

Business Intelligence Analyst

Virtual Internship Experience

By Rasyidah Maulida Putri Andini



Rasyidah Maulida Putri Andini

About Me

An active student in both internal and external to campus, with a high interest in Data Analyst, BI Analyst, Machine Learning, and Data Science. Often performs the data mining process by analyzing data, representing data, preprocessing data, making predictions. Excellent in understanding business operations, various data analytics tools (Python, SQL, Pandas, Numpy, Matplotlib, Seaborn, Sklearn, BigQuery) and dashboarding using Google Data Studio. Experienced with organizations and being a leader on several projects, good analytical thinking and problem solver, detail-oriented, responsive, good performance, and communication skills, eager to learn something new and challenging things, able to work in a team and adapt quickly to a new environment.



Case Study

Process the raw data from end to end with the output in the dashboard version.








Data Gathering

Link Dataset: <https://drive.google.com/file/d/1RwsBQ1FriNfz6qiq0V5nD7gF7jO81To3/view?usp=sharing>

Dataset Task 5.rar 4 item

Name	Terakhir diubah	Ukuran file
 Customers.xlsx	17 Apr 2022	217 KB
 Orders.xlsx	17 Apr 2022	108 KB
 ProductCategory.xlsx	17 Apr 2022	9 KB
 Products.xlsx	17 Apr 2022	11 KB

▼  Final_Task ☆ ⋮

 Customers	☆	⋮
 Orders	☆	⋮
 ProductCategory	☆	⋮
 Products	☆	⋮

1. The dataset was downloaded from the Google Drive link and then saved to a local file.
2. Next, the dataset was uploaded to Google BigQuery for further processing.
3. The dataset consists of four tables and is ready for further analysis and manipulation.



Data Understanding

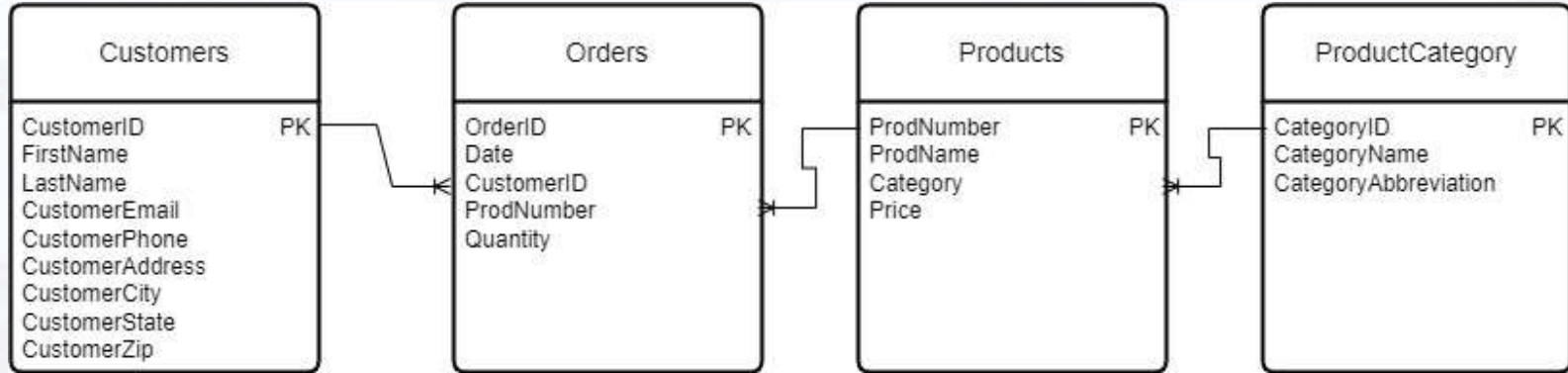
Primary Key

A primary key is a column that uniquely identifies each row of data in a table. It can distinguish one row from another and is often used as a reference for other tables. The primary key value in each row must be unique, and there must be no duplicates. This primary key ensures that the data in the table is organized in a structured and efficient way.

Determine each primary key in the 4 sales tables.

1. Primary key table Customer : CustomerID
2. Primary key table Products : ProdNumber
3. Primary key table Orders : OrderID
4. Primary key table ProductCategory : CategoryID

Relationship Table



1. Table Customer.CustomerID = Orders.CustomerID => One to Many relationship
2. Table Products.ProdNumber = Orders.ProdNumber => One to Many relationship
3. Table ProductCategory.CategoryID = Products.Category => One to Many relationship

One-to-many means that one data in a table can be associated with one or more data in another. For example, each customer can have many sales transactions.



Data Processing

Create Detail Transaction Table

Row	order_date ▾	category_name ▾	product_name ▾	product_price ▾	order_qty ▾	total_sales ▾	cust_email ▾	cust_city
1	2020-01-01	Drone Kits	BYOD-220	69.0	1	69.0	edew@nba.com#mailto:edew...	Honolulu
2	2020-01-01	eBooks	Polar Robots	23.99	2	47.98	fvaslerqt@comsenz.com#mailt...	Jackson
3	2020-01-01	Robots	RWW-75 Robot	883.0	3	2649.0	tmckernot@tinyurl.com#mailto...	Katy
4	2020-01-01	eBooks	SCARA Robots	19.5	5	97.5	llespercx@com.com#mailto:lle...	Des Moines

Results per page: 50 ▾ 1 – 50 of 3339 |< < > >|

Master_Table  RUN  SAVE QUERY ▾  SHARE ▾  SCHEDULE  MORE ▾

```
1 SELECT
2     o.date AS order_date,
3     pc.categoryname AS category_name,
4     p.prodname AS product_name,
5     p.price AS product_price,
6     o.quantity AS order_qty,
7     (p.price * o.quantity) AS total_sales,
8     c.customeremail AS cust_email,
9     c.customercity AS cust_city
10 FROM
11     'Final_Task.Customers' AS c
12 JOIN
13     'Final_Task.Orders' AS o ON c.customerid = o.customerid
14 JOIN
15     'Final_Task.Products' AS p ON o.prodnumber = p.prodnumber
16 JOIN
17     'Final_Task.ProductCategory' AS pc ON p.category = pc.categoryid
18 ORDER BY o.date, o.quantity;
```

Link Query:

<https://console.cloud.google.com/bigquery?sq=3695549098:faaabfcbf833425f8dfccb937f8b0383>



Data Analyzing & visualization

Data Visualization

Link Dashboard:

<https://lookerstudio.google.com/reporting/aeaf8e1e-4dca-4b29-8414-415f90e1b15d>

Total Sales:

1,754,750.57

Total Order Quantity:

11,654

Total Sales by Product Category

	Product Category ▾	Total Sales
1.	Blueprints	16,434.51
2.	Drone Kits	161,242.5
3.	Drones	477,447
4.	Robot Kits	216,437
5.	Robots	763,506
6.	Training Videos	80,716.15
7.	eBooks	58,968.41

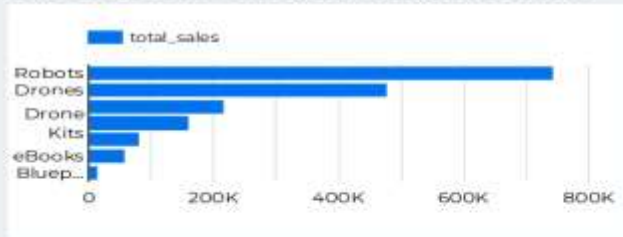
1 - 7 / 7 < >

Total Order Quantity by Product Category

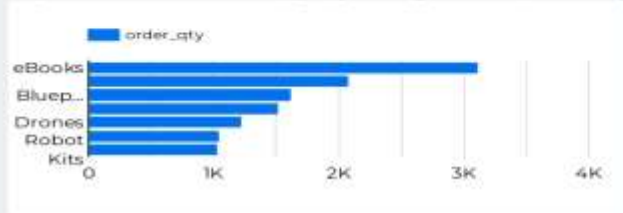
	product_name ▾	Order Quantity
1.	Virtual Reality Basics	132
2.	Understanding Raspberry Pi	159
3.	Understanding Drone Regulatio...	143
4.	Understanding Automation	183
5.	Understanding Artificial Intellige...	151
6.	Understanding Arduino	152
7.	Understanding 3D Printing	169

1 - 69 / 69 < >

Top 5 Product Category : Highest Sales



Top 5 Product Category : Highest Quantity



Total Sales by City

	City	Total Sales ▾
1.	Washington	55,381.94
2.	Houston	33,761.49
3.	Sacramento	33,380.2
4.	San Diego	29,228.59
5.	Albany	25,405.88

1 - 100 / 361 < >

Total Order Quantity by City

	City	Order Quantity ▾
1.	Washington	308
2.	Houston	249
3.	San Diego	203
4.	Sacramento	153

1 - 100 / 361 < >

Insights

- The total sales amount is 1.75 million dollars. The "Robots" product category contributed the most to sales, totaling 743.5k dollars, representing 42% of all sales. Regarding the highest number of orders, "eBooks" was the most popular product category, with 3.12k orders.
- There were five product categories with the highest sales, which are "Robots," "Drones," "Robot Kits," "Drone Kits," and "Training Videos." On the other hand, the five product categories with the highest total orders are "eBooks," "Training Videos," "Blueprints," "Drone Kits," and "Drones."
- Washington state generated the highest sales and total orders, with 55k dollars and 308 orders.

Suggestions for Sales Improvement

- Increase stock inventory for the product categories with the highest sales and total orders. This will help meet the high demand for these products in the future.
- Adjust the customer segmentation for marketing campaigns by targeting cities with high sales and total order values.
- Analyze the data of cities with low sales to identify the reason behind their low sales and total orders. This will help in finding solutions to improve the sales in those cities.

Thank You

