# Statistical analysis for quick stab collection agency

Statistical Analysis for Quick Stab Collection Agency Executive Summary The purpose of the paper is to provide a statistical analysis of overdue bills for Quick Stab Collection Agency (QSCA). The data will be taken from accounts closed over a six month period. The goal is to determine if a correlation between the type of account, the amount of the bill and the days to collection exists. To determine the existence of a correlation, regression analysis of several variables was completed. This regression analysis also included predictions. Further study also included descriptive statistical analysis, together with graphs.

This analysis will show that the correlation exists between the type of account and the days to collection. It will also show that the dollar amount the bill did not play a significant role in the days until collection. Introduction Quick Stab Collection Agency (QSCA) is a bill collection agency specializing in small less risky accounts. QSCA buys the rights to collect debts from the original owner of the debt at a significant discount. The right to collect a $50 debt may be purchased for as little as $10.

This example would indicate a profit of $40 not accounting for costs of doing business. Based on this example QSCA will need to be selective when purchasing debt as there the potential for profit loss with non-payment. In order for QSCA to remain successful they must collect the greatest amount of payments in the shortest number of days. In an effort to accomplish this task, a review of 96 accounts which were closed out in the months of January through June was completed. This review provided the data set which will be used in this analysis. This statistical analysis will attempt to answer the following questions: ? What is the average dollar amount for each account

type over the six month period? ? What is the average collection time for each bill type? ? Is there a correlation between the size of bill and the number of days until collection? ? Does the type of bill matter in terms of days until collection? Data A random sample of accounts closed out during the months of January through June was provided.

From this file titled OVERDUE (EXHIBIT A), we obtained the following: ? The number of days to collect the payment. The number of days to collect the payment varied greatly across both account types. The minimum days until collection was 5 days and the maximum days until collection was 99 days. ? The amount of the overdue bill oThe amount of bills also varied greatly across both account types. The minimum amount due was $46 and the maximum amount due was $311 ? Commercial account or Type 0 oThere were 48 commercial accounts present during the six month period analyzed ? Residential account or Type 1. oThere were 48 residential accounts present during the six month period analyzed.

This created an equal representation for the two account types present. Results ? The average bill for both residential and commercial accounts was $174. 27 during the months of January through June. oThe significance of the average being equal for both account types remains undetermined as the variables present were inadequate. ? The statistical model (Exhibit B) of the relationship between the type of bill, days late and the amount of the bill indicates that the type of account is statistically relevant to the days until collection.

oAccount Type 0 higher relevance then Account Type 1 The statistical model indicates the following: oThe residential account or Type 1 took an average of 31 days to collect. oThe commercial account or Type 0 took an average of 68 days to collect. oThe size of the bill does not relate to the number of days the bill is late. oType 0 accounts take longer to collect regardless of the dollar amount of the bill. Generally speaking, the statistical model (Exhibit B) used would be considered valid as 63.

5% of the variation in days is explained. As Exhibit B depicts a standard error of over 14%, another analysis with more variables ay provide a more accurate result. A further study with added variables broken down into more data elements may significantly increase or decrease the variation in days. Recommendations The findings of this statistical analysis indicate that account Type 1 can be collected in an average of 31 days, regardless of the dollar amount of the outstanding bill. This study further indicated that the Type 0 accounts have more days until collection when the dollar amount is smaller.

The larger the bill due in Type 0 accounts the payment time decreased by one to two days. Based on this information it is recommended that QSCA increase the purchase of Type 1 accounts as these accounts generally pay faster regardless of the amount due. It is also recommended that Type 0 having a value over $300 be purchased, as the analysis revealed that a bill with a value of $310 would be paid in sixty-six days. The purchase of these specific Type 0 accounts should be purchased for a period of six months, at which time another regression analysis should be completed. For example

(Exhibit C), a residential account or Type 1 bill in the amount of $310 will have a range of 23 to 36 days until collection.

A commercial account or Type 0 bill in the amount of $310 will have a range of 60 to 73 days until collection. Limitations This statistical analysis was limited by the data elements included in the study. The four variables presented did not provide enough information to test several anomalies present in the regression analysis. The regression analysis indicated a standard of error greater then 14% and an unusually high p-value. This analysis would benefit from the addition of more variables. The addition of variables would allow a more accurate study.

These variables should also being broken down into several data elements. Some suggestions for variables for future analysis include the discount for the bill purchased and a variety of account demographics. These demographics might include credit status, geographic location, bankruptcy/foreclosure, net worth or net income. This study should also be for a longer time period. The additional variables would identify the greatest possible accounts with the shortest time for collection.

EXHIBIT A Descriptive statistics DAYS BILL TYPE count96 96 96 mean49. 78 174. 27 0. 50 ample variance557. 86 6, 056. 58 0.

25 sample standard deviation23. 62 77. 82 0. 50 minimum5 46 0 maximum99 311 1 range94 265 1 population variance552. 05 5, 993.

49 0. 25 population standard deviation23. 50 77. 42 0. 50 standard error of the mean2.

41 7. 94 0. 05 confidence interval 95. % lower45. 00 158. 50 0.

40 confidence interval 95. % upper54. 57 190. 04 0. 60 half-width4. 79 15.

77 0. 10 1st quartile32. 50 97. 75 0. 00 median47.

50 170. 50 0. 50 3rd quartile69. 00 221. 25 1.

00 interquartile range36. 50 123. 50 1. 00 mode47. 00 205.

00 1. 00 EXHIBIT B Regression AnalysisR? 0. 635 Adjusted R? 0. 627 n 96 R 0. 797 k 2 Std. Error 14.

421 Dep. Var. DAYS ANOVA table SourceSS df MSFp-value Regression 33, 654. 6771 2 16, 827. 3385 80.

914. 41E-21 Residual 19, 341. 7292 93 207. 9756 Total 52, 996. 4063 95 Regression outputconfidence interval variables coefficientsstd. error t (df= 93)p-value95% lower95% upper Intercept70.

6818 3. 9129 18. 064 3. 60E-3262. 9116 78.

4520 BILL-0. 0126 0. 0190 -0. 665 . 5078-0. 0504 0.

0251 TYPE-37. 3958 2. 9437 -12. 703 4. 92E-22-43. 2415 -31.

5501 EXHIBIT C Regression Analysis R? 0. 635 Adjusted R? 0. 27 n 96 R 0. 797 k 2 Std. Error 14. 421 Dep.

Var. DAYS ANOVA table SourceSS df MSFp-value Regression 33, 654. 6771 2 16, 827. 3385 80. 914. 41E-21 Residual 19, 341.

7292 93 207. 9756 Total 52, 996. 4063 95 Regression outputconfidence interval variables coefficientsstd. error t (df= 93)p-value95% lower95% upper Intercept70. 6818 3.

9129 18. 064 3. 60E-3262. 9116 78. 4520 BILL-0. 0126 0.

0190 -0. 665 . 5078-0. 0504 0. 0251 TYPE-37.

3958 2. 9437 -12. 703 4. 92E-22-43. 2415 -31. 5501 Predicted values for: DAYS 95% Confidence Intervals 95% Prediction IntervalsBILLTYPE Predicted lower upper lower upper Leverage 175 0 68.

470 64. 336 72. 604 39. 535 97. 405 0.

021 175 1 31. 074 26. 940 35. 208 2. 139 60.

009 0. 021 Regression Analysis R? 0. 635 Adjusted R? 0. 627 n 96 R 0. 797 k 2 Std.

Error 14. 421 Dep. Var. DAYS ANOVA table SourceSS df MSFp-value Regression 33, 654. 6771 2 16, 827. 3385 80.

914. 41E-21 Residual 19, 341. 7292 93 207. 9756 Total 52, 996. 4063 95 Regression outputconfidence interval variables coefficientsstd.

error t (df= 93)p-value95% lower95% upper Intercept70. 6818 3. 9129 18. 064 3. 60E-3262. 9116 78.

520 BILL-0. 0126 0. 0190 -0. 665 . 5078-0.

0504 0. 0251 TYPE-37. 3958 2. 9437 -12. 703 4.

92E-22-43. 2415 -31. 5501 Predicted values for: DAYS 95% Confidence Intervals 95% Prediction Intervals BILLTYPE Predicted lower upper lower upper Leverage 210 0 68. 028 63. 680 72.

376 39. 061 96. 994 0. 023 210 1 30. 632 26.

284 34. 980 1. 666 59. 598 0. 023 Regression Analysis R? 0.

635 Adjusted R? 0. 627 n 96 R 0. 797 k 2 Std. Error 14. 421 Dep. Var.

DAYS ANOVA table SourceSS df MSFp-value Regression 33, 654. 6771 2 16, 827. 3385 80. 914. 41E-21 Residual 19, 341.

292 93 207. 9756 Total 52, 996. 4063 95 Regression outputconfidence interval variables coefficientsstd. error t (df= 93)p-value95% lower95% upper Intercept70. 6818 3. 9129 18.

064 3. 60E-3262. 9116 78. 4520 BILL-0. 0126 0.

0190 -0. 665 . 5078-0. 0504 0. 0251 TYPE-37.

3958 2. 9437 -12. 703 4. 92E-22-43. 2415 -31. 5501 Predicted values for: DAYS 95% Confidence Intervals 95% Prediction Intervals BILLTYPE Predicted lower upper lower upper Leverage 310 0 66.

764 60. 180 73. 347 37. 379 96. 149 0.

053 310 1 29. 368 22. 784 35. 952 -0. 017 58.

753 0. 053