|  |  |
| --- | --- |
| **MS5105 Exercise Number:** 3 | **Total Number of Pages:** 14 |
| **Student Number:** 19230487 | **Student Name:** Jayakarthi Boovendran |

|  |
| --- |
| **Exercise Number:** Ch 08 – Exercise 02A |
| **Dataset Used :** [DataSet1] G:\MS5105 - Statistical Techniques for Business Analytics\Assignments\3\ Ch 08 - Exercise 02A.sav |
| **Description of the Test**  The primary objective of the given exercise is to find the **correlation** between the **continuous variables**, non-work hours spent on computer and the social relationship satisfaction. Therefore, in this case, **Pearson correlation test** is performed to statistically assess the correlation between the below-mentioned continuous variables.  **Variables**  Continuous Variable 1 : Number of non-work hours spent on computer  Continuous Variable 2 : Social relationship satisfaction score (ASSI score)  **Hypotheses**  **H0 –** There is **no statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score)  **H1 -** There **is a statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score)  **Pre-test Checklist for Pearson Test**  Following are the pre-test conditions for correlation and regression using Pearson test,   1. Normality 2. Linearity 3. Homoscedasticity   **Pre-test Condition - 1 : Normality**  FREQUENCIES VARIABLES=comphrs assi  /FORMAT=NOTABLE  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM MEAN MEDIAN MODE  /HISTOGRAM NORMAL  /ORDER=ANALYSIS     * **Descriptive Statistics**  |  |  |  |  | | --- | --- | --- | --- | | **Statistics** | | | | |  | | comphrs | assi | | N | Valid | 20 | 20 | | Missing | 0 | 0 | | Mean | | 19.0375 | 59.10 | | Median | | 21.0000 | 56.50 | | Mode | | 21.00 | 53a | | Std. Deviation | | 6.08940 | 15.145 | | Variance | | 37.081 | 229.358 | | Range | | 23.00 | 65 | | Minimum | | 6.00 | 31 | | Maximum | | 29.00 | 96 | | a. Multiple modes exist. The smallest value is shown | | | |  * **Histogram with a normal curve for ASSI**      * **Histogram with a normal curve for Computer hours**     **Discussion on Normality**   * From the inspection of both the histograms it is evident that the distribution of ASSI and computer hours is reasonably normal * Therefore, the pre-test criterion of **Normality is satisfied**   **Pre-test Condition -2 : Linearity**  \* Chart Builder.  GGRAPH  /GRAPHDATASET NAME="graphdataset" VARIABLES=comphrs assi MISSING=LISTWISE REPORTMISSING=NO  /GRAPHSPEC SOURCE=INLINE  /FITLINE TOTAL=NO.  BEGIN GPL  SOURCE: s=userSource(id("graphdataset"))  DATA: comphrs=col(source(s), name("comphrs"))  DATA: assi=col(source(s), name("assi"))  GUIDE: axis(dim(1), label("comphrs"))  GUIDE: axis(dim(2), label("assi"))  GUIDE: text.title(label("Simple Scatter of assi by comphrs"))  ELEMENT: point(position(comphrs\*assi))  END GPL.   * **Scatter plot with regression line for the two continuous variables**     **Discussion on Linearity**   * The points in the scatter plot forms a relatively straight line. Furthermore, the regression line takes the middle of the road path through the cloud of points. * Therefore, the criterion of **Linearity is satisfied**.   **Pre-test Condition - 3 : Homoscedasticity**  \* Chart Builder.  GGRAPH  /GRAPHDATASET NAME="graphdataset" VARIABLES=comphrs assi MISSING=LISTWISE REPORTMISSING=NO  /GRAPHSPEC SOURCE=INLINE  /FITLINE TOTAL=NO.  BEGIN GPL  SOURCE: s=userSource(id("graphdataset"))  DATA: comphrs=col(source(s), name("comphrs"))  DATA: assi=col(source(s), name("assi"))  GUIDE: axis(dim(1), label("comphrs"))  GUIDE: axis(dim(2), label("assi"))  GUIDE: text.title(label("Simple Scatter of assi by comphrs"))  ELEMENT: point(position(comphrs\*assi))  END GPL.   * **Scatter plot with regression line for the two continuous variables**     **Discussion on Homoscedasticity**   * The density of points along the regression line is high in the middle and tapers off at the ends * Since the cloud of points are focused in the middle of the regression line, it is assumed that the criterion of **Homoscedasticity is** **satisfied**   **Documenting Results of Pre-test Conditions**   * Normality * Linearity * Homoscedasticity   Since all the required pre-test conditions are satisfied, we proceed with Pearson’s bivariate correlation test.  **Pearson’s Bivariate Test for Correlation**  CORRELATIONS  /VARIABLES=comphrs assi  /PRINT=TWOTAIL NOSIG  /MISSING=PAIRWISE    **Discussion**   * The Pearson correlation **(r) of .317** with a **p > 0.05** indicates **statistically insignificant positive correlation** between the no. of non-work hours spent on computers and the ASSI score. * Furthermore, the value of correlation (r) close to 0 indicates a **weak correlation by strength.** * Since r = +0.317 is a **positive correlation by direction**, the two continuous variables move in the same direction * Computer hours : ASSI score * Computer hours : ASSI score   **Hypothesis Resolution**   * **Accept => H0 –** There is **no statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) * **Reject => H1 -** There **is statistically significant a correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) |

|  |
| --- |
| **Discussion of the Test** |
| In order to find the correlation between the number of non-work hours spent on computer and the social relationship satisfaction score (ASSI score). We recorded the non-work hours of twenty individuals (**n=20**) The recorded computer hours had a **mean** (**µ**) of **19.037** and a **Standard deviation** (**σ**) of **6.089**. Each participant is then assessed for their social relationship satisfaction score using ASSI (0 to 80, with 0 being very low social satisfaction and 80 indicating a very high social satisfaction). The recorded ASSI score had a **mean** (**µ**) of **59.10** and a **Standard deviation** (**σ**) of **15.145**. In order to statistically assess the correlation between the above-mentioned continuous variables we perform Pearson correlation test.  To proceed with Pearson correlation test, the pre-test conditions of Normality, Linearity and Homoscedasticity should be met. From the visual interpretation of the histograms of both the continuous variables, it is assumed that the data is approximately normally distributed. Furthermore, on regression analysis, the cloud of points on the scatter plot graph forms a relatively straight line, and is also dense along the middle of the regression line. Thus satisfied the pre-test conditions of linearity and homoscedasticity. Since all the pre-test criteria for correlation test are met. We proceeded with Pearson bivariate correlation test and discovered a **statistically insignificant positive correlation** between the ASSI score and computer hours (**r=0.317** and p>α; where **p=0.173** and α=0.05). Since p > α, we **reject** **H1**and **accept H0**. Therefore, there **is no statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) |

|  |
| --- |
| **Exercise Number:** Ch 08 – Exercise 02B |
| **Dataset Used :** [DataSet1] G:\MS5105 - Statistical Techniques for Business Analytics\Assignments\3\ Ch 08 - Exercise 02B.sav |
| **Description of the Test:**  The primary objective of the given problem is to find the **correlation** between the **continuous variables**, non-work hours spent on computer and the social relationship satisfaction. Therefore, in this case, **Pearson correlation test** is performed to statistically assess the correlation between the below-mentioned continuous variables.  **Variables**  Continuous Variable 1 : Number of non-work hours spent on computer  Continuous Variable 2 : Social relationship satisfaction score (ASSI score)  **Hypotheses**  **H0 –** There is **no statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score)  **H1 -** There **is a statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score)  **Pre-test Checklist for Pearson Test**  Following are the pre-test conditions for correlation and regression using Pearson test,   1. Normality 2. Linearity 3. Homoscedasticity   **Pre-test Condition - 1 : Normality**  FREQUENCIES VARIABLES=comphrs assi  /FORMAT=NOTABLE  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM MEAN MEDIAN MODE  /HISTOGRAM NORMAL  /ORDER=ANALYSIS.     * **Descriptive Statistics**  |  |  |  |  | | --- | --- | --- | --- | | **Statistics** | | | | |  | | comphrs | assi | | N | Valid | 23 | 23 | | Missing | 0 | 0 | | Mean | | 19.0543 | 58.78 | | Median | | 18.7500 | 59.00 | | Mode | | 14.00 | 54a | | Std. Deviation | | 5.27158 | 16.014 | | Variance | | 27.790 | 256.451 | | Range | | 21.25 | 75 | | Minimum | | 8.00 | 21 | | Maximum | | 29.25 | 96 | | a. Multiple modes exist. The smallest value is shown | | | |  * **Histogram with a normal curve for ASSI**      * **Histogram with a normal curve for Computer hours**     **Discussion on Normality**   * Through the visual interpretation of both the histograms it is evident that ASSI score and computer hours is approximately normally distributed * Therefore, the pre-test criterion of **Normality is satisfied**   **Pre-test Condition - 2 : Linearity**  \* Chart Builder.  GGRAPH  /GRAPHDATASET NAME="graphdataset" VARIABLES=comphrs assi MISSING=LISTWISE REPORTMISSING=NO  /GRAPHSPEC SOURCE=INLINE  /FITLINE TOTAL=NO.  BEGIN GPL  SOURCE: s=userSource(id("graphdataset"))  DATA: comphrs=col(source(s), name("comphrs"))  DATA: assi=col(source(s), name("assi"))  GUIDE: axis(dim(1), label("comphrs"))  GUIDE: axis(dim(2), label("assi"))  GUIDE: text.title(label("Simple Scatter of assi by comphrs"))  ELEMENT: point(position(comphrs\*assi))  END GPL.    **Discussion on Linearity**   * The points in the scatter plot forms a relatively straight line. Furthermore, the regression line takes the middle of the road path through the cloud of points. * Therefore, the criterion of **Linearity is satisfied**.   **Pre-test Condition - 3 : Homoscedasticity**  \* Chart Builder.  GGRAPH  /GRAPHDATASET NAME="graphdataset" VARIABLES=comphrs assi MISSING=LISTWISE REPORTMISSING=NO  /GRAPHSPEC SOURCE=INLINE  /FITLINE TOTAL=NO.  BEGIN GPL  SOURCE: s=userSource(id("graphdataset"))  DATA: comphrs=col(source(s), name("comphrs"))  DATA: assi=col(source(s), name("assi"))  GUIDE: axis(dim(1), label("comphrs"))  GUIDE: axis(dim(2), label("assi"))  GUIDE: text.title(label("Simple Scatter of assi by comphrs"))  ELEMENT: point(position(comphrs\*assi))  END GPL.    **Discussion on Homoscedasticity**   * The density of points along the regression line is high in the middle and tapers off at the ends * Since the cloud of points are focused in the middle of the regression line, it is assumed that the criterion of **Homoscedasticity is satisfied**   **Documenting Results of Pre-test Conditions**   * Normality * Linearity * Homoscedasticity   Since all the required pre-test conditions are satisfied, we proceed with Pearson’s bivariate correlation test.  **Pearson’s Bivariate Test for Correlation**  CORRELATIONS  /VARIABLES=comphrs assi  /PRINT=TWOTAIL NOSIG  /MISSING=PAIRWISE    **Discussion**   * The Pearson correlation **(r) of -0.497** with a **p < 0.05** indicates **statistically significant negative correlation** between the no. of non-work hours spent on computers and the ASSI score. * The value of r near to 0 indicates a **weak correlation by strength** * Since r= -0.497 is a **negative correlation by direction**, the two continuous variables move in the opposite direction * Computer hours : ASSI score * Computer hours : ASSI score   **Hypothesis Resolution**   * **Reject => H0 –** There is **no statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) * **Accept => H1 -** There **is a statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) |

|  |
| --- |
| **Discussion of the Test** |
| In order to find the correlation between the number of non-work hours spent on computer and the social relationship satisfaction score (ASSI score). We recorded the non-work hours of twenty three individuals (**n=23**). The recorded computer hours had a **mean** (**µ**) of **19.054** and a **Standard deviation** (**σ**) of **5.271**. Each participant is then assessed for their social relationship satisfaction score using ASSI (0 to 80, with 0 being very low social satisfaction and 80 indicating a very high social satisfaction). The recorded ASSI score had a **mean** (**µ**) of **58.78** and a **Standard deviation** (**σ**) of **16.014**. In order to statistically assess the correlation between the above-mentioned continuous variables we perform Pearson correlation test.  To proceed with Pearson correlation test, the pre-test conditions of Normality, Linearity and Homoscedasticity should be met. From the visual interpretation of the histograms of both the continuous variables, it is assumed that the data is approximately normally distributed. Furthermore, on regression analysis, the cloud of points on the scatter plot graph forms a relatively straight line, and is also dense along the middle of the regression line. Thus satisfied the pre-test conditions of linearity and homoscedasticity. Since all the pre-test criteria for correlation test are met. We proceeded with Pearson bivariate correlation test and discovered a **statistically significant negative correlation** between the ASSI score and computer hours (**r=-0.497** and p>α; where **p=0.016** and α=0.05). Since p < α, we **reject** **H0**and **accept H1**. Therefore, there **is a statistically significant correlation** between non-work hours spent on computer and social relationship satisfaction (ASSI score) |