“What’s the Difference Between a Data Scientist and an Analyst?”:

Some Findings From a Recent Study

Jeanne Harris

Vijay Mehrotra

Presentation Outline

- Who Are We, What Did We Study, and Why/How
- Summary of Findings
- Data and Details
- Summary and Discussion

Who We Are

Jeanne Harris

- Global Managing Director, Accenture Institute for High Performance
- Co-Author, ***Competing on Analytics*** and ***Analytics at Work***
- Faculty Member, Columbia University

Vijay Mehrotra

- Professor, University of San Francisco & Columnist, Analytics Magazine
- Co-Founder and CEO, Onward Inc. (acquired by Advertising.com)
- Investor/Advisor: Resonate Networks, Tapjoy, LiftMedia, Fabless Labs, Rubicon Project

Motivation For Research Study "*What's a Data Scientist?*"

SOME COMMON ANSWERS

- A Statistician That Lives in California
- Someone Without Enough Personality to be a Real Scientist
- A Systems Administrator That Knows How to Program in R
- A Better Programmer Than Most Statisticians &
  A Better Statistician Than Most Programmers

What We Considered:
A (Slightly) Different Question

What Distinguishes a

Data Scientist

from a Traditional Analyst?

"Data Scientist is as Data Scientist Does"

Our Methodology

1. Preliminary Interviews
2. Survey Development and Implementation
3. Focus Groups
4. Data Analysis and Summarization

Our Methodology

- Data Scientists vs Traditional Analysts
  - Identified through Job Title (Initial Wave of Respondents) OR
  - Self-Identified (Subsequent Waves of Respondents) with Manual Cross-Checking of Job Titles
- Sample Sizes
  - Data Scientists n=109
  - Traditional Analysts n=208

Presentation Outline

- Who Are We, What Did We Study, and Why/How
- Summary of Findings
- Data and Details
- Summary and Discussion

Data Scientists are Different [Statistically Speaking]

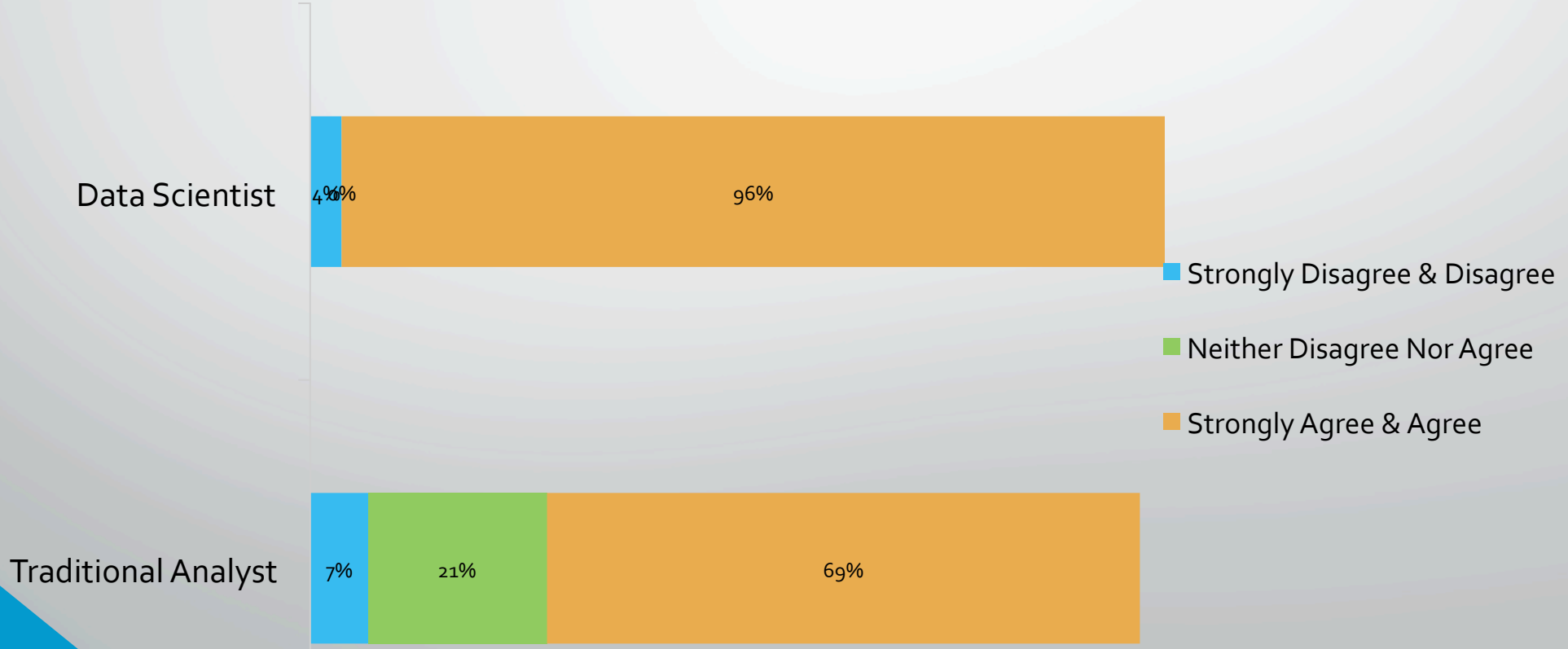
- Have a More Business-Centric Mindset
- Are More Focused on Innovation and Customer Understanding
- Work with Larger, More Diverse Data Integrated Across More Sources
- Seem Less Skeptical About Big Data Hype
- Are More Likely to Report Up to CIO

Presentation Outline

- Who Are We, What Did We Study, and Why/How
- Summary of Findings
- Data and Details
- Summary and Discussion

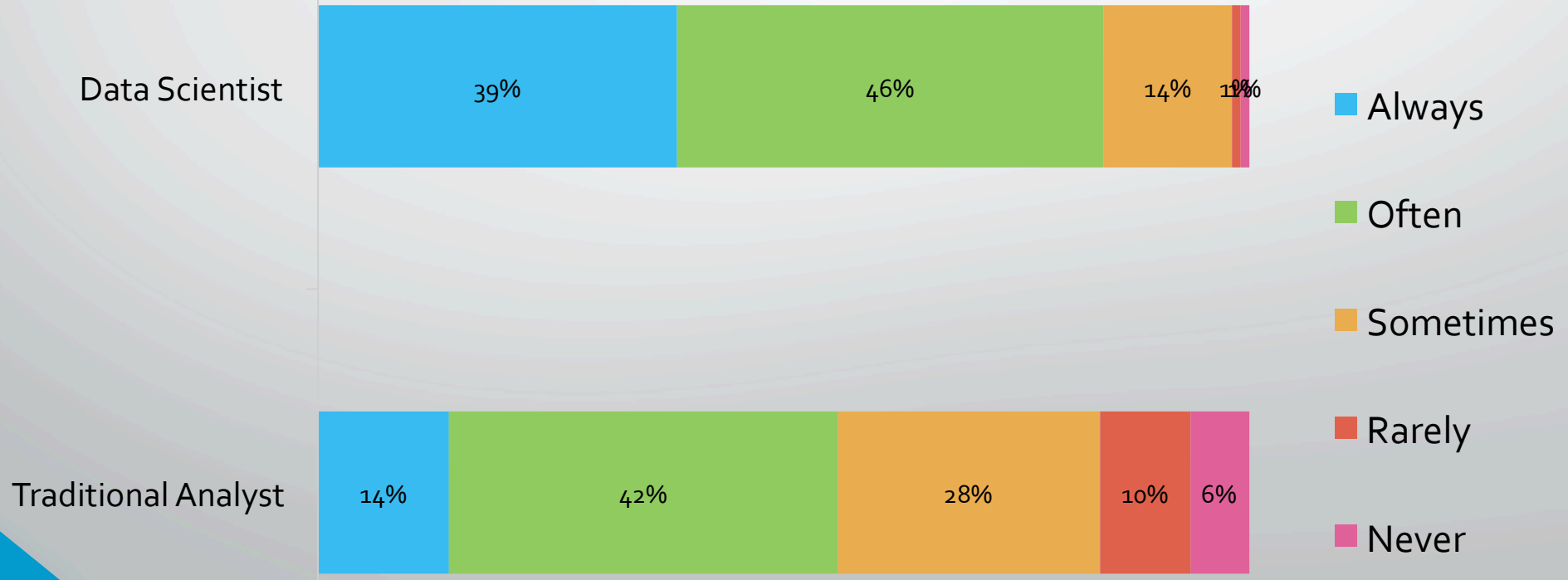
Finding: More Business-Centric Mindset

"Our group is very entrepreneurial in creating new sources of value from analytics"



Finding: More Business-Centric Mindset

"We create prototypes to help get support for our projects from key stakeholders"



Finding: More Business-Centric Mindset

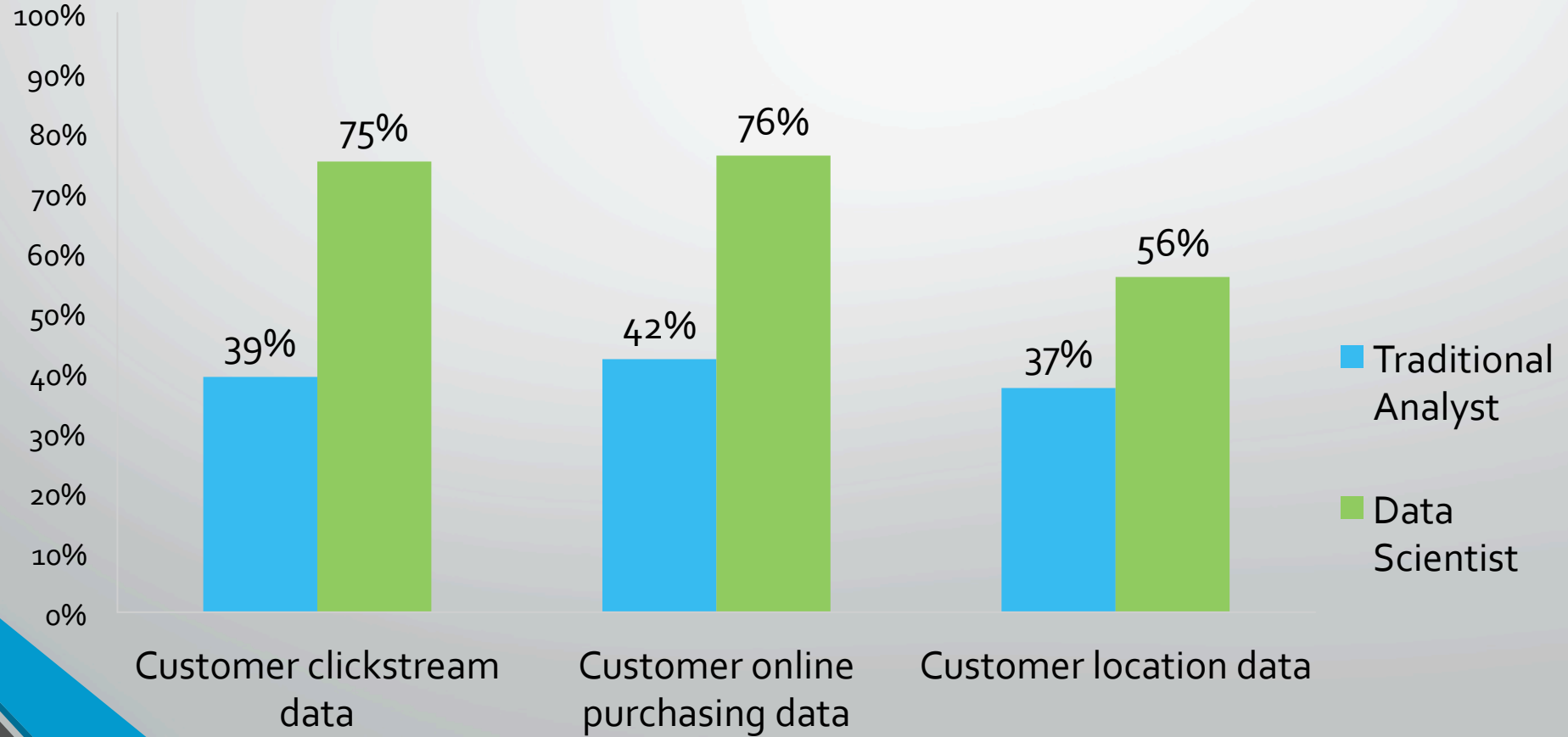
- "Data scientists are being asked to go further into new areas, rather than just traditional well-defined projects like marketing automation"
- "We have embraced the key problem of translating complex analytic results to business users"
- "We are seeing tight integration with our product teams and our product managers"
- "Data scientists are much more reluctant to simply be analysts stuck in a closet"

Finding: Customer-Focused

"The success of our group is based on our ability to:"

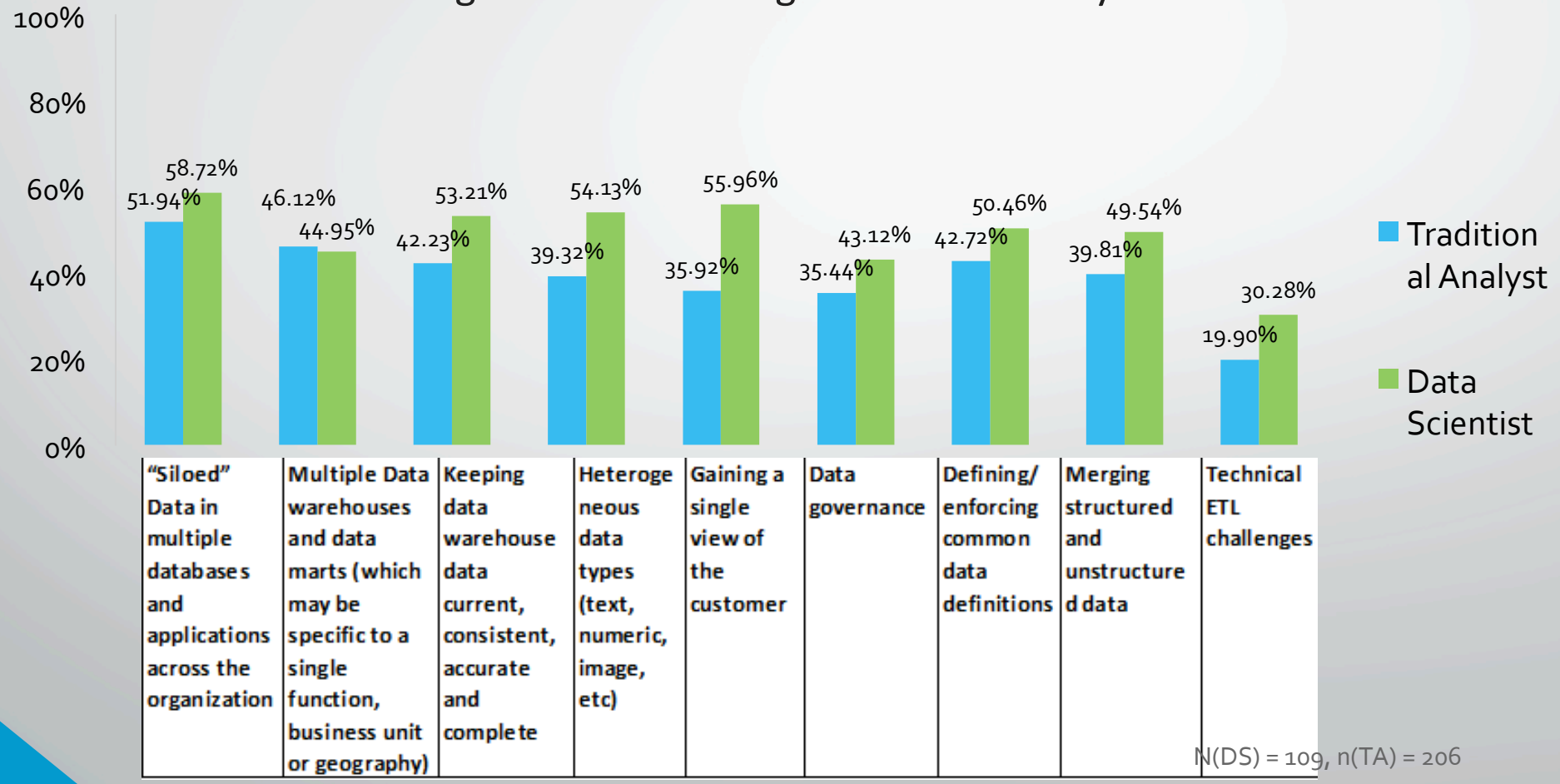
Finding: Customer Focus

"We use the following types of data in our analysis"



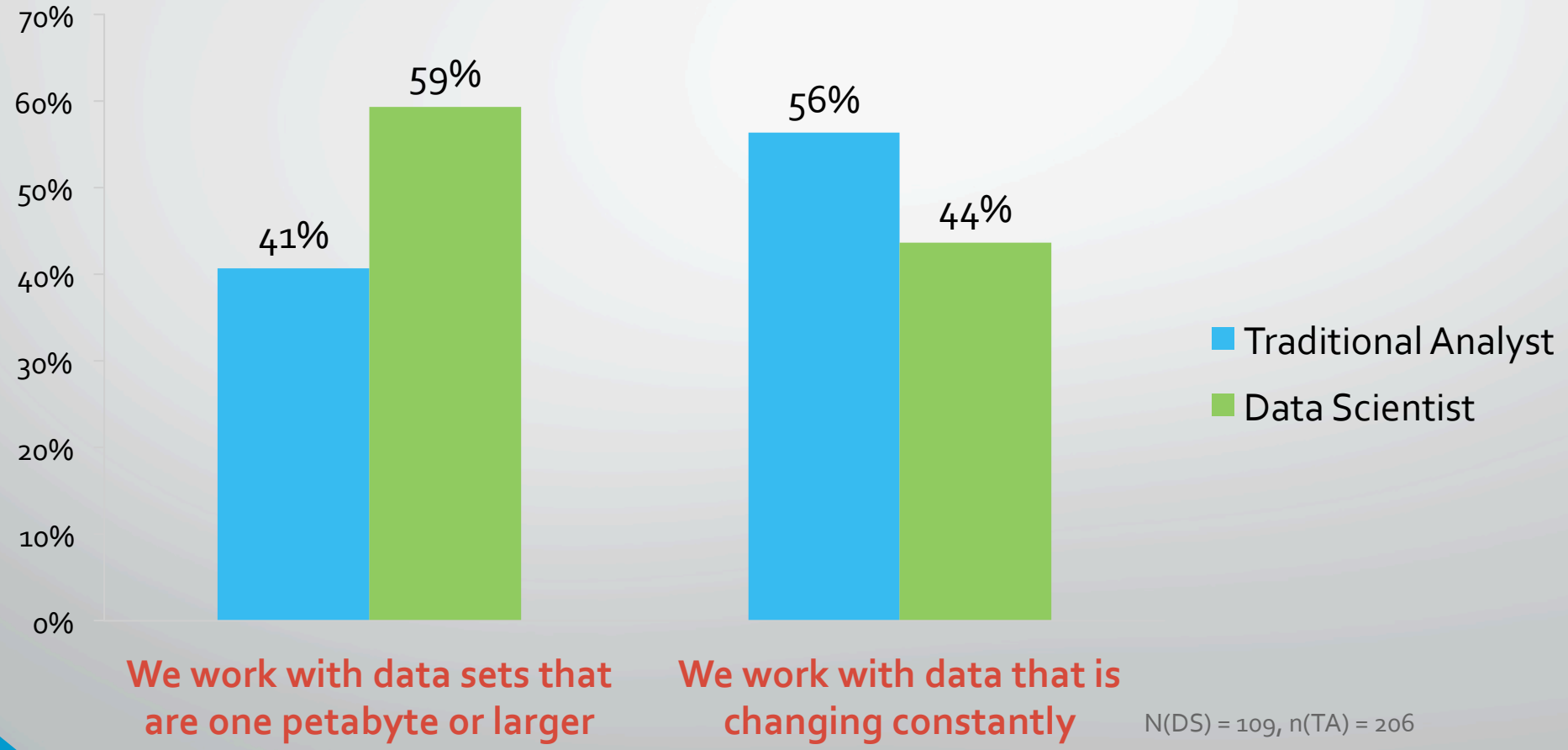
Findings: Customer Focus, More Diverse Data Sets

What are the greatest data integration obstacles you face?



N(DS) = 109, n(TA) = 206

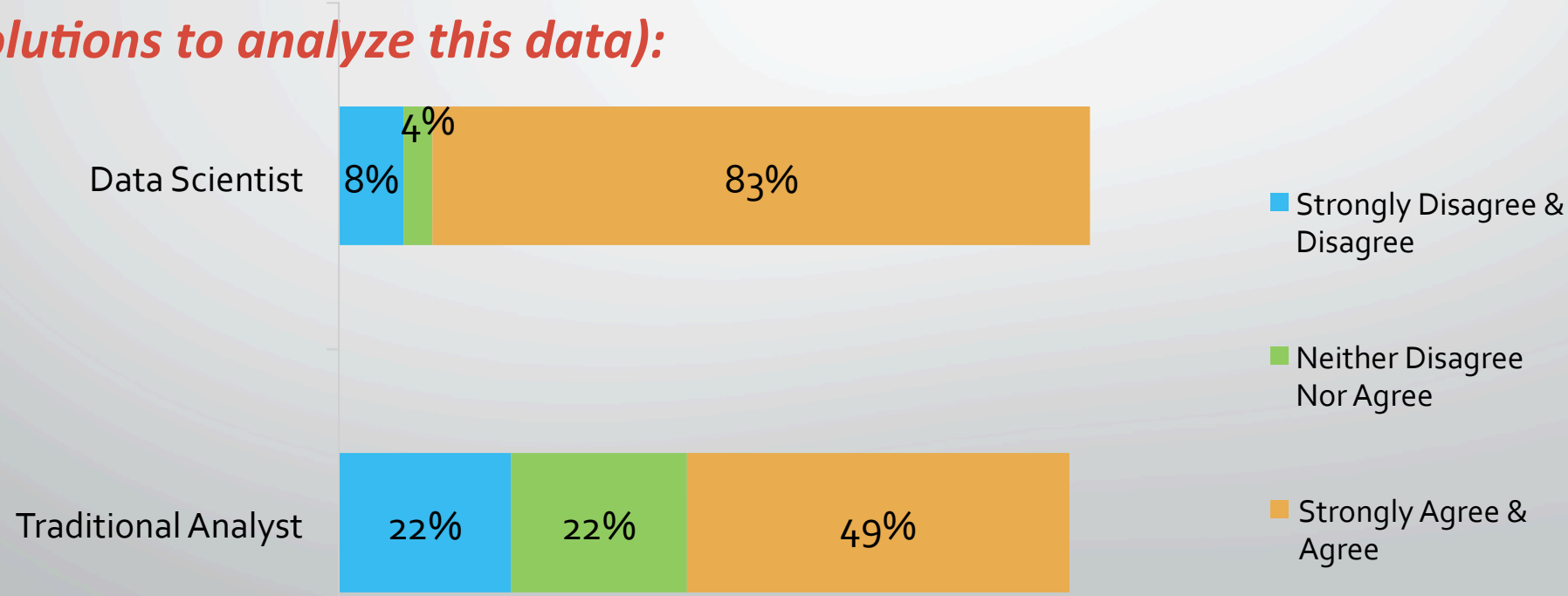
Finding: Larger, More Diverse Data Sets



N(DS) = 109, n(TA) = 206

Finding: Less Skepticism About Big Data Among Data Scientists

The volume and variety of data my group deals with is causing us to consider using Hadoop-based solutions within the next 12 months for storing our data (and distributed computing solutions to analyze this data):



N(DS) = 109, n(TA) = 206

Data Scientists Less Skeptical About Big Data, But Still Measured

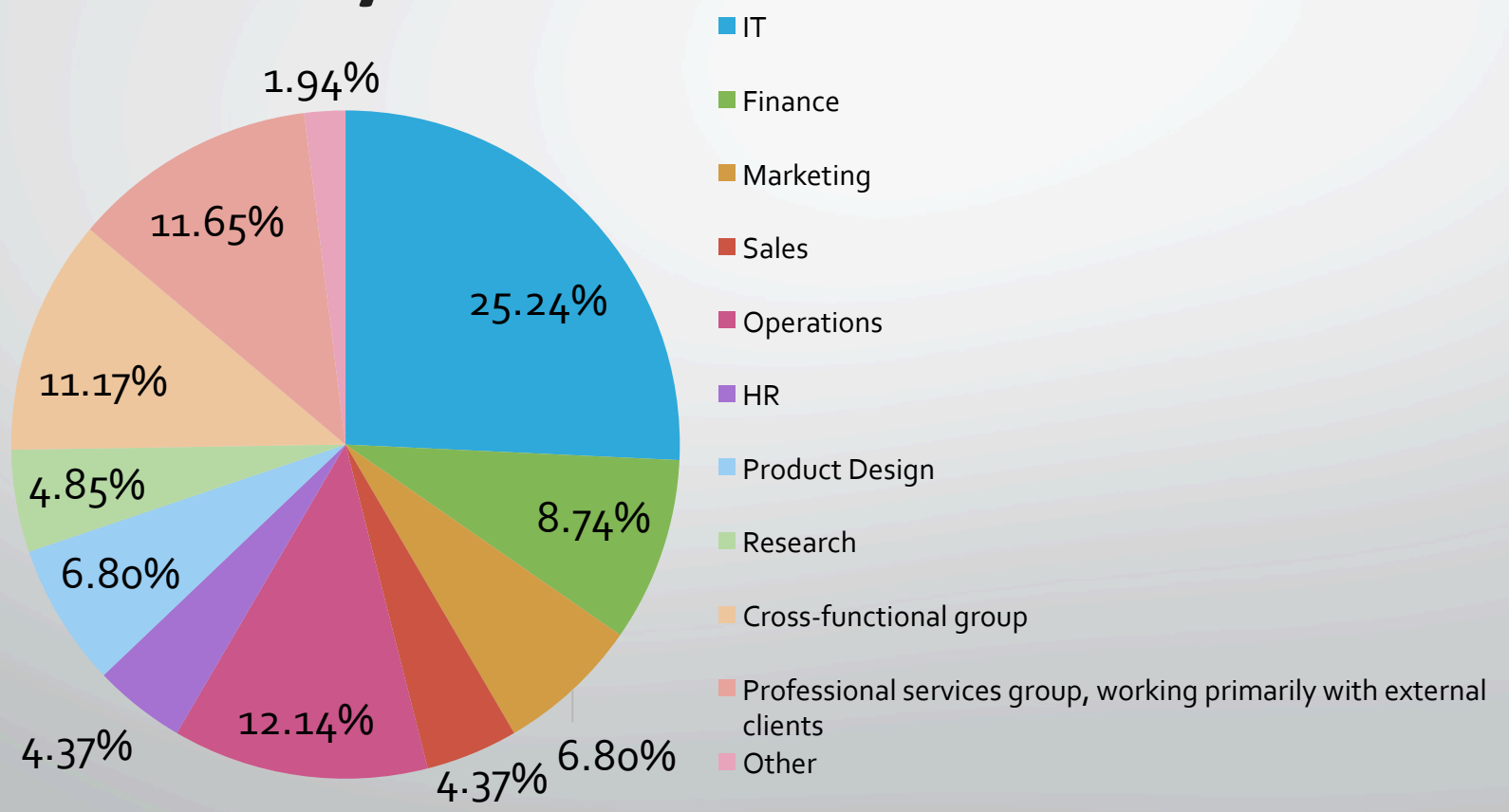
- **Big data is not replacing small data, but instead we see that the problems that we encounter are growing in scope and complexity.**
- **We are just barely starting to understand the potential of Big Data**
- **We face a lot of pressure in our marketplace to appear to be leveraging Big Data, but often our reality does not match up with our rhetoric…**

Traditional Analysts Are More Skeptical About Big Data

- **Big Data is coming, but the hype is still way out ahead. Most people using Hadoop today are researchers and platform companies.**
- **Big Data will take longer than anyone thinks to create value on any kind of scale**
- **Big Data is sexy, but problem solving is beautiful**

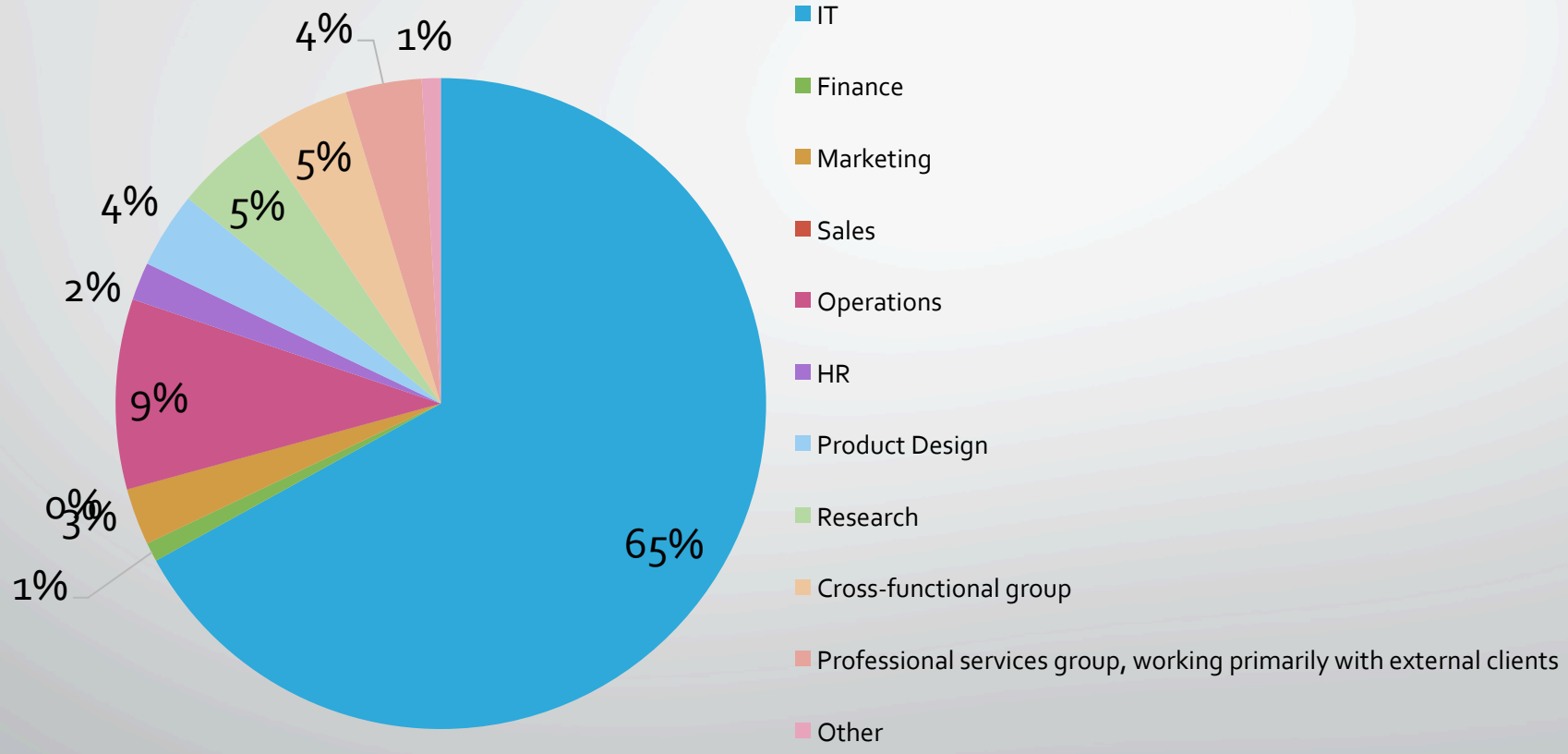
"What Part of the Organization Are You In?"

Traditional Analysts



“What Part of the Organization Are You In?”

Data Scientists



Presentation Outline

- Who Are We, What Did We Study, and Why/How
- Summary of Findings
- Data and Details
- Summary and Discussion

Yes, Data Scientist Are Different!

- **More Data (VVV), More Mathematical Tools**
- **More Forward Looking, More Focus on Prediction**
- **More Likely to Embrace Big Data Solutions**

BUT MOST SIGNIFICANTLY

- **Different Mindset Than Traditional Analysts**

EMERGING IDEAL OF A DATA SCIENTIST: "A superb business consultant who also has a broad array of technical skills for data management, analysis, and modeling"

Questions? Thoughts? Comments? Ideas? Objections?

Jeanne Harris
Accenture IHP
jeanne.g.harris@accenture.com

Vijay Mehrotra
University of San Francisco
vmehrotra@usfca.edu