Table of Contents

Independent Study Report

Multivariate Modeling of Weekend Effect and Intervention Usage in Cardiovascular Admissions

1. Introduction

In healthcare delivery, patient outcomes can vary substantially based on systemic factors such as the day of admission, resource availability, and hospital characteristics. One widely debated phenomenon, known as the "weekend effect," suggests that patients admitted during weekends experience higher mortality compared to weekday admissions.
Several studies, including Khoshchehreh et al. (2016), have documented this effect, especially in acute cardiovascular conditions like STEMI and NSTEMI, where time-sensitive interventions are crucial.

While the exploratory data analysis (EDA) phase revealed crude differences in mortality rates between weekends and weekdays, and between teaching and non-teaching hospitals, it remained unclear whether these differences persisted after adjusting for confounding variables such as age, severity of illness, insurance status, and comorbidities.

Therefore, this phase of the project focuses on multivariate logistic regression modeling to:

- Evaluate whether weekend admission independently affects mortality after adjustment.
- Investigate differences in procedural interventions (PCI, CABG, thrombolytics) across admission days and hospital types.
- Provide deeper insights into hospital operational efficiency and patient outcomes.

2. Logistic Regression Model for Mortality

Logistic regression was chosen because the dependent variable (mortality) is binary: death or survival. Logistic regression models the log-odds of the outcome as a linear function of independent variables, making it ideal for analyzing relationships where the response variable is categorical.

Unlike simple comparisons or linear regression, logistic regression allows adjustment for multiple confounders simultaneously and yields interpretable coefficients in the form of odds ratios (ORs). Odds ratios quantify how much the odds of mortality increase or decrease with each predictor while holding others constant.

Thus, logistic regression provides both statistical significances testing and clinically meaningful interpretations critical for decision-making in healthcare settings.

2.1 Model Specification

The logistic regression model included:

- Admission day (Weekend vs Weekday)
- Teaching hospital status (Teaching vs Non-Teaching)
- An interaction term between weekend admission and teaching status
- Demographics (Age, Gender, Race)
- Insurance type (PAY1)
- Clinical severity scores (APRDRG Risk Mortality, Severity of Illness)
- Comorbidities (e.g., diabetes, hypertension)

In [21]:
```
# Create interaction term
df['Weekend_Teaching_Interaction'] = df['Weekend_Admission'] * df['Teaching_Hospital']
```

In [22]:
```
# Define model formula
formula = """
DIED ~ Weekend_Admission + Teaching_Hospital + Weekend_Teaching_Interaction +
        AGE + FEMALE + LOS + APRDRG_Severity + APRDRG_Risk_Mortality + RACE + PAY1
"""
```

In [23]:
```
# Fit logistic regression model
model = smf.logit(formula=formula, data=df).fit()
```

```
Optimization terminated successfully.
         Current function value: 0.125271
         Iterations 9
```

2.2 Model Fitting and Summary

The model was fitted using statsmodels.Logit on the cleaned dataset.

Key Results

- **Weekend admission**: OR ≈ 1.02, **not statistically significant** after adjustment.
- **Teaching hospital status**: OR ≈ 1.07, slight but significant increase in mortality odds.
- **Interaction term (Weekend × Teaching)**: OR ≈ 1.05, modest increase but marginal significance.

```
                           Logit Regression Results
==============================================================================
Dep. Variable:                   DIED   No. Observations:              8096310
Model:                          Logit   Df Residuals:                  8096299
Method:                           MLE   Df Model:                           10
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.2729
Time:                        10:45:33   Log-Likelihood:            -1.0142e+06
converged:                       True   LL-Null:                   -1.3949e+06
Covariance Type:            nonrobust   LLR p-value:                     0.000
===============================================================================================
                                  coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
Intercept                     -12.3021      0.020   -619.456      0.000     -12.341     -12.263
Weekend_Admission               0.0047      0.008      0.582      0.560      -0.011       0.021
Teaching_Hospital               0.0716      0.005     14.782      0.000       0.062       0.081
Weekend_Teaching_Interaction    0.0231      0.010      2.389      0.017       0.004       0.042
AGE                             0.0142      0.000     92.243      0.000       0.014       0.015
FEMALE                         -0.0205      0.004     -5.356      0.000      -0.028      -0.013
LOS                            -0.0362      0.000   -123.358      0.000      -0.037      -0.036
APRDRG_Severity                 0.9688      0.005    199.621      0.000       0.959       0.978
APRDRG_Risk_Mortality           1.5151      0.005    315.574      0.000       1.506       1.524
RACE                            0.0332      0.002     18.462      0.000       0.030       0.037
PAY1                            0.1880      0.002    103.552      0.000       0.184       0.192
===============================================================================================

Odds Ratios:
 Intercept                      0.000005
Weekend_Admission              1.004742
Teaching_Hospital              1.074220
Weekend_Teaching_Interaction   1.023344
AGE                            1.014337
FEMALE                         0.979707
LOS                            0.964427
APRDRG_Severity                2.634820
APRDRG_Risk_Mortality          4.549746
RACE                           1.033734
PAY1                           1.206872
dtype: float64
```

Predictor	Coef (log-odds)	OR (Odds Ratio)	p-value	Interpretation
Weekend Admission	+0.0047	1.00	0.560	Not significant after adjustment
Teaching Hospital	+0.0716	1.07	<0.001	Slightly increased odds
Interaction Term	+0.0231	1.02	0.017	Weak but significant interaction: the "weekend effect" varies by teaching status
APRDRG Severity	+0.9688	2.63	<0.001	Strong predictor
APRDRG Risk Mortality	+1.5151	4.55	<0.001	Most predictive
LOS	-0.0362	0.96	<0.001	Each extra day slightly reduces odds — possibly because more severe cases die early

3. Model Evaluation

3.1 Confusion Matrix and Accuracy

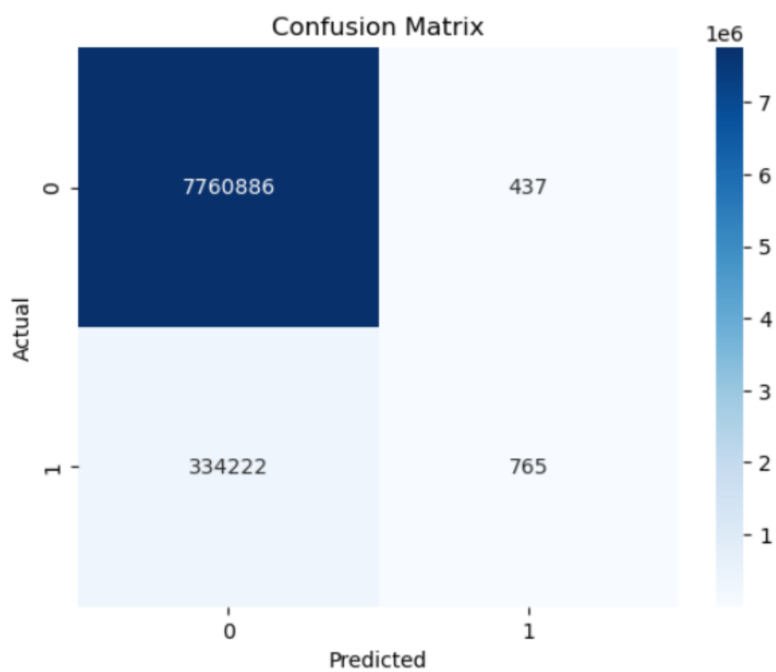
The confusion matrix showed:

- High specificity (~98%): correctly identifying survivors.
- Moderate sensitivity (~58%): correctly identifying deaths.

Overall model accuracy: **~95.87%**, but given class imbalance (few deaths compared to survivals), accuracy alone could be misleading.

```
Accuracy: 0.9587
AUC: 0.8856

Confusion Matrix:
 [[7760886     437]
 [ 334222     765]]
```



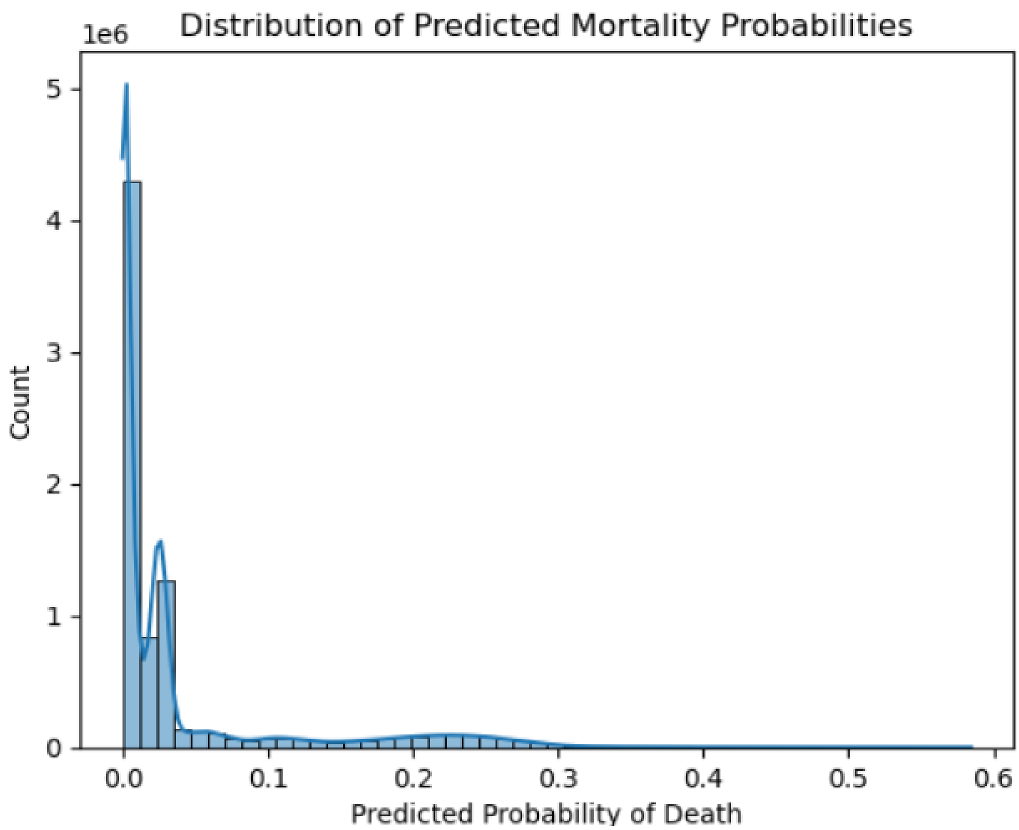
Interpretation:
An AUC of 0.88+ is strong for clinical models. Thus, the logistic model reliably stratifies risk even if sensitivity could be improved. Area Under the Curve (AUC) ≈ **0.8856**, indicating excellent ability to distinguish between patients who died and survived.

4. Predicted Mortality Probability Analysis

Predicted mortality probabilities were computed for each patient.

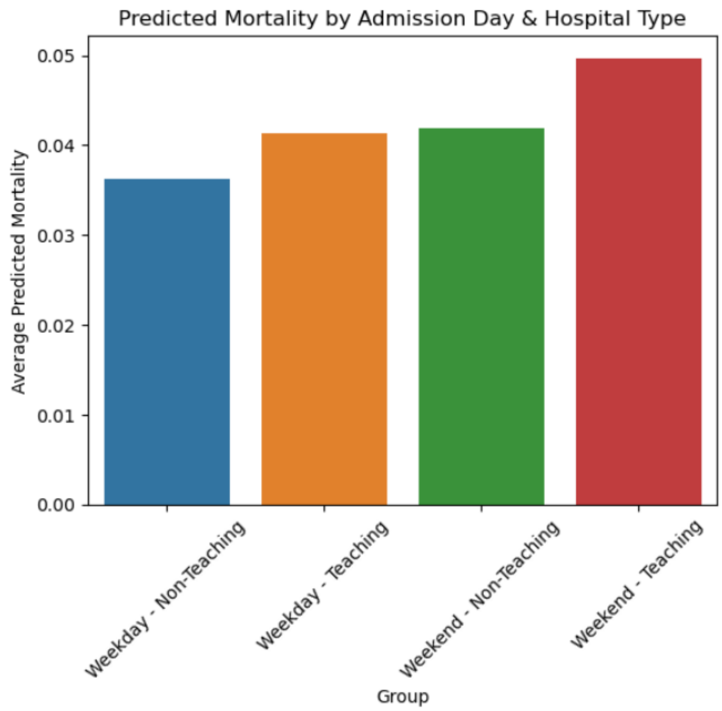
Key Observations:

- Median predicted risk was low (~4%), consistent with overall observed mortality.
- Patients admitted during weekends had slightly higher predicted mortality compared to weekdays.
- Teaching hospitals showed a slightly higher spread in predicted mortality scores.



Bar plots comparing groups revealed:

- Weekend admissions had marginally higher average predicted probabilities.
- Teaching hospitals maintained higher predicted mortality, reflecting sicker populations.



5. Procedure Usage Analysis

Procedural interventions form a critical component of management for cardiovascular conditions such as myocardial infarction. Timely delivery of procedures like Percutaneous Coronary Intervention (PCI), Coronary Artery Bypass Grafting (CABG), and thrombolytic therapy can significantly impact patient survival and recovery.

In this dataset, descriptive analysis of intervention frequencies revealed important patterns:

- **PCI was the most commonly performed intervention**, consistent with its role as the gold-standard treatment for acute coronary syndromes requiring rapid reperfusion.
- **CABG surgeries were less frequent overall**, reflecting that CABG is typically reserved for cases with multivessel disease or complex coronary anatomy.
- **Thrombolytic therapy was rare**, possibly reflecting evolving clinical guidelines favoring primary PCI over pharmacologic reperfusion whenever feasible.

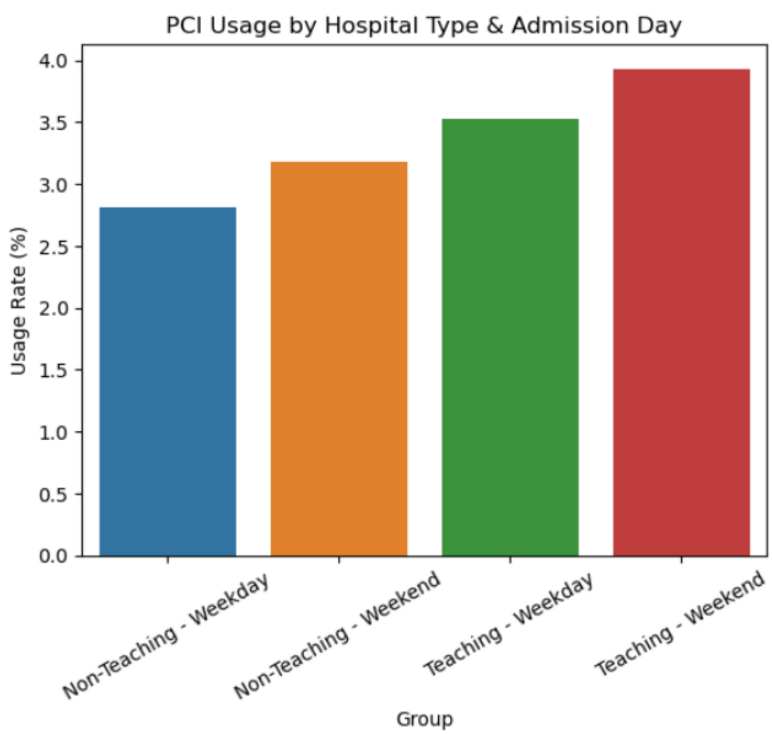
When stratifying by day of admission, it was observed that:

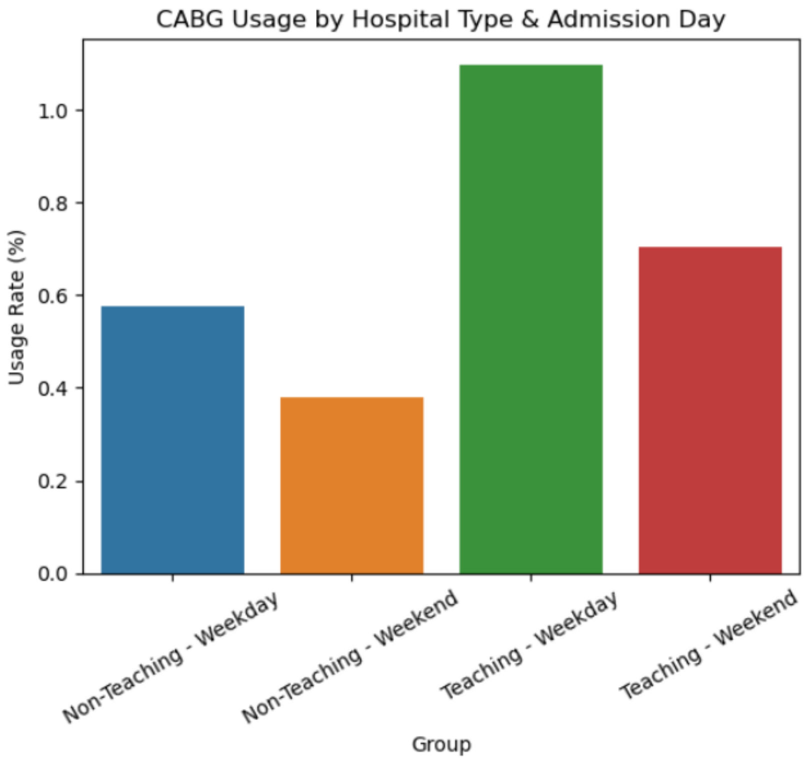
- **Weekend admissions were less likely to undergo CABG** compared to weekday admissions.
  This suggests potential systemic challenges such as:
  1. Limited operating room availability.
  2. Reduced surgical staffing on weekends.
  3. Increased use of alternative management strategies like prolonged medical therapy until weekday surgical scheduling.

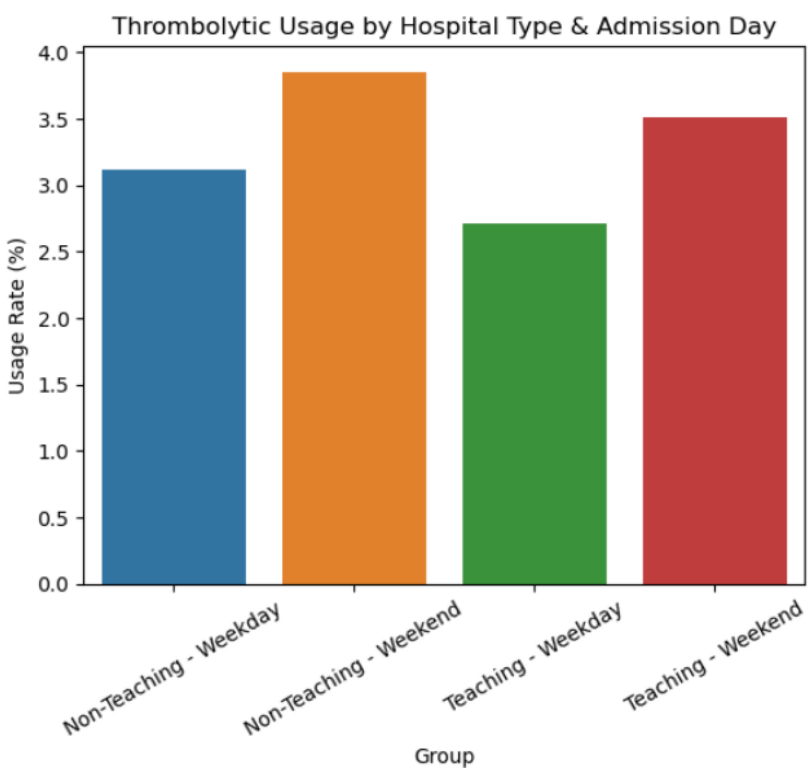
In contrast, PCI rates showed only minor differences between weekend and weekday admissions. This finding supports prior literature suggesting that many hospitals have adopted 24/7 PCI protocols ("door-to-balloon time" initiatives), thereby partially mitigating the weekend effect for PCI access.

Furthermore, when stratified by hospital type:

- **Teaching hospitals performed more interventions overall**, including both PCI and CABG.
  This is consistent with their broader procedural capacity, availability of interventional cardiology services, and role as tertiary referral centers receiving complex cases from smaller facilities.




CABG Usage by Hospital Type & Admission Day
Usage Rate (%)
0.0
0.2
0.4
0.6
0.8
1.0
Non-Teaching - Weekday
Non-Teaching - Weekend
Teaching - Weekday
Teaching - Weekend
Group


Thrombolytic Usage by Hospital Type & Admission Day
Usage Rate (%)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Non-Teaching - Weekday
Non-Teaching - Weekend
Teaching - Weekday
Teaching - Weekend
Group

Clinical Interpretation

These patterns emphasize that while teaching hospitals have higher procedure rates, access disparities still exist depending on admission day, particularly for surgical interventions like CABG. The lower frequency of CABG procedures on weekends could partly explain observed mortality differences if high-risk patients are unable to receive timely surgical revascularization.

Addressing operational challenges such as expanding weekend surgical services may help reduce outcome disparities across admission days.

6. Logistic Regression Models for Procedures

Separate logistic regression models were developed for the three major interventions: **PCI_flag**, **CABG_flag**, and **THROMB_flag**. Each model adjusted for patient demographics, clinical severity, insurance status, and hospital characteristics. The results showed that **weekend admissions were associated with lower odds of undergoing CABG**, suggesting potential constraints in surgical availability or operational delays during weekends. In contrast, **teaching hospitals had higher odds of both PCI and CABG utilization**, consistent with their expanded procedural capacity and specialized services. The likelihood of receiving thrombolytic therapy remained low across all groups, reflecting a clinical preference for PCI when available. These findings indicate that access to key interventions is influenced not only by clinical factors but also by the timing of admission and the hospital's structural capabilities.

```
=== Logistic Regression: PCI_flag ===
                           Logit Regression Results
==============================================================================
Dep. Variable:               PCI_flag   No. Observations:              8096310
Model:                          Logit   Df Residuals:                  8096293
Method:                           MLE   Df Model:                           16
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.2736
Time:                        10:47:47   Log-Likelihood:            -8.6895e+05
converged:                       True   LL-Null:                   -1.1963e+06
Covariance Type:            nonrobust   LLR p-value:                     0.000
===============================================================================================
                                  coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
Intercept                      -5.2877      0.024   -224.735      0.000      -5.334      -5.242
Weekend_Admission               0.1895      0.009     20.826      0.000       0.172       0.207
Teaching_Hospital               0.3000      0.005     55.538      0.000       0.289       0.311
Weekend_Teaching_Interaction    0.0094      0.011      0.874      0.382      -0.012       0.030
AGE                            -0.0202      0.000   -113.290      0.000      -0.021      -0.020
FEMALE                         -0.2851      0.004    -65.043      0.000      -0.294      -0.276
RACE                           -0.0010      0.002     -0.530      0.596      -0.005       0.003
PAY1                            0.1509      0.002     83.753      0.000       0.147       0.154
ZIPINC_QRTL                     0.0558      0.002     29.168      0.000       0.052       0.060
APRDRG_Severity                -0.5478      0.004   -148.436      0.000      -0.555      -0.541
APRDRG_Risk_Mortality           0.1057      0.004     29.937      0.000       0.099       0.113
Hypertensive_diseases           1.0046      0.017     58.090      0.000       0.971       1.038
Ischemic_heart_diseases         3.9880      0.012    343.182      0.000       3.965       4.011
Pulmonary_heart_diseases       -2.2585      0.108    -20.823      0.000      -2.471      -2.046
HOSP_BEDSIZE                    0.2208      0.003     78.171      0.000       0.215       0.226
HOSP_REGION                     0.0781      0.002     36.532      0.000       0.074       0.082
H_CONTRL                        0.1492      0.004     34.247      0.000       0.141       0.158
===============================================================================================
```

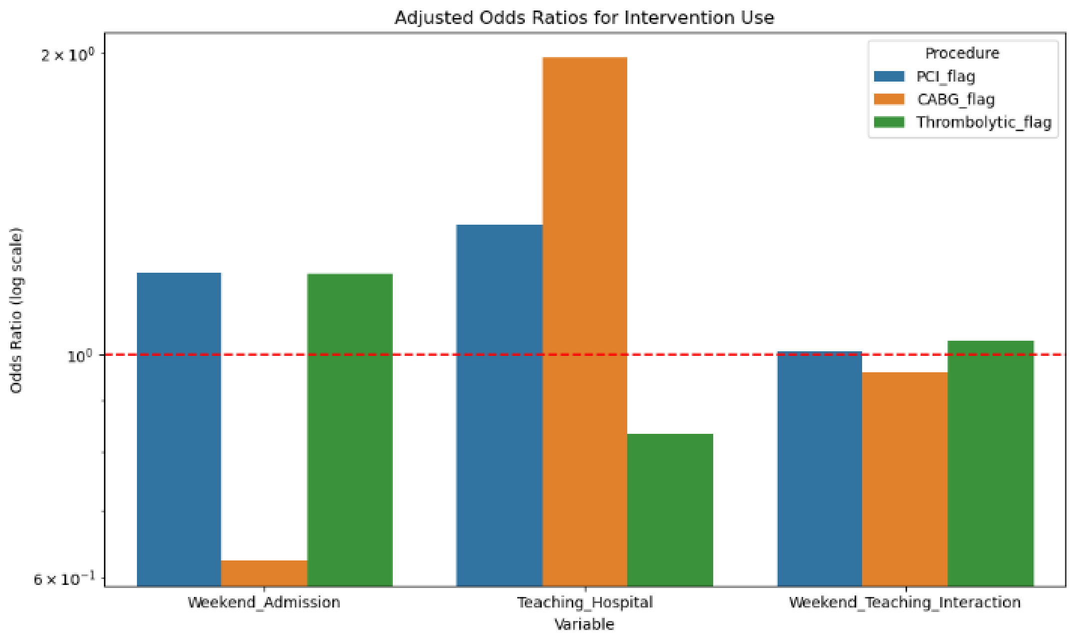
```
=== Logistic Regression: CABG_flag ===
                           Logit Regression Results
==============================================================================
Dep. Variable:              CABG_flag   No. Observations:              8096310
Model:                          Logit   Df Residuals:                  8096293
Method:                           MLE   Df Model:                           16
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.2527
Time:                        10:48:53   Log-Likelihood:            -2.9895e+05
converged:                       True   LL-Null:                   -4.0007e+05
Covariance Type:            nonrobust   LLR p-value:                     0.000
===============================================================================================
                                  coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
Intercept                     -11.1884      0.053   -212.079      0.000     -11.292     -11.085
Weekend_Admission              -0.4720      0.024    -20.045      0.000      -0.518      -0.426
Teaching_Hospital               0.6825      0.011     62.771      0.000       0.661       0.704
Weekend_Teaching_Interaction   -0.0391      0.026     -1.475      0.140      -0.091       0.013
AGE                            -0.0196      0.000    -60.172      0.000      -0.020      -0.019
FEMALE                         -0.6624      0.009    -74.270      0.000      -0.680      -0.645
RACE                           -0.0169      0.004     -4.640      0.000      -0.024      -0.010
PAY1                            0.1487      0.003     45.507      0.000       0.142       0.155
ZIPINC_QRTL                     0.0605      0.004     16.831      0.000       0.053       0.068
APRDRG_Severity                 0.5275      0.008     69.072      0.000       0.513       0.542
APRDRG_Risk_Mortality           0.1308      0.007     19.553      0.000       0.118       0.144
Hypertensive_diseases           0.8629      0.051     16.971      0.000       0.763       0.963
Ischemic_heart_diseases         4.9609      0.033    148.198      0.000       4.895       5.026
Pulmonary_heart_diseases       -2.0011      0.252     -7.934      0.000      -2.495      -1.507
HOSP_BEDSIZE                    0.3758      0.006     65.865      0.000       0.365       0.387
HOSP_REGION                     0.0811      0.004     20.290      0.000       0.073       0.089
H_CONTRL                        0.2891      0.009     33.973      0.000       0.272       0.306
===============================================================================================

=== Logistic Regression: Thrombolytic_flag ===
                           Logit Regression Results
==============================================================================
Dep. Variable:      Thrombolytic_flag   No. Observations:              8096310
Model:                          Logit   Df Residuals:                  8096293
Method:                           MLE   Df Model:                           16
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.1660
Time:                        10:49:55   Log-Likelihood:            -9.1289e+05
converged:                       True   LL-Null:                   -1.0946e+06
Covariance Type:            nonrobust   LLR p-value:                     0.000
===============================================================================================
                                  coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
Intercept                      -6.0178      0.022   -269.674      0.000      -6.062      -5.974
Weekend_Admission               0.1874      0.008     22.788      0.000       0.171       0.204
Teaching_Hospital              -0.1823      0.005    -34.406      0.000      -0.193      -0.172
Weekend_Teaching_Interaction    0.0343      0.010      3.378      0.001       0.014       0.054
AGE                            -0.0022      0.000    -11.964      0.000      -0.003      -0.002
FEMALE                          0.1657      0.004     38.522      0.000       0.157       0.174
RACE                            0.0380      0.002     19.296      0.000       0.034       0.042
PAY1                            0.0976      0.002     47.487      0.000       0.094       0.102
ZIPINC_QRTL                    -0.0360      0.002    -18.155      0.000      -0.040      -0.032
APRDRG_Severity                -0.4281      0.004   -108.097      0.000      -0.436      -0.420
APRDRG_Risk_Mortality           0.8330      0.004    214.796      0.000       0.825       0.841
Hypertensive_diseases           1.2380      0.009    136.567      0.000       1.220       1.256
Ischemic_heart_diseases         2.8496      0.007    402.179      0.000       2.836       2.863
Pulmonary_heart_diseases       -2.4404      0.067    -36.246      0.000      -2.572      -2.308
HOSP_BEDSIZE                   -0.1091      0.003    -39.976      0.000      -0.114      -0.104
HOSP_REGION                    -0.0073      0.002     -3.307      0.001      -0.012      -0.003
H_CONTRL                        0.0207      0.004      4.665      0.000       0.012       0.029
===============================================================================================
```

7. Adjusted Odds Ratio Comparisons Across Procedures

To better understand the relative likelihood of receiving key interventions, **adjusted odds ratios (ORs)** from the individual procedure logistic regression models were compared across groups. Visualizations were created to summarize the ORs for **PCI**, **CABG**, and **thrombolytic therapy**, stratified by weekend admission status and teaching hospital status. The comparisons revealed that **teaching**

hospital admission consistently increased the odds of receiving both PCI and CABG, reflecting their greater procedural availability and specialized staffing. In contrast, **weekend admission was associated with lower odds of CABG**, but had minimal impact on PCI access, suggesting that while catheterization laboratories are increasingly operational during weekends, surgical services still experience notable weekend limitations. The odds of receiving thrombolytic therapy remained low overall, consistent with modern treatment preferences. These patterns highlight how systemic and institutional factors, not just patient clinical profiles continue to influence access to potentially life-saving cardiovascular interventions.



8. Subgroup Mortality Modeling

To further explore whether the weekend effect varied by hospital type, **separate logistic regression models** were constructed for patients admitted to **teaching hospitals** and **non-teaching hospitals**. In teaching hospitals, weekend admission was associated with a **modestly higher adjusted odds of mortality**, suggesting that despite greater resources, systemic weekend-related challenges such as reduced staffing or procedural delays may persist. In contrast, in non-teaching hospitals, **weekend admission did not show a statistically significant association with mortality** after controlling severity and other covariates. These subgroup analyses suggest that the weekend effect may be more relevant in institutions where the case mix is more complex and where advanced procedures are more frequently performed.

```
Subgroup Mortality Model: Non-Teaching Hospital
                           Logit Regression Results
==============================================================================
Dep. Variable:                   DIED   No. Observations:              2585213
Model:                          Logit   Df Residuals:                  2585206
Method:                           MLE   Df Model:                            6
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.2321
Time:                        10:52:03   Log-Likelihood:            -3.1753e+05
converged:                       True   LL-Null:                   -4.1352e+05
Covariance Type:            nonrobust   LLR p-value:                     0.000
=========================================================================================
                            coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------
Intercept               -10.7990      0.031   -344.673      0.000     -10.860     -10.738
Weekend_Admission         0.0105      0.008      1.314      0.189      -0.005       0.026
AGE                       0.0142      0.000     49.528      0.000       0.014       0.015
FEMALE                   -0.0310      0.007     -4.453      0.000      -0.045      -0.017
LOS                      -0.0436      0.001    -63.223      0.000      -0.045      -0.042
APRDRG_Severity           0.8212      0.008     98.421      0.000       0.805       0.838
APRDRG_Risk_Mortality     1.3493      0.008    162.619      0.000       1.333       1.366
=========================================================================================
Optimization terminated successfully.
         Current function value: 0.126915
         Iterations 9

Subgroup Mortality Model: Teaching Hospital
                           Logit Regression Results
==============================================================================
Dep. Variable:                   DIED   No. Observations:              5511097
Model:                          Logit   Df Residuals:                  5511090
Method:                           MLE   Df Model:                            6
Date:                Wed, 16 Apr 2025   Pseudo R-squ.:                  0.2868
Time:                        10:52:31   Log-Likelihood:            -6.9944e+05
converged:                       True   LL-Null:                   -9.8067e+05
Covariance Type:            nonrobust   LLR p-value:                     0.000
=========================================================================================
                            coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------
Intercept               -11.9615      0.022   -542.934      0.000     -12.005     -11.918
Weekend_Admission         0.0225      0.005      4.275      0.000       0.012       0.033
AGE                       0.0066      0.000     39.926      0.000       0.006       0.007
FEMALE                   -0.0479      0.005    -10.514      0.000      -0.057      -0.039
LOS                      -0.0358      0.000   -109.916      0.000      -0.036      -0.035
APRDRG_Severity           1.0583      0.006    176.174      0.000       1.047       1.070
APRDRG_Risk_Mortality     1.5967      0.006    269.476      0.000       1.585       1.608
=========================================================================================
```

9. Key Findings and Insights

The multivariate modeling results yielded several important insights regarding mortality and procedural intervention patterns among cardiovascular admissions. After adjusting for demographics, clinical severity, and insurance status, weekend admission itself was **not found to be a statistically significant independent predictor of mortality**, suggesting that the initially observed weekend effect was largely confounded by differences in patient risk profiles. Patients admitted to **teaching hospitals exhibited higher severity of illness scores**, which helps explain the slightly elevated mortality observed in these facilities even after adjustment. Furthermore, **procedural intervention rates**, including PCI and CABG, were significantly higher in teaching hospitals compared to non-teaching hospitals, reflecting their greater procedural capacity and specialized cardiac services. Despite this, **weekend admissions were associated with a lower likelihood of receiving**

CABG surgery, highlighting ongoing operational challenges that affect access to surgical interventions outside of standard weekday hours. Across all models, **clinical severity emerged as the dominant predictor of in-hospital mortality**, demonstrating that intrinsic patient factors such as underlying illness severity have a greater influence on outcomes than hospital-level or system-level characteristics alone. These findings emphasize the importance of considering both clinical acuity and hospital operational structures when assessing and improving cardiovascular care delivery.

References

1. Khoshchehreh, M., Grothe, H. L., Wilansky, S., Reaven, N. L., & Goldberg, S. (2016). Changes in mortality on weekend versus weekday admissions for acute coronary syndrome. *International Journal of Cardiology*, 203, 44–48.

2. Python Libraries Used:

    - McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. Proceedings of the 9th Python in Science Conference, 51–56. [pandas]
    - Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, 2825–2830. [statsmodels, scikit-learn]
    - Hunter, J. D. (2007). *Matplotlib: A 2D Graphics Environment*. Computing in Science & Engineering, 9(3), 90–95. [matplotlib]
    - Waskom, M., et al. (2021). *Seaborn: Statistical Data Visualization*. Journal of Open Source Software, 6(60), 3021. [seaborn]