DART Training for PAMS and CSN

(Data Analysis and Reporting Tool for Photochemical Assessment Monitoring Station and Chemical Speciation Network Data)

Jennifer DeWinter, Steve Brown, and Hilary Hafner
Sonoma Technology, Inc.

Beth Landis and Kevin Cavender
U.S. EPA

for

National Ambient Air Monitoring Conference

Data Validation Training Session
St. Louis, MO

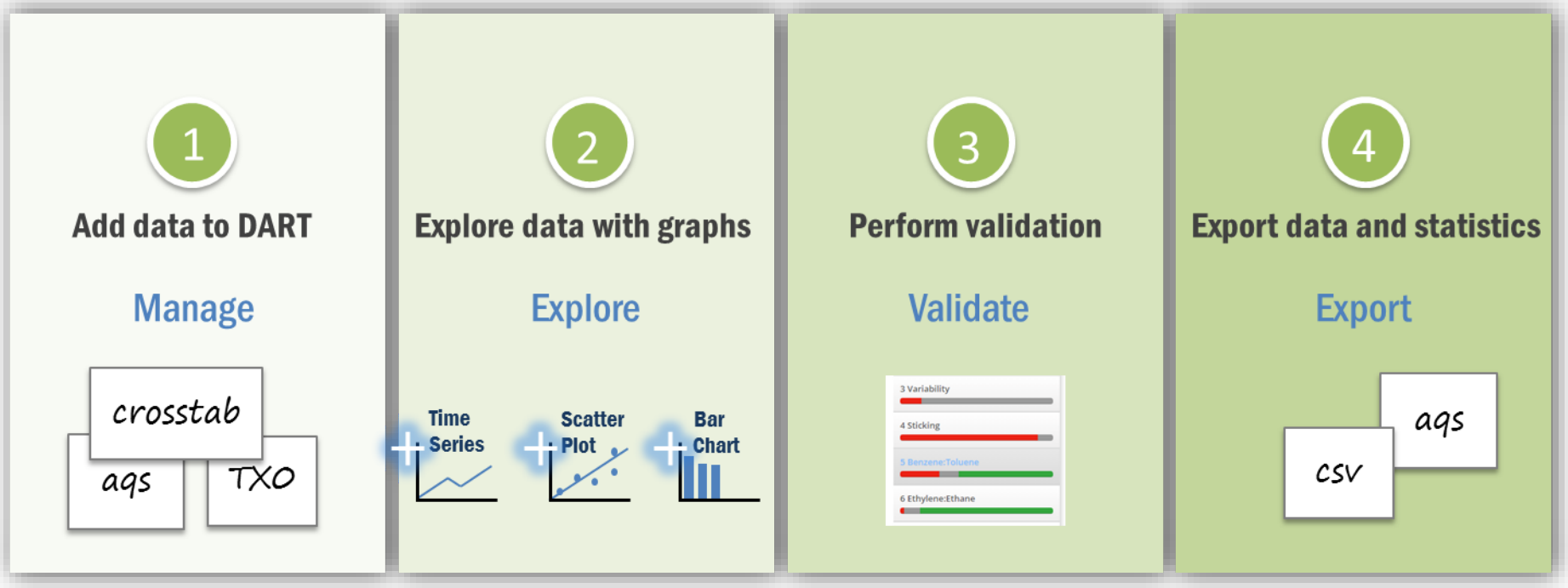
August 8, 2016

Sonoma Technology, Inc.

Outline for Training Session

- DART overview
  - Accessing and navigating DART
  - Bringing data into DART
- Using DART as part of the new CSN data review process (with live demo)
- Using DART for PAMS (and other) data validation and analysis (with live demo)
- Future enhancements
- Q&A

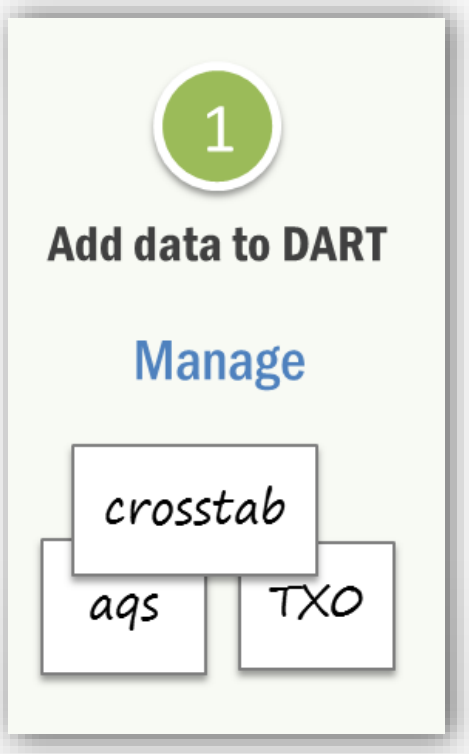
Data Analysis & Reporting Tool



- DART is a web-based data validation and analysis tool that allows monitoring agencies to manage, explore, validate, and flag data.
- DART includes automated screening, statistical summaries, and AQS-formatted data.

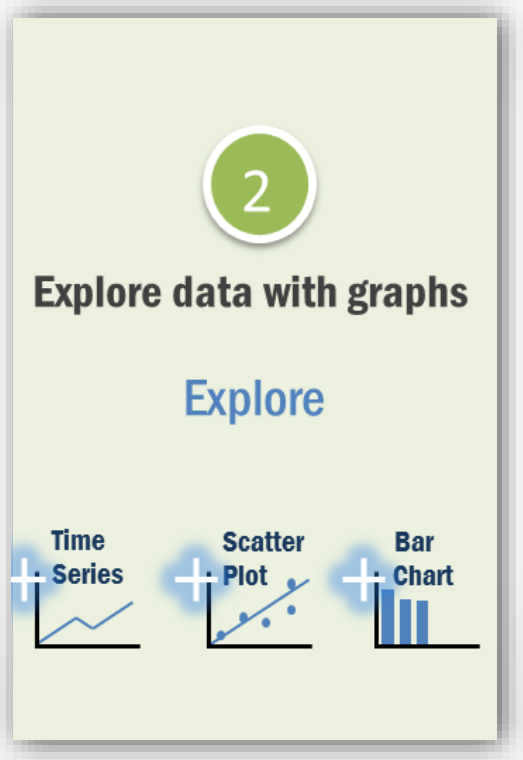
DART Features

- Manage data
  - Upload data files for validation
  - Request data from AQS
  - Convert units
  - Aggregate data



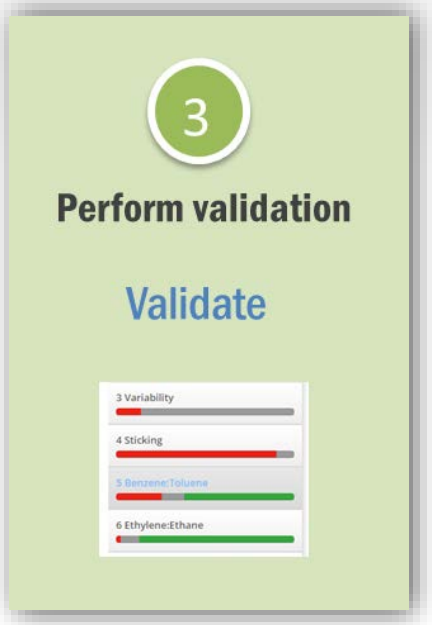
DART Features

- Explore data
  - Create time-series plots
  - Edit data
  - Create scatter plots
  - Create bar plots



DART Features

- Validate data
  - Screen data
  - "DART Smarts"

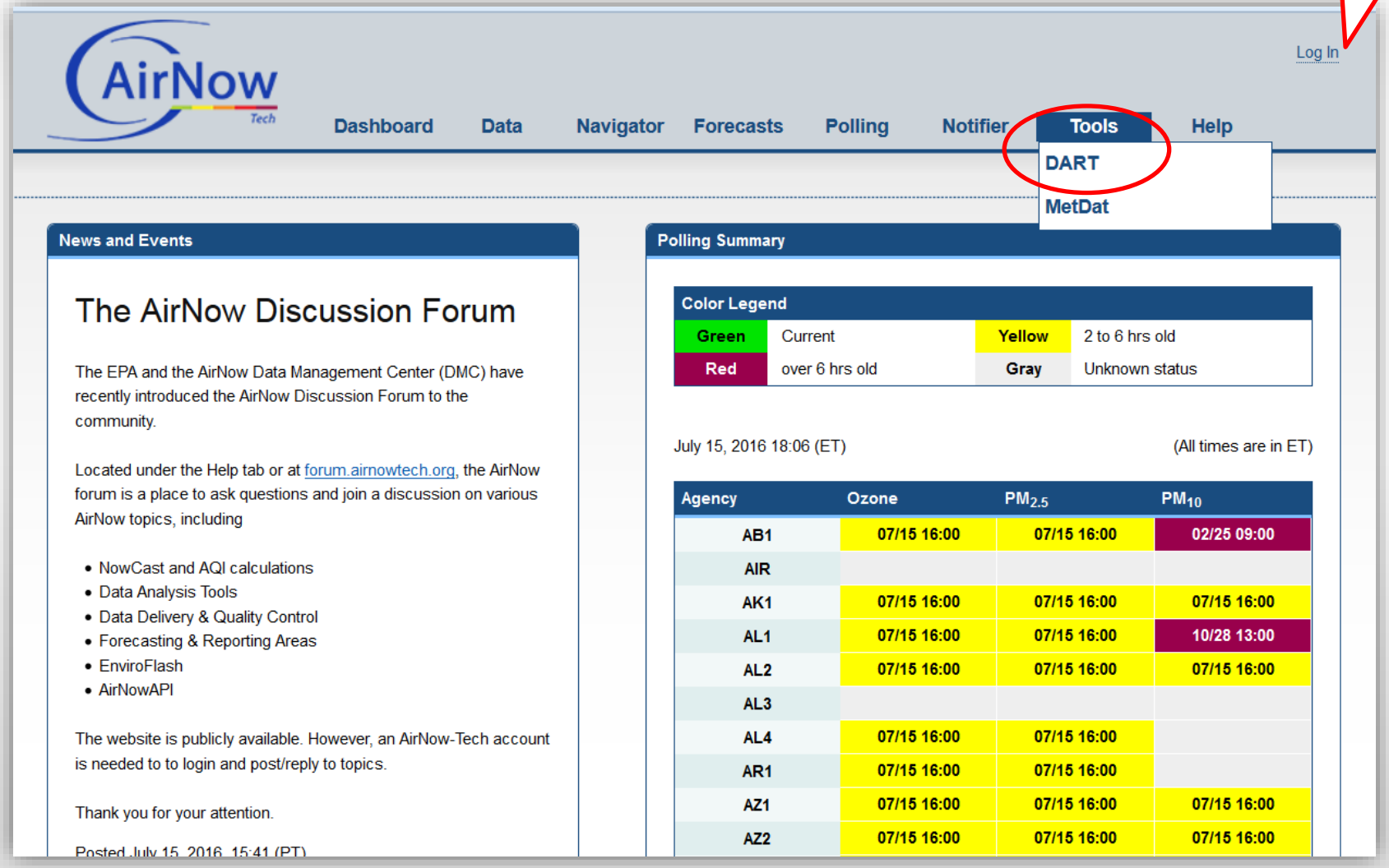


- Export data
  - Create AQS-ready files
  - Obtain summary statistics

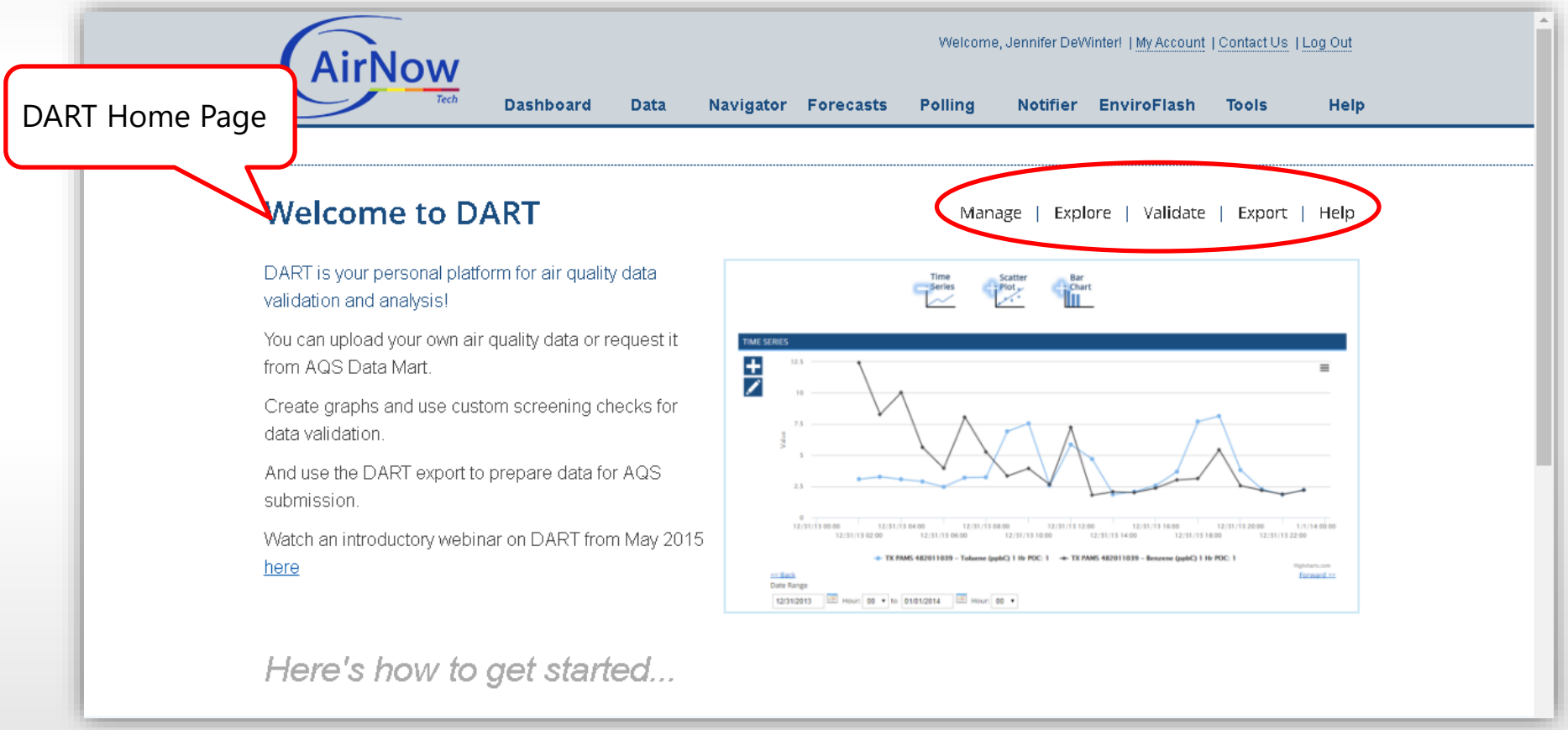
Accessing DART

http://www.airnowtech.org

Requires an AirNow-Tech account



Navigating DART

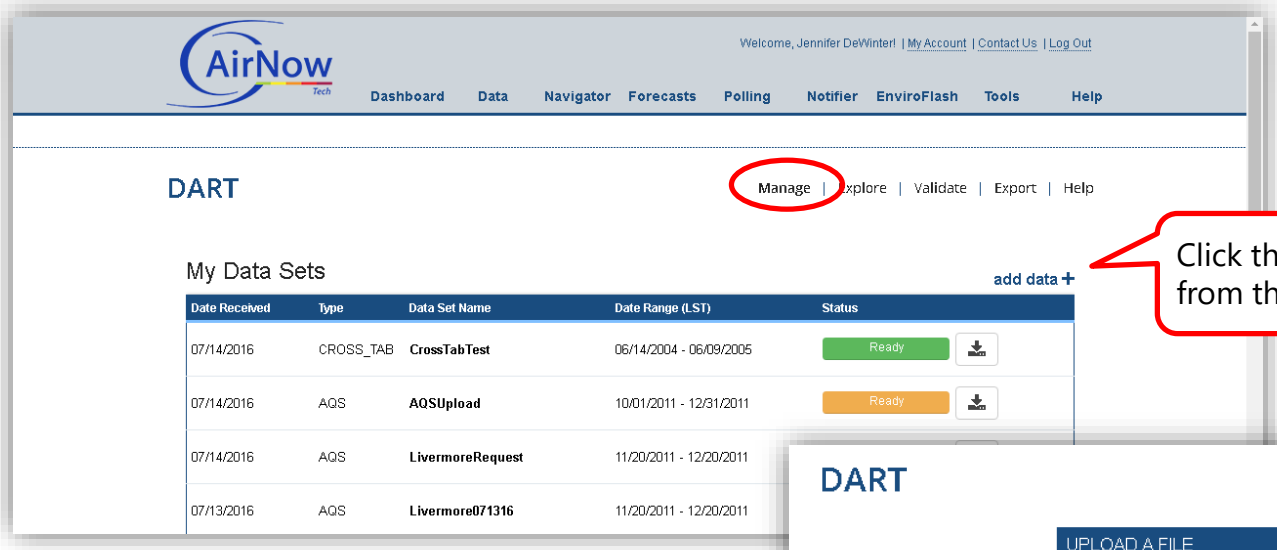


Bringing Data into DART

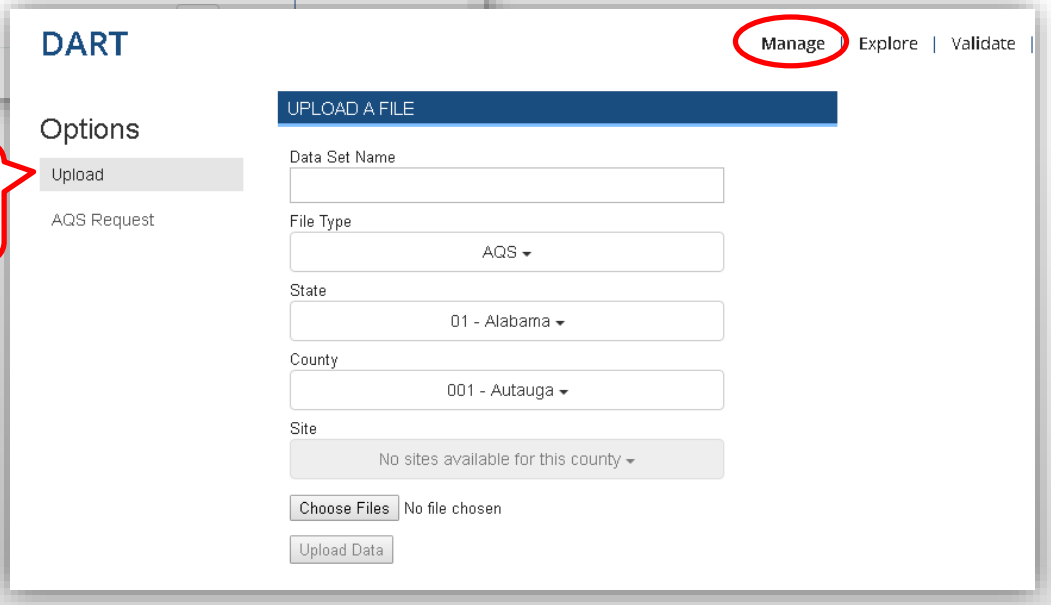
Three ways to add data to DART

1. Upload data files from your computer
   - AQS format
   - CSV format
   - TX0 format
2. Request data from AQS
3. Receive data automatically from a laboratory transfer
   - Currently available for CSN data from UC Davis (UCD)

Bringing Data into DART



Click the "add data" link from the "Manage" page

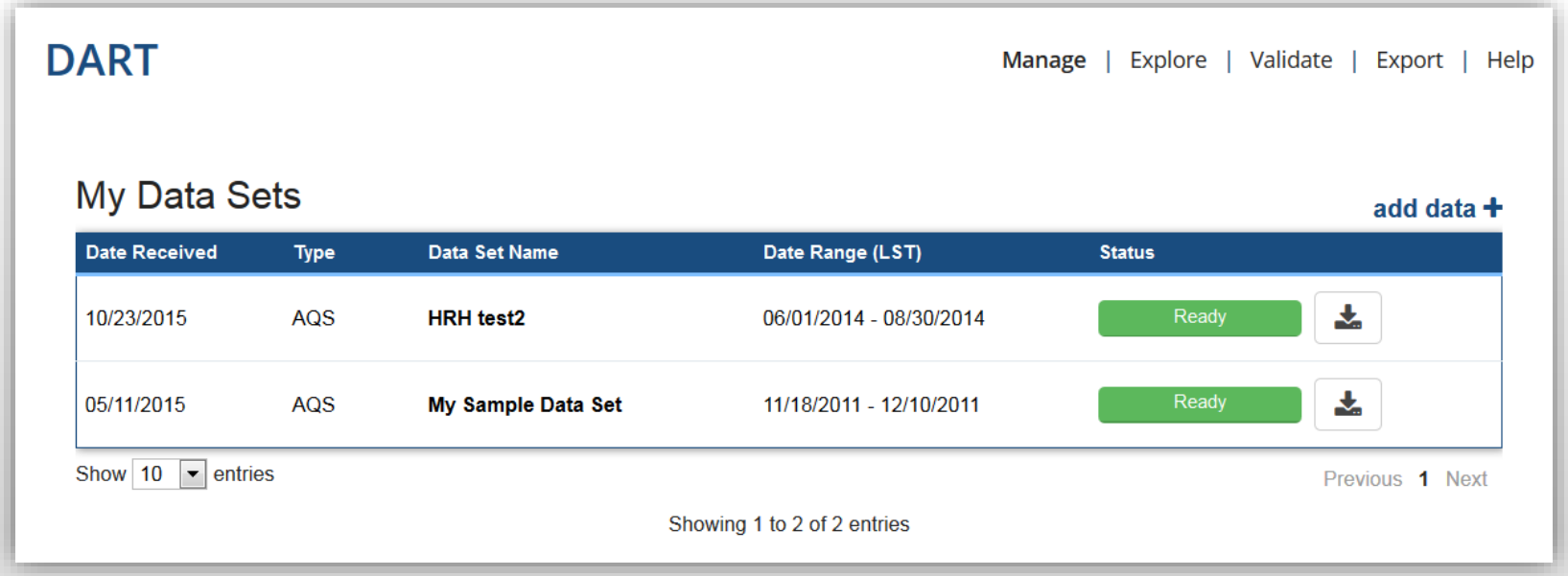


Choose "Upload" or "AQS Request"

Your Data in DART

When you upload your PAMS (or other) data, or make an AQS Request:

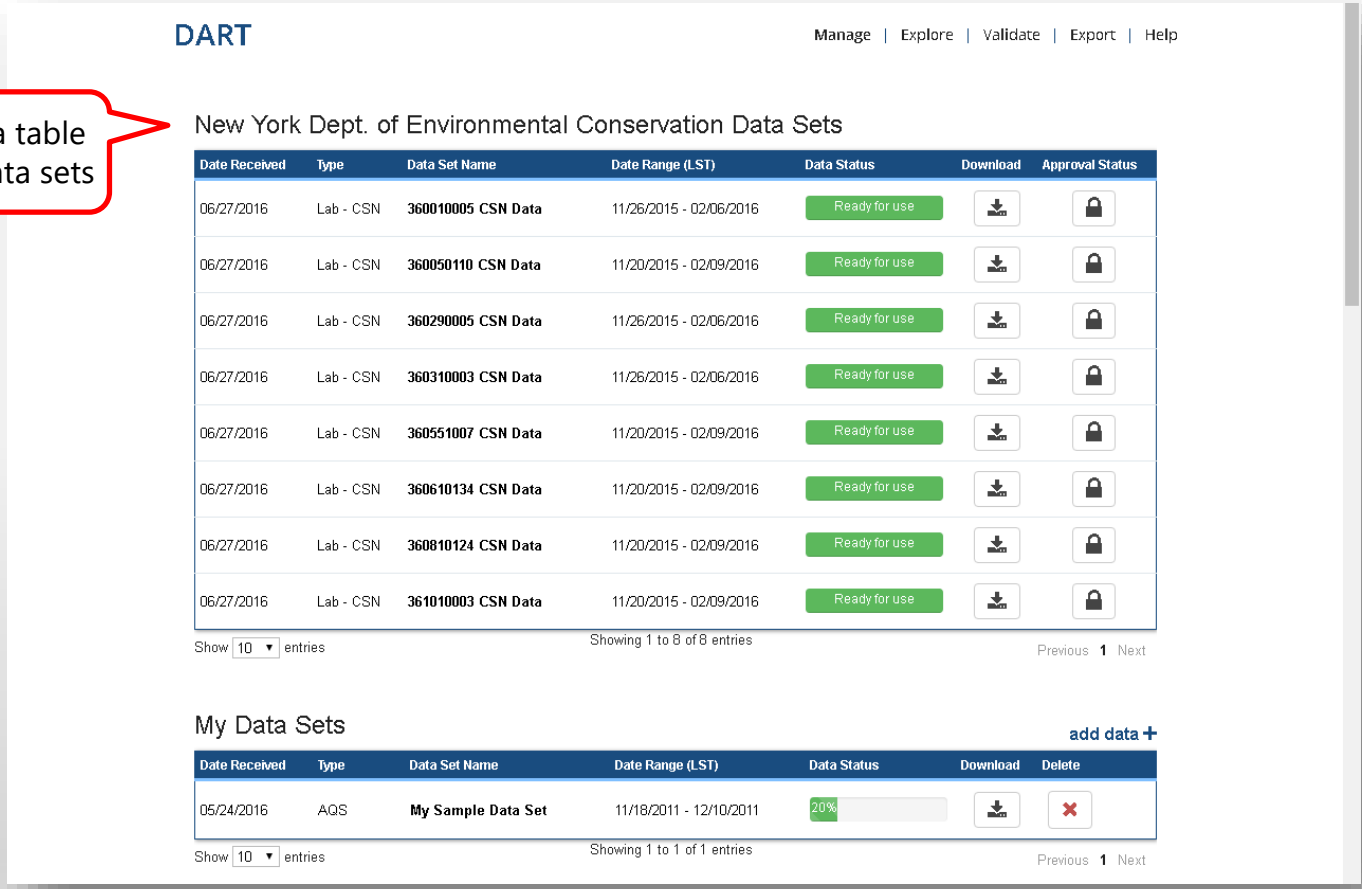
- Currently, it's unique to you
- In future, you will be able to share your data within your agency (and potentially with other AirNow-Tech users)



Your Data in DART

CSN data from UC Davis are available to approved CSN validators in your agency

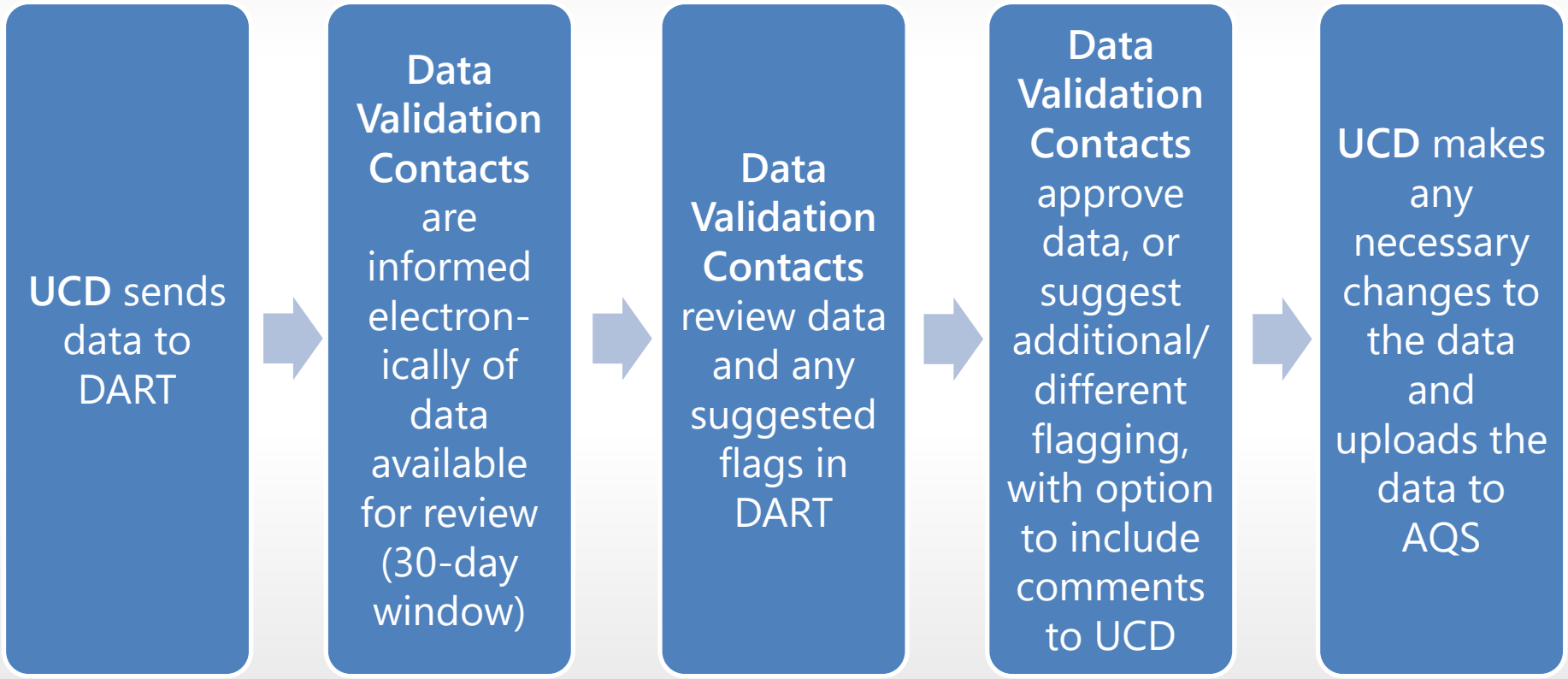
Agency data table with CSN data sets



CSN Data Review Process

1. UC Davis generates laboratory data; acquires mass data from continuous FEMs (as available) from AirNow-Tech.
2. UCD conducts data validation (details on next slide).
3. UCD sends the data batch to DART.
4. DART ingests the batch data, then automatically alerts CSN data reviewers (details on following slides) that data are available in DART.
5. CSN data reviewers then have 30 days to review and approve the data.
6. After 30 days, the data are sent back to UCD for final review and submission to AQS.

DART Data Flow for CSN



CSN Data Validators

- One agency is responsible for data validation for each site (e.g., New York DEC reviews Rochester site data).
- For each site, EPA provided a list of data validators and their associated agency.
- All listed data validators for an individual agency have access to the agency data in DART; they do not have access to data that are not for their agency.
- Once data are available in DART for review, the data validators will be contacted automatically via email.

CSN Data Review in DART

- Once data are in DART, validators have 30 days to review and approve the CSN batch.
- DART provides summary statistics
  - flags and comments that UCD applied
  - other data characteristics (percent complete, percent above detection limit, etc.)
- In DART, validators can flag and comment on data for UCD to review, and use sortable tables to review the CSN batch.
- Time-series and bar plots are linked to the data table so validators can also graph the data.

CSN Data Review in DART

From: Dart Email Notification [mailto:noreply@airnowtech.org]
Sent: Friday, June 24, 2016 10:33 AM
To: Angela L. Ekstrand
Subject: Current batch of CSN data expires in 14 days!

Automated Notification Via Email

Dear DART User,

Please disregard this email if you have already completed your review.

You currently have a batch of CSN data waiting for you to approve in DART. These data are available until 11:59 pm on **Saturday, July 23, 2016**; upon expiration, the data are returned to the laboratory and submitted to AQS.

"Review by"

Your current batch has 2 sites.

QUEENS COLLEGE 2 (360810124) has 14 samples from 11/20/2015 00:00:00 to 12/29/2015 00:00:00.

PINNACLE STATE PARK (361010003) has 14 samples from 11/20/2015 00:00:00 to 12/29/2015 00:00:00.

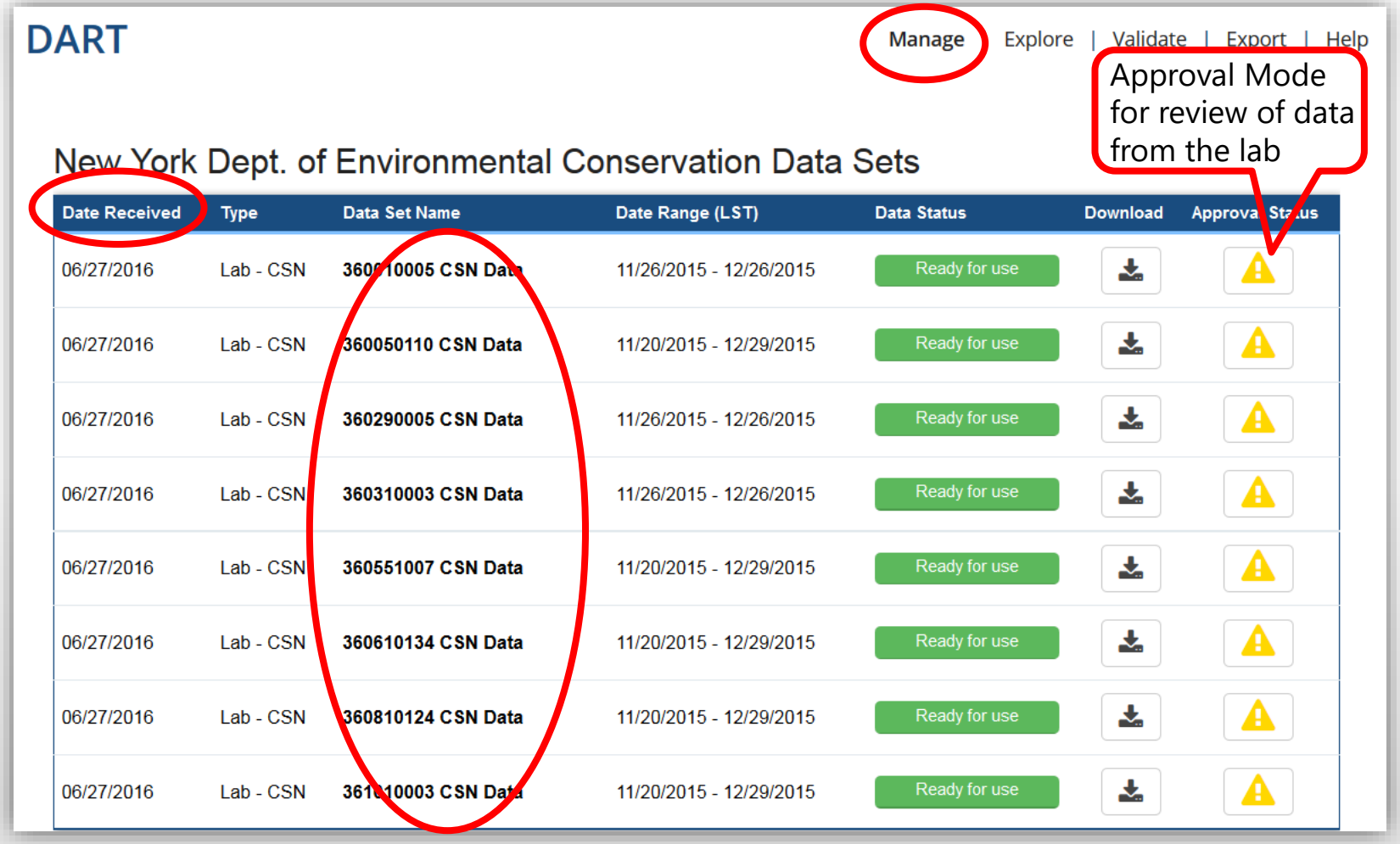
Data sets

Please email CSNsupport@sonomatech.com if you have questions or trouble accessing your data.

Help!

Thanks,
DART Support Team

CSN Data Review in DART



CSN Data Review in DART

Batch Created: 07/29/2016 - Review By: 08/28/2016 23:59

"Review by" Date

Total Samples: 6 **Date Range:** 01/07/2016 - 02/06/2016

Date	Species	Total Qualifiers	Total Null Codes	Data Completeness	Data Above Detection
01/07/2016	47	32	11	77%	32%
01/13/2016	47	45	1	98%	34%
01/19/2016	47	28	1	98%	40%
01/25/2016	47	26	1	98%	45%
01/31/2016	47	30	1	98%	36%
02/06/2016	47	25	1	98%	47%

Flag as reviewed

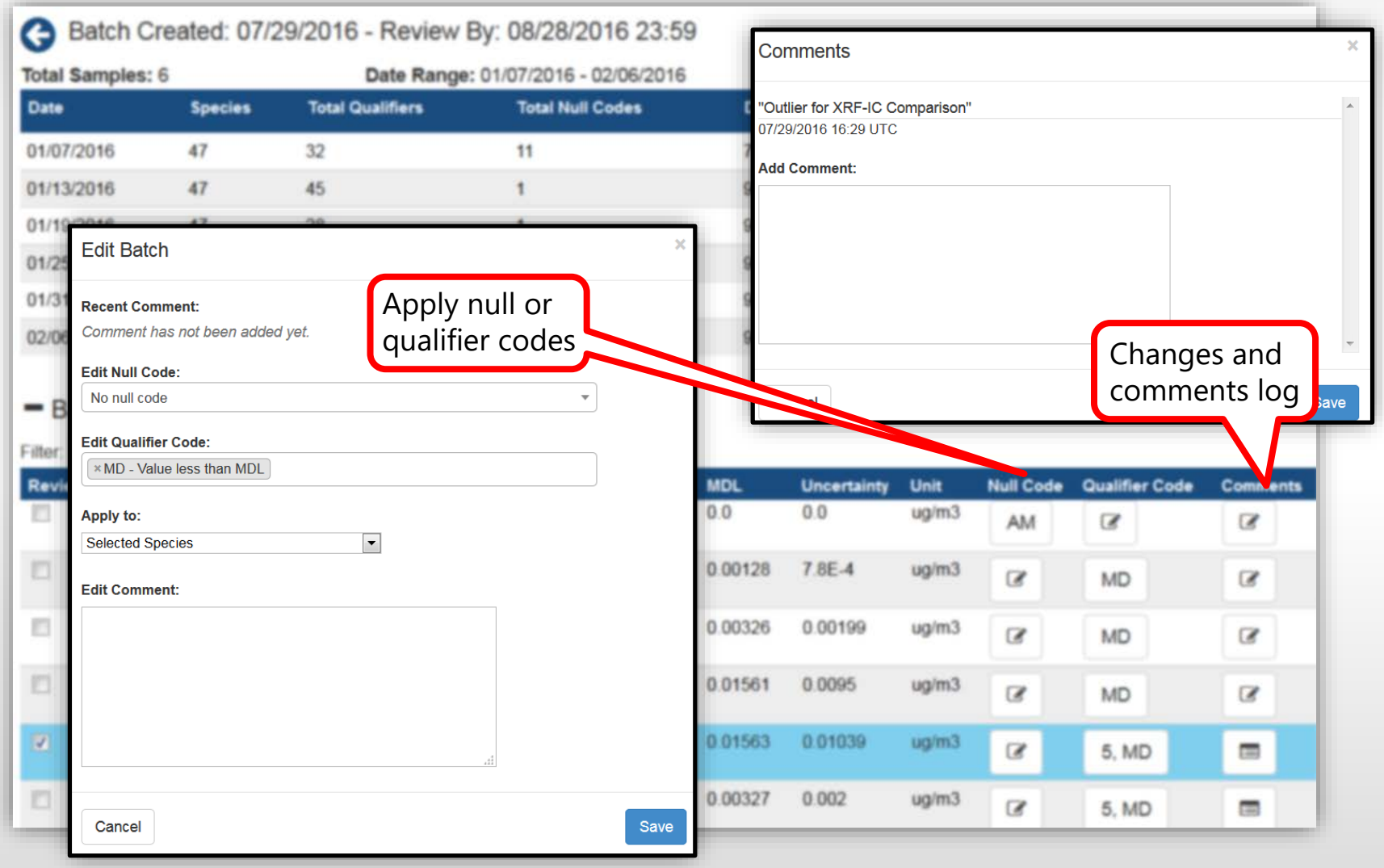
Batch Data

Filter:

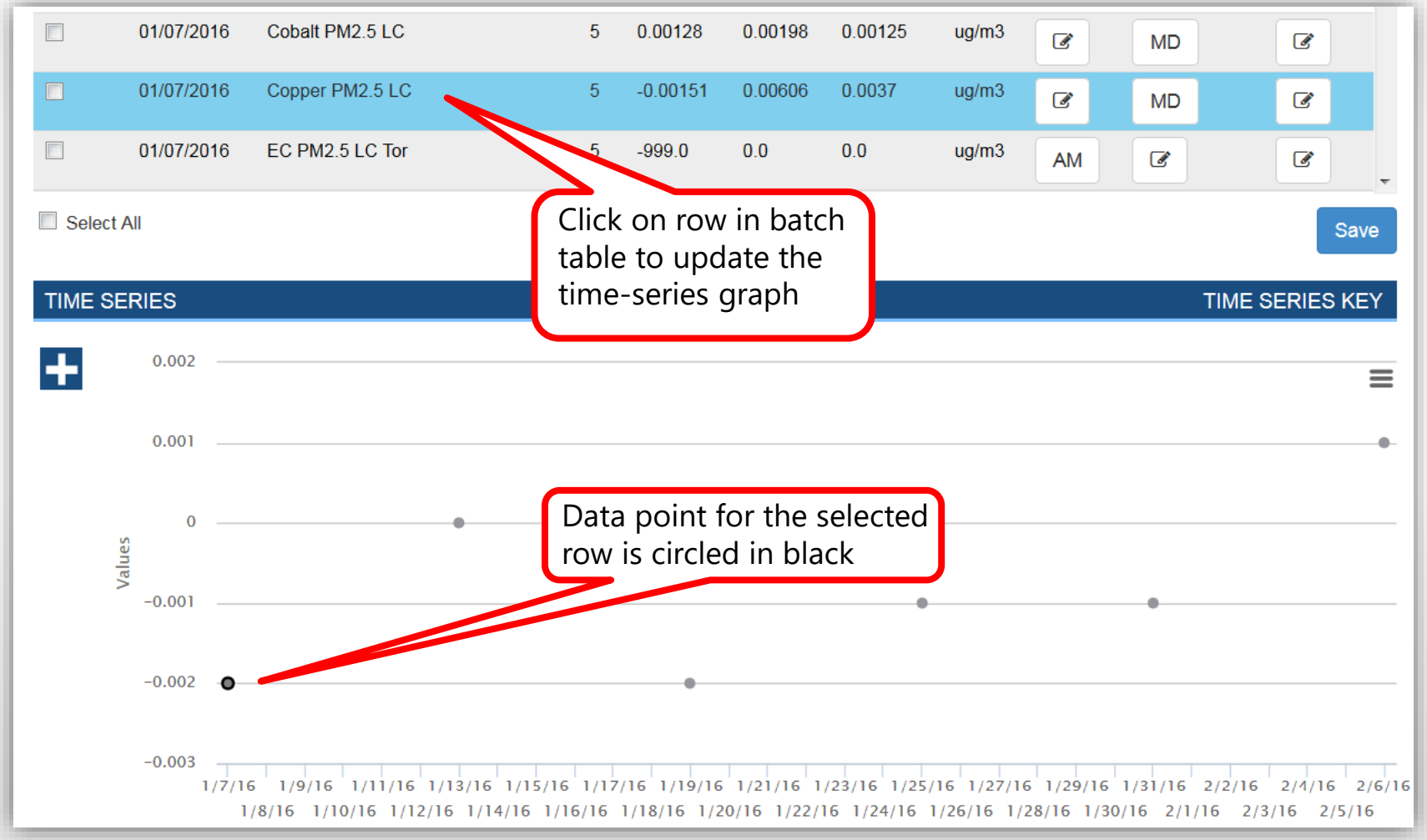
Reviewed	Date	Parameter	POC	Value	MDL	Uncertainty	Unit	Null Code	Qualifier Code	Comments
☐	01/07/2016	Total Nitrate PM2.5 LC	5	-999.0	0.0	0.0	ug/m3	AM		
☐	01/07/2016	Vanadium PM2.5 LC	5	-1.2E-4	0.00128	7.8E-4	ug/m3		MD	
☐	01/07/2016	Zinc PM2.5 LC	5	0.00162	0.00326	0.00199	ug/m3		MD	
☐	01/07/2016	Zirconium PM2.5 LC	5	-0.00198	0.01561	0.0095	ug/m3		MD	
☑	01/13/2016	Zirconium PM2.5 LC	5	-0.01687	0.01563	0.01039	ug/m3		5, MD	
☐	01/13/2016	Zinc PM2.5 LC	5	0.00189	0.00327	0.002	ug/m3		5, MD	

Save

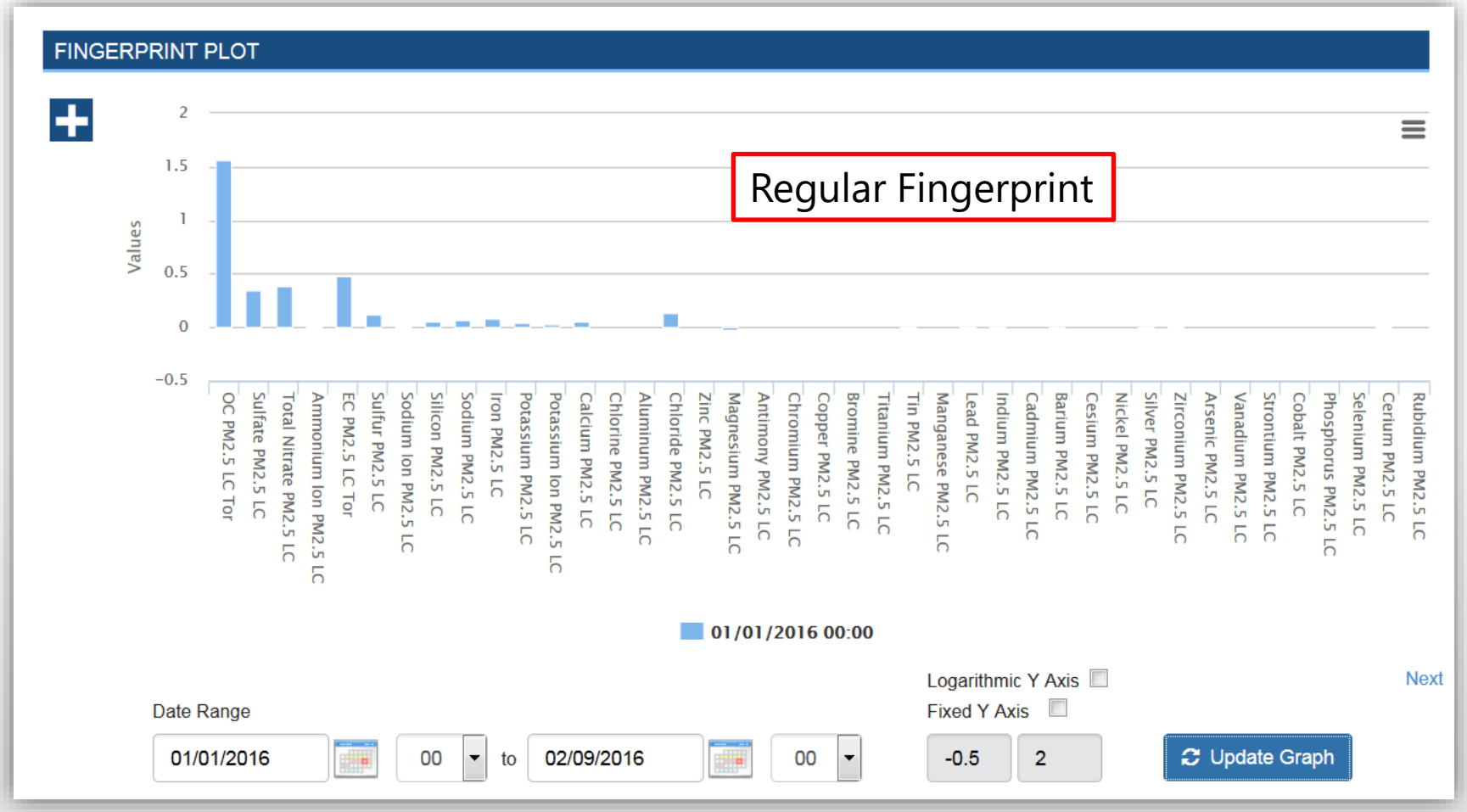
CSN Data Review in DART



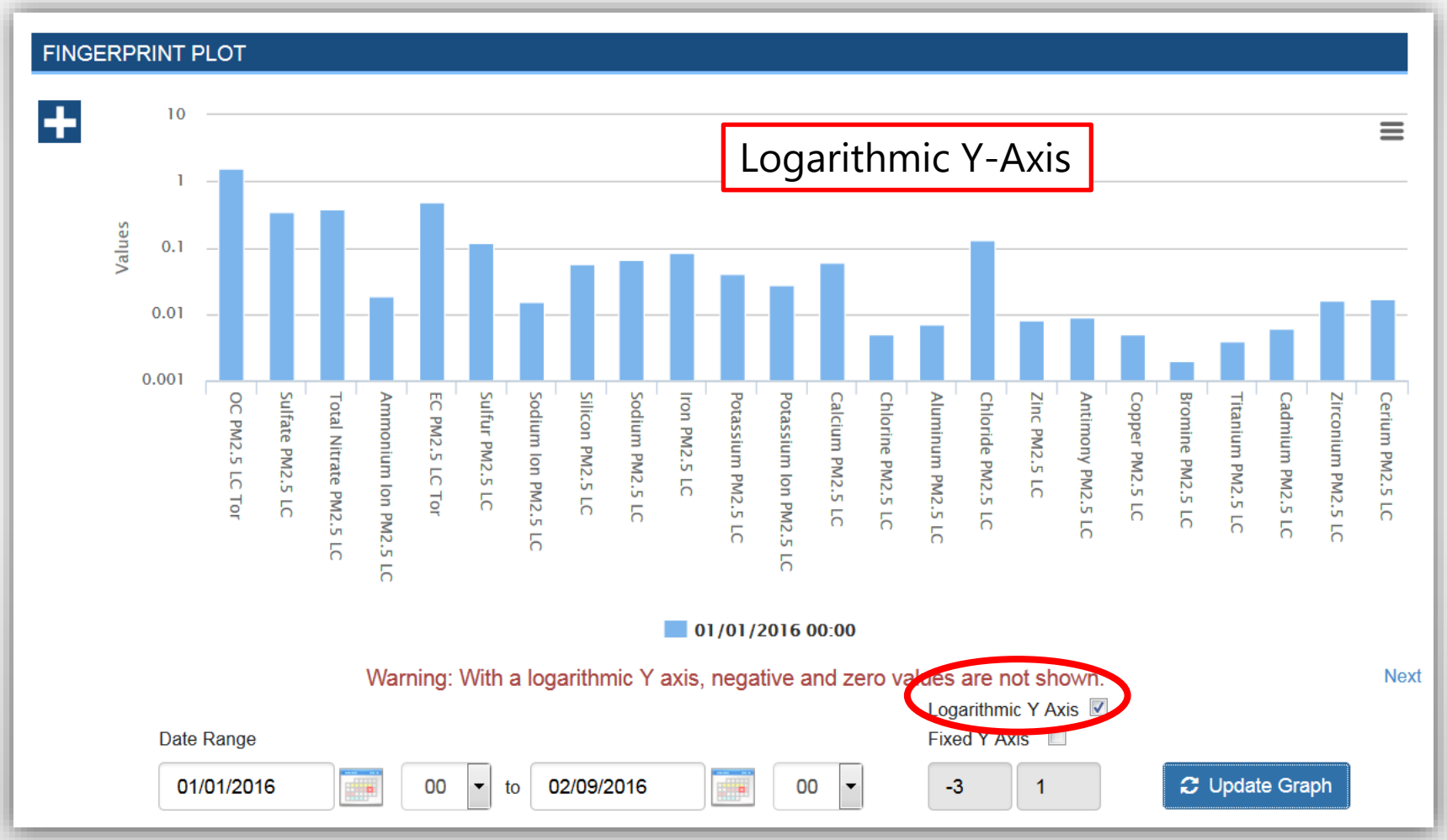
CSN Data Review in DART



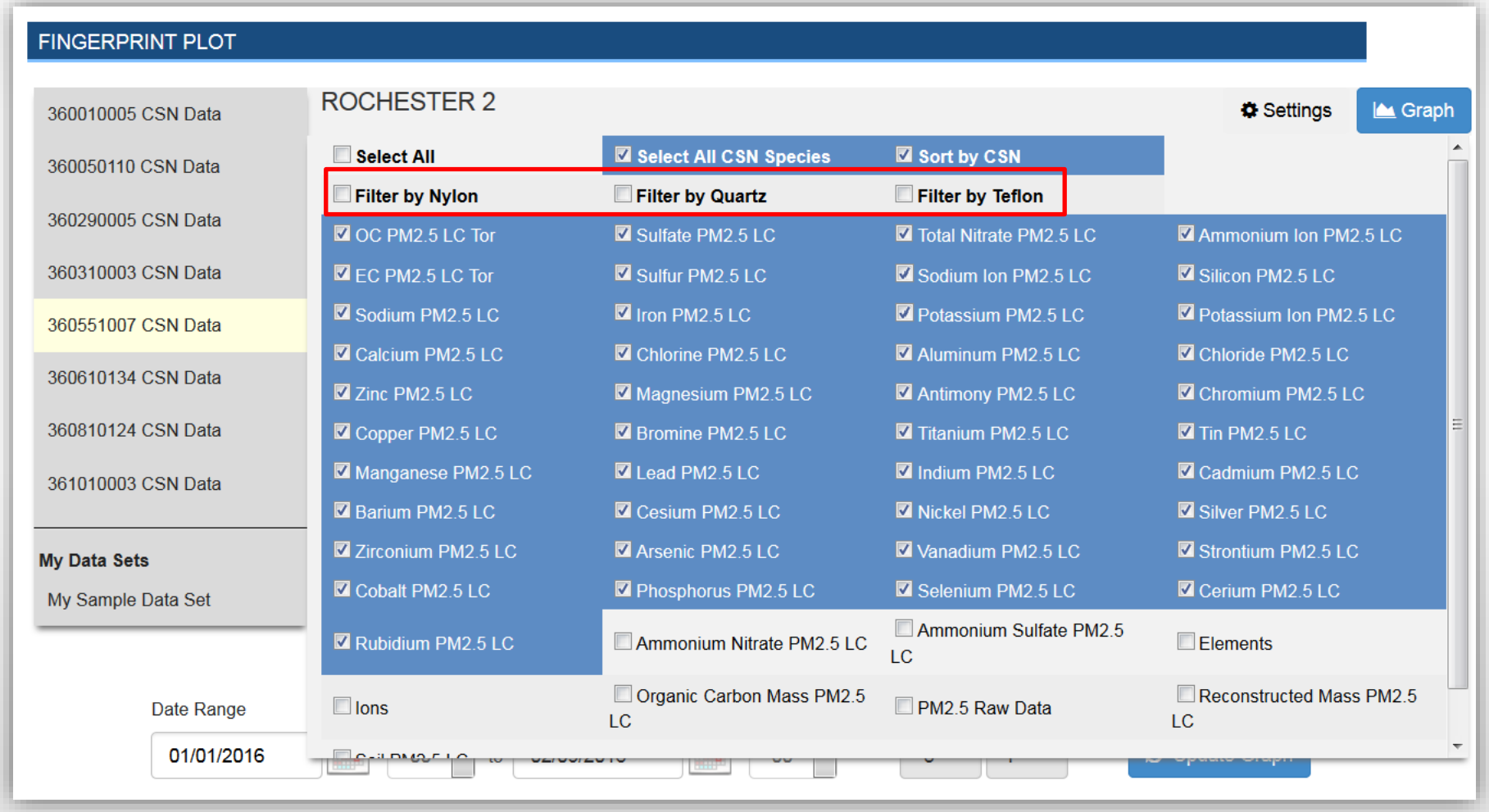
CSN Data Review in DART



CSN Data Review in DART

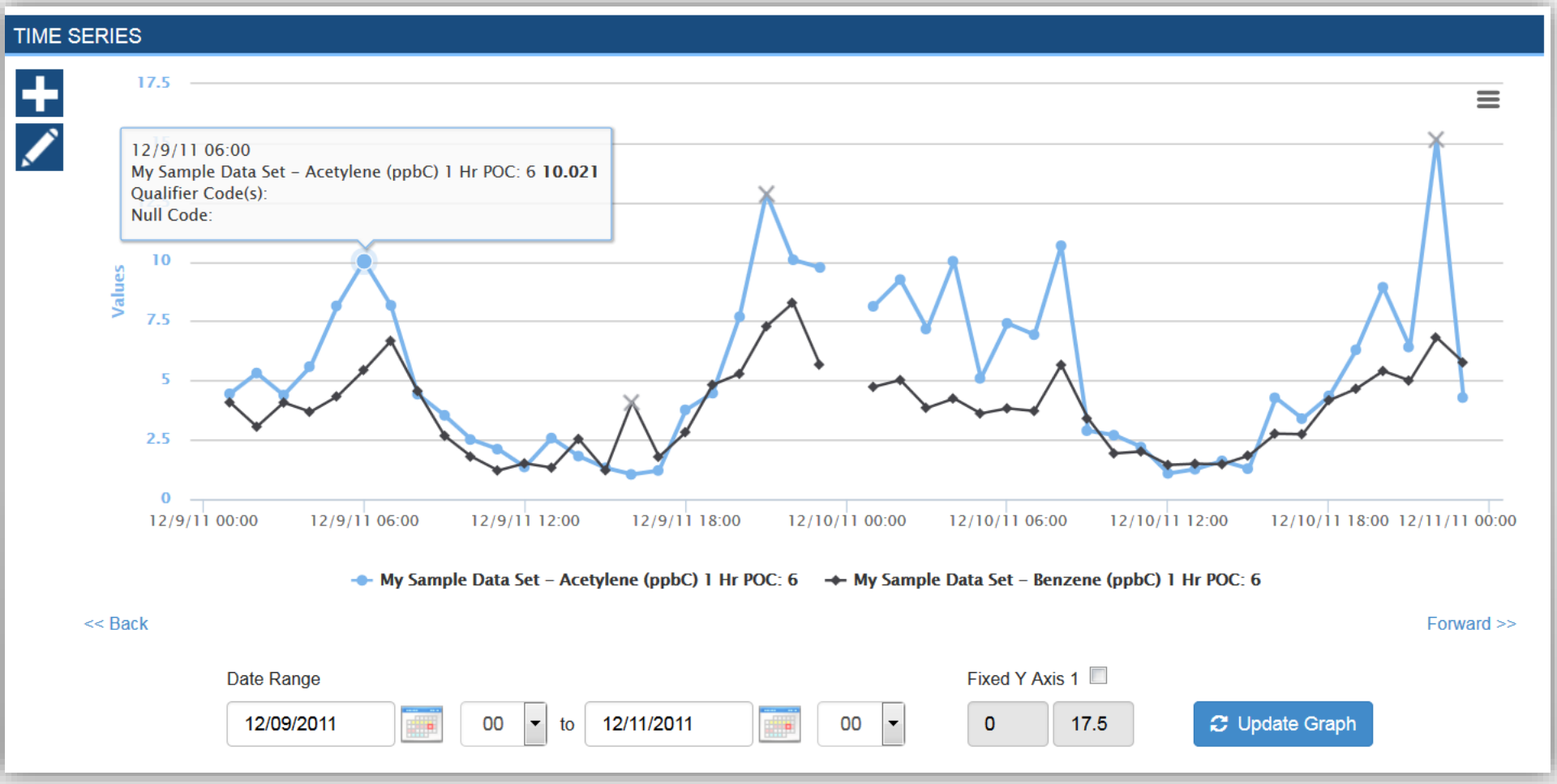


CSN Data Review in DART

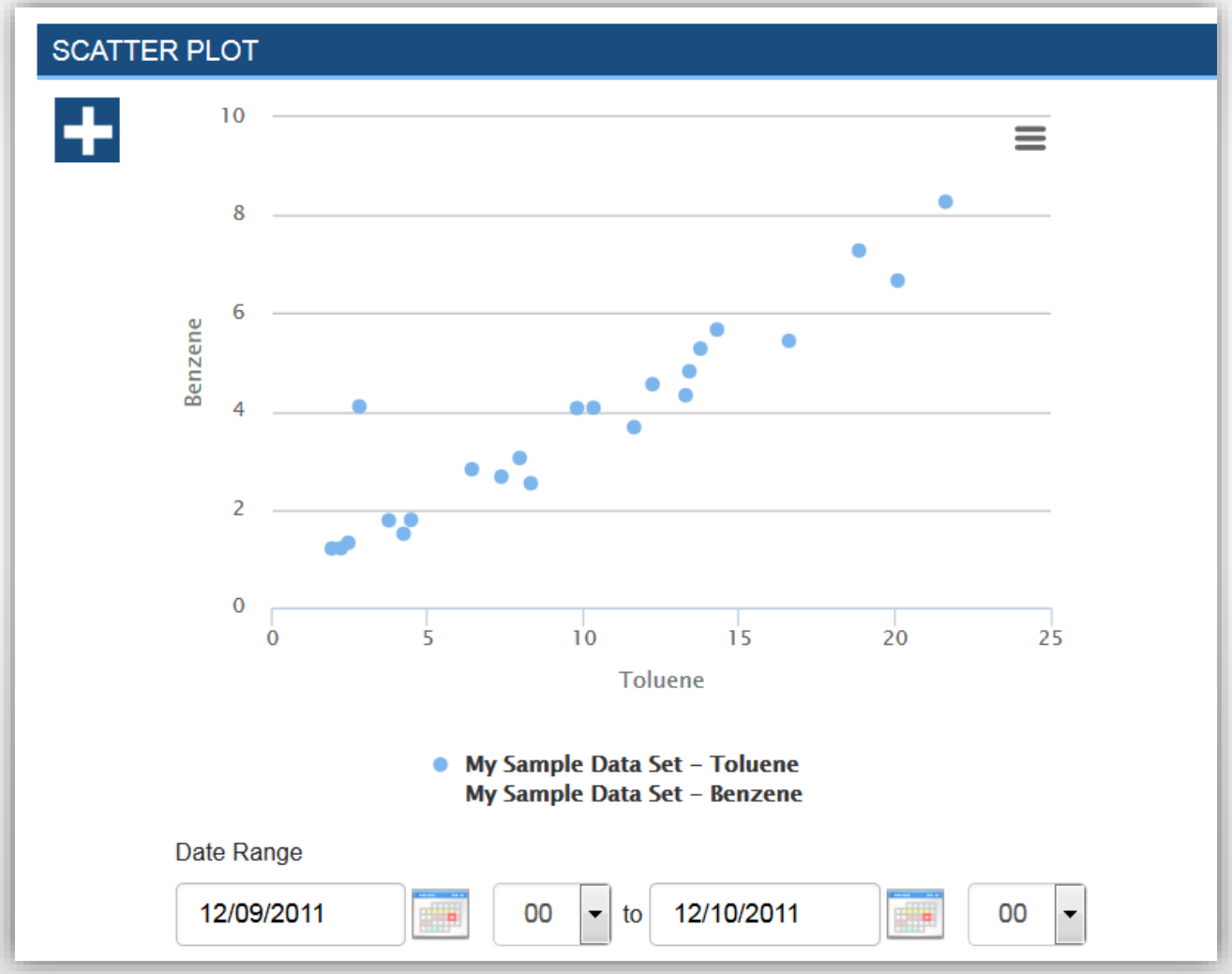


Transition To PAMS Training

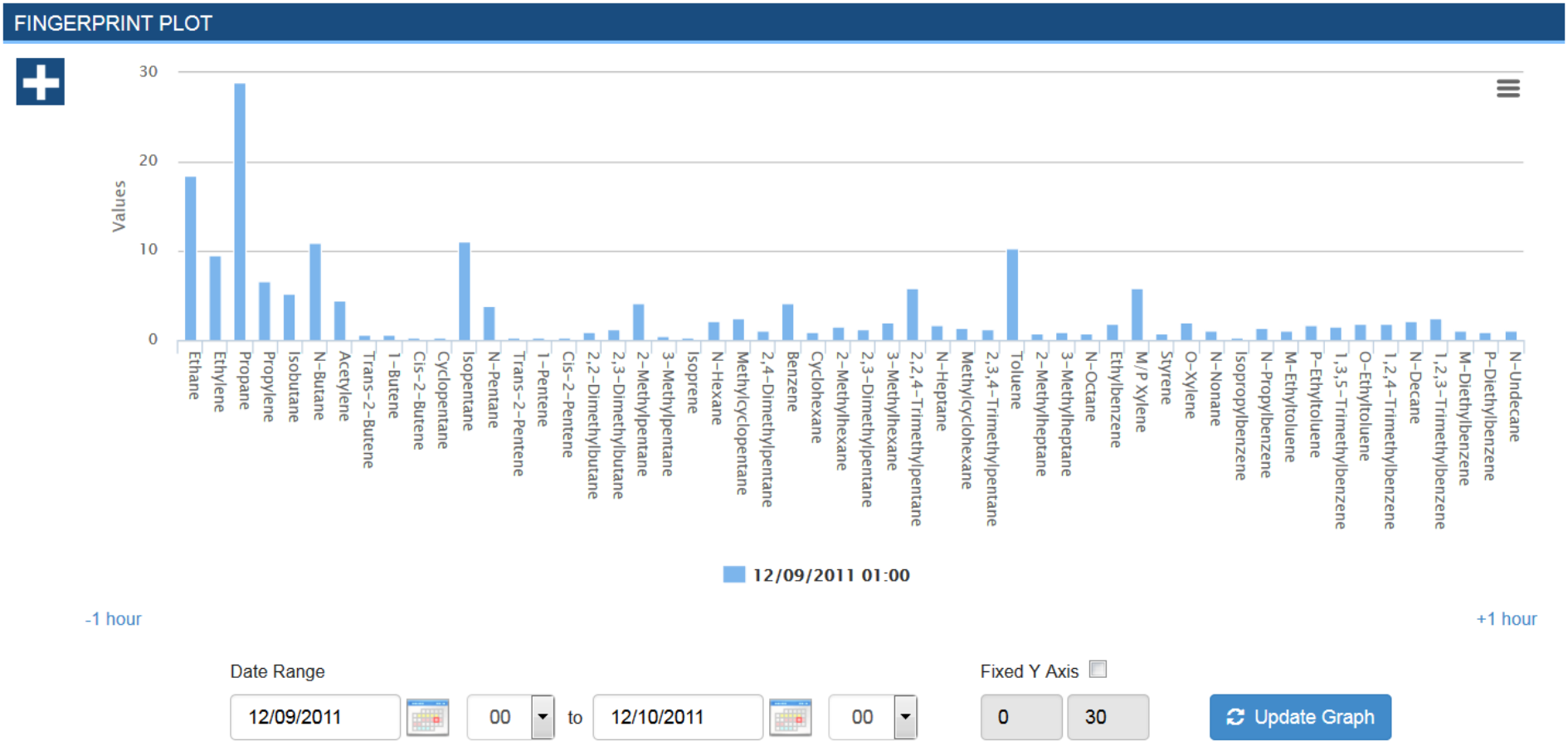
Time-Series Graphs



Scatter Plots



Bar Plots


FINGERPRINT PLOT
30
20
10
0
Values
Ethane
Ethylene
Propane
Propylene
Isobutane
N-Butane
Acetylene
Trans-2-Butene
1-Butene
Cis-2-Butene
Cyclopentane
Isopentane
N-Pentane
Trans-2-Pentene
1-Pentene
Cis-2-Pentene
2,2-Dimethylbutane
2,3-Dimethylbutane
2-Methylpentane
3-Methylpentane
Isoprene
N-Hexane
Methylcyclopentane
2,4-Dimethylpentane
Benzene
Cyclohexane
2-Methylhexane
2,3-Dimethylpentane
3-Methylhexane
2,2,4-Trimethylpentane
N-Heptane
Methylcyclohexane
2,3,4-Trimethylpentane
Toluene
2-Methylheptane
3-Methylheptane
N-Octane
Ethylbenzene
M/P Xylene
Styrene
O-Xylene
N-Nonane
Isopropylbenzene
N-Propylbenzene
M-Ethyltoluene
P-Ethyltoluene
1,3,5-Trimethylbenzene
O-Ethyltoluene
1,2,4-Trimethylbenzene
N-Decane
1,2,3-Trimethylbenzene
M-Diethylbenzene
P-Diethylbenzene
N-Undecane
12/09/2011 01:00
-1 hour
+1 hour
Date Range
12/09/2011
00
to
12/10/2011
00
Fixed Y Axis
0
30
Update Graph

Stacked Bar Charts

Auto-Validation: Screening

Check	Fails If …	DART Smarts Action If Check Fails
Abundant Species	Any of **Benzene**, **Propane**, **N-Butane**, **Isoprene**, **N-Hexane**, **Ethylbenzene** are missing or 0	If two or more species missing or = 0, flag sample with code "AQ"
TNMOC	-**TNMOC** is missing or 0; or -**Unidentified** exceeds 50% of TNMOC; or -**Sum of PAMS** exceeds TNMOC	-Flag **TNMOC** and **Unidentified** with code "AN" -Flag **Unidentified** with code "DA" -Flag **TNMOC** and **Sum of PAMS** with code "DA"
Variability	Species concentration exceeds the mean + 4*standard deviation	None
Sticking	Species has same non-zero value for 3 or more consecutive samples	Flag species with code "DA"
Benzene : Toluene	**Benzene** exceeds 0.2 and exceeds **Toluene**	Flag **Benzene** and **Toluene** with code "DA"
Ethylene : Ethane	**Ethylene** exceeds 0.5 and exceeds **Ethane**	Flag **Ethylene** and **Ethane** with code "DA"
Propylene : Propane	**Propylene** exceeds 0.5 and exceeds **Propane**	Flag **Propylene** and **Propane** with code "DA"

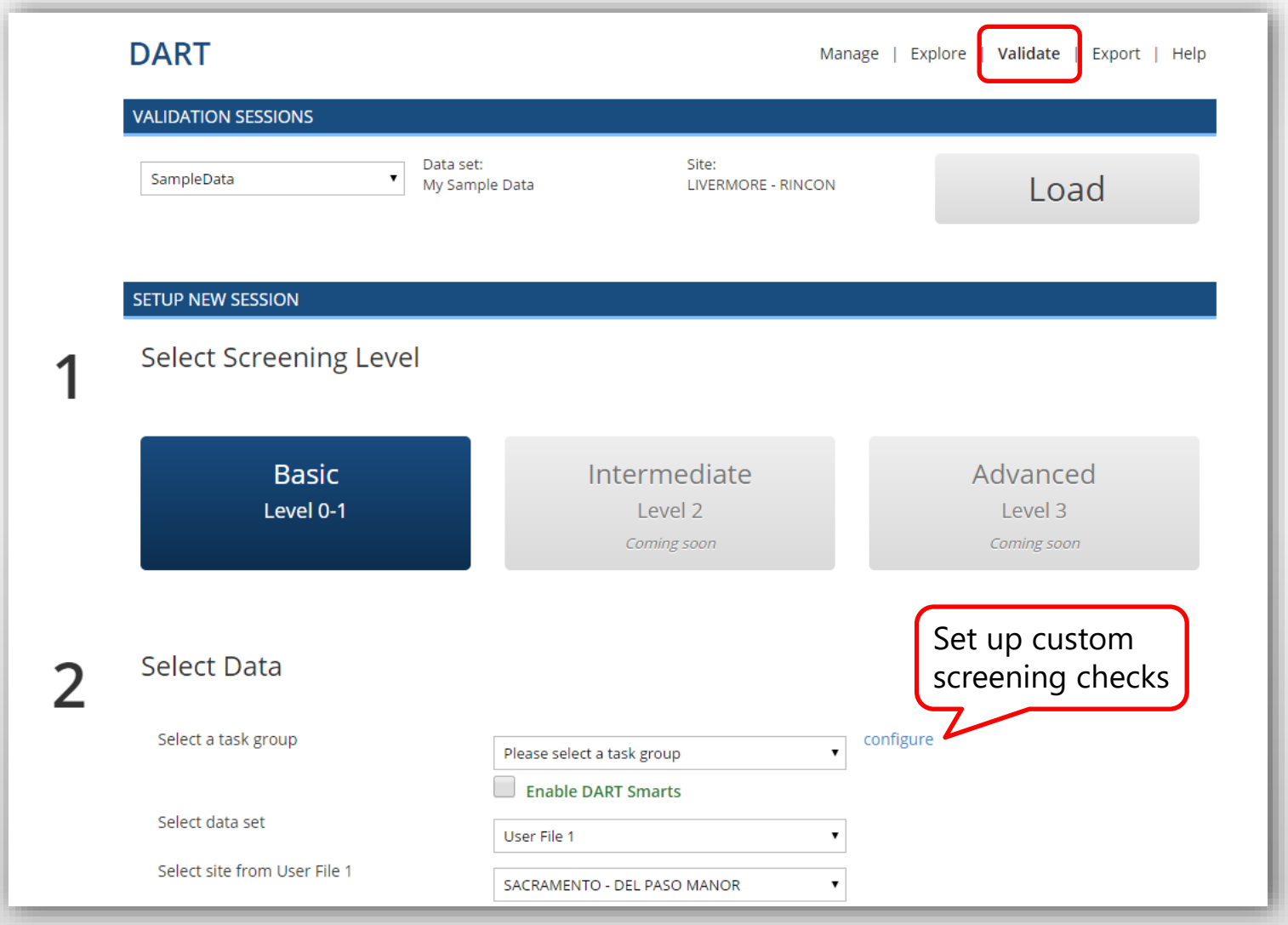
All checks done in ppbC. AQ = collection error; AN = machine malfunction; DA = aberrant data; BH = interference/coelution/misidentification. .

Auto-Validation: Screening

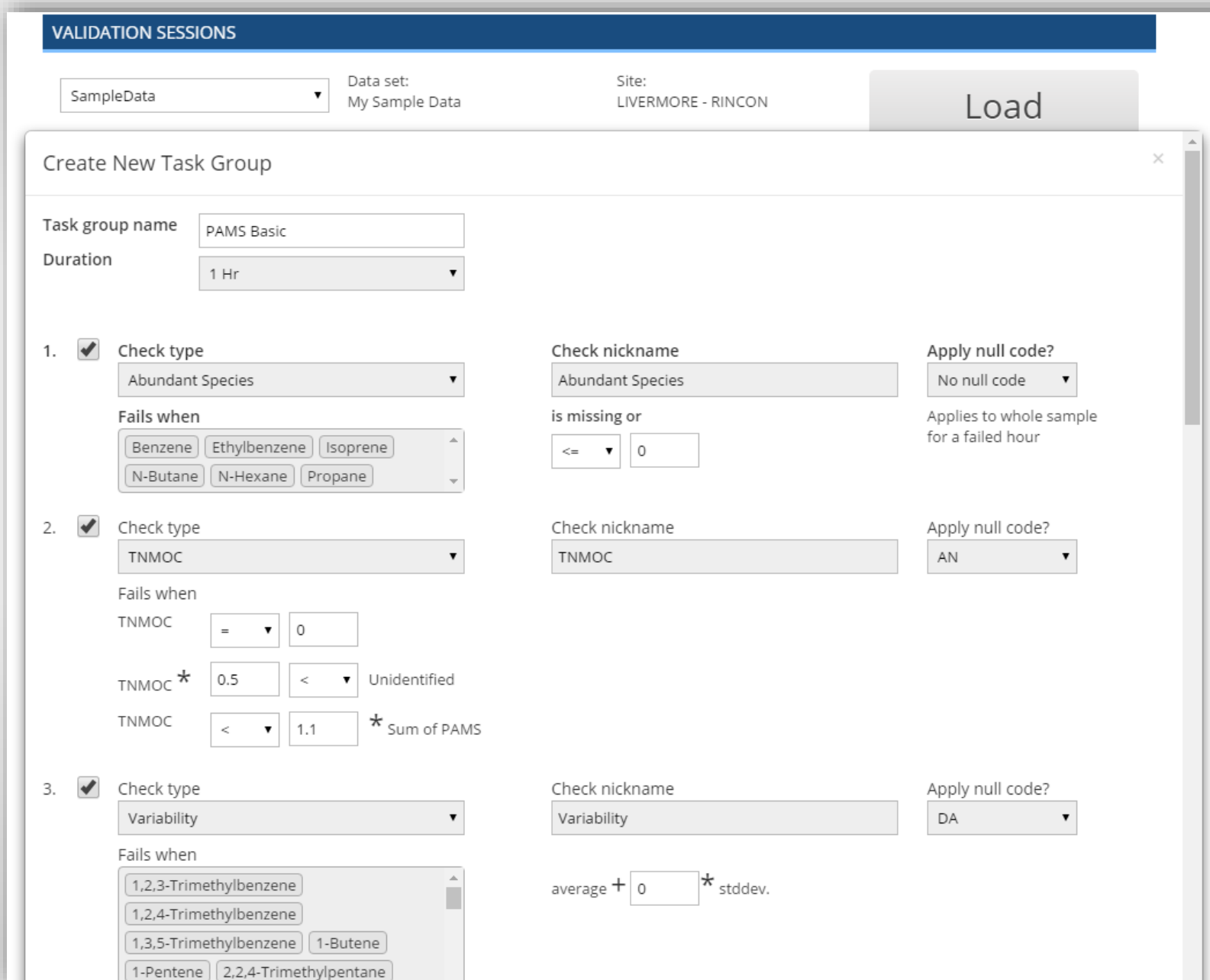
Check	Fails If …	DART Smarts Action If Check Fails
O-Xylene : M/P Xylene	**O-Xylene** exceeds 0.5 and exceeds **M/P Xylene**	Flag **Xylenes** with code "DA"
Methylpentanes	**3-Methylpentane** exceeds 0.1 and exceeds **0.6*2-Methylpentane**	If **3-Methylpentane** exceeds 0.1 and exceeds **0.65*2-Methylpentane**, then flag **Methylpentanes** with code "BH"
Undecane : Decane	**N-Undecane** exceeds 0.5 and exceeds **N-Decane**	Flag **N-Undecane** and **N-Decane** with code "DA"
Olefins : Paraffins	**Sum of Olefins** exceeds **Sum of Paraffins**	Flag **Olefins** and **Paraffins** with code "DA"
Carbon Tetrachloride	**Carbon Tetrachloride** exceeds 0.16 ppb	Flag **Carbon Tetrachloride** with code "AQ"
Formaldehyde	**Formaldehyde** is less than 0.3 ppb	None
Nighttime Isoprene	**Isoprene** increases between 8 pm and 3 am local time	Flag **Isoprene** with code "DA"

All checks done in ppbC, except carbon tetrachloride and formaldehyde checks, which use ppb.. AQ = collection error; AN = machine malfunction; DA = aberrant data; BH = interference/coelution/misidentification.

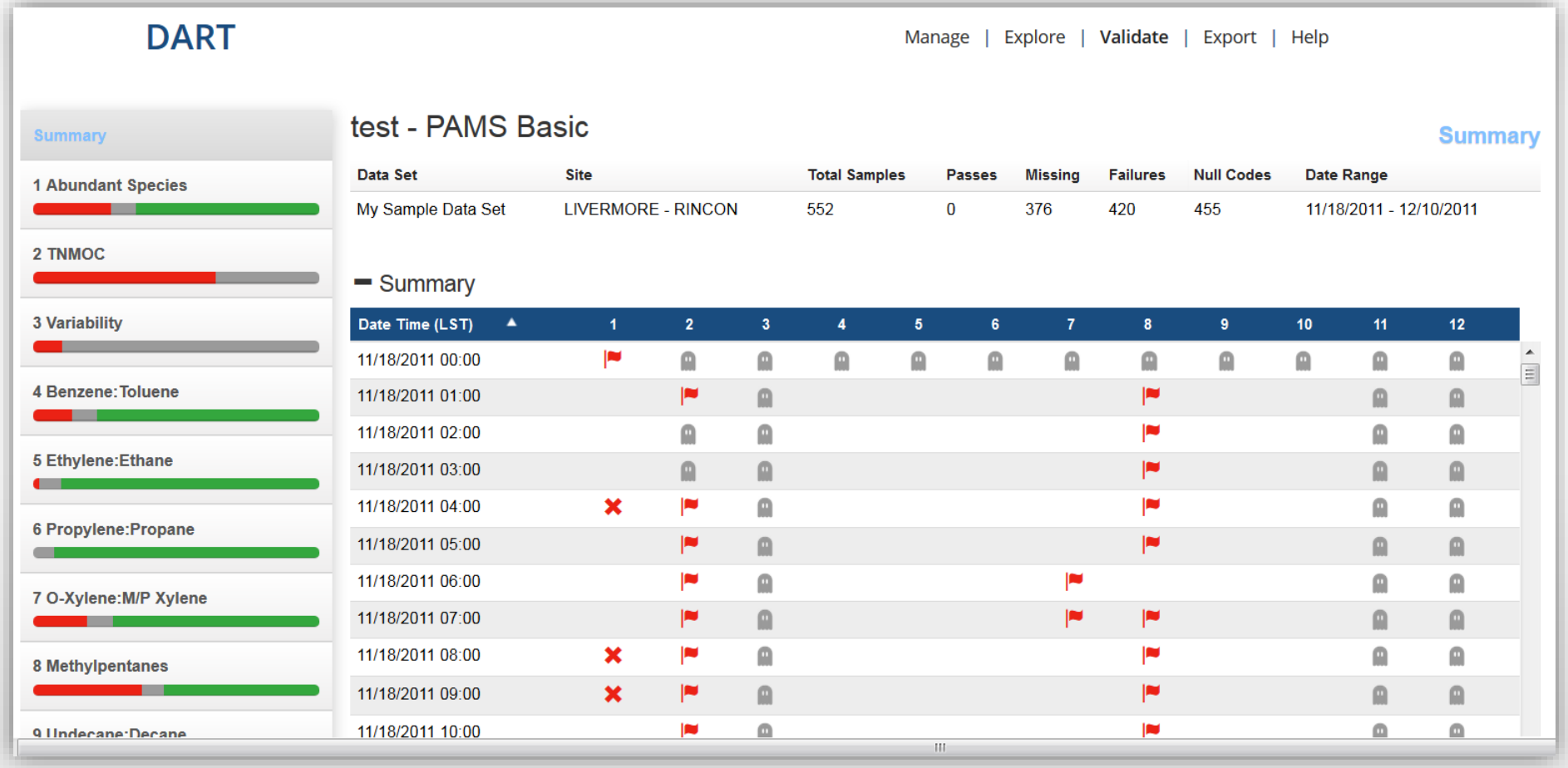
Auto-Validation: Workflow



Auto-Validation: Workflow



Auto-Validation: Workflow



Auto-Validation: Workflow

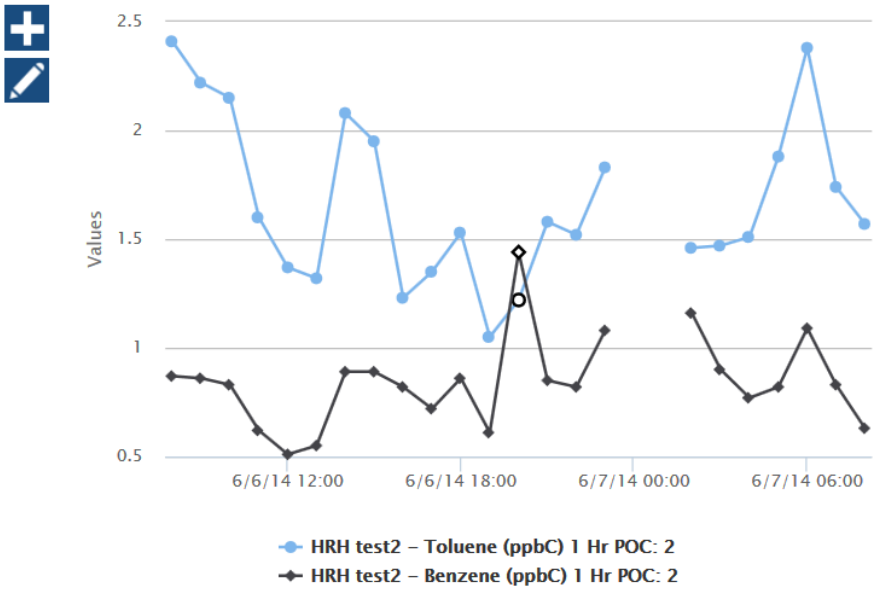
Summary

- 1 Abundant Species
- 2 TNMOC
- 3 Variability
- 4 Benzene:Toluene
- 5 Ethylene:Ethane
- 6 Propylene:Propane
- 7 O-Xylene:M/P Xylene
- 8 Methylpentanes
- 9 Undecane:Decane
- 10 Olefins:Paraffins
- 11 Carbon Tetrachloride
- 12 Formaldehyde

DV clinton drive - PAMS Basic

Benzene:Toluene

Data Set	Site	Total Samples	Passes	Missing	Failures	Null Codes	Date Range
HRH test2	CLINTON	1832	0	175	309	356	06/01/2014 - 08/30/2014



<< Back Forward >>

Date Range

06/06/2014 Hour: 08 to 06/07/2014 Hour: 08

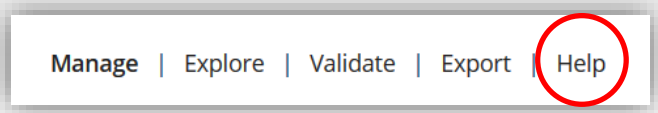
Y Axis

Minimum: 0.5 Maximum: 2.5 Fixed min/max

Filter:

Reviewed	Date Time (LST)	DART Smarts	Failed Param
☐	06/06/2014 20:00		Benze Toluer
✔	07/04/2014 01:00	⚑	Toluer Benze
☐	07/04/2014 14:00		Benze Toluer
☐	07/04/2014 15:00		Toluer Benze
☐	07/04/2014 18:00		Benze Toluer
☐	07/05/2014 09:00	⚑	Toluer Benze
☐	07/05/2014 10:00		Benze Toluer
☐	07/05/2014 11:00	⚑	Toluer Benze
☐	07/05/2014 13:00	⚑	Benze Toluer
☐	07/06/2014 08:00		Toluer Benze
☐	07/12/2014 15:00	⚑	Benze

Getting Help



DART

Search

Contents

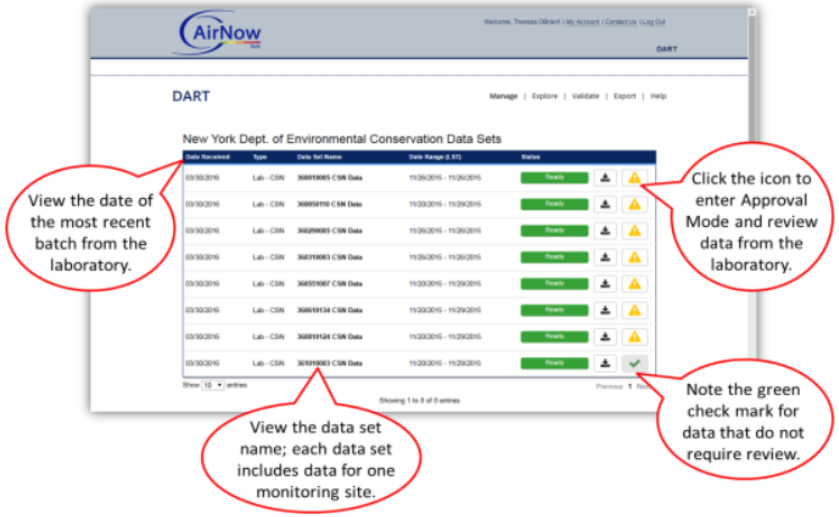
1. About DART
2. The Importance of Data Validation
3. Getting Started
4. System Overview
4.1 Manage
4.1.1 Supported File Formats
4.1.2 Agency Data Sets from a Labo
4.2 Explore
4.3 Validate
4.4 Export
5. More Information
Acknowledgments
Glossary and Acronyms

(Previous Topic: 4.1.1 Supported File Formats)

4.1.2 Agency Data Sets from a Laboratory

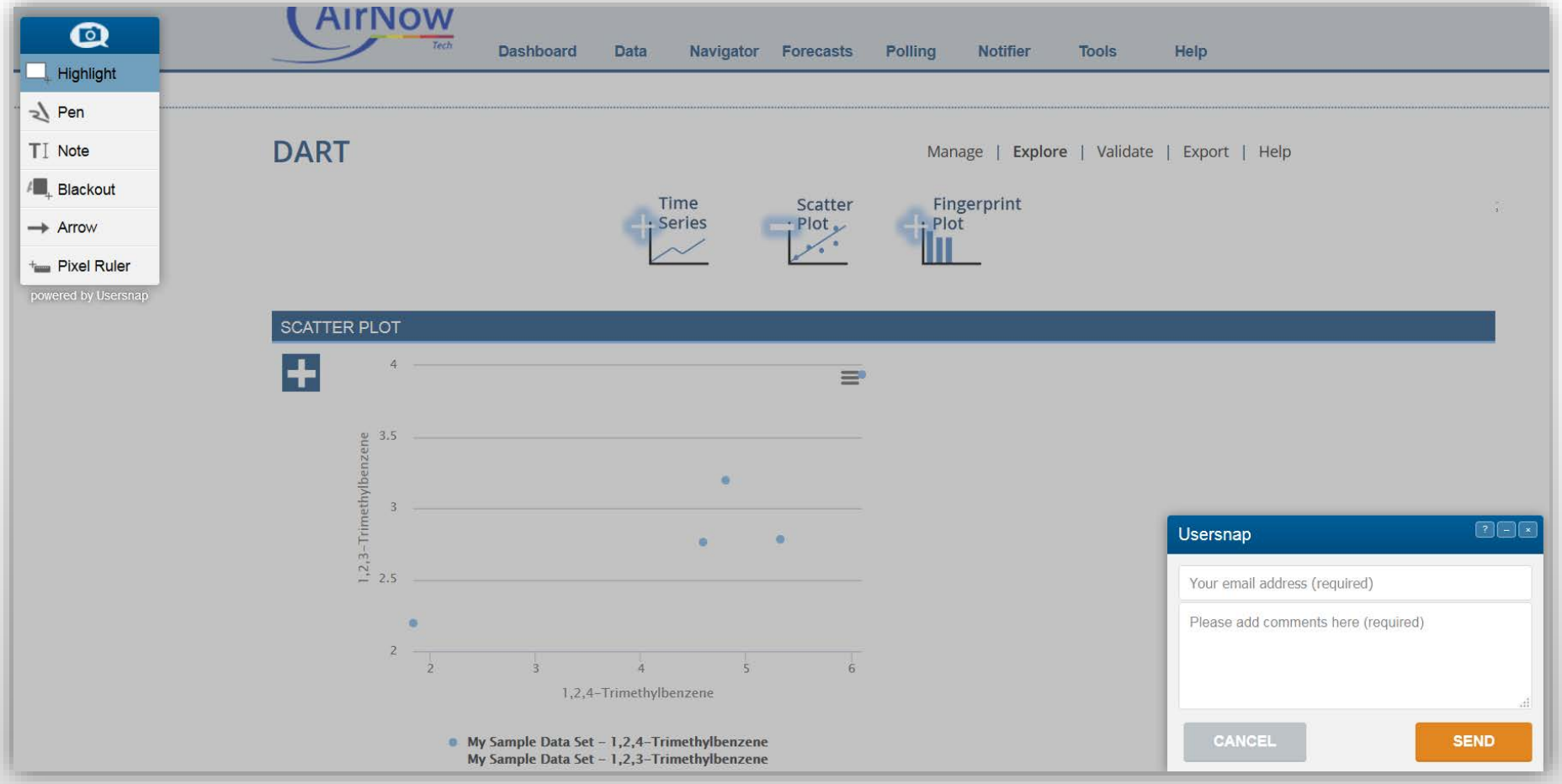
In addition to ingesting file uploads and AQS requests, DART can ingest data directly from an air quality laboratory. Data from a laboratory can be provided to DART via File Transfer Protocol (FTP) and are automatically made available in the correct DART user accounts.

Currently, $PM_{2.5}$ speciation data collected as part of the Chemical Speciation Network (CSN) program are transferred to DART from the Crocker Nuclear Laboratory at the University of California, Davis. CSN data are listed in the **Agency** table on the **Manage** screen, in the DART user accounts that are registered to the appropriate agency. Data are received from the laboratory in batches and are available for review and validation using the DART **Approval Mode** screen.



Getting Help

Feedback button



Important CSN Contacts

CSNsupport@sonomatech.com

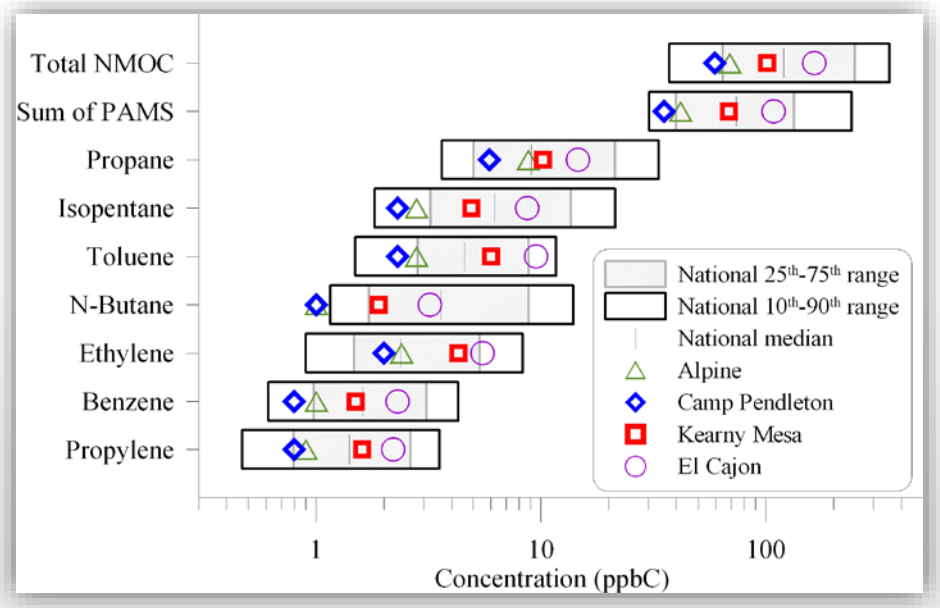
Role	*Contact*	*Phone Number*	*Email*
EPA Project Officer	Jeff Yane	919-541-2962	yane.jeff@epa.gov
EPA Project Manager	Elizabeth Landis	919-541-2262	landis.elizabeth@epa.gov
EPA QA Officer	Jenia Tufts	919-541-0371	tufts.jenia@epa.gov
Shipping & Handling (Amec)	Justin Knoll	352-333-6621	justin.knoll@amecfw.com
Gravimetric Analysis (Amec)	Bill Barnard	352-333-6617	bill.barnard@amecfw.com
Laboratory Analysis (UCD)	Nicole Hyslop	530-754-8979	nmhyslop@ucdavis.edu
DART (STI)	Jennifer DeWinter Steve Brown	707-665-9900	jdewinter@sonomatech.com steveb@sonomatech.com

CSN Regional Representatives

R1 - Alan VanArsdale & Catie Taylor
R2 – Mazeeda Khan
R3 – Lori Hyden
R4 – Keith Harris
R5 – Scott Hamilton
R6 – Frances Verhalen
R7 – Leland Grooms
R8 – Joshua Rickard
R9 – Anna Mebust & Dena Vallano
R10 – Chris Hall & Keith Rose

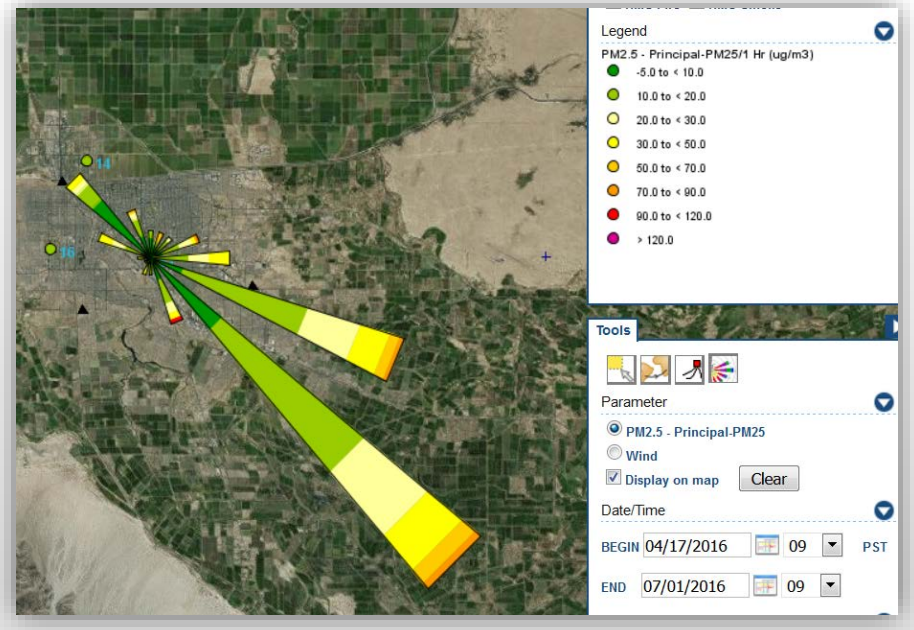
Potential New DART Features

- Interactive map for Data Mart AQS requests
- Suite of automated screening checks for air toxics, CSN data
- Compare site data to national statistics
- New analyses and plot types



Potential New DART Features

- Plot concentrations with MDL values
- Plot concentrations with annual averages
- Support for more import file formats
- Connection to AirNow-Tech features
- Pollution and wind roses



Potential CSN Features in DART

- Plot concentrations with MDL values
- Plot concentrations with annual averages
- Compare site data to national statistics
- Stacked bar and scatter plots in approval mode
- Pollution and wind roses

Acknowledgments

- Joann Rice, CSN Program Support, EPA
- National Association of Clean Air Agencies Steering Committee
- Nick Mangus and Robert Coats, AQS Support, EPA

Contact Us

Jennifer DeWinter

DART Project Manager (STI)
jdewinter@sonomatech.com

Kevin Cavender

EPA PAMS Lead
cavender.kevin@epa.gov

Steve Brown

Senior Atmospheric Scientist
sbrown@sonomatech.com

Beth Landis

EPA CSN Lead
landis.elizabeth@epa.gov

STi
Sonoma Technology, Inc.
Environmental Science and Innovative Solutions

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY