

TeleConnect Customer Churn and Revenue Risk Analysis

Complete SQL Portfolio Project Documentation

Project detail	Value
Author	Ernest-Cyril Divine Onyeze
Database	MySQL Server 8.0
Client	MySQL Workbench 8.0 CE
Dataset	7,043 fictional telecom customers in California
Completed	18 September 2026

Portfolio: cyrilanalytics.com

Executive Summary

This project uses MySQL to analyse customer churn, historical customer value, service adoption, and retention priorities for TeleConnect, a fictional telecommunications company. A flat 38-column customer file was converted into a normalized five-table relational database. The workflow includes staged data loading, integrity checks, cleaning decisions, basic and advanced SQL analysis, reusable views, and business recommendations.

The verified data contains 7,043 customers: 1,869 churned, 4,720 stayed, and 454 recently joined. Churn was 26.54% across all customers and 28.37% when recent joins were excluded. The strongest descriptive patterns involved contract type, early tenure, competitor pressure, payment method, internet type, and service-bundle depth.

Metric	Verified result
Total customers	7,043
Churned customers	1,869
Overall churn rate	26.54%
Churn rate excluding recent joins	28.37%
Historical revenue from churned customers	\$3,684,459.82
Priority-retention segment	442 active customers

Business Problem and Objectives

TeleConnect needs to understand who is leaving, why churn occurs, which customer groups require attention, and how SQL analysis can support retention decisions. The project addresses the following objectives:

- Measure customer status and define churn-rate denominators clearly.
- Compare churn across contract, payment method, internet type, tenure, service adoption, revenue tier, and location.
- Identify the most frequent churn categories and specific reasons.
- Quantify historical revenue associated with churned customers without misrepresenting it as future loss.
- Create a transparent rule-based segment for retention outreach.
- Demonstrate reproducible SQL, relational modelling, validation, and documentation practices.

Data Source and Scope

The project uses the Maven Analytics Telecom Customer Churn dataset, based on IBM's fictional telecommunications sample. The source represents a customer snapshot at the end of Q2 2022 and includes demographics, location, subscribed services, billing, customer status, churn categories, and churn reasons.

Source file	Rows	Purpose
telecom_customer_churn.csv	7,043	Customer demographics, services, billing, and churn outcome
telecom_zipcode_population.csv	1,671	California population by ZIP code
telecom_data_dictionary.csv	40 definitions	Field descriptions and business meaning

Dataset source: <https://mavenanalytics.io/data-playground/telecom-customer-churn>

Database Design

The flat source was normalized into five core tables so that demographics, services, billing, churn outcomes, and ZIP-level population could be maintained and queried separately. The design supports meaningful joins and reduces repeated analytical logic.

Table	Purpose	Key
customers	Demographics and customer location	customer_id
customer_services	Phone, internet, streaming, and support services	customer_id
customer_billing	Tenure, contract, payment, charges, and historical revenue	customer_id
customer_churn_status	Status, churn category, and churn reason	customer_id
zip_population	Population by California ZIP code	zip_code

SQL Analysis

Six ordered SQL scripts create the schema, load the data, validate quality, answer business questions, apply advanced techniques, and create reusable views.

Technique	Application
Conditional aggregation	Customer counts and churn rates within grouped segments
Joins	Combined customer, billing, service, status, and population records
CTEs	Contract averages, tenure buckets, and location summaries
Subqueries	Company-wide and segment-level comparison thresholds
Window functions	Revenue quartiles, ranks, and cumulative historical revenue
Views	Customer 360, standardized segment rates, and retention targeting

Verified Results

Customer status and churn rate

Of 7,043 customers, 1,869 churned, 4,720 stayed, and 454 recently joined. Churn is 26.54% when all customers are included. Excluding customers classified as newly joined produces a 28.37% rate across customers who had stayed or churned.

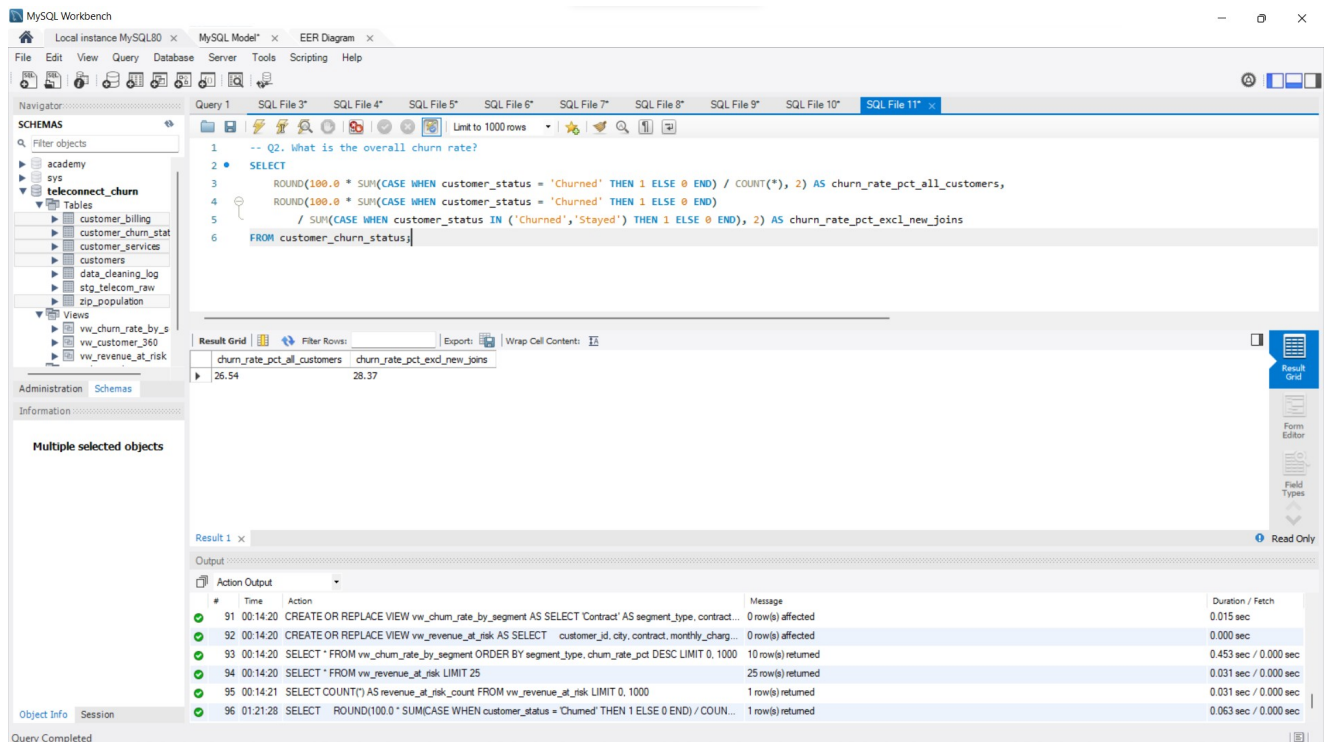


Figure 2. MySQL query and result showing both churn-rate definitions.

Churn by contract

Contract	Customers	Churned	Churn rate
Month-to-Month	3,610	1,655	45.84%
One Year	1,550	166	10.71%
Two Year	1,883	48	2.55%

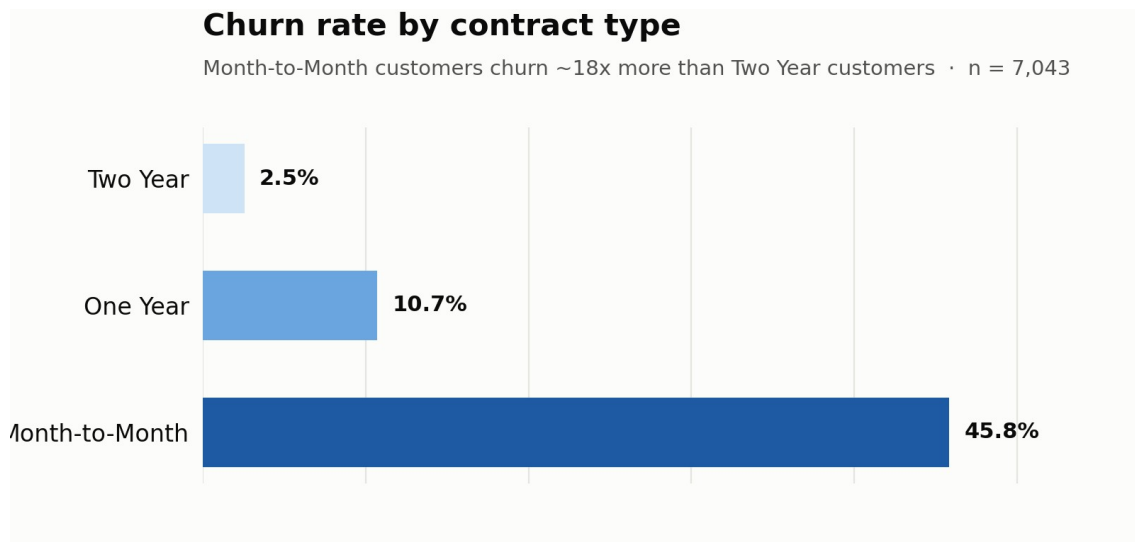


Figure 3. Churn rate by contract type.

Churn by tenure

Tenure	Customers	Churned	Churn rate
0-12 months	1,732	1,037	59.87%
13-24 months	1,024	294	28.71%
25-48 months	1,594	325	20.39%
49+ months	2,239	213	9.51%

Churn rate by tenure stage

Nearly 60% of first-year customers churn; risk drops sharply after year one

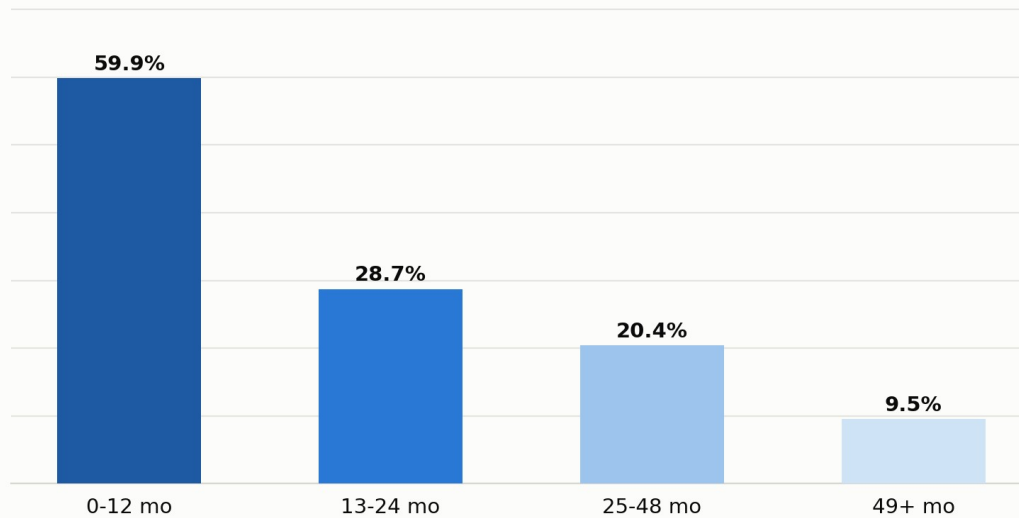


Figure 4. Churn rate by customer tenure.

Payment method and internet type

Segment	Churn rate
Mailed Check	36.88%
Bank Withdrawal	34.00%
Credit Card	14.48%
Fiber Optic	40.72%
Cable	25.66%
DSL	18.58%
No Internet	7.40%

Churn reasons

Competitor-related reasons accounted for 841 customers, or 45.00% of churn. The leading specific reasons were better devices from a competitor (313), a better competitor offer (311), attitude of a support person (220), an unknown reason (130), and more data from a competitor (117).

Service add-ons

Among internet customers, churn decreased as adoption increased across online security, online backup, device protection, and premium technical support. The pattern is descriptive and may also reflect tenure, contract, or customer-need differences.

Measured add-ons	Customers	Churned	Churn rate
0	1,130	718	63.54%

1	1,370	570	41.61%
2	1,340	326	24.33%
3	935	117	12.51%
4	470	25	5.32%

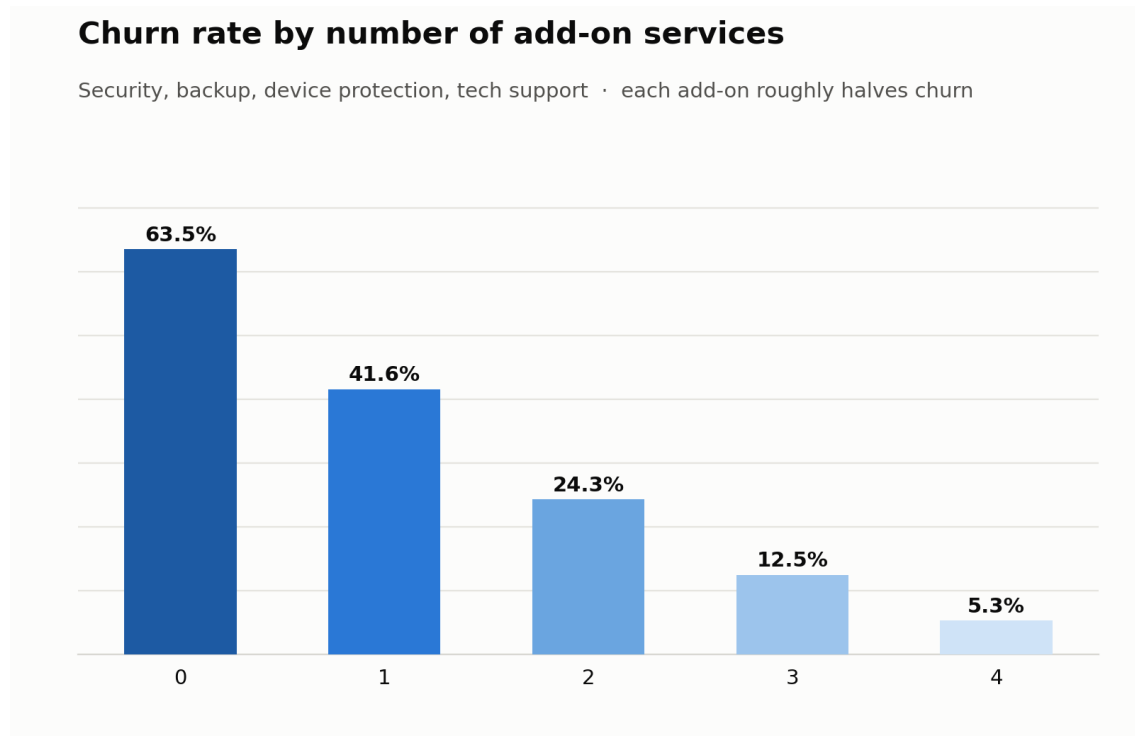


Figure 5. Churn rate by measured service add-on count.

Historical customer value

Churned customers had generated \$3,684,459.82 in cumulative historical revenue, averaging \$1,971.35 per customer. Customers who stayed had generated \$17,632,392.12, averaging \$3,735.68. These amounts describe revenue accumulated through the snapshot date; they are not forecasts of revenue that will be lost after churn.

The top ten churned customers by historical revenue had generated \$107,437.62 combined. The advanced SQL uses `RANK()` and a running `SUM() OVER()` to preserve row-level detail while showing rank and cumulative value.

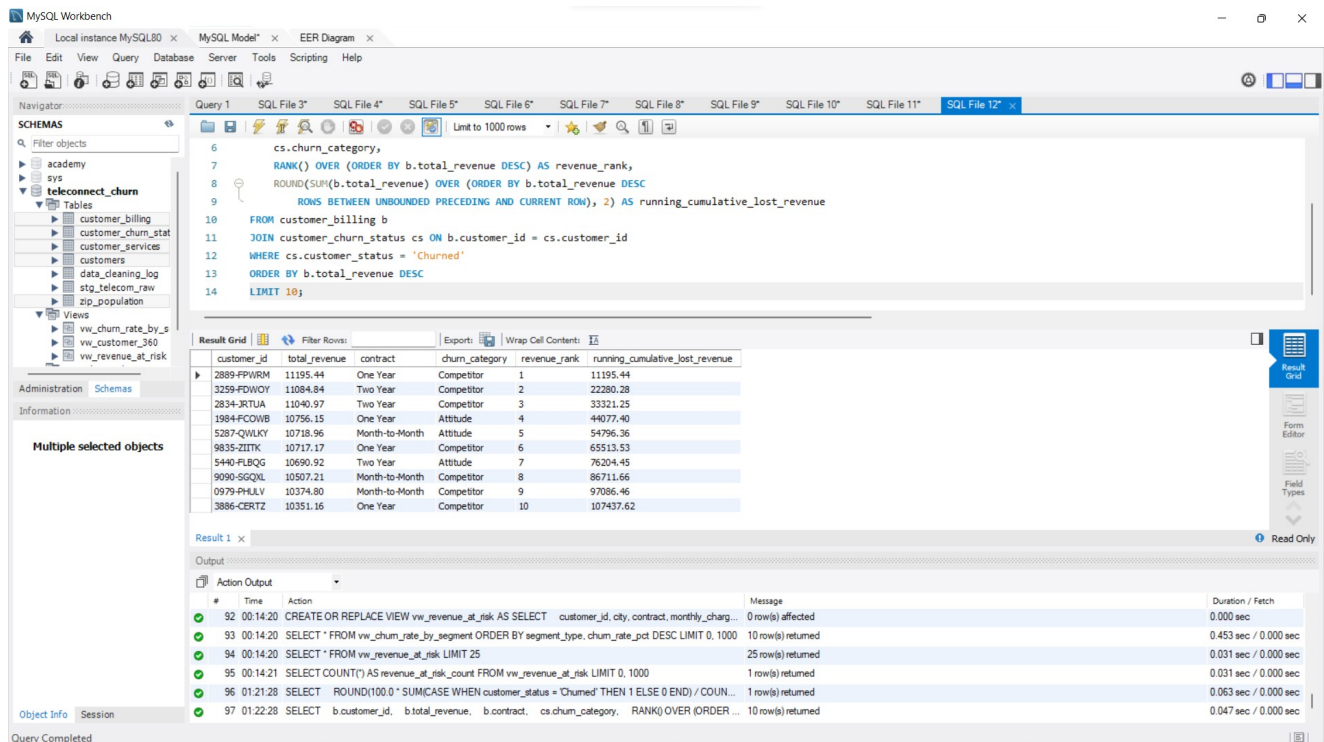


Figure 6. MySQL Workbench evidence for the ranking and cumulative historical-revenue query. The original result alias uses “lost revenue”; the audited interpretation is historical revenue from churned customers.

Priority-retention segment

A transparent rule identified 442 customers who were classified as Stayed, held Month-to-Month contracts, and had total historical revenue above the company average. This segment is useful for targeted outreach but is not a predictive model and does not establish that every selected customer will churn.

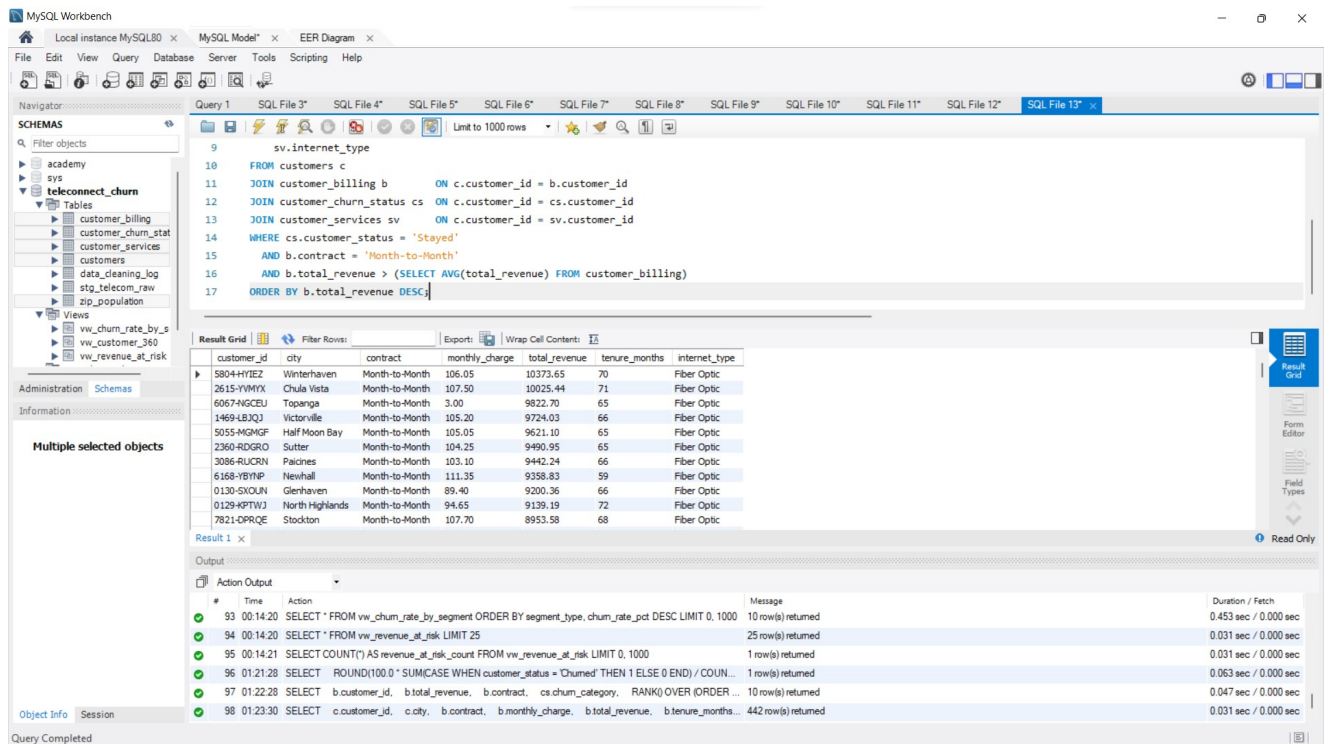


Figure 7. Multi-table SQL query returning the 442-customer priority-retention segment.

Business Recommendations

1. Introduce structured onboarding and first-year retention contacts before months 3, 6, and 12.
2. Test suitable incentives for Month-to-Month customers to move to longer contracts.
3. Review Fiber Optic reliability, support cases, pricing, and competitor comparisons.
4. Develop competitor-response offers around devices, plan value, and data allowances.
5. Test add-on bundles and measure both retention impact and contribution margin.
6. Contact the 442 priority customers with tailored outreach and compare outcomes with a control group.
7. Investigate concentrated ZIP-code patterns while accounting for small local sample sizes.

Implementation Challenges

Challenge	Resolution
Workbench import wizard was extremely slow	Used server-side `LOAD DATA INFILE` for bulk loading
`LOAD DATA LOCAL INFILE` returned Error 2068	Used the server `secure_file_priv` directory
Blank usage fields failed numeric import	Loaded through variables and converted blanks with `NULLIF`
Correlated subquery exceeded the read timeout	Calculated contract averages once in a CTE and joined them
Ambiguous negative monthly charges	Preserved, flagged, and excluded from charge averages

Limitations

- The dataset is a single-quarter snapshot and cannot measure individual customer changes through time.
- The findings identify associations and do not prove causality.
- Historical total revenue is not equivalent to future recurring revenue loss.
- The priority-retention segment is a rule-based shortlist rather than a predicted probability of churn.
- The true values of 120 negative monthly-charge records cannot be established from the supplied source.
- ZIP-code hotspot results contain relatively small customer counts and require cautious interpretation.

Portfolio Deliverables

The repository includes six SQL scripts, three source datasets, query-result exports, a data-quality summary, the MySQL EER diagram, genuine MySQL Workbench screenshots, analytical charts, supporting Markdown documentation, this complete report, and publishing copy for the portfolio website, CV, and LinkedIn.

The screenshot displays the MySQL Workbench interface. On the left, the 'SCHEMAS' pane shows a tree view of the database structure, including tables like 'customer_billing', 'customer_churn_stat', 'customer_services', 'customers', 'data_cleaning_log', 'stg_telecom_raw', 'zip_population', and views like 'vw_churn_rate_by_segment', 'vw_customer_360', 'vw_revenue_at_risk', and 'Stored Procedures'. The main window shows a SQL query in the 'Query 1' tab, which is a complex query involving multiple tables and views. Below the query, the 'Result Grid' shows the output of the query, which is a table with columns: 'segment_type', 'segment_value', 'customers', 'churned', and 'churn_rate_pct'. The results are grouped by 'segment_type' and 'segment_value'. The bottom pane shows the 'Output' tab, which displays the execution log of the query, including the time taken for each step and the number of rows affected or returned.

segment_type	segment_value	customers	churned	churn_rate_pct
Contract	Month-to-Month	3610	1655	45.84
Contract	One Year	1550	166	10.71
Contract	Two Year	1883	48	2.55
Internet Type	Fiber Optic	3035	1236	40.72
Internet Type	Cable	830	213	25.66
Internet Type	DSL	1652	307	18.58
Internet Type	No Internet	1526	113	7.40
Payment Method	Mailed Check	385	142	36.88
Payment Method	Bank Withdrawal	3909	1329	34.00
Payment Method	Credit Card	2249	398	14.48

Figure 8. MySQL Workbench schema, tables, views, and verified segment results.

Conclusion

The project demonstrates an end-to-end SQL workflow: relational design, staged data loading, transparent data-quality decisions, reproducible analysis, advanced SQL, business interpretation, and communication. The results support focused retention testing around early tenure, Month-to-Month contracts, competitor pressure, service experience, and targeted outreach, while the documented limitations prevent descriptive findings from being overstated.

References

Maven Analytics. Telecom Customer Churn. <https://mavenanalytics.io/data-playground/telecom-customer-churn>

IBM Cognos Analytics. Telco customer churn sample.

<https://www.ibm.com/docs/en/cognos-analytics/12.1.x?topic=samples-telco-customer-churn>