Crash Data Introduction:

Crash data are information that comes from a reportable crash. A reportable crash according to Title 75, Pennsylvania Consolidated Statutes, Section 3746(a) is: an incident that occurs on a highway or traffic way that is open to the public by right or custom and involved at least one motor vehicle in transport with:

- Injury to or death of any person, and/or
- Damage to any vehicle to the extent that it cannot be driven under it's own power in it's customary manner without further damage or hazard to the vehicle, other traffic elements, or the roadway, and therefore requires towing.

When reviewing data in the tables, keep in mind the following:

- Crash data does not include non-reportable crashes or near misses
- Crash data may not contain complete information, some elements may be unknown
- Crash data is dynamic:
    - The Department receives crash reports in paper and electronic formats; however, not in chronological order.
    - By law, police agencies may submit crash report forms up to 15 days after the crash event. However, this often takes longer.
    - The Department does not process reports in chronological order. For efficiency, data analysts may process reports by region or geographic area rather than date sequence.

Defining Criteria:

When writing your query you will want to decide upon a few things before you begin:

- What exactly do I want to count or lookup?
- What tables contain the data I need?
- How do these tables relate to one another?
- Do I want any calculated fields?
- Do I want to limit my results with any specific criteria?
- How do I want to sort my records?

These questions may seem relatively basic, but it is very important to know exactly what you want and are able to extract it correctly as your decisions may be made on the basis of the results. The data in these tables was compiled mainly to make informed decisions to reduce the number and severity of crashes. It cannot analyze and interpret the data for you. Therefore you must analyze the data to determine its value in helping you make decisions with your own set of criteria. Before accepting the results of your queries try and verify its accuracy before making any decisions based upon the output. Things you may ask yourself:

- Does this data make sense?

- Are these results what I expected?
- Do the crash types match the roadway's geometry?

Additionally, you should:

- Compare data to older reports for the same location
- Compare data to other reports for similar locations
- Look for known incidents or specific crashes (i.e., publicized fatal crashes, etc.)
- Use maps and other GIS tools to confirm locations
- Start with small pieces of a report and combine them together to build a complex report. (For example, if you are looking for causations along an entire route look at one causation first, and then another. Alternatively, you can look at road segments rather than the entire road.)

Crash Location Criteria:

Intersection versus **Mid-block** locations:

- *Mid-block Locations:* PennDOT will populate the database with <u>one and only one</u> roadway. Normally it will be numbered as roadway **number 3.**
- *Intersections:* PennDOT will populate the database with a minimum of two roadways. Generally the roads will be numbered as 3 and 4; however it could go up to roadway 9 depending on the configuration of the intersection.

Categories of Roadways:

State Routes are PennDOT maintained roadways which include Interstates, US (Federal) Routes, and State Roads. These will contain:

- Two digit County
- Four digit State Route Number
- Four digit Segment (or 9999 for unknown)
- Four digit Offset (or 9999 for unknown)
- They May or may not have a Street Name and Street Ending

Local Roads include roads maintained by the County, Local Municipality, or privately maintained roadway open to the public. These will contain:

- Two digit County
- Route number (usually blank but might contain a 3 digit route number preceded by a letter, most commonly T for Township Routes or C for County Routes)
- Street Name
- Street Ending

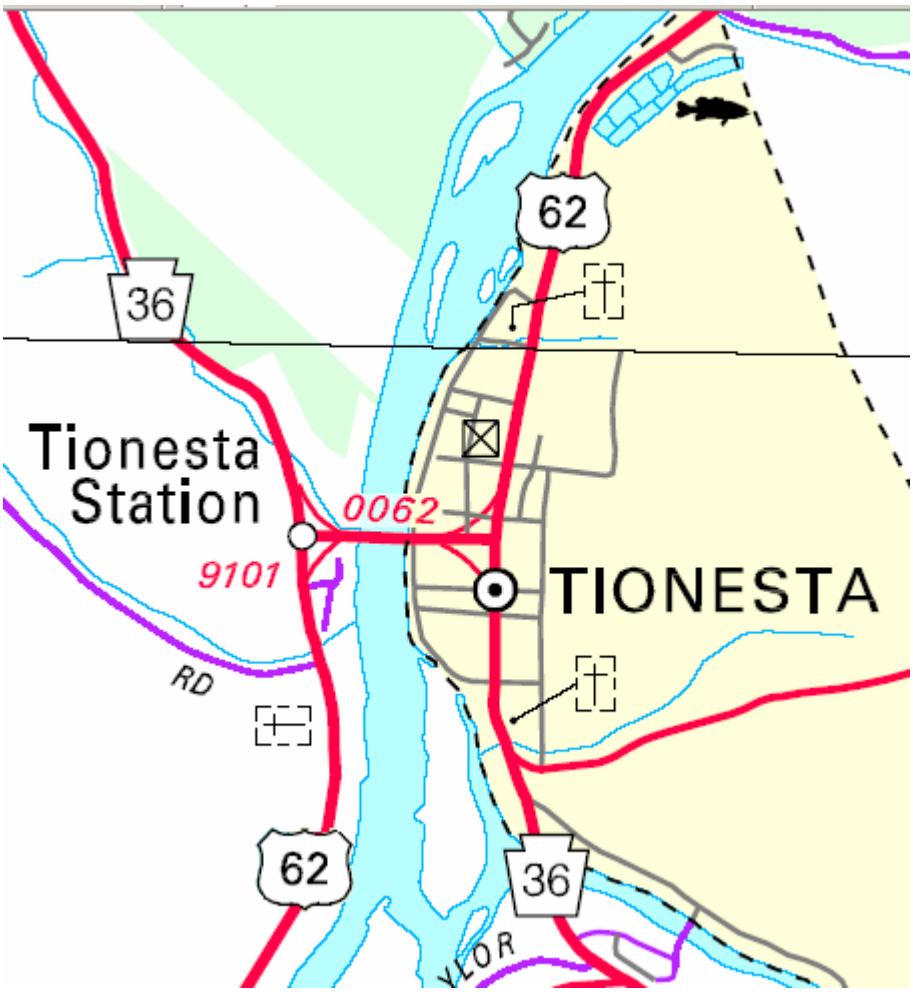
Defining location on a State Route:

Each State Route is separated into distinct but not necessarily contiguous <u>*segments*</u>. The segments commonly begin with segment 0010 and increment by 10 until the end of each route within each county. The segments and offset within the segments increase as you head either east or north. Upon reaching a county line, the segment will usually start again at 0010 (except for interstates).

Divided Highways are separated by a barrier, land, or a minimum of 4 feet of painted lines and are heading in opposing directions will have two separate segment designations, one for each direction. These are known as *even/odd segment pairs*. The odd segment corresponding to an even segment will always be one digit higher than the even segment. For example, US Route 15 in Adams County, starting at the Maryland state line, has 2 northbound lanes and 2 southbound lanes separated by a median. The northbound lanes are considered segment 0010 and the southbound lanes are considered segment 0011.

Null Segments:

There are numerous examples of State Routes that are disjoint. The point from which the first section of a state route ends to the point where the next section of the same state route picks up again within any given county is referred to as a *null segment*. This can occur where either a roadway is shared by multiple state routes (example 22/322 north of Harrisburg) or where a section of state route has been turned back to a local municipality (example Route 72 in the City of Lebanon)



In Tionesta Borough, Forest County, US Route 62 heads north into Tionesta Station, crosses the Allegheny River then heads north towards Warren. Route 36 heads north into Tionesta, crosses the river and then heads north towards Titusville.

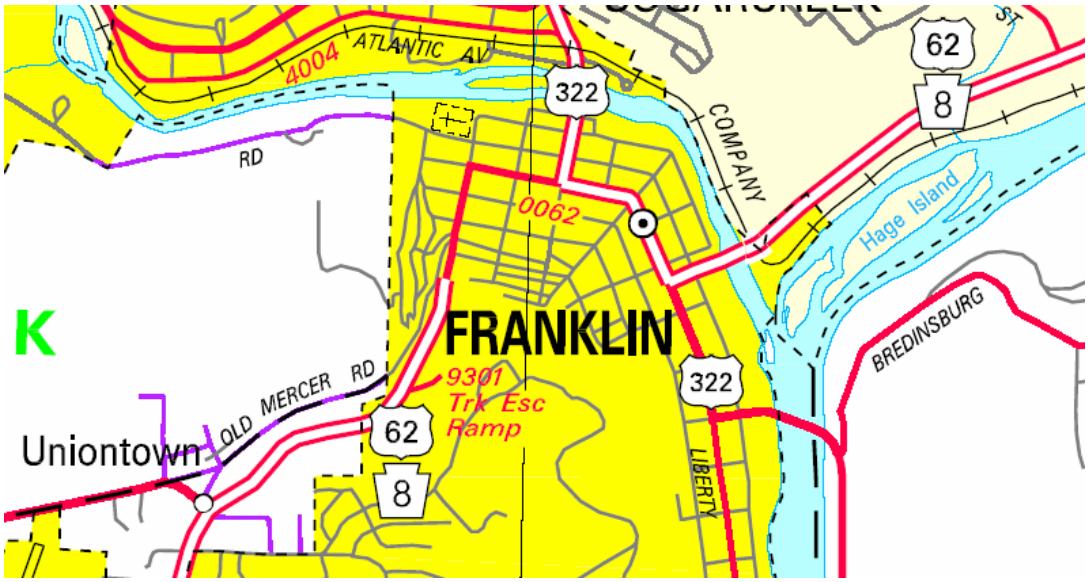
- Any crashes that would occur on the bridge but not at an intersection on either side would be coded as County 27 Route 0062.

- If the crash were to happen at an intersection at either end of the bridge, then both roadways would be included.

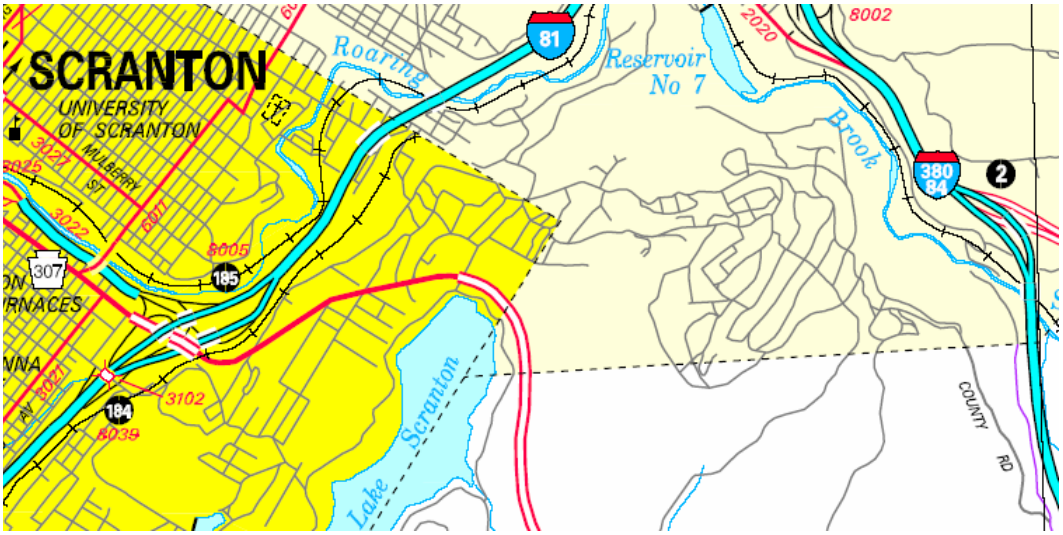
- Between the two intersections, a ***null segment*** would exist for route 36, and no crashes would be coded against that null segment.

When a roadway has multiple route designations, the way to determine which will be the null segment is based on 2 factors. The first factor is route designation. Interstates always take priority, followed by US Routes, and finally State Highways. If there are multiple routes that are in the same category then the lowest number prevails.

For Example:



For US Route 62/US Route 322/Route 8 in the city of Franklin. SR0062 and SR0322 are US highways and SR0008 is a state highway. That section of roadway is considered SR0062 and would be a ***null segment*** for both SR0008 and SR0322.

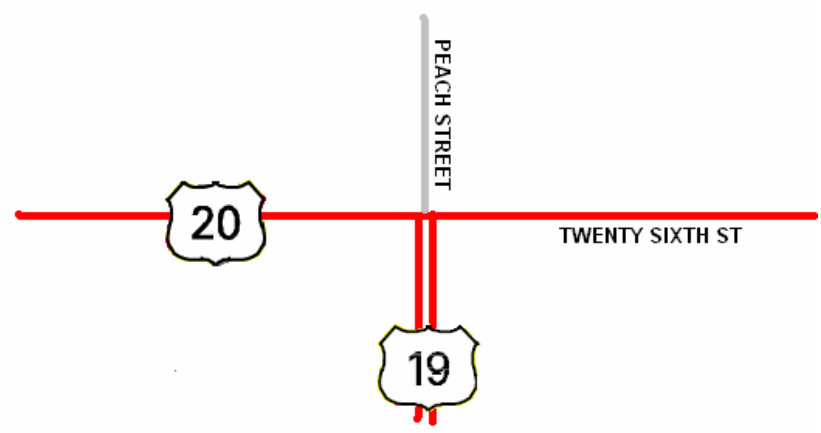


Interstate 380/84 in Lackawanna County would be coded as SR0084 and would be a ***null segment*** for Interstate 380 (SR0380).

Roadway Data at Intersections:

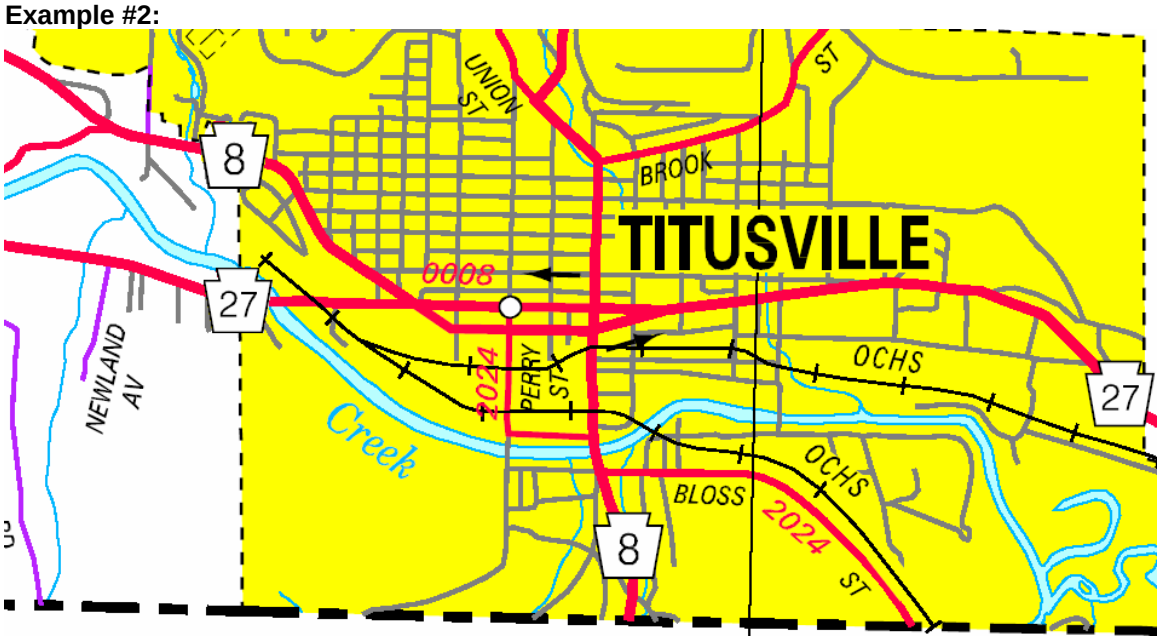
When a crash is located at an intersection with a divided highway, both the even and odd segment will be included with the roadway data. The exception would be when the divided highway is made up of two separate named roadways, usually within a city or borough.

Example #1



In the City of Erie, US Route 19 ends at US Route 20 at the corner of 26th and Peach. US Route 19 is a divided highway and US Route 20 is a non-divided highway. Peach Street becomes a local one-way street north of US20. The 4 roadways that will be included with this location are the even and odd segments of the divided highway (SR0019, segments 0410 and 0411), the non-divided highway (SR0020, segment 0580) and the local road (Peach Street).

Example #2:



In the City of Titusville, Route 8 splits and becomes two separate one way streets, one carrying north bound traffic (Central Avenue) and the other carrying south bound traffic (Spring Street).

- The intersection of Perry Street (SR2024) and Spring Street (SR0008) would have only 2 roadways listed. SR2024 segment 0010 offset 0231 and SR0008 segment 0031 offset 0896.

- The intersection of Perry Street (SR2024) and Center Avenue would have 3 roadways listed. SR2024 segment 0010 offset 0000, SR0008 segment 0030 offset 0896, and the local portion of Perry Street.

Non-typical Roadway data:

- Local Roads:

Local roads do not follow the same conventions as State Routes in that they do not get stored with a numeric route, segment and offset. They also do not have corresponding latitude and longitude coordinates. Compiling data on local roads cannot be done with as much accuracy in regard to location as with state roads. Finding an intersection between 2 local roads can be achieved by finding all crashes that happen on each road separately and combining the subset. Be sure to narrow down your search to only a specific municipality (see municipal code table).

Finding crashes for a portion of a local road cannot be determined since there is no systematic method used for determining the distinct crash location. Definitively finding all crashes that happened on Maple Street between Front Street and Pennsylvania Avenue is simply not possible. The only way this might be achieved is to determine that no crashes happened on Maple Street, or if there were crashes on Maple Street, to determine the location of each one, crash by crash, and determine if they are within the location range. If any of those crashes were midblock crashes… all bets are off.

Another problem is that street names are not always consistent. Though the majority of streets are easily kept consistent (no one ever misspells OAK STREET) there are cases where streets are referred to by multiple names, commonly misspelled, have a space or hyphen in the name that aren't used consistently, or are referred to by both a local name and township/county route number.

- Interchanges and Ramps

Ramps at interchanges are a separate category of state roads and are normally designated with a state route designation between 8001 and 8999. These route numbers are distinct for each interchange within each county (virtually every county has a state route 8001). Each of the ramps within one interchange has the same SR number, but each ramp will have a different segment number. The segments are generally assigned in a clock wise manner.

Crashes that occur at interchanges that involve a ramp will either be a midblock crash on the ramp or an intersection between the ramp and another roadway.

Determining location ranges:

When PennDOT is asked to produce a crash history for a specific location we generally follow certain guidelines.

- For intersections, we include all crashes that occur within 100 feet of the intersection.

- For interchanges, we include all crashes on all ramps and on the state route within 100 feet of those ramps.

- Point A to point B requests are done for each route separately when multiple routes are involved (common where null segments exist, or when a named local street has multiple state route designations).

- For those that involve interchanges (such as Interstate X from Exit 10 to Exit 15) the report is typically run to include 100 feet beyond all entrance and exit ramps. For multiple sections (i.e. Interstate X

from Exit 10 to Exit 11 and Exit 11 to Exit 12), the delineation between the two requests is where the main roads cross each other.

- When a request spans a county line, each county is run separately.

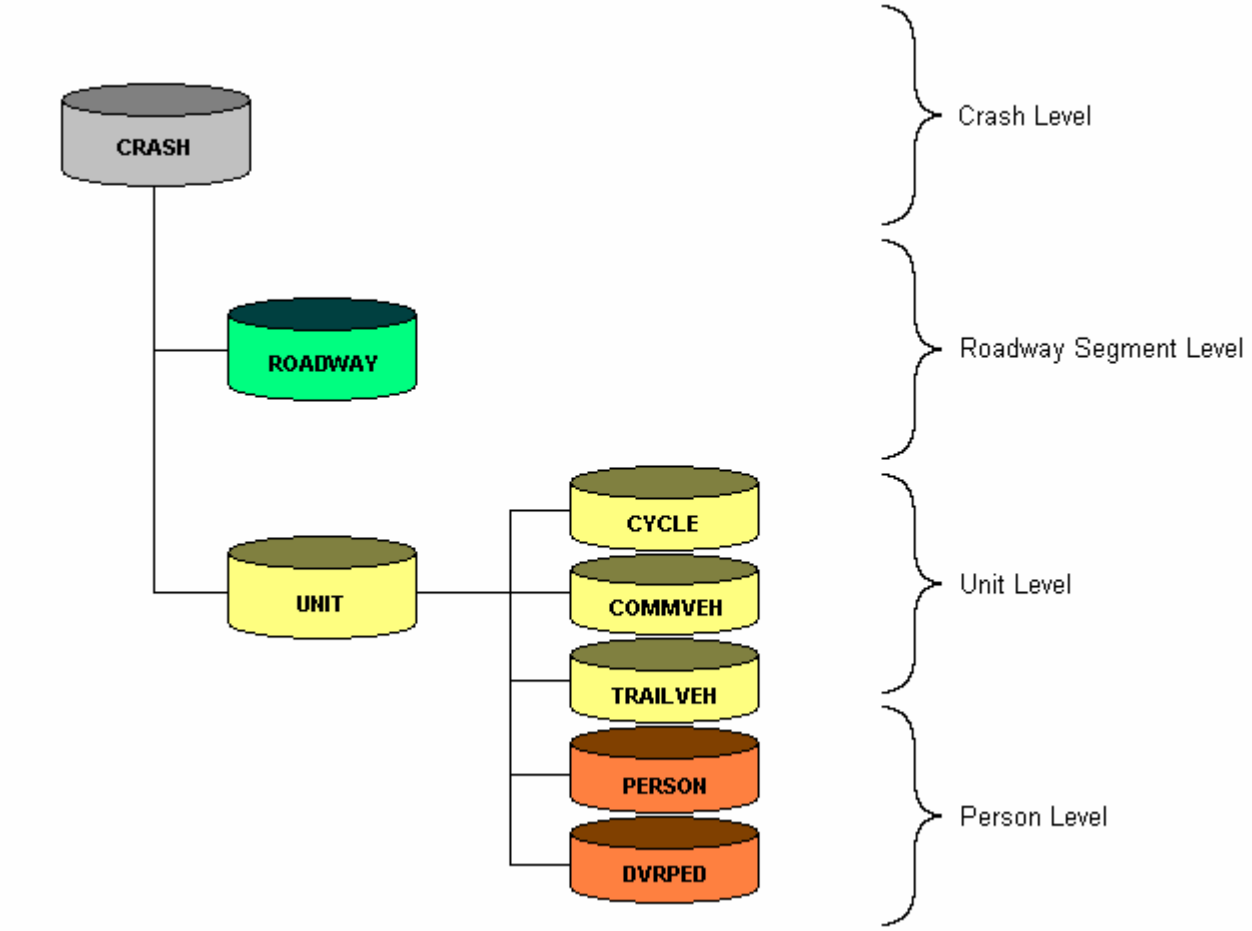
The Database:

The data that has been sent is in MSAccess XP/2002. Whether you keep the data and work with it in MSAccess or export it to another database you will want to extract your data based on specific criteria by using SQL. You can use Structured Query Language (SQL) to query, update, and manage relational data. Packages such as MSAccess or Crystal Reports (reporting software) also have query builders that aid in the query building process. Before you begin writing your queries it is important to familiarize yourself with the data elements so you can better define your criteria so that the results you desire are accurate.

The tables that have been sent to you are shown here with a general description about the data elements contained within.

Table Name	Description
CRASH	Information about the crash such as: **Where:** Latitude, Longitude, County, Municipality, Work zone **When:** Date, Time, Day of Week, Hour of Day, Month of Year **Item Counts:** People, Vehicles, Unbelted, Fatal, etc.
ROADWAY	Information about all the roadways involved in the crash such as: Route number or name, Segment, Offset, Type of Roadway, Rating, and many other Roadway defining elements.
UNIT	Information about all vehicles involved in the crash such as: Body Type, Most Harmful events and all Harmful Events, Movement, Position, Unit number in the crash And other vehicle related information.
CYCLE	Information that pertains to motorcycle/pedal cycles, such as helmet usage and appropriate attire and other accessories such as side bags
COMMVEH	Information about commercial vehicles, such as carrier information, the cargo body type, Hazmat information, and official agency registration numbers.
TRAILVEH	Information about all the roadways involved in the crash such as: Route number or name, Segment, Offset, Type of Roadway, Rating, and many other Roadway defining elements.
PERSON	Information about all people from all units related to the crash such as: Age, Sex, Where they sat and in which vehicle, Were they ejected from the vehicle? etc.
DVRPED	Information about drivers and pedestrians from all units related to the crash such as: Drug and alcohol results, Actions in the crash, Driver history information.

The next visual will illustrate the level of the information and a general understanding of how the tables relate to one another. When you remain within the CRASH table you will normally be counting crashes. As you move down through the table structure you will very often be counting other attributes involved in the crash.



You should familiarize yourself with the data dictionary and the column code or constraints criteria for specific fields and the tables in which they occur. The data dictionary and constraint tables were included as DataDictionary.doc.

The constraints or fields that have a Column Code were intended to limit you to only valid values on an application level. Since you will be working directly with the raw data you will want to be careful and only apply valid criteria as your filter to a field.

The Bureau of Highway Safety and Traffic Engineering would like to wish you well on the use of the data, and thanks you for your cooperation.

Most data requests are relatively straightforward. Others may take a bit of time to get sorted out before writing your query. If you encounter a problem such as a location range that is unusually complicated, we would be more than willing to assist you with your request. You can contact us at (717) 787-2855 or email us at penndotcrashhelp@state.pa.us