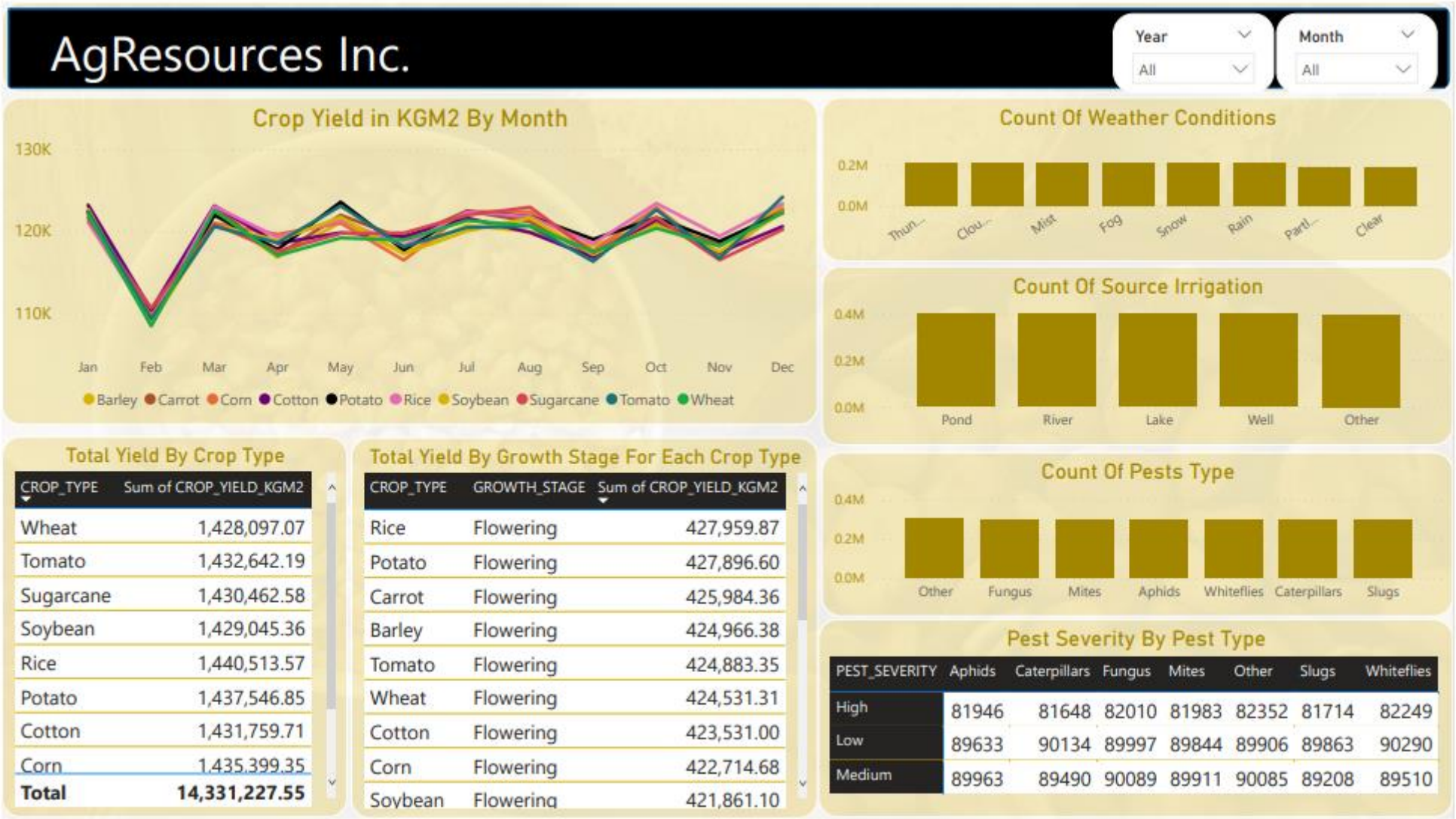
Team X: Report



Data Cleaning and Preprocessing: A Comprehensive Overview

Data is often considered the new oil in today's data-driven world. However, much like unrefined oil, raw data is seldom ready for analysis. It requires meticulous cleaning and preprocessing to ensure its quality, relevance, and coherence. Here, we delve into the steps we took to prepare our dataset for a seamless analysis.

1. **Merging Tables**:

-**Crop Type and Growth Stage with Crop**:

- In many datasets, information can be spread across multiple tables. To consolidate our data and ensure a uniform dataset, we merged the 'Crop Type' and 'Growth Stage' tables using the 'Crop' as the connecting key. This brings together critical data about each crop and its corresponding growth stage, thus facilitating analyses that consider the interaction between these two elements.

- Irrigation Method with Irrigation:

- Efficient water management is vital in agriculture. To gain deeper insights into irrigation practices, we merged the 'Irrigation Method' with the 'Irrigation' table. This provides a comprehensive view of how different irrigation methods correspond with various irrigation events.

- Pest Type with Pest:

- A robust understanding of pests is essential for sustainable agriculture. By merging 'Pest Type' with 'Pest', we centralized information about different pests and their specific categories, making it easier to understand the relationships and patterns between different pests.

- Weather and Weather Condition:

- Environmental conditions play a pivotal role in agriculture. To avoid redundant or missing data, we executed an inner join between 'Weather' and 'Weather Condition'. This ensures that only rows with complete data in both tables are retained, eliminating entries with 'NaN' or missing values. This step results in a dataset that's both comprehensive and free from gaps.

2. **Column Renaming**:

A dataset can have columns with ambiguous or unclear names, making it challenging to understand the data. We revisited the column titles and renamed some for clarity. Standardizing column names also aids in ensuring consistency across datasets and making future merges or joins more straightforward.

3**. Time-stamp Conversion**:

Temporal data is a treasure trove for analysis but can be tricky if not standardized. Our dataset contained timestamps that were converted into a 'datetime' format. This transformation not only facilitates easier temporal analyses but also ensures compatibility when merging with other datasets that might also have datetime fields.

Data Model

Overview of the Data Model

The provided data model appears to represent the various components and factors involved in agricultural and environmental data management. This model is built around seven primary tables: IRRIGATION, LOCATION, SENSOR, CROP, PEST, WEATHER, and SOIL. Additionally, there's a table named 'Calendar', which could serve as a lookup or reference table for time-related data.

IRRIGATION: Contains details about irrigation with attributes such as IRRIGATION_DURATION_SEC, IRRIGATION_METHOD, WATER_SOURCE, and timestamps.

CROP: Contains information about different crops. Attributes include CROP_ID, CROP_TYPE, CROP_YIELD_KG/M2, GROWTH_STAGE_ID, and timestamps.

PEST: Provides details about pests affecting the crops with attributes such as PEST_DESCRIPTION, PEST_ID, PEST_SEVERITY, and PEST_TYPE.

SOIL: This table captures the soil characteristics essential for crop growth. The attributes are ORGANIC_MATTER, PHOSPHORUS_LEVEL, SOIL_COMP, SOIL_TEMP, SOIL_MOISTURE, and SOIL_PH.

WEATHER: Maintains records of weather conditions with attributes like MONTH_NAME, PRECIPITATION_MM, TIMESTAMP, WEATHER_CONDITION, and WEATHER_CONDITION_ID.

LOCATION: Gives details about the geographical and regional information of the fields. Attributes include ELEVATION, LATITUDE, LONGITUDE, REGION, and LOCATION_NAME.

SENSOR: Stores data related to sensors, capturing various environmental metrics. Attributes include BATTERY_LEVEL, HUMIDITY, LIGHT_INTENSITY, SENSOR_ID, SOIL_MOISTURE, and timestamps.

Calendar: a support table for date-based queries, including Date, Month, Month_name, Year, and Weekday.

Analysis:

1. Crop Yields:

- Based on the "Total Yield By Crop Type" table, all crops have very similar yields, with most hovering just above 1.43 million KGM2.

- Throughout the year, from the "Crop Yield in KGM2 By Month" graph, we observe a generally consistent yield for most crops with a minor dip in the mid-year.

- No single crop stands out as having a significantly higher yield than the others, which means we would need to factor in production costs, market demand, and selling price to determine profitability accurately.

2. Environmental Factors:

a) Weather Conditions:

- The "Count Of Weather Conditions" doesn't indicate which condition is more favorable for which crop, but it provides an overall sense of the prevailing conditions.

- Some crops might have decreased yields during specific weather conditions. For instance, continuous rain might be detrimental to crops like wheat due to potential fungal infections. On the other hand, sunny or clear weather might benefit sun-loving crops like corn.

b) Source of Irrigation:

- The dashboard indicates that various sources like Pond, River, Lake, Well, etc., are used for irrigation, with no significant source dominating over the others.

- Depending on the water quality and mineral content from these sources, some crops might benefit more than others. For example, river water might be more nutrient-rich than well water, affecting the growth of specific crops.

c) Pest Types and Severity:

- Multiple pests affect the crops, with Aphids, Caterpillars, and Whiteflies having relatively high severity counts across all severity levels.

- The profitability of crops can be severely ed by these pests, especially if they target high-yield crops. Integrated pest management systems should be employed to mitigate these effects.

3. Projecting into Future Profitability:

a) Soil:

Changes in soil pH, nutrient content, and structure can significantly affect crop yields. For instance, tomatoes prefer slightly acidic soil, while carrots prefer sandy soil.

Regular soil testing and necessary amendments can ensure optimal soil health for each crop, leading to increased yields and profitability.

b) Pests:

An increase in pest severity or the emergence of a new pest can drastically reduce yields.

Constant monitoring, introduction of natural predators, and rotation of crops can help in managing pest infestations.

c) Weather Conditions:

Unpredictable weather patterns due to climate change can affect crop yields. For instance, unexpected frosts can damage crops like rice and tomatoes.

Diversifying crops, investing in weather prediction tools, and perhaps considering controlled environment agriculture (like greenhouses) can help in mitigating the adverse effects of changing weather patterns.

Risk Analysis:

Certainly! Creating a comprehensive risk profile involves analyzing both the current data from your dashboard and incorporating external factors that could influence farming operations over the next decade. Here's a risk profile based on the provided information and broader agricultural considerations:

1. Environmental Risks:

a) Weather Patterns:

: Climate change has led to unpredictable weather patterns. Droughts, unexpected frosts, heavy rainfall, or prolonged periods of extreme temperatures can threaten crop yields.

: Invest in weather prediction tools, diversify crop types to include drought or flood-resistant variants, and consider controlled environment agriculture.

b) Water Scarcity:

: As global water resources become strained, irrigation from traditional sources like rivers or ponds might become limited.

: Explore water-saving irrigation techniques, such as drip irrigation. Consider rainwater harvesting and the reuse of treated wastewater for irrigation.

2. Pest-Related Risks:

a) Increase in Pest Severity or New Pests:

: As seen in the dashboard, several pests can affect crops. The emergence of new pests or an increase in existing pest populations can reduce yields.

: Implement integrated pest management systems, use natural predators, and engage in crop rotation. Stay informed about global pest trends.

b) Pesticide Resistance:

: Over-reliance on specific pesticides can lead to pests developing resistance, rendering the pesticides ineffective.

: Rotate pesticide use, utilize organic pest control methods, and develop genetically resistant crop strains.

3. Economic Risks:

a) Market Fluctuations:

: Changes in consumer preferences, international trade policies, or oversupply can lead to reduced market prices for crops.

: Diversify crops, explore niche markets, and consider forward contracts or futures to lock in prices.

b) Increased Operational Costs:

: Costs of seeds, fertilizers, equipment, and labor might increase over the decade.

: Invest in technological innovations to optimize resource use and explore sustainable farming practices that reduce dependency on expensive inputs.

4. Resource-Related Risks:

a) Soil Degradation:

: Intensive farming can lead to soil degradation, reducing its nutrient content and fertility.

: Adopt sustainable farming practices like crop rotation, no-till farming, and organic fertilization to maintain soil health.

b) Labor Shortages:

: Dependence on manual labor can lead to shortages, especially during peak seasons.

: Invest in mechanization and automation, provide training and incentives for seasonal workers, and consider vertical farming where feasible.

5. Regulatory and Social Risks:

a) Changing Agricultural Policies:

: Governments might introduce new agricultural policies, taxes, or regulations that could affect operations.

: Stay updated on local and international agricultural policies. Engage in agricultural associations to lobby for favorable terms.

b) Social Awareness and Perception:

: Increasing consumer awareness might lead to demands for organic, non-GMO, or sustainably produced crops.

: Transition to sustainable farming practices, obtain necessary certifications, and actively engage in public relations and consumer education.

We analyzed AgResources Inc.'s farming data, specifically focusing on crop yields and environmental factors. While crop yields were found to be consistent across types, various environmental factors, such as weather conditions, source of irrigation, and pest severity, played significant roles in influencing these yields.

To ensure sustained profitability, it's essential to consider production costs, market demand, and the selling price of crops. Environmental factors, including changing weather patterns and water scarcity due to climate change, are major risks. Pests present another set of challenges, with the potential for increased severity, emergence of new pests, and pesticide resistance.

Economic challenges encompass market fluctuations and increased operational costs. The soil's health is paramount for continued crop yields, making soil degradation a concern. As farming relies heavily on labor, potential labor shortages present another risk. Finally, changing agricultural policies and shifting consumer perceptions towards sustainability and organic farming could affect farming operations.

In conclusion, while AgResources Inc. has consistent yields, a proactive approach to addressing environmental, economic, resource-related, and regulatory challenges is crucial for ensuring long-term profitability and sustainability.