Cecil Thompson

Omega Graduate School

COM 968-52 Statistics for Social Research III

Prof. Richard Nti

Dr. Sean Talady

February 17, 2024

Assignment #3 – Essay In this advanced statistics assignment, you will continue working with the sample dataset from the fictional study "Religiosity and Social Behavior in a Diverse Community." You will utilize PSPP statistical software to perform complex statistical analyses.

Follow the steps below:

1. Introduction - Compose an overview of the sections of this assignment and justify why they are important for statistical analysis in research.
2. Splitting Data Based on Independent Variables (1 page) - Start by splitting the

dataset into groups based on selected independent variables. Explain your rationale for selecting these variables and how the dataset has been divided. Ensure you have distinct groups for comparison in subsequent analyses (for example, two racial or gender categories).

3. Conducting Significant Difference and Correlational Procedures (2 pages) –

Choose appropriate statistical procedures to analyze the relationship between

the independent and dependent variables.

1. Conduct analyses such as independent t - tests, ANOVA, or correlation

analysis using PSPP.

b. Clearly describe the statistical procedures you are using.

c. Present the results, including statistical significance and effect sizes if

applicable.

1. Conducting Post - hoc Procedures (1 page) - If your analyses reveal significant

differences between groups, conduct post - hoc procedures (e.g., Tukey's HSD, Bonferroni correction) to identify specific group differences. Explain the post - hoc tests used and their purpose.

1. Interpreting Results (2 pages) - Interpret the results of you r statistical procedures comprehensively.
2. Discuss the practical significance of significant differences or correlations and how they relate to your research questions. Provide insights into the implications of your findings. 6. Conclusion and Reflection (1 page) - Summarize your findings and the key takeaways from your analysis. Reflect on the process of conducting complex statistical analyses using PSPP. Discuss how this analysis informs your research and contributes to understa nding the dataset. Ensure your assignment is well - structured, concise, and supported with appropriate statistical output generated using PSPP. This assignment will help you apply your knowledge of data splitting, advanced statistical procedures, post - hoc analysis, and results interpretation in a practical context using statistical software. The assignment should be 7 - 10 pages long. Cite research design experts from your developmental readings and use proper APA formatting.

Although the samples in the “Fictional Study” provided are expected to

have been chosen randomly, and probably were, the effort to do the sampling from a “diverse” population appeared to have been done purely on the basis of race. The sample comprised of 42% of the majority race, 33% of the second largest racial group, and 25% of the smallest group. One would think that the sampling of a typical racially diverse American community, would likely have contained a tiny percentage of “others”. But rather than us making an off-the-cuff assumption that there must have been some kind of shortcoming in the extent to which bias had been mitigated during the sample collection process, the absence of very small racial minorities, even if lumped together as “other”, might very probably be an indication that the sample size is so very small, at twelve.

To guide the sample selection, process, the researcher would, obviously have begun with the formulation of a primary hypothesis and proceed to split the data into independent (x**1**) and dependent (y1) variables. In the fictional study provided, the statement about the object is that it is meant to explore whether there is a relationship between “Religiosity” and “Social Behavior” of people living in a diverse society. The question is, if such a relationship exists, what might the strength of that relationship be? Once the diverse population is selected, it would seem necessary for the researcher to draw a “random sample” of 30 or more participants (a sample size of 12 participants is provided by the fictional study). With the primary hypothesis firmly in place, the researcher would, thus, be able to make an assumption about how the data collected suggest that “Religiosity” of people living in such diverse communities is the independent variable (x1) and their “Social Behavior” is the dependent variable (y1). As such, this major hypothesis could be that, “There is a statistically significant relationship between the religiosity and social behavior of people living in a racially diverse community” (Sharma, 2019).

By providing Columns D, E, and F, the fictional study, has, in effect, already “split” the major dependent variable (Social Behavior) into three dependent variables. Since the dependent variable shown in the major hypothesis above broadly covers different forms of “Social Behavior,” it does seem natural to treat Social Behavior as something of a “universal set” containing several “sub-sets” or “categories” of social behaviors. For that reason, the researcher has apparently chosen to split up the major dependent variable (Social Behavior) into three categories (The Fictional Study).

By treating the dependent variable—Social Behavior— as the universal set, it can thus be splintered into a number of sub-sets which would still fall under “Social Behavior”. In this way, the strength of the relationship between each of these subordinate dependent variables and the single independent variable—“Religiosity—in the major hypothesis, should be roughly consistent, if the argument in the hypothesis has any real merit. That is to say that the proposition that “Social Behavior” is a function of “Religiosity” should hold even though one of the sub-sets of Social Behavior, belonging to a particular type, namely, “Community Service,” may not be a social behavior that is applicable to the large majority of this racially diverse society. Further, by splitting the dependent variable, as the fictional study has done, the door is pushed wide open to a more detailed investigation of the major thesis through the generation of specific minor subordinate hypotheses (Fictional Study; Reichard, 2021; Sharma, 2019).

The headings of Columns A through F, in the fictional study, range from “Participants ID”, through “Race,” “Community Service (hours),” “Social Justice Attitudes (1-5),” to “Social Cohesion Score (1-10),” respectively. The headings, from C through F, are all examples of how the fictional researcher had split the given dependent variable into three types of social behaviors which could still have been split even further. For example, “Community Service (hours)”, for one, may even be split further into “Food Pantry Service” and “Education of the Homeless”.

However, if we focus on just the three types of already splintered pieces from the broad “Social Behavior” of people living in racially diverse communities, we should understand that those three, by themselves, should suffice in our effort to “. . . [c]onduct significant difference and correlational procedures using statistical software to examine the relationship between variables in the dataset” (Fictional Study). To pair the independent variable, “Religiosity,” with each of the other three splintered pieces from the major dependent variable, (“Social Behavior”) we, first, need to hypothesize in a manner that presses the assumption which sets up the pairing, and hence, the possible relationship between the independent variable, in Column C (“Religiosity),” and each newly splintered piece of the major dependent variable (Social Behavior).

The first hypothesis drawn from the fictional study could be that, “There is a statistically significant relationship between the religiosity of people in a racially diverse community and the number of hours they invest in community service.” The second would probably be, “There is a statistically significant relationship between the religiosity of people living in a racially diverse community and their Social Justice Attitudes.” And, the third could be that, “There is statistically significant relationship between the religiosity of people living in a racially diverse community and their Social Cohesion Score.”

Working with these three hypotheses, the researcher’s objective is to determine whether there is any relationship between the independent variable (x**1**) and its paired partnership with each of the three dependent variables—yd, ye, and y**f**, (McCarty’s 2016-2024 video tutorial). The strength of the relationship between each of the three types of dependents that was split off the primary dependent variable, “Social Behavior,” being paired together with the independent variable, “Religiosity,” should further help to prove or disprove the strength of the broader relationship between “Religiosity” and “Social Behavior”.

To recap, in this fictional study, the researcher has split up the universal set of “Social Behavior” into subsets of, “Community Service (hours),” “Social Justice Attitudes (1-5),” and “Social Cohesiveness Score (1-10)”. Properly speaking, then, this fictional study has generated three minor hypotheses containing one independent variable that is paired with a different dependent variable for each of the three different hypotheses.

First, “Religiosity (x**1**),” the independent variable, is paired with the dependent variable, “Community Service (y**d**)” to form the first subordinate hypothesis from the dataset in the fictional study. The proposed hypothesis is, “There is statistically significant relationship between the religiosity of people living in a racially diverse community and the number of hours they put into community service.” When we look at the histogram that depicts the relationship between the paired variables (x and y), we get a visual sense of the evenness versus the degree of disproportionality between the blue and orange pairs that tell a graphic story about the correlation coefficient that links the two variables.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | |  | |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Once the population sample is selected and the data from the participants, as to their “Religiosity Levels” and “Community Service”, are paired, the data connected to the independent and dependent variables are highlighted in Excel and the correlation coefficient is produced. In this case, the correlation coefficient turns out to be r = 0.545538686. This is the number rounded off to three decimal places in the bottom right- hand corner of this histogram. If rounded to the nearest hundredth, r = 0.55, indicates that there is a 55% probability of there being a statistical relationship between “Religiosity Level” and “Community Service” (Bradburn; McCarty, 2016-2024).

The scatter plot below visually presents a picture which shows that only 6 of the 12 datapoints are relatively close to the line passing among and between them. This proximity of the points (x, y) to the line (y = mx + b) falls in the class of “standard error”.

Religiosity Level (XC) and Community Service (YD)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | |  | | --- | |  | |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  | |  |

In this graph, the “line of best fit” or the “line of regression” is positive because it lies in the first quadrant of the Cartesian plane. In this quadrant, y increases as x increases. Thus, everything about the nature of this plot is described by the linear form, y= mx + b. As such, “m” is the slope and “b” is the y-intercept. Therefore, since m = 1.7358 and b = 5.1873, we should be able to pick an x-value and plug it in so as to “predict” the value of y, in the equation, y = 1.7358x + 5.1873. The expectation is that the “predicted” value of y should be relatively close to the value of the original y, paired with the substituted x given in the fictional dataset.

Taking the first of the 12 x-values under the independent variable (Column C) and the first of the 12 y-values under the dependent variable (Column D), in this relationship, we could validate the “line of best fit” in our effort to show the strength of the relationship between “Religiosity Level” and “Community Service (hours)”. In these two Columns, C and D, the first datapoint is x = 4 and y = 10. So, by substituting the 4 into x, into y = 1.7358(x) + 5.1873, the predicted value of y should be a number relatively close to the actual value of y, which is10. Of course, in a case such as this, where the correlation coefficient is as moderate as r = 0.545538686, any of the 12 datapoints could be way off collinearity because of “standard error”.

Here, substituting 4 into the x, of y = 1.7358(x) + 5.1873, it just happens that the “predicted” value of y turns out to be reasonably close to 10, at 12.1305.

Upon pairing all the x and y-values of the datapoints between “Religiosity” and “Community Service”, Excel produced a correlation coefficient of r = 0.545538686. Based on Jason Fernando’s (2023) explanation of the function of the correlation coefficient, Pearson r should indicate to the statistician what the “strength and direction” of the relationship between the paired independent and dependent variables of a hypothesis is depictable, graphically, by the near collinearity of the datapoints with the “line of best fit” (Fernando, 2023, p.n.).

Although a “correlation coefficient”, by itself, is not an indication that the behavior of one variable affects the behavior of the other, the dependent variable (y) is frequently spoken of mathematically as a “function” of the independent variable (x). So, at least, mathematically, the behavior of x affects the behavior of y. What the correlation coefficient is supposed to help us to see is whether or not we can make a reasonable assumption about the strength of the relationship between the x and the y variables. Also, the convention that sets up how Pearson r helps the researcher to determine the strength of the relationship between the two variables differ among the families of disciplines (Fernando, 2023, p.n.).

According to Jason Fernando (2023), the strength of the correlation coefficient often depends on the discipline to which Pearson is being applied field. He says that in physics and chemistry, Pearson r has to be less than -0.9 or greater than 0.9 in order for it to be considered meaningful. While, in the social sciences, a correlation coefficient of -0.5 or +0.5 is considered marginally significant (Fernando, 2023, p.n.).

Whereas a correlation coefficient such as r = 0.545538686, which is what we have in this relationship between “Religiosity Level” and “Community Service”, is generally considered to be of “fair” strength in the social sciences, Fernando (2023) writes that the threshold of acceptance is r = 0.94 in the physical or chemical sciences. The question is can one’s finding of a correlation coefficient help one to determine how to increase or decrease a certain behavior of a dependent variable during its relationship with the independent variable?

Cohen (1988) writes that, presently, “Effect Size” has become the preferred approach over the use of test-statistic in medical research, when an effort is being made to separate the effect that an independent variable (x) may have upon a dependent variable (y) from outside effects caused by variances. He describes it as (Eta squared) which he says “. . . is analogous to r2 but is used for t-tests and ANOVA. Like r2, it reflects the percentage of variability in the dependent variable that can be explained by the independent variable” (Cohen, 1988). While Fernando (2023), in speaking of the process he entitles “Coefficient of Determination,” which he identifies as R-squared, he writes that the process shows “. . .what portion of the variation in the dependent variable is attributableto the independent variable” (Fernando).

In the above regression graph, R2 is one of the values presented by Excel.

The Excel software has given us R2 = 0.2976. That number happens to be the square of the correlation coefficient— (0.545538686)2 – in the statistical relationship between “Religiosity Levels” and “Community Service”. If we, then, convert 0.2976 to percent, we get 29.76%. Rounded to the nearest hundredth, the result is 30%.

What Cohen and Fernando are telling us is that “Religiosity Level (x1)” accounts for 30% of the effect on “Community Service(yd)”. And, by that very fact, the remaining 70% of the effect on “Community Service (yd)” comes from variances such as standard error, biases, intervening variables, and the likes. So, whatever Pearson r, itself, can do or cannot do, it still appears to be the cop on point duty whose gesticulatory commands are everywhere involved in the directing of traffic in the correlational universe, as it is, here, in the matter of the effect from variances on the dependent variable in the form of what is not within the bounds of R2.

Based on the relationship between the paired data of “Religiosity Level” and “Community Service”, it is fair to assume that the data are normally distributed. Through the calculations by Excel, it is shown that there is an “alpha-value” equal to 0.05. To understand the meaning of “alpha value”, it is important to look at it in its relationship to the “p-value” and how that latter value relates to the probability of “Statistical Significance”.

In order to properly apply these terms, it is necessary to, first, present a null hypothesis (Ho). This is tantamount to saying that it is essential for the researcher to begin by testing the viability of a proposition that negates the very proposition he intends to argue.

Since the null hypothesis, Ho, is set up to argue the exact inverse position of the alternate hypothesis, H0, that the researcher wishes to prove, he must begin by running a test to determine whether H0 is rejected. Ordinarily, it would be necessary to figure out the sample mean, and hence, the spread of the data (the standard deviation) and locate the z-value in the z-standard table. But by being given the alpha-value of 0.05 by Excel, we are able to determine, with a high degree of “Confidence”, that the p-value is less than 0.05. And, by that fact, the “Probability of Statistical Significance” will be determined to be 95% or greater.

By using Excel, test statistics devices such as t-tests, ANOVA, F-tests, and so forth, can be used to find variances and statistical measures for determining statistical significance. For example, the F-test table below is generated by Excel in respect to the relationship between “Religiosity Level” and “Community Service”.

|  |  |  |  |
| --- | --- | --- | --- |
| F-Test Two-Sample for Variances | | |  |
|  |  |  |  |
|  | *Variable 1* | *Variable 2* |  |
| Mean | 10.25 | 2.91666667 |  |
| Variance | 22.9318182 | 2.26515152 |  |
| Observations | 12 | 12 |  |
| df | 11 | 11 |  |
| F | 10.1237458 |  |  |
| P(F<=f) one-tail | 0.00029734 |  |  |
| F Critical one-tail | 2.81793047 |  |  |

SD = 2

n

Standard Deviation = 1.440966579

SD

SE =

Standard Error = 0.4159712212

The pairing of “Religiosity Level” with the second dependent variable (Column E) splintered off the primary dependent variable (“Social Behavior”), taken from the table in the fictional study, is “Social Justice Attitudes”. In response to this pairing of these two variables, the histogram shows the number, 0.7693, in the bottom right-hand corner. This number happens to be the correlation coefficient rounded to the ten thousandth place. By this fact, it is clear that the relationship between the pairs of colored bars as well as among the numerical pairings of those in “Religiosity Level” and “Social Justice Attitudes” is much stronger than that between “Religiosity Level” and “Community Service”.

|  |
| --- |
|  |
|  |
|  |
|  |
|  |

Given the fictional dataset, when “Religiosity Level” is paired with the “Social Justice

The data from this second pairing involving the relationship between “Religiosity Level” and “Social Justice Attitudes” of people living in a racially diverse community, produces a correlation coefficient of r = 0.769283575. It also confirms the number in the bottom right-hand corner of the histogram. In round numbers, this correlation score approximates, r = 0.77, which tells us that there is a probable 77% relationship between “Religiosity Level” and “Social Justice Attitudes”. When looked at in terms of the comparative strength of the relationship between these two variables and that between those of the earlier pairing—“Religiosity Level (x)” and “Community Service (y)”—the relationship between “Religiosity (x)” and “Social Justice Attitudes (y)” is considerably stronger. Like the relationship between the first pairing of variables, the fact that the coefficient is positive indicates that the line of regression is in the first quadrant of the Cartesian plane and should display all the attendant characteristics accordingly with a few important exceptions. One of the important exceptions is the fact that with a Pearson r that is considered to be reasonably strong, at 0.769283575, the “Determination of Effect” is significant.

Here, with the coefficient r = 0.769283575, the square of “R” would be 0.5917972188. And, when r2 is converted to percentage, we could say that “Religiosity Level” would have approximately a 59% effect on the “Social Justice Attitudes” of people living in a racially diverse community as that in this fictional study. In which case, the effect from variances such as biases, errors, and intervening variables would account for 41% of the variances that impacted “Social Justice Attitudes”.

Religiosity and Social Justice Attitudes

The pairing of “Religiosity Level (x1)” with the third form of “Social Behavior (y1)”, from the fictional dataset, is that pairing which involves “Social Cohesion (Column F)”. The correlation coefficient between “Religiosity Level (Column C)” and “Social Cohesion (Column F)” closely resembles that between the “Religiosity Level (Column C)” and “Social Justice Attitudes (Column E)”.

So, by Excel’s calculation, Pearson r, in this third pairing, is 0.756759345. This number rounded to the nearest hundredth is 0.76. And the relationship between “Religiosity Level (C)” and “Social Cohesion (F)”, at 76%, is roughly in the same ballpark as that of “Religiosity (C)” and “Social Justice Attitudes (F)”. Since it is now generally recognized that the “effect” that an independent variable has upon the dependent variable can be determined by R2, we may proceed to conclude that with R2 = 0.5726847062, the effect of “Religiosity Level” on “Social Cohesion” is about 57%.

*Ipsofacto*, the collective outside-effect from variances would be 43%. And, with that, one might expect to see a histogram that shows some “visual” resemblance to the previous pair which purports a relationship between the same independent variable, “Religiosity Level,” and that previous dependent variable, “Social Justice Attitudes” with the collective variances amounting to 41%. Yet, despite the similarity in the numbers, the two histograms do not show any real resemblance to each other.



While, in comparing the extraneous effects on the dependent variables in the two situations, we find that the pile of variances stacked up in the relationship between the “Religiosity Level” and “Social Justice Attitudes” case, being 41%, and the other pile in the “Social Cohesion”, case being 43%, implies a similarity that is not so obvious in the histograms when viewed side by side. The wonder is whether we should expect to see greater similarity in the behavior of the two different pairs of variables in linear regressions.

Well, at least, all we can tell from comparing the r- values and the R2-values is that all the linear regressions are occurring in the first quadrant with each dependent variable being a function of the independent variable. What we can sale is that in spite of the few obvious similarities we find between the two sets of results and their corresponding graphs, there is no guarantee that the datapoints and their loci, within the scatterplot, will be anything close to mirror images of each other.

The scatterplot below is an illustration of the relational behavior of the datapoints and the line of best fit for the pairing of “Religiosity Level (Column C)” and “Social Cohesion (Column F)”. With this we should be able to make a visual comparison between this and the scatter plot for the second pair of variables above.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  |  | |  | |
|  |  | |  | |
|  |  | |  | |
|  |  | |  | |
|  |  | |  | |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |
|  | |  | |  | |  |

At this point, when we contrast and compare the statistical relationship between the independent variable, “Religiosity Level,” and each of the three dependent variables that result from splitting the primary dependent variable, “Social Behavior,” we are better positioned to recognize the importance of the relative strength of the correlation coefficient among the three pairs of variables. In the first situation, “Religiosity Level” is paired with “Community Service”. Here, the correlation coefficient of r = 0.55 indicates that the statistical relationship between the variables stands at 55%. In such a case, the strength of the relationship between the independent variable (Column C) and the dependent variable (Column E) is moderate. At the same time, the extent of the statistical significance in the relationship between “Religiosity” and “Social Justice Attitudes” is quite strong at r =0.77, or 77%. The correlation in the third pairing—between “Religiosity” and “Social Cohesion”—is also considered to have strong statistical significance at r = 0.76, or 76%.

The fact that, in all three of these cases, the value of r is positive indicates that both the independent (x) and the dependent (y) variables are positive. Clearly, since the increase of y as x is not a constant proportion for all datapoints, the software depicts the situation as a “line of best fit” in a nest of scattered points. Whenever the y-value increases in response to an increase in the x-value, at an average of 55% rate of increase, as it is in the case of the first paired variables, approximately one half of the scattered datapoints will be relatively close to the line of best fit. In similar fashion, the disproportionate increase will occur between the X s and Y s with the second pair of variables. Only that, in this case, the average rate at which y-values increase as x-values do, is 77%, which is reflected in the scatter plot. Likewise, in respect to the third pair of variables, y-value will also increase as x-value does, at an average rate of 76%.

So, in each of these situations, the closer Pearson r is to 1, the closer will the increase of y-values be to a proportional increase of the x- values. If the increases in y-values are all “directly proportional” to the increases in the x-values, the correlation coefficient will be r =1, which is 100%. In such a case, we will have perfect collinearity, with no scattered datapoints about the line of “best fit”. If, however, the x-value decreases as the y-value increases, the correlation coefficient will be negative and will, therefore, be depicted as a single line in the second quadrant. If the x-values decrease in “direct proportionality” as y-values increase, the correlation coefficient will be -1, and the statistical relationship between the independent and dependent variables will be 100%.

The fictional study has taken care to draw its sample from what it terms to be a diverse population. Yet, as we put out the effort to determine what is meant by a “diverse population, we see that it is categorized simply as Race. However, it does seem a bit problematic that only three races are sampled. The three races are reprsented as 1, 2, and 3.

One would have expected that a “random sampling” of a racially diverse population, such as would be found in a typically diverse American community, would contain Whites, Blacks, Latinos, Asians, Middle Eastern folk, and others. Even if we decided to excuse the suspicion of bias in the sampling process, here, we have to acknowledge that there is too “small” a sample to be considered a truly respectable representation of a racially diverse American population. At any rate, it is interesting to see what the outcome is when we proceed to test the reasonableness of the sampling by proposing the absolutely absurd idea of a hypothesis in which “Race (Column B” is treated as the independent variable and “Religiosity Level (Column C)” as the dependent variable.

Given the data provided in this fictional study, we try to test the absurdity of a correlation coefficient for the relationship between “Race (B)”, the independent variable, and “Religiosity Level (C)” the dependent variable. And there is no surprise at what the software tells us. What Excel tells us is that in this relationship, Pearson r is equal to -0.156763099. This correlation coefficient tells us that the relationship between “Race” and “Religiosity Level”, based on this fictional data, is extremely weak. A graphic representation of this weak relationship between the independent (“Race”) and dependent (“Religiosity Level”) variables, in the histogram below, tells us precious little outside of its confirmation of the fact that r, rounded to the nearest hundredth, is -0.16.

|  |  |  |
| --- | --- | --- |
| |  | | --- | |  | | |
|  | |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | | |  | | --- | |  | | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | | And, at length, when we use Eta-squared to determine the effect that “Race” has on “Religiosity  Level”, in a racially diverse population, we get R2 = 0.0245746692. What this tells us is that the effect of  “Race” on “Religiosity Level” in such a population is only about 2.4%. It follows, then, that 97.6% of what  affects the “Religiosity Level” a person living in such a community, to would come, not from “Race”,  but from a host of variances. For that reason, there should be little surprise as to what the scatterplot  for this relationship will show.   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | |  | | --- | |  | |  |  |  |  |  | | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  |  |  | |  | |  | |  | |  | |  | |  | | | |  | | --- | |  | |  |  | |  | |  | |  | |  | |  | |  | | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | | Works Cited | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | | Contrastingly, these datapoints are mostly scattered away from the line of regression in a manner  very unlike what we have seen in the relationships between the variables of the previous three cases.  And finally, the line, lying in the second quadrant, has a negative slope as indicated by the liner equation  y = -0.2826x + 3.4348 which is also quite unlike the previous three pairings of relationships between the  independent variable, “Religiosity Level,” and the dependent variables of Columns D, E, and F, that were  split off from “Social Behaviors”.  Works Cited | | |  | |  | |  | |  | |  | |  | |  | |  |  | | Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L.R., Johnson, J.D.  Linz, D. G., Malamuth, N.M., & Wartella, E. (2003). The influence of media violence on  youth. *Psychological Science in the Public Interest*, 4 (3), 18-110.  Bradburn, Steven. (n.d.). *How to perform a simple regression in* Excel. *Youtube* [Video].  Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, N.J.,  Lawrence Erlbaum Associates.  Fernando, Jason. (2023, Oct. 18). The correlation coefficient: What it is, what it tells investors.  *Investopedia*. Reviewed by Margaret James and Fact checked by Vikki Velasquez.  Hartin, Travis and Botts, Vanessa (2023, Jan. 11). Degrees of Freedom formula, uses and  examples.  McCarty (2016- 2024). *Using* Excel *to calculate a correlation coefficient and interpret*  *relationships between variables*. *Youtube*. [Video]  https://alphabench.com/data/excel-corr.  Reichard, Joshua (2021). *Statistics as a language: Overview of statistical tests and hypothesis*  *Testing* (Tutorial 2). *Youtube* [Video]. <https://www.youtube.com/watch?v=F17ek>.  Sharma, Vaibhav. (2019, Jan. 15). Importing and splitting data into Dependent and independent  features for machine learning. *Pluralsight*.  Stevens, Scott (2015). Excel *quick and simple charts* (Tutorial). *Youtube* [Video].  Zatz, David (2021). *Installing PSPP for Mac*. *Youtube* [Video].   |  | | --- | |  | | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |  | | |  | |  | |  | |  | |  | |  | |  | |  |  | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| |  | | --- | |  | |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |