Help the Ag Industry

Aim:

To understand trends in APMC (Agricultural produce market committee)/mandi price & quantity arrival data for different commodities in Maharashtra.

Objective:

1. Test and filter outliers.

2. Understand price fluctuations accounting the seasonal effect.

    1. Detect seasonality type (multiplicative or additive) for each cluster of APMC and commodities.

    2. Deseasonalize prices for each commodity and APMC according to the detected seasonality type.

3. Compare prices in APMC/Mandi with MSP (Minimum Support Price)- raw and deseasonalised.

4. Flag set of APMC/mandis and commodities with highest price fluctuation across different commodities in each relevant season, and year.

Methodology

Filtering outliers

Following filters were applied to remove outliers from the dataset given –

- max_price > modal_price > min_price
- max_price > 0 & modal_price > 0
- max_price < 1000000

Structuring data for analysis

- In order to analyse given market price data, we slice it on the basis APMC and commodities.
- As a result, of the 349 APMCs and 204 different commodities, we form 4826 non empty dataframes, each containing unique APMC and commodity combination for analysis.
- These individual dataframes form a dictionary of dataframes, where each dataframe is associated to a unique integer value from 0 to 4825.

Understanding seasonality

- In order to determine seasonality of market price data, auto seasonal decompose function from Python' s statistical modelling library is used.
- Initial exploratory analysis shows that min_price, max_price and modal_price follow similar price trends, as seen in price trends graph below.
- As such, seasonal decompose is run only on modal_price series to avoid unnecessary computation requirements.
- Further, data is tested to seasonality tests of both additive and multiplicative models.
- Since the time period of observations is little over 24 months or 2 years, the frequency parameter in analysis is set as 2.
- If the is not perfectly seasonal, i.e. with mean of seasonal variations being 0, over the duration of observations, the data is deemed not to be seasonal in that model.
- modal_price data which fails seasonality test with both multiplicative and additive models is deemed to be not seasonal.
- If seasonality is found in data, deseasonalized modal price and peak seasonal price fluctuation from mean price is also calculated.

Results

- All dataframes have been saved as CSV files in the Output folder, with APMC and commodity in filename.
- The key trends in price data can be seen below for a few combinations of APMCs & commodities.