Problem Set 1: PSID - Labor Outcomes

Tate Mason

```
df <- read_dta("~/SchoolWork/Y2S1/Macro/Data/PSID/PSID.dta")
```

```
df <- df %>%
  mutate(
    famid = coalesce(!!!rlang::syms(c(
      "ER66009","ER60009","ER53009","ER47309","ER42009",
      "ER36009","ER25009","ER21009","ER17022","ER13019"
    )))
  )

long <- map_dfr(1:nrow(pivot), function(i){
  sel <- pivot[i, ]
  df %>%
    transmute(
      famid,
      year = sel$year,
      sex = .data[[sel$sex]],
      age = .data[[sel$age]],
      inc = .data[[sel$earnings_annual]],
      labor_par = .data[[sel$labor_par]],
      hourly = .data[[sel$hourly]],
      hr_worked = .data[[sel$hr_worked]],
      educ_HS = .data[[sel$educ_HS]],
      educ_coll = .data[[sel$educ_coll]],
      ind = .data[[sel$ind]],
      wealth = .data[[sel$wealth]],
      cpi_ratio = sel$cpi_ratio
    )
})

glimpse(long)
```

```
Rows: 188,900
Columns: 13
$ famid     <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, NA, 4, 4, 4, 4,~
$ year      <dbl> 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, ~
$ sex       <dbl> 1, 1, 2, NA, 2, 1, 2, NA, 1, 1, 2, NA, NA, NA, NA, NA, 1, 1,~
$ age       <dbl> 76, 47, 45, NA, 39, 34, 25, NA, 23, 22, 20, NA, NA, NA, NA, ~
$ inc       <dbl> 0, 400, 0, NA, 0, 600, 0, NA, 0, 0, 0, NA, NA, NA, NA, NA, 5~
$ labor_par <dbl> 0, 0, 0, NA, 0, 0, 0, NA, 0, 0, 0, NA, NA, NA, NA, NA, 0, 0,~
$ hourly    <dbl> 0.00, 0.00, 0.00, NA, 15.35, 0.00, 0.00, NA, 8.50, 8.00, 5.0~
$ hr_worked <dbl> 0, 40, 65, NA, 45, 50, 0, NA, 40, 30, 40, NA, NA, NA, NA, NA~
$ educ_HS   <dbl> 3, 3, 1, NA, 1, 3, 1, NA, 3, 1, 1, NA, NA, NA, NA, NA, 1, 1,~
$ educ_coll <dbl> 0, 0, 0, NA, 0, 0, 0, NA, 0, 5, 0, NA, NA, NA, NA, NA, 1, 0,~
$ ind       <dbl> 0, 628, 669, NA, 907, 69, 0, NA, 139, 67, 669, NA, NA, NA, N~
$ wealth    <dbl> 91500, 26000, 413500, NA, 42000, 20000, 12700, NA, 3000, 125~
$ cpi_ratio <dbl> 1.928571, 1.928571, 1.928571, 1.928571, 1.928571, 1.928571, ~
```

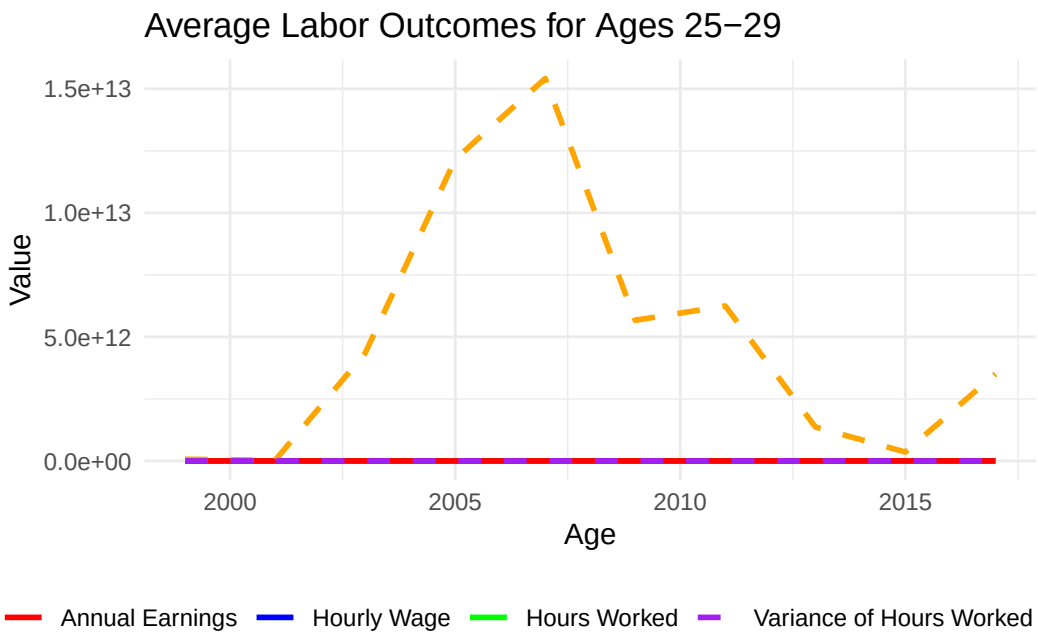
```r
long <- long %>%
  filter(
    age >= 25 & age <= 60,
    sex == 1
  ) %>%
  group_by(year) %>%
  mutate(
    wealth_adj = wealth * cpi_ratio,
    inc_adj = inc * cpi_ratio,
    hourly_adj = hourly * cpi_ratio
  )

long <- long %>%
  group_by(year) %>%
  mutate(
    labor_par = as.numeric(labor_par == 0),
    blue_col = as.numeric(ind %in% c(range(67:77), range(47:57), range(17:28), 107:398)),
    white_col = as.numeric(ind %in% c(range(407:479), range(707:718), range(727:759), range(8
    educ_HS = as.numeric(educ_HS == 1),
    educ_coll = as.numeric(educ_coll == 1),
  )
glimpse(long$age)
```

```
 num [1:44819] 47 34 50 39 55 41 36 32 44 54 ...
```

```r
prof1 <- long %>%
  group_by(year) %>%
  summarise(
    avg_hourly = mean(hourly_adj, na.rm = TRUE),
    avg_hours = mean(hr_worked, na.rm = TRUE),
    avg_earnings = mean(inc_adj, na.rm = TRUE),
    var_inc = var(inc_adj, na.rm = TRUE),
    var_hours = var(hr_worked, na.rm = TRUE)
  ) %>%
  ggplot(aes(x = year)) +
  geom_line(aes(y = avg_hourly, color = "Hourly Wage"), size = 1) +
  geom_line(aes(y = avg_hours, color = "Hours Worked"),
            size = 1) +
  geom_line(aes(y = avg_earnings, color = "Annual Earnings"),
            size = 1) +
  geom_line(aes(y = var_inc, color = "Variance of Income"),
            size = 1, linetype = "dashed") +
  geom_line(aes(y = var_hours, color = "Variance of Hours Worked"),
            size = 1, linetype = "dashed") +
  labs(
    title = "Average Labor Outcomes for Ages 25-29",
    x = "Age",
    y = "Value",
    color = "Outcome"
  ) +
  scale_color_manual(values = c("Hourly Wage" = "blue", "Hours Worked" = "green", "Annual Ea
    "Variance of Income" = "orange", "Variance of Hours Worked" = "purple")) +
  theme_minimal() +
  theme(legend.position = "bottom")
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```

```r
prof1
```


Average Labor Outcomes for Ages 25−29
1.5e+13
1.0e+13
5.0e+12
0.0e+00
Value
2000
2005
2010
2015
Age
Annual Earnings
Hourly Wage
Hours Worked
Variance of Hours Worked