Crash Data Analysis 101
DOTD Highway Safety Section

Safety Management Process

Objectives & Safety Goals

STAGE 0: LRSP APPLICATION

Systemic Identification

CRASH DATA

STAGE 1: Network Screening

Identify Risk Factors

ROADWAY DATA

STAGE 2: Countermeasure Selection

Site Selection

Data Collection

STAGE 3: Site Specific Identification

Data Analysis

Implementation

Evaluations

STAGE 4: Economic Evaluation

Implementation

Innovations & Prioritization

Key Points

- Perform these task honestly without fear of litigation
- Interest of the public to protect safety information
- Not the function of the judge or jury to second guess decisions

Agenda

- Safety Program Overview, Safety Management Process, & 23 USC 409
- General Crash Data Information
- Crash Querying
- Crash Data Analysis
- General CAT Scan Information
- CAT Scan Process
- Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations

Crash Report Etiquette

- Personal information should be ignored
- Do not print crash reports
- Do not include copies of crash reports in studies or Stage 0

Duration

- 30 minutes
- 30 minutes
- 15 minutes
- 30 minutes
- 15 minutes
- 45 minutes
- 15 minutes

(minute)
Report vs. Data

- Crash Report: form
  - Completed by Law Enforcement Officer (LEO)
  - Owned by LEO’s Agency
  - Entry Options
    - Open notes – free form
    - Certain formats (number, letters, time)
    - List Selection
- Crash Data: warehouse of elements
  - Subset of report
  - List – Codes

Electronic Reporting

- Crash Report
  - Current Version: 2005
  - Minimum: 4 pages
  - Maximum: no limit
  - Format: Paper; Electronic
  - Applications:
    - paper
    - LACrash
    - vendors

Improved Timeliness

- Crash Report
  - Crash Event – 2 pages
  - Each Vehicle – 2 pages
  - Additional Occupants (8, optional)
  - Narrative Supplement (optional)
  - Alternative Grid (optional)
  - Driver/Witness Voluntary Statement
    - (optional; at least 1 page)
  - Railroad Grade Crossing Supplement
    - (1 page)
Vehicle

Crash Report Elements
• Administration – 8
• Crash – 72
• Vehicle – 72
• Commercial – 17
• People – 43
• Passengers – 15
• Pedestrians – 6
• Train – 106

Crash Data @ DOTD
• HSRG -> DOTD
• Apply current year Road Profile
  – May be different than the year before or after
  – User must research location to ensure no changes

Crash Data Elements
• Crash Number – unique identifier
• Intersection flag – 1 is True; 0 is False
• Collision Manner
• (Human) Severity
• Most Harmful Event
• Event Sequence
• Prior Movement

Crash Data Elements
• Pedestrian flag – 1 is True; 0 is False
• Train flag – 1 is True; 0 is False
• Vehicle Configuration
• Coordinates
  – LEO
  – DOTD

Collision Manner
• Top of Diagram
• Primary pattern 2-vehicle
### (Human) Severity

**LA Codes** | **National Safety Council**
--- | ---
A | K – Killed
B | A – Incapacitating Injury
C | B – Evident Injury
D | C – Possible Injury
E | O – No Injury

### Derived Elements

- Road Departure
- Intersection ID – geography based
- Location Type
- Crash Type – replaced Accident Type
- Contributing Factor
- Vehicle Severity – uses Human Severity scale

### Location Type

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>related to an intersection (signal, roundabout, stop, etc.)</td>
</tr>
<tr>
<td>B</td>
<td>related to a median crossover</td>
</tr>
<tr>
<td>C</td>
<td>related to a driveway</td>
</tr>
<tr>
<td>D</td>
<td>related to an on-ramp or off-ramp</td>
</tr>
<tr>
<td>E</td>
<td>related to a non-road path (sidewalk, bike, rail, golf-cart, etc.)</td>
</tr>
<tr>
<td>F</td>
<td>related to a merging area</td>
</tr>
<tr>
<td>G</td>
<td>related to a bus-stop</td>
</tr>
<tr>
<td>H</td>
<td>related to a shoulder</td>
</tr>
<tr>
<td>J</td>
<td>related to a turn-lane</td>
</tr>
<tr>
<td>Z</td>
<td>not related to any feature</td>
</tr>
</tbody>
</table>

### Crash Type

- Addressed issues
  - Accident Type: Greater detail on single vehicle
  - Difficult to assess non-motorized users, vulnerable users
- Non-motorized User codes: A, B
- Vulnerable User codes: C, D

### Vehicle Severity

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>bus</td>
</tr>
<tr>
<td>N</td>
<td>animal</td>
</tr>
<tr>
<td>P</td>
<td>other fixed</td>
</tr>
<tr>
<td>Q</td>
<td>transport</td>
</tr>
<tr>
<td>R</td>
<td>3+ vehicles</td>
</tr>
<tr>
<td>S</td>
<td>2 vehicles</td>
</tr>
<tr>
<td>T</td>
<td>miscellaneous</td>
</tr>
</tbody>
</table>

- Derived from:
  - Most Harmful Event
  - Vehicle Configuration
  - Movement Reason
  - Roadway Conditions

### Combined Severity

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.4%</td>
<td>0.4%</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.7%</td>
<td>6.2%</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5.9%</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>22%</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>71%</td>
<td>51%</td>
<td></td>
</tr>
</tbody>
</table>
Exercise 1
• Go to <crashdata.lsu.edu>
• Select: Interactive Crash Manual
  bottom left
• Select the website
• Go to Vehicle 2nd page
• Select Sequence of Events
• Select Railway Vehicle
• Explore crash report manual

Agenda
✓ Safety Program Overview, Safety Management Process, & 23 USC 409
✓ General Crash Data Information
  • Crash Querying
  • Crash Data Analysis
  • General CAT Scan Information
  • CAT Scan Process
  • Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations

Crash Query Tools

Crash Data Query
• One data file
  – 1 entry per crash
  – Consolidates
    • Vehicles
    • People
    • Years

  • Crash1 – query by state
  • Crash3 – query by local
  • Crash2 – query by individual

Crash1

Crash3
Crash2

LADOTD Highway Crash Reports

- Beginning Year: 2017
- Ending Year: 2018

Options:
- Count of Crashes by City
- Search for a crash number
- Fatal Crashes Only

Limits - All
- Insufficient Crashes
  - Intersections: <5 crashes / year
  - Segments: <5 crashes / year / mile
- Remedy
  - Increase geographic span
  - Increase time span

Limits - Time
- Most recent
- Minimum: 3 years
- Sometimes: 4 or 5 years
- If LOSS 3 or 4 and:
  - Insufficient crashes
- Consistency
  - Similar operations
  - No major construction
Limits - Segments

- Considerations
  - Too small may be too close to randomness
  - Too large may be too close to average

- Suggested extremes

| (miles) | Urban | Rural | Highways
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.4</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

- If an end-point is at an intersection, trim it down to avoid the intersection’s functional area

Limits - Intersections

- Considerations
  - Too small, may not capture all crashes
  - Too large, may perform excessive QA

- Preferably at least 150 feet*
  *but do not include other intersections or their functional area(s)

- If too close to another intersection, then split distance between intersections

Segment v Intersection

Exercise 2

- Open Crash1 [http://engrapps/crash1/]
- Input project information
  - Years: 2014 to 2016
  - Control Section 246-01
  - Log-mile from: 0.77
  - Log-mile to: 2.38
- How many crashes?

Agenda

- Safety Program Overview, Safety Management Process, & 23 USC 409
- General Crash Data Information
- Crash Querying
  - Crash Data Analysis
  - General CAT Scan Information
  - CAT Scan Process
  - Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations
- SMP’s Site Specific
  - SMP = Safety Management Process
  - Identify project and limits
  - Query crash data
  - Conduct quality assurance
  - Calculate Safety Service Level
  - Review crash patterns
  - Determine mitigation strategies
  - Develop planning level cost estimates
  - Calculate crash reduction cost
  - Determine benefit-cost ratio
**Engineering Judgment**

- Dividing Routes into Segments
  - AADT difference >50%
  - Classification change
- Determine Intersection’s limits
- Determine AADT – complicated
- Interpreting Output
- Correcting Crash Data
- Decipher correctable crash pattern
- Designing Mitigation Strategy

**Getting Started**

**Find Coordinate Converter**

**Find Location**

**Observe Location**

**Refine Location**
**Location Familiarization**

- Most data elements from LEOs ~70% - 80% accurate
- Collision Manner – 76%
- Location at 0.05 mile threshold – 75%
- Without Quality Assurance – Answers ≈ Maybe True
- With Quality Assurance – Answers ≈ Likely True

**Why partial investigation?**

- No need to review
  - error free crashes
  - not road’s fault crashes
  - not over-represented crashes
- Determining mitigation strategies – theory of diminishing returns

**Why investigate?**

- Most data elements from LEOs ~70% - 80% accurate
- Collision Manner – 76%
- Location at 0.05 mile threshold – 75%
- Without Quality Assurance
  - Answers ≈ Maybe True
- With Quality Assurance
  - Answers ≈ Likely True
**Agenda**

- Safety Program Overview, Safety Management Process, & 23 USC 409
- General Crash Data Information
- Crash Querying
- Crash Data Analysis
  - General CAT Scan Information
  - CAT Scan Process
  - Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations

**What is CAT Scan?**

- Crash Analysis Tool
  - Site Level
  - MS Excel based
  - Quantifies
    - Average Daily Traffic
    - Existing Crash Data

**CAT Scan Benefits**

- Uses Highway Safety Manual methodology
- Use of Safety Performance Functions (SPF)
- Empirical Bayes to account for regression to the mean (RTM) bias

**Analysis Use**

- Site Level Analysis
  - Traffic Studies
  - Transportation Management Plans
- File
  - Each segment
  - Each intersection
- Not
  - New alignment
  - Unique locations – no model

**Where to find?**

Internal via Crash1:

External via Website:
CAT Scan Limitations

• Dependent upon great crash data queries
• Dependent upon high quality crash data
• Study location must match a model

Over-coming Limitations

Conducting great crash data queries
  – Time
    • At least 3 most recent years
    • Consistent operations
  – Geography
    • Capture enough area
    • Segments: start with longest of similar classification
    • Intersection: include turn-lanes

Over-coming Limitations

• Crash Data Quality
  – QA (Quality Assurance) section
    • Identifies potential errors
    • Guides partial examination
    • Provides space for correction
  – Goal: 90+% accuracy

CAT Scan Prerequisites

• General knowledge of MS Excel
  – Using Filters
  – Entering data
  – Manipulating data
  – Conditional Formatting – remove duplicates

CAT Scan

• Visual Optimization
  – Spreadsheets hidden
  – Columns minimized
• Open – Nothing locked
  – Careful not to delete
  – Careful to modify

CAT Scan Spreadsheets

• Start: Introductory Information, References, Instructions; Guidance
• Inputs: Place to make global inputs for this workbook.
• QA: (Quality Assurance) Manage crash data review process
• Outputs: Near print ready summary
• CM: Counter-Measure guidance
CAT Scan - Start

- Overview of tool
- Workbook Instructions
- Instruction Guidance
- QA Spreadsheet Overview
- QA Spreadsheet Guidance

Segment Model Development

Example diagram showing the process of CAT Scan and Segment Model Development.

Intersection Model Development

A table with details about Rural and Urban conditions, Unsignalized and Signalized intersections.

Agenda

- Safety Program Overview, Safety Management Process, & 23 USC 409
- General Crash Data Information
- Crash Querying
- Crash Data Analysis
- General CAT Scan Information
- CAT Scan Process
- Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations

Project Process (1of2)

- Start
  - One Segment file
  - Perform QA
- Develop Project
  - Divide Segment (if necessary)
  - Initiate Intersection file(s)
  - Copy appropriate crashes from Segment file(s) to Intersection file(s)

Project Process (2of2)

- Intersection(s)
  - Perform QA
  - Copy appropriate updates or crashes from Intersection file(s) to Segment file(s)
- Publish Results
  - Segment(s)
  - Intersection(s)
Instructions (1 of 2)

0. Location Familiarization ~1 hour
   1. Download Data ~10 min
   2. Import Crash Data ~2 min
   3. Initial Documentation ~15 min
   4. Outputs – Safety Comparison ~2 min
   5. Outputs – Pattern Recognition Analysis ~5 min

Instructions (2 of 2)

6. Quality Assurance: Use if
   – LOSS 3,
   – LOSS 4, or
   – Correctable Crash Pattern
   6. Quality Assurance < 10 min/crash
   7. Printing ~2 min
   8. Reporting Errors ~2 min

1. Ensuring good location

2. Import Data

3. Initial Documentation

Segments:

Helping Hint: use limits to develop project
Exercise 3: Initialize
- Open Folder “Basic Crash Class files…”
- Open Document “CAT Scan – Segments”
- Open Folder “Basic Crash Analysis Class - Exercise 3”
- Open Document “Basic Crash Analysis Class – Reference…”
- Open Folder “Segment”
- Open Document “crash1 (84)”

Exercise 3: Step 2
- Select All & Copy from “crash1 (84)”
- Go to “CAT Scan – Segments”
- Go to spreadsheet “Data”
- Select All & Paste
- Save “CAT Scan – Segments” as “CAT Scan - LA 57 C-S 246-01 LM 0.77 to 2.38 – 20180425”

Exercise 3: Step 3
- Close Document “crash1 (84)”
- Update values to match reference file
  - Log-miles
  - Highway Classification
  - AADT in cell 132
Macro Trust Issues (1 of 2)

- Click the File tab
- Click Options (bottom left). A new dialog box should open.
- Click Trust Center (bottom left)
- Click Trust Center Settings (center right). A new dialog box should open.
- Click Trusted Locations (second option from top left)
- Click Add new locations... (bottom center). A new dialog box should open.

Macro Trust Issues (2 of 2)

- Click Browse... A new dialog box should open. Find the location that you will save these files. Click Ok.
  [You may have to repeat this step if you save your files in multiple locations. Note trusting a parent file will enable trust for each child file within.]
- Click Ok.
- Click Ok.
- Click Ok. (yes three time)
- Close your file.
- Reopen your file. It should be work properly now.

4. Outputs – Filters

- Filters out crashes that does not meet the below criteria
  - Segment
    - Intersection = 0
  - Intersection
    - Intersection = 1
    - Intersection ID of interest

Safety Service Level

LOS5 1: High potential for safety improvements
LOS5 2: Moderate potential for safety improvements
LOS5 3: Low potential for safety improvements
LOS5 4: Nil to low potential for safety improvements

4. Outputs – Safety Comparison
Pattern Recognition Analysis

- Crash attribute as Binomial Trial

\[ P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \]

- Each trial compares: subject % vs. class %

114 5. Outputs – Pattern Recognition Analysis

115 5. Outputs – Pattern Recognition Analysis

116 5. Outputs – Pattern Recognition Analysis

117 5. Outputs – Pattern Recognition Analysis

118 5. Outputs – Pattern Recognition Analysis

119 5. Outputs – Pattern Recognition Analysis

120 5. Outputs – PRA

- Gray entries: Difficult to mitigate with infrastructure or traffic operations strategies
  - Same direction movements
    - Rear-ends
    - Left-over-takes
    - Graze with flow
  - Driver behaviors: impairment, etc.
  - Non-road issues: lost cargo, etc.
5. Cumulative Time of Day

![Cumulative Time of Day Graph]

6. Quality Assurance

**Columns**
- A:I
- CH:CV
- DA:DR
- DU:EP
- ES:EH
- FI:GJ
- GM:HP

**Errors Detection:** Functions that detect possible errors.
**Analysis:** Functions that provide further analysis.
**User Entry:** Space to upgrade the data, provide notes on the crash, provide potential solutions, and provide more specific location information. Changes in this section will be reflected in the Output.
**Error Notation:** Space to consolidate errors for reporting.
**Calculations:** Functions that perform the work for other parts of this spreadsheet.
**Research:** Space to consolidate crash data for research.

6. QA Management

- **Authenticity**
  - A = Authentic
  - B = Beyond Limits
  - Z = Not Authentic
- **Status 1: view**
  - Completed crashes
  - Need Reviewing
- **Status 2: review**
  - Review location
  - Update coordinates
  - Evaluate location
  - Propose solution
- **Status 3: quality**
  - Resolve errors
  - Over-write errors (very sparingly)

6. QA User Entry

- **Update Lat/Long or Location Evaluation** (not both)
- **Review Notes**
- **Possible Solutions**
- **Others depends on review and error mitigation.**

- **Intersection ID**
  - Use "0" when not related to intersection
  - When related to Intersection:
    - Intersection, use [specific id]
    - Segment, use "?"

6. QA User Entry

- **Review Notes**
  - Can't view: no crash report available to review
  - not within study area: removes crash from analysis
- **Intersection ID**
  - Use "0" when not related to intersection
  - When related to Intersection:
    - Intersection, use [specific id]
    - Segment, use "?"

6. QA EDA

- **Error Detection Algorithms**
  - Compare data and note issues
  1. Collision Manner & Crash Type
  2. Collision Manner & Angle
  3. Intersection & Intersection ID
  4. Collision Manner & Location Type & Access
  5. Same as Original
  6. Might be Night
  7. Crossed Center & White Line
  8. Road Depart & Crash Type
QC vs. QA

• QC: Quality Control
  – Performed before and after QA
  – Before: restrict inputs
  – After Decide:
    • Need more QA
    • Use or Not

• QA: Quality Assurance
  – Network/Global level
  – Project level
  – Sample and correct

6. QA Analysis

• Solution Count
• Evaluation Score
  – heuristic for evaluation management
  – Factors
    • Segment – 18
    • Intersection – 13

6. QA Error Notation

• Mostly Automatic
• “why”
  – Corrected value = Original value
  – Resolve: delete corrected value
• Other: for notation of errors not otherwise captured

6. QA Process (1 of 4)

A. After establishing parameters – Input
B. After running PRA – Output
C. Initiate QA spreadsheet
  1. Filter Authenticity by Z – delete entries
  2. Filter Status1 by Yes
D. Start at bottom – work upward
E. Run PRA – if not first time and segment

6. QA Process (2 of 4)

F. Copy crash number
G. Paste into Crash2 & run query
   1. Open location
   2. Open report
      a. If Investigating Agency <> A (LSP)
         – open crash report
      b. If Investigating Agency = A (LSP), Then go to ThinkStream to open crash report

6. QA Process (3 of 4)

H. Review location and either:
   1. Update Lat/Long or
   2. Location Evaluation
I. Review Notes
J. Possible Solutions
K. Resolve Errors
6. QA Process (4 of 4)

L. Loop back to F – until no new crashes appear
M. Reapply filter
N. Do more crashes appear
   1. Yes: Loop back to E
   2. No: proceed to Step 7

7. Printing

- Output spreadsheet
- Select Print Preview
  - Segment: 4 pages
  - Intersection: 3 pages
  - TMP: +2 pages (manual addition)

8. Reporting Errors

E-mail CAT Scan files:
Bryan.Costello@LA.gov

Agenda

- Safety Program Overview, Safety Management Process, & 23 USC 409
- General Crash Data Information
- Crash Querying
- Crash Data Analysis
- General CAT Scan Information
- CAT Scan Process
  - Interpreting the Data, Selecting Mitigation Strategies, BCA, and Other Considerations

Mitigation Strategies

- Engineering Judgement
  - Decipher correctable crash pattern
  - Design mitigation strategy
- CM spreadsheet
- CMF Clearinghouse
  <http://www.cmfclearinghouse.org/>
Economic Appraisal

- Safety Benefit Cost
  - Benefits of reduced crashes
    - CMF * Facility Average
    - When using multiple CMF, one should not reduce the same crash more than once
  - Cost of modifications
    - Estimated construction
    - Not right-of-way acquisition

Other Considerations

- ADA Transition Plan
- Complete Streets Policy
- Bicycle Planning Tool
  - Network Analysis
  - Recommended Facility Type
- Local Plans
- Stakeholder Input

DOTD Maps

ADA Transition Plan

- State routes with existing sidewalks only
- ADA Program Funds available
- Required to fix deficiencies if within project limits (even PRRR)

Bicycle Planning Tool

Recommended Facility Type:

Questions

Bryan.Costello@LA.gov