

---

# ANALYZING AIS DATA WITH THE UN GLOBAL PLATFORM

ANN-KRISTIN KREUTZMANN

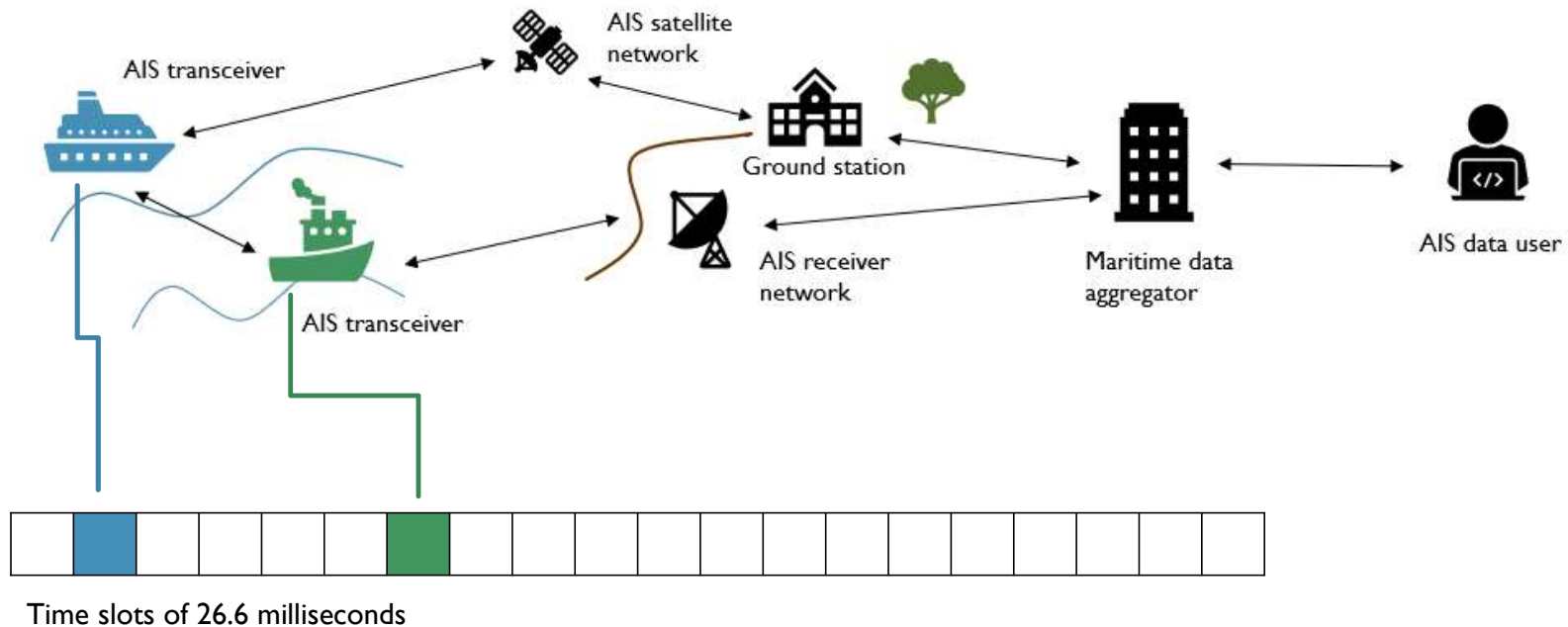


UNITED NATIONS



## WHAT IS AIS DATA – BASIC INFORMATION

- The automatic identification system (AIS) is a tracking system for ships, originally developed for collision avoidance



# WHAT IS AIS DATA – BASIC INFORMATION

## Dynamic information

- Automatically transmitted
- Every 2 to 10 seconds when moving
- Every 3 to 6 minutes when anchored

- Maritime Mobile Service Identity number (MMSI) – nine-digit number
- AIS navigational status
- Rate of turn
- Speed over ground
- Position coordinates (longitude/latitude)
- Course over ground
- Heading
- Bearing at own position
- UTC second

## Static information

- Manually transmitted
- Every 6 minutes

- International Maritime Organisation number (IMO)
- Call sign
- Name
- Type
- Dimensions
- Location of the positioning system's antenna on board the vessel
- Type of positioning system
- Draught
- Destination
- ETA (estimated time of arrival)

Reference: [MarineTraffic \(2018\). What kind of information is AIS-transmitted?](#)

---

## WHAT IS THE UN GLOBAL PLATFORM

- Global platform for
  - Exchanging ideas and methods for processing, analyzing and visualizing Big Data
  - Storing Big Data sources and enable processing, analyzing and visualizing the data
  - Demonstrating the value of Big Data in better decision making through pilots and case studies
  - Providing training materials and workshops for capacity building



---

## MOST IMPORTANT SYSTEMS FOR AIS ANALYSIS AT THE UN GLOBAL PLATFORM

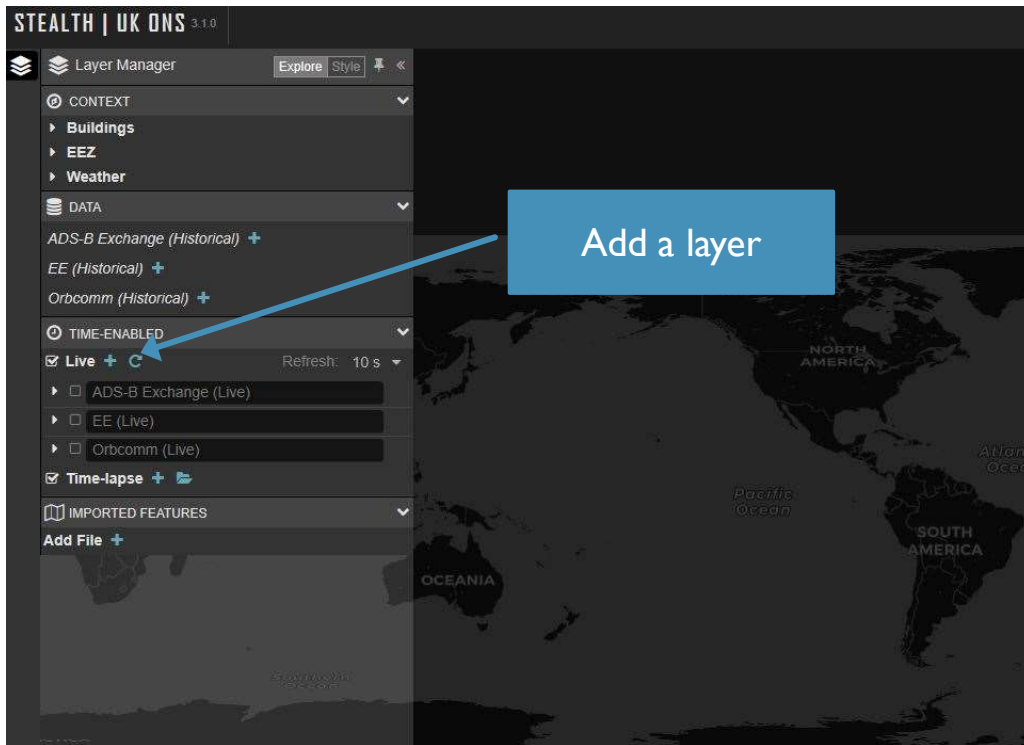
### Stealth

- <https://location.officialstatistics.org//stealth/>
- browser-based dynamic spatial visualization tool that can represent many layers of geospatial data along with their relationships, often with animation of up to millions of entities in your browser

### JupyterHub

- <https://location.officialstatistics.org/jupyter>
- Console-based, interactive web application that streamlines developing, documenting and executing code.

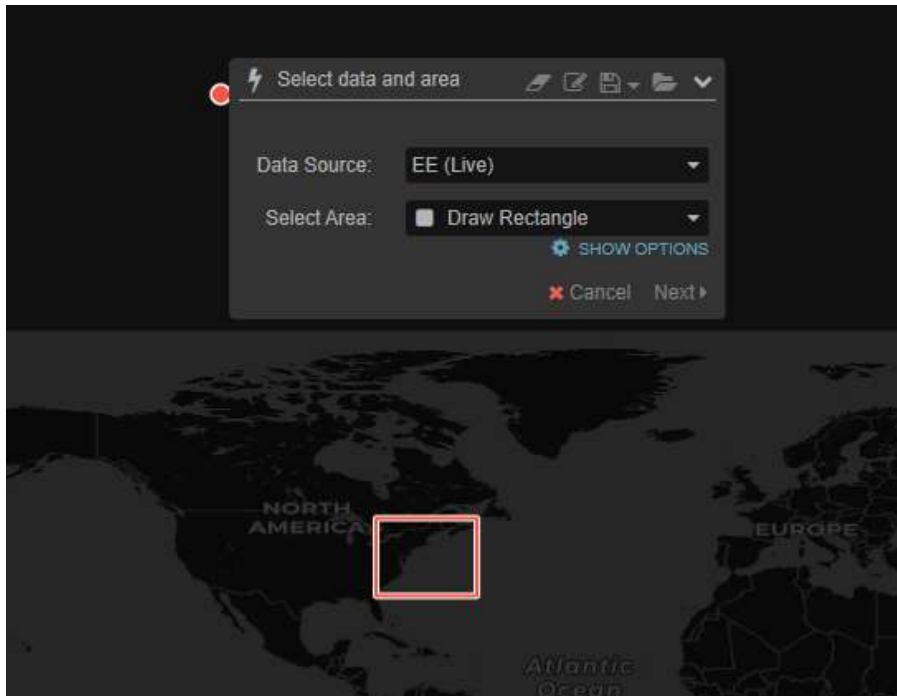
# STEALTH: HOW DOES IT WORK



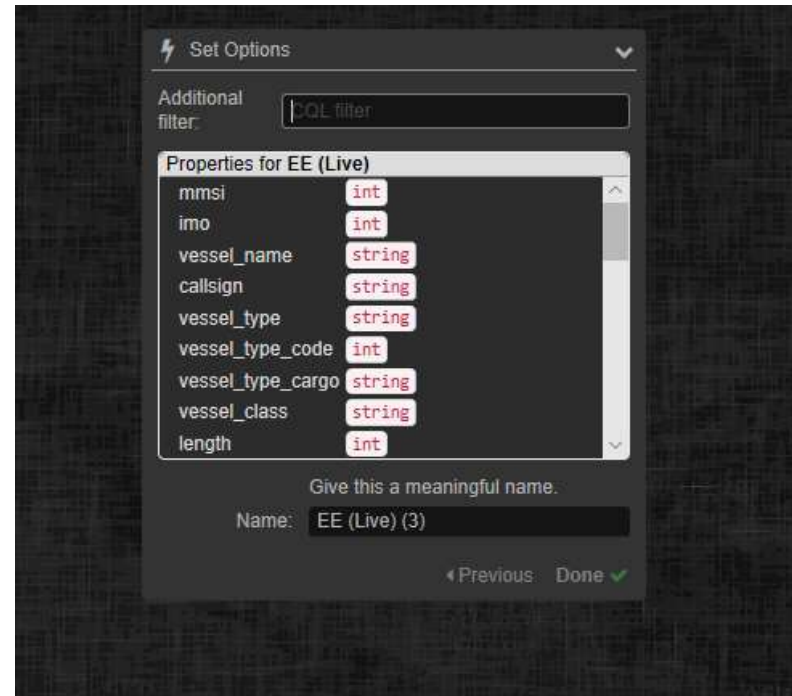
- ADS-B: current information on aircrafts
- EE: AIS data from the provider exactEarth
- Orbcomm: AIS data from the provider Orbcomm
  
- For all data sources, historical and live data is available

# STEALTH: HOW DOES IT WORK

1. Choose data source
2. Select area

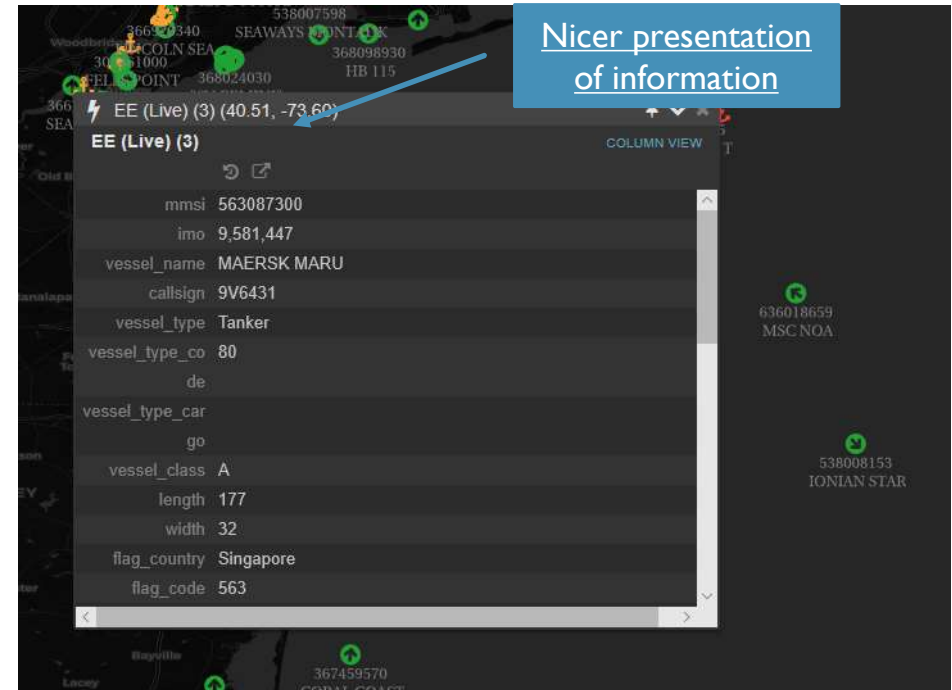
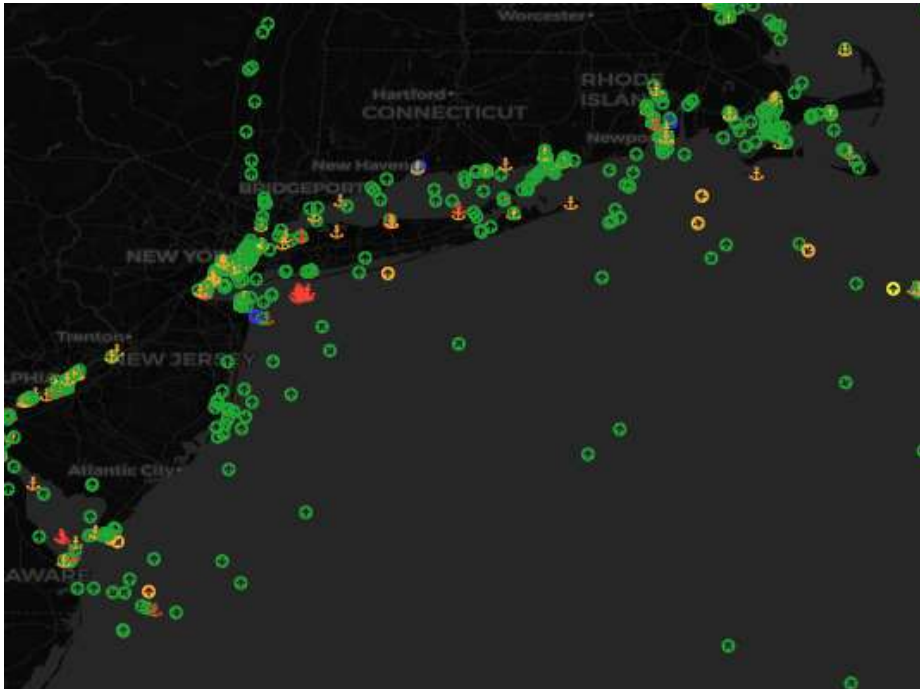


3. See and set options



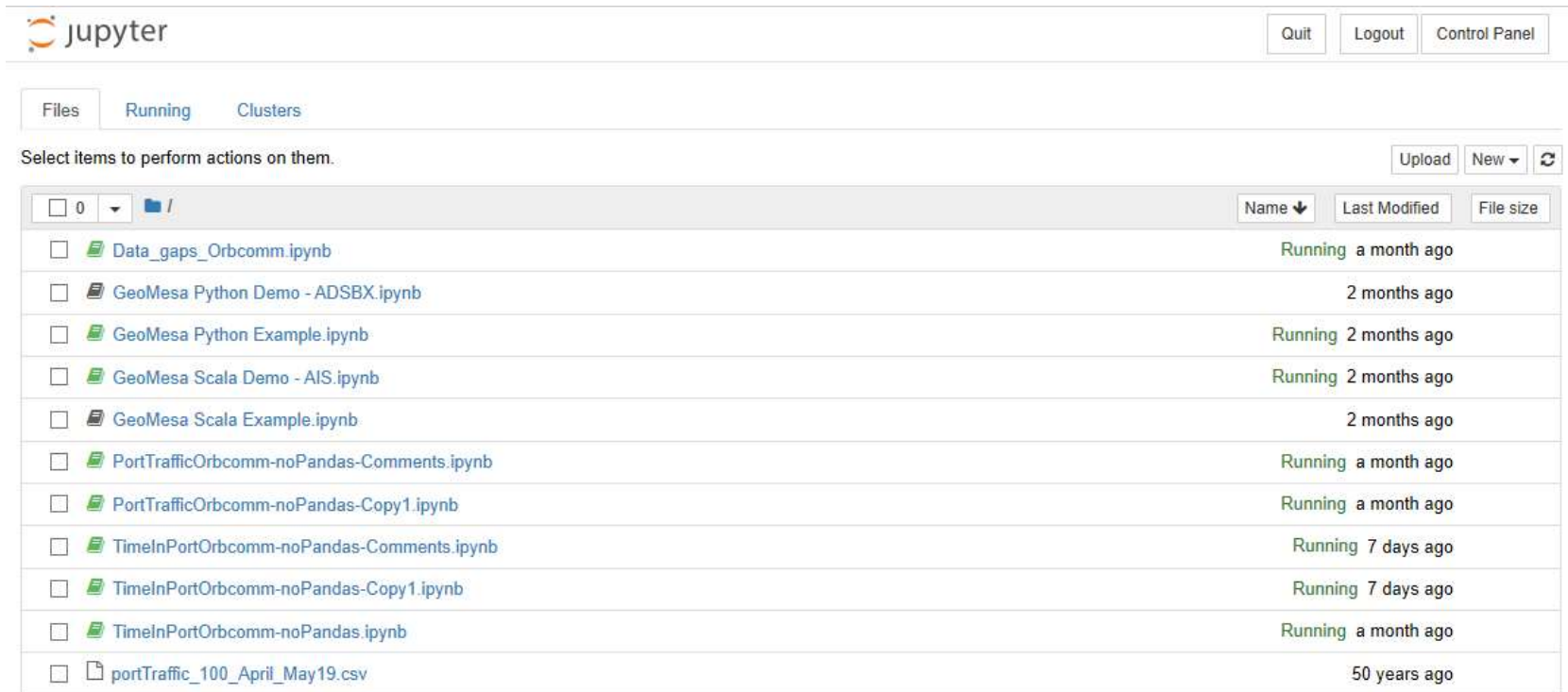
## STEALTH: HOW DOES IT WORK

- Get a first overview on ships in the selected area
- Get more information by clicking on a single vessel





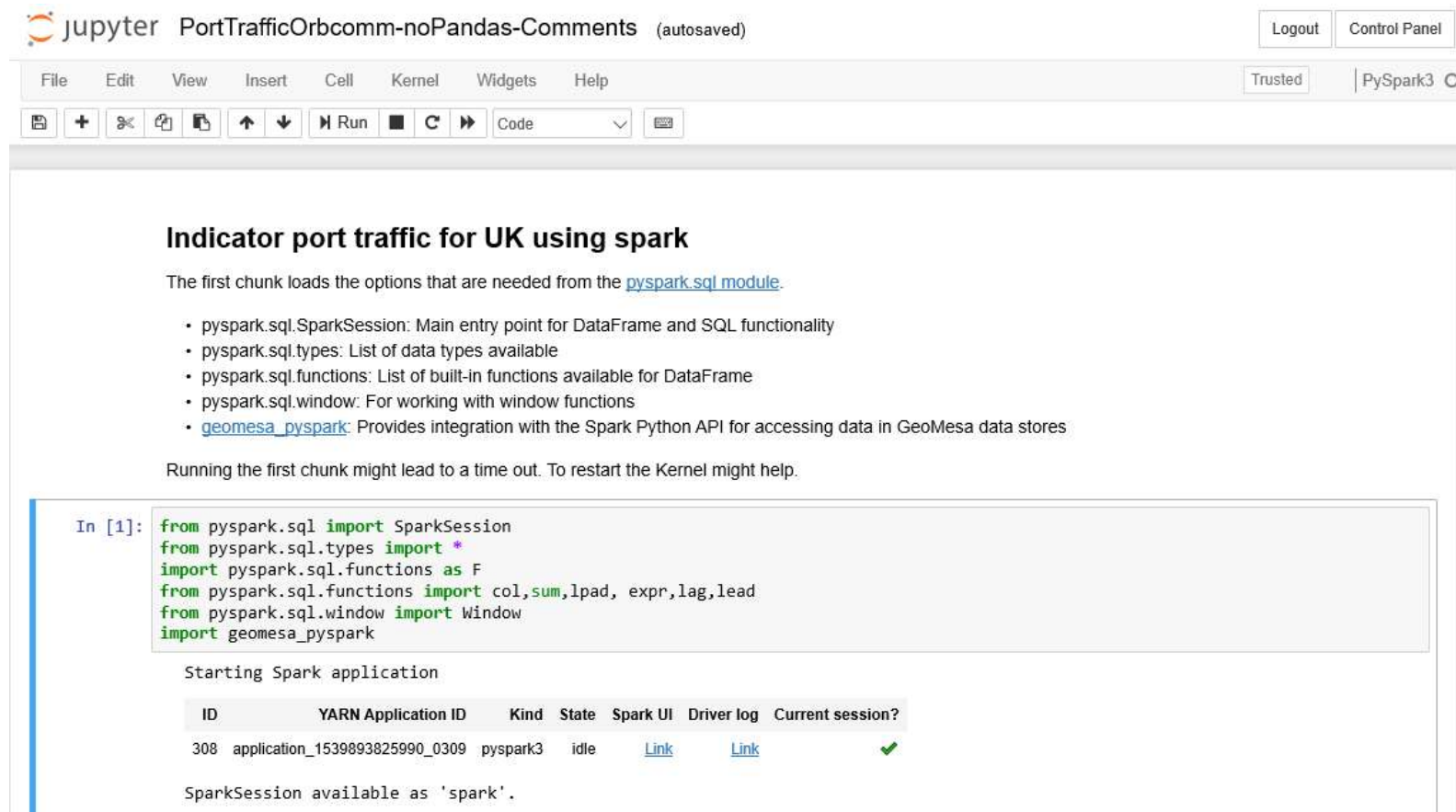
# JUPYTER HUB: HOW TO WORK WITH



The screenshot displays the Jupyter Hub interface. At the top left is the Jupyter logo. On the top right, there are buttons for 'Quit', 'Logout', and 'Control Panel'. Below the logo, there are tabs for 'Files', 'Running', and 'Clusters'. A message says 'Select items to perform actions on them.' followed by 'Upload', 'New', and a refresh icon. The main area shows a file list with columns for 'Name', 'Last Modified', and 'File size'. The file list includes several .ipynb files and one .csv file.

<input type="checkbox"/>	0	/	Name	Last Modified	File size
<input type="checkbox"/>			Data_gaps_Orbcomm.ipynb	Running a month ago	
<input type="checkbox"/>			GeoMesa Python Demo - ADSBX.ipynb	2 months ago	
<input type="checkbox"/>			GeoMesa Python Example.ipynb	Running 2 months ago	
<input type="checkbox"/>			GeoMesa Scala Demo - AIS.ipynb	Running 2 months ago	
<input type="checkbox"/>			GeoMesa Scala Example.ipynb	2 months ago	
<input type="checkbox"/>			PortTrafficOrbcomm-noPandas-Comments.ipynb	Running a month ago	
<input type="checkbox"/>			PortTrafficOrbcomm-noPandas-Copy1.ipynb	Running a month ago	
<input type="checkbox"/>			TimeInPortOrbcomm-noPandas-Comments.ipynb	Running 7 days ago	
<input type="checkbox"/>			TimeInPortOrbcomm-noPandas-Copy1.ipynb	Running 7 days ago	
<input type="checkbox"/>			TimeInPortOrbcomm-noPandas.ipynb	Running a month ago	
<input type="checkbox"/>			portTraffic_100_April_May19.csv	50 years ago	

# JUPYTER HUB: HOW TO WORK WITH



Jupyter PortTrafficOrbcomm-noPandas-Comments (autosaved) Logout Control Panel

File Edit View Insert Cell Kernel Widgets Help Trusted | PySpark3

Indicator port traffic for UK using spark

The first chunk loads the options that are needed from the [pyspark.sql module](#).

- `pyspark.sql.Session`: Main entry point for DataFrame and SQL functionality
- `pyspark.sql.types`: List of data types available
- `pyspark.sql.functions`: List of built-in functions available for DataFrame
- `pyspark.sql.window`: For working with window functions
- [geomesa\\_pyspark](#): Provides integration with the Spark Python API for accessing data in GeoMesa data stores

Running the first chunk might lead to a time out. To restart the Kernel might help.

```
In [1]: from pyspark.sql import SparkSession
from pyspark.sql.types import *
import pyspark.sql.functions as F
from pyspark.sql.functions import col,sum,lpad, expr,lag,lead
from pyspark.sql.window import Window
import geomesa_pyspark
```

Starting Spark application

ID	YARN Application ID	Kind	State	Spark UI	Driver log	Current session?
308	application_1539893825990_0309	pyspark3	idle	<a href="#">Link</a>	<a href="#">Link</a>	✓

SparkSession available as 'spark'.

# OVERVIEW AND MORE SYSTEMS/TOOLS



Platform providing data, methods and code for various fields of official statistics

Receiving and providing AIS data

Storage of data

Processing and visualizing data

Tools for analysis

**ORBCOMM**  
Data aggregator

**exactEarth**  
Data aggregator

**APACHE HBASE**  
Database storing AIS data

**CCRI**  
DATA TO KNOWLEDGE  
Service provider for advanced analytical techniques

**STEALTH**

Browser-based visualization

**APACHE SPARK**  
Analytics engine for large-scale data processing

Same credentials

**geomesa**

Tool for large-scale geospatial analytics

**jupyter**

Interactive web-application for creating documents

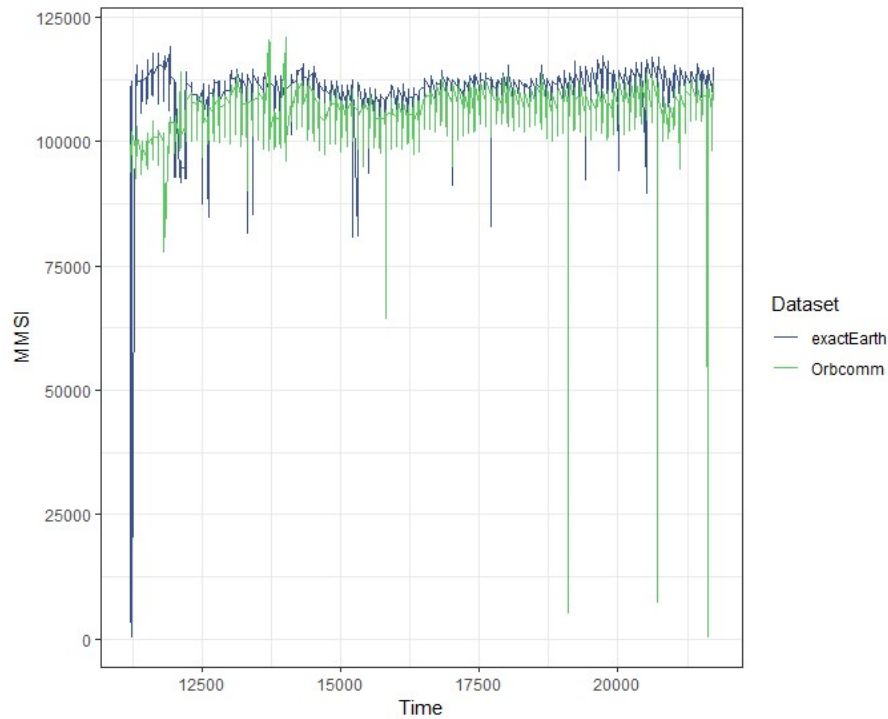
Credentials differ

**GitLab**

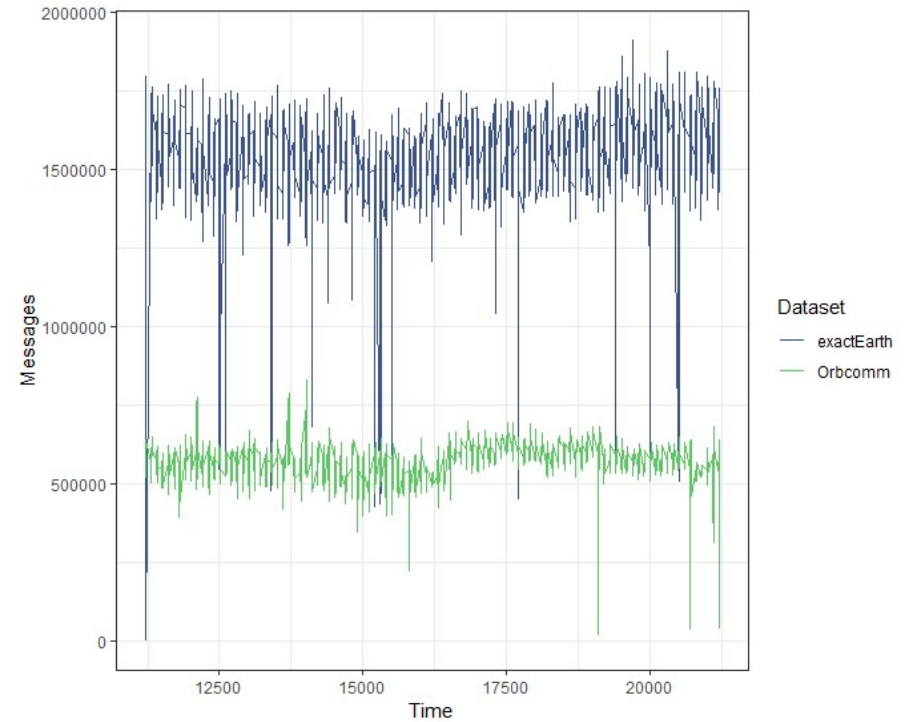
Version control and code sharing

Most apparent tools

# COMPARISON OF MARITIME DATA AGGREGATOR



	Median	Mean
exactEarth	109894	109089
Orbcomm	105550	104976



	Median	Mean
exactEarth	1548598	1527333
Orbcomm	574140	571359