# **Survey Report (Waste Disposal System)**

We surveyed random sampling (In Lucknow) to collect the primary data and secondary data collected from the govt website and Google search. Based on the primary and secondary data of 500 households in the city we try to find out the KPI.

Here satisfaction level due to Govt. The waste disposal system is a Dependent Variable (DV) as the NGO has to find out the utilization of this system. We take other Variables as Independent variables (IV).

**Regression Analysis:**

Sample size = 500

Null Hypothesis (H0): The Model is insignificant.

Alternative Hypothesis (H1): The Model is significant.

ANOVA Test:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 11 | 456.3273557 | 41.48430506 | 123.694 | 2.0556E-133 |
| Residual | 488 | 163.6646443 | 0.33537837 |  |  |
| Total | 499 | 619.992 |  |  |  |

From the ANOVA table, the Significance value is 2.0556E-133<0.05 (p-value). So Null Hypothesis is rejected therefore Regression model is significant.

**Model Strength:**

|  |  |
| --- | --- |
| *Regression Statistics* | |
| Multiple R | 0.857916873 |
| R Square | 0.736021361 |
| Adjusted R Square | 0.730071023 |
| Standard Error | 0.579118614 |
| Observations | 500 |

From the R Square value (Explained Variation/Total Variation) we can say that the regression model is good or 73.6% of the variation in DV can be explained by the IVs.

There may be some predictors which have an impact on the regression model. They are

* Issue in the frequency of collecting waste.
* Govt. Provide a separate dustbin to segregate waste.

Now every IV is not a good predictor, we have to find out the most significant predictors/KPIs.

Most of the IVs are categorical data so to make a regression model we create a Dummy variable to convert it into numeric data.

**Significance of KPIs:**

We took a Hypothesis for each IV to find out the most significant IVs.

Null Hypothesis (H0): Coefficient is insignificant (β=0)

Alternative Hypothesis (H1): Coefficient is significant (β≠0).

**Coefficient Table:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept | -0.338637199 | 0.181532171 | -1.865439043 | 0.062719732 |
| X Variable 1 | -0.065272963 | 0.023839223 | -2.738049135 | 0.006406337 |
| X Variable 2 | 0.198822672 | 0.08491444 | 2.341447137 | 0.019610601 |
| X Variable 3 | 0 | 0 | 65535 | #NUM! |
| X Variable 4 | 0.688528445 | 0.106857032 | 6.443454701 | #NUM! |
| X Variable 5 | 1.590625665 | 0.100938527 | 15.75836021 | 1.58907E-45 |
| X Variable 6 | 0.365679926 | 0.053517682 | 6.832880482 | 2.48895E-11 |
| X Variable 7 | 1.925904011 | 0.078533572 | 24.5233213 | 3.86523E-87 |
| X Variable 8 | 0.6293941 | 0.08919494 | 7.056387964 | 5.89331E-12 |
| X Variable 9 | 0.784106084 | 0.088435135 | 8.866454354 | 1.42363E-17 |
| X Variable 10 | -0.118202932 | 0.059648938 | -1.981643516 | 0.048079845 |
| X Variable 11 | -0.085859453 | 0.081935541 | -1.047890222 | 0.295207897 |

From Coefficient Table we can clearly see Variable1 (HH Member), Variable5 (Dummy4 Door to door), Variable6 (Dummy5 Yes(segregate)), Variable7 (Dummy6 (Everyday)), Variable8 (Dummy7 (Once Every Two Day)), Variable9 (Dummy8 (Morning)), Variable10 (Dumm9 (Yes (Awareness))) these all IVs are significant as p-value of all these IVs are less than 0.05.

**Hierarchy of KPIs:**

Among these 7 significant variables, we then identified the ranking of the variables (KPIs). The highest rank means it is highly impacting our dependent variable. This is identified by seeing the values of the beta coefficients. The higher the beta higher the impact on DV based on that concept we make a hierarchy of IVs. So, according to it, we identified the top 3 KPIs i.e. Key Performing Indicators which are Variable 7 (Dummy6(Everyday) **[Rank 1]**, then Variable 5(Dummy4 (Door to door)) **[Rank 2]** and Variable 9(Dummy8(Morning)) **[Rank 3].**

**Assumptions:**

1. It is assumed that the sample size of 500 gives an inference about the population.
2. The sampling we take during the survey is random sampling.
3. The respondent gives a response with honesty and without any bias.
4. There may be errors in data as respondents are not interested during the survey.
5. The survey questions are reliable and valid.
6. Linear relationship between DV and coefficients of IVs.
7. There is the independence of observation of each household.
8. Presence of Homogeneity of variance.
9. The residuals from the regression model are assumed to be normally distributed.
10. The IVs are not highly correlated with each other.
11. Proper conversion of the categorical variable into a dummy variable.
12. The confidence interval for this regression model is 95%.
13. A higher R square value indicates a better model fit.

Links Related to Data:

* [Better Waste Management in Lucknow](https://www.hindustantimes.com)​([Hindustan Times](https://www.hindustantimes.com/cities/lucknow-news/better-waste-mgmt-lmc-to-set-up-10-material-resource-recovery-centres-101668972573110.html))
* [CPCB Annual Report](https://cpcb.nic.in/uploads/MSW/MSW_AnnualReport_2020-21.pdf)​([Central Pollution Control Board](https://cpcb.nic.in/uploads/MSW/MSW_AnnualReport_2020-21.pdf))

**Recommendation (Based on 3 important KPIs):**

The 3 important KPIs are

1. Frequency of Waste Collection (every day).
2. Disposal Method (Door to door).
3. Time of Disposal (Morning).

**1. Frequency of Waste Collection (Everyday)**

* **Issue Identified**: Households with everyday waste collection report higher satisfaction compared to those whose waste is collected less frequently. This was identified as the most significant factor influencing satisfaction levels.

**Recommendation:**

* Use technology like GPS to make the waste collection process more efficient.
* Implement a digital alarm system (Via Mobile app, SMS) to remind residents of the daily collection schedule to ensure they do not miss collection times.
* Increase daily collection coverage.

**2. Disposal Method (Door to door):**

* **Issue Identified**: Households with access to door-to-door collection are more satisfied with the waste management system. This service was found to be crucial for overall satisfaction, as it offers convenience and reduces the need for residents to transport waste to central collection points or public bins.

**Recommendations:**

* Regular and punctual collection is key to maintaining satisfaction. Ensure that waste collection teams adhere to strict schedules for door-to-door pickups.
* Increase the number of dustbins or size of dustbins where households through the waste.
* Implement sensor-equipped waste bins to alert collection teams when they are full, ensuring timely pickups.

**3. Time of Disposal (Morning):**

* **Issue Identified:** Households that dispose of waste in the morning report higher satisfaction, suggesting that the current collection schedules, which predominantly occur in the morning, are preferred by most residents.

**Recommendations:**

* Continue the emphasis on morning waste collection, as it aligns with the preferences of the majority of households.
* To accommodate households that may not be able to dispose of waste in the morning (due to work schedules, etc.), offer limited afternoon collection slots.
* Develop a mobile application that allows residents to report issues, track collection schedules, and receive updates.
* Create a reward system that provides discounts on municipal fees or other benefits for consistent participation in recycling programs.