A Path Analysis of Citations Received by Patents:

Investigating the Drivers of Technology Transfer Outcomes

Malcolm S. Townes

Saint Louis University

Malcom,

This assignment is professionally done as always. The use of transformation seems to be appropriate in this case, and thank you for pp-plots which help me understand the relationship. I wrote some comments in the text below. This paper is presentable if not publishable. You might want to consider the same line of study for your dissertation by obtaining a similar but newer data.

Excellent work! A

HM

**Abstract**

Technology transfer is the transition of technology or intellectual property from one person or entity to another person or entity. Improving the transfer of technology derived from federally-funded research and development (R&D) to the private sector is a priority for the public policy of the United States of America (OMB, 2018). University technology transfer (UTT) is a subcategory of the broader technology transfer field. It focuses on the transfer of technology derived from research conducted at universities to the private sector. Identifying the drivers of technology transfer outcomes is an important topic for study. There are several potential benefits to developing predictive models describing technology transfer and understanding the factors associated with successful technology transfer. Such knowledge would be useful for managing technological innovation and identifying high potential technologies for further development (Choi, Jang, Jun & Park, 2015). This study used path analysis to investigate the potential drivers of the technology transfer process in an effort to assess whether the number of citations received by a patent is a useful and feasible measure of technology transfer outcomes. Insights from the broader field of technology transfer should be applicable to the narrower field of UTT.

Keywords: technology transfer, university technology transfer, technology commercialization, federally funded research and development, patents, patent citations, path analysis

**Introduction**

This study continues the investigation of how technology transfer success can be defined and measured that I began on Assignments 01 and 02 for SOC 6100 in the Fall 2018 semester. In this study, I conducted a path analysis to evaluate direct and indirect effects among various patent citation data variables in an effort to better understand the drivers of technology transfer outcomes. As in the previous analyses, I used patents issued by the United States Patent and Trademark Office (USPTO) as a proxy for units of technology and the number of citations a given U.S. patent receives from other U.S. patents as a measure of technology transfer.

**Literature Review**

Several researchers have investigated the relationship between various factors and technology transfer outcomes including patent citations and other patent data. In a cursory review of the recent literature published in the last five years, I found a few studies that used path analysis or structural equation models (SEM) to examine drivers and processes of technology transfer. However, none of them explicitly focused on patent citation data. Vagnani & Volpe (2017) integrated SEM with a meta-analysis to study the associations between various attributes of innovation, manager behaviors, and organizational decisions regarding the adoption of innovation. They found that innovation attributes have indirect effects on the innovation adoption decisions of organizations via the behavioral preferences of managers. Raut, Priyadarshinee, Garda & Jha (2018) incorporated SEM in a hybrid three-stage approach to analyze factors that influence the adoption of cloud computing technology by private organizations in India. Yan & Yu (2016) used a path-based method to examine the structure of time-dependent, discipline-level citation networks. Markman & Gionidis (2016) used hierarchical regression analysis to isolate various drivers of technology transfer. They found that several factors were associated with patent license income and university spin-out company formation. However, their analysis indicated that patent data variables were either not significant or did not improve their existing model. Ji, Lim, & Park (2016) used patent citation data to identify potential cases of technology transfer. Sharma (2017) conducted a survey of patent citation analysis and presented a methodology for generating patent citation networks. Park, Yoon & Kim (2013) used a function-based patent analysis to identify potential opportunities to apply technologies in various fields.

**Data and Methods**

**Data Sources**

This study uses a subset of 2,000 observations taken patent data obtained from the National Bureau of Economic Research (NBER) website. The source data contains both original and constructed variables. The data file included all utility patents granted in the U.S. from January 1, 1963 to December 30, 1999 listed in the Technology Assessment and Forecast (TAF) database of the USPTO. The source file contained data on 2,923,922 patents across 23 variables.

Table 1 details the original USPTO variables of the source data and explanations of their meanings. PATENT indicates the number assigned by the USPTO to the allowed patent. GYEAR is the year the USPTO allowed the patent. APPYEAR is the year the patent application was submitted to the USPTO. GDATE is the number of weeks elapsed since January 1, 1960 to the date the USPTO allowed the patent. COUNTRY is the country of citizenship for the first inventor listed on the patent application. POSTATE is the state of residency for the first inventor listed on the patent application. ASSIGNEE indicates to whom the patent is assigned and is unique to each assignee. ASSCODE indicates the type of assignee. CLAIMS is the number of independent and dependent claims listed on the patent. NCLASS indicates the broad classification for the patent.

Table 1

Original USPTO Variables of Source Data

| Variable | Variable Type | Extended Name | Description |
| --- | --- | --- | --- |
| PATENT | Numeric  Nominal | Patent Number | The number assigned to the allowed patent by the USPTO.  Takes on values integer values between 3070801 and 6009554. |
| GYEAR | Numeric  Interval | Grant Year | The year the USPTO allowed the patent.  Takes on integer values between 1963 – 1999. |
| GDATE | Numeric  Interval | Grant Date | The date the USPTO allowed the patent expressed in terms of the number of weeks elapsed since  January 1, 1960.  Takes on integer values between 156 and 2,028. |
| APPYEAR | Numeric  Interval | Application Year | The year the patent application was submitted to the USPTO.  Takes on integer values between 1963 – 1999. |
| COUNTRY | Character  Nominal | Country of First Inventor | The country of citizenship for the first inventor listed on the patent application.  Takes on values of two character string data. |
| POSTATE | Character  Nominal | State of First Inventor (US) | The state of residency for the first inventor listed on the patent application if the country of citizenship is the United States of America.  Takes on values of two character string data. |
| ASSIGNEE | Numeric  Nominal | Assignee Identifier | Unique identifier for the assignee of the patent.  Takes on values from 10950 to 99550. |
| ASSCODE | Numeric  Nominal | Assignee Code | A one character code categorizing the type of assignee.  Takes on values from 1 to 7. |
| CLAIMS | Numeric  Interval | Number of Claims | Number of independent and dependent claims on the patent.  Takes on integer values from 1 to . |
| NCLASS | Numeric  Nominal | Main Patent Class | A code that categorizes the patent into one of several broad classifications.  Takes on integer values from 1 to 800. |

Table 2 provides information about the source data constructed variables and explanations of their meanings. CAT is a higher-level classification of the main patent class. SUBCAT is a sub-category of the main patent class. CMADE indicates the number of citations made by the patent. CRECEIVE indicates the number of citations in other patents that reference the patent. RATIOCIT is the ratio of the number of citations made by all patents granted since 1963 to the total number of citations made by the patent. GENERAL is a measure of how broad the influence of a patent spans across fields. ORIGINAL is a measure of the originality of the patent. FWDAPLAG measures forward citations lag. BCKGTLAG measures backward citations lag. SELFCTUB is the upper bound for the share of citations the patent makes to other patents assigned to the same assignee (i.e., self-citations made). SELFCTLB is the lower bound for the share of citations the patent makes to other patents assigned to the same assignee. SECUPBD is the upper bound for the share of citations the patent receives from other patents assigned to the same assignee (i.e., self-citations received). SECDLWBD is the lower bound for the share of citations the patent receives from other patents assigned to the same assignee.

Table 2

Source Data Constructed Variables

| Variable | Variable Type | Extended Name | Description |
| --- | --- | --- | --- |
| CAT | Numeric  Nominal | Technological Category | A higher-level classification of the Main Patent Class.  Takes on integer values from 1 to 6. |
| SUBCAT | Numeric  Nominal | Technological Sub-category | The sub-category of the primary technological category to which the patent is assigned.  Takes on integer values from 1 to 69. |
| CMADE | Numeric  Interval | Number of Citations Made | The number of citations made by the patent.  Takes on integer values from 1 to . |
| CRECEIVE | Numeric  Interval | No. of Citations Received | The number of citations in other patents that reference the patent.  Takes on integer values from 1 to . |
| RATIOCIT | Numeric  Ratio | Percent of Citations Made to Patents Granted Since 1963 | The ratio of the number of citations made by all patents granted since 1963 to the total number of citations made by the particular patent.  Takes on continuous values between 0 and 1. |
| GENERAL | Numeric  Ratio | Measure of Generality | A measure of how broad the influence of a patent spans across fields as determined by the number of different fields of all patents that cite the patent of interest.  Calculated as the following:  Generalityi = 1 - , where *sij* denotes the percentage of citations received by patent *i* that belong to patent class *j*, out of *ni* patent classes.  Takes on continuous values between 0 and 1. |
| ORIGINAL | Numeric  Ratio | Measure of Originality | A measure of the originality of a patent as determined by the number of different fields for all patents cited by the patent of interest.  Calculated as the following:  Originalityi = 1 - , where *sij* denotes the percentage of citations made by patent *i* that belong to patent class *j*, out of *ni* patent classes.  Takes on continuous values between 0 and 1. |
| FWDAPLAG | Numeric  Ratio | Mean Forward Citation Lag | The mean time difference between the application or grant date of the patent and that of the other patents citing this patent.  Takes on continuous values between 0 and 1. |
| BCKGTLAG | Numeric  Ratio | Mean Backward Citation Lag | The mean time difference between the application or grant date of the patent and that of the patents it cites.  Takes on continuous values between 0 and 1. |
| SELFCTUB | Numeric  Ratio | Share of Self-Citations Made – Upper Bound | The number of citations made by the patent to other patents with the same assignee divided by the total number of citations made by all patents with assignee codes.  Takes on continuous values between 0 and 1. |
| SELFCTLB | Numeric  Ratio | Share of Self-Citations Made – Lower Bound | The number of citations made by the patent to other patents with the same assignee divided by the total number of citations made by all patents.  Takes on continuous values between 0 and 1. |
| SECUPBD | Numeric  Ratio | Share of Self-Citations Received – Upper Bound | The number of citations received by the patent from other patents with the same assignee divided by the total number of citations received by all patents with assignee codes.  Takes on continuous values between 0 and 1. |
| SECDLWBD | Numeric  Ratio | Share of Self-Citations Received – Lower Bound | The number of citations received by the patent from other patents with the same assignee divided by the total number of citations received by all patents.  Takes on continuous values between 0 and 1. |

**Data Selection and Modification**

For this study, I only used the GYEAR, CLAIMS, CRECEIVE, RATIOCIT, GENERAL, and ORIGINAL variables in the path analysis. Based on the results of my previous analyses, I made several modifications to the data and incorporated several previous observations into the initial theoretical path analysis model. I removed the following variables because of high multicollinearity: APPYEAR, BCKGTLAG, FWDAPLAG, SELFCTLB, and SECDLWBD. Based on a scatter plot of the CRECEIVE variable against the CLAIMS variable, I removed observations with CLAIMS greater than 90 claims and CRECEIVE greater than 40 citations received as outliers. This resulted in 42 outlier observations being removed from the analysis for a final sample count of 1,958 observations. Additionally, I created a new variable (CRECEIVEln) using the Transform > Compute Variable function of SPSS Statistics 25. The CRECEIVEln variable is the natural logarithm transformation of the CRECEIVE variable. I performed a transformation on the CRECEIVE variable because the data was skewed to the right (i.e., positively skewed) based on a visual inspection of a histogram. I chose a natural logarithm transformation because it appeared to bring out potential linear relationships between the CRECEIVE variable and primary independent variable (IV) of interest (CLAIMS) better than other transformations that I considered, which included base 10 logarithm and reciprocal transformations. The CRECEIVEln variable is what I used as the dependent variable (DV) of interest in the analysis.

**Theoretical Model**

Figure 1 shows the theoretical path model that I developed from logical consideration of the relationships among the variables.



Figure 1. Logically Derived Theoretical Path Model

The theoretical path model uses the CRECEIVEln variable as the primary DV of interest. Sub-model 1 uses GENERAL as the DV and ORIGINAL as the IV. It is likely that the originality of a patent will influence whether or not it will have broad applicability. The more original the patent, the more likely that other innovators in various fields will identify applications of the technology over time. Patents that rank low in originality are likely to be specific or specialized to a narrower range of applications within closely related fields.

Sub-model 2 uses CLAIMS as the DV and ORIGINAL, GENERAL, GYEAR, and RATIOCIT as the IVs. The claims of a patent define the scope of the subject that it asserts to be novel, nonobvious, and useful. Patents that rank higher in originality are likely to generate more claims because they stake out new innovation territory. Patents the rank higher in generality are likely to generate more claims because the scope of their applicability. In general, patents are likely to have more claims as the grant year increases because of the temporal nature of advances in sciences and the cumulative effects of scientific knowledge. I suspect that as the ratio of the number of citations made by all patents granted since 1960 to the total number of citations made by a particular patent increases the number of claims for a patent will increase because of the general increase in scientific knowledge due to network effects.

Sub-model 3 uses CRECEIVEln as the DV and ORIGINAL, GENERAL, CLAIMS, GYEAR, and RATIOCIT as the IVs. Previous analysis indicated an inverse relationship between the originality of a patent and the number of citations it received. This may be because the full capabilities of highly original patents are less readily apparent than patents that rank lower on originality. Patents that rank high in generality probably receive higher numbers of citations because the broader scope of their applicability creates more opportunities to be cited. Likewise, patents that have more claims probably have more opportunities to be cited than patents with fewer claims. In general, patents are likely to receive more citations over time because scientific knowledge accumulates and spreads over time. I suspect that as the ratio of the number of citations made by all patents granted since 1960 to the total number of citations made by a particular patent increases the number of citations a patent receives will likely increase because of network effects.

**Analysis**

I used IBM SPSS Statistics 25 to analyze the theoretical path model. I used the Analyze > Regression > Linear function to prepare regression analyses for each sub-model. The options I selected included model fit, R square change, part and partial correlations, and collinearity diagnostics for the regression statistics; estimates, a confidence level of 95 percent, and covariance matrix for the regression coefficients; and Durbin-Watson, casewise diagnostics for outliers beyond 3 standard deviations for the residuals. For each model I used the enter method. The complete SPSS Statistics 25 output file for the analysis is shown in Appendix A.

**Results and Findings**

Figure 2 shows theoretical path model with standardized coefficients and p-values from the various regression analyses.



Figure 2. Path Model with Standardized Coefficients and P-Values

The associations between GENERAL and ORIGINAL and of ORIGINAL, GENERAL, GYEAR, with CLAIMS were all significant. GENERAL, ORIGINAL, and GYEAR all had indirect effects on CRECEIVEln through CLAIMS. I initially assumed that RATIOCIT would have an indirect effect on CRECEIVEln through the CLAIMS variable, but the p-value of the association was just above the threshold for significance at the 0.05 level. Interestingly, GYEAR had an inverse relationship with CRECEIVEln, which was counter to my original assumption about the relationship between these two variables.

Figure 3 shows the final path model with standardized coefficients and p-values. I removed the RATIOCIT variable from the final model because it was not significant. The   
p-value was 0.055, which was just above the threshold for significance at the 0.05 level. After removing the RATIOCIT variable, I re-calculated the linear regression for CLAIMS as the DV with ORIGINAL, GENERAL, and GYEAR as the IVs. However, the standardized coefficients changed only slightly and did not change the indirect effects calculations for these variables.

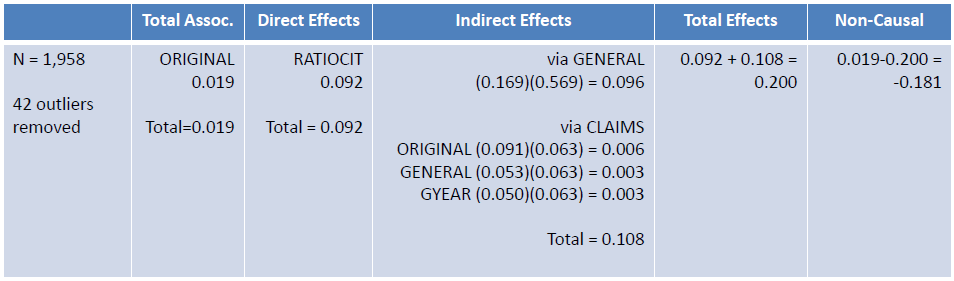


Figure 3. Final Path Model with Standardized Coefficients and P-Values

Table 1 summarizes the direct, indirect, and total effects for the final path model. Total effects were 0.200 while non-causal effects were -0.181. The indirect effects on CRECEIVEln through the CLAIMS variables were relatively miniscule.

Table 1

Path Model Direct and Indirect Effects



**Discussion**

**Policy Implications**

There are several possible policy implications of this study. The analysis provides insight into a topic that is of considerable interest to policymakers. It provides information to help policymakers better understand the drivers of the technology transfer outcomes and identify possible factors that should be considered when forming public policy regarding technology transfer. As such, this study may help policymakers formulate public policy that leads to greater levels of transfer of federally-funded research to the private sector.

**Limitations of the Analysis**

As with any research project or study, this analysis has limitations. Since this analysis was focused on patent data for a five year period from 1995 to 1999, findings based on the data may not be relevant to time frames before or after this period. Additionally, there is a truncation effect in the data. Patents issued in the earliest part of the study period have the potential of receiving citations from patents over a longer period than patents issued in the latter part of the study period. This could potentially be contributing to the skewness observed in the sample distribution.

**Future Study**

There are several opportunities to improve upon and extend the analysis presented in this paper. To begin, it might prove useful to secure more recent data and to examine a subset of data buffered by at least 5 years of data on both sides of the period of study to minimize truncation effects. By merging the data with data containing information about patent assignees, it should be possible to further subset the patent data specifically for university technologies. It might also be useful to introduce the category and subcategory of patents into the analysis to determine if the type of technology is associated with technology transfer outcomes. Finally, comparing an analysis of various baseline path analysis models could help identify a more optimized path analysis model.

**Conclusion**

In this paper, I have continued to explore an alternative conceptualization of technology transfer and an approach to measuring technology transfer based on patent citations received. Using patent data, I conducted a path analysis using a variable measuring patent citations received as the dependent variable and measures of the patent’s originality, generality, claims, the year the patent was granted, and citations ratio as independent variables. The path analysis model that I developed indicated that the generality of a patent was the overwhelming contributor of effects on the number of citations the patent received. Moreover, the study revealed an inverse relationship between the year a patent was granted and the number of citations the patent received.

Finally, I identify potential policy implications for studying this topic. It provides information to help policymakers identify factors that they should potentially consider when forming public policy regarding technology transfer. As such, this study may help policymakers craft public policy that increases the amount of federally-funded research that is transfer to the marketplace to benefit the public interest.

References

Aldieri, L. ( 1 ), & Vinci, C. P. ( 2 ). (n.d.). Technological Spillovers through a Patent Citation Analysis. *International Journal of Innovation Management*, 20(2). https://doi.org/10.1142/S1363919616500286

Chávez, G. A. G., & Víquez, H. G. (2015). Patterns of knowledge flow from industrialized to Latin American and Asian countries in the pharmaceutical industry: a patent citation analysis. *Contaduría y Administración*, 60 (Supplement 1), 31–56. https://doi.org/10.1016/j.cya.2015.08.008

Choi, J., Jang, D., Jun, S., & Park, S. (2015). A Predictive Model of Technology Transfer Using Patent Analysis. *Sustainability* (2071-1050), 7(12), 16175

Markman, G. D., Gianiodis, P. T., & Phan, P. H. (2009). Supply-Side Innovation and Technology Commercialization. *Journal of Management Studies*, 46(4), 625-649. doi:10.1111/j.1467-6486.2009.00835.x

Hall, B. H., Jaffe, A. B. and Trajtenberg, M. (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." *NBER Working Paper 8498*. Retrieved from http://www.nber.org/patents/

Ji, I., Lim, H., & Park, T.-Y. (2016). Exploring Potential Users of Patents for Technology Transfer: Utilizing Patent Citation Data. *Procedia Computer Science*, 91, 211–220. https://doi.org/10.1016/j.procs.2016.07.059

National Bureau of Economic Research. (2018). Patent data, including constructed variables [data file]. Retrieved from http://www.nber.org/patents/

Office of Management and Budget [OMB]. (2018). *The President's Management Agenda*. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2018/03/Presidents-Management-Agenda.pdf

Park, T.-Y., Lim, H., & Ji, I. (n.d.). Identifying potential users of technology for technology transfer using patent citation analysis: a case analysis of a Korean research institute. *SCIENTOMETRICS*, 116(3), 1541–1558. https://doi.org/10.1007/s11192-018-2792-9

Raut, R. D., Priyadarshinee, P., Gardas, B. B., & Jha, M. K. (2018). Analyzing the factors influencing cloud computing adoption using three stage hybrid SEM-ANN-ISM (SEANIS) approach. *Technological Forecasting & Social Change*, 134, 98–123. https://doi.org/10.1016/j.techfore.2018.05.020

Sharma, P., & Tripathi, R. C. (2017). Patent citation: A technique for measuring the knowledge flow of information and innovation. *World Patent Information*, 51, 31–42. https://doi.org/10.1016/j.wpi.2017.11.002

Vagnani, G., & Volpe, L. (2017). Innovation attributes and managers’ decisions about the adoption of innovations in organizations: A meta-analytical review. *International Journal of Innovation Studies*, 1, 107–133. https://doi.org/10.1016/j.ijis.2017.10.001

Yan, E., & Yu, Q. (2016). Using path-based approaches to examine the dynamic structure of discipline-level citation networks: 1997-2011. *Journal of the Association for Information Science & Technology*, 67(8), 1943–1955. https://doi.org/10.1002/asi.23516

Appendix A. IBM Statistics SPSS 25 Output

**Frequencies**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 16:39:40 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 2000 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on all cases with valid data. |
| Syntax | | FREQUENCIES VARIABLES=CRECEIVE CRECEIVEln CRECEIVElog10 CRECEIVErecip  /STATISTICS=STDDEV RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN SKEWNESS SESKEW KURTOSIS SEKURT  /HISTOGRAM NORMAL  /ORDER=ANALYSIS. |
| Resources | Processor Time | 00:00:07.36 |
| Elapsed Time | 00:00:03.49 |

[DataSet1] D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Statistics** | | | | | |
|  | | CRECEIVE | CRECEIVEln | CRECEIVElog10 | CRECEIVErecip |
| N | Valid | 2000 | 2000 | 2000 | 2000 |
| Missing | 0 | 0 | 0 | 0 |
| Mean | | 3.18 | .7788 | .3382 | .589921152058347 |
| Std. Error of Mean | | .096 | .01770 | .00769 | .007877009960755 |
| Median | | 2.00 | .6931 | .3010 | .500000000000000 |
| Std. Deviation | | 4.309 | .79147 | .34373 | .352270594633808 |
| Skewness | | 10.292 | .797 | .797 | .095 |
| Std. Error of Skewness | | .055 | .055 | .055 | .055 |
| Kurtosis | | 214.022 | .090 | .090 | -1.630 |
| Std. Error of Kurtosis | | .109 | .109 | .109 | .109 |
| Range | | 111 | 4.72 | 2.05 | .99107142857142860 |
| Minimum | | 1 | .00 | .00 | .00892857142857143 |
| Maximum | | 112 | 4.72 | 2.05 | 1.00000000000000000 |

**Frequency Table**

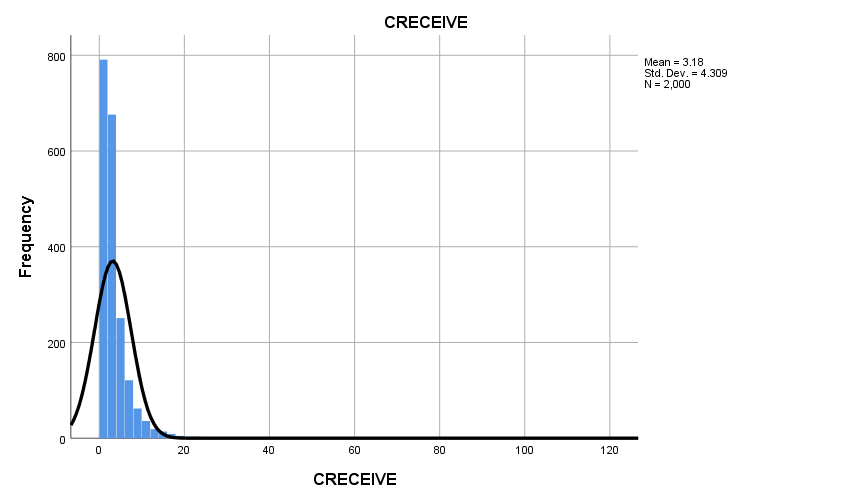
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CRECEIVE** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 1 | 791 | 39.6 | 39.6 | 39.6 |
| 2 | 433 | 21.7 | 21.7 | 61.2 |
| 3 | 243 | 12.2 | 12.2 | 73.4 |
| 4 | 146 | 7.3 | 7.3 | 80.7 |
| 5 | 105 | 5.3 | 5.3 | 85.9 |
| 6 | 75 | 3.8 | 3.8 | 89.6 |
| 7 | 46 | 2.3 | 2.3 | 92.0 |
| 8 | 36 | 1.8 | 1.8 | 93.8 |
| 9 | 26 | 1.3 | 1.3 | 95.1 |
| 10 | 22 | 1.1 | 1.1 | 96.2 |
| 11 | 14 | .7 | .7 | 96.9 |
| 12 | 10 | .5 | .5 | 97.4 |
| 13 | 9 | .4 | .4 | 97.8 |
| 14 | 10 | .5 | .5 | 98.3 |
| 15 | 4 | .2 | .2 | 98.5 |
| 16 | 4 | .2 | .2 | 98.7 |
| 17 | 5 | .3 | .3 | 99.0 |
| 18 | 5 | .3 | .3 | 99.2 |
| 20 | 1 | .1 | .1 | 99.3 |
| 21 | 2 | .1 | .1 | 99.4 |
| 22 | 1 | .1 | .1 | 99.4 |
| 23 | 2 | .1 | .1 | 99.5 |
| 25 | 1 | .1 | .1 | 99.6 |
| 26 | 2 | .1 | .1 | 99.7 |
| 30 | 1 | .1 | .1 | 99.7 |
| 31 | 1 | .1 | .1 | 99.8 |
| 32 | 1 | .1 | .1 | 99.8 |
| 33 | 1 | .1 | .1 | 99.9 |
| 38 | 1 | .1 | .1 | 99.9 |
| 42 | 1 | .1 | .1 | 100.0 |
| 112 | 1 | .1 | .1 | 100.0 |
| Total | 2000 | 100.0 | 100.0 |  |

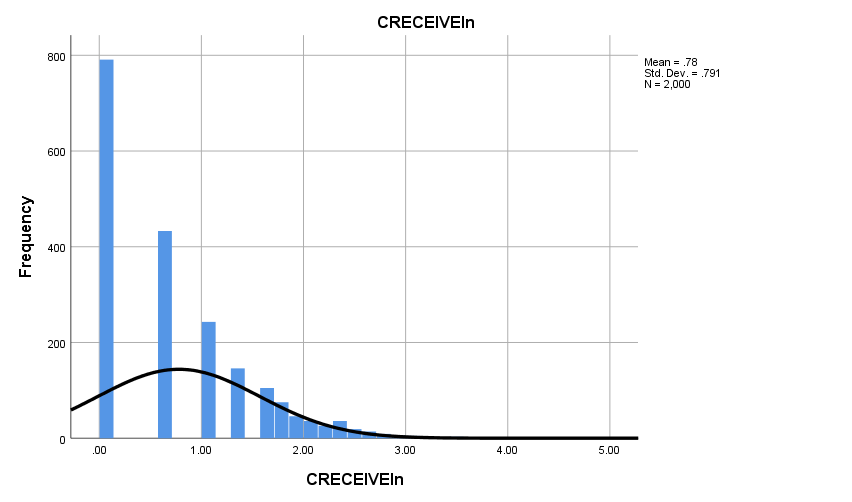
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CRECEIVEln** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | .00 | 791 | 39.6 | 39.6 | 39.6 |
| .69 | 433 | 21.7 | 21.7 | 61.2 |
| 1.10 | 243 | 12.2 | 12.2 | 73.4 |
| 1.39 | 146 | 7.3 | 7.3 | 80.7 |
| 1.61 | 105 | 5.3 | 5.3 | 85.9 |
| 1.79 | 75 | 3.8 | 3.8 | 89.6 |
| 1.95 | 46 | 2.3 | 2.3 | 92.0 |
| 2.08 | 36 | 1.8 | 1.8 | 93.8 |
| 2.20 | 26 | 1.3 | 1.3 | 95.1 |
| 2.30 | 22 | 1.1 | 1.1 | 96.2 |
| 2.40 | 14 | .7 | .7 | 96.9 |
| 2.48 | 10 | .5 | .5 | 97.4 |
| 2.56 | 9 | .4 | .4 | 97.8 |
| 2.64 | 10 | .5 | .5 | 98.3 |
| 2.71 | 4 | .2 | .2 | 98.5 |
| 2.77 | 4 | .2 | .2 | 98.7 |
| 2.83 | 5 | .3 | .3 | 99.0 |
| 2.89 | 5 | .3 | .3 | 99.2 |
| 3.00 | 1 | .1 | .1 | 99.3 |
| 3.04 | 2 | .1 | .1 | 99.4 |
| 3.09 | 1 | .1 | .1 | 99.4 |
| 3.14 | 2 | .1 | .1 | 99.5 |
| 3.22 | 1 | .1 | .1 | 99.6 |
| 3.26 | 2 | .1 | .1 | 99.7 |
| 3.40 | 1 | .1 | .1 | 99.7 |
| 3.43 | 1 | .1 | .1 | 99.8 |
| 3.47 | 1 | .1 | .1 | 99.8 |
| 3.50 | 1 | .1 | .1 | 99.9 |
| 3.64 | 1 | .1 | .1 | 99.9 |
| 3.74 | 1 | .1 | .1 | 100.0 |
| 4.72 | 1 | .1 | .1 | 100.0 |
| Total | 2000 | 100.0 | 100.0 |  |

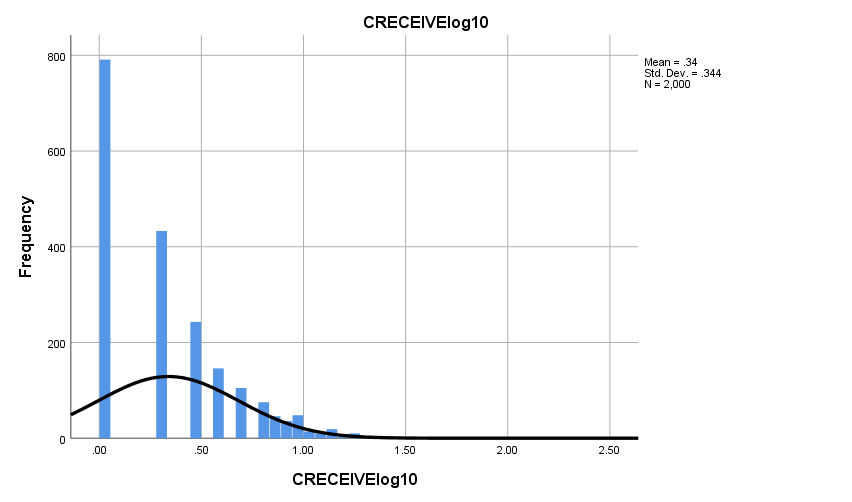
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CRECEIVElog10** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | .00 | 791 | 39.6 | 39.6 | 39.6 |
| .30 | 433 | 21.7 | 21.7 | 61.2 |
| .48 | 243 | 12.2 | 12.2 | 73.4 |
| .60 | 146 | 7.3 | 7.3 | 80.7 |
| .70 | 105 | 5.3 | 5.3 | 85.9 |
| .78 | 75 | 3.8 | 3.8 | 89.6 |
| .85 | 46 | 2.3 | 2.3 | 92.0 |
| .90 | 36 | 1.8 | 1.8 | 93.8 |
| .95 | 26 | 1.3 | 1.3 | 95.1 |
| 1.00 | 22 | 1.1 | 1.1 | 96.2 |
| 1.04 | 14 | .7 | .7 | 96.9 |
| 1.08 | 10 | .5 | .5 | 97.4 |
| 1.11 | 9 | .4 | .4 | 97.8 |
| 1.15 | 10 | .5 | .5 | 98.3 |
| 1.18 | 4 | .2 | .2 | 98.5 |
| 1.20 | 4 | .2 | .2 | 98.7 |
| 1.23 | 5 | .3 | .3 | 99.0 |
| 1.26 | 5 | .3 | .3 | 99.2 |
| 1.30 | 1 | .1 | .1 | 99.3 |
| 1.32 | 2 | .1 | .1 | 99.4 |
| 1.34 | 1 | .1 | .1 | 99.4 |
| 1.36 | 2 | .1 | .1 | 99.5 |
| 1.40 | 1 | .1 | .1 | 99.6 |
| 1.41 | 2 | .1 | .1 | 99.7 |
| 1.48 | 1 | .1 | .1 | 99.7 |
| 1.49 | 1 | .1 | .1 | 99.8 |
| 1.51 | 1 | .1 | .1 | 99.8 |
| 1.52 | 1 | .1 | .1 | 99.9 |
| 1.58 | 1 | .1 | .1 | 99.9 |
| 1.62 | 1 | .1 | .1 | 100.0 |
| 2.05 | 1 | .1 | .1 | 100.0 |
| Total | 2000 | 100.0 | 100.0 |  |

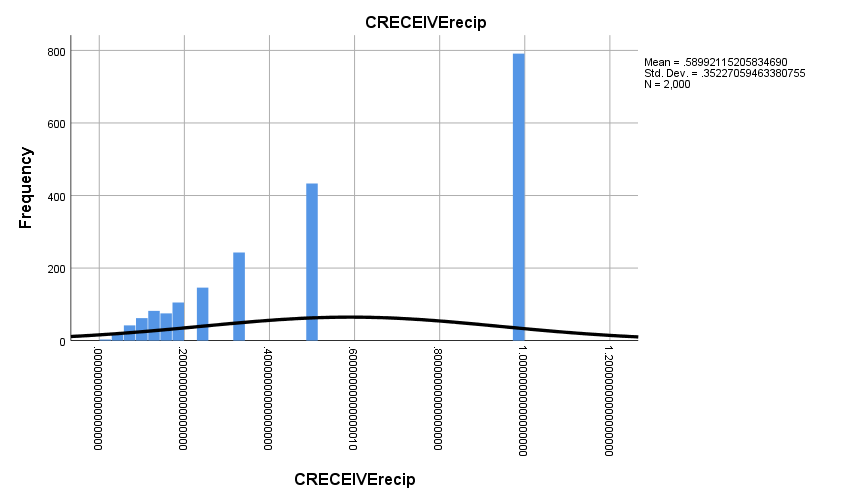
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CRECEIVErecip** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | .00892857142857143 | 1 | .1 | .1 | .1 |
| .02380952380952380 | 1 | .1 | .1 | .1 |
| .02631578947368420 | 1 | .1 | .1 | .2 |
| .03030303030303030 | 1 | .1 | .1 | .2 |
| .03125000000000000 | 1 | .1 | .1 | .3 |
| .03225806451612900 | 1 | .1 | .1 | .3 |
| .03333333333333330 | 1 | .1 | .1 | .4 |
| .03846153846153850 | 2 | .1 | .1 | .4 |
| .04000000000000000 | 1 | .1 | .1 | .5 |
| .04347826086956520 | 2 | .1 | .1 | .6 |
| .04545454545454550 | 1 | .1 | .1 | .7 |
| .04761904761904760 | 2 | .1 | .1 | .8 |
| .05000000000000000 | 1 | .1 | .1 | .8 |
| .05555555555555560 | 5 | .3 | .3 | 1.1 |
| .05882352941176470 | 5 | .3 | .3 | 1.3 |
| .06250000000000000 | 4 | .2 | .2 | 1.5 |
| .06666666666666670 | 4 | .2 | .2 | 1.7 |
| .07142857142857140 | 10 | .5 | .5 | 2.2 |
| .07692307692307690 | 9 | .4 | .4 | 2.7 |
| .08333333333333330 | 10 | .5 | .5 | 3.2 |
| .09090909090909090 | 14 | .7 | .7 | 3.9 |
| .10000000000000000 | 22 | 1.1 | 1.1 | 5.0 |
| .11111111111111100 | 26 | 1.3 | 1.3 | 6.3 |
| .12500000000000000 | 36 | 1.8 | 1.8 | 8.1 |
| .14285714285714300 | 46 | 2.3 | 2.3 | 10.4 |
| .16666666666666700 | 75 | 3.8 | 3.8 | 14.1 |
| .20000000000000000 | 105 | 5.3 | 5.3 | 19.4 |
| .25000000000000000 | 146 | 7.3 | 7.3 | 26.7 |
| .33333333333333300 | 243 | 12.2 | 12.2 | 38.8 |
| .50000000000000000 | 433 | 21.7 | 21.7 | 60.5 |
| 1.00000000000000000 | 791 | 39.6 | 39.6 | 100.0 |
| Total | 2000 | 100.0 | 100.0 |  |

**Histogram**



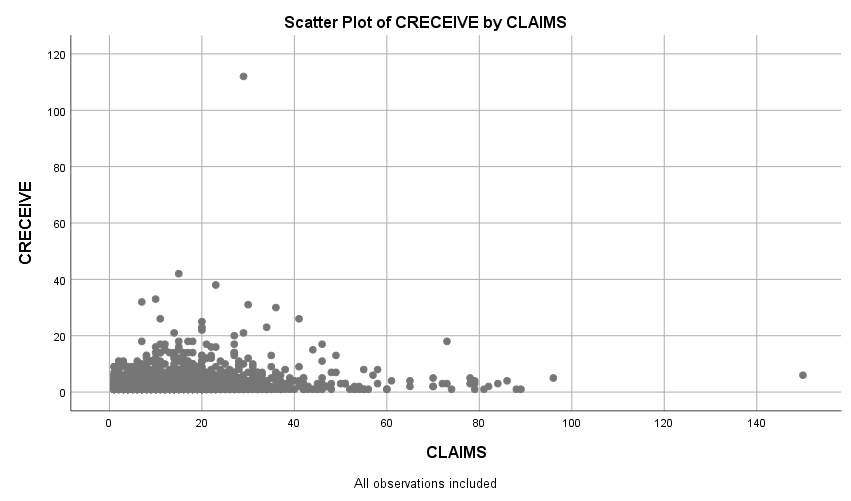






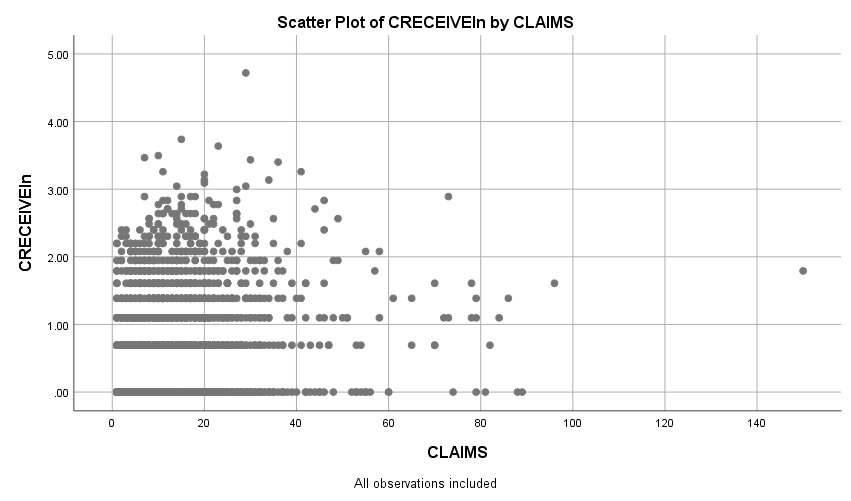
**Graph**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 16:50:51 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 2000 |
| Syntax | | GRAPH  /SCATTERPLOT(BIVAR)=CLAIMS WITH CRECEIVE  /MISSING=LISTWISE  /TITLE='Scatter Plot of CRECEIVE by CLAIMS'  /FOOTNOTE='All observations included'. |
| Resources | Processor Time | 00:00:01.69 |
| Elapsed Time | 00:00:00.88 |



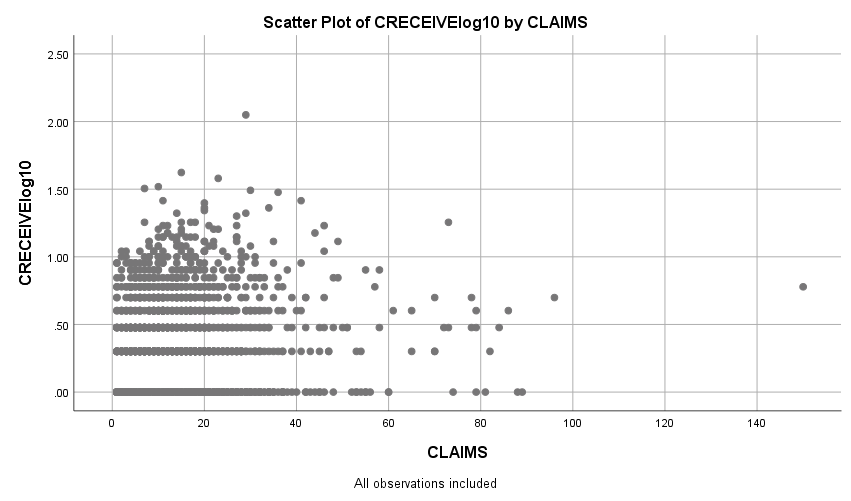
**Graph**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 16:52:07 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 2000 |
| Syntax | | GRAPH  /SCATTERPLOT(BIVAR)=CLAIMS WITH CRECEIVEln  /MISSING=LISTWISE  /TITLE='Scatter Plot of CRECEIVEln by CLAIMS'  /FOOTNOTE='All observations included'. |
| Resources | Processor Time | 00:00:00.97 |
| Elapsed Time | 00:00:00.70 |



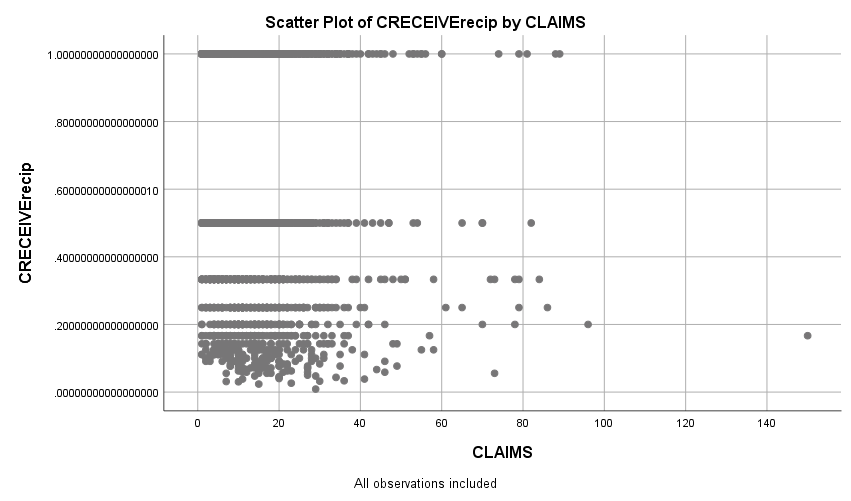
**Graph**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 16:52:57 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 2000 |
| Syntax | | GRAPH  /SCATTERPLOT(BIVAR)=CLAIMS WITH CRECEIVElog10  /MISSING=LISTWISE  /TITLE='Scatter Plot of CRECEIVElog10 by CLAIMS'  /FOOTNOTE='All observations included'. |
| Resources | Processor Time | 00:00:01.19 |
| Elapsed Time | 00:00:00.60 |



**Graph**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 16:53:51 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 2000 |
| Syntax | | GRAPH  /SCATTERPLOT(BIVAR)=CLAIMS WITH CRECEIVErecip  /MISSING=LISTWISE  /TITLE='Scatter Plot of CRECEIVErecip by CLAIMS'  /FOOTNOTE='All observations included'. |
| Resources | Processor Time | 00:00:01.11 |
| Elapsed Time | 00:00:00.59 |



**Regression**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 19:27:42 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | CRECEIVE < 40 AND CLAIMS < 90 (FILTER) |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 1996 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any variable used. |
| Syntax | | REGRESSION  /DESCRIPTIVES MEAN STDDEV CORR SIG N  /MISSING LISTWISE  /STATISTICS COEFF OUTS CI(95) BCOV R ANOVA COLLIN TOL CHANGE ZPP  /CRITERIA=PIN(.05) POUT(.10)  /NOORIGIN  /DEPENDENT GENERAL  /METHOD=ENTER ORIGINAL  /SCATTERPLOT=(\*ZRESID ,\*ZPRED)  /RESIDUALS DURBIN HISTOGRAM(ZRESID) NORMPROB(ZRESID)  /CASEWISE PLOT(ZRESID) OUTLIERS(3). |
| Resources | Processor Time | 00:00:08.63 |
| Elapsed Time | 00:00:04.53 |
| Memory Required | 3488 bytes |
| Additional Memory Required for Residual Plots | 680 bytes |

[DataSet1] D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| GENERAL | .194597 | .2589553 | 1958 |
| ORIGINAL | .398282 | .2739932 | 1958 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | GENERAL | ORIGINAL |
| Pearson Correlation | GENERAL | 1.000 | .169 |
| ORIGINAL | .169 | 1.000 |
| Sig. (1-tailed) | GENERAL | . | .000 |
| ORIGINAL | .000 | . |
| N | GENERAL | 1958 | 1958 |
| ORIGINAL | 1958 | 1958 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | ORIGINALb | . | Enter |

|  |
| --- |
| a. Dependent Variable: GENERAL |
| b. All requested variables entered. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | |
| R Square Change | F Change | df1 |
| 1 | .169a | .029 | .028 | .2552980 | .029 | 57.473 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Summaryb** | | | |
| Model | Change Statistics | | |
| df2 | Sig. F Change |  |
| 1 | 1956 | .000 | 2.053 |

|  |
| --- |
| a. Predictors: (Constant), ORIGINAL |
| b. Dependent Variable: GENERAL |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 3.746 | 1 | 3.746 | 57.473 | .000b |
| Residual | 127.486 | 1956 | .065 |  |  |
| Total | 131.232 | 1957 |  |  |  |

|  |
| --- |
| a. Dependent Variable: GENERAL |
| b. Predictors: (Constant), ORIGINAL |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | .131 | .010 |  | 12.867 | .000 |
| ORIGINAL | .160 | .021 | .169 | 7.581 | .000 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | |
| Model | | 95.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics |
| Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance |
| 1 | (Constant) | .111 | .151 |  |  |  |  |
| ORIGINAL | .118 | .201 | .169 | .169 | .169 | 1.000 |

|  |  |  |
| --- | --- | --- |
| **Coefficientsa** | | |
| Model | | Collinearity Statistics |
| VIF |
| 1 | (Constant) |  |
| ORIGINAL | 1.000 |

|  |
| --- |
| a. Dependent Variable: GENERAL |

|  |  |  |  |
| --- | --- | --- | --- |
| **Coefficient Correlationsa** | | | |
| Model | | | ORIGINAL |
| 1 | Correlations | ORIGINAL | 1.000 |
| Covariances | ORIGINAL | .000 |

|  |
| --- |
| a. Dependent Variable: GENERAL |

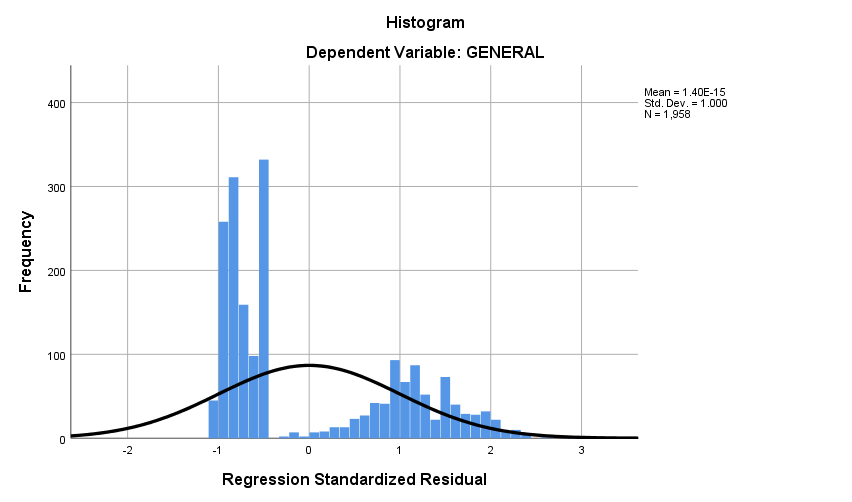
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | | | |
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | |
| (Constant) | ORIGINAL |
| 1 | 1 | 1.824 | 1.000 | .09 | .09 |
| 2 | .176 | 3.219 | .91 | .91 |

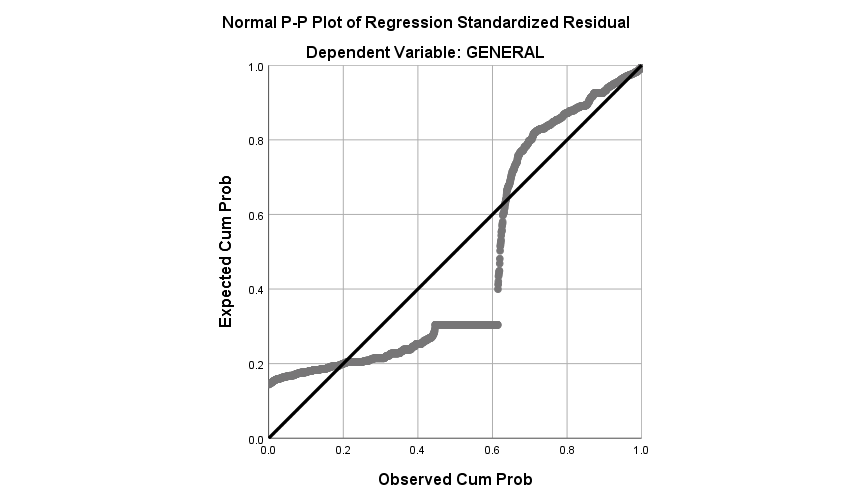
|  |
| --- |
| a. Dependent Variable: GENERAL |

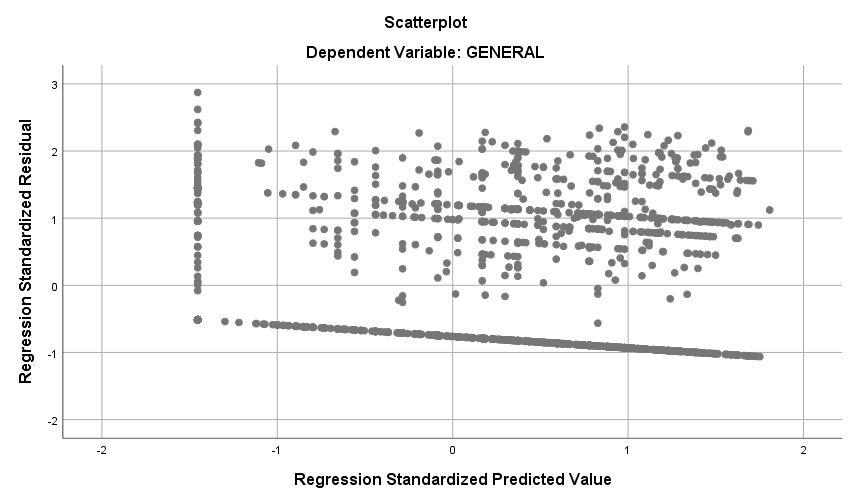
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | .131001 | .273656 | .194597 | .0437504 | 1958 |
| Residual | -.2711814 | .7331990 | .0000000 | .2552328 | 1958 |
| Std. Predicted Value | -1.454 | 1.807 | .000 | 1.000 | 1958 |
| Std. Residual | -1.062 | 2.872 | .000 | 1.000 | 1958 |

|  |
| --- |
| a. Dependent Variable: GENERAL |

**Charts**







**Regression**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 19:32:32 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | CRECEIVE < 40 AND CLAIMS < 90 (FILTER) |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 1996 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any variable used. |
| Syntax | | REGRESSION  /DESCRIPTIVES MEAN STDDEV CORR SIG N  /MISSING LISTWISE  /STATISTICS COEFF OUTS CI(95) BCOV R ANOVA COLLIN TOL CHANGE ZPP  /CRITERIA=PIN(.05) POUT(.10)  /NOORIGIN  /DEPENDENT CLAIMS  /METHOD=ENTER ORIGINAL GENERAL GYEAR RATIOCIT  /SCATTERPLOT=(\*ZRESID ,\*ZPRED)  /RESIDUALS DURBIN HISTOGRAM(ZRESID) NORMPROB(ZRESID)  /CASEWISE PLOT(ZRESID) OUTLIERS(3). |
| Resources | Processor Time | 00:00:02.46 |
| Elapsed Time | 00:00:01.57 |
| Memory Required | 5072 bytes |
| Additional Memory Required for Residual Plots | 632 bytes |

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| CLAIMS | 14.97 | 11.689 | 1958 |
| ORIGINAL | .398282 | .2739932 | 1958 |
| GENERAL | .194597 | .2589553 | 1958 |
| GYEAR | 1996.27 | 1.075 | 1958 |
| RATIOCIT | .939529 | .1404329 | 1958 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | | |
|  | | CLAIMS | ORIGINAL | GENERAL | GYEAR | RATIOCIT |
| Pearson Correlation | CLAIMS | 1.000 | .101 | .056 | .039 | .054 |
| ORIGINAL | .101 | 1.000 | .169 | .017 | .052 |
| GENERAL | .056 | .169 | 1.000 | -.238 | .043 |
| GYEAR | .039 | .017 | -.238 | 1.000 | .083 |
| RATIOCIT | .054 | .052 | .043 | .083 | 1.000 |
| Sig. (1-tailed) | CLAIMS | . | .000 | .006 | .042 | .008 |
| ORIGINAL | .000 | . | .000 | .225 | .011 |
| GENERAL | .006 | .000 | . | .000 | .028 |
| GYEAR | .042 | .225 | .000 | . | .000 |
| RATIOCIT | .008 | .011 | .028 | .000 | . |
| N | CLAIMS | 1958 | 1958 | 1958 | 1958 | 1958 |
| ORIGINAL | 1958 | 1958 | 1958 | 1958 | 1958 |
| GENERAL | 1958 | 1958 | 1958 | 1958 | 1958 |
| GYEAR | 1958 | 1958 | 1958 | 1958 | 1958 |
| RATIOCIT | 1958 | 1958 | 1958 | 1958 | 1958 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | RATIOCIT, GENERAL, ORIGINAL, GYEARb | . | Enter |

|  |
| --- |
| a. Dependent Variable: CLAIMS |
| b. All requested variables entered. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | |
| R Square Change | F Change | df1 |
| 1 | .127a | .016 | .014 | 11.607 | .016 | 7.952 | 4 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Summaryb** | | | |
| Model | Change Statistics | | |
| df2 | Sig. F Change |  |
| 1 | 1953 | .000 | 1.966 |

|  |
| --- |
| a. Predictors: (Constant), RATIOCIT, GENERAL, ORIGINAL, GYEAR |
| b. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 4285.217 | 4 | 1071.304 | 7.952 | .000b |
| Residual | 263126.455 | 1953 | 134.729 |  |  |
| Total | 267411.672 | 1957 |  |  |  |

|  |
| --- |
| a. Dependent Variable: CLAIMS |
| b. Predictors: (Constant), RATIOCIT, GENERAL, ORIGINAL, GYEAR |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | -986.372 | 504.599 |  | -1.955 | .051 |
| ORIGINAL | 3.820 | .974 | .090 | 3.921 | .000 |
| GENERAL | 2.274 | 1.062 | .050 | 2.142 | .032 |
| GYEAR | .499 | .253 | .046 | 1.973 | .049 |
| RATIOCIT | 3.608 | 1.880 | .043 | 1.919 | .055 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | |
| Model | | 95.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics |
| Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance |
| 1 | (Constant) | -1975.981 | 3.237 |  |  |  |  |
| ORIGINAL | 1.909 | 5.730 | .101 | .088 | .088 | .966 |
| GENERAL | .192 | 4.357 | .056 | .048 | .048 | .911 |
| GYEAR | .003 | .995 | .039 | .045 | .044 | .932 |
| RATIOCIT | -.080 | 7.296 | .054 | .043 | .043 | .987 |

|  |  |  |
| --- | --- | --- |
| **Coefficientsa** | | |
| Model | | Collinearity Statistics |
| VIF |
| 1 | (Constant) |  |
| ORIGINAL | 1.035 |
| GENERAL | 1.098 |
| GYEAR | 1.073 |
| RATIOCIT | 1.013 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficient Correlationsa** | | | | | | |
| Model | | | RATIOCIT | GENERAL | ORIGINAL | GYEAR |
| 1 | Correlations | RATIOCIT | 1.000 | -.057 | -.040 | -.094 |
| GENERAL | -.057 | 1.000 | -.175 | .248 |
| ORIGINAL | -.040 | -.175 | 1.000 | -.056 |
| GYEAR | -.094 | .248 | -.056 | 1.000 |
| Covariances | RATIOCIT | 3.536 | -.113 | -.073 | -.044 |
| GENERAL | -.113 | 1.128 | -.181 | .067 |
| ORIGINAL | -.073 | -.181 | .949 | -.014 |
| GYEAR | -.044 | .067 | -.014 | .064 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | | | | | |
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | | |
| (Constant) | ORIGINAL | GENERAL | GYEAR |
| 1 | 1 | 4.172 | 1.000 | .00 | .01 | .02 | .00 |
| 2 | .558 | 2.734 | .00 | .00 | .91 | .00 |
| 3 | .255 | 4.048 | .00 | .98 | .01 | .00 |
| 4 | .015 | 16.893 | .00 | .00 | .00 | .00 |
| 5 | 1.351E-7 | 5557.627 | 1.00 | .00 | .06 | 1.00 |

|  |  |  |
| --- | --- | --- |
| **Collinearity Diagnosticsa** | | |
| Model | Dimension | Variance Proportions |
| RATIOCIT |
| 1 | 1 | .00 |
| 2 | .00 |
| 3 | .01 |
| 4 | .98 |
| 5 | .01 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

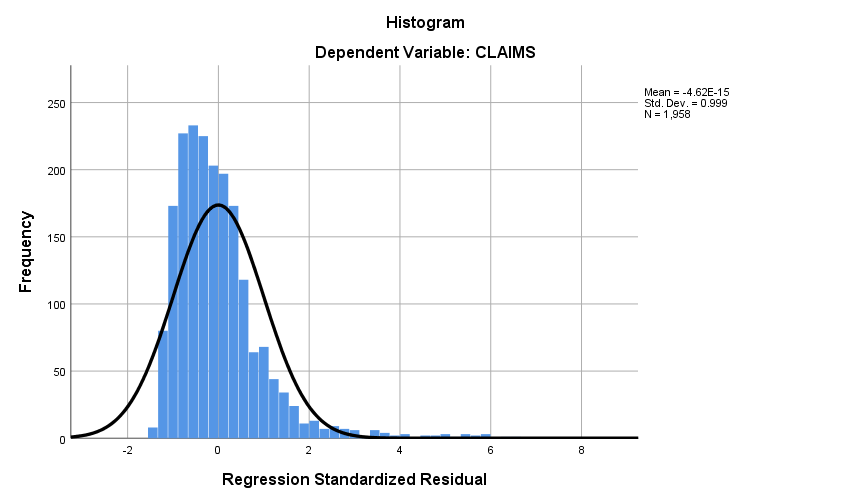
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Casewise Diagnosticsa** | | | | |
| Case Number | Std. Residual | CLAIMS | Predicted Value | Residual |
| 33 | 5.444 | 79 | 15.81 | 63.192 |
| 102 | 6.496 | 88 | 12.60 | 75.404 |
| 114 | 4.478 | 70 | 18.02 | 51.981 |
| 117 | 3.576 | 57 | 15.49 | 41.510 |
| 149 | 4.972 | 72 | 14.29 | 57.707 |
| 187 | 3.016 | 53 | 17.99 | 35.012 |
| 255 | 5.088 | 73 | 13.94 | 59.061 |
| 290 | 4.781 | 70 | 14.51 | 55.490 |
| 382 | 4.097 | 61 | 13.45 | 47.551 |
| 501 | 3.060 | 49 | 13.49 | 35.515 |
| 515 | 3.909 | 60 | 14.63 | 45.375 |
| 541 | 4.146 | 65 | 16.87 | 48.128 |
| 605 | 4.999 | 74 | 15.98 | 58.019 |
| 668 | 3.610 | 56 | 14.09 | 41.907 |
| 766 | 3.054 | 52 | 16.55 | 35.445 |
| 794 | 3.606 | 58 | 16.14 | 41.857 |
| 860 | 3.771 | 60 | 16.23 | 43.768 |
| 868 | 5.362 | 79 | 16.77 | 62.233 |
| 951 | 4.127 | 65 | 17.10 | 47.903 |
| 1010 | 3.075 | 51 | 15.30 | 35.696 |
| 1024 | 5.912 | 84 | 15.38 | 68.617 |
| 1029 | 3.382 | 54 | 14.74 | 39.258 |
| 1128 | 5.289 | 79 | 17.60 | 61.395 |
| 1160 | 5.807 | 81 | 13.59 | 67.406 |
| 1210 | 3.789 | 58 | 14.02 | 43.981 |
| 1248 | 4.772 | 73 | 17.61 | 55.390 |
| 1272 | 3.367 | 55 | 15.92 | 39.084 |
| 1336 | 3.439 | 55 | 15.08 | 39.916 |
| 1381 | 6.375 | 89 | 15.00 | 73.995 |
| 1451 | 3.266 | 51 | 13.09 | 37.905 |
| 1461 | 3.375 | 54 | 14.83 | 39.174 |
| 1480 | 5.592 | 78 | 13.09 | 64.905 |
| 1507 | 5.635 | 82 | 16.60 | 65.402 |
| 1557 | 3.104 | 53 | 16.97 | 36.029 |
| 1655 | 5.537 | 78 | 13.73 | 64.274 |
| 1671 | 3.438 | 53 | 13.09 | 39.905 |
| 1774 | 5.883 | 86 | 17.72 | 68.280 |
| 1884 | 3.488 | 55 | 14.51 | 40.492 |
| 1991 | 4.615 | 70 | 16.43 | 53.565 |

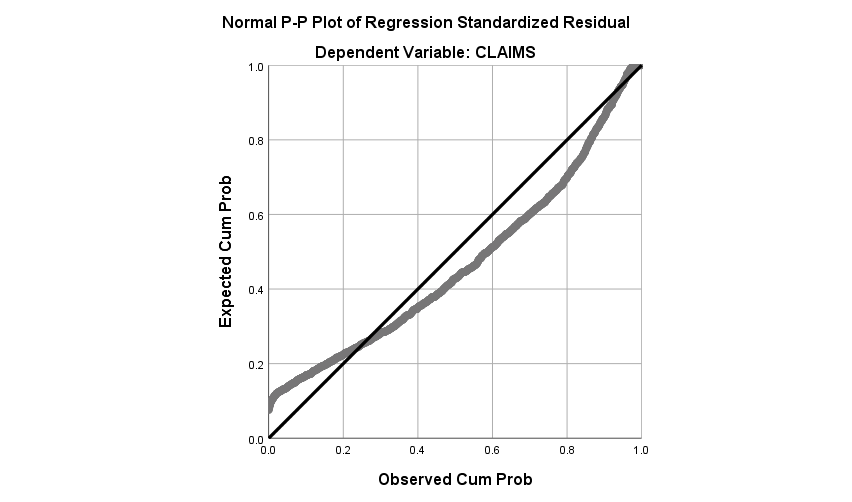
|  |
| --- |
| a. Dependent Variable: CLAIMS |

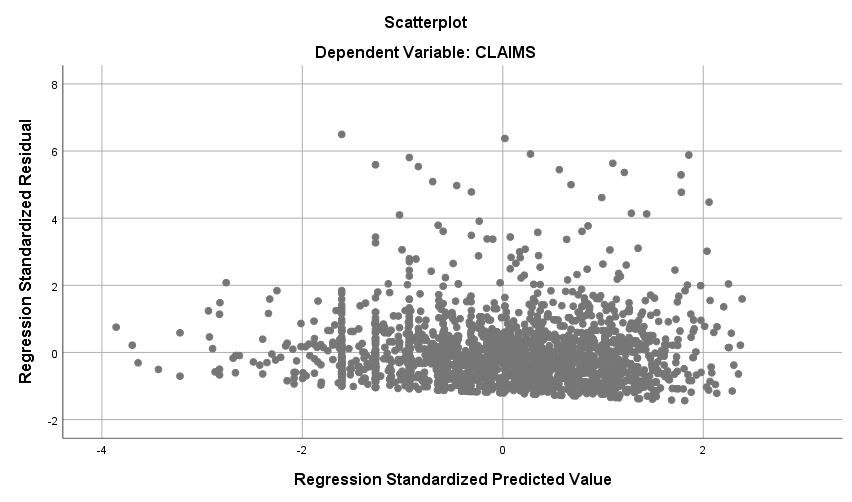
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | 9.27 | 18.51 | 14.97 | 1.480 | 1958 |
| Residual | -16.660 | 75.404 | .000 | 11.595 | 1958 |
| Std. Predicted Value | -3.858 | 2.388 | .000 | 1.000 | 1958 |
| Std. Residual | -1.435 | 6.496 | .000 | .999 | 1958 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

**Charts**







**Regression**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 28-OCT-2018 19:34:32 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | CRECEIVE < 40 AND CLAIMS < 90 (FILTER) |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 1996 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any variable used. |
| Syntax | | REGRESSION  /DESCRIPTIVES MEAN STDDEV CORR SIG N  /MISSING LISTWISE  /STATISTICS COEFF OUTS CI(95) BCOV R ANOVA COLLIN TOL CHANGE ZPP  /CRITERIA=PIN(.05) POUT(.10)  /NOORIGIN  /DEPENDENT CRECEIVEln  /METHOD=ENTER ORIGINAL GENERAL GYEAR RATIOCIT CLAIMS  /SCATTERPLOT=(\*ZRESID ,\*ZPRED)  /RESIDUALS DURBIN HISTOGRAM(ZRESID) NORMPROB(ZRESID)  /CASEWISE PLOT(ZRESID) OUTLIERS(3). |
| Resources | Processor Time | 00:00:01.95 |
| Elapsed Time | 00:00:01.54 |
| Memory Required | 5728 bytes |
| Additional Memory Required for Residual Plots | 616 bytes |

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| CRECEIVEln | .7773 | .78494 | 1958 |
| ORIGINAL | .398282 | .2739932 | 1958 |
| GENERAL | .194597 | .2589553 | 1958 |
| GYEAR | 1996.27 | 1.075 | 1958 |
| RATIOCIT | .939529 | .1404329 | 1958 |
| CLAIMS | 14.97 | 11.689 | 1958 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | | |
|  | | CRECEIVEln | ORIGINAL | GENERAL | GYEAR | RATIOCIT |
| Pearson Correlation | CRECEIVEln | 1.000 | .019 | .612 | -.339 | .098 |
| ORIGINAL | .019 | 1.000 | .169 | .017 | .052 |
| GENERAL | .612 | .169 | 1.000 | -.238 | .043 |
| GYEAR | -.339 | .017 | -.238 | 1.000 | .083 |
| RATIOCIT | .098 | .052 | .043 | .083 | 1.000 |
| CLAIMS | .083 | .101 | .056 | .039 | .054 |
| Sig. (1-tailed) | CRECEIVEln | . | .204 | .000 | .000 | .000 |
| ORIGINAL | .204 | . | .000 | .225 | .011 |
| GENERAL | .000 | .000 | . | .000 | .028 |
| GYEAR | .000 | .225 | .000 | . | .000 |
| RATIOCIT | .000 | .011 | .028 | .000 | . |
| CLAIMS | .000 | .000 | .006 | .042 | .008 |
| N | CRECEIVEln | 1958 | 1958 | 1958 | 1958 | 1958 |
| ORIGINAL | 1958 | 1958 | 1958 | 1958 | 1958 |
| GENERAL | 1958 | 1958 | 1958 | 1958 | 1958 |
| GYEAR | 1958 | 1958 | 1958 | 1958 | 1958 |
| RATIOCIT | 1958 | 1958 | 1958 | 1958 | 1958 |
| CLAIMS | 1958 | 1958 | 1958 | 1958 | 1958 |

|  |  |  |
| --- | --- | --- |
| **Correlations** | | |
|  | | CLAIMS |
| Pearson Correlation | CRECEIVEln | .083 |
| ORIGINAL | .101 |
| GENERAL | .056 |
| GYEAR | .039 |
| RATIOCIT | .054 |
| CLAIMS | 1.000 |
| Sig. (1-tailed) | CRECEIVEln | .000 |
| ORIGINAL | .000 |
| GENERAL | .006 |
| GYEAR | .042 |
| RATIOCIT | .008 |
| CLAIMS | . |
| N | CRECEIVEln | 1958 |
| ORIGINAL | 1958 |
| GENERAL | 1958 |
| GYEAR | 1958 |
| RATIOCIT | 1958 |
| CLAIMS | 1958 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | CLAIMS, GYEAR, RATIOCIT, ORIGINAL, GENERALb | . | Enter |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |
| b. All requested variables entered. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | |
| R Square Change | F Change | df1 |
| 1 | .658a | .433 | .432 | .59178 | .433 | 298.204 | 5 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Summaryb** | | | |
| Model | Change Statistics | | |
| df2 | Sig. F Change |  |
| 1 | 1952 | .000 | 2.008 |

|  |
| --- |
| a. Predictors: (Constant), CLAIMS, GYEAR, RATIOCIT, ORIGINAL, GENERAL |
| b. Dependent Variable: CRECEIVEln |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 522.161 | 5 | 104.432 | 298.204 | .000b |
| Residual | 683.599 | 1952 | .350 |  |  |
| Total | 1205.759 | 1957 |  |  |  |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |
| b. Predictors: (Constant), CLAIMS, GYEAR, RATIOCIT, ORIGINAL, GENERAL |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 310.279 | 25.751 |  | 12.049 | .000 |
| ORIGINAL | -.243 | .050 | -.085 | -4.874 | .000 |
| GENERAL | 1.723 | .054 | .569 | 31.794 | .000 |
| GYEAR | -.155 | .013 | -.213 | -12.046 | .000 |
| RATIOCIT | .516 | .096 | .092 | 5.382 | .000 |
| CLAIMS | .004 | .001 | .063 | 3.650 | .000 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | |
| Model | | 95.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics |
| Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance |
| 1 | (Constant) | 259.776 | 360.782 |  |  |  |  |
| ORIGINAL | -.341 | -.145 | .019 | -.110 | -.083 | .959 |
| GENERAL | 1.617 | 1.830 | .612 | .584 | .542 | .908 |
| GYEAR | -.181 | -.130 | -.339 | -.263 | -.205 | .930 |
| RATIOCIT | .328 | .705 | .098 | .121 | .092 | .986 |
| CLAIMS | .002 | .006 | .083 | .082 | .062 | .984 |

|  |  |  |
| --- | --- | --- |
| **Coefficientsa** | | |
| Model | | Collinearity Statistics |
| VIF |
| 1 | (Constant) |  |
| ORIGINAL | 1.043 |
| GENERAL | 1.101 |
| GYEAR | 1.075 |
| RATIOCIT | 1.015 |
| CLAIMS | 1.016 |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficient Correlationsa** | | | | | | | |
| Model | | | CLAIMS | GYEAR | RATIOCIT | ORIGINAL | GENERAL |
| 1 | Correlations | CLAIMS | 1.000 | -.045 | -.043 | -.088 | -.048 |
| GYEAR | -.045 | 1.000 | -.091 | -.052 | .250 |
| RATIOCIT | -.043 | -.091 | 1.000 | -.036 | -.055 |
| ORIGINAL | -.088 | -.052 | -.036 | 1.000 | -.170 |
| GENERAL | -.048 | .250 | -.055 | -.170 | 1.000 |
| Covariances | CLAIMS | 1.331E-6 | -6.640E-7 | -4.802E-6 | -5.084E-6 | -3.027E-6 |
| GYEAR | -6.640E-7 | .000 | .000 | -3.315E-5 | .000 |
| RATIOCIT | -4.802E-6 | .000 | .009 | .000 | .000 |
| ORIGINAL | -5.084E-6 | -3.315E-5 | .000 | .002 | .000 |
| GENERAL | -3.027E-6 | .000 | .000 | .000 | .003 |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | | | | | |
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | | |
| (Constant) | ORIGINAL | GENERAL | GYEAR |
| 1 | 1 | 4.843 | 1.000 | .00 | .01 | .01 | .00 |
| 2 | .574 | 2.905 | .00 | .00 | .89 | .00 |
| 3 | .321 | 3.886 | .00 | .18 | .03 | .00 |
| 4 | .248 | 4.415 | .00 | .80 | .01 | .00 |
| 5 | .015 | 18.199 | .00 | .00 | .00 | .00 |
| 6 | 1.348E-7 | 5993.179 | 1.00 | .00 | .06 | 1.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | |
| Model | Dimension | Variance Proportions | |
| RATIOCIT | CLAIMS |
| 1 | 1 | .00 | .01 |
| 2 | .00 | .03 |
| 3 | .00 | .85 |
| 4 | .01 | .10 |
| 5 | .98 | .00 |
| 6 | .01 | .00 |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |

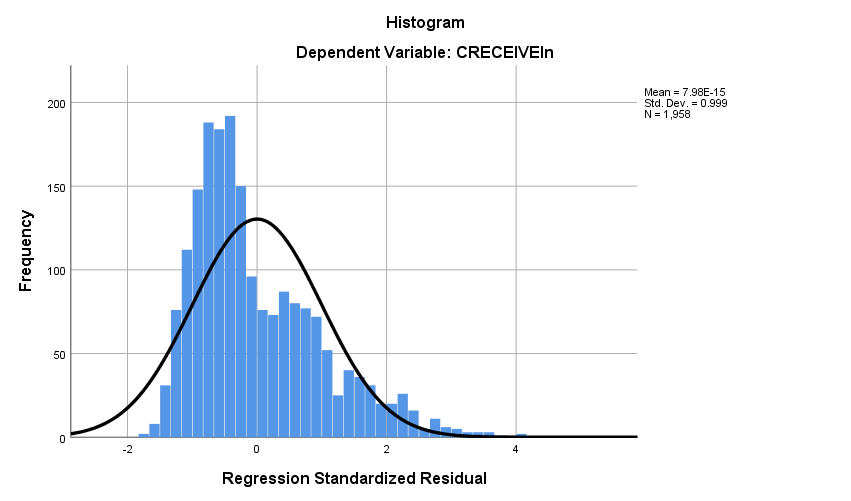
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Casewise Diagnosticsa** | | | | |
| Case Number | Std. Residual | CRECEIVEln | Predicted Value | Residual |
| 159 | 3.367 | 2.20 | .2046 | 1.99260 |
| 198 | 3.034 | 2.48 | .6894 | 1.79550 |
| 243 | 3.559 | 3.14 | 1.0295 | 2.10601 |
| 269 | 3.441 | 2.64 | .6028 | 2.03628 |
| 446 | 3.513 | 3.22 | 1.1397 | 2.07915 |
| 457 | 4.068 | 3.43 | 1.0264 | 2.40757 |
| 776 | 3.046 | 2.71 | .9052 | 1.80282 |
| 832 | 3.199 | 2.64 | .7461 | 1.89295 |
| 859 | 3.632 | 2.56 | .4157 | 2.14925 |
| 879 | 3.013 | 3.26 | 1.4748 | 1.78326 |
| 1058 | 3.130 | 1.95 | .0934 | 1.85254 |
| 1063 | 4.125 | 3.64 | 1.1964 | 2.44122 |
| 1116 | 3.895 | 3.09 | .7859 | 2.30512 |
| 1220 | 3.375 | 3.50 | 1.4990 | 1.99752 |
| 1347 | 3.126 | 2.83 | .9832 | 1.85002 |
| 1379 | 3.251 | 2.48 | .5613 | 1.92362 |
| 1680 | 3.815 | 3.26 | 1.0002 | 2.25791 |
| 1744 | 3.179 | 2.77 | .8914 | 1.88122 |

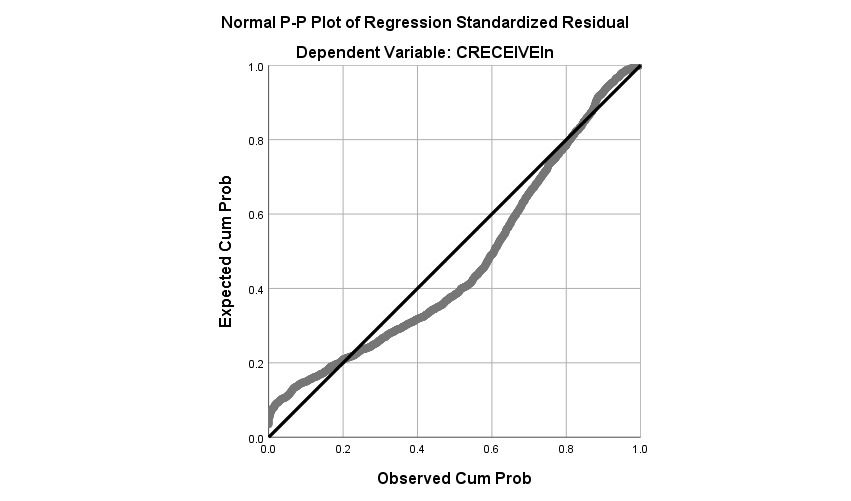
|  |
| --- |
| a. Dependent Variable: CRECEIVEln |

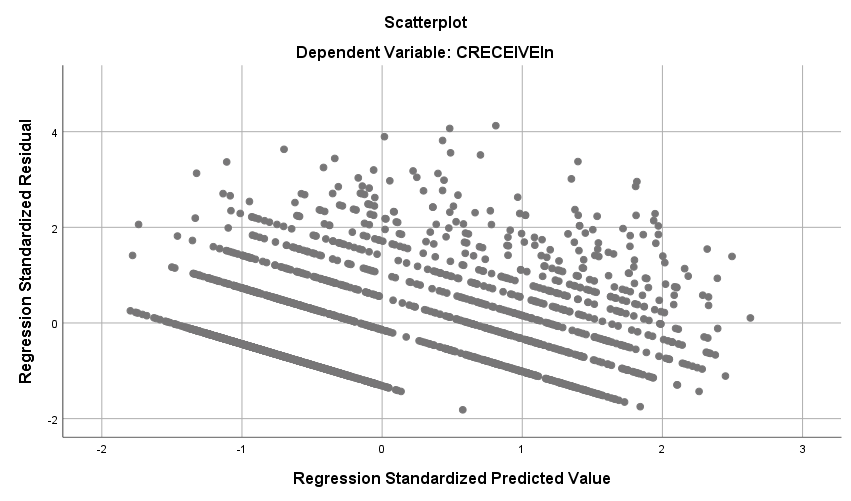
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | -.1505 | 2.1346 | .7773 | .51654 | 1958 |
| Residual | -1.07455 | 2.44122 | .00000 | .59102 | 1958 |
| Std. Predicted Value | -1.796 | 2.628 | .000 | 1.000 | 1958 |
| Std. Residual | -1.816 | 4.125 | .000 | .999 | 1958 |

|  |
| --- |
| a. Dependent Variable: CRECEIVEln |

**Charts**







**Regression**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 30-OCT-2018 18:57:59 |
| Comments | |  |
| Input | Data | D:\SOC6100\Assignments\Assignment03\Data\DataClean\Townes\_SOC6100\_Assignment03\_Data.sav |
| Active Dataset | DataSet1 |
| Filter | CRECEIVE<40 AND CLAIMS<90 (FILTER) |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 1996 |
| Missing Value Handling | Definition of Missing | User-defined missing values are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any variable used. |
| Syntax | | REGRESSION  /DESCRIPTIVES MEAN STDDEV CORR SIG N  /MISSING LISTWISE  /STATISTICS COEFF OUTS CI(95) BCOV R ANOVA COLLIN TOL CHANGE ZPP  /CRITERIA=PIN(.05) POUT(.10)  /NOORIGIN  /DEPENDENT CLAIMS  /METHOD=ENTER GENERAL GYEAR ORIGINAL  /SCATTERPLOT=(\*ZRESID ,\*ZPRED)  /RESIDUALS DURBIN HISTOGRAM(ZRESID) NORMPROB(ZRESID)  /CASEWISE PLOT(ZRESID) OUTLIERS(3). |
| Resources | Processor Time | 00:00:04.10 |
| Elapsed Time | 00:00:01.87 |
| Memory Required | 4480 bytes |
| Additional Memory Required for Residual Plots | 648 bytes |

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| CLAIMS | 14.97 | 11.689 | 1958 |
| GENERAL | .194597 | .2589553 | 1958 |
| GYEAR | 1996.27 | 1.075 | 1958 |
| ORIGINAL | .398282 | .2739932 | 1958 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
|  | | CLAIMS | GENERAL | GYEAR | ORIGINAL |
| Pearson Correlation | CLAIMS | 1.000 | .056 | .039 | .101 |
| GENERAL | .056 | 1.000 | -.238 | .169 |
| GYEAR | .039 | -.238 | 1.000 | .017 |
| ORIGINAL | .101 | .169 | .017 | 1.000 |
| Sig. (1-tailed) | CLAIMS | . | .006 | .042 | .000 |
| GENERAL | .006 | . | .000 | .000 |
| GYEAR | .042 | .000 | . | .225 |
| ORIGINAL | .000 | .000 | .225 | . |
| N | CLAIMS | 1958 | 1958 | 1958 | 1958 |
| GENERAL | 1958 | 1958 | 1958 | 1958 |
| GYEAR | 1958 | 1958 | 1958 | 1958 |
| ORIGINAL | 1958 | 1958 | 1958 | 1958 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables Entered/Removeda** | | | |
| Model | Variables Entered | Variables Removed | Method |
| 1 | ORIGINAL, GYEAR, GENERALb | . | Enter |

|  |
| --- |
| a. Dependent Variable: CLAIMS |
| b. All requested variables entered. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | |
| R Square Change | F Change | df1 |
| 1 | .119a | .014 | .013 | 11.615 | .014 | 9.362 | 3 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Summaryb** | | | |
| Model | Change Statistics | | |
| df2 | Sig. F Change |  |
| 1 | 1954 | .000 | 1.963 |

|  |
| --- |
| a. Predictors: (Constant), ORIGINAL, GYEAR, GENERAL |
| b. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 3789.155 | 3 | 1263.052 | 9.362 | .000b |
| Residual | 263622.516 | 1954 | 134.914 |  |  |
| Total | 267411.672 | 1957 |  |  |  |

|  |
| --- |
| a. Dependent Variable: CLAIMS |
| b. Predictors: (Constant), ORIGINAL, GYEAR, GENERAL |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | -1073.630 | 502.890 |  | -2.135 | .033 |
| GENERAL | 2.390 | 1.061 | .053 | 2.253 | .024 |
| GYEAR | .544 | .252 | .050 | 2.161 | .031 |
| ORIGINAL | 3.894 | .974 | .091 | 3.998 | .000 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | |
| Model | | 95.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics |
| Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance |
| 1 | (Constant) | -2059.888 | -87.372 |  |  |  |  |
| GENERAL | .310 | 4.471 | .056 | .051 | .051 | .913 |
| GYEAR | .050 | 1.038 | .039 | .049 | .049 | .940 |
| ORIGINAL | 1.984 | 5.805 | .101 | .090 | .090 | .968 |

|  |  |  |
| --- | --- | --- |
| **Coefficientsa** | | |
| Model | | Collinearity Statistics |
| VIF |
| 1 | (Constant) |  |
| GENERAL | 1.095 |
| GYEAR | 1.064 |
| ORIGINAL | 1.033 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Coefficient Correlationsa** | | | | | |
| Model | | | ORIGINAL | GYEAR | GENERAL |
| 1 | Correlations | ORIGINAL | 1.000 | -.060 | -.178 |
| GYEAR | -.060 | 1.000 | .244 |
| GENERAL | -.178 | .244 | 1.000 |
| Covariances | ORIGINAL | .949 | -.015 | -.184 |
| GYEAR | -.015 | .063 | .065 |
| GENERAL | -.184 | .065 | 1.125 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | | | | | |
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | | |
| (Constant) | GENERAL | GYEAR | ORIGINAL |
| 1 | 1 | 3.237 | 1.000 | .00 | .03 | .00 | .02 |
| 2 | .531 | 2.468 | .00 | .91 | .00 | .02 |
| 3 | .232 | 3.737 | .00 | .00 | .00 | .95 |
| 4 | 1.362E-7 | 4874.267 | 1.00 | .06 | 1.00 | .00 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Casewise Diagnosticsa** | | | | |
| Case Number | Std. Residual | CLAIMS | Predicted Value | Residual |
| 33 | 5.454 | 79 | 15.65 | 63.351 |
| 102 | 6.520 | 88 | 12.27 | 75.732 |
| 114 | 4.485 | 70 | 17.90 | 52.096 |
| 117 | 3.466 | 57 | 16.75 | 40.253 |
| 149 | 4.994 | 72 | 14.00 | 58.002 |
| 187 | 3.021 | 53 | 17.91 | 35.092 |
| 255 | 4.983 | 73 | 15.12 | 57.880 |
| 290 | 4.797 | 70 | 14.28 | 55.722 |
| 382 | 4.118 | 61 | 13.16 | 47.836 |
| 501 | 3.081 | 49 | 13.21 | 35.789 |
| 515 | 3.931 | 60 | 14.34 | 45.663 |
| 541 | 4.158 | 65 | 16.70 | 48.301 |
| 605 | 5.011 | 74 | 15.79 | 58.210 |
| 668 | 3.624 | 56 | 13.90 | 42.099 |
| 766 | 3.064 | 52 | 16.41 | 35.589 |
| 794 | 3.617 | 58 | 15.99 | 42.008 |
| 860 | 3.781 | 60 | 16.08 | 43.918 |
| 868 | 5.375 | 79 | 16.57 | 62.434 |
| 951 | 4.136 | 65 | 16.96 | 48.045 |
| 1010 | 3.094 | 51 | 15.06 | 35.939 |
| 1024 | 5.913 | 84 | 15.32 | 68.684 |
| 1029 | 3.404 | 54 | 14.46 | 39.544 |
| 1128 | 5.297 | 79 | 17.48 | 61.522 |
| 1160 | 5.824 | 81 | 13.36 | 67.644 |
| 1210 | 3.807 | 58 | 13.78 | 44.217 |
| 1248 | 4.764 | 73 | 17.67 | 55.331 |
| 1272 | 3.374 | 55 | 15.81 | 39.186 |
| 1336 | 3.457 | 55 | 14.84 | 40.160 |
| 1381 | 6.392 | 89 | 14.76 | 74.241 |
| 1451 | 3.288 | 51 | 12.81 | 38.188 |
| 1461 | 3.367 | 54 | 14.89 | 39.106 |
| 1480 | 5.612 | 78 | 12.81 | 65.188 |
| 1507 | 5.649 | 82 | 16.38 | 65.616 |
| 1557 | 3.114 | 53 | 16.84 | 36.165 |
| 1655 | 5.557 | 78 | 13.46 | 64.544 |
| 1671 | 3.460 | 53 | 12.81 | 40.188 |
| 1774 | 5.887 | 86 | 17.63 | 68.374 |
| 1884 | 3.418 | 55 | 15.30 | 39.696 |
| 1991 | 4.627 | 70 | 16.25 | 53.747 |
| 1994 | 3.018 | 50 | 14.95 | 35.055 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | 12.27 | 18.45 | 14.97 | 1.391 | 1958 |
| Residual | -16.534 | 75.732 | .000 | 11.606 | 1958 |
| Std. Predicted Value | -1.945 | 2.498 | .000 | 1.000 | 1958 |
| Std. Residual | -1.423 | 6.520 | .000 | .999 | 1958 |

|  |
| --- |
| a. Dependent Variable: CLAIMS |

**Charts**

