BT 5110 Database and Warehouse

Project Maritime Q2 Report

Ke Ma, A0212524U

Mingzhe Xu, A0232022A

Nanhai Zhong, A0231953E

Xiao Liang, A0232007X

Introduction

In Question 2, we extend the reference website by creating a star schema and an interactive interface to show different dimension tables and visualize in order to do exploratory analysis online.

The data used is consolidated and publicly available by the European Maritime Safety Agency (https://www.emsa.europa.eu). The dataset used contains ship details, technical efficiency data, verifier data and CO_2 emission records for different ships across year 2018-2020.

The data is available from https://mrv.emsa.europa.eu/#public/emission-report. This report gives a briefing to the star schema we created and deployed in database in the framework of Django and how to conduct interactively exploratory analysis on the website interface.

Data Warehouse

1. Star Schema

In this project, we design this database as a star schema with fact table and four dimension tables. The fact table records the measures of the design efficiency of vessels, the total fuel consumption, the total CO2 emission and so on, which are all additive facts. The four dimension tables record more detailed information about each dimension to support analysis, shown as below *Figure1*. Each dimension table has a surrogate key to represent each entry, which is used by the fact table to link to different dimension tables.

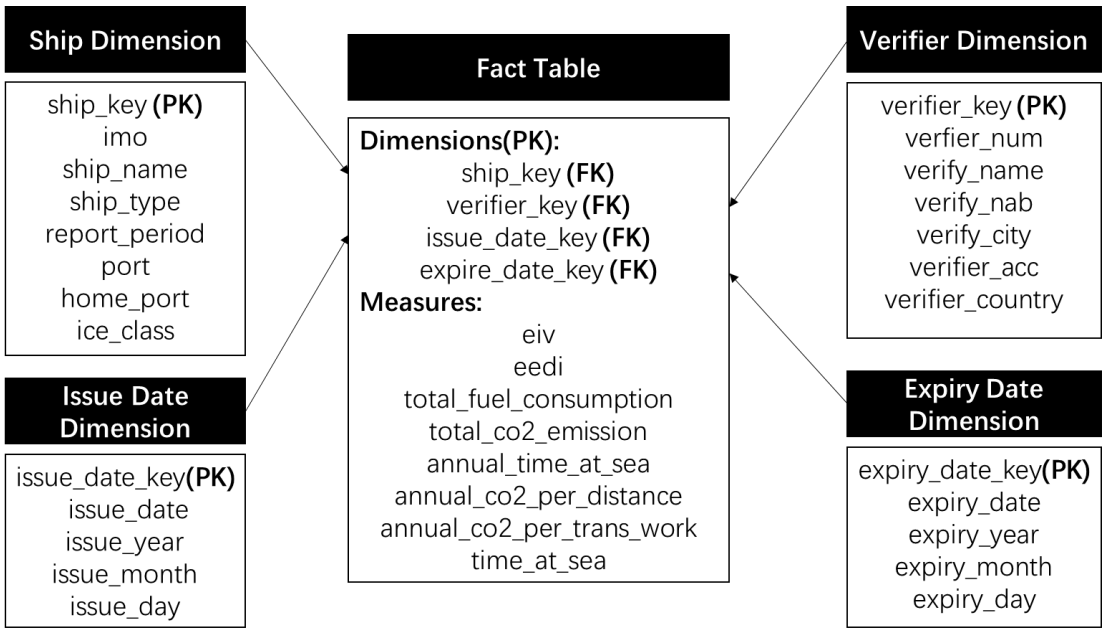


***Figure 1.* Star Schema Sketch**

2. Data Structure

For this question 2, we structure 5 tables which are the *fact table, explore*, *explore_verifier, expire_date, issue_date*. Here are one fact table and 4 dimension tables to structure our star schema. Moreover, we create 4 surrogate keys which are *ship_key, verifier_key, issue_date_key, expire_date_key* and assign each to dimension tables. For fact table, we linked with other dimension tables by surrogate keys, so there are 4 primary keys in the fact table. Also, in fact table, all the features' property are numeric.

The ERM of data structure is shown below.

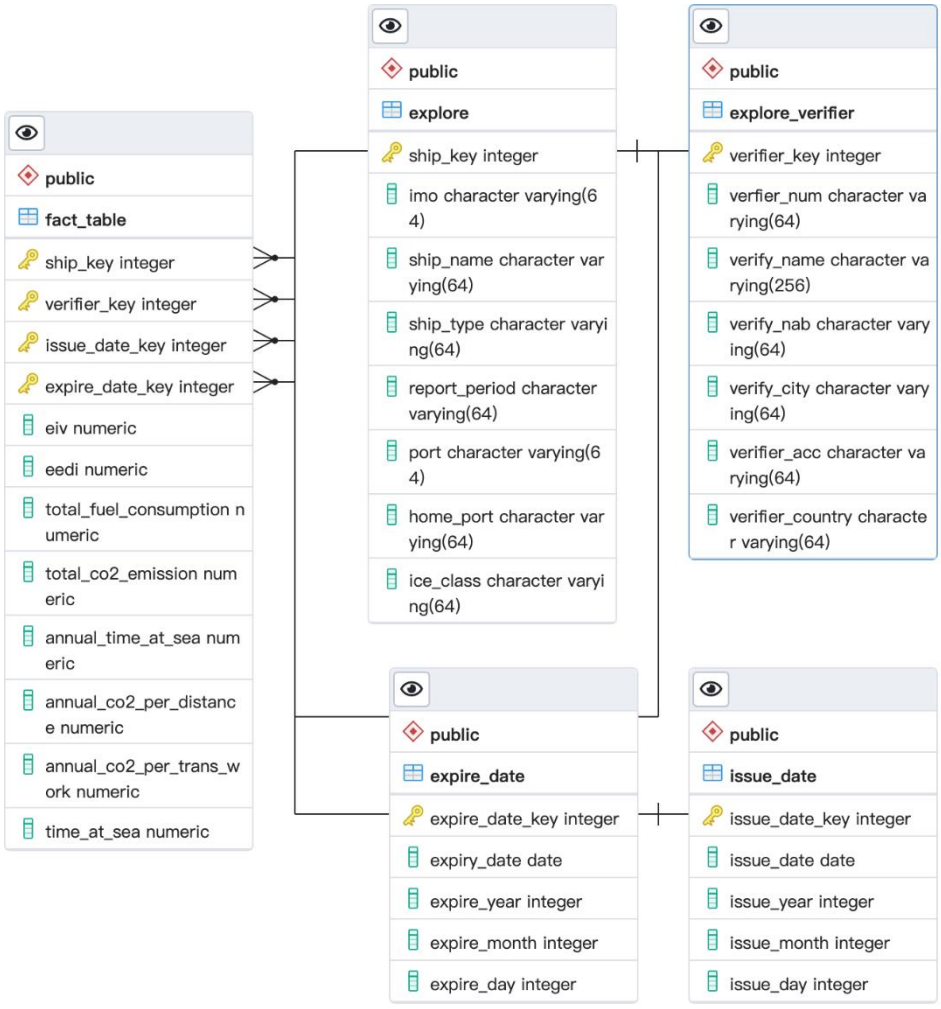
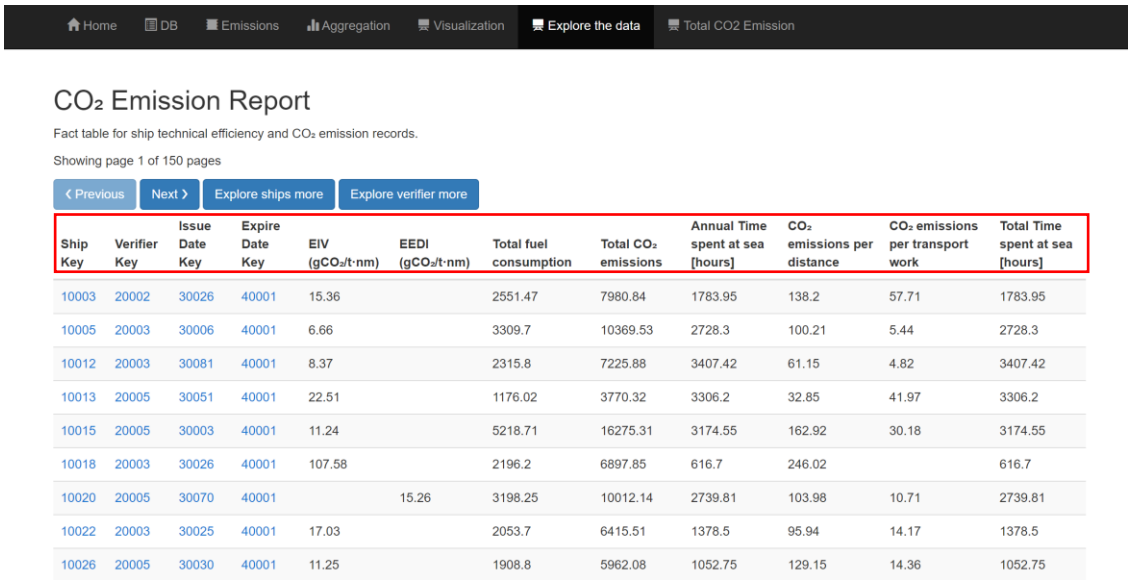


Figure 2. ***Data Structure ERM***

Functions Achieved

1. Explore the Data

a) Main Interface (Fact table)



Ship Key	Verifier Key	Issue Date Key	Expire Date Key	EIV ($gCO_2/t{\cdot}nm$)	EEDI ($gCO_2/t{\cdot}nm$)	Total fuel consumption	Total CO_2 emissions	Annual Time spent at sea [hours]	CO_2 emissions per distance	CO_2 emissions per transport work	Total Time spent at sea [hours]
10003	20002	30026	40001	15.36		2551.47	7980.84	1783.95	138.2	57.71	1783.95
10005	20003	30006	40001	6.66		3309.7	10369.53	2728.3	100.21	5.44	2728.3
10012	20003	30081	40001	8.37		2315.8	7225.88	3407.42	61.15	4.82	3407.42
10013	20005	30051	40001	22.51		1176.02	3770.32	3306.2	32.85	41.97	3306.2
10015	20005	30003	40001	11.24		5218.71	16275.31	3174.55	162.92	30.18	3174.55
10018	20003	30026	40001	107.58		2196.2	6897.85	616.7	246.02		616.7
10020	20005	30070	40001		15.26	3198.25	10012.14	2739.81	103.98	10.71	2739.81
10022	20003	30025	40001	17.03		2053.7	6415.51	1378.5	95.94	14.17	1378.5
10026	20005	30030	40001	11.25		1908.8	5962.08	1052.75	129.15	14.36	1052.75

***Figure 3* Main Interface**

Main interface we created is shown above in the block *Explore the data*. This interface contains two page buttons, two explore more buttons which are introduced later, and a fact table contains basic information of ships emission. Four surrogate keys represent records in dimension table about ship, verifier, issue date and expire date respectively.

Click on table headers in red box, can order the table by each column.

b) Dimension tables for four surrogate keys

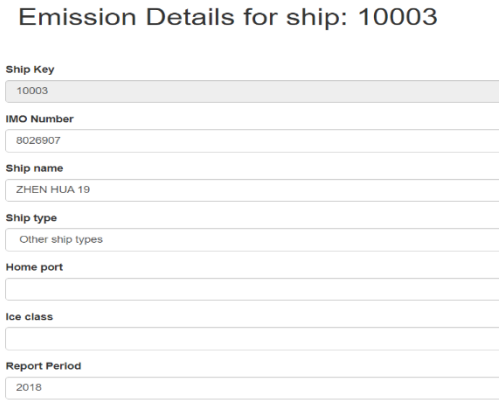


***Figure 4.* ship record**

***Figure 4.* Verifier record**

All the four surrogate keys (Blue font color as shown below) in the table can be clicked. And it will respectively return different record accordingly from dimension tables.

Dimension record displayed after clicking *Ship_Key, Verifier_Key*.
There are also two buttons on main page: *Explore ships more*, *Explore verifier more*.

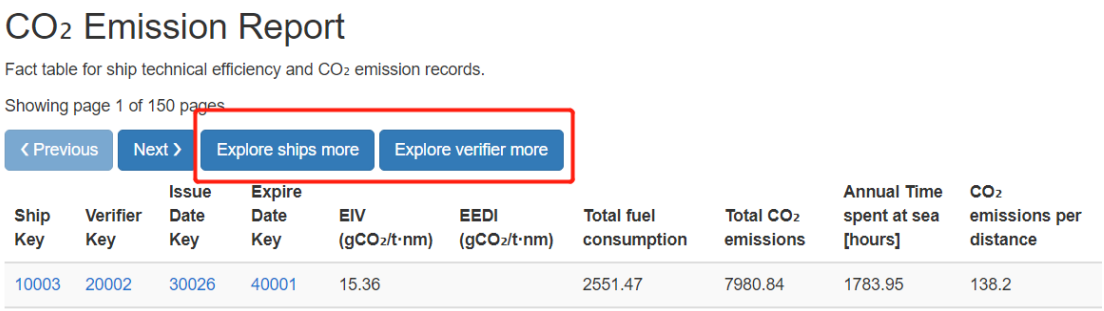


***Figure 5.* Explore more function**

After clicking these two buttons, statistical graphs and tables about ships and verifiers will be displayed respectively.

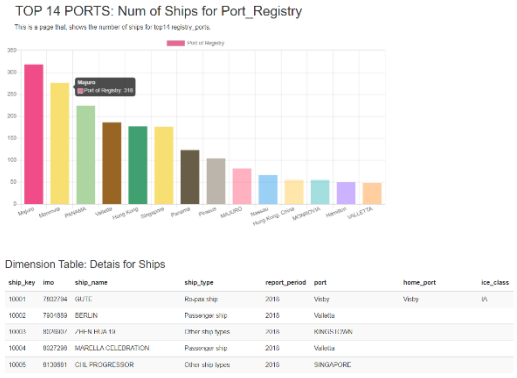


***Figure 6.* Explore the ship dimension**

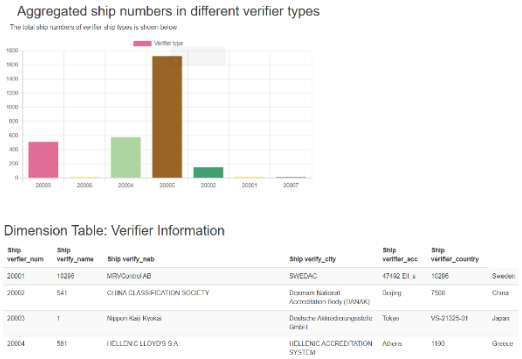


***Figure 7.* Explore the verifier dimension**

The statistical graph ***Figure 8*** and dimension table displayed after clicking Explore ships more are as follows. The graph shows the number of ships of top 14 registry ports, and the table contains more details for each ship.

The statistical graph ***Figure 9*** and dimension table displayed after clicking Explore verifiers more are as follows. The graph shows the number of ships of different verifier types, and the table contains more details for verifier information.

2. Total CO_2 Emission

In this section, we try to calculate and visualize the sum of total CO_2 emission in total and in different groups, including different ship types, different verifier countries and different ship types in each verifier country.

a) Visualization

There are three bar charts showing the sum of total CO_2 emission in different groups.

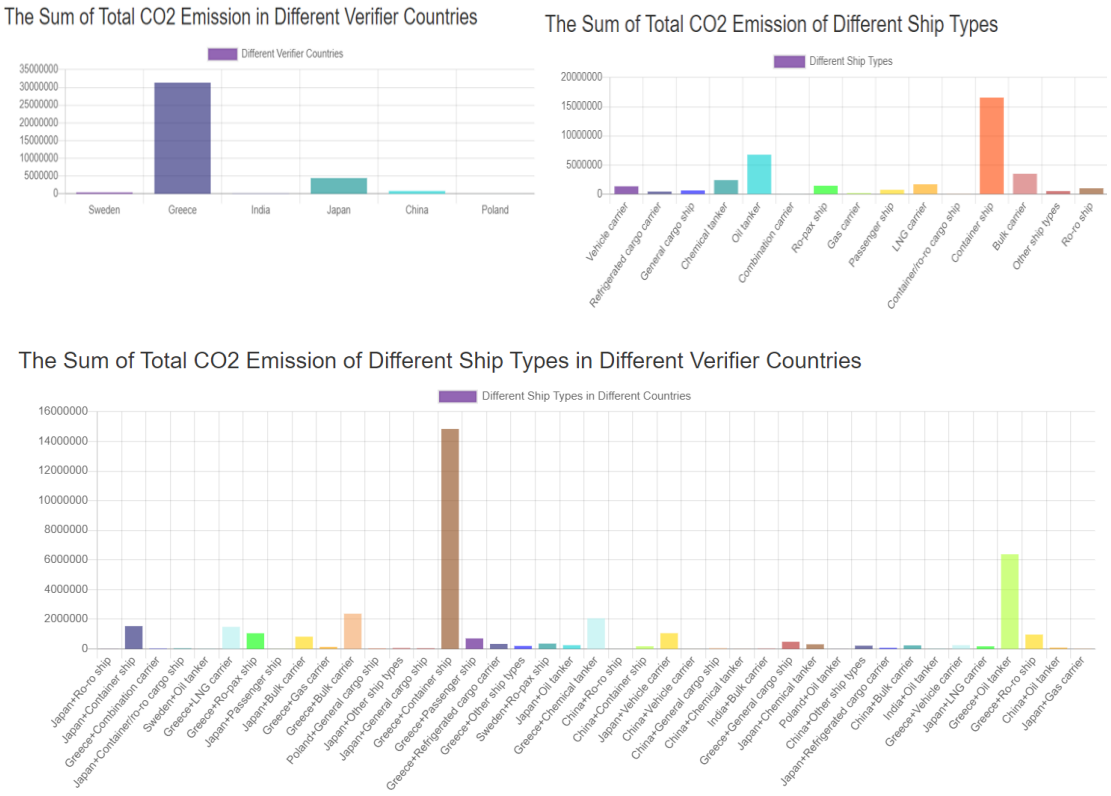


Figure 9. ***Roll up and cube queries for CO_2 Emission Statistics***

b) Calculation Tables

There are two tables showing the calculation results with different calculation methods. The first one uses the CUBE query, which includes all possible grouping sets, while the second one only generates all grouping sets in the order of the input columns.

Total CO2 Emission (with cube function)

Verifier Country	Ship Type	Sum of CO2 Emission
		36980743.85
Greece		31325401.59
	Container ship	16551728.19
Greece	Container ship	14835661.95

Total CO2 Emission (with rollup function)

Verifier Country	Ship Type	Sum of CO2 Emission
		36980743.85
Greece		31325401.59
Greece	Container ship	14835661.95
Greece	Oil tanker	6383243.28