New statistical method to improve the quality of Official statistics in CAPMAS (Opportunities and Challenges)

By: Said Kamal Mohammed

Statistician

C.A.P.M.A.S

New statistical method to improve the quality of Official statistics in CAPMAS

Said Kamal Mohammed

Statistician

Central Agency for Public Mobilization and Statistics (CAPMAS)

www.capmas.gov.eg
Cairo – Egypt

Email: said_1984_33@yahoo.com

Abstract

Quality control is a process which is used to ensure a certain level of quality in a statistical product. The quality control unit in CAPMAS has an integrated role to improve the quality of official statistics and surveys, as well as evaluating research's methodology and improving the performance of employees. CAPMAS has been using several methods to improve the quality of official statistics, one of the most important method an experiment was conducted in corroboration with university of Michigan on data of labor force survey 2015 (the fourth quarter).The experiment proven that using cell phones (callback) for first time, In addition to quality control field work a highly effectiveness in quality of survey data.

The purpose of this experiment was to use mobile phones along with the field work through survey auditing and quality control processes. This national experiment of labor force survey (n=2348) addressed that 58% of the representative population (n=2348) have cellular phones. The experiment also demonstrated that some cities among the 58%- have relatively high response rate than others. Al Wadi El Jadid governorate has the highest response rate 85%. Luxor governorate has the lowest response rate 15%. The results showed that this new method Led to improve the basic indicators of labor force survey by 7%. The objective of this study views the most important results of using new method in data quality, which showed deep errors in the important fields of the labor force survey.

Keywords
Mobile phone, data quality, callback

1. Introduction:

Quality control is a process which is used to ensure a certain level of quality in a statistical product. The quality control unit in CAPMAS has an integrated role to improve the quality of official statistics and surveys, as well as evaluating surveys' methodology and improving the performance of employees.

Scientific sources, which focus on data quality, showed the importance of this topic, as the statistical data usually used in various areas, including planning, development, policy and decision-making.

There are standards and a dimension of quality has to be taken into account in the production process of statistical figure: (Relevance - Accuracy -Timeliness and Punctuality - Accessibility and clarity - Comparability – Coherence -Completeness).

As for quality control process, it's the tool which be used to secure a certain level of quality in the statistical outputs. CAPMAS is using several ways to improve the quality of official statistics, Call-back is the most important method which has been applied in cooperation with the Michigan University on the data of Labor Force Survey 2015 (the fourth quarter).

One of the important results of using the new method in improving the statistical data and other surveys other than labor force survey is to improve the quality of this data and surveys when applied for the first time, it has proven very effective at:

1 - Household Income, Expenditure, and Consumption Survey, HIECS.

2 - time-use survey

3- E-indicators of the Egyptian society.

ETC.

2. The objective of the study:

This study aims to show the importance of using call- back in quality control which showed errors in the important fields when use in the labor force survey and also display the results of the comparison between it and the use of field work on the quality of surveys.

3. Methodology and data sources:

A random sample was drawn with 20% of the households that were interviewed during the fourth quarter of 2015, consisting of 35 sampling areas at field and 116 areas in the (call-back) unit. Annex (1)

The methodology is based on using:

1. Field work: where **CAPMAS** quality control unit is drawing a random sample of five households (105 questionnaires) covering five sampling areas for (7 governorate).

2. Call-back: drawing a random sample of all the governorates (After excluding of the planned areas for the field work) with almost 40% of the interviewed households in the phase of collecting data (Labor Force survey) 134 sampling areas have been drawn with total of (2348) questionnaires representing all Egypt.

We should take into account non-recurrence of sampling areas according to the methodology of the study. So, when applying call-back method, the field examined sampling areas shall be excluded.

4. Finding:

The auditing stage showed that there are multiple errors in the data collection of the questionnaire, mainly the telephone number field where about 294 (13%) out of 2348 questionnaires not recorded fields.

The methodology of call-back depending on using telephone call. So, the following table shows some telephone number are as follows:

Table (1-1) shows the situation of telephone and the response of households

Calculated by researcher

S	**Telephone situation**	**Numbers**	**%**
1	Working number and households are respond	1187 out of 2054	58%
2	Wrong numbers.	169 out of 2054	8.2%
3	Out of service numbers.	74 out of 2054	3.6 %
4	No answer after repeated calls	297 out of 2054	14.5 %
5	not available/ switched off	252 out of 2054	12 %
6	Households refused to respond	75 out of 2054	3.7 %

1- After using call-back we found total errors conducted by the field and call-back controller:

Table (1-2) shows the total number and the percentage of main errors.

Total households	**total number of errors**	**%**	**Matching errors by respondent**		**Difference errors by respondent**	
			Number	**%**	**Number**	**%**
1187	**85**	**7%**	**32**	**37%**	**53**	**63%**

The above mention table shows that best governorates in completing phone numbers (100%) are (Damietta, So hag, Al Wadi El Jadid, and Red Sea). Where Luxor governorate is 10% and worst governorate in completing telephone number is Al-Sharqia with about 2%.

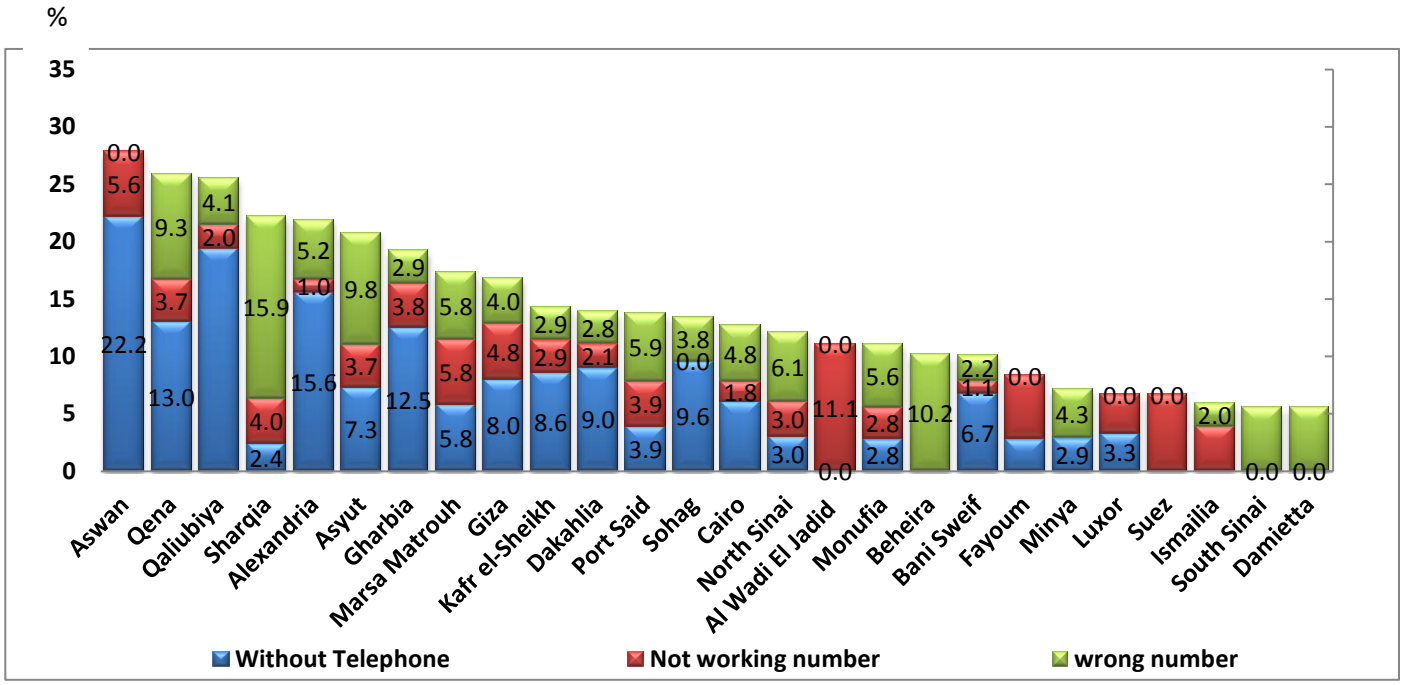
Also table (1-1) shows that the main errors in completing and recording data in the questionnaires as it's clear in Annex (2) are:

A. Record wrong phone number.

B. Record a not working phone number.

C. Do not Record any phone numbers.

Figure (1-1) shows the percentage of previous three errors according to governorate.



5. The main results:

First: Results of call-back:

The study results show the errors affecting the Labor force data:

A. Some questionnaires have more individuals than the household's members.

B. The Data Collector did not record 15 members in 9 households belongs to these households.

1- Errors affecting the employed and the unemployed:

The differences between the field researcher and Call-back researcher were (16) errors which equal (35%) of the total errors distributed as following:

A. The field researcher record 8 persons outside labor force who are already having a job.

B. The Data Collector record 1 individual as inside labor force who is having a job outside the labor force**.**

C. The Data Collector record 1 individual as unemployed who is outside the labor force**.**

D. The Data Collector record 5 individuals as unemployed persons which in fact they are working.

2- Errors affecting the Educational status and the difficulties and the profession:

A. Numbers of differences between The Data Collector and Callback controller have been recorded regarding some household members in basic data (educational status and difficulties). There are (14) differences which represents 33% of the total errors.

B. Numbers of differences between The Data Collector Callback researchers have been registered regarding profession nature of some household members. There is (1) difference which represents 2.3% of the total errors.

Second: Results of field work:

1. A number of 100 households have been interviewed, and we could not interview 5 households, (three of them in Cairo and two in Qaliubiya).

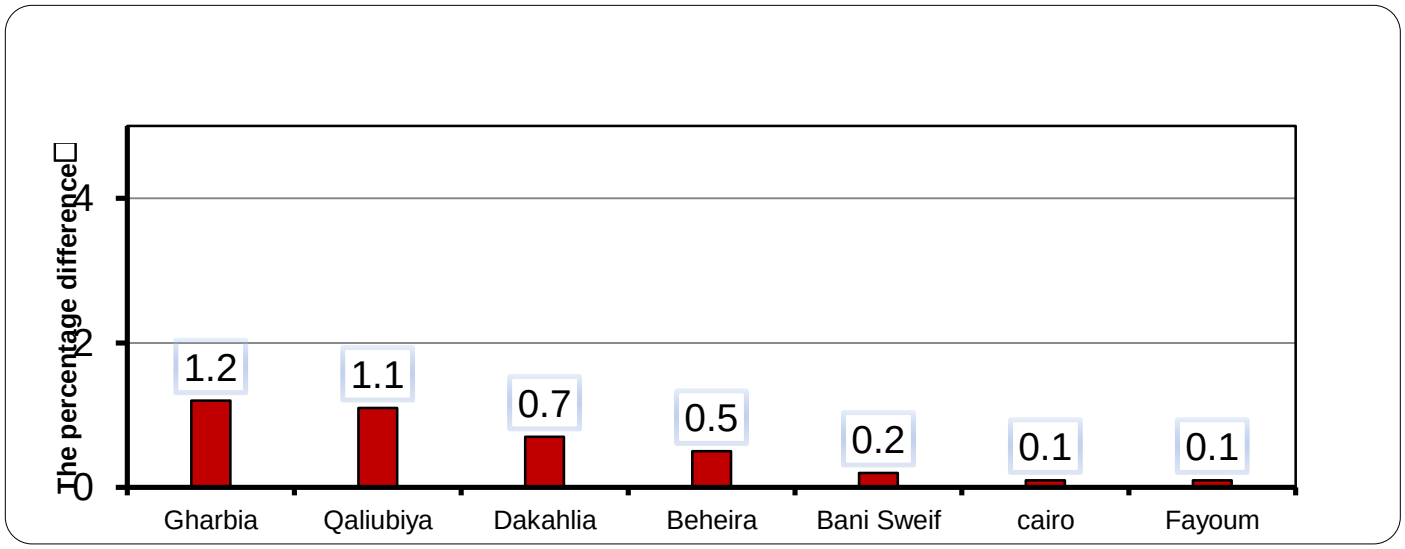
2. The most important errors are concentrated in recording the basic data, main labor force, education questions, difficulties, and the calculation of employed and unemployed.

As it's clear from next Table (annex 3):

Table (2-1)) shows the percentage of main errors according to governorate

Governorate	**Cairo**	**Qaliubiya**	**Beheira**	**Gharbia**	**Dakahlia**	**Fayoum**	**Beni Suef**
Percentage	**0.10 %**	**1.10 %**	**0.50 %**	**1.20%**	**0.70%**	**0.10 %**	**0.20%**

Figure (2-1) the following diagram shows the results for each governorate:



3. The highest difference percentage between the data collector and quality controller is 2% in (Qaliubiya, Dakahlia, Beni Suef).
4. The best governorates are (Cairo - Beheira - Fayoum).

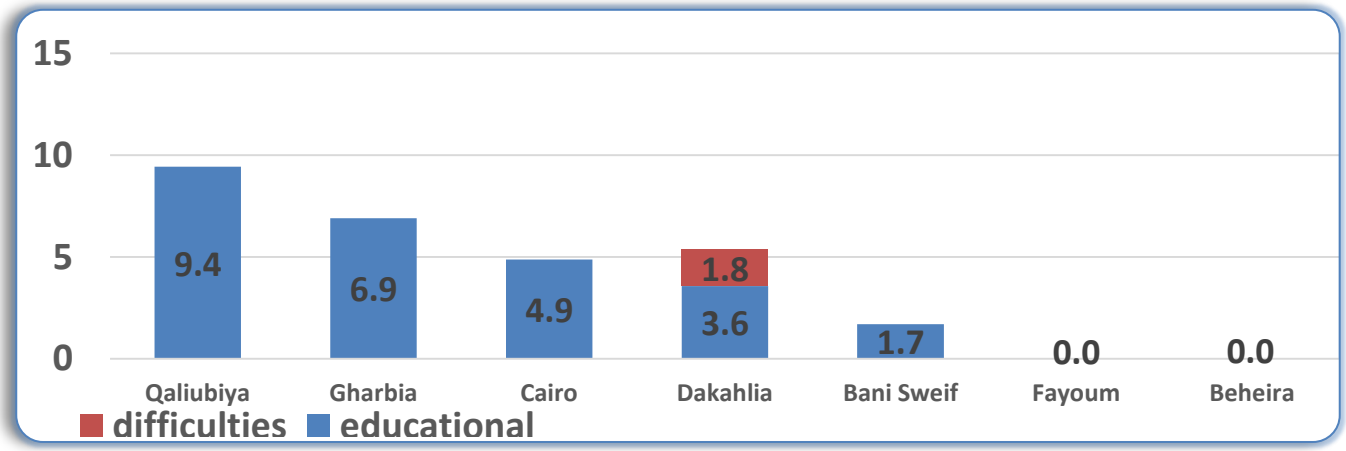
Error affecting basic data (educational status and difficulties):

A- Best governorates in education were (Beheira- Fayoum), highest difference percentage was in Qaliubiya.

B- Best governorates in difficulties were (Cairo – Qaliubiya - Beheira - Gharbia - Fayoum - Beni Suef), and highest difference percentage in difficulties was Dakahlia.

Table (2-2) :errors affecting basic data (educational status and difficulties):					
Governorate	Number of Households' members	Education		Difficulties	
		Number of individuals that have difference	%	Number of individuals that have difference	%
Cairo	**41**	**2**	**5**	**0**	**0**
Qaliubiya	**53**	**5**	**9**	**0**	**0**
Beheira	**63**	**0**	**0.0**	**0**	**0**
Gharbia	**58**	**4**	**7**	**0**	**0**
Dakahlia	**56**	**2**	**4**	**1**	**1.8**
Fayoum	**73**	**0**	**0.0**	**0**	**0**
Bani Sweif	**59**	**1**	**2**	**0**	**0**
Total	**403**	**14**	**3.5**	**1**	**1.8**

Figure (2-2) the difference parentage in difficulties& educational status.



SWOT analysis of using call-back:

Call-back unit is the new methodology used to monitor the quality of the data, It showed in the first use of good results in the verification of data quality, which will clear through the following analysis.

Strengths

1. High coverage and reliability.
2. Decreasing in the cost.
3. Decreasing in the time.
4. Coverage some dangerous areas.

Weaknesses

1. The difficulty of communication in some areas.
2. Lack of trained workers.
3. Lack of electronic equipment.

Opportunities

1. Availability of qualified senior campaign.
2. Coverage all the surveys and research.

Threats

1. Weak financial remuneration.
2. Slowdown in senior management decision-making.

6. **Recommendations:**

1. Statistical product process shall cover the quality control process within any statistical system on the two processes (the internal audit process - external audit process).

2. The internal audit process shall cover all stages of the research within the organization (desk review - data entry –automatic processing - Results).

3- The external audit process shall include accompanying field researchers and train them in the field.

4- Instructing the data collector about the necessity of filling the phone numbers in the questionnaires.

5. Continuous improvement and development: continuous improvement and development should be a permanent objective of the organization, employees and for the operations management system, products and services within the Quality Management System

References:

1. Abraham, B. and Ledolter, J. (1983). Statistical Methods for Forecasting, Wiley, New York, NY.
2. Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). Time Series Analysis, Forecasting and Control, 3rd ed. Prentice Hall, Englewood Clifs, NJ.
3. Duncan, A. J. (1986). Quality Control and Industrial Statistics, 5th ed., Irwin, Homewood, IL.
4. Hotelling, H. (1947). Multivariate Quality Control. In C. Eisenhart, M. W. Hastay, and W. A. Wallis, eds. Techniques of Statistical Analysis. New York: McGraw-Hill.

Annex(1)

Governorate	Number of sampling areas (call-back)	Number of sampling areas (In the field)	Number of sampling families (call-back)	Number of sampling families (In the field)
Cairo	21	3	345	30
Giza	9	0	154	0
Qaliubiya	8	3	144	35
Alexandria	7	0	115	0
Sharqia	8	0	144	0
Gharbia	7	3	126	45
Beheira	8	3	144	45
Monufia	6	0	107	0
Dakahlia	8	0	144	0
Kafr el-Sheikh	4	0	72	0
Damietta	2	0	36	0
Suez	2	0	36	0
Ismailia	2	0	36	0
Port Said	3	0	50	0
North Sinai	2	0	35	0
South Sinai	1	0	18	0
Marsa Matrouh	1	0	15	0
Fayoum	4	3	72	45
Bani Sweif	4	3	72	45
Minya	6	0	108	0
Asyut	5	0	90	0
Sohag	6	0	106	0
Qena	3	0	54	0
Luxor	3	0	54	0
Aswan	2	0	35	0
Al Wadi El Jadid	1	0	18	0
Red Sea	1	0	18	0
	134	18	2348	245

Annex(٢)

Results of telephone call according the number of households

Governorate	Number of Questionnaires (number of households)	Correct phone numbers			unconfirmed phone numbers			Researcher error						
								number of errors			errors percentage %			Total Percentage
		call-back	Refusal to respond by telephone	Percentage	No answer	Not Available/ switched off	Percentage	Without Telephone	Not working number	wrong number	Without Telephone	Not working number	wrong number	
Cairo	345	163	18	52.5	56	33	25.8	44	14	17	12.8	4.1	4.9	21.7
Giza	154	73	2	48.7	19	12	20.1	40	3	5	26.0	1.9	3.2	31.2
Qaliubiya	144	95	5	69.4	13	11	16.7	10	4	6	6.9	2.8	4.2	13.9
Alexandria	115	60	3	54.8	13	4	14.8	24	9	2	20.9	7.8	1.7	30.4
Sharqia	144	83	6	61.8	16	16	22.2	3	0	20	2.1	0.0	13.9	16.0
Gharbia	126	63	0	50.0	11	15	20.6	23	2	12	18.3	1.6	9.5	29.4
Beheira	144	74	1	52.1	25	14	27.1	17	3	10	11.8	2.1	6.9	20.8
Monufia	107	59	1	56.1	12	17	27.1	12	1	5	11.2	0.9	4.7	16.8
Dakahlia	144	77	3	55.6	14	19	22.9	21	1	9	14.6	0.7	6.3	21.5
-Kafr el Sheikh	72	41	2	59.7	10	6	22.2	6	3	4	8.3	4.2	5.6	18.1
Damietta	36	20	4	66.7	3	6	25.0	0	1	2	0.0	2.8	5.6	8.3
Suez	36	26	0	72.2	5	1	16.7	2	0	2	5.6	0.0	5.6	11.1
Ismailia	36	21	0	58.3	5	4	25.0	1	1	4	2.8	2.8	11.1	16.7
Port Said	50	33	2	70.0	8	1	18.0	4	0	2	8.0	0.0	4.0	12.0
North Sinai	35	16	0	45.7	8	9	48.6	1	0	1	2.9	0.0	2.9	5.7
South Sinai	18	11	1	66.7	1	1	11.1	1	2	1	5.6	11.1	5.6	22.2
Marsa Matrouh	15	5	3	53.3	3	0	20.0	1	0	3	6.7	0.0	20.0	26.7
Fayoum	72	36	2	52.8	10	9	26.4	13	1	1	18.1	1.4	1.4	20.8
Bani Sweif	72	41	0	56.9	8	9	23.6	11	0	3	15.3	0.0	4.2	19.4
Minya	108	40	5	41.7	16	15	28.7	22	2	8	20.4	1.9	7.4	29.6
Asyut	90	46	4	55.6	9	10	21.1	11	2	8	12.2	2.2	8.9	23.3
Sohag	106	39	3	39.6	15	19	32.1	0	12	18	0.0	11.3	17.0	28.3
Qena	54	20	5	46.3	5	8	24.1	7	2	7	13.0	3.7	13.0	29.6
Luxor	54	6	0	11.1	5	8	24.1	15	8	12	27.8	14.8	22.2	64.8
Aswan	35	15	2	48.6	6	2	22.9	5	0	5	14.3	0.0	14.3	28.6
Al Wadi El Jadid	18	15	0	83.3	0	1	5.6	0	0	2	0.0	0.0	11.1	11.1
Red Sea	18	9	3	66.7	1	2	16.7	0	2	1	0.0	11.1	5.6	16.7
total	2348	1187	75	53.75	297	252	23.38	294	73	170	12.5	3.1	7.2	22.87

Annex(3)

percentage between the field researcher and quality researcher

statement / governorates	The first section (some of the demographic characteristics of household members) %	The second section (working characteristics) %	the third Section (individual characteristics ofunemployed) %	the fourth Section (individualJobless characteristics) %	average %
Cairo	0.3	0.0	0.0	0.0	0.1
Qaliubiya	1.0	0.0	3.2	0.0	1.1
Beheira	0.1	0.2	1.8	0.0	0.5
Gharbia	1.1	0.0	2.3	1.4	1.2
Dakahlia	1.0	0.8	1.1	0.0	0.7
Fayoum	0.2	0.0	0.0	0.0	0.1
Bani Sweif	0.3	0.0	0.6	0.0	0.2

Annex(2)

Results of telephone call according the number of households

Governorate	Number of Questionnaires (number of households)	Correct phone numbers			unconfirmed phone numbers			Researcher error						
								number of errors			errors percentage ٪			Total Percentage
		callback	Refusal to respond by telephone	Percentage	No answer	Not Available/ switched off	Percentage	Without Telephone	Not working number	wrong number	Without Telephone	Not working number	wrong number	
Cairo	345	163	18	52.5	56	33	25.8	44	14	17	12.8	4.1	4.9	21.7
Giza	154	73	2	48.7	19	12	20.1	40	3	5	26.0	1.9	3.2	31.2
Qaliubiya	144	95	5	69.4	13	11	16.7	10	4	6	6.9	2.8	4.2	13.9
Alexandria	115	60	3	54.8	13	4	14.8	24	9	2	20.9	7.8	1.7	30.4
Sharqia	144	83	6	61.8	16	16	22.2	3	0	20	2.1	0.0	13.9	16.0
Gharbia	126	63	0	50.0	11	15	20.6	23	2	12	18.3	1.6	9.5	29.4
Beheira	144	74	1	52.1	25	14	27.1	17	3	10	11.8	2.1	6.9	20.8
Monufia	107	59	1	56.1	12	17	27.1	12	1	5	11.2	0.9	4.7	16.8
Dakahlia	144	77	3	55.6	14	19	22.9	21	1	9	14.6	0.7	6.3	21.5
-Kafr el Sheikh	72	41	2	59.7	10	6	22.2	6	3	4	8.3	4.2	5.6	18.1
Damietta	36	20	4	66.7	3	6	25.0	0	1	2	0.0	2.8	5.6	8.3
Suez	36	26	0	72.2	5	1	16.7	2	0	2	5.6	0.0	5.6	11.1
Ismailia	36	21	0	58.3	5	4	25.0	1	1	4	2.8	2.8	11.1	16.7
Port Said	50	33	2	70.0	8	1	18.0	4	0	2	8.0	0.0	4.0	12.0
North Sinai	35	16	0	45.7	8	9	48.6	1	0	1	2.9	0.0	2.9	5.7
South Sinai	18	11	1	66.7	1	1	11.1	1	2	1	5.6	11.1	5.6	22.2
Marsa Matrouh	15	5	3	53.3	3	0	20.0	1	0	3	6.7	0.0	20.0	26.7
Fayoum	72	36	2	52.8	10	9	26.4	13	1	1	18.1	1.4	1.4	20.8
Bani Sweif	72	41	0	56.9	8	9	23.6	11	0	3	15.3	0.0	4.2	19.4
Minya	108	40	5	41.7	16	15	28.7	22	2	8	20.4	1.9	7.4	29.6
Asyut	90	46	4	55.6	9	10	21.1	11	2	8	12.2	2.2	8.9	23.3
Sohag	106	39	3	39.6	15	19	32.1	0	12	18	0.0	11.3	17.0	28.3
Qena	54	20	5	46.3	5	8	24.1	7	2	7	13.0	3.7	13.0	29.6
Luxor	54	6	0	11.1	5	8	24.1	15	8	12	27.8	14.8	22.2	64.8
Aswan	35	15	2	48.6	6	2	22.9	5	0	5	14.3	0.0	14.3	28.6
Al Wadi El Jadid	18	15	0	83.3	0	1	5.6	0	0	2	0.0	0.0	11.1	11.1
Red Sea	18	9	3	66.7	1	2	16.7	0	2	1	0.0	11.1	5.6	16.7
total	2348	1187	75	53.75	297	252	23.38	294	73	170	12.5	3.1	7.2	22.87