

Understanding the landscape of inbound tourism measurement

A report by

Jason Li Chen, Gang Li, Anyu Liu, and Allan Williams

School of Hospitality and Tourism Management, University of Surrey

Guildford, United Kingdom

July 2021

Executive Summary

Context. The tourism industry increasingly requires detailed inbound tourism measures. The measurement approaches used across international destinations have commonly faced challenges in a number of areas such as limited regional sample size, increasing difficulty to recruit respondents especially during situations such as the COVID-19 pandemic, as well as the financial and time costs of data collation.

This study has been commissioned by VisitBritain to analyse the approaches of a range of selected destinations to measuring inbound tourism and identify the strengths and limitations of different methods. The report overviews the International Passenger Survey (IPS) used in the UK as a tool to measure inbound arrivals and expenditure. It then summarises the approaches used by twelve destinations which share some similarities to the UK: Australia, New Zealand, Japan, Ireland, South Korea, USA, South Africa, Hong Kong, the EU as a whole as well as France, Austria and Saudi Arabia. Details of official measures such as the data collecting methods, data processing, publishing schedules are provided and mapped. Additionally, given the interruption during the COVID-19 crisis, the study explores complementary data sources that can be applied as mitigation when traditional data collection is substantially interrupted. Alternative potential or emerging types of measures considered or used by these destinations, including mobile positioning data, bank card transaction data and Google trends data are investigated. Further details of the approaches used by each of the selected destinations are provided in the form of case studies.

Approach. A mixed-method approach has been applied in this study including both desk research and interviews with relevant tourism or statistical authorities. Specifically, the study has been carried out in four stages including destination selection, desk research, interviews and the exploration of complementary data sources.

Visitor arrivals. It has been found that three main sources of data have been used to measure visitor arrivals across destinations: 1) data from passport control or entry ports; 2) sample or census surveys collected from accommodation establishments; and 3) data from visitor surveys. For destinations with stricter border control, visitor arrivals are usually collected from the administrative source at the entry ports or passport control. For destinations within the Schengen area, due to the absence of internal border controls, accommodation statistics are often used to measure visitor arrivals. Other destinations conduct various forms of visitor surveys, usually at airports and seaports, to estimate visitor arrivals. All the selected destinations publish the arrivals data monthly.

Visitor expenditure is normally estimated based on visitor surveys, complemented with additional sources such as administrative data or census data to stratify or weight the survey sample. Most destinations conduct the survey at departure ports or border-crossing points in the format of face-to-face interviews. Among the selected destinations, New Zealand and Austria are the only exceptions that employ self-administered online surveys. Due to the disruption of the COVID-19 pandemic, Hong Kong has also experimented with this method. Most of the destinations publish expenditure data quarterly in the format of compiled survey results or as part of the travel Balance of Payments.

A summary of the measurement approaches is presented in the table below. The strengths and limitations of each approach as well as the recommendations are summarised in Section 5 of the report.

Summary of approaches to measuring inbound tourism

Destination	Visitor arrivals					Visitor expenditure			
	Data source			Monthly	Face-to-face interviews	Self-administered online survey	Publishing frequency (highest)		
	Passport control/Entry ports	Accommodation statistics	Visitor survey				Monthly	Quarterly	Annually
UK			✓	✓	✓		✓		
Australia	✓			✓	✓			✓	
New Zealand	✓			✓		✓		✓	
Ireland			✓	✓	✓			✓	
Japan	✓		✓	✓	✓			✓	
South Korea	✓			✓	✓		✓		
Austria		✓		✓		✓		✓	
France		✓		✓	✓		✓		
EU-wide	Most states			✓	Most states			✓	
USA	✓		✓	✓	✓			✓	
South Africa	✓			✓	✓			✓	
Hong Kong	✓			✓	✓	During COVID-19 pandemic			✓
Saudi Arabia	✓			✓	✓		Planned		

Table of Contents

1 Introduction	7
2 Methodology	9
2.1 Destination selection	9
2.2 Desk research	10
2.3 Interviews	10
2.4 Exploration of complementary data sources	10
3 Overview of approaches	11
3.1 The UK International Passenger Survey	11
3.2 Approaches used in selected destinations – visitor arrivals	12
3.3 Approaches used in selected destinations – visitor expenditure	15
3.4 Alternative data sources	18
3.4.1 Mobile positioning data.....	18
3.4.2 Bank card transaction data.....	19
3.4.3 Google trends data.....	19
4 Case Studies.....	21
Australia.....	21
1. Visitor Arrivals and Departures	21
1.1 Data source.....	21
1.2 Data sampling	21
1.3 Publishing schedules and format.....	21
2. Visitor Expenditure	22
2.1 Survey collection and sampling	22
2.2 Data processing	23
2.3 Publishing schedule and format	23
2.4 Changes due to COVID-19	23
2.5 Alternative plans/future plans	23
New Zealand.....	24
1. Visitor Arrivals	24
2. Visitor Expenditure	24
2.1 Sampling and data collection method.....	24
2.2 Data processing	24
2.3 Publishing schedules and format.....	24
2.4 Regional level data.....	25
2.5 Changes due to COVID-19	25

Ireland.....	26
1. Visitor Arrivals	26
1.1 Sampling and data collection method.....	26
1.2 Data processing	26
1.3 Publishing schedules	26
1.4 Changes due to COVID-19	26
2. Visitor Expenditure	27
2.1 Sampling and data collection method.....	27
2.2 Data Processing.....	27
2.3 Publishing schedules and format.....	27
2.4 Alternative sources/future plans	27
Japan.....	29
1. Visitor Arrivals	29
2. Visitor Expenditure	29
2.1 Sampling and Data collection method	29
2.3 Publishing schedules and format.....	29
South Korea.....	30
1. Visitor Arrivals	30
2. Visitor Expenditure	30
2.1 Sampling and data collection methods	30
2.2 Data Processing.....	30
2.3 Publishing schedules	30
2.4 Alternative sources.....	30
Austria	32
1. Visitor Arrivals	32
1.1 Data source.....	32
1.2 Publishing schedules	34
1.3 Lessons learned from the interview.....	34
2. Visitor Expenditure	34
2.1 Survey collection and sampling	34
2.2 Publishing schedule and format	35
2.3 Alternative plans/future plans	35
France	36
1. Visitor Arrivals	36
1.1 Data source.....	36
1.2 Sampling	36
1.3 Publishing schedule	36

1.4 Changes due to COVID-19	36
2. Visitor Expenditure	36
2.1 Sampling and data collection method.....	36
2.3 Publishing schedules and format.....	37
2.2 Alternative sources.....	37
2.3 PiLOT project	37
European Union	39
1. Visitor Arrivals	39
1.1 Data source.....	39
1.2 Publishing schedules and format.....	39
2. Visitor Expenditure	39
2.1 Sampling.....	39
2.2 Data collection method.....	40
2.3 Publishing schedule and format	40
USA	42
1. Visitor Arrivals	42
1.1 Data source.....	42
1.2 Sampling.....	42
1.3 Publishing schedules and format.....	43
2. Visitor Expenditure	43
2.1 Data collection and sampling	43
2.2 Data processing	43
2.3 Publishing schedule and format	43
2.4 Changes due to COVID-19	44
2.5 Alternative plans/future plans	44
South Africa	45
1. Visitor arrivals	45
1.1 Data source.....	45
1.2 Publishing schedules and format.....	45
2. Visitor Expenditure	45
2.1 Sampling and data collection method.....	45
2.2 Survey design	45
2.3 Data processing	46
2.4 Publishing schedules and format.....	46
2.5 Alternative data sources.....	46
Hong Kong.....	47
1. Visitor Arrivals	47

1.1 Data source.....	47
1.2 Publishing schedules	47
2. Visitor Expenditure	47
2.1 Sampling and data collection	47
2.2 Publishing schedules and format.....	47
2.3 Changes due to COVID-19	47
2.4 Alternative data sources/future plans	48
Saudi Arabia	49
1. Visitor arrivals.....	49
1.1 Data source.....	49
1.2 Publishing schedule and format	49
1.3 Alternative sources/future plans	49
2. Visitor expenditure.....	49
2.1 Data collection and sampling	49
2.2 Data processing	50
2.3 Publishing schedule and format	50
2.4 Alternative data sources/future plans	50
5. Conclusion	52
References	53
Summary of interview sources	54

1 Introduction

Tourism is one of the largest economic sectors in the world and contributes about 10% to global GDP when taking account of direct, indirect and induced effects together (World Travel and Tourism Council, WTTC, 2020). Tourism is also one of the most important sectors in the UK's economy (VisitBritain, 2020). A pre COVID-19 forecast noted it was likely to continue to grow to a value of over £257 billion by 2025 (VisitBritain). This highlights the importance of inbound tourism, as an export and stimulus to the British economy. Therefore, measuring inbound tourism accurately has crucial implications to monitoring and understanding the UK's visitor economy, and for designing policies to support it.

Tourism or, in a broader sense, the visitor economy is complex. A full understanding of the landscape of inbound tourism measurement should take account of the complexity of the landscapes of the tourism industry and tourism boards. From a visitor experience perspective, tourism is a composite product that involves a combination of a variety of goods and services provided by different sectors such as visitor attractions, tour operators, travel agencies, accommodation, food and beverage, transport, and retailing (Song, Dwyer, Li and Cao, 2012). This means that tourism statistics should aim to take full account of these economic activities in addition to enumerating visitor numbers.

From a tourism supply and destination management perspective, the tourism industry is a dynamic affiliation of private and public organisations ranging from small- to medium-size enterprises (SMEs) and international businesses, to destination management organisations (DMOs) at local and regional levels (VisitBritain, 2020). Tourism statistics should take account of the needs of these different stakeholders.

The complexity of tourism can also be seen from its intra-regional spill-overs and regional disparities. The intra-regional spill-over effects arise from tourists' multi-destination travel itineraries, cross-region supply chains, labour mobility and investment penetration (Li, Chen, Li and Goh, 2016). Meanwhile, major regional disparities exist in inbound tourism which are substantially, but not only, due to imbalanced tourism resource endowment and tourism infrastructure across the different regions of the UK. The production and presentation of tourism statistics should take these complex regional relationships into account in order to inform the development of effective tourism policies which address such regional inequalities. In the UK, VisitBritain/VisitEngland operates alongside independent tourist boards in Scotland, Wales and Northern Ireland: Visit Wales, VisitScotland and the Northern Ireland Tourist Board, respectively; in addition, London has its official promotional agency: London & Partners. Furthermore, there are a wide range of regional DMOs and local authorities in the UK with responsibility for tourism at different scales (e.g., Tourism Southeast, Visit Surrey and Visit Guildford). Effective collaborations among the national, regional and local DMOs are not only dependent on but also necessary to obtain, high quality data. Improved tourism measurements are essential to serve the needs of different levels of regional/local destination management and help address regional inequalities in tourism development.

Although the significance of tourism statistics, especially inbound tourism measures, in informing decision-making, performance evaluation, planning and forecasting has been widely acknowledged, tourism and statistics authorities face persistent challenges (Volo, 2019):

- 1) deficiencies in the systematic collection of elementary data; and

- 2) difficulties in accounting for the complex nature of tourism which requires appropriate, and sophisticated, operationalisation and measurements of the multi-dimensional nature of inbound tourism.

In the UK, the tourism industry increasingly requires detailed inbound tourism measures. For businesses, the real value comes in understanding visitors to a specific region or local area, or from a specific country of origin. But the International Passenger Survey (IPS) sample size places limits on the robustness of such analysis, as its overall sampling methodology was not initially designed with this purpose in mind.

In particular, there are two major challenges related to UK inbound tourism statistics:

- 1) relatively small sub-sample size by country of residence, quarter, purpose, or area visited, which generates wide confidence intervals, especially for spend; and
- 2) increasingly difficult to recruit respondents for the IPS, especially during situations such as the pandemic.

This study therefore analyses the approaches of a carefully selected group of destinations to measuring inbound tourism and will identify strengths and limitations of different methods. Given the interruption of the IPS during the COVID-19 crisis, mitigation measures that can be employed in similar situations are explored.

2 Methodology

The study aims to achieve the following four research objectives:

- 1) To identify appropriate examples of destinations and how they measure inbound tourism;
- 2) To source information about what official statistics/measures are reported publicly;
- 3) To review what other data are looked at/have been considered by the selected destinations; and
- 4) To identify best practice for VisitBritain.

First, an interview was conducted with the Office for National Statistics (ONS) team with detailed knowledge of the IPS in the UK. This provides a deeper understanding of the data collection process and of any changes under consideration, which facilitates the analysis of the project. Secondly, we undertook a rapid assessment by email survey with local tourism organisations to confirm and identify their needs as leading users of inbound tourism statistics. Then, to achieve the above objectives, a mixed-methods approach including both desk research and primary research (i.e., interviews with relevant statistical or tourism bodies) were applied. Specifically, the study was carried out in four stages including destination selection, desk research, interviews and the exploration of complementary data sources. We had extensive discussion with the VisitBritain inbound tourism team at the kick-off meeting to confirm the details of the project design.

2.1 Destination selection

Following the guidelines provided by VisitBritain, twelve tourism destinations were selected including the following categories:

- 1) Similar destinations, such as island destinations or limited land borders: Australia, New Zealand, Japan, Ireland and South Korea;
- 2) Similar challenges to the UK for inbound tourism measurement in terms of destination appeal: the EU destination group as a whole, as well as France and Austria. Austria is cited by the Eurostat methodological manual as a good practice case study, while France has similarities to the UK in terms of scale and product diversity¹;
- 3) Similar challenges to the UK for inbound tourism measurement in terms of access (air links only with most visitor markets): USA, South Africa, Hong Kong; and
- 4) A representative Mideast destination: Saudi Arabia.

¹ Ireland, France, Germany and Austria broadly follow Eurostat guidance on data collection, but also have some specific national and regional data collection features, as well as different approaches to engaging with stakeholders and users of these statistics.

2.2 Desk research

Desk research was employed to analyse the selected destinations' approaches to measuring inbound tourism. Details of official measures such as the measurements, sampling approach and/or data collecting methods, data processing, and publishing schedules were collated, mapped and presented in our report. Alternative potential and/or emerging types of measures considered by these destinations were also investigated where applicable. The advantages and disadvantages of each method were also analysed.

2.3 Interviews

After the desk research and analysis of the collected methodological documents, key informants were interviewed from relevant national tourism or statistical organisations in the above selected destinations. Some academics in the field of tourism economics were also approached for interviews to obtain further insight into the measurement of inbound tourism from their perspectives. The key issues covered include the following:

- To understand their evaluation of the performance of existing measurements, and their strengths and weaknesses;
- To understand other measures considered by the organisations and any lessons learned;
- To explore how these organisations communicate with stakeholders about the data: including the type of information, format and frequency shared outside their organisation;
- To explore mitigation measures when traditional data collection is interrupted in situations such as the COVID-19 crisis.

2.4 Exploration of complementary data sources

There have been a range of emerging alternative data sources that may complement official tourism statistics at both national and regional/local levels. However, there has been much speculation about but very little evaluation of these potentially important data sources. Through the Granger causality test, the research team explored the potential of Google Trends data as a complementary source in measuring inbound tourism performance.

3 Overview of approaches

3.1 The UK International Passenger Survey

The main source of information on the UK inbound tourism performance is the International Passenger Survey (IPS). The IPS collects information about passengers entering and leaving the UK at major airports, sea routes, Eurostar terminals and on Eurotunnel shuttle trains. The weighted survey results are used for multiple purposes, such as feeding into the Travel Account of the Balance of Payments and Tourism Satellite Account (TSA), and estimating the inbound visitor arrivals and expenditure.

The IPS is conducted by the ONS through face-to-face interviews throughout the year. Each year, between 700,000 and 800,000 interviews are conducted, of which approximately 250,000 interviews are used to estimate the statistics for travel and tourism. The travel and tourism data include overseas residents who are departing the UK and UK residents returning to the UK. As a result, a proportion of the sample forms the basis of inbound tourism statistics. The sample for inbound tourism in the last 10 years (2010-2019) was around 42,000 per annum.

The sampling of the IPS follows a multi-stage design. Depending on the transport mode, time periods/shifts or sea crossings are selected at the first stage. Passengers are then systematically chosen at fixed intervals from a random start within sampled shifts or crossings at the second stage.

The inbound visitor arrivals and expenditure data are published monthly, quarterly and annually with different level of details. The monthly data provide provisional estimates which are published approximately 7 weeks after the end of the reporting month. The quarterly publication contains more accurate and detailed estimates and is published approximately 3 and a half months after the end of the reporting period. An annual publication is published approximately 4 months after the end of the year providing annual estimates.

The IPS is aligned with recognised best practice and international standards set out by the United Nations World Tourism Organization (UNWTO) and the Organisation for Economic Co-operation and Development (OECD). A key strength of the IPS is that it provides a comprehensive time series of tourism performance data that are useful for identifying long-term trends and patterns. The IPS data have long been used by academics and industry practitioners for tourism demand modelling and forecasting. The importance of the IPS has been emphasised by local DMOs which regard it as the single consistent source of data which is available for a wide range of purposes such as identifying key source markets, creating market profiles, benchmarking tourism performance, assessing the impacts of Brexit and the COVID-19 pandemic, and informing strategic direction and planning. In addition, the use of a dedicated field force helps ensure that respondents have a consistent understanding of survey questions. The clear documentation of methodology also provides valuable information for users.

As with any survey approach, there are challenges mainly related to the costs and difficulty in recruiting survey respondents. The travel and tourism interviews select passengers who are ending their visit. As a result, the large proportion of passengers approached who are starting their visit are not eligible for the interview. Another issue faced globally is that the IPS was suspended during the COVID-19 pandemic. IPS data are also subject to both

sampling and non-sampling errors. Around 95 per cent of passengers entering and leaving the UK have a chance of being selected to participate in the survey. The remainder include passengers travelling at night when the survey is suspended, and on those routes too small in volume or too expensive to be covered. Another challenge is that the sample size for certain regions/cities or particular routes are sometimes small which consequently generate higher sampling errors and less reliable estimates.

3.2 Approaches used in selected destinations – visitor arrivals

This section summarises the methods that the selected destinations used to compile visitor arrivals data. Overall, three main types of approaches have been covered. For destinations with stricter border control (i.e., Australia, New Zealand, Saudi Arabia, USA, Japan, South Africa, South Korea and Hong Kong), visitor arrivals are collected at the entry port or passport control before entering the border of the destination. For Schengen destinations such as Austria and France, due to the absence of internal border controls, accommodation statistics are usually used to estimate the aggregated visitor arrivals at the local level and national level. Statistics Austria receives data from municipalities, who collect accommodation statistics from each accommodation establishment. The National Institute of Statistics and Economic Studies (INSEE) in France instead send monthly attendance surveys to tourist accommodations. Both of the procedures in Austria and France are compulsory. Other destinations measure visitor arrivals through surveys. For example, Ireland conducts the country of residence (CRS) survey at departure gates of airports and seaports to estimate visitor arrivals. Japan also uses an overnight Travel Statistics Survey to complement the administrative data obtained at the border. From that particular survey, the number of stays of foreign tourists in each prefecture is reported on a monthly basis, according to Japan Tourism Agency (JTA).

One potential disadvantage of collecting visitor arrivals data at border control points is that visitors from certain source markets might be excluded from the system. For example, international travellers from New Zealand are not included in the data collection system at the border of Australia. The I-94 programme in the US exempts travellers from Canada and Mexico. In Ireland, trips made by residents of Northern Ireland into and out of the Republic of Ireland are excluded. But this disadvantage can be potentially addressed by complementing the data collection system with data shared by the source markets. The I-94 programme in the US is complemented with the data from Statistics Canada's International Travel Survey and the Instituto Nacional de Estadística y Geografía/INEGI (Banco de Mexico) to estimate full international arrivals to the US. The Central Statistics Office (CSO) in Ireland also exchanges information with the Northern Ireland Statistics and Research Agency (NISRA) to collect data on visitors who travel across the border between the Republic of Ireland and Northern Ireland.

Accommodation statistics also have some disadvantages according to the interview with Statistics Austria. While the majority of visitors should be captured by this measure, one of the limitations is that informal accommodation is not covered. Another issue is related to multiple counting. If tourists visit multiple regions within the destination, or stay in multiple accommodations, while the number of nights stayed would not be a problem, their arrivals would be counted multiple times. So far, no filtering or adjustment solution has been established to solve this issue in the accommodation statistics.

The approaches used in the UK along with 12 selected destinations to measuring visitor arrivals are summarised in Table 1. More details for each selected destination are presented in the case studies in Section 4 of this report.

Table 1. Summary of measurement approaches in selected destinations – visitor arrivals

Destination	Approach	Details of approach	Publishing schedule
UK	IPS	The IPS is conducted by the ONS at major airports and sea routes etc. through face-to-face interviews. Passengers are sampled as they enter or leave the UK following a multi-stage sampling design.	Monthly estimates are published 7 weeks after the end of the reporting month; more accurate and detailed quarterly estimates are published with a lead time of 3.5 months; annual estimates are published approximately 4 months after the end of the year.
Australia	Passport control	Administrative information on arrivals and departures are collected via passport documents, visa information and incoming passenger cards by the Home Affairs department and then supplied to ABS.	Overseas Arrivals and Departures (OAD): monthly, approximately 6 weeks after the reference period. Disaggregated OAD: quarterly.
New Zealand	Custom card data	Custom card data are collected by the International Travel & Migration Statistics to count passengers arriving at and departing from New Zealand.	The arrivals data are published on a monthly basis.
Ireland	Country of Residence Survey (CRS)	The CRS survey is conducted face-to-face at the departure gates at airports and seaports by a Central Statistics Office (CSO) interviewer. A sample of sailings and flights is selected and a systematic sample of passengers on each is surveyed.	Provisional results are not published for the Country of Residence Survey and final results are published on a monthly basis within 28 days of the end of the reference period.
Japan	Border and Overnight Travel Statistics Survey	The visitor arrivals data are collected by the Ministry of Justice. An additional 'Overnight Travel Statistics Survey' is conducted by JTA, which collects the number of stays of foreign tourists in each prefecture on a monthly basis.	Breakdown of total number of overnight guests by place of residence in each month and breakdown of total number of foreign guests staying in each month by nationality are published monthly.
South Korea	Immigration office	Statistics of visitor arrivals, Korean departures and cruise visitors are aggregated and compiled by the Ministry of Justice, Korea Immigration Service.	The tourism statistics in South Korea are published monthly.

Destination	Approach	Details of approach	Publishing schedule
Austria	Monthly accommodation statistics	The basis of the monthly accommodation statistics are the reports from municipalities. The reporting municipalities themselves receive the data from local accommodation establishments. The aggregated municipal results are then transmitted to Statistics Austria.	Accommodation statistics are available monthly. The preliminary results are due to be available 24 days after the reference period, whereas the final results would be available in a 50-day lead time.
France	Monthly attendance survey	The visitor arrivals data are compiled through the monthly attendance survey in tourist accommodation by the National Institute of Statistics and Economic Studies (INSEE). The survey is collected by sending a mail or email to the tourist accommodations at the beginning of month M+1 for the data of month M.	The provisional data dissemination is released at the beginning of the month M+2 for the month M. Final data dissemination is released at the end of the month M+3 for the month M
EU-wide	Accommodation statistics	The reporting unit of each EU member state collects arrivals statistics from accommodation establishments and transmits the data to the national statistical institutes. The national level data are then compiled and published by Eurostat.	Monthly accommodation statistics have a lead time of: 8 weeks: nights spent (national level); or 3 months: nights spent, arrivals, net occupancy rates (national level). Annual accommodation statistics have a lead time of: 6 months: establishments, bed places, arrivals etc. (national level); or 9 months: nights spent in non-rented accommodation (national level).
USA	I-94 arrivals programme	The I-94 arrivals programme consists of three parts: Department of Homeland Security (DHS)/U.S. Customs and Border Protection (CBP), Statistics Canada's International Travel survey and the Instituto Nacional de Estadística y Geografía/INEGI (Banco de Mexico). A complete I-94 arrivals record includes information transmitted to the CBP using Advanced Passenger Information System (APIS), information provided by the air or sea carrier about the flight or voyage and visa information by the U.S. Department of State.	Summary data of visitor volume by both country of residence (COR) and country of citizenship (COC) are posted monthly to trade.gov

Destination	Approach	Details of approach	Publishing schedule
South Africa	Entry ports	Data are routinely collected by immigration officers at all entry ports on all travellers arriving at or departing from South Africa directly thorough the travel documents either by scanning or capturing these into the ports' electronic database.	The statistical release of the Tourism and Migration report is schedule with an advanced release calendar available on the Statistics South Africa website monthly.
Hong Kong	Immigration office	Statistics of visitor arrivals in Hong Kong are collected by immigration officers at all entry ports.	Visitor arrivals are published monthly with a one-month lead time.
Saudi Arabia	Passport control	Administrative data such as nationalities, date and port of arrival and gender are obtained on a regular basis.	Inbound data for the previous month is released on the 15th of each month.

3.3 Approaches used in selected destinations – visitor expenditure

Visitor expenditure is estimated mainly from visitor surveys conducted at the destination, along with some additional data sources such as administrative data or census data to stratify or weight the survey sample. Most destinations conduct the survey in departure lounges of airports, seaports or border-crossing points in the format of face-to-face interviews. A potential issue with the survey is that respondents may have different understanding of some terminologies and definitions in the survey questions. Face-to-face interviews carried out by well-trained interviewers can help ensure a universal and consistent understanding of the questions. Face-to-face interviews also help to control the sample size taken in each departure port. For example, the number of questionnaires distributed by ports is proportional to the traffic in each port in Saudi Arabia and the US. Other destinations including Australia and Ireland sample departing flights and sailings first and then distribute the surveys within these samples. Among the selected destinations, New Zealand and Austria are the only exceptions that use online questionnaires, but utilising different approaches. In New Zealand, travellers are sampled from international visitors departing from major international airports. They will be asked if they are willing to take the survey and, if eligible, the interviewer will ask for their email address to send a link to the International Visitor Survey (IVS). In Austria, the T-MONA (Tourism Monitor Austria) survey is collected via various channels such as national tourism organisations, participating destinations and other partners. The survey is also accessible through WLAN hotspot in Austria. Due to the disruption of the COVID-19 pandemic, Hong Kong recently moved to the use of online surveys from face-to-face interviews. Postcards with a QR code are distributed at airports to visitors with some incentives such as an e-coupon to boost the response rate. One issue with online surveys is that visitors might not choose to complete the questionnaire immediately. The time lag between their travel and the completion of the survey might cause some level of biases, especially in terms of questions related to expenditure which is

relatively difficult to recall correctly. Low response rate is another concern. Australia has previously trialled online surveys, but the response rate was not very satisfying. Meanwhile, some visitors might give an invalid email address or change their email address. It has also been found that students and respondents from certain countries are more willing to take part in the survey than others, which may cause representation issues. The response rate of the IVS survey in New Zealand was about 30%.

The approaches used in the UK along with 12 selected destinations to measuring visitor expenditure are summarised in Table 2.

Table 2. Summary of measurement approaches in selected destinations – visitor expenditure

Countries	Approach	Details of approach	Publishing schedule
UK	IPS	The IPS is conducted by the ONS at major airports and sea routes etc. through face-to-face interviews. Passengers are sampled as they enter or leave the UK following a multi-stage sampling design.	Monthly estimates are published 7 weeks after the end of the reporting month; more accurate and detailed quarterly estimates are published with a lead time of 3.5 months; annual estimates are published approximately 4 months after the end of the year.
Australia	IVS	The IVS is conducted face-to-face in the departure lounges of eight major international airports in Australia by Tourism Research Australia (TRA). Monthly samples of departing flights are selected first to target visitors to be interviewed.	The survey results are published on a quarterly basis.
New Zealand	IVS	Travelers are sampled from international visitors departing from Auckland, Wellington, Christchurch and Queenstown international airports. Departing visitors are asked if they are willing to take the survey. If eligible, the interviewer asks for their email address to send a link to the IVS.	The quarterly outputs from the analysis are released at 11 am on a predetermined date which can be found on the Ministry of Business, Innovation and Employment (MBIE) website.
Ireland	PCI (Passenger Card Inquiry)	The PCI surveys passengers at airports and seaports. The PCI card is distributed to all passengers on sampled flights and sailings.	The expenditure data are published at national level quarterly.
Japan	IVS	The IVS is conducted in departure lobbies at airports or seaports. Interviewers approach international visitors and ask them to reply to a set of questions while showing a tablet- or paper-based questionnaire.	The IVS results are published quarterly. First and secondary preliminary results are released at the end of the first and fourth months following the survey period, respectively. Final annual aggregates are published at the end of

			March in the following year.
South Korea	IVS	The IVS is a face-to-face or self-administered questionnaire survey that is conducted at airports and harbours on the last week every month.	The Balance of Payments data are published monthly by the Bank of Korea with a lead time of 2 months.
Austria	T-MONA survey	The T-MONA survey is exclusively electronically collected via various channels such as national tourism organisations, participating destinations and other partners.	Quarterly tourism receipts are published in the format of travel Balance of Payments.
France	EVE (L'Enquête auprès des Visiteurs venant de l'Étranger, Survey of Visitors from Abroad)	The survey tracks activity by foreign tourists in France by surveying tourists as they leave the metropolitan area. Vehicles are observed at border crossing points and more than 135,000 airline passengers are surveyed face-to-face when they embark.	The total inbound tourism receipts are published monthly and quarterly in the format of travel Balance of Payments. The breakdowns by source country are also available in the annual travel Balance of Payments.
EU-wide	Visitor surveys	The specific sampling and data collection method varies from country to country. Most EU states conduct face-to-face interviews to collect visitor expenditure data. National level data are then compiled and published in the form of travel receipts in the Balance of Payments.	The travel account of Balance of Payments is published by Eurostat on a quarterly basis.
USA	SIAT (Survey of International Air Travelers)	The survey is conducted on selected flights which have departed, or are about to depart, from the major U.S. international gateway airports, which are distributed either by flight attendants on-board or by contractors in the departing gate area. The number of flights sampled proportionately reflects the total number of flights offered by carriers and the dispersion of flights by destination.	Data are available on a quarterly and annual basis for either non-resident inbound or resident outbound.
South Africa	Departure Survey	The Departure surveys are conducted at airports and land border posts monthly through face-to-face interviews.	Tourism performance report is published quarterly, which provides a summary of the travel patterns of foreign tourists.
Hong Kong	DVS (Departing Visitor Survey)	The DVS is a face-to-face survey commissioned by the HKTB at departure gates at airports and piers. Appointed agents will approach visitors to ask if they are willing to take the survey. If the visitor agrees, then the agent will ask them to reply to questions shown on a tablet.	Annual visitor profile is published annually, normally in Quarter 2.

Saudi Arabia	Survey	The survey is conducted at the border exit points in the format of a face-to-face interview. The number of surveys distributed is based on the traffic in each exit point.	Monthly publications are planned to be made available in 2021.
--------------	--------	--	--

3.4 Alternative data sources

The tourism industry has experienced unprecedented challenges from the COVID-19 outbreak globally. Thus, the collection of tourism statistics has also been disrupted in most destinations. As surveys have been suspended in many destinations due to the lack of tourist flows and regulations such as social-distancing, alternative data sources have been explored to help analyse tourism statistics, mostly as complementary data. Three potential data sources have been mentioned frequently in the interviews with most of the destinations, including mobile positioning data, bank card transaction data, Google trends data and hotel booking data. But for most destinations, the exploration remains in an early stage, due to some common issues associated with each data source. The big data sources are perceived as having potential to reflect changing trends instead of real volume in the near future for most destinations. Some destinations, including New Zealand, France and South Korea, have already used these data sources practically. Details are presented in the following sections in terms of each data source.

3.4.1 Mobile positioning data

Mobile positioning data have been mentioned by many destinations as a potential data source that is worth further exploration. But for now, little useful output has been published for most destinations. One common issue with mobile positioning data is the representativeness of the users in a particular mobile phone operator company. Apart from mobile positioning data, the demographics of the users are also required to assess the sample data's representativeness and reliability, whereas that information is not always available (revealed in Australia and Ireland interviews). Another common issue mentioned frequently is that people may have multiple phones, while some people may not have one. In this case, the mobile positioning data again are not reliable as particular types of visitors/residents might be excluded. Finally, cross-border issues also influence the quality of mobile positioning data. For example, there might exist some level of noise near the land border of countries such as Austria. It is difficult to capture whether the phone user is within the border and outside the border of Austria. Thus, mobile positioning data should be more reliable and stable in destinations with a clear maritime border (e.g., UK and New Zealand).

Some of the destinations have started using mobile positioning data to analyse tourists' behaviour. For example, South Korea has used mobile positioning data to analyse local visitor behaviour. But the outputs need to be interpreted with caution according to the Korean big data lab. Data should be used for trend analysis as the output data of visitors are an estimated value instead of actual visits. The US International Trade Administration has used cellular pin data to produce drive travel data, which are available from the Travel

Industry Monitor. Mobile positioning data obtained from mobile network operators are used to extract trip itineraries, but these data have not been applied to estimate tourism statistics.

3.4.2 Bank card transaction data

Bank card transaction data are another data source that has been considered as a potential complementary data source for many destinations. For destination such as Australia, Saudi Arabia and Ireland, the use of card transaction data is still in an early stage, where the DMOs or statistical offices are exploring the potential to relate the transaction data to the tourist statistics. For destinations such as Austria, bank card transaction data provided by the central bank are fed into the Balance of Payments and TSA.

New Zealand appears to be one of the early adopters of bank card transaction data. This data source is combined with data from New Zealand TSA and IVS results to estimate regional tourism expenditure. Although the US has not used bank transaction data in their official statistics, Visit Florida has used the transaction data provided by VisaVue to calculate the number of overseas visitor arrivals to Florida by using credit card data, which offers a baseline for various countries and highlights the countries that they should be targeting. According to the interview, the VisaVue product would also show expenditure by commodity at a very detailed level, such as retail and jewellery, which provide another way to show the differentiated spending of international travellers at the state-level or regional level. Similarly, France used bank transaction data to monitor the differentiated effects of the pandemic on economic sections that usually welcome tourists and to analyse changes in residents' consumption during the pandemic.

But the use of bank transaction data still has many limitations, addressed by multiple destinations. The limitations are listed as follows:

- 1) Some visitors might have multiple cards, whereas some other visitors do not have bank cards.
- 2) There is considerable delay between the date of transaction and the date of consumption especially in accommodation and transportation.
- 3) Spending by foreign tourists is not able to be fully covered as they tend to possess bank cards from their country of residence.
- 4) It is hard to distinguish whether the card user is a tourist or not. Personal information is not available due to privacy issues.
- 5) The difference between a card user's citizenship and the issuing place of the card makes it difficult to correctly identify their nationality.

3.4.3 Google trends data

While post-visit data provide relatively accurate information about tourism demand, the tourism industry has strong needs for forecasting future demand. Forecasting based on historical data may not be able to capture changes in trend. As raised by some local DMOs, one possible solution is to complement post-visit data with big data sources such as forward bookings or Google trends. Google trends data have also been mentioned as an alternative source of data during the interviews with national DMOs. But so far, only France has already put the data into use by analysing potential tourists' search enquiries. As with mobile positioning data, the Google trends data provide some insight into travel and spending patterns. But one of the critical limitations of Google trends data is the search happens

before the visiting behaviour and thus it is more a proxy of visiting intention, rather than effective demand.

Tourism academics have started to explore the effectiveness of Google trends data on improving demand forecasting performance since 2012. As the pioneer study in this topic, Pan et al. (2012) revealed the superiority of tourism demand forecasting models with search enquiries. Li et al. (2020) show the effectiveness of using big data in improving forecasting accuracy based on multiple sources including search enquiries and online reviews. However, the superiority of Google trends data does not work on all occasions. Volchek et al. (2019) found that the models with search enquiries as an explanatory variable cannot always outperform the models without the variable in forecasting tourism demand.

To examine the feasibility of using Google trends data as an alternative data source to proxy tourism demand, the research team has conducted an explorative study to test the long-run relationship between actual visitor arrivals and Google search enquiries from the top 10 source markets of UK inbound tourism including France, the USA, Germany, Ireland, Spain, Italy, Belgium, the Netherlands, Switzerland and Canada. The Granger causality test (Granger, 1969) is used to investigate the predictability of Google trends data with respect to the actual tourism demand for the UK. In line with previous academic research, Granger causality has been found within the sampling period between the search enquiries from the USA, Germany and Italy and the actual visitor arrivals from those source markets. However, no such causality has been discovered for the other seven origin countries.

Although there is the potential to apply Google trends data to predict the tourism demand from a particular source market when the actual data are not available, caution is needed as only search enquiries from particular source markets have enough information to forecast the tourism demand. Further contributions from both academics and practitioners are needed to explore the potential and summarise the experience of Google trends data as an alternative data source for measuring effective tourism demand.

4 Case Studies

Australia

1. Visitor Arrivals and Departures

1.1 Data source

Administrative information on arrivals and departures is collected via various processing systems, including passport documents, visa information and incoming passenger cards. For international travellers (not including New Zealand citizens), additional information is required in visa applications and these data are collected by the Home Affairs department. All the administrative information is supplied to the Australian Bureau of Statistics (ABS) to compile overseas arrivals and departure (OAD). OAD are classified into permanent movement, long-term movement (i.e., with a duration of one year or more) and short-term movement (i.e., with a duration of less than one year), according to length of stay, as recorded by information on passenger cards, or derived with reference to previous border crossings.

1.2 Data sampling

OAD statistics are derived from a combination of full enumeration and sampling. All permanent arrivals, long-term arrivals and all departures have always been fully enumerated. Prior to April 2020, short-term arrivals were fully sampled as well, whereas only the total number of short-term travellers and their citizenships were fully known. Statistics on other characteristics (i.e., territory of stay/residence) were sampled through a fixed skip process based on a threshold by country of citizenship to minimise the effect of sampling error associated with the available statistics. The passenger card is selected and recorded once the threshold of the country was reached. Thus, countries of citizenship associated with a small number of movements were fully sampled, whereas countries with a large number of movements were sampled based on a higher threshold. However, since April 2020, a full sample has been undertaken for all OAD statistics due to the small volume of international travel movements caused by the impact of COVID-19.

1.3 Publishing schedules and format

In terms of publishing schedules, OAD statistics are released monthly, approximately six weeks after the end of the reference period. OAD data are available in many different formats on the ABS website, including the release of 'The Overseas Arrivals and Departures, Australia', and various outputs such as .pdf, ABS. stat datasets, spreadsheets and data files. Graphs on the monthly change to short-term arrivals and departures are available as well. More detailed or customised OAD data is available upon request to ABS Consultancy Services.

At the disaggregated level, visitor number/nights by purpose of visit are estimated using the results from the International Visitor Survey (IVS), which is published on a quarterly basis. Details on the IVS will be presented in the following section.

2. Visitor Expenditure

Other tourism statistics including visitor spend, as well as disaggregated visitor arrivals and visitor nights, are computed by the results from IVS conducted by Tourism Research Australia (TRA) and administrative data from the Home Affairs department. Detailed information on the IVS is presented as follows.

2.1 Survey collection and sampling

2.1.1 Survey collection procedure

According to the TRA website, the IVS is conducted by Computer Assisted Personal Interviewing (CAPI) in the departure lounges of the eight major international airports: Sydney, Melbourne, Brisbane, Cairns, Perth, Adelaide, Darwin and the Gold Coast.

The interview with the Australian Trade and Investment Commission (Austrade) has provided supplementary information on the procedure of survey data collection. TRA first collects data from the Department of Immigration on all departing flights. The flight details of the previous 12 months were used to model the sample in order to obtain some characteristics of the passengers travelling on different flights. This information is sent out to the interviewing teams at each of the airports around the country to identify targets (i.e., the type of person they want to interview) from targeted flights. After conducting face-to-face interviews, administrative data such as data collected from incoming passenger cards and passports scans will be matched to the information collected through interviews to create one complete record for each traveller.

Some other comments on the format of the IVS have been provided by Austrade. As tourism is a very complex industry with connections to all other industries, tourism statistics are difficult to measure. From that point of view, given that a universally agreed definition of a tourist or related tourism terms are not available, it would be difficult to define the phenomenon being measured in any accurate way without these customised face-to-face interviews. Some of the definitions and concepts are considered to be quite difficult, even for an interviewer who has been working for a couple of months to understand. Therefore, having a stable, well-trained interview team to conduct the interviews is crucial to the stability and quality of the data series.

2.1.2 Sample size/Coverage

The IVS samples 40,000 departing short-term international travellers aged 15 years and over who have been visiting Australia. The total number of interviews conducted with particular residents of each country or region is distributed among airports by selecting monthly samples of departing flights and visitors on those flights to achieve acceptable sample sizes in various categories. The response rate for the international survey is 60% to 70%.

2.1.3 Survey design

The IVS contains the following information.

- Usual place of residence
- Repeat visitation
- Group tours
- Travel party
- Sources for obtaining information about Australia
- Purpose of visit and places visited
- Transportation and accommodation

- Activities
- Expenditure

2.2 Data processing

Survey results are weighted according to the administrative data on international visitor arrivals over the period, provided by the Department of Immigration and Citizenship (DIAC), with the assistance of the Australian Bureau of Statistics (ABS). Variables used in weighting are country of residence, state of arrival, main purpose, airport of departure, age and sex.

2.3 Publishing schedule and format

The survey results are published on a quarterly basis. In addition to visitor spend, visitor number, visitor nights and disaggregated estimates by purpose of visit are published as well.

2.4 Changes due to COVID-19

Due to the outbreak of COVID-19, the IVS has been suspended since 1 April 2020. As an alternative, the TRA took trip characteristics from past IVS records with similar profiles to the current visitors based on information from the incoming passenger card to estimate values for a range of visitor characteristics from a small sample since 1 April 2020. According to the interview, the TRA also used the change in domestic travel behaviour as a sensitivity test for the international visitors, as they are exposed to the sample quarantine rules and regulations.

2.5 Alternative plans/future plans

According to the DMO, the incoming passenger card will be removed some time later in 2021. Alternative plans will be more similar to an administrative data source, rather than a sampling process. Besides, there will not be a major change in the IVS in the near future. The TRA will not consider using online surveys as a complementary survey source as in some other countries (e.g., New Zealand) in the near future. They have trialled the use of an online survey previously, but the results were not very satisfactory. For example, international students studying in Australia are more likely to give their contact details than a tourist. Meanwhile, it is considered that without the assistance of an interviewer, inconsistent understanding of the questions about tourism is inevitable even with explanatory notes.

New Zealand

1. Visitor Arrivals

Custom card data are compiled in International Travel and Migration statistics to count passengers arriving at and departing from New Zealand on a monthly basis².

2. Visitor Expenditure

Visitor expenditure data are measured using the IVS questionnaire in New Zealand. In addition to total annual expenditure, the survey provides data for determining the exports of travel services in the Balance of Payments, and tourism expenditure in the Tourism Satellite Account, measures the amount of expenditure per visitor for the top six key international markets (Australia, China, USA, UK, Germany and Japan). It also determines the activities international visitors participate in, the transport and accommodation types used, and places visited and provides demographic information about international visitors, their motivation for visiting New Zealand, and their satisfaction with their visit to New Zealand.

2.1 Sampling and data collection method

Travellers are sampled from international visitors departing from Auckland, Wellington, Christchurch and Queenstown international airports. The collection of IVS data is a two-step process. Firstly, departing visitors are screened by randomly asking them if they are willing to fill in the survey. Each person is provided with a brochure explaining the purpose of the IVS and benefits to New Zealand. If the individual agrees, the interviewer will check their eligibility and ask for some details including an email address. Secondly the IVS questionnaire is sent via email. The target sample size is 8,900 per year with an expected response rate of 30%.

2.2 Data processing

The sample is weighted to represent the survey's target population by the following criteria in four stages: ages group by airport, gender by age groups within country of residence, purpose of visit and length of stay bands within country of residence and totals for each of the top six country of residence markets. Detected outliers are handled using the Banff outlier treatment developed by Statistics Canada.

2.3 Publishing schedules and format

The quarterly outputs from the analysis are released at 11am on a predetermined date which can be found on the Ministry of Business, Innovation and Employment (MBIE) website, including an infographic on the release detailing market trend, a key data table detailing market trends in the past year, pivot tables detailing market trends over time, performance against pre-determined quality target and some microdata. The output also provides data for determining the exports of travel services in the Balance of Payments, and tourism expenditure in the TSA.

² Source: <https://www.mbie.govt.nz/immigration-and-tourism/tourism-research-and-data/tourism-data-overview/tourism-data-sources-diagram/>

2.4 Regional level data

According to the Ministry of Business, Innovation and Employment (MBIE, 2013)³, the Monthly Regional Tourism Estimates (MRTEs) provide an estimate of regional monthly expenditure on tourism from both international and domestic consumers. Three data sources are used to derive the Regional Tourism Estimates (RTEs), including the Regional Tourism Indicators (RTIs), Statistics New Zealand's TSA and IVS.

The RTIs use unit record electronic card transactions (ECT) data as the main source, obtained from Paymark via Marketview. The RTI data include all debit, credit and charge card transactions with New Zealand based merchants, all card-present transactions at the point of sale, whether authorised by PIN or signature and 15% Goods and Services Tax (GST). New Zealand is one of the pioneers that use card transaction data in compiling tourism statistics. TSA data are released by Statistics NZ in October each year, which provide the margin totals for deriving the RTEs by industry. The margins that are used from the IVS to derive the RTEs are the visitors' country of origin totals for tourism expenditure aggregated from the unit record data.

The Iterative Proportional Fitting (IPF) method is used to derive the RTEs in New Zealand. The IPF is more commonly known as the 'raking' technique. IPF is a procedure for adjusting a table of data cells so that they add up to selected totals for both the columns and rows (in the two-dimensional case) of the table. The method scales the cells in a contingency table formed from the RTI electronic card transactions unit record data so that their marginal totals equal totals estimated from the IVS and the TSA. Customised one-page regional summaries for each Regional Tourism Organisation (RTO) are published after the consultation with the RTOs.

2.5 Changes due to COVID-19⁴

The MRTE series has been disrupted during the COVID-19 outbreak, because the methodology for weighting the measured electronic card transaction spend up to the whole of industry spend became unusable. The Tourism Electronic Card Transactions (TECTs) have been established by MBIE as an interim replacement for the MRTEs, which aim to present the measured electronic card transactions (ECT) attributable to tourism but without any attempt to represent the total spend. This method is considered to be the best way to provide spend data to the industry while travel is severely affected by COVID-19 border restrictions. The MBIE has recommended to use the TECTs to look at trends and comparison figures and focus on the domestic market.

³ Source: <https://www.mbie.govt.nz/assets/02416ac5f3/regional-estimates-tourism-expenditure-method-concepts-results.pdf>

⁴ Source: <https://www.mbie.govt.nz/immigration-and-tourism/tourism-research-and-data/tourism-data-releases/tourism-electronic-card-transactions/>

Ireland

1. Visitor Arrivals

The visitor arrivals data are compiled from the Country of Residence Survey (CRS) conducted by the Central Statistics Office (CSO) in Ireland. Overseas trips to Ireland by area of residence and Irish residents' trips overseas are collected through the monthly survey. However, trips made by residents of Northern Ireland into or out of the Republic of Ireland and domestic travel within the Republic of Ireland are excluded.

1.1 Sampling and data collection method

1.1.1 Locations and collection of survey

The CRS survey is conducted at departure gates of Dublin, Cork, Shannon, Knock and Kerry airports and Rosslare, Ringaskiddy and Dublin seaports. Surveys are also conducted on behalf of the CSO at Holyhead Port by the UK's Office for National Statistics. Residency data are collected from passengers by a Central Statistics Office (CSO) interviewer. Detailed information on transportation is also collected from passengers' travel tickets and boarding passes. The survey is voluntary. The completed CRS forms are returned directly to the Tourism and Travel section of the CSO by tourist enumerators at the airports and seaports.

1.1.2 Sampling

A sample of sailings and flights is selected and a systematic sample of passengers on each of these is surveyed. On the selected sailings/flights, a 1-in-5 systematic sample of passengers is selected and their country of residence is recorded. Sample results are grossed up to total passenger numbers provided by the airports and ferry companies. However, cross-border data are not available.

The sample size is 267,000 passengers for Feb 2020 and seeks to cover all overseas travel into and out of Ireland by Irish residents and non-residents. Travel by residents of Northern Ireland into or out of the Republic of Ireland and domestic travel within the Republic of Ireland are excluded.

1.2 Data processing

Where there is no survey coverage in a particular time period for an airport/seaport pairing, results are estimated using the 'nearest neighbour' estimation methodology. Seasonally adjusted data are made available using the X-12-ARIMA model, developed by the U.S. Census Bureau, which is applied to the unadjusted data.

1.3 Publishing schedules

Provisional results are not published for the Country of Residence Survey and final results are published on a monthly basis within 28 days of the end of the reference period. The monthly Overseas Travel release is available on the CSO website. Local and regional level data are not available in Ireland. The survey results have been used externally by many organisations including Failte Ireland, Tourism Ireland, National Accounts, Balance of Payments, Eurostat, as well as by the government, economists and academics.

1.4 Changes due to COVID-19

The CRS survey has been suspended since March 2020 and there is no timeline for the resumption. No additional sources have been used to estimate related tourism statistics for this period.

2. Visitor Expenditure

2.1 Sampling and data collection method

Visitor expenditure data are estimated from the Passenger Card Inquiry (PCI) conducted by the CSO in Ireland.

2.1.1 Locations and collections of survey

The PCI surveys passengers at Dublin, Cork, Shannon and Knock airports and at Rosslare, Ringaskiddy and Dublin seaports. Data are collected from passengers who complete a questionnaire provided by one of a team of Central Statistics Office (CSO) interviewers. The ONS staff also collect data on behalf of CSO at Holyhead seaport covering passengers coming into the Republic of Ireland from Holyhead.

2.1.2 Data sampling

A sample of flights and sailings is selected to facilitate accurate representation of airport/seaport pairings (e.g., Dublin-London Heathrow), day, night, weekday and weekend flights/sailings. The PCI card is then distributed by CSO interviewers to all passengers on these flights and sailings. The population covered in the PCI survey are all passengers on overseas flights and sailings travelling into and out of Ireland. The effective sample size for the PCI was 38,070 passengers in Quarter 3 of 2019.

2.2 Data Processing

Where expenditure or fare data are missing, results are estimated using the 'nearest neighbour' estimation method. In other words, these expenditures and fares are estimated using the average fares of other observations in the sample with similar characteristics.

2.3 Publishing schedules and format

The data are published quarterly. In terms of Irish residents outbound travelling overseas, the total number of trips, average length of stay and expenditure are published for all Irish residents. For inbound international travel, the total number of trips, average length of stay and expenditure are published by the passengers' country of residence. Regional-level breakdown is not available.

2.4 Alternative sources/future plans

Due to the outbreak of COVID-19, the PCI survey has been suspended since March 2020, the same as the CRS survey. According to the interview, no alternative data sources have been used to address this tourist statistics compilation issue. No changes regarding the survey collection method and survey design are foreseen by the CSO. However, it is recognised that there are some disadvantages with the current method. A major gap mentioned during the interview is related to tourists coming from Northern Ireland across the border, as the CSO has no measurement of these cross-border flows. There is some level of collaboration between the tourism statistical offices in the island of Ireland. From Ireland, the CSO shares data quarterly with the Northern Ireland Statistics & Research Agency (NISRA) to provide some estimates of Northern Ireland residents who take the survey, although those statistics are excluded in the published national statistics in Ireland.

The CSO in Ireland has investigated some alternative data sources. Even before COVID-19, the CSO has recognised the need to modernise the process by considering moving away from this interviewer related survey-based data collection at the airports and seaports towards using mobile devices to capture information. But the common issues associated with mobile data persist. For example, most mobile network companies are reluctant to provide the data for free. The second issue is related to representativeness. For a particular mobile company, besides the mobile data usage, the demographic distributions of the users

are needed to assess the representativeness of the users. The CSO has also mentioned the credit card transaction data and bookings data, but limited further information is available. For now, future plans are more likely to centre on a blended approach, by using the alternative sources as supplementary data, or as a calibration tool rather than direct use of such information.

Tourism Northern Ireland is looking into using one of the credit card companies' details on spend in Northern Ireland by international visitors. They are also examining the mobile phone network, which will give an indication on trends and probably will be more reliable for assessing movement within Northern Ireland instead of measuring the volume of visitors. A potential challenge is related to the cross-border issue. When people stay in places near the border between Northern Ireland and the Republic of Ireland, the mobile positioning data are not reliable to capture which side of the border they are currently in.

Japan

1. Visitor Arrivals

The visitor arrivals data are collected by the Ministry of Justice. The population covers inbound international visitors referring to foreign visitors entering Japan. This group consists of foreign nationals legally entering Japan minus foreign nationals who are residents of Japan (both according to the nationality-based statistics published by the Ministry of Justice), plus those in transit. The visitor arrivals data are available monthly.

An additional 'Overnight Travel Statistics Survey' is conducted by the Japan Tourism Agency (JTA), which collects the number of stays of foreign tourists in each prefecture on a monthly basis, according to the E-mail interview with the JTA. The breakdown of the total number of overnight guests by place of residence in each month (by prefecture or outside the prefecture) and the breakdown of the total number of foreign guests staying in each month by nationality are also estimated from the survey results. Private contractors are responsible to distribute the survey by mail or online to the reporters and send the completed survey back to the JTA. Monthly reports will be available by the middle of the following month. A white paper on Tourism in Japan is published annually, presenting global tourism trends and tourism trends in Japan (i.e., international visitors travelling to Japan, Japanese overseas/domestic travel, overnight trips and tourism trends in regional areas).

2. Visitor Expenditure

2.1 Sampling and Data collection method

The IVS is aimed at obtaining baseline information for designing and evaluating measures to attract foreign tourists by assessing trends in consumption by international visitors to Japan. The IVS is made up of three quarterly surveys for different purposes. The national survey is designed to characterise international visitors to Japan as a whole by their attributes as well as by what they did and how much they spent in Japan. The regional survey is designed to characterise international visitors to each prefecture in Japan by their attributes as well as by what they did and how much they spent in Japan. The cruise survey is designed to characterise international visitors to Japan with landing permission for cruise ship tourists by their attributes as well as by what they did and how much they spent in Japan.

The IVS is conducted in departure lobbies at airports or seaports. Interviewers approach international visitors and ask these visitors to reply to a set of questions while showing a tablet or paper-based questionnaire in an interview style, whereby the answers are recorded by the interviewer.

The target sample size is 34,964 in total for each quarter, with 7,830 responses for the national survey, 26,174 responses for the regional survey and 960 for the cruise survey (an annual total of 139,856).

2.3 Publishing schedules and format

The IVS results are published quarterly. The first and secondary preliminary results are released at the end of the first and fourth months following the survey period, respectively. Final annual aggregates are published at the end of March in the following year.

South Korea

1. Visitor Arrivals

Statistics of visitor arrivals, Korean departures and cruise visitors are aggregated and compiled by Ministry of Justice, Korea Immigration Service according to the Korean Tourism Statistics Guidebook published by the Korean Tourism Organisation (KTO).

The tourism statistics in South Korea are published monthly. Final figures of the Tourism Balance of each year are released before the end of the following year (figures are replaced with provisional figures before the release of final figures).

2. Visitor Expenditure

2.1 Sampling and data collection methods

Visitor spending data are estimated from the IVS, which aims to provide statistical information that can be widely used for policy making, promoting tourist destinations, and improving tourism readiness through analysis, tracking travel trends and an overall evaluation of traveling to Korea. The IVS is a face-to-face or self-administered questionnaire survey that is conducted at airports and harbours on the last week of every month. The sample size was 16,076 persons in 2019 (1,340 persons per month). The sample population is international visitors of foreign nationality departing from Korea. The survey includes questions in terms of the context of the trip (e.g., number of visit to Korea; travel companions; countries visited before/after; main purpose), patterns of behaviour during the trip (e.g., main sources of tour information during travel; activities in Korea; visited sites; and length of stay), expenditure, evaluation of the trip and others (e.g., use of smartphones during travel).

2.2 Data Processing

A valid sample was extracted using the modified proportional distribution method, and then post-weighted based on the foreign tourist entrance by country, month, gender and ages.

2.3 Publishing schedules

The Balance of Payments data are published monthly by the Bank of Korea, where travel is the second item in the services category of the current account.

2.4 Alternative sources

Mobile communication data are an alternative data source that KTO has used to analyse local visitor movement behaviour and the number of visitors and increase/decrease rate by local government. Visitors are classified into different types by the use of mobile positioning data. Based on the relevant metropolitan local government area, the visitors are classified into resident (local) and foreign visitors. Based on specific locations, visitors are analysed in terms of visitors to tourist destinations and visitors to festivities.

However, it is considered that the output from the mobile communication data should be interpreted with caution. To begin with, the number of visitors is an estimated value instead of the actual number. Thus, it is advisable to use this for trend analysis, such as an increase or decrease rate. In particular, the number of visitors derived from mobile communication data tends to be underestimated or overestimated for several reasons. Firstly, the algorithm and data extraction process may vary depending on the mobile communication method (2G,

3G, 4G, 5G, etc.) of the data extraction period and changes in the base station radio wave area. Secondly, visitors passing through the area such as on expressways, roads, and railroads can be counted. There are some common issues with mobile positioning data that have been noted in other countries, such as miscounting of individuals who do not have a mobile phone or have multiple phones. In the case of foreigners, only foreigners using roaming service can be counted. Secondly the exact purpose of the visit is unknown. Meanwhile, the number of visitors is enumerated on the basis of the number of unique visits per day. Therefore, for visitors who stay more than one night in a tourist destination, the number of those visits would be counted as multiple different visitors. Finally, the local level data cannot be aggregated for the 11 basic local governments arbitrarily. Thus, higher level data cannot be retrieved from mobile communication data.

Austria

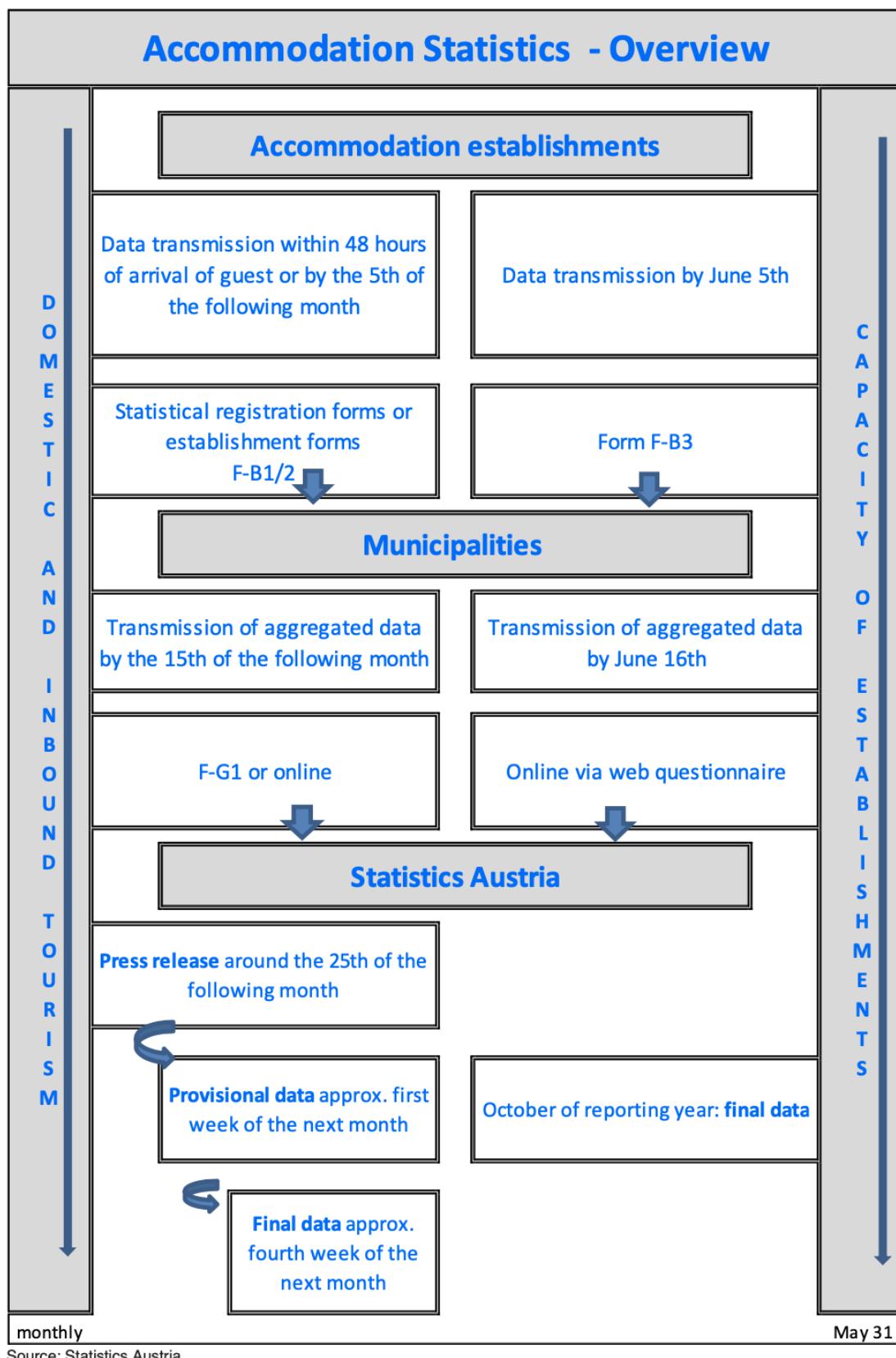
1. Visitor Arrivals

1.1 Data source

Monthly accommodation statistics are compiled by Statistics Austria as a measurement of tourism demand. Annual accommodation capacity statistics are also compiled annually as a measure of tourism supply. Monthly accommodation statistics collect data on arrivals and nights spent in Austria by type of accommodation establishment and by country of origin, whereas annual accommodation capacity statistics record number of accommodation establishments and bed capacity only by type of accommodation establishment.

The basis of the two indicators are the reports from 1,558 (in 2017) municipalities. Those reporting municipalities themselves receive the data from accommodation establishments (64,478 in 2017) located in their respective municipalities. The aggregated municipal results are then transmitted to Statistics Austria. The whole procedure is compulsory for both accommodation establishments and municipalities. Figure 1 presents the detailed procedure of data collection. This whole procedure is carried on throughout the year.

As the data collection is through a bottom-up approach, municipal level data are available because municipalities are responsible for collecting accommodation statistics from the accommodation establishments. No sampling adjustments have been made regarding the accommodation statistics. The sample size for the nine federal states in Austria are sufficiently reliable, whereas there are insufficient sample size issues at NUTS III level. However, according to the interview with Central Statistics Office in Austria, those destinations could possibly conduct additional surveys by paying extra fees.



Source: Statistics Austria

Figure 1 Accommodation statistics in Austria

1.2 Publishing schedules

Accommodation statistics are available monthly. The preliminary results are due to be available 24 days after the reference period, whereas the final results should be available in a 50-day lag period. The final results of accommodation compacity statistics are reported at the end of the calendar year, in terms of the tourism year, the winter season and the summer season. Data in smaller tourism regions (i.e., NUTS III level) are not published due to small sample sizes.

1.3 Lessons learned from the interview

Some issues or biases have been brought up during the interview with Statistics Austria related to monthly accommodation statistics. Accommodation establishments other than hotels such as Airbnb are also legally required to collect and send data to the municipalities, at least in Vienna. Thus the quality of data should be reliable. However, while the majority of visitor arrivals should be captured, there are gaps, e.g., same-day visitors, visitors staying in informal accommodation such as staying with friends and family, visitors staying in rental properties that are not legally registered. Another issue is related to multiple counting. Tourists who visit multiple regions or stay in multiple hotels would be counted as multiple visits. No further adjustments have been made to correct this bias due to difficulties in developing a reliable estimation method. But in Austria this issue is not regarded as very significant because Vienna is the major destination for long-haul tourists. Tourists from nearby European countries, such as Germany and Netherlands, tend to stay in a single location, for example, during the skiing season.

2. Visitor Expenditure

Visitor Expenditure is compiled by Austria Tourism through the Tourism Monitory (T-MONA) Austria survey which includes extensive information about the travel behaviour of tourists in Austria. According to the interview, the main purpose of the survey is to evaluate satisfaction, whereas spending is only a small element at the end of the survey.

2.1 Survey collection and sampling

2.1.1 Survey collection procedure

The T-MONA survey in Austria is exclusively electronically collected since 2017/18 via various b2c contact channels such as national tourism organisations, participating destinations and other partners. Before 2017, the survey was carried out through self-competed paper questionnaires accompanied by interviewers. But all face-to-face interviews have been suspended due to the costs. The survey is also available via public WLAN hotspots at locations relevant to tourism throughout Austria and via social media channels (e.g., Facebook and Instagram), various print products and newsletters. The survey is designed to include questions on the following sections:

- Tourist structure (country of origin, age, gender, occupation, education, income, visiting experience, travel companion)
- Motivation, information, decision-making and booking behaviour before the trip
- Travel behaviour on-site (arrivals, types and activities of vacation, accommodation, travel duration, vacation expenses)
- Satisfaction and recommendations
- Image

2.1.2 Sample size/Coverage

The target population is holiday visitors who are currently on vacation at the time of the survey (at least one night) and past visitors who have departed within the last two months. Expenditure data are stratified according to country of origin, based on the monthly accommodation statistics.

2.2 Publishing schedule and format

The data on inbound tourism receipts (excluding international passenger transport) are published quarterly in the format of travel balance of payments. An online interactive database is available to browse the data.

2.3 Alternative plans/future plans

In terms of the survey quality, as the questionnaire is mainly designed to measure tourist satisfaction in Austria, with questions related to expenditure at the end of a long survey, people may be reluctant to answer questions concerning expenditure. Also, the timing when the tourists answer the questionnaire will affect the quality of the expenditure data. Since they will be given a link to fill out the survey, if they choose to complete the survey some time after the trip, then the data is likely to be less accurate than if collected at the point of departure. Therefore, there are concerns about the accuracy of the expenditure data.

Bank card transactions are an alternative data source that is being considered by the CSO in Austria. The credit and debit card transaction data are received from the Austria Central Bank at the national level to compile the Balance of Payments, with both inbound international visitors and outbound Austrian travellers included. But there still exist some common issues related to credit card transaction data such as multiple cards within a household or differentiation between business card and personal card. No adjustments were made to accommodate these issues. Mobile positioning data are also considered to measure outbound visits by Austrian residents. Due to contractual issues with the mobile operators, inbound visits are not taken into account in terms of mobile positions. There also exist some noise around the border area in terms of the mobile positioning data. Changes or alternative plans in compiling visitor arrivals are not considered in the near future according to Statistics Austria.

France

1. Visitor Arrivals

1.1 Data source

The visitor arrivals data are compiled through the monthly attendance survey in tourist accommodation by the National Institute of Statistics and Economic Studies (INSEE), the national statistics bureau of France. The survey is conducted by sending a mail or email to tourist accommodations at the beginning of month M+1 for the data of month M. The main topics included in the survey are characteristics of the establishment, reception capacities, daily occupancy of rooms, arrivals and overnight stays according to the country of residence of the tourists, duration of stay and proportion of business tourism. The survey is carried out each month, from January to December for three types of establishments.

1.2 Sampling

Tourist accommodations are categorised into hotels, campsites and other collective accommodations. Each category has its own sampling frame and stratification criteria. Under each stratum, the draw of sample is random.

In terms of hotels, the sampling frame is the index of tourism hotels. The stratification criteria of the national sample are the geographical axis, hotel category, and the type of hotel. The sampling frame of camping is the file of establishments offered at least at one location. In small regions, the survey is comprehensive. For other regions, the stratification criteria are the geographical axis, campsites category and the size in number of locations. For other collective tourist accommodations, the sampling frame is the file of other collective accommodation. The stratification criteria are the geographical axis and the type of establishment.

1.3 Publishing schedule

The provisional data dissemination is released at the beginning of the month M+2 for the month M. Final data dissemination is released at the end of the month M+3 for the month M. Publications include a 'information rapides' at a quarterly basis. A summer season report is published in December and a winter season report is published in June. An annual report is published in April of the following year.

1.4 Changes due to COVID-19

In the context of COVID-19, lean operations were put in place in spring 2020 (April, May, June) and from autumn 2020 onward. It intended that they will be maintained for as long as necessary. This involves retrieving, at a minimum, information on the number of open establishments, the number of rooms offered, the number of rooms occupied and the number of total nights.

2. Visitor Expenditure

2.1 Sampling and data collection method

The visitor expenditure data are estimated from the EVE survey of visitors from abroad, which is conducted by the Directorate General for Enterprise (DGE) and the Banque de France. The survey tracks activities by foreign tourists in France by surveying tourists as they leave metropolitan France, irrespective of their form of transportation – whether by

road, air, rail or sea. Vehicles are observed at border crossing points and more than 135,000 airline passengers are surveyed face-to-face when they embark. The survey is also administered to 60,000 non-residents to identify key characteristics of their stays, such as length, reason and expenditures.

2.3 Publishing schedules and format

The total inbound tourism receipts are published monthly and quarterly in the format of travel Balance of Payments. The breakdowns by source country are also available in the annual travel Balance of Payments.

2.2 Alternative sources

Due to the COVID-19 pandemic which has significantly influenced the tourism industry globally, some alternative sources are used to analyse domestic tourist activity in summer 2020, which are presented as follows.

Tourist Demand Survey

The survey monitoring tourist demand (SDT) has been conducted monthly with 22,000 participants aged 15 or older, which provides information consistent with the European regulation on tourism statistics: defining tourist trips by residents (personal or professional, departure rate, number of trips, number of nights spent, tourist spending, etc.).

CB bank card transactions

Bank transaction data, which were extracted from the anonymised and aggregated transactions with coverage of the majority of transactions by residents, have been used to monitor the differentiated effects of the pandemic on economic sections that usually welcome tourists. Departmental data were produced based on the location of the establishment where the transaction took place and distanced transactions (i.e., internet transactions) were not taken into account. These transaction data were used to analyse changes in residents' consumption during the pandemic.

There are some limitations regarding bank card transactions data. The first important limitation is the considerable delay between the date of the transaction and the date of consumption for some sectors such as accommodation and transportation. Secondly, people may have changed the way of using bank cards during this pandemic. Thus, changes in residents' consumption behaviour only partially reflect changes in the turnover of the concerned establishments. Finally, spending by foreign tourists is not covered by this mode of payment whereas foreign visitors represented 38% of total international tourist consumption in previous years (although, of course, much lower during the pandemic).

Online searches monitored by Google Trends

Google Trends data are used to analyse users' interest in different search enquiries through calculation of a weekly aggregated score from a representative and filtered sample of searches. This represents the proportion of searches for a given key word in a region and for a specific period, compared to the data when the usage rate of this key word is the highest. Although, normally, Google Trends data cannot improve forecasts of household spending significantly, they do provide some insight into spending patterns. Another advantage of the Google Trends data is the quicker availability of the data even for countries other than France.

2.3 PiLOT project

The PiLOT project is an online platform for DMOs and territory performance benchmarking instrument for the local level destinations. Two key indicators are developed including key Socio-Economic Indicators (ISEC) which measure the financial performance of local

destinations, and Tourism Potential Index (ITT). The data collection procedure follows a bottom-up approach, in which local DMOs are responsible for collecting related data and feeding them into the PilOT platform. The quality of the data is controlled by experts before the data are recorded into the system. Four dimensions of data have been covered, including stay, daytrip, accessibility and attractiveness. At the time of writing, around 25% of local destinations in France have participated in this project. Although the main purpose of the project is local destination benchmarking rather than harvesting of tourist statistics, the data collection procedure can provide some insights into local level tourist data.

European Union

1. Visitor Arrivals

1.1 Data source

The data of visitor arrivals published by the Eurostat are mainly compiled using accommodation statistics, which is a key part of the system of tourism statistics in the EU and has a long history of data collection. Arrivals are registered by country of residence of the guest and by month when a tourist stays at a tourist accommodation establishment. The reporting unit of each country collects arrivals statistics from the accommodation establishments and then transmits these to the national statistical institutes.

1.2 Publishing schedules and format

In the Regulation concerning European statistics on tourism, transmission deadlines are as follows:

- Monthly data
 - 8 weeks: nights spent (national level)
 - 3 months: nights spent, arrivals, net occupancy rates (national level)
- Annual data
 - 6 months: establishments, bed places, arrivals etc. (national level)
 - 9 months: nights spent in non-rented accommodation (national level)

2. Visitor Expenditure

2.1 Sampling

Different European destinations applied different sampling frames to collect survey data. The following type of sources for the frame population were reported for the European countries:

- Population register, central registration register
- Census of population and housing
- Register of census districts
- Telephone directory
- Register of electors
- Passengers in airports and ports
- Border records

In terms of sampling design, there are major differences among national statistical institutes. Sampling designs used across Europe include:

- Stratified simple random sampling
- Two-stage cluster sampling
- Three-stage cluster sampling
- Stratified systematic sample
- Stratified sample of port/a port/airport routes, systematic sample of passengers (or cluster sampling at gate or arrival belt)
- Two-stage area sampling
- Quota sampling, and others

Given the different sources used, the sample sizes are not directly comparable across countries. In 2008, the sample size ranged from 1200 persons in a quarterly survey to over 300,000 persons in a border survey over a reference year.

2.2 Data collection method

The main sources for tourism expenditure data are surveys. The visitor surveys are often in the form of border surveys or departure/entry information, and household surveys are used to measure domestic and outbound tourism. In EU states, the tourism demand surveys are mandatory.

Face-to-face interviews, telephone interviews, postal surveys and web surveys are possible techniques used to collect data for outbound and domestic tourism. Face-to-face interviews and telephone interviews are preferred to postal surveys because interaction between the respondent and the interviewer can be important for understanding the questions. Postal surveys may be used to lower survey costs when budgets are limited. They can also give better possibilities to include clear explanatory notes for the respondents. Table 3 presents the number of countries using each method of data collection. It can be seen that most countries used a face-to-face interview or telephone interview to collect data.

Table 3 Data collection methods

Means of data collection	Number of countries
Postal questionnaire	4
Face-to-face interview	11
PAPI ⁵	3
CATI	13
CAPI+CATI	6
CAWI	1
CAPI	1
Others/Unknown	5

2.3 Publishing schedule and format

According to Directive 95/57/EC (Art. 7), Member States should transmit provisional quarterly data within three months of the end of the reference period and final data within six months of the end of the reference period. For the annual data, this is six months and twelve months for provisional and final data respectively. A majority of the European countries (26 surveys) used tourism demand surveys with a reference period of three months (one quarter). This choice of a 3 months' reference period is a trade-off between recall bias and collection or organization costs. Another 10 surveys had a reference period of one month and, in three cases, an annual reference period was used. For border surveys, the reference period is not relevant since the interview refers to the trip that starts or ends at the time of the interview. Across Europe, major differences can be observed in terms of timetable and timelines of data collection and transmission. The time between the end of the reference

⁵ PAPI: Pen-and-Paper Personal Interviews; CATI: Computer Assisted Telephone Interviews; CAPI: Computer Assisted Personal Interviews; CAWI: Computer Assisted Web Interviews

period and the end of the data validation ranges from 20 days to 7 months. The time between the end of the reference period and the first publication of the data ranges from 40 days to 7 months.

Apart from the statistics, Member States should provide Eurostat with a report on the quality of the data relating to the reference periods in the reference year and on any methodological changes that have been made every year. The report should be provided within 9 months after the end of reference period.

Inbound tourist expenditure data are compiled and published in the form of travel receipts in the Balance of Payments. Services are the second major category of the current account in the Balance of Payments, which includes three sub-items: transportation, travel, and other services. The credit side of the item "travel" includes purchases made by foreign travellers on the national territory. This item contains both business travel and personal travel. The costs of international transportation to the destination are recorded under the item "transportation", but all movements within the country, including cruises, are reported under "travel".

USA

1. Visitor Arrivals

1.1 Data source

Inbound visitor arrivals are collected via the I-94 arrivals programme, which consists of three parts: Department of Homeland Security (DHS)/U.S. Customs and Border Protection (CBP), Statistics Canada's International Travel survey and the Instituto Nacional de Estadística y Geografía/INEGI (Banco de Mexico) because travellers from Canada and Mexico are basically exempted from the border system, according to the interviewee in the National Travel and Tourism Office (NTTO). NTTO manages the ADIS/I-94 arrivals programme (I-94) in cooperation with the Department of Homeland Security DHS/CBP utilising a variety of systems to create a unique person-centric record with complete travel history. A complete I-94 arrivals record includes information transmitted to CBP using Advanced Passenger Information System (APIS) (i.e., biological data page of travel documents and information collected via the passenger reservation and check-in processes), information provided by an air or sea carrier about the flight or voyage and visa information by the U.S. Department of State via the Consular Consolidated Database (CCD). The detailed data processing procedure (according to the interview) is as follows:

Subtotal – Overseas Visitors to the US

- Total records received from DHS/CBP
- Substitute country of citizenship (COC) for any records missing a valid entry in the country of residence field
- Remove duplicate records
- Remove zero-night stays
- Remove Mexico residents (land entry)
- Remove “non-visitors” based on visa type (DHS’s “class of admission”)
- Remove duplicate entries by land
- Remove Mexico residents (air, sea, not reported entry)
- Remove Canadian residents

Total International Visitors to the US

- Add Mexican Visitors to the US (Banco de Mexico)
- Add Canadian Visitors to the US (Stats Canada)

1.2 Sampling

The US follows the United Nations World Tourism organization (UNWTO) guideline to gather arrivals data based on country of residence (COR). The UNWTO's definition of an overnight visitor is any traveller taking a trip to a main destination outside their residence, for at least one night - but less than one year, and for any main purpose (business, pleasure, or other personal purpose) other than to be employed in the country or place visited. Thus to calculate visitation to the US, filters include Visa Class of Admission, One or More Nights and COR. In 2020, DHS/CBP provided NTTO with 547,047,790 ADIS/I-94 raw records of which only 8,684,990 records represented travel from overseas countries (all modes) and Mexico air and sea.

1.3 Publishing schedules and format

Summary data of visitor volume by both COR and COC are posted monthly to trade.gov. The reason of having both sets of visitor volume data is that the monthly volume reporting was challenging because their source files increasingly missed a valid code in the COR field, caused by the base of country of origin on residence. According to the interviewee, destination-level data are available mostly in percentages when the sample size of that particular destination reaches 100. But full destination-level data are available upon request. The Summary of International Travel to the United States report is available by monthly, quarterly, or annual fee-based subscriptions.

Categories published based on both COR and COC include month of arrivals with percent change, percent change comparisons year over year, market shares, type of visa, mode of transportation, US port of entry and limited Mexican and Canadian visitor statistics. Annual publications are released no later than March of the calendar year.

2. Visitor Expenditure

Visitor expenditure is mainly estimated using the Survey of International Air Travelers (SIAT), which is an on-going primary research programme that gathers statistical data about air passengers in the US - overseas and US - Mexican markets. The survey includes information on passenger trip planning, travel patterns, demographics and spending for non-US residents traveling to the US and for US residents traveling from the US.

2.1 Data collection and sampling

2.1.1 Survey collection procedure

The Survey was conducted on selected flights which have departed, or are about to depart, from major US international gateway airports, which are distributed either by flight attendants during the flight or by contractors in the departing gate area. The survey is self-administered by passengers who volunteer to take it. The contractor processes the return of the survey kits, which include verifying receipt of survey, coding, data entry and editing of the completed questionnaires.

2.1.2 Sample size and coverage

All international air travellers to or from the US (except Canadians) are covered. A computer programme selects a random sample of flights to be surveyed from an electronic database of scheduled airline flights. The number of flights sampled proportionately reflects the total number of flights offered by the carriers and the dispersion of flights by destination. According to the interviewee, the data are sampled roughly in proportion across two variables: country of origin and port of entry.

2.2 Data processing

The Survey data are “weighted” to the census data. For example, non-resident inbound survey responses are weighted to the “100%” population of DHS I-94 arrival records to adjust for over and under sampling.

2.3 Publishing schedule and format

Data are available on a quarterly and annual basis for either non-resident inbound or resident outbound. Twelve country reports are published for inbound data. State and city reports (including 12 states and top 12 cities) are available.

2.4 Changes due to COVID-19

Survey interviews were difficult to collect. Only seven interviews have been completed so far.

2.5 Alternative plans/future plans

2.5.1 Cellular pin data

Cellular pin data were mentioned in the interview, as a potential alternative source to produce car-based travel data, which are available from the Travel Industry Monitor published by International Trade Administration. Mobile positioning data are available from the phone operator companies and thus individuals' trip itinerary can be extracted from tracking the change of mobile position. Whether the user is a US resident or not can also be interpolated based on their usual mobile positions. But limited details are available about the integration of cellular pin data and other tourism statistics⁶.

2.5.2 Credit card transaction data (VisaVue)

Credit card transaction data have been considered. Credit card companies would sell subscriptions of a product called VisaVue to destinations. Although the Bureau of Economics at the federal level has been looking into the data, not much output has been generated. According to Donna Keren, the Vice-President of Research NYC & Company, they have purchased VisaVue for three years now and many analyses of the data and comparisons to other data including SIAT have been undertaken. But to date, they have found no relationship between VisaVue data and any other data series they track pertaining to visitor volume and visitor spending. On the other hand, Visit Florida has made some use of the VisaVue data, but not for the measurement of tourism statistics such as tourist arrivals and expenditure. Visit Florida used VisaVue data to calculate the number of overseas visitors to Florida by using credit card data, which offered a baseline from various countries and highlights the countries that they should be targeting. But overseas transaction data are not always available for some specific countries in South America and Europe because of the different bank card companies operating in those regions. According to the interview, the VisaVue product should be more useful to estimate the volume by country of origin than for estimating spending. The VisaVue product would also show data by commodity at a very detailed level, such as retail and jewellery. Thus, the VisaVue data provide another way to show the spending differentiation of international travellers in the state-level or regional level.

⁶ Source: <https://www.trade.gov/data-visualization/covid-19-travel-industry-monitor>

South Africa

1. Visitor arrivals

1.1 Data source

The only comprehensive source of information on foreign arrivals and departures, documented immigrants and emigrants are the administrative data sources from the Department of Home Affairs (DHA) in South Africa.

Data are routinely collected by immigration officers at all road, air and sea entry ports on all travellers (South African residents and foreign travellers) arriving at or departing from South Africa directly through their travel documents either by scanning or capturing these onto the ports' electronic database. Individual ports regularly transmit data onto the national database at the head office of the DHA.

1.2 Publishing schedules and format

The statistical release of the Tourism and Migration Report is scheduled with an advanced release calendar available on the Statistics South Africa website, which includes reports on tourist arrivals and departures as well as transit of South African residents and foreign travellers. Detailed information is also provided on the age and sex distribution, mode of travel, national and regional distribution, and purpose of travel of overnight visitors or tourists.

2. Visitor Expenditure

The Departure Survey is conducted to capture the travel patterns of international visitors in South Africa.

2.1 Sampling and data collection method

The Departure survey is conducted monthly at airports and land border posts through a face-to-face interview assisted by the use of tablets (i.e., tablet-assisted personal interviews). At airports, interviewers will approach visitors in the seating area. Stratified probability sampling is used, with random 1-in-5 systematic sampling within each stratum. At land border posts, interviewers will intercept and randomly sample respondents across three interception points using the 1-in-5 sampling method. The sample size at airports is 3,800 per month and the sample size at land border posts is 1,000 per month.

2.2 Survey design

The Departure Survey includes questions in the following critical areas:

- Reasons for visiting South Africa
- Time spent in South Africa
- Attractions visited when in South Africa
- Expenditure patterns while in South Africa
- Decision-making process to visit South Africa
- Arrangements during travel
- Overall impression of their visit to South Africa
- Travel to other African countries
- Personal demographic information

2.3 Data processing

The results are subsequently statistically weighted up to the tourist arrivals by air and road data as released by Stats SA, in a given time period. This allows continuous tracking into expenditure patterns, accommodation usage, travel behaviour, experiences, purchasing and travel patterns.

2.4 Publishing schedules and format

The tourism performance report is published quarterly, which provides a summary of travel patterns of foreign visitors in South Africa. Tourism statistics on international tourist arrivals by major markets, purpose of visit, total foreign direct spending, average spending, and so on are reported at the national level. Regional breakdown of those statistics by province is also available in the quarterly reports.

2.5 Alternative data sources

The current data collection platforms have several advantages. The comprehensive survey covers a range of areas as listed in Section 2.2. The setup of the whole procedure is relatively easy. The completeness of the surveys is generally high and the non-sampling errors are minimal, as the stratified probability sampling ensures a representative sampling frame. However, the disadvantage is that the methodology is costlier due to high involvement process.

Some secondary sources are considered to supplement the primary data obtained from the Departure Survey. But there are some limitations to these sources. Data from Department of Home Affairs may be difficult to obtain, and they only answer some of the questions listed in Section 2.2. Data from banks only contain spending information, which is likely to have high quality. However, those spending data will only be as accurate as the numbers of tourists who use bank cards for their purchases. Data from accommodation establishments have also been considered. But inaccuracies occur when visitors stay in multiple accommodations within the same trip.

Hong Kong

1. Visitor Arrivals

1.1 Data source

The visitor arrivals data are collected at the borders by immigration officers and then transferred to the Hong Kong Tourism Board (HKTB). Visitors by origin are compiled by the arrival cards using systematic sampling method.

1.2 Publishing schedules

The collection of visitor arrivals was not interrupted greatly during the COVID-19 pandemic. The arrivals data are published monthly, with a month lag after the reference period. Publishing contents include visitor arrivals by country/region of residence and mode of transport.

2. Visitor Expenditure

2.1 Sampling and data collection

The HKTB Departing Visitor Research Project is conducted throughout the year to collect visitor spending data, travel details and activities and satisfaction. Target respondents are overseas visitors aged 16 or above departing from ten major control points in Hong Kong. Appointed agents will approach visitors to ask if they are willing to take the Departing Visitor Survey (DVS). If the visitor agrees, then the agent will ask him or her to reply to questions showing on a tablet, according to the interview with the HKTB. The DVS includes two separate surveys, a Travel Pattern Survey (TPS), which is a shorter survey with an annual sample size of 204,000, whereas the Main Visitor Survey (MVS) is more comprehensive and normally will last around 20 minutes, with an annual sample size of 57,000. In terms of the sampling approach, the TPS adopts systematic random sampling to study the visitors' main purpose of travelling to Hong Kong. The MVS employs disproportionate stratified quota sampling to collect data, including visitors' travel itineraries, spending, satisfaction of their trip, and so on.

2.2 Publishing schedules and format

Two major publications are produced regularly by the HKTB. Tourism Expenditure Associated to Inbound Tourism is published bi-annually, including total tourism expenditure associated to inbound tourism, visitor spending by major market areas, visitor spending by spending category and visitor arrivals. The visitor profile report is published annually (mostly in Quarter 2), including analysis on demographic data, travel details and activities, visitor spending details and satisfaction indicators.

Local level data are not an issue for Hong Kong, as Hong Kong is a city instead of a country. Some specific attractions, such as Hong Kong Disneyland and Ocean Park, have their own surveys. Thus not much local level (i.e., district-level or attractions-level) analysis has been conducted by the HKTB.

2.3 Changes due to COVID-19

Due to the outbreak of COVID-19, the DVS has been suspended since February 2020. An alternative plan that the HKTB is currently piloting is to conduct online surveys by distributing postcards with a QR code to travellers at the airport. To encourage more responses, the

HKTB also provides some incentives (e.g., E-coupons) to people who have completed the online survey. The data collection is still ongoing, so the quality of those surveys is not yet known. But some potential biases can be foreseen regarding the sample representativeness. For example, young people might be more willing to fill in the online survey than elderly people. Thus the HKTB has proposed a hybrid approach to collecting survey data in future when the visitor flows return to normal, with a combination of an online survey and face-to-face interviews to ensure the sample quality.

2.4 Alternative data sources/future plans

As in most destinations, some alternative data sources such as mobile positioning data and credit card transaction data are examined by the HKTB to generate some analysis in tourism statistics. The Hong Kong government is currently running a project named Smart Tourism to promote the use of big data or online surveys. The HKTB is also seeking collaboration with Visa and Mastercard, to obtain data on visitor spending in Hong Kong. However, those explorations are still at an early stage. Issues and concerns regarding these big data sources have been mentioned during the interview, such as privacy in both mobile positioning data and bank card transaction data. Normally only aggregated level data without any personal information would be provided. Therefore, it would be difficult to examine whether the sample is representative and unbiased.

Saudi Arabia

Limited public information of tourism statistics in Saudi Arabia is available. Thus, the following information was obtained from the interview with the national DMO of Saudi Arabia.

1. Visitor arrivals

1.1 Data source

The primary source of inbound visitor arrivals is the data from passport control. Administrative data such as nationalities, date and port of arrival and gender are obtained on a regular basis. The system of records on individual arrivals is aggregated at national level so that no personal information is involved. The primary advantage for visitor arrivals is that Saudi Arabia has a strong passport control system that all individuals will go through when they enter the country. This guarantees the quality of inbound visitor arrivals data.

The number of Saudis residing abroad is also estimated based on the combination of administrative data from passport control and data from surveys which are conducted at the border exit. Details on the survey will be illustrated in the following sections.

In terms of regional level data, the region where the visitors visited is considered as their primary region of visitation in most cases. In the survey, the visitors are asked about the places visited, the number of nights stay, and their spending in the regions, which captures arrivals and spending data at the regional level. The subsample sizes are generally considered sufficient due to a relatively large overall sample size (about 36,000 completed interviews and 48,000 trips in any given year).

1.2 Publishing schedule and format

At the time of writing (March 2021), data are published quarterly mainly as PDF documents. Inbound data for the previous month specifically will be released on the 15th of each month. However, according to the DMO, the website will be updated to publish more open and detailed information later 2021 on a monthly basis. More detailed monthly information in both PDF format and online dashboards will be accessible internally within the ministry and for some key stakeholders.

1.3 Alternative sources/future plans

Currently the DMO is working on the improvement of the passport control data to include additional information such as the length of stay in order to estimate the Saudis residing abroad. The survey will be continued because considerable additional information is covered.

2. Visitor expenditure

2.1 Data collection and sampling

2.1.1 Survey collection procedure

Visitor expenditure data are estimated by the combination of administrative data from the passport control and the survey conducted at the border exits in the format of face-to-face interviews. The survey is comprehensive (15-20 minutes) and has been used to calculate

spending and calibrate visits. These surveys are distributed by port based on the volume of traffic in each port.

2.1.2 Sample size and coverage

About 3,000 surveys are completed every month, with around 36,000 in any given year. Only overnight visitors are captured by the survey. Visitors who come to the country more than 4 times a month are excluded from the definition of tourists.

2.2 Data processing

The combination of survey data and the administrative data from the border control is used to estimate inbound arrivals, spending, average length of stay, average spending per trip and average spending per night.

Besides the survey data, some alternative data such as points of sale, credit cards and money transfers are explored as complementary data to understanding of spending amount in calibration with the spending amount obtained from the surveys.

2.3 Publishing schedule and format

As mentioned earlier, limited information is currently available in terms of visitor spending. However, there has been a plan to publish more detailed monthly statistics in the near future.

2.4 Alternative data sources/future plans

Due to the outbreak of COVID-19, the Saudi borders have been closed through 2020 with very limited access. As a result, the inbound surveys have been interrupted, whereas the border data from passport control are still available. As the number of international visitors is quite small, the DMO of Saudi Arabia mostly focused on domestic visitors and started to explore some alternative data sources such as mobile positioning data, credit card data, point-of-sale data and flight data. They are working to further strengthen the current methodology with the integration of these additional big data. These additional data sources have been explored in terms of domestic tourism. Strong correlation has been detected in relation to the domestic survey data and the accuracy of domestic travel data has been improved with the incorporation of these data. Thus, if the border is opened in the second half of 2021, they plan to explore how these new data sources can be further infused into their system of international tourism statistics.

In the meantime, some concerns and challenges relating to these big data sources have been expressed. For credit card data and point-of-sale data, one of the biggest challenges is that accommodation payments sometimes occur via online transactions. Thus, it would be difficult to trace the transactions accurately to attribute them to accommodation payments in the country. Another issue is the difference between a card user's citizenship and the issuing place of the card, which makes it difficult to identify his or her nationality correctly. Meanwhile, it is difficult to distinguish whether the card user is a tourist or not, as only normalised and aggregated data are available to protect personal identifiable information. Another challenge with the point-of-sale data is the attribution to the tourism industry. A strong system that maps merchant category codes (MCC) with International Standard Industrial Classification of All Economic Activities (ISIC) codes is required, whereas that mapping is not always straightforward. With mobile positioning data, the main challenge is associated with the network operators' data availability including the penetration rate, market share, demographics and distribution of customers of the company. But those data are not all available for every network operator. Meanwhile, some people may have multiple mobile phones for different operators, whereas some people (i.e., young children) do not have mobile phones. Thus, mobile phone position has important constraints.

In terms of the survey, one limitation is that the understanding of questions or definitions is different across countries. Therefore, an accurate and agreed understanding is difficult to achieve when interviewing visitors from different countries. Another challenge is associated with the estimation of city-level tourist statistics. As the survey is designed to reflect statistics at the national level, city-level statistics cannot be fully captured through the survey.

The use of border data does not have such issues. The border control systems are quite accurate, compared to some countries in the European Union which have open 'internal' borders, as every passport scan is recorded in the system.

5. Conclusion

Through a mixed-method approach including both desk research and stakeholder interviews, and based on a case study strategy including twelve selected destinations, this study analyses and compares the approaches to measuring inbound tourism. It has been found that three main sources of data are used to measure visitor arrivals across destinations: 1) data from passport control or entry ports; 2) sample or census surveys collected from accommodation establishments; and 3) data from visitor surveys.

Collecting data from passport control or entry ports is considered to be the most straightforward approach. However, a potential disadvantage is that visitors from certain countries could be excluded from the system. The issue can be potentially addressed by complementing it with the outbound data reported by the origin countries. When obtaining data from border control is not an option, accommodation statistics are often used to estimate visitor arrivals. The limitation of this approach is that same-day visitors and visitors staying in informal accommodation are not covered, while visitors staying in multiple accommodation establishments would be counted multiple times. With a good sampling method, the visitor survey approach conducted at departure ports tends to generate reliable data to estimate both arrivals and expenditure statistics. However, the associated costs and potentially limited local level sample size for particular regions can be limitations. Some adjustment such as a rolling average may be necessary to account for small sub-samples. Another potential solution is developing partnership with local tourism authorities to conduct visitor surveys as a complementary source of the local level data.

In terms of visitor expenditure, this is normally estimated based on visitor surveys, supplemented with additional sources such as administrative data or census data to stratify or weight the survey sample. Common strengths and limitations in relation to surveys exist as discussed above. Some of the solutions, such as local partnerships, may still apply. With regard to the data collection method of the survey, most destinations conduct face-to-face interviews, while some destinations have employed self-administered online surveys. The main benefit of the latter approach is cost saving. It may also be the only feasible way to conduct the survey in situations such as the COVID-19 pandemic. However, the lower reliability due to the time lag, difficulty or inconsistency in interpreting questions, lack of control in respondent groups and lower response rates are important challenges. Face-to-face interviews carried out by well-trained interviewers can help address these issues. But higher costs and interruptions during the pandemic are disadvantages of this approach.

Given the above limitations of the commonly used approaches, alternative data sources and measures have been explored across destinations. Mobile positioning data, bank card transaction data, Google trends data and other forms of big data such as hotel and flight booking data are emerging options. Lower costs and higher efficiency are the main benefits. On the other hand, issues such as privacy and legal concerns as well as lower reliability due to the difficulty of filtering the “noise” from big data mean that these measures may be used to reflect trends instead of actual levels. There have been some efforts from both academics and practitioners to incorporate big data as one of the explanatory variables into models to help estimate tourism demand. However, far from a “a one-size-fits-all” solution, the exploration remains at its early stage. Caution is required when investigating the features of the data and interpreting the results.

References

- Granger, C.W.J. (1969) Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, 37(3). pp.424–438. doi: 10.2307/1912791.
- Li, H., Chen, J.L., Li, G. and Goh, C. (2016) Tourism and regional income inequality: Evidence from China, *Annals of Tourism Research*, 58. pp.81–99. doi: 10.1016/j.annals.2016.02.001.
- Li, H., Hu, M. and Li, G. (2020) Forecasting tourism demand with multisource big data. *Annals of Tourism Research*, 83, p.102912. doi: 10.1016/j.annals.2020.102912.
- Pan, B., Wu, D.C. and Song, H. (2012) Forecasting hotel room demand using search engine data, *Journal of Hospitality and Tourism Technology*, 3(3), pp.196–210. doi: 10.1108/17579881211264486.
- Song, H., Dwyer, L., Li, G. and Cao, Z. (2012). Tourism economics research: A review and assessment. *Annals of Tourism Research*, 39 (3), pp.1653–1682. doi: 10.1016/j.annals.2012.05.023.
- UNWTO (2020) *Methodological Notes to the Tourism Statistics Database, 2020 Edition*. World Tourism Organization (UNWTO). doi: 10.18111/9789284421473.
- VisitBritain (2020) *Britain's visitor economy facts*. Available at: <https://www.visitbritain.org/visitor-economy-facts> (Accessed 28 December 2020).
- Volchek, K., Liu, A., Song, H. and Buhalis, D. (2019) Forecasting tourist arrivals at attractions: Search engine empowered methodologies, *Tourism Economics*, 25(3), pp.425–447.
- Volo, S. (2019) Tourism statistics, indicators and big data: a perspective article, *Tourism Review*, 75(1). pp.304–309. doi: 10.1108/TR-06-2019-0262.
- WTTC (2020) Travel & Tourism: Global Economic Impacts & Trends 2020. World Travel and Tourism Council. Available at: <https://wttc.org/Research/Economic-Impact> (Accessed: 30 December 2020).

Summary of interview sources

Destination	Organisation
UK	Office for National Statistics (ONS)
UK (IPS user)	Tourism Northern Ireland Marketing Manchester VisitWales
Australia/New Zealand	Australian Trade and Investment Commission (Austrade)
Ireland/EU	Central Statistics Office (CSO)
Northern Ireland/Ireland	Northern Ireland Statistics and Research Agency (NISRA)
Japan	Japan Tourism Agency (JTA)
South Korea	Korean Tourism Organisation (KTO)
Austria/EU	Statistics Austria
France/EU	National Federation of Institutional Tourism Organisations University of Perpignan Via Domitia
USA	U.S. Department of Commerce
South Africa	North-West University, South Africa
Hong Kong	Hong Kong Tourism Board (HKTB)
Saudi Arabia	Ministry of Tourism