



IBM



Data Transformation-IBM HR Analytics

By: Frederick Zoreta

Data Transformations Using 3 Methods



SQL Server -
SSMS



MS Excel



Power Pivot

Microsoft



Microsoft®
SQL Server®



Ex

Create A Table: Employee_Sales

◆ Using SQL Server:

Query3.sql - LA...NE0CN4\ericz (57))*

SQLQuery2.sql - LA...NE0CN4\ericz (54))*

SQLQuery1.sql - LA...NE0CN4\ericz (53))*

```
1 CREATE TABLE employee_sales (  
2     Attrition VARCHAR(50),  
3     Department VARCHAR(50),  
4     JobSatisfaction INT,  
5     MonthlyIncome DECIMAL(10,2)  
6 );  
7
```

Messages

Commands completed successfully.

Completion time: 2025-02-19T18:56:44.1469359-05:00

Creating Employee_Sales: Using Excel

- ◇ Manually creating a blank sheet, and copying only the 4 attributes (columns) needed:


| | A | B | C | D | E | F |
|----|-----------|------------------------|-----------------|---------------|---|---|
| 1 | Attrition | Department | JobSatisfaction | MonthlyIncome | | |
| 2 | Yes | Sales | 4 | 5993 | | |
| 3 | No | Research & Development | 2 | 5130 | | |
| 4 | Yes | Research & Development | 3 | 2090 | | |
| 5 | No | Research & Development | 3 | 2909 | | |
| 6 | No | Research & Development | 2 | 3468 | | |
| 7 | No | Research & Development | 4 | 3068 | | |
| 8 | No | Research & Development | 1 | 2670 | | |
| 9 | No | Research & Development | 3 | 2693 | | |
| 10 | No | Research & Development | 3 | 9526 | | |
| 11 | No | Research & Development | 3 | 5237 | | |
| 12 | No | Research & Development | 2 | 2426 | | |
| 13 | No | Research & Development | 3 | 4193 | | |
| 14 | No | Research & Development | 3 | 2911 | | |
| 15 | No | Research & Development | 4 | 2661 | | |
| 16 | Yes | Research & Development | 3 | 2028 | | |
| 17 | No | Research & Development | 1 | 9980 | | |
| 18 | No | Research & Development | 2 | 3298 | | |
| 19 | No | Research & Development | 4 | 2935 | | |
| 20 | No | Sales | 4 | 15427 | | |
| 21 | No | Research & Development | 4 | 3944 | | |
| 22 | No | Research & Development | 3 | 4011 | | |
| 23 | Yes | Sales | 1 | 3407 | | |
| 24 | No | Research & Development | 2 | 11994 | | |
| 25 | No | Research & Development | 4 | 1232 | | |
| 26 | Yes | Research & Development | 1 | 2960 | | |

< >

IBM DataSet

Employee Sales

+

Ready  Accessibility: Unavailable

Creating Employee_Sales: Pivot Table Way

◇ An initial view using Excel's Pivot Table:

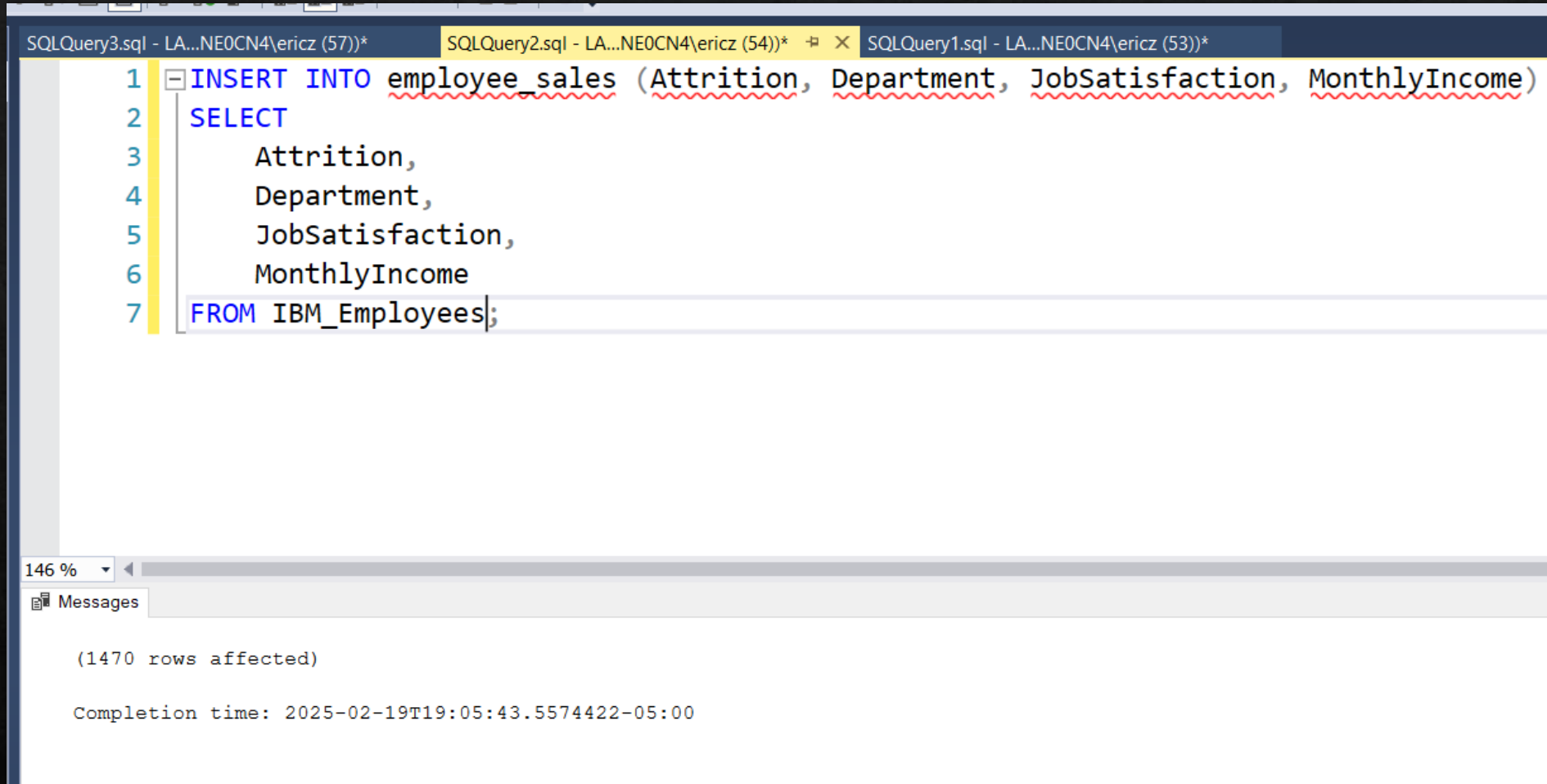
The screenshot shows an Excel PivotTable with the following data:

| Department | Attrition | Sum of JobSatisfaction | Sum of MonthlyIncome |
|------------------------|-----------|------------------------|----------------------|
| Human Resources | No | 138 | \$ 374,645.00 |
| | Yes | 26 | \$ 44,589.00 |
| Research & Development | No | 2293 | \$ 5,489,910.00 |
| | Yes | 327 | \$ 546,374.00 |
| Sales | No | 995 | \$ 2,560,213.00 |
| | Yes | 232 | \$ 543,578.00 |
| Grand Total | | 4011 | \$ 9,559,309.00 |

The PivotTable Fields task pane on the right shows the following configuration:

- Choose fields to add to report:** JobSatisfaction, MonthlyIncome (checked); JobInvolvement, JobLevel, JobRole, MaritalStatus, MonthlyRate (unchecked).
- Filters:** (Empty)
- Columns:** Values (Sum of JobSatisfaction, Sum of MonthlyIncome)
- Rows:** Department, Attrition
- Defer Layout Update:** (Unchecked)

Loading the IBM_Employee Table into 'Employee_Sales' : SQL Server



The screenshot displays the SQL Server Enterprise Manager interface. At the top, there are three tabs for SQL queries: 'SQLQuery3.sql - LA...NE0CN4\ericz (57))*', 'SQLQuery2.sql - LA...NE0CN4\ericz (54))*', and 'SQLQuery1.sql - LA...NE0CN4\ericz (53))*'. The active window shows an SQL query with line numbers 1 through 7 on the left margin. The query is an INSERT statement that loads data from the 'IBM_Employees' table into the 'employee_sales' table, specifically into columns 'Attrition', 'Department', 'JobSatisfaction', and 'MonthlyIncome'. Below the query editor, a 'Messages' pane shows the execution result: '(1470 rows affected)' and the 'Completion time: 2025-02-19T19:05:43.5574422-05:00'.

```
1 INSERT INTO employee_sales (Attrition, Department, JobSatisfaction, MonthlyIncome)
2 SELECT
3     Attrition,
4     Department,
5     JobSatisfaction,
6     MonthlyIncome
7 FROM IBM_Employees;
```

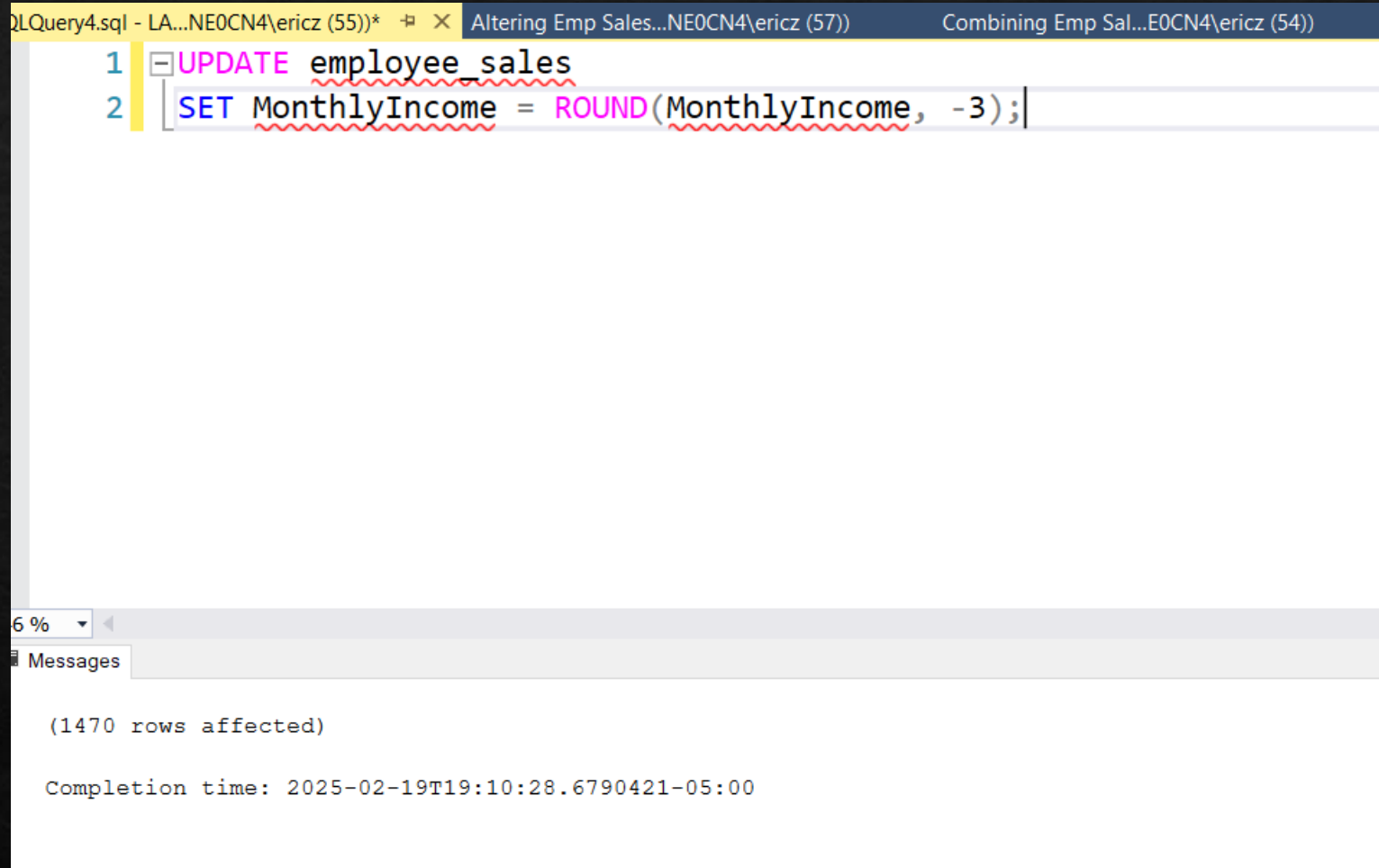
146 %

Messages

(1470 rows affected)

Completion time: 2025-02-19T19:05:43.5574422-05:00

Rounding the Monthly income in SQL Server



The screenshot displays the SQL Server Enterprise Manager interface. At the top, three tabs are visible: 'DLQuery4.sql - LA...NE0CN4\ericz (55))*', 'Alterning Emp Sales...NE0CN4\ericz (57))', and 'Combining Emp Sal...E0CN4\ericz (54))'. The active window shows a SQL query with two lines: '1 UPDATE employee_sales' and '2 SET MonthlyIncome = ROUND(MonthlyIncome, -3);'. Below the query editor, a progress bar indicates '6 %' completion. A 'Messages' tab is active, showing the execution results: '(1470 rows affected)' and 'Completion time: 2025-02-19T19:10:28.6790421-05:00'.

```
1 UPDATE employee_sales
2 SET MonthlyIncome = ROUND(MonthlyIncome, -3);
```

6 %

Messages

(1470 rows affected)

Completion time: 2025-02-19T19:10:28.6790421-05:00

Rounding the Monthly Income in Excel: Using the formula =ROUND(D2,-3)
* I added an extra column that serves as the 'Rounded to the nearest 1k'

Excel Ribbon:

- Clipboard:** Paste
- Font:** Aptos Narrow, Size 11, Bold (B), Italic (I), Underline (U), Color (A)
- Alignment:** Wrap Text, Merge & Center
- Number:** Accounting, \$, %, , #,0.00, #.00
- Conditional Formatting:** Conditional Formatting
- Formulas:** Formulas

Formula Bar: E2 =ROUND(D2,-3)

| A | B | C | D | E |
|-----------|------------------------|-----------------|---------------|-----------------------|
| Attrition | Department | JobSatisfaction | MonthlyIncome | Rounded_MonthlyIncome |
| Yes | Sales | 4 | \$ 5,993.00 | \$ 6,000.00 |
| No | Research & Development | 2 | \$ 5,130.00 | \$ 5,000.00 |
| Yes | Research & Development | 3 | \$ 2,090.00 | \$ 2,000.00 |
| No | Research & Development | 3 | \$ 2,909.00 | \$ 3,000.00 |
| No | Research & Development | 2 | \$ 3,468.00 | \$ 3,000.00 |
| No | Research & Development | 4 | \$ 3,068.00 | \$ 3,000.00 |
| No | Research & Development | 1 | \$ 2,670.00 | \$ 3,000.00 |
| No | Research & Development | 3 | \$ 2,693.00 | \$ 3,000.00 |
| No | Research & Development | 3 | \$ 9,526.00 | \$ 10,000.00 |
| No | Research & Development | 3 | \$ 5,237.00 | \$ 5,000.00 |
| No | Research & Development | 2 | \$ 2,426.00 | \$ 2,000.00 |
| No | Research & Development | 3 | \$ 4,193.00 | \$ 4,000.00 |
| No | Research & Development | 3 | \$ 2,911.00 | \$ 3,000.00 |
| No | Research & Development | 4 | \$ 2,661.00 | \$ 3,000.00 |
| Yes | Research & Development | 3 | \$ 2,028.00 | \$ 2,000.00 |
| No | Research & Development | 1 | \$ 9,980.00 | \$ 10,000.00 |
| No | Research & Development | 2 | \$ 3,298.00 | \$ 3,000.00 |
| No | Research & Development | 4 | \$ 2,935.00 | \$ 3,000.00 |
| No | Sales | 4 | \$ 15,427.00 | \$ 15,000.00 |
| No | Research & Development | 4 | \$ 3,944.00 | \$ 4,000.00 |
| No | Research & Development | 3 | \$ 4,011.00 | \$ 4,000.00 |
| Yes | Sales | 1 | \$ 3,407.00 | \$ 3,000.00 |
| No | Research & Development | 2 | \$ 11,994.00 | \$ 12,000.00 |
| No | Research & Development | 4 | \$ 1,232.00 | \$ 1,000.00 |
| Yes | Research & Development | 1 | \$ 2,960.00 | \$ 3,000.00 |

Status Bar: Pivot | IBM DataSet | Employee Sales

Filtering only the 'Sales' Dept.

SQLQuery5.sql - LA...NE0CN4\ericz (53))* X Altering Emp Sales...NE0CN4\ericz (57))

```
1 SELECT *
2 FROM employee_sales
3 WHERE Department = 'Sales';
```

146 %

Results Messages

| | Attrition | Department | JobSatisfaction | MonthlyIncome |
|----|-----------|------------|-----------------|---------------|
| 1 | Yes | Sales | Very High | 2000.00 |
| 2 | No | Sales | Very High | 14000.00 |
| 3 | Yes | Sales | High | 5000.00 |
| 4 | No | Sales | Very High | 7000.00 |
| 5 | Yes | Sales | High | 3000.00 |
| 6 | Yes | Sales | Medium | 3000.00 |
| 7 | No | Sales | Medium | 2000.00 |
| 8 | No | Sales | Very High | 14000.00 |
| 9 | Yes | Sales | Low | 2000.00 |
| 10 | No | Sales | High | 3000.00 |
| 11 | No | Sales | High | 6000.00 |
| 12 | Yes | Sales | Medium | 2000.00 |
| 13 | No | Sales | Medium | 10000.00 |
| 14 | Yes | Sales | Medium | 2000.00 |
| 15 | Yes | Sales | Medium | 13000.00 |
| 16 | No | Sales | High | 18000.00 |
| 17 | No | Sales | Very High | 7000.00 |
| 18 | No | Sales | Very High | 11000.00 |
| 19 | No | Sales | Very High | 7000.00 |
| 20 | No | Sales | Medium | 7000.00 |
| 21 | No | Sales | Low | 3000.00 |
| 22 | No | Sales | Very High | 3000.00 |
| 23 | Yes | Sales | Very High | 2000.00 |

| A | B | C | D | E |
|-----------|------------|-----------------|----------------|-----------------------|
| Attrition | Department | JobSatisfaction | MonthlyIncome | Rounded_MonthlyIncome |
| Yes | Sales | | 4 \$ 5,993.00 | \$ 6,000.00 |
| No | Sales | | 4 \$ 15,427.00 | \$ 15,000.00 |
| Yes | Sales | | 1 \$ 3,407.00 | \$ 3,000.00 |
| No | Sales | | 2 \$ 6,825.00 | \$ 7,000.00 |
| No | Sales | | 1 \$ 18,947.00 | \$ 19,000.00 |
| Yes | Sales | | 4 \$ 2,086.00 | \$ 2,000.00 |
| Yes | Sales | | 3 \$ 2,683.00 | \$ 3,000.00 |
| No | Sales | | 4 \$ 2,014.00 | \$ 2,000.00 |
| No | Sales | | 1 \$ 5,376.00 | \$ 5,000.00 |
| No | Sales | | 3 \$ 8,726.00 | \$ 9,000.00 |
| No | Sales | | 3 \$ 4,568.00 | \$ 5,000.00 |
| No | Sales | | 4 \$ 5,772.00 | \$ 6,000.00 |
| No | Sales | | 1 \$ 5,454.00 | \$ 5,000.00 |
| No | Sales | | 4 \$ 4,157.00 | \$ 4,000.00 |
| No | Sales | | 1 \$ 9,069.00 | \$ 9,000.00 |
| No | Sales | | 1 \$ 7,637.00 | \$ 8,000.00 |
| No | Sales | | 3 \$ 5,473.00 | \$ 5,000.00 |
| No | Sales | | 1 \$ 4,312.00 | \$ 4,000.00 |
| No | Sales | | 4 \$ 10,239.00 | \$ 10,000.00 |
| No | Sales | | 1 \$ 2,322.00 | \$ 2,000.00 |
| Yes | Sales | | 4 \$ 9,619.00 | \$ 10,000.00 |
| No | Sales | | 4 \$ 5,441.00 | \$ 5,000.00 |
| No | Sales | | 2 \$ 5,209.00 | \$ 5,000.00 |
| No | Sales | | 3 \$ 5,010.00 | \$ 5,000.00 |
| No | Sales | | 3 \$ 4,999.00 | \$ 5,000.00 |

> Pivot IBM DataSet Employee Sales +

Ordering the data by DESCending 'Job Satisfaction' and DESCending 'Job Satisfaction & Monthly Income' - SQL Version

SQLQuery6.sql - LA...NE0CN4\ericz (54))* SQLQuery5.sql - LA...NE0CN4\ericz

```
1 SELECT *
2 FROM employee_sales
3 WHERE Department = 'Sales'
4 ORDER BY JobSatisfaction DESC;
```

146 %

Results Messages

| | Attrition | Department | JobSatisfaction | MonthlyIncome |
|----|-----------|------------|-----------------|---------------|
| 1 | Yes | Sales | Very High | 2000.00 |
| 2 | No | Sales | Very High | 14000.00 |
| 3 | No | Sales | Very High | 7000.00 |
| 4 | No | Sales | Very High | 14000.00 |
| 5 | No | Sales | Very High | 7000.00 |
| 6 | No | Sales | Very High | 11000.00 |
| 7 | No | Sales | Very High | 7000.00 |
| 8 | No | Sales | Very High | 3000.00 |
| 9 | Yes | Sales | Very High | 3000.00 |
| 10 | Yes | Sales | Very High | 1000.00 |
| 11 | No | Sales | Very High | 18000.00 |
| 12 | No | Sales | Very High | 4000.00 |
| 13 | No | Sales | Very High | 3000.00 |
| 14 | No | Sales | Very High | 5000.00 |
| 15 | Yes | Sales | Very High | 2000.00 |
| 16 | No | Sales | Very High | 5000.00 |
| 17 | No | Sales | Very High | 3000.00 |

Query executed successfully.

SQLQuery6.sql - LA...NE0CN4\ericz (54))* SQLQuery5.sql - LA...NE0CN4\ericz (53))* Altering Emp Sal

```
1 SELECT *
2 FROM employee_sales
3 WHERE Department = 'Sales'
4 ORDER BY JobSatisfaction, MonthlyIncome DESC;
```

146 %

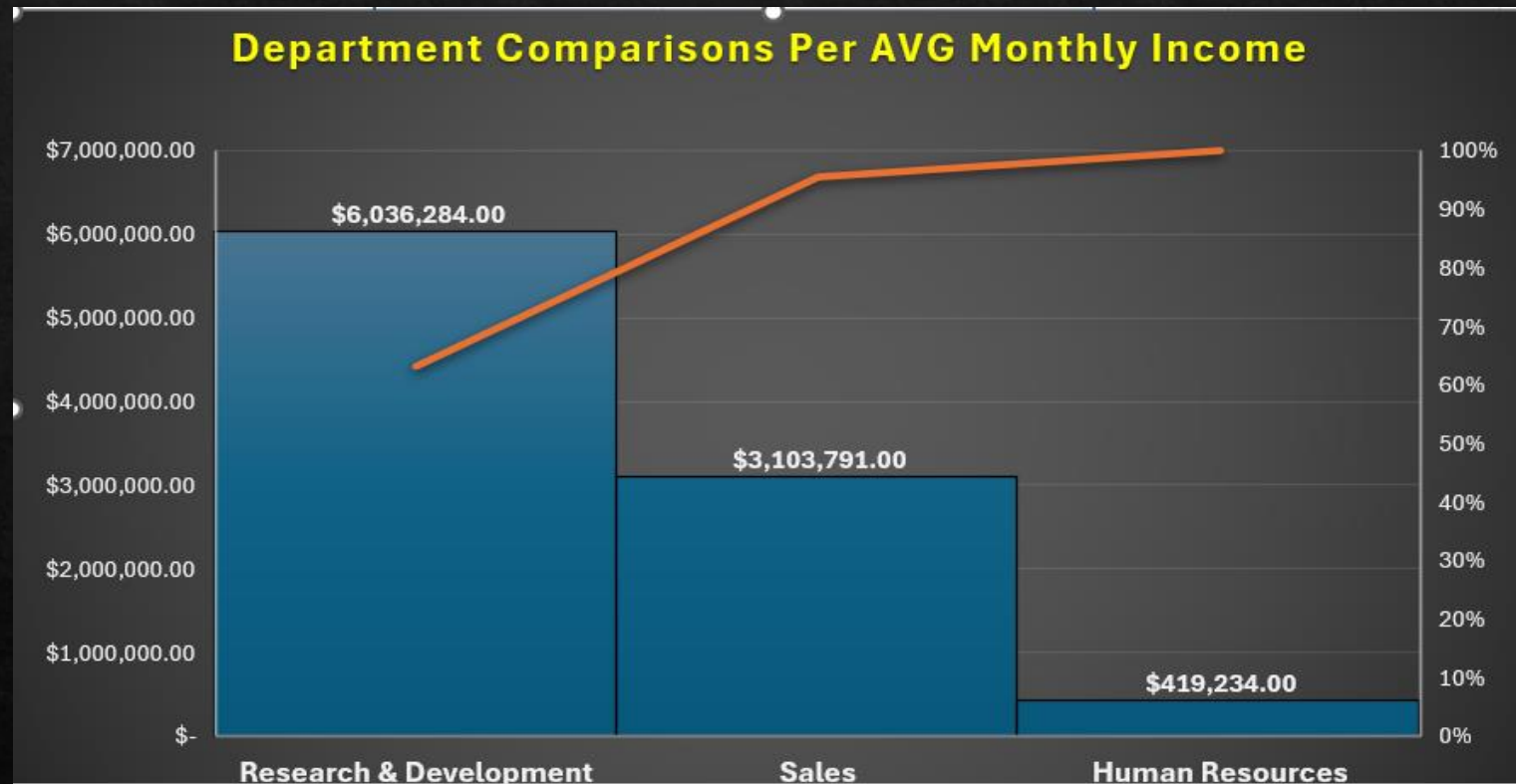
Results Messages

| | Attrition | Department | JobSatisfaction | MonthlyIncome |
|----|-----------|------------|-----------------|---------------|
| 1 | No | Sales | High | 19000.00 |
| 2 | No | Sales | High | 19000.00 |
| 3 | No | Sales | High | 18000.00 |
| 4 | No | Sales | High | 18000.00 |
| 5 | No | Sales | High | 17000.00 |
| 6 | No | Sales | High | 17000.00 |
| 7 | No | Sales | High | 16000.00 |
| 8 | No | Sales | High | 15000.00 |
| 9 | No | Sales | High | 14000.00 |
| 10 | No | Sales | High | 14000.00 |
| 11 | Yes | Sales | High | 14000.00 |
| 12 | Yes | Sales | High | 13000.00 |
| 13 | No | Sales | High | 13000.00 |
| 14 | Yes | Sales | High | 11000.00 |
| 15 | No | Sales | High | 11000.00 |
| 16 | No | Sales | High | 11000.00 |
| 17 | Yes | Sales | High | 10000.00 |

Query executed successfully.

LAPTO

Additional Insights using Data Visualizations:



A Brief Overview of Monthly Income Comparisons VS Job Roles

| JobRole ▼ | Total Monthly Income | Average Monthly Income | Max Monthly Income | Min Monthly Income |
|---------------------------|----------------------|------------------------|--------------------|--------------------|
| Healthcare Representative | \$ 986,268.00 | \$ 7,528.76 | \$ 13,966.00 | \$ 4,000.00 |
| Human Resources | \$ 220,259.00 | \$ 4,235.75 | \$ 10,725.00 | \$ 1,555.00 |
| Laboratory Technician | \$ 838,427.00 | \$ 3,237.17 | \$ 7,403.00 | \$ 1,102.00 |
| Manager | \$ 1,752,531.00 | \$ 17,181.68 | \$ 19,999.00 | \$ 11,244.00 |
| Manufacturing Director | \$ 1,057,795.00 | \$ 7,295.14 | \$ 13,973.00 | \$ 4,011.00 |
| Research Director | \$ 1,282,684.00 | \$ 16,033.55 | \$ 19,973.00 | \$ 11,031.00 |
| Research Scientist | \$ 946,072.00 | \$ 3,239.97 | \$ 9,724.00 | \$ 1,009.00 |
| Sales Executive | \$ 2,257,315.00 | \$ 6,924.28 | \$ 13,872.00 | \$ 4,001.00 |
| Sales Representative | \$ 217,958.00 | \$ 2,626.00 | \$ 6,632.00 | \$ 1,052.00 |
| Grand Total | \$ 9,559,309.00 | \$ 6,502.93 | \$ 19,999.00 | \$ 1,009.00 |

Max Monthly Income Min Monthly Income

Monthly Income Per Job Roles (Max & Min Comparisons)

