

A Datapalooza for Measured Building Performance: The DOE Buildings Performance Database

Paul Mathew

Staff Scientist, Commercial Building Systems Group

Lawrence Berkeley National Laboratory

November 4, 2013

BPD Team



Rich Brown
Claudine Custudio
Laurel Dunn
Paul Mathew
John Mejia
Andrea Mercado
Michael Sohn
Travis Walter

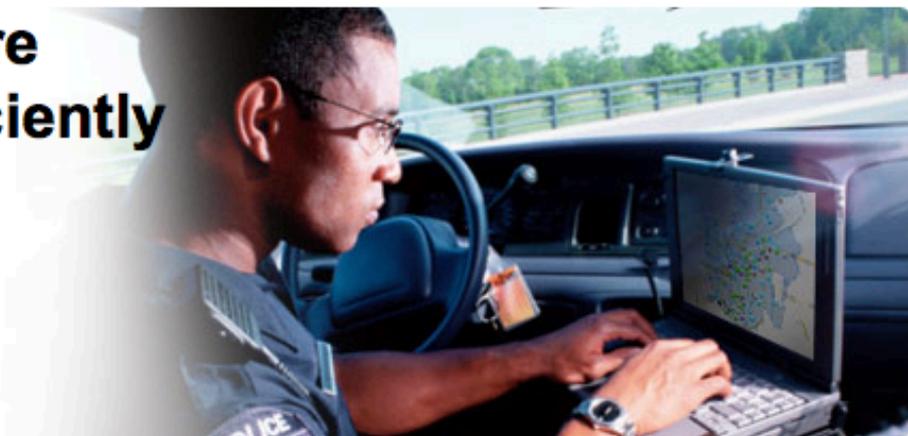
Software partner:



Sponsor:



Deploy Officers More Effectively and Efficiently with DDACTS



What is DDACTS?

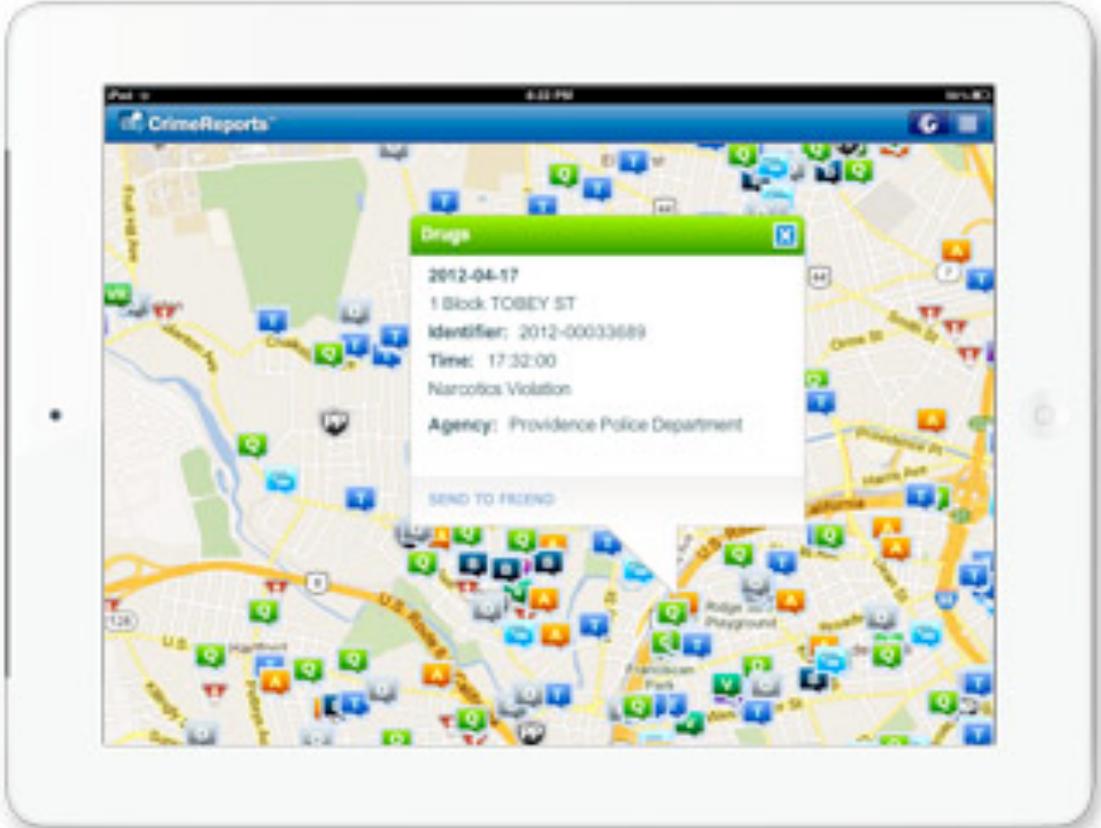
Data-Driven Approaches to Crime and Traffic Safety (DDACTS) is one of the most effective and efficient methods for deploying law enforcement officers to areas that have high incidences of crime and vehicle crashes, drug crime, as well as reducing crashes and traffic violations. (DDACTS) is a vehicle, the primary objectives of DDACTS are to reduce

How Does Incident Mapping Technology Methodology?

An ever-increasing number of law enforcement agencies are using traffic data to dramatically lower traffic accidents and victimization, drop in residential and commercial burglaries and a 17%

How Can PublicEngines Support

Public Engines offers two industry-leading solutions that allow traffic safety agencies proactively tackle crime and minimize traffic safety resources can be deployed much more effi





FiveThirtyEight

Nate Silver's Political Calculus

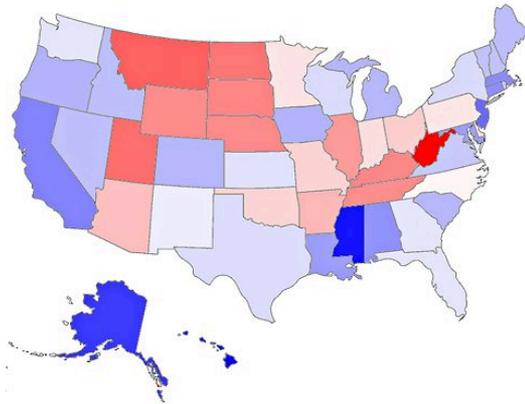
November 21, 2012, 4:28 pm | [Comment](#)

Where Obama and Romney Beat Their Polls

By NATE SILVER

I'm traveling for Thanksgiving, so we'll keep this relatively brief. But I thought this map was worth sharing. It shows how President Obama and Mitt Romney performed on Election Day relative to the FiveThirtyEight forecasts in each state, based on the ballots counted so far.

States colored in blue represent those where Mr. Obama beat his forecast — the deeper the blue color, the larger the margin by which he did so — while those in red are the states where Mr. Romney bested his.



A few things jump out here.

First, there are some pretty clear regional patterns in which each candidate beat his forecast (and, by extension, beat the polls). States where Mr. Obama beat the polls (Oregon, for example) tended to border others where he also

Search This Blog

All NYTimes.com Blogs »

FOLLOW THIS BLOG [Twitter](#) [RSS](#)

FiveThirtyEight Forecast

Updated 10:10 AM ET on Nov. 6

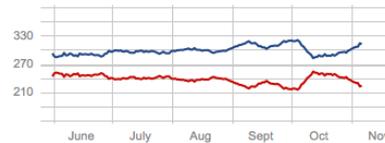
President Nov. 6 Forecast	President Now-cast	Senate Nov. 6 Forecast
------------------------------	-----------------------	---------------------------

Barack Obama

Mitt Romney

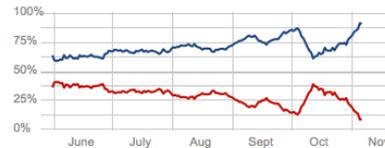
313.0 **Electoral vote** **225.0**
+14.0 since Oct. 30 -14.0 since Oct. 30

270 to win



90.9% **Chance of Winning** **9.1%**
+13.5 since Oct. 30 -13.5 since Oct. 30

50%



50.8% **Popular vote** **48.3%**
+0.4 since Oct. 30 -0.2 since Oct. 30

50%



..... analytical revolution upending the way campaigns political are run in the 21st century... the smartest campaigns now believe they know who you will vote for even before you do...

The New York Times

OP-ED CONTRIBUTOR

Don't LEED Us Astray

By ALEC APPELBAUM

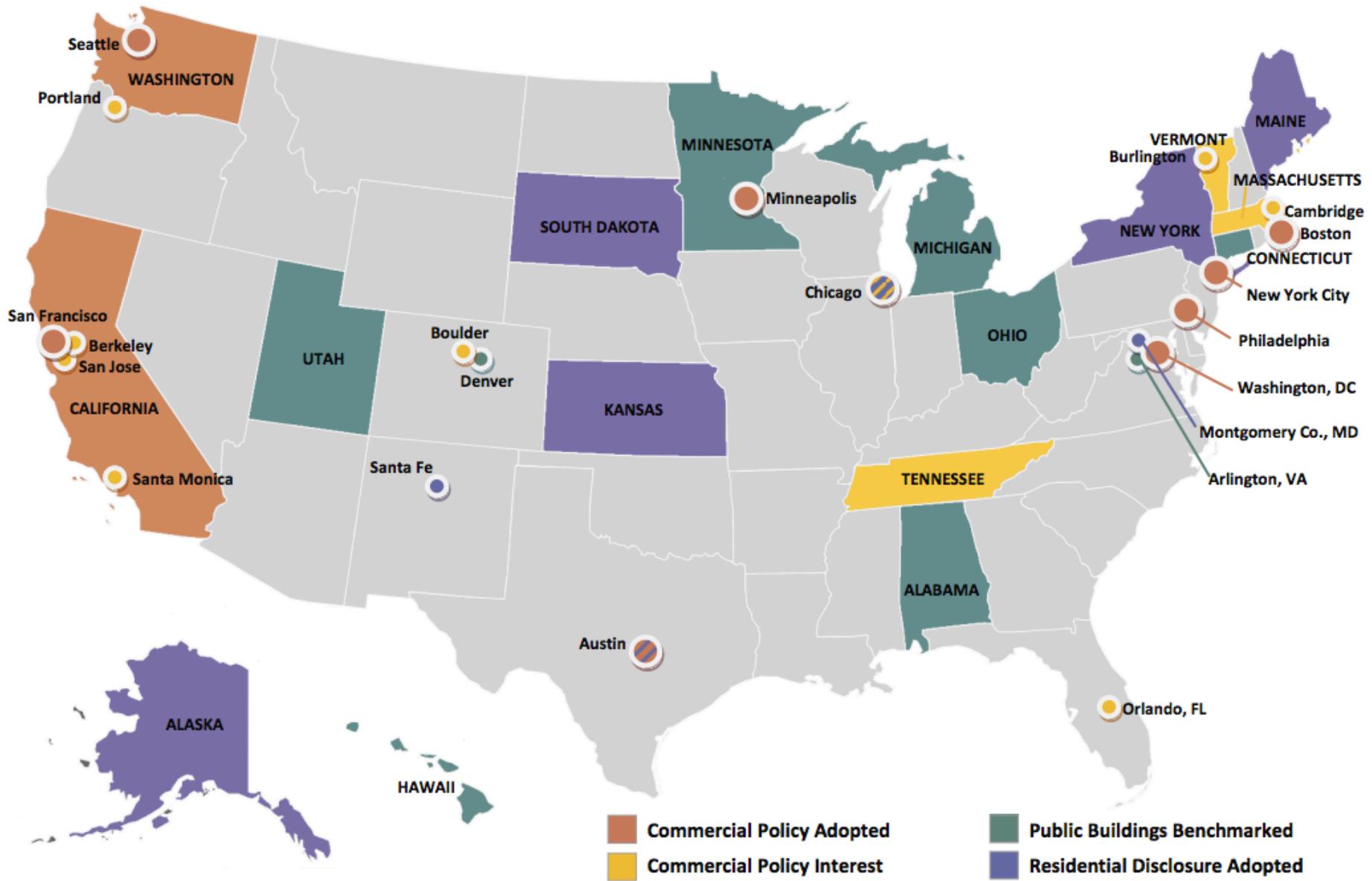
Published: May 19, 2010

There are plenty of buildings that boast of being “built to LEED standards” but might then leak excessive carbon once in operation. With so many developers, municipalities and federal agencies embracing environmental awareness, the government needs to make sure that a green building doesn't go gray after its grand opening.

Energy Benchmarking Policies (selected)

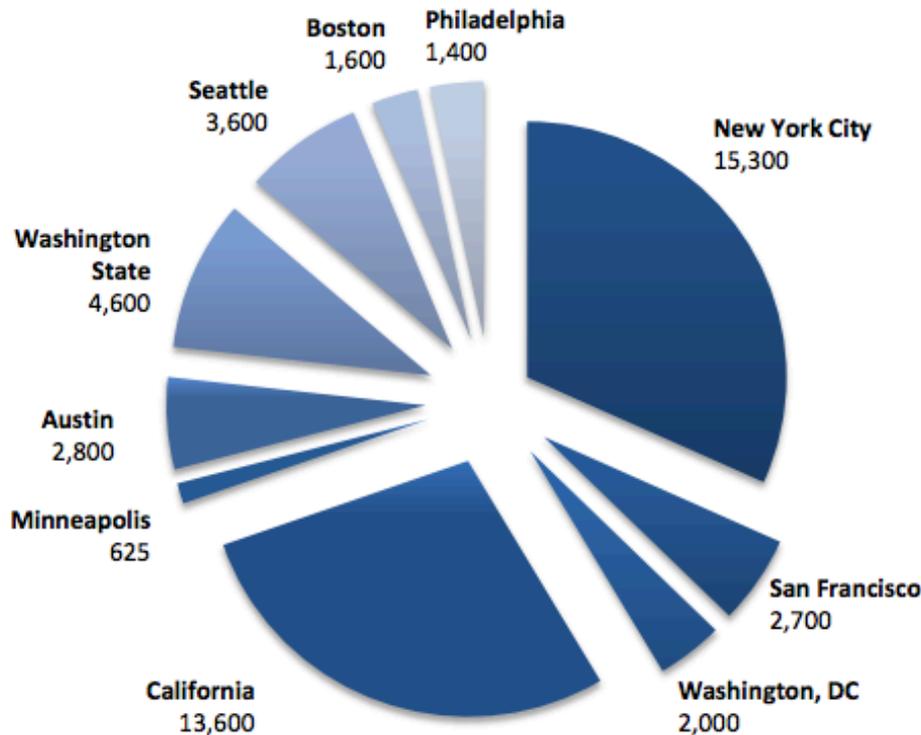
- **California**
 - AB1103 requires benchmarking of all commercial buildings at time of lease or sale.
 - Executive order S-20-04 requires benchmarking of all state buildings.
 - SB1 requires buildings applying for solar incentives to benchmark energy use intensity.
- **Federal**
 - EISA 2007 requires benchmarking of federal buildings to track performance of energy goals.
- **Disclosure laws**
 - Austin, NYC, Philadelphia, San Francisco, Seattle, Wash DC,
 - Many organizations requiring LEED-EB, which requires benchmarking.
- **Europe**
 - Energy Performance of Buildings Directive requires energy performance to be publicly displayed.

U.S. Building Benchmarking and Disclosure Policies



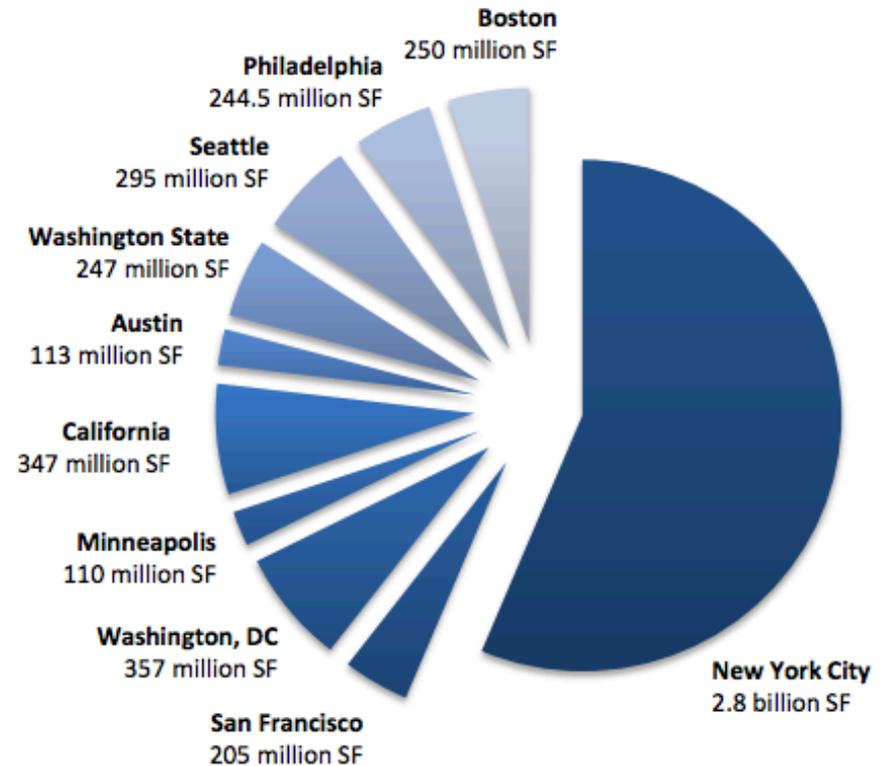
U.S. Building Benchmarking and Disclosure Policies

NUMBER OF PROPERTIES COVERED ANNUALLY



Source: IMT

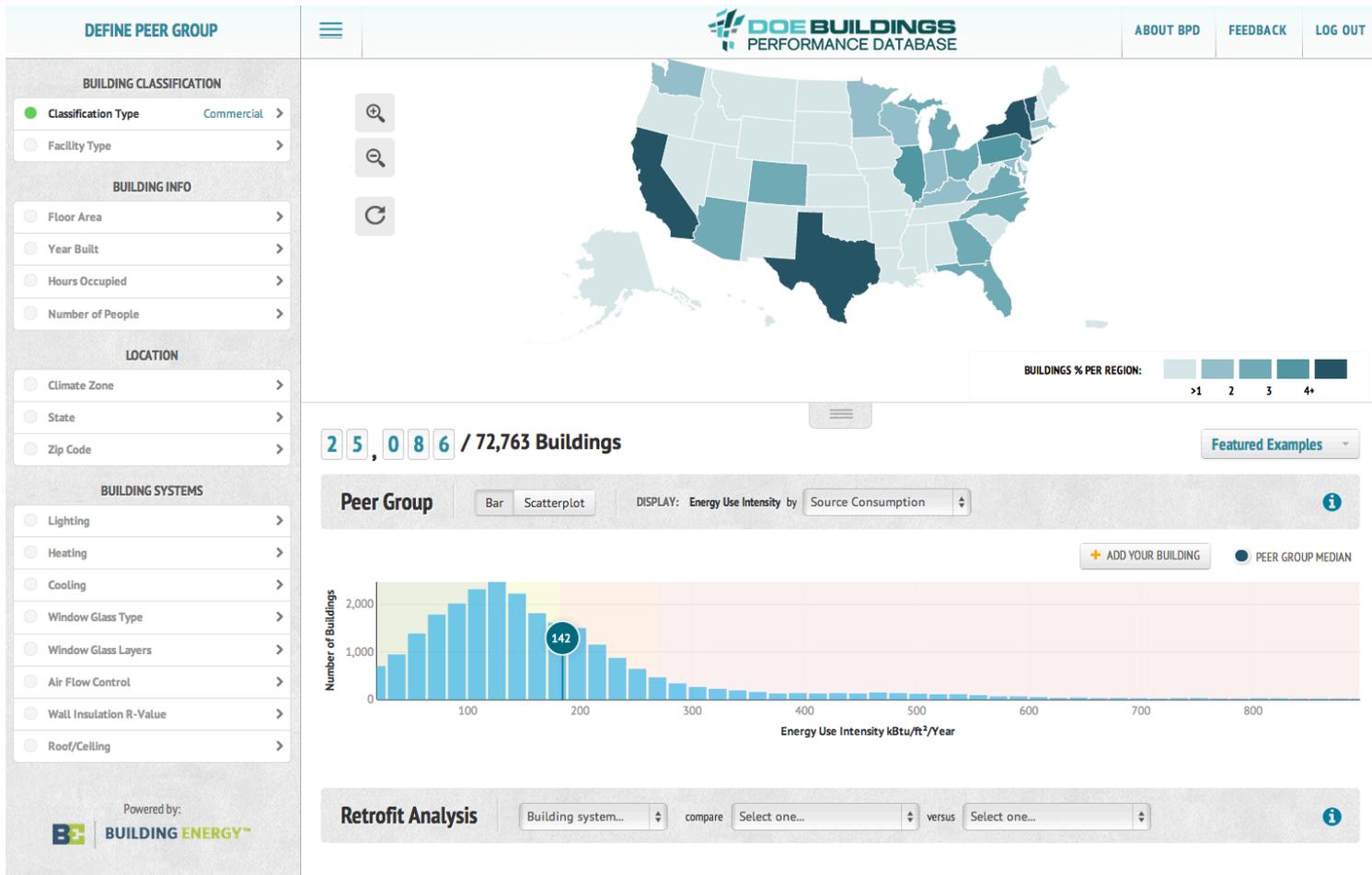
BUILDING AREA (IN SQUARE FEET) COVERED ANNUALLY



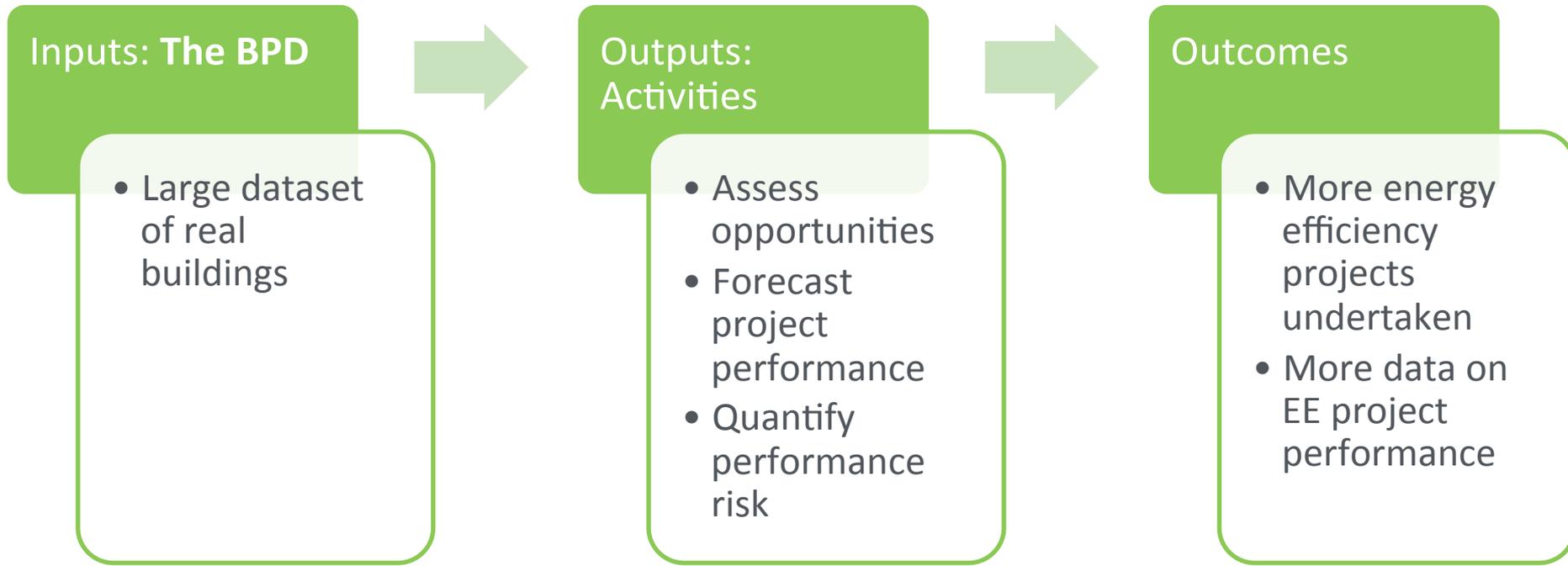
Source: IMT

The Buildings Performance Database

The BPD is the largest publicly-available dataset of information about the energy performance of real commercial and residential buildings.



BPD unlocks the power of building energy performance data.



Design Principles

- The BPD contains *actual data* on existing buildings - not modeled data or anecdotal evidence.
- The BPD enables statistical analysis without revealing information about individual buildings.
- The BPD cleanses and validates data from many sources and translates it into a standard format.
- In addition to the BPD's analysis tools, third parties will be able to create applications using the database.

Assess opportunities

- Identify high or low performing buildings, and identify improvements that will likely have a significant savings impact

Understand performance risk

- Analyze the range of likely returns from an investment

Evaluate investment performance

- Compare efficiency project performance to similar projects

Assess opportunities

- Identify high or low performing buildings, and identify improvements that will likely have a significant savings impact

Understand performance risk

- Analyze the range of likely returns from an investment

Evaluate investment performance

- Compare efficiency project performance to similar projects

Influence local real estate markets

- Enable public access to general statistical information about buildings, without sharing building-level information

Help participants assess opportunities

- Help building owners, managers, and contractors identify improvements that will likely have a significant savings impact

Target program design

- Identify buildings and efficiency measures with the greatest savings potential

Support M&V

- Optimize M&V requirements based on measured savings uncertainty and persistence

Assess opportunities

- Identify high or low performing buildings, and identify improvements that will likely have a significant savings impact

Increase confidence in returns

- Analyze actual building performance (as opposed to modeled or predicted performance)

Conduct performance risk analysis

- Quantitatively distinguish between expected returns and performance risk

Support portfolio-level investment strategy

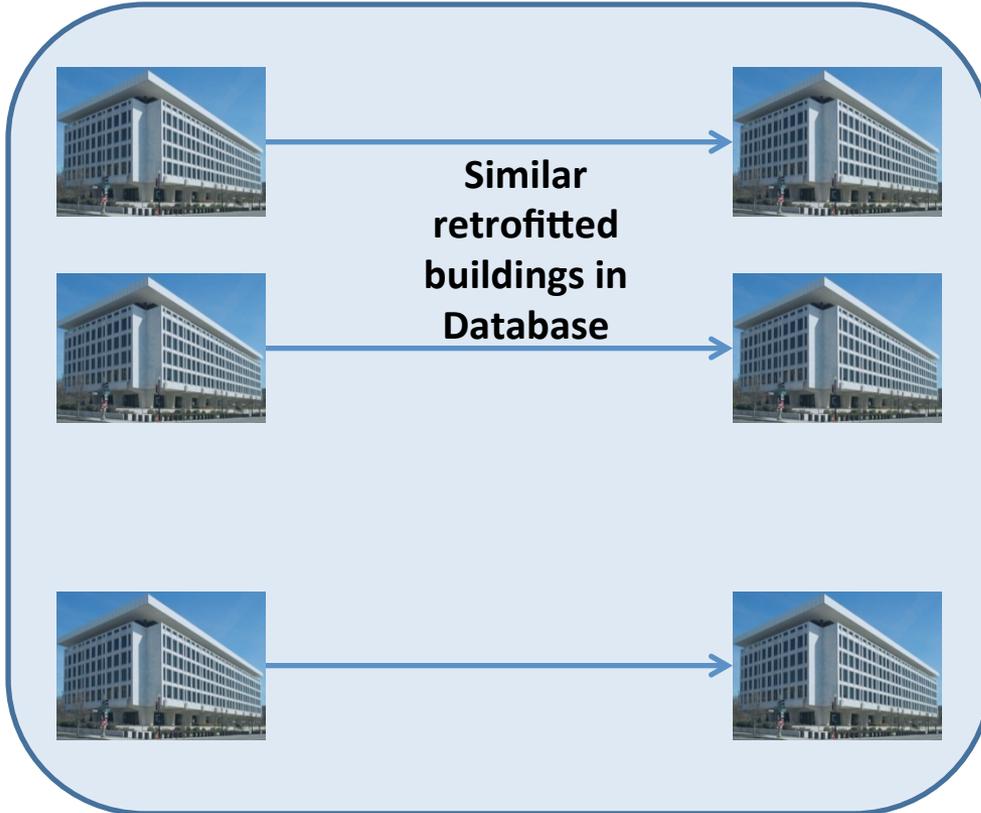
- Diversify risk by investing in a range of buildings and measures

Conventional Method for Retrofit Savings

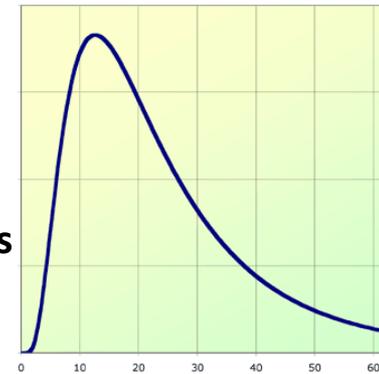
**My building:
now**



**My building:
after retrofit**



**Number of
Buildings**



Savings

“Actuarial” Method of Savings Analysis

My building:
now



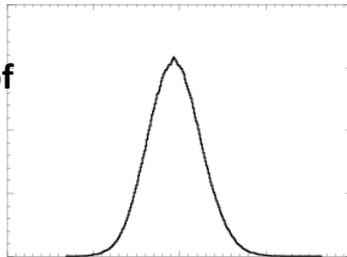
My building:
after retrofit



Peer group in
Database

Peer group in
Database

Number of
Buildings



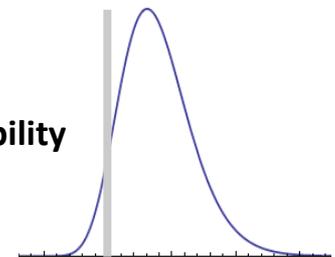
Energy/SqFt

Number of
Buildings



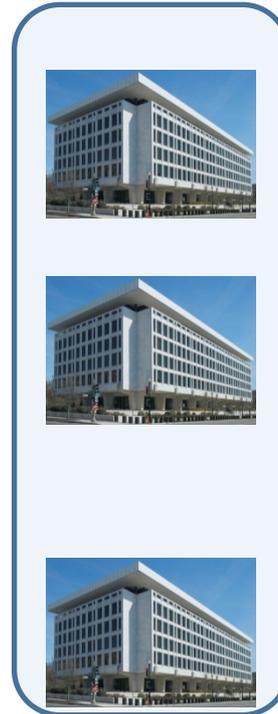
Energy/SqFt

%
Probability



Savings/SqFt

“Subtract” the two
histograms to get savings
histogram



Florida Homes

Building Count by Source kBtu/SF/yr

DEFINE PEER GROUP

BUILDING CLASSIFICATION

Classification Type 1 2 >

Facility Type All >

BUILDING INFO

Floor Area 0 - 2,000,000 >

Year Built 1800 - 2013 >

Hours Occupied 0 - 168 >

Number of People 0 - 1,000 >

LOCATION

Climate Zone All >

State 1 52 >

Zip Code All >

BUILDING SYSTEMS

Lighting All >

Heating All >

Cooling All >

Window Glass Type All >

Window Glass Layers All >

Air Flow Control All >

Wall Insulation R-Value 0 - 80 >

Roof/Ceiling All >

ABOUT BPD FEEDBACK LOG OUT

Buildings % Per Region: >1 2 3 4+

33,315 / 66,954 Buildings

Peer Group Bar Scatterplot DISPLAY: Energy Use Intensity by Source Consumption

+ ADD YOUR BUILDING PEER GROUP MEDIAN

Peer Group Tool

Washington DC Benchmarking Data

Office Buildings <1M SF, built since 1900

Source Consumption by Gross SF

3 4 8 / 66,954 Buildings

Featured Examples

Peer Group

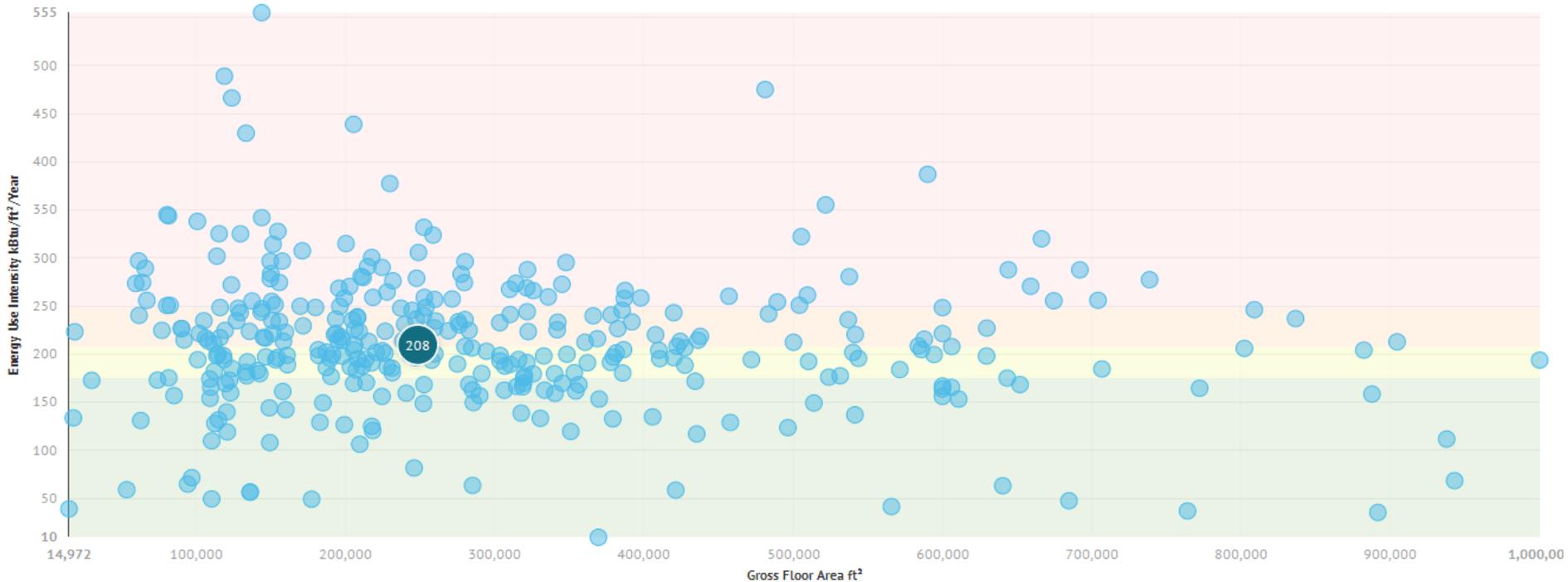
Bar Scatterplot

DISPLAY: Energy Use Intensity by Source Consumption and Gross Floor Area



+ ADD YOUR BUILDING

● PEER GROUP MEDIAN



Peer Group Tool

Washington DC Benchmarking Data

Office Buildings <1M SF, built since 1900

Source Consumption by Year Built

3 4 8 / 66,954 Buildings

Featured Examples

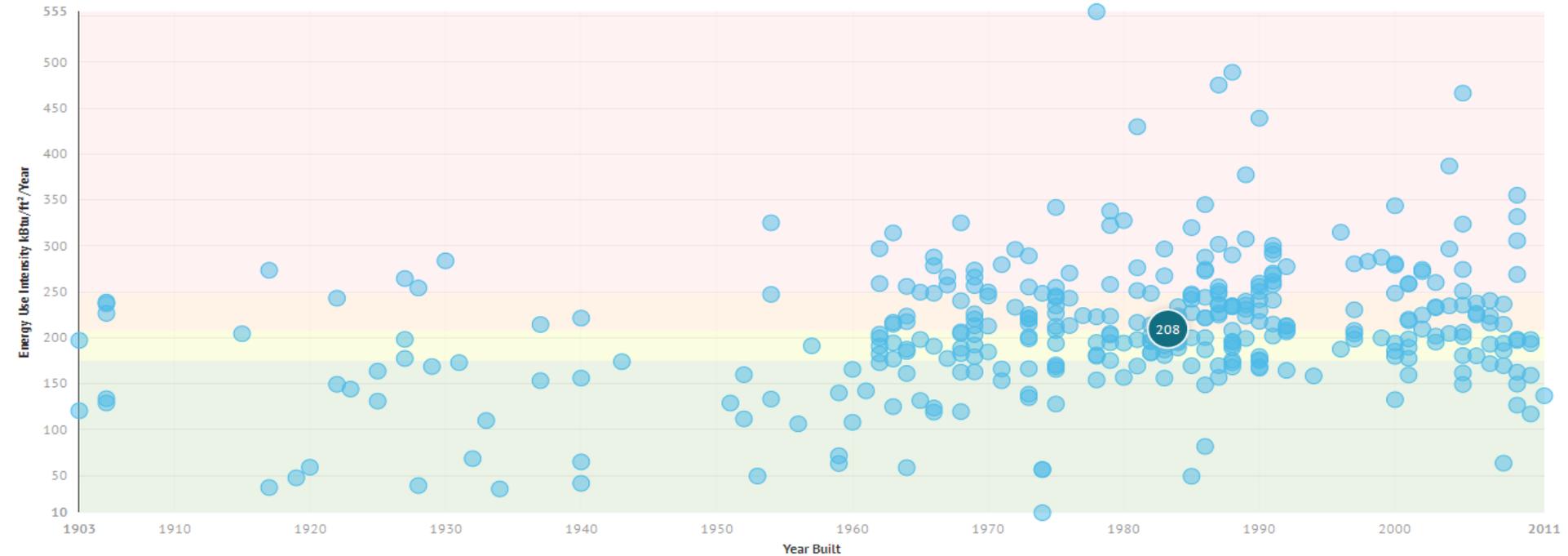
Peer Group

Bar Scatterplot

DISPLAY: Energy Use Intensity by Source Consumption and Year Built



+ ADD YOUR BUILDING ● PEER GROUP MEDIAN



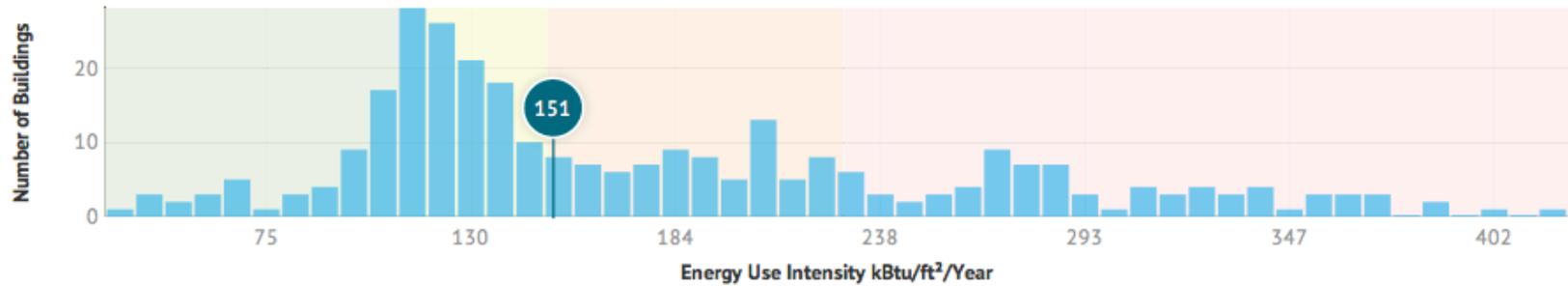
19 Note: Does not yet include Federally owned buildings

CA Retail Buildings >50,000 SF; N = 320

Compare Packaged Direct Expansion to Air Source Heat Pump

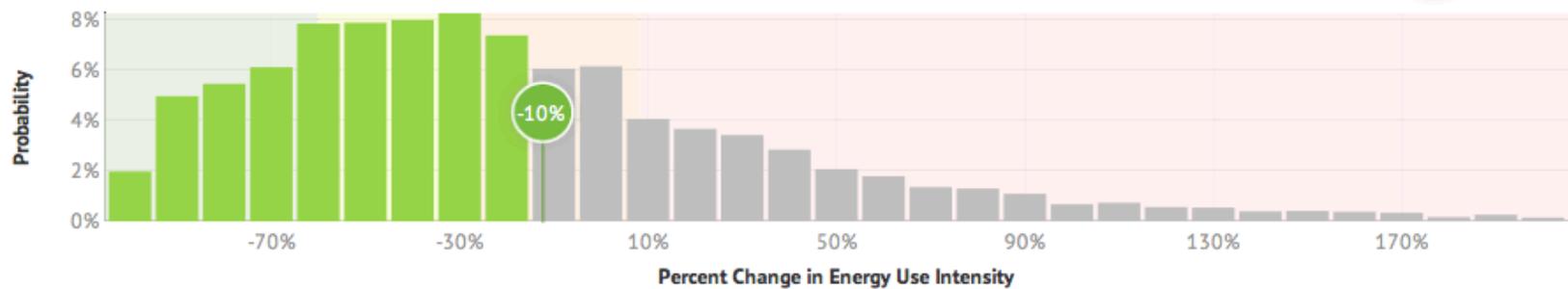
Peer Group | | DISPLAY: Energy Use Intensity by

PEER GROUP MEDIAN



Retrofit Analysis |

58% PROBABILITY THAT EUI WILL DECREASE **10%** OR MORE

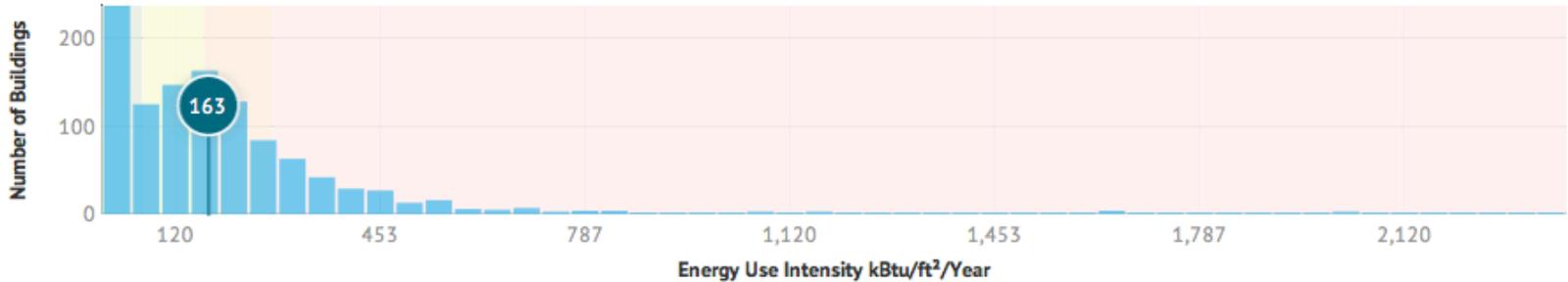


Office Buildings; N = 2,022

Compare Packaged Hot Water Boiler to Air Source Heat Pump

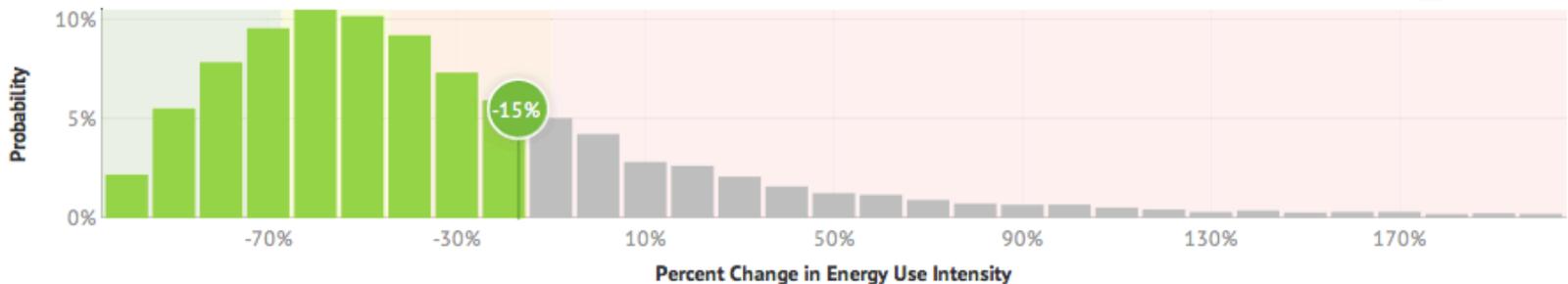
Peer Group | Bar Scatterplot | DISPLAY: Energy Use Intensity by Source Consumption + i

+ ADD YOUR BUILDING ● PEER GROUP MEDIAN



Retrofit Analysis | Heating compare Boiler - Hot Water versus Heat Pump - Air Source + i

68% PROBABILITY THAT EUI WILL DECREASE **15%** OR MORE



- >100,000 buildings, from both public and private datasets.
- More datasets are being added regularly. There is no upper limit for the number of buildings the BPD can hold.
- More analyses will become possible as data is added.

Public Sector

U.S. Energy Information Administration
U.S. General Services Administration
U.S. Environmental Protection Agency
New York City Dept. of Citywide Administrative Services
Pennsylvania Keystone HELP Home Energy Loan Program
San Francisco Department of the Environment
State of California Public Utilities Commission
State of California Energy Commission
University of Arizona
University of Dayton
District Department of the Environment: Washington, DC
Vermont Energy Investment Corporation
Virginia Beach City Public Schools

Private Sector

Brandywine Realty Trust
Connexion Asset Group
Kohl's
EnergyIT.com
Liberty Property Trust
Lucid Design Group
Prudential
Related
Tishman Speyer
Transwestern
USAA
Vornado Realty Trust

Data Mapping Tool

Map your data for sdge_commercial_2.csv

Field Mappings					Source Data			
Import	Table	Field	← Mapping Confidence →	Spreadsheet header	Row 1	Row 2	Row	
1	<input checked="" type="checkbox"/>	CommercialFacility	source_facility_id	High	Source Facility ID	34096	34244	351
2	<input checked="" type="checkbox"/>	CommercialFacility	complete_total_energy	High	Complete Total Energy	TRUE	TRUE	TRU
3	<input checked="" type="checkbox"/>	CommercialFacility	primary_facility_type	High	Primary Facility Type	Store Buil	Light Indu	Store
4	<input checked="" type="checkbox"/>	CommercialFacility	year_completed	Medium	Year Complete		1945	
5	<input checked="" type="checkbox"/>	CommercialFacility	gross_floor_area	High	Gross Floor Area	3448	811	112
6	<input checked="" type="checkbox"/>	CommercialFacility	num_floors	High	Num_Floors	2	1	1
7	<input checked="" type="checkbox"/>	CommercialFacility	num_activity_areas	High	Number of Activity Areas	1	1	1
8	<input checked="" type="checkbox"/>	CommercialFacility	facility_num	High	Facility Number	1	1	1

BPD data fields

Mapping Confidence
(based on learning
algorithm)

Source data fields

Data Cleansing Rules

Check for:

- Data type
- Allowed values
- In-range validation

Significant effort!

Under continuous development...

Field	Data Type	Allowed Values	In-range check 1	In-range check 2
Source Facility ID	ALPHANUMERIC		Match Source Facility ID in Site table	
Year Completed	INTEGER(4)	1600 - present		
Year Occupied	INTEGER(4)	1600 - present	>= Year Completed	
Owner Type	CONSTRAINED LIST	List		
Percent Occupied by Owner	DOUBLE	0 - 100		
Operator Type	CONSTRAINED LIST	List	If Percent Occupied by Owner = 100, then Operator Type = "Owner"	
Surroundings	CONSTRAINED LIST	List		
Orientation	CONSTRAINED LIST	List		
Building Footprint Area	DOUBLE	100 - 2,000,000	<= Gross Floor Area	If Number of Floors = 1, then Building Footprint Area = Gross Floor Area.
Footprint Shape	CONSTRAINED LIST	List		
Perimeter	DOUBLE	40 - 30,000		
Gross Floor Area	DOUBLE	100 - 7,000,000		
Net Floor Area	DOUBLE	100 - 7,000,000	<= Gross Floor Area	
Rentable Floor Area	DOUBLE	0 - 7,000,000	<= Gross Floor Area	
Occupied Floor Area	DOUBLE	0 - 7,000,000	<= Gross Floor Area	
Lighted Floor Area	DOUBLE	0 - 7,000,000	<= Gross Floor Area	
Heated Floor Area	DOUBLE	0 - 7,000,000	<= Gross Floor Area	>= Sum of Heated Floor Area in all Activity Areas
Cooled Floor Area	DOUBLE	0 - 7,000,000	<= Gross Floor Area	>= Sum of Cooled Floor Area in all Activity Areas

Check out the BPD: [Buildings.energy.gov/BPD](https://buildings.energy.gov/BPD)



- At: BuildingsPerformanceDatabase@ee.doe.gov

- The BPD staff can help you identify datasets that contain the minimum and optional fields.

- You can share your Portfolio Manager account, or transfer files via an FTP site or email.
- The BPD accepts all electronic file formats including .csv, .xls and .accdb.
- Include the words "PROPRIETARY DATA" in the file name or email body.

- The BPD team will reformat, cleanse and anonymize your data before entering it into the BPD.
- You may request a cleansed copy of your own dataset.

- The BPD team will let you know when your data has been uploaded!
- BPD data is stored under stringent privacy and security protocols.

Contact:

BuildingsPerformanceDatabase@ee.doe.gov

BPD is seeking datasets of >50 buildings that include:

- **Required Fields**

- ✓ *Basic Building Characteristics*

- City, State, Zip Code
- Usage type (office, retail, home)
- Building floor area
- Year completed
- Electricity/fuel use for at least one year

- **Optional Fields**

- ✓ *Detailed Building Characteristics*

Operational information (Portfolio Manager data),
such as:

- Types of activities and associated floor area
- Operating hours
- Number of occupants

Equipment & Asset information, such as:

- Lighting type and controls
- Air distribution configuration, controls, etc
- Heating and cooling equipment types & efficiencies
- Hot water equipment type & efficiency
- Wall, roof and window characteristics

6,000 users since June 2013 launch

Building
Owners &
Managers

Contractors &
Software
Developers

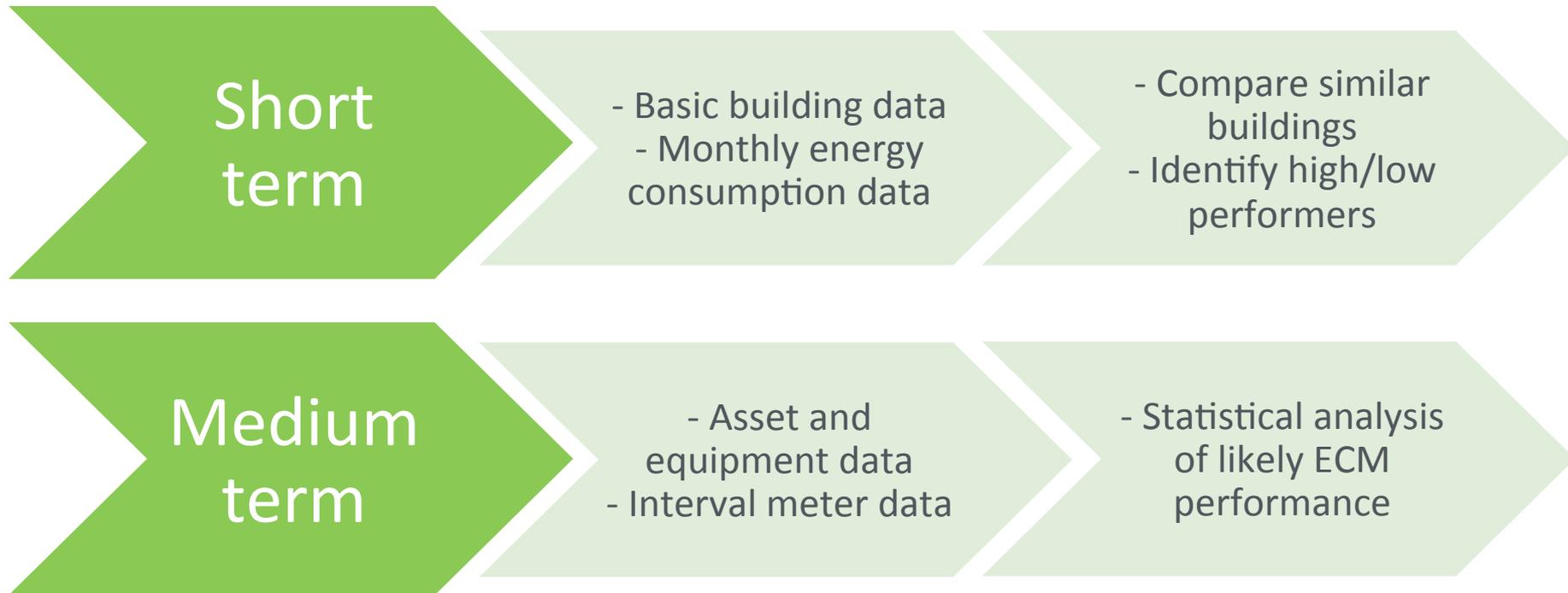
Financial
Institutions

Public
Agencies

EE Program
Administrators

Research
Institutions

- The BPD demonstrates the value of aggregating the kind of data that is commonly collected today.
- As stakeholders begin to collect and contribute richer data, the BPD will support more advanced analysis.



What to Expect in FY14

- More data (and improved data quality)
- New features
 - Linear regression analysis
 - Weather normalization
 - More filters
 - “Advanced user” analysis features
- Case studies
- Public API

Big Data for Building Energy: Key R&D challenges: Data

- Cost-effective ways of collecting characteristics data (system types, operations schedules, etc.).
- Cost-effective sub-metering of energy use.
- Data cleansing protocols and methods.
- Data standardization, especially characteristics data.

Big Data for Building Energy: Key R&D challenges: Analysis

- Data analysis and visualization approaches optimized to the building energy context.
- “Value of information” approach to prioritize data collection and analysis.
- Deeper (more rigorous) understanding of how decision-makers in the building industry comprehend and use energy information.

Questions?



Paul Mathew
(510) 486-5116
pamathew@lbl.gov