**Sri Lanka Institute of Information**

**Technology**

A picture containing text, clipart, vector graphics

Description automatically generated

**Data Warehousing & Business Intelligence**

**Assignment 01**

**Submitted by:**

**Rathnasooriya P.U (IT20156206)**

Contents

[1.Data Set Selection. 3](#_Toc105358450)

[01. Background: - 3](#_Toc105358451)

[02. Content: - 3](#_Toc105358452)

[03. ER Diagram 4](#_Toc105358453)

[2. Preparation of Data Sources. 5](#_Toc105358454)

[3. Solution Architecture. 6](#_Toc105358455)

[4. Data warehouse design & development. 7](#_Toc105358456)

[5. ETL Development 8](#_Toc105358457)

[5.1 ETL –Source to Staging 8](#_Toc105358458)

[5.1.1 Load Data Department Staging Table 8](#_Toc105358459)

[5.1.2 Load Data Doctor Staging Table 9](#_Toc105358460)

[5.1.3 Load Data Location Staging Table 9](#_Toc105358461)

[5.1.4 Load Data Hospital Staging Table 10](#_Toc105358462)

[5.1.5 Load Data patient Staging Table 10](#_Toc105358463)

[5.1.6 Load Data Branch Staging Table 11](#_Toc105358464)

[6. Staging To Data Warehouse 12](#_Toc105358465)

[6.1Load Department Data 12](#_Toc105358466)

[6.2 Load Doctor Data 13](#_Toc105358467)

[6.3 Load Patient Data 13](#_Toc105358468)

[6.4 Load Location Data 14](#_Toc105358469)

[6.5 Load FactHospital Data 15](#_Toc105358470)

[7. Data Warehouse Update SISS Package 16](#_Toc105358471)

[7.1 Load accm\_txn\_complete\_time Data 16](#_Toc105358472)

[7.2 Load\_txn\_process\_time\_houres Data 17](#_Toc105358473)

[8. Accumulated Fact Table 18](#_Toc105358474)

# 1.Data Set Selection.

## Background: -

With rapidly growing demand for healthcare in hospitals and intensive care units (ICUs) throughout the world, the COVID-19 epidemic has put unprecedented strain on health systems. As the epidemic spreads, evaluating the related healthcare resource requirements (beds, personnel, and equipment) has become a top issue for many countries. Estimates of how long patients with COVID-19 require various levels of hospital care are needed to forecast future demand.

## Content: -

The data set was downloaded from -:

[COVID-19 Hospitals Treatment Plan | Kaggle](https://www.kaggle.com/datasets/arashnic/covid19-hospital-treatment?select=host_train.csv)

## ER Diagram

Diagram

Description automatically generated

# Preparation of Data Sources.

In order to data extraction need to prepare the data sources. From my main data source, I have extracted to type of data sources.

1. Text file (.txt)
2. Excel files (.xlsx)
   * Hospital Excel Data File.
   * Department Excel Data File.
   * Doctor Excel Data File.
   * Customer Excel Data File.
   * Patient Excel File.
   * Accm\_txn\_complete\_time Excel File.

From these dataset, I import above excel files to SQL SERVER MANAGEMENT STUDIO and create the database called **Hospital\_sourceDB** Database.

**Hospital\_sourceDB** have following tables -:

* + Hospital table.
  + Department Table.
  + Doctor Table.
  + Customer Table.
  + Patient Table.

**Text file** –This text file include all the hospital address details including address, city, state, postal code and the country.

# Diagram Description automatically generatedSolution Architecture.

**Hospital\_stagingDB**

* Hospital\_Staging
* Location\_Staging
* Department\_Staging
* Doctor\_Staging
* Location\_Staging
* Patient\_Staging
* accm\_txn\_complete\_time\_Staging

**Hospital\_DW**

* DimLocation
* DimDepartmen
* DimDoctor
* DimLocation
* FactHospital

# Data warehouse design & development.

**Relational Diagram – Star Schema.**

**A screenshot of a computer

Description automatically generated with medium confidence**

DimLocation is slowly changing dimension. Hospital Address may be changed in future. Therefore, I get it as slowly changing attribute.

Country-> City-> Postal Code-> State-> Address = This is the Hierarchies.

# **ETL Development**

## 5.1 ETL –Source to Staging

**Graphical user interface

Description automatically generated**

### 5.1.1 Load Data Department Staging Table

**Graphical user interface, text, application, chat or text message

Description automatically generated**

### 5.1.2 Load Data Doctor Staging Table

Graphical user interface, text

Description automatically generated

### 5.1.3 Load Data Location Staging Table

**Graphical user interface, text

Description automatically generated**

### 5.1.4 Load Data Hospital Staging Table

**Graphical user interface, text, chat or text message

Description automatically generated**

### 5.1.5 Load Data patient Staging Table

Graphical user interface, text

Description automatically generated

### 5.1.6 Load Data Branch Staging Table

Graphical user interface, text, application, chat or text message

Description automatically generated

# Staging To Data Warehouse

**Graphical user interface

Description automatically generated**

### 6.1Load Department Data

**Graphical user interface, text

Description automatically generated**

### 6.2 Load Doctor Data

Text

Description automatically generated

6.3 Load Patient Data

Graphical user interface, text, application

Description automatically generated

### 6.4 Load Location Data

A picture containing graphical user interface

Description automatically generated

### 6.5 Load FactHospital Data

A picture containing graphical user interface

Description automatically generated

# Data Warehouse Update SISS Package

**A picture containing diagram

Description automatically generated**

### 7.1 Load accm\_txn\_complete\_time Data

**Text

Description automatically generated**

### 7.2 Load\_txn\_process\_time\_houres Data

**Text

Description automatically generated with medium confidence**

# Accumulated Fact Table

**Graphical user interface, application, table

Description automatically generated**

**Graphical user interface

Description automatically generated**