Excel files (.xlsx) are a zip file containing several XML files. We will need to note two in particular for adding data too. One will contain the strings such as names and notes, the others will contain all the formatting.

The best way to do this would be to keep a template Excel file on the server that is already unzipped. Containing all the formatting etc, just that has no data in it.

The archive will contain a file called “sharedStrings.xml” that will contain all the strings used in the spreadsheet.

Text

Description automatically generated

The archive will also contain a folder “worksheets” that will contain each and every worksheet in its own file named “sheet1xml, sheet2xml…”

A picture containing chart

Description automatically generated

This is an example of the string we’ll need to create for each row.

It’s probably best to obtain all the data from the database, parse all the strings into some kind of map, then iterate through the map to create the string and save it to the two files.

The sheet names are stored in the “workbook.xml” file stored in the root.

Text

Description automatically generated

So when parsing the Strings into the spreadsheet we should create an array of the strings and they will need to be mapped to the relevant cells. Since all the strings for names should be in one column only, as are the notes, the row numbers should be relevant to map to all the strings. It is unlikely that any notes should be duplicates, but on the small chance there are duplicated notes, the same procedure should apply.

We need to check with the client if they only use a single worksheet or multiple worksheets. If they do want multiple sheets, how do they want them named.

We’ll then need to repackage the files into a zip archive, rename it to a .xlsx file and upload it to the users device

It may be easier to build this as a desktop app, have the user’s PC pull the necessary data from the database and then parse and convert the files locally. Since there’s four phases to the process. Collecting the data from the data base, converting that data into the spreadsheet, zipping, and renaming the archive and finally downloading the file. The processing will cost the aquarium if done on a server, where downloading the data to the users machine and then doing all of the conversion locally will save money in that respect.