

**Name:** Akshay Hareshbhai Sitlani  
**RUID:** 197006017  
**Subject:** Data Wrangling and Husbandry  
**Term:** Spring 2020  
**Instructor:** Jason Klusowski

**Project Title:** Novel Coronavirus (COVID-19) India Analysis

**Brief Description**

We all are familiar with the crisis we are going through right now. Although, it started with China, USA is the most affected country of all right now. In India, although the cases of COVID-19 surfaced a little late, the current situation is getting worse. The rate at which the cases are increasing on daily basis is high. I came across a dataset which gathers data from Ministry of Health and Family Welfare (MoHFW) of India regarding the current cases, deaths and recovered cases of novel coronavirus. Along with this, India also has its own dedicated coronavirus tracker website namely, [www.covid19india.org](http://www.covid19india.org). Similar data is taken from this website and a comprehensive dataset has been made. I am using this dataset for my project. Based on a preliminary view of the dataset, I plan to do the following analysis for this project:

- Analyze the current situation of disease on country-wide level
- Get gender and age wise insights (Active cases, deaths and recovered cases)
- Figure out various trends. For example:
  - Number of tests conducted per day
  - Percentage of people tested positive out of the tests conducted
  - New cases reported on daily and weekly basis
  - State wise cases analysis
  - Recovery and Mortality rate analysis
  - Develop demographic visualization

The project will include these specifics but will not be limited to it. I plan to add more features as I progress through the project.

Data source: [www.kaggle.com](http://www.kaggle.com)

**NOTE:** I wish to do a longer final report for the project and no presentation.