Microsoft Access and Excel for Data-Driven Crime Analysis: A 5-Part Series

Part 3: Building Master Tables and Ensuring Data Quality

Welcome to **Part 3** of our 5-part analytical series. Today, we dive into creating master tables and ensuring the quality of your data. This session continues the partnership between IADLEST and TxDOT, which has been supporting Texas law enforcement for nearly a decade in developing comprehensive data strategies to reduce traffic fatalities and crime. Building master tables will help consolidate and organize large amounts of data, while ensuring consistency and accuracy through validation techniques will empower agencies to make informed decisions. The work we do here, through the lens of the TxDOT and IADLEST partnership, will strengthen your agency's ability to clean, manage, and maintain high-quality datasets for analysis. By improving data quality, you'll be able to trust your data more fully and make more impactful decisions based on reliable and accurate information.

Dawn Reeby
Senior Analytical Specialist

Welcome to **Part 3: Mastering Tables and Data Quality Cleaning Tips**

I'm Dawn Reeby, and I'm honored to guide you through this transformative training. With over 25 years in law enforcement analysis, I've spent nearly 14 of those years collaborating with IADLEST to deliver impactful analytical trainings, webinars, and technical assistance nationwide. Our work focuses on helping agencies implement smarter, data-driven strategies while fostering high-performing teams. As the author of *Bigger Than Data* and the *Building a Crime Analysis Legacy* books, my mission is to empower crime analysts and supervisors to strengthen their careers, build team capacity, and leave lasting legacies in the field.

In this session, we'll dive into tools and techniques using Microsoft Access and Excel to help you manage and analyze data effectively, creating strategies that drive results. I'm here to share my experience and insights with you, and I'm excited to help you take your analytical skills to the next level!

Let's get started!

The SAFER (Strategic Analysis for Focused Engagement with Results: Crime and Crash Reduction) model, developed by the International Association of Directors of Law Enforcement Standards and Training (IADLEST) and supported by the Texas Department of Transportation (TxDOT), integrates location-based crime and crash data to create long-term strategies for reducing social harms. It builds on the earlier DDACTS 2.0 model, initially supported by the National Highway Traffic Safety Administration (NHTSA). IADLEST offers a variety of resources, including workshops, literature, webinars, and training, to help law enforcement agencies implement data-driven models like SAFER for more effective crime and traffic safety analysis and deployment.

Part 2: Helpful Queries and Reports in Microsoft Access ®
RECAP

We've now completed **Part 2** of our 5-part series, where we focused on creating effective queries and reports in Microsoft Access® to streamline your analysis process and automate repetitive tasks. You learned how to pull relevant data, build clear reports, and improve workflow efficiency, even for agencies without direct data connections.

Remember, while this is part of an ongoing series, each session is designed to stand alone, so you can always refer back to individual segments based on your needs.

As we continue to build on these foundational skills, **Part 3** will guide you through the process of creating master tables and applying data quality control techniques. These steps will ensure your data is organized, accurate, and ready for deeper analysis.

Let's jump into **Part 3**, where we'll focus on organizing and maintaining high-quality datasets to make your analysis even more powerful!

Learning Objectives

Objective 1: Create and manage master tables.

Objective 2: Apply data cleaning techniques for quality control.

**Objective 3**: Evaluate accuracy and consistency by implementing validation processes.

In this session, we're going to dive into the crucial role that master tables play in crime analysis. Master tables are key to consolidating data from various sources, ensuring quality control, and streamlining the analysis process. By the end of this module, you'll understand how to effectively create and manage these tables, how to apply data cleaning techniques to ensure accuracy, and how to use validation processes to maintain consistency in your data.

The learning objectives for this session include:

1. Create and manage master tables to organize and consolidate data.
2. Apply data cleaning techniques to ensure data quality and eliminate errors.
3. Evaluate accuracy and consistency by implementing validation processes.

Let's get started!
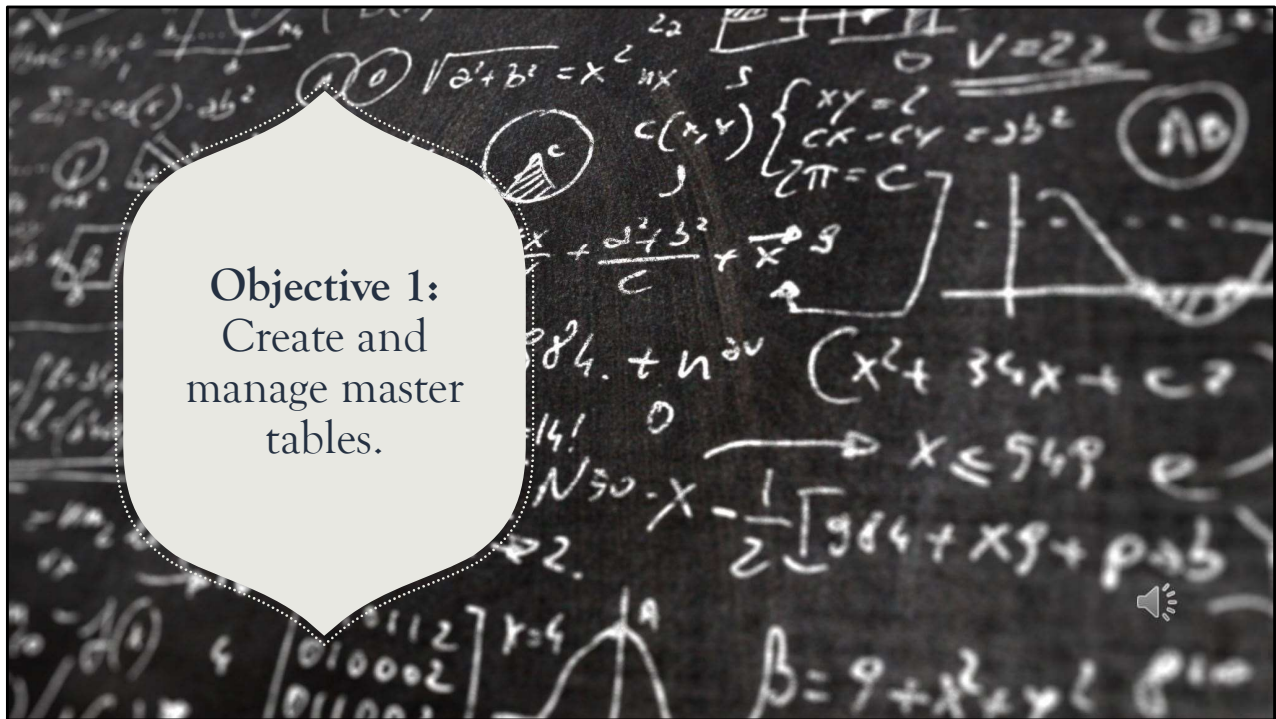
Objective 1: Create and manage master tables to organize and consolidate data.

Benefits of Master Static Tables

- Can store fields from multiple data sources all in one place
- Can have manually added data (MO, Story, Notes)
- Can be used in various queries
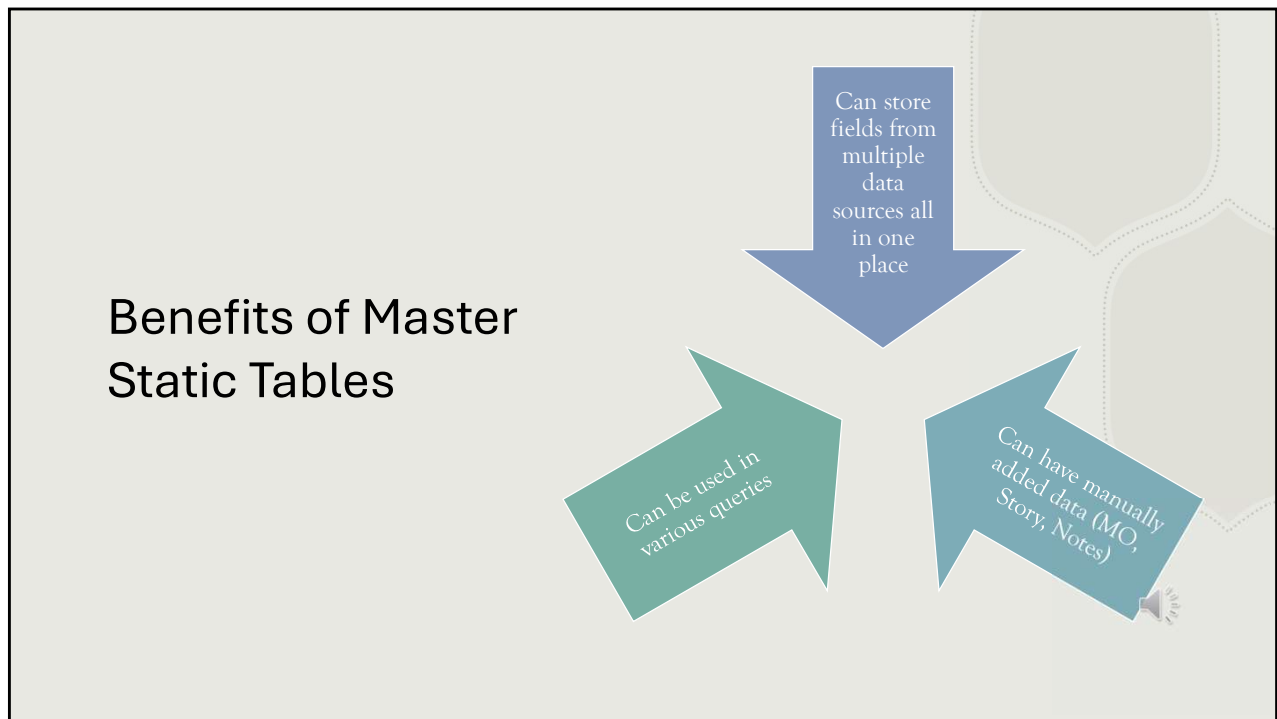
There are numerous benefits of having master static tables:

-Can store fields from multiple data sources all in one place
-Can manually add data (MO, Story, Notes)
-Can be used in various queries

With "Static tables" – data is no longer "live" on the network. Within a static table, data can be cleaned, updated, supplemented, added to and queried in ways that cannot be done to the "live" tables the data was pulled from. The analysis will now be done off of the network and will not bog down the CAD or RMS network servers and systems. The static master tables essentially create a "mirror" database to what exists within the RMS and CAD systems.

Benefits of Master Static Tables

Can store fields from multiple data sources all in none place

Here's an example of how queries can store fields from multiple data sources all in one place.

In this case, once we linked our tables from our RMS system into Microsoft Access®. We found that 3 tables contained the entirety of the crash data that we wanted: one table had the crash location and time, another table had injury, and the third table had road conditions. To make querying easy and accurate, we wanted to combine all these tables into one table. The output of the "Make Table" query gave us all the data in one accessible and editable location.

Benefits of Master Static Tables

Can store fields from multiple data sources all in none place

CallNum	AccidentDat	AccidentTim	ReportDate	OUIinvolved	Intersection	LightCond	WeatherCor
18-278	20180105	0849	20180105	N		1	1
18-2529	20180207	1151	20180207		Y	1	4
18-2520	20180207	0910	20180207			1	2
18-2493	20180206	1919	20180206			5	2
18-2403	20180205	1612	20180205	N		1	1
18-2400	20180205	1516	20180205		Y	1	1
18-2355	20180205	0720	20180205	N	N	1	1
18-2354	20180205	0716	20180205	N		1	1
18-2358	20180205	0745	20180205	N	N	1	1
18-2349	20180205	0636	20180205		Y	1	1
18-2351	20180205	0704	20180205	N	N	1	1
18-276	20180105	0745	20180105			1	1
18-2345	20180204	2318	20180205	N	N	4	3
18-2319	20180204	1053	20180204			1	2
18-2239	20180202	2226	20180203	N	Y	4	1
18-2220	20180202	1623	20180202			1	1

Here is the data view of our query. The results can bring data from multiple sources into one complete file for analysis.

Benefits of Master Static Tables

Can add crime analysis notes

Story	Crime Bulletin	Narcotics Bulletin	Gang Bulletin	Reviewed	Bu
Dawn Reeby overdosed on heroin; 3rd time in 2 months, m	☑	☑	☑	☑	
John Goffstown (11/11/75) reportedly put a large metal pie	☐	☐	☐	☑	
Evan Collins (3/9/97) reported that that Michael Manley (7/	☐	☐	☐	☑	
MV stop; Operator, Bethany Grassi (6/30/88); passenger Ta	☐	☑	☐	☑	
HEROIN OVERDOSE: JOSEPH SAMPSON (3/18/87) overdosec	☑	☑	☐	☑	
Jeffrey Jones (6-9-64) passed away; lung cancer – home hos	☐	☐	☐	☑	
Arena Sports Bar (53 High St.) – employee Adam Marcus (11	☐	☐	☐	☑	

Here's an example of how a table can play a role when we want to add additional data to our table such as MO, Story, Notes. Let's say we want automated crime bulletins every Friday or narcotics reports every morning. By adding a check box to the master table, we are later able to create automated reporting based off the new check box. One can add analyst notes.

Benefits of Master Static Tables

Can be used in various queries

IncidentType	2018	2019	2020	2021	2022	2023	Avg	Normal Range	Change
Rape	2	1	2	1	2	1	1.6	1-2	-38%
Indecent Assault	6	1	1	4	5	1	3.4	1-5	-71%
Peeping & Spying	0	0	1	0	0	2	0.2	0-1	900%
Aggravated Assault	11	14	6	5	8	11	8.8	5-12	25%
Simple Assault	44	41	40	33	33	49	38.2	34-43	28%
Robbery	5	0	3	2	3	2	2.6	1-4	-23%
Threats	32	23	26	12	16	31	21.8	15-29	42%
Violation of RO	15	14	9	14	9	4	12.2	10-15	-67%
Housebreak	11	12	16	16	18	18	14.6	12-17	23%
Commercial Break	15	11	12	14	12	8	12.8	11-14	-38%
Larceny from MV	61	73	67	62	43	90	61.2	51-71	47%
Larceny from Building	28	24	19	23	21	35	23	20-26	52%
Larceny from Person	11	11	12	4	2	3	8	4-12	-63%
Larceny from Residence	15	13	8	8	24	9	13.6	8-19	-34%
Larceny of Bicycle	15	9	13	9	9	17	11	8-14	55%
Larceny of Services	4	9	7	2	7	2	5.8	3-8	-66%
Shoplifting	51	58	55	77	63	72	60.8	52-70	18%

Creating static master tables will also facilitate the execution of other queries. For instance, since we needed six tables to set up our master crimes data, the query results would have been way too complicated to run summary stats by year by crime type. By developing a master crimes table through a 'Make Table' query, we simplified the results and were able to use them in this 'click of a button' query.

Additionally, it is crucial for new analysts to have or develop a strong foundational understanding of the data being entered. This includes being able to distinguish between similar yet distinct crime categories, such as burglary and theft, as these differences significantly impact the accuracy and relevance of their analysis. For instance, burglary typically involves unlawful entry into a structure to commit a crime, whereas theft refers to the unlawful taking of someone's property without necessarily involving entry. Misclassifying these crimes could lead to inaccurate reporting, skewed crime trends, and ultimately ineffective resource allocation. Analysts must also familiarize themselves with the legal definitions, agency-specific coding practices, and contextual nuances of these offenses to ensure that their findings accurately reflect the nature of the crimes being examined. Such understanding is essential for producing reliable analyses that can guide decision-making and strategic planning.

Linking vs. Importing

Link and Append Live Tables

Import and Append Static files
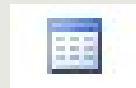
Check-In Question #1

Which of the following is a benefit of using master static tables in crime analysis?

a) They only store live data from network servers

b) They cannot be used in queries

c) They consolidate data from multiple sources in one location

d) They are difficult to update and maintain

Check-In Question #1

Which of the following is a benefit of using master static tables in crime analysis?

a) They only store live data from network servers

b) They cannot be used in queries

c) They consolidate data from multiple sources in one location

d) They are difficult to update and maintain

Build These Master Static Tables

Master Crash Table

Provides an unduplicated count of crashes with details

Master Crimes Table
Provides an unduplicated count of crimes with details and provides
FI and events data.

Master CAD Table

Provides an unduplicated count of calls for service

Master Arrest Table

Provides all arrests reports and bases the IBR Description on hierarchy (duplicated if more than one person was arrested during one incident)

Master Activity Table

Provides all police activity

Master Drug Arrest Table

Provides all Drug Arrests regardless of other crime committed

Creating Master Tables

| LIVE CONNECTION | IMPORT STATIC DATA | MANUAL ENTRY |

As we discuss the methods for managing data—whether through ODBC connections, importing static data, or even manual entry—the ultimate goal remains the same: to create master tables that support flexible, dynamic querying.

Master tables act as the foundation for your analysis. They consolidate and clean data into a reliable format, ensuring that no matter the source, you have a consistent structure to work from.

This flexibility allows you to quickly filter, segment, and analyze data, adapting to the unique demands of each case or request. The method you choose depends on your agency's resources, but always keep the focus on building a reliable framework for deeper, more actionable insights."

Method 1: ODBC using 'Make Table' and 'Append' Queries

Make Table Query

Append Query

GO TO YOUR VIDEO LIBRARY: "MAKE TABLE AND APPEND QUERIES"

An append query adds new records to the existing table so that the table is always up to date. Analysts will typically run an append query once a day or more to ensure that any "master table" includes the most up to date records extracted from CAD or RMS. For example, as officers and others enter new reports into the CAD or RMS, the analyst can extract those records and append them to a master table.

For more information on the topics discussed, visit:

1. "Make Table Query". https://support.office.com/en-us/article/Create-a-make-table-query-96424F9E-82FD-411E-ACA4-E21AD0A94F1B
2. "Append Query". https://support.office.com/en-us/article/Add-records-to-a-table-by-using-an-append-query-98a5bd66-2190-4243-9638-8ef649cf3596?ui=en-US&rs=en-US&ad=US
3. "Microsoft Access 2016 Append Query: Append Query". Retrieved from https://www.youtube.com/watch?v=nx_-z-uuw5Y

You use a "make table" query when you want to retrieve data from one or more tables, and then load the resulting dataset into a new table.
You use an "append" query when you need to add new records to an existing table by using data from other sources. *Adds current data to tables.*

Pause the video here and check out this 8-minute video on how create "Make Table" and "Append" queries.

Method 2: Appending Static Files

Appending Static Files

Appending Static Files to Create a Searchable Master Crime Analysis Database

"When creating a master crime analysis database, appending static files is a common and practical method to consolidate data into a centralized, searchable format. This process involves importing datasets—such as incident reports, CAD logs, or historical records—from external sources and combining them into a single master table.

The process typically includes the following steps:
1.**Standardizing Data:** Ensure that each static file follows the same structure, such as uniform column names and formats.
2.**Data Cleaning:** Remove duplicates, correct errors, and align inconsistent values to maintain data quality.
3.**Appending Files:** Combine the datasets, carefully aligning data fields to maintain integrity and usability.
4.**Validation:** Run checks to confirm the data is complete and consistent after appending.

The benefits are significant:
•**Centralized Access:** A master table allows you to query all your data in one place, saving time and effort.

•**Enhanced Analysis:** With all datasets consolidated, patterns and trends are easier to identify.
•**Consistency:** Standardized data ensures reliable results across all analyses.

By appending static files, you create a solid foundation for conducting detailed and efficient crime analysis, enabling you to turn raw data into actionable insights to better support your agency's goals."

Pause the video here and check out these videos on importing Excel into Access and on how append static data queries.

Method 3: Manual Entry

Manual Entry

GO TO YOUR VIDEO LIBRARY: "MANUAL ENTRY"

Why a Small-Agency Analyst Might Use Manual Entry

In smaller agencies, where resources or technological infrastructure might be limited, manual data entry can play a crucial role in maintaining a robust crime analysis database. Analysts in these settings often face unique challenges, such as incomplete data capture by their Records Management System (RMS) or the absence of automated tools for integrating datasets.

Manual entry allows these analysts to fill the gaps by adding critical information that might otherwise be overlooked. This could include:
•**Contextual Details:** Notes from officers, witness statements, or case-specific nuances not captured in structured fields.
•**Non-RMS Data:** Data from community complaints, external databases, or intelligence briefings.
•**Specialized Tracking:** Custom fields for tracking local trends, chronic offenders, or non-criminal activity patterns.

Even if your agency has advanced systems, every analyst should have access to a crime analysis database where they can enter their own data. Why? Because not everything gets recorded in the RMS or other systems. Think of trends like loitering, emerging drug

activity, or informal observations from patrol that could become significant over time.

By maintaining control over your data, you create a flexible, tailored resource that reflects the nuances of your jurisdiction. For smaller agencies, this manual process ensures that no valuable information slips through the cracks, while empowering analysts to deliver actionable insights that support proactive policing and community safety.

Pause the video here and check out this minute video on manual entry.

Linking Tables

Incident Number	Dispatch Type	Date of Report	Address
2023-01011540	Assault	1/23/2023	250 Main St.
2023-05776452	Domestic Verbal	1/23/2023	198 Haverhill St.
2023-03756487	Assault	1/23/2023	254 Lowell St.

Incident Number	Role	PersonNum
2023-01011540	Victim	15456
2023-01011540	Witness	13473
2023-01011540	Suspect	17843
2023-05776452	Victim	25904
2023-03756487	Victim	44949
2023-03756487	Suspect	44890

Incident Number	Role	PersonNum	First Name	Last Name	DOB	Dispatch Type	Date	Address
2023-01011540	Victim	15456	Sara	Richards	1/22/1987	Assault	1/23/2023	250 Main St.
2023-01011540	Witness	13473	Michael	Yates	3/23/2000	Assault	1/23/2023	250 Main St.
2023-01011540	Suspect	17843	Teresa	McDugal	2/2/2000	Assault	1/23/2023	250 Main St.
2023-05776452	Victim	25904	Dawn	Sweeney	7/7/1987	Domestic Verbal	1/23/2023	198 Haverhill St.
2023-03756487	Victim	44949	Deb	Speel	8/7/1986	Assault	1/23/2023	245 Lowell St.
2023-03756487	Suspect	44890	Jon	Sampson	12/23/1967	Assault	1/23/2023	245 Lowell St.

In a relational database, every data type appears in its own table, and the tables are related through common fields such as CAD number. This minimizes duplication of entry and ensures accurate results when querying.

For instance, some crimes have multiple victims and offenders, such as a robbery. In these cases, relational databases are much more helpful. A relational database is a database structured to recognize relationships among stored items of information.

These tables communicate and share information, which facilitates data search ability, organization and reporting. In English, relational databases bring data sets together by a common factor (such as CAD number) so that searches can be much more comprehensive.

See Part 2 – Joining Tables video as a refresher.

Table Structures
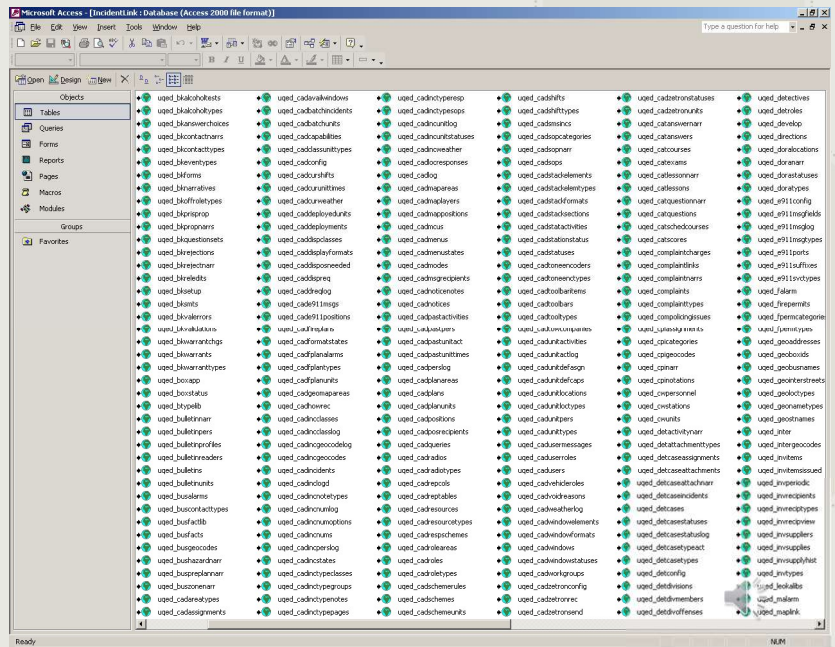
The globes represent the live connection to your RMS/CAD or other database. If you have imported the data, rather than linked the data source, instead of globes you will see mini excel-like looking sheets. These are static data and does not change or update.

However, the linked files, noted by the globes here, are LIVE data, and are automatically updated in your Microsoft Access® database when updated in your RMS system. In other words, when the officer submits a new report on Sunday, and on Monday you look at the data, the Sunday report with automatically be there.

You will probably be confronted by hundreds of data tables. If you're lucky, you'll either get a data dictionary from your vendor or the tables will be named sensibly enough that you can understand them. It's best that you identify which tables contain the data you desire. Mostly, you should be looking for a file or files with your CAD, crashes, crimes, arrests, and citations data. See if you can get your DDF files from your vendor to make this process easier.

Table Structures

- Core data tables
- Library tables
- System Tables
- Link/Activity Tables
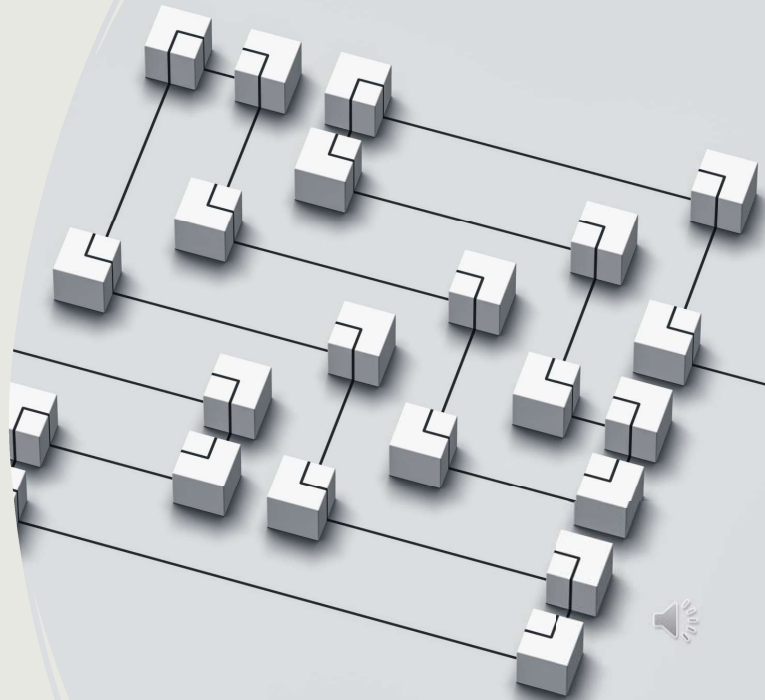
You will find that the many tables in your database break down into one of four types:

- CORE DATA TABLES: which store data about incidents, locations, people, vehicles, property, etc.
- LIBRARY TABLES: which store "look up" values for drop down menus.
- SYSTEM TABLES: that store values needed by the system to perform various tasks.
- LINK/ACTIVITY TABLES: that serve as a bridge between two tables with no common fields.

Core Tables

- Incident
- Offense and *modus operandi*
- Person
- Property
- Vehicles
- Citations

Examples of CORE TABLES are incident tables, address tables, offenses and MO tables, master name files, property tables, vehicle tables, and citation tables. Records management systems and tables vary. They all have similar core tables, but they may have different names and may be structured differently.

Library Tables

- Incident or Offense Types
- Person "Roles"
- IBR Descriptions
- Colors
- Vehicle Makes & Models

Examples of LIBRARY TABLES are incident or offense type, IBR codes, roles of persons (i.e., victim, offender, witness), colors, vehicle makes and models. Library tables can be helpful as one can select or link to the category to more specifically describe data elements. Library tables are often used to populate dropdown lists. They are typically not in free text formats and using library tables can cut down on data entry mistakes and help to standardize the data.

Table Structures

Figure 8: A table's design view

Field data types available in Access

Data Type	Description	Uses in Policing/CA
Text (*Short text in 2013*)	Alphanumeric combinations, up to 255 characters	Crime type, *modus operandi* factor, street name, last name, serial number
Memo (*long text in 2013*)	Text up to 63,999 characters	Narratives, stories, lengthy comments and notes
Number	Real numbers, including decimal numbers	Est. number of members in a gang, number of lanes on a street, number of months of a sentence.
Date/Time	Dates, times, or both	Date and time of report, date of birth, release date for parolees

Once the table is opened in the design view, you can see that there are different field structures for different types of data. For instance, a field that you want to write a short description in may be structured as a TEXT field. Or a date field, such as date of birth, date of incident, etc. might be structured to represent a date in a specific format, such as 1/1/2023.

Table Structures

Data Type	Description	Uses in Policing/CA
Currency	Specially-formatted number field with four decimal places, up to eight integers, and the country's currency symbol	Value of property stolen, bail amount
AutoNumber	Sequential or random number, always unique for each record	A primary key when the table lacks another option for one
Yes/No	One of two values, designated by yes/no, true/false, on/off, etc.	"Include this incident in weekly report," other flags
OLE Object	A file linked or embedded	Person photographs, fingerprints
Hyperlink	Hyperlink to computer, network, or web location	Link to record in the official RMS; link to property record in online assessor's database
Attachment	Like OLE objects, but allows multiple file attachments in a single record	Multiple photographs for an offender, multiple crime scene photographs for a crime, victim and witness statements, related crime bulletins
Calculated	Calculates a value based on another field	Calculate an "Address" field based on multiple discrete fields: StNo, Street, etc.
Lookup Wizard	Not a separate field type. This is a text field that auto-populates lookup properties. Best to set to "Text" and do it yourself.	

==GO TO YOUR VIDEO LIBRARY: "CHANGE DATA TYPE FORMAT"==

There are many data structures to choose from depending on the type of field you have. Here are a few more taken from Bruce's *Microsoft Access for Crime Analysis*. *(See Resource Guide for list of all webinar resources.)*

Pause the video here and check out this minute video on changing the data type and formatting.

Check-In Question #2

When creating master tables, what type of queries are commonly used?

a) Sorting queries

b) Make Table and Append queries

c) Filtering queries

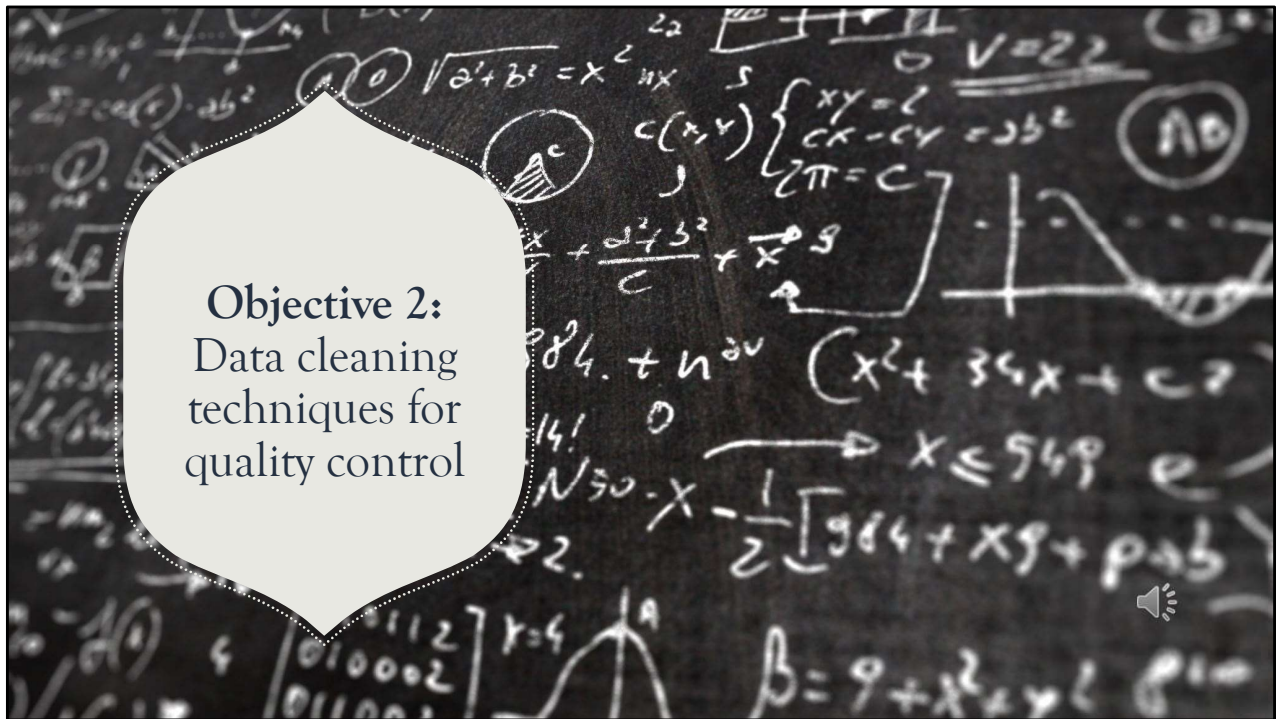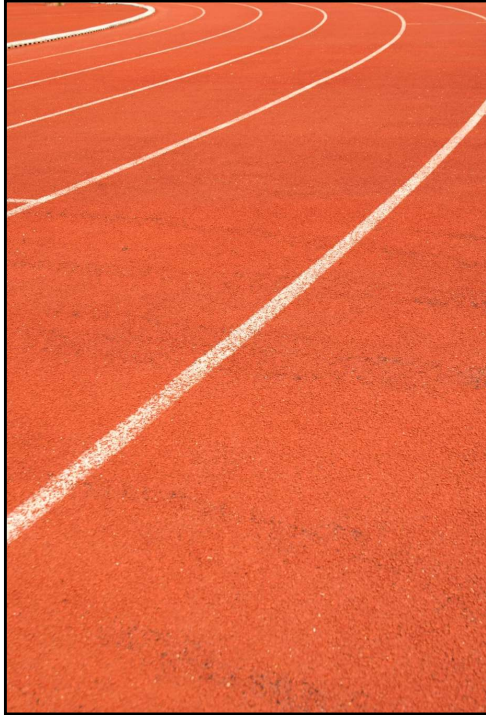d) Reporting queries

Check-In Question #2

When creating master tables, what type of queries are commonly used?

a) Sorting queries

b) Make Table and Append queries

c) Filtering queries

d) Reporting queries

Objective 2: Apply data cleaning techniques to ensure data quality and eliminate errors.

What goes in must come out!

- **What goes in?** – tacos, tacos, tacos, martini, pan dulce. Nighttime chocolate snack.
- **Start running** – no warmup, no stretching
- 10 minutes **FRUSTRATION** sets in
- Who do I **BLAME**?
- Results in a **BAD CYCLE** of poor quality

It's a marathon, not a sprint. Start with: Building blocks!!!

Let's compare building personal fitness to building quality in crime analysis. Imagine committing to shedding a few pounds—you're excited, determined, and ready to dive in. But without preparation—no warm-up or stretching—you quickly hit a wall. Frustration sets in, and you look for someone to blame. The cycle of unmet expectations leads to giving up before real progress begins.

The same thing happens in crime analysis when we don't prioritize **data quality**. If we dive into analysis without first cleaning and preparing our data—our "warm-up"—we'll quickly find ourselves overwhelmed with errors, inconsistencies, and missing pieces. This frustration can lead to inefficiencies, inaccurate findings, and analysis that falls short of its potential.

How do we avoid this? By focusing on the building blocks:
- **Agency Culture:** Establish a commitment to quality through supportive policies and training for officers, supervisors, and dispatchers.
- **Vendor Validations:** Work with RMS vendors to ensure accurate and reliable data collection.
- **Analyst Tools:** Use effective querying methods to streamline your workflow and focus on analysis.

Just like in fitness, it's about consistency and preparation. Starting with small improvements—better training, stronger systems, and cleaner data—creates a

foundation for high-quality analysis. Over time, this allows analysts to spend more time uncovering insights rather than correcting errors.

Remember, success in crime analysis isn't a sprint—it's a marathon built on intentional, incremental progress.
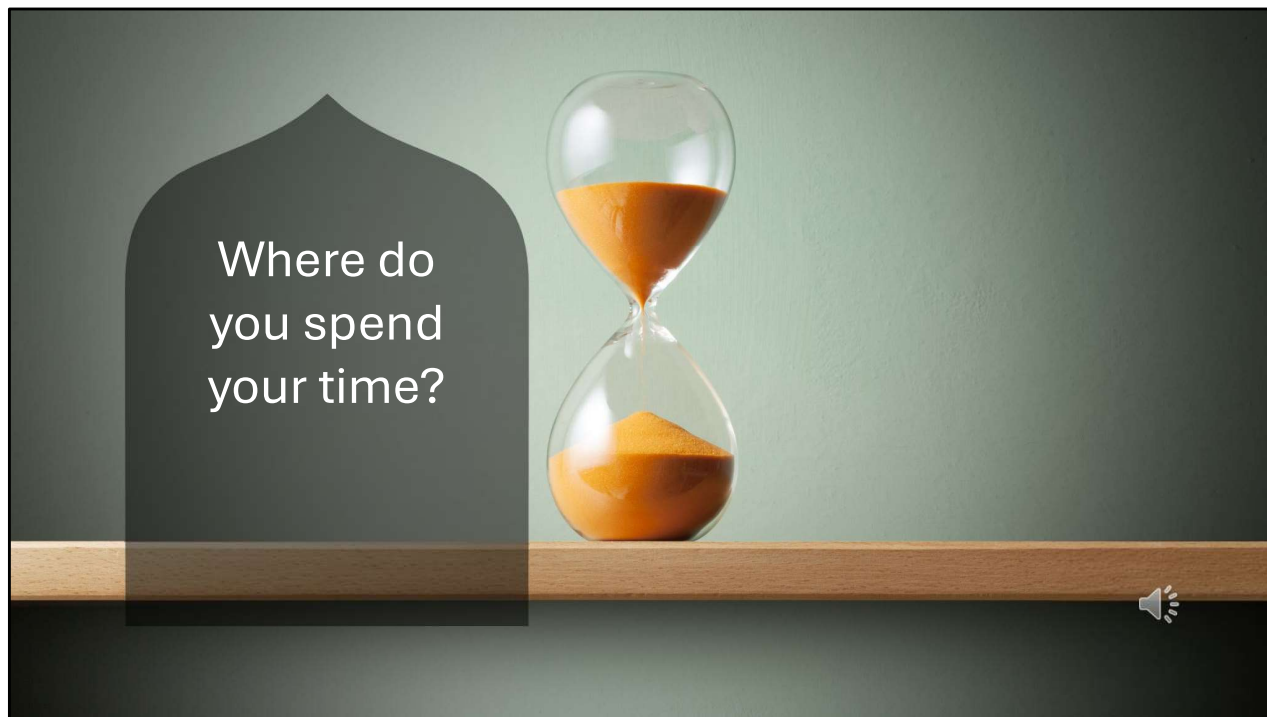
Let's talk about data quality and how it directly impacts your role as a crime analyst. Take a moment to reflect: How much of your time is spent cleaning data versus actually analyzing it? For many analysts, poor data quality is a significant hurdle, consuming valuable time that could be used for deeper analysis and delivering actionable insights.

To shift this balance, we need to focus on creating systems of improvement at every level of the agency. This includes:
•**Policy Design:** Clear, data-focused policies that set standards for data entry and quality.
•**Training:** Equipping supervisors, officers, and dispatchers with the knowledge and tools to enter accurate, consistent data from the start.
•**Vendor Validations:** Ensuring your RMS or CAD systems are set up to reduce errors and streamline data entry.
•**Querying Methods for Analysts:** Developing efficient queries and workflows to identify and address data issues quickly.

When we address data quality as a system-wide priority, we not only improve the reliability of our datasets but also foster a culture where accurate data is valued agency-wide. This means you, as the analyst, spend less time fixing errors and more time focusing on what matters most—producing high-quality analysis that drives strategic

decision-making and operational success.
Better systems lead to better outcomes, and better outcomes allow you to have a greater impact on your agency and your community.

Common Data Errors

Common Data Errors
1. Most of our crime occurred at our police department
2 .HWY/Major thoroughfare – officers aren't clear what address to record; or missing street #s
3. Free typing/mis-spellings
4. Missing data
5. Duplicate master data
6. Address file doesn't match mapping files

Common Data Errors
1. Most of our crime occurred at our police department
2. HWY/Major thoroughfare – officers aren't clear what address to record; or missing street #s
3. Free typing/misspellings
4. Missing data
5. Duplicate master data
6. Address file doesn't match mapping files

Common Error #1
Most Crime Occurs at Police Department

Top individual locations, with three-year totals
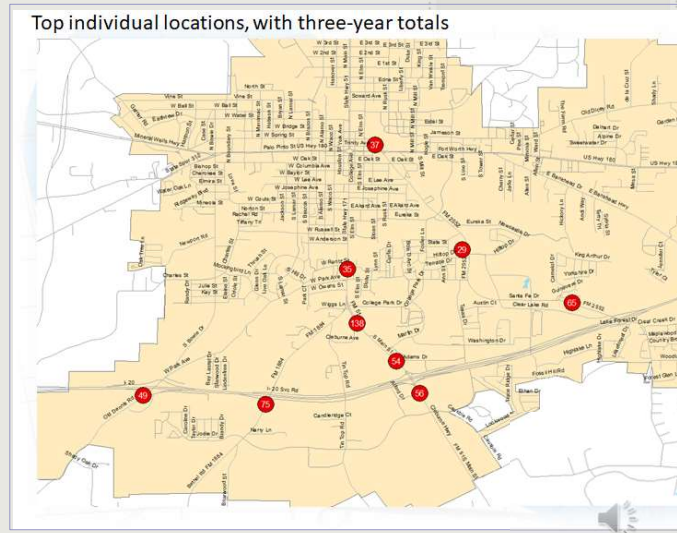
Does most of your crime *appear* to occur at your police department?

Common Error #2
HWY/No Street
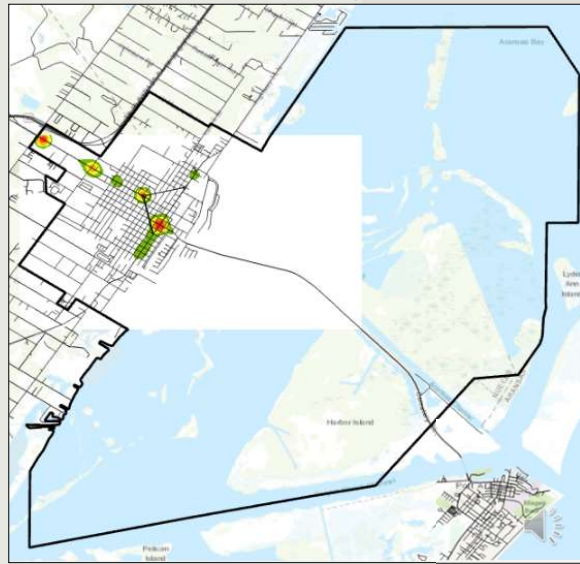
Hit rate is 56% because the data indicates "highway" as the traffic crash location but does not determine a specific location on the highway.

Common Error #3
Free Typing

STREET
Wal-Mart
123 Main Street/Wal-Mart
Wally World
123 M St
South of Town Walmart
Wallmart
123 Main St

Is free typing an issue for your department? Free-typing can make data difficult to ask questions.

Common Error #4
Blank fields

Address	FREQUENCY
(blank)	1,499
150 S Main Street	51
2607 W Daniel Webster Hwy	37
2103 Hickory LN	33
Hwy 75	24
Hwy 222	23
2702 W Daniel Webster Hwy	17
700 W Daniel Webster Hwy	14
322 MARKET ST	13

Do your officers leave fields blank?

Common Error #5
Duplicate Names

FIRST NAME	LAST NAME	DOB
Dawn	Reaby	10/12/77
Dawn	Reeby	10/12/77
Dawn	Reiby	10/12/77
Don	Reeby	77/10/12
Reeby	Dawn	10/12/77
DawnN	Reeby	10/12/17
Dawna	Reaby	10/12/07

Technically, these are not duplicates because every name is spelled differently. This is an accuracy issue, not a system issue. Does your department have an issue with duplicate master names?

Common Error #6
Unmapped HWY
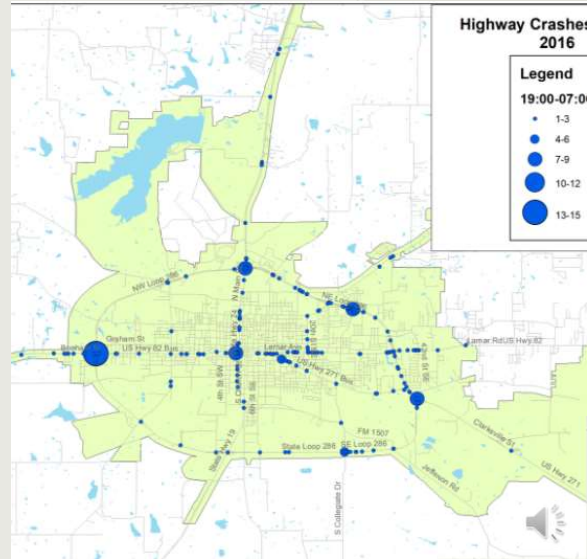
This goes back to inaccurate/incomplete addresses that did not validate. It is not specific to highways or any other type of address.
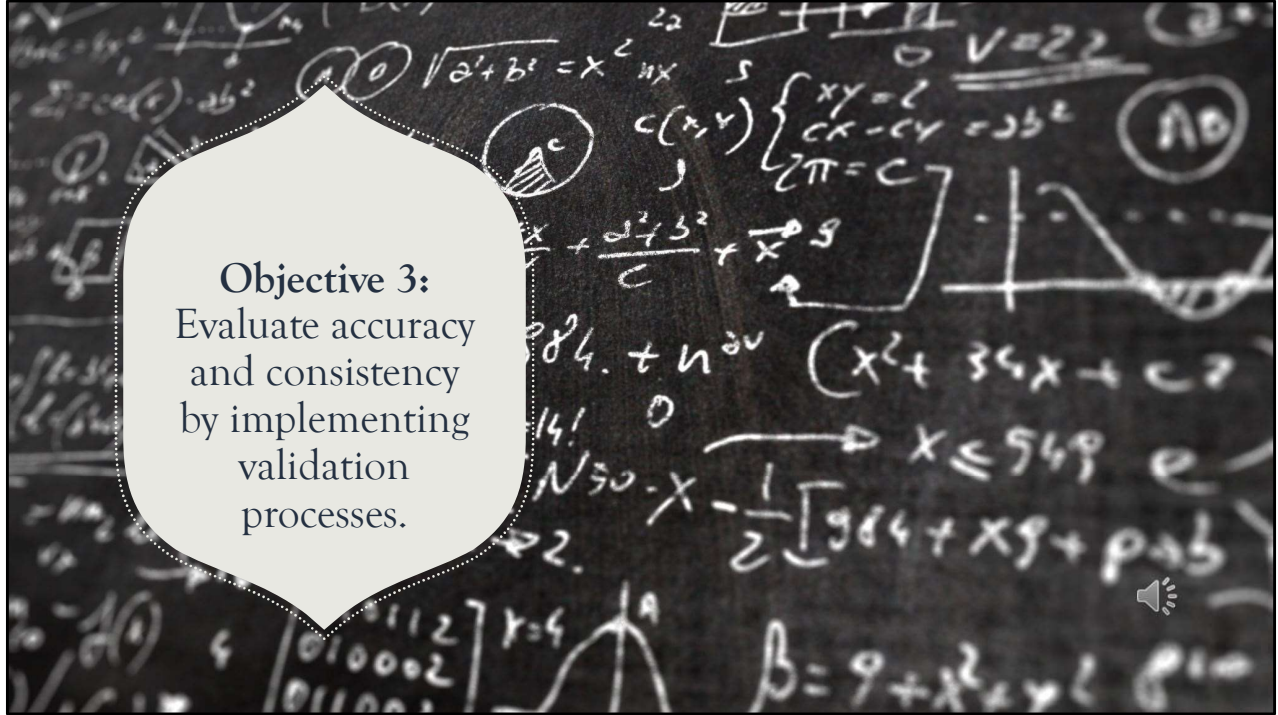
Objective 3: Evaluate accuracy and consistency by implementing validation processes.

Tenants of Data Integrity

COMPLETENESS   TIMELINESS   ACCURACY   FLEXIBILITY OF
                                                            QUERYING

Data quality is the foundation of an effective analytical function. The ultimate goal is to provide actionable insights that drive success, and this can only be achieved with high-quality information.

To meet this goal, agencies must prioritize the integrity of their data. While analyses can include external sources like probation records, pawn data, or open-source information, this discussion focuses on **agency-generated data**—records, dispatch, crash reports, and similar datasets.
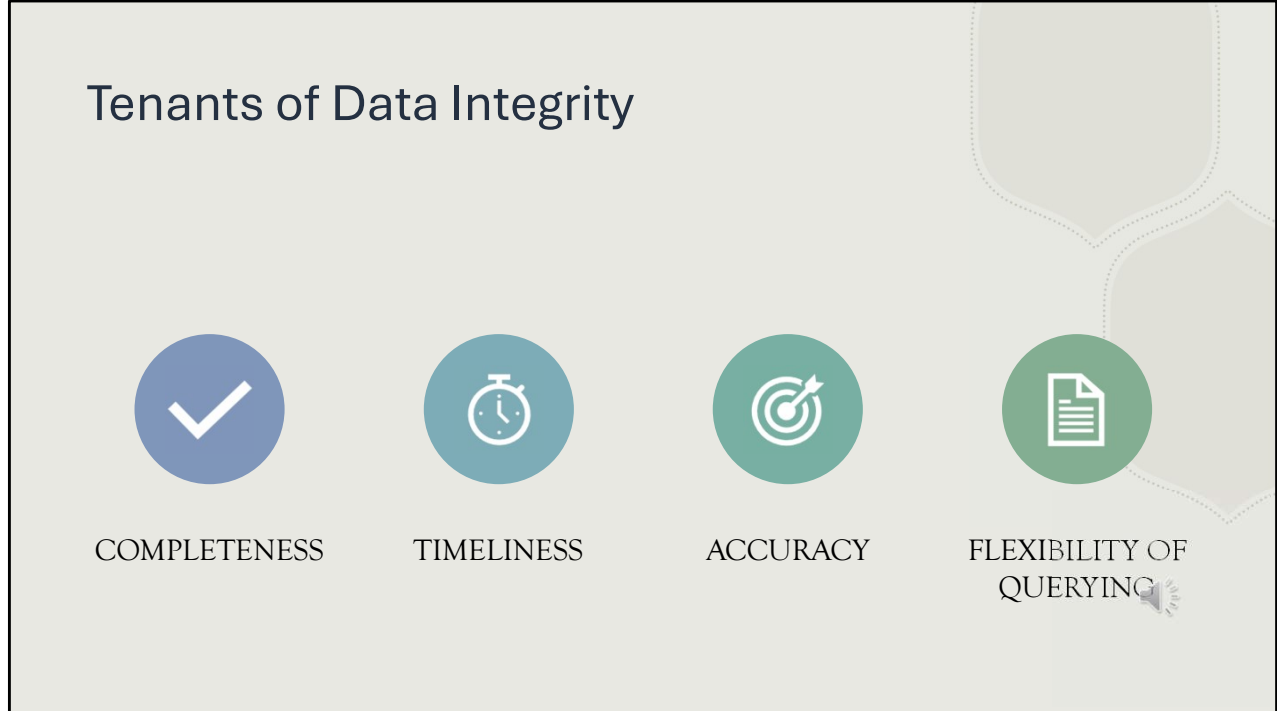
The Four Pillars of Quality Data:
- **Completeness:** Ensure all necessary data is captured and no critical details are missing.
- **Timeliness:** Data must be submitted promptly to allow for relevant and timely analysis.
- **Accuracy:** Eliminate errors and inconsistencies to ensure reliable results.
- **Flexible Querying:** Organize data in a way that supports dynamic and efficient analysis.

When these pillars are in place, analysts can produce impactful insights that support informed decision-making and proactive strategies.

Systematic Methods to Ensure Data Quality

✓ Policy level

🗄️ **Database level**: Required fields, validation rules

👮 **Supervisory level:** Review and approval of new reports

🖥️ **Administrator level:** Single person or unit responsible for running quality checks against database and fixing errors

Policy Level

- How does an agency manage data at the policy level?

- *Identify a working team* from multiple units (executive, analyst, patrol, detectives) to meet and identify top five data quality challenges and solutions. Then plan to address these issues through collaboration and assistance throughout the agency

- *Training patrol and detectives* to help them understand the purpose of analysis.

How does an agency manage data at the policy level?

Identify a working team from multiple units (executive, analyst, patrol, detectives) to meet and identify top five data quality challenges and solutions. Then plan to address these issues through collaboration and assistance throughout the agency

Training patrol and detectives to help them understand the purpose of analysis.

Solutions to Common Data Errors

Common Data Errors	Possible Solutions
1. Most of our crime occurred at our police department	Send back to officers to fix; train officers and supervisors
2. Duplicate master data	Identify duplicates; Train officers

Policy-level solutions look something like this:

When the common data error is most of our crime occurred at our police department, the policy can be to send the report back to the officers to adjust. Training for officers and supervisors is also helpful.

If your common data error is having numerous duplicate master data, you can create a policy-level solution by identifying duplicates systematically and training officers accordingly.

Database Level

Work with the vendor to establish drop down menus and validation processes that ensure more quality reporting.

Example: If crime type is burglary, the following fields must be completed: victim type, date to, date from, time to, time from, entry, exit...
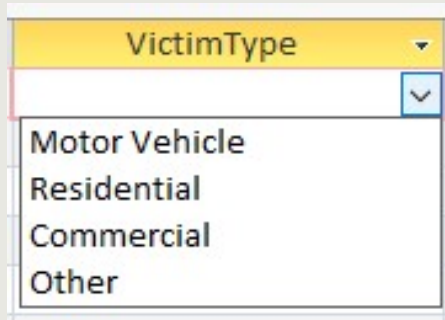
Work with the vendor to establish drop down menus and validation processes that ensure more quality reporting.

Example: If crime type is burglary, the following fields must be completed: victim type, date to, date from, time to, time from, entry, exit...

Solutions to Common Errors

Common Problems	Possible Solutions
4. Free typing/mis-spellings	Work with vendor to create drop down options
5. Missing data	Work with vendor to set up validation rules

VictimType dropdown options:
- Motor Vehicle
- Residential
- Commercial
- Other

What types of validation rules would be helpful for your department? VALIDATION RULES:

- If crime is burglary, then premise type must be Residential, Commercial, or Other.

Supervisory Level

Tighten up Review Process

Example: In most agencies, Sergeants will be the first set of eyes to review the data entry. In some agencies, commanders will serve as the second set of eyes to review the data entry. Records staff will be the final set of eyes to review the data entry send back to the supervisors as they decide.

Work with every supervisor to tighten up the review processes.

Improvement Strategy Timeline

Every 3 months (quarterly) the CAU will review and assess the data quality improvement progress.

Every 4 months, the team will discuss the progress and mistakes of the data improvement during command staff meetings.

Annually a comparison of bad data provided.

Every 3 months (quarterly), the Crime Analysis Unit (CAU) will review and assess the progress of the data quality improvement efforts.

Every 4 months, the team will meet during command staff meetings to discuss the progress and challenges related to data improvement. These meetings should involve key personnel in the data improvement process, including analysts, supervisors, command staff, and relevant department heads. The purpose is to provide an opportunity for stakeholders to review the effectiveness of current data strategies, identify any issues or areas of concern, and collaborate on solutions for continuous improvement.

In these discussions, the team will evaluate the data collection, analysis, and reporting processes, address setbacks, and highlight successful outcomes. Feedback from staff who directly interact with the data—such as officers, data entry clerks, and IT personnel—will be essential for understanding real-world challenges and identifying areas for enhancement. Involving command staff ensures alignment between data-driven initiatives and broader organizational objectives.

By holding these meetings regularly, the team can maintain focus on the overall goal of improving data quality and usage, spot trends or obstacles early, and implement

corrective measures. This collaborative approach fosters transparency, accountability, and a shared commitment to improving public safety through data-driven decision-making.

Annually, the CAU will provide command staff with a comparison of poor data quality (by user or department) versus high-quality data, along with a cost-benefit analysis of importing evidence, body cam footage, and surveillance video into the RMS. This report will be shared for acknowledgment and may prompt policy revisions or updates to training bulletins.

Analyst/Administrative Level

Single person or unit responsible for running quality checks against database and fixing errors

The analyst may have to:

- Create new fields (such as month, year, age, etc.) to produce more in-depth analysis.
- Create a library table for codes or addresses
- Create simple processes that uniform data and allow for a more robust analysis.

It should be noted that an analyst is not responsible for data quality, but because it impacts the analysis
so greatly, many analysts must find
ways to make improvements.

At the administrative level, you may have a single person or unit responsible for running quality checks against database and fixing errors

The analyst may have to:

- Create new fields (such as month, year, age, etc.) to produce more in-depth analysis.
- Create a library table for codes or addresses
- Create simple processes that uniform data and allow for a more robust analysis.

It should be noted that an analyst is not responsible for data quality, but because it impacts the analysis so greatly, many analysts must find
ways to make improvements.

Solutions to Common Errors

Create an Address Library

Analysts in a Texas agency went through the 3,500 address data errors that we identified with the mapping summation tool and created new addresses in Microsoft Access = mapping address library.
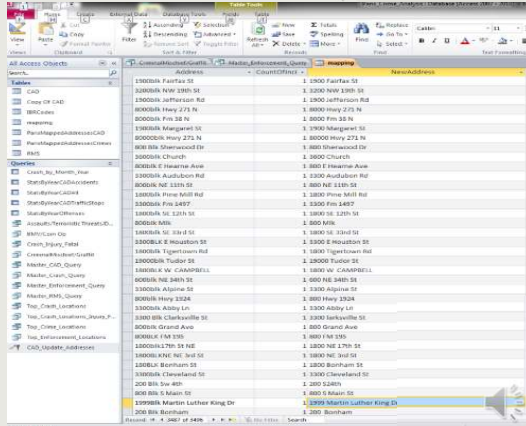
Another solution can be to create an Address Library and add to narratives. Analysts in a Texas agency went through the 3,500 address data errors that we identified with the mapping summation tool and created new addresses in Microsoft Access = mapping address library.
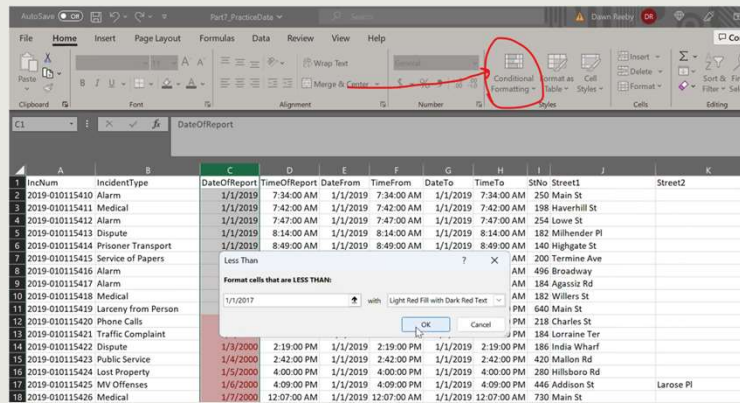
Pause and view the video "Manual Entry Video" to learn more tips.
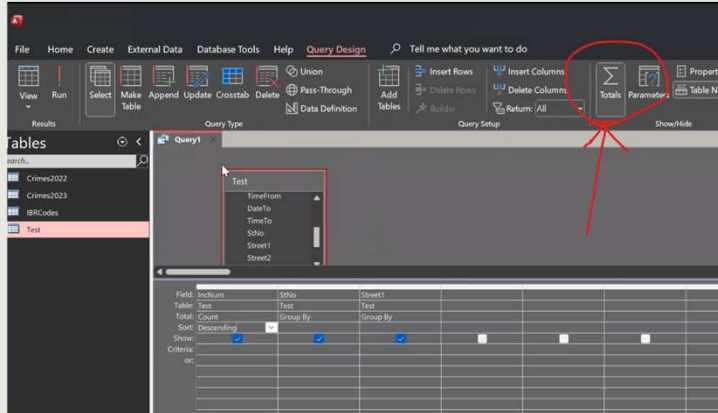
Use Excel to identify data quality issues

GO TO YOUR VIDEO LIBRARY: "FINDING DIRTY DATA USING EXCEL"

Use conditional formatting and sorting options in Excel to identify date and other errors.

Pause and view the video "Finding Dirty Data in Excel" to learn more tips.

Use Access to identify data quality issues

GO TO YOUR VIDEO LIBRARY: "IDENTIFYING DIRTY DATA IN ACCESS"

Use Aggregation Queries to identify repeat locations, bad addresses, and more.

Pause and view the video "Identifying Dirty Data in Access" to learn more tips.

Check-In Question #3

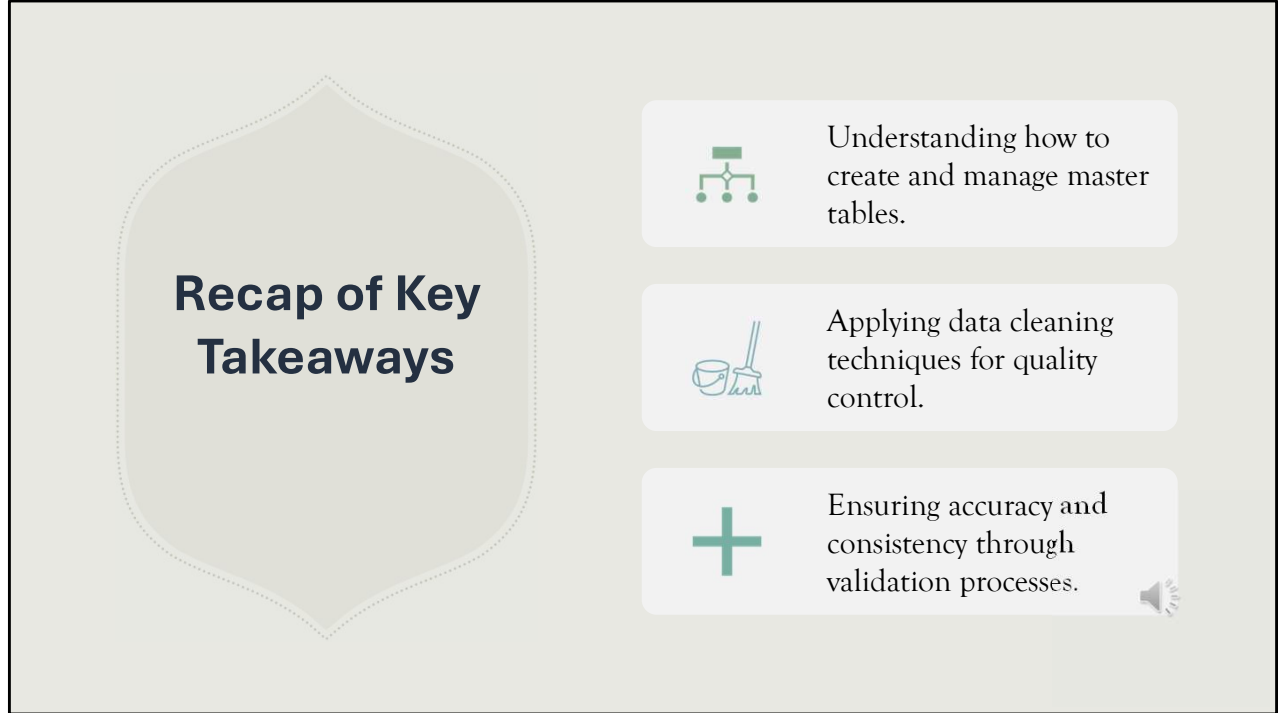**Which is NOT an example of a system-level solution to improve data quality?**

a) Sending reports back to officers for correction

b) Identifying duplicates and training officers

c) Having a single person responsible for quality checks against the database

d) Creating training sessions for officers on crime analysis

Check-In Question #3

Which is NOT an example of a system-level solution to improve data quality?

a) Sending reports back to officers for correction

b) Identifying duplicates and training officers

c) Having a single person responsible for quality checks against the database

d) Creating training sessions for officers on crime analysis

Recap

Recap of Key Takeaways

- Understanding how to create and manage master tables.
- Applying data cleaning techniques for quality control.
- Ensuring accuracy and consistency through validation processes.

Today, we delved into the critical importance of master tables in the field of crime analysis. These tables form the backbone of efficient and effective data management, providing a centralized framework to consolidate information from diverse sources, such as CAD, RMS, and external databases. By integrating these datasets into master tables, analysts can eliminate redundancies, identify discrepancies, and create a clean, standardized foundation for all analytical work.

We explored the practical steps to creating and managing these tables, focusing on structuring them in a way that accommodates flexible querying and analysis. Through data cleaning techniques, we highlighted how to resolve errors, standardize formats, and remove duplications to maintain the integrity of the data. Additionally, we discussed the role of validation processes in ensuring ongoing accuracy and consistency, which is especially critical in law enforcement, where decisions often rely on precise and reliable information.

Master tables are more than just a tool—they are a strategic advantage. By centralizing and organizing your data effectively, you gain the ability to uncover trends, identify patterns, and generate actionable insights with confidence. Whether you're conducting hotspot analysis, staffing studies, or investigative support, a robust master table structure simplifies your workflow and enhances the accuracy of your findings.

As crime analysts, our ultimate goal is to support our agencies and communities with data-driven solutions. By prioritizing strong data systems through master tables, we position ourselves to provide insights that lead to informed decisions and impactful outcomes. Remember: the strength of your analysis starts with the strength of your data.

Next Steps: Advanced Expressions and Automation Processes

In **Part 4**, we will focus on using advanced expressions in Microsoft Access® to streamline complex calculations and automate reporting processes. This session is designed to enhance your ability to work efficiently, reducing manual effort and improving your workflow by utilizing advanced expressions and automation techniques. You'll learn to create reusable templates for repetitive analysis tasks, enabling you to perform more precise analyses with less time and effort. This session will provide you with the tools to automate reports and processes, freeing up your time for more strategic analysis.

IADLEST is proud to support law enforcement agencies with a variety of resources designed to promote the implementation of data-driven operational policing. In addition to this training series, IADLEST offers in-person and virtual workshops, literature, webinars, and other valuable resources aimed at enhancing the effectiveness of law enforcement through data-driven strategies. We encourage you to explore these opportunities to continue your professional development and further strengthen your agency's ability to make informed, impactful decisions based on reliable data.

Thank you for your participation in this session, and we look forward to seeing you in **Part 4**.

Evaluation Survey