

SUMMARY AND RECOMMENDATIONS

This analysis focuses on understanding the factors driving customer churn within a dataset of **7,043** customer records and **21** variables. The core objective of the exploratory data analysis (EDA) was to uncover patterns, validate assumptions, and prepare the data for predictive modeling.



Data Overview

- **Size:** 7,043 rows and 21 columns.
- **Initial Data Types:** Most columns are of `object` (categorical) type, with `tenure`, `SeniorCitizen` (`int64`), `MonthlyCharges` (`float64`), and `TotalCharges` (`object`, due to initial blanks) being numeric.
- **Data Cleaning:**
 - **Data Issue Detected:** The `TotalCharges` column was initially an `object` type because it contained blank (space) values, which corresponded to customers with a `tenure` of 0.
 - **Resolution:** The blank values were **replaced with '0'** and the column was successfully **converted to a `float64`** type for numerical analysis.
 - **Missing Values:** After cleanup, a check (`df.isnull().sum().sum()`) confirmed **zero missing values** across the entire dataset.
 - **Duplicates:** The analysis of the `customerID` column confirmed **zero duplicate customer records**.
- **Feature Transformation:** The binary `SeniorCitizen` variable was converted from its original integer format (0/1) to more intuitive **'no'/'yes'** categorical labels for improved visualization and interpretation.

Descriptive Statistics Insights

- **Tenure:** The average customer tenure is **32.37 months**, with a minimum of 0 (new customers) and a maximum of 72 months.
 - **Monthly Charges:** The mean monthly charge is **\$64.76**. The 25th percentile is \$35.50 and the 75th percentile is \$89.85, indicating a wide range of pricing plans.
 - **Total Charges:** The mean total charge is **\$2,279.73**. The large standard deviation (\$2,266.79) is expected, as it is heavily influenced by the `tenure` length.
-

Trends and Patterns in Churn

Overall Churn Rate

The overall churn rate in the dataset serves as the baseline for the analysis:

- **Churned: 26.54%** (1,869 customers)
- **Retained: 73.46%** (5,174 customers)

Key Churn Drivers (Validated by Visualizations)

Factor	Churning Sub-group	Churn Rate in Group	Trend/Hypothesis Test
Contract	Month-to-month	42.71% (1,655 out of 3,875)	Critical Driver. Customers on flexible, short-term contracts are significantly less loyal.
	One year	11.27% (166 out of 1,473)	
	Two year	3.01% (48 out of 1,695)	
Payment Method	Electronic check	45.29% (1,294 out of 2,850)	Strong Driver. This payment method is the riskiest, potentially due to processing friction or the type of customer it attracts.
	Mailed check	19.11% (308 out of 1,612)	
	Bank transfer (automatic)	16.71% (258 out of 1,544)	

	Credit card (automatic)	15.25% (232 out of 1,522)	
Tenure	Short-term (1-2 months)	Highest Concentration of churn.	Strong Driver. Churn is an acute issue in the first few months of service, suggesting potential onboarding or initial service quality issues.
Internet Service	Fiber optic	41.89% (1,297 out of 3,096)	Key Service Issue. The high churn for this premium service suggests major service quality/reliability problems specific to the fiber infrastructure.
	DSL	18.99% (459 out of 2,421)	
	No Internet Service	7.40% (113 out of 1,526)	
Senior Citizen	Yes (Senior Citizen)	41.68% (476 out of 1,142)	Noteworthy Demographic. Senior citizens are a disproportionately high-churn demographic.
	No	23.61% (1,393 out of 5,901)	
Gender	Female	26.60% (939 out of 3,555)	Not a Driver. Churn rates are nearly identical across genders.

	Male	26.16% (930 out of 3,488)	
--	------	---------------------------	--

The analysis highlights that **Fiber optic internet, month-to-month contracts, Electronic check payments, and lack of foundational support services** are the biggest drivers of customer churn. Senior Citizens are also a high-risk demographic group.

To reduce the overall churn rate, the company should consider:

1. **Contract Strategy:** Incentivize all customers, especially those on a month-to-month plan, to transition to **one-year or two-year contracts**.
2. **Service Quality:** Investigate potential reliability issues and improve service for **Fiber optic** internet users.
3. **Customer Protection & Support:** Push bundled packages that include **Online Security** and **Tech Support** to significantly reduce the risk of churn.
4. **Payment Method Migration:** Encourage the adoption of **automatic payment methods** (Bank Transfer or Credit Card) to move customers away from Electronic and Mailed checks.

Service Subscription Impact (Hypothesis: More services = Higher Loyalty)

The visualizations largely validate the hypothesis that **a lack of value-added services correlates with increased churn risk**.

- Customers **without Online Security** had a churn rate of **41.77%**.
- Customers **without Tech Support** had a churn rate of **41.64%**.
- Customers **with Online Backup** churned at **21.50%**, vs. **39.93%** for those *without*.