AIS

AIS (Automatic Identification System) data provides real-time information about a vessel's position, speed, and course, among others. AIS is essential for navigation and collision avoidance, maritime traffic monitoring, and environmental protection. This data enhances safety, efficiency, and security in maritime operations, but its limitations include potential inaccuracies due to human error and signal interference, as well as vulnerabilities to spoofing and data manipulation.

TIME:

This represents the precise date and time when the AIS (Automatic Identification System) data was recorded. It's a timestamp in a standard format.

COG (Course Over Ground):

This variable indicates the actual path the vessel is following over the Earth's surface, measured in degrees from 0 to 360. Unlike the vessel's heading, which shows where the bow is pointing, COG shows the true trajectory considering factors like current and wind. It's essential for navigation and understanding the vessel's movement direction in data analysis.

Interpretation:
- Course over ground in 1/10 = (0-359)
- 360 = not available = default.
- 360.1 - 409.5 should not be used

SOG (Speed Over Ground):

This measures the vessel's actual speed relative to the ground (seafloor), usually in knots (nautical miles per hour). It's a crucial variable for assessing the vessel's performance and progress. SOG can differ from the speed through water (STW) due to currents and tides.

Interpretation:
- Speed over ground in 1/10 knot steps (0-102.2 knots)
- 1 023 = not available, 1 022 = 102.2 knots or higher

ROT (Rate of Turn):

This indicates how quickly the vessel is changing its heading, measured in degrees per minute. A positive value indicates a turn to the right (starboard), while a negative value indicates a turn to the left (port). It's useful for understanding maneuvering behaviors and is critical in collision avoidance systems. Its calculated taking $4.733 * \sqrt{ROT\_sensor}$

Interpretation:

- 0 to +126 = turning right at up to 708 deg per min or higher
- 0 to -126 = turning left at up to 708 deg per min or higher Values between 0 and 708
- +127 = turning right at more than 5 deg per 30 s (No TI available)
- -127 = turning left at more than 5 deg per 30 s (No TI available)
- -128 (80 hex) indicates no turn information available (default).

HEADING:

This shows the direction in which the vessel's bow is pointing, measured in degrees from 0 to 360. Unlike COG, heading is not influenced by drift. It's a vital navigation parameter used alongside COG and SOG to understand vessel dynamics and control.

Interpretation:
Degrees (0-359) (511 indicates not available = default)

NAVSTAT (Navigational Status):

This variable indicates the current operational state of the vessel, such as 'Underway using engine', 'At anchor', 'Not under command', or 'Restricted maneuverability'. This status is crucial for situational awareness and compliance with maritime regulations. Worth mentioning is that some vessels use more than one "code" describing current activity, eg. most vessels use code 0 ("Under way using engine"), but from time to time also use 8 ("Under way sailing"), although only one code per given time.
More on the different codes meaning NAVSTAT

ETARAW (Estimated Time of Arrival - Raw):

This provides the expected arrival time at the destination in a raw format, typically requiring parsing into a standard datetime format. It's used for planning and operational logistics, allowing for efficient scheduling and resource allocation at ports. This estimation is set locally on the vessel and can be not updated / according to plan, or for other reasons not correct.

LATITUDE:

This is a geographic coordinate that specifies the vessel's north-south position on the Earth's surface, measured in degrees. Values range from -90° (south) to +90° (north). Latitude is essential for mapping and geospatial analysis of vessel movements.

LONGITUDE:

This is a geographic coordinate that specifies the vessel's east-west position on the Earth's surface, measured in degrees. Values range from -180° (west) to +180° (east). Longitude, together with latitude, is fundamental for accurate geolocation and route tracking.

More knowledge about the AIS standard: https://www.navcen.uscg.gov/ais-class-a-reports

vesselId:

This is a unique identifier assigned to each vessel. It's a reference key that can be used to look up detailed information about the vessel in an external file named "vessels.csv". This ID is for linking AIS data with additional vessel-specific information, such as size, weight and build year.

portId:

This is a unique identifier for ports, mapped for convenience, it's set by the captain in each vessel so it can be wrong or misleading in some cases, which will lead to wrong mapping. It's used as a proposal to identify the destination or origin port of a vessel's voyage. It is derived based on a prioritized search on the "dest" column in the original AIS data by matching on name or port codes. In the case where multiple port codes are listed in the column, the last value is used. This rule can potentially lead to some inconsistencies.

Vessels

shippingLineId:

This is a unique identifier assigned to a shipping line. It links a vessel to its owner, which can be useful for analyzing operational patterns, fleet management, and company-specific performance metrics. This company is not always the one operating the vessel.

id:

This is a unique identifier for each record in the dataset.

CEU (Car Equivalent Unit):

This measures the car-carrying capacity of a vessel. One CEU represents one standard car. It's particularly relevant for roll-on/roll-off (RoRo) ships and is used to gauge the vessel's cargo capacity in terms of vehicle units.

DWT (Deadweight Tonnage):

This represents the total weight a vessel can safely carry, including cargo, fuel, passengers, crew, and provisions. Measured in metric tons, DWT is a critical measure of a vessel's carrying capacity and overall size.

GT (Gross Tonnage):

This is a volumetric measurement of the vessel's overall internal volume, including all enclosed spaces. GT is used to determine regulatory requirements, port fees, and safety rules, reflecting the vessel's size rather than weight.

NT (Net Tonnage):

This represents the usable volume of a vessel, excluding spaces such as crew quarters and machinery spaces. NT is used to calculate harbor dues and is a measure of the earning capacity of the vessel.

vesselType:

This variable categorizes the vessel according to its primary use, such as container ship, bulk carrier, tanker, passenger ship. In this case the difference is not that significant between the selected vessels, because of similar characteristics between PCTC (Pure car and truck carriers), PCC (Pure car carrier) within the RORO (Roll on roll off) segment.

breadth:

This is the width of the vessel at its widest point, usually measured in meters. Breadth is important for stability calculations, docking requirements, and navigating through narrow channels. Most channels have strict rules for what maximum dimensions are allowed, which is the reason why most vessels' length do not surpass 200 meters.

depth:

This is the vertical distance from the keel (bottom of the hull) to the deck, measured in meters. Depth is a structural dimension that affects the vessel's capacity and stability, playing a role in the design and load distribution.

draft:

This is the vertical distance between the waterline and the bottom of the hull (keel). It indicates how deep the vessel sits in the water and is crucial for ensuring that the vessel can safely navigate through various water depths without running aground.

length:

This is the overall length of the vessel, usually measured in meters from the bow (front) to the stern (back). Length affects docking requirements, navigability, and the overall design and capacity of the vessel.

Many channels (e.g., Panama, Suez, ++) have strict rules about vessel dimensions, so the vessels' lengths do not typically exceed 200 meters for vessels within RORO.

Vessels typically fall into one of two categories: deep-sea and short-sea. Deep-sea vessels travel between continents, while short-sea vessels operate in smaller regions, often completing the final leg of a journey from larger terminals to the end customer or port.

homePort:

This denotes the port where the vessel is registered, often tied to the vessel's country of registration. The home port can influence regulatory compliance, taxation, and operational logistics.

yearBuilt:

This indicates the year the vessel was constructed. The age of the vessel can impact its operational efficiency, compliance with environmental regulations, and maintenance requirements. It's also used in market analysis and fleet modernization studies.

Ports

portId:

A unique identifier assigned to each port, used for easy reference and linking with other data, such as schedules, or mapped ais-data.

name:

The official name of the port.

portLocation:

A description or name of the specific location or area where the port is situated.

longitude

The geographic coordinate that specifies the east-west position of the port on the Earth's surface.

latitude:

The geographic coordinate that specifies the north-south position of the port on the Earth's surface.

UN_LOCODE:

The United Nations Code for Trade and Transport Locations, a unique code assigned to the port for international trade and logistics. The code is always a 5 character code, where the 2 first represent the country, and the last 3 represent the port locally. Eg. NOOSL, first two NO, stands for Norway, last three "OSL", for Oslo.

countryName:

The name of the country where the port is located.

ISO:

A two-letter code representing the country where the port is located.

Schedule

The schedule_to_may_2024.csv file contains the planned destinations, arrival and sailing time for 252 vessels. This data denotes the planned routes of each vessel as presented by the shipping lines. Each row has at least either *arrivalDate* or *sailingDate* defined, and some have both.

Vessels operate on schedules similar to bus routes, whether they are short-sea or deep-sea. They follow set routes and timings to transport goods efficiently. It's also important to note that vessels can be leased, co-shared, or otherwise managed, which can cause changes in their usual patterns and routes.

vesselId

The ID of the vessel this concerns.

shippingLineId

The shipping line that owns this vessel.

shippingLineName

The name of the shipping line that owns this vessel.

arrivalDate

The time the vessel is scheduled to arrive at the given port.

sailingDate

Time time the vessel is scheduled to depart from the port.

portName

The name of the port the vessel is scheduled for.

portId

The ID of the port the vessel is scheduled for.

portLatitude

The latitude of the port the vessel is scheduled for.

portLongitude

The longitude of the port the vessel is scheduled for.