SQL PROJECT

Basic Analysis

1. Find the total revenue (sum of Quantity * Unit Price) generated from all invoices.

```
6    -- Basic Analysis
7     -- 1. Find the total revenue (sum of Quantity * UnitPrice) generated from all invoices
8    Select sum(Quantity * UnitPrice) as total_revenue
9    from retail;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

total_revenue
9747747.859567512

2. Count the number of unique products (Stock Code) sold.

```
11    -- 2. Count the number of unique products (StockCode) sold
12   Select count(distinct Stockcode) as unique_product
13   from retail;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

unique_product
3958

3. Identify the total number of invoices in the dataset

```
15    -- 3. Identify the total number of invoices in the dataset
16    Select Count(distinct InvoiceNo) as inv_total
17   from retail;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

inv_total
25900

4. Find the total quantity of products sold for each Stock Code and sort them in descending order.

```
19    -- 4. Find the total quantity of products sold for each StockCode and sort them in descending order
20    Select StockCode,
21        SUM(Quantity) as total_quantity
22    from retail
23    group by Stockcode
24    order by total_quantity desc;
25
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

StockCode	total_quantity
22197	56450
84077	53847
85099B	47363
85123A	39122
84879	36221
21212	36039
23084	30646
22492	26437
22616	26315
21977	24753
22178	23854
17003	23053
15036	22552
21915	22066

Result 15

5. Count the number of transactions (distinct Invoice No) per customer (Customer ID)

```
26    -- 5. Count the number of transactions (distinct InvoiceNo) per customer (CustomerID)
27    Select CustomerID,
28        Count(distinct InvoiceNo) as Transactions
29    from retail
30    group by CustomerID
31    Order by Transactions;
32
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

CustomerID	Transactions
12715	1
14641	1
17274	1
12349	1
12350	1
13464	1

Customer Analysis

1. Identify the top 5 customers who have generated the highest revenue

```
33      /* Customer Analysis */
34
35      -- 1. Identify the top 5 customers who have generated the highest revenue
36      SELECT CustomerID,
37          SUM(Quantity * UnitPrice) as total_revenue
38      from retail
39      group by CustomerID
40      Order by total_revenue desc
41      limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

CustomerID	total_revenue
14646	282038.7394294441
18102	272390.59846553206
17450	187523.77097034454
14911	147375.14925409108
14156	126339.51951554418

2. Find the average number of products purchased per customer.

```
43      -- 2. Find the average number of products purchased per customer
44      Select CustomerID,
45          avg(Quantity) as avg_quantity
46      from retail
47      group by CustomerID
48      order by avg_quantity desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

CustomerID	avg_quantity
13135	4300.0000
16754	2140.0000
16308	2000.0000
15195	1404.0000
13256	1141.0000
14609	838.4000
13848	740.0000

Result 5

3. Retrieve all transactions made by the customer who has purchased the most products in total.

```
50      -- 3. Retrieve all transactions made by the customer who has purchased the most products in total.
51  •   Select CustomerID,
52          SUM(Quantity) as total_quantity
53      from retail
54      group by CustomerID
55      order by total_quantity desc
56      limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

CustomerID	total_quantity
14646	198072

```
58  •   Select * from retail
59      Where CustomerID = (
60              Select CustomerID
61              from retail
62              group by CustomerID
63              order by sum(Quantity) desc
64              limit 1
65      );
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
539491	21981	PACK OF 12 WOODLAND TISSUES	12	2010-12-20 10:09:00	0.29	14646	Netherlands
539491	21986	PACK OF 12 PINK POLKADOT TISSUES	12	2010-12-20 10:09:00	0.29	14646	Netherlands
539491	22720	SET OF 3 CAKE TINS PANTRY DESIGN	2	2010-12-20 10:09:00	4.95	14646	Netherlands
539491	21931	JUMBO STORAGE BAG SUKI	1	2010-12-20 10:09:00	1.95	14646	Netherlands
539491	22613	PACK OF 20 SPACEBOY NAPKINS	2	2010-12-20 10:09:00	0.85	14646	Netherlands
539491	20751	FUNKY WASHING UP GLOVES ASSORTED	1	2010-12-20 10:09:00	2.1	14646	Netherlands
539491	21246	RED RETROSPOT BIG BOWL	2	2010-12-20 10:09:00	4.95	14646	Netherlands
539491	22960	JAM MAKING SET WITH JARS	1	2010-12-20 10:09:00	4.25	14646	Netherlands
539491	22355	CHARLOTTE BAG SUKI DESIGN	2	2010-12-20 10:09:00	0.85	14646	Netherlands

retail 12 ×

4. Identify the country with the highest number of unique customers.

```
67      -- 4. Identify the country with the highest number of unique customers.
68  •   Select Country,
69              Count(distinct CustomerID) as unique_customer
70      from retail
71      group by Country
72      order by unique_customer desc
73      Limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Country	unique_customer
United Kingdom	4084

5. Find the customer who made the maximum number of transactions

```
75      -- 5. Find the customer who made the maximum number of transactions
76      Select CustomerID,
77          Count(distinct InvoiceNo) as Transactions
78      from retail
79      group by CustomerID
80      order by Transactions desc
81      limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

CustomerID	Transactions
14911	279

Product-Based Analysis

1. List the top 5 most frequently purchased products (based on total quantity sold).

```
85      -- 1. List the top 5 most frequently purchased products (based on total quantity sold).
86      Select Stockcode,
87          sum(Quantity) as total_quantity
88      from retail
89      group by stockcode
90      order by total_quantity desc
91      limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Stockcode	total_quantity
22197	56450
84077	53847
85099B	47363
85123A	39122
84879	36221

2. Find the product that generated the highest revenue

```
93      -- 2. Find the product that generated the highest revenue
94      Select stockcode,
95          sum(Quantity * UnitPrice) as total_revenue
96      from retail
97      group by stockcode
98      order by total_revenue desc
99      limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

stockcode	total_revenue
DOT	206245.48028597236

3. Identify products that have been sold in exactly 10 or more different invoices.

```
101     -- 3. Identify products that have been sold in exactly 10 or more different invoices.
102     Select stockcode,
103         Count(distinct InvoiceNo) as unique_invoices
104     from retail
105     group by stockcode
106     Having unique_invoices >= 10
107     order by unique_invoices;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

stockcode	unique_invoices
16162L	10
16162M	10
17012E	10
17090A	10
20778	10

4. Count how many times each product has been sold and list those that have been purchased more than 5 times

```
109     -- 4. Count how many times each product has been sold and list those that have been purchased more than 5 times
110     Select Stockcode,
111         Count(distinct InvoiceNo) as unique_invoices
112     from retail
113     group by stockcode
114     having unique_invoices > 5
115     order by unique_invoices;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Stockcode	unique_invoices
17109D	6
20785	6
20793	6
20892	6
21655	6

Result 20

5. Retrieve all distinct product descriptions purchased by a specific customer (Customer ID = 17850).

```
117     -- 5. Retrieve all distinct product descriptions purchased by a specific customer (CustomerID = 17850).
118     Select distinct description
119     from retail
120     where CustomerID = 17850;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

description
WHITE HANGING HEART T-LIGHT HOLDER
WHITE METAL LANTERN
CREAM CUPID HEARTS COAT HANGER
KNITTED UNION FLAG HOT WATER BOTTLE
RED WOOLLY HOTTIE WHITE HEART.

etail 21

Time-Based Analysis

1. Find the total revenue generated per month.

```
122     --  Time-Based Analysis
123
124     -- 1. Find the total revenue generated per month
125     Select
126         Year(InvoiceDate) as invoice_year,
127         Month(InvoiceDate) as invoice_month,
128         sum(Quantity * UnitPrice) as total_revenue
129     from retail
130     group by invoice_year, invoice_month;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

invoice_year	invoice_month	total_revenue
2010	12	748957.0380319059
2011	1	560000.2885112464
2011	2	498062.648327291
2011	3	683267.0775260255
2011	4	493207.1188206653

Result 24

2. Identify the hour of the day when the highest number of transactions occurred.

```
132     -- 2. Identify the hour of the day when the highest number of transactions occurred.
133     Select
134         hour(InvoiceDate) as invoice_hour,
135         Count(Distinct InvoiceNO) as transactions
136     from retail
137     group by invoice_hour
138     order by transactions desc
139     limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

invoice_hour	transactions
12	3962

3. Count the number of invoices generated per day

```
141     -- 3. Count the number of invoices generated per day
142     Select
143         Date(InvoiceDate) as invoice_date,
144         Count(distinct InvoiceNo) as num_invoices
145     from retail
146     group by invoice_date;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

invoice_date	num_invoices
2010-12-01	143
2010-12-02	167
2010-12-03	108
2010-12-05	95
2010-12-06	133

Result 30

4. Identify the date when the highest number of products were sold.

```
148     -- 4. Identify the date when the highest number of products were sold.
149     Select
150         Date(InvoiceDate) as invoice_date,
151         sum(quantity) as total_qty
152     from retail
153     group by invoice_date
154     order by total_qty desc
155     limit 1;
156
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

invoice_date	total_qty
2011-10-05	46161

5. Find the number of transactions that happened before 12 PM vs. after 12 PM

```
157     -- 5. Find the number of transactions that happened before 12 PM vs. after 12 PM
158     Select
159         case
160             when hour(InvoiceDate) < 12 then "Before 12 PM"
161             else "After 12 PM"
162         end as Timeperiod,
163         Count(distinct InvoiceNo) as transactions
164     from retail
165     group by Timeperiod;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Timeperiod	transactions
After 12 PM	17274
Before 12 PM	8627