Statistical Analysis Plan:

**Purpose/Objective**: To study the mortality trends of cancers in the Laos population over a decade. (for example: 2007-2017)

**Hypotheses:** N/A for above paper

**Data sources and primary variables**:

Main exposure: N/A

Main Outcome: N/A

Other relevant variables: All variables needed to explain the trends for cancers over time.

Data sources would include: population registry (demographic variables), cancer registry (ICD-10 codes); medical registry (diagnosis, index date, treatment); prescription registry; hospital registry, death registry

**Crude methods, derived variables and role of covariates**

**Skeleton tables:**

Baseline characteristics of the study population

Figures of the trends of mortality for different cancers in the study population over a decade

**Schedule**: a time plan for the project (literature review; data collection and analysis; writing the article and submitting it)

**Storage**: Directory, work server

For the project, a general data management plan would include,

**Step 1:** I would start by making a statistical analysis plan for the project. (so the aim, hypothesis, variables needed, statistical methods and software that will be used, and a project timeline)

**Step 2:** Study the data and check for the variables that are really relevant for the project that will be used. Keep only the required data and drop the rest. Also, create self-explanatory variable names.

**Step 3:** Reduce space taken by the data by creating an efficient data set- avoid redundancy of the data, reduce length of strings by labeling them, rounding up decimals, storing in UTF-8 format. Set up a denormalized data set, so it will be easy to join them when needed or only merge those needed for the analysis.

**Step 4:** Write codes; codes that are

Easy to understand to both self and others

Short and precise

Adding comments as to why the code is done

Creating pseudocodes

Having several versions

**Step 5:** Have good documentation and control of the different versions, can be done

i. manually - (directory structure, read-me files, self-explaining file name with date or version no:, templates, create back-ups)

ii. automatic - using Git for example which creates back ups on a web hotel, easy to share files with others, the files are protected, changes can be performed and all codes that are changed will be documented.