|  |
| --- |
| **Summary Report: Credit Data** |
| **CAPP 30254** |
| **April 20, 2015**   |  |  | | --- | --- | | Macintosh HD:Users:Joshua:Pictures:Logos and Signatures:University of Chicago:uchicago_rgb_maroon-copy2.jpg | Macintosh HD:Users:Joshua:Pictures:Logos and Signatures:University of Chicago:Chicago Seal transparent.gif |  |  |  |  | | --- | --- | --- | |  |  |  |   **Macintosh HD:Users:Joshua:Pictures:Logos and Signatures:University of Chicago:Chicago Seal transparent.gif** |
| **PRimary Investigator:** |
| **Joshua Mausolf** |

1. **Summary Data for Serious Delinquencies in the Last Two Years:**

|  |  |
| --- | --- |
| **FIGURE 1** | **Serious Delinquencies in the Last Two Years** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:serious_dlqin2yrs_bar.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 1** | **Serious Delinquencies in the Last Two Years** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Percent Delinquent = | | 6.68% |
| Mean = | | 0.067 |
| Median = | | 0.00 |
| Min, Max = | | (0, 1) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Revolving Utilization of Unsecured Lines:**

|  |  |
| --- | --- |
| **FIGURE 2** | **Revolving Utilization of Unsecured Lines** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:revolving_utilization_of_unsecured_lines_histogram2_1.5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 2** | **Revolving Utilization of Unsecured Lines** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 6.05 |
| Median = | | 0.15 |
| Standard Deviation = | | 249.76 |
| Min, Max = | | (0, 50,708) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Age:**

|  |  |
| --- | --- |
| **FIGURE 3** | **Age** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:age_histogram1_5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 3** | **Age** | |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 52 |
| Median = | | 52 |
| Mode = | | 49 |
| Standard Deviation = | | 15 |
| Min, Max = | | (0, 109) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Times 30-59 Days Past Due (Not Worse):**

|  |  |
| --- | --- |
| **FIGURE 4** | **Number of Times 30-59 Days Past Due (Not Worse)** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_of_time30-59_days_past_due_not_worse_histogram2_1.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 4** | **Number of Times 30-59 Days Past Due (Not Worse)** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 0.42 |
| Median = | | 0.00 |
| Standard Deviation = | | 4.19 |
| Min, Max = | | (0, 98) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Times 60-89 Days Past Due (Not Worse):**

|  |  |
| --- | --- |
| **FIGURE 5** | **Number of Times 60-89 Days Past Due (Not Worse)** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_of_time60-89_days_past_due_not_worse_histogram2_1.5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 5** | **Number of Times 60-89 Days Past Due (Not Worse)** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 0.24 |
| Median = | | 0.00 |
| Standard Deviation = | | 4.16 |
| Min, Max = | | (0, 98) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Debt Ratio:**

|  |  |
| --- | --- |
| **FIGURE 6** | **Debt Ratio** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:debt_ratio_histogram2_1.5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 6** | **Debt Ratio** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 353.01 |
| Median = | | 0.37 |
| Standard Deviation = | | 2037.82 |
| Min, Max = | | (0, 329,664) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Monthly Income:**

|  |  |
| --- | --- |
| **FIGURE 7** | **Monthly Income** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:monthly_income_histogram2_1.5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 7** | **Monthly Income** | |
|  | |  |
| N = | | 120,269 |
| Missing = | | 29,271 |
| Mean = | | 6,670.22 |
| Median = | | 5,400.00 |
| Standard Deviation = | | 14,384.67 |
| Min, Max = | | (0, 3,008,750) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Open Credit Lines and Loans:**

|  |  |
| --- | --- |
| **FIGURE 8** | **Number of Open Credit Lines and Loans** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_of_open_credit_lines_and_loans_histogram1_5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 8** | **Number of Open Credit Lines and Loans** | |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 8.45 |
| Median = | | 8.00 |
| Mode = | | 6.00 |
| Standard Deviation = | | 5.15 |
| Min, Max = | | (0, 58) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Times 90 Days Late:**

|  |  |
| --- | --- |
| **FIGURE 9** | **Number of Times 90 Days Late** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_of_times90_days_late_histogram2_1.5.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 9** | **Number of Times 90 Days Late** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 0.27 |
| Median = | | 0.00 |
| Standard Deviation = | | 4.17 |
| Min, Max = | | (0, 98) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Real Estate Loans or Titles:**

|  |  |
| --- | --- |
| **FIGURE 10** | **Number of Real Estate Loans or Titles** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_real_estate_loans_or_lines_histogram2_1.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 10** | **Number of Real Estate Loans or Titles** | |
|  | |  |
| N = | | 150,000 |
| Missing = | | 0 |
| Mean = | | 1.02 |
| Median = | | 1.00 |
| Standard Deviation = | | 1.13 |
| Min, Max = | | (0, 54) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Number of Dependents:**

|  |  |
| --- | --- |
| **FIGURE 11** | **Number of Dependents** |
| Macintosh HD:Users:Joshua:Desktop:jmausolf:pa2:Histograms:number_of_dependents_histogram2_1.jpg | |

|  |  |  |
| --- | --- | --- |
| **TABLE 11** | **Number of Dependents** | |
| N = | | 146,076 |
| Missing = | | 3,924 |
| Mean = | | 0.76 |
| Median = | | 0.00 |
| Standard Deviation = | | 1.12 |
| Min, Max = | | (0, 20) |
| *Source: cs-training.csv* | | |

1. **Summary Data for Logistic Prediction Classifier for Serious Delinquency:**

|  |  |  |  |
| --- | --- | --- | --- |
| **TABLE 12** | **Logistic Regression Classifier Predicting Likelihood of**  **Serious Delinquency Last Two Years** | | |
| Serious Delinquency Last Two Years | | | |
| Variable | | Odds Ratio | P>|z| |
| Age | | 0.951\*\*\* | 0.000 |
| Debt Ratio | | 0.999\* | 0.022 |
| Monthly Income | | 0.999\*\*\* | 0.000 |
| Number of Dependent | | 1.033\*\*\* | 0.000 |
| Number of Open Credit Lines and Loans | | 0.979\*\*\* | 0.000 |
| Number of Times 30-59 Days Past Due Not Worse | | 1.634\*\*\* | 0.000 |
| Number of Times 60-89 Days Past Due Not Worse | | 0.413\*\*\* | 0.000 |
| Number of Times 90 Days Late | | 1.523\*\*\* | 0.000 |
| Number of Real Estate Loans or Lines | | 1.106\*\*\* | 0.000 |
| Revolving Utilization of Unsecured Lines | | 0.999 | 0.373 |
| Psuedo = | | 0.069 | |
| Log-Likelihood = | | -34,276.00 | |
| Mean Classifier Accuracy =  (10-Fold Cross Validation) | | 93.38% | |
| N = | | 150,000 | |
| Source : | | cs-training.csv | |
| *Note: +*p<0.1, \*p<.05, \*\*p<.01, \*\*\*p<.001.  *T-test values.* P > | Z |  Logistic Model with MLE using 10-Fold Cross Validation for Mean Classifier Accuracy | | | |