MOBILITY
ANALYTICS

Foreign Tourists on the Lyngen Alps

a Mobility Analytics report for
Verneområdestyret for Lyngsalpan, Fylkesmannen i Troms

CONTENTS

1   Introduction

Mobility Analytics pilot for the Lyngen Alps

Verneområdestyret for Lyngsalpan has been in touch with Telenor Norway to investigate opportunities for a pilot in the area of Mobility Analytics. In a joint workshop, several ideas were discussed and evaluated and a first pilot's scope was proposed to include the primary goal of estimating number and origin of tourists traveling to the Lyngen Alps during a single month. Based on the proposed project plan, and the importance of the Lyngen Alps project for Telenor, the period of observations was extended without an impact on the price of the pilot. The measurements were therefore performed between February the 11th and May the 1st of 2017.

2   Mobility Analytics pilot

2.1   Mobility Analytics as a service

Mobility Analytics is Telenor's solution to provide relevant insights on human mobility patterns. Telenor Norway is leveraging techniques developed at Telenor Group Research over the past years in order to extract mobility patterns from Telenor's network data.
Knowing how people's mobility is affected by different factors, such as time of the day, day of the week, weather or special events can help businesses make better decisions and prepare and tailor their offerings.

2.2   Methodology and Privacy

Signaling data stemming from Telenor's mobile network, once processed in a secure way with respect to end-user's privacy, can be aggregated and extrapolated to create measurements and visualizations of human mobility patterns.

Collecting and processing of location data is a highly sensitive topic. Processing and storing data on individual level is tightly regulated in Norway.

Telenor utilized a three-step "Privacy by Design" framework to assure the privacy of end users as well as to provide an estimate for the whole population based on extrapolation techniques. Based in the experience from other pilots where a tested ground truth was available, considering Telenor's market share of cca. 60% and SSB data on Norwegian population Telenor has obtained an extrapolating factor that allows us to go from the number of people connected to our network to an estimate on the total number of people present in any area.

«Privacy by Design» framework

I. Masking (Pseudonymization)   II. Aggregation (Anonymization)   III. Prediction (Extrapolation)

Figure 1. «Privacy by Design» framework steps overview

2.2.1. K-Anonimity effects

In according to Telenor's privacy methodology, only measurements for countries with a high presence of visitors in Norway can be reported, avoiding reporting on the presence of visitors from "rare" countries. In addition to this measure, 20 or more Norwegians and 5 or more non-Norwegian subscriptions in single measurements were promoted into the analysis. This on its own limits the reporting possibilities of the rare occurrence of nationalities per region (four or less visitors from a country were not reported).

In the case of the measurements running in the Lyngen Alps, necessary for this pilot, the large regions of the Alps ensured a high presence of tourists all together. The cases where the number of visitors was less than 5 was very limited.

2.3   Objective and scope

The objective of the delivery of this pilot was to provide measurements on the number of tourists present in the Lyngen Alps region. The statistics are as detailed as possible with the boundaries set due to the Privacy framework. Telenor delivered an overview of estimated counts of foreign tourists using the country of origin information (based on the mobile subscription's country of origin).
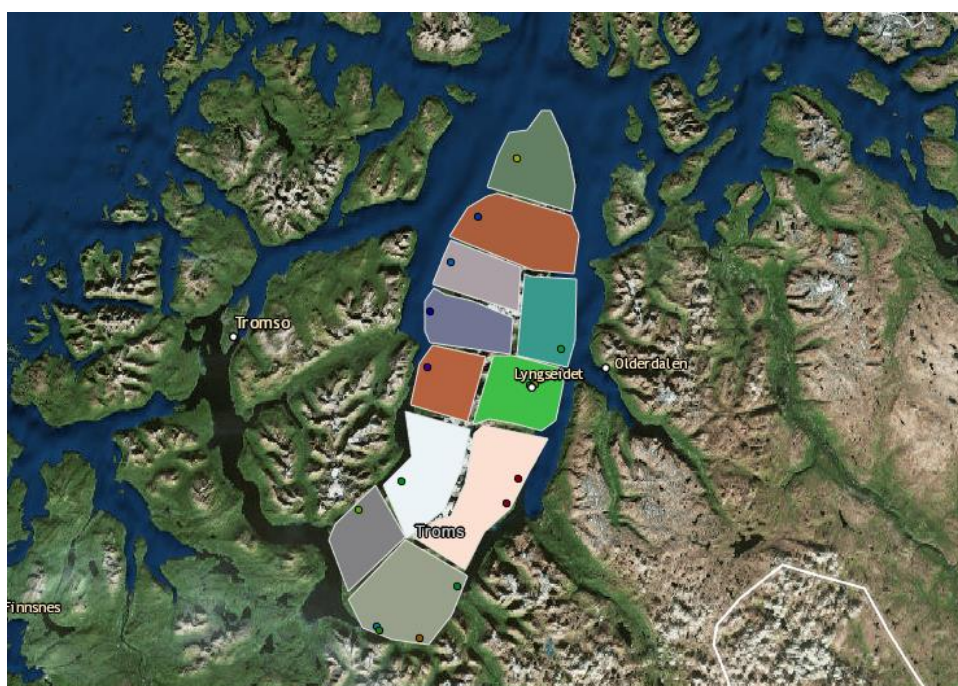

Figure 2. Lyngen Alps region divided according to the coverage of 15 base stations present in the region.

The measurements were conducted utilizing the cellular infrastructure present in the Lyngen Alps: 15 base stations covering the whole region as shown in Figure 2. In accordance with Verneområdestyret for Lyngsalpan the measurements were taken with an hourly frequency for the whole period.

The scope of this pilot is to present as a result of the measurements snapshots on the presence of tourists every hour. It was out of scope to measure the number of unique vs. recurrent visitors. Tracking of tourists was also out of scope for this pilot.

2.4 Telenor's measurements

Telenor has conducted measurements in total at sixteen base stations, at one hour intervals in the time window between February the 11th and May the 1st of 2017. The format of the output of the measurements is as shown in Figure 3.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	time	easting	northing	Norway	Sweden	Denmark	Germany	Other	Switzerlar	United Kir	Lithuania	Latvia	Finland	Italy
2	11.02.2017 14:00	682692	7709983	544	21	19	18	31	6	15	20	14	5	12
3	11.02.2017 15:00	682692	7709983	542	21	19	15	31	7	13	19	14	6	10
4	11.02.2017 16:00	682692	7709983	544	21	19	19	29	7	14	18	15	5	11
5	11.02.2017 17:00	682692	7709983	549	22	19	19	36	6	15	20	16	5	16
6	11.02.2017 18:00	682692	7709983	556	21	20	43	26	12	19	20	17	5	18
7	11.02.2017 19:00	682692	7709983	559	21	20	56	28	17	26	20	19	5	19
8	11.02.2017 20:00	682692	7709983	564	21	20	70	40	28	67	19	18	6	37
9	11.02.2017 21:00	682692	7709983	559	20	20	65	33	29	76	19	17	5	55
10	11.02.2017 22:00	682692	7709983	560	20	20	67	45	27	76	20	16	0	53
11	11.02.2017 23:00	682692	7709983	557	18	20	51	50	21	75	22	16	0	70
12	12.02.2017 00:00	682692	7709983	555	18	19	44	54	18	56	21	15	0	38
13	12.02.2017 01:00	682692	7709983	554	17	20	40	50	12	45	21	15	0	29
14	12.02.2017 02:00	682692	7709983	551	17	19	34	40	8	24	20	17	0	26
15	12.02.2017 03:00	682692	7709983	551	18	19	28	31	8	17	21	15	0	26
16	12.02.2017 04:00	682692	7709983	549	17	19	22	36	6	12	20	13	0	26

Figure 3: Header and sample of mobile sourced dataset

The measurements delivered from this pilot were presented in sixteen different spreadsheets, one for each of the base stations in the Lyngen Alps region.

2.4.1. Dataset fields

- **time.** Represents a time stamp indicating when the measurements of the row were taken in the fomat dd.mm.yyyy hh:mm.
- **easting.** Represents an approximate location of the "x-coordinate" of the base station taking the measurements. The projection used for the x-coordinate is a standard for Norway: UTM 33N.
- **northing.** Represents an approximate location of the "y-coordinate" of the base station taking the measurements. The projection used for the y-coordinate is a standard for Norway: UTM 33N.
- **Norway, Sweden, Denmark, Germany, etc.** These fields indicate the country of origin of the people present in the region and their total number estimated. Countries are ordered in a descending manner. When the field indicates 0, it means that the number of present tourists was four or less.

2.6.1. The overspill effect

One "challenge" during the collection of measurements came from the overspill of coverage in the base stations present in the Lyngen Alps region. It was in general not possible to filter

out people present only within the Alps. It was a common thing that base stations were providing a wide coverage for urban areas outside the Alps.

This effect is easily appreciated due to the fact that during night time there was a remanence of people reported as "present" in the Alps. This is an effect we are aware of and are working to overcome as part of the future service.

One of the consequences of this effect is that it is not possible to distinguish, at this stage, between people camping in the Alps region and people spending the night in their houses in urban areas near to the Lyngen Alps.

2.5   Quality of network measurements

Telenor has conducted "best-effort" measures to deliver the pilot and ensure the data quality going into the piloting case. There were no downtimes recorded for the measurements during the time window between February the 11th and May the 1st of 2017. There were 360 measurements daily (24 measurements times 15 measuring points, base stations).

3. Results

In this section we will present the results of the measurements taken during the pilot's time window. The results will be divided in sixteen sections, corresponding to the measurements, at time series level, taken by the different base stations.

At the end of this Results section, we will present they key results of the pilot.

3.1 Description of measurements per base station

We present now the key output of each one of the sixteen excel files "Base0-15.exe": a time series displaying the presence of people, broken down per nationality, per time of measurement (time window as specified above for the pilot).

It is advised for the reader to keep Figure 4 at hand: the correspondence between the base station numbering and the geography cover by this, is represented in Figure 4.
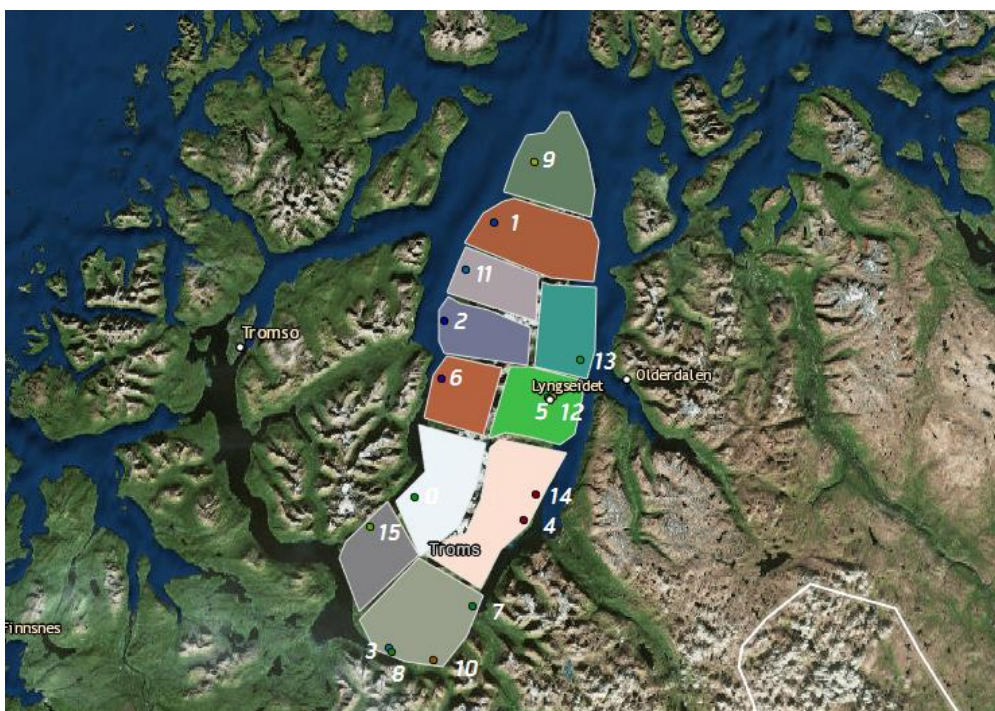
Figure 4: Lyngen Alps region. This figure presents the Lyngen Alps divided according to the approximate coverage regions of the sixteen base stations. The base stations present in the area are numbered.

3.1.1 Base station 0

- In this region, there were tourists coming from 44 different countries.
- Maximum presence of Norwegians: ~750
- Maximum presence of tourists: ~350
- 9th place for National visitors (out of the 16 measured)
- 9th place for Foreign visitors (out of the 16 measured)
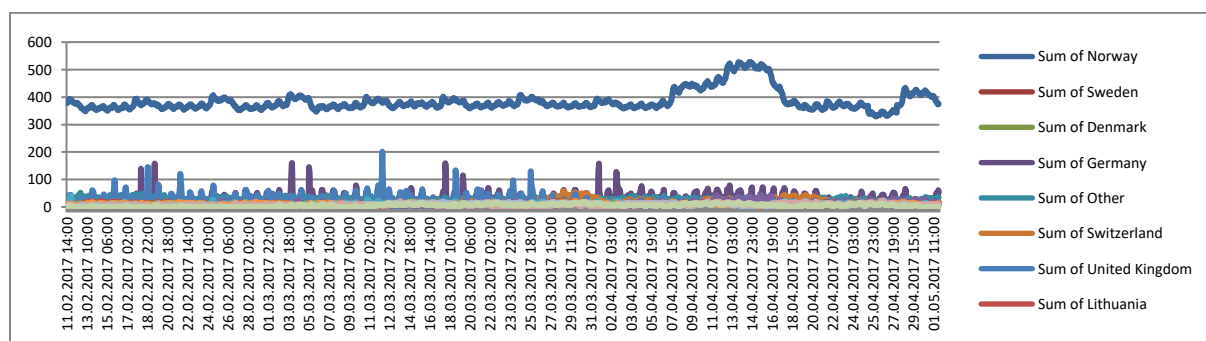- %Max Norwegian /Max Foreign: 46% foreign tourists
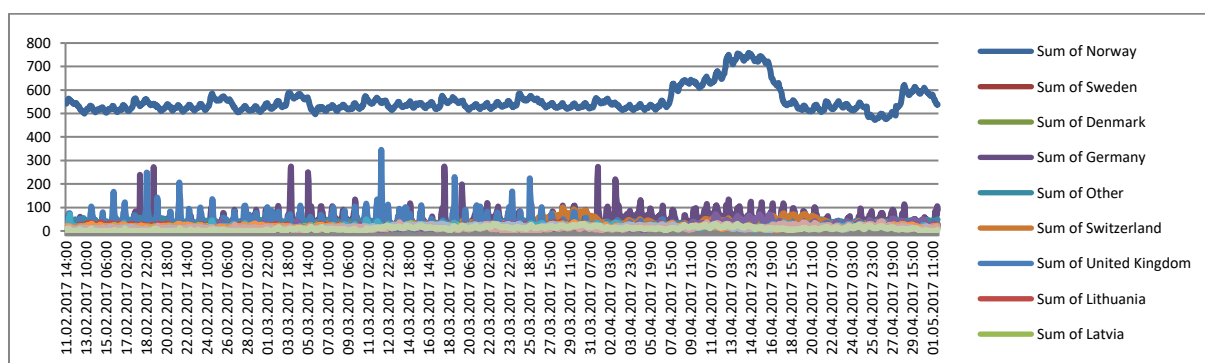
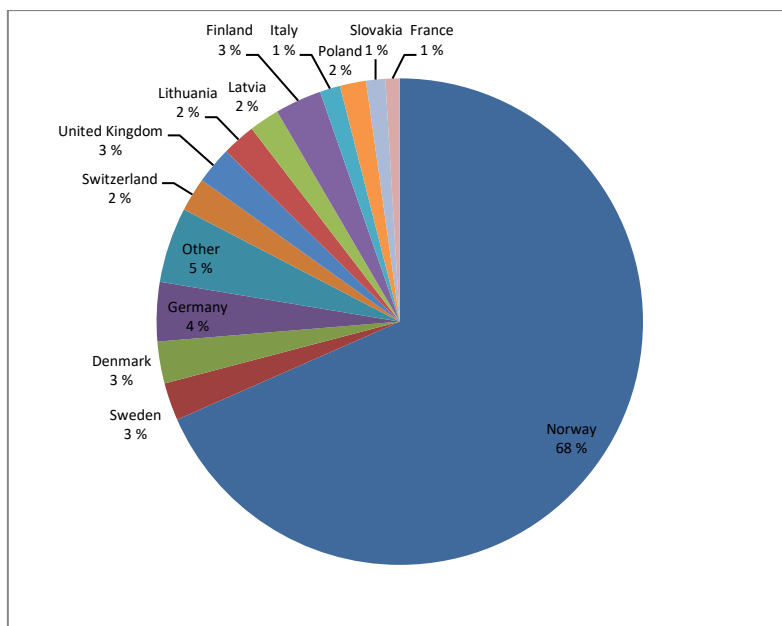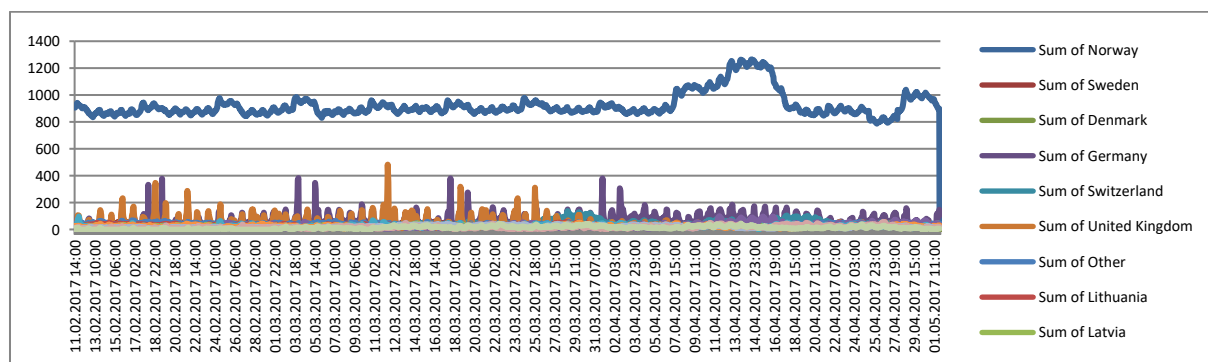Figure 5: Time-series base station 0.
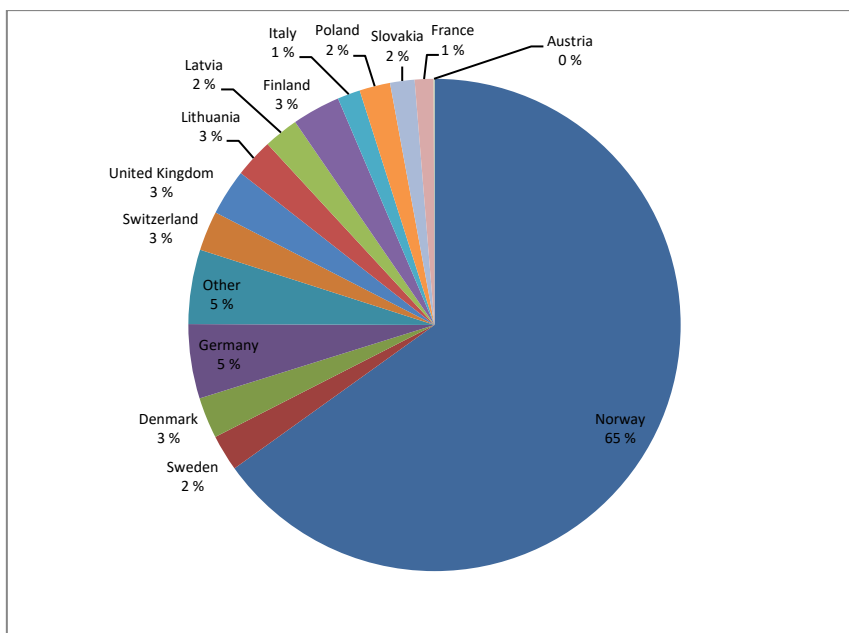
Figure 6: Overall distribution of nationalities during the duration of the pilot for base station 0.

3.1.2 Base station 1

- In this region, there were tourists coming from 37 different countries.
- Maximum presence of Norwegians: ~500
- Maximum presence of tourists: ~200
- 12th place for National visitors (out of the 16 measured)
- 13th place for Foreign visitors (out of the 16 measured)
- %Max Norwegian /Max Foreign: 40% foreign tourists
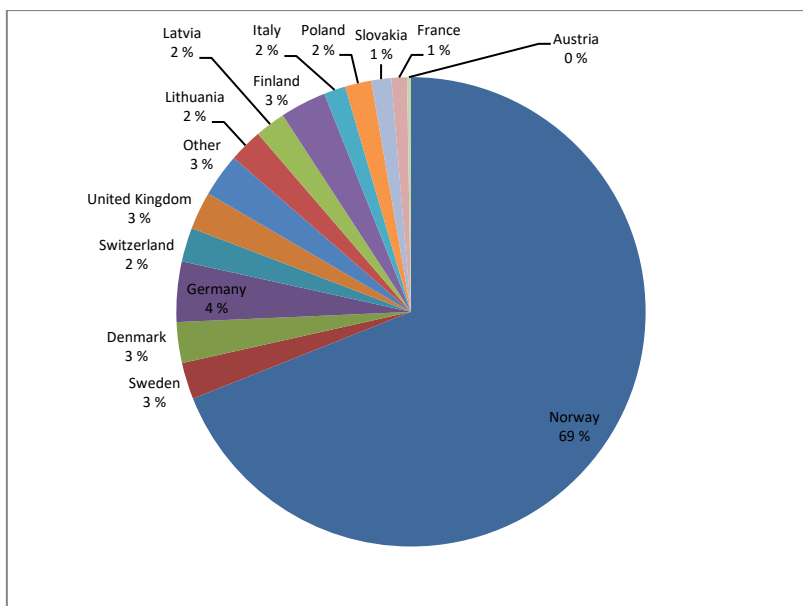

Figure 7: Time-series base station 1.

31.03.2017

Figure 8: Overall distribution of nationalities during the duration of the pilot for base station 1.

3.1.3 Base station 2

- In this region, there were tourists coming from 50 different countries.
- Maximum presence of Norwegians: ~1262
- Maximum presence of tourists: ~482
- 6^h place for National visitors (out of the 16 measured)
- 5th place for Foreign visitors (out of the 16 measured)
- %Max Norwegian /Max Foreign: 38% foreign tourists

Figure 9: Time-series base station 2.

Figure 10: Overall distribution of nationalities during the duration of the pilot for base station 2.

3.1.4 Base station 3

- In this region, there were tourists coming from 68 different countries.
- Maximum presence of Norwegians: ~1692
- Maximum presence of tourists: ~776
- 3rd place for National visitors (out of the 16 measured)
- 2nd place for Foreign visitors (out of the 16 measured)
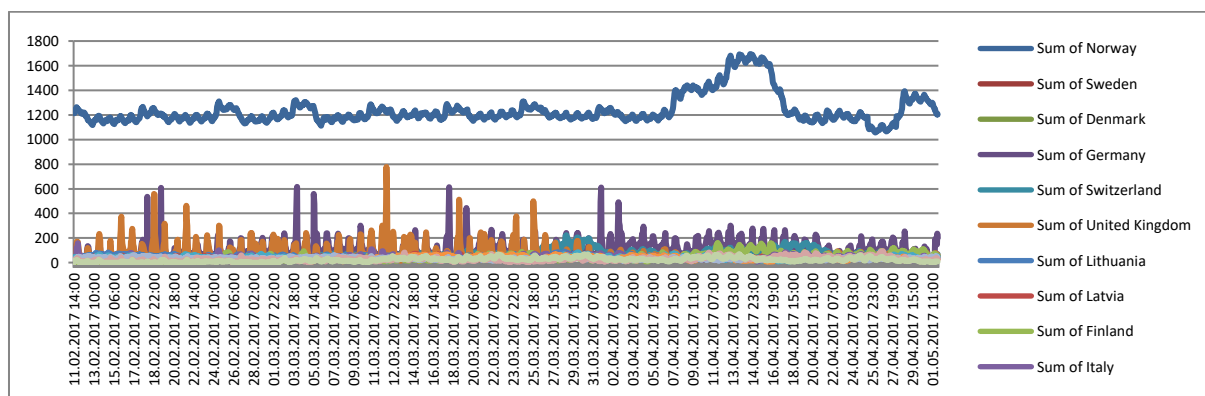- %Max Norwegian /Max Foreign: 48% foreign tourists
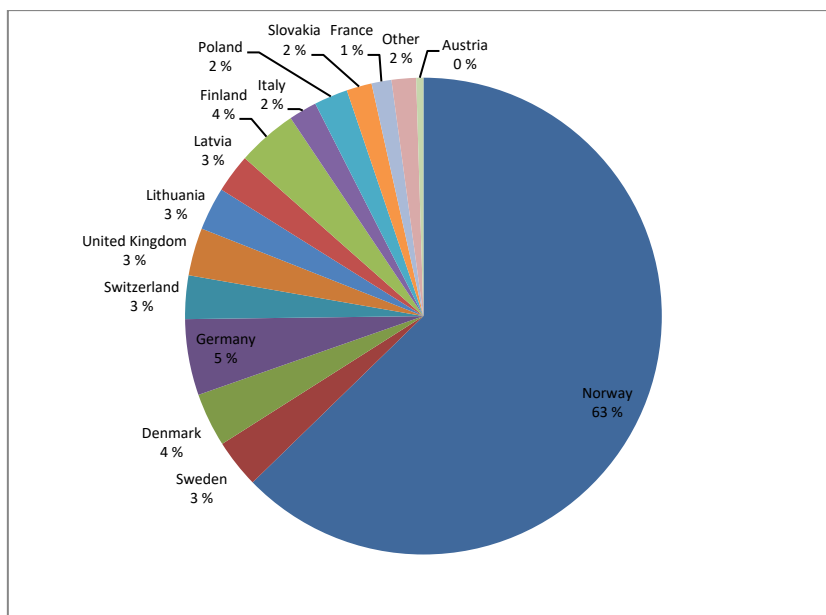

Figure 11: Time-series base station 3.

Figure 12: Overall distribution of nationalities during the duration of the pilot for base station 3.

3.1.5 Base station 4

- In this region, there were tourists coming from 59 different countries.
- Maximum presence of Norwegians: ~1629
- Maximum presence of tourists: ~776
- 4th place for National visitors (out of the 16 measured)
- 3rd place for Foreign visitors (out of the 16 measured)
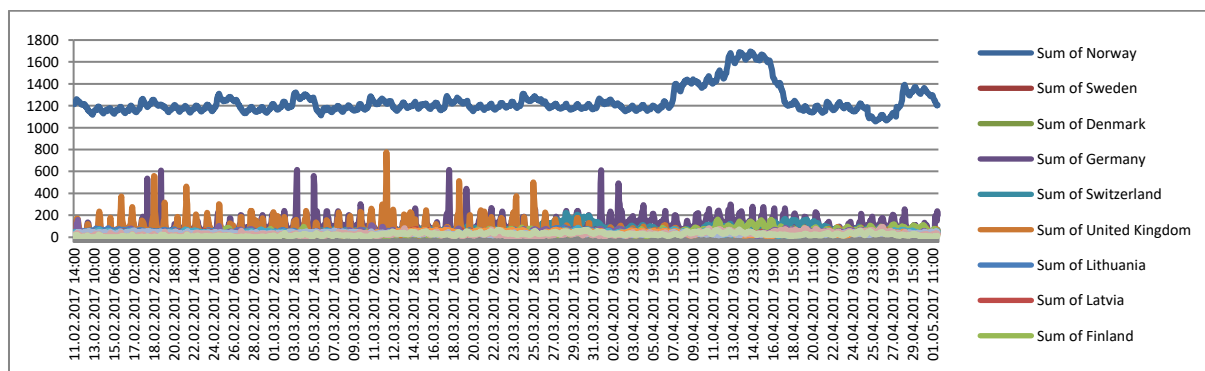- %Max Norwegian /Max Foreign: 48% foreign tourists
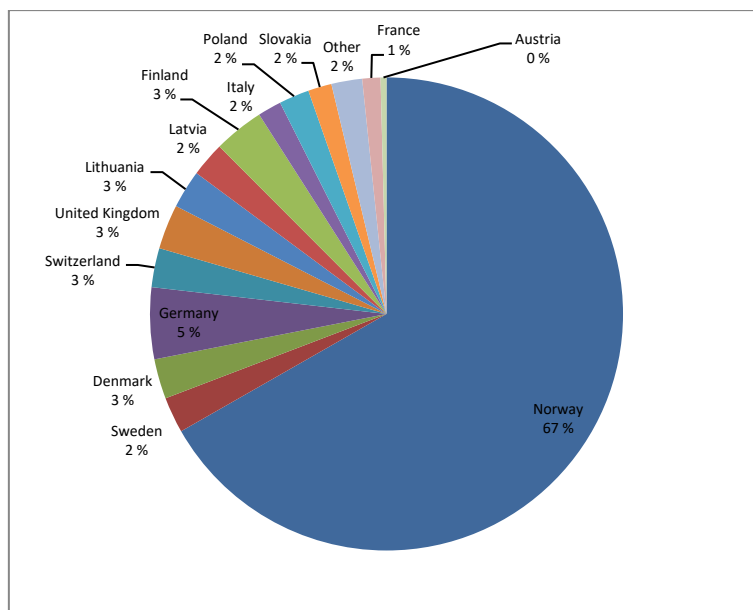

Figure 13: Time-series base station 4.

Figure 14: Overall distribution of nationalities during the duration of the pilot for base station 4.

3.1.6 Base station 5

- In this region, there were tourists coming from 50 different countries.
- Maximum presence of Norwegians: ~959
- Maximum presence of tourists: ~448
- 8th place for National visitors (out of the 16 measured)
- 7th place for Foreign visitors (out of the 16 measured)
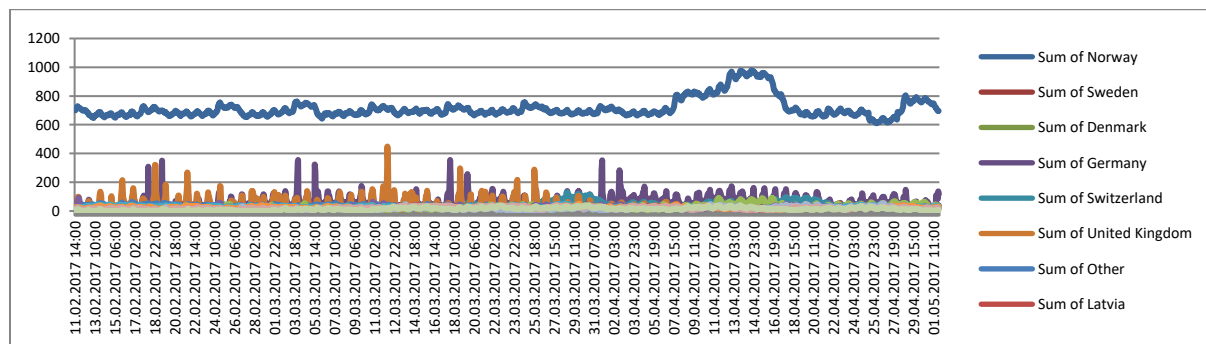- %Max Norwegian /Max Foreign: 47% foreign tourists
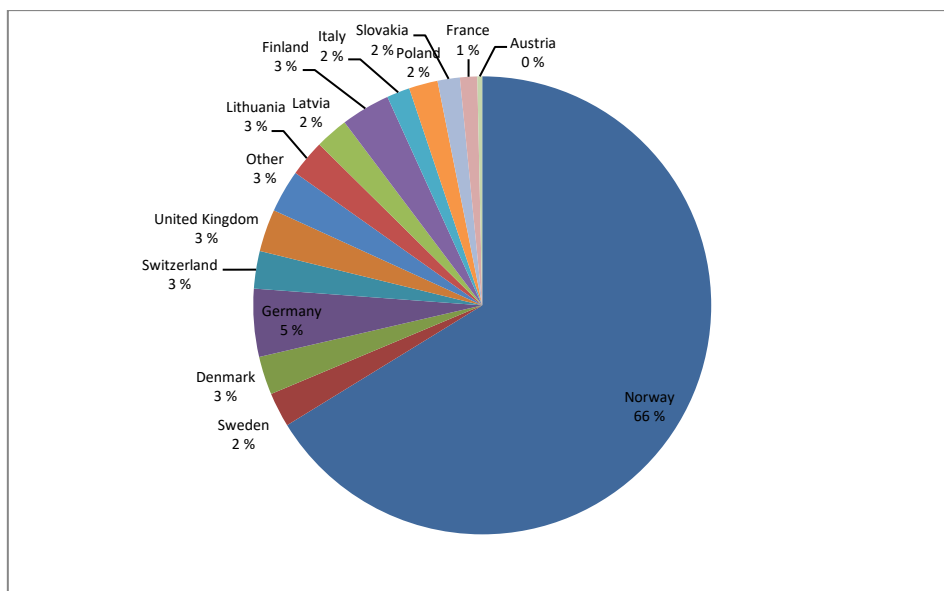

Figure 15: Time-series base station 5.

Figure 16: Overall distribution of nationalities during the duration of the pilot for base station 5.

3.1.7 Base station 6

- In this region, there were tourists coming from 50 different countries.
- Maximum presence of Norwegians: ~1250
- Maximum presence of tourists: ~477
- 7th place for National visitors (out of the 16 measured)
- 6th place for Foreign visitors (out of the 16 measured)
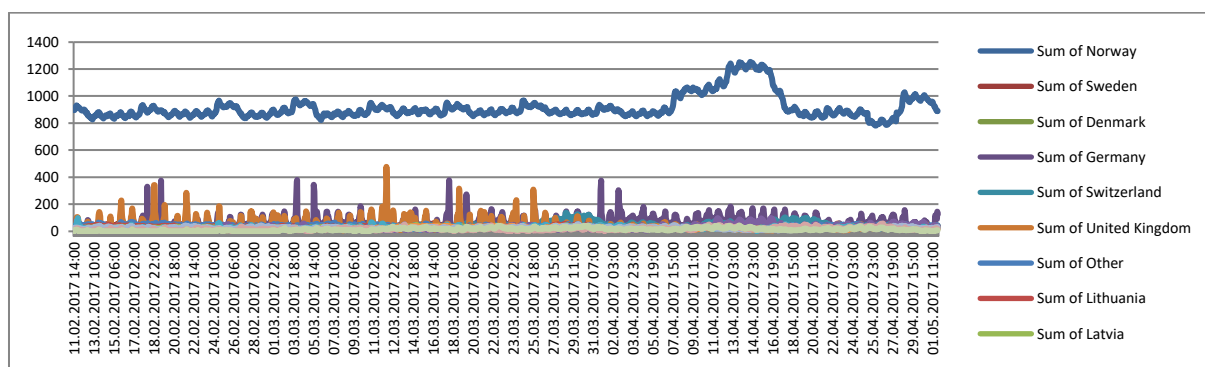- %Max Norwegian /Max Foreign: 38% foreign tourists
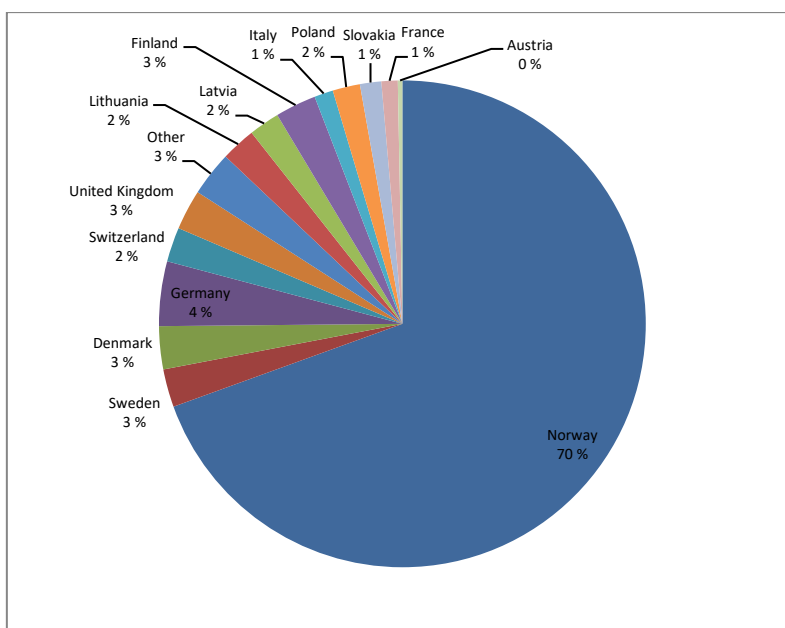

Figure 17: Time-series base station 6.

Figure 18: Overall distribution of nationalities during the duration of the pilot for base station 6.

3.1.8 Base station 7

- In this region, there were tourists coming from 68 different countries.
- Maximum presence of Norwegians: ~1955
- Maximum presence of tourists: ~1176
- 2nd place for National visitors (out of the 16 measured)
- 1st place for Foreign visitors (out of the 16 measured)
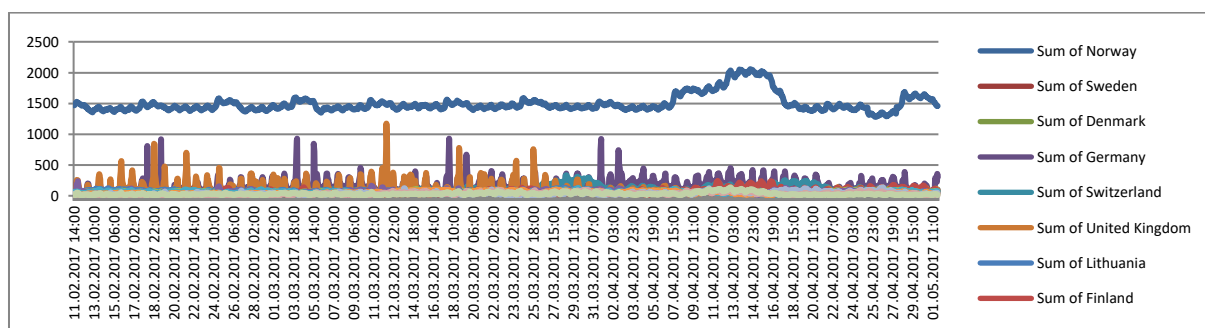- %Max Norwegian /Max Foreign: 60% foreign tourists
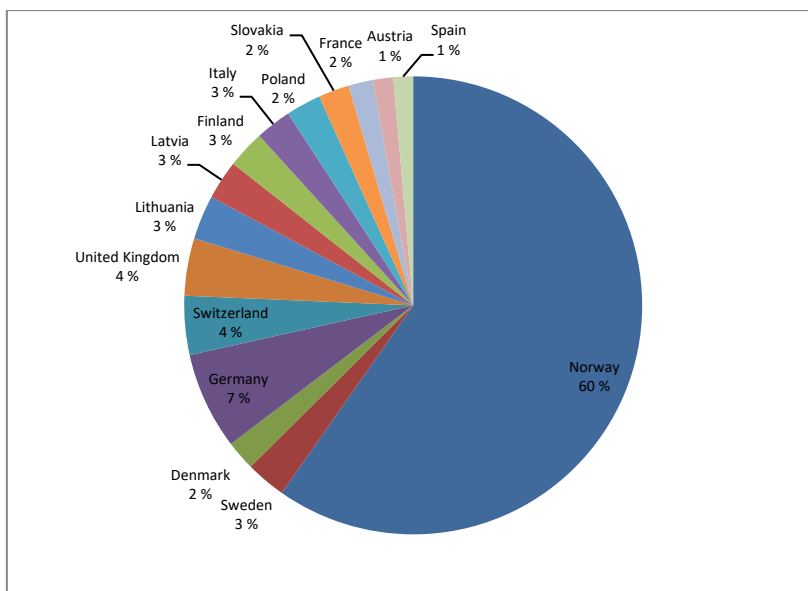

Figure 19: Time-series base station 7.

Figure 20: Overall distribution of nationalities during the duration of the pilot for base station 7.

3.1.9 Base station 8

- In this region, there were tourists coming from 43 different countries.
- Maximum presence of Norwegians: ~523
- Maximum presence of tourists: ~300
- 11th place for National visitors (out of the 16 measured)
- 11th place for Foreign visitors (out of the 16 measured)
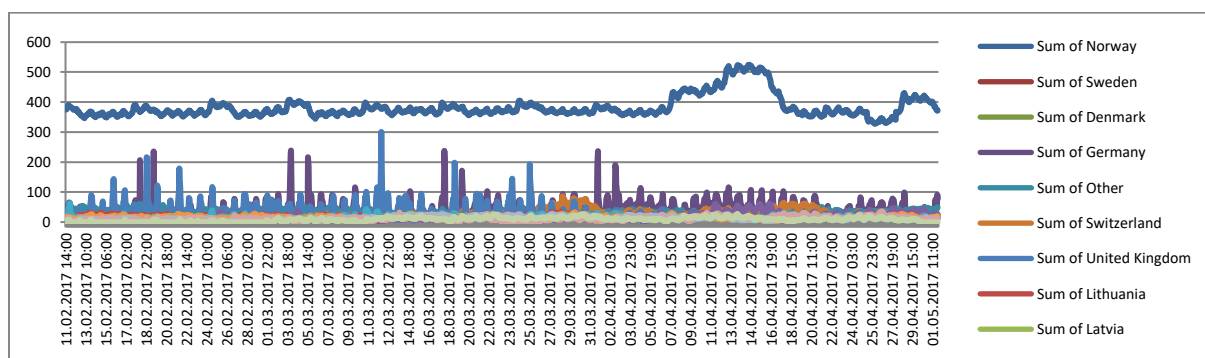- %Max Norwegian /Max Foreign: 57% foreign tourists
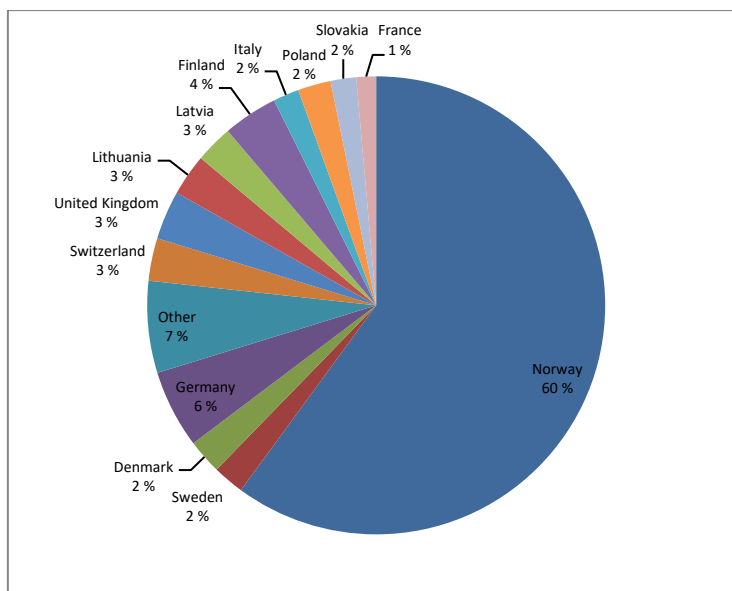

Figure 21: Time-series base station 8.

Figure 22: Overall distribution of nationalities during the duration of the pilot for base station 8.

3.1.10 Base station 9

- In this region, there were tourists coming from 47 different countries.
- Maximum presence of Norwegians: ~701
- Maximum presence of tourists: ~402
- 10th place for National visitors (out of the 16 measured)
- 8th place for Foreign visitors (out of the 16 measured)
- %Max Norwegian /Max Foreign: 57% foreign tourists

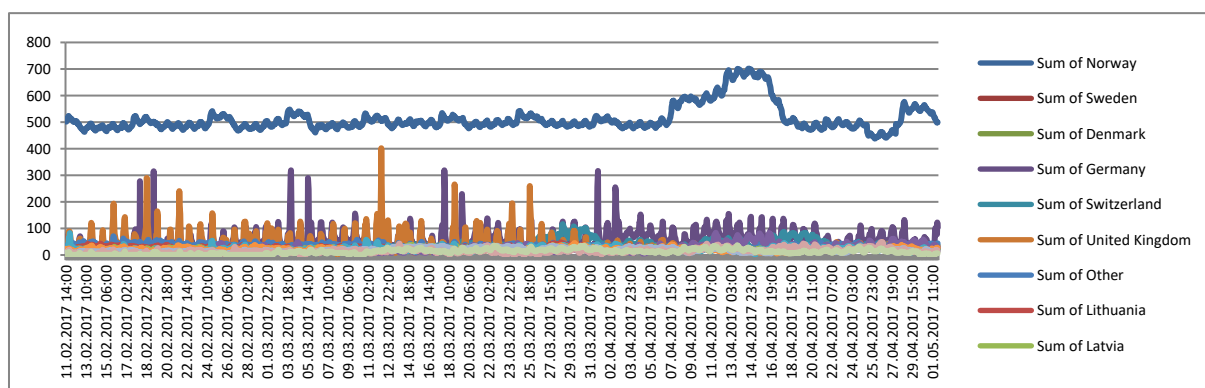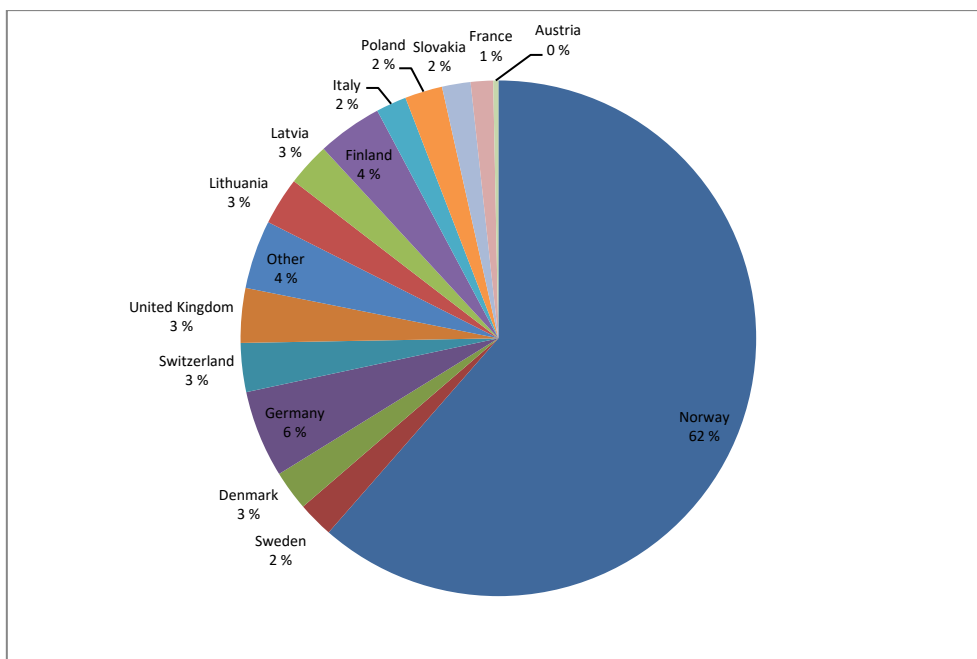Figure 23: Time-series base station 9.

Figure 24: Overall distribution of nationalities during the duration of the pilot for base station 9.

3.1.11 Base station 10

- In this region, there were tourists coming from 35 different countries.
- Maximum presence of Norwegians: ~295
- Maximum presence of tourists: ~177
- 14th place for National visitors (out of the 16 measured)
- 15th place for Foreign visitors (out of the 16 measured)
- %Max Norwegian /Max Foreign: 60% foreign tourists
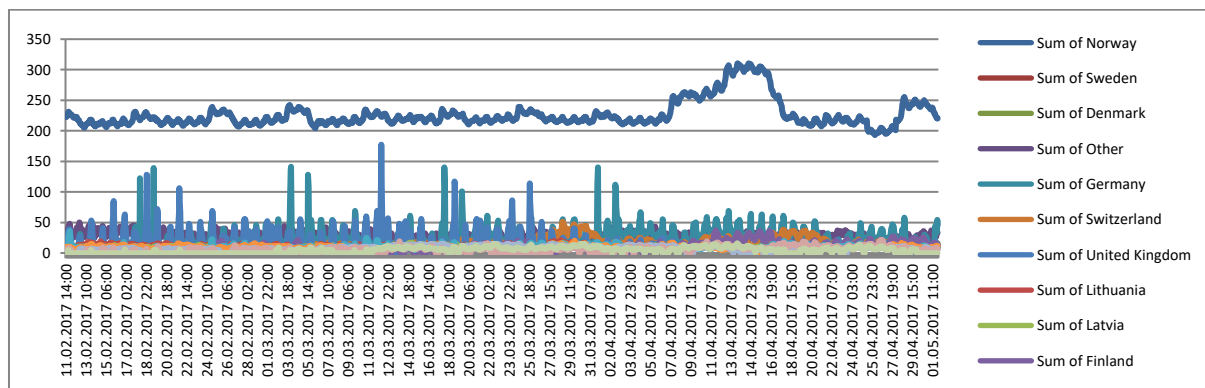

Figure 25: Time-series base station 10.

Figure 26: Overall distribution of nationalities during the duration of the pilot for base station 10.

3.1.12 Base station 11

- In this region, there were tourists coming from 40 different countries.
- Maximum presence of Norwegians: ~220
- Maximum presence of tourists: ~260
- 15th place for National visitors (out of the 16 measured)
- 12th place for Foreign visitors (out of the 16 measured)
- %Max Norwegian /Max Foreign: 118% foreign tourists

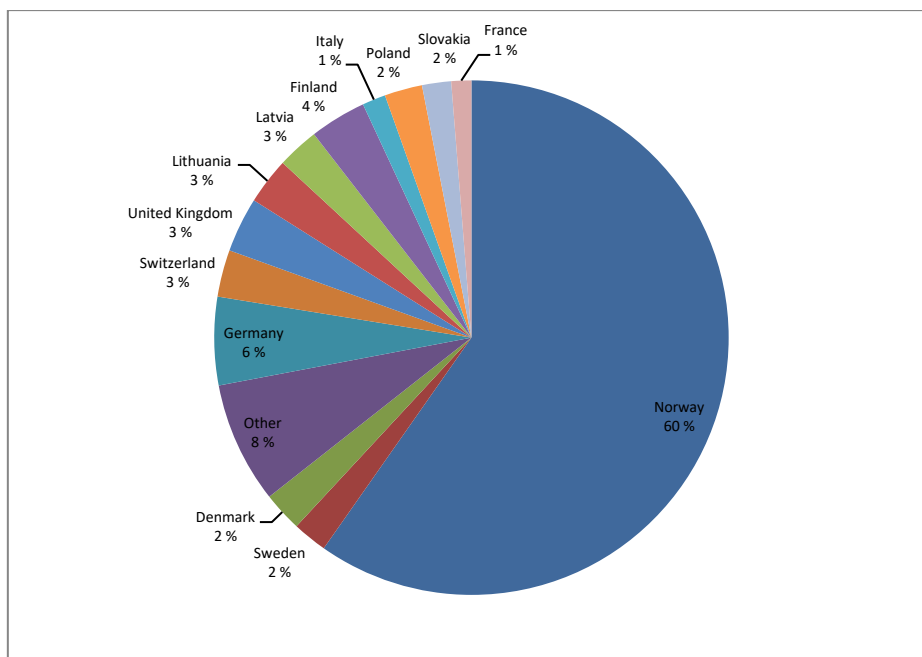Figure 27: Time-series base station 11.

31.03.2017

Figure 28: Overall distribution of nationalities during the duration of the pilot for base station 11.

3.1.13 Base station 12

- In this region, there were tourists coming from 44 different countries.
- Maximum presence of Norwegians: ~1979
- Maximum presence of tourists: ~309
- 1st place for National visitors (out of the 16 measured)
- 10th place for Foreign visitors (out of the 16 measured)
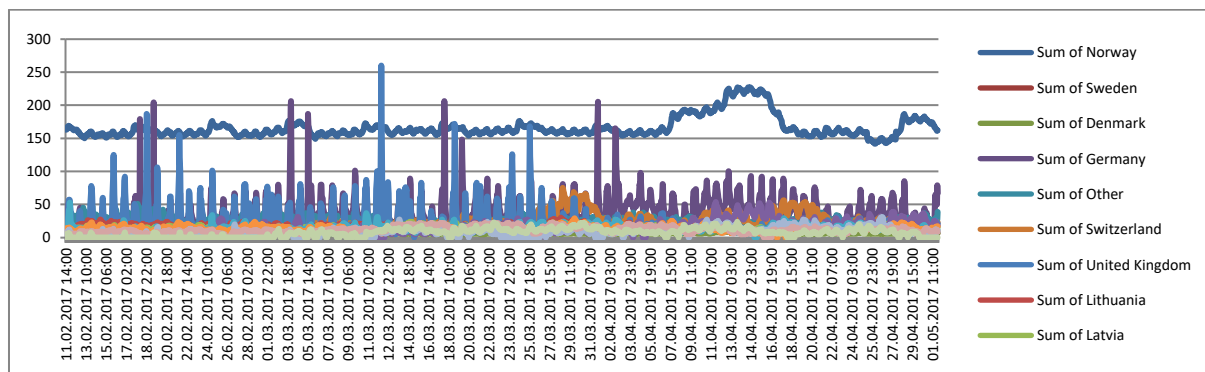- %Max Norwegian /Max Foreign: 15% foreign tourists
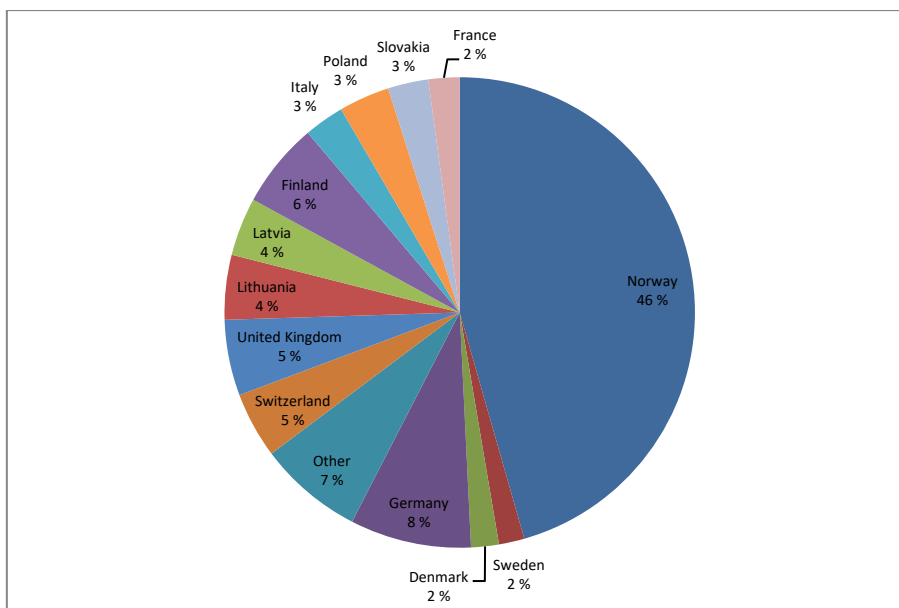

Figure 29: Time-series base station 12.

Figure 30: Overall distribution of nationalities during the duration of the pilot for base station 12.

3.1.14 Base station 13

- In this region, there were tourists coming from 15 different countries.
- Maximum presence of Norwegians: ~195
- Maximum presence of tourists: ~26
- 16th place for National visitors (out of the 16 measured)
- 16th place for Foreign visitors (out of the 16 measured)
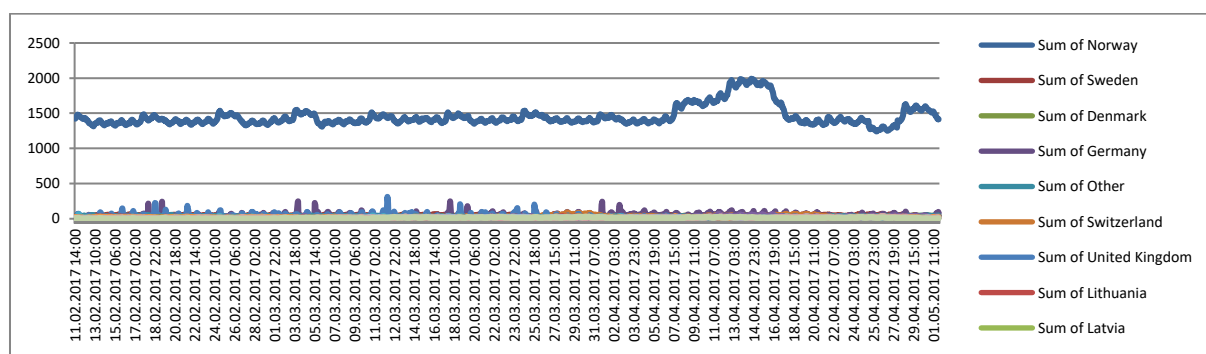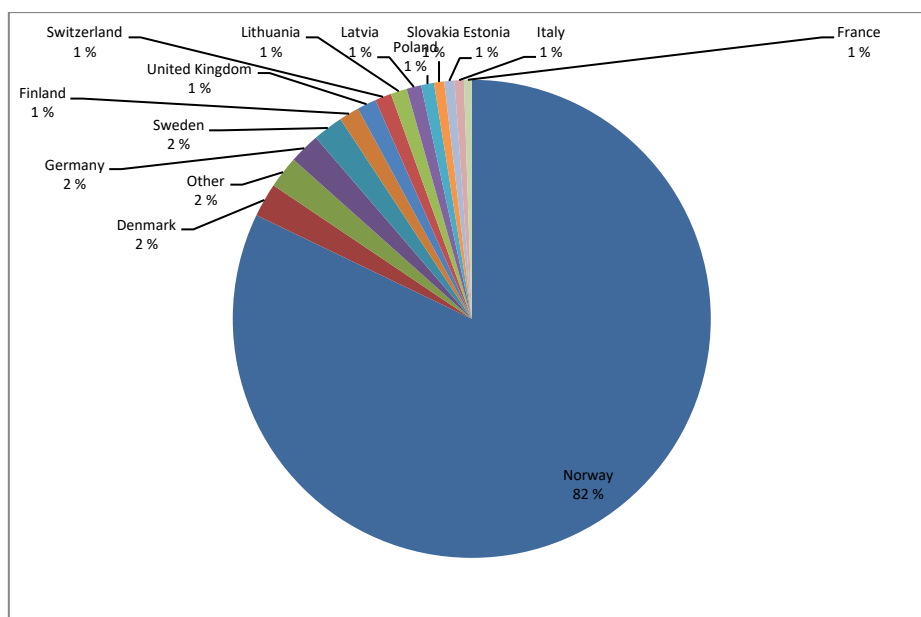- %Max Norwegian /Max Foreign: 13% foreign tourists

Figure 31: Time-series base station 13.

Figure 32: Overall distribution of nationalities during the duration of the pilot for base station 13.

3.1.15 Base station 14

- In this region, there were tourists coming from 55 different countries.
- Maximum presence of Norwegians: ~1303
- Maximum presence of tourists: ~501
- 5th place for National visitors (out of the 16 measured)
- 4th place for Foreign visitors (out of the 16 measured)
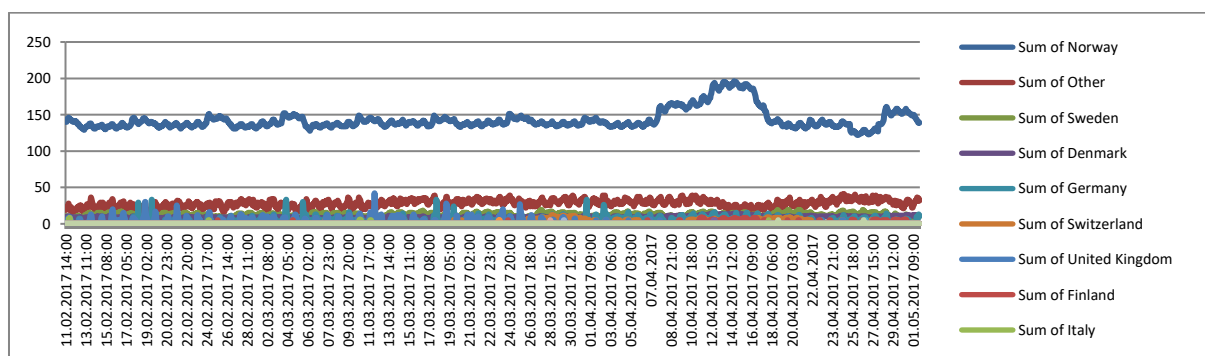- %Max Norwegian /Max Foreign: 38% foreign tourists

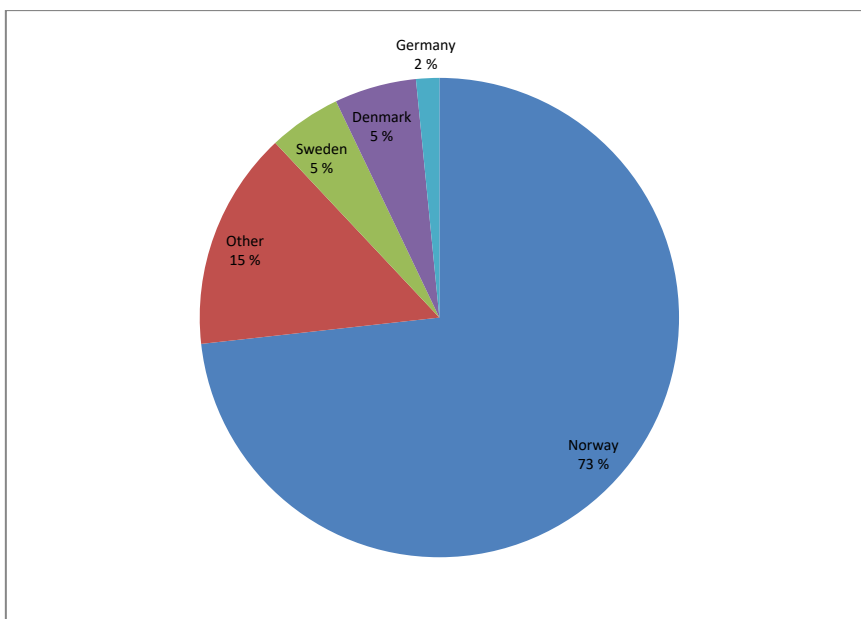Figure 33: Time-series base station 14.

Figure 34: Overall distribution of nationalities during the duration of the pilot for base station 14.

3.1.16 Base station 15

- In this region, there were tourists coming from 29 different countries.
- Maximum presence of Norwegians: ~427
- Maximum presence of tourists: ~120
- 13th place for National visitors (out of the 16 measured)
- 15th place for Foreign visitors (out of the 16 measured)
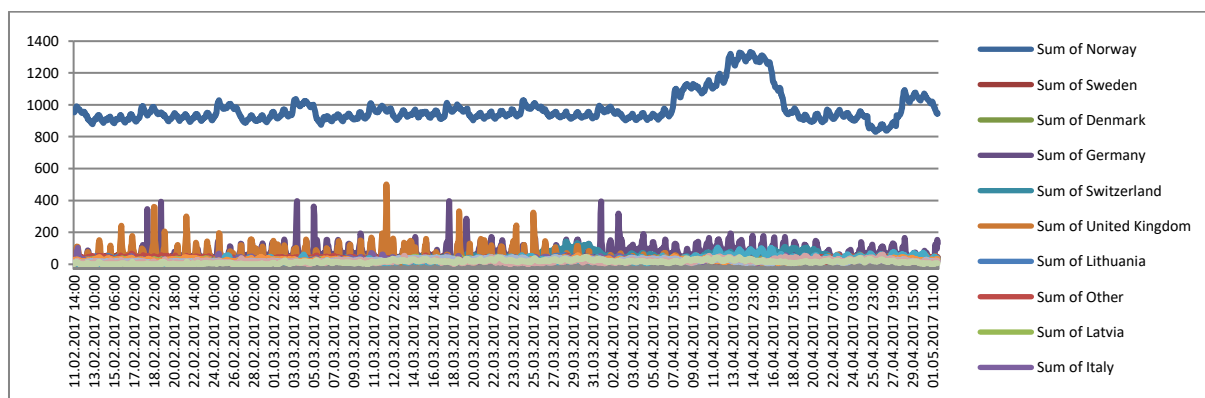- %Max Norwegian /Max Foreign: 28% foreign tourists
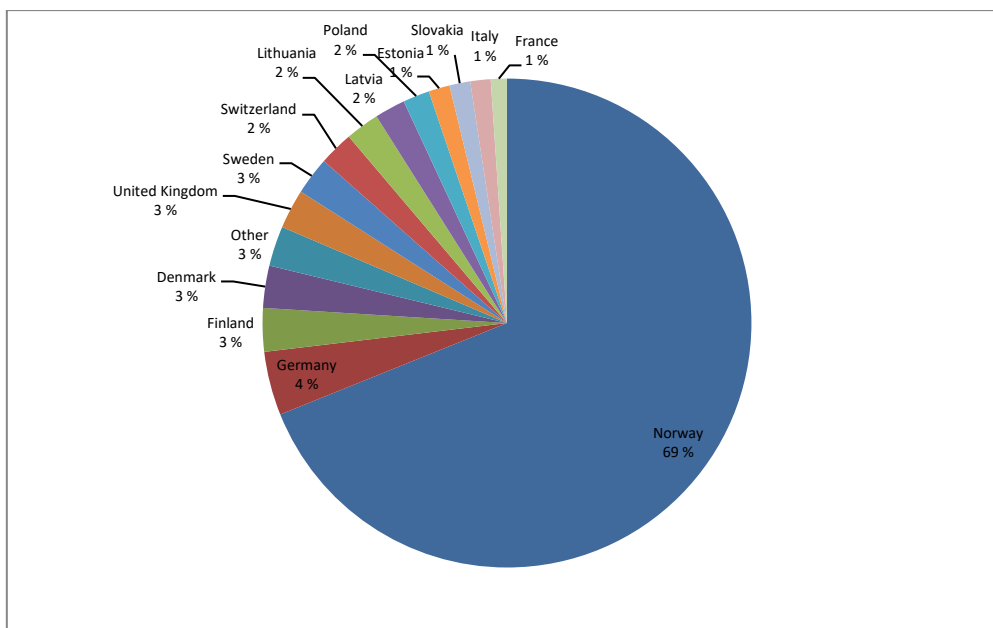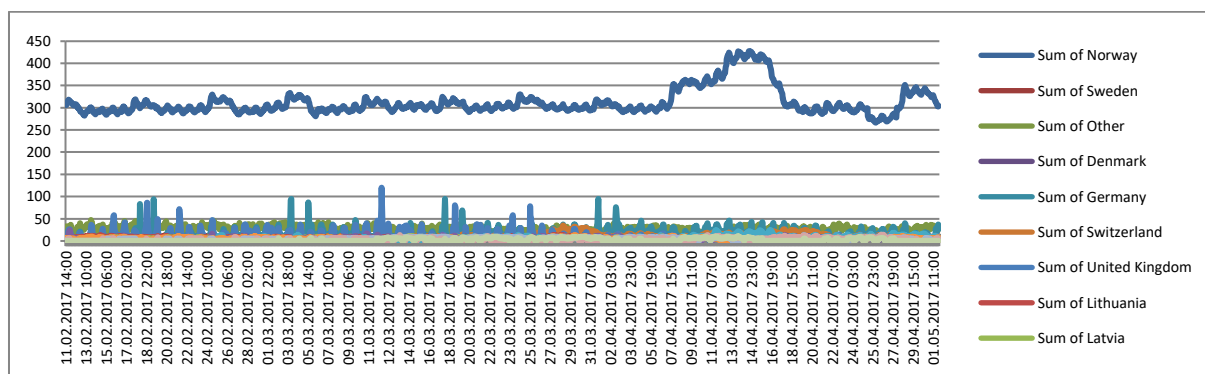

Figure 35: Time-series base station 15.
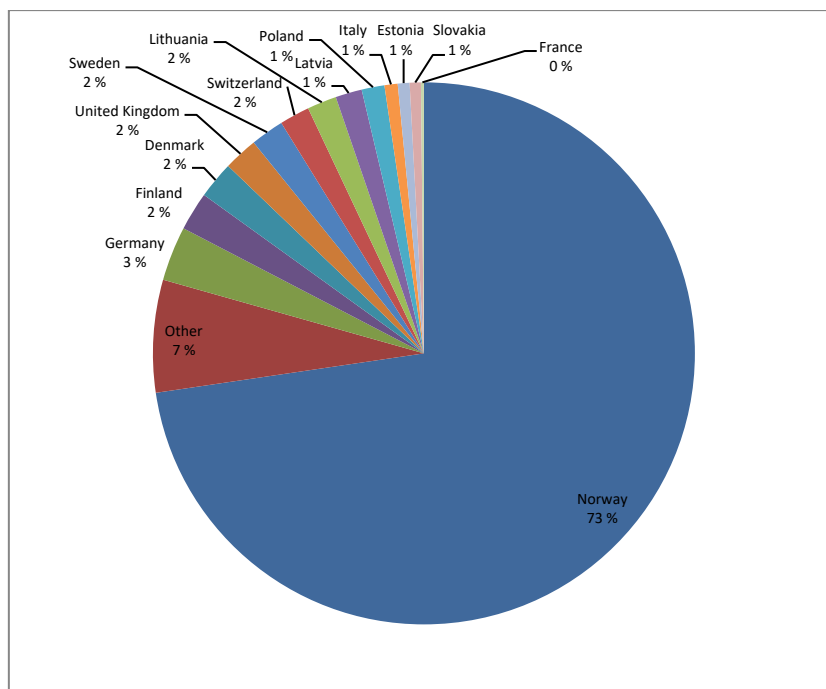
Figure 36: Overall distribution of nationalities during the duration of the pilot for base station 15.

3.2 Lyngen's coverage as a tourists attraction

Figure 37 and 38 represent the coverage of countries the Lyngen Alps are attracting.

Figure 37: Coverage of countries of origin of tourists in the Lyngen Alps region from February to May 2017. All the "rare" cases where the k-anonymity level was not met are filtered out.

Figure 38: Heat map of countries of origin of tourists in the Lyngen Alps region from February to May 2017. All the "rare" cases where the k-anonymity level was not met are filtered out.

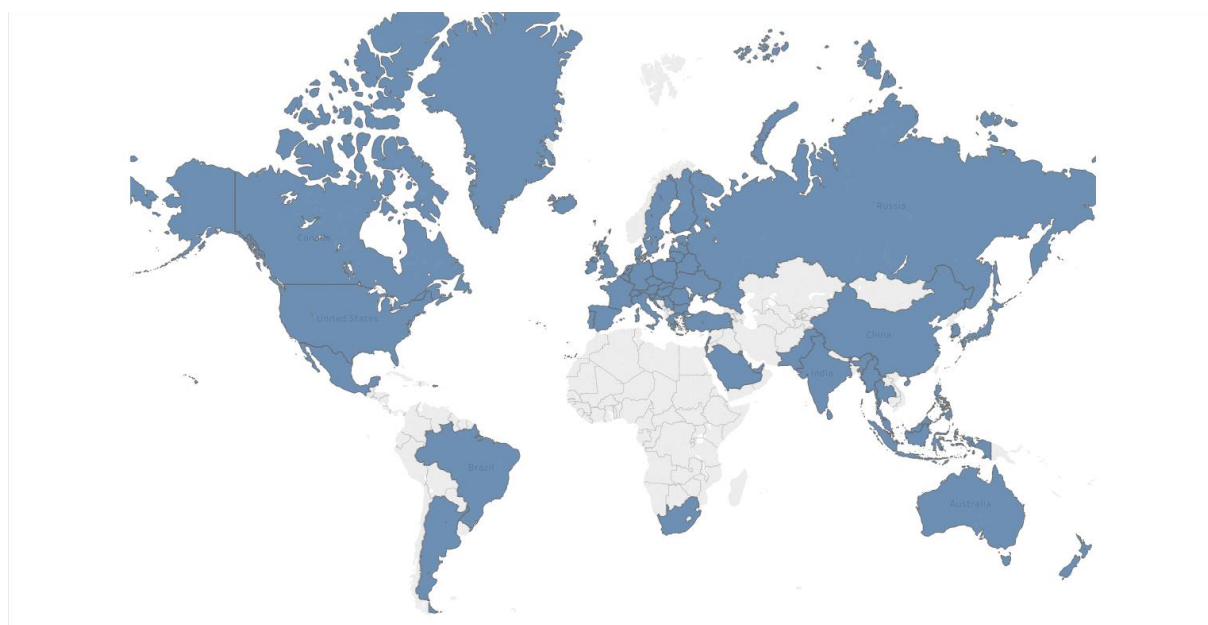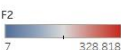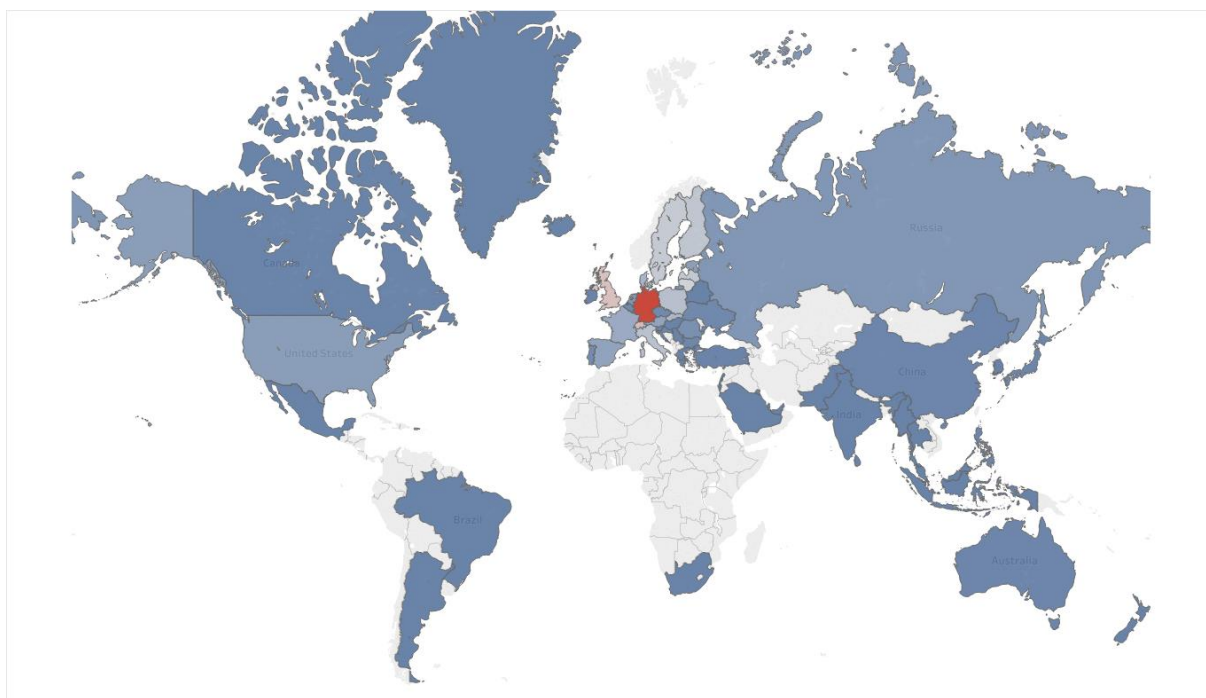## 3.3 Extrapolation methodology

In this region we overview the method used to extrapolate Telenor's customer base (handsets interacting with Telenor's cellular network) to an estimate on the total population present in regions covered by our base stations.

Figure 37 is a representation of the multiple inputs used for the extrapolation algorithms:

- **Regional market share.** Telenor continuously updates its surveys to determine its approximate market share, per Fylke, gender and age band. Figure 37 below exemplifies the kind of information we use as one of the extrapolation's inputs.

			Østfold	Akershus	Oslo	Hedmark	Oppland	Buskerud	Vestfold
Gender	Men	Yes Freq	329	621	832	252	216	330	297
		Yes % Valid	51,70 %	50,80 %	51,10 %	52,50 %	49,60 %	52,30 %	47,50 %
	Women	Yes Freq	308	601	797	228	219	301	328
		Yes % Valid	48,30 %	49,20 %	48,90 %	47,50 %	50,40 %	47,70 %	52,50 %
Stdalder	15-29	Yes Freq	103	168	456	88	81	129	116
		Yes % Valid	17,20 %	14,60 %	29,30 %	19,60 %	19,50 %	21,60 %	19,30 %
	30-44	Yes Freq	179	343	555	114	107	173	137
		Yes % Valid	29,90 %	29,80 %	35,70 %	25,40 %	25,80 %	29,00 %	22,80 %
	45-59	Yes Freq	183	379	328	134	129	147	190
		Yes % Valid	30,60 %	32,90 %	21,10 %	29,90 %	31,10 %	24,60 %	31,80 %
	60+	Yes Freq	133	262	218	113	98	149	156
		Yes % Valid	22,30 %	22,70 %	14,00 %	25,10 %	23,60 %	24,90 %	26,00 %
What is your main provider brand for voice and sms? Fi	Telenor	Yes Freq	251	497	570	215	184	251	215
		Yes % Valid	39,40 %	40,60 %	35,00 %	44,80 %	42,30 %	39,70 %	34,40 %

Figure 39: Example of regional market share tables. Information is broken down per Gender and Age Band.

- **SSB Statistics.** When extrapolating we are trying to recover not only 100% of all handsets in Norway, therefore we also use SSB statistics. By using statistics such as the ones present in Figure 38, we aim to recover the total of Norwegian population as output from our Extrapolation Algorithm. As we will present in Figure 41, after using our extrapolation algorithms we are able to recover the whole Norwegian population with good accuracy.

Population and area. Urban settlements.1 January

	Population	Change	Change in per cent
	2016	2015 - 2016	2015 - 2016
Residents in urban settlements	4 229 849	57 045	1.4
Residents in rural settlements	968 576	-5 236	-0.5
Area of urban settlement (km²)	2 172.50	14.22	0.66
Number of residents pr km2 in urban areas	1 947	14	0.7
Proportion residents in urban settlements, per cent	81.13	0.18	0.22
Residents in the five largest urban settlements			
Oslo	975 744	17 366	1.8
Bergen	252 772	2 352	0.9
Stavanger/Sandnes	213 313	2 439	1.2
Trondheim	177 617	2 549	1.5
Drammen	115 137	1 603	1.4

Figure 40: SSB statistics available at: http://www.ssb.no/en/befolkning/statistikker/beftett .

- **International roaming agreements.** When estimating the total number of foreign tourists roaming in a region, we use as an additional input a set of roaming agreements we have at Telenor Group global level. Figure 39 presents a few rows of the agreements we have with international operators. An example of how these are used: Roamers coming from Sweden have to be extrapolated by using the roaming agreements Telenor has in Sweden plus Telenor's market share in Sweden.

Operatør kode	Land	Operatør
		Live operatør oversigt inbound for TnM - per Nov 2016
41201	AFGANISTAN	AWCC
41220	AFGANISTAN	Telecom Development Company Afganistan (TDCA)
27601	ALBANIA	Telekom Albania (Albanian Mobile,(AMC)
27602	ALBANIA	Vodafone Albania
60301	ALGERIA	ATM Mobilis
60302	ALGERIA	Optimum Telecom Algerie S.p.A (Djezzy) Orascom Telecom Algerie (Djezzy)
60303	ALGERIA	Wataniya Telecom (Ooredoo) (Algérie)
21303	ANDORRA	Andorra Telecom (Servei de Telecomm. d'Andorra (STA))
63102	ANGOLA	Unitel
63104	ANGOLA	Movicel
338050	ANGUILLA	Digicel Caribbean
365840	ANGUILLA	LIME (Cable&Wireless)
34403	ANTIGUA & BARBUDA	APUA PCS Ltd.
338050	ANTIGUA & BARBUDA	Digicel Caribbean
344920	ANTIGUA & BARBUDA	LIME (Cable&Wireless)
72207	ARGENTINA	Telefónica Móviles Argentina SA (TCP-Unifon)
72231	ARGENTINA	Claro (AMX Argentina S.A. - CTI Móvil)

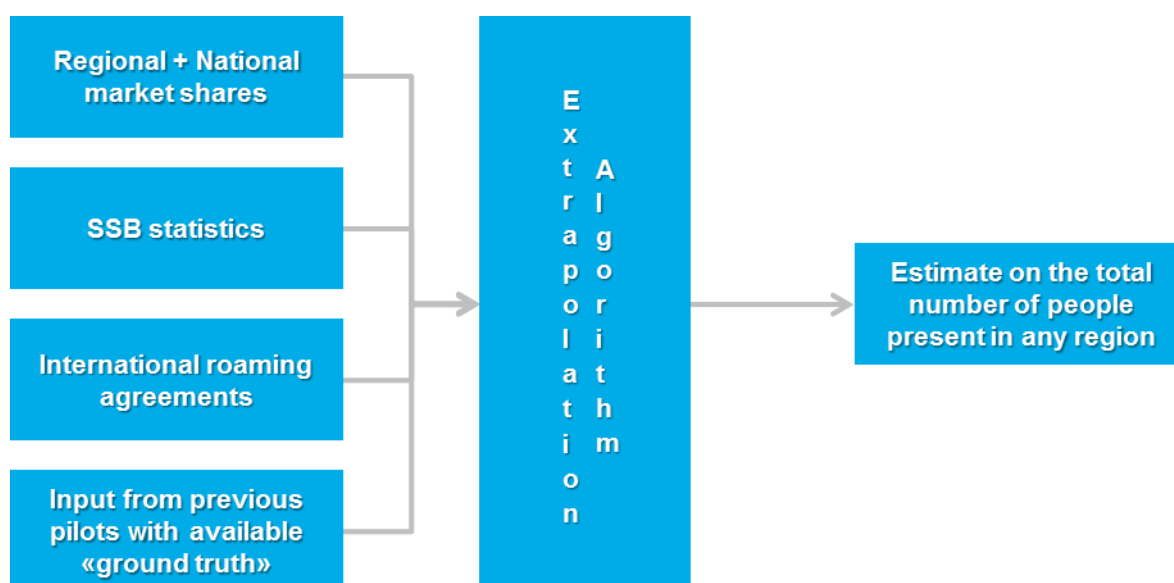Figure 41: Example input coming from international roaming agreements.

Figure 42: Extrapolation mechanism. On the left panel we present the inputs used for the estimation of an extrapolating factor. The extrapolating factor is used to go from Telenor's customer base to an estimate on the total population present in any region measured.
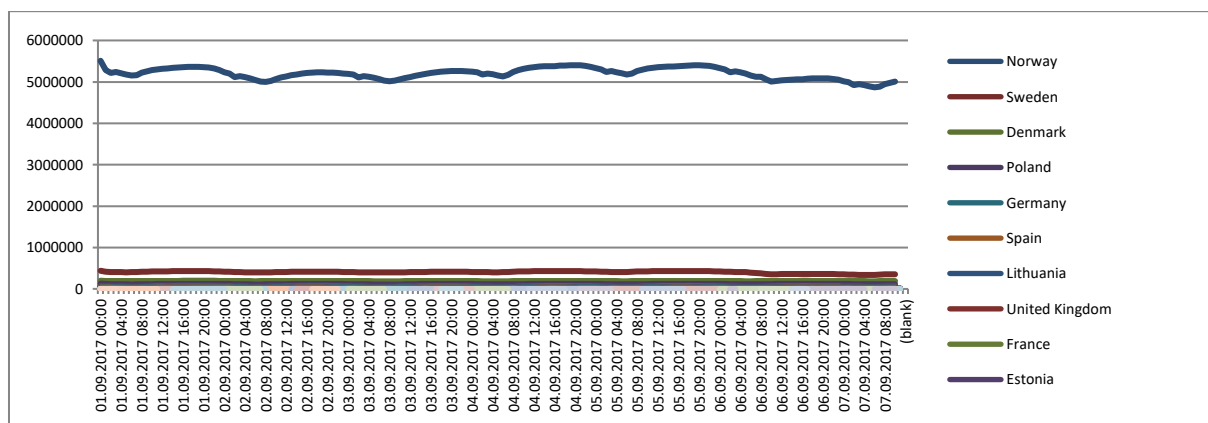
Figure 43: Time-series base station 1.

Figure 41 above illustrates the output of our extrapolation algorithm when estimating the total number of people in Norway. Our values for national census oscillate (due to handsets turning on and off) around SSB's census data within a 3% range.

4. Conclusions and recommendations

The results of the pilot seem to indicate to support several key assumptions.

It is in fact possible to use mobile network signaling data to estimate total number of tourists present in an area and their countries of origin, this based on the country that the mobile subscription is registered in.

The pilot would also indicate that the operator's market share is a good starting point for extrapolating the number of registered subscription to the total number of visitors for national parks.

For a potential follow up pilot, it would be advised to start measurements a couple of weeks before and after the period of interest. This, to measure the increase and decrease before and after, for example, high seasonal effects.

To further validate the data on actual tourist counts, some ground of truth data (e.g. from surveys) for the distribution of foreign tourists per region in Lyngen.

4.1 Using the results in the line of business

This type of information could be potentially operationalized in several areas/ways going forward:

- Conducting partner assessment (e.g. contribution of partners benefiting from the presence of tourists in the region)
- Facilitating better service to foreign travelers (e.g. in-app, signage, customer service and frontline language support etc.)
- 1:1 marketing towards foreign audiences when they arrive to
- Introducing new KPIs to measure improvements in attracting foreign visitors to the Lyngen Alps region
- Build better business cases to fund overseas campaigns with Tourism industry players in Norway and allow for measurable follow-up
- More detailed insight into two key driving forces of megatrends that could impact the Lyngen Alps at an ecological level.