

# Business EDA

Saif Amer

# Team Member

Saif Amer

[saifamer67@gmail.com](mailto:saifamer67@gmail.com)

United States

The Ohio State University

## Project Objective

- Understand the key factors impacting patient persistency.
- Support strategic decision-making for drug adherence programs.

## Data Summary

- Total Records: Approx. 10,000
- Features: 69
- Target Variable: Persistency\_Flag
- No missing values detected.

# Data Cleaning

- Verified no null values
- Encoded categorical variables (Gender, Ethnicity, etc.)
- Created 'Count\_Of\_Risks' variable from multiple binary risk flags

## Key Demographic Insights

- Most patients are Female, aged 65-75
- Dominant regions: Southeast & Northeast
- Non-persistent patients more likely to have joint disorders

## Behavioral Patterns

- Prior fracture history and low T-score patients show higher persistency
- Vitamin D deficiency common in non-persistent patients

## Risk Factor Distribution

- Visualization shows more risks often means higher persistency
- Possibly due to increased patient awareness or clinical attention



## Business Recommendations

- Target demographics with low persistency for outreach
- Raise awareness about fracture risk
- Collaborate with high-performing specialists

## Next Steps

- Hyperparameter tuning (GridSearchCV)
- Evaluation (Precision, Recall, F1, ROC-AUC)
- Feature selection to reduce redundancy