Intermediate Analytics in Excel for System Diagnosis and Improvement

April 2019

Andrea Miller, Cloudburst
Mary Schwartz, Abt Associates

Learning Objectives

- Retrieve LSA data sets from HDX 2.0
- Use intermediate analytic strategies to assess, parse, transform and restructure large datasets into meaningful subsets of data
- Create an Excel-based data dashboard

Rationale:
Dynamic use of Excel is required for many of the visualizations and dashboards that are being produced locally to examine homeless systems of care. HUD encourages all HMIS Leads and other stakeholders to become proficient in macros, advanced pivot tables, slicers, and dashboards, and how to integrate data across multiple sources using built-in or add-on tools.

Data Analytics Overview

Beginner

- Informal/ad hoc planning
- Counts, sums, averages
- Descriptive
- Bars, columns, lines, waffles/pies

Intermediate

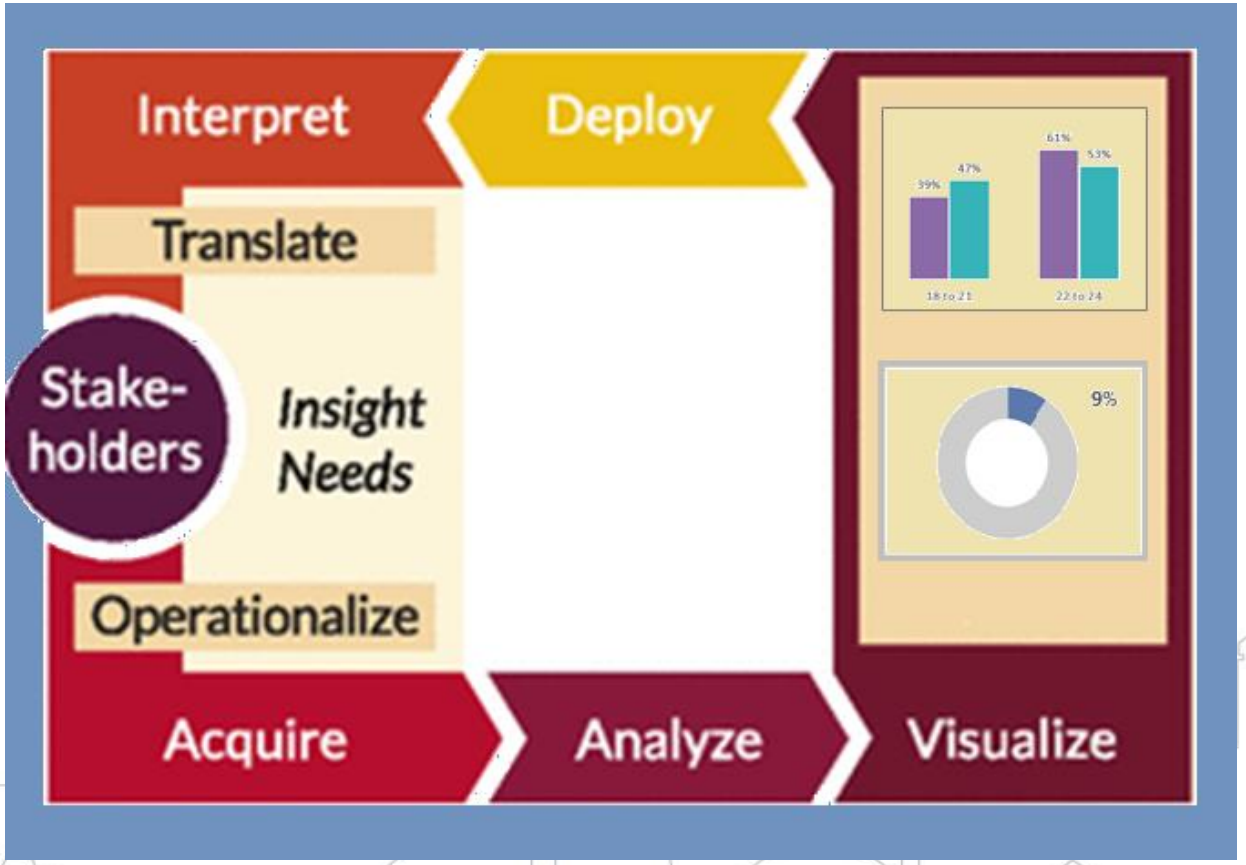
- Formal data analysis plan
- Summary statistics, comparisons
- Descriptive and Diagnostic
- Box-and-whisker, stacked columns/bars, scatter, bubble, dual axis

Expert

- Data project life cycle
- Co-dependencies, detailed variables/statistical analyses
- Inferential statistics
- Predictive and Prescriptive
- Scatter matricies, violins, sankeys

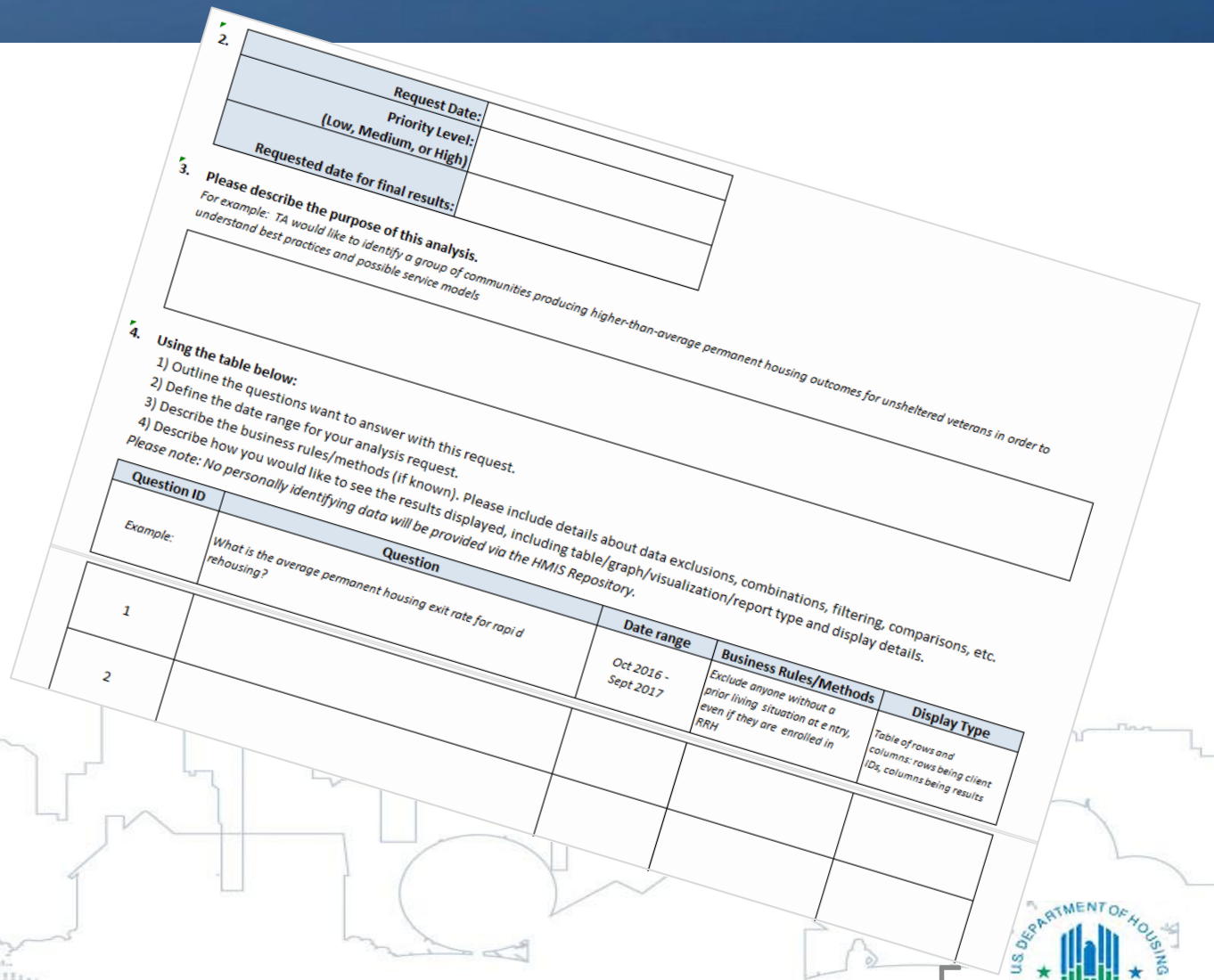
The Framework

The data project life cycle



Planning Processes

- What questions need to be answered?
- What elements/variables do you need to answer those questions?
- What datasets do you have access to that include that info?
- Project Plan: visualize all the tools, time and resources needed



2.

Request Date:	
Priority Level: (Low, Medium, or High)	
Requested date for final results:	

3. **Please describe the purpose of this analysis.**
For example: TA would like to identify a group of communities producing higher-than-average permanent housing outcomes for unsheltered veterans in order to understand best practices and possible service models

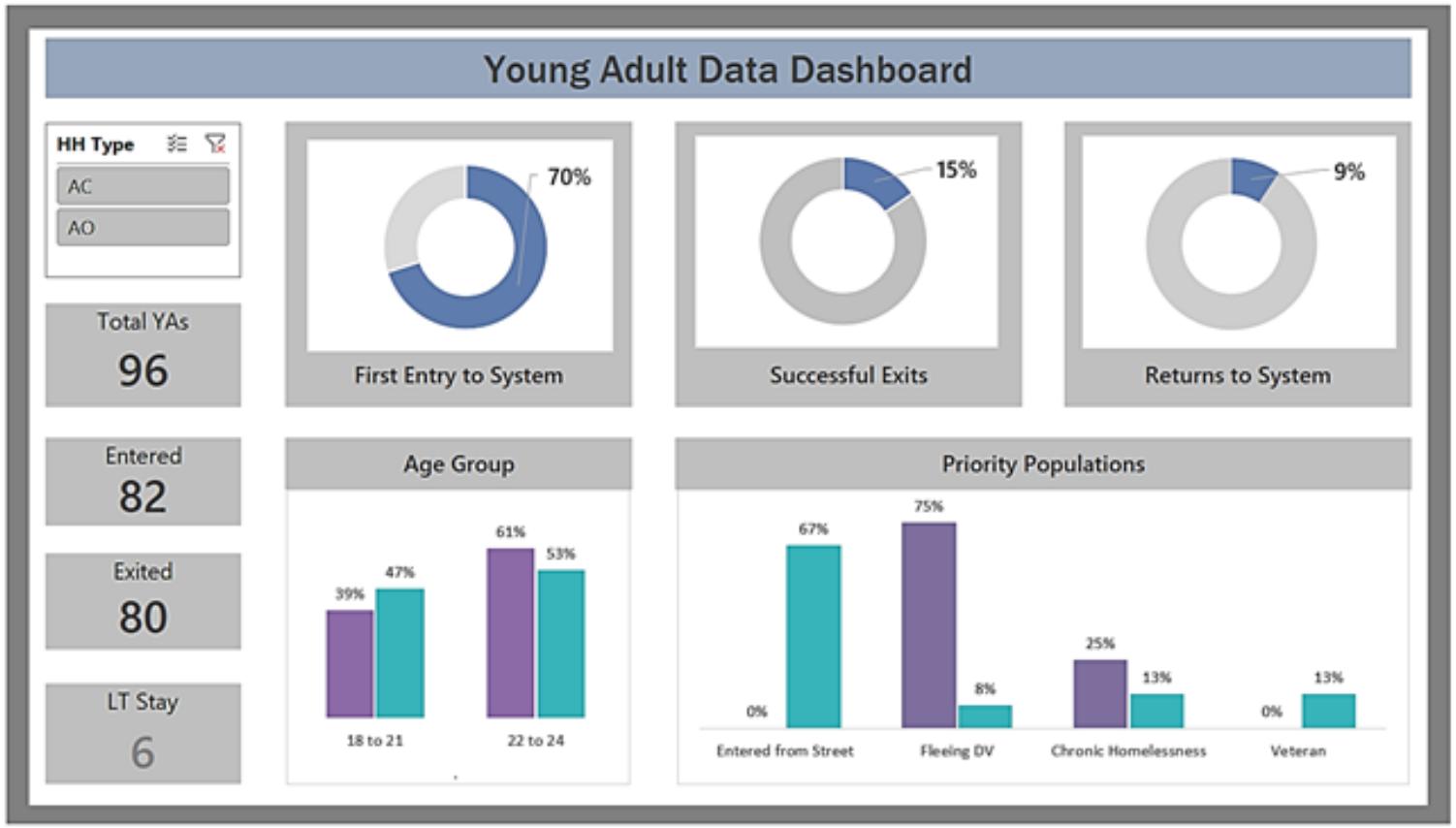
4. **Using the table below:**
1) Outline the questions want to answer with this request.
2) Define the date range for your analysis request.
3) Describe the business rules/methods (if known). Please include details about data exclusions, combinations, filtering, comparisons, etc.
4) Describe how you would like to see the results displayed, including table/graph/visualization/report type and display details.
Please note: No personally identifying data will be provided via the HMIS Repository.

Question ID	Question	Date range	Business Rules/Methods	Display Type
Example:	*What is the average permanent housing exit rate for rapid rehousing?*	*Oct 2016 - Sept 2017*	*Exclude anyone without a prior living situation at entry, even if they are enrolled in RRH*	*Table of rows and columns: rows being client IDs, columns being results*
1				
2				

The End Goal!

A useful, functional
data dashboard
that can be updated regularly
to provide timely information
to staff, stakeholders
and others!

Ctrl+Click to follow link and access the data dashboard workbook>>



The Analytic Process: A Framework

Step 1: Acquire the data
Step 2: Analyze the data
Step 3: Visualize the data
Step 4: Deploy a dashboard
Step 5: Interpret the findings

The Analytic Process

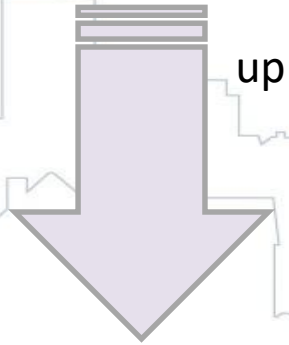
Step 1: Acquire the data

1. Acquire the Data

- Foraging: Where are your datasets? Can you add context with others?
- Sense-making: What is in each dataset and what is relevant? Narrow the variables.

  It is helpful to both expand out to all the possibilities and contract after that to the essential elements. Rinse and repeat.

  Sense-making includes thinking through what tools will help accomplish the task.

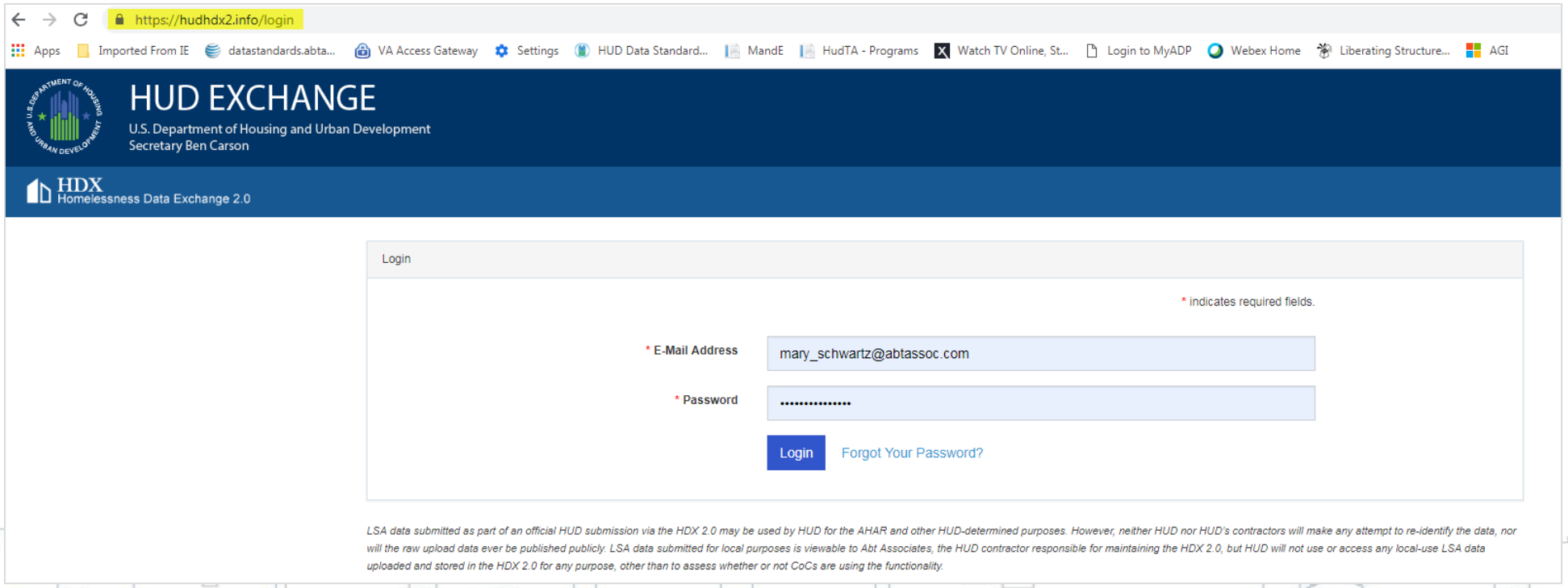


up next: pro tip!

1. Acquire the Data

PRO TIP

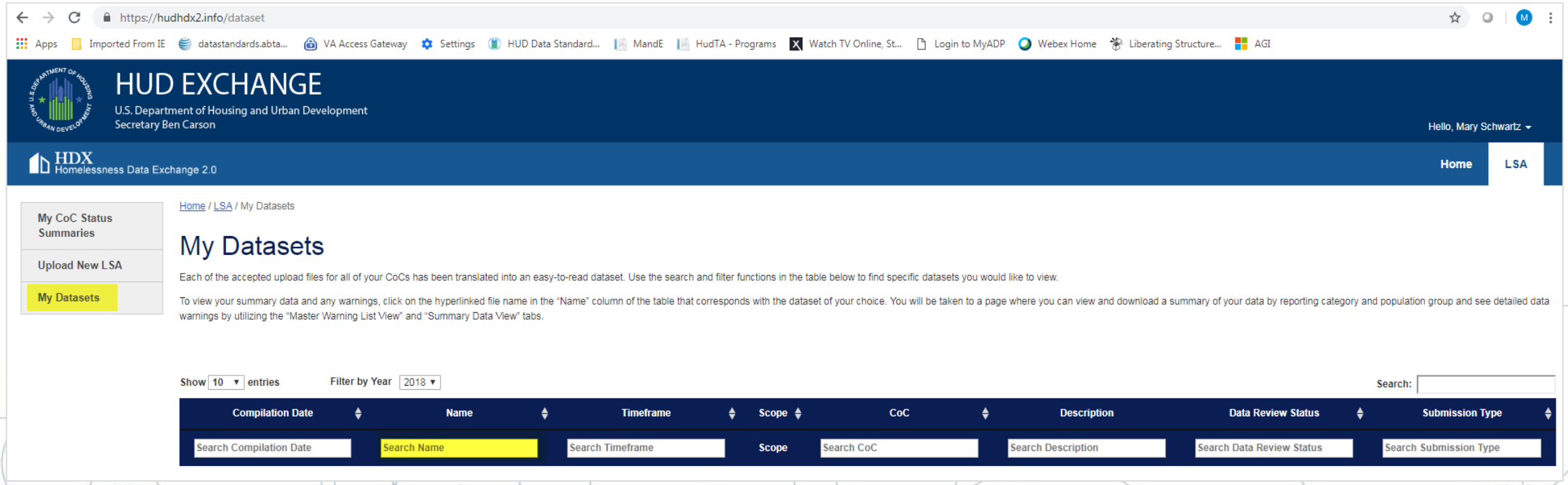
HDX 2.0 Foraging



1. Acquire the Data

PRO TIP

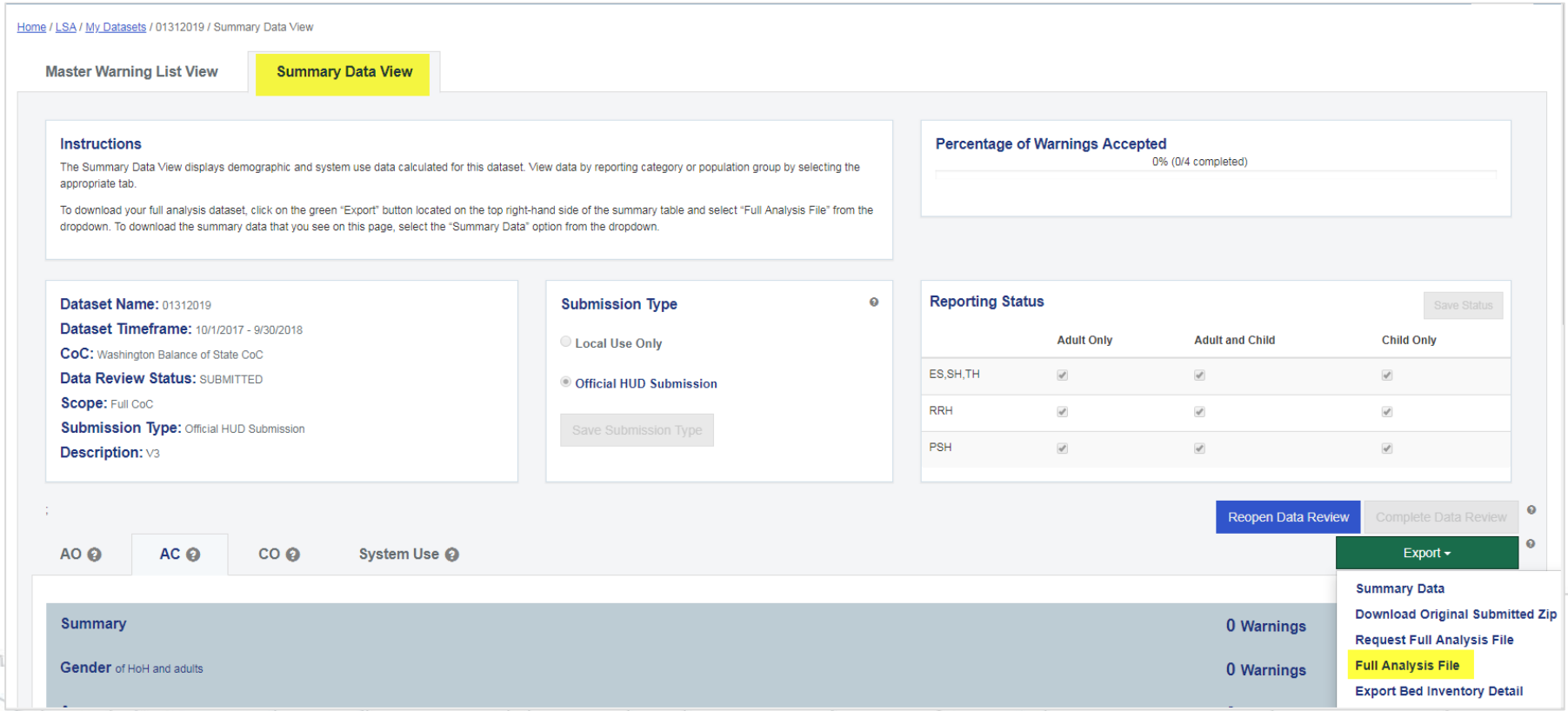
HDX 2.0 Foraging



1. Acquire the Data

PRO TIP

HDX 2.0 Foraging



1. Acquire the Data

RESOURCES

Data workflow tools

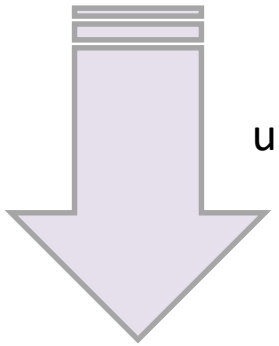
- SQL, such as MySQL or PostgreSQL
- Self service data analytics platforms
- Statistical software scripts, such as those scripted in R
- Data integration plug-ins

The Analytic Process

Step 2: Analyze the data

2. Analyze the Data

- **Clean and prep the data**
- **Conduct the analysis**
- **Make use of your analytics toolkit**



up next: pro tip!

2. Analyze the Data

PRO TIP

Create a Data Analysis Plan

HMIS Data Analysis Plan Template

Data Project: _____

Objective: _____

Key Questions *What are you trying to learn? What are the key outcomes of interest?*	1) 2) 3)
Target Population *with sample size*	
Selected Time Frame	Start Date: End Date:

2. Analyze the Data

PRO TIP

The Data Analysis Plan

Unit(s) of Analysis	Primary:	By:
Dataset Variables	• Universal data elements with all person/household identifiers • Project and household type • • •	
Data Prep	• Delete exact duplicates • Identify missing data and temporarily manage - such as HHIDs, destination, etc. • Examine outliers to determine if they represent invalid data – and correct • Reformat data that lacks consistency or that is not structured properly	
Data Transformation		

2. Analyze the Data

- **Clean and prep**
  - De-duplicate
  - Handle missing values
  - Assess outliers
  - Restructure the data set as needed
  - Transform the data: calculate, recode/geocode

2. Analyze the Data

- **Use your analyst's toolkit**
  - At least one data analysis platform
  - Customizations and plug-ins to enhance the workflow
  - Go-to resources

2. Analyze the Data

PRO TIP

Customization: Use a customized Macro for simple automations

Demonstration

1. Enable the **Developer Tab**
2. Select “Record Macro” and name your script
3. Complete the activities that will be saved as a macro script
4. Select “Stop Recording”
5. Presto! *

*er, with countless caveats

2. Analyze the Data

RESOURCES

for productivity

- Data prep tools
- Analysis or query tools: Excel; Tableau Public
- No Cost statistical tools
  - Real Statistics in Excel
  - R Studio
  - Epi Info (from the CDC)
  - SOFA ("Stastistics Open For All")

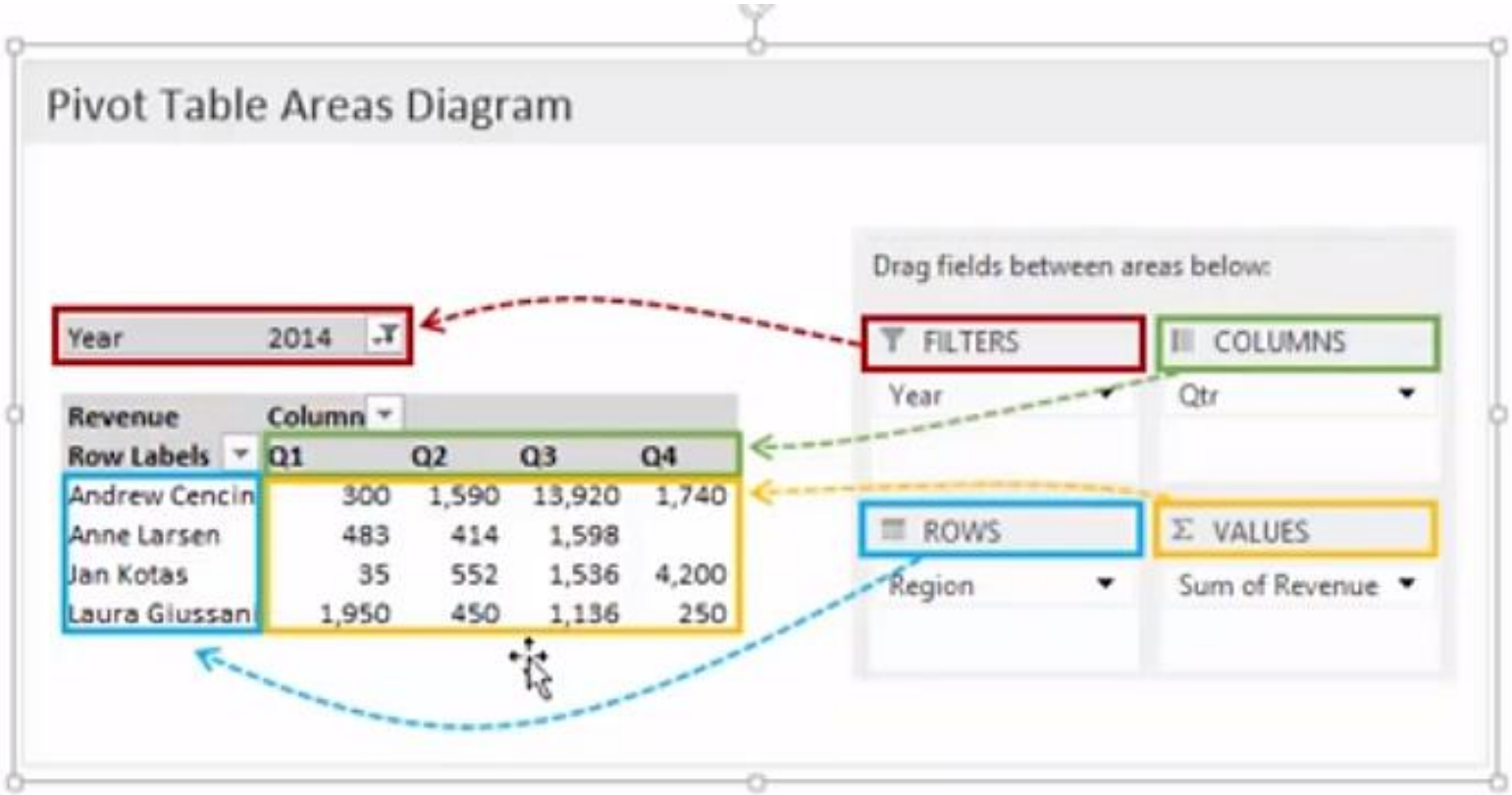
2. Analyze the Data

Demonstration of Analysis Phase

1. Create Young Adult subset files from the LSA Full Analysis File
   - Utilize Excel functions, or a macro, or a plug-in
   - Yield 3 subanalyses data sets: YA demos, YA LOS, and YA outcomes
2. Recode and/or transform select variables
3. Create PivotTables from each subset file
   - First, use filtered tables to review, inspect, de-duplicate and validate the data
   - Next, use filtered tables to analyze the KPIs that will be included on the data dashboard

2. Analyze the Data

Demonstration:
The Anatomy of a Pivot Table



The Analytic Process

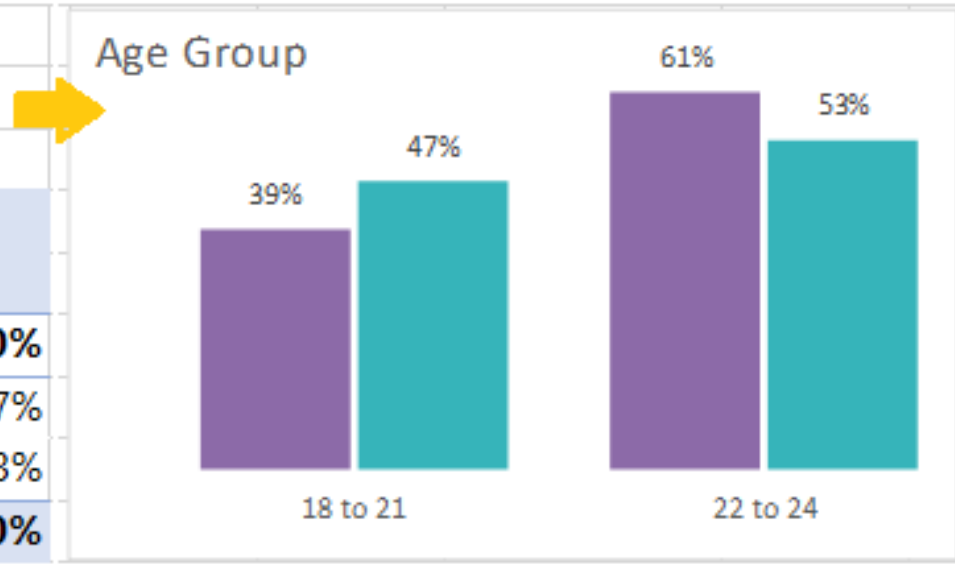
Step 3: Visualize the data

3. Visualize the Data

Demonstration of the Visualization Phase

1. Create PivotCharts utilizing the analytic PivotTables
2. Customize!

Project Type	ES,SH,TH	
Household Subtype	(Multiple I ms)	
Sum of Value	**Column L**	
Row Labels	**AC**	**AO**
Age Group	**100%**	**100%**
18 to 21	39%	47%
22 to 24	61%	53%
Grand Total	**100%**	**100%**



3. Visualize the Data

PRO TIP

Data Viz Best Practices

1. Layout
   - Dashboard design
2. Color
   - Color theory and accessibility
   - Where are your eyes drawn?
3. The data-ink ratio
   - Remove to improve / less is more
4. Which chart when

3. Visualize the Data

PRO TIP

Which Chart to Use When?

- Simple (sometimes best): Columns, bars, pies
  Good for easily comparing values (Pies – up to 3, columns/bars up to 8) – add color for easy analysis
- Other visualizations: box-and-whisker, histograms, scatterplots
  Good for statistical analysis, distribution, multiple variables – you must know your data to use these well
- Line charts are the most common ways to visualize variable(s) over time, with time as the X (bottom) axis

3. Visualize the Data

RESOURCES

Master best practices…

- Interactive Chart Chooser
- Data Viz Checklist and Rate Your Viz
- How to Choose the Right Visualization

For a deeper dive…

- Data Visualization Catalogue
- Data Visualization Project/Function
- Visualize Vocabulary

The Analytic Process

Step 4: Deploy the Dashboard

4. Deploy a Dashboard

- Create a mock up
  Utilize dashboard design best practices

- Utilize low-tech automation
  Particularly useful for regularly updating the dashboard

- Customize the data output and level of interactivity to meet the needs of the audience

4. Deploy a Dashboard

- Create a mock up

4. Deploy a Dashboard

- Streamline and automate

4. Deploy a Dashboard

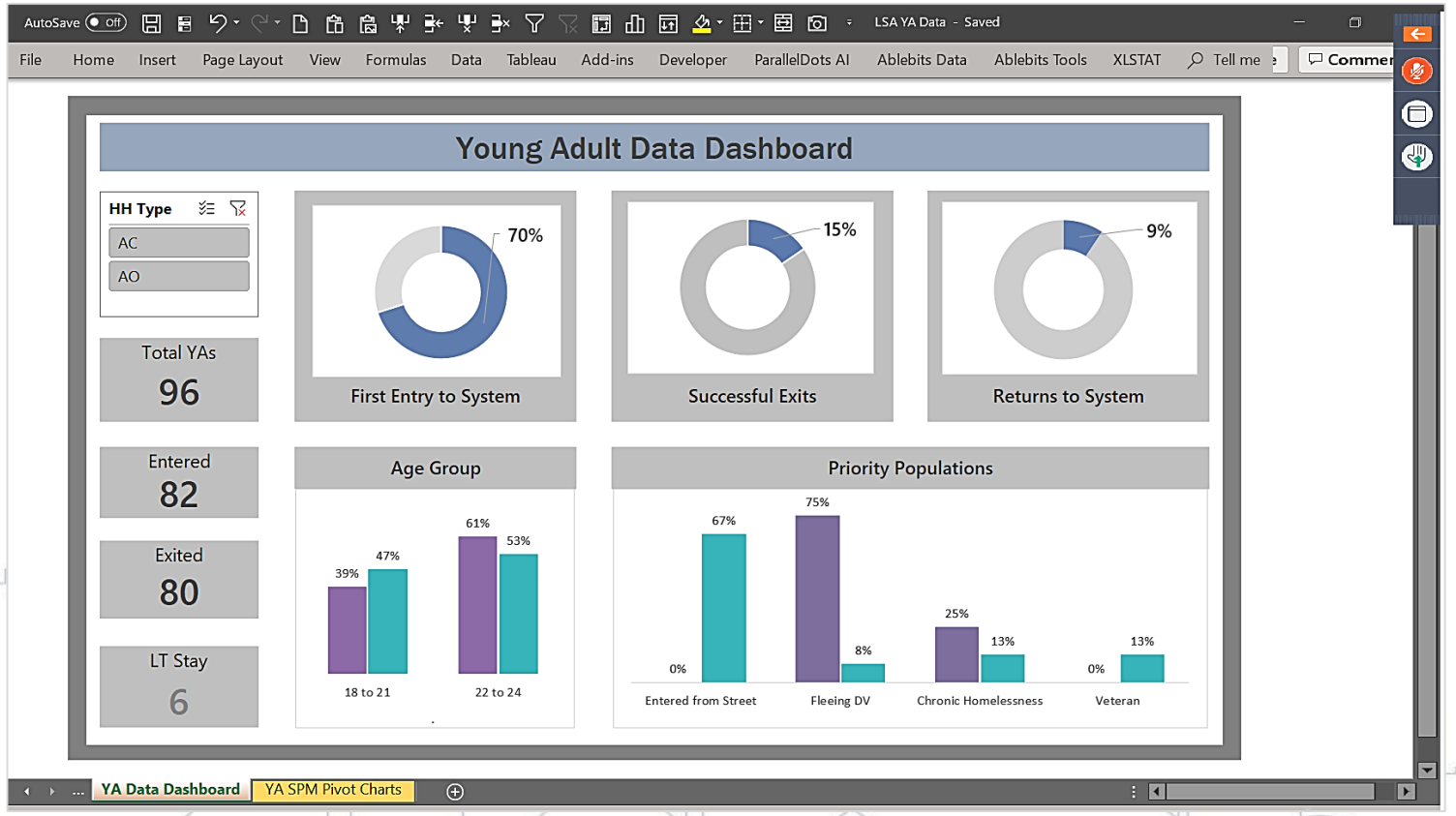
- Customize interactivity to meet the needs of the audience

Data User Role	Description
Data Consumer	Interested and engaged. If the data trigger action, may move into a more active role. Typically have little data and domain expertise.
Data Actor	Act on and leverage the data to drive change. May have significant clout, staff and domain knowledge but limited time.
Data Promoter	Leverage data to create additional value: inform, educate or build products around data. Multiply the audience and may influence consumers and actors.
Data Analyst	Use data to create deeper understanding. Have deep domain knowledge and extensive data knowledge.
Data Researcher	Work in the trenches to collect, analyze, and synthesize data for the groups above. May perform data collection and analysis themselves.

4. Deploy a Dashboard

Demonstration of the deployment phase:

How to build an interactive dashboard focused on Young Adults using Downloaded LSA data



The Analytic Process

Step 5: Interpret/translate/communicate

5. Interpret

No demonstration of the Interpret/translate phase!

- **Interpret**
  - > Identify insights
  - > Provide context
- **Translate and Communicate**
  - > Apply data communication principles
  - > Use data engagement strategies

5. Interpret and Translate

PRO TIP

Data communication principles

1. Know your goal
2. Use the right data
3. Select the right visualizations
4. Design for aesthetics
5. Choose the right medium and channel
6. Check the results

Summary

- Revisit learning objectives

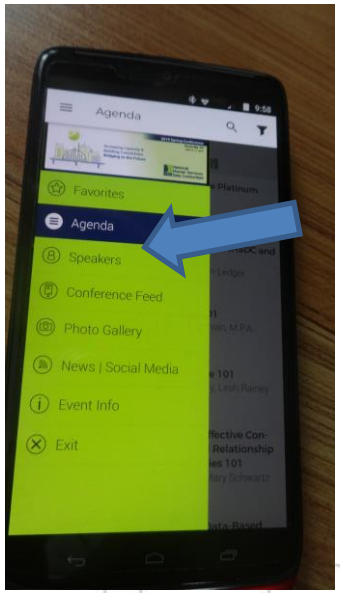
- Retrieve LSA data sets from HDX 2.0
- Use intermediate Excel strategies to assess, parse, transform and restructure large datasets into meaningful subsets of data
- Create an Excel-based data dashboard

Local Action Planning

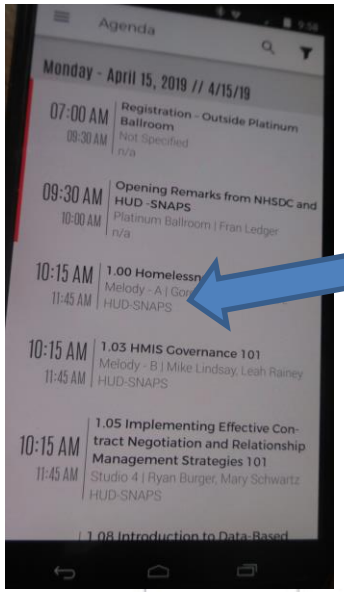
- Can your CoC benefit from ad hoc analytics?
- Do you foresee being able to draw from downloaded LSA data?

Evaluate This Session on Your Conference App!
(It takes 5 minutes to complete)

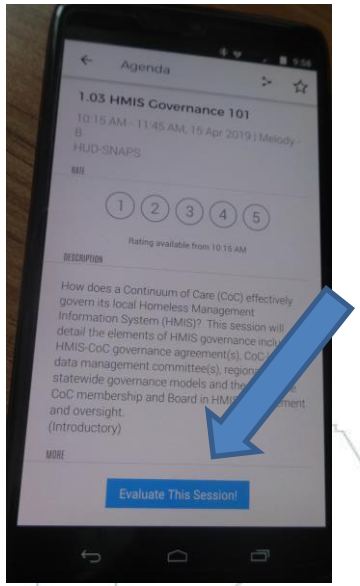
1) Select “Agenda” from the navigation menu.



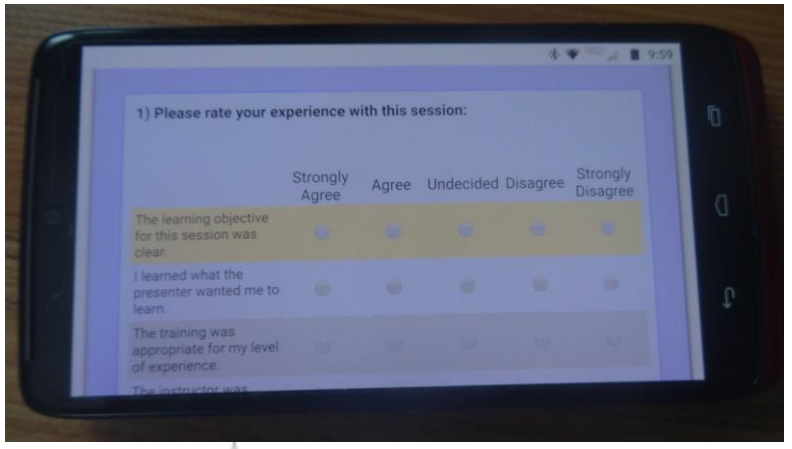
2) Select the name of the session.



3) Select the blue “Evaluate This Session”.



4) Complete the Evaluation and Select “Finish”.



TIP:

Turn your phone horizontally to see rating options.

U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

COMMUNITY PLANNING DEVELOPMENT

HUD Certificate-of-Completion

Reminder: HUD is offering a Certificate-of-Completion for completing at least 4 sessions within either track:

1) HMIS Fundamentals Track
2) System Planning with Data Track

To earn credit for completion of this session, please complete the evaluation on the conference app and include contact details when prompted

Thank you!

- Questions & Answers
- Contact Info

Andrea Miller	Mary Schwartz
amiller.rotondi@gmail.com	Mary_Schwartz@AbtAssoc.com