Data Quality Assessment Framework (DQAF) Report

Analysis of qualitative aspects of the education statistical system in Mozambique

Draft Version 0.3

Marc BERNAL
Thierry LAIREZ
Mélanie SMUGA
# Table of Contents

1. **Background** ................................................. 3
2. **General analytical framework** ........................................... 3
3. **Pre-requisites of quality** .................................................. 4
   - 3.1. Legal and institutional environment ........................................ 4
   - 3.2. Resources: Resources are commensurate with needs of statistical programs ........................................ 5
   - 3.3. Quality awareness: Quality is a cornerstone of statistical work ........................................ 7
   - 3.4. Synthesis and score .................................................... 7
   - 3.5. Recommendations .................................................... 8
4. **Integrity: The principle of objectivity in the collection, processing, and dissemination of statistics is firmly adhered to** ........................................ 9
   - 4.1. Professionalism: Statistical policies and practices are guided by professional principles ........................................ 9
   - 4.2. Transparency: Statistical policies and practices are transparent ........................................ 9
   - 4.3. Ethical standards: Policies and practices are guided by ethical standards ........................................ 10
   - 4.4. Synthesis and score .................................................... 10
   - 4.5. Recommendations .................................................... 10
5. **Methodological soundness: The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices** ........................................ 11
   - 5.1. Concepts and definitions: Concepts and definitions used are in accord with standard statistical frameworks ........................................ 11
   - 5.2. Scope: The scope is in accord with internationally accepted standards, guidelines, or good practices ........................................ 12
   - 5.3. Classification / sectorization: Classification and sectorization systems are in accord with internationally accepted standards, guidelines, or good practices ........................................ 12
   - 5.4. Basis for recording: Data are recorded according to internationally accepted standards, guidelines, or good practices ........................................ 14
   - 5.5. Database structure .................................................... 17
   - 5.6. Synthesis and score .................................................... 20
   - 5.7. Recommendations .................................................... 21
6. **Accuracy and reliability: Source data and statistical techniques are sound and statistical outputs sufficiently portray reality** ........................................ 22
   - 6.1. Source data available provide an adequate basis to compile statistics ........................................ 22
   - 6.2. Statistical techniques: Statistical techniques employed conform to sound statistical procedures, and are documented ........................................ 23
   - 6.3. Assessment and validation of source data: Source data are regularly assessed and validated ........................................ 23
   - 6.4. Revision studies: Revisions, as a gauge of reliability, are tracked and mined for the information they may provide ........................................ 23
   - 6.5. Synthesis and score .................................................... 24
   - 6.6. Recommendations .................................................... 24
7. **Serviceability: Statistics are relevant, timely, consistent, and follow a predictable revisions policy** ........................................ 25
   - 7.1. Periodicity and timeliness: Periodicity and timeliness follow internationally accepted dissemination standards ........................................ 25
   - 7.2. Consistency: Statistics are consistent within a dataset and over time, and with other major data sets ........................................ 25
   - 7.3. Revision policy and practice: Data revisions follow a regular and publicized procedure ........................................ 33
   - 7.4. Synthesis and score .................................................... 33
   - 7.5. Recommendations .................................................... 34
8. **Accessibility: Data and metadata are easily available and assistance to users is adequate** ........................................ 35
   - 8.1. Data accessibility: Statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis ........................................ 35
   - 8.2. Metadata accessibility: Up-to-date and pertinent metadata are made available ........................................ 36
   - 8.3. Assistance with the users: Prompt knowledge support service is available ........................................ 36
   - 8.4. Synthesis and score .................................................... 36
   - 8.5. Recommendations .................................................... 37
9. **Conclusion and overall recommendations** ........................................ 37
10. **Annexes** .............................................................. 39
1. **Background**

1. Education Management Information Systems (EMIS) are one of the priorities of the African Union's action plan for the Second Decade of Education for Africa and the South African Development Community (SADC) education programme. In addition, the availability of reliable information at the continental, regional and national levels is necessary for policy planning, monitoring and evaluation, as well as decision-making, and should essentially be based on efficient reporting and well documented statistical information systems.

2. Many countries in Africa face several challenges such as poor quality data that are unreliable and difficult to validate, weak statistical systems which do not facilitate the collection, processing, analysis and dissemination of data, high staff turnover in Ministries of Education (particularly for key roles including statisticians and Information Technology (IT) staff), limited collaboration between the different data-producing agencies, and a lack of statistical awareness within the Ministries of Education.

3. The Strategic Plan for Education and Culture 2006 – 2010/11 (ref.: Plano Estratégico de Educação e Cultura 2006 – 2010/11 (PEEC)) is the key to all strategies that are and will be implemented for the development of the Mozambican education and culture sectors until 2011. To reflect and support the decentralization which is proposed in the PEEC, it is necessary to strengthen all the processes in order to monitor and report better at the provincial, district and school levels. To help in this matter, the PEEC also stipulates that there is a need to improve the EMIS and to implement a better use of Information and Communication Technology (ICT) tools. All the objectives / targets set for each sub-sector of education as well as related indicators have been oriented and defined by taking into account various elements such as the Millennium Development Goals (MDGs), the Action Plan for the Reduction of Absolute Poverty (ref.: Plano de Acção para a Redução da Pobreza Absoluta (PARPA)), and the Education for All / Fast Track Initiative (EFA / FTI).

4. The Ministry of Education and Culture (MoEC) of Mozambique has a Strategic Plan for the Statistics Sector (ref.: Plano Estratégico do Sector de Estatística (2008-2012)) for the period 2008-2012. This strategic plan comes within the framework of the Strategic Plan of the National Statistical System (ref.: Plano Estratégico do Sistema Estatístico Nacional (PESEN)), which is aligned with government policies to reduce poverty. One of its objectives is to collect and publish statistical information on the main sub-sectors of the national education system. Two data quality dimensions are also being put forward: increasing the coverage and the timeliness of data collected.

5. The United Nations Educational, Scientific and Cultural Organization (UNESCO) is engaged in proposing a methodology to evaluate the quality of data produced by education sub-sector. Currently, seven pilot countries in the SADC region – including Mozambique – have been selected for the purposes of establishing a Data Quality Assessment Framework (DQAF) in 2008 and 2009. The results of these findings are expected to be synthesized at the regional level with the possibility of extending the DQAF activities to other countries on the African continent.

2. **General analytical framework**

6. The generic DQAF evaluation has the principal objective to provide a flexible approach to ensure that a qualitative assessment of the statistics produced by the education sector in Mozambique is available. The tools and methodologies applied have been adapted from initial evaluations undertaken by the International Monetary Fund (IMF) and the World...
Bank, and further adapted by the UNESCO Institute for Statistics (UIS) to ensure a comprehensive evaluation of the quality of education statistics.

7. During the DQAF evaluation, a participatory and need-based approach has been adopted to conduct a series of interviews with national and sub-national agencies responsible for the production of statistics.

8. The evaluation framework has covered all the different steps included in the statistical business process model at the national and sub-national levels and has assessed the strengths and weaknesses of the available structures using the six DQAF dimensions: (i) pre-requisites of quality; (ii) integrity; (iii) methodological soundness; (iv) accuracy and reliability; (v) serviceability and; (vi) accessibility.

9. The report is organized around these six dimensions and their sub-components. For each one of them, a score has been assigned, as an indication only, whose sole purpose is to give the position of Mozambique in relation to an ideal international standard.

3. **Pre-requisites of quality**

10. Pre-requisites of quality are not as such a qualitative dimension, but the evaluation and understanding of the institutional context in which the statistical process exists is essential to the other dimensions. This dimension presents the context in which available statistical laws, including essential human and technical resources, impact other quality dimensions.

3.1. **Legal and institutional environment**

11. Following a national statistical legislation which appeared to be inadequate due to economic and social changes that have occurred in Mozambique, a National Statistical System (NSS) has been created in the country (ref.: Article 1 of Lei n.º 7 / 96 de 5 de Julho). The NSS mainly aims to (i) ensure the collection, processing, analysis and dissemination of statistical information necessary for the country to guide its socio-economic development in its different levels; (ii) optimize the use of human, technical, financial and material resources in the production of official statistics and the development of national statistical activities, avoiding duplication of efforts and consequent wastage of resources and; (iii) promote the interests of the population, public and private institutions, and enterprises in the national statistical activity, and to promote their participation and collaboration in the collection of relevant, reliable and timely statistical data.

12. The official statistical activity of the NSS is based on the following principles: (i) statistical authority; (ii) confidentiality; (iii) technical autonomy; (iv) impartiality; (v) transparency; (vi) reliability; (vii) relevance and; (viii) statistical coordination.

13. The principle of the statistical authority is the power conferred on the National Statistics Institute (NSI) in carrying out surveys which respect the deadlines in order to publish timely data and make all necessary arrangements for the production of statistics.

14. The NSI delegates to the MoEC its authority in terms of collection, processing, analysis and dissemination of administrative data collected at the school level. The NSI is in charge of household surveys and population censuses which include items related to education. The MoEC is involved in the design managed by the NSI. However, the MoEC is not involved in the design of household surveys carried out by other agencies such as the United Nations Bank.
Children's Fund (UNICEF) and the World Bank, even if such surveys include education components.

15. There are no systematical procedures that have been put in place by the MoEC to compare the data that the Ministry collects with other sources of data (e.g. data from household surveys and population censuses). However, administrative data provided by the MoEC are checked by the NSI (i.e. time series verification is done at the national and provincial levels) but no imputation methods are applied.

16. The Ministry of Education (MINED)\(^1\) has elaborated a document named Policies and Procedures Management of Information Systems / Information Technologies (ref: Políticas e Procedimentos de Gestão de SI / TI) which is very well defined and complete regarding what is put in place to preserve the confidentiality of individual respondent's data. Formal authorizations from the MINED or the Department of Communication and Information Technologies (DTIC) are required for any external data requests or hardware/software installation for example. Violation of these policies and procedures will result in disciplinary action.

17. In reference to the Lei n.º 7 / 96 de 5 de Julho mentioned previously, respondents are clearly informed about their rights and obligations with regard to the provision of information as well as the penalties / fines associated with the refusal to provide information and / or the provision of inaccurate information. In addition, it is clearly stated in Article 7 on statistical confidentiality that the information collected is only for statistical purposes and it is the responsibility of the NSS to protect individual data. However, at the MoEC level, there are no penalties / fines issued to anyone who does not provide data or provides inaccurate data and no system as such is foreseen in the future.

18. All the necessary measures have been put in place to restrict the access to individual data.

3.2. Resources: Resources are commensurate with needs of statistical programs

19. The Directorate of Planning and Cooperation (DIPLAC) of the MoEC is divided into four departments: (i) statistics; (ii) cooperation; (iii) planning and; (iv) school buildings and equipment. However, there is an imbalance regarding the number of staff working in these different sections.

20. For the Statistics Department, there are only four people working in this area: two persons in charge of basic education, technical/vocational and adult education, one person responsible for higher education, and one director/head of unit, while there is twice as many staff working for the Planning Department. At the end of 2008, DIPLAC planned to recruit an additional public servant to help in statistics matters and two expatriates for short term contracts to complement and capacitate the team on financial issues. In general, staff involved in statistical production has university degrees in education planning and has been following different short trainings in statistics and planning.

21. DTIC, like DIPLAC, is also heavily involved in the production of education statistics. Regarding IT staff, more specifically for hardware maintenance, there is a sufficient number of staff at the national level but this number is particularly inadequate at the provincial level.

\(^1\) Former name given to the Ministry of Education before it merged with the Ministry of Culture in 2005 (i.e. MoEC).
22. In most of the provinces, four public servants are working with statistics: one statistician and three clerks. Capacities in terms of data processing are generally fairly high, with an average level in using Microsoft Office tools such as Excel, Word and Access, but with low capacities in education performance analysis.

23. There are no efforts made to ensure the retention at any point of time of a core contingent of trained staff. Therefore, there is a high staff turnover.

24. At the MoEC, there are approximately 400 workstations which are connected to a high speed Local Area Network (LAN). Servers are well administrated and allow a secure backup policy. Two high speed Internet connections are available. In most of the provinces, four computers connected to a small LAN are available for statistical purposes.

25. There are sufficient resources allocated for hardware but not enough for software. Different information systems are being developed in parallel, such as the General Education Statistical Information System (GESIS) which is the only one to be fully operational. It has been decentralized to the provinces since 1994 and the current version is in use since 2004. This version was developed by external expertise and was afterwards handed over to IT staffs which have left their jobs since then. Some changes have been made by the new IT head that had to work without any available technical documentation.

26. The data capture interface of the GESIS is quite user-friendly with good correlation to the questionnaire formats and with some error tracking facilities / data entry checks. Reporting functionalities are provided by using pivot tables produced by Excel accessing the database via in-built queries.

27. Pivot tables would benefit from being structured under a unique Excel folder in order to improve statistical abstract automation. The user-friendliness of this reporting facility also suffers from some weaknesses in the database structure, which results in its accessibility being limited to statisticians.

28. No relations between tables and, consequently, no relational integrity hindering the database consistency. Fields naming is not compliant with standards: some tables are empty and some nomenclature tables are absent (e.g. grades, years, etc.).

29. There is no technical documentation and only an installation guide is available for the provinces.

30. The higher education information system is made up of a list of tables under Access, but it is not automated and many data are duplicated. Reports are produced using SPSS. The whole process requires at least four months to process data from 23 institutions.

31. The financial resources for compiling statistics are generally adequate to execute the activities that have been planned. However, many projects are not included in this plan because funds are insufficient to cover all of them.

32. At the moment, the codification system used by the MoEC and the NSI is not yet shared between both organizations. Nevertheless, a project is planned to take effect in 2010 in order to resolve this situation.

33. Education questionnaires have been designed internally by a joint effort of various directors belonging to different sections. However, no consultation with the users/stakeholders has been made.
34. There are no specific guidelines which exist to track discrepancies during data compilation procedures. At the national level, data processing and validation are done manually (i.e. through pivot table reports created from the database) but this procedure is not performed in a systematic way.

35. Periodic reviews of working processes are not undertaken to ensure that they are improved upon.

3.3. Quality awareness: Quality is a cornerstone of statistical work

36. The director of planning is heavily involved in key working groups dealing with data quality such as the national council of statistics. A quality / steering group has also been set up. However, the processes for implementing data quality controls are still not systematic and based on informal communication.

37. At the national level, consistency checks and time series verifications are performed. This process is automated but not systematized.

38. NSI conducts public opinion surveys but they are at the general level and therefore, not specific for education. No surveys as such are performed regarding education administrative data. However, for data users who consult the MoEC website, there is a mechanism for sending general inquiries (such as feedback on data quality issues) via e-mail which is available to users.

3.4. Synthesis and score

39. Based on an assessment of all the DQAF sub-dimensions, a global score of 60% has been assigned for the pre-requisites of quality dimension.

Figure 1: Results of pre-requisites of quality
3.5. Recommendations

40. Educate / sensitize the different partners on the necessary involvement of the MoEC in any survey design which includes education components.

41. The quality / steering group that has been put in place by the MoEC must set up a procedure manual including penalties for violation of the published calendar of operations. The National Strategy for Development of Statistics (NSDS) should strengthen the relationship between the MoEC and the NSI.

42. The quality / steering group should elaborate quality norms and standards (e.g. confidence level, imputation methods for missing data, etc.).

43. Ensure that the reports of technical meetings are easily available in order to follow the recommendations arising from these meetings.

44. There should be a balance in the recruitment process between statisticians and education planners and it is important to establish parity for each education sub-sector. Qualified IT developers should be recruited to design and develop IT tools. There should be a development of skills regarding maintenance at the provincial level.

45. Recruit additional staff at the national level. In general, staff is overburdened and does not undertake more in-depth research on planning and education sector performance analysis. Furthermore, and consequently, capacities are relatively low in these two domains.

46. Build capacities on sector performance analysis at the national level and extend this at the regional level. This is to be done following a “learning by doing” exercise including identification of financial data sources and construction of an “in-house constructed” simulation model.

47. Establish a continuing education plan to address staff needs.

48. Establish systematized procedures and access to these. Even if the development of ICT is professionally managed compared to many other countries in which the IT function is found within the Department of Policy and Planning (DPP), there is a resultant gap that exists in terms of accessibility to DTIC by DIPLAC.

49. If possible, upgrade the salaries.

50. Develop technical and functional documentation of computer applications in order to enhance the work and to consolidate achievements.

51. There should be the implementation of a training plan which would be oriented on the management of databases. In addition, measures have to be taken to ensure their sustainability and make sure that the system covers all the education sub-sectors.

52. Periodic reviews of working processes should become more systematic and organized.

53. The quality / steering group must dynamize its activities. It should put in place a functional database containing raw data and must produce a set of guidelines describing data validation procedures.
4. **Integrity: The principle of objectivity in the collection, processing, and dissemination of statistics is firmly adhered to**

54. Integrity is a quality dimension which is based on the principal of objectivity or impartiality in the collection, processing, analysis and dissemination of data. This dimension includes as well the institutionalized provisions which guarantee or improve the professionalism of the different actors in their practices.

4.1. **Professionalism: Statistical policies and practices are guided by professional principles**

55. There is no law or other formal provision which addresses the general need for the professional independence of the MoEC or prohibits interference from others, including other government agencies in the compilation and / or dissemination of statistical information. However, the NSI does not interfere in the work done by the MoEC.

56. No strict rules have been established to specify the minimum required qualifications / level of study to work at the DPP. Some information exists at the Human Resources (HR) Department, but nothing is available at the DPP. Nevertheless, there is a strong desire to systematize the recruitment based on career criteria (terms of references for instance) but it seems that there are only few candidates who fit with this specific profile.

57. Professionalism is encouraged by the participation of various international meetings such as those organized by the UIS, the National Education Statistical Information Systems (NESIS), etc.

58. The MoEC does not use any data resulting from household surveys produced in the country. Education data collected by the NSI are completely ignored by the MoEC, thus depriving itself of a very important comparison tool for data quality improvement.

59. The MoEC does not comment publicly on erroneous interpretations or misuses of its statistical data in the media and in other fora, or seek to prevent this kind of situation by providing explanatory materials and briefings, and by following closely the press and other media.

4.2. **Transparency: Statistical policies and practices are transparent**

60. The terms and conditions under which statistics are collected, processed and disseminated are not available to the public in any publication and / or website.

61. Access to statistics prior to release is not made public in terms of who has access, and at what point of the compilation process access is given. Even if the head of statistics and the director of planning have working sessions about the coherence of data before they decide to publish them, nothing is made available to the public.

62. Statistical products are clearly identified by the MoEC name and logo. However, the MoEC does not request attribution when its statistics are used or reproduced.

63. Users of statistics are not made aware of major changes in methodology, source data, and statistical techniques. For instance, when new population data are introduced in the
calculation of indicators, no information regarding this change is made available to data users.

4.3. Ethical standards: Policies and practices are guided by ethical standards

64. In the framework of the status of public servant, there are guidelines outlining correct staff behavior.

65. Management acknowledges its status as role model and is vigilant in following the guidelines. However, staff members are not reminded periodically of these guidelines.

4.4. Synthesis and score

66. Based on an assessment of all the DQAF sub-dimensions, a global score of 30% has been assigned for the integrity dimension.

![Figure 2: Results of integrity](image)

4.5. Recommendations

67. Promote the emergence of an independent statistical culture.

68. Cultivate the assessment and regular use of secondary sources of data (i.e. education data extracted from household surveys).

69. Develop a scientific approach regarding statistics by promoting professionalism among staff which should be encouraged to present their reasoning for the choice of methodologies.
70. There should be a public presentation of MoEC data before its final publication.

71. Develop an educational communication policy towards the public and include it in the agenda of the management. Internal comments should be done regarding the articles on education published in the national press. In addition, the reading and the dissemination of international articles should be encouraged among staff members.

72. Implement the material about the terms and conditions under which official statistics are compiled and disseminated. Establish and make available a statistical business process model to ensure internal traceability and the reverse engineering of the model.

73. Put in place an internal and external validation process prior to data release.

74. Even if a framework of the status of public servant exists, there should be a section which must emphasize on the specificities related to the status of a “national statistician”.

75. Sensitize staff regarding ethical standards / guidelines with sessions of continuing education.

5. Methodological soundness: The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices

76. This dimension of quality is based on the principle that the production of statistics should be established on a solid methodological basis with the application of well-defined standards, principles and good practices which are recognized internationally.

77. Therefore, this dimension is specific for each data set since different methodologies are applied depending on the specificities of each one of them.

5.1. Concepts and definitions: Concepts and definitions used are in accord with standard statistical frameworks

78. Production of statistics in Mozambique is based on internationally compliant methodological practices and principles.

79. Concepts and definitions follow methodologies used by the UIS which are described in the International Standard Classification of Education – 1997 version (ISCED97). This classification system provides a framework for the comprehensive statistical description of national educational systems and a methodology that translates national educational programmes into internationally comparable levels of education.

80. A complete ISCED97 mapping\(^2\) has been produced for Mozambique and is available on the UIS website\(^3\). Apart from the correspondence of national education programmes with the different ISCED97 levels, it provides information such as the programme orientation, the theoretical entrance age, the theoretical duration, etc.

\(^{2}\) Please refer to Annex 2.
### Name of the national education programme | ISCED97 level | Programme orientation | Theoretical entrance age | Theoretical duration
--- | --- | --- | --- | ---
Ensino pré-escolar | Pre-primary education | Not applicable | 3 | 3
Ensino primário do 1º grau (EPI) | Primary education | Not applicable | 6 | 7
Ensino primário do 2º grau (EP2) | | | | |
Ensino técnico elementar (ETE) | | | | |
Ensino secundário geral do 1º ciclo (ESG1) | Lower secondary education | General | 13 | 3
Ensino técnico profissional básico (ETB) | Vocational | | | |
Ensino secundário geral do 2º ciclo (ESG2) | Upper secondary education | Vocational | 16 | 2
Ensino técnico profissional médio (ETM) | General | | | |
Formação de professores (nível médio) | Post-secondary non-tertiary education | General | 18 | 1
Bachelor | Tertiary education (first stage) | First degree | 18 | 7-5
Masters | | Second degree | 23+ | 2
Licentiate | Tertiary education (second stage) | | | |
Doctorate | | | 25+ | 3-5

81. Even if the concept of part-time teaching staff is only applicable at the tertiary education level, there is no conversion factor for calculation of full-time equivalent.4

82. There is even no measure regarding the number of hours spent teaching per year and per full-time teacher.

83. At the primary level, there is no policy on promotion and repetition since education at this level is compulsory. At the secondary level, the law has been established that repetition is limited at two times. Since 2004, the MoEC introduced new curriculum where there are automatic promotions.

### 5.2. Scope: The scope is in accord with internationally accepted standards, guidelines, or good practices

84. In general, there are sufficient data for measuring the education system, in regard of its structure and performance (i.e. access, enrolment, progression, completion and student learning achievements). However, the main problem is the lack of quality data on finances, especially on human and financial resources invested in education.

85. The scope of statistics is adequate in terms of other relevant variables for analytical purposes, such as geographical boundaries (i.e. data are collected through the schools belonging to the 128 districts of the 10 Mozambican provinces) and socioeconomic groups (i.e. male / female, public / private, trained / untrained and full-time / part-time).

### 5.3. Classification / sectorization: Classification and sectorization systems are in accord with internationally accepted standards, guidelines, or good practices

86. The classification used by the MoEC is consistent with the ISCED97 which allows international comparability.

87. The MoEC does not publish an instruction manual to support users on how to use the classification and sectorization systems.

---

4 The full-time equivalence of part-time teachers is determined by calculating the ratio of their hours worked to the statutory hours worked by a full-time teacher during the school year.
88. The latest data submission from the MoEC to the UIS in regard to statistics of pre-primary, primary, secondary and post-secondary non-tertiary education is for the academic year 2008 (ref.: Questionnaire UIS/E/2009A). The data coverage is not fully complete since data for pre-primary education are reported as missing (i.e. “m”) due to the fact that it is the Ministry of Women and Social Affairs which is totally in charge of this sub-sector of education.

Enrolment by age (public and private)

<table>
<thead>
<tr>
<th>Age</th>
<th>All children enrolled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-primary education programmes</td>
</tr>
<tr>
<td></td>
<td>Both sexes</td>
</tr>
<tr>
<td>&lt; 3</td>
<td>m</td>
</tr>
<tr>
<td>3</td>
<td>m</td>
</tr>
<tr>
<td>4</td>
<td>m</td>
</tr>
<tr>
<td>5</td>
<td>m</td>
</tr>
<tr>
<td>6</td>
<td>m</td>
</tr>
<tr>
<td>7</td>
<td>m</td>
</tr>
<tr>
<td>&gt; 7</td>
<td>m</td>
</tr>
<tr>
<td>Age unknown</td>
<td>m</td>
</tr>
<tr>
<td>TOTAL</td>
<td>m</td>
</tr>
</tbody>
</table>

of which in:

<table>
<thead>
<tr>
<th></th>
<th>Public institutions</th>
<th>Private institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m</td>
<td>m</td>
</tr>
</tbody>
</table>

89. Regarding statistics of tertiary education, the latest data submission from the MoEC to the UIS is for the academic year 2007 (ref: Questionnaire UIS/E/2008C). One of the issue encountered, as mentioned previously, is the impossibility of providing data for full-time equivalent number of teachers. As shown in the table below, data have been reported as missing (i.e. "m"). However, it would be interesting to have a measure of this concept to allow international comparisons.

Teaching staff by type of programme (public and private)^5

<table>
<thead>
<tr>
<th>ISCED97 level</th>
<th>Type of Programme</th>
<th>Full and part-time</th>
<th>Part-time only</th>
<th>Full-time equivalent number of teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Both sexes</td>
<td>Female</td>
<td>Both sexes</td>
</tr>
<tr>
<td>5A</td>
<td>First stage (leading to entry into advanced research programmes)</td>
<td>3292</td>
<td>773</td>
<td>1764</td>
</tr>
<tr>
<td>5B</td>
<td>First stage (not leading to entry into advanced research programmes)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>6</td>
<td>Second stage (leading to an advanced research qualification)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>5+6</td>
<td>TOTAL</td>
<td>3292</td>
<td>773</td>
<td>1764</td>
</tr>
<tr>
<td></td>
<td>of which in:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5+6</td>
<td>public institutions</td>
<td>2210</td>
<td>695</td>
<td>901</td>
</tr>
<tr>
<td>5+6</td>
<td>Government dependent private institutions</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>5+6</td>
<td>Independent private institutions</td>
<td>1082</td>
<td>173</td>
<td>863</td>
</tr>
</tbody>
</table>

^5 Please note that "a" (i.e. category is not applicable) has been reported for ISCED97 level 6 in the UIS questionnaire while such a programme has been reported in the ISCED97 mapping. In fact, there is only one programme of doctorate that exists in the country (i.e. linguistic). However, not more than five students have subscribed to this programme. Most of the Mozambicans at the doctorate level are usually linked with universities outside the country.
5.4. **Basis for recording: Data are recorded according to internationally accepted standards, guidelines, or good practices**

90. All questionnaires used to collect education data are in accordance with international standards.

91. The MoEC is responsible for almost the entire education sector from primary to tertiary education. Non-formal education programme is also run by the MoEC including primary schooling opportunities provided to out-of-school children as well as functional adult literacy programme.

92. In Mozambique, the academic year begins at the end of January and ends in November. The data collection for enrolment is done on March 3rd and data for graduates are collected in December. The process of data collection is very well organized (not many countries in Africa meet such deadlines and conduct two rounds in a year). However, decentralization does not extend to the district level despite the fact that there is a strong demand for it.

93. Data collection instruments are very well structured, with an easy and readable layout, well written footnotes, explanations and advices for data control verification, as well as a clear timeframe for meeting deadlines.

94. From ISCED97 level 1 to 4, data collected are detailed and disaggregated at various levels. Data are generally broken down by sex, age, grade, public/private institutions, level of orientation (i.e. general and technical/vocational), etc. There are no data collected on the number of part-time pupils as well as the number of part-time and full-time equivalent teachers since this concept is simply not applicable for these sub-sectors of education.

95. The rural/urban component is not systematically considered in the data collection instruments. This information could be provided but it needs to be mapped with the NSI classification and introduced in the MoEC database.

96. Questionnaires are not divided by each sub-sector of education. For instance, there are questionnaires collecting data only for first degree of primary education (EP1), questionnaires which will cover second degree of primary education (EP2) in addition to first cycle of general secondary education (ESG1), questionnaires collecting data for second cycle of general secondary education (ESG2), and finally, questionnaires covering technical and vocational programmes.

97. For the questionnaires collecting data for second degree of primary education (EP2) and first cycle of general secondary education (ESG1), as well as for the questionnaires covering second cycle of general secondary education (ESG2), there are two kinds of questionnaires: one for the day shift (*turno diurno*) and another one for the night shift (*turno nocturno*). Differences between both kinds in terms of data collected are enumerated in the tables that follow.

98. The type of information collected for primary education (ISCED97 level 1) is the following:
### Topic | Type of information collected through the questionnaires | First degree of primary education (EP1) | Second degree of primary education (EP2) | Day shift | Night shift |
--- | --- | --- | --- | --- | --- |
#### Pupils | Number of pupils, repeaters, turmas, and alunos internos by sex, age, and grade | x | x | x |
#### Pupils | New entrants to Grade 1 which have 6 years old by sex | x |
#### Pupils | Number of pupils with books by discipline and grade | x | x |
#### Pupils | Number of orphans (mother, father and both sexes) by sex | x | x |
#### Pupils | Number of pupils at the end of the school year, of which passed the final exam, and by the final average note by discipline, sex, and grade | x | x | x |
#### Pupils | Number of pupils by cause of withdrawal, sex, and grade | x | x |
#### Pupils | Number of pupils transferred and which arrived after the survey of March 3rd by sex and grade | x | x |
#### Pupils | Number of turmas by shift (morning, afternoon) and grade | x |
#### Pupils | Number of external pupils who applied for examination by sex | x |
#### Teachers | Number of teachers by pedagogical qualification and sex | x | x | x |
#### Teachers | Number of teachers by number of shifts they teach | x |
#### Teachers | Time (entry, exit) and number of pupils by shift | x |
#### Teachers | Number of teachers by grade | x | x |
#### Teachers | Number of teachers with manuals by discipline and grade | x | x |
#### Teachers | Number of teachers that left during the school year by reason (death, disease or resignation/abandonment) and sex | x | x |
#### Teachers | Number of teachers at the end of the school year (with or without pedagogical training) by sex | x | x | x |
#### Teachers | Number of teachers by discipline and sex | x | x |
#### Others | Number of classrooms by type of construction | x | x |
#### Others | Number of employees, other than teachers by sex | x | x |
#### Others | Number of classrooms in the school | x |

**Turmas:** Groups of students taught, in general, by the same teacher at the same time.

**Alunos internos:** Students living in school and who are in charge of the administrative structures of education.

99. The type of information collected for secondary education (ISCED97 levels 2 and 3) is the following:

### Topic | Type of information collected through the questionnaires | First cycle of general secondary education (ESG1) | Second cycle of general secondary education (ESG2) | Day shift | Night shift |
--- | --- | --- | --- | --- | --- |
#### Pupils | Number of pupils, repeaters, turmas, and alunos internos by sex, age, and grade | x | x | x | x | x |
#### Pupils | Number of pupils with books by discipline and grade | x |
#### Pupils | Number of orphans (mother, father and both sexes) by sex | x |
#### Pupils | Number of pupils at the end of the school year, of which passed the final exam, and by the final average note by discipline, sex, and grade | x | x | x | x |
#### Pupils | Number of pupils by cause of withdrawal, sex, and grade | x |
#### Pupils | Number of pupils transferred and which arrived after the survey of March 3rd by sex and grade | x |
#### Pupils | Number of turmas by shift (morning, afternoon) and grade | x | x |
#### Pupils | Number of external pupils who applied for examination by sex | x | x | x | x |
#### Pupils | Number of pupils attending Grade 10 by section and sex | x | x |
#### Teachers | Number of teachers by pedagogical qualification and sex | x | x | x | x | x |
#### Teachers | Number of teachers by grade | x |
#### Teachers | Number of teachers with manuals by discipline and grade | x |
#### Teachers | Number of teachers that left during the school year by reason (death, disease or resignation/abandonment) and sex | x | x |
#### Teachers | Number of teachers at the end of the school year (with or without pedagogical training) by sex | x | x | x | x | x |
#### Teachers | Number of teachers by discipline and sex | x | x | x | x | x |
#### Others | Number of classrooms by type of construction | x |
#### Others | Number of employees, other than teachers by sex | x | x | x | x | x |
#### Others | Number of classrooms in the school | x | x | x |
#### Others | Number of residential facilities | x |

**Turmas:** Groups of students taught, in general, by the same teacher at the same time.

**Alunos internos:** Students living in school and who are in charge of the administrative structures of education.

100. The type of information collected for technical / vocational education is the following:
<table>
<thead>
<tr>
<th>Topic</th>
<th>Type of information collected through the questionnaires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupils</td>
<td>Number of pupils, repeaters, turmas, and alunos internos by sex, age, and grade</td>
</tr>
<tr>
<td></td>
<td>Number of pupils at the end of the school year, of which passed the final exam, and by the final average note by discipline, sex, and grade</td>
</tr>
<tr>
<td>Teachers</td>
<td>Number of teachers by nationality, shift (morning, evening), discipline, pedagogical training, and sex</td>
</tr>
<tr>
<td></td>
<td>Number of teachers that left during the school year by reason (death, disease or resignation/abandonment) and sex</td>
</tr>
<tr>
<td>Others</td>
<td>Number of employees, other than teachers by sex</td>
</tr>
<tr>
<td></td>
<td>Number of classrooms in the school</td>
</tr>
<tr>
<td></td>
<td>Number of residential facilities</td>
</tr>
</tbody>
</table>

Turmas: Groups of students taught, in general, by the same teacher at the same time.
Alunos internos: Students living in school and who are in charge of the administrative structures of education.

101. The type of information collected for literacy and adult education is the following:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Type of information collected through the questionnaires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>Number of students, repeaters, and turmas by sex, stage of learning, and name of center</td>
</tr>
<tr>
<td></td>
<td>Number of students with books by discipline and stage of learning</td>
</tr>
<tr>
<td></td>
<td>Number of students which have less than 14 years old by sex and stage of learning</td>
</tr>
<tr>
<td></td>
<td>Number of students by sex and language they learn (Portuguese, other language)</td>
</tr>
<tr>
<td></td>
<td>Number of students at the end of the school year and which were approved by sex, stage of learning, and name of center</td>
</tr>
<tr>
<td>Teachers</td>
<td>Number of teachers / tutors by pedagogical qualification (with or without) and sex</td>
</tr>
<tr>
<td>Others</td>
<td>Responsible for the preparation of the education programmes (MoEC, the provider)</td>
</tr>
<tr>
<td></td>
<td>Duration of the programme if it has been developed by the education provider</td>
</tr>
<tr>
<td></td>
<td>Other type of activities developed other than read, write, and count</td>
</tr>
<tr>
<td></td>
<td>Number of classrooms in the school</td>
</tr>
<tr>
<td></td>
<td>Number of outdoor classrooms</td>
</tr>
</tbody>
</table>

Turmas: Groups of students taught, in general, by the same teacher at the same time.

102. The type of information collected for tertiary education is the following:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Type of information collected through the questionnaires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>Functioning courses description (ISCED, city, prepared degree, duration)</td>
</tr>
<tr>
<td></td>
<td>Enrolment / course</td>
</tr>
<tr>
<td></td>
<td>Entrants</td>
</tr>
<tr>
<td></td>
<td>Enrolment - entrants - graduates / citizenship</td>
</tr>
<tr>
<td></td>
<td>Enrolment - entrants - graduates / region of school origin</td>
</tr>
<tr>
<td></td>
<td>Enrolment - entrants - graduates / region of birth origin</td>
</tr>
<tr>
<td></td>
<td>Graduates / course / 1st year of registration</td>
</tr>
<tr>
<td></td>
<td>Students registered / course</td>
</tr>
<tr>
<td></td>
<td>New entrants / course / degree</td>
</tr>
<tr>
<td>Teachers</td>
<td>Teachers / type of contract (full-time, part-time) / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Full-time teachers / citizenship / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Part-time teachers / citizenship / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Teachers / type of contractual relation / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Teachers / age / type of contract (full-time, part-time) / citizenship</td>
</tr>
<tr>
<td>Others</td>
<td>Technic-admin / type of contract (full-time, part-time) / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Full-time technic-admin / citizenship / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Part-time technic-admin / citizenship / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Technic-admin / type of contractual relation / pedagogical level</td>
</tr>
<tr>
<td></td>
<td>Technic-admin / age / type of contract (full-time, part-time) / citizenship</td>
</tr>
<tr>
<td></td>
<td>Financial revenues / source</td>
</tr>
<tr>
<td></td>
<td>Financial expenses / type of expenses (current, investment)</td>
</tr>
</tbody>
</table>
103. For primary, secondary, technical/vocational, literacy and adult education, questionnaires are printed at the central level and sent to the provinces. The provinces update the list of schools in the database. They prepare paper format list of schools and sets of questionnaires and send them to the districts. The districts send out questionnaires to be filled at the school level. One copy is kept at the school and two others copies are sent to the district. Questionnaire recovery is monitored at the district level according to the list of schools issued by the provinces. Each district keeps one copy and does its own manual processing. Afterwards, it sends a copy to the provinces. The provinces capture data, control quality and provide the central level with decentralized database on a CD-Rom or flash disk. At the central level, the national database is then consolidated.

104. For higher education, the whole process is centralized at the Higher Education Department.

105. Data on education finances are absent, hindering publication and analyses on cost efficiency and equity.

5.5. **Database structure**

106. Data collected from school census are processed by a relational database using Access. The database is multi-annual which constitutes an asset for data treatment. Generally, the database is well structured but is poorly documented.

107. Several tables or fields seem not to be used while doing data processing. There are at least three tables which are not used in any query: `EducationSubsystem`, `MarchOtherGrade` and `Professores`. Some tables contain fields that are not utilized and that are likely to introduce a misunderstanding of the data model such as the field `TypeId`, which provides information on the school status (i.e. public, private, community, etc.) in the table `School`.

108. Name fields are not standardized. All tables use systematically the field name `Id` for the unique identifier (i.e. primary key) (please refer to Figure A). This is fundamentally a source of error for maintenance operations and generally do not help in the understanding of the model by non-IT specialists, since these names are not reused in the data tables (please refer to Figure B).

Figure A:
109. The relational power of the Database Management System (DBMS) is not used as it should be. This could cause a loss of referential integrity for data. Most of the tables are not linked together as shown in Figure C where only six tables are related on a total of 70.

110. As an example, Figure D shows an extract of table MarchAges for which more than 2,000 records are not linked to the table School, while several queries use a “1/n relation type” as it is shown in Figure E which implies a referential integrity in the data.

---

6 For instance, it is possible to modify a code assigned for a specific school without such change being reflected in all other relevant tables.
Similarly, there are nearly 700 schools that are referenced in the table School which do not have associated data on pupils or are dummy records (please refer to Figure F).
112. The table School, which is the key for the establishment of reliable lists, is not maintained properly. A small analysis demonstrates the possible existence of duplicates (please refer to Figure G).

Figure G:

![Table showing data from School](image)

113. The current database refers to 70 tables, which are poorly documented, and to 303 requests for which none of them are documented as it is showed in Figure H (ref.: Qry_Alunos).

Figure H:

```
SELECT QryAlunosSemIdade.Year AS Ano, Province.Name AS Provincia, Province.Id AS IDProv,
School.DistrictId, District.Name AS Distrito, QryAlunosSemIdade.SchoolId, School.Name AS Escola,
SchoolType.Id AS IDEns, SchoolType.Name AS Ensino, EducationBranch.BranchId,
EducationBranch.BranchAbbreviation AS Nivel, QryAlunosSemIdade.ShiftId, Shift.Name AS Turno,
QryAlunosSemIdade.Grade AS Classe, QryAlunosSemIdade.SumOfNumberFemale AS Alu_M,
QryAlunosSemIdade.SumOfNumberMale AS Alu_H, MarchRepClassBoard.RepeatersFemale AS Rep_M,
MarchRepClassBoard.RepeatersMale AS Rep_H, MarchRepClassBoard.PureClasses AS TP,
MarchRepClassBoard.MixedClasses AS TM, MarchRepClassBoard.BoardersFemale AS Int_M,
MarchRepClassBoard.BoardersMale AS Int_H, [Alu_H]+[Alu_M] AS Alu_HM, [Rep_H]+[Rep_M] AS Rep_HM,
[TP]+[TM] AS Tur, [Int_H]+[Int_M] AS Int_HM, SchoolTypeHistory.SchoolTypeId,
SchoolTypeHistory.StartYear
FROM MarchRepClassBoard
RIGHT JOIN (((QryAlunosSemIdade INNER JOIN Shift ON
QryAlunosSemIdade.ShiftId = Shift.Id) INNER JOIN (Province INNER JOIN (School INNER JOIN District
ON School.DistrictId = District.Id) ON Province.Id = District.ProvinceId) ON QryAlunosSemIdade.SchoolId =
School.Id) INNER JOIN (EducationBranch INNER JOIN Education ON EducationBranch.BranchId =
Education.BranchId) ON QryAlunosSemIdade.EducationId = Education.Id) INNER JOIN (SchoolTypeHistory
INNER JOIN SchoolType ON SchoolTypeHistory.SchoolTypeId = SchoolType.Id) ON School.Id =
SchoolTypeHistory.SchoolId) ON (MarchRepClassBoard.Year = QryAlunosSemIdade.Year) AND
(MarchRepClassBoard.SchoolId = QryAlunosSemIdade.SchoolId) AND (MarchRepClassBoard.EducationId =
QryAlunosSemIdade.EducationId) AND (MarchRepClassBoard.ShiftId = QryAlunosSemIdade.ShiftId) AND
(MarchRepClassBoard.Grade = QryAlunosSemIdade.Grade)
WHERE (((SchoolTypeHistory.SchoolTypeId)=(SELECT Top 1 SchoolTypeId FROM SchoolTypeHistory
WHERE schoolId=School.Id AND StartYear<=QryAlunosSemIdade.Year ORDER BY StartYear DESC)) AND
((SchoolTypeHistory.StartYear)=(SELECT Top 1 StartYear FROM SchoolTypeHistory WHERE
schoolId=School.Id AND StartYear<=QryAlunosSemIdade.Year ORDER BY StartYear DESC)) AND
((School.Local)=0));
```

5.6. Synthesis and score

114. Based on an assessment of all the DQAF sub-dimensions, a global score of 68% has been assigned for the methodological soundness dimension.
5.7. Recommendations

115. Adapt the data collection instruments in order to introduce variables which are necessary to measure the number of hours spent teaching during an academic year.

116. Establish a national policy regarding repetition in order to allow a better planning of pupils / students flows over the years.

117. Develop a long term plan for decentralization, starting with a medium term pilot.

118. Even if new items have been introduced recently in the national questionnaires (i.e. teachers’ attrition and orphans in 2006 as well as dropouts by reason of dropout in 2007), there should be a planned and periodical review of data collection instruments with users and partners.

119. Develop an EMIS capacity development plan and a Monitoring and Evaluation (M&E) framework. There is no capacity development plan with a list of activities, time frame, budget and M&E framework for EMIS improvement.

120. Document the data model in order to make it available and usable by non-IT specialists who will have to use it, as well as maintainable over time.
121. Ensure data integrity using the management features of the referential data integrity (e.g., relations between tables, clearance of existing data and in particular the lists of references, etc.).

122. Document systematically queries used by the annual data processing.

123. Stabilize the list of establishments which is supported by the table School to ensure quality of the coded lists of establishments.

6. Accuracy and reliability: Source data and statistical techniques are sound and statistical outputs sufficiently portray reality

124. This dimension of quality is based on the principle that data produced give an adequate picture of the reality of the education sector in Mozambique.

125. Therefore, this dimension is specific for each data set and reflects the specificity of its sources and treatments.

6.1. Source data available provide an adequate basis to compile statistics

126. Statistics on enrolment and education resources, except educational expenditures, are collected through a regular administrative school census program.

127. The MoEC does not have a central list of all schools in order to evaluate the coverage at the national level. Therefore, it is impossible to know if data collected through administrative school questionnaires are fully representative of the Mozambican situation. There is a school codification that has been established in the database but it is strictly for technical reasons.

128. Education data collected through household surveys and population censuses are detailed regarding data on school enrolment and educational attainment. However, data on access to school, educational expenditures, recurrent expenditures (i.e. teachers’ salaries and administrative costs) and features of education expenditures are not collected through these types of data collections.

129. Statistics on the quality of learning outcomes are collected through assessments of student achievement. The National Institute for Educational Development (NIED) works in collaboration with the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ7) on a regular programme assessment of student achievement, at one or more ages or levels of education.

130. Statistics on the environment within schools that impact on quality of education collected using school surveys or censuses are generally detailed regarding school and teacher characteristics. However, data on school characteristics such as the level of community involvement, funding sources, estimated average family income and actual length of instruction time per year are not collected, neither the ones on teacher characteristics such as the age, experience as a teacher, pre-service teacher training and in-service training.

---

7 SACMEQ is an international non-profit developmental organization of 15 Ministries of Education which decided to work together to share experiences and expertise in developing the capacities of education planners to apply scientific methods to monitor and evaluate the conditions of schooling and the quality of education.
Teacher characteristics are available at the HR Department but these data are not used for statistical purposes.

6.2. Statistical techniques: Statistical techniques employed conform to sound statistical procedures, and are documented

131. The report forms are designed in a way that makes them easy to complete and appropriate for computer processing using Excel. However, procedures are not standardized.

132. There are no imputation methods and estimation techniques (e.g. sampling and calibration weights, etc.) which are used to treat the data.

6.3. Assessment and validation of source data: Source data are regularly assessed and validated

133. Accuracy of information is not routinely assessed. Administrative and survey data are not audited to check the accuracy of source data (e.g. inspection of field collections, random post-enumeration checks, etc.). However, there is a project under way in regard to this issue. In addition, information on coverage, non-response errors and the percentage of missing and/or imputed data by methods of imputation is not compiled.

134. Appropriate measures are not taken to validate data sources. For instance, no training is provided to improve data accuracy.

135. Concerning administrative data, since there is no central school register/list of all schools available, it is not possible to assess, for example, the number of students which are dropping out, moving or changing schools, or even students currently enrolled using this tool.

136. Intermediate/provisional results are not validated against other information where applicable.

137. Statistical discrepancies and other potential indicators of problems in statistical outputs are rarely investigated. For instance, there are no systematic processes which are in place to monitor errors and omissions, as well as to address data problems. As mentioned previously, results are not checked against demographic data and other survey/census results. In addition, data are sometimes compared with data from earlier years to examine reasonableness of year-to-year changes and trends; however, there are no systematic procedures which have been put in place.

6.4. Revision studies: Revisions, as a gauge of reliability, are tracked and mined for the information they may provide

138. Studies and analyses of revisions are not carried out routinely and used to inform statistical processes. For instance, methodologies are not regularly assessed and revised; therefore, the findings from these investigations can not be taken into account when data are compiled for subsequent periods.

139. Analyses of preliminary versus revised data are sometimes conducted but these are not done systematically.


6.5. **Synthesis and score**

140. Based on an assessment of all the DQAF sub-dimensions, a global score of 26% has been assigned for the accuracy and reliability dimension.

*Figure 4: Results of accuracy and reliability*

![Radar chart showing accuracy and reliability results with Mozambique and International norms]

6.6. **Recommendations**

141. Establish procedures to evaluate the school coverage. There should be the implementation of a national codification of all Mozambican institutions in order to maintain administrative lists up-to-date.

142. Generalize the use of Geographic Information System (GIS). An expertise in school mapping is available at the national level but such technology is much more relevant and useful at the sub-national level.

143. Data especially produced for the UIS should be used for steering and management.

144. Improve capacities in dealing with missing data (i.e. imputation of missing values) and quality control.
145. Data collected should be compared with data from different sources which also measure the same or closely related phenomena. This data validation would ensure a better quality of data by detecting inconsistencies that may arise from this process.

7. Serviceability: Statistics are relevant, timely, consistent, and follow a predictable revisions policy

146. This dimension is based on the principle that statistics should be published at appropriate intervals and in reasonable time, be consistent within a dataset and over time, as well as with other major datasets, and be subject to a revision policy and practice.

7.1. Periodicity and timeliness: Periodicity and timeliness follow internationally accepted dissemination standards

147. Respondents are made aware of the deadlines set for reporting. The MoEC employs follow-up procedures to ensure the timely receipt of respondent's data. However, if respondents fail to submit the questionnaires due to lack of knowledge or lack of resources, there are no appropriate adjustments which are made (e.g. missing data treatment).

148. Source data from the school census on enrolments and teachers are provided to the MoEC no later than four months after the beginning of the school year. On the other hand, source data on educational expenditures are not collected from within the MoEC no later than four months after the end of the school year.

149. At the central and provincial levels, there are no publications publishing preliminary data on education derived from the administrative school census.

150. Final publications of education statistics are disseminated annually. However, only 400 copies of these publications are produced which is a small number in order to have a good visibility.

151. The timeliness of statistics follows internationally accepted good practices. For final and international publications, statistics derived from the administrative school census are disseminated within 6 to 12 months after the beginning of the school year. However, data on pre-primary education are not submitted to the UIS for publication.

7.2. Consistency: Statistics are consistent within a dataset and over time, and with other major data sets

152. Accounting identities between aggregates and their components are not observed for all involved data. For instance, there are no verifications / comparisons of data from schools to schools.

153. Accounting identities between enrolments, repeaters, drop-outs and demographic data are not observed. Even if it is said that the coverage of schools is near 100%, without any real possibility to verify and confirm this fact, it was found that many schools do not complete the entire questionnaire which make difficult the analysis of internal consistency of the data.
154. Statistics are not cross-checked across geographic areas and sub-groups of population.

155. Consistent time series data are available for a period of eleven years. However, there is no revision policy.

156. When changes in source data, methodology and statistical techniques are introduced, historical data are not reconstructed as far back as reasonably possible. Detailed methodological notes which identify and explain the main breaks and discontinuities in time data, their causes as well as adjustments made to maintain consistency over time are not published.

157. Education statistics are reconciled with data from other sources but only internally with data coming from the HR Department. For instance, they are not reconciled with data collected through household surveys and population censuses.

158. From 2004 to 2009, data available show a continuous increase of school enrolment over this five years period, whether for boys or girls, with an average annual rate of 9% (please refer to Figure I). The provinces of Zambézia and Nampula are mainly responsible of this significant growth rate since they represent almost 40% of enrolment.

Figure I:

159. In addition, Figure J shows an atypical and significant increase of school enrolment for all grades in the provinces of Zambézia and Nampula.
160. The analysis of changes which occurred in school enrolment between 2007 and 2008 shows that all provinces do not have the same stability in terms of data collected (please refer to Figure J). There are reasons to believe that the quality of data is doubtful in the province of Zambézia and to a lesser extent in the provinces of Nampula and Tete given the significant variation in numbers from one year to another. It should be noted that the province of Zambézia covers over 20% of all institutions and school enrolment in Mozambique. Therefore, the quality of data in this particular province has a significant impact on the overall quality at the national level for both raw data and indicators.
161. Population data from the last General Population and Housing Census 2007 were supplied by the NSI and were disaggregated by province, district and single year of age. As shown in Figure K, data seem to be consistent by looking at their structure which is an important element regarding the reliability of denominators for the main intake ratios.
162. Regarding EP1, school administrative data available over the last five years put into perspective with the latest population data allow to produce a series of gross enrolment ratio (GER). Figure M shows that all ratios are higher than 100%, which suggests a poor quality of data collected and/or high repetition rates. The provinces of Zambézia, Tete and Nampula have again an atypical profile which deserves further analysis.
163. Net enrolment rates are very sensitive to the reliability of population and / or pupil age distribution. As it has been mentioned above, data on population do not jeopardize the quality of the age distribution of school-age children. However, by following a cohort (for instance, the one for school children of six years old in 2004), important inconsistencies can be clearly identified about the distribution in certain provinces such as Zambézia and Nampula (please refer to Figure N). This finding makes questionable the reliability of data by age collected through school census.
The net enrolment rate (NER) calculated for the first five years of EP1 is 102%. By disaggregating this rate at the provincial level, it appears that half of the provinces have rates beyond the technical limit of 100% (please refer to Figure O). These provinces are Zambézia, Tete, Niassa, Manica and Maputo, as well as Maputo City (i.e. Cidade de Maputo).
165. It is likely that the problem encountered in the province of Maputo is mainly due to a misallocation of educational entities between the province of Maputo and Maputo City because typically the main intake ratios are generally better in urban areas. For instance, the aggregation of these two entities into one reduces the overall rate of the Maputo “zone” to 97%.

166. The calculation of NER requires the distribution of population by single year of age. The database structure can not fully meet this requirement. The ISCED97 mapping indicates that the theoretical entrance age for EP1 is 6 years old but the codification in the database does not use this specific age and uses less or equal to 6 (i.e. 6 ou menos) instead (please refer to Figure P). Therefore, it is likely that children of five years old are actually included in the NER when they should not be part of the calculation.

Figure P:

167. As shown in Figure Q, there is a relatively homogeneous student age distribution at the national level.

Figure Q:

168. However, there is likely an over-reporting of pupils of ten years old (classical phenomenon of attraction of round numbers) resulting in an unjustified increase of pupils which are taken into account in the calculation of the NER (please refer to Figure R).
169. Although longitudinal data are available, dropout and repetition rates are evaluated by comparing data collected in March and December. This method is very dependent on the quality of the data collection over a year. The availability of two consecutive years allows producing these rates in a more compliant manner with international methodologies.

7.3. Revision policy and practice: Data revisions follow a regular and publicized procedure

170. Revisions do not follow a regular, well established and transparent schedule. There is no adequate documentation of revisions included in the publication of the statistical data and in the database accessible to users.

171. Preliminary data or first estimates are not clearly identified in statistical releases. Users are not alerted that the initially published data are preliminary and subject to revision. In addition, the revised data are not disseminated with the same level of details as previously published for the data being revised. There is no revision policy in place. Nevertheless, a project is in progress for 2010.

7.4. Synthesis and score

172. Based on an assessment of all the DQAF sub-dimensions, a global score of 32% has been assigned for the serviceability dimension.
**Figure 5: Results of serviceability**

![Graph showing results of serviceability with values 0.64, 0.08, and 0.24 for Periodicity and timeliness, Revision policy and practice, and Consistency respectively. Mozambique and International norms are compared.]

### 7.5. Recommendations

173. Implement missing data treatment at the decentralized level.

174. Financial information at the school level (e.g., capitation grant) should be collected, particularly in order to cross-check these data with macro data.

175. Promote preliminary publications on education statistics.

176. There should be an increase of the number of final publications published in order to obtain a wider dissemination of education statistics.

177. Develop automated procedures for controls of coherence at the central and decentralized levels.

178. Establish systematic procedures at the school and district levels in order to identify and explain the important variations of the number of students.

179. Identify problems related to the data collection of student ages and probably implement technical tools to perform the rectification of distributions.

180. Conduct a general review of the classification of establishments by province and, more specifically, for Maputo and Maputo City.

181. Establish strong methodological procedures for data collection and robust technical procedures for treating student ages.
182. Investigate in a more extensive way the quality of data of Zambézia, Nampula and Tete since they showed an “atypical behavior” compared to other provinces. Since these three provinces represent almost 50% of the school aged population in the country, it is worth to dedicate more time on them in order to improve the overall data quality.

183. Use and implement the necessary technology to produce rates of internal efficiency using standard methodology (i.e. promotion, repetition and dropout rates).

8. Accessibility: Data and metadata are easily available and assistance to users is adequate

184. This dimension is based on the principle that data and metadata should be presented in a clear and understandable way and should be easily available to users. Metadata should also be relevant and regularly updated. In addition, assistance to users should be available, efficient and performed in a reasonable time frame.

8.1. Data accessibility: Statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis

185. Education data are published in a clear manner; charts and tables are disseminated with the data to facilitate the analysis. Three major publications are published every year and data offer adequate details as well as analysis of current period estimates.

186. Depending on the intended audience and purposes, data of different degree of aggregation (e.g. school, region), sub-components (e.g. gender, level of education, age, public / private, full-time / part-time) and additional data (e.g. demographic, socioeconomic, geographic information) are made available.

187. Dissemination media and formats are not adequate. Data are not first released via an information release, which is then followed by the release of a more comprehensive publication. In addition, there are no recently released data and longer time series data which can be accessed through an electronic database maintained by the MoEC.

188. Annual education statistical yearbooks are made available through the MoEC website. However, the statistical yearbook is available for the current year and the ones from the previous years are only available on demand.

189. The statistical data are not released according to a pre-announced schedule.

190. Non-published and non-confidential specialized tabulations (e.g. sub-aggregates of units of analysis) are made available upon request and non-confidential micro data files (e.g. with information allowing the identification of individual respondents removed) are available to permit analytical use by researchers and other users.

191. The availability of non-published data and the terms and conditions on which they are made available are not publicized.
8.2. **Metadata accessibility: Up-to-date and pertinent metadata are made available**

192. The metadata for the data produced provide users with adequate information about what the data mean and about the methodology used to collect and process them. Some metadata are attached to the education data which are available on the NSI website. However, no metadata exist on the MoEC website.

193. Metadata do not provide information on biases in the data, information about response rates to the main surveys used, main linkages with other major data systems highlighting important differences from these data systems, neither other information users may need to assess the data.

194. Deviations from internationally accepted standards, guidelines, or good practices are not documented in the metadata.

195. Neither brochure nor booklet has been prepared to inform general users about the data. In addition, there is no comprehensive sources and methods document produced to inform analysts and other users of statistics about how statistics are compiled.

8.3. **Assistance with the users: Prompt knowledge support service is available**

196. Prompt and knowledgeable service and support are not available to users of statistics. All statistical releases do not identify specific individuals who may be contacted by mail, telephone, facsimile, or by email.

197. There is no documentation developed (e.g. brochure, booklet) to educate users of related datasets.

198. Assistance to users is not monitored through periodic surveys of users.

8.4. **Synthesis and score**

199. Based on an assessment of all the DQAF sub-dimensions, a global score of 29% has been assigned for the accessibility dimension.
8.5. Recommendations

200. Make available recently released data and longer time series data, preferably through an electronic database maintained by the MoEC. Since data on the website are accessible through PDF files, which is not convenient for research purposes, providing Excel files would be a first step towards improving the quality of the service given to data users.

201. There should be metadata available on the MoEC website since nothing as such is available to education data users.

202. Make available a more complete and precise set of metadata regarding data which are currently available on the NSI website.

203. Expand the diversity of the publications, like for instance, by producing specific brochures for vulgarization purposes.

204. Produce a methodological document describing the statistical business process model used at the MoEC.

205. Invest more efforts and resources on the assistance with data users by increasing the quality of the service which is actually provided.

9. Conclusion and overall recommendations

206. Based on an assessment of all the six DQAF dimensions, a global score of 41% has been assigned for the different steps currently in place for the collection, processing, analysis and dissemination of Mozambican education data at the national and sub-national levels.
207. The principal dimensions where there is a great potential for improvement are integrity (particularly in the areas of professionalism and transparency), accuracy and reliability (specifically in assessment of data source, statistical techniques and revision studies), serviceability (in particular consistency and revision policy and practice) as well as accessibility (especially on metadata accessibility and assistance with the users).

208. For pre-requisites of quality, improvements should be particularly made on the human resources and communication aspects. At the MoEC, there is a lack of personnel mainly in the field of statistics. The recruitment of additional staff at the national level is also crucial since more in-depth research on planning and education sector performance analysis is required. On the communication issue, there should be a better collaboration between the MoEC and other education data producing agencies. More particularly, the MoEC and the NSI would benefit from strengthening their relationship. Another aspect of communication which must be improved is the production of technical documentation and statistical procedures manuals.

209. For integrity, some efforts could be made in enhancing the objectivity or impartiality in the collection, processing, analysis and dissemination of data, as well as in the institutionalized provisions which guarantee or improve the professionalism of the different actors in their practices. The MoEC should cultivate the assessment and regular use of secondary sources of data to benefit from a very important comparison tool for data quality improvement. In addition, important progress should be made regarding transparency. The MoEC should implement its own documentation about the terms and conditions under which official statistics are compiled and disseminated. The development of an educational communication policy towards the public is a key component to improve transparency.

210. For methodological soundness, one of the major issues is the lack of quality data on finances, especially on human and financial resources invested in education. In addition, data coverage is not complete since data for pre-primary education are collected by the
Ministry of Women and Social Affairs and these are not reported to the UIS to allow international comparisons. The MoEC should develop a long term plan for decentralization, an EMIS capacity development plan and a M&E framework.

211. For accuracy and reliability, the exactness of information is not routinely assessed, appropriate measures are not taken to validate data sources and the accuracy of school registers is not periodically evaluated. Concerning soundness of data procedures, improvement should be made especially regarding measures on coverage and procedures on how to manage missing data. In addition, the MoEC would benefit from comparing the data they collect with data from different sources which also measure the same or closely related phenomena. This data validation would ensure a better quality of the data by detecting inconsistencies that may arise from this process.

212. For serviceability, the main issues that should be addressed are the lack of missing data treatment and preliminary data published at the central and provincial levels. While some efforts are made regarding consistency over time and consistency obtained through other data sources and / or statistical frameworks, greater efforts should be made regarding internal consistency. No checks are done to verify if statistics are consistent within the dataset. The MoEC should put in place automated procedures to control / check the coherence of data at the central, provincial and district levels.

213. Finally, for accessibility, statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis. However, the dissemination means are not commensurate with users' needs and there is no schedule for data release which is announced in advance. In terms of metadata accessibility, major improvements should be done to make available up-to-date and pertinent metadata. They should provide more detailed information such as biases in the data, information about response rates to the main surveys used, etc., as well as levels of details that are adapted to the needs of the intended audience. Regarding assistance with the users, improvements should be made to provide a prompt and knowledgeable support service to data users.

10. Annexes

Annex 1: Acknowledgments – Participation list

<table>
<thead>
<tr>
<th>Province</th>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maputo</td>
<td>Mr. Raphael CHAUQUE</td>
<td>Statistician</td>
</tr>
<tr>
<td></td>
<td>Mr. Samuel MENESSES</td>
<td>Head of Planning Department</td>
</tr>
<tr>
<td>Bagamoyo</td>
<td>Mr. Aloyce GABRIEL</td>
<td>District Planning and Logistical Officer</td>
</tr>
<tr>
<td></td>
<td>Ms. Cheka OMARI</td>
<td>District Education Officer</td>
</tr>
<tr>
<td></td>
<td>Ms. Judika KISENGE</td>
<td>Statistical and Logistical Officer</td>
</tr>
<tr>
<td>Temeke</td>
<td>Ms. Beatrice NYANGIRI</td>
<td>IT Specialist</td>
</tr>
<tr>
<td></td>
<td>Mr. Frank MAKINGI</td>
<td>Statistical and Logistical Officer</td>
</tr>
<tr>
<td></td>
<td>Ms. Halima SONJE</td>
<td>Adult Education Officer</td>
</tr>
<tr>
<td></td>
<td>Ms. Honorina MUMBA</td>
<td>District Education Officer</td>
</tr>
<tr>
<td>Dar es Salaam Regional Office</td>
<td>Ms. Assumpta NDIMBO</td>
<td>Regional Administrative Secretary</td>
</tr>
<tr>
<td>EMIS Unit</td>
<td>Ms. Assela LUENA</td>
<td>Unit Head</td>
</tr>
<tr>
<td></td>
<td>Mr. Godfrey PONERA</td>
<td>Senior Statistician - SACMEQ Specialist</td>
</tr>
<tr>
<td></td>
<td>Mr. P. KIRUMBA</td>
<td>Senior Statistician</td>
</tr>
<tr>
<td>Ministry of Education and Culture</td>
<td>Mr. Adelino CONSTANCIO</td>
<td>Statistics Department</td>
</tr>
<tr>
<td></td>
<td>Mr. Ildio BUDUIA</td>
<td>Head of Statistics Department</td>
</tr>
<tr>
<td></td>
<td>Mr. Kauxique MAGANLAL</td>
<td>Head of Communication and Information Technologies Department</td>
</tr>
<tr>
<td></td>
<td>Mr. Manuel REGO</td>
<td>National Director of Planning and Cooperation</td>
</tr>
<tr>
<td>National Statistics Institute</td>
<td>Mr. KWESSI GABO</td>
<td>Head of Social Sectors</td>
</tr>
<tr>
<td>Philippe Samuel Magaia School</td>
<td>Mr. MWARAPALA</td>
<td>Education Statistician</td>
</tr>
<tr>
<td></td>
<td>Patenguana</td>
<td>Director</td>
</tr>
<tr>
<td>UNICEF - Mozambique</td>
<td>Ms. Anjana MA NGALAGIRI</td>
<td>Chief of Education and Child Development</td>
</tr>
<tr>
<td></td>
<td>Mr. Stefano VISANI</td>
<td>Programme Specialist Assessment</td>
</tr>
<tr>
<td></td>
<td>Mr. Tomoko SHIBUYA</td>
<td>Education Programme Specialist</td>
</tr>
<tr>
<td>National Statistics Institute</td>
<td>Ms. Fatima ZACARIAS</td>
<td>Director for Demographic and Social Statistics</td>
</tr>
<tr>
<td>ID</td>
<td>Name of the education programme</td>
<td>Minimum entrance requirements</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Ensino Pré-escolar</td>
<td>3 years old</td>
</tr>
<tr>
<td>2</td>
<td>EP1 - Ensino Primário do 1º grau</td>
<td>6 years old</td>
</tr>
<tr>
<td>3</td>
<td>EP2 - Ensino Primário do 2º grau</td>
<td>EP1 - Ensino Primário do 1º grau</td>
</tr>
<tr>
<td>4</td>
<td>ETE - Ensino Técnico Elementar</td>
<td>EP1 - Ensino Primário do 1º grau</td>
</tr>
<tr>
<td>5</td>
<td>Ensino Secundário, 1º Ciclo (ESG1)</td>
<td>Primary Education Certificate</td>
</tr>
<tr>
<td>6</td>
<td>Ensino Secundário, 2º ciclo (ESG2)</td>
<td>Primary Education Certificate</td>
</tr>
<tr>
<td>7</td>
<td>Ensino Secundário, 2º ciclo (ESG2)</td>
<td>Lower secondary education certificate</td>
</tr>
<tr>
<td>8</td>
<td>Ensino Técnico profissional (ETP)</td>
<td>Secondary education certificate</td>
</tr>
<tr>
<td>9</td>
<td>Formação de professores (Nível médio)</td>
<td>ESG 1 (10º classes)</td>
</tr>
<tr>
<td>10</td>
<td>Formação de professores (Nível superior)</td>
<td>ESG 2 (12º classes)</td>
</tr>
<tr>
<td>11</td>
<td>Bachelor</td>
<td>Upper secondary education certificate (general or vocational)</td>
</tr>
<tr>
<td>12</td>
<td>Masters</td>
<td>Bacharelato</td>
</tr>
<tr>
<td>13</td>
<td>Licenciatura</td>
<td>Bacharelato</td>
</tr>
<tr>
<td>14</td>
<td>Doctorate</td>
<td>Master’s Degree</td>
</tr>
</tbody>
</table>

**Notes:**
- G = General
- P = Pre-vocational
- V = Vocational
- na = not applicable
- A, B, C = Destination categories

**N.B.** For some programmes, the entrance age and duration presented in columns E and F for ISCED levels 0-3 may be different from that in columns K and L because these reflect the parameters used by UIS in the calculation of participation indicators. The tertiary gross enrolment ratio (GER) uses a population of the age-group corresponding to five years following on from the secondary school leaving age (e.g. if the ending of secondary education is 17, then the age-group used for the tertiary GER will be 18-22).