**H1 and H2 Code:**

/\* Step 1: Box Plot for Sales Distribution by Product Category \*/

proc sgplot data=work.retail\_cleaned;

vbox "Retail Sales"n / category="Item Type"n;

xaxis label="Product Category" fitpolicy=stagger; /\* Staggered labels for clarity \*/

yaxis label="Retail Sales (USD)" min=0 max=1000; /\* Adjust axis range to reduce compression \*/

title "Sales Distribution by Product Category (Adjusted)";

run;

/\* Step 2: Perform ANOVA for Product Categories and Sales \*/

proc anova data=work.retail\_cleaned PLOTS(MAXPOINTS=15000);

class "Item Type"n;

model "Retail Sales"n = "Item Type"n;

means "Item Type"n / tukey;

title "ANOVA for Product Category and Sales";

run;

/\* Step 3: Add Season Variable \*/

data work.retail\_seasonal;

set work.retail\_cleaned;

length Season $10; /\* Ensure Season variable has character length \*/

if MONTH in (3, 4, 5) then Season = "Spring";

else if MONTH in (6, 7, 8) then Season = "Summer";

else if MONTH in (9, 10, 11) then Season = "Fall";

else if MONTH in (12, 1, 2) then Season = "Winter";

run;

/\* Step 4: Summarize Sales by Season and Product Category \*/

proc sql;

create table work.seasonal\_sales as

select "Item Type"n as Category,

Season,

sum("Retail Sales"n) as Total\_Sales,

mean("Retail Sales"n) as Avg\_Sales

from work.retail\_seasonal

group by "Item Type"n, Season

order by Category, Season;

quit;

/\* Step 5: Perform Two-Way ANOVA \*/

proc anova data=work.retail\_seasonal;

class Season "Item Type"n;

model "Retail Sales"n = Season | "Item Type"n;

means Season "Item Type"n / tukey;

title "Two-Way ANOVA for Seasonal Trends and Product Categories";

run;

/\* Step 6: Bar Chart: Total Sales by Season and Product Category \*/

proc sgplot data=work.seasonal\_sales;

vbar Season / response=Total\_Sales group=Category groupdisplay=cluster datalabel;

xaxis label="Season";

yaxis label="Total Sales (USD)";

title "Total Sales by Season and Product Category";

run;

/\* Step 7: Interaction Plot: Seasonal Trends by Product Category \*/

proc sgplot data=work.seasonal\_sales;

series x=Season y=Avg\_Sales / group=Category markers lineattrs=(pattern=solid);

xaxis label="Season";

yaxis label="Average Sales (USD)";

title "Seasonal Trends by Product Category";

run;

/\* Step 10: Recode SKU and Sales\_Category to Numeric \*/

data work.sales\_categorized\_numeric;

set work.import;

/\* Ensure SKU is numeric \*/

SKU\_Code = input(SKU, best12.); /\* If SKU cannot be converted, assign numeric codes manually \*/

/\* Recode Sales\_Category \*/

if Sales\_Category = "Low" then Sales\_Code = 1;

else if Sales\_Category = "Medium" then Sales\_Code = 2;

else if Sales\_Category = "High" then Sales\_Code = 3;

run;

/\* Step 11: Verify Recoded Dataset \*/

proc contents data=work.sales\_categorized\_numeric;

title "Verify Recoded Dataset";

run;

/\* Step 13: Prepare Data for Time-Series Plot \*/

proc sql;

create table work.sales\_trends as

select YEAR, MONTH,

sum("Retail Sales"n) as Total\_Sales

from work.import

group by YEAR, MONTH

order by YEAR, MONTH;

quit;

/\* Step 14: Line Chart: Monthly Sales Trends \*/

proc sgplot data=work.sales\_trends;

series x=MONTH y=Total\_Sales / markers lineattrs=(color=blue pattern=solid thickness=2);

xaxis label="Month";

yaxis label="Total Sales (USD)";

title "Monthly Retail Sales Trends";

run;