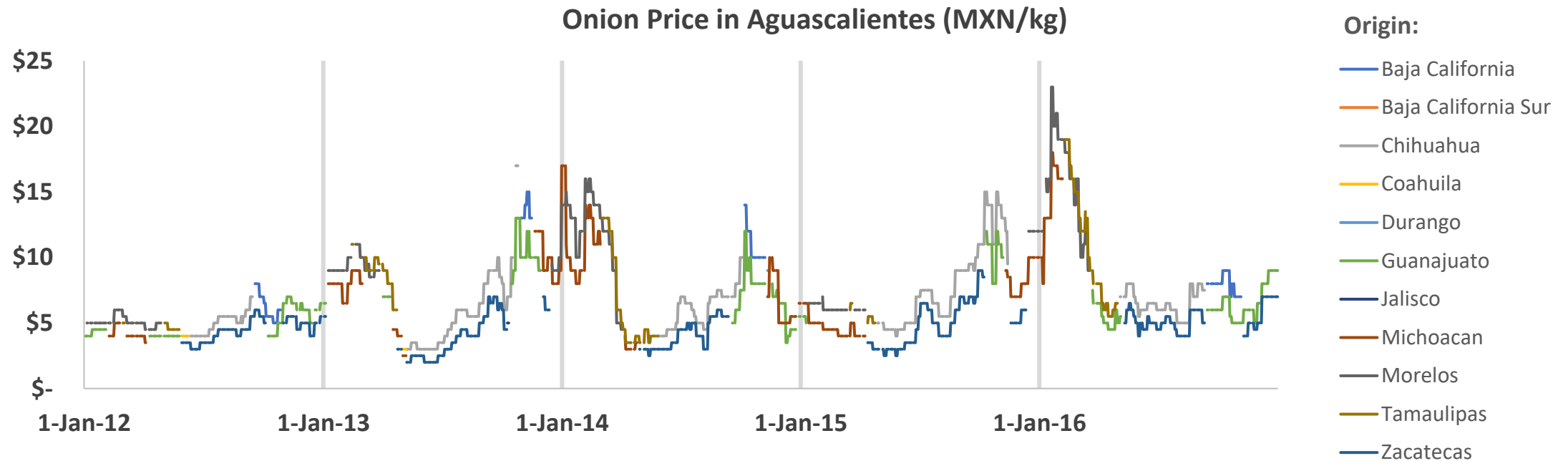# Onion Analytics

Business Analytics

APRIL 26, 2017

Ramsey Hamadeh
Hassan Ahmed Mortagy
Daniel Mac Gregor
Kavana Rudresh
Ramon Rodriganez Riesco

Farmers are exposed to spot market prices...

"Most Mexican onion producers sell their crops in the wholesale market. Since prices can change greatly in the short term, there is a lot of uncertainty for producers, and a year's work depends on selling when the price is right."

-**Ernesto Grossmann, Mexican onion producer**

Onion Price in Aguascalientes (MXN/kg)

Origin:
- Baja California
- Baja California Sur
- Chihuahua
- Coahuila
- Durango
- Guanajuato
- Jalisco
- Michoacan
- Morelos
- Tamaulipas
- Zacatecas

...which provides an incredible opportunity for improvement

Data*:

◦ Yearly production of onions: **1.3 million tons**

Hypothesis:

◦ Optimal price arbitrage over 28 days: **1.4 pesos / kg**

◦ Switching **5% of production** will not affect prices (i.e. farmers will continue being price takers)

◦ We will charge **30% of value generated**
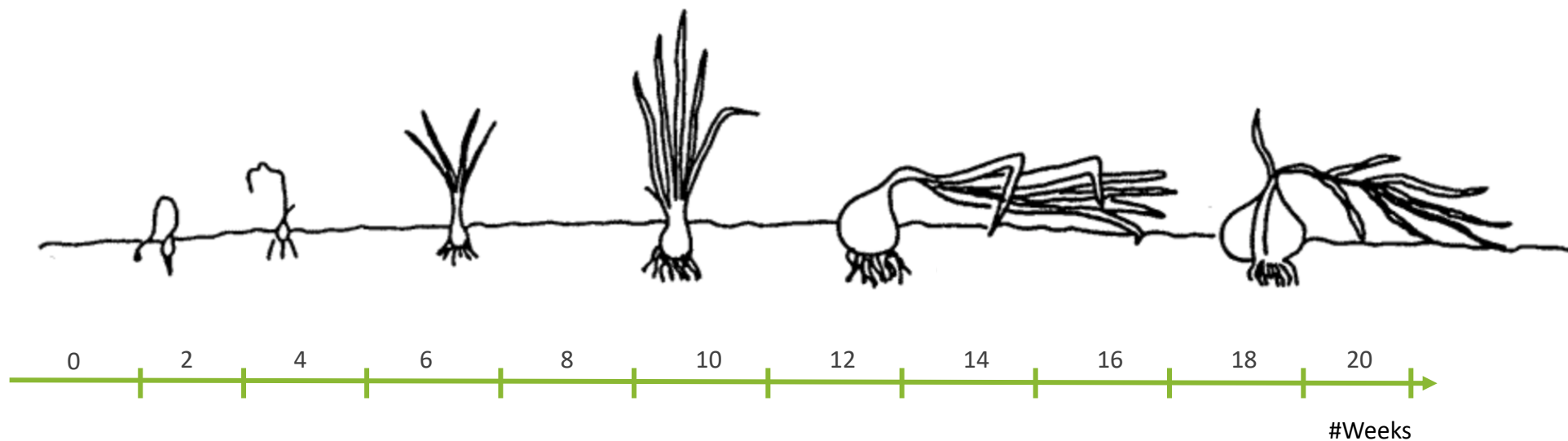
TAM = 5% Yearly Prod x Price Spread x 30%

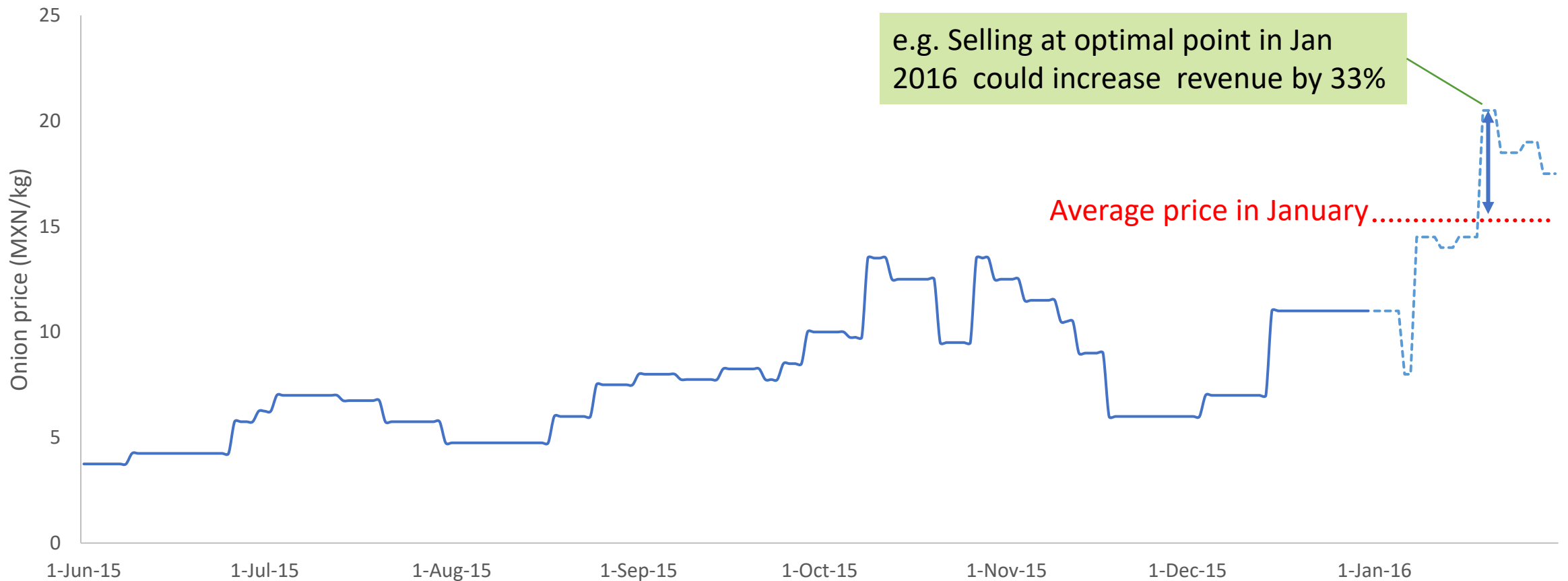An onion farm in Mexico

TAM = $1.4 Million

Some facts about Onions!

- The cycle from planting to harvesting is around 18 weeks

- Produced year-round in Mexico (March, May and August produce the most)

- Disease can affect crops if soil remains wet for more than 2 weeks

- Onions can be stored for a few weeks without loss of properties (**28 days is normal**)

- Typical yield 30 tons/ha

0   2   4   6   8   10   12   14   16   18   20

#Weeks

Value Idea: Selling Onions at Optimal Price

Predict optimal price during next month

e.g. Selling at optimal point in Jan 2016 could increase revenue by 33%

Average price in January ....

Onion price (MXN/kg)

25

20

15

10

5

0

1-Jun-15   1-Jul-15   1-Aug-15   1-Sep-15   1-Oct-15   1-Nov-15   1-Dec-15   1-Jan-16

Web-scraping and Data Cleaning involved..

WEATHER DATA

Source: National centers for environmental information at NOAA (US validated data source)

Data gathering: requests at their website (2 days to receive)

Raw data: Daily Min, Max and Average Temperatures. Precipitation for selected weather stations

Features created: over the plant to harvest period:
- Maximum Temperature
- Minimum Temperature
- Average Temperature
- Humid for > 14 days
- Max. # days without rain
- Average precipitation
- …

HISTORICAL ONION PRICES

Source: Ministry of economics in Mexico (validated source that gathers daily prices: ,max, min and average)

Data gathering: web scrapping with Python (BeautifulSoup)

Raw data: Daily price for white onions in all major wholesale markets in Mexico

Features created:
- Mean, Min and Max prices 28 days ago.
- Mean price of the origin 1 year before
- Mean price of the origin 2 years before
- Change in Prices over 1 week (starting 28 days ago)
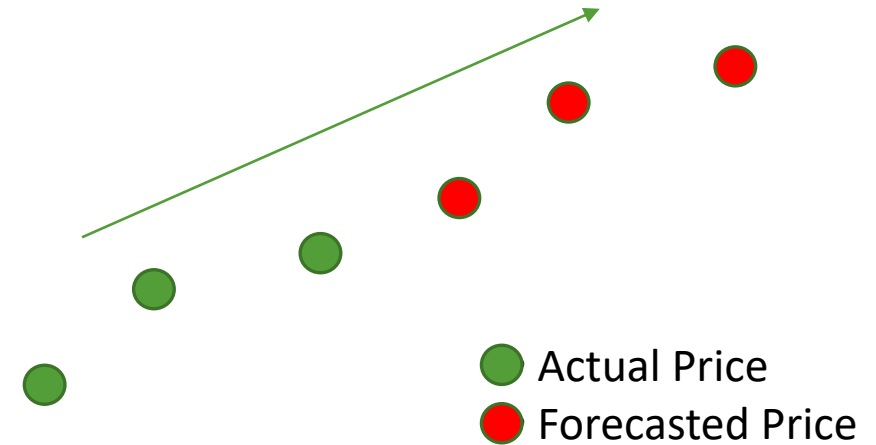- Change in Prices over 1 month (starting 28 days go).
- …

A TOTAL OF 22 FEATURES WERE CHOSEN

Our two approaches: Method 0 vs Method 1

METHOD 0

- **1 day ahead** price prediction
- **Compounded error** with time
- Unique trend: **no maximum** / minimum
- **Predictors**: last **4 days** (TSA)

● Actual Price
● Forecasted Price

METHOD 1

- **28 days ahead** prediction
- Next 28 days with **same error**
- Shows **variability** with maximums and minimums
- **Predictors** based on **crop** features

CHOSEN METHOD

Method 1 – Compare MSE different models

Algorithm	MSE (Test set)
Linear Regression	4.51
Lasso	3.98

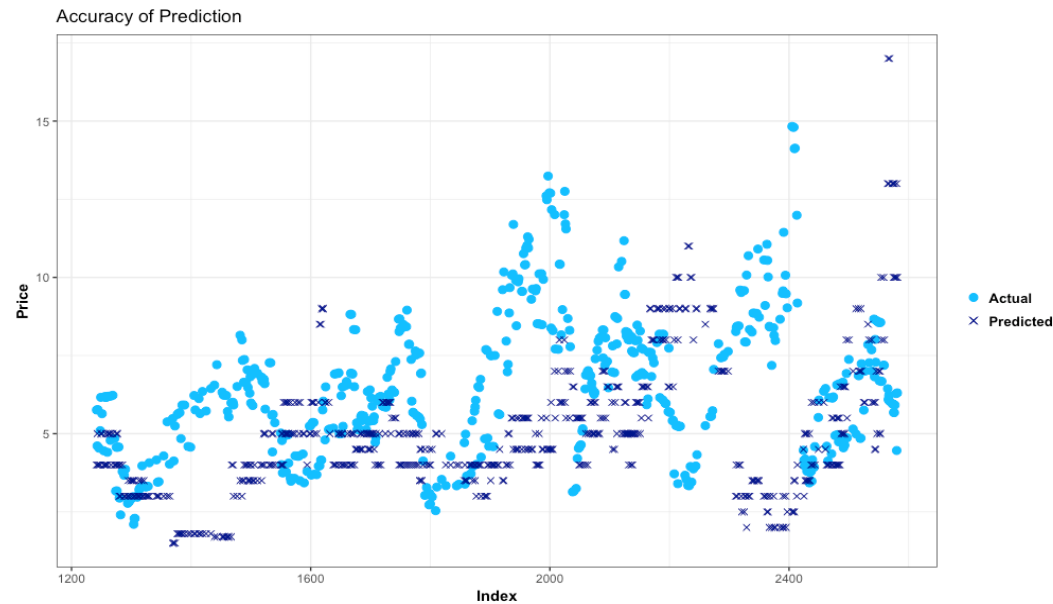| Algorithm	MSE (Test set)
Support Vector Machine (SVM)	0.395
Neural networks	0.909

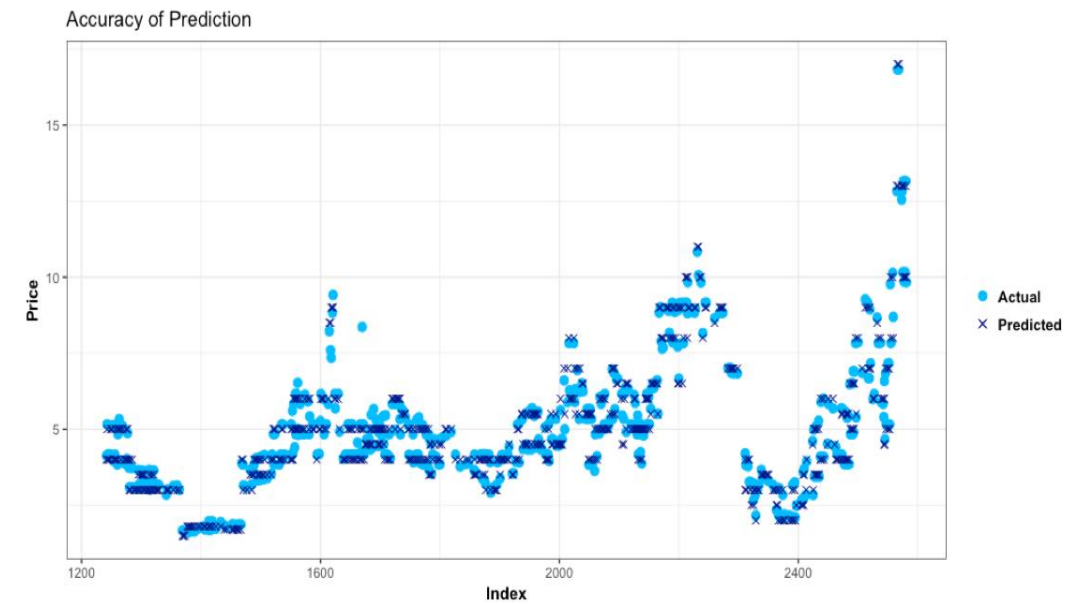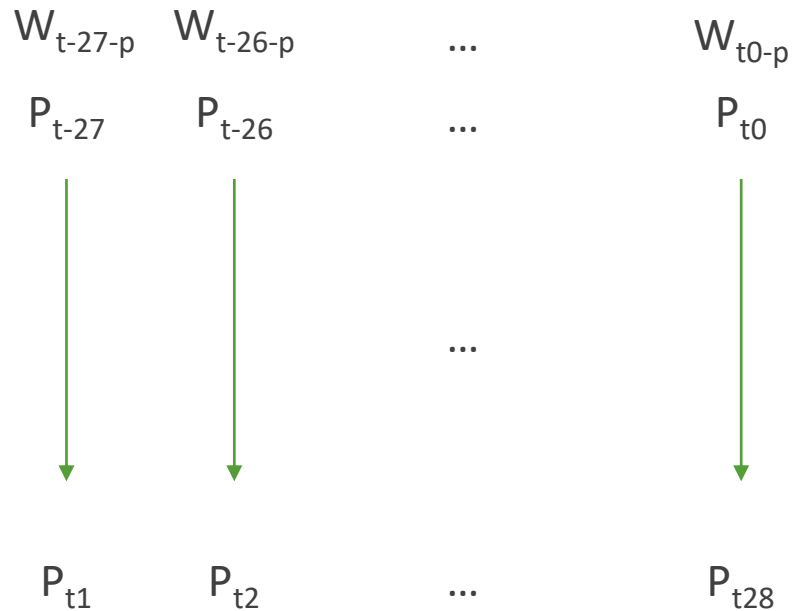Figure: Lasso predictions vs actual price


Figure: SVM predictions vs actual price

Model – Final Tool

EACH DAY t0 WE COMPUTE THE PREDICTION FOR t+28

W_{t-27-p} W_{t-26-p} ... W_{t0-p}

P_{t-27} P_{t-26} ... P_{t0}

...

P_{t1} P_{t2} ... P_{t28}

| Zones Prediction Matrix | ➕ | Data 2011-2014 | ➕ | SVM Trained 2011-2015 |
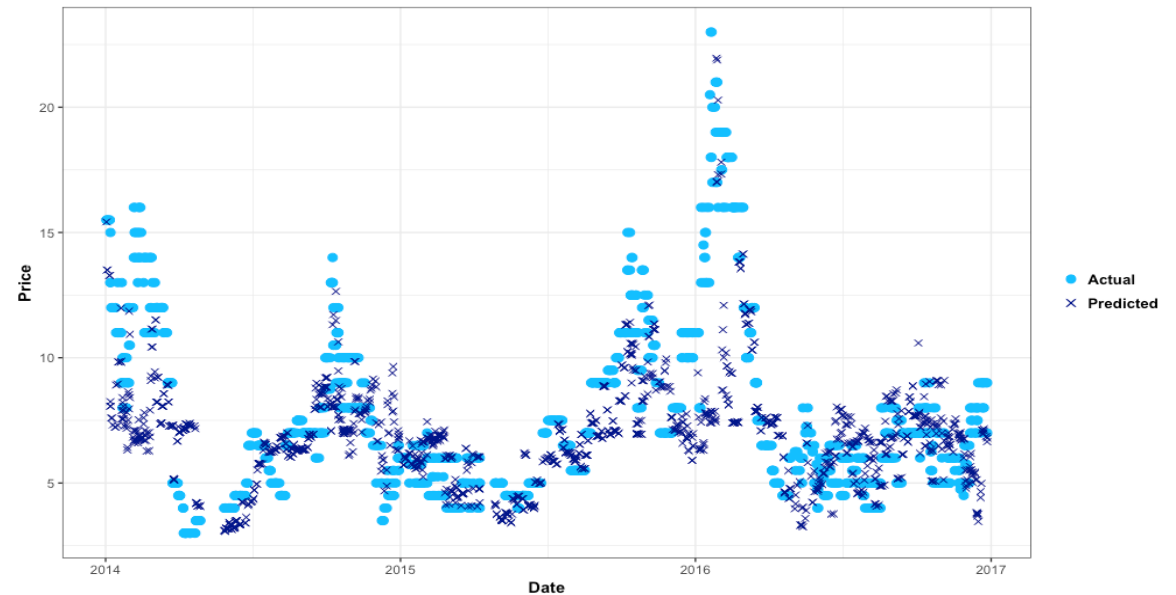
HAVING THE NEXT 28 DAYS PREDICTION WE CAN CHOOSE THE BEST POSSIBLE DAY

Potential profit is $31k/y for a mid-sized farmer

Price Aguascalientes (MXN/kg)

— Optimal seller   — Random seller
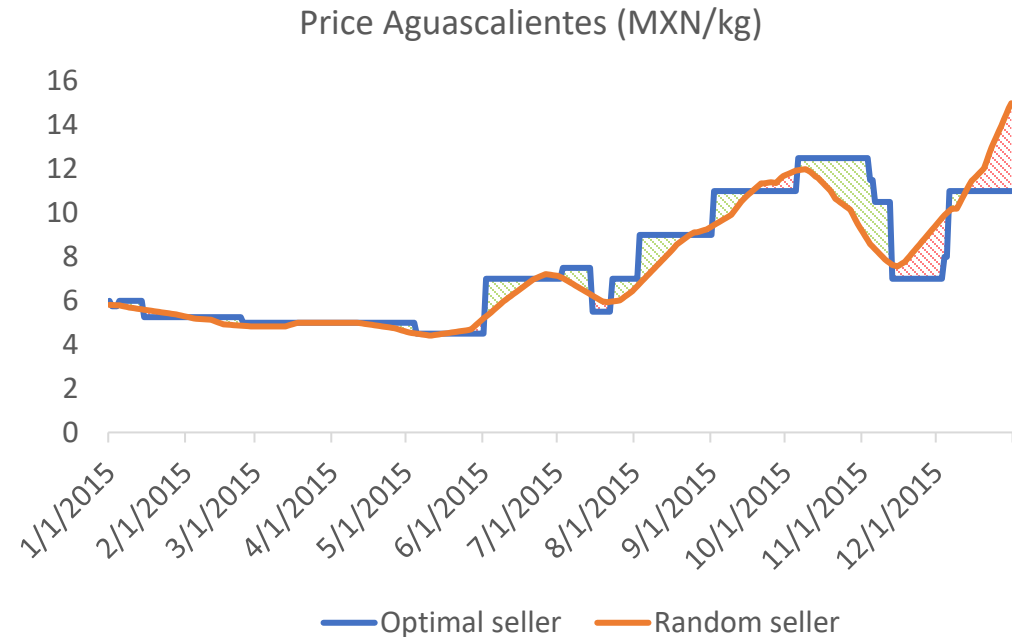
Business Value

We ran the model every day from **Jan 2014 to Dec 2016**

For each day we compared if the farmer sold at the model's optimal date within 28 day period vs selling in a random day on the same period.

On average, a farmer using the model gains **0.3 MXN/kg**

A farm yielding 35 ton/ha would **profit $310 per ha with model**

	Date	Mean_daliy_price	Avg_next_30	Best_forecast	forecast	model_sale_price	value
76	2015-03-17	5.00	5.000000	4	7.054787	5.75	0.75000000
77	2015-03-18	5.75	4.976190	3	7.370739	5.75	0.77380952
78	2015-03-19	5.75	4.937500	2	7.381049	5.75	0.81250000
79	2015-03-20	5.75	4.894737	1	7.389639	5.75	0.85526316

Example:

Farmer harvests March 17th:
- **Randomly** sell her produce in the next month: **Sells at 5 MXN/kg**
- **With model**, waits 4 days for max price: **Sells at 5.75 MXN/kg**
- **Value created** for farmer= **0.75 MXN/kg**

The Value of Onion Analytics

Value OA = 5% Yearly Prod x Price Spread x OA fee

= 5% * 1.3 million tons * 15.9 USD/ton * 30%

= **$310k**

We are Onion Analytics

Thanks for your Attention!

Questions?