Data Preprocessing Documentation: Detailed steps and methods used for data cleaning and preparation.

**1. Churn Rate and Customer Demographics**

* **Queries**: Identify the total number of customers and calculate the churn rate.
* **Data Preparation**:
  + Filtered CustomerStatus to separate churned from active customers.
  + Ensured completeness in CustomerID for an accurate count.
* **Purpose**: Establishes baseline churn rate for demographic analysis.

**2. Average Age of Churned Customers**

* **Queries**: Calculate the average age of churned customers.
* **Data Preparation**:
  + Validated Age entries for consistency and accuracy.
  + Filtered to include only churned customers in calculations.
* **Purpose**: Provides insights into the age demographic most likely to churn.

**3. Most Common Contract Types Among Churned Customers**

* **Queries**: Identify the most common contract types among churned customers.
* **Data Preparation**:
  + Standardized Contract terms for uniformity.
  + Filtered data by CustomerStatus to focus on churned customers.
* **Purpose**: Highlights contract types that could be contributing to churn.

**4. Monthly Charges Distribution for Churned Customers**

* **Queries**: Analyze the distribution of monthly charges among churned customers.
* **Data Preparation**:
  + Validated Monthly Charges entries for non-null values.
  + Handled outliers to avoid skewed distribution results.
* **Purpose**: Understands typical spending ranges of churned customers.

**5. Contract Types Most Prone to Churn**

* **Queries**: Identify contract types that are most prone to churn.
* **Data Preparation**:
  + Ensured completeness in Contract and CustomerStatus fields.
  + Grouped data by Contract to assess correlation with churn.
* **Purpose**: Helps in assessing which contracts require attention for churn management.

**6. High Total Charges Among Churned Customers**

* **Queries**: Identify customers with high total charges who have churned.
* **Data Preparation**:
  + Validated entries in Total Charges for accuracy.
  + Filtered for churned customers, ensuring no nulls in Total Charges.
* **Purpose**: Identifies high-spending churned customers, useful for retention strategies.

**7. Total Charges Distribution by Churn Status**

* **Queries**: Calculate total charges distribution for churned and non-churned customers.
* **Data Preparation**:
  + Segmented data by CustomerStatus for comparison.
  + Checked for outliers in Total Charges to avoid skewed results.
* **Purpose**: Analyzes spending differences between churned and active customers.

**8. Average Monthly Charges by Contract Type for Churned Customers**

* **Queries**: Calculate average monthly charges for different contract types among churned customers.
* **Data Preparation**:
  + Standardized Contract entries to avoid duplication.
  + Verified non-null entries in Monthly Charges.
* **Purpose**: Shows how contract type influences churned customers’ monthly expenses.

**9. Non-Churned Customers with Online Security and Backup Services**

* **Queries**: Identify non-churned customers with both online security and backup services.
* **Data Preparation**:
  + Replaced nulls with ‘No’ in service-related fields for consistency.
  + Filtered for CustomerStatus = active and OnlineSecurity and OnlineBackup = ‘Yes’.
* **Purpose**: Evaluates if specific service combinations aid in retention.

**10. Common Service Combinations Among Churned Customers**

* **Queries**: Determine the most common combinations of services among churned customers.
* **Data Preparation**:
  + Standardized yes/no entries in service columns; replaced nulls with ‘No’.
  + Filtered by CustomerStatus = churned to isolate trends.
* **Purpose**: Identifies service patterns potentially influencing churn.

**11. Average Total Charges by Gender and Marital Status**

* **Queries**: Calculate average total charges for customers grouped by gender and marital status.
* **Data Preparation**:
  + Checked for consistency in Gender and Marital Status.
  + Ensured no nulls in Total Charges for accurate averages.
* **Purpose**: Analyzes customer spending patterns by demographic factors.

**12. Monthly Charges by Age Group for Churned Customers**

* **Queries**: Calculate the average monthly charges for different age groups among churned customers.
* **Data Preparation**:
  + Created age group segments for classification.
  + Ensured Age and Monthly Charges data integrity.
* **Purpose**: Assesses how monthly charges vary by age group among churned customers.

**13. Age and Charges for Customers with Multiple Lines and Online Backup**

* **Queries**: Determine average age and total charges for customers with multiple lines and online backup.
* **Data Preparation**:
  + Verified multiple lines and backup service fields for accuracy.
  + Filtered by service combination and cleaned data for Age and Total Charges.
* **Purpose**: Explores the impact of bundled services on customer demographics and spending.

**14. Churn Rate by Contract Type for Senior Citizens**

* **Queries**: Identify the contract types with the highest churn rate among senior citizens (65+).
* **Data Preparation**:
  + Filtered data for Age >= 65 and validated Contract entries.
* **Purpose**: Analyzes contract preferences and churn rate for older customers.

**15. Monthly Charges for Customers with Multiple Lines and Streaming TV**

* **Queries**: Calculate average monthly charges for customers with multiple lines and streaming TV.
* **Data Preparation**:
  + Checked Multiple Lines and Streaming TV fields for consistency.
  + Ensured Monthly Charges values are complete and accurate.
* **Purpose**: Understands how bundled services affect monthly charges.

**16. Online Service Usage Among Churned Customers**

* **Queries**: Identify churned customers using the most online services.
* **Data Preparation**:
  + Created count of active online services; replaced nulls with 'No'.
* **Purpose**: Highlights service engagement among churned customers.

**17. Age and Total Charges by Streaming Service Combinations**

* **Queries**: Calculate average age and total charges for different combinations of streaming services.
* **Data Preparation**:
  + Verified Age and Total Charges; filtered data by distinct streaming services.
* **Purpose**: Understands the relationship between streaming service combinations and spending.

**18. Gender Distribution Among Yearly Contract Churned Customers**

* **Queries**: Identify gender distribution among customers who have churned on yearly contracts.
* **Data Preparation**:
  + Filtered by Contract type and churned status; standardized gender entries.
* **Purpose**: Investigates if gender influences churn in yearly contracts.

**19. Monthly and Total Charges by Contract Type and Internet Service (Churned)**

* **Queries**: Calculate average monthly and total charges by contract and internet service type for churned customers.
* **Data Preparation**:
  + Ensured consistent entries in Contract, Internet Service, Monthly Charge, and Total Charges.
* **Purpose**: Analyzes the impact of contract and service types on spending patterns among churned customers.

**20. Churned Customers Without Online Services and Their Charges**

* **Queries**: Identify churned customers without online services and calculate their average total charges.
* **Data Preparation**:
  + Set nulls to 'No' in online services fields; filtered by non-users.
* **Purpose**: Investigates churn reasons among minimal-service users.

**21. Average Monthly and Total Charges by Number of Dependents (Churned)**

* **Queries**: Calculate average monthly and total charges for churned customers grouped by the number of dependents.
* **Data Preparation**:
  + Ensured complete entries in dependents and charges fields.
* **Purpose**: Assesses financial patterns linked to family size.

**22. Contract Duration for Monthly Contracts (Churned)**

* **Queries**: Identify churned customers’ contract duration in months for monthly contracts.
* **Data Preparation**:
  + Filtered by monthly contracts; calculated duration in months.
* **Purpose**: Understands contract duration trends among churned monthly customers.

**23. Average Age and Total Charges by Internet and Phone Service**

* **Queries**: Calculate average age and total charges for churned customers grouped by internet and phone service.
* **Data Preparation**:
  + Verified data integrity in service-related fields.
* **Purpose**: Analyzes the impact of combined internet and phone services on age and spending.

**24. View for Customers with Highest Monthly Charges by Contract Type**

* **Queries**: Create view to find the customers with the highest monthly charges in each contract type.
* **Data Preparation**:
  + Defined highest charges by contract type; validated monthly charge data.
* **Purpose**: Facilitates quick comparison of high-value contracts by type.

**25. View for Average Monthly Charges Comparison Among Churned Customers**

* **Queries**: Create view to compare churned customers’ average monthly charges with the overall average.
* **Data Preparation**:
  + Ensured consistency in Monthly Charges and calculated overall average.
* **Purpose**: Allows comparison of churned customers’ spending relative to all customers.

**26. Cumulative Total Charges Over Time for Churned Customers (View)**

* **Queries**: Create a view to track churned customers’ cumulative total charges over time.
* **Data Preparation**:
  + Ensured time-series data completeness and accurate Total Charges.
* **Purpose**: Tracks spending growth over time for churned customers.

**27. Stored Procedure for Churn Rate Calculation**

* **Queries**: Develop a stored procedure to calculate churn rate.
* **Data Preparation**:
  + Defined a churn flag for clarity; checked for CustomerStatus completeness.
* **Purpose**: Automates the calculation of the churn rate.

**28. Stored Procedure to Identify High-Value Customers at Risk of Churning**

* **Queries**: Develop a stored procedure to identify high-value customers at risk of churning.
* **Data Preparation**:
  + Defined criteria for high-value status; flagged at-risk customers.
* **Purpose**: Identifies potential high-revenue customers who may churn.