

BOWLEY

ELEMENTS
OF
STATISTICS

1902

B

HA
29
B685
1902
c. 1
ROBA



3 1761 04945084 4

ELEMENTS
OF *~~~~~*
STATISTICS
~~~~~

ARTHUR L. BOWLEY.

List of the MEMBERS of this FAMILY, of VISITORS, of BOARDERS, and of SERVANTS, who SLEPT or ABODE in this dwelling on the
NIGHT of SUNDAY, APRIL 5th 1891.

Cols.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
	NAME and SURNAME.	RELATION to Head of Family.	CONDITION as to Marriage.	SEX.	AGE last Birthday.	PROFESSION or OCCUPATION.				WHERE BORN.	If (1) Deaf and Dumb. (2) Blind. (3) Lunatic Imbecile or Idiot.
	<p><i>No persons absent on the night of Sunday, April 5th, to be entered here; except those who may be travelling or put at Work during that night (and are not elsewhere returned), and who Return Home on Monday, April 6th.</i></p> <p>Write after the Name of the Head of the Family, the Names of his Wife, Children, and other Relatives; then Visitors, Boarders, and Servants.</p>	State whether Head, or Wife, Son, Daughter, or other Relative, Visitor, Boarder, or Servant.	Write either "Married," "Widower," "Widow," or "Single," opposite the Names of all Persons, except Young Children.	Write "M" opposite Males and "F" opposite Females.	For Infants under One Year state the age in Months, writing "Under 1 Month," "1 Month," "2 Months," &c.	<p><i>Before filling up columns 6, 7, 8, and 9, you are requested to read carefully the Special Instructions printed on the other side.</i></p> <p>Employer.</p> <p>Employed.</p> <p>Neither Employer nor Employed, but working on own account.</p>			<p>Opposite the Names of those born in ENGLAND and WALES, write the COUNTY, and TOWN or PARISH.</p> <p>If born in SCOTLAND, IRELAND, the BRITISH COLONIES or the EAST INDIES, state the Country or Colony.</p> <p>If BORN IN FOREIGN PARTS, write the particular State or Country; and if also a BRITISH Subject, add "British Subject," or "Naturalised British Subject," as the case may be.</p>	Write the precise Infirmary, if any, opposite the name of the person; and if the Infirmary dates from childhood, add "from childhood." Do not use such a general term as "Afflicted" or "Infirm."	
			<p>I declare the foregoing to be a true Return, according to the best of my knowledge and belief.</p> <p>Witness my hand, (Signature).....</p>								
	If you occupy less than five rooms, write in this space the number of rooms occupied by you.										

List of the MEMBERS of this FAMILY on the

Cols.	1.	2.
	NAME and SURNAME.	RELATION to Head of Family and Dumb. Imbecile idiot.
	No persons absent on the night of Sunday, April 5th, to be entered here; except those who may be travelling or out at Work during that night (and are not elsewhere returned), and who Return Home on Monday, April 6th.	State whether precise Head, or Wife, or Son, or Daughter, or other Relative; if the Visitor, Boarder, or Servant. Do not add "from a general fictitious" or
	Write after the Name of the Head of the Family, the Names of his Wife, Children, and other Relatives; then Visitors, Boarders, and Servants.	
	If you occupy less than five rooms, write in this space the number of rooms occupied by you.

CHAPTER III.

ILLUSTRATIONS OF METHOD.

SECTION I.—THE POPULATION CENSUS.

THE population census will provide good illustrations of the principles laid down in the last chapter, both because we shall

be at first on familiar ground, since every one knows its scheme, purpose, and details, and because the

form of inquiry used for the collection of the original data brings out very prominently the difficulties met with in detailed statistical investigations.

The first thing to be considered is the exact object for which the census is undertaken. It is for demographical purposes; to supply information as to the numbers and

local distribution of the population, the numbers of each sex and age, their so-called civil condition (*i.e.*, whether single, married, or widowed), and their nationality. This is the

minimum information necessary for administrative purposes. In addition to these facts there are very many others which the

statesman and the economist wish to know about each member of the population, and the census form is the only means in

England of collecting universal data; the question as to which of these shall be investigated and which neglected, is decided

more by expediency than on principle. Of these desiderata the following may be mentioned: the

size and structure of the family, its position in the social scale, the economic position of its head; the nature of employment

of its members, the wage or income of each member and of the family as a whole, the rent and size of their house, their educational condition, the ages at which they commenced or retired

from work, their migrations, their combination in religious or other bodies, and their infirmities. It is clear that some of

this information must be dispensed with, if the form is not to be overcrowded, and if the tabulation is to be finished in any

reasonable time; and an examination of the general nature of the questions which can suitably be put will show how the necessary selection is made.

First, the questions must be those which the informant is able to answer. Now, if the questions were only to be put to educated and methodical persons, doubtless a full account could be given of the family migrations and of the ages at which each member had been at work; but the peculiarity of the census is that it is universal, and the questions must be such that the least educated and most unthrifty householder shall be able to answer; in many cases such facts would have been unrecorded and forgotten.

Secondly, the questions must be perfectly definite, so that there can be no doubt as to what the right answer should be.

✓ **Ability to answer.** The only answers which are of value to the statistician are "yes," "no," or a simple number. Adjectives and adverbs such as many, often, partly, &c., bear different numerical meanings to different people, and, though they may express fairly clearly the position of an individual, are nearly useless for tabulation,* which is their only purpose so far as the census is concerned. Thus the question as to education would have to be, not "state whether well, moderately, or badly educated," but "state at what age school was left," or "how many years at school?" But even if such questions were not excluded by our first test, by the forgetfulness of the informant, the statements given would be of little practical value, and very often incorrect. An inquiry as to wage and income could not be made sufficiently definite without so many questions as to require a form to itself; for wages, as we shall see when considering the Wage Census, require very careful definition, and many subsidiary questions must be put to get a proper estimate; the simple query, "what is your weekly wage or annual income?" would be answered on so many varying principles that the result would be valueless.

Thirdly, the questions must be such as will be answered truthfully and without bias. There is hardly a demand on the census form which would not be excluded, if ✓ **Veracity.** this rule was too rigorously enforced, as we shall see immediately. The worst offender in this respect is the

* But see p. 138, *infra*.

question, *Employer or employed?* For though there are many cases in which a man is both employer and employed so that this question should be excluded by our second test, many persons consciously exaggerate their social importance by erroneously replying the former. Questions relating to social position must generally be excluded by this rule.

Fourthly, the questions must be those which will be answered willingly, and must therefore not be inquisitorial, or such as **Reluctance to answer.** to raise apprehension of a change of law or an imposition of taxes. Questions as to membership of trade unions, or of friendly societies, or as to insurance, would be thought inquisitorial. Many would refuse to state their incomes, holding it to be no one's concern but their own. Questions as to rent might be regarded as possibly leading to taxation. Questions as to religion are badly answered, as was shown in the evidence before the Census Committee of 1890,* and should be excluded by each of these four rules. Some persons do not know what their religion should be named, others would find the question indefinite, others would deliberately answer wrongly, and many not at all.

The questions on the census form† not excluded on one or other of these grounds are Nos. 1, 2, 3, 4, 5, and 10; these are fairly definite, and householders are generally able and willing to give correct answers to them. Questions 6, 7, 8, 9, and 11 compete with many others, which lead to equal inaccuracies, for a space on the census sheet. No. 6 has long held its place because of its great importance; Nos. 7, 8, and 9 are on their trial. A further discussion of the merits of some of these is to be found in the Report of the Committee already mentioned; here it is only intended to indicate the general grounds of inclusion or exclusion.

So far we have not discussed the important question as to who should fill in the form. If, as in the English Census, **Filling up of the form.** it is to be filled in by the householder, the questions must be much simpler in matter and words than if it is to be filled in by an official teller. In the latter case the form may be much more complicated, the questions more inquisitorial and such as might lead to indefinite answers on the part of ignorant people; for the teller would insist on

* Report of Committee on the Census, 1890 (C.—6071). † Facing p. 23.

an answer, be able to exclude those obviously wrong, and cross-question till the indefinite answers were so altered as to allow definite tabulation. In a great and complex undertaking like the Census, where many tellers must be impressed for a single day's work, their instructions and the general plan must be sufficiently simple; but as the extent of an inquiry contracts, the tellers can receive more complete instructions, and the information requisitioned may be more complex. This is of most importance in connection with columns 6-9.

The general shape and appearance of the sheet needs attention. If the structure of the family is to be shown, the answers are best given on a single sheet, which must contain enough lines for the largest ordinary household, so that the trouble of fastening together of many couples may be avoided, and tabulation not be hindered. The spaces must contain plenty of room for answers in uneducated handwriting, without making the whole so large as not to lie easily on a desk. The instructions must be distinct and visible, and placed in close connection with the answers; to further this, a skilful use may be made of capitals, italics, and different founts of type. On the form facing p. 23, those in use are roughly reproduced in miniature.

The form should always show for what purpose the figures are collected, and how they will be used, in order to enlist the support of the informant and allay misapprehension. The extent to which this should be done depends on whether the filling-up is compulsory, as in the population census, or voluntary, as in the wage census. In the case before us no preamble is necessary, since every one knows the main features of a census, and most are willing to further its objects; but it must be shown that the inquiry is sanctioned by Parliament, and that compliance is compulsory. This is done on the back, on the fold which is outside before the form is opened; and even though penalties are threatened against absence of or falsification of returns, the last sentence describes the object of the inquiry and guarantees the informant against malicious use of his answers. Where information is voluntary, a careful letter should be printed and circulated with the form, persuading the informant to give his assistance.

While the main part of the form is filled in by the house-

holder, other parts are filled in by the officials, and with very little trouble a good deal of subsidiary information can be collected in this way. On the outside the Parish, Town, Sanitary District, Street, and Number are endorsed, so that the answers can be tabulated for any of these districts. The teller could also, as he took the form, enter the number of stories to a house, which is not done in the English Census, and other information as to the style of house and street might be endorsed. In a more intensive investigation, Mr Charles Booth's assistants, for instance, could be trusted to come out of a house with an accurate knowledge of many interesting details.

We can now proceed to the individual criticism of the form in the light of the rules suggested above. In the first place, even the arrangement of columns is not perfect. To labourers who are not in the habit of writing at all, and who have (to judge from election posters) to be instructed how to put their mark in the right place on a ballot paper (many papers being destroyed simply through ignorance), this arrangement of horizontal and vertical columns would be confusing, and without help they would not gather at all what they were to do. They would fill up more easily a paper in which the answers were to follow the questions immediately:—

State your Name _____

State your Age _____

State your Sex _____

Unmarried, Married, or Widowed _____

and so on.

This form, however, could only be used if a separate paper were to be filled in for every individual, children and all. Other elementary matters might be improved. On looking through the form a great number of words and phrases will be found which are not in common use, *e.g.*, abode, dwelling (as a noun), elsewhere, East Indies, imbecile, "precise" infirmity, general term, column, the foregoing, condition as to marriage. In column 1 the phrase, "name and surname" reads as though surname were not a name, and perhaps the word "surname" is not in general

use, so that the printed word might be taken to mean title, and the confusing answer "none" written under it. Does the instruction "write after" mean to the right, or below?

The first question, which for the general purpose of the census should be the most definite of all, leaves some room for doubt. What of a night-watchman returning at 4 A.M., or a printer at 2 A.M.? What constitutes a traveller: does a man who leaves the house before midnight, or

Criticism of the questions.

"Slept or abode." a man who goes down to Brighton by the theatre train come under the term? Is midnight or 2 A.M. the critical time? What of a person who dies at 1 A.M., or a birth at midnight? How is the householder to know whether any of his establishment are returned elsewhere? Since too many instructions only lead to confusion, the tellers should be specially taught the answers to such questions.

The very meaning of the phrase "population of a district" is open to much doubt. In France "la population de fait," which consists of all present in the given district at the given moment, is distinguished from "la population de droit," which consists of all usually resident in the district, including those temporarily absent, and excluding those only momentarily present, and from "la population municipale," which is "la population de droit," less prisoners, hospital patients, scholars resident in schools, members of convents, the army, and so on.* The English Census counts "la population de fait." In the United States we find a "constitutional population," which excludes residents in Indian Reservations, the Territories, and the District of Columbia; the "general population," which includes in addition the Territories (except the Indian Reservation, Indian Territory, and Alaska); and the "total population," which includes all excluded in the former.† In the future questions will arise as to the inclusion of the Philippines and Cuba. Notice that the Channel Islands and the Isle of Man are included in the English Census.

* See Bertillon, *ibid.*, p. 146.

† Willcox: *Area and Population of the United States at the XI. Census*, a book which gives a very useful criticism of the accuracy of the most elementary data of statistics. It is a pity that space is wasted in a useless attempt to supplant the word "statistician," which has now a definite meaning, by the word "statist," which has another equally definite meaning. Does Dr Willcox wish to substitute "statics" for "statistics"?

It is possible to find difficulties in filling up all the columns except No. 4. For illustration, consider how column 2 should be filled in in the case of a cousin who was a "paying guest," or a relation who was a visitor; for column 3, is a divorced person single or a widower, and what of a woman who is doubtful whether her husband is lost at sea? Errors come from No. 3 because many unmarried people call themselves married.

It is well known that column 5 is wrongly filled in for two reasons—one, that elderly people often do not know their ages

accurately and enter them to the nearest round

Age.

number, so that the returns congregate at 40, 50, 60: the error thus arising is eliminated by tabulation in the groups 35-45, 45-55 years, &c., and for more minute tabulation the groups 3-7, 8-12, 13-17, &c., are suggested: the other is that many ladies habitually enter their ages too low; in this case also the Registrar-General is able to deduce nearly correct totals.

It is to be noticed that, since the ages stated are those "last birthday," the age will on the average be given six months too low, and, in fact, the ages given as 17, *e.g.*, should be scattered nearly uniformly over the months to the eighteenth year.

The most important criticisms of the census-schedule are to be made on columns 6-9. It will not be expedient here to go

into all the questions raised before the Committee on the Census as regards an industrial census.

Occupation.

While there can be little doubt that a thorough census of occupations would be best undertaken separately, and on somewhat different principles from the population census, it is certainly better, till opinion is ripe for so radical a change, to include in the present census the best questions we can as to occupations, than to omit them altogether in despair of accurate results.

The objects aimed at, which we must always keep in mind when criticising special questions, are two: to find the number employed in each trade and industry, that is, so to say, to form vertical divisions; and to find the number in each rank or grade of employment (labourer, artisan, employer, &c.) in horizontal divisions; so that the tabulation may give some such result as—

TEXTILE INDUSTRIES.

	Cotton.	Wool.	Linen.	Totals.
Employers -				
Managers -				
Overlookers -				
Spinners -				
Weavers -				
Labourers -				
Children -				
Totals -				

The necessary minimum of information would be given by such answers as

Legal—Solicitor—Managing clerk.
Mining—Coal—Hewer.
Metal-worker—Iron—Smith's striker.

Now the simple instruction, "State your occupation," would of course not lead to information of this sort. The coal-hewer would simply say miner; the clerk, managing clerk; the striker, very likely smith. To explain what is wanted and avoid mistakes, the question is not put on the face of the form at all, but the informant is referred to the back, half of which is devoted to instructions relating to this column. These are lucid, carefully picked out with capitals and italics, comprehensive, brief and to the point. No one who wishes to fill in the form rightly, and is sufficiently educated to understand simple instructions, can easily go wrong. Yet, as a matter of fact, these instructions are in very many cases neither read nor followed; and this fact is very important in connection with the general study of blank forms of inquiry. Forms issued to people uninterested in the object in view will generally be filled in with the least possible expenditure of time and intelligence. Hence two courses are open: to reduce

the question to the simplest possible form, and make the best of the result; or not to allow the informants to write in their own answers, but to take them *viva voce* by means of a teller, who has mastered the instructions, and has the necessary legal force behind him to compel information. The latter course entails time and expense.

The result of the present system of inquiry, combined with a faulty method of tabulation, which it to some extent makes necessary, is that we have no reliable census of occupations for the United Kingdom. The present figures break down both from faulty data and from insufficient tabulation directly we attempt to make any calculations depending on them.

An attempt has been made to correct to some extent our ignorance of the relative numbers of unskilled and skilled labourers, employers and employed, by columns The result of the new questions. 7, 8, and 9. The headings are not a model of clearness; there is not the ordinary imperative "state" or "write," nor is one told on the front of the form whether to write Yes or No or to make a mark in the appropriate column, nor is the distinction between the three headings a perfectly definite one; but still one is hardly prepared for the following statement in the report: *—

"In numerous instances, no cross at all was made; in many others, crosses were made in two or even all three columns, and, even when only one cross was made, there were often very strong reasons for believing that it has been made in the wrong column. Oftentimes this use of the wrong column can scarcely have been other than intentional; being dictated by the foolish but very common desire of persons to magnify the importance of their occupational condition. This desire must have led many subordinates to return themselves as employers rather than as employed, for it is only on this supposition that we can account for the otherwise unintelligible fact that, under several headings, there are actually, according to the returns, more employers than employed, more masters than men. . . . We hold [these returns] to be excessively untrustworthy, and shall make no use whatsoever of them in our remarks."

This attempt and its result are of the greatest importance to all who try to draw up forms of inquiry.

* *General Report on the Census of 1891*, p. 36 (C.—7222 of 1893).

Before leaving the subject, it should be mentioned in passing that we cannot deduce directly from our census the number of persons dependent on a particular trade for their living; that is to say, the number of employers, their families (not otherwise returned) and domestic servants, and the number of employes and their dependent families. This, the most important total for estimating the relative importance of different trades of the country, is not tabulated, though such tabulation has been found possible in other countries, and we are dependent on the estimates of statisticians for such totals.*

To see how the information given by the answers on the census schedule can be worked up into detailed specific numbers, it is only necessary to look at the diagram and table prefixed to each of the sections relating to special trades in Mr Booth's *Life and Labour of the People* (e.g., vol. v., p. 46).†

* See Booth in *Statistical Journal*, vol. xlix.

† See p. 78, *infra*.

SECTION 2.—THE WAGE CENSUS.

The main differences between the wage census, taken in 1886, and the general population census are—(1) That the filling up the forms in the wage census was voluntary; (2) that their correct filling up required a higher degree of intelligence and education. As before, we must consider first

The object. the object which the wage census was intended to fulfil: it was to describe the earnings of the

people of the United Kingdom, to compare the rates of wages trade by trade, and to find the relative numbers earning at each rate. What is the best quantity to measure with this object in view? As a preliminary question should we take the

The unit of time. day, week, or year as the unit of time? Clearly we shall not be able to compute weekly wages if we

only obtain daily, for the week's work varies from four to seven days in different occupations. The week's wage is a more definite quantity; but the simple comparison of weekly wages in different trades will be deceptive, because most trades are busier at one season of the year than at another, and in many the difference between season and season is very great; in any particular week, then, we may be comparing the best season of one industry with the worst of another. To avoid this error, and because we do not know how many full weeks' wages are obtained in a year, except in a few non-intermittent trades, it would seem best to take the year as unit; but the direct calculation of an individual's annual earnings is practically impossible. The employer is not acquainted with this sum, for in large establishments the hands are continually changing, and one man will be paid by two or more masters in the same year; and even in a factory with a nearly constant personnel, the weekly amounts paid to individuals are not in general so tabulated as to be easily summed, and the working out of the totals would require a prohibitive amount of clerical labour. If we turn to the workman, on the other hand, we shall find in the majority of cases that no accurate account has been kept of earnings through the year, and it would only be by careful individual examination, impracticable on any large scale, that