Exploiting small and big data for machine learning based incident prevention

08/06/2023

Topics for today

- BT Business – Who we are?
- The first challenge - Incident volume prediction
  - The troubles
  - The solution

The second challenge - Device failure prediction

- The troubles
- The infrastructure
- The solution

Who we are….

The company

BT Group plc

- BT is one of the UK's biggest telecommunications and network providers. We've also got a global presence in around 180 countries.

BT Business

- … is one of the leading providers of B2B connectivity and related services, with customers ranging from small businesses to multinational corporations and governments worldwide.

The team

BT Business Data Science Team

- 20 data scientists, data analysts and ML engineers
  ~9 data scientists in one team

The first challenge - Incident volume prediction

Situation

- "Incidents" are events when a service situation requires action on side of the BT Operations Team
- Incident resolution support is provided by teams that are dedicated to a fixed set of customers (1 or a few)

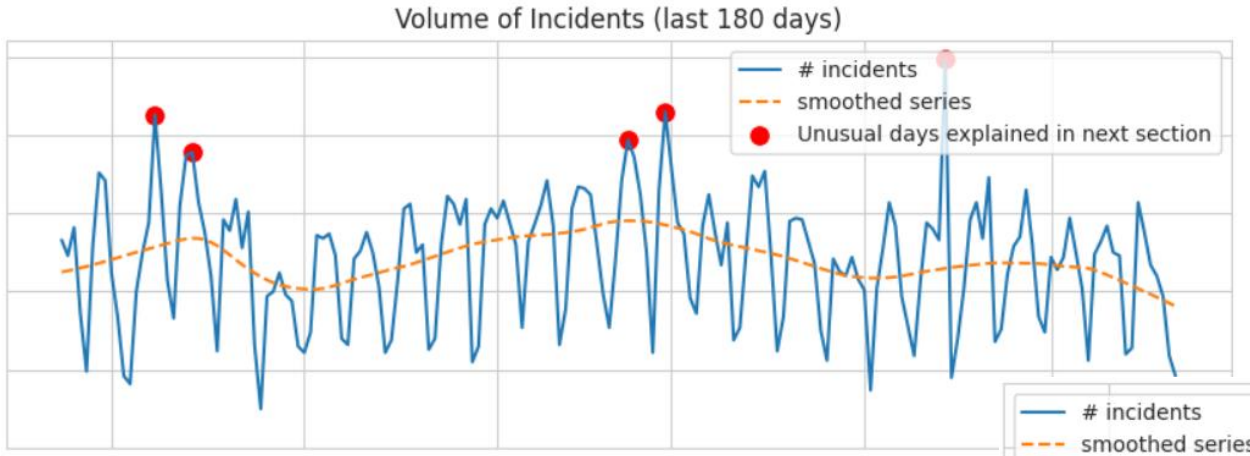
- There is some flexibility in scheduling the members for work

Mission

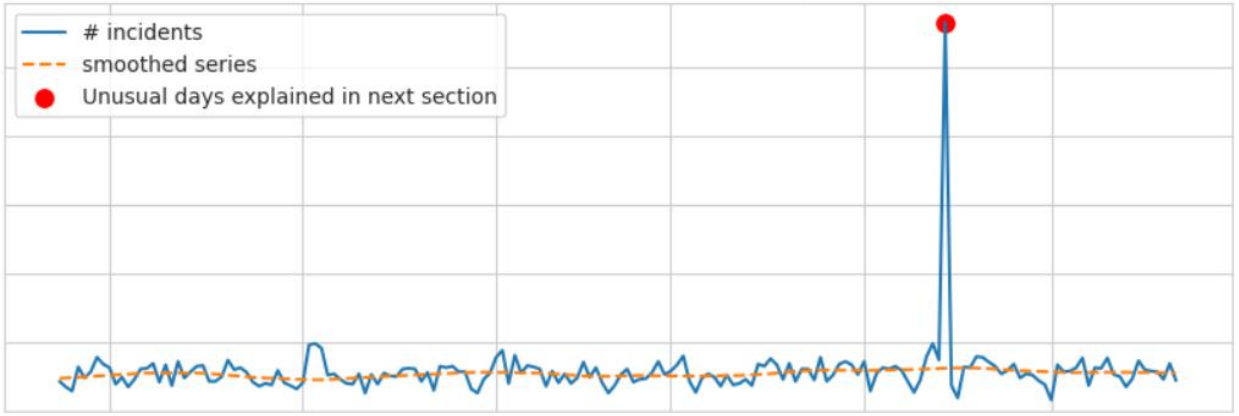
- Predict the number of incidents for the coming 6 weeks
  - Team total / separately for major customers
  - Total / broken down by major dimensions (severity, workday/weekend day, work shifts, etc)
- Within a short time

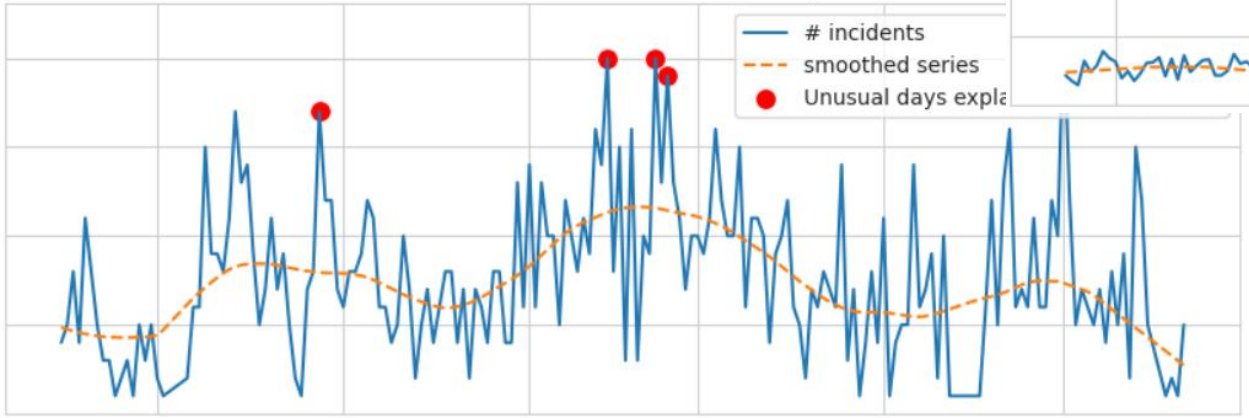
- (Highlight "trouble spots")
- (Provide information about process problems)

Volume forecasting: the troubles



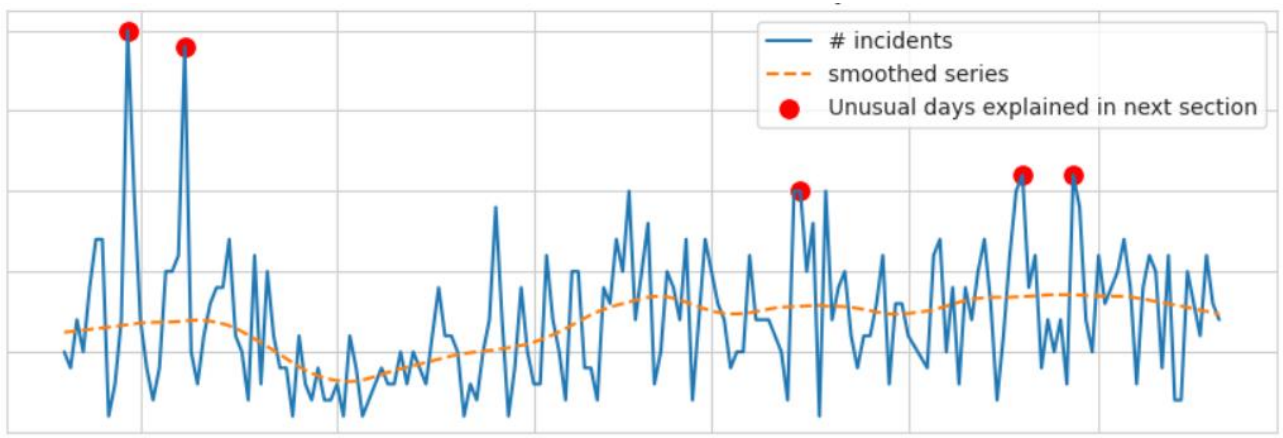
- Highly volatile incident number series
- Total lack of any seasonality, patterns
- Total lack of any forward looking information
  - Nothing at hand, except the time series

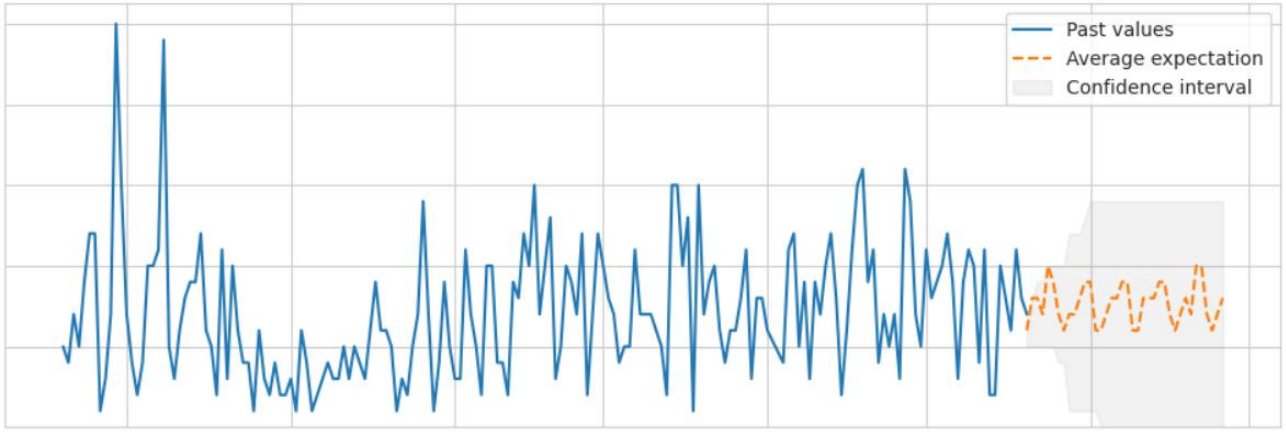




- Overlapping shift times
- Set of managers with different interests
- "I want one more report, that …."

Volume forecasting: the solution





- Not building a “model”
- Not attempting to predict (only) the raw time series
- Applying smoothing procedures and predicting both raw and smooth series
- Trying multiple forecasting methods
- Choosing the smoother * method combination that worked the best for the last month of actual data
- Simple “flatness” checks to avoid large number of forecasts
- Cheap proof-of-concept infrastructure: BigQuery – Jupyter Notebook – HTML report (plan for production: Qlik)

The second challenge - Device failure prediction

Situation

- "Device" = routers and switches
- Device statuses are monitored an (almost) uniform set of sensors reporting on a 5 minute basis
- "Failure" = device does not respond to health check pings for 15 mins
- There is supplementary health data coming from automatic alert systems (power supply status, communication protocol events)

Mission

- Predict the upcoming device failures with some lead time ahead
- … knowing that an external team of data scientist had been working on the problem for more than a year and failed….

Device failure forecasting: the troubles

Amount of sensor data

- 1 customer / 4k devices / 1 day sensor data = 15 GB, 32M reading values

Variations of sensor data

- E. g. 11 different indicators for CPU % utilisation

Time inconsistency of readings

- ± 10 secs easily, ± 2.5 mins happens

Hard availability of non-sensor data

- These are existing in their own process / database silos

Missing is normal

- What is more, a nonresponsive device can work normally otherwise

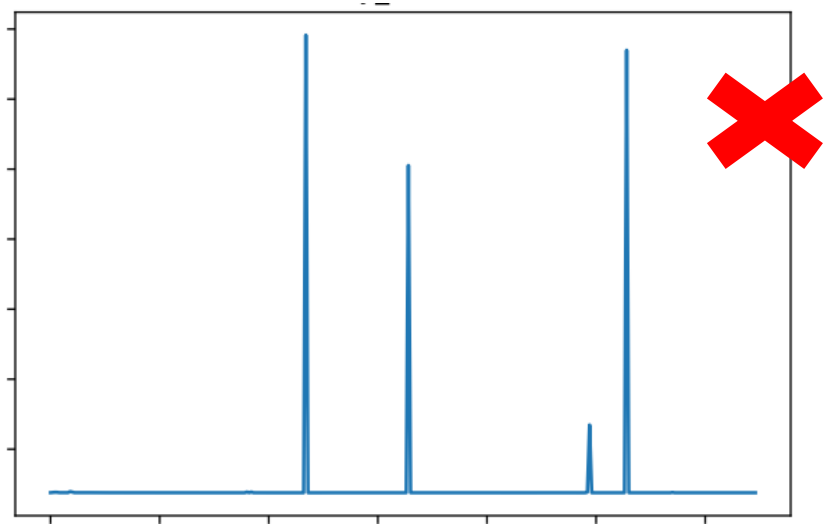
What is a “failure”?

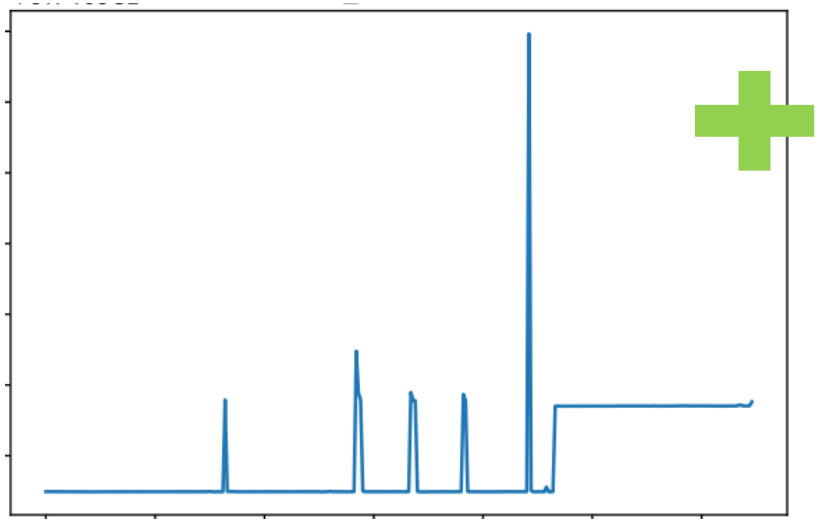
- Many definitions – but what does the data support?

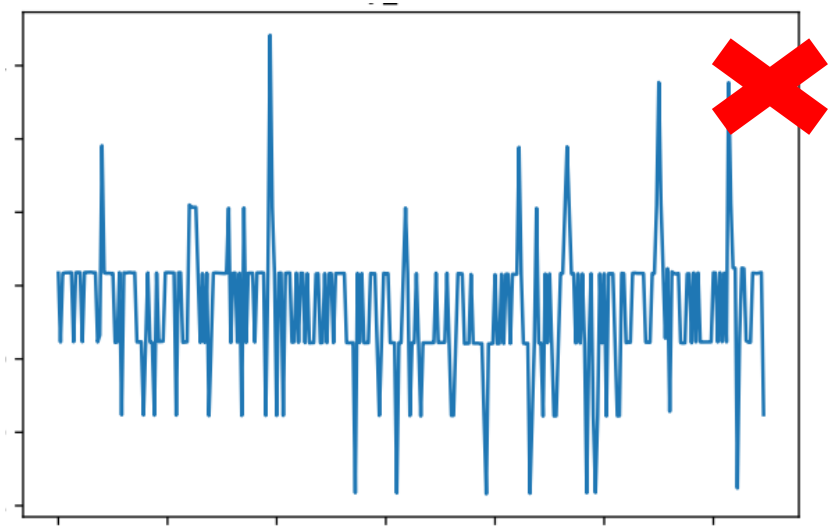
No one clear signal

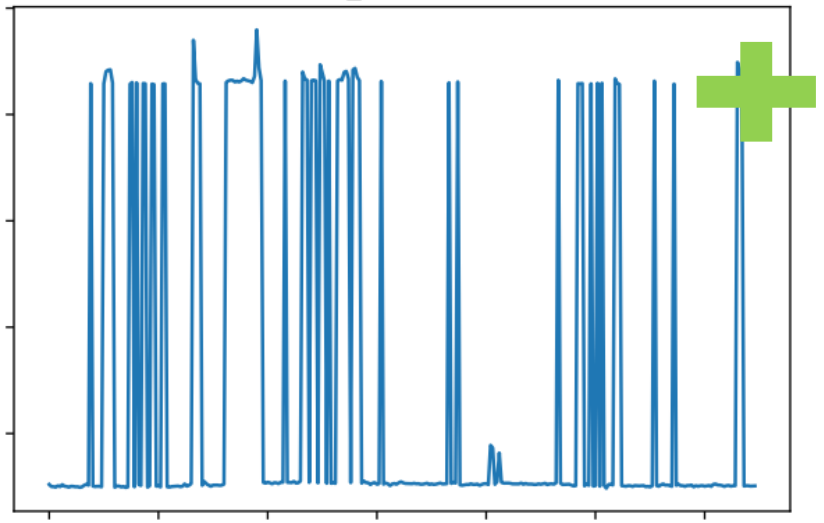
No one clear signal …

2 devices failed, 2 had no problem. Spot them!









Infrastructure for doing data science: current and proposed

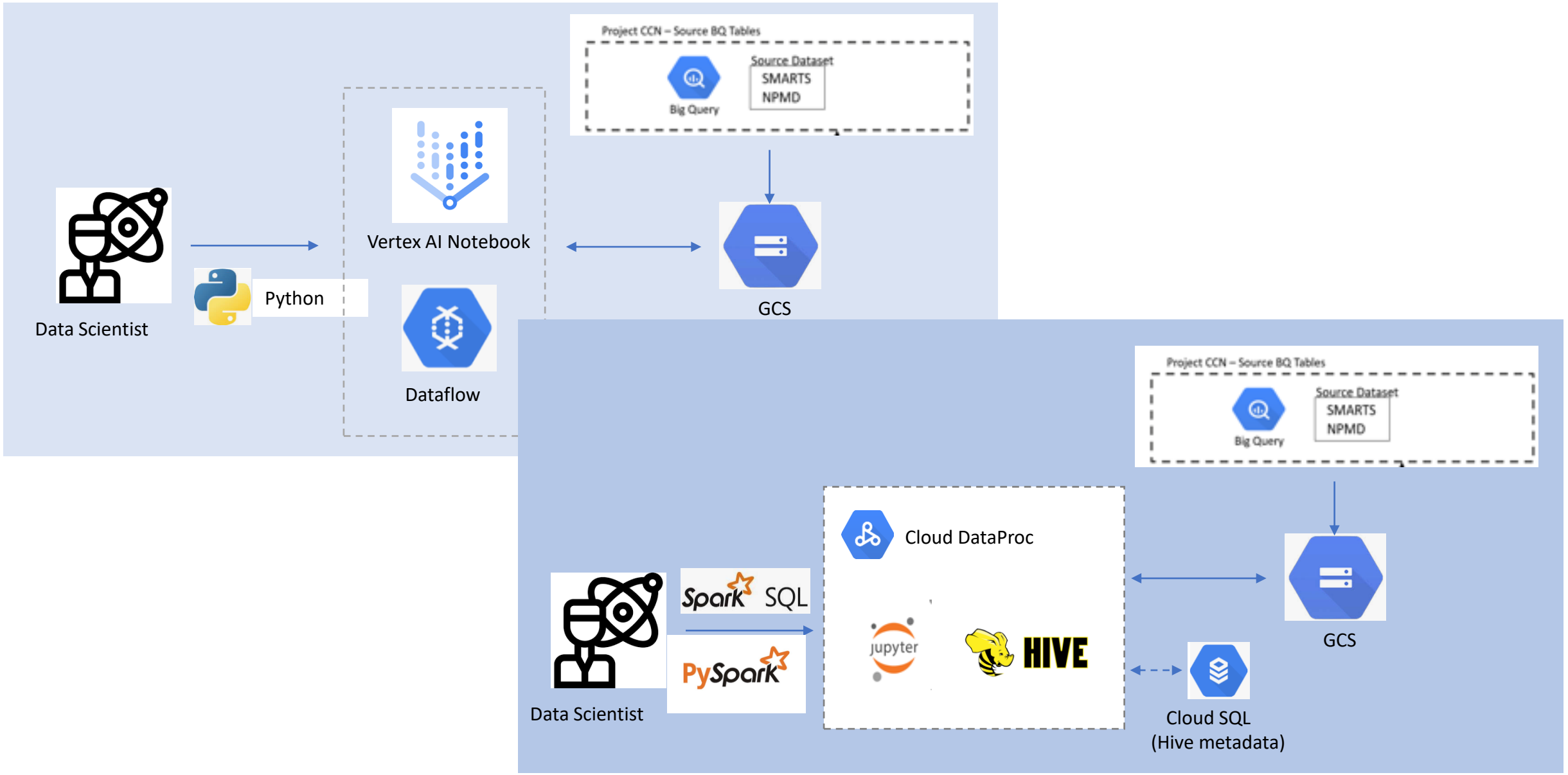


Image generated by Infosys experts providing data engineering support to the project

GCP options considered for data manipulation

Dataflow ? **Dataproc** ? **Bigquery with slots**

- "The fully-managed, serverless, auto-optimized data processor that simplifies development and management of stream and batch pipelines"
- ≈ Apache Beam in GCP environment
- "Auto graph-optimized for best execution path, Autoscaling mid-job, Dynamic Work Rebalancing mid-job"

- "Cloud Dataproc is a managed service for running Hadoop and Spark data processing workloads"
- ≈ Apache Spark in GCP environment
- "Managed YARN Cluster for Hadoop, Spark, Pig, Hive"
- "Deploy clusters in 90 seconds and pay per-second
- "Add or remove workers with zero downtime"
- "Simplified version management"

- "slot" = unit of computational capacity
- "A slot commitment is equivalent to buying a dedicated BigQuery cluster."

Price = P or P * 0.92 depending on flavor

Price = P * 1.09

Decisive factor: flexibility in modelling

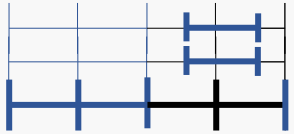
Decisive factor: the learning curve

Statistical modelling – the model fitting part

Classification

Input features → Tsfresh features → Sampling Preprocessing → Classifier → Anomaly score

Input features	Tsfresh features	Sampling Preprocessing
Selected set of raw variables	With Python tsfresh package	Oversampling / SMOTE Train / Test / Holdout



Anomaly detection

Input features → Data scaler → Anomaly detection algo → Post-processor → Anomaly score

Input features	Data scaler	Post-processor
Selected set of raw variables	Some cleaning and bringing to common scale.	Signal amplifier classifier model cross-validated

Device failure forecasting: the solution – how well is it working?

Business goals were defined in terms of precision and recall
During model building AUC is observed
Reporting here F1 ratios for some secrecy

	External team	Classification	Anomaly detection
F1 score	x	4 x	14 x

So … what's next?

	External team	
F1 score cust1	4 x	Reported recalls much higher than the best result above
F1 score cust2	11 x	

$$F1 = 2\frac{precision \times recall}{precision + recall}$$

Some extra slides here….

Appendix – data ingestion from source systems to GCP

Infrastructure – data ingestion -1

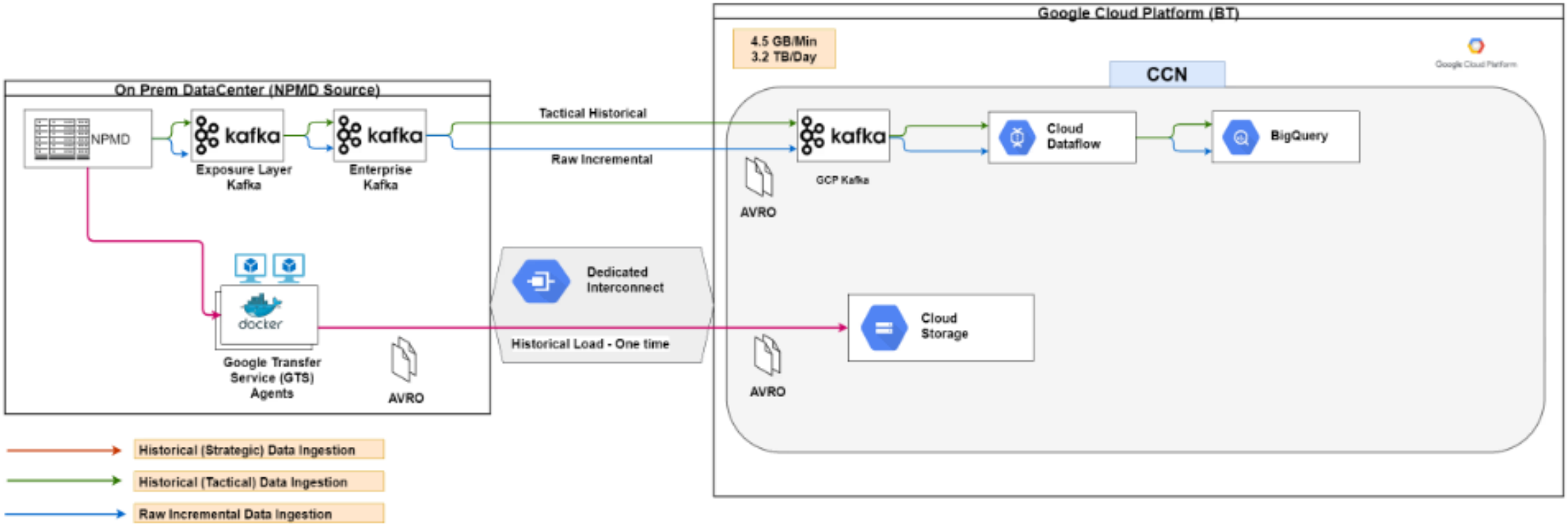
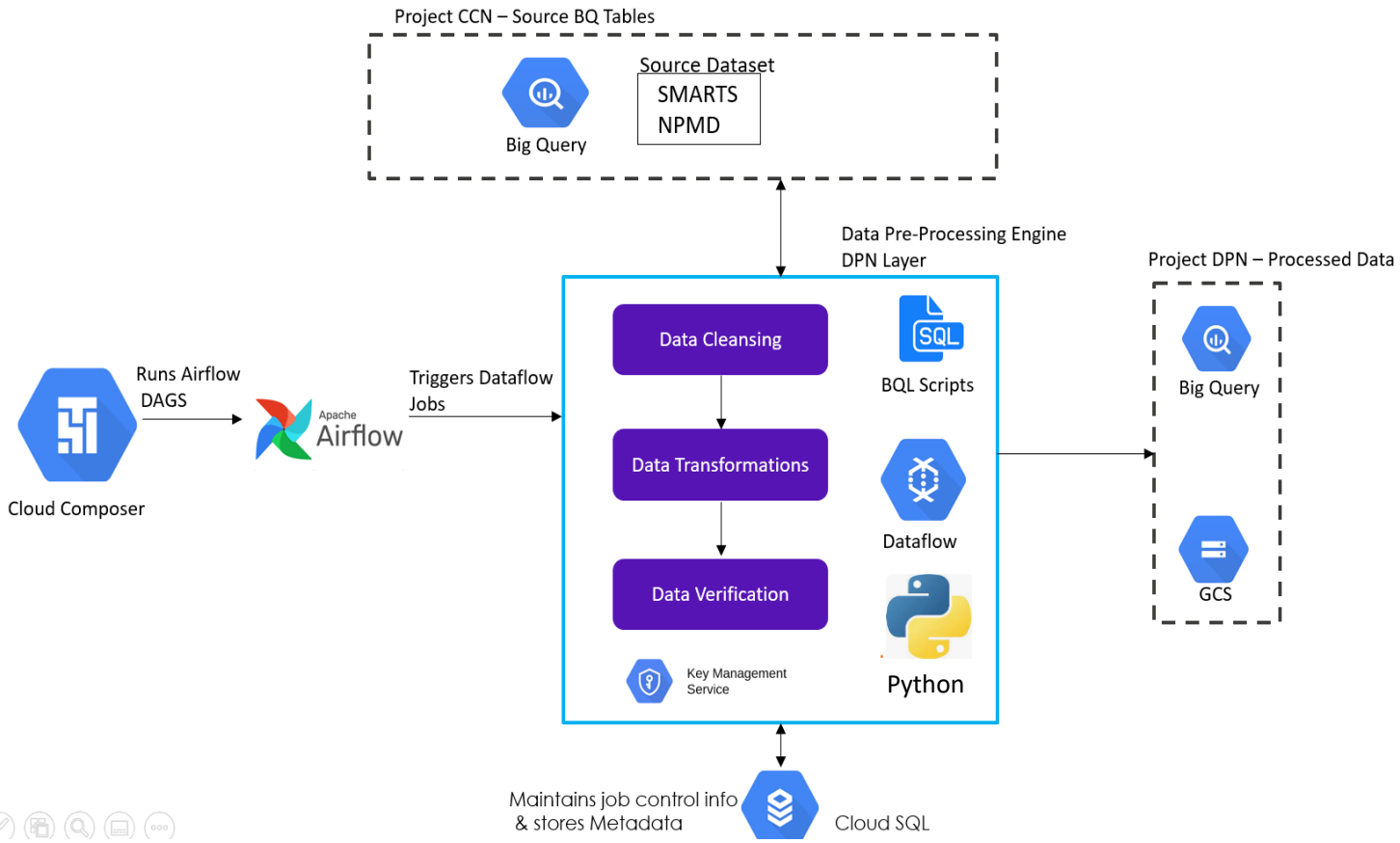


Image generated by Infosys experts providing data engineering support to the project

Infrastructure – data ingestion - 2



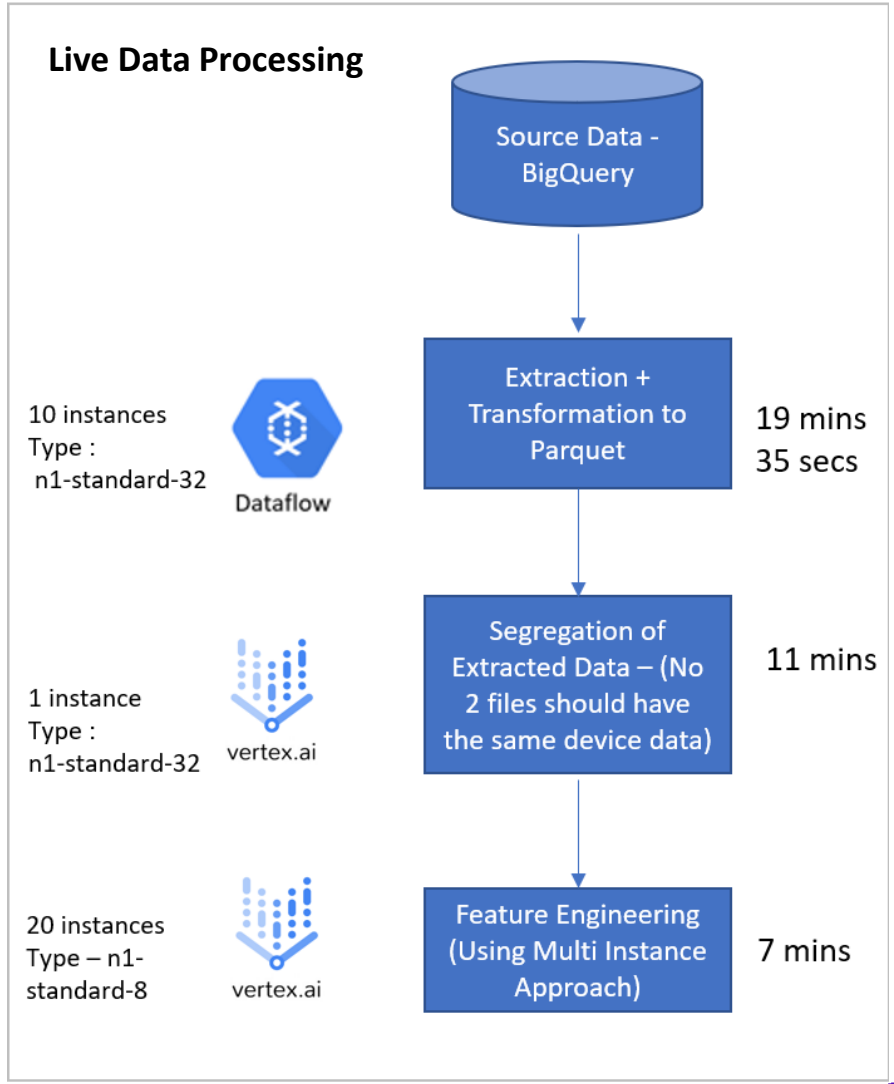


Image generated by Infosys experts providing data engineering support to the project

Appendix – sampling for training failure detection models

Statistical modelling – our working sample

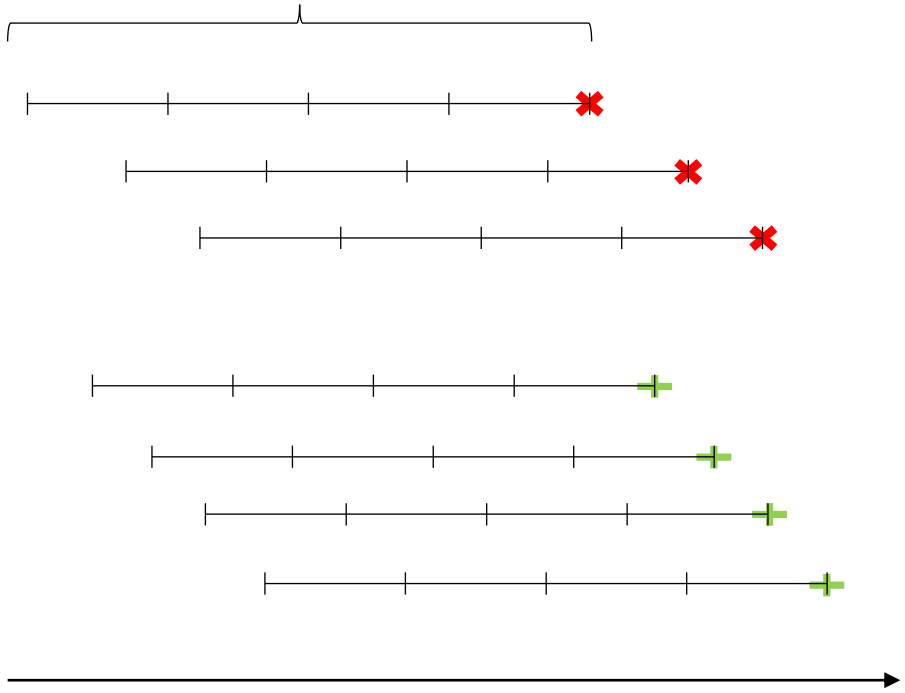
For sake of speed & due to tooling constraints we worked on a subset of device stories. 1/5 of all devices were included, among them all failed ones. The usage of sample made quick experimentation possible.

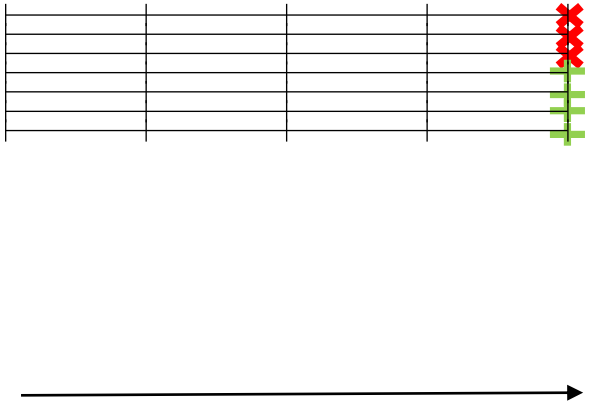
Original data - 4 weeks somewhen between 01/09/2022 and 31/01/2023

Our data – 4 week long stories

All device histories before a failure

Several other non-failure stories, but not all





Thank You