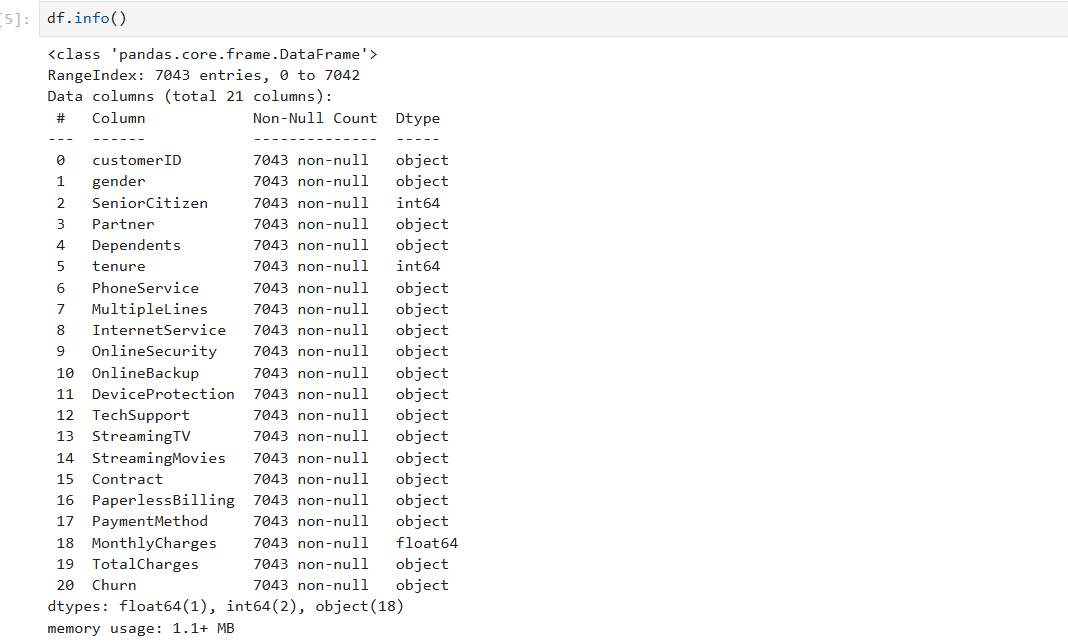


We have imported the essential Python libraries for data analysis:

* **pandas** for data manipulation and analysis.
* **numpy** for numerical operations.
* **matplotlib.pyplot** and **seaborn** for creating charts and visualizations.
* Defines the path to the **customer churn dataset**.
* Loads the CSV into a DataFrame called df using pandas.read\_csv().
* Uses UTF-8 encoding to ensure all text is correctly read.
* Each row is a **customer**.
* Each column is a **feature** about that customer.

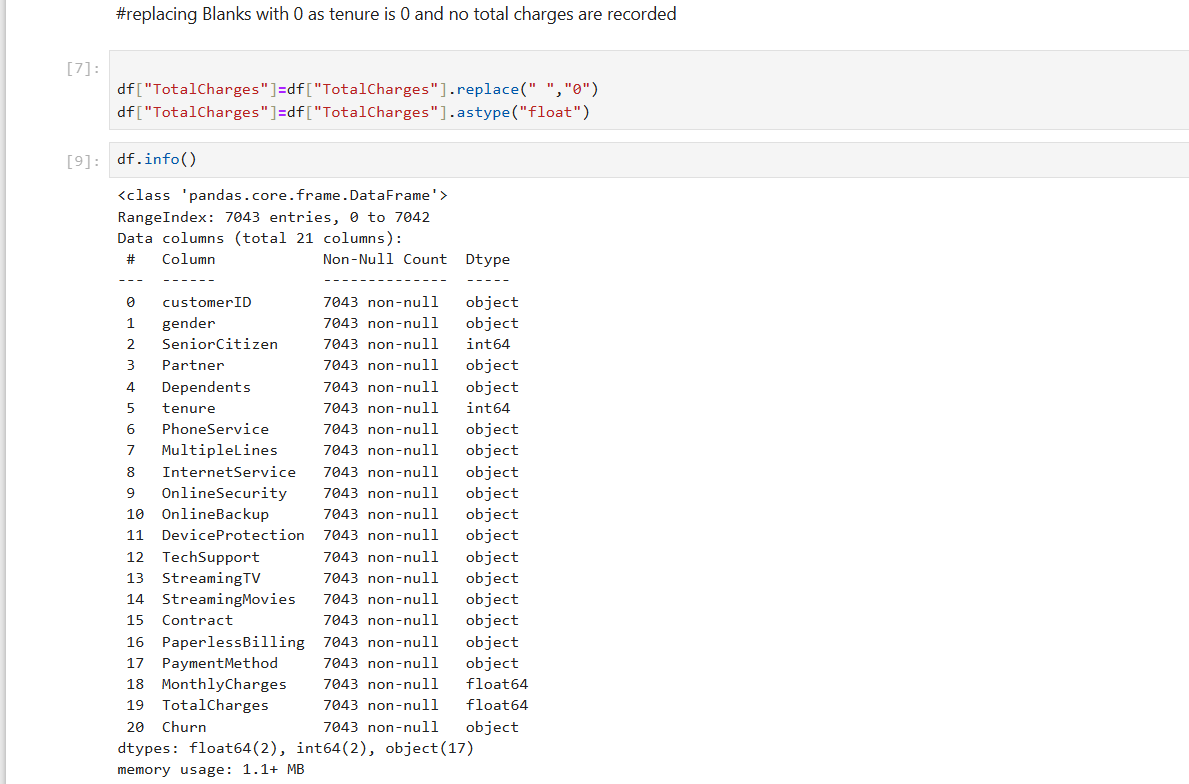


The dataset contains **7,043 records** and **21 columns**, with **no missing values** across any column.

Most columns are of type **object**, indicating categorical/string data, which may need encoding for modeling.

The columns MonthlyCharges and TotalCharges are numeric, but **TotalCharges is incorrectly typed as object** — likely due to spaces or formatting.

Overall, the dataset is clean and ready for further preprocessing and exploratory



 The TotalCharges column initially had blank spaces for some rows where tenure was 0. These were replaced with "0" using .replace(" ", "0").

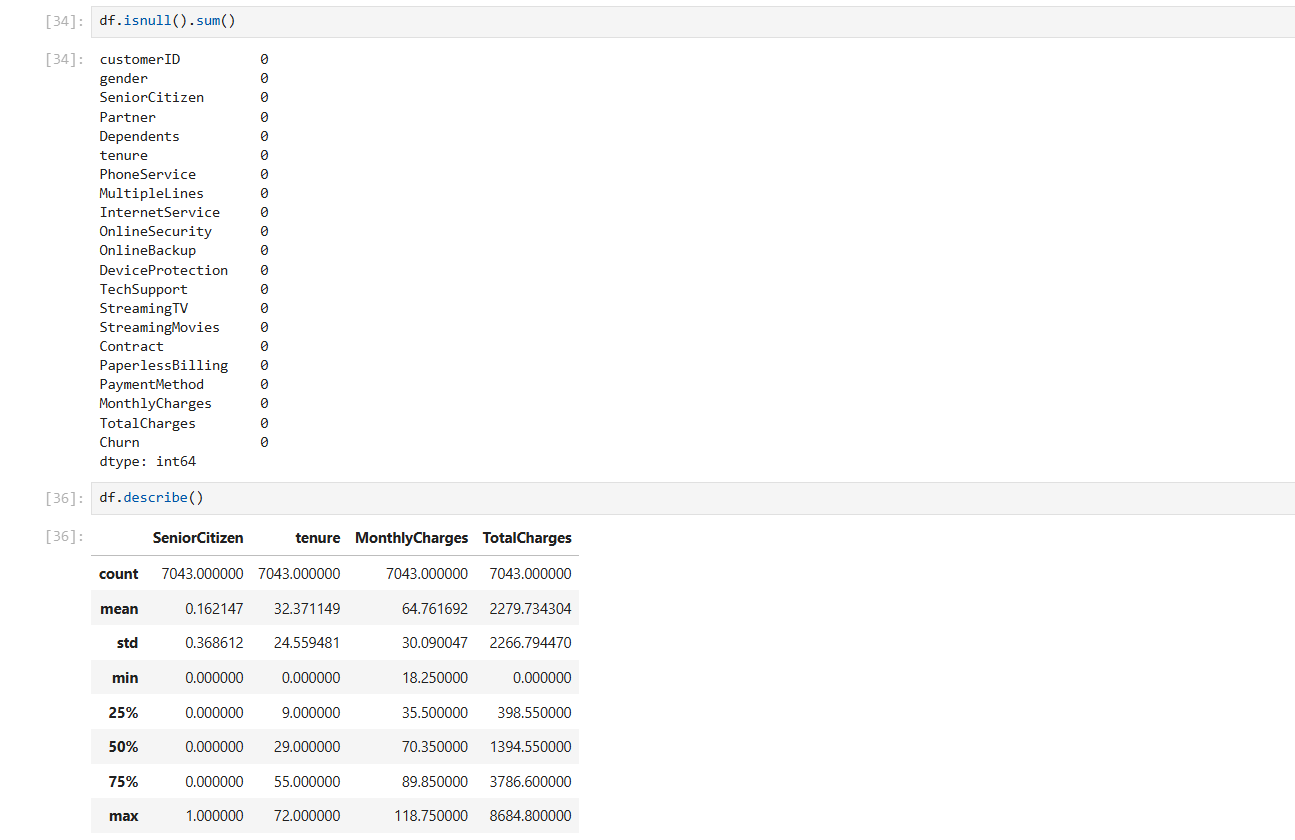
 The column was then converted from object to float64 using .astype("float") to ensure it's numerically usable for analysis.

 Running df.info() confirms that the data types are now correct, with both MonthlyCharges and TotalCharges as float64.

 All 7043 entries are non-null, indicating that missing values have been successfully handled.

 This step was essential for ensuring numerical integrity in correlation analysis and visualizations involving charges.

 Proper data typing like this improves modeling accuracy and prevents runtime errors in statistical operations.



 The output of df.isnull().sum() confirms that **there are no missing values** in any of the 21 columns—ensuring a clean dataset for analysis.

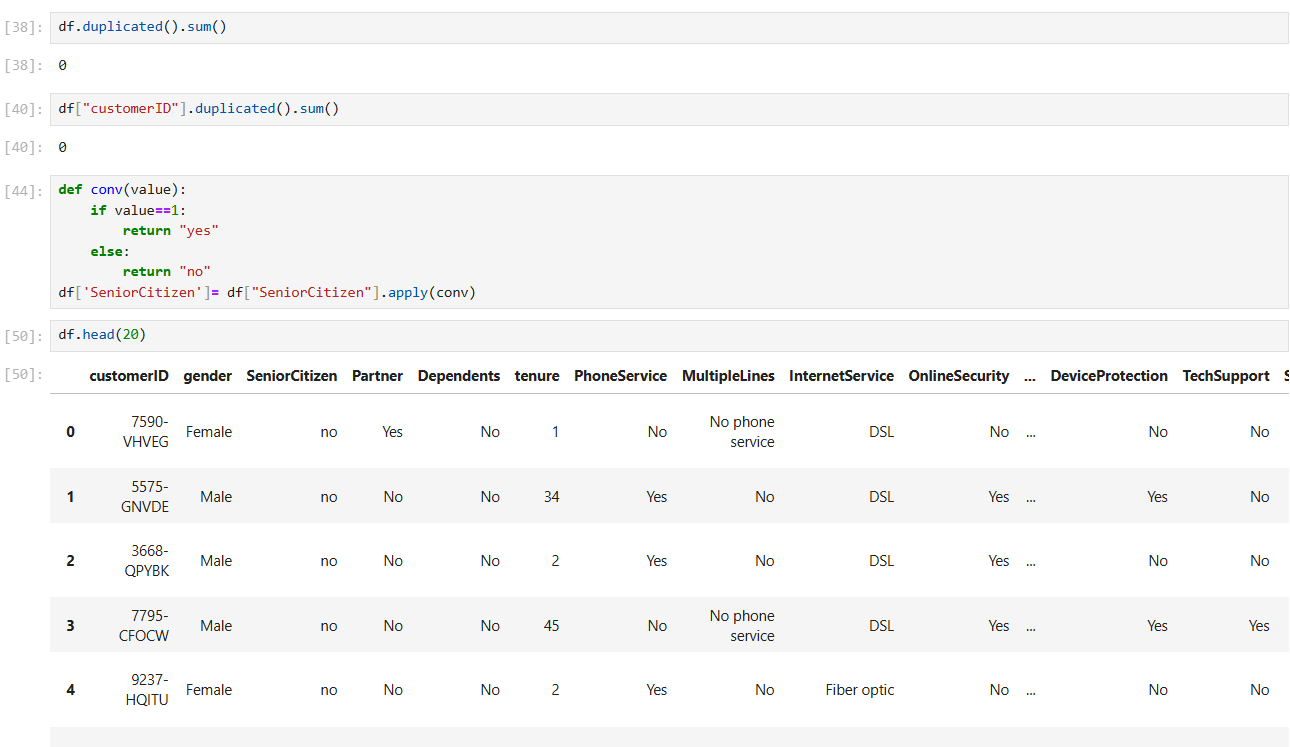
 df.describe() gives statistical summaries of numerical columns like SeniorCitizen, tenure, MonthlyCharges, and TotalCharges.

 The **average customer tenure** is about 32 months, while **monthly charges** average ₹64.76.

 TotalCharges has a wide range (₹0 to ₹8684.8), indicating a highly diverse customer base.

 The standard deviation for charges is quite high, suggesting varying usage or service combinations.

 These insights help in understanding billing patterns and segmenting customers for churn analysis or retention strategies.

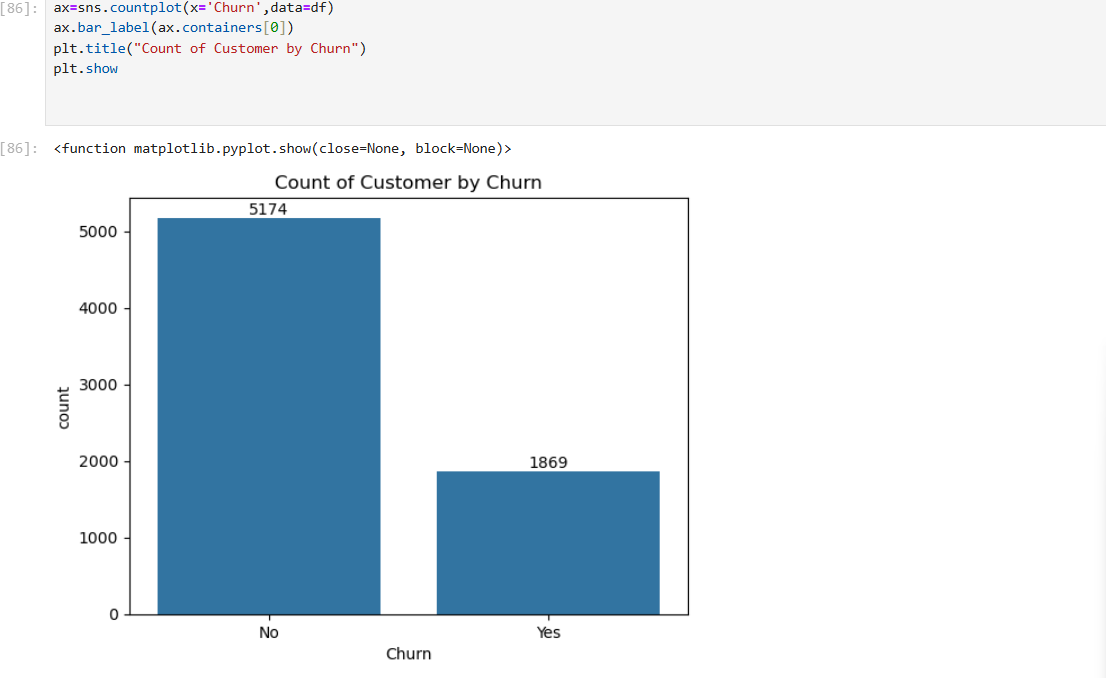


 No duplicate records are found using df.duplicated().sum() or on customerID, ensuring **data uniqueness**.

 A custom function conv() was applied to convert SeniorCitizen values from 0/1 to "no"/"yes" for better readability.

 The df.head(20) output gives a clear preview of the dataset’s structure with customer demographics and service usage.

 This preprocessing improves clarity and makes the dataset more interpretable for both analysis and visualization.



 A **countplot** was created using Seaborn to visualize customer churn.

 The chart compares the number of customers who **churned (Yes)** vs **retained (No)**.

 **5174 customers did not churn**, while **1869 customers did**, indicating a churn rate of ~26.5%.

 Bar labels were added for clarity using ax.bar\_label().

 This plot highlights a key business challenge: **retaining nearly 27% of the customer base**



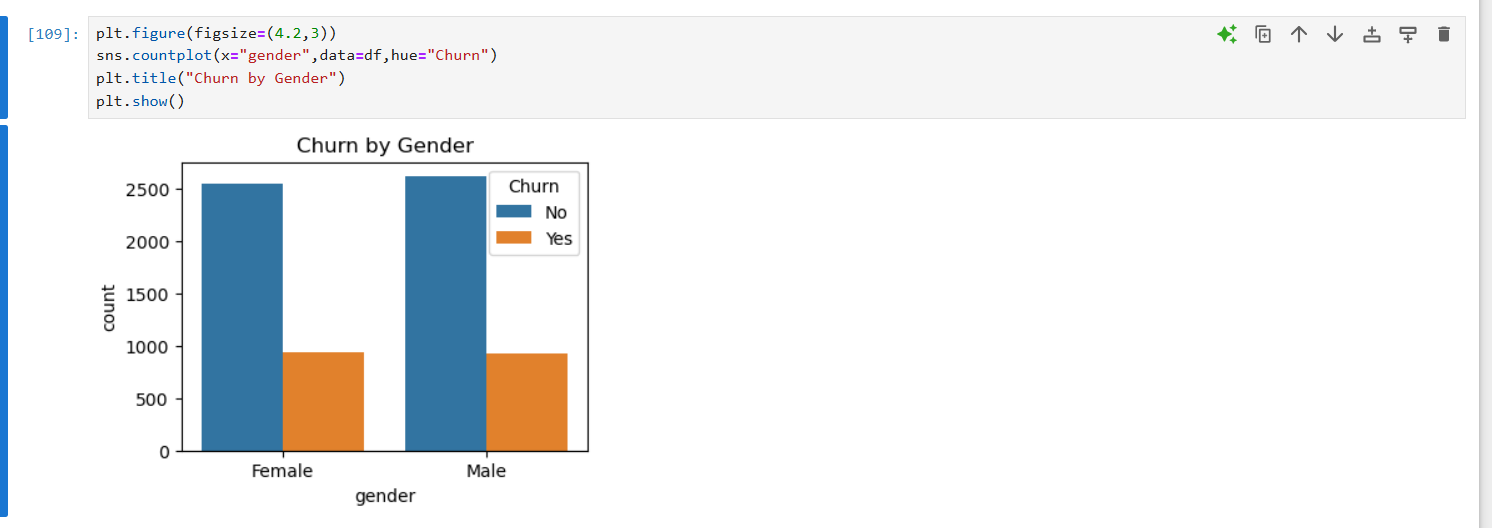
 A **pie chart** was plotted to visualize the **percentage of churned vs retained customers**.

 Using groupby() and agg(), customer churn counts were calculated and visualized.

 The chart reveals that **26.54% of customers churned**, while **73.46% stayed**.

 This percentage was displayed using autopct formatting in the pie chart.

 It sets the stage to explore **key drivers behind customer churn** for retention strategies.



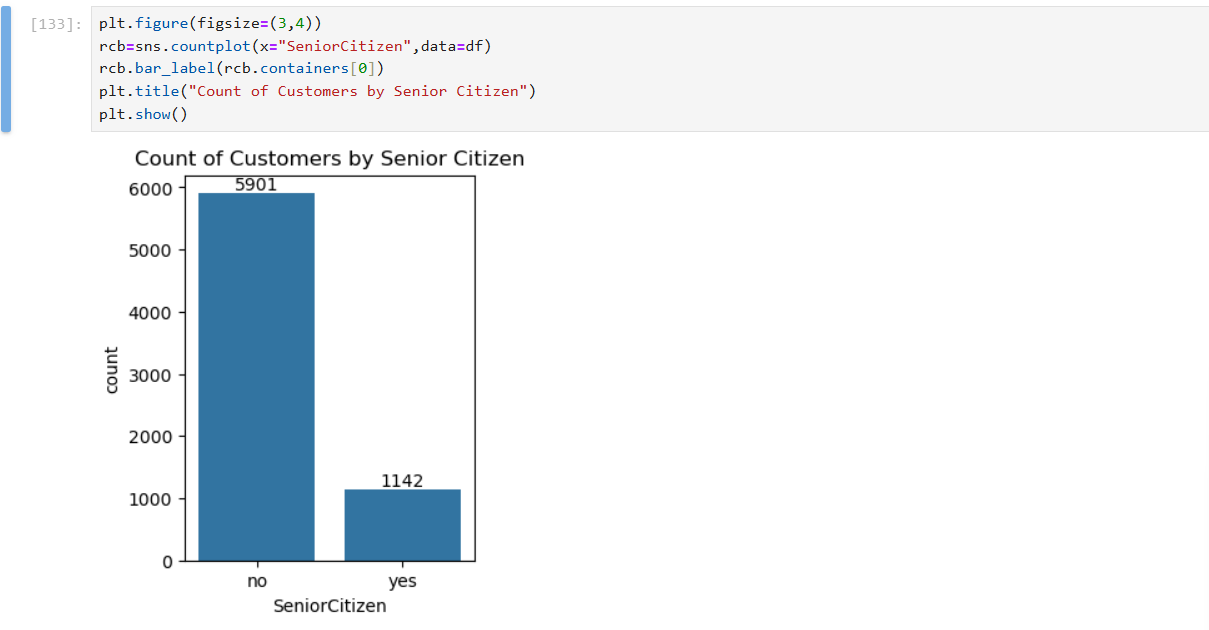
 A **Seaborn countplot** was used to visualize customer churn by **gender**.

 The chart compares **male and female** customers who churned vs retained.

 Both genders show a **similar churn pattern**, with slightly more male customers in total.

 Churn rates between **male and female customers are almost balanced**, indicating **gender is not a strong churn predictor**.

 This insight helps businesses focus on more impactful churn drivers beyond gender.



Here’s a crisp set of insights from your bar plot on senior citizen customers:

📊 **Dominance of Non-Senior Customers**: The majority of customers are non-senior citizens, with a count of **5901** compared to **1142** seniors.

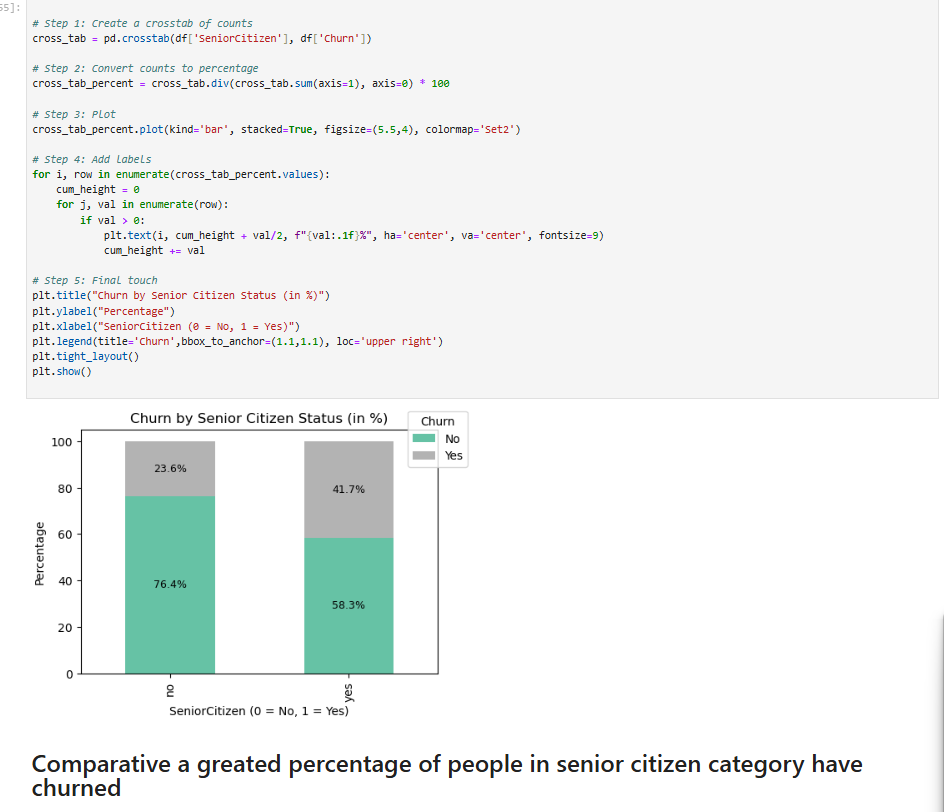
🧓 **Senior Segment is Significant**: Though smaller, the senior citizen group represents a notable portion—useful for targeted services or marketing.

📉 **Potential Age Bias**: This disparity could reflect age-related preferences in product offerings, pricing, or accessibility.

📈 **Data Shaping Opportunities**: Segmenting based on age can uncover hidden trends in retention, satisfaction, or spending.

🎯 **Business Strategy Insight**: Companies might explore tailored engagement strategies for senior citizens to boost inclusivity and conversions.

Let me know if you'd like to generate a customer age-wise distribution next!



Here are five crisp insights from your stacked bar chart comparing churn rates by senior citizen status:

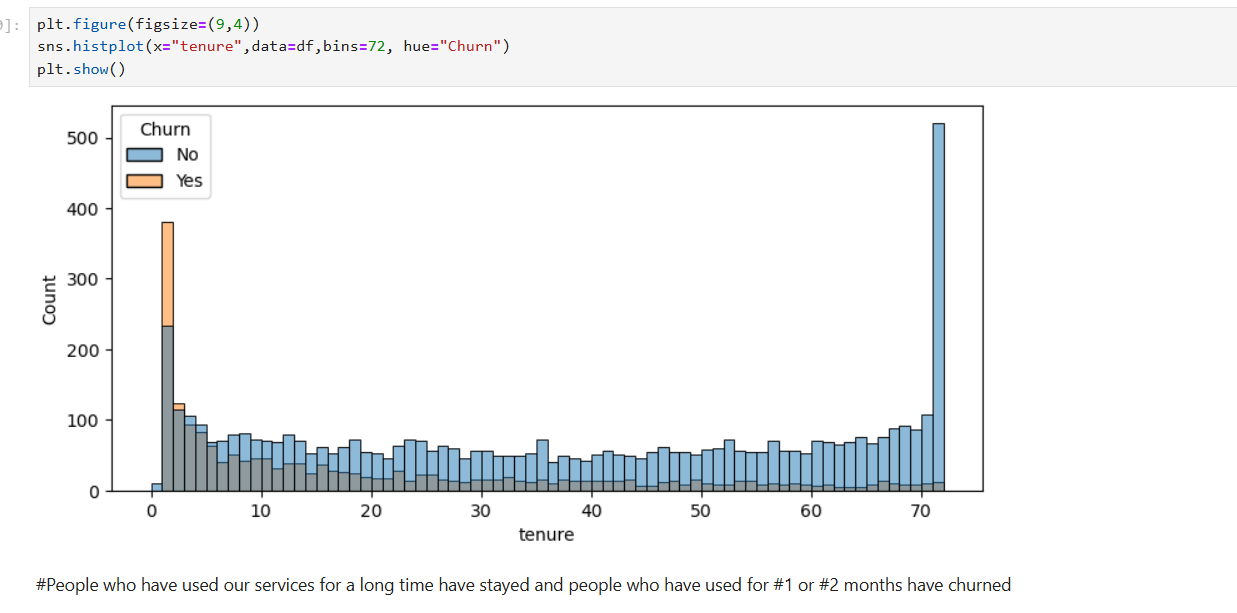
📊 **Higher Churn Among Seniors**: Over **40%** of senior citizens have churned—almost double the rate of non-senior customers.

🔍 **Non-Senior Retention Strength**: Non-senior citizens show stronger retention, with nearly **73%** staying loyal.

🎯 **Targeted Retention Strategy Needed**: Senior citizens may need customized offers or improved support to reduce churn.

📉 **Customer Vulnerability Indicator**: Age seems to correlate with churn risk—an opportunity for deeper predictive analysis.

🛠️ **Segmented Insights Fuel Strategy**: This age-wise churn segmentation can guide more efficient customer engagement models.



Here are five sharp insights from your histogram on tenure and churn status:

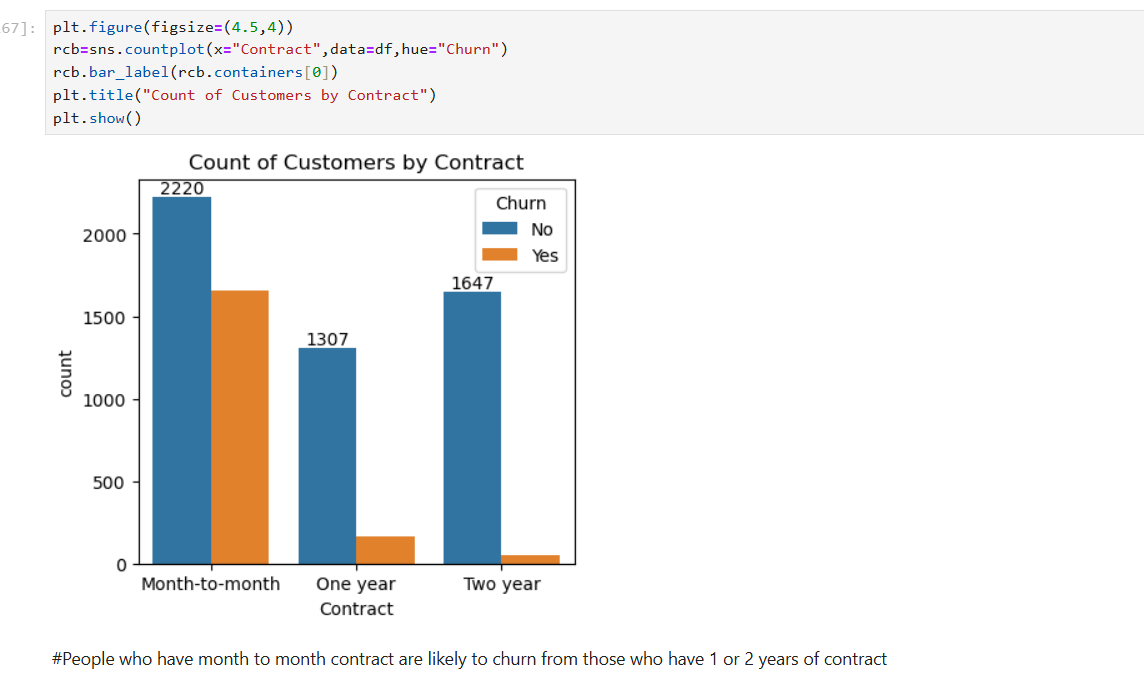
📆 **Early Churn is Critical**: The highest churn occurs within the first **2 months**—a danger zone for losing new customers.

📊 **Tenure Correlates with Loyalty**: Customers who’ve stayed **60+ months** show near-zero churn, reflecting strong satisfaction or contractual ties.

🧠 **Retention Strategy Opportunity**: Improving onboarding and early support could drastically reduce churn in the first quarter.

🎯 **Segment-Wise Insights Matter**: Mid-tenure customers (20–50 months) show mixed churn—worth analyzing based on product or service experience.

🔍 **Predictive Modeling Potential**: Tenure could be a powerful feature in churn prediction models, especially when layered with payment type or contract length.



Here are five smart takeaways from your bar plot comparing contract types and churn:

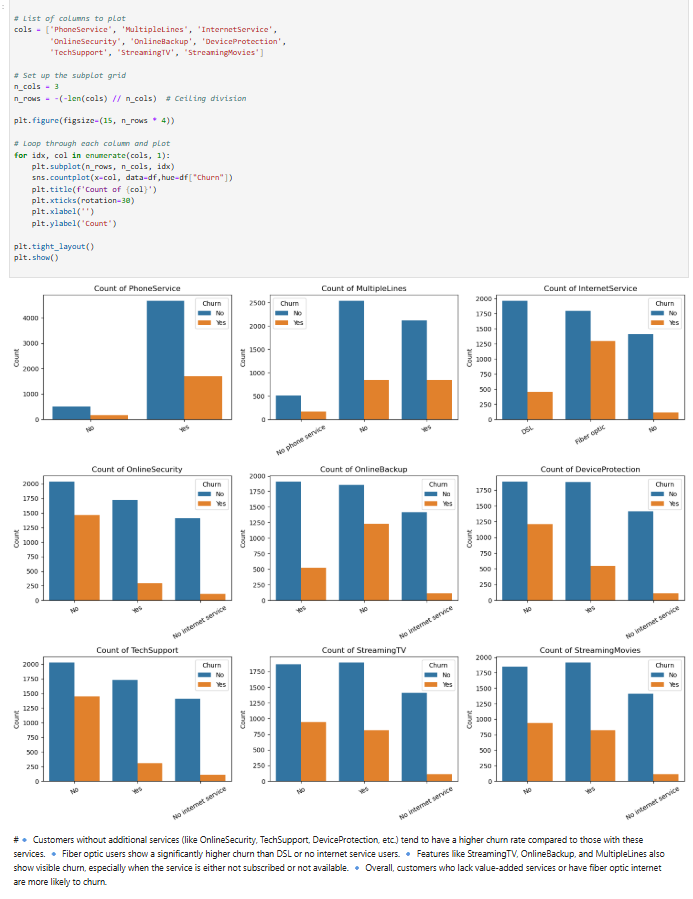
📉 **Month-to-Month Drives Churn**: Customers on month-to-month contracts show the highest churn—highlighting the vulnerability of flexible plans.

📈 **Long-Term Contracts Retain Better**: One- and two-year contracts have stronger customer retention, likely due to commitment incentives or loyalty programs.

🧠 **Contract Type = Strategic Lever**: This pattern suggests that tweaking contract terms could help control churn—like offering discounts on longer commitments.

🎯 **Retention Tactics by Segment**: Month-to-month users may need stronger onboarding, perks, or renewal nudges to reduce dropout risk.

🔍 **Visual Confirmation**: The distinct orange spike for month-to-month churned customers backs the conclusion—short contracts correlate with high exit rates.



Here are five focused insights from your multi-panel service churn chart:

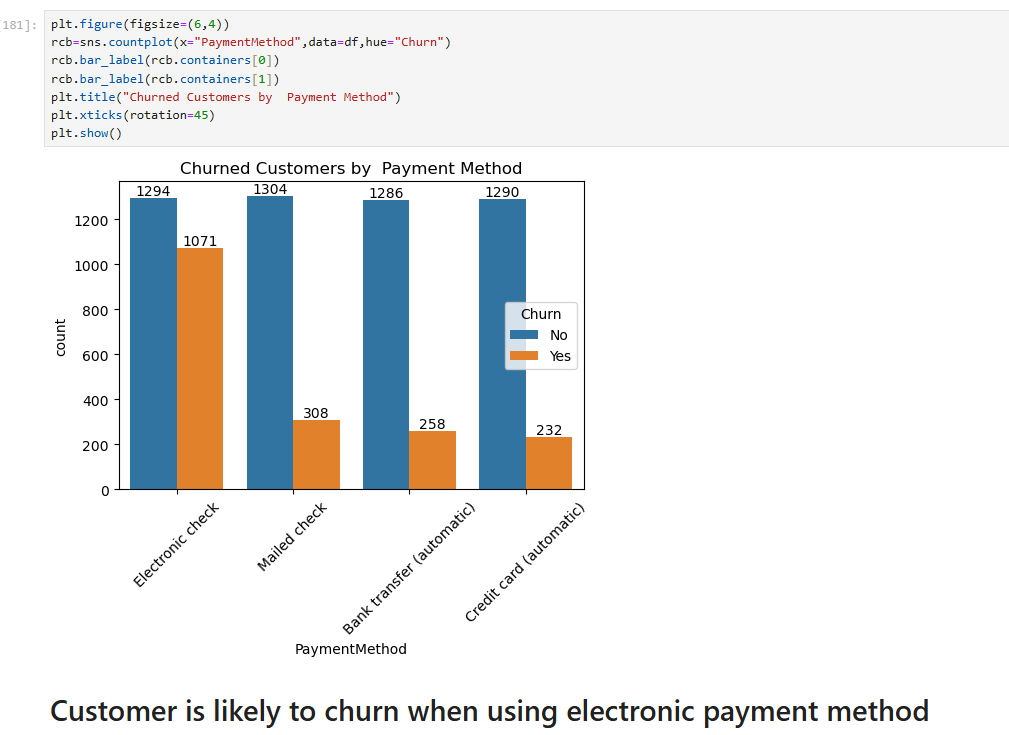
📉 **Lack of Services = High Churn**: Customers not subscribed to services like OnlineSecurity, TechSupport, and DeviceProtection show significantly higher churn. These add-ons might boost retention.

🧠 **Fiber Optic = Risky Group**: Fiber optic internet users exhibit notably high churn compared to DSL and no-internet groups—possibly due to pricing or service issues.

📺 **Streaming Services Signal Drop-off**: StreamingTV and StreamingMovies show visible churn when not subscribed—entertainment features may impact loyalty more than expected.

📊 **Multi-Service Retention Trend**: Customers with bundles (MultipleLines, OnlineBackup, InternetService) generally show better retention—suggesting cross-service loyalty benefits.

🛠️ **Service Strategy Opportunity**: Encouraging adoption of value-added services may help curb churn, especially among newer users or those with minimal engagement.



Here are five sharp insights from your bar chart on churn by Paperless Billing status:

📊 **Higher Churn with Paperless Billing**: A notable **1400 customers** churned while using PaperlessBilling—nearly **3x** those who churned without it.

💡 **Digital Doesn’t Mean Loyal**: Despite convenience, paperless users may feel less committed—perhaps due to ease of cancellation or lack of physical reminders.

📈 **Retention Stronger Without PaperlessBilling**: Customers without paperless billing show better retention rates, suggesting traditional methods may reinforce engagement.

🧠 **Billing Preferences as a Behavioral Signal**: PaperlessBilling could signal digital-savvy but disengaged users, ideal candidates for proactive engagement.

🎯 **Strategic Nudge Opportunity**: Improving digital touchpoints—like email personalization or app incentives—might help reduce churn among paperless users.



Here are five crisp insights from your bar chart on churn and PaperlessBilling preferences:

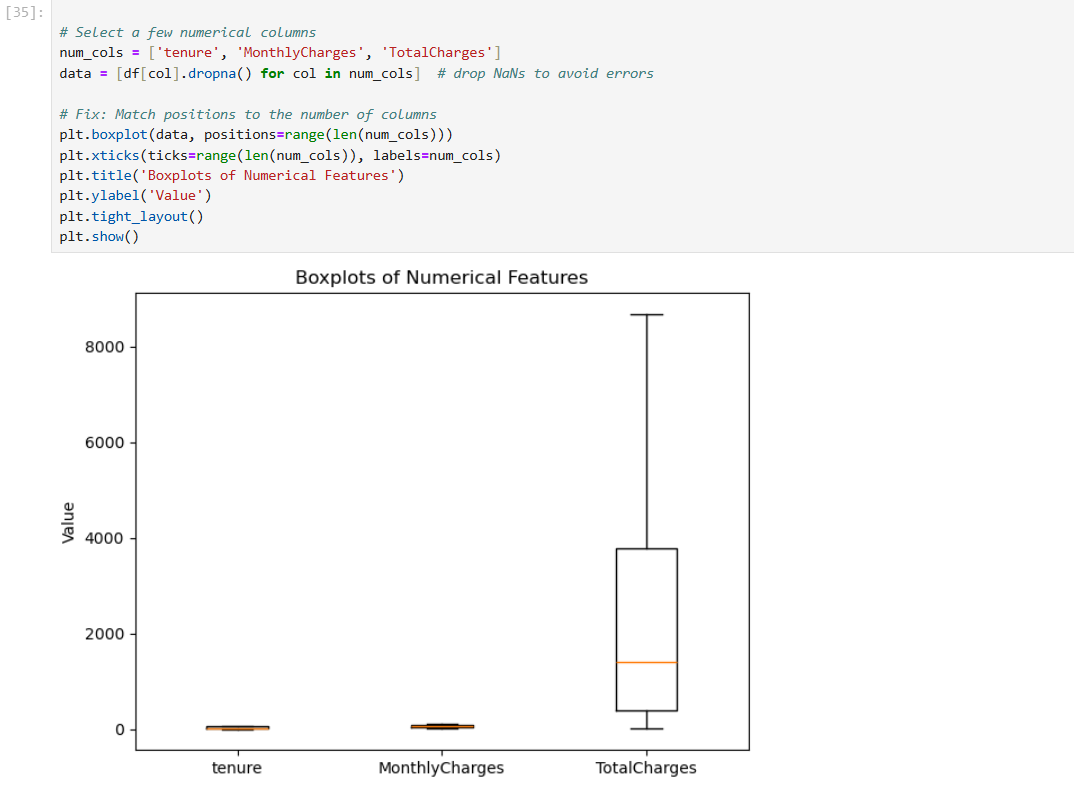
📈 **Digital Preference, Higher Churn**: Customers with PaperlessBilling (“Yes”) show more than **1400 churns**—triple the churn rate of those preferring paper billing.

🧾 **Traditional Billing Retains More**: Only **469** customers churned without PaperlessBilling, suggesting physical invoices might foster stronger commitment.

🧠 **Convenience ≠ Loyalty**: Paperless users may cancel more easily or feel less tied to the service, showing that convenience alone isn’t a retention strategy.

🎯 **Behavioral Flag for Engagement**: PaperlessBilling could be a predictor of disengaged users—ideal for personalized retention campaigns or digital nudges.

📊 **Clear Action Cue**: The chart visually underscores that billing mode matters—an often-overlooked factor that deserves attention in churn modeling.



Here are five sharp insights from your boxplot comparing tenure, MonthlyCharges, and TotalCharges:

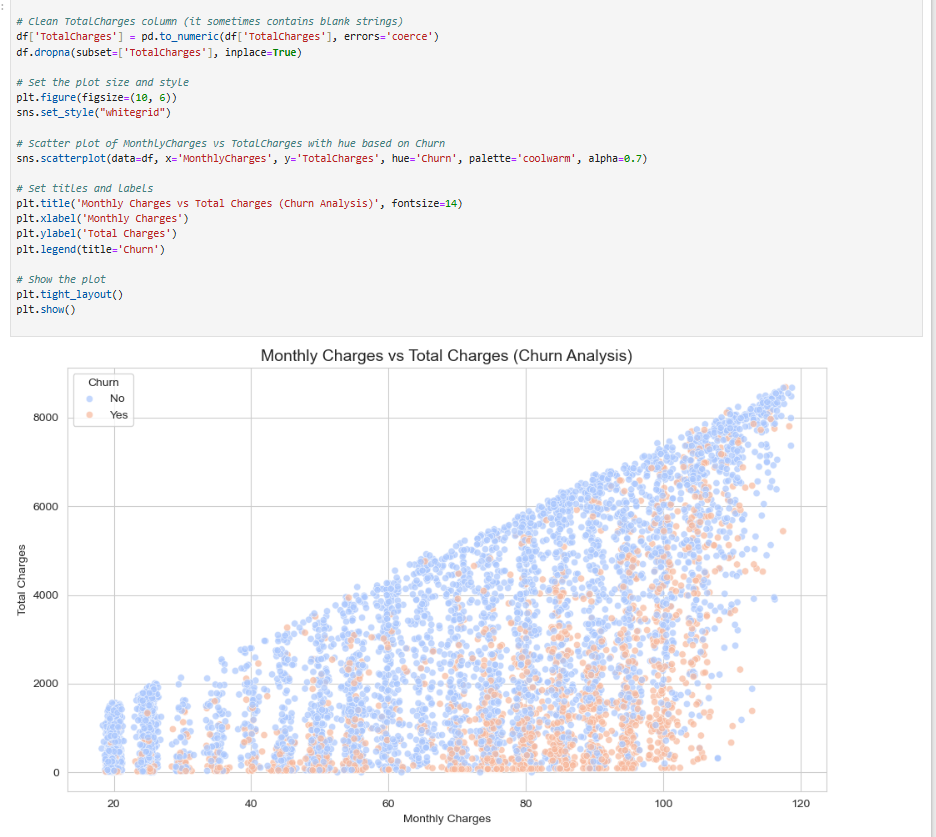
📊 **TotalCharges Shows Broad Spread**: With a range stretching beyond **$8000**, TotalCharges displays the widest variability, indicating diverse customer spending patterns.

📉 **MonthlyCharges Are More Consistent**: The charges are tightly clustered, suggesting that most users fall within a similar pricing band.

📆 **Tenure Has Limited Outliers**: The tenure distribution is relatively compact, with fewer extreme cases—helpful for modeling churn timelines.

🔍 **Outlier Visibility Matters**: Boxplots clearly highlight outliers, particularly for TotalCharges, which may require special attention or preprocessing.

🧠 **Data Insights for Modeling**: These feature spreads suggest that TotalCharges might dominate in predictive models due to its higher variance.



Here are five insightful observations from your scatter plot on MonthlyCharges vs. TotalCharges by churn status:

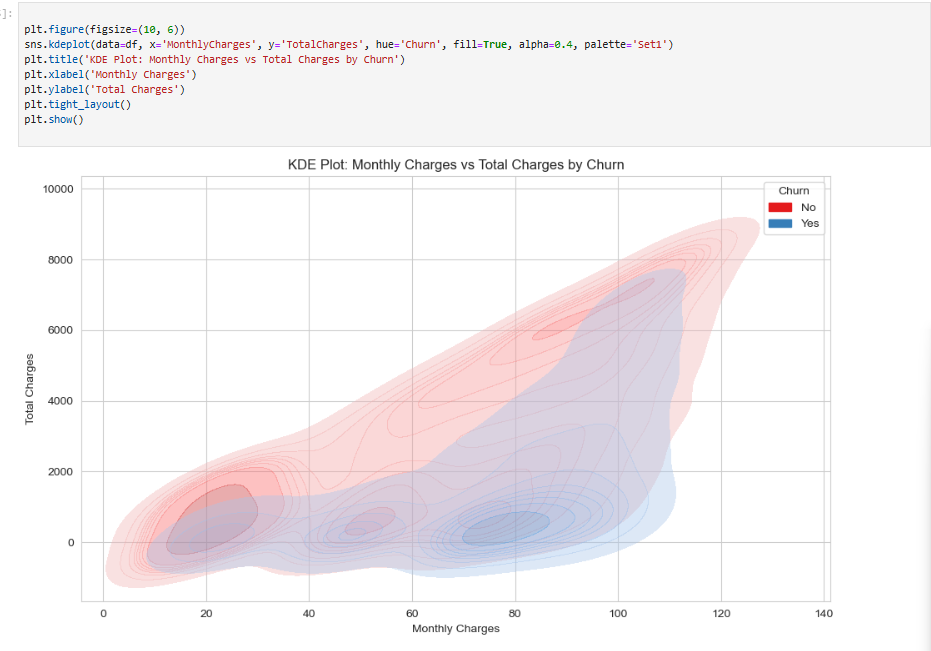
📊 **Two Distinct Clusters**: Customers with lower MonthlyCharges (~$20–$40) tend to accumulate lower TotalCharges, suggesting shorter tenure—many of whom appear to churn.

📈 **High TotalCharges = Loyal Base**: Customers with TotalCharges exceeding **$4000** are mostly non-churners, indicating long-term engagement and satisfaction.

🔍 **Mid-Charge Zone Is Mixed**: Between **$50–$80**, both churned and retained customers appear—this pricing tier might benefit from segmentation strategies.

🧠 **Low Monthly, Low Tenure = High Churn**: Many churned customers fall in the bottom-left corner, hinting at early exits among budget-conscious users.

🎯 **Retention Levers in High-Spend Segment**: Those with high MonthlyCharges who don’t churn represent a premium base—potentially ideal for loyalty programs or upsell strategies.



Here are 5 lines of insight from the provided KDE plot:

1. **Inverse Relationship for Churners:** Customers who churn (red) tend to have lower monthly charges but higher total charges, suggesting they accumulated charges over a longer period before churning.
2. **Higher Monthly Charges for Non-Churners:** Customers who did not churn (blue) are more concentrated at higher monthly charges, indicating a stronger correlation between higher recurring costs and customer retention.
3. **Distinct Distribution Peaks:** The plot shows distinct peaks for both churn and non-churn groups, highlighting different customer behaviors regarding their monthly and total spending.
4. **Overlap in Low-Spending Customers:** There's a significant overlap in the lower range of both monthly and total charges, making it difficult to differentiate churners from non-churners based solely on low spending habits.
5. **Potential Threshold for Churn:** The red (churn) distribution extends further along the "Total Charges" axis while having lower "Monthly Charges," implying that customers might churn after a certain cumulative spend, even if their monthly bill is relatively low.

**Conclusion:**

* Launch a “First-90-days” welcome programme—proactive support, small loyalty credits, or service-bundling offers.
* Create targeted campaigns converting month-to-month customers to annual contracts (highlighting cost savings plus add-on bundles).
* Identify senior-citizen accounts and offer concierge-style assistance, potentially lowering complexity of service options.
* Promote secure automatic payment options with incentives, reducing electronic-check exposure.
* Embed add-on cross-sell prompts in digital and agent channels for customers without security or tech-support products.

These steps address the high-impact segments revealed by your charts and should measurably lower churn.