Poli 101
HW #2

HW#2 – Indexing and Inference

Dataset: Anes 2016 Pilot – Weighted

Hypothosis:

The Concept that I am measuring is the optimism of economic advancement. I've indexed five indicators surveying the levels of optimism people have about the ability for individuals to advance economically as well as general feelings about the future of the economy. The strong alpha suggests these ideas are essentially the same concept.

I will be measuring this concept against three independent variables: Education, Political Ideology, and Race. I believe that any of these variables could have a relationship to general economic optimism.

Independent Variable 1:

Indicator: Education

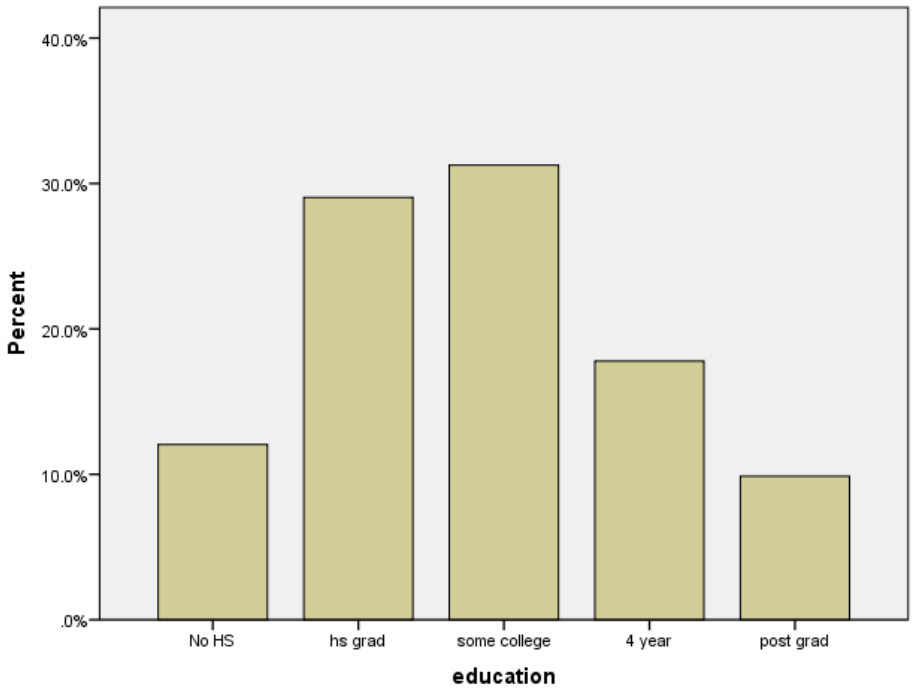
Indicator Type: Ordinal

Missing: Skipped, Not Asked

Recode: Because '2 year' was a small group and it stood out from the rest of the data, I recoded to combine those who answered with '2 Year' with 'Some College.'

Hypothosis: Optimism regarding economic opportunity will increase with education level.

The data has a normal distribution with a slight skew toward less education. The mode and median are both 3 (some college) and the mean is 2.8439, indicating the skewness of .222.



Independent Variable 2:

Indicator: Ideology

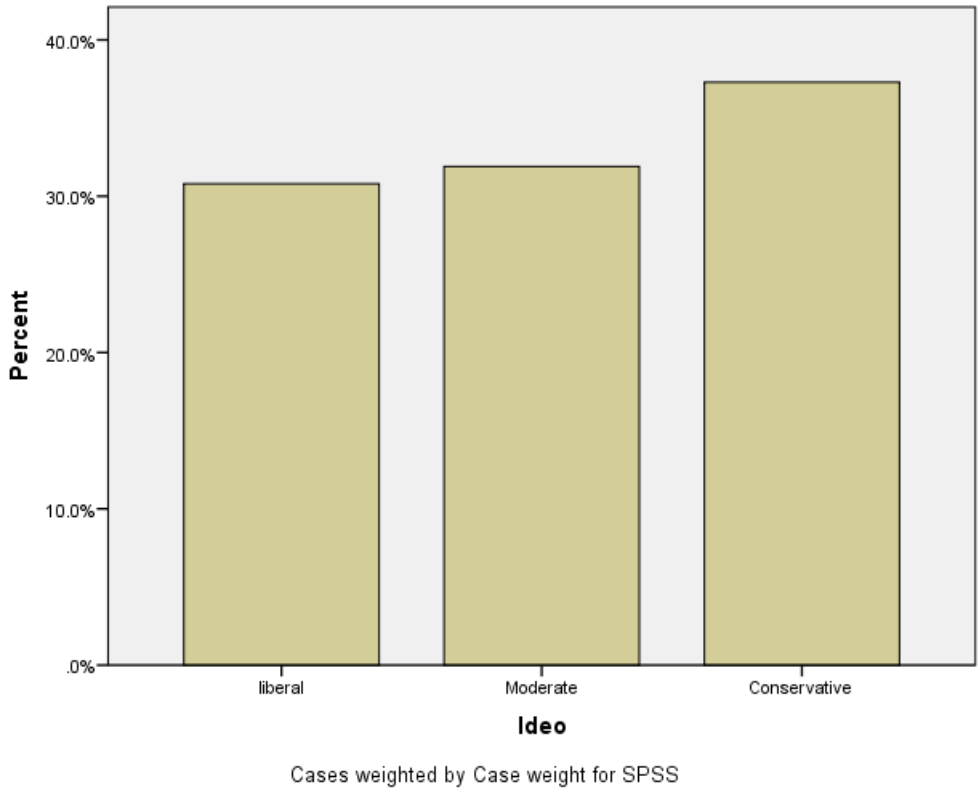
Indicator Type: Ordinal

Missing: Skipped, Not Asked

Recode: Since the question was asked in terms of liberal and conservative, I combined all liberal values together from 'Slightly Liberal' to 'very liberal.' I did the same for conservative values. I left 'neither liberal nor conservative' alone.

Hypothesis: People who identify as liberal will be more optimistic about economic opportunity.

The data for Ideology has a negative skew. While the mode is 3 (Conservative) there seems to be a mostly even representation of liberal, moderate and conservative responses. The medial remains 2 (moderate).



Independent variable 3:

Indicator: Race

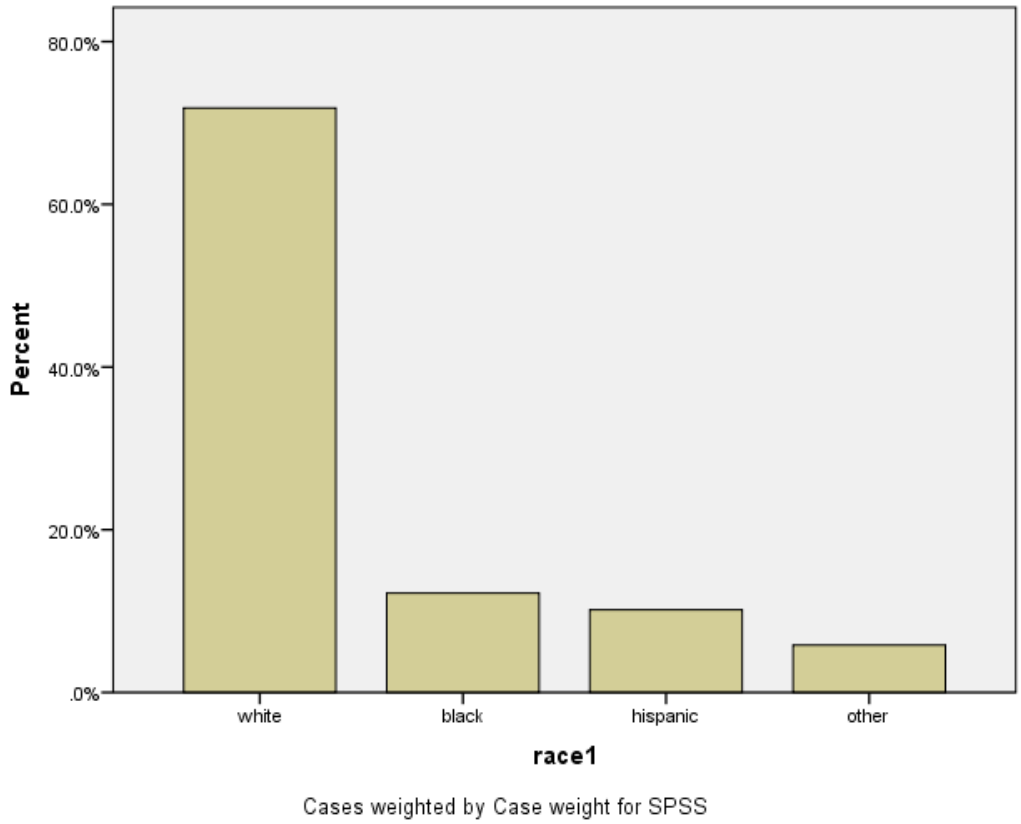
Indicator Type: Nominal

Missing: Skipped, Not Asked

Recode: The survey asked for many different races, though there was as significantly higher portion of white responses. I felt it was more interesting to look at the difference between 'white' and 'non-white,' so I hoped to simplify the data. However, there seemed to be differences between non-whites, so I left categories for 'Black' and 'Hispanic' in addition to 'other.' The 'other' category contains all other responses to the question, but not the missing categories.

Hypothesis: People who identify as white will feel more optimistic about economic opportunity.

The data is skewed heavily toward 'White.' Even after recoding to combine several categories into 'other.' White still dominates the representation.



Building the index for the dependent variable:

I selected five questions that are indicators of two ideas that essentially measure the same thing. 'Getahead' asks how much opportunity for the average person to get ahead today. 'Ladder' asks if it is easier/harder to move up the economic ladder today than in the past. 'Finwell' asks whether people's ability for someone to improve their financial well-being is now better/worse. These ideas relate directly to individuals—sort of the 'Bootstrap' mentality.

But it seems that people feel their own ability is largely tied to greater economic prosperity as well. 'Econnow' asks how people feel about the economy now, compared to the previous year. And 'Econ12mo' asks how people feel about the economy 12 months from now.

1) 'Getahead' - Opportunity in America today for the average person to get ahead
   (recode:Opport1)
   Missing: skipped, not asked.

Recode: `recode getahead (1=0) (2=.25) (3= .5) (4= .75) (5=1) into Opport1.`
`VALUE LABELS Opport1 0 'None' .25 'a little' .5 'a moderate amount' .75 'a lot' 1 'a great deal'.`

2) 'Ladder' - Move up the income ladder - Easier/harder (Opport2)

Missing: skipped, not asked

Recode: `recode ladder (1=1) (2= .833) (3= .667) (4= .5) (5=.333) (6= .167) (7= 0) into Opport2.`
`value labels Opport2 1 'a great deal easier' .833 'moderately easier' .667 'a little easier' .5 'neither easier or harder' .333 'a little harder' .167 'moderately harder' 0 'a great deal harder'.`

3) 'Finwell' - People's ability to improve their financial well-being is now better/worse (Opport3)

Missing: skipped, not asked

Recode: `recode finwell (1=1) (2= .833) (3= .667) (4= .5) (5=.333) (6= .167) (7= 0) into Opport3.`
`value labels Opport3 1 'a great deal better' .833 'moderately better' .667 'a little better' .5 'same as 20 years ago' .333 'a little worse' .167 'moderately worse' 0 'a great deal worse'.`

4) 'Econnow' - Economy compared to a year ago (Opport4)

Missing: skipped, not asked

Recode: `recode econnow (1=1) (2= .75) (3= .5) (4= .25) (5= 0) into Opport4.`
`value labels Opport4 1 'much better' .75 'somewhat better' .5 'about the same' .25 'somewhat worse' 0 'much worse'.`

5) 'Econ12mo' - Economy 12 months from now—compared to now (Opport5):

Missing: Skipped, Not asked

Recode: `Recode econ12mo (1=1) (2=.75) (3=.5) (4=.25) (5=0) into opport5.`
`value labels Opport5 1 'much better' .75 'somewhat better' .5 'about the same' .25 'somewhat worse' 0 'much worse'.`

Reliability test:

With all five indicators returning an alpha score of .667, running a reliability test returned predictable results. Removing other variables seemed to not significantly alter the alpha. I kept all 5 indicators of the concept in the final index.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.667	.662	5

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Opport1	1.7871	.598	.348	.210	.646
Opport2	1.7873	.498	.520	.530	.565
Opport3	1.7944	.487	.587	.545	.532
Opport4	1.7337	.601	.300	.396	.668
Opport5	1.7216	.601	.356	.406	.642

Coding the raw index(RawIndex):

In coding the raw index, the values of the responses were scaled from 0 for less optimistic to 1 for more optimistic about the economy.

Recoding the index (opportind):

To simplify the information, I recoded the index into 3 categories at just about one third of the population into each category:

```
*Recoding the Index*.
```

```
recode RawIndex (0 thru 1.83 =1) (1.84 thru 2.58 =2) (2.59 thru 5
= 3) into OpportInd.
value labels OpportInd 1 'low’ 2 'med' 3 'high'.
fre var OpportInd
   /statistics mean median stddev skew kurosis.
```

Central measures of tendency:

Statistics

RawIndex

N	Valid	684
	Missing	0
Mean		2.2060
Median		2.2500
Mode		2.50
Std. Deviation		.89634
Variance		.803
Skewness		-.066
Kurtosis		-.015

Statistics

OpportInd

N	Valid	684
	Missing	0
Mean		2.0357
Median		2.0000
Mode		2.00
Std. Deviation		.80245
Variance		.644
Skewness		-.064
Kurtosis		-1.444

The central measures have shifted slightly. The Kurtosis is more negative, indicating that the distribution would be more flat. The skewness hasn't changed much but the mode and median have both shifted down to 2, instead of 2.5. The data is now more compact.

Crosstab for Education:

The first independent variable measures the highest level of education that the participants in the survey have achieved. By looking at the crosstab you can immediately read that there is little if any relationship between education and optimism about economic opportunity. However, Kendall's Tau does almost reach .5, which would suggest a relationship. But, even if we were feeling generous, the ChiSq is too high for it to be statistically significant. We must accept the null hypothesis that Education is not an indicator for optimism for economic self-betterment.

Still, this is very interesting, because if education is supposed to be a gateway to economic advancement, you would think that those who have achieved higher education might feel more optimistic about the possibility for financial advancement.

OpportInd * education Crosstabulation

			education					
			No HS	hs grad	some college	4 year	post grad	Total
OpportInd	Low	Count	25	64	75	25	19	208
		% within education	30.1%	32.5%	35.0%	20.5%	28.4%	30.5%
	Med	Count	32	61	70	53	27	243
		% within education	38.6%	31.0%	32.7%	43.4%	40.3%	35.6%
	High	Count	26	72	69	44	21	232
		% within education	31.3%	36.5%	32.2%	36.1%	31.3%	34.0%
Total		Count	83	197	214	122	67	683
		% within education	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Phi	.129			.186
	Cramer's V	.091			.186
Ordinal by Ordinal	Kendall's tau-c	.024	.034	.706	.480
N of Valid Cases		683			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	11.284[a]	8	.186
Likelihood Ratio	11.637	8	.168
Linear-by-Linear Association	.483	1	.487
N of Valid Cases	683		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 20.40.

Crosstab for Ideology:

The independent variable in this crosstab measures how people perceive their own political ideology. By setting up the cross tab for Political Ideology with Economic Optimism, we can begin to understand the different ways, if any, that liberals, moderates or conservatives may view economic opportunity. The crosstab does slightly indicate that the hypothesis is correct. Liberals seem to have a slight tendency toward more economic optimism, but the indicator seems to only apply to people who identify as liberal. Moderates and conservatives seem to have an even distribution across how optimistic they are about the ability for someone to advance economically.

However, according to the tau value, this relationship is not a strong relationship. In fact at .011 it is very weak. With a ChiSq at .024, we can feel very confident that the relationship is very week. According to the evidence, we can accept the null hypothesis that there can be no inference between ideology and optimism for economic opportinuty.

OpportInd * Ideo Crosstabulation

			Ideo			
			liberal	Moderate	Conservative	Total
OpportInd	Low	Count	48	73	87	208
		% within Ideo	22.9%	33.6%	34.3%	30.5%
	Med	Count	76	81	85	242
		% within Ideo	36.2%	37.3%	33.5%	35.5%
	High	Count	86	63	82	231
		% within Ideo	41.0%	29.0%	32.3%	33.9%
Total		Count	210	217	254	681
		% within Ideo	100.0%	100.0%	100.0%	100.0%

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Phi	.128			.024
	Cramer's V	.091			.024
Ordinal by Ordinal	Kendall's tau-c	-.086	.034	-2.531	.011
N of Valid Cases		681			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Crosstab for race:

Quite simply, I put up the hypothesis that minorities would be less likely to feel optimistic about the opportunity for economic advancement based on preconceived notions about systematic racism that makes it harder for minorities to advance. Contrary to the hypothesis, when crosstabbed against the general feelings for economic opportunity, much like political ideology, people who Identify as Black and Hispanic seem to feel much more optimistic about the prospect for economic advancement.

However, when put to the test, Kendall's Tau returns .000. ChiSq is also .000.

OpportInd * race1 Crosstabulation

			race1				
			white	black	hispanic	other	Total
OpportInd	Low	Count	175	13	10	11	209
		% within race1	35.6%	15.7%	14.3%	27.5%	30.6%
	Med	Count	192	20	17	14	243
		% within race1	39.1%	24.1%	24.3%	35.0%	35.5%
	High	Count	124	50	43	15	232
		% within race1	25.3%	60.2%	61.4%	37.5%	33.9%
Total		Count	491	83	70	40	684
		% within race1	100.0%	100.0%	100.0%	100.0%	100.0%

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Phi	.314			.000
	Cramer's V	.222			.000
Ordinal by Ordinal	Kendall's tau-c	.193	.028	6.877	.0000000000
N of Valid Cases		684			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Conclusion:

In addition to these independent variables, I tested several others that all returned null hypothesis. There seems to be some traction along race lines in regards to optimism with non-whites, but there is not enough of a relationship to make an inference. Since minorities are such a small part of the sample, perhaps with more data we may find a more relationship between race and optimism about advancing economically.

With five indicators supporting this concept of optimism for financial advancement, there should be some other variables that we could infer a relationship, but I was unable to find any strong relationships.

SYNTAX—Dataset: ANES 2016 Pilot:

```
* Encoding: UTF-8.
weight by weight_spss.

missing values getahead (8, 9).
fre var=getahead
  /statistics stdev skew kurtosis.

recode getahead (1=0) (2=.25) (3= .5) (4= .75) (5=1) into Opport1.
VALUE LABELS  Opport1  0 'None' .25 'a little' .5 'a moderate amount' .75 'a lot' 1 'a great deal'.
fre var=Opport1
   /statistics stdev skew kurtosis.

missing values ladder (8, 9).
fre var=ladder
   /statistics stdev skew kurtosis.

recode ladder (1=1) (2= .833) (3= .667) (4= .5) (5=.333)  (6= .167) (7= 0) into Opport2.
value labels Opport2 1 'a great deal easier' .833 'moderately easier' .667 'a little easier' .5 'neither easier
or harder' .333 'a little harder' .167 'moderately harder' 0 'a great deal harder'.
fre var=Opport2
   /statistics stdev skew kurtosis.

missing values finwell (8,9).
fre var=finwell
   /statistics stdev skew kurtosis.

recode finwell (1=1) (2= .833) (3= .667) (4= .5) (5=.333)  (6= .167) (7= 0) into Opport3.
value labels Opport3 1 'a great deal better' .833 'moderately better' .667 'a little better' .5 'same as 20
years ago' .333 'a little worse' .167 'moderately worse' 0 'a great deal worse'.
fre var opport3
```

```
   /statistics stdev skew kurtosis.

missing values econnow (8,9).
fre var=econnow
   /statistics stdev skew kurtosis.

recode econnow (1=1) (2= .75) (3= .5) (4= .25) (5= 0) into Opport4.
value labels Opport4 1 'much better' .75 'somewhat better' .5 'about the same' .25 'somewhat worse' 0
'much worse'.
fre var opport4
   /statistics stdev skew kurtosis.

missing values econ12mo (8,9).
fre var=econ12mo
   /statistics stdev skew kurtosis.

Recode econ12mo (1=1) (2=.75) (3=.5) (4=.25) (5=0) into opport5.
value labels Opport5 1 'much better' .75 'somewhat better' .5 'about the same' .25 'somewhat worse' 0
'much worse'.
fre var opport5
   /statistitcs stdev skew kurtosis.

*Conducting Reliabiility Analysis*.
reliability
  /variables= Opport1 Opport2 Opport3 Opport4 Opport5
  /scale(Opport1) all
  /summary = all.

*Constructing the Index*.
compute RawIndex = Opport1 + Opport2 + Opport3 + Opport4 + Opport5.
fre var=RawIndex
  /statistics = variance mean median mode stddev skew kurtosis.
```

```
*Recoding the Index*.
recode RawIndex (0 thru 1.83 =1) (1.84 thru 2.58 =2) (2.59 thru 5 = 3) into OpportInd.
value labels OpportInd 1 'Low' 2 'Med' 3 'High'.
fre var OpportInd
  /statistics = variance mode mean median stddev skew kurosis.

*Creating an indicator of Ideology*.
missing values lcself (8,9).
fre var=lcself
   /statistics = mean median stddev skew kurtosis.

missing values lcself (8,9).
recode lcself (1, 2, 3= 1) (4 = 2) (5,6,7 = 3) into Ideo.
value labels Ideo 1 'liberal' 2 'Moderate' 3 'Conservative'.
fre var = Ideo
   /statistics = mode  mean median stddev skew kurtosis.

GRAPH
 /BAR(SIMPLE)=PCT BY ideo.

*Crosstabulation of OpportInd by ideo*.
crosstabs tables = OpportInd by Ideo
  /cells = column count
  /statistics = phi ctau chisq.

oneway opportind by ideo
 /statistics = descriptives
  /ranges=scheffe
  /plot means.

*Creating an indicator of Education*.
```

```
missing values Educ (8,9).
fre var=educ
   /statistics = mode mean median stddev skew kurtosis.

missing values educ (8,9).
recode educ (1 = 1) (2 = 2) (3,4 = 3) (5=4) (6=5) into education.
value labels education 1 'No HS' 2 'hs grad' 3 'some college'  4 '4 year' 5 'post grad'.
fre var = Education
   /statistics = mode mean median stddev skew kurtosis.

GRAPH
 /BAR(SIMPLE)=PCT BY Education.

*Crosstabulation of OpportInd by Education*.
crosstabs tables = OpportInd by Education
  /cells = column count
  /statistics = phi ctau chisq.

oneway opportind by education
 /statistics = descriptives
  /ranges=scheffe
  /plot means.

*creating and indicator for Race*.
missing values race (98,99).
fre var = race
   /statistics mean median stddev skew kurtosis.

missing values race (98,99).
recode race (1 =1) (2=2) (3=3) (4,5,6,7,8=4)  into race1.
value labels race1 1 'white' 2 'black' 3'hispanic' 4'other'.
fre var= race1
```

```
    /statistics mode mean median stddev skew kurtosis.

crosstabs tables =opportind by race1
    /cells column count
    /statistics phi ctau chisq.

oneway opportind by Race1
  /statistics = descriptives
   /ranges=scheffe
   /plot means.


GRAPH
  /BAR(SIMPLE)=PCT BY race1.
```