**Project Initialization and Planning Phase**

|  |  |
| --- | --- |
| Date | 18 August 2025 |
| Name | Mohammed Farhan |
| Project Title | India Agriculture Crop Production Analysis (1997–2021) Using Tableau |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | Analyze India’s agricultural crop production data (1997–2021) to surface long‑term trends, regional variations, and crop performance, and to deliver interactive Tableau dashboards and a Tableau Story that support policy, research, and farm‑level decisions. |
| Scope | • Years: 1997–2021 • Geography: Indian states/UTs available in the dataset • Crops: Major food & cash crops (e.g., Rice, Wheat, Maize, Pulses, Sugarcane, Cotton) • Analyses: State-wise & crop-wise production, yield trends, YoY growth, and top/bottom performers • Outputs: Interactive Tableau dashboards with filters (State, Crop, Year) and a Tableau Story for narrative insights • Out of scope: Predictive modeling beyond 2021 and causal attribution studies |
| **Problem Statement** | |
| Description | Agriculture production data in India is published across multiple years and regions in static reports/spreadsheets. This fragmentation makes manual analysis slow and error‑prone, limiting the ability to quickly discover patterns in crop performance, yield, and regional disparities. |
| Impact | • Enables evidence‑based policy decisions for crop planning and resource allocation • Helps identify underperforming crops/regions to prioritize interventions • Assists farmers/advisors with seasonal planning by visualizing historical patterns • Improves transparency and reproducibility using version‑controlled workflows |
| **Proposed Solution** | |
| Approach | • Acquire crop production dataset from official open data portals (CSV format) • Clean & preprocess: handle missing values, remove duplicates, standardize crop/state names and units • Engineer metrics: total production, area, yield (tons/ha), YoY growth, multi‑year averages • Build Tableau dashboards: time‑series (1997–2021), state comparisons, crop‑wise views, top/bottom rankings • Add interactivity: global filters (State, Crop, Year), drill‑downs, and a Tableau Story for end‑to‑end narration • Validate results and document insights; publish packaged workbook and final report |
| Key Features | • State‑wise and crop‑wise interactive dashboards • Time‑series trends across 1997–2021 with YoY growth indicators • Top‑N/Bottom‑N crops and state comparisons with quick filters • Optional choropleth maps (if geo fields available) for regional insight • Tableau Story integrating multiple dashboards into a single narrative • Clean, responsive layout with performance‑optimized extracts |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | CPU/GPU specifications, number of cores | Quad‑core CPU (Intel i5/i7 or equivalent); GPU not required |
| Memory | RAM specifications | 16 GB RAM (8 GB minimum) |
| Storage | Disk space for data, models, and logs | ≥100 GB SSD free space (datasets, extracts, exports) |
|  | | |
| Frameworks | Visualization & analytics platform | Tableau Desktop / Tableau Public |
| Software | Visualization & analytics platform | Tableau Desktop / Tableau Public |
| Libraries | Python libraries for data preprocessing/validation | pandas, numpy, matplotlib |
| Development Environment | IDE & version control | Jupyter Notebook, GitHub |
| **Data** | | |
| Data | Source, size, format | India Agriculture Crop Production Data (1997–2021); Source: Government Open Data Portal / Directorate of Economics & Statistics; CSV; ~10–200 MB |