Agricultural Economics: Analysis of Inputs, Outputs, Expenditures, and Price Spreads

March 9, 2025

Abstract

Agriculture is not only the backbone of global food production but also a crucial component of the economy. This report analyzes the relationships between agricultural inputs, outputs, food expenditures, and price spreads. Leveraging publicly available datasets from the USDA and utilizing advanced data processing and dashboarding techniques, the study provides actionable insights into the agricultural supply chain—from the farm gate to the consumer. The analysis supports predictive modeling for agricultural productivity and price forecasting, which can guide decision-making for farmers, investors, and policymakers.

1 Introduction

Agriculture plays a pivotal role in both global food production and economic stability. Understanding the dynamics of agricultural inputs, outputs, and price fluctuations is essential for optimizing farming practices, guiding economic policies, and informing investment decisions. This study employs data science techniques to decode the complex economic forces at play in agricultural markets. The report is structured following a scientific methodological approach: from data collection and preprocessing to exploratory analysis, dashboarding, and predictive modeling.

1.1 Motivation of the Work

The project is driven by the need to:

- Analyze relationships between agricultural inputs (e.g., labor, fertilizers, pesticides) and outputs (e.g., crop and livestock production).
- Understand food expenditure trends and price fluctuations, which reflect the broader economic impact of agricultural production.
- Identify and quantify the spread price between farm gate prices and retail prices, revealing inefficiencies in the supply chain.

By utilizing USDA datasets, this study aims to provide insights that can lead to optimized farming strategies, improved economic policies, and informed investment decisions.

2 Data Source

The analysis uses several publicly available datasets from USDA.gov, including:

- **Agricultural Output and Input Data**: Details inputs such as labor, fertilizers, and pesticides alongside outputs like crop and livestock production.
- **Agricultural Output Price Index Over Time**: A time-series dataset showing fluctuations in agricultural prices.
- **Normalized Output and Input Data**: Normalized values enabling cross-sector comparisons.

3 Data Understanding

The datasets are segmented into three main aspects:

3.1 Agricultural Output and Input Analysis

This dataset explores the relationship between the volume of inputs and the resulting outputs:

- **Labor Inputs**: Information on self-employed and unpaid family workers.
- **Intermediate Inputs**: Fertilizers, pesticides, and other purchased services.
- **Output Data**: Quantities of crops and livestock produced.

Data cleaning involves handling missing values and normalizing different scales before merging the datasets for a unified analysis.

3.2 Food Expenditure Analysis

This aspect examines consumer spending on agricultural products, highlighting:

- **Consumer Expenditures**: Breakdown by product type (e.g., livestock, crops, processed foods).
- **Food Inflation and Price Changes**: Time-series trends that link agricultural production with consumer prices.

Time-series analysis and regression techniques help establish correlations between production levels and price variations.

3.3 Spread Price Analysis

Spread price is the difference between the price received by farmers and the retail price paid by consumers. This analysis addresses:

- **Price Spread**: Measurement of the gap between farm and retail prices.
- **Supply Chain Efficiency**: Evaluation of processing, transportation, and retailing costs.

4 Interrelationships and the Bigger Picture

The three analysis aspects are interconnected:

- **Inputs & Outputs**: Efficient use of inputs directly impacts production costs and output volumes.
- **Food Expenditure**: Rising production costs due to inefficiencies lead to higher consumer prices.
- **Spread Price**: Highlights value addition and inefficiencies along the supply chain.

A change in one domain (e.g., increased fertilizer use) can influence the entire chain—from production costs to consumer pricing.

5 Business and Data Science Objectives

5.1 Agricultural Output and Input Analysis

Business Objective: Optimize production by identifying cost-effective input combinations.
Data Science Objective: Develop predictive models to optimize input use and maximize output.

5.2 Food Expenditure Analysis

Business Objective: Forecast food price trends to assist in budgeting and planning.
Data Science Objective: Apply time-series analysis and regression models to predict consumer price changes based on production data.

5.3 Spread Price Analysis

Business Objective: Uncover inefficiencies in the supply chain to reduce costs.
Data Science Objective: Build models to quantify the price spread and pinpoint cost reduction opportunities.

6 Data Processing and Dashboarding Techniques

To transform raw data into actionable insights, we employed various data processing and dashboarding techniques. Below are creative examples drawn from our project:

6.1 Data Processing Example

The following Python code snippet demonstrates data cleaning, merging, and the creation of a date column for time-series analysis:

Listing 1: Data Cleaning and Merging for Sales Data

```
import pandas as pd
import streamlit as st
```

```
# Load datasets
monthly_sales = pd.read_csv('monthly_sales.csv')
monthly_sales_by_outlet = pd.read_csv('monthly_sales_by_outlet.csv')

# Clean data function to remove commas and colons
def clean_data(df):
    for column in df.columns:
        if df[column].dtype == 'object':
            df[column] = df[column].str.replace(',', '').str.replace(':', '').astype(
                float, errors='ignore')
    return df

monthly_sales = clean_data(monthly_sales)
monthly_sales_by_outlet = clean_data(monthly_sales_by_outlet)

# Merge datasets on Year and Month
merged_data = pd.merge(monthly_sales, monthly_sales_by_outlet, on=['Year', 'Month'],
    how='outer')

# Create a Date column for time-series analysis
merged_data['Date'] = pd.to_datetime(merged_data['Month'] + ' ' + merged_data['Year'].
    astype(str))
st.dataframe(merged_data.head())
```

6.2 Dashboarding and Visualization Example

The next snippet demonstrates how to create an interactive dashboard using Streamlit and Plotly. Users can select specific columns to visualize trends over time:

Listing 2: Interactive Dashboard for Sales Trends

```
import matplotlib.pyplot as plt

st.markdown("### Visualizing Sales Trends")
# Identify columns that include the term 'sales'
columns_to_plot = [col for col in merged_data.columns if 'sales' in col.lower()]
selected_columns = st.multiselect("Select columns to visualize", columns_to_plot)

if selected_columns:
    fig, ax = plt.subplots(figsize=(10, 6))
    for column in selected_columns:
        ax.plot(merged_data['Date'], merged_data[column], marker='.', linestyle='-',
            label=column)
    ax.set_title("Sales Trends Over Time")
    ax.set_xlabel("Date")
    ax.set_ylabel("Sales (in millions)")
    ax.legend()
    ax.grid(True)
    st.pyplot(fig)
else:
    st.write("Please select at least one column to plot.")
```

6.3 Predictive Modeling and Price Spread Analysis Example

This snippet outlines how a linear regression model is implemented to predict price trends and analyze the spread between retail and farm prices:

Listing 3: Predictive Modeling and Price Spread Analysis

```
import plotly.express as px
from sklearn.linear_model import LinearRegression

# Load dataset for price spread analysis
dataset1 = pd.read_csv('table01a.csv')

# Filter data for Price and Quantity attributes
price_data = dataset1[dataset1['Attribute'].str.contains('Price')]
quantity_data = dataset1[dataset1['Attribute'].str.contains('Quantity')]

# Compute rolling average of price (5-year window)
price_data['Rolling_Avg'] = price_data.groupby('Attribute')['Value'].rolling(window=5)
    .mean().reset_index(0, drop=True)

# Calculate Year-over-Year Price Change
price_data['Price_Change'] = price_data.groupby('Attribute')['Value'].pct_change() *
    100

# Build Linear Regression Model
X = price_data['Year'].values.reshape(-1, 1)
y = price_data['Value'].values
model = LinearRegression()
model.fit(X, y)
slope, intercept = model.coef_[0], model.intercept_

# Create interactive Plotly dashboard for trend analysis
trend_line_fig = px.scatter(price_data, x='Year', y='Value', title="Price␣Trend␣with␣
    Regression␣Line")
trend_line_fig.add_scatter(x=price_data['Year'], y=model.predict(X), mode='lines',
    name='Trend␣Line')
st.plotly_chart(trend_line_fig)

# Example of Price Spread Calculation for Flour Data
flour_data = pd.read_csv('flour.csv')
flour_data['Price_Spread'] = flour_data['Retail␣Price-Dollars'] - flour_data['Farm␣
    Value-Dollars']
st.write("Price␣Spread␣Analysis␣for␣Flour:")
st.dataframe(flour_data[['Year', 'Price_Spread']])
```

7 Insights and What We Gain from Understanding the Data

Analyzing these datasets provides the following insights:

- **Farmers**: Optimized resource allocation and improved yield predictions.
- **Investors**: Better forecasts of agricultural market trends, guiding investment decisions.
- **Consumers**: A deeper understanding of food price inflation and supply chain inefficiencies.

The integration of data processing and dashboarding techniques transforms raw datasets into actionable intelligence, enabling stakeholders to make informed decisions.

8 Conclusion

This report presents a comprehensive analysis of agricultural inputs, outputs, food expenditures, and price spreads. By integrating advanced data processing methods with interactive dashboarding techniques, we have uncovered key insights into the agricultural supply chain. Efficient input utilization leads to cost-effective production, which in turn influences consumer prices and reveals supply chain inefficiencies. The methodologies and insights presented here form a framework for further research and policy development in agricultural economics, ultimately benefiting farmers, businesses, and consumers alike.