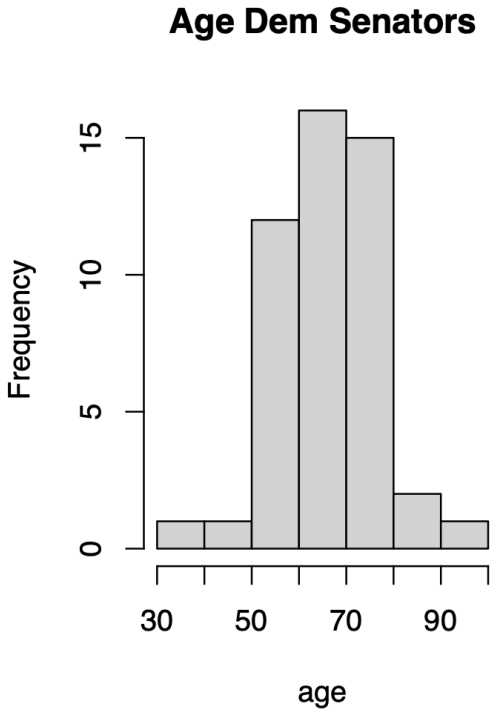
#Key Assumptions of Wilcoxon rank sum test(Hypothesis of Comparisons Version)

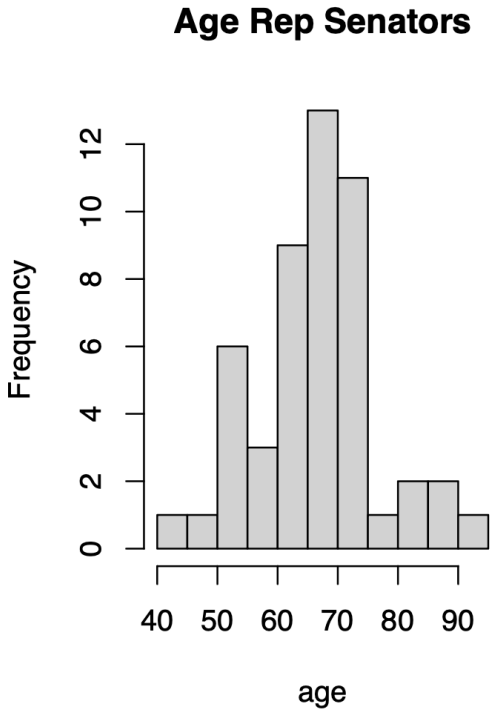
1. (At least) Ordinal scale.
2. IID data. Each of Xi is drawn from the same distribution, each Yi is drawn from the same distribution, and all Xi and Yi are mutually independent.

##Assumption Analysis

- Assumption 1: The two population samples X and Y are ages of Democratic and Republican senators, respectively. Age is a metric variable, so the condition of at least ordinal scale is met.
- IID assumption, part a: Independence of Xi and Yi. This assumption is questionable. The minimum voting age to run for senate is 30 years, which suggests some dependency between X and Y. Similarly, the population of voters may be similar among states. Among red and blue states, states may also have similar political tendencies, leading to similar age characteristics for Democratic and Republican senators. All these factors suggest a dependency between X and Y. Thus, the assumption of independence between X and Y is questionable.
- IID assumption, part b: Identical distribution of X and Y. Histograms of the two populations look different, suggesting different distributions for X and Y. X ranges from 37 years to 91 years, and Y ranges from 44 years to 91 years, indicating different ranges for X and Y. The means are similar, but the distribution of X seems to be centered around its mean, whereas Y's distribution is more dispersed. Thus, X and Y are not identically distributed.

#Conclusion: Because the IID assumption is violated, the Wilcoxon rank sum test does not apply here.






```
## # A tibble: 3 x 7
##   party       count mean_age median_age min_age max_age sd_age
##   <chr>       <int>    <dbl>      <dbl>   <dbl>   <dbl>  <dbl>
## 1 Democrat       48     66.3       66.5      37      91  10.7
## 2 Independent     2     81.5       81.5      80      83   2.12
## 3 Republican     50     67.0       68.5      44      91  10.1
```