

nielsen  
.....

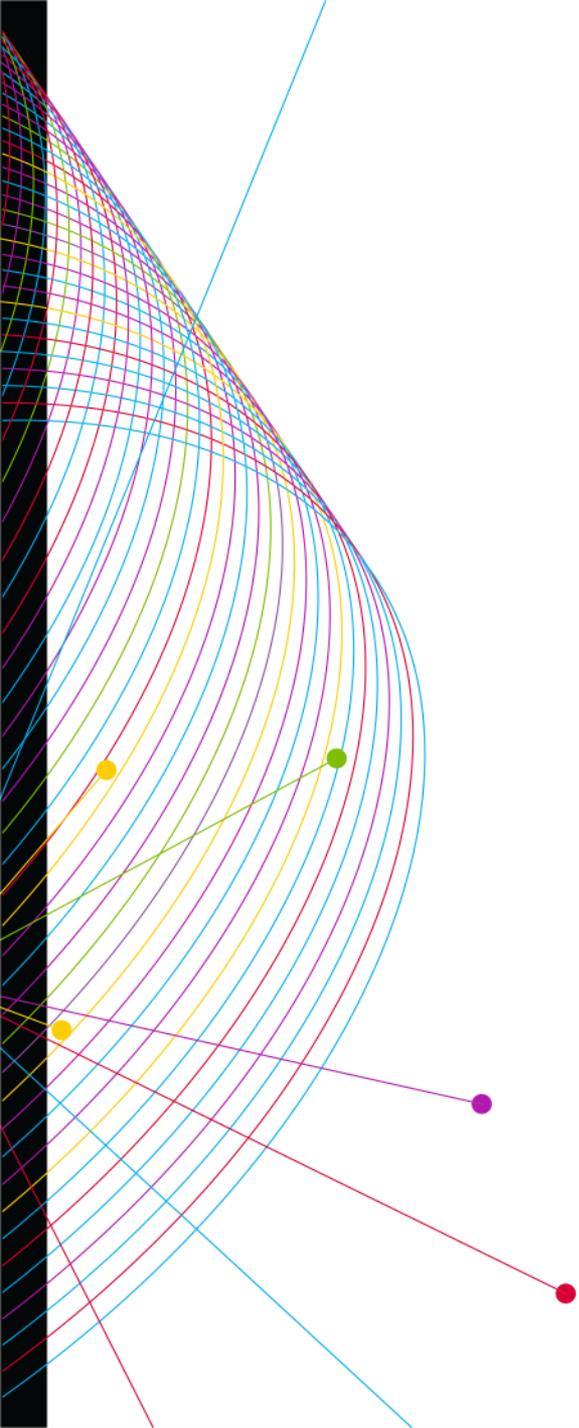
# MODERN CONSUMER PANEL RESEARCH

TRENDS, TECHNIQUES, AND INNOVATION IN  
CONSUMER MEASUREMENT

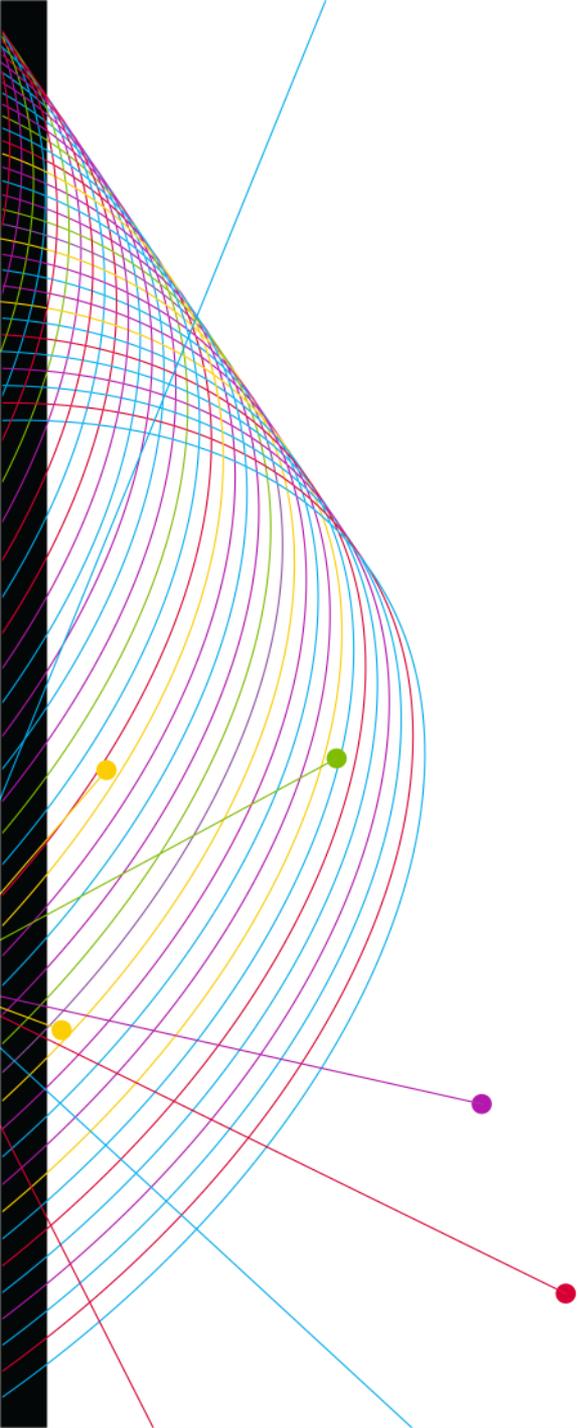
Ben Wood  
October 9, 2015

# OUTLINE

1. Introduction to Nielsen
2. Panel Research Challenges
3. Innovation in behavioral projection
4. Big Data with Panels
5. Using a truth source to evaluate quality
6. Calibrating and completing third party data

An abstract graphic on the left side of the slide. It features a vertical black bar on the far left. To its right, a series of curved, overlapping lines in various colors (blue, green, yellow, red, purple) form a spherical shape. Several colored dots (yellow, green, purple, red) are placed on these lines, with thin lines extending from them towards the right side of the slide.

# INTRODUCTION TO NIELSEN



## OUR MISSION

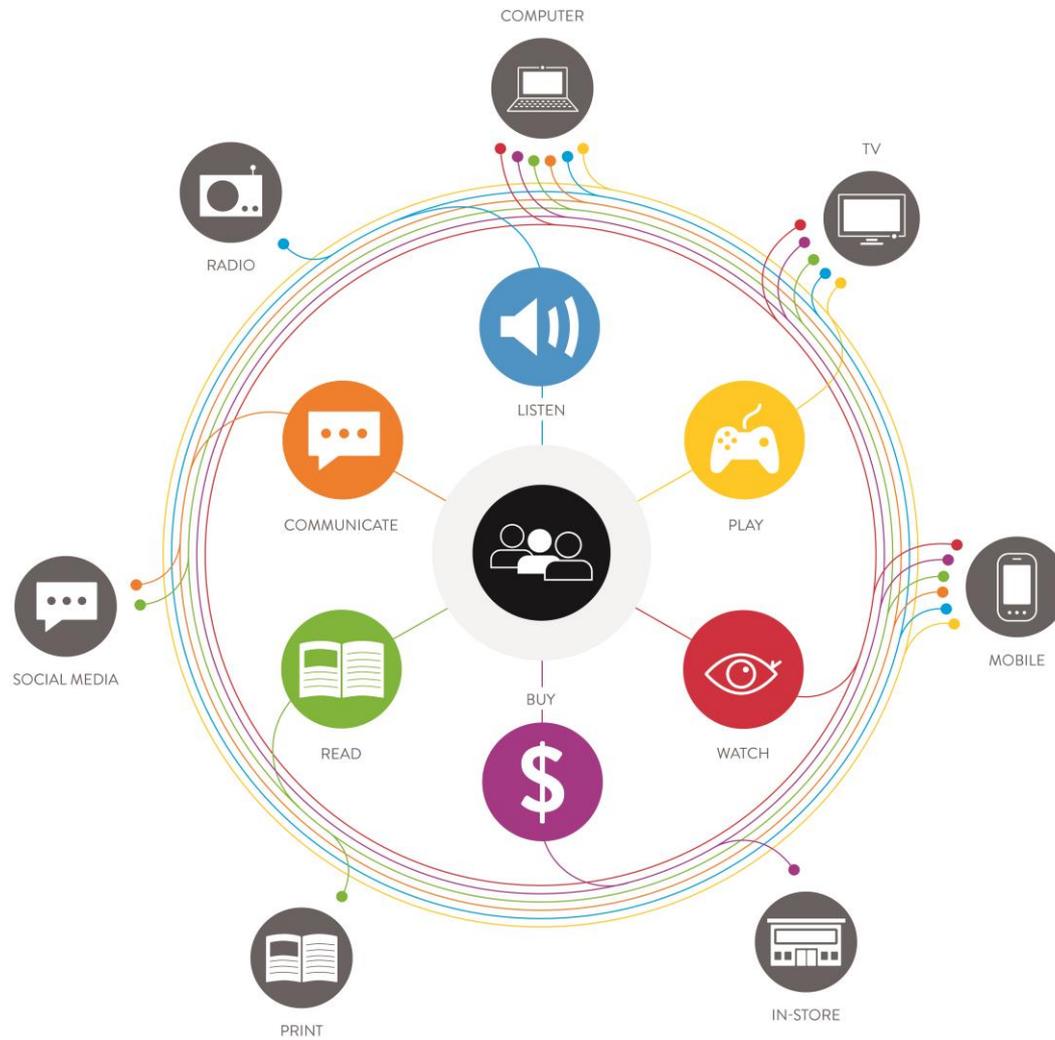
The most complete understanding of consumers worldwide.

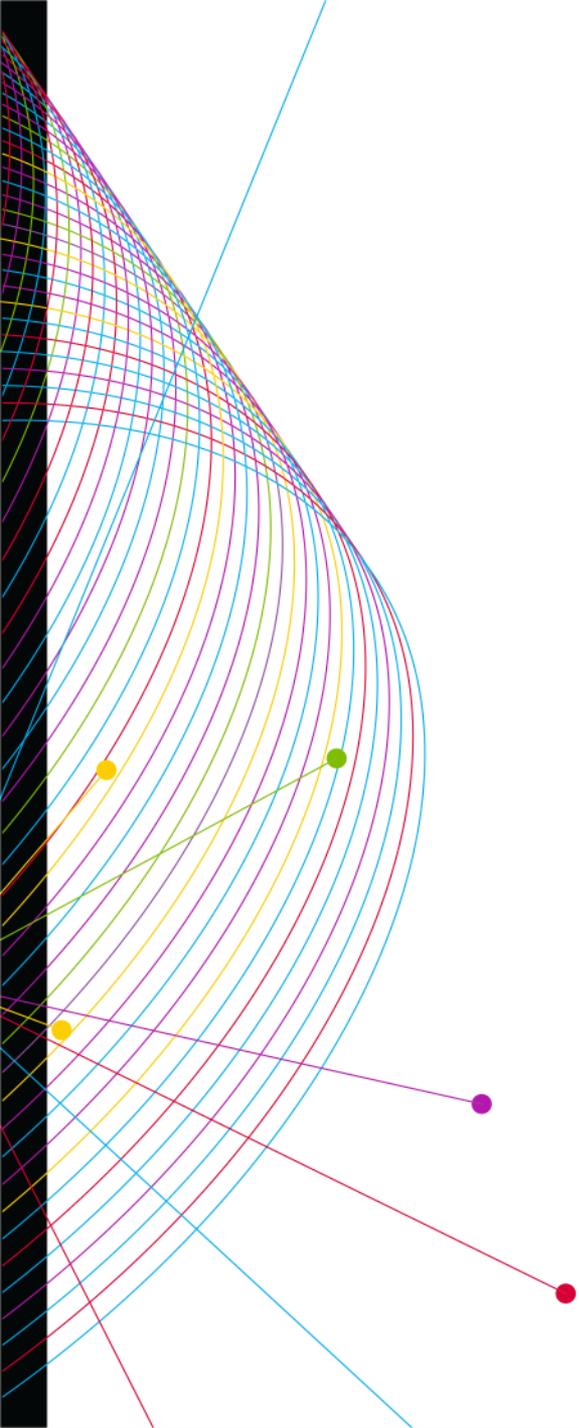
## OUR PROMISE

An uncommon sense of the consumer for faster, smarter, more confident decisions that *drive growth*

# FROM DATA TO MEASUREMENT

*Simplifying the Complex*



An abstract graphic on the left side of the slide. It features a vertical black bar on the far left. From this bar, a dense grid of thin, curved lines in various colors (blue, green, yellow, orange, red, purple) extends to the right. The lines are more densely packed on the left and become more sparse as they move right. Several lines end in small colored dots (yellow, green, purple, red) that are scattered across the white background.

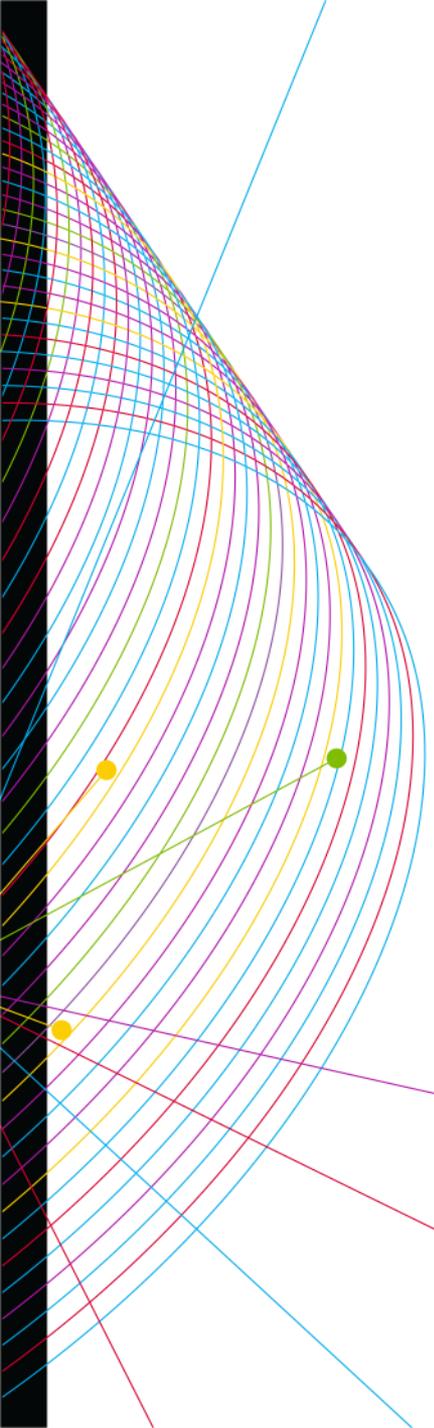
# PANEL RESEARCH CHALLENGES

# TRADITIONAL PANEL MEASUREMENT

Facing pressure from several factors – Trying to measure MORE while maintaining statistical reliability



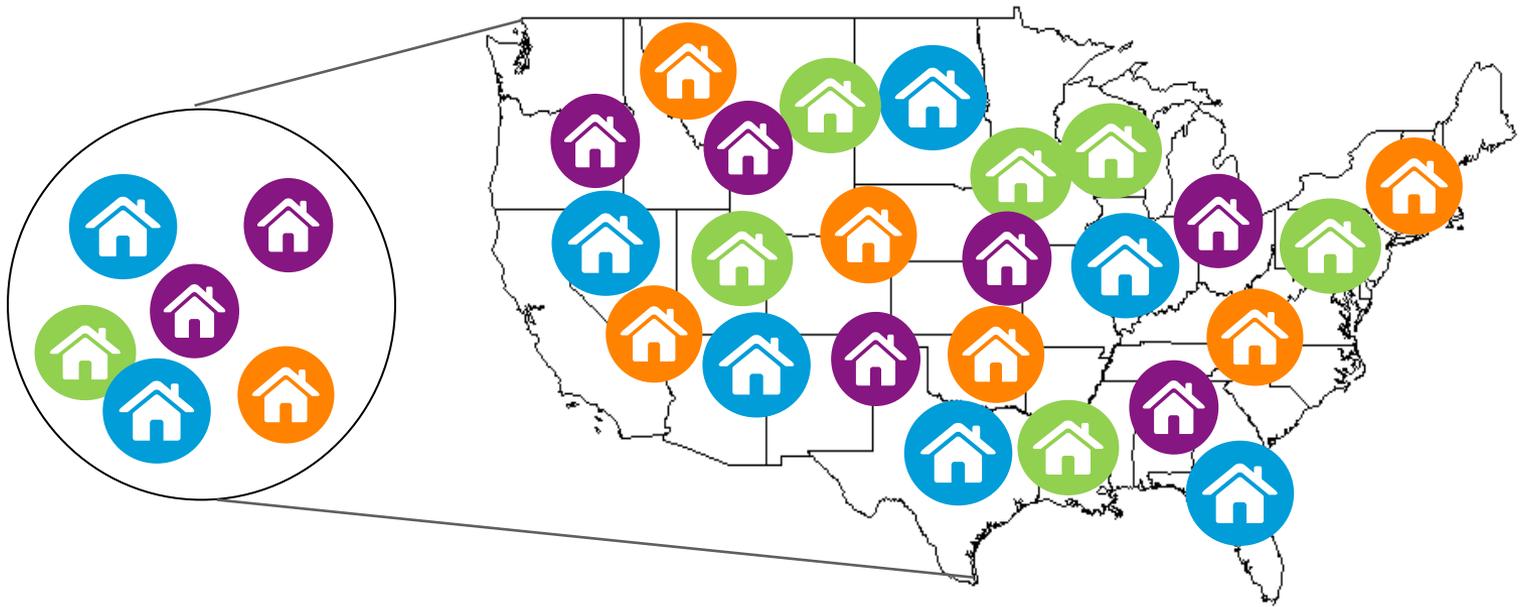
1. Declining cooperation and participation rates
2. Challenges in obtaining true probability samples
3. Fragmentation of behavior



# INNOVATING PANEL PROJECTIONS

# PROJECTION BASICS

Project sample households to a universe of households



# PROJECTION DESIGN

Demographically and Geographically represented



- Demographically
  - HH Size, Income, Head Age, Female Head Education, Male Head Education, Race, Hispanic, Head Occupation, Presence of Children
- Geographically
  - Total US, Major Market, Census Regions

# MISSING BEHAVIOR

2,000 panelists  
Population of 2,000,000  
4 Coca-Cola purchases/panelist  
=8,000,000 Coca-Cola units

10,000,000 units of Coca-Cola sold

Data from panelists  
showing **what** they  
bought, **when** they bought  
it, and from **where**

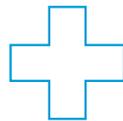


Data from stores showing  
how much they sold

# SCAN, GRANULARITY, AND TRIP PROJECTION



PANEL DATA



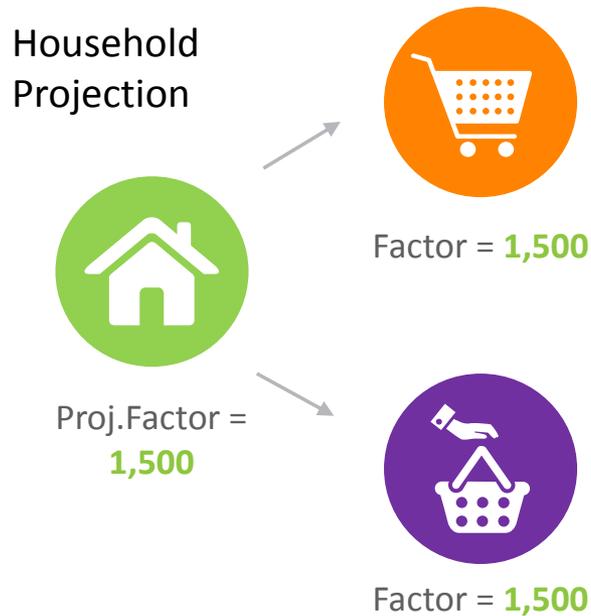
SCAN  
TARGETS



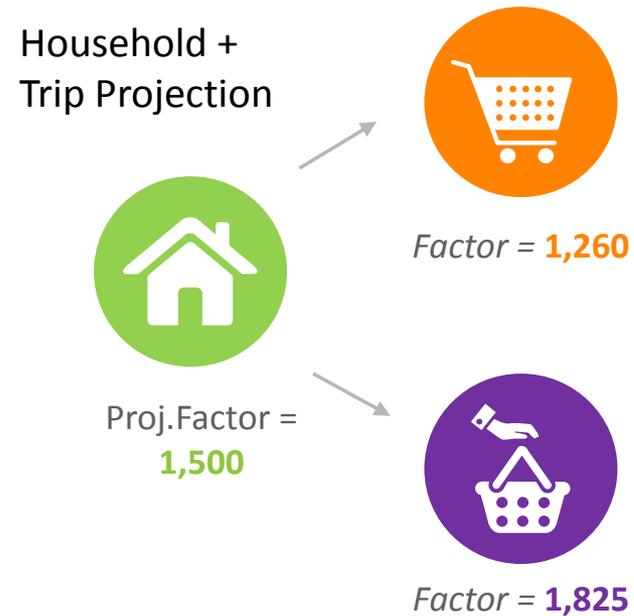
TRIP FACTORS

# GRANULARITY

## FROM

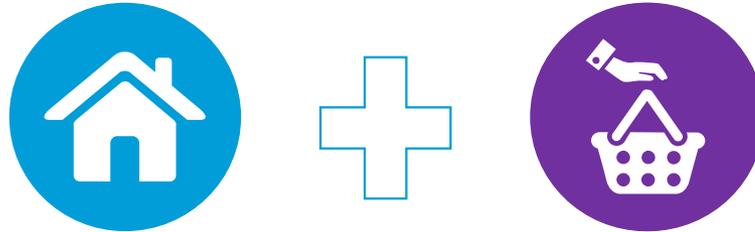


## TO

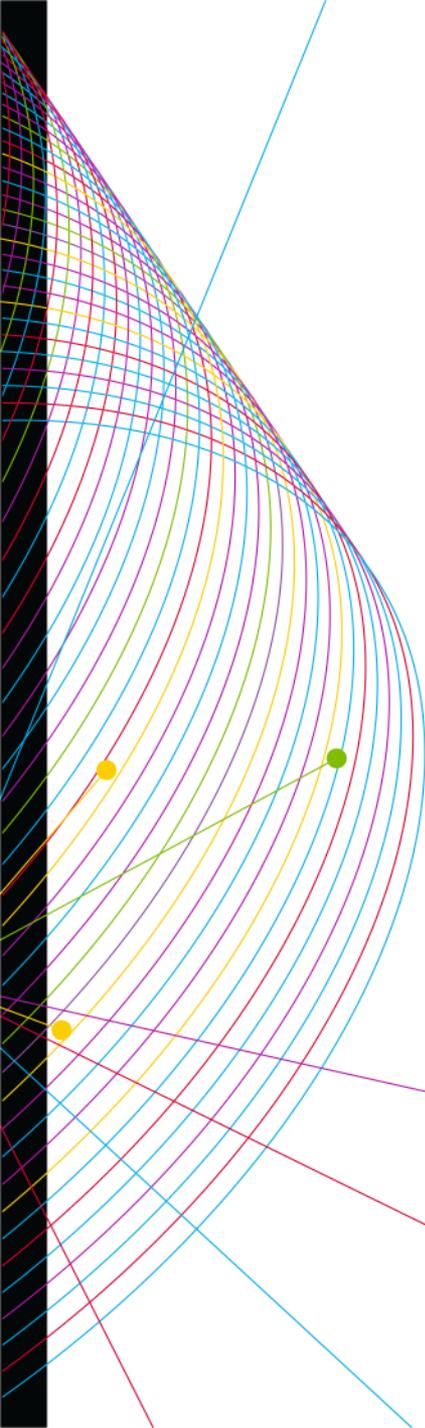


Using each individual trips as the basis to create projection factors significantly increases flexibility to align with scan

# TRIP PROJECTION



- Merging two concepts
  - Dual projection design; Household factor AND Trip factor
    - One household factor per household, one trip factor per trip
    - Creating a methodology that continues to utilize household factors helps in keeping the demographic integrity of our insights
  - Remaining question - are additional purchases allocated to panel homes, or missing homes? If purchases are projected downward, which panelists do they come from?

An abstract graphic on the left side of the slide. It features a vertical black bar on the far left. From this bar, a dense grid of thin, curved lines in various colors (blue, green, yellow, orange, red, purple) extends to the right, creating a funnel-like shape that narrows as it moves away from the bar. Several lines end in small colored dots (yellow, green, purple, red) that are scattered across the lower half of the grid.

# THE FUTURE: BIG DATA WITH PANELS

# TRADITIONAL PANEL MEASUREMENT

Organic big data becoming more viable for measurement

## -Explosion of data collected-

Facebook alone has over 30 petabytes of user data generated through user behavior \*

## -Storage and processing availability-

Cost of 1GB on April 1, 2000 - \$14.90  
Cost of 1GB on March 4, 2015 \$0.03 \*\*

## -New analytic technology-

Cluster computing enables processing orders of magnitude faster \*\*\*

\*<http://www.ibmbigdatahub.com/gallery/quick-facts-and-stats-big-data>

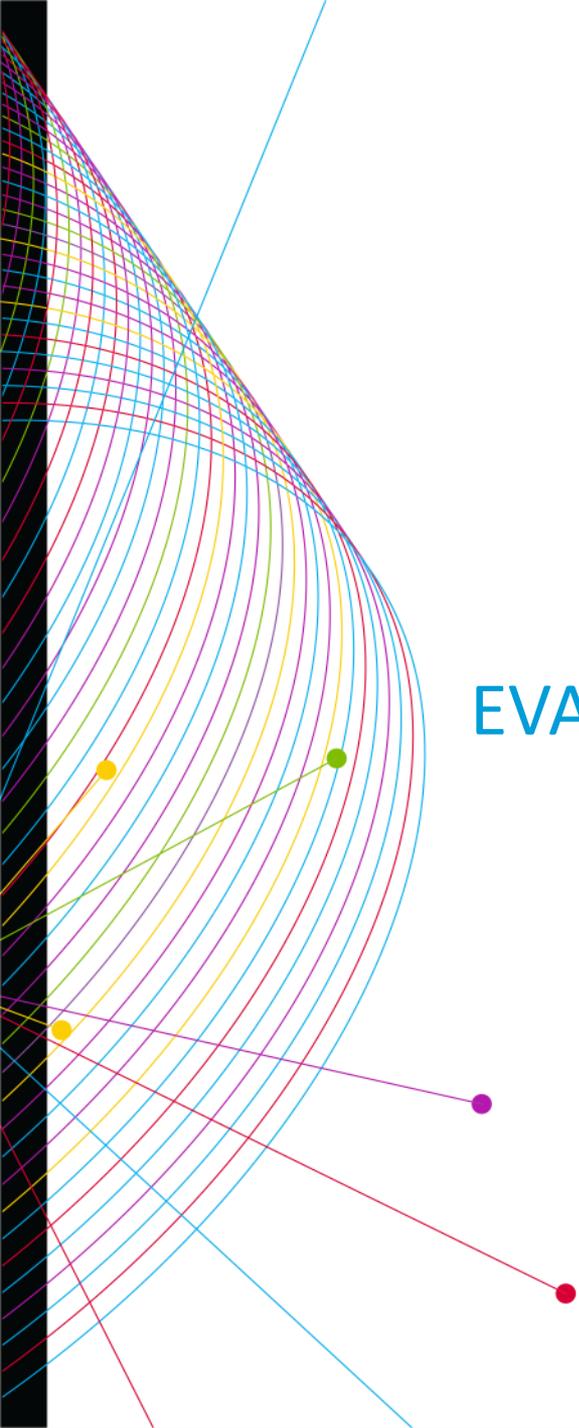
\*\*<http://www.mkomo.com/cost-per-gigabyte-update>

\*\*\*<http://spark.apache.org/>

# USING BIG DATA

1. Cost of collection: Many organic panelists may be cheaper than a single sampled panelist
2. Coverage: hard-to-reach respondents can be measured organically
3. Stability: larger data sources do not see a great impact from change in a few respondents or panelists
4. Projectability: confidence increases with larger panels

1. Representativeness: Many big data sources are self-selected, and may not represent the total population of interest
2. Control: The method of data collection is often not controlled at the point of collection
3. Comparability: The data itself may not be comparable with data used for measurement purposes
4. Accuracy: Data may not be accurate due to missing controls, sources of truth, etc.

An abstract graphic on the left side of the slide. It features a vertical black bar on the far left. From this bar, a series of curved, overlapping lines in various colors (blue, green, yellow, orange, red, purple) fan out towards the right. Several colored dots (yellow, green, purple, red) are placed at various points along these lines, with thin lines extending from them towards the right edge of the graphic area.

# EVALUATING DATA QUALITY USING A PANEL TRUTH SOURCE

# INSIGHT INFLECTION

Critical points in using third party data

- When evaluating quality of a third party data, it's critical to balance individual and aggregate metrics, keeping in mind that the purpose of the data is to help make a decision
- The point where accuracy of a data source is unsuitable is the point where it informs a different action than correct data

# INDEX

Provides an Idea of Aggregate Level Accuracy

<u>HOH Education</u>	Positive NPM HHs	%	Positive 3rd Party HHs	%	Index: 3rd Party/Nielsen
Less than High School Diploma	1,492	7.5	2,735	13.7	184
4 Years High School	4,376	21.9	5,832	29.2	133
1-3 Years College	6,833	34.2	5,367	26.9	79
4+ Years College	7,285	36.5	6,027	30.2	83

3rd Party skews  
toward less  
educated  
homes

How does the distribution of the vendor's data compare with the Nielsen truth?

# SENSITIVITY

Shows Us if the Vendor is Capturing the Traits We Need

<u>Presence of Children</u>	Positive NPM Households	Positive 3rd Party Households	True Positive	True Positive Accuracy	Sensitivity: True Positive/Positive NPM Households
Any Children	6,850	4,733	3,348	70.7	48.9
No Children	13,136	15,253	11,751	77.0	89.5

Although the accuracy of the any children identifier is high, 3rd Party isn't capturing over half of families with children

Of all homes with a particular trait, what percentage did the vendor correctly capture?

# BIN DISTANCE AS AN ACCURACY METRIC

HH Size:		Nielsen			
		1	2	3	4+
Third Party	1	11129	53	14	18
	2	99	5042	18	21
	3	50	37	1670	20
	4+	32	32	16	1735

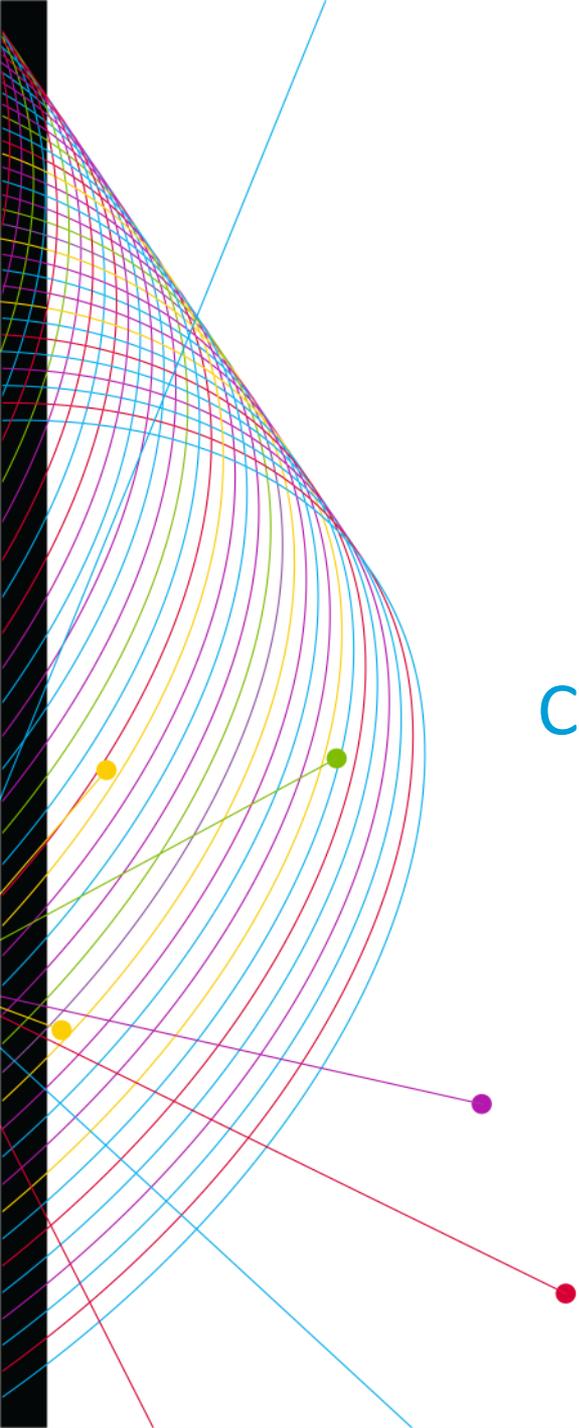
Bin Distance = 0

Bin Distance = 1

Bin Distance = 2

Bin Distance = 3

Bin Distance is helpful for ordinal characteristics where accuracy does not tell the complete story

An abstract graphic on the left side of the slide. It features a vertical black bar on the far left. From this bar, a series of curved lines in various colors (blue, green, yellow, orange, red, purple) fan out to the right, forming a cone-like shape. Several colored dots are placed at the ends of these lines, and thin lines connect some of these dots to each other, creating a network of connections.

# CALIBRATING AND COMPLETING THIRD PARTY DATA

# SET TOP BOX DATA OVERVIEW



Cable and Satellite providers collect TV tuning information from TV Sets in the home

## Benefits



Volume and Granularity of data

Increases Rating Stability

Fewer minutes with no audience

Report more stations

## Challenges



Data Quality

Coverage

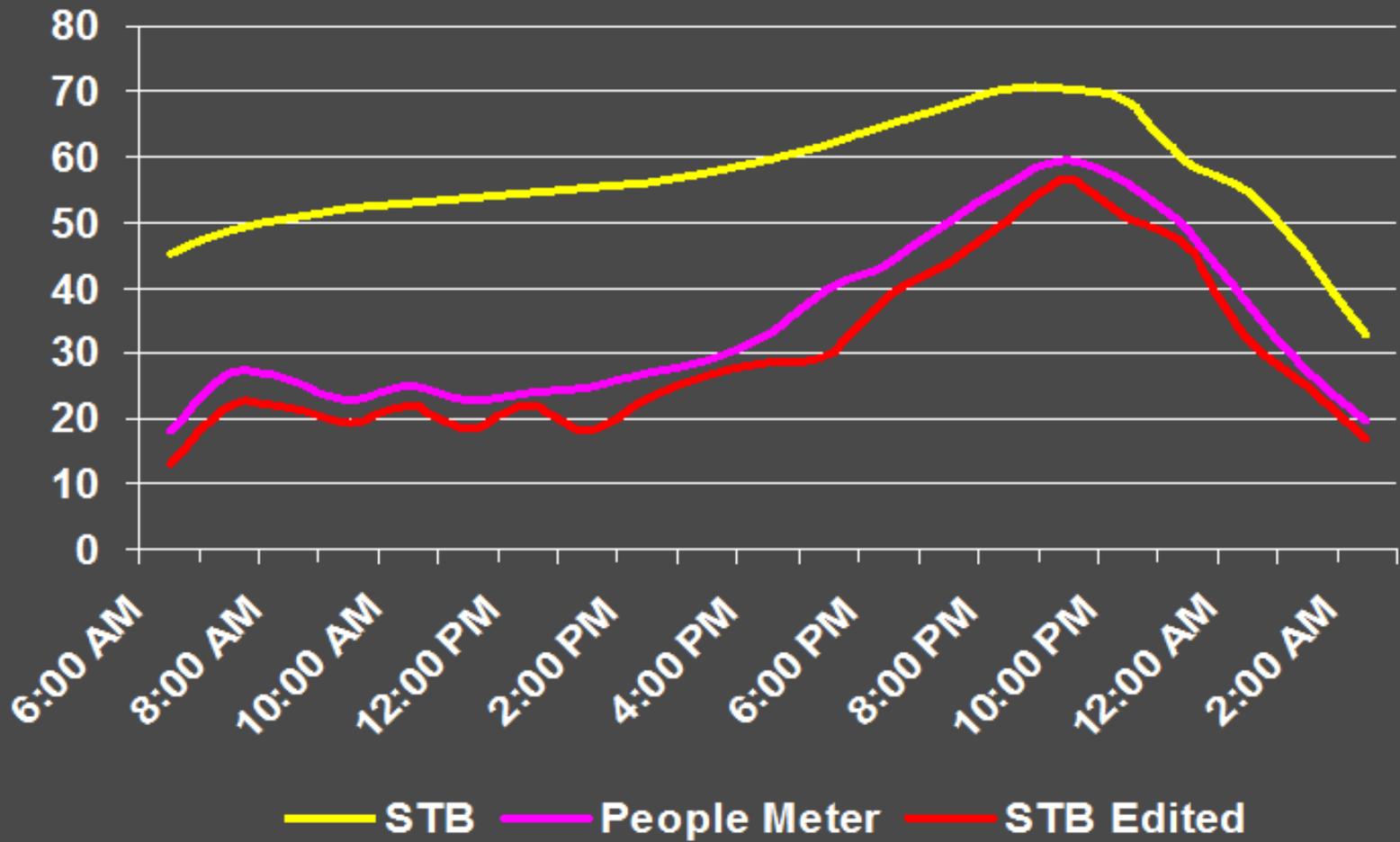
Bias

Person Level detail

# EVALUATING BEHAVIOR – SET OFF MODELING



## STB and People Meter Tuning Levels Compared

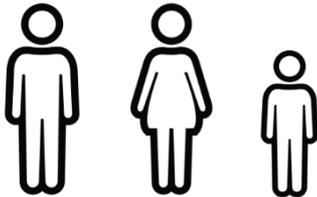


Copyright © 2012 The Nielsen Company. Confidential and proprietary.

# COMPLETING THE DATA

Panel Home  
-Known Demographics  
-Measured Tuning

Composition



Set Top Box Tuning



Set Top Box Home  
-Unknown Demographics  
-Measured Tuning

Composition



Set Top Box Tuning



# USING PANEL DATA TO FIND DEMOGRAPHICS

- We use behavioral patterns that are similar between panel and non-panel homes to infer the unknown demographics of the non-panel homes
- Proportionate usage is promoted through penalizing re-use



nielsen  
.....

AN UNCOMMON SENSE  
OF THE CONSUMER™

