Business Value of Advanced Analytics and Current Trends

Dr. Gül Ege

Senior Director, Advanced Analytics R&D

SAS Institute

§sas

Key Points

- What is business analytics?
- New era for business analytics
  - New challenges
  - New opportunities
- Success factors for the practitioners and executives

§sas

Business Analytics

The use of statistical analysis, data mining, forecasting, machine learning and optimization to make critical business decisions based on customer and operational data.

§sas

Technology: SAS® ADVANCED ANALYTICS

FORECASTING / ECONOMETRICS

Leveraging historical time series data to drive better insight into decision-making
for the future

MACHINE LEARNING/ DATA MINING

Understand and find relationships in data to make accurate predictions about the future
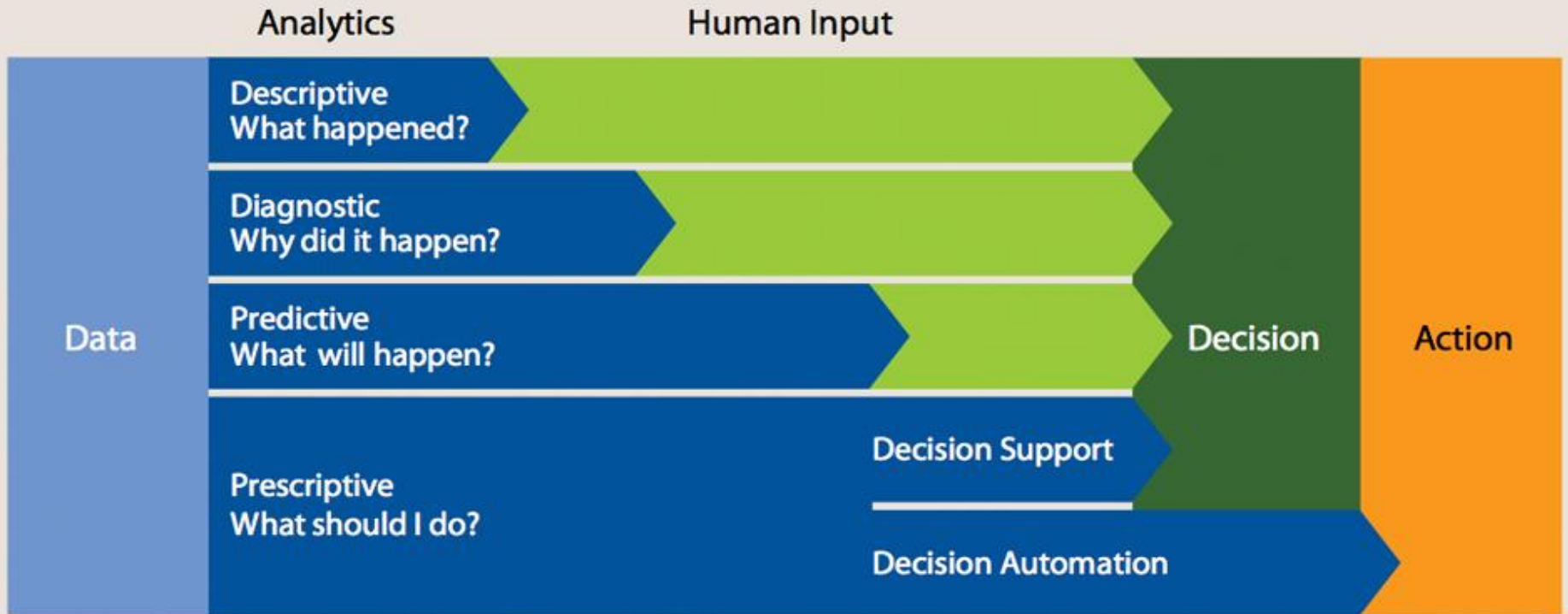
OPTIMIZATION

Make appropriate business decisions by understanding dynamics and utilize resources the best way

TEXT ANALYTICS

Discover relevant themes and relationships in social media, call notes and email for deeper insights and improved business management
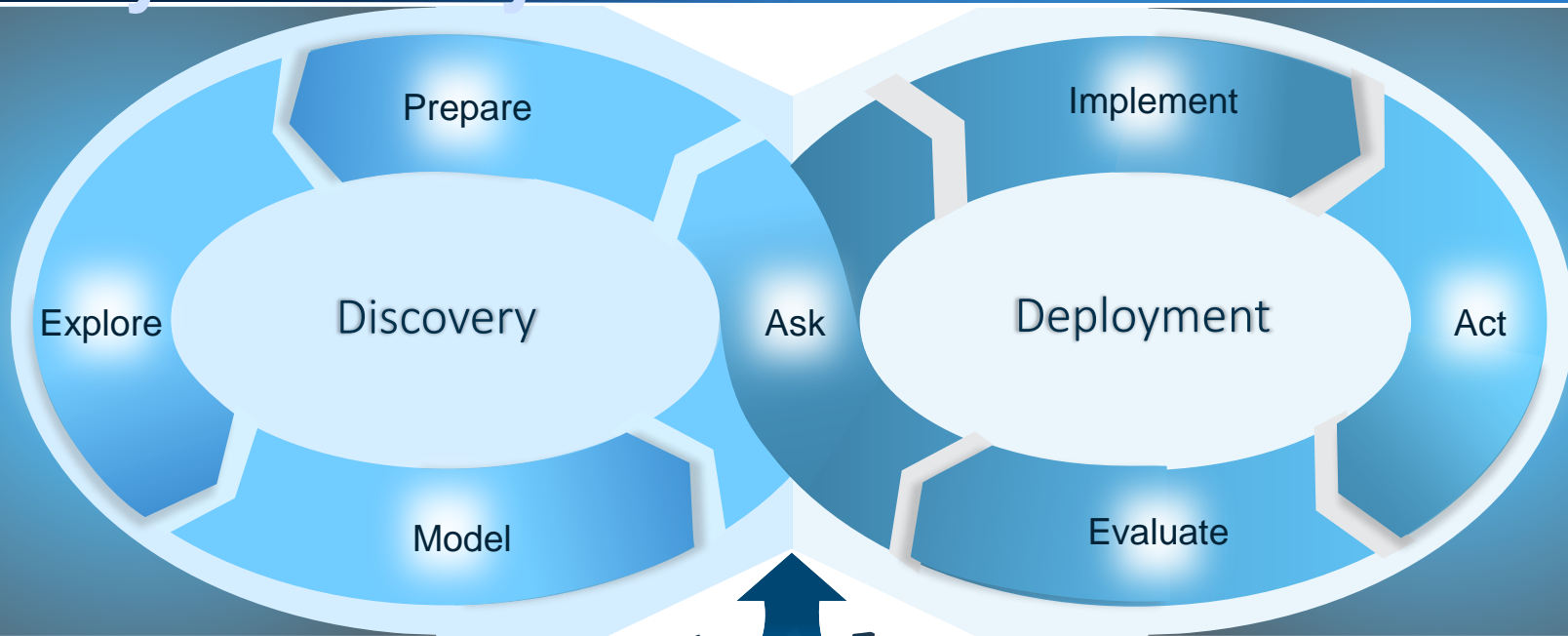
DISOVERY

BUSINESS INTELLIGENCE

DATA MANAGEMENT

VISUALIZATION

REPORTING

FOUNDATION ANALYTICS

Gartner: Analytics Continium

Analytics | Human Input

Descriptive
What happened?

Diagnostic
Why did it happen?

Predictive
What will happen?

Prescriptive
What should I do?

Data

Decision Support
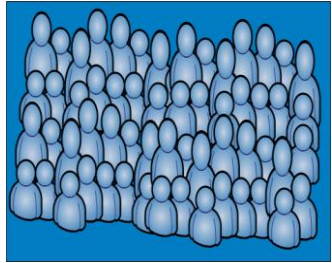
Decision Automation

Decision

Action

Source: Gartner, #G00254653 (September 2013)

§sas

Analytics Lifecycle
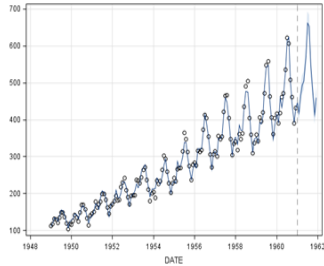
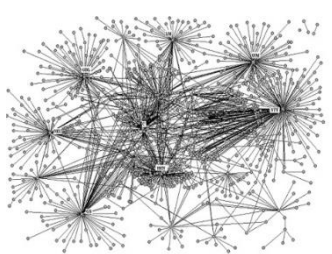# Data Galore

Cross-Sectional

Time Series
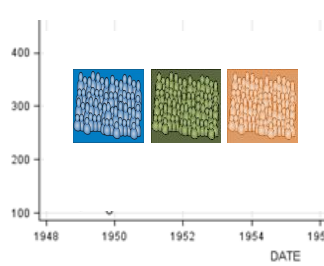
Panel
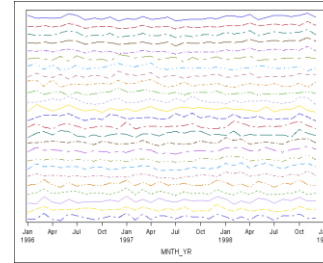
Streaming

Spatial

Network

Link

Text

Sound

Image/Video

Business Questions

Who will respond
to a campaign?

What will future
demand look like?

Will regulations
have an impact?

Are there
anomalies?

Where should we
look for oil?

Who are
influencers?

Which item should I
recommend?

How is our product
perceived?

Emerging problem in
our product?

Can we prevent
insurance claims?

§sas

Trends

- Big data/streaming data

- Complex problems requiring innovation in algorithms

- Unstructured data – call center logs, warranty information, social media

- Real time analysis: sensors and analytics behind them – aircraft, turbines, terrorist alerts, outbreak of disease, health wearables

- Multi-disciplinary teams

- Democratization of analytics

§sas

MITSloan Management Review: 2017

- 2602 managers, executives and data professionals around the globe

- Four key findings:
  - More companies report competitive advantage from their use of data and analytics
  - Innovation from analytics is surging: processes, products/services, business models
  - Data governance fosters innovation
  - Smart machines create opportunity for innovative thinking

S. Ransbotham, D. Kiron, "Analytics as a Source of Business Innovation," MIT Sloan Management Review, February 2017.

§sas

BIG DATA

2009 — 0.8 ZETTABYTES

2020 — 35.2 ZETTABYTES 44x

| Value	Metric
1000	Kilobyte
1000^2	Megabyte
1000^3	Gigabyte
1000^4	Terabyte
1000^5	Petabyte
1000^6	Exabyte
1000^7	Zettabyte
1000^8	Yottabyte

§sas

Innovations

- Technology innovations
  - » Exploiting the computing power
  - » Streaming data processing
- Analytical innovations:
  - » New Algorithms
  - » Automation – exploiting parallelism
  - » Computationally 'intractable' techniques
- Multi disciplinary teams

§sas

THE INTERNET OF THINGS

www.comsoc.org/blog

§sas

IOT Economic Impact: $4-11 Trillion by 2025
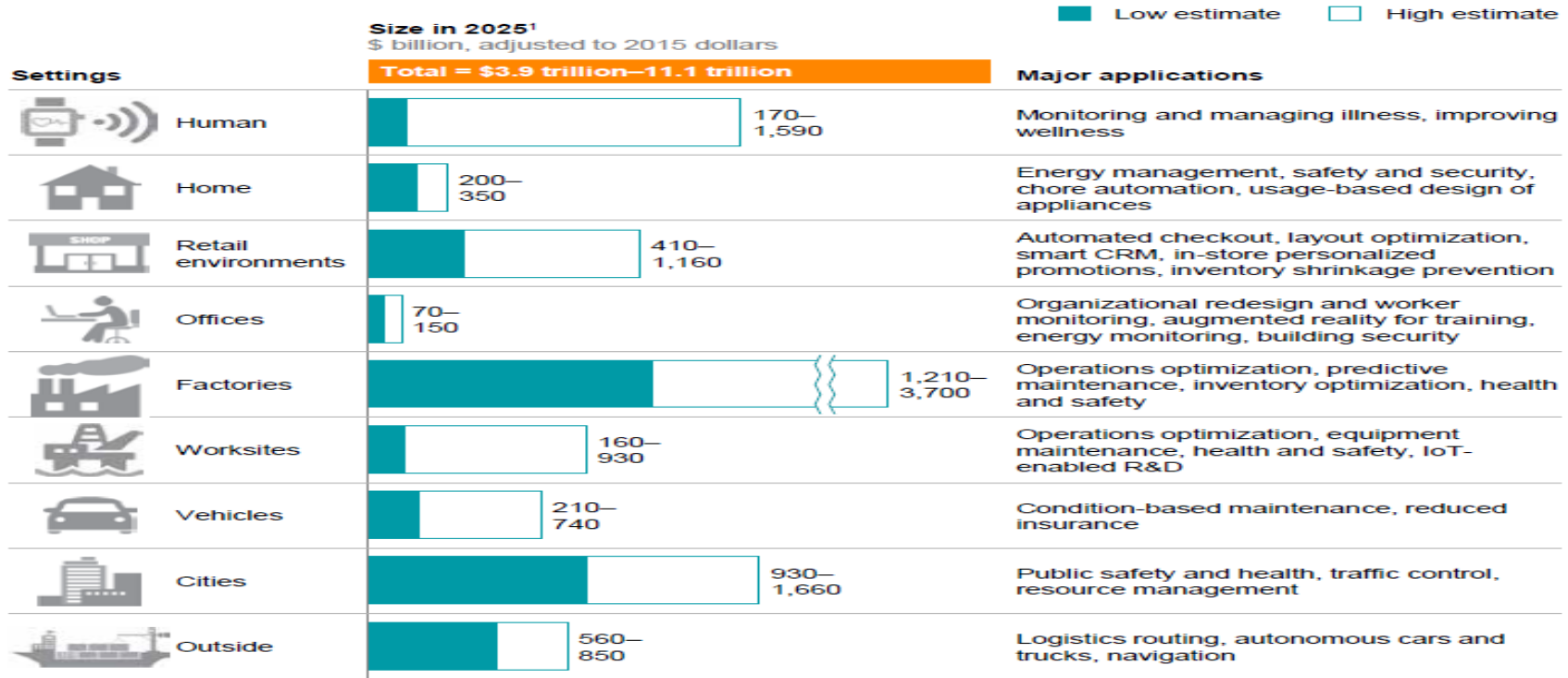
Analytics of Things (AoT)

- Making better business decisions more rapidly in a highly connected world

- It is the new frontier in scale, in speed, in predictive modeling and automated data-driven decisions

§sas

Common Challenges with IoT Data

- Lots and lots of variables

- High frequency, high volume, high speed

- Noisy data & missing values

- Unexplored to date

- Largely reflective of normal operations

LOOKING FOR PREDICTIVE POWER IN A HUGE HAYSTACK

§sas

Common Business and Analytical Themes

- Early warning of degrading systems/human health
- Predictive maintenance of high value assets
- Improved manufacturing quality
- Fraud and Cybersecurity

- Dimension & data reduction
- Transformations
- Anomaly detection
- Pattern recognition

§sas

Streaming Analytics Ecosystem

Edge Analytics

Network Systems, Surveillance

Monitor equipment on the platform for failures and safety issues, and take action.

In-Motion Analytics

Transactions, Logs, Clickstreams

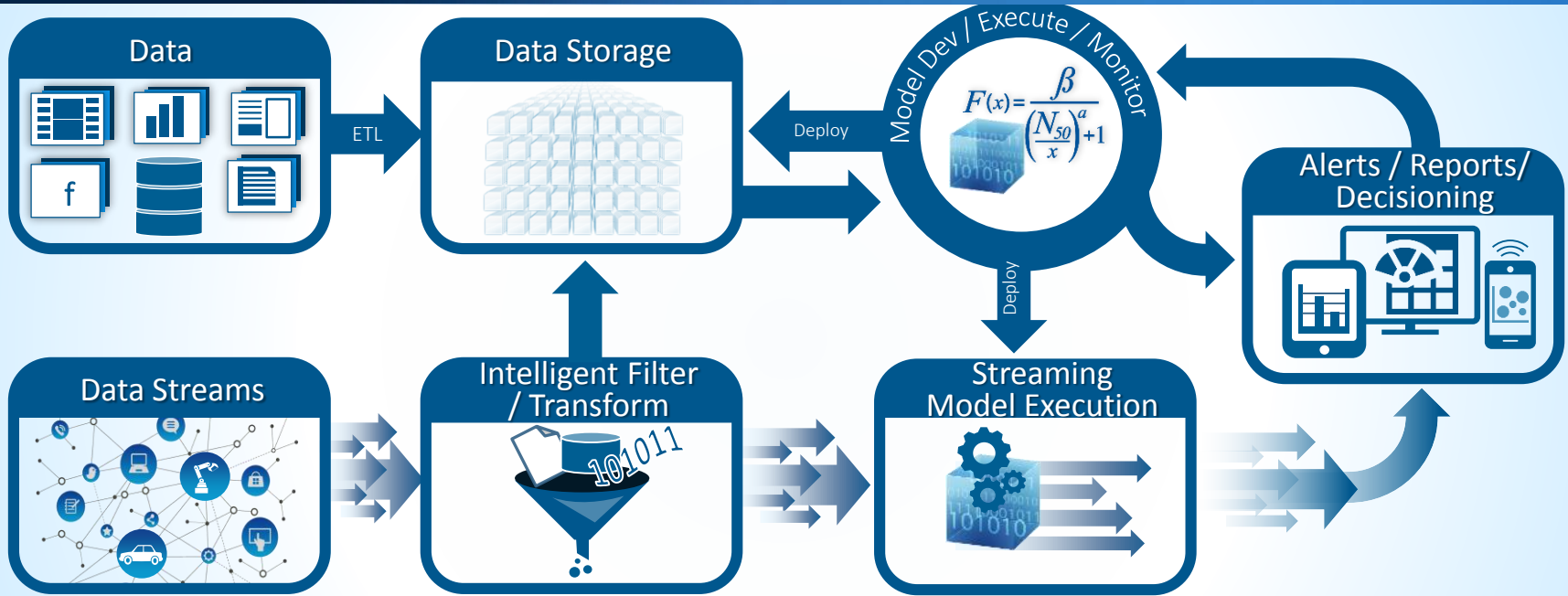Identify fraudulent transactions and be alerted in real-time.

At-Rest Analytics

Strategic Data Integration

Intelligently integrate data & develop predictive models

§sas

The IoT Analytical Lifecycle

Stream it, Filter it, Score it, Store it

Machine Learning in the Era of AoT

Machine learning is a branch of artificial intelligence that **automates** the building of systems that learn from data, identify patterns, and predict future results – with **minimal human intervention.** It shares many approaches with statistical modeling, data mining, information retrieval and other related fields.

Applications:

- Automotive -- driverless cars and automatic emergency response systems

- Banking – big data sources for marketing, fraud detections

- Government – pattern recognition

- Manufacturing – sensor data

- Retail -- micro-segmentation

§sas

Determining Electrical Grid Disturbances

- Using Phasor Measurement Unit readings scattered over a grid to determine where a disturbance started and how it spreads through the grid

- Measuring the disturbance by monitoring the voltage volatility by estimated range ratio transform

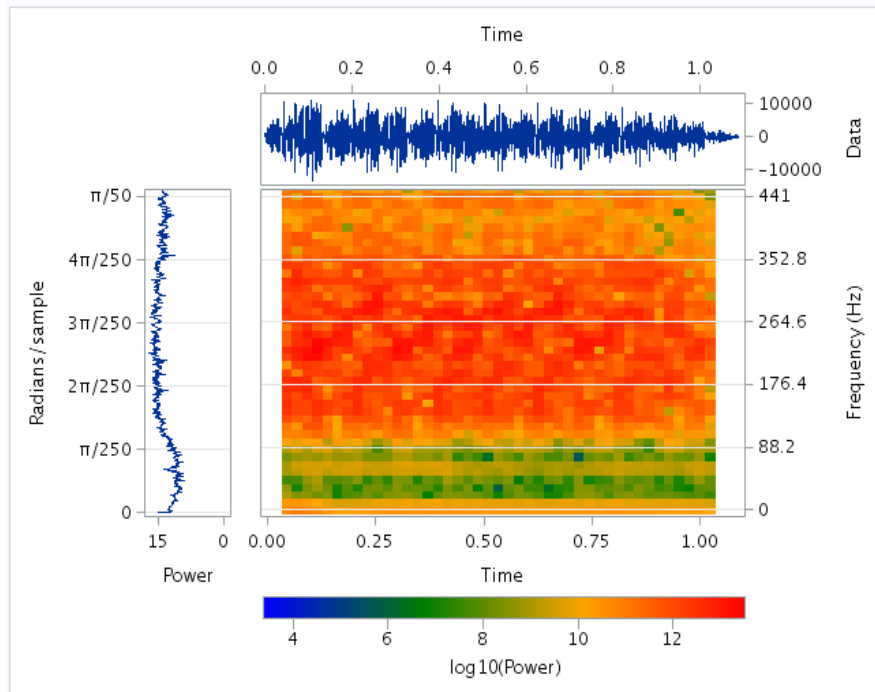- Real time monitoring of disturbance in the grid
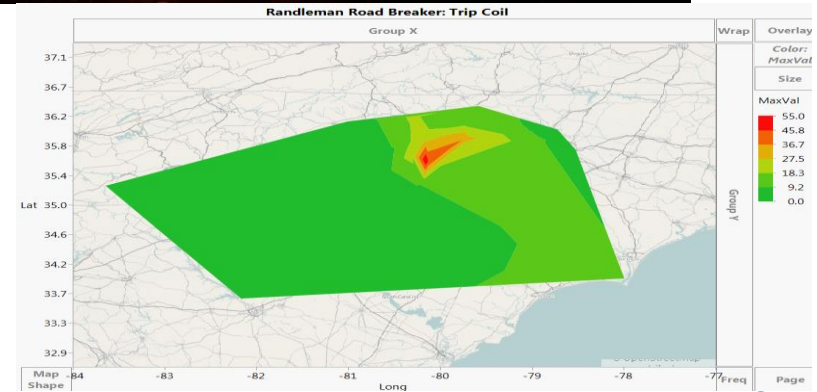
Detecting heart murmur
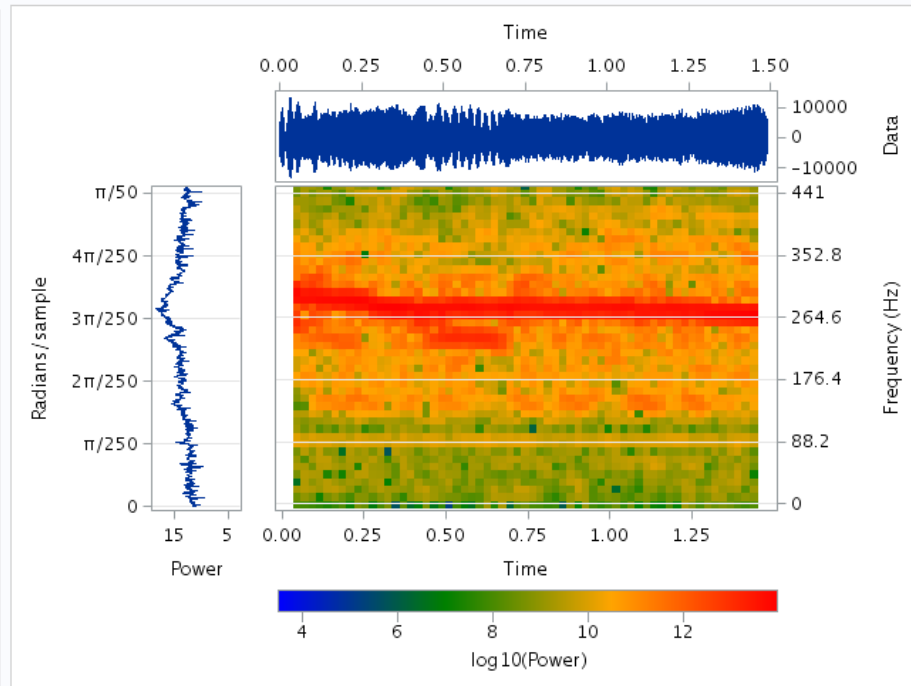
Short Time Fourier Transformations (STFT) on audio data
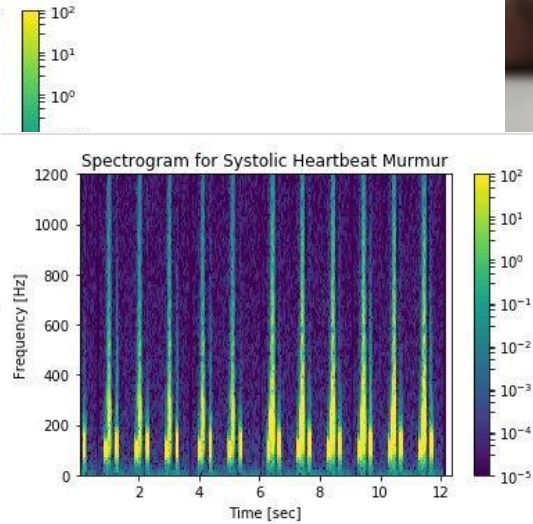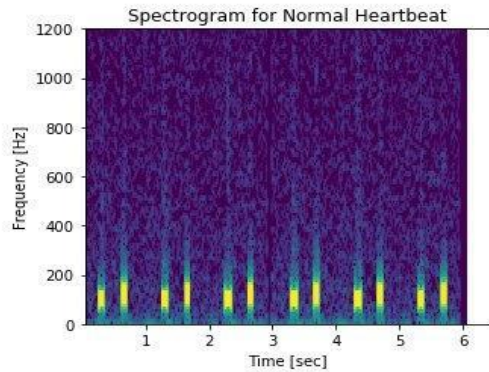
Signatures of healthy breathing vs asthma

Predicting Solar Farm Output

- 77 features output on an hourly basis

- Robust Principle Component Analysis (RPCA)

- Reduce data dimensionality & size

- Create more accurate models

- Identify anomalies

§sas

RPCA in Surveillance Video


Original video

Foreground

Background

Monitoring Engine Degradation

- Support Vector Data Description (SVDD) to detect engine degradation:

  - Single class classification technique

  - Multi modal operating conditions

  - Real time scoring for anomaly detection

- NASA Turbofan Engine simulated data set – 218 engines operated until failure, 26 variables

§sas

Results for Randomly Selected Engines
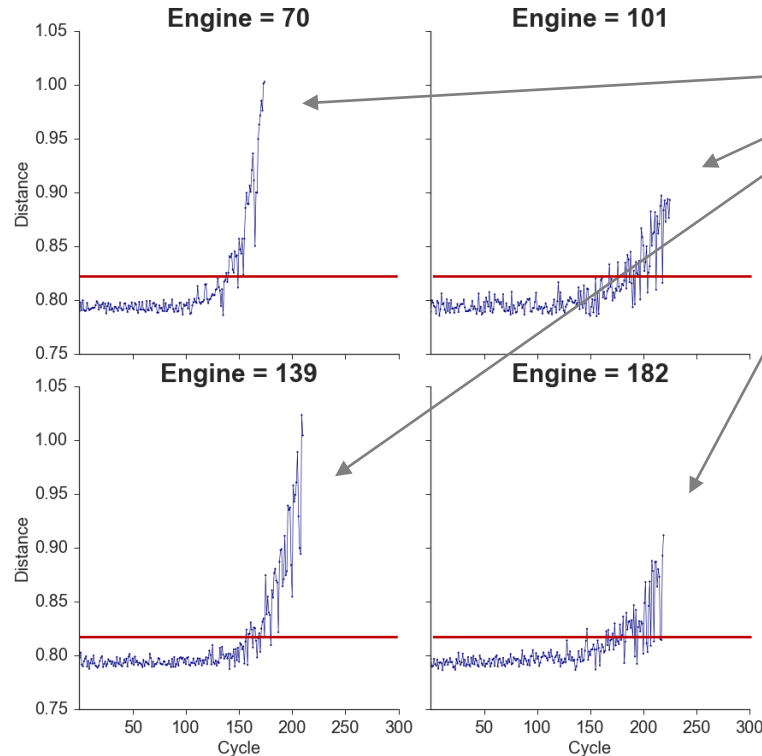
Different patterns of degradation

- Distance measure scored by SVDD for each flight
- Compared to the R for normal operation
- Alerts and filters can be set on how the distance measure compares to the R of normal classification
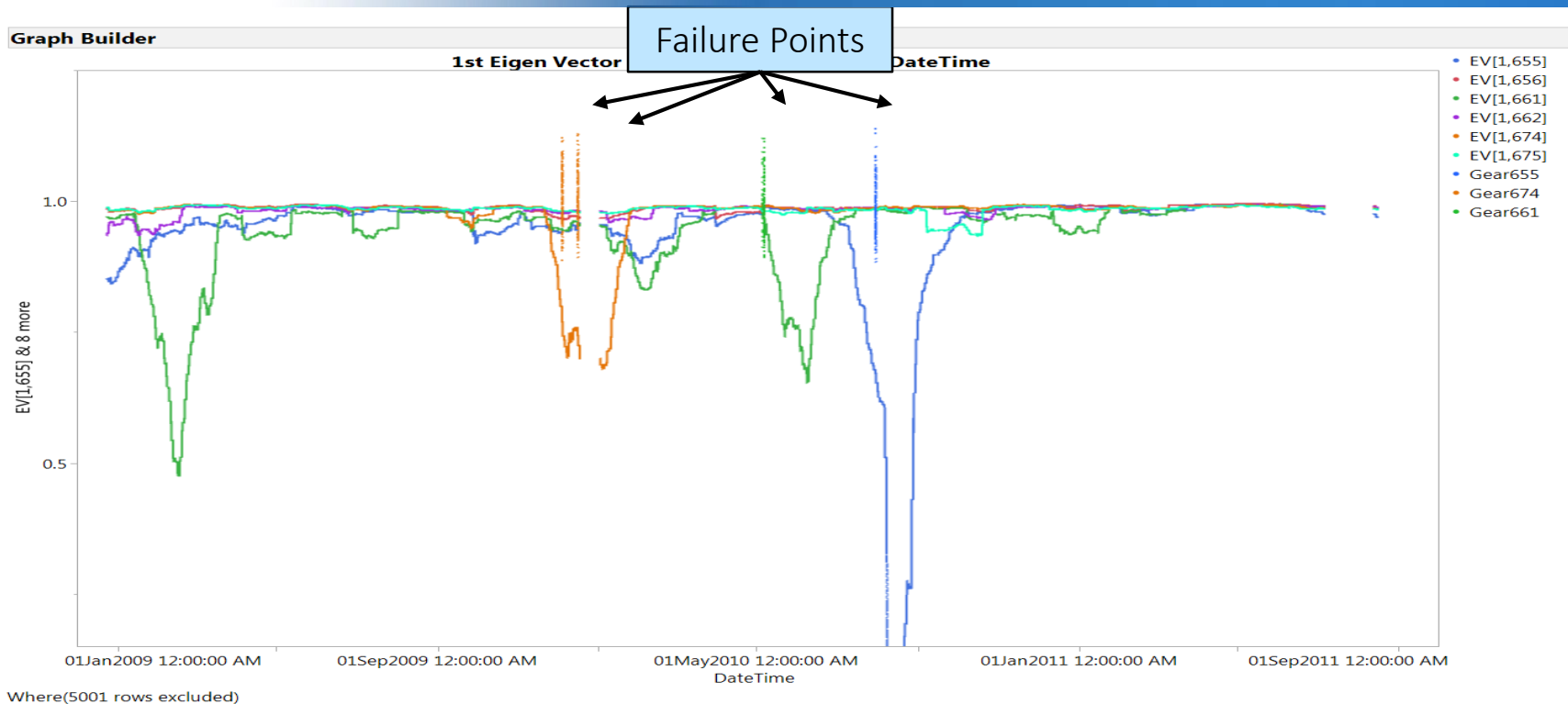
Detecting Windmill Degradation

Intuition: "Birds of a feather flock together"

- Similar machines facing the same environment should be behaving similarly
- If a machine is degrading, you should be able to see it in comparison to the other machines
- Moving Windows Principal Components (MWPCA)

§sas

Detecting Windmill Degradation with MWPCA

Success Criteria for Advanced Analytics

- Insight
- Influence
- Impact

§sas

Thank you.