

KEDGE BUSINESS SCHOOL

# PROJECT REPORT

FISHING VESSEL DATA ANALYSIS

**Course:** Fundamentals of Data Analytics  
for Business

**Professor:** Nafe Moradkhani

**Team member:** Ngoc Thao Uyen LE  
Aditi Prakash JIBHAKATE  
Lin YU-CHIAO  
Vaibhav RAJPUT

## 1. DESCRIPTION:

This project aims to analyze a dataset of fishing vessels to gain insights into their movement patterns, fishing activities, and behavioral anomalies. The dataset includes variables such as vessel speed, distance traveled, timestamp, and geographical coordinates (longitude and latitude). By applying data analytics techniques, we aim to understand the operational characteristics of fishing vessels, including their speed distributions, fishing hotspots, and temporal trends.

## 2. METHODOLOGY:

### 2.1. Descriptive Data Analysis:

- Calculated basic statistics (mean, median, mode, etc) for `distance_from_shore`, `distance_from_port`, `speed`, `course`, `latitude` and `longitude`.
- Identified key movement patterns and common behaviors of fishing vessels.

### 2.2. Data Visualization:

- Created histograms to examine speed and distance distributions.
- Developed scatter plots to explore the relationship between vessel speed and distance from shore.
- Constructed heatmaps to identify the most active fishing vessels and their locations.

### 2.3. Time Series Analysis:

- Analyzed trends in fishing activity over time by aggregating `is_fishing` counts per day and week.
- Identified peak fishing hours and seasonal variations in fishing activity.

### 2.4. Spatial Analysis:

- Mapped fishing activity using heat maps based on latitude and longitude.
- Visualized fishing hotspots with geographical mapping tools to highlight regions with high vessel density.

## 3. RESULTS:

### 3.1. Interpretation of Descriptive Data:

- **Distance from shore:**
  - + *Mean:* ~ 324,438 meters

- + *Standard deviation*: ~ 512,176 meters → High variation → Some vessels operate close to shore, while others go far offshore.
- + *Min*: 0 meters → Some vessels are at shore.
- + *25th Percentile (Q1)*: 0 meters → A significant number of vessels are near shore.
- + *50th Percentile (Median)*: ~38,012 meters → Half the vessels operate within 38 km from shore.
- + *75th Percentile (Q3)*: ~519 km → Many vessels operate further offshore.
- + *Max*: 3.5 million meters (~3,500 km) → Some vessels are extremely far from shore.

👉 Key Takeaway: Most vessels operate close to shore, but some travel very far offshore.

#### - Distance from Port:

- + *Mean*: ~478,630 meters
- + *Standard deviation*: 701,425 meters → High variation → Large differences in vessel proximity to ports.
- + *Min*: 0 meters → Some vessels are at port.
- + *Q1 (25%)*: ~17 km → A quarter of vessels are within 17 km of a port.
- + *Median*: ~97 km → Half of the vessels are within 97 km of a port.
- + *Q3 (75%)*: ~739 km → Some vessels operate far from ports.
- + *Max*: ~10,959 km → Some vessels operate far from any port.

👉 Key Takeaway: While many vessels operate within 100 km of a port, some are significantly farther away.

#### - Speed:

- + *Mean*: ~3.79 knots → Average vessel speed is relatively low, suggesting many are stationary or moving slowly.
- + *Standard deviation*: 5.27 → High variation, meaning some vessels move significantly faster.
- + *Min*: 0 knots → Some vessels are stationary.
- + *Q1*: 0 knots → 25% of vessels are completely still.
- + *Median*: 2.1 knots → Half of the vessels move at 2.1 knots or slower.
- + *Q3*: 7.2 knots → Most vessels travel below 7.2 knots.
- + *Max*: 102.3 knots → Some vessels travel extremely fast (possibly errors or high-speed boats).
- + *Mode*: 0 knots → The most common speed is 0 knots, meaning many vessels are stationary at times.

👉 Key Takeaway: A large portion of vessels are either stationary or moving very slowly, with only a few traveling at high speeds.

**- Course (Direction in Degrees):**

- + *Mean*:  $\sim 183.5^\circ$  → Average direction is close to due south ( $180^\circ$ ).
- + *Standard deviation*:  $114.73^\circ$  → High variation, indicating vessels travel in various directions.
- + *Min*:  $0^\circ$  → Some vessels travel directly north.
- + *Q1*:  $81.5^\circ$  → 25% of vessels travel eastward.
- + *Median*:  $193.5^\circ$  → Majority move southward.
- + *Q3*:  $282.5^\circ$  → Many vessels move westward.
- + *Max*:  $511^\circ$  → Possible erroneous values (course should range from 0-360°).
- + *Mode*:  $0^\circ$  → Most common course is  $0^\circ$  (northward movement).

👉 Key Takeaway: Vessels move in all directions, but the general trend is southward.

**- Latitude (Lat):**

- + *Mean*:  $\sim 20.11^\circ$  → Most vessels are in tropical regions.
- + *Standard deviation*:  $38.95^\circ$  → High variation (vessels operate in various latitudes).
- + *Min*:  $-83.67^\circ$  → Some vessels are near Antarctica.
- + *Q1*:  $-20.15^\circ$  → 25% of vessels are in the Southern Hemisphere.
- + *Median*:  $35.69^\circ$  → Half of the vessels are in mid-latitude regions.
- + *Q3*:  $56.00^\circ$  → Many vessels operate in the Northern Hemisphere.
- + *Max*:  $84.69^\circ$  → Some vessels operate near the Arctic.
- + *Mode*:  $57.65^\circ$  → Most common latitude suggests a high concentration of vessels around this region.

👉 Key Takeaway: Most vessels are concentrated in mid-latitude and tropical waters, but some operate near polar regions.

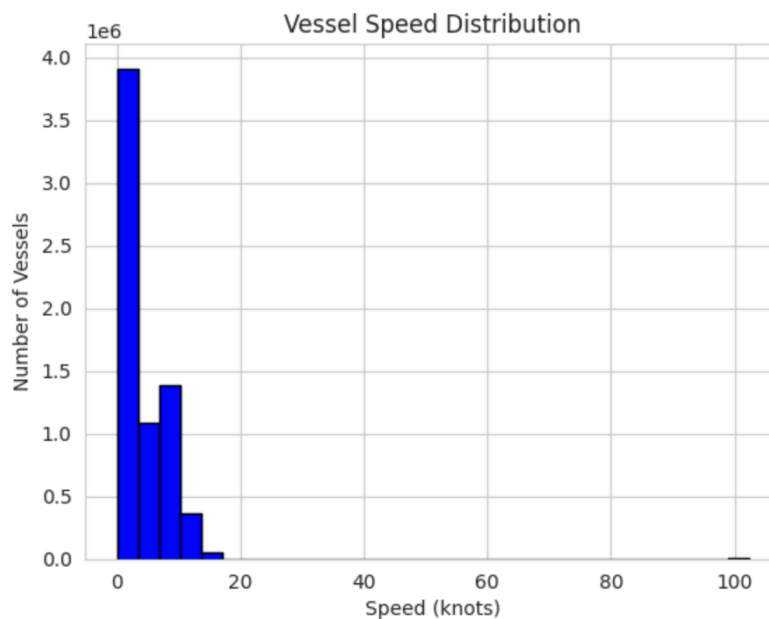
**- Longitude (Lon):**

- + *Mean*:  $\sim 16.67^\circ$  → Most vessels are in the Eastern Hemisphere.
- + *Standard deviation*:  $82.96^\circ$  → Wide global distribution.
- + *Min*:  $-179.99^\circ$  → Some vessels are near the International Date Line (Pacific Ocean).
- + *Q1*:  $-24.99^\circ$  → Many vessels are in the Atlantic Ocean.
- + *Median*:  $\sim 6.28^\circ$  → Half the vessels are near the Prime Meridian (Africa/Europe).
- + *Q3*:  $57.50^\circ$  → Many vessels are in the Indian Ocean / Asia.
- + *Max*:  $179.99^\circ$  → Some vessels are near the Pacific Ocean / Japan.
- + *Mode*:  $147.98^\circ$  → Most common longitude suggests a high concentration of vessels in the Western Pacific / Asian waters.

👉 Key Takeaway: Vessels are spread across the globe, with a notable concentration around the Western Pacific and Atlantic Oceans.

⇒ **Overall Findings:**

- Many vessels stay near shore and ports, but some operate far offshore.
- Most vessels are stationary or moving slowly, with only a few traveling fast.
- General movement trend is southward, with widespread variation in direction.
- Vessels are concentrated in mid-latitude regions, particularly in the Atlantic and Pacific Oceans.

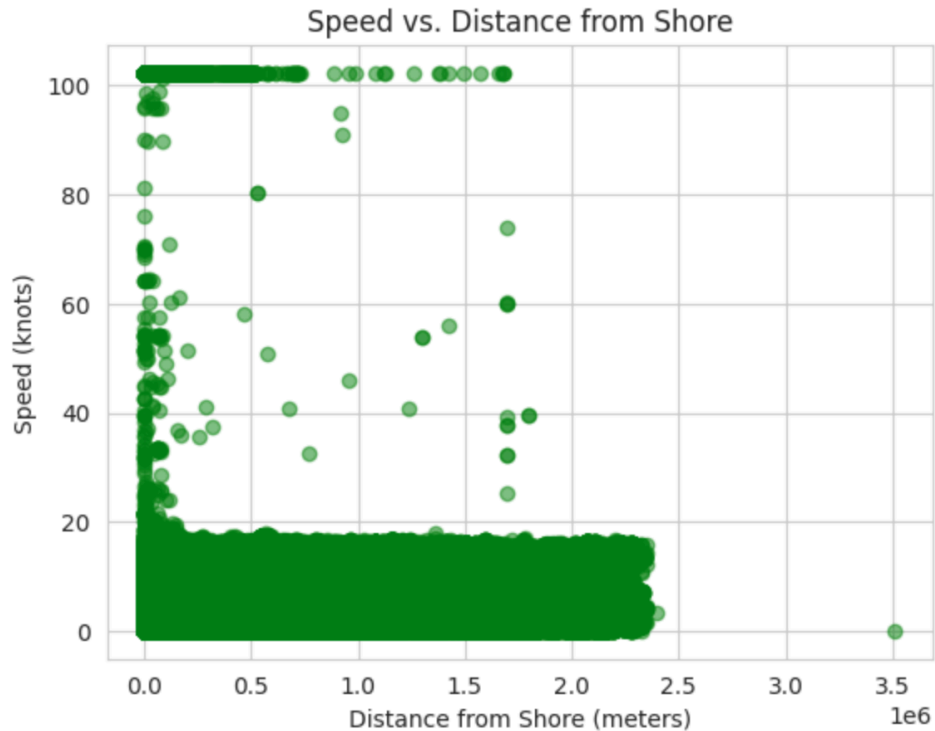
**3.2. Interpretation of Graphs and Maps:****- Histogram: Vessel Speed Distribution**

- + Most vessels are moving at low speeds (around 0-5 knots).
- + A few vessels move above 20 knots, but they are rare.

👉 **Key Takeaway:**

- + Many vessels might be stationary or moving slowly, likely fishing or waiting.
- + Only a small percentage of vessels travel at high speeds, possibly moving between fishing areas or heading to ports.

**- Scatter Plot: Speed vs. Distance from Shore**

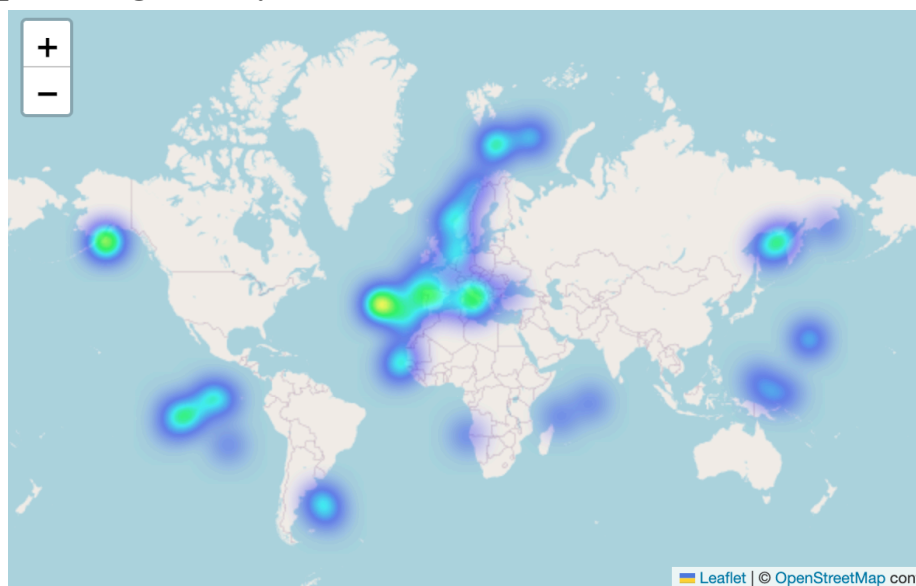


- + Vessels near the shore tend to move slower (low speeds).
- + Some vessels far from shore are moving faster, but there's variation.

👉 Key Takeaway:

- + Fishing vessels tend to operate at low speeds, even if they are far from shore.
- + Transit vessels (not fishing) may move at higher speeds, especially in open waters.
- + There could be fishing zones close to shore where vessels slow down to operate.

**- Heatmap: Fishing Activity Locations**



- + Red/yellow areas → High fishing activity (many vessels fishing here).
- + Blue/green areas → Less fishing activity.

👉 **Key Takeaway:**

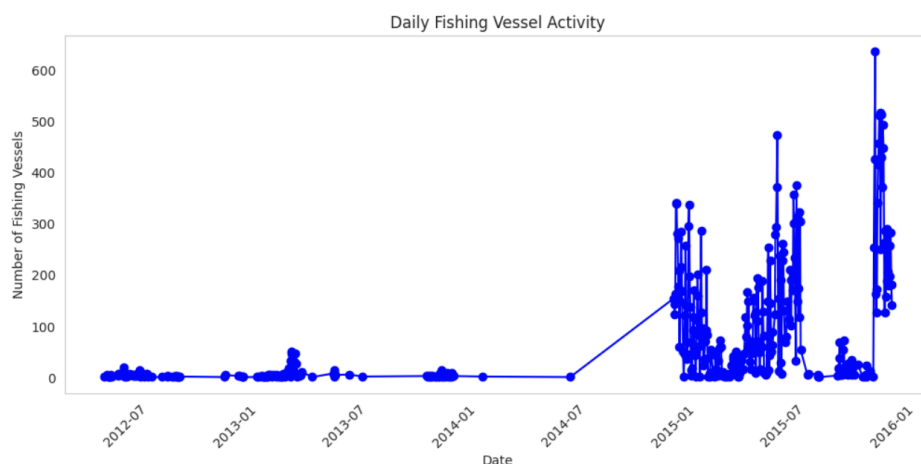
- + Clusters of fishing indicate important fishing zones.
- + If fishing areas are close to ports, they might be coastal fisheries.
- + If fishing happens far from shore, these could be deep-sea fishing grounds.

⇒ **Overall Findings:**

- Most vessels move slowly, supporting the idea that many are engaged in fishing rather than high-speed travel.
- Fishing activity is concentrated in certain areas, indicating possible regulated or preferred fishing zones.
- Distance from shore affects vessel behavior, with slower speeds near shore and mixed speeds farther out.

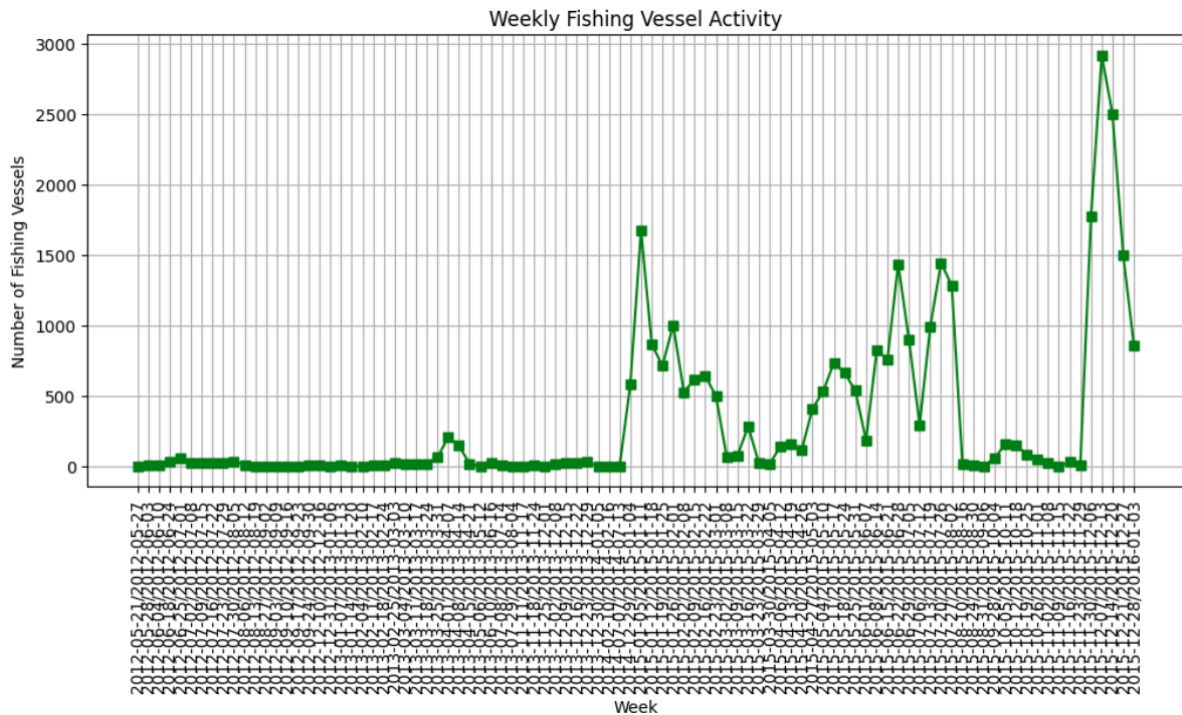
### 3.3. Interpretation of Time Series Analysis:

**- Daily Fishing Activity:**



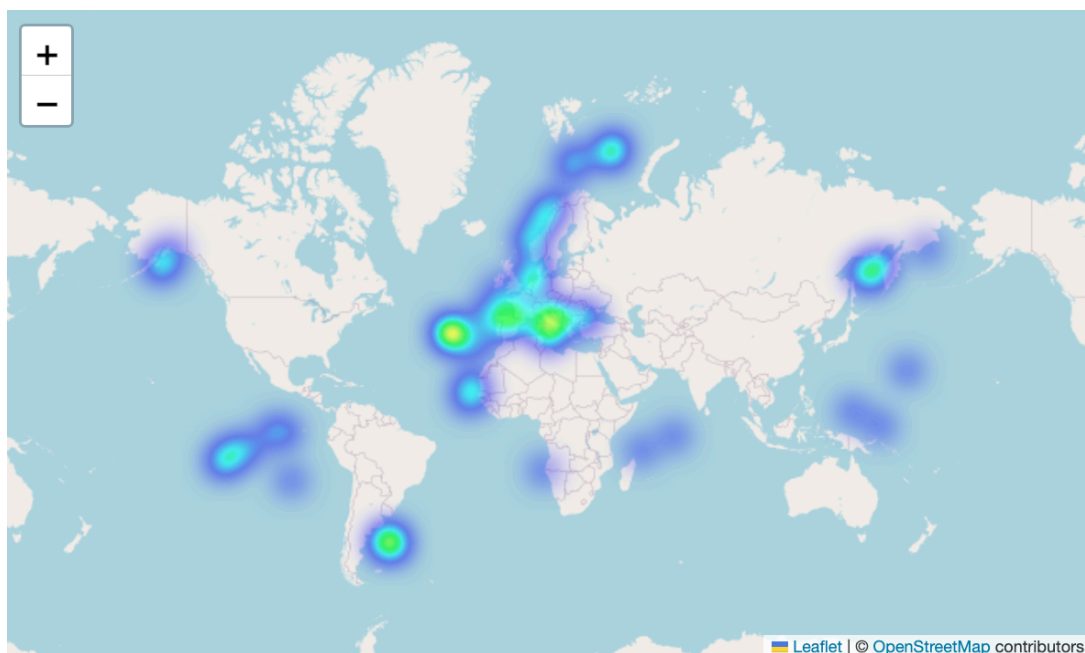
- + The daily fishing vessel count fluctuates, indicating varying fishing patterns.
- + Peaks suggest high fishing activity on certain days, possibly due to weather conditions, fishing quotas, or seasonal trends.
- + A consistent pattern could indicate operational routines of fishing fleets.

**- Weekly Fishing Activity:**



- + The weekly trend smooths out daily fluctuations, making it easier to identify long-term patterns.
- + If there is a rise in fishing vessel counts in specific weeks, this could correlate with favorable fishing conditions, regulatory changes, or seasonal fish migrations.
- + A sudden drop might indicate bad weather, closed fishing zones, or fuel price impacts.

### 3.4. Interpretation of Spatial Analysis:





**- Fishing Hotspots:**

- + The heatmap reveals concentrated areas of fishing activity, often near coastal regions, continental shelves, or known fish-rich areas.
- + High-density areas might correspond to economic fishing zones where vessels cluster for better yields.
- + If some areas show little to no fishing, they could be marine protected areas (MPAs), deep-sea zones, or areas with poor fish stocks.

**- Geographical Trends:**

- + If fishing activity is concentrated near shore, it might suggest small-scale fishing fleets operating in coastal areas.
- + If activity is offshore, it indicates industrial fishing vessels targeting deep-sea fish populations.