Telecom Churn Analysis

1. Dataset Overview

- The dataset includes customer details for a telecom company, with features such as:
  - Customer demographic information (e.g., age, gender, tenure).
  - Service details (e.g., type of internet service, phone service).
  - Payment and contract details
  - Churn information as the target variable (indicates if a customer has left the service).

2. Exploratory Data Analysis (EDA)

- **Churn Rate**: Approximately X% of customers in the dataset has churned.
- **Key Variables Affecting Churn**:
  - Higher churn rates are observed among customers with short tenure.
  - Customers on month-to-month contracts are more likely to churn.
  - Higher monthly charges are associated with increased churn.
  - Fiber optic internet service users show higher churn rates compared to DSL users.
- **Correlation Analysis**:
  - Monthly charges and tenure are negatively correlated with churn likelihood.
  - Strong correlations observed between service features and churn.

3. Data Preprocessing

- **Handling Missing Values**: Techniques used for imputing missing data.
- **Encoding**: Categorical variables are one-hot encoded for model compatibility.
- **Scaling**: Numerical features are normalized or standardized for better performance in model training.

4. Model Building and Evaluation

- **Models evaluated include:**
  - Logistic Regression
  - Decision Trees
  - Random Forest
  - Gradient Boosting (e.g., XGBoost)
- **Performance Metrics**:
  - Accuracy, Precision, Recall, and F1-Score were calculated for each model.
  - Best model identified: [Model Name] with an F1-Score of approximately [Score].
- **Feature Importance**:
  - The most significant predictors of churn include:
    - Monthly Charges
    - Tenure
    - Contract Type
    - Internet Service Type

5. Recommendations

- Offer incentives to customers with month-to-month contracts to switch to longer-term contracts.
- Develop loyalty programs targeting customers with short tenure.
- Monitor and address issues related to high charges for fiber optic users.

Visualization of specific results.

1. Churn Rate

- **Visualization**: Bar chart or pie chart showing the proportion of churned vs. retained customers.
- **Purpose**: Highlights the overall churn rate in the dataset.

2. Churn by Contract Type

- **Visualization**: Grouped bar chart comparing churn rates across different contract types (e.g., month-to-month, one-year, two-year contracts).
- **Purpose**: Shows the impact of contract type on churn likelihood.

3. Monthly Charges and Churn

- **Visualization**: Box plot comparing monthly charges for churned vs. non-churned customers.
- **Purpose**: Illustrates how higher charges correlate with churn.

4. Tenure and Churn

- **Visualization**: Histogram or KDE plot showing the distribution of tenure for churned vs. non-churned customers.
- **Purpose**: Demonstrates that short tenure is associated with higher churn rates.

5. Feature Importance

- **Visualization**: Horizontal bar chart displaying the feature importance scores from the best-performing model (e.g., Random Forest or Gradient Boosting).
- **Purpose**: Highlights the most significant predictors of churn.

6. Model Performance

- **Visualization**:
    - Confusion matrix for the best model.
    - Bar chart showing precision, recall, F1-Score, and accuracy for different models.
- **Purpose**: Compares model performance and explains the final model's evaluation metrics.