TASK 3 REPORT: DATA COLLECTION METHODOLOGIES

EUFireStat - Closing data gaps and paving the way for pan-European Fire Safety Efforts

Contractor European Commission
Directorate General for Internal Market, Industry, Entrepreneurship and SMEs

Reference of the contract SI2.830108

Prepared by:
National Fire Protection Association
Danish Institute of Fire and Security Technology (DBI) - Leader of Task 4
Bundesanstalt für Materialforschung und –prüfung (BAM)
Centre for Fire Statistics of CTIF (CFS-CTIF)
Efectis - Consortium leader in the project
Lund University (LU)
National Fire Protection Association (NFPA)
School of Engineering, The University of Edinburgh (UoE)
The European Fire Safety Alliance (EuroFSA)
Vereinigung zur Förderung des Deutschen Brandschutzes (VFDB)

Date: 01-12-2021
Revision index:
Number of pages: 89

TRACK OF MODIFICATION

Indices of revision	Date	Modifications
A	14/09/2021	Initial version
B	01/12/2021	Revised version with comments from the European Commission

Table of Contents

Executive Summary

Data about the causes and circumstances of fires play an important role in informing fire safety efforts and guiding policy decisions. When data are harmonized among countries, they can promote the sharing of fire prevention strategies and provide access to novel programs and interventions. It is essential that data from fire incidents reliably capture the experiences of underlying populations if they are to serve these public purposes.

This report for the EUFireStat project, "Closing data gaps and paving the way for pan-European Fire Safety Efforts," provides background on data collection methodologies as they pertain to the collection of data on building fires among Member States of the European Union.

This report includes several components:

- A brief overview of data sources in the reporting of fire incidents in order to highlight unique features that merit attention in considering methodology and design of fire incident data collection.
- An overview of data collection strategies, focusing on the collection of data through census surveys, sample surveys, and convenience sampling methodologies. The report discusses advantages and disadvantages of these methods in relation to collecting information on fire incidents and identifies cost components of the respective methodologies. This section of the report also provides background on the CARES road accident database in the European Union and reviews points of comparison between data collection of road accidents and fire data collection systems.
- A discussion of missing data in survey research and its potential to undermine the reliability of data, leading to invalid conclusions about phenomena related to a target population. This section reviews common types of missing data, including a complete failure to respond to a survey and failure to answer individual items in a survey. The section identifies issues arising from missing data that impact the validity of results and common methods of handling missing data. The discussion also reviews the potential impact of missing data in fire incident data collection and approaches to missing data in select national fire data collection systems.
- A section on uncertainty analysis and overview of statistical reliability, including factors which influence the relationship between true values and estimated values of a statistical measure. The analysis illustrates how greater accuracy and precision in measurement increase the correspondence between observed values and the true values. The section also includes a qualitative assessment of twelve variables proposed for inclusion in a harmonized fire data collection system, as previously presented in Task 2. The assessment ranks the variables according to "high," "medium," and "low" uncertainty and suggests potential ways to reduce the uncertainty around high and medium uncertainty variables. This section also identifies areas for future study that may better quantify uncertainties which arise in fire statistics.
- A detailed review of the prospective financial costs of fire data collection at the national level. The section outlines data collection for a harmonized data collection system as comprised of a series of steps which include data collection at the local level, local data management, reporting of local data to the national level, data management at the national level, and reporting of national data to the EU level. The section examines cost components of local data collection at the local level, then compiles the costs for data collection at the national level for census and sample surveys respectively. Using these cost components, the analysis provides cost estimates for census and sample surveys for each of the EU member states. No cost estimates were calculated for convenience sampling due to the subjectivity and variability inherent in this form of data collection.

List of Figures

List of Tables

List of Abbreviations

CADaS	Community Accident Data Set
CARE	Community database on Accidents on the Roads in Europe
CPSZ	Consumer Product Safety Commission
EC	European Commission
EU	European Union
FES	Fire Experience Survey
GPS	Global Positioning System
ICD-10	International Classification of Diseases, 10th Revision
MSB	The Swedish Civil Contingencies Agency
NFIRS	National Fire Incident Reporting System
NFPA	National Fire Protection Association
UKSA	United Kingdom Statistics Authority
USFA	United States Fire Administration

1. Introduction

Data on fire incidents is essential for identifying trends on how and why fires occur and planning intervention efforts. There are nevertheless some substantial challenges in collecting data on fire incidents that require careful planning in the design and implementation of data collection systems. These challenges include the reliance upon firefighters for principal data collection, time lags in the availability of certain key data elements (such as causal information), and the scattering of some information, such as details on casualties or monetary losses, among different informants. We attempt to address some of these challenges in this review of methods for fire data collection and to identify potential issues for introducing a uniform system of fire data collection across member states of the European Union.

The body of the report is comprised of six sections related to data collection methods and analysis in fire incident data collection. The report also included three annexes which provide detailed information on the data collection options discussed in the report and cost estimates for fire incident data collection for each of the Member States of the European Union.

Section 2 opens the report with brief overview of data collection and its function as a surveillance system. We include here a review of the common mechanics involved in the data collection of fire incidents and some of the complicating aspects that uniquely distinguish fire incident data collection.

In Section 3, we review three major methods for collecting fire data: census, survey, and convenience sampling. These methods represent general data collection strategies and are not specific to the collection of fire incident data, although they could serve as potential models for the collection of fire incident data. We present distinguishing features of each data collection system and highlight strengths and weaknesses of each approach. In addition to these formal models of data collection, we present in this opening section information on a data collection system on road accidents in Europe. This database, the Community database on Accidents on the Roads in Europe (CARE), is potentially instructive for a data collection system of fire incidents in Europe since the CARE database is comprised of data collected by Member States of the European Union and implementation has involved challenges around data harmonization and inclusion of common variables.

Section 4 of the report directs attention to missing data and its impact upon data quality in data collection. This section also includes a review of uncertainty of data and its implications for the quality and reliability of data. We introduce this section with a general overview of missing data in survey research – including a discussion of what missing data is, how it affects data quality, and some of the techniques for handling it to optimize the utility of results.

In Section 5, we turn our focus on missing data to missing data in fire incident reporting. This discussion identifies some of the factors that influence the reporting of data on fire incidents, as well as efforts to deal with missing data in existing fire incident data collection systems. The attention here primarily focuses upon an approach to missing data in the U.S. National Fire Incident Reporting System, where the problem seems to have received the greatest attention. We also review existing approaches to the treatment of missing data in member countries of the European Union, where information is available. The section concludes with an analysis of uncertainty in data collection, including ways in which uncertainty arises and the corresponding impact upon the results.

Section 6 of the report examines uncertainty in data collection and factors influence statistical reliability. overview of statistical reliability, including factors which influence the relationship between true values and estimated values of a statistical measure. The analysis illustrates how greater accuracy and precision in measurement increase the correspondence between observed values and the true values. The section also includes qualitative assessment of twelve variables proposed for inclusion in a harmonized fire data collection system, as previously presented in Task 2.

Section 7 of the report provides cost estimates of implementing a fire incident data collection system at the national level. The section begins with an overview of data collection through a series of steps: data collection at the local level, local data management, reporting of local data to national level, national data management, and reporting of national data to the EU level. The section cost components of local data collection at the local level, then reviews costs for data collection at the national level for census and sample surveys respectively. Using these cost components, the analysis provides cost estimates for census and sample surveys for each of the EU member states. No cost estimates were calculated for convenience sampling due to the subjectivity and variability inherent in this form of data collection.

We summarize key findings from the research and offer concluding observations in Section 8.

Information in the report annexes is as follows. Annex 1 provides a comparative assessment of census, survey sample, and conveniences sampling data collection methodologies and highlights advantages and disadvantages in the collection of fire incident data. Annex 2 provides an overview of data used to calculate data collection costs in EU Member States. In Annex 3, we provide a breakdown of cost estimates for fire incident data collection for each of the 27 Member States of the European Union.

2. Data Sources in Fire Incident Data Collection

Data collection systems for fire incidents effectively serve as a type of surveillance system, a tool for collecting data on adverse events and key factors associated with them in order to help plan prevention activities (1). While surveillance systems often receive little public attention, they serve an important role by identifying vulnerable populations, pinpointing hazards, establishing priorities for intervention, and helping to mobilize the efficient allocation of resources.

A distinguishing feature of fire incident data collection is that the primary sources of critical information are local fire departments. This reliance on fire departments for data collection is understandable since they have unique access to critical details, such as when the fire occurred (month, day, hour), type of building (if a building fire), whether the building was occupied, the area of fire origin, presence or absence of detection equipment, and so on. Information in incident records is typically entered by firefighters at some point following a service call.

It should be noted that there is substantial variation in the mechanics of data collection systems among fire departments. Some local fire departments in Canada still use paper records, while others rely on local computerized systems and still others have more sophisticated systems linked to larger networks (2). Local fire departments in the United States are also characterized by notable differences in the availability and types of technology used in data collection (3). Such disparities also exist in fire data collection among member states of the European Union.

The issue of information technology in fire incident data collection is particularly significant in relation to the capacity to establish linkages with external databases that contain data not recorded by or available to the local fire department. While local fire departments are likely to be responsible for generating records and providing essential incident information, some key data elements will have to be provided by other primary data sources. For instance, information of deaths or injuries may need to be supplied or supplemented by medical records, while information on financial losses may be provided by insurers. In addition, the determination of fire cause is frequently made only after the completion of an investigation and will need to be transmitted to the incident record. It should be noted that a recommendation to consult other data sets is also discussed in Task 7 as part of the fire data journey.

It is important to take account of these features of fire incident data collection due to their influence upon the selection of appropriate methods for collecting information and strategies for analyzing results. As observers have pointed out in the context of fire data collection in the U.S. and Canada, local fire departments vary in their levels of funding and resources, influencing the capacity to engage in data collection, and thereby the completeness and accuracy of data (2,3). A related challenge is that the firefighters tasked with filing incident reports perform this duty as a side activity to their important public safety function. This issue will be compounded in smaller departments staffed by volunteer firefighters, for whom data collection is even less likely to be a motivating factor for service (2).

An overview of who collects the data, the entity that processes the data and the entity that reports the data in each of countries studied in task 0 and 1 is presented in Annex Table 1.4. For all countries studied except Austria, data collection originates with the fire department. How data are processed and who they are ultimately reported to nationally varies considerably with only two countries Hungary and Italy using a national statistical agency. This will be further discussed in Task 7 where recommendations for achieving a harmonized EU wide system is provided.

3. Data collection methods

It is apparent from information collected in earlier stages of this research project that there is no single approach to fire incident data collection by countries within and outside the European Union. Differences in data collection methodologies may be influenced by a variety of factors, including distribution of administrative authority, data collection traditions, available technologies, size and distribution of the population, levels of funding, and other factors. The achievement of more harmonized data among fire data collection systems in the European Union will be substantially influenced by the consistency and completeness of information produced by the data collection systems of member states. In this section, we review three general approaches to data collection – census surveys, sample surveys, and convenience samples -- as potential approaches to collecting information on fire incidents, with some commentary on pros and cons of each method. In Annex 1, we augment this information with detailed tables comparing the respective data collection methods, as well as an overview of variables proposed for data collection and a comparison of current fire data collection systems reviewed in the course of this research.

Data collection by census
A census is a method of collecting information for every unit in a population or group within a specific defined territory and within a clearly defined point in time, with the population defined as all units within an investigation area (4,5). A census is usually taken at regular intervals. If a census fails to include the entire population, the results may be referred to as an "incomplete census" (6). Data collection by census is appropriate when population units are heterogeneous, when researchers seek detailed information from all constituent subgroups, and when all units of the population is necessary (4-6,).

A key benefit of data collected by census is that the estimates are not subject to sampling error (5,6). In addition, census data should capture sub-groups of the population and provide information that is unique to those sub-groups and might go unrecognized through alternative means of data collection (6). Information from a properly conducted census should be highly accurate.

The comprehensive nature of the census carries with it certain disadvantages. Administration of a census can be complicated, time-consuming, and more costly enterprise than other data collection methods (5,6). A large staff is required to plan and administer a census and analyze results. Collecting information from all members of the target population and maintaining control over the complex array of procedures represent substantial logistical challenges that are beyond the capacity of some research projects (6).

Although data collection by census is most associated with national population surveys, it is a technique applicable to other data collection purposes. In the United States, for instance, the Bureau of Labor Statistics uses a census approach to compile a comprehensive annual count of fatal occupational injuries and descriptive information of their nature and circumstances (7). For purposes of fire incident data collection, use of a census approach to data collection would involve the preparation and submittal of a fire incident form for every eligible fire incident in the country. This approach might have to be replicated in all member states of the European Union for inter-country comparisons of fire incident data.

A clear advantage of the census approach for fire data collection is that it seeks to collect data from the entire population of fire incidents in a country. Data collection by census should thereby minimize the possibility of failing to include reports of rare incidents or those associated with special populations, such those who live in rural areas or areas of economic disadvantage. However, census surveys can also be expensive, and they require substantial administrative and other resources to conduct outreach and ensure the inclusion of all eligible incidents.

Data collection by sample survey
Sample surveys are a common form of data collection which differ from a census by gathering data from a subset of a population. Rather than attempt to collect information from the entire population, researchers can infer the information sought by creating estimates based on samples of the population.

Sample surveys fall into two broad categories: probability samples and non-probability samples.

Probability sampling uses all eligible units from the population (8). A basic probability sampling technique is simple random sampling. Here, information is collected from units that are randomly chosen from the population entirely by chance, with each unit having an equal chance of selection (9). To select the sample, it is first necessary to create a sampling frame comprised of all units from which the selection will be made. A common technique for drawing the sample is to assign a number to every unit in the frame and to select numbers from random number tables or through the use of statistical software (10). The sample is seen to represent the larger population from which it is drawn and to provide reasonably representative information on the measures of interest.

When data is collected from populations that are more diverse, researchers divide the target population into sub-groups, or strata, and then draw probability samples from each (7,8). Such stratified sampling of research subjects is appropriate when the measures of interest can be reasonably expected to vary between sub-groups (8). Stratified sampling seeks to ensure that information collected from the sample is reliably representative of the entire population. Stratified sampling may select unequal samples from each stratum in order to reflect the relative sizes of the strata (8).

An advantage of stratified sampling is that it reduces sampling bias and improves the accuracy of results. If samples are sufficiently large, information from stratified sampling should have sufficient statistical power to reveal differences between sub-groups (9). Disadvantages of probability sampling include being less precise than a census due to the reliance upon a sample population and potentially failing to provide information about smaller sub-groups (7). Sample surveys may be favored because they are less expensive, less time-consuming, and less operationally complicated than data collection by census. For stratified sampling in particular, however, substantial effort is needed to identify sub-groups and create samples that are sufficiently large to be statistically meaningful (9).

Non-probability sampling: probability and convenience sample surveys

Unlike information collected from probability samples, non-probability sampling methods do not use all eligible units in the sampling frame. Instead, the sample is created on the basis of some criterion of research interest or objectives. Since some population units have no chance of being selected, information collected by non-probability samples cannot be considered to produce generalizable results, and it is not possible to estimate sampling error. Two representative forms of non-probability sampling include purposive sampling and convenience sampling, each of which is briefly reviewed below.

Purposive sampling

Purposive sampling is a technique for collecting information from research subjects that are selected on the basis of the judgment of the researcher (10). Also referred to as selective sampling, purposive sampling allows researchers to identify informants who are viewed as representative of research needs, based upon certain specific characteristics or criteria. Purposive sampling is often used in qualitative research when the goal is to utilize cases that are rich in information within the context of limited research resources (11).

Purposive sampling facilitates the ability to collect information from a range of respondents, while also being relatively time and cost effective. However, information collected the purposively selected sample may not fully represent the underlying population and will be more likely to be subject to bias. This limitation may not be a problem if the results are not intended to be generalizable for the entire population (10), and the information collected from purposively-selected research cases may be useful for providing clues about phenomena of research interest and guiding interventions.

In the case of fire incident data collection, purposive sampling might be used to collect information on select fires of particular interest (such as fires meeting a certain loss threshold), on a cross-section of fires in specific building categories, or on the basis of some other criteria. This approach would potentially facilitate the collection of particularly meaningful information with the most efficient use of scarce resources. There would nevertheless be limited opportunities in this approach to use the data in inter-country comparisons, to analyze trends, or to apply it to other data surveillance applications concerning the nature and distribution of fire incidents.

Convenience sampling
Convenience sample surveys are a common method of non-probability sampling. As the name implies, convenience sampling is a method that collects information from informants who are convenient to the researcher, with selection likely to be based on such factors as proximity to the research or accessibility (8). There are no inclusion criteria involved in selecting respondents and the information collected from the sample is unlikely to be generalizable to the general population.

Convenience sampling may be useful for exploratory research with limited purposes, but the method is subject to a high degree of bias and consequently the data has limited credibility. Findings from the research are limited to the sample itself (8). Differences within the underlying population that impact the phenomenon of interest will go undetected and fail to be addressed by conclusions and recommendations of the research.

The advantages of convenience sampling include the simplicity of respondent recruitment, the ability to collect data in a short period, and the lower cost in comparison to other methods (8,10). However, for research that seeks to collect information representative of an entire population, the convenience sample has notable disadvantages of being vulnerable to selection bias and carrying a high level of sampling error (8,10).

For purposes of this project, a convenience sample would fail to include all fire incident reports, making it less inclusive than a census and less statistically rigorous than a sample survey. Fire incident data collected through convenience sampling methods might rely on incident reports based upon relationships with local fire service officers or location in familiar cities or regions. If the sample was solely drawn from fire departments in wealthy districts, for instance, hazards associated with substandard housing or lower income neighborhoods would go unrecognized in policy recommendations.

Implementation of a Common Data Collection System in the EU: Road accidents in Europe
In considering the prospect of introducing a common data collection system for fire incidents across Member States of the European Union (EU), it is worth noting that the EU has experience with data collection for road accidents that appears in certain respects to follow a census approach based on its determination to include every vehicle incident that meets inclusion criteria.

The Community database on Accidents on the Roads in Europe (CARE) is comprised of detailed data on individual accidents collected by the Member States of the European Union using a structure that allows for maximum flexibility in the analysis of information contained in the system (12). The database includes information recorded from all road accidents involving at least one moving vehicle and one injury or death. The inclusion of information on injury victims in the accident is optional. Data for accidents involving only material damage are not considered (13). At the local level, each European Union country transmits the data from its national collection to the European Commission. The data are then transferred from the European Commission to CARE.

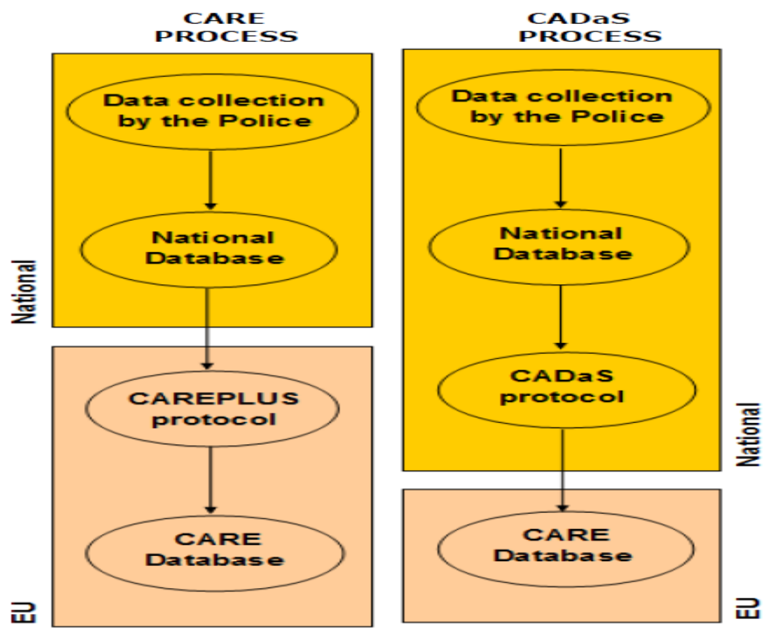
Quality checks are applied by CARE and the respective national datasets are merged into a unique database. The national datasets are partly integrated into the CARE database in their original national structure and definitions. However, existing national accident data collection systems are not always comparable across countries. Data are made compatible through a framework of transformation rules to the national datasets. Definitions are available and provided in a glossary (13).

The quality and availability of road accident data are affected by differences in data collection form structures and the relevant data formats among the existing national databases, recorded variables and available definitions. This lack of accident data uniformity among and within EU countries limits the possible analyses and comparisons at the EU level (13).

It is in this light that recommendations for a Common Accident Data Set (CADaS) have been developed to compare road accident data, with the inclusion of additional variables and values with a common definition to those contained in the previous models of the CARE database. CADaS contains a minimum set of standardised data elements that can be implemented voluntarily. The EU Member States are not obliged to adopt CADaS and they transmit the data at the EU level choosing the level of detail (13).

The methodology can be subdivided into two main processes: the CARE or the CADaS process, as illustrated in Figure 1 (13). Data are collected by the police and merged into the national database. If the country adopts the CADaS protocol, this is managed at the local level and consequently transmitted at the EU level. Otherwise, the national dataset is transmitted at the EU level and uploaded to the CARE database.

Figure 1: Accident data transformation processes.



The transmission of the data from local to EU level is based on the following steps:

0. Each country has a contact person in the European Commission (EC). This person provides support in the process to adopt the CADaS model and transmit the data
1. Four files are created (1 per category) for each country by transforming the national country data structure into the CADaS structure
2. The four files are transmitted using a web portal to the EC
3. EC receives the files and transfers them to the CARE team. The CARE team uploads the files to the CADaS database and performs quality checks. In case of any issue, an email is sent to the specific country for clarifications
4. Once data are successfully uploaded, the country is informed.

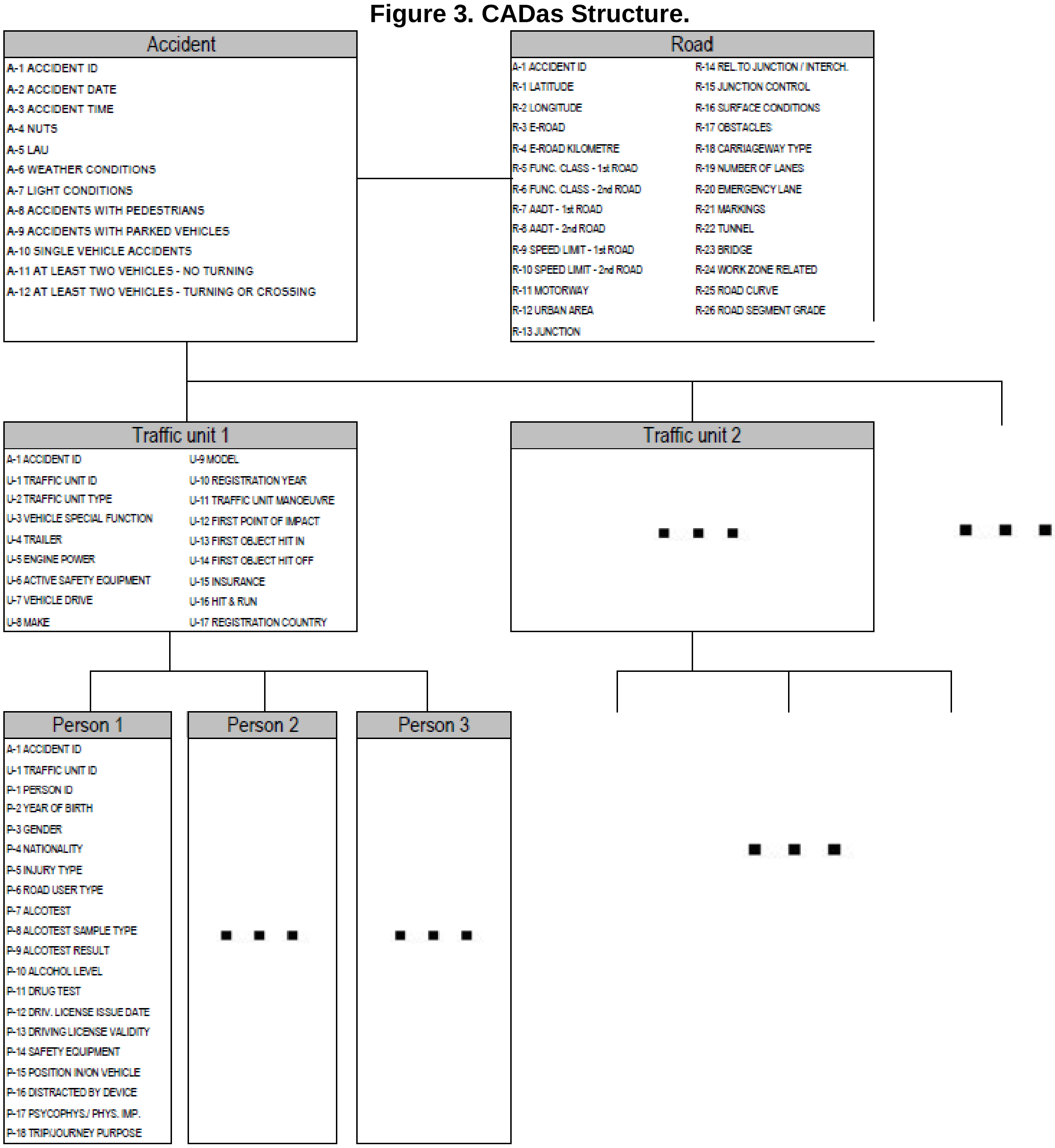
The variables collected are subdivided into four categories: A: Accident; R: Road; U: Traffic Unit (vehicle and pedestrian) and P: Person. Figure 2 depicts the classification of variables is classified into High (H) and Low (L) levels of importance (13).

Figure 2: CADaS variables and values in numbers.

category	Code	Number of Variables			Number of Values		
		High (H) importance	Lower (L) importance	Total	Detailed values	Alternative values (A)	Total
Accident	A	7	6	13	91	13	104
Road	R	12	13	25	92	13	105
Traffic Unit	U	8	10	18	181	15	196
Person	P	13	8	21	92	10	102
Total		40	37	77	456	51	507

Values are attributed to the variables and each value is identified by the code of the variable followed by the number which corresponds to each value and name, as shown in Figure 3 (12). Data formats concern the possibility to attribute one or more values to a variable; the format of the value is given by code, number and text. Values can be detailed values at the highest level of detail or alternative values related to information at a more aggregate level when detailed values are not available.

Figure 3. CADas Structure.



Data are publicly available and published in the annual report. Finally, the users are represented by the European Commission, EU Member States, national police road and the general public.

Points of comparison with fire incident data collection

Several points merit attention in considering EU road accident data collection in relation to the collection of data on fire incidents. Like the EU road accident data collection model, fire incident data collection generally relies on voluntary reporting, a feature that offers a starting point for comparison of the two data collection efforts. In this regard, it would be useful to know whether the mechanics of the CARE database, including its flexibility and gradual approach to implementation, are effective in facilitating compliance with reporting. However, we were unable to identify any published assessment of the completeness of road accident data, which are likely to vary from country to country. We can nevertheless note that EU road safety data are actively serving as a rich database for accident research and safety outreach (14-17).

A point of comparison between data collection in the CARE system and fire incident data collection is that preparing reports of vehicle accidents resulting in casualties is a more traditional practice for police officers and is likely to be an expected part of job responsibilities. For firefighters, filling out paperwork on fire incidents is a comparatively novel part of the job, one with less tradition behind it. This discrepancy suggests that there may be greater motivation to comply with reporting requirements in road accident data collection than is the case with fire incident reporting.

An additional factor likely to facilitate the reporting of road accident data in the CARE system is that the inclusion criteria are quite restrictive compared to fire incident data collection systems. Data collection in CARE is limited to road accidents that result in death or injury, specifically excluding incidents that only involve material damage. In contrast, fire incident data collection generally includes all fires attended by fire departments, even those resulting in minimal damage. Such a narrower focus in some respects relates to the impact of missing data in survey research and is discussed in the section which follows.

Although the CARE system does not call itself a census, its inclusion of all members of a population (fatal road accidents) at defined intervals in its dataset appears to meet basic census criteria. It seems likely that the narrower criteria for in-scope incidents in CARE data collection will result in more complete reporting than the reporting of fire incidents, whose broader inclusion criteria could be considered onerous for data collectors and provide greater opportunity for non-reporting. Consequently, this also suggests that if the goal of complete reporting of in-scope incidents is to be taken to fire incident data collection in the European Union, consideration should be given to limiting data collection to fatal fires, as is the practice in Spain.

4. Missing Data in Data Collection

This section of the report focuses upon missing data and its impact upon data quality in data collection. We introduce this section with a general overview of missing data in survey research – including a discussion of what missing data is, how it affects data quality, and some of the techniques for handling it in order to optimize the utility of results. The discussion then turns specific attention to missing data in fire incident reporting. This discussion identifies some of the factors that influence the reporting of data on fire incidents, as well as efforts to deal with missing data in existing fire incident data collection systems. The attention here primarily focuses upon an approach to missing data in the U.S. National Fire Incident Reporting System, where the problem seems to have received the greatest attention. We also review existing approaches to the treatment of missing data in member countries of the European Union, where information is available.

Survey research is an important method for collecting information from a population to learn more about issues or phenomena of interest. However, it's normal for research results to include some responses that go unanswered, and this can happen in a variety of ways and for a number of reasons (18-20). While missing data represent the norm rather than the exception in survey research, they pose a significant challenge for researchers. When data are missing from a survey, they can compromise inferences about the population under study, introduce bias into the estimates of data parameters, undermine the generalizability of results, and lead to errors of interpretation (1,4,19). Data estimates that are biased will convey information about a population that differs from that of the target population (19). It follows that conclusions based on that information will not be responsive to the phenomena of interest in the study population.

Data can go missing at different levels. Dong and Peng note that missing data can occur either at the unit level or item level, with unit level missing data representing the failure to obtain any data from a respondent, while item non-response refers to the failure of a respondent to complete all items in a survey (18). Newman goes somewhat further by providing for a third form of missing data in surveys whose measures include multi-item scales (19). He refers to "item-level" missingness as a situation in which a respondent leaves one or more items blank on a multi-item scale, while "construct-level" missingness refers to a respondent's failure to answer any items on a scale or construct and "person-level" missingness to a failure to respond to any part of a survey. The latter situation is equivalent to unit level non-response identified by Dong and Peng (18).

Missing data at the unit- or person-level is viewed as the most problematic because the researcher is provided with no information that could be used to improve estimates or reduce bias arising from missing data.

Although it is almost inevitable that survey results will experience some form of missing data, analysts have observed that researchers frequently fail to acknowledge missing data in presenting their results (19,20). The credibility and quality of research is argued to require the discussion of missing data and the conditions under which they occur, as well as the methods used to problems associated with missing data (20).

Missing data: How does it occur and how much data is missing?
The extent to which missing data pose a problem for research is a function of how much data is missing and the ways in which data are missing, which is to say, underlying patterns in which respondents fail to report data.

A review of missing data techniques by Schafer argues that missing data in the range of five percent or less has little to no impact upon the validity of data (21). Another view proposes that data analysis is likely to be biased if ten percent or more of data are missing (22). However, other commentators assert that the issue of how much data is missing is less important than patterns in the way in which data go missing (20,23). Here, the critical factor is the degree of randomness in missing data.

Studies of missing data commonly identify three missing data patterns that influence the potential for bias and subsequently impact the selection of appropriate techniques for dealing with incomplete data (19-20,22).

- Missing completely at random: Data that is missing completely at random refers to missing data that is unrelated to the variable of interest. Hence, research subjects with complete data are indistinguishable from those with incomplete data, and missing values can be presumed to be a random sub-sample of the actual data values (22). Data that is missing completely at random can be treated as unbiased. However, meeting the assumption that data is missing completely at random is seen to be extremely difficult in practice (22).

- Missing at random: The missingness of data that are missing at random is partly dependent upon other data, but not dependent on any of the values that are missing (19). In this case, respondents providing incomplete data do differ from respondents providing complete data, but the pattern in which data goes missing can be predicted from other variables in the dataset and is not contingent upon the specific variable with missing data (22). When data are missing at random, conclusions about population parameters under study are not contingent upon the way in which data are missing, and this mechanism of missing data can be ignored in the analysis stage as a result (22).

- Missing not at random: Data that are missing not at random refers to situations in which the missing value is related to the reason that it is missing. Non-random missing data cannot be predicted on the basis of other variables in a dataset. Dong and Peng illustrate the case of data missing not at random by referring to a situation where high-income earners might be more likely to withhold data on income than respondents with low or middle incomes (18). Data that is missing not at random has been referred to as “non-ignorable” missing data because it cannot be ignored in the analytical process (22). In order to deal with data missing not at random, researchers should be equipped to specify the missing mechanism and incorporate it into analysis in order to produce estimates that are not biased (18).

Methods for dealing with missing data

There are several statistical techniques that are recommended for dealing with missing data. Because reviews of these techniques tend to be highly technical and beyond the scope of the general reader, we will restrict our overview to basic descriptions of the primary methods and their relative merits. A discussion of specific techniques in current use for dealing with missing data in fire data collection is covered in an accompanying section of this report.

An approach to missing data known as listwise deletion involves the exclusion of records from analysis if any value is missing. Critics of listwise deletion consider it a blunt instrument that discards considerable amounts of information from respondents who have provided responses to some (perhaps most) but not all questions on a survey (19,24). It is also based on an assumption that missing data are missing completely at random (19,25). Consequently, listwise deletion is faulted for introducing bias into results through the exclusion of information that may not be random and for weakening statistical power by limiting the sample size (24,25).

An alternative approach, pairwise deletion, also excludes certain data from analysis but retains data that would be lost through listwise deletion practices. Pairwise deletion procedures exclude from analysis only variables which are missing values but use data for other variables with non-missing data. Not all respondents (cases) will fail to provide responses to the same variables, so different observations may be based on different cases, and the number of cases may vary from one analysis to another (20). Pairwise deletion also operates under the assumption that data is missing completely at random and is seen to increase the potential for bias and to reduce statistical power in ways similar to listwise deletion (20).

Imputation is another strategy for addressing missing data that attempts to avoid some of the weaknesses associated with deletion by replacing missing values with imputed values (19,26).

Single imputation involves replacing missing data with a value that is derived from other responses (18,26). One method is to replace missing values with the group mean for the relevant variable. Another method, referred to as "hot-deck" imputation, replaces missing values with values derived from respondents whose responses were a match on other variables (19,22). A third technique uses values predicted from multiple regression of observed cases to replace the missing data (23). Finally, "cold-deck" imputation uses externally derived information on respondents to impute a missing value based on responses from similar research subjects (22). Single imputation is viewed as a serviceable technique when small amounts of data are missing, but not when a dataset has large amounts of missing data or when data is missing completely at random (19). An additional disadvantage associated with single imputation is that analysis will treat imputed values as if they were identical to observed values, leading to underestimates of true variance among respondents.

Multiple imputation is an approach to missing data that imputes missing values multiple times in order to reduce the uncertainty associated with imputed values (18,19). Multiple imputation creates a set of m plausible values for each missing data point, producing m complete data sets. Bennett indicates that m typically ranges between five and ten (19). The researcher then analyzes the new datasets, each with a unique estimate of the missing values, and then incorporates the results into a single summary parameter and its associated standard error. Multiple imputation is recognized for minimizing bias associated with estimates that rely on single imputation techniques (18,19).

A final approach to missing data that merits attention is the maximum likelihood method. The maximum likelihood method does not impute any missing values, but instead estimates parameters of interest from an incomplete dataset (18,19). Dong and Peng found in their review of literature that 26.1 percent of studies that had missing data used maximum-likelihood procedures to resolve the missing data problem (19). Maximum likelihood uses structural equation modeling to create parameter estimates by maximizing the likelihood function using all available information, both complete and incomplete, from the dataset (19,24). These parameter values are assumed to have the highest probability of producing the sample data (27).

5. Approaches to Missing Data in Fire Incident Data Collection

Although missing data is the subject of a substantial research literature, it has received little attention in relation to fire incident data collection. There appears to be little acknowledgment that missing data is an issue in fire incident data collection systems, nor is there significant discussion of mechanisms for its treatment. Available research on missing data in fire incident data collection appears to largely focus on how it is handled in the United States, whose National Fire Incident Reporting System (NFIRS) is one of the largest and most complex reporting systems for fire incidents.

Missing data in the U.S. National Fire Incident Reporting System

In the United States, the National Fire Incident Reporting System (NFIRS) was introduced by the United States Fire Administration (USFA) in the mid-1970s to serve as a database capable of addressing national patterns for fires of all sizes by specific property use and specific fire cause. It seeks to provide the most detailed incident information on the nation's fire problem without being limited only to large fires. It also seeks to assist state and local authorities in identifying information about the fire problem for their own use. Given these ambitions and the heterogeneity of geographic, economic, infrastructure, and other fire-related influences in the United States, it should not be surprising that missing data in NFIRS data has been a critical issue since the system's inception.

Missing data arises at different levels in NFIRS reports. NFIRS is a voluntary system through which participating fire departments report detailed factors about the fires to which they respond. Roughly two-thirds of U.S. fire departments participate, although not all of these departments provide data every year. Different states and jurisdictions have different reporting requirements and practices and some fire departments don't submit reports for all fires that they have attended. Fire departments may fail to report fires due to time or resource constraints, including budgetary or personnel limitations. Some types of fires may also go routinely unreported because they are not considered meaningful. Such missing data is equivalent to failure of respondents to respond to any information in survey research, or unit-level missingness and may lead to an underreporting in the true number of fires.

The second form of missing data in NFIRS consists of incomplete or inaccurate reporting of data elements in individual incident records, a form of item-level missingness. NFIRS has a wide variety of data elements and codes, including variables for specific type of property, size of fire, equipment involved in ignition of a fire, heat sources, area of fire origin, factors contributing to the ignition of a fire, the extent of flame spread, automatic detection and suppression equipment, and victim information. In this case, data gaps may result from missing information in individual fire incident records.

NFPA approach to unreported fires in NFIRS

The National Fire Protection Association (NFPA) employs a "national estimates approach" to adjust for fires that are likely to be attended by fire departments but not reported to NFIRS (28). This approach represents an effort to deal with respondent level missingness in NFIRS by adjusting the number of reported fires and associated losses upwards through the application of a multiplier, which is derived from a second dataset that captures the broad national fire experience.

To create the multiplier, also referred to as a "scaling ratio," NFPA each year administers its Fire Experience Survey (FES) to a sample of fire departments around the country (28). Surveys are sent to all municipal departments protecting populations of 5,000 or more and a random sample, stratified by community size, of the smaller departments. Scaling ratios are obtained by comparing NFPA's projected totals of residential structure fires, non-residential structure fires, vehicle fires, and outside and other fires, and associated civilian deaths, civilian injuries, and direct property damage with comparable totals in NFIRS.

Analysts at the NFPA, the USFA and the Consumer Product Safety Commission (CPSC) developed the analytical rules used in analyzing data from the two data sets (28). The scaling ratio based on these two data sets produces a multiplier in which the projected numbers from NFPA's broad fire experience survey provide the numerator and numerical results from NFIRS act as the denominator, as indicated below:

$$\frac{\textit{NFPA's fire experience survey projections}}{\textit{NFIRS totals}}$$

Estimates of the total number of fires and losses are obtained by multiplying the NFIRS data results by the scaling ratios.

NFPA approach to missing data at the unit level

The variables in NFIRS frequently include a number of code choices that represent unknown data. In addition to variables that contain missing values, NFIRS includes code choices that indicate that the requested values are undetermined, unknown, or under investigation. NFPA treats variables with unknown values and those with missing values in the same way and applies a procedure which redistributes unknown values in order to produce an estimate of the true values (29).

For most fields, NFPA allocates unknown data proportionally among known data by calculating percentages based on the known data and then multiplying the totals by the percentages known. This approach assumes that if the missing data were known, it would be distributed in the same manner as the known data. NFPA makes additional adjustments to several fields.

The allocation of unknowns must be done separately for a category of fires that are classified as "confined fires" in NFIRS. These are small fires that do not spread beyond their object of origin, such as a cooking pan. Reports on these fires do not require detailed information, but certain analyses examine confined fires in greater detail when causal scenarios are of particular interest, as is the case with fires caused by heating equipment, smoking materials or playing with fire. Because the confined fire incident types describe certain scenarios, the distribution of unknown data differs from that of all fires, and allocation of unknowns must be done separately, while following the same proportional allocation methodology.

Other considerations

The counsel for researchers to minimize the potential for missing data through careful design of survey instruments and in data collection procedures is also an issue of some relevance to the fire data collection experience in the U.S. A central consideration here is that NFIRS is a large and detailed data collection instrument, while those who submit the reports are generally firefighters for whom data collection may be a distraction.

A study examining the problem of unknown data in NFIRS in 2014 found that information containing causal information from fire investigations was often not included in NFIRS incident reports because the reports were not updated after the information became available (3). The research also suggested that firefighters received insufficient training about the importance of data collection, that the reporting system was viewed by respondents as overly complex and not user-friendly, and that regular work responsibilities of firefighters were likely to undermine dedication to data reporting. The authors concluded that NFIRS program managers at the state level could be valuable resources for improving data quality by interacting with fire departments, supporting fire department participation, and providing quality control and feedback.

In a second study, researchers in 2016 examined the three-digit coding structure for types of fire incidents in NFIRS (30). In the NFIRS coding system, the first digit in incident codes represents a broad category (such as fire, explosion, etc.) while the second is narrower, (such as fire in structure or fire in mobile property used as fixed structure), and the third provides the most detail (such as building fire or fire in a structure other than a building). The researchers compared NFIRS narrative information from three large municipal fire departments with corresponding codes for type of incident, property use, and actions taken by the fire department. The authors were most likely to agree at the broadest level code level, but there was little agreement at the three-digit level. The authors pointed out that long lists reduce accuracy and that too many choices can result in indecision or a default choice. They concluded that the NFIRS coding structure was based upon the researcher preferences rather than taking account of the way firefighters think and process information.

Research has also looked at the potential impact of report-level missing data in the NFIRS data base. A 2012 study found that there were differences in socioeconomic and fire department characteristics between those cities that submit reports to NFIRS and those that do not (31). The authors argued that if those factors also affected the risk of fire, then generalizations made about fire safety and risk based only on NFIRS data will not apply to non-reporting areas of the United States.

In sum, a variety of factors help to explain missing data in NFIRS reporting. The voluntary nature of NFIRS reporting and different state-level requirements for fire incident reporting help to explain unit-level missingness. In addition, research has shown that multiple considerations influence the incomplete reporting of incident variables on reporting forms. Firefighters are not trained researchers and have substantial responsibilities that can diminish attention to data collection and reporting. Liability concerns may lead fire departments to discourage complete reporting of certain types of information, such as causes of the fire. Funding and resource limitations can undermine support for data collection, including participation in training and access to computer and software support. Another issue with respect to the completeness and accuracy of reporting is that codes are seen to be overly complex, resulting in frustration that can deter reporting.

Missing data in the United Kingdom

The Community database on Accidents on the Roads in Europe published a regulatory standard in 2015 for the quality assurance of administrative data which applies a number of quality assurance checks. In keeping with that standard, the Home Office has adopted several procedures to promote the quality of data reported in the Incident Reporting System for incidents attended by fire and rescue services. A variety of the procedures directly or indirectly attempt to address problems relating to missing data or data that is incorrectly recorded (32).

The Incident Reporting System, for instance, includes automatic checks to ensure that only applicable questions are answered, all dates and times are complete and entered in the correct format, and that only appropriate options are displayed to system users. A quality assurance team conducts a check of submitted information.

In addition, statisticians at the Home Office carry out a monthly monitoring process that looks for unusual patterns in the data and examines data gaps and conducts variance checks in order to identify entries that seem unusually large or small relative to figures for the same month in the prior year.

In order to address unit-level missingness, tables are footnoted when a significant number of incidents are determined to be missing. The level of missing data in data fields is said to be very low (32). Missing data are reported as unknown, and no imputation or other estimation methods are used to deal with such item-level missingness.

In Wales, statisticians occasionally record data as missing and seek to impute data if there are time and resource constraints with the submission of incident reports (33). Data providers are informed of any imputation or changes to data and provided an opportunity to challenge or comment upon the changes. This is considered to be a compromise to resolve validation issues and to minimally impact the usability of the dataset.

Missing Data in the European Union

In general, the issue of missing data in fire data collection systems has not received the same level of attention among member states of the European Union as it has in the United States. It is not clear if this is because missing data is less of a problem in these countries because the reporting forms are less complex and reporting compliance is more complete or because problems with missing data are less likely to be acknowledged or to be handled at the analytical stage. We were able to identify efforts in the United Kingdom, Sweden, and France that attempt to address missing data in their fire data collection systems and we briefly highlight those below.

Approaches to missing data in Sweden

The Swedish Civil Contingencies Agency (MSB) has a national system for collecting data from incident reports that are recorded by local fire departments in their own records management systems. Prior to 2018, MSB could not be sure that it had received all relevant incident reports, and MSB on occasion discovered underreporting for a specific fire department during a specific period. As a result, a procedure was introduced in 2018 to identify underreporting at the central level. The local data collection system sends a message to the national system when a report is initiated. Every month MSB sends feedback to fire departments on reports that have been initiated but remain uncompleted in the national dataset.

The law governing the work of fire departments was revised on January 1, 2020, and it is now mandatory for all fire departments to send their incident reports to MSB (34). It is currently assumed that there is no underreporting of actual fires, though it is apparent that a small number of initiated reports are not completed. In the latest statistics published for the year 2020, 645 out of 138,000 reports were incomplete (35). MSB contacts fire departments with incomplete incident reports and is often told that they involve incoming telephone calls that did not require a fire department response or were dealt with by a neighbouring department.

Both the local and national fire data collection systems implement logical checks which prevent a report being sent with missing values. The local system should not send a report with missing values, and MSB's system will not read a report with missing values into the national database. In the latest statistics for 2020, three reports were submitted to MSB with missing values and were not included in the national statistics (35). It should be noted that some variables have a relatively high proportion of values that are recorded as "undetermined" or "unknown," but MSB publish the data as received.

When publishing statistics, MSB has never made estimates to mitigate for missing data. MSB makes the judgement that it is better to accept this loss of precision in the statistics, rather than run the risk of introducing bias in the material. The approach adopted is in line with the practice for European statistical authorities as formulated by Eurostat, where the quality in the data is described in a quality declaration. It is then up to the user of the statistics to decide how to deal with missing data from underreporting or other shortcomings in data quality.

Approaches to missing data in France
France uses a weighted average to deal with unit-level missing data in the number of fire interventions reported by fire departments. When data on fire interventions are not reported by a fire department, the Ministry of Interior calculates the weighted average number of interventions by fire departments protecting populations of similar size and applies that number to the data for the non-reporting fire department.

Discussion
Missing data takes different forms and varies by amount of incomplete data in a dataset. Appropriate methods for dealing with missing data in any specific situation will hinge on the patterns and degree missing data, and it is important that the technique of choice is one that meets methodological assumptions in addressing the specific form of missingness. While there is no generalizable best approach to dealing with missing data, there is considerable agreement that doing nothing about missing data is a serious mistake, inasmuch as missing data can introduce bias into results and lead to conclusions that are inappropriate for an actual study population.

Deleting cases with any amount of missing data is a strategy that appears to be no longer in vogue among analysts due to advances in statistical procedures and the problems associated with biased results. Researchers are encouraged to make use of all the data that they collect, while exercising caution if a substantial percentage of values are missing for a variable of interest. In deciding how to handle missing data, researchers should be familiar with issues relating to their technique of choice, including its statistical assumptions and proper implementation. With the application of suitable techniques, researchers can utilize datasets with missing data and still produce valid results.

There are practical reasons with potential real-life consequences to address missing data in fire data collection. For instance, data on the number and distribution of fires in a town, region, or country can influence funding for fire protection, including siting of fire stations and number of personnel. When data reporting systems fail to capture all fires, budgets may be inadequate to the level of fire protection needs. Other problems may arise when data on fires don't accurately reflect the nature of the fire problem. Data collection systems that only include fires from large cities or more developed areas may miss unique fire experiences in rural areas or poor neighborhoods. Fire safety hazards in underrepresented areas, including needs of special populations, would subsequently go unaddressed in planning and prevention programs.

Finally, consistent and relatively complete data is important for evaluating the efficacy of prevention efforts, for without it, it is difficult to assess the accuracy of fire trends.

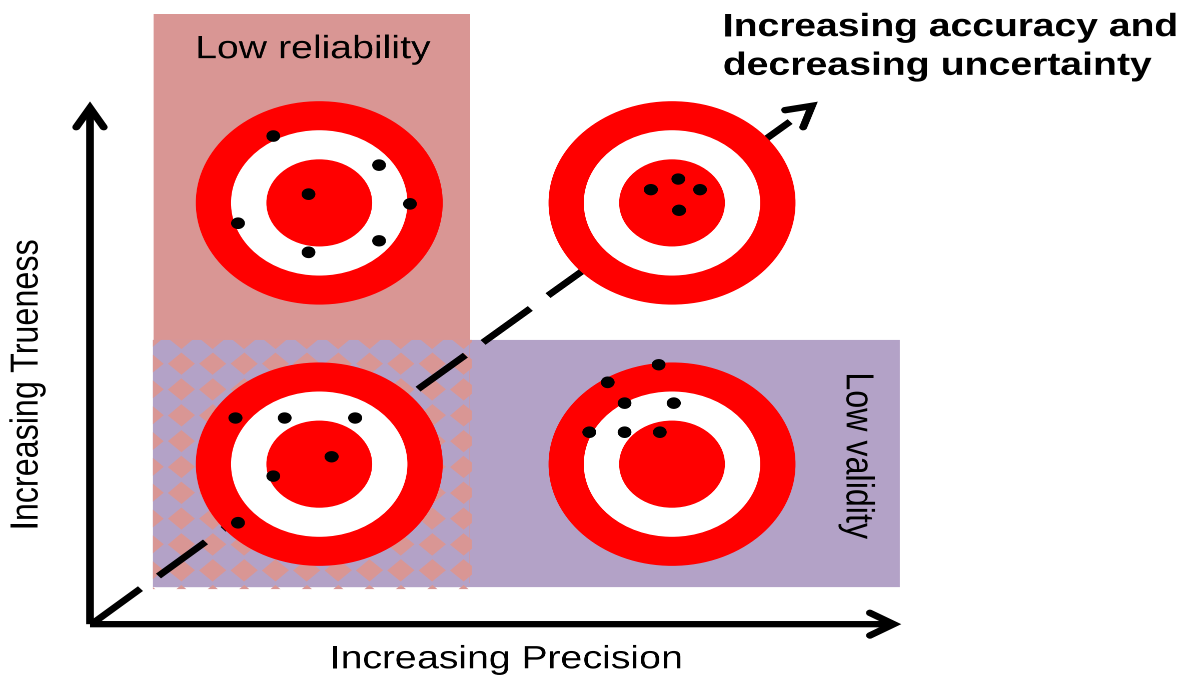
6. Uncertainty Analysis and Statistical Reliability

This section of the report utilizes qualitative analysis to identify points of uncertainty from the collection through analysis phases of fire data collection. In addition, we offer suggestions for potential ways to reduce uncertainties in fire data collection and identify areas for future study to better quantify uncertainties in fire statistics.

What is uncertainty?

The Cambridge Dictionary defines uncertainty as "a situation in which something is not known, or something that is not known or certain" (36). In its broadest sense, "uncertainty of measurement" refers to doubt about the validity of the result of a measurement (37). Uncertainty relates to how an estimate might differ from a "true value"; a reduced uncertainty will result in an increased accuracy, where accuracy is defined as the closeness of agreement between a measured value and the true value (38). Uncertainty in relation to measurements of some physical phenomena can be illustrated with the help of the terms "precision" and "trueness." The relationship between precision, trueness and accuracy is illustrated in **Erreur ! Source du renvoi introuvable.** (39). The centre of each circle in Figure 4 corresponds to the "true value" of the variable being measured while the black dots represent the observations. Precision increases as the scatter of the measurement becomes smaller, while trueness increases as the average of measurements is closer to the centre. Low reliability is related to a large scatter due to random errors, while low validity is related to systematic errors (or bias). Even though Figure 4 is connected to physical measurements in science and engineering, the concepts are considered useful to illustrate uncertainty and error in surveys.

Figure 4: A description of the relation between precision, trueness, accuracy, and uncertainty.



Sample errors and non-sample errors

A cost-effective way of conducting a statistical survey is to collect information from a sample of the population. Since sample surveys don't collect information from the entire population, the results will represent estimates of the unknown population values. Thus, it is important that the sample represents the population to as high a degree as possible. The types of error that can occur in a sample survey are categorized as sampling errors and non-sampling errors (40).

Sampling errors arise because the estimates derived from a sample (sub-set of fires) are likely to differ from the unknown population (all fires) value, and this can affect both the reliability and the validity of the data. Careful design of the sample helps to ensure representativeness and increase reliability (41).

Other types of errors (non-sampling errors) are usually very difficult to quantify, and specific studies are usually required to make such assessments (42). These errors are also associated with census surveys. Examples of non-sampling errors include:

- Cases when data is not collected
- Faulty inputs due to ignorant reporter, or fatigue when reporting
- Inaccurate answers (forms filled in inaccurately, questions misinterpreted, definitions misunderstood)
- Errors when processing, presenting, or analysing data

Three different data collection techniques were considered in Section 3 of this report and the relationship between them, and types of sampling error are presented in Table 1. Sampling errors are only to a limited degree present in census studies because the survey is done on the entire population (for example all fatal building fires). Sample errors will be present in sample surveys, but these can be addressed by designing the sample using a scientific approach. The standard error, coefficient of variation and confidence interval are measures that can be used to help interpret the possible sampling error. Statistical significance tests can be used to investigate if difference between estimates from different samples is caused by a real change in the population, or whether it is likely due to the effects of random sampling (42).

Convenience collection will be associated with both sampling errors and non-sampling errors, and it will be difficult or even impossible to reduce sampling error. Thus, it will be very difficult extrapolate any result to the whole population.

Table 1: Relationship between different data collection techniques and types of error.

Data collection method	Sampling errors	Non-sampling errors
Census	Limited	Yes
Sample survey	Yes, but can be treated	Yes
Convenience collection	Yes	Yes

Increasing accuracy and reducing uncertainty in fire statistics

The ESS Handbook for Quality and Metadata Reports (43) supports the *European Statistics Code of Practice* (44) by providing recommendations on how to prepare comprehensive quality reports for the full range of statistical processes and their outputs. The handbook includes a section on accuracy and reliability. Five different types of non-sampling errors are raised in the handbook:

- Coverage error
- Measurement error
- Response error
- Processing error
- Model assumption error

Coverage error is due to a divergence between the survey population and the target population. This can be problematic in all the data collection methods in Table 1. Matching with different registers is one way of minimizing the problem that is mentioned in the handbook. There are examples when this has been utilized in specific studies of fire statistics. Research by Jonsson and Bergqvist found that 20% of the fire fatalities in the fire fatalities database for years 1999-2007 in Sweden were missing (47). A similar study in France found possible underreporting of 30% (45-49). A possible way to either reduce or quantify the coverage error in a census system is to establish a connection between fire service dispatch system and the reporting

system, as is done in the Nordic countries. As an example, 645 reports were initiated but never finalized in Sweden during 2020, representing 0.5% of the total amount of reports that year (50).
Errors that occur during data collection and cause the recorded values of variables to be different from their true values are called measurements errors. The main issues regarding measurement error and fire statistics are likely to be associated with the survey instrument and the reporter. For instance, the wording of questions in the survey instrument or the order or context in which they are presented might lead to measurement errors. Alternatively, the reporter might provide erroneous data due to confusion, ignorance, or carelessness. Inadequate training and knowledge of the reporter might also lead to measurement errors. As an example, the NFPA study of NFIRS coding referenced earlier found that the incident type instructions did not include a clear definition of "fire," undermining accuracy (30). This illustrates the need for those involved in developing surveys to understand the target group and their language needs. The definitions of terms and the wording of instructions should be appropriate for survey reporters.

Measurement errors can be either random or systematic. Random errors will affect precision (see Figure 4) and the effect can generally be reduced if the dataset is large. Random errors may have a large influence on results in small countries or for fire events that are relatively rare, such as example fatal fires or resource-demanding fires. In such cases, a longer time series of incidents will need to be studied or specific attention drawn to the event (for example by performing more detailed fire investigations). Systematic errors (or bias) affect trueness (see Figure 4) and can be even more difficult to treat. Systematic errors are in general treated with specific evaluation studies or parallel data collection by different reporters, but this is not feasible for all types of fires. An approach to this problem in some European countries, such as Austria, Sweden and the Netherlands (see Task 0 report), is to use information from more detailed investigations to complement or update the initial data collection. In Sweden, MSB performs some re-examination of all reported fatal fires in order to address the potential for systematic error.

Response errors occur either when no data is collected from a unit or when some but not all data from the variables are collected. The response error becomes more problematic when it is associated with a bias. Specifically developed studies or connections to other registers could be used to reduce the problem of non-response, together with the previous mentioned strategies for reducing coverage and measurement errors. A further discussion on response errors and missing data is available in Section 5.

Processing errors refer to errors arising from the faulty implementation of correctly planned implementation methods in the final data collection (43). Processing errors include all post-collection operations, such as errors of transcription, data editing, data coding, aggregating, weighting, and errors in programming (43, 51). Processing errors affecting individual observations will, as with measurement errors, give rise to systematic errors. Careful verification of data processing is necessary to limit potential processing errors.

Model assumption errors are connected to specific models that are used in a domain, for example to improve the precision or adjust for measurement errors or non-response. A relevant area in fire statistics is the use of adjustments for "unknown cause of fire". "Unknown" could be assumed to be distributed in the same way as all other causes, or possibly in some more profound manner. When implementing such models, it is necessary to carefully assess the validity of the model.

In conclusion, the different non-sampling errors needs to be addressed separately. In some cases (like with random errors) the error can be reduced by increasing the sample size. However, in many cases specific studies or investigations are needed to be able to estimate the error.

Uncertainties connected with variables proposed for collection

A total of twelve variables have been suggested for inclusion in harmonized European fire statistics in Task 2. Below we present a qualitative assessment of possible uncertainty issues connected to these variables.

Number of fatalities

As seen in studies in both Sweden (45) and France (46), underreporting of fatalities can be as high as 20-30%. Underreporting may occur when a victim dies after transport from the fire scene or because the fire service is never called to the scene, or for some other reason. Studies of hospital records can help in determining the actual number of fire fatalities, but automated solutions can be complicated since personal information (like social security number) is seldom recorded by the fire service. Even so, the best data quality check is to regularly perform specific studies which compare hospital records of fire fatalities with the outcomes from fire statistics. Indeed, medical data (for example those based on ICD10) should be cross-referenced with other sources to find an agreement. Similar initiatives for injury related mortalities have been examined across Europe (47, 49). In Sweden, MSB hopes to follow up all reported fatal fires by collecting supplementary information from the Police and the Board of Forensic Medicine.

Number of injuries
Fire injuries are even more difficult to systematically record than fatalities. It is likely that the fire service will keep track of how many people they rescue (52), but people with injuries might evacuate by themselves or with the assistance of others than the fire service. Although firefighters will be able to collect some injury data on the scene, they might not be competent to evaluate injury severity, complicating data quality. In cases where people are transported by ambulance from the scene, the data can be used to perform specific studies of the accuracy of fire service reports of injuries, as in the case of fire fatalities referenced above. However, there are likely to be situations in which injury victims may be transported by family or friends before fire service arrival. Complications in recording injury is illustrated by comparing France and Italy, which are similar with respect to populations, building methods and fatalities per 100.000 inhabitants, but which have completely different outcomes for injuries, leading to a doubt about the difference in the definitions. High level comparisons between countries could be an appropriate tool for this variable that can help identifying major discrepancies.

Fire cause
Fire cause is likely to be prone both to measurement and response errors, since confusion, ignorance, or carelessness of the reporter might result in faulty inputs. Another complicating factor is that evidence at the scene may have been destroyed by the fire. The reporter may also feel an uncertainty or unease when assigning the fire cause, which results in assigning it as unknown. In NFPA analyses of NFIRS data, the unknown fires are distributed in the same proportion as the fires for which the data are known (29). However, this might lead to model assumption errors. When detailed fire investigations occur, they can be used to update the fire cause first assigned, and thus improve the accuracy. Even so, the destructive nature of fires can result in it being impossible to determine the fire cause.

Type of building
This variable may also be prone to measurement errors due to confusion or ignorance. As an example, there are different views in different countries regarding what is included in the term “residential building”. Holiday homes are considered residential in some countries but not others. A category like “public building” might also be interpreted differently in different countries. Clear definitions and instructions to the reporter are needed on how to interpret the variable and the different categories. An additional possibility is that building information can be double-checked with real estate information or records at a municipal level.

Incident location
If the report of a fire incident reporting is connected to a dispatch system where the location is recorded (address and/or coordinates), the uncertainties of incident location can be reduced. However, there might be problems with measurement errors (faulty inputs by dispatcher or reporter at the scene) and there might also be non-responses (for example address missing in the report). Possible errors can be reduced if both address and Global Positioning System (GPS) coordinates are reported, as seen in a study where fatal fires in Sweden were connected to real estate information by utilizing both information on address and coordinate (53). In some cases, address information was lacking and data on coordinates could be used, in other cases the coordinates were wrong, and the address could be used.

Incident date
If incident reporting is connected to a dispatch system where the time and date for call received, unit dispatched and unit at fire scene are automatically recorded, the uncertainties regarding this variable are considered small. If the variable is entered manually, it will potentially be prone to measurement and response errors (incidents that occur close to midnight will likely be most affected). Systematic errors may also occur but are most often likely to be random in nature.

Incident time
The uncertainty connected to incident time is considered to be small if the time and date is collected and recorded automatically. Errors are more likely if incident time is recorded manually, but errors are again likely to be random. In cases where incident time is recorded as a rough estimate of the time (e.g., night, morning, noon, afternoon, evening) the error will most likely be small.

Age of victim
There are a number of uncertainties regarding the ability of fire service to record the age of a victim at a fire scene. For example, there may be no one at the scene to attest to the age of the deceased in the event of a fatal fire. Age information will be available in other databases if the victim has been hospitalised or is deceased. Cross references to such databases can be made in order to quantify and evaluate the information in the fire service database.

Number of floors
The number of floors in a building should be quite straightforward to report if the variable is well defined and understood by the reporter. It must be clear for the reporter how to interpret basement floors, attic floors, mezzanine floors and ground level for uneven floors. Studies of the accuracy of this variable can be done by studying documentation and images of fire-exposed buildings.

Room of origin
The room of origin will most likely be associated with less uncertainty than fire cause. If the building is still standing or there are some cues based on eyewitness information, the room of origin should be easy to determine. Still, distinct categories are necessary to avoid systematic errors. As an example, a category labelled as “storage” could be interpreted as designated storage room or as a room used for storage in a basement. The latter can be confusing if “basement” is itself a possible category.

Source of ignition (or heat source)
Similar to fire cause, it might be difficult to determine the source of ignition (or heat source) due to the destructive potential of the fire. The reporter might need to rely on second-hand data, such as information from residents or other eyewitnesses if the fire itself has destroyed cues to the source of ignition. The category “unknown” might cause issues with source of ignition for similar reasons as fire cause (see above). Problems can also arise if the reporter is confused or unable able to distinguish between the fire cause and the source of ignition.

Material mainly responsible for fire development
The uncertainty connected to this variable is considered similar to fire cause and source of ignition. As long as the fire is kept in the room of origin, the damage will most likely not be too severe to be able to determine the material mainly responsible for the fire development. However, there might be situations when relevant knowledge in fire development and fire dynamics is required by the reporter in order to accurately categorise this variable.

Summary of uncertainties connected to the described variables
Based on the description and discussion above, we can make some general estimates of the uncertainty associated with the different variable assigned in
Table ***2***. The estimates are rough qualitative estimates that indicate which variables can be expected to be affected by the largest degree of uncertainty. It should also be stressed that these uncertainties can be reduced by applying different measures, like the measures discussed above.

Table 2: Estimated associated uncertainties with the variables suggested in Task 2.

Variable	Estimated associated uncertainty
Number of fatalities	Medium
Number of injuries	High
Fire cause	High
Type of building	Medium
Incident location	Low
Incident date	Low
Incident time	Low
Age of victim	High
Number of floors	Low
Room of origin	Low
Source of ignition (or heat source)	High
Material mainly responsible for fire development	Medium

Suggestions for future work

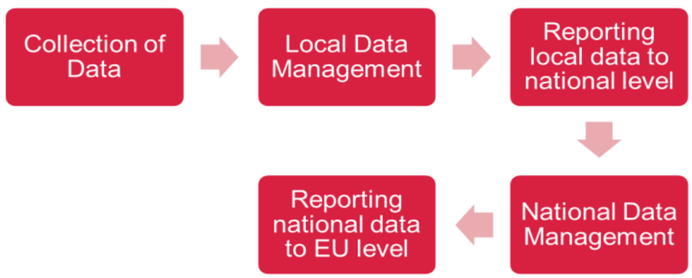
It is considered very difficult to quantify uncertainties and errors connected to fire statistics without doing specific studies and detailed evaluation of the different variables. Regarding fire fatalities, as previously mentioned, such studies have been conducted in Sweden and France. T Further study of variables more prone to error is needed in order to better quantify uncertainties.

An experimental methodology, so-called round robin tests, are sometimes used for inter-laboratory tests to determine the reproducibility of an experiment or test method. In a round robin, multiple independent laboratories or individual scientists perform the same task to quantify the reproducibility or the possible spread in the result. Such round robin studies have also been conducted to study variation in fire safety analysis (54, 55) and reviews of fire safety designs by the fire service (56). Such studies could possibly also be performed to quantify the variation in reporting data from fire incidents, although this exercise will not allow evaluating the measurement and response errors.

7. Cost Estimates for Data Collection

Regardless of the data collection methodology proposed, reports must be collected at the response level for each incident, according to the data collection threshold. Local policies might dictate that all incidents are reported into the system. Figure 5 provides a graphical representation of the data journey from collection at the point of incident to reporting of data at EU level. Note that the fire data journey as well as recommended timing for each step is discussed in detail in Task 7.

Figure 5. Graphical representation of data journey from collection at point of incident to reporting of data at EU level.



Local level

Reporting forms should be filled out at the local level. Reports could be completed by responding firefighters or officers within hours of the incident, with or without additional information added by investigators. This requires a computer in each station, a firefighter to fill out the form, an officer to review the form for completeness, and a database or incident reporting system to hold the file and produce reports at the local level. Training on the system would occur on an on-going basis.

It has been difficult to identify a cost point for data collection and management at the local level. The only identified source is USFA,[1] which estimates an average time 0.45 hours per incident to fill out an NFIRS report and the management cost is related only to the investment in and operations and maintenance of computer systems ($48 and $84 per fire station per year). The need for continuous training was not included in the estimate.

Since the cost at local level will be the same regardless of the data collection methodology, this section will focus on the cost of data management and reporting at national level where differences between the methodologies are apparent. The cost estimates presented do not take into account any potential savings from the use of existing systems in each member state. It was not possible to obtain this information within the time and cost constraints of the project.

National level: Census

To estimate the national costs of EU-wide fire statistics, the following parameters are taken into account:

Number of fires	(events)	[FIRE]
Average annual hours worked per worker	(hours per year)	[AHW]
Labour cost from Eurostat (57)	(€ per hour)	[LCH]
Price level indices EU from Eurostat (58)	(%)	[PLI]
Average hours worked per fire event	(hours per event)	[AHFE]

[1] Supporting Statement for Paperwork Reduction Act Submissions, OMB Control Number: 1660 - 0069
Title: National Fire Incident Reporting System (NFIRS) v5.0; March 10, 2021

The following cost factors are calculated from these parameters:

Total hours for performance of yearly fire statistics	(hours per year)	[THP]
Total labour costs	(€)	[TLC]
Number of staff	(workers)	[NSTF]

The total hours for performance of yearly fire statistics are determined with the following relation:
THP = FIRE x AHFE

The corresponding national values for [FIRE] can be found in Annex 2 Table 2-1. The value for [AHFE] is an estimate by the CTIF, based on the 25 years of experience of the Center of Fire Statistics. Official information on this parameter cannot be determined. The [AHFE] is set at 0.25 hours per fire event for all EU member states. This value is significantly lower than what is estimated by USFA for data collection at the local level.

The total labour costs (in €) are determined with the following relation: TLC = THP x LCH

The corresponding national values for the total hours for performance of yearly fire statistics [THP] calculated as shown above. The labour cost (€ per hour) [LCH] can be found in Annex 2 Table 2-2.

The number of national fire statistics staff [NSTF] is determined with the relation: NSTF = THP/AHW

The corresponding national values for the total hours for performance of yearly fire statistics [THP] calculated as shown above, and the average annual hours worked per worker [AHW] can be found in Annex 2, Table 2-2.

The determined total labour costs [TLC] (see 2) are the most important component of the national costs of fire statistics. The following cost categories are to be added: Hardware, Software, Other costs, including training, refresher courses, etc.

From surveys by the CTIF in recent years, the following ratios are used to estimate the total costs:

> [TLC] - the total labour costs (in €) are included in the calculation with 80%.
> [HARD] - the procurement and renewal of hardware is included in the calculation with 5%.
> [SOFT] software is a very sensitive cost parameter and should be included in the calculation with 10%.
> [OTH] - the remaining 5% are accounted for by other costs, including training, refresher courses, etc.

The total costs of the national fire statistics are ultimately determined from the following relation:

TOTAL = TLC + HARD + SOFT + OTH

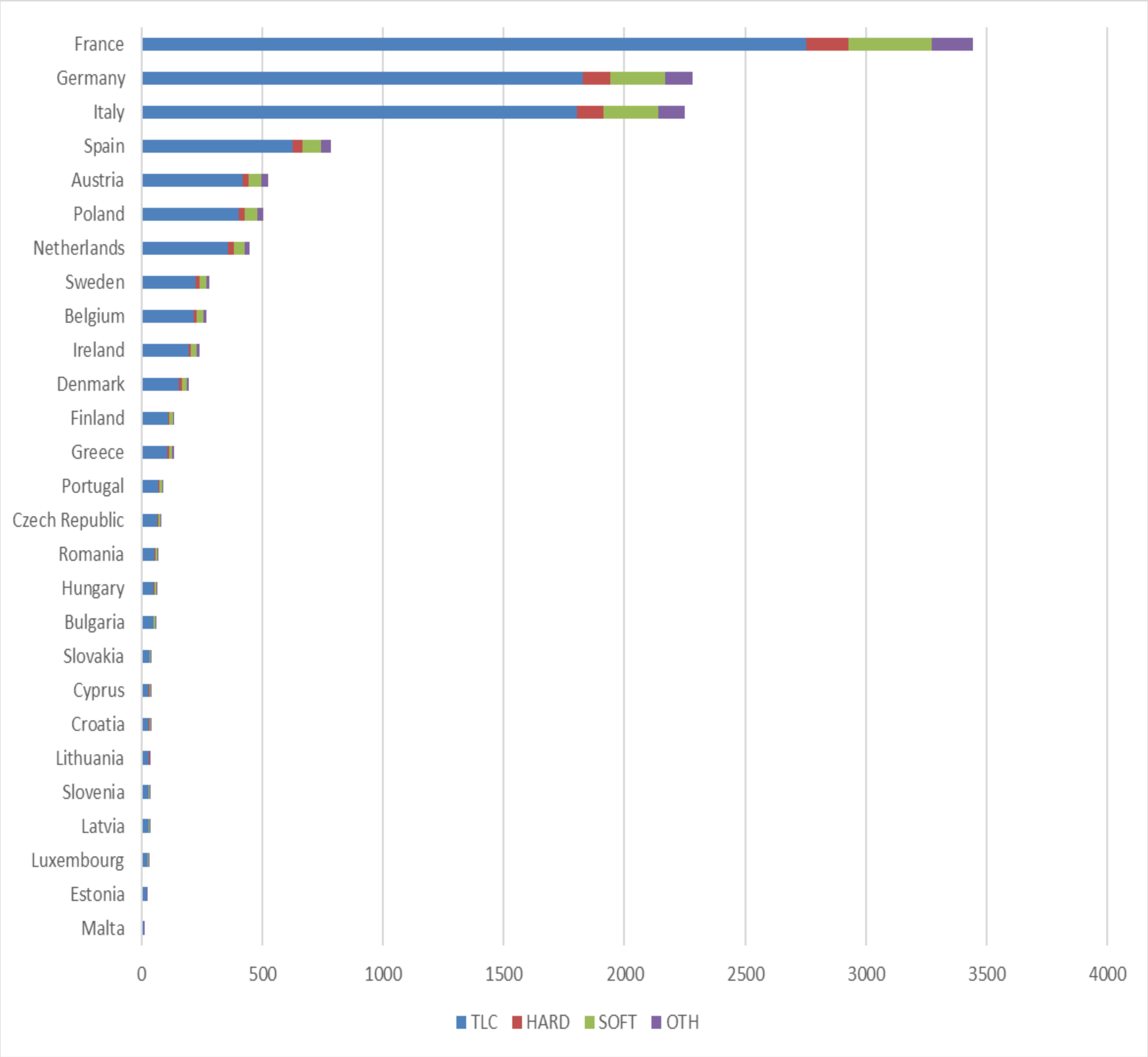
Calculations for each member state are shown in detail in Annex 3. Table 3 and figure 6 provide an overview of the estimated cost at national level for a census survey. It should be noted however that these estimates assume no collecting system is in place. If a nation already has such a system in place that collects the data that fits directly into the European statistics the cost will be very different and much lower. When comparing the cost between member states, it should be kept in mind that the result for each country is directly related to the number of fires in each member state as reported to CTIF. It is expected that these numbers will change significantly when definitions are harmonised. It was unfortunately not possible to extract the number of structure fires for all member states from the data available causing this calculation to be related to all fires and not just structure fires.

Table 3. Estimated cost of census data collection for EU Member States

No.	Member state	Cost category (in €)				NSTF (worker)
		TLC	HARD	SOFT	OTH	
1	Austria	417921.25	26120.08	52240.16	26120.08	8
2	Belgium	215775.00	13485.94	26971.88	13485.94	4
3	Bulgaria	47853.00	2990.81	5981.63	2990.81	5
4	Croatia	29748.60	1859.29	3718.58	1859.29	2
5	Cyprus	29750.00	1859.38	3718.75	1859.38	1
6	Czech Republic	64401.75	4025.11	8050.22	4025.11	3
7	Denmark	156418.45	9776.15	19552.31	9776.15	3
8	Estonia	18757.80	1172.36	2344.73	1172.36	1
9	Finland	108773.88	6798.37	13596.73	6798.37	2
10	France	2755059.38	172191.21	344382.42	172191.21	52
11	Germany	1826358.30	114147.39	228294.79	114147.39	37
12	Greece	107158.68	6697.42	13394.83	6697.42	4
13	Hungary	50878.58	3179.91	6359.82	3179.91	3
14	Ireland	193291.28	12080.70	24161.41	12080.70	3
15	Italy	1801164.15	112572.76	225145.52	112572.76	39
16	Latvia	27339.38	1708.71	3417.42	1708.71	2
17	Lithuania	29350.60	1834.41	3668.83	1834.41	2
18	Luxembourg	23807.55	1487.97	2975.94	1487.97	0
19	Malta	6684.50	417.78	835.56	417.78	0
20	Netherlands	357880.00	22367.50	44735.00	22367.50	7
21	Poland	402402.00	25150.13	50300.25	25150.13	21
22	Portugal	70875.00	4429.69	8859.38	4429.69	5
23	Romania	55073.93	3442.12	6884.24	3442.12	4
24	Slovakia	32414.60	2025.91	4051.83	2025.91	2
25	Slovenia	28223.18	1763.95	3527.90	1763.95	1
26	Spain	627000.00	39187.50	78375.00	39187.50	17
27	Sweden	225581.08	14098.82	28197.63	14098.82	4
Total		**9709941.88**	**606871.367**	**1213742.73**	**606871.367**	**232**

* The value has been rounded up to a positive whole number

Figure 6. Total cost of census at national level in 1000 Euros.



National level: Survey sample

For a survey sample, it is necessary to start with the design of the survey and identify which fire departments will be asked to report data. As explained previously designing the survey should be done based on significant local research. The design of the survey is directly linked to the estimation method applied to the collected data to achieve national estimates. There are no reliable estimates of this initial cost, which is largely dependent on local populations and socioeconomic factors. After the initial survey design, it should only be necessary to check its continued appropriateness every 5 – 10 years. Hence the annual cost is only related to the operation of the survey and analysis of results.

The Fire Experience Survey (FES) administered by NFPA is the sole example of a sample survey used for the projection of national estimates of fires and losses (28). The annual cost of maintaining and executing this survey are as follows according to data from NFPA for 2019:

- System for collecting survey responses (e.g., Snap Survey) = $3,000
- Maintaining directory of fire departments requires 0.5 Full Time Employee (FTE)/year @ $60,000/year = $30,000

- Staff to quality control responses and contact Fire Departments for missing responses requires 1.5 FTE/year @ $60,000/year = $ 90,000
- Project mgmt. and statistician for analysis and reporting requires 0.5 FTE/year @ $120,000/year = $ 60,000
- Operating cost including printing, postage, mailing, phoning = $40,000
- Total cost per year is $ 223,000.
- The U.S. had an estimated 1,300,000 fires in 2019.

Using the data from NFPA's Fire Experience Survey, the parameters needed to calculate cost in each EU member state can be derived as follows:

- Average hours worked per fire event [AHFE] = 2.5 FTE x 1,800 hours per year/1,300,000 = 0.0035
- The operational cost [OPR] amounts to approximately 100 x 43,000/180,000 = 24% of the labor cost.

Using these values, the cost at the national level for a survey sample collection of fire data per country can be calculated as follows:

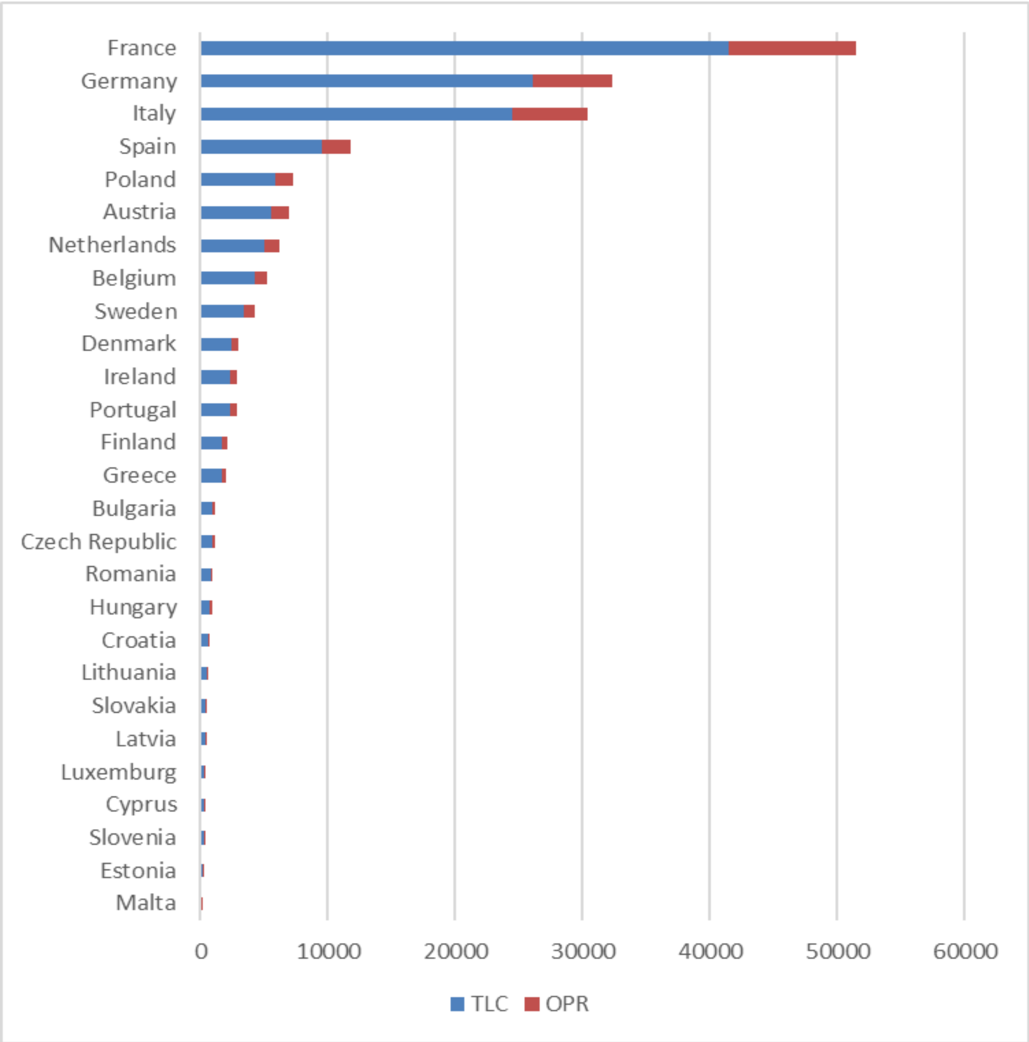
Total hours for performance of yearly fire survey sample	THP = AHFE x FIRE
Total Labour Cost	TLC = THP x LCH
Operational Cost	OPR = 24% of TLC
The number of national fire statistics staff [NSTF] needed	NSTF = THP/AHW

The results per country are shown in Table 4 and figure 7.

Table 4. Estimated cost of survey data collection for EU Member States

	Country	TLC	OPR	Total cost	NSTF
1	Austria	5571	1337	6908	0.11
2	Belgium	4261	1023	5284	0.07
3	Bulgaria	959	230	1189	0.09
4	Croatia	566	136	702	0.03
5	Cyprus	321	77	398	0.01
6	Czech Republic	902	216	1118	0.04
7	Denmark	2417	580	2998	0.04
8	Estonia	223	53	276	0.01
9	Finland	1712	411	2123	0.03
10	France	41488	9957	51445	0.79
11	Germany	26058	6254	32312	0.53
12	Greece	1643	394	2038	0.06
13	Hungary	725	174	899	0.04
14	Ireland	2346	563	2910	0.04
15	Italy	24477	5874	30351	0.53
16	Latvia	371	89	460	0.02
17	Lithuania	465	112	577	0.03
18	Luxemburg	328	79	407	0.01
19	Malta	89	21	110	0.00
20	Netherlands	5010	1202	6213	0.10
21	Poland	5911	1419	7329	0.30
22	Portugal	2330	559	2889	0.09
23	Romania	788	189	977	0.06
24	Slovakia	450	108	558	0.02
25	Slovenia	308	74	382	0.01
26	Spain	9488	2277	11765	0.26
27	Sweden	3452	829	4281	0.06
	Total	142660	34238	176899	3.40

Figure 7. Total cost of sample survey at national level in Euros.



Convenience sample

As explained earlier a convenience sample would fail to include all fire incident reports, making it less inclusive than a census and less statistically rigorous than a sample survey. It will not be possible to give a reliable estimate of the cost of this methodology as it will depend on how much each country would decide to include. The only appropriate assumption is that convenience sampling of fire data will be less costly than a full census but depending on size of the sample could be more expensive than a survey sample.

Conclusion

The cost was estimated for two of the three methodologies discussed while it is not possible to give an estimate of the convenience sample methodology. When calculating the cost for Census and Survey Sample methodologies the focus was on the cost of managing and operating the methodology at national level. While it's possible to provide estimate of cost of local level it is assumed that this will be more or less the same for the two methodologies as the results of previous tasks has shown that data are already collected in all countries and most likely at all fire departments.

The cost of the census methodology is by far the highest with an estimated total cost for all 27 EU member states of 12.7 million Euro. The cost of the survey sample methodology is considerably lower with a total cost for all 27 EU member states of 190,682 Euro. It should be kept in mind when comparing these numbers that a sample survey will need an up-front investment in research leading to appropriate survey sample being designed.

While every effort has been made to provide as good an estimate as possible it should be kept in mind that some of the basic assumptions applied can significantly skew the results. This includes the number of fires and time needed per incident. In particular, the latter variable has been based on experience available but have not been able to be confirmed by any quantitative studies.

8. Conclusion

In this report, we have reviewed critical issues involved in the design and implementation of fire incident data collection systems. The review proceeds from the assumption that fire incident data can serve a number of important purposes -- helping to reduce fires and losses, identifying opportunities for safety interventions and education programs, guiding the allocation of public resources to areas of greatest need and impact, and monitoring progress of safety initiatives. Data collection systems can also facilitate opportunities to share experiences and successes across regions and between countries, promoting a broader diffusion of technical and other innovations that increase fire safety. To achieve these objectives, it is important that data collection systems produce data that is reliable. We summarize some of the key factors related to data collection design and practice here.

The selection of data collection method should be determined by research needs and capabilities. In data collection with systematic intervention purposes, as is the case with fire incident data collection, it is important to create a sufficiently robust data base that can help identify risk factors and document fire incidence with reasonable confidence. Data collection systems that rely on voluntary reporting will almost certainly fall short of a complete census, while data collected by convenience sampling methods might have selective utility but would be insufficient to capture the broad range of fire incidents at the national level.

It appears that most countries currently employ a voluntary approach to data collection, with expectations that fire departments should participate in filing reports, but with mixed efforts by national programs to encourage and evaluate compliance. Whatever form the data collection system takes, it is important that it reliably capture the experiences of the populations it seeks to measure. To this end, data collection systems should be prepared to conduct follow up with non-respondents, assess the completeness of reporting, and identify any systematic patterns of non-reporting.

The potential for missing data is an issue that should be addressed in all phases of the research. We were able to find little discussion of missing data among the fire incident data collection systems in the European Union, as well as most fire incident data collection systems more generally. It may be the case that missing data receives the greatest attention in the United States because its data collection system is the most extensively detailed, with the greatest potential to produce items with unknown values, and potentially to discourage submission of reports altogether. Missing data may be less problematic in reporting systems that require less detail and whose population groups may be more uniform with respect to fire experiences. However, it is critical that effort be made to identify the extent of missing data and the patterns it takes if data on fire incidents is to be considered reliable.

The impact of missing data is likely to be especially problematic if it fails to account for differences in the populations that experience fire incidents. Such differences might include regional differences in the built environment, differences in neighborhood conditions, including housing quality and social conditions, or differences in age demographics. Assessment of missing data will accordingly be especially important in countries that are characterized by diverse regional levels of economic development and diversity of economic and social conditions.

On this point, it is important to note that the fire data collection systems examined in this research appear to be generally regarded as census systems of data collection. We cannot say if this is a view held by key users of fire data in these systems. However, there is a danger in assuming that data collection systems capture all or most fire incidents absent any examination of the degree and form of unreported fires or other missing data. Any systematic failure to collect data that is not randomly distributed runs the risk of failing to identify risk factors associated with social and economic disadvantage. Accordingly, it is important that the implementation of fire data collection systems include plans for data quality checks and procedures for handling missing data in order to verify the validity and reliability of data findings.

Financial costs will vary by country and be influenced by existing state of fire data collection practices and resources. It is important that there be some realistic appraisal of the economic costs of fire incident data collection if any harmonized system is to be sustainable over time. Countries and regions with stronger national traditions of data collection in support of policy objectives will require substantially less investment in supporting a harmonized fire incident data collection system than those in which data collection efforts are less mature or concentrated in specific areas. It is important to note here that substantial costs can be accrued in coming to decisions about what data to include and how to collect it, and this may be an unforeseen cost in seeking to achieve harmonized data collection in countries with decentralized and non-uniform systems, even if those systems are mature.

It is clear that the cost of implementing a comprehensive data collection system will be greatest in countries that have the least experience and fewest resources. Countries with less established or comprehensive data collection systems will assume significantly greater training costs in seeking to introduce data collection in fire departments nationwide. The cost burden will also be influenced by the availability and sophistication of computer hardware and software. Considering such differences, as well as relative differences in certain costs between Member States of the European Union, we have identified the core cost components of data collection as a starting point for assessments of financial commitment.

Additional observations

Our review of data collection methods and systems provides a foundation for several additional concluding observations relative to national systems of fire incident data collection.

- Data collection systems should be designed with sustainability in mind. Public funding for data collection systems can lag if they fail to generate recognition as a public good or commitment among key principals.
- Overly ambitious and detailed data collection systems may tax the patience of participants and undermine data quality. To encourage compliance and build competence and interest among participants, it may be useful for the architects of data collection systems to begin with comparatively modest reporting requirements and to introduce additional details incrementally as participants gain experience.
- Align data collection content with realistic policy goals and use data to promote safety interventions and practices.
- Use data to chart and publicize trends, demonstrate the utility of data collection, and build public recognition and support.

References

[1] Horan JM. Injury surveillance. Epidemiologic Reviews 2003; 25:24-42. Available at: https://doi.org/10.1093/epirev/mxg010. Accessed 2021-05-22.

[2] Maxim P, Plecas D, Garis L. Report on the feasibility of a Canadian National Fire Information Database. Centre for Public Safety and Criminal Justice Research, University of Fraser Valley 2012.

[3] National Association of State Fire Marshals. Fire Research and Education Foundation. Conquering the Unknowns 2014.

[4] Trivedi M. Unit 1: Introduction to sample surveys. Indira Gandhi National Open University; 2017.

[5] Dodge Y. (ed.) The Oxford dictionary of statistical terms. 6th ed. Oxford: Oxford University Press; 2006.

[6] Australian Bureau of Statistics. Samples and Censuses. Available at: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Basic+Survey+Design+-+Samples+and+Censuses. Accessed 2021-05-22.

[7] United States Bureau of Labor Statistics. Handbook of Methods: Census of Fatal Occupational Injuries. Available at: https://www.bls.gov/opub/hom/cfoi/pdf/cfoi.pdf. Accessed 2021-05-25.

[8] Bornstein MH, Jager J, Putnick DL. Sampling in developmental science: Situations shortcomings, solutions, and standards. Developmental Review 2013; 33:357-370.

[9] Moore DS, McCabe GP. Introduction to the Practice of Statistics, 5th ed. New York: Freeman; 2006

[10] Banerjee A, Chaudhury S. Statistics without tears: Populations and samples. Industrial Psychiatry Journal 2010; 19:60-65. doi: 10.4103/0972-6748.77642.

[11] Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and Policy in Mental Health and Mental Health Services Research 2015; 42:533-544.

[12] European Road Safety Observatory, Annual Accident Report 2018, 2018.

[13] Directorate-General for Mobility and Transport, Common Accident Data Set - Reference Guide, 2018.

[14] Levulyte L, Baranyai D, Török Á, Sokolovskij E. Bicycles' role in road accidents: A review of the literature. Transport and Telecommunication 2016; 17:122-127.

[15] Bauer R, Machata K, Brandstaetter C, Yannis G, Laiou A, Folla K, Road traffic accidents in European urban areas, Proceedings of the 1st European Road Infrastructure Congress, Leeds, October 2016.

[16] European Commission. Road safety: Data show improvements in 2018 but further concrete and swift actions are needed, Press Release 2019. Available at: https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1951. Accessed 5 July, 2021.

[17] Nikolaou P, Folla K, Dimitriou L, Yannis G. European countries' road safety evaluation by taking into account multiple classes of fatalities. Transportation Research Procedia 2021; 52:284-291.

[18] Dong Y, Peng CY. Principled missing data methods for researchers. Springer Plus 2013; 2:222. doi: 10.1186/2193-1801-2-222

[19] Newman DA. Missing data: Five practical guidelines. Organizational Research Methods 2014; 17:372-411. doi: 10.1177/1094428114548590.

[20 Langkamp DL, Lehman A, Lemeshow S. Techniques for handling missing data in secondary analyses of large surveys. Academic Pediatrics 2010; 10: 205-210. doi:10.1016/j.acap.2010.01.005.

[21] Schafer JL. Multiple imputation: a primer. Statistical Methods in Medicine 1999; 8:3-15. doi: 10.1191/096228099671525676.

[22] Bennett DA. How can I deal with missing data in my study? Australian and New Zealand Journal of Public Health. 2001; 25:464-469.

[23] Tabachnick BG, Fidell LS. Using multivariate statistics. 7th ed. New York: Pearson; 2019.

[24] Myers TA Goodbye, Listwise deletion: presenting hot deck imputation as an easy and effective tool for handling missing data. Communication Methods and Measures 2011; 5:297-310.

[25] Roth PL. Missing data: a conceptual review for applied psychologists. Personnel Psychology 1994; 47:537-560.

[26] Zhang Z. Missing data imputation: focusing on single imputation. Annals of Translational Medicine 2016; doi: 10.3978/j.issn.2305-5839.2015.12.38.

[27] Baraldi AN, Enders CK. An introduction to modern missing data analysis. Journal of School Psychology 2010; 48:5-37.

[28] Hall Jr. JR, Harwood B. The national estimates approach to U.S. fire statistics. Fire Technology 1989; 25:99-113.

[29] Ahrens M, Frazier P, Heeschen J. Use of fire incident data and statistics. In: AE Cote (Ed). Fire Protection Handbook, 19th Edition. Quincy MA: National Fire Protection Association, 2003.

[30] Kinsey, K. and Ahrens, M. NFIRS Incident Types: Why aren't they telling a clearer story? Quincy, MA: National Fire Protection Association, 2016.

[31] Butry, DT, Thomas, DS. Evaluating Potential Bias in Non-Randomly Reported Fire Incident Data. National Institute for Standards and Technology. Technical Note 1770, 2012.

[32] Home Office. Fire and rescue incident statistics. Methodology and quality report. 2020.

[33] Welsh Government, Quality Report for Welsh Fire Statistics, 2012.

[34] Civil Protection Act, Chapter 3. 2003.

[35] MSB. Kvalitet I MSB:s insatsstatistik 2020. April, 2021.

[36] Cambridge dictionary, "Uncertainty" https://dictionary.cambridge.org/dictionary/english/uncertainty. Accessed 2021-06-04.

[37] Guide to the Expression of Uncertainty in Measurement. International Organisation for Standardisation; Geneva, Switzerland, 1993. https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6

[38] JCGM International vocabulary of metrology (VIM) — Basic and general concepts and associated terms. JCGM 200:2012, Joint Committee for Guides in Metrology, 2012.

[39] Johansson, N. (2013) *Numerical experiments - a research method in fire science*, Department of Fire Safety Engineering and Systems Safety, Lund University 2013.

[40] Tillé Y. Théorie des sondages - Échantillonnage et estimation en populations finies, 2nd Edition, 2019.

[41] Ardilly P. Les techniques de sondage, 2nd Edition, 2006.

[42] Office of National Statistics, *Uncertainty and how we measure it for our surveys* https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/uncertaintyandhowwemeasureit. Accessed 2021-06-04

[43] Eurostat, *ESS Handbook for Quality and Metadata Reports*, https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-19-006. Accessed 2021-06-04

[44] Eurostat, *European Statistics Code of Practice*, https://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice. Accessed 2021-06-04

[45] MSB, *Dödsbränder i Sverige Kvalitetsgranskning av MSB:s dödsbrandsdatabas*, https://rib.msb.se/filer/pdf/26318.pdf

[46] Lasbeur L, Thélot B. Mortalité par accident de la vie courante en France métropolitaine, 2000-2010 Septembre 2014 – Congrès de l'Association des épidémiologistes de langue française (Adelf)- Epiter – Nice, France

[47] Lasbeur L, Rigou A, Thélot B. Application de la méthode capture-recapture aux victimes décédées par incendie après transfert en milieu hospitalier 2007. http://www.invs.sante.fr

[48] Belanger F, Ung A-B, et al. Analysis of injury related mortality in Europe, The ANAMORT project. Institut de veille sanitaire - Saint-Maurice, 2008.

[49] Carlotti P, Parisse D, Risler N. Statistical analysis of intervention reports for fires resulting in casualties deceased on the spot in Paris area. Fire Safety Journal 2017; 92:77–79. https://doi.org/10.1016/j.firesaf.2017.05.017.

[50] MSB, *Kvalitet i MSB:s insatsstatistik 2020*, https://ida.msb.se/dokument/insatsstatistik/kvalitet2020/Kvalitetsdeklaration2020.pdf visited 2021-06-04

[51] Glossary of Statistical terms, *Processing errors*, https://stats.oecd.org/glossary/detail.asp?ID=2138 visited 2021-06-04

[52] Runefors, M. Measuring the Capabilities of the Swedish Fire Service to Save Lives in Residential Fires, Fire Technology, 2020 56(2), 583-603. https://doi.org/10.1007/s10694-019-00892-y

[53] Johansson, N. *Dödsbränder i bostäder 2012-2015: En studie av fastighetsinformation*, Lund University 2017.

[54] Rein G, Torero-Cullen J, Jahn W, et al. Round-robin study of a priori modelling predictions of the Dalmarnock fire test one. Fire Safety Journal 2009 44(4), 590- 602.

[55] Johansson N, Anderson J, McNamee R., Pelo C. A Round Robin of fire modelling for performance-based design, Fire and Materials 2020. https://doi.org/10.1002/fam.2891

[56] Johansson N. *Round Robin av ett byggprocessärende: En studie av remissvar från räddningstjänsten*, Lund University 2019.

[57] Eurostat; Labour cost levels by NACE Rev. 2 activity [LC_LCI_LEV$DEFAULTVIEW]; Data extracted on 20/05/2021 21:21:05 from https://ec.europa.eu/eurostat/databrowser/product/page/LC_LCI_LEV

[58] Eurostat; Purchasing power parities (PPPs), price level indices and real expenditures for ESA 2010 aggregates; Data extracted on 20/05/2021 21:33:56 from https://ec.europa.eu/eurostat/databrowser/product/page/PRC_PPP_IND$DEFAULTVIEW

Annex 1. Comparative Assessment of Data Collection Methods

Table 1.1: Comparison of Data Collection Methods

	Census	Sample survey	Convenience collection
Definition	We are using the term 'census' to describe data collection methods where ALL cases for a specific jurisdiction (e.g., nation, city, region, etc.) are collected.	We are using the term 'sample survey' to describe data collection methods where a statistically-designed sample of data is collected and then analyzed to produce an estimate of the total for some category of cases (e.g., fires, deaths, financial loss, etc.) for a specific jurisdiction (e.g. nation, city, region, etc.).	We are using the term 'convenience collection' to describe data collection that involves identifying cases in a manner that is not structured in a way that allows a statistical methodology to calculate national totals for any category of cases. This would include, for example, collection of incidents from media reports as well as collection of easily-obtained data available in non-statistical sampling.
Output	This method, if successful, will result in a count of total number cases (e.g., all fires, all deaths, all injuries, etc.). It will also provide a full database of all incidents in addition to total case numbers.	A properly designed sample will result in an estimate (usually within error bounds and to a certain statistical significance) for the total number of cases (e.g., estimated total fires, estimated total deaths, estimated injuries, etc.).	This method will NOT result in a total (estimated or not) for the number of cases. Descriptive statistics for the collected data can be reported, but those results cannot be considered representative for the entire population. The full database of what was collected will be available.
Limitations/challenges	* Total coverage of the population is required for a true census. * Can be costly.	* Proper statistical sample design is required. * The sample must be representative of the population for whom the estimate will apply (e.g., cannot sample within urban areas and expect the estimate to apply to rural communities). * Inclusion or omission of serious incidents can skew the loss measures. * Does not include full database of incidents.	* Completeness of the dataset cannot be determined. * The collected data cannot be considered representative of the whole. * May unintentionally exclude large parts of population of interest * National estimate for totals cannot be calculated from a convenience collection of reports.

Table 1.1: Comparison of Data Collection Methods (Continued)

	Census	Sample survey	Convenience collection
Advantages	* A true census will capture all cases, though details for each case might be incomplete.	* Representative. * Can be easier and less costly than a census.	* 'Convenient'. * Can be least expensive option.
Can results be compared to results using other methods?	Totals calculated with this method can be compared to totals estimated through the use of a sample survey.	Totals calculated with this method can be compared to totals reported from a census.	This method does not calculate totals that can be compared to results from the other methods.

Who collects?	Census	Sample survey	Convenience collection
Firefighters	Incident data preferably using a Common Fire Incident Data Set.	Incident data, preferably using a common fire incident data set.	Incident data, preferably using a common fire incident data set.
Police	Investigation reports	Investigation reports.	Investigation reports.
Hospitals/Medical Examiner	Death certificates and injury records.	Death certificates and injury records.	Death certificates and injury records.
Insurance	Investigation reports.	Investigation reports.	Investigation reports.

Local data management	Census	Sample survey	Convenience collection
Who?	Fire Department	Fire Department	Fire Department
Missing Data	Use reports from police and insurance and death and injury data from hospitals. Building departments , tax offices, or other agencies may have building data. Ideally, it would be good to link to such sources.	Use reports from police and insurance and death and injury data from hospitals.	Use reports from police and insurance and death and injury data from hospitals.
Unknowns	Always allow for unknowns.	Always allow for unknowns.	Always allow for unknowns.
Formatting	Ensure data are in required National Format or can be readily converted to the national format.	Ensure data are in required national format.	Ensure data are in required national format or can be readily converted to the national format.

Table 1.1: Comparison of Data Collection Methods (Continued)

Reporting of local data to national level	Census	Sample survey	Convenience collection
Full Database	Full database is forwarded to national fire data collection agency, although data collected by local option or only of relevance at the local level (e.g., street address, victim name, fire department response details, etc.) may not be required at the national level.	The agency filling out the survey can also forward the underlying database if required. This will lead to a database at national level that includes only the detailed data from the agencies responding to the survey.	Full database is forwarded to national fire data collection agency, although data collected by local option or only of relevance at the local level (e.g., street address, victim name, fire department response details, etc.) may not be required at the national level.
Totals	Totals calculated directly from data.	Based on sample design. Local entities will report total numbers as a response to a survey. Only total numbers reported for each variable and only from those surveyed.	Only summation of data provided.

National data management	Census	Sample survey	Convenience collection
Who?	One central agency needed to handle the national level and to be responsible for reporting data to the EU. Preferably National Statistical entity or alternatively the ministry responsible for the fire service.	One central agency needed to handle the national level and to be responsible for reporting data to the EU. Preferably National Statistical entity or alternatively the ministry responsible for the fire service.	One central agency needed to handle the national level and to be responsible for reporting data to the EU. Preferably National Statistical entity or alternatively the ministry responsible for the fire service.
Unknowns	Presented as unknowns and not distributed among the other data fields for the variable.	As part of the national estimate methodology, distribution of unknowns can be considered. This will require additional statistical evaluation to ensure that a large number of unknowns don't skew the total estimate.	Presented as Unknowns and not distributed among the other data fields for the variable.

Table 1.1: Comparison of Data Collection Methods (Continued)

National data management	Census	Sample survey	Convenience collection
National Totals	Calculated directly from final dataset.	National total estimates calculated using defined methodology. This is closely linked to the sample design for the survey.	No national totals possible. Only summation of data collected and if possible estimate of size of sample.
Transformation from National dataset to Harmonised Dataset	If the national dataset does not correspond directly to the harmonised fire incident data set transformation rules are needed. The transformation rules needed depend on the national variables collected, definitions used for each of those and how those corresponds to the harmonised variables.	If the underlying dataset does not correspond directly to the Harmonised Fire Incident Data Set transformation rules are needed as described for the census method. National totals can then be calculated for each of the EU variables.	Same as for census method.
Formatting	Database formatted to comply with harmonised fire incident data set	Totals reported for each of the variables in the harmonised incident data set.	Database formatted to comply with harmonised fire incident data set

Reporting of national data to EU	Census	Sample survey	Convenience collection
Full Database	Full database provided.	The database from the sampled agencies can be reported with the understanding that this is only a sampled survey and statistical rules are to be used to calculate national totals for each variable.	Full database provided.
Totals	Total numbers provided for each variable for quality control purposes at EU level.	Total numbers for each variable reported to EU.	Only summation of data provided.

Table 1.2: Variables to be Collected Tier 1 & 2

Category of Interest	Variable	Data Fields	Example values (for illustration)
Intervention Characteristics	Incident date		
		Incident date	dd/mm/yyyy
Intervention Characteristics	Incident time		
		Incident time	hh/mm
Intervention Characteristics	Incident location		
		Incident location	City or postal code
Human Characteristics	Number of deaths		
		Number of deaths	Numeric value
		Or, could depend on definition of deaths - if some causes of injury are not counted and ff/civ	
		Cause of fatal injury	*Exposure to smoke, exposure to heat/flame, fall, struck in collapse, etc.*
		Role	*Civilian, firefighter, etc.*
Human Characteristics	Number of injuries		
		Number of injuries	Numeric value or age group
		Or, could depend on definition of injuries - if some causes of injury are not counted and ff/civ	
		Cause of injury	*Exposure to smoke, exposure to heat/flame, fall, struck in collapse, etc.*
		Role	*Civilian, firefighter, etc.*
Human Characteristics	Age of victim		
		Age of victim	Numeric value
		Age categories	*e.g., infant, child, youth, adult, elderly.*
Building Characteristics	Type of building		
		Type of building	House, apartment, hospital, office building, factory, etc.
Building Characteristics	Number of floors		
		Number of floors	Numeric value
Fire Characteristics	Fire Cause		

Table 1.2: Variables to be Collected Tier 1 & 2 (Continued)

Category of Interest	Variable	Data Fields	Example values (for illustration)
		Cause of Ignition	Simple - intentional, unintentional, failure of equipment or heat source, under investigation, not determined.
		Cause Hierarchy:	*Detailed - intentional, child play, smoking, heating, cooking, electrical distribution, appliances, open flame, other heat, other equipment, natural, exposure, unknown.*
		Ignition Factor	*More detailed than 'Cause of Ignition' - act or omission that explains why the heat and material were able to combine to start the fire.*
		Source of Ignition	*Lighter, match, heat from cooking equipment, fireplace, portable heater, properly operating equipment, open flame from candle, etc.*
		Equipment involved	*Heating system, cooking equipment, A/C or refrigeration, electrical distribution eqpt, appliances, special eqpt, processing eqpt, service/maintenance eqpt., etc.*
Fire Characteristics	Room of origin		
		Room of origin	Lounge/living room, bedroom, kitchen, closet, machine room, operating room, etc.
Fire Characteristics	Source of Ignition		
		Source of Ignition	Lighter, match, heat from cooking equipment, fireplace, portable heater, properly operating equipment, open flame from candle, etc.
Fire Characteristics	Material mainly resp.		
		Material mainly resp	Gas, flammable/combustible liquid, volatile solid/chemical, plastic, natural product, wood/paper, fabric/textile, material compounded with oil, other.

Table 1.3: Variables to be Collected Tier 3 (>70% stakeholders)

Building Characteristics	Construction type (e.g. reinforced concrete, steel).		
		Type of Construction	Fire resistive, heavy timber, non-/limited combustible (protected/unprotected), ordinary (protected/unprotected), wood-frame (protected/unprotected).
Building Characteristics	Fire safety measures present (e.g. alarm system, compartmentation)		
		Either check boxes for each type of safety measure (alarm system y/n, sprinkler system y/n, smoke control system y/n, compartmentation y/n, etc.) or a separate set of fields to detail each system.	
Consequences	Effectiveness of fire safety measures in reducing the fire		
		This would need an entry for each type of safety measure.	
		Effect/performance	Worked as designed, device failed, blocked/impaired, manual but not operated, ineffective due to interference, not proper for hazard, insufficient number or size, structurally failed, installation not complete, other/undetermined.

Table 1.4: Overview of Current Fire Data Collection Systems

Country	Who Collects the data?	What entity processes the data?	What entity reports the data?
France	Firefighters	Ministry of Interior	
Spain	Firefighters	MAPFRE (insurance) and APTB (professional association of fire services) collect data on fatal fires only from the fire services	MAPFRE and APTB publish Yearly reports
Hungary	Firefigthers collect Basic data. County Directorate for Disaster Management collects fire causes	The National Inspectorate General for Fire Services at the National Directorate General for Disaster Management (NDGDM) annually issues the requested data for the Central Statistics Bureau.	County Directorate for Disaster Management
Switzerland	Firefighters provide intervention data. Building damage statistics and building fire fatalitites from APIRE (Association of Public Insurance Companies for Real Estate) or Vereinigung kantonaler Gebäudeversicherungen (VKG).	The insurance association (APIRE)	The insurance association (APIRE)
Austria	Austrian Fire Protection Association. SIZ (Sicherheitsinformationszentrum). BVS - Brandverhütungsstelle für Oberösterreich for property loss. Landesstelle Steiermark.	In general, the Austrian Fire Prevention Associations collect data from the police stations and insurers for each federal state (excepting Vienna) and publish them yearly. The several fire statistics are gathered by the Upper Austrian Fire Prevention Association for creating and publishing an Austrian Fire Statistic, which is also published once a year.	Austrian Fire Protection Association
Poland	Firefighters	Fire brigades for building fires.	The State Fire Service (SFS) of Poland, supervised by the Minister of the Interior and Administration

Table 1.4: Overview of Current Fire Data Collection Systems (Continued)

Country	Who Collects the data?	What entity processes the data?	What entity reports the data?
Germany	Firefighters	Ministry of the interior in the regions.	Ministry of the interior in the regions.
Luxembourg	Firefighters for incident data. Police for victim data.	The Luxembourgish Fire and Rescue Corps process the data.	Only internal reports
Russia	Firefighters	State Fire service	State Fire Service
Czech Republic	Firefighter	Fire Service	Fire Service
Greece	Firefighters	The Hellenic Fire Corps	The Hellenic Fire Corps
Bulgaria	Firefighters	Fire service	Fire service
Latvia	Firefighters	Fire service	Fire service
Croatia	Firefighters	Fire service	Fire service
Lithuania	Firefighters	Fire service	Fire service
UK	Firefighters	England - Home Office. Scotland - Scottish Fire and Rescue servies data team. Wales - Welsh Government	England - Home Office. Scotland - Scottish Fire and Rescue servies data team. Wales - Welsh Government
Ireland	Firefighters	The National Directorate for Fire and Emergency Management (NDFEM) contact each authority to confirm that the information is correct.	Department of Housing, Local Government and Heritage produces statistics about fire and other emergency calls dealt with by local authority fire brigades during that year.
Australia	Firefighters (officer of first appliance)	Fire Service	The Fire and Rescue Organization of each State yearly publishes some information on the dataset.
Italy	Firefighters (Crew Commander)	The data are collected and elaborated at the national, regional and provincial level by: - Central Statistic Service at the Cabinet of the Head of the CNVVF (Bureau of Direct Collaboration of the Head of the C.N.VV.F); - Regional Statistic Services inside the Regional Directions of the C.N.VV.F; - Statistical Services inside the Provincial Fire Departments.	The Central Statistic Service

Table 1.4: Overview of Current Fire Data Collection Systems (Continued)

Country	Who Collects the data?	What entity processes the data?	What entity reports the data?
Sweden	Firefighters	Swedish Civil Contingencies Agency (MSB)	MSB
Denmark	Firefighters	Danish Emergency Management Agency (DEMA)	DEMA
Norway	Firefighters	The Norwegian Directorate for Civil Protection (DSB)	DSB
Slovakia	Fire investigators under the fire department	District Head-Offices of the Fire and Rescue Corps via fire investigators process data on fires that have occurred in their territory. At the beginning of each month the above-mentioned data are sent to the Fire Research Institute of the Ministry of Interior of the Slovak Republic, where they are further processed and analyzed according to various indicators.	Fire & Rescue Corps
US	Firefighters	USFA and NFPA	USFA and NFPA
Canada	Firefighters	Varies by province.	Provincial fire marshals and commissioners

Annex 2. Data used for calculating cost of data collection methodologies.

On average, 1.5 million fires are registered in the EU every year. The numbers fluctuate significantly among the Member States: Malta with around 1700 and France with around 316,100 fires.

Table 2-1: Number of fires in the EU Member States (mean values for 2014-2020)

No.	Member state	Population, 1000. inh.	Number of fires	Fires per 1000 inhabitants
1	Austria	8.859	43.370	4,9
2	Belgium	10.667	29.622	2,8
3	Bulgaria	7.050	42.141	6,0
4	Croatia	4.058	14.980	3,7
5	Cyprus	858	5.400	6,3
6	Czech Republic	10.650	18.270	1,7
7	Denmark	5.786	15.081	2,6
8	Estonia	1.329	4.675	3,5
9	Finland	5.474	14.264	2,6
10	France	66.628	316.100	4,7
11	Germany	83.020	203.419	2,5
12	Greece	10.788	27.784	2,6
13	Hungary	9.772	20.913	2,1
14	Ireland	4.581	20.756	4,5
15	Italy	61.000	234.675	3,8
16	Latvia	1.920	10.095	5,3
17	Lithuania	2.794	13.163	4,7
18	Luxemburg	602	2.228	3,7
19	Malta	420	1.749	4,2
20	Netherlands	17.282	38.900	2,3
21	Poland	38.411	153.520	4,0
22	Portugal	11.000	42.398	3,9
23	Romania	20.121	27.804	1,4
24	Slovakia	5.458	9.602	1,8
25	Slovenia	2.095	4.427	2,1
26	Spain	46.157	118.892	2,6
27	Sweden	10.328	26.445	2,6
	Total	447.108	1.546.973	3,5

Source: CTIF (www.ctif.org): World Fire Statistics, Report 2016 (No. 21), Report 2021 (No. 26).

Approximately 447 million people live in the 27 Member States. The population per country fluctuates between 420,000 inhabitants (Malta) and 83 million inhabitants (Germany). Figure 2-1 illustrates the population figures in the EU member states (mean values for 2014-2020).

Figure 2-1. Population of EU Member States, 1,000 Inhabitants
(Mean values for 2014-2020)

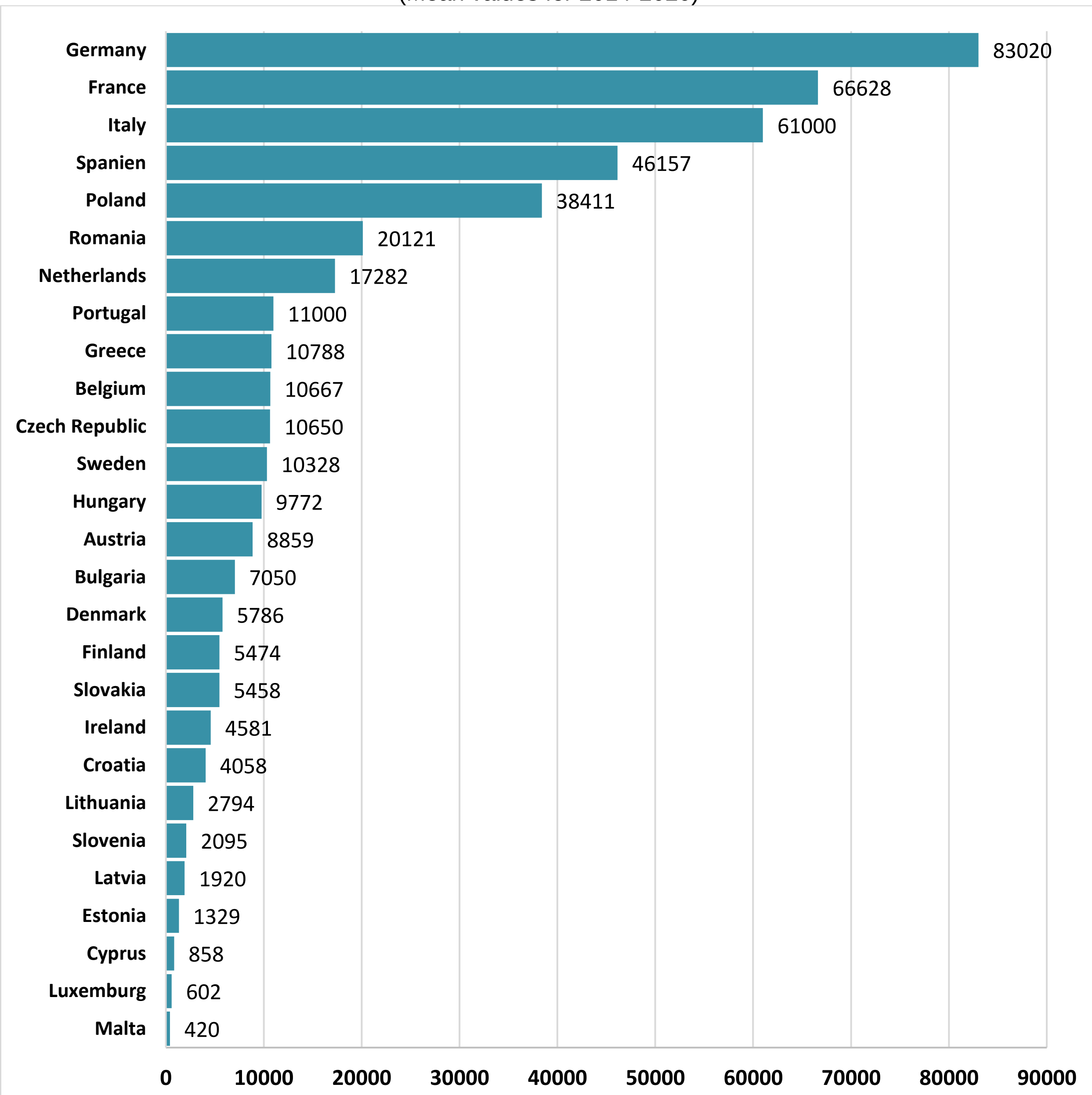


Figure 2-2 shows the number of fires in the EU Member States (mean values for 2014-2020). Figure 2-3 shows the distribution of the number of fires per 1000 inhabitants in the EU member states (mean values for 2014-2020).

Figure 2-2: Number of fires in the EU Member States (mean values for 2014-2020)

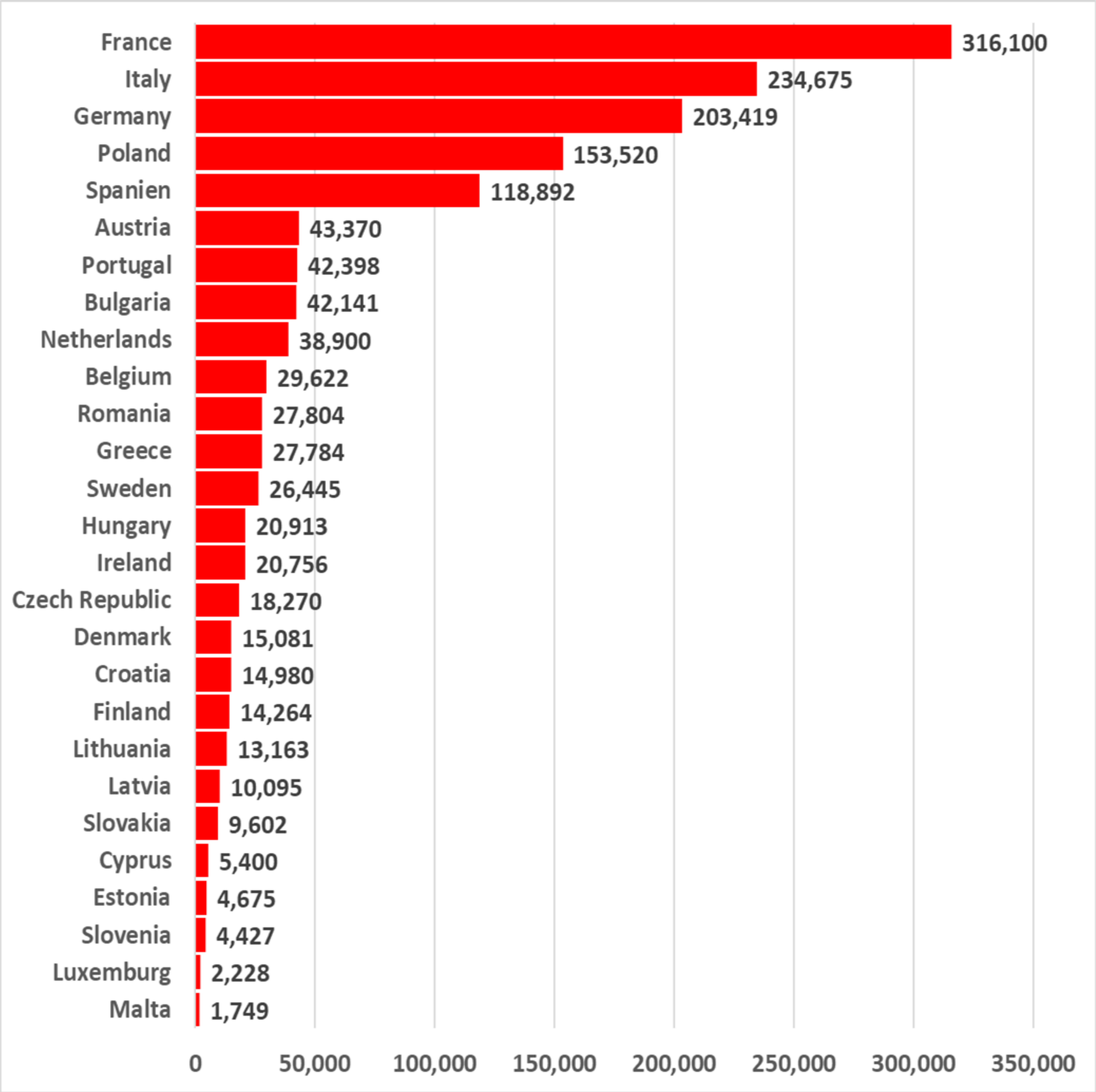
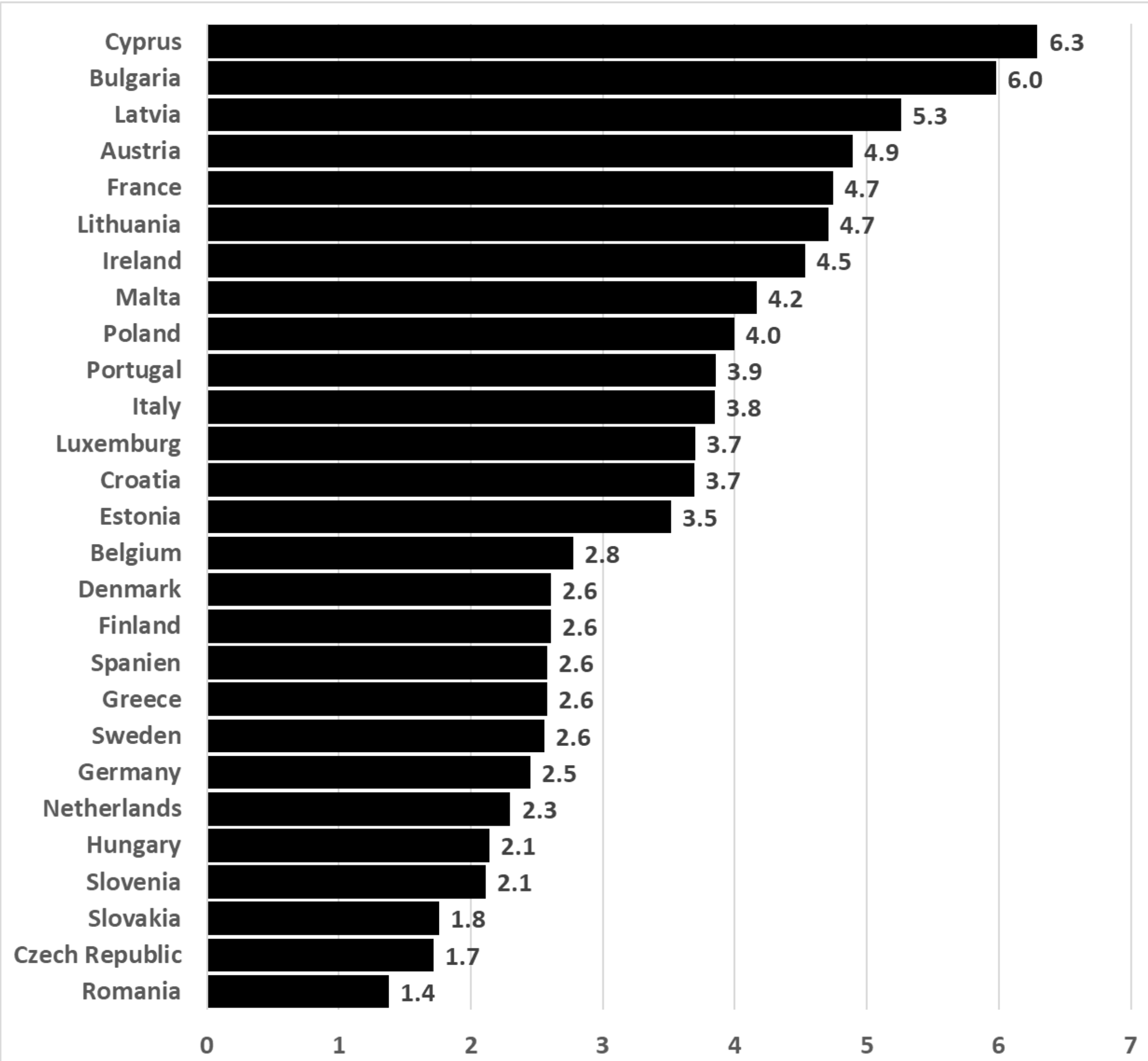


Figure 2-3 shows the distribution of the number of fires per 1000 inhabitants in the EU Member States (mean values for 2014-2020).

Figure 2-3: Number of fires per 1000 inhabitants in the EU Member States (Mean values for 2014-2020)



The numerical values in EU Member States for average annual hours worked[1], labour costs[2], and price level indices[3] are presented in Table 2-2 and figures 2-4, 2-5 & 2-6.

Table 2-2: Economic indicators in the EU Member States (mean values for 2014-2020).

No.	Member state	Average annual hours actually worked per worker (2020) in €	Labour cost level (2020) in €	Price level indices (2019), EU average is 100
1	Austria	1400,00	36,70	118,50
2	Belgium	1481,00	41,10	117,10
3	Bulgaria	1600,00	6,50	49,10
4	Croatia	1500,00	10,80	66,90
5	Cyprus	1700,00	17,00	94,10
6	Czech Republic	1705,00	14,10	69,60
7	Denmark	1346,00	45,80	141,50
8	Estonia	1654,00	13,60	82,30
9	Finland	1531,00	34,30	126,40
10	France	1402,00	37,50	110,00
11	Germany	1331,70	36,60	107,80
12	Greece	1728,00	16,90	83,70
13	Hungary	1660,30	9,90	63,30
14	Ireland	1746,00	32,30	137,20
15	Italy	1558,70	29,80	102,90
16	Latvia	1577,00	10,50	74,20
17	Lithuania	1595,00	10,10	64,90
18	Luxembourg	1427,00	42,10	147,00
19	Malta	1500,00	14,50	87,20
20	Netherlands	1399,00	36,80	119,70
21	Poland	1766,00	11,00	57,30
22	Portugal	1613,00	15,70	85,70
23	Romania	1600,00	8,10	50,50
24	Slovakia	1572,00	13,40	79,20
25	Slovenia	1514,60	19,90	86,50
26	Spain	1577,20	22,80	96,90
27	Sweden	1424,00	37,30	131,40

[1]Data extracted on 24 Jun 2021 06:37 UTC (GMT) from OECD.Stat, Average annual hours actually worked per worker.
[2] Data extracted on 20/05/2021 21:21:05 from [ESTAT], Labour costs structure (2020)[3]Price level indicies EU from EuroStat (2020)

For the average annual hours worked per worker (2020), the figures vary between 1331 hours in Germany and 1766 hours in Poland (Figure 2-4).

Figure 2-4: Average annual hours worked per worker in the EU Member States (2020).

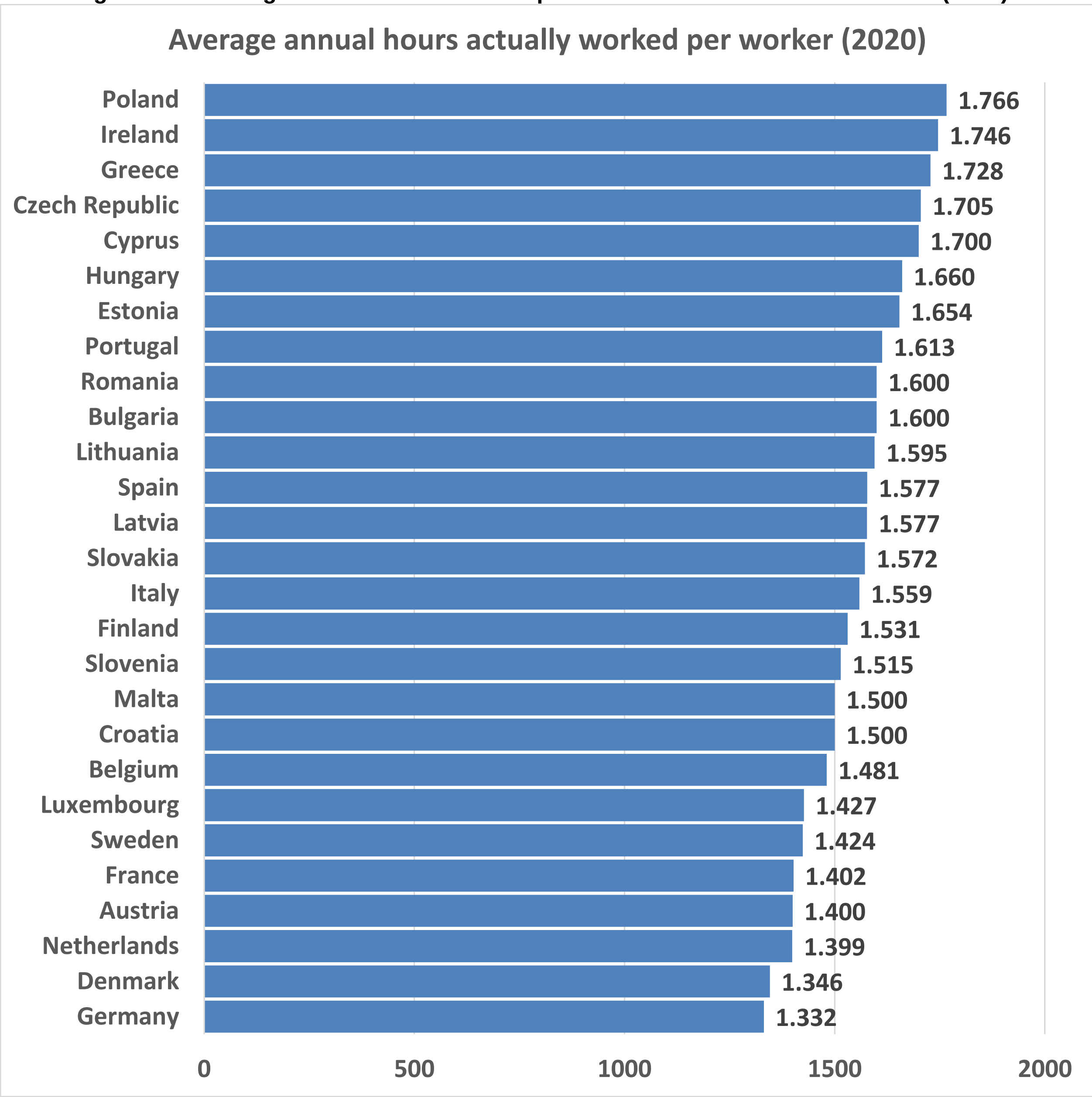


Figure 2-5: Labour cost level in the EU Member States (2020

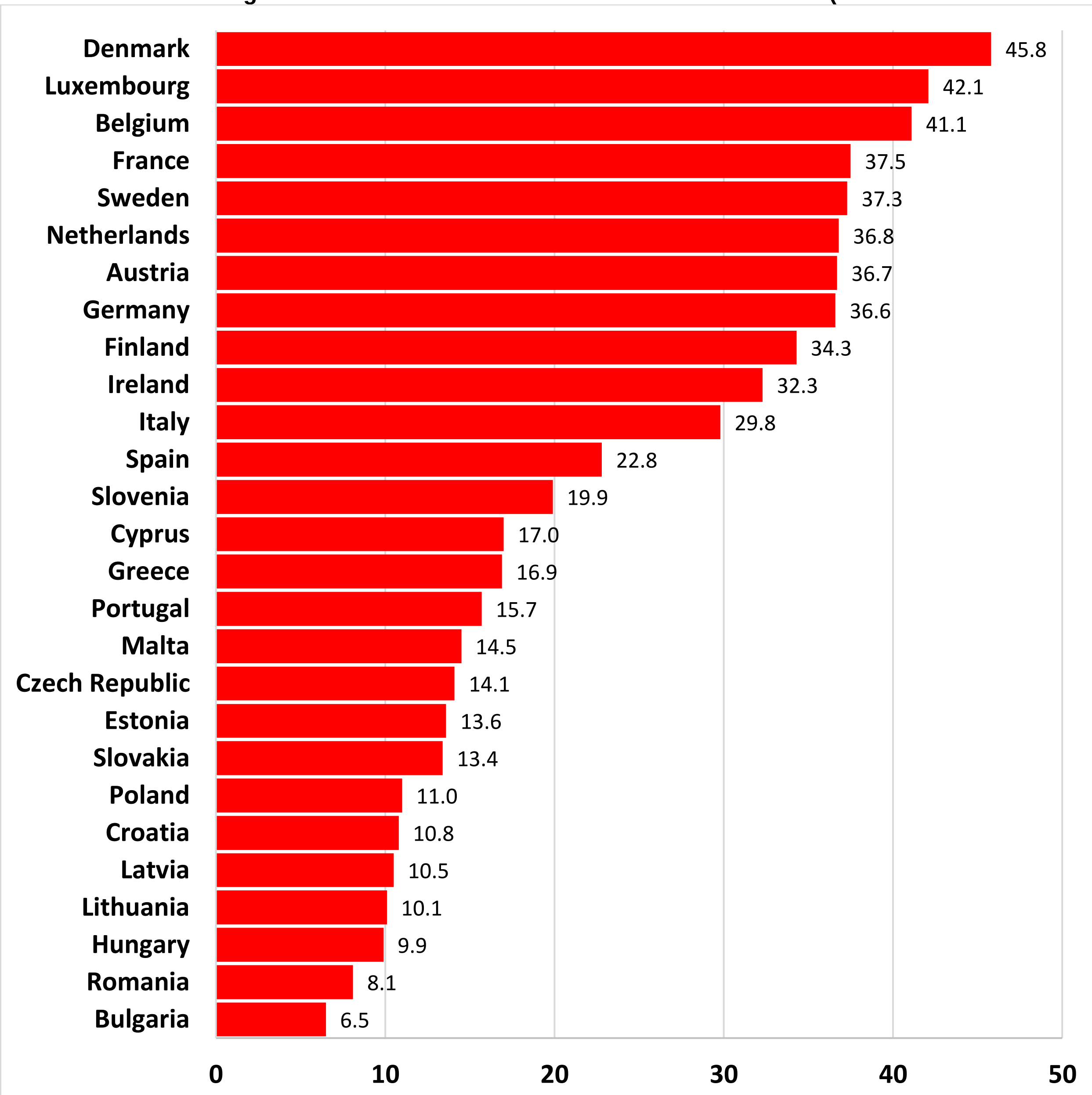
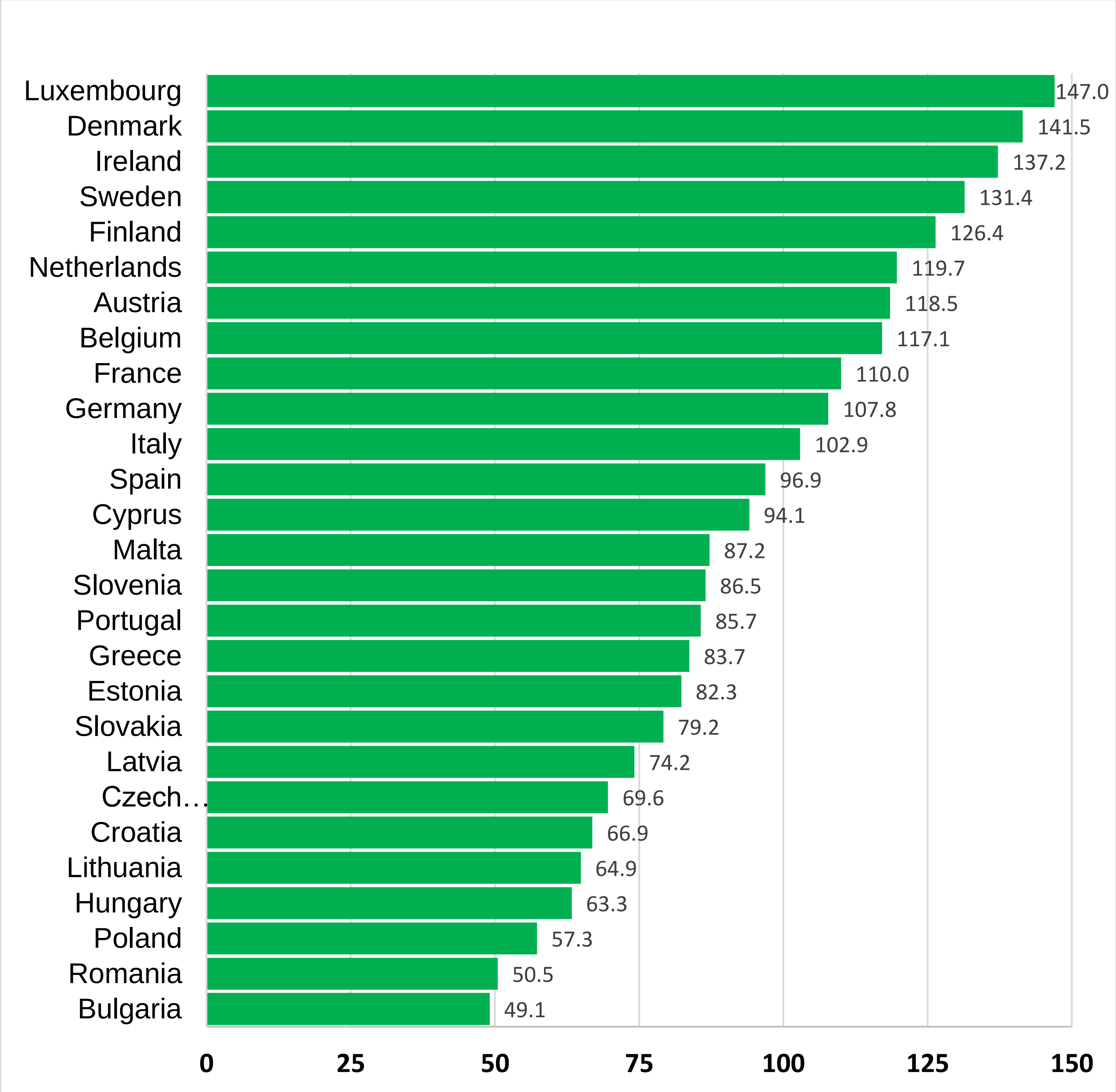


Figure 2-6: Price level indices in the EU Member States (2020)



Annex 3. Breakdown of costs for individual EU Member States.

This Annex lists the calculation of the national fire statistics for each of the 27 EU Member States (Tables 3-1 to 3-27).

Table 3-1: Calculation of the national costs of EU fire statistics for Austria.

		Area	Population density		Capital
AUSTRIA		83855	106		Vienna
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		8858775	45550	46	184

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1400,00
D	Labour cost (€ per hour) [LCH]	36,70
E	Price level indicies EU [PLI]	118,50
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	11387,50
H	Total Labour costs (€),[TLC]	417921,25
K	Number of Staff [NSTF]	8

L	Cost category	€
M	Personnel costs (80%)	417921,25
N	Hardware (5%)	26120,08
O	Software (10%)	52240,16
P	Other costs, including training, refresher couces, etc. (5%)	26120,08
Q	Total, €	522401,56

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Austria, the financing of the national EU fire statistics should be supported with 520000 € and 8 full-time employees.

Table 3-2: Calculation of the national costs of EU fire statistics for Belgium.

BELGIUM		Area	Population density		Capital
		30528	376		Brussels
		sq.km	inh./sq.km		
No.	**Parameter fo Calculation**	**Basic Numbers for Calculation...**			
		Population (2019)	**Fires (Average 2014-2018)**	**Fire Deaths (Average 2014-2018)**	**Fire Injuries (Average 2014-2018)**
A		**11467923**	**21000**	**60**	**240**

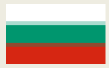
B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1481,00
D	Labour cost (€ per hour) [LCH]	41,10
E	Price level indicies EU [PLI]	117,10
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	5250,00
H	Total Labour costs (€),[TLC]	215775,00
K	Number of Staff [NSTF]	4

L	Cost category	€
M	Personnel costs (80%)	215775,00
N	Hardware (5%)	13485,94
O	Software (10%)	26971,88
P	Other costs, including training, refresher couces, etc. (5%)	13485,94
Q	Total, €	269718,75

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Belgium, the financing of the national EU fire statistics should be supported with 269000 € and 4 full-time employees.

Table 3-3: Calculation of the national costs of EU fire statistics for Bulgaria.

BULGARIA 		Area	Population density		Capital
		110994	63		Sofia
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		7000039	29448	126	504

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1600,00
D	Labour cost (€ per hour) [LCH]	6,5
E	Price level indicies EU [PLI]	49,1
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	7362,00
H	Total Labour costs (€),[TLC]	47853,00
K	Number of Staff [NSTF]	5

L	Cost category	€
M	Personnel costs (80%)	47853,00
N	Hardware (5%)	2990,81
O	Software (10%)	5981,63
P	Other costs, including training, refresher couces, etc. (5%)	2990,81
Q	Total, €	59816,25

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Bulgaria, the financing of the national EU fire statistics should be supported with 60000 € and 5 full-time employees.

Table 3-4: Calculation of the national costs of EU fire statistics for Croatia.

No.	Parameter fo Calculation				
CROATIA 		Area	Population density		Capital
		56594	72		Zagreb
		sq.km	inh./sq.km		
		Basic Numbers for Calculation...			
No.	Parameter fo Calculation	Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		4076246	11018	24	96
B	Basic data for cost estimations				
C	Average annual hours worked per worker [AHW]				1500,00
D	Labour cost (€ per hour) [LCH]				10,80
E	Price level indicies EU [PLI]				66,90
F	Average hours worked per fire event [AHFE]				0,25
G	Total hours for performance of yearly fire statistics [THP]				2754,50
H	Total Labour costs (€),[TLC]				29748,60
K	Number of Staff [NSTF]				2
L	Cost category				€
M	Personnel costs (80%)				29748,60
N	Hardware (5%)				1859,29
O	Software (10%)				3718,58
P	Other costs, including training, refresher couces, etc. (5%)				1859,29
Q	Total, €				37185,75

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Croatia, the financing of the national EU fire statistics should be supported with 37000 € and 2 full-time employees.

Table 3-5: Calculation of the national costs of EU fire statistics for Cyprus.

CYPRUS		Area	Population density		Capital
		9251	95		Nicosia
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		875898	7000	3	12
B	Basic data for cost estimations				
C	Average annual hours worked per worker [AHW]				1700,00
D	Labour cost (€ per hour) [LCH]				17,00
E	Price level indicies EU [PLI]				94,10
F	Average hours worked per fire event [AHFE]				0,25
G	Total hours for performance of yearly fire statistics [THP]				1750,00
H	Total Labour costs (€),[TLC]				29750,00
K	Number of Staff [NSTF]				1
L	Cost category				€
M	Personnel costs (80%)				29750,00
N	Hardware (5%)				1859,38
O	Software (10%)				3718,75
P	Other costs, including training, refresher couces, etc. (5%)				1859,38
Q	Total, €				37187,50

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Cyprus, the financing of the national EU fire statistics should be supported with 37000 € and 1 full-time employee.

Table 3-6: Calculation of the national costs of EU fire statistics for the Czech Republic.

CZECH REPUBLIC 		Area	Population density		Capital
		78866	135		Prague
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		10649800	18270	109	436

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1705,00
D	Labour cost (€ per hour) [LCH]	14,10
E	Price level indicies EU [PLI]	69,6
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	4567,50
H	Total Labour costs (€),[TLC]	64401,75
K	Number of Staff [NSTF]	3

L	Cost category	€
M	Personnel costs (80%)	64401,75
N	Hardware (5%)	4025,11
O	Software (10%)	8050,22
P	Other costs, including training, refresher couces, etc. (5%)	4025,11
Q	Total, €	80502,19

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Czech Republic, the financing of the national EU fire statistics should be supported with 80000 € and 3 full-time employees.

Table 3-7: Calculation of the national costs of EU fire statistics for Denmark.

DENMARK		Area	Population density		Capital
		43075	135		Copenhagen
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation…			
		Population (2019)	Fires 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		5806081	13661	67	268

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1346,00
D	Labour cost (€ per hour) [LCH]	45,80
E	Price level indicies EU [PLI]	141,50
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	3415,25
H	Total Labour costs (€),[TLC]	156418,45
K	Number of Staff [NSTF]	3

L	Cost category	€
M	Personnel costs (80%)	156418,45
N	Hardware (5%)	9776,15
O	Software (10%)	19552,31
P	Other costs, including training, refresher couces, etc. (5%)	9776,15
Q	Total, €	195523,06

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Denmark, the financing of the national EU fire statistics should be supported with 195000 € and 3 full-time employees.

Table 3-8: Calculation of the national costs of EU fire statistics for Estonia.

ESTONIA		Area	Population density		Capital
		45227	29		Tallinn
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		1324820	5517	46	184

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1654,00
D	Labour cost (€ per hour) [LCH]	13,60
E	Price level indicies EU [PLI]	82,30
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	1379,25
H	Total Labour costs (€),[TLC]	18757,80
K	Number of Staff [NSTF]	1

L	Cost category	€
M	Personnel costs (80%)	18757,80
N	Hardware (5%)	1172,36
O	Software (10%)	2344,73
P	Other costs, including training, refresher couces, etc. (5%)	1172,36
Q	Total, €	23447,25

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Estonia, the financing of the national EU fire statistics should be supported with 23000 € and 1 full-time employee.

Table 3-9: Calculation of the national costs of EU fire statistics for Finland.

FINLAND		Area	Population density		Capital
		338424	16		Helsinki
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		5517919	12685	72	288

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1531,00
D	Labour cost (€ per hour) [LCH]	34,30
E	Price level indicies EU [PLI]	126,40
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	3171,25
H	Total Labour costs (€),[TLC]	108773,88
K	Number of Staff [NSTF]	2

L	Cost category	€
M	Personnel costs (80%)	108773,88
N	Hardware (5%)	6798,37
O	Software (10%)	13596,73
P	Other costs, including training, refresher couces, etc. (5%)	6798,37
Q	Total, €	135967,34

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Finland, the financing of the national EU fire statistics should be supported with 135000 € and 2 full-time employees.

Table 3-10: Calculation of the national costs of EU fire statistics for France.

FRANCE		Area	Population density		Capital
		640679	105		Paris
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		67028048	293873	289	1156

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1402,00
D	Labour cost (€ per hour) [LCH]	37,50
E	Price level indicies EU [PLI]	110,00
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	73468,25
H	Total Labour costs (€),[TLC]	2755059,38
K	Number of Staff [NSTF]	52

L	Cost category	€
M	Personnel costs (80%)	2755059,38
N	Hardware (5%)	172191,21
O	Software (10%)	344382,42
P	Other costs, including training, refresher couces, etc. (5%)	172191,21
Q	Total, €	3443824,22

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state France, the financing of the national EU fire statistics should be supported with 3440000 € and 52 full-time employees.

Table 3-11: Calculation of the national costs of EU fire statistics for Germany.

GERMANY		Area	Population density		Capital
		357021	233		Berlin
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		83019214	199602	362	1448

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1331,70
D	Labour cost (€ per hour) [LCH]	36,60
E	Price level indicies EU [PLI]	107,80
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	49900,50
H	Total Labour costs (€),[TLC]	1826358,30
K	Number of Staff [NSTF]	37

L	Cost category	€
M	Personnel costs (80%)	1826358,30
N	Hardware (5%)	114147,39
O	Software (10%)	228294,79
P	Other costs, including training, refresher couces, etc. (5%)	114147,39
Q	Total, €	2282947,88

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Germany, the financing of the national EU fire statistics should be supported with 2282000 € and 37 full-time employees.

Table 3-12: Calculation of the national costs of EU fire statistics for Greece.

GREECE		Area	Population density		Capital
		131990	81		Athens
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		10722287	25363	44	176

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1728,00
D	Labour cost (€ per hour) [LCH]	16,90
E	Price level indicies EU [PLI]	83,70
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	6340,75
H	Total Labour costs (€),[TLC]	107158,68
K	Number of Staff [NSTF]	4

L	Cost category	€
M	Personnel costs (80%)	107158,68
N	Hardware (5%)	6697,42
O	Software (10%)	13394,83
P	Other costs, including training, refresher couces, etc. (5%)	6697,42
Q	Total, €	133948,34

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Greece, the financing of the national EU fire statistics should be supported with 133000 € and 4 full-time employees.

Table 3-13: Calculation of the national costs of EU fire statistics for Hungary.

HUNGARY		Area	Population density		Capital
		93030	105		Budapest
		sq.km	inh./sq.km		
No.	**Parameter fo Calculation**	**Basic Numbers for Calculation...**			
		Population (2019)	**Fires (Average 2014-2018)**	**Fire Deaths (Average 2014-2018)**	**Fire Injuries (Average 2014-2018)**
A		**9797561**	**20557**	**109**	**436**

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1660,30
D	Labour cost (€ per hour) [LCH]	9,90
E	Price level indicies EU [PLI]	63,30
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	5139,25
H	Total Labour costs (€),[TLC]	50878,58
K	Number of Staff [NSTF]	3

L	Cost category	€
M	Personnel costs (80%)	50878,58
N	Hardware (5%)	3179,91
O	Software (10%)	6359,82
P	Other costs, including training, refresher couces, etc. (5%)	3179,91
Q	Total, €	63598,22

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Hungary, the financing of the national EU fire statistics should be supported with 63000 € and 3 full-time employees.

Table 3-14: Calculation of the national costs of EU fire statistics for Ireland.

IRELAND		Area	Population density		Capital
		70273	70		Dublin
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		4904226	23937	31	124

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1746,00
D	Labour cost (€ per hour) [LCH]	32,30
E	Price level indicies EU [PLI]	137,20
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	5984,25
H	Total Labour costs (€),[TLC]	193291,28
K	Number of Staff [NSTF]	3

L	Cost category	€
M	Personnel costs (80%)	193291,28
N	Hardware (5%)	12080,70
O	Software (10%)	24161,41
P	Other costs, including training, refresher couces, etc. (5%)	12080,70
Q	Total, €	241614,09

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Ireland, the financing of the national EU fire statistics should be supported with 241000 € and 3 full-time employees.

Table 3-15: Calculation of the national costs of EU fire statistics for Italy.

ITALY 		Area	Population density		Capital
		301338	200		Rome
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		60359546	241767	237	948

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1558,70
D	Labour cost (€ per hour) [LCH]	29,80
E	Price level indicies EU [PLI]	102,90
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	60441,75
H	Total Labour costs (€),[TLC]	1801164,15
K	Number of Staff [NSTF]	39

L	Cost category	€
M	Personnel costs (80%)	1801164,15
N	Hardware (5%)	112572,76
O	Software (10%)	225145,52
P	Other costs, including training, refresher couces, etc. (5%)	112572,76
Q	Total, €	2251455,19

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Italy, the financing of the national EU fire statistics should be supported with 2251000 € and 39 full-time employees.

Table 3-16: Calculation of the national costs of EU fire statistics for Latvia.

LATVIA		Area	Population density		Capital
		64589	30		Riga
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		1919968	10415	87	348

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1577,00
D	Labour cost (€ per hour) [LCH]	10,50
E	Price level indicies EU [PLI]	74,20
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	2603,75
H	Total Labour costs (€),[TLC]	27339,38
K	Number of Staff [NSTF]	2

L	Cost category	€
M	Personnel costs (80%)	27339,38
N	Hardware (5%)	1708,71
O	Software (10%)	3417,42
P	Other costs, including training, refresher couces, etc. (5%)	1708,71
Q	Total, €	34174,22

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Latvia, the financing of the national EU fire statistics should be supported with 34000 € and 2 full-time employees.

Table 3-17: Calculation of the national costs of EU fire statistics for Lithuania.

LITHUANIA		Area	Population density		Capital
		65200	43		Vilnius
		sq.km	inh./sq.km		
No.	**Parameter fo Calculation**	**Basic Numbers for Calculation...**			
		Population (2019)	**Fires (Average 2014-2018)**	**Fire Deaths (Average 2014-2018)**	**Fire Injuries (Average 2014-2018)**
A		**2794184**	**11624**	**100**	**400**

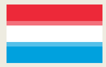
B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1595,00
D	Labour cost (€ per hour) [LCH]	10,10
E	Price level indicies EU [PLI]	64,90
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	2906,00
H	Total Labour costs (€),[TLC]	29350,60
K	Number of Staff [NSTF]	2

L	Cost category	€
M	Personnel costs (80%)	29350,60
N	Hardware (5%)	1834,41
O	Software (10%)	3668,83
P	Other costs, including training, refresher couces, etc. (5%)	1834,41
Q	Total, €	36688,25

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Lithuania, the financing of the national EU fire statistics should be supported with 37000€ and 2 full-time employees.

Table 3-18: Calculation of the national costs of EU fire statistics for Luxembourg.

LUXEMBOURG 		Area	Population density		Capital
		2586	237		Luxembourg City
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		613894	2262	1	4

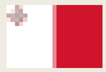
B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1427,00
D	Labour cost (€ per hour) [LCH]	42,10
E	Price level indicies EU [PLI]	147,00
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	565,50
H	Total Labour costs (€),[TLC]	23807,55
K	Number of Staff [NSTF]	0,4

L	Cost category	€
M	Personnel costs (80%)	23807,55
N	Hardware (5%)	1487,97
O	Software (10%)	2975,94
P	Other costs, including training, refresher couces, etc. (5%)	1487,97
Q	Total, €	29759,44

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Luxembourg, the financing of the national EU fire statistics should be supported with 30000€ and 1 full-time employee.

Table 3-19: Calculation of the national costs of EU fire statistics for Malta.

MALTA 		Area	Population density		Capital
		316	1562		Valletta
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		493559	1844	1	4

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1500,00
D	Labour cost (€ per hour) [LCH]	14,50
E	Price level indicies EU [PLI]	87,20
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	461,00
H	Total Labour costs (€),[TLC]	6684,50
K	Number of Staff [NSTF]	0,3

L	Cost category	€
M	Personnel costs (80%)	6684,50
N	Hardware (5%)	417,78
O	Software (10%)	835,56
P	Other costs, including training, refresher couces, etc. (5%)	417,78
Q	Total, €	8355,63

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Malta, the financing of the national EU fire statistics should be supported with 8500 € and 1 full-time employees.

Table 3-20: Calculation of the national costs of EU fire statistics for the Netherlands.

NETHERLANDS		Area	Population density		Capital
		41543	416		Amsterdam
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		17282163	38900	58	232

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1399.00
D	Labour cost (€ per hour) [LCH]	36.80
E	Price level indicies EU [PLI]	119.70
F	Average hours worked per fire event [AHFE]	0.25

G	Total hours for performance of yearly fire statistics [THP]	9725.00
H	Total Labour costs (€),[TLC]	357880.00
K	Number of Staff [NSTF]	7

L	Cost category	€
M	Personnel costs (80%)	357880.00
N	Hardware (5%)	22367.50
O	Software (10%)	44735.00
P	Other costs, including training, refresher couces, etc. (5%)	22367.50
Q	Total, €	447350.00

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Netherlands, the financing of the national EU fire statistics should be supported with 447350.00 € and 7 full-time employees.

Table 3-21: Calculation of the national costs of EU fire statistics for Poland.

POLAND		Area	Population density		Capital
		312685	121		Warsaw
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		37972812	146328	499	1996

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1766,00
D	Labour cost (€ per hour) [LCH]	11,00
E	Price level indicies EU [PLI]	57,30
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	36582,00
H	Total Labour costs (€),[TLC]	402402,00
K	Number of Staff [NSTF]	21

L	Cost category	€
M	Personnel costs (80%)	402402,00
N	Hardware (5%)	25150,13
O	Software (10%)	50300,25
P	Other costs, including training, refresher couces, etc. (5%)	25150,13
Q	Total, €	503002,50

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Poland, the financing of the national EU fire statistics should be supported with 503000€ and 21 full-time employees.

Table 3-22: Calculation of the national costs of EU fire statistics for Portugal.

PORTUGAL 		Area	Population density		Capital
		92390	111		Lisbon
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		10276617	35000	70	700

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1600,00
D	Labour cost (€ per hour) [LCH]	8,10
E	Price level indicies EU [PLI]	50,50
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	8750,00
H	Total Labour costs (€),[TLC]	70875,00
K	Number of Staff [NSTF]	5

L	Cost category	€
M	Personnel costs (80%)	70875,00
N	Hardware (5%)	4429,69
O	Software (10%)	8859,38
P	Other costs, including training, refresher couces, etc. (5%)	4429,69
Q	Total, €	88593,75

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Portugal, the financing of the national EU fire statistics should be supported with 88000€ and 5 full-time employees.

Table 3-23: Calculation of the national costs of EU fire statistics for Romania.

ROMANIA		Area	Population density		Capital
		238391	81		Bucharest
		sq.km	inh./sq.km		
No.	**Parameter fo Calculation**	**Basic Numbers for Calculation...**			
		Population (2019)	**Fires (Average 2014-2018)**	**Fire Deaths (Average 2014-2018)**	**Fire Injuries (Average 2014-2018)**
A		**19401658**	**27197**	**400**	**1600**

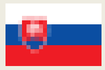
B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1600,00
D	Labour cost (€ per hour) [LCH]	8,10
E	Price level indicies EU [PLI]	50,50
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	6799,25
H	Total Labour costs (€),[TLC]	55073,93
K	Number of Staff [NSTF]	4

L	Cost category	€
M	Personnel costs (80%)	55073,93
N	Hardware (5%)	3442,12
O	Software (10%)	6884,24
P	Other costs, including training, refresher couces, etc. (5%)	3442,12
Q	Total, €	68842,41

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Romania, the financing of the national EU fire statistics should be supported with 69000 € and 4 full-time employees.

Table 3-24: Calculation of the national costs of EU fire statistics for Slovakia.

SLOVAKIA 		Area	Population density		Capital
		49035	111		Bratislava
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		5450421	9676	51	204

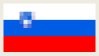
B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1572,00
D	Labour cost (€ per hour) [LCH]	13,40
E	Price level indicies EU [PLI]	79,20
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	2419,00
H	Total Labour costs (€),[TLC]	32414,60
K	Number of Staff [NSTF]	2

L	Cost category	€
M	Personnel costs (80%)	32414,60
N	Hardware (5%)	2025,91
O	Software (10%)	4051,83
P	Other costs, including training, refresher couces, etc. (5%)	2025,91
Q	Total, €	40518,25

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Slovakia, the financing of the national EU fire statistics should be supported with 40000 € and 2 full-time employees.

Table 3-25: Calculation of the national costs of EU fire statistics for Slovenia.

SLOVENIA 		Area	Population density		Capital
		20273	103		Ljubljana
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2019)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		2080908	5673	3	12

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1514,60
D	Labour cost (€ per hour) [LCH]	19,90
E	Price level indicies EU [PLI]	86,50
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	1418,25
H	Total Labour costs (€),[TLC]	28223,18
K	Number of Staff [NSTF]	1

L	Cost category	€
M	Personnel costs (80%)	28223,18
N	Hardware (5%)	1763,95
O	Software (10%)	3527,90
P	Other costs, including training, refresher couces, etc. (5%)	1763,95
Q	Total, €	35278,97

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Slovenia, the financing of the national EU fire statistics should be supported with 32000 € and 1 full-time employees.

Table 3-26: Calculation of the national costs of EU fire statistics for Spain.

SPAIN 		Area	Population density		Capital
		505990	94		Madrid
		sq.km	inh./sq.km		
No.	Parameter fo Calculation	Basic Numbers for Calculation...			
		Population (2018)	Fires (Average 2014-2018)	Fire Deaths (Average 2014-2018)	Fire Injuries (Average 2014-2018)
A		47.351.567	110000	190	760

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1577,20
D	Labour cost (€ per hour) [LCH]	22,80
E	Price level indicies EU [PLI]	96,90
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	27500,00
H	Total Labour costs (€),[TLC]	627000,00
K	Number of Staff [NSTF]	17

L	Cost category	€
M	Personnel costs (80%)	627000,00
N	Hardware (5%)	39187,50
O	Software (10%)	78375,00
P	Other costs, including training, refresher couces, etc. (5%)	39187,50
Q	Total, €	783750,00

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Spain, the financing of the national EU fire statistics should be supported with 783000€ and 17 full-time employees.

Table 3-27: Calculation of the national costs of EU fire statistics for Sweden.

SWEDEN 		Area	Population density		Capital
		449964	23		Stockholm
		sq.km	inh./sq.km		
No.	**Parameter fo Calculation**	**Basic Numbers for Calculation...**			
		Population (2019)	**Fires (Average 2014-2018)**	**Fire Deaths (Average 2014-2018)**	**Fire Injuries (Average 2014-2018)**
A		**10230185**	**24191**	**94**	**376**

B	Basic data for cost estimations	
C	Average annual hours worked per worker [AHW]	1424,00
D	Labour cost (€ per hour) [LCH]	37,30
E	Price level indicies EU [PLI]	131,40
F	Average hours worked per fire event [AHFE]	0,25

G	Total hours for performance of yearly fire statistics [THP]	6047,75
H	Total Labour costs (€),[TLC]	225581,08
K	Number of Staff [NSTF]	4

L	Cost category	€
M	Personnel costs (80%)	225581,08
N	Hardware (5%)	14098,82
O	Software (10%)	28197,63
P	Other costs, including training, refresher couces, etc. (5%)	14098,82
Q	Total, €	281976,34

Recommendation: Considering the national characteristics of the organization of fire safety in the EU member state Sweden, the financing of the national EU fire statistics should be supported with 282000€ and 4 full-time employees.