

Table creation query

```
DROP TABLE
CREATE TABLE
postgres=# -- Create staging table
CREATE TABLE telco_churn_staging (
    customerID TEXT,
    gender TEXT,
    SeniorCitizen INT,
    Partner TEXT,
    Dependents TEXT,
    tenure INT,
    PhoneService TEXT,
    MultipleLines TEXT,
    InternetService TEXT,
    OnlineSecurity TEXT,
    OnlineBackup TEXT,
    DeviceProtection TEXT,
    TechSupport TEXT,
    StreamingTV TEXT,
    StreamingMovies TEXT,
    Contract TEXT,
    PaperlessBilling TEXT,
    PaymentMethod TEXT,
    MonthlyCharges NUMERIC,
    TotalCharges TEXT,
    Churn TEXT
);
```

```

postgres=# INSERT INTO telco_churn_clean
SELECT
    customerID,
    gender,
    CASE WHEN SeniorCitizen = 1 THEN TRUE ELSE FALSE END,
    CASE WHEN Partner = 'Yes' THEN TRUE ELSE FALSE END,
    CASE WHEN Dependents = 'Yes' THEN TRUE ELSE FALSE END,
    tenure,
    CASE WHEN PhoneService = 'Yes' THEN TRUE ELSE FALSE END,
    MultipleLines,
    InternetService,
    OnlineSecurity,
    OnlineBackup,
    DeviceProtection,
    TechSupport,
    StreamingTV,
    StreamingMovies,
    Contract,
    CASE WHEN PaperlessBilling = 'Yes' THEN TRUE ELSE FALSE END,
    PaymentMethod,
    MonthlyCharges,
    CASE WHEN TRIM(TotalCharges) = '' OR TotalCharges IS NULL THEN 0
        ELSE CAST(TRIM(TotalCharges) AS NUMERIC(10,2))
    END,
    CASE WHEN Churn = 'Yes' THEN TRUE ELSE FALSE END,
    CURRENT_TIMESTAMP
FROM telco_churn_staging
WHERE customerID IS NOT NULL;

```

## Output

feature	coef	std_err	z_stat	p_value	odds_ratio
Intercept	-1.9733015908810936	0.08280357036519138	-23.83117211228634	1.5896696244087652e-125	0.13899718596936406
tenure	-0.03572872301662112	0.002040638680513788	-17.508598341194535	1.2319016412802922e-68	0.9649020136897138
monthly_charges	0.02861745575326007	0.0013494229297107516	21.207180583032845	8.198184412447421e-100	1.02993808693325072
contract_month	0.37059218917506653	0.059359105683461226	6.34329110792758	4.282502674753827e-10	1.448592012322061
contract_one_year	-0.6918712523873665	0.07901673548284897	-8.756009067693526	2.0228545190853954e-18	0.5006383712576202
contract_two_year	-1.6520225276687843	0.12261160294513213	-13.473623115489758	2.2361798846139074e-41	0.1916618749035207
(6 rows)					

## Total quantity sold query

```

2025-10-25T00:43:56,646 INFO [718bd7c4-40ad-4c82-bd2b-72a97bcc7319 Main] org.apache.hadoop.hive.qe.session.SessionState
hive> SELECT
    >     StockCode,
    >     SUM(Quantity) AS total_quantity
    > FROM online_retail
    > GROUP BY StockCode
    > ORDER BY total_quantity DESC
    > LIMIT 20;

```

## Accuracy

```
FROM churn_predictions;
 true_positive | true_negative | false_positive | false_negative
-----+-----+-----+-----+
      917 |        4628 |       546 |       952
(1 row)

postgres=# SELECT
  ROUND(
    100.0 * SUM(CASE WHEN churn = predicted THEN 1 ELSE 0 END)::numeric / COUNT(*),
    2
  ) AS accuracy_percent
FROM churn_predictions;
 accuracy_percent
-----
      78.73
(1 row)

postgres=#
```