NATIONAL REFERENCE METADATA IN SINGLE INTEGRATED METADATA STRUCTURE (SIMS)

CONCEPT 1 – CONTACT

Sub-Concept 1.1: Contact organisation

National Statistics Office

Sub-Concept 1.2: Contact organisation unit

Tourism and Education Statistics

Sub-Concept 1.3: Contact name

Lara Pace

Sub-Concept 1.4: Contact person function

Principal Statistician

Sub-Concept 1.5: Contact mail address

National Statistics Office (NSO),
Tourism and Education Statistics,
Lascaris,
Valletta, VLT 2000,
Malta

Sub-Concept 1.6: Contact e-mail address

lara.e.pace@gov.mt

Sub-Concept 1.7: Contact phone number

25997630

CONCEPT 2 – METADATA UPDATE

Sub-Concept 2.1: Metadata last certified

21/12/2021

Sub-Concept 2.2: Metadata last posted

21/12/2021

Sub-Concept 2.3: Metadata last update

21/12/2021

CONCEPT 3 – STATISTICAL PRESENTATION

Sub-Concept 3.1: Data description

The National Tourism survey is conducted on a quarterly basis amongst the Maltese residents (excluding institutions) aged 15 years and over, residing in private households to collect information on:

- Outbound and domestic tourism trips with at least one overnight stay, which ended during the reference period.
- Outbound and domestic same-day visits during the reference period.

In the fourth quarter, an ad-hoc module is added to the questionnaire to collect information on participation in tourism for personal purposes during the reference year.

The questionnaire consists of five sections:

Section A – Number of Overnight Trips
Section B – Characteristics of Overnight Trips
Section C – Same-Day Visits
Section D – Participation in Tourism (Ad-hoc Module for Quarter 4 2019)
Section E – Socio-Demographic Characteristics

Sub-Concept 3.2: Classification system

The following statistical classifications are used for the National Tourism survey as a model instrument for coding variables and presenting data in an ordered manner:

- List of Countries - ISO 3166, which s the International Standard for country codes that are used internationally.

- List of Accommodation Establishments - NACE Rev. 2, which is the classification of economic activities corresponding to ISIC Rev. 4 (International Standard Industrial Classification of All Economic Activities) at European level.

- List of Localities - The local administrative units, abbreviated as LAUs, form a system for dividing up the economic territory of the European Union (EU) for the purpose of statistics at local level based on the geographical location of units. They have been set up by Eurostat and they are compatible with NUTS (Nomenclature of Territorial Units for Statistics). At local level, two levels of LAU have been defined: The upper level (LAU1, formerly NUTS level 4) is defined for most, but not all, of the countries. The lower level (LAU2, formerly NUTS level 5) consists of municipalities or equivalent units in the EU Member States.

- Level of Education - The International Standard Classification of Education (ISCED-2011) has been designed as an instrument suitable for assembling, compiling, and presenting statistics of education both within individual countries and internationally. It is expected to facilitate international compilation and comparison of education statistics.

• Labour status - The concepts and definitions used are based on those used in the Labour Market Unit (NSO) and as recommended by the International Labour Organisation (ILO).

Sub-Concept 3.3: Sector coverage

Maltese residents aged 15 and over, residing in private households (excluding institutions), stratified by gender, age group (15-24, 25-34, 35-54, 55+) and district of residence.

Sub-Concept 3.4: Statistical concepts and definitions

See Regulation 692/2011 and Methodological Manual for Tourism Statistics

Sub-Concept 3.5: Statistical unit

Statistical unit

- Participation in tourism: The statistical unit is the individual.
- Tourism trips with overnights stays: The statistical unit is the trip with at least one overnight stay made by the individual.
- Same-day visits: The statistical unit is the same-day trip made by the individual.

Reporting unit

One person in the selected household (within the age scope)

Sub-Concept 3.6: Statistical population

Overall target population

Maltese residents aged 15 and over residing in private households stratified by gender, age group and district of residence.

Participation in tourism

Maltese residents, aged 15 or over, participating in tourism for personal purposes.

Tourism trips and visitors making the trip

All tourism trips of at least one overnight stay made outside the usual environment by the residents aged 15 or over.

Same-day visits

All outbound and domestic same-day visits made outside the usual environment by the residents aged 15 or over.

Sub-Concept 3.7: Reference area

Persons aged 15 and over residing in private households (excluding institutions). NUTS Level 3.

Sub-Concept 3.8: Time coverage

Participation in tourism (Yearly) - Data are available from 2013-2019.
Tourism trips (Yearly) – Data are available from 2014-2019.
Same-day visits (outbound) (Yearly) – Data are available from 2014-2019.
Same-day visits (domestic) (Year) – Data are available for 2018 only.

Sub-Concept 3.9: Base period

Not applicable.

CONCEPT 4 – UNIT OF MEASURE

Tourism trips with overnights stays

The respondent selected for the interview will report on the overnight trip(s) made during the reference period and give the details for the different characteristics of the trip and his/her socio-demographic profile (the visitor making the trip). Each observed trip is an individual record in the transmitted micro-data file. Data for expenditure on tourism trips is expressed in Euros.

Participation in tourism

Data are transmitted in the form of aggregated tables with absolute values representing Maltese residents, aged 15 and over, participating in tourism for personal purposes.

Same-day visits

Characteristics of same-day visits (both outbound and domestic) are transmitted in the form of aggregated data with absolute values representing the number of same-day trips by the Maltese residents. Expenditure figures are expressed in Euros.

CONCEPT 5 – REFERENCE PERIOD

2019

CONCEPT 6 – INSTITUTIONAL MANDATE

Sub-Concept 6.1: Legal acts and other agreements

The Malta Statistics Authority Act. The Act empowers the NSO to collect, compile, extract and release official statistics related to demographic, social, environment, economic and general activities, and conditions of Malta.

At European level the following regulations are followed: Regulation 692/2011 and Regulation 1051/2011

Sub-Concept 6.2: Data sharing

There are no particular data sharing agreements in place. Users can request information through the NSO website:
https://nso.gov.mt/en/Services/Pages/Request-for-Information.aspx

CONCEPT 7 – CONFIDENTIALITY

Sub-Concept 7.1: Confidentiality – Policy

At National level:

The NSO requests information for the compilation of official statistics according to the articles of the MSA Act – Cap. 422 and the Data Protection Act – Cap. 586 of the Laws of Malta implementing the General Data Protection Regulations (GDPR).

Article 40 of the MSA Act stipulates the restrictions on the use of information while Article 41 stipulates the prohibition of disclosure of information. Furthermore, Section IX of the Act (Offences and Penalties) lays down the measures to be taken in case of unlawful exercise of any officer of statistics regarding confidentiality of data.

Since its inception, the NSO has always assured that all data collected remains confidential and that it is used for statistical purposes only according to the articles and derogations stipulated in the laws quoted above. The Office is obliged to protect the identify of data providers and refrain from divulging any data to third parties that might lead to the identification of persons or entities.

During 2009, the NSO has set up a Statistical Disclosure Committee to ensure that statistical confidentiality is observed, especially when requests for microdata are received.

Upon employment, all NSO employees are informed of the rules and duties pertaining to confidential information and its treatment. In line with stipulations of the MSA Act, before commencing work, every employee is required to take an oath of secrecy whose text is included in the same Act.

An internal policy on anonymisation and pseudo-anonymisation is in place to ascertain that adequate methods are used for the protection of data which the office collects and shares with the public in its capacity as the National Statistics Office. The policy is meant to safeguard confidentiality of both personal and business data entrusted to the NSO. The document provides guidance for all NSO employees who process data on a daily basis as to how anonymisation and pseudo-anonymisation methods should be applied. The policy applies to all confidential, restricted and internal information, regardless of form (paper or electronic documents, applications and databases) that is received, processed, stored and disseminated by the NSO.

At European level:

Regulation (EC) No 223/2009 on European statistics (recital 24 and Article 20(4) of 11 March 2009 (OJ L 87, p. 164), stipulates the need to establish common principles and guidelines ensuring the confidentiality of data used for the production of European statistics and the access to those confidential data with due account for technical developments and the requirements of users in a democratic society.

Sub-Concept 7.2: Confidentiality – Data Treatment

In order to guarantee confidentiality of personal data, only aggregated data are published (observation units are not recognizable either directly or indirectly). However, to avoid the publication of figures which are statistically unreliable, NSO applies the following basis guideline when disseminating aggregate tables:

- Estimation based on 20 to 49 sample observations => flag 'u'
- Estimation based on fewer than 20 observations => flag 'c'

Concept 8 – Release Policy

Sub-Concept 8.1: Release Calendar

At National level

All releases are published and disseminated at 11:00 a.m. as scheduled in the Advance Release Calendar. The calendar is published on the NSO website and includes a three-month advance notice (the current month and the forthcoming two months).

At European level

Data are released via Eurostat public database (https://ec.europa.eu/eurostat/data/database) shortly after reception of the data are transmitted. Data are released when all validation rules are respected and when data are reliable.

Sub-Concept 8.2: Release Calendar access

https://nso.gov.mt/en/News_Releases/Release_Calendar/Pages/News-Release-Calendar.aspx

Sub-Concept 8.3: User access

An internal policy on dissemination is in place to govern the dissemination of official statistics in an impartial, independent and timely manner, making them available simultaneously to all users.

The NSO's primary channel for the dissemination of official statistics is the NSO website. Tailored requests for statistical information may also be submitted through the NSO website.

Moreover, dedicated news releases are available in electronic format on the NSO website.

Concept 9 – Frequency of Dissemination

At National level

Figures on domestic tourism trips and visitors making the trip are published once a year in the form of the news release: Regional Tourism Demand

At European level

In the Regulation concerning European statistics on tourism, a transmission of the annual demand side data (Annex II of the Regulation, data on participation, trips, same-day visits) is foreseen within six months after the end of the reference year (T+6 months).

Concept 10 – Accessibility and Clarity

Sub-Concept 10.1: News release

On a national level, domestic tourism statistics are published in the following dissemination product:

Regional Tourism Demand news release:
https://nso.gov.mt/en/News_Releases/View_by_Unit/Unit_C3/Tourism_Statistics/Pages/Regional-Tourism.aspx

Sub-Concept 10.2: Publications

Not applicable.

Sub-Concept 10.3: Online Database

At European level, tourism statistics of the demand side data (Annex II of the Regulation, data on participation, trips, same-day visits) are published on Eurostat online statistical database in the form of aggregated data which can be assessed on the following website (Databases by themes – Industry, trade and services - Tourism (tour): http://ec.europa.eu/eurostat/data/database

Sub-Concept 10.4: Micro-data access

Microdata in anonymised form may be provided, upon request, under strict conditions to a selected number of institutions or persons accredited as research entities or researchers respectively. Further information on access to anonymised microdata is available on the NSO website through:
http://nso.gov.mt/en/Services/Microdata/Pages/Access-to-Microdata.aspx.

Sub-Concept 10.5: Other

Not applicable.

Sub-Concept 10.6: Documentation on methodology

The Methodological manual for tourism statistics provided by the Eurostat is used as guidelines for the compilation of the tourism statistics:
http://ec.europa.eu/eurostat/documents/3859598/6454997/KS-GQ-14-013-EN-N.pdf.

Data collection guidelines based on the Tourism Methodological manual (provided by Eurostat) is also available for the interviewers.

Sub-Concept 10.6.1: Metadata completeness rate

Information about all required metadata concepts (and sub-concepts thereof) are provided.

Sub-Concept 10.7: Quality Documentation

The NSO has developed an internal Quality Management Framework (QMF) which is built on common requirements of the ESS Code of Practice (ESS CoP). A document was prepared to include a set of general quality guidelines spanning over all statistical domains. Assuring methodological soundness is an integral part of the QMF, nonetheless, the document spans also on other areas related to institutional aspects.

CONCEPT 11 – QUALITY MANAGEMENT

Sub-Concept 11.1: Quality Assurance

The Methodological manual for tourism statistics provided by the Eurostat is used as guidelines for the compilation of the tourism statistics:
http://ec.europa.eu/eurostat/documents/3859598/6454997/KS-GQ-14-013-EN-N.pdf.

The NSO ensures the accuracy of data released to the public and prepares clear methodological notes which explain the processes involved in the collection and production of official statistics.

The NSO has developed an internal Quality Management Framework (QMF) which is built on common requirements of the ESS Code of Practice (ESS CoP). A document was prepared to include a set of general quality guidelines spanning over all statistical domains. Assuring methodological soundness is an integral part of the QMF, nonetheless, the document spans also on other areas related to institutional aspects.

Every five to seven years, the NSO participates in a Peer Review exercise through which the compliance of its operations with principles of the ESS CoP is assessed by an expert team. Peer Reviews are indeed part of the European Statistical System (ESS) strategy to implement the ESS CoP. Each NSI is expected to provide information as requested by a standard self-assessment questionnaire. Following this an expert team visits the office to meet NSI representatives and main stakeholders. Peer Reviews result in a compliance report and the listing of a set of Improvement Actions which need to be followed up by the NSI. The next round of Peer Reviews is planned to be carried out in 2022.

Sub-Concept 11.2: Quality Assessment

- Quality assurance in survey on tourist trips is carried out in accordance with Regulation 1051/2011 implementing Regulation 692/2011 of the European Parliament and of the Council concerning European statistics on tourism, as regards the structure of the quality reports and the transmission of the data.
- Data collection, treatment and publication follow Eurostat Methodological Manual. This improves comparability with other countries.
- For data collection, interviewers are employed and trained by NSO to ensure quality and comparability.
- Monitoring and supervising the interviewers allows detecting where there is room for improvement.
- Data are checked regularly during data collection, to improve data collection methods where necessary.
- Data treatment is clearly documented.
- Metadata report is delivered to Eurostat.
- The Quality Management Framework (QMF) for the NSO in order to ensure consistency with the European Statistics Code of Practice.
- The European Statistics Code of Practice sets the standard for developing, producing and disseminating European statistics

Concept 12 – Relevance

Sub-Concept 12.1: User needs

The data are highly sought after by the media, researchers, students and policy makers. Such users can request information through the NSO website and data are provided at aggregated level.

Sub-Concept 12.2: User satisfaction

The last User Satisfaction Survey was held in 2014 with the aim to collect information about key users' satisfaction with statistical output.

The NSO keeps record of the number of News Releases and publications disseminated on its website; the users to whom statistical products are provided; as well as the number of requests that are processed every year.

News Releases and tailor-made statistical outputs were assessed on account of their quality, timeliness, and on their ability to meet users' needs.

Sub-Concept 12.3: Data Completeness

Compliant with the requirements of Regulation on tourism statistics 692/2011 as well as recommendations found in the Methodological Manual for tourism statistics.

Concept 13 – Accuracy and reliability

Sub-Concept 13.1: Overall accuracy

- The main possible source of error comes from the memory effect (respondents might forget trips and expenditure). This would lead to an underestimation of trips and expenses.

- Sampling errors vary among variables. The main variables are relatively accurate (e.g. Number of trips with purpose "leisure, recreation and holidays") whereas more detailed variables have a larger sampling error (e.g. expenditure related questions).

- Non-response errors: Significant item non-response appears for variables referring to expenditures. For dealing with item non-response the regression imputation methods are used.
- Measurement errors: (1) by interviewers: possibility of not understanding or misinterpreting the respondent's answers, typing errors and (2) by respondents: possibility of not understanding the questions, not knowing the answers, not interested in the survey, not willing to provide proper information on sensitive questions such as expenditure. Main possible source of error comes from recall bias (respondent might forget some details on trips, such as expenditure). This would lead to an underestimation of trips and expenses.

Sub-Concept 13.2: Sampling errors

Sampling errors

- Participation in tourism: number of residents, aged 15 or over, having made at least 1 trip of at least 1 overnight stay (all age groups) – 284,028 (CV: 2.21%)

- Tourism trips - Total number of trips – 908,931 (CV: 1.62%)

- Domestic trips – 237,237 (CV: 3.28%)

- Outbound trips - 671,694 (CV: 1.49%)

- Private/Personal trips – 785,430 (CV: 1.15%)

- Professional/business trips – 123,502 (CV: 5.29%)

- Tourism trips - Total expenditure – 706,063 (CV: 4.49%)

Note: Expenditure estimates are rounded to the nearest thousand.

Outlier detection techniques

Two variables that are most likely to be affected with outliers are the number of nights spent and the expenditure categories. The standardized residuals, which are obtained by fitting a multiple linear regression model, are the most commonly used measures for detecting outliers. Observations with resulting standardized residuals exceeding +2 or -2, indicates that one should investigate these data points. An alternative measure that is used for indicating data points that are particularly worth checking for validity is Cook's distance. A conventional cut-off value that is used for spotting highly influential points is when the resulting Cook's distance is greater than 4/n, where n is the sample size. The two methods are used to detect outliers for the number of nights spent and for the different expenditure categories.

Sub-Concept 13.3: Non-sampling error

The following are the main sources of non-sampling errors:

- Interviewers' bias – Questions laced with interviewer bias can influence respondents in such a way that distorts the outcome of the interview;
- In case of business trips organized by employers using services of travel agencies, respondents do not know the costs of transport and other services;
- Sometimes respondents are reluctant to provide data on expenditures;
- Respondents do not remember on details;

Bias associated with imputation.

Sub-Concept 13.3.1: Coverage error

The sampling frame being used covers private households. Hence persons living in institutional households are not being covered. Since the 2011 Census is being used as a sampling frame, households created after 2011 are not well represented. A degree of under-coverage may exist as not all immigrants are immediately captured in the register.

Sub-Concept 13.3.1.1: Over Coverage

There is over-representation for households created prior to 2011 and for households which were present in 2011 and no longer exist after 2011. Units which are not eligible for the survey and were included in the sample due to lack of information (e.g. still registered as permanent residents in Malta but living abroad).

Sub-Concept 13.3.1.2: Common Units Proportion

Not applicable.

Sub-Concept 13.3.2: Measurement error

<u>Measurement errors by respondents</u>: possibility of not understanding the questions, not knowing the answers (do not remember details), not interested in the survey, not willing to provide proper information on sensitive questions such as expenditure. Main possible source of error comes from recall bias (respondent might forget some details on trips, such as expenditure). This may lead to over- or underestimation of trips and expenses.

<u>Measurement errors by interviewers</u>: possibility of not understanding or misinterpreting the respondent's answers, typing errors, recall bias, interviewers' bias (Questions laced with interviewer bias can influence respondents in such a way that distorts the outcome of the interview.)

In case of business trips organized by employers using services of travel agencies, respondents do not know the costs of transport and other services. Significant item non-response appears for variables referring to travel ticket/accommodation/other expenditures. For dealing with item non-response, regression imputation methods are used.

Households using only mobile phone connections cannot be reached since only landline phone numbers are available. Also, some of the landline phone numbers tend to be old as such information on the contact of individuals were obtained from the census.

Quality improvements

- New in-built validations in the data collection program.
- Improvement in the wording of the questions to minimize interviewers' bias especially in the type of accommodation question.
- The length, design and complexity of the questionnaire.

Intensive interviewer's training.

Sub-Concept 13.3.3: Non-response error

The strategy used for the data collection was stratified random sampling where the strata were constructed using District, Gender and Age Group. Each stratum is adequately represented by a relevant quota. These need to be met for the sample to be representative. The net sample should contain a minimum of 1,823 individuals per quarter.

During the year 2019, the gross sample was 29,279 persons. Of these, a total of 15,803 persons were contacted to participate in this survey and 13,476 were never contacted since all quotas were reached. Out of the total of 15,803 individuals, 7,362 participated, 671 were refusals, 2,460 were not eligible to participate and 5,310 didn't answer their call. Persons who were contacted at least once to no avail were eventually not contacted again due to exhausted quotas. This resulted in a net effective response rate of 91.4 per cent.

The main causes of unit non-response are: respondents might not understand the questions, do not remember the details of the trip, not interested in the survey or not willing to provide proper information on sensitive questions such as expenditure. Main possible source of error comes from recall bias (respondent might forget some details on trips, such as expenditure). This may lead to over- or underestimation of trips and expenses.

Methods used for dealing with unit non-response: Post-stratification weighting basing on respondents' sex, age and district (LAU 1). Figures were inflated over the total population despite being collected from households.

Significant item non-response appears for variables referring to expenditures. For dealing with item non-response the regression imputation methods are used.

Sub-Concept 13.3.3.1: Unit non-response

Unit non-response

Number of ineligible units/ elements: 7,750
Number of eligible units/elements: 8,053
Number of non-contacts: 20
Number of refusals: 671
Number of rejected questionnaires: 0
Number of other types of non-response: 0
Total non-response: 691

Unit non-response rate for Trip's dataset: 8.6%

Unit non-response rate for Participation dataset: 7.7%

Unit non-response rate for same-day visitors' dataset: 8.6%

Methods used for dealing with/minimising unit non-response: The strategy used for the data collection was stratified random sampling where the strata were constructed using District, Gender and Age Group. Each stratum is adequately represented by a relevant quota. These need to be met for the sample to be representative. The net sample should contain a minimum of 1,823 individuals per quarter.
During the year 2019, the gross sample was 29,279 persons. Of these, a total of 15,803 persons were contacted to participate in this survey and 13,476 were never contacted since all quotas were reached. Out of the total of 15,803 individuals, 7,362 participated, 671 were refusals, 2,460 were not eligible to participate and 5,310 didn't answer their call. Persons who were contacted at least once to no avail were eventually not contacted again due to exhausted quotas. This resulted in a net effective response rate of 91.4 per cent.

Methods used for dealing with unit non-response: Post-stratification weighting based on respondents' sex, age and district (LAU 1). Figures were inflated over the total population despite being collected from households.

Note: Gross sample consists of all units who have been contacted only. Ineligible units consider all units who have been contacted at least once to no avail, but which have not been attempted at a later stage due to quota being reached. Ineligible units consist of: (1) Ineligible, (2) No answer, (3) Does not live at home, (4) Disabled line, (5) Voicemail. Eligible units consist of: (1) Good response, (2) Refusal, (3) Disabled person. Non-contacts consist of: (1) Disabled person. Refusal consists of: (1) Refusals.

Sub-Concept 13.3.3.2: Item non-response

Overnight Trips

The imputation rates are as follows:

- Package expenditure – 5.5%
- Transport (Travel ticket) expenditure – 9.8%
- Accommodation expenditure – 7.1%
- Other expenditure – 4.7%

Outbound Same day trips

The imputation rates are as follows:

- Transport (Travel ticket) expenditure – 12.0%
- Other expenditure – 10.0%

Sub-Concept 13.3.4: Processing error

In-built validations are available in the data collection program. We continue to do improvements for the program (Blaise) and the controls for data entry in order to minimize errors and missing values for the core variables (i.e. expenditures and socio-demographic characteristics).
The main variables are relatively accurate (e.g. Number of trips with purpose "leisure, recreation and holidays") whereas more detailed variables have a larger sampling error (e.g. Total Expenses).
No imputation is needed for the core variables (purpose of visit, destination, type of accommodation used etc.) as interviewers are instructed to get complete answers on the latest 3 trips. Therefore, respondents will be able to remember the characteristics and details of their trip.
Imputations are only carried out on expenditure categories, which provide us with very accurate expenditure estimates.

Proxy interviews are kept at a minimum in order to minimize the effects of non-response due to ageing and ageing-related health-conditions.

Sub-Concept 13.3.5: Model assumption error

The methodology used for weighting assumes that behavioural characteristics of persons living in households are very similar to those living in institutional households.

CONCEPT 14 – TIMELINESS AND PUNCTUALITY

Sub-Concept 14.1: Timeliness

The average production time for each quarterly data is two weeks. Timeliness (first results available days after the end of the reference year) – 183 days

Sub-Concept 14.2: Punctuality

The final, validated, and complete results were submitted to Eurostat on 14.05.2020.

Concept 15 – Coherence and Comparability

Sub-Concept 15.1: Comparability – Geographical

No geographical comparability problems. The sample is stratified by district and the data are collected uniformly at national level.

Sub-Concept 15.2: Comparability – Over Time

During 2014, the National Statistics Office has introduced a new tool to measure National Tourism in line with Eurostat recommendations in the methodological manual. Up to 2013, data for total outbound trips was collected by means of the ongoing border survey, while information on the last three domestic trips was collected by means of an extra module attached to the Labour Force Survey. As from reference year 2014, information on the last three trips (outbound + domestic) is collected via quarterly CATI surveys. Due to this change in methodology, National Tourism trips microdata for 2014 were marked with a 'break in series' flag.

Sub-Concept 15.3: Coherence – Cross Domain

Outbound tourism trips collected in the National Tourism survey are calibrated with the ongoing border Tourstat survey. In addition, the official population figures by sex, age, and district were used to calibrate the survey, and get more accurate raising factors.

Sub-Concept 15.3.1: Coherence – Sub-Annual and Annual statistics

Not applicable.

Sub-Concept 15.3.2: Coherence – National Accounts

Not applicable.

Sub-Concept 15.4: Coherence – Internal

Not applicable.

Concept 16 – Cost and Burden

The burden on the respondents is very low since the survey takes only a few minutes to be completed.

Significant high costs are associated with the data collection and production of the statistical product. These include:

- Ten interviewers and one supervisor are to be present simultaneously to carry out the calls;
- Overtime costs for one member from IT Systems who is present in case of any IT related issues;
- IT costs for the development and maintenance of the CATI program;
- A full-time statistician is responsible for data cleaning/analysis and for the final statistical product.

Concept 17 – Data revision

Sub-Concept 17.1: Data revision – Policy

The National Statistics Office has an internal data revision policy which we adhere to.

Sub-Concept 17.2: Data revision – Practice and Data Revision

All past National Tourism survey data, including trips micro-data and aggregated data for participation in tourism for the reference years 2014 to 2016, were re-weighted following the publication of revised population figures issued by the National Statistics Office on 12th February 2018 (Release No: 022/2018):

https://nso.gov.mt/en/News_Releases/View_by_Unit/Unit_C5/Population_and_Migration_Statistics/Pages/Population%20Revisions.aspx

The revised population figures affected the grossing up methodology used for the National Tourism survey.

Concept 18 – Statistical Processing

Sub-Concept 18.1: Source data

Source data

Survey on domestic and outbound tourism trips, participation in tourism, outbound and domestic same-day visits.

Survey vehicle

Stand-alone survey

Population frame

Population census - Persons aged 15+ residing in private households (excluding individuals living in institutions).

Coverage errors of population frame

- Difficulties locating persons that they are no longer living in the same household.
- Residents that were temporarily away during the Census period and now they returned to Malta.
- Non-residents that were living in Malta during the Census period and now they left Malta.
- Persons that during census were living in the households and now have moved to an institution.
- Occupied or unoccupied dwellings.

Gross sample size

Gross sample size for trips: 15,803

Gross sample size for participation in tourism: 3,894

Gross sample size for (outbound and domestic) same-day visits: 15,803

Net sample size

Net sample size for trips: 7,362

Net sample size for participation in tourism: 1,842

Net sample size for (outbound and domestic) same-day visits: 7,362

Sampling design

Stratified, simple random sampling by Sex, Age group and District

Sub-Concept 18.2: Frequency of data collection

Overnight and same-day trips are collected once every quarter. Participation in tourism is collected only once a year during the data collection of the reference quarter 4.

Sub-Concept 18.3: Data Collection

- Every quarter, the domain unit do regular meetings with the Data Management unit to discuss the requirements and enhancements on the National Tourism survey.
- The Data Management Unit inform the IT Unit, if necessary, to apply any developments and enhancements in the data collection tool.
- The Methodology and Quality Unit prepares the quotas and the sampling frame.
- Data collection is carried out at the Data Management Unit via CATI (Computer-Assisted Telephone Interview) by around ten interviewers.
- A supervisor should be present during the data collection to assist in any queries that might arise. The supervisor is also responsible for the quotas as once the desired sample in each stratum is reached, the supervisor must cancel any pending appointments with respondents that are classified in the stratum where the quota has been reached.
- One member from IT Systems should also be present in case of any IT related issues.
- Once the data collection is ready, the raw data are extracted from the Blaise program.
- Data are cleaned in terms of value labels and checked for quality issues by the Methodology and Quality Unit, which is then sent to the domain unit.

- The domain unit will start the validating the data, do consistency checks, impute the missing data in the expenditure categories and finally compute the grossing-up weights with the mid-year household population and outbound tourists from the Tourstat survey.

- The mid-year household population is provided by the Population and Migration Statistics Unit.

- Micro-data and aggregated data are then sent to Eurostat and the Regional Tourism Demand news release is compiled

Sub-Concept 18.4: Data Validation

The completed raw data are further checked for inconsistencies and missing values. Some of the main checks include:

- Did the respondent undergo the right route in the questionnaire?
- Are the answers logical and consistent?
- Are the time periods possible?
- Are there any overlapping trips?

The electronic questionnaire facilitates the inclusion of validation and consistency checks and thus validation of data is possible during the data collection phase. Once the data are collected, it is transferred to the responsible statistician for data cleaning and statistical analysis. The main aim of this second stage is to assess completeness and item non-response. Special attention is given to completed fields on expenditures. The missing values for expenditures are imputed using regression models basing on the destination, departure month, type of accommodation and transportation variables (these variables are mandatory fields in the questionnaire). Monitoring of outliers including length of stay and expenditure categories. Data are thoroughly validated using of SPSS to ensure quality and consistency of results. These rules are mainly based on the definitions provided in the methodological manual prepared by Eurostat.

Sub-Concept 18.5: Data Compilation

- Data cleaning/editing: The completed raw data are checked for inconsistencies and missing values. Refer to Sub-Concept 18.4: Data Validation.

- Imputation: Multiple linear regression models are carried out on all expenditure categories. Linear regression is an approach to model the relationship between a scalar dependent variable and one or more explanatory variables. The method requires that the values of one or more auxiliary variables are known for both the complete cases on which the variable of interest is recorded and for the missing cases. A linear regression model relating the variable of interest to the set of auxiliary variables is set up.

The independent variables which are used in the linear regression models to estimate the missing values in our expenditure categories are the following:

Travel Ticket Expenditure
Dependent variable: Air/Sea Expenditure
Independent variables: Type of trip, Purpose of visit, Destination, Airline type

Accommodation Expenditure
Dependent variable: Accommodation Expenditure
Independent variables: Type of trip, Purpose of visit, Type of accommodation used, and Nights spent.

Other Expenditure
Dependent variable: Other Expenditure
Independent variables: Type of trip, Nights spent, and Type of accommodation used.

- Outlier detection: Two variables that are most likely to be affected with outliers are the number of nights spent and the expenditure categories. Refer to Sub-Concept 13.2: Sampling errors for a detailed description of the outlier treatment.

- Weighting: (1) Adjustment for non-response, (2) Calibration with the actual mid-year, household, population figures and outbound tourism trips from the Tourstat survey and (3) Re-align the last three trips on the total number of trips that were made during the reference period.

- Compilation of the Regional Tourism news release.

- Micro-data for tourism trips with overnights stays is transmitted to Eurostat.

- Data on participation in tourism are transmitted in the form of aggregated tables with absolute values representing Maltese residents, aged 15 and over, participating in tourism for personal purposes.

- Outbound and domestic same-day visits by purpose of visit (personal/business) are transmitted in the form of aggregated data with absolute values representing the number of same-day trips by the Maltese residents, together with their corresponding expenditure figures.

Sub-Concept 18.5.1: Imputation

Refer to Sub-Concept 13.3.3.2: Item non-response.

Sub-Concept 18.6: Adjustment

Weighting procedure

1. The mid-year household population (excluding institutions) of 2019, categorized by District (Northern Harbour, Southern Harbour, Western, South Eastern, Northern, Gozo and Comino), Age group (15-24, 25-34, 35-54, 55+) and Gender was used to compute the ratio of persons in the population to the respondents in the sample.

2. The software R was used as a supplementary tool for the calibration of the survey in order to re-align the survey results with the actual mid-year population figures and outbound tourism trips from the Tourstat survey. While National Tourism survey collects information on trips depending on the month of arrival, the border survey collects information by month of departure. Outbound tourism trips were calibrated by month of arrival instead by month of departure given that the scope of observation refers to trips that ended during the reference period, not trips that started during the reference period.

3. The total domestic and outbound trips were estimated on the total number of trips that were made during the reference period.

Note: Tourstat survey is used to calibrate outbound tourism trips (overnight) and outbound same-day visits of the National Tourism survey. Outbound tourism trips and outbound same-day trips are adjusted with the ongoing frontier national survey known as Tourstat. Tourist air departures are collected through a continuous survey carried out at the departure lounge of the Malta International Airport (MIA). A two-stage sampling design is used to collect air passengers. In the first stage, alternate days and nights are selected. In the second stage, within each shift, a sample of passengers is selected systematically. Every crossing passenger is counted, and respondents are selected using a pre-defined interval of 1:20 for air. Tourist sea departures are supplemented by administrative data provided by ferry operators.

Sub-Concept 18.6.1: Seasonal Adjustment

Not applicable

CONCEPT 19 – COMMENT

No further comments.