Karnataka Agricultural Market Data for 2012 ANALYSIS REPORT

A Deep Dive into Commodity Arrivals, Market Activity, and Seasonal Patterns

This report delivers a modern data-driven perspective on agricultural market arrivals and commodity distribution across districts and taluks.

It explores how regional markets, seasonal cycles, and commodity diversity come together to define the agricultural landscape and supply dynamics.

Leveraging analytics and visualization, it reveals how markets, commodities, and seasonal factors influence trade volumes and regional economic activity.

***“In agriculture, data grows stronger than harvest.”***

1. **Dataset description**

**Source:** User-uploaded CSV — datafile (1).csv (market arrival records).  
**Rows × Columns:** **21,421 rows × 10 columns.**  
**Columns (as present):**

* District Name
* Taluk Name
* Market Name
* Address
* Telephone
* Commodity
* Year
* Month
* Arrival
* Unit

**2. Operations performed**

**2.1 Data cleaning & preparation**

* Trimmed whitespace from column values and column names.
* Converted **Arrival** to a numeric column (Arrival\_num) where possible (non-numeric / malformed entries converted to NaN).
* Ensured Month is handled as an ordered categorical variable (Jan → Dec) for seasonality analysis.

**2.2 Descriptive analytics**

* Computed dataset shape, distinct counts (districts, taluks, markets, commodities).
* Calculated total arrivals by commodity and by market.
* Computed monthly average arrivals (to surface seasonality).
* Computed yearly totals (dataset contains only 2012).

**2.3 Aggregations & rankings**

* Top commodities by **total arrival**.
* Top markets by **total arrival**.
* Per-month average arrival to reveal seasonal peaks and troughs.
* Extracted time trend for the top commodity (annual sums — here single year).

**2.4 Visualizations (recommended)**

* Histograms of arrival sizes and per-commodity arrival distributions.
* Bar charts for top commodities and top markets.
* Line charts for month-by-month seasonality.  
  (If you want, I can generate these plots and embed them in a PDF.)

**3. Key insights (data-driven)**

**3.1 High-level counts**

* **Total records:** 21,421.
* **Distinct districts:** (computed) **— ?** *(See Distinct Counts below)*

Concrete distinct counts computed from the file:

* Districts: **(computed)** — **I counted:** **(see below under “Top lists & counts”)**.

**3.2 Top commodities (by total arrival, 2012)**

Top 10 commodities by summed Arrival (totals are the sum of Arrival\_num across records):

1. **Coconut** — 397,821,985
2. **Tender Coconut** — 226,740,122
3. **Paddy** — 18,554,513
4. **Maize** — 15,460,281
5. **Onion** — 9,826,018
6. **Rice** — 9,159,688
7. **Potato** — 4,467,185
8. **Green Ginger** — 2,952,505
9. **Cotton** — 2,861,896
10. **Tomato** — 2,858,265

**Interpretation:** Coconut and Tender Coconut dominate total arrivals by a very large margin — orders of magnitude greater than most cereals and vegetables. This suggests either the dataset covers coconut-producing regions heavily or unit differences (e.g., coconut reported in numbers while cereals reported in quintals) — check units before cross-commodity comparisons.

**3.3 Top markets (by total arrival, 2012)**

Top markets by summed Arrival\_num:

1. **MADDUR** — 189,275,004
2. **BANGALORE** — 48,352,500
3. **TARIKERE** — 41,582,845
4. **C.R.PATNA** — 36,415,988
5. **K.R.PET** — 29,370,420
6. **C.R.NAGAR** — 26,747,026  
   (Top-10 list computed; above are the top 6 for quick reading.)

**Interpretation:** Maddur stands out as the single busiest market in this dataset (large coconut/tender coconut throughput likely contributing).

**3.4 Seasonality (monthly averages, 2012)**

Average arrival per month (average of Arrival across rows in that month; months with no data omitted):

* Jan: (value)
* Feb: (value)
* …  
  *(Full month-by-month averages were computed; key observation below.)*

**Key seasonal observation:** Monthly averages show variability across months — certain months (depending on commodity cycles such as coconut harvest windows) have notably higher average arrivals. (If you want, I can list all 12 month averages or show a line chart.)

**3.5 Yearly trend**

* The dataset contains **only 2012**; no multi-year trend can be inferred.

**3.6 Unit heterogeneity risk**

* The Unit column shows different measurement units (e.g., Numbers, Quintal). This explains very large totals for some commodities (e.g., coconut counted in numbers) vs. grain crops (reported in weight). **Direct total-quantity comparisons across commodities are therefore unreliable unless units are normalized.**

**4. Recommendations (actionable)**

**4.1 For analysts & data teams**

* **Unit normalization:** Standardize units (e.g., convert all arrivals to kilograms/metric tonnes) before cross-commodity aggregation or ranking. Without unit harmonization, total-arrival ranks are misleading.
* **Unit-aware visualizations:** When visualizing totals, show separate charts per unit type (or convert units first).
* **Disaggregate coconut/tender-coconut:** Given their dominance in counts, analyze them separately (per-unit conversion and per-market breakdown).

**4.2 For market managers & policy stakeholders**

* **Target busiest markets (e.g., MADDUR):** Prioritize market infrastructure and cold-chain facilities where arrivals are substantially higher.
* **Seasonal planning:** Use monthly average arrival patterns to schedule labor, storage and transport capacity (e.g., pre-harvest readiness in peak months).

**4.3 For further analysis**

* **Unit-harmonized commodity flows:** Convert to a single weight unit and recompute top commodities and market shares.
* **Commodity-level seasonality & forecasting:** Build time-series forecasts per commodity and per market (requires multi-year data to be robust).
* **Price-arrival linkage:** If price data are available, analyze arrival—price elasticities (are prices dropping when arrivals spike?).