Data Quality Review of Crash Reports Accepted With Warning

Prepared for

MassDOT Registry of Motor Vehicles

Prepared by

University of Massachusetts
142 Marston Hall
Amherst, MA 01003
Tel 413.545.1035
UMassSafe@umass.edu

www.umasssafe.org

August 30, 2019

Table of Contents

List of Tables

List of Figures

BACKGROUND & PROJECT OVERVIEW

This project was developed as a result of findings from a 2014 MA Traffic Records assessment conducted by a National Highway Traffic Safety Administration (NHTSA) team. The team examined the core traffic records system in Massachusetts, including the MassDOT Registry of Motor Vehicles (RMV) Crash Data System (CDS), which scored poorly at 68.1%. Specifically, the data quality control programs scored 55.8%. The RMV and UMassSafe used these findings to examine specific crash report errors, track those errors, and provide feedback to police departments regarding those errors.

A review of crash reports submitted to the RMV for the period of 2012 to 2014 found that 17% of crash reports had been fully accepted, 72% had been accepted with warning (AWW), and 11% had been rejected and sent back to police departments for further information. The RMV was already in the process of developing methods for both tracking rejected crash reports and determining whether those reports were resubmitted. However, further review of the 72% of crash reports that were accepted with warning was needed.

The data submission warning process at the RMV exists to provide immediate feedback to the submitting law enforcement agency about the quality and completeness of their crash reports. The system identifies both empty fields and values that are invalid (do not fit the character of the field). This procedure is in lieu of fully rejecting the crash report, instead requesting that the police department revise and resubmit the report.

UMassSafe developed and implemented a process for reviewing crash reports that had been 'accepted with warning' by the RMV. Specifically, fields that were problematic for all crash reports accepted with warnings were studied, as well as those for individual police departments/barracks, police officers, and Record Management Systems (RMS). The RMV Law Enforcement Liaison (LEL) then used these findings to work with individual police departments to expand their understanding of data quality issues specific to their department in order to make improvements accordingly. Additionally the LEL identified that many issues occurring at the department level were a result of a RMS issue, and therefore technical assistance was additionally provided to their associated RMS vendors.

DATA SOURCES & CONSIDERATIONS

The RMV maintains a database of all issued warnings to the various law enforcement agencies. When any crash report is submitted electronically to the RMV, it is assessed for completeness and any invalid values. If applicable, a warning is issued to the department and this crash would be recorded as 'Accepted with Warning' (AWW). All departments have different submission processes, varying from weekly to on-request only. Departments often resubmit crashes many times due to date overlap and attempts to ensure that RMV has received the most recent/improved report. However, this practice may bias the sample for these analysis purposes, as warnings will be issued for each re-submission. Therefore, UMassSafe analyzed CDS data residing in the UMassSafe Highway Safety Data Warehouse to conduct most analyses for ease of normalizing and cleaning. It must be noted that while these analyses are based on the interpretation of the AWW criteria's intent, they may not align identically due to differing data management entities.

UMassSafe worked with the RMV to examine each existing and proposed AWW field. In the process, challenges with the AWW process were recognized for several fields described below.

Truck & Bus

Due to a flaw in the determination of a truck/bus reportable crash within CDS, those related fields were excluded from this analysis. It is recommended that key changes be made to the crash report form in order to more accurately determine which crashes are truly reportable as truck/bus. In the meantime, truck/bus warnings will still be issued if at least one of the truck/bus fields are completed, whereas previously any vehicle with a configuration of 4-13 would receive warnings for missing truck/bus fields, even when they are not actually relevant or required. For the purpose of conveying AWW program information and education to law enforcement and RMS providers, the truck/bus fields are indicated on data summary sheets but do not contain data values due to the conflicting criteria verse actual definition.

Unknown Driver Last Name

The RMV requires that all vehicle records contain an associated driver, including instances when the driver is unknown due to a hit/run, or alternatively when no driver was involved due to the vehicle being parked. This is problematic at the data entry level, as officers are often required by the RMS to enter a value into the Driver Last Name field. Upon further analysis, it was understood that the words 'UNKNOWN' or 'PARKED' had been used to fill this requirement, accounting for 10% of the crash-vehicle records during the period of Jan 2017-Aug 2018. In situations wherein departments used this method, they were then required to complete the corresponding occupant/person/driver related fields or would be issued a warning through the AWW process. These corresponding data fields would accurately be indicated as 'not applicable', with the 'unknown' option being accurate only in the event of a true hit/run. However, the 'not applicable' option does not exist for most fields.

Although the ultimate directive from RMV is that all fields should be completed, and to instruct departments to enter 'Unknown' (99) for any person fields that are not relevant, the existing practice at many large departments is to leave those related fields incomplete when not applicable. Due to this conflict, a short-term remedy was implemented by RMV/MassDOT IT to bypass the warning feature when the last name was 'UNKNOWN' or 'PARKED'. UMassSafe recommended a long-term solution of a parked vehicle flag and adherence to other MMUCC guidelines for the next revision of the crash report.

Speed Limit

During the ongoing analysis and technical assistance for the AWW project, an unusual pattern of AWWs over time appeared for the Speed Limit field. What originally started as one of the most improved fields in 2018 had risen to above 90% incomplete in February 2019, as seen in Table 1. Due to the sudden issue and severity of the problem, this field was removed for the remainder of the analysis, while the RMV examined why this error was occurring.

Table 1: Examination of Speed Limit Crash Fields by Month/Year

	Invalid/Incomplete	Valid	Grand Total	% Invalid/Incomplete
2018	**25545**	**100970**	**126515**	**20%**
1	4946	7378	12324	40%
2	3396	5356	8752	39%
3	3524	6402	9926	36%
4	2446	6630	9076	27%
5	2096	8447	10543	20%
6	1758	8938	10696	16%
7	1374	8820	10194	13%
8	1232	8977	10209	12%
9	1106	8750	9856	11%
10	1159	10318	11477	10%
11	1175	11136	12311	10%
12	1333	9818	11151	12%
2019	**54970**	**10664**	**65634**	**84%**
1	3047	7342	10389	29%
2	9018	765	9783	92%
3	9154	573	9727	94%
4	8690	500	9190	95%
5	9614	541	10155	95%
6	9413	566	9979	94%
7	6034	377	6411	94%

AWW FIELD SELECTION

To assess which crash data fields would have the most significant impact on crash data quality, and thus be included in the AWW system, UMassSafe analyzed known data quality issues, utilizing and updating findings from the 2014 Massachusetts Crash Data Audit along with the existing and planned revision of AWW fields from the RMV. Although some fields were initially selected for the AWW system, in an effort to be comprehensive and make the most progress on improving data quality, all crash fields were considered. As noted in the following table, fields were quantified by the frequency they were left empty or contained invalid entries (crash data was reflective of 1/1/2018-9/30/2018 as of 1/14/2019). Not all fields could be numerically assessed due to limitations of the data structure.

Table 2: Existing & Proposed AWW Fields for Consideration

Category	Field	Empty/ Invalid %	RMV Status	UMassSafe Recommendation
Crash	Crash Diagram	2%	Existing AWW Field	
	First Harmful Event Code	2%	RMV Planned Addition	
	First Harmful Event Location Code	2%	RMV Planned Addition	
	Light Conditions	1%	RMV Planned Addition	
	Road Contributing Circumstances	6%	RMV Planned Addition	
	Road Surface Condition Code	1%	RMV Planned Addition	
	Roadway Intersection Type Code	2%	RMV Planned Addition	
	School Bus Related Code	1%	RMV Planned Addition	
	Traffic Control Device	2%	RMV Planned Addition	
	Traffic way Description Code	2%	RMV Planned Addition	
	Weather Conditions Primary	2%	RMV Planned Addition	
	Work Zone Related Code	1%	RMV Planned Addition	
	Manner Collision	1%		Proposed Addition
	Crash Narrative	12%	RMV Planned Addition	
	Speed Limit	29%	RMV Planned Addition	
	Traffic Control Device Function	1%		Proposed Addition
Vehicle	Hit/Run		RMV Planned Addition	Remove Until Criteria is Y/N
	Moped		RMV Planned Addition	Remove Until Criteria is Y/N
	Vehicle Configuration Code	2%	Existing AWW Field	
	Event Sequence [Event Sequence 1]	2%	RMV Planned Addition	
	Towed	11%	RMV Planned Addition	
	Vehicle Registration #	1%	RMV Planned Addition	
	Damage Area	7%		Proposed Addition
	Emergency Use Code	21%		Proposed Addition
	Travel Direction	9%		Proposed Addition
Person/ Occupant	Occupant Seating Position	2%		Proposed Addition
	Injury Status Code	10%	Existing AWW Field	
	Ejection Code	11%	RMV Planned Addition	
	Transport Code	13%	RMV Planned Addition	
	Airbag Status	11%		Proposed Addition
	Occupant Protection System Use	13%		Proposed Addition
	Medical Facility			Proposed addition if injury status and transport code is affirmative

Category	Field	Empty/ Invalid %	RMV Status	UMassSafe Recommendation
Driver	Driver Cited	25%	RMV Planned Addition	
	Driver Contributing Code [Code 1]	12%	RMV Planned Addition	
	Driver Distracted By	22%	RMV Planned Addition	
	Driver's License #	11%	RMV Planned Addition	
	Driver's License Class Code [Class Code 1]	16%	RMV Planned Addition	
	Citation #		RMV Planned Addition	Remove – Driver Cited would need to be Y/N
	License Restrictions		RMV Planned Addition	Remove – need 'none' attribute
	Violation Code [Violation Code 1]		RMV Planned Addition	Remove – Driver Cited would need to be Y/N
NM	Last Name		RMV Planned Addition	
	NM Action	24%	RMV Planned Addition	
	NM Condition	25%	RMV Planned Addition	
	NM Location	25%	RMV Planned Addition	
	NM Type	21%	RMV Planned Addition	
	NM Safety Equp Code	67%		Proposed addition
T&B	Truck & Bus related information Mandatory		Existing AWW Field	Remove - revision dependent on crash report modification
	Carrier Name		RMV Planned Addition	
	Gross Vehicle Weight Rating		RMV Planned Addition	
	HazMat Placard		RMV Planned Addition	
	Interstate		RMV Planned Addition	
	Truck/Bus Body Type Code		RMV Planned Addition	
	US DOT		RMV Planned Addition	
	Commercial Driver's License Endorsements		RMV Planned Addition	Revise to be dependent on Configuration & Body Type
	Issuing State		RMV Planned Addition	Remove, phasing out
	MC/MX/Interstate Commerce Commission#		RMV Planned Addition	Remove, phasing out
	State Number		RMV Planned Addition	Remove, phasing out

As noted in Table 2, UMassSafe made recommendations for consideration by the RMV. A balance between comprehensiveness and ease of system implementation, as well as the scale of these changes affecting RMS & departments were considered. Ultimately Hit/Run, Moped, Citation #, Driver Cited, License Restrictions, and Violation Code were removed due to being either free-form or check box configuration, which would prevent a definitive invalid/incomplete designation. Additionally, NM Safety Equip Code, Damage Area, Emergency Use Code, Vehicle Travel Direction, Occupant Safety System, and Airbag Status were added to the Accepted with Warning implementation plan, with final list documented below in Table 3.

Table 3: AWW Fields for Project Implementation

Category	Field Name	Category	Field Name
Crash	Light Conditions	**Occupant**	Safety System
	Weather Conditions		Airbag Status
	Traffic Control Device Type		Ejection Code
	Road Surface		Injury Status
	Roadway Intersection Type		Transported by Code
	Trafficway Description	**Driver**	License Class
	School Bus Related		Driver Contributing Code
	Work Zone Related Code		Driver Distracted By
	First Harmful Event Location		License #
	First Harmful Event	**Non-Motorist**	NM Type
	Road Contributing Circumstances		NM Action
	Crash Narrative		NM Location
	Speed Limit		NM Condition
	Crash Diagram		NM Safety System
Vehicle	Vehicle Configuration		NM Injury Status
	Sequence of Events		NM Transported by Code
	Damaged Area Code	**Truck & Bus**	Carrier Name
	Towed from Scene?		Interstate
	Registration #		Cargo Body Type Code
	Responding to Emergency?		GVWR/GCWR
	Vehicle Travel Direction		HazMat Placard
			US DOT #

After examining existing data quality research and specific fields that led to crash reports being accepted with warning, the number of invalid/incomplete occurrences were ranked by field. Analyzing the data simply by volume identifies which fields would result in the immediate improvement, as shown in Table 4.

Table 4: Top 10 AWW Fields by % Invalid Incomplete, July 2018-June 2019

Crash Report Field	# Invalid/Incomplete	Sample Size	% Invalid/Incomplete
Vehicle: Responding to Emergency?	37897	238941	16%
Driver: Driver Distracted By	28810	216945	13%
Vehicle: Damaged Area Code	15902	238941	7%
Driver: License Class	15035	216945	7%
Vehicle: Vehicle Travel Direction	13336	238941	6%
Occupant: Safety System	11447	270925	4%
Driver: Driver Contributing Code	11037	216945	5%
Occupant: Transported by Code	9442	270933	3%
Occupant: Airbag Status	6253	270925	2%
Occupant: Ejection Code	4206	270925	2%

DATA DISSEMINATION

Local Police

An analysis by individual police department was conducted in order to determine the departments that would benefit most from RMV Law Enforcement Liaison (LEL) technical assistance.

Local police departments were ranked by overall invalid/incomplete percentage, accounting for all AWW fields, as shown in Table 5. Brookline, Somerville, Cambridge and Arlington had the highest overall invalid/incomplete AWW field rate. Invalid/incomplete AWW fields were also categorized by crash, vehicle occupant, driver, and non-motorist level to examine for trends.

Table 5: Ranking of Local Police Departments by % AWW Fields Invalid/Incomplete

Ranking	Police Department	Analyzed Report Count	Percent Invalid/Incomplete of AWW Fields by Category					
			Overall	Crash	Vehicle	Occupant	Driver	Non-Motorist
1	BROOKLINE	209	23%	20%	21%	25%	28%	43%
2	SOMERVILLE	314	19%	18%	18%	17%	28%	24%
3	CAMBRIDGE	831	18%	15%	17%	29%	12%	28%
4	ARLINGTON	256	18%	16%	16%	22%	17%	43%
5	EDGARTOWN	29	17%	13%	16%	10%	41%	22%
6	MEDFORD	393	17%	17%	16%	11%	28%	39%
7	BELMONT	206	16%	15%	15%	4%	36%	35%
8	LAWRENCE	986	15%	16%	15%	7%	25%	40%
9	LYNN	1143	15%	16%	15%	2%	30%	41%
10	W TISBURY	44	15%	12%	14%	5%	42%	-
11	SWAMPSCOTT	97	15%	15%	15%	4%	27%	21%
12	MALDEN	534	14%	14%	15%	7%	18%	37%
13	STONEHAM	133	14%	15%	15%	3%	26%	44%
14	HAVERHILL	991	14%	15%	15%	3%	25%	46%
15	DANVERS	282	13%	15%	14%	2%	27%	31%
16	NEWTON	854	13%	15%	15%	2%	19%	44%
17	WALTHAM	781	13%	15%	15%	3%	15%	29%
18	QUINCY	1083	12%	15%	15%	1%	18%	33%
19	N ATTLEBORO	431	12%	3%	3%	23%	23%	67%
20	CHELMSFORD	258	12%	7%	7%	20%	15%	7%
STATEWIDE		**62915**	**5%**	**5%**	**5%**	**2%**	**6%**	**24%**

The 20 local police departments with the greatest percentage of invalid/incomplete AWW fields included small, medium and large departments. Discussions with the RMV staff and key stakeholders indicated that departments of varying sizes had different crash reporting issues as seen below in Table 6. As a result, a separate ranking was developed for varying sized police departments. Small departments were defined as those submitting less than 100 crash reports, medium departments as those submitting 100-500 crash reports, and large departments as those submitting more than 500 crash reports electronically, in the 6-month period, 6/1/18-12/31/18.

Table 6: Invalid/Incomplete Field Analysis by Department Size and Crash Report Section

Department Size	Overall Field % Error	Crash % Error	Vehicle % Error	Occupant % Error	Driver % Error	Non-Motorist % Error
LARGE	6%	7%	7%	3%	8%	24%
MEDIUM	4%	4%	4%	3%	5%	25%
SMALL	3%	3%	4%	1%	4%	15%

Aiming to guide outreach efforts of the RMV Law Enforcement Liaison (LEL), all local law enforcement agencies with electronically submitted records were ranked by frequency of invalid/incomplete fields within their respective size groups, as outlined in Table 7. To achieve a balance between a variety of departments with high AWW occurrences and crash volume, the RMV LEL strategically met with departments from across all variances. The complete list of each department size ranking can be found in Appendix B.

Table 7: Top Ranking of Police Departments by % AWW Fields Invalid/Incomplete and Department Size

Small Police Departments	Count	% Invalid/ Incomplete	Medium Police Departments	Count	% Invalid/ Incomplete	Large Police Departments	Count	% Invalid/ Incomplete
EDGARTOWN	29	17.3%	BROOKLINE	209	23.0%	CAMBRIDGE	831	17.8%
W TISBURY	44	15.0%	SOMERVILLE	314	19.3%	LAWRENCE	986	15.1%
SWAMPSCOTT	97	14.5%	ARLINGTON	256	17.8%	LYNN	1143	15.1%
AQUINNAH	4	11.4%	MEDFORD	393	17.2%	MALDEN	534	14.2%
LEVERETT	8	10.3%	BELMONT	206	15.7%	HAVERHILL	991	13.7%
BERNARDSTON	2	6.7%	STONEHAM	133	14.2%	NEWTON	854	13.0%
MIDDLETON	94	5.7%	DANVERS	282	13.3%	WALTHAM	781	12.6%
HARDWICK	23	5.4%	N ATTLEBORO	431	12.0%	QUINCY	1083	12.3%
CARLISLE	30	5.2%	CHELMSFORD	258	11.7%	METHUEN	719	8.7%
HULL	62	4.4%	MEDWAY	114	9.7%	FRAMINGHAM	903	6.5%
WILLIAMSBURG	23	4.3%	EVERETT	117	9.2%	CHICOPEE	1031	5.3%
MANCHESTER	26	4.3%	BEVERLY	297	8.6%	ATTLEBORO	567	4.3%
WARWICK	4	4.2%	NORTON	209	8.0%	NEW BEDFORD	2119	4.2%
MERRIMAC	35	4.1%	WINCHESTER	114	7.9%	TAUNTON	709	3.7%
BOXFORD	31	4.0%	PLAINVILLE	146	7.4%	MILFORD	663	3.7%

Individual Police Department Analysis

Department specific data quality reports were developed for 62 local police departments – 10 small, 32 medium and 20 large. These reports included specific invalid/incomplete rate for each AWW field, as compared to the statewide average in order to identify areas and fields needing attention. Color shading of the fields highlight the best data quality in green, progressively worse to red, an example of which is shown in Figure 1 for one individual police department. See Appendix C for all printable department summary sheets.

MassDOT Registry of Motor Vehicles
Crash Report Invalid/Incomplete Field Analysis
Accepted With Warnings 'AWW' Project

Arlington Local Police

Category	Field Name	Field #	Criteria Notes	Invalid/Incomplete %	
				Statewide	Arlington
Crash	Light Conditions	1		0.6%	1.6%
	Weather Conditions	2		0.7%	2.0%
	Traffic Control Device Type	4		0.8%	3.5%
	Road Surface	6		0.7%	4.7%
	Roadway Intersection Type	7		0.8%	4.7%
	Trafficway Description	8		1.0%	5.9%
	School Bus Related	9		0.8%	8.2%
	Work Zone Related Code	10		0.8%	7.0%
	First Harmful Event Location	12		0.8%	1.6%
	First Harmful Event	13		0.7%	6.3%
	Road Contributing Circumstances	14		4.2%	13.3%
	Crash Narrative			1.5%	0.8%
	Speed Limit			16.6%	30.5%
	Crash Diagram			0.7%	0.0%
Vehicle	Vehicle Configuration	21		0.9%	0.9%
	Sequence of Events	23		1.1%	1.8%
	Damaged Area Code	27		7.7%	2.9%
	Towed from Scene?	33		1.5%	0.2%
	Registration #			0.4%	0.2%
	Responding to Emergency?			19.4%	99.8%
	Vehicle Travel Direction			6.6%	9.4%
Occupant	Safety System	35	excluding 'Unknown'/'Parked'	2.5%	7.1%
	Airbag Status	36		2.1%	8.0%
	Ejection Code	37		2.0%	8.9%
	Injury Status	39		0.9%	11.9%
	Transported by Code	40		4.9%	72.1%
Driver	License Class	19	excluding 'Unknown'/'Parked'	7.3%	16.8%
	Driver Contributing Code	25		4.5%	17.3%
	Driver Distracted By	26		10.4%	31.7%
	License #			1.3%	2.1%
Non-Motorist	NM Type	15		18.2%	26.9%
	NM Action	16		20.2%	30.8%
	NM Location	17		21.1%	26.9%
	NM Condition	18		21.6%	26.9%
	NM Safety System	35		59.8%	100.0%
	NM Injury Status	39		8.0%	34.6%
	NM Transported by Code	40		14.1%	53.8%
Truck & Bus	Carrier Name		Warning only issued/applicable if one of the other T&B fields are completed.		
	Interstate	43			
	Cargo Body Type Code	44			
	GVWR/GCWR	45			
	HazMat Placard	47			
	US DOT #				

Figure 1: Sample Department Specific AWW Field Data Quality Summary

Officer specific analysis was conducted for 15 law enforcement agencies on an as-needed basis for the LEL efforts. It was discovered that towns which ranked with a high invalid/incomplete rate may not always benefit from officer specific analysis due to universal RMS-level errors (i.e. Responding to Emergency does not exist for QED submitted crashes), resulting in non-significant results at the officer level. Alternatively, the departments were chosen on a case-by-case basis. A sample is shown in Figure 2. The fields included in each analysis vary based on those most problematic relative to each department. See Appendix D for all officer specific data summary sheets.

Cambridge Police Crash Data Field Analysis | Data: June 1 - Dec 31, 2018 | MassDOT Registy of Motor Vehicles Crash Reports Accepted With Warning

		% Invalid/Incomplete Entry by Field														
		Crash	Vehicle	Occupant				Driver			Non-Motorist				
Name	# Analyzed Crash Reports	Speed Limit	Vehicle Travel Direction	Safety System	Airbag Status	Ejection Code	Transported by Code	License Class	Driver Contributing Code	Driver Distracted By	Type	Action	Location	Condition	Transported by Code
TOTAL Cambridge Police	**832**	6%	13%	24%	24%	25%	35%	18%	15%	15%	15%	19%	18%	20%	82%
JEAN-BAPTISTE, JERRY	87		26%	92%	92%	92%	96%	9%	13%	13%	4%	17%	13%	4%	100%
CALLINAN, JASON	72		18%	6%	6%	6%	19%	18%	17%	17%	13%	13%	13%	13%	83%
DONAHUE, MARK	28		2%	80%	87%	93%	7%	20%	13%	13%	100%	100%	100%	100%	100%
MICELI, MELISSA	24		3%	82%	82%	82%	18%	25%	19%	19%	71%	71%	71%	71%	71%
CIRIELLO, ROBERT	23		8%	15%	15%	19%	26%	33%	33%	33%	33%	33%	33%	33%	67%
VALENTIN, SIMON	22	95%		14%	14%	14%	5%	10%	10%	10%	50%	50%	50%	50%	50%
ROSA, DAVID	22		33%	28%	44%	22%	11%	11%	11%	17%					50%
EDWARDS, GARY	20			16%	16%	16%		64%	59%	59%					
GRASSI, JOSEPH	18			21%	21%	21%	16%	19%	13%	13%	33%	33%	33%	33%	100%
JOSEPH, DONYELL	15		4%				15%	24%	24%	19%					50%
ALLEN, STEVEN	15	100%		4%	4%	4%	92%	4%	4%	4%					100%
BARTLETT, DANIEL	14		16%	5%	5%	5%	73%	11%	11%	11%					100%
LOWE, SEAN	13		20%	6%	13%	13%	75%	20%						100%	
BROWN, ZACHARY	13		4%	25%	25%	25%	13%	21%	21%	21%	67%	67%	67%	67%	33%
CLAVETTE, MARK	12		9%				87%	23%	15%	15%					100%
CHERUBINO, MICHAEL	12		11%				14%	17%							67%
BROWN, RICHARD	11		29%	11%	11%	11%	56%	38%	13%	13%	100%	100%	100%	100%	100%
COSTA, EDMUND	11		5%	37%	37%	37%		9%	9%	9%	50%	50%	50%	50%	
DIGGINS, JAMES	10		6%	17%	17%	17%	17%	36%	27%	27%					100%
SMITH, MARK	10			21%	21%	21%		25%	17%	17%					100%
BUILES, LUIS	10		14%	27%	27%	27%		18%	9%	9%					100%
CAZEAU, ANDY	9			13%	6%	6%		50%	50%	50%	67%	67%	67%	67%	67%
MORRISSEY, MICAHEL	9		7%	100%	100%	100%	20%	40%	40%	40%					
CROWLEY, JOHN	9	11%					29%			14%				50%	100%
FOSTER, EDDIE	9		6%	15%	15%	15%	31%	10%	10%	10%					100%
VIEIRA, LEE	9	100%	14%	100%	100%	100%	11%								
AYOUB, NICHOLAS	9	11%					22%								67%
ALI, ASIF	9		28%	14%	14%	14%	36%			10%					100%
GALUSKI, KYLE	9		25%					14%	14%	14%					100%
AMES, CHRISTOPHER	8		25%	29%	29%	29%	50%	9%							
HUDSON, LAWRENCE	8		40%	14%	14%	14%	29%	14%	14%	14%					100%
PADGETT, IVELISE	8		25%				46%	11%			50%	50%	50%	50%	100%
ANTONOPOULOS, MILTIADES	8						13%	8%	8%	8%					100%
CROWLEY, JOSEPH	8		15%				33%	44%	44%	44%					
BUXBAUM, JOSHUA	8						15%	17%	17%	17%					
O'REGAN, BRIAN	8			18%	18%	18%	91%	11%	11%	11%					100%
MCMAHON, DEVIN	8		21%	25%	25%	25%	50%								

Figure 2: Sample Officer Specific Analysis of AWW Fields of Interest

State Police

An analysis of State Police crash records was conducted in order to determine any variances when compared to local police departments as shown below in Table 8. Identifying overall trends where State Police differ can help guide LEL efforts more effectively.

Table 8: Invalid/Incomplete % of AWW Fields by Police Type

Category	Field Name	Empty/Invalid %	
		All Local Police	All State Police
Crash	Light Conditions	0.6%	0.1%
	Weather Conditions	0.7%	3.7%
	Traffic Control Device Type	0.8%	0.9%
	Road Surface	0.7%	0.1%
	Roadway Intersection Type	0.8%	1.2%
	Trafficway Description	1.0%	0.7%
	School Bus Related	0.8%	0.3%
	Work Zone Related Code	0.8%	0.4%
	First Harmful Event Location	0.8%	1.2%
	First Harmful Event	0.7%	1.1%
	Road Contributing Circumstances	4.2%	2.4%
	Crash Narrative	1.5%	0.4%
	Speed Limit	16.6%	2.2%
	Crash Diagram	0.7%	0.0%
Vehicle	Vehicle Configuration	0.9%	0.7%
	Sequence of Events	1.1%	1.1%
	Damaged Area Code	7.7%	1.9%
	Towed from Scene?	1.5%	0.2%
	Registration #	0.4%	0.4%
	Responding to Emergency?	19.4%	1.4%
	Vehicle Travel Direction	6.6%	2.0%
Occupant	Safety System	2.5%	13.5%
	Airbag Status	2.1%	4.9%
	Ejection Code	2.0%	1.6%
	Injury Status	0.9%	1.5%
	Transported by Code	4.9%	4.2%
Driver	License Class	7.3%	8.8%
	Driver Contributing Code	4.5%	6.3%
	Driver Distracted By	10.4%	27.0%
	License #	1.3%	1.3%
Non-Motorist	NM Type	18.2%	38.4%
	NM Action	20.2%	46.0%
	NM Location	21.1%	49.8%
	NM Condition	21.6%	52.3%
	NM Safety System	59.8%	84.8%
	NM Injury Status	8.0%	7.2%
	NM Transported by Code	14.1%	8.4%

State Police barracks were then ranked by percent of invalid/incomplete AWW fields, as shown in Table 9. Barracks H1, B2, and A6 had the highest overall invalid/incomplete AWW field rate, each of significant size difference. Invalid/incomplete AWW fields were also categorized by crash, vehicle occupant, driver, and non-motorist level to examine for unique problems at the barrack level.

Table 9: Ranking of State Police Barracks by Rate of Overall Invalid/Incomplete AWW Fields and Category

Rank	MSP Entity	Total Crashes	Total % Error	Crash Fields %	Vehicle Fields %	Occupant Fields %	Driver Fields %	Non-Motorist Fields %
Rank	**ALL**	**15067**	**3.7%**	**1%**	**1%**	**5%**	**11%**	**41%**
1	H1	12	6.1%	6%	1%	8%	13%	
2	B2	122	5.3%	2%	1%	10%	14%	25%
3	A6	859	5.2%	0%	1%	8%	12%	71%
4	C9	22	5.0%	1%	3%	9%	10%	
5	A4	825	4.9%	1%	1%	7%	16%	23%
6	H5	349	4.8%	1%	1%	10%	10%	23%
7	A2	280	4.7%	2%	2%	7%	12%	0%
8	C5	176	4.6%	1%	1%	8%	13%	64%
9	D7	315	4.3%	3%	1%	5%	12%	71%
10	B4	73	4.2%	2%	1%	6%	9%	86%
11	B6	214	4.2%	1%	2%	8%	9%	21%
12	H4	362	4.2%	1%	1%	5%	14%	26%
13	D4	486	4.2%	1%	1%	7%	11%	71%
14	B5	77	4.1%	1%	2%	8%	9%	39%
15	A5	338	4.1%	1%	1%	5%	13%	29%
16	D2	186	3.9%	2%	1%	5%	10%	
17	C4	454	3.9%	2%	1%	6%	10%	71%
18	H6	859	3.8%	1%	1%	4%	12%	20%
19	D3	685	3.7%	1%	1%	4%	13%	43%
20	C7	26	3.7%	4%	1%	5%	8%	
21	H3	786	3.5%	1%	1%	3%	13%	69%
22	C3	37	3.4%	2%	2%	5%	8%	
23	C2	341	3.3%	1%	1%	6%	9%	
24	D1	427	3.2%	1%	1%	5%	8%	21%
25	H7	684	3.2%	0%	1%	4%	10%	32%
26	A1	1012	3.2%	1%	1%	5%	9%	49%
27	B1	43	3.0%	1%	1%	7%	5%	
28	H2	480	2.7%	1%	1%	3%	8%	43%
29	C6	521	2.7%	0%	1%	4%	8%	71%
30	B3	475	2.7%	1%	0%	3%	9%	51%
31	C1	146	2.1%	1%	1%	3%	7%	14%
32	A3	746	2.0%	1%	0%	2%	7%	14%
	CAR	24	2.7%	1%	0%	1%	12%	
	GH	2615	3.9%	2%	1%	5%	11%	31%
	D5	4	5.9%	5%	0%	0%	17%	14%
	D6	2	1.2%	0%	0%	0%	8%	
	MV2	3	5.7%	7%	3%	8%	5%	
	MV5	1	10.0%	0%	0%	40%	25%	

Thirty-eight barrack specific data quality reports were developed for dissemination. These reports included their specific invalid/incomplete rate for each AWW field, as compared to the entirety of the State Police. See Appendix E for all printable barrack summary sheets.

Records Management Systems

Invalid/incomplete data entries may result from varying factors, such as RMSs without current RMV specifications, or technological upload/processing errors. In this specific analysis, RMS crash reporting trends were examined in order to better understand possible global RMS issues, as well as those for specific RMSs which could be remedied at the software capture, export or import level.

The analysis began with an examination of how many crash reports were submitted through each RMS. As shown in Figure 3, IMC was utilized for close to 60% of all crash records electronically submitted in the period analyzed, followed by RAMS with almost 19%, and QED with 11.5%. There were also several smaller RMSs each submitting less than 5% of crash reports. These include Pamet (4.9%), Nexgen (2.3%), Microsystems (2.1%), Keystone (1.2%), MAACs (0.5%), and Larimore (0.3%). However, considering the number of invalid/incomplete fields submitted by vendor, IMC had the lowest relative proportion while Larimore and QED had the most problematic ratio.

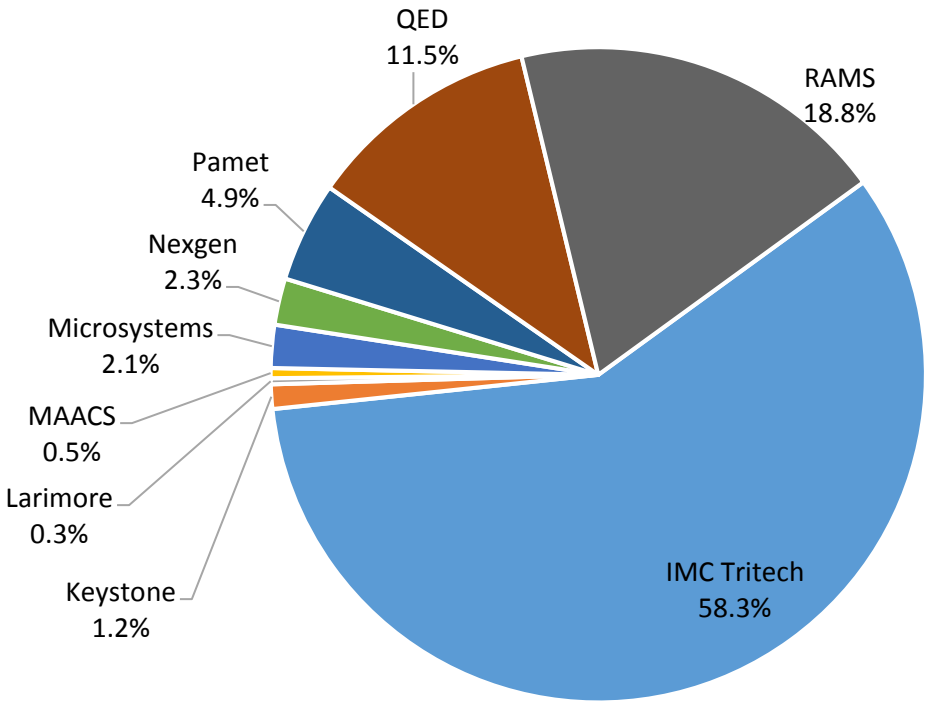

Figure 3: Crash Reports by RMS Submission, July 2018 – June 2019

Examining the total number of invalid/incomplete field entries for each RMS by crash report section revealed that crash-level fields were not a significant problem, and alternatively, invalid/incomplete entries existed more often in the vehicle, occupant, driver and non-motorist sections, as shown in Table 10.

Table 100: Percent Invalid/Incomplete by AWW Field Category and RMS, Jan 2019-May 2019

RMS	Crash	Vehicle	Occupant	Driver	Non-Motorist
IMC	0%	4%	0%	2%	7%
QED	2%	15%	4%	23%	41%
Larimore	2%	21%	24%	26%	44%
RAMS	1%	1%	5%	11%	46%
Pamet	2%	2%	9%	9%	75%
Keystone	0%	9%	1%	2%	37%
Microsystems	1%	6%	4%	7%	15%
MAACS	6%	2%	3%	10%	7%
Nexgen	1%	2%	6%	4%	31%
State Total	**0.8%**	**4.5%**	**2.5%**	**6.5%**	**28.7%**

A detailed RMS Analysis report can be found in Appendix F, along with a full matrix, which examined every field and RMS combination. Among the most notable findings, crash reports submitted via Larimore were discovered to have 100% incomplete for the Vehicle Damaged Area Code, while crash reports submitted via QED had almost 100% incomplete for the Responding to Emergency field.

PROGRESS TRACKING

As described earlier, the entirety of the above analysis used data from the Massachusetts CDS housed in the UMassSafe Highway Safety Data Warehouse. In contrast, initial analysis attempts utilized log records of the AWW system retrieved from the RMV. A notable difference between the two data sources was the possibility that a crash report may be submitted to the RMV multiple times and be over-represented if utilizing warning records. Due to this caveat, the AWW data was unable to be effectively normalized, while the CDS crash records proportionally identified and created trends and guidelines for strategic implementation statewide. With that in mind, progress tracking was conducted using CDS data, therefore counting the AWWs for each crash only once.

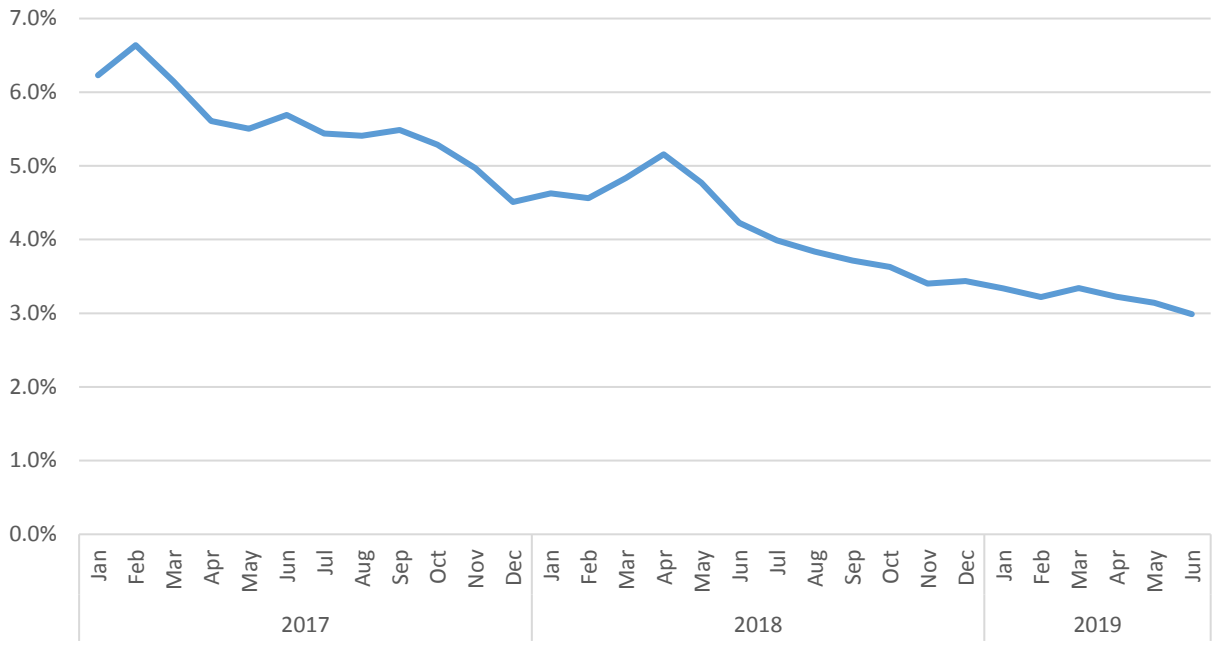

Figure 4: Invalid/Incomplete of All AWW Fields over Time

Considering all crash report fields addressed in the AWW data quality initiative, the frequency of invalid/incomplete fields reduced significantly over time, from above 6% in early 2017 to 3% in June 2019, as shown in Figure 4 below. Alternatively, when considering the rate of crash reports that had any invalid/incomplete AWW field over time, as seen in Figure 5, there was a reduction from 57% to 46%. In this view, if a crash report had one invalid/incomplete field, it was just as significant as if it had ten.

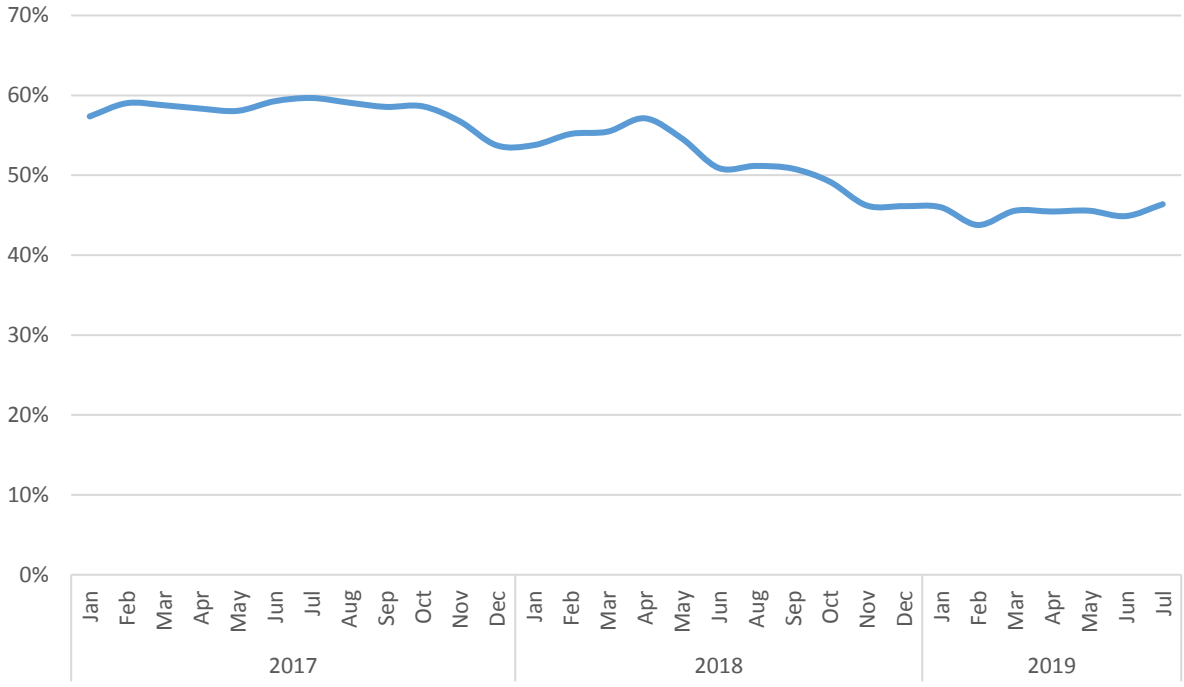

Figure 5: Percent of Crash Reports with Any Invalid/Incomplete AWW Field over Time

PERFORMANCE MEASURES & CONSIDERATIONS

Goal: To improve the accuracy and completeness of the Registry of Motor Vehicles Crash Data System by reducing the percentage of crash reports accepted with warning from 14% to 12% in year one, and to 10% by 6/30/19.

Result: Reports Accepted with Warning in June 2019 = 882

Reports Received in June 2019 = 10,056

Percent Accepted with Warning in June 2019 = 8.77%

RMV education and outreach efforts to police agencies and vendors have been ongoing with great feedback. Anticipation of continued success on performance measures is expected with a greater lapse in time for the efforts to actualize. However, as long as any fields remain 100% incomplete by department or RMS, gains will be hindered.

Also worth discussion is the applicability of the goal due to the variability in the specifications and number of active warnings in the AWW system. Upon initiation of the project there were four active warnings - vehicle configuration code, injury status code, crash diagram, and truck/bus, which then underwent numerous field additions and criteria changes, currently with over 30 fields, therefore creating a challenge in accurately documenting the even greater successes achieved. A secondary consideration is the lack of normalization when crash reports are submitted more than once, altering the typical distribution.

Additionally, success is even more notable when examining the same criteria and fields over time, as illustrated previously in Figures 4 & 5, wherein the frequency of invalid/incomplete AWW fields reduced from above 6% in January 2017 to 3.0% in June 2019. While using the same method of analysis of crash reports that have any invalid/incomplete AWW fields over time, there is a slight reduction from 57% to 46%.