FinalProject_Question2

Sid Tekumalla

2025-11-19

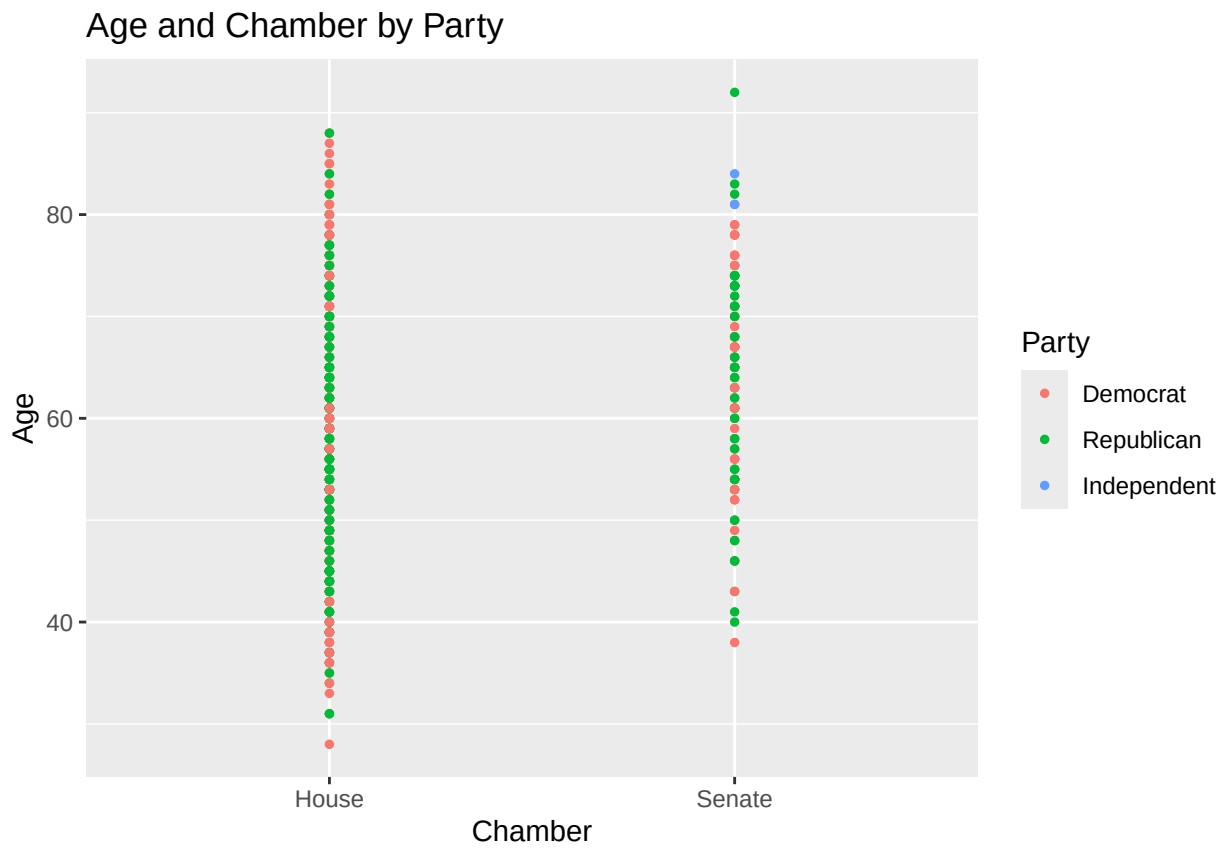
Age and Ideology

Using age data, is there a correlation between member age and ideological positioning in 2025, controlling for party? If so, are younger Republicans/Democrats different from older ones, and in what ways?

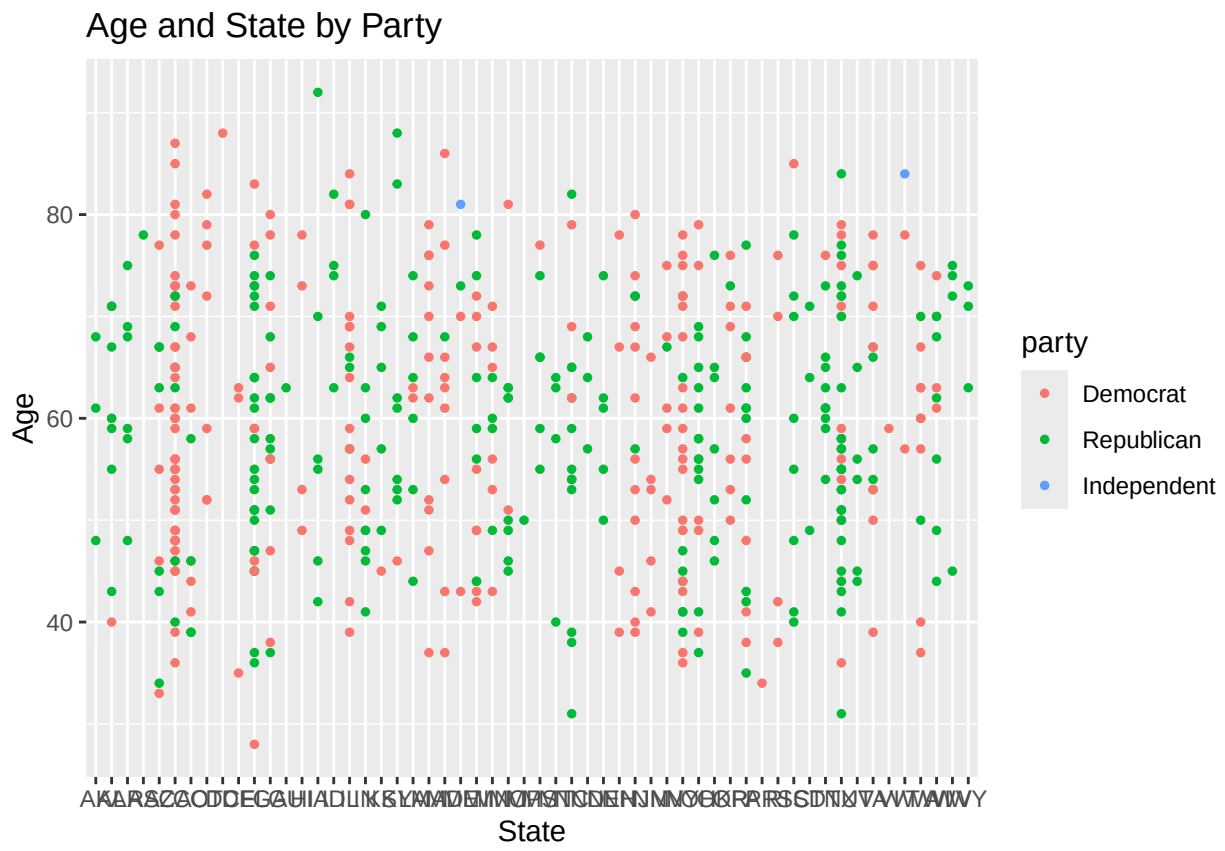
Testing

We can begin our investigation with some exploratory data analysis, using scatterplots to compare age with every other predictor (controlling for political party) to see if there any relationships which we can capture with regression.

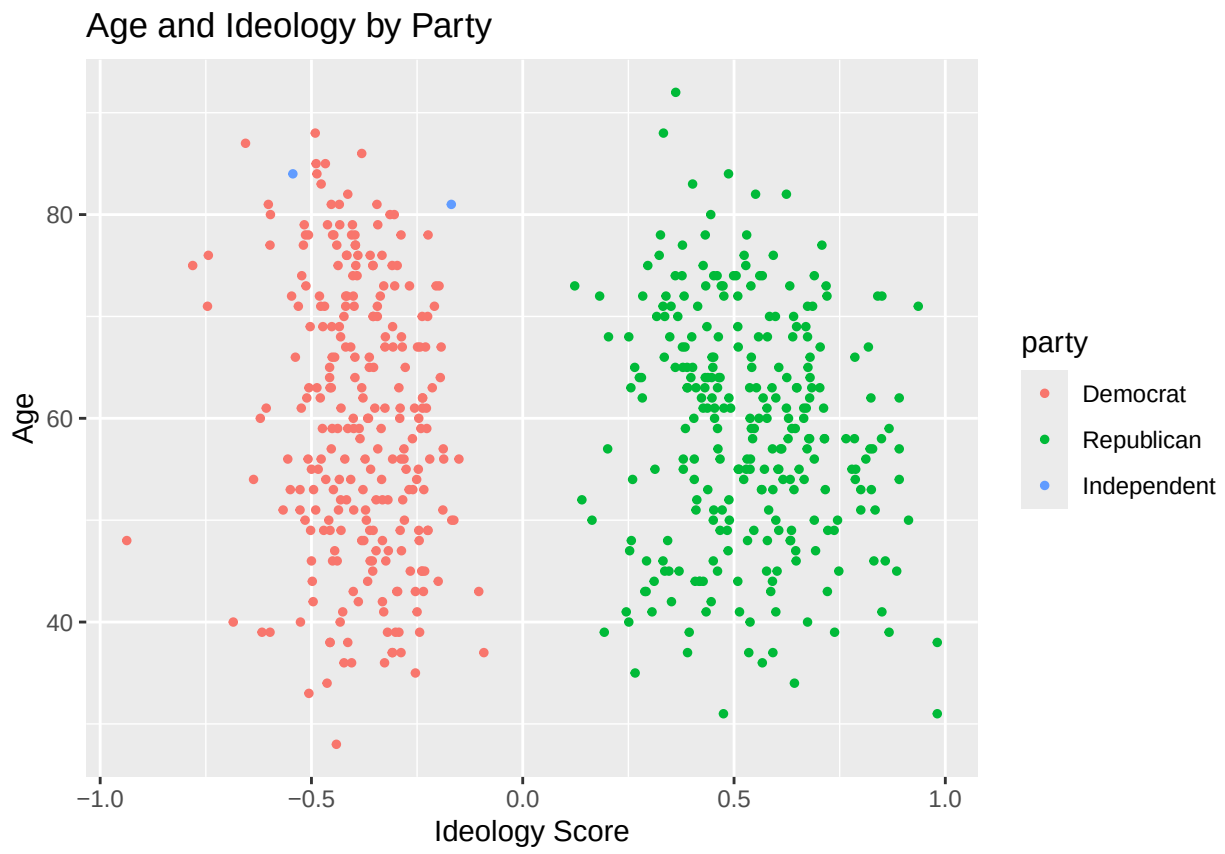
```
# Compare with chamber
ggplot(data, aes(x = chamber, y = age, color = party)) +
  geom_point(size = 1) +
  labs(title = "Age and Chamber by Party",
       x = "Chamber",
       y = "Age",
       color = "Party")
```



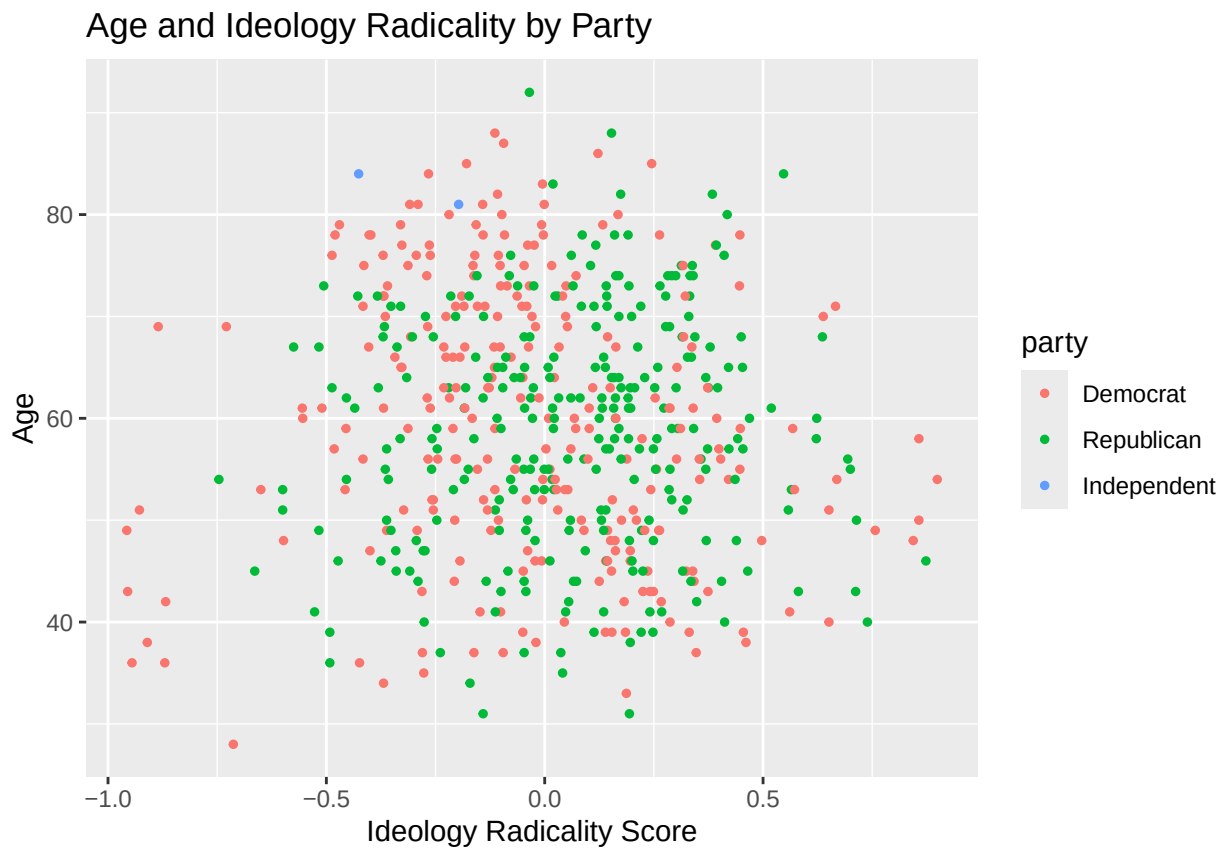

```
# Compare with State
ggplot(data, aes(x = state_abbrev, y = age, color=party)) +
  geom_point(size = 1) +
  labs(title = "Age and State by Party",
       x = "State",
       y = "Age")
```



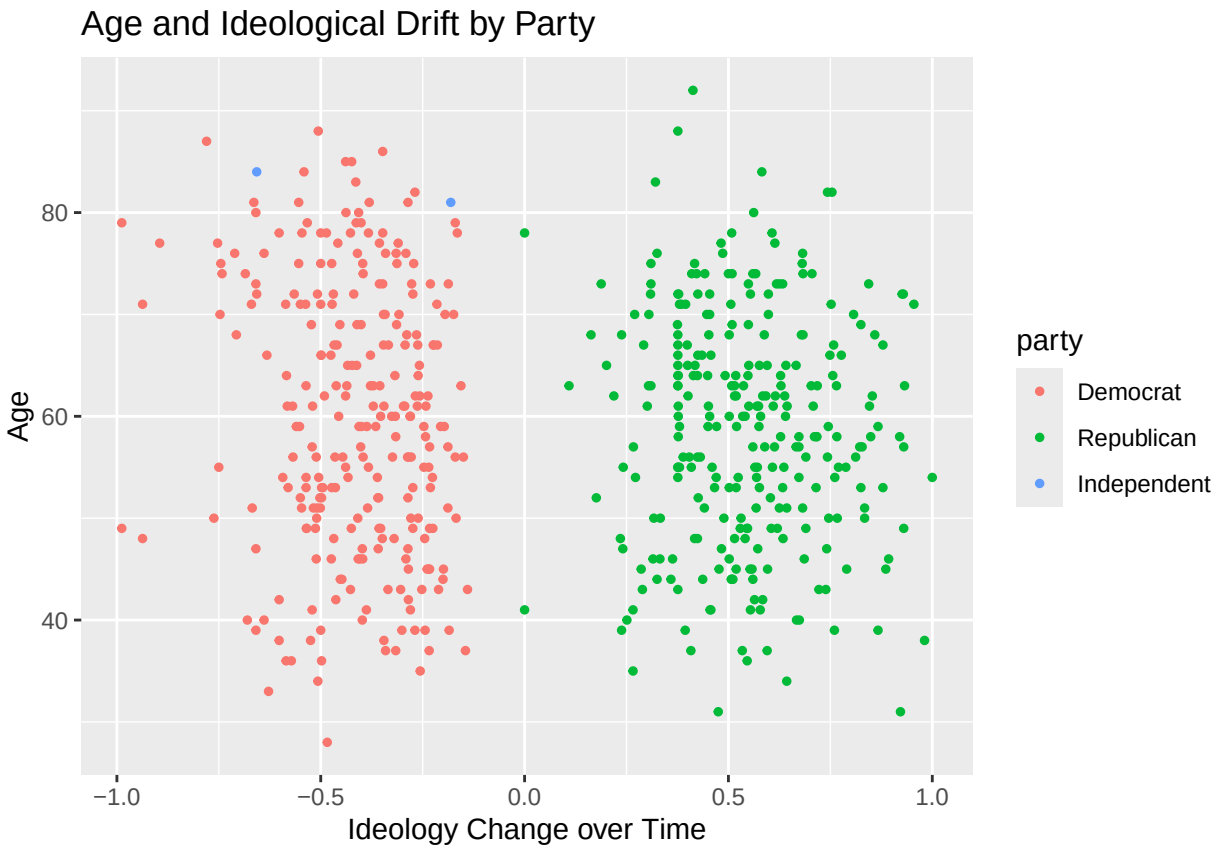

```
# Compare with Ideological Score
ggplot(data, aes(x = nominate_dim1, y = age, color = party)) +
  geom_point(size = 1) +
  labs(title = "Age and Ideology by Party",
       x = "Ideology Score",
       y = "Age")
```




```
# Compare with Ideological Radicality Score
ggplot(data, aes(x = nominate_dim2, y = age,color = party)) +
  geom_point(size = 1) +
  labs(title = "Age and Ideology Radicality by Party",
       x = "Ideology Radicality Score",
       y = "Age")
```




```
# Compare with Ideological Score over Time
ggplot(data, aes(x = nokken_poole_dim1, y = age, color = party)) +
  geom_point(size = 1) +
  labs(title = "Age and Ideological Drift by Party",
       x = "Ideology Change over Time",
       y = "Age")
```



As the EDA shows, controlling for party affiliation, there is no visible correlation between congress member age and ideological positioning in the current Congress.

This is further supported if we were to try linear regression, as shown below. We choose a range of important predictors like chamber, state, and ideological scores, and run a regresion on them.

```
data <- data %>%
  select(
    age,                          # outcome variable
    chamber,                      # House or Senate
    state_abbrev,                 # state abbreviation
    nominate_dim1,                # ideology dimension 1
    nominate_dim2,                # ideology dimension 2
    nokken_poole_dim1,           # alternative ideology measure 1
    nokken_poole_dim2            # alternative ideology measure 2
  ) %>%
  na.omit() %>%  # Remove any remaining missing values
  mutate(
    chamber = as.factor(chamber),
    state_abbrev = as.factor(state_abbrev)
  )

set.seed(380)

train_index <- createDataPartition(data$age, p = 0.8, list = FALSE)
train_data <- data[train_index, ]
test_data <- data[-train_index, ]
```

```r
# Linear regression
linear_model <- lm(age ~ ., data = train_data)

summary(linear_model)
```

```
##
## Call:
## lm(formula = age ~ ., data = train_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.028  -8.623   0.000   8.251  29.624
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        56.4183     7.1172   7.927 2.56e-14 ***
## chamberSenate       5.7329     1.6432   3.489 0.000543 ***
## state_abbrevAL      3.7651     8.2618   0.456 0.648849
## state_abbrevAR      4.8575     9.9018   0.491 0.624021
## state_abbrevAS     28.8405    14.1524   2.038 0.042262 *
## state_abbrevAZ     -3.1205     8.0243  -0.389 0.697581
## state_abbrevCA      2.4224     7.3818   0.328 0.742975
## state_abbrevCO     -4.6186     8.1230  -0.569 0.569979
## state_abbrevCT     11.0223     8.7071   1.266 0.206328
## state_abbrevDC     31.0151    14.1016   2.199 0.028456 *
## state_abbrevDE     -7.2799     9.9677  -0.730 0.465631
## state_abbrevFL     -0.6158     7.4945  -0.082 0.934559
## state_abbrevGA      0.2294     7.8609   0.029 0.976737
## state_abbrevHI      2.7770    11.1681   0.249 0.803761
## state_abbrevIA     -0.8927     9.2786  -0.096 0.923403
## state_abbrevID     14.3573     9.2974   1.544 0.123373
## state_abbrevIL      3.6412     7.7441   0.470 0.638492
## state_abbrevIN     -1.7041     8.1255  -0.210 0.833994
## state_abbrevKS      0.9000     9.2964   0.097 0.922928
## state_abbrevKY      7.0441     8.3713   0.841 0.400629
## state_abbrevLA      2.9940     8.5899   0.349 0.727619
## state_abbrevMA      4.3901     8.0854   0.543 0.587473
## state_abbrevMD     -2.3337     8.7177  -0.268 0.789075
## state_abbrevME      4.9095     9.9635   0.493 0.622480
## state_abbrevMI     -0.9047     7.8114  -0.116 0.907857
## state_abbrevMN      0.6747     8.2715   0.082 0.935030
## state_abbrevMO     -0.1030     8.3788  -0.012 0.990198
## state_abbrevMP    -19.4637    14.5163  -1.341 0.180789
## state_abbrevMS     11.8543     8.9004   1.332 0.183700
## state_abbrevMT      0.9884    11.1196   0.089 0.929218
## state_abbrevNC      2.0013     7.7707   0.258 0.796902
## state_abbrevND      2.4272    14.0169   0.173 0.862618
## state_abbrevNE      5.2279     8.9014   0.587 0.557343
## state_abbrevNH     -2.4982     9.3298  -0.268 0.789025
## state_abbrevNJ      0.1763     7.8926   0.022 0.982196
## state_abbrevNM    -10.6961    11.1400  -0.960 0.337597
## state_abbrevNV      4.1227     8.6574   0.476 0.634201
## state_abbrevNY     -1.5774     7.6436  -0.206 0.836612
```

```
## state_abbrevOH     -1.3505     7.7902  -0.173 0.862462
## state_abbrevOK      1.7166     8.3816   0.205 0.837836
## state_abbrevOR      4.9346     8.4701   0.583 0.560514
## state_abbrevPA     -2.6211     7.6412  -0.343 0.731778
## state_abbrevPR    -25.5026    14.1156  -1.807 0.071606 .
## state_abbrevRI     -7.1006    10.0364  -0.707 0.479703
## state_abbrevSC      5.7117     8.5826   0.665 0.506140
## state_abbrevSD      0.9563     9.8995   0.097 0.923095
## state_abbrevTN      5.9980     8.2472   0.727 0.467506
## state_abbrevTX      0.7549     7.3807   0.102 0.918591
## state_abbrevUT     -0.5727     8.9069  -0.064 0.948763
## state_abbrevVA      4.2526     7.8920   0.539 0.590308
## state_abbrevVI     10.2150    14.2873   0.715 0.475069
## state_abbrevVT      7.1795    11.1985   0.641 0.521840
## state_abbrevWA      6.4597     8.6725   0.745 0.456825
## state_abbrevWI      6.3786     8.2619   0.772 0.440564
## state_abbrevWV     15.2592    11.0838   1.377 0.169415
## state_abbrevWY      7.5389    11.1685   0.675 0.500083
## nominate_dim1     -15.9976     6.6310  -2.413 0.016318 *
## nominate_dim2     -10.6996     3.2871  -3.255 0.001236 **
## nokken_poole_dim1  13.3113     6.1585   2.161 0.031291 *
## nokken_poole_dim2   9.2798     2.8832   3.219 0.001400 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.09 on 377 degrees of freedom
## Multiple R-squared:  0.1811, Adjusted R-squared:  0.05292
## F-statistic: 1.413 on 59 and 377 DF,  p-value: 0.03108
```

The multiple R-squared value is a measly 0.1811, and the adjusted R-squared is an even lower 0.05292. This communicates that a very small amount of the variance in the data is covered by our predictor variables, making them unfit to explain age.

Interpretation

It is well known that American Congressmembers skew on the old side. The median age is 59, and plenty of congressmembers are far older than this. Our analysis shows that this age is not related to location or ideology, party or chamber, suggesting that this high age is inbuilt into the institution itself.

Further studies would be needed to examine the aforementioned institution, probably through a study of national/state laws and party election rules. These, however, are beyond the scope of our project.