National Records of Scotland

Population and Migration Statistics Conference
26th October 2011 – Victoria Quay, Edinburgh

Feedback from Attendees

1.   Introduction

The Census has long been the main way to build a comprehensive picture of the population. However, user needs are changing and as the 10 year census cycle means that the statistics can quickly become out of date, alternative approaches to producing population and socio-demographic statistics are being explored.

In October 2011 National Records of Scotland (NRS) organised a conference for stakeholders which included a session to describe how administrative sources are being used to inform population statistics, quality assure the Census and contribute to potential alternatives to the Census. A workshop was held to gather feedback to help shape the requirements for any future demographic system.

Approximately 100 delegates representing many different organisations and interests attended the conference. Local Government accounted for almost half of those attending. The NHS, Central Government, Further and Higher Education, religious bodies, Government agencies, Emergency Fire and Rescue, and voluntary organisations were also well represented.

This report summarises the feedback received in the course of the conference. The two sources were:

(i) a questionnaire asking delegates for their views on the relevance of particular topics and their relative priority in terms of accuracy, frequency and geography (Annex A);
(ii) workshop discussions.

2.   Questionnaires

We received completed questionnaires from around 40% of the delegates.

2.1. Use made of Census data

Policy Development, Monitoring and Evaluation together accounted for the greatest use with 46% of respondents stating that they used Census outputs for these purposes. Resource Allocation and Service Planning accounted for 36% and Academic/Social Research 13%. The remaining 5% included uses such as economic analysis, epidemiology, labour market analysis, population health surveillance and market research. Delegates were able to select more than one type of use.

2.2. Topics and quality characteristics

Introduction
Delegates were offered a list of traditional census topic areas (population, household and family structure, housing, ethnicity and migration) and asked to rate them in terms of relevance to their work ('high', 'medium' or 'low'). For each topic they were asked to apportion 10 points between three quality characteristics (accuracy, geography, frequency) according to relative importance. These scores were added

to generate an overall score for each quality aspect and topic area. The results are presented in Annex B.

Current census standards were taken as a reference point for the high quality end on accuracy and geography (data zones were given as an example for low level geography). The mid-year estimates were used as a high quality standard in terms of frequency (i.e. annual).

Relevance

Population statistics were most frequently seen as an area of high importance: with 38 of 42 delegates giving this response. This was followed by migration with 21 delegates ranking its importance as high and housing with 19 delegates considering it of high importance. Household and family structure was least likely to be seen as high in importance with only 15 delegates expressing this view.

Scoring

Although many individual respondents expressed clear preference in terms of relative priority across the 3 quality aspects, the overall distribution within each topic area was fairly balanced (Annex B). The total points available for each category were split roughly into thirds across the 3 quality characteristics.

Figures 1 to 6 in Annex C show the distribution of the scores illustrated by box plots. A box plot is a visual representation of the dispersion of the data. More information on how to interpret the box plots can be found in Annex C. Figures 1 to 5 show the distribution of the scores given by respondents to each Census topic. Figure 6 shows overall distribution of the scores given to all five topics.

Figure 2 and Figure 3 show that the Household and Family Structure and Housing topics have the highest maximum score for low geography (score 8). Figure 5 show that respondents gave the highest score for high frequency under Migration (score 8). Figure 1 and Figure 5 shows that more than 50% of the respondents gave a score of 4 or higher for high frequency for population and migration. From Figure 4 it can be seen that Ethnicity has the most variation in scoring. Figure 6 shows that on combining scores for all topics, the dispersion of scores was the same for high frequency and high accuracy.

Figures 7 to 12 in Annex D are radar charts showing average scores given to each category within each topic. The charts allow a direct comparison of the mean scores given to each quality characteristic. Figure 7 show that for Population, high frequency has a slightly higher average score compared to high accuracy and low geography. Similarly, from Figure 9 it can be seen that for Housing, low geography received the highest average score compared to the other two categories. Figures 10 and 11 show that low geography was seen as slightly less important than high accuracy or high frequency for both Migration and Ethnicity statistics.

2.3. Summary of delegate comments by topic

Population

As the category most relevant to the majority of delegates who responded, the creation and maintenance of accurate population counts is vital to a wide range of organisations.

Delegates who specified requirements on frequency indicated that annual was regular enough, and one delegate requested a preference for more frequent than annually.

Of the 11 delegates who specifically mentioned geography, nine specified data zone level, output area or another small geography smaller than Local Authority level. Two indicated population figures at Local Authority level are sufficient. Other breakdowns that were noted as being helpful were age, gender, religion and single year of age.

Three delegates highlighted that accuracy was of greatest importance to their needs. A small number of delegates assigned equal and high importance to all of the categories accuracy, frequency and geography.

Household and family structure
Feedback from the 31 delegates who provided comments indicated that a high priority should be given to identifying lone parent households, overcrowding, pensioner households and older people living alone for care provision purposes.

From the comments recorded these data are used in highly specific ways - some users requiring accuracy, some low geographies and some greater frequency. Delegates who expressed a preference indicated that smaller geographies were preferable over higher frequency.

Housing
Several delegates stated that Housing data were not critical to their needs. However, for users that do use Housing statistics, tenure data and property type were noted as being particularly useful. One respondent attached a low priority to Housing data but noted that this was on the assumption that Council Tax (and associated data collection) would continue in its current format and not be replaced by a local income tax. More frequent data was suggested along with a request for data on housing adaptations e.g. telecare, community alarms. Low Geography was seen as key. This was reflected in the high scores allocated to Housing at Low Geography in the questionnaire responses.

Ethnicity
Many of the 23 respondents commented that the Census is a key data source for complete ethnicity data. Priority was given to all of accuracy, frequency and small geography by different users, however accuracy was noted as being particularly important because of the small numbers in some categories. In addition, delegates notes that rapid changes in Scotland's population require more frequent information on the composition of the ethnic population.

Migration
Twenty-six delegates provided comments on migration statistics. Accuracy was highlighted as a priority, however, requirements for data at small geographies and a higher frequency of output were also mentioned.

Recent improvements in the data were appreciated and lack of an alternative source for these data noted.

Other statistics
There were 31 varied responses to this open question. Suggestions ranged from highlighting the importance of some Census data already collected (e.g. self assessed health, educational attainment, employment status) to requesting new information (e.g. crime and income) and requests to reconsider including an income question (11 delegates).

Several respondents stated that the Census was their only source of self assessed health information, qualifications, local travel to work patterns, overcrowding and religion not available elsewhere.

The data are used for various purposes including community fire safety modelling, calculating Carstairs mental health scores, and healthcare planning.

3. Workshop discussions

Delegates were allocated to one of four groups and views on priorities for an alternative system for population statistics discussed. These discussions were recorded and are summarised below.

3.1. Traditional Census

The discussions revealed strong support for the continuation of the Census in its current format. It is widely used as a baseline for other statistics and seen as the only source of vital data on topics such as self assessed health and religion.

As an organisation NRS is generally trusted by people to safeguard their data and treat it with respect. Traditional census records are important for genealogical research and ancestor tracking.

Although only two negative aspects to the Census were mentioned, collectively they carry some weight. Namely the cost of a Census is high and the data are soon out of date.

3.2. Alternatives to the Census

A lot of discussion took place about possible alternatives to a traditional Census, however it was acknowledged that no one source could provide the complete picture that the Census does.

The following points were made in the discussions.

- Additional questions could be added to Council Tax forms to collect household data.
- One Local Authority is already using Council Tax data for household structure data.
- Some groups traditionally missed from a Census may be picked up through Local Authority data sources e.g. Council Tax systems and Citizens Account.
- Citizens Account could be used to link people to property.
- The Electoral Register and the Landlord Registration Scheme are also possible data sources with partial coverage.
- Regular sample surveys could provide a replacement for Census baseline as used in Local Authority statistics. At the same time there was concern that Local Authorities might be required to carry these out themselves.

- Government IT systems were seen as important. There were some concerns about potential risks of leaking confidential data, especially where several organisations were involved in data sharing. This was contrasted with the sound reputation of the Census and the NRS who are trusted by the public.

3.3. Census data attributes

With respect to the trade-off between accuracy, frequency and geographical level of statistical outputs, delegates appreciated that there was a need for a balance However, the optimal balance is different depending on user requirements. Local Authorities would prefer small geographies at the expense of frequency, whereas national organisations could use higher geographies but at a higher frequency.

The balance between accuracy, frequency and geography is also different depending on the topic. For example, it was noted that high accuracy was preferred for equality statistics and data zone level information would be most useful for health provision planning purposes.

Delegates mentioned the following aspect of census type statistics that were important in their work.

- Population figures broken down by age and sex.
- Information on household structure was important for Local Authority planning around health and housing provision.
- Information on the social characteristics of the population was necessary in addition to population counts.
- Health data was required at data zone level.
- Accurate migration information is essential.
- Ethnicity data are important for equality agenda. Accuracy is the priority but frequency can be more important at times.

3.4. General comments and questions

A number of more general comments were made in the workshops.

- There was disappointment that the income question was not included in the 2011 census.
- There was interest in whether data from administrative sources used instead of a census would be subject to the same disclosure and confidentiality protection as the current Census.
- Delegates pointed out the public attitudes are becoming more favourable towards the sharing of personal information (e.g. Facebook and social networking).
- It was suggested that establishing unique personal identifiers for everyone in the Scottish population would be helpful.
- It was suggested that, in the absence of a Census, some of the Census funding could be used to improve the quality of local authority databases. For example for the purposes of household low level comparisons to identify households with children living in poverty.
- Change to the Census will impact on the processes that are reliant on it, e.g. planning and service delivery.

4.   Conclusions

A wide range of interested parties were represented at the conference and shared their views on the future of the Census.

The balance between accuracy, frequency and geographical level of Census outputs is a complex one and requirements vary between users and Census topic areas. Users expect a reasonable level of accuracy for all statistics. Preferences on frequency depend on the subject area, however at least annual was preferable for most subject areas. The geographical level of outputs required is dependent on the nature of the user requirement. Data zone was suggested as an ideal geography for most outputs.

Population figures were deemed to be the most relevant Census output to the majority of delegates and migration was deemed the second most important topic. Use made of housing, housing structure and ethnicity statistics varied.

There are pros and cons to the Census in its current format, but users were currently in favour of its retention. It is a rich data source of the entire population and collects some data that are not currently available from any other source. In addition to statistical uses, Census records are used in genealogy and this important use should not be under estimated. On the other hand, traditional census taking is costly and the data are soon out of date. User needs are changing and there is demand for both more information and more frequent statistical outputs.

Annex A – Example of Delegate Questionnaire

PAMS Conference Workshop 26 October 2011: Alternatives to the Census

Participants to complete individually before the conference, hand in on the day and be prepared to discuss.

The **Census** currently provides the base for many of the population and socio-demographic statistics in Scotland, providing comparable information from national to local level on a range of topics, and acting as a benchmark for many other statistics. However the Census is becoming **increasingly costly**, a **more mobile** population makes enumeration more difficult, and a snapshot **every 10 years** means data becomes less accurate and relevant over the 'inter-censal' period.

NRS are investigating the use of **integrated data sources** to replace or complement existing approaches to the production of population and socio-demographic statistics in Scotland. A range of **statistical options or hybrids** are being considered, under the following three high-level categories:
- Census type solutions
- Administrative data solutions
- Survey solutions

The options will be assessed against a set of **criteria** such as public acceptability, cost, fitness for purpose, accuracy, frequency, geography, technical feasibility etc.

A recommendation is required to be made in **2014** as to whether to continue with a **2021 census** or whether to go with an **alternative option**. The final recommendation will need to balance user needs against the set of criteria chosen and it is inevitable that there are going to be some **compromises** in the solution adopted. For example, for any given topic a more frequent supply of outputs may require us to compromise on a higher level of geography or lower quality - or vice versa.

The different statistical options that we are investigating will offer different advantages and this form asks for your views on the **key trade-offs** between **accuracy, frequency and geography** for different socio-demographic topics.

Information about you and your organisation
1. Contact details

Contact name _____

Organisation name (if appropriate) _____

Email address _____

2. What type of organisation do you represent?		3. What are your main uses of population and socio-demographic data?	
Please tick one		*Please tick all that apply*	
Central government		Academic/social research	
Charity/voluntary organisation		Marketing research	
Government agency		Policy development	
Local government		Policy monitoring and evaluation	
Higher/further education		Resource allocation	
Private/commercial organisation		Service planning	
Other (please specify)		Other (please specify)	

*We may make individual responses publicly available and may refer to the response of specific individuals or organisations in our workshop report. Please indicate here if you do **NOT** wish your response to be used in this way.*_____

Please apportion 10 points across the three quality characteristics (no decimal places please!)

Theme	Importance/ relevance to you (High, Medium, Low)	High frequency e.g. annual	Low geography e.g. datazone	High accuracy e.g. current census standards
***Example -** Population*	H	7	1	2
Comments	*We want all of these but on balance regularly updated outputs of reasonable quality are more important to us than detailed geography. In-migration puts real pressure on local services and we need to know more about how the population profile of the LA is changing. Annually is good enough.*			
Population e.g. age, sex, marital status				
Comments				
Household and family structure e.g. no., relationships				
Comments				
Housing e.g. tenure, type, no. of rooms				
Comments				
Ethnicity e.g. ethnicity, national identity				
Comments				
Migration e.g. immigration, internal migration				
Comments				
Other statistics (please specify)				
Comments				

Annex B – Summary of Questionnaire Scores

This table shows the distribution of the total scores assigned to each quality characteristic within each topic.

Headings	Ratings	Percent
Population	High Frequency	36.99
	Low Geography	33.41
	High accuracy	29.59
	Total	**100.00**
Household and family structure	High Frequency	31.78
	Low Geography	32.76
	High accuracy	35.45
	Total	**100.00**
Housing	High Frequency	32.52
	Low Geography	36.92
	High accuracy	30.56
	Total	**100.00**
Ethnicity	High Frequency	35.21
	Low Geography	30.32
	High accuracy	34.47
	Total	**100.00**
Migration	High Frequency	34.46
	Low Geography	31.81
	High accuracy	33.73
	Total	**100.00**

Annex C – Distribution of scores assigned to each quality characteristics by topic

Box plots show the minimum and maximum scores allocated to each quality characteristic, the median score (or middle value, denoted by an 'X') and the range of the middle 50% of the data (the 'quartile range', the upper and lower bounds of the rectangle).

Figure 1: Distribution of scores - Population
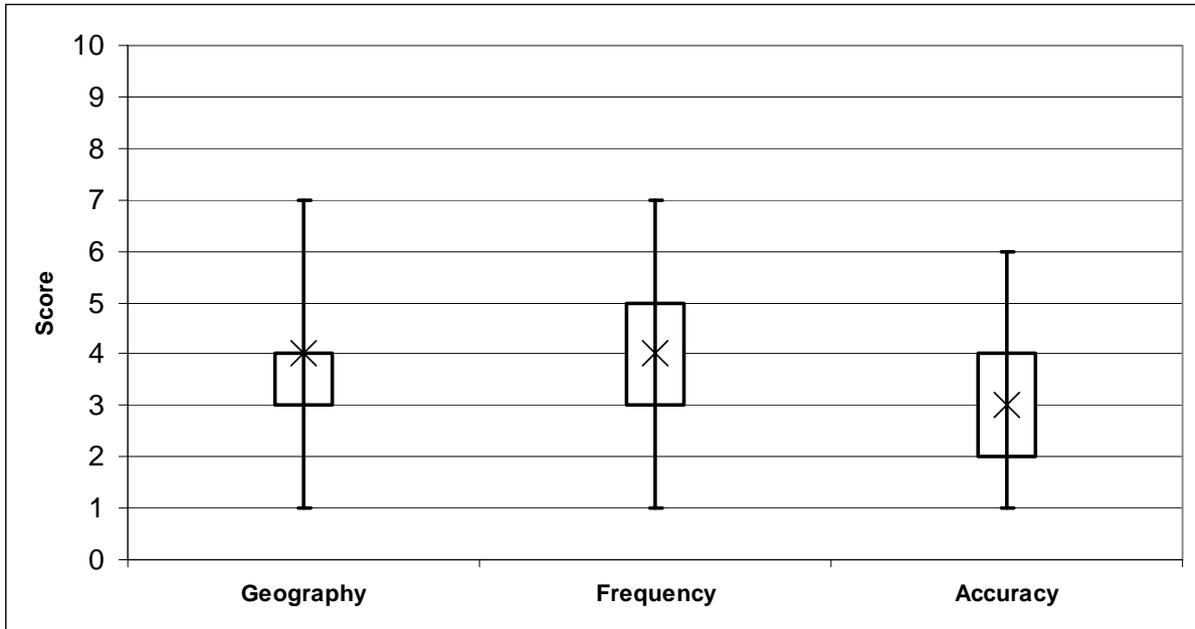
Figure 2: Distribution of scores for Household - Family Structure
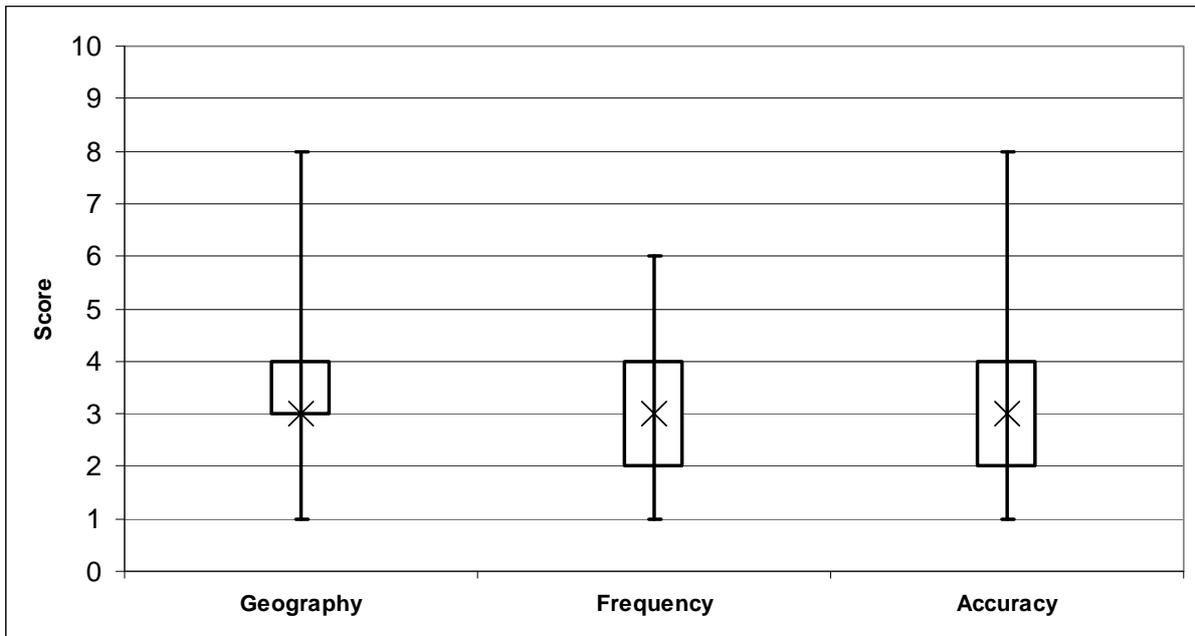
Figure 3: Distribution of scores - Housing
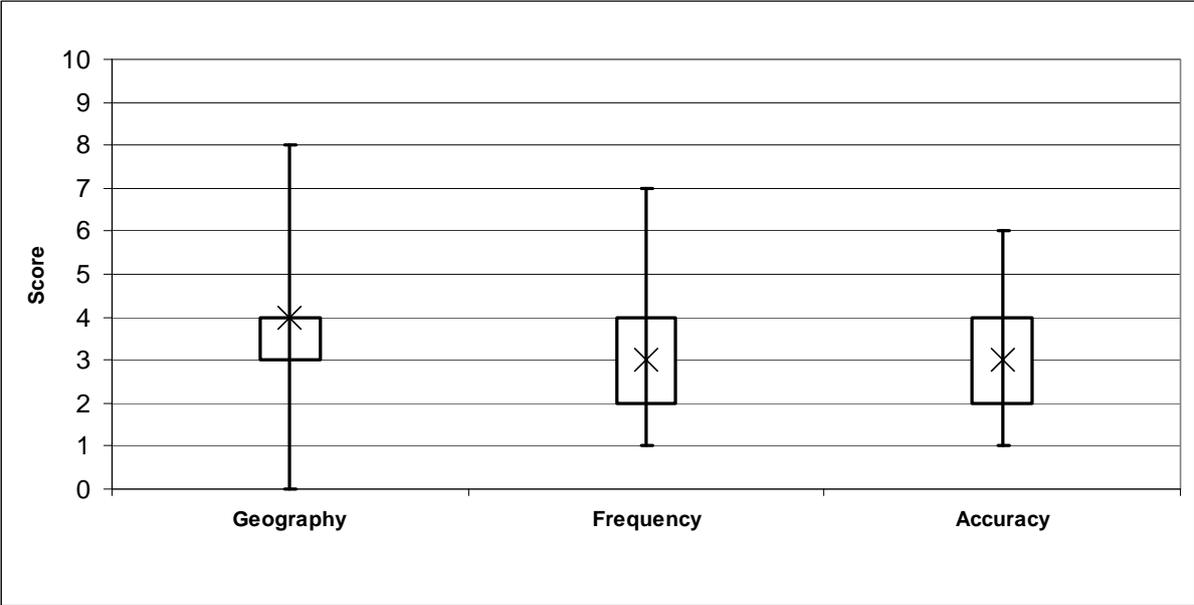
Figure 4: Distribution of scores - Ethnicity
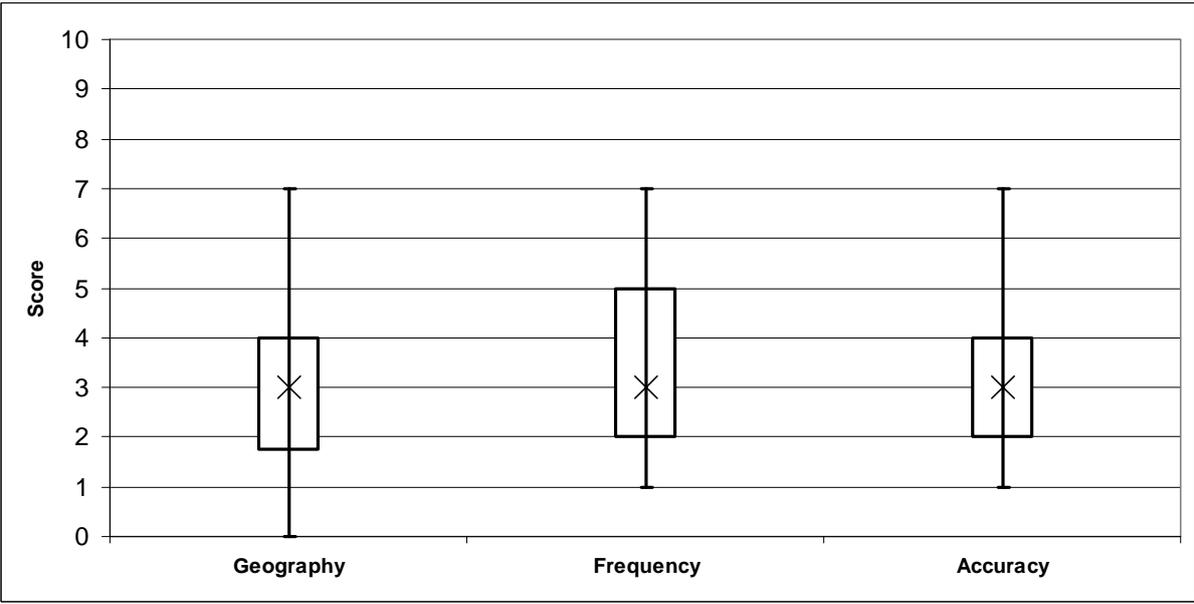
Figure 5: Distribution of scores - Migration
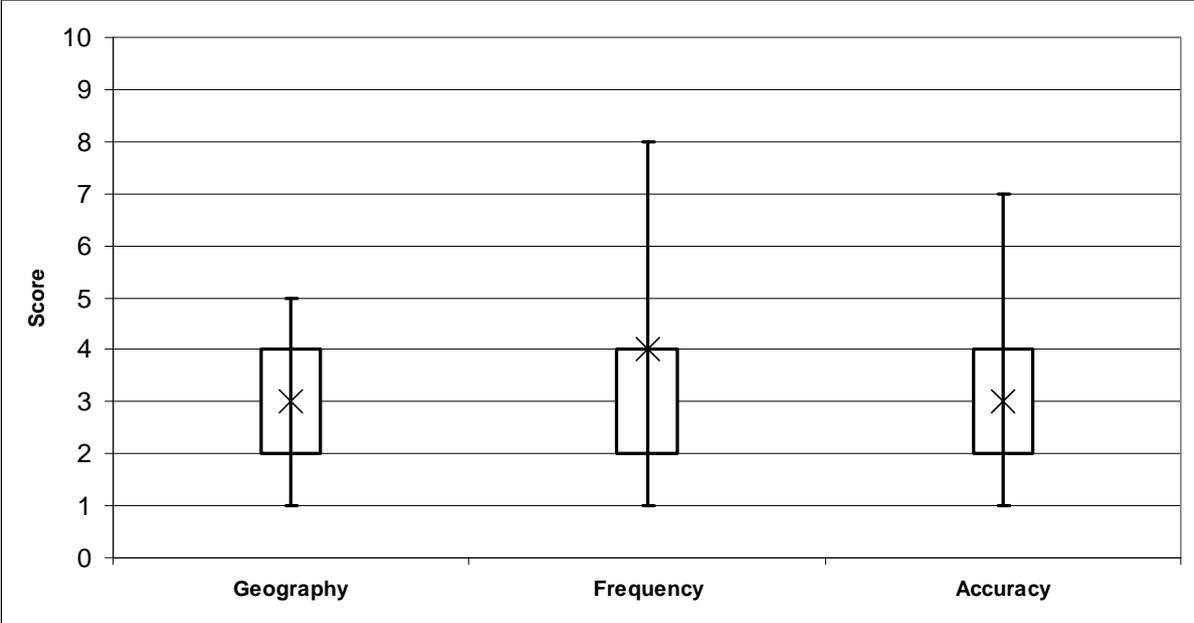
Figure 6: Distribution of scores – all topics
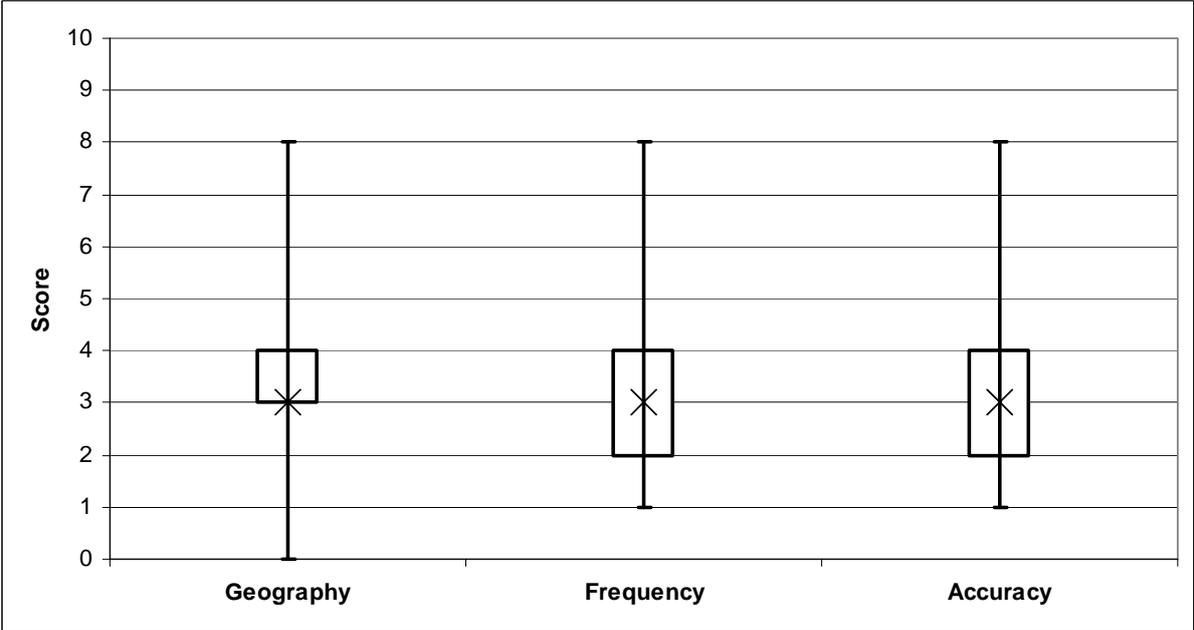
Annex D – Average scores assigned to each quality characteristics by topic

Figure 7: Comparison of average scores - Population
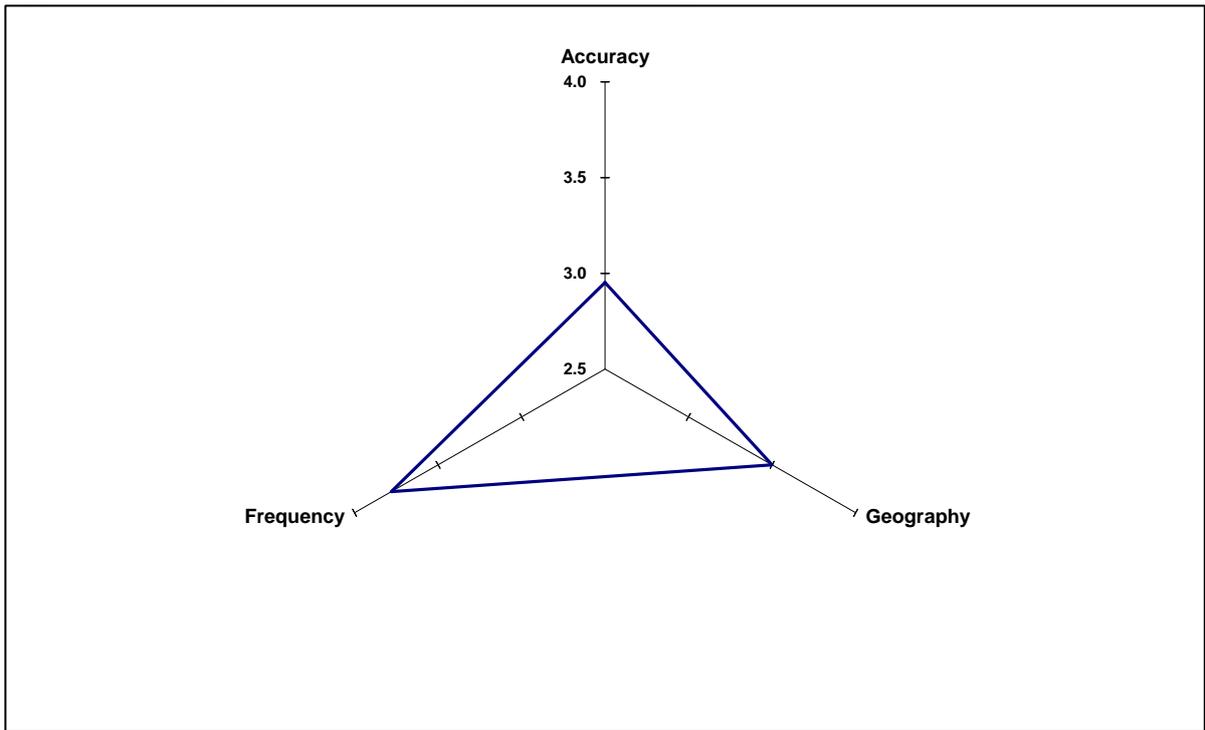
**Figure 8: Comparison of average scores - Household and Family structure**
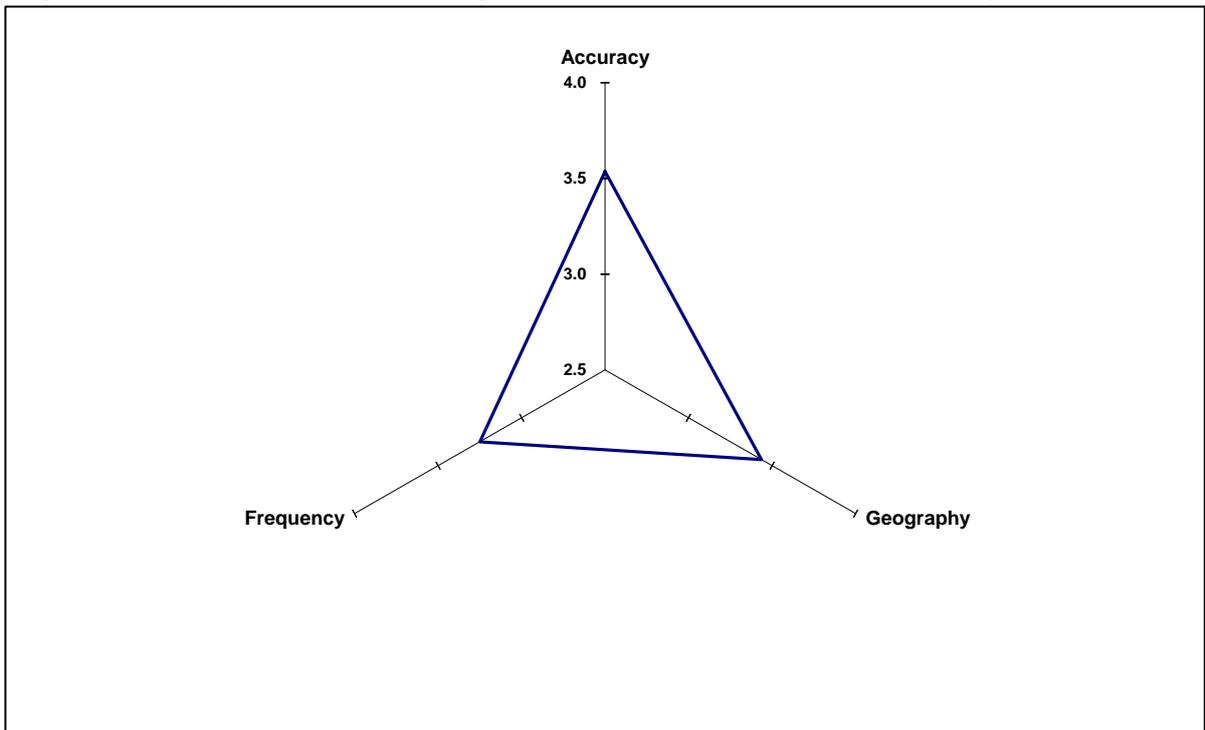
Figure 9: Comparison of average scores - Housing
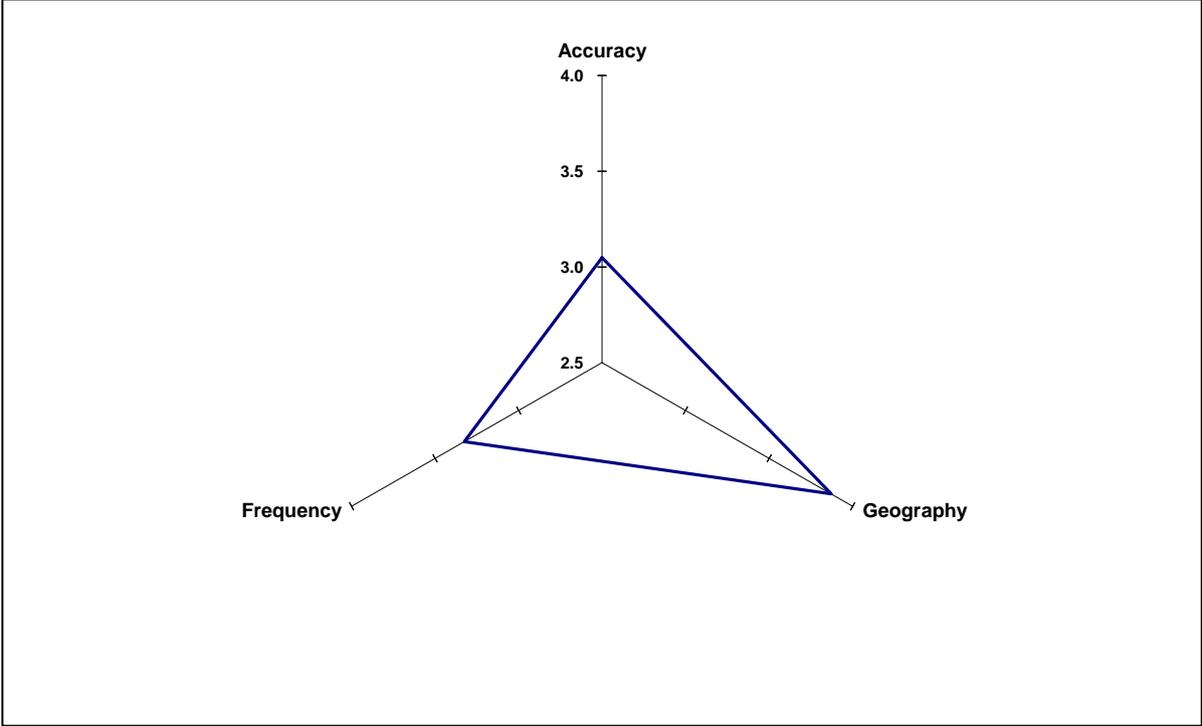
**Figure 10: Comparison of average scores - Ethnicity**
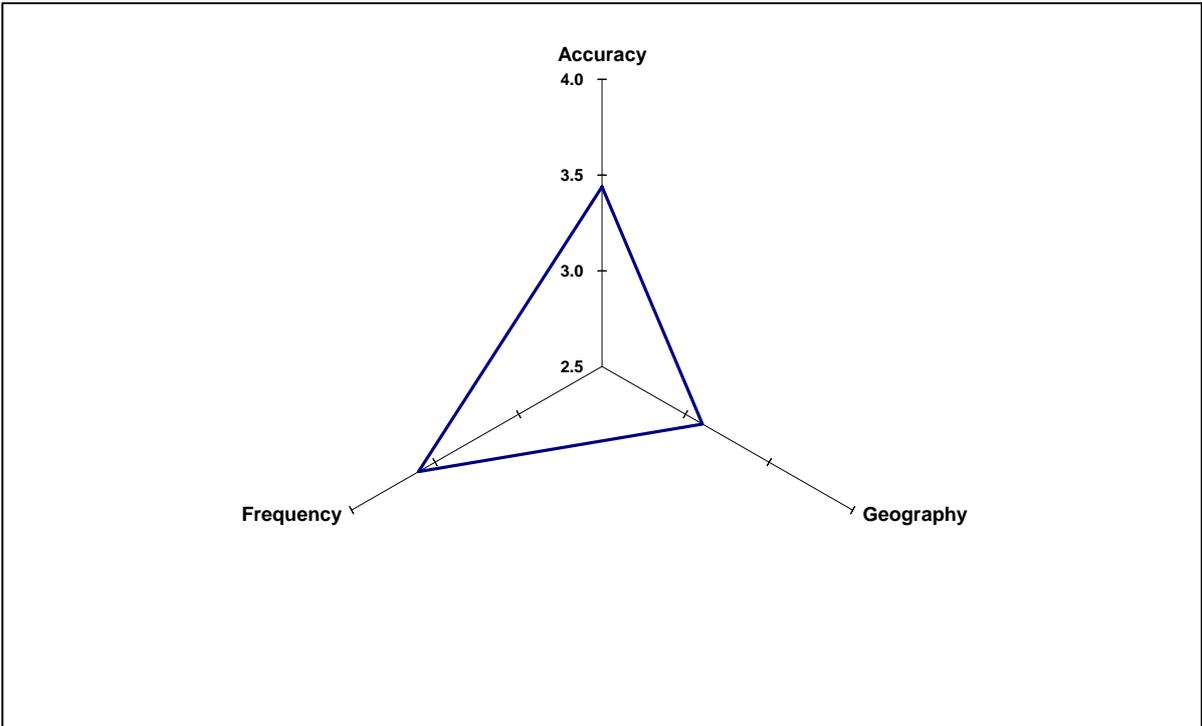
Figure 11: Comparison of average scores given for migration
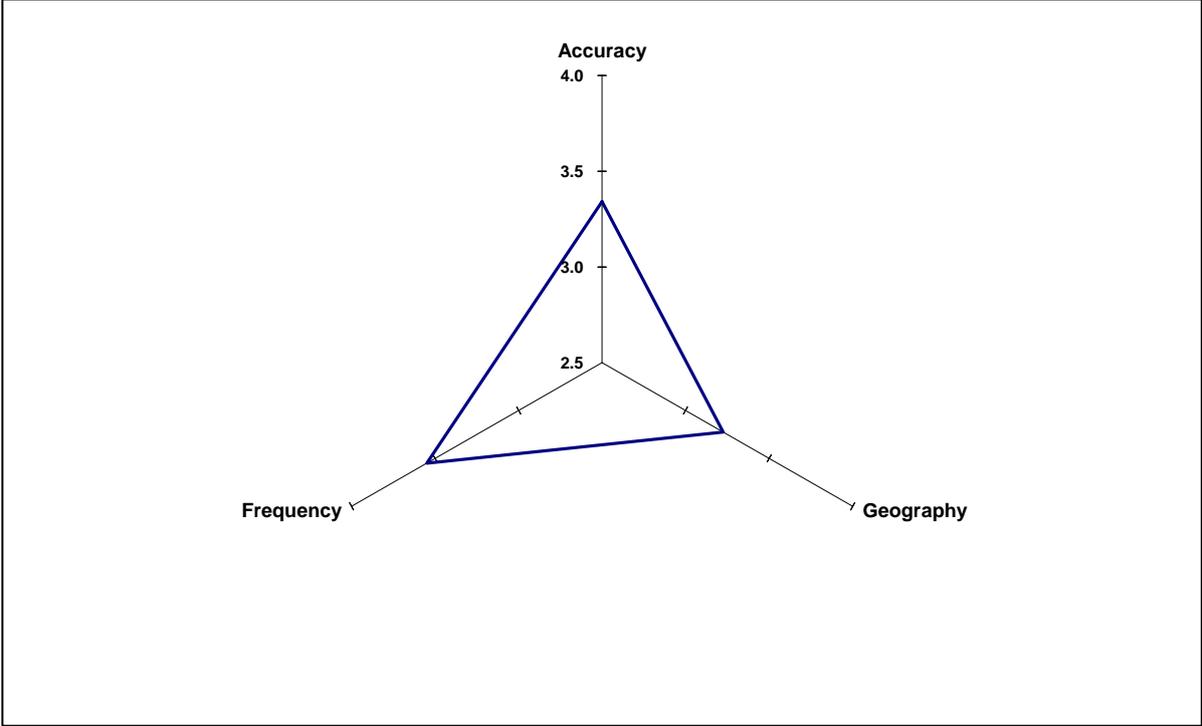
Figure 12: Comparison of average scores – all topics