SQL PROJECT BY NIKHIL GUPTA



Project Missing Money Matters

# Challenges

**The situation:** For a long time, Adams Andrew, Manager of WSDA Music, has been unable to account for a discrepancy in his company’s financials.

The furthest he has gotten in his own attempts to analyze the company data is figuring out that the discrepancy occurred between the years 2011 and 2012. But that’s about all that Adams knows for certain.

You have been called in to do what you do best—apply your SQL skills:

* Analyze WSDA Music’s Data to:
  + Get a list of suspects
  + Narrow your list
  + Pinpoint your prime suspect(s)

The Management team at WSDA Music is eager to review your findings! Work through each part of the project below, then watch the corresponding final project movie to compare your answers.

# Considerations

Hmmm. OK, so where do we start? We have access to WSDA Music’s Data and a great first step is to revisit the situation, with the aim of zeroing in on the specific tables that may be relevant to start looking at first.

## Hint:

Money is missing so, did ***someone*** take it? Did ***something*** account for the discrepancy? What can be done? Starting with the assumption that the missing funds can be attributed to *someone*, what tables will most likely contain information about **people**? What related activity in this database do people do that creates a transaction record (since we are after a missing sum of money—in what tables would this type of business store this information?).

## Tables: Customers, Employees, Invoices CHALLENGE 1

General queries that begin to give you some high-level context

1. How many transactions took place between the years 2011 and 2012?

SELECT COUNT(\*) AS "NO. OF TRANSACTION"

FROM Invoice  
WHERE strftime('%Y', InvoiceDate) = '2011' OR strftime('%Y', InvoiceDate) = '2012'

1. How much money did WSDA Music make during the same period?

SELECT SUM(TOTAL) AS "NO. OF TRANSACTION"

FROM Invoice

WHERE strftime('%Y', InvoiceDate) = '2011' OR strftime('%Y', InvoiceDate) = '2012’

## CHALLENGE 2

More targeted questions that query tables containing data about customers and employees

1. Get a list of customers who made purchases between 2011 and 2012.

SELECT C.FirstName,C.LastName,I.InvoiceDate,I.total FROM Customer C

JOIN Invoice I ON I.CustomerId = C.CustomerId

WHERE InvoiceDate >= '2011-01-01' AND InvoiceDate <= '2012-12-31'

ORDER BY InvoiceDate

1. Get a list of customers, sales reps, and total transaction amounts for each customer between 2011 and 2012.

SELECT c.FirstName ,c.LastName ,e.FirstName ,e.LastName ,i.total FROM Invoice i

INNER JOIN Customer c ON i.CustomerId = c.CustomerId

INNER JOIN Employee e ON e.EmployeeId = c.SupportRepId

WHERE InvoiceDate >= '2011-01-01' AND InvoiceDate <='2012-12-31'

ORDER BY i.total DESC

1. How many transactions are above the average transaction amount during the same time period?

SELECT COUNT(\*) AS [Num of Transactions Above Avg] FROM Invoice

WHERE total > (SELECT ROUND(AVG(Total),2) from Invoice

WHERE InvoiceDate >= '2011-01-01' AND InvoiceDate <='2012-12-31')

1. What is the average transaction amount for each year that WSDA Music has been in business?

SELECT strftime('%Y',InvoiceDate) AS [Year],round(AVG(TOTAL),2) AS [Avg\_Total] from Invoice

GROUP BY strftime('%Y',InvoiceDate)

## CHALLENGE 3

Queries that perform in-depth analysis with the aim of finding employees who may have been financially motivated to commit a crime

1. Get a list of employees who exceeded the average transaction amount from sales they generated during 2011 and 2012.

SELECT C.FirstName,C.LastName,SUM(I.total) AS [Total\_Sum]

FROM

Customer C

JOIN

Invoice I ON I.CustomerId = C.CustomerId

WHERE

InvoiceDate >= '2011-01-01'

AND InvoiceDate <= '2012-12-31'

GROUP BY C.FirstName, C.LastName

HAVING

SUM(I.total) > (

SELECT ROUND(AVG(total), 2) AS [Avg Transaction Amount]

FROM Invoice

WHERE InvoiceDate >= '2011-01-01'

AND InvoiceDate <= '2012-12-31'

)

ORDER BY C.FirstName,C.LastName;

1. Create a Commission Payout column that displays each employee’s commission based on 15% of the sales transaction amount.

SELECT

e.FirstName,

e.LastName,

sum(i.total) AS [Total Sales],

round(sum(i.total) \*.15,2) AS [Commission Payout]

FROM

Invoice i

INNER JOIN

Customer c

ON i.CustomerId = c.CustomerId

INNER JOIN

Employee e

ON e.EmployeeId = c.SupportRepId

WHERE

InvoiceDate >= '2011-01-01' AND InvoiceDate <='2012-12-31'

GROUP BY

e.FirstName,

e.LastName

ORDER BY e.LastName

1. Which employee made the highest commission?

SELECT

e.FirstName,

e.LastName,

sum(i.total) AS [Total Sales],

round(sum(i.total) \*.15,2) AS [Commission Payout]

FROM

Invoice i

INNER JOIN

Customer c

ON i.CustomerId = c.CustomerId

INNER JOIN

Employee e

ON e.EmployeeId = c.SupportRepId

WHERE

InvoiceDate >= '2011-01-01' AND InvoiceDate <='2012-12-31'

GROUP BY

e.FirstName,

e.LastName

ORDER BY round(sum(i.total) \*.15,2) DESC

LIMIT 1;

1. List the customers that the employee identified in the last question.

SELECT CustomerId,FirstName,LastName from Customer

where SupportRepId in (SELECT

e.EmployeeId

FROM

Invoice i

INNER JOIN

Customer c

ON i.CustomerId = c.CustomerId

INNER JOIN Employee e ON e.EmployeeId = c.SupportRepId

WHERE InvoiceDate >= '2011-01-01' AND InvoiceDate <='2012-12-31'

GROUP BY e.EmployeeId

ORDER BY round(sum(i.total) \*.15,2) DESC

LIMIT 1)

1. Which customer made the highest purchase?

SELECT C.CustomerId,C.FirstName,C.LastName,SUM(total) as [Total\_Purchase] FROM Customer c

JOIN Invoice I on c.CustomerId = I.CustomerId

GROUP by C.CustomerId,C.FirstName,C.LastName

ORDER BY SUM(total) DESC

LIMIT 1;

1. Look at this customer record—do you see anything suspicious?

SELECT \* FROM Customer c

WHERE c.LastName = 'Doeein'

1. Who do you conclude is our primary person of interest?

Jane Peacock

Well done! You’ve uncovered insights by analyzing the data. WSDA Music management thanks you. Use these same SQL skills to analyze future data sets.

Additional learning resources are also shared in the final chapter of this course.