Data Quality Project Methodology

DQM Data Quality Manager

DQ survey

Statement of needs
- problem definition
- assumed level of duplicates
- expected data quality level

Data quality audit
- data profile
- data investigation
- data quality reports

Project definition
- stakeholder management
- reference materials

Software install
- hardware verification
- DQM install
- reference data set up

DQ analysis

Data sampling
- define scope of sampling
- build sampling from source

Define business problem
- business perception
- current process
- business impact

Collect data source
- source data inventory
- sample data
- study external data

Profile analyse data
- profile data
- token & pattern analysis
- investigate data quality

Data governance

Prepare workshop
- engage stakeholders
- recap project objectives
- summarise DQ issues

Measure KPI
- build dashboard
- collect sample issues

Process redesign
- process implications
- systems impact

Data governance
- data governance workshop
- cost & business case

Data re-engineer

Refine match rule
- build token list
- standardise rules
- dedupe & match rules

Cleanse match job
- build parse rules
- build match rules
- dev survival process

Systems integration
- integrate to ops system
- automate batch process
- assess thruput performance

Online data entry
- online process
- design prototype UI
- setup online match job

DQ implementation

Documentation
- process documentation
- match result documentation

Integration testing
- end-to-end testing
- audit result

DQ implementation
- present to project sponsor
- product training
- snapshot back up