TABLEAU CONFERENCE

TABLEAU CONFERENCE

Welcome

TABLEAU CONFERENCE

Using R and Tableau at Worthington Industries:
Price optimization for high-mix, low-volume environments

Steve Bartos

Manager, Predictive Analytics

Worthington Industries

Wil Davis

Analyst, Predictive Analytics

Worthington Industries

D.R.E.A.M.

Agenda

TABLEAU CONFERENCE

Why pricing? *A Brief History of Analytics at Worthington Industries*

A Machine Learning ~~Approach~~ Misfire.

Power to the People! *Let them choose with Tableau and R*

The Complete Stack. *Rserve, CI/CD and Version Control*

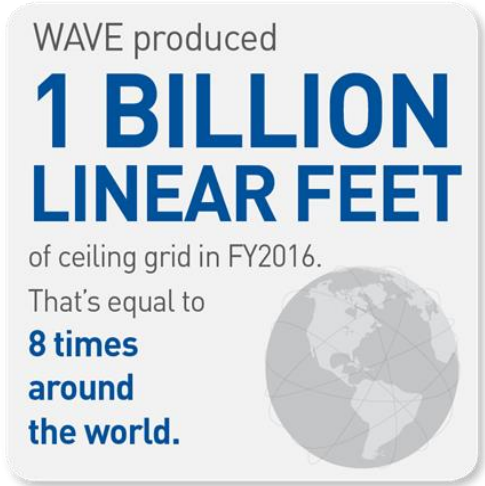
TABLEAU CONFERENCE

A Brief History of Analytics at Worthington Industries

Worthington Industries

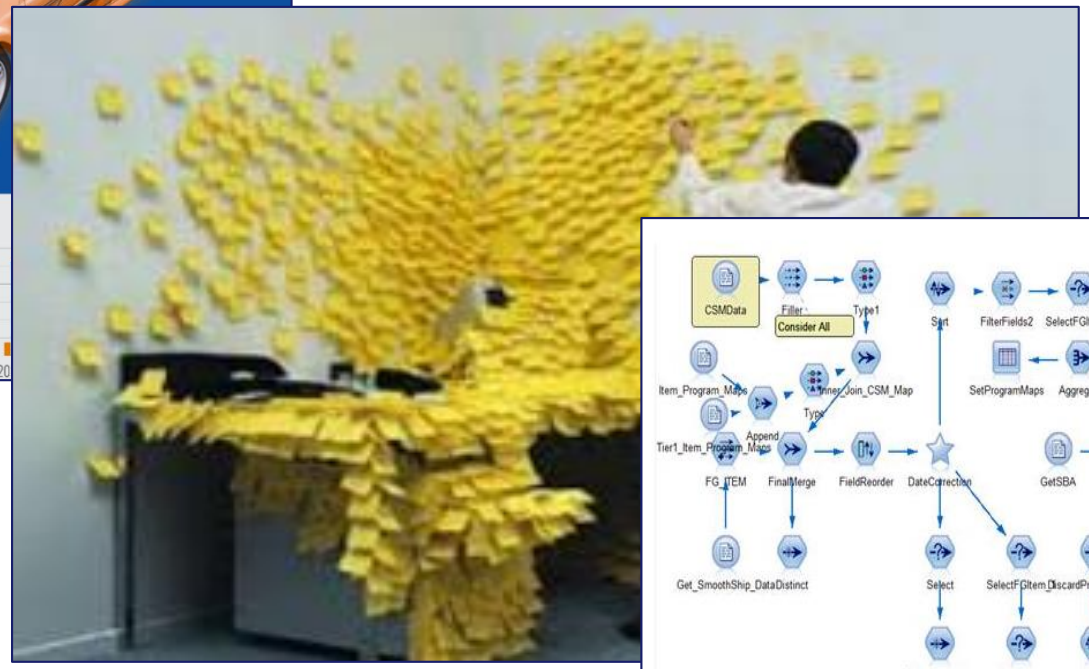
TABLEAU CONFERENCE

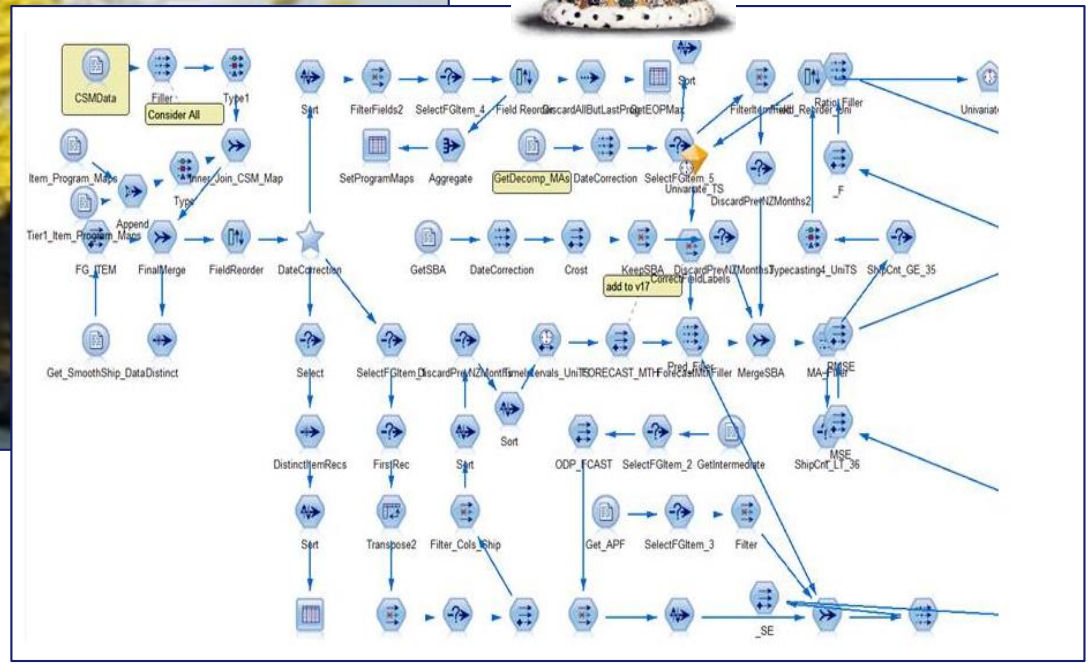
- ✓ Founded in 1955 and headquartered in Columbus, OH
- ✓ Publicly traded on the NYSE under the ticker WOR
- ✓ 10,000 employees & 5,000 customers; 80 facilities in 11 countries
- ✓ Employee, customer, supplier and investor-centered philosophy
- ✓ Leader in safety management and injury prevention – company wide goal of zero accidents and injuries
- ✓ Named one of "America's Safest Companies" by Occupational Hazards magazine, 2008
- ✓ Named to Fortune's "100 Best Companies to Work For" list five times



In the beginning: one analyst, one idea

TABLEAU CONFERENCE





Opportunities for another crown jewel of analytics?

TABLEAU CONFERENCE

Key Questions Addressed by Analytics

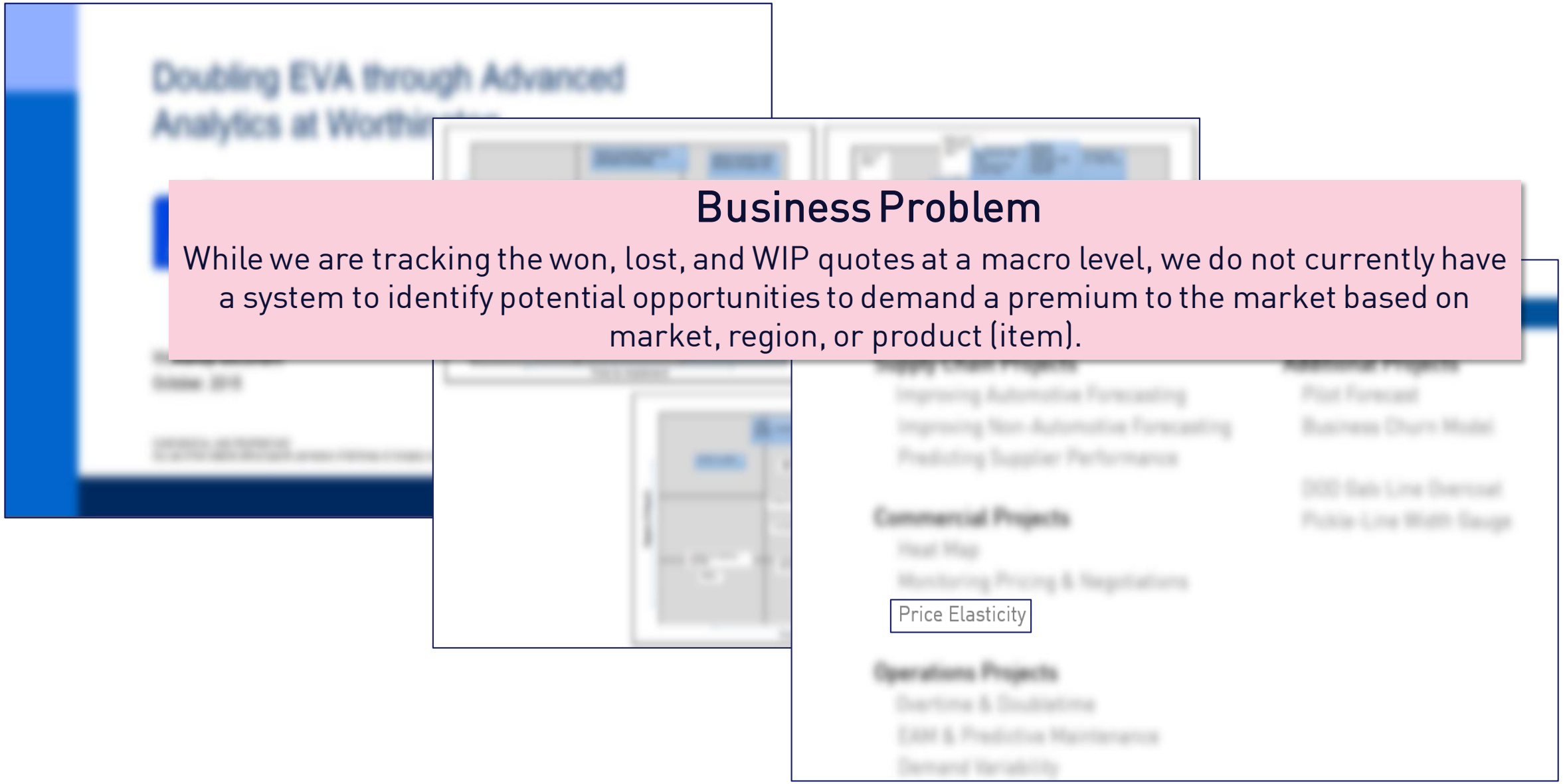
TABLEAU CONFERENCE

	Past	Present	Future
Information	**What happened?** Reporting	**What is happening?** Alerts	**What will happen?** Extrapolation
Insight	**How and why did it happen?** Modeling	**What's the best action?** Optimization	**What's the best/worst that can happen?** Prediction

Source: Davenport et al. *Analytics at Work: Smarter Decisions, Better Results*

The Price Elasticity Project

TABLEAU CONFERENCE



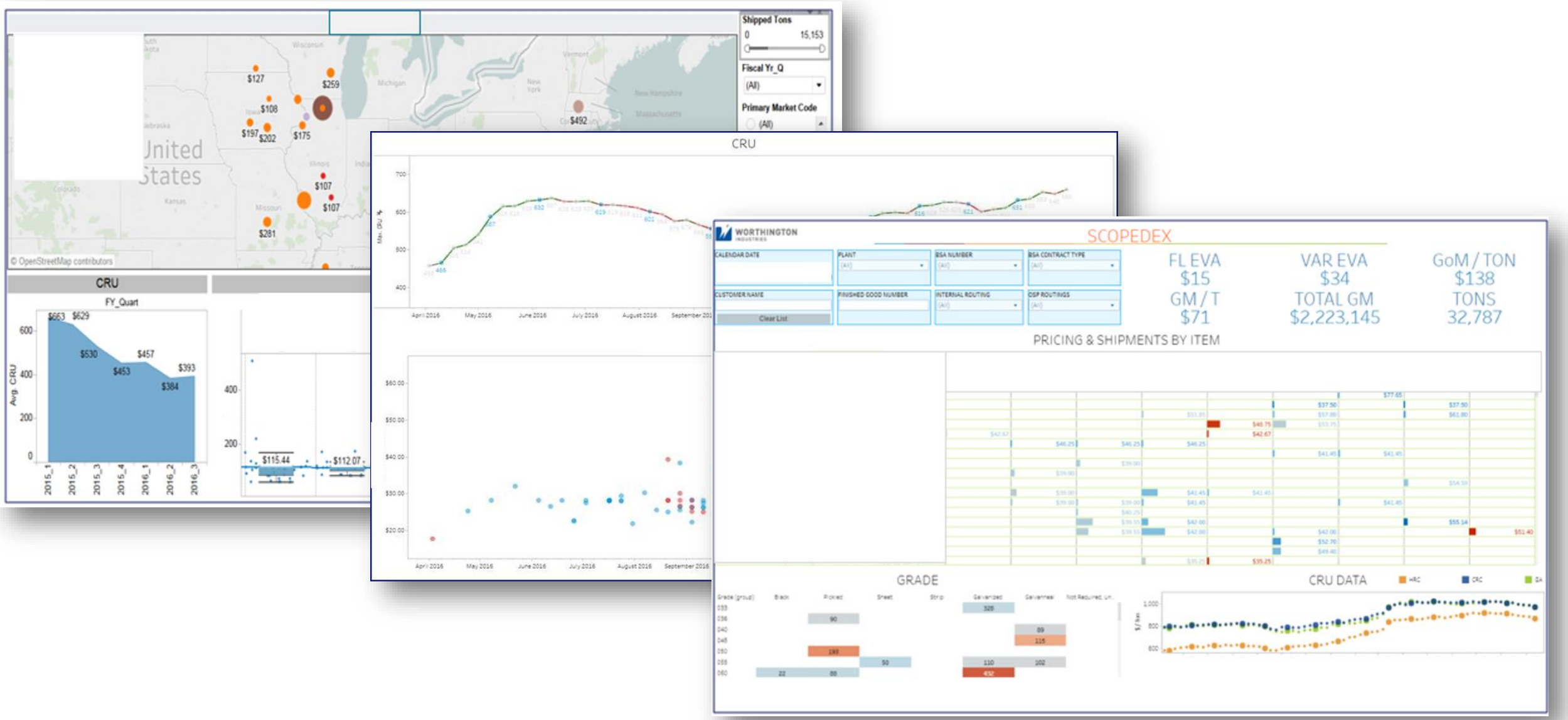
SCOPE tools: from historical pricing data to our won/lost quote data

TABLEAU CONFERENCE

> “We have to find a way of making the important measurable, instead of making the measurable important.”
>
> Robert McNamara, U.S. Sec. of Defense

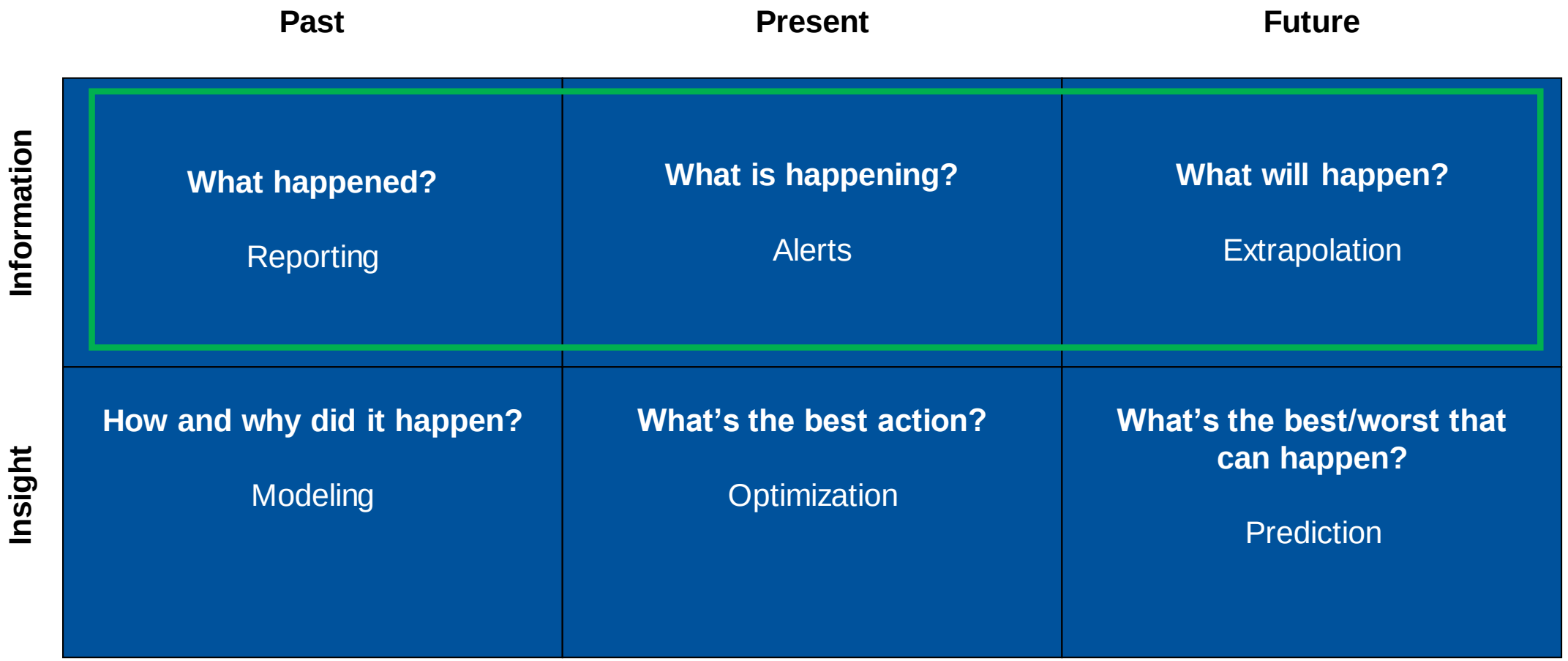
SCOPE tools: from historical pricing data to our won/lost quote data

TABLEAU CONFERENCE



Key Questions Addressed by Analytics

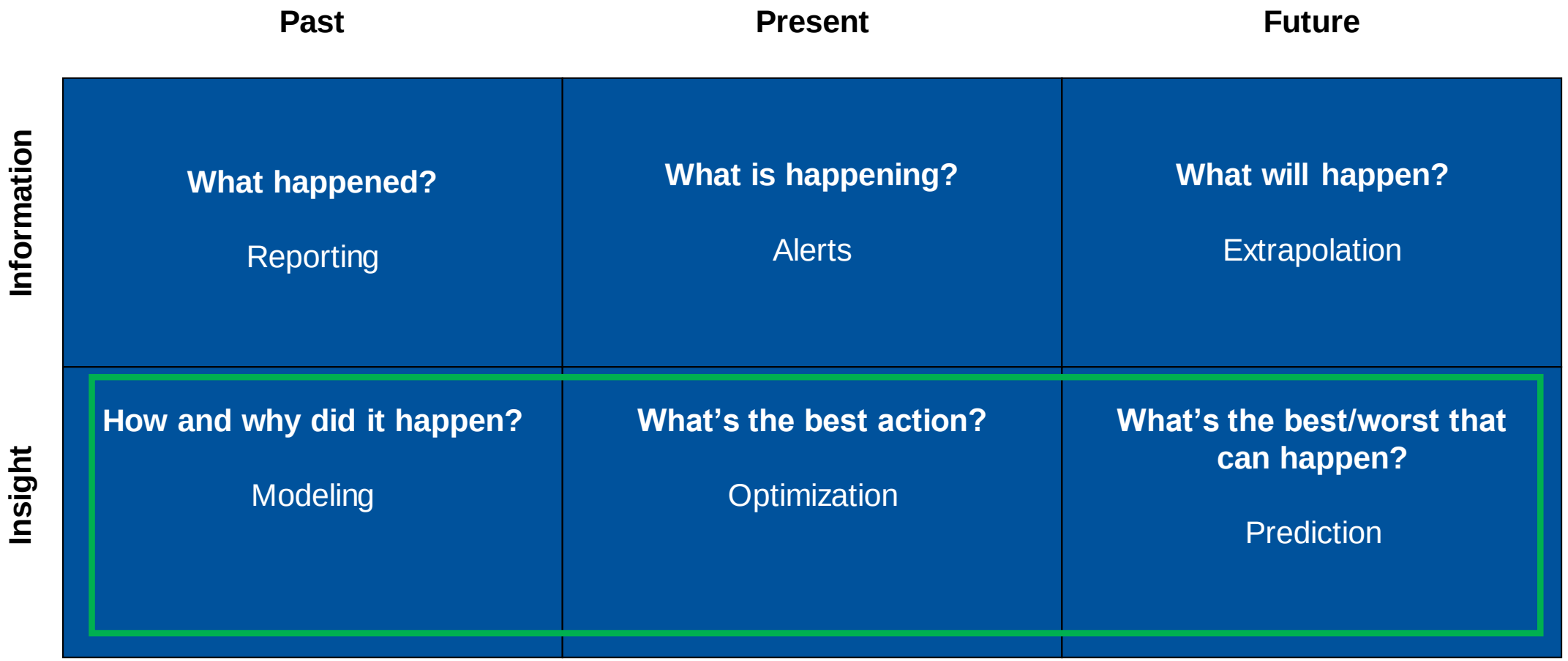
TABLEAU CONFERENCE



Source: Davenport et al. *Analytics at Work: Smarter Decisions, Better Results*

Key Questions Addressed by Analytics

TABLEAU CONFERENCE



Source: Davenport et al. *Analytics at Work: Smarter Decisions, Better Results*

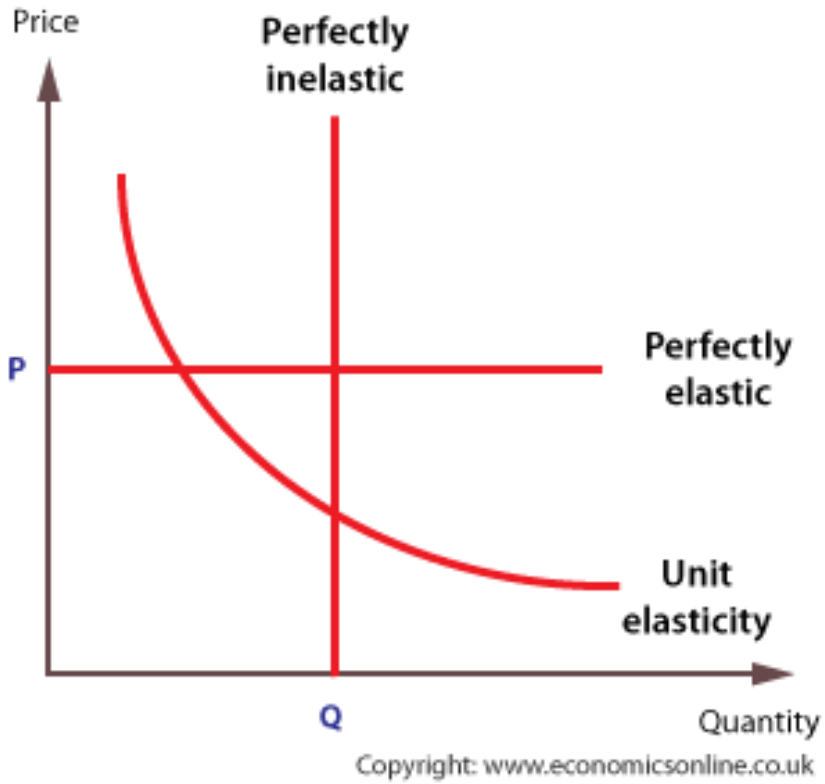
TABLEAU CONFERENCE

A Machine Learning ~~Approach~~ Misfire

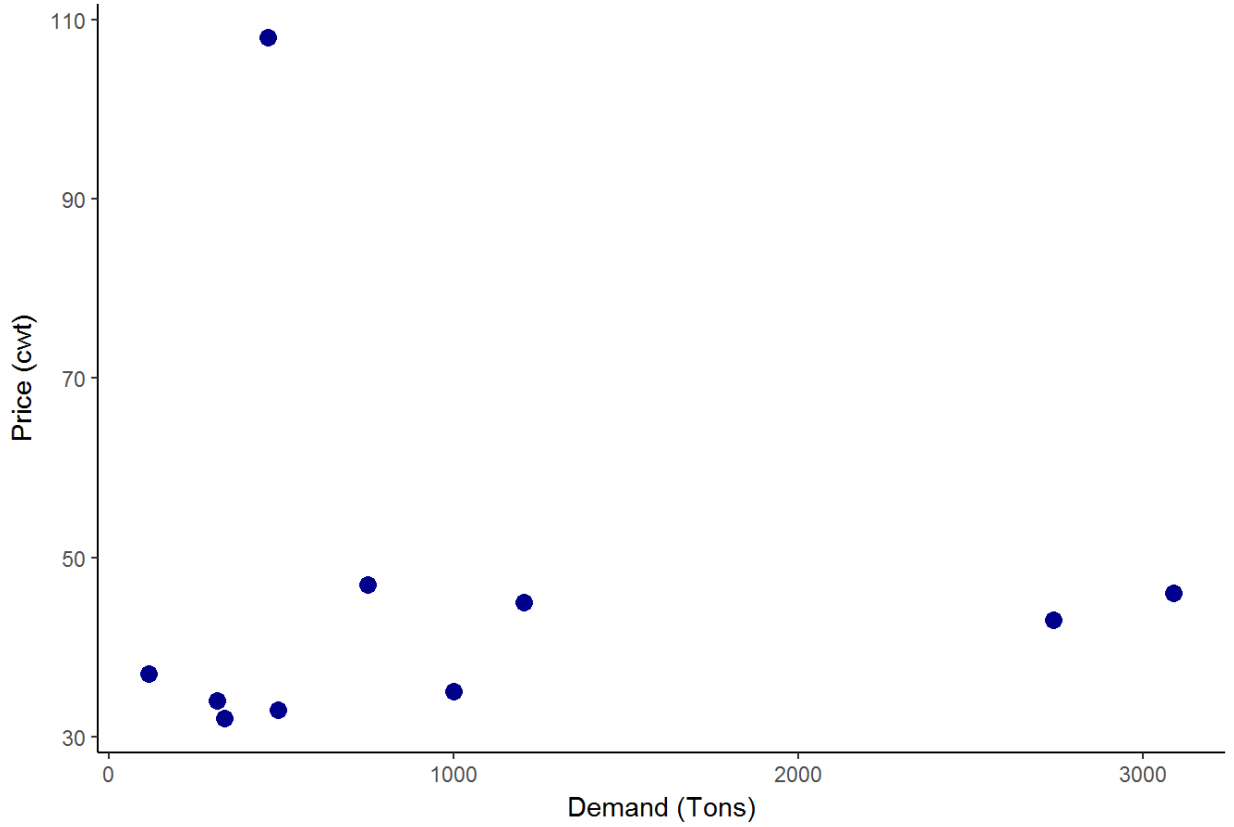
Price Elasticity and Optimization

TABLEAU CONFERENCE

A traditional Price Elasticity Curve...



Price Elasticity at Worthington Steel...



Model Design – What is steel?

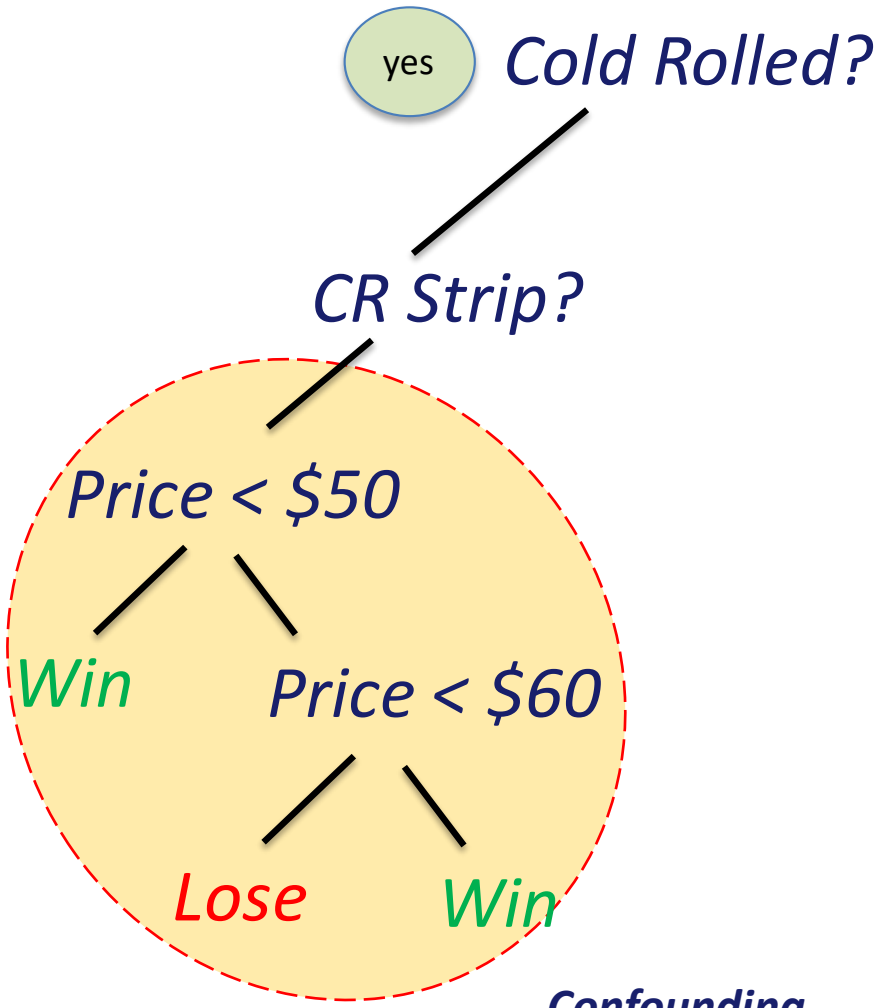
- Alloy (carbon content)
- Thickness
- Width
- Other elements (N, Si, Mb, etc)
- Shape (coil, sheet, blank)

Model Design

TABLEAU CONFERENCE

- ~8,000 observations
- 150 original variables (50 useable)
- >500 engineered variables

Overfitting



Confounding variables!

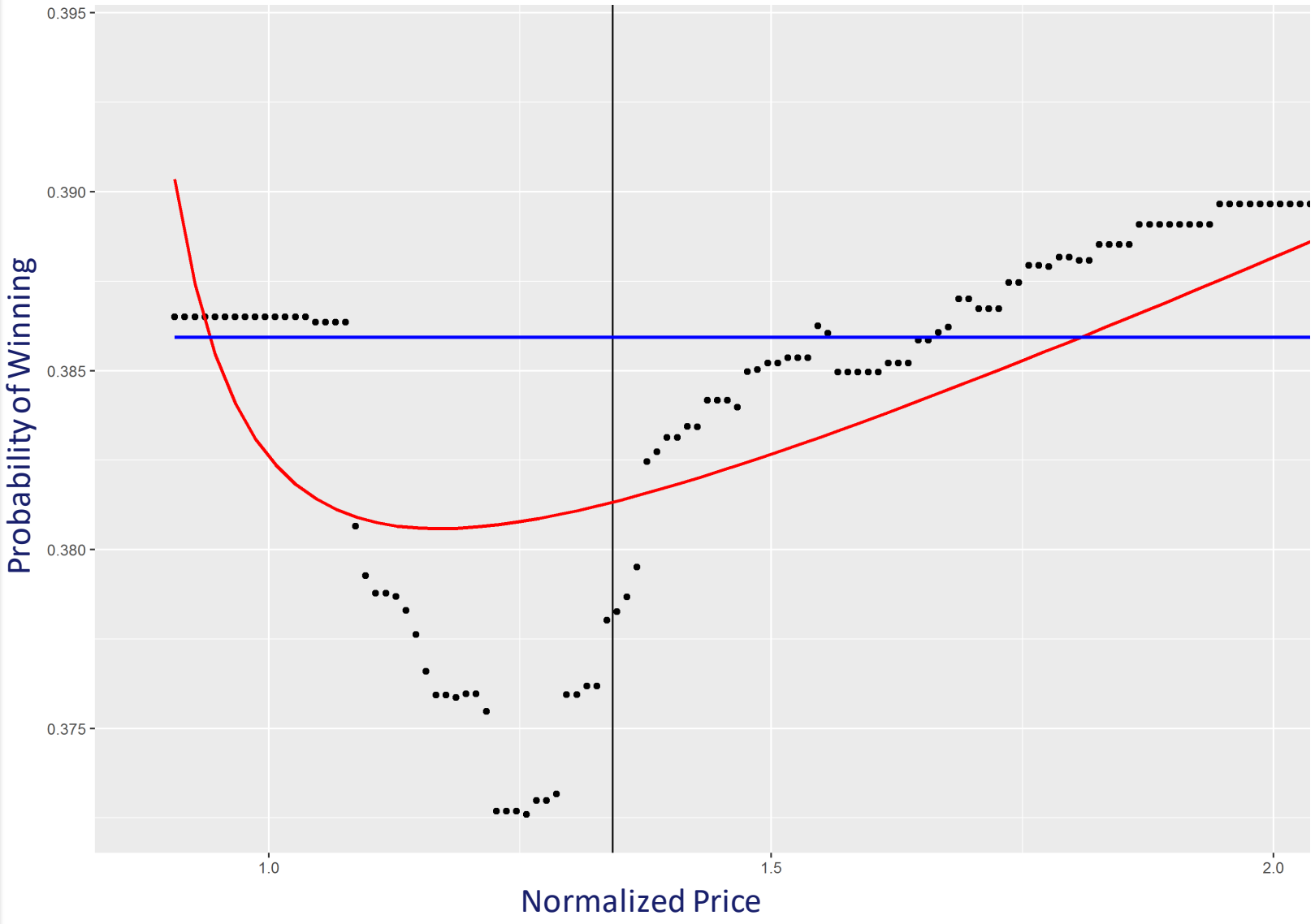
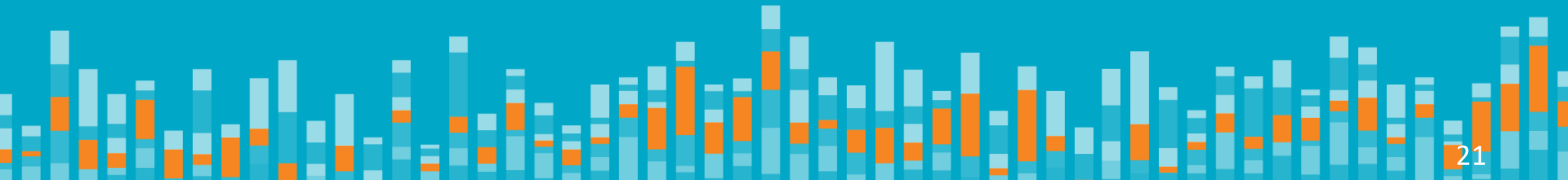


TABLEAU CONFERENCE

How should features be chosen to maximize accuracy and minimize over-fitting?

TABLEAU CONFERENCE

Let the User Choose with Tableau (and R)!



Model Status:

Variance Type
Pct

Current Market Price
$400.00

Proposed (Test) Price
$475.00

Product Type
Type 1

Product Type 2
(All)

Plant
(All)

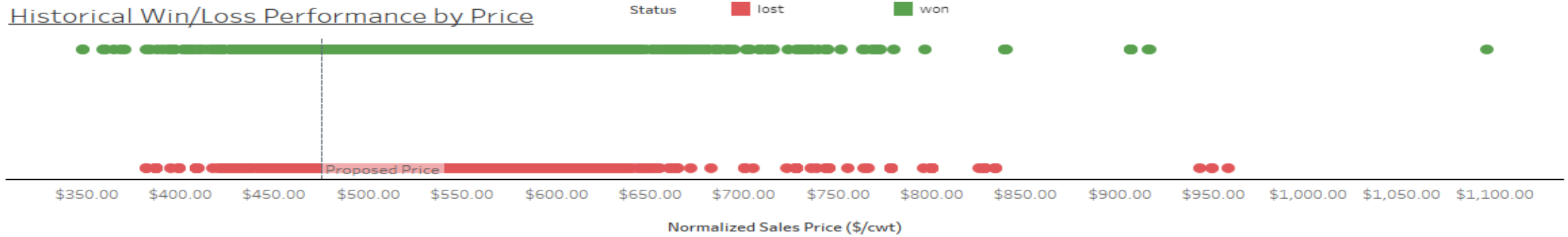
Contract Type
(All)

Customer
(All)

Date (Due)
4/26/2016 1/9/2018

Exclude Quotes
(Multiple values)

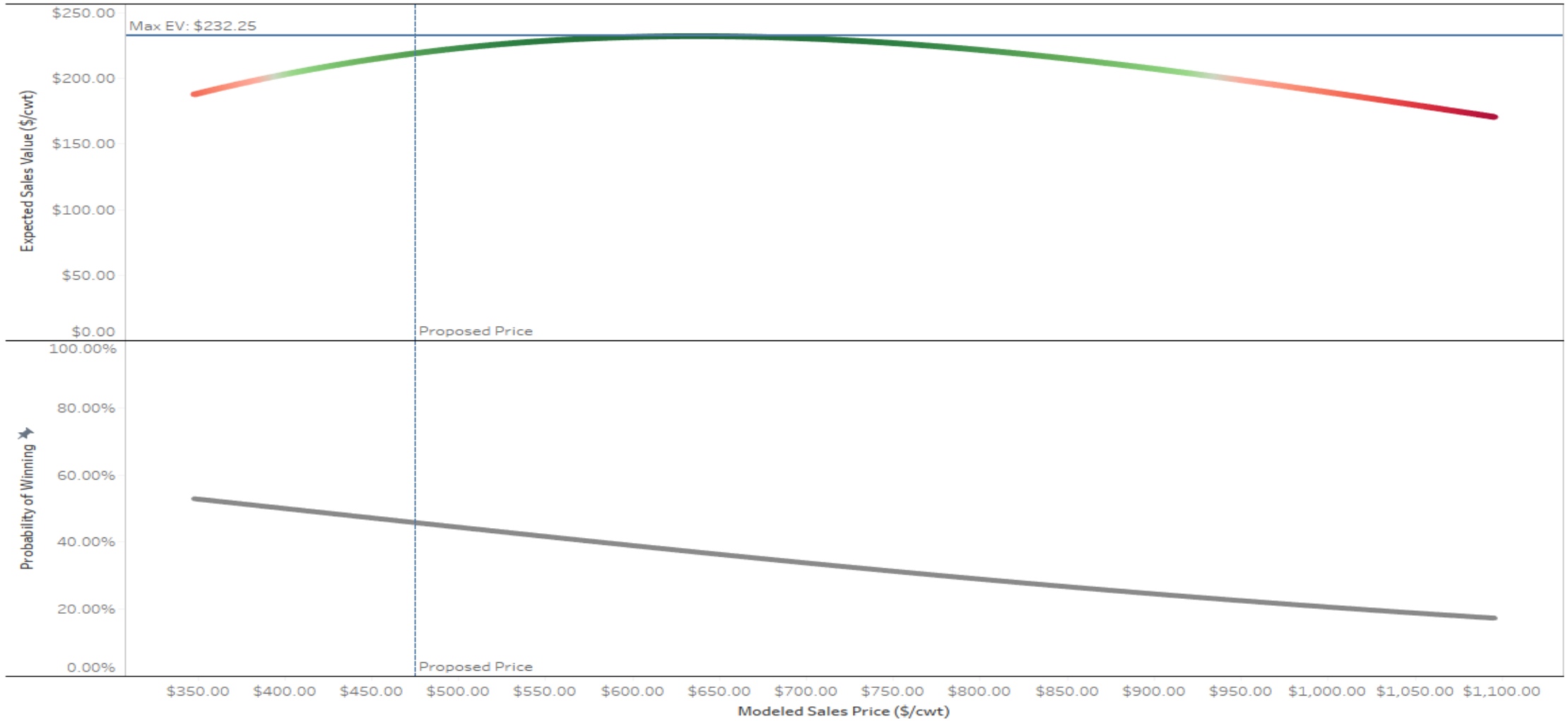
Historical Win/Loss Performance by Price

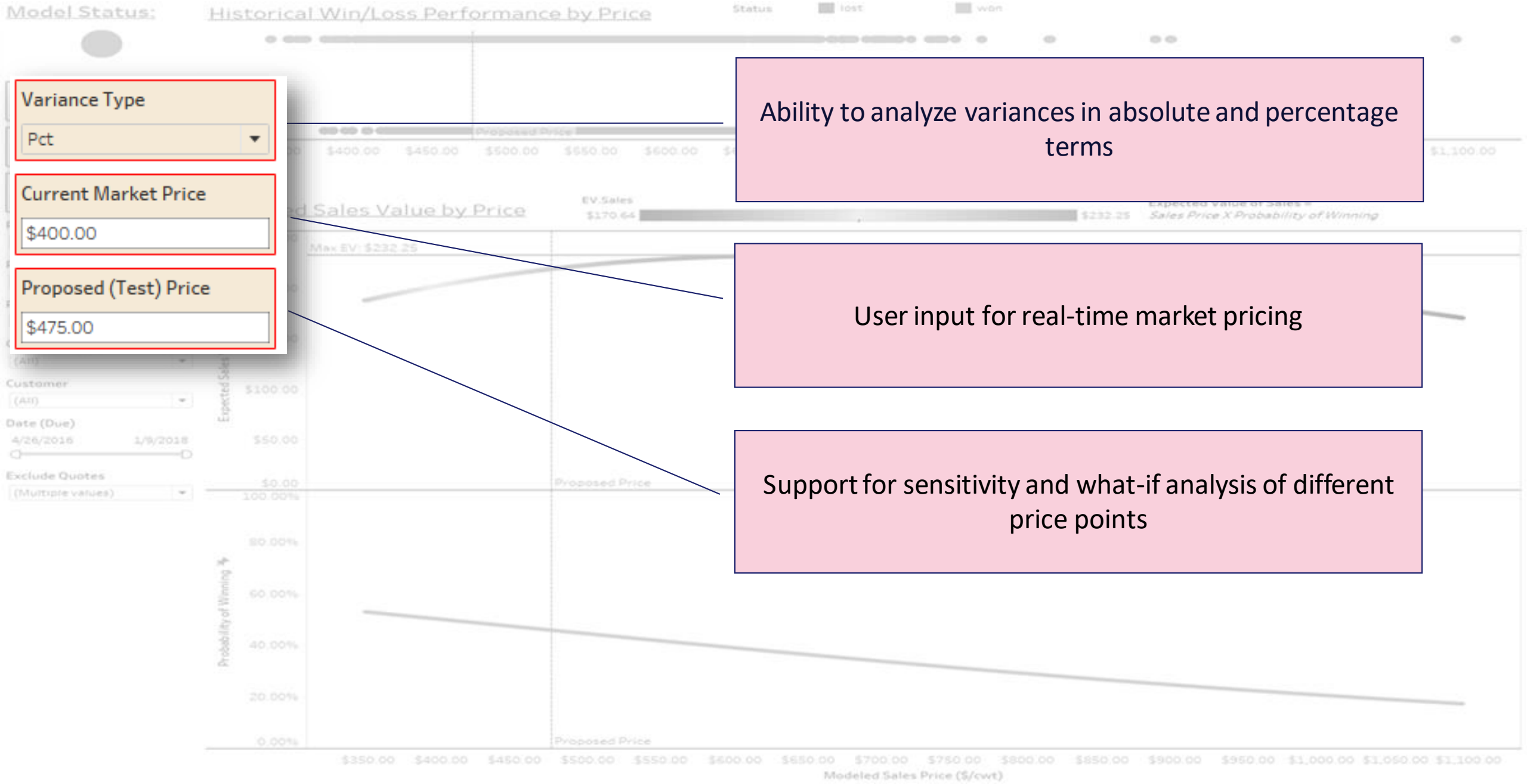


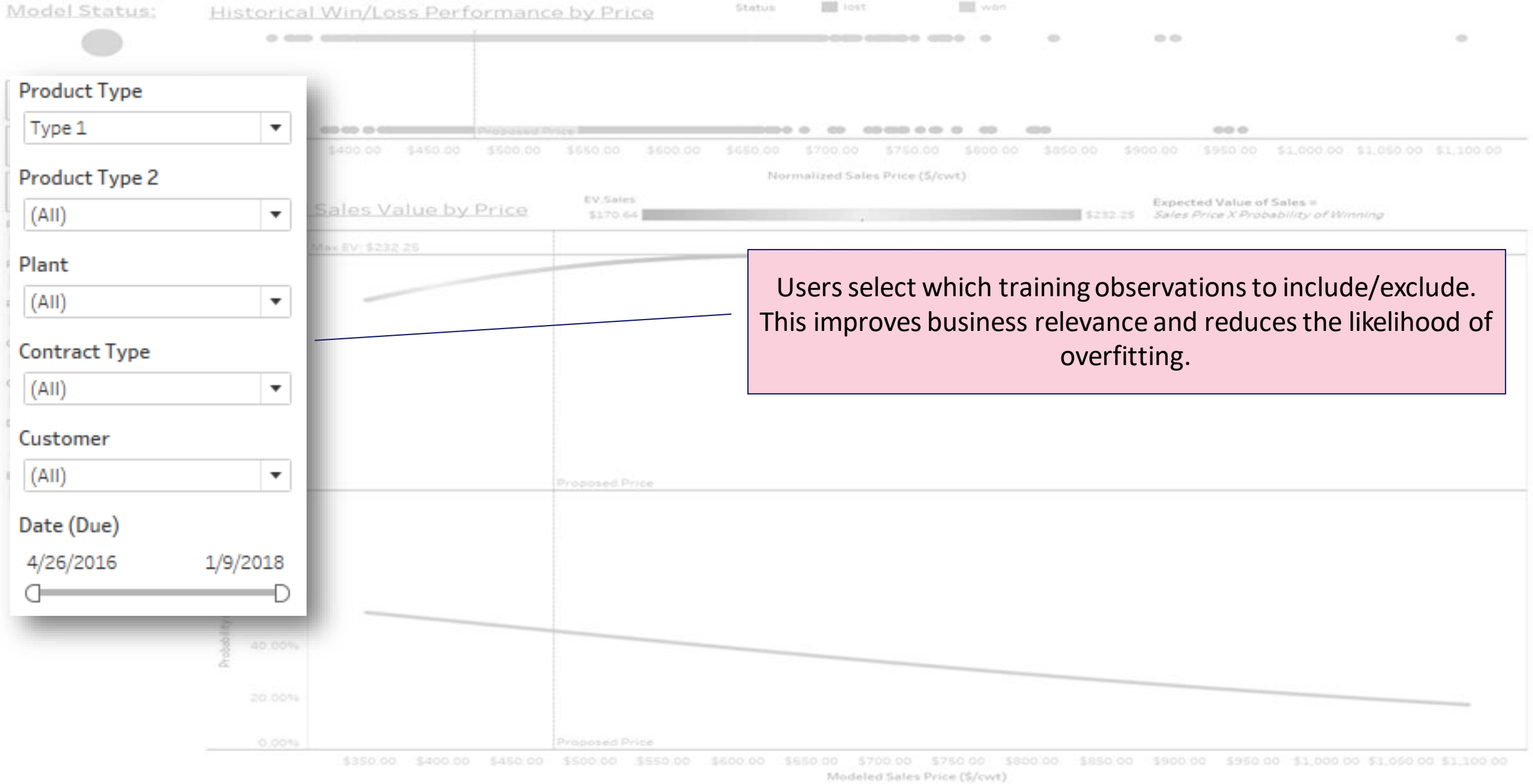
Expected Sales Value by Price

EV.Sales
$170.64 $232.25

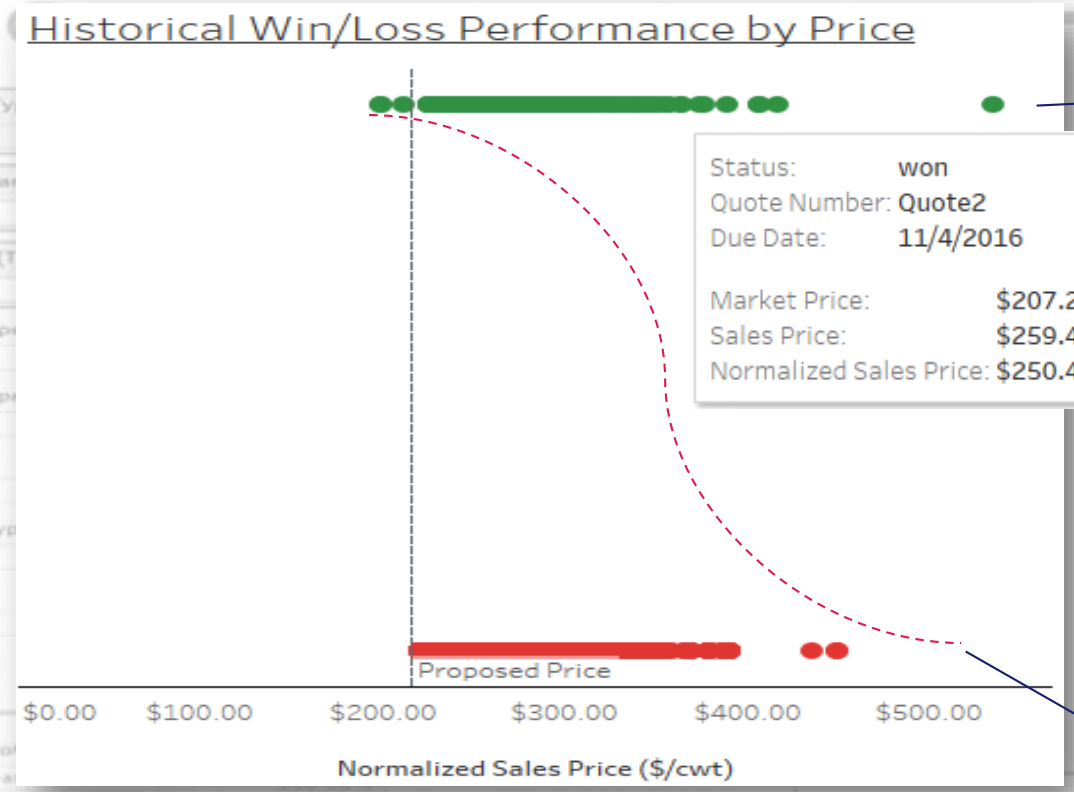
Expected Value of Sales =
Sales Price X Probability of Winning




Model Status:
Historical Win/Loss Performance by Price
Variance Type
Pct
Current Market Price
$400.00
Proposed (Test) Price
$475.00
Ability to analyze variances in absolute and percentage terms
User input for real-time market pricing
Support for sensitivity and what-if analysis of different price points
Modeled Sales Price ($/cwt)



Users select which training observations to include/exclude. This improves business relevance and reduces the likelihood of overfitting.



Historical wins and losses by price are provided for context

Tooltip provides individual quote data

Chart type and layout chosen to imply logistic regression context

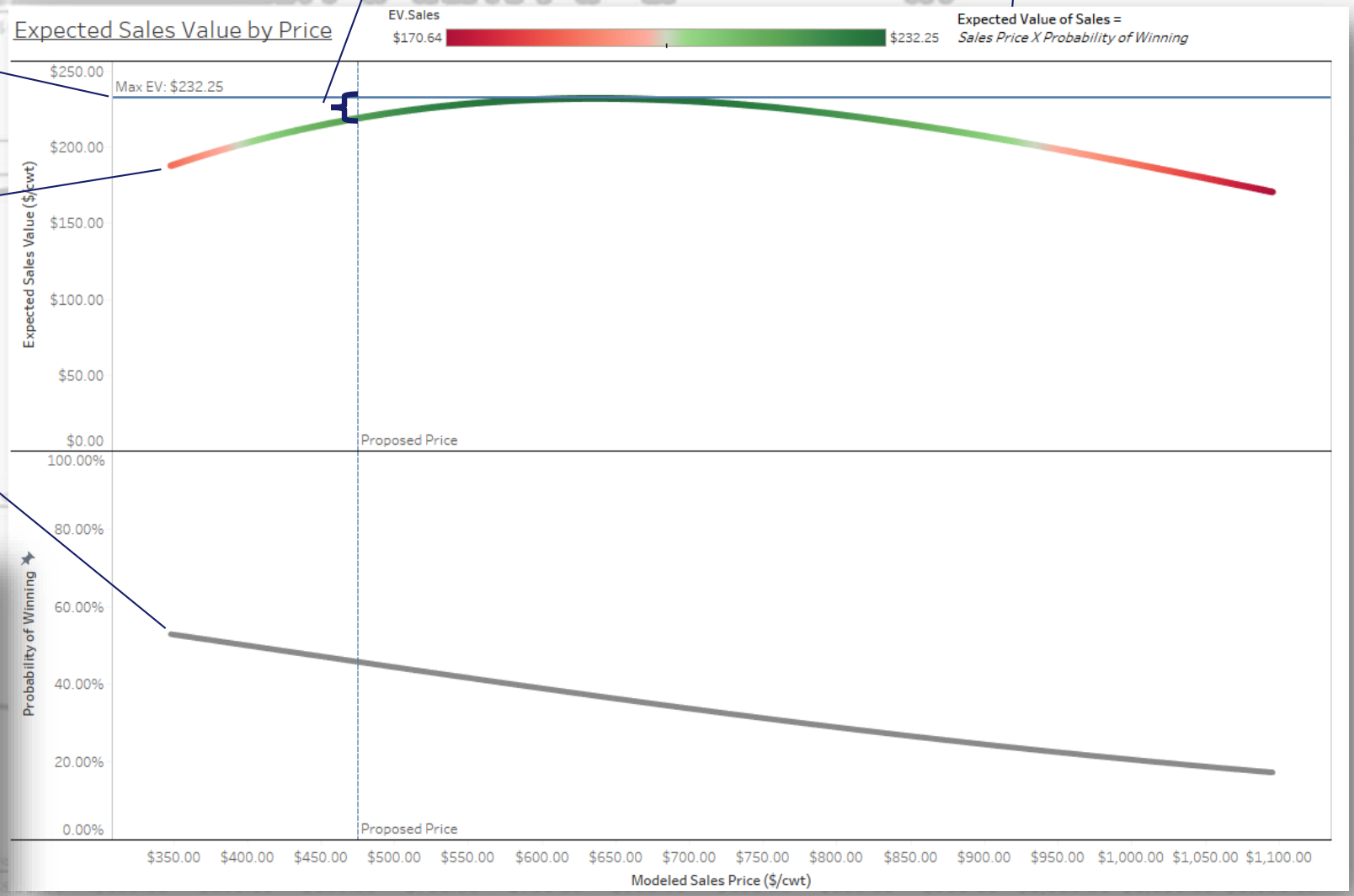
Leakage from sub-optimal pricing

Expected value defined as:
$EV = Price \times P(Winning)$

Model identifies price that maximizes expected value

Expected value as a function of price

Model predicts win probability given simulated price



Results are computed along Quote.Key.

```
SCRIPT_REAL('
library(pe_lm)
x <- pe_lm(
var = .arg1,
win_status = .arg2
)
x$ndf$win_prob
',AVG([Variance.ParamCtrl]),ATTR([Status]))
```

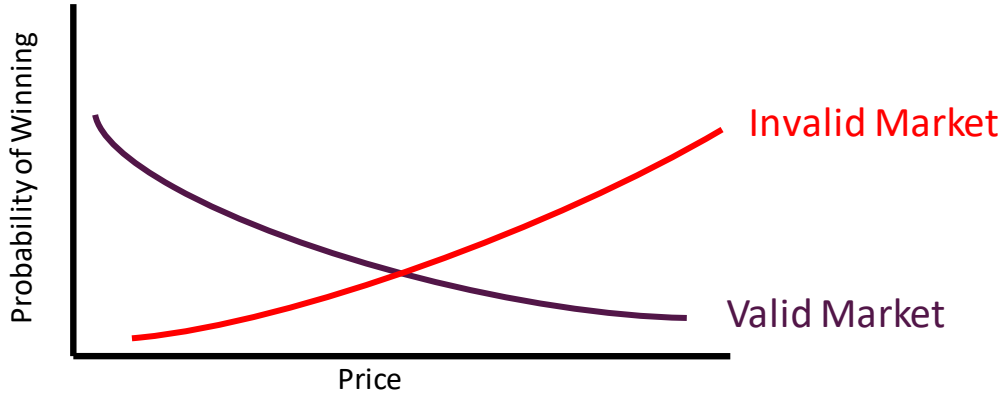
Model Status:

Traffic light indicates model quality

Model is valid!

: `fit.sum` only contains data when a model is fit. || b0:0.0994 || b1:-1.1945 || b0.p:0.0075 || b1.p:0 || dwdv:Min.:-0.299, 1st Qu.:-0.292, Median:-0.249, Mean:-0.236, 3rd Qu.:-0.187, Max.:-0.128, sd: 0.056

Tooltip provides additional context and statistical measures




```
# derivative
# dwdv is the derivative of win probability with respect to price variance
var <- out$pred_frame$var
dwdv <- deriv(eq,"var") %>% eval() %>% attr("gradient") %>% as.numeric()
```

Model.Message

price_elasticity_pilot_dataset

Results are computed along Table (across).

```
SCRIPT_STR('
library(pe_lm)
x <- pe_lm(
var = .arg1,
win_status = .arg2
)
x$status$msg
',AVG([Variance.ParamCtrl]),ATTR([Final.Status.of.Quote]))
```

The calculation is valid.

Sheets Affected

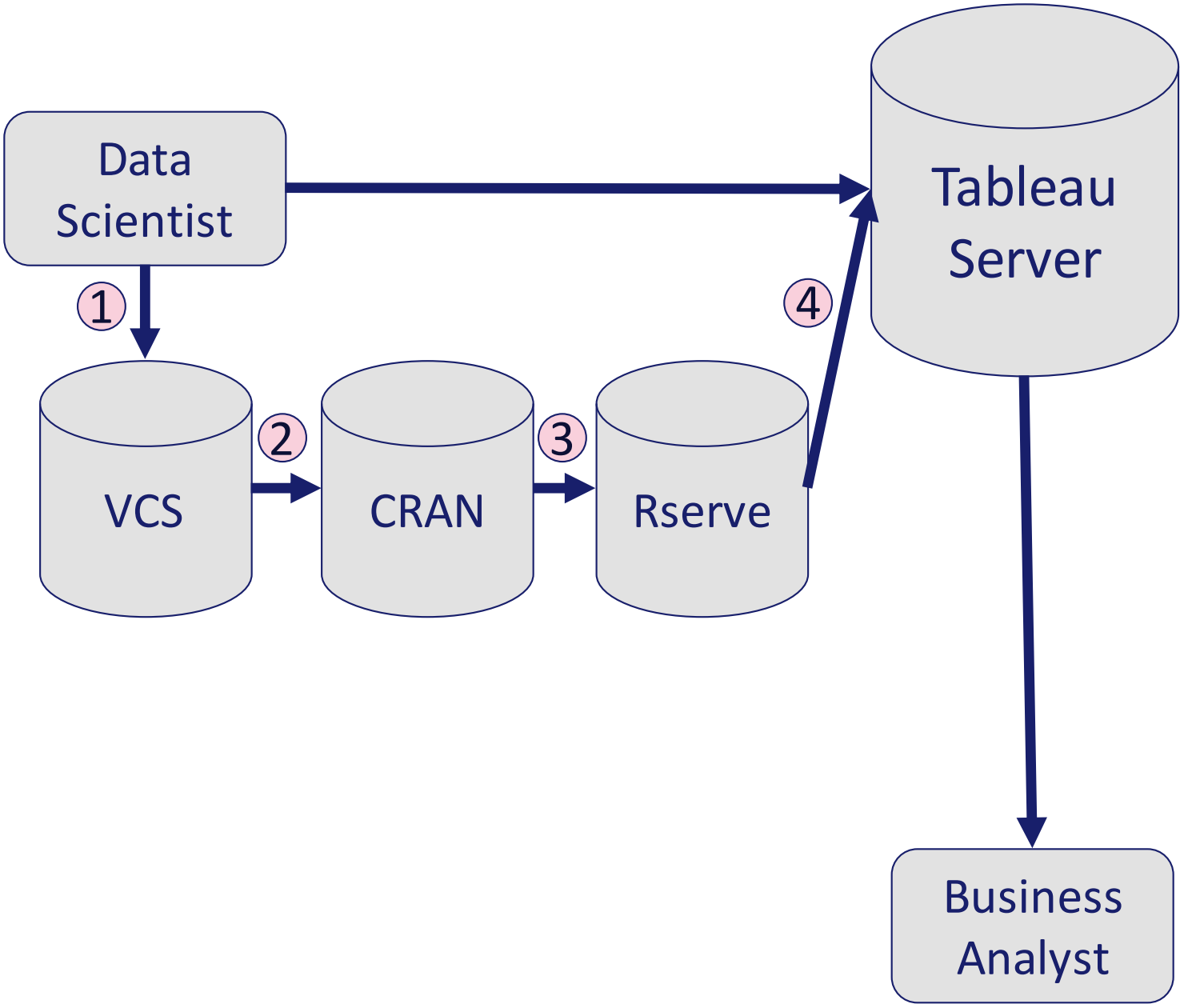
What makes a successful journey?

TABLEAU CONFERENCE

- Strong **engagement** from the business
- **Alignment** across functional areas (sales, analytics, IT) – are we solving a real business problem?
- Introduce complexity **in pieces** to reduce learning curve
- Allow users to **see their data** (in addition to the model)

TABLEAU CONFERENCE

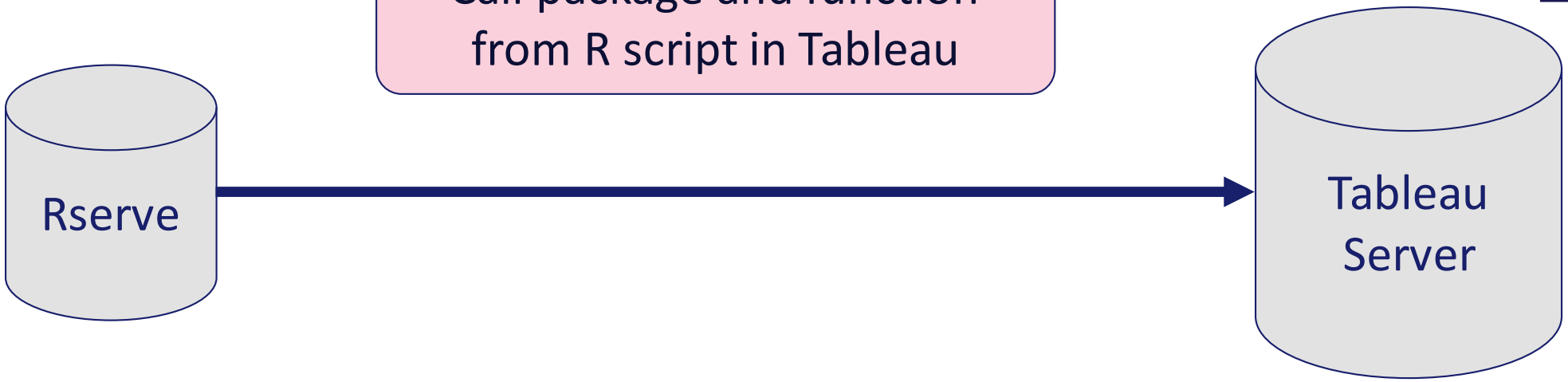
Tableau and R Technology Stack



1. Develop and push model as an R package
2. On merge to master, build to an internal CRAN mirror
3. Install/update package on Rserve server from CRAN
4. Call package and function from R script in Tableau

TABLEAU CONFERENCE

Call package and function from R script in Tableau




```
pe_lm <- function(var,win_status) {
  wl <- factor(win_status,levels = c("lost","won"))
  df <- data.frame(var,wl)
  wts <- round((1-(table(wl)/length(wl)))*100)
  df$wts <- as.numeric(wts[wl])
  fit <- glm(wl ~ var,family = "binomial",data=df,weights = wts)
  price_var_vec <- seq(min(var),max(var),by = (max(var)-min(var))/(length(var)-1))
  price_var_vec <- price_var_vec[rank(fit$data$var,ties.method = "first")]
  ndf <- data.frame(var=price_var_vec)
  ndf$win_prob <- predict.glm(fit,newdata = ndf,type="response")
  out <- list(ndf = ndf, fit = summary(fit))
  out$status <- pe_lm_status(out)
  return(out)
}
```

Results are computed along Quote.Key.

```
SCRIPT_REAL('
library(pe_lm)
x <- pe_lm(
var = .arg1,
win_status = .arg2
)
x$ndf$win_prob
',AVG([Variance.ParamCtrl]),ATTR([Status]))
```

Why build and maintain this infrastructure?

TABLEAU CONFERENCE

- **Efficiency** of continuous integration and continuous deployment
- **Stability** and **reproducibility** of version control
- Automated **testing** in R (`testthat` package)
- **Performance** gains from distributed and concurrent processing – R and Tableau are running on separate servers
- **Security** of a CRAN repository that is behind the firewall

TABLEAU CONFERENCE

Thank you!

Steve Bartos

@OldDirtyBarGraph

Steve.Bartos@worthingtonindustries.com

Wil Davis

github.com/wkdavis

William.Davis@worthingtonindustries.com