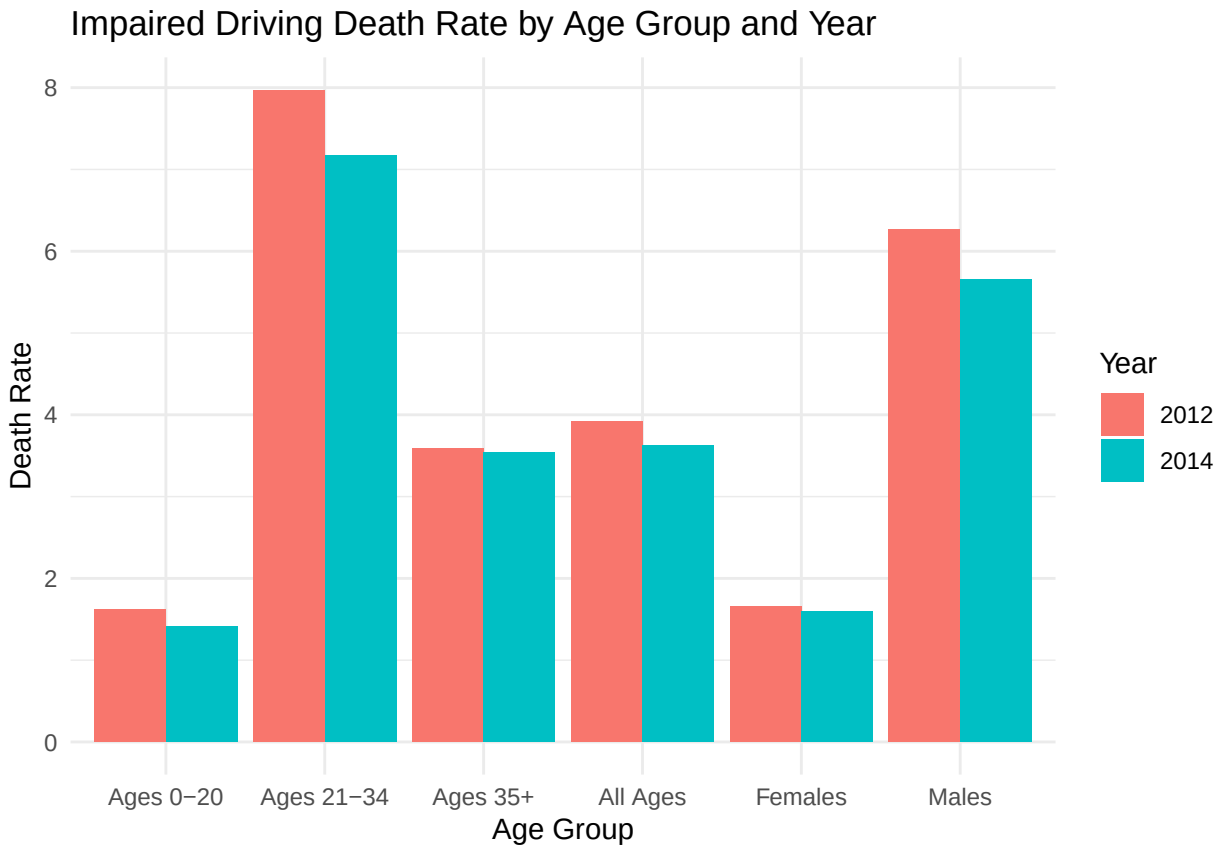
```
library("tidyr")
library("reshape2")
```

```
##
## Attaching package: 'reshape2'
```

```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
#Read data into coding environment

driving_death <- read.csv("Impaired_Driving_Death_Rate__by_Age_and_Gender__2012___2014__All_States.csv"

#rename annoying column names

new_dd <- driving_death %>%
  rename(
    State = State,
    Location = Location,
    All_2012 = `All.Ages..2012`,
    Age0to20_2012 = `Ages.0.20..2012`,
    Age21to34_2012 = `Ages.21.34..2012`,
    Age35plus_2012 = `Ages.35...2012`,
    Male_2012 = `Male..2012`,
    Female_2012 = `Female..2012`,
    All_2014 = `All.Ages..2014`,
    Age0to20_2014 = `Ages.0.20..2014`,
```

```
    Age21to34_2014 = `Ages.21.34..2014`,
    Age35plus_2014 = `Ages.35...2014`,
    Male_2014 = `Male..2014`,
    Female_2014 = `Female..2014`)

#Take the average for each age group column to aggregate data for entire US

US_dd <- new_dd %>%
  summarise(
    All_2012 = mean(All_2012, na.rm = TRUE),
    Age0to20_2012 = mean(Age0to20_2012, na.rm = TRUE),
    Age21to34_2012 = mean(Age21to34_2012, na.rm = TRUE),
    Age35plus_2012 = mean(Age35plus_2012, na.rm = TRUE),
    Male_2012 = mean(Male_2012, na.rm = TRUE),
    Female_2012 = mean(Female_2012, na.rm = TRUE),
    All_2014 = mean(All_2014, na.rm = TRUE),
    Age0to20_2014 = mean(Age0to20_2014, na.rm = TRUE),
    Age21to34_2014 = mean(Age21to34_2014, na.rm = TRUE),
    Age35plus_2014 = mean(Age35plus_2014, na.rm = TRUE),
    Male_2014 = mean(Male_2014, na.rm = TRUE),
    Female_2014 = mean(Female_2014, na.rm = TRUE))

#Transform data into long format with category and year as columns

long_dd <- US_dd %>%
  tidyr::pivot_longer(
    everything(),
    names_to = c("Category", "Year"),
    names_sep = "_",
    values_to = "Value")

#Visualize data

ggplot(long_dd, aes(x = Category, y = Value, fill = Year)) +
  geom_bar(stat = "identity", position = "dodge") + # Dodge will place bars side by side
  labs(title = "Impaired Driving Death Rate by Age Group and Year", x = "Age Group", y = "Death Rate") +
  scale_x_discrete(labels = c(All = "All Ages", Age0to20 = "Ages 0-20", Age21to34 = "Ages 21-34", Age35p
```


Impaired Driving Death Rate by Age Group and Year
Death Rate
8
6
4
2
0
Ages 0−20
Ages 21−34
Ages 35+
All Ages
Females
Males
Age Group
Year
2012
2014