**H4 Code:**

/\* Step 1: Aggregate Sales Data by Customer \*/

proc sql;

create table customer\_data as

select Customer\_ID,

sum(Sales\_Amount) as Total\_Sales,

count(Transaction\_ID) as Number\_of\_Transactions

from work.IMPORT1 /\* Use the correct dataset \*/

group by Customer\_ID;

quit;

/\* Step 2: Define Retention Status \*/

data customer\_data\_retention;

set customer\_data;

if Number\_of\_Transactions > 1 then Retention\_Status = "Retained";

else Retention\_Status = "Non-Retained";

run;

/\* Step 3: Categorize Total Sales \*/

data customer\_data\_retention;

set customer\_data\_retention;

if Total\_Sales <= 100 then Sales\_Category = "Low";

else if Total\_Sales <= 1000 then Sales\_Category = "Medium";

else Sales\_Category = "High";

run;

/\* Step 4: Chi-Square Test for Sales Volume and Retention Status \*/

proc freq data=customer\_data\_retention;

tables Sales\_Category\*Retention\_Status / chisq;

title "Chi-Square Test for Sales Volume and Retention Status";

run;

/\* Step 5: Correlation Analysis \*/

proc corr data=customer\_data\_retention;

var Total\_Sales Number\_of\_Transactions;

title "Correlation Analysis Between Sales Volume and Customer Retention";

run;

/\* Step 6: Scatter Plot of Sales Volume and Number of Transactions \*/

proc sgplot data=customer\_data\_retention;

scatter x=Number\_of\_Transactions y=Total\_Sales / group=Retention\_Status;

xaxis label="Number of Transactions";

yaxis label="Total Sales (USD)";

title "Scatter Plot of Sales Volume and Number of Transactions";

run;