

## Le calendrier pèse plus que le programme

Le maïs médian vaut 75 gourdes à la récolte de décembre et 130 au pic de juillet. Comparez deux passages d'enquête pris à des points différents de ce cycle et vous avez mesuré le calendrier — et le panel d'un district a changé de composition exactement quand cela comptait.

---

Pierre Robentz Cassion

Leçon 5 sur 8

Analyse de la sécurité alimentaire et IPC — Le marché est aussi un résultat

## Ce que couvre cette leçon

- Une ligne de preuve indépendante
- Nettoyez trois choses avant tout tracé
- Le cycle saisonnier
- Le panel qui a changé de composition
- Rapportez la série avec son panel
- La suite

# Une ligne de preuve indépendante — En Python

```
import pandas as pd

prices = pd.read_csv("market-prices-2024.v1.csv")
print(f"{len(prices):,} observations")
print(f"{prices['market_id'].nunique()} markets, "
      f"{prices['commodity'].nunique()} series, "
      f"{prices['period'].nunique()} months")
```

## Une ligne de preuve indépendante — En R

```
library(dplyr)
```

```
prices |> summarise(n = n(), markets = n_distinct(market_id),  
                  series = n_distinct(commodity), months = n_distinct(period))
```

## Nettoyez trois choses avant tout tracé — En Python

```
prices["price_htg"] = pd.to_numeric(prices["price_htg"])

dollars = prices["price_htg"] < 10
marmite = (prices["commodity"] == "maize") & (prices["market_id"] == "MK005")
thin = prices["trader_quotes"] == 1

print(f"dollar entries: {dollars.sum()}")
print(f"marmite market maize rows: {marmite.sum()}")
print(f"single-quote observations: {thin.sum()}")
```

## Nettoyez trois choses avant tout tracé — En R

```
prices |> summarise(  
  dollars = sum(price_htg < 10),  
  thin = sum(trader_quotes == 1)  
)
```

## Nettoyez trois choses avant tout tracé

- Huit prix sont en dollars — avec la colonne de devise indiquant toujours HTG
- Un marché relève le maïs à la marmite — mesure de volume d'environ 2,7 kg, et le classe dans une colonne libellée au. . .

## Nettoyez trois choses avant tout tracé — En Python

```
by_market = prices[prices["commodity"] == "maize"].groupby("market_id")["price_htg"].median()  
print(by_market.round(1).sort_values())
```

## Nettoyez trois choses avant tout tracé — En R

```
prices |> filter(commodity == "maize") |>  
  summarise(median = median(price_htg), .by = market_id) |> arrange(median)
```

## Nettoyez trois choses avant tout tracé

- L'erreur d'unité est invisible dans un marché et évidente sur douze — C'est la règle générale du décalage d'unité : il...
- Treize observations reposent sur une seule cotation — Ce ne sont pas des erreurs

## Le cycle saisonnier — En Python

```
clean = prices[~dollars & ~marmite]
maize = clean[clean["commodity"] == "maize"]
series = maize.groupby("period")["price_htg"].median()
print(series.round(0))
```

```
prices |>  
  filter(commodity == "maize", price_htg >= 10, market_id != "MK005") |>  
  summarise(median = median(price_htg), .by = period) |> arrange(period)
```

## Le cycle saisonnier

| Mois                            | 2023       | 2024       |
|---------------------------------|------------|------------|
| Janvier                         | 66         | 70         |
| Avril                           | 97         | 110        |
| <b>Juillet (pic de soudure)</b> | <b>130</b> | <b>159</b> |
| Octobre                         | 92         | 104        |
| <b>Décembre (récolte)</b>       | <b>75</b>  | <b>81</b>  |

- Le maïs coûte 74 % de plus au pic de juillet qu'à la récolte de décembre en 2023, et 96 % de plus en 2024 — C'est. . .
- La conséquence pour un plan d'enquête est directe — Une situation de référence en décembre et une évaluation finale en. . .

```
peaks = series.loc[["2023-07", "2024-07"]]  
print(f"lean peak 2023: {peaks['2023-07']:.1f}")  
print(f"lean peak 2024: {peaks['2024-07']:.1f}")  
print(f"year on year: {peaks['2024-07'] / peaks['2023-07'] - 1:+.1%}")
```

## Le cycle saisonnier — En R

*# Same point in the season, one year apart. Everything else is the calendar.*

- 130,1 contre 159,3 — une hausse de 22 % au même point de la saison — *Voilà* la dégradation, et elle vaut le quart de...

## Le panel qui a changé de composition — En Python

```
coverage = maize.pivot_table(index="period", columns="market_id",  
                              values="price_htg", aggfunc="size")  
print(coverage.isna().sum(axis=1)[lambda s: s > 0])
```

## Le panel qui a changé de composition — En R

```
prices |> filter(commodity == "maize") |> count(period, market_id) |>  
  count(period) |> filter(n < 11)
```

## Le panel qui a changé de composition

- MK001 cesse de rapporter de juin à septembre 2024 — quand sa route est coupée

## Le panel qui a changé de composition — En Python

```
nord = maize[maize["district"] == "Nord-Ouest"]
reported = nord.groupby("period")["price_htg"].mean()

balanced = nord[nord["market_id"] != "MK001"].groupby("period")["price_htg"].mean()
print(pd.DataFrame({"reported": reported, "balanced": balanced}).round(1)
      .loc["2024-04":"2024-10"])
```

## Le panel qui a changé de composition — En R

```
prices |>
  filter(commodity == "maize", district == "Nord-Ouest") |>
  summarise(reported = mean(price_htg), .by = period)
```

## Le panel qui a changé de composition

| Mois              | Ceux qui ont rapporté | Les deux ayant rapporté chaque mois |
|-------------------|-----------------------|-------------------------------------|
| Mai 2024          | 147,0                 | 123,2                               |
| Juin 2024         | 172,1                 | 172,1                               |
| <b>Mai à juin</b> | <b>+17 %</b>          | <b>+40 %</b>                        |

## Le panel qui a changé de composition

- La série rapportée sous-estime de plus de moitié la hausse de soudure — et chacun de ses nombres est une moyenne. . .
- Employez un panel équilibré — seulement les marchés présents à chacun des mois comparés — ou chaîne des variations. . .

## Rapportez la série avec son panel — Exemple

Maize price, gourdes per kg, median across markets

|                                     |      |                      |      |
|-------------------------------------|------|----------------------|------|
| 2023 harvest (Dec)                  | 75   | 2024 harvest (Dec)   | 81   |
| 2023 lean peak (Jul)                | 130  | 2024 lean peak (Jul) | 159  |
| Seasonal swing                      | +74% | Seasonal swing       | +96% |
| Year on year at the lean peak: +22% |      |                      |      |

Balanced panel of 10 markets throughout. MK001 (Nord-Ouest) did not report June to September 2024 and is excluded from all periods; MK005 records maize by the marmite and is excluded from maize.

8 dollar-denominated entries converted at 132 HTG; 13 single-quote observations retained and flagged.

- Un prix dit ce que la nourriture coûtait.

Lire la leçon complète, avec le code exécutable

```
https://data-analysis.cassion.dev/fr/cours/food-security-ipc/  
fs-05-price-series-and-seasonality
```