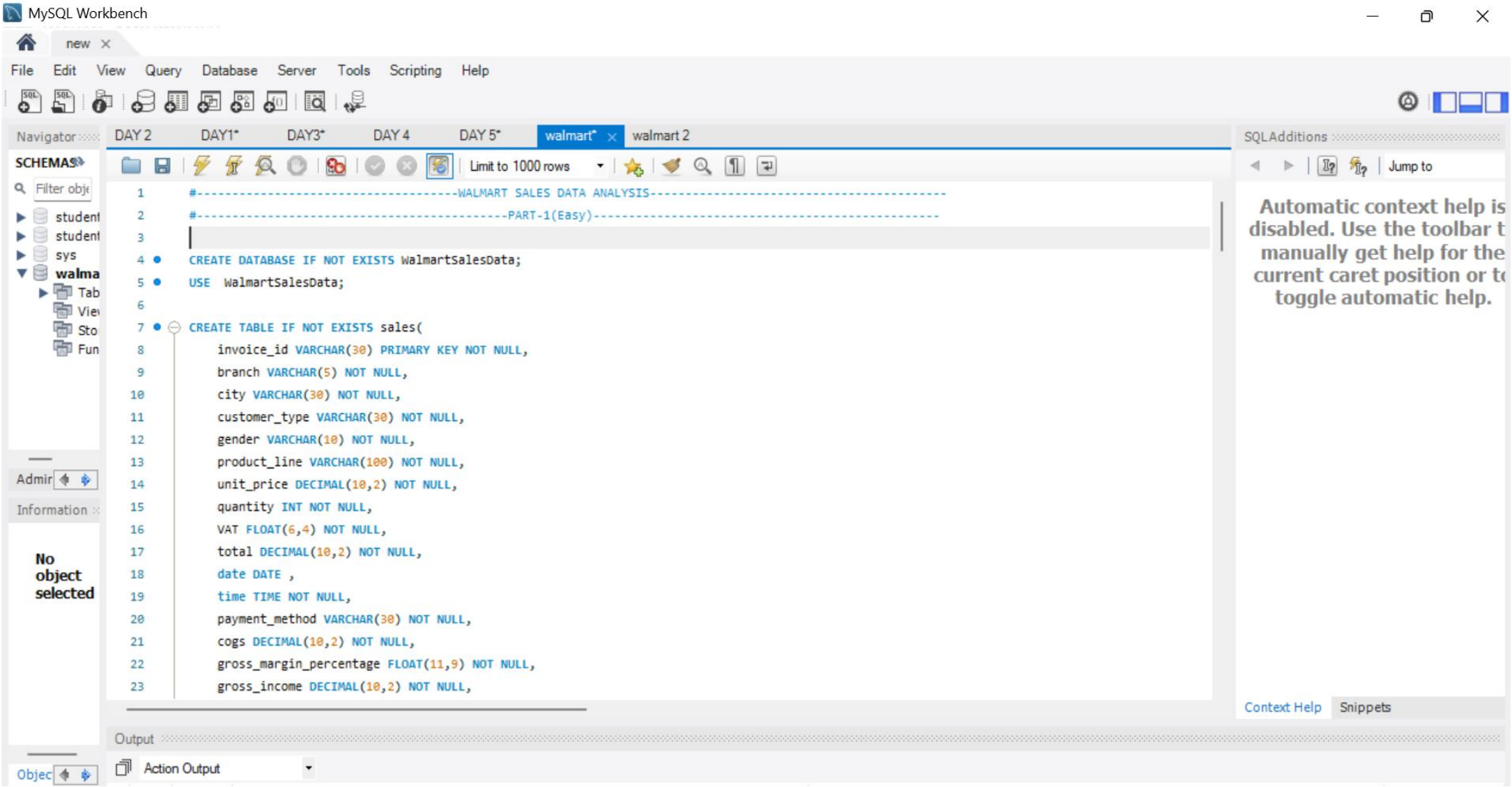
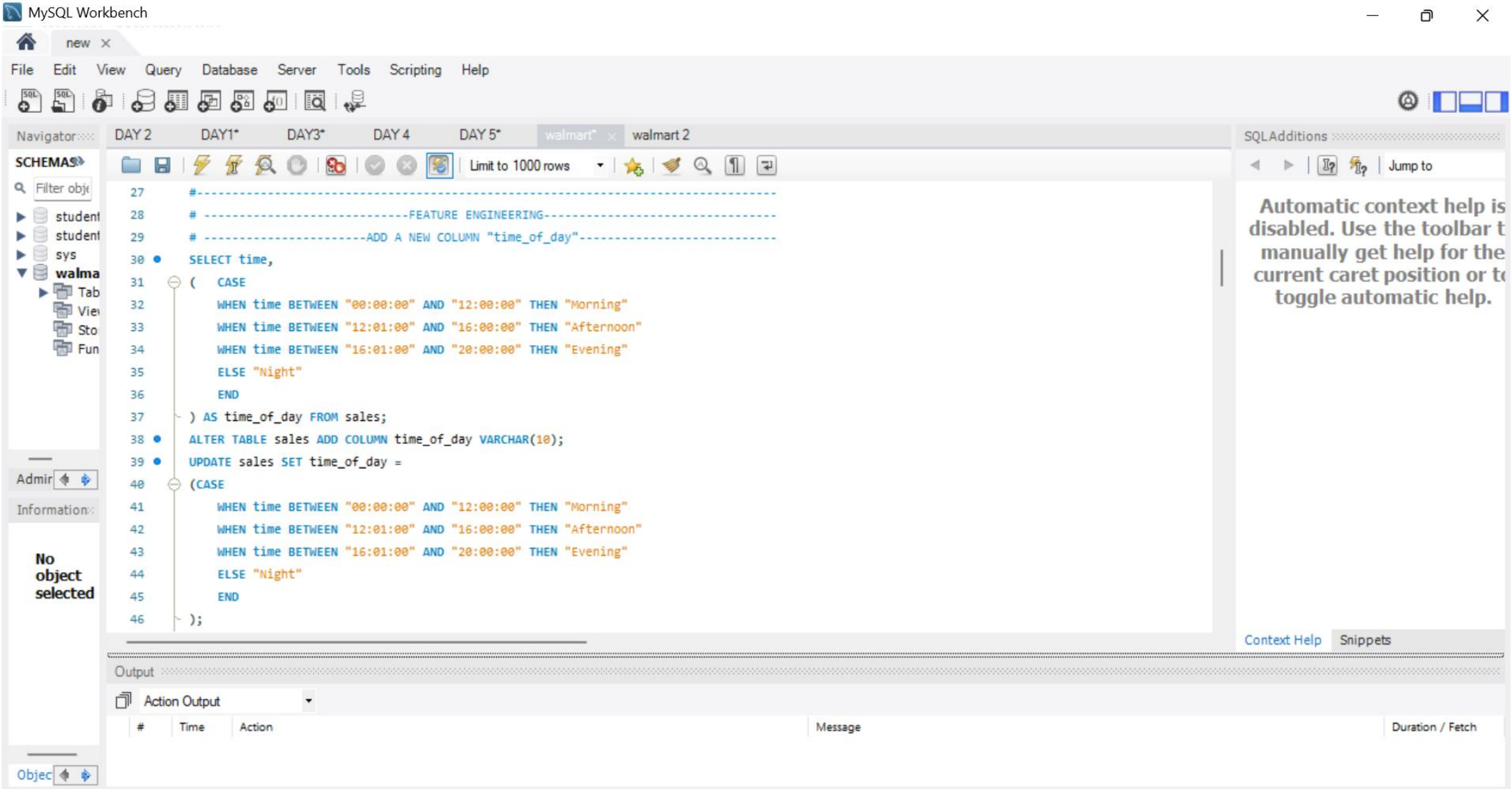
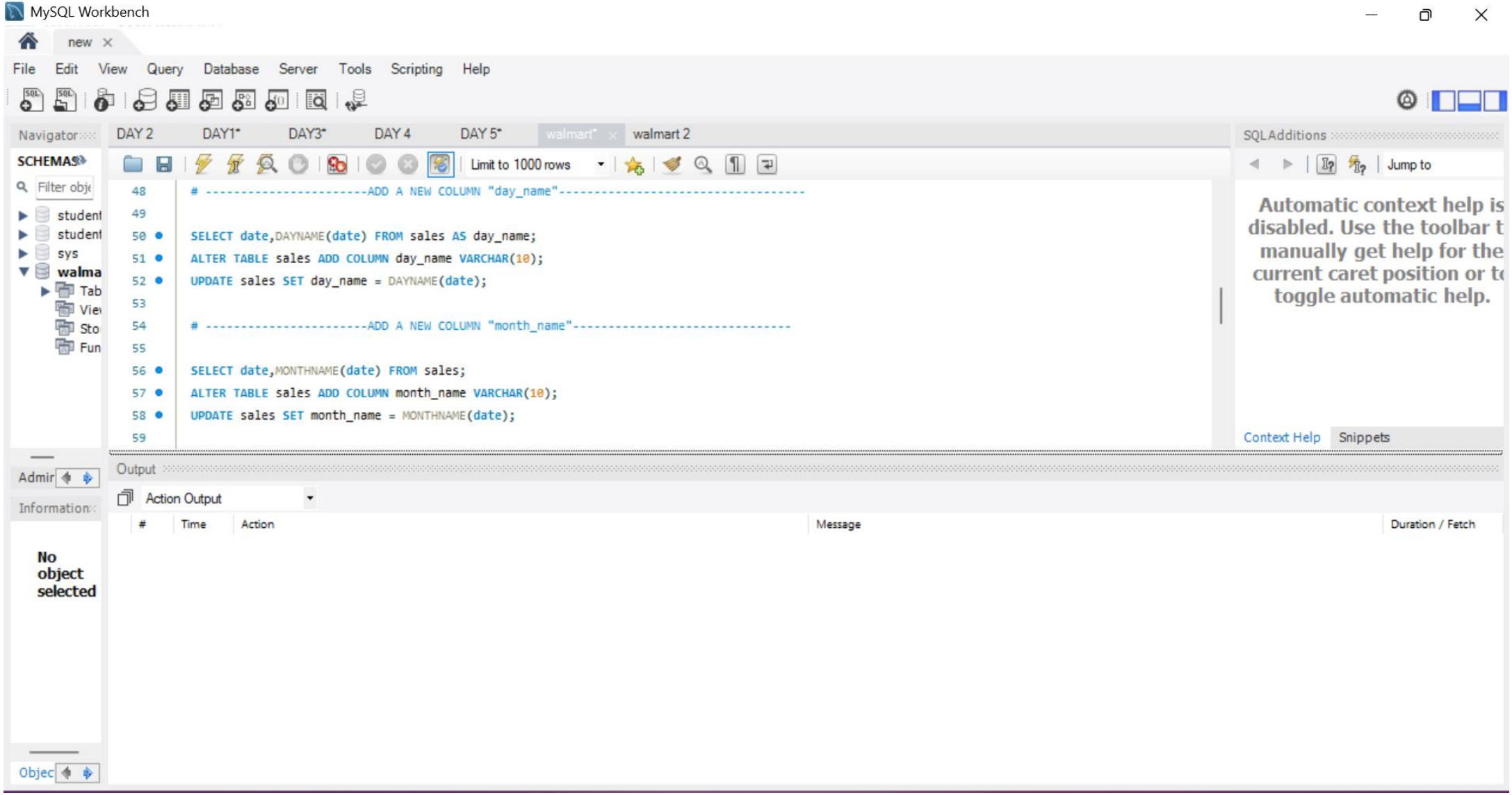
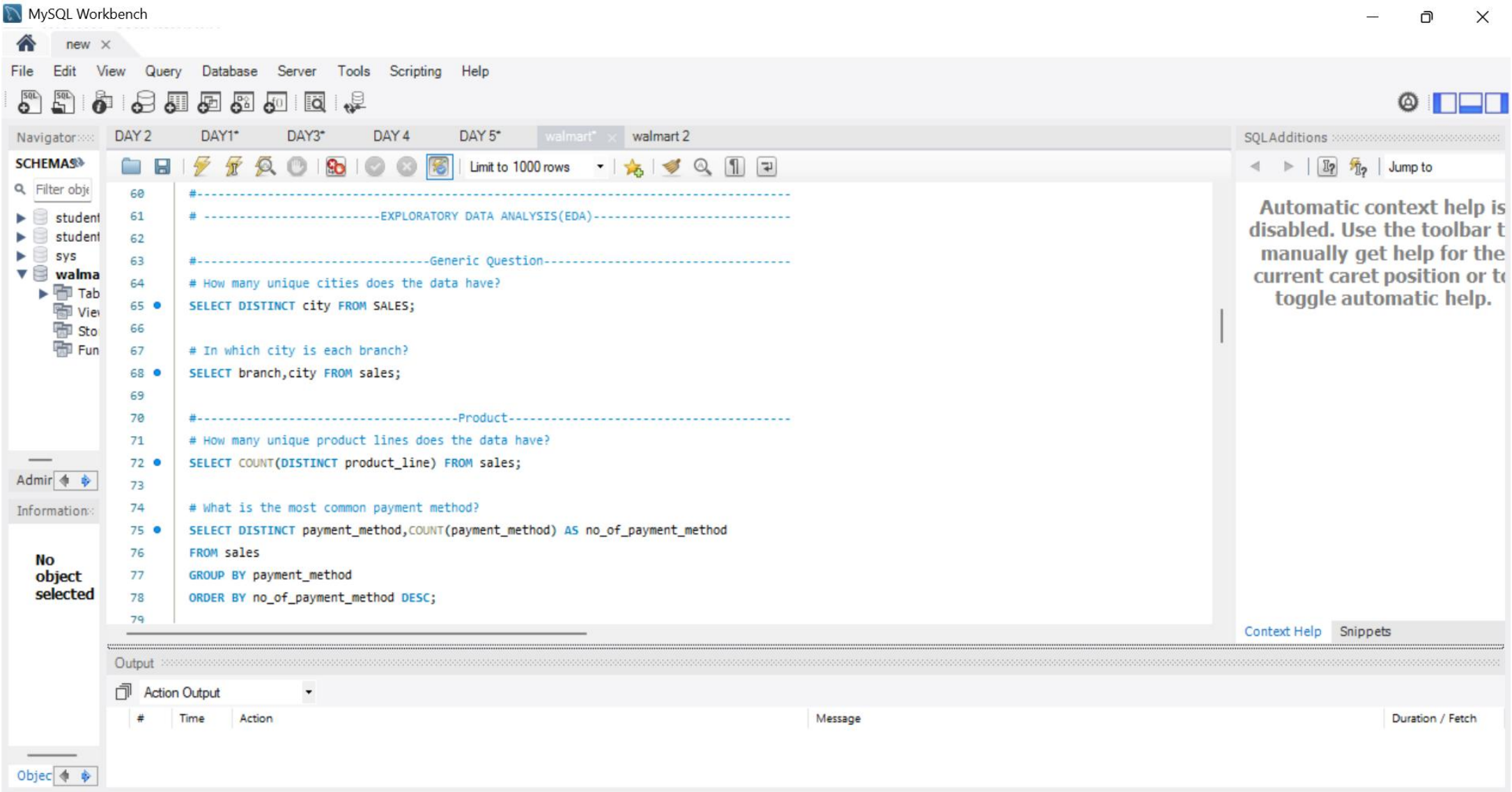
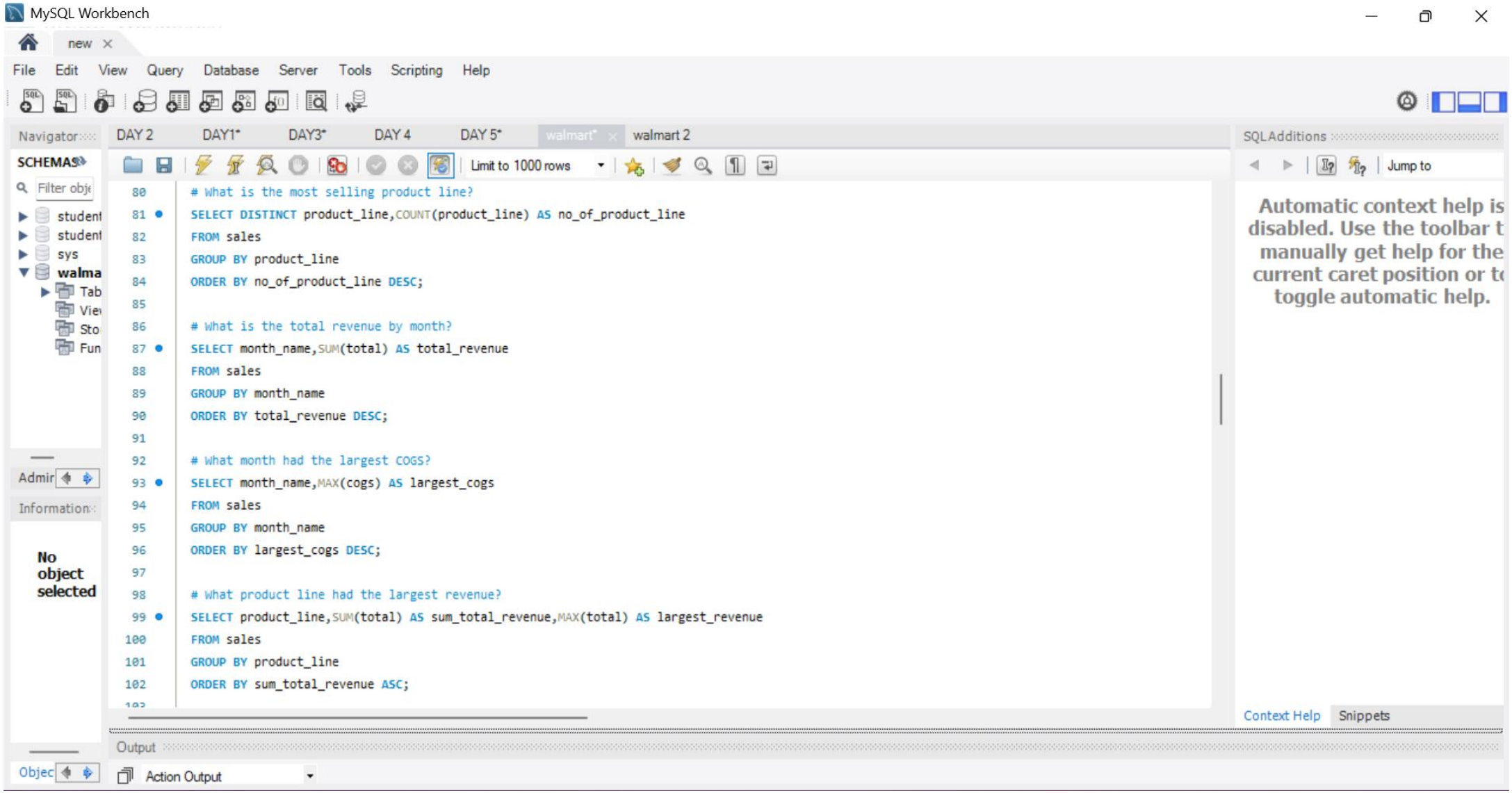
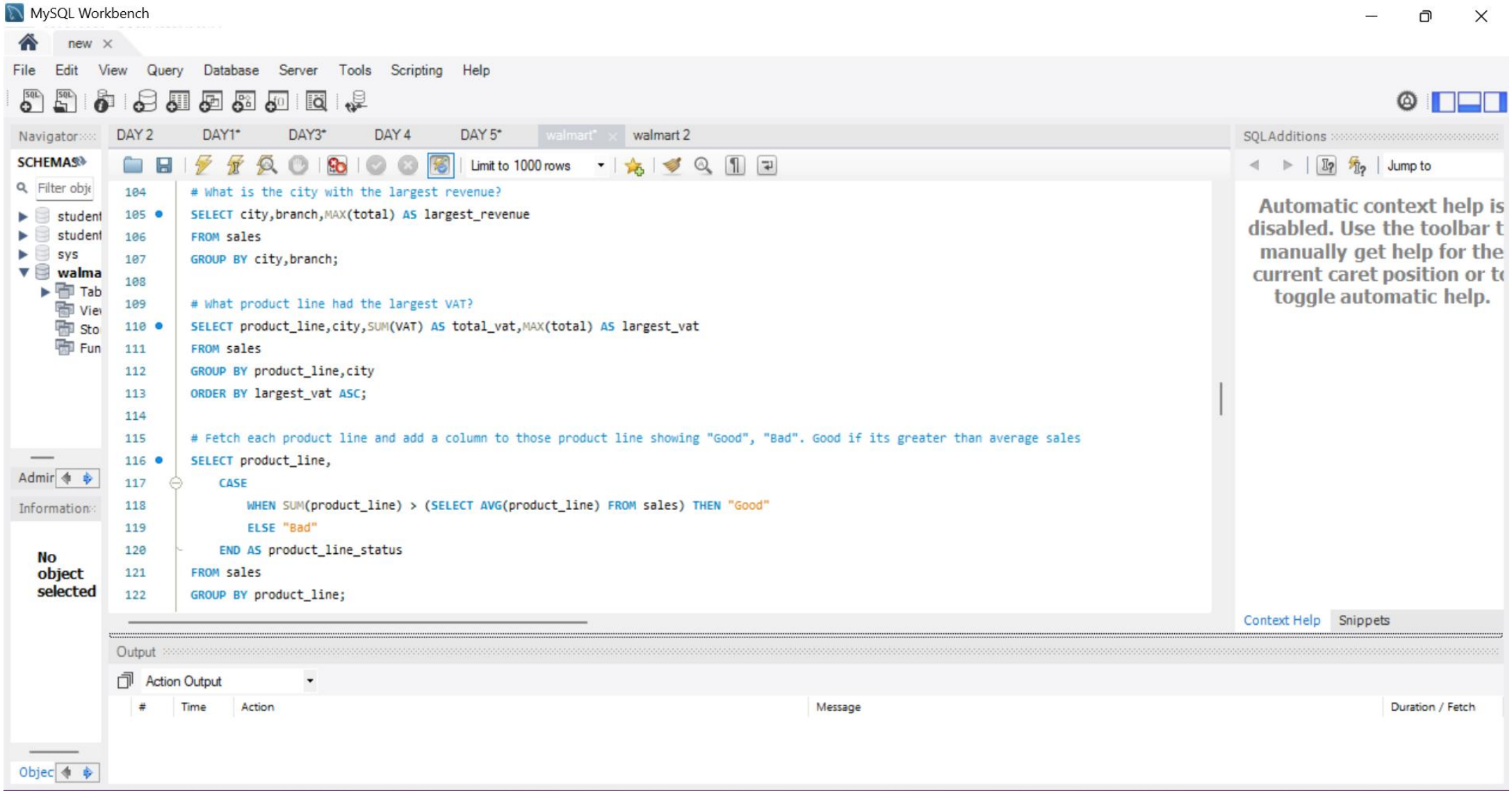

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
#----------WALMART SALES DATA ANALYSIS----------
#----------PART-1(Easy)----------
CREATE DATABASE IF NOT EXISTS WalmartSalesData;
USE WalmartSalesData;
CREATE TABLE IF NOT EXISTS sales(
invoice_id VARCHAR(30) PRIMARY KEY NOT NULL,
branch VARCHAR(5) NOT NULL,
city VARCHAR(30) NOT NULL,
customer_type VARCHAR(30) NOT NULL,
gender VARCHAR(10) NOT NULL,
product_line VARCHAR(100) NOT NULL,
unit_price DECIMAL(10,2) NOT NULL,
quantity INT NOT NULL,
VAT FLOAT(6,4) NOT NULL,
total DECIMAL(10,2) NOT NULL,
date DATE ,
time TIME NOT NULL,
payment_method VARCHAR(30) NOT NULL,
cogs DECIMAL(10,2) NOT NULL,
gross_margin_percentage FLOAT(11,9) NOT NULL,
gross_income DECIMAL(10,2) NOT NULL,
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output

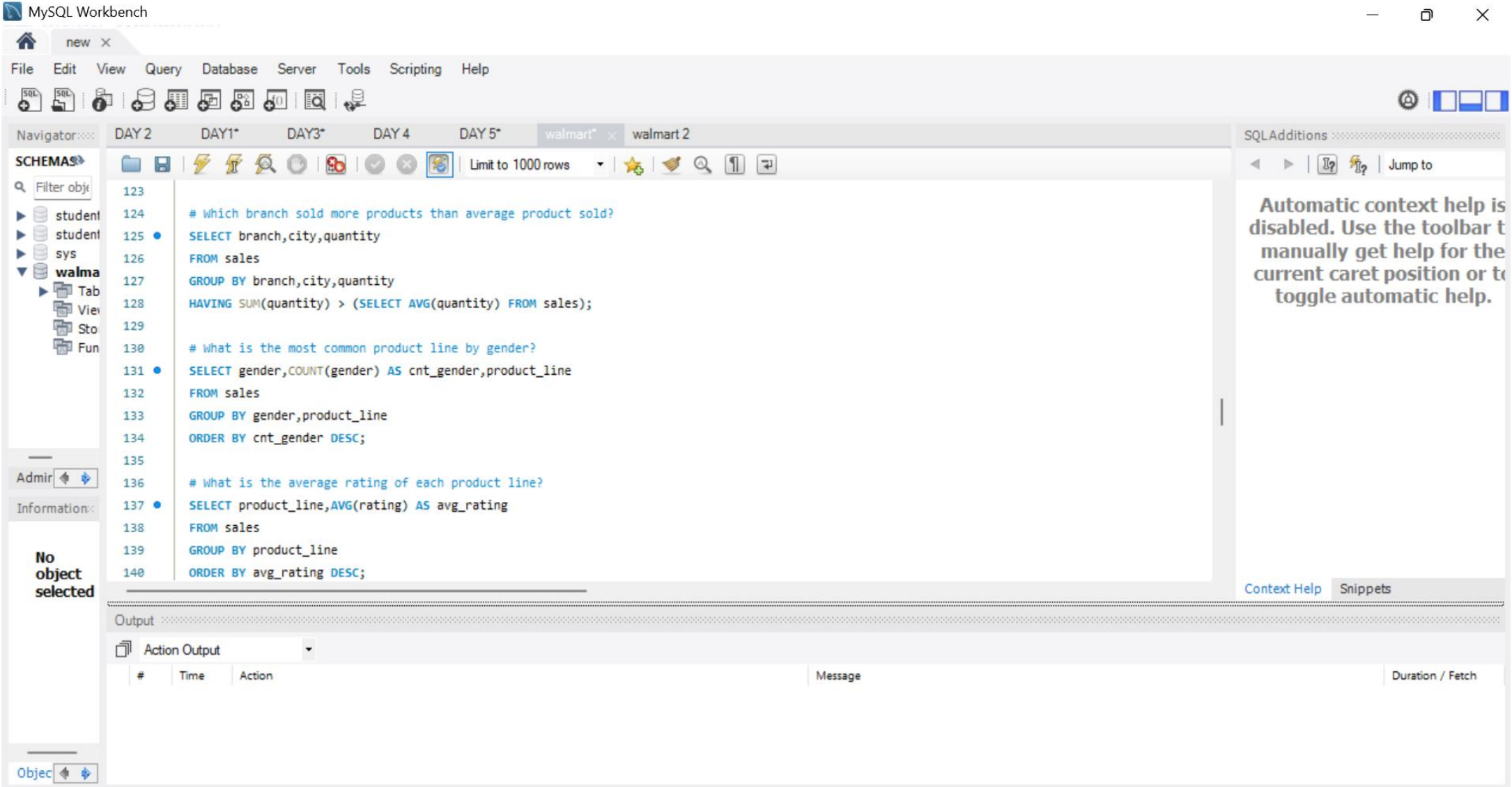

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
#----------
----------FEATURE ENGINEERING----------
----------ADD A NEW COLUMN "time_of_day"----------
SELECT time,
(CASE
WHEN time BETWEEN "00:00:00" AND "12:00:00" THEN "Morning"
WHEN time BETWEEN "12:01:00" AND "16:00:00" THEN "Afternoon"
WHEN time BETWEEN "16:01:00" AND "20:00:00" THEN "Evening"
ELSE "Night"
END
) AS time_of_day FROM sales;
ALTER TABLE sales ADD COLUMN time_of_day VARCHAR(10);
UPDATE sales SET time_of_day =
(CASE
WHEN time BETWEEN "00:00:00" AND "12:00:00" THEN "Morning"
WHEN time BETWEEN "12:01:00" AND "16:00:00" THEN "Afternoon"
WHEN time BETWEEN "16:01:00" AND "20:00:00" THEN "Evening"
ELSE "Night"
END
);
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch

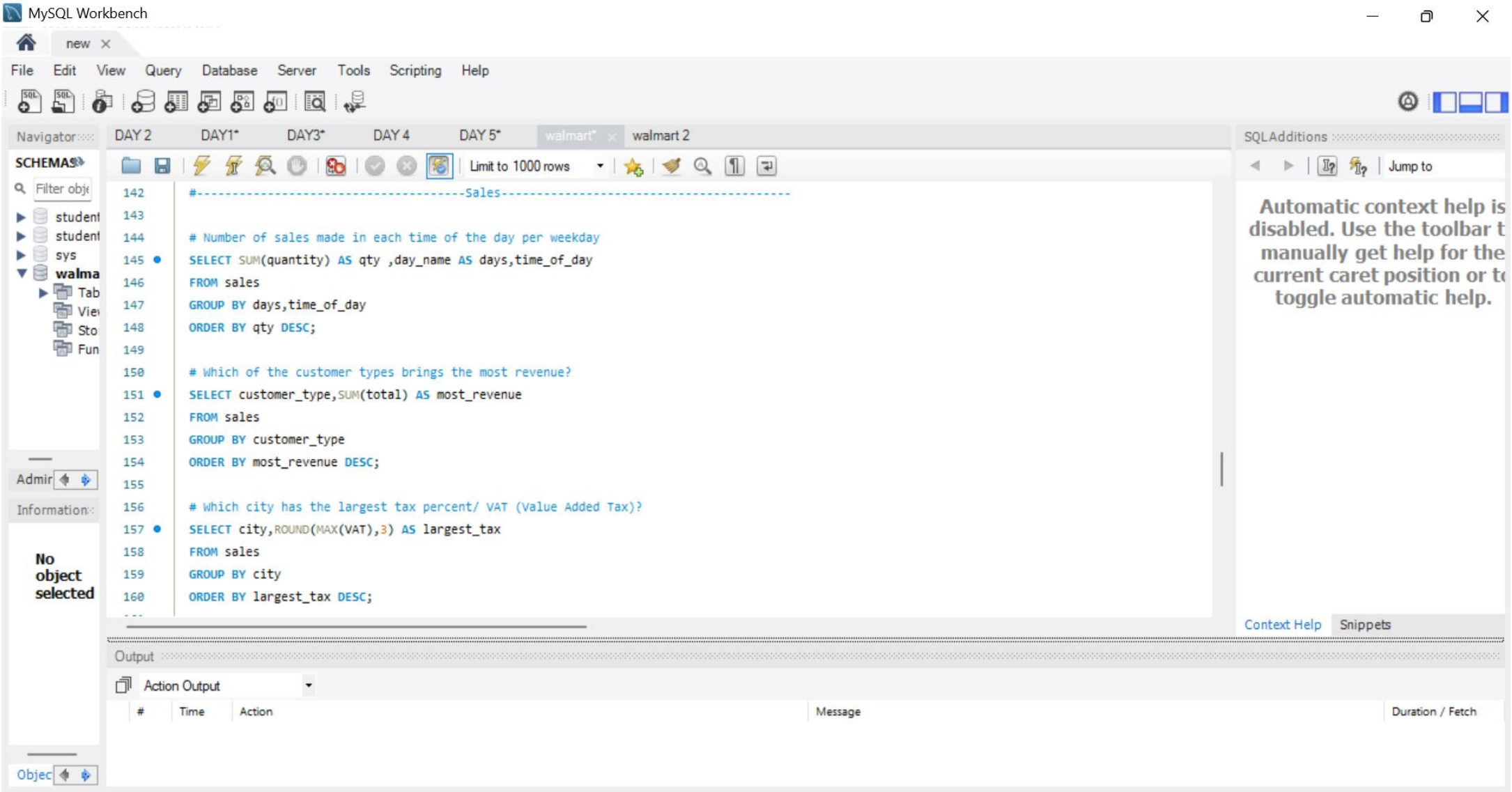

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
48 # ----------ADD A NEW COLUMN "day_name"----------
49
50 SELECT date,DAYNAME(date) FROM sales AS day_name;
51 ALTER TABLE sales ADD COLUMN day_name VARCHAR(10);
52 UPDATE sales SET day_name = DAYNAME(date);
53
54 # ----------ADD A NEW COLUMN "month_name"----------
55
56 SELECT date,MONTHNAME(date) FROM sales;
57 ALTER TABLE sales ADD COLUMN month_name VARCHAR(10);
58 UPDATE sales SET month_name = MONTHNAME(date);
59
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch

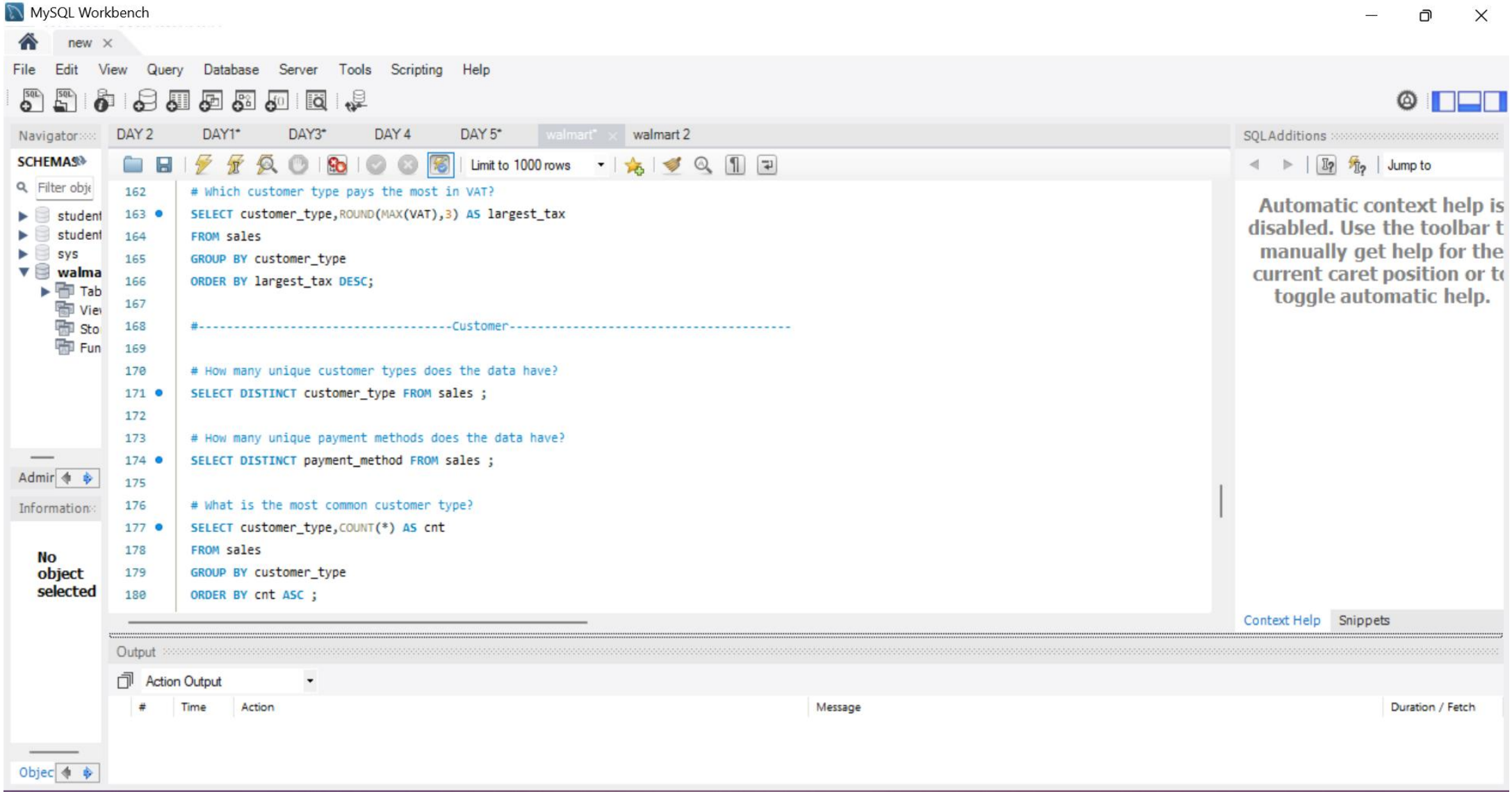

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
#----------EXPLORATORY DATA ANALYSIS(EDA)----------
#----------Generic Question----------
How many unique cities does the data have?
SELECT DISTINCT city FROM SALES;
In which city is each branch?
SELECT branch,city FROM sales;
#----------Product----------
How many unique product lines does the data have?
SELECT COUNT(DISTINCT product_line) FROM sales;
What is the most common payment method?
SELECT DISTINCT payment_method,COUNT(payment_method) AS no_of_payment_method
FROM sales
GROUP BY payment_method
ORDER BY no_of_payment_method DESC;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch

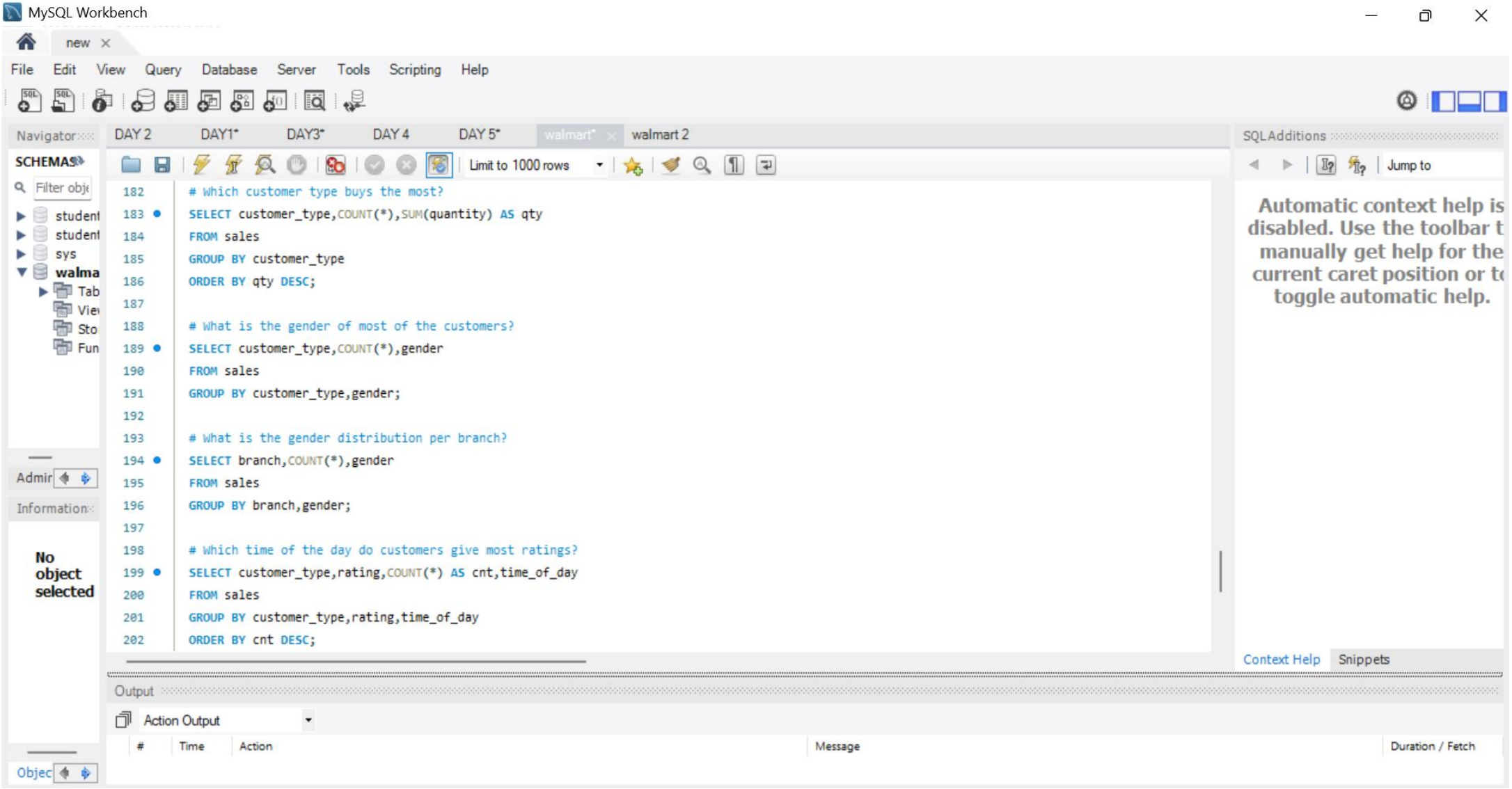

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
80 # What is the most selling product line?
81 SELECT DISTINCT product_line,COUNT(product_line) AS no_of_product_line
82 FROM sales
83 GROUP BY product_line
84 ORDER BY no_of_product_line DESC;
85
86 # What is the total revenue by month?
87 SELECT month_name,SUM(total) AS total_revenue
88 FROM sales
89 GROUP BY month_name
90 ORDER BY total_revenue DESC;
91
92 # What month had the largest COGS?
93 SELECT month_name,MAX(cogs) AS largest_cogs
94 FROM sales
95 GROUP BY month_name
96 ORDER BY largest_cogs DESC;
97
98 # What product line had the largest revenue?
99 SELECT product_line,SUM(total) AS sum_total_revenue,MAX(total) AS largest_revenue
100 FROM sales
101 GROUP BY product_line
102 ORDER BY sum_total_revenue ASC;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output

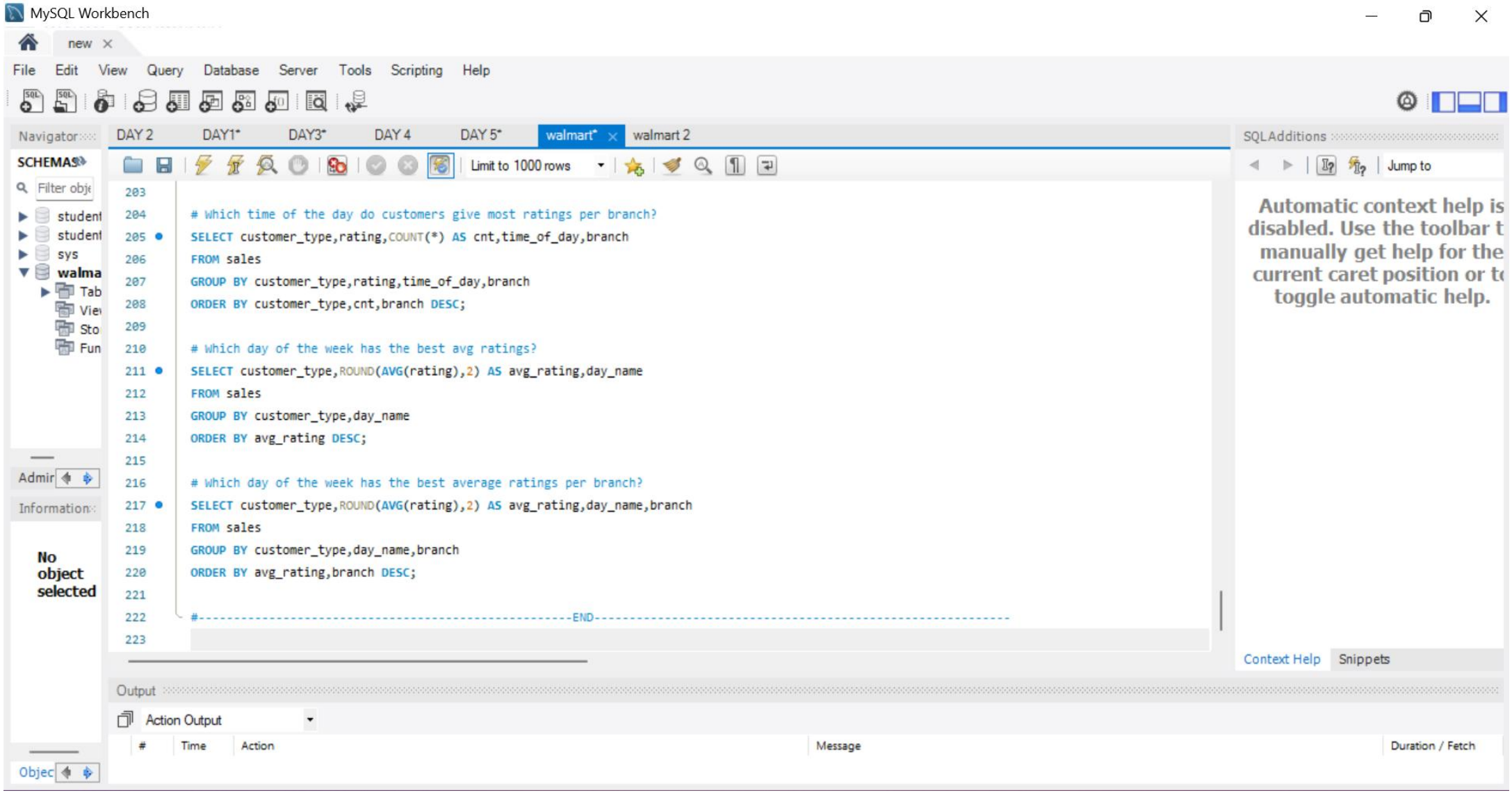

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
104 # What is the city with the largest revenue?
105 SELECT city,branch,MAX(total) AS largest_revenue
106 FROM sales
107 GROUP BY city,branch;
108
109 # What product line had the largest VAT?
110 SELECT product_line,city,SUM(VAT) AS total_vat,MAX(total) AS largest_vat
111 FROM sales
112 GROUP BY product_line,city
113 ORDER BY largest_vat ASC;
114
115 # Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average sales
116 SELECT product_line,
117 CASE
118 WHEN SUM(product_line) > (SELECT AVG(product_line) FROM sales) THEN "Good"
119 ELSE "Bad"
120 END AS product_line_status
121 FROM sales
122 GROUP BY product_line;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch

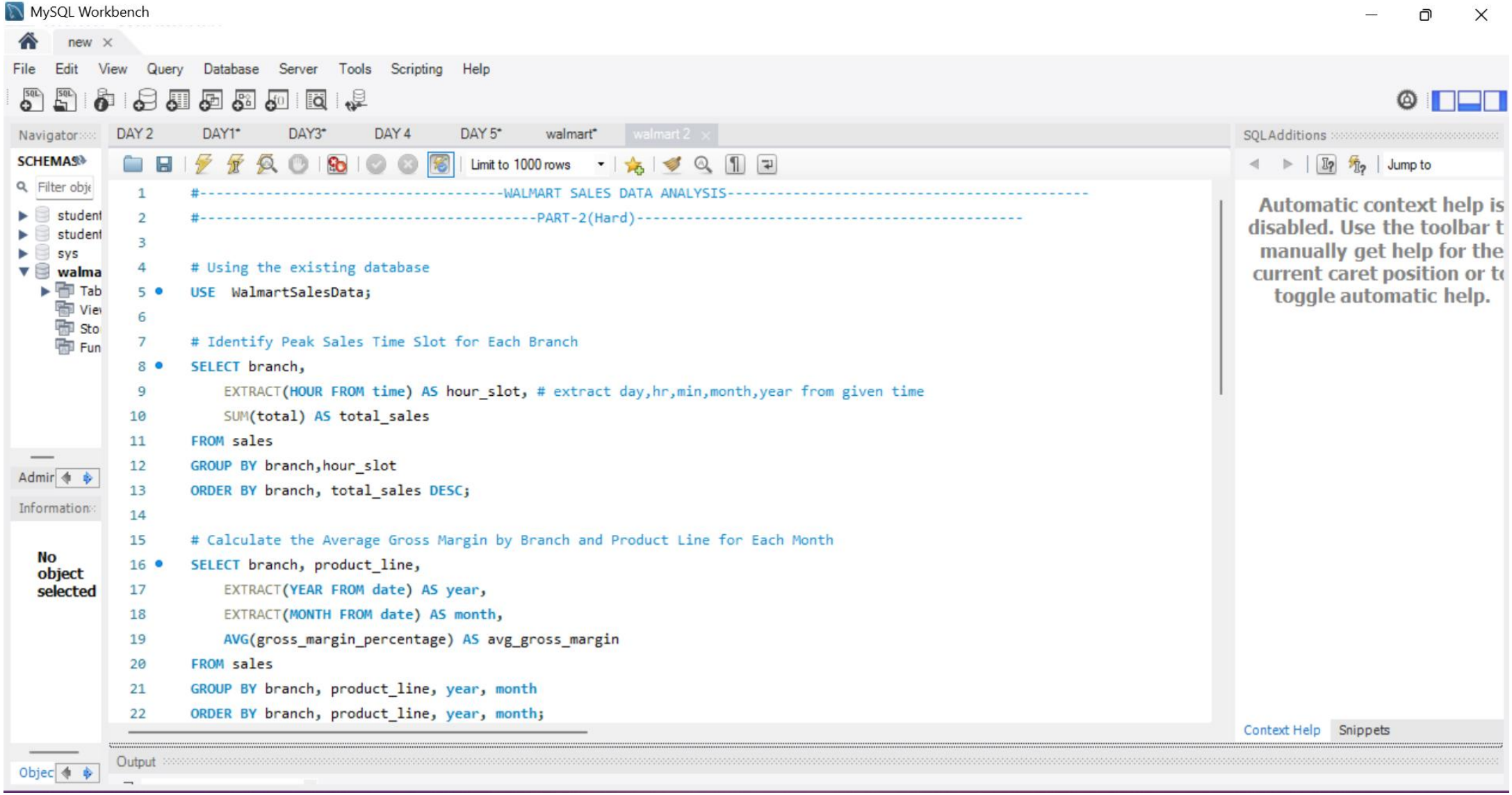

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
123
124 # Which branch sold more products than average product sold?
125 SELECT branch,city,quantity
126 FROM sales
127 GROUP BY branch,city,quantity
128 HAVING SUM(quantity) > (SELECT AVG(quantity) FROM sales);
129
130 # What is the most common product line by gender?
131 SELECT gender,COUNT(gender) AS cnt_gender,product_line
132 FROM sales
133 GROUP BY gender,product_line
134 ORDER BY cnt_gender DESC;
135
136 # What is the average rating of each product line?
137 SELECT product_line,AVG(rating) AS avg_rating
138 FROM sales
139 GROUP BY product_line
140 ORDER BY avg_rating DESC;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch

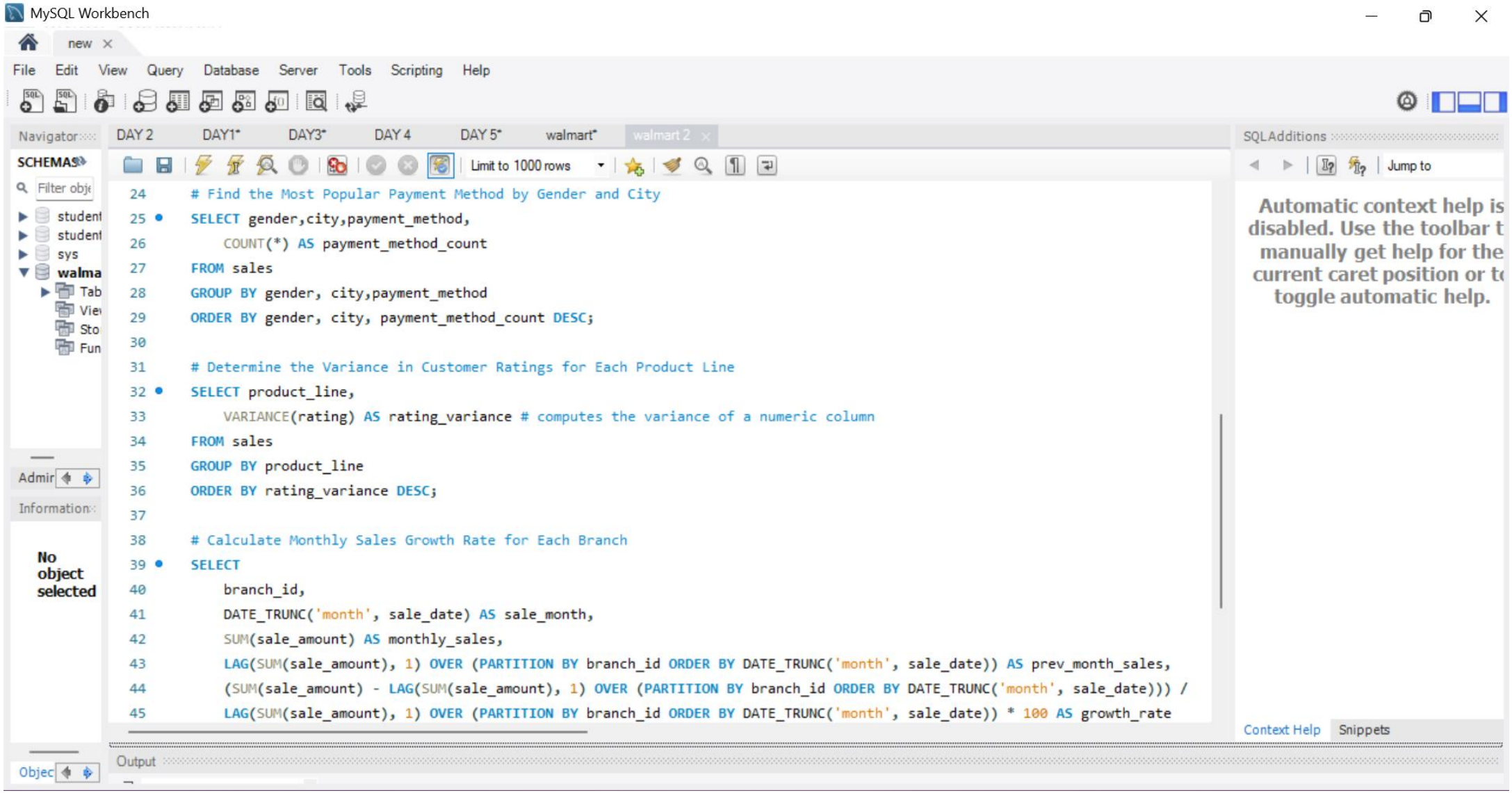

MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
142 #----------Sales----------
143
144 # Number of sales made in each time of the day per weekday
145 SELECT SUM(quantity) AS qty ,day_name AS days,time_of_day
146 FROM sales
147 GROUP BY days,time_of_day
148 ORDER BY qty DESC;
149
150 # Which of the customer types brings the most revenue?
151 SELECT customer_type,SUM(total) AS most_revenue
152 FROM sales
153 GROUP BY customer_type
154 ORDER BY most_revenue DESC;
155
156 # Which city has the largest tax percent/ VAT (Value Added Tax)?
157 SELECT city,ROUND(MAX(VAT),3) AS largest_tax
158 FROM sales
159 GROUP BY city
160 ORDER BY largest_tax DESC;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch


MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
162 # Which customer type pays the most in VAT?
163 SELECT customer_type,ROUND(MAX(VAT),3) AS largest_tax
164 FROM sales
165 GROUP BY customer_type
166 ORDER BY largest_tax DESC;
167
168 #----------Customer----------
169
170 # How many unique customer types does the data have?
171 SELECT DISTINCT customer_type FROM sales ;
172
173 # How many unique payment methods does the data have?
174 SELECT DISTINCT payment_method FROM sales ;
175
176 # What is the most common customer type?
177 SELECT customer_type,COUNT(*) AS cnt
178 FROM sales
179 GROUP BY customer_type
180 ORDER BY cnt ASC ;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch


MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
182 # Which customer type buys the most?
183 SELECT customer_type,COUNT(*),SUM(quantity) AS qty
184 FROM sales
185 GROUP BY customer_type
186 ORDER BY qty DESC;
187
188 # What is the gender of most of the customers?
189 SELECT customer_type,COUNT(*),gender
190 FROM sales
191 GROUP BY customer_type,gender;
192
193 # What is the gender distribution per branch?
194 SELECT branch,COUNT(*),gender
195 FROM sales
196 GROUP BY branch,gender;
197
198 # Which time of the day do customers give most ratings?
199 SELECT customer_type,rating,COUNT(*) AS cnt,time_of_day
200 FROM sales
201 GROUP BY customer_type,rating,time_of_day
202 ORDER BY cnt DESC;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch


MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
203
204 # Which time of the day do customers give most ratings per branch?
205 SELECT customer_type,rating,COUNT(*) AS cnt,time_of_day,branch
206 FROM sales
207 GROUP BY customer_type,rating,time_of_day,branch
208 ORDER BY customer_type,cnt,branch DESC;
209
210 # Which day of the week has the best avg ratings?
211 SELECT customer_type,ROUND(AVG(rating),2) AS avg_rating,day_name
212 FROM sales
213 GROUP BY customer_type,day_name
214 ORDER BY avg_rating DESC;
215
216 # Which day of the week has the best average ratings per branch?
217 SELECT customer_type,ROUND(AVG(rating),2) AS avg_rating,day_name,branch
218 FROM sales
219 GROUP BY customer_type,day_name,branch
220 ORDER BY avg_rating,branch DESC;
221
222 #----------END----------
223
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output
Action Output
Time Action Message Duration / Fetch


MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2
DAY1*
DAY3*
DAY 4
DAY 5*
walmart*
walmart 2
Limit to 1000 rows
#----------WALMART SALES DATA ANALYSIS----------
#----------PART-2(Hard)----------
Using the existing database
USE WalmartSalesData;
Identify Peak Sales Time Slot for Each Branch
SELECT branch,
EXTRACT(HOUR FROM time) AS hour_slot, # extract day,hr,min,month,year from given time
SUM(total) AS total_sales
FROM sales
GROUP BY branch,hour_slot
ORDER BY branch, total_sales DESC;
Calculate the Average Gross Margin by Branch and Product Line for Each Month
SELECT branch, product_line,
EXTRACT(YEAR FROM date) AS year,
EXTRACT(MONTH FROM date) AS month,
AVG(gross_margin_percentage) AS avg_gross_margin
FROM sales
GROUP BY branch, product_line, year, month
ORDER BY branch, product_line, year, month;
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output


MySQL Workbench
new
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter obje
student
student
sys
walma
Tab
Vie
Sto
Fun
Admir
Information
No object selected
Objec
DAY 2 DAY1* DAY3* DAY 4 DAY 5* walmart* walmart 2
Limit to 1000 rows
Find the Most Popular Payment Method by Gender and City
SELECT gender,city,payment_method,
COUNT(*) AS payment_method_count
FROM sales
GROUP BY gender, city,payment_method
ORDER BY gender, city, payment_method_count DESC;
Determine the Variance in Customer Ratings for Each Product Line
SELECT product_line,
VARIANCE(rating) AS rating_variance # computes the variance of a numeric column
FROM sales
GROUP BY product_line
ORDER BY rating_variance DESC;
Calculate Monthly Sales Growth Rate for Each Branch
SELECT
branch_id,
DATE_TRUNC('month', sale_date) AS sale_month,
SUM(sale_amount) AS monthly_sales,
LAG(SUM(sale_amount), 1) OVER (PARTITION BY branch_id ORDER BY DATE_TRUNC('month', sale_date)) AS prev_month_sales,
(SUM(sale_amount) - LAG(SUM(sale_amount), 1) OVER (PARTITION BY branch_id ORDER BY DATE_TRUNC('month', sale_date))) /
LAG(SUM(sale_amount), 1) OVER (PARTITION BY branch_id ORDER BY DATE_TRUNC('month', sale_date)) * 100 AS growth_rate
SQLAdditions
Jump to
Automatic context help is disabled. Use the toolbar t manually get help for the current caret position or t toggle automatic help.
Context Help
Snippets
Output