**SDS-Mini-Project**

**Sem-3**

**Group members:**

**Report:**

**Description of dataset**

* This dataset consists of the weather records for a month taken with the observations being taken down and noted per hour.
* It consists of various aspects of weather like precipitation, humidity, wind speed, wind bearing, visibility, loud cover, pressure, temperature, apparent temperature against a particular instant of time.
* The summary and the daily summary are noted down appropriately as per the parameters collected.
* Most of the parameters are numerical with the remaining few being categorical.
* The objective of this project is to make optimum use of the dataset to extract maximum and most relevant information out of the dataset for various purposes.
* The first section of the project would consist of descriptive statistics while the second half would contain inferential statistics.
* With the help of the relationships generated between various variables, we can appropriately put this to a great use by inferring various parameters from it.
* While this is just a sample of 720 records, we can easily put this to a great use to infer about various weather parameters on the other given days with optimal accuracy and precision along with minimal margin of error.

## **Inferences**

**With the assistance of data and results proved above, we obtain various inferences on the nature of the dataset as mentioned below:**

* Most of the days were rainy and cloudy in nature since the observations in the dataset have been from the month of September, during which monsoon/rainy season is prevalent in most parts of the Northern Hemisphere in and around the Tropic of Cancer.
* We can also easily infer and predict the weather in the fore-coming couple of months, using the given statistics to find the parameters.
* We can also predict the weather in the similar months in the upcoming few years, but at the same time, we would also need to take into account the rising rate of various environmental effects like greenhouse effect leading to global warming thereby leading to imbalanced seasons which result in excess droughts and floods. The average temperature of the earth is also rising slowly but steadily.
* So, it wouldn't be wrong to say that the inferences and predictions made in the future i.e. after 5 years and more... wouldn't be plausible as they are for the near future, thereby leading to an increase in the degree of uncertainty and many more assumptions and null hypotheses being rejected.
* But we can predict the weather to some extent in general even in the upcoming years, just that the degree of uncertainty would be witnessing an exponential rise which isn’t going to be of any help to the statisticians.