

BIG QUERY PRACTICAL EXERCISE – TSHIKOSI TSHIFHIWA

QUESTION 1

The screenshot shows the Google Cloud BigQuery console interface. The left sidebar contains navigation options: Overview, Studio, Pipelines & Integration, Data transfers, Dataform, Scheduled queries, Scheduling, Governance, Sharing (Analytics Hub), Policy tags, Metadata curation, Administration, Monitoring, Jobs explorer, Partner Center, Settings, and Release Notes. The main area displays an 'Untitled query' with the following SQL code:

```
1 SELECT * FROM `practical-exercise-478516.retail.sales` LIMIT 1000;
2 -----
3 --Q1. Filter all transactions that occurred in the year 2023. Expected output: All columns
4 SELECT *
5 FROM `practical-exercise-478516.retail.sales`
6 WHERE EXTRACT (YEAR FROM DATE) = 2023;
```

Below the query, it states: 'This script will process 145.38 KB when run. Using on-demand processing quota'. The 'Query results' section shows a table with 7 rows and 8 columns: Transaction ID, Date, Customer ID, Gender, Age, Product Category, and Quantity. The data is as follows:

Row	Transaction ID	Date	Customer ID	Gender	Age	Product Category	Quantity
1	191	2023-10-18	CUST191	Male	64	Beauty	
2	204	2023-09-28	CUST204	Male	39	Beauty	
3	230	2023-04-23	CUST230	Male	54	Beauty	
4	232	2023-02-06	CUST232	Female	43	Beauty	
5	309	2023-12-23	CUST309	Female	26	Beauty	
6	310	2023-10-12	CUST310	Female	28	Beauty	
7	363	2023-06-03	CUST363	Male	64	Beauty	

At the bottom, it indicates 'Results per page: 50' and '1 - 50 of 998'.

QUESTION 2

The screenshot shows the Google Cloud BigQuery console interface. The left sidebar contains navigation options: Explorer, Home, Starred, Shared with me, Job history, and practical-exercise-478516. The main area displays an 'Untitled query' with the following SQL code:

```
1 --Q2. Display all transactions where the Total Amount is more than the average Total Amount of the entire dataset. Expected output: All
2 columns
3 SELECT *
4 FROM `practical-exercise-478516.retail.sales` AS sales
5 WHERE
6 sales.Total Amount > (
7 SELECT
8 AVG(sales_avg.Total Amount)
9 FROM
10 `practical-exercise-478516.retail.sales` AS sales_avg )
11 ORDER BY
12 sales.Date;
```

Below the query, it states: 'This query will process 72.69 KB when run. Using on-demand processing quota'. The 'Query results' section shows a table with 2 rows and 8 columns: Transaction ID, Date, Customer ID, Gender, Age, Product Category, and Quantity. The data is as follows:

Row	Transaction ID	Date	Customer ID	Gender	Age	Product Category	Quantity
1	180	2023-01-01	CUST180	Male	41	Clothing	
2	559	2023-01-01	CUST559	Female	40	Clothing	

At the bottom, it indicates 'Results per page: 50' and '1 - 50 of 350'.

QUESTION 3

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
12 sales.Date;  
13  
14 --Q3. Calculate the total revenue (sum of Total Amount). Expected output: Total_Revenue  
15 SELECT SUM (sales.'Total Amount') AS TOTAL_REVENUE  
16 FROM 'practical-exercise-478516'.'retail'.'sales';
```

Below the query, the 'Query results' section shows a table with one row:

Row	TOTAL_REVENUE
1	456000

The interface also includes a sidebar with 'Explorer' and 'Job history' tabs, and a top navigation bar with 'Google Cloud' and 'practical exercise' labels.

QUESTION 4

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
18 --Q4. Display all distinct Product Categories in the dataset. Expected output: Product_Category  
19 SELECT DISTINCT sales.'product category'  
20 FROM 'practical-exercise-478516'.'retail'.'sales';
```

Below the query, the 'Query results' section shows a table with three rows:

Row	product category
1	Beauty
2	Clothing
3	Electronics

The interface also includes a sidebar with 'Explorer' and 'Job history' tabs, and a top navigation bar with 'Google Cloud' and 'practical exercise' labels.

QUESTION 5

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
--Q5. For each Product Category, calculate the total quantity sold. Expected output: Product_Category, Total_Quantity
SELECT sales.`product category`,
       SUM(sales.quantity) AS TOTAL_QUANTITY_SOLD
FROM `practical-exercise-478516`.`retail`.`sales`
GROUP BY sales.`product category`;
```

Below the query, a message states: "This script will process 108.65 KB when run. Using on-demand processing quota".

The 'Query results' section shows a table with the following data:

Row	product category	TOTAL_QUANTITY_SOLD
1	Beauty	771
2	Clothing	894
3	Electronics	849

The interface includes a sidebar with navigation options like Explorer, Home, Starred, Shared with me, Job history, and practical-exercise-478516. The top navigation bar shows the Google Cloud logo and a search bar.

QUESTION 6

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
--Q6. Create a column called Age_Group that classifies customers as 'Youth' (<30), 'Adult' (30-59), and 'Senior' (60+). Expected output:
Customer_ID, Age, Age_Group
SELECT
  sales.`customer id`,
  sales.age,
  CASE
    WHEN sales.age < 30 THEN 'Youth'
    WHEN sales.age BETWEEN 30 AND 59 THEN 'Adult'
    WHEN sales.age >= 60 THEN 'Senior'
  END AS Age_Group
FROM `practical-exercise-478516`.`retail`.`sales`;
```

Below the query, a message states: "This query will process 16.6 KB when run. Using on-demand processing quota".

The 'Query results' section shows a table with the following data:

Row	customer id	age	Age_Group
1	CUST191	64	Senior
2	CUST204	39	Adult
3	CUST230	54	Adult
4	CUST232	43	Adult

The interface includes a sidebar with navigation options like Explorer, Home, Starred, Shared with me, Job history, and practical-exercise-478516. The top navigation bar shows the Google Cloud logo and a search bar.

QUESTION 7

The screenshot shows the Google Cloud Console interface. The main area displays a BigQuery query titled "Untitled query". The query is as follows:

```
--Q7. For each Gender, count how many high-value transactions occurred (where Total Amount > 500). Expected output: Gender, High_Value_Transactions
SELECT sales.`gender`,
       COUNT(*) AS high_value_transactions
FROM `practical-exercise-478516`.retail.sales
WHERE sales.`total amount` > 500
GROUP BY sales.`gender`;
```

Below the query, the "Query results" section shows a table with 2 rows:

Row	gender	high_value_transactions
1	Female	155
2	Male	144

The interface also shows a sidebar with navigation options like Explorer, Home, Starred, Shared with me, Job history, and a search bar. The top navigation bar includes the Google Cloud logo and a search bar.

QUESTION 8

The screenshot shows the Google Cloud Console interface. The main area displays a BigQuery query titled "Untitled query". The query is as follows:

```
--Q8. For each Product Category, show only those categories where the total revenue exceeds 5,000. Expected output: Product_Category, Total_Revenue
SELECT sales.`Product Category`,
       SUM(sales.`total amount`) AS total_revenue
FROM `practical-exercise-478516`.retail.sales
GROUP BY sales.`Product Category`
HAVING SUM(sales.`total amount`) > 5000;
```

Below the query, the "Query results" section shows a table with 3 rows:

Row	Product Category	total_revenue
1	Beauty	143515
2	Clothing	155580
3	Electronics	156905

The interface also shows a sidebar with navigation options like Explorer, Home, Starred, Shared with me, Job history, and a search bar. The top navigation bar includes the Google Cloud logo and a search bar.

QUESTION 9

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
--Q9. Display a new column called Unit_Cost_Category that labels a transaction as: 'Cheap' if Price per Unit < 50 'Moderate' if Price per Unit between 50 and 200 'Expensive' if Price per Unit > 200. Expected output: Transaction_ID, Price_per_Unit, Unit_Cost_Category

SELECT sales.Transaction_ID,
       sales.Price_per_Unit,
       CASE
         WHEN sales.Price_per_Unit < 50 THEN 'Cheap'
         WHEN sales.Price_per_Unit BETWEEN 50 AND 200 THEN 'Moderate'
         WHEN sales.Price_per_Unit > 200 THEN 'Expensive'
         ELSE 'Unknown'
       END AS Unit_Cost_Category
FROM `practical-exercise-478516`.`retail`.`sales` AS sales;
```

Below the query, the 'Query results' section shows a table with 4 rows of data:

Row	Transaction ID	Price per Unit	Unit_Cost_Category
1	191	25	Cheap
2	204	25	Cheap
3	230	25	Cheap
4	232	25	Cheap

The interface also shows a sidebar with navigation options and a top bar with search and account information.

QUESTION 10

The screenshot shows the Google Cloud console interface. The main area displays an 'Untitled query' with the following SQL code:

```
otherwise 'Low'. Expected output: Customer_ID, Age, Total_Amount, Spending_Level

SELECT sales.Customer_ID,
       sales.Age,
       sales.Total_Amount,
       CASE
         WHEN sales.Total_Amount > 1000 THEN 'High'
         ELSE 'Low'
       END AS Spending_Level
FROM `practical-exercise-478516`.`retail`.`sales`
WHERE sales.Age >= 40
ORDER BY sales.Customer_ID;
```

Below the query, the 'Query results' section shows a table with 3 rows of data:

Row	Customer ID	Age	Total Amount	Spending_Level
1	CUST003	50	30	Low
2	CUST006	45	30	Low
3	CUST007	46	50	Low

The interface also shows a sidebar with navigation options and a top bar with search and account information.