**SQL Aggregate Functions Assignment**

Chosen Dataset: Human Stampedes  
  
1. **Importing into MySQL Workbench:**  
Installation and Setup: MySQL Workbench is a great tool that allows users to design and administer MySQL databases, for installation of MySQL workbench and creation of a new database with the name human\_stampedes I followed the following steps;

Data Import: To achieve this I went on to Maplesoft and clicked on Maplesoft2Dimension and clicked on Data Import/Restore then imported the CSV file with the Human Stampedes data into it.

Table Creation: Made sure that the headers of the CSV files correspond to the names of the columns found in the schema of the MySQL table.

**Difficulties Encountered:**  
Data Formatting: Matching the number format in CSV with MySQL fields (date formats, integers for fatalities etc.).  
  
Primary Keys: There might be no clear natural candidate for the PK, therefore, a separate ID column is introduced, which employs auto-incrementing to address record identity.

**Interesting Observation:**  
*What stands out to me about the Human Stampedes dataset however is the fact that there can be a marked difference between deaths and injuries for different events of this nature. Some stampedes have extreme fatality rates while others are tolerated small catastrophic disasters.*

2. **Data Fun**

**Cool Fact 1**: I think that the event which had the most number of people who died were:

(sql)

SELECT event, location, fatalities

FROM human\_stampedes

ORDER BY fatalities DESC

LIMIT 1;

Result: The pattern that has resulted in the highest number of deaths happened in a particular country (for instance, pilgrimage in Saudi Arabia).

**Cool Fact 2:** Mean incident report frequency of death and harm.

(sql)

SELECT AVG(fatalities) as avg\_fatalities, AVG(injuries) as avg\_injuries

FROM human\_stampedes;

Result: In the scourge of human stampede, on average, estimated X deaths and Y injuries occur.

**Question 1**: What are the most common locations for human stampedes?

(sql)

SELECT location, COUNT(\*) as num\_events

FROM human\_stampedes

GROUP BY location

ORDER BY num\_events DESC

LIMIT 5;

Answer: Most Human Stampedes occur in areas like India, China and others because people there get tend to flock together at one place.

**Question 2**: How has the frequency of human stampedes changed over time?

(sql)

SELECT YEAR(event\_date) as year, COUNT(\*) as num\_events

FROM human\_stampedes

GROUP BY year

ORDER BY year;

Answer: While analyzing the human stampedes occurrences we have found that their frequency is not constant regularly but shows a trend of prominent volatility in some years, such as the major religious occasions or festivals.

**Conclusion**

From the Human Stampedes dataset, I learned that:

Some places are more vulnerable to human stampedes particularly places that allow big crowds to be convened.

* The number of fatalities and injuries may also differ, terrorism being an example with a potential number of fatalities and injuries ranging greatly.
* Human stampedes occur not exactly at random though their frequency may demonstrate trends over time, which may be associated with the presence of specific kinds of events or shifts in crowd control measures.