

### **Task 3 at Elevate Labs**

**Name of the Intern: Nikhil Kumar Nigam**

**Platform Used: Bigquery**

**Dataset Name: E-Commerce Data from Keggle**

## 1. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

⌂ X Keggle\_...rce X Q \*Untitled...ery X + ▾

Keggle\_Ecom... Query Open in ▾ + Share Copy Snapshot Delete Export ↻

Schema Details Preview Table explorer Preview Insights Lineage Data profile Data Quality

Filter Enter property name or value ?

<input type="checkbox"/>	Field name	Type	Mode	Key	Collation	Default value	Policy tags ?	Description
<input type="checkbox"/>	InvoiceNo	STRING	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	StockCode	STRING	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	Description	STRING	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	Quantity	INTEGER	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	InvoiceDate	TIMESTAMP	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	UnitPrice	FLOAT	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	CustomerID	INTEGER	NULLABLE	-	-	-	-	-
<input type="checkbox"/>	Country	STRING	NULLABLE	-	-	-	-	-

## 2. Preview of the dataset:

```
select * from `Keggle_Ecommerce.Keggle_Ecommerce`
```

### Query results

Save results ▾

Job information		Results	Chart	JSON	Execution details	Execution graph		
Row	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
1	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00 UTC	2.55	17850	United Kingdom
2	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00 UTC	3.39	17850	United Kingdom
3	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00 UTC	2.75	17850	United Kingdom
4	536365	84029G	KNITTED UNION FLAG HOT WATER BOT...	6	2010-12-01 08:26:00 UTC	3.39	17850	United Kingdom
5	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00 UTC	3.39	17850	United Kingdom
6	536365	22752	SET 7 BABUSHKA NESTING BOXES	2	2010-12-01 08:26:00 UTC	7.65	17850	United Kingdom
7	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	2010-12-01 08:26:00 UTC	4.25	17850	United Kingdom
8	536366	22633	HAND WARMER UNION JACK	6	2010-12-01 08:28:00 UTC	1.85	17850	United Kingdom

### 3. Total Revenue

```
SELECT SUM(Quantity * UnitPrice) AS total_revenue  
from `Keggle_Ecommerce.Keggle_Ecommerce`
```

Query results		
Job information		
Results		
Chart		
JSON		
Execution details		
Row	total_revenue	
1	9747747.93	

#### Inference:

This is the total earnings from all sales recorded in the dataset. It gives a high-level metric to assess the financial scale of the business.

#### 4. Top 10 Selling Products

```
SELECT
  Description,
  SUM(Quantity) AS total_sold
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
GROUP BY Description
ORDER BY total_sold DESC
LIMIT 10;
```

##### Query results

Job information		Results	Chart	JSON	Execution details
Row	Description	total_sold			
1	WORLD WAR 2 GLIDERS ASSTD DESIGNS	53847			
2	JUMBO BAG RED RETROSPOT	47363			
3	ASSORTED COLOUR BIRD ORNAMENT	36381			
4	POPCORN HOLDER	36334			
5	PACK OF 72 RETROSPOT CAKE CASES	36039			
6	WHITE HANGING HEART T-LIGHT HOLDER	35317			
7	RABBIT NIGHT LIGHT	30680			
8	MINI PAINT SET VINTAGE	26437			
9	PACK OF 12 LONDON TISSUES	26315			
10	PACK OF 60 PINK PAISLEY CAKE CASES	24753			

##### Inference:

These are the most popular products based on volume. The company can prioritize these items for promotions, stock management, or bundling strategies.

## 5. Monthly Sales Trend

```
SELECT
    FORMAT_TIMESTAMP('%Y-%m', InvoiceDate) AS month,
    ROUND(SUM(Quantity * UnitPrice),2) AS monthly_revenue
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
GROUP BY month
ORDER BY month;
```

Query results			
Job information		Results	Chart
JSON			
Row	month	monthly_revenue	
1	2010-12	748957.02	
2	2011-01	560000.26	
3	2011-02	498062.65	
4	2011-03	683267.08	
5	2011-04	493207.12	
6	2011-05	723333.51	
7	2011-06	691123.12	
8	2011-07	681300.11	
9	2011-08	682680.51	
10	2011-09	1019687.62	

### Inference:

Query result helps identify **seasonal patterns**, periods of growth or decline, and the impact of campaigns. It's essential for forecasting and trend analysis.

## 6. Top Customers by Spend

```
SELECT
    CustomerID,
    ROUND(SUM(Quantity * UnitPrice), 2) AS total_spent
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
WHERE CustomerID IS NOT NULL
GROUP BY CustomerID
ORDER BY total_spent DESC
LIMIT 10;
```

### Query results

Job information		Results	Chart
Row	CustomerID	total_spent	
1	14646	279489.02	
2	18102	256438.49	
3	17450	187482.17	
4	14911	132572.62	
5	12415	123725.45	
6	14156	113384.14	
7	17511	88125.38	
8	16684	65892.08	
9	13694	62653.1	
10	15311	59419.34	

### Inference:

These customers contribute significantly to total revenue. The company can target them with loyalty programs or exclusive offers to retain them.

## 7. Revenue by Country

```
SELECT
  Country,
  ROUND(SUM(Quantity * UnitPrice), 2) AS total_revenue
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
GROUP BY Country
ORDER BY total_revenue DESC;
```

Query results		
Job information		
Results		
Chart		
JSON		
Row	Country	total_revenue
1	United Kingdom	8187806.36
2	Netherlands	284661.54
3	EIRE	263276.82
4	Germany	221698.21
5	France	197403.9
6	Australia	137077.27
7	Switzerland	56385.35
8	Spain	54774.58
9	Belgium	40910.96
10	Sweden	36595.91

### Inference:

Useful for **geographical market analysis**. It highlights top-performing regions and potential areas for expansion or marketing focus. **United Kingdom** dominates in revenue, suggesting it's the primary market. Countries like **Netherlands**, **EIRE**, **Germany**, and the **France** follow. These markets show potential for expanding logistics or marketing.

## 8. Peak Sales Periods

```
SELECT
  FORMAT_TIMESTAMP('%A', InvoiceDate) AS day_of_week,
  COUNT(DISTINCT InvoiceNo) AS total_orders,
  ROUND(SUM(Quantity * UnitPrice), 2) AS total_revenue
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
GROUP BY day_of_week
ORDER BY total_revenue DESC;
```

### Query results

Job information		Results	Chart	JSON	Execution details
Row	day_of_week ▼	total_orders ▼	total_revenue ▼		
1	Thursday	5660	2112519.0		
2	Tuesday	4722	1966182.79		
3	Wednesday	4815	1734147.01		
4	Monday	4138	1588609.43		
5	Friday	4184	1540610.81		
6	Sunday	2381	805678.89		

### Inference:

Sales peak midweek, typically **Tuesday to Thursday**, which may align with B2B bulk buyers placing orders. Fewer sales on weekends suggest lower consumer traffic, pointing to a potential B2B-centric model or weekday promotions.



## 9. Revenue and order distribution By Hour of the Day

```
SELECT
    EXTRACT(HOUR FROM InvoiceDate) AS hour_of_day,
    COUNT(DISTINCT InvoiceNo) AS total_orders,
    ROUND(SUM(Quantity * UnitPrice), 2) AS total_revenue
FROM `Keggle_Ecommerce.Keggle_Ecommerce`
GROUP BY hour_of_day
ORDER BY hour_of_day;
```

Query results				
Job information	Results	Chart	JSON	Exec
Row	hour_of_day	total_orders	total_revenue	
1	6	22	-497.35	
2	7	31	31009.32	
3	8	624	281840.86	
4	9	1824	766734.05	
5	10	2961	1329056.52	
6	11	3165	1147437.92	
7	12	3962	1362484.29	
8	13	3369	1177506.37	
9	14	3137	1095212.9	
10	15	3069	1189458.28	
11	16	1952	729140.82	
12	17	1205	435444.11	
13	18	333	140574.48	
14	19	219	46324.99	
15	20	28	16020.37	

### Inference:

Peak activity occurs between **9 AM and 3 PM**, aligning with standard working hours. This reinforces the hypothesis of business orders. Knowing this helps schedule server loads, customer service hours, or ad placements.