

# DMDD document flow:

## 1<sup>st</sup> Folder- Data Gathering,Cleaning and verifying.

### 1) Data Source Files:

This file contains all the data that we have gathered from Kaggle.

### 2)Data Analysis and Visualization:

This file contains analysis of data gathered.

### 3)Data Definition:

This file scripts for insertion of data in SQL and visibility of data.

### 4)Data Gathering:

SQL code

### 5)Data Gathering, Cleaning and Visualization:

Files contains all the data gathered, cleaned.

### 6)Data Insertion Script:

File contains the Python code to insert the data in MySQL tables.

## 2<sup>ND</sup> Folder- ER diagrams.

File contains diagram before and after normalization.

## 3<sup>rd</sup> Folder- Normalization and Data Formatting

### 1) Executing NF

Pdf shows how NF has been executed along with screenshots.

### 2) NF Normalization Script

This file contains normalization script.

## 4<sup>th</sup> Folder- Use Cases and Output of our project:

This folder contains all the use cases file to showcase our project.