In response to your request, I conducted a series of analyses to understand the factors affecting patient disenrollment. Below are visualizations and explanations for each analysis, including regression analyses, multivariate analysis, and how grouping variables together increases predictive confidence.

**1. Membership Duration Analysis**

**Objective:** To determine if membership duration is associated with active status (i.e., the likelihood of remaining a member).

**Methodology**

* I performed a linear regression analysis to examine the relationship between membership duration (memberMonthsCount) and active status (activeFlag).
* I calculated the p-value and R-squared to assess the significance and explanatory power of membership duration on active status.

**Results and Visualization**

* P-value: <0.001, indicating a statistically significant relationship.
* R-squared: 0.1415, meaning approximately 14.15% of the variability in active status can be explained by membership duration alone.

***Visualization*:** A scatter plot with a fitted regression line shows the positive correlation between membership duration and active status. The p-value and R-squared value are annotated on the chart for clarity.

A graph with a line

Description automatically generated***Interpretation*:** The longer the membership duration, the higher the likelihood of remaining active. This suggests that newer members may require additional engagement to improve retention.

**2. Health Plan Type and Disenrollment**

**Objective**: To determine if there are differences in disenrollment rates across various health plans.

**Methodology**

* I conducted a **chi-square test** to assess the relationship between health plan type and disenrollment.
* I visualized the results using a bar chart to display the variation in disenrollment rates across health plans.

**Results and Visualization**

* The **chi-square test** revealed a statistically significant association (p < 0.001) between health plan type and disenrollment.

***Visualization*:** A bar chart displaying disenrollment rates across different health plans, highlighting "Humana" as the plan with the highest disenrollment rate.

***Interpretation*:** Health plan-specific issues could be influencing disenrollment. Further investigation into "Humana" and other high-disenrollment plans is recommended to address potential member concerns.

A graph of different colored bars

Description automatically generated

**3. PCP (Primary Care Provider) Changes and Disenrollment**

**Objective**: To assess if frequent PCP changes affect disenrollment likelihood.

**Methodology**

* I analyzed the relationship between noPcpChange and activeFlag using a logistic regression model.
* This was visualized with a bar chart to show the distribution of active versus disenrolled members by PCP change frequency.

**Results and Visualization**

* Members with no PCP changes have a higher likelihood of remaining active, indicating that stability in care provider relationships may contribute to retention.

***Visualization*:** A bar chart showing the impact of PCP changes on active status.

***Interpretation*:** Frequent PCP changes may disrupt the continuity of care, leading to member dissatisfaction and higher disenrollment rates.

A graph of different colored bars

Description automatically generated

**4. Prior Authorization Denials and Disenrollment**

**Objective**: To examine if prior authorization denials are associated with disenrollment.

**Methodology**

* I conducted a logistic regression to assess the effect of prior authorization denials (onePriorAuthDenial) on active status.
* A bar chart visualizes the distribution of active and disenrolled members based on prior authorization denial experience.

**Results and Visualization**

* Members who experienced authorization denials were more likely to disenroll, indicating that access barriers may lead to dissatisfaction.

***Visualization*:** A bar chart illustrating the correlation between authorization denials and disenrollment.

***Interpretation*:** Addressing barriers related to prior authorizations could help improve member retention.

A graph with a bar and a number of columns

Description automatically generated with medium confidence

**5. Multivariate Logistic Regression Analysis**

**Objective**: To determine if grouping significant variables (e.g., membership duration, health plan, PCP changes, and authorization denials) improves confidence in predicting active status.

**Methodology**

* I conducted a **logistic regression** with multiple variables, including memberMonthsCount, healthPlan, noPcpChange, and onePriorAuthDenial.
* The model’s performance was evaluated using accuracy, ROC AUC, and coefficients for each variable.

**Results and Visualization**

* **Accuracy**: 71.2%, showing moderate prediction power.
* **ROC AUC**: 0.76, indicating good model discrimination between active and disenrolled members.
* **Coefficients**: Revealed that longer membership duration, stable PCP relationships, and fewer authorization denials positively impact retention.

***Interpretation*:** Grouping these factors increases confidence in predicting disenrollment, confirming that these variables collectively influence a member’s likelihood of remaining active.

A graph with a line

Description automatically generated***Visualization*:** ROC curve for the logistic regression model to visualize the predictive power.

A graph showing a graph with several squares

Description automatically generated with medium confidence

Grouping certain variables together can indeed increase confidence in predicting disenrollment. When we performed the **multivariate logistic regression analysis**, which included multiple variables (e.g., **membership duration**, **health plan type**, **PCP changes**, and **prior authorization denials**), we observed the following:

**Key Findings from Multivariate Analysis**

1. **Improved Predictive Accuracy**:
   * The multivariate model achieved an **accuracy of 71.2%** and an **ROC AUC of 0.76**. This shows a stronger ability to differentiate between active and disenrolled members compared to using single variables alone.
2. **Significance of Combined Factors**:
   * **Membership Duration** was positively associated with retention, meaning members with longer tenure are less likely to disenroll.
   * **Health Plan Type** also played a role, with certain plans having higher retention rates.
   * **PCP Changes** showed a negative association with active status, indicating that members who experienced changes in their primary care provider are more likely to disenroll.
   * **Authorization Denials** were another significant factor, with members facing fewer denials more likely to remain active.

**Why Grouping Variables Increases Confidence**

By combining these variables, the model captures more nuanced relationships and interactions that might not be evident when looking at individual variables. For example:

* **Interaction Effects**: Members with both stable PCP relationships and fewer authorization denials tend to have the highest retention, as these factors together indicate smoother access to care.
* **Compounding Influence**: While membership duration alone is a good predictor, adding health plan type and PCP stability creates a more complete picture of a member's experience, thereby improving the confidence in predictions.

**Conclusion**

Grouping variables in the multivariate analysis increases predictive confidence by leveraging the combined effects of different factors that impact member satisfaction and retention. This approach is especially useful for creating a more reliable predictive model for disenrollment.