

Welcome!



KEVIN TEETS

TBM Advisor

Apptio



CHUCK HARNISH

Director, Data Management

Global Foundation Services

Microsoft



CHRIS NORBY

Operations Program Manager

Data Management

Global Foundation Services

Microsoft

Hallmarks of **Wildly Successful Customers**



TBM OFFICE

PROGRAM – GOVERNANCE – TBMA



ROADMAP

ALIGNED WITH INITIATIVES



EVOLUTION

DATA – ALLOCATIONS – CAPABILITY



SPANS IT & FINANCE

KEY STAKEHOLDERS



TBM PROCESSES

2+ OPERATIONALIZED



INSIGHTS

DECISION MAKING ENGINE

Introduction

1. TBM ADVISORS – MISSION STATEMENT



2. WHY AM I TALKING ABOUT DATA QUALITY?



3. WHAT YOU WILL GET OUT OF TODAY'S PRESENTATION?



Challenges

Don't want to start TBM
because of poor data quality

Do not have confidence
in the level of data quality
to defend the allocations
within the Apptio model

Data Stewards do not have visibility into how
big the problem is or how to quantify the data
quality problem, making it difficult to convince
management to fund clean up initiatives

START NOW!

“A journey of a thousand miles begins
with a single step.”

– Lao-tsu, The Way of Lao-tsu – Chinese philosopher (604 BC – 531 BC)

EMBRACE FAILURE!

“Ever tried? Ever failed? No matter, try again,
fail again, fail better.”

– Samuel Beckett – Irish author, dramatist, novelist in France (1906 – 1989)

COMMUNICATION!

“If you cannot convince them, confuse them.”

– Harry S. Truman (1884 – 1972)

What We Do



Multi-faceted approach: Data Quality Dimensions, Adaptive Data Management, and Data Quality Allocations



Raise visibility to data quality issues across 5 core dimensions



Use Adaptive Data Management methods to validate, cleanse, enrich, and relate data within Apptio so you don't need to fix all of your DQ issues in the source systems



Identify breakage or “unallocated” costs in the model by providing a robust set of data quality allocation reports

Apptio Data Quality Dimensions – Defined

APPTIO ON DEMAND



MAINTAINABILITY

A measure of the degree to which data can be accessed and used, and the degree to which data can be updated, maintained, and managed.

UNIQUENESS

A measure of the number of distinct or unique values within a field.

COMPLETENESS

A measure of the count and percentage of the fields that contain a value. Example: A server Operating System has a value in 90 out of 100 records, Data Quality Completeness = 90%.

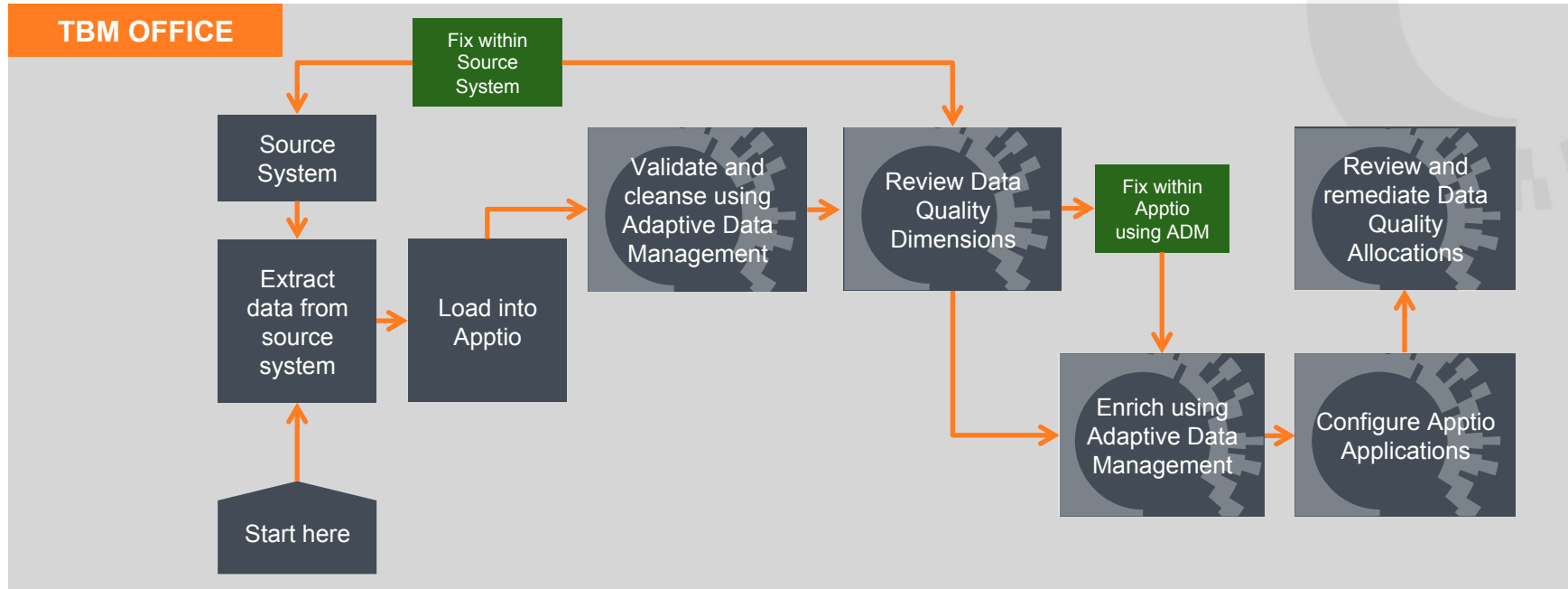
VALIDITY

A measure of the values within the data set compared to a set of allowed valid values. Example: Test all values in OS to a master set of valid values to ensure there are no outliers.

FREQUENCY DISTRIBUTION

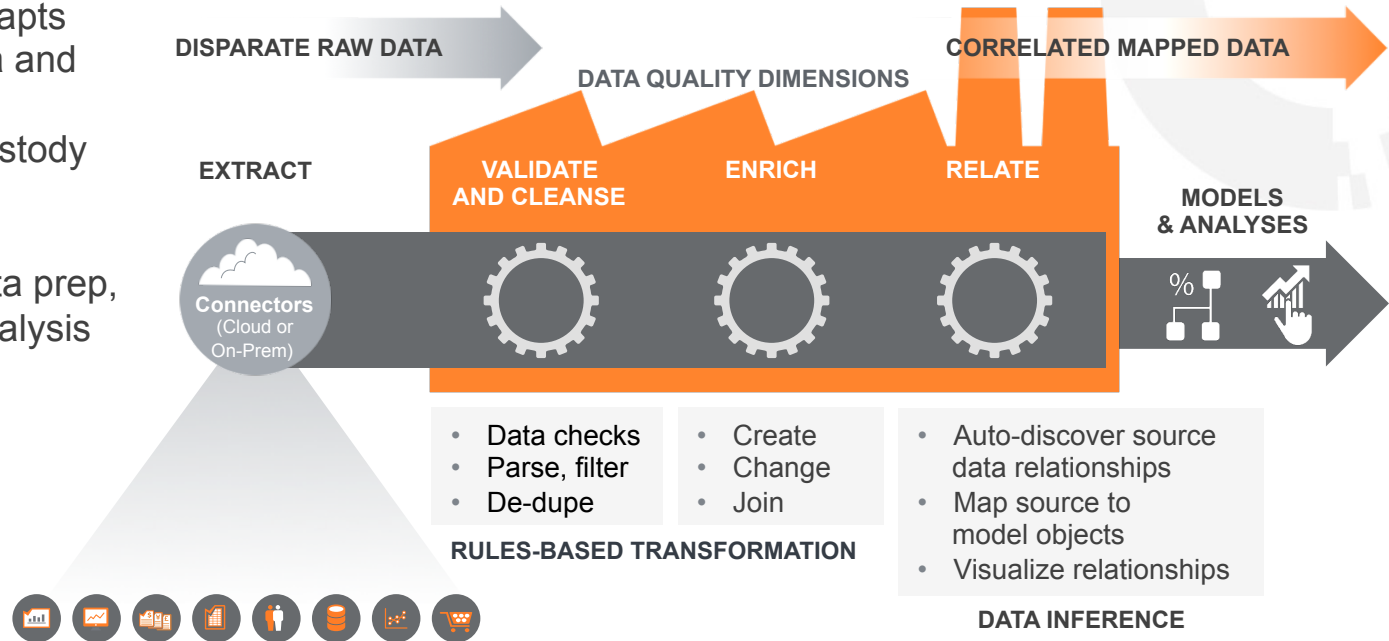
The distribution of unique values within a field by count and percentage. For those values with a low frequency, consider dropping them and changing to another comparable value.

Data Quality Process



Apptio Adaptive Data Management

- Continuously adapts to changing data and analytics needs
- Data chain of custody preserved and documented
- Less time on data prep, more time on analysis and decisions



Data Quality Dimensions – Dashboard




Data Quality Dimensions Dashboard

	Maintainability	Uniqueness	Completeness	Validity	Frequency Distribution
Applications	App Maintainability % 98.00% App Total Rows: 1,243	App Uniqueness % 99.68% App Duplicate Rows #: 4	App Completeness % 79.08% App Blanks #: 1,820	App Validity % 72.97% App Invalid #: 2,016	App Freq Dist % 95.98% App Threshold Rows #: 37
Storage	Storage Maintainability % 84.00% Storage Total Rows #: 7,776	Storage Uniqueness % 99.65% Storage Duplicate Rows #: 27	Storage Completeness % 88.57% Storage Blanks #: 7,112	Storage Validity % 77.41% Storage Invalid #: 7,027	Storage Freq Dist % 99.26% Storage Duplicate Rows #: 27
Servers	Server Maintainability % 98.00% Server Total Rows #: 6,874	Server Uniqueness % 99.88% Server Duplicate Rows #: 8	Server Completeness % 82.38% Server Blanks #: 18,166	Server Validity % 78.22% Server Invalid #: 16,466	Server Freq Dist % 91.25% Server Threshold Rows #: 137
GL	GL Maintainability % 76.00% GL Total Rows: 30,881	GL Uniqueness % 14.78% GL Duplicate Rows #: 26,317	GL Completeness % 89.57% GL Blanks #: 32,211	GL Validity % 48.24% GL Invalid #: 159,834	GL Freq Dist % 80.09% GL Threshold Rows #: 618

Drill into
any KPI for
additional
detail

Data Quality Dimensions – Maintainability



DQ Dimensions - Servers

Home | Maintainability | Uniqueness | Completeness | Validity | Frequency

Definition - Maintainability: A measure of the degree to which data can be accessed and used, and the degree to which data can be updated, maintained, and managed.

Is the Source System Database, Excel, or Other?:

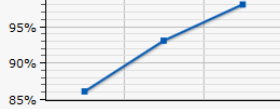
☒ Database ☐ Excel ☐ Other

Rate from 1 being the lowest to 5 being the highest the amount of time it takes to extract data from the source system:

☐ 1 ☒ 2 ☐ 3 ☐ 4 ☐ 5

Rate from 1 being the lowest to 5 being the highest the amount of data manipulation that occurs between the source and Apptio:

☒ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5



Server Maintainability %
98.00%
Server Total Rows #: 6,874

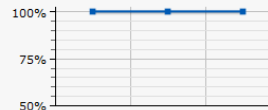
Answer 3
questions
to evaluate
score

Data Quality Dimensions – Uniqueness



DQ Dimensions - Servers

[Home](#) | [Maintainability](#) | [Uniqueness](#) | [Completeness](#) | [Validity](#) | [Frequency](#)



Server Uniqueness %

99.88%

Server Duplicate Rows #: 8

Definition - Uniqueness: A measure of the number of distinct or unique values within a field. It is the inverse of duplication.

Overview of Duplication in Apptio:

If there are duplicate rows in the dataset, they will appear in the table below. To understand how to manage duplication, there are different ways to review and fix this issue.

1. Cardinality setting within the original dataset. You can always review the Cardinality setting with any column within Apptio. Navigate to the Databab --> your original dataset --> Edit Columns Tab --> Choose the column you would like to review cardinality --> Cardinality will either say "Duplicates Found" or "Unique" for each column.

2. Manage Unique Row Filter: If you cannot modify the source data to remove duplicates, Apptio allows you to filter out duplicates using the "Manage Unique Row Filter" for any dataset. See the Help Guide for more information.

Server ID	Purpose	Class	Platform	Location	OS	Status	Vendor	CPU Cores	Memory Capacity	Virtualization Profile	BU Consumer	Hosted On	Application	Service Consumer	DuplicateCount
MKLXSNA01SESX33	Production	Physical	VMware Host	SATC	VMWare	Installed	Single-tenant Host	24	64	Hypervisor	MarketLinx	MKLXSNA01CAP003	Other	Other	2
MKLXSNA01SESX34	Production	Physical	VMware Host	SATC	VMWare	Installed	Single-tenant Host	24	64	Hypervisor	MarketLinx	MKLXSNA01CAP003	Other	Other	2
MKLXSNA01SESX35	Production	Physical	VMware Host	SATC	VMWare	Installed	Single-tenant Host	24	64	Hypervisor	MarketLinx	MKLXSNA01CAP003	Other	Other	2
MKLXSNA01SESX36	Production	Physical	VMware Host	SATC	VMWare	Installed	Single-tenant Host	24	64	Hypervisor	MarketLinx	MKLXSNA01CAP003	Other	Other	2

Total number of duplicate records

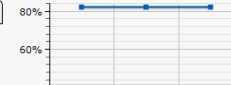
Data Quality Dimensions – Completeness

blanks
per column



DQ Dimensions - Servers

Home | Maintainability | Uniqueness | Completeness | Validity | Frequency



Server Completeness %

82.38%

Server Blanks #: 18,166

Definition - Completeness: A measure of the count and percentage of the fields that contain a value.

Example: A server Operating System has a value in 90 out of 100 records, Data Quality Completeness = 90%

▼ Number of Blanks in each applicable Column

blank_Purpose	blank_Class	blank_Platform	blank_Location	blank_OS	blank_Status	blank_Vendor	blank_CPU Cores	blank_Memory Capacity	blank_Date Purchased	blank_Virtualizat Profile	blank_BU Consumer	blank_Hosted On	blank_Application	blank_Service Consumer	Total
335	1	75	1	205	1	6469	1474	1970	1893	1	13	770	2479	2479	18166

▼ Purpose

DQ Error

▼ Class

DQ Error

▼ Platform

DQ Error

▼ Location

DQ Error

▼ OS

DQ Error

▼ Status

DQ Error

▼ Vendor

DQ Error

▼ CPUCore

DQ Error

▼ MemCap

DQ Error

▼ DatePur

DQ Error

▼ VirtPro

DQ Error

▼ BUCons

DQ Error

▼ Hosted

DQ Error

▼ App

DQ Error

▼ SerCons

DQ Error

Server ID ▲	Purpose	Class	Platform	Location	OS	Status	Vendor	CPU Cores	Memory Capacity	Date Purchased	Virtualization Profile	BU Consumer	Hosted On	Application	Service Consumer
10.03_U10_MDS 9216	Production	Physical	Physical	SATC	Unknown	Installed		0	0	5/2/2008 0:00	Physical	Enterprise Technology Services			
2074 CONTROLLER ALPHA	Production	Physical	Physical	SATC	Unknown	Installed		0	0	5/2/2008 0:00	Physical	Enterprise Technology Services			
2074 CONTROLLER BETA	Production	Physical	Physical	SATC	Unknown	Installed		0	0	5/2/2008 0:00	Physical	Enterprise Technology Services			
2107 CONSOLE	Production	Physical	Physical	SATC	Windows	Installed		0	0	5/2/2008 0:00	Physical	Enterprise Technology Services			
2UX8500157		Physical	Physical	SATC		Received		0	0	12/23/2010 0:00	Physical	Enterprise Technology Services		Other	Other
2UX91104KV		Physical	Physical	SATC		Received		0	0	11/23/2010 0:00	Physical	Enterprise Technology Services		Other	Other
2UX91104L9		Physical	Physical	SATC		Received		0	0	4/28/2011 0:00	Physical	Enterprise Technology Services		Other	Other
2UX91104LJ		Physical	Physical	SATC		Received		0	0	11/29/2010 0:00	Physical	Enterprise Technology Services		Other	Other
2UX91201RC		Physical	Physical	SATC		Received		0	0	12/23/2010 0:00	Physical	Enterprise Technology Services		Other	Other
2UX91201RL		Physical	Physical	SATC		Received		0	0	12/23/2010 0:00	Physical	Enterprise Technology Services		Other	Other

◀

Page

1

of 687

▶

⌂

Displaying 1 - 10 of 6870

Page 1 of 687

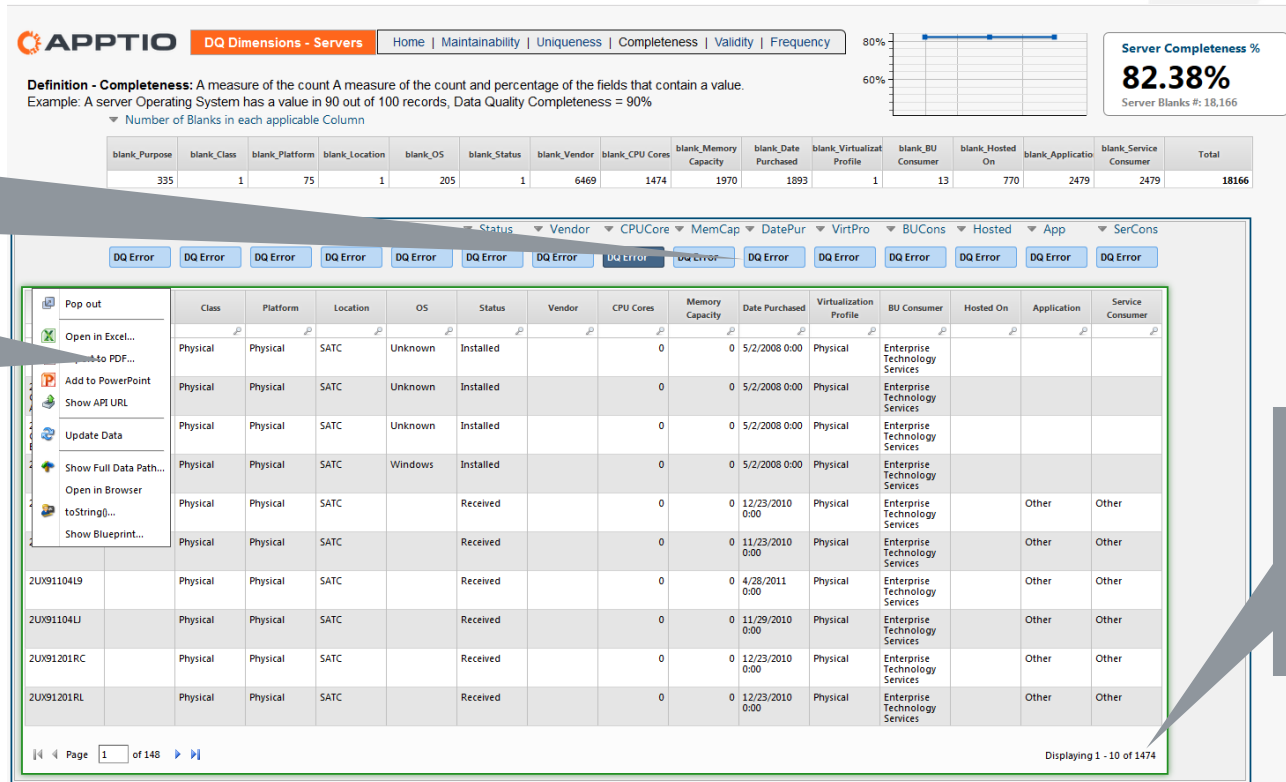
Displaying 1 - 10 of 6870

Data Quality Dimensions – Completeness

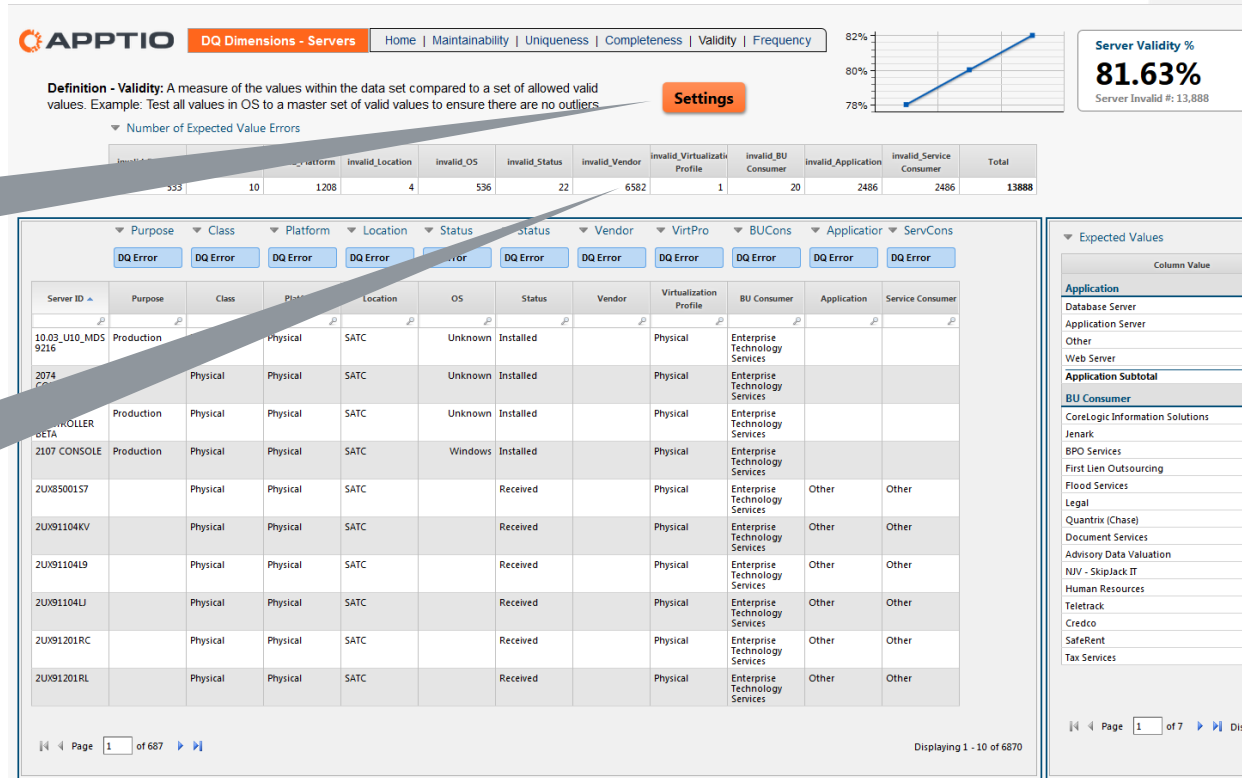
Slice into each column to discover the rows causing the errors

Export out for remediation

Row count is the same as # blanks above



Data Quality Dimensions – Validity



First –
configure
Settings

Invalid cells
are based
on values set
in Settings

Data Quality Dimensions – Validity Settings

Instructions

APPTIO DQ Dimensions - Servers Validity Settings | Validity

Overview Purpose Class Platform Location OS Status Vendor Virtualization Profile BU Consumer Application Service Consumer

Overview of Validity Settings:

Each tab in this section represents a specific column from the master data set to be tested for Validity. Values have been populated from the master dataset. The user must select whether or not the value is "Valid" or "Invalid", testing for whether or not the value should be a standard valid value. Functionality is as follows:

1. Unique values are derived from the master data set
2. New values that may show up in future months will appear as BLANK and show up in the exception list table in the Overview tab
3. All BLANK Column values default to "Invalid" and cannot be changed
4. All BLANK Validity values default to "Invalid" until changed
5. Users can change all values (Either "Valid" or "Invalid") so the user will be able to determine the new values that have come in on a monthly basis by reviewing the exception list.

Values Unassigned "Valid" or "Invalid"

Column Value	Validity
Class	
Multi-tenant - Virtualization	
Class Subtotal	

Unique Values in each Column

Column Value	Validity
Application	
Database Server	Valid
Application Server	Valid
Other	Valid
Web Server	Valid
Proxy Server	Invalid
ApPlication Server	Invalid
Application Subtotal	
BU Consumer	
CoreLogic Information Solutions	Valid
Jenark	Valid
BPO Services	Valid
First Lien Outsourcing	Valid
Flood Services	Valid
Legal	Valid
Quantrix (Chase)	Valid
Document Services	Valid
Advisory Data Valuation	Valid
NJV - SkipJack IT	Valid
Human Resources	Valid

Validity Slicer

Invalid
Valid

Column Slicer

Application
BU Consumer
Class
Location
OS
Platform
Purpose
Service Consumer
Status
Vendor
Virtualization Profile

Total

Page 1 of 9 Displaying 1 - 18 of 154

Values that are new and have not been assigned

Data Quality Dimensions – Validity Settings

APPTIO DQ Dimensions - Servers Validity Settings | Validity

Overview Purpose Class Platform Location OS Status Vendor Virtualization Profile BU Consumer Application Service Consumer

▼ DQ Servers Validity zconf Class (Raw Table)

Class	Count	Validity
Physical	2539	Valid
Shared	115	Valid
Cloud - Prod	1400	Valid
Multi-Tenant - Virtualization	1453	Valid
Single-Tenant - Virtualization	1264	Valid
Multi-tenant - Virtualization	9	
Cloud - Staging	93	Valid

Save Cancel

Assign “Valid” or “Invalid” as appropriate. These values serve as the “Master” lookup values

Change values and click Save before moving on

Data Quality Dimensions – Validity

Invalid values per column



DQ Dimensions - Servers

Home | Maintainability | Uniqueness | Completeness | Validity | Frequency

Definition - Validity: A measure of the values within the data set compared to a set of allowed valid values. Example: Test all values in OS to a master set of valid values to ensure there are no outliers.

Settings



Server Validity %

81.63%

Server Invalid #: 13,888

▼ Number of Expected Value Errors

Invalid_Purpose	Invalid_Class	Invalid_Platform	Invalid_Location	Invalid_OS	Invalid_Status	Invalid_Vendor	Invalid_Virtualization_Profile	Invalid_BU_Consumer	Invalid_Application	Invalid_Service_Consumer	Total
533	10	1208	4	536	22	6582	1	20	2486	2486	13888

▼ Purpose ▼ Class ▼ Platform ▼ Location ▼ Status ▼ Status ▼ Vendor ▼ VirtPro ▼ BUCons ▼ Application ▼ ServCons

DQ Error DQ Error DQ Error DQ Error DQ Error DQ Error DQ Error DQ Error DQ Error DQ Error

Server ID	Purpose	Class	Platform	Location	OS	Status	Vendor	Virtualization Profile	BU Consumer	Application	Service Consumer
2UX9S001S7		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I104KV		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I104L9		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I104LJ		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I201RC		Physical	Physical			Received		Physical	Enterprise Technology Services	Other	Other
2UX9I201RL				SATC		Received		Physical	Enterprise Technology Services	Other	Other
		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I304V8		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I304V9		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other
2UX9I304VB		Physical	Physical	SATC		Received		Physical	Enterprise Technology Services	Other	Other

Page 1 of 3

Displaying 1 - 10 of 22

▼ Expected Values

Column Value

Status

Installed

Installed-Offline

Status Subtotal

▼ Column Slicer

Application

BU Consumer

Class

Location

OS

Platform

Purpose

Service Consumer

Status

Vendor

Virtualization Profile

Compare against expected values

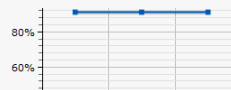
Data Quality Dimensions – Frequency Distribution



DQ Dimensions - Servers

Home | Maintainability | Uniqueness | Completeness | Validity | Frequency

Definition - Frequency Distribution: The distribution of unique values within a field by count and percentage. For those values with a low frequency, consider dropping them and changing to another comparable value.



Server Freq Dist %

91.25%

Server Threshold Rows #: 137

Select the Threshold % of records that would qualify for inspection:

2%

Save

Frequency Distribution

Column Value	Count
Application	
Application Server	2479
Other	1
Proxy Server	1314
Database Server	6
Application Server	1084
Web Server	1509
Application Subtotal	481
BU Consumer	
BPO Services	6874
First Lien Outsourcing	13
Flood Services	63
Legal	220
Quantrix (Chase)	129
Advisory Data Valuation	15
Human Resources	5
Teletrack	36
Credco	7
Tax Services	186
Commercial Real Estate Services	38
Total	236

Page 1 of 50 Displaying 1 - 23 of 1132

Count Slicer

Greater Than 0

Column Slicer

- Application
- BU Consumer
- Class
- CPU Cores
- Date Purchased
- Location
- Memory Capacity
- OS
- Platform
- Purpose
- Service Consumer
- Status
- Vendor
- Virtualization Profile

Instructions: The KPI (Server Freq Dist %) is calculated based on the number of rows that fall below the Threshold Percent that is set in Step 1 below. By setting a threshold %, all Frequency Distribution columns values are evaluated to see if the number of occurrences of that value are less than the Threshold Rows KPI (This KPI is simply calculating the Number of rows that would apply based on the Threshold % set and the total number of records in the dataset).

Step 1: Review the Distributions of all the column values in the table. Notice that each SubTotal = the Number of records in the dataset.


Step 2: Select the desired Threshold % that will be evaluated for this KPI and hit Save. NOTE: The higher the tolerance, the more records will fall below for inspection, the lower the overall KPI will be.

Step 3: Change the Count Slicer from Greater Than 0 to Less Than (Value shown in the "Server Threshold Rows KPI")

Step 4: Review all records in the table that should be evaluated for cleanup based on the LOW number of distribution of that value within that particular column.

Instructions and overview of calculation

Data Quality Dimensions – Frequency Distribution

**APPTIO**

DQ Dimensions - Servers

Home | Maintainability | Uniqueness | Completeness | Validity | Frequency

Definition - Frequency Distribution: The distribution of unique values within a field by count and percentage. For those values with a low frequency, consider dropping them and changing to another comparable value.

Select the Threshold % of records that would qualify for inspection: 2% Save

Frequency Distribution

Column Value	Count
Application	
ApPlication Server	2479
Other	1
Proxy Server	1314
Database Server	6
Application Server	1084
Web Server	1509
Application Subtotal	481
BU Consumer	
BPO Services	13
First Lien Outsourcing	63
Flood Services	220
Legal	129
Quantrix (Chase)	15
Advisory Data Valuation	5
Human Resources	36
Teletrack	7
Credco	186
Tax Services	38
Commercial Real Estate Services	236
Total	1
96236	

Page 1 of 50 Displaying 1 - 23 of 1132

Count Slicer

Greater Than 0

Column Slicer

- Application
- BU Consumer
- Class
- CPU Cores
- Date Purchased
- Location
- Memory Capacity
- OS
- Platform
- Purpose
- Service Consumer
- Status
- Vendor
- Virtualization Profile

Server Freq Dist %

91.25%

Server Threshold Rows #: 137

Instructions: The KPI (Server Freq Dist %) is calculated based on the number of rows that fall below the Threshold Percent that is set in Step 1 below. By setting a threshold %, all Frequency Distribution columns values are evaluated to see if the number of occurrences of that value are less than the Threshold Rows KPI (This KPI is simply calculating the Number of rows that would apply based on the Threshold % set and the total number of records in the dataset).

Step 1: Review the Distributions of all the column values in the table. Notice that each SubTotal = the Number of records in the dataset.

Step 2: Select the desired Threshold % that will be evaluated for this KPI and hit Save. NOTE: The higher the tolerance, the more records will fall below for inspection, the lower the overall KPI will be.

Step 3: Change the Count Slicer from Greater Than 0 to Less Than (Value shown in the "Server Threshold Rows KPI")

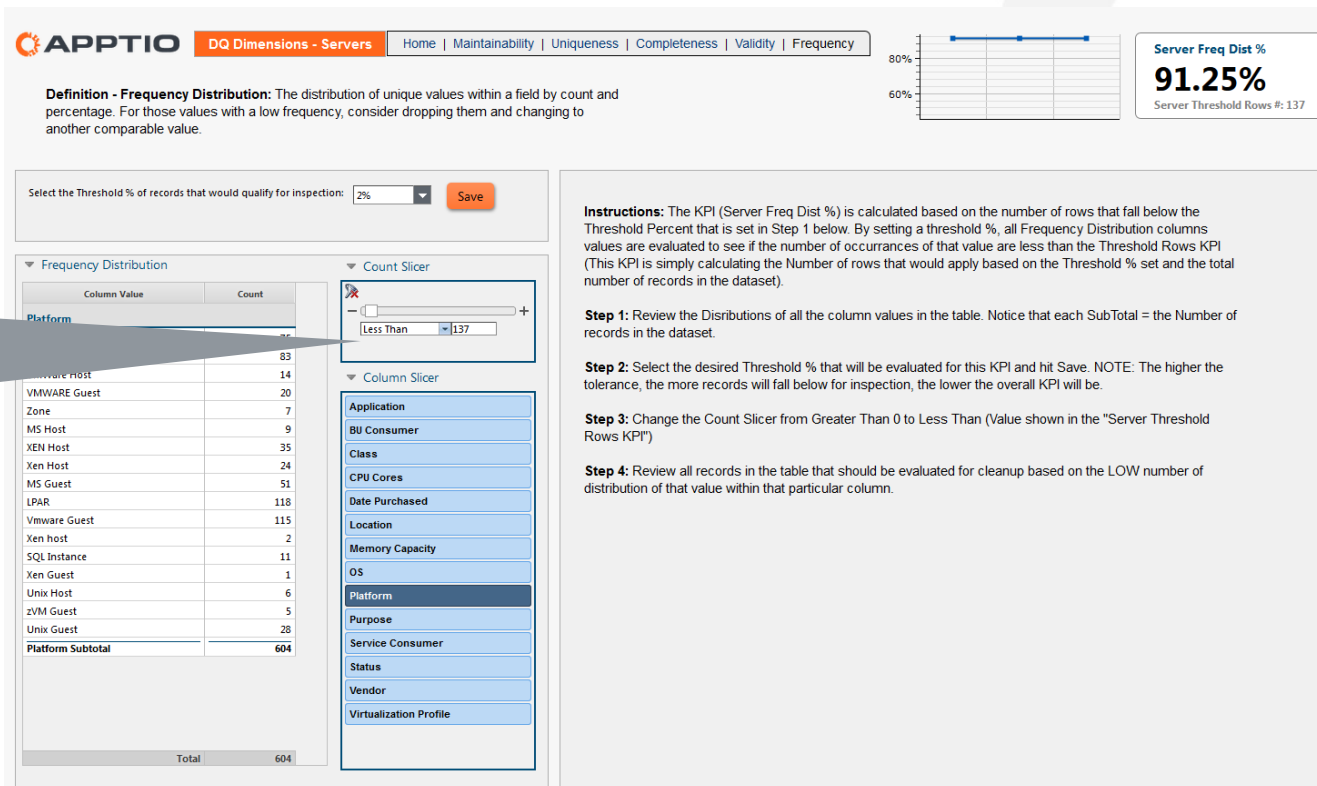
Step 4: Review all records in the table that should be evaluated for cleanup based on the LOW number of distribution of that value within that particular column.

Set
Threshold
Value

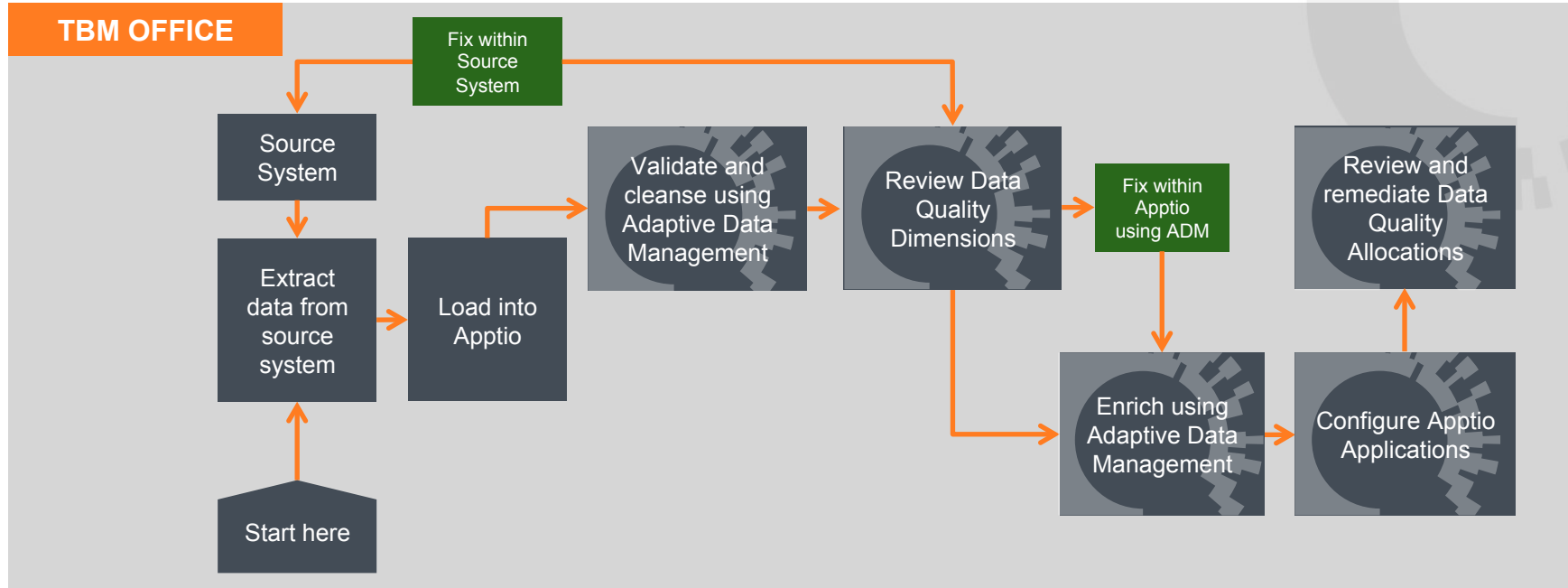
Review
Threshold
Row number

Data Quality Dimensions – Frequency Distribution

Adjust slicer settings and insert Threshold Rows KPI to review items for inspection

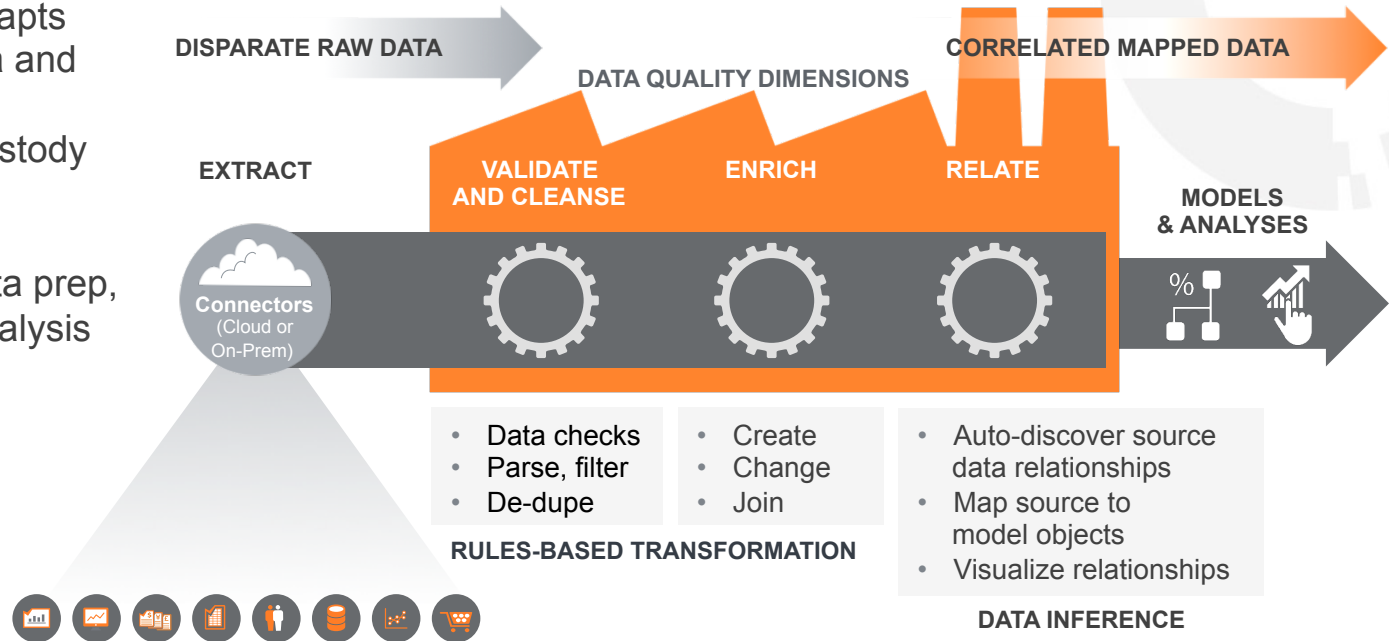


Data Quality Process



Apptio Adaptive Data Management

- Continuously adapts to changing data and analytics needs
- Data chain of custody preserved and documented
- Less time on data prep, more time on analysis and decisions



Data Quality Allocations Summary

The screenshot displays the Apptio dashboard interface. At the top, the Apptio logo and a 'Home' button are visible. The dashboard is organized into a grid of modules. A callout box on the left points to the 'Data Quality Owner' module, with the text: 'Navigate to the Data Quality Dashboard to review and remediate issues'.

Cost Transparency

- IT Finance**
 - Drive awareness of the cost of technology services
 - Provide defensible cost allocations
 - Manage your IT financial plan
- TBM Scorecard**
 - Understand the operating cost of IT Services and Resources
 - Identify the source of spend variance to plan
 - Automate analysis of IT staffing efficiency
- IT Operations**
 - Right size infrastructure capacity to usage
 - Rationalize your hardware platforms
 - Optimize your server and storage environments
- IT Management**
 - Understand the operating cost of IT Services and Resources
 - Identify the source of spend variance to plan
 - Automate analysis of IT staffing efficiency
- Data Quality Owner**
 - Understand how to make best use of the data you have
 - Identify places to improve source data
 - Understand the value better data can bring
- IT Service Owner**
 - Understand Application cost, quality, and value
 - Rationalize your Application Portfolio
 - Right size Application Infrastructure

Data Quality Allocations Summary



Data Quality

Summary

Financials

IT Resources

Services

Servers

Storage

Service Desk

Data Centers



Different tabs for different data owners

Total Unallocated Model Costs

\$7.13M

Total Spend Under Management: \$21.03M

Total User Uploaded Datasets Count

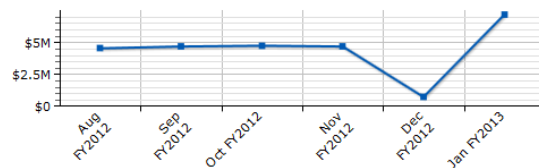
54 Datasets

[DataLink URL](#)

Total Model Object Count

28 Objects

Total Unallocated Model Cost



Finance Layer

1%

Unallocated

IT Resource Layer

8%

Unallocated

Service Layer

4%

Unallocated

Quickly understand largest gaps in unallocated \$ by model layer

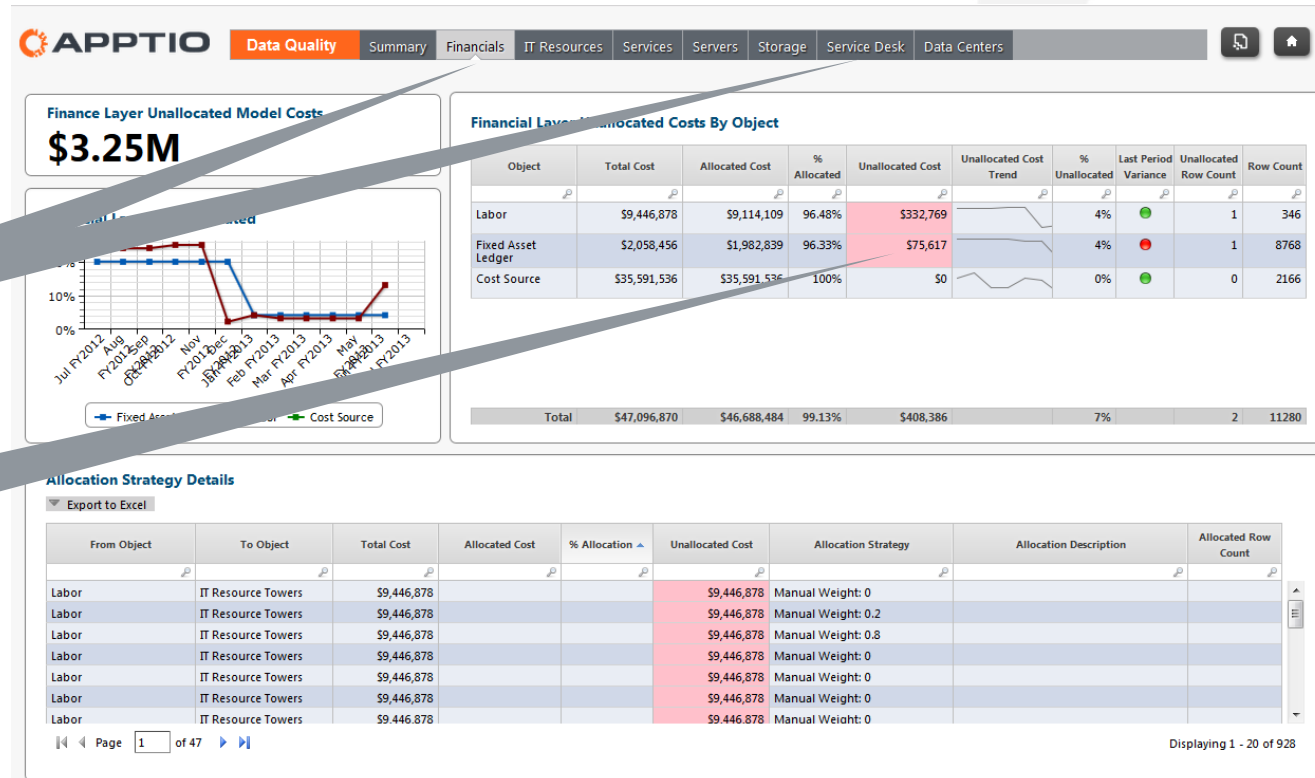
Unallocated Cost by Object Category

Object Category	Total Cost	Allocated Cost	% Allocated	Unallocated Cost	Unallocated Cost Trend	% Unallocated	Last Period Variance	Unallocated Row Count	Row Count
Services and Applications	\$26,474,630	\$25,419,422	96.01%	\$1,055,208		4%	●	20	279
Infrastructure	\$17,822,222	\$17,822,222	100%	\$0		0%	●	0	864
IT Resources	\$33,875,513	\$32,076,829	94.69%	\$2,682,689		8%	●	1871	12333
Financial	\$47,096,870	\$46,688,484	99.13%	\$408,386		1%	●	669	11280
Total	\$125,269,234	\$122,006,957	97.4%	\$4,146,283		13%		2560	24756

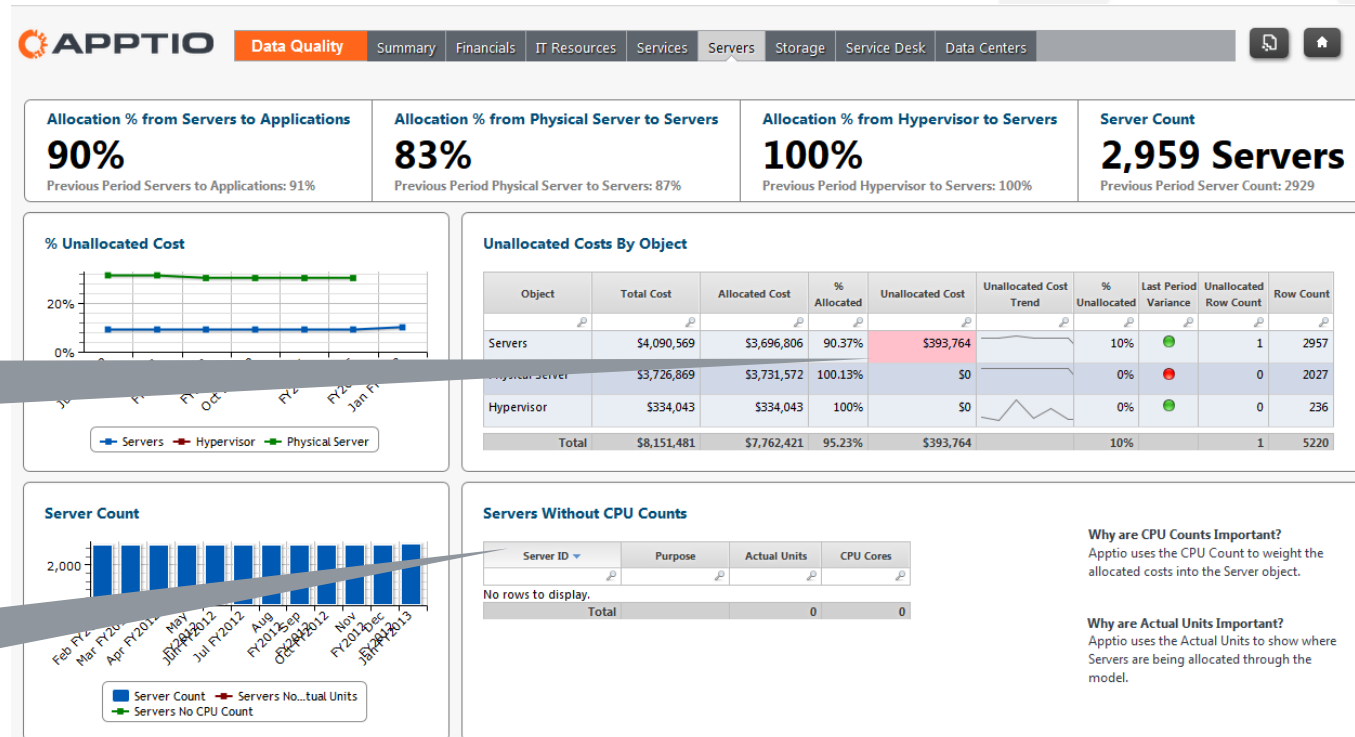
Data Quality Allocations Summary

Finance, IT Resources, and Services tabs represent model layer dashboards. The other tabs are for each resource domains data owner.

Review unallocated cost by object for each model layer



Data Quality Allocations Summary



Colored cells quickly allow you to pinpoint issues

Exception list for servers without CPU Counts

Get Started Today – TBM Starter Kit

An exclusive collection of assets to help you get started and accelerate your TBM road to success



- I. Improving Business Processes with TBM
- II. Building a TBM Office
- III. Data Quality
- IV. Looking Ahead

Kevin Teets
TBM Advisor
M: 949-375-7400
Kteets@apptio.com



Worldwide Headquarters

11100 NE 8th Street, Suite 600
Bellevue, Washington 98004 USA

Info@Apptio.com
Apptio.com

