

CSC433 Scripting for Data Analysis Assignment, Part 4a)

Matt Triano

May 1, 2017

1 Data Source (This Dataset is Too Hot To Handle Comfortably)

This is a dataset that you absolutely cannot talk about at parties; it focuses on all of the prohibited topics (religion, money, politics, race). But those things are all extremely influential in shaping human behavior, so they're worth studying, even if it made me uncomfortable. Even still, I didn't dig into the data collection methods and I'm just interpreting the numbers.

1.1 Bible Interpretation By Gender

This was the most vanilla question I investigated. I plotted bible interpretation grouped by gender, and per the Chi-Square value of 22.9669 (p -value : 0.0001), there is a statistically significant association of gender and bible interpretation.

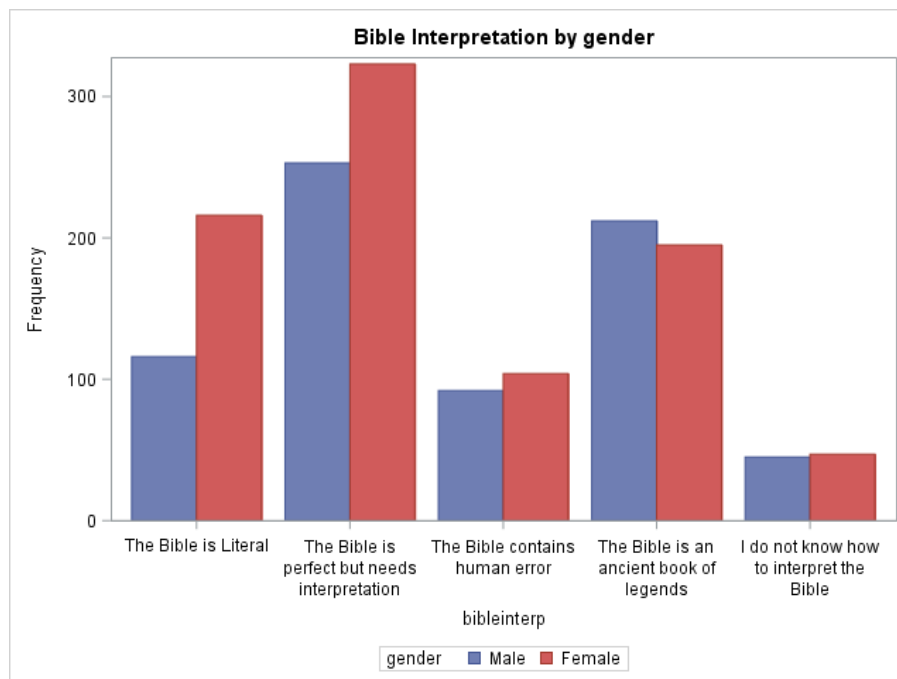


Figure 1: Bible Interpretation By Gender: Barplot

Bible Interpretation by gender

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of bibleinterp by gender			
	bibleinterp	gender		
		Male	Female	Total
	The Bible is Literal	116 7.24 34.94 16.16	216 13.47 65.06 24.41	332 20.71
	The Bible is perfect but needs interpretation	253 15.78 43.92 35.24	323 20.15 56.08 36.50	576 35.93
	The Bible contains human error	92 5.74 46.94 12.81	104 6.49 53.06 11.75	196 12.23
	The Bible is an ancient book of legends	212 13.23 52.09 29.53	195 12.16 47.91 22.03	407 25.39
	I do not know how to interpret the Bible	45 2.81 48.91 6.27	47 2.93 51.09 5.31	92 5.74
	Total	718 44.79	885 55.21	1603 100.00
Frequency Missing = 45				

Statistics for Table of bibleinterp by gender

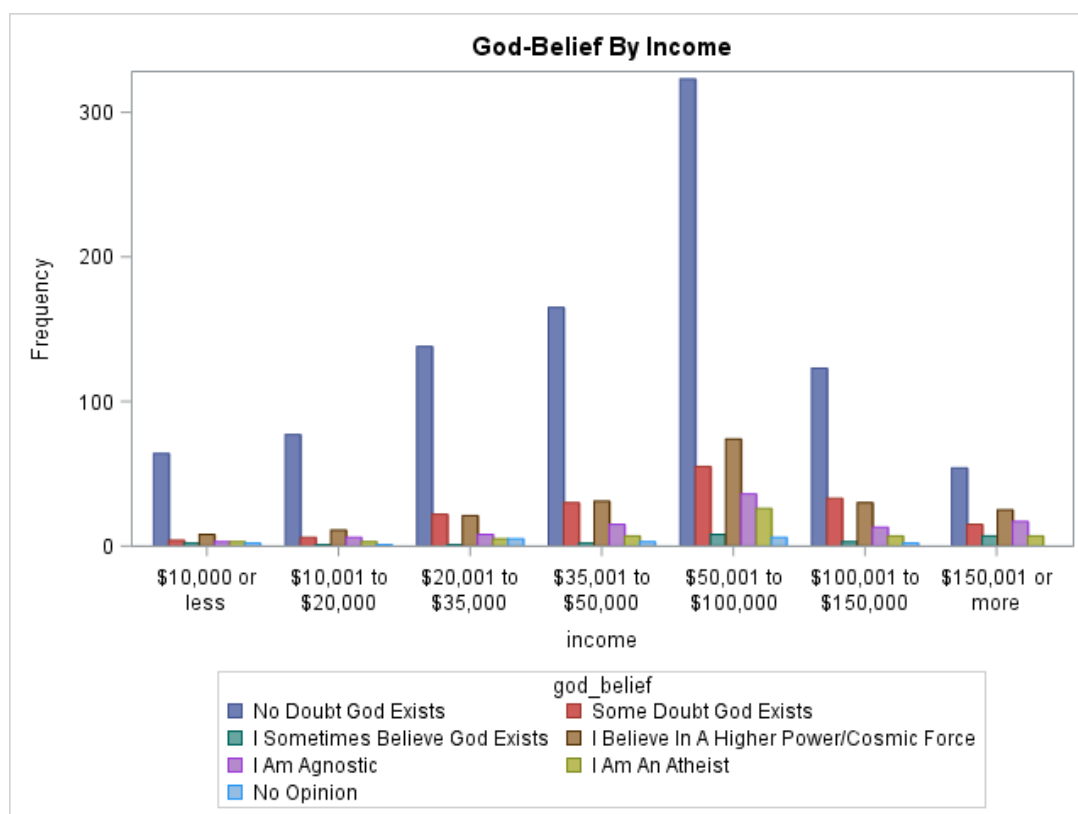
Statistic	DF	Value	Prob
Chi-Square	4	22.9669	0.0001
Likelihood Ratio Chi-Square	4	23.1806	0.0001
Mantel-Haenszel Chi-Square	1	19.3985	<.0001
Phi Coefficient		0.1197	
Contingency Coefficient		0.1188	
Cramer's V		0.1197	

Effective Sample Size = 1603
Frequency Missing = 45

1.2 God-Belief By Income

To explore the relationship between belief-in-god and income, I plotted the people's statement on their faith and their asserted incomes, using the bins already provided. For all income brackets, firm faith in god is the most frequent belief. From the row percents on the table, we see that as income increases, the percentage of responses indicating firm faith decreases, but remains clearly dominant. Interestingly, at the highest income bracket, the proportion of respondents that believe in a 'Higher Power or Cosmic Force' peaks at nearly half the proportion that describes that higher power as 'God'.

From the $0.0005p$ - *value* on the Chi-Square analysis table, we can infer that there is a statistically significant association between income and belief-in-god.



God-Belief By Income

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of income by god_belief								
	income	god_belief							
		No Doubt God Exists	Some Doubt God Exists	I Sometimes Believe God Exists	I Believe In A Higher Power/Cosmic Force	I Am Agnostic	I Am An Atheist	No Opinion	Total
	\$10,000 or less	64 4.24 74.42 6.78	4 0.27 4.65 2.42	2 0.13 2.33 8.33	8 0.53 9.30 4.00	3 0.20 3.49 3.06	3 0.20 3.49 5.17	2 0.13 2.33 10.53	86 5.70
	\$10,001 to \$20,000	77 5.11 73.33 8.16	6 0.40 5.71 3.64	1 0.07 0.95 4.17	11 0.73 10.48 5.50	6 0.40 5.71 6.12	3 0.20 2.86 5.17	1 0.07 0.95 5.26	105 6.96
	\$20,001 to \$35,000	138 9.15 69.00 14.62	22 1.46 11.00 13.33	1 0.07 0.50 4.17	21 1.39 10.50 10.50	8 0.53 4.00 8.16	5 0.33 2.50 8.62	5 0.33 2.50 26.32	200 13.26
	\$35,001 to \$50,000	165 10.94 65.22 17.48	30 1.99 11.86 18.18	2 0.13 0.79 8.33	31 2.06 12.25 15.50	15 0.99 5.93 15.31	7 0.46 2.77 12.07	3 0.20 1.19 15.79	253 16.78
	\$50,001 to \$100,000	323 21.42 61.17 34.22	55 3.65 10.42 33.33	8 0.53 1.52 33.33	74 4.91 14.02 37.00	36 2.39 6.82 36.73	26 1.72 4.92 44.83	6 0.40 1.14 31.58	528 35.01
	\$100,001 to \$150,000	123 8.16 58.29 13.03	33 2.19 15.64 20.00	3 0.20 1.42 12.50	30 1.99 14.22 15.00	13 0.86 6.16 13.27	7 0.46 3.32 12.07	2 0.13 0.95 10.53	211 13.99
	\$150,001 or more	54 3.58 43.20 5.72	15 0.99 12.00 9.09	7 0.46 5.60 29.17	25 1.66 20.00 12.50	17 1.13 13.60 17.35	7 0.46 5.60 12.07	0 0.00 0.00 0.00	125 8.29
	Total	944 62.60	165 10.94	24 1.59	200 13.26	98 6.50	58 3.85	19 1.26	1508 100.00
Frequency Missing = 140									

Statistics for Table of income by god_belief

Statistic	DF	Value	Prob
Chi-Square	36	70.6015	0.0005
Likelihood Ratio Chi-Square	36	66.7489	0.0014
Mantel-Haenszel Chi-Square	1	18.2254	<.0001
Phi Coefficient		0.2164	
Contingency Coefficient		0.2115	
Cramer's V		0.0883	
WARNING: 31% of the cells have expected counts less than 5. Chi-Square may not be a valid test.			

Effective Sample Size = 1508
Frequency Missing = 140

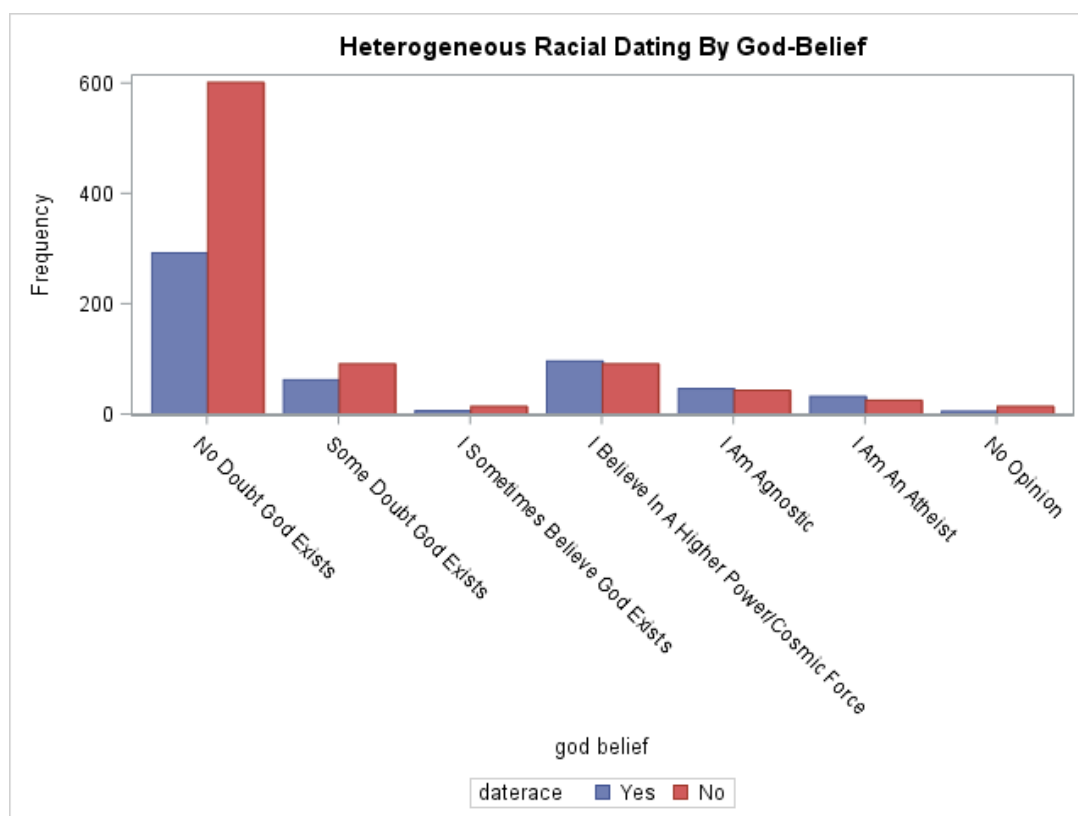
1.3 Heterogeneous Dating Experiences by God-Belief

This question made me the most uncomfortable, but while reviewing the study booklet, I noticed question 55, which kind of offended me. This reaction made me reflect and wonder why (presumably) real data would offend me, so I decided to see what the data said and test a hypothesis: is there a statistically significant association between racially heterogeneous dating and belief in god.

Granted, this analysis says nothing about causation and can't disentangle factors like local demographics, local culture, motivations for belief in god, or (most importantly) honesty of the respondent.

That being said, there was near parity in the (self-reported) yes and no responses of Agnostics or believers in a Higher Power/Cosmic Force (both slightly on the affirmative side). Atheists were even more progressive in their answers, with the largest proportion responding in the affirmative. Aside from those with 'No Opinion' (only 19 respondents, so potentially not a representative sample of any population), those who firmly believe in God were the least likely to have a racially heterogeneous romantic partner.

From the $< .0001p - value$ on the Chi-Square analysis table, we can infer that there is a statistically significant association between reporting a racially heterogeneous dating experience and belief-in-god.



Heterogeneous Racial Dating By God-Belief

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of god_belief by daterace			
	god_belief	daterace		
		Yes	No	Total
	No Doubt God Exists	292 20.59 32.70 54.17	601 42.38 67.30 68.37	893 62.98
	Some Doubt God Exists	62 4.37 40.52 11.50	91 6.42 59.48 10.35	153 10.79
	I Sometimes Believe God Exists	6 0.42 30.00 1.11	14 0.99 70.00 1.59	20 1.41
	I Believe In A Higher Power/Cosmic Force	96 6.77 51.34 17.81	91 6.42 48.66 10.35	187 13.19
	I Am Agnostic	46 3.24 51.69 8.53	43 3.03 48.31 4.89	89 6.28
	I Am An Atheist	32 2.26 56.14 5.94	25 1.76 43.86 2.84	57 4.02
	No Opinion	5 0.35 26.32 0.93	14 0.99 73.68 1.59	19 1.34
	Total	539 38.01	879 61.99	1418 100.00
Frequency Missing = 230				

Statistics for Table of god_belief by daterace

Statistic	DF	Value	Prob
Chi-Square	6	41.8593	<.0001
Likelihood Ratio Chi-Square	6	41.1986	<.0001
Mantel-Haenszel Chi-Square	1	29.1074	<.0001
Phi Coefficient		0.1718	
Contingency Coefficient		0.1693	
Cramer's V		0.1718	

Effective Sample Size = 1418
Frequency Missing = 230

WARNING: 14% of the data are missing.