Reinventing Freight Logistics with Data Science
1300+ locations, 100+ countries, 70k+ employees, $20bn revenue

#1
#2
#3
10m sqm
4.1m/y
“predicting and optimizing logistics and supply chains through data analysis”
Big Data
Internet of Things
Predictive Analytics
Data Science
Artificial Intelligence
Machine Learning
China exports jump 44.5% in February
Forecasting port activity

Port of Los Angeles Imports Surged in February

Container volume jumped 30% from February 2014 as trucking companies also see reported accelerating shipping demand
Forecasting port activity (cont.)

1. Measure draught
2. Think
3. Forecast
Practical case: sea-freight tracking and benchmarking
Questions relevant to our business:

- When does Emma Maersk’s voyage 39W arrive in Hamburg? (“No! I mean really...”)
- Which carrier is more accurate on their ETAs?
- Which ports show larger delays? Generically or individual terminal handlers?
- Whose fleet has the least carbon footprint in S. Asia - N. Europe lane in the last 6m?
- Which carriers have the most stable schedules?
- What are the load factors on individual lanes?
- Which connections are more sensitive to delays?
- How can cargo insurance costs be optimized?
- How can trade financing costs be optimized?
- ....
AIS position messages
3 billion messages received in 2017 for the about 4000 container ships
Use of geofencing
From positions to routes
The “Highways of the Sea”

Each individual voyage between 2 ports was monitored, creating a universe of about 16,000 distinct routes.
Vessel Tracker: Forecast

- **Emma Maersk**
  - Heading towards: Salalah
  - ETA (UTC+4): 2018-09-02 22:30

- **Bremerhaven**
  - Arrival: 827W (UTC+1)
  - Departure: 8331 (UTC+1)

- **Rotterdam**
  - Arrival: 827W (UTC+2)
  - Departure: 8331 (UTC+2)

- **Tanger Med**
  - Arrival: 8331 (UTC+1)
  - Departure: 8331 (UTC+1)

- **Salalah**
  - Arrival: 8331 (UTC+4)
  - Departure: 8331 (UTC+4)

- **Colombo**
  - Arrival: 8331 (UTC+5.5)
  - Departure: 8331 (UTC+5.5)

- **Ningbo**
  - Arrival: 8331 (UTC+8)
  - Departure: 8331 (UTC+8)

- **Xiamen**
  - Arrival: 8331 (UTC+8)
  - Departure: 8331 (UTC+8)

Schedule retrieved: 2018-07-10 16:17
Questions relevant to our business (again):

- When does Emma Maersk’s voyage 39W arrive in Hamburg? (“No! I mean really...”)
- Which carrier is more accurate on their ETAs?
- Which ports show larger delays? Generically or individual terminal handlers?
- Whose fleet has the least carbon footprint in S. Asia - N. Europe lane in the last 6m?
- Which carriers have the most stable schedules?
- What are the load factors on individual lanes?
- Which connections are more sensitive to delays?
- How can cargo insurance costs be optimized?
- How can trade financing costs be optimized?
- ....
Backtesting

Start by reducing population to about 60 million points within 5 months.
Typical snowflake model
Measures versus Dimensions

- Schedule error
- Estimate error

Aggregation Function: Quantile
How many times we have a different ETA from carrier?

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>All</td>
</tr>
<tr>
<td>Port</td>
<td>All</td>
</tr>
<tr>
<td>Carrier Serv</td>
<td>All</td>
</tr>
<tr>
<td>KN Service</td>
<td>All</td>
</tr>
<tr>
<td>Lane</td>
<td>All</td>
</tr>
<tr>
<td>Voyage</td>
<td>Type voyage number</td>
</tr>
<tr>
<td>Vessel</td>
<td>Type IMO here</td>
</tr>
</tbody>
</table>

Total number of observations: 30078515

Adjusted: 49%
Original: 51%
When we adjust how successful are we?
What are the delays?
How does delay evolve with horizon?
### Schedule Error (90 Quantile) & Adjustment Improvement

- **Schedule Error (90 Quantile)**: 3.22 days
- **Estimate Error (90 Quantile)**: 1.93 days
- **Adjustment Improvement (90 Quantile)**: -40.07%

<table>
<thead>
<tr>
<th>Carrier</th>
<th>Estimate Error (days)</th>
<th>Schedule Error (days)</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allianca Navegacao e Logistica Ltda (BWR)</td>
<td>1.629259</td>
<td>3.120120</td>
<td>-47.8 %</td>
</tr>
<tr>
<td>CMA CGM (BWR)</td>
<td>1.840198</td>
<td>2.196921</td>
<td>-16.2 %</td>
</tr>
<tr>
<td>CNC Line-Cheng Lie Navigation Co. Ltd. (BWR)</td>
<td>1.953926</td>
<td>2.665264</td>
<td>-26.7 %</td>
</tr>
<tr>
<td>COSCO Shipping Lines Co. Ltd (BWR)</td>
<td>1.822327</td>
<td>3.605351</td>
<td>-49.5 %</td>
</tr>
<tr>
<td>Hamburg Sudamerikanische Dampfschiffahrts-Gesellschaft KG (BWR)</td>
<td>1.631890</td>
<td>3.121657</td>
<td>-47.7 %</td>
</tr>
<tr>
<td>Hyundai Merchant Marine Co Ltd (BWR)</td>
<td>1.655806</td>
<td>3.542143</td>
<td>-53.3 %</td>
</tr>
<tr>
<td>KMTC Line (BWR)</td>
<td>1.988658</td>
<td>4.729010</td>
<td>-58.0 %</td>
</tr>
</tbody>
</table>
Kühne+Nagel Sea-Explorer

Port Search

Origin(s):
- Shanghai (CNSHA)

Destination(s):
- Hamburg (DEHAM)

Search

9 options on 9 services match your requirement

9 options:
- Manta Ray 3
- Blue Whale 4
- Minke Whale 1
- Beluga 1
- Beluga 3
- Beluga 4
- Beluga 5
- Beluga 6

Details:
- **Manta Ray 3**
  - Transit Time: 28
  - Reliability Index: A
  - Emission Rating: A

- **Blue Whale 4**
  - Transit Time: 29
  - Reliability Index: B
  - Emission Rating: A

- **Minke Whale 1**
  - Transit Time: 30
  - Reliability Index: B
  - Emission Rating: D

- **Beluga 1**
  - Transit Time: 32
  - Reliability Index: B
  - Emission Rating: A

- **Beluga 3**
  - Transit Time: 32
  - Reliability Index: A
  - Emission Rating: A

- **Beluga 4**
  - Transit Time: 32
  - Reliability Index: A
  - Emission Rating: A

- **Beluga 5**
  - Transit Time: 33
  - Reliability Index: A
  - Emission Rating: A

- **Beluga 6**
  - Transit Time: 34
  - Reliability Index: B
  - Emission Rating: A
Lessons learned

- Keep data!
- Invest in real-time analytics: any backtesting comes for “free”
- “Know your science”
- Be nimble
- Understand the limitations of corporations
- Avoid “top-down” approach
- If you’re too old to study statistics, convince your kids to study statistics
- Hire well
"The forecasting accuracy of ML forecasting methods is lower than the worst of statistical ones while the accuracy of more than half the ML methods is lower than a random walk"
logindex

thanks!