How To:
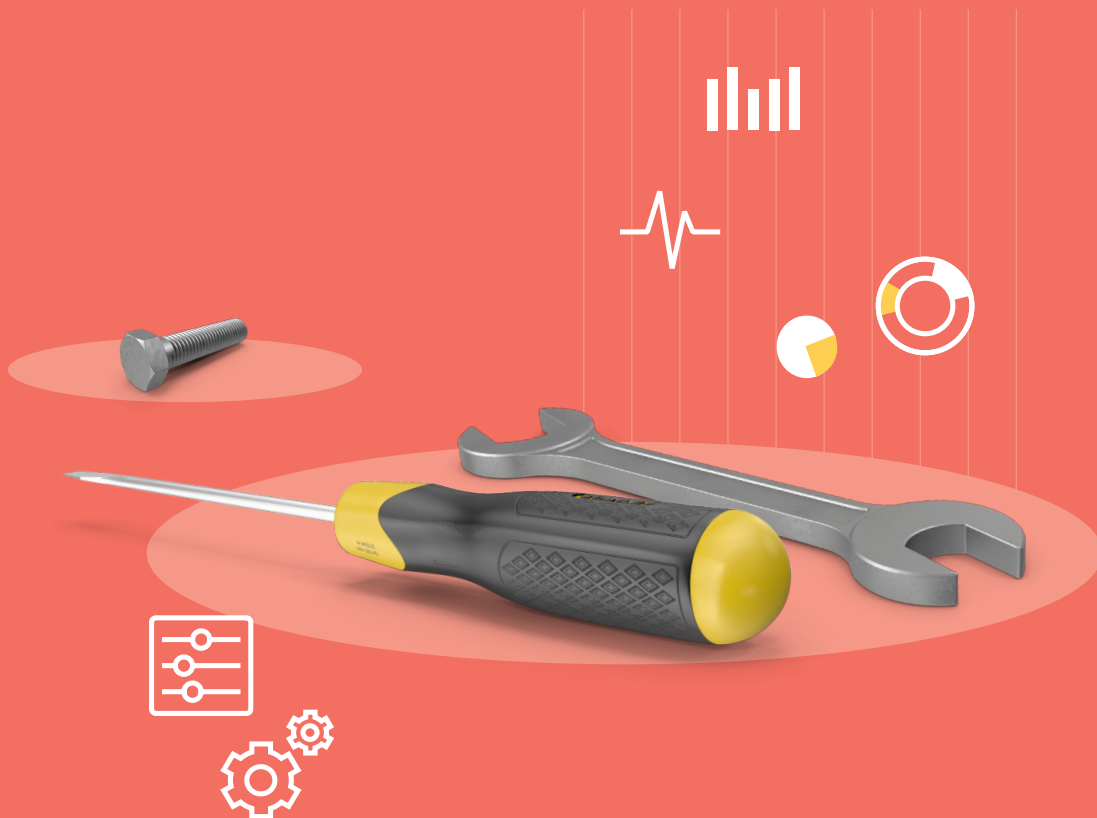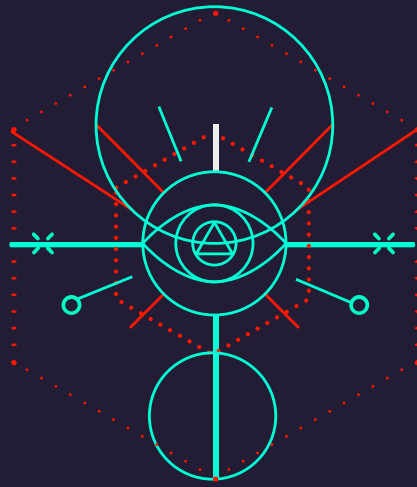# Future-Proof Your Operations With Predictive Maintenance

Leverage Real Time IoT Data to Anticipate
and Optimize Equipment Repair

data
iku

S U M M A R Y

ABOUT THE GUIDEBOOK

This guide is intended to provide a **high-level overview of the process involved in setting up a predictive maintenance program for any type of high-capital asset**, whether heavy machinery, fleets of any kind, manufacturing equipment, and more. By the end, readers should have an understanding of:

- How predictive maintenance differs from (and provides advantages over) data analysis techniques.

- Use cases and potential sources of data for those use cases.

- Specific challenges to exploring, cleaning, and modeling data for predictive maintenance use cases.

- Secondary analytics and how this technique can further optimize operations for a complete predictive maintenance strategy.

- Practical next steps for getting started.

For those completely unfamiliar with data science in the context of predictive maintenance, this guide will provide a short introduction to the topic and walk through the core aspects. But on top of that, for those already familiar, the guide includes a more advanced look at predictive maintenance elements and techniques for the most optimal strategy.
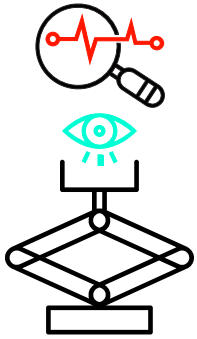
What?

Today, predictive maintenance is widely considered to be the obvious next step for any business with high-capital assets, harnessing machine learning to control rising equipment maintenance costs.

Predictive maintenance takes data from multiple and varied sources, combines it, and uses machine learning techniques to anticipate equipment failure before it happens.

Lots of businesses are already using continuous monitoring technologies - like Internet of Things (IoT) connected devices - which is a good start; but **the key lies in not just simply monitoring the output of various data (which is how many companies use it today), but in taking the next step and employing advanced algorithms and machine learning to take action from real-time insights**.

Going one step further, the most innovative enterprises (no matter what type of high-capital assets they maintain), see the largest cost savings from predictive maintenance not just by putting a system in place that returns simple predictive outputs but by rethinking and optimizing their entire maintenance strategy as a whole from top to bottom. This means:
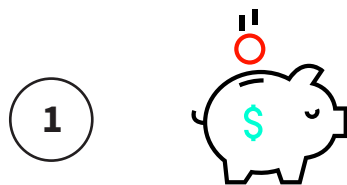
1 Considering a combination of maintenance strategies to determine the optimal cost-saving combination of predictive and traditional maintenance, perhaps even on an asset-by-asset basis.
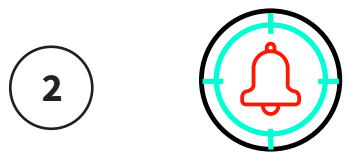
2 Paving the way for artificial intelligence (AI) and self maintenance by optimizing for (and automating) the immediate next steps once predictive systems point to imminent failure, whether this automatically triggers a work order, notifies a technician or certain team, places an order for a replacement part, etc.

3 Identifying how to best execute necessary repairs through second-order or secondary analytics. This means having a process in place for an entire deeper layer of analysis to determine the best time to actually remove the asset from service and which additional repairs - if any - should be conducted simultaneously to minimize the cost of having to remove the asset again for a different failure within a short window.
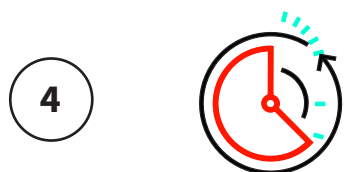
4 Determining if, through predictive maintenance, assets now have extra capacity due to decreased overall downtime and whether that time can be sold to other (generally smaller) businesses.
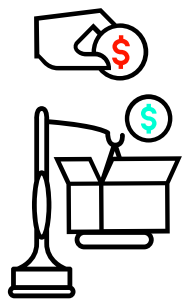
Why?

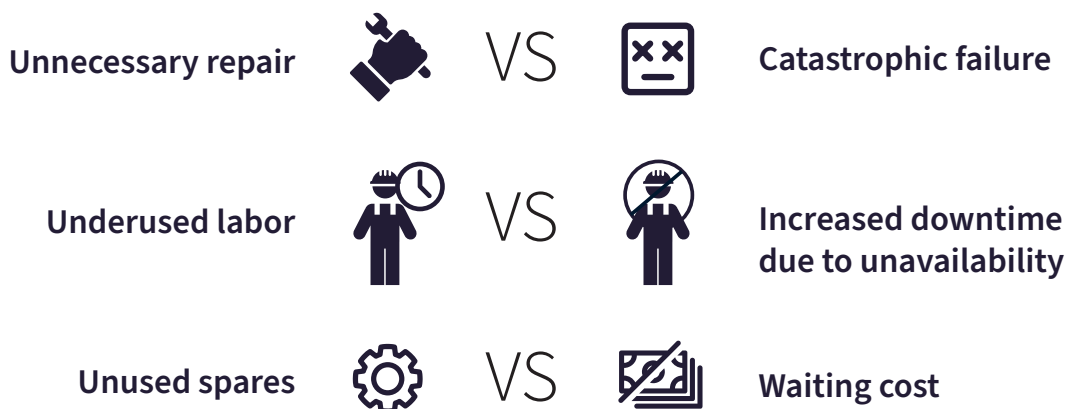Across all industry, balancing and benchmarking maintenance costs is a major challenge, and reducing those costs isn't a simple equation - managing maintenance operations comes down to weighing a series of cost-related tradeoffs:

Unnecessary repair	VS	Catastrophic failure
Underused labor	VS	Increased downtime due to unavailability
Unused spares	VS	Waiting cost

Most businesses have accepted the balancing of these tradeoffs as necessary costs of doing business and employ a variety of techniques (manual checks, scheduled maintenance, data from individual assets, etc.) to optimize manually as best as possible. But even combining these less-than-ideal options in the most logical way possible is manual and extraordinarily costly, with some industries reporting staggering maintenance costs of up to 40 percent (or more) of total costs.

Predictive maintenance often allows for the detection of impending failures that could never be detected by human eyes - take, for example, imaging that looks for microcracks in heavy machinery, even while in use. With predictive maintenance, downtime and repairs are directly tied to likely failure, minimizing cost (less downtime, less labor time, less chance of unexpected failure) and maximizing asset life:

Traditional Maintenance

Predictive Maintenance

- ● Start of life
- ⚪ Maintenance / Downtime
- ❗ Failure
- ❗ High Failure Probability
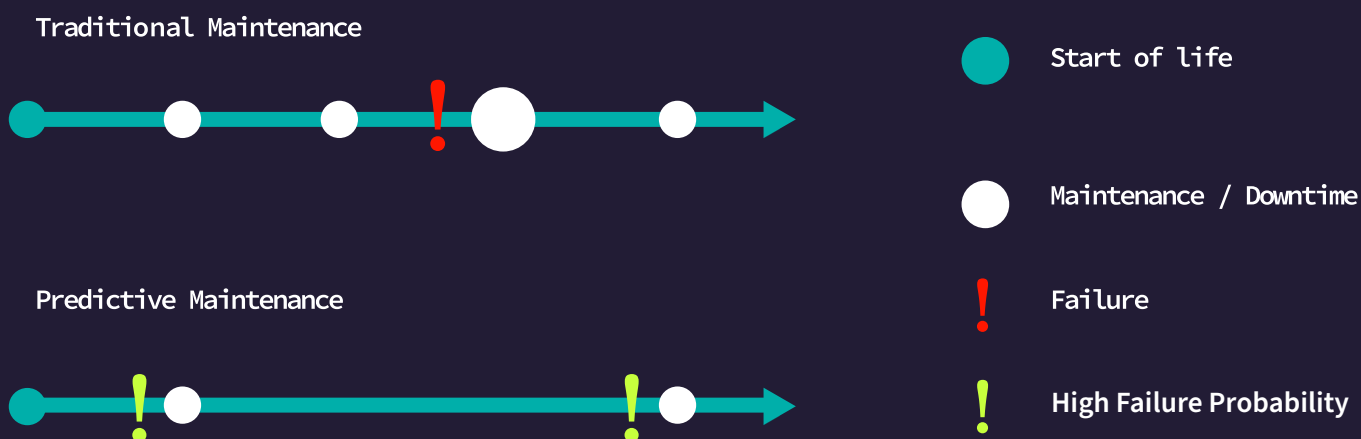
By contrast, traditional maintenance techniques (run-to-failure, preventative, or some combination of the two) inevitably mean unexpected repair, which leads to longer downtime on top of unnecessary downtime due to regular inspection. **Due to a lack of a complete picture surrounding the state of a given asset, traditional maintenance presents a slew of issue**s that contribute to both the breakdown of equipment and rising costs - the very things maintenance efforts are in place to prevent:

	Run-to-Failure	Preventative	Combination (Run-to-Failure + Preventative)	Predictive
Amount of Planned Downtime	$ $$$$ Low	$$$$ $ High, but potentially ineffective/ unnecessary (or technician could cause damage)	$$ $$$ Medium	$ $$$$ Only as-needed: Effective and targeted
Amount of Unplanned Downtime	$$$$$ Very High	$$ $$$ Very High	$$$ $$ High	$ $$$$ Minimal, but non-zero; models are never 100% accurate

Of course, it's not such a binary question as traditional versus or against predictive maintenance, as even with a predictive maintenance strategy, some traditional inspections may be performed - for example, after a storm or major event. But in general, the move toward predictive maintenance - even partially - has huge associated cost savings and paves the way toward the future, where AI plays a larger role in the industrial world.

How?
Understand the Need

The first step in moving toward predictive maintenance is to understand pain points (namely drivers of costs, waste, or inefficiency) and identify the best use case for your business. To date, there are use cases that span across multiple industries, but the primary differentiator between predictive maintenance and other data-driven strategies is, well - the predictive component. In other words, not just taking past data and doing static analysis, but **using real time data to predict future asset performance or issues at any given time and having a real-time feedback loop** in place to act on those predictions.
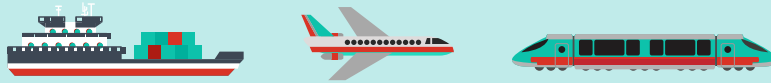
TRANSPORTATION

Maintaining a fleet, whether vehicles, barges, aircraft, or trains, presents additional challenges (not to mention costs and downtime) when repairs are required mid-voyage, so there is additional incentive to implement predictive maintenance strategies in this sector.

One large shipping company operating in more than 35 countries and with a fleet of more than 400 vessels uses a network of 4,000 sensors per boat taking readings three times every second. The sensors measure the temperature, exerted pressure, etc., on a variety of components. This data allows fleet managers to predict component failure and fix them before vessels face larger breakdown issues and, sometimes weeks before the failure occurs.

Other common examples being employed today include recommending the proper time for oil changes for large fleets of trucks, busses, or cars; preventing mechanical failures that result in delays for aircraft; and optimizing driving technique (e.g., speed) to reduce wear and extend fleet life.

MANUFACTURING AND INDUSTRY

(ENERGY, OIL/GAS, ETC.)

Manufacturing and industry are hands down the most obvious wins when it comes to predictive maintenance, and they are already ahead of the game, representing a large portion of the current use cases for this technology.

For example, a manufacturer and distributor of medical equipment worldwide with more than 25 production sites and 10 distribution centers from which they fulfill customer and provider orders was losing productivity due to premature wear on their machines. By ingesting data from a variety of sources (including IoT-connected sensors, machine logs, and manual quality measurements), the company is now able to predict the key variables leading to premature wear. From there, they use the real time data to optimize machine settings and minimize the key variables indicative of future performance issues and wear. Another use case is simple machine maintenance (vs. optimization) and is common not only in manufacturing of all products but also the energy industry.

CONSTRUCTION AND AGRICULTURE

These industries present some challenges when it comes to predictive maintenance, namely the higher cost of sensors due to the need for more rugged equipment and, often, lack of infrastructure (for example, WiFi for IoT devices).
Most commonly, predictive maintenance use cases in these industries align closely with fleet management use cases, predicting failure and optimizing uptime of heavy equipment. The largest difference is not in the way the data is used but in the hardware. The feasibility of predictive maintenance in construction and agriculture often hinges on demanding requirements for operating temperature, ingress protection, and vibration ranges of IoT devices.

REAL ESTATE

The future of real estate is IoT-enabled building management systems that measure everything from temperature to pressure and detection of chemicals or gasses. These systems would then be able to detect anomalies and allow for preventative maintenance by building management to fix things like gas or water leaks before they cause significant damage to other parts of the building or, worse, to tenants (whether residential or commercial).

With lots of high-capital equipment, namely freezers and refrigerators, food retailers can employ predictive maintenance techniques using IoT-connected devices to monitor energy consumption. By using energy level data, retailers can repair malfunctioning equipment and prevent even more costly complete failure of freezers and refrigerators in stores or warehouses.
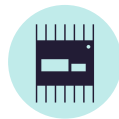
Get Data

Of course, the proliferation of IoT plays a large role in predictive maintenance, especially with cheap sensors and data storage combined with more powerful data processing that has made the technology accessible. So a portion of the data feeding predictive models comes from IoT devices monitoring all aspects of an asset - anything from temperature to vibration, pressure and fluid levels, even sound. **Some real time data from IoT sensors is the bare minimum and an essential component to true predictive maintenance because unlike the other types of data listed below, it directly monitors asset conditions and allows for up-to-the-minute predictions.**

But data from IoT devices on the assets themselves is not the only source. The advantage of predictive maintenance is the ability to combine data from a large variety of sources for the most accurate predictions. Other data sources might include:

Data from programmable controllers	Manufacturing execution systems	Building management systems
Manual data from human inspection	Static data, like manufacturer service recommendations for each asset	External data from APIs, like weather
Geographical data	Equipment usage history data	Parts composition

Essentially, any source of data on an asset can augment IoT data and be used to build and test a predictive maintenance algorithm. From here, you can determine which data sources are the best indicators of failure, wear, or breakdown and add or remove features to refine the final model.

GO FURTHER

When selecting actual training and testing data sets for predictive maintenance, note that you may or may not include data where the assets are allowed to run until failure. If you already have data from where assets are allowed to run until failure, you'll later be able to compute a "time before failure" window.

But if you're just getting started implementing a network of IoT devices and therefore don't have this data, you'll need to employ survival analysis to build a model without first having run-to-failure data.

The primary challenge is that data from different sensors can vary immensely. This means not only does each one measure a different aspect of the environment or of the present conditions, but they can have completely different levels of noise (i.e., data from some sensors barely changes over time while others are extremely noisy) and, on top of that, levels of quality. Because of the frequency of measurements, there will likely be lots (hundreds or even thousands) of values that are uninformative. So a large portion of this phase will be figuring out how to cut through this noise by handling missing data and removing redundancies in the test/training sets (and, of course, automatically going forward).
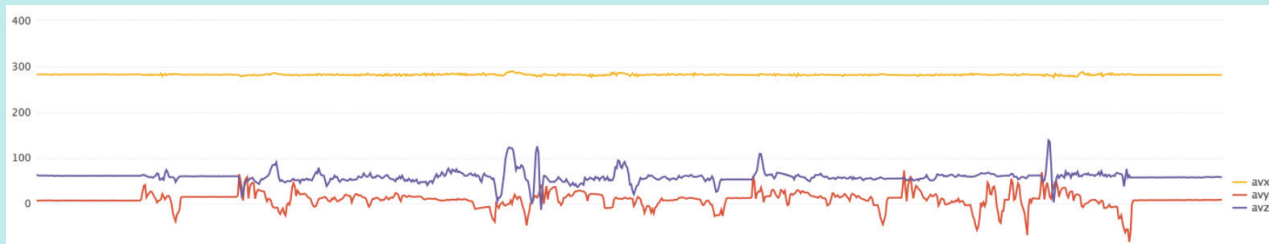
During this stage of data exploration, if you know when each asset in the training set will fail, it's possible to compute a "time before failure" (elapsed life at that time minus its total lifetime). This "countdown to failure" will allow you to align different engines' data to a common end point, which will simplify analysis. If you don't know when an asset will fail or don't have failure information available, it will be up to you to first detect failure based on available signals without a countdown to failure/time before failure available.

It's also important to note at this stage that it's likely the assets in your data sets have each started with a different (unknown) level of wear and that they wear differently over time, so it may not be possible to generalize and predict failure even across otherwise identical assets.

Enrich Data

Manipulating data at this stage means adding more features and joining it in meaningful ways so that each data set, or data from multiple sensors, can be taken as a whole instead of in parts.

GO FURTHER

Multiple sensors can measure various aspects or conditions of the exact same asset, so it might be helpful to join data for the same asset to get a complete picture of one particular piece of equipment. Or combine data from the same sensors on different assets to compare how certain sensors perform regardless of the specific asset. Aligning the data in several different ways and leveraging visualization tools during your analysis will allow you to uncover more patterns about how sensors behave immediately preceding or leading up to a failure.

Enriching IoT data with static data at this point can also provide valuable clues. For example, joining data from manual inspection and repair with sensor data will provide a more complete picture of an asset's life cycle and how performance was impacted by human intervention.

Get Predictive

It is precisely this combination of a variety of sources and data types that allows for the most robust and accurate predictive models. **The more sources and types of data available, generally the better the complete picture of a particular asset** and the better the prediction.

Once the data is combined, predictive algorithms, analytics, and machine learning can be applied to produce an accurate model that successfully predicts failures, and the model is tested to ensure an acceptable degree of accuracy that minimizes false negatives and false positives.

In the exploration phase, you might have been able to use a "countdown to failure" in your analysis since you knew precisely when the assets in the training set would fail. But when creating models, the actual moment of failure is unknown, so the challenge is to predict the amount of time the engine will continue to function before it fails - known as the Remaining Useful Life (RUL). Correctly predicting an asset's RUL is the crux of predictive maintenance.

A consideration specific to predictive maintenance is the cost of incorrect predictions. It would be ideal to have a model that is 100 percent accurate, but of course, this will never happen. So with predictive maintenance, it's important to consider the costs of these failures and whether it's more optimal to over- or under-estimate the RUL, thus leading to either catastrophic failure or unnecessary maintenance, respectively.

Ultimately, this is a business decision that needs to be made with and discussed as a team before deploying a model with such consequences into production.

Visualization

Visualization is an important tool in predictive maintenance as it often closes the feedback loop, **allowing maintenance managers and staff to see the outputs of predictive models and direct their attention accordingly**. Robust data science or data team tools today allow maintenance managers and staff on-the-ground to easily access and digest outputs in a familiar format so that the entire team - from analysts to technicians - receive the same feedback.

GO FURTHER

On top of visualizations delivered as a final asset to maintenance managers, visualization can be helpful also at previous stages, especially data exploration. The massive amount of data from sensors is often best explored in a visual format, especially when combining data from multiple sensors on the same asset for a complete, visual picture of possible (or actual) failure.

Iterate and Deploy

Deploying a predictive maintenance model into production means working with real time data, but one step further than simply providing visual real time dashboards for on-the-ground maintenance teams. For some use cases, **feedback can be integrated directly into the predictive maintenance process**, requiring no (or little) human interaction.

GO FURTHER

Businesses using an API-driven solution for predictive analytics can do anything from instruct machines to directly make changes or adjustments based on data, automatically order parts or check if there are spare parts already available, notify a team or technician, etc.

At this stage, especially if automation is in place, it's critical to ensure the model's viability in production and to continue to monitor its performance over time.

What's Next
Secondary Analytics

Once it's clear repair is necessary and initial - perhaps automated - first steps or processes have been kicked off, that's where secondary analytics come in. Because taking high-capital assets out of service can be extremely costly in and of itself (even when compared to the benefits of identifying necessary maintenance before run to failure), the next questions are: when and how?

Take, for example, a truck from a large fleet with a part identified by your predictive maintenance system as being N days away from failure. Once identified, a member of the data team should be ready to send a secondary follow-up report to the maintenance team detailing:

Estimated time out of service for assets with similar maintenance issues.
How long is the asset expected to be out of service?

The likelihood that the asset's part fails before or after N days.
Realistically, how much time does the maintenance team have to consider and execute the maintenance plan?

Other maintenance tasks that make financial sense to tackle simultaneously while the asset is already out of service.
Are there other tasks that should be completed at the same time to prevent the aseet from being taken out of service again soon?

Optimal logistics for completing the maintenance.
If currently in use, where or when is the ideal place/time to take the asset out of service for repair based on availability of parts and labor?

Ultimately, the goal of secondary analytics following predictive maintenance is to determine a plan of action for exactly when the asset should be taken out of service so as to minimize disruption and loss (both imminent and future) and maximize resources.

Self-Maintenance

Predictive maintenance especially lends itself to the future of artificial intelligence (AI), where operations will be entirely self-maintenance with very little human interaction whatsoever. AI in the predictive maintenance space would go one step beyond the steps discussed above, which would still require some manual analysis of models and outputs.

These systems will watch thousands of variables and apply deep learning to find information that could otherwise be undetected that might lead to failure. Ultimately, predictive maintenance isn't so far off from AI, and businesses that get started with predictive maintenance programs now will be well poised as market leaders in the future.

More Opportunity

Of course, the biggest initial win with predictive maintenance initiatives is cost savings. But after implementing a larger, more robust, and more mature predictive maintenance strategy, larger opportunities begin to open from a business perspective, and **high-value assets can bring in some additional revenue instead of just being pure costs**. For example, decreased down time could free up assets for additional loads or allow for additional capacity. If that capacity isn't required for the business, in turn, that asset time can be sold to other smaller businesses for added revenue.
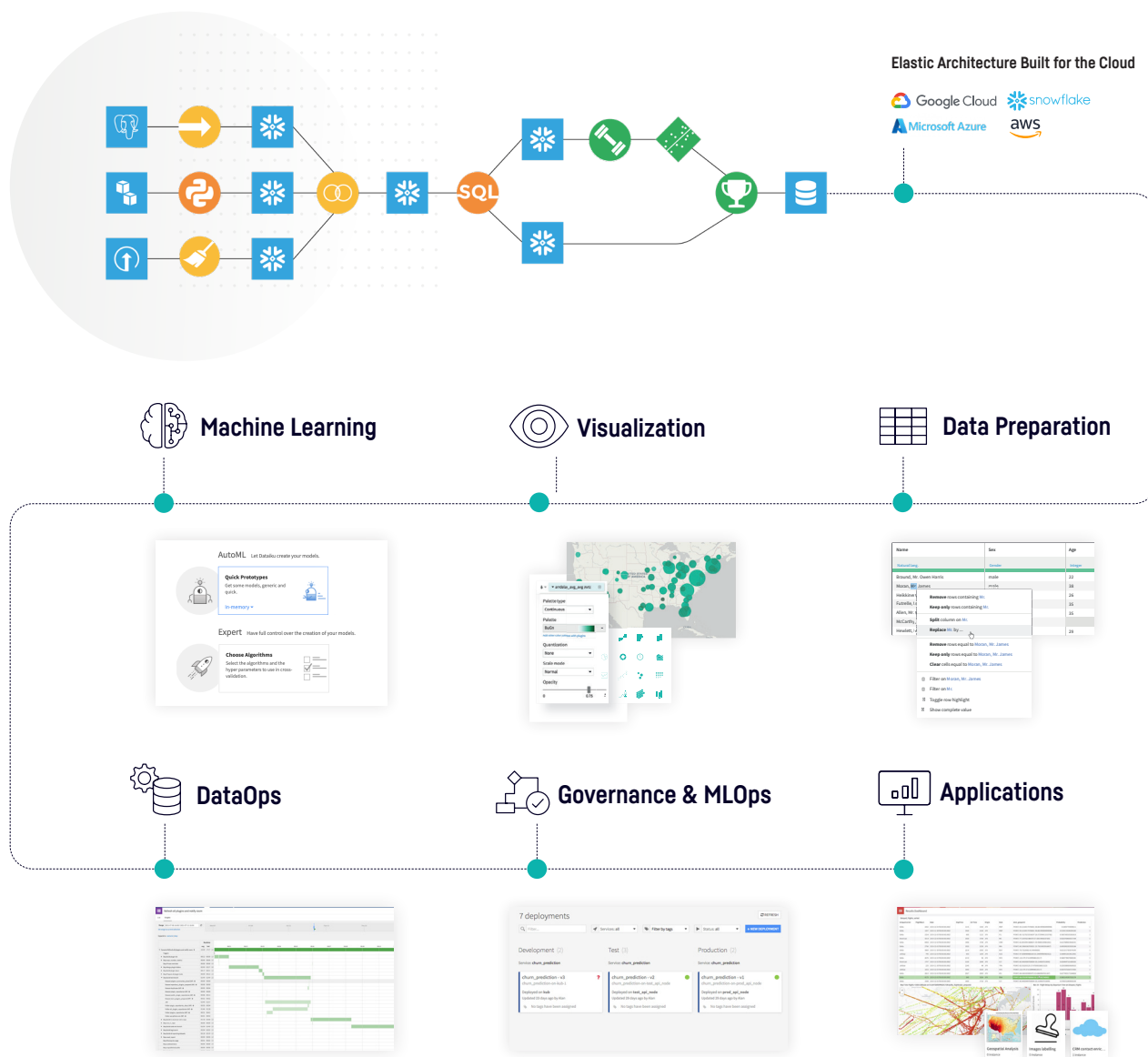
CONCLUSION

Given the resources spent on maintenance across industry today, **predictive maintenance is a critical step in reducing costs, increasing productivity, and - ultimately - staying relevant in an era where cutting-edge, data-to-the-core businesses are gaining market share**. Though the initial cost of implementing a predictive maintenance strategy may be steep, companies today are already proving that the returns in increased productivity plus reduced costs are a valuable return.

So How to Begin?

- [ ] Conduct a cost/benefit analysis to determine the savings predictive maintenance might provide for your assets.

- [ ] Determine if a predictive maintenance strategy makes sense for all assets or only a portion of them based on cost analysis.

- [ ] Conduct analysis to determine the optimal combination of predictive vs. preventative maintenance per asset.

- [ ] Assemble a cross-functional team (maintenance management, data science, business analysts, IT, etc.) to prototype and lead the project.

- [ ] If automatic equipment monitoring isn't already in place, target specific assets on which to get started based on historical maintenance costs.

- [ ] Allocate funding for predictive maintenance systems, including IoT-connected devices, data storage, and a data science platform for ETL, modeling, visualization, and deploying to production.

- [ ] After prototyping, present analysis and get buy-in from management to roll out the program into production.

- [ ] With a successful proof-of-concept, roll out predictive maintenance to additional assets.

- [ ] If your business already has a predictive maintenance program in place but wants to go further, start moving toward a second-order or secondary analytics program to determine the most cost-effective maintenance plan when taking assets out of service for predicted failure. Additionally, start identifying opportunities for further AI automation with the goal of getting closer to self-maintenance.

Everyday AI,
Extraordinary People

Elastic Architecture Built for the Cloud

Google Cloud ❄ snowflake Microsoft Azure aws

Machine Learning

Visualization

Data Preparation

DataOps

Governance & MLOps

Applications

45,000+
ACTIVE USERS

450+
CUSTOMERS

Dataiku is the platform for Everyday AI, systemizing the use of data for exceptional business results. Organizations that use Dataiku elevate their people (whether technical and working in code or on the business side and low- or no-code) to extraordinary, arming them with the ability to make better day-to-day decisions with data.

GUIDEBOOK