Data entering, data management and QC

Dr. Matieu Henry & Dr. Laurent Saint-André & Dr. Nicolas Picard
Hanoi, June 18 - 22th 2012
How data resulting from inventory and monitoring efforts will be acquired, managed, processed, checked, analyzed, distributed, and archived.

**Objective:**

To transfer field data into ONE electronic database

**Different types of software:**

- Microsoft Office
- Access
- MySQL

http://science.nature.nps.gov/im/datamgmt/index.cfm
In order to avoid errors:

To check all the data several times: double entry

Comparison of the two files to detect the errors

To well define the variables to be measured (text of numbers)

Exemple: Geographic coordinates

<table>
<thead>
<tr>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lat 25.456</td>
<td>Long 12.567</td>
</tr>
<tr>
<td>Lat 7°28’55.1</td>
<td>Long 13°41’25.9</td>
</tr>
</tbody>
</table>
METAINFORMATION

When the measurements were performed

Details of the code and acronyms

Definition of the variables

Height of the measured diameter or circumference etc.

Units of the measured parameters

Any additional information

Crucial to allow the use of the data at different time period and different users
If you have different level of information, you need to have different tables.

This structure of information minimize the repetition of the information.
Objective: Delete the artifacts/ impossible data
The difficulty depends on identifying what is normal or not normal
Very often, the errors come from a change of units
QUALITY CONTROL

Graphic analysis

Descriptive statistics

DATA STRUCTURE

One line for one individual of the sampled population

One column -> one descriptive variable