**Project: Exploring the Relationship Between Income and VCR Ownership: Descriptive Statistics and Chi-Square Analysis.**

**1.0. Introduction**

The analysis delves into the interplay between personal electronics purchases and individuals' income levels. It serves as a comprehensive overview of statistical insights drawn from two key variables: individual purchases (the independent variable) and the corresponding income levels (the dependent variable). Through rigorous statistical procedures, the study aims to illuminate the nuanced relationship between these factors.

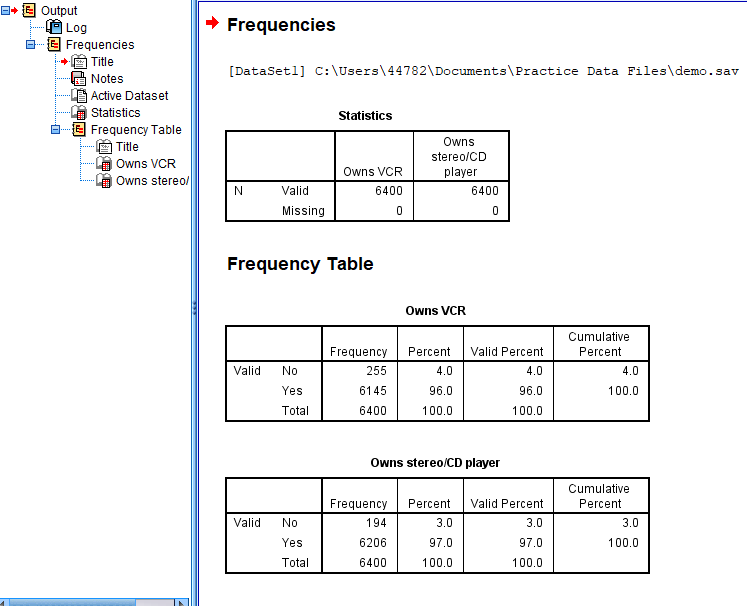
To deepen our understanding, a third variable was introduced as a control parameter. This additional element allows for a more nuanced examination of the significance level inherent in the relationship between personal electronics expenditures and income stories.

Utilizing the powerful analytical capabilities of the SPSS application, the study employed Chi-square as a robust statistical tool. Chi-square testing stands as a reliable method for gauging the strength and direction of relationships within datasets. This statistical approach provides a clear framework for evaluating the correlation between variables, offering valuable insights into consumer behaviour and income dynamics.

For further exploration and detailed findings, please refer to the data file provided alongside this analysis. This dataset serves as the foundation upon which the statistical conclusions and insights are built, offering a tangible representation of the intricate connections between personal electronics purchases and individuals' income profiles.

*demo.sav.* This is a hypothetical data file that concerns a purchased customer database, for the purpose of mailing monthly offers. Whether or not the customer responded to the offer is recorded, along with various demographic information.

**2.0. Frequency of VCR (own vcr)and Owns stereo CD player (own SCD)**



The frequency tables present a fascinating insight into ownership patterns among the surveyed population. Notably, the data unveils that 96% of individuals possess VCRs, indicating a high prevalence of this technology. Similarly, a substantial 97% of respondents are found to have stereo CD players, reflecting a widespread adoption of this audio equipment.

However, amidst these high ownership rates lies an intriguing revelation: a small percentage of individuals do not own VCRs or stereo CD players. Exploring further into this minority group could yield valuable insights into alternative consumer preferences, lifestyle choices, or perhaps economic considerations.

To visually represent these findings, the data is artfully displayed in a frequency table complemented by a bar chart. This visual aid not only provides a clear snapshot of the ownership percentages but also offers a compelling illustration of the disparity between the majority and the minority groups. Delving deeper into the reasons behind the choices of this smaller cohort promises to unveil a rich tapestry of consumer behaviours and preferences, adding depth and context to the overall analysis.

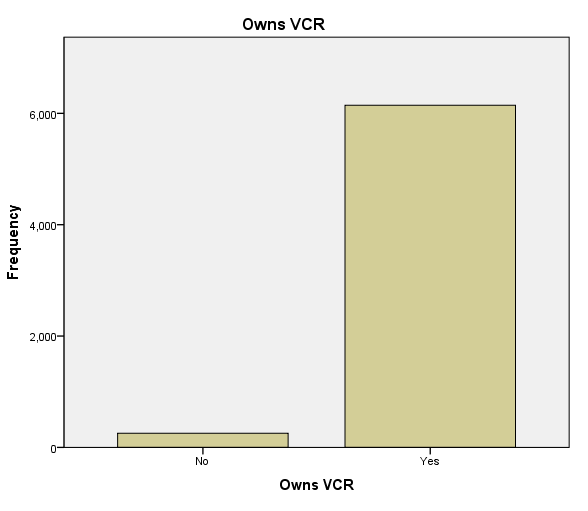
**3.0. Bar Chart of VCR (own vcr) and Owns stereo CD player (own CD)**

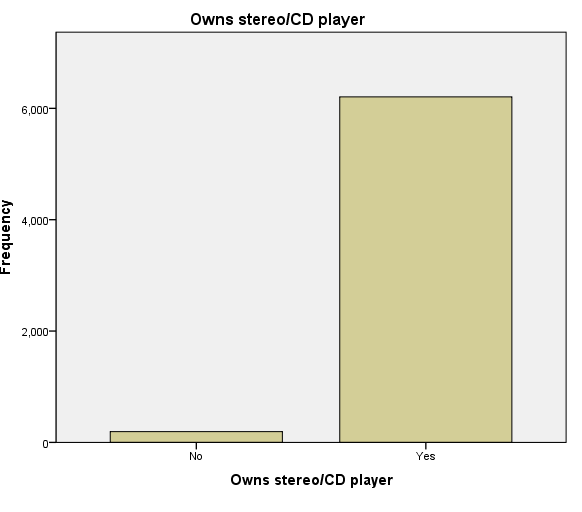
The Bar charts accompanying the frequency tables further illuminate the overwhelming prevalence of VCRs and stereo CD players among the surveyed population. The visual representation vividly portrays a stark reality: nearly everyone in the study cohort possesses both a VCR and a stereo CD player.

The Bar charts succinctly illustrate the distribution of ownership, with towering bars representing most individuals who own these electronic devices. Conversely, the corresponding minuscule bars reveal the tiny fraction of respondents who do not own either a VCR or a stereo CD player.

This graphical depiction reinforces the earlier findings, emphasizing the near-universal adoption of these technologies within the surveyed demographic. The visual impact of the Bar charts serves as a compelling testament to the widespread integration of VCRs and stereo CD players into the daily lives of the participants.

Moreover, the stark contrast between the dominant bars and the negligible ones prompts further curiosity about the motivations and circumstances of the small minority without these devices. Exploring the reasons behind their lack of ownership could unveil valuable insights into evolving consumer preferences, technological trends, or socio-economic dynamics within the studied population.





**4.0. Summary Measures for Scale Variables**

To provide a comprehensive analysis of the dataset, measures of central tendency and dispersion are essential components. These statistical metrics offer valuable insights into the distribution and variability of the data regarding personal electronics ownership.

Starting with measures of central tendency, the mean and median offer distinct perspectives on the average ownership within the surveyed population. The mean provides the numerical average of VCR and stereo CD player ownership, offering a balanced representation of the dataset's central value. Meanwhile, the median offers the middle point of the ownership data, disregarding extreme values and providing a robust measure of the dataset's central tendency.

Supplementing these measures are the measures of dispersion, which illuminate the extent of variability or spread within the ownership data. The standard deviation quantifies the average amount by which each data point deviates from the mean, providing a clear understanding of the data's overall variability. Additionally, the minimum and maximum values pinpoint the range within which ownership values fluctuate, offering insights into the breadth of ownership experiences within the surveyed population.

By incorporating these statistical measures, the analysis gains depth and precision. It not only highlights the prevalent ownership of VCRs and stereo CD players but also provides nuanced insights into the range and distribution of ownership levels. Understanding both the central tendency and dispersion of ownership data enriches our comprehension of consumer behaviours and preferences within the context of personal electronics acquisition.

|  |  |  |
| --- | --- | --- |
| **Statistics** | | |
| Household income in thousands | | |
| N | Valid | 6400 |
| Missing | 0 |
| Mean | | 69.4748 |
| Median | | 45.0000 |
| Std. Deviation | | 78.71856 |
| Minimum | | 9.00 |
| Maximum | | 1116.00 |

The statistical table presented above offers a detailed snapshot of key metrics, including the mean, median, standard deviation, minimum, and maximum values. Notably, the mean stands out prominently, showcasing a significant difference of approximately 24,475 units compared to the median. This discrepancy in central measures indicates a notable skewness in the distribution towards the higher end of the ownership values.

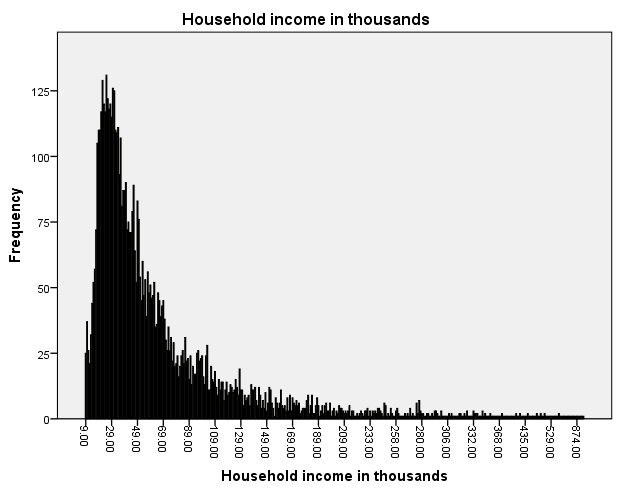
The mean's substantial divergence from the median suggests a positive skewness, indicating that a minority of respondents report much higher ownership values, pulling the mean upwards. In contrast, most respondents likely fall within a lower ownership range, as indicated by the median.

This positive skewness implies that the dataset's distribution is not symmetrical or normally distributed. Rather, it reveals a pattern where most individuals possess lower-than-average ownership values, while a smaller yet significant group reports much higher ownership levels.

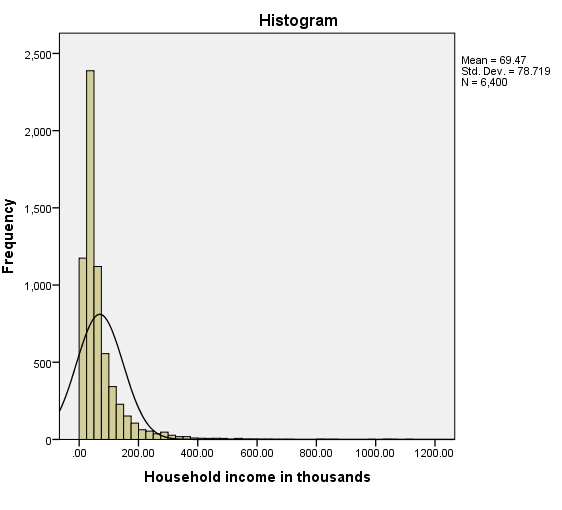
To visually confirm this observation, a histogram would serve as an excellent tool. By plotting the ownership values on a histogram, the distribution's shape would become more apparent. A positively skewed distribution would exhibit a longer tail towards the higher values, reinforcing the statistical inference drawn from the mean and median values.

Thus, the combination of statistical analysis and visual representation through a histogram provides a comprehensive understanding of the dataset's distribution. This approach allows for a nuanced exploration of the ownership patterns, highlighting the existence of a minority with significantly higher ownership levels amidst a predominantly lower ownership majority.

**Right Skewed Distribution(positive)**



**5.0. Histograms for Scale Variables**



The histogram generated from the dataset vividly illustrates the distribution of personal electronics ownership values among the surveyed population. A striking observation is the clustering of most cases towards the lower end of the scale, with a notable concentration falling below the 100,000 marks. This concentration of values paints a clear picture of the predominant ownership patterns within the dataset, indicating a prevalent trend towards lower ownership levels among the respondents.

However, the histogram also reveals a handful of outlier cases at the higher end of the scale, particularly in the 600,000 range and beyond. These extreme values, while few, exert a significant influence on the mean due to their substantial magnitude. The mean, being sensitive to outliers, is notably pulled upwards by these extreme cases, reflecting the impact of these high values on the dataset's average ownership.

In contrast, the median remains largely unaffected by these outliers. As a robust measure of central tendency, the median disregards extreme values and instead focuses on the middle point of the dataset. In this example, the median serves as a more reliable indicator of the central ownership value. Its resilience against the influence of outliers makes it a stable measure, accurately representing the typical ownership experience within the surveyed population.

This discrepancy between the mean and median underscores the importance of considering the dataset's distribution when selecting an appropriate measure of central tendency. In cases where outliers significantly skew the mean, the median emerges as a more suitable choice for capturing the typical or central ownership value. Thus, the histogram not only visually confirms the positive skewness of the distribution but also highlights the practical implications for selecting the most appropriate measure of central tendency within this dataset.

**6.0. Crosstabulation**

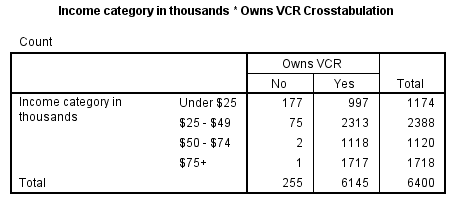
To delve deeper into the factors influencing personal electronics purchases, simple crosstabulation emerges as a powerful analytical tool. This method allows for a meticulous examination of the relationship between two variables—in this case, income levels and VCR ownership.

Through crosstabulation, the dataset is organized into a contingency table, presenting a clear overview of how VCR ownership varies across different income brackets. The income variable, acting as the focal point, serves to segment the surveyed population based on their income levels. Meanwhile, the VCR ownership variable provides insights into the prevalence of VCRs within each income group.

By cross-referencing these variables, the crosstabulation uncovers nuanced patterns and correlations. It reveals whether there exists a significant association between income levels and the likelihood of owning a VCR. This analysis goes beyond mere statistics, shedding light on the socio-economic dynamics that might influence consumer behaviours towards VCR purchases.

For instance, the crosstabulation might reveal that individuals in higher income brackets exhibit a higher propensity to own VCRs, perhaps due to disposable income, lifestyle preferences, or generational factors. Conversely, lower income groups might demonstrate a lower prevalence of VCR ownership, hinting at potential affordability constraints or evolving technology preferences.

Ultimately, this simple yet insightful analysis through crosstabulation allows researchers and analysts to glean actionable insights. It enables the identification of target demographics, the formulation of tailored marketing strategies, or even the exploration of broader socio-economic trends in consumer electronics adoption. By leveraging the relationship between income levels and VCR ownership, businesses and policymakers alike can make informed decisions to better serve their target audiences.



As we delve into the crosstabulation table detailing the relationship between income levels and VCR ownership, each cell represents the count or number of cases falling within a specific combination of values. For instance, we observe that 2,313 individuals within the income range of $25,000–$49,000 report owning VCRs.

Upon initial inspection, none of the numbers within this table immediately stand out as indicative of a clear relationship between the variables. The counts distributed across the various income brackets do not demonstrate a distinct pattern that would suggest a straightforward association between income levels and VCR ownership.

This lack of an obvious trend or relationship highlights the complexity of consumer behaviours and purchase decisions. While it might be tempting to expect a linear progression—such as higher incomes correlating with higher rates of VCR ownership—the reality is often more nuanced.

Several factors could contribute to this seemingly disparate distribution of VCR ownership across income brackets. For instance, cultural preferences, generational influences, technological advancements, or even regional differences might all play a role in shaping these patterns.

Despite the absence of a glaring trend, this crosstabulation table remains a valuable tool for analysis. It prompts further exploration into the nuances of consumer behaviours within different income groups. Researchers might delve deeper into specific subsets of the data, conduct additional statistical tests, or explore qualitative insights to uncover hidden relationships or factors at play.

By leveraging the information within this crosstabulation, analysts can refine their understanding of the intricate interplay between income levels and VCR ownership. This nuanced approach lays the groundwork for more targeted marketing strategies, product development initiatives, or policy decisions tailored to the diverse preferences and behaviours of the consumer base.

**7.0. Counts vs. Percentages**

Analysing a crosstabulation solely based on the raw counts in each cell can often lead to misinterpretations, especially when dealing with varying sample sizes across categories. For instance, observing that there are more than twice as many VCR owners in the $25,000–$49,000 income category compared to the under $25,000 category might not provide meaningful insights on its own. This is because there are also more than twice as many individuals in the $25,000–$49,000 income bracket.

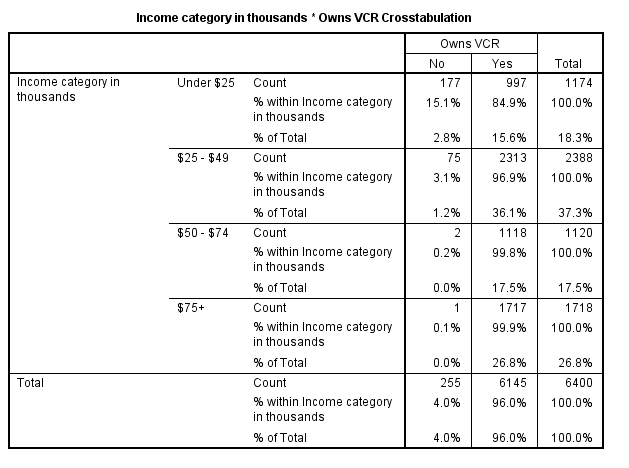
To address this issue and gain a clearer understanding of the relationship between income levels and VCR ownership, converting the counts into percentages proves to be invaluable. By comparing percentages within each income category, a more standardized and equitable comparison emerges. This approach allows for a fair assessment of ownership rates relative to the size of each income group.

Upon converting the counts into percentages, a more nuanced and revealing pattern begins to unfold. The crosstabulation table now paints a clearer picture, showcasing how the percentage of individuals owning VCRs rises as income categories ascend.

For example, the table might reveal that while the $25,000–$49,000 income bracket has more VCR owners in absolute terms, the percentage of VCR owners within this group is lower than in higher income brackets. Conversely, the under $25,000 income category might exhibit a higher percentage of VCR ownership despite its smaller absolute count, highlighting a potentially stronger association between lower income levels and VCR ownership.

This shift from counts to percentages offers a more standardized and comparative view, allowing analysts to discern meaningful trends and patterns. It aids in uncovering subtle relationships that might not be immediately apparent when focusing solely on raw counts.

By embracing this approach, the analysis gains depth and clarity, empowering researchers to draw informed conclusions about the impact of income levels on VCR ownership. This nuanced understanding can then inform strategic decisions, marketing campaigns, or policy initiatives tailored to the diverse consumer landscape across income groups.



**8.0. Significance Testing for Crosstabulations.**

At the core of crosstabulation analysis lies the aim to illuminate the relationship between two variables. In the context of the current study, the focus is on understanding how income levels influence VCR ownership. However, the mere observation of counts or percentages within a crosstabulation table might not suffice to ascertain the significance of this relationship. This is where statistical tests, such as the chi-square test, come into play.

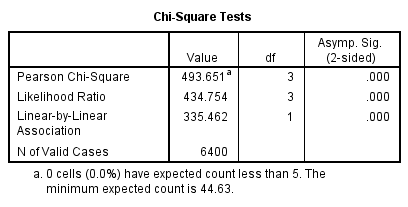
The chi-square test serves as a robust tool to determine whether the observed relationship between two cross-tabulated variables is statistically significant or merely the result of random chance. This statistical test assesses the difference between the observed frequencies in the crosstabulation table and the frequencies that would be expected if there were no true relationship between the variables.

One notable advantage of the chi-square test is its versatility and applicability to various types of data. Whether dealing with nominal, ordinal, or interval-level data, the chi-square test remains a reliable method for assessing associations between variables.

By conducting a chi-square test on the crosstabulated data of income levels and VCR ownership, analysts can ascertain whether the observed differences in ownership rates across income categories are statistically meaningful. A significant chi-square result indicates that the relationship between income levels and VCR ownership is unlikely to have occurred by random chance alone. Instead, it suggests a legitimate association between these variables.

This statistical validation offers crucial insights for researchers, marketers, or policymakers. It confirms whether income levels indeed play a significant role in determining VCR ownership patterns within the studied population. Armed with this knowledge, stakeholders can make informed decisions, tailor marketing strategies, allocate resources efficiently, or design interventions to better cater to the preferences and needs of specific income groups.

In essence, the chi-square test serves as a powerful ally in the realm of crosstabulation analysis, enabling researchers to move beyond descriptive statistics and uncover meaningful relationships between variables. Its versatility and ability to handle various types of data make it a go-to method for determining the significance of observed associations in diverse research contexts.



Within the realm of statistical analysis, the significance value (often denoted as Asymp. Sig.) holds the key to unlocking the meaning behind the relationship between two variables, as revealed through tests like the chi-square test. This value provides a measure of the probability that the observed relationship between variables is due to random chance alone.

In the context of the current study examining the relationship between income levels and VCR ownership, the significance value obtained from the chi-square test is a critical piece of information. A lower significance value indicates a stronger case against the null hypothesis, suggesting that the observed relationship is unlikely to be a product of random variation.

Upon conducting the chi-square test, the significance value in this instance is so low that it is displayed as .000. This tiny value essentially signifies that the probability of the observed relationship between income levels and VCR ownership occurring by chance alone is extremely unlikely.

In practical terms, this result signifies a clear and statistically significant relationship between income levels and VCR ownership within the studied population. It provides robust evidence that individuals' income levels indeed play a role in determining whether they are likely to own a VCR.

With a significance value of .000, we can confidently conclude that the two variables—income levels and VCR ownership—are related. This statistical validation reinforces the earlier observations gleaned from the crosstabulation and percentage analysis.

This newfound clarity opens doors for deeper insights and informed decision-making. Researchers, marketers, or policymakers can now rely on this concrete evidence to develop targeted strategies, allocate resources effectively, or tailor interventions to cater to specific income groups' needs and preferences.

In summary, the remarkably low significance value obtained from the chi-square test serves as a beacon of statistical strength, affirming the significant relationship between income levels and VCR ownership. This insight, backed by robust statistical analysis, empowers stakeholders to make data-driven decisions with confidence and precision.

Top of Form

**8.0. Control variable.**

In statistical analysis, the variable often referred to as the "control variable" plays a pivotal role in understanding the nuanced relationships between the main variables of interest. Sometimes also known as a "covariate," this variable helps researchers delve deeper into the effects of the primary variables under investigation.

The control variable acts as a moderator, revealing how the relationship between the row and column variables changes or varies when the effects of this third variable are considered. By "controlling" for the influence of the control variable, analysts can isolate and examine the direct relationship between the primary variables of interest.

This approach allows researchers to discern whether the observed relationship between the row and column variables holds true even when the effects of the control variable are considered. Essentially, the control variable serves as a lens through which to view the interaction effects between the primary variables, offering insights into how external factors might impact this relationship.

For instance, in the context of the analysis on income levels and VCR ownership, a potential control variable could be demographic factors such as age or geographic location. By controlling for the effects of age, researchers can ascertain whether the relationship between income levels and VCR ownership remains consistent across different age groups.

In essence, the control variable provides a means to refine and deepen the analysis, offering a more nuanced understanding of the interplay between variables. By accounting for potential confounding factors or external influences, researchers can better isolate and interpret the true relationship of interest.

Therefore, the inclusion of a control variable in the analysis enriches the study by revealing how the primary variables interact in the presence of this moderating factor. It allows for a more sophisticated exploration of the underlying dynamics, providing valuable insights that might have otherwise remained obscured.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Income category in thousands \* Owns VCR \* Level of education Crosstabulation** | | | | | | |
| Level of education | | | | Owns VCR | | Total |
| No | Yes |
| Did not complete high school | Income category in thousands | Under $25 | Count | 46 | 276 | 322 |
| % of Total | 3.3% | 19.9% | 23.2% |
| $25 - $49 | Count | 12 | 525 | 537 |
| % of Total | 0.9% | 37.8% | 38.6% |
| $50 - $74 | Count | 0 | 224 | 224 |
| % of Total | 0.0% | 16.1% | 16.1% |
| $75+ | Count | 0 | 307 | 307 |
| % of Total | 0.0% | 22.1% | 22.1% |
| Total | | Count | 58 | 1332 | 1390 |
| % of Total | 4.2% | 95.8% | 100.0% |
| High school degree | Income category in thousands | Under $25 | Count | 50 | 328 | 378 |
| % of Total | 2.6% | 16.9% | 19.5% |
| $25 - $49 | Count | 26 | 704 | 730 |
| % of Total | 1.3% | 36.4% | 37.7% |
| $50 - $74 | Count | 1 | 325 | 326 |
| % of Total | 0.1% | 16.8% | 16.8% |
| $75+ | Count | 0 | 502 | 502 |
| % of Total | 0.0% | 25.9% | 25.9% |
| Total | | Count | 77 | 1859 | 1936 |
| % of Total | 4.0% | 96.0% | 100.0% |
| Some college | Income category in thousands | Under $25 | Count | 45 | 196 | 241 |
| % of Total | 3.3% | 14.4% | 17.7% |
| $25 - $49 | Count | 16 | 495 | 511 |
| % of Total | 1.2% | 36.4% | 37.6% |
| $50 - $74 | Count | 0 | 248 | 248 |
| % of Total | 0.0% | 18.2% | 18.2% |
| $75+ | Count | 0 | 360 | 360 |
| % of Total | 0.0% | 26.5% | 26.5% |
| Total | | Count | 61 | 1299 | 1360 |
| % of Total | 4.5% | 95.5% | 100.0% |
| College degree | Income category in thousands | Under $25 | Count | 32 | 164 | 196 |
| % of Total | 2.4% | 12.1% | 14.5% |
| $25 - $49 | Count | 15 | 475 | 490 |
| % of Total | 1.1% | 35.1% | 36.2% |
| $50 - $74 | Count | 1 | 258 | 259 |
| % of Total | 0.1% | 19.0% | 19.1% |
| $75+ | Count | 0 | 410 | 410 |
| % of Total | 0.0% | 30.3% | 30.3% |
| Total | | Count | 48 | 1307 | 1355 |
| % of Total | 3.5% | 96.5% | 100.0% |
| Post-undergraduate degree | Income category in thousands | Under $25 | Count | 4 | 33 | 37 |
| % of Total | 1.1% | 9.2% | 10.3% |
| $25 - $49 | Count | 6 | 114 | 120 |
| % of Total | 1.7% | 31.8% | 33.4% |
| $50 - $74 | Count | 0 | 63 | 63 |
| % of Total | 0.0% | 17.5% | 17.5% |
| $75+ | Count | 1 | 138 | 139 |
| % of Total | 0.3% | 38.4% | 38.7% |
| Total | | Count | 11 | 348 | 359 |
| % of Total | 3.1% | 96.9% | 100.0% |
| Total | Income category in thousands | Under $25 | Count | 177 | 997 | 1174 |
| % of Total | 2.8% | 15.6% | 18.3% |
| $25 - $49 | Count | 75 | 2313 | 2388 |
| % of Total | 1.2% | 36.1% | 37.3% |
| $50 - $74 | Count | 2 | 1118 | 1120 |
| % of Total | 0.0% | 17.5% | 17.5% |
| $75+ | Count | 1 | 1717 | 1718 |
| % of Total | 0.0% | 26.8% | 26.8% |
| Total | | Count | 255 | 6145 | 6400 |
| % of Total | 4.0% | 96.0% | 100.0% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Chi-Square Tests** | | | | |
| Level of education | | Value | df | Asymp. Sig. (2-sided) |
| Did not complete high school | Pearson Chi-Square | 110.526b | 3 | .000 |
| Likelihood Ratio | 102.962 | 3 | .000 |
| Linear-by-Linear Association | 72.021 | 1 | .000 |
| N of Valid Cases | 1390 |  |  |
| High school degree | Pearson Chi-Square | 117.315c | 3 | .000 |
| Likelihood Ratio | 114.069 | 3 | .000 |
| Linear-by-Linear Association | 88.002 | 1 | .000 |
| N of Valid Cases | 1936 |  |  |
| Some college | Pearson Chi-Square | 143.961d | 3 | .000 |
| Likelihood Ratio | 123.569 | 3 | .000 |
| Linear-by-Linear Association | 91.202 | 1 | .000 |
| N of Valid Cases | 1360 |  |  |
| College degree | Pearson Chi-Square | 116.688e | 3 | .000 |
| Likelihood Ratio | 93.259 | 3 | .000 |
| Linear-by-Linear Association | 74.647 | 1 | .000 |
| N of Valid Cases | 1355 |  |  |
| Post-undergraduate degree | Pearson Chi-Square | 13.554f | 3 | .004 |
| Likelihood Ratio | 13.485 | 3 | .004 |
| Linear-by-Linear Association | 10.784 | 1 | .001 |
| N of Valid Cases | 359 |  |  |
| Total | Pearson Chi-Square | 493.651a | 3 | .000 |
| Likelihood Ratio | 434.754 | 3 | .000 |
| Linear-by-Linear Association | 335.462 | 1 | .000 |
| N of Valid Cases | 6400 |  |  |
| a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 44.63. | | | | |
| b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.35. | | | | |
| c. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 12.97. | | | | |
| d. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.81. | | | | |
| e. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.94. | | | | |
| f. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 1.13. | | | | |

**9.0. Conclusion**

In conclusion, the analysis of the chi-square statistics table within the education categories reveals a notable relationship between income levels and VCR ownership. Typically, a significance value below 0.05 is considered "significant," indicating that the observed relationship is likely not due to random chance.

The results strongly suggest a significant association between income levels and VCR ownership within the surveyed population. Individuals with higher incomes are more likely to own VCRs, while those with lower incomes exhibit lower rates of ownership.

Moreover, the analysis unveils an underlying relationship between education level and VCR ownership. As education levels rise, there is a corresponding increase in income. This trend indicates that apparent relationships observed between income and other variables may, in fact, be influenced by educational differences.

The link between education, income, and VCR ownership underscores the complex interplay of socio-economic factors in consumer behaviors. Individuals with higher educational attainment are more likely to have higher incomes, which, in turn, enables them to afford and invest in technologies such as VCRs.

This nuanced understanding sheds light on the broader dynamics at play within the consumer landscape. It emphasizes the need to consider multiple factors when analyzing consumer behaviors and preferences, rather than viewing variables in isolation.

Moving forward, these findings hold valuable implications for market researchers, policymakers, and businesses. Tailoring marketing strategies, product offerings, or educational initiatives to specific income and education demographics can enhance effectiveness and target audiences more precisely.

In essence, the study highlights the intricate relationships between income, education, and consumer technology adoption. It underscores the importance of considering these multifaceted variables when seeking to understand and cater to diverse consumer needs and preferences within a dynamic socio-economic landscape.