

Business Intelligence Analyst

MINI PROJECT

PORTFOLIO

Benedicta Cardalia Ginting

ABOUT ME

A fresh graduate who is interested in data and visualization. Graduated from Food Science and Technology has formed her analytical skill by doing some research that requires data to generate essential and important information from certain conditions. Seek data skills through several courses that result in experience in data analysis and data visualization using Excel, SQL and Tableau. Furthermore, experienced in communication skills, able to work independently as well as in a team, detail oriented with problem solving and critical thinking, and eager to learn and loves challenges.



Benedicta Cardalia Ginting

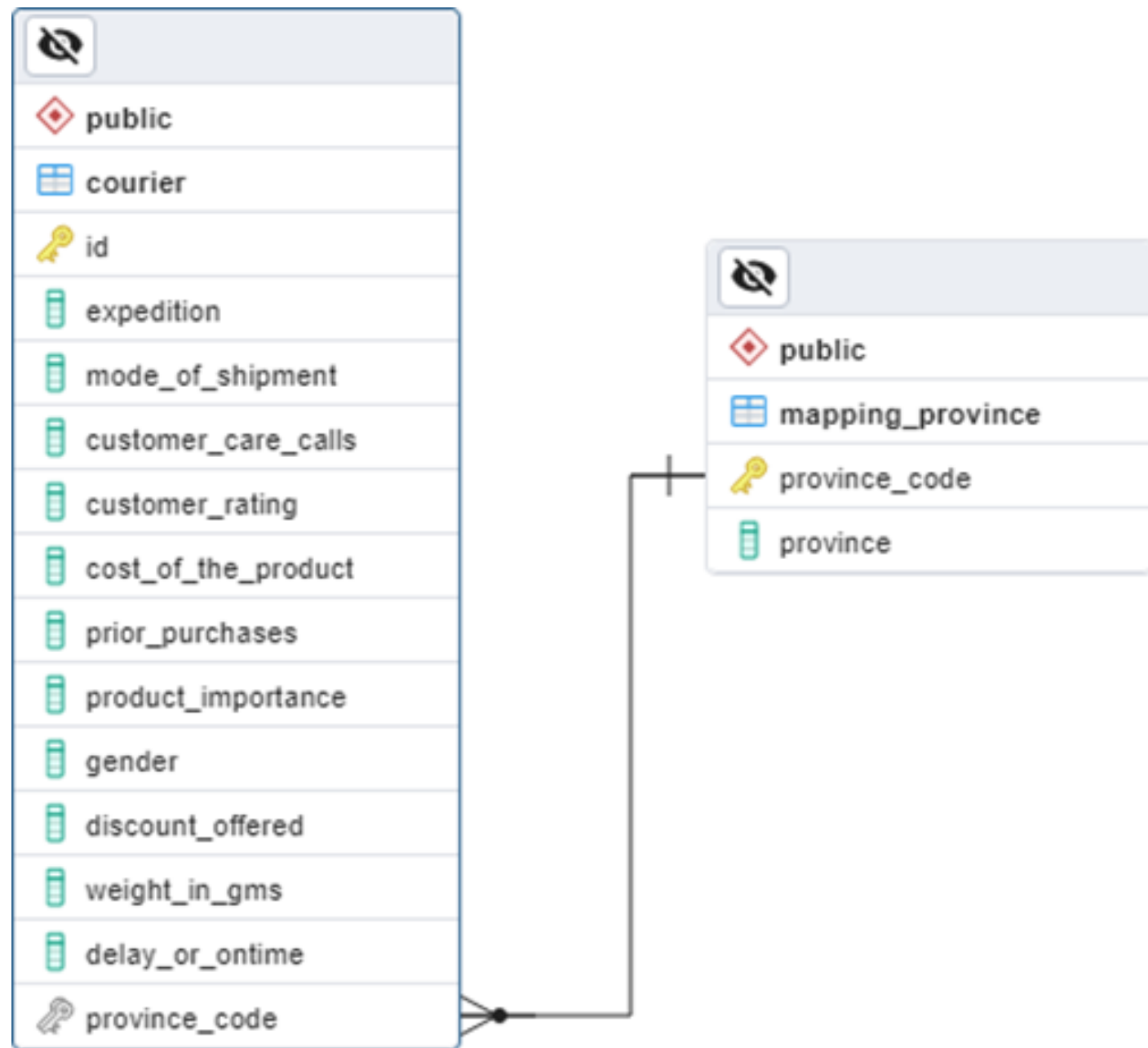
OVERVIEW

"To support a business continuity, both goods and services, quality is needed that needs to be maintained. In the goods delivery (courier) industry, the main aspect that can determine the quality of an expedition is whether or not the delivery time is received by the customer. Thus, several factors that allegedly affect the timeliness need to be analyzed periodically. In this Mini project, I will first do a brief descriptive elaboration to determine the characteristics of shipping data using expeditions in Indonesian territory. After determining the parameters that are highly correlated, we can carry out further analysis centered on these variables, in order to improve the quality of freight forwarding in the future."

DATA PRE-PROCESSING

Before processing the data, the first thing to do if you get raw data is to prepare the raw data (data pre-processing), so that the data becomes structured and ready for processing.

ENTITY RELATIONSHIP DIAGRAM



Create
Database

Create
database in
pgAdmin 4
software from
PostgreSQL.

Import the
Data

Create a table in a
database that was
previously created
with the CREATE
TABLE function and
adjust the data type
that will be used for
each column, and
import data.

Generate
ERD

Defines the
relationship
between tables
by specifying
the primary key
and foreign key
using the
ALTER
function.

DATA PRE-PROCESSING

DATA UNDERSTANDING

The data used states the characteristics of the data on several types of expeditions used by freight forwarders. There are two tables in the courier company data that will be used, namely the Courier and Mapping Province tables.

Explanation of columns in each table are:

ID: Customer serial number

Expedition: The shipping company used

Customer Care Calls: Number of calls made by customers during the delivery process

Customer Rating: Rating from customers, 1 is the worst and 5 is the best

Cost of the Product: Shipping rates in IDR

Prior Purchases: The number of orders that have previously been made

Product Importance: Importance bds product categories (low, medium, high)

Gender: the sex of the customer

Discount Offered: a discount on the price of the item that was successfully obtained

Weight in gms: Weight of goods in grams

Reached on Time: 1 indicates late and 0 indicates the item was delivered on time

Province Code: Provincial code

**Courier
Table**

Province Code: Provincial code

Province: The province to which goods are sent

**Mapping
Province Table**

DATA CLEANING

1

Name	Data type
id	character varying v
expedition	character varying v
mode_of_shipment	character varying v
customer_care_cal	integer v
customer_rating	integer v
cost_of_the_produ	numeric v

2

```
1 SELECT
2     CAST(expedition AS character varying)
3 FROM courier
```

Data type	expedition
	character varying
1	Ninja Express
2	JNE

3

```
1 SELECT
2     c1.*,
3     m1.*
4 FROM courier c1
5 JOIN mapping_province m1
6 ON c1.province_code = m1.province_code
7 where m1.province_code IS NULL
```

In numbers 1 and 2 are two ways how to change the data type. In this mini project I used PostgreSQL, so when I first created a table I immediately identified the data type that would be used for each column, but if you want to change the data type in a column you can use the CAST function (column name AS data type) or the ALTER function.

Number 3 shows that how to find NULL values in a dataset. The result of this query is that the Courier and Mapping Province tables do not have NULL values.

DATA CLEANING

Find Outliers

1

```
WITH percentile AS(
  SELECT percentile_cont(0.25) within group (order by cost_of_the_product) q1,
         percentile_cont(0.75) within group (order by cost_of_the_product) q3
  FROM courier),

iqrtable AS(
  SELECT q3-q1 AS iqr
  FROM percentile),

outlier AS(
  SELECT
    q1-(1.5*iqr) minimum,
    q3+(1.5*iqr) maximum
  FROM percentile
  JOIN iqrtable
  ON 1=1)

SELECT typeoutlier, count(typeoutlier)
FROM(
  SELECT
    cost_of_the_product,
    CASE
      WHEN cost_of_the_product <= 660744 THEN 'negative outlier'
      WHEN cost_of_the_product >= 5372136 THEN 'positive outlier'
      ELSE 'inlier'
    END AS typeoutlier
  FROM courier) t2
group by 1
```

typeoutlier	count
text	bigint
inlier	10999

2

```
WITH tableuses AS(
  SELECT expedition, count(expedition) AS numberofuses
  FROM courier
  GROUP BY 1),

percentile AS(
  SELECT
    percentile_cont(0.25) within group (order by numberofuses) q1,
    percentile_cont(0.75) within group (order by numberofuses) q3
  FROM tableuses),

iqrtable AS(
  SELECT q3-q1 AS iqr
  FROM percentile),

outlier AS(
  SELECT
    q1-(1.5*iqr) minimum,
    q3+(1.5*iqr) maximum
  FROM percentile
  JOIN iqrtable
  ON 1=1)

SELECT
  expedition,
  CASE
    WHEN numberofuses <= 1831.5 THEN 'negative outlier'
    WHEN numberofuses >= 1835.5 THEN 'positive outlier'
    ELSE 'inlier'
  END AS typeoutlier
FROM tableuses
```

expedition	typeoutlier
character varying	text
TIKI	inlier
Ninja Express	inlier
SiCepat	inlier
JNE	positive outlier
J&T	inlier

Checking outliers in the table can use the interquartile formula, namely by subtracting Q3 from Q1. The first outlier check uses the cost_of_the_product column, it can be seen if that column has no data that is biased against other data. However, if we look at the number of users using various types of expeditions, it is clear that the JNE expedition has a positive type of outlier, that is, the total value of the expedition users far exceeds that of other expedition users.

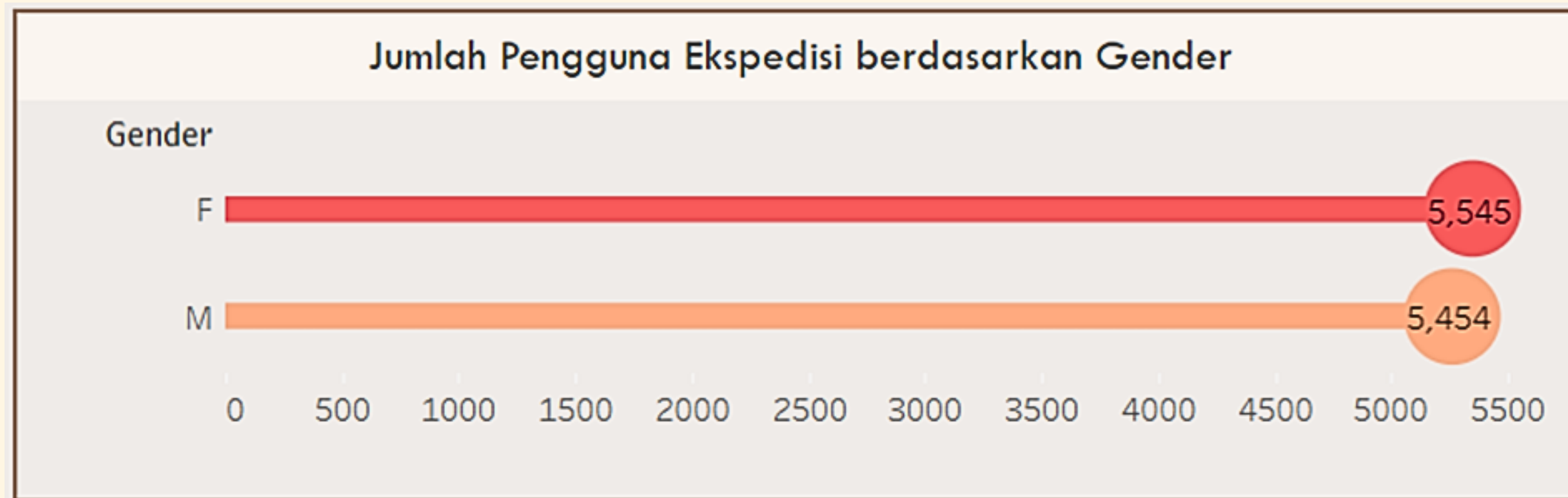
Number of Users and Average Delivery Rates by Expedition Type

Jumlah Pengguna Tiap Ekspedisi				
J&T	JNE	Ninja Express	SiCepat	TIKI
1,833	3,666	1,834	1,833	1,833

Rata-Rata Tarif Pengiriman Per-Ekspedisi				
J&T	JNE	Ninja Express	SiCepat	TIKI
3,033,555	3,003,185	3,029,480	2,998,738	3,047,456

- The highest number of users is found in JNE expeditions with a value 2x higher than the average number of other expedition users.
- The highest average shipping rates are on TIKI expeditions.

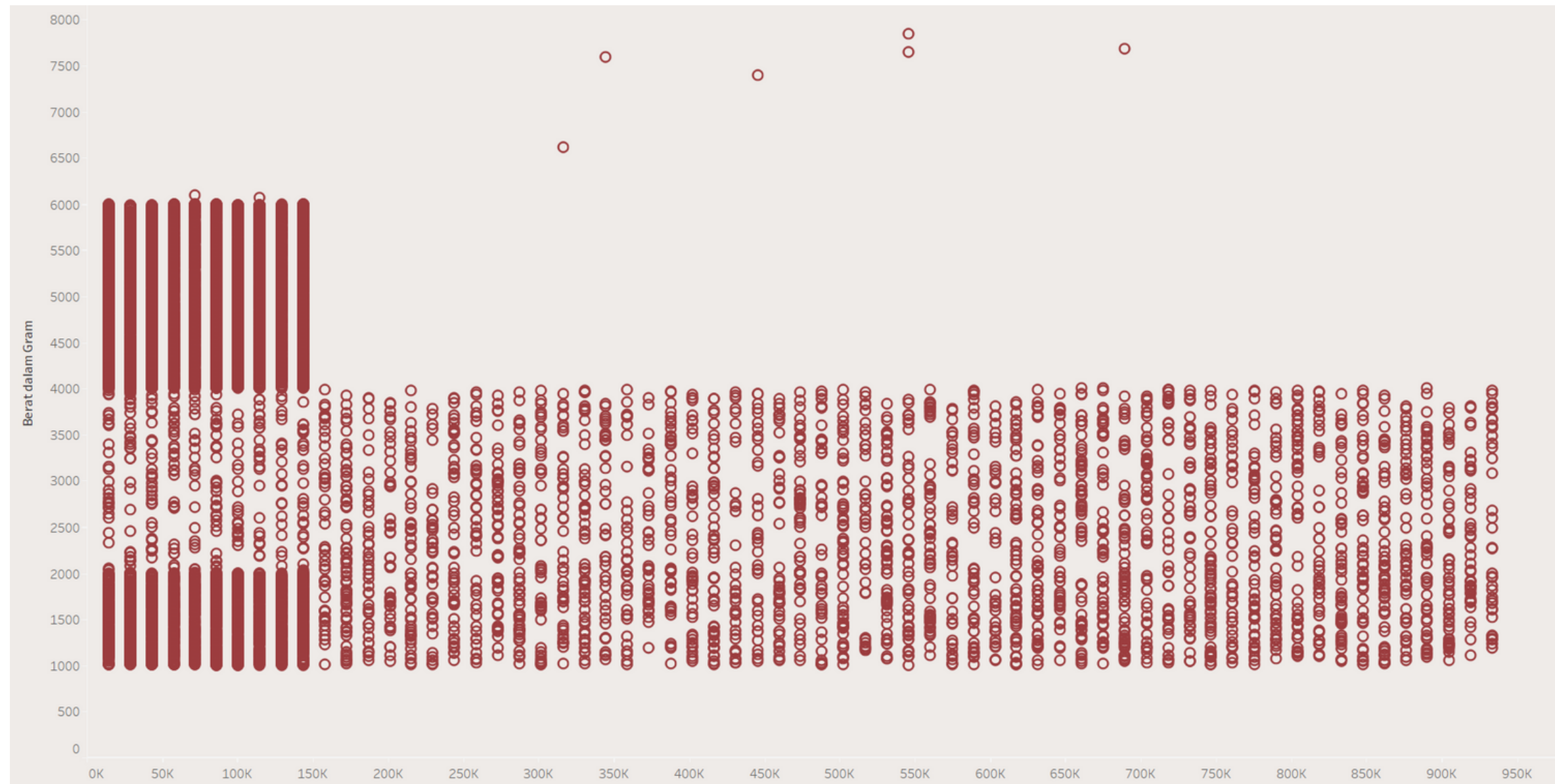
Number of Expedition Users by Gender



Based on the data above, women use more expeditions than men, this is allegedly due to the lifestyle of women who shop online more often so they will also use expedition services more often.

For details, you can see the tableau workbook [here](#)

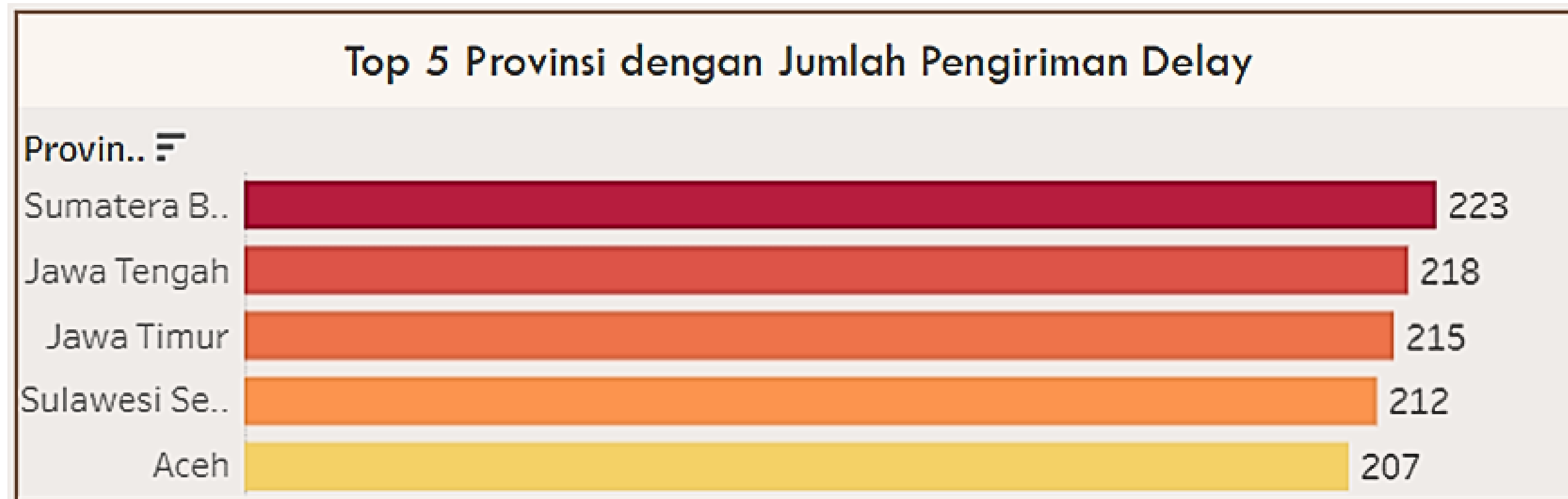
Correlation Between Product Weight and Discount Offered



Goods weighing above 4 kilograms on average get a discount below 150 thousand rupiah, and no goods get a discount above the price of 150 thousand for goods weighing over 4 kilograms.

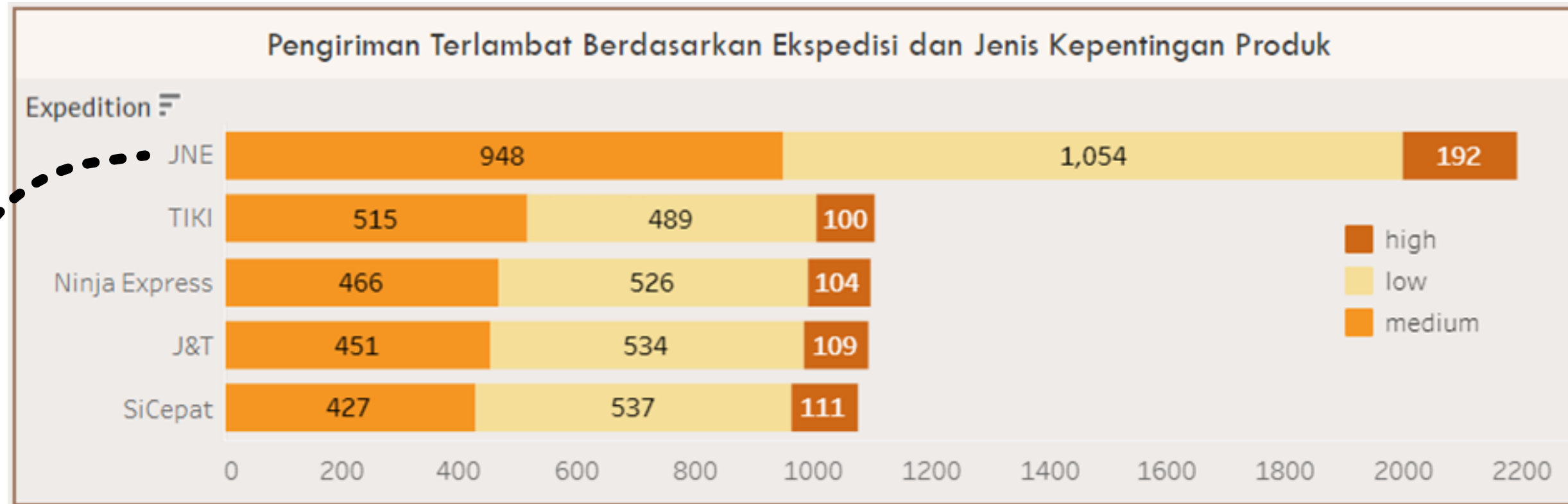
For details, you can see the tableau workbook [here](#)

Top 5 Provinces with Number of Delivery Delays



West Sumatra is the province with the most delivery delays in Indonesia compared to 34 other provinces in the data, even though it ranks 12th in the number of expedition users but the number of delivery delays reaches 66.8% compared to the total shipments to West Sumatra destination cities.

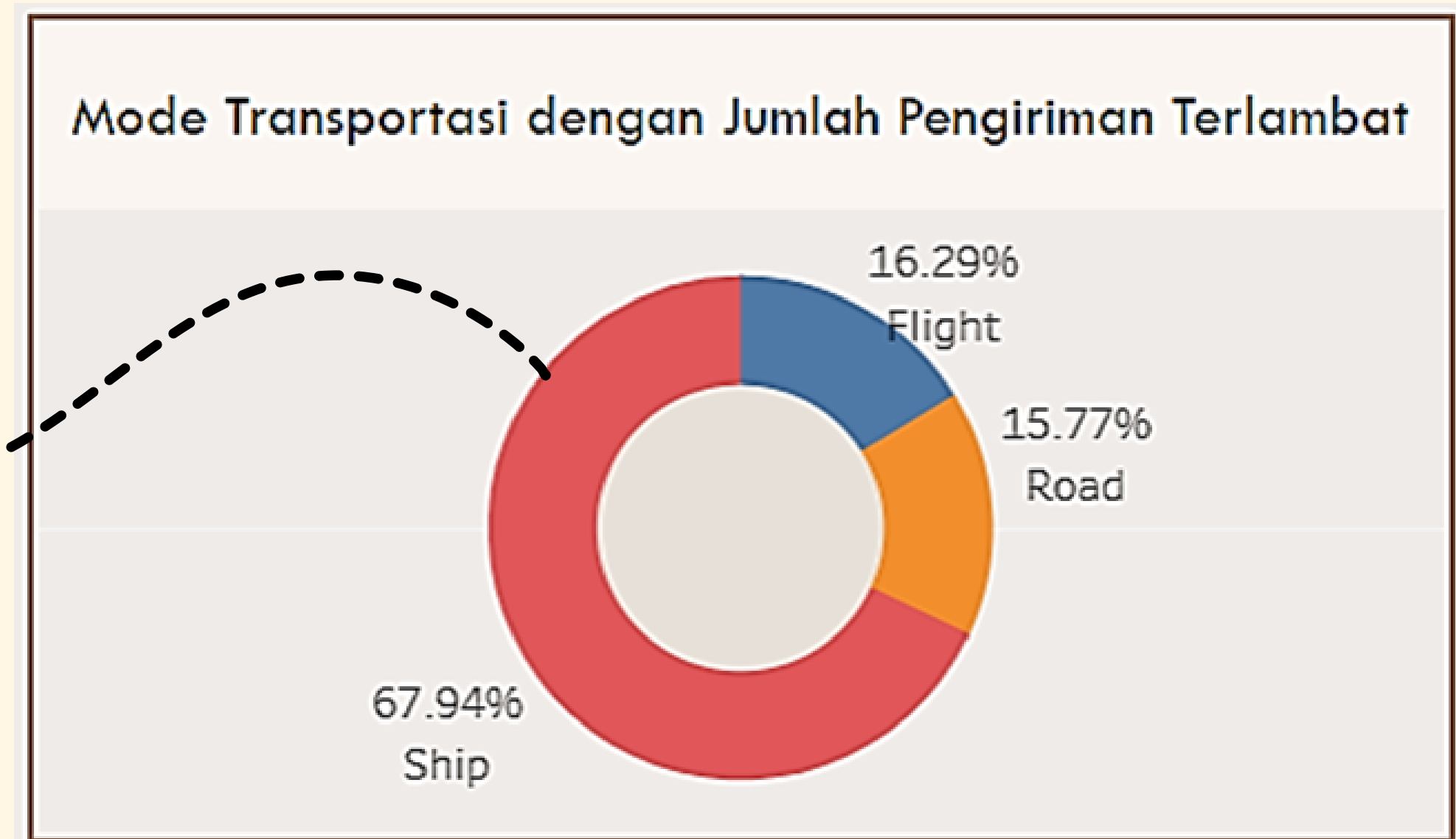
Late Delivery Based on Expedition and Type of Product Interests



JNE expeditions had a higher total delay than other expeditions, with a total of 2,194 shipments, with medium category products which experienced more delays.

Mode of Transport with Number of Late Shipments

Shipping transportation using ships is more prone to experiencing delays or delays in delivery, this is allegedly due to situations that occur when shipments are unstable.



CONCLUSION

- The highest number of expedition users is found in JNE expeditions with the highest average rates found in TIKI expeditions.
- Female and male users do not have a significant difference regarding the number of expedition users based on gender.
- Goods weighing over 4 kilograms have a discount offer below 150 thousand rupiah.
- Most of the delivery delays are in West Sumatra Province and in the expedition category there are JNE expeditions, so the average expedition in Indonesia, the quality of the expedition in terms of delivery time is still quite low.
- The mode of transportation that is prone to delays in delivery is shipping by ship, and good shipping can use the road.

Mini Project

Completion Certification



has been presented to

Benedicta Cardalia Ginting

For successfully completing business intelligence
Mini Project How to Optimize Performance from Non-Delayed
into Super On-Time Process for Shipping Courier in Indonesia

CEO Rakamin Academy

A handwritten signature in black ink, appearing to read "Andika".

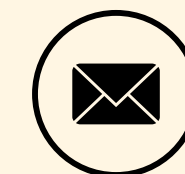
Andika Deni Prasetya

Business Analyst,
Bank Mandiri

A handwritten signature in black ink, appearing to read "Dimas".

Dimas Rahadiyuza

*Keep in touch
with me*



liaacardalia@gmail.com



<https://www.linkedin.com/in/benedicta-cardalia>



<https://github.com/cardalia>