Data Model Design: Initial Notes

Right now, it appears that no port in the United States is even close to being as tech savvy and optimized as the Port of Rotterdam. Our case can be getting a US port, such as the Port of Baltimore, optimized through a data warehouse which tracks time, location, date, etc; To get us closer to this idea of a ‘Smart’ Port.

https://www.ibm.com/blogs/industries/port-of-rotterdam-digitization-energy-transition/

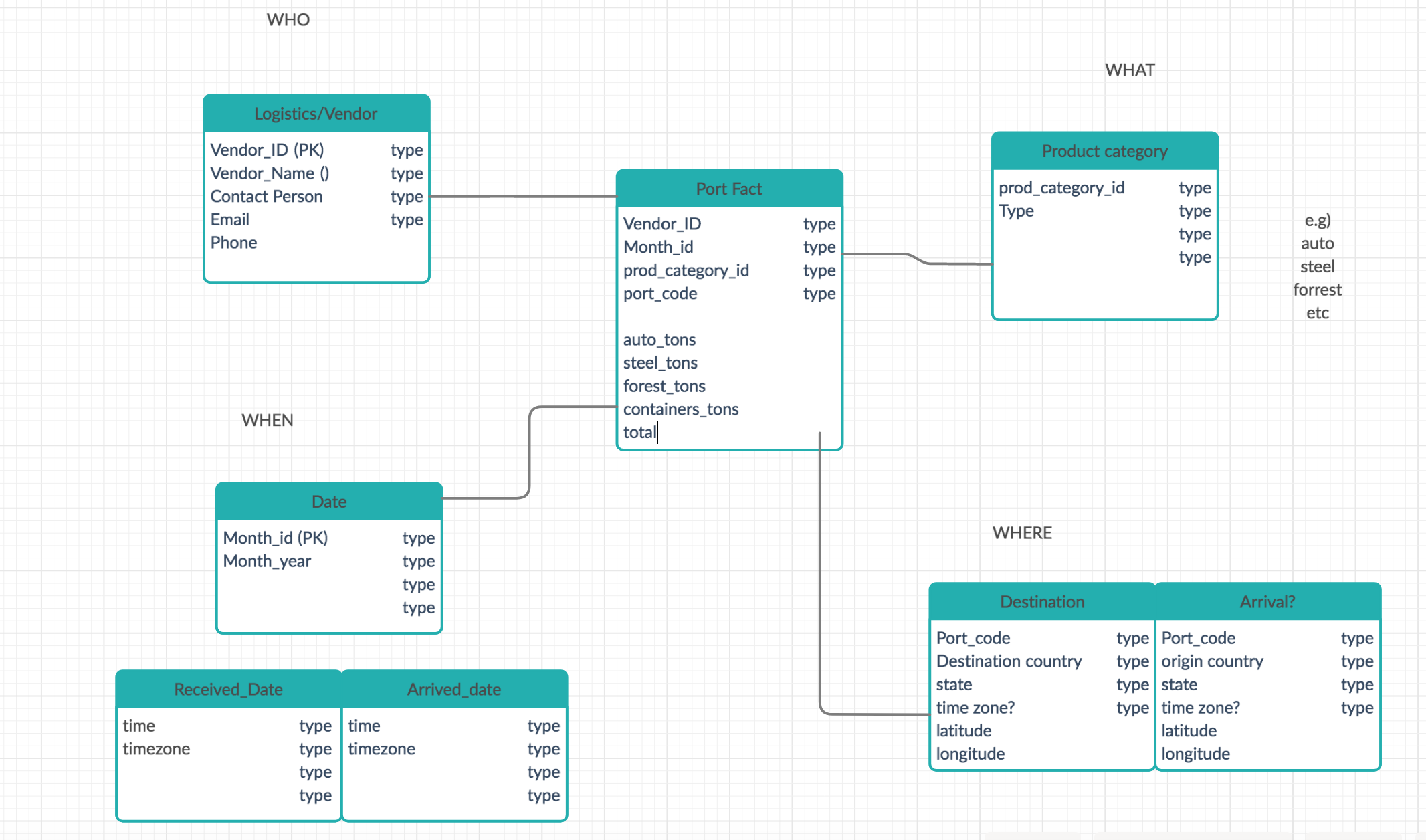
**Potential pain points for us to address:**

-Having a consolidated single view

-Cross functionality outlined between government port authority, logistics companies, etc

-Addressing the issue of congestion, and optimizing workflow by tracking time, date, revenue

**Initial design of our dimensional model based on research of attributes relevant to shipping:**



**Research sources and how they relate to our potential dimensional schema:**

Sources which serve as the basis for our schema design:

<https://opendata.maryland.gov/Transportation/Maryland-Port-Administration-General-Cargo/2ir4-626w>

Maryland ports **dataset**, mostly just aggregate sums, grouped by different material types for each month of the year. This serves as the initial inspiration for our fact table. It’s mostly numeric aggregate data, involving cargo types and cargo weights. So we will make a fact table which echoes a similar design. Except in our case, we may also want to outline attributes such as cost, times, etc. (Fact Table, as well as our WHAT Dimension and WHEN dimension)

<https://data.imap.maryland.gov/datasets/c7e2b2c07d9043f9b4c34aa1b9815224_20/data?geometry=-76.797%2C39.215%2C-76.302%2C39.308&orderBy=FACILITY_T&page=2>

**Dataset** for location data for Maryland ports**.** Contains numerous attributes which relate to exact latitudinal and longitudinal coordinates for ports of arrival. Again, it is a little bit different from our use case, but it can serve as a template we can build upon. (WHERE Dimension)

<https://www.freightos.com/freight-resources/seaport-code-name-finder/>

Contains unique Port ID, latitude, and longitude of global ports. Useful for us tracking the location of arrival and destination ports. (WHERE Dimension)

<https://pobdirectory.com/directory/?category=Steamship%20and%20Barge%20Lines%20and%20Agents>

Provides a comprehensive list of logistics companies that provide container and vessel services to the Port of Baltimore. (WHO Dimension)

Sources which serve as the basis for our business case:

<http://rvlasveld.github.io/blog/2013/09/11/maelstrom-using-port-of-rotterdam-open-data/>

Experiment from a few years ago, about traffic into the port of Rotterdam.

<https://www.portofrotterdam.com/en/news-and-press-releases/self-learning-computer-predicts-vessel-arrival>

Port of Rotterdam site, talks about tackling the problem of congestion by recording arrival times, etc and predicting them.

<https://www.mckinsey.com/industries/travel-logistics-and-transport-infrastructure/our-insights/ports-and-shipping-the-need-for-solutions-that-cross-lines>

McKinsey writeup about the shipping industry.

“In the past, predicting the arrival of vessels in port was difficult… But new tools combine publicly available GPS data on the current position of vessels with the latest machine-learning techniques. Such tools can estimate the deviation between the current and scheduled position of a ship and therefore provide a better estimated time of arrival for any port on its route. ”

<https://www2.deloitte.com/content/dam/insights/us/articles/4954_FoM-and-movement-of-goods/4954_FoM-and-movement-of-goods.pdf>

Writeup by Deloitte also brings up the topic of congestion.

“Apparel retailer Lululemon estimated a US$10 million loss in holiday sales due to port congestion in 2015 ”

<https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/consumer-business/deloitte-nl-cb-global-port-trends-2030.pdf>

Another good Deloitte writeup pretty much tackles all optimization aspects.

<https://www.maritimedatacompany.com/product/ais-dataset-port-of-rotterdam/>

Port of Rotterdam dataset. It costs €300 a month! But it outlines some good business Q’s on the website, one of which brings up congestion.

<https://www.census.gov/foreign-trade/reference/products/catalog/port.html>

US Census site, brings up some interesting potential attributes for ur model.

<https://blog.shipuwl.com/east-coast-port-congestion-updates-from-baltimore-chicago-and-newark>

Another writeup on congestion, mentions Port of Baltimore.

<https://www.techrepublic.com/article/ai-on-the-high-seas-digital-transformation-is-revolutionizing-global-shipping/>

Prof’s article, on how Digital Transformation can disrupt the shipping industry.

**Our data sources:**

ports-metadata: Potential attributes we can generate

Maryland\_Port\_Administration\_General\_Cargo: Basis of our fact table. Aggregate weight totals for Maryland Port Authority.

Maryland\_Transit\_-\_Sea\_Ports: Location data for Maryland ports.

Additionally, we may simply generate data, and use these as a basis.