

# Business Intelligence

PEMBANGUNAN PROYEK BUSINESS INTELLIGENCE PADA E-COMMERCE ABCD UNTUK  
MENGANALISIS TREN PENJUALAN DAN PERILAKU PELANGGAN

---

TUGAS BESAR  
Assignment 2

GROUP  
Kelompok 9

KELAS  
Business Intelligence A

# Topics

---

01

## Information Architecture

WHAT  
WHO  
WHEN  
WHERE  
WHY  
HOW

---

02

## Information Requirements

---

03

## Data Architecture

DATA WORKFLOW  
DATA WAREHOUSE DESIGN  
ETL DESIGN

---

04

## Technique & Product Architecture

---

# Members Team



---

**Nadia Nur Saida**

2111521007



---

**Karin Oktariani E**

2111522019



---

**Ahmad Ali Zulfiqar**

2111522007

# Information Architecture

WHAT	WHO	WHEN
<ul style="list-style-type: none"><li>1. <b>What business process or functions are going to be supported?</b> → Proses bisnis penjualan</li><li>2. <b>What types of analytics will be needed?</b> → Analisis tren penjualan, segmentasi pelanggan, dan memprediksi tren penjualan di masa depan</li><li>3. <b>What types of decision are affected?</b> → Keputusan strategis display stok produk, dan strategi pemasaran bisnis</li></ul>	<ul style="list-style-type: none"><li>1. <b>Who will have access?</b> → Stakeholder khususnya seller pada e-commerce, serta tim pemasaran dan analytics data pada e-commerce</li></ul>	<ul style="list-style-type: none"><li>1. <b>Where the data is now?</b> → Data diperoleh dari kaggle dan tersimpan dalam file csv terpisah untuk setiap datasetnya</li><li>2. <b>Where it will be integrated?</b> → Data diintegrasikan dalam alat analisis data menggunakan python</li><li>3. <b>Where it'll be consumed in analytical application?</b> → Data akan digunakan dalam aplikasi analitis, dashboard BI</li></ul>

# Information Architecture

WHEN	WHY	HOW
1.Data E-commerce Order tahun 2017-2018	1.Why the BI Solution(s) will be built-what the business and technical requirements are? → Solusi BI yang akan dibangun untuk mengoptimalkan operasional E-Commerce, meningkatkan pengalaman pelanggan, dan meningkatkan pendapatan dengan memahami tren dan pola dalam data penjualan.	1. Menggunakan metode analisis untuk memahami tren penjualan serta faktor-faktor yang dapat mempengaruhi penjualan produk, dan analisis prediktif untuk memprediksi tren penjualan di masa depan

# Information Requirements

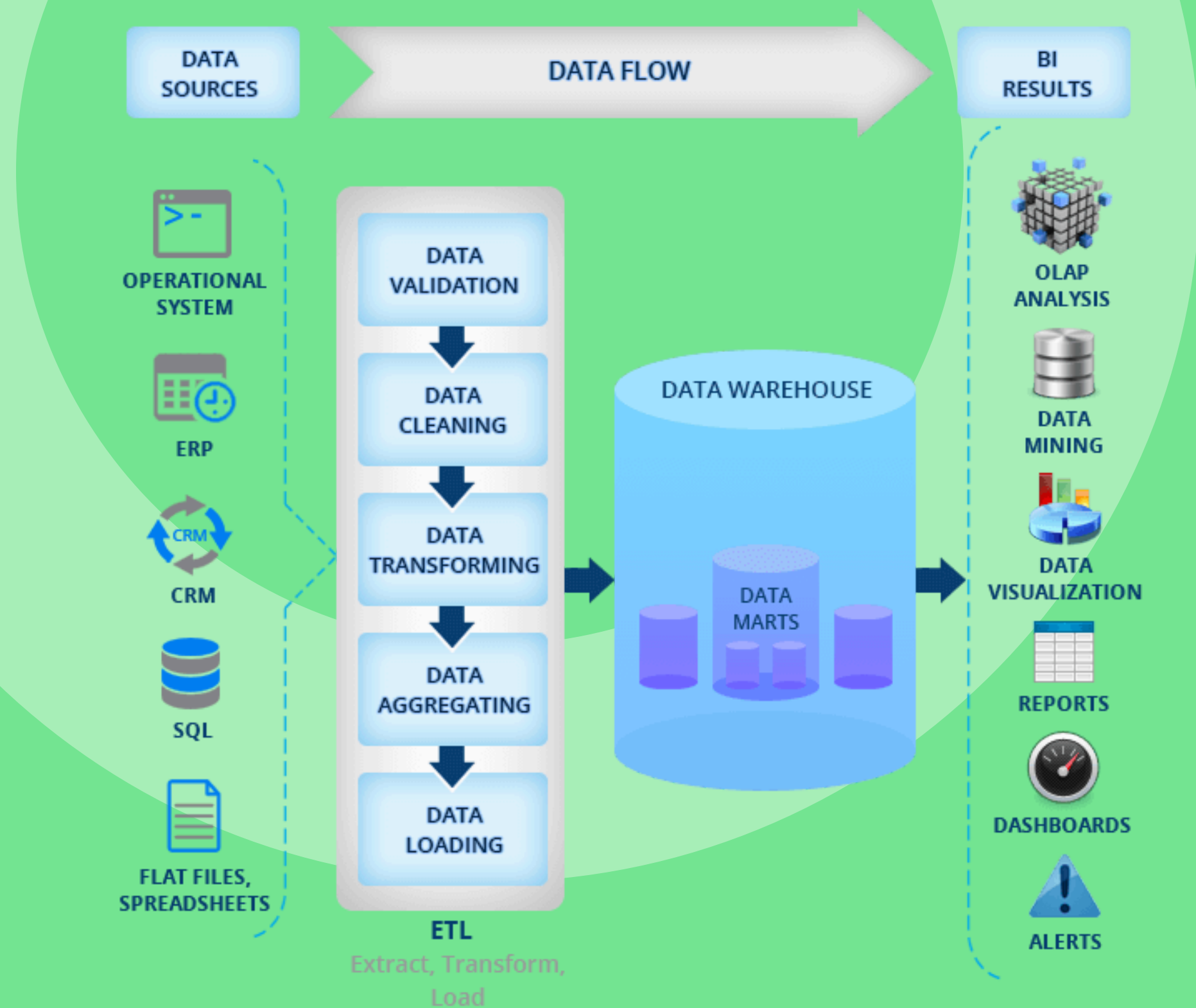
Dalam menganalisis kebutuhan informasi, dilakukan observasi terhadap data yang telah di peroleh. Berikut kebutuhan informasi yang diperlukan:

- Analisis Penjualan dan Tren Produk berdasarkan waktu atau periode penjualan.
  - Informasi Top produk
  - Informasi Top metode pembayaran
  - Informasi rata-rata biaya pengiriman
  - Informasi produk yang di beli beserta qty
  - Informasi tren produk di setiap bulan
  - Informasi produk termahal yang terjual di setiap bulan
- Analisis Segmentasi Pelanggan.
  - Informasi lokasi pelanggan
  - Informasi pelanggan terbanyak di setiap lokasi
  - Informasi riwayat transaksi pelanggan
  - Informasi nilai setiap transaksi pelanggan
  - Informasi jarak/distribusi produk ke pelanggan
  - Informasi perilaku pembelian pelanggan
- Analisis Prediktif menggunakan Forecasting.
  - Menggunakan data historis penjualan untuk menganalisis tren produk dengan pembutan model prediksi. Sehingga dapat dilakukan peramalan tren produk yang mungkin paling banyak disukai pelanggan



# Data Architecture

## DATA WORKFLOW



Sumber data yang digunakan adalah E-commerce Orders merupakan sekumpulan dataset yang berisikan informasi terkait aspek-aspek penjualan yang ada di sebuah situs E-Commerce. Dataset tersebut meliputi, dataset order, dataset customer, dataset payment, dataset, product, dataset geo, dan dataset sellers.

*Data ini bersumber dari kaggle.*

1	order_id	order_item_id	product_id	seller_id	shipping_limit_date	price	freight_value
2	00010242fe8c5a6d1ba2dd792cb16214	1	4244733e06e7ecb4970a6e2683c13e61	48436dade18ac8b2bce089ec2a041202	19/09/2017 09:45	58.90	13.29
3	00018f77f2f0320c557190d7a144bdd3	1	e5fd52b802189ee658865ca93d83a8f	dd7ddc04e1b6c2c614352b383efe2d36	03/05/2017 11:05	239.90	19.93
4	000229ec398224ef6ca0657da4fc703e	1	c777355d18b72b67abbeef9df44fd0fd	5b51032eddd242adc84c38acab88f23d	18/01/2018 14:48	199.00	17.87
5	00024acbcd0a6daa1e931b038114c75	1	7634da152a4610f1595efa32f14722fc	9d7a1d34a5052409006425275ba1c2b4	15/08/2018 10:10	12.99	12.79
6	00042b26cf59d7ce69dfabb4e55b4fd9	1	ac6c3623068f30de03045865e4e10089	df560393f3a51e74553ab94004ba5c87	13/02/2017 13:57	199.90	18.14
7	00048cc3ae777c65dbb7d2a0634bc1ea	1	ef92defde845ab8450f9d70c526ef70f	6426d21aca402a131fc0a5d0960a3c90	23/05/2017 03:55	21.90	12.69
8	00054e8431b9d7675808bcb819fb4a32	1	8d4f2bb7e93e6710a28f34fa83ee7d28	7040e82f899a04d1b434b795a43b4617	14/12/2017 12:10	19.90	11.85
9	000576fe39319847cbb9d288c5617fa6	1	557d850972a7d6f792fd18ae1400d9b6	5996cddab893a4652a15592fb58ab8db	10/07/2018 12:30	810.00	70.75
10	0005a1a1728c9d785b8e2b08b904576c	1	310ae3c140ff94b03219ad0adc3c778f	a416b6a846a11724393025641d4edd5e	26/03/2018 18:31	145.95	11.65
11	0005f50442cb953dcd1d21e1fb923495	1	4535b0e1091c278dfd193e5a1d63b39f	ba143b05f0110f0dc71ad71b4466ce92	06/07/2018 14:10	53.99	11.40
12	00061f2a7bc09da83e415a52dc8a4af1	1	d63c1011f49d98b976c352955b1c4bea	cc419e0650a3c5ba77189a1882b7556a	29/03/2018 22:28	59.99	8.88
13	00063b381e2406b52ad429470734ebd5	1	f177554ea93259a5b282f24e33f65ab6	8602a61d680a10a82cccee0da0d99ea3d	31/07/2018 17:30	45.00	12.98
14	0006ec9db01a64e59a68b2c340bf65a7	1	99a4788cb24856965c36a24e339b6058	4a3ca9315b744ce9f8e9374361493884	26/07/2018 17:24	74.00	23.32
15	0008288aa423d2a3f00fcb17cd7d8719	1	368c6c730842d78016ad823897a372db	1f50f920176fa81dab994f9023523100	21/02/2018 02:55	49.90	13.37
16	0008288aa423d2a3f00fcb17cd7d8719	2	368c6c730842d78016ad823897a372db	1f50f920176fa81dab994f9023523100	21/02/2018 02:55	49.90	13.37
17	0009792311464db532ff765bf7b182ae	1	8cab8abac59158715e0d70a36c807415	530ec6109d11eaaaf87999465c6afee01	17/08/2018 12:15	99.90	27.65
18	0009c9a17f916a706d71784483a5d643	1	3f27ac8e699df3d300ec4a5d8c5cf0b2	fc5ace8bcc92f75707dc0f01a27d269	02/05/2018 09:31	639.00	11.34
19	000aed2e25dbad2f9ddb70584c5a2ded	1	4fa33915031a8cde03dd0d3e8fb27f01	fe2032dab1a61af8794248c8196565c9	16/05/2018 20:57	144.00	8.77
20	000c3e6612759851cc3cbb4b83257986	1	b50c950aba0dcead2c48032a690ce817	218d46b86c1881d022bce9c68a7d4b15	21/08/2017 03:33	99.00	13.71
21	000e562887b1f2006d75e0be9558292e	1	5ed9eaf534f6936b51d0b6c5e4d5c2e9	8cbac7e12637ed9cffa18c7875207478	28/02/2018 12:08	25.00	16.11
22	000e63d38ae8c00bbcb5a30573b99628	1	553e0e7590d3116a072507a3635d2877	1c129092bf23f28a5930387c980c0dfc	29/03/2018 20:07	47.90	8.88
23	000e906b789b55f64edcb1f84030f90d	1	57d79905de06d8897872c551bfd09358	ea8482cd71df3c1969d7b9473ff13abc	27/11/2017 19:09	21.99	11.85
24	000f25f4d72195062c040b12dce9a18a	1	1c05e0964302b6cf68ca0d15f326c6ba	7c67e1448b00f6e969d365cea6b010ab	21/03/2018 11:10	119.99	44.40
25	001021efaa8636c29475e7734483457d	1	5d7c23067ed3fc8c6e699b9373d5890b	6560211a19b47992c3666cc44a7e94c0	05/03/2018 09:35	49.00	15.10
26	0010b2e5201cc5f1ae7e9c6cc8f5bd00	1	5a419dbf24a8c9718fe522b81c69f61a	3504c0cb71d7fa48d967e0e4c94d59d9	15/09/2017 18:04	48.90	16.60
27	00119ff934e539cf26f92b9ef0cdfed8	1	21b1c2f67a9aafb5af0eb06c13b9dbda	c864036feaab8c1659f65ea4faebe1da	11/08/2017 00:35	219.90	16.98
28	0011d82c4b53e22e84023405fb467e57	1	c389f712c4b4510bc997cee93e8b1a28	bfd27a966d91cfaafdb25d076585f0da	29/01/2018 21:51	289.00	26.33
29	00125cb692d04887809806618a2a145f	1	1c0c0093a48f13ba70d0c6b0a9157cb7	41b39e28db005d9731d9d485a83b4c38	29/03/2017 13:05	109.90	25.51
30	00130c0eee84a3d909e75bc08c5c3ca1	1	89321f94e35fc6d7903d36f74e351d40	16090f2ca825584b5a147ab24aa30c86	14/06/2018 05:16	27.90	7.94
31	0013503b13da1eac686219390b7d641b	1	38afdf723b95d455b418a0f57d623c6b	1554a68530182680ad5c8b042c3ab563	13/12/2017 03:16	119.90	17.32



# Data Architecture

## DATA WAREHOUSE DESIGN

### 4-STEPS OF KIMBALL'S TECHNIQUE

1. Select the business process.
2. Declare the grain.
3. Identify the dimension.
4. Identify the facts.





Ruang lingkup proyek ini adalah proses bisnis penjualan pada E-commerce ABCD. Pada proses bisnis penjualan ada proses transaksi penjualan barang yang terjadi antara pelanggan dengan seller pada situs E-commerce.

BUSINESS PROCESS	GRAIN	FACT	FACT MEASURE
Analisis Penjualan – Menganalisis pola pembelian dan perilaku pelanggan	Order Level – Setiap baris mewakili satu transaksi	Order	Subtotal Freight_Cost Total Qty

DIMENSION				FACT			
Tabel Order	Atribut Order	Fact	Fact Attribute & Measure	Tabel Master	Atribut Transaksi	Dimen sion	Dimension Attribute
Order	<ul style="list-style-type: none"><li>• Order_id</li><li>• Customer_id</li><li>• Order_status</li><li>• Order_purchase_timestamp</li><li>• Order_approved_at</li><li>• Order_delivered_carrier_date</li><li>• Order_delivered_customer_date</li><li>• Order_estimated_delivery,</li><li>• Date</li></ul>	Order	<ul style="list-style-type: none"><li>• Order_id</li><li>• Customer_id,</li><li>• Product_id</li><li>• Seller_id</li><li>• Order_date</li><li>• Payment_id</li><li>• Subtotal</li><li>• Freight_cost</li><li>• Total</li><li>• Qty</li></ul>	Customer	<ul style="list-style-type: none"><li>• Customer_id</li><li>• Customer_unique_id</li><li>• Customer_zip_code_prefix</li><li>• Customer_city</li><li>• Customer_state</li></ul>	Dim Customer	<ul style="list-style-type: none"><li>• Customer_id</li><li>• Customer_zip_code</li><li>• Customer_city</li><li>• Customer_state</li></ul>

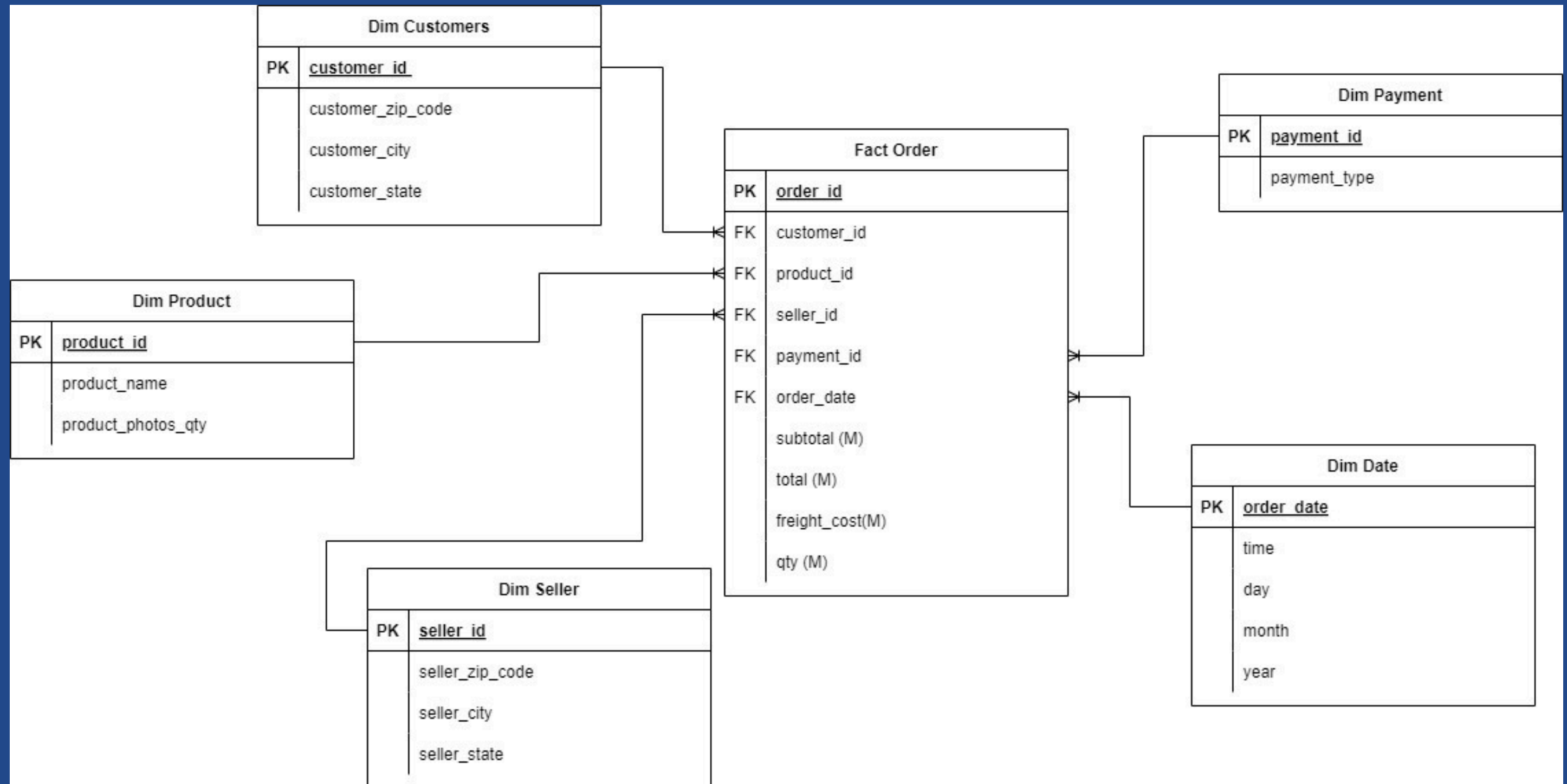
DIMENSION				FACT			
Tabel Order	Atribut Order	Fact	Fact Attribute & Measure	Tabel Master	Atribut Transaksi	Dimension	Dimension Attribute
				Product	<ul style="list-style-type: none"><li>• Product_id</li><li>• Product_category_name</li><li>• Product_name_length</li><li>• Product_description_length</li><li>• Product_photos_qty</li><li>• Product_weight_g</li><li>• Product_length_cm</li><li>• Product_height_cm</li><li>• Product_width_cm</li></ul>	Dim Product	<ul style="list-style-type: none"><li>• Product_id</li><li>• Product_name</li><li>• product_photos_qty</li></ul>

DIMENSION				FACT			
Tabel Order	Atribut Order	Fact	Fact Attribute & Measure	Tabel Master	Atribut Transaksi	Dimension	Dimension Attribute
				Seller	<ul style="list-style-type: none"><li>• Seller_id,</li><li>• Seller_city,</li><li>• Seller_state</li></ul>	Dim Seller	<ul style="list-style-type: none"><li>• Seller_id,</li><li>• Seller_city,</li><li>• Seller_state</li><li>•</li></ul>

DIMENSION				FACT			
Tabel Order	Atribut Order	Fact	Fact Attribute & Measure	Tabel Master	Atribut Transaksi	Dimension	Dimension Attribute
				Dim Payment	<ul style="list-style-type: none"><li>• Payment_id</li><li>• Payment_type</li><li>• Payment_installments</li><li>• Payment_sequential</li></ul>	Dim Payment	<ul style="list-style-type: none"><li>• Payment_id</li><li>• Payment_type</li></ul>
						Dim Date	<ul style="list-style-type: none"><li>• Order_date</li><li>• Time</li><li>• Day</li><li>• Month</li><li>• Year</li></ul>



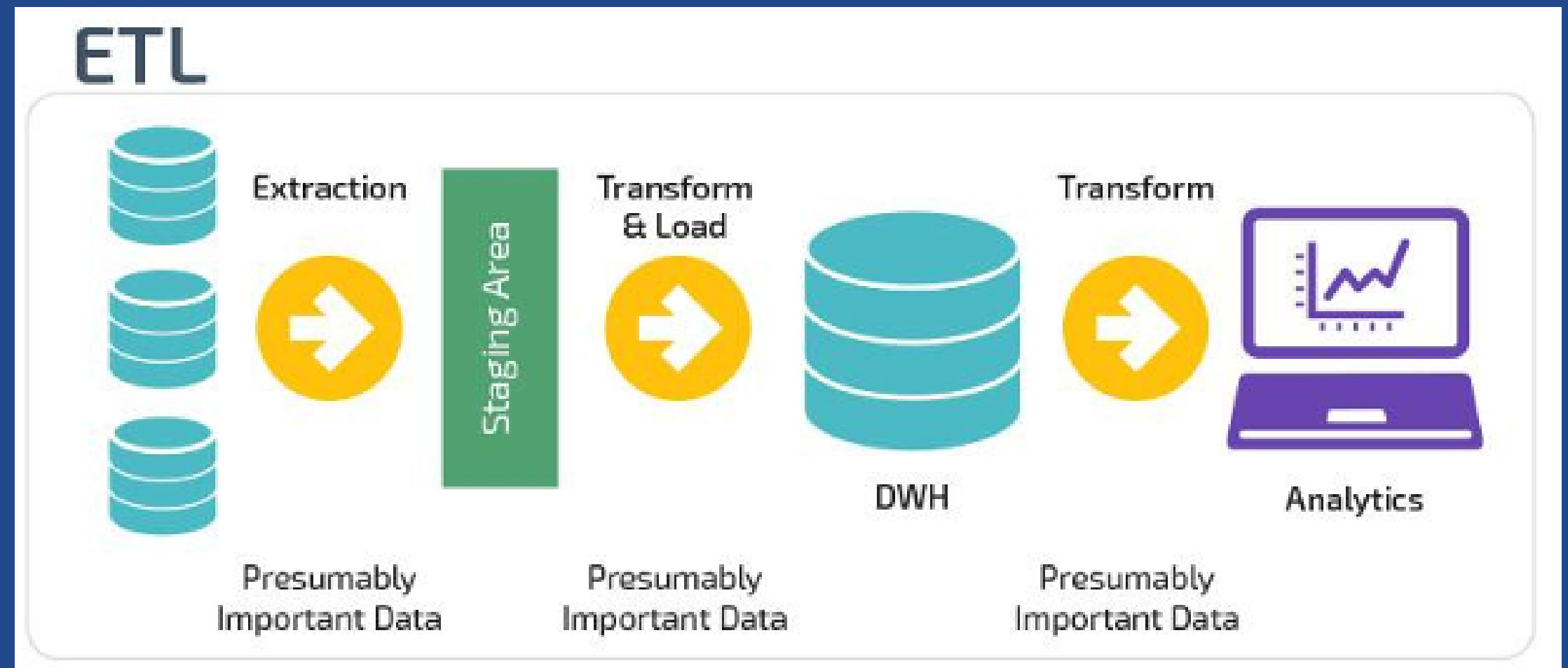
# Schema Data Warehouse



# Data Architecture

## ETL DESIGN

Proses ETL akan dilakukan menggunakan bantuan alat business intelligence bahasa pemrograman python pada jupyter notebook. Data yang di peroleh akan di integrasikan dan disatukan kedalam sebiah satu dataset besar. Kemudian, data akan di cek (noise, null, dll) dan di bersihkan. Data yang telah bersih kemudian akan disimpan kedalam data warehouse untuk di analisis dan dilakukan prediksi



# Tecnique & Product Architecture

01

---

## Predictive Analytics:

Dengan menganalisis data historis, dapat meramalkan tren penjualan di masa depan. Prediksi ini dilakukan berdasarkan variabel-variabel seperti waktu, total biaya pengiriman, dan lokasi.

02

---

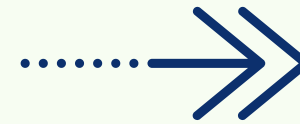
## Data Visualization

Visualisasi data dapat membantu dalam memahami tren dan pola secara intuitif. Dengan menggunakan visualisasi grafik dapat menampilkan kebutuhan informasi yang diperlukan stakeholder.



Microsoft Excel

**Data Source**



Jupyter Notebook



Python

**ETL**

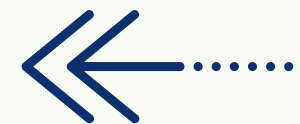


Jupyter Notebook



Power BI

**Business  
Intelligence**



MySQL DB

**Data Warehouse**

The background features a solid green field with two large, overlapping circles. The inner circle is a medium green, and the outer circle is a lighter shade of green. Overlaid on the left side of these circles is the text 'Thank You.' in a bold, dark blue, sans-serif font. The text is split into two lines: 'Thank' on the top line and 'You.' on the bottom line.

# **Thank You.**