

Smart Water Usage in Urban Homes: Trend Analysis & Cost Optimization Using NYC Data (2013–2025)

Prepared By: GODI PAVAN
Submission For: Data Analyst Internship Challenge
Date: 21-06-2025

1. Executive Summary

This project addresses water usage inefficiencies in New York City urban homes using a public dataset of water consumption and billing (2013 – 2025). Through exploratory and statistical analysis, we identified borough-level patterns, usage anomalies, and cost disparities. The findings support prescriptive strategies to promote efficient water usage, lower utility costs, and assist city officials in decision-making.

2. Key Findings

- **High Usage Areas:** Bronx and Queens had the highest average monthly consumption among boroughs.
 - **Cost Disparities:** Some buildings pay over 2x per HCF due to higher charges despite average usage. Likely causes include inaccurate readings or outdated rate classes.
 - **Anomalies Detected:** Multiple records flagged for extreme values (negative charges, long billing periods >180 days).
 - **Billing Frequency Impact:** Longer service periods (>90 days) are linked to higher total consumption and unpredictable costs.
-

3. Recommendations

- **Install Smart Meters** in buildings with frequent 'Estimated' readings to reduce billing errors and encourage real-time tracking.
 - **Audit Buildings** with high cost-per-HCF to identify possible leaks, inefficiencies, or data/reporting issues.
 - **Shorter Billing Cycles** to enable quicker anomaly detection and prevent overuse.
 - **Targeted Education Campaigns** in high-usage boroughs to promote water-saving practices.
-

4. Tools & Methodologies Used

- Python (pandas, seaborn, sklearn)
 - Descriptive stats, outlier detection, regression modeling
 - Visualization: histograms, boxplots, heatmaps
-

5. Next Steps / Future Work

- Develop an interactive Power BI dashboard for borough-level monitoring
 - Incorporate seasonal weather data to correlate climate with usage
 - Simulate potential savings from applying smart-metering in high-cost buildings
-

6. Project Files Submitted

- Google Colab File(.py)
- PDF Report (this document)
- Power BI dashboard
- Cleaned dataset (.csv)