



Tid, plats och statistik

Om sjötrafik och positionsdata som informationskälla

Marcus Justesen, geodataanalytiker SCB

GISS årsmöte 2017-02-16



facebook.com/statisticssweden



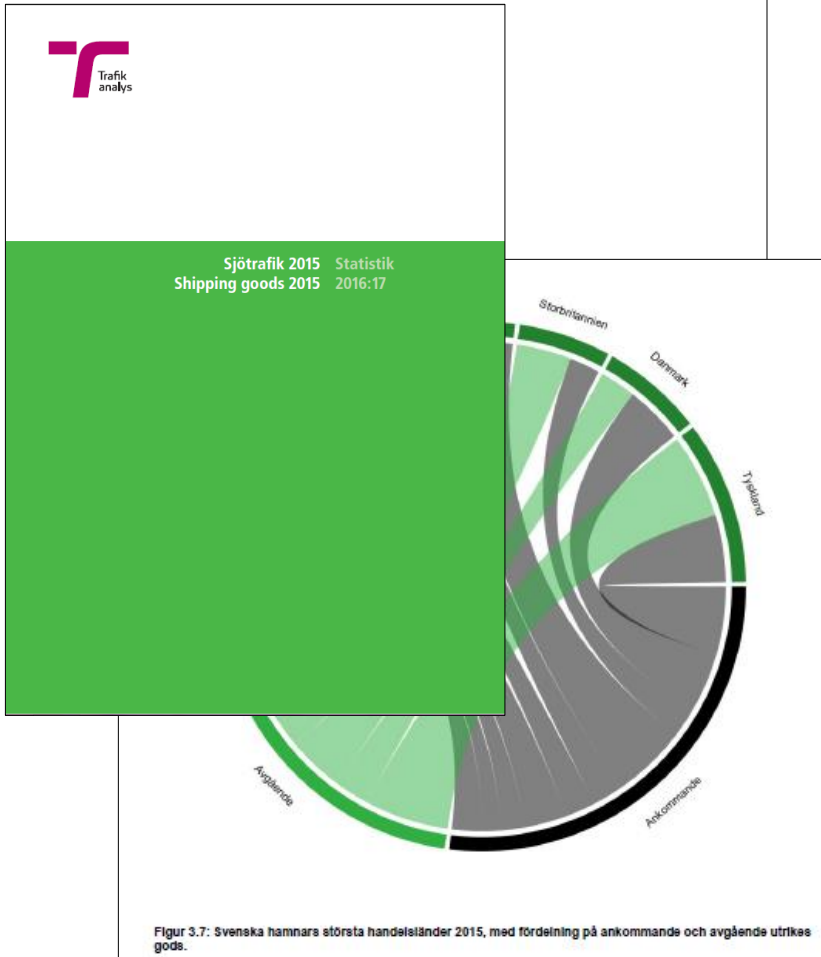
[@SCB_nyheter](https://twitter.com/SCB_nyheter)



[statistiska_centralbyran_scb](https://instagram.com/statistiska_centralbyran_scb)



Sjötrafiksstatiken och Trafikanalys



s). Översikt av sjöfarten år
II. Avgångna

Y F	Summa		Med last	
	Antal	Ton	Antal	Ton
3 098	71 216	632	70 367	
963	83 804	144	49 415	
122	11 569	384	107 027	
92	1 510	1	2 119	
5 139	317 034	2 936	862 142	
3 132	226 262	2 777	2 117 563	
373	35 329	1 005	572 859	
34	5 434	492	684 802	
8	829	206	196 292	
625	216 415	1 894	1 729 559	
90	18 901	456	378 977	
1	119	13	12 476	
13	3 992	16	38 118	
1	368	22	19 389	
1	135	11	9 545	
3	1 327	9	9 381	
—	—	3	8 719	
—	—	1	2 409	
—	—	25	61 293	
—	—	5	6 571	
—	—	14	13 361	
—	—	5	4 134	
10	2 206	—	—	
—	—	18	22 054	
3	725	2	1 821	
1	31 640	10	31 251	
4	4 186	19	66 168	
—	—	4	6 341	
—	—	9	7 573	
1	679	—	—	
—	—	17	59 253	
—	—	37	96 268	
16	28 067	—	—	
80	59 129	—	—	
9	27 386	—	—	
46	118 351	—	—	
—	—	—	—	
—	—	—	—	
3	3 123	—	—	
40	Summa år 1911	10 823	875 103	3 405
111 097	14 238	980 290	11 189	7 456 223
41	» 1910	10 254	840 448	3 614
133 286	13 848	978 774	10 411	6 724 260
42	» 1909	9 765	796 368	3 473
140 834	13 297	837 432	8 647	5 429 829
43	» 1908	10 381	854 454	4 199
139 832	14 580	994 296	9 452	5 533 062
44	» 1907	12 120	1 002 235	4 745
141 669	16 865	1 143 904	9 370	5 679 123



Tung insamling från hamnarna (TA:s hjältar! ♥)

Detaljerade uppgifter:

- Fartygsanlöp
- Fartyg
- Passagerare
- Fordon
- Gods



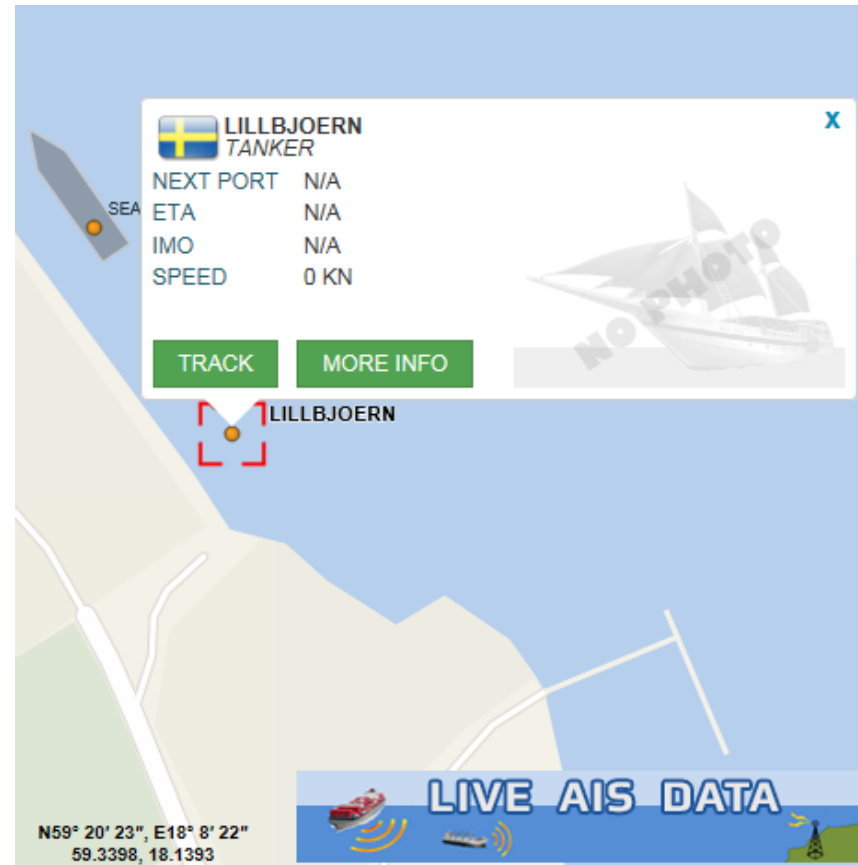
Statistikens utvecklingsbehov – och möjligheter

- endast angörande fartyg → alla fartyg i våra vatten
- saknar trafikarbete (fartygskilometer) → kan beräknas
- beräknar transportarbete mycket grovt → noggrannare
- stor börda → kan lindras något, med uppgifter om anlöp
- Samt mycket mer! (miljöbelastning mm)



Lösningen: AIS (Automatic Identification System)

- Ett obligatoriskt säkerhetssystem för handelssjöfarten
- Digital information delas via radio
- Sänder egen information till andra fartyg och landstationer, bl.a. position, hastighet, kurs
- Endast några sekunder mellan signalerna



www.marinetraffic.org




Beviljat stöd från VINNOVA

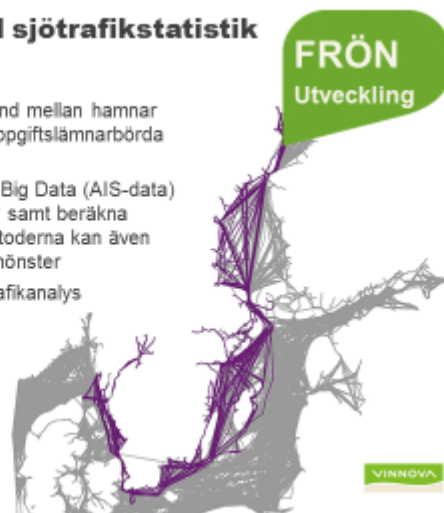
Motivering:

Projektet adresserar ett viktigt behov vars lösning kan leda till goda samhällseffekter i linje med utlysningens mål. Projektets resultat kommer att vara av stort värde för beslutsfattare samt minskad börda både för sjöfartsbranschen och hamnarna.

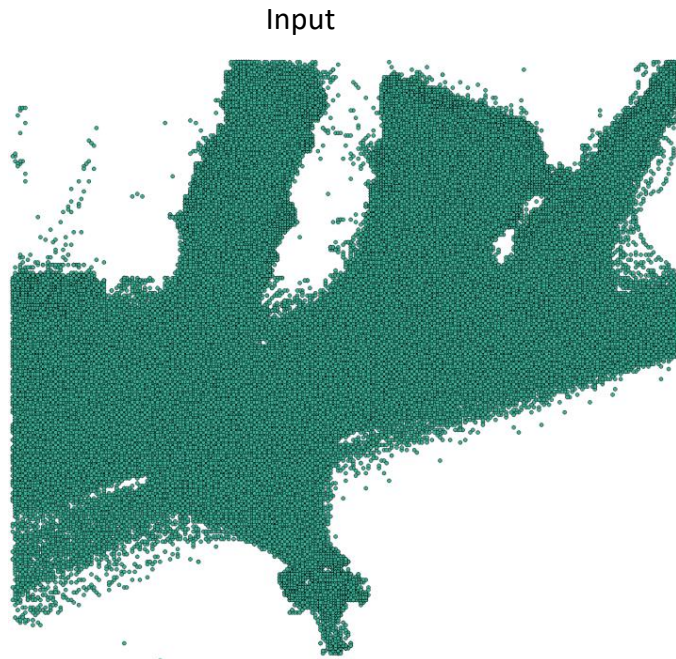
Vidare ser VINNOVA positivt på att detta projekt har en generaliserbarhet på andra områden än sjöfarten.

Metodutveckling för förbättrad sjötrafikstatistik med Big Data

-  Data om fartygsanlöp till hamnar samt avstånd mellan hamnar samlas idag in via en enkät, vilket innebär uppgiftslämnarbörda samt risk för mätfel
-  Utveckla metod för att använda geografiska Big Data (AIS-data) för att lista vilka fartyg som anländer hamnar samt beräkna noggrannare avstånd mellan hamnarna. Metoderna kan även användas för transportflöden och pendlingsmönster
-  Statistiska centralbyrån tillsammans med Trafikanalys



Förutsättningar: in- och utdata

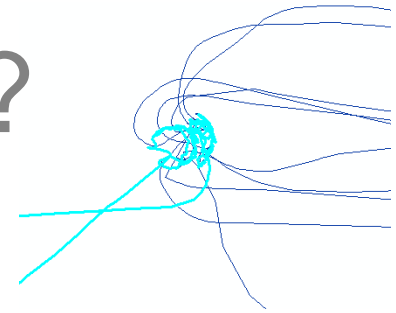


Output

- Avstånd mellan alla hamnar uppdelat på:
 - nationellt/internationellt/inlandsvatten
 - Fartygstyp (last/passagerare)
- Anlöpslistor per hamn
- Passagelinjer
- Trafikdensitet per län/övrig region.

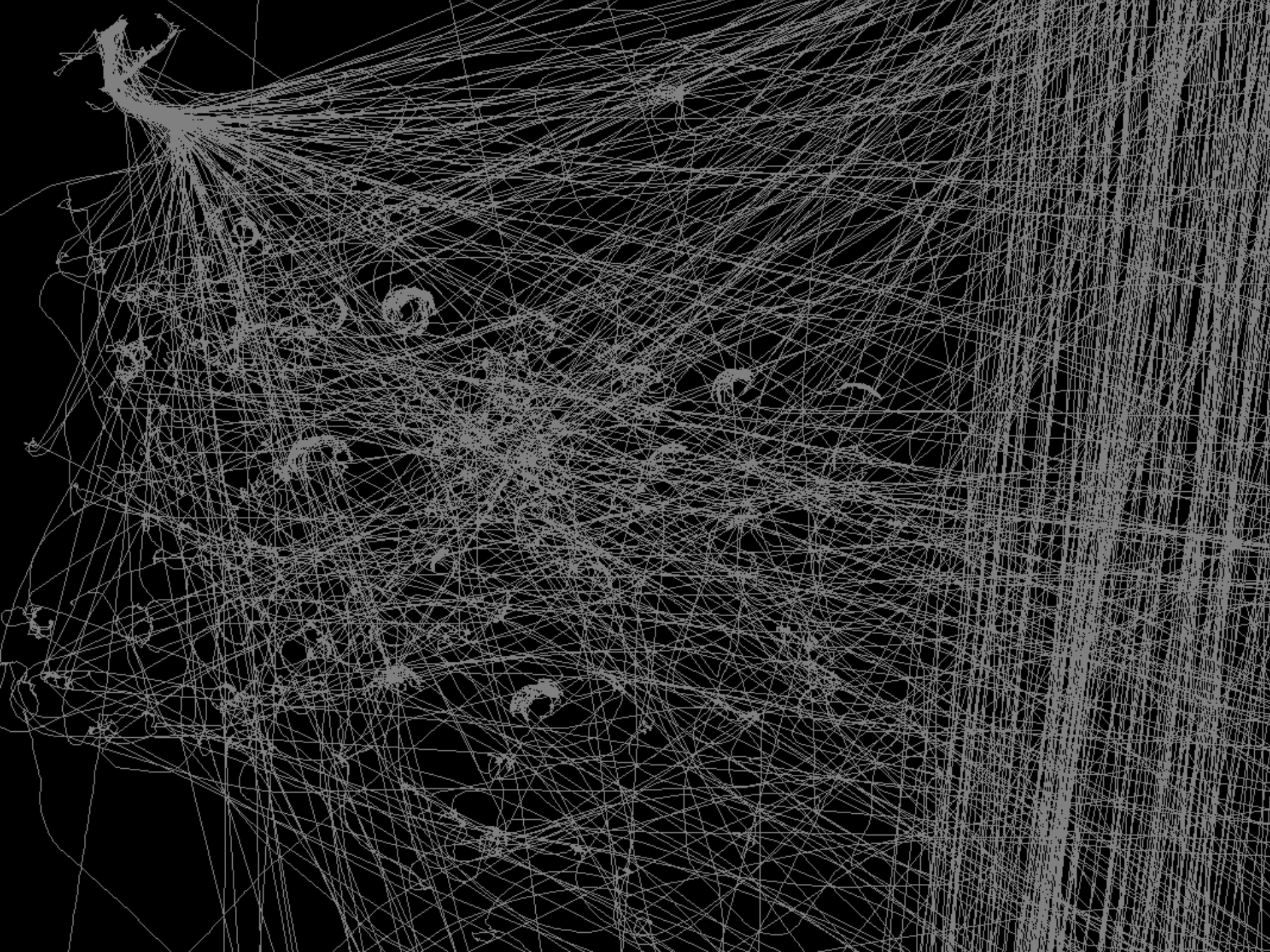
- AIS data för hela Östersjön för 2013-2015 från Sjöfartsverket/HELCOM
 - Ca 330 miljoner observationer per/år (2015)
 - Tidsintervall ca 6 sek. / 6 min. (svensk/Internationell data) , geografisk upplösning ca 5-10 cm.
- Varje punkt är en rapporterad position med uppgifter om:
 - Fartygsid, fartygstyp, storlek mm.
 - Tid, hastighet, kurs mm

Vilka utmaningar finns?



- Data måste omstruktureras
 - Minska ner datamängden för att kunna hantera den
 - Vi ska bara slutligen använda en liten del av all data(?!)
- Del av datafiltreringen och bearbetningar måste göras geografiskt (i GIS- geografiska informationssystem)
 - Geografiska beräkningar kan bli tunga
 - Viktigt med genomtänkt system
- Vi måste identifiera transporterna till och från svenska hamnar
 - Linjer måste skapas av punktinformationen
 - Angöringsområden (hamnar) måste avgränsas

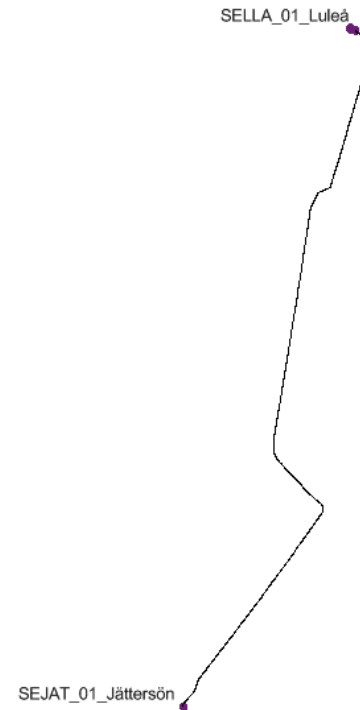




Omstrukturera data: skapa linjer

Fyra attribut behövs för att skapa transportlinjer mellan hamnar/angöringsområden:

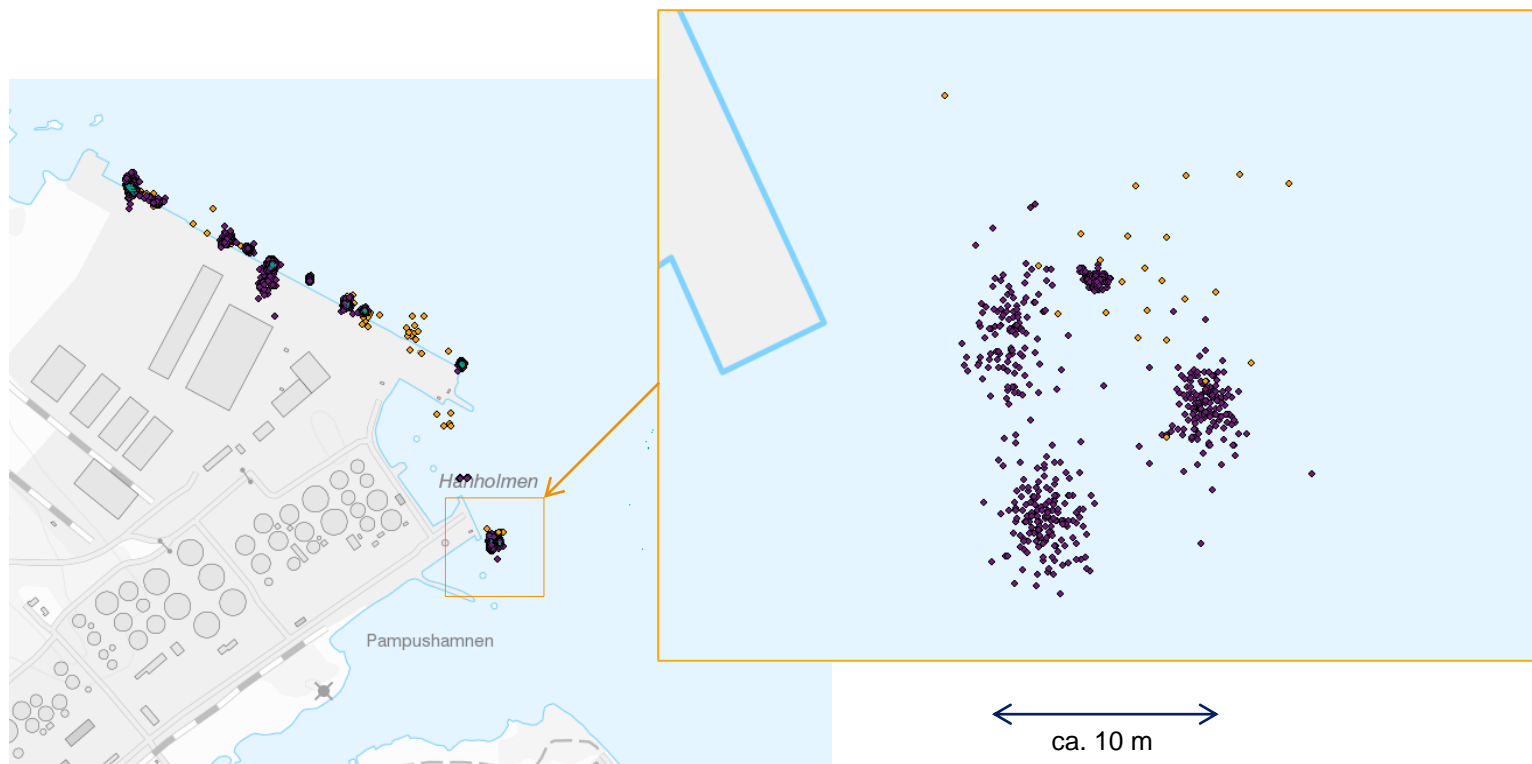
- Unikt id för fartyg (MMSI/IMO)
- Position
- Tid
- Angöringsområde
 - Finns "gömt" i data och kan plockas fram (data mining!)



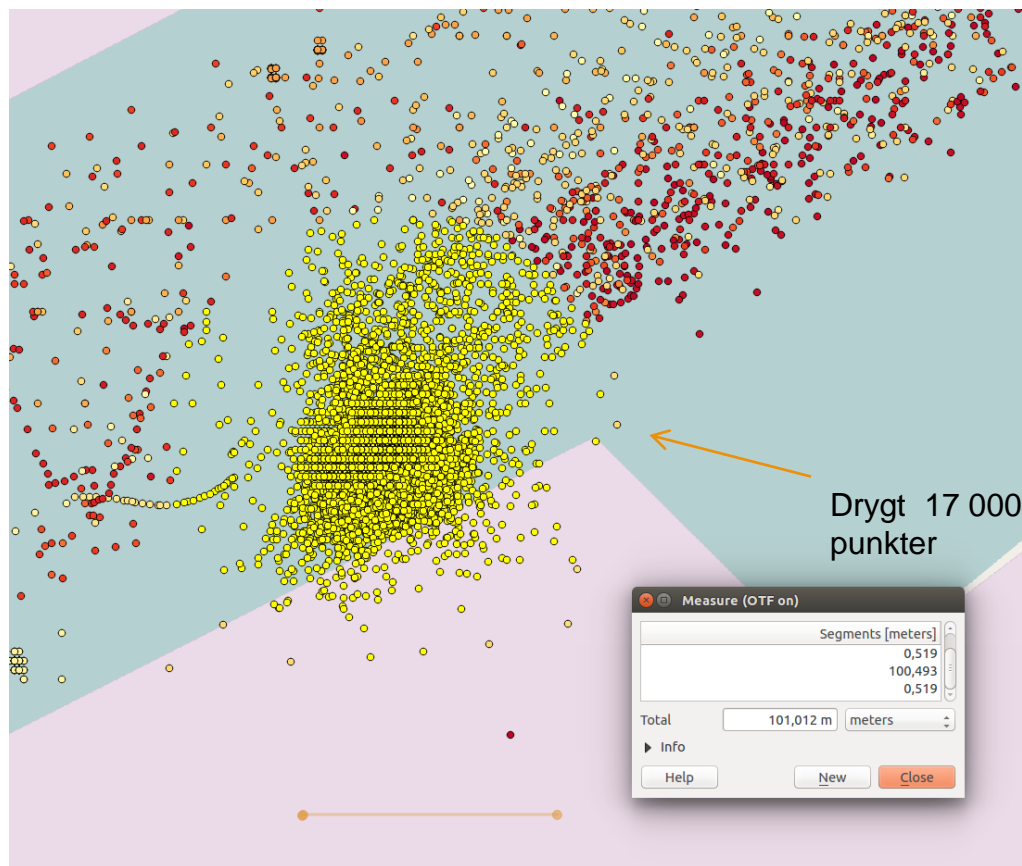
Angöringsområden utifrån AIS



Högupplöst data: Överflödigt – eller användbart?



Högupplöst data: Överflödigt – eller användbart?

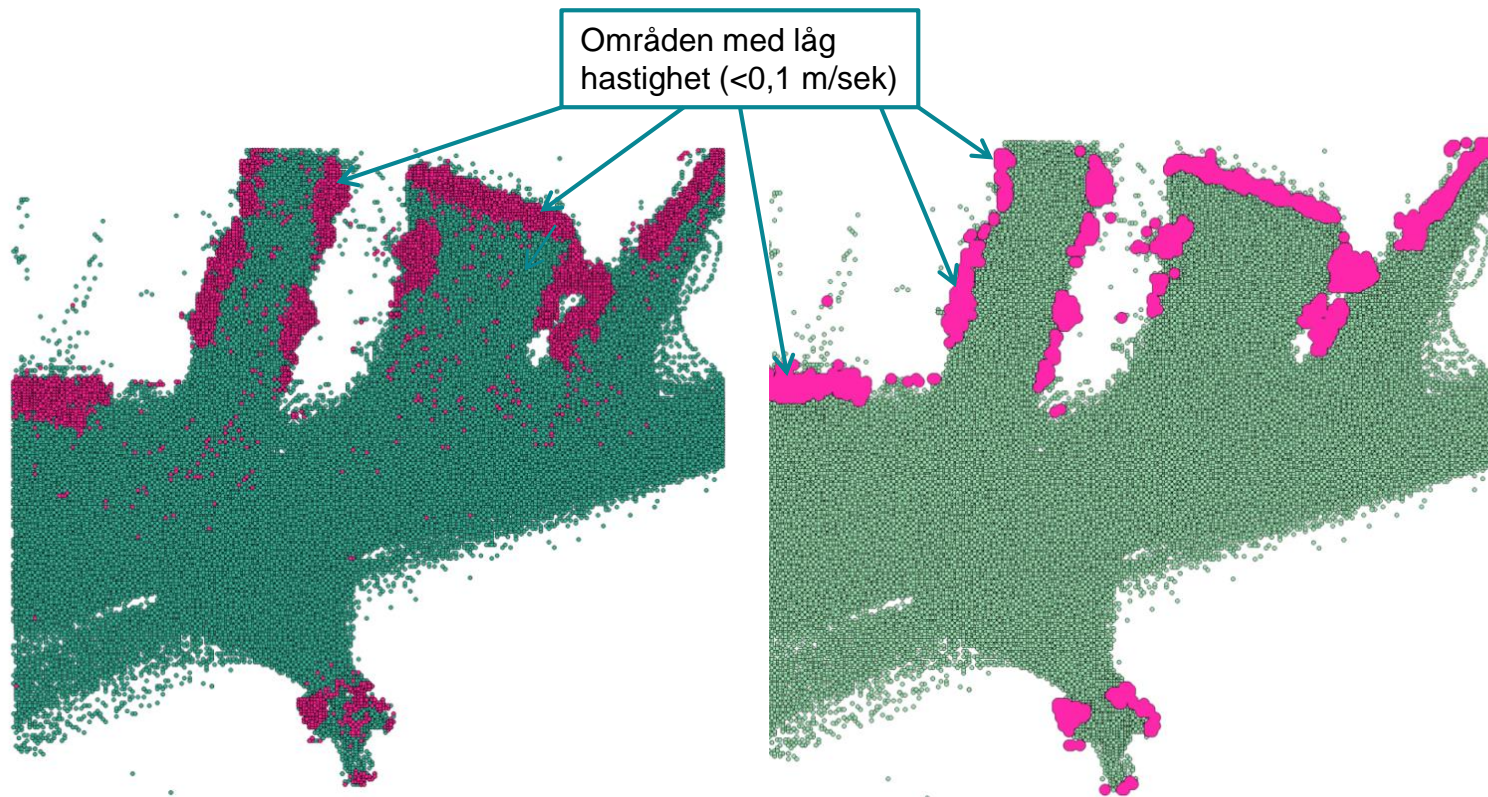


[galaxy test 2\frihamn all3 8.html](#)

[galaxy test 2\frihamn all3 6.html](#)



Högupplöst data: Överflödigt – eller användbart?



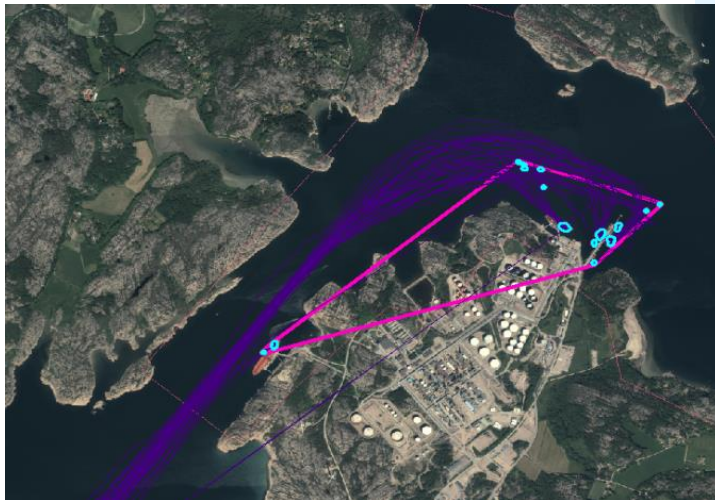
Originaldata (ca 10 cm upplösning)
Ca 1 miljon punkter, del av Göteborgs hamn

Aggregerad data (10 m upplösning)
medelhastighet per 10 m ruta



Angöringsområden, status och resultat

- Ett pågående arbete, än så länge version 1b:
 - Grova angöringsområden skapade för samtliga länder runt Östersjön
 - Mer bearbetade, noggranna angöringsområden skapade för Sverige

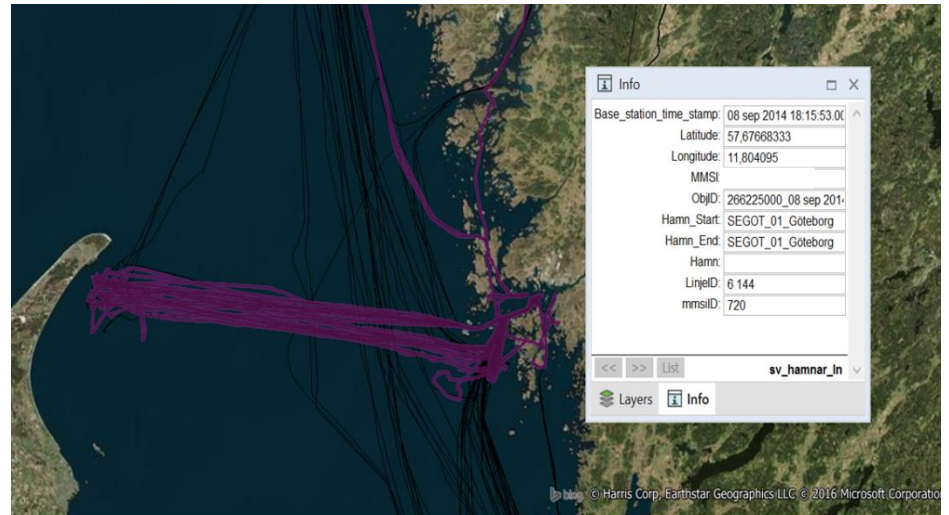


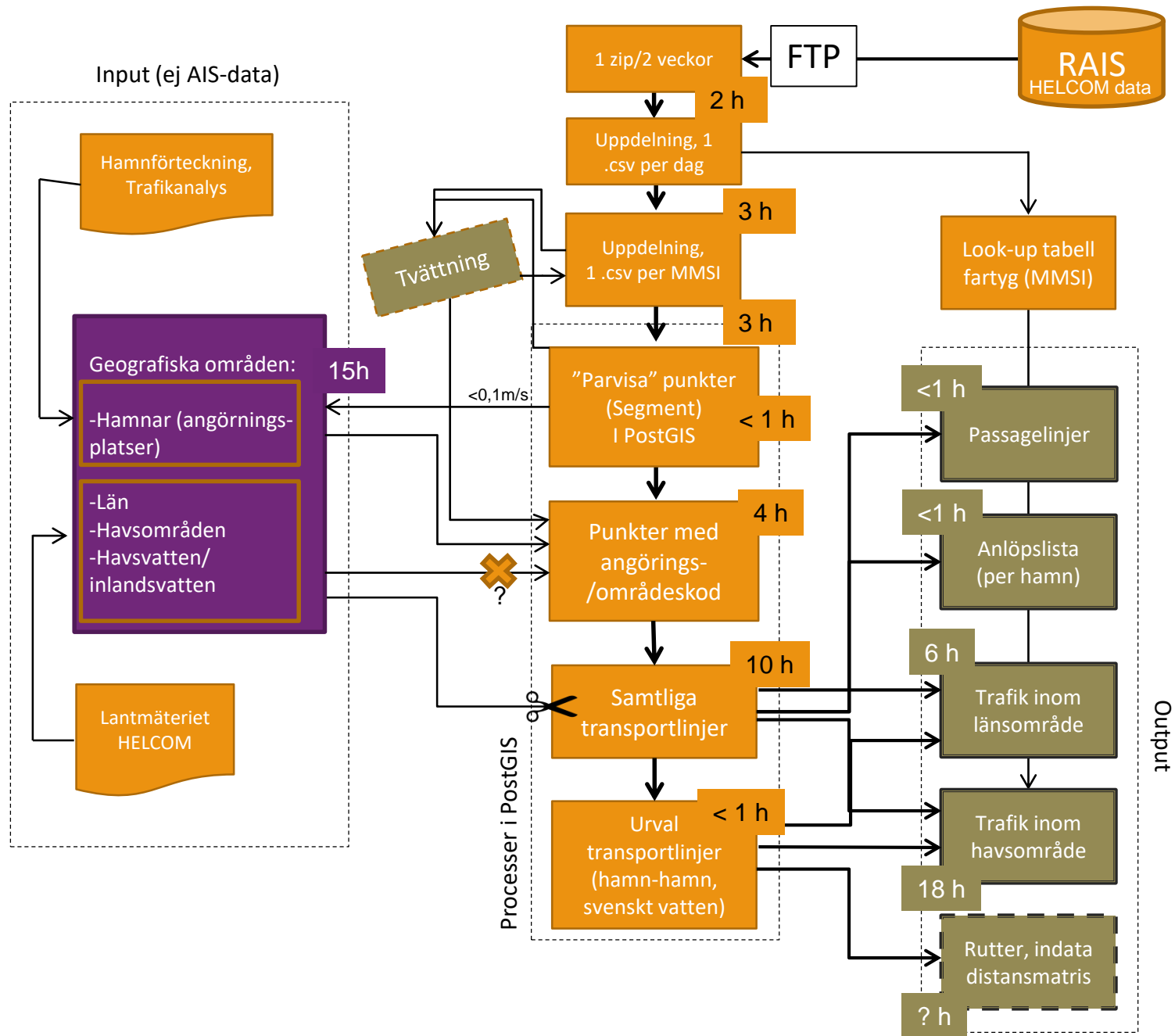
Angöringsområden = hot spots med stillastående fartyg



Vi får med:

- Ankringsplatser
- Slussar
- Diverse väntplatser





Punkter med angöringskod

Output pane

Data Output Explain Messages History

	ts timestamp without time zone	ts2 timestamp without time zone	tdelta double precision	speed real	port character varying(18)
1	2015-03-09 16:34:31	2015-03-09 16:43:34	543	0.00453088	SE0100 4723
2	2015-03-09 16:43:34	2015-03-09 16:59:51	977	0.00475342	SE0100 4723
3	2015-03-09 16:59:51	2015-03-09 17:13:50	839	0.00926859	SE0100 4723
4	2015-03-09 17:13:50	2015-03-09 17:17:53	243	0.00659043	SE0100 4723
5	2015-03-09 17:17:53	2015-03-09 17:35:20	1047	0.00943702	SE0100 4723
6	2015-03-09 17:35:20	2015-03-09 17:35:50	30	1.23674	SE0100 4723
7	2015-03-09 17:35:50	2015-03-09 17:36:21	31	4.80451	SE0100 4723
8	2015-03-09 17:36:21	2015-03-09 17:37:20	59	7.12447	
9	2015-03-09 17:37:20	2015-03-09 17:38:51	91	5.88451	
10	2015-03-09 17:38:51	2015-03-09 17:39:19	28	11.5544	
11	2015-03-09 17:39:19	2015-03-09 17:39:49	30	16.7758	
12	2015-03-09 17:39:49	2015-03-09 17:40:20	31	18.6489	
13	2015-03-09 17:40:20	2015-03-09 17:42:49	149	19.487	
14	2015-03-09 17:42:49	2015-03-09 17:43:21	32	20.1612	
15	2015-03-09 17:43:21	2015-03-09 17:43:49	28	22.3896	
16	2015-03-09 17:43:49	2015-03-09 17:44:20	31	21.166	Transport
17	2015-03-09 17:44:20	2015-03-09 17:45:19	59	22.0609	
18	2015-03-09 17:45:19	2015-03-09 17:45:52	33	21.8253	
19	2015-03-09 17:45:52	2015-03-09 17:47:19	87	18.0234	
20	2015-03-09 17:47:19	2015-03-09 17:47:52	33	6.15023	
21	2015-03-09 17:47:52	2015-03-09 17:48:19	27	5.09267	
22	2015-03-09 17:48:19	2015-03-09 17:48:49	30	10.7341	
23	2015-03-09 17:48:49	2015-03-09 17:49:20	31	7.97329	
24	2015-03-09 17:49:20	2015-03-09 17:49:49	29	5.39264	
25	2015-03-09 17:49:49	2015-03-09 17:50:19	30	0.912759	SE0100 4420
26	2015-03-09 17:50:19	2015-03-09 17:56:20	361	0.0162565	SE0100 4420
27	2015-03-09 17:56:20	2015-03-09 18:02:19	359	0.0204591	SE0100 4420
28	2015-03-09 18:02:19	2015-03-09 18:08:19	360	0.0160997	SE0100 4420
29	2015-03-09 18:08:19				
30	2015-03-09 18:08:19				
31	2015-03-09 18:08:19				
32	2015-03-09 18:08:19				
33	2015-03-09 18:08:19				
34	2015-03-09 18:08:19				
35	2015-03-09 18:08:19				
36	2015-03-09 18:08:19				
37	2015-03-09 18:08:19				
38	2015-03-09 18:08:19				

← Lämnar ett område

← Kommer in i nästa område
← Står still i området



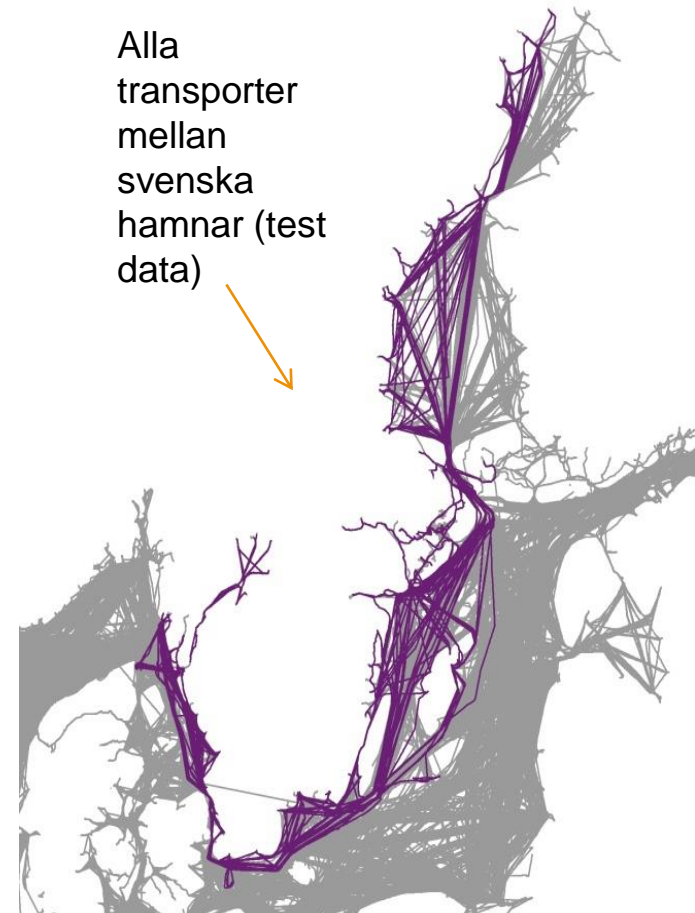
Från punkter med angöringskod till linje

En linje skapas av de punkter som ligger tidsmässigt mellan det att ett fartyg lämnar ett område och till dess det kommer till ett nytt område.

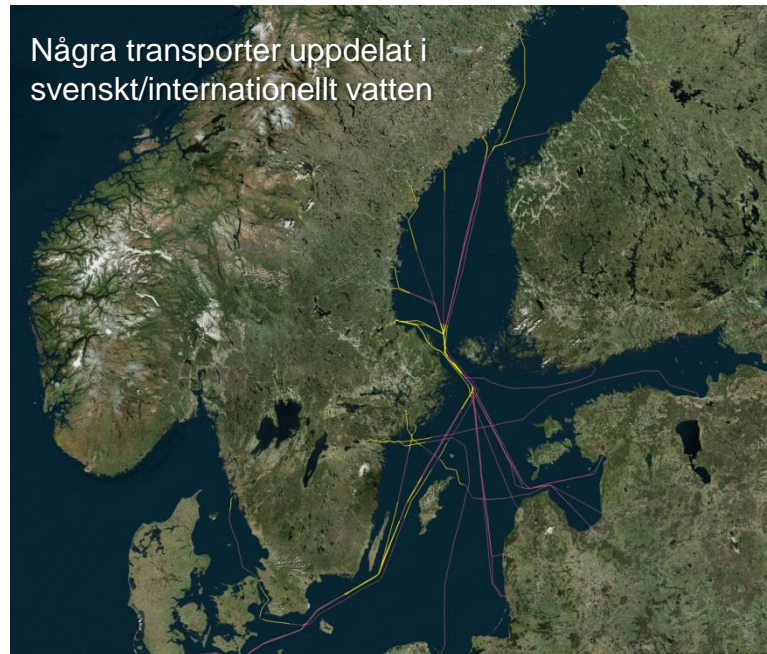
Några fördelar:

- Vi har reducerat data vi jobbar med
- Vi behöver inte skapa onödiga linjer inom ett område
- Vi blir flexibla och kan plocka fram de transporter vi är intresserade av.
- Vi sparar lagringsplats – vi kan generera linjer ad hoc

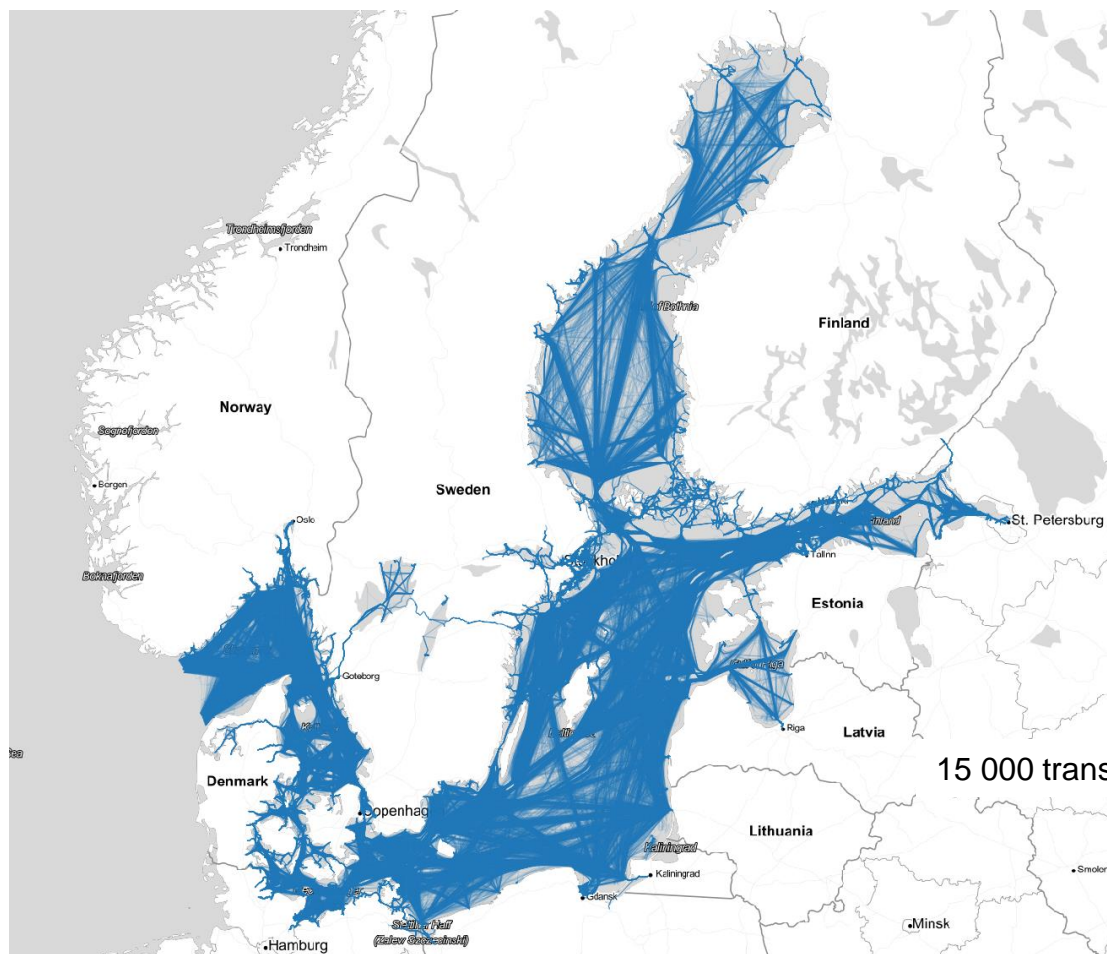
Och vi kan använda linjerna som input till modell för avståndsberäkningar i distansmatrisen!



Uppdelning av transporter



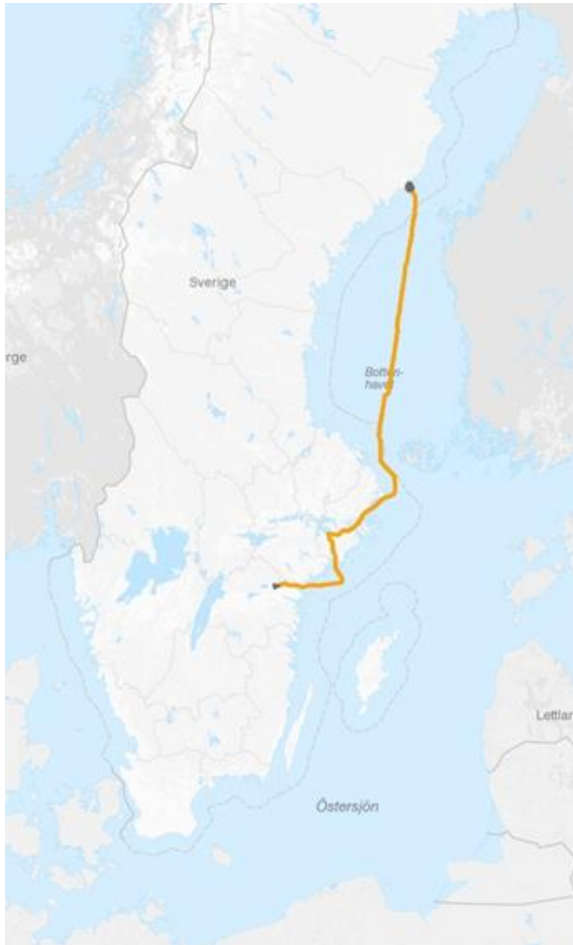
Distansmatrisen



15 000 transporter från 2015



Distansmatrisen



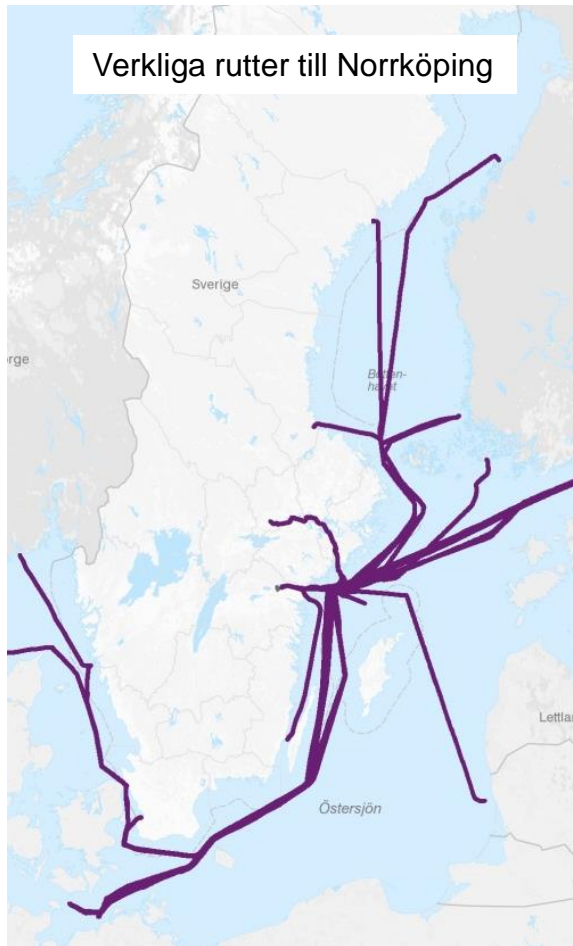
- Avstånd mellan hamnar används för att beräkna transportarbete (tonkilometer och personkilometer)
- Avståndet bör motsvara den vanligaste rutten.
- Ett transportnätverk skapas
 - Byggt på de linjer vi tidigare skapade, konverterade till ett raster (rutor) med 1 km upplösning
 - Varje ruta i nätverket är viktat efter
 - trafikdensitet
 - destination
- Rutten beräknas efter viktningen så att den prioriterade rutten blir där flest fartyg redan åkt.

Distansmatrisen



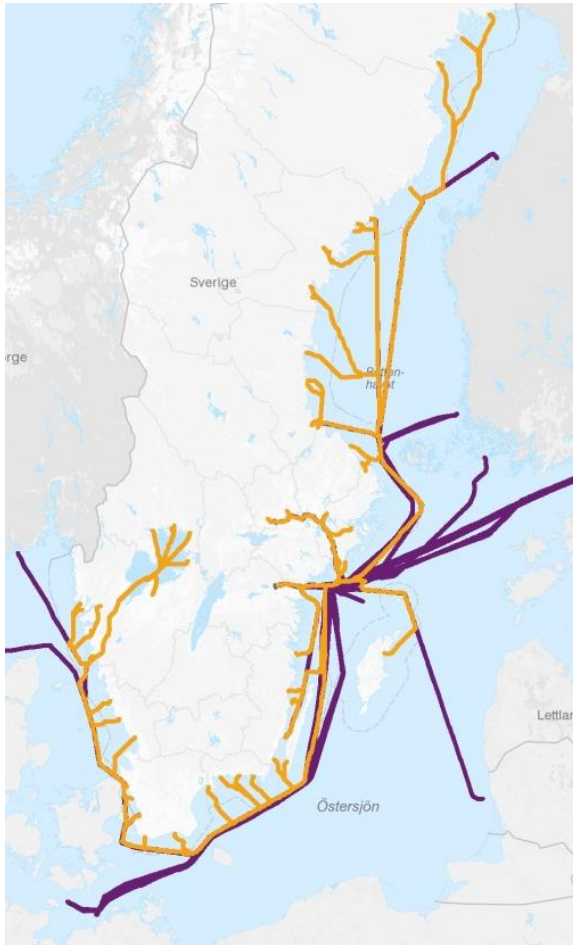
- Avstånd mellan hamnar används för att beräkna transportarbete (tonkilometer och personkilometer)
- Avståndet bör motsvara den vanligaste rutten.
- Ett transportnätverk skapas
 - Byggt på de linjer vi tidigare skapade, konverterade till ett raster (rutor) med 1 km upplösning
 - Varje ruta i nätverket är viktat efter
 - trafikdensitet
 - destination
- Rutten beräknas efter viktningen så att den prioriterade rutten blir där flest fartyg redan åkt.

Distansmatrisen



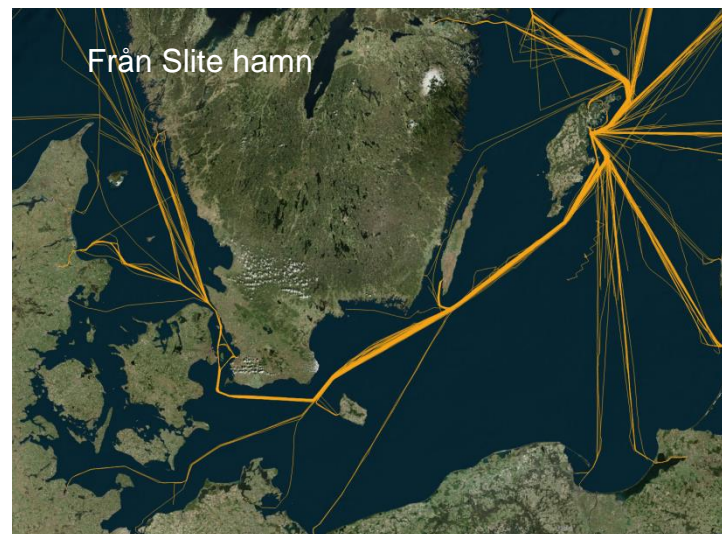
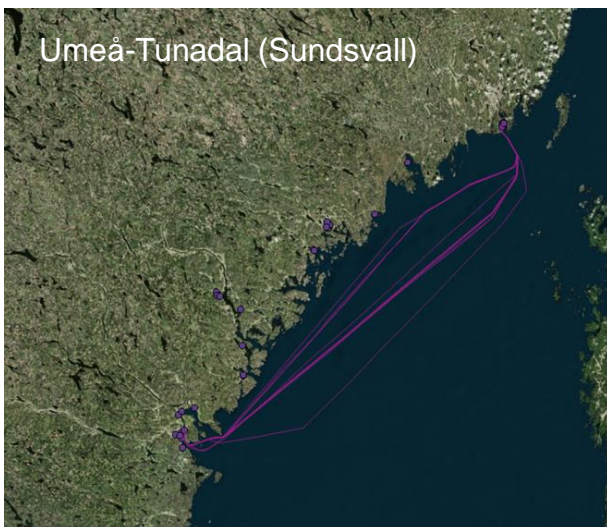
- Avstånd mellan hamnar används för att beräkna transportarbete (tonkilometer och personkilometer)
- Avståndet bör motsvara den vanligaste ruten.
- Ett transportnätverk skapas
 - Byggt på de linjer vi tidigare skapade, konverterade till ett raster (rutor) med 1 km upplösning
 - Varje ruta i nätverket är viktat efter
 - trafikdensitet
 - destination
- Ruten beräknas efter viktningen så att den prioriterade ruten blir där flest fartyg redan åkt.

Distansmatrisen



- Avstånd mellan hamnar används för att beräkna transportarbete (tonkilometer och personkilometer)
- Avståndet bör motsvara den vanligaste rutten.
- Ett transportnätverk skapas
 - Byggt på de linjer vi tidigare skapade, konverterade till ett raster (rutor) med 1 km upplösning
 - Varje ruta i nätverket är viktat efter
 - trafikdensitet
 - destination
- Rutten beräknas efter viktningen så att den prioriterade rutten blir där flest fartyg redan åkt.

Den rätta vägen?

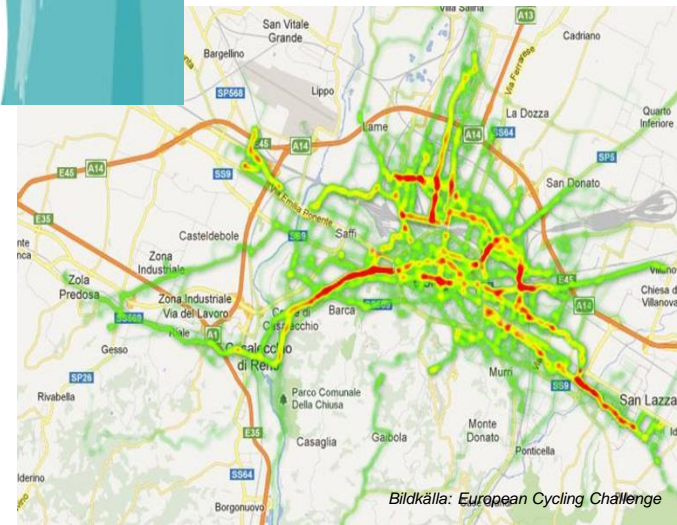


Andra användningsområden?



Bildkälla: <http://gisgeography.com/10-global-transportation-maps/>

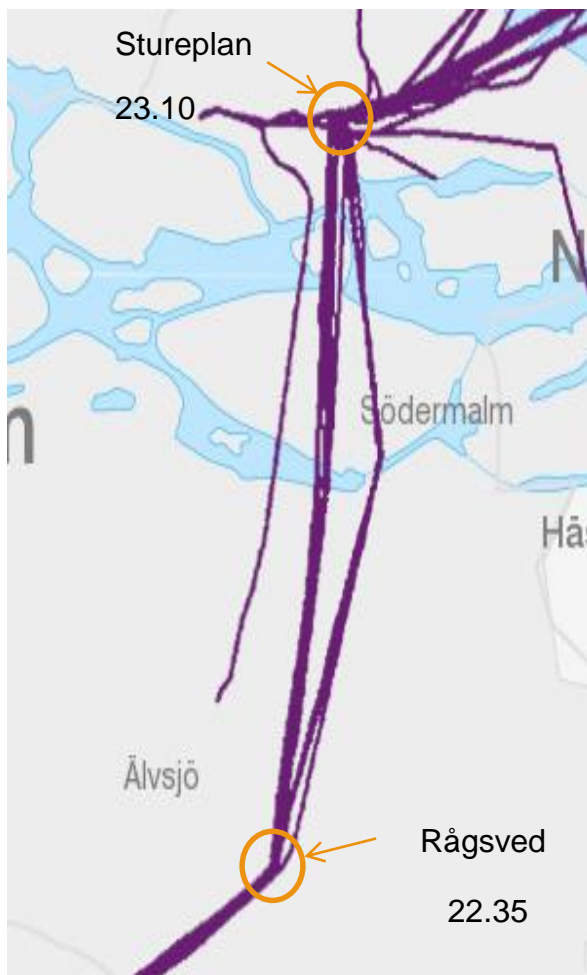
t.ex. med mobildata?



Bildkälla: European Cycling Challenge



Stureplan fredag kl 23

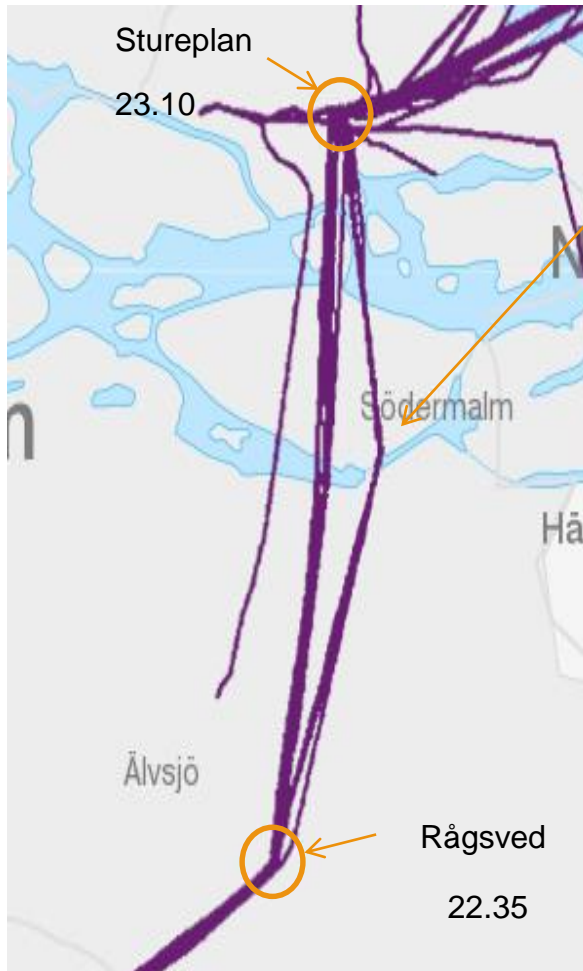


Mobildata innehåller all grundinformation som behövs :

- Id
- Tid
- Position



Till Stureplan från Rågsved

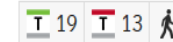


Ca 35 min = tunnelbana?

22:36 » 23:10

Restid 34 min

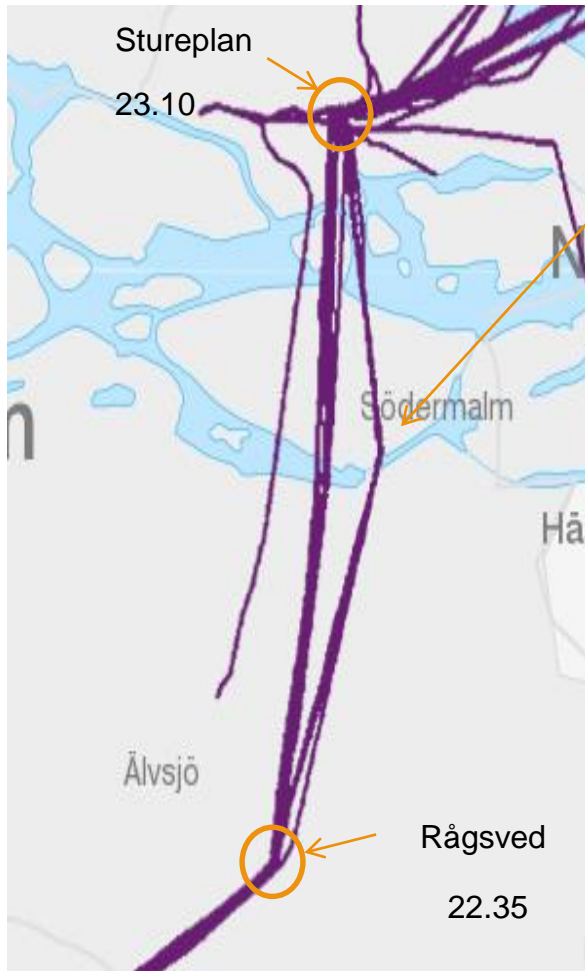
Rågsved » Östermalmstorg



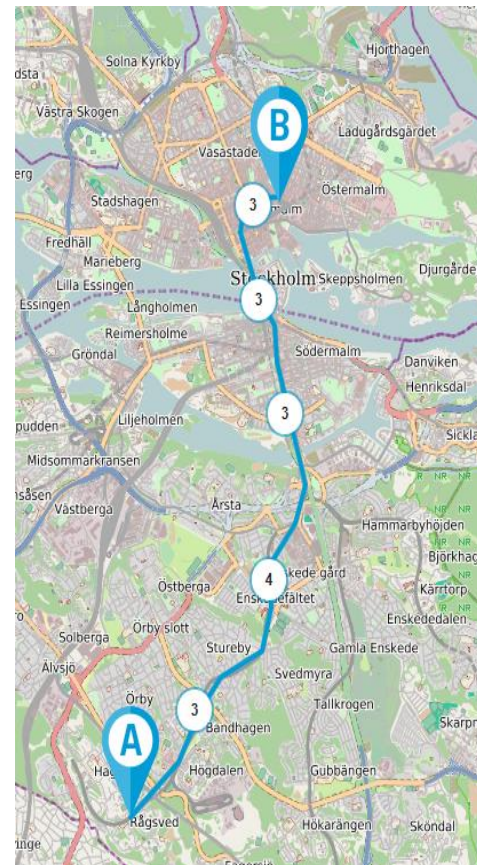
- 22:36 Rågsved
 - Tunnelbanans gröna linje 19 mot Hässelby strand
- 22:54 Slussen
- 22:57 Slussen
 - Tunnelbanans röda linje 13 mot Ropsten
- 23:04 Östermalmstorg
- 23:05 Östermalmstorg Gå 179 meter
- 23:10 Stureplan (på Birger Jarlsgatan)



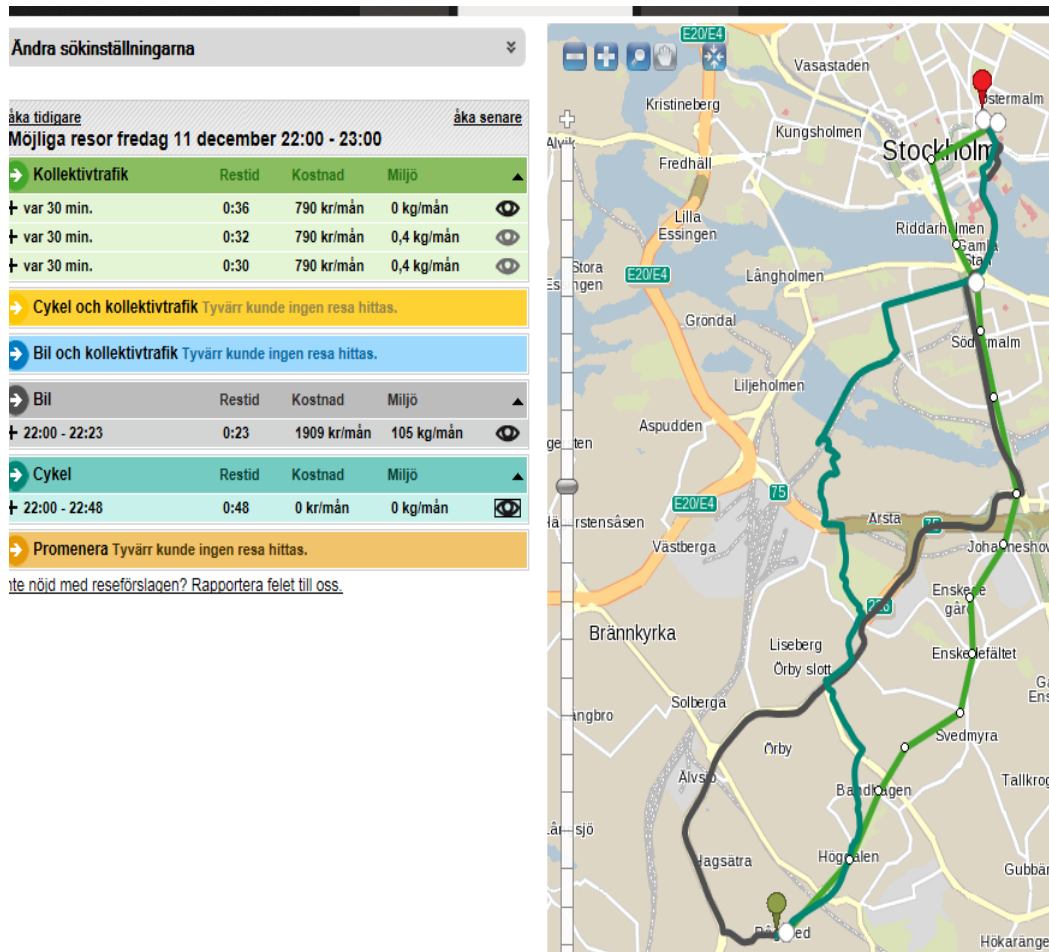
Till Stureplan från Rågsved



Ca 35 min = tunnelbana?



Tid och linjegeometri ger transportslag



FRÖN – För ökad innovation i offentligt finansierad verksamhet

1. Planeringsprojekt

- Ett planeringsprojekt 2015 bedömde att AIS-data är användbart som källa till ny och förbättrad sjötrafikstatistik
- Projektparter: SCB (koordinator) och Trafikanalys
- Budget 0,6 mkr, varav 65 procent finansierades av Vinnovas utlysning FRÖN

2. Utvecklingsprojekt

- 2016–2017 pågår utvecklingsprojektet *Metodutveckling för förbättrad sjötrafikstatistik med Big Data*
- Budget 1,4 mkr, varav 50 procent finansieras av Vinnovas utlysning FRÖN
- Referensgrupp med Sveriges hamnar, Sjöfartsverket, Sjöfarts-högskolan, Havsmiljöinstitutet, KTH, SMHI, Energimyndigheten

3. Införandeprojekt

- Ska resultaten användas i statistiken kommer Trafikanalys att behöva ett införandeprojekt i nästa steg



Hård- och mjukvara

Datorn:

64 GB RAM

2 TB SSD (Nya datorn)

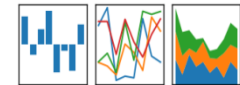
4 kärnor, 8 trådar

Ca 700 gb data i databas

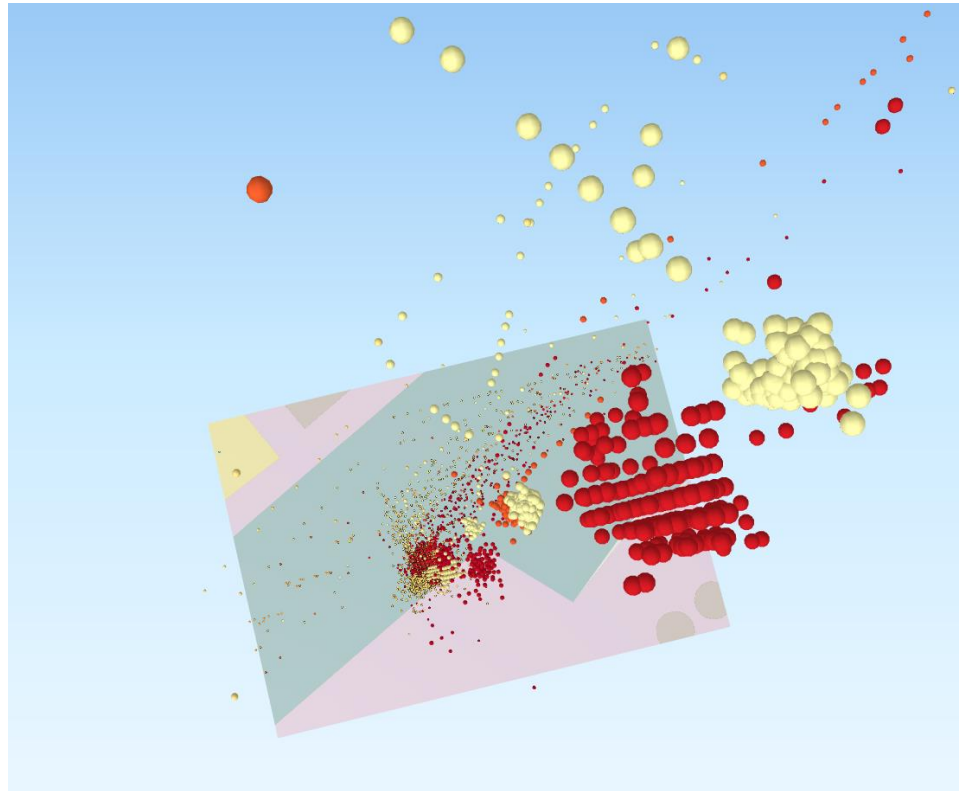


pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Tack för er uppmärksamhet!



Frågor?

