DATA DRIVEN ANALYTICS PROJECT PART 1 AND 2

MENTOR : MUNNA PANDEY

SUBMITTED BY : DASARI DEEPAK VENKAT SAIRAM

Part 1

Task 1: Customer Data Analysis.

1. Find the top 10 customers by credit limit.

customerNumber	customerName	creditLimit
141	Euro+ Shopping Channel	227600
124	Mini Gifts Distributors Ltd.	210500
298	Vida Sport, Ltd	141300
151	Muscle Machine Inc	138500
187	AV Stores, Co.	136800
146	Saveley & Henriot, Co.	123900
286	Marta's Replicas Co.	123700
386	L'ordine Souveniers	121400
227	Heintze Collectables	120800
259	Toms SpezialitÃ¤ten, Ltd	120400

Interpretation:

The command queries the modelcarsdb database, retrieves the customerNumber, customerName, and creditLimit for each customer in the customers table, sorts the results by creditLimit in descending order, and returns only the top 10 customers with the highest credit limits.

Part 1

Task 1: Customer Data Analysis.

2. Find the average credit limit for customers in each country.

country	average_credit
France	77691.66667
USA	78102.77778
Australia	86060
Norway	91200
Poland	0
Germany	19776.92308
Spain	73971.42857
Sweden	84750
Denmark	102100
Singapore	67233.33333
Portugal	0
Japan	87800
Finland	95266.66667
UK	88740
Ireland	34700
Canada	76200

Interpretation:

The command calculates the average credit limit (average_creditLimit) for customers in each country. It groups the results by country, so you get one row per country showing the average credit limit of customers from that country.

Part 1
Task 2: Office Data Analysis

1. Count the number of employees working in each office.

officeCode	numberofemployees
1	6
2	2
3	2
4	5
5	2
6	4
7	2

Interpretation:

The command retrieves the count of employees for each office code. It groups the employees by officeCode and counts the number of employees in each group, providing a total count per office. The result will include each office code and the corresponding number of employees working at that office.

Part1
Task 2: Office Data Analysis

2. Identify the offices with less than a certain number of employees.

officeCode	numberofemployees
2	2
3	2
5	2
7	2

Interpretation:

The command retrieves office codes and the count of employees for each office code where the number of employees is less than 3. It groups employees by officeCode, counts the number of employees in each group, and then filters to include only those groups where the count is fewer than 3.

Part 1
Task 3: Product Data Analysis

1. Count the number of products in each product line

productline	numberofproducts
Classic Cars	38
Vintage Cars	24
Motorcycles	13
Planes	12
Trucks and Buses	11
Ships	9
Trains	3

Interpretation:

The command retrieves the count of products for each product line from the products table. It groups the results by productline, counts the number of products within each product line, and orders the results in descending order based on the count. This allows you to see which product lines have the most products.

Part 1
Task 3: Product Data Analysis

2.Find the product line with the highest average product price.

productline	averageprice
Classic Cars	64.446316

Interpretation:

The command retrieves the product line with the highest average buy price from the products table. It calculates the average buy price for each product line, groups the results by productline, orders the results in descending order of average price, and limits the result to the top product line. This allows you to identify which product line has the highest average buy price.

Part2
Task 1: Employee Data Analysis

1. Find the total number of employees.

total_employees
23

Interpretation:

The command retrieves the total number of employees from the table and labels this count as 'total_employees ' in the result. If the table has 150 employees, for instance,the result will be 'total_employees ' = 150.

Part2

Task 1: Employee Data Analysis

2. List all employees with their basic information.

employee	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle
1002	Murphy	Diane	x5800	dmurphy@	1	NULL	President
1056	Patterson	Mary	x4611	mpatterso	1	1002	VP Sales
1076	Firrelli	Jeff	x9273	jfirrelli@cl	1	1002	VP Marketing
1088	Patterson	William	x4871	wpatterso	6	1056	Sales Manager (APAC)
1102	Bondur	Gerard	x5408	gbondur@	4	1056	Sale Manager (EMEA)
1143	Bow	Anthony	x5428	abow@cla	1	1056	Sales Manager (NA)
1165	Jennings	Leslie	x3291	ljennings@	1	1143	Sales Rep
1166	Thompson	Leslie	x4065	lthompsor	1	1143	Sales Rep
1188	Firrelli	Julie	x2173	jfirrelli@cl	2	1143	Sales Rep
1216	Patterson	Steve	x4334	spatterson	2	1143	Sales Rep
1286	Tseng	Foon Yue	x2248	ftseng@cl	3	1143	Sales Rep
1323	Vanauf	George	x4102	gvanauf@	3	1143	Sales Rep
1337	Bondur	Loui	x6493	lbondur@	4	1102	Sales Rep
1370	Hernande	Gerard	x2028	ghernande	4	1102	Sales Rep
1401	Castillo	Pamela	x2759	pcastillo@	4	1102	Sales Rep
1501	Bott	Larry	x2311	lbott@clas	7	1102	Sales Rep
1504	Jones	Barry	x102	bjones@c	7	1102	Sales Rep
1611	Fixter	Andy	x101	afixter@cl	6	1088	Sales Rep
1612	Marsh	Peter	x102	pmarsh@	6	1088	Sales Rep
1619	King	Tom	x103	tking@cla	6	1088	Sales Rep
1621	Nishi	Mami	x101	mnishi@c	5	1056	Sales Rep
1625	Kato	Yoshimi	x102	ykato@cla	5	1621	Sales Rep
1702	Gerard	Martin	x2312	mgerard@	4	1102	Sales Rep

Interpretation:

The command will return a result set containing every row and every column in the employees table. If the table has columns like employee_id,name,position etc., all these columns will be included in the output for every row in the table.

Part2
Task 1: Employee Data Analysis

3. Count the number of employees holding each job title.

jobTitle	employee_count
President	1
VP Sales	1
VP Marketing	1
Sales Manager (APAC)	1
Sale Manager (EMEA)	1
Sales Manager (NA)	1
Sales Rep	17

Interpretation:

The SQL command is a powerful tool for summarizing and analyzing the distribution of job titles within the table. By grouping the data by job title and counting the number of employees in each group, this query provides a clear view of how many employees hold each job title. This can be useful for understanding workforce composition, identifying common roles, and making data-driven decisions related to staffing and organizational structure.

Part2

Task 1: Employee Data Analysis

4. Find the employees who don't have a manager (reports To is NULL).

employeeNumber	lastName	firstName
1002	Murphy	Diane

Interpretation:

The query returns a list of employees who do not report to anyone, including their employee number, last name, and first name. This might be used to identify top-level executives or employees who are at the highest level in the organizational hierarchy.

Part2

Task 2: Order Analysis

1. Find the average order amount for each customer.

customer	customerName	Avg_order_am
103	Atelier graphiqu	3187.765714
112	Signal Gift Store	2764.861379
114	Australian Colle	3283.364909
119	La Rochelle Gift	2991.94566
121	Baane Mini Imp	3257.024688
124	Mini Gifts Distri	3287.929667
128	Blauer See Auto	3451.716364
129	Mini Wheels Co	3176.693333
131	Land of Toys In	3042.554082
141	Euro+ Shopping	3168.685483
144	Volvo Model Re	3510.253684
145	Danish Wholes	3585.697778
146	Saveley & Henri	3178.179268
148	Dragon Souven	3633.744884

Interpretation:

The query retrieves the average order amount for each customer. It calculates the total amount for each order by multiplying 'quantityOrdered' by 'priceEach', then computes the average of these amounts for each customer. The result is a list of customers with their customer numbers, names, and the average order amount they have placed. The list is ordered by customer number.

Part2

Task 2: Order Analysis

2. Find the number of orders placed in each month.

month_number	order_count
1	25
2	26
3	27
4	29
5	29
6	19
7	18
8	17
9	20
10	31
11	63
12	22

Interpretation:

The query provides a summary of the number of orders placed each month. It returns a result set where each row represents a month (from 1 to 12), along with the count of orders placed in that month.

Part2

Task 2: Order Analysis

13.Identify orders with delayed shipping (shippedDate > requiredDate).

ordernumber	orderdate	customernumber	customername
10165	22-10-2003	148	Dragon Souveniers, Ltd.

Interpretation:

The query retrieves details about orders that were shipped after their required dates, including the order number, order date, customer number, and customer name. It lists these details in ascending order of order number.

Part1 Conclusion:

- Comprehensive Coverage: Your queries cover essential aspects of customer, office, and product data, providing a well-rounded analysis.
- Actionable Insights: The results can guide strategic decisions in sales, marketing, and resource management.
- Efficiency: Queries are generally well-structured, though some could be optimized for performance with alternative approaches.Overall, your SQL queries are effective tools for deriving meaningful insights and making informed business decisions.

Part 2 Conclusion;

- Overall, these analyses offer valuable insights into both employee performance and operational efficiency. By leveraging this data, the company can:
- Optimize workforce management and improve employee performance through targeted support and recognition.
- Enhance operational efficiency by addressing issues in the order fulfillment process and refining inventory management.
- Increase revenue and profitability by understanding customer behavior, popular products, and high-value orders, allowing for strategic adjustments in sales and marketing efforts.
- In essence, these analyses support data-driven decision-making, enabling the company to improve its overall performance and achieve its business objectives more effectively.

THANK YOU.