

# Analysis of Graduate Student Data

Author: Godwin Festus (ExploreAI Academy Graduate Student)

---

**GitHub Link:** [[https://github.com/RaveEx12/community\\_intern\\_application.git](https://github.com/RaveEx12/community_intern_application.git)]

## Task 1: Data Cleaning Report

### Objective:

The purpose of Task 1 was to clean a dataset containing graduate data, which included duplicates and inconsistencies. The task aimed to identify and remove these duplicates to produce a clean, accurate dataset for further analysis.

### Approach:

We initiated the process by loading the dataset into a Pandas DataFrame and analyzing its structure. Our primary focus was to identify duplicates across the dataset. After identifying and examining over 300 duplicate records, they were successfully removed using Python's Pandas library. Additionally, we performed a quality check to ensure that there were no missing values or anomalies that might affect future analyses.

Upon completion of the cleaning, the dataset was saved in a new Excel file. The cleaned dataset, which now contains only unique graduate records, is ready for further analysis.

### Tools Used:

- Python (pandas): Used for data manipulation and cleaning.
- Excel: Used to save the cleaned version of the dataset.

### Findings:

- Over 300 duplicate entries were removed.
- No missing values or significant anomalies were found during the cleaning process.
- The dataset was successfully cleaned, reducing the risk of inaccuracies in future analyses.

This data cleaning process ensured that the dataset is now reliable for further insights, which are essential for community engagement planning.

---

## Task 2: Dashboard Update Request

### Initial Response

Thank you for assigning me the task to fix the broken dashboard.

Request for Clarification

To ensure accurate completion of the task, I require the following information:

### Essential Resources

1. Existing Dashboard: Access to the current dashboard (or a copy/mock-up)
2. Issue Clarification: Specific issues with the current dashboard
3. Relevant Documentation: Guidelines or documentation related to the dashboard

### **Task Objectives**

With the provided information, I will:

1. Update Dashboard: Refresh the dashboard with the new dataset, ensuring accuracy and no duplicates.
2. Identify Manual Update Issues: Detect potential problems with manual updates
3. Propose Automation Solutions: Develop a comprehensive report outlining automation strategies.

### **Deliverables**

I am committed to delivering:

1. Updated Dashboard: In Excel or a visualization tool
2. Brief Report: 300-500 words discussing manual update challenges and automation solutions.

Additional Information: Please notify me if there's any additional context or information required to complete this task.

---

## **Task 3: Data Insights Report**

### **Objective:**

The goal of Task 2 was to analyze the cleaned dataset to provide actionable insights for the community team, helping them plan more targeted outreach and engagement strategies.

### **Approach:**

Using the cleaned dataset, we conducted five key analyses to derive insights that would inform engagement strategies:

1. **Distribution of Graduates by Country:**

We analyzed the geographical distribution of graduates. Kenya emerged as the country with the highest number of graduates. This finding suggests that Kenya should be a focal point for outreach efforts, including local events and community engagement initiatives. The analysis supports region-based strategies for optimizing engagement.

2. **Popular Rewards and Brands:**

We explored which rewards and brands were most popular among graduates. Rewards like Amazon Gift Cards and brands like Uber and Coursera were among the most frequently redeemed. This indicates that digital and e-learning services are highly valued, and future campaigns should emphasize these offerings to maintain or increase engagement.

3. **Satisfaction Rating Analysis:**

Satisfaction ratings for the rewards were generally high, with a mean satisfaction score of 3.92 (on a scale of 1-5). However, the data also revealed that some rewards had slightly

lower satisfaction ratings. This suggests that while the community is largely satisfied, certain rewards could be optimized based on feedback to improve the overall experience.

4. Redemption Frequency:

The analysis showed that graduates redeemed rewards multiple times, with an average of 5.29 redemptions per user, ranging from 1 to 9. While many users are actively engaged, there is an opportunity to target less active users with personalized campaigns to increase their participation.

5. Cost vs. Satisfaction:

Interestingly, there was no strong correlation between the cost of a reward and the satisfaction it produced. This indicates that higher-cost rewards do not necessarily result in higher satisfaction, suggesting that user satisfaction is more dependent on the perceived value or utility of the reward rather than its monetary value.

Descriptive Statistics:

The descriptive statistics provided further clarity on the dataset:

- Mean reward value: \$48.35, with a range from \$10 to \$100.
- Mean satisfaction rating: 3.92, with a range from 3.01 to 4.96.
- Redemption frequency: Users redeemed rewards an average of 5.29 times.
- Mean cost per redemption: \$52.42, with a range from \$11 to \$99.

Recommendations:

- City-based Outreach: Prioritize engagement in cities with high graduate populations, such as Kenya, through local events or virtual meet-ups.
- Popular Rewards Focus: Promote the most popular rewards (Amazon Gift Cards, Uber) in outreach campaigns to maintain interest.
- Optimize Satisfaction: Gather user feedback to improve lower-rated rewards, ensuring that satisfaction levels remain high across the board.
- Engagement Campaigns for Less Active Users: Develop targeted campaigns to encourage less frequent redeemers to increase their participation.

Tools Used:

- Python (pandas and matplotlib) for data analysis and visualization.
- Excel for cleaned dataset storage.

To summarize, these findings provide a data-driven foundation for strengthening community engagement tactics, allowing the team to successfully target key groups, optimize rewards, and sustain high levels of user happiness.

The complete project's details are available in the GitHub repository at [\[https://github.com/RaveEx12/community\\_intern\\_application.git\]](https://github.com/RaveEx12/community_intern_application.git). The repository is configured for public view, so you can take a look for grading and assessments.