Climbing the causal ladder for fun and profit

drbenvincent/pydata-global-2024

Benjamin T. Vincent, DPhil

PyData Global

PyMC Labs

PyMC

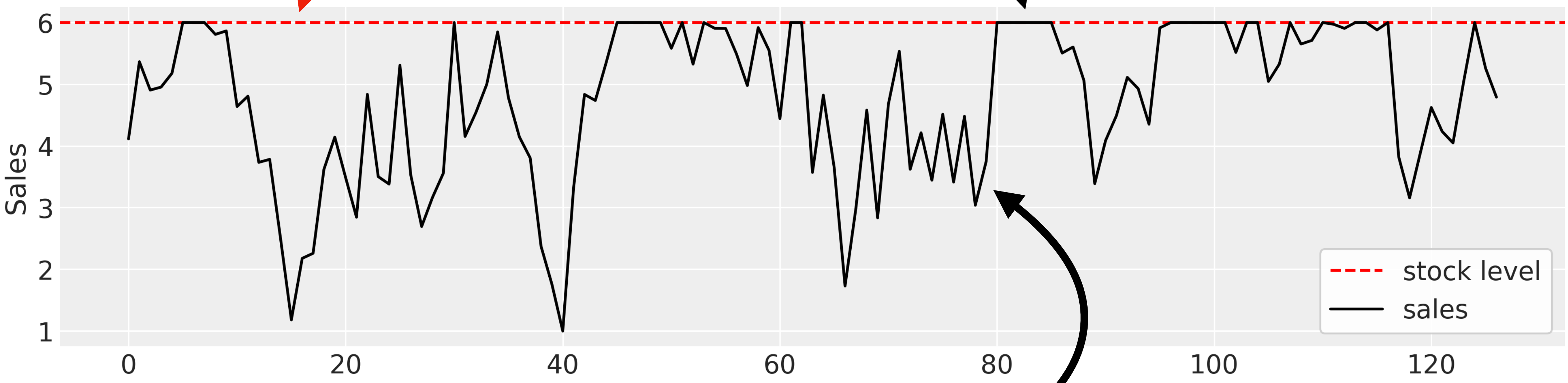
PyMC-Marketing

CausalPy

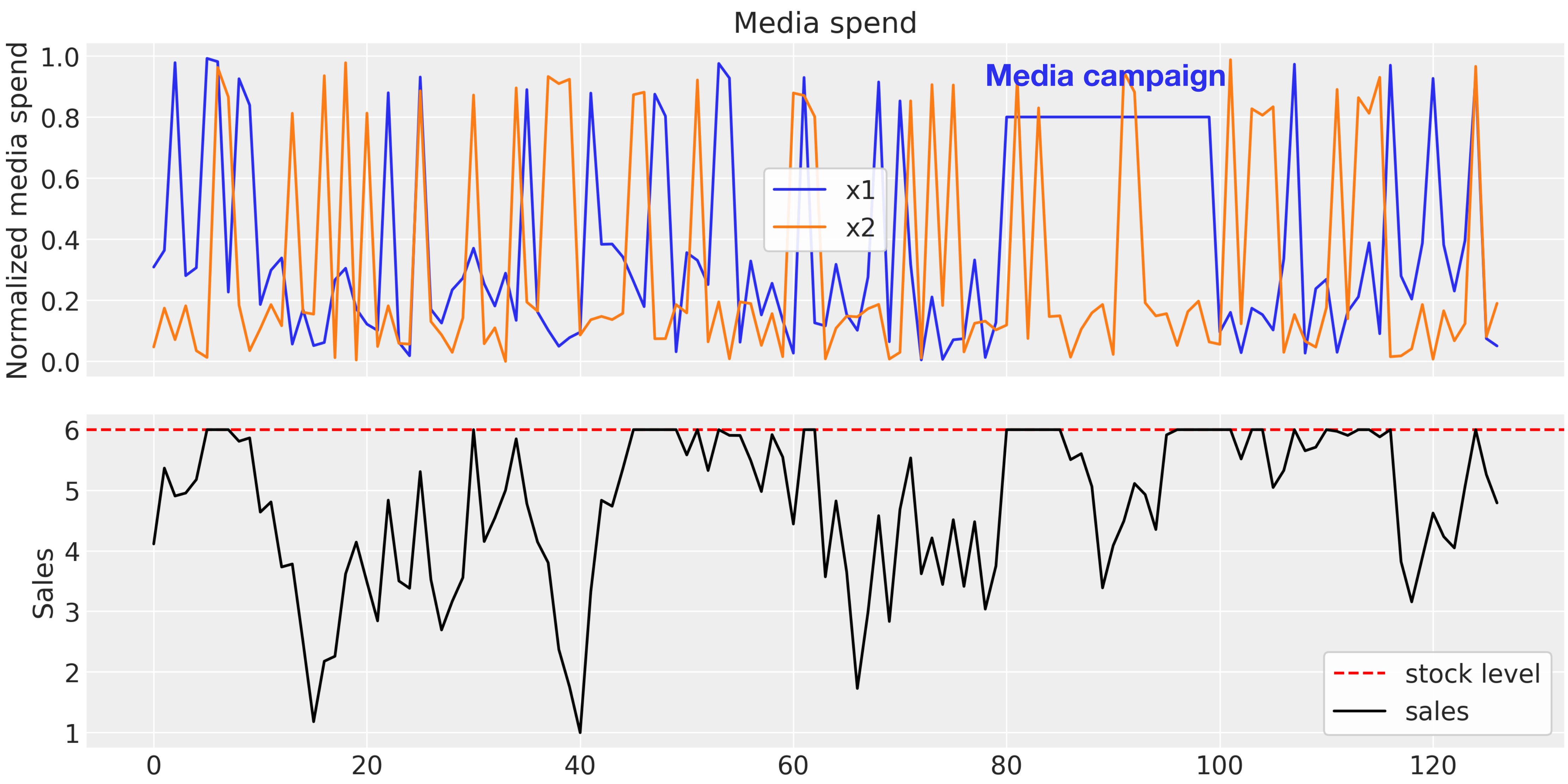
Motivating example:
Latent demand estimation

Limited stock levels

Stockouts
Sales = stock level
Demand?

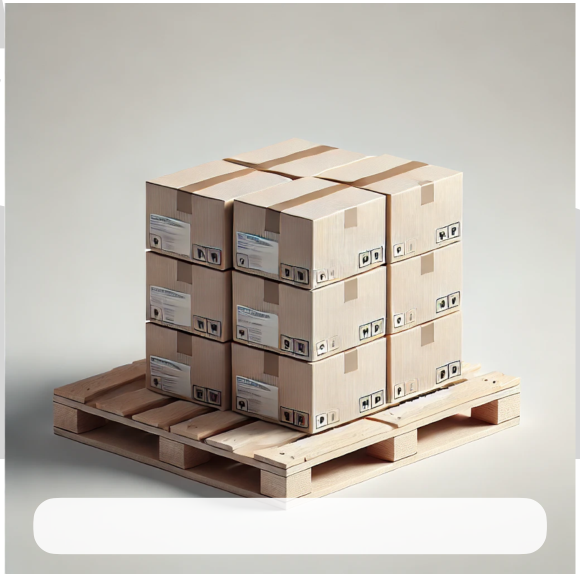


Overstock
Sales < stock level
Demand ≈ Sales

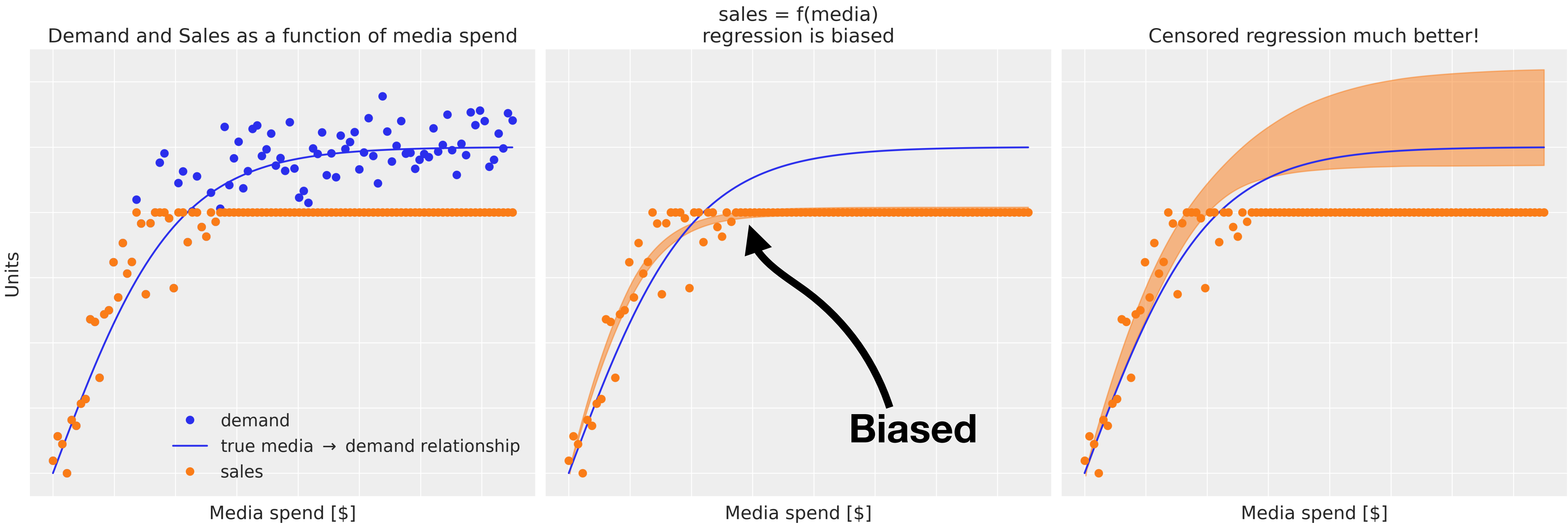


Media spend
Media campaign
x1
x2
Normalized media spend
Sales
stock level
sales

SNAKE
FLAVOR



Stockouts cause problems



We think statistics can solve our stockout problem (censored regression)

But what kind of business insights do we want?

Can we forecast based on previous observations?

By how much did last quarter's marketing campaign drive up demand and sales?

What is the relationship between media spend and demand?

Can we forecast what will happen if we run another media campaign?

Can we forecast based on previous observations?

By how much did last quarter's marketing campaign drive up demand and sales?

What is the relationship between media spend and demand?

Can we forecast what will happen if we run another media campaign?

Pearl's causal ladder

Step 3
Counterfactuals

By how much did last quarter's marketing campaign drive up demand and sales?

Step 2
Intervention

Can we forecast what will happen if we run another media campaign?

Step 1
Correlation / Forecasting

What is the relationship between media spend and demand?

Can we forecast based on previous observations?

Pearl's causal ladder

Step 3
Counterfactuals

By how much did last quarter's marketing campaign drive up demand and sales?

Step 2
Intervention

Can we forecast what will happen if we run another media campaign?

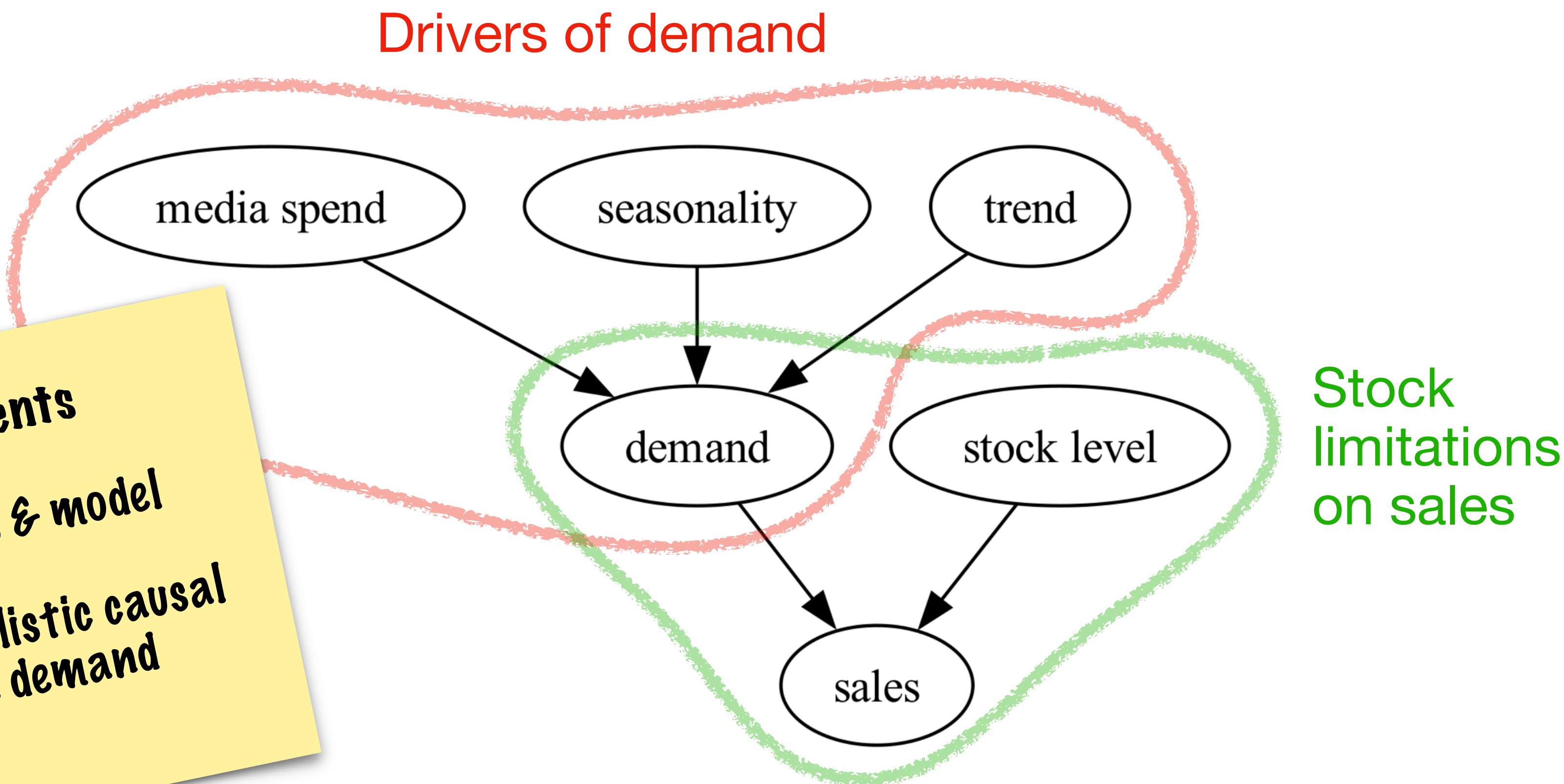
Step 1
Correlation / Forecasting

What is the relationship between media spend and demand?

Can we forecast based on previous observations?

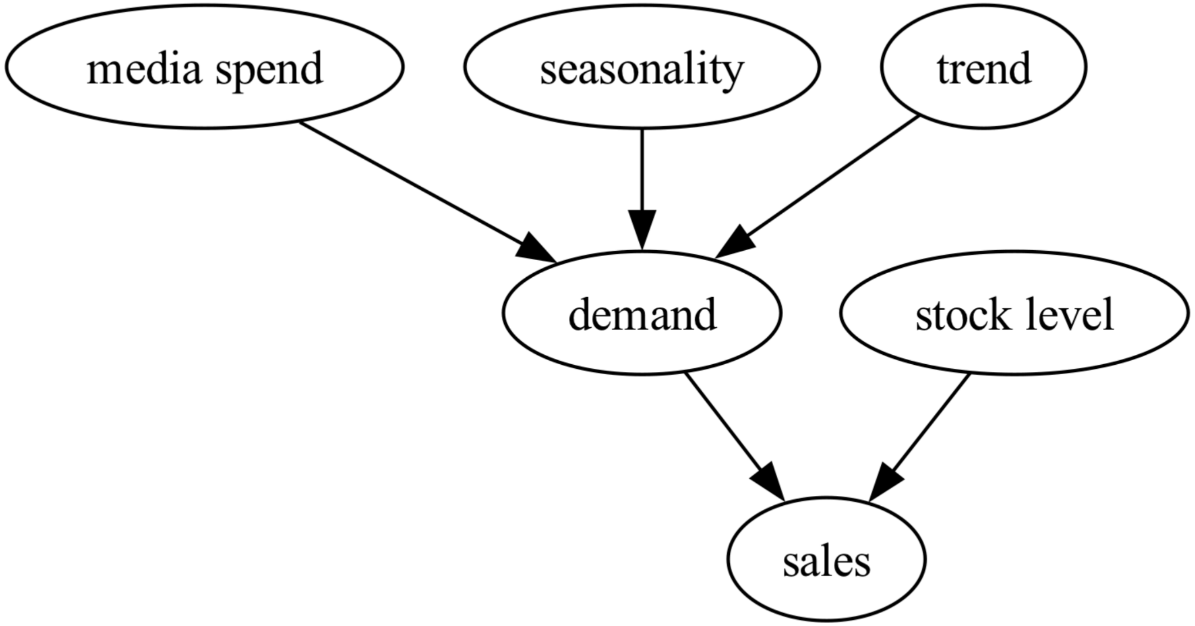
Causal thinking
&
Structural Causal Models

Let’s build a Structural Causal Model



Improvements

- Add price drivers & model price elasticity
- Incorporate realistic causal relations in the demand funnel



And make it Bayesian by placing priors on the parameters

intercept

media spend + media transformations

other predictors (seasonality, trend)

$$\text{demand}_t = \alpha + \Sigma_{m=1}^{M} \beta_m f(x_{m,t}) + \Sigma_{c=1}^{C} \gamma_c z_{c,t}$$

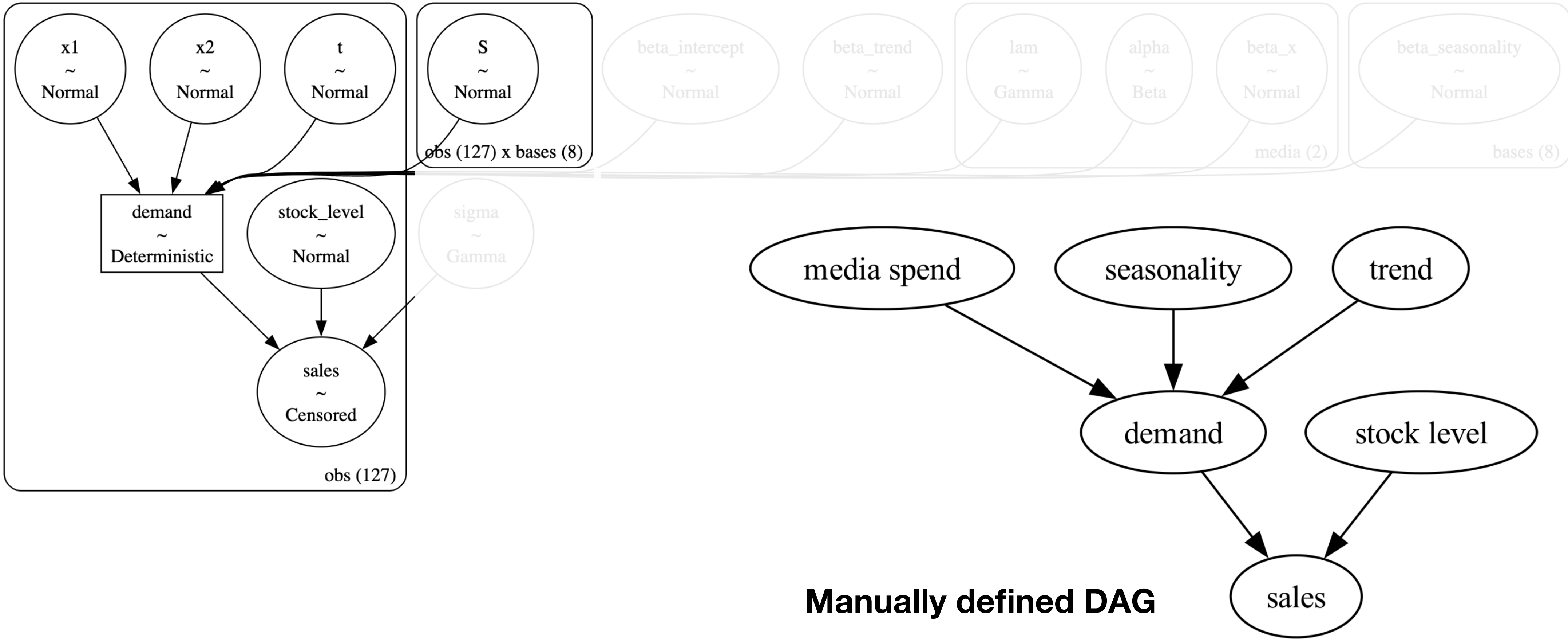
$$\text{sales}_t \sim \text{CensoredNormal}(\text{demand}_t, \sigma, \text{upper} = \text{stock}_t)$$

```python
with pm.Model(coords=coords) as model:
    t = pm.Normal("t", dims="obs")
    x1 = pm.Normal("x1", dims="obs")
    x2 = pm.Normal("x2", dims="obs")
    S = pm.Normal("S", dims=("obs", "bases"))
    stock_level = pm.Normal("stock_level", dims="obs")

    beta_intercept = pm.Normal("beta_intercept", mu=0, sigma=1)
    beta_trend = pm.Normal("beta_trend", mu=0, sigma=1)
    beta_x = pm.Normal("beta_x", mu=0, sigma=1, dims="media")
    beta_seasonality = pm.Normal("beta_seasonality", mu=0, sigma=1, dims="bases")
    sigma = pm.Gamma("sigma", 1, 1)
    alpha = pm.Beta("alpha", alpha=1, beta=1, dims="media")
    lam = pm.Gamma("lam", alpha=1, beta=1, dims="media")

    x1_saturated = apply_media_transformations(x1, alpha=alpha[0], lam=lam[0], l_max=8)
    x2_saturated = apply_media_transformations(x2, alpha=alpha[1], lam=lam[1], l_max=8)

    demand = pm.Deterministic(
        "demand",
        beta_intercept
        + beta_trend * t
        + beta_x[0] * x1_saturated
        + beta_x[1] * x2_saturated
        + pm.math.dot(S, beta_seasonality),
        dims="obs",
    )

    normal_dist = pm.Normal.dist(mu=demand, sigma=sigma)
    censored_normal = pm.Censored(
        "sales",
        normal_dist,
        lower=0.0,
        upper=stock_level,
        dims="obs",
    )
```

PyMC

pymc.io

model.to_graphviz()

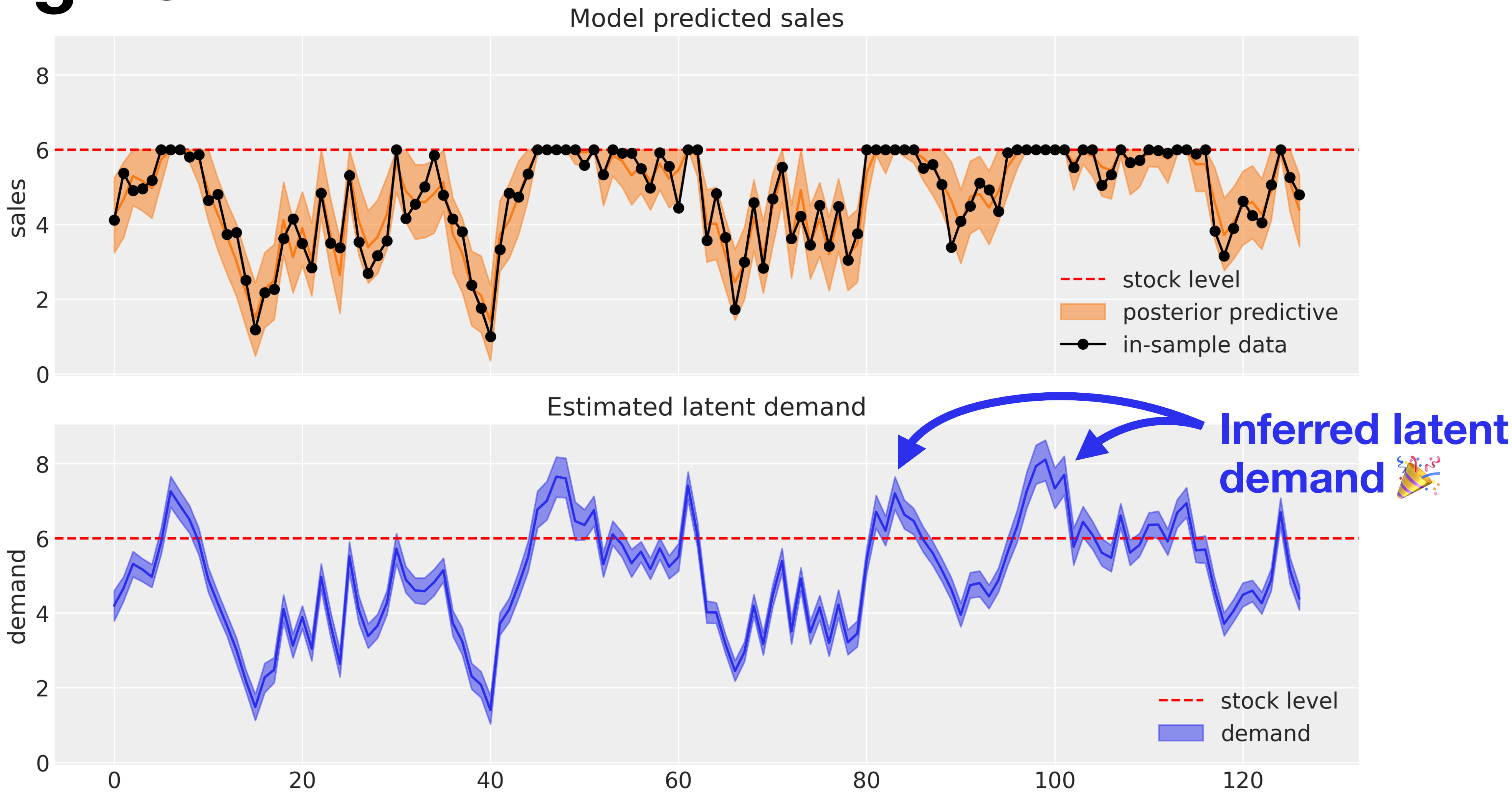


Manually defined DAG

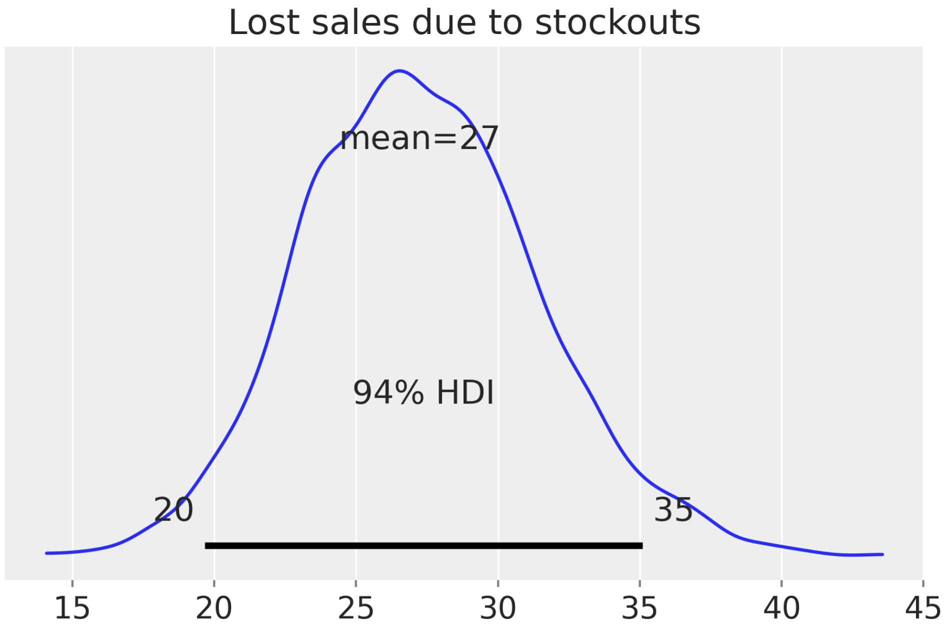
Fit the model

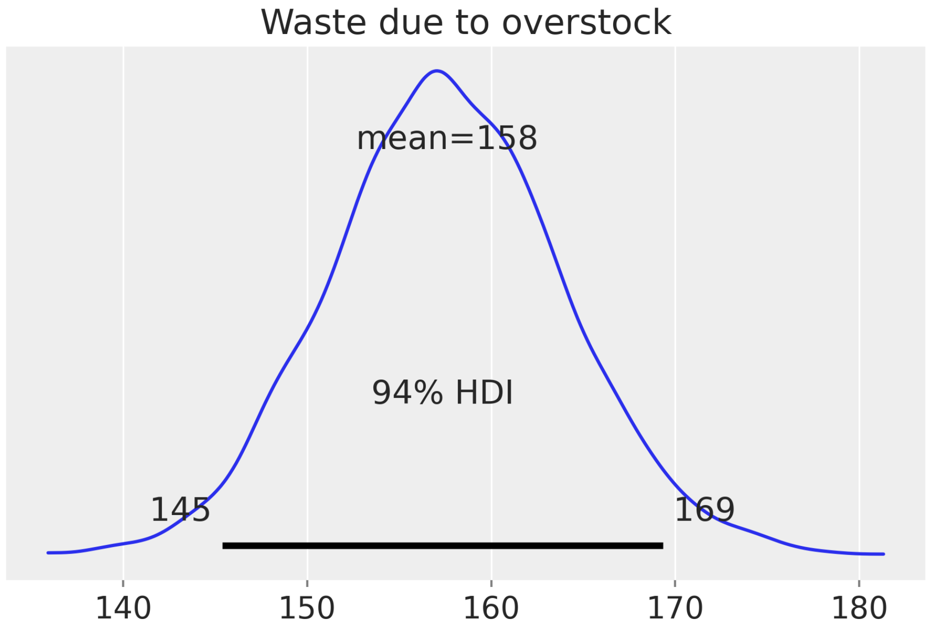
```python
obs = {
    "t": exogenous["t"],
    "x1": exogenous["x1"],
    "x2": exogenous["x2"],
    "S": extract_seasonality_to_array(exogenous),
    "stock_level": exogenous["stock_level"],
}

# do operator for exogenous variables
inference_model = pm.do(model, obs)

# observe operator for outcome
inference_model = pm.observe(inference_model, {"sales": exogenous["sales"]})

with inference_model:
    idata = pm.sample()
    idata.extend(pm.sample_posterior_predictive(idata, var_names=["demand", "sales"]))
```

Insights


Model predicted sales
sales
stock level
posterior predictive
in-sample data
Estimated latent demand
demand
stock level
demand
0
2
4
6
8
0
20
40
60
80
100
120
Inferred latent demand 🎉

Insights





calculate lost revenue

calculate wasted investment

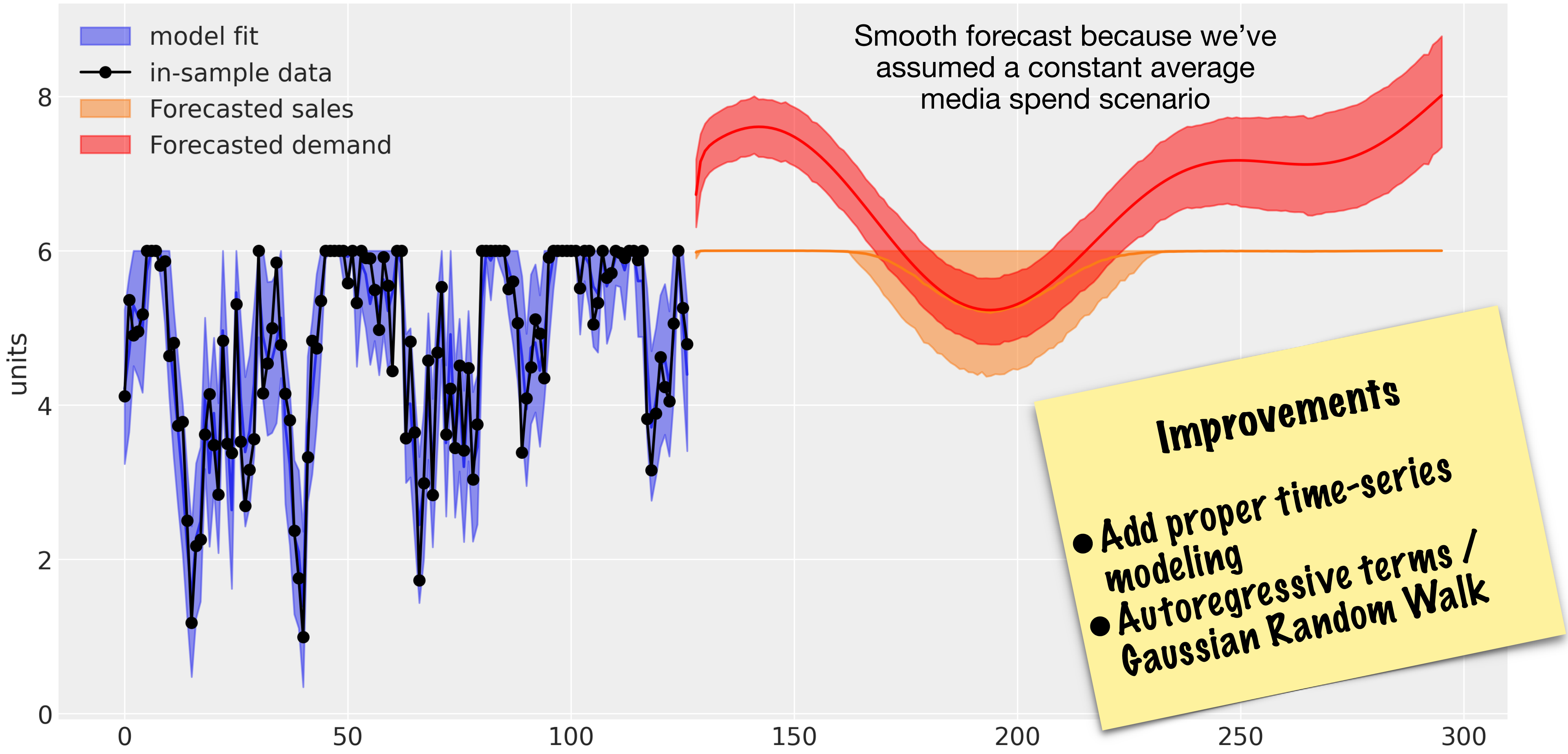
profit optimisation through demand and supply management

Step 1: Forecasting demand

We don’t know future marketing spend, so we will guess at the average.

```
intervention = {
    "t": t,
    "x1": exogenous["x1"].mean() * np.ones(n_new_periods),
    "x2": exogenous["x2"].mean() * np.ones(n_new_periods),
    "S": extract_seasonality_to_array(forecast_df),
    "stock_level": exogenous["stock_level"].mean() * np.ones(n_new_periods),
}

new_model = pm.do(model, intervention)

with new_model:
    mcmc_samples = pm.sample_posterior_predictive(mcmc_samples)
```

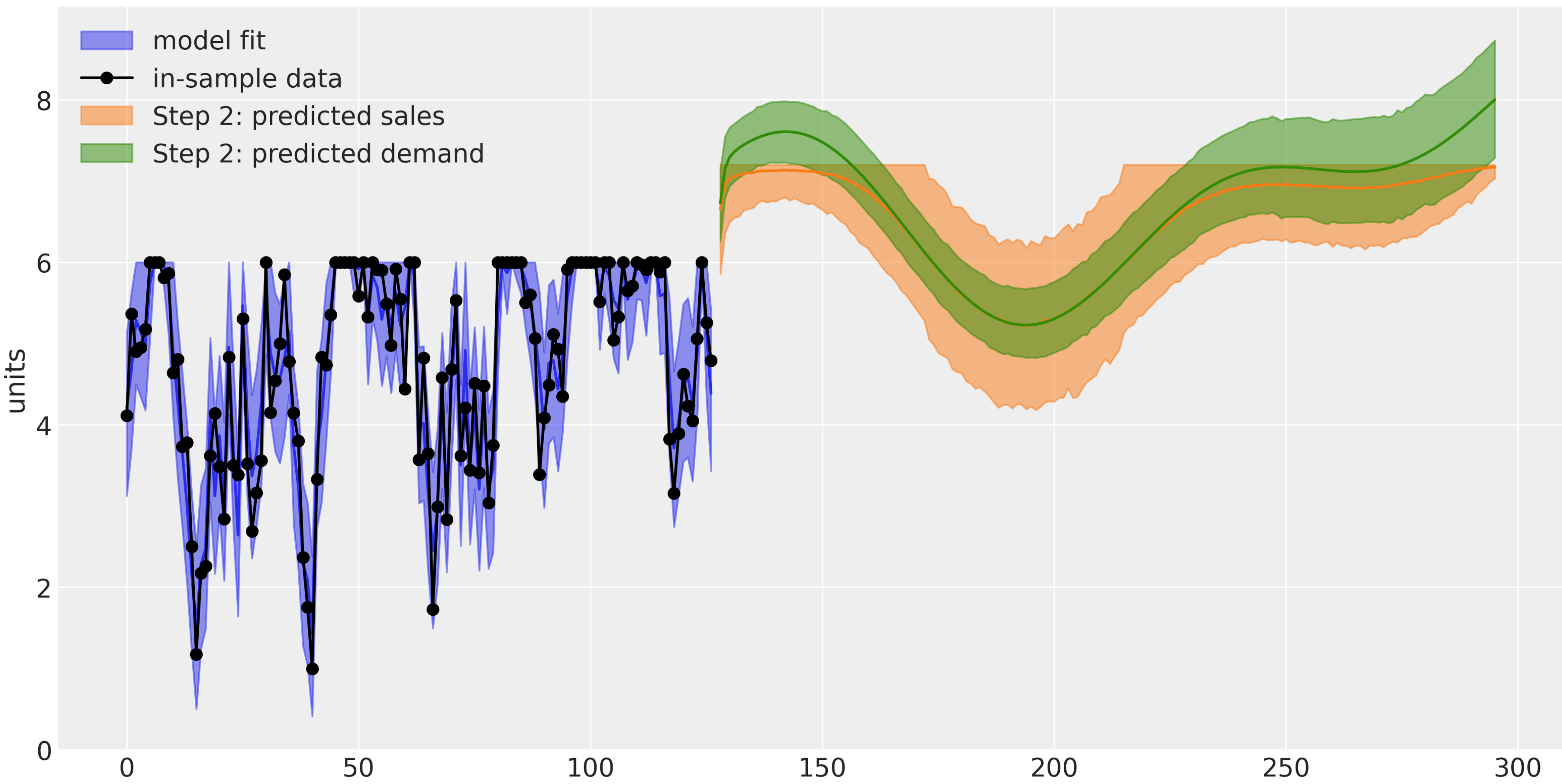
Step 1: Forecasting demand



Step 2: What if we increase stock levels?

```
intervention["stock_level"] = params["stock_level"] * 1.2 * np.ones(n_new_periods)

model_step2 = pm.do(model_step1, intervention)

with model_step2:
    idata_step2 = pm.sample_posterior_predictive(idata, var_names=["demand", "sales"])
```

Step 2: What if we increase stock levels?

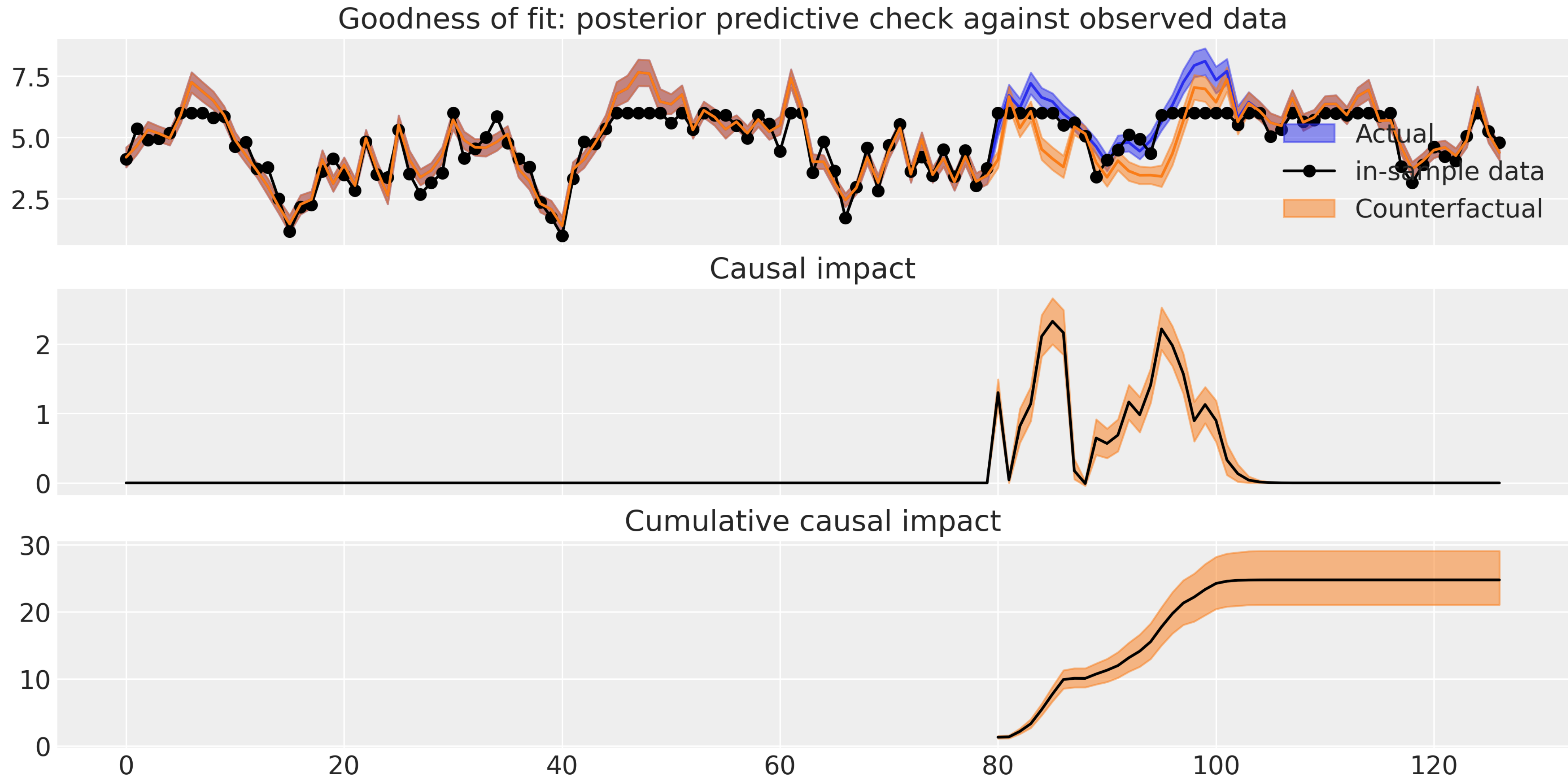


Step 3: What was the impact of the campaign?

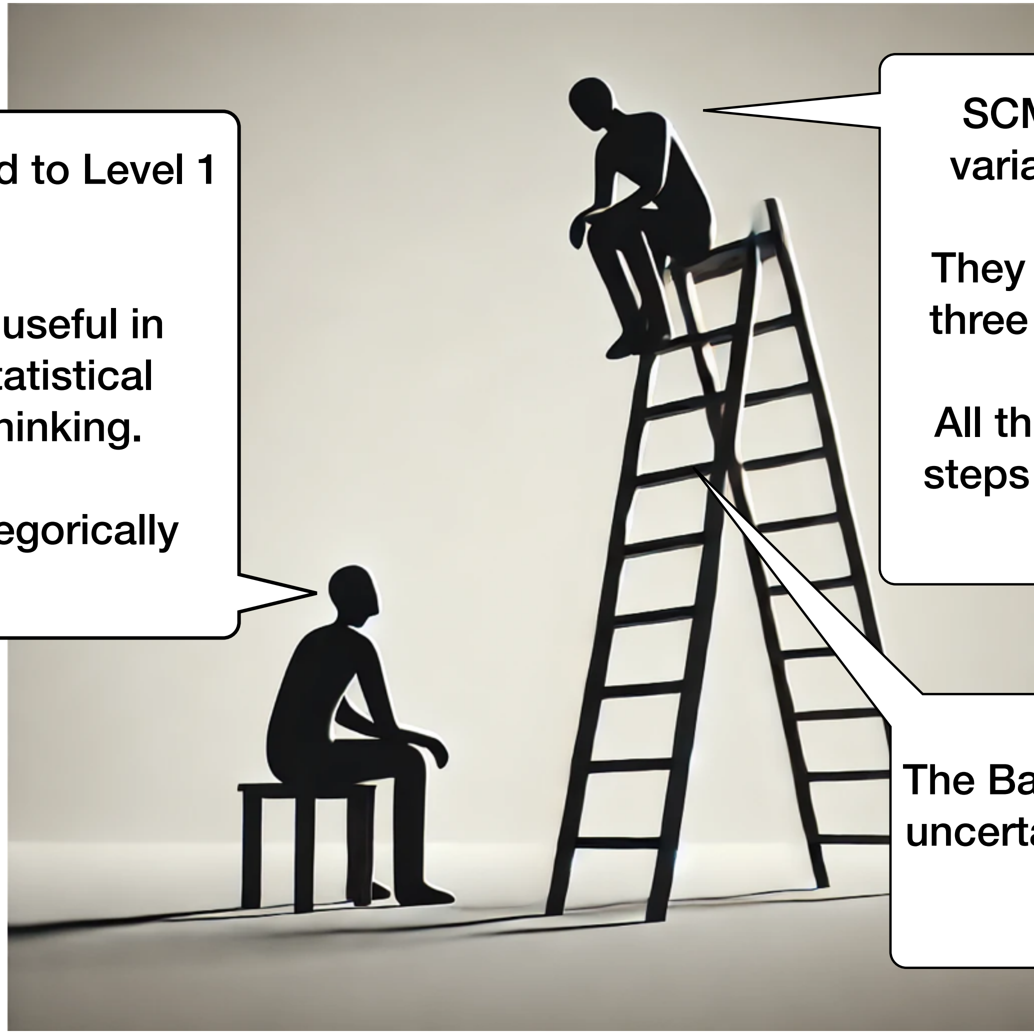
What the marketing spend would have been, in the absence of the marketing campaign

```
counterfactual = {"x1": exogenous[“x1_counterfactual"]}

new_model = pm.do(model, counterfactual)

with new_model:
    mcmc_samples = pm.sample_posterior_predictive(fitted_mcmc_samples)
```

Level 3: Counterfactuals



Summary



I am sad to be restricted to Level 1 questions.

The causal ladder is useful in moving from from statistical thinking to causal thinking.

Each step seems categorically different.

SCM's fully describe how variables causally interact.

They allow you to explore all three levels of Pearl's ladder.

All that changes on different steps is the way in which you query your SCM

The Bayesian element quantifies uncertainty and injects business knowledge.

Thanks!
Any questions?

Reach out to me on LinkedIn
to explore how these approaches can
help answer your business questions

drbenvincent/pydata-global-2024