



Seeqing Automated Model Retrains with Data Lab

Stephanie Ma
Invista Propylene

Stephanie Ma

- **Invista Propylene, Houston | Modeling & Analytics Engineer (~1 yr)**
- B.S. Chemical Engineering, Kansas State University '24

Past Experience (~2 yrs)

- Goodyear | Extrusion/Mixing Quality Co-op
- Chevron Phillips Chemical | Olefins Process Monitoring Intern *
- Chevron Phillips Chemical | NGL Process Engineer

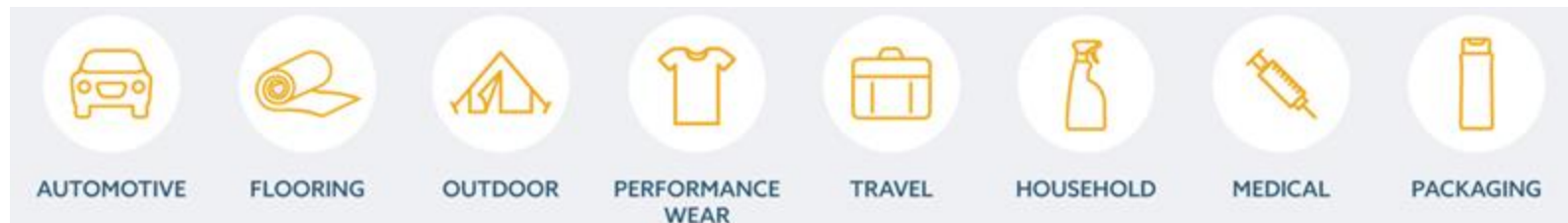


Who is INVISTA?

We make the chemical ingredients –

including adiponitrile (ADN), hexamethylene diamine (HMD), adipic acid (AA), polyamide 66 (PA66) polymer-grade propylene (PGP), and polypropylene (PP)

– that are essential for many different products.



Within our polyamide business, INVISTA has a **global footprint** to serve the market with **local production** and we continue to build our production capabilities around the world to meet future market demand.

With our propylene value chain, we are intentional about delivering **high-performing solutions** that meet your specific needs and helping you get your products to market as quickly as possible.

We are pioneers in **innovative research** and create advanced process and product technologies – ensuring lasting, **long-term value** for our customer.



ABOUT

3,500

EMPLOYEES
WORLDWIDE



ABOUT

250

PATENT FAMILIES



ABOUT

20

LOCATIONS



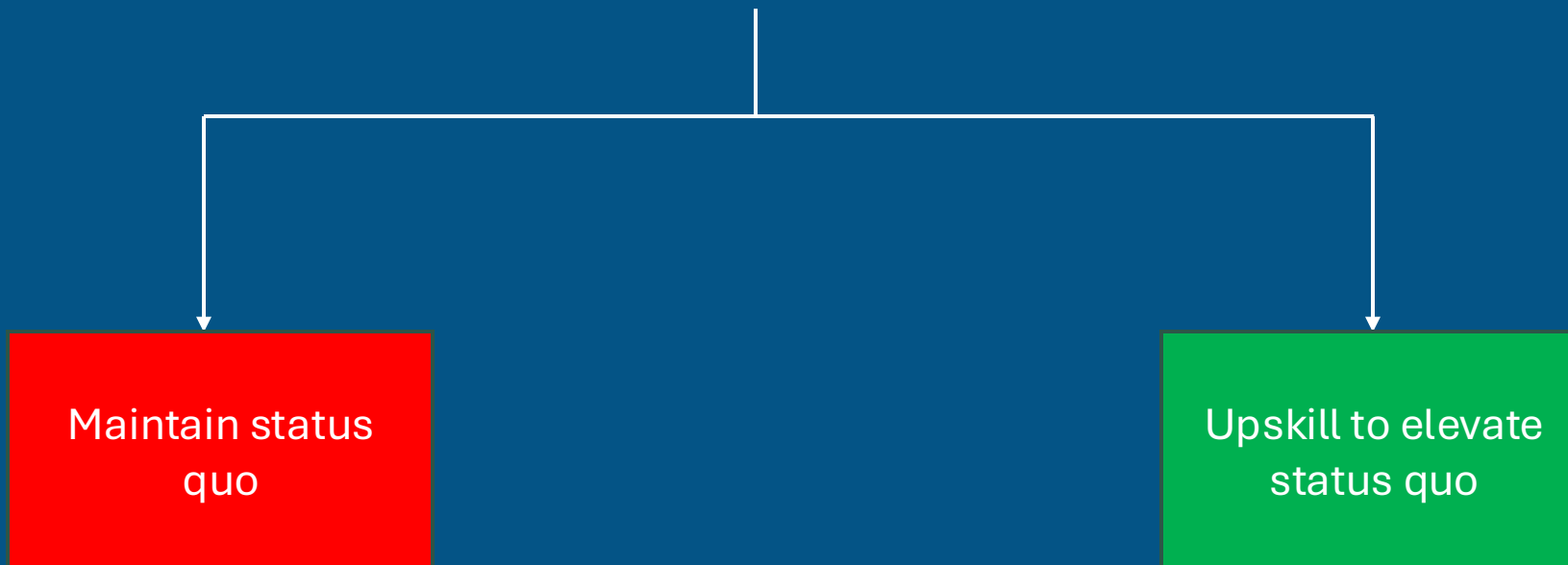
Seeing Automated Model Retrains w/ Data Lab

1. Challenge: inefficient workflow
2. Solution: automation
 1. Engine changeout check/training range determination
 2. On/off check
 3. Outlier check
 4. Final bounds determination
3. My approach as a non-computer scientist
4. Next steps
5. Key takeaways

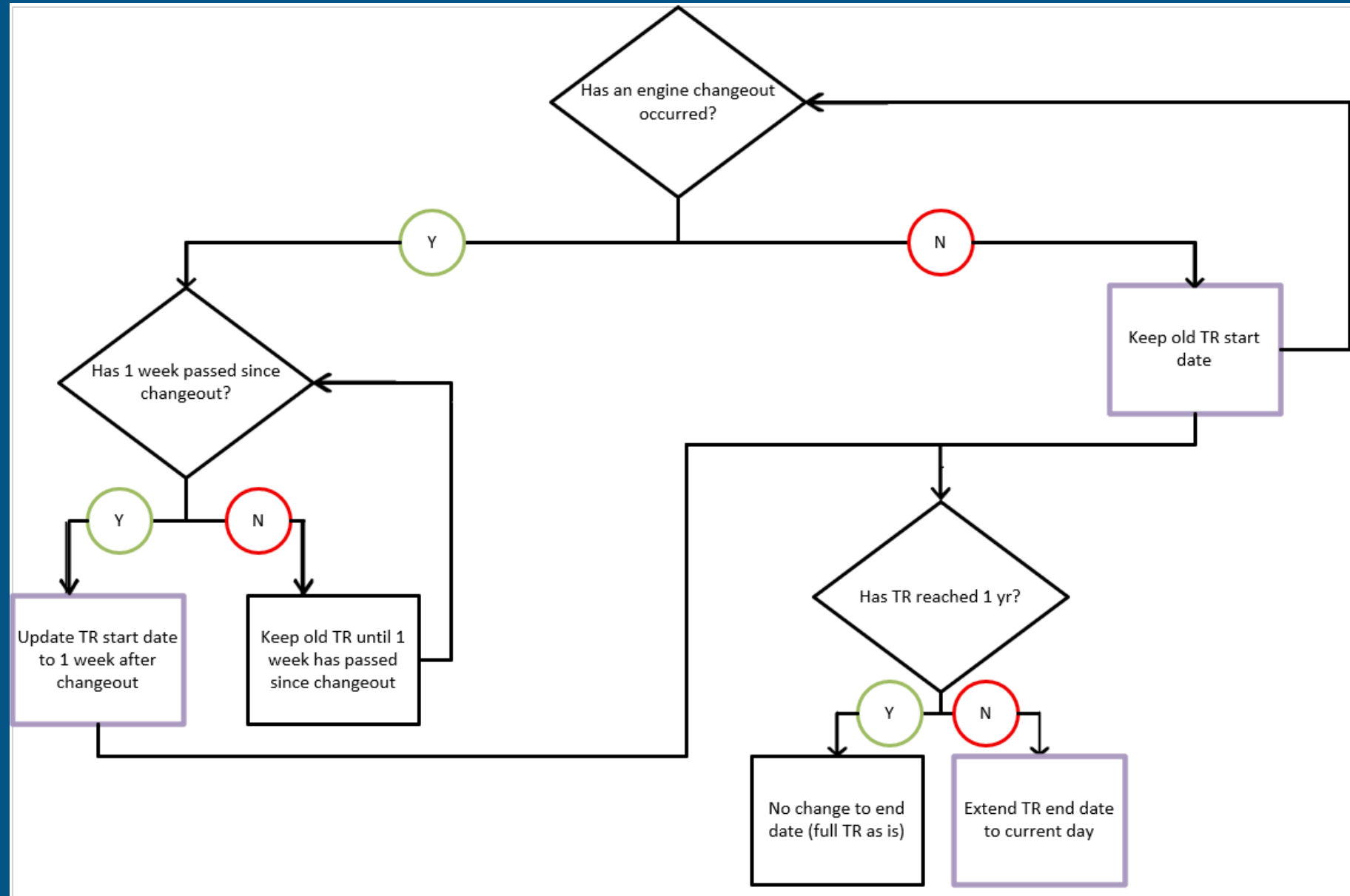
Challenge: Inefficient Workflow

- Built 120 models, ~90% in nuisance alert due to incorrect training range
 - Equipment overhaul strategy almost entirely independent from outage/TA schedule

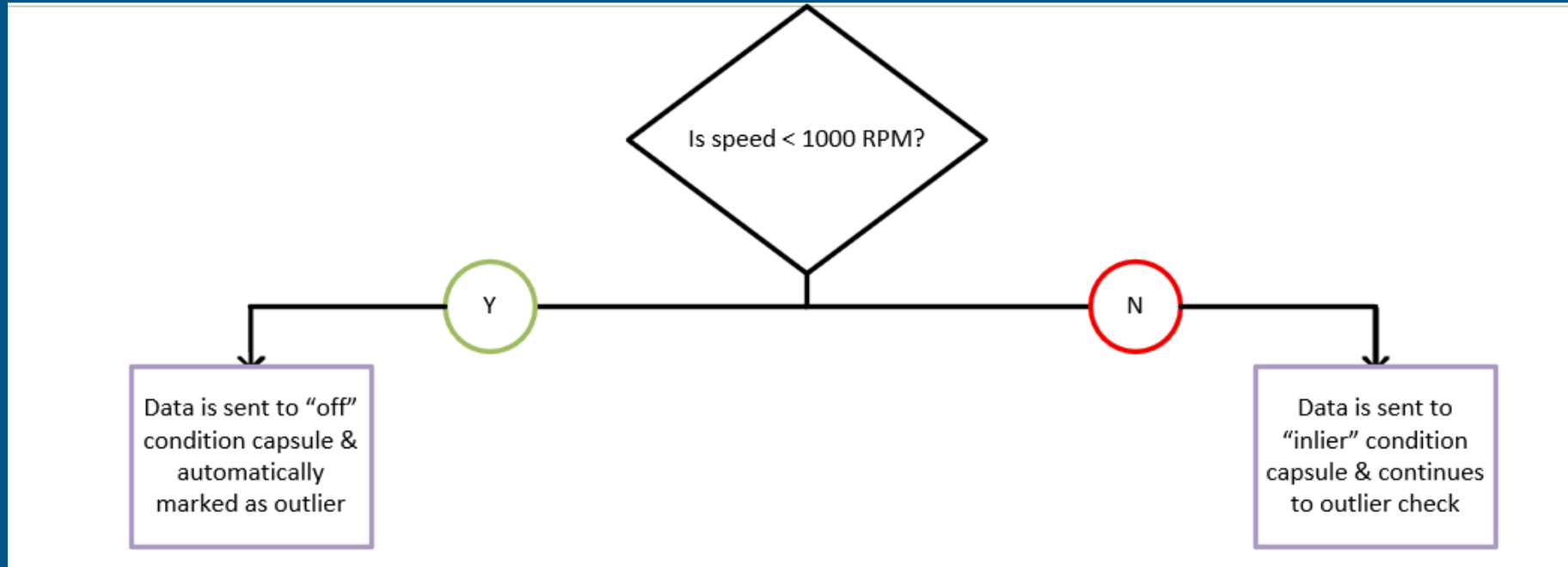
Problem: I doubled the alert count I looked through each morning & added a workload I didn't want to carry



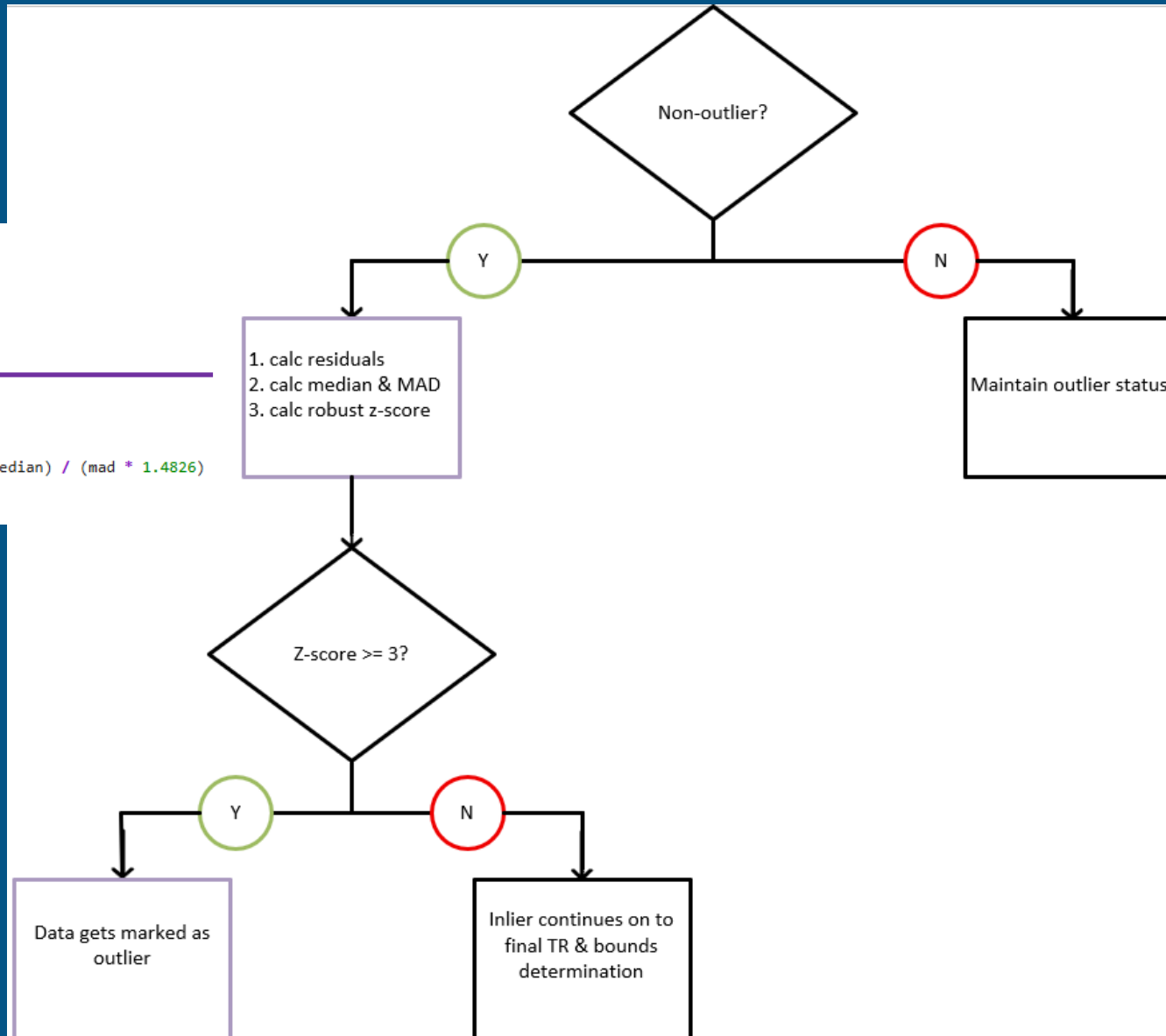
1. Engine Changeout Check/TR Determination



2. On/Off Check



3. Outlier Check

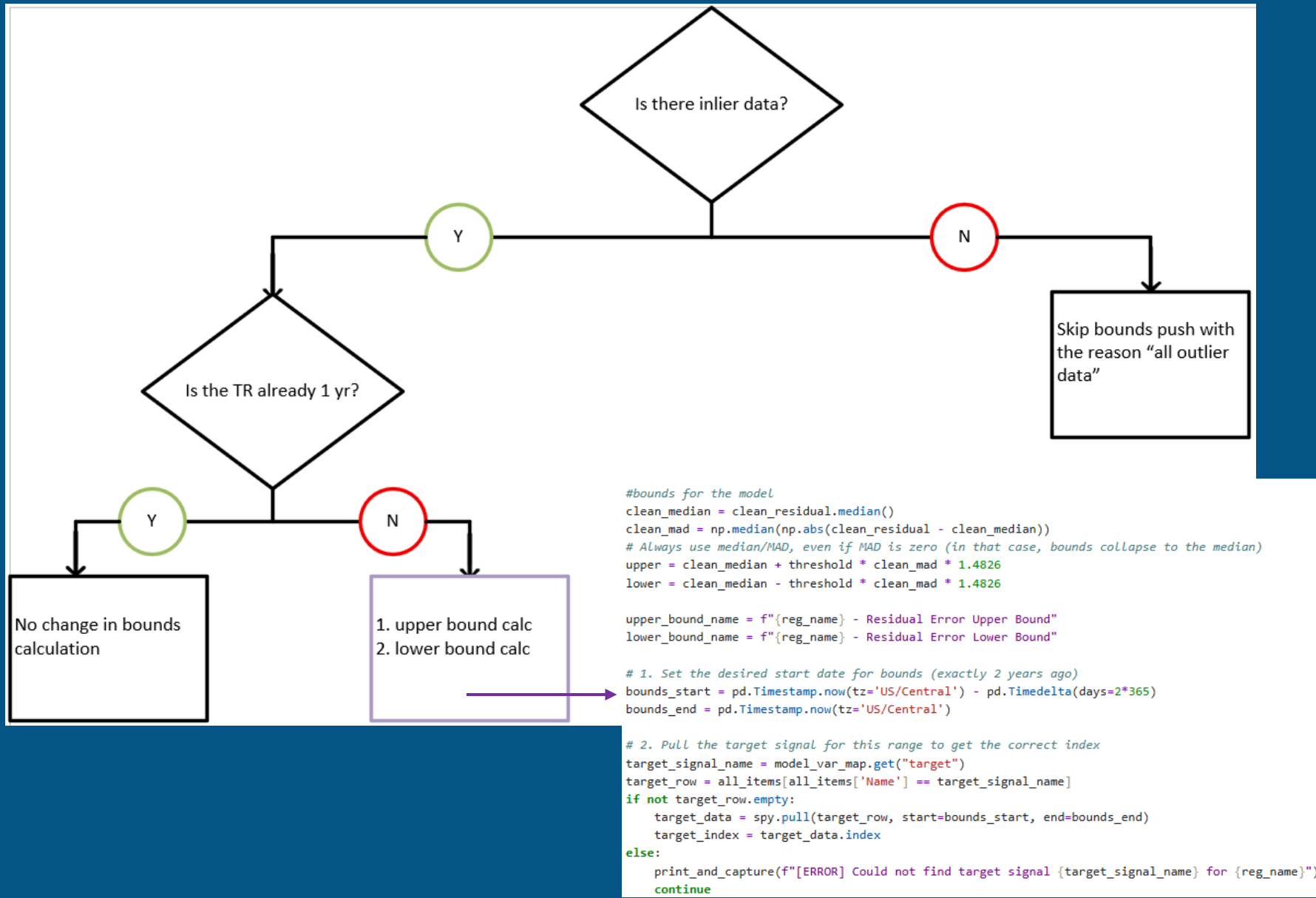


```
residual = data[target_signal_name] - data[reg_name]
gt_results[reg_name]['Residuals'] = residual
```

```
# --- DEBUG: Print and collect residuals/z-scores ---
residual_no_nan = residual.dropna()
median = residual_no_nan.median()
mad = np.median(np.abs(residual_no_nan - median))
mad = max(mad, 1e-6) # Avoid division by zero
```

```
robust_zscores = pd.Series(np.nan, index=residual.index)
robust_zscores[residual_no_nan.index] = (residual_no_nan - median) / (mad * 1.4826)
non_outlier_mask = robust_zscores.abs() <= threshold
non_outlier_mask = non_outlier_mask.fillna(False)
```

4. Final Bounds Determination

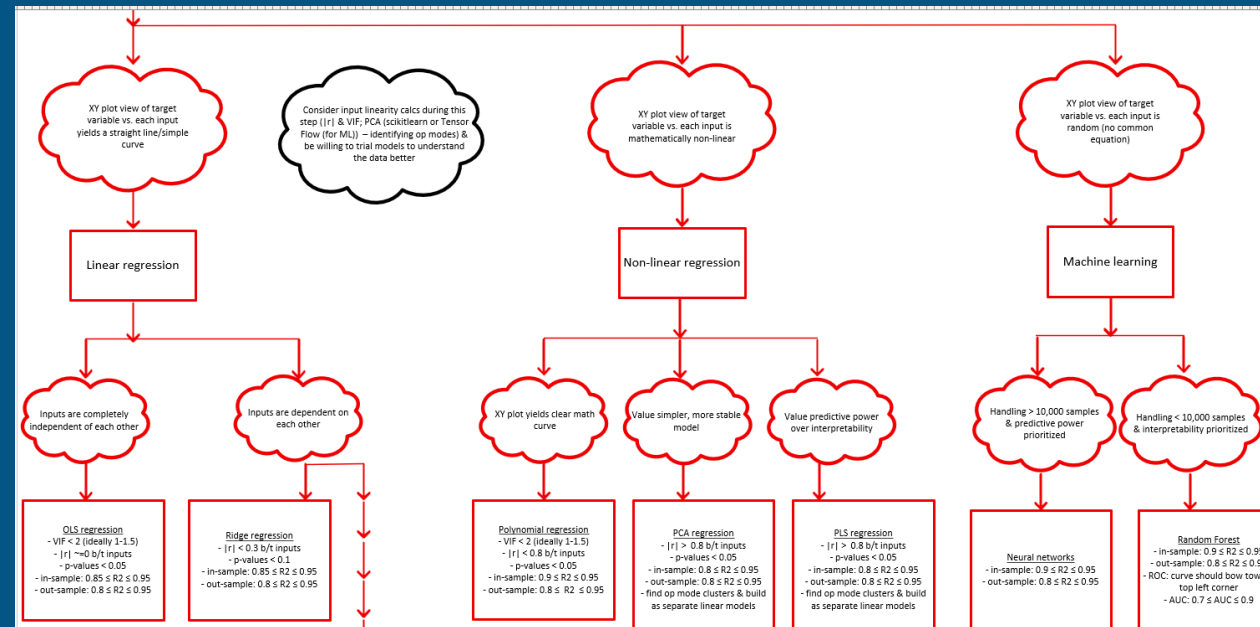


My Approach (Non-Computer Scientist)

1. Take Data Lab Training
 1. Variable to signal per model dictionary
 2. `spy.search()`*
 3. Calcs w/normal Python syntax
 4. `spy.push()`*
 5. `spy.notifications.send_email()`*
 6. `spy.jobs.schedule()`*
2. Use diagnostic data sent via `spy.notifications.send_email()` & Data Lab AI Agent to help debug
3. Have a clear idea of what outcome you want
4. Learn best practices

Next Steps

1. Apply TR automation code to the 120 GT models I built
2. Finetune continuous Prediction Interval & SHAP calculations code to help with model health diagnostics
3. Find other equipment that would benefit from TR automation
4. Finetune “model type decision tree” code to streamline model building in Seeq while accounting for multicollinearity, input significance, TR fit, and prediction accuracy



Key Takeaways

1. The Data Lab training is worthwhile
2. Learning the specific Seeq/Python code structure is worthwhile
3. The Data Lab AI Agent is your best friend

- Copy & paste error message *
- Why is this error happening and how can I fix it?
- Copy & paste code *
- Copy & paste output emailed to myself *

- I want to accomplish a specific task, but I keep running into this specific problem
- What is the best practice for doing this on Data Lab?

Seeq® DATA LAB



Initializing Project ...



Data Lab

I can help you do magic with Data Lab and Python code

Thank you for your time!

connect
ORLANDO JUNE 1-3, 2026

John Cox, Michael Schemm, Janelle Armstead-English
Invista Monitoring & Machinery
Seeq support in general



SeeQ®