



# ASK

## Analytics for Sandia Knowledge

Judy Spomer

Manager, Software Systems R&D (9525)



# What is ASK?



## ***Vision***

A decision support system  
that will enable timely,  
informed, and integrated  
decisions



***Discover, Analyze,  
Visualize Information***

**C I O** SERVICES EXPO



Sandia National Laboratories



# Why ASK?

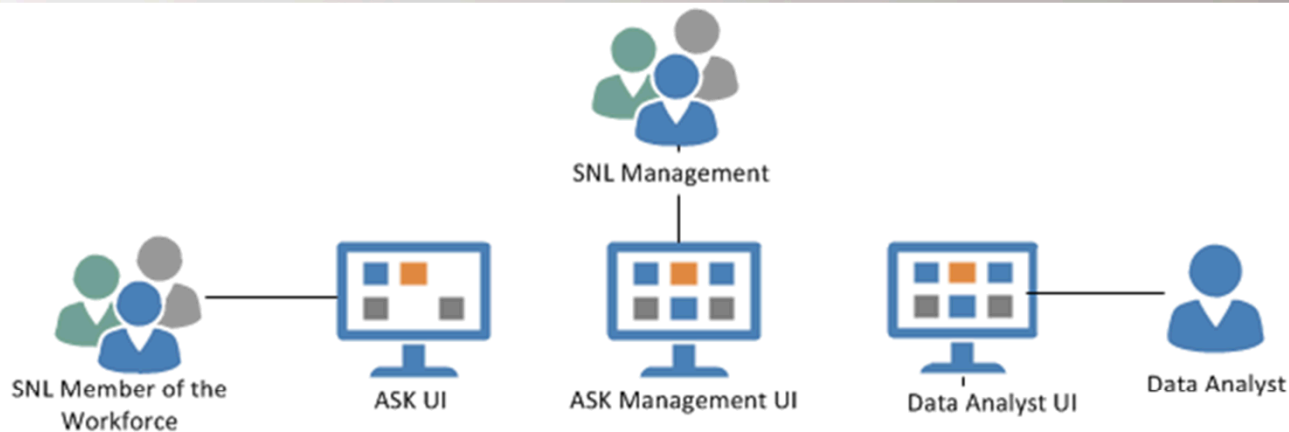
## ***PROBLEMS***

- Can't find information
- Can't integrate data
- Can't get access to data
- Must Protect Info
- Poor Data Quality
- Analyses take too long



## ***SOLUTIONS***

- Streamline Data Discovery & Acquisition
- Establish Governance
- Enable Analytics
- Provide Natural Language Search



### ASK Data Security & Policy Enforcement

Alerts

Visualization

Query Drilldown

Analytic File Generation

Standard Reports

Ad-hoc Reporting

Statistical Analysis

Data Catalog

### ASK Data Services (e.g. Integration, Transformation, Virtualization, etc.)



CIO SERVICES EXPO



Sandia National Laboratories







# Analytics for Sandia Knowledge

*Enabling timely, informed, and integrated decisions involving complex and interdependent issues.*

**FY14**

Understand Data Analytics Needs

Develop Proof-of-Concept

Predictive Model

Foundational Technologies

**FY15**

Implement Initial Framework and Applications

Establish Data Governance

Deploy Foundational Technologies

**FY16**

Business Intelligence and Self-Service Analytics

Prototype Predictive and Big Data Analytics

Non-Traditional Data Sources

**FY17**

Seamless Integration to Advanced Analytics

Analytics for Big Data

External Data Sources

Big Data Sources

Text Analytics

*data ANALYTICS*

*data GOVERNANCE*

*data MANAGEMENT*

# A Quick Look: the FY14 Proof-of-Concept

Provide capabilities to  
*find, collect, assess, and analyze data*  
to support development of an  
ES&H Predictive Model



Selection	Preview	Quality	Data Analysis	
	Available Data		Description	
<input type="checkbox"/>	<input type="checkbox"/>	<b>Foreign Travel Locations</b>		A view containing summary data of employees that have visited a foreign country
<input type="checkbox"/>	<input type="checkbox"/>	foreign_trv_loc_12mo		Number of foreign travel locations visited in the last 12 months
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	foreign_trv_loc_24mo		Number of foreign travel locations visited in the last 24 months
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	foreign_trv_loc_36mo		Number of foreign travel locations visited in the last 36 months
<input type="checkbox"/>	<input type="checkbox"/>	foreign_trv_loc_48mo		Number of foreign travel locations visited in the last 48 months
<input type="checkbox"/>	<input type="checkbox"/>	foreign_trv_loc_60mo		Number of foreign travel locations visited in the last 60 months
<input type="checkbox"/>	<input type="checkbox"/>	<b>ES&amp;H Incidents</b>		A view containing summary data of ES&H incidents by employee and by the set of a manager's subordinates.
<input type="checkbox"/>	<input type="checkbox"/>	emp_incidents_12mo		Number of ES&H incidents this person has had in the prior 12 months to the OBSERVATION_DATE.
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	emp_incidents_24mo		Number of ES&H incidents this person has had in the prior 24 months to the OBSERVATION_DATE.
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	emp_incidents_36mo		Number of ES&H incidents this person has had in the prior 36 months to the OBSERVATION_DATE.
<input type="checkbox"/>	<input type="checkbox"/>	emp_incidents_48mo		Number of ES&H incidents this person has had in the prior 48 months to the OBSERVATION_DATE.
<input type="checkbox"/>	<input type="checkbox"/>	emp_incidents_60mo		Number of ES&H incidents this person has had in the prior 60 months to the OBSERVATION_DATE.
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_12mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the prior 12 months to the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_24mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the prior 24 months to the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_36mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the prior 36 months to the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_48mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the prior 48 months to the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_60mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the prior 60 months to the OBSERVATION_DATE
<input type="checkbox"/>	<input checked="" type="checkbox"/>	emp_incidents_future_6mo		Number of ES&H incidents this person has had in the 6-month period following the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	mgr_emp_incidents_future_6mo		If this person is or has been a manager, the number of ES&H incidents that occurred in their organization in the 6-month period following the OBSERVATION_DATE
<input type="checkbox"/>	<input type="checkbox"/>	<b>Badging and Hosting</b>		A view representing badge counts and visitor hosting counts within a period of time.
<input type="checkbox"/>	<input type="checkbox"/>	<b>Person and Org Information</b>		A view containing employee human resources data by observation date.
<input type="checkbox"/>	<input type="checkbox"/>	<b>Training Compliance</b>		A view representing training compliance within a period of time.
<input type="checkbox"/>	<input type="checkbox"/>	<b>Expense Report Travel Days</b>		A view containing summary data of employee business travel from the ExpenseReport application.
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<b>NWIS Device Ownership</b>		A view containing summary data of employees and their device ownership
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	all_device_count		Total number of devices assigned to person in NWIS
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	computer_device_count		Total number of computer devices assigned to person in NWIS
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	mobile_device_count		Total number of mobile devices assigned to person in NWIS

Select data from observation date: 06/30/2012 through: 03/31/2013

Export Selected Data



Preview of selected data for observation dates from 06/30/2012 to 03/31/2013

SNL Random Id	Observation Date	average, no. inv./obs	average, no. inv./obs	avg. incidents./obs	avg. incidents./obs	avg. incidents./obs, 95% CI	all
130996569	06/30/2012	1	1	0	0		0
130996569	09/30/2012	1	1	0	0		0
130996569	12/31/2012	1	1	0	0		0
130996569	03/31/2013	1	1	0	0		0
743726412	06/30/2012	0	0	0	0		0
743726412	09/30/2012	0	0	0	0		0
743726412	12/31/2012	0	0	0	0		0
743726412	03/31/2013	0	0	0	0		0
1470973369	06/30/2012	0	0	0	0		0
1470973369	09/30/2012	0	0	0	0		0
1470973369	12/31/2012	0	0	0	0		0
1470973369	03/31/2013	0	0	0	0		0
2048135934	06/30/2012	0	0	0	0		0
2048135934	09/30/2012	0	0	0	0		0
2048135934	12/31/2012	0	0	0	0		0
2048135934	03/31/2013	0	0	0	0		0
2943228614	06/30/2012	0	0	0	0		0
2943228614	09/30/2012	0	0	0	0		0
2943228614	12/31/2012	0	0	0	0		0
2943228614	03/31/2013	0	0	0	0		0
3263327908	09/30/2012						0
3263327908	12/31/2012						0
3263327908	03/31/2013						0
4020867458	06/30/2012	14	14	0	0		0







# ASK!

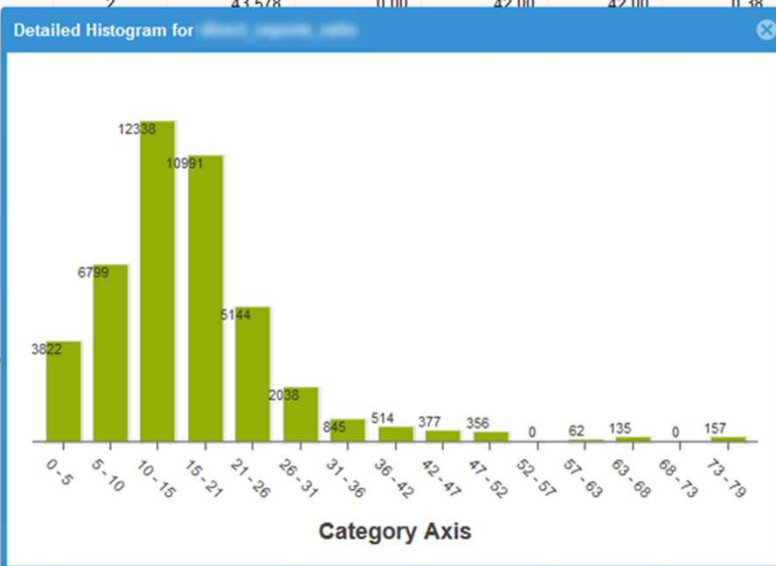
## ANALYTICS FOR SANDIA KNOWLEDGE

user   role: *data analyst*

Selection Preview **Quality** Data Analysis

Summary of quality for selected data for observation dates from 06/30/2012 to 03/31/2013

Field	Histogram	Concerns ↓	Observations	Minimum	Maximum	Range	Mean	Std. Deviation	Unique	Outliers	Extremes	% Complete
Range_in_inches		2	43,578	0.00	42.00	42.00	0.37	2.03	36	516	329	87.91
Range_in_inches		2	43,578	0.00	42.00	42.00	0.38	2.04	35	475	303	76.83
avg_incidents_per_year								0.12	3	563	563	100.00
time_in_year_job								21.58	221	1216	518	100.00
time_in_year_job								9.89	56	902	226	100.00
avg_time_in_year								6.44	2373	359	0	100.00
all_data_count								18.65	197	352	158	100.00
complete_data_count								11.20	145	430	213	100.00
missing_data_count								0.87	20	384	52	100.00



Data analysis for observation dates from 06/30/2012 to 03/31/2013

Data Analysis: Scatter Plot (2 numerical variables)

▲ X-axis

☐ show only predictor fields

view:

Person and Org Information

☒ show all fields

field:

New\_JobDate

▲ Y-axis

☐ show only predictor fields

view:

Person and Org Information

☒ show all fields

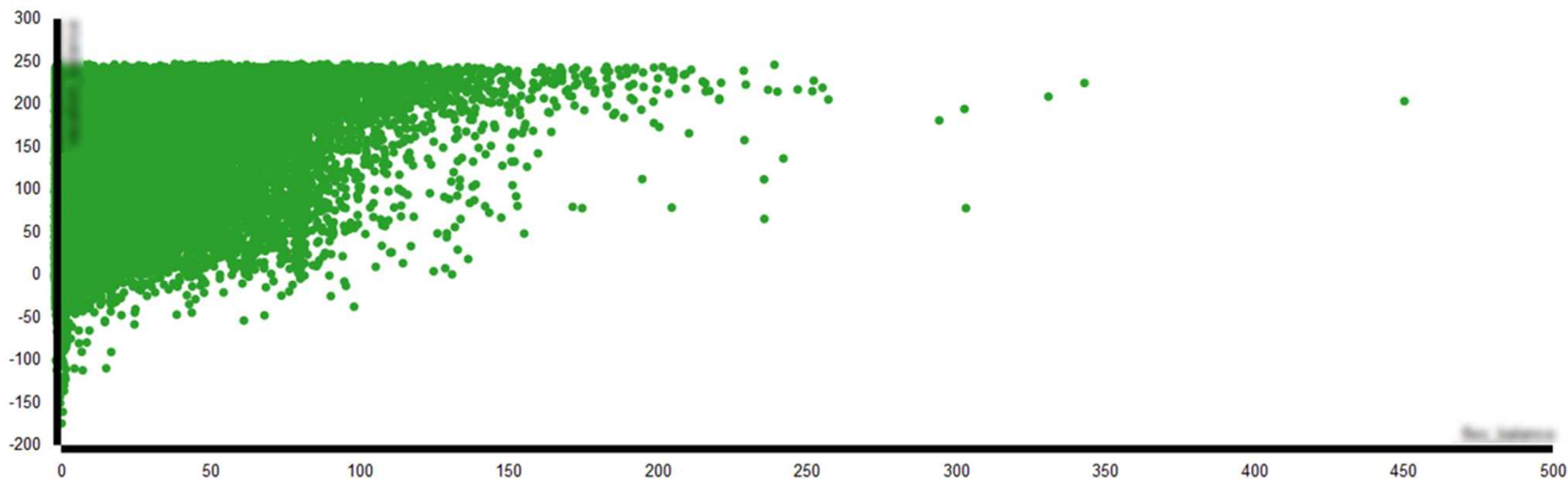
field:

vacation\_days

Perform Analysis

Scatter Plot with X-axis: New\_JobDate Y-axis: vacation\_days

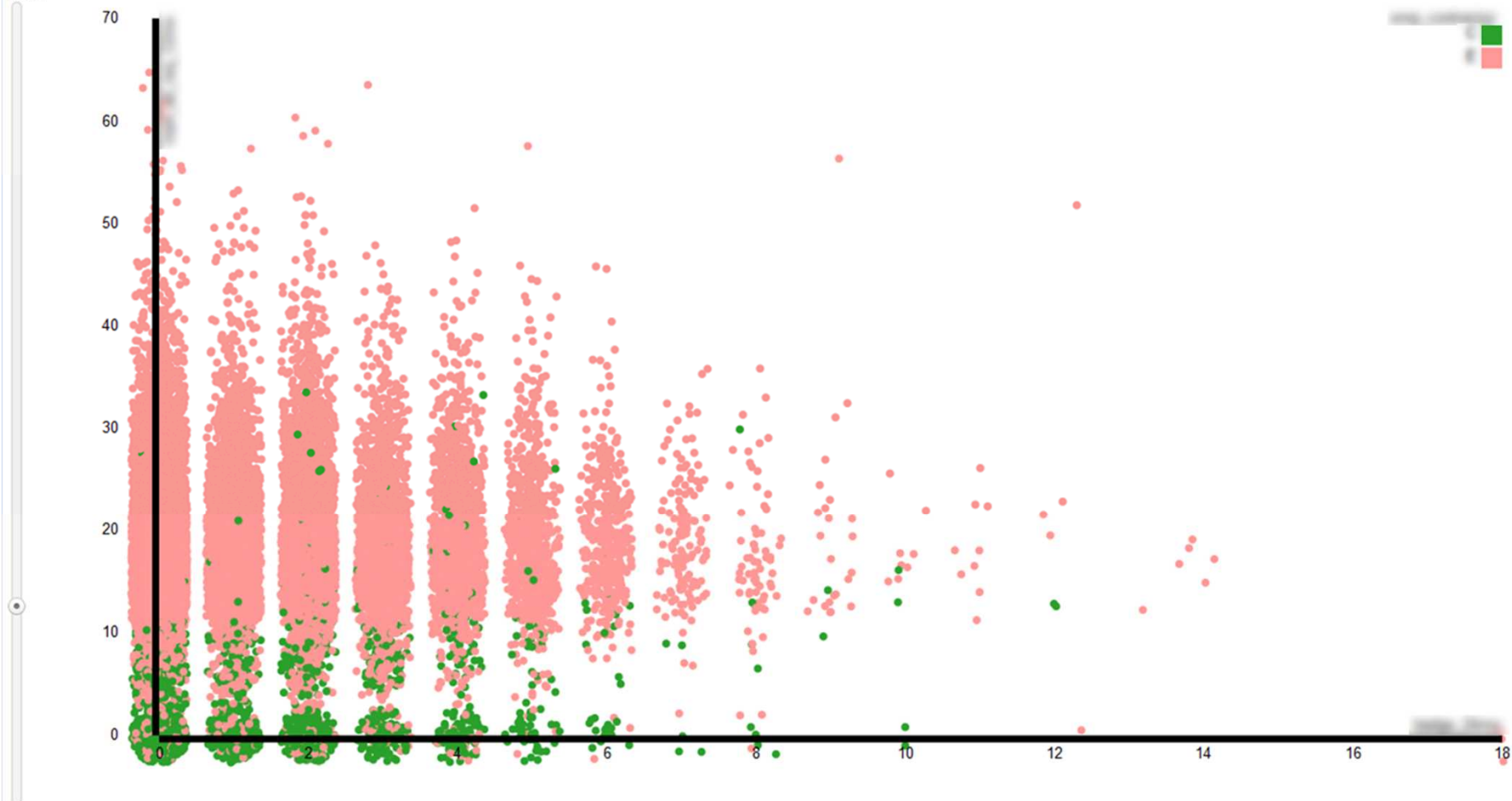
Jitter:



Data analysis for observation dates from 06/30/2012 to 03/31/2013

Scatter Plot with X-axis:  Y-axis:  Color:

Jitter:



 Sandia National Laboratories

C I O SERVICES EXPO

 Sandia National Laboratories



Data analysis for observation dates from 06/30/2012 to 03/31/2013

Bivariate Grouping Percentage Analysis with Predictor:  Target:

Show Details (for now, assumption is that NO is a desirable outcome)

Bin Contents	Desirable Outcome Ratio
0.005479 to 0.986301	-1.44
0.994521 to 2.060274	-1.18
2.071233 to 3.268493	-1.06
3.282192 to 5.052055	-1.10
5.068493 to 7.024658	-1.01
7.043836 to 8.484932	1.09
8.49589 to 10.230137	1.08
10.235616 to 12.410959	1.29
12.416438 to 18.430137	1.16
18.435616 to 31.019178	1.41



Selection

Preview

Quality

Data Analysis

Data analysis for observation dates from 06/30/2012 to 03/31/2013

Bivariate Grouping Percentage Analysis with Predictor: **avg\_age\_2000** Target: **avg\_education\_2000**

Show Details

(for now, assumption is that NO is a desirable outcome)

Bin Contents

Desirable Outcome Ratio

0.0 to 1.0

1.19

2.0 to 2.0

1.08

3.0 to 3.0

-1.26

4.0 to 5.0

-1.24

6.0 to 14.0

-1.37

C I O SERVICES EXPO



Sandia National Laboratories



# Coming Soon: Answers to Questions

ASK | Analytics for Sandia Knowledge

HOME

BUSINESS INTELLIGENCE CENTER

DATA EXPLORATION

DATA ANALYTICS APPLICATION

How many slips, trips and falls in my organization?

ASK

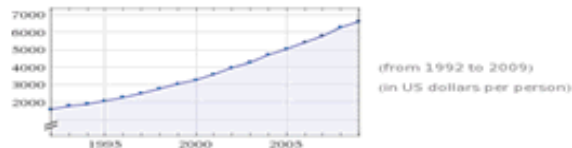
## Your Input:

[search term] within [all catalogs] ([Change?](#))

## Summary:

median	\$6161
highest	\$15010 (Pennsylvania)
lowest	\$0 (Hawaii and District of Columbia)
distribution	 (2009)

## Distribution Plots:



## Related Searches:

[query](#), [eh&s](#), [accidents](#), [vacation](#), [org 0000](#), [incidents](#)

CIO SERVICES EXPO

# ASK – Ambitious, Transformational

## Learn More About ASK at the Analytics Booth



VP Champion:

Mike Vahle, CIO and VP 9000

Milestone Owner:

Joselyne Gallegos, Director 9500

C I O SERVICES EXPO



Sandia National Laboratories

