**Analysis**

**Overview**

This note provides the findings from two analytic exercises based on the survey of Jordanian businesses on their experience with government inspectors. The first analysis covers only the General Amman Municipality (GAM) based inspections. The GAM data is singled out because GAM conducted the most inspections by far and, thus, deserved to be examined in a separate study. The other analysis was based on the aggregate dataset of all inspected businesses covered by the survey.

**Data and logical framework**

The GAM dataset contained several factors that were potentially significant in explaining the quality of inspections: the average number of inspection visits in 2019, number of concerned employees involved in the inspection, the time spent by the employees in the inspections, whether the business was aware of the inspection procedures, and of their legal rights, whether the inspection was a surprise visit, whether the inspector was aware of the inspection procedures, and any signs of inspector bias. These were treated as independent variables in the analysis. While the first three variables were discrete numbers the other five were binary in nature with Yes/No as the responses. The dependent variables which reflected the quality of inspections were whether the business was satisfied with the inspector’s performance and conduct and the business perception on whether the inspection process had improved or not in the recent past. The first dependent variable was binary, with Yes/No as the responses while the second was ordinal, with a score of 1 if the quality of the process had declined, 2 if it had remained the same and 3 if it had improved.

The aggregate dataset, i.e., dataset on all inspections, also had several potentially significant factors, which were defined as changes in the business awareness of their legal rights and of inspection procedures and their obligations, the inspector’s awareness of the technical aspects, and the frequency and duplication of visits. The two variables which were a proxy for inspection quality were inspector performance and conduct, and the fairness of their decisions. The two models used to analyze the aggregate dataset used these two as dependent variables. All variables were ordinal in nature, with a score of 1 if the quality of the process had declined, 2 if it had remained the same and 3 if it had improved.

**Models Chosen**

|  |  |  |
| --- | --- | --- |
| **Independent Variables** | **Dependent Variables** | **Dataset** |
| Employee Working Time + Inspector Awareness | Satisfaction | GAM |
| Average number of visits + Satisfaction | Change | GAM |
| Performance + Frequency | Fairness | Aggregate |
| Legal Awareness + Procedure Awareness | Performance | Aggregate |

**Analytic approach**

Correlation analysis

First, a set of correlation analysis was conducted to examine the relationships between the factors in both datasets (*Figures A1 and A2 in Appendix)*. To determine the level of significance, the range for a strong positive or negative correlation was when the correlation was either less than -0.70 or more than 0.70. The GAM dataset revealed slightly more positive relationships than negative ones. The only pair of variables that had a significant degree of correlation was the one between the average number of inspection visits and amount of employee time spent on inspections (0.71). The next potential strong relationships between variables was between the business awareness of their legal rights, and their awareness of inspection procedures and their obligations (0.66) and the inspector awareness of technical aspects and business satisfaction with inspections (0.50).

For the aggregate dataset, there were no negative relationships between factors nor was there multiple strong relationship between variables. The strongest relationship was between the changes in inspector awareness of the technical aspects and changes in their performance and conduct (0.76). After that relationship, the next two strongest ones was between changes in the business awareness of their legal rights and changes in business awareness of inspection procedures and their obligations (0.68) and changes in inspector’s performance and conduct and frequency of inspections (higher score here means less frequent inspections (0.51).

Model selection

The “all best subsets” command was used to find the three best models to analyze via a logistic regression. To determine which models moved further, the analysis revealed which had the highest adjusted R square figures and the lowest scores on the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Amemiya Prediction Criteria (APC).

Logistic regression

The logistic regression used the p-value of each factor to determine which model should be the best for the prediction stage. The best prediction model was always the one, where all p-values were less than 0.05. Once the best logistic regression model was chosen, I created a series of predictions and compared it to the actual data to check for the reliability of it.

**Findings**

GAM data

*Model 1*

The first model on the GAM dataset showed that using a logistic regression with the “total employee working time during inspections” and “inspector awareness of the business and legal rights and procedures” would be the best model to accurately predict the level of satisfaction that the business had about the inspection (*Figure A3 in Appendix)*. This was because the best subset models showed that the second, third and fourth models had the highest adjusted r squared as well as the lowest AIC, BIC, and APC. The following independent variables were included in these three models: total employee working time and the inspector awareness (2); total employee working time, the inspector awareness, and the business awareness (3); and total employee working time, the inspector awareness, the legal awareness, and the business awareness (4) (Figure 1).

When all three models were tested through logistic regression, it was discovered that the second model was the best because all the p-values were significant (Figure 1) and had a low AIC value of 90.319. As shown in Figure 1 below, this logistic regression indicates that the level of business satisfaction with the inspection process improves significantly with improvements in inspector awareness of the technical aspects of the inspection process (at 99.9% level of significance) and improvement (i.e. reduction) in the employee time spent in dealing with inspections (at 95% level of significance).

*Figure 1*

Table

Description automatically generated

**Text

Description automatically generated**

This model is also accurate enough to predict the satisfaction of future inspections (*Figure A4 in Appendix)*. When I ran the logistic regression and took the distribution of satisfied business to not, I discovered that 85 businesses were happy with the inspection. This is close to the actual data that revealed that 81 businesses were happy and 21 was not.

*Model 2*

In this model, the dependent variable was the perceived change in the inspection process over the past few years. This variable had three possible values: 1 (negative, i.e., worsening of the process), 2 (neutral, i.e., no change), 3 (positive, i.e., improvement). This model, which was identified through the model selection process as explained above, included the average number of visits and whether the businesses were satisfied with the inspection as the independent variables (*Figure 2)*. The logistic regression showed that there was an indirect relationship between the average number of visits (coefficient value of -0.02; significant at the 90% level)) and a direct relationship between whether the businesses were satisfied with the inspection and change of opinion between inspection (coefficient value of 0.7821; significant at the 99.9% level)).

Figure 2

**Text

Description automatically generated**

Like the previous model that used whether the businesses were satisfied with the inspection as a dependent variable, this model was also accurate (*Figure A6 in Appendix)*. The prediction model showed that 78 businesses felt that the current inspection did not change from the previous one. This is like the actual dataset, where 79 businesses did not feel that anything changed.

***Conclusions from GAM analysis***

The analysis was developed in several stages. Several steps were taken to ensure the model was accurate. Several prediction criteria were used to help choose the best regression model for this study. The models that were chosen via logistic regression seem to be correct as the predictions almost matched the actual dataset. Therefore, it is fair to assume that in order to improve the number of satisfied businesses, the best factors to look at would be the total time that employees spend during the inspections and the level of awareness that the inspector has about the technical aspects of inspection.

The first model analyzed in the GAM dataset indicated that the amount of time that employees spent in the inspection had an adverse relationship with whether the business was happy with the inspection, while the inspector’s level of awareness of the technical aspects of inspection had a direct relationship with business satisfaction. This means that a good way to improve business perception of the quality of inspection is to reduce the number of visits so that less employee time is involved with the inspection. Also, a good way to have a more positive inspection experience is to ensure that the inspectors are adequately trained so that they are more aware of the technical aspects of inspection.

The second model on the GAM dataset showed that the average number of visits per business and whether the business were satisfied were the best areas to focus on if they want to make the next inspection better than the last. While the average number of visits has an adverse relationship with the change of feelings between inspections, whether the businesses were satisfied had a direct relationship with the change factor. Therefore, increasing the number of visits would worsen the businesses perception of the inspection and vice versa. Also, inspectors should ensure that the businesses were satisfied if they want to create more positive inspections. One way to improve the satisfaction levels is to look at the first model, which illustrates how inspectors can increase the number of satisfied businesses.

Aggregate inspection data

*Model 1*

I chose to use the fairness of the inspection as the dependent variable in the first model that analyzed the aggregate data (*Figure A7 in Appendix)*. After running multiple analysis, I discovered that the best logistic model to visualize trends in the fairness of the inspection as well as the right model to base predictions on was one where the performance and frequency of the inspections was the independent variables (*Figure 3)*. Unlike the GAM datasets, this model is not accurate as the model predicts that there would be 88 inspections that would be deemed fair (*Figure A8 in Appendix)* and the actual dataset has 128 fair ones.

Figure 3

**Text

Description automatically generated**

*Model 2*

The 2nd model focused on the performance of the inspector during the inspection and through logistic regression, determined that the factors researchers should look at where the legal awareness and knowledge of inspection procedure from both the business and the inspector (*Figure A9 in Appendix)*. The second model was far more reliable, when it came to predicting whether the inspector performed positively (*Figure 4)*. When I ran the prediction analysis, I ended up with 80 inspections, where the inspector’s performance during the inspection improved. This was close to the actual number of 78 inspections as such *(Figure A10 in Appendix)*.

Figure 4

Text

Description automatically generated

***Conclusions from aggregate inspection data analysis***

Both factors in the first model has a direct relationship to the fairness of the inspection and that means that improving the performance and decreasing the frequency of the inspections would create more fair inspections, theoretically. However, the limited accuracy that I saw caused me to determine that this model is not good for predictions and further analysis must be made.

I found out that the second model was like the first, in the sense that both factors had a direct relationship with the performance of the inspector. My conclusion from the second model is that if you want the performance of the inspector during the inspection to increase, you should make both parties more aware of the legal possibilities and the inspection procedures. Not only did I find correlation, I also concluded that this is a reliable model to use to predict the number of inspections where the performance of the inspector has improved as the predicted data came close to the actual one.

**Implications**

If a similar study were to be repeated to decipher even more information, one area to focus on would be situational probability. Here, a decision tree could be used to both verify the results from this study as well as create even more accurate models.

**Appendix**

**Figure A1: GAM data: Correlation Matrix**

Chart, scatter chart

Description automatically generated

**Table

Description automatically generated**

**Figure A2: Aggregate data: Correlation Matrix**

**Chart

Description automatically generated**

A picture containing table

Description automatically generated

**Figure A3:**

**Table

Description automatically generated**

**Figure A4:**

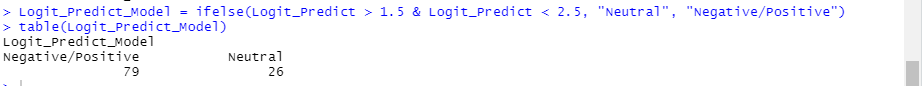
**A picture containing text

Description automatically generated**

**Figure A5:**

**Table

Description automatically generated**

**Figure A6:**

**Figure A7:Table

Description automatically generated**

**Figure A8:**

**Text

Description automatically generated**

**Figure A9:Table

Description automatically generated**

**Figure A10:**

**Text

Description automatically generated**