

Data Management

[Home](#)[Exercises](#)[Skill builders](#)[Slide exercises](#)[Support](#)[Instructor](#)

Additional SQL queries

[ClassicModels](#) is a fictitious company. Use the [ClassicModels database](#) to answer the following requests ([SQL to create the database](#)). Your instructor has the answers to all queries.

The latitude and longitude are specified for office and customer locations in the Offices and Customers tables, respectively. The SRID is set to 4326 to indicate the Spatial Reference System (SRS) is geographic (see Chapter 11 for more details on SRID and SRS).

You can access ClassicModels at [richardtwatson.com](https://www.richardtwatson.com) with accountid=db1 and password=student.

Single entity

1. Prepare a list of offices sorted by country, state, city.
2. How many employees are there in the company?
3. What is the total of payments received?
4. List the product lines that contain 'Cars'.
5. Report total payments for October 28, 2004.
6. Report those payments greater than \$100,000.
7. List the products in each product line.
8. How many products in each product line?
9. What is the minimum payment received?
10. List all payments greater than twice the average payment.
11. What is the average percentage markup of the MSRP on buyPrice?
12. How many distinct products does ClassicModels sell?
13. Report the name and city of customers who don't have sales representatives?
14. What are the names of executives with VP or Manager in their title? Use the CONCAT function to combine the employee's first name and last name into a single field for reporting.
15. Which orders have a value greater than \$5,000?

One to many relationship

1. Report the account representative for each customer.
2. Report total payments for Atelier graphique.
3. Report the total payments by date
4. Report the products that have not been sold.
5. List the amount paid by each customer.
6. How many orders have been placed by Herkku Gifts?
7. Who are the employees in Boston?
8. Report those payments greater than \$100,000. Sort the report so the customer who made the highest payment appears first.
9. List the value of 'On Hold' orders.
10. Report the number of orders 'On Hold' for each customer.

Many to many relationship

1. List products sold by order date.
2. List the order dates in descending order for orders for the 1940 Ford Pickup Truck.
3. List the names of customers and their corresponding order number where a particular order from that customer has a value greater than \$25,000?
4. Are there any products that appear on all orders?
5. List the names of products sold at less than 80% of the MSRP.
6. Reports those products that have been sold with a markup of 100% or more (i.e., the priceEach is at least twice the buyPrice)
7. List the products ordered on a Monday.
8. What is the quantity on hand for products listed on 'On Hold' orders?

Regular expressions

1. Find products containing the name 'Ford'.
2. List products ending in 'ship'.
3. Report the number of customers in Denmark, Norway, and Sweden.
4. What are the products with a product code in the range S700_1000 to S700_1499?
5. Which customers have a digit in their name?
6. List the names of employees called Dianne or Diane.
7. List the products containing ship or boat in their product name.
8. List the products with a product code beginning with S700.
9. List the names of employees called Larry or Barry.
10. List the names of employees with non-alphabetic characters in their names.
11. List the vendors whose name ends in Diecast

General queries

1. Who is at the top of the organization (i.e., reports to no one).
2. Who reports to William Patterson?
3. List all the products purchased by Herkku Gifts.
4. Compute the commission for each sales representative, assuming the commission is 5% of the value of an order. Sort by employee last name and first name.
5. What is the difference in days between the most recent and oldest order date in the Orders file?
6. Compute the average time between order date and ship date for each customer ordered by the largest difference.
7. What is the value of orders shipped in August 2004? ([Hint](#)).
8. Compute the total value ordered, total amount paid, and their difference for each customer for orders placed in 2004 and payments received in 2004 (Hint; Create views for the total paid and total ordered).
9. List the employees who report to those employees who report to Diane Murphy. Use the CONCAT function to combine the employee's first name and last name into a single field for reporting.
10. What is the percentage value of each product in inventory sorted by the highest percentage first (Hint: Create a view first).
11. Write a function to convert miles per gallon to liters per 100 kilometers.
12. Write a procedure to increase the price of a specified product category by a given percentage. You will need to create a product table with appropriate data to test your procedure. Alternatively, load the ClassicModels database on your personal machine so you have complete access. You have to change the [DELIMITER](#) prior to creating the procedure.
13. What is the value of orders shipped in August 2004? ([Hint](#)).
14. What is the ratio the value of payments made to orders received for each month of 2004. (i.e., divide the value of payments made by the orders received)?
15. What is the difference in the amount received for each month of 2004 compared to 2003?

16. Write a procedure to report the amount ordered in a specific month and year for customers containing a specified character string in their name.
17. Write a procedure to change the credit limit of all customers in a specified country by a specified percentage.
18. **Basket of goods analysis:** A common retail analytics task is to analyze each basket or order to learn what products are often purchased together. Report the names of products that appear in the same order ten or more times.
19. **ABC reporting:** Compute the revenue generated by each customer based on their orders. Also, show each customer's revenue as a percentage of total revenue. Sort by customer name.
20. Compute the profit generated by each customer based on their orders. Also, show each customer's profit as a percentage of total profit. Sort by profit descending.
21. Compute the revenue generated by each sales representative based on the orders from the customers they serve.
22. Compute the profit generated by each sales representative based on the orders from the customers they serve. Sort by profit generated descending.
23. Compute the revenue generated by each product, sorted by product name.
24. Compute the profit generated by each product line, sorted by profit descending.
25. **Same as Last Year (SALY) analysis:** Compute the ratio for each product of sales for 2003 versus 2004.
26. Compute the ratio of payments for each customer for 2003 versus 2004.
27. Find the products sold in 2003 but not 2004.
28. Find the customers without payments in 2003.

Correlated subqueries

1. Who reports to Mary Patterson?
2. Which payments in any month and year are more than twice the average for that month and year (i.e. compare all payments in Oct 2004 with the average payment for Oct 2004)? Order the results by the date of the payment. You will need to use the [date functions](#).
3. Report for each product, the percentage value of its stock on hand as a percentage of the stock on hand for product line to which it belongs. Order the report by product line and percentage value within product line descending. Show percentages with two decimal places.
4. For orders containing more than two products, report those products that constitute more than 50% of the value of the order.

Spatial data

The Offices and Customers tables contain the latitude and longitude of each office and customer in officeLocation and customerLocation, respectively, in POINT format. Conventionally, latitude and longitude are reported as a pair of points, with latitude first.

1. Which customers are in the Southern Hemisphere?
2. Which US customers are south west of the New York office?
3. Which customers are closest to the Tokyo office (i.e., closer to Tokyo than any other office)?
4. Which French customer is furthest from the Paris office?
5. Who is the northernmost customer?
6. What is the distance between the Paris and Boston offices?
To be precise for long distances, the distance in kilometers, as the crow flies, between two points when you have latitude and longitude, is $(\text{ACOS}(\text{SIN}(\text{lat1} * \text{PI}() / 180) * \text{SIN}(\text{lat2} * \text{PI}() / 180) + \text{COS}(\text{lat1} * \text{PI}() / 180) * \text{COS}(\text{lat2} * \text{PI}() / 180) * \text{COS}((\text{lon1} - \text{lon2}) * \text{PI}() / 180)) * 180 / \text{PI}()) * 60 * 1.8532$

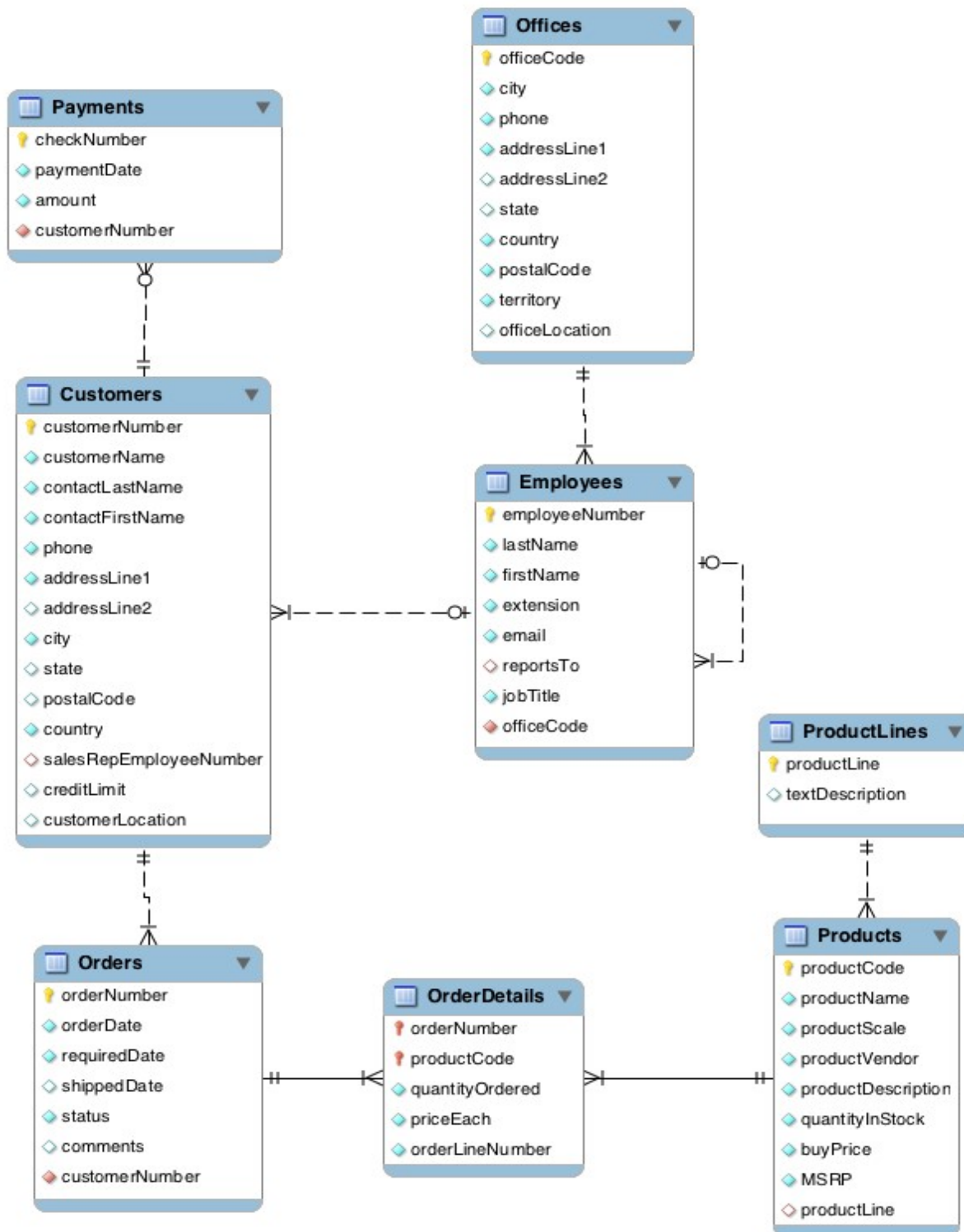
Data visualization

For each problem, use R to run an SQL query and then visualize the results. The first step in all cases

is to load the MySQL library and access the database.

1. Visualize in blue the number of items for each product scale.
2. Prepare a line plot with appropriate labels for total payments for each month in 2004.
3. Create a histogram with appropriate labels for the value of orders received from the Nordic countries (Denmark,Finland, Norway,Sweden).
4. Create a heatmap for product lines and Norwegian cities.
5. Create a parallel coordinates plot for product scale, quantity in stock, and MSRP in the Products table.

[Workbench file](#) for the following image.



This page is part of the promotional and support material for *Data Management (sixth edition)* by Richard T. Watson
For questions and comments please [contact the author](#)

Date revised: 02-Jan-2020