Indian Agriculture Data

Compiled from various Indian government sources

Contributed by

Mr. Digvijay Singh	Ms. Usha Viswanathan,
Project Research Associate	Sr. Project Manager
R Team, FOSSEE	FOSSEE
IIT Bombay	IIT Bombay

13 October 2022

1. Introduction

This document involves a detailed description of the **Indian Agriculture Data** dataset. **Indian Agriculture Data** is a combination of multiple datasets related to soil moisture, soil nutrition, agricultural land area, agricultural production of various horticulture crops, and Gross State Domestic Product or GSDP for 27 Indian states (except Goa and Sikkim) in the year 2018-19.

This document lists all the data sources used, along with the duration of the extracted data. It also mentions the column headers of the **Indian Agriculture Data** dataset along with the description of their content.

2. Data Sources

The following table gives a summary of data sources along with their associated web links used to construct the dataset:

Table 1: List of data sources used to construct the dataset.

Source Name	Duration
[Soil Moisture (District-wise)](#)	June 2018 - Dec 2018

[Gross State Domestic Product (Current Prices) with base year 2011-12](#)	2018-19
[List of Indian States by Zones](#)	1956, 1972 & 2002
[Soil Nutrition (State-wise)](#)	2017-19
[Soil Nutrition (District-wise)](#)	2017-18 to 2018-19
[Area and Production of various Horticulture Crops (State-wise)](#)	2018-19

These datasets contain the status of soil, area and production of horticulture crops, GSDP (Gross State Domestic Product), etc. of Indian states. The information in the above-mentioned data sources is in PDF/CSV/HTML format. We compiled the relevant information from these sources and constructed a dataset, and named it **Indian Agriculture Data.** In the next chapter, we elaborate on the columns of the constructed dataset.

3. Data Description

The description of headers/column-names of the constructed dataset, as described in the previous chapter, is given in the table below.

Table 2: Description of each header of the constructed dataset.

Header	Description	Data type
State	Name of an Indian state	character
Zone	Zone in which the state resides	character
GSDP	Gross State Domestic Product for 2018-19 at current prices with 2011-12 as the base year	integer
Area	State-wise area in '000 Ha of various horticulture crops during the year 2018-19	numeric
Prod.	State-wise production in '000 MT of various horticulture crops during the year 2018-19	numeric
Prod./Area	Ratio of production and area in MT per Ha	numeric
NFS	Nitrogen Fertility Status of the soil from a particular Indian state during the 2017-2019 cycle	character
PFS	Phosphorus Fertility Status of the soil from a particular Indian state during the 2017-2019 cycle	character
KFS	Potassium Fertility Status of the soil from a particular Indian state during the 2017-2019 cycle	character
Dist.	Name of a district associated with the corresponding state	character

SML	Distrist-wise mean Aggregate Soil Moisture Level Percentage (at 15cm)	numeric
SMV	Distrist-wise mean Aggregate Soil Moisture Volume Percentage (at 15cm)	numeric
SDN	District-wise Soil Deficiency percentage in terms of nutrients in Nitrogen for 2017-18 to 2018-19 cycle	numeric
SDP	District-wise Soil Deficiency percentage in terms of nutrients in Phosphorus for 2017-18 to 2018-19 cycle	numeric
SDK	District-wise Soil Deficiency percentage in terms of nutrients in Potassium for 2017-18 to 2018-19 cycle	numeric

4. Data file format

This dataset is stored in three different file formats as listed below:

a) **CSV:** Indian_Agriculture_Data.csv
b) **XLS:** Indian_Agriculture_Data.xls
c) **XLSX:** Indian_Agriculture_Data.xlsx

A random subset of size 5 is selected from the constructed dataset and printed below for clarity.

Table 3: A random subset of the constructed dataset.

State	Zone	GSDP	Area	Prod.	Prod. /Area	NFS	PFS	KFS	Dist.	SML	SMV	SDN	SDP	SDK
TAMIL NADU	Southern	16302 0915	1309 .76	1654 6.07	12.63	LOW	MED IUM	MED IUM	PUD UKK OTT AI	32.29	24.89	99.16	24.65	53.38
KAR NAT AKA	Southern	14906 2378	2114 .99	1877 5.66	8.88	LOW	MED IUM	MED IUM	GAD AG	55.77	23.1	89.58	98.96	0.9
UTT AR PRA DES H	Central	15847 6388	2286 .97	3875 3.71	16.95	LOW	LOW	MED IUM	CHA NDA ULI	20.02	22.96	99.74	97.45	16.97
TAMIL NADU	Southern	16302 0915	1309 .76	1654 6.07	12.63	LOW	MED IUM	MED IUM	PER AMB ALU R	31.61	30.73	100	75.67	6.19

| GUJARAT | Western | 149215571 | 1606.16 | 22908.74 | 14.26 | LOW | MEDIUM | HIGH | TAPI | 6.38 | 18.27 | 92.11 | 19.04 | 0.08 |