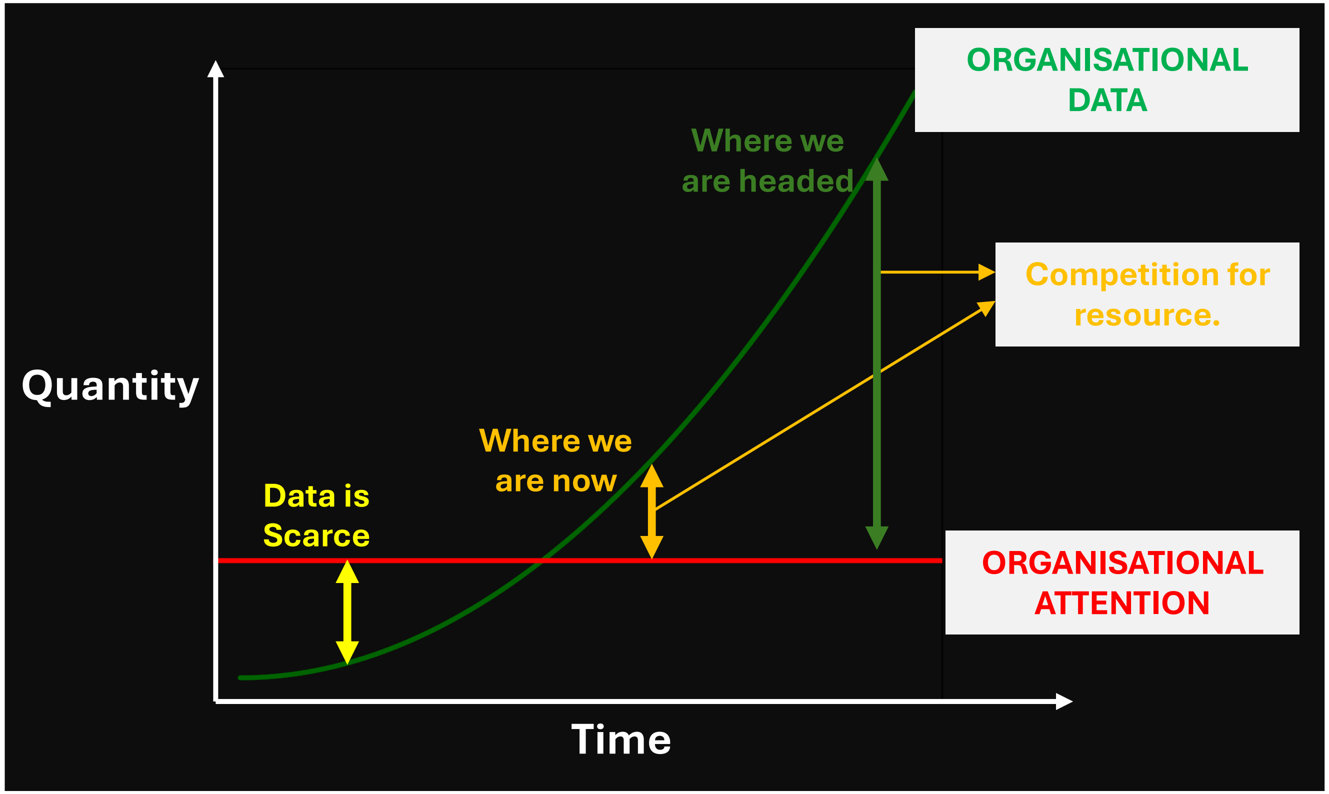
Lessons from the wild: Data Science @ Triumph Motorcycles

Dr. Shashwat M. Pande

Things I will talk about

1. Some preliminaries, setting the scene and so on.
2. Some things we did at Triumph.
3. Some things we learnt (by having done some things) at Triumph.
4. Some questions, concluding remarks and acknowledgements.

1.1) Data is...the new oil?



Volume

Lots and lots and lots...and lots of data!

Variety

Not just spreadsheets, not just numbers, not just held in relational databases.

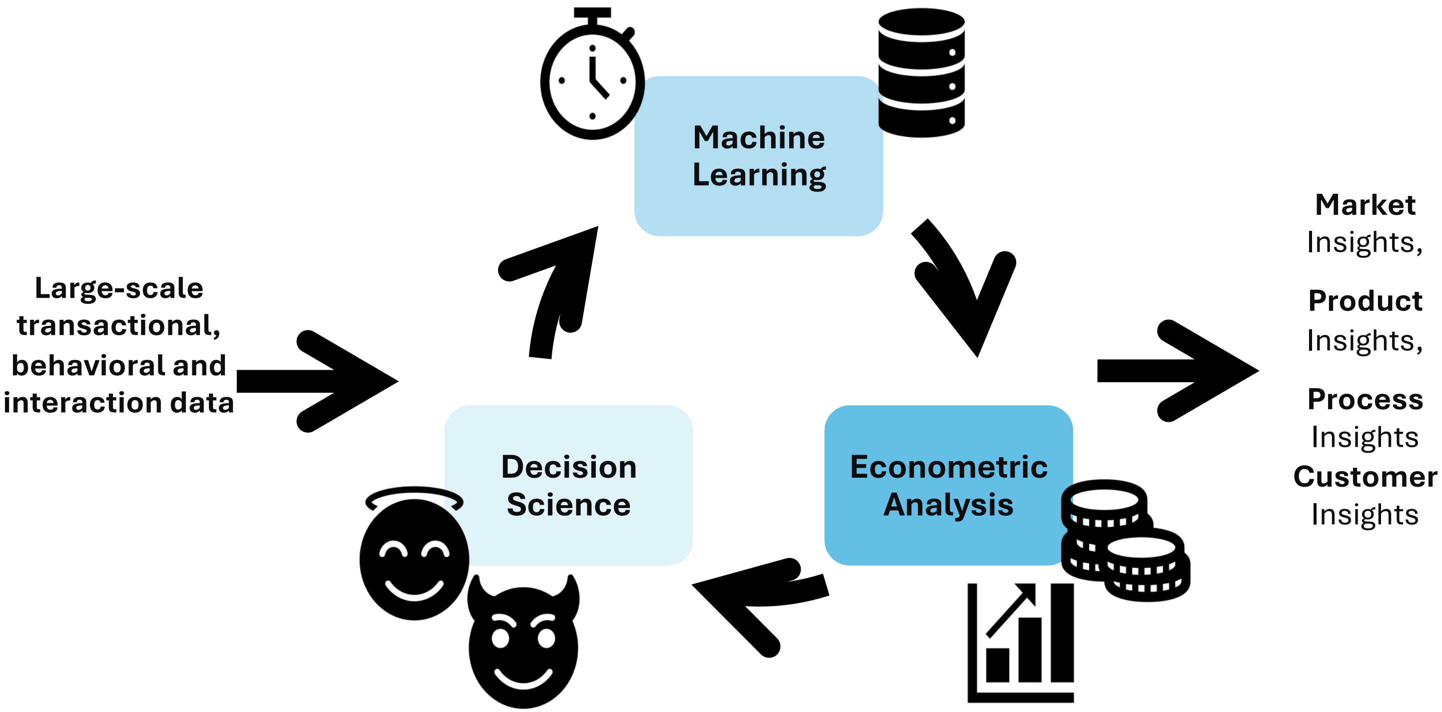
Velocity

Collected every day, hour, minute, second, transaction, interaction etc.

(**V**eracity)

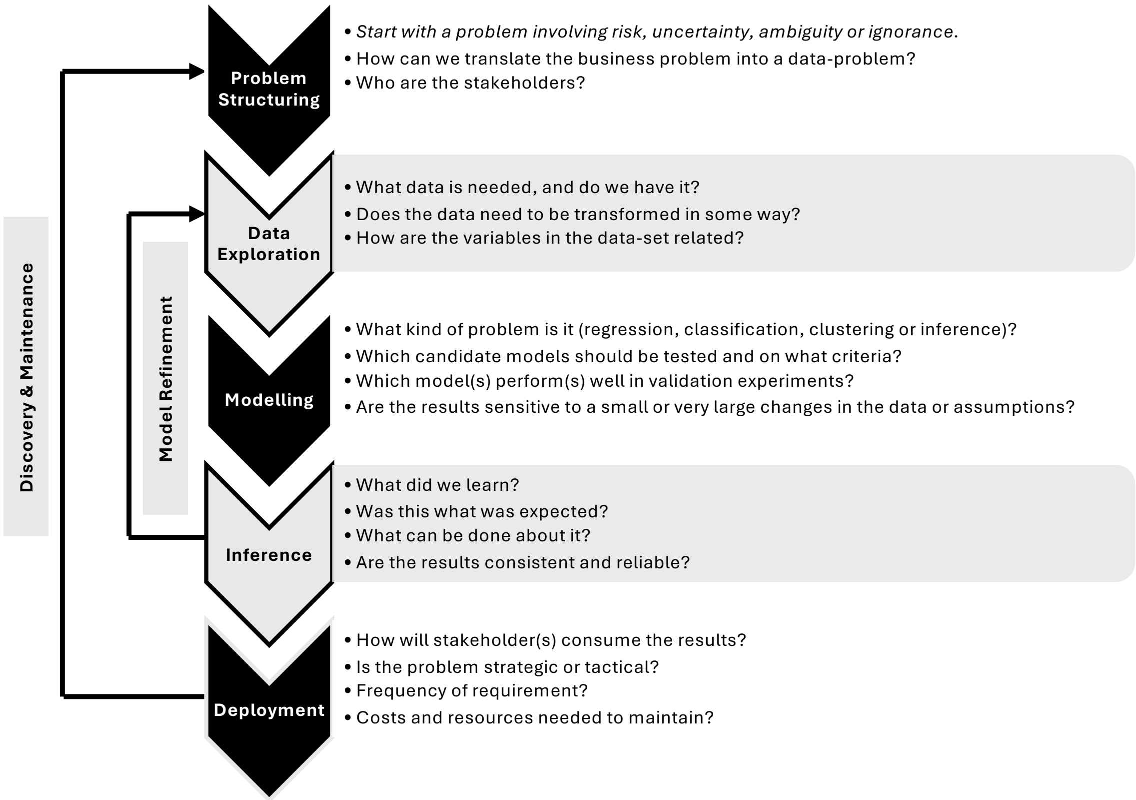
The ugly step-sister that eventually gets the guy, determines everything but, whom we don't hear as much about until the final act.

1.2) The elevator pitch

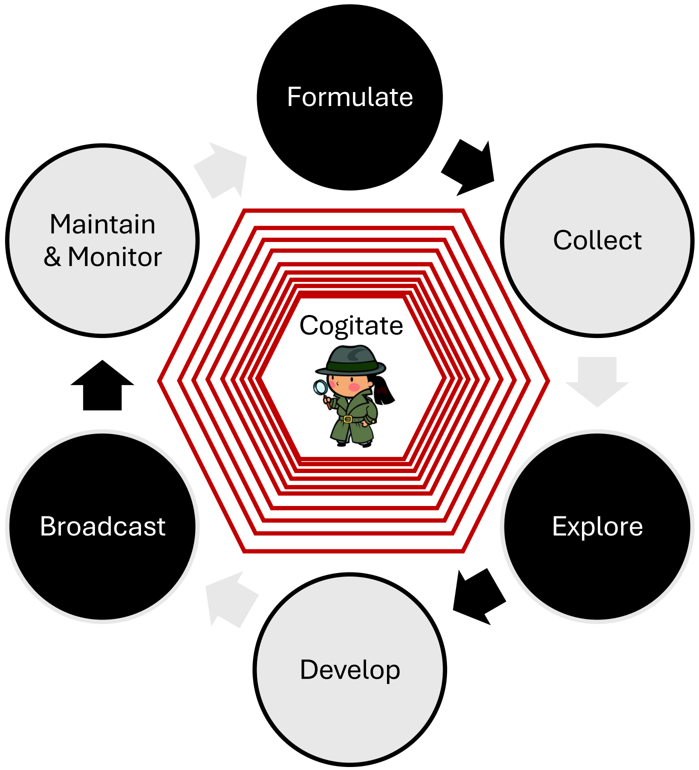


Adapted from: Mortenson et al. (2015). Operational research from Taylorism to Terabytes: A research agenda for the analytics age. *European Journal of Operational Research*, *241*(3), 583-595.

1.3) To whom it may concern



Archetypal Process (see also CRISP-DM)



Typical Workflow

2) Some things we did at Triumph

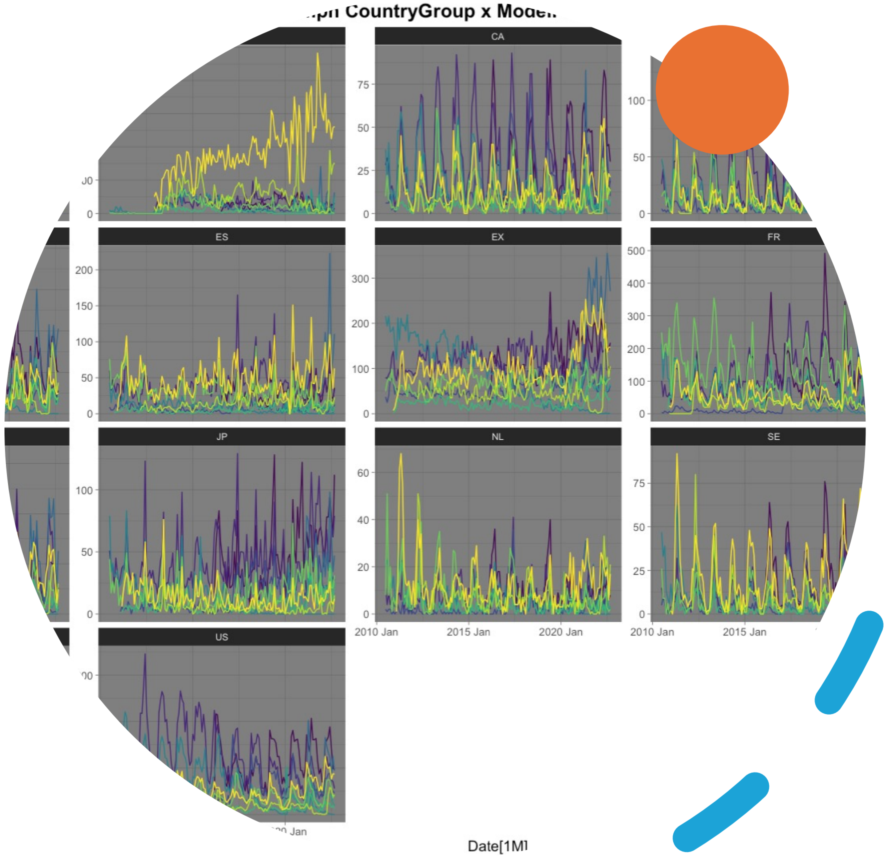
1. **Forecasting**: *Hierarchical sales forecasting for demand prediction*.
2. **Unsupervised-Learning**: *Detecting critical safety-threats and reputational-risk in online rider communities*.
3. **Geo-spatial Analysis**: *Mapping census-scale demographics and commuter networks*.
4. **Product Analytics**: *What's the word on the Tiger 1200?*

2.1) Forecasting

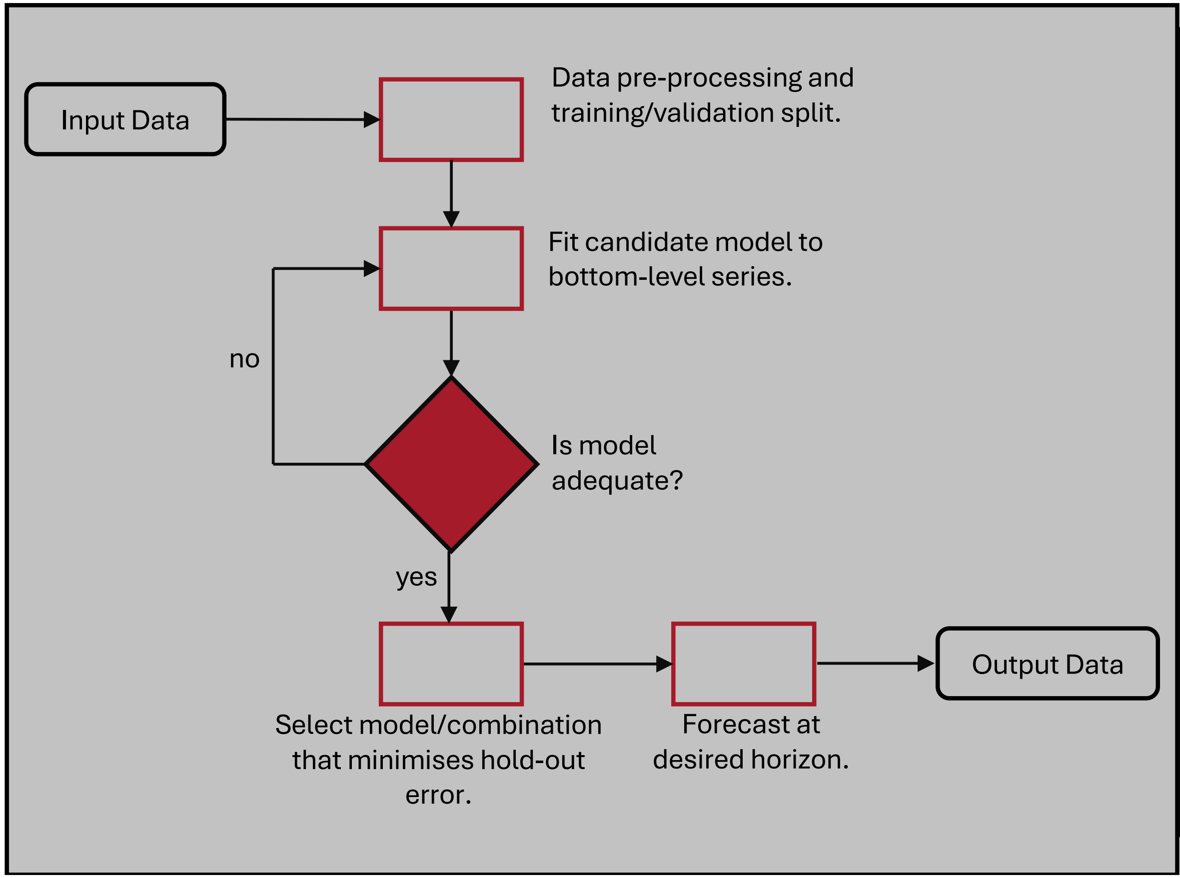
Problem Definition:

1. Triumph's expansionary targets signal an exciting period of growth for the company.
2. Rapid expansion of core product-portfolio alongside a growing dealer network.
3. Judgement alone cannot *reliably* and *reproducibly* forecast product-demand on a continuous basis.

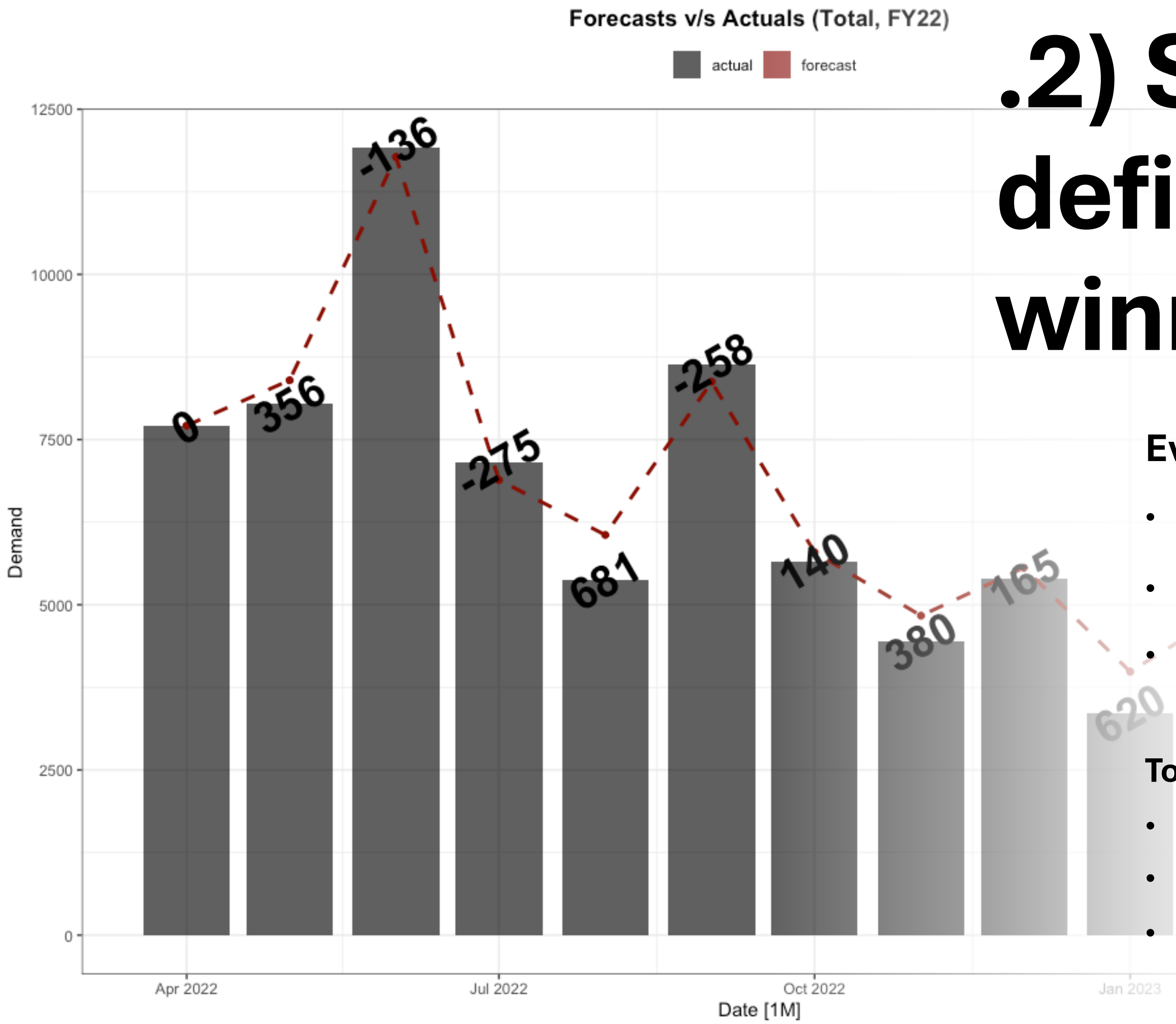
Use-Case: An automated forecasting procedure that can reliably estimate monthly demand and serve as a supplement to sales-force consensus negotiations.



.1) Schematic of the approach



SPLIT	*TRAIN, TEST AND VALIDATION SAMPLES.*
TRAIN	*ESTIMATE MODEL PARAMETERS.*
EVALUATE	*HOW WELL DID EACH MODEL DO ON THE TEST DATA?*
SELECT	*CHOOSE A MODEL THAT SATISFIES THE OBJECTIVE.*
PREDICT	*USE THE MODEL TO PREDICT FUTURE VALUES.*



.2) Some definitions and a winning chart

Evaluation

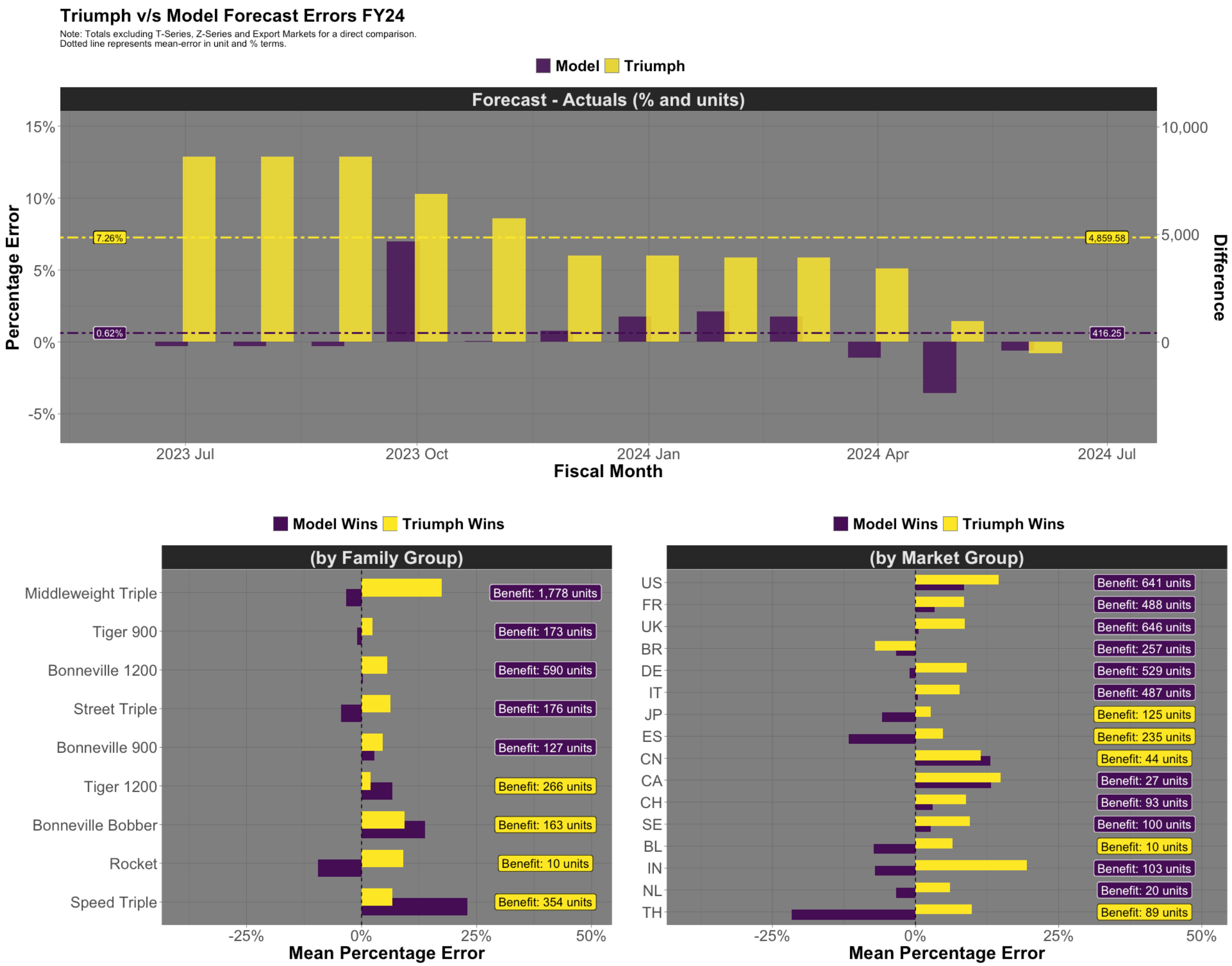
- **MAPE** = $\frac{100}{n}\sum_{t=1}^{n}\left|\frac{Forecast_t - Actual_t}{Actual_t}\right|$
- **MAE** = $\frac{1}{n}\sum_{t=1}^{n}|Forecast_t - Actual_t|$
- **Bias** = $\frac{1}{n}\sum_{t=1}^{n} Forecast_t - Actual_t$

Top-Level Forecast Error

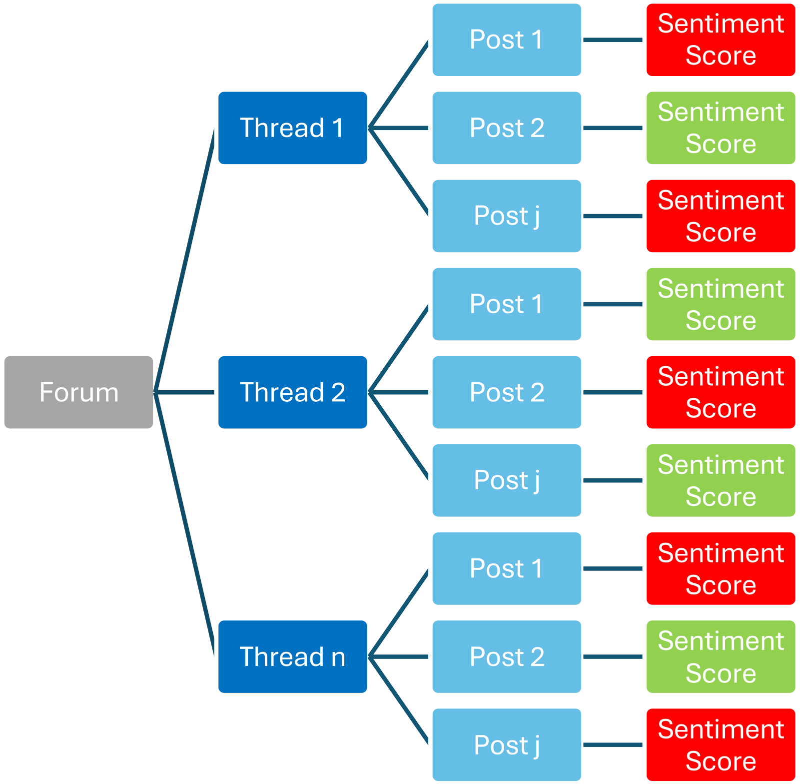
- **MAPE** = (+/-) **6.4%**
- **MAE** = (+/-) **392 units**
- **ME (Bias)** = (+) **55 units**

.3) Human v/s model judgement

- Putting a **£** on confidence.
- Testing your predictions
  - ➢ What's the appropriate benchmark?
  - ➢ What does (over)-optimism look like?
  - ➢ Who is the right audience?



2.2) Unsupervised learning



Problem Definition:

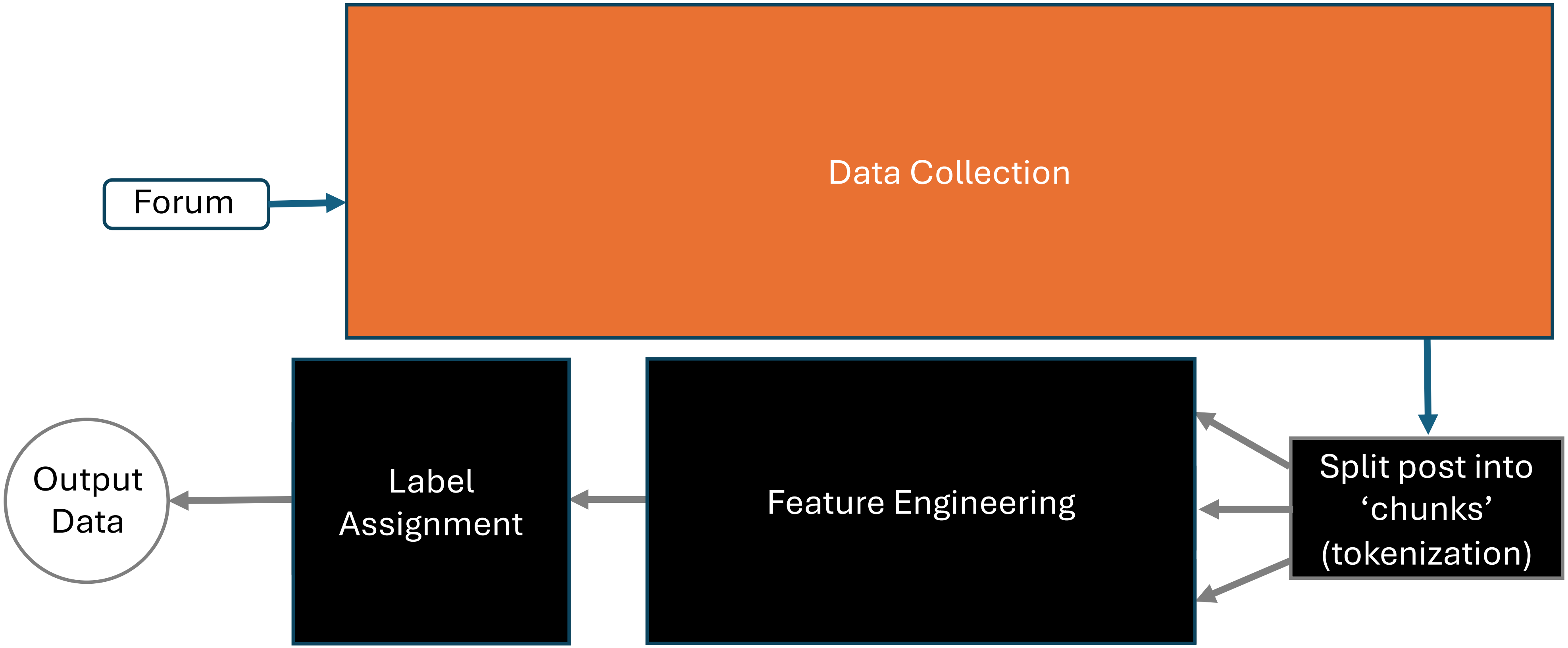
1. Customers discuss product experiences online and often report issues with products and services.

2. Useful information is prohibitively difficult to identify in the 'blizzard of buzz' by human intervention alone.

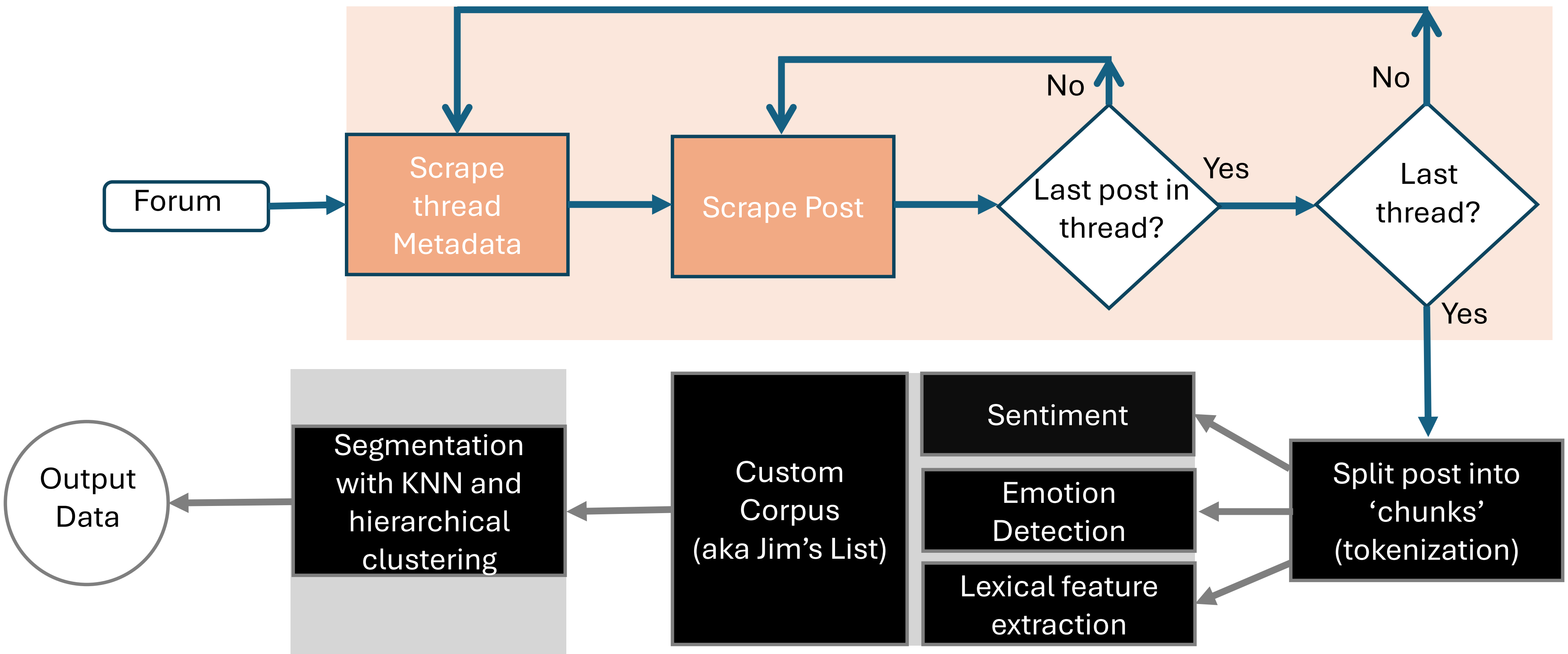
3. We can supplement human judgement by providing a method to target problem areas using text analytics and unsupervised learning.

Use-Case: *A proactive approach to mitigating reputational-risk by detecting real or perceived safety-threats and litigious intent in online rider-community discussions.*

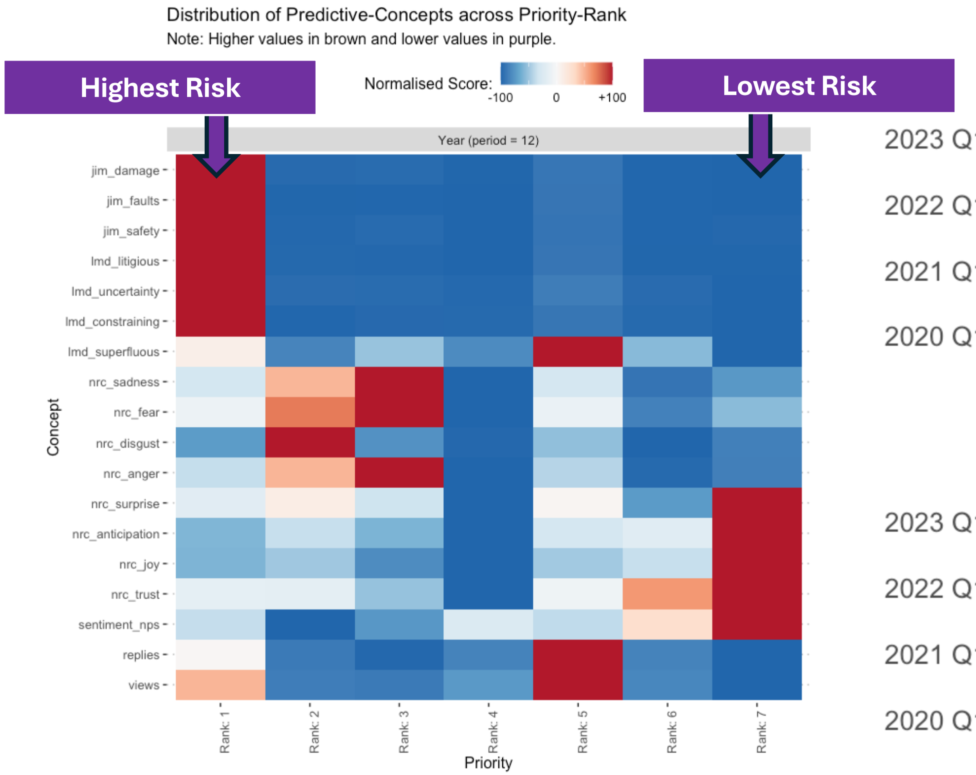
.1) Schematic Representation of the Approach



.2) The Nuts and Bolts



.3) Measured Speech



2.3) Geospatial analysis

<u>Motorcycle Nation</u>: Motorcycle Trips in England and Wales
(source: 2011 origin-destination dataset, ONS)

193K trips daily, 0.9% of the commuting population ride-to-work.

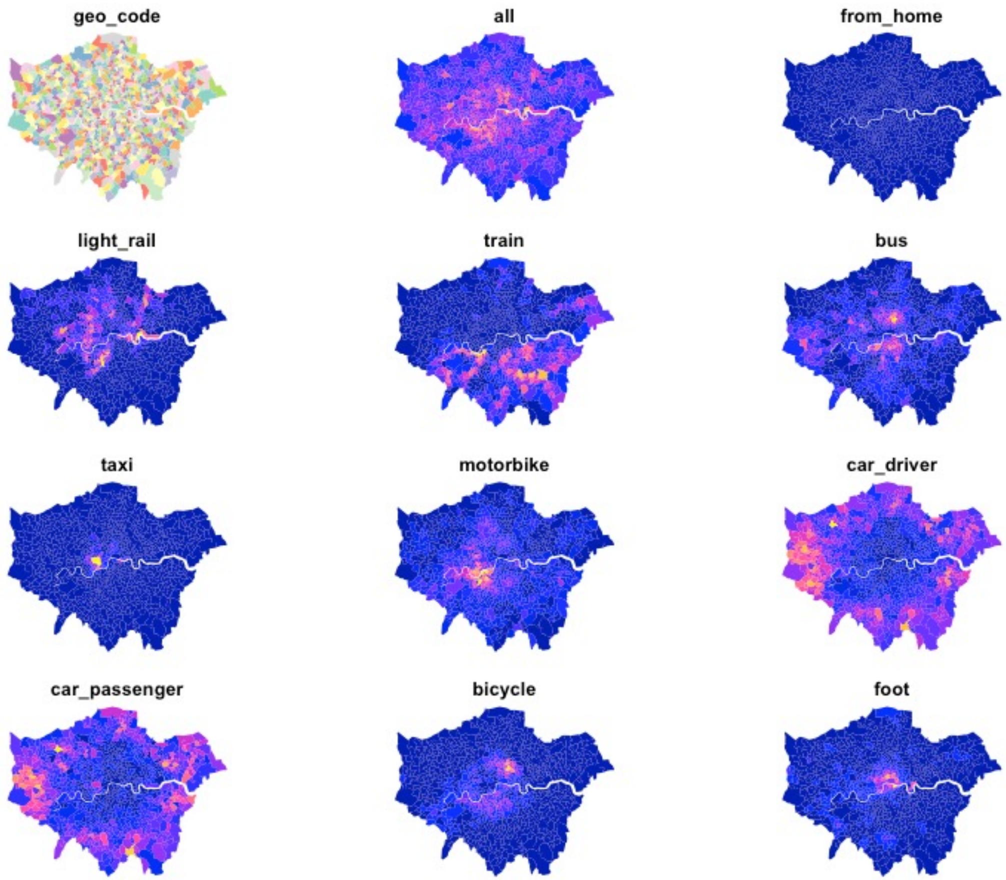
Problem Definition:

1. Triumph's recently announced entrance to the small-displacement commuter segment.
2. New market in which Triumph lacks experience and market insight.
3. Bringing external data to bear to help target field research.

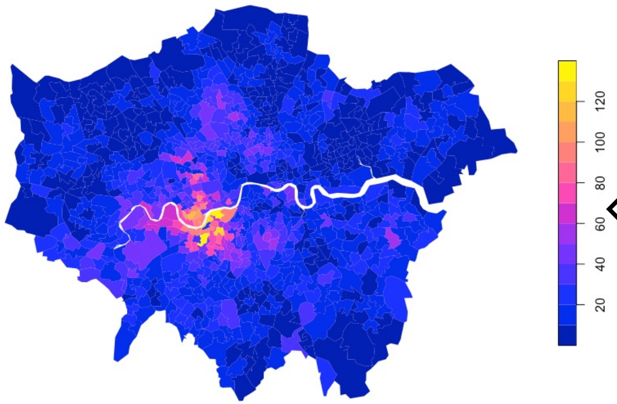
Use-Case: *Explore a new market and segment to generate a local ideal of what a 'customer' might look like.*

.1) Looking at London

- ❑ Travel flows **reveal aspects about demographic and infrastructural makeup**.
- ➢ Train-users come from the south and north-east.
- ➢ Car drivers depart from the city periphery.
- ➢ Foot traffic is concentrated in central areas.
- ➢ Some taxi-hailers and motorcyclists seem to have a few things in common!

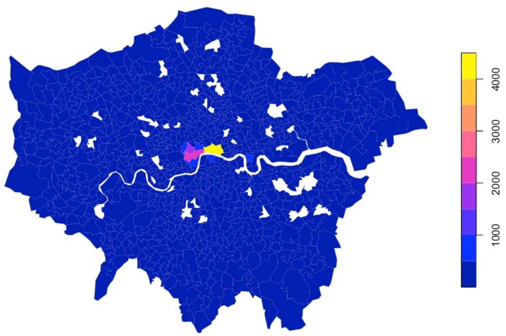


.2) London riders: way they come and where they go...



Departing from

Arriving here



DEPARTURES

Hammersmith & Fulham.

Kensington & Chelsea.

Wandsworth.

Westminster.

ARRIVALS

City of London.

Westminster.

Camden.

Southwark.

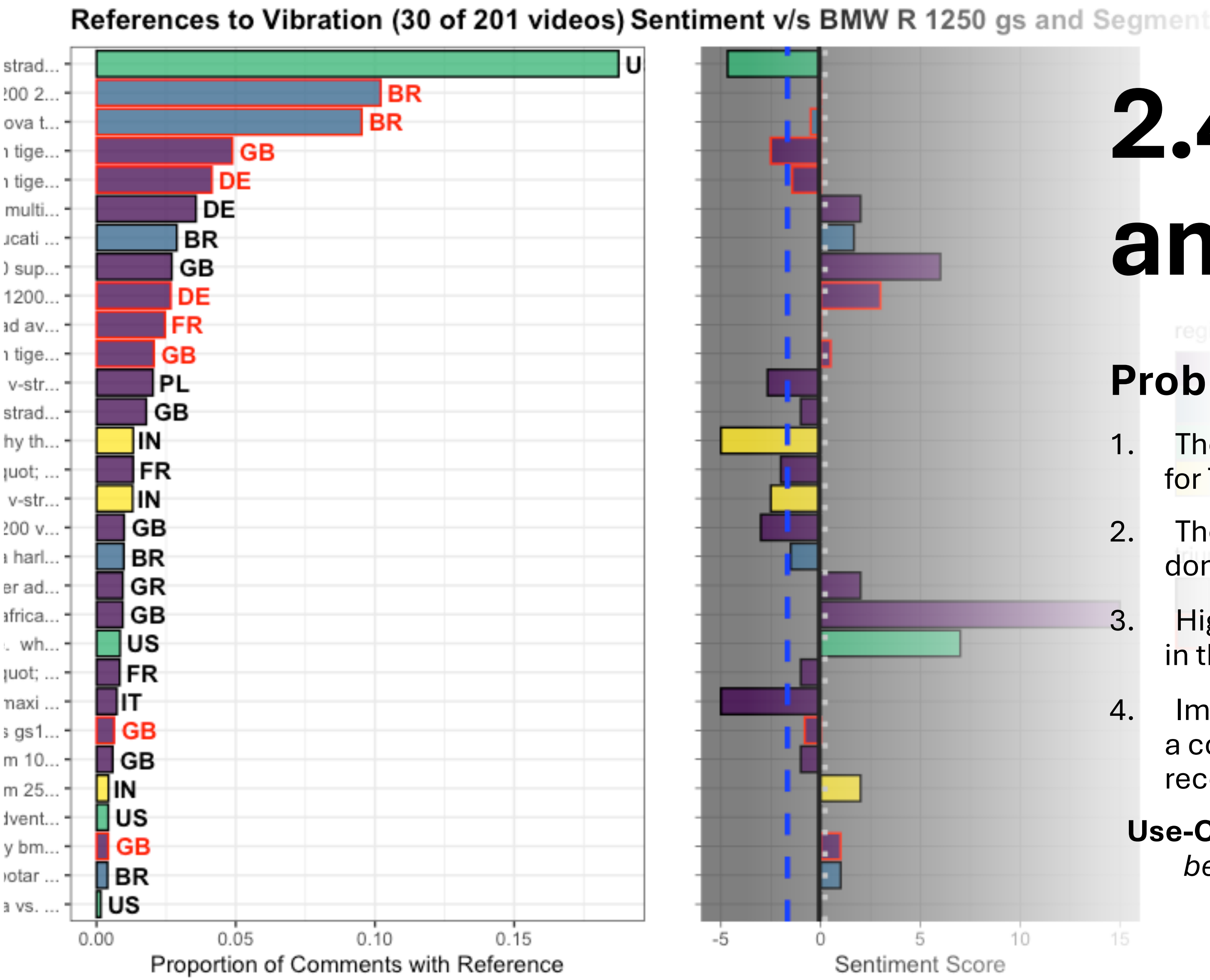
Descriptors

- ❑ Ethnically diverse
- ❑ High working-age population.
- ❑ High CO_2 emissions.
- ❑ High 'Gross Value Added'.
- ❑ High house-prices and low housing-stock.

Demographics

- ❑ > 2-4x disposable income v/s national-average.
- ❑ Established tech-workers (23%)
- ❑ Highly-qualified professionals (13%).
- ❑ White-collar professionals (13%)

Any marketers in the room?



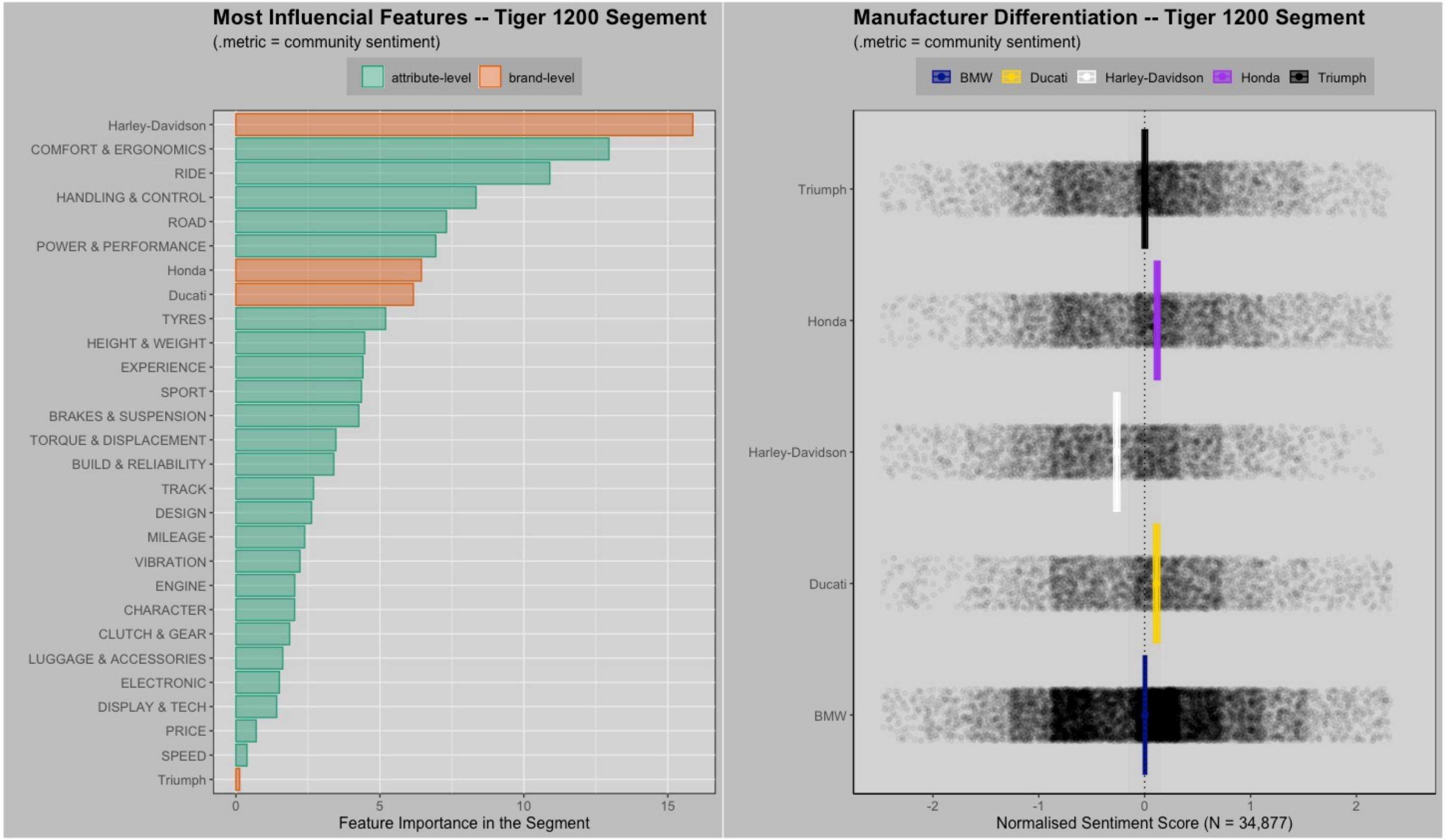
2.4) Product analytics

Problem Definition:

1. The Tiger 1200 was a big release for Triumph.
2. The segment is historically dominated by BMW.
3. High-margins and high volumes in the segment.
4. Important strategic product from a commercial and brand-recognition standpoint.

Use-Case: *Event based competitor benchmarking and product intelligence.*

.1) What's the sample size?

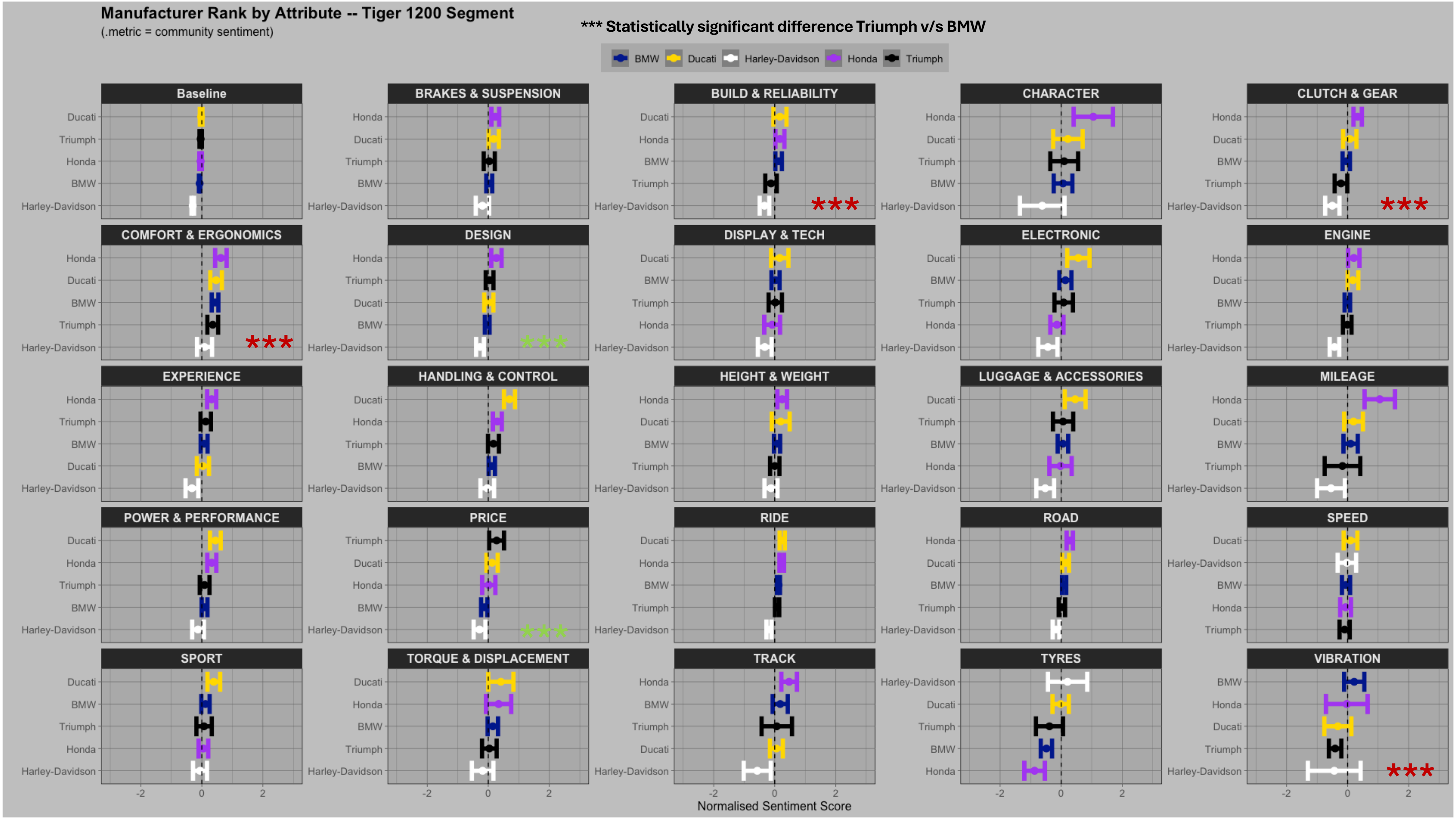


The panel to the left are results from a regression model predicting **comment-level sentiment** based on an attribute x manufacture combination.

- The level of analysis is a comment.
- 1-week post-launch.

Observations

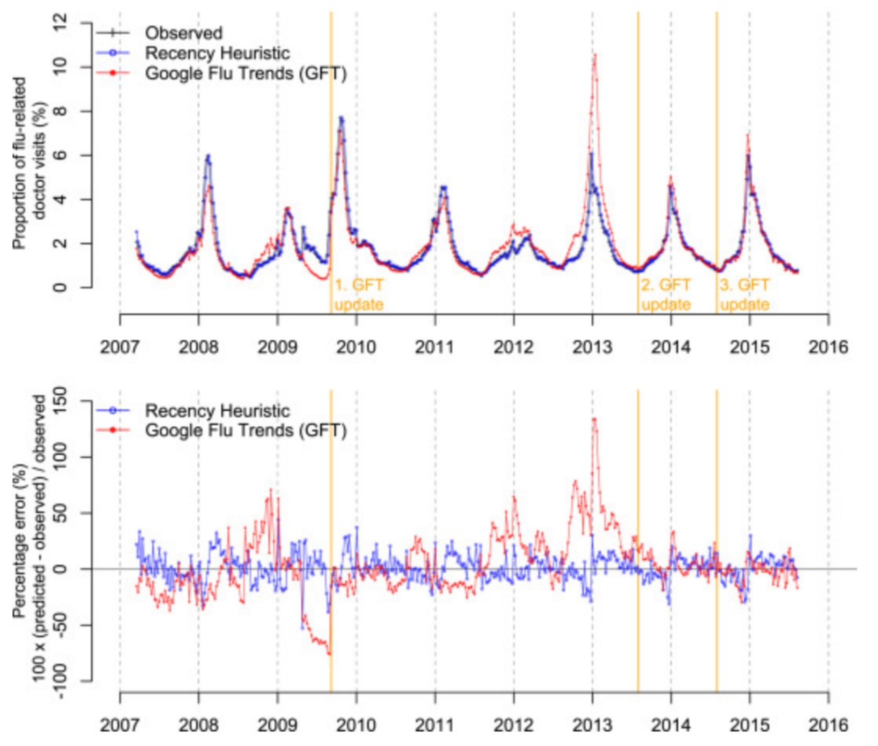
- Comfort is the most important feature.
- Price ranks low in the list suggesting the segment may not be particularly price-sensitive.
- Brand differentiation is high.
- Triumph **is perceived less differentiated from BMW than the remaining three competitors** in the segment. – the right panel shows the statistical distribution.



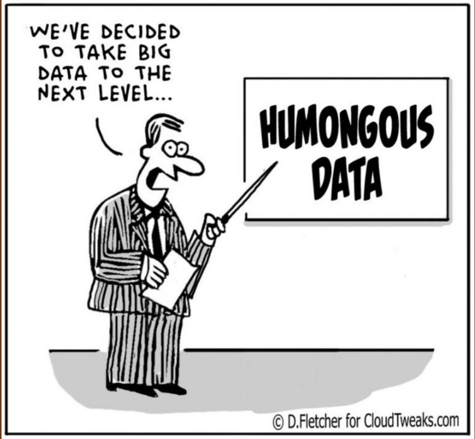
Manufacturer Rank by Attribute -- Tiger 1200 Segment
(.metric = community sentiment)
*** Statistically significant difference Triumph v/s BMW
BMW
Ducati
Harley-Davidson
Honda
Triumph
Baseline
BRAKES & SUSPENSION
BUILD & RELIABILITY
CHARACTER
CLUTCH & GEAR
COMFORT & ERGONOMICS
DESIGN
DISPLAY & TECH
ELECTRONIC
ENGINE
EXPERIENCE
HANDLING & CONTROL
HEIGHT & WEIGHT
LUGGAGE & ACCESSORIES
MILEAGE
POWER & PERFORMANCE
PRICE
RIDE
ROAD
SPEED
SPORT
TORQUE & DISPLACEMENT
TRACK
TYRES
VIBRATION
***
Normalised Sentiment Score


3) Some things we have learnt
(by having done some things) at
Triumph Motorcycles

3.1) "…not everything that can be counted counts, and not everything that counts can be counted."– *William Cameron (1963)*



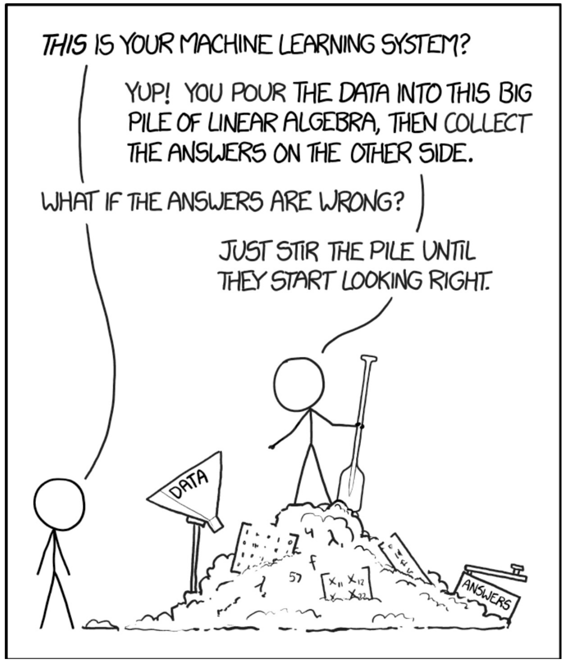
Influenza Prediction: Google Flu Trends v/s seasonal naïve model ("recency heuristic") – *Katsikopoulos et.al. (2022)*

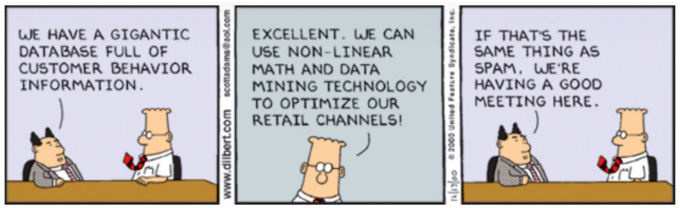


- Big Data v/s the right data
  - More data is not always better!
  - Finding the right data is a challenging task and involves a great deal of participative modelling.
  - Going from the right data to good data even more so.
  - Business users may often have a poor understanding and visibility over key-drivers and KPIs – managing 'the art of possible'.

3.2) “I’d prefer not to.” – *Bartleby (Melville, 1853)*

- **The significance of problem-structuring and decision science.**
  - Leaders want to solve the most difficult challenges using DS/ML.
  - Strategic challenges are often fuzzy, ***loosely defined*** and require substantial effort to formalize.
  - Involve large numbers of stakeholders with ***differing incentives*** and expectations.
  - ***Overweighting of features*** that might be complex and time-consuming to implement in practice and may only yield incremental (if any) accuracy improvements.





3.3) "All models are wrong, but some are useful." – *George Box*

- **Parsimony is often surprisingly effective.**

- 70% of development time is spent on data-collection, cleaning and management.
- 25% is spent communicating with stakeholders and understanding the decision-context.
- Model development is a small part of the story and for many challenges simple statistical algorithms perform surprisingly well.
- '***Fail fast and fail up***'
  - Focus on an initial minimum-feasible solution (PoC) and add complexity afterwards.
  - If you can't explain why it works, assume it doesn't, especially if the stakes are high.
  - Each failed experiment is another successfully collected data point.

THE CONVERSATION

Academic rigour, journalistic flair

COVID-19 Arts + Culture Business + Economy Cities Education Environment Health Politics + Society Science + Technology Podcasts

Rationality: research shows we're not as stupid as we have been led to believe

3.4) "Our deep learning neural network can predict the correct outcome 9 in 10 times." – *hyperbolic tech-sales exec.*

- **Identifying a good baseline and remaining vigilant for arbitrary metrics.**
  - What's the appropriate benchmark?
  - Constructing, selecting and tracking the correct KPIs can make a **BIG** difference in developing business-users' **intuition and trust** around model-supported insights.
  - Human judgment is a far more formidable opponent than H0.
    - Beating the experts can be very hard!
    - Why not combine forces instead?

Dog Cat

1 in 10?

Thank you for your attention!

Feel free to reach out: shashwatpande101@gmail.com