Technical Task

Report

Task “Missing Money Matters”

Prepared by

Spitkovska Vladyslava

Kyiv – 2024

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.

It's important to:

− Get a list of suspects
− Narrow your list
− Pinpoint your prime suspect(s)

Task list:

Challenge 1

General queries that begin to give you some high-level context ● How many transactions took place between the years 2011 and 2012?

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

Challenge 2 More targeted questions that query tables containing data about customers and employees

- Get a list of customers who made purchases between 2011 and 2012.
- Get a list of customers, sales reps, and total transaction amounts for each customer between 2011 and 2012
- How many transactions are above the average transaction amount during the same time period?
- What is the average transaction amount for each year that WSDA Music has been in business?

Challenge 3 Queries that perform in-depth analysis with the aim of finding employees who may have been financially motivated to commit a crime

- Get a list of employees who exceeded the average transaction amount from sales they generated during 2011 and 2012.
- Create a Commission Payout column that displays each employee's commission based on 15% of the sales transaction amount.
- Which employee made the highest commission?
- List the customers that the employee identified in the last question. ● Which customer made the highest purchase?

- Look at this customer record—do you see anything suspicious? ● Who do you conclude is our primary person of interest?

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

WSDA_Music.db

```
1 SELECT COUNT(*) AS transaction_count
2 FROM Invoice
3 WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31';
4
```

transaction_count
167

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

```
1 SELECT SUM(Total) || '$' AS total_money
2 FROM Invoice
3 WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31';
```

total_money
1947.97$

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

```
11 SELECT DISTINCT c.CustomerId, c.firstname || ' ' || c.lastname AS FullName
12 FROM Customer c
13 JOIN Invoice i ON c.CustomerId = i.CustomerId
14 WHERE i.InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31';
15
```

CustomerId	FullName
1	Luís Gonçalves
2	Leonie Köhler
3	François Tremblay
4	Bjørn Hansen
5	František Wichterlová
6	Helena Holý

(List contains 60 rows of data)

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

```
20 SELECT c.CustomerId, c.FirstName || ' ' || c.LastName AS CustomerName,
21 e.FirstName || ' ' || e.LastName AS SalesRepName, SUM(i.Total) AS TotalAmount
22 FROM Customer c
23 JOIN Invoice i ON c.CustomerId = i.CustomerId
24 LEFT JOIN  Employee e ON c.SupportRepId = e.EmployeeId
25 WHERE i.InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
26 GROUP BY c.CustomerId, CustomerName, SalesRepName;
27
```

CustomerId	CustomerName	SalesRepName	TotalAmount
1	Luís Gonçalves	Jane Peacock	16.83
2	Leonie Köhler	Steve Johnson	12.870000000000001
3	François Tremblay	Jane Peacock	5.9399999999999995
4	Bjørn Hansen	Margaret Park	26.75
5	František Wichterlová	Margaret Park	19.83
6	Helena Holý	Steve Johnson	12.870000000000001
7	Astrid Gruber	Steve Johnson	11.88
8	Daan Peeters	Margaret Park	24.75

(List contains 60 rows of data)

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

```
 5 WITH AvgTransaction AS (
 6     SELECT AVG(Total) AS avg_transaction_amount
 7     FROM Invoice
 8     WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
 9 )
10 SELECT COUNT(*) AS transactions_above_average
11 FROM Invoice, AvgTransaction
12 WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
13   AND Total > AvgTransaction.avg_transaction_amount;
14
```

transactions_above_average
26

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

```sql
SELECT strftime('%Y', InvoiceDate) AS Year, AVG(Total) AS avg_transaction_amount
FROM Invoice
GROUP BY Year
ORDER BY Year;
```

Year	avg_transaction_amount
2009	5.415180722891566
2010	5.800602409638554
2011	17.505238095238095
2012	5.753373493975904
2013	5.63225

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

```sql
WITH AvgTransaction AS (SELECT AVG(Total) AS avg_transaction_amount
FROM Invoice
WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31')
SELECT e.EmployeeId, i.InvoiceId, i.Total, e.FirstName || ' ' || e.LastName AS EmployeeName,
FROM Invoice i
JOIN Customer c ON i.CustomerId = c.CustomerId
JOIN Employee e ON c.SupportRepId = e.EmployeeId
WHERE i.InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
AND i.Total > (SELECT avg_transaction_amount FROM AvgTransaction)
ORDER BY e.EmployeeId, i.Total DESC;
```

EmployeeId	EmployeeName	InvoiceId	Total
3	Jane Peacock	413	1000.86
3	Jane Peacock	194	21.86
3	Jane Peacock	313	16.86
3	Jane Peacock	193	14.91
3	Jane Peacock	180	13.86
3	Jane Peacock	215	13.86
3	Jane Peacock	229	13.86
3	Jane Peacock	236	13.86

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

```sql
WITH AvgTransaction AS (SELECT AVG(Total) AS avg_transaction_amount
FROM Invoice
WHERE InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31')
SELECT e.EmployeeId, e.FirstName || ' ' || e.LastName AS EmployeeName, i.InvoiceId,
i.Total, (i.Total * 0.15) || '$' AS CommissionPayout
FROM Invoice i
JOIN Customer c ON i.CustomerId = c.CustomerId
JOIN Employee e ON c.SupportRepId = e.EmployeeId
JOIN AvgTransaction at
WHERE i.InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
AND i.Total > at.avg_transaction_amount
ORDER BY e.EmployeeId, i.Total DESC;
```

EmployeeId	EmployeeName	InvoiceId	Total	CommissionPayout
3	Jane Peacock	413	1000.86	150.129$
3	Jane Peacock	194	21.86	3.279$
3	Jane Peacock	313	16.86	2.529$
3	Jane Peacock	193	14.91	2.2365$
3	Jane Peacock	180	13.86	2.079$
3	Jane Peacock	215	13.86	2.079$
3	Jane Peacock	229	13.86	2.079$
3	Jane Peacock	236	13.86	2.079$
3	Jane Peacock	278	13.86	2.079$

Which employee made the highest commission?

```sql
WITH EmployeeCommissions AS (SELECT e.EmployeeId, e.FirstName || ' ' || e.LastName AS EmployeeName,
SUM(i.Total * 0.15) AS TotalCommission
FROM Invoice AS i
JOIN Customer AS c ON i.CustomerId = c.CustomerId
JOIN Employee AS e ON c.SupportRepId = e.EmployeeId
WHERE i.InvoiceDate BETWEEN '2011-01-01' AND '2012-12-31'
GROUP BY e.EmployeeId, e.FirstName,e.LastName)
SELECT EmployeeId,EmployeeName, TotalCommission
FROM EmployeeCommissions
ORDER BY TotalCommission DESC
LIMIT 1;

```

EmployeeId	EmployeeName	TotalCommission
3	Jane Peacock	199.76999999999998

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

```sql
SELECT c.CustomerId, c.FirstName || ' ' || c.LastName AS CustomerName
FROM Customer c
WHERE c.SupportRepId = 3;
```

CustomerId	CustomerName
1	Luís Gonçalves
3	François Tremblay
12	Roberto Almeida
15	Jennifer Peterson
18	Michelle Brooks
19	Tim Goyer
24	Frank Ralston

Which customer made the highest purchase?

```sql
SELECT c.CustomerId, c.FirstName || ' ' || c.LastName AS CustomerName,
SUM(i.Total) AS TotalSpent,(SUM(i.Total) )|| '$' AS TotalSpentDollars
FROM Customer c
JOIN Invoice i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId, c.FirstName,c.LastName
ORDER BY TotalSpent DESC
LIMIT 1;
```

CustomerId	CustomerName	TotalSpent	TotalSpentDollars
60	John Doeein	1000.86	1000.86$

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

C…	Fir…	La…	Co…	Ad…	City	State	Co…	Po…	Ph…	F…	Email
60	John	Doe…		NULL	NULL	NULL	NULL	NULL	NULL	N…	

I have decided to analyze the account of this customer and didn't find any contact information nor anything about his company. This fact made me feel suspicious about the user and I decided to continue researching info about other users (especially users connected with Jane Peacock,

C…	Cust…	Com…	Addr…	City	State	Cou…	Post…	Phone	Fax	Email	TotalSpent
38	Niklas…	NULL	Barba…	Berlin	NULL	Germ…	107…	+49 0…	NULL	nschr…	37.62
52	Emm…	NULL	202 H…	London	NULL	Unite…	N1 5LH	+44 0…	NULL	emm…	37.62
53	Phil H…	NULL	113 L…	London	NULL	Unite…	SW1…	+44 0…	NULL	phil.h…	37.62
59	Puja …	NULL	3,Raj …	Bang…	NULL	India	560…	+91 0…	NULL	puja_…	36.64

but didn't find any similar-looking accounts among them and the total spent was rather less them John's one.

Therefore I tried to became familiar with John's transactions.

```
3 SELECT *
4 FROM Invoice AS i
5 WHERE i.CustomerId = 60
6 ORDER BY i.InvoiceDate;
7
```

Invoi…	Custom…	Invoice…	BillingA…	BillingCity	BillingS…	BillingC…	BillingP…	Total
413	60	2011-11-…	NULL	NULL	NULL	NULL	NULL	1000.86

There's only one record about his transaction with total 1000.86$ without any billing information mentioned, which looks really suspicious.

Who do you conclude is our primary person of interest?

Our analysis reveals that this employee earns a 15% commission on sales transactions. The large transaction amount for the customer with missing information suggests the possibility that the employee created a fake customer record and transaction to boost their commission earnings. As I found no similar anomalies in the data of other customers handled by the same employee, which points towards a deliberate attempt to manipulate records for financial gain.

My assumption according to the previously mentioned information: employee ID 3 is suspected of fraudulently increasing their commission by creating a fake customer with missing contact details, resulting in a single large transaction over $1000. This manipulation suggests deliberate falsification of records for personal financial gain.