

HW4 Answers

1.

Using the California and Texas HW4 syntax files, complete the **linked table** or HW4-24 Table. Enter the appropriate coefficients using 3 decimal places.

Summary measures for 3 Indexes of Attitudes toward Immigration by IVs

	California		Texas			
	ImmCal3	Chi-sq	ImmIncl3		ImmExcl3	
	MoA	p =	MoA	p =	MoA	p=
female	.089 v	.003	-.030 v	.647	.102 v	.006
ethn	.197 v	<.001	.116 v	<.001	.226 v	<.001
age	-.160 T _c	<.001	-.228 T _c	<.001	.268 T _c	<.001
educR	.016 T _c	<.001	.079 T _c	.016	-.107 T _c	.053
income	-.045 T _c	<.001	.117 T _c	.002	-.005 T _c	.419
interest	-.004 T _c	<.001	.034 T _c	<.001	.167 T _c	<.001
Dem3	.408 T _b	<.001	.331 T _b	<.001	-.435 T _b	<.001
lib3	.421 T _b	<.001	.430 T _b	<.001	-.520 T _b	<.001

California Data—June2023 PPIC survey.

Texas Data—December 2023 Univ of Texas.

California Index:

ImmCal3 asks about the benefits of immigrants, support for daca, providing a path to citizenship & health care for immigrants. Alpha = .823 recoded to 3 categories.

Texas Indices:

ImmIncl3: asks about admitting skilled immigrants, those with graduate training and offering a path to citizenship, Alpha = .756, recoded to 3 categories.

ImmExcl3 asks about tightening admission and checking on immigrants and building a wall. Alpha = .793, recoded to 3 categories.

Notes:

- as coded, female may be considered ordinal, hence coefficients =
 ImmCal3 by female = -.002 T_c
 ImmIncl3 by female = -.021 T_c
 ImmExcl3 by female = -.108 T_c
- High scoring Ethnicity depends on DV:
 Hispanics score high on ImmCal3 & ImmIncl3;
 Whites score high on ImmExcl3.
- Signs reverse on Age, Dem3 and lib3.

2.

Based on the first crosstabulation produced by the data entry syntax included in the California syntax file, how well would you say the September 2021 PPIC survey responses reflect the 2021 California recall election results?

Column 1 entries are derived from the valid percentages produced by missing values q5 (8,9).

fre var q5.

Column 2 results were obtained at:

<https://www.washingtonpost.com/elections/election-results/california/governor-recall/>

(before proceeding close but not save the small untitled data file created by this syntax).

*Data entry on Recall election.

*PPIC Survey vs Recall Election results.

data list free / Recall DataType count.

begin data.

1 1 563

1 2 4894473

2 1 937

2 2 7944092

end data.

variable labels Recall "Recall Newsome".

value labels Recall 1 'yes' 2 'no'.

variable labels DataType "Survey vs Vote Result".

value labels DataType 1 'Survey' 2 'Vote Result'.

weight by count.

crosstabs tables = Recall by DataType

/cells = column count

/statistics = phi chisq.

Recall Newsome * Survey vs Vote Result Crosstabulation

		Survey vs Vote Result			Total
		Survey	Vote Result		
Recall Newsome	yes	Count	563	4894473	4895036
		% within Survey vs Vote Result	37.5%	38.1%	38.1%
	no	Count	937	7944092	7945029
		% within Survey vs Vote Result	62.5%	61.9%	61.9%
Total		Count	1500	12838565	12840065
		% within Survey vs Vote Result	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.221 ^a	1	.638		
Continuity Correction ^b	.197	1	.657		
Likelihood Ratio	.222	1	.638		
Fisher's Exact Test				.651	.329
Linear-by-Linear Association	.221	1	.638		
N of Valid Cases	12840065				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 571.85.

b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	.000	.638
	Cramer's V	.000	.638
N of Valid Cases		12840065	

The September 2021 PPIC survey results very closely match the California recall election results. They differ by less than 1%. The phi coefficient confirms that the difference between the survey and election results is negligible. The probability of chi-square also suggests no significant difference between the survey and election results (p=.64), with a very high chance of obtaining such a percentage difference due solely to sampling error. Phi =.000; p = .638

3.

With the syntax used above as a model and referring to page 13 of *The Economist* YouGov data set [Link](#), write and run the data entry syntax for Immigration by PartyID. Submit your syntax and report the appropriate measure of association and its associated p-value.

Party ID		
Dem	Ind	Rep
42%	46%	73%
39%	32%	20%
15%	13%	5%
4%	8%	1%
100%	99%	99%
(515)	(560)	(425)

*Data entry on Economist 2023 Data.

*ImmigImp vs PartyID.

uses raw numbers.

data list free / Immlmp PartyID count.

begin data.

1 1 216

1 2 258

1 3 310

2 1 201

2 2 179

2 3 85

3 1 77

3 2 73

3 3 21

4 1 21

4 2 45

4 3 4

end data.

variable labels Immlmp "Importance of Immigration".

value labels Immlmp 1 'very' 2 'somewhat' 3 'not very' 4 'not at all'.

variable labels PartyID "PartyID".

value labels PartyID 1 'Dem' 2 'Ind' 3 'Rep'.

weight by count.

crosstabs tables = Immlmp by PartyID

/cells = column count

/statistics = phi ctau chisq.

Importance of Immigration * PartyID Crosstabulation

		PartyID				
			Dem	Ind	Rep	Total
Importance of Immigration	very	Count	216	258	310	784
		% within PartyID	41.9%	46.5%	73.8%	52.6%
	somewhat	Count	201	179	85	465
		% within PartyID	39.0%	32.3%	20.2%	31.2%
	not very	Count	77	73	21	171
		% within PartyID	15.0%	13.2%	5.0%	11.5%
not at all	Count	21	45	4	70	
	% within PartyID	4.1%	8.1%	1.0%	4.7%	
Total		Count	515	555	420	1490
		% within PartyID	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	126.282 ^a	6	<.001
Likelihood Ratio	132.145	6	<.001
Linear-by-Linear Association	67.765	1	<.001
N of Valid Cases	1490		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 19.73.

Symmetric Measures

		Value	Asymptotic Standard Error ^a	Approximate T ^b	Approximate Significance
Nominal by Nominal	Phi	.291			<.001
	Cramer's V	.206			<.001
Ordinal by Ordinal	Kendall's tau-c	-.199	.020	-9.777	<.001
N of Valid Cases		1490			

a. Not assuming the null hypothesis.

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

Tauc = -.199; p <.001

Note sign as well as magnitude.

Interpretation

There is a moderately weak negative relationship (-.199) between respondents' Party Id and their assessment of the importance of immigration as a political issue. It is very unlikely this relationship is due to sampling error as relationship is statistically very significant ($p < .001$) according to chi-square. There is less than one chance in one-thousand to obtain such an association due to chance.

These data are from the US as a whole. The magnitude of the association here between Party ID and attitude toward immigration is markedly less than what we found in either California or Texas using different variables. This may suggest that opinion on immigration is less polarized along partisan lines in the rest of the country than it is in California or Texas.

Given the coding of the variables (Republican = hi & immig being very important = low) the negative ordinal measure of association says the Republicans are more likely to say immigration is very important as a concern and Democrats are more likely to say immigration is less important. This shows the importance of understanding the direction of coding of the variables. Moreover, it suggests recoding the variables to facilitate more ready interpretation. If immigration were coded with concern being hi, the coefficient would be positive.

Consistent with the coding used in the Calif and Texas data I would code Democrat as hi and concern over immigration as high. This would still produce a negative coefficient with Democrats being less concerned over immigration and Republicans being more so.