

# CDS101 Final Project

Tyson Johnson

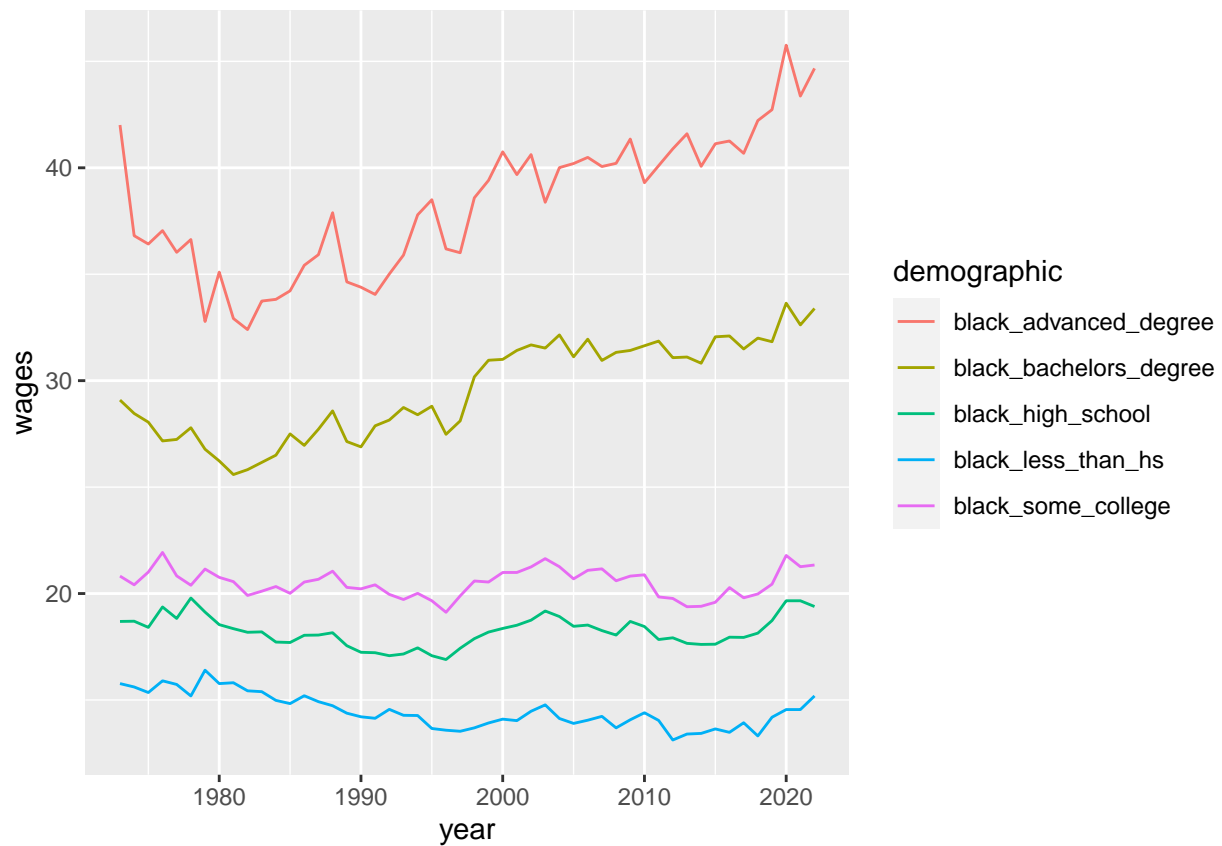
2023-10-25

## Summary Table

```
wages_tidy %>%
  group_by(demographic) %>%
  summarize(
    mean = mean(wages, na.rm=TRUE),
    median = median(wages, na.rm=TRUE),
    min = min(wages, na.rm=TRUE),
    max = max(wages, na.rm=TRUE),
    sd = sd(wages, na.rm=TRUE),
    iqr = IQR(wages, na.rm=TRUE)
  )
```

```
## # A tibble: 60 x 7
##   demographic      mean median   min   max    sd   iqr
##   <chr>          <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 advanced_degree  43.9   44.1  35.3  53.7  5.31  8.80
## 2 bachelors_degree 34.8   34.2  30.0  41.6  3.31  5.25
## 3 black_advanced_degree 38.3   38.5  32.4  45.8  3.27  4.76
## 4 black_bachelors_degree 29.6   29.6  25.6  33.6  2.27  3.96
## 5 black_high_school 18.2   18.2  16.9  19.8  0.712 0.985
## 6 black_less_than_hs 14.5   14.3  13.1  16.4  0.803 1.21
## 7 black_men_advanced_degree 41.2   40.7  34.3  52.9  4.53  6.38
## 8 black_men_bachelors_degree 31.7   31.9  27.9  37.6  2.43  3.91
## 9 black_men_high_school 20.0   20.0  18.1  22.8  1.11  1.53
## 10 black_men_less_than_hs 16.0   15.5  14.1  18.6  1.29  1.97
## # i 50 more rows
```

```
wages_tidy %>%
  filter(grepl("black", demographic)) %>%
  filter(!grepl("men", demographic)) %>%
  ggplot() +
  geom_line(
    mapping = aes(
      x = year,
      y = wages,
      color = demographic
    )
  )
)
```



```
wages_ethnicity <- wages_tidy %>%
  filter(grepl(c("black|hispanic|white"), demographic)) %>%
  separate(demographic, into = c("ethnicity", "demographic"),
    sep = "_", extra = "merge") %>%
  arrange(ethnicity, demographic, year)

wages_ethnicity_inverse <- wages_tidy %>%
  filter(!grepl(c("black|hispanic|white"), demographic)) %>%
  mutate(ethnicity = NA) %>%
  relocate(ethnicity, .after = year)

wages_ethnicity_combined <- wages_ethnicity %>%
  bind_rows(wages_ethnicity_inverse) %>%
```

```
arrange(ethnicity, demographic, year)
```

```
wages_gender <- wages_ethnicity_combined %>%  
  filter(grepl(c("men|women"), demographic)) %>%  
  separate(demographic, into = c("gender", "demographic"),  
    sep = "_", extra = "merge") %>%  
  arrange(ethnicity, gender, demographic, year)
```

```
wages_gender_inverse <- wages_ethnicity_combined %>%  
  filter(!grepl(c("men|women"), demographic)) %>%  
  mutate(gender = NA) %>%  
  relocate(gender, .after = ethnicity)
```

```
wages_gender_combined <- wages_gender %>%  
  bind_rows(wages_gender_inverse) %>%  
  arrange(ethnicity, gender, demographic, year)
```

```
wages_sep <- wages_gender_combined
```

```
wages_sep %>%  
  write.csv("wages_sep.csv", row.names=FALSE)
```