Building Energy Data Analytics: Current Status and Future Directions

Brock Glasgo, Postdoctoral Research Associate

7th Annual High Performance Buildings Summit

Cincinnati, OH | October 4, 2018

Carnegie
Mellon
University

Motivation

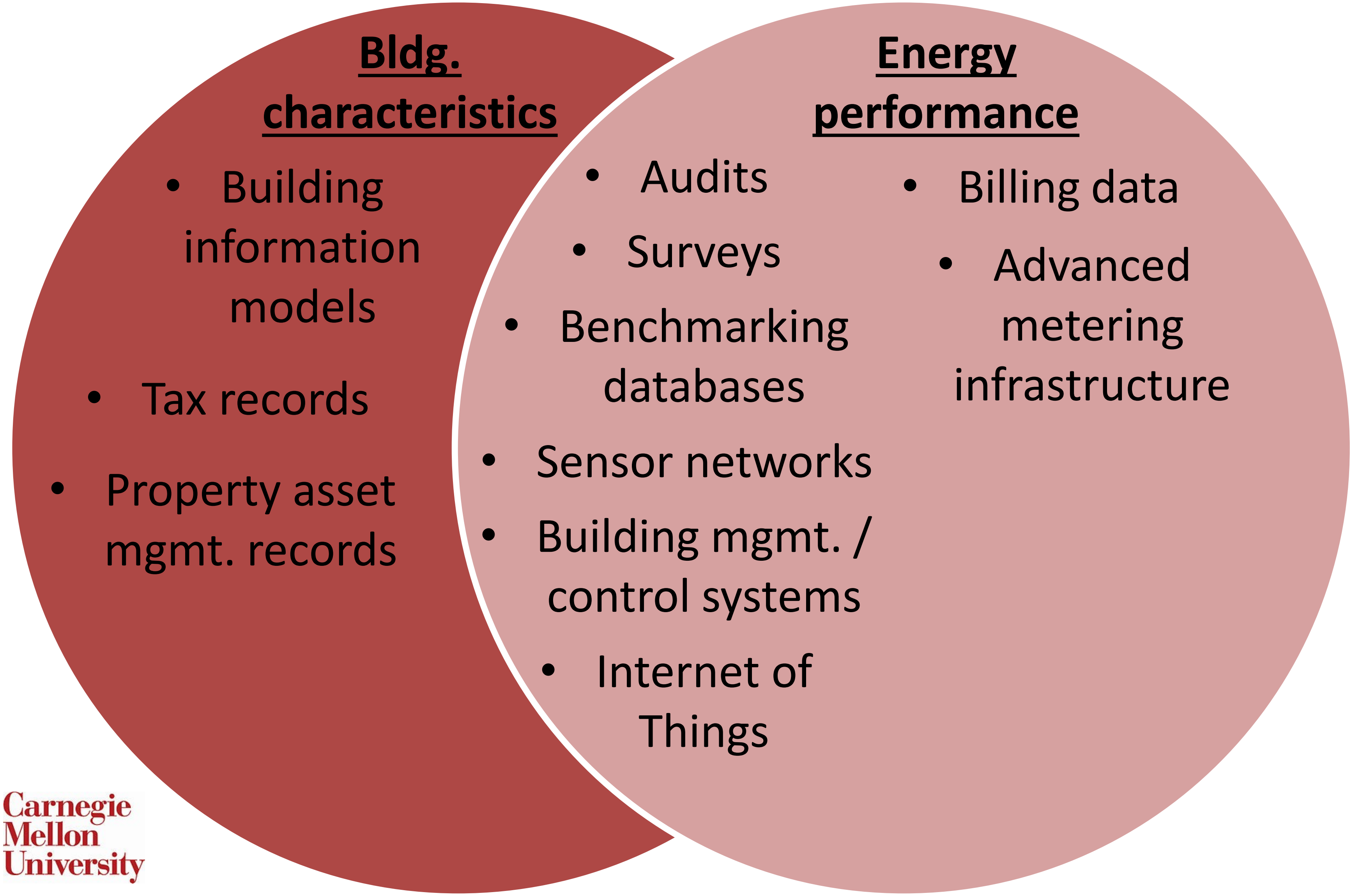
- Buildings generate more data than operators can process
  - Vendors and consumers have confused data with information
- Analytics and technologies are now catching up to all of that data
  - These methods can be overwhelming
- Goal: Provide a high-level introduction to the data, methods, and applications of those methods

Learning objectives

- What are the main sources of building and energy data?
- What are the methods being used to translate that data into actionable information?
- How are these datasets and methods being applied today?
- Where is building energy data analysis headed in the future?

Carnegie
Mellon
University

Where’s this data coming from?



Bldg. characteristics

- Building information models
- Tax records
- Property asset mgmt. records

- Audits
- Surveys
- Benchmarking databases
- Sensor networks
- Building mgmt. / control systems
- Internet of Things

Energy performance

- Billing data
- Advanced metering infrastructure

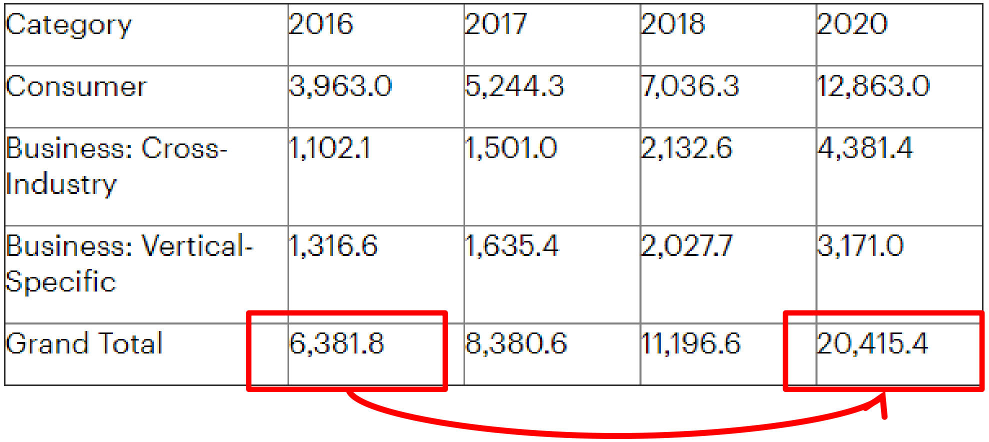
Advanced metering infrastructure: "The Smart grid"

- Smart meters now serve ~50% of US customers
- Utilities' focus has been on transmission and distribution analytics
  - Quickly identifying and responding to outages
  - Power quality analytics to identify faulty equipment
- Application to buildings lags behind the technology
  - Online dashboards
  - Time-of-use pricing
  - Load control
  - Targeting and M&V of efficiency interventions
- Burden for extracting value is left to customers

Source(s): Gartner

The Internet of Things (IoT)

- Essentially the movement toward connecting more devices to the internet
- 6.4 billion connected devices in 2016, expected to increase to over 20 billion by 2020



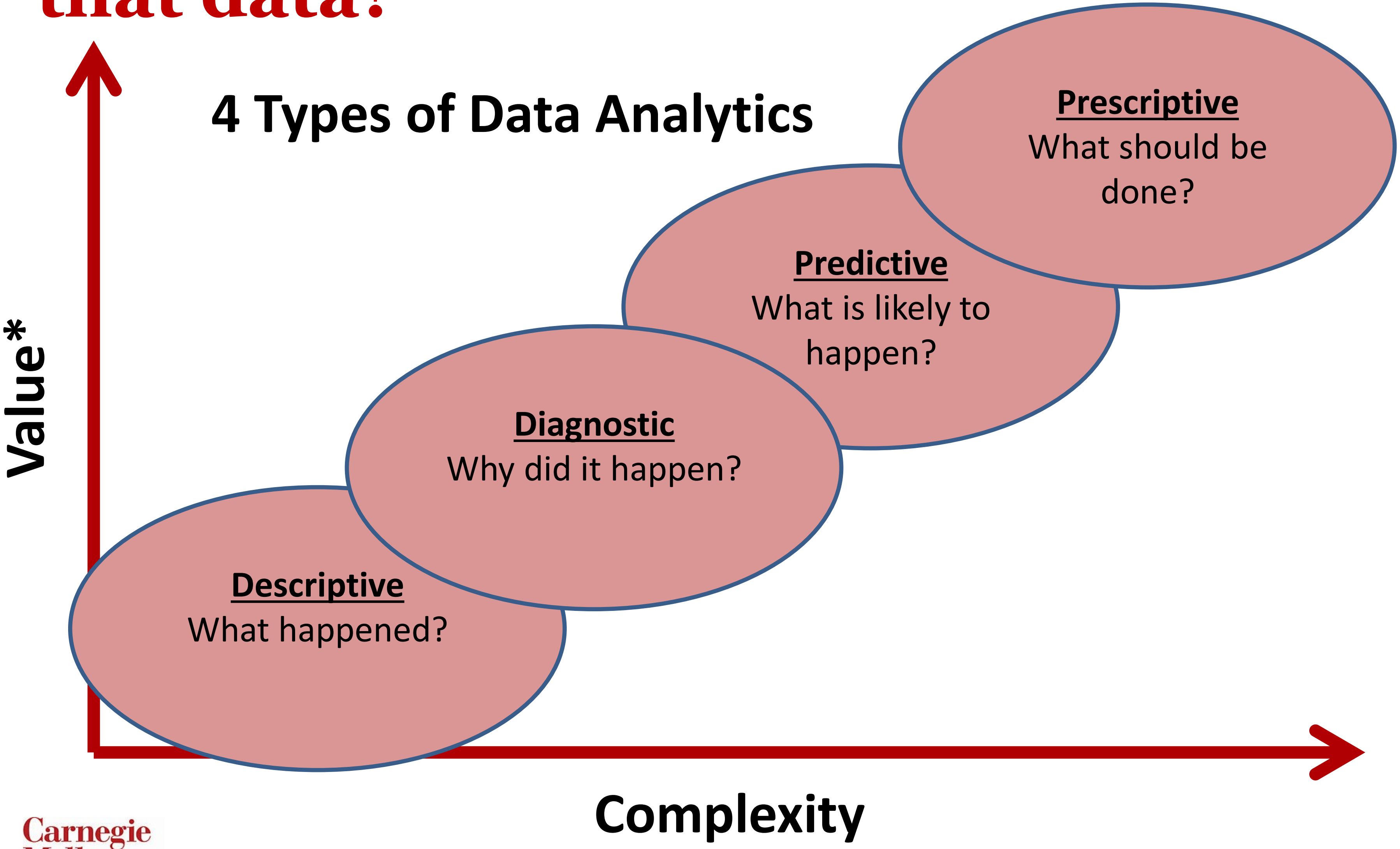
Category	2016	2017	2018	2020
Consumer	3,963.0	5,244.3	7,036.3	12,863.0
Business: Cross-Industry	1,102.1	1,501.0	2,132.6	4,381.4
Business: Vertical-Specific	1,316.6	1,635.4	2,027.7	3,171.0
Grand Total	6,381.8	8,380.6	11,196.6	20,415.4

- The future of IoT is uncertain, but some things are clear
- Connectivity will:

1. Make more devices remotely controllable
2. Generate more data about how devices operate
3. Create more opportunities for understanding and optimizing that operation

Source(s): Gartner

How do we extract meaning from that data?



4 Types of Data Analytics

Descriptive
What happened?

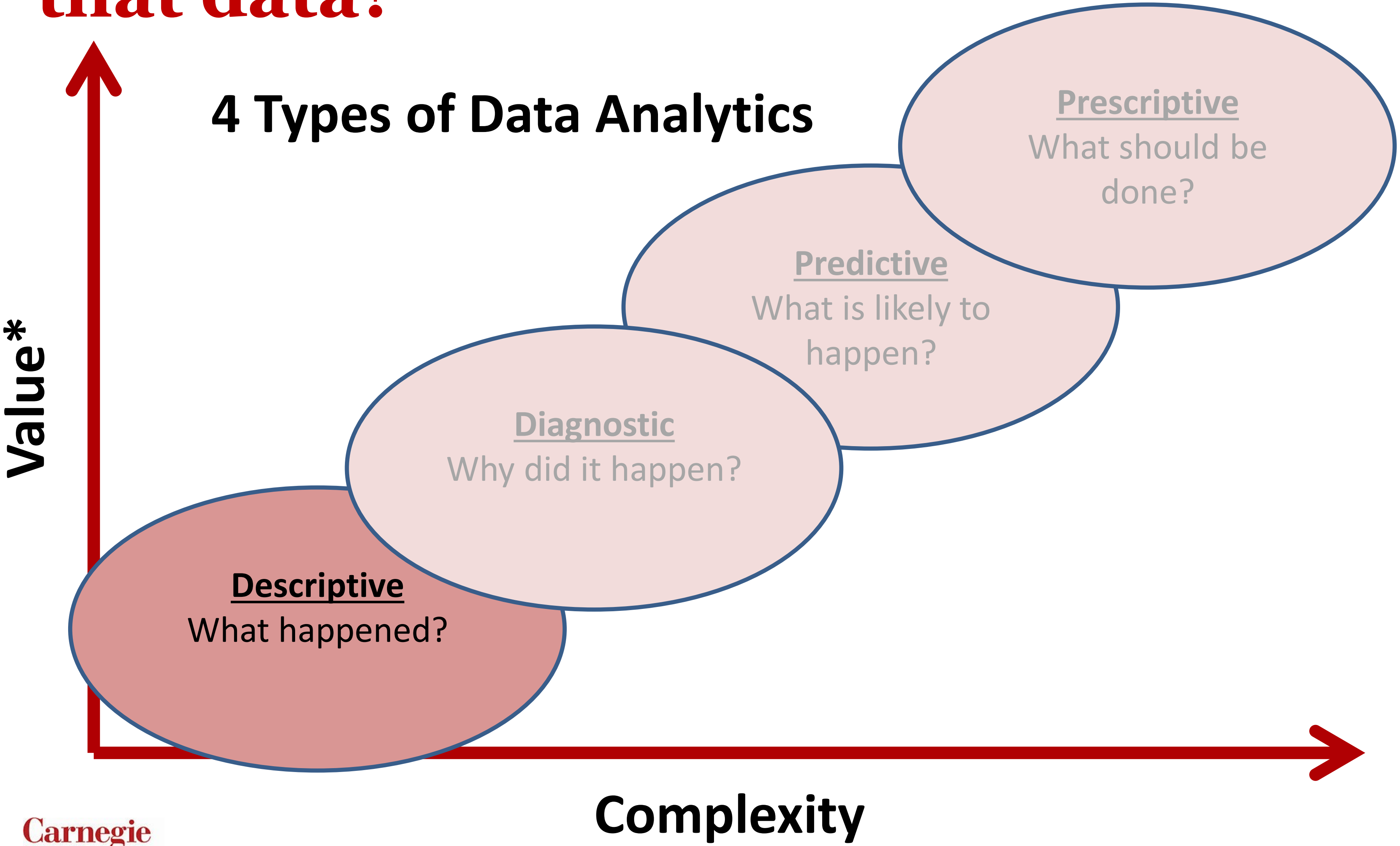
Diagnostic
Why did it happen?

Predictive
What is likely to happen?

Prescriptive
What should be done?

Carnegie Mellon University

How do we extract meaning from that data?



4 Types of Data Analytics

Descriptive
What happened?

Diagnostic
Why did it happen?

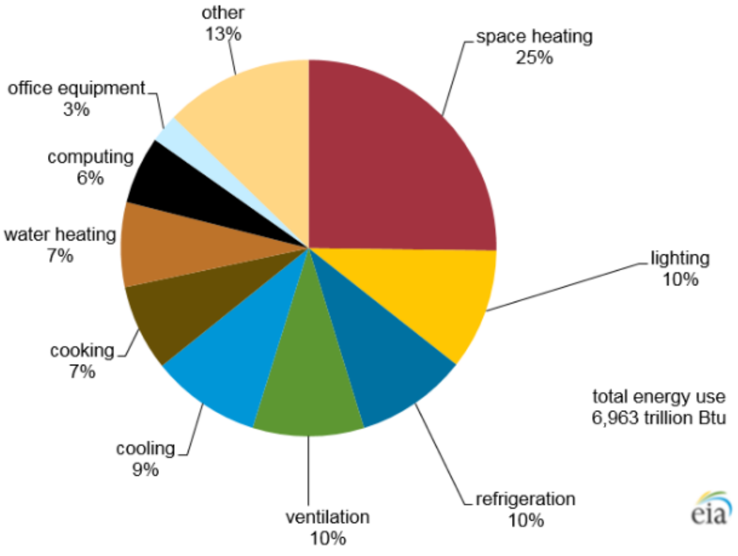
Predictive
What is likely to happen?

Prescriptive
What should be done?

Carnegie Mellon University

Descriptive analytics: What happened?

- Key questions:
  - How much energy is my building consuming?
  - When and where is that energy being consumed?
- Data involved:
  - Historic energy use data
  - Building and system characteristics
- Methods:
  - Summary statistics
- Applications:
  - Baselining and benchmarking
  - Simple dashboards



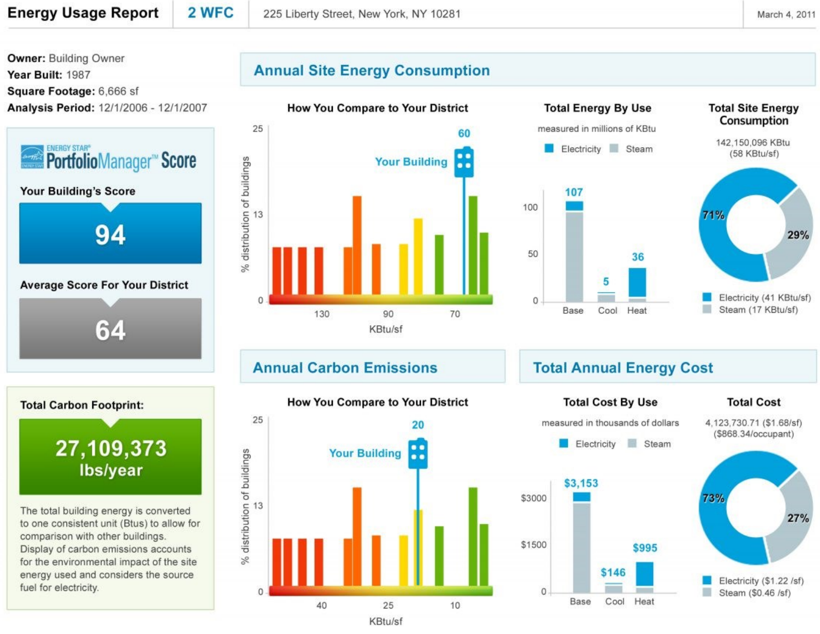
Source(s): Northwest Mechanical, CBECS

Case Study #1 – ENERGY STAR Portfolio Manager

- Online tool to simplify and standardize building energy baselining and benchmarking

- Floor area
- Year built
- Occupancy
- Energy bills
- Operating hours
- # workers
- # computers
- % cooled
- More…

ENERGY STAR



Energy Usage Report | 2 WFC | 225 Liberty Street, New York, NY 10281 | March 4, 2011

Owner: Building Owner
Year Built: 1987
Square Footage: 6,666 sf
Analysis Period: 12/1/2006 - 12/1/2007

ENERGY STAR® PortfolioManager™ Score

Your Building's Score: 94

Average Score For Your District: 64

Total Carbon Footprint: 27,109,373 lbs/year

The total building energy is converted to one consistent unit (Btus) to allow for comparison with other buildings. Display of carbon emissions accounts for the environmental impact of the site energy used and considers the source fuel for electricity.

Annual Site Energy Consumption

Total Site Energy Consumption: 142,150,096 KBtu (58 KBtu/sf)

Annual Carbon Emissions

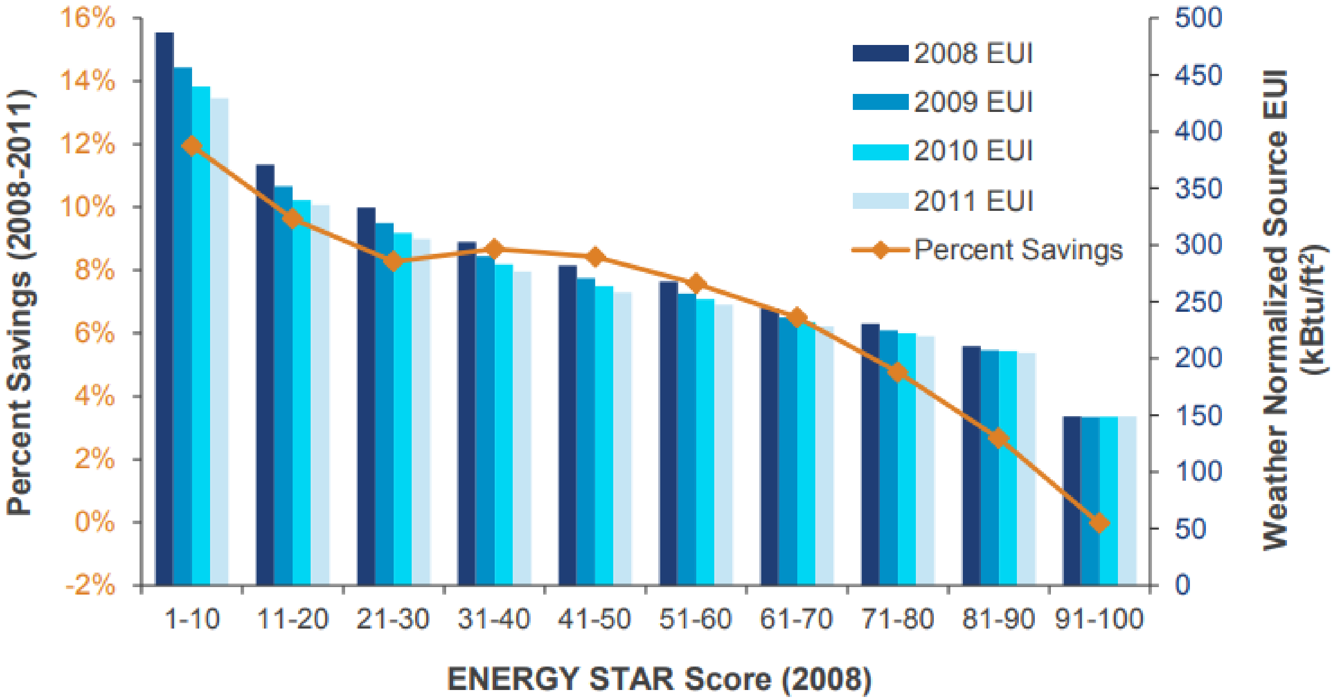
Total Annual Energy Cost

Total Cost: 4,123,730.71 ($1.68/sf) ($868.34/occupant)

Source(s): EPA, Taylor

Case Study #1 – ENERGY STAR Portfolio Manager

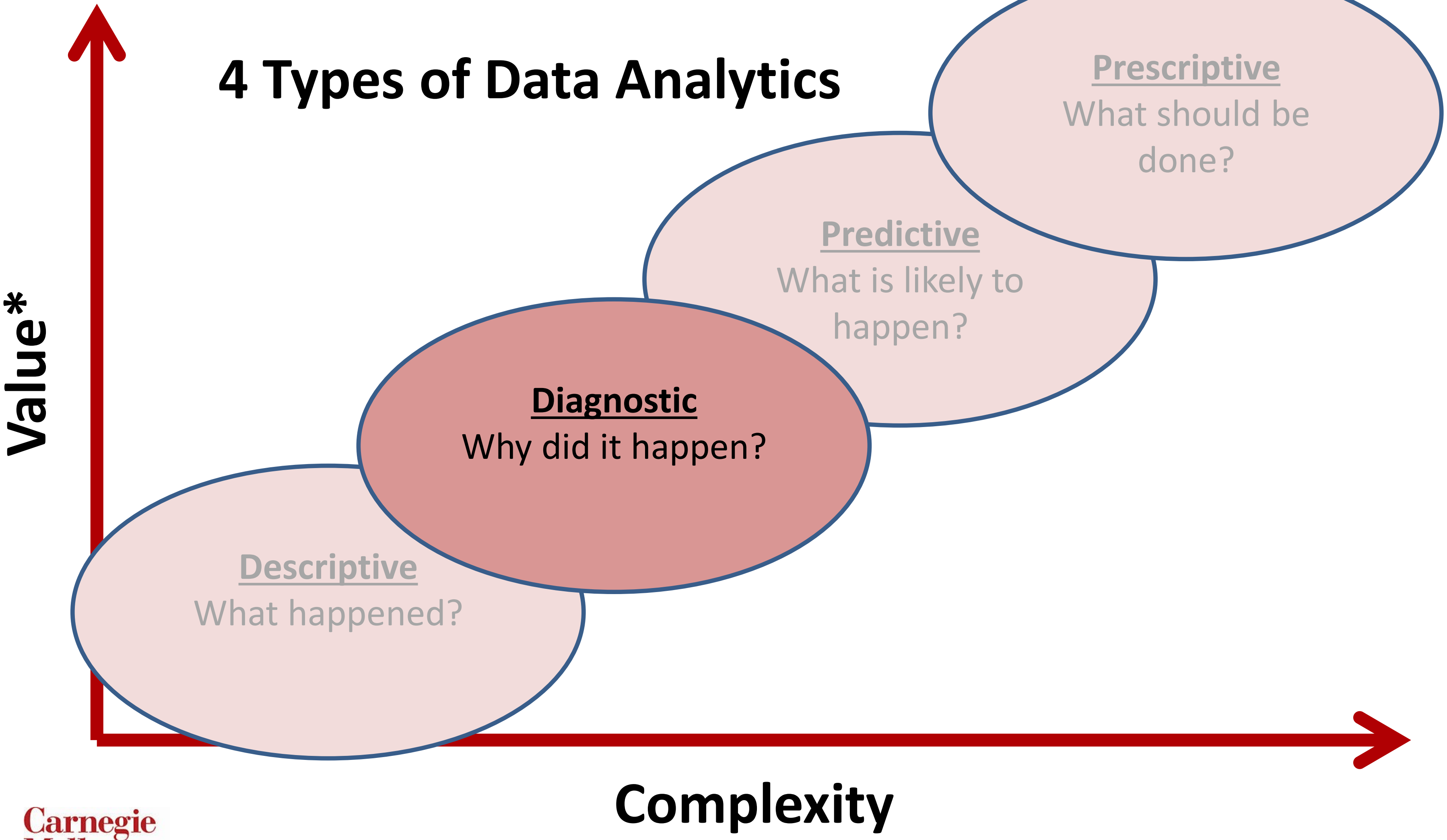
- Buildings that benchmark are saving energy



- 26 cities and 12 states have mandatory benchmarking laws

Source(s): EPA

How do we extract meaning from that data?



4 Types of Data Analytics

Descriptive
What happened?

Diagnostic
Why did it happen?

Predictive
What is likely to happen?

Prescriptive
What should be done?

Carnegie Mellon University

Diagnostic analytics: Why did it happen?

- Key questions:
  - How is my building’s energy use affected by weather, occupancy, time of day, production output, efficiency measures, operational changes, and other variables?
- Data involved:
  - Interval data
  - Operational data
  - Weather data
- Methods:
  - Regression
  - Machine learning
- Applications:
  - Identifying key drivers of energy use
  - M&V of efficiency measures
  - Load disaggregation

Diagnostic analytics: Regression

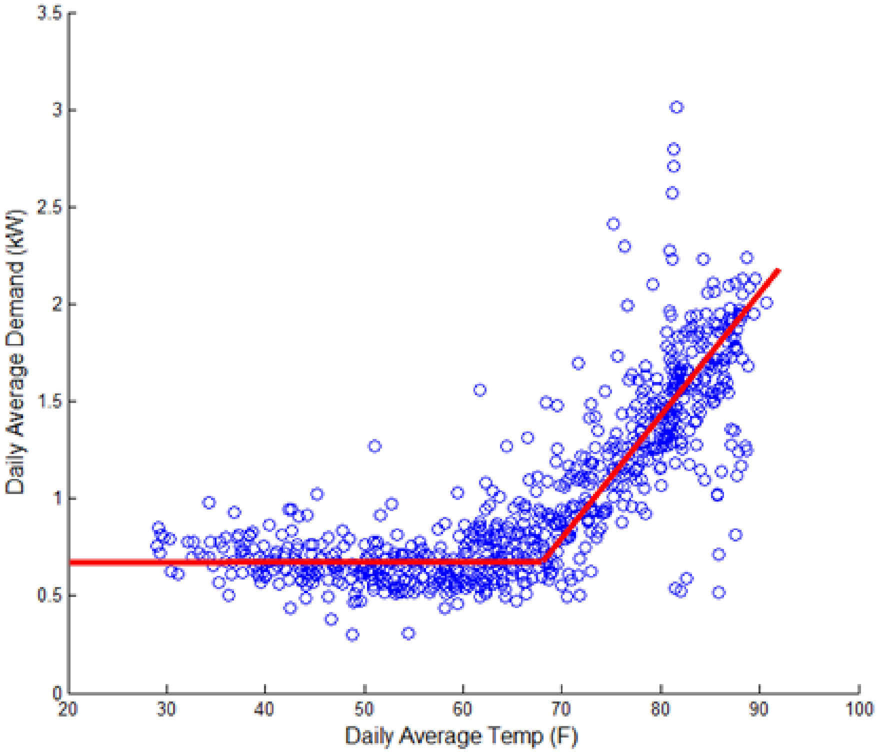
- Generally follow the form:

$$y = \beta_0 + \beta_n \cdot X_n + \mu$$

y is the variable being described (energy use or demand)

X are predictor variables (weather, occupancy, time of day, etc.)

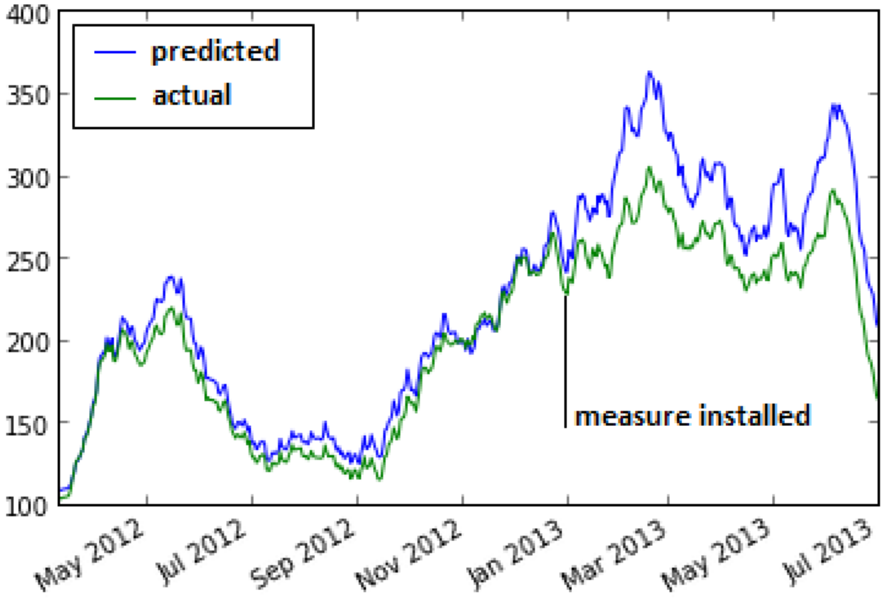
β are the strengths of effects



Case Study #2: M&V of ECMs

- Regression is a widely accepted metric for M&V of efficiency measures
- Steps

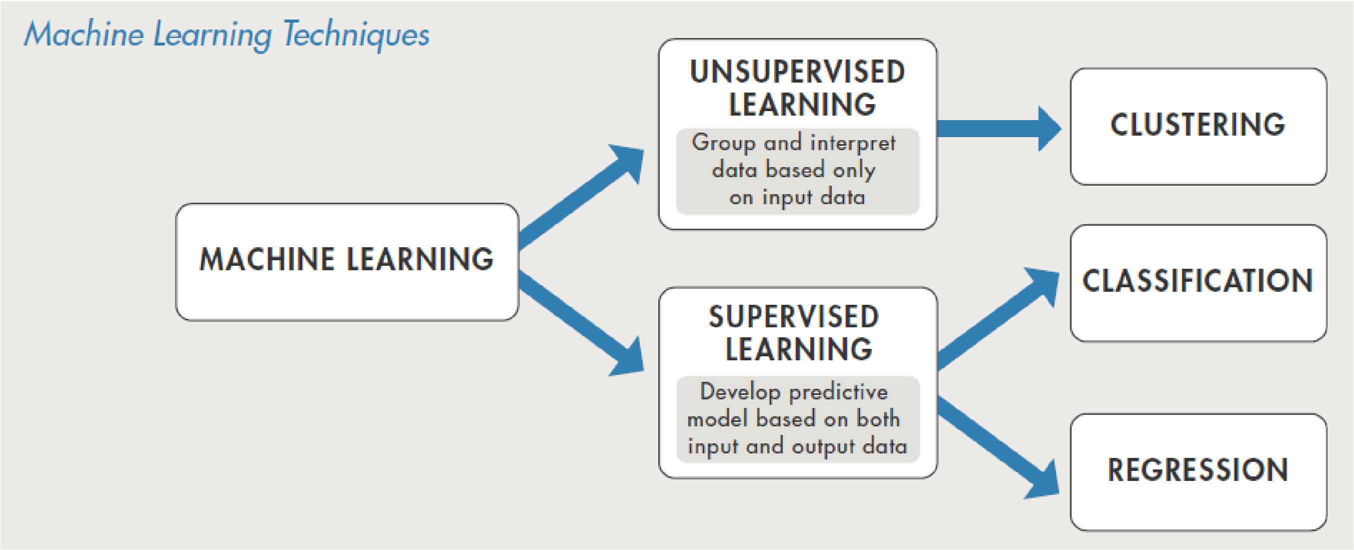
1. Identify all predictor variables
2. Collect data before and after the measure was implemented
3. Estimate the β's from data before the ECM
4. Calculate predicted energy consumption (based on existing operation)
5. Compare predicted consumption to actual consumption after the measure was installed



Carnegie Mellon University

Diagnostic analytics: Machine Learning

- Regression: specify a model, then add data
- Machine learning methods start with data and estimate the underlying model
- Useful when a problem has a large amount of data, and messy/no equations defining relationships



Machine Learning Techniques

MACHINE LEARNING

UNSUPERVISED LEARNING

Group and interpret data based only on input data

SUPERVISED LEARNING

Develop predictive model based on both input and output data

CLUSTERING

CLASSIFICATION

REGRESSION

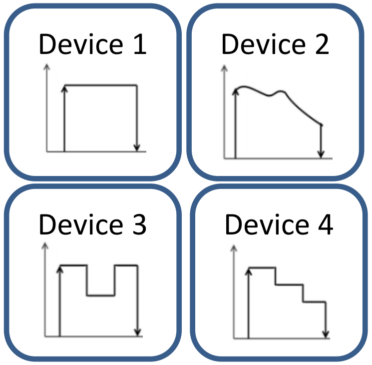
Carnegie Mellon University

Source(s): Mathworks

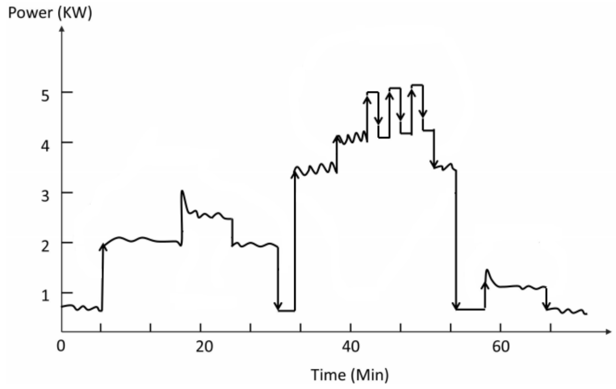
Case Study #3: Load disaggregation

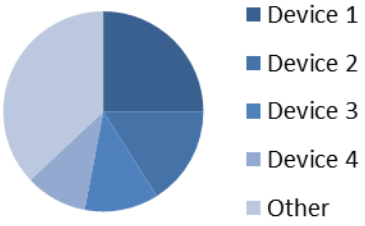
- Supervised learning is the basis for most disaggregation platforms

Training data



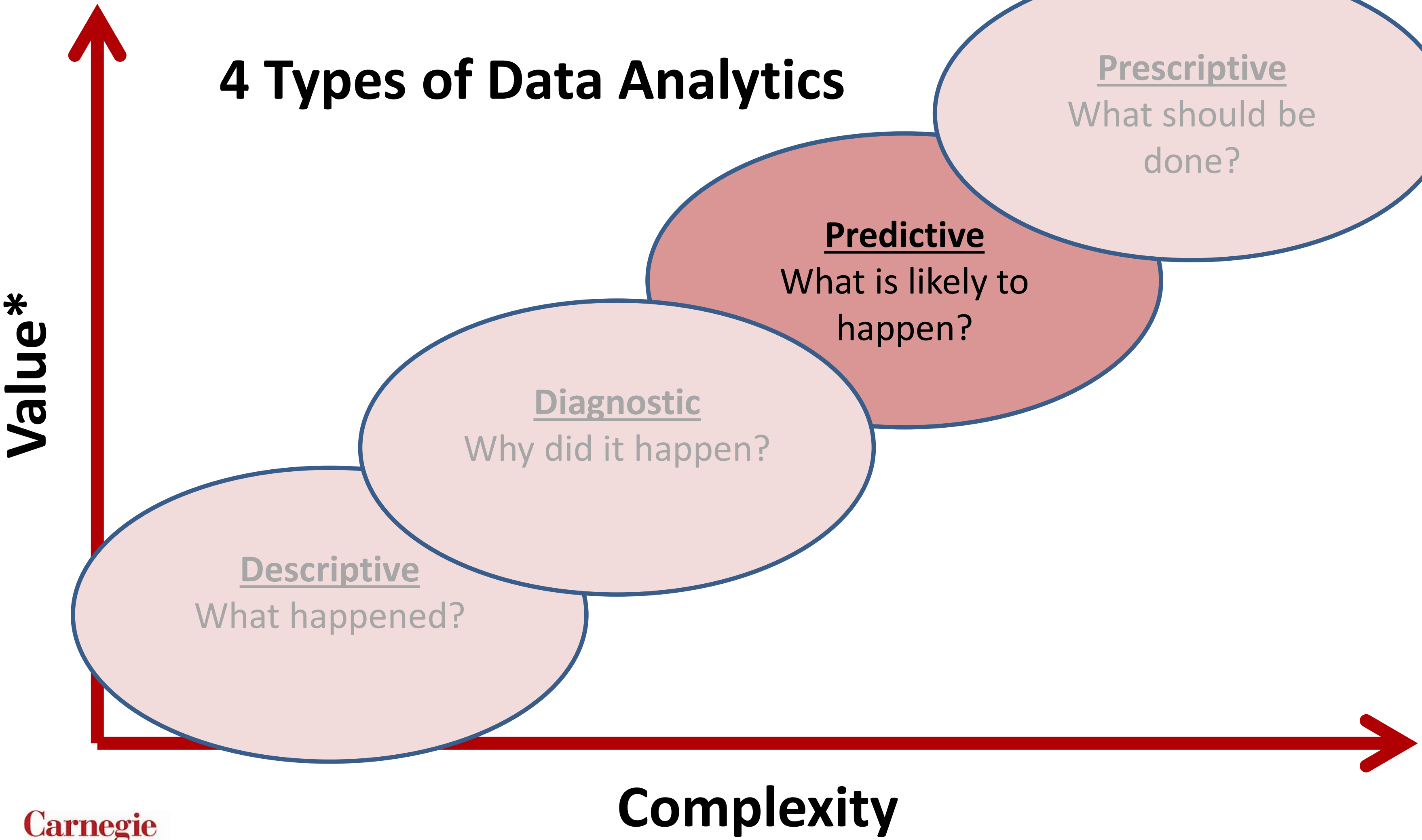
Real-time, kHz frequency demand data





Source(s): Zoha, Sense

How do we extract meaning from that data?



Descriptive
What happened?

Diagnostic
Why did it happen?

Predictive
What is likely to happen?

Prescriptive
What should be done?

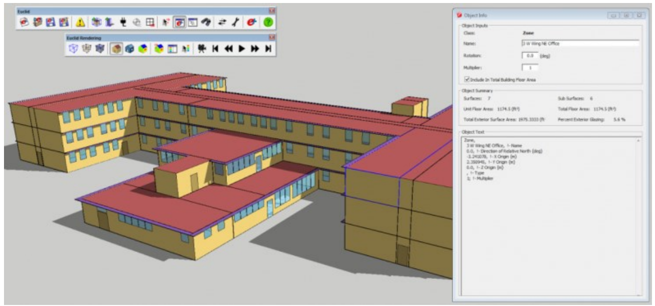
Carnegie Mellon University

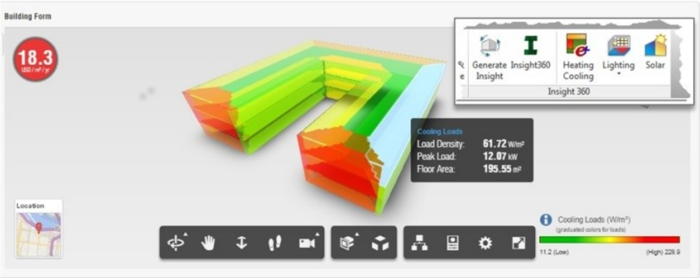
Predictive analytics: What is likely to happen?

- Key questions:
  - How will a certain design decision affect a future building’s energy performance?
  - How will an equipment or operational change affect energy performance?
- Data involved:
  - Detailed building characteristics
  - Historic energy data
  - Forecast data: weather, occupancy, production output, etc.
- Methods:
  - Physics-based building simulation
  - Regression
  - Statistical simulation
  - Machine learning
- Applications:
  - Evaluating design options
  - Estimating savings from ECMs

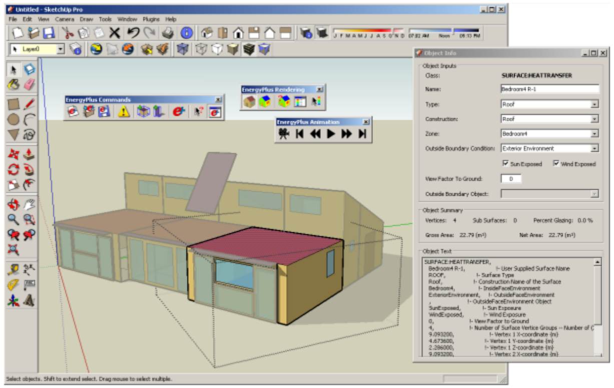
Case Study #4: Building energy simulation

- Building energy simulation models – EnergyPlus, eQuest, Trane Trace, etc. – are seeing increased use across building sectors
- Take inputs of detailed building characteristics, materials, location, occupancy, simulate energy performance
- More user-friendly + less computationally intensive = increased use
  - Residential sector
  - Energy code analysis, design, and verification





Cooling Loads
Load Density: 61.72 W/m²
Peak Load: 12.07 kW
Floor Area: 195.55 m²

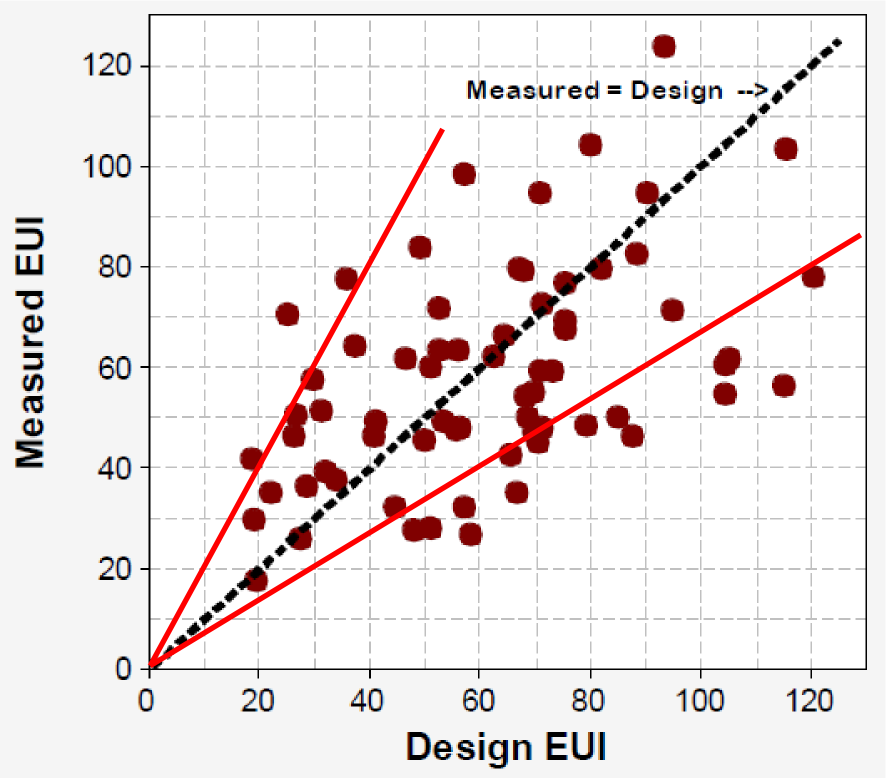


Carnegie Mellon University

Source(s): EERE, Crawley

Case Study #4: Building energy simulation

- Tools are improving, but accuracy is not
- Attention is now turning to increasing the accuracy of these models through validation and calibration
  - Simulations for retrofits and renovations validated using historical data
  - Simulations for new buildings should present uncertainty
- ASHRAE Guideline 14 lays out metrics and tolerance limits to define a calibrated simulation

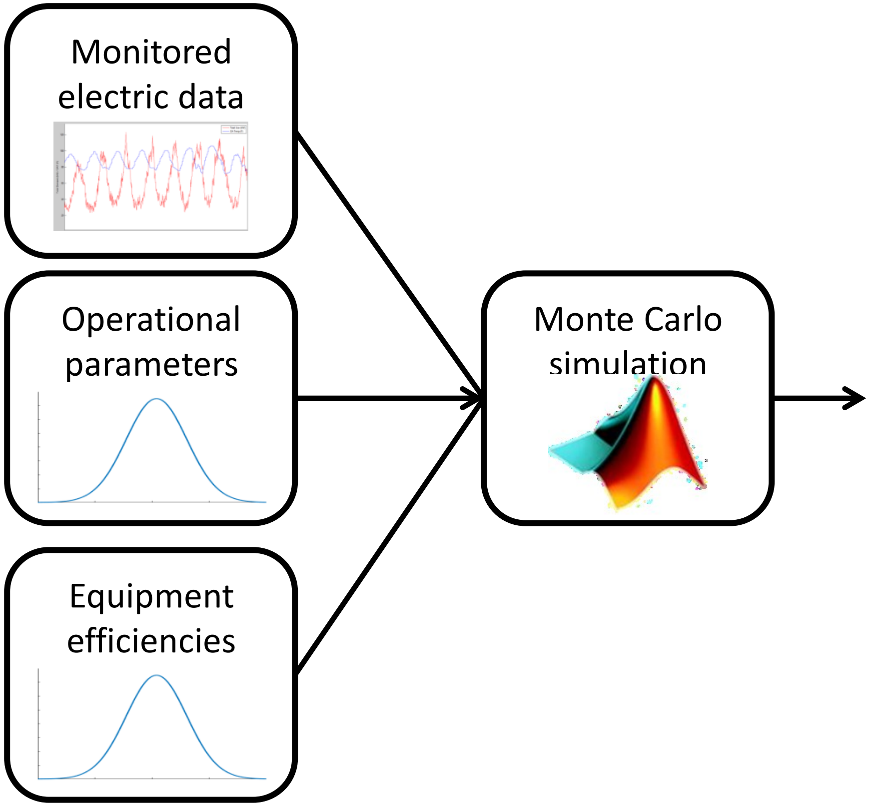


Carnegie Mellon University

Source(s): Turner

Case Study #5: Statistical simulation

- When detailed historical demand data is available, estimates of existing and proposed equipment operating parameters can be used to simulate interventions
- Monte Carlo simulation methods handle uncertainty

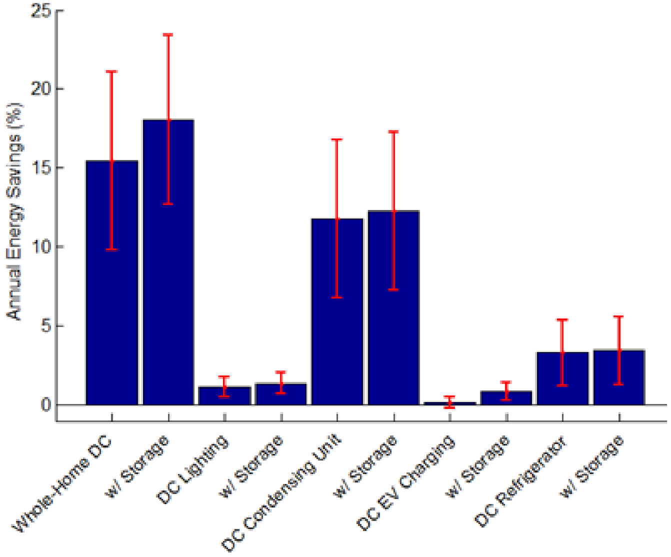


Monitored electric data

Operational parameters

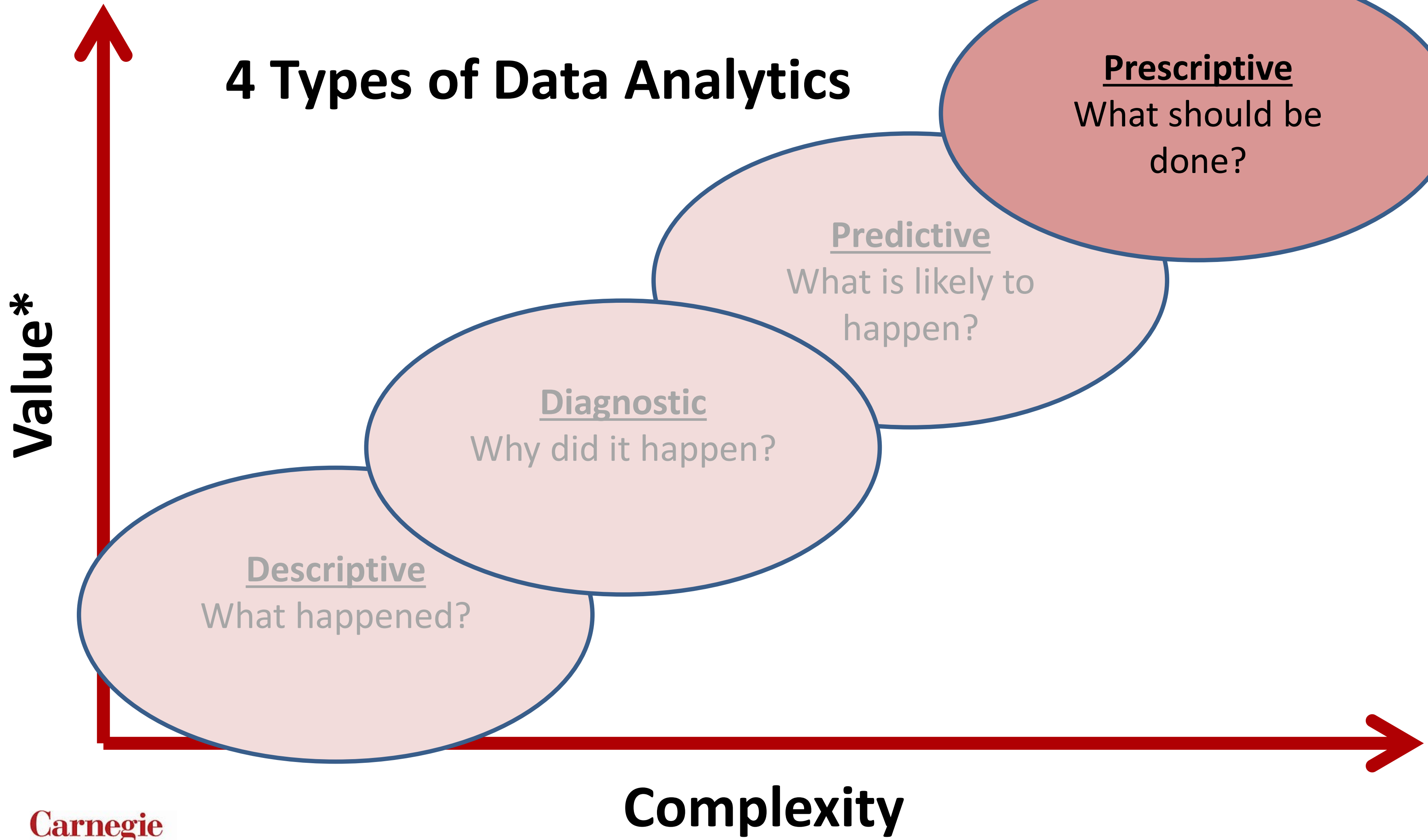
Equipment efficiencies

Monte Carlo simulation



Source(s): Glasgo

How do we extract meaning from that data?



4 Types of Data Analytics

Prescriptive
What should be done?

Predictive
What is likely to happen?

Diagnostic
Why did it happen?

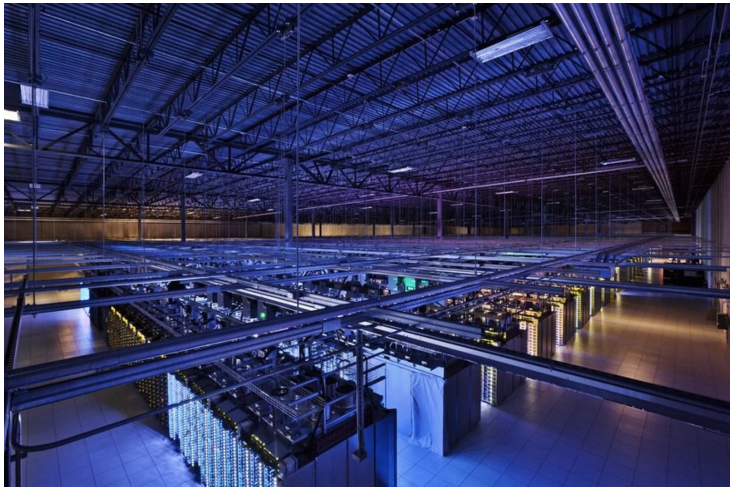
Descriptive
What happened?

Prescriptive analytics: What should be done?

- Key questions:
  - How should a building operate to optimize its energy performance?
- Data involved:
  - Multiple, large datasets
  - Controls system trend data
  - Indoor and outdoor environmental sensor data
  - Submeter power data
  - Energy cost data
- Methods:
  - Supervised machine learning
- Applications:
  - Building control optimization

Case Study #6: DeepMind

- Machine learning firm DeepMind running Google's data centers
- Google's data centers were already efficient
  - PUE of 1.12 (12% overhead energy)
  - Industry average is around 1.7 (70% overhead)



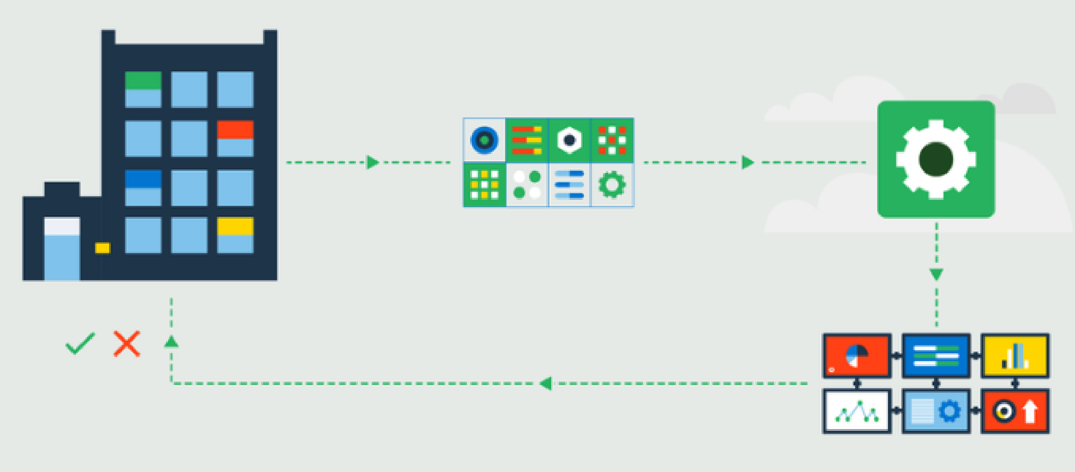
Source(s): DeepMind

Case Study #6: DeepMind

Historical trend data used to train algorithms to identify relationships between variables

Snapshot of current operations sent to the cloud-based algorithm

Future parameters are predicted based on possible actions



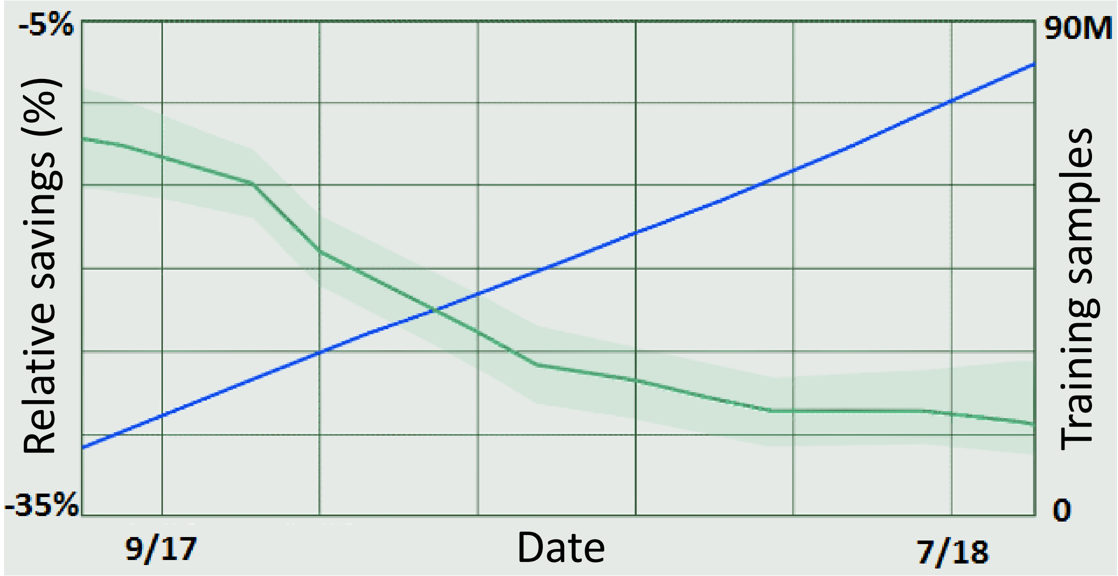
Actions are chosen to meet safety constraints and optimize energy performance

Setpoints are validated against safety checks and sent to the equipment

Source(s): DeepMind

Case Study #6: DeepMind

- Over time, added training data improves performance
  - Relative savings increase from 12% to around 30%



- Long-term plans to expand beyond data centers

Carnegie Mellon University

Source(s): DeepMind

What’s next?

Machine learning advances

- Real-time control optimization
- Load disaggregation
- Fault detection and diagnostics
- Automated building energy model calibration

Open source BMS

- Existing building management systems limit custom analytics and innovation potential
  - Equipment, networks, and software are owned and controlled by vendors
  - Different control systems don’t integrate well
    - Lighting controls vs HVAC controls
- More flexible, open source platforms that allow for communication, data sharing, and programming between networks, system types, and technologies
  - Allow third party research, analytics, and innovation

Source(s):Weng

Transactive energy

- New form of coordinated control of grid supply and demand
  - Elements of direct load control and price response control
- Uses dynamic energy pricing as operational parameter to control flexible demand and generation
- Enables consumers to produce, buy, and sell electricity using automated control
- Benefits:
  - Reduce peak demand and grid constraints
  - Wholesale price purchases by utilities
  - Building energy cost savings

Carnegie
Mellon
University

THANK YOU. QUESTIONS?

Contact:
Brock Glasgo
bpglasgo@gmail.com
513.519.1008

Carnegie Mellon University

Resources

1. The Edison Foundation Institute for Electric Innovation. *Electric Company Smart Meter Deployments: Foundation for a Smart Grid.* 2017.
2. Gartner. *Press release: Gartner says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016*. https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016
3. Northwest Mechanical. *Automation and Controls.* https://www.northwestmech.com/automation-controls/energy-advice/
4. Taylor Consulting and Contracting. *Use ENERGYSTAR Portfolio Manager.* https://www.taylorcc.com/use-energy-star-portfolio-manager/
5. EPA. *ENERGYSTAR Portfolio Manager*. https://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager?s=mega
6. Mathworks. *Introducing Machine Learning.* 2018.
7. Zoha, et al. *Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey*. Sensors, 2012.
8. Turner et al. *Energy Performance of LEED for New Construction Buildings*. New Buildings Institute, 2008.
9. Glasgo et al. How much electricity can we save by using direct current circuits in homes?. Applied Energy.
10. EERE. Buildings Technologies Office. https://www.energy.gov/eere/buildings/building-technologies-office
11. Crawley et al. *Energy Design Plugin: An EnergyPlus Plugin for Sketchup*. 2008. https://www.nrel.gov/docs/fy08osti/43569.pdf
12. DeepMind. *Safety-first AI for autonomous data center cooling and industrial control*. https://deepmind.com/blog/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control/
13. Weng. *Building Depot 2.0: An Integrated Management System for Building Analysis and Control*. Buildsys '13.
14. PNNL. Transactive System Part 1: Theoretical underpinnings of payoff functions, control decisions, information privacy, and solution concepts. 2017. https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-27235Part1.pdf