# India In Pixels Assignment

**Problem Statement :**

In this assignment we need to prepare a dataset of all ODI players and create 3 datasets as :
1. A list of all cricket players who have ever played ODI matches
2. Runs they have made every year in their career
3. Cummulify the scores by year for each player

**Strategy Used :**

The Strategy used to solve the problem is to scrap the data from reliable sources and then save it, preprocess it and transform it as per the need.

**Programming Language Used :**

Python

**Modules Used :**

1. requests : for sending HTTP requests to the server
2. BeautifulSoup : to scrap data
3. pandas : to perform data manipluation and structuring
4. numpy : to replace null values with NaN
5. csv : to create CSV file

**Source of data :**

The data is scraped from http://www.howstat.com/cricket/

**Files Attached :**

**Jupyter notebook files :**
1. data_scraping.ipynb : This file contains the codes used for scraping the data
2. data_cleaning.ipynb : This file contains the codes used for claening and restructuring the data
3. cumulative_data.ipynb : This file contains the codes used for cumulating the data

**CSV files :**
1. ODI_Players.csv : This file has the list of all ODI players with their countries for which they have played
2. ODI_Players_runs.csv : This file as the data of scores in each year

3. ODI_Players_runs_Cleaned.csv : This file has restructured and cleaned data for convenience
4. cumulative_data.csv : This file has the cumulative scores of each player


**<u>Report Prepared By :</u>**
Deepanshu
Contact no. :- 9911234151
Email_ID :- pundhird@gmail.com