# YRBS_Analysis

## 2023-08-31

```r
library(openxlsx)
```

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

```r
library(dplyr)
```

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

```
##
## 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
```

```r
library(ggplot2)
```

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

```r
library(tidyr)
```

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

```r
get_data = function(file){
  data = read.xlsx(file, fillMergedCells = TRUE, sep.names = " ",
                   startRow = 2)
  colnames(data)[1] = gsub(" %", "", colnames(data)[1])
  colnames(data)[2] = gsub(" %", "", colnames(data)[2])
  data %>%
    tidyr::pivot_longer(cols = 3:length(data), names_to = c("Year", "Type"), values_to = "Value",
                        names_pattern = "([0-9]+)(.*)") %>%
    mutate(Value = as.numeric(gsub("[^0-9.-]", "", Value)),
           Type = trimws(Type)) %>%
    filter(!is.na(.[[1]]) & !is.na(.[[2]])) %>%
    tidyr::pivot_wider(names_from = "Type", values_from = "Value")%>%
    filter(Race %in% c("Total", "Asian", "Black", "Hispanic", "White")) %>%
```

```
    mutate(Year = as.numeric(Year)) %>%
    group_by(Grade) %>%
    group_split()
}

data = get_data("C:/Users/willi/Downloads/YRBS1.xlsx")
data_NYC = get_data("C:/Users/willi/Downloads/YRBS_NYC.xlsx")
data_MA = get_data("C:/Users/willi/Downloads/YRBS_Massachusetts.xlsx")
data_SF = get_data("C:/Users/willi/Downloads/YRBS_SF.xlsx")
data_LA = get_data("C:/Users/willi/Downloads/YRBS_LA.xlsx")
data_4 = get_data("C:/Users/willi/Downloads/YRBS_4.xlsx")
data_NYC_4 = get_data("C:/Users/willi/Downloads/YRBS_NYC_4.xlsx")
data_LA_4 = get_data("C:/Users/willi/Downloads/YRBS_LA_4.xlsx")
data_Active = get_data("C:/Users/willi/Downloads/YRBS_Active.xlsx")
data_Active_NYC = get_data("C:/Users/willi/Downloads/YRBS_Active_NYC.xlsx")
data_Alcohol = get_data("C:/Users/willi/Downloads/YRBS_Alcohol.xlsx")
data_Alcohol_NYC = get_data("C:/Users/willi/Downloads/YRBS_Alcohol_NYC.xlsx")


plot_data = function(data, place, titl){
  if (data$Grade[1] == "Total"){
    data = data %>%
      mutate(Grade = "9th-12th")
  }
  title = paste0(titl,
                 data$Grade[1], " Grade\n", " in ", place, " (YRBS Data)")
  ggplot(data, aes(x = Year, y = `%`, group = Race, color = Race)) +
    geom_errorbar(aes(ymin = `Lower CI Limit`, ymax = `Upper CI Limit`), width = 0.5) +
    scale_y_continuous(breaks = scales::breaks_pretty(5), limits = c(0, NA)) +
    geom_point(size = 2) + geom_line(size = 1) + theme_bw() + ggtitle(title) +
    ylab("Percentage")
}
```

```
plots_1 = lapply(data, plot_data, "United States", "Percentage that Ever Had Sexual Intercourse by Race
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```
plots_2 = lapply(data_NYC, plot_data, "New York City", "Percentage that Ever Had Sexual Intercourse by
plots_3 = lapply(data_MA, plot_data, "Massachusetts", "Percentage that Ever Had Sexual Intercourse by Ra
plots_4 = lapply(data_SF, plot_data, "San Francisco", "Percentage that Ever Had Sexual Intercourse by Ra
plots_5 = lapply(data_LA, plot_data, "Los Angeles", "Percentage that Ever Had Sexual Intercourse by Race
plots_6 = lapply(data_4, plot_data, "United States", "Percentage that Have 4 Or More Sex Partners by Rac
plots_7 = lapply(data_NYC_4, plot_data, "New York City", "Percentage that Have 4 Or More Sex Partners by
plots_8 = lapply(data_LA_4, plot_data, "Los Angeles", "Percentage that Have 4 Or More Sex Partners by Ra
plots_9 = lapply(data_Active, plot_data, "United States", "Percentage Sexually Active Past 90 Days by Ra
plots_10 = lapply(data_Active_NYC, plot_data, "New York City", "Percentage Sexually Active Past 90 Days
plots_11 = lapply(data_Alcohol, plot_data, "United States", "Percentage that Drank Alcohol Past 90 Days
```
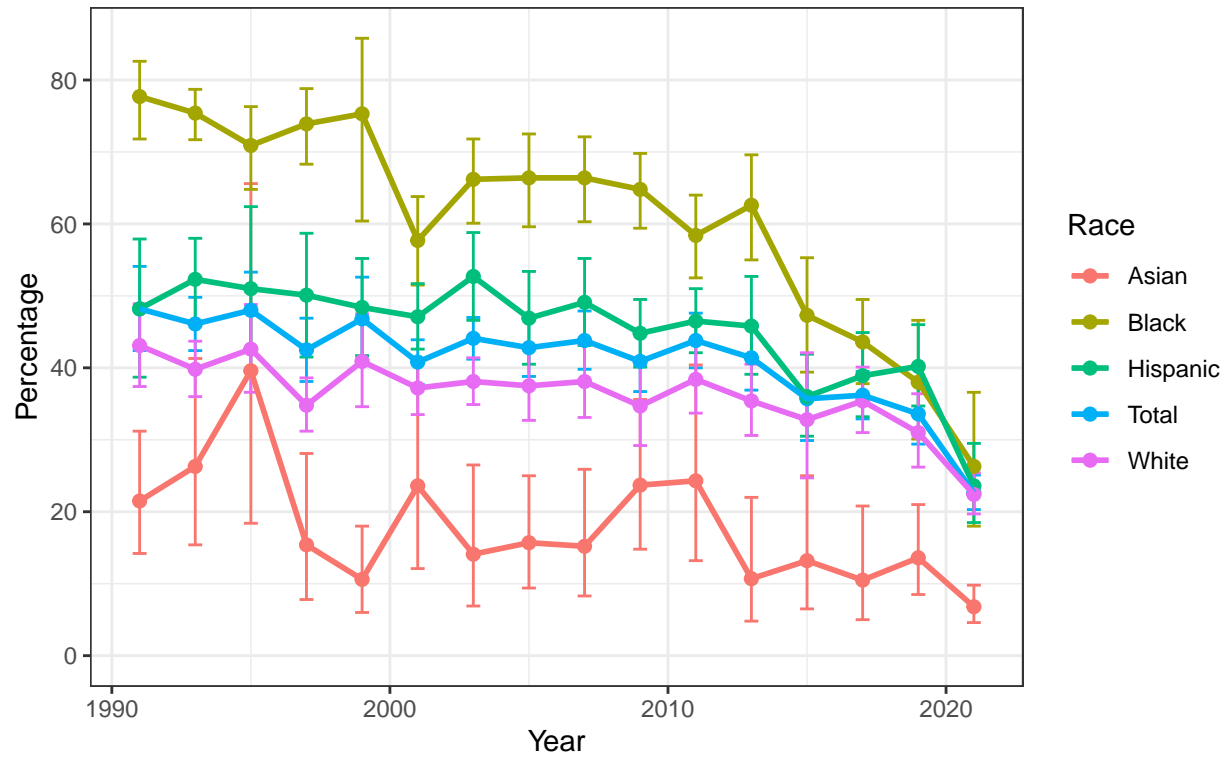
```
plots_12 = lapply(data_Alcohol_NYC, plot_data, "New York City", "Percentage that Drank Alcohol Past 90 I

plots_1
```

## [[1]]
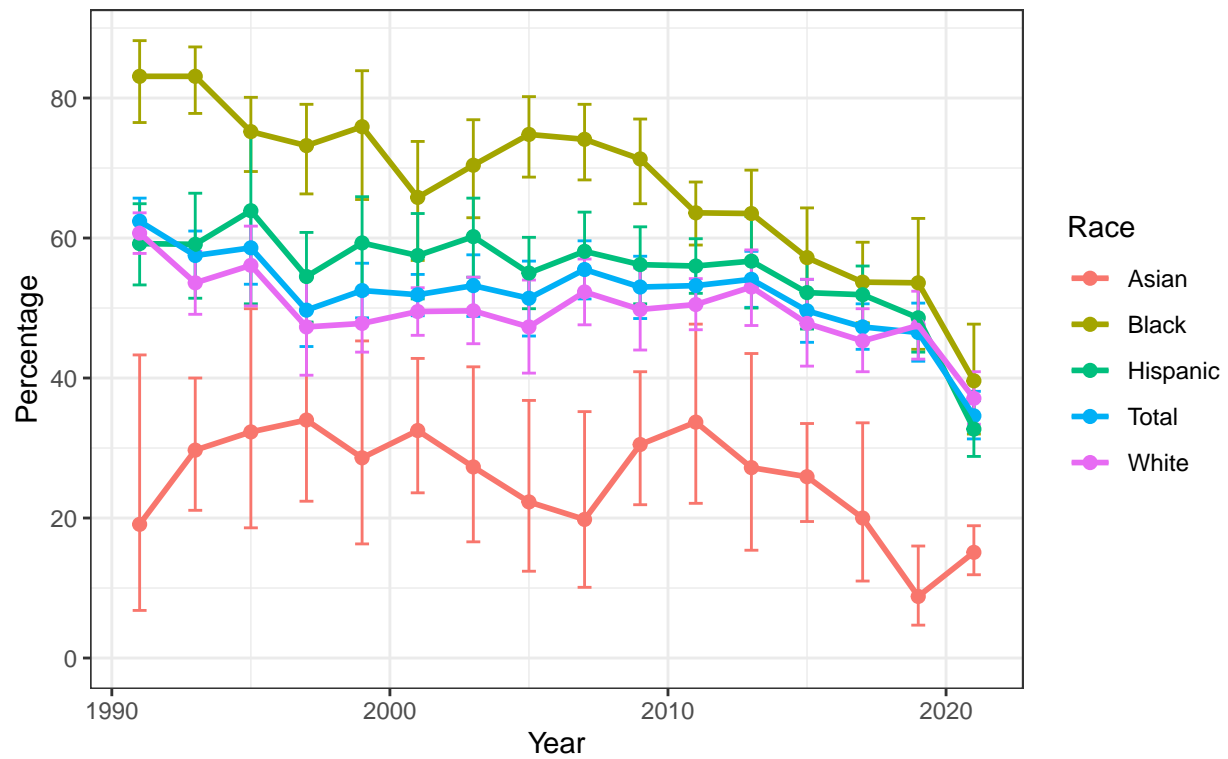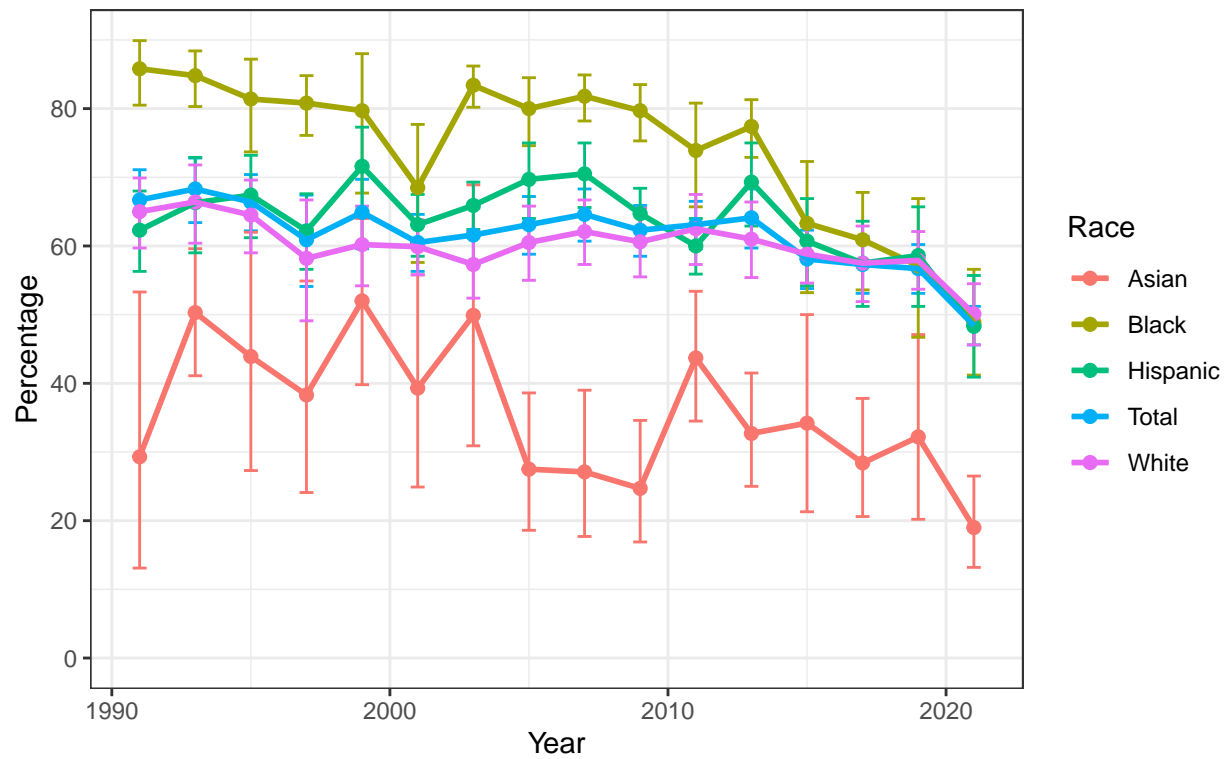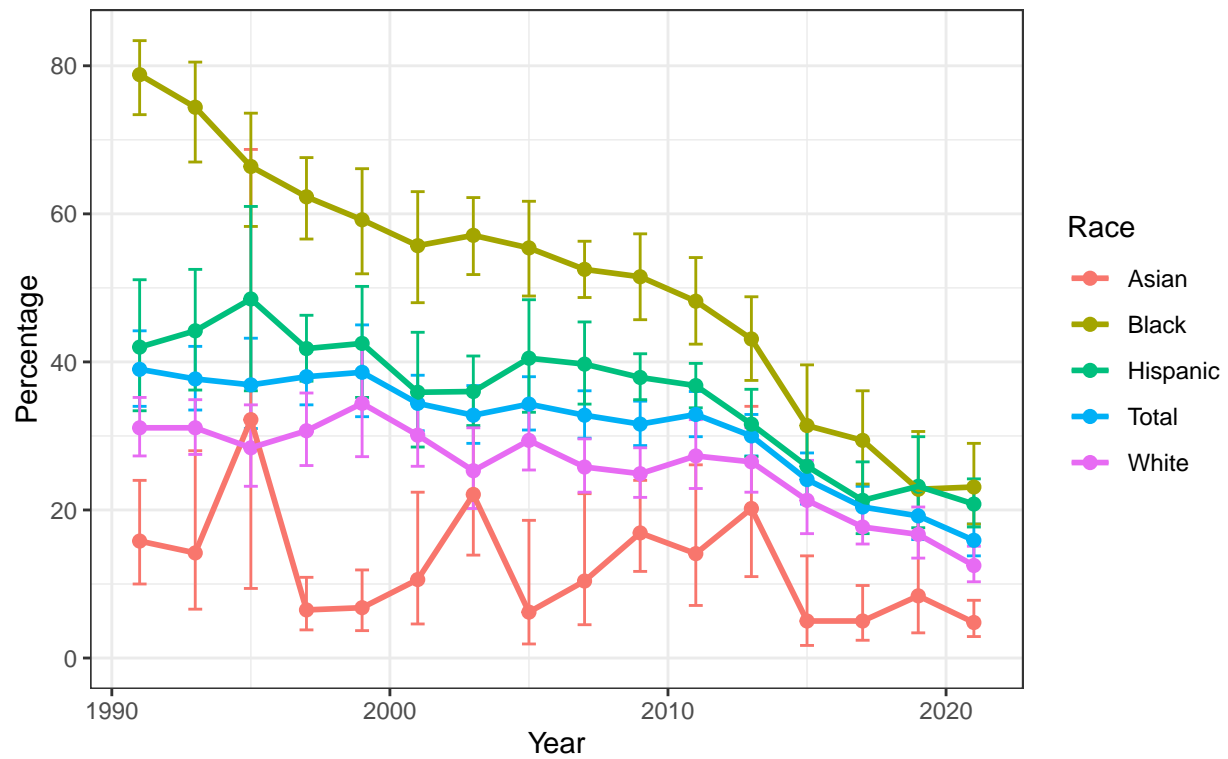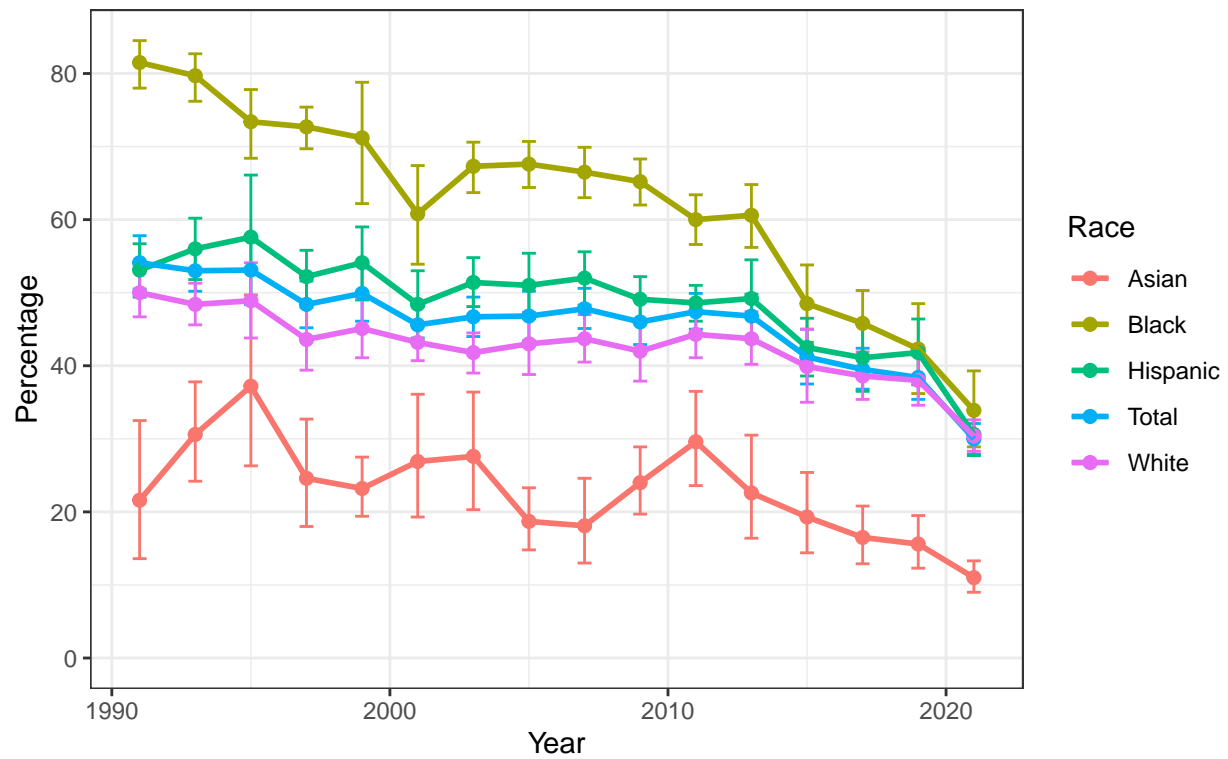
Percentage that Ever Had Sexual Intercourse by Race for 10th Grade
 in United States (YRBS Data)



##
## [[2]]

Percentage that Ever Had Sexual Intercourse by Race for 11th Grade
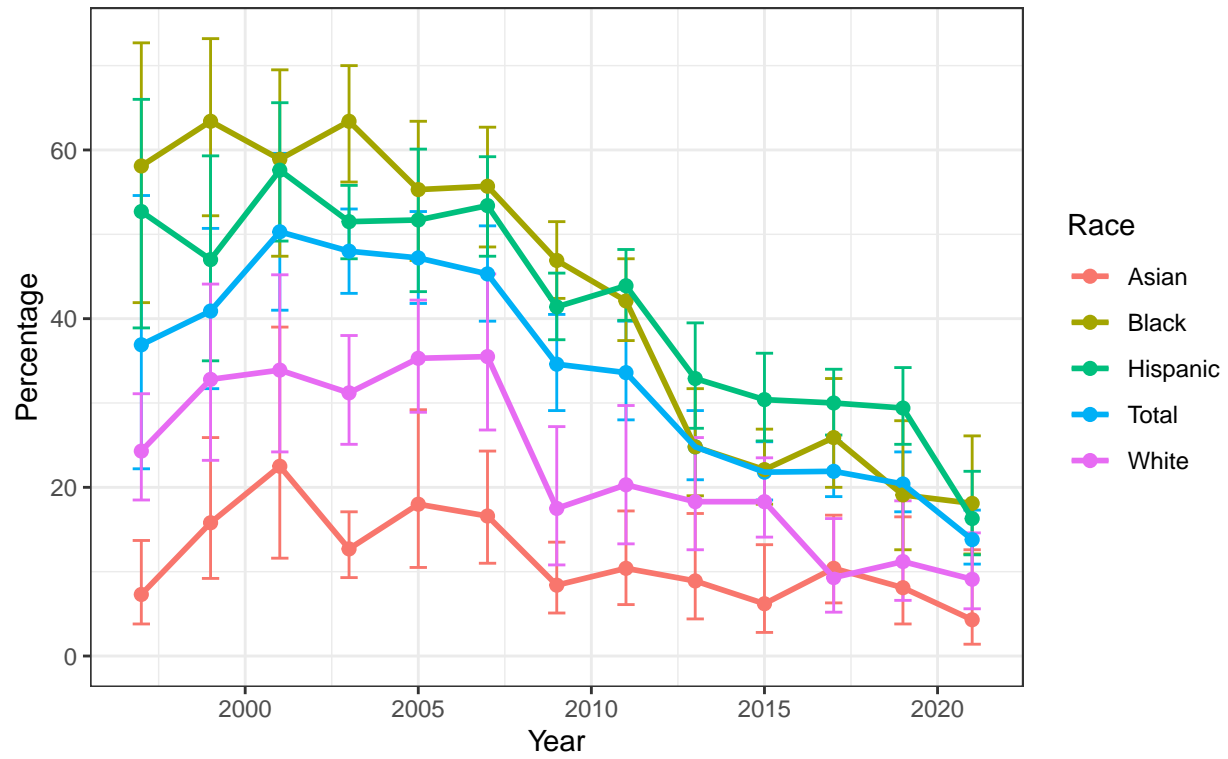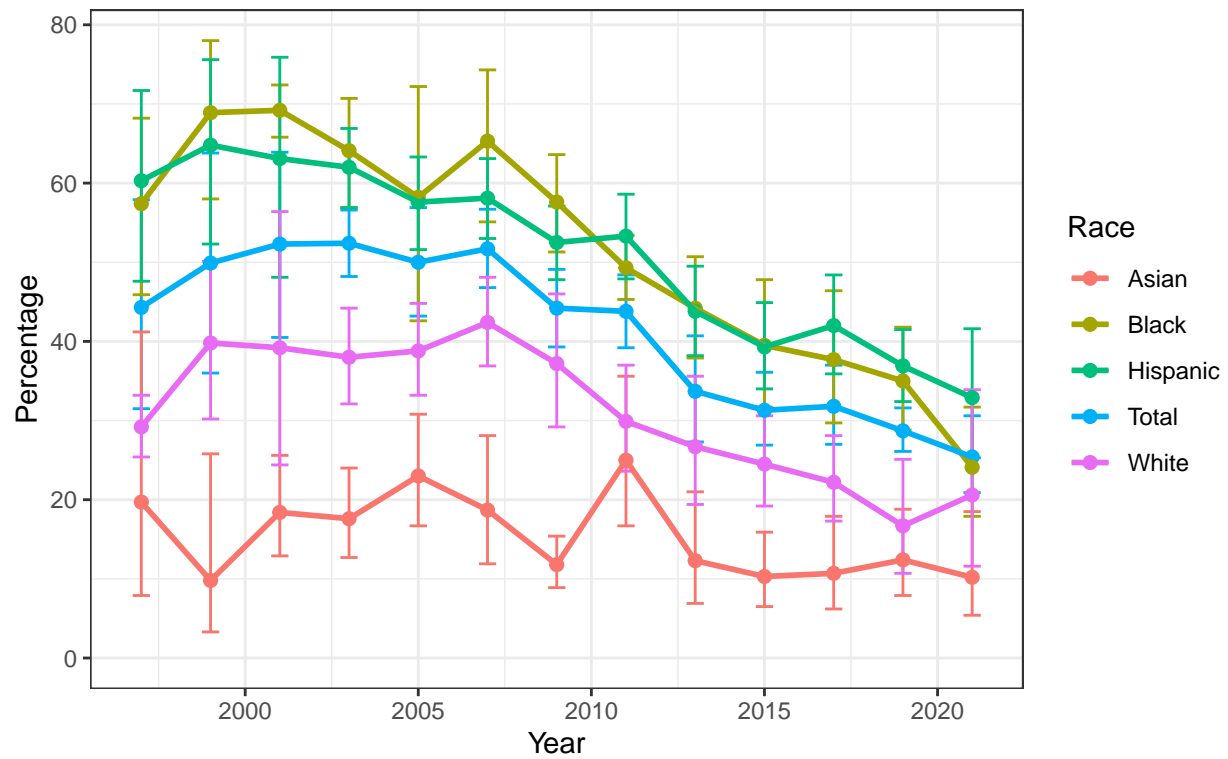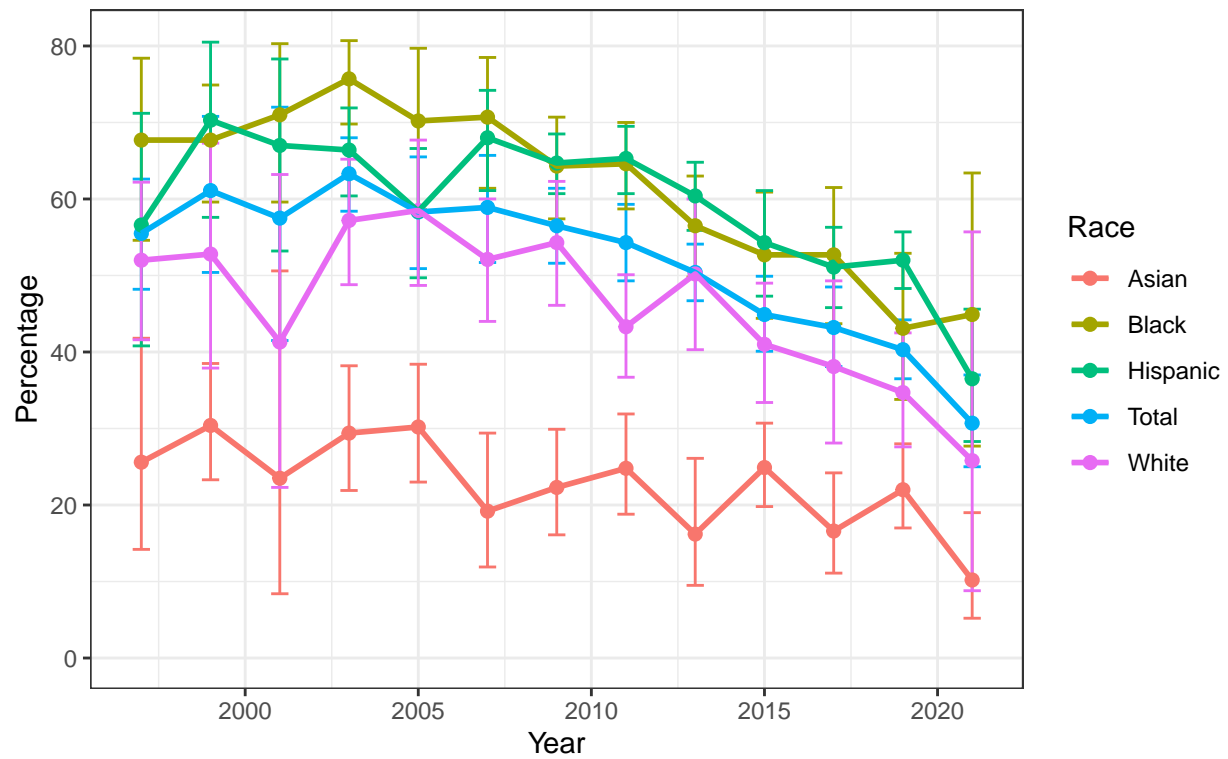in United States (YRBS Data)



```
##
## [[3]]
```

# Percentage that Ever Had Sexual Intercourse by Race for 12th Grade in United States (YRBS Data)



```
## 
## [[4]]
```

Percentage that Ever Had Sexual Intercourse by Race for 9th Grade in United States (YRBS Data)



```
##
## [[5]]
```

Percentage that Ever Had Sexual Intercourse by Race for 9th–12th Grade in United States (YRBS Data)

```
plots_2
```

```
## [[1]]
```

Percentage that Ever Had Sexual Intercourse by Race for 10th Grade in New York City (YRBS Data)

```
## 
## [[2]]
```

Percentage that Ever Had Sexual Intercourse by Race for 11th Grade in New York City (YRBS Data)

```
## 
## [[3]]
```

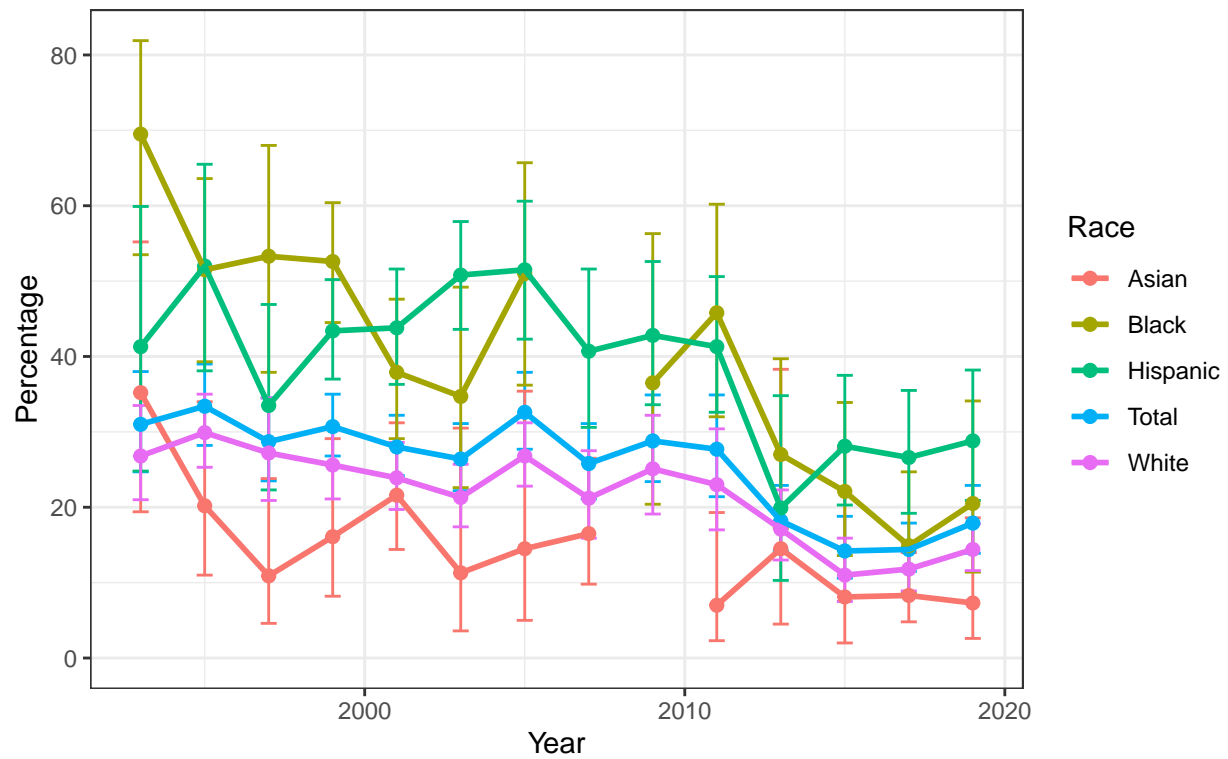## Percentage that Ever Had Sexual Intercourse by Race for 12th Grade in New York City (YRBS Data)
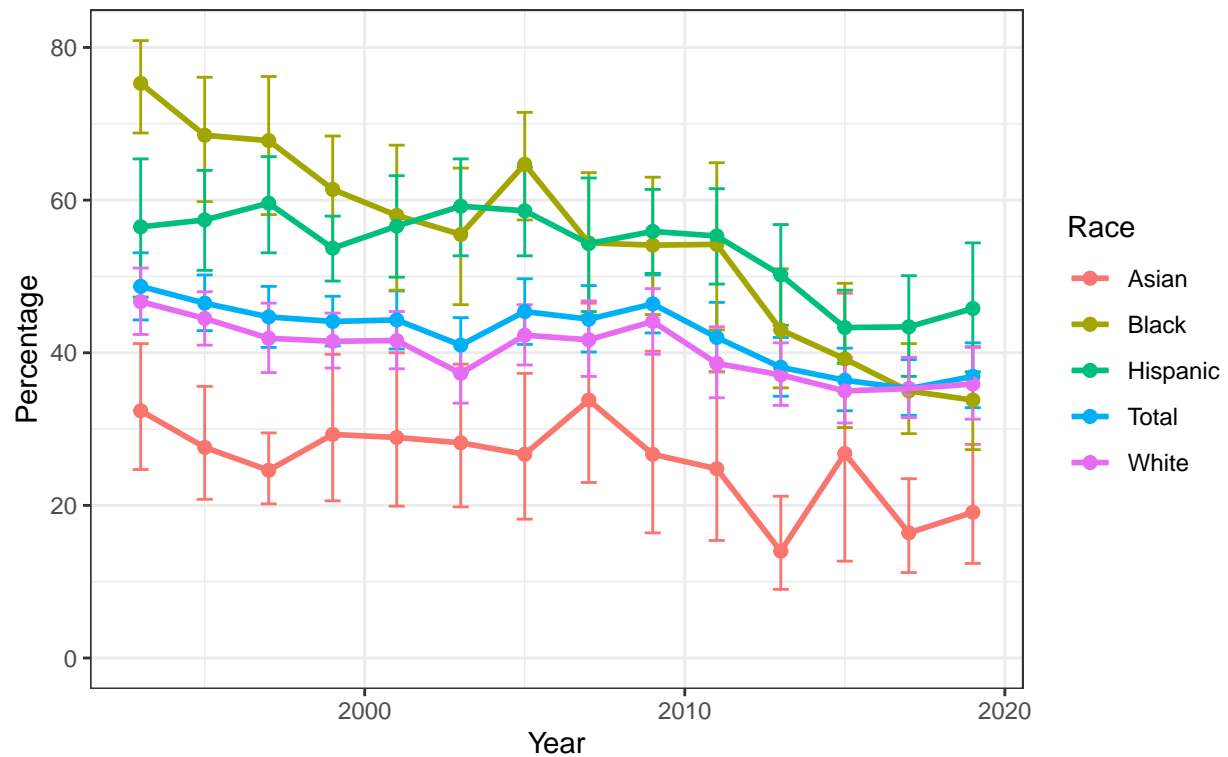


```
## 
## [[4]]
```

```
## Warning: Removed 2 rows containing missing values ('geom_point()').
```

Percentage that Ever Had Sexual Intercourse by Race for 9th Grade in New York City (YRBS Data)

```
##
## [[5]]
```

# Percentage that Ever Had Sexual Intercourse by Race for 9th–12th Grade in New York City (YRBS Data)
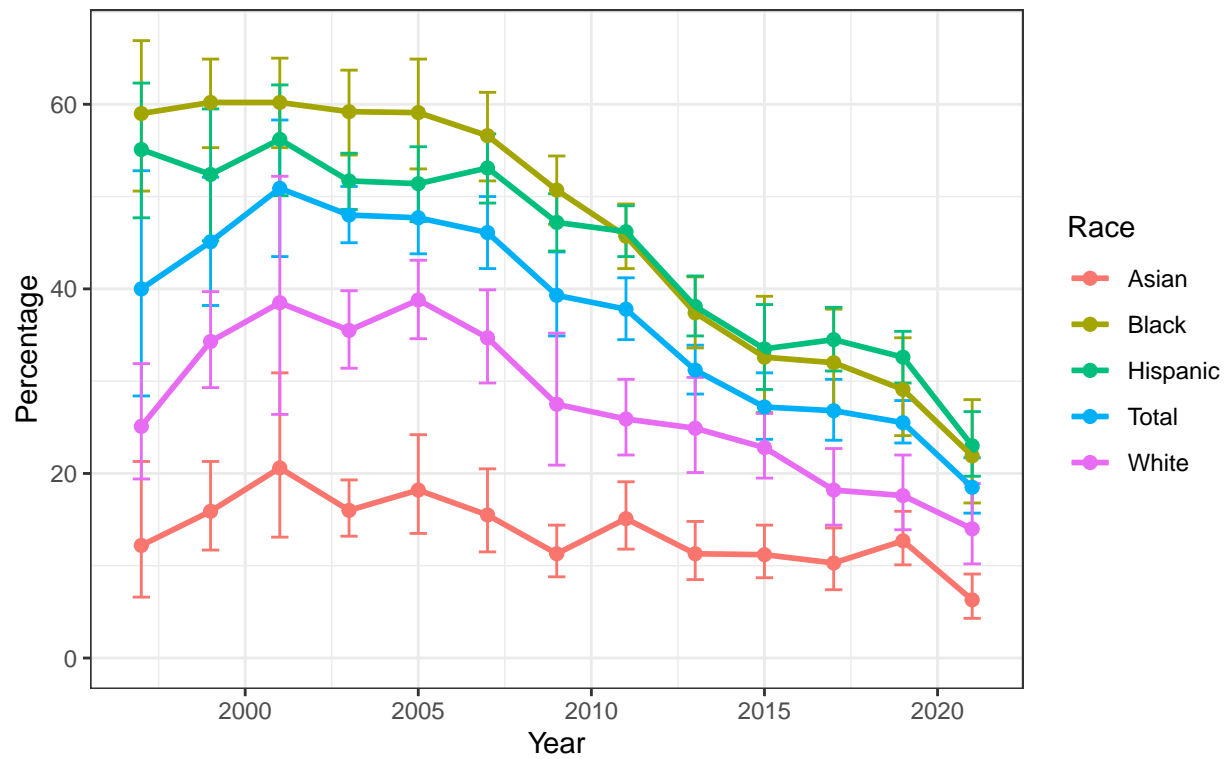


```
plots_3
```

```
## [[1]]
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```

## Percentage that Ever Had Sexual Intercourse by Race for 10th Grade in Massachusetts (YRBS Data)
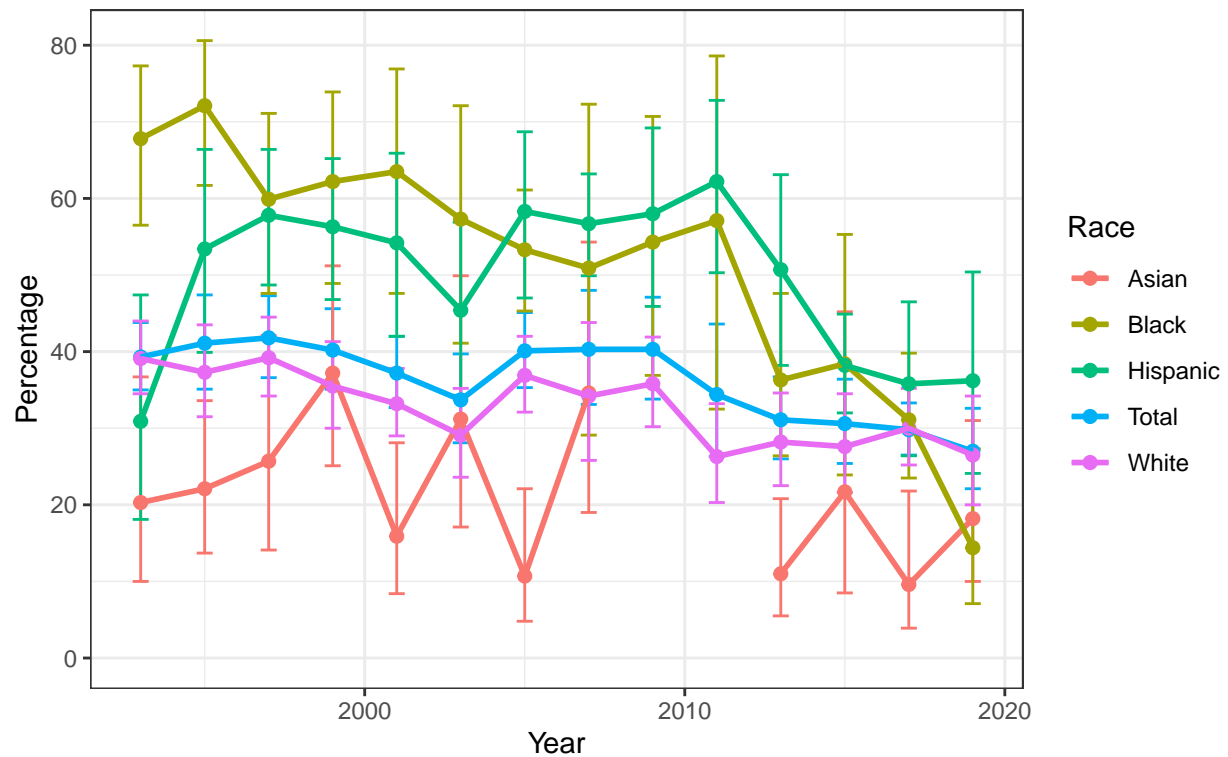

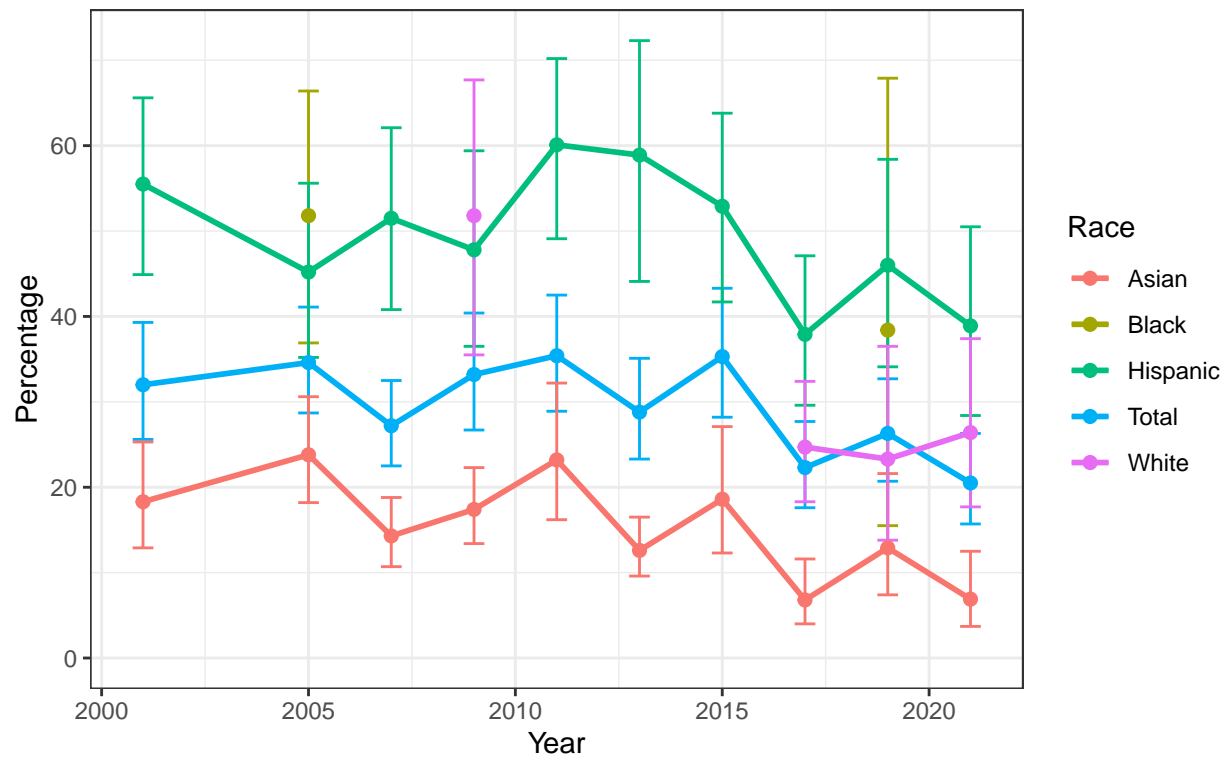
```
##
## [[2]]
```

```
## Warning: Removed 3 rows containing missing values (`geom_point()`).
```

```
## Warning: Removed 1 row containing missing values (`geom_line()`).
```

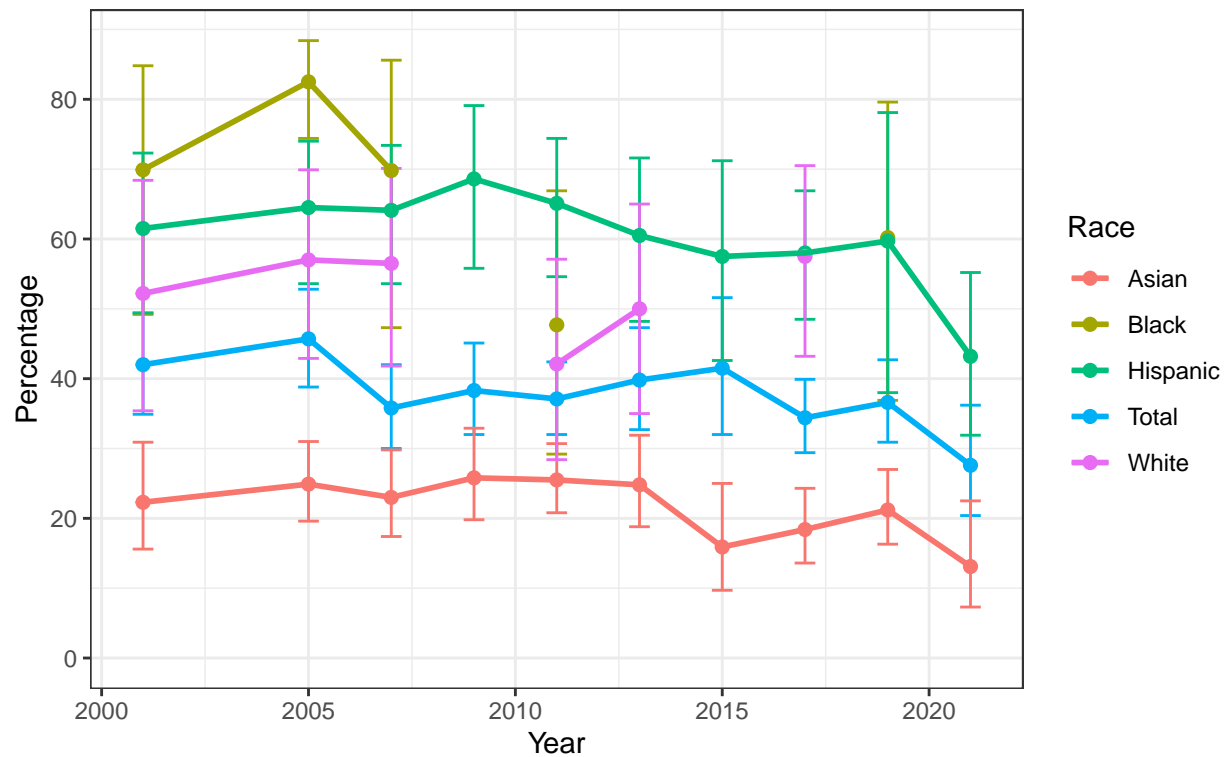## Percentage that Ever Had Sexual Intercourse by Race for 11th Grade in Massachusetts (YRBS Data)
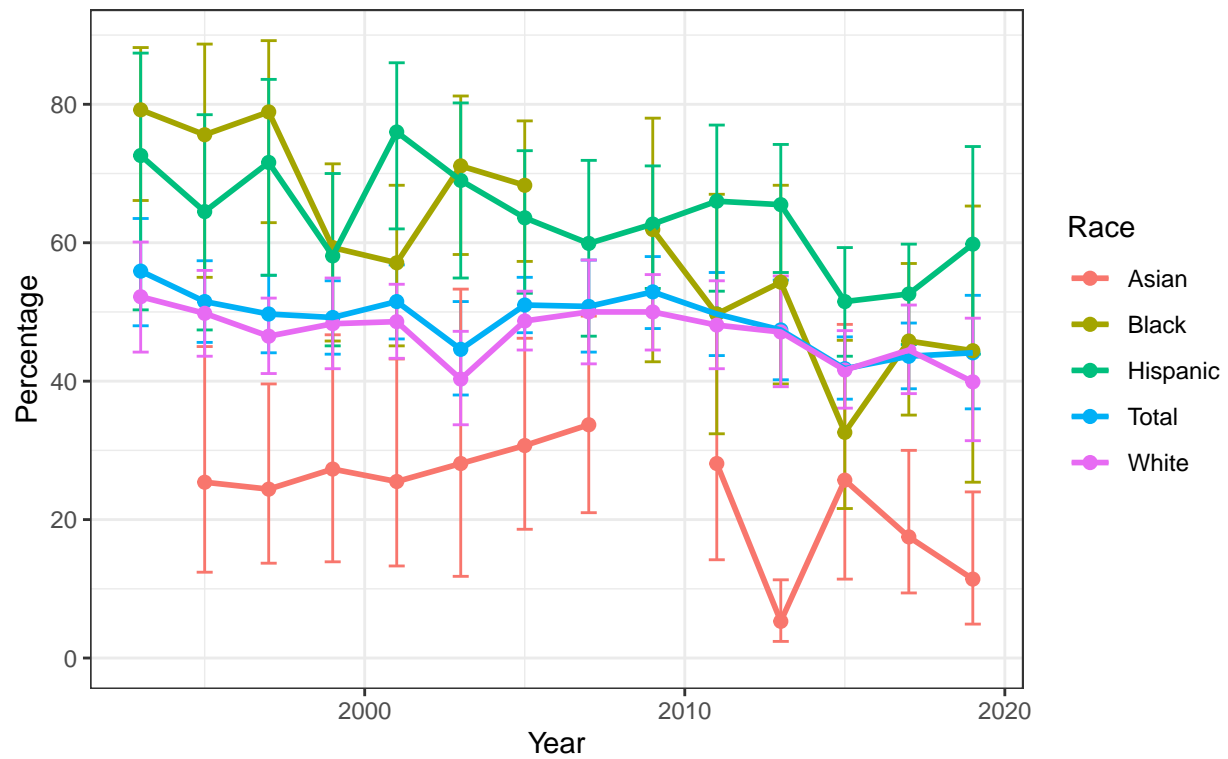


```
##
## [[3]]

## Warning: Removed 7 rows containing missing values ('geom_point()').

## Warning: Removed 2 rows containing missing values ('geom_line()').
```

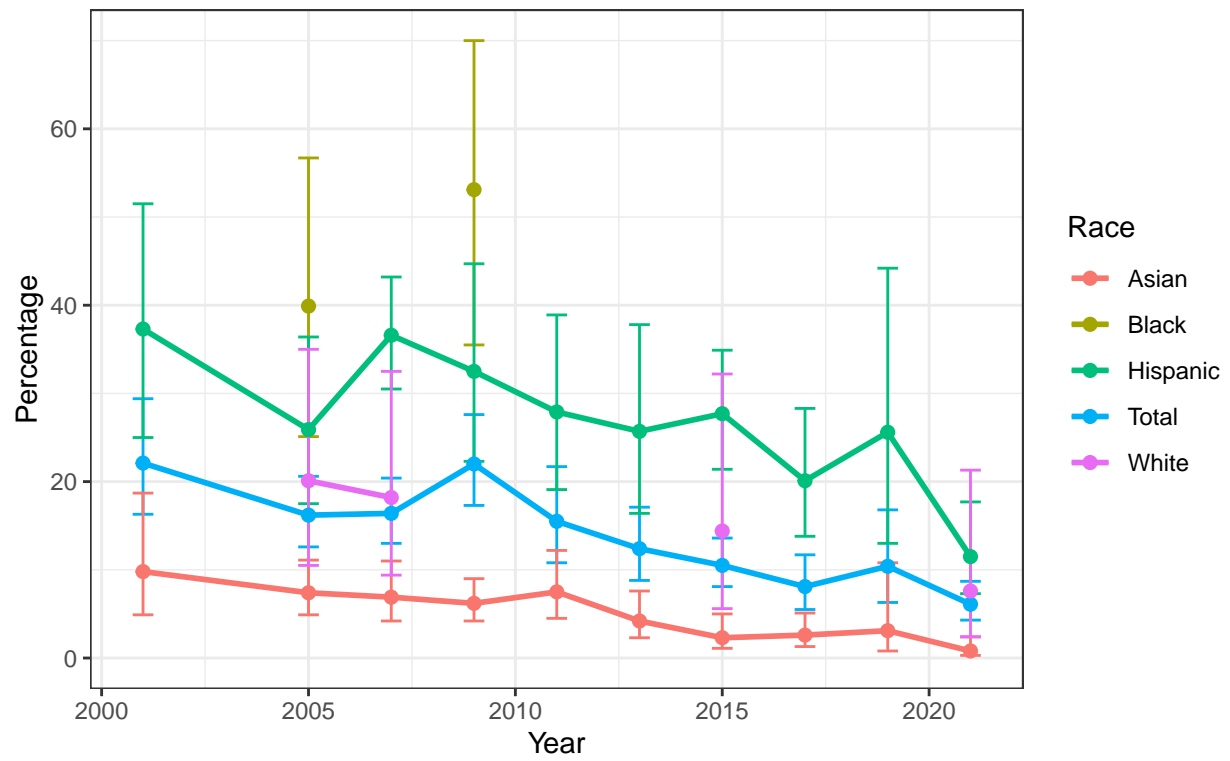Percentage that Ever Had Sexual Intercourse by Race for 12th Grade in Massachusetts (YRBS Data)
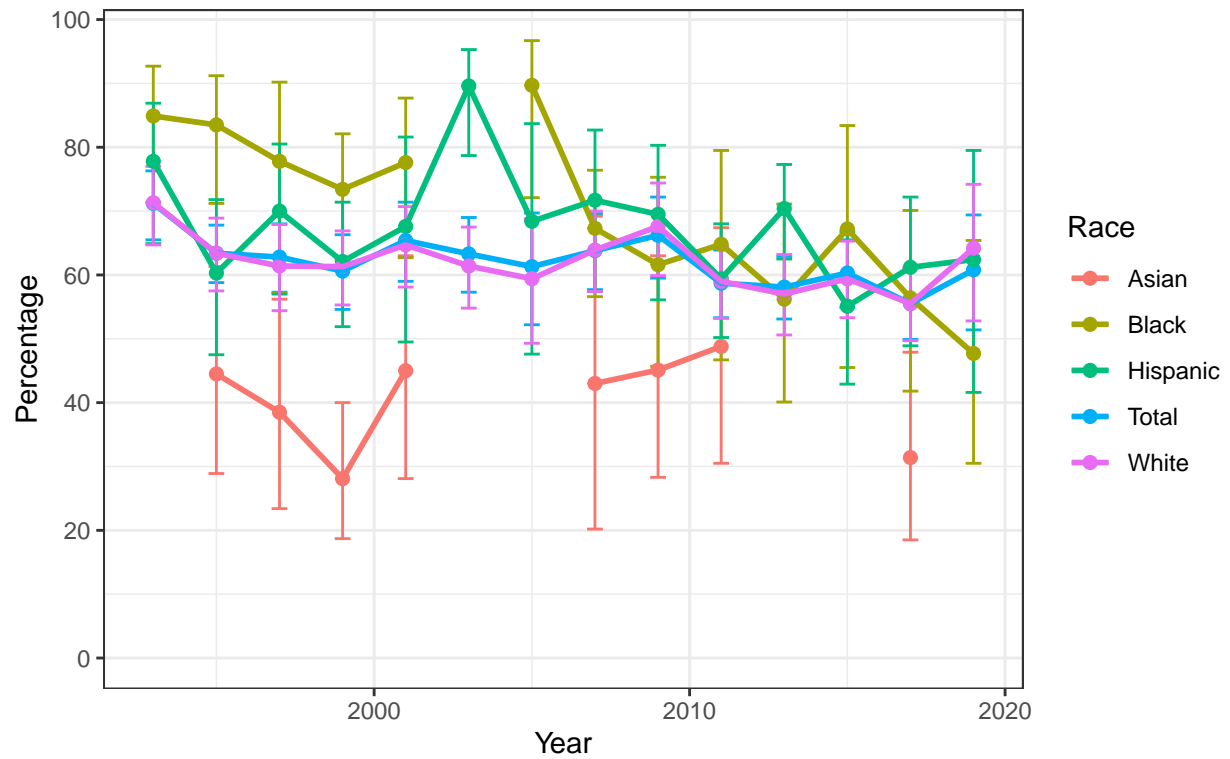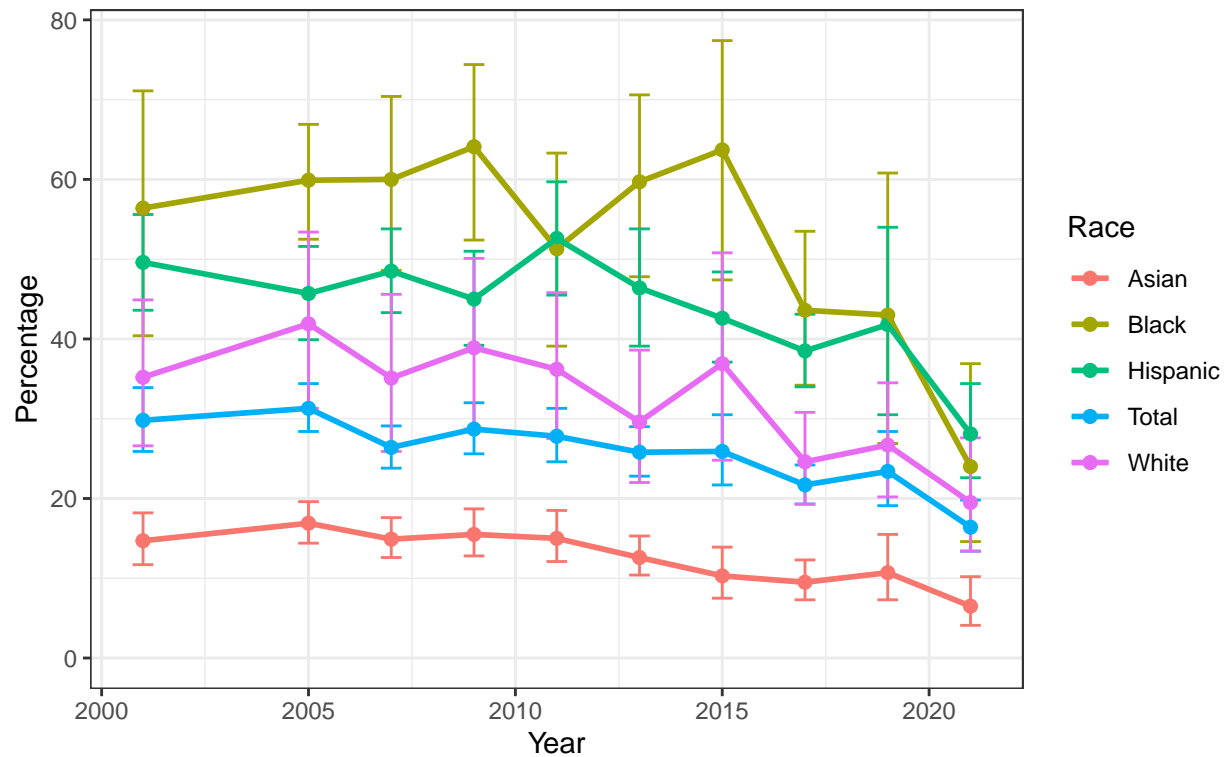
```
##
## [[4]]

## Warning: Removed 2 rows containing missing values ('geom_point()').
```

Percentage that Ever Had Sexual Intercourse by Race for 9th Grade in Massachusetts (YRBS Data)

```
## 
## [[5]]
```

Percentage that Ever Had Sexual Intercourse by Race for 9th–12th Grade
 in Massachusetts (YRBS Data)



```
plots_4
```
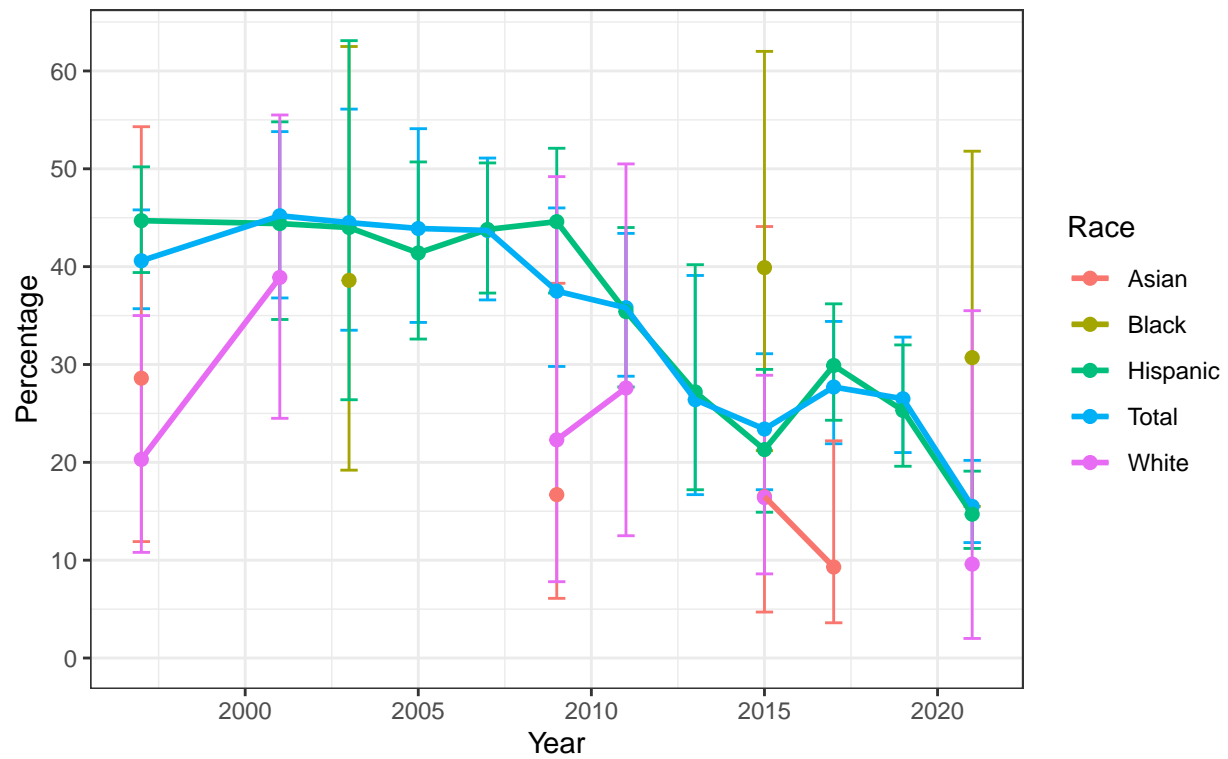
```
## [[1]]
```

```
## Warning: Removed 10 rows containing missing values ('geom_point()').
## Removed 2 rows containing missing values ('geom_line()').
```

Percentage that Ever Had Sexual Intercourse by Race for 10th Grade in San Francisco (YRBS Data)
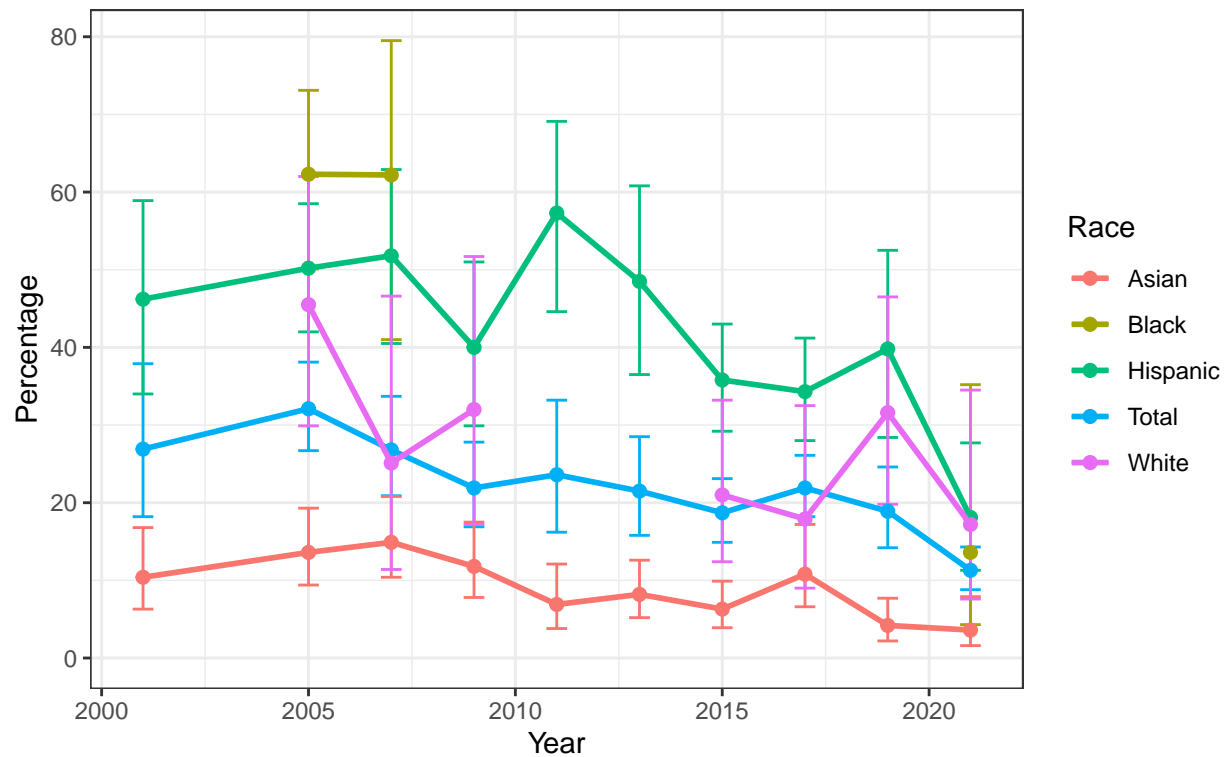
```
##
## [[2]]

## Warning: Removed 14 rows containing missing values ('geom_point()').

## Warning: Removed 5 rows containing missing values ('geom_line()').
```

Percentage that Ever Had Sexual Intercourse by Race for 11th Grade
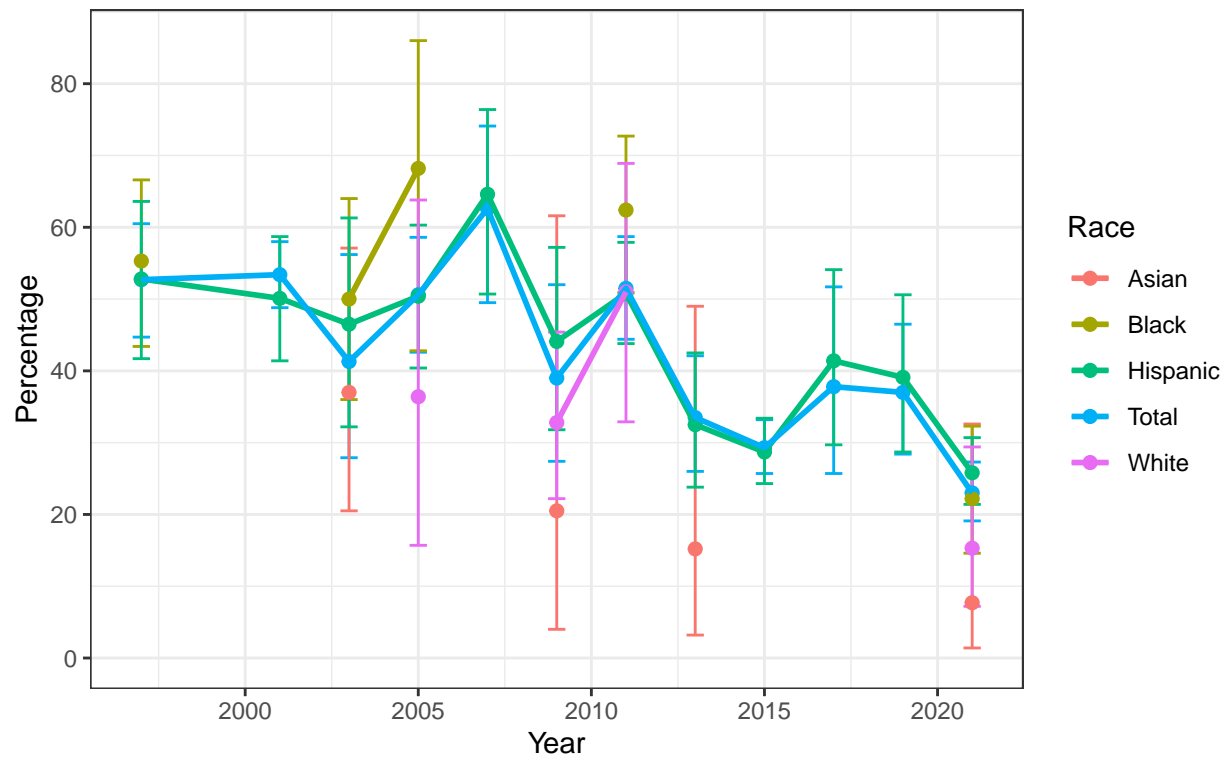 in San Francisco (YRBS Data)



```
##
## [[3]]

## Warning: Removed 9 rows containing missing values ('geom_point()').

## Warning: Removed 3 rows containing missing values ('geom_line()').
```

Percentage that Ever Had Sexual Intercourse by Race for 12th Grade in San Francisco (YRBS Data)
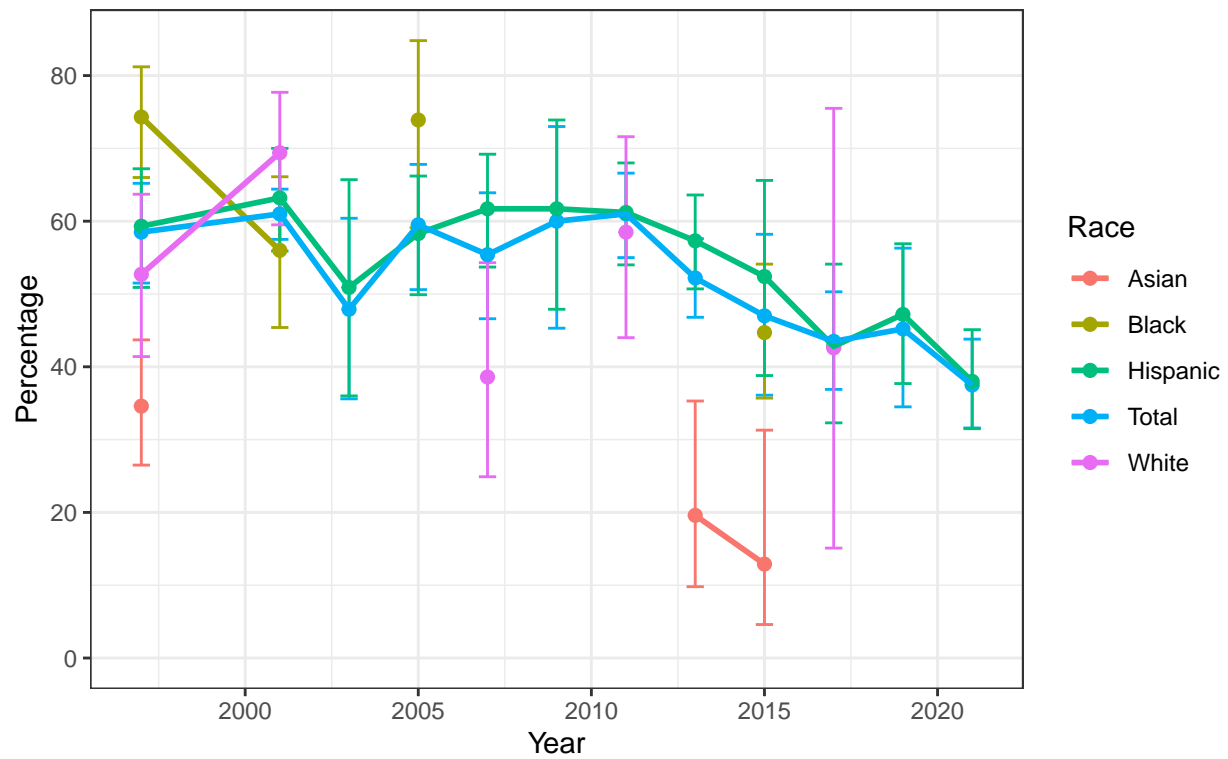
```
##
## [[4]]

## Warning: Removed 14 rows containing missing values (`geom_point()`).

## Warning: Removed 8 rows containing missing values (`geom_line()`).
```

Percentage that Ever Had Sexual Intercourse by Race for 9th Grade in San Francisco (YRBS Data)
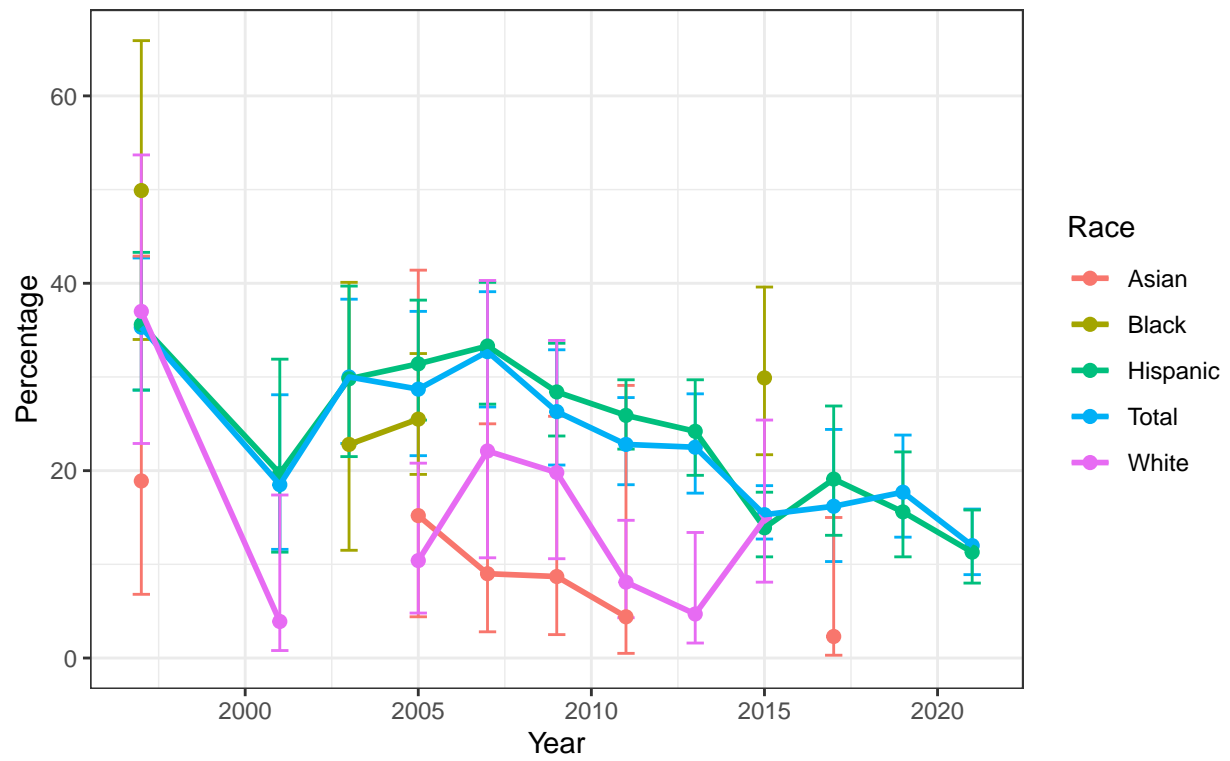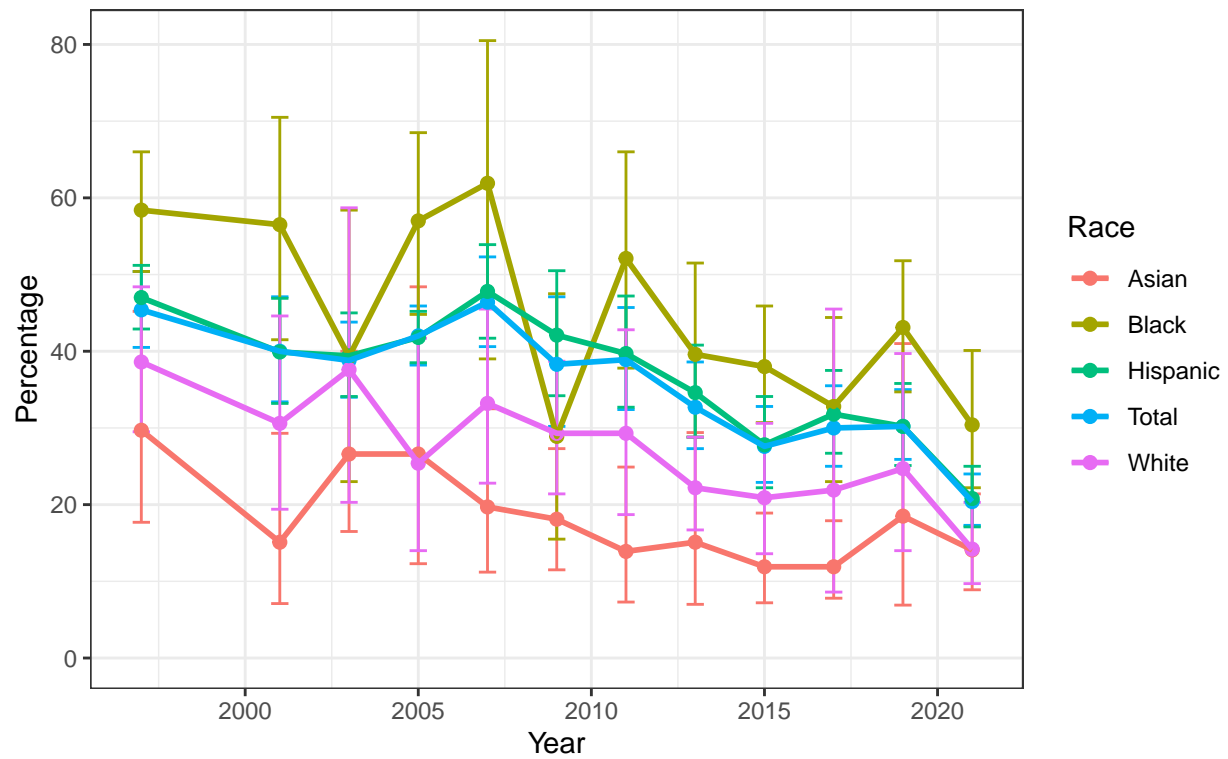
```
##
## [[5]]
```

Percentage that Ever Had Sexual Intercourse by Race for 9th–12th Grade
 in San Francisco (YRBS Data)



```
plots_5
```

```
## [[1]]
```

```
## Warning: Removed 23 rows containing missing values (`geom_point()`).
```

```
## Warning: Removed 4 rows containing missing values (`geom_line()`).
```

Percentage that Ever Had Sexual Intercourse by Race for 10th Grade in Los Angeles (YRBS Data)

```
##
## [[2]]

## Warning: Removed 23 rows containing missing values (`geom_point()`).

## Warning: Removed 5 rows containing missing values (`geom_line()`).
```

Percentage that Ever Had Sexual Intercourse by Race for 11th Grade
in Los Angeles (YRBS Data)



```
## 
## [[3]]

## Warning: Removed 24 rows containing missing values (`geom_point()`).

## Warning: Removed 8 rows containing missing values (`geom_line()`).
```

Percentage that Ever Had Sexual Intercourse by Race for 12th Grade in Los Angeles (YRBS Data)

```
##
## [[4]]

## Warning: Removed 18 rows containing missing values (`geom_point()`).
## Removed 8 rows containing missing values (`geom_line()`).
```

Percentage that Ever Had Sexual Intercourse by Race for 9th Grade in Los Angeles (YRBS Data)

```
##
## [[5]]
```

Percentage that Ever Had Sexual Intercourse by Race for 9th–12th Grade in Los Angeles (YRBS Data)

```
plots_6
```

```
## [[1]]
```

## Percentage that Have 4 Or More Sex Partners by Race for 10th Grade in United States (YRBS Data)



```
## 
## [[2]]
```

Percentage that Have 4 Or More Sex Partners by Race for 11th Grade
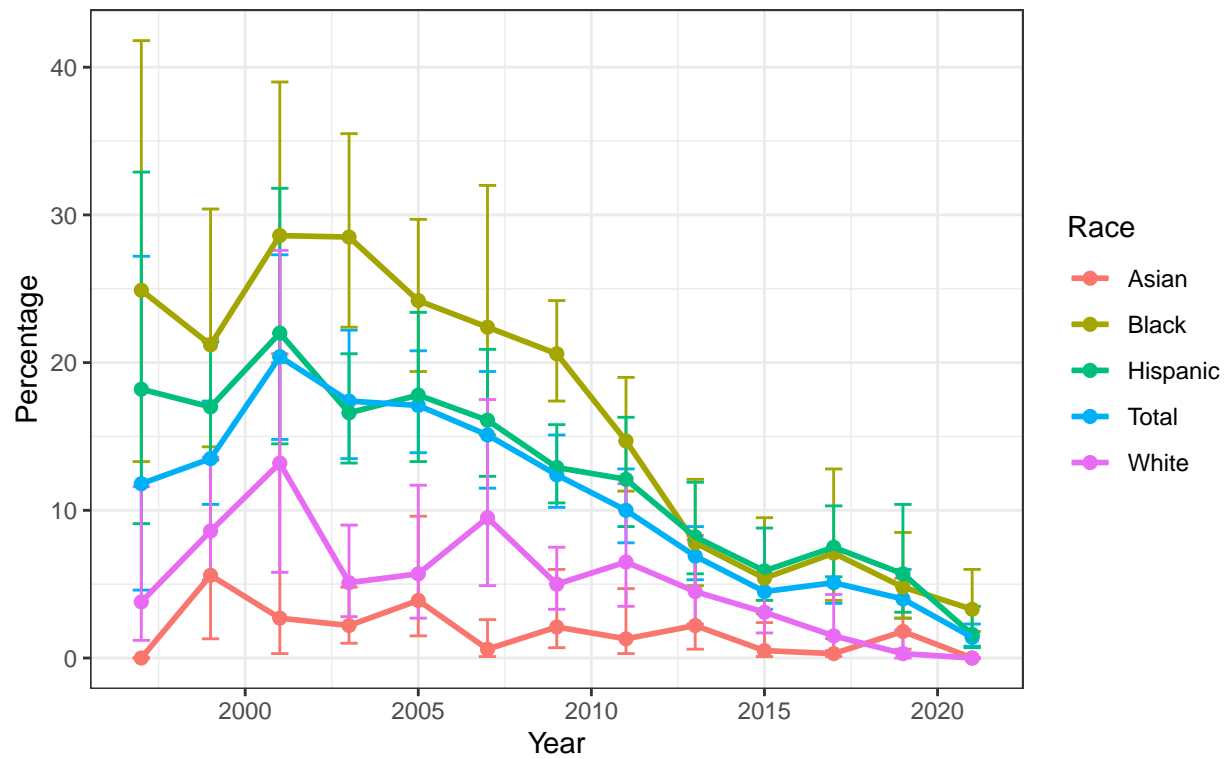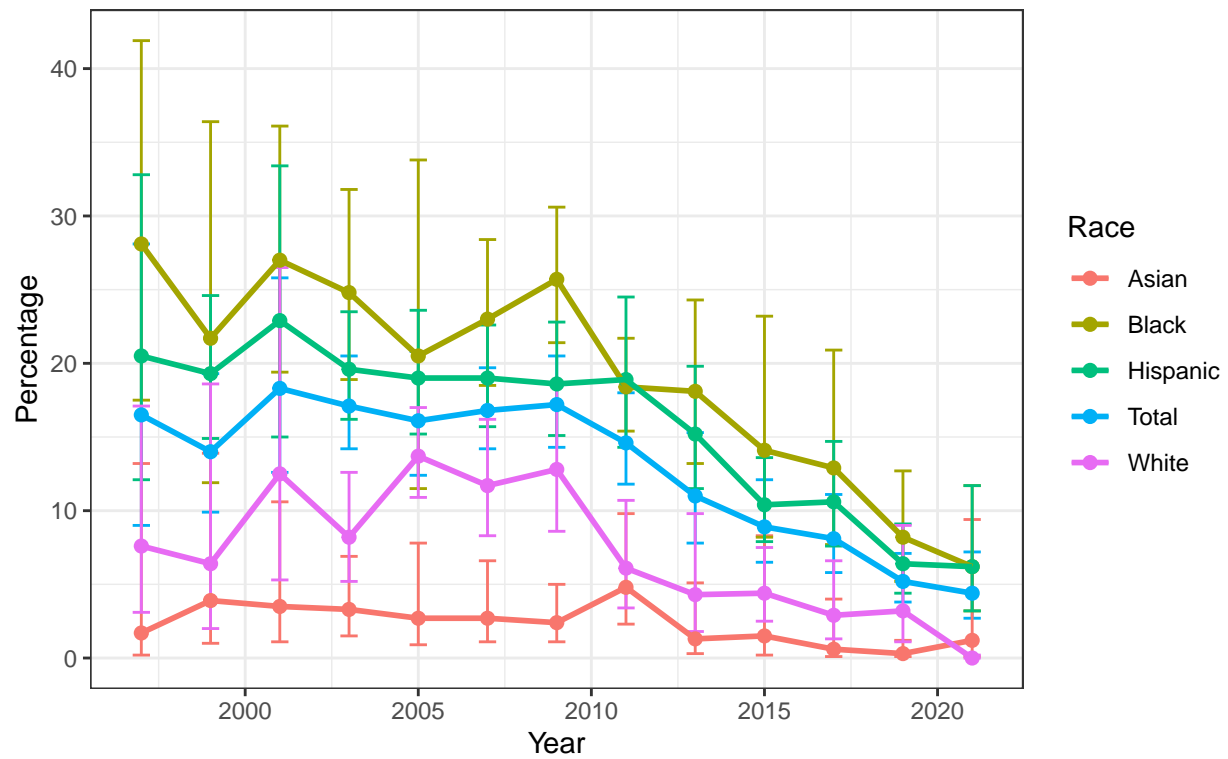 in United States (YRBS Data)



```
##
## [[3]]
```
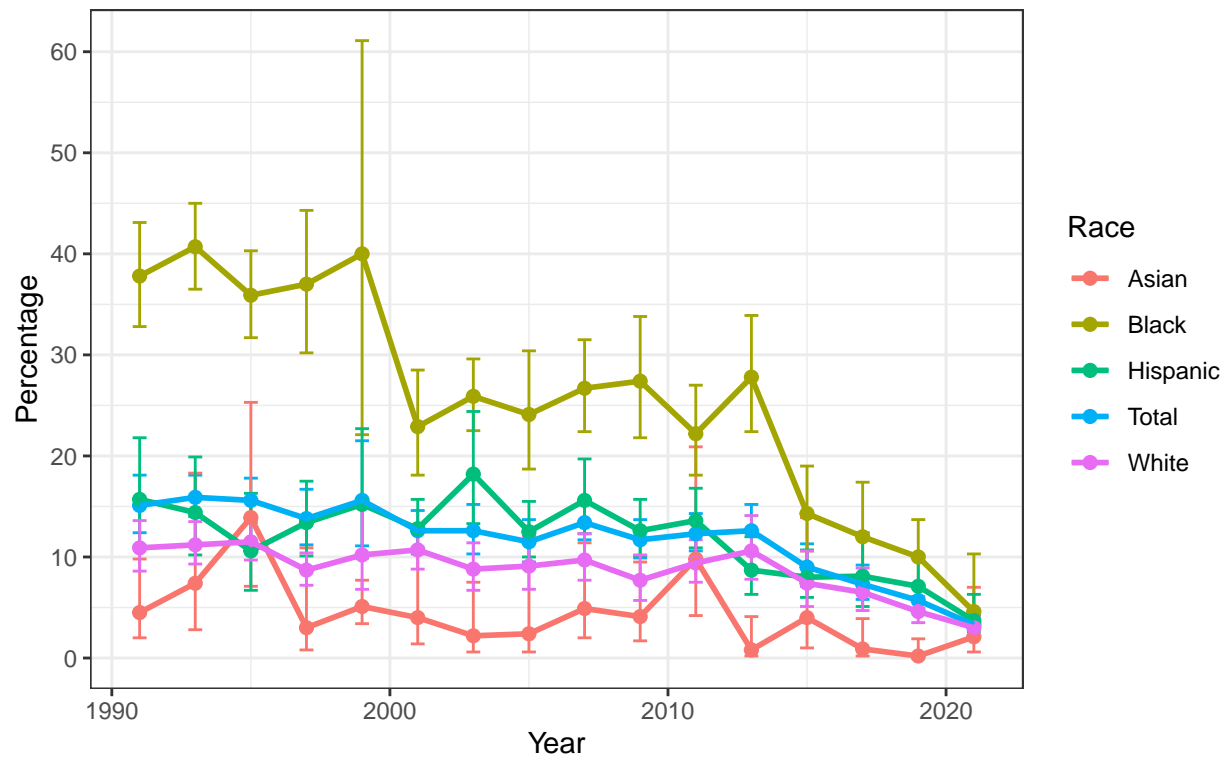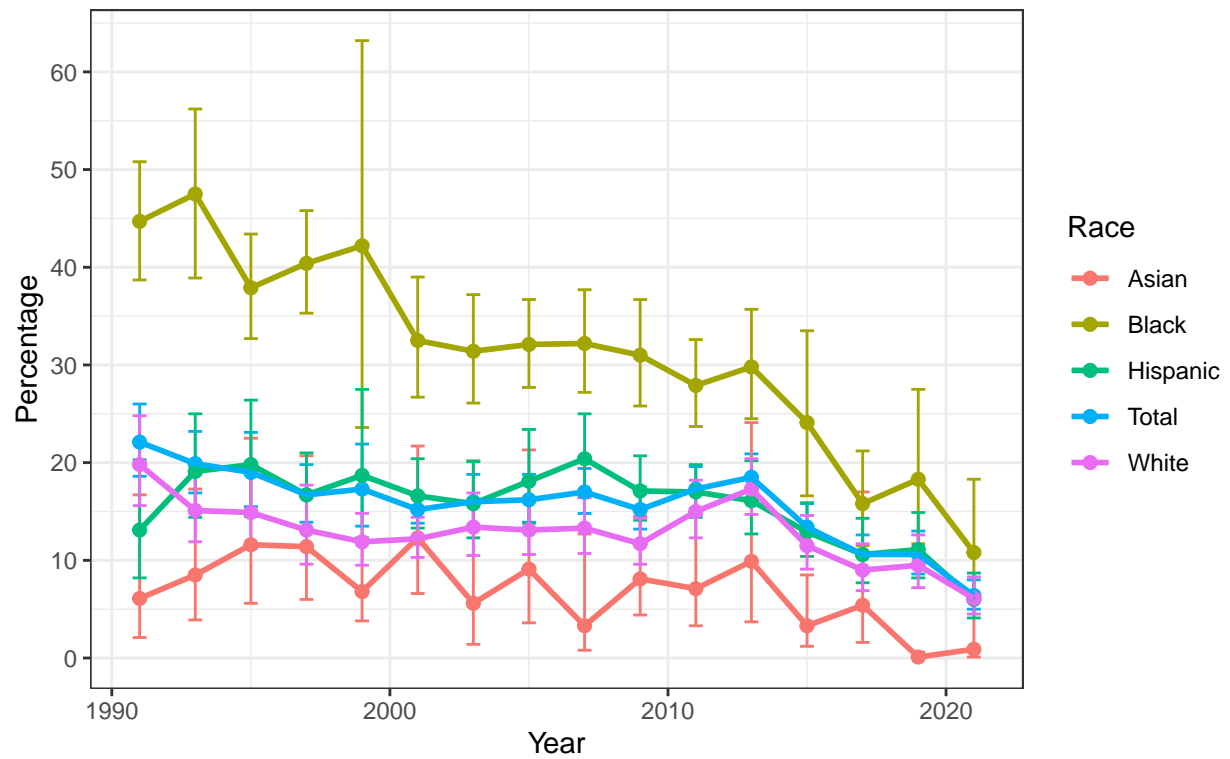
Percentage that Have 4 Or More Sex Partners by Race for 12th Grade
 in United States (YRBS Data)
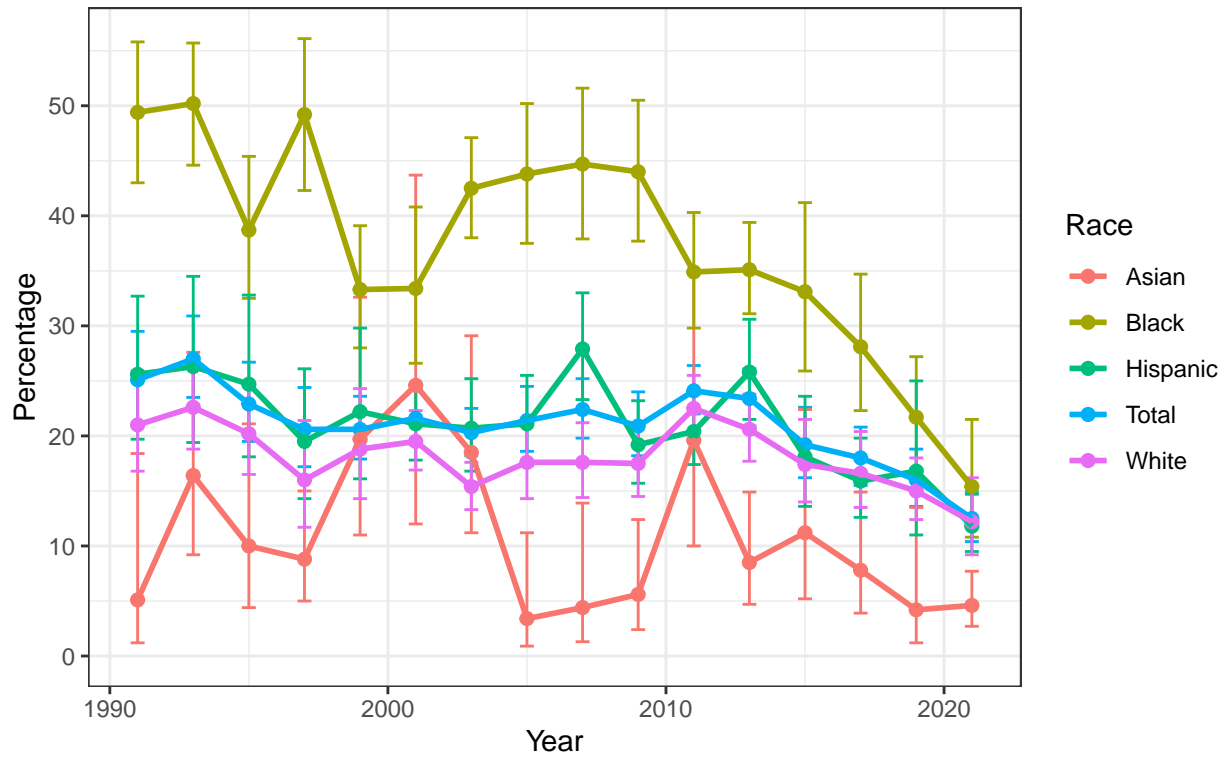


```
## 
## [[4]]
```

Percentage that Have 4 Or More Sex Partners by Race for 9th Grade in United States (YRBS Data)
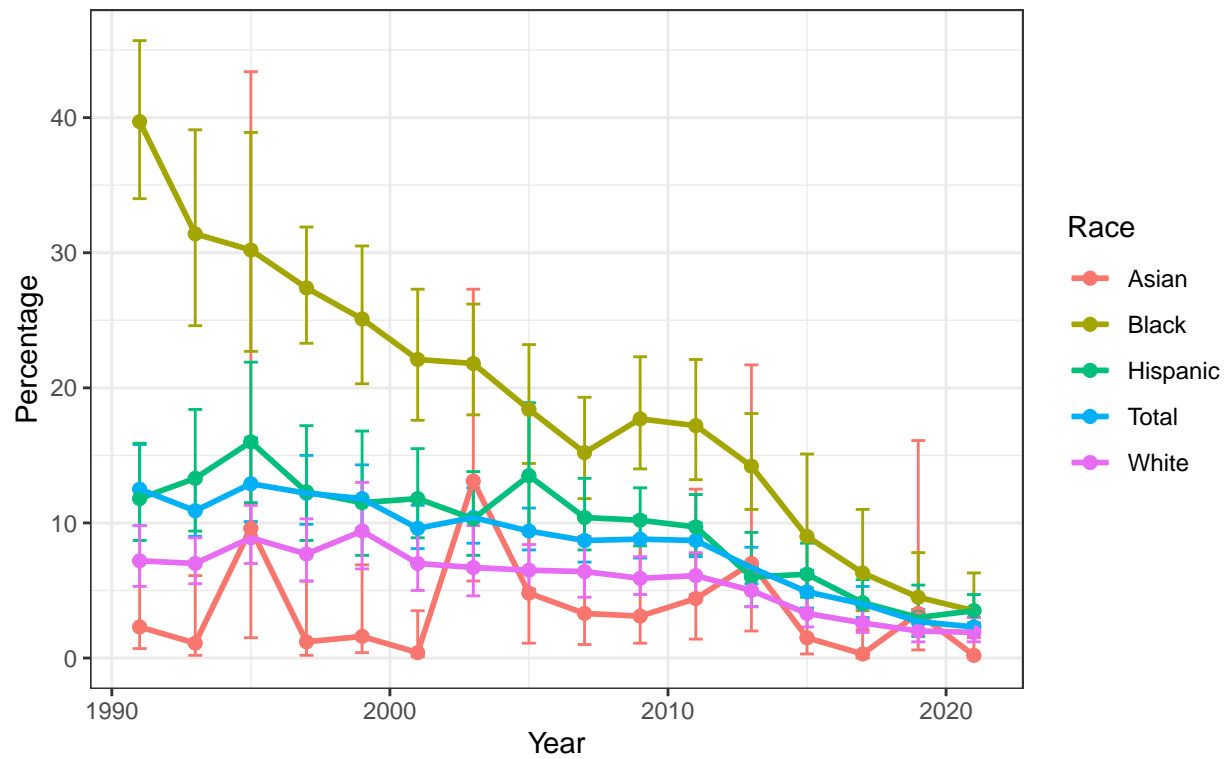
```
##
## [[5]]
```

Percentage that Have 4 Or More Sex Partners by Race for 9th–12th Grade
 in United States (YRBS Data)
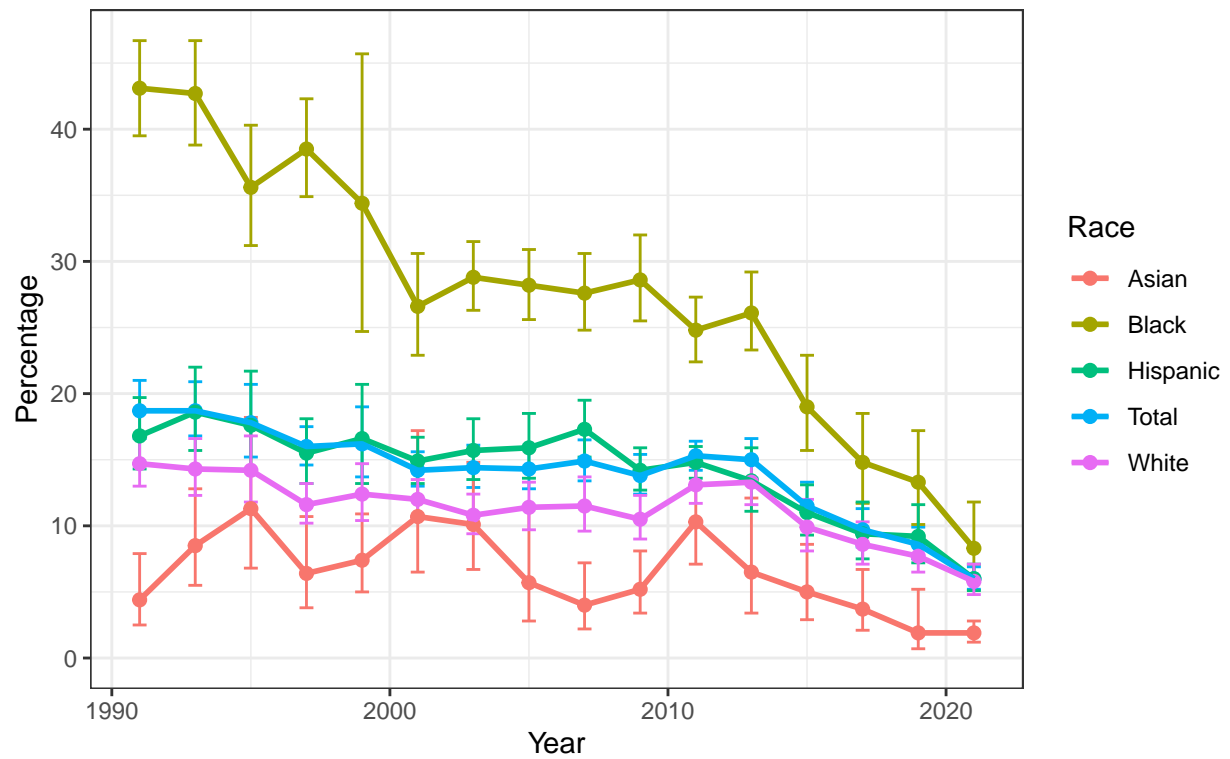


```
plots_7
```

```
## [[1]]
```

# Percentage that Have 4 Or More Sex Partners by Race for 10th Grade in New York City (YRBS Data)



```
## 
## [[2]]
```

Percentage that Have 4 Or More Sex Partners by Race for 11th Grade in New York City (YRBS Data)

```
##
## [[3]]
```

Percentage that Have 4 Or More Sex Partners by Race for 12th Grade
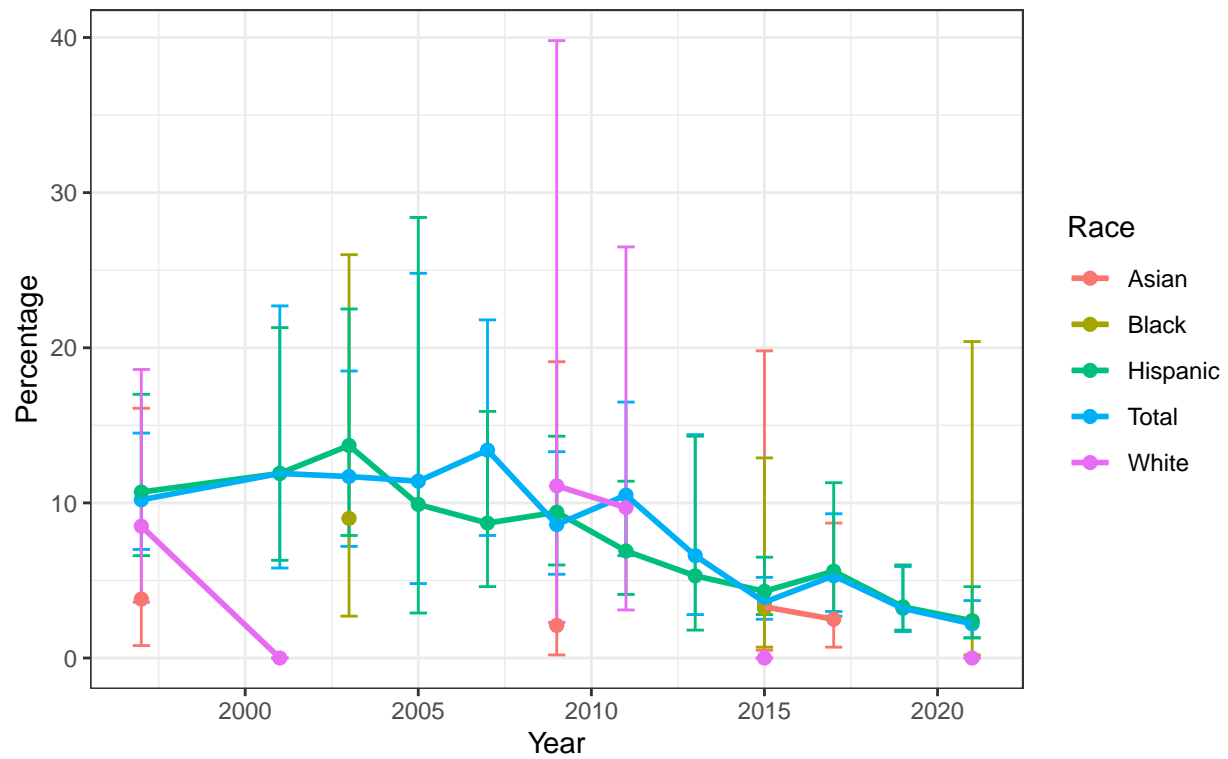 in New York City (YRBS Data)



```
## 
## [[4]]
```
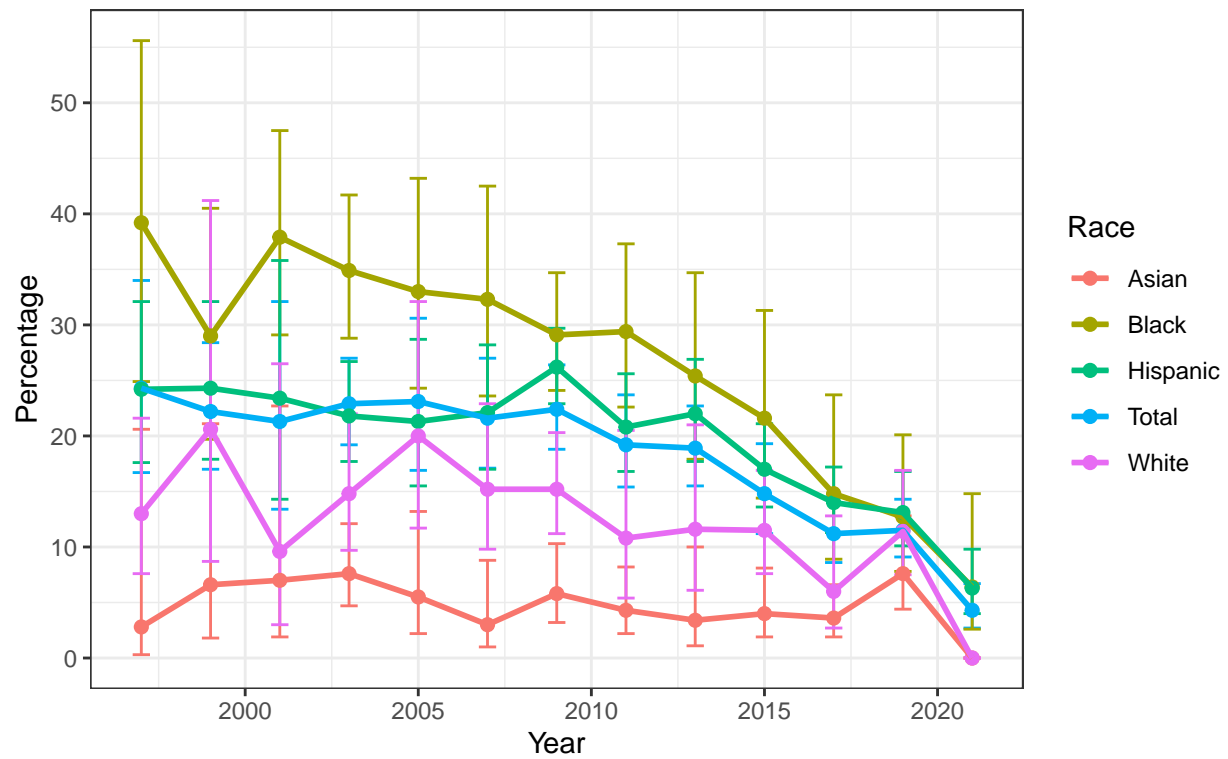
```
## Warning: Removed 2 rows containing missing values ('geom_point()').
```

Percentage that Have 4 Or More Sex Partners by Race for 9th Grade in New York City (YRBS Data)
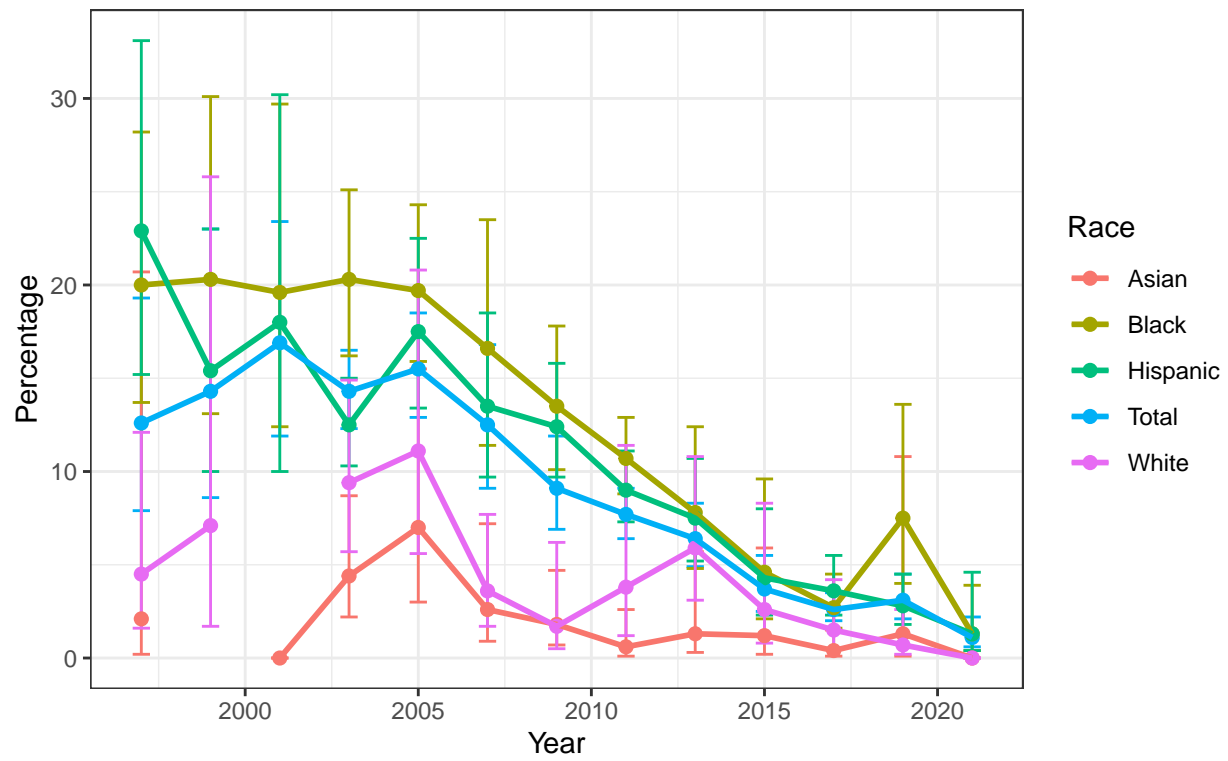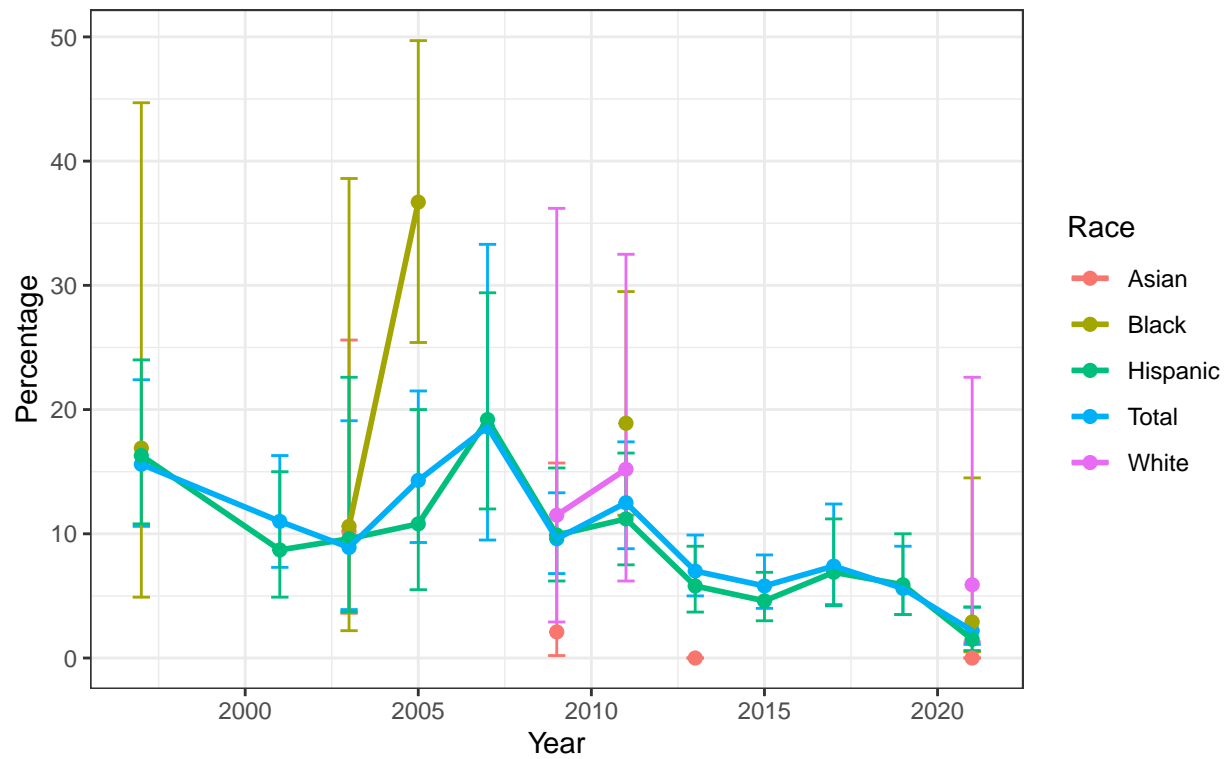
```
##
## [[5]]
```

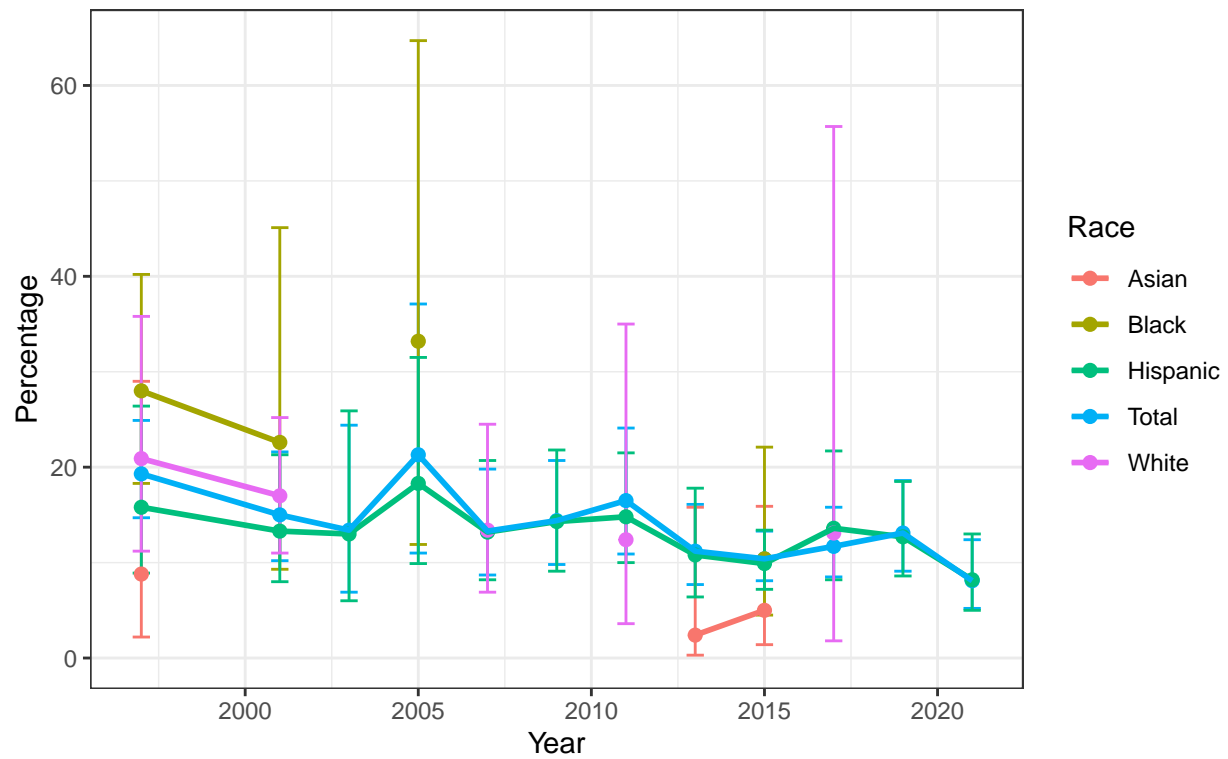Percentage that Have 4 Or More Sex Partners by Race for 9th–12th Grade in New York City (YRBS Data)
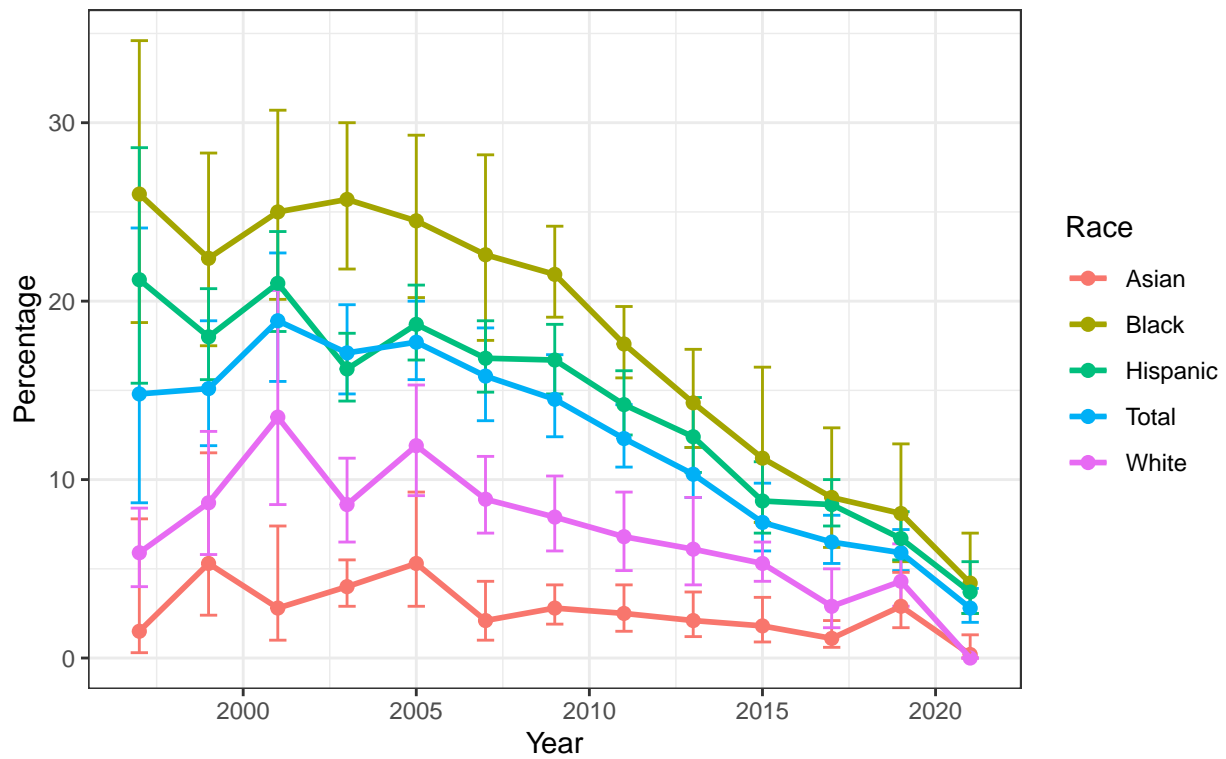
plots_8

## [[1]]

## Warning: Removed 23 rows containing missing values (`geom_point()`).

## Warning: Removed 4 rows containing missing values (`geom_line()`).

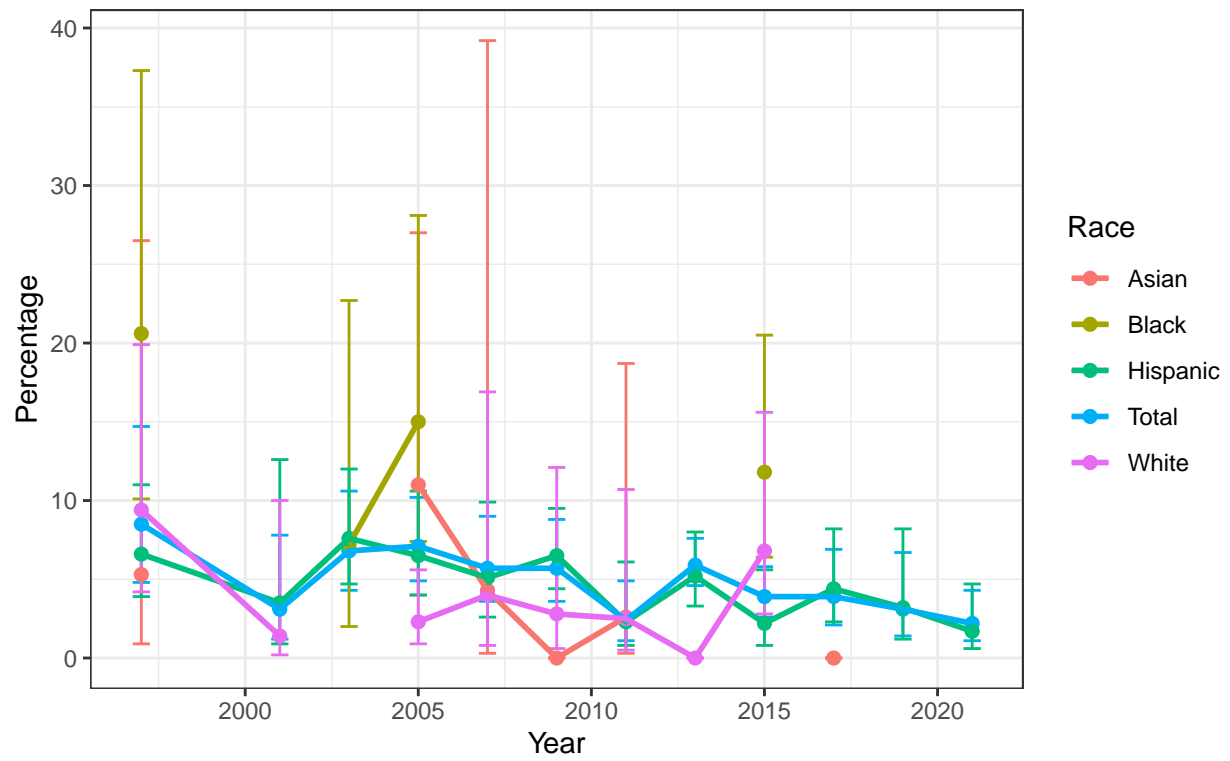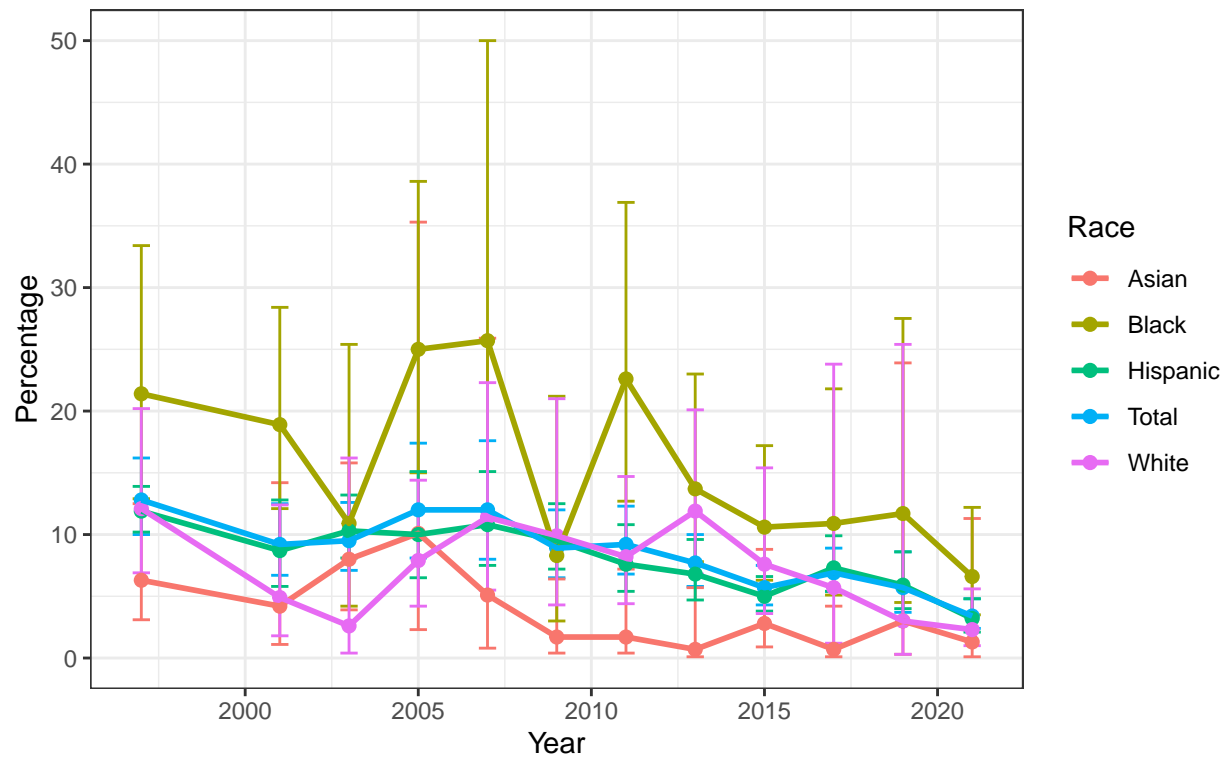Percentage that Have 4 Or More Sex Partners by Race for 10th Grade in Los Angeles (YRBS Data)

```
## 
## [[2]]
```

```
## Warning: Removed 24 rows containing missing values (`geom_point()`).
```

```
## Warning: Removed 7 rows containing missing values (`geom_line()`).
```

Percentage that Have 4 Or More Sex Partners by Race for 11th Grade in Los Angeles (YRBS Data)

```
##
## [[3]]

## Warning: Removed 24 rows containing missing values (`geom_point()`).

## Warning: Removed 8 rows containing missing values (`geom_line()`).
```

Percentage that Have 4 Or More Sex Partners by Race for 12th Grade in Los Angeles (YRBS Data)

```
##
## [[4]]
```

```
## Warning: Removed 18 rows containing missing values (`geom_point()`).
## Removed 8 rows containing missing values (`geom_line()`).
```

Percentage that Have 4 Or More Sex Partners by Race for 9th Grade in Los Angeles (YRBS Data)

```
##
## [[5]]
```

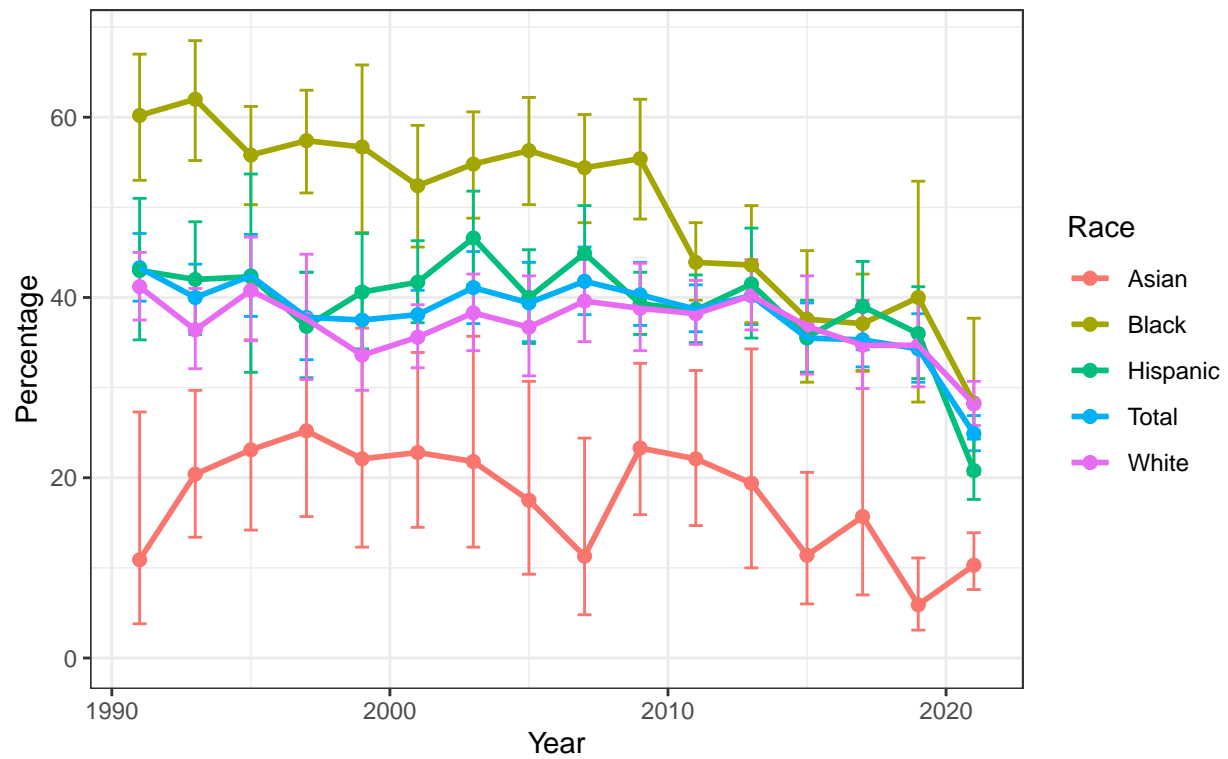Percentage that Have 4 Or More Sex Partners by Race for 9th–12th Grade in Los Angeles (YRBS Data)
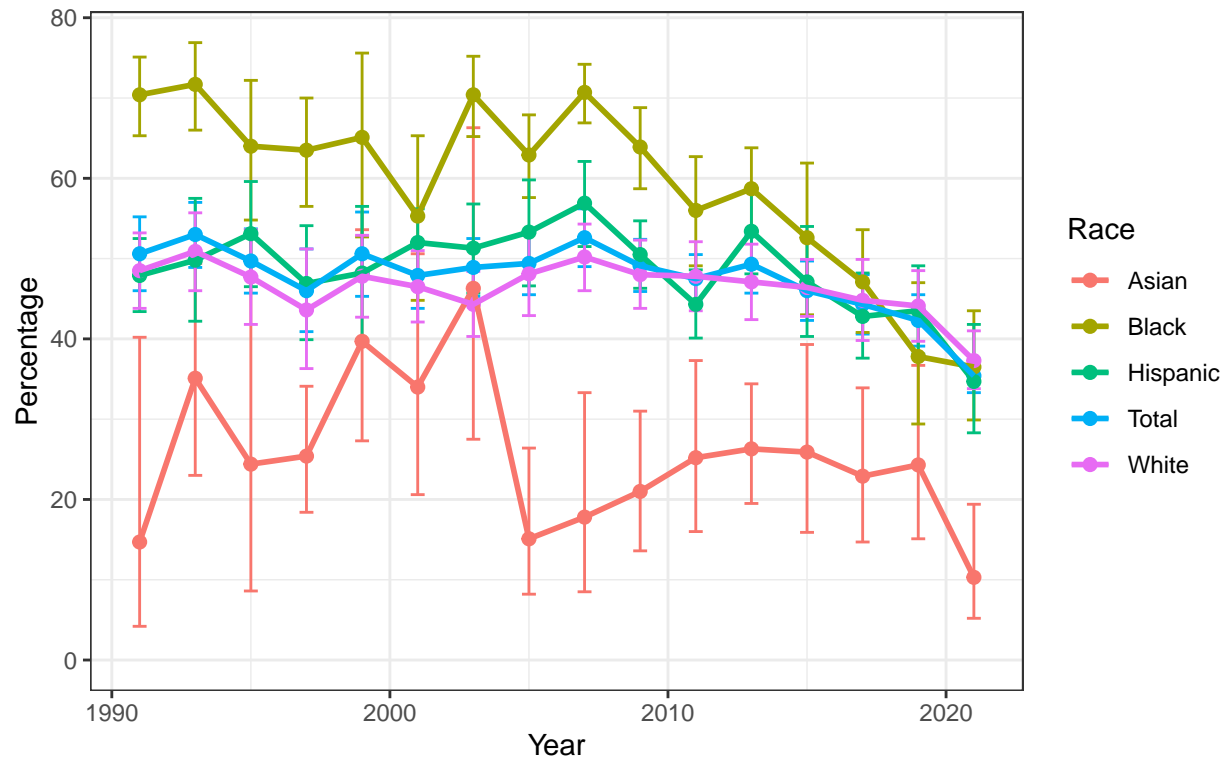
```
plots_9
```

```
## [[1]]
```

Percentage Sexually Active Past 90 Days by Race for 10th Grade
in United States (YRBS Data)
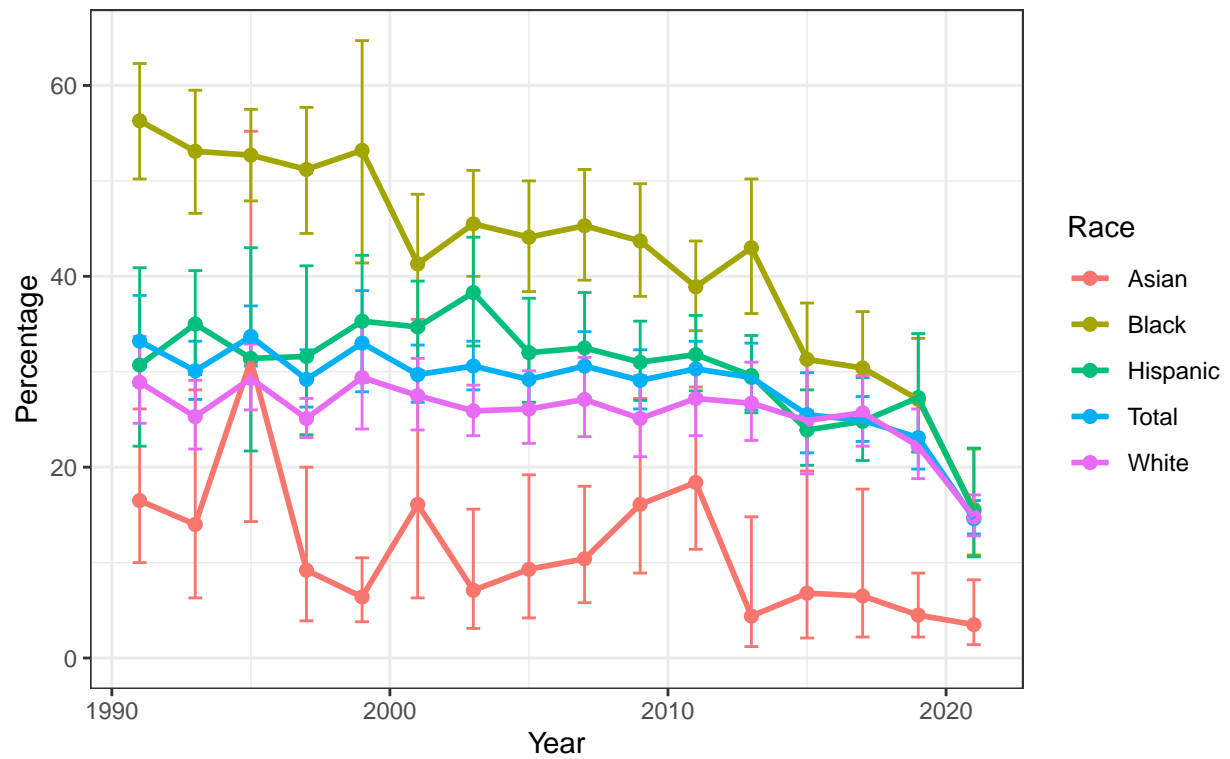
```
##
## [[2]]
```

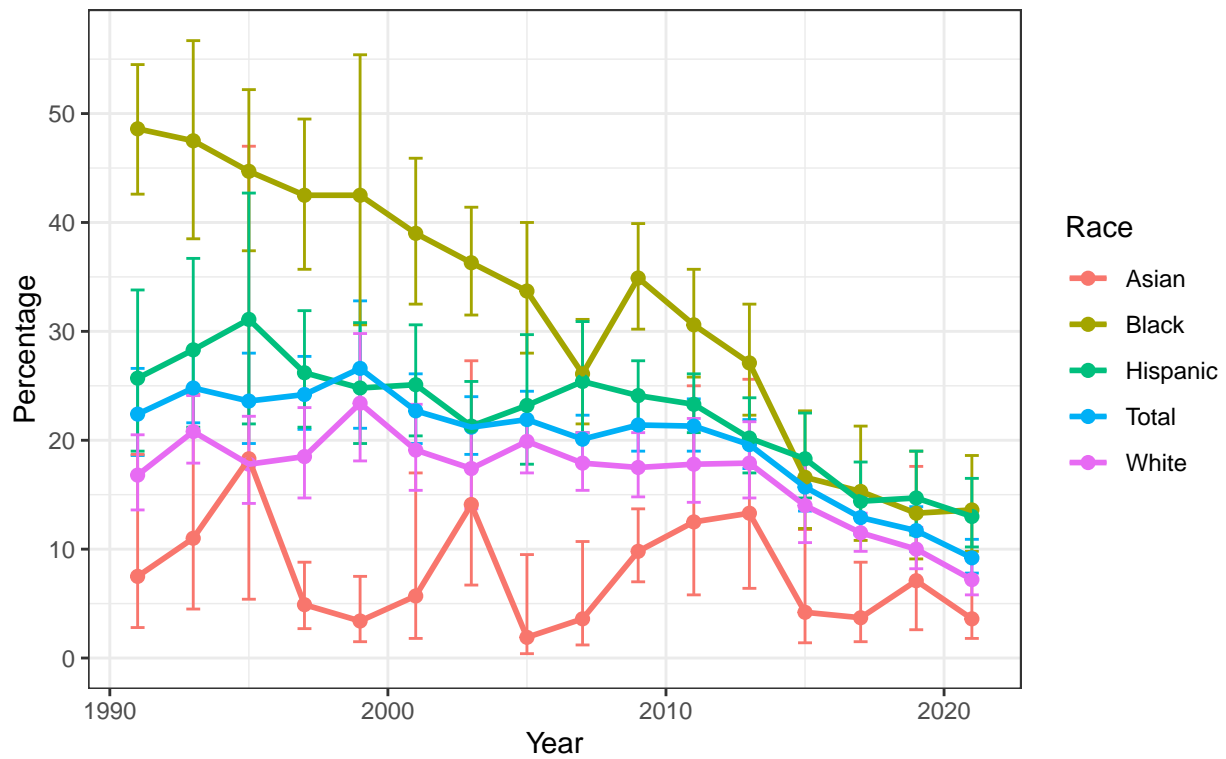Percentage Sexually Active Past 90 Days by Race for 11th Grade
in United States (YRBS Data)

```
##
## [[3]]
```

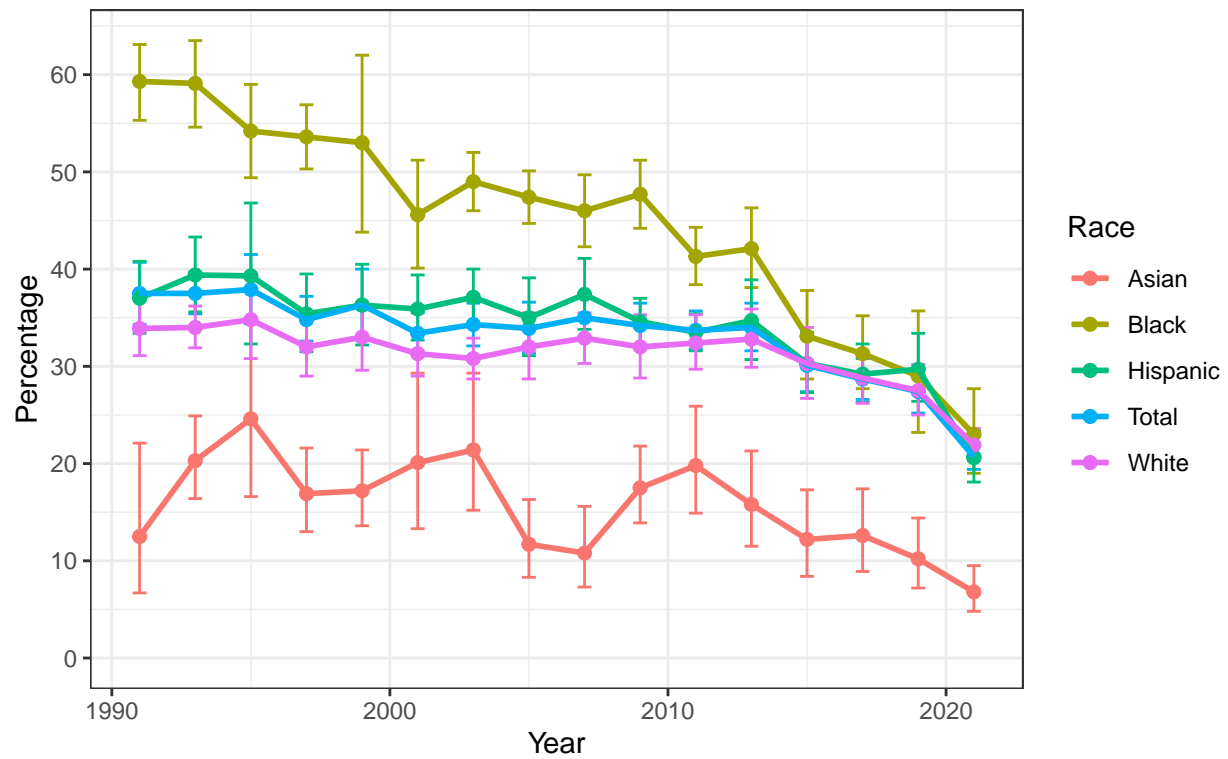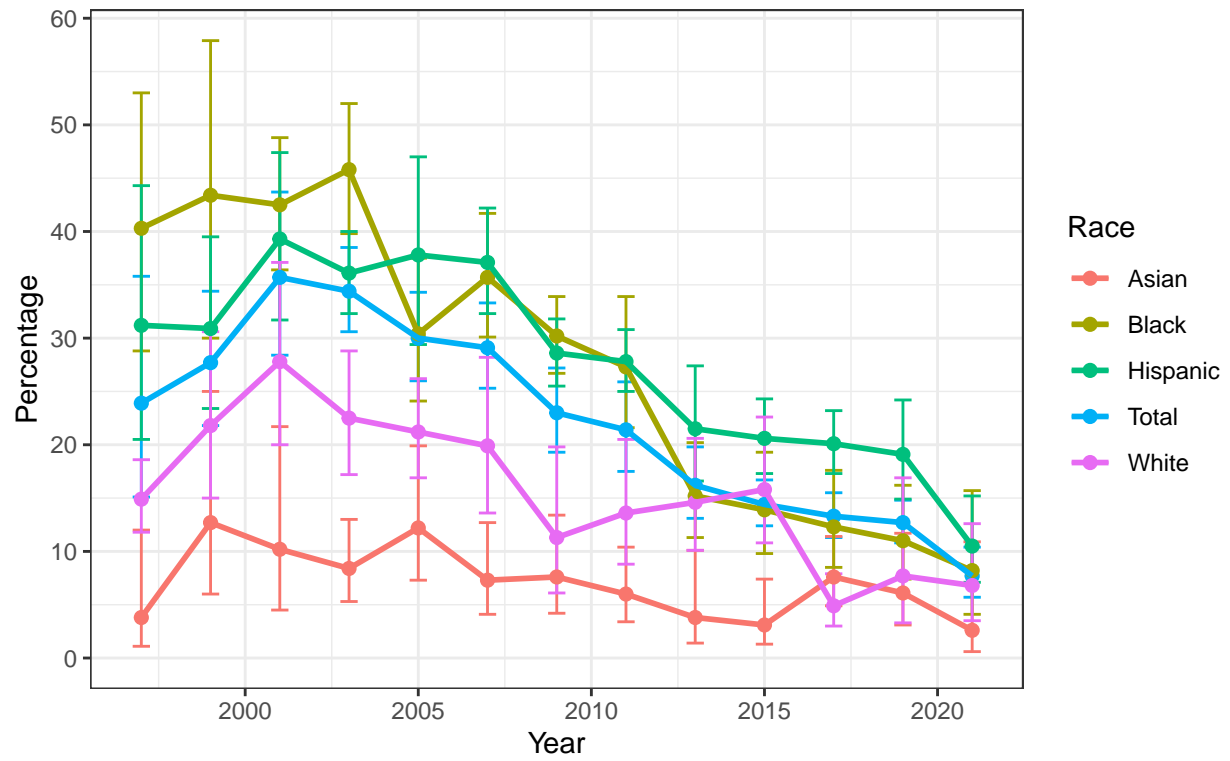Percentage Sexually Active Past 90 Days by Race for 12th Grade in United States (YRBS Data)
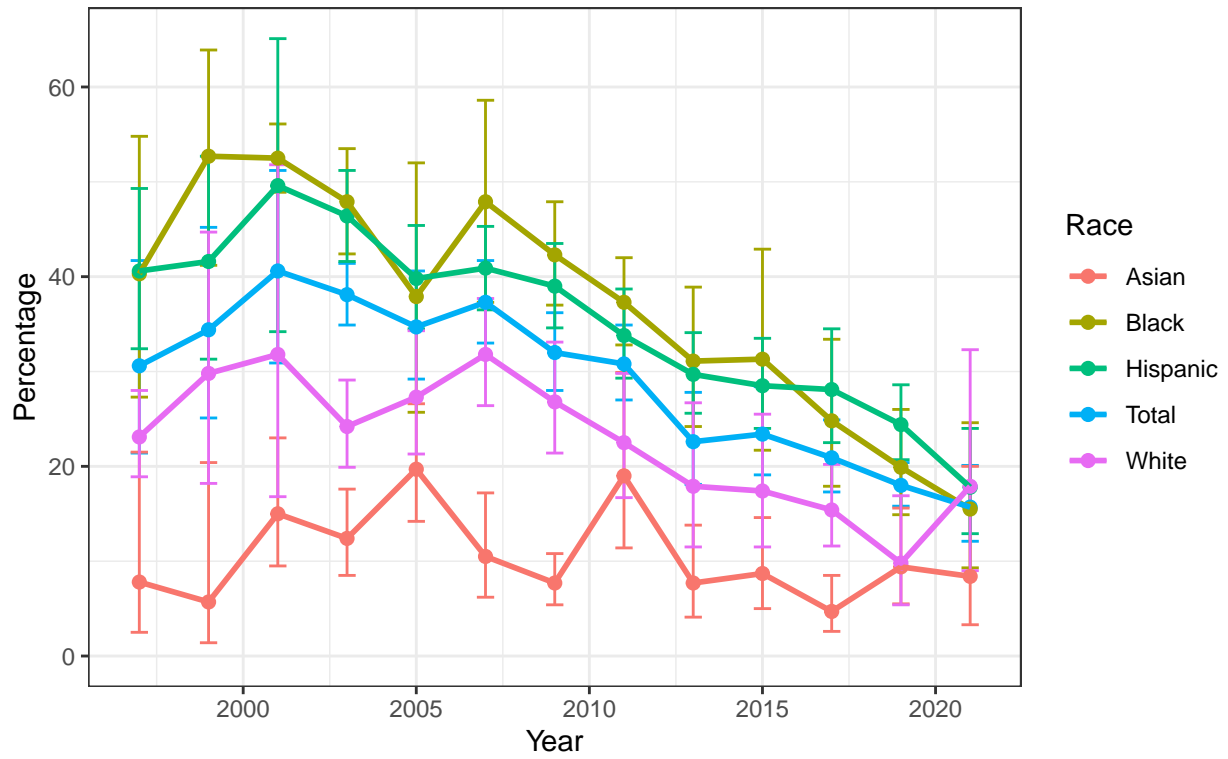
```
##
## [[4]]
```

Percentage Sexually Active Past 90 Days by Race for 9th Grade in United States (YRBS Data)

```
##
## [[5]]
```

Percentage Sexually Active Past 90 Days by Race for 9th–12th Grade
 in United States (YRBS Data)
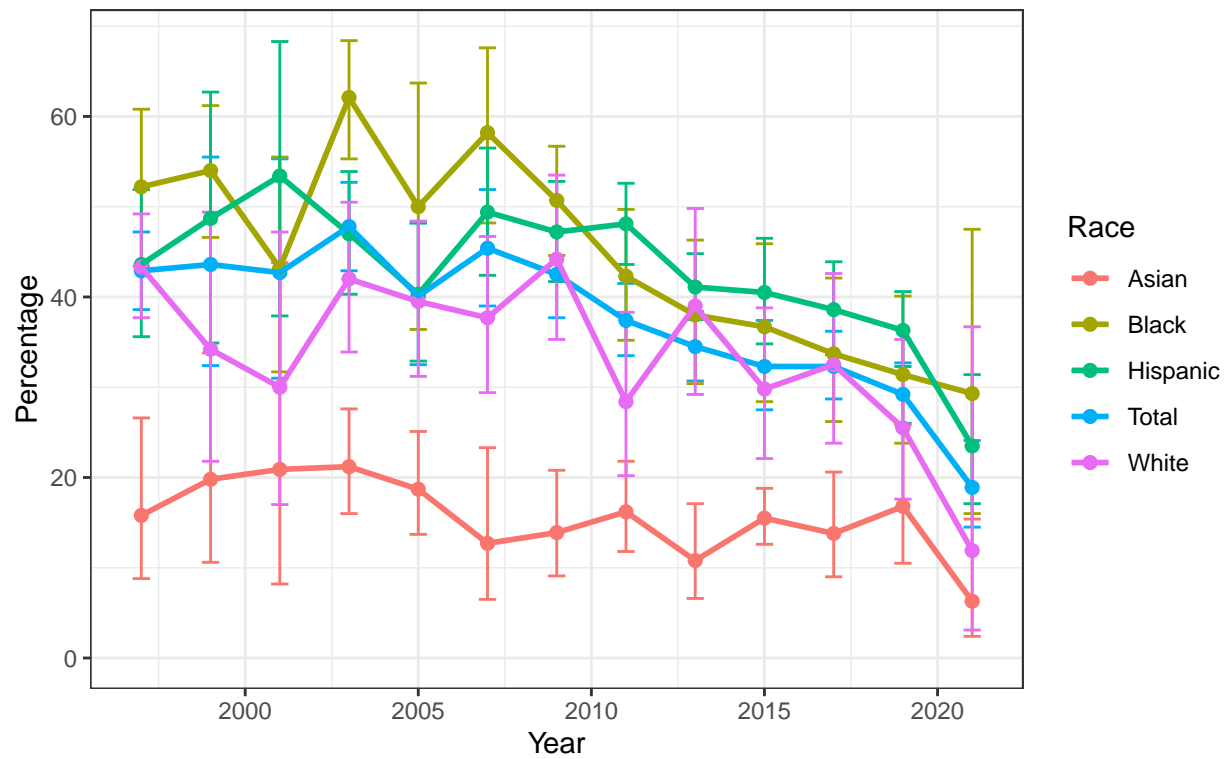


```
plots_10
```

```
## [[1]]
```

Percentage Sexually Active Past 90 Days by Race for 10th Grade in New York City (YRBS Data)

```
##
## [[2]]
```

Percentage Sexually Active Past 90 Days by Race for 11th Grade in New York City (YRBS Data)

```
##
## [[3]]
```

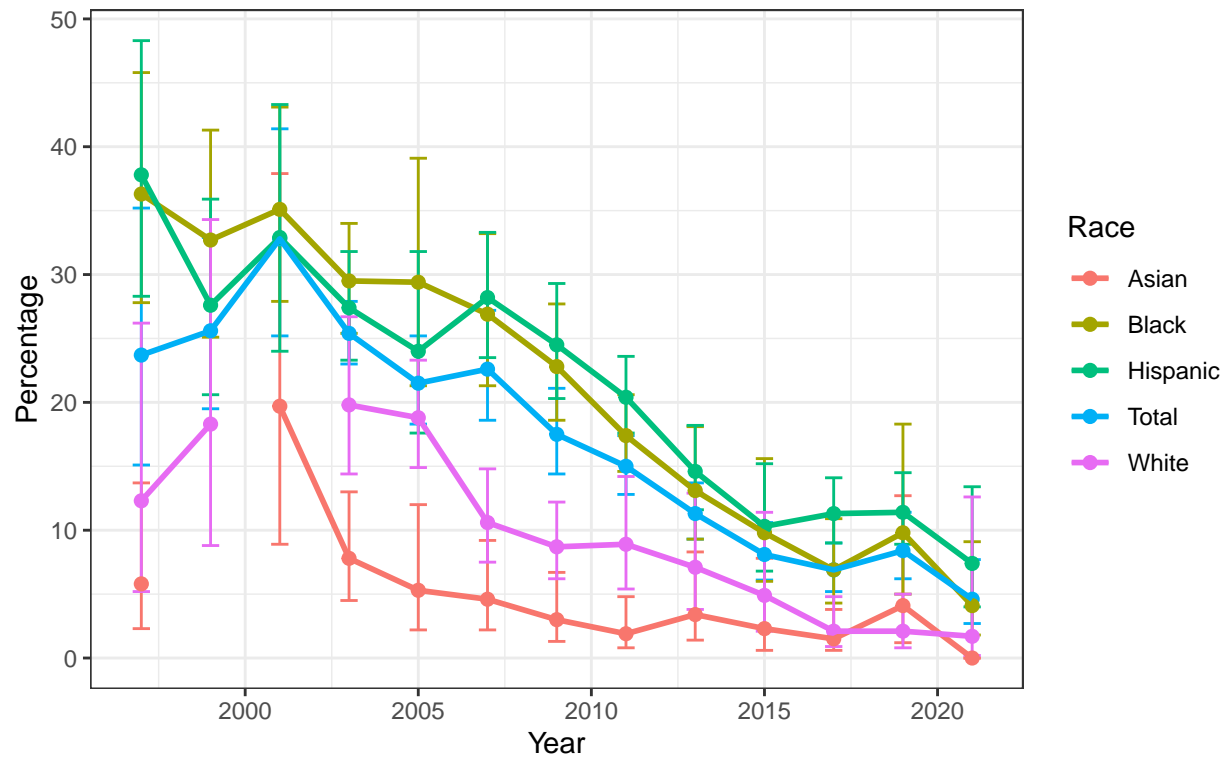Percentage Sexually Active Past 90 Days by Race for 12th Grade
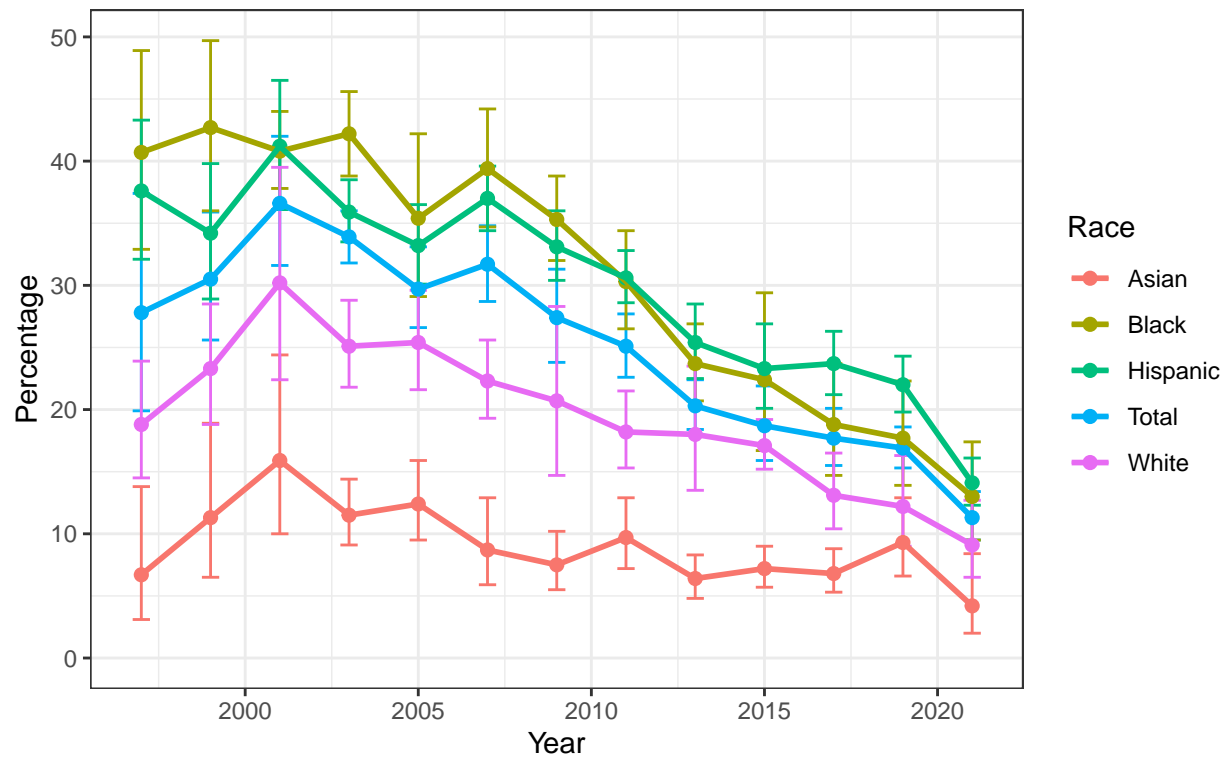in New York City (YRBS Data)



```
## 
## [[4]]
```

```
## Warning: Removed 2 rows containing missing values ('geom_point()').
```

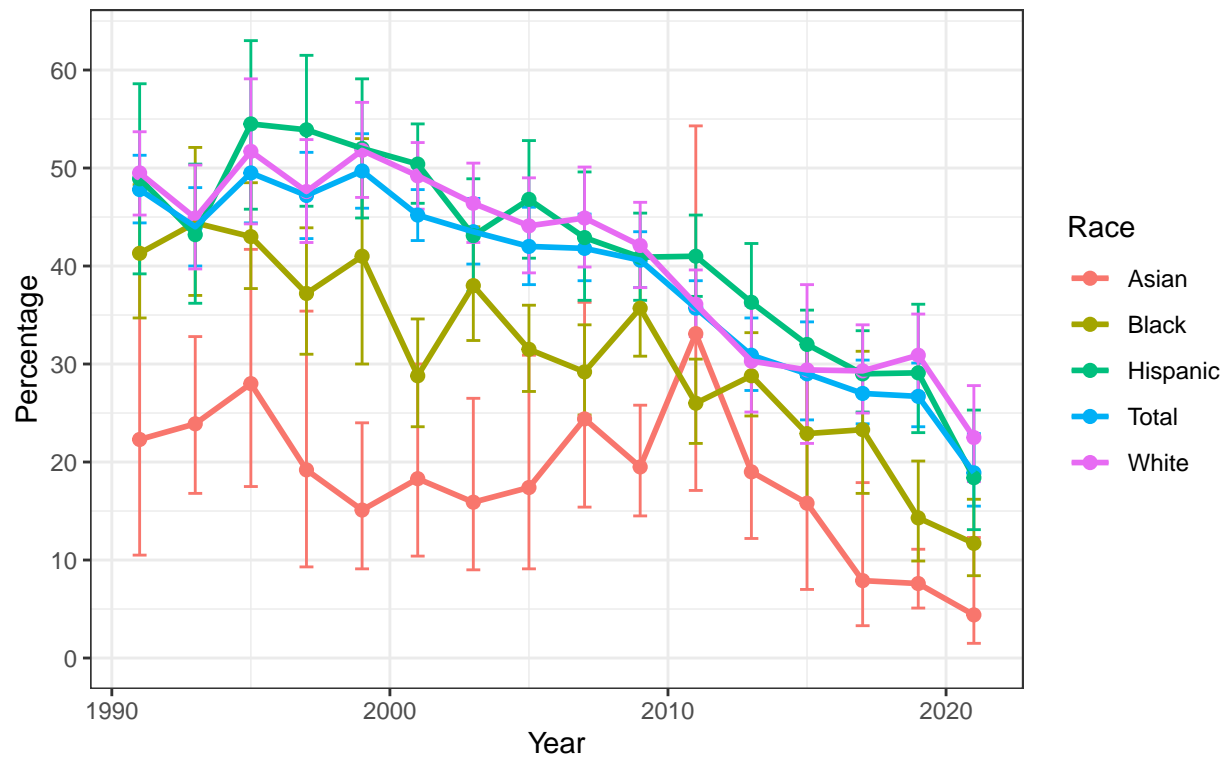Percentage Sexually Active Past 90 Days by Race for 9th Grade in New York City (YRBS Data)

```
## 
## [[5]]
```

Percentage Sexually Active Past 90 Days by Race for 9th–12th Grade in New York City (YRBS Data)

```
plots_11
```
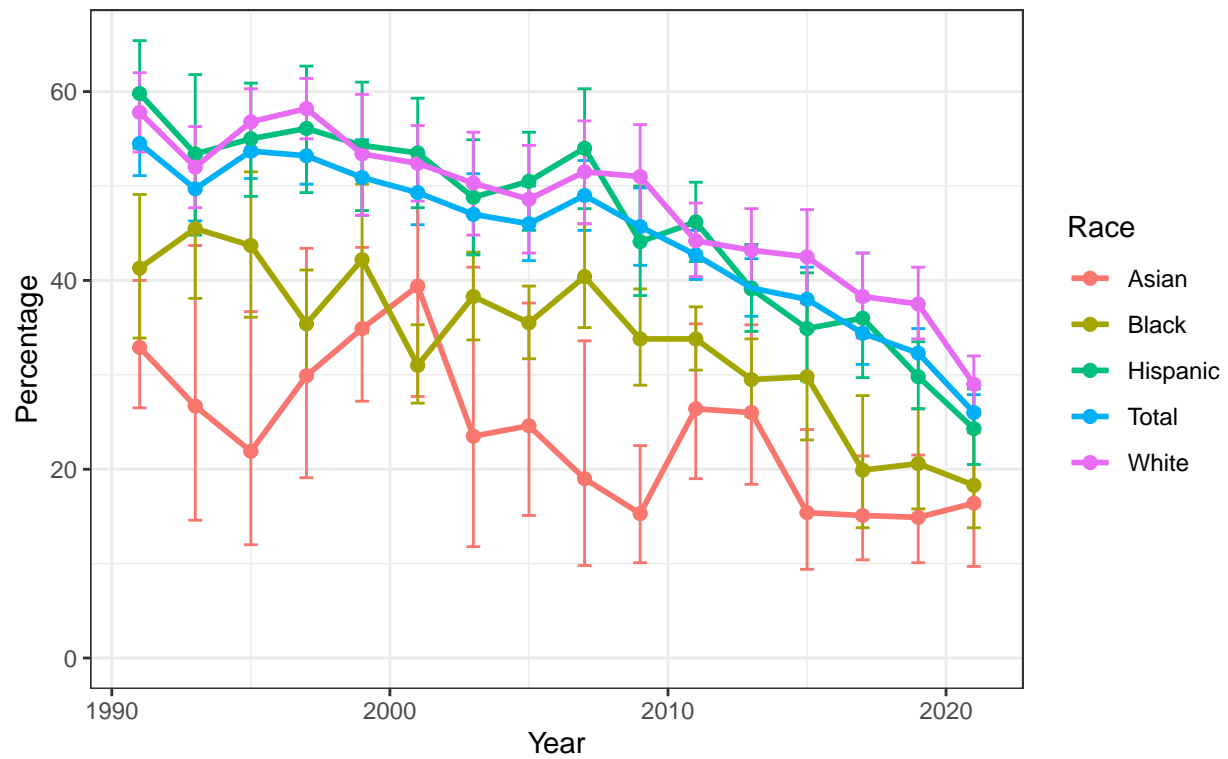
```
## [[1]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 10th Grade
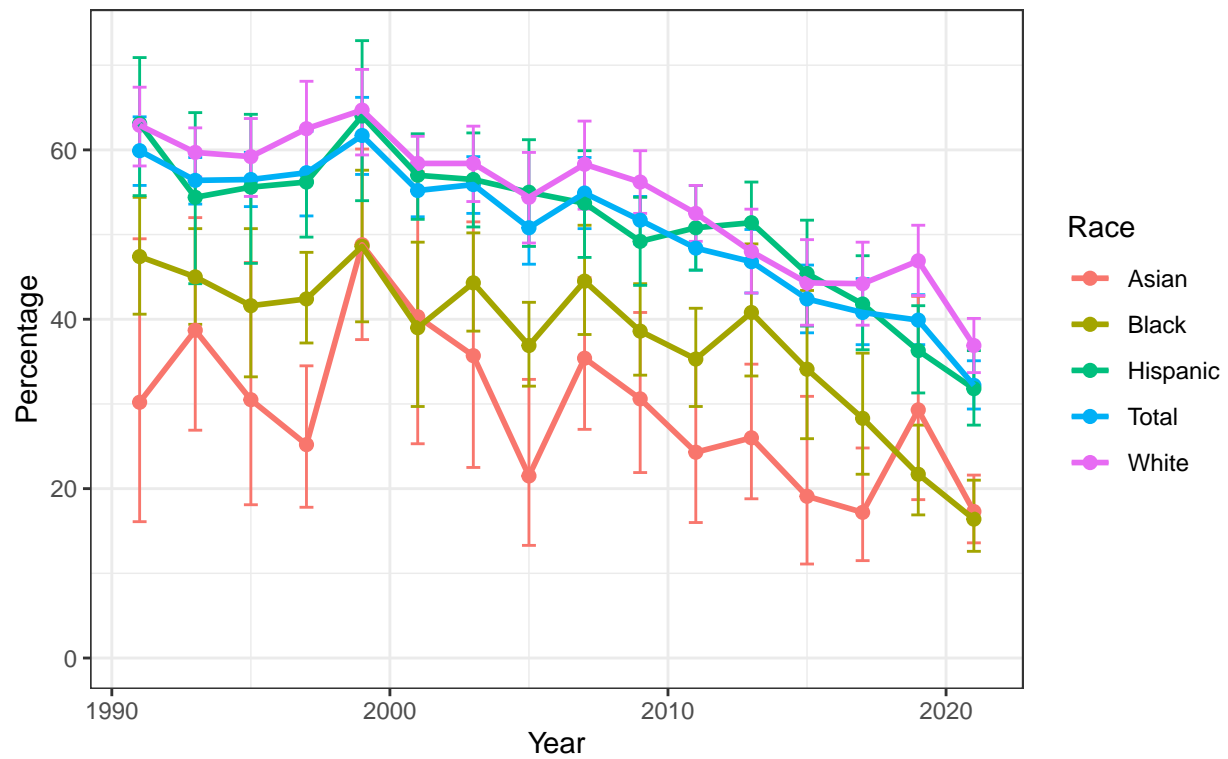 in United States (YRBS Data)



```
## 
## [[2]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 11th Grade
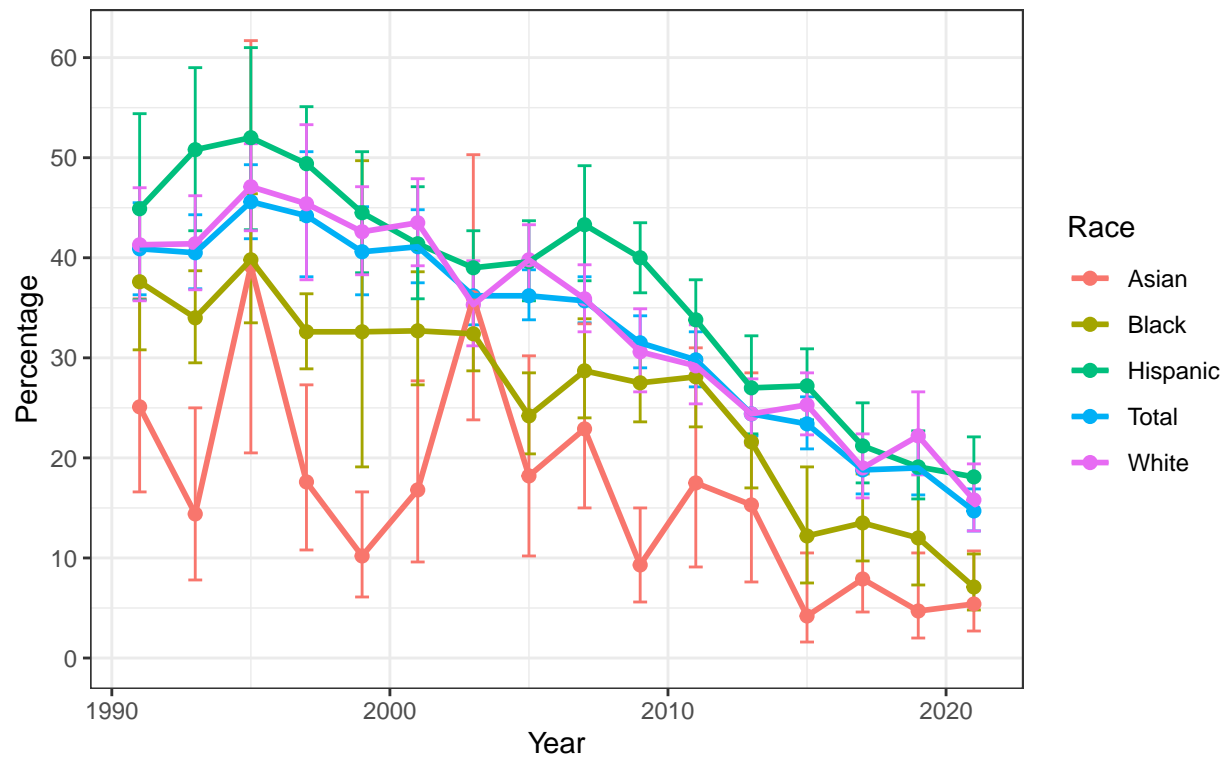in United States (YRBS Data)



```
##
## [[3]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 12th Grade in United States (YRBS Data)
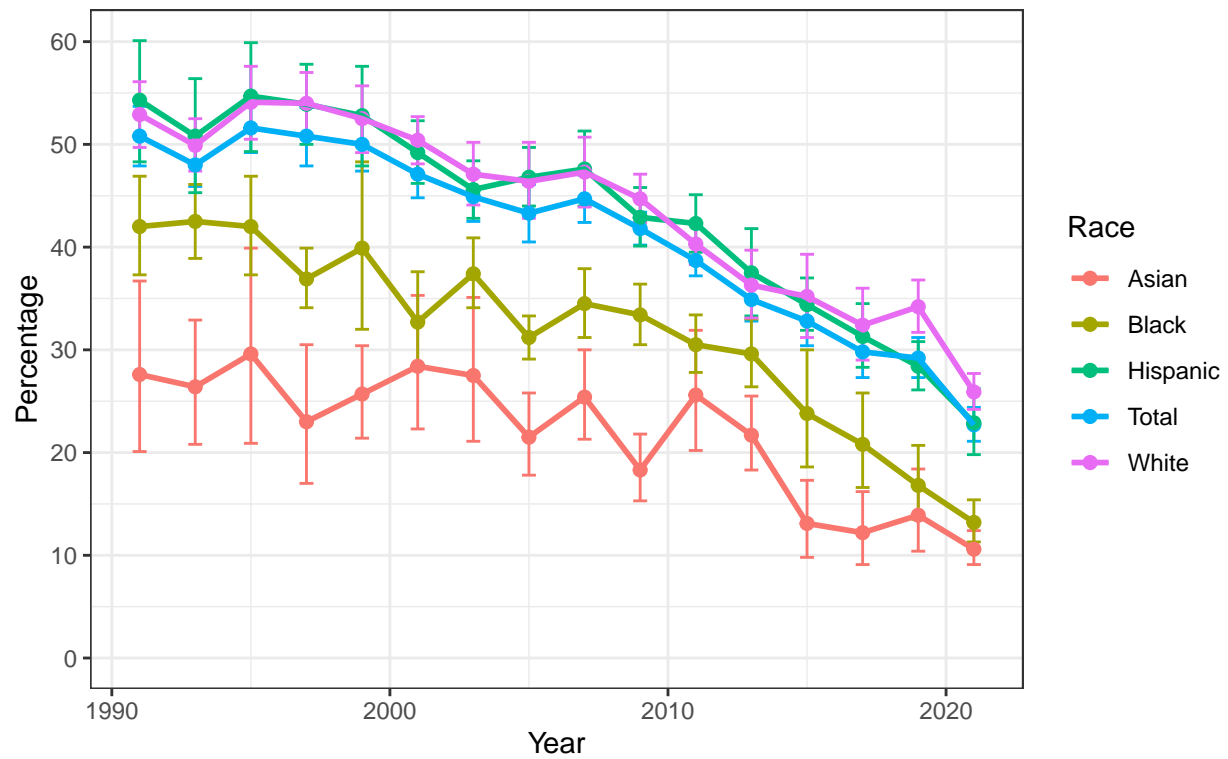
```
## 
## [[4]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 9th Grade
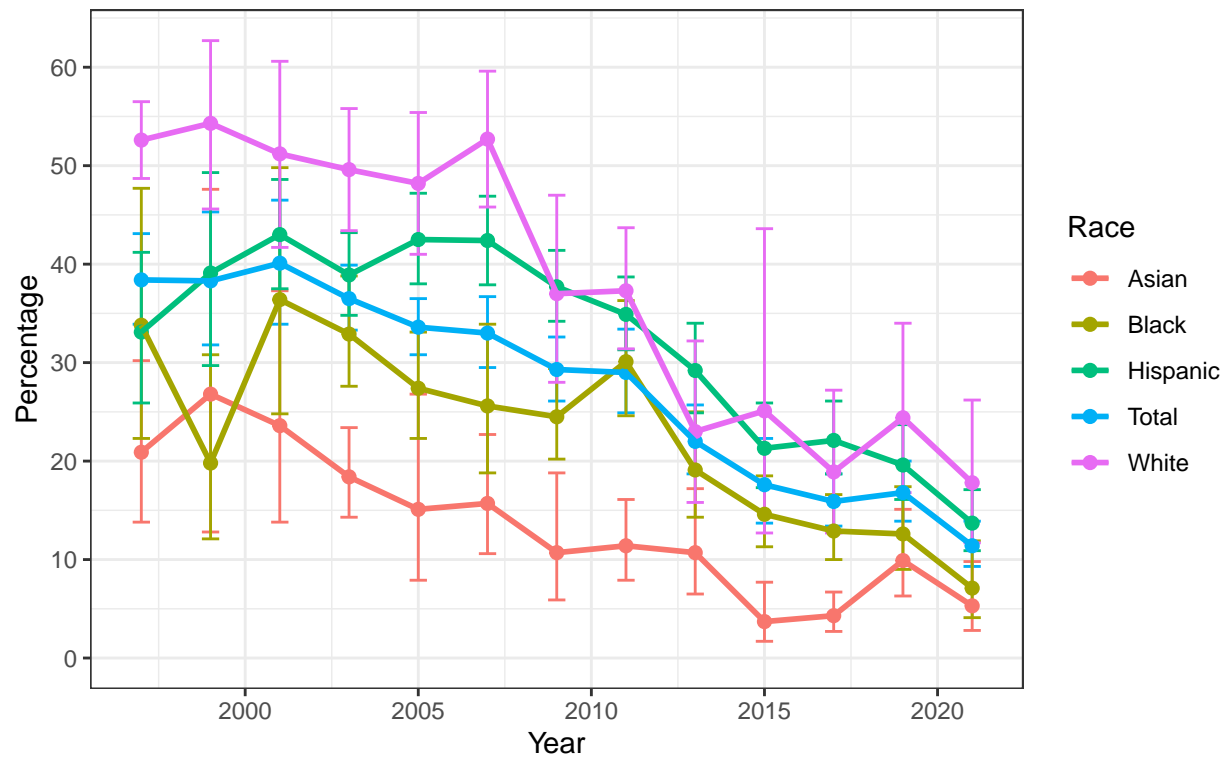in United States (YRBS Data)

```
##
## [[5]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 9th–12th Grade in United States (YRBS Data)
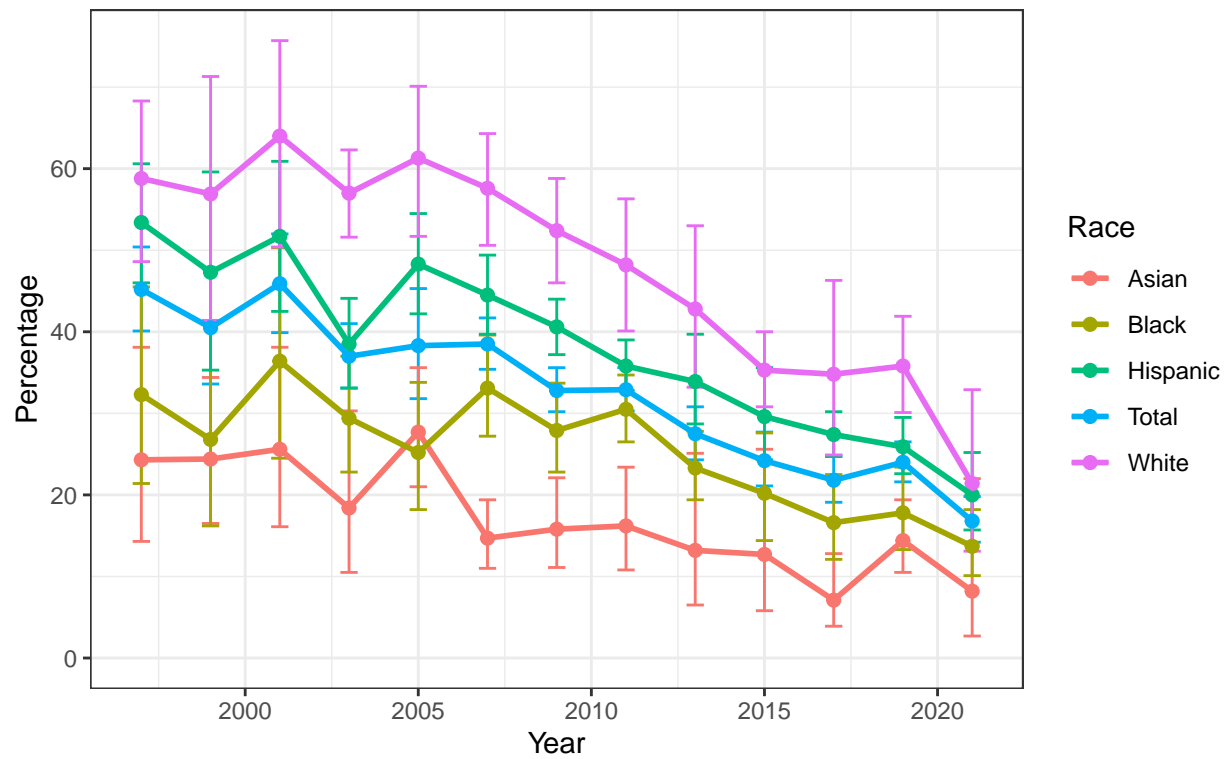
```
plots_12
```

```
## [[1]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 10th Grade in New York City (YRBS Data)
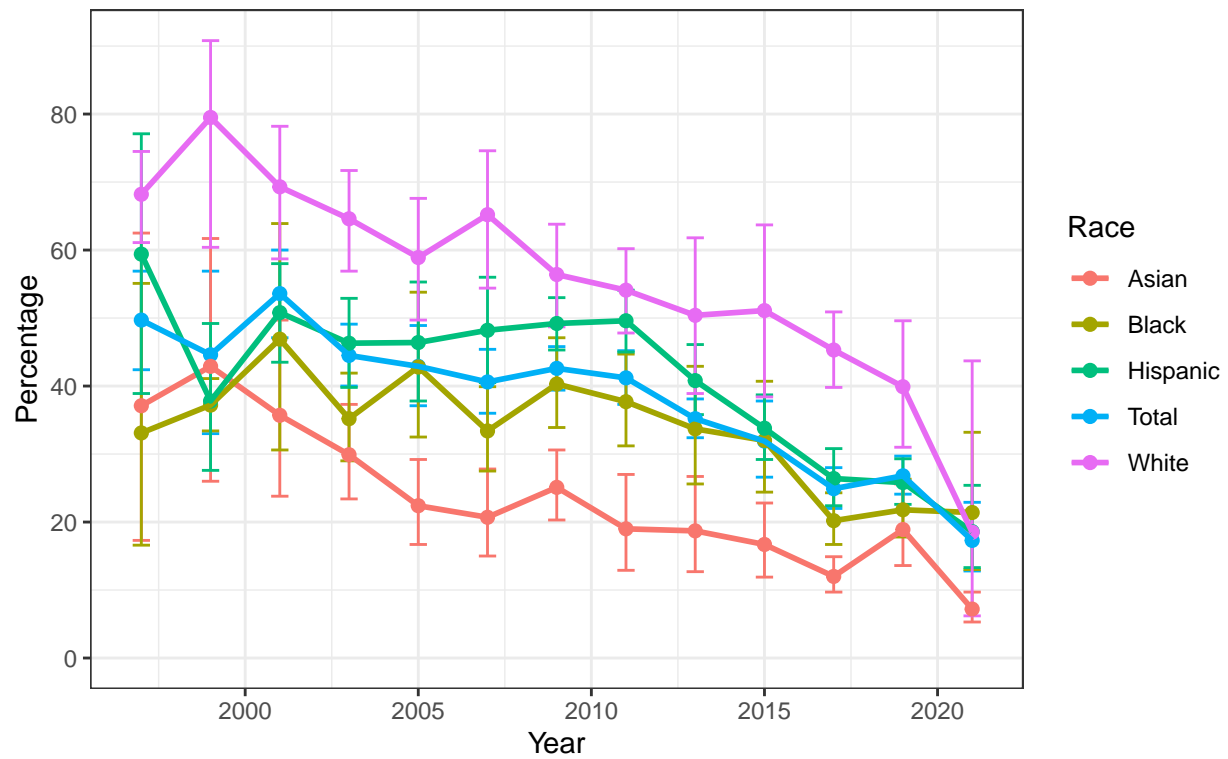
```
##
## [[2]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 11th Grade in New York City (YRBS Data)

```
## 
## [[3]]
```

Percentage that Drank Alcohol Past 90 Days by Race for 12th Grade in New York City (YRBS Data)

```
##
## [[4]]

## Warning: Removed 2 rows containing missing values ('geom_point()').
```

Percentage that Drank Alcohol Past 90 Days by Race for 9th Grade in New York City (YRBS Data)

```
##
## [[5]]
```

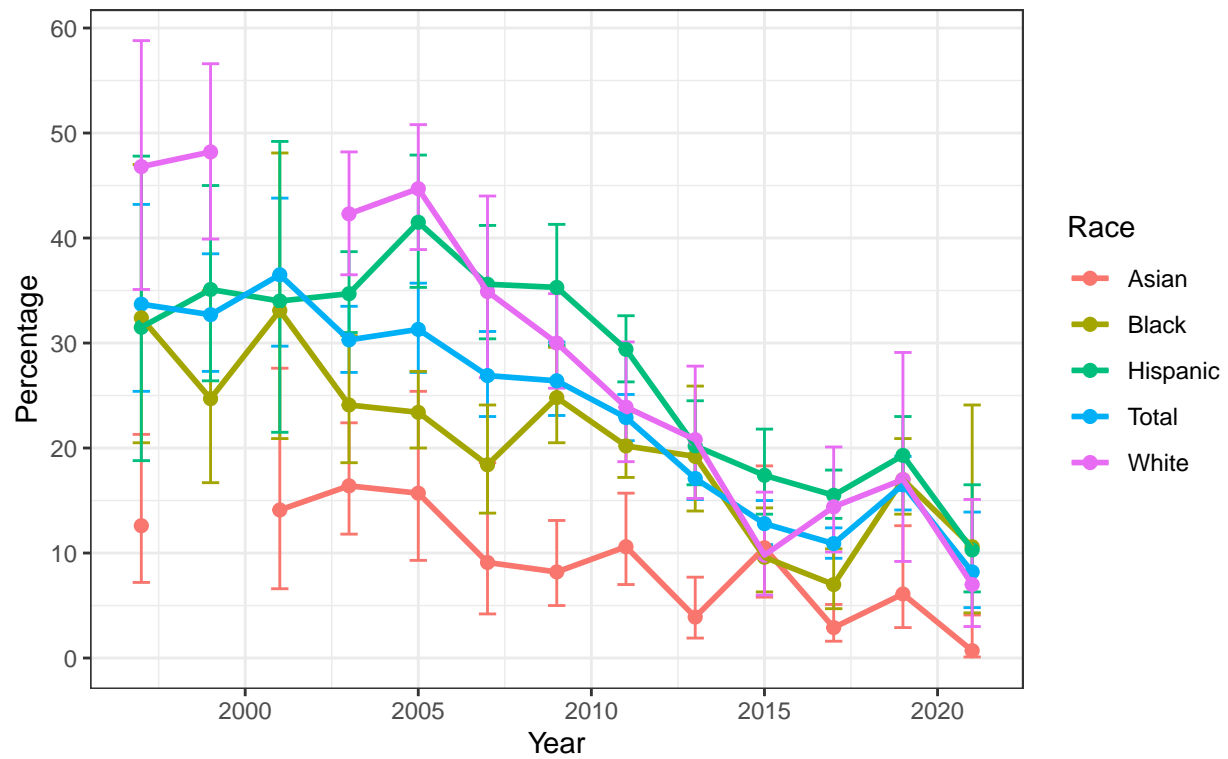Percentage that Drank Alcohol Past 90 Days by Race for 9th–12th Grade in New York City (YRBS Data)