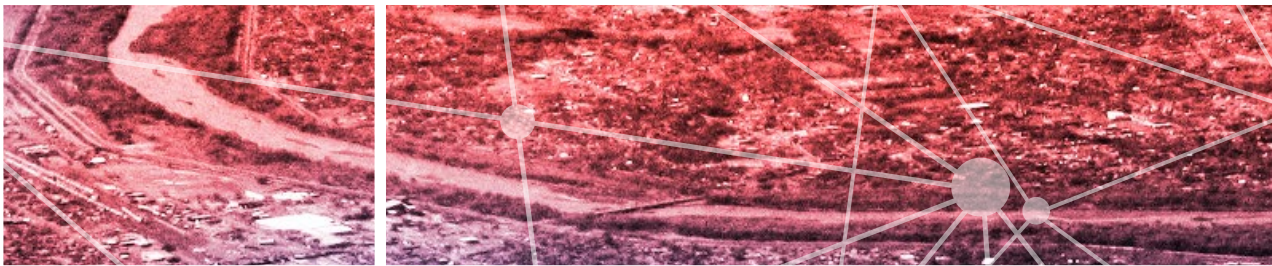


# Albuquerque Police Department Data Warehouse & Reports Tables Assessment Final Report

November 15, 2017

An Independent Observation of Paragraph 298 Data Reporting Mechanisms  
Conducted by RS21



# Table of Contents

## **Executive Summary 1**

Introduction 3

About This Report 4

## **OBSERVATIONS & RECOMMENDATIONS**

Data Management Infrastructure: The APD Data Warehouse 8

Reports Tables Data Quality Assessment Overview 8

Reports Tables Data Quality Assessment Observations 15

Reports Tables Data Fitness Assessment 36

SQL and SQL Codes 40

Data Quality Assurance Plans 41

Metadata 48

Modification Policy 49

Data Collection 50

Reporting and Communicating Data Insights 51

Conclusion 54

References 55

Appendices 57

About the Authors 58

## EXECUTIVE SUMMARY

As required by paragraph 298 of the Court Approved Settlement Agreement (CASA), the Albuquerque Police Department (APD/the Department) has been collecting, maintaining, and reporting performance data detailing operational patterns in areas such as use of force incidents, training, and recruitment. In recent years, the Department has made significant progress in its approach to data quality, but some challenges remain as the Department continues to evolve and mature its data management practices.

Previously, APD 298 performance reports have been highlighted for possible data quality deficiencies. As a result, the City of Albuquerque has contracted RS21 to provide an independent quality assessment of APD's data management infrastructure (the data warehouse) and the paragraph 298 data reports tables generated from the warehouse for data from January 1, 2016 to the most recent available at the time of observation.

This assessment does not include direct observations of Department practices of collecting or entering data into the system of record, nor does it include ground validation of data values at the point of data collection (e.g., field or administrative reports).

In recent years, APD has made notable progress in implementing improved data management infrastructure and procedures to advance data quality. Specifically, the Department has:

- 1.) Implemented a database and reporting architecture centered on a single data warehouse, which is expected to reduce the potential for clerical errors and conflicting data sets. Data warehouses are considered industry best practice in data management and help to ensure that everyone accessing the data is viewing and reporting on the same content.
- 2.) Developed data quality assurance plans specific to the majority of paragraph 298 reporting data sets. These plans are vital for ensuring that data sets are maintained and validated in a consistent and reliable manner over time.
- 3.) Developed metadata (data about data) describing the context and nuanced characteristics specific to each data set. The Department's metadata includes many important notations such as disclaimers and other notations about data collection and use—all of which are highly useful in maintaining diligence in data quality and communicating across a large number of data analysts and audiences both inside and outside the Department.
- 4.) Hired additional personnel (analysts) to aid in the internal management and validation of 298 reporting data. Data quality assurance takes time, expertise, and continued effort to maintain. The allocation of these additional resources is expected to improve data quality by allowing for more thorough data management activities such as internal validation conducted on a more frequent basis.

During the assessment, the authors have reviewed the APD data warehouse and reports tables against best practices in data quality and have highlighted a series of potential issues which may benefit from continued refinement. This report outlines a series of recommendations to assist the Department in improving data management and reporting practices into the future:

- 1.) While the vast majority of 298 reports contained data quality assurance plans, a small number did not yet have these plans assigned to them. To promote sustainable data quality over time, the Department should continue to develop, maintain, and closely follow a well-crafted quality assurance plan for every report.
- 2.) At the time of the assessment, multiple versions of some reports existed on the APD SharePoint site. To minimize potential for confusion and miscommunication, APD should strive to maintain only a single current version of any one report. Where practical, prior versions should be archived to a separate location outside the SharePoint site where they cannot be easily mistaken for current versions of the reports. APD has expressed to the authors that it is currently in the process of removing older versions of the reports.
- 3.) During the assessment, the APD team made modifications to the reports tables and metadata over a relatively short period of time. To minimize the potential for mistakes or miscommunications in this process, the Department should continue to review its current policy (or develop new guidelines if policy does not exist) to ensure that all proposed modifications are validated and approved by the appropriate authoritative sources before being implemented.
- 4.) A number of reports observed contained columns titled “No Data” or “Unknown.” The percentage of “No Data” or “Unknown” entries out of the total counts ranged from less than 0.5% to more than 75%. Indications of “No Data” and “Unknown” do not necessarily suggest errors in the data storage or reporting system. However, high percentages of either can make it difficult to effectively assess operational change. APD should continue to review existing data collection practices and look to further bolster a culture and practice of thorough data capture at the initial level of field or administrative reporting.

## INTRODUCTION

It is estimated that at present, the world creates more data every two years than has been created in the entirety of human history. Through increasingly sophisticated technology, we can capture and study immense quantities of highly detailed records of innumerable events occurring around us. By analyzing this data, we can begin to see intricate patterns in contextual characteristics correlated with events – patterns that might otherwise go unseen.

Increasingly, organizations of all varieties are collecting data about the performance of their own operations and harnessing the power of analytics to improve everything from policy and decision-making to operational effectiveness. Organizations leveraging analytics regularly to continually assess operations and inform changes to policies and procedures are sometimes referred to as Learning Organizations.

At this point in time, APD appears to be making notable progress in its transition to a Learning Organization, and it is our understanding that the ultimate goal is to harness robust data collection and analytics as a consistent tool to enhance decision-making and drive operational change. The Department's goal is to become a national leader in data collection, management, and analytics.

In law enforcement, operational analytics can empower law enforcement organizations to increase their efficacy in maintaining public safety while simultaneously strengthening relationships with members of the communities they serve. Agencies that become data-driven organizations can more easily assess and discuss the positive and negative outcomes of past performance, develop plans to maintain areas of strength, and identify and target operational opportunities for refining performance in areas needing improvement. This practice of perpetual self-assessment and organizational learning has the potential to lead to reduced crime and safer communities.

To be most effective, data collection and analytics procedures require thoughtful planning, well-crafted procedural practices across the organization, and diligent maintenance over time. Hardware, software, and personnel must all work together in a closely coordinated operation.

These dynamic aspects of a data management and reporting system must be carefully managed and overcome to ensure reliability of data and the resulting insights generated through analytics. And while the initial transition to a Learning Organization can require a good deal of time and organizational change to implement, the outcomes and improvements in decision-making intelligence can be greatly beneficial as analytics methods mature over time.

## ABOUT THIS REPORT

Since late August 2017, RS21—a data science, analytics, and visualization service provider—has been contracted by the City of Albuquerque to conduct an independent data quality assessment of APD’s data warehouse and paragraph 298 data reports tables.

In an enterprise level data management system such as APD’s, there are generally a large number of moving parts and elements to be maintained. Because of their potential impacts on the quality of data being reported in response to paragraph 298, the authors have developed a set of best practices, observations, and recommendations for future action with regard to the following components of the APD warehousing and reports system:

### **Data Management Infrastructure: The Data Warehouse**

- Data quality can be greatly affected by the infrastructure used to organize and manage the data
- The authors have reviewed the APD data warehouse for its effectiveness as a data management infrastructure

### **Reports Tables Data Quality Assessment**

- The reports tables capture curated data from the data warehouse to populate the APD 298 reports. It is crucial that the reports tables are constructed with quality in mind.
- The authors have evaluated the APD 298 reports tables for solvency across 6 dimensions of data quality

The methodology for assessing the data reports tables is based on best practices as published by independent sources for measuring data quality across six different dimensions. The specific terms of the methodology are described in further detail in Addendum A of this report.

Due to changes in APD data collection and management procedures in recent years, only data in the system since January 1, 2016 was included in this assessment. The authors used the date range of 01/01/2016 to 12/31/2017 to access the reports tables from the SharePoint site.

### **SQL (Database) Code Verification**

- The SQL code documented in the metadata of each report are the computer instructions for the reports tables to pull the data from the data warehouse into the reports tables. It is important that this code be kept up-to-date and reliable to drive accurate data transfer into the 298 reports.
- The authors have tested SQL codes to verify that they pull the same matching data from the database.

### **Data Quality Assurance Plans (DQA)**

- DQAs are a critical component of data quality management over time as they serve to standardize procedures to be followed by personnel in maintaining and validating data quality over time.
- The authors have evaluated existing APD DQAs against best practices from academia and industry.



### **Metadata**

- Metadata (data about data) is vital for documenting and communicating how, when, and by whom a data set has been collected, modified, and maintained over time.
- The authors have evaluated the metadata associated with APD 298 reports.

### **Modification Policy**

- To preserve data quality measures already implemented on a given data set, report, or metadata, necessary modifications to existing content must be approached and executed with caution.
- The authors have highlighted a series of recommendations for APD to consider as it moves forward.

### **Data Collection**

- The scope of this assessment does not include verification related to data collection procedures, but it is worth noting a few observations regarding data collection which will be valuable as APD continues its transition to a data-driven organization.

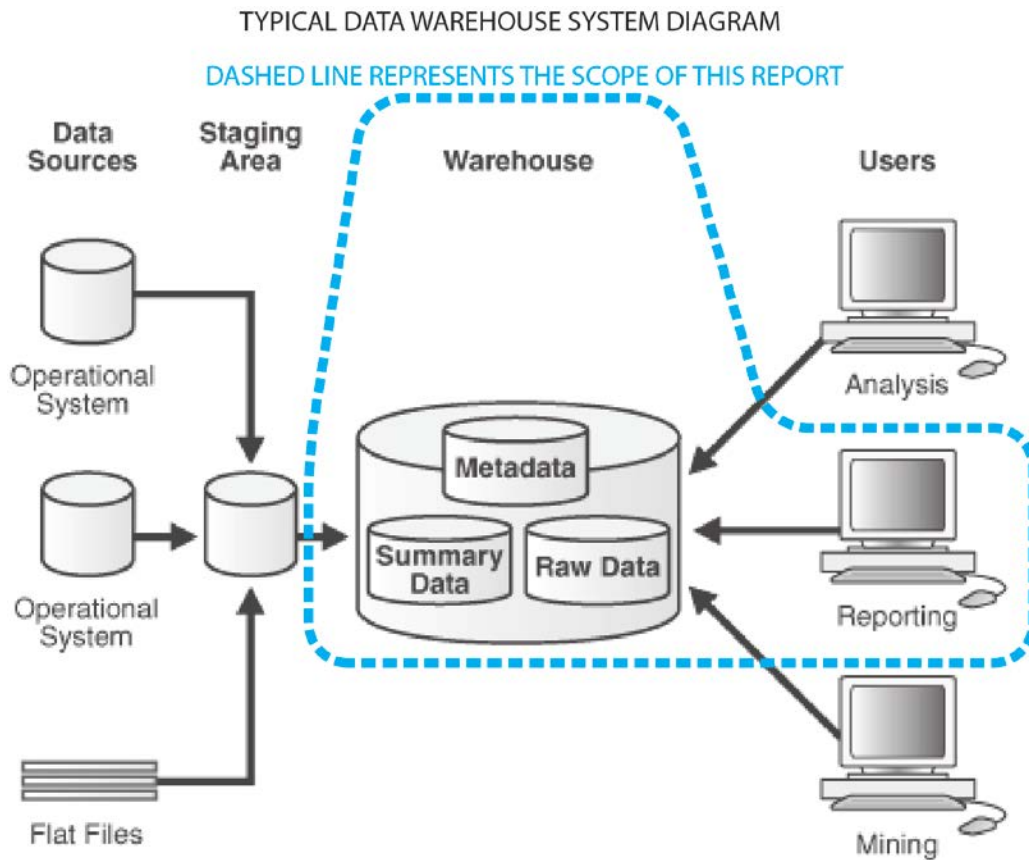
### **Fitness**

- Fitness refers to the ability of a data set to effectively answer a given analytical question—in the context of this report, the ability of APD reporting data sets to satisfy the specifications of paragraph 298 of the CASA.
- The authors have documented observations of data fitness for each of APD's 298 reporting tables.

### **Reporting Procedures and Communicating Data Insights**

- Communicating the valuable insights from a data set is often equally important as technically managing data quality.
- The authors have documented a series of observations and tradeoff decisions regarding the manner in which the reports tables are structured for communication to outside parties.

Given the timeframe of the assessment, it was not feasible to validate the ground truth of data at the initial source of data creation (e.g., field reports, administrative reports, and data entry of this information into the system of record). Moreover, the assessment does not include review of 298 reports as submitted to the Monitor or DoJ. As outlined in the diagram on the following page, the assessment has been focused on the portion of the APD data environment from the data warehouse to the reports tables used to generate 298 data reports. To facilitate the assessment, APD provided the authors with direct access to the data warehouse and reports tables.



**Typical data warehouse system diagram with staging area (Oracle Database Data Warehousing Guide 12c Release 1 (12.1.0.2))**

The data reports tables are “live” curated views of data referenced from the data warehouse and are used to construct the 298 reports. The graphic above is an illustration provided by APD of the data collection, management, and reporting system. The portion of the system observed by the authors is highlighted.

For the purposes of this assessment, the authors accessed and observed the reports tables via URL hyperlinks housed on the APD SharePoint site. The SharePoint site is a digital access point to the reports tables; it also serves as a repository of supplemental information (metadata), including a list of the data sets involved in each report, names of individuals associated with the data sets, database SQL codes, and a series of narrative disclaimers which provide additional context for the data observed in the reports tables.



The following table describes the APD reports tables which were reviewed as part of this assessment:

<b>PARAGRAPH 298 SECTION</b>	<b>REPORTS REVIEWED</b>	<b>NOT REVIEWED</b>
<i>298-a-i</i>	A1	A2 – Data managed by City of Albuquerque Citizen Police Oversight Agency (CPOA)  A7 - Reports tables not accessible at the time of the assessment
<i>298-a-iii</i>	A3	
<i>298-a-iv</i>	A4	
<i>298-a-v</i>	A5	
<i>298-a-vi</i>	A6 - Citizen	
<i>298-a-vi</i>	A6 - Officer	
<i>298-b-i</i>	B1	
<i>298-b-ii</i>	B2	
<i>298-c</i>	C - (Paragraph 129 & 137)	
<i>298-d-i</i>	D1	
<i>298-d-ii : 298-d-vi</i>	D2-6	
<i>298-e</i>	E	
<i>298-f-i</i>		F1 – Reports tables not accessible to authors at the time of assessment
<i>298-f-ii</i>	F2	
<i>298-g-i</i>	G1.a	
<i>298-g-i</i>	G1.b	
<i>298-h</i>	H – BlueTeam Communications	
<i>298-h</i>	H – Line Inspection	
<i>298-h</i>	H – Video Inspection Summary	
<i>298-i</i>	I CPC	
<i>298-i</i>	I IA	

## SUMMARY OF FINDINGS

The following sections describe the findings of the assessment in greater detail.

### **DATA MANAGEMENT INFRASTRUCTURE – THE APD DATA WAREHOUSE**

Since the 1990's, data warehouses have become commonplace in the industry. While a system of record maintains “golden” copies of data records, data warehouses are used to “pull” data from the system of record to populate reports, dashboards, and other types of analytics. This data warehouse concept has long been considered valuable for minimizing errors due to unintentional deletion, modification, or other by-products of users accessing golden records directly to generate reports (Kimball, Reeves, Ross, & Thornthwaite, 1998). Data warehouses are often conceptualized as a “single point of truth” for different users across an organization looking to access and report the same data without risk of conflicting data values from one user to the next (Eckerson, 2017).

The Department has developed and implemented a data warehouse that serves as the single point of truth for operations data storage and reporting access. Without a data warehouse, it would be far more likely that records could be negatively altered, degrading quality and reliability of the data. While there are many possible configurations for a data warehouse, it appears that the APD warehouse is organized as a top-down framework (Kimball, Reeves, Ross, & Thornthwaite, 1998). This assessment did not include an in-depth assessment of the warehouse configuration, but at a high level, the top-down configuration appears to meet the Department's needs.

Looking forward, it will be critical that the data warehouse management team continue to improve and evolve the warehouse to keep pace with changes in best practices and reporting or analytics needs of the Department over time.

### **REPORTS TABLES**

The reports tables are “live views” of the data housed in the data warehouse and are used to pull data from the warehouse (without modification) to populate APD's paragraph 298 reports. Because the reports tables pull directly from the data warehouse in an automated fashion, it is critical that the computer code driving this linkage is properly directing the database to pull the intended data from the warehouse into the reports tables. The image below illustrates a common set of reports tables:

Start Date  End Date

1 of 1 | 100% | Find | Next



## Paragraph 298B2 Specialized Units UOF on Activations Report v2 1/1/2016 - 12/31/2017

### UOF Incident By Month

Occurred Date	Count
01-2017	2
02-2017	1
04-2017	1
06-2017	1
07-2017	1
08-2017	1
09-2017	2
Total	9

### UOF by Force Type by Month

Occurred Date	Empty Hand Techniques	Impact - 40mm	K9 Apprehension - Bite	OC Fogger	Other Explain in Summary	Takedowns - Solo	Total
01-2017	0	1	0	1	1	0	3
02-2017	0	1	0	0	0	0	1
04-2017	0	0	1	0	0	0	1
06-2017	0	1	0	0	0	0	1
07-2017	0	0	0	0	1	0	1
08-2017	1	0	0	0	0	1	2
09-2017	0	2	1	0	0	0	3
Total	1	5	2	1	2	1	12

Typical Reports Tables layout and organization as observed by the authors.

During initial observations, the authors reviewed the reports tables for potential indications of issues such as the following:

- **Tables with empty cells where values would otherwise be expected**
- **Rows and/or columns without headers (titles) to describe the types of values displayed**
- **Total and sub-total tabulations which do not agree with manually-tabulated values**
  - o Some instances of disagreement between displayed tabulations and manually-calculated tabulations do exist, but in many cases, these instances appear reasonable as described in the disclaimers section of the metadata on the SharePoint site.
  - o As an example of a tabulation disagreement, in the table shown below from Report D2-6 (Recruiting Details), the manual tabulation of the bottom row is 31, but the column titled "Number of People Fluent in Other [Languages]" shows a value of 29. Due to the structure of the table, one would expect the column on the far right to reflect the total of all values to the left. In this case, it is possible that 2 of the 29 people speak 2 languages or 1 of the 29 people speaks 3 languages. Due to the structure of the table, a person speaking more than 1 language other than English would appear in more than 1 column. This could explain the discrepancy between the manually tabulated languages

(31) and the displayed total number of people (29). While this table may not contain an error, the structure and presentation of the table may confuse some readers.

**Academy Recruitment - Fluency in languages other than English**

Cohort Class Name	American Sign Language	French	Haitian	Japanese	Sarankule	Spanish	Turkish	Vietnamese	Wolof	Number of People Fluent in Other Lang.
117	1	2	1	1	1	22	1	1	1	29
<b>Total</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>22</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>29</b>

Typical Reports Table with reasonable difference between manually-tabulated total and total displayed in table.

**Reports Tables Data Quality Assessment Methodology**

While definitions of data quality vary slightly between publishers, the following methodology has been constructed on principles of data quality consistent across multiple sources of publication and directly applicable to the APD 298 reporting data within the context of this report.

In addition to other sources cited in the References section of this report, similar definitions of data quality are described in the following relevant publications:

- *U.S. Federal Bureau of Investigation (FBI) National Crime Information Center (NCIC) 2000 Operating Manual*
- *The Data Validation and Verification Assessment Matrix, published 2003 by the U.S. Department of the Interior (LeRon E. Bielak, Office of Planning and Performance Management)*
- *The Six Primary Dimensions for Data Quality Assessment (DAMA UK Working Group on "Data Quality Dimensions")*

RS21 has evaluated the APD 298 report tables for solvency across the following 6 dimensions of data quality according to the following definitions:

**1.) Structure and Consistency**

Structure refers to the formatting and organization of a given table. The tables visible on the Reports Server have been reviewed for consistency in their organizational construction. Generally, all tables reporting on similar types of data should be consistent in formatting. For example, in the case of a table reporting on officer race and a table reporting on citizen race, both tables should include the same number of columns for the various categories of race, and these columns should also occur from left to right in the same order in both tables. Discrepancies in this type of structural consistency should not be considered major errors, but are worth noting because inconsistent table structures could lead to limitations in future analytics exercises.

#### Observational Codes

Pass	The structure of the table appears consistent and appropriate for the table
*Pass	Flag for instances in which two or more tables with corresponding categories of data have been organized in a different order of columns or rows from one table to the next
	Other nuanced observations have been described directly in the observational tables

## 2.) Completeness

Within the context of this report, tables visible on the Reports Server have been examined for missing field values. In the scope of this assessment, the examination of completeness does not evaluate for whether critical rows or columns are missing; instead, it looks exclusively at the completeness of rows and columns present in the tables.

In some tables, there are columns or rows labeled “No Data” or “Unknown”. In either case, these tables have been flagged with an asterisk denoting that while the table itself can be deemed “complete” because all fields display a value, there could be a condition upstream of the table in which data is not being captured. In some cases, there may be a reasonable explanation for “No Data” or “Unknown” entries at the point of initial data collection, but within the context of this assessment, occurrences of “No Data” or “Unknown” entries have been noted for documentation purposes. “Unknown” may be a legitimate selection by an officer entering data. However, high prevalence of “Unknown” entries can minimize the value of a data set when analyzed to drive operational change.

#### Completeness Codes

Pass	A value is displayed in every cell of the table
*Pass	Table displays a value in every cell, but a column or row is labeled “No Data” or “Unknown”. The presence of these categories does not deem the table incomplete but does suggest that insights gleaned from the table could benefit from additional clarification of counts in these categories
'x'%	For any tables in which values appear to be missing, the percentage of complete cells has been documented
	Other nuanced observations have been described directly in the observational tables

### 3.) Uniqueness

Uniqueness evaluates whether there are any overlaps in data categorization which could cause over counting, under counting, or duplicated values within a given category (column or row header). Issues of uniqueness are often the result of syntactical inconsistency in categorical definitions. For example, if a data category is titled “Non-Hispanic,” we would flag an issue of uniqueness if a second overlapping category (especially within the same table) were present, such as “Non Hispanic,” or “Non\_Hispanic.” Conceptually, each title is intended to describe the same category of data, but due to syntactical discrepancies, what should be a single category of data with a single count value has been split into three separate categories, each with its own separate count value, thereby altering the reading of the table and potentially leading to complications of further analytics in this category of data.

In instances where a table includes rows or columns titled “No Data” or “Unknown,” it is not possible to effectively evaluate the table for uniqueness because there are no immediate means for concluding whether the values under those headings could be duplicative of other existing categories already present within the table.

#### Uniqueness Codes

Pass	Data categories do not suggest overlap, double counting, or under counting
*Pass	Other nuanced observations have been described directly in the observational tables

### 4.) Accuracy

For the purposes of this section of the report, accuracy has been evaluated only at the level of the reporting tables visible in the Reports Server. This assessment does not include evaluation with regards to the accuracy of values during initial data capture (e.g., an officer completing a field report), nor does it include verification of accuracy at any level of data processing upstream to the reporting tables on the Reports Server.

Accuracy of the tables on the Reports Server is calculated by manually tabulating the subtotals (or row and/or column counts) within those tables and comparing the manually tabulated totals to the total counts displayed in the Reports Server tables. While there are number of possible explanations for discrepancies between displayed totals and manually tabulated totals, disagreements between these two values are cause for uncertainty in reading the table. In some cases, explanations have been provided on the SharePoint site for a given report.

#### Accuracy Codes

Pass	Total values displayed in the report table are reproducible when counts or subtotals are tabulated manually
Tabulation Discrepancy	Total values displayed in the report table are different than the total when counts or subtotals are tabulated manually



	Other nuanced observations have been described directly in the observational tables
--	---

**5.) Validity**

Here, validity is defined as a given value’s structural appropriateness with respect to its relative position within a table. For example, if a column is labeled “Number of Arrests,” valid values would be whole, positive numbers. In this column, values like percentages, fractions, or negatives would be flagged as invalid in their structural appropriateness for describing “Number of Arrests”.

Within the scope of this assessment, data has not been evaluated for its validity beyond its structural appropriateness within the tables on the Report Server. In other words, no truth-finding or field validation has been conducted to assess truthfulness of data entry or accuracy of processing upstream of the APD reports tables on the Reports Server.

Validity codes

Pass	Total values displayed in the report table are reproducible when counts or subtotals are tabulated manually
	Other nuanced observations have been described directly in the observational tables

**6.) Fitness**

Fitness is a measure of a data category’s appropriateness in addressing the specific requirements of paragraph 298. For example, if paragraph 298 requires reporting on the gender of both officers and citizens involved in Use of Force incidents, we would expect to clearly see metrics adequately describing gender characteristics for both groups of people. Results of the fitness assessment are included after the observations of the reports tables.

**Additional Methodological Details**

RS21’s observations across the six dimensions of data quality have been organized into observational tables on the following pages.

It is important to note that the APD reports tables have been evaluated as they appeared to the authors on the APD Reports Server on the APD Real Time Crime Center (RTCC) SharePoint site as accessed through credentials and permissions provided by APD to the authors. APD provided guidance to the authors on how to access the appropriate channels for the assessment.

To the extent practical, all reasonable efforts have been made to access and evaluate the most current version of each report table at the time of assessment. Because the data warehouse was under active development during the assessment, multiple URL hyperlinks for various reports were often present, each directing to a different version of a given report. In these instances, the authors accessed the reports tables through the URL hyperlink listed under “Report Description” on the SharePoint site. Typically, the URL hyperlinks in this location correspond to “version 2” of a given report. In instances where no link was present in the “Report Description” location, the authors accessed the report table

through the primary report link on the SharePoint site. APD has expressed that it is currently in the process of removing duplicate reports links.

The authors were made aware that many changes and improvements had been made in past years to APD's policies for collecting, organizing, and reporting the 298 data. As such, RS21 was advised to limit the scope of the assessment to include only data beginning January 1, 2016.

Because changes were made to the SharePoint site during the assessment period, each of the following observational tables reflects the date on which the data set was observed and evaluated.

A1		UOF						
Reviewed: 10/25/2017 & 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
		*Columns or rows appear inconsistent between correlary tables where we would expect them to be similar from one table to the next	*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"				Reported Total	Manually Tabulated Total
		-	-	-	-	-	-	-
1	UOF Incident by Month	Pass	Pass	Pass	Tabulation Discrepancy	743	746	Pass
2	UOF by Force Type by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
3	UOF by Area Command by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
4	UOF by Arrest Type by Month	Pass	Pass	Pass	Tabulation Discrepancies (Multiple)	-	-	Pass
	UOF by Demographics by Month	-	-	-	-	-	-	-
5	Officer Gender by Month	*Pass	*Pass	Pass	Pass	-	-	Pass
6	Citizen Gender by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
7	Officer Age by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
8	Citizen Age by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
9	Officer Race by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
10	Citizen Race by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
11	Officer Ethnicity by Month	*Pass	*Pass	Pass	Tabulation Discrepancy	773	776	Pass
12	Citizen Ethnicity by Month	*Pass	*Pass	Pass	Not Tested	-	-	Pass
13	Citizen Gender Expression by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
14	Citizen Sexual Orientation by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
15	Citizen English Proficiency by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass

A3		UOF Report v2						
Reviewed: 10/25/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
		*Columns or rows appear inconsistent between correlary tables where we would expect them to be similar from one table to the next	*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	UOF Incident by Month	Pass	Pass	Pass	Pass	-	-	Pass
2	UOF by Force Type by Month	Pass	*Pass	Duplicate Column Headers: 'Hand/Feet Impact' + 'Hands-Feet-Impact'	Pass	-	-	Pass
3	UOF by Area Command by Month	Pass	Pass	Pass	Pass	-	-	Pass
4	UOF by Arrest Type by Month	Pass	Pass	Pass	Pass	-	-	Pass
	UOF by Demographics by Month	-	-	-	-	-	-	-
5	Officer Gender by Month	Pass	Pass	Pass	Pass	-	-	Pass
6	Citizen Gender by Month	Pass	Pass	Pass	Pass	-	-	Pass
7	Officer Age by Month	Pass	Pass	Pass	Pass	-	-	Pass
8	Citizen Age by Month	Pass	Pass	Pass	Pass	-	-	Pass
9	Officer Race by Month	*Pass	Pass	Pass	Pass	-	-	Pass
10	Citizen Race by Month	*Pass	Pass	Pass	Pass	-	-	Pass
11	Officer Ethnicity by Month	Pass	Pass	Pass	Pass	-	-	Pass
12	Citizen Ethnicity by Month	Pass	Pass	Pass	Pass	-	-	Pass
13	Citizen Gender Expression by Month	Pass	*Pass	Pass	Pass	-	-	Pass
14	Citizen Sexual Orientation by Month	Pass	*Pass	Pass	Pass	-	-	Pass
15	Citizen English Proficiency by Month	Pass	*Pass	Pass	Pass	-	-	Pass

A4	UOF Report v2							
Reviewed: 10/26/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Administrative Investigations Supported by Preponderance of the Evidence	Pass	Pass	Pass	Pass	-	-	Pass

<b>A5 UOF Report v2</b>								
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"		For Discrepancies, refer to APD "Disclaimer" for this report	Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Officers Identified in EIS for UOF (A5 pt 1)	Pass	Pass	Pass	Tabulation Discrepancies (Multiple)	-	-	Pass
2	Force Found to violate policy (A5 pt 2)	Pass	Pass	Pass	N/A	-	-	Pass



<b>A6 UOF Citizen Injury Report v2</b>								
Reviewed: 10/26/2017 & 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"					
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Citizen Injury Incident by Month	Pass	Pass	Pass	Pass	-	-	Pass
2	Citizen Injury by Injury Type by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
3	Citizen Injury by Area Command by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
4	Force Type Used by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
5	Citizen Injury by Arrest Type by Month	Pass	Pass	Pass	Tabulation Discrepancies	-	-	Pass
	Citizen Injury by Demographics by Month	-	-	-	-	-	-	-
6	Officer Gender by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
7	Citizen Gender by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
8	Officer Age by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
9	Citizen Age by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
10	Officer Race by Month	Pass	*Pass	Pass	Pass	-	-	Pass
11	Citizen Race by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
12	Officer Ethnicity by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
13	Citizen Ethnicity by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
14	Citizen Gender Expression by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
15	Citizen Sexual Orientation by Month	Pass	*Pass	Pass	Pass	-	-	Pass
16	Citizen English Proficiency by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass

<b>A6 UOF Officer Injury Report v2</b>								
Reviewed: 10/26/2017 & 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"					
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Officer Injury Incident by Month	Pass	Pass	Pass	Pass	-	-	Pass
2	Officer Injury by Injury Type by Month	Pass	Pass	Pass	Not Tested	-	-	Pass
3	Officer Injury by Area Command by Month	Pass	Pass	Pass	Not Tested	-	-	Pass
4	Force Type Used by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
5	Officer Injury by Arrest Type by Month	Pass	Pass	Pass	Pass	-	-	Pass
	Officer Injury by Demographics by Month	-	-	-	-	-	-	-
6	Officer Gender by Month	Pass	Pass	Pass	Not Tested	-	-	Pass
7	Citizen Gender by Month	Pass	*Pass	Pass	Pass	-	-	Pass
8	Officer Age by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
9	Citizen Age by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
10	Officer Race by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
11	Citizen Race by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
12	Officer Ethnicity by Month	Pass	*Pass	Pass	Pass	-	-	Pass
13	Citizen Ethnicity by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
14	Citizen Gender Expression by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
15	Citizen Sexual Orientation by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass
16	Citizen English Proficiency by Month	Pass	*Pass	Pass	Not Tested	-	-	Pass

<b>B1</b>		<b>Specialized Units Report</b>						
Reviewed: 10/25/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"					
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Monthly Specialized Unit Deployments and Activations	Pass	Pass	Pass	Pass	-	-	Pass

<b>B2 Specialized Units UOF on Activations Report v2</b>								
Reviewed: 10/26/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	UOF Incident by Month	Pass	Pass	Pass	Pass	-	-	Pass
2	UOF by Force Type by Month	Pass	Pass	Pass	Pass	-	-	Pass
3	UOF by Area Command by Month	Pass	Pass	Pass	Pass	-	-	Pass
4	UOF by Arrest Type by Month	Pass	Pass	Pass	Pass	-	-	Pass
	UOF by Demographics by Month	-	-	-	-	-	-	-
5	Officer Gender by Month	Pass	Pass	Pass	Pass	-	-	Pass
6	Citizen Gender by Month	Pass	Pass	Pass	Pass	-	-	Pass
7	Officer Age by Month	Pass	Pass	Pass	Pass	-	-	Pass
8	Citizen Age by Month	Pass	*Pass	Pass	Pass	-	-	Pass
9	Officer Race by Month	Pass	Pass	Pass	Pass	-	-	Pass
10	Citizen Race by Month	Pass	*Pass	Pass	Pass	-	-	Pass
11	Officer Ethnicity by Month	Pass	Pass	Pass	Pass	-	-	Pass
12	Citizen Ethnicity by Month	Pass	*Pass	Pass	Pass	-	-	Pass
13	Citizen Gender Expression by Month	Pass	*Pass	Pass	Tabulation Discrepancies: '01-2017' and '09-2017'	-	-	Pass
14	Citizen Sexual Orientation by Month	Pass	*Pass	Pass	Tabulation Discrepancies: '01-2017' and '09-2017'	-	-	Pass
15	Citizen English Proficiency by Month	Pass	*Pass	Pass	Tabulation Discrepancies: '01-2017'	-	-	Pass

<b>C Paragraph 129 - Behavioral Health Incidents Report v2</b>									
Reviewed: 10/25/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity	
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total		
		-	-	-	-	-	-	-	
1	Behavioral Health Incidents by Month	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
2	Behavioral Health Incidents by Month and Shift	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
3	Behavioral health Incidents by Month and Area Command	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
4	Behaviorial health Incidents by Month and Supervisor Responded	Pass	*Pass	Pass	Not Tested	-	-	Column Header 321?	
	Subject Demographics								
5	Subject Gender	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
6	Subject Age	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
7	Subject Race	Pass	*Pass	Pass	Not Tested	-	-	Not Tested	
8	Subject Ethnicity	Pass	*Pass	Pass	Not Tested	-	-	Not Tested	
9	Subject Veteran Status	Pass	*Pass	Pass	Not Tested	-	-	Not Tested	
10	Subject Armed	Pass	*Pass	Pass	Not Tested	-	-	Not Tested	
	Subject Demographics								
11	Behavioral Health Incident w/ UOF by Month	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
12	Type of Force Used by Month	Pass		Pass	Not Tested	-	-	Not Tested	
13	Type of Officer Injury by Month	Pass	Pass	Duplicate column header: 'Abrasions'	Not Tested	-	-	Not Tested	
14	Type of Subject Injury by Month	Pass	Pass	Pass	Not Tested	-	-	Not Tested	
		NOTE: Some fields listed "Not Tested" because reports tables were unaccessible at the time of observation							

D1		Recruiting Events Report v2						
Reviewed: 10/25/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Recruitment Interest Summary	Pass	Pass	Pass	Pass	-	-	Pass
2	Recruiting Events Summary	Pass	Pass	Pass	N/A	-	-	Pass
3	Recruiting Events Detail	Pass	52%	Pass	N/A	-	-	Pass



**D2-6 Well Qualified Recruit Details Report (Cohorts 117,118)**

Reviewed: 10/26/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"		*While tabulation discrepancies were observed, they are assumed to be the result of table organization and presentation, not of any database error	Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Academy Recruitment - Selection Process	Pass	Pass	Pass	Pass	-	-	Pass
	Well Qualified Applicant	-	-	-	-	-	-	-
2	Academy Recruitment - Exceptions Granted	Pass	Pass	Pass	Pass	-	-	Pass
3	Academy Recruitment - Fluency in Languages other than English	Pass	Pass	Pass	*Pass	-	-	Pass
4	Academy Recruitment - Former Law Enforcement Experience	Pass	Pass	Pass	N/A	-	-	Pass
5	Academy Recruitment - Education	Pass	*Pass	Pass	Pass	-	-	Pass
6	Academy Recruitment - Military	Pass	Pass	Pass	Pass	-	-	Pass
NOTE: At the time of this review, RS21 understands that data entry for cohorts other than 117 and 118 have not yet been completed.								
For this reason, cohorts other than 117 and 118 have been omitted from this review.								

<b>E Summary Report v2</b>								
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
						Reported Total	Manually Tabulated Total	
Force Review Board Summary		-	-	-	-	-	-	-
Cases Reviewed by Month and Type								
1	Training Deficiencies	Pass	Pass	Pass	Pass	-	-	Pass
2	Tactical Deficiencies	Pass	Pass	Pass	Pass	-	-	Pass
3	Policy Deficiencies	Pass	Pass	Pass	Pass	-	-	Pass
4	Equipment Deficiencies	N/A	N/A	N/A	N/A	N/A	N/A	N/A
NOTE: Equipment Deficiencies table was not displaying values at time of observation.								

<b>F2</b>		<b>Report v2</b>							
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy				Validity
						Reported Total	Manually Tabulated Total		
		-	-	-	-	-	-	-	-
1	Untitled	Pass	Pass	Pass	Pass	-	-	-	Pass

<b>G1.a BSS Activities Report</b>								
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"					
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Behavioral Science Section Activities Summary	Pass	Pass	Pass	Pass	-	-	Pass
2	Behavioral Science Section Activities by Person(s) Served	Pass	Pass	Pass	Pass	-	-	Pass
3	Behavioral Science Section Activities by Service Mandated	Pass	Pass	Pass	Pass	-	-	Pass

<b>G1.b Peer Support Report v2</b>								
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Peer Supported Services Provided	Pass	Pass	Pass	Pass	-	-	Pass
2	Method of Contacting Peer Support Services	Pass	Pass	Pass	Pass	-	-	Pass
3	Initiating Party	Pass	Pass	Pass	Pass	-	-	Pass

<b>H BlueTeam Communication Review v2</b>								
Reviewed: 10/14/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Communication Summary by Month Incident Occurred	Pass	Pass		Pass	-	-	Pass
2	Not Approved Reason Detail by Month	Pass	*Pass		Pass	-	-	Pass
3	Rejected by Rank Detail by Month	Pass	*Pass		Not Tested	-	-	Pass
4	Sent Back to Rank Detail by Month	Pass	*Pass		Not Tested	-	-	Pass

H Line Inspection Report								
Reviewed: 10/14/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Line Inspection Test - Summary	Pass	Pass	Pass	Pass	-	-	Pass
	Line Inspection Test - Duty Pistol	-	-	-	-	-	-	-
2	Duty Pistol Serial Number Matches File	Pass	*Pass	Pass	Not Tested	-	-	Pass
3	Duty Pistol Correct Department Ammunition	Pass	Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Shotgun	-	-	-	-	-	-	-
4	Shotgun Equipped	Pass	Pass	Pass	Not Tested	-	-	Pass
5	Shotgun Serial Number Matches File	Pass	Pass	Pass	Not Tested	-	-	Pass
6	Shotgun Correct Department Ammunition	Pass	Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Rifle	-	-	-	-	-	-	-
7	Rifle Equipped	Pass	Pass	Pass	Not Tested	-	-	Pass
8	Rifle Serial Number Matches File	Pass	Pass	Pass	Not Tested	-	-	Pass
9	Rifle Correct Department Ammunition	Pass	Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Less Lethal Shotgun	-	-	-	-	-	-	-
10	Less Lethal Shotgun Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
11	Less Lethal Shotgun Serial Number Matches File	Pass	*Pass	Pass	Not Tested	-	-	Pass
12	Less Lethal Shotgun Correct Department Ammunition	Pass	*Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - BB Shotgun	-	-	-	-	-	-	-
13	BB Shotgun Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
14	BB Shotgun Serial Number Matches File	Pass	*Pass	Pass	Not Tested	-	-	Pass
15	BB Shotgun Correct Department Ammunition	Pass	*Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Backup Weapon	-	-	-	-	-	-	-
16	Backup Weapon Equipped	Pass	Pass	Pass	Not Tested	-	-	Pass
17	Backup Weapon Serial Number Matches File	Pass	Pass	Pass	Not Tested	-	-	Pass
18	Backup Weapon Correct Department Ammunition	Pass	Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Electronic Control Weapon	-	-	-	-	-	-	-
19	Electronic Control Weapon Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
20	Electronic Control Weapon Serial Number Matches File	Pass	Pass	Pass	Not Tested	-	-	Pass
21	Expired Electronic Control Weapon Cartridges	Pass	*Pass	Pass	Not Tested	-	-	Pass
22	Holstered on Support Side	Pass	Pass	Pass	Not Tested	-	-	Pass
23	Current Charge Percentage	Pass	Pass	Pass	Not Tested	-	-	Pass
24	Quarterly Upload Conducted	Pass	Pass	Pass	Not Tested	-	-	Pass

<b>H Line Inspection Report</b>								
Reviewed: 10/14/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
	Line Inspection Test - Other Weapons	-	-	-	-	-	-	-
25	40 mm Impact Launcher Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
26	Breaching Kit Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
27	Oleoresin Capsicum Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
28	Baton Equipped	Pass	*Pass	Pass	Not Tested	-	-	Pass
29	Good Repair	Pass	*Pass	Pass	Not Tested	-	-	Title of table unclear
	Line Inspection Test - On-Body Camera	-	-	-	-	-	-	-
30	On-Body Camera Equipped	Pass	Pass	Pass	Not Tested	-	-	Pass
31	On-Body Camera Test	Pass	Pass	Pass	Not Tested	-	-	Pass
32	Ancillary Parts in Good Repair	Pass	Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Vehicle	-	-	-	-	-	-	-
33	Vehicle Mileage PM Past Due	Pass	*Pass	Pass	Not Tested	-	-	Pass
34	Interior Clean	Pass	*Pass	Pass	Not Tested	-	-	Pass
35	Exterior Damage	Pass	*Pass	Pass	Not Tested	-	-	Pass
36	Backseat Clear	Pass	*Pass	Pass	Not Tested	-	-	Pass
37	Trunk Authorized Items Only	Pass	*Pass	Pass	Not Tested	-	-	Pass
	Line Inspection Test - Other	-	-	-	-	-	-	-
38	Untitled	Pass	Pass	Pass	Not Tested	-	-	Table not titled
39	Citizen Complaint Forms	Pass	Pass	Pass	Not Tested	-	-	Pass
NOTE: Due to time constraints, tables in this report were note tested for accuracy.								



H		Video Inspection Report						
Reviewed: 10/14/2017 & 11/08/201		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Video Inspection - Summary	Pass	Pass	Pass	Pass	-	-	Pass
2	Video Inspection - Positive Feedback and Training Issues and Deficiencies Categories	Pass	Pass	Pass	Pass	-	-	Pass
3	Video Inspection Detail	Pass	Pass	Pass	Pass	-	-	Pass

I		CPC Report v2						
Reviewed: 11/08/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"					
						Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	Civilian Polic Complaints Summary (I1)	Pass	Pass	Pass	N/A	-	-	Pass
	Misconduct Complaint Disposition Summary (I2)	-	-	-	-	-	-	-
2	CPOA Findings	Pass	*Pass	Pass	Pass	-	-	Pass
3	Repeated Misconduct Complaints Summary (I4 pt 1)	Pass	Pass	Pass	N/A	-	-	Pass
4	Repeated Instances of Sustained Misconduct Complaints Summary (I4 pt 2)	Pass	Pass	Pass	N/A	-	-	Pass
5	Misconduct Complaint Incident Detail	Pass	99%	Pass	N/A	-	-	Pass

I IA Investigations Report v2								
Reviewed: 10/26/2017		Structure / Consistency	Completeness	Uniqueness	Accuracy			Validity
			*All fields appear "Complete", but with presence of field(s) labeled as "No Data" or "Unknown"			Reported Total	Manually Tabulated Total	
		-	-	-	-	-	-	-
1	IA Investigation Allegation Disposition Summary (I2)	Pass	*Pass	Pass	Pass	-	-	Pass
2	Repeated IA Investigation Summary (I4 pt 1)	Pass	Pass	Pass	N/A	-	-	Pass
2	Repeated Instances of Sustained IA Investigation Summary (I4 pt 2)	Pass	Pass	Pass	N/A	-	-	Pass
3	Criminal Prosecutions of Officers (I5)	Pass	Pass	Pass	N/A	-	-	Pass

## REPORTS TABLES FITNESS ASSESSMENT

As part of this assessment, the authors have reviewed the data sets for perceived fitness in responding to the requirements of paragraph 298 performance reporting. In this exercise, fitness is considered a qualitative term and as such, the observations below are subject to the reasoning and interpretations of the authors and should be considered with regard to the context of each 298 reporting section.

For reporting specific elements specified in paragraph 298 (such as demographics), the authors assume that the Department, the Monitor, and the Department of Justice have worked together to appropriately define the data sets necessary for adequate reporting. In the case of demographics, the reports tables display data related to gender, age, race, ethnicity, gender expression, sexual orientation, and English proficiency. For the purposes of this assessment, it is assumed that this list is comprehensive, fitting, and need not include additional demographic indicators.

### **A1**

Appears fit

**A2** – Not evaluated for fitness

### **A3**

Appears fit

NOTE: Table titles should include more specific language matching paragraph 298 “number of uses of force that violate policy....”.

### **A4**

Appears fit

### **A5**

Appears fit

### **A6 (Citizen)**

Appears fit

NOTE: The following tables are not specified by paragraph 298:

- Force Type Used
- Citizen Injury by Arrest Type

## **A6 (Officer)**

Appears fit

NOTE: The following tables are not specified by paragraph 298:

- Force Type Used
- Officer Injury by Arrest Type

**A7** – Not evaluated for fitness

## **B1**

Appears fit

NOTE: Paragraph 298 specifies “activations and deployments...”, suggesting that the two categories of data should be reported as separate counts.

## **B2**

Appears fit

NOTE: The following tables are not specified by paragraph 298:

- Use of Force by Arrest Type

**C** – Not evaluated for fitness

## **D1**

“Recruiting Events Detail” table does not seem to capture all events listed in the “Recruiting Events Summary” table. Also, while community organizations are mentioned in the title and description columns of the “Recruiting Events Detail” table, the report does not directly respond to the 298 specification of “development and leveraging community partnerships”, which could be interpreted as requiring separate counts for development of community partnerships and for leveraging those partnerships.

## **D2-6**

Reports D2-6 were reviewed for only Cohorts 117, 118 because APD stated data entry was still in progress at the time of the assessment.

Paragraph 298.d.ii specifies “number of recruit applicants who failed to advance through the selection process after having been identified as well qualified, grouped by the reason for failure to advance...”. While the corresponding report table displays the groupings of recruit applicants who failed to advance, it does not display the total number of recruit applicants who failed. The groupings of applicants who failed can be manually tabulated to generate a total, but the authors recommend the addition of a totals column to more directly respond to 298.d.ii.

The table responding to 298.d.v “number of well-qualified recruit applicants with previous law enforcement experience, grouped by former agencies and years of service” lists “Police Department” and “Sheriff Department” as former agencies. For additional clarity, the authors recommend specifying which Police Departments and which Sheriff Departments these applicants served (e.g. Albuquerque Police Department and Bernalillo County Sheriff’s Office).

NOTE: While each of the tables in D2-6 responds to a separate section (2-6) of 298.d, it may be useful to also highlight any recruits which span multiple categories. For example, it would be difficult for the reader of the reports tables to assess whether any recruits speak multiple languages AND have former law enforcement experience AND attained education of a certain level AND possess former military experience. This level of detail is perhaps not explicitly necessary for responding to paragraph 298, but could be valuable for better measuring and communicating the characteristics of recruit cohorts.

## **E**

Appears fit

NOTE: The following tables are not specified by paragraph 298:

- Equipment Deficiencies

**F1** – Not evaluated for fitness

## **F2**

Due to the esoteric nature of the column headings, it is difficult for the authors to evaluate the fitness of this report.

## **G1**

Appears fit

## **H – BlueTeam Communication**

Appears fit (based on assumptions)

NOTE: APD should consider revising report tables titles and column headers to more directly align with the language of paragraph 298.

## **H – Line Inspection**

Appears fit (based on assumptions)

NOTE: APD should consider revising report tables titles and column headers to more directly align with the language of paragraph 298.

## **H – Video Inspection**

Appears fit (based on assumptions)

NOTE: APD should consider revising report tables titles and column headers to more directly align with the language of paragraph 298.

## **I – CPC & I – IA**

Data sets do not appear to respond to the paragraph 298.I.i specification for “whether any increase or decrease appears related to access to the complaint process;”.

Data sets do not appear to respond to the paragraph 298.I.iii specification for “number of misconduct complaint allegations supported by a preponderance of the evidence;”.

## SQL AND SQL CODES

As part of this observational assessment, RS21 has attempted to replicate data found in the 298 SharePoint responses based on the version two SQL code. The authors did not examine every data warehouse data table and its relationship, only those tables listed in Appendix 1. APD's SharePoint drive is a 'live' document, meaning that any updates to the SharePoint content without updating the SQL code responsible will result in irreproducible data. This can be an issue when outside sources attempt to replicate the data being presented in a 298 response with outdated SQL code, as was the case with RS21's SQL and data warehouse assessment.

Lastly, the authors examined the structure of the data warehouse as it relates to the parent and child relationship. Parent-child relationships describe a hierarchy where parent datasets are joined to child datasets. Under typical data structure conditions, it is expected that each child has a parent and that there are no children without a parent database relationship.

Because updates to the SharePoint data do not automatically update the SQL code, better documentation of what data changes are made in relationship to changing SQL code are necessary. RS21 executed the SQL code in the data warehouse to examine ease of execution, structure of the code, and terminology used. The code was executable and structured in such a manner that an intermediate SQL user could decipher the code. However, some of the terminology might be easily misconstrued. For example, the difference between "case" and "incident" was unclear. Better documentation of terminology would provide a clearer picture of how the SQL code creates the SharePoint data outputs.

The replication test found that the "CITResourceProvided" and "IncidentActionsTaken" tables contained rows which did not belong to a parent table. The authors also examined childless parents in the data warehouse but determined that cataloging any childless parent data tables would rely too heavily on speculation at this juncture, lacking a more definitive data dictionary.

The authors recommend that APD identify if there is a parent relationship with any orphan child or childless parent dataset. If there is, a relationship connection needs to be established. Further data warehouse management should include quality checks for orphan child and childless datasets. RS21 recommends that APD further clarifies their data definitions. As stated above, it was impractical to assess the number of childless parents because data definitions describing this relationship were too vague.



## DATA QUALITY ASSURANCE (DQA) PLANS

There are three main components in best practices for Data Quality Assurance Plans: Framework, Internal Quality, and External Quality.

### **Framework**

The framework of a DQA plan should include the purpose, goals, definitions of terms, and qualifications of data personnel (Jones, Ross, Ruusalepp, & Dobрева, 2009). Using these four elements, the DQA framework establishes important data management and communication procedures to help guide internal and external data quality. The DQA framework also establishes a resolution for the intent of the DQA plan and its desired outcome. In addition, the DQA plan should include detailed qualifications of DQA authors and those interacting with the data at the different data life cycle stages, including data entry, processing, storage, usage, reporting, and deletion (Goben & Raszewski, 2015; United States Geological Survey, 2010).

### **Internal Quality – Data Quality Review, Correction, and Verification**

A plan for assessing the quality of data is the heart of DQA. Data quality monitoring should include three main aspects: review, correct, and verify (EPA, 2002; USDOJ, 2003).

First, DQA plans should include examination and review of data quality dimensions in each data life cycle stage since data errors can arise in any of these stages. Monitoring the data based on these data quality dimensions ensures the appropriateness of the data. Furthermore, this review process provides a framework to identify exactly where data quality issues arise (e.g., upstream or downstream of the database) and, consequently, where they need to be addressed (e.g., during entry, processing, storage, usage, reporting, or deletion). For instance, identifying a data field which is consistently missing a value in the data entry stage suggests that the data quality issue may have occurred from personnel entering data incorrectly or through a computer malfunction (assuming the personnel electronically report data).

Second, after data quality dimensions have been reviewed, appropriate corrective actions need to be taken and documented. Documenting corrective actions creates a roadmap describing changes made to data. This step ensures accurate data.

Third, verification plans help ensure credibility and integrity of data. Data verification practices are often based on establishing standards and procedures, documenting data entry and transfers, creating protocols for data security and integrity, and understanding and documenting data limitations.

### **External Quality – Public’s Role in Data Transparency and Metadata**

DQA plans should contain elements that describe the public’s role in data transparency and standards for metadata creation. Transparency can be defined as the disclosure of information which allows external actors to assess and ultimately monitor decisions and performance of the information authors (Welch, 2012). A study by Grimmeliikhuijsen (2010) found that increased transparency increased trust among citizens with little subject matter knowledge and little

initial trust. However, their study also found that increased transparency decreased trust among citizens with little subject matter knowledge but initial high trust in the government. Many cities are beginning to share law enforcement data with the public, a form of government transparency which is likely to enhance public trust (O’Hara, 2012). Such practices have had many positive outcomes, including public-private partnerships, increased accountability among law enforcement, improved criminal justice systems, and greater understanding of the complexities of law enforcement data among members of the public (O’Hara, 2012). However, data transparency alone may not be entirely effective at creating data accountability (Lindstedt & Naurin, 2010). The relationship between government transparency and public participation is discussed by numerous academics; the APD can use this discussion as a guide to help frame their approach to transparency in the DQA plan.

### **APD DQA - Overview**

RS21 has reviewed ADP’s DQA plans (as of 10/20/2017) as included on the APD’s SharePoint website for each of their DOJ 298 responses. Several of APD’s 298 responses did not contain a DQA plan (shown in Appendix X). Each reviewed DQA plan is outlined in a similar structure: Introduction, Quality Objectives, Management, Documentation, and Documentation Reviews. Many of the DQA plans contain identical language (copy and paste approach), with minor changes applied to reflect different data content between APD’s 298 responses. This may not be a concern as it may be a legacy of a broader DQA plan that governs APD’s 298 responses.

### **RS21 Observations – Overall Content, Structure, and Readability**

RS21 first examined the different elements found within APD’s 298 response DQA plans. APD has included a purpose and scope as to why they are generating a DQA plan (Section 1). This not only provides context for the report, but also helps frame the objectives or goals. Section 2 of APD’s DQA plan focuses on external Standard Operating Procedure (SOP) documents and BlueTeam data entry instructions that act as quality objectives. Clear data quality objectives are an important element in a DQA plan as they provide a roadmap for how data errors are recorded and corrected. Section 3 involves data management, specifically defining members of the audit group, their responsibilities, points of contact, and review processes. This, too, is an important section, providing readers assurance about the credibility of the DQA plan and approach to ensuring data quality. Section 4 is the data documentation procedure, which helps establish data transparency. Finally, APD’s DQA plan discusses documentation reviews (Section 5), including a review plan, review schedule, organization responsibilities, problem resolution and correction, tools, techniques, methodologies, and quality records. Again, this is a key section in helping guide data quality assurance.

RS21 next examined each element of APD’s DQA plan to assess how APD engages with data quality practices.

For some 298 responses, the DQA plan relies on the user viewing an external report as a supplemental document. For example, Section 2 – Quality Objectives states, “The quality objectives are laid out in the SOPs and in the Blue Team entry instructions. The quality of the data will

be determined based on such.” Our DQA plan assessment does not examine or critique any of these external supporting documents in detail. However, it would be beneficial if the DQA plan contained a hyperlink directing the reader to the appropriate supporting document. There are several additional issues that become clear when examining this section. First, there are some DQA plans that cite SOPs or BlueTeam data entry instructions but supporting documentation is not provided for that specific 298 response. Second, SOP supporting documentation found in the SharePoint 298 reports is not clearly labeled. This makes it difficult for the reader to identify which supporting document is the SOP. Third, it is unclear how the BlueTeam data entry instructions meet data quality objectives. Fourth, the SOPs provided by APD are not SOPs for data quality management, but rather SOPs for law enforcement practices. While this helps ensure police are practicing the correct procedures in the field, it does not help ensure the best data quality.

All of APD’s DQA plans define several law enforcement terms, respective to each 298 response, which would otherwise be unclear to lay people. These definitions ensure that clear data communication is achieved between the APD and outside audiences. However, there are still instances in which the DQA should include definitions for acronyms and terminology that are not universally known. For instance, in the DQA plan APD uses the term “SOP” (standard operating procedure). This acronym may not be commonly understood by the public and thus should be spelled out in the DQA definitions. In addition, Section 5.4 – Problem Resolution and Correction Action includes the following passage: “Positive identification of any discrepancies will be promptly reported to the proper chain of command for further problem resolution and corrective action.” While the verbiage of this section is likely in development, there are opportunities for definition clarification, additional depth, and overall refinement. For example, what is the specific definition of “prompt” reporting? What do “problem resolution and corrective action” entail? Other subsections under Section 5 contain similarly vague language which should be further refined over time.

### **RS21 Observations: Framework**

Practices for framing a DQA plan usually include four main elements: the purpose, goals, definitions of acronyms and terms, and qualifications of data personnel. However, many of the acronyms and terms are not clearly defined, nor was it clear how the listed qualifications of data personnel related to the DQA plan.

### **RS21 Observations: Internal Quality - Data Quality Review, Correction, and Verification**

DQA plans often include three main components for ensuring data quality: reviewing, correcting, and verifying data quality. The content of each APD DQA plan does not contain a detailed plan for these three components. This will necessarily become one of the APD’s DQA centerpieces as they further refine their DQA plan, as it creates not only accurate data, but also mechanisms for internal data accountability.

The first component involves reviewing the different dimensions of data quality at each data life cycle stage, including entry, processing, storage, usage, reporting, and deletion. In terms of data entry, it is unclear if APD’s data entry software helps minimize data quality dimensions. In terms of data storage and processing, APD uses SQL databases. As mentioned above, the

current SQL processing does not check for data quality errors. Lastly, the APD uses SharePoint to access, share, and report data. Unfortunately, the data presented in their SharePoint is generated by an SQL process which in its current state does not take data quality dimensions into consideration.

The second component involves recording and correcting identified data quality errors found during review (component one). APD's DQA plan briefly states their approach toward data correction.

The third component includes a verification plan to check the reliability of the data at all stages of the data life cycle. APD has conducted data verification on several 298 responses. The scope and level of detail in this verification process remains unknown to RS21, as the verification process is not documented.

Combined, these three components create a mechanism for internal data accountability by documenting the details of data changes (reviewing and then correcting inaccuracies, incompleteness, etc.) and where recurring errors are located. For instance, there might be consistent data errors in the entry stage of the data life cycle. This may represent improper data entry training or other issues, in which case appropriate training or disciplinary action are needed, respectively.

#### **RS21 Observations: External Quality – Transparency and Public Participation**

Ultimately, fully-matured DQA plans should contain detailed explanations of data transparency, public participation, and metadata, all of which help create external data credibility. Currently, APD's DQA plans do not contain specific documentation on data transparency and public participation, but the authors understand that the Department is currently developing a public-facing website (APDreform.com) where, among other information, paragraph 298 reporting data will be made available for viewing and download. As this website develops, the authors recommend that the Department consider including documentation of APDreform.com and other activities related to transparency and public participation directly in its DQA plans.

## **RS21 Suggested Improvements – Overall Content, Structure, and Readability**

RS21 has three recommendations for the overall content, structure and readability.

First, RS21 recommends that APD clearly label all supporting documents, provide hyperlinks in each DQA plan to appropriate supplemental documentation, and clearly present the data quality objectives in all SOPs and BlueTeam data entry instruction documentation.

Second, terms and acronyms that are more clearly defined are more easily accessible. More refined definitions of what is now vaguely defined terminology in the DQA plan will bolster the plan's ability to guide data quality. Some of the supporting documentation contains definitions for several terms but, as mentioned above, it is frequently difficult to identify the relevant document holding these definitions.

Third, RS21 suggests that careful proofreading of the DQA plan is required for a more reader-friendly presentation of DQA plan elements and content. For example, in the 298 G1 Behavioral and Sciences Section DQA plan, the description of the collected data is at the end of the 1.2 Scope section. Due to this placement, the reader is not easily able to parse this section of the document. Providing detail on the data collected is a helpful component to include in any DQA plan, but the placement of this information should be logical and clear to the reader.

### **RS21 Suggested Improvements: Framework**

**Statement of Purpose** – Each DQA plan for APD's 298 response has a general purpose stated that rationalizes its creation. RS21 has no recommendations regarding APD's DQA plan's purpose.

**Statement of DQA goals** – DQA plans should clearly state their goals as these goals can vary significantly. RS21 recommends that APD establish and document a long-term goal for structuring, managing, and implementing a DQA plan. Some of the content provided by RS21 in this report may be useful in drafting such a goal.

**Full scope of definitions** – terminology and acronyms that are unfamiliar to lay people need to be clearly defined. Furthermore, documentation that houses collections of terms and acronyms and their definitions needs to be easily accessible. Currently, APD's DAQ plans recommend that readers view external supporting documentation that defines a number of terms. However, this documentation is difficult to access. Supplemental documents that define multiple terms need to be accessible within the DQA plan, and therefore, RS21 recommends that as APD refines their DQA plan, they create more accessible documentation.

**Qualification of data personnel** – APD's DQA plans give little detail on the qualifications of those individuals analyzing APD's data. It is vital that the qualifications of staff involved with any process of ensuring data quality are explicitly laid out. This helps create DQA plan credibility

with the public. RS21 suggests that APD increase the depth to which they describe DQA staff qualifications in addition to creating qualification and training standards for those involved.

### **RS21 Suggested Improvements: Internal Quality - Data Quality Review, Correction, and Verification**

As APD further refines their DQA plans, they should consider several key elements that will help promote a higher standard of data quality.

*Data Review:* Currently, APD's DQA plan does not contain a plan for reviewing different data quality dimensions in the various data life cycle stages. We suggest that APD create a strategy to review and identify any data quality errors, including incompleteness, duplicity, invalidity, inaccuracy, and inconsistencies at different data life cycle stages. As APD continues to develop their DQA plan, they can determine the importance and definitions of additional data quality dimensions.

To help manage data quality, it would also be beneficial for APD to seek a single centralized piece of software that reviews data quality through the different data life cycle stages. APD currently uses BlueTeam for data entry, SQL databases for data storage and processing, and SharePoint for data access, sharing, and reporting. As a standalone process, each of these has many strengths, but the combination creates avenues where the integrity of data quality dimensions can dissolve. For instance, in its current state the SQL code used to generate APD's 298 responses does not check for data quality errors. Furthermore, using three different data systems requires training personnel on how to properly interact with each piece of software to avoid possible data errors. This is time consuming, expensive, and more likely to increase data errors. Instead, centralized software should be employed at all stages of the data life cycle (entry, processing, storage, usage, reporting, and deleting) helping eliminate the possibility of data errors since personnel will be trained on a single data management platform. This platform will also need built-in functions that not only minimize errors of the data quality dimensions, but also help process, report, share, and ultimately access the data. For instance, after the software processes field officer's data, it would begin the data review process that identifies data errors and then automatically generate a report for the appropriate audiences to communicate data errors.

*Data Correction:* APD's DQA plan briefly states their approach towards data correction. RS21 recommends that APD expand on the detail of their corrective procedures. This expansion would include methods for correcting each of the different data quality elements at each data life cycle stage, a schedule for when corrections are to be completed, the party responsible for correcting data errors, and any potential data misinterpretations that could result from uncorrected data.

*Data Verification:* APD currently has a data verification approach, but the details of this approach remain unknown to RS21. Detailed documentation of this process is necessary since it tests the reliability of an organization's data and data management practices. Sound verification

approaches help answer questions around items such as established procedures for checking the quality of data, data error correction, and what, if any, software is used. RS21 recommends that APD document their verification process in greater depth and detail. This will increase data credibility since the verification process acts as a tool to check data reliability. As with all components of this report, APD should examine the vast quantity of scientific literature available on data verification elements. Some verification reporting likely to be of interest to APD include the United States Department of Interior (2003) and the US Environmental Protection Agency (EPA). After examining these and other data verification reports, APD should collaborate with members of the public to create a robust data verification plan with appropriate documentation.

Once APD has established data recording, correction, and verification documentation, they will be able to better construct accountability plans. Some recommended elements of the accountability plan include data training standards, appropriate disciplinary procedures, and public disclosure procedures. These and other elements should be incorporated into a standard operating procedure plan for data management.

### **RS21 Suggested Improvements: External Quality – Transparency, Public Participation, and Metadata**

In recent years, academia and government agencies have been rigorously assessing the impact of government open data initiatives, transparency, public participation, and metadata. The relationships between these initiatives is highly complex. Broadly speaking, transparency, public participation, and metadata are all DQA elements that increase data accountability and therefore data quality. RS21 understands that APD is currently working to develop a public-facing website (APDreform.com) as a public-engagement effort. RS21 recommends that the Department leverage this effort to build a collaborative environment between APD and the public that helps flesh out a mutually agreed upon approach for data transparency. This collaborative effort helps ensure that the data released to the public meets their needs and follows a trusted and rigorous DQA plan. For instance, the public may want access to data about use of force cases by specific location (point/latitude and longitude data) and who committed the use of force for public analysis. APD, on the other hand, may want to report that information based on the total number of city-wide cases to protect the identities of victims. In this situation, a possible compromise might be reached to report use of force cases at the neighborhood level, preserving the anonymity of victims but providing enough detail for police accountability.

## METADATA

Metadata can be defined as information that describes a dataset, helping add context to the data. Well-maintained metadata provides users with the necessary context to understand what the data represents, how it was collected, any potential inaccuracies, data definitions, and other elements (Greenberg, White, Carrier, & Scherle, 2009; Palmer & Knutson, 2004; Park & Tosaka, 2010; Smithsonian Libraries, n.d.; University of North Carolina, n.d.).

APD provides metadata with each 298 response as shown on the SharePoint website. Each metadata report contains the following fields: Report Title, Link to Report, Paragraph Reference Number, 298 Contents, Original Data Sources, Alternative Data Sources, Data Contracts, Contact Divisions, Date Report Created, Group Approval Date, Notes, Data Production Methodology, SQL Code Version 1, SQL Code Version 2, Data Auditor, Assumptions, Assurance Plan, Assurance Plan Contact, Disclaimers, Report Description, Warehouse Refresh Frequency, Report Reference Number, View Tasks, Create Tasks, Monitor Report Page Number, Attachments, Created at, and Modified at.

Values are often missing in several of these fields, including Alternate Data Sources, Group Approval Date, and Assurance Plan Contact. It is uncertain if these fields are missing due to the occurrence of internal discussions about the field's content, a lack of appropriateness, or some other reason. In this situation, "missing" refers to the fact that the field is "blank" and not filled-in with "N/A" or some other spacer value. This gives a sense that the metadata file is incomplete.

Metadata documentation has become more crucial and standard practice in recent years as it helps add important context for outside consumers such as the structure, quality, confidentiality, and terms of use. RS21 has several recommendations to improve APD metadata documentation. First, metadata fields should not have incomplete entries; if a field within the metadata file does not have an applicable value, then it should be filled in accordingly with "N/A". Second, if metadata files are going to contain information on two different versions of the data, this needs to be clearly communicated to the audience. This was not clear in APD's reports as the metadata related to separate versions of the reports were not clearly labeled. The authors suggest that APD only include metadata which is relevant to the data being reported. For example, the most current version of SQL code should be included in the metadata file for the most current 298 report tables. 'Old' versions of the 298 report tables and SQL code should not be included in current versions. Rather these 'old' data versions should have their own separate location where their metadata and other documentation can be archived and referenced as needed. The authors assume that APD is aware of this issue and is in the process of mitigating these duplications. Third, RS21 suggests that it might be useful for APD to identify additional elements of metadata by engaging with the public, which may provide insight into what metadata elements best suit their needs as external data auditors.



## MODIFICATION POLICY

In addition to maintaining and adhering to robust data quality assurance plans, it is crucial for an organization to follow strict procedures for how and when the contents of a report or associated metadata may be modified. Such policy helps to minimize potential miscommunication or clerical errors associated with changes to content within the reporting system.

During the assessment period, the authors observed changes made to the content of the SharePoint in a seemingly short period of time. Because the authors of this report are not intimately familiar with existing APD policy governing modifications or with the specific roles and responsibilities of all APD personnel assigned to coordinate and approve such modifications, it is difficult at this point in time to assess the fitness of existing modifications policy and procedures.

In either case, the authors recommend that APD review and, if necessary, strengthen the policy in place for approving and implementing any modifications to the reports tables and associated metadata on the SharePoint site.

Conceptually speaking, the following should be considered when developing or reviewing a modification policy:

- Formal review of proposed modifications should be conducted before any change and formal validation should be conducted after the change is implemented
- Multiple personnel should participate in the review and validation; where appropriate, this should include the data contact associated with the data set(s) in question
- Modifications should not be made based on “memory” or anecdotal discussion
- For any modification, three activities should be documented:
  - 1.) Proposed modification
  - 2.) Formal approval by the appropriate supervisors and/or data contact(s)
  - 3.) Implementation of the modification

## DATA COLLECTION

While, as previously mentioned, the scope of this assessment does not directly include investigation specific to APD data collection and data entry processes, it is worth noting the authors' observations in the reports tables which could suggest room for improvement in future data collection procedures.

Data quality begins at the point of initial data capture (e.g. field reports). When using data to illuminate operational patterns and drive operational change, it is crucial to use the most thorough data available at the time. More thorough data collection can facilitate more robust analytics with greater reliability and trustworthiness. More robust analytics, in turn, can be used to drive more effective decisions with regard to operational change strategies.

At the time of writing, a number of reports included data values in columns titled "No data" or "Unknown". It is important to draw the distinction that the presence of "No data" or "Unknown" values does not necessarily suggest any error in the data management or reporting system.

Additionally, APD expressed to the authors that many instances of "No data" may be attributed to the circumstance that the Department did not collect certain points of data before the CASA was implemented in 2014 and that, since then, there has been a period of organizational change during which data collection has been expanded over time to include new types of data. In the scope of this assessment, it was not feasible to validate this pattern beyond noting the Department's comments.

As observed by the authors, the number of values in these categories relative to the total values in a given table ranged from less than 0.5% to over 75%. When this percentage is low, it should not pose a substantial hindrance to analytics conducted on the data set. However, when the percentage is high, operational analytics generated from the data are less useful because the analytical findings become increasingly uncertain. In turn, highly uncertain findings are less actionable and less likely to identify true priorities for operational change.

While it may not be possible to decrease the presence of "No data" or "Unknown" values for data sets collected in the past, the authors recommend that APD review existing data collection policies and, if necessary, consider strategies to minimize these occurrences in the future. Because there is no such thing as a perfect data set, we would not expect to see a complete elimination of "No data" or "Unknown" values, but improvement in this area will serve to support increasingly robust analytics moving forward.

Generally speaking, "No data" values are more problematic in analytics than "Unknown" values. As such, the authors recommend that APD prioritize occurrences of "No data" before "Unknown" values in its policy review.

## REPORTING AND COMMUNICATING DATA INSIGHTS

Although the subsections of paragraph 298 do not explicitly specify the manner in which performance metrics are to be formatted or communicated, the authors offer the following narrative addressing possible trade-off decisions in communicating insights from raw data. There are a number of different approaches to communicating data for different types of consumption by different types of audiences and analytical purposes. For example, formatting demands are different if data is used for documenting raw numbers and progress over time than they would be if data were being used to drive daily management. Similarly, we would expect it commonplace to see different formatting strategies for internal analysts versus the general public, for which APD is currently developing a public-facing website (APDreform.com), which will feature highly accessible visuals to communicate paragraph 298 data to the community.

In a data-driven organization, communicating insights from data is just as important as collecting, managing, and analyzing it. To drive effective operational change decisions, data (symbols) must be converted to information (processed data), and then to knowledge (information applied to answer questions), and then to understanding (Ackoff, 1989). Effective data communication bridges the gap between data and understanding, taking into account the insights available in the data and the unique needs and personalities of the audience(s).

In the case of APD's 298 reporting, there seem to be a number of possible audiences—each with unique and potentially conflicting informational needs. This can make it difficult to format the data most clearly for all audiences. For example, an auditor may need to see every possible row and column to verify that all cells in a table are complete, but a citizen viewing the data from a less technical perspective could be easily overwhelmed and confused by seeing unnecessary cells (e.g. cells with zero counts). The Department has expressed to the authors that formatting decisions have been made with a priority for simplifying the reading of the tables for non-technical audiences.

Currently, the 298 reporting data is organized into a series of matrixed tables (rows and columns) as shown below:

### UOF Incident By Month

Investigation Completed Date	Count
07-2016	1
09-2016	1
10-2016	1
01-2017	1
07-2017	1
Total	5

**Typical Reports Table with intermediate dates with assumed zero values strategically omitted to simplify viewing.**

This method of communicating the data is one of the more basic approaches. While this type of matrixed structure may be simple to manage, it can be difficult for some audiences to ascertain

understanding of the data beyond the direct numbers of occurrences per month. For example, in nearly all of the reports tables, the primary goal is to track operational change over time (trends). While a consumer of this data table could look, row-by-row, through this table and begin to understand some indications of a trend, the format of the table itself makes this process a challenge.

Firstly, if a reader is interested in quickly identifying trends, there are a number of visualization types which are better equipped to effectively communicate changes over time. Simple line graphs and bar charts (histograms) generally present the same information in a more accessible and intuitive manner, making it easier for the reader to clearly see trends.

Secondly, in all tables observed in this assessment, a structural decision has been made to omit any rows or columns which contain all zero values. This strategy comes with a series of tradeoffs. On one hand, it minimizes visual clutter by only displaying rows and columns which have non-zero values. Theoretically, this can make it easier for the reader to process the values which are displayed, because there is less “noise” from row after row of zeros.

However, this strategy can complicate verification exercises. In the table above, it is clear to see that many months in the date range are not displayed. As independent observers, we assume that, to simplify the reading of the table, the months not displayed have been intentionally omitted because the counts associated with those months are all zero. This may well be true, but if looking at this table alone, it is not possible to verify whether all data for this date range is being displayed. As an example, if there were an error in the production of the table which incorrectly omitted a month with a count more than zero, then the report would be inaccurate. If the common strategy for organizing this table were to display every month in the given range (even those with counts of zero), an auditor reviewing the reports tables would have greater evidence to verify that data existing for all months in the date range was being displayed.

In some instances, there could be an unmanageable number of possible rows and columns if every possible category was displayed. For example, in the table of injury types below, there could be hundreds of possible injury types. Displaying every possible injury type as its own column with all zero values would create an unnecessary burden to both the APD team and the readers of the data reports. In this case, it makes sense to omit categories with only zero values.

**Officer Injury by Injury Type by Month**

Occurred Date	Abrasions	Biohazard Contaminatio	Bite Marks	Broken Bones	Bruises	Gun Shot	Lacerations	OC exposure	C
01-2016	5	0	0	1	0	0	0	0	0
02-2016	9	0	0	0	0	0	0	0	0
03-2016	7	0	0	0	0	0	2	0	0
04-2016	3	0	0	0	0	0	1	0	0
05-2016	8	0	0	0	0	0	0	0	1
06-2016	9	0	0	0	0	1	0	0	0
07-2016	6	0	0	0	0	0	1	0	0
08-2016	17	3	0	0	1	0	3	0	0
09-2016	6	0	1	0	0	0	2	0	0
10-2016	3	1	0	0	1	0	3	0	0
11-2016	2	0	0	0	0	0	1	0	0
12-2016	8	0	0	0	3	0	2	0	0
01-2017	4	1	0	0	2	0	1	2	2
02-2017	2	0	0	0	0	0	1	1	1
03-2017	5	0	1	1	4	0	1	1	1

**Typical Reports Table for which it would be unmanageable to display every possible column for which there are only zero values.**

However, in other instances, it may be more manageable to curate tables to include a more comprehensive set of rows and columns, even where entire rows or columns display only zero values. For example, in a table of genders, we would expect it to be feasible to include (at a minimum) separate columns for male, female, and other gender. In this case, the number of additional columns added to the table would be both manageable and legible for the reader without excessive visual clutter.

Ultimately, the decision of which rows and columns to include in a given table should be driven by the nature of the information at hand, the needs of the reader (context), and the level of effort required to provide adequate context. When it comes to effective visual communication, excessive visual clutter can be just as problematic as sparse information. Curators of data reports must continually strive for a balance in this realm.

Where appropriate for a given audience, the authors recommend that the Department look to provide charts, graphs, or other visual mechanisms to improve communication of the data to the appropriate audiences. Reports tables for 298, section G have begun to include histograms, which make it far easier for a reader to assess for trends over time. Visuals need not be complicated, but could make a big difference in the way the data can be processed by a reader.

## CONCLUSION

To reiterate, the scope of this assessment did not include validating the full spectrum of data management from the initial point of collection to final reporting. Rather, this report reflects observations by the authors looking only at the data management and reporting mechanisms which have been implemented by APD in recent years—specifically, the data warehouse and paragraph 298 reports tables.

As previously mentioned, the Albuquerque Police Department has made great strides in fortifying its data management infrastructure and data quality practices by adopting many industry standard practices such as implementing the data warehouse, developing data quality assurance plans, recording and maintaining crucial metadata, and boosting staffing to increase internal auditing activities and data quality scans.

As far as the authors can see, these accomplishments have required a substantial level of organizational change over the past years and represent a fundamental shift in the Department toward becoming a data-driven organization. These improvements should help to increase the sustainability of data quality over time and should serve the Department well as it continues to build upon this solid foundation.

In addition to the improvements the Department has already implemented, the authors have observed and highlighted a number of remaining issues which we believe deserve refinement by the Department moving forward.

Some of the potential issues identified are perhaps more minor and many are subject to contextual conditions such as the timeframe in which APD began collecting certain data sets. For example, the authors' observations of multiple versions of reports on the SharePoint site may be of little concern in the very near future, as APD has communicated that it is already in the process of addressing these occurrences and they should be resolved shortly. Some issues discussed in this report, such as the occurrence of "No Data" and "Unknown" values, may be infeasible to improve upon, because it is likely impossible to retroactively replace these values which were entered into the system years ago.

Other recommendations outlined in this report—such as continuing to develop increasingly robust data quality assurance plans for every report—should be considered more critical to long term sustainable data quality management practices and should be prioritized by the Department. Likewise, APD has already generated valuable metadata for its paragraph 298 reports, but should continue to keep the creation and maintenance of thorough and clear metadata as a top priority in quality assurance as it looks to the future.

Already, the Department appears to be responding quickly to many of the issues observed during the assessment. As APD strives to become a national leader in police operations data management and analytics, the authors intend that the recommendations offered in this report will serve as guiding provocations to assist the Department by helping to further develop data collection, data management, and data reporting practices into the future.

## REFERENCES

- Ackoff, R.L. (1989). From data to wisdom. *Journal of Applied Systems Analysis* 16 (pp. 3-9).
- Cykana, P., Paul, A., & Stern, M. (1996). DoD Guidelines on Data Quality Management. In *IQ* (pp. 154–171).
- Eckerson, W., (2017). Four Ways to Build a Data Warehouse. *Business Intelligence Best Practices*.
- EPA. (2002). *Guidance on Environmental Data Verification and Data Validation*. Retrieved from <https://www.epa.gov/sites/production/files/2015-06/documents/g8-final.pdf>.
- Gardyn, E. (1997). A Data Quality Handbook for a Data Warehouse. In *IQ* (pp. 267–290).
- Goben, A., & Raszewski, R. (2015). The data life cycle applied to our own data. *Journal of the Medical Library Association: JMLA*, 103(1), 40.
- Gostin, L. O. (2006). Federal executive power and communicable disease control: CDC quarantine regulations. *Hastings Center Report*, 36(2), 10–11.
- Greenberg, J., White, H. C., Carrier, S., & Scherle, R. (2009). A metadata best practice for a scientific data repository. *Journal of Library Metadata*, 9(3–4), 194–212.
- Grimmelikhuijsen, S. G. (2010). Transparency of Public Decision-Making: Towards Trust in Local Government? *Policy & Internet*, 2(1), 5–35.
- Jones, S., Ross, S., Ruusalepp, R., & Dobreva, M. (2009). Data audit framework methodology. *Humanities Advanced Technology and Information Institute (HATII), University of Glasgow*, 26.
- Kaiser, M., Klier, M., & Heinrich, B. (2007). How to measure data quality?-a metric-based approach. *ICIS 2007 Proceedings*, 108.
- Kimball, R., Reeves, L., Ross, M., Thornthwaite, W. (1998). *The Data Warehouse Lifecycle Toolkit Expert Methods for Designing Developing and Deploying Data Warehouses*.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: a methodology for information quality assessment. *Information & Management*, 40(2), 133–146.
- Lindstedt, C., & Naurin, D. (2010). Transparency is not enough: Making transparency effective in reducing corruption. *International Political Science Review*, 31(3), 301–322.
- Loshin, D. (2006). Monitoring data quality performance using data quality metrics. *Informatica Corporation*.
- O’Hara, K. (2012). Transparency, open data and trust in government: shaping the infosphere. In *Proceedings of the 4th annual ACM web science conference* (pp. 223–232). ACM.
- Palmer, C. L., & Knutson, E. M. (2004). Metadata practices and implications for federated collections. *Proceedings of the Association for Information Science and Technology*, 41(1), 456–462.
- Park, J., & Tosaka, Y. (2010). Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. *Information Technology and Libraries*, 29(3), 104.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211–218.
- Smithsonian Libraries. (n.d.). Metadata best practices. Retrieved from <https://library.si.edu/research/metadata-best-practices>

United States Geological Survey. (2010). *Data Lifecycle Models and Concepts*.

University of North Carolina. (n.d.). Metadata for Data Management: A Tutorial. Retrieved from <http://guides.lib.unc.edu/c.php?g=8749&p=44503>

USDOI. (2003). *DATA VALIDATION AND VERIFICATION ASSESSMENT MATRIX*. Retrieved from [https://www.fws.gov/Planning/Documents/Validation/DOI\\_Performance\\_Data\\_V&V\\_Guide\\_Jan03.pdf](https://www.fws.gov/Planning/Documents/Validation/DOI_Performance_Data_V&V_Guide_Jan03.pdf)

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5–33.

Welch, E. W. (2012). The relationship between transparent and participative government: A study of local governments in the United States. *International Review of Administrative Sciences*, 78(1), 93–115.



## APPENDICES

### **Appendix 1 – SQL Data Tables Examined (10-25-2017)**

CITIncident  
CITArrest  
CITInjury  
CITMentalIllness  
CITResourceProvided  
CITRisk  
CITSubstance  
CITTechniqueEquipmentUsed  
CITWeapon

Incident  
IncidentActionsTaken  
IncidentAllegation  
IncidentCitizenCharge  
IncidentCitizenInjury  
IncidentCommunicationTracking  
IncidentForceType  
IncidentOfficerInjury

### **Appendix 2 – Reports Absent DQA Plans (11-03-2017)**

298 Reporting A5  
298 Reporting A7  
298 Reporting B1  
298 Reporting B2

## ABOUT THE AUTHORS

**RS21** is a data science, analytics, and visualization service company providing custom analytics solutions to help organizations of all shapes and sizes solve complex human problems. Our diverse team brings together the best in today's technology with experts in data science, policy analysis, visual interaction design, and web-development to empower decision-makers with intuitive, data-driven insights which often challenge and outperform conventional wisdom.

Working with clients from large federal government institutions to local businesses and non-profits, RS21 has built a powerful ability to create new, powerful ways for organizations to understand and harness their data to inform better decisions and drive effective operational change.

**Josh Vertalka, Ph.D.** is the Chief Data Scientist at RS21. Josh has published several articles in top peer-reviewed scientific journals and taught college level courses. As the CDS, he leverages expert level geography, geostatistical, and urban planning knowledge, methods, and data management approaches. This allows him to create actionable conclusions to the most challenging urban planning, resiliency, and healthcare problems. Josh also has extensive experience with social media data for predictive analytics in the healthcare industry. Josh's responsibility is to ensure all data related matters meet the highest level of sophistication and scientific rigor.

**Brian Barnes, M.Arch., R.A., NCARB** is the RS21 Emerging Markets Program Manager with a specialty in policy analysis, planning, and design to address a wide range of operational challenges through strategic data analytics. With over ten years of professional experience solving complex problems with data analytics, Brian has published research through the NOW Institute and manages multi-disciplinary teams across a successful portfolio of RS21 analytics projects from assessing root causes of youth violence in urban communities to evaluating and improving supply chain resilience to building new analytical methods for healthcare providers to understand social determinants of health in a patient's every-day environmental conditions.