

DATA 607 Project 3 - Part 1

TEAM NAME: DATA DOMINATION

Team Members

1. Nfn Tenzin Dakar
2. Nwe Oo Mon
3. Crystal Quezada

Project Description

Our group, DATA DOMINATION, will use Zoom and Slack for communication. For code sharing and documentation, we're using RStudio Cloud with R Markdown published on RPubS. All files, including data and documentation, are stored in a GitHub repo. We've created an ERD using Mysql DataBase and identified our data sources, which will be loaded via CSV files from GitHub.

Tools for the Project

Our team leverages several tools to streamline collaboration and project development:

- RStudio Cloud (rstudio.cloud): This platform facilitates collaborative code development, allowing all team members to view, edit, and share code seamlessly in real-time.
- R Markdown in RStudio Cloud: We use R Markdown to document our project, and publish the results through RPubS (rpubs.com), making our work easily accessible.
- MySQL : This tool was employed to design the entity-relationship diagram (ERD) for the project.
- GitHubRepository (https://github.com/Nweoomon/Project3_607) : All source CSV files and R Markdown (RMD) files are stored centrally on GitHub, allowing the team to access and manage files efficiently.
- Communication: Our discussions and updates happen over Zoom for meetings and Slack for daily communication, ensuring smooth coordination.

We identified our sources of data as the following:

1. Kaggle ML & DS Survey:

Retrieved from <https://www.kaggle.com/code/kerneler/starter-ds-job-listing-technology-04cdb0b7-0/input> It is also stored in our collaborative [GitHub repository](https://github.com/Nweoomon/Project3_607)

Entity Relationship Diagram

Our diagram can be found on our shared

https://github.com/Nweoomon/Project3_607/blob/main/NOM_Job_listing_ER_Diagram.docx

Load libraries

```
# load in the library  
  
library(tidyverse)  
library(lubridate)  
library(wordcloud)  
library(RMySQL)
```

Importing Data through MySQL Database

Data clean up and transformation

Analysis

Conclusion