Week 10: Deliverables

Group Name: Medical Data Science

Name: Peter Abban

Email: abbanpeter12@gmail.com

Country: Hungary

Specialization: Data Science

- **Problem Description**

A critical challenge for pharmaceutical companies is **understanding and monitoring drug persistency**, which reflects the extent to which patients adhere to prescribed medication schedules over time. Poor adherence compromises treatment efficacy, increases healthcare costs, and negatively impacts patient outcomes.

To address this challenge, **ABC Pharmaceutical Company engaged an analytics firm** to develop an automated system based on some feature variables of the patients

By leveraging advanced analytics, the solution aims to enhance **patient care, support physicians in monitoring compliance, and inform strategic decisions for the company's pharmaceutical portfolio** by making predictions on whether the patient is persistent or non-persistent

- **Exploratory Data Analysis on the Dataset**

1. Introduction

This analysis investigates patterns of drug persistency and non-persistency among patients based on demographic, clinical, and behavioral variables. Bar plots were generated to illustrate the distribution of persistent and non-persistent patients across several categorical features, including Gender, Race, Ethnicity, Region, Age_Bucket, Ntm_Speciality_Bucket, Risk_Segment_Prior_Ntm, Tscore_Bucket_Prior_Ntm, Risk_Segment_During_Rx, Tscore_Bucket_During_Rx, Adherent_Flag, and Injectable_Experience_During_Rx. The following are the figures of the selected categories plotted with the persistency flag.

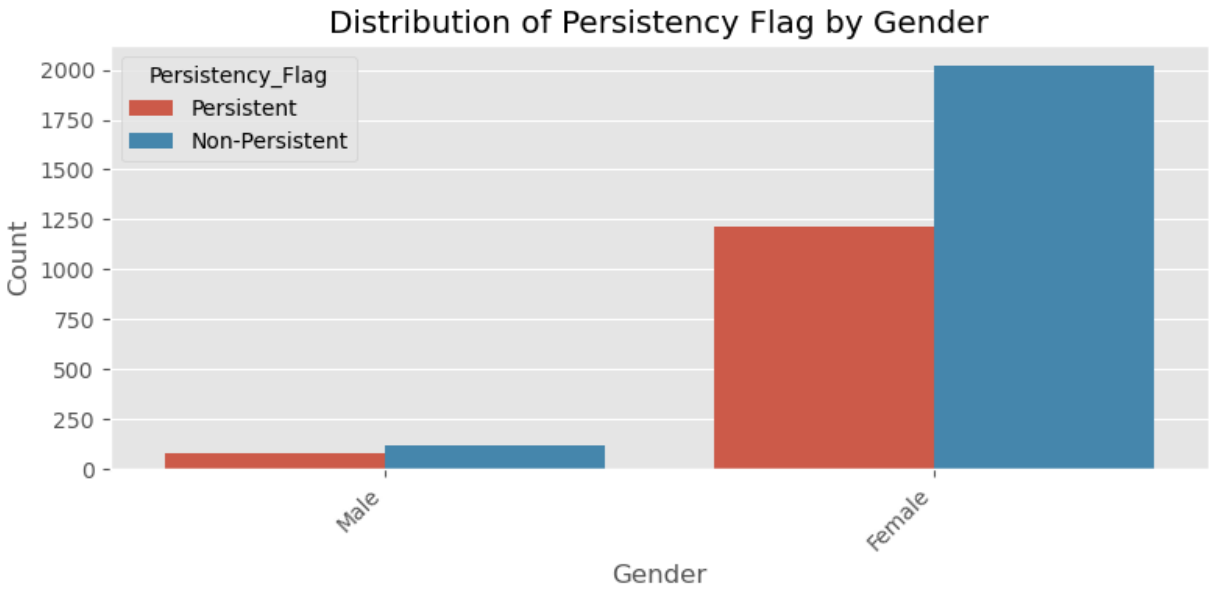

Figure 1. Plot of Gender with Persistency flag

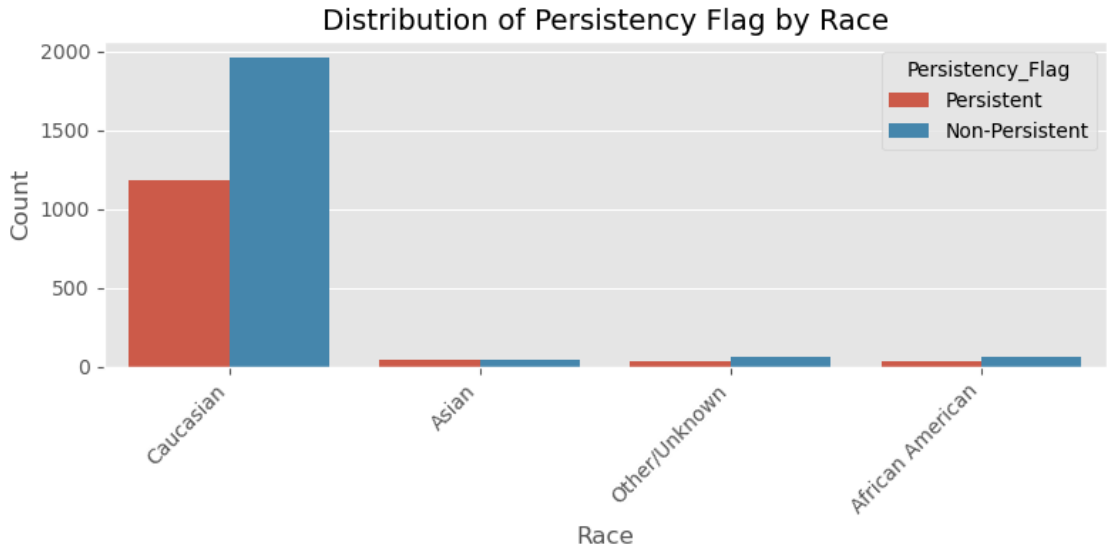

Figure 2. Plot of Race with Persistency flag

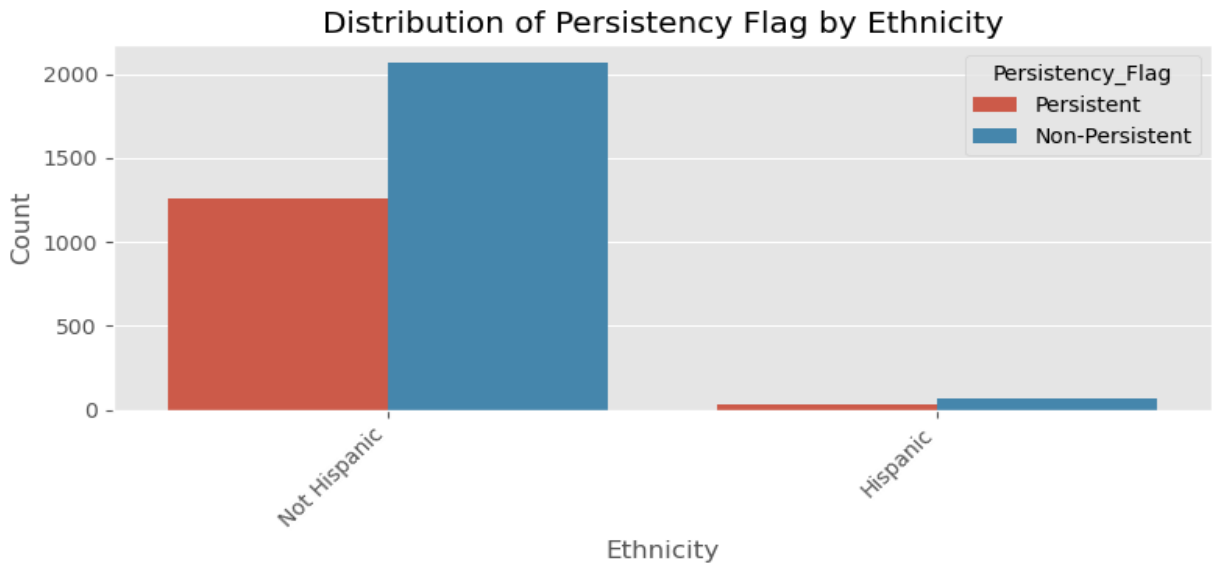

Figure 3. Plot of Ethnicity with Persistency flag

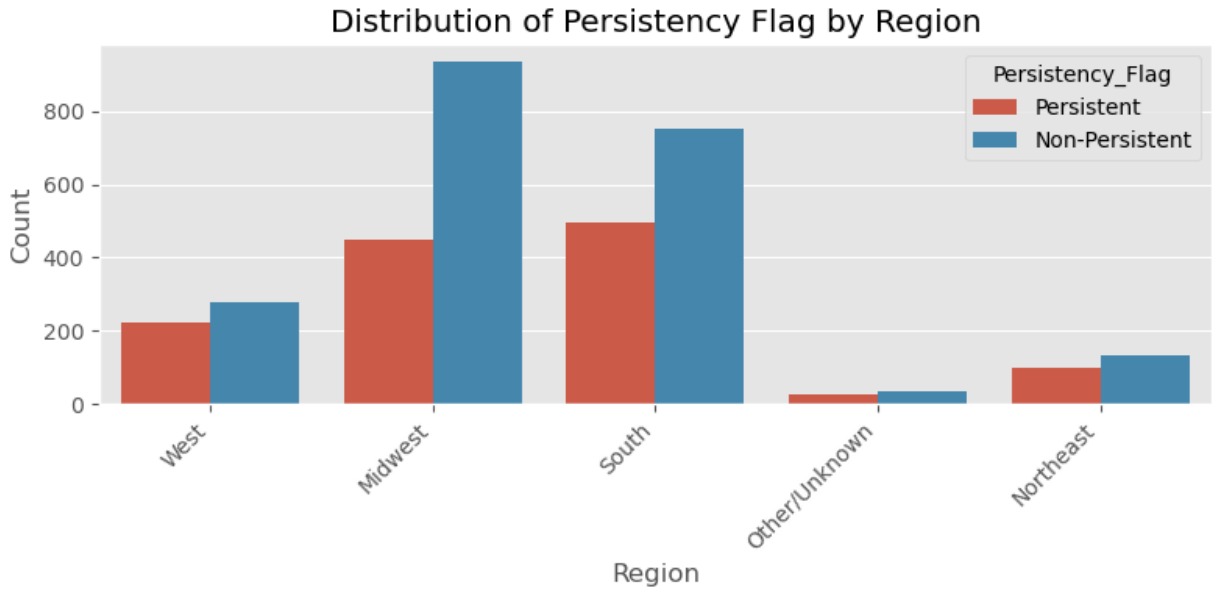

Figure 4. Plot of Region with Persistency flag

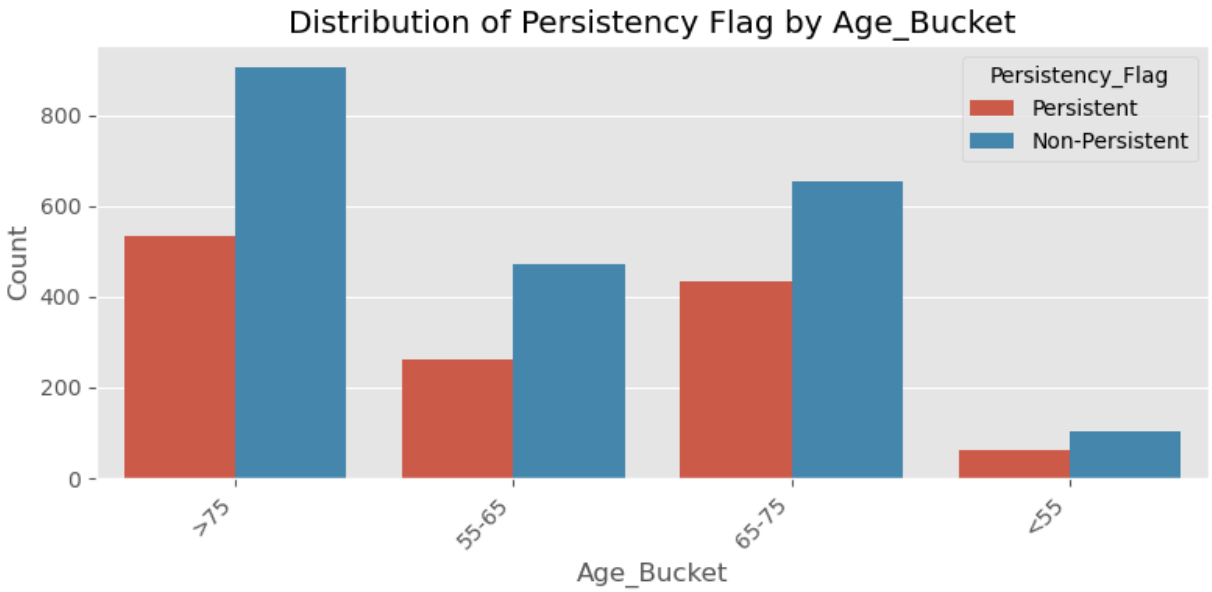

Figure 5. Plot of Age_Bucket with Persistency flag

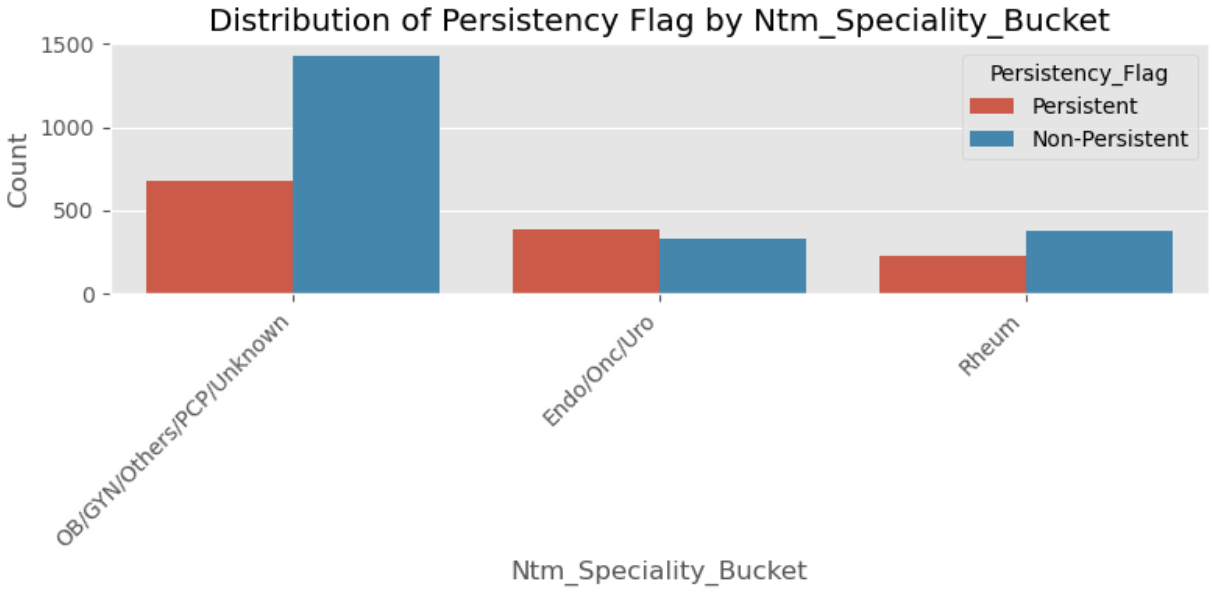

Figure 6. Plot of Ntm_Speciality with Persistency flag

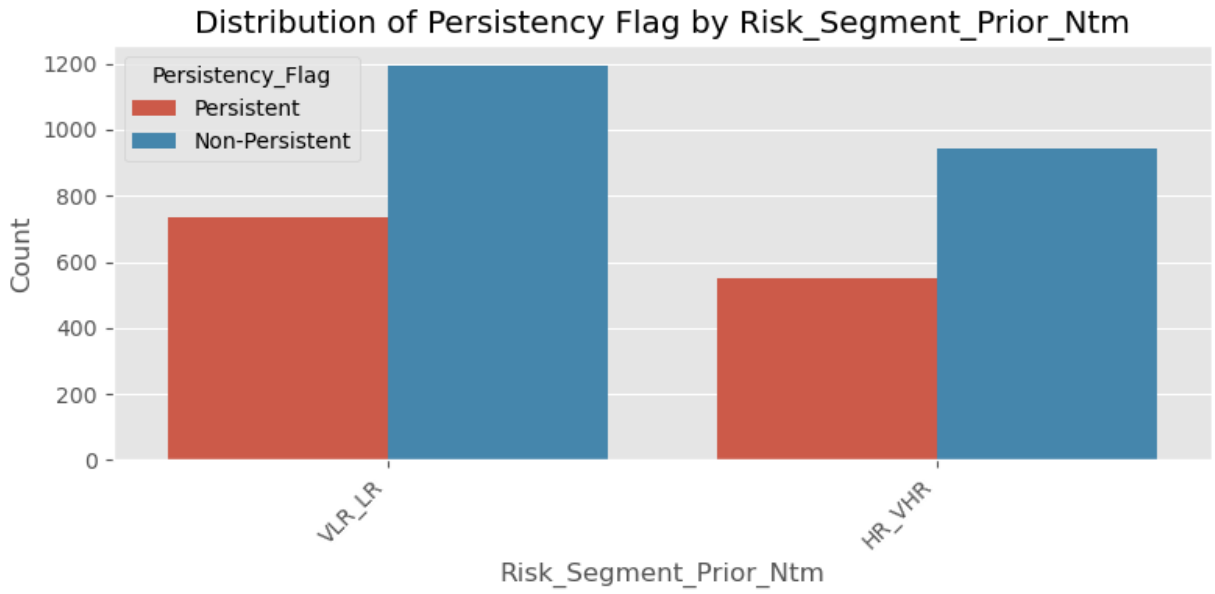

Figure 7. Plot of Risk_Segment_Prior_Ntm with Persistency flag

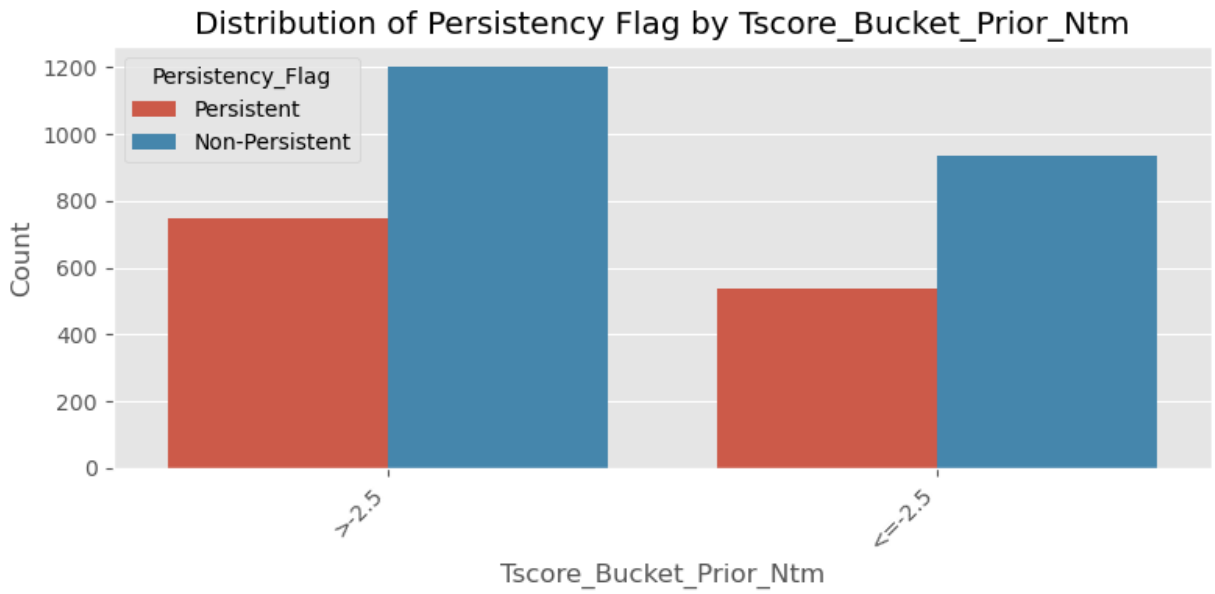

Figure 8. Plot of Tscore_Bucket_Prior_Ntm with Persistency flag

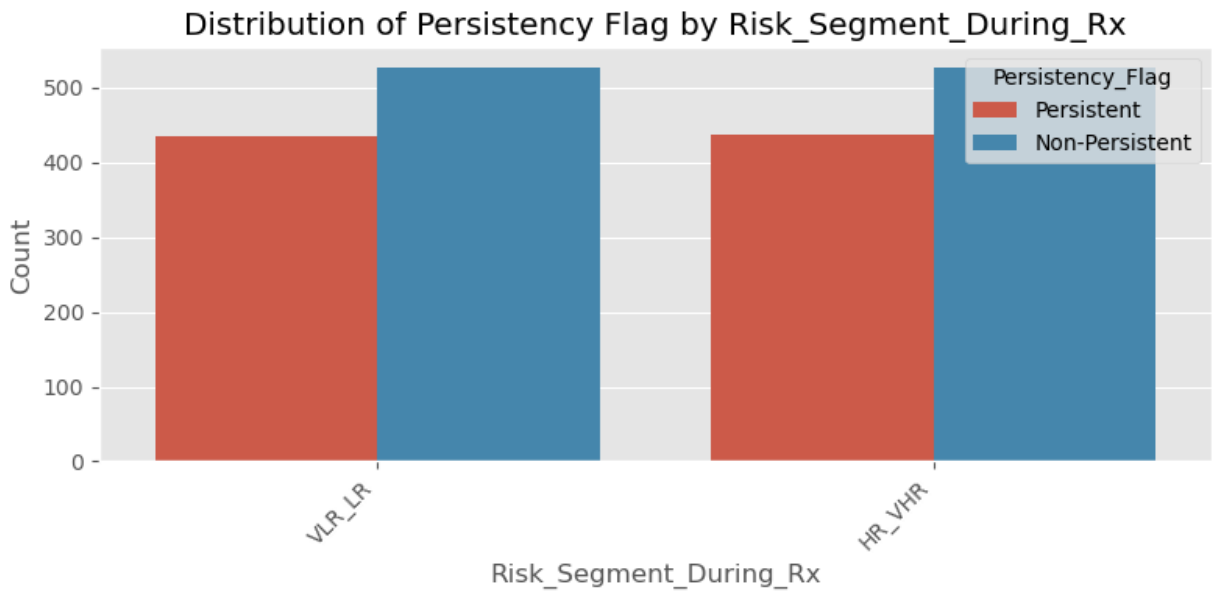

Figure 9. Plot of Risk_Segment_During_Rx with Persistency flag

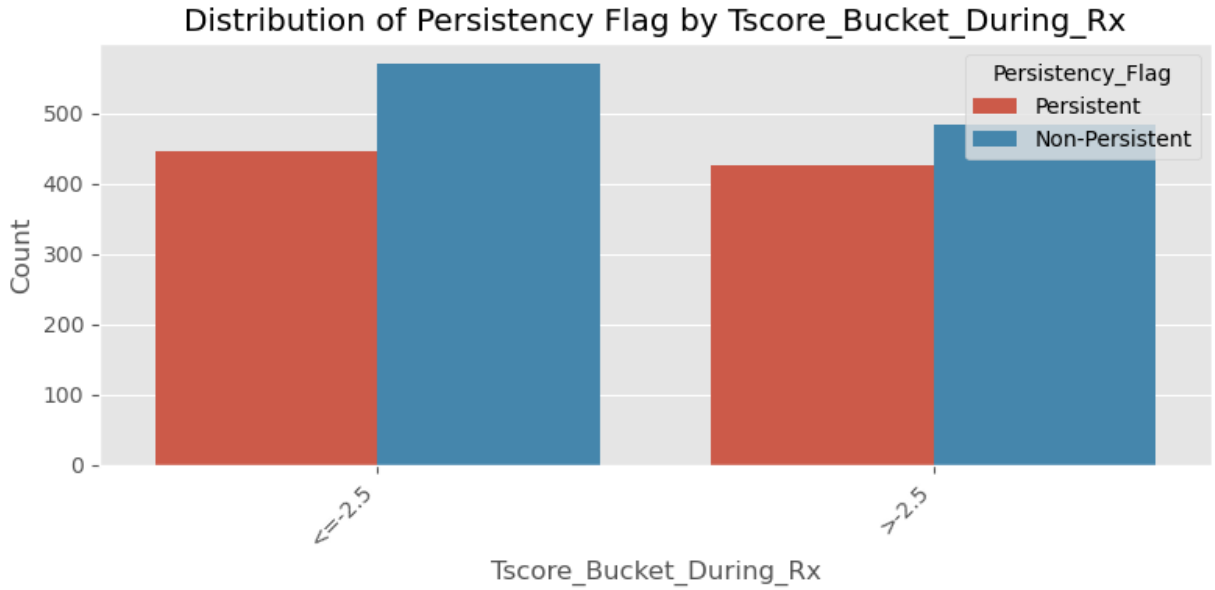

Figure 10. Plot of Tscore_Bucket_During_Rx with Persistency flag

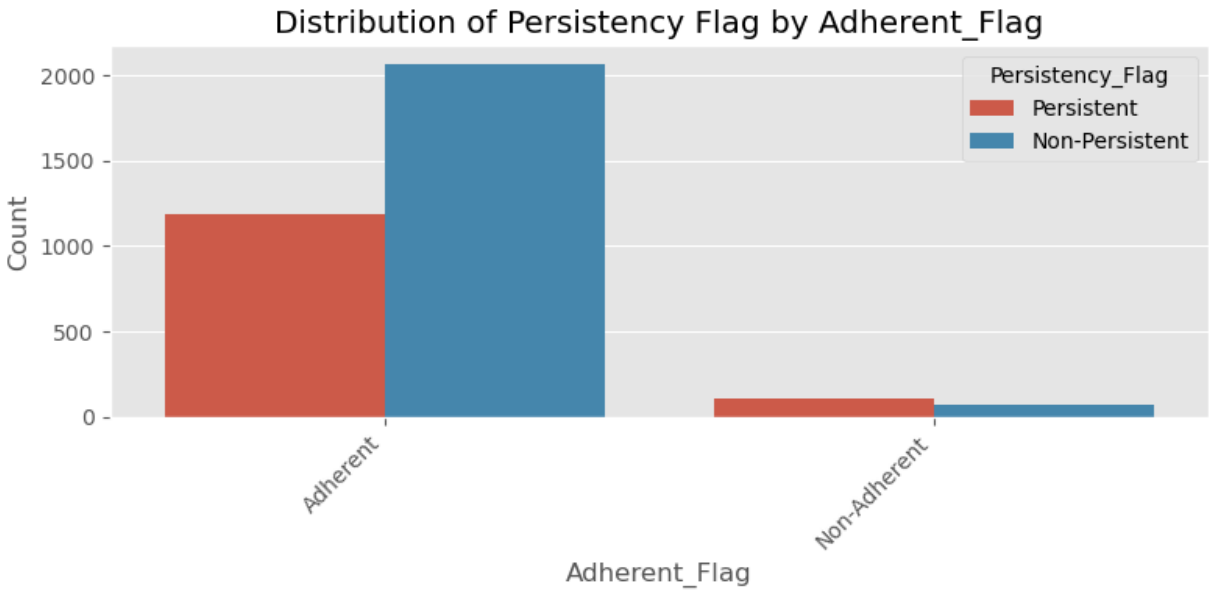

Figure 11. Plot of Adherent flag with Persistency flag

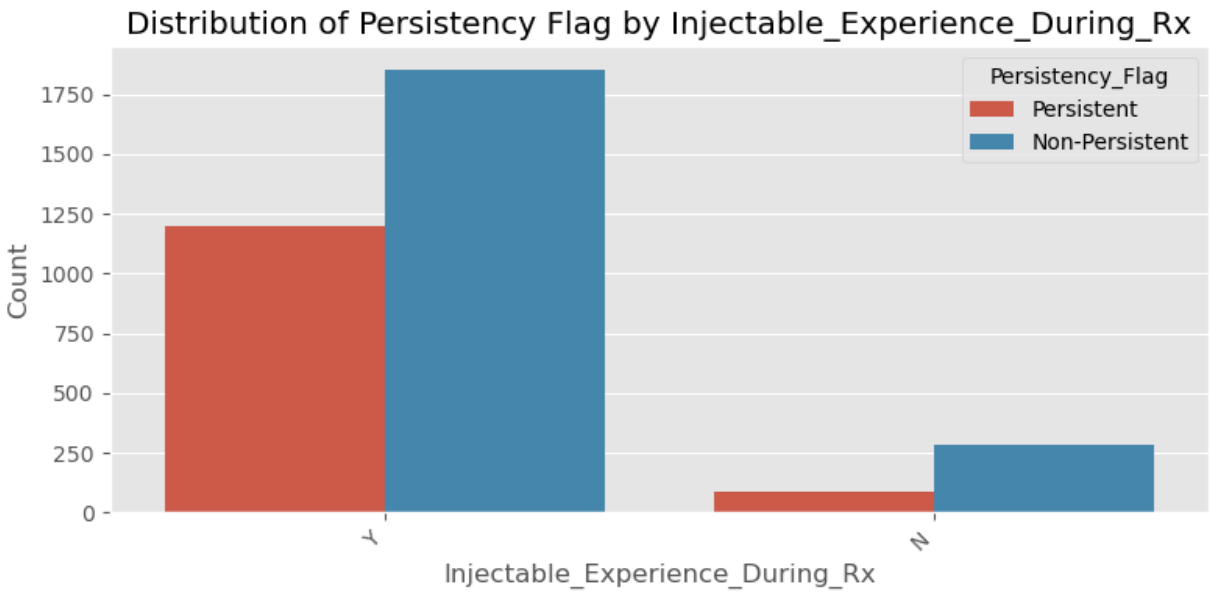

Figure 12. Plot of Injectable_Experience_During_Rx with Persistency flag

2. Key Observations

a. Demographic Patterns

• **Gender:** Female patients represent the majority across all categories, with a higher proportion of non-persistent cases.
• **Race:** Caucasian patients show the highest representation, again with elevated non-persistency rates.
• **Ethnicity:** Non-Hispanic individuals display the largest patient count and the highest non-persistency levels.
• **Region:** The Midwest records the highest number of non-persistent cases.
• **Age_Bucket:** Patients aged 75 years and above exhibit the greatest non-persistency, suggesting age-related adherence challenges.

b. Clinical and Risk Factors

• **Ntm_Speciality_Bucket:** OB/GYN and other specialties show the highest counts for both persistence and non-persistence.
• **Risk_Segment_Prior_Ntm:** The VLR_LR (Very Low Risk–Low Risk) group has the largest count for both persistency flag outcomes.
• **Tscore_Bucket_During_Rx:** Patients with T-scores ≤ -2.5 dominate both persistent and non-persistent groups.
• **Adherent_Flag:** The Adherent group shows the highest counts overall, yet includes many non-persistent cases.
• **Injectable_Experience_During_Rx:** Patients with prior injectable experience (Y) have the highest counts, with non-persistency also prominent.

The remaining figures show the plots of categorical features and their behavior.

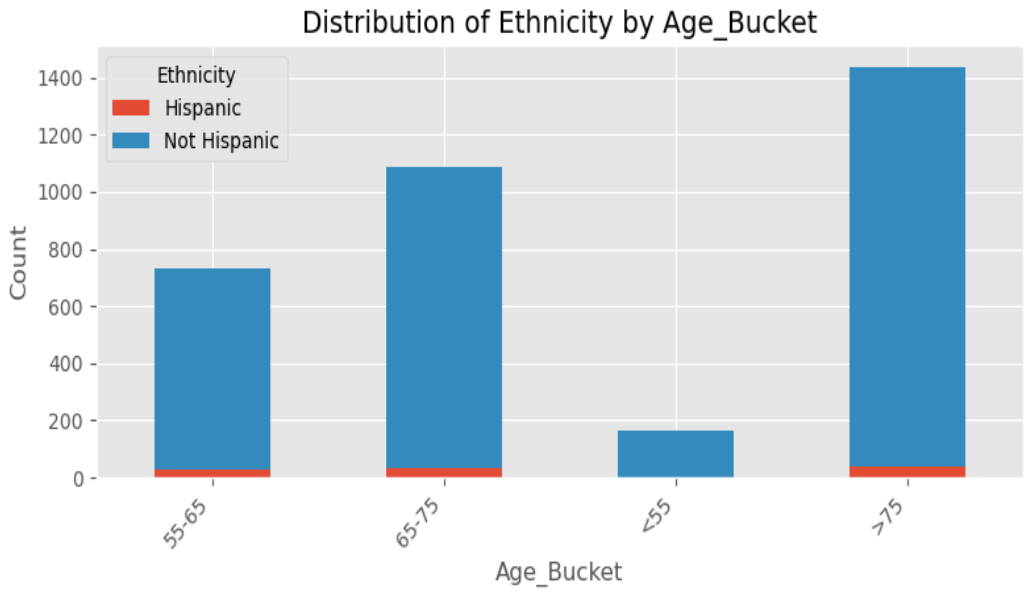

Figure 13. Plot of Age_bucket by Ethnicity

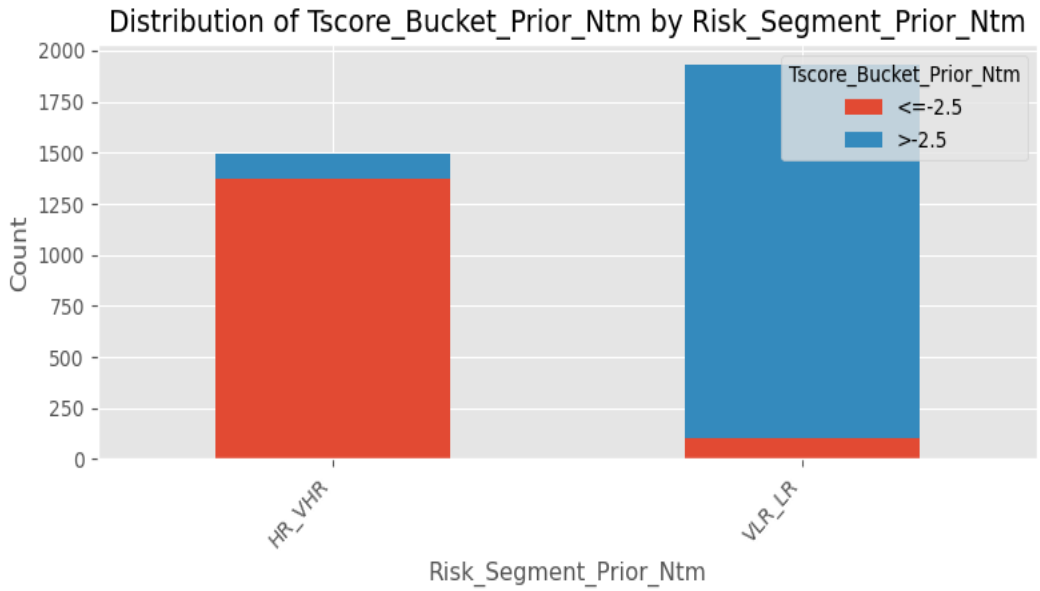

Figure 14. Plot of Tscore_Bucket_Prior_Ntm by Risk_Segment_Prior_Ntm

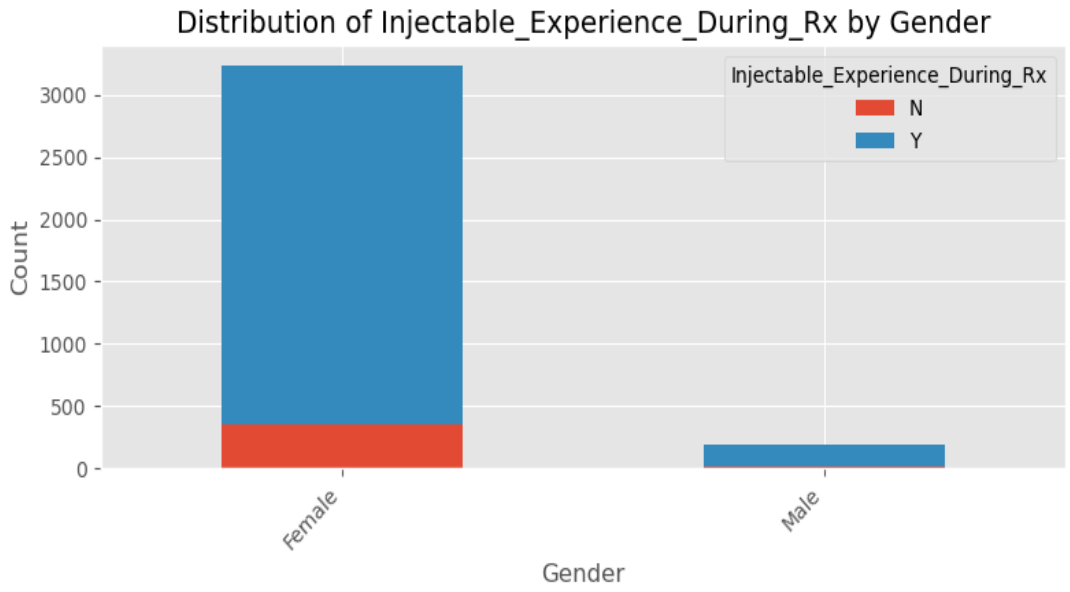

Figure 15. Plot of Injectable_Experience_During_Rx by Gender

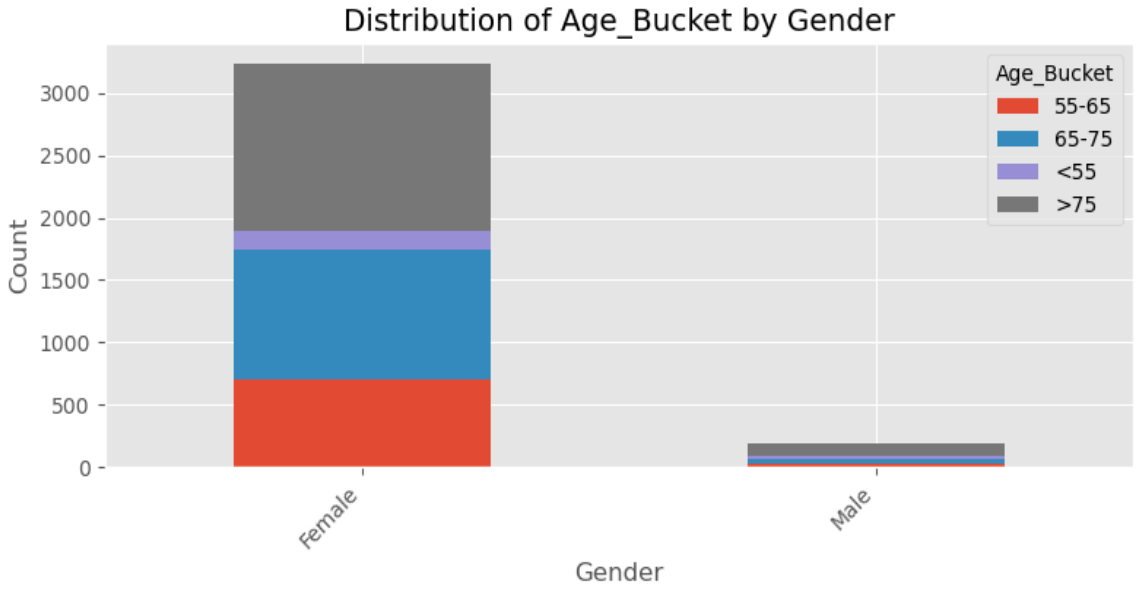

Figure 16. Plot of Age_Bucket by Gender

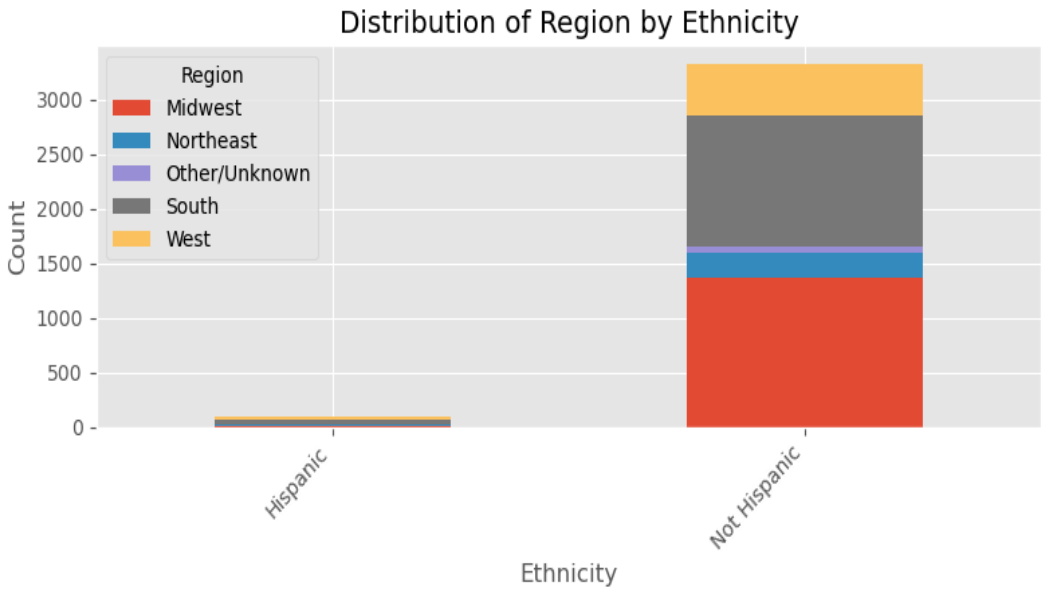

Figure 17. Plot of Region by Ethnicity

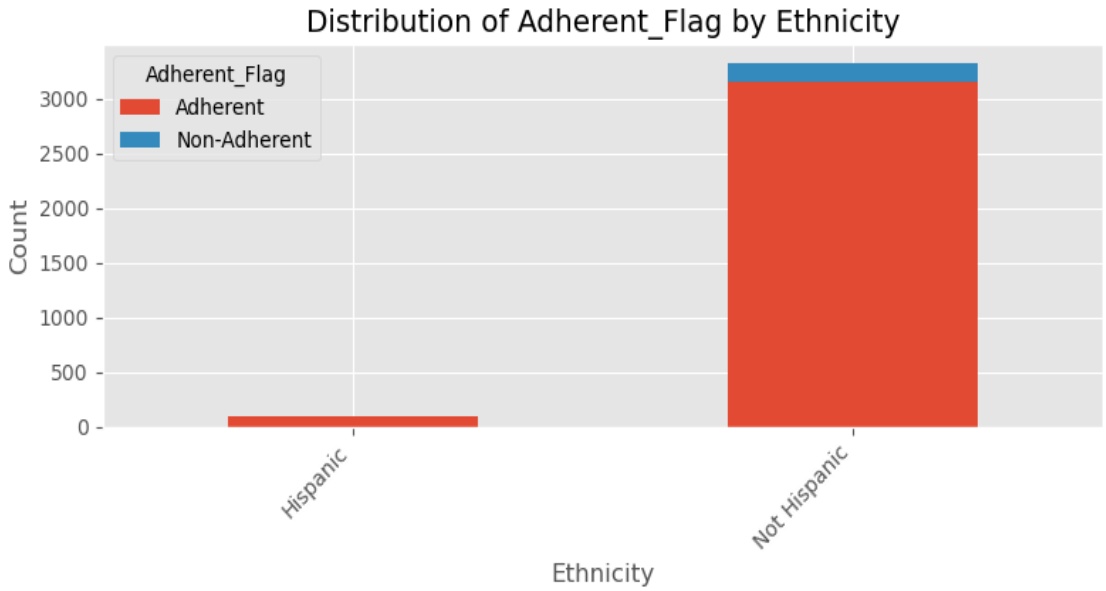

Figure 18. Plot of Adherent flag by Ethnicity

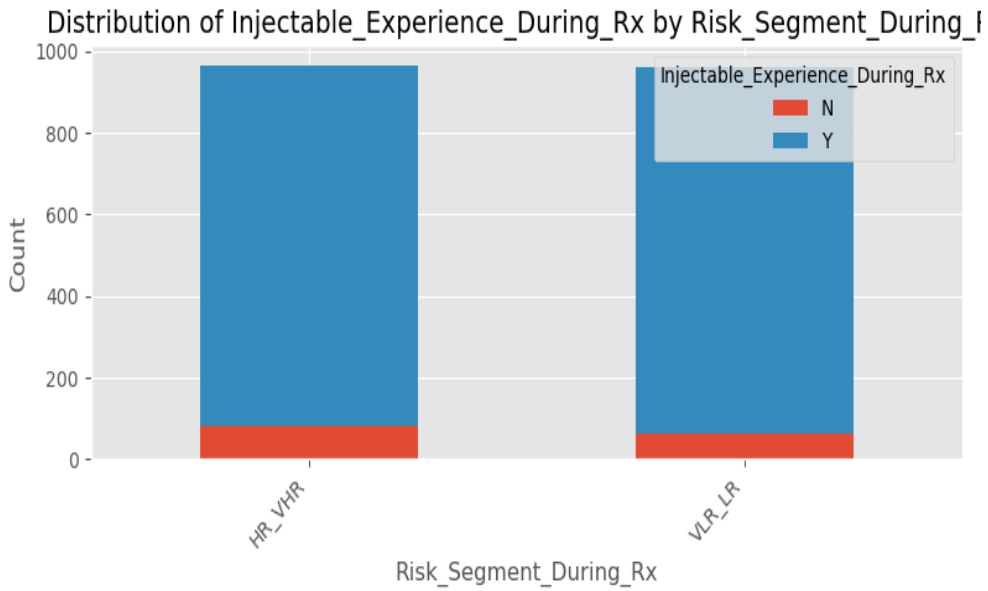

Figure 19. Plot of Injectable_Experience_During_Rx by Risk_Segment_During_Rx

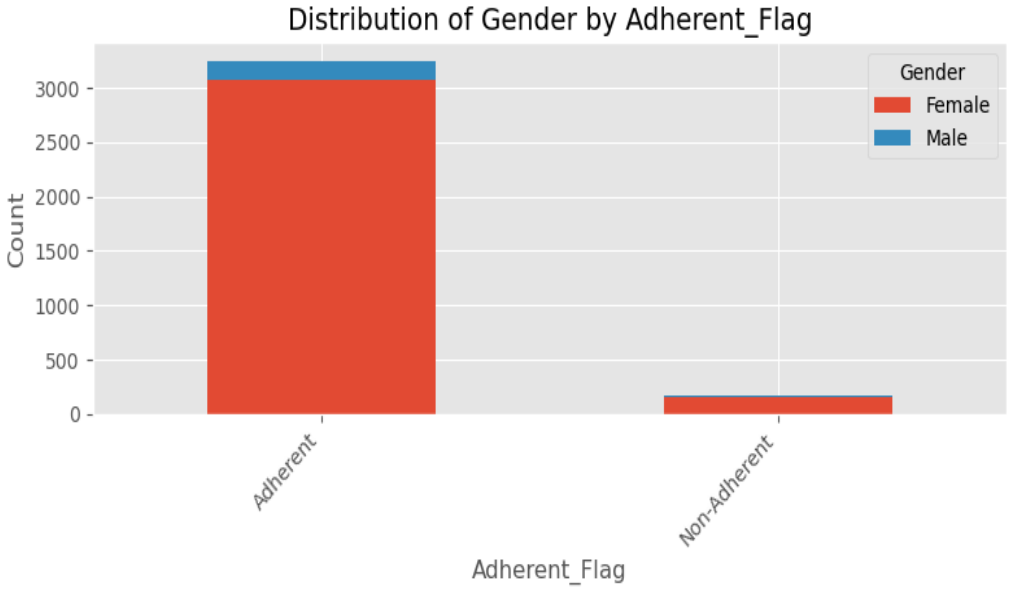

Figure 20. Plot of Gender by Adherent flag

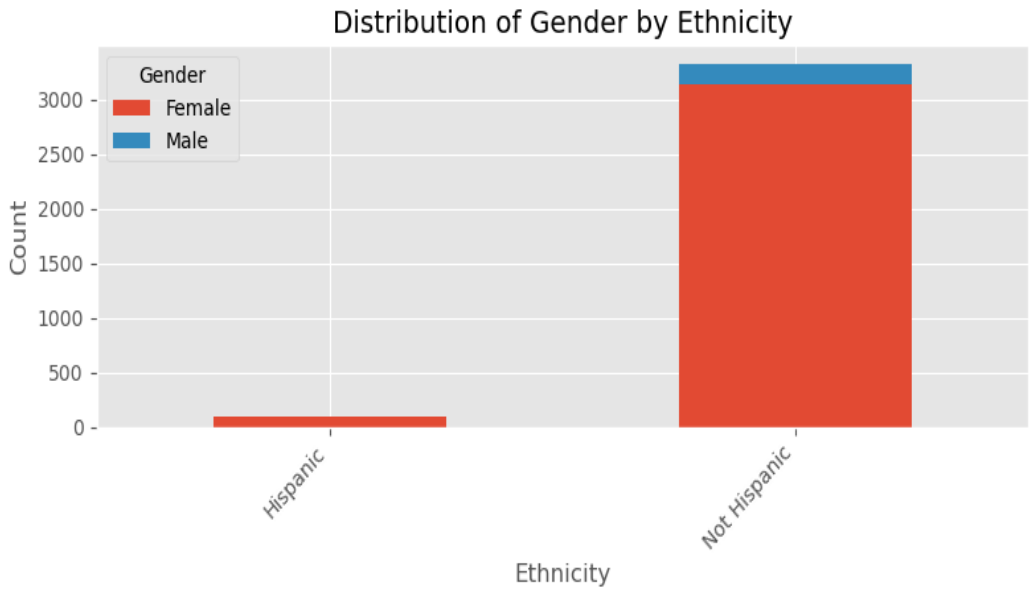

Figure 21. Plot of Gender by Ethnicity

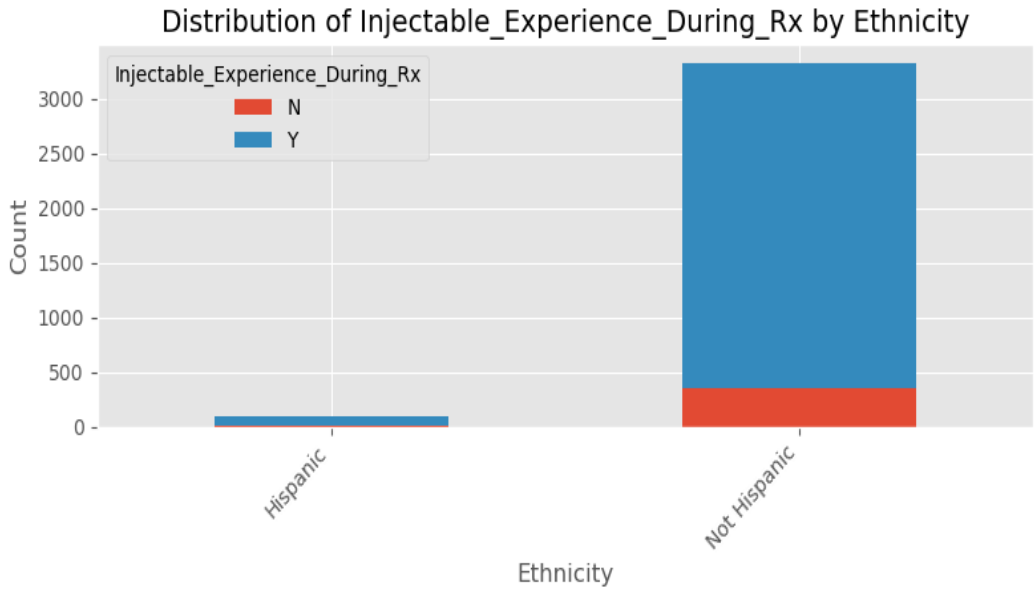

Figure 21. Plot of Injectable_Experience_During_Rx by Ethnicity

c. Categorical Comparisons

• **Ethnicity × Age_Bucket:** Not-Hispanic individuals dominate the >75 years age group.

• **Tscore_Bucket_Prior_Ntm × Risk_Segment_Prior_Ntm:** The VLR_LR group generally presents T-scores > -2.5.

• **Injectable_Experience_During_Rx × Gender:** Female patients show the highest

proportion with prior injectable experience.
• **Age_Bucket × Gender:** Female patients dominate across all age categories.
• **Adherent_Flag × Ethnicity:** Not-Hispanic individuals are the most adherent.
• **Gender × Ethnicity:** The majority of female patients are Not-Hispanic.

3. Deductions and Insights

1. Gender bias is evident, with female patients dominating across all variables, possibly reflecting differences in health-seeking behavior or disease prevalence.
2. Non-persistency increases with age (>75 years), likely due to comorbidities, cognitive decline, or complex regimens, with the potential majority being females.
3. Adherence does not guarantee persistence, indicating a gap between short-term compliance and sustained treatment.
4. Regional and ethnic disparities (Midwest and Non-Hispanic dominance) may reflect socioeconomic or healthcare accessibility factors.
5. Patients with low bone density (T-score ≤ -2.5) are more engaged but still experience non-persistency, potentially due to side effects or treatment fatigue.

4. Summary

The analysis reveals consistent demographic and behavioral patterns influencing drug persistency flag. Females, Non-Hispanic patients, older adults, and those in the Midwest constitute the majority of the patient population. While adherence rates are initially high, persistence over time remains a challenge, especially among older patients It also suggests that females are the gender category who are highly non-persistent.

5. Recommendations

1. Implement targeted engagement programs for older adults with the focus on female non-Hispanics.
2. Develop culturally and regionally tailored communication strategies, focusing on the Midwest and Non-Hispanic populations.
3. Introduce behavioral adherence tools, such as mobile reminders or pharmacist counseling for female non-Hispanics.
4. Provide persistence-focused training for OB/GYN and primary care providers.