5. [.2, .23, .17, .05]

mean = .163

median = .185

$v = \frac{.001 + .004 + 0 + .013}{3} = .006$

sd = .077

se = .039

CT, Median maybe, due to outlier of 5%.. It is within SE range. So I'd go with 18.5%..

Assignments

Now that we have introduced some tools for describing populations, let's try them out. First do these drills by hand, then use the Python code we've provided in the previous assignments to check your work. Keep track of your work in a Google document or markdown file that you can submit below and share with your mentor.

1. Greg was 14, Marcia was 12, Peter was 11, Jan was 10, Bobby was 8, and Cindy was 6 when they started playing the Brady kids on The Brady Bunch. Cousin Oliver was 8 years old when he joined the show. What are the mean, median, and mode of the kids' ages when they first appeared on the show? What are the variance, standard deviation, and standard error?
2. Using these estimates, if you had to choose only one estimate of central tendency and one estimate of variance to describe the data, which would you pick and why?
3. Next, Cindy has a birthday. Update your estimates- what changed, and what didn't?
4. Nobody likes Cousin Oliver. Maybe the network should have used an even younger actor. Replace Cousin Oliver with 1-year-old Jessica, then recalculate again. Does this change your choice of central tendency or variance estimation methods?
5. On the 50th anniversary of The Brady Bunch, four different magazines asked their readers whether they were fans of the show. The answers were: TV Guide 20% fans Entertainment Weekly 23% fans Pop Culture Today 17% fans SciPhi Phanatic 5% fans

   Based on these numbers, what percentage of adult Americans would you estimate were Brady Bunch fans on the 50th anniversary of the show?

Discuss your answer to each of these questions, along with your code, with your mentor.

When you've given it a try, you can find a solution here (https://github.com/Thinkful-Ed/data-201-resources/blob/master/solutions/Prep%20course/3.1.4.ipynb).

1. [14, 12, 11, 10, 8, 6, 8] = 69

Mean: 69/7 = 9.857

Median → [6, 8, 8, 10, 11, 12, 14]

10

Mode: 8

Var: $\frac{(17.164 + 4.592 + 1.306 + .020 + 3.448 + 14.876 + 3.448)}{7 - 1}$ → Sample Variance

$= 44.854 / 6 = 7.476$

Std Dev: $\sqrt{7.476} = 2.734$

Std Err: $2.734 / \sqrt{7} = 1.033$

2. In this first instance, I would probably go with the mean and standard deviation as there are no crazy outliers and the stadard error shows our mean can't be too far off. Median could work better as measure of central tendency, simply because we are dealing with age and ~~whole~~ whole numbers, but mean is likely more useful still.

3. Mean = 10
Median = 10
Mode = 6

$v = \frac{16+4+1+0+4+9+4}{6}$ [14, 12, 11, 10, 8, 7, 6] Sum: 70

v = 6.333
sd = 2.517
se = 0.951

Median & mode stayed exactly the same. Mean, variance, standard deviation & standard error remained very close to the original values with very slight changes.

4. [14, 12, 11, 10, 8, 7, 1]
mean = 9.286
median = 10
No Mode

$v = \frac{22.222+7.366+2.938+.510+1.654+5.226+68.658}{6}$

v = 18.096
sd = 4.254
se = 1.604

Median may become choice for central tendency, but our confidence in the mean is still between 7.678 and 10.894. The one year old Jessica is definitely an outlier as our variance is much higher than the other sets.