**Project Proposal Form**

**Name:**  Maxim Rohit

**Student Number:**

**Course:**  MS in Data Science

**Project Title/Area 1:**

Understanding Air pollution trends in various Part of Maharashtra, India.

**Description:**

**Motivation:**

Indian Government in a press release on 8th December released few disturbing facts about how the increase in air Pollution had a degrading impact on Indian Living conditions.

1. Air pollution can be attributed to of 12.5% of deaths in India. [1]
2. 12.4 lakh deaths in India attributed to air pollution in 2017, of which 50+% were in individuals younger than 70 years.[1]
3. Life expectancy in India has gone down by 1.7 years due to health loss caused by higher (than accepted) air pollution level. [1]

We will attempt to analyse localize impact on region of Maharashtra in detail.

**Summary:**

1. Analyse & Categorize the rate at which air pollution is increasing in Various cities & Towns of Maharashtra using clustering techniques like hierarchical clustering and K-mean.
   1. Validate how city progression maps to its Air pollution levels?
2. Understand how various pollutants and other features co-relates to each other using PCA
3. Is their time based (seasonal & otherwise) variation in the spread of various pollutant, via EDA.
4. Use timeseries analysis to project the pollution growth pattern in the coming year and predict when the air pollutant level goes beyond breathable standards.

**Objective:**

1. Categorize various regions within Maharashtra, India based on Air pollution trend.
2. Analyse the difference between these regions w.r.t
   1. Concentration Type
   2. Rate of Change
   3. Population
   4. Elevation from Sea level
   5. Time of Year
   6. Buying trends of automobiles/Impact of Additional Vehicles
   7. Rainfall levels
   8. Others
3. How does cities like Pune (An IT Hub in India) & Mumbai (An Overall commercial Hub) fair against Neighbouring upcoming Towns (Nasik, Nagpur) versus the less progressive one like Ahmednagar, Solapur over the last couple of decades?

**Data :-**

1. **Air Quality data** [2] : -
   1. Year wise
      1. 2004-2015 – 57529 data points
      2. 1987-2003 – 2891 data points
   2. City wise 2500+ records each for 11 different cities & more
   3. Pollutant types
      1. SO2
      2. NO2
      3. RSPM/PM10
2. Population data [3] : -
   1. 2001 & 2011 District wise Sensex data
3. Automobiles [4]: -
   1. Citi wise registration data
   2. Year range: 2001-2015

**Project Title/Area 2:**

Analyse the impact of formation across the top 6 European football leagues in the 21th Century

**Description:**

1. English
2. Spanish
3. Italian
4. German
5. French
6. Portuguese

The said leagues have different playing styles, some of the analysis of the difference are along these lines: -

1. How their champions play

2. How team trying to survive in the league play

3. How team trying to qualifying for Europe play

4. Plus there is always that new team a1st season great; 2nd season abysmal syndrome

We would analyse these leagues from following angles & more and see which footballing formation 4-4-2,4-3-3, 4-2-2-2 etc in their attacking or defensive flavours have brought most success to these leagues.

**Assumptions: -**

The formation under consideration will be from an overall mostly used formation during the season, we will ignore the tactical change that the team might have made to address a certain opponent in specific situation in a match.

**Project Title/Area 3:**

Analyze the relation of Economy rate (Runs given per over)  of a bowler w.r.t the probability of its team winning the IPL Match.

**Description:**

Cricket specifically the 20-20 format is considered a Batsman game.

We would try to find the depth of truth in that statement by analyzing the impact of an economical bowler in the team victory in the IPL matches.

Member of Staff:                                                                                   \_\_\_\_

If you have found a supervisor then you and the member of staff who agreed to supervise your project should sign below.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_                                                                        \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student Signature                                                                         Supervisor Signature

\_\_\_\_\_\_\_\_\_\_\_\_\_                                                                            \_\_\_\_\_\_\_\_\_\_\_\_

Date                                                                                               Date