



POWERED BY SIMPLON

Exploratory study of Port of Antwerp inland dataset for years 2015 & 2019.

Elaborated by:

Becode. Students AI training program.

Cohort ANT-Theano-1-21

August, 2020

1. Data description

The following report contains an exploratory analysis of the data contained in the file “CASD-4113 Smart Shipping Studie Anomized.xlsx”. A total of 174,977 rows and 12 columns provide a register of the operations for Antwerp port in the years 2015 and 2019. The dataset was studied using Python tools and libraries with the intention of finding insights and correlations between data that could be helpful for stakeholders.

The dataset “CASD-4113 Smart Shipping Studie Anomized.xlsx” contains information associated with operations for Antwerp’s port in the years 2015 and 2019. The following data is provided:

Aanvang jaar	Recorded year, Data available for 2015 and 2019.
Verblijf	Visit ID 1
Reis	Visit ID 2
FD nummer ANO	Ship’s Unique Identification number
Max lengte	Max. length of the ship
Productcode	Unique product identifier for transported goods
Product	Name of the transported goods
DWT	Deadweight tonnage, max. carry capacity of the ship
Haven herkomst	Origin of current route
Haven bestemming	Destination of current route
Route categorie	Kind of route (arriving / departing / in transit)
Scheepstype	Type of ship (19 different types)

Table 1.1. Label names and the meaning and/or interpretation of those labels

In order to facilitate the visualization of results due to the existence of different ship types and thousands of transported goods, in some cases these are classified into the groups described below:

Ship Class	Categorization of ships by length
Klasse 1 (SPITS)	.. - 38,5 m
Klasse 2 (KEMP)	38,51 - 55 m
Klasse 3	55,01 - 85 m
Klasse 4	85,01 - 105 m
Klasse 5	105,01 - 135,00 m
Klasse 6-7	135,01 - .. m

Table 1.2 Categorization of ships by length

The data in the Type of ship column (*Scheepstype*) contains 19 unique types of ship. We were advised to reduce this list to just 9 ship types, in the following fashion:

Type of ship	Simplified Type
MOTORTANKSCHIP	Liquid bulk
BUNKERSCHIP (BINV)	Liquid bulk
MOTORVRACHTSCHIP	Dry bulk
VRACHTDUWBAK (VDB)	Push boat (Dry bulk)
TANKDUWBAK (TDB), DROGE LADING	Push boat (Liquid bulk)
CONTAINERSCHIP	Container ship
MOTORTANKSCHIP, VL. LADING TYPE C	Liquid bulk
WERKVAARTUIG (BINV)	Offshore ship
PONTON	Other
MOTORTANKSCHIP, DROGE LADING	Dry bulk
SLEEP-VRACHTSCHIP	Push boat (Dry bulk)
MOTORTANKSCHIP, VL. LADING TYPE N	Liquid bulk
DUWBAK, CHEMISCH	Push boat (Liquid bulk)
DUWBOOT LOSVAREND	Push boat (Dry bulk)
GAS-TANKSCHIP	Liquid bulk
TANKDUWBAK (TDB), DROGE LADING	Dry bulk
TANKER, (GEEN GAS) (ZEE)	Liquid bulk
SLEEPBOOT LOSVAREND	Tugboat

Table 1.3 Categorization of ship types by carried product

Anomalies in the data

Ship mapping data have two values for "TANKDUWBAK (TDB) DROGE LADING", Dry Bulk & Push boat:

Ship	Ship_new	Category
MOTORTANKSCHIP	Liquid bulk	1
BUNKERSCHIP (BINV)	Liquid bulk	1
MOTORTANKSCHIP, VL. LADING TYPE C	Liquid bulk	1
MOTORTANKSCHIP, VL. LADING TYPE N	Liquid bulk	1
GAS-TANKSCHIP	Liquid bulk	1
TANKER, (GEEN GAS) (ZEE)	Liquid bulk	1
CONTAINERSCHIP	Container ship	2
MOTORVRACHTSCHIP	Dry bulk	3
MOTORTANKSCHIP, DROGE LADING	Dry bulk	3
TANKDUWBAK (TDB), DROGE LADING	Dry bulk	3
WERKVAARTUIG (BINV)	Offshore ship	4
VRACHTDUWBAK (VDB)	Push boat (Dry bulk)	5
SLEEP-VRACHTSCHIP	Push boat (Dry bulk)	5
DUWBOOT LOSVAREND	Push boat (Dry bulk)	5
TANKDUWBAK (TDB), DROGE LADING	Push boat (Liquid bulk)	6
DUWBAK, CHEMISCH	Push boat (Liquid bulk)	6
SLEEPBOOT LOSVAREND	Tugboat	7
PONTON	Other	8

Fig 1.4 Duplicate interpretations for the same ship type

Missing Values

Aanvang jaar	0
Verblijf	0
Reis	0
FD nummer ANO	0
Max lengte	0
Productcode	0
Product	0
DWT	0
Haven herkomst	1
Haven bestemming	10
Route categorie	0
Scheepstype	0
dtype:	int64

Fig 1.5 Two labels contain rows where no data was recorded for that ship

Different ship lengths were found for the same ship

Example: ship FD008 with two different lengths reported (52.61 meters and 73.6 meters)

The dataset contains 5096 unique ships and 194 out of 5096 ships have more than 1 value for length. This doesn't necessarily mean the physical length of the ship changed all of a sudden, but can be explained by the additional length of a possible tug/push boat being recorded by the port.

2. Data analysis and results

Top products registered in the port

The port's traffic is associated with a large set of products classified by name and ID. The biggest contributions to port traffic can be attributed to only a few products,

For instance:

- 8609000000 CONTAINERS
- 1202002 DIESELOLIE OF GASOLIE OF STOOKOLIE

Together, Containers and Dieselolie, account for 40% of total products that passed through the harbour.

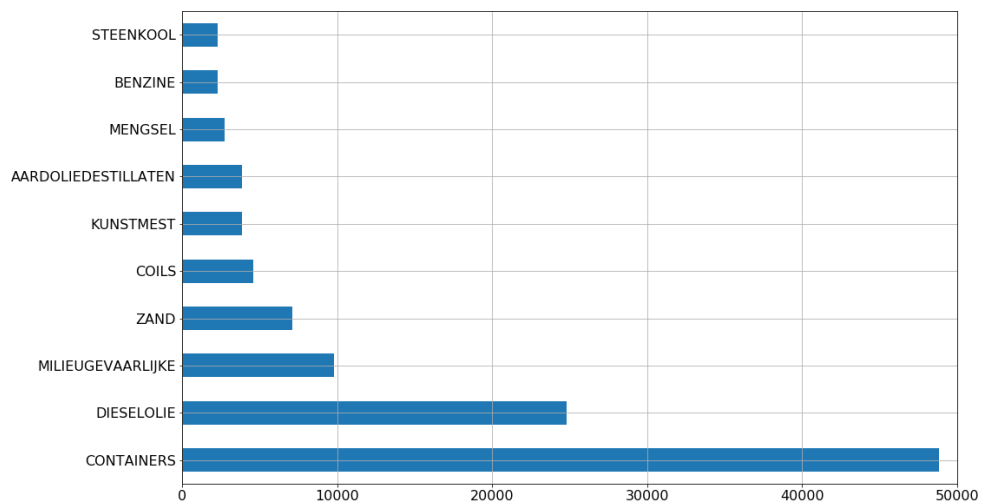


Fig 2.1 Principal products registered historically in the port

Note that these values are representative of the amount of times a ship carrying that kind of cargo was observed. It doesn't account for the amount of tons of products, since the actual weight of the ship's cargo is missing in the dataset.

How do the most observed products change between 2015 and 2019?

good_id	good_name	quantity_2015
8609000000	CONTAINERS	24261
1202002	DIESELOLIE OF GASOLIE OF STOOKOLIE, LICHT (VLAMPUNT HOGER DAN 61°C, DOCH TEN HOOGSTE 100°C)	11622
3082000	MILIEUGEVAARLIJKE VLOEISTOF, N.E.G.	4238
2505900000	ZAND OVERIG, ZEE, ZOUT ETC [O.A. ZIRCONIUMZAND]	3169
7213000000	COILS, WALSDRAAD VAN IJZER OF VAN NIET-GELEGEERD STAAL	1960
1965000	MENGSEL VAN KOOLWATERSTOFGASSEN, VLOEIBAAR GEMAAKT, N.E.G. (MENGSEL A, A01, A02, A0, A1, B1, B2, B OF C)	1413
3103000000	KUNSTMEST, FOSFAATMESTSTOFFEN	1272
1268005	AARDOLIEDESTILLATEN, N.E.G. OF AARDOLIEPRODUCTEN, N.E.G. [INCLUSIEF NAFTA - NAPHTA]	1265
1203000	BENZINE OF MOTORBRANDSTOF	1242
1001000000	TARWE EN MENGKOREN	1118

Fig 2.2 Top ten products registered in 2015

good_id	good_name	quantity_2019
8609000000	CONTAINERS	22721
1202002	DIESELOLIE OF GASOLIE OF STOOKOLIE, LICHT (VLAMPUNT HOGER DAN 61°C, DOCH TEN HOOGSTE 100°C)	11617
3082000	MILIEUGEVAARLIJKE VLOEISTOF, N.E.G.	5524
2505900000	ZAND OVERIG, ZEE, ZOUT ETC [O.A. ZIRCONIUMZAND]	2810
7213000000	COILS, WALSDRAAD VAN IJZER OF VAN NIET-GELEGEERD STAAL	2643
1268001	AARDOLIEDESTILLATEN, N.E.G. OF AARDOLIEPRODUCTEN, N.E.G. (DAMPDRUK BIJ 50°C HOGER DAN 110 KPA, DOCH TEN HOOGSTE 175 KPA)	1577
3103000000	KUNSTMEST, FOSFAATMESTSTOFFEN	1355
1965000	MENGSEL VAN KOOLWATERSTOFGASSEN, VLOEIBAAR GEMAAKT, N.E.G. (MENGSEL A, A01, A02, A0, A1, B1, B2, B OF C)	1309
1203000	BENZINE OF MOTORBRANDSTOF	1067
2530900000	GROND, AARDE, MINERALE STOFFEN NIET ELDERS ONDER BEGREPEN [INCL. BAGGERSPECIE; AL DAN NIET VERVUILD]	1043

Fig 2.3 Top ten products registered in 2019

- Containers remain as the top product registered in the port for 2015 and 2019, however a decrease in units of 6% can be perceived in 2019 with respect to 2015.
- There is no remarkable difference in the top 5, but after that a few new categories of product appear in the table, while others seem to have dropped in frequency.
- As described previously, most of the cargo is classified as containers & diesel oil. Analysis of data according with a simplified classification (*Table 1.3*), shows that almost 50% of the total ship cargo is listed as dry bulk (*this result might be due to containers being classified as dry bulk*). See Fig 2.4

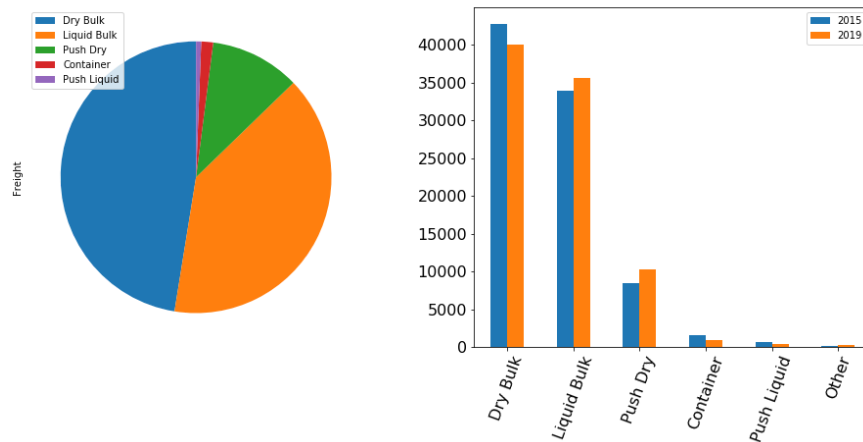


Fig 2.4 Distribution of ships by carried goods

Ship visits to the Port of Antwerp

When grouped together according to the classification in *Table 1.2*, the following information can be found:

Class Year	1	2	3	4	5	6
2015	1436	5690	23575	19138	37502	41
2019	2892	5345	19364	18998	40639	177

Table 2.5 Ship class of port visitors in 2015 and 2019

- Class 5 ships were the most common visitors to the Port, both in 2015 and 2019.
- The larger Class 6 ships were the least common visitor.
- Between 2015 and 2019, there appears to have been a drop in Class 3 ships visiting the Port.
- The middle length ships (Class 2 - Class 3 - Class 4) seem to have visited less often in 2019.
- The largest differences seem to be in the smallest and longest ship classes, with Class 1 doubling, and Class 6 quadrupling frequencies of visits in 2019 in reference to 2015.

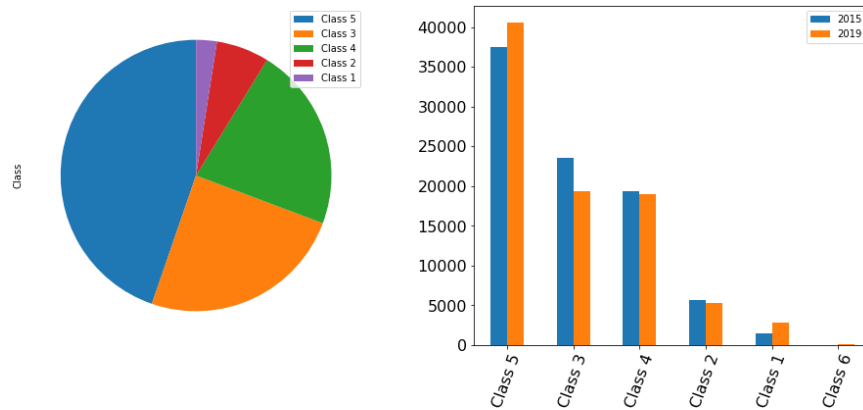


Fig 2.6 Frequency of visits per ship class

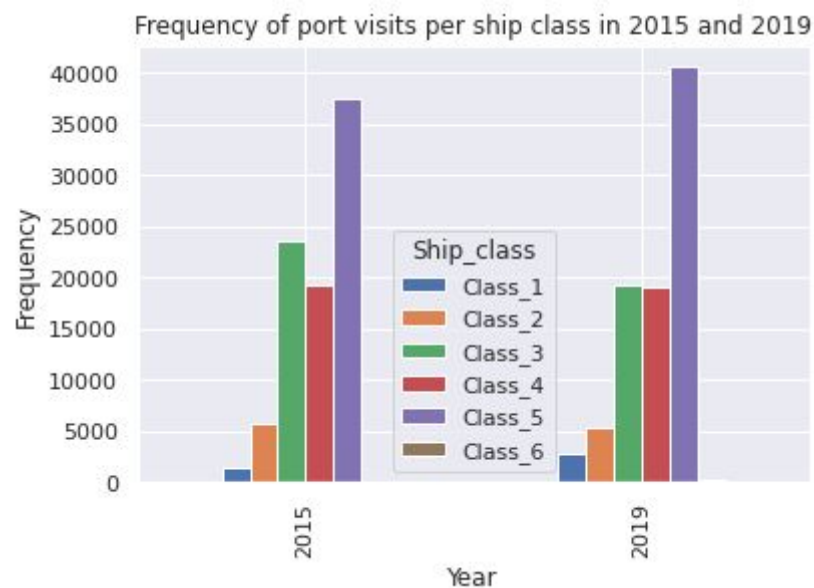


Fig 2.7 Frequency of visits per ship class

According to *Table 2.5*, the class 6 experimented in 2019 an increase on its visits by a factor of 4x respect to 2015. However, in *Figures 2.6* and *2.7* this growth cannot be observed because the Class 6 ships still make up for less than 1% of the total ships.

By looking more closely at the class of ship, *Table 2.8*, whether they depart, stop or are just in transit, It was observed that the majority of ships on through-passage are Class 3. Whereas, the majority of ships unloading/loading in the port are Class 5.

Route_category	Ship class Year	Class_1	Class_2	Class_3	Class_4	Class_5	Class_6
Arrival	2015	713	1602	7906	7133	15539	30
	2019	1430	1948	6093	7589	16717	95
Depart	2015	644	1750	8700	8388	17740	9
	2019	1390	1538	7480	8083	19599	74
Through- passage	2015	79	2338	6969	3797	4223	2
	2019	72	1859	5791	3326	4323	8

Table 2.8 Frequency of port visitors in 2015 and 2019 by ship class and route category:

Estimated daily ship visits

An estimation of the daily ship visits was made on the assumption that daily ship traffic is evenly distributed over the 365 days in a year. On average there are approximately 90 arrivals and 100 departures per day.

Route category	Arrival	Depart	Through -passage
2015	90.2	102.0	47.6
2019	92.8	104.5	42.1

Table 2.9 Estimated daily arrival, departs and through per day.

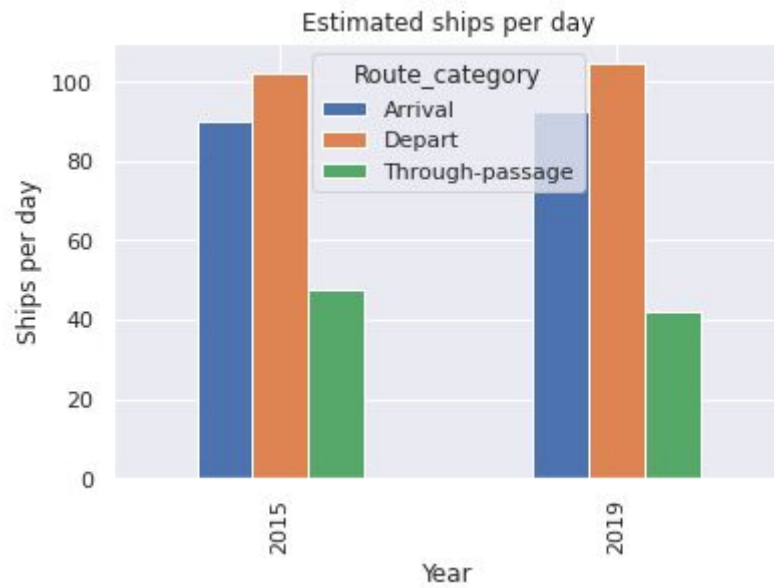


Fig 2.10 Estimated frequency of visits per day

Distribution by type

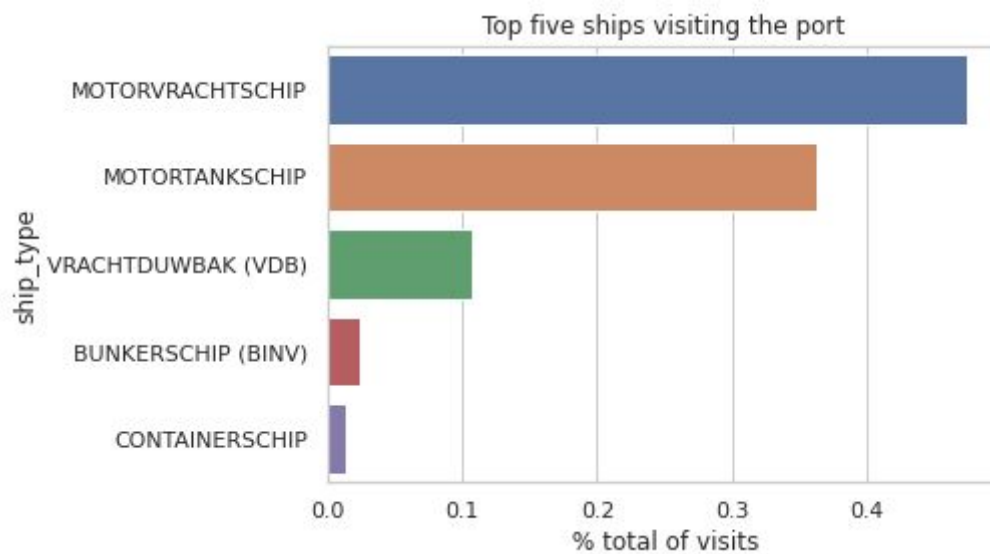


Figure 2.11 Top five ship types registered in the port historically

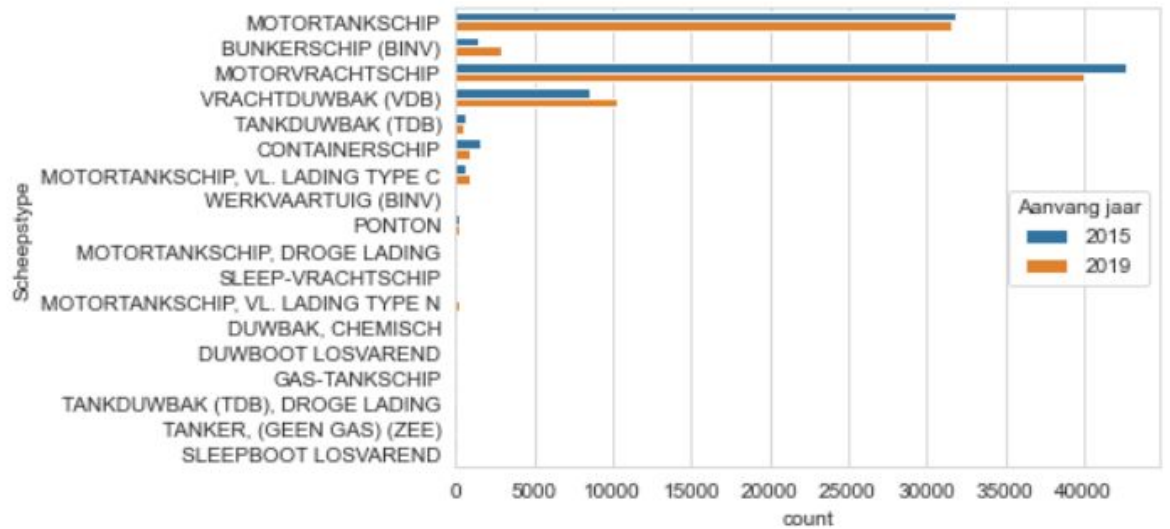


Figure 2.12 Ship type visiting the port by year

By comparing values for 2015 and 2019, Motortankship & Motorvrachtship count register a decrease but Bunkership & Vrachtduwbak are increased.

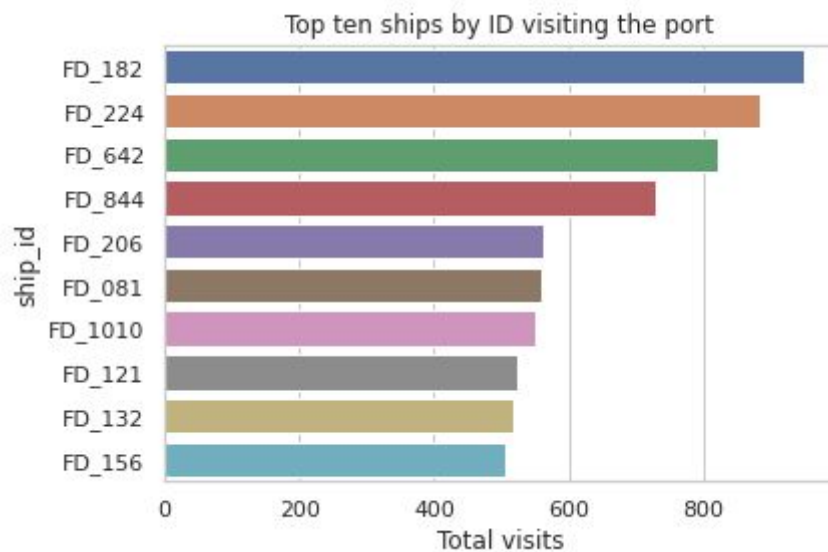


Figure 2.13 Total of visits per boat ID (Top ten)

	ship_type	quantity
0	MOTORVRACHTSCHIP	0.473176
1	MOTORTANKSCHIP	0.362253
2	VRACHTDUWBAK (VDB)	0.106734
3	BUNKERSCHIP (BINV)	0.024420
4	CONTAINERSCHIP	0.014185
5	MOTORTANKSCHIP, VL. LADING TYPE C	0.008030
6	TANKDUWBAK (TDB)	0.006081
7	PONTON	0.001972
8	MOTORTANKSCHIP, VL. LADING TYPE N	0.001434
9	GAS-TANKSCHIP	0.000531
10	WERKVAARTUIG (BINV)	0.000343
11	SLEEP-VRACHTSCHIP	0.000331
12	MOTORTANKSCHIP, DROGE LADING	0.000263
13	TANKER, (GEEN GAS) (ZEE)	0.000166
14	DUWBAK, CHEMISCH	0.000051
15	DUWBOOT LOSVAREND	0.000017
16	SLEEPBOOT LOSVAREND	0.000006
17	TANKDUWBAK (TDB), DROGE LADING	0.000006

Table 2.14 Procentual distributions of ships visiting the port

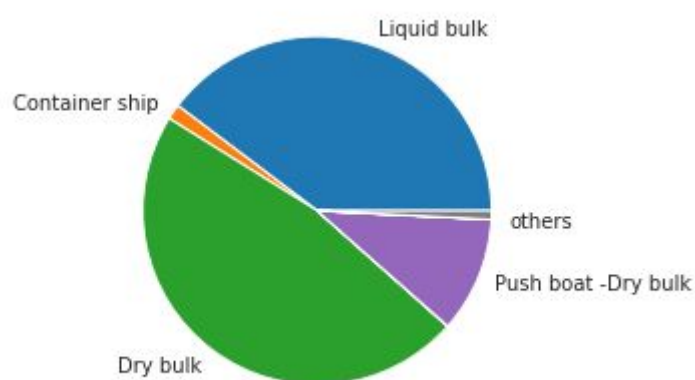


Figure 2.15 Dry bulk, Liquid bulk, Push boat -Dry bulk and Container ship are representing 99.15% of total ships type.

Distribution by DWT and length

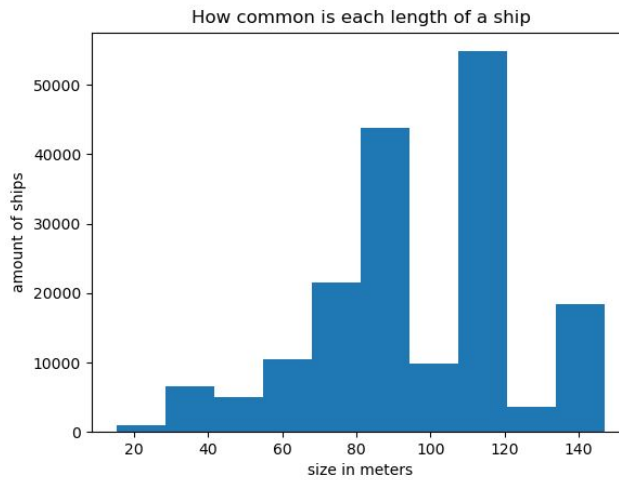


Figure 2.16 Distribution of ships by size in meters

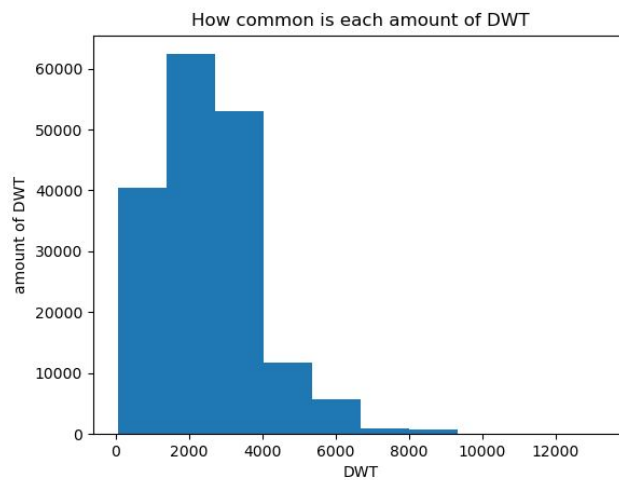


Figure 2.17 Distribution of ships by DWT

The above graphs represent the distributions of boats by DWT and length, with the intention of exploring the presence of 'outliers' in the data.

- There are no outliers in the max length of a ship.
- The distribution of DWT is skewed to the left. There are 31 outliers with a DWT higher than 12000.

Route profile and relevant destinations

Total destinations in 2015: <i>Doorvaart excluded</i>	815
Total destinations in 2019: <i>Doorvaart excluded</i>	765

Table 2.18 Total of destinations registered by year.

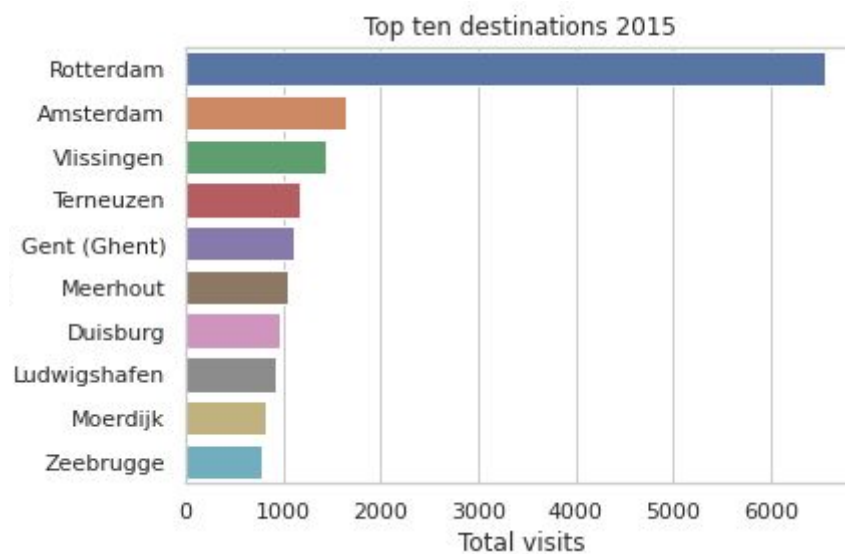


Fig. 2.19 Most frequent destinations in 2015.

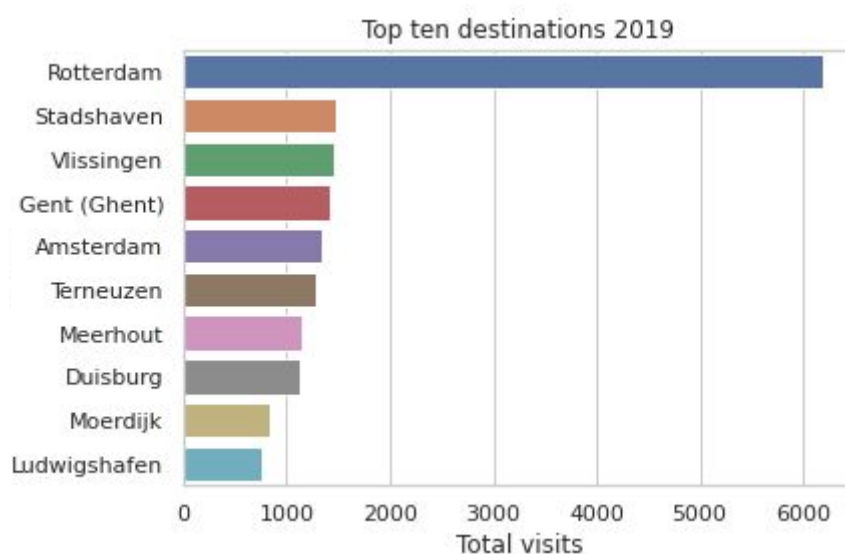


Fig 2.20 Most frequent destinations in 2019.

Rotterdam remains as the most significant destination over time.

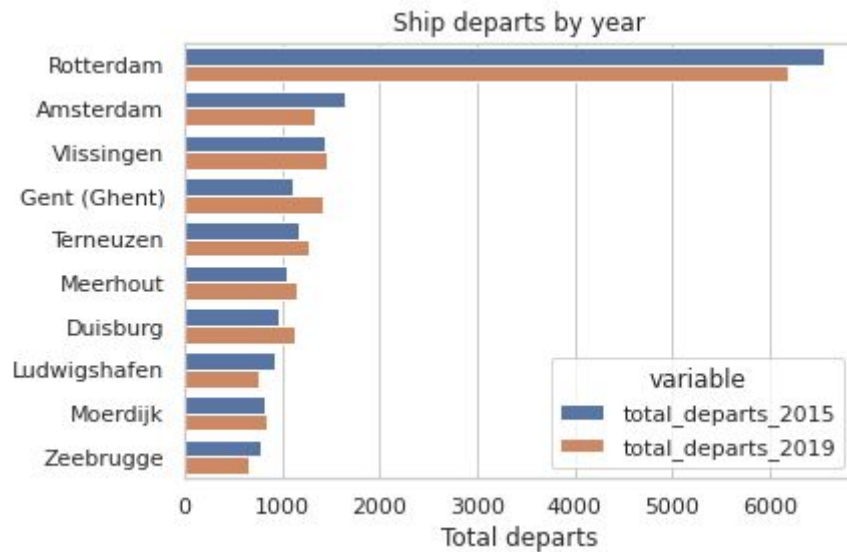


Fig 2.21 Most frequent destinations in 2015 vs 2019.

As shown in the below figure, transit trips are decreasing, whereas the origin and total destinations have increased when comparing 2015 to 2019 data.

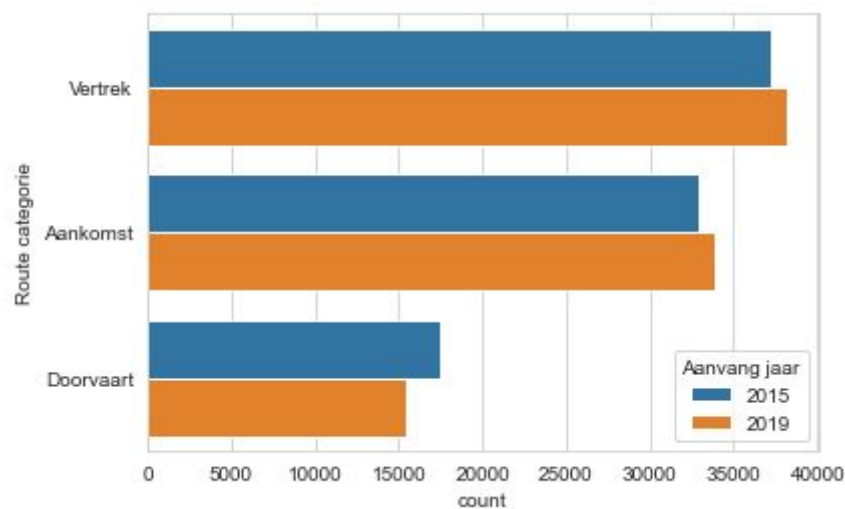


Fig 2.22 Route categories by year.

Arrival and departure information reported by each ship was studied to discover which were the most common route profiles. Below, *Table 2.23* shows the top 25 routes based on 2019 data. It indicates many ships make round-trips, rather than one-way or multi-port trips. Further data including the ship registry ID and date/time of each visit could be helpful to predict which ship class and ship type will visit the Port each day based on previous years' data.

Similarly, a study of ships' destinations according to their port of origin was done and it's illustrated in the *Figure 2.24* The graph shows the prevalence of round trips as well as the relevance of Rotterdam as a destination but also insights regarding how the goods are transported among the cities.

Departures	Arrivals	Visits
Meerhout	Meerhout	4082
Stadshaven	Stadshaven	2769
Vlissingen	Vlissingen	2585
Terneuzen	Terneuzen	1980
Willebroek	Willebroek	1815
Deurne	Deurne	1687
Amsterdam	Amsterdam	1679
Genk	Genk	1670
Zeebrugge	Zeebrugge	1589
Rotterdam	Vlissingen	1455
Vlissingen	Rotterdam	1274
Rotterdam	Amsterdam	1176
Gent (Ghent)	Gent (Ghent)	1071
Alphen aan den Rijn	Alphen aan den Rijn	977
Duisburg	Rotterdam	784
Duisburg	Duisburg	761
Ludwigshafen	Ludwigshafen	748
Amsterdam	Rotterdam	714
Liege	Liege	710
Straatsburgdok	Straatsburgdok	705
Rotterdam	Gent (Ghent)	638
Grobbendonk	Grobbendonk	632
Brussel (Bruxelles)	Brussel (Bruxelles)	628
Bergen op Zoom	Bergen op Zoom	604

Table 2.23 The most common routes by departure and arrival (2019)

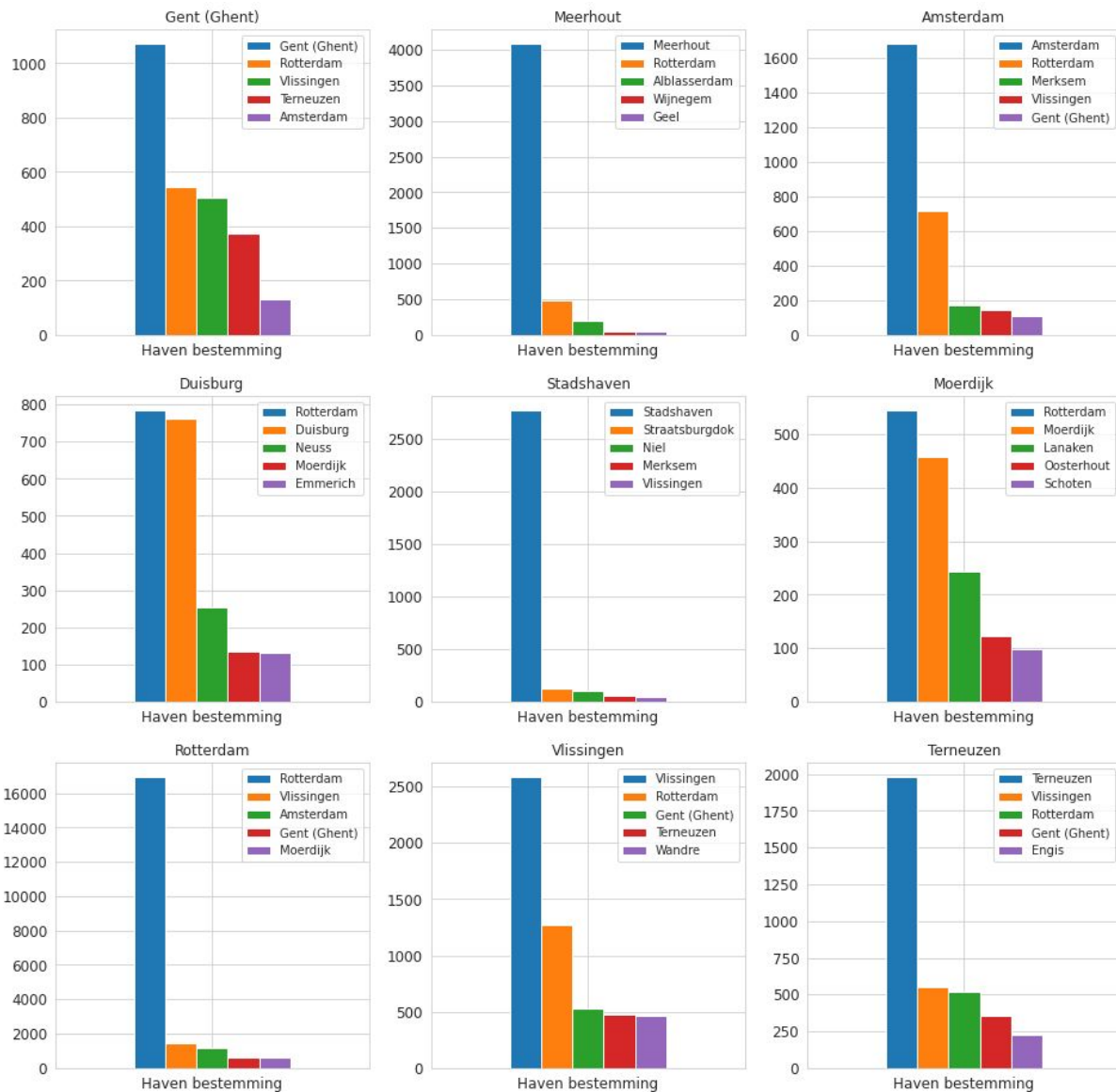


Fig 2.24 Route destinations according with port of origin

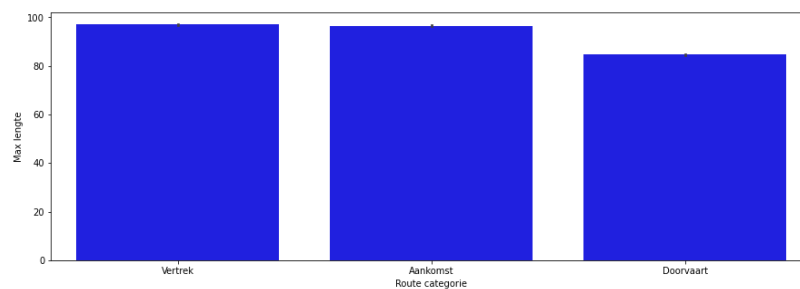


Figure 2.25 Distribution of ships by DWT

By comparing the relationship Max-length and route category, it's possible to identify that for boats at the top 10% of *Max lengte* ships don't have *Doorvaart* as their route category.

Ship type and their route represented geographically

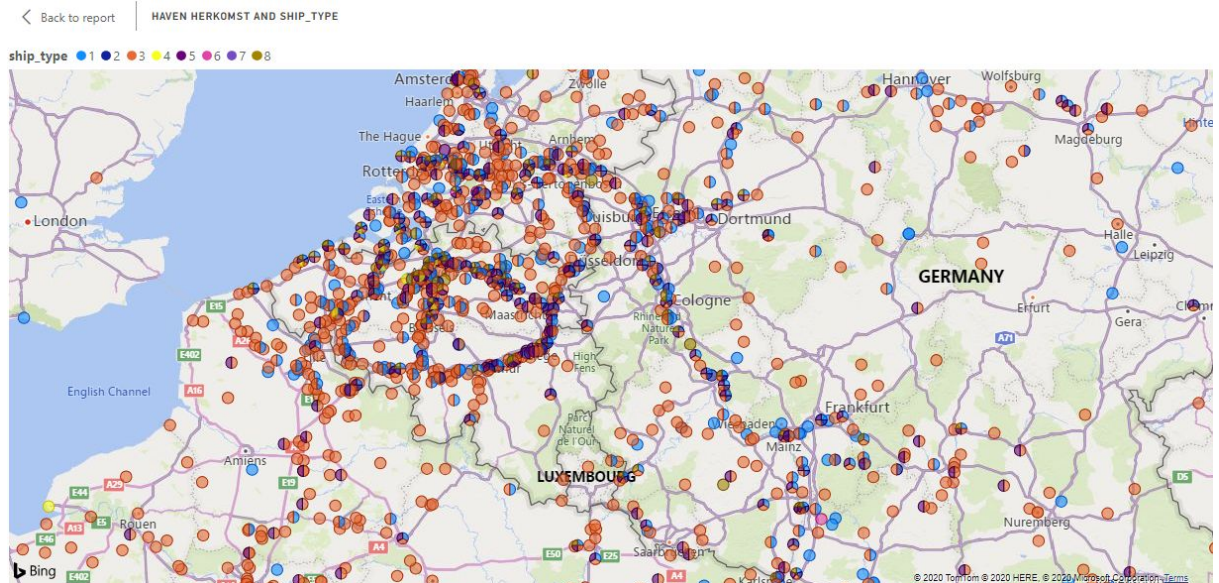


Fig 2.26 Trip category vs. Route (colored by ship type)

In the below figure Fig 2.27, the size of the dots indicates the relevance of the port according to the quantity of visits. Rotterdam can be seen as the busiest port.



Fig 2.27 Trip category vs. Route (Bin size by product code count)

Exploring possible correlations

The below *figure 2.28* is the result of exploring the dataset to find correlations along its values. For example, the combination of “**DWT**” at the Y-axis and “**Max Lengte**” at the X-axis, indicate a linear correlation between the values; “As the ship’s capacity in tons increases then so will the boat’s maximum length increase”.

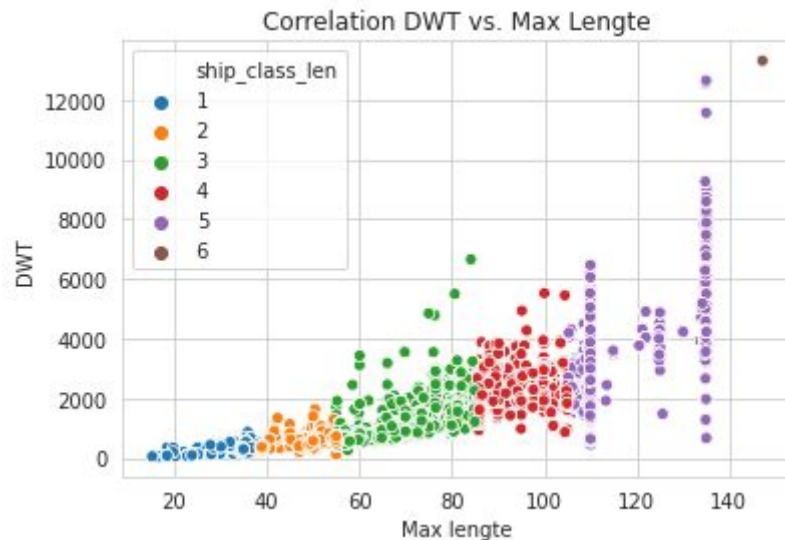


Fig 2.28 Exploratory graphs to find data correlations

Clear correlations among the other values in the dataset were not found, however the analysis is not conclusive since the study comes just from exploring raw data without cross-referencing with other sources of information.

About the report

This study and its associated report was elaborated as an academic exercise by students of the training program in Artificial Intelligence led by Bencode on its Antwerp campus. The insights found are the result of a general exploration to the dataset provided; more detailed information might be found by deeper analysis, definition of the main fields of interest and exploration of additional data.

Contributors:

Albermar Dominguez

Chirag Sharma

Chrysanthi Katrini. **Mentor**

Hedia Bougi

Kristof Bosmans

Maarten Fish

Robin McGregor

Saba Hasan

Santosh Ahirrao

Seppe Seghers

Steven Verkest

Tomas Verrelst

Wim Christiaansen

Contact information:

Chrysanthi Katrini.

chrysanthi.katrini@bencode.org