test

sn

2024-03-25

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

Import data

```
data <- read.csv("substance_use.csv")
head(data)
```

```
##   measure                  location    sex      age                  cause
## 1  Deaths East Asia & Pacific - WB   Male 25 to 29 Alcohol use disorders
## 2  Deaths East Asia & Pacific - WB Female 25 to 29 Alcohol use disorders
## 3  Deaths East Asia & Pacific - WB   Male 30 to 34 Alcohol use disorders
## 4  Deaths East Asia & Pacific - WB Female 30 to 34 Alcohol use disorders
## 5  Deaths East Asia & Pacific - WB   Male 35 to 39 Alcohol use disorders
## 6  Deaths East Asia & Pacific - WB Female 35 to 39 Alcohol use disorders
##    metric year         val       upper       lower
## 1 Percent 1990 0.004355489 0.005574785 0.003579575
## 2 Percent 1990 0.002316023 0.002622133 0.002052042
## 3 Percent 1990 0.006539015 0.007974114 0.005392593
## 4 Percent 1990 0.002667792 0.002950154 0.002417720
## 5 Percent 1990 0.007597508 0.010585770 0.006359210
## 6 Percent 1990 0.002744876 0.003049935 0.002468063
```

In 2019, what region of the world has the highest rate of alcohol-related deaths among men aged 40-44?

```
alcohol_death <- data[data$year == '2019' & data$cause == 'Alcohol use disorders' & data$sex == 'Male' &
head(alcohol_death)
```

```
##        measure                       location  sex      age
## 7315   Deaths               South Asia - WB Male 40 to 44
## 7351   Deaths Middle East & North Africa - WB Male 40 to 44
## 7387   Deaths         East Asia & Pacific - WB Male 40 to 44
## 7423   Deaths                North America Male 40 to 44
## 7459   Deaths         Sub-Saharan Africa - WB Male 40 to 44
## 7495   Deaths       Europe & Central Asia - WB Male 40 to 44
##                         cause  metric year        val       upper       lower
## 7315 Alcohol use disorders Percent 2019 0.012215856 0.014481335 0.008484016
```

```
## 7351 Alcohol use disorders Percent 2019 0.003040330 0.003688087 0.002506647
## 7387 Alcohol use disorders Percent 2019 0.012726958 0.014213882 0.008809356
## 7423 Alcohol use disorders Percent 2019 0.029002889 0.031514494 0.026391834
## 7459 Alcohol use disorders Percent 2019 0.003210615 0.004246450 0.002634772
## 7495 Alcohol use disorders Percent 2019 0.053798538 0.058466137 0.047957598
```

```r
highest_val <- alcohol_death[which.max(alcohol_death$val), ]
highest_region <- highest_val$location
print(highest_region)
```

```
## [1] "Europe & Central Asia - WB"
```

Looking at the prevalence of alcohol-related disease in the East Asia and Pacific region, how has this changed over time and in the different age groups? Is there a difference between men and women?

select the East Asia and Pacific region

```r
Asia_Pacific <- data[data$location == 'East Asia & Pacific - WB' & data$cause == 'Alcohol use disorders
```

how the prevalence of alcohol-related diseasechange over time

```r
time <- ggplot(Asia_Pacific, aes(x = year, y = val, color = age, shape = sex)) +
  geom_line() +
  geom_point()
print(time)
```
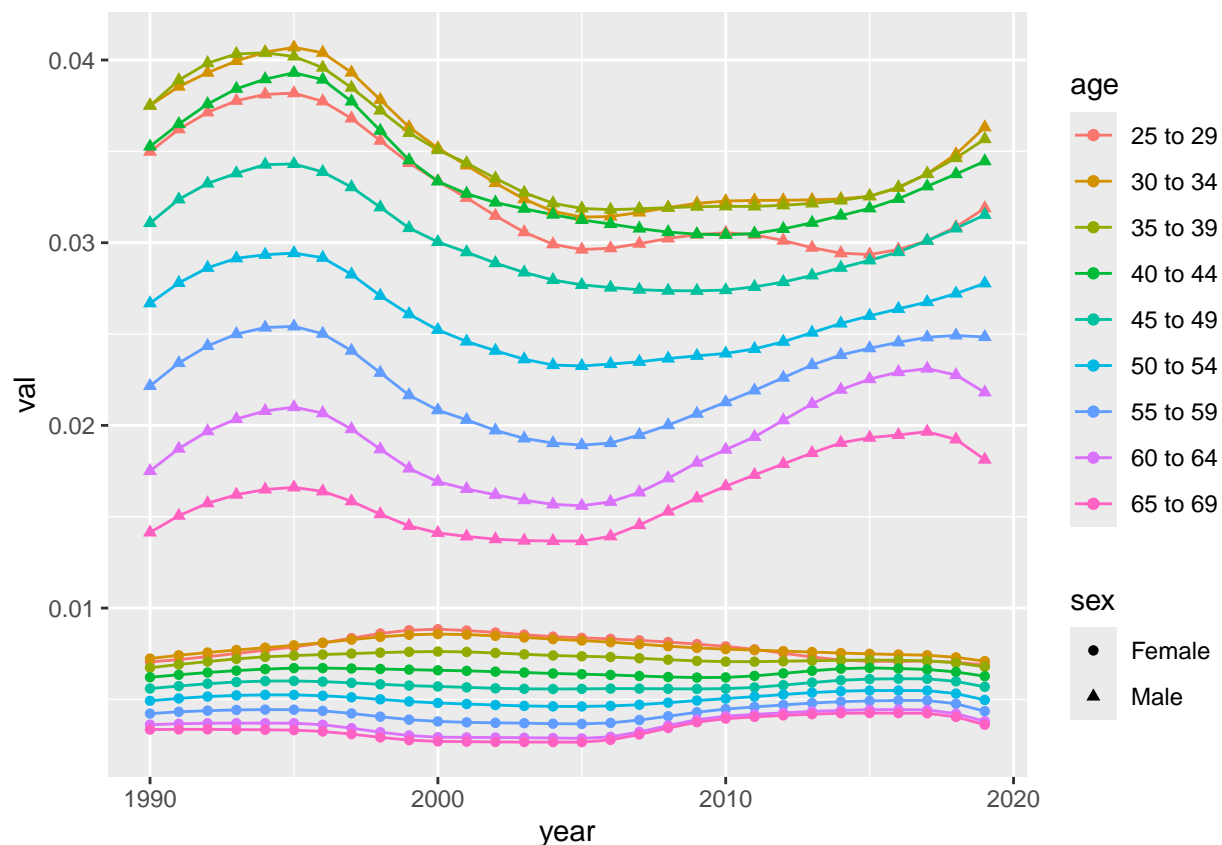
differences between males and females

```
Male <- Asia_Pacific[Asia_Pacific$sex == 'Male', ]
Female <- Asia_Pacific[Asia_Pacific$sex == 'Female', ]
wilcox.test(Male$val, Female$val, alternative = 'two.sided')
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  Male$val and Female$val
## W = 72900, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

Looking at the data from the United States, can you confirm an increase in the prevalence of diseases related to opioid use? What age group is the most affected?

```
US <- data[data$location == 'North America' & data$cause == 'Opioid use disorders' & data$measure == 'Pr

opioid <- ggplot(US, aes(x = year, y = val, color = age)) +
  geom_line() +
```

```
  geom_point()
print(opioid)
```