**Are age and education independent of voting party preference?**

The topic I am exploring in my IA is between age and voting party preference and education and voting party preference. This topic was chosen because I was interested in exploring the relationship between age, education, and voting. This is interesting because of the recent election results. The data for this assessment was collected by CNN. The questions I will ask are:

 Is age independent of voting party preference?

Is education independent of voting party preference?

The type of mathematics I will be using is probability and chi-squared testing. I expect to find if my hypothesis is accepted or rejected.

The process I will be taking is:

* Collect data from CNN
* Organize the data into Google Docs charts create contingency table
* Change percentages into whole numbers
* Organize the data into frequencies to begin a chi-squared test
* Create probability problems
* Analyze the results of the chi-squared test and probability problems
* Reflect on my IA
* Create a bibliography list

**24558** respondents

|  |  |  |  |
| --- | --- | --- | --- |
| **Age** | **Clinton** | **Trump** |  Others/no answer |
| 18-2410% | 56% | 34% | 10% |
| 25-299% | 54% | 38% | 8% |
| 30-3917% | 51% | 39% | 10% |
| 40-4919% | 46% | 49% | 5% |
| 50-6430% | 44% | 52% | 4% |
| 65 and older16% | 45% | 52% | 3% |

|  |  |  |  |
| --- | --- | --- | --- |
| **Education** | **Clinton** | **Trump** | Other/no answer |
| high school or less18% | 46% | 51% | 3% |
| some college32% | 43% | 51% | 6% |
| college graduate32% | 49% | 44% | 7% |
| postgraduate18% | 58% | 37% | 5% |

**Mathematical Process:**

I will be doing a frequency chart as well as chi squared and probability.

**Frequency:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Age** | Clinton | Trump | Other | Frequency | Relative Frequency | Cumulative Frequency |
| 18-25 (2,447) | 1,370 | 832 | 245 | 2,447 | 2,447/24,794= .1 | 2,447 |
| 25-29 (2,210) | 1,193 | 840 | 177 | 2,210 | 2,210/24,794=.09 | 4,657 |
| 30-39 (4,175) | 2,129 | 1,628 | 418 | 4,175 | 4,175/24,794=.17 | 8,832 |
| 40-49 (4,665) | 2,146 | 2,286 | 233 | 4,665 | 4,665/24,794=.19 | 13,497 |
| 50-64 (7,367) | 3,241 | 3,831 | 295 | 7,367 | 7,367/24,794=.3 | 20,864 |
| 65 and older (3,929) | 1,768 | 2,043 | 118 | 3,929 | 3,929/24,794=.16 | 24,793 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Education** | Clinton | Trump | Other | Frequency | Relative Frequency | Cumulative Frequency |
| High school or less (4,420) | 2,033 | 2,254 | 133 | 4,420 | 4,420//24,557= .18 | 4,420 |
| Some college (7,859) | 3,379 | 4,008 | 472 | 7,859 | 7,859/24,557= .32 | 12,279 |
| College graduate (7,858) | 3,850 | 3,458 | 550 | 7,858 | 7,858/24,557= .32 | 20,137 |
| Postgraduate (4,420) | 2,564 | 1,635 | 221 | 4,420 | 4,420//24,557= .18 | 24,557 |

**The Chi-Squared:**

My null hypotheses ($H\_{0}$) are age is independent of voting party preference and, education is independent of voting party preference. My alternative hypotheses ($H\_{1}$) are age is not independent of voting party preference and education is not independent of voting party preference.

Observed

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Age** | Clinton | Trump | Other | Total |
| 18-25  | 1,370 | 832 | 245 | 2,447 |
| 25-29  | 1,193 | 840 | 177 | 2,210 |
| 30-39  | 2,129 | 1,628 | 418 | 4,175 |
| 40-49  | 2,146 | 2,286 | 233 | 4,665 |
| 50-64  | 3,241 | 3,831 | 295 | 7,367 |
| 65 and older  | 1,768 | 2,043 | 118 | 3,929 |
| Total | 11,847 | 11,460 | 1,486 | 24,794 |

Observed

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Education** | Clinton | Trump | Other | Total |
| High school or less  | 2,033 | 2,254 | 133 | 4,420 |
| Some college  | 3,379 | 4,008 | 472 | 7,859 |
| College graduate  | 3,850 | 3,458 | 550 | 7,858 |
| Postgraduate ( | 2,564 | 1,635 | 221 | 4,420 |
| Total | 11,826 | 11,355 | 1,376 | 24,557 |

Calculate the Expected Value

|  |  |  |  |
| --- | --- | --- | --- |
| **Age** | Clinton | Trump | Other |
| 18-25  | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 |
| 25-29  | $\frac{(2,210)(11,847)}{24,793}$=1055.89 | $\frac{(2,210)(11,847)}{24,793}$=1055.89 | $\frac{(2,210)(11,847)}{24,793}$=1055.89 |
| 30-39  | $\frac{(4,175)(11,847)}{24,793}$=1994.97 | $\frac{(4,175)(11,847)}{24,793}$=1994.97 | $\frac{(4,175)(11,847)}{24,793}$=1994.97 |
| 40-49  | $\frac{(4,665)(11,847)}{24,793}$=2229.11 | $\frac{(4,665)(11,847)}{24,793}$=2229.11 | $\frac{(4,665)(11,847)}{24,793}$=2229.11 |
| 50-64  | $\frac{(7,367)(11,847)}{24,793}$=3520.22 | $\frac{(7,367)(11,847)}{24,793}$=3520.22 | $\frac{(7,367)(11,847)}{24,793}$=3520.22 |
| 65 and older  | $\frac{(3,929)(11,847)}{24,793}$=1877.42 | $\frac{(3,929)(11,847)}{24,793}$=1877.42 | $\frac{(3,929)(11,847)}{24,793}$=1877.42 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Education** | Clinton | Trump | Other |
| High school or less  | $\frac{(4,420)(11,826)}{24,557}$=2128.55 | $\frac{(4,420)(11,355)}{24,557}=$2043.78 | $\frac{(4,420)(1,376)}{24,557}$= $247.66$ |
| Some college  | $\frac{(7,859)(11,355)}{24,557}$=3633.95 | $\frac{(7,859)(11,355)}{24,557}$=3633.95 | $\frac{(7,859)(1,376)}{24,557}$=440.36 |
| College graduate  | $\frac{(7,858)(11,826)}{24,557}$=3784.20 | $\frac{(7,858)(11,355)}{24,557}$=3633.48 | $\frac{(7,858)(1,376)}{24,557}$=440.31 |
| Postgraduate  | $\frac{(4,420)(11,826)}{24,557}$=2128.55 | $\frac{(4,420)(11,355)}{24,557}=$2043.78 | $\frac{(4,420)(1,376)}{24,557}$= $247.66$ |

**Education**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 𝑓o | 𝑓e | (𝑓o-𝑓e) | (𝑓o-𝑓e$)^{}^{2}$ | $$\frac{(f\_{o}-fe\_{})^{}^{2}}{fe}$$ |
| 2,033 | $\frac{(4,420)(11,826)}{24,557}$=2128.55 | -95.55 | 9129.80 | 4.28 |
| 3,379 | $\frac{(7,859)(11,355)}{24,557}$=3633.95 | -254.95 | 64999.50 | 17.9 |
| 3,850 | $\frac{(7,858)(1,376)}{24,557}$=440.31 | 3409.7 | 11626054.09 | 26404.24 |
| 2,564 | $\frac{(4,420)(1,376)}{24,557}$=247.66 | 2316.34 | 5365430.99 | 21664.5 |
| 2,254 | $\frac{(4,420)(11,355)}{24,557}=$2043.78 | 210.22 | 44192.44 | 21.6 |
| 4,008 | $\frac{(7,859)(11,355)}{24,557}$=3633.95 | 374.05 | 139913.40 | 38.5 |
| 3,458 | $\frac{(7,858)(11,355)}{24,557}$=3633.48 | -175.48 | 30793.23 | 8.47 |
| 1,635 | $\frac{(4,420)(11,355)}{24,557}=$2043.78 | -408.78 | 167101.08 | 81.8 |
| 133 | $\frac{(4,420)(1,376)}{24,557}$= $247.66$ | -114.66 | 13146.91 | 53.1 |
| 472 | $\frac{(7,859)(1,376)}{24,557}$=440.36 | 31.64 | 1001.08 | 2.27 |
| 550 | $\frac{(7,858)(1,376)}{24,557}$=440.31 | 109.69 | 12031.89 | 27.3 |
| 221 | $\frac{(4,420)(1,376)}{24,557}$= $247.66$ | -26.66 | 710.75 | 2.86 |

$x^{2}calc =^{}\sum\_{}^{}\frac{(f\_{o}-fe\_{})^{}^{2}}{fe}$= 48,327

To find our critical value, we need to find the degrees of freedom:

*Degrees of freedom= (number of rows- 1) (number of columns -1)*= (4-1)(3-1) = 6

The critical value at df 6 at 5% significance level is 12.59.

48,327 > 12.59

We reject $H\_{0}$.

**Age**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 𝑓o | 𝑓e | |𝑓o-𝑓e| | (𝑓o-𝑓e$)^{}^{2}$ | $$\frac{(f\_{o}-fe\_{})^{}^{2}}{fe}$$ |
| 1,370 | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 | 200.73 | 40292.53 | 34.5 |
| 1,193 | $\frac{(2,210)(11,847)}{24,793}$=1055.89 | 137.11 | 18799.15 | 17.8 |
| 2,129 | $\frac{(4,175)(11,847)}{24,793}$=1994.97 | 134.03 | 17964.04 | 9 |
| 2,146 | $\frac{(4,665)(11,847)}{24,793}$=2229.11 | 83.11 | 6907.27 | 3.1 |
| 3,241 | $\frac{(7,367)(11,847)}{24,793}$=3520.22 | 279.22 | 77963.8 | 22.15 |
| 1,768 | $\frac{(3,929)(11,847)}{24,793}$=1877.42 | 109.42 | 11972.73 | 6.38 |
| 832 | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 | 337.27 | 113569 | 97.13 |
| 840 | $\frac{(2,210)(11,847)}{24,793}$=1055.89 | 215.89 | 46608.49 | 44.1 |
| 1,628 | $\frac{(4,175)(11,847)}{24,793}$=1994.97 | 366.97 | 134666.98 | 67.5 |
| 2,286 | $\frac{(4,665)(11,847)}{24,793}$=2229.11 | 56.89 | 3236.47 | 1.45 |
| 3,831 | $\frac{(7,367)(11,847)}{24,793}$=3520.22 | 310.78 | 96584.20 | 27.44 |
| 2,043 | $\frac{(3,929)(11,847)}{24,793}$=1877.42 | 165.58 | 27416.73 | 14.6 |
| 245 | $\frac{(2,447)(11,847)}{24,793}$= 1169.27 | 924.27 | 854275.03 | 730.6 |
| 177 | $\frac{(2,210)(11,847)}{24,793}$=1055.89 | 878.89 | 772447.63 | 731.6 |
| 418 | $\frac{(4,175)(11,847)}{24,793}$=1994.97 | 1576.97 | 2486834.38 | 1246.6 |
| 233 | $\frac{(4,665)(11,847)}{24,793}$=2229.11 | 1996.11 | 3984455.13 | 1787.5 |
| 295 | $\frac{(7,367)(11,847)}{24,793}$=3520.22 | 3225.22 | 10402044.05 | 2954.94 |
| 118 | $\frac{(3,929)(11,847)}{24,793}$=1877.42 | 1759.42 | 3095558.73 | 164.9 |

$x^{2}calc =^{}\sum\_{}^{}\frac{(f\_{o}-fe\_{})^{}^{2}}{fe}$= 7,961

Degrees of freedom: 10

Critical value at 5%: 18.31

7,961 > 18.31

We reject $H\_{0}$.

**Probability:**

“T” stands for someone who voted for Trump.

“C” stands for someone who voted for Clinton.

“O” stand for someone who voted other/no answer.

“H” stands for high school or less.

“G” stands for college graduate.

“D” stands for postgraduate.

Age:

P (18-25 $∪$ T) = $\frac{2,447 + 11,460 - 832}{24,794}$= 0.53 = 53%

P(40-49 $∩$ C) = $\frac{2,146}{24,794}$= 0.09 = 9%

P (50-64 | O) = $\frac{3,831}{11,460}$= 0.33= 33%

Education:

P (G $∪$ C) = $\frac{4,420 +11,826 - 2,033}{24,557}$=0.58 = 58%

P(D $∩$ T) = $\frac{1,635}{24.,557}$= 0.07 = 7%

P (H | O) = $\frac{133}{1,376}$= 0.10 = %10

**Interpretation:** In my chi squared test, I tested to find if age was independent of voting preference and if education is independent of voting preference. The test showed us that at a 5% significance level, neither age nor education are independent on voting party preference. This proves that there is a correlation between people’s age and education when it comes to voting.

In my probability portion of this assessment, it shows that people in the age range of 18 to 25, or a Trump supporter, 53% of that age range did not vote for Trump. Also, The people between the ages of 40 to 49 and voting for Clinton was a 9% probability. Given that the person was a Trump supporter between the ages of 50 to 64 the percentage was 33%. For education, the people who reached high school education or less given that they voted other/no answer was 10%. The probability that someone with a college graduate education or a Clinton supporter had a probability of 58%. Someone who reached postgraduate education and voted for Trump was a 7% chance. Based on my testing, age and education are not independent when it comes to voting.

**Validity:**

CNN is considered secondary data so, therefore, there are a few factors that have influenced my calculations/testing to become less accurate. This is because CNN was rounding the percentage number of voters. I was rounding to the whole number of voters. That is why my numbers were slightly larger than the data I retrieved from the source. Also, some voters decided not to answer certain questions on their ballot because it was optional therefore, there was a different number of correspondents in the education portion of the data. To ensure the validity of my assessment, I have double checked the entire chi-squared portion.

**Reflections:** I believed there were several factors that determined how people voted. Since I explored this topic, I learned that age and education are not independent of voting party preference. I was unaware that age and education was apart of it as well. The most common ones were race, gender and religion. These things should not factor in the decision of picking the next leader unless it is a danger to the wellbeing of society. Unfortunately, these results are disappointing and affecting the lives of many. People should vote based on what best for the sake of the country.

.Bibliography:<http://www.cnn.com/election/results/exit-polls>