# Caribbean Netherlands: a different method of making regional tourism statistics

Shirley Ortega Azurduy[1]

Nico Heerschap

## 1.    Introduction

The Kingdom of the Netherlands does not only consist of an European part. It also consists of several islands in the Caribbean area. The Dutch Antilles were dissolved in October 2010. Two of these islands (Curaçao and Sint Maarten) gained a special status as independent countries whereas three of them (Bonaire, St Eustatius, and Saba) became specific public bodies (special municipalities) of the kingdom. This means that since 2010 Statistics Netherlands (SN) is responsible for the statistics of these three islands, including those on tourism.

To produce key tourism statistical indicators for the three islands, in the first instance, the idea was to use existing arrival- and departure cards. However, in the new situation, the Customs and Border Protection were not willing to cooperate anymore. They did not see the collection of the cards as their core duty. Also airlines were not very willing to cooperate on the distribution of the cards on a voluntary basis. A second idea was to employ the same method as used in the Netherlands, that is to produce the number of tourists and nights spent based on questionnaires sent to tourism accommodations. However, here the main problem was that the Chamber of Commerce could not provide a good and up-to-date list of tourism accommodations. Also, a big part of accommodation renting takes place outside the normal group of hotels and resorts, that is holiday homes, often rented in the grey circuit.

SN experimented with the use of internet as a source of accommodation statistics for Caribbean Netherlands (CN). The idea was to gather accommodation units and their characteristics through the use of dedicated web crawlers [Heerschap et al, 2014]. In this way, a first inventory of the supply side of the tourism industry for Bonaire, St Eustatius, and Saba was achieved. The database contains types of accommodations, accommodation capacities, and contact details.

Some conclusions of the pilot were that 1) technically, it is not difficult to build, test and use dedicated robots; 2) the challenge of replicating this pilot (and to turn it into a production process) is actually the deduplication procedure across the web crawlers databases[2]; 3) approaching tourism accommodations and seeking cooperation with especially Rental Agencies was difficult as well [Heerschap et al, 2014]. Altogether, it meant that a new method had to be found to make tourism statistics for CN.

This paper describes a different method of making regional tourism statistics. The key element is that the method is totally based on existing registers. The paper is organized as follows. In section 2 the data sources and developed methodology are presented. In section 3 we show the produced

---

[1] Corresponding author: Statistics Netherlands, Department of Arts, Tourism and Technology, P.O. Box 24500, 2490 HA The Hague, The Netherlands (e-mail: S.OrtegaAzurduy@cbs.nl) .

[2] During deduplication, it is controlled whether accommodations are in business. Data are also compared and completed with auxiliary information, e.g., if accommodations are connected to rental agencies.

indicators which include the number of inbound day-trippers and tourists' length of their stay. Also indicators on age and gender of the visitors can be extracted, as well as their nationality as a proxy for their country of residence. In section 4 some conclusions on the advantages and disadvantages of this approach are drawn. Additionally, the design and setup of an exit survey is shortly described. Ideas on the use of mobile telephone data are also discussed along with the redesigning of the previous web scraping approach. Finally, some remarks are made on how the local tourism policy makers perceive this change in producing tourism statistics.

## 2. Methodological Approach of Using Registers Data

Since 2012, SN produces and publishes an overview of a number of important indicators of tourism-industry in the CN. Firstly, the estimates of the number of visitors arriving by air are divided according to their nationality and the number of overnight stays. Notice that a visitor who visits an island twice counts for two visitors. Secondly, the estimates on the number of passengers that arrived by sea. All these figures are entirely based on available registers only. It means that the estimated statistics depend on the quality of the registers and that SN must carry out some editing and comprehensive plausibility checks to these registers, among others, by matching registers to each other. Once data of registers are available, the processing time is about two months (excluding reporting and publication in Dutch, English, and Papiamento).

### 2.1 Data sources

The following registers are used in the statistical process:

1. The Border Management System (BMS). This register contains data from scanned passports of all passengers entering and leaving the islands of the CN. It might sometimes not be complete, i.e., some data are not included in the system for some incoming and outgoing passengers because passports have not been scanned successfully.
2. The Tax and Customs Administration (Customs). This register contains data on the number of flight movements and the total number of passengers per aircraft per island. It concerns aggregated data on residents and non-residents, and contains also information about shipping traffic including cruises, ferries, yachts, and other vessels.
3. The General Declaration System (Gendecs). This register is similar to the Customs register and has the same source. However, it only contains data about air traffic.
4. The population register of the Caribbean Netherlands (PIVA). This register is used in the processing system to determine whether an airplane passenger is a resident of CN or not. (Brought into use since 2014.)
5. The population register of the Netherlands (GBA). This register is also used in the processing system to determine whether an airplane passenger is a resident of the (European part of the) Netherlands or, in other words, to distinguish between Dutch passport holders from Europe and those from Aruba, Curaçao, and Sint Maarten. (Brought into use since 2015.)
6. The Bonaire International Airport (BIA) and datasets of the Harbour Master of Bonaire (HMB)[3] are registers that are mainly used for plausibility checks.

### 2.2 Inbound tourism by air: Processing

Firstly, the total number of incoming (and departing) air passengers per month are determined based on the Gendecs- and Customs data. For Bonaire, these data are checked and possibly improved by comparing the output with the information from Flamingo Airport on the number of incoming and departing passengers per airline. Secondly, the so-called BMS (inbound passengers) is used to determine which part of the passengers is visitor and which part is resident. By linking these data to the PIVA and the GBA, it can be determined whether a passenger is a resident or a non-resident (visitor). The ratio between the number of residents and number of non-residents of the air passengers is calculated (*tourism-ratio*). Depending on the number of arriving and departing passengers, the calculation results in a monthly ratio (for Bonaire) or a yearly ratio (for Saba and St Eustatius). The total number of incoming visitors can then be determined by multiplying these ratios by the total number of air passengers. Sometimes and for several reasons, not all the passports of the inbound passengers arriving in the CN can be scanned. Therefore, figures obtained based on the BMS-system are estimates. It means that the error margin of these estimates increases as the

---

[3] Notice that the BIA and HMB datasets are only available for Bonaire.

coverage of the number of the scanned passports of the arriving passenger decreases. BMS data are seen as if they were a random data sample.

Thirdly, based on the proportions of the arriving passengers in het BMS, the visitors are subdivided according to "nationality"[4], which is used as a proxy for country of residence. Finally, by linking the data of a visitor from the moment of entry to the details of his / her departure, an estimate can also be made of the number of overnight stays, including day-trippers. For practical reasons, the linking is limited to those matches where the difference between arrival and departure does not exceed two months. This differs from the international definition, involving tourism, i.e., *"Any person who travels to a country other than that in which s/he has his/her usual residence but outside his/her usual environment for a period not exceeding 12 months,…"*[UNWTO, 2016]. The difference between total number of airplane visitors and the number of overnight stays allows to estimate day trippers. The day trippers do not count when calculating the average number of overnight stays. Next, we derive the distribution of the visitors ages and gender. The target gender derivation is not shown in this article, however, it can be straightforwardly determined based on the BMS-data. An overview of the processing steps required to estimate inbound tourism statistics is provided in Scheme 1.

Scheme 1. Processing steps required by inbound tourism statistics
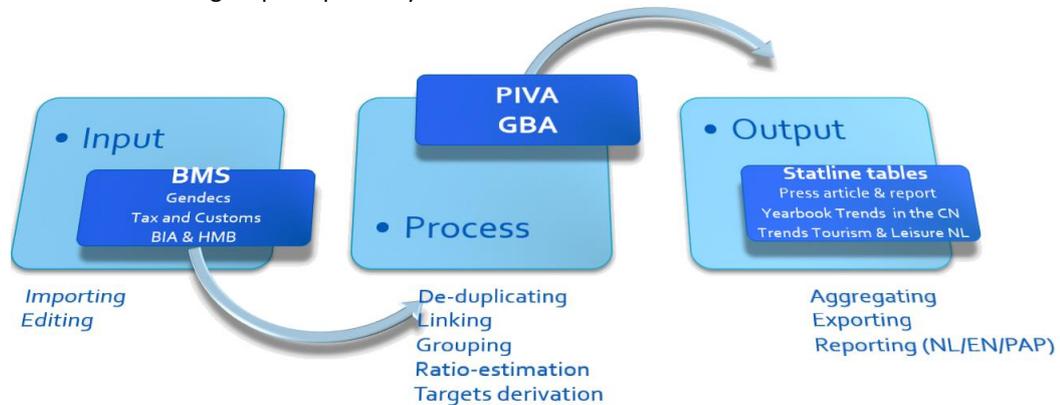


## 2.3 Inbound tourism by sea: Processing

With regard to shipping traffic, basically, the statistics are based on the Customs data only. These data involve large and small (motor) yachts, cruise ships (Bonaire only), ferries (only Saba), and other ships. Passengers of cruise ships and motor (yachts) can, roughly speaking, all be seen as visitors (excluding crew). This does not necessarily apply to passengers on the ferries, since these passengers may concern both residents and non-residents. The HMB file contains only the number of cruise passengers connected to Bonaire. It is unclear which part of the visitors on board actually disembarks. BREA reports that during the 2017/2018 cruise year[5], 79% of the cruise passengers disembarked and visited Bonaire [BREA, 2018].

A point of improvement on the inbound tourism by sea is related to incoming passengers arriving to Saba by ferry. The idea is to discriminate between residents- and non-residents passengers, and to identify their countries of origin and their lengths of stay.

# 3. Results

This section gives an overview of the tourism industry in Caribbean Netherlands and shows, in particular, how current tourism statistics can be totally based on available registers. No questions are asked to inbound visitors.

---

[4] The UNWTO divides inbound visitors by country of residence. However, in the Caribbean Netherlands, visitors often come from overseas territories of the United States of America, France or Great Britain, but also from their continental parts. Hence to refer to these overseas countries/territories, it can become extensive to name them all. Tables can become not legible. Thus, for these regional statistics, the nationality French often concerns residents from Saint Martin, Martinique, Guadeloupe, Saint Barthélemy; US Americans refer also to residents from the neighboring island of Puerto Rico and the US Virgin Islands. In the case of the British nationality, it concerns also residents of Anguilla, British Virgin Islands, Cayman Islands and Montserrat.

[5] The 2017/2018 cruise year includes the twelve months beginning in May, 2017 and ending in April, 2018.

## 3.1 Estimated number of inbound tourists

In this paragraph, the estimated number of visitors arriving by air and the number of passengers arriving by sea are shown.

### 3.1.1 Tourists by air

Figures 1a and 1b show the yearly total inbound tourism by air in the period In 2014-2017. In Figure 1a the yearly development of the number of visitors are compared among the islands of the CN using as base year 2014. The levels of Bonaire's inbound tourism have remained almost constant in this period, whereas the inbound tourism by air on St Eustatius and Saba has declined by -7 and -18 percent, respectively. Note that the drop in tourism on St Eustatius and Saba is partly related to 2017 Hurricane Irma. About Bonaire, the -5 percent decline in inbound tourism compared to 2016 was likely due to operational problems of a regional airline. Figure 1b shows the yearly total absolute number of arriving visitors on Bonaire, St Eustatius, and Saba. Bonaire receives by far the highest number of visitors. Account should be taken of the fact that figures published here also include business tourism. This means, for example, that aid workers and civil servants who do not live on the CN are also counted
as visitors.

Figure 1. Yearly total inbound tourism by air, Caribbean Netherlands 2014-2017*

a)



b)
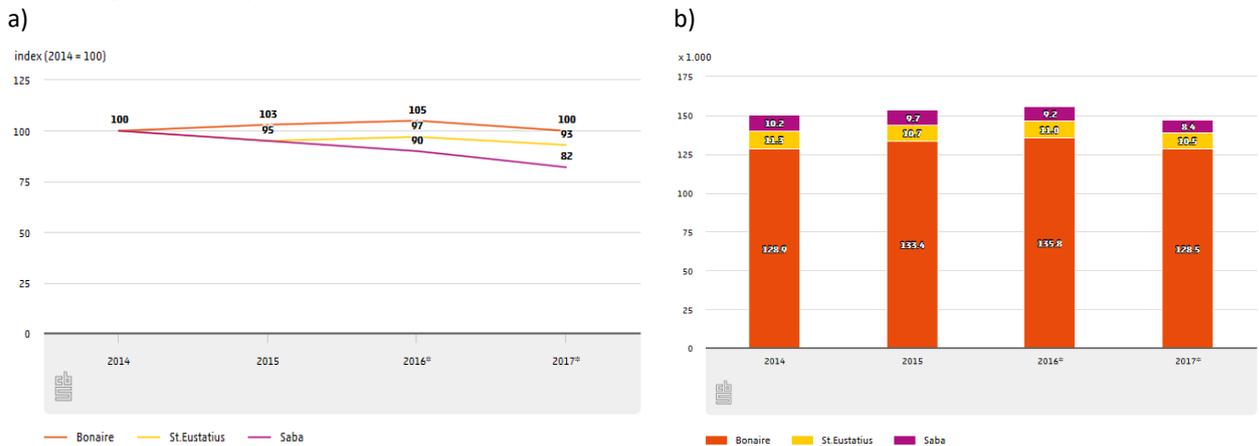


The monthly inbound tourism by air and the yearly inbound tourism by air and by nationality are shown in Figures 2 to 4 for Bonaire, St Eustatius, and Saba, respectively. These figures show that tourism in the CN is season- and holidays-related with peaks in the months of December / January, March, July / Augustus, and October. As for the nationalities of the inbound visitors, most visitors arriving by air in the CN are Dutch or US citizens.

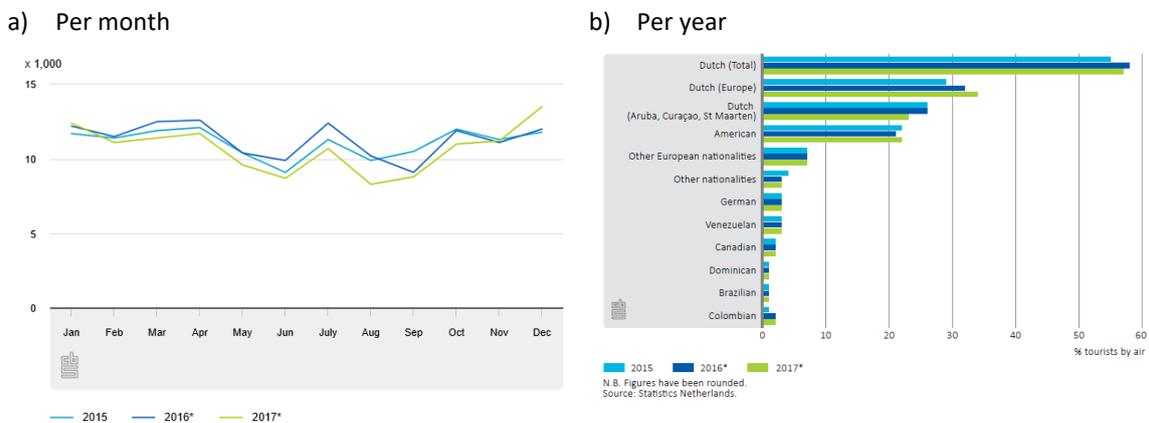Figure 2. Inbound tourism by air (left) and by air and by nationality (right), Bonaire

a)  Per month



b)  Per year



N.B. Figures have been rounded.
Source: Statistics Netherlands.

Figure 3. Inbound tourism by air (left) and by air and by nationality (right), St Eustatius

a)  Per month

b)  Per year
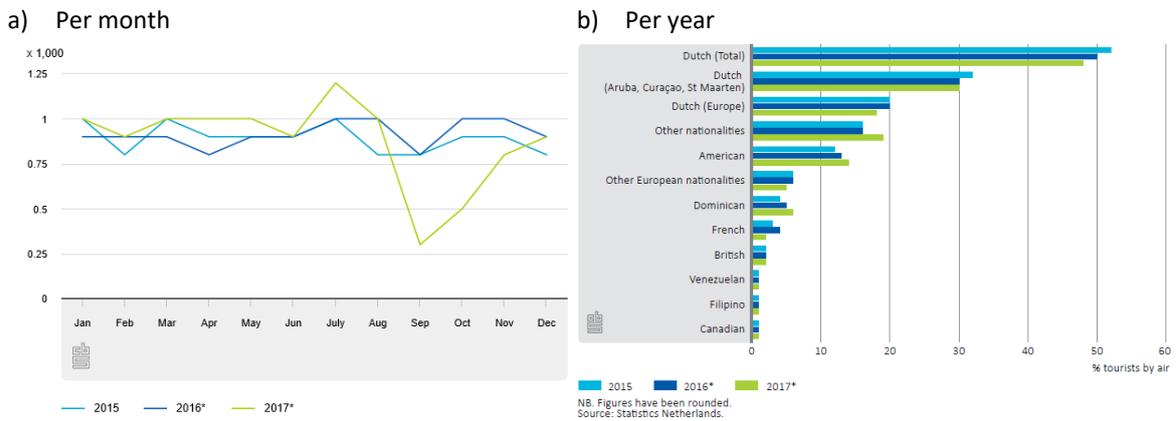


Figure 4. Inbound tourism by air (left) and by air and by nationality (right), Saba

a)  Per month

b)  Per year
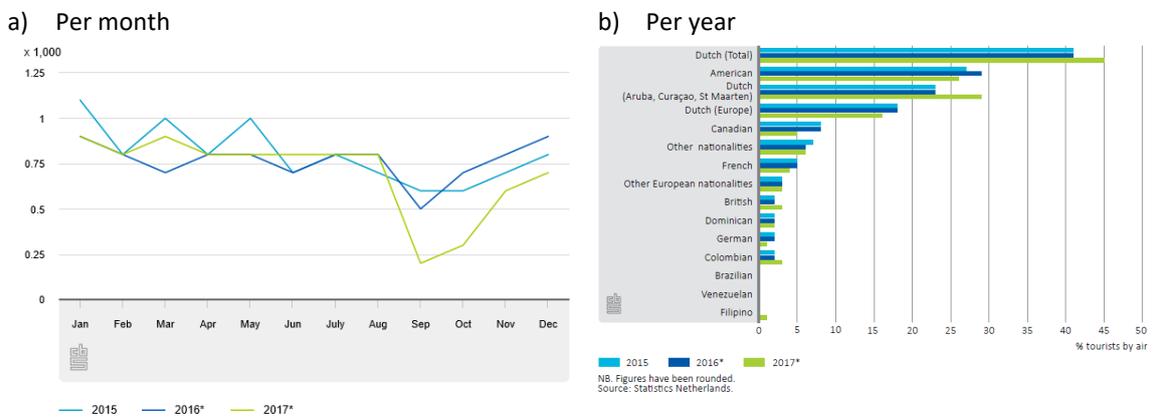


### 3.1.1  Passengers by sea

Figure 5a shows the seasonality of cruise arrivals on Bonaire. Until 2016, the cruise season started every year in October / November after the hurricane period and ended in May the year after. Since 2017, the cruise season has been expanded and cruises arrive almost all year round. Figure 5b shows that the number of cruise passengers on Bonaire more than doubled in 2017. In contrast to tourist air travel, 2017 was an exceptionally good year for Bonaire's cruise tourism. The number of cruise passengers increased by no less than 88 percent: 407.3 thousand cruise passengers in 2017 compared to 216.5 thousand in 2016 [CBS, 2018].

Figure 5. inbound cruise passengers per month (left) and per year (right), Bonaire
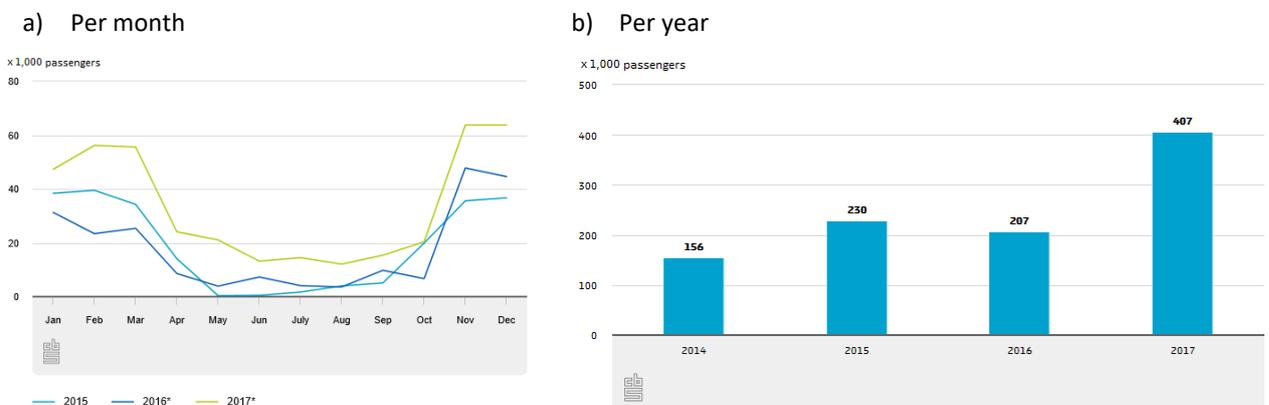
a)  Per month

b)  Per year



Figure 6a presents also the seasonality of the inbound passengers arriving by sea on Saba. Figure 6b shows that the number of passengers arriving by ferry or yacht to Saba is as important as the number of visitors arriving by air. The majority of these arrivals are tourists. In 2017, approximately 12.1 thousand passengers arrived on Saba by ferry or yacht. When adding together the traffic

arriving by air and sea, 2017 shows only a slight decline: approximately 20.8 thousand passenger arrivals in 2017 versus approximately 20.9 thousand in 2016. Finally, the large number of passenger arrivals by sea in December is striking [CBS, 2018].

Figure 6. Inbound passengers by sea per month (left) and per year (right), Saba



x 1,000 passengers

N.B. Figures have been rounded.
Source: Statistics Netherlands.
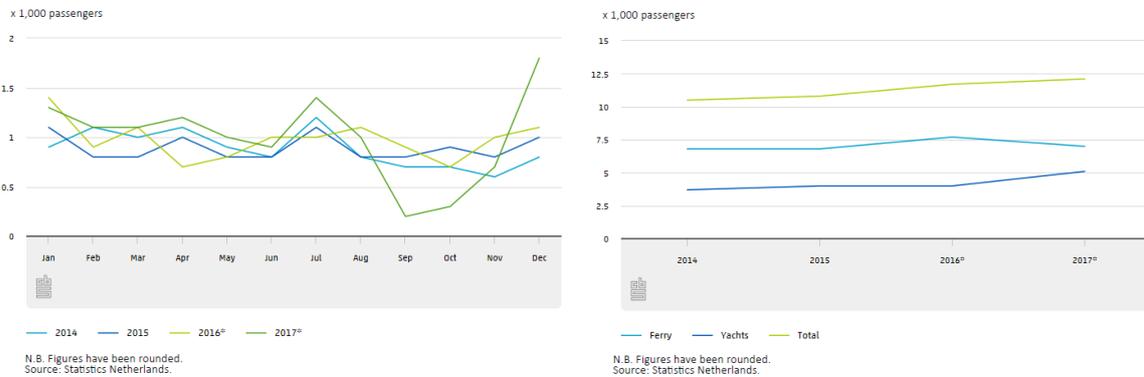
## 3.2 Estimated number of inbound day trippers, overnight stays and age distribution

Table 1 gives and overview of the number of inbound day trippers, length of stay, and age distribution of the tourists arriving by air in CN in 2016. Notice that the presented distributions have not changed drastically over the years.

1. One in ten visitors arriving by air on Bonaire are day trippers. This proportion is higher for visitors visiting St Eustatius and Saba by air, where one of the five and one of the four, respectively, are day trippers.
2. Just over half of overnight visitors flying in Bonaire stay for up to one week. In St Eustatius and Saba the proportions of tourists staying 1 to 7 nights are (almost) two of three tourists and (almost) four of five tourists, respectively.
3. On average, overnight visitors spent 9.6 and 9.5 nights on Bonaire and St Eustatius, with a total of 1.1 million and 85 thousand overnight stays, respectively, in 2016. The average of nights spend by visitors on Saba is below 9, namely 6.7 nights with a 48 thousand overnight stays in 2016. See also Figures 7a to 9a.
4. Most tourists arriving by air in the CN are between 40 to 59 years old. St Eustatius, in particular, is an important holiday destination for 40 to 59-year-olds and, to a lesser extent, for 30 to 39-year-olds. On the other hand, Saba attracts next to 40 to 59 year-old also relatively many tourists in the age category 20-39. In general, the age distribution of tourists in the CN has hardly changed over the years. See also Figures 7b to 9b [CBS, 2017a]. Note that the tail of the age distribution of St Eustatius is longer than those of Bonaire and Saba. The cumulative percentage of 35+ nights ranges between 6 to 8 percent over the last years.

Table 1. Summary overnight tourism and age distribution, Caribbean Netherlands 2016

| CN | % DAY TRIPPERS | % 1 - 7 NIGHTS *Exclusive day trippers* | AVERAGE # NIGHTS | TOTAL OVERNIGHTS | AGE VISITORS |
|---|---|---|---|---|---|
| **BONAIRE** | 10% | 57% | 9.6 | 1 100 000 | 40 – 59 years old |
| **ST EUSTATIUS** | 20% | 64% | 9.5 | 85 000 | 40 – 59 years old 30 – 39 years old (also) |
| **SABA** | 25% | 78% | 6.7 | 48 000 | 40 – 59 years old 20 – 39 years old (also) |

Another target variable that could be also extracted is the number of repeat visits to the CN.

Figure 7. Overnight tourism by air by length of stay (left) and age distribution(right), Bonaire
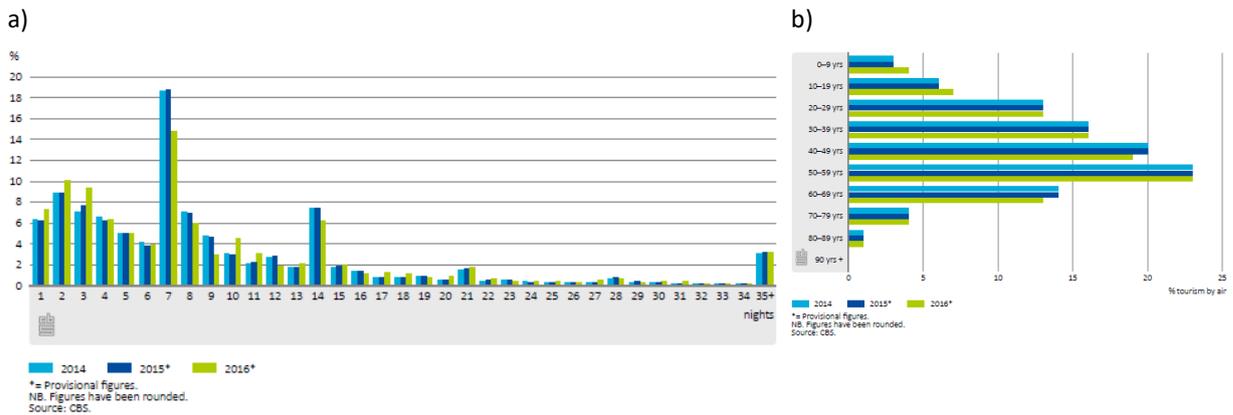
a)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.

b)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.

Figure 8. Overnight tourism by air by length of stay (left) and age distribution(right), St Eustatius

a)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.

b)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.
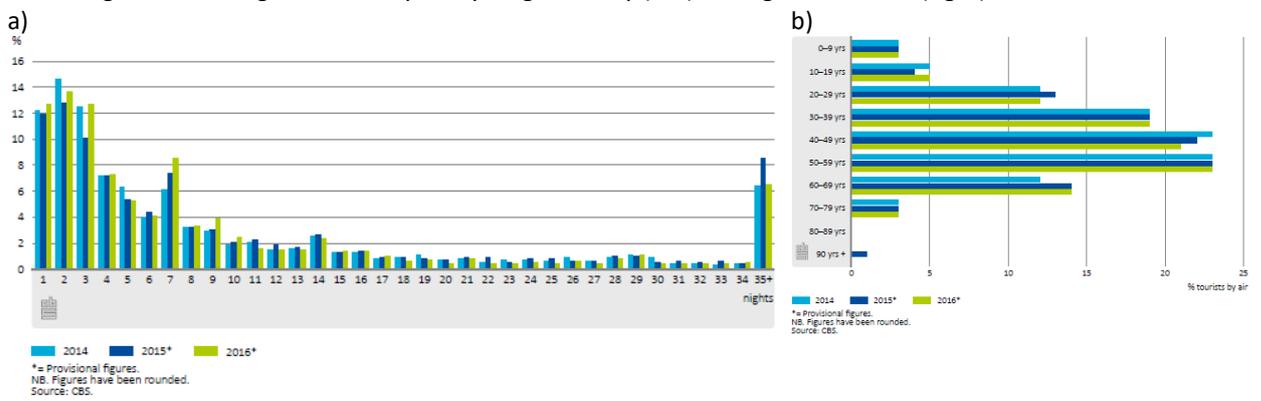
Figure 9. Overnight tourism by air by length of stay (left) and age distribution(right), Saba

a)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.

b)



*= Provisional figures.
NB. Figures have been rounded.
Source: CBS.
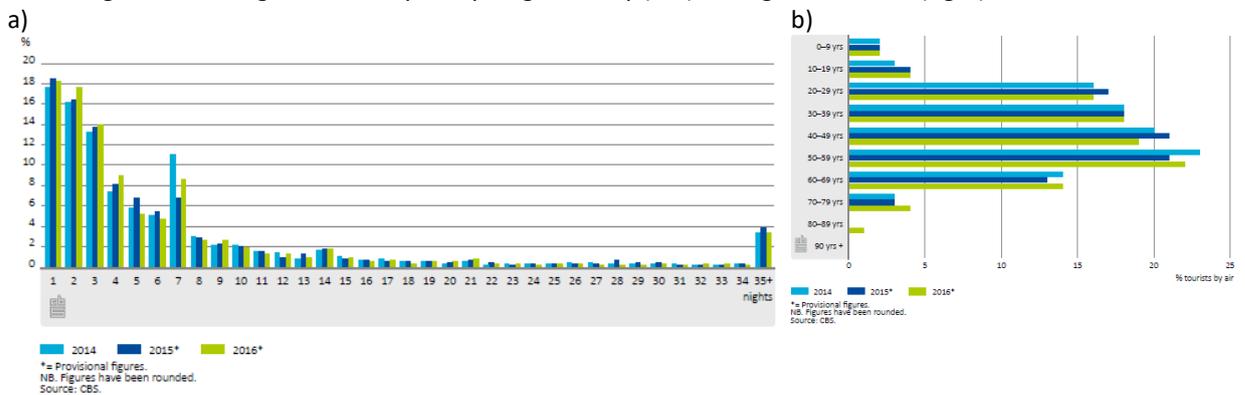
# 4.    Conclusions

**Main conclusions:**
Most important pros and cons of making regional tourism statistics based on registers are as follows:

- Until now SN is satisfied with the results of the production of basic tourism indicators based on registers. The main advantage is the reduction of the burden of statistical enquiries on respondents and also the lower-cost involved when compared with the traditional survey approach.
- However, the quality of the estimated statistics depends strongly on the quality of the registers. SN carries out structurally quality controls of the BMS-data, in particular. These data have shown issues connected to the functioning of the passport scanning facilities at the airports, dependence on internet connection, the update of their IT-system by a third-party, and especially the dependence on Border Police policies causing selection bias. For instance, cases in which some

airplanes are more strictly scanned than others based on their precedence or time of the day. The implementation of policies requires much communication.

- Rules of editing and plausibility checks of the BMS-data requires, among others, efficient matching of all available registers.
- The availability of all the registers at the same time requires effort of CBS-Caribbean Netherlands. They carry out the follow up and the set-up of service levels agreements with register owners. Note that currently, the data agreements related to the BMS-register is being revised in connection to the implementation of the General Data Protection Regulation (GDPR).
- If the proportion of missing data in the BMS-register surpasses predefined limits, the derivation of target variables like nationality and length of stay is dropped, in particular, when it concerns the small islands (St Eustatius and Saba). This also holds if a selection bias is detected, e.g., intensive scanning of some flights and shallow scanning of other flights has occurred, for instance during peak arrival hours at the airport(s).
- All types of visitors are taken into account in the estimates. Based on the current registers, we cannot discriminate visitors' travel intentions (e.g. business in the case of Bonaire), students (e.g. connected to the Saba University School of Medicine) or health (e.g. patients leaving the CN to seek specialized health services in other regional hospitals). The current registers contain less information on issues like expenditure, type of accommodation used, activities and the purpose of the trips, including business visitors. Since there is a policy need for this kind of information, this has to be derived from other sources, like an exit survey.
- Also, the estimation of the usage of airports, ports, tourist attractions and other facilities is an important issue that cannot be dealt with the current registers. Services' companies could offer their services more efficiently by knowing the touristic routes of e.g., cruise passengers upon disembarking in Bonaire. This also in connection to traffic jams when cruises reach Bonaire's sea ports.  Mobile phone data could be used in this case.
- SN pays attention to how the local tourism policy makers perceive this change on how tourism statistics are produced nowadays. The perception of the estimation quality is diverse. The main reasons are the new estimation method, the different data sources and, specially, regional statistics needs. The ED-cards remain an issue. Local policy makers still think this is the best method to gather data on visitors.  However, regional day trippers, frequent travellers, and family visits from the region were likely underestimated as visitors feel often themselves as locals while they are not residents of the CN. This method requires extra work and cooperation of many parties involved.
- Finally, sustainable tourism remains high on the agenda. Saba and St Eustatius are already engaged in driving the transition to green energy and thus to a greener economy [Bouma, 2017; The Daily Herald, 2018]. Bonaire faces the challenge of enabling sustainable growth of their cruise tourism. Some negative impacts have been already spotted by locals [Reporter, 2018]. Local policy makers could use overviews of the degree of pressure of tourism such as tourism densities (number of overnight stays per $km^2$) and tourism intensities (number of overnight stays of visitors per inhabitants) to monitor and tune the implementation of their policies in the CN, e.g., increase of tourist tax, promotion of cruise tourism.

**Other possibilities**

*Exit Survey (Pilot, phase 2)*
Besides the basic statistics, a first pilot was carried out to set up an exit survey. The electronic (tablet) exit survey system was designed, setup, and tested in St Eustatius. This survey concerned, among others, questions about expenditure before and during a trip to St Eustatius, the number of traveling companions (e.g. none, family, friends, or organized travel group), expenses related to the booking of the airplane tickets, choice and costs of accommodation, costs of food and beverage, expenses on transportation while on St Eustatius (including rental car and taxi's), money spent on shopping (including gifts and souvenirs), expenditures on sport activities like diving (including equipment, lessons and fees), sailing (including motor yachting), other sport activities and eventually also recreation and entertainment expenses (e.g. entrance fees). The enquiry also includes questions on the used accommodations, number of repeat visits, and overall visitors' satisfaction with their trip [CBS, 2017b].

Main conclusions of this pilot were that a) a good technical layout of the screens and usability of the survey are important elements for the willingness of the respondents to participate; b) the financial questions remain a burden and often difficult to understand for the respondents despite all the efforts to simplifying the questions and preparing the interviewers, and c) the cooperation and communication with of the local interviewers is essential. Instead of using an electronic version, St Eustatius Tourism Development Foundation (STD) also reintroduced a Visitor Survey Card (paper version) in cooperation with the local government and the immigration office since July/Augustus 2018.

*Mobile telephone*
Another option is to investigate the use of mobile phone data. The supply side of the telecom sector in the CN includes a very few large regional companies and relatively small local ones. In fact, these regional companies operate across the Caribbean region. The main telecom provider in Bonaire holds a share as big as 60% of the telecom market [SEO, 2016]. Partnerships with non-governmental organizations that analyse mobile data in the Caribbean region are relevant. These organizations have specific knowledge on the use of telecom data since they already engaged on projects of disaster response, population displacement, epidemiology, and need of humanitarian assistance after Haiti hurricane Matthew in 2016 and also after the 2010 Haiti Earthquake [Flowminder, 2018]. SN has knowledge on structural and methodological processing and analysis of mobile telephone data for Tourism-related purposes [Offermans et al, 2013].

*Redesign Web scraping*
The use of scraping methods combined with machine learning tools is also interesting for the CN in order to produce relevant and updated figures on the number of accommodations and the number of bed places. The idea is to use aggregated sites to update the 2013 accommodations database. The crucial steps are: a) find groups of websites which cover the whole domain as best as possible, b) to use new data sources as social media choosing semantically efficient sets of words in an automatic, strategic, refined and structured manner, and c) identify machine learning techniques with the best discriminating performance. For example, a tourist accommodation often will post its information on more than one website, namely its own and aggregation sites. This means that after the collection of the data, a de-duplication process has to be implemented, e.g. based on names or addresses of the accommodations [Heerschap et al 2014]. The idea is to use machine learning to compare not only text but also images.

### *References*

1.  Heerschap, N. Ortega S., Priem A. and Offermans M. (2014), "Innovation of tourism statistics through the use of new big data sources". In 12th Global Forum on Tourism Statistics, Prague, CZ.
2.  CBS (2018) "Tourism in the Caribbean Netherlands down in 2017", 8 May 2018.
3.  BREA (2018) "Economic Contribution of Cruise Tourism to the Destination Economies: A Survey-based Analysis of the Impacts of Passenger, Crew and Cruise Line Spending", Destination Report: Bonaire, Florida Caribbean Cruise Association, Business Research & Economic Advisors, October 2018.
4.  Bouma, J.(2017) "Sint-Eustatius haalt nu al de helft van de stroombehoefte uit zon", Trouw, In Dutch, 3 November 2017.
5.  The Daily Herald (2017) "Share of solar power on the rise in Statia, Saba ", 17 February 2017.
6.  CBS (2017a), "Inbound tourism in Caribbean Netherlands 2014 - 2016", 4 September 2017.
7.  CBS (2017b), "Exit Survey Tourists St Eustatius 2.3".
8.  UNWTO (2006), "Glossary of Statistical Terms", United Nations and World Tourism Organization. Recommendations on Tourism Statistics. Series M, No. 83. (para. 29).
9.  SEO (2016) "Telecommunicatie in Caribisch Nederland", In Dutch.
10. Flowminder (2016) "Haiti Hurricane Matthew 2016", Case study.
11. Offermans, M., Priem, A., and Tennekes, M. (2013), "Rapportage Project Impact ICT; Mobiele Telefonie", In Dutch, 16 April 2013.
12. Reporter (2018), "Questioning Bonaire Cruise Ship Economics", 7 May 2018.