AgriTech : AI & Analytics in Agriculture

Overview

This comprehensive dataset encompasses various aspects of agriculture, which includes detailed information on farmers, crops, locations, agricultural technologies, and market dynamics. By leveraging data analytics and artificial intelligence, this project aims to uncover valuable insights that can drive informed decision-making in the agricultural sector.
The goal is to identify patterns and trends that can enhance crop yields, optimize resource utilization, and improve farmers' livelihoods. Through advanced visualization techniques, interactive dashboards will be created to illustrate the potential impact of AI and modern farming techniques on agricultural productivity and sustainability. Here, the data is considered for India.

NOTE : This is fake data and is generated using Python.

Who are the audiences ?

Agricultural Experts and Researchers, Supply Chain Managers, Agriculture-Business Companies, Corporate MNCs who want to invest in CSR initiatives

Dataset Description

Dimension Tables

1. **Farmer**

Field	Measure/Dimension DataType	Description
Farmer_ID	Dimension; Integer	Unique identifier for each farmer
Farmer Name	Dimension; String	Name of the farmer, representative of the regional names by state
Age	Dimension; Integer	Age of the farmer, ranging between 18 and 95 years
Experience	Dimension;Integer	Years of farming experience, ranging between 1 and 40 years
Education	Dimension; String	Education level, categorized as Illiterate, Primary, Secondary, Higher Secondary, or Graduate
Village	Dimension; String	Village name, specific to the farmer’s state

State	Dimension; String	State where the farmer resides, selected from Indian states
Nearby_City	Dimension; String	Closest major city to the farmer's village

2. **Crop**

Field	Measure/Dimension DataType	Description
Crop_ID	Dimension;Integer	Unique identifier for each crop
Name	Dimension;String	Name of the crop (e.g., Rice, Wheat, Maize), common in Indian agriculture
Category	Dimension;String	Crop category (e.g., Cereal, Pulses, Cash Crop, Oilseed)
Growing_Season	Dimension;String	Primary season when the crop grows (Kharif, Rabi, or Zaid)
Water_Requirement	Dimension;Float	Average water requirement per growing season in millimeters (mm)
Soil_type	Dimension;String	Type of soil best suited for the crop (e.g., Alluvial, Black, Red, Laterite, Arid)

3. **Location**

Field	Measure/Dimension DataType	Description
Location_ID	Dimension;Integer	Unique identifier for each location
District	Dimension;String	Name of the district within the state
State	Dimension;String	Name of the state where the location is situated, chosen from major Indian states
Region	Dimension;String	Geographic region of India where the location lies (e.g., North, South, East, West, Central)
Climate_Zone	Dimension;String	Climate classification for the location (e.g., Tropical Wet, Tropical Dry, Subtropical Humid)

Soil Quality	Dimension;String	Quality rating of soil in the location (e.g., Good, Average, Poor)

4. **Technology**

Field	Measure/Dimension DataType	Description
Technology_ID	Dimension; Integer	Unique identifier for each agricultural technology
Name	Dimension;String	Name of the agricultural technology (e.g., Drip Irrigation, Soil Sensors)
Category	Dimension;String	Category of technology, specifying its purpose (e.g., Irrigation, Monitoring, Machinery, Software, Infrastructure)
Complexity	Dimension;String	Level of complexity in implementing the technology (e.g., Low, Medium, High)
Initial Cost	Dimension; Float	Initial setup cost for the technology in INR, estimated based on the type and purpose of the technology
Maintenance Cost	Dimension;Float	Estimated annual maintenance cost in INR for using or operating the technology

5. **Market**

Field	Measure/Dimension DataType	Description
Market_ID	Dimension; Integer	Unique identifier for each market
Name	Dimension; String	Name of the agricultural market (e.g., "XYZ Mandi" or "ABC Agro Market")
Type	Dimension; String	Type of market, indicating the operational model (e.g., APMC, Farmers Market, Cooperative, E-commerce, Contract Farming)
Distance	Measure; Float	Distance of the market from nearby villages or towns in kilometers

Demand_Level	Dimension; String	Estimated demand level of the market, indicating buyer interest (e.g., High, Medium, Low)
Price_Volatility	Measure; Float	Price fluctuation factor in the market, indicating how much prices can vary

Fact Tables

1. **Production**

Field	Measure/Dimension DataType	Description
Production_ID	Dimension; Integer	Unique identifier for each production record
Farmer_ID	Dimension; Integer	Foreign key linking to `dim_farmer`, identifying the farmer associated with the production
Crop_ID	Dimension; Integer	Foreign key linking to `dim_crop`, identifying the crop being produced
Location_ID	Dimension; Integer	Foreign key linking to `dim_location`, identifying the location of the production
Yield Kg / hectare	Measure; Float	Crop yield per hectare in kilograms, indicating productivity of the crop
Production cost / hectare	Measure; Float	Cost of production per hectare in INR, capturing expenses such as labor, fertilizer, etc
Land_Used_Hectares	Measure; Float	Land area used for the crop in hectares
Water_Used_Liters	Measure; Float	Water consumption in liters based on crop and area, reflecting water used for irrigation
Fertilizer_Used_Kg	Measure; Float	Fertilizer quantity used in kilograms, indicating the amount applied to the land for crop yield

2. **Technology Adoption**

Field	Measure/Dimension DataType	Description
Tech/Adoption_ID	Dimension; Integer	Unique identifier for each technology adoption record
Farmer_ID	Dimension; Integer	Foreign key linking to `dim_farmer`, identifying the farmer who adopted the technology
Technology_ID	Dimension; Integer	Foreign key linking to `dim_technology`, identifying the technology adopted
Location_ID	Dimension; Integer	Foreign key linking to `dim_location`, identifying the location where the technology is adopted
Adoption Level	Measure; String	Level of adoption indicating the extent of technology utilization (e.g., Low, Medium, High)
Implementation Cost	Measure; Float	Cost incurred by the farmer for implementing the technology in INR
Productivity_Increate_Percent	Measure; Float	Percentage increase in productivity due to technology adoption
ROI_Percent	Measure; Float	Return on investment percentage reflecting financial benefits of the technology
Training_Hours	Measure; Float	Number of hours invested in training to use the technology

3. **Market Sales**

Field	Measure/Dimension DataType	Description
MarketSales_ID	Dimension; Integer	Unique identifier for each market sales record
Famer_ID	Dimension; Integer	Foreign key linking to `dim_farmer`, identifying the farmer involved in the sale
Crop_ID	Dimension; Integer	Foreign key linking to `dim_crop`, identifying the crop sold
Market_ID	Dimension; Integer	Foreign key linking to `dim_market`, identifying the market where the crop was sold
Quantity Sold (Kg)	Measure; Float	Quantity of the crop sold in kilograms
Sale Price / Kg	Measure; Float	Price per kilogram of the crop in INR
Total Revenue	Measure; Float	Total revenue generated from the sale, calculated as `Quantity_Sold_Kg` × `Sale_Price_Per_Kg`
Transportation Cost	Measure; Float	Cost incurred for transporting the crop to the market in INR
Market Fees %	Measure; Float	Percentage of the sale price charged as market fees

Key Performance Indicators (KPIs)

1. Crop Yield Efficiency
   KPI: Average Yield per Hectare

   Research Questions:
   a. How does crop yield vary across different regions and soil types?
   b. How does yield per hectare compare across crop types and growing seasons?
   c. What role does technology adoption play in increasing average yield?

2. Water Usage Efficiency
   KPI: Liters of Water Used per Kg of Crop Produced

Research Questions:

a. Which crops and regions require the most water per hectare?

b. Are farmers in certain regions adopting water-efficient technologies, and how does that impact usage?

c. How does water usage vary across different growing seasons?

d. How does technology adoption impact water usage efficiency?

3. Transportation Cost as Percentage of Revenue
   KPI: Percentage of total revenue spent on transportation costs for crop sales.

   Research Questions:

   a. Which regions or crop types have the highest transportation costs relative to revenue?

   b. Does proximity to the market reduce transportation cost percentages?

   c. How do transportation costs vary across different market types (e.g., APMC vs. e-commerce)?

4. Market Sales Performance
   KPI: Average Revenue per Sale

   Research Questions:

   a. How does market distance affect sales revenue and transportation costs?

   b. What is the relationship between market demand levels and sale prices?

5. Return on Investment for Technology
   KPI: Average ROI Percentage for Technology Adoption

   Research Questions:

   a. Which technologies provide the highest ROI across different farm sizes?

   b. How long does it take for different technologies to achieve a positive ROI?

   c. Are there specific crops or regions where technology adoption has shown a higher ROI?

6. Market Profitability
   KPI: Net Profit Margin (Total Revenue - Costs) / Total Revenue

   Research Questions:

a. How do market fees and transportation costs impact profitability across different regions?
b. What is the relationship between crop type and market profitability?

7. Adoption Rate of Advanced Agricultural Technologies
   KPI: Percentage of farmers using advanced technologies like AI-based tools, GPS-guided tractors, and soil sensors.

   Research Questions:
   a. Which states or regions show the highest adoption rate for advanced technologies?
   b. What are the top technologies adopted by farmers, and how do they impact productivity?
   c. How does adoption differ between crop categories (e.g., cash crops vs. cereals)?