MUNI UNIVERSITY

FACULTY: TECHNO-SCIENCE

DEPARTMENT: COMPUTER AND INFORMATION SCIENCE

**DEVELOPING OF AGRIWARE DATA WAREHOUSE FOR RESEARCH CENTERS IN UGANDA**

ISM 3202: DATA WAREHOUSE AND BUSINESS INTELLIGENCE.

COURSE WORK 1

OREE DENISH od1@muni.ac.ug

ADRIKO DEBO JOEL debojoel@gmail.com

ENZEBO EMMANUEL ee1@muni.ac.ug

CHANDIRU ROSEMARY

ADAKU JOHN aj2@muni.ac.ug

ENZAMA CEASER ec1@muni.ac.ug

OCATRE FRANK of1@muni.ac.ug

**BACKGROUND**

This data warehouse (AGRIWARE) is to serve all agricultural research centers in Uganda namely Abi zardi. (Abi Zonal Agricultural Research and Development Institute), Kawanda, Namulonge, etc. Agriculture is one of the major economic activities in Uganda and the major crops include: cassava, beans, groundnuts, sweet potatoes, sesame, tobacco, sorghum, rice, maize, soya beans, finger millet, banana, etc. besides livestock. The government in conjunction with NGOs has established Agricultural research centers, therefore AGRIWARE is expected to integrate these centers through data sharing.

**PROBLEM STATEMENT.**

For all these research centers to effectively and efficiently share information and data, there should bea shared data warehouse.

However currently there hasn’t been a successfully executed efforts in establishing a data warehouse to integrate these disparate data centers.

Therefore AGRIWARE development is to enable the agricultural research centers to respond to agricultural research situations and advice the agricultural community based on the real time data.

**INFORMATION NEEDS.**

1. FAO database (FAOSTAT). This database can be accessed through [www.FAO.org/FAOSTAT](http://www.FAO.org/FAOSTAT) and contains a massive data set based on categorical domains such as; production(crops processed, live animals, etc.); inputs which improves fertilizers by nutrients, fertilizers by products, etc.; Trade (crop and livestock products, live animal, detailed trade matrix, and trade indices); Population (annual population), etc. By the time of writing this proposal, the following data under FAOSTAT were found to be of more interest.

Data set about the crop domain, which describes trends based on area harvested, yields, and production elements for various crop species for the year 1961 to 2016. Analysis and visualization of this data set could probably give an insight to the Agricultural research teams on how to influence the agricultural practices for the better.

Other data sets of interest include live animals, livestock primary,Livestock processed land use, air and climate change, however others will be selected in future during execution of the project.

2. Data.ug. This can be accessed from catalog.data.ug/dataset. Even though FAO data set forms the core, there should be a secondary data set to give more insight to the trends, for example contrasting climatic factors against the trends. Furthermore, the FAO data set focuses on Uganda as whole, therefore data from data.ug is to further disperse the data sets to various elements such as districts in particular distances to local markets, number of households, proportions of households related to agricultural trends, dispersion of major crops across districts, etc. By the time of writing this proposal, the following data sets were already identified to be of interest.

Districts major crops, districts major crop production, estimated district population, percentage distribution of households by distance. It’s very important to know the following;

1. More data sources will be identified during the course of execution of the project. Therefore the data sources listed above does not imply that AGRIWARE will be solely developed in them.
2. All the data sets (including their descriptions) and to be discovered data sets will be accessed from the team’s Github repository.

**DIMENSIONS AND FACTS**

**Dimension Tables**

**Key:**

**Naming convention: camel case**

**PK = Primary Key**

**SK = surrogate Key**

**U = Unique**

**FAOSTAT dimension table**

|  |
| --- |
| FAOSTATDomain |
| domainId (PK)(SK)  domainCode(U)  domainName  dateOfRecord |

|  |
| --- |
| FAOSTATItem |
| ItemId (PK)(SK)  ItemCode(U)  ItemName  dateOfRecord |

|  |
| --- |
| FAOSTATFlag |
| FlagId (PK)(SK)  Flag  FlagDesscription  dateOfRecord |

**FAOSTAT FACT table**

|  |
| --- |
| FAOSTATFaoData |
| DataId (PK)(SK)  domainId(FK)  ElementId(FK)  ItemId(FK)  FlagId(FK)  YearId(FK)  Unit  Value  dateOfRecord |

**Data.ug**

1. **District Dimension table.**

|  |
| --- |
| **DataUgAge** |
| AgeId(PK)(SK) |
| Age |
| DateOfRecord |

|  |
| --- |
| **DataUgGender** |
| GenderId(PK)(SK) |
| Gender |
| DateOfRecord |

|  |
| --- |
| **DataUgDistrict** |
| DistrictId(PK)(SK) |
| DistrictName |
| DateOfRecord |

|  |
| --- |
| **DataUgYear** |
| YearId(PK)(SK) |
| Year |
| DateOfRecord |

**FACT TABLE**

|  |
| --- |
| **DataUgPopulation** |
| PopulationId(PK)(SK) |
| DistrictId(FK) |
| AgeId(FK) |
| GenderId(FK) |
| YearId(FK) |
| Population |
| DateOfRecord |

**2. Percentage distribution of households by distance.**

**DIMENSION**

|  |
| --- |
| **DataUgDistanceToLocalMarket** |
| DistanceToLocalMarketId(PK)(SK) |
| DistanceToLocalMarket |
| DateOfRecords |

**FACT TABLE**

|  |
| --- |
| **DataUgData** |
| DataId(PK)(SK) |
| DistanceToLocalMarketId(FK |
| NumberOfHouseholds |
| ProportionOfHouseholds |
| DateOfRecord |