IMPROVING DATA COLLECTION FOR THE QUALITY OF LIFE IN EUROPEAN CITIES SURVEY

SAMPLING METHOD, INTERVIEWING METHODOLOGY AND STRUCTURE OF THE QUESTIONNAIRE

RUDOLF SELJAK, LUKASZ ARENDT

2021 edition

STATISTICAL WORKING PAPERS

eurostat

IMPROVING DATA COLLECTION FOR THE QUALITY OF LIFE IN EUROPEAN CITIES SURVEY

SAMPLING METHOD, INTERVIEWING METHODOLOGY AND STRUCTURE OF THE QUESTIONNAIRE

RUDOLF SELJAK, LUKASZ ARENDT

2021 edition

Manuscript completed in December 2021

Luxembourg: Publications Office of the European Union, 2022

Theme: General and regional statistics
Collection: Statistical working paper

ISBN 978-92-76-46110-4 doi: 10.2785/359627 ISSN 2315-0807 KS-TC-21-008-EN-N

Abstract

The Quality of Life in European Cities Survey (QLEC Survey) is a statistical survey that the European Commission has been carrying out since 2004. It aims at measuring the quality of life in 83 European (greater) cities through a perception survey. This report introduces the results of the review of the 2019 implementation of the QLEC survey. The findings are accompanied with recommendations for improvements to implement in the next survey wave.

The emphasis of the revision is placed on sampling and data collection modes. It largely focuses on the current interviewing mode (CATI) as well as on the sampling issues resulting from this exclusive data collection mode.

The paper consists of three parts. The first part summarises the main methodological elements of the previous survey implementation and describes the main problems identified during the revision. The second part provides detailed recommendations for necessary improvements. The third part brings those recommendations in a short and concise form. The Annexes present the results of the microanalysis as well as the proposals for CATI and CAWI questionnaires.

Keywords: QLEC Survey, data collection mode, CAPI, CAWI, sample size, non-response, under-coverage.

Authors: Rudolf Seljak, Lukasz Arendt

Acknowledgments: We would like to particularly thank the coordinators of this study at DG-REGIO, Ms Laura de Dominicis, and at Eurostat – Unit A5, Mr Matyas Meszaros.

The study was carried out by GOPA as a contractor of Eurostat in the framework of the specific contract on ‘Quality of Life in European Cities Survey’ (Specific contract N° 000100 Ref. N° 2021.0074 under framework contract Ref. 2018.0086).

Table of contents

List of Tables

List of Figures

Abbreviations

AI	Artificial Intelligence
CAPI	Computer Assisted Personal Interviewing
CATI	Computer Assisted Telephone Interviewing
CAWI	Computer Assisted Web Interviewing
CEDEOP	European Centre for the Development of Vocational Training
CV	Coefficient of Variation
DESI	Digital Economy and Society Index
EC	European Commission
ESQRS	Standard for Quality Reports Structure
ESS	European Statistical System
EU	European Union
GOPA	Gesellschaft für Organisation, Planung und Ausbildung
ICT	Information and Communications Technologies
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ITU	International Telecommunication Union
LAU	Local Area Units
LFS	Labour Force Survey
NMS	New Member States
OECD	Organisation for Economic Co-operation and Development
PIAAC	Programme for the International Assessment of Adult Competencies
QLEC	Quality of Life in European Cities
SAS	Statistical Analysis System
SILC	Statistics on Income and Living Conditions
SMS	Short Message Service

Introduction

Introduction

Cities play an important role in enhancing the socio-economic development of regions and countries. They form part of European Union strategies and policies (Rauhut, Humer 2020). From the point of view of EU policy, the role of cities as drivers of growth and development is discussed in the 2007 Territorial Agenda, in the 2020 Territorial Agenda, and in the Urban Agenda for the EU. These clearly emphasise the need for an urban-oriented approach in EU actions, which translates explicitly into the growth objectives of the Europe 2020 Strategy.

With the emergence of new technologies – e.g. ICT, Big Data, Internet of Things – and the concept of the 'smart city', cities are perceived as being entities that should combine different 'smart' approaches, in order to improve the quality of life of their inhabitants (Romanelli et al., 2019). The quality of life has been monitored by the European Commission since 2004 through the Quality of Life in European Cities survey (QLEC). The survey covers 83 European cities, that is all capital cities and up to six greater cities in larger Member States. It is conducted every three years. The last wave of the QLEC was held in 2019 in the form of a CATI-based survey, in which 700 inhabitants of each city were interviewed.

The aim here is to review the methodology currently being used to perform the QLEC survey, with a specific focus on advising on the current interviewing mode (CATI), the current sampling methodology, and the questionnaire's current structure to improve for the future waves taking into account new mode of data collection.

The paper is structured as follows:

- the first section summarises the survey methodology used in 2019 and presents the main problematic issues of the methodology, already identified in the Evaluation report (European Commission, 2020) as well as within the analysis conducted in this study;
- the second section provides a proposal for the revision of the QLEC methodology for the next wave;
- the third section contains a structured list of recommendations to be implemented in the next wave of the survey;
- the paper closes with the Annexes, which provide the results of the microanalysis as well as a proposal of CATI and CAWI questionnaires.

1 Overview of QLEC 2019

1. Overview of QLEC 2019

1.1 Overview of the current survey methodology

This paper first summarises the main elements of the survey methodology, as described in the Evaluation report and as explained in the correspondence with Eurostat and with DG REGIO. The focus of the review is placed mainly on the sampling and on the interviewing part of the survey's implementation:

- **Data collection mode**
  The 2019 edition of the Quality of Life in European Cities survey used the CATI interviewing mode. It is widely understood that surveys based on a CATI data collection mode only face ever-growing difficulties with ensuring their representativeness. This is mostly because the coverage of landline telephone numbers is constantly decreasing. The methodology for 'mobile phone surveys' also faces a large number of methodological issues that are still the subject of research and exploration - e.g. large amounts of disclosed numbers, large amounts of duplicates, difficulties with getting approval for interviews. It is also widely known that mobile-phone surveys tend to underrepresent certain population groups, i.e. older people, the less educated, and people in rural areas (Lamanna et al. 2019). Such an issue is addressed in more detail in the next sections of this paper.

CATI – Computer Assisted Telephone Interviewing

CATI – Computer Assisted Telephone Interviewing – is a mode of survey data collection in which a respondent is contacted by an interviewer by telephone. The interviewer follows the questionnaire according to the script developed using specialised computer software and records the respondent's answers in digital form during the interview. CATI provides less flexibility compared to face-to-face interviews as there is no visual contact between the interviewer and the respondent. However, it enables good quality control, sufficient flexibility to adapt the questionnaire to local languages and to assist respondents, in case of complex questions, and it is less expensive than 'face-to-face' interviews. It is advisable that the maximum time-frame for a survey should not exceed 20–30 minutes (de Leeuw 2008; Lamanna et al. 2019).

- **Questionnaire**
  The questionnaire used in the 2019 wave of the QLEC survey consisted of screening questions, some of which are in fact socio-demographic questions. There were 14 questions in the main part of the questionnaire relating to the respondents' perceived quality of life,

followed by 14 questions describing socio-demographic characteristics of the surveyed persons. This resulted in 79 variables that could be analysed. The average length of conducting the interview varied between 09:48 minutes in Ljubljana and 13:54 minutes in Athens, which is in line with the methodological recommendations.

The screening procedure aimed at the verification of eligibility by checking the respondent's age and the city of residence. As for regional eligibility, in most cases the respondent was asked to provide the postcode of her/his place of residence. If verification with the use of the postcode was not possible – as in Bulgaria, Ireland, Portugal, Romania and the United Kingdom) – the respondent was directly asked for her/his region of residence.

- **Definition of target population**

The theoretical target population was defined according to the following criteria:

  - Resident of one of the 83 European cities;
  - Persons that have sufficient command of (one of) the respective national/regional language(s) or of English, which allows them to comfortably answer the questionnaire;
  - Persons aged 15+ that live in a private household, which means that the target population excludes prisoners, residents of retirement homes, etc. who are hard to reach via a telephone survey.

- **Sample size**
  - The target net sample size was 700 responses in each of the cities. The achieved net sample size was exactly 700 responses in each of the cities.

- **Sample methodology**
  - Sample design was largely determined by the fact that it was a telephone (CATI) survey.
  - A 'Dual frame' approach was taken in sampling frame creation. Two separate, overlapping sample frames, mobile phones list and landline telephones list, resulted in three disjoint sub-sets of the frame: persons who only have access to a mobile phone (i.e. 'mobile only'), persons who only have access to a fixed line telephone (i.e. 'fixed only') and persons who have access to both mobile and fixed-line telephones (i.e. 'mixed').
  - For the list of landline telephone numbers, a list of publicly available numbers was used, together with the combination of Random digit dialling (RDD). The list of primary numbers was created on the basis of publicly available telephone books/registers. Based on this list, additional randomly generated numbers were created.
  - For the list of mobile numbers, an RDD approach was initially used to obtain the first list of mobile phone numbers. That list was then merged with the publicly available online data, mainly through the social media, in order to limit the list of numbers pertaining to the city in question. Details of this procedure are not provided in the Evaluation report.
  - The gross sample size was determined by applying a factor of 1:24 to the planned net sample. This amounts to a gross sample of 16 800 numbers per city. The separate random samples were drawn from two frames, i.e. list of mobile numbers and list of fixed numbers. Allocation to the fixed or to the mobile sample was determined by estimating phone type ownership data for each country and then applying that allocation key to each city in the country.

- One-stage stratified systematic random sampling was used. The following strata were used:

- City
- Phone type (mobile-fixed)
- Local Area Units (LAUs).
  - Implicit stratification: within each stratified LAU block, numbers were sorted according to the first three digits following the prefix.

1.2 Main methodological and practical problems of QLEC 2019

The Evaluation report of the QLEC 2019 survey (European Commission 2020) describes several problematic issues in the QLEC survey methodology, the most outstanding being:

- The slight under-representation of youth in the survey;
- Screening the eligibility of respondents participating in the survey through using post codes;
- The correct categorisation of the respondent's occupation, in line with the International Standard Classification of Occupations (ISCO)-08 classification;
- Another minor issue was related to the quality of translations into national languages – mainly into Albanian.
- One of the main issues with the questionnaire's structure identified during the fieldwork was related to a relatively high cognitive respondent burden at the beginning of the survey, resulting from the sequence of large-grid questions, which caused a large dropout rate around the Q2–Q4 block. Since it is expected for the scope of the survey in the current wave to be comparable to the 2019 wave, there is no option to shorten the questionnaire, but rather to introduce changes to the sequence of questions.
- The other issue, reported in the Evaluation report, relates to high item non-response rates, in the case of questions related to sensitive topics, such as corruption, social minorities etc., to the respondent's professional situation, and to some aspects of quality of life. Interestingly, in the case of questions regarding health status, the item non-response rate appeared to be relatively low. In general, the structure of the questionnaire has a very moderate impact on the item non-response rate in the case of questions on sensitive topics([1]). However, it is argued that, in the self-completion CAWI mode, this non-response rate is lower compared to modes in which there is personal contact with the interviewer. It can consequently be assumed that introducing a CAWI-CATI mixed mode should have a positive influence on item non-response rates.

The other potential methodological problems of the QLEC (2019) survey identified in this study encompass:

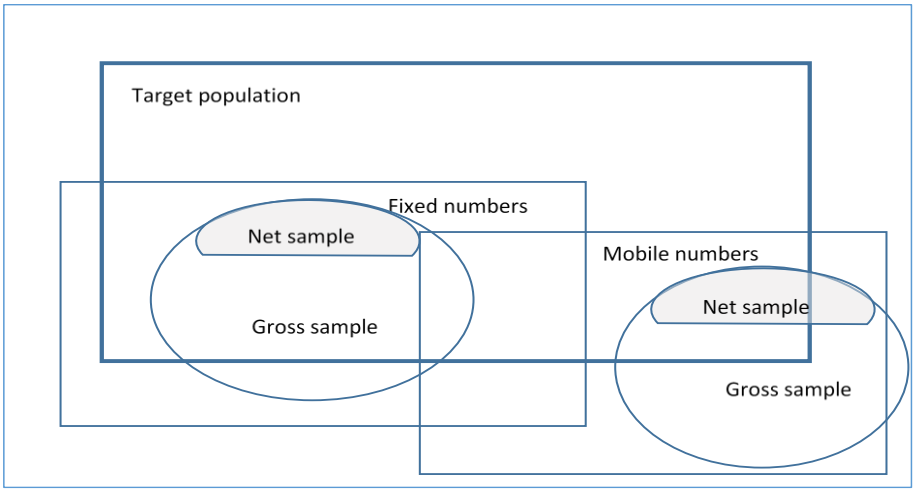
- The definition of a (greater) city is non-specific. The current definition is as follows: 'The greater city is an approximation of the urban centres when this stretches far beyond the administrative city boundaries" (EC 2020, p. 9). Moreover, there are no clear requirements to be met in order to include the city in the survey. This may influence the reliability of the results. Looking at Poland, for example, four cities were included in the survey: Białystok, Gdańsk, Kraków and Warsaw. There is however no logic pattern behind this choice. The ten largest cities in Poland listed by decreasing population size are: Warsaw, Kraków, Łódź,

([1]) It is also argued, that the use of different devices to participate in a survey (computer versus smartphone) in case of web-based surveys does not affect the answers with regard to sensitive topics (Toninelli, Revilla 2016).

Wrocław, Poznań, Gdańsk, Szczecin, Bydgoszcz, Lublin, and Białystok. In comparison, the ranking of city sizes by area: Warsaw, Krakow, Szczecin, Lodz, Wroclaw, Zielona Gora, Gdansk, Poznan, Swinoujscie, and Dabrowa Gornicza, does not correspond to the list of surveyed cities.

- In the QLEC survey, one of the requirements that must be satisfied in order to be eligible to respond is to be a resident of the city. Since the survey includes greater cities, which play the role of regional centres, the phenomenon of daily/weekly commuting is not negligible, when making a reliable assessment of the quality of life in selected cities. The QLEC survey does not take such cases into account.
- One of the criteria defining the target population is that 'Persons […] have sufficient command of (one of) the respective national/regional language(s) or English, which allows them to comfortably answer the questionnaire'. It is not clear-cut, whether this exclusion criterion is really part of the definition of the target population or whether this should be treated as an under-coverage problem. Are the persons living in a given city, who do not have a good command of the respective language (e.g. recent immigrants) be excluded from the target population?
- Concerning the conceptual definition of the target population, another question needs to be addressed: Are persons that do not own a landline or mobile telephone conceptually part of the target population or not?
- The gross sample was 16 800 for each city. When considering sample surveys, the varying response and eligibility rates in different cities would usually result in different net-sample sizes. Here there were exactly 700 responses in each of the cities. This indicates that the approach taken in the practical realisation of data collection was close to what is known as 'quota-sampling'.
- **Sample size.** The figures provided in the Evaluation report as well as additional analysis carried out on the basis of the micro-data provided (see Annex I) indicate that the standard errors are in the acceptable range. It can further be concluded that the net sample size of 700 units per city is sufficient for this type of the survey.
- Definitely, the largest problem of the QLEC survey is that of **non-response/under-coverage**. The 'dual frame approach' described in the Evaluation report can be depicted through the following graphical presentation:

Figure 1: 'Dual frame approach' in QLEC 2019



- **Response rate**. The response rate([2]) is generally defined as being the share of units that responded with respect to the eligible units) in the sample. By eligible units, we mean the units that are actually part of the target population. The fact is that this ratio is usually hard to estimate thoroughly with precision. The main problem is representing the non-contacted units, for which it is difficult to assess, which proportion thereof are really eligible. A number of different variants can be used to estimate the non-response rate. In the Evaluation report, three different types of response rates (referenced as Type 1, 3 and 4 – the definitions of those three types are provided in the Evaluation report) are calculated. As the estimation methods differ, depending on which units are taken to be eligible, their output values are related in a consistent manner, in the sense that Type 1 always yields the lowest and Type 4 the highest response rate. One must also note that, in the case of the QLEC survey, the three response rates are quite close together. It therefore does not matter much, which one is used. By any standard, it can be ascertained that the response rates are quite low as, in most of the cities, they are below 5 %. As pointed out also in the Evaluation report, the fact is that very low response rates do not automatically cause problems with representativeness, as it will be shown below. Still, it can be concluded that improving the response rate should be one of the goals of the further implementation.
- **Under-coverage**. The 'under-coverage problem' concerns that part of the target population that is not covered by phone lists. In the Evaluation report, there is no quantitative estimation of this phenomenon. There is also no information on the size of the sampling frame. It is therefore hard to estimate the extent of the problem. However, based on the description of the procedures applied to create the lists, it can be assumed that it is a non-negligible problem.
- **Over-coverage**. Theoretically, the phenomenon of over-coverage deals with the part of the sampling frame that falls outside the target population. In practice, only the units in the sample are observed; hence, the over-coverage can also only be estimated based on the sample data. Therefore, over-coverage as it is considered here represents the telephone numbers that have turned out to be ineligible - e.g. those of non-residents of the city, businesses. As reported in the Evaluation report, this phenomenon varied significantly among the cities, ranging from 62.5 % in Cluj-Napoca to 3 % in Reykjavík.
- **Representativeness**. As shortly described in the introduction to the micro-analysis (see Annex I), in the statistical surveys, non-response and coverage errors are usually the main source of the final net-sample's lack of representativeness, which can then result in the bias of the statistical results. In the implementation of statistical surveys, the non-response, over- and under-coverage rates are standard quality indicators that show how good the performance of certain processes like sampling frame creation, data collection/interviewing was. However, the fact is that those indicators do not show directly what the achieved rates' contribution to an eventual bias is and by how much those errors have distorted the population structure, when certain key socio-demographic characteristics are being measured. When looking at a city's gender-age distribution, which is certainly one of the most common socio-demographic variables, net- samples showed some deviances from the population structure, especially an under-representation of younger populations. However, given low eligibility and response rates, it still can be said that, with respect to these variables, in most cities, the sample can be claimed to be representative. Further analysis of the attained level of education however showed that there could be a much more outstanding problem with other socio-demographic variables. Namely, the analysis showed that the 'low-educated' part of the population could be very significantly under-represented. This fact can then lead to the more-general question of the under-representation of the

([2]) Most studies focus on measuring non-response and measurement error (Cehovin et al. 2018).

'lower classes' of the population. According to the results of the microanalysis, foreign-born persons are also a quite significantly under-represented (see results presented in Annex I).

- **Additional mode**. The observed low response and eligibility rates are more and more common phenomena in the case of telephone surveys, especially if lists of telephone subscribers are not readily available and if proxy sources need to be used. Although certain procedures can be applied to improve eligibility and response rates, the main means of improving representability is that of introducing an additional data collection mode. The following section assesses a number of considerations with regards to introducing the CAWI collection mode into the next implementation of the survey.

2 Revision of the QLEC methodology

2. Revision of the QLEC methodology

2.1 Introduction of the new data collection mode

A significant share of the drawbacks described in the previous section arises from the fact that only the CATI data-collection mode was used. In order to reduce these drawbacks and to improve the representativeness of the final net-sample, it is recommended to introduce another collection mode. What follows is a discussion of various aspects of the transition to a mixed-mode approach.

2.1.1 GENERAL CONSIDERATIONS ON THE CAWI COLLECTION MODE

It is suggested to introduce a mixed-mode of data collection by combining CATI and CAWI. Data collection methods should be flexibly distributed between both modes in an approximate ratio of 40% to 60 % respectively. There are no universal guidelines for the allocation between the two modes in the gross sample. The above proposal is based on the expert knowledge and experience of Statistics Norway (see Berg et al. 2020)[3]. The proposal is based also on the assumption that the distribution between the two modes should take into account the structure of the population according to age, education, and digital skills, as they influence the feasibility of conducting a CAWI survey.

There are two main prerequisites to introducing the CAWI mode effectively:

a) sufficient internet coverage among households;
b) sufficient digital skills of the respondents.

The first prerequisite does not seem to be a problem. The share of households that have a broadband connection has been growing constantly and it reached 89.4 % in the EU on average in 2020, with uptake ranging from 78.6 % in Bulgaria to 97.0 % in the Netherlands. Most of the New Member States (NMS) fell below the EU average, based on Digital Economy and Society Index (DESI) sub-indices([4]). Even the perceptible range between highest and lowest broadband coverage does not seem to be problematic, as the data pertain to country level. A more detailed analysis does however show that rural areas are lagging while larger cities usually perform above country average([5]) (see: Organisation for Economic Co-operation and Development (OECD) 2020; International Telecommunication Union (ITU) 2021). Furthermore, as many municipalities

([3]) In the Norwegian pilot CATI-CAWI SILC survey, the net CATI sample accounted for 64 % and the net CAWI sample for 36 % of total.
([4]) see https://digital-agenda-data.eu/charts/analyse-one-indicator-and-compare-countries
([5]) The rural/urban gap persists in terms of access to internet. – In Europe in 2019 the share of households with internet access accounted for 77.9 % in rural areas and 87.9 % in urban areas (ITU, 2021).

provide free-of-charge WiFi hotspots, access to the CAWI survey should be almost universal in the cities surveyed.
The other issue relates to the use of internet enabled by digital skills, which is largely driven by socio-demographic features of individuals: age, sex and education, as they are closely linked to digital skills. Gender and age gaps remain in Europe. With internet use, there is a slight bias towards the male part of the population (85.1 % versus 80.1 % of the female part). More significantly, there is a bias towards younger age groups; 96.2 % of individuals aged 15–24 were using Internet in 2019, compared to the European population average of 77.4 % (see ITU, 2021). It is also reasonable to expect significant differences in internet use, where different education levels are concerned. Thus, when the CAWI mode of the survey is introduced, priority should be given to those groups that may potentially be under-represented.
The major issue to be taken into account is however the variability of digital skills among countries and among surveyed cities, as low levels of digital skills may result in low response rates, in the CAWI mode. A general outlook on digital skills defined in line with Eurostat's approach reveals perceptibly large variations among countries. For example, in 2019, Romania, Latvia, Cyprus, Bulgaria, and Ireland recorded the highest shares of individuals with a low level of digital skills - between 42.7 % and 35.8 % respectively. On the other hand, Member States such as Finland, the Netherlands, Denmark, and Sweden have low shares of individuals with low digital skills and high shares of individuals with an above-basic level of digital skills. The practical implication of this observation is that cities located in countries with higher digital skills should aim at getting more surveys completed through using the CAWI mode, instead of using CATI. This recommendation is supported by the results from the Programme for the International Assessment of Adult Competencies (PIAAC) study. The PIAAC study shows that there is a relationship between income, educational level, age, and computer problem-solving skills, with the 16-24 age group having the highest skills in all countries participating in PIAAC (OECD 2013a, OECD 2013b). It has also been shown that digital skills are linked to labour market status, with better computer problem-solving skills in the group of the employed and lower skills among those who are not employed (Bjørkeng 2013).

Introducing the CAWI mode may potentially solve the problem of under-representation of youth in the sample, but the experience of Statistics Finland shows that the response rate in the 15–24 age group was disappointingly low (Larja, Taskinen, 2014). However, the use of smartphones is widespread, with even greater coverage in the group of persons aged 15-35. Providing access to the CAWI questionnaire in a form that is convenient for smartphone users thus seems to be crucial. To solve this issue, a relative over-representation of younger age groups, as well as women and low-educated persons, should be ensured in the gross sample.

CAWI – Computer Assisted Web-based Interviewing

CAWI – Computer Assisted Web-based Interviewing – is a self-administered mode of survey data collection, in which a respondent is invited to fill in the questionnaire, which is available online. In the CAWI mode, there is no interaction between the interviewer and the respondent. It is one of the least expensive and time-consuming modes of data collection and it can be used to reach respondents from many locations (countries, cities). However, there are also several drawbacks to this mode, the main ones being:

a) Potential coverage deficiencies, mostly related to differences in internet access and problems with the availability of e-mail addresses of the target population;

b) Questionnaire must be reasonably short in order to prevent a large drop-out of respondents.(de Leeuw 2008).

In a Mixed-mode survey, a combination of multiple methods of data collection is introduced in order to decrease the drawbacks of a single mode (e.g. CATI or CAWI), to reduce costs, and to increase the response rate and data quality. In other words, using a mixed-mode enables to compensate for the drawbacks of individual modes at an affordable cost (de Leeuw 2005; Mauz et

al. 2018; Vannieuwenhuyze 2014). The history of mixed-mode surveys dates back to the 1960s. Their use has however recently grown in popularity and scale, with the development of web-based data collection tools. There are two main implementation strategies of the mixed-mode: concurrent and sequential([6]) — (de Leeuw, Toepoel 2018).

2.1.2 POSSIBLE STRATEGIES FOR INTRODUCING THE MIXED MODE APPROACH

The two basic approaches in the implementation of the mixed mode are designated by the terms sequential and concurrent design. They can be described as follows:

- Sequential design. One mode (e.g. web) is used in the first phase for all selected persons. Non-respondents are then followed up in other modes. CAWI, which is in most implementations the cheapest mode, is usually taken as the first mode while, most frequently, CATI or Computer-Assisted Personal Interviewing (CAPI) is used in a follow-up phase.
- Concurrent design. Here, all modes are carried out simultaneously. There are two further possibilities for implementing this approach:
    - Selected persons have all modes at their disposal; the mode used is thus the respondent's choice. There are some considerations against this approach, arguing that an additional 'pre-phase' of mode selection can increase the refusal rate.
    - Two separate samples are taken, each of them being intended for a specific collection mode. Here, there are considerations that it is not an optimal option that each person should only have one possible mode at her or his disposal. In order to overcome this consideration (at least to a certain degree), two possible variants to this approach can be introduced:
        - The second mode can be introduced in the subsequent phase, with the aim of increasing the response rate in each of the modes.
        - When approaching the selected persons with the 'primary' mode, the alternative option can already be offered, in case they refuse the primary mode.

All of the options introduced above can include further variants. It is therefore clear that a wide range of different strategies can be defined in the search for the most suitable option for the survey in question. The final decision on the strategy depends on many different survey characteristics, such as the sources available for sampling frame creation, the time available for the interviewing phase, the survey budget, whether interviewing or self-completion is more suitable, etc. At this stage of considering the introduction of additional mode(s) into the QLEC survey, it is not possible to propose the optimal solution. This is mostly due to the fact that there are still too many open questions, especially where CAWI is concerned. Two possible options are therefore discussed here. Description is made of how we see the implementation of the concurrent and of the consecutive approach, and what, in our opinion, would be the most realistic, when taking into account the characteristics ascertained through the previous waves of surveys. The assumption is that the second mode will be CAWI, i.e. a web-administered self-completion survey. Hence, the CATI-CAWI mixed mode is discussed:

- Sequential design. In the first phase, the web survey is carried out with the aim of collecting a target number of responses, e.g. 40 % (280 responses) of the entire sample. It must be

([6]) More details on these strategies are given in Section 2.1.2.

pointed out that in the CAWI mode control over the collected responses is quite limited, so one cannot expect to obtain the exact target number of responses. The final results are expected to vary around the target number, though. In the second phase, the telephone mode is used, with the aim of completing the set of responses, to reach the total target (e.g. 700 responses). At this stage, the sample structure (e.g. by age-gender groups), which will very likely have been 'distorted' during the 'CAWI phase', can be rebalanced, in order to improve the representativeness of the final net sample.

- Concurrent design. Two parallel, independent samples are selected, one for each mode. Their sizes are calculated according to the target share of each mode in the net sample as well as according to the assumed response rate for each of the modes. If, for instance, the target share is 60 % CATI: 40 % CAWI, and the assumed response rates are 10 % for CATI and 5 % for CAWI, the gross sample sizes would be:

$$\text{CATI} = 700 * 0.6 * 1/0.1 = 4200$$

$$\text{CAWI} = 700 * 0.4 * 1/0.05 = 5600$$

Of course, since it is not possible to control the final net size of the CAWI sample, the final net sample cannot be fixed, but it varies around the target number.

To our best knowledge, no databases that contain both the telephone numbers and the e-mail addresses of the same persons are available. Therefore, to pool the sample, concurrent design seems to be more appropriate([7]). Sampling for CATI may be done in way similar to that used for the 2019 survey, while CAWI sampling should take advantage of:

- e-mail addresses of respondents collected by city authorities or other public institutions;
- social media, through which the QLEC survey can be promoted. Using this approach, the link to the CAWI questionnaire could be uploaded on the social-media fanpages administered by city authorities.

Both with sequential and concurrent design, a mechanism that prevents a person from participating in both CATI and CAWI mode must be implemented.

2.2 Revision of the sampling methodology

Revision of the sampling methodology is very strongly connected to the introduction of the additional data collection mode. Given that the introduction of the CAWI collection mode still has many open issues that need to be resolved during the course of concrete implementation, some of the aspects of sampling methodology, especially those related to the sampling frame construction, will also be left open. Only a restricted number of recommendations will therefore be formulated as direct instructions to be followed, while some of the recommendations will be provided on a quite general level.

2.2.1 SAMPLING FRAME CONSTRUCTION

- Although the criteria for the inclusion of a given greater city in the sample are vague, it is recommended to stick to the list of cities used in the QLEC 2019 wave.
- The conceptual definition of the target population needs to be more precise. We propose to keep only two conditions defining target population:

([7]) As Becker (2021) showed, sequential design that uses a 'push to web' strategy does not bring the expected results.

- Residence in one of the 83 European cities;
- Persons aged 15+ that live in a private household, which means that the target population will exclude prisoners, residents of retirement homes, etc. who are hard to reach via a telephone survey.

If the idea of including persons who are residents of other municipalities, meaning that their permanent address being outside the surveyed city, but whose private/professional activities are pursued mainly in the surveyed city (see Section 2.1) were to be accepted, this condition should be added to the definition of the target population.

All other groups that cannot be reached through CATI([8]) () nor through CAWI([9]) () should be treated as forming part of an 'under-coverage problem', the extent of which should be reduced as much as possible.

- In introducing the CAWI mode, the main issue is the construction of the sampling frame for this mode. Since many countries are involved in this survey, it is difficult to predict which sources should be used. In general, one can however envisage two options for the physical realisation of the sampling frame:
  - List of persons with addresses. The invitation letter, containing the respective web link, would be sent to those persons, inviting them to complete the web survey. Such lists could probably only be obtained through cooperation with government administration bodies (e.g. National Statistical Institute).
  - List of e-mail addresses. How to obtain this list will certainly be a big challenge. Probably the only feasible option is to use social media and other publicly available sources.
- What is missing in the Evaluation report of the last survey wave is the estimation of the under-coverage problem, i.e. an estimation of the part of the target population that is not covered by the sampling frame(s). It is hence recommended that at least a rough estimation is made of the share of the target population covered by the three established sampling frames.

2.2.2 SAMPLING

- Sample size. As implied by the findings of the microdata analysis, the sample size of 700 responses/per city is sufficient, where the precision of the statistical results is concerned. However, the following points shall be raised:
  - If the idea of including commuters (see Section 2.5) is accepted, the target net sample size should be increased accordingly.
  - A fixed sample size applied to each city is not, in our opinion, a feature that should necessarily to be kept. This target number should rather be set as a minimal requirement and, if additional responses were to increase the representativeness of certain groups, larger numbers of responses would be welcomed.
- Allocation. Once the CAWI mode is introduced, there will be a triple-frame, instead of a double-frame situation. The gross target sample will therefore need to be allocated to three

([8]) For instance, persons that do not have a telephone, persons that do not have sufficient command of one of the respective national/regional language(s) or English.
([9]) For instance, persons that do not use internet.

strata: web, fixed phones, mobile phones. The allocation between CAWI and CATI should be performed in the first stage. Some recommendations on this allocation are already provided in Section 2.1.1. For the 'CATI part', the allocation procedure between fixed and mobile phones can remain the same as in the 2019 wave.

- Representativeness. As implied by the results of the micro-data analysis, the final sample of the previous implementation suffers from the problem of under-representing certain segments of society — e.g. low-educated persons, foreign-born persons. Therefore, in addition to the age-group structure, these groups should also be taken into account when controlling the net-sample structure.

2.3 Improvements of the response rates

Certain actions are recommended to be carried out, to ensure higher response rates than was the case in the last survey implementation:

- A special emphasis should be placed on managing the CAWI data collection and on attracting the respondents to use this mode of survey delivery. Ways of increasing the response rate for CAWI are:
    - A standard way of promoting CAWI is to prepare and send the letters in advance. Up-to-date experience shows that the notice given should not be too short, and the letter should contain information about the purpose of the survey, the importance of participating in the survey, as well as instructions on how to log in and security measures. The advance letters may also include information — a mild warning — that if the respondent does not follow the CAWI mode, she/he will be contacted by the interviewer and surveyed in the CATI mode.
    - Advance letters should preferably be sent on a Friday, making it possible for respondents to complete the survey online during the weekend (see Gravem et al. 2014),
    - Usually two reminder letters are sent — the first, 3–5 days after invitation, the second approx. 10 days after the initial invitation. Reminders should always include information on how to log in to the web survey. It is not advisable to send more than two reminders, as it may affect the survey participation rate.
    - If notifications are sent via SMS, the smartphone and tablet compliance of the questionnaire should be ensured, as some respondents only have internet access through smartphone-type devices.

- Also at general survey level, including both modes, certain actions could be implemented, in order to improve response rate:
    - Some effort and budget resources are expected to be put into promotional activities. The survey's implementation should be advertised through the media. The aim of promoting the QLEC is to distinguish it from market polls. Promotion through the appropriate presentation of the previous survey's results would be a good option.
    - Some mixed-mode surveys have used incentives to influence the response rate. In general two sorts of incentives can be used:
        - Monetary incentives. Some money is offered to be paid.
        - Non-monetary incentives – gifts,lottery.

Incentives can further be distinguished by:

- Pre-paid incentives. All the participants receive gifts, usually in advance with the invitation letter.
- Post-paid incentives. Only respondents receive money, a gift or they are included in the drawing of a lottery.

However, as the impact of incentives on the response rate is ambiguous across different surveys (e.g. Singer 2018), it is not recommended to introduce such an approach into the forthcoming wave of the QLEC survey.

2.4 Revision of the questionnaire and its structure

Introducing the mixed-mode requires adapting the questionnaire that was previously used in CATI mode for the 2019 wave. In case a CAWI-CATI mixed-mode is used, the ideal solution is to follow Dillman's (2000) () guidelines([10]) and to provide a 'unimode' questionnaire (an approach preferred in most mixed-mode surveys). This reduces the problem of the mode effect. However, since the CAWI questionnaire shall potentially be adjusted to mobile devices like smartphones or tablets, it may be required to take advantage of the specific mode's questionnaire design. This decision is to be taken once pilot testing of the questionnaire in both modes — CAWI and CATI — has been completed, and mode effects have been analysed.

Proposed amendments to the questionnaire take into account the large cognitive burden resulting in a high dropout rate around the Q2–Q4 block of the initial CATI questionnaire. This is a less pronounced problem, when the CAWI mode is used, as that mode provides the possibility of 'freezing' the questionnaire and of answering the remaining questions later. Since it is virtually impossible to avoid grids in the questionnaire, the following sequence of questions is recommended([11]) :

Table 1: List of sequence of questions recommended

CATI questionnaire	CAWI questionnaire
D1	D1
D2	D2
D3 sequence	D3.aa
Q1	D3 sequence
Q5	Q1
Q6	Q5
Q2	Q6
Q4	Q2
Q7	Q4
Q8	Q7
Q9	Q8
Q10	Q9
Q11	Q10
Q12	Q11
Q13	Q12
Q3	Q13

([10]) see also de Leeuw 2005; 2018
([11]) The question numbers refer to the numbers as used in the CATI questionnaire, QLEC 2019 wave.

CATI questionnaire	CAWI questionnaire
Q14	Q3
D5	Q14
D6	D5
D7	D6
D9	D7
D9b	D9
D8	D9b
D10	D8
D11a	D10
D11	D11a
D12	D11
D13	D12
D14*	D13
D15*	Q15a
Q15a	Q15
Q15	

* Questions concerning having a mobile/landline telephone – D14 and D15 – are seemingly not relevant in the CAWI mode.

Question Q3 has intentionally been moved towards the end of the main section of the questionnaire, as it was the one with the highest break-off rate in the QLEC 2019 wave.

The following changes in the wording of questions for the CAWI mode are proposed:
Q1. In general, please mark how satisfied you are with each of the following issues in your city or area
Q6. Concerning public transport in your city, based on your experience or perceptions, please mark to what extent you agree with each of these statements
Q2. Please mark to what extent you agree with each of these statements
Q4. Overall, how satisfied are you with
Q13. Please mark to what extent you agree with each of these statements about the local public administration in your city
Moreover, screening question **D3.aa** — 'Which country do you live in?' is added.

It also seems reasonable to introduce the following changes, in both CAWI and CATI modes:
Q8. Within the last 12 months, was any money or property stolen from you or another household member, in your city?
D12. (open question) What is your current job position/occupation?

The proposal for the CATI/CAWI questionnaires is presented in Annex III.

2.5 Other methodological issues

1. Introducing a mix-mode approach does not solve the problem of measuring respondents' occupations in line with the ISCO-08 classification, at major group level. Respondents usually do not know the ISCO-08 classification, and matching the individual's job position with a major ISCO occupational group is not intuitive. Two options are available to cope with this issue:

a. to provide examples of names of occupations/job positions linked to each response option (occupational group) — through an attachment to answer categories in the CAWI mode, and through suggestions from the interviewer in CATI mode;

b. to make this an open question — each respondent puts the name of the job/occupation she/he performs, all answers are then re-coded into major ISCO-08 groups. Re-coding may be done on the basis of prepared dictionaries, or through using the Artificial Intelligence (AI) solution developed by CEDEOP to analyse the content of on-line job postings. The latter option is highly recommended.

2. In the forthcoming waves of the QLEC survey, it would be reasonable to work on extending the scope of eligible persons by including those who are residents of other municipalities and their permanent address is outside the surveyed city, but whose private/professional activities are pursued mainly in the surveyed city (e.g. commuters). The eligibility criteria may be:

    a. being a resident in the surveyed city, or

    b. being a resident in another municipality (non-resident), but staying in the surveyed city for minimum 10 hours per day, for at least 4 days in a week, on a permanent basis for at least 3 months prior to the survey. Eligibility would be screened by providing the postcode of the person's temporary place of residence or the postcode of the organisation in which the person works/has other activity in the surveyed city.

Due to the limited availability of the phone numbers/e-mail addresses of non-residents, this group would mainly be surveyed in the CAWI mode.
To assure full result comparability between the 2019 and future editions of the survey, the data analysis should be done separately for residents and non-residents, as well as for the entire surveyed population. Expanding the sample number to cover 700 residents as in 2019 edition, and an additional number of non-residents, would be desirable.

3 Conclusions and recommendations

3. Conclusions and recommendations

This concluding section summarises the practical actions that are recommended to be implemented in the forthcoming wave of the QLEC survey in terms of interviewing mode, sampling approach and questionnaire design. These proposals are based on expert analysis of previous QLEC waves, with a special focus on the 2019 wave. The rationale for these actions is detailed in the paper's main sections. They are:

- Introduction of the mixed-mode of data collection by combining CATI and CAWI methods, with the flexible distribution between both modes in 40 %: 60 % parts of the entire sample. The distribution provided is only an indicative proposal, based on the experiences from National Statistical Institutes, and it can be a subject of further discussion.
- Introducing the CAWI mode may potentially solve the problem of under-representation of the younger part of the population, but it is still to be expected that response in age group 15–24 will be low. To overcome this issue, two additional measures can be taken:
    - Create a version of the CAWI questionnaire that can be used on smartphones;
    - Over-sample the age groups, for which a lower response rate is expected — e.g. young persons, low-educated persons.
- The mixed-mode strategy should be determined before the actual sample design is determined and the sample is selected. The two basic approaches are:
    - **Sequential design**. Data collection using the two modes will be carried out in two consecutive phases. In the first phase only one mode will be used, and in the second phase, only the other.
    - **Concurrent design.** Both modes are carried out simultaneously in the same phase.

  Additional details about the possible concrete implementation of both approaches are provided in Section 2.1.2.

- The conceptual definition of the target population needs to be more precise. The proposal is to keep only two conditions that define target population:
    - Resident of one of the 83 European cities;
    - Persons aged 15+ that live in a private household.

  Although expanding the sample by including commuters seems potentially valuable, it is not recommended to introduce such a concept in the forthcoming QLEC wave.

- The geographical scope of the survey — the list of greater cities — shall remain unchanged.
- Creating the sampling frame for the 'CAWI sample' is of crucial importance. Two options can be foreseen for the physical realisation of the sampling frame:
    - List of persons with addresses. The invitation letter with the respective web link would be sent to those persons, inviting them to complete the web survey. This should be a priority option as it enables exact location of the person.

    - List of e-mail addresses. Similar to the case with mobile phones, also in this case a big challenge will be to limit the list to city inhabitants only.
- The sample size of 700 responses/per city is considered to be sufficient, where the precision of the statistical results is concerned. The fixed sample size/per city is in our opinion not a feature that it is necessary to keep. 700 responses could be set as a minimal requirement.
- The allocation procedure of the 'CATI part', between fixed and mobile phones, can remain the same as in the previous wave. A procedure for the allocation between CAWI and CATI is to be introduced. Some suggestions for this procedure are provided in Sections 2.1.2. and 2.2.1.
- In addition to the age-group structure, some other population groups, such as low-educated persons and foreign-born persons, should be taken into account when controlling the net-sample structure.
- Special focus should be placed on managing the CAWI data collection and on attracting respondents to use this mode of survey delivery.
- In order for respondents to distinguish between the QLEC survey and market polls, we recommend to put some effort and budget resources into promotional activities. The survey's implementation should be advertised through different media, and in cooperation with city authorities.
- Advance letters, preferably sent on a Friday, should not be too short, and they should contain information about the survey's purpose, the importance of taking part in the survey, the instructions on how to log in, and security measures. It is advised to send two reminder letters: the first after 3-5 days after invitation, the other approx. 10 days after the initial invitation (see Section 2.3. of the paper).
- The questionnaire should follow Dillman's guidelines for unimode questionnaires, to reduce the problem of the mode effect. However, it may be required to take advantage of the questionnaire's mode-specific design, in the case of CAWI questionnaire, which should also be made available for mobile devices. Such decisions are to be taken once pilot testing of the questionnaire in mixed-mode is complete (see Annex III for a proposal of CATI and CAWI questionnaires).
- The question relating to the respondent's occupation should be an open question. To re-code answers into ISCO-08 major occupational groups the tool developed by CEDEFOP can potentially be used.

References

References

1. Becker, R. (2021), *The Effects of a Special Sequential Mixed-mode Design, and Reminders, on Panellists' Participation in a Probability-based Panel Study, Quality & Quantity*, https://doi.org/10.1007/s11135-021-01126-6

2. Berg, N., Snellingen Bye, L., Rossbach, K., Lande With, M. (2020), *Report from the Norwegian SILC web pilot 2019*, Statistics Norway

3. Bjørkeng, B. (ed.) (2013), Ferdigheter i voksenbefolkningen. Resultater fra den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC), Statistics Norway

4. Cehovin, G., Bosnjak, M., Lozar Manfreda, K. (2018), 'Meat-Analyses in Survey Methodology. A Systematic Review', *Public Opinion Quarterly*, 82(4): 641–660.

5. Dillman, D.A. (2000), *Mail and Internet surveys: The Tailored Design Method*, New York: John Wiley

6. European Commission (2020), *Report on the Quality of Life in European Cities, 2020*, Luxembourg: Publications Office of the European Union

7. European Commission (2020a), *Perception Survey on the Quality of Life in European Cities 2019,* Evaluation Report, Brussels.

8. Graven, D.F., Bashus, T., Lagerstrøm, B.O. (2014), 'The organisation of mixed-mode data collection', *ESSnet Data Collection for Social Surveys Using Multiple Modes*,

https://ec.europa.eu/eurostat/cros/system/files/D3_4_Orga_of_mixed_mode.pdf

9. Lamanna, C, Hachhethu, K, Chesterman, S, Singhal, G, Mwongela, B, Ng'endo, M, et al. (2019), 'Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya', PLoS ONE 14(1): e0210050 https://doi.org/10.1371/journal.pone.0210050

10. Larja, L., Taskinen, P. (2014), Mixed mode in LFS: questionnaire design and mode-effects, mimeo

11. De Leeuw, E.D. (2018), 'Mixed-Mode: Past, Present, and Future', *Survey Research Methods*, 12(2), pp. 75–89.

12. De Leeuw, E.D. (2008), 'Choosing the Method of Data Collection' [in]: E.D. de Leeuw, J.J. Hox, D.A. Dillman (eds.), *International Handbook of Survey Methodology*, Taylor & Francis Group/Lawrence Erlbaum Associates, pp. 113–135.

13. De Leeuw, E.D. (2005), 'To Mix or Not to Mix Data Collection Modes in Surveys', *Journal of Official Statistics*, 21(2), pp. 233–255.

14. De Leeuw, E.D., Toeopel, V. (2018), 'Mixed-Mode and Mixed-Device Surveys' [in:] D.L.

Vannette, J.A. Krosnick (eds.), *The Palgrave Handbook of Survey Research*, Palgrave Macmillan, pp. 51–61.

15. Mauz, E., von der Lippe, E., Allen, J., Schilling, R., Muters, S., Hoebel, J., Schmich, P., Wetzstein, M., Kamtsiuris, P., Lange, C. (2018), 'Mixing Modes in a Population-based Interview Survey: Comparison of a Sequential and a Concurrent Mixed-mode Design for Public Health Research', *Archives of Public Health*, pp. 76(8).

16. ITU (2021), Digital Trends in Europe 2021. *ICT Trends and Developments in Europe*, 2017-2020, International Telecommunication Union

17. OECD (2013a), *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris, OECD Publishing

18. OECD (2013b), *The OECD Survey of Adult Skills International Report* (Volume II). Paris, OECD Publishing

19. OECD (2020), *OECD Regions and Cities at Glance*, OECD Publishing, Paris, https://doi.org/10.1787/959d5ba0-en

20. Rauhut, D., Humer, A. (2020), 'EU Cohesion Policy and spatial economic growth: trajectories in economic thought', *European Planning Studies*, 28(11), pp. 2116–2133,

DOI: 10.1080/09654313.2019.1709416

21. Romanelli, M., Metallo, C., Agrifoglio, R., Ferrara, M. (2019), 'Cities, Smartness and Participation Towards Sustainability' [in]: A. Lazazzara, R. Nacamulli, C. Rossignoli, S. Za (eds), *Organizing for Digital Innovation. Lecture Notes in Information Systems and Organisation*, vol 27. Springer, Cham.

https://doi.org/10.1007/978-3-319-90500-6_10

22. Singer, E. (2018), 'The Use and Effects of Incentives in Surveys' [in:] D.L. Vannette, J.A. Krosnick (eds.), *The Palgrave Handbook of Survey Research*, Palgrave Macmillan, pp. 63-70.

23. Toninelli, D., Revilla, M. (2016), 'Smartphones vs PCs: Does the Device Affect the Web Survey Experience and the Measurement Error for Sensitive Topics? - A Replication of the Mavletova & Couper's 2013 Experiment', *Survey Research Methods*, 10(2): pp. 153–169. https://doi.org/10.18148/srm/2016.v10i2.6274

24. Vannieuwenhuyze, J.T.A. (2014), 'On the Relative Advantage of Mixed-Mode versus Single-Mode Surveys', *Survey Research Methods*, 8(1), pp. 31–42.

Annex I: Micro-analysis of the QLEC 2019 survey output

Introduction

This annex aims to present the main findings of the micro-analysis of the 2019 implementation of the Quality of Life in European Cities Survey (hereinafter QLEC). The present analysis provides support for the first deliverable, where a broader analysis of the main elements of survey methodology will be carried out and some recommendations for future improvements will be given.

The present analysis uses some elements of the so-called European Statistical System (ESS) Standard for Quality Reports Structure (ESQRS) to indicate the quality of the survey data and survey processes concerning different quality dimensions. However, before we deal with the quality dimensions, we provide some basic information about the analysed micro-data.

Micro-data

Since the net-sample size was exactly 700 responses per city and the survey included 83 cities, the micro-data includes 58,100 records. The dataset includes 79 variables that can be roughly divided into three subsets:

- Data (variables) from the questionnaire. Large share of these variables are categorical variables, coming out from the questions with a four-point scale. One part of the questionnaire data is devoted to the socio-demographic data, e.g. age, gender, education.
- Data from the sample: City, Local Area Units (LAUs), Telephone type (fixed-mobile).
- Data coming out from the weighting procedure: design weight, calibrated weight, population weight.

Accuracy

As a quality component, accuracy deals with assessing the total survey error, most commonly disaggregated into two main components: standard error and bias.

Standard error

By its basic definition, standard error is a standard deviation of the theoretical distribution of the statistical estimate. In other words, this is the error of the estimates, which is caused by the random factors. In practice, it is estimated as the square root of the sampling error. In our case, we estimated the standard error for the city estimates of the proportions, deriving from all the ‘opinion questions’ with the 3-5-point scale. We used SAS procedure PROC SURVEYFREQ to calculate the standard errors, taking into account the survey design as described in the Evaluation report, namely one-stage stratified sample, where the stratification was determined by the following variables: City, Local area unit, Type of telephone (mobile/fixed). The category 99 (‘Don't know/No Answer/ Refuses’) was

transformed into Null values, so these values were excluded from the survey analysis. The following table shows 10 highest standard errors for the estimated proportions on the city level.

Table 2: Standard error of the estimated proportions

City	Variable	Category	Percent	StdErr
Antalya	q3_03	A good place to live	56.6	2.7
Antalya	q3_03	Not a good place to live	43.4	2.7
Antalya	q1_04	Rather satisfied	40.7	2.7
Antalya	q13_5	Strongly disagree	36.8	2.6
Miskolc	q3_03	Not a good place to live	29.6	2.6
Miskolc	q3_03	A good place to live	70.4	2.6
Riga	q3_03	A good place to live	62.5	2.6
Riga	q3_03	Not a good place to live	37.5	2.6
Antalya	q1_03	Rather satisfied	36.1	2.6
Leipzig	q13_5	Somewhat disagree	48.5	2.6

The coefficients of variations are very suitable indicator to set-up the threshold value for sufficient precision (e.g. CV=10 %) in the case of totals and proportions, but due to non-symmetry for p,$1-p$, they can't be used in the case of proportions. However, we can still determine an analogue criterion, using the following logic: As it is known, considering only the estimated proportion the highest standard error is for value $\hat{p} = 0.5$. If we transform the 'CV threshold' of 10 % to 'standard error threshold', we get the threshold $se_t = 0.05$. We can use then this 'standard error threshold' for all the estimated proportions. Considering this criterion, we can say that all the standard errors are well below the threshold. Even if we take a stricter baseline criterion CV=5 %, which leads to 'standard error threshold' 0.025, we can see from the complete set of standard errors, that 99 % of estimated standard errors are below this threshold. General conclusion of this analysis can be that the estimates are of sufficient precision and the sample size of 700 per city is large enough when the precision of the statistical results is in question.

Bias

The bias of the statistical estimates is the phenomenon that is much more difficult to estimate than standard errors. This is because of the many factors that can cause bias and many of them are not easy to discover. Therefore, what is usually done in practice is that we do not estimate bias directly, but we estimate different types of non-sampling errors, which are most common sources of bias.

Item non-response

We treated as item non-response values 99 ('Don't know/No Answer/Refuses') for all variables. Missing values that also appear in some cases (e.g. occupation for non-occupied persons), were considered as 'eligible missing' and were not counted as item non-response.

If we calculate the item non-response on the overall dataset, considering all cities together, the highest item non-response reaches 20 %. The following table shows 10 highest values of item non-response.

Table 3: Item non-response on the whole dataset level (in %)

Variable	Description	Item non-response rate (%)
q13_5	There is corruption in my local public administration	20.2
q3_03	Is the city where you live a good place to live for gay or lesbian people.	15.4
q1_07	Satisfaction with Schools and other educational facilities	13.9
q1_03	Satisfaction with Sport facilities such as sport fields and indoor sports halls	13.4
q2_02	It is easy to find a good job in my city.	11.0
q13_4	Information and services of my local public administration can be easily accessed online	10.3
q13_1	Satisfaction with the amount of time it takes to get a request solved by my local public administration	9.9
q2_05	It is easy to find good housing in my city at a reasonable price.	9.6
q3_04	Is the city where you live a good place to live for Immigrants from other countries.	9.4
q3_02	Is the city where you live a good place to live for Racial and ethnic minorities.	8.1

If we calculate the item non-response on the city level, the highest item non-response rates range over 40 %. The following table shows 10 highest values of item non-response on the city level.

Table 4: Item non-response on the city level (in %)

City	Variable	Description	Item non-response rate (%)
Miskolc	q3_03	Is the city where you live a good place or not a good place to live for gay or lesbian people.	42.6
Leipzig	q13_5	There is corruption in my local public administration	41.4
Essen	q13_5	There is corruption in my local public administration	38.4
Burgas	q3_03	Is the city where you live a good place or not a good place to live for gay or lesbian people.	37.9
Sofia	q3_03	Is the city where you live a good place or not a good place to live for gay or lesbian people.	36.1
Rotterdam	q13_5	There is corruption in my local public administration	35.6
Berlin	q13_5	There is corruption in my local public administration	34.6
Rostock	q1_07	Satisfaction with schools and other educational facilities.	34.0
Krakow	q13_5	There is corruption in my local public administration	33.6
Hamburg	q13_5	There is corruption in my local public administration	32.9

Coverage errors

As usually in telephone surveys, coverage is, in our opinion, the crucial element, when the quality of the QLEC results are in question. There is certainly under-coverage of the sampling frame, but it is very difficult to directly estimate how large it is and how large influence it has on the bias of the estimates. What we do here is to estimate how biased is the final sample according to a few socio-demographic variables (age, gender and education). We have to stress again that even if certain 'distortion' of the net-sample structure is indicated, it is almost impossible to see to what extent this is due to coverage problems and not to other factors, like non-response.

Gender and age groups

We don't have direct data on the age, gender population structure in the selected cities, but we will take advantage of the fact that, according to the Evaluation report, during the weighting procedure, weights were calibrated to the gender, age population structure. Hence, we can assume that the population weights (variable weight population) reflect the actual population structure. We, hence, calculated the difference in the age, gender structure if we take into account sample number (sample structure) and if we take into account population weights (population structure). We calculated this difference first separately in gender and age-group, and then by taking gender age-group as cross-cutting domain. The following table shows the 10 highest absolute differences in each domain.

Table 5: Sample deviation from population structure (gender and age group)

City	Domain	Domain_value	Sample	Population	Absolute difference
Skopje	Gender	Female	0.44	0.51	0.07
Skopje	Gender	Male	0.56	0.49	0.07
Košice	Gender	Male	0.43	0.47	0.04
Košice	Gender	Female	0.57	0.53	0.04
Oviedo	Gender	Male	0.50	0.46	0.04
Oviedo	Gender	Female	0.50	0.54	0.04
Ljubljana	Gender	Female	0.56	0.53	0.04
Ljubljana	Gender	Male	0.44	0.47	0.04
Ostrava	Gender	Female	0.56	0.52	0.04
Ostrava	Gender	Male	0.44	0.48	0.04
Malmo	Age_group	25–34	0.17	0.23	0.06
Cardiff	Age_group	55–64	0.18	0.12	0.06
Malmo	Age_group	55–64	0.18	0.12	0.06
Cardiff	Age_group	25–34	0.14	0.20	0.06
Skopje	Age_group	45–54	0.22	0.17	0.06
Zurich	Age_group	35–44	0.19	0.13	0.06
Stockholm	Age_group	55–64	0.18	0.13	0.06
Valletta	Age_group	25–34	0.13	0.18	0.05
Rennes	Age_group	25–34	0.14	0.19	0.05
Athina	Age_group	55–64	0.19	0.14	0.05
Skopje	Gender*Age_group	Male*45–54	0.12	0.07	0.04
Tirana	Gender*Age_group	Male*25–34	0.20	0.15	0.04

City	Domain	Domain_value	Sample	Population	Absolute difference
Cardiff	Gender*Age_group	Female*25–34	0.07	0.11	0.04
Košice	Gender*Age_group	Female*55–64	0.12	0.09	0.04
Skopje	Gender*Age_group	Male*35–44	0.15	0.11	0.03
Podgorica	Gender*Age_group	Female*25–34	0.07	0.10	0.03
Zurich	Gender*Age_group	Female*35–44	0.11	0.08	0.03
Cardiff	Gender*Age_group	Male*55–64	0.10	0.06	0.03
Malmo	Gender*Age_group	Female*25–34	0.09	0.13	0.03
Budapest	Gender*Age_group	Male*45–54	0.11	0.07	0.03

We can conclude, from a general viewpoint, that there are some cases where the deviation from the population structure is a bit higher, but still (according to the fact that telephone survey is in question), the differences are still quite moderate. Especially, if we take into account the fact that the differences exceeds 2 percentage points only in 20 % of the cases.

Education

Since education was not one of the variables, used or calibration, we couldn't use the weighted structure as the population structure. In fact, we don't have a population structure according to the education and we can't provide a comprehensive analysis with this regard. We therefore provide only one example, taking publicly available population for education structure for Ljubljana, Slovenia and compare it to the net-sample structure. The educational structure for Ljubljana is available at https://pxweb.stat.si/SiStatData/pxweb/en/Data/-/05G2014S.px. The following table shows the structure concerning the three categories of education attained.

Table 6: Sample vs population – education
(in %)

Education	Sample	Population
Below upper secondary education	4.4	16.9
Upper secondary education	35.7	47.6
Tertiary education	59.8	35.5

What it is clearly outstanding in this comparison is a significant underrepresentation of the low-educated part and overrepresentation of the high-educated part of the population. For the other cities, we show the percentage of the first category (International Standard Classification of Education (ISCED = 0.1) in the sample, just to indicate whether this 'underrepresentation' can be a general problem.

Table 7: Share of persons with 'basic' education in the sample

City	Count	Percent
Graz	62	8.9
Wien	50	7.1
Antwerpen	54	7.7
Bruxelles	61	8.7
Liege	69	9.9
Burgas	43	6.1
Sofia	40	5.7
Zagreb	63	9.0

City	Count	Percent
Lefkosia	77	11.0
Ostrava	46	6.6
Praha	54	7.7
Aalborg	46	6.6
Kobenhavn	42	6.0
Tallinn	108	15.5
Helsinki	32	4.6
Oulu	26	3.7
Bordeaux	113	16.3
Lille	144	20.7
Marseille	123	17.7
Rennes	82	11.8
Strasbourg	104	15.0
Paris	104	14.9
Berlin	101	14.4
Dortmund	134	19.2
Essen	118	17.0
Hamburg	142	20.3
Leipzig	118	17.0
Munchen	113	16.2
Rostock	118	16.9
Athena	88	12.6
Irakleio	79	11.3
Budapest	70	10.0
Miskolc	90	12.9
Dublin	59	8.7
Bologna	65	9.3
Napoli	74	10.6
Palermo	87	12.4
Roma	83	11.9
Torino	86	12.3
Verona	68	9.7
Vilnius	60	8.6
Luxembourg	136	19.5
Riga	40	5.7
Valletta	149	21.3
Amsterdam	76	10.9
Groningen	90	12.9
Rotterdam	96	13.8
Bialystok	70	10.0
Gdansk	54	7.7

City	Count	Percent
Krakow	77	11.0
Warszawa	53	7.6
Braga	142	20.3
Lisboa	91	13.1
Bucurest	101	14.5
Cluj	109	15.6
Piatra Neamt	129	18.5
Bratislava	50	7.1
Košice	55	7.9
Ljubljana	31	4.4
Barcelona	91	13.1
Madrid	79	11.3
Malaga	77	11.2
Oviedo	70	10.1
Malmo	52	7.4
Stockholm	44	6.3
Belfast	116	16.7
Cardiff	126	18.3
Glasgow	119	17.1
London	116	16.7
Manchester	126	18.4
Tyneside conurbation	131	18.8
Reykjavik	70	10.0
Oslo	38	5.4
Geneve	49	7.1
Zurich	63	9.1
Tirana	90	12.9
Skopje	81	11.6
Podgorica	34	4.9
Beograd	32	4.6
Ankara	85	12.1
Istanbul	106	15.1
Antalya	90	12.9
Diyarbakir	87	12.4

Activity status

We provide comparison between sample and population or the 'activity status structure' again for Ljubljana. The activity status structure for Ljubljana is available at https://pxweb.stat.si/SiStatData/pxweb/en/Data/-/05G3016S.px/

The following table shows the structure for the most frequent 4 categories.

Table 8: Differences in share of foreign-born inhabitants

Activity status	Sample (in %)	Population (in %)
At work	58.5	50.0
Unemployed	4.2	5.0
In full-time education	4.4	12.3
Retired	26.9	27.3

Share of foreign-born inhabitants

In the QLEC survey there was also a question, asking country of birth of the respondent. Based on the responses to this question, we have calculated the share of foreign-born persons in the sample and compared it to the data disseminated by Eurostat.

Table 9: Outliers by Tukey method

City	Year	Eurostat (in %)	QLEC (in %)
Barcelona	2019	21.5	14.9
Bratislava	2019	5.8	2.0
Bruxelles	2019	45.8	24.6
Budapest	2019	10.5	2.3
Glasgow	2018	15.4	3.9
Helsinki	2018	14.3	5.6
Ljubljana	2018	20.2	4.3
Madrid	2019	20.2	9.7
Manchester	2018	28.6	7.6
Paris	2017	25.0	12.1
Riga	2019	18.0	3.7
Sofia	2018	3.4	1.0
Stockholm	2018	25.6	8.0
Tallinn	2019	20.9	3.0
Zagreb	2019	17.7	6.0
Zürich	2018	41.7	16.9

Source: https://ec.europa.eu/eurostat/web/cities/data/database

As it can be seen, the shares in the QLEC are in all cases much lower than the shares that can be derived from the Eurostat data. This is a clear indication that we are dealing with the under-representation of foreign-born citizens in this case.

Measurement error

Measurement error is in fact a phenomenon which is quite difficult to be estimated with the 'posterior analysis', performed only on the basis of the final micro-data. Therefore, what we provide here is only the analysis of the outliers, which can indicate possible measurement error in some cases. Since the key variables in this survey are categorical variables, we cannot consider outliers on the micro-level. For this reason, we calculated proportions of each of the variable categories on the city level and searched outliers among these 83 proportions. To detect outliers, we used two approaches.

The first approach is to use the well-known 'Tukey method', where lower and upper limit of the acceptable region are calculated as Q1-k*(Q3–Q1), where Q1 and Q3 are first and third quartile in the data distribution and k is appropriately selected parameter that defines the width of the acceptable interval. The most commonly used parameter values are 1.5 for 'ordinary' and 3 for 'extreme' outliers. We took value 5 in order to find the really outstanding outliers.

Table 10: Outliers by distance to the closest value

City	Variable	Form	Percent	Median	Q1	Q2
Tirana	q1_01	Very unsatisfied	46.4	5.8	2.9	9.6
Tirana	q6_02	Strongly disagree	27.4	2.5	1.2	5.3
Napoli	q6_04	Strongly disagree	49.0	5.1	2.2	9.6
Athena	q2_01	Strongly disagree	17.8	1.7	0.9	3.6
Skopje	q1_08	Very unsatisfied	69.7	9.3	5.1	15.7
Roma	q6_02	Strongly disagree	26.2	2.5	1.2	5.3
Valletta	q1_04	Very unsatisfied	35.4	4.0	2.4	7.8
Istanbul	q2_01	Strongly disagree	17.3	1.7	0.9	3.6

By the second approach, we took the highest/lowest value in each variable/category and then calculated the difference to the second highest/lowest value. The following table shows 10 cases with the highest difference.

Table 11: Coherence of the results between two surveys
(in %)

City	Variable	Category	Percent	Next	Difference
Roma	q14	Decreased	71.9	53.9	18.0
Miskolc	q15	Fair (neither good or bad)	45.0	31.6	13.4
Skopje	q3_04	Not a good place to live	65.7	52.4	13.3
Tirana	q1_03	Very unsatisfied	43.0	31.1	12.0
Oviedo	q1_10	Very satisfied	58.6	46.9	11.7
Athena	q09	No	71.8	82.1	10.3
Zurich	q1_01	Very satisfied	79.6	69.4	10.2
Istanbul	q3_01	Not a good place to live	44.3	34.2	10.1
Braga	q2_02	Somewhat disagree	62.0	52.1	10.0
Skopje	q1_08	Very unsatisfied	69.7	60.2	9.5

Comparability and coherence

We made a comparison of the results coming out from the question '*Q05m1: On a typical day, which mode(s) of transport do you use most often*?' with publicly available similar statistical results.

We made first a comparison with the mobility survey that was carried out in autumn 2017. We made a comparison for the city Ljubljana. Results of the Mobility survey for Ljubljana are available at https://pxweb.stat.si/SiStatData/pxweb/sl/Data/-/2281275S.px

We have to stress out that there is a number of methodological differences on how the results in both surveys have been obtained. In the QLEC survey, we have only one question and the respective shares have been calculated based on this question. In the mobility survey, the trips and associated trips' information (e.g. mode, length) for the selected persons were recorded for several days and the shares were calculated based on this data. In addition, the reference year differs for two years.

Table 12: Coherence of the results between two surveys
(in %)

Transportation mode	Quality of life	Mobility survey
1 Car	49.3	61.9
3 Bicycle	13.0	7.2
4 Foot	13.4	19.4
5 Train	1.7	1.4
6 Urban public transport (bus, tram or metro)	15.0	8.3

The second comparison was performed with the publicly available Eurostat data. At the Eurostat site https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/mobility there are few data on the cities where inhabitants most frequently go to work/training by certain transportation mean. We compared these figures with the shares of means (questions Q05m1, Q05m2) which employed persons and persons in education use on a typical day in QLEC survey.

Table 13: Highest shares of used transportation means

	Eurostat	QLEC
People going to work with public transport		
Paris	80.0	50.9
Vienna	73.0	49.8
Prague	67.0	54.8
People going to work by car		
Nicosia	91.0	56.1
Luxembourg	67.0	48.8
Riga	68.0	48.0
Valetta	68.0	61.0
People going to work by bike		
Copenhagen	58.0	39.3
Amsterdam	53.0	43.6
Ljubljana	26.0	20.1

As we can see, some values are quite close together, while other differ significantly. The reason of this is difficult to be established, since we do not have enough information about the methodology

that is behind the Eurostat statistics. What it can be said is that the differences are the most outstanding in the case of large proportions. It is quite difficult to establish whether we have upward bias in the case of Eurostat data, and downward bias in the case of QLEC survey, or the differences are merely generated by methodological differences.

Annex II: List of (perception) variables, subject of the micro-analyses

Question	Description
q07	In the city where you live, do you have confidence in the local police force?
q08	Within the last 12 months, was any money or property stolen from you or another household member in your city?
q09	Within the last 12 months, have you been assaulted or mugged in your city?
q10	Within the last 12 months, would you say you had difficulties to pay your bills at the end of the month ...
q11	Do you feel that if you needed material help (e.g. money, loan or an object) you could receive it from relatives, friends, neighbours or other persons you know?
q12	Do you feel that if you needed non-material help (e.g. somebody to talk to, help with doing something or collecting something) you could receive it from relatives, friends, neighbours or other persons you know?
q14	Compared to five years ago, would you say the quality of life in your city or area has:
q15	In general, how is your health?
q05m1	On a typical day, which mode(s) of transport do you use most often?... | 1st Mention
q05m2	On a typical day, which mode(s) of transport do you use most often?... | 2nd Mention
q13_1	I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, ... — I am satisfied with the amount of time it takes to get a request solved by my local public administration
q13_2	I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, ... — The procedures used by my local public administration are straightforward and easy to understand
q13_3	I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, ... — The fees charged by my local public administration are reasonable
q13_4	I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, ... — Information and services of my local public administration can be easily accessed online
q13_5	I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, ... — There is corruption in my local public administration
q15a	The next question is about your health status. Please remember that all your responses will be treated confidentially. You do not have to answer

Question	Description
	this question if you do not want to. Are you happy to proceed?
q1_01	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Public transport, for example the bus, tram or metro.
q1_02	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Health care services, doctors and hospitals.
q1_03	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Sport facilities such as sport fields and indoor sports halls.
q1_04	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Cultural facilities such as concert halls, theatres, museums and libraries.
q1_05	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Green spaces such as parks and gardens.
q1_06	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Public spaces such as markets, squares, pedestrian areas.
q1_07	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Schools and other educational facilities.
q1_08	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — The quality of the air.
q1_09	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — The noise level.
q1_10	Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or not at all satisfied with each of the following issues in your city or area. — Cleanliness.
q2_01	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — I'm satisfied to live in my city.
q2_02	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — It is easy to find a good job in my city.
q2_03	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — I feel safe walking alone at night in my city.
q2_04	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — I feel safe walking alone at night in my

Question	Description
	neighbourhood.
q2_05	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — It is easy to find good housing in my city at a reasonable price.
q2_06	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Generally speaking, most people in my city can be trusted.
q2_07	I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Generally speaking, most people in my neighbourhood can be trusted.
q3_01	Is the city where you live a good place or not a good place to live for the following groups? — People in general.
q3_02	Is the city where you live a good place or not a good place to live for the following groups? — Racial and ethnic minorities.
q3_03	Is the city where you live a good place or not a good place to live for the following groups? — Gay or lesbian people.
q3_04	Is the city where you live a good place or not a good place to live for the following groups? — Immigrants from other countries.
q3_05	Is the city where you live a good place or not a good place to live for the following groups? — Young families with children.
q3_06	Is the city where you live a good place or not a good place to live for the following groups? — Elderly people.
q4_01	On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with: — The neighbourhood where you live
q4_02	On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with: — Your personal job situation.
q4_03	On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with: — The financial situation of your household.
q4_04	On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with: — The life you lead.
q6_01	Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Affordable
q6_02	Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Safe
q6_03	Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. —

Question	Description
	Easy to get
q6_04	Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Frequent (comes often)
q6_05	Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements. — Reliable (comes when it says it will)

Annex III: Proposal for CATI/CAWI questionnaires (examples)

CATI Questionnaire

Screener

D1. [PROG: Quantity, 3 digits, range min. 0 – max. 115 + 999]
What is your age?
999. Don't know/No Answer/Refuses (DO NOT READ OUT)
D1_recode. [PROG: HIDDEN VARIABLE; recode the response from D1 into the corresponding age category]
1. 15–19
2. 20–24
3. 25–34
4. 35–44
5. 45–54
6. 55–64
7. 65–74
8. 75+
999. Don't know/No Answer/Refuses
[PROG: IF D1 < 15 : Screen out]
[PROG: IF D1 = 999 : Screen out]

D2. [PROG: SINGLE RESPONSE]
What is your sex? (DO NOT READ OUT, to be observed by interviewer)
1. Male
2. Female

D3. SEQUENCE

Main Questionnaire

Q1. [PROG: SINGLE RESPONSE GRID]
Generally speaking, please tell me if you are very satisfied, rather satisfied, rather unsatisfied or very unsatisfied with each of the following issues in your city or area.
Rows [PROG: Randomise items 1–10]
1. Public transport, for example the bus, tram or metro.
2. Health care services, doctors and hospitals.
3. Sport facilities such as sport fields and indoor sports halls.
4. Cultural facilities such as concert halls, theatres, museums and libraries.
5. Green spaces such as parks and gardens.
6. Public spaces such as markets, squares, pedestrian areas.
7. Schools and other educational facilities.
8. The quality of the air.
9. The noise level.
10. Cleanliness.

Columns
4. Very satisfied
3. Rather satisfied
2. Rather unsatisfied

1. Very unsatisfied
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q5. [PROG: MULTIPLE RESPONSE; max. 2 responses allowed]
On a typical day, which mode(s) of transport do you use most often?…
Interviewer instruction: allow 2 responses if offered spontaneously by the respondent, but do not probe if only 1 is given.
1. Car
2. Motorcycle
3. Bicycle
4. Foot
5. Train
6. Urban public transport (bus, tram or metro)
7. Other
98. Do not commute **[PROG: Single Response]**
99. Don't know/No Answer/Refuses (DO NOT READ OUT) **[PROG: Single Response]**

Q6. [PROG: SINGLE RESPONSE GRID]
Thinking about public transport in your city, based on your experience or perceptions, please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements.
Public transport in your city is:
Rows [PROG: Randomise items 1–5]
1. Affordable
2. Safe
3. Easy to get
4. Frequent (comes often)
5. Reliable (comes when it says it will)
Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q2. [PROG: SINGLE RESPONSE GRID]
I will read you a few statements. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements.
Rows [PROG: Randomise items 1–7; Treat 3–4 and 6–7 as fixed pairs: Make sure that item 4 always comes right after 3, and item 7 right after 6]
1. I'm satisfied to live in my city.
2. It is easy to find a good job in my city.
3. I feel safe walking alone at night in my city.
4. I feel safe walking alone at night in my neighbourhood.
5. It is easy to find good housing in my city at a reasonable price.
6. Generally speaking, most people in my city can be trusted.
7. Generally speaking, most people in my neighbourhood can be trusted.
Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q4. [PROG: SINGLE RESPONSE GRID]

On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with:
Rows [PROG: Randomise items 1–4]
1. The neighbourhood where you live
2. Your personal job situation.
3. The financial situation of your household.
4. The life you lead.
Columns
4. Very satisfied
3. Fairly satisfied
2. Not very satisfied
1. Not at all satisfied
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q7. [PROG: SINGLE RESPONSE] (REGIO)
In the city where you live, do you have confidence in the local police force? (7)
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q8. [PROG: SINGLE RESPONSE]
Within the last 12 months, was any money or property stolen from you or another household member, in your city?
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q9. [PROG: SINGLE RESPONSE]
Within the last 12 months, have you been assaulted or mugged in your city?
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q10. [PROG: SINGLE RESPONSE]
Within the last 12 months, would you say you had difficulties to pay your bills at the end of the month …
1. Most of the time
2. From time to time
3. Almost never/never
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q11. [PROG: SINGLE RESPONSE]
Do you feel that if you needed material help (e.g. money, loan or an object) you could receive it from relatives, friends, neighbours or other persons you know?
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q12. [PROG: SINGLE RESPONSE]
Do you feel that if you needed non-material help (e.g. somebody to talk to, help with doing something or collecting something) you could receive it from relatives, friends, neighbours or other persons you know?
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q13. [PROG: SINGLE RESPONSE GRID]
I will read you a few statements about the local public administration in your city. Please tell me whether you strongly agree, somewhat agree, somewhat disagree or strongly disagree with each of these statements.

Rows [PROG: Randomise items 1–5]
1. I am satisfied with the amount of time it takes to get a request solved by my local public administration.
2. The procedures used by my local public administration are straightforward and easy to understand
3. The fees charged by my local public administration are reasonable
4. Information and services of my local public administration can be easily accessed online
5. There is corruption in my local public administration

Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q3. [PROG: SINGLE RESPONSE GRID]
Is the city where you live a good place or not a good place to live for the following groups?
Rows [PROG: Randomise Rows; Keep item 1 always first, randomise items 2–6] (5)
1. People in general. **[PROG: Fixed]**
2. Racial and ethnic minorities.
3. Gay or lesbian people.
4. Immigrants from other countries.
5. Young families with children.
6. Elderly people.

Columns
1. A good place to live
2. Not a good place to live
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Q14. [PROG: SINGLE RESPONSE]
Compared to five years ago, would you say the quality of life in your city or area has:
1. Decreased
2. Stayed the same
3. Increased
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

Socio Demographic questions

D5. [PROG: SINGLE RESPONSE; insert answer list 'D5 – Countries' as drop down]
In which country were you born?

D6. [PROG: SINGLE RESPONSE]
Have you ever lived in another city for at least 1 year?
1. Yes
2. No
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D7. [PROG: Quantity; only if D6 = 1; min. 0; max. 115]
How many years have you been living in your current city since last moving here?
Interviewer instruction: If respondent answers 'less than 1 year', code as 0
999. Don't know/No Answer/Refuses (DO NOT READ OUT)

D9. [PROG: Quantity; min. 1; max. 15]
How many people usually live in your household? Please include yourself.

D9b. [PROG: Quantity; only if D9 > 1; min.1.; max. = answer given in D9]
How many of these are aged 15 and older? Please include yourself.
[PROG: autocode D9b = 1 if D9 = 1]

D8. [PROG: SINGLE RESPONSE. ONLY IF D9 > 1]

Which of the following best describes your household composition? With household, we mean all people that typically live with you in the same residence. Please include anyone who is temporarily away for work, study or vacation
[PROG: autocode D8 = 1 if D9 = 1]
1. One-person household **[PROG: do not show. If D9 = 1, autocode D8 = 1]**
2. Lone parent with at least one child aged less than 25
3. Lone parent with all children aged 25 or more
4. Couple without any child(ren)
5. Couple with at least one child aged less than 25
6. Couple with all children aged 25 or more
7. Other type of household
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D10local. [PROG: SINGLE RESPONSE; insert answer list 'D10 – education'; use the value and show 'Educ categories ENGLISH' in the master questionnaire and the 'Educ categories LOCAL' for the local translations]
What is the highest level of education you have successfully completed?
Interviewer instruction: DO NOT READ OUT response options unless needed to proceed
99. Don't know/No Answer/Refuses (DO NOT READ OUT)
D10ISCED. [PROG: HIDDEN VARIABLE; recode the response from D10local into the corresponding ISCED level as indicated in the column 'ISCED code']
1. Less than Primary education (ISCED 0)
2. Primary education (ISCED 1)
3. Lower secondary education (ISCED 2)
4. Upper secondary education (ISCED 3)
5. Post-secondary non-tertiary education (ISCED 4)
6. Short-cycle tertiary education (ISCED 5)
7. Bachelor or equivalent (ISCED 6)
8. Master or equivalent (ISCED 7)
9. Doctoral or equivalent (ISCED 8)
10. Don't know/No Answer/Refuses

D11a. [PROG: SINGLE RESPONSE]
Do you currently have a job?
1. Yes
2. No (GO TO D14)
99. Don't know/No Answer/Refuses (DO NOT READ OUT; GO TO D14)

D11. [PROG: SINGLE RESPONSE]
Which of the following best describes your current working status?
1. At work as employee or employer/self-employed/relative assisting on family business
2. Unemployed, not looking actively for a job
3. Unemployed, looking actively for a job
4. Retired
5. Unable to work due to long-standing health problems
6. In full-time education (at school, university, etc.) / student
7. Full-time homemaker/responsible for ordinary shopping and looking after home
8. Compulsory military or civilian service
9. Other
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D12. [PROG: SINGLE RESPONSE; only ask if D11 =1]
What is your current job position/occupation?
Interviewer instruction: DO NOT READ OUT response options. If respondent is unsure, ask to state their exact job/function. If a respondent is in the military, always code as 'armed forces occupation', regardless of their job within the military.
D12_recode. [PROG: HIDDEN VARIABLE; recode the response from D12 into the corresponding ISCO category]
1. Manager
2. Professional

3. Technician and associate professional
4. Clerical support worker
5. Services and sales worker
6. Agricultural, forestry or fishery worker
7. Craft or related trade worker
8. Plant or machine operator or assembler
9. Elementary occupation
10. Armed forces occupation **[PROG: autocode D12 = 10 if D11 = 8]**
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D13. [PROG: SINGLE RESPONSE; ask if D11 = 1 or D11 = 8]
Which of the following best describes your job?
1. Full-time job
2. Part-time job
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D14. [PROG: SINGLE RESPONSE; ask if SampleType = 2 (Fixed)]
Do you personally own a mobile phone?
1. Yes
2. No
[PROG: autocode D14 = 1 if SampleType = 1 (mobile sample)]

D15. [PROG: SINGLE RESPONSE; ask if SampleType = 1 (Mobile)]
Do you have a landline phone in the household?
1. Yes
2. No
[PROG: autocode D15 = 1 if SampleType = 2 (fixed sample)]
Mobfix. [PROG: HIDDEN VARIABLE; recode the response from D14 and D15 into the corresponding category]
1. Fixed only: If (SampleType = 2 and D14 = 2)
2. Mobile only: if (SampleType = 1 and D15 = 2)
3. Mixed: if (SampleType = 2 and D14 = 1) OR or (SampleType = 1 and D15 = 1)

Q15a [PROG: SINGLE RESPONSE]
The next question is about your health status. Please remember that all your responses will be treated confidentially. You do not have to answer this question if you do not want to.
Are you happy to proceed?
1. Yes
2. No

Q15. [PROG: SINGLE RESPONSE, ask if Q15a=1]
In general, how is your health?
[PROG: autocode Q15=99 if Q15a = 2]
5. Very good
4. Good
3. Fair (neither good or bad)
2. Bad
1. Very bad
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

CAWI Questionnaire

Screener

D1. [PROG: Quantity, 3 digits, range min. 0 – max. 115 + 999]
What is your age?
D1_recode. [PROG: HIDDEN VARIABLE; recode the response from D1 into the corresponding age category]
1. 15–19
2. 20–24
3. 25–34
4. 35–44
5. 45–54
6. 55–64
7. 65–74
8. 75+
999. Don't know/No Answer/Refuses
[PROG: IF D1 < 15 : Screen out]
[PROG: IF D1 = 999 : Screen out]

D2. [PROG: SINGLE RESPONSE]
What is your sex?
1. Male
2. Female

D3.aa [PROG: SINGLE RESPONSE]
Which country do you live in?
D3.aa_recode. [PROG: HIDDEN VARIABLE; recode the response from D3.aa into the corresponding country name]

01. Belgium
02. Bulgaria
03. Czech Republic
04. Denmark
05. Germany
06. Estonia
07. Ireland
08. Greece
09. Spain
10. France
11. Italy
12. Cyprus
13. Latvia
14. Lithuania
15. Luxembourg
16. Hungary
17. Mata
18. The Netherlands
19. Austria
20. Poland
21. Portugal
22. Romania
23. Slovenia
24. Slovak Republic
25. Finland
26. Sweden
27. Ukraine
28. Croatia

29. Albania
30. Turkey
31. Macedonia
32. Serbia
33. Montenegro
34. Norway
35. Iceland
999. Don't know/No Answer/Refuses
[PROG: IF D3.aa = 999 : Screen out]

D3. SEQUENCE

Main Questionnaire

Q1. [PROG: SINGLE RESPONSE GRID]
In general, please mark how satisfied you are with each of the following issues in your city or area.
Rows [PROG: Randomise items 1–10]
1. Public transport, for example the bus, tram or metro.
2. Health care services, doctors and hospitals.
3. Sport facilities such as sport fields and indoor sports halls.
4. Cultural facilities such as concert halls, theatres, museums and libraries.
5. Green spaces such as parks and gardens.
6. Public spaces such as markets, squares, pedestrian areas.
7. Schools and other educational facilities.
8. The quality of the air.
9. The noise level.
10. Cleanliness.

Columns
4. Very satisfied
3. Rather satisfied
2. Rather unsatisfied
1. Very unsatisfied
5. Don't know
99. No Answer

Q5. [PROG: MULTIPLE RESPONSE; max. 2 responses allowed]
On a typical day, which mode(s) of transport do you use most often?…
1. Car
2. Motorcycle
3. Bicycle
4. Foot
5. Train
6. Urban public transport (bus, tram or metro)
7. Other
8. Don't know
98. Do not commute **[PROG: Single Response]**
99. No Answer **[PROG: Single Response]**

Q6. [PROG: SINGLE RESPONSE GRID]
Thinking about public transport in your city, based on your experience or perceptions, please mark to what extent you agree with each of these statements.
Public transport in your city is:
Rows [PROG: Randomise items 1–5]
1. Affordable
2. Safe
3. Easy to get
4. Frequent (comes often)
5. Reliable (comes when it says it will)

Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
5. Don't know
99. No Answer

Q2. [PROG: SINGLE RESPONSE GRID]
Please mark to what extent you agree with each of these statements.
Rows [PROG: Randomise items 1–7; Treat 3–4 and 6–7 as fixed pairs: Make sure that item 4 always comes right after 3, and item 7 right after 6]
1. I'm satisfied to live in my city.
2. It is easy to find a good job in my city.
3. I feel safe walking alone at night in my city.
4. I feel safe walking alone at night in my neighbourhood.
5. It is easy to find good housing in my city at a reasonable price.
6. Generally speaking, most people in my city can be trusted.
7. Generally speaking, most people in my neighbourhood can be trusted.

Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
5. Don't know
99. No Answer

Q4. [PROG: SINGLE RESPONSE GRID]
Overall, how satisfied are you with:
Rows [PROG: Randomise items 1–4]
1. The neighbourhood where you live
2. Your personal job situation.
3. The financial situation of your household.
4. The life you lead.

Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
5. Don't know
99. No Answer

Q7. [PROG: SINGLE RESPONSE] (REGIO)
In the city where you live, do you have confidence in the local police force?
1. Yes
2. No
3. Don't know
99. No Answer

Q8. [PROG: SINGLE RESPONSE]
Within the last 12 months, was any money or property stolen from you or another household member, in your city?
1. Yes
2. No
3. Don't know
99. No Answer

Q9. [PROG: SINGLE RESPONSE]
Within the last 12 months, have you been assaulted or mugged in your city?
1. Yes
2. No
3. Don't know
99. No Answer

Q10. [PROG: SINGLE RESPONSE]
Within the last 12 months, would you say you had difficulties to pay your bills at the end of the month …
1. Most of the time
2. From time to time
3. Almost never/never
4. Don't know
99. No Answer

Q11. [PROG: SINGLE RESPONSE]
Do you feel that if you needed material help (e.g. money, loan or an object) you could receive it from relatives, friends, neighbours or other persons you know?
1. Yes
2. No
3. Don't know
99. No Answer

Q12. [PROG: SINGLE RESPONSE]
Do you feel that if you needed non-material help (e.g. somebody to talk to, help with doing something or collecting something) you could receive it from relatives, friends, neighbours or other persons you know?
1. Yes
2. No
3. Don't know
99. No Answer

Q13. [PROG: SINGLE RESPONSE GRID]
Please mark to what extent you agree with each of these statements about the local public administration in your city:
Rows [PROG: Randomise items 1–5]
1. I am satisfied with the amount of time it takes to get a request solved by my local public administration.
2. The procedures used by my local public administration are straightforward and easy to understand
3. The fees charged by my local public administration are reasonable
4. Information and services of my local public administration can be easily accessed online
5. There is corruption in my local public administration

Columns
4. Strongly agree
3. Somewhat agree
2. Somewhat disagree
1. Strongly disagree
5. Don't know
99. No Answer

Q3. [PROG: SINGLE RESPONSE GRID]
Is the city where you live a good place or not a good place to live for the following groups?
Rows [PROG: Randomise Rows; Keep item 1 always first, randomise items 2–6] (5)
1. People in general. **[PROG: Fixed]**
2. Racial and ethnic minorities.
3. Gay or lesbian people.
4. Immigrants from other countries.
5. Young families with children.
6. Elderly people.

Columns
1. A good place to live
2. Not a good place to live
3. Don't know
99. No Answer

Q14. [PROG: SINGLE RESPONSE]
Compared to five years ago, would you say the quality of life in your city or area has:
1. Decreased
2. Stayed the same
3. Increased
4. Don't know
99. No Answer

Socio Demographic questions

D5. [PROG: SINGLE RESPONSE; insert answer list 'D5 – Countries' as drop down]
In which country were you born?

D6. [PROG: SINGLE RESPONSE]
Have you ever lived in another city for at least 1 year?
1. Yes
2. No
99. No Answer

D7. [PROG: Quantity; only if D6 = 1; min. 0; max. 115]
How many years have you been living in your current city since last moving here?
Interviewer instruction: If respondent answers 'less than 1 year', code as 0
999. Don't know/No Answer

D9. [PROG: Quantity; min. 1; max. 15]
How many people usually live in your household? Please include yourself.

D9b. [PROG: Quantity; only if D9 > 1; min.1.; max. = answer given in D9]
How many of these are aged 15 and older? Please include yourself.
[PROG: autocode D9b = 1 if D9 = 1]

D8. [PROG: SINGLE RESPONSE. ONLY IF D9 > 1]
Which of the following best describes your household composition? With household, we mean all people that typically live with you in the same residence. Please include anyone who is temporarily away for work, study or vacation
[PROG: autocode D8 = 1 if D9 = 1]
1. One-person household **[PROG: do not show. If D9 = 1, autocode D8 = 1]**
2. Lone parent with at least one child aged less than 25
3. Lone parent with all children aged 25 or more
4. Couple without any child(ren)
5. Couple with at least one child aged less than 25
6. Couple with all children aged 25 or more
7. Other type of household
99. Don't know/No Answer/Refuses (DO NOT READ OUT)

D10. [PROG: SINGLE RESPONSE]
What is the highest level of education you have successfully completed?
1. Less than Primary education (ISCED 0)
2. Primary education (ISCED 1)
3. Lower secondary education (ISCED 2)
4. Upper secondary education (ISCED 3)
5. Post-secondary non-tertiary education (ISCED 4)
6. Short-cycle tertiary education (ISCED 5)
7. Bachelor or equivalent (ISCED 6)
8. Master or equivalent (ISCED 7)
9. Doctoral or equivalent (ISCED 8)
10. Don't know
99. No Answer

D11a. [PROG: SINGLE RESPONSE]
Do you currently have a job?
1. Yes
2. No (GO TO Q15a)
99. Don't know/No Answer/Refuses (DO NOT READ OUT; GO TO Q15a)

D11. [PROG: SINGLE RESPONSE]
Which of the following best describes your current working status?
1. At work as employee or employer/self-employed/relative assisting on family business
2. Unemployed, not looking actively for a job
3. Unemployed, looking actively for a job
4. Retired
5. Unable to work due to long-standing health problems
6. In full-time education (at school, university, etc.) / student
7. Full-time homemaker/responsible for ordinary shopping and looking after home
8. Compulsory military or civilian service
9. Other
10. Don't know
99. No Answer

D12. [PROG: SINGLE RESPONSE; only ask if D11 =1]
What is your current job position/occupation?
D12_recode. [PROG: HIDDEN VARIABLE; recode the response from D12 into the corresponding ISCO category]
1. Manager
2. Professional
3. Technician and associate professional
4. Clerical support worker
5. Services and sales worker
6. Agricultural, forestry or fishery worker
7. Craft or related trade worker
8. Plant or machine operator or assembler
9. Elementary occupation
10. Armed forces occupation **[PROG: autocode D12 = 10 if D11 = 8]**
11. Don't know
99. No Answer

D13. [PROG: SINGLE RESPONSE; ask if D11 = 1 or D11 = 8]
Which of the following best describes your job?
1. Full-time job
2. Part-time job
3. Don't know
99. No Answer

Q15a [PROG: SINGLE RESPONSE]
The next question is about your health status. Please remember that all your responses will be treated confidentially. You do not have to answer this question if you do not want to. Are you happy to proceed?
1. Yes
2. No

Q15. [PROG: SINGLE RESPONSE, ask if Q15a=1]
In general, how is your health?
[PROG: autocode Q15=99 if Q15a = 2]
5. Very good
4. Good
3. Fair (neither good or bad)
2. Bad
1. Very bad
6. Don't know
99. No Answer

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications at: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

KS-TC-21-008-EN-N

IMPROVING DATA COLLECTION FOR THE QUALITY OF LIFE IN EUROPEAN CITIES SURVEY

The Quality of Life in European Cities Survey (QLEC Survey) is a statistical survey that the European Commission has been carrying out since 2004. It aims at measuring the quality of life in 83 European (greater) cities through a perception survey. This study introduces the results of the revision of the 2019 implementation of the QLEC survey, in the form of a scientific paper. The findings of the revision are accompanied with recommendations for the improvements to make in the next survey wave.

For more information
https://ec.europa.eu/eurostat/

Publications Office
of the European Union

ISBN 978-92-76-46110-4