SQL: Final Project

This project is using the Classic Models database and MySQL. It was completed using MySQL Workbench.

1. Which three product lines were the most profitable in 2004?

SELECT productLine, SUM(quantityOrdered \* priceEach) AS productLineProfit

FROM products P

JOIN orderdetails D ON P.productCode = D.productCode

JOIN orders O ON D.orderNumber = O.orderNumber

WHERE year(orderDate) = 2004

GROUP BY productLine

ORDER BY 2 DESC LIMIT 3;

Answer:

|  |  |
| --- | --- |
| **productLine** | **productLineProfit** |
| Classic Cars | $1,763,136.73 |
| Motorcycles | $527,243.84 |
| Vintage Cars | $854,551.85 |

1. Which office location brought in the most profit per employee?

SELECT officeCode, ROUND(SUM(quantityOrdered \* priceEach)/COUNT(employeeNumber), 2) AS averageEmployeeSales

FROM employees E

JOIN customers C ON E.employeeNumber = C.salesRepEmployeeNumber

JOIN orders O ON C.customerNumber = O.customerNumber

JOIN orderdetails D ON D.orderNumber = O.orderNumber

GROUP BY officeCode

ORDER BY 2 DESC LIMIT 1;

Answer:

|  |  |
| --- | --- |
| **officeCode** | **averageEmployeeSales** |
| 5 | $3,336.57 |

1. List the employees with customers that have no orders.

SELECT CONCAT(firstName, ' ', lastName) AS employee

FROM customers C

JOIN employees E ON C.salesRepEmployeeNumber = E.employeeNumber

LEFT JOIN orders O ON O.customerNumber = C.customerNumber

WHERE O.ordernumber IS NULL;

Answer:

|  |
| --- |
| **employee** |
| Foon Yue Tseng |
| Martin Gerard |

**Reflection**:

I have learned the basics of SQL in this course. I do not know how to do extremely complex SQL statements, but I have all the tools I need to be able to figure those things out. I also know ways I can preserve the integrity, privacy, and security of the data I am working with. This means that, while I am still a novice, I will be able to continue to safely hone my skills and gain more intuition with working with databases. SQL is a powerful tool for data scientists because it allows them to intimately handle large amount of related data more efficiently. Things that would take days to figure out with a folder of spreadsheets can take minutes with SQL.