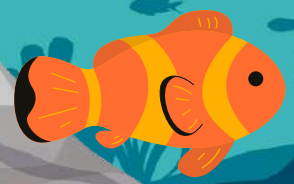
BIG DATA ANALYTICS

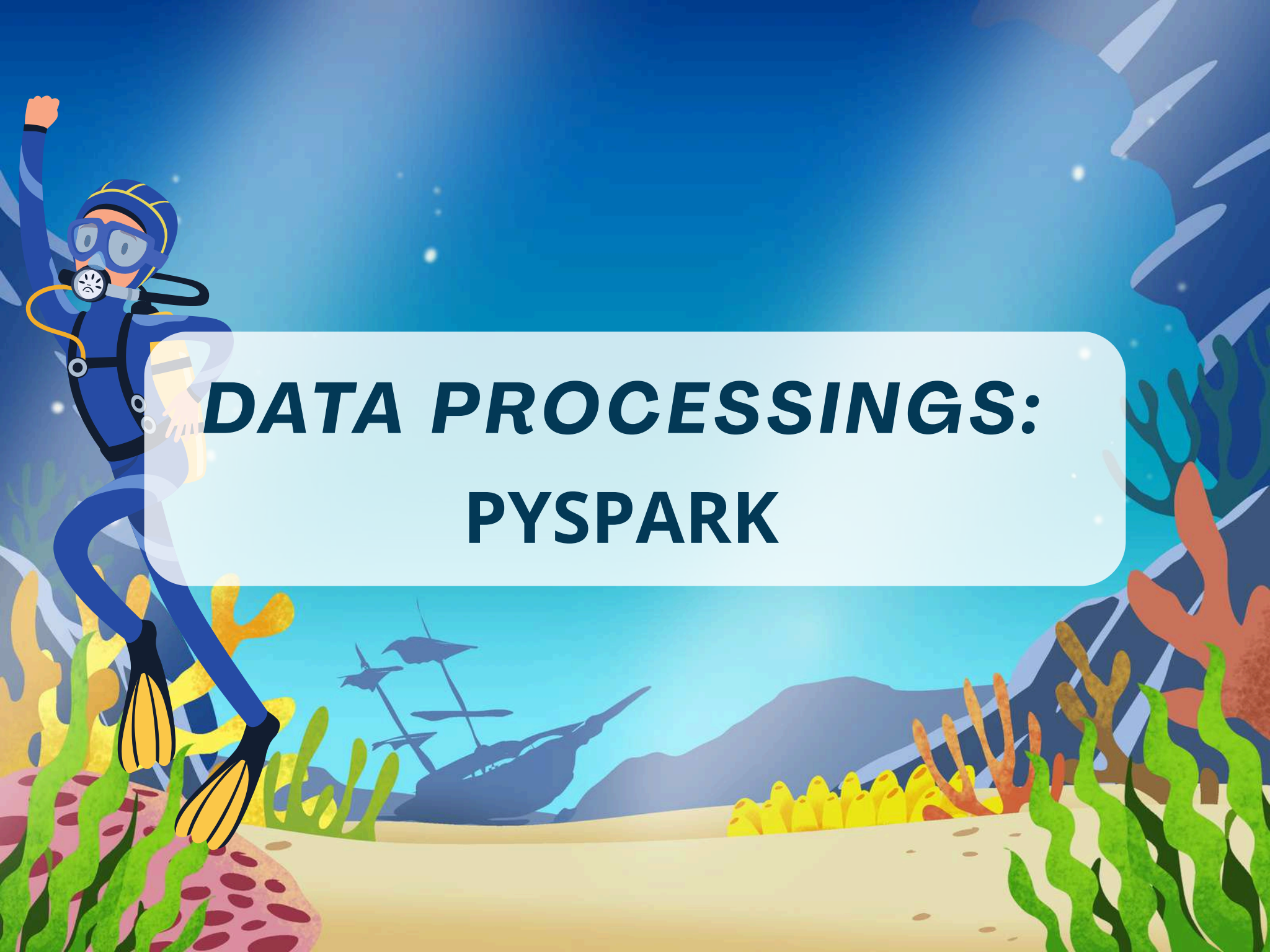
"ANALISIS PEMANFAATAN BIG DATA UNTUK PENGELOLAAN DATASETS PERIKANAN"

PRESENTED BY:
KELOMPOK 1

ABOUT DATASET

- fish_dataset.csv (Dataset Utama)
- fishers_dataset.csv
- fish_locations_dataset.csv
- fish_categories_dataset
- fish_market_prices.csv




DATA PROCESSINGS:
PYSPARK

PySpark Setup Process

Langkah-langkah nya:

1. Install Java & PySpark
2. Set Environment Variables
3. Inisialisasi SparkSession

Load Datasets

1. Dataset Utama: fish_dataset

Menyimpan hasilnya ke dalam variabel fish_df

fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id	fisher_id
1	Catfish	19.83	115.1	7/6/2023	Papua	Trap	3	4	4	381
2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line	10	4	12	337
3	Anchovy	43.1	184.4	5/1/2023	Jawa Timur	Purse Seine	4	2	18	144
4	Sole	36.27	106.9	5/25/2023	Maluku	Gill Net	10	1	13	271
5	Sole	15.7	162.8	3/29/2023	Jawa Timur	Trawling	2	3	20	244

only showing top 5 rows

2. fishers_dataset

Menyimpan hasilnya ke dalam variabel fishers_df

fisher_id	fisher_name	experience_years	boat_type	boat_capacity_kg	home_port	license_number
1	Nelayan_001	26	Motor Boat	3359	Lampung	LIC901799
2	Nelayan_002	33	Traditional Boat	860	Jawa Timur	LIC612729
3	Nelayan_003	3	Traditional Boat	932	Sumatera Utara	LIC650826
4	Nelayan_004	29	Motor Boat	2303	Papua	LIC942039
5	Nelayan_005	36	Trawler	2438	Sulawesi Selatan	LIC488256

only showing top 5 rows

Load Datasets

3. fish_categories_dataset

Menyimpan hasilnya ke dalam variabel fish_categories_df

```
+-----------+---------------+--------------+----------------+
|category_id|  category_name|       habitat|avg_price_per_kg|
+-----------+---------------+--------------+----------------+
|          1|   Pelagic Fish|    Open Ocean|           85000|
|          2|  Demersal Fish|    Sea Bottom|          120000|
|          3|      Reef Fish|    Coral Reef|          150000|
|          4|Freshwater Fish|  Rivers/Lakes|           45000|
|          5|Anadromous Fish|Sea/Freshwater|          200000|
+-----------+---------------+--------------+----------------+
```

4. fish_locations_dataset

Menyimpan hasilnya ke dalam variabel fish_locations_df

```
+-----------+--------------------+---------+----------+-------------+------------------+-----------+
|location_id|       location_name| latitude| longitude|water_depth_m|temperature_celsius|salinity_ppt|
+-----------+--------------------+---------+----------+-------------+------------------+-----------+
|          1|  Laut Flores Zone A| 1.612803|  99.89888|         1276|              29.2|       33.5|
|          2|   Laut Timor Zone B|-6.261366|116.469252|         1153|              29.3|       33.6|
|          3|  Laut Natuna Zone C| 2.105336|123.630308|         2780|              26.0|       33.8|
|          4|    Laut Jawa Zone D|-8.014815|100.237065|         2780|              29.8|       33.5|
|          5|Laut Sulawesi Zone E| 1.763838|111.580859|          424|              25.9|       32.6|
+-----------+--------------------+---------+----------+-------------+------------------+-----------+
only showing top 5 rows
```

Data Cleaning

1. Dataset Utama: fish_dataset

Check Missing Values:

Hasil pengecekan disimpan di variabel missing_values

fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id	fisher_id
0	1	3	1	0	0	0	1	0	0	0

Kolom yang terdapat missing values:

- fish_name = 1
- weight_kg = 3
- length_cm = 1
- freshness_score = 1

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 1: Isi missing value di kolom fish_name dengan nilai modus

Langkah 2: Isi missing value di kolom weight_kg dengan median

Langkah 3: Isi missing value di kolom length_cm dengan median

Langkah 4: Isi missing value di kolom freshness_score dengan modus

Langkah 5: Cek lagi sisa missing values keseluruhan kolom

fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method
0	0	0	0	0	0	0

freshness_score	category_id	location_id	fisher_id
0	0	0	0

Seluruh nilai kosong pada dataset fish_dataset berhasil ditangani, data sudah bersih dan siap untuk dianalisis lebih lanjut.

Langkah 6: Hapus data duplikat berdasarkan fish_id

Data Cleaning

Check Missing Values

2. fishers_dataset

fisher_id	fisher_name	experience_years	boat_type	boat_capacity_kg	home_port	license_number
0	0	0	0	0	0	0

3. fish_categories_dataset

category_id	category_name	habitat	avg_price_per_kg
0	0	0	0

Semua kolom di dataset fishers_dataset dan fish_categories_dataset tidak mengandung missing values

Data Cleaning

4. fish_locations_dataset

location_id	location_name	latitude	longitude	water_depth_m	temperature_celsius	salinity_ppt
0	0	0	0	0	0	0

Semua kolom di dataset fishers_dataset, fish_categories_datset, fish_locations_dataset tidak mengandung missing values

Join Operations

(Inner Join antar dataset)

1. Inner Join antara **‘fish_dataset’** dengan **‘fish_categories_dataset’**

Hasilnya disimpan di DataFrame baru bernama joined_df

2. Inner Join hasil joined_df sebelumnya dengan **‘fishers_dataset’** dan **‘fish_locations_dataset**

Hasilnya disimpan di DataFrame baru (hasil penggabungan empat dataset) bernama final_joined_df

Hasil setelah melakukan inner join antar dataset:
seluruh dataset berhasil tergabung.

location_id	fisher_id	category_id	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method
4	381	4	1	Catfish	19.83	115.1	7/6/2023	Papua	Trap
12	337	4	2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line
18	144	2	3	Anchovy	43.1	184.4	5/1/2023	Jawa Timur	Purse Seine

only showing top 3 rows

freshness_score	category_name	habitat	avg_price_per_kg	fisher_name	experience_years	boat_type	boat_capacity_kg
3	Freshwater Fish	Rivers/Lakes	45000	Nelayan_381	26	Motor Boat	3264
10	Freshwater Fish	Rivers/Lakes	45000	Nelayan_337	19	Motor Boat	4236
4	Demersal Fish	Sea Bottom	120000	Nelayan_144	10	Motor Boat	3465

home_port	license_number	location_name	latitude	longitude	water_depth_m	temperature_celsius	salinity_ppt
Sulawesi Selatan	LIC250393	Laut Jawa Zone D	-8.014815	100.237065	2780	29.8	33.5
Sulawesi Selatan	LIC493492	Laut Flores Zone B	-3.052861	139.649345	728	24.0	33.4
Sumatera Utara	LIC411388	Laut Jawa Zone H	0.689541	101.236256	2306	24.8	33.7


ANALISIS SEDERHANA:
PYSPARK

1. Hitung rata-rata berat per jenis ikan

fish_name	avg(weight_kg)
Swordfish	24.973698296837004
Halibut	25.360780141843968
Tilapia	25.64151589242053
Pomfret	25.862076372315048
Tuna	26.995311720698258
Catfish	25.07497560975611
Herring	23.86791962174938
Salmon	25.29116331096196
Sardine	25.286651376146814
Trout	25.407820224719096
Barramundi	25.350092592592596
Yellowtail	25.187732696897378
Carp	25.677430730478573
Carp	25.677430730478573
Mackerel	26.08572178477694
Cod	25.079391100702587
Mahi-mahi	25.29258536585366
Snapper	25.663428571428586
Sea Bass	26.04451456310679
Grouper	25.59466981132073
Anchovy	24.666567901234576
Flounder	24.85926436781611
Sole	25.288005050505067
Amberjack	25.33071960297766
King Fish	24.408171021377687

2. Hitung rata-rata panjang ikan per jenis kategori ikan

category_name	avg(length_cm)
Demersal Fish	108.3855511420061
Reef Fish	107.52776375567876
Anadromous Fish	108.60477849676451
Freshwater Fish	105.76244274809163
Pelagic Fish	108.8427589592537 9

3. Hitung berapa total berat ikan yang berhasil ditangkap oleh setiap nelayan

fisher_name	boat_type	total_tangkapan_kg
Nelayan_038	Motor Boat	865.2799999999999
Nelayan_391	Traditional Boat	860.44
Nelayan_344	Trawler	844.4899999999999
Nelayan_332	Fishing Vessel	836.4399999999999
Nelayan_304	Motor Boat	836.4199999999998
Nelayan_075	Traditional Boat	831.8
Nelayan_106	Trawler	822.7800000000002
Nelayan_250	Traditional Boat	817.98
Nelayan_335	Motor Boat	816.45
Nelayan_139	Traditional Boat	783.57
Nelayan_174	Fishing Vessel	781.09
Nelayan_343	Trawler	777.8000000000001
Nelayan_120	Traditional Boat	777.0900000000001
Nelayan_385	Trawler	771.6299999999999
Nelayan_017	Trawler	769.7400000000004
Nelayan_194	Trawler	769.6099999999999
Nelayan_108	Traditional Boat	768.9700000000001
Nelayan_197	Traditional Boat	758.87
Nelayan_164	Trawler	758.6300000000001
Nelayan_015	Traditional Boat	746.0599999999998

only showing top 20 rows

4. Hubungan pengalaman nelayan dengan volume tangkapan

fisher_name	experience_years	total_weight_kg
Nelayan_439	40	409.6700000000001
Nelayan_422	40	500.56
Nelayan_141	40	422.82000000000005
Nelayan_457	40	380.4
Nelayan_264	40	502.21000000000004
Nelayan_429	40	316.75
Nelayan_205	40	670.45
Nelayan_045	40	433.6
Nelayan_219	40	465.05999999999995
Nelayan_310	40	384.09000000000003
Nelayan_424	40	492.03000000000003
Nelayan_398	39	546.41
Nelayan_383	39	359.00000000000006
Nelayan_351	39	625.62
Nelayan_221	39	494.32
Nelayan_386	39	298.04999999999995
Nelayan_462	39	583.29
Nelayan_342	39	400.0299999999999
Nelayan_392	39	261.94
Nelayan_500	39	439.39000000000004

only showing top 20 rows

5. Karakteristik lingkungan tempat ikan ditangkap

fish_name	avg_depth	avg_temp	avg_salinity
Swordfish	1465.094890510949	27.48807785888075	32.5841849148418
Halibut	1489.983451536643	27.389361702127626	32.63711583924347
Tilapia	1455.8508557457212	27.345476772616138	32.61735941320292
Pomfret	1557.6825775656325	27.39880668257753	32.66396181384245
Tuna	1511.9775561097256	27.188279301745613	32.780798004987496
Catfish	1469.3658536585365	27.37121951219509	32.67024390243899
Herring	1565.5862884160756	27.420094562647744	32.813238770685516
Salmon	1480.0738255033557	27.29932885906034	32.68165548098428
Sardine	1429.6605504587155	27.45665137614674	32.665137614678855
Trout	1446.5573033707865	27.292134831460626	32.532584269662884
Barramundi	1566.0625	27.5092592592592	32.66828703703699
Yellowtail	1538.3054892601433	27.405966587112133	32.749164677804266
Carp	1469.0982367758186	27.564735516372803	32.620906801007486
Mackerel	1400.6850393700788	27.31154855643042	32.7133858267716
Cod	1552.4262295081967	27.34730679156906	32.61194379391094
Mahi-mahi	1550.5341463414634	27.24048780487801	32.69682926829264
Snapper	1475.1166666666666	27.485952380952337	32.557380952380896
Sea Bass	1605.8398058252428	27.42548543689318	32.711165048543656
Grouper	1434.0306603773586	27.56485849056602	32.59245283018865
Anchovy	1447.3753086419754	27.41407407407403	32.53160493827157

only showing top 20 rows

6. Metode penangkapan paling efektif untuk kategori ikan tertentu

category_name	fishing_method	total_catch
Anadromous Fish	Purse Seine	9753.560000000001
Demersal Fish	Purse Seine	9482.409999999998
Pelagic Fish	Trawling	9423.660000000003
Reef Fish	Gill Net	9102.06
Demersal Fish	Gill Net	8946.980000000005
Pelagic Fish	Purse Seine	8880.890000000005
Freshwater Fish	Trap	8862.420000000007
Demersal Fish	Hand Line	8810.899999999998
Demersal Fish	Trawling	8754.460000000006
Freshwater Fish	Trawling	8538.71
Pelagic Fish	Gill Net	8525.81
Freshwater Fish	Hand Line	8520.299999999992
Anadromous Fish	Trawling	8490.9
Reef Fish	Purse Seine	8483.919999999996
Reef Fish	Long Line	8480.869999999999
Pelagic Fish	Trap	8478.370000000003
Anadromous Fish	Trap	8421.220000000003
Freshwater Fish	Purse Seine	8382.630000000001
Reef Fish	Trawling	8364.959999999995
Anadromous Fish	Long Line	8325.4

only showing top 20 rows


DATA PROCESSINGS:
PANDAS

Pandas Setup Process

Langkah-langkah nya:

1. Import Library

   import pandas, Import file

2. Load Dataset

   Di langkah ini, kita mulai membaca semua file data yang kita butuhkan.
   Kita pakai perintah read_csv dari Pandas untuk buka file-file itu, lalu masing-masing file disimpan ke dalam variabel. Setelah itu semua data ini akan digabung jadi satu untuk dianalisis lebih lanjut.

3. Cek Missing Values & Cleaning Datasets

   Melakukan pengecekan pada asetiap datasets dan membersihkan setiap datasets.

4. Join Operation

5. Analisis dan Visualisasi

Load Datasets

Fish Datasets

Dalam File colab ini ada 5 dataset CSV yang di upload ke dalam Python menggunakan library pandas. Dataset tersebut berisi berbagai informasi penting:

- fish_datasets.csv
- fish_categories_dataset.csv
- fish_locations_dataset.csv
- fishers_dataset.csv
- fish_market_prices_dataset.csv

Data Cleaning

1. Dataset Utama: fish_dataset

Check Missing Values:

- Menggunakan fungsi .isnull().sum() dari pandas.
- Hasil akhir disimpan di variabel missing_values

	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id	fisher_id
0	0	1	3	1	0	0	0	1	0	0	0

Kolom yang terdapat missing values:

- fish_name = 1
- weight_kg = 3
- length_cm = 1
- freshness_score = 1

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 1: Isi missing value di kolom leght_cm dengan median

- Menggunakan fungsi median() karena kolom lenght_cm bersifat numerik
- Mengisi semua nilai kosong pada kolom length_cm dengan nilai median yang sudah dihitung.
- Menggunakan metode fillna, Akan mengganti semua nilai kosong dengan nilai median dan Menyimpan hasil pengisian kembali ke kolom length_cm.

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 2: Isi missing value di kolom fish_name dengan modus

- Menggunakan fungsi mode() karena kolom fish_name bersifat kategori
- Mengisi semua kosong pada kolom fish_name dengan nilai modus kategori nama ikan yang sering muncul.
- Menggunakan metode fillna, Akan mengganti semua nilai kosong dengan nilai modus dan Menyimpan hasil pengisian kembali ke kolom fish_name.

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 3: Isi missing value di kolom weight_kg dengan median

- Menggunakan fungsi median() karena kolom weight_kg bersifat numerik
- Mengisi semua nilai kosong pada kolom weight_kg dengan nilai median yang sudah dihitung.
- Menggunakan metode fillna, Akan mengganti semua nilai kosong dengan nilai median dan Menyimpan hasil pengisian kembali ke kolom weight_lg

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 4: Isi missing value di kolom catch_data dengan modus

- Menggunakan fungsi mode() karena kolom catch_data bersifat kategorikal bertipe tanggal.
- Mengisi semua kosong pada kolom catch_data dengan nilai modus kategori nama ikan yang sering muncul.
- Menggunakan metode fillna, Akan mengganti semua nilai kosong dengan nilai modus dan Menyimpan hasil pengisian kembali ke kolom catch_data.

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 5: Isi missing value di kolom freshness_score dengan modus

- Menggunakan fungsi mode() karena kolom freshness_score bersifat kategori tersembunyi.
- Mengisi semua kosong pada kolom freshness_scoredengan nilai modus kategori freshness_score yang sering muncul.
- Menggunakan metode fillna, Akan mengganti semua nilai kosong dengan nilai modus dan Menyimpan hasil pengisian kembali ke kolom freshness_score.

Data Cleaning

1. Dataset Utama: fish_dataset

Langkah 6: Menghapus duplikat data pada fish_id

Untuk memastikan setiap id ikan hanya muncul satu kali, dilakukan penghapusan data duplikat berdasarkan baris fish_id.
Langkah ini bertujuan untuk menghindari pengulangan data yang dapat memengaruhi hasil analisis, terutama saat membuat ringkasan atau visualisasi data.

	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id	fisher_id
0	1	Catfish	19.83	115.1	7/6/2023	Papua	Trap	3.0	4	4	381
1	2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line	10.0	4	12	337
2	3	Anchovy	43.10	184.4	5/1/2023	Jawa Timur	Purse Seine	4.0	2	18	144
3	4	Sole	36.27	106.9	5/25/2023	Maluku	Gill Net	10.0	1	13	271
4	5	Sole	15.70	162.8	3/29/2023	Jawa Timur	Trawling	2.0	3	20	244

Data Cleaning

2. Dataset Selanjutnya: fish_categories, fish_locations, fish_fishers.

Check Missing Values:

- Menggunakan fungsi .isnull().sum() dari pandas.
- Hasil akhir disimpan di variabel missing_values.

	category_id	category_name	habitat	avg_price_per_kg
0	0	0	0	0

	location_id	location_name	latitude	longitude	water_depth_m	temperature_celsius	salinity_ppt
0	0	0	0	0	0	0	0

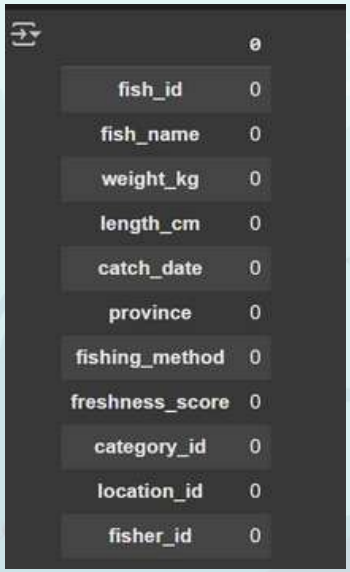
	fisher_id	fisher_name	experience_years	boat_type	boat_capacity_kg	home_port	license_number
0	0	0	0	0	0	0	0

Data Cleaning

Cek apakah masih ada missing values

Langkah Akhir: Cek apakah masih ada missing values

Menggunakan fungsi df_fish.isnull().sum()
untuk mengecek apakah masih ada data yang kosong (missing values) di setiap kolom setelah proses pembersihan data.



	0
fish_id	0
fish_name	0
weight_kg	0
length_cm	0
catch_date	0
province	0
fishing_method	0
freshness_score	0
category_id	0
location_id	0
fisher_id	0

Join Operation

1. Menggabungkan Fish_datasets dengan fish_category

	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id
0	1	Catfish	19.83	115.1	7/6/2023	Papua	Trap	3.0	4	4
1	2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line	10.0	4	12
2	3	Anchovy	43.10	184.4	5/1/2023	Jawa Timur	Purse Seine	4.0	2	18
3	4	Sole	36.27	106.9	5/25/2023	Maluku	Gill Net	10.0	1	13
4	5	Sole	15.70	162.8	3/29/2023	Jawa Timur	Trawling	2.0	3	20

fisher_id	category_name	habitat	avg_price_per_kg
381	Freshwater Fish	Rivers/Lakes	45000
337	Freshwater Fish	Rivers/Lakes	45000
144	Demersal Fish	Sea Bottom	120000
271	Pelagic Fish	Open Ocean	85000
244	Reef Fish	Coral Reef	150000

- Menggunakan fungsi pd.marge()
- Menggunakan paremeter how=inner untuk mengetauhi mana saja yang di inner join.
- Hasil dari operasi penggabungan ini akan disimpan dalam DataFrame baru bernama joined_df.

Join Operation

2. Menggabungkan Fish_location

	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id	fisher_id	category_name
0	1	Catfish	19.83	115.1	7/6/2023	Papua	Trap	3.0	4	4	381	Freshwater Fish
1	2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line	10.0	4	12	337	Freshwater Fish
2	3	Anchovy	43.10	184.4	5/1/2023	Jawa Timur	Purse Seine	4.0	2	18	144	Demersal Fish
3	4	Sole	36.27	106.9	5/25/2023	Maluku	Gill Net	10.0	1	13	271	Pelagic Fish
4	5	Sole	15.70	162.8	3/29/2023	Jawa Timur	Trawling	2.0	3	20	244	Reef Fish

habitat	avg_price_per_kg	location_name	latitude	longitude	water_depth_m	temperature_celsius	salinity_ppt
Rivers/Lakes	45000	Laut Jawa Zone D	-8.014815	100.237065	2780	29.8	33.5
Rivers/Lakes	45000	Laut Flores Zone B	-3.052861	139.649345	728	24.0	33.4
Sea Bottom	120000	Laut Jawa Zone H	0.689541	101.236256	2306	24.8	33.7
Open Ocean	85000	Laut Banda Zone C	-6.811777	137.613943	2682	27.2	32.1
Coral Reef	150000	Laut Natuna Zone J	4.826068	110.349063	59	28.7	31.0

- Menggunakan fungsi pd.marge()
- Menggunakan paremeter how=inner untuk mengetauhi mana saja yang di inner join.
- Hasil dari operasi penggabungan ini akan disimpan dalam DataFrame baru bernama joined_df.

Join Operation

3. Menggabungkan Fisher

	fish_id	fish_name	weight_kg	length_cm	catch_date	province	fishing_method	freshness_score	category_id	location_id
0	1	Catfish	19.83	115.1	7/6/2023	Papua	Trap	3.0	4	4
1	2	Carp	21.02	101.8	11/19/2023	Sulawesi Selatan	Hand Line	10.0	4	12
2	3	Anchovy	43.10	184.4	5/1/2023	Jawa Timur	Purse Seine	4.0	2	18
3	4	Sole	36.27	106.9	5/25/2023	Maluku	Gill Net	10.0	1	13
4	5	Sole	15.70	162.8	3/29/2023	Jawa Timur	Trawling	2.0	3	20

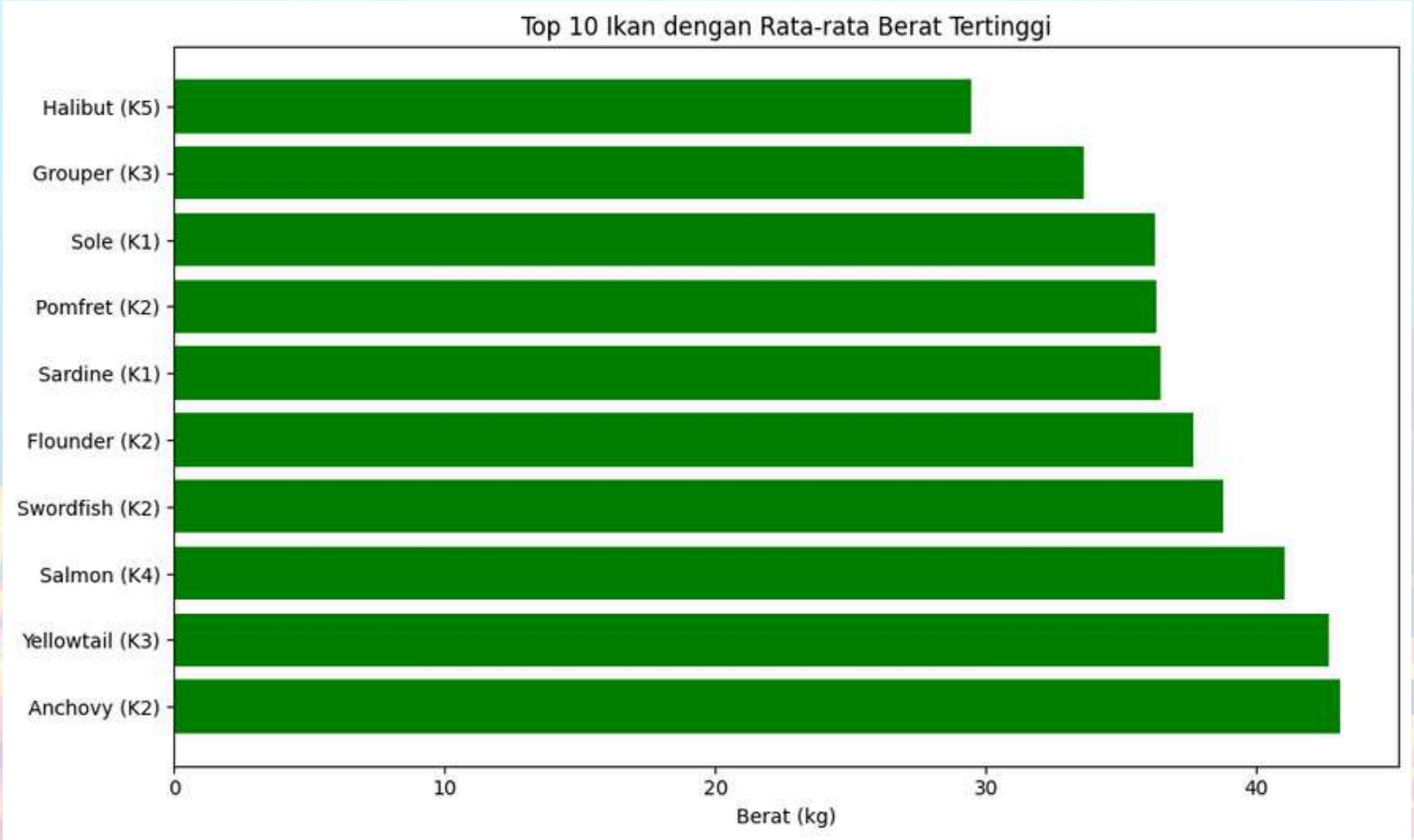
longitude	water_depth_m	temperature_celsius	salinity_ppt	fisher_name	experience_years	boat_type	boat_capacity_kg	home_port	license_number
100.237065	2780	29.8	33.5	Nelayan_381	26	Motor Boat	3264	Sulawesi Selatan	LIC250393
139.649345	728	24.0	33.4	Nelayan_337	19	Motor Boat	4236	Sulawesi Selatan	LIC493492
101.236256	2306	24.8	33.7	Nelayan_144	10	Motor Boat	3465	Sumatera Utara	LIC411388
137.613943	2682	27.2	32.1	Nelayan_271	5	Motor Boat	1004	Bali	LIC129483
110.349063	59	28.7	31.0	Nelayan_244	13	Trawler	379	Papua	LIC188541

- Menggunakan fungsi pd.marge()
- Menggunakan paremeter how=inner untuk mengetauhi mana saja yang di inner join.
- Hasil dari operasi penggabungan ini akan disimpan dalam DataFrame baru bernama joined_df.


ANALISIS SEDERHANA
DAN VISUAlISASI:
PANDAS

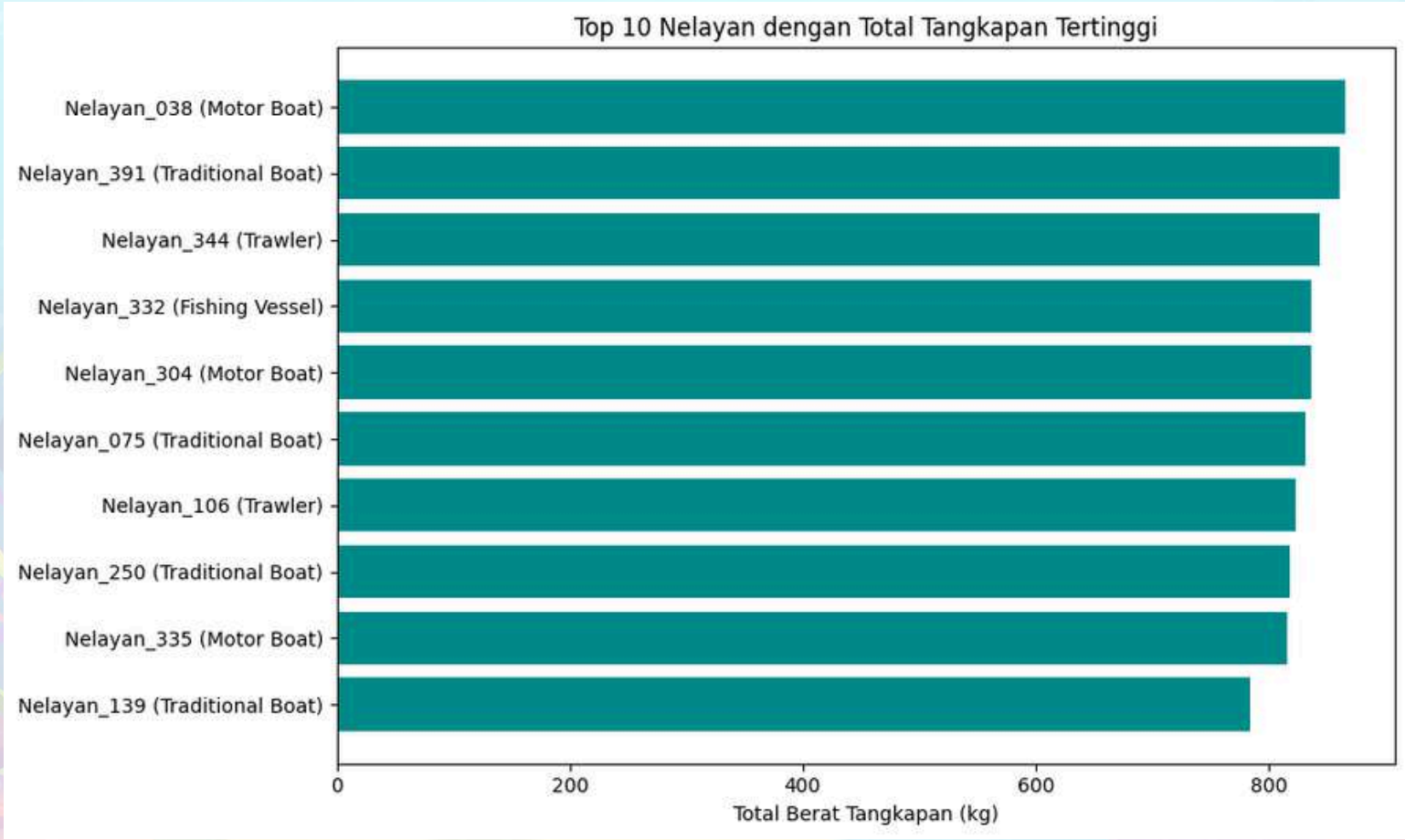
1. Hitung Rata rata berat perjenis ikan

	fish_name	category_id	weight_kg	length_cm
0	Amberjack	5	7.44	128.3
1	Anchovy	2	43.10	184.4
2	Barramundi	4	12.13	88.0
3	Carp	4	21.02	101.8
4	Catfish	4	19.83	115.1
5	Cod	2	21.57	190.8
6	Flounder	2	37.67	115.6
7	Grouper	3	33.63	59.9
8	Halibut	5	29.48	72.1
9	Herring	1	3.70	21.9
10	King Fish	2	12.20	58.2
11	Mackerel	1	9.22	193.3
12	Mahi-mahi	4	16.64	98.9
13	Pomfret	2	36.31	49.7
14	Salmon	4	41.02	101.4
15	Sardine	1	36.45	95.3
16	Sea Bass	3	10.30	78.4
17	Snapper	2	28.42	74.7
18	Sole	1	36.27	106.9
19	Swordfish	2	38.80	32.8
20	Tilapia	1	16.85	68.7
21	Trout	3	0.71	170.4
22	Tuna	4	27.57	69.9
23	Yellowtail	3	42.66	187.4


Top 10 Ikan dengan Rata-rata Berat Tertinggi
Halibut (K5)
Grouper (K3)
Sole (K1)
Pomfret (K2)
Sardine (K1)
Flounder (K2)
Swordfish (K2)
Salmon (K4)
Yellowtail (K3)
Anchovy (K2)
0
10
20
30
40
Berat (kg)

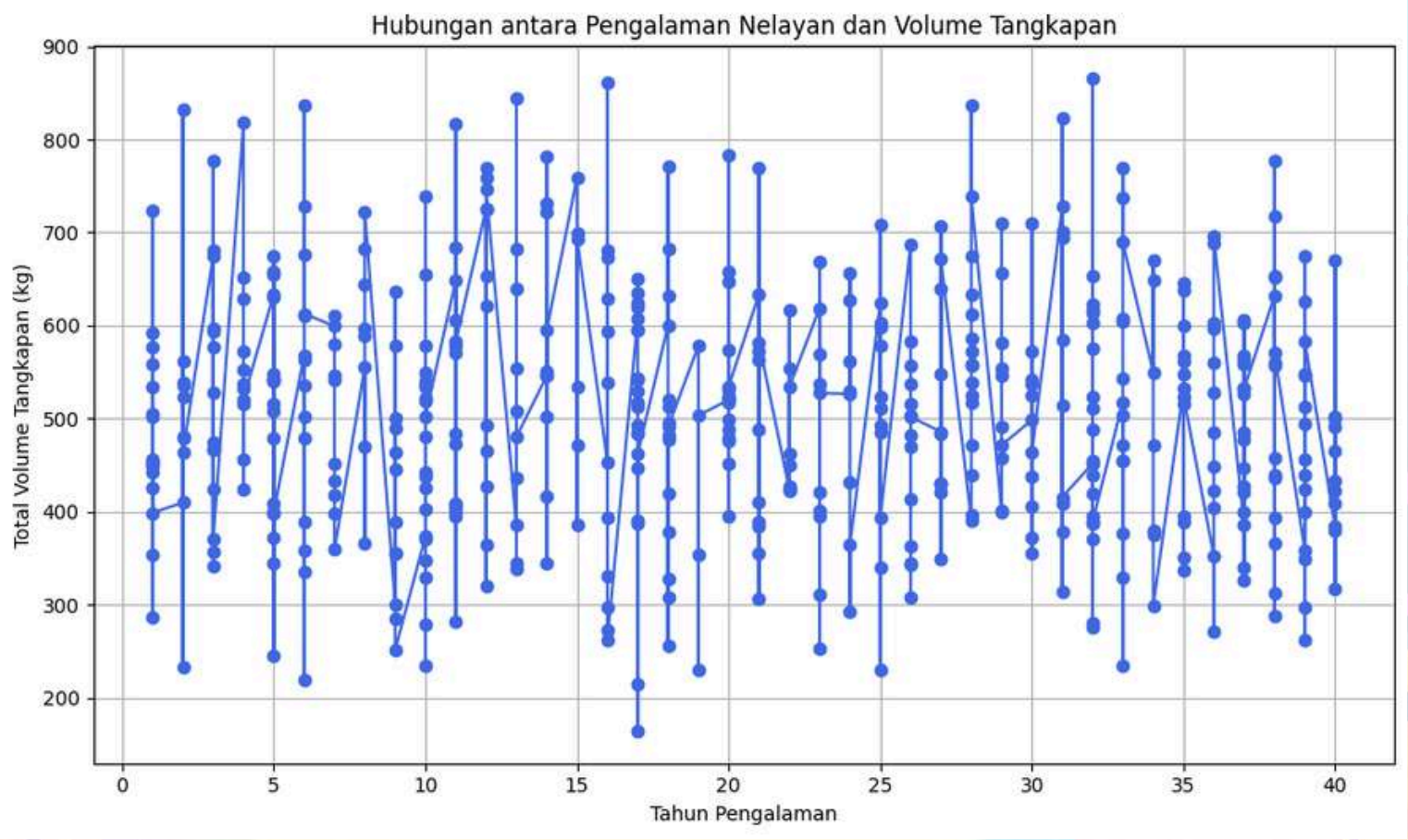
2. TOTAL TANGKAPAN NELAYAN

	fisher_name	boat_type	weight_kg
0	Nelayan_001	Motor Boat	686.58
1	Nelayan_002	Traditional Boat	604.61
2	Nelayan_003	Traditional Boat	597.49
3	Nelayan_004	Motor Boat	491.68
4	Nelayan_005	Trawler	486.05
...	...	...	...
495	Nelayan_496	Fishing Vessel	463.92
496	Nelayan_497	Motor Boat	457.12
497	Nelayan_498	Motor Boat	603.67
498	Nelayan_499	Traditional Boat	416.90
499	Nelayan_500	Traditional Boat	439.39


Top 10 Nelayan dengan Total Tangkapan Tertinggi
Nelayan_038 (Motor Boat)
Nelayan_391 (Traditional Boat)
Nelayan_344 (Trawler)
Nelayan_332 (Fishing Vessel)
Nelayan_304 (Motor Boat)
Nelayan_075 (Traditional Boat)
Nelayan_106 (Trawler)
Nelayan_250 (Traditional Boat)
Nelayan_335 (Motor Boat)
Nelayan_139 (Traditional Boat)
0
200
400
600
800
Total Berat Tangkapan (kg)

3. HUBUNGAN PENGALAMAN NELAYAN DENGAN TANGKAPAN

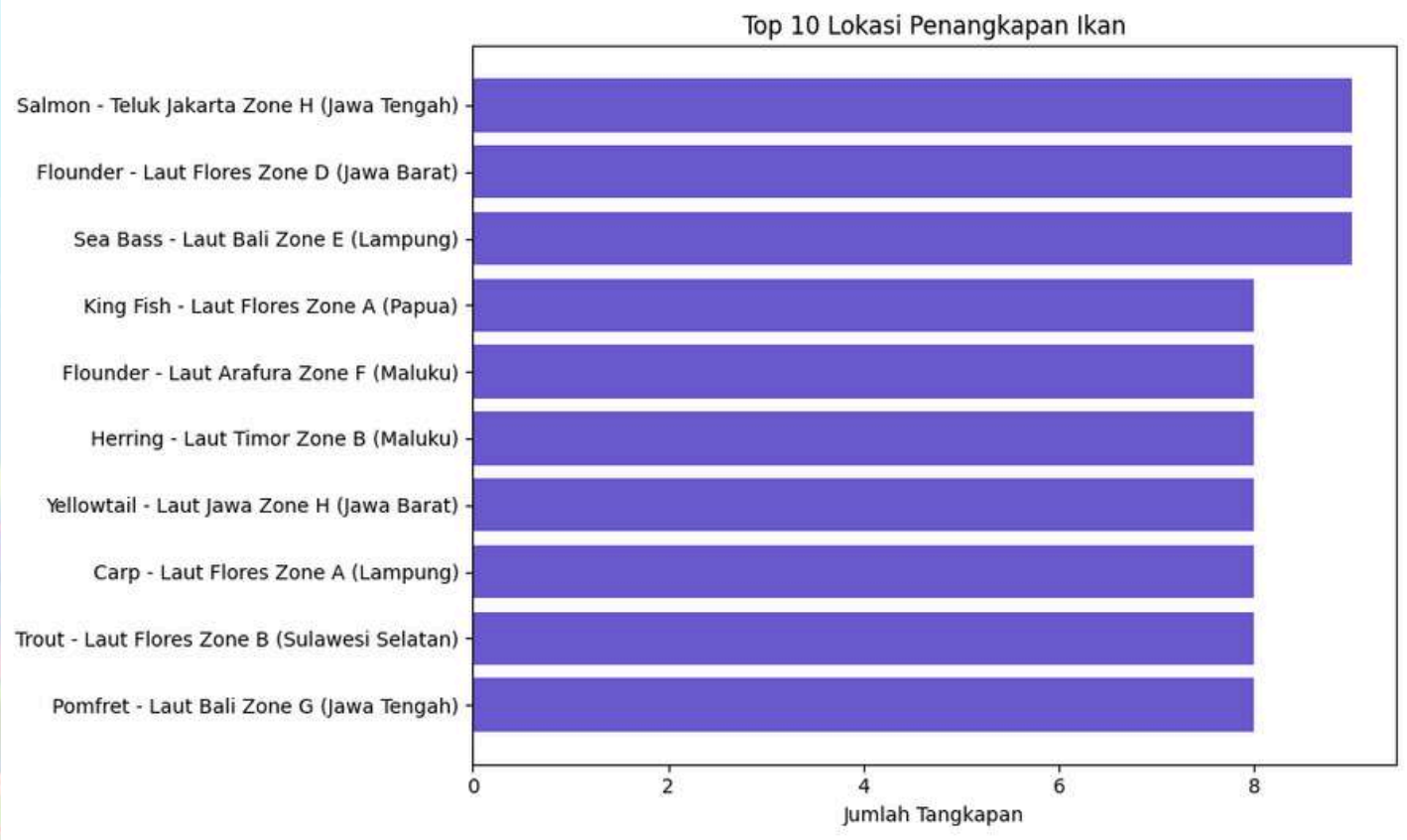
	experience_years	weight_kg
0	1	7710.83
1	2	5059.79
2	3	7332.55
3	4	6212.94
4	5	9099.47
5	6	7418.16
6	7	4941.04
7	8	4628.23
8	9	5053.69
9	10	9801.26
10	11	7741.14
11	12	7069.73
12	13	5214.02
13	14	5738.64
14	15	3543.00
15	16	5985.92
16	17	10501.93
17	18	7356.88
18	19	1665.97
19	20	8602.40
20	21	5454.66
21	22	3892.27

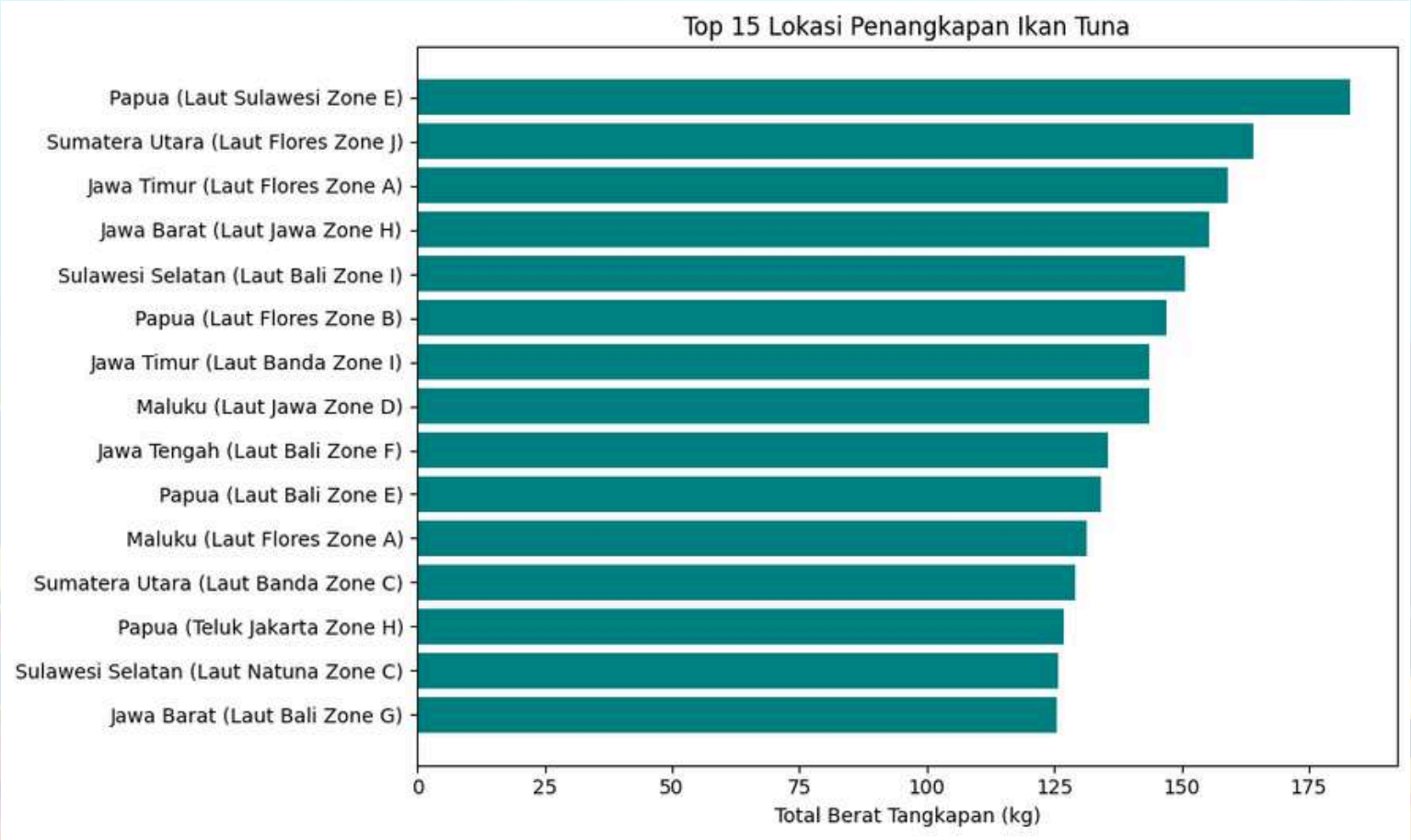

Hubungan antara Pengalaman Nelayan dan Volume Tangkapan
Total Volume Tangkapan (kg)
900
800
700
600
500
400
300
200
0
5
10
15
20
25
30
35
40
Tahun Pengalaman

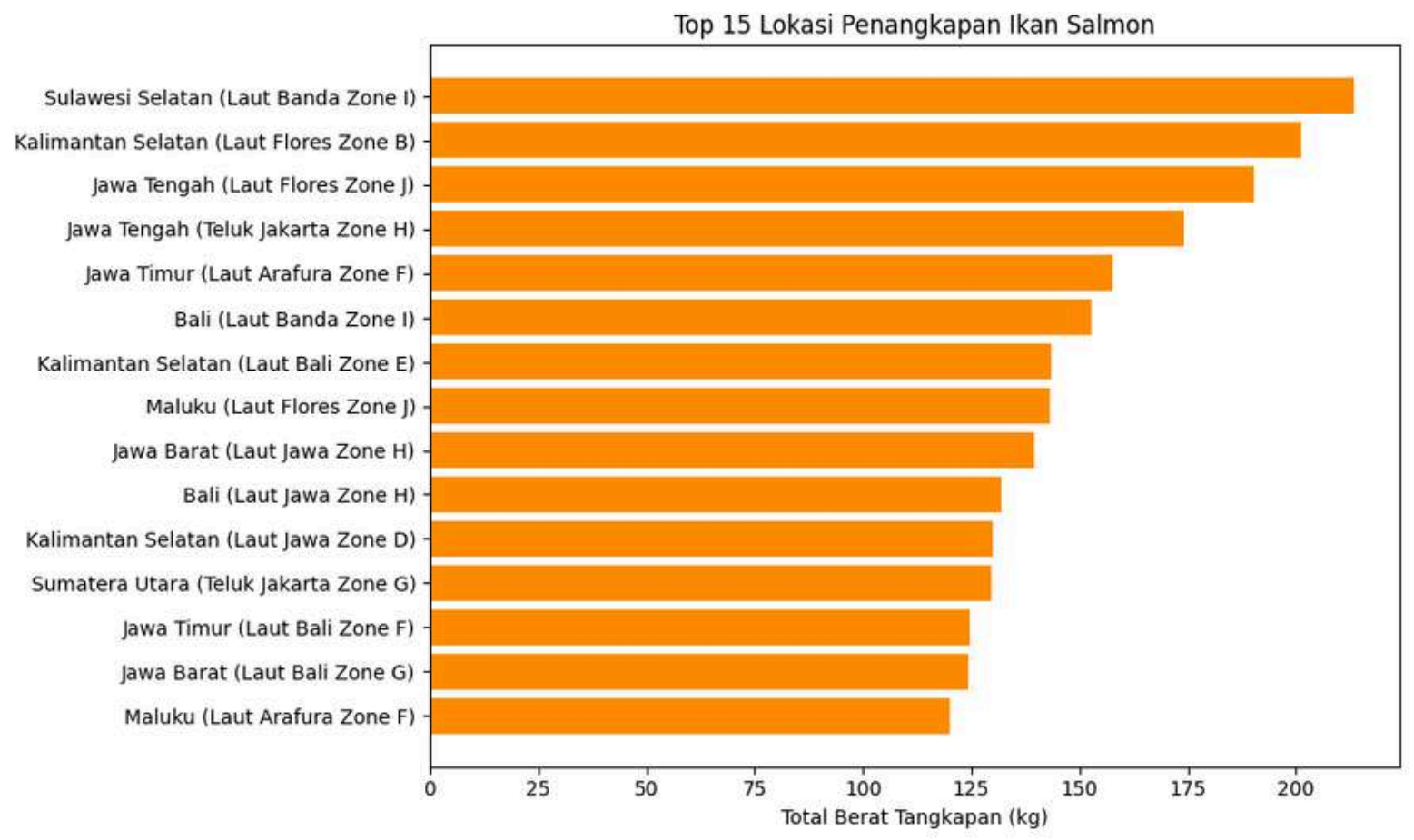
4. Distribusi Lokasi Penangkapan

	fish_name	province	location_name	jumlah_tangkapan
0	Amberjack	Bali	Laut Arafura Zone F	4
1	Amberjack	Bali	Laut Bali Zone E	5
2	Amberjack	Bali	Laut Bali Zone F	2
3	Amberjack	Bali	Laut Bali Zone G	3
4	Amberjack	Bali	Laut Bali Zone I	4
...	...	...	...	...
4206	Yellowtail	Sumatera Utara	Laut Natuna Zone C	3
4207	Yellowtail	Sumatera Utara	Laut Natuna Zone J	1
4208	Yellowtail	Sumatera Utara	Laut Sulawesi Zone E	2
4209	Yellowtail	Sumatera Utara	Laut Timor Zone B	2
4210	Yellowtail	Sumatera Utara	Teluk Jakarta Zone H	3

4211 rows × 4 columns

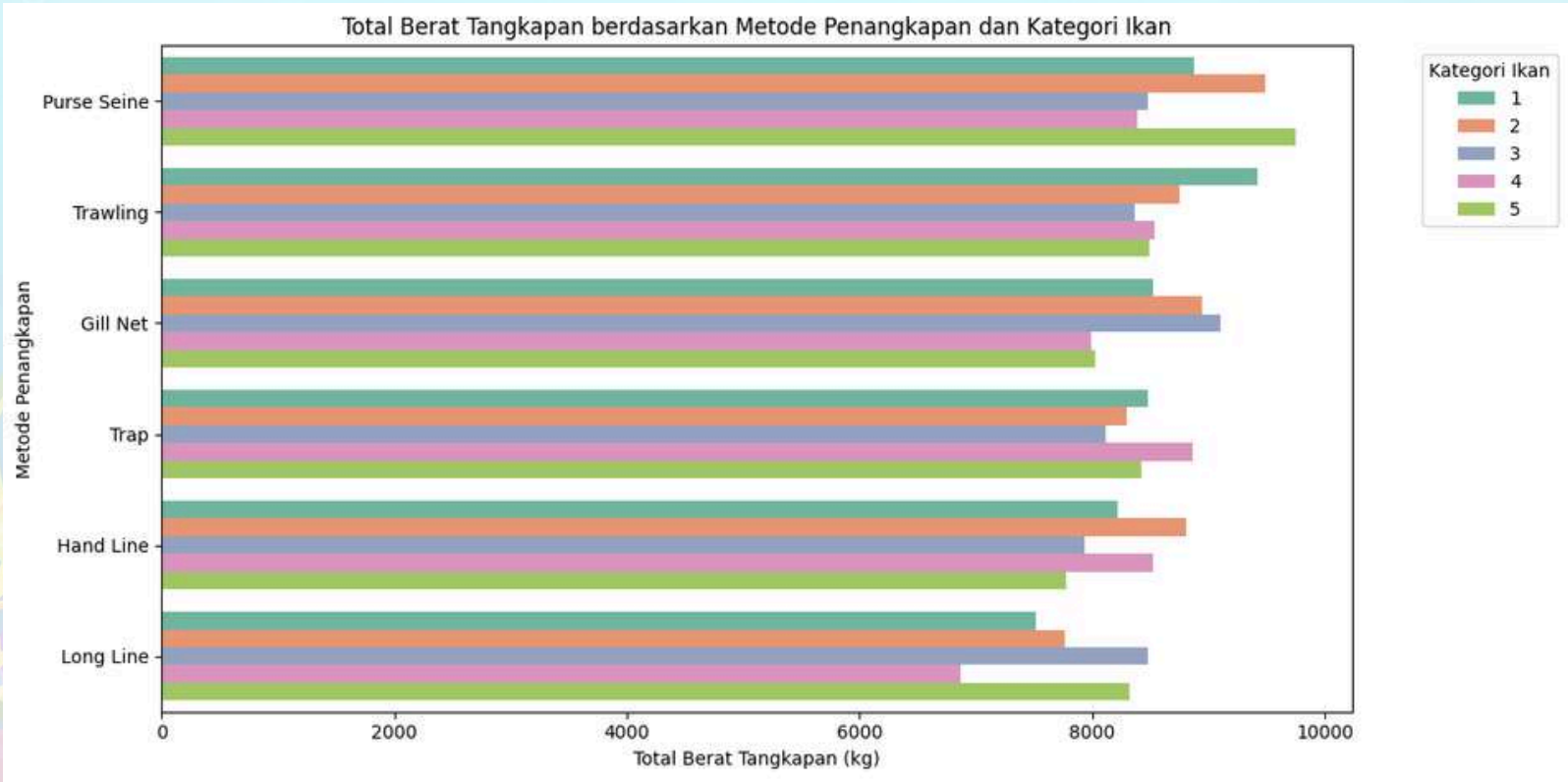

Top 10 Lokasi Penangkapan Ikan
Salmon - Teluk Jakarta Zone H (Jawa Tengah)
Flounder - Laut Flores Zone D (Jawa Barat)
Sea Bass - Laut Bali Zone E (Lampung)
King Fish - Laut Flores Zone A (Papua)
Flounder - Laut Arafura Zone F (Maluku)
Herring - Laut Timor Zone B (Maluku)
Yellowtail - Laut Jawa Zone H (Jawa Barat)
Carp - Laut Flores Zone A (Lampung)
Trout - Laut Flores Zone B (Sulawesi Selatan)
Pomfret - Laut Bali Zone G (Jawa Tengah)
0
2
4
6
8
Jumlah Tangkapan


Top 15 Lokasi Penangkapan Ikan Tuna
Papua (Laut Sulawesi Zone E)
Sumatera Utara (Laut Flores Zone J)
Jawa Timur (Laut Flores Zone A)
Jawa Barat (Laut Jawa Zone H)
Sulawesi Selatan (Laut Bali Zone I)
Papua (Laut Flores Zone B)
Jawa Timur (Laut Banda Zone I)
Maluku (Laut Jawa Zone D)
Jawa Tengah (Laut Bali Zone F)
Papua (Laut Bali Zone E)
Maluku (Laut Flores Zone A)
Sumatera Utara (Laut Banda Zone C)
Papua (Teluk Jakarta Zone H)
Sulawesi Selatan (Laut Natuna Zone C)
Jawa Barat (Laut Bali Zone G)
0
25
50
75
100
125
150
175
Total Berat Tangkapan (kg)


Top 15 Lokasi Penangkapan Ikan Salmon
Sulawesi Selatan (Laut Banda Zone I)
Kalimantan Selatan (Laut Flores Zone B)
Jawa Tengah (Laut Flores Zone J)
Jawa Tengah (Teluk Jakarta Zone H)
Jawa Timur (Laut Arafura Zone F)
Bali (Laut Banda Zone I)
Kalimantan Selatan (Laut Bali Zone E)
Maluku (Laut Flores Zone J)
Jawa Barat (Laut Jawa Zone H)
Bali (Laut Jawa Zone H)
Kalimantan Selatan (Laut Jawa Zone D)
Sumatera Utara (Teluk Jakarta Zone G)
Jawa Timur (Laut Bali Zone F)
Jawa Barat (Laut Bali Zone G)
Maluku (Laut Arafura Zone F)
0
25
50
75
100
125
150
175
200
Total Berat Tangkapan (kg)

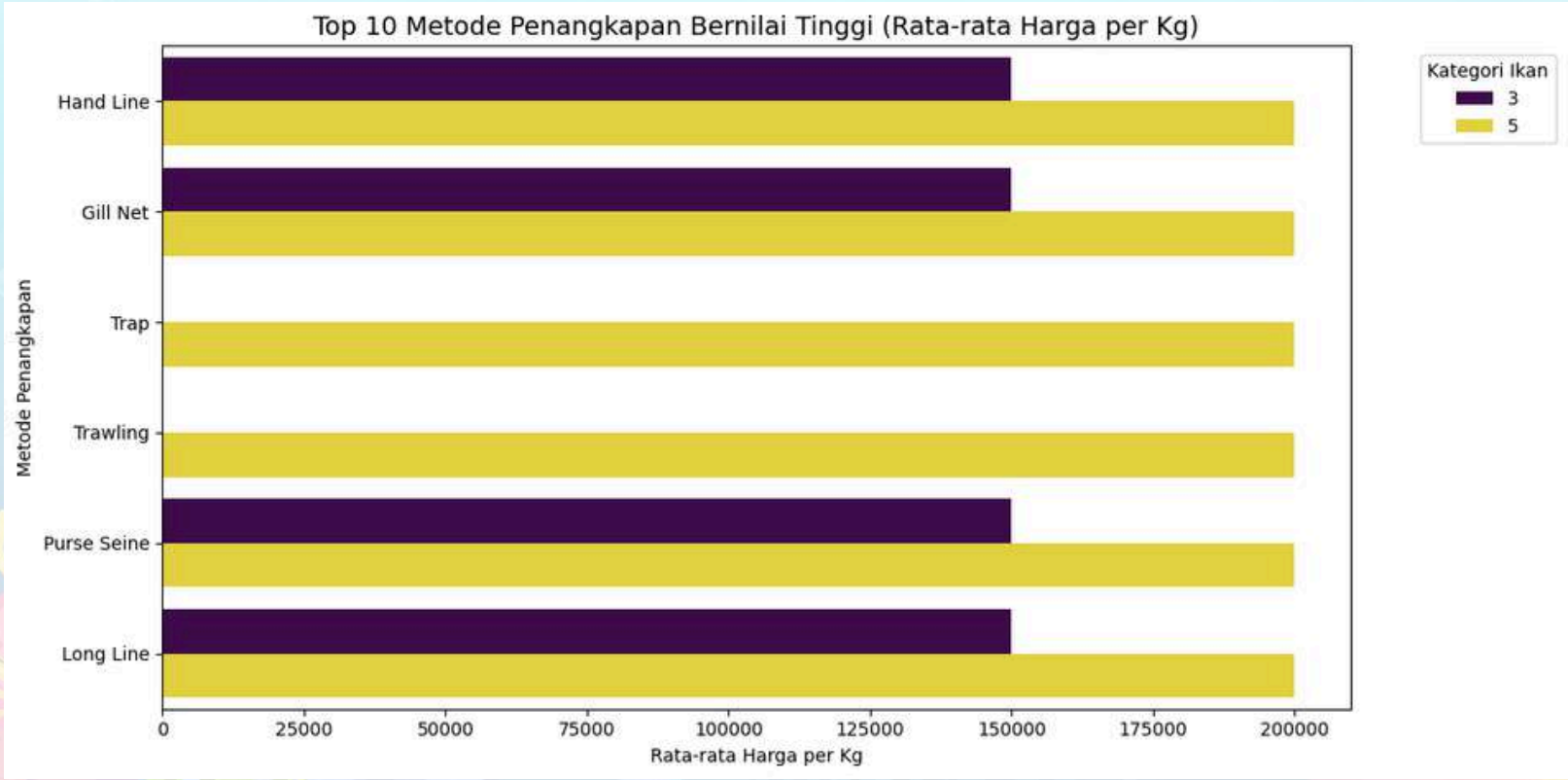
5. METODE PENANGKAPAN

	category_name	fishing_method	weight_kg
3	Anadromous Fish	Purse Seine	9753.56
9	Demersal Fish	Purse Seine	9482.41
23	Pelagic Fish	Trawling	9423.66
24	Reef Fish	Gill Net	9102.06
6	Demersal Fish	Gill Net	8946.98
21	Pelagic Fish	Purse Seine	8880.89
16	Freshwater Fish	Trap	8862.42
7	Demersal Fish	Hand Line	8810.90
11	Demersal Fish	Trawling	8754.46
17	Freshwater Fish	Trawling	8538.71
18	Pelagic Fish	Gill Net	8525.81
13	Freshwater Fish	Hand Line	8520.30
5	Anadromous Fish	Trawling	8490.90
27	Reef Fish	Purse Seine	8483.92
26	Reef Fish	Long Line	8480.87
22	Pelagic Fish	Trap	8478.37
4	Anadromous Fish	Trap	8421.22
15	Freshwater Fish	Purse Seine	8382.63
29	Reef Fish	Trawling	8364.96
2	Anadromous Fish	Long Line	8325.40
10	Demersal Fish	Trap	8292.59
19	Pelagic Fish	Hand Line	8220.87
28	Reef Fish	Trap	8116.25
0	Anadromous Fish	Gill Net	8026.29
12	Freshwater Fish	Gill Net	7993.16
25	Reef Fish	Hand Line	7929.09


Total Berat Tangkapan berdasarkan Metode Penangkapan dan Kategori Ikan
Purse Seine
Trawling
Gill Net
Trap
Hand Line
Long Line
Metode Penangkapan
0
2000
4000
6000
8000
10000
Total Berat Tangkapan (kg)
Kategori Ikan
1
2
3
4
5

6. ANALISIS KATEGORI BERNILAI TINGGI DAN METODE PENANGKAPAN

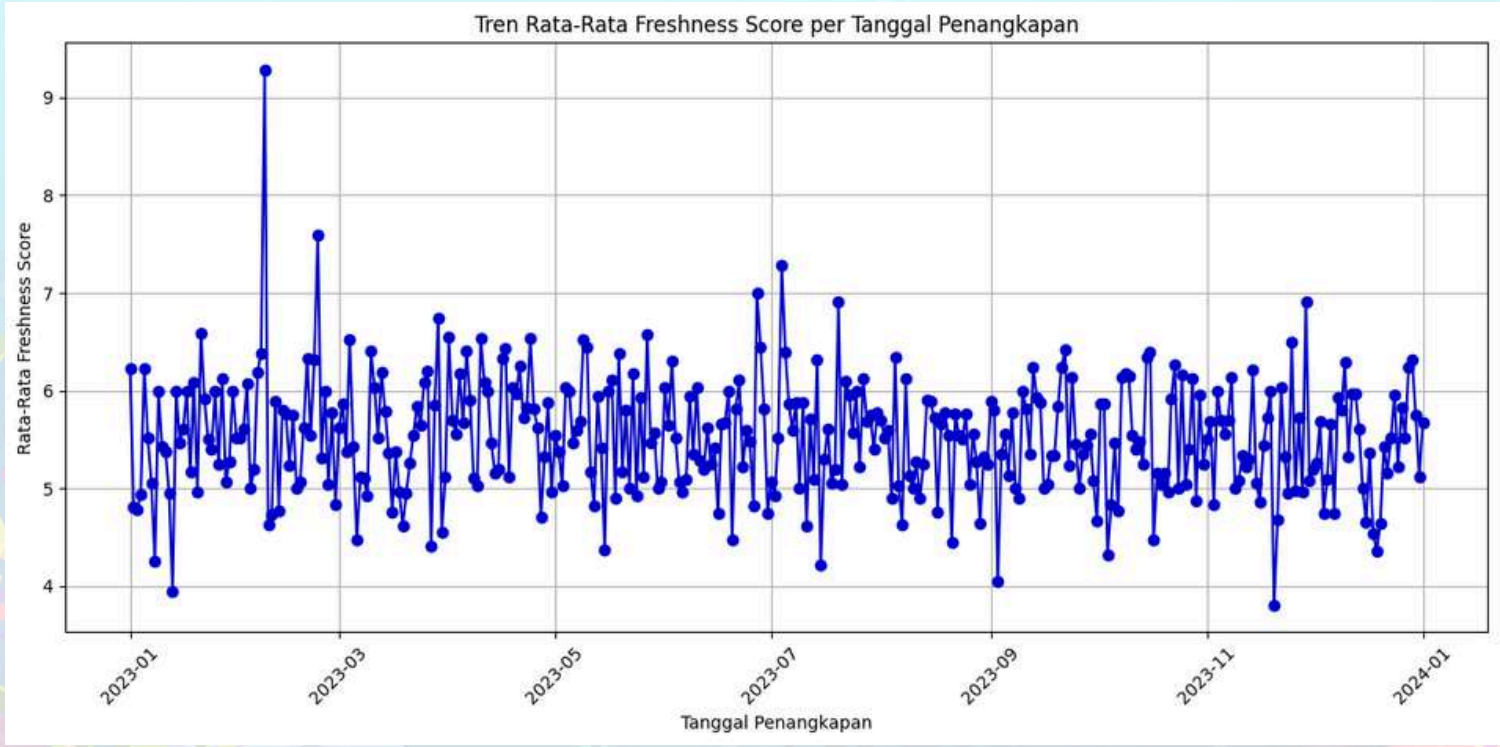
	category_name	fishing_method	avg_price_per_kg
0	Anadromous Fish	Gill Net	200000.0
1	Anadromous Fish	Hand Line	200000.0
2	Anadromous Fish	Long Line	200000.0
3	Anadromous Fish	Purse Seine	200000.0
4	Anadromous Fish	Trap	200000.0
5	Anadromous Fish	Trawling	200000.0
28	Reef Fish	Trap	150000.0
29	Reef Fish	Trawling	150000.0
25	Reef Fish	Hand Line	150000.0
24	Reef Fish	Gill Net	150000.0
27	Reef Fish	Purse Seine	150000.0
26	Reef Fish	Long Line	150000.0
6	Demersal Fish	Gill Net	120000.0
7	Demersal Fish	Hand Line	120000.0
9	Demersal Fish	Purse Seine	120000.0
8	Demersal Fish	Long Line	120000.0
11	Demersal Fish	Trawling	120000.0
10	Demersal Fish	Trap	120000.0
18	Pelagic Fish	Gill Net	85000.0
19	Pelagic Fish	Hand Line	85000.0
21	Pelagic Fish	Purse Seine	85000.0
20	Pelagic Fish	Long Line	85000.0
22	Pelagic Fish	Trap	85000.0
23	Pelagic Fish	Trawling	85000.0
13	Freshwater Fish	Hand Line	45000.0
12	Freshwater Fish	Gill Net	45000.0
14	Freshwater Fish	Long Line	45000.0


Top 10 Metode Penangkapan Bernilai Tinggi (Rata-rata Harga per Kg)
Metode Penangkapan
Hand Line
Gill Net
Trap
Trawling
Purse Seine
Long Line
0
25000
50000
75000
100000
125000
150000
175000
200000
Rata-rata Harga per Kg
Kategori Ikan
3
5

7. ANALISIS KUALITAS (FRESHNESS SCORE VS TANGGAL TANGKAP)

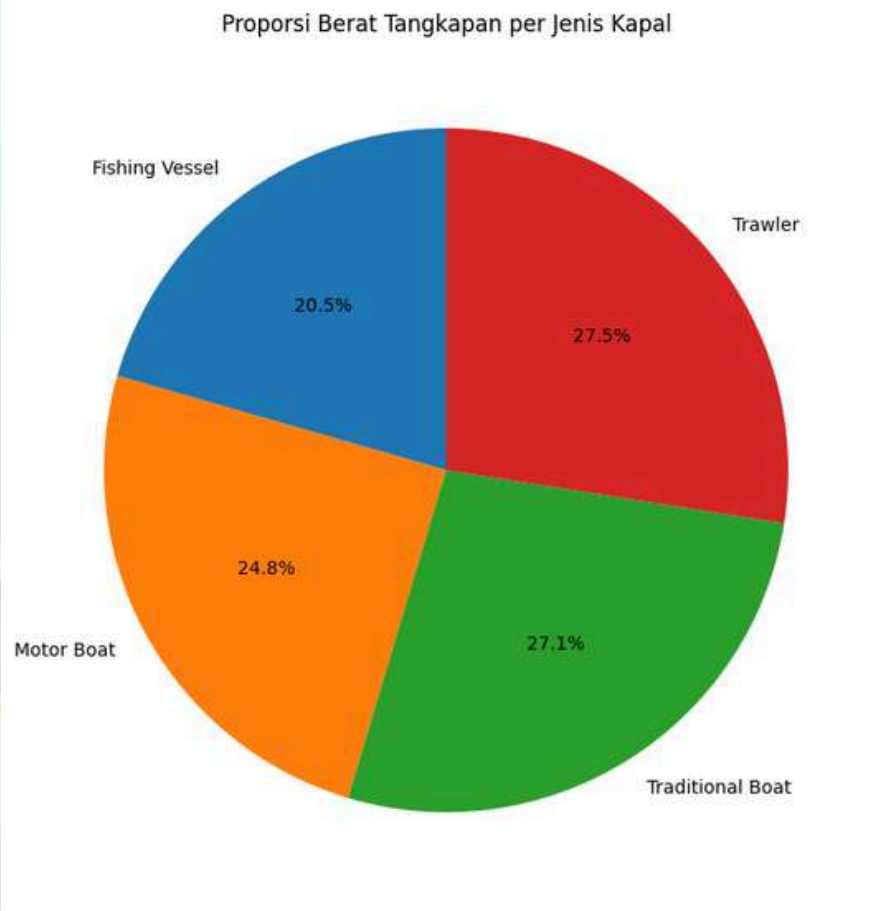
	catch_date	freshness_score
0	1/1/2023	6.222222
1	1/1/2024	5.677419
2	1/10/2023	5.428571
3	1/11/2023	5.375000
4	1/12/2023	4.954545
...	...	...
361	9/5/2023	5.552632
362	9/6/2023	5.136364
363	9/7/2023	5.777778
364	9/8/2023	5.000000
365	9/9/2023	4.906250

366 rows × 2 columns


Tren Rata-Rata Freshness Score per Tanggal Penangkapan
Rata-Rata Freshness Score
9
8
7
6
5
4
2023-01
2023-03
2023-05
2023-07
2023-09
2023-11
2024-01
Tanggal Penangkapan

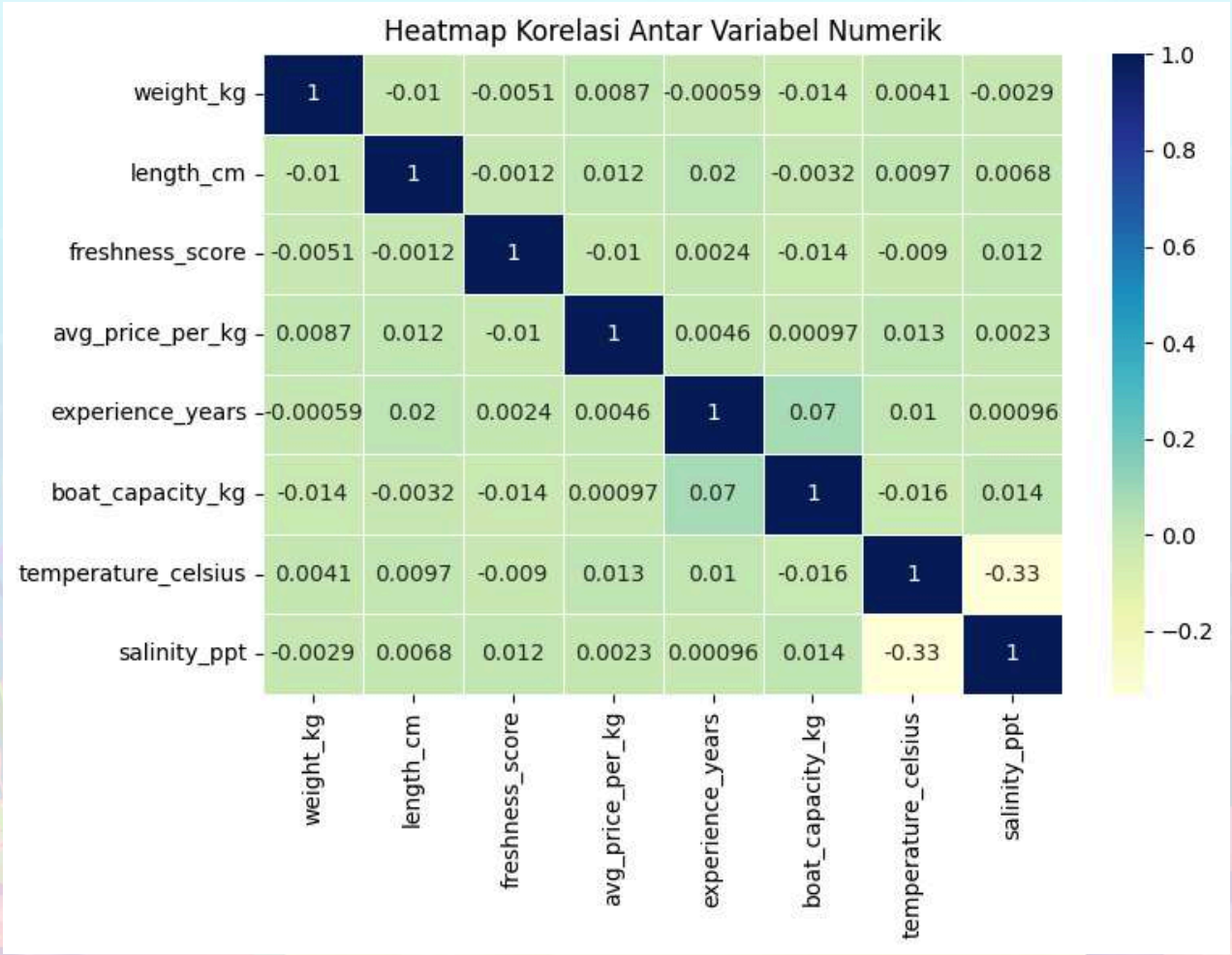
8. PROPORSI BERAT TANGKAPAN PER BERAT KAPAL

	boat_type	weight_kg
3	Trawler	69682.63
2	Traditional Boat	68810.61
1	Motor Boat	62963.26
0	Fishing Vessel	52076.88


Proporsi Berat Tangkapan per Jenis Kapal
Fishing Vessel
Trawler
20.5%
27.5%
24.8%
Motor Boat
27.1%
Traditional Boat

9. HEATMAP KOLERASI ANTAR KOLOM NUMERIK.

	salinity_ppt	experience_years	boat_capacity_kg
fish_id	0.004161	-0.007804	-0.000311
weight_kg	-0.002931	-0.000588	-0.013778
length_cm	0.006771	0.020341	-0.003182
freshness_score	0.012019	0.002443	-0.014321
category_id	-0.001092	0.020467	0.014869
location_id	-0.358897	-0.019373	0.004122
fisher_id	0.005415	0.083957	0.063009
avg_price_per_kg	0.002332	0.004589	0.000967
latitude	-0.099657	0.007545	0.002676
longitude	0.094411	-0.000829	0.005175
water_depth_m	0.393494	0.016470	0.001670
temperature_celsius	-0.330705	0.010154	-0.016027
salinity_ppt	1.000000	0.000965	0.013790
experience_years	0.000965	1.000000	0.069702
boat_capacity_kg	0.013790	0.069702	1.000000


Heatmap Korelasi Antar Variabel Numerik
weight_kg
length_cm
freshness_score
avg_price_per_kg
experience_years
boat_capacity_kg
temperature_celsius
salinity_ppt
1 -0.01 -0.0051 0.0087 -0.00059 -0.014 0.0041 -0.0029
-0.01 1 -0.0012 0.012 0.02 -0.0032 0.0097 0.0068
-0.0051 -0.0012 1 -0.01 0.0024 -0.014 -0.009 0.012
0.0087 0.012 -0.01 1 0.0046 0.00097 0.013 0.0023
-0.00059 0.02 0.0024 0.0046 1 0.07 0.01 0.00096
-0.014 -0.0032 -0.014 0.00097 0.07 1 -0.016 0.014
0.0041 0.0097 -0.009 0.013 0.01 -0.016 1 -0.33
-0.0029 0.0068 0.012 0.0023 0.00096 0.014 -0.33 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

Terima Kasih