OTREC

FINAL REPORT

OREGON TRANSPORTATION RESEARCH AND EDUCATION CONSORTIUM

Development of the Oregon Traffic Safety Data Archive: Phases 1 and 2

OTREC-RR-11-28
March 2012

A National University Transportation Center sponsored by the U.S. Department of Transportation's Research and Innovative Technology Administration

DEVELOPMENT OF THE OREGON TRAFFIC SAFETY DATA ARCHIVE: PHASES 1 AND 2

Final Report

OTREC-RR-11-28

by

Christopher M. Monsere, Ph.D., P.E., Assistant Professor
Department of Civil & Environmental Engineering
Portland State University
Myenwoo Lim
Chengxin Dai
Xiaowei Wu

for

Oregon Transportation Research
and Education Consortium (OTREC)
P.O. Box 751
Portland, OR 97207

OTREC
OREGON TRANSPORTATION RESEARCH
AND EDUCATION CONSORTIUM

March 2012

Technical Report Documentation Page

1. Report No. OTREC-RR-11-28	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF THE OREGON TRAFFIC SAFETY DATA ARCHIVE: PHASE 1 AND 2		5. Report Date March 2012 6. Performing Organization Code
7. Author(s) Chris Monsere, Myeonwoo Lim, Chengxin Dai, Xiaowei Wu		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil & Environmental Engineering PO Box 751 Portland State University Portland, OR 97207		10. Work Unit No. (TRAIS) 11. Contract or Grant No. 2008-176, 2009-232
12. Sponsoring Agency Name and Address Oregon Transportation Research and Education Consortium (OTREC) P.O. Box 751 Portland, OR 97207		13. Type of Report and Period Covered 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This report describes the preliminary work to develop the Oregon Traffic Safety Data Archive (OrTSDA). The mission of OrTSDA is to "build the knowledge base of traffic safety data in Oregon". The archive hopes to become a valuable traffic safety data resource with researchers, practitioners, and community supporters. As envisioned, the archive will provide current and historical versions of related data, encompass complete documentation, and contain an online interface to the variety of traffic safety related data. Working with partner state agencies, the archive contains reported motor vehicle crashes statewide, population data, estimates of vehicle miles traveled, traffic crime, and information on driving under the influence measures. These data have been loaded in an online web-based archive for users to query and analyze. The archive produces safety performance metrics for statewide, county and city aggregations. Users can also access a set of pre-built queries which they can filter for crash type, month, time, and severity. Experienced users are given a one-step feature to obtain raw crash data. One key feature of the archive is a Google-map based tool that allows users to select crash data for analysis by drawing a selection boundary on the map. The report also includes a brief review of other online highway safety data portals. This review highlights that interactive crash portals are restricted to a handful of states agencies or research university partners.

17. Key Words Safety Data, Safety Analysis, Data Archive		18. Distribution Statement No restrictions. Copies available from OTREC: www.otrec.us	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 46	22. Price

ACKNOWLEDGEMENTS

Match and support for this research project was provided by the City of Portland, the Oregon Department of Transportation (Research and Traffic and Roadway Services), and Portland State University Provost's Foundation Faculty Development Awards. Robin Ness and Theresa Heyn at Oregon DOT's Crash Analysis and Reporting Unit have been instrumental in data assembly for the tool.

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

EXECUTIVE SUMMARY **1**
1.0 INTRODUCTION **3**
1.1 RESEARCH OBJECTIVES 4
1.2 REPORT ORGANIZATION 4
2.0 SURVEY OF SAFETY PORTALS **5**
2.1 FEDERAL 5
2.1.1 Fatality Analysis Reporting System (FARS) 5
2.1.2 SafeRoad Maps 7
2.2 STATE 7
2.2.1 Iowa 8
2.2.2 Kentucky 9
2.2.3 Michigan 10
2.2.4 New Mexico 11
2.2.5 North Carolina 12
2.2.6 Oregon 13
2.3 SUMMARY 14
3.0 ASSEMBLY OF DATA **15**
3.1 OREGON STATEWIDE CRASH DATA SYSTEM 15
3.1.1 Geo-location Information 15
3.1.2 Disclaimer 17
3.2 POPULATION 17
3.3 TRAFFIC VOLUMES 17
3.4 TRAFFIC CRIME 17
3.5 DUII MEASURES 17
4.0 ONLINE DATA ARCHIVE **19**
4.1 SOFTWARE 19
4.2 USER LOGIN 19
4.3 DASHBOARD 19
4.3.1 Crash Outcomes 20

4.3.2 Behavioral 21
4.3.3 Emergency Services 21
4.3.4 Emphasis Crash Outcomes 21
4.4 CRASH PATTERNS 24
4.5 GOOGLE MAPS TOOL 26
4.6 RAW DATA 29
4.6.1 Crash Data 29
4.6.2 GIS Archive 29
5.0 CONCLUSIONS 31
6.0 REFERENCES 33
APPENDIX A 35

LIST OF TABLES

Table 3.1 Crash Geo-Location Summary 16

LIST OF FIGURES

Figure 2.1 Screen capture of FARS cross-tab query built with query builder tool 6
Figure 2.2 Screen capture of SafeRoad Map output for Oregon, 2008 7
Figure 2.3 Screen capture of Iowa Traffic Safety Data Service 9
Figure 2.4 Screen capture of Kentucky Collision Analysis for the Public 10
Figure 2.5 Screen capture of Michigan TCF Data Query Tool 11
Figure 2.6 Screen capture of New Mexico's web-based GIS map of crash data..... 12
Figure 2.7 Screen capture of North Carolina query tool..... 13
Figure 2.8 Screen capture of Oregon TransViewer crash site 14
Figure 4.1 Sample index chart for crash severity showing user interaction 22
Figure 4.2 Screen capture of OrTSDA dashboard 23
Figure 4.3 Screen capture of crash severity time series. 24
Figure 4.4 Screen capture of result of crash patterns type and severity 25
Figure 4.5 Sample crash map from Google map 27
Figure 4.6 Step 1: Drawing the selection polygon..... 28
Figure 4.7 Step 2: Selection of crashes within enclosed polygon region 28
Figure 4.8 Screen capture of download raw data user interface 29
Figure 4.9 Screen capture of GIS archive 30

EXECUTIVE SUMMARY

This report describes the preliminary work to develop the Oregon Traffic Safety Data Archive (OrTSDA). The mission of OrTSDA is to "build the knowledge base of traffic safety data in Oregon." It is hoped the archive will become a valuable traffic safety data resource with researchers, practitioners and community supporters. As envisioned, the archive will provide current and historical versions of related data, encompass complete documentation, and contain an online interface to a variety of traffic safety-related data.

Working with partner state agencies, the archive contains reported motor vehicle crashes statewide, population data, estimates of vehicle miles traveled, traffic crime, and information on driving under the influence measures. These data have been loaded in an online, web-based archive for users to query and analyze. The archive produces safety performance metrics for statewide, county and city aggregations. Users can also access a set of pre-built queries which they can filter for crash type, month, time and severity. Experienced users are given a one-step feature to obtain raw crash data. One key feature of the archive is a Google map-based tool that allows users to select crash data for analysis by drawing a selection boundary on the map.

The report also includes a brief review of other online highway safety data portals. This review highlights that interactive crash portals are restricted to a handful of states agencies or research university partners. Concerns of liability most likely contribute to a reluctance to share collision data more broadly. Of those states providing direct access to the crash data, there is a common theme of maps, data filters and crash summaries.

1.0 INTRODUCTION

As early as 1924, the need for safety knowledge was recognized. At a national conference on street and highway safety, it was reported that safety data “are so vital to any comprehensive understanding and treatment of the safety problem that their collection and analysis in every State and community are essential" (Griffith et al, 2003). Many decades later, this simple statement has, in fact, taken on increased significance. In the information age, the challenge is now to make effective decisions based on a wealth of safety-related data. Hauer (1997, 2003 and 2005) makes many arguments that improvement in transportation safety requires a substantial advancement in the demand for knowledge-based applications. These applications require appropriate research methods, adequate data, and a demand for those with skills to put these two tools together.

Traffic records such as driver files, crash data, enforcement records, highway traffic and geometric information, court records, and emergency medical records are the common data needed to make effective safety-related decisions. Often these data are in various formats, maintained by distinct agencies, and require specialized knowledge to use and link together information to achieve maximum use of the data. In their comprehensive review of the U.S. crash records systems, Delucia and Scoptaz (2005) made a number of recommendations for a successful, useful, robust traffic records system. One of those recommendations was that development of a “knowledge base” of traffic records information is imperative to fully extract the expense of creating the individual data systems. They stressed, quite correctly, that while the benefits of integrating various data sources are substantial, the challenges and institutional limitations have prevented many states from vigorously pursuing this fruitful path.

Many agencies are being called on to concentrate on core business functions and the added resources to develop a knowledge base are difficult to obtain. There are number of states (e.g., Kentucky) that have found the resources to develop such a safety-related data clearinghouse (Delucia and Scoptaz, 2005). In other transportation data areas such as freeway management systems, successful data archives such as the Portland Oregon Regional Transportation Archive (PORTAL) and the Freeway Performance Measurement System (PeMS) have academic roots (Bertini, 2005). Other states (e.g., New Mexico, Wisconsin, Iowa and Michigan) have pursued a research university as the developer of the data archive and knowledge base.

In Oregon, the situation with traffic records is similar to many other states. While nearly all traffic safety-related data (enforcement, crash, highway, court records and emergency management) are available on request from the various agencies that maintain the data (Oregon Department of Transportation, Oregon Justice Department, U.S. Department of Transportation, Human Services Department), there is no coordinated clearinghouse to integrate these data or provide a central location where other interested researchers, private citizens and professionals can easily access the data in a processed, consistent and useable form.

This report summarizes the first steps towards the development of such a clearinghouse. Two OTREC-funded research projects are summarized in this report. Both of these phases were

focused primarily on tools and analysis using motor vehicle crash data. Phase 1 was partial seed funding; Phase 2 was the complete funding of the Phase 1 scope. A third phase, focused on enhanced visualization tools, is also funded.

1.1 RESEARCH OBJECTIVES

This research had three modest objectives:

1) *Crash data collection* – Assemble the most recent statewide crash data files as possible, over a long time series, to provide snapshots of system performance.

2) *Create an online data archive* – Load the data collected in an online, open-source, web database and create a web-based user interface. An export function will allow high-level users to receive disaggregate data for detailed analyses. Simple form-based queries to display results in graphs, tables or other combinations will be developed.

3) *Explore additional data for inclusion in the archive* - Gather additional linkable data sources from the primary sources (enforcement, highway, court records) to enhance a crash-based archive.

1.2 REPORT ORGANIZATION

This report is organized as follows: Chapter 2 presents a brief overview of other available online crash archives, and Chapter 3 describes the source of the data collected, how it was collected and how it was loaded in the archive. Chapter 4 is essentially a user's manual for the online archive – all the functionality of the site is described. Finally, the next steps for the archive are described.

2.0 SURVEY OF SAFETY PORTALS

While not intended to be a comprehensive review, we attempt to briefly summarize some of the available safety archives or data portals for traffic safety data. We focus on crash systems since these are the primary sources of traffic safety data. Due to the dynamic nature of the web-based material, these summaries should be considered snapshots that describe the overview of the data present and user interface tools in use. The data portals are organized by federal and state sources.

2.1 FEDERAL

2.1.1 Fatality Analysis Reporting System (FARS)

The FARS data system includes all vehicle crashes in the U.S. that occurred on a public roadway and involved a fatality. The data are collected by each state and forwarded to FARS in a consistent format in a long time series. The FARS query website[1] is a robust analysis tool that allows the user to query the main FARS data tables. The tool allows users to summarize fatal crash data from 1994 to 2009. Query-building tools allow the users to do uni-variate (one variable summary) and cross tabulations of most variables in the FARS data. A screen capture is shown in Figure 2.1 of a cross-tab query counting fatal crashes by hour of day and day of week for all states in 2009. The query tool, however, does not allow users to do multiyear queries. User can also obtain the raw data by purchasing a data CD or downloading from the Internet via FTP.

[1] http://www-fars.nhtsa.dot.gov/QueryTool/QuerySection/SelectYear.aspx

FARS FTP ▪ NASS GES FTP ▪

NHTSA www.nhtsa.gov National Highway Traffic Safety Administration
Our Mission: Save lives, prevent injuries, reduce vehicle-related crashes

NCSA DATA RESOURCE WEBSITE
fars FATALITY ANALYSIS REPORTING SYSTEM ENCYCLOPEDIA

Pubs/Data Requests | FARS Data Tables | Query FARS Data | State Traffic Safety Info | Help

Report:

STATE: All | YEAR: 2009 | COUNT: Number of Crashes

OUTPUT OPTIONS: EXPORT (TXT) View Cases

Crash Hour	Day of Week							Total
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
0:00am-0:59am	312	103	96	147	149	169	320	1,296
1:00am-1:59am	346	98	80	89	136	173	318	1,240
2:00am-2:59am	414	115	94	92	137	183	386	1,421
3:00am-3:59am	303	82	67	84	94	113	257	1,000
4:00am-4:59am	189	66	60	69	69	96	207	756
5:00am-5:59am	143	97	94	106	113	128	150	831
6:00am-6:59am	136	158	144	137	143	158	139	1,015
7:00am-7:59am	92	145	125	135	158	143	105	903
8:00am-8:59am	96	148	119	128	127	128	134	880
9:00am-9:59am	92	141	130	110	118	140	143	874
10:00am-10:59am	134	127	142	121	138	153	146	961
11:00am-11:59am	166	152	137	142	139	159	204	1,099
12:00pm-12:59pm	168	172	144	157	158	189	185	1,173
1:00pm-1:59pm	187	191	190	150	164	200	228	1,310
2:00pm-2:59pm	204	209	208	186	202	233	256	1,498
3:00pm-3:59pm	220	227	208	237	207	241	235	1,575
4:00pm-4:59pm	228	207	222	221	210	248	259	1,595
5:00pm-5:59pm	252	253	247	205	242	278	297	1,774
6:00pm-6:59pm	278	208	204	224	224	289	310	1,737
7:00pm-7:59pm	259	211	201	221	226	254	301	1,673
8:00pm-8:59pm	231	208	189	183	230	264	273	1,578
9:00pm-9:59pm	217	185	187	227	206	294	279	1,595
10:00pm-10:59pm	163	154	159	181	222	307	278	1,464
11:00pm-11:59pm	143	109	138	132	164	291	323	1,300
Unknown	31	31	30	40	28	36	53	249
TOTAL	5,004	3,797	3,615	3,724	4,004	4,867	5,786	30,797

Search Criteria:
Year 2009

NHTSA USA.gov NCSA fars

PRIVACY POLICY ▪ SITE MAP ▪ EMAIL NCSA

Figure 2.1 Screen capture of FARS cross-tab query built with query builder tool

2.1.2 SafeRoad Maps

Using FARS data, researchers at the University of Minnesota, Claremont Graduate University and the Center for Excellence in Rural Safety (CERS) have developed SafeRoad Maps[2]. The tool has a number of interesting features designed to allow users to view the location of fatal crashes in a Google maps platform. Various filters are available for the user to customize the type of crashes displayed. A hot-spot analysis of fatal crashes as well as a route-based selection of crashes is available. A screen capture is shown in Figure 2.2. The site has been mentioned in a number of press articles.

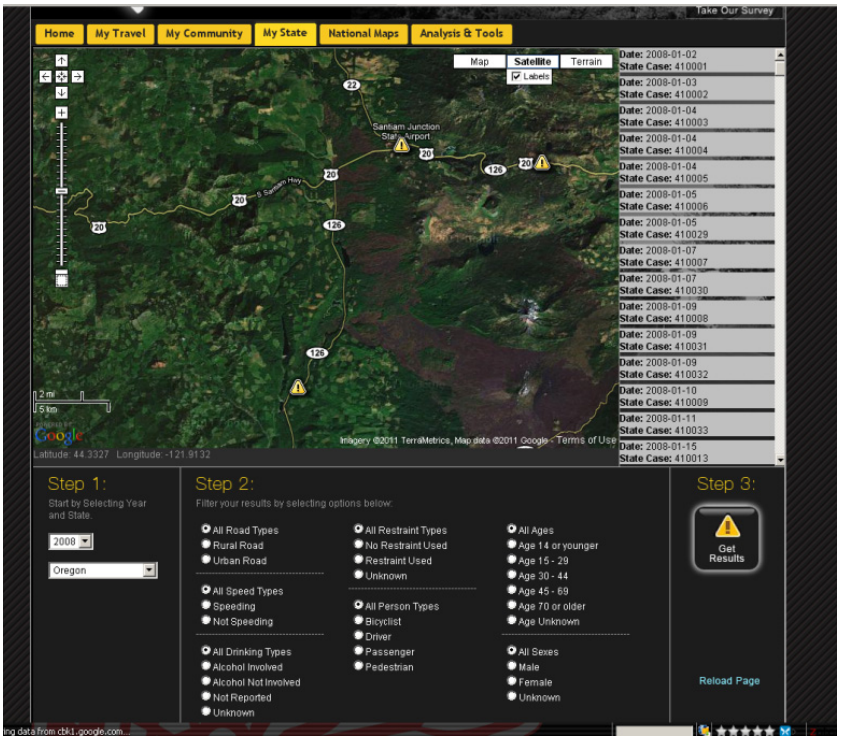


Figure 2.2 Screen capture of SafeRoad Map output for Oregon, 2008

2.2 STATE

The following present sample safety portals from a number of states. The sample is not comprehensive and given the dynamic nature of online offerings there may be additional sites not reviewed here.

[2] http://www.saferoadmaps.org/home/

2.2.1 Iowa

One example partnership between a DOT and research university is located in Iowa. The Iowa Department of Transportation website contains static crash profiles for Iowa counties/cities and a corresponding traffic safety manual, safety analysis and safety programs. A list of crash analysis resources and a list of safety improvement candidate locations are provided. Data is archived in the form of PDF format files and mainly available in the form of a map. There is no access to raw data.

However, the Iowa DOT has also partnered with CTRE to house the Iowa Traffic Safety Data Service. ITSDS is equipped to handle a wide variety of crash analysis requests, including crash histories, fatalities and/or injuries, alcohol-related crashes, seatbelt status, cross-median crashes, pedestrian crashes and weather conditions[3]. While the ITSDS is not an archive, it is a venue where users can request specialized analysis of crash data. The output is in the form of maps, diagrams or reports. Completed projects and publications (brochures, manuals and short reports) are archived and categorized based on the requesting agency or completion date. These projects can only be accessed after sending a request and getting approval. There is no access to raw data. Data is available in the form of maps, diagrams and reports. Upon request, a user can get a copy of a completed project similar to the one they are interested in or send a request for a new crash analysis. A screen capture of the ITDS website is shown in Figure 2.3.

[3] http://www.ctre.iastate.edu/itsds/index.htm

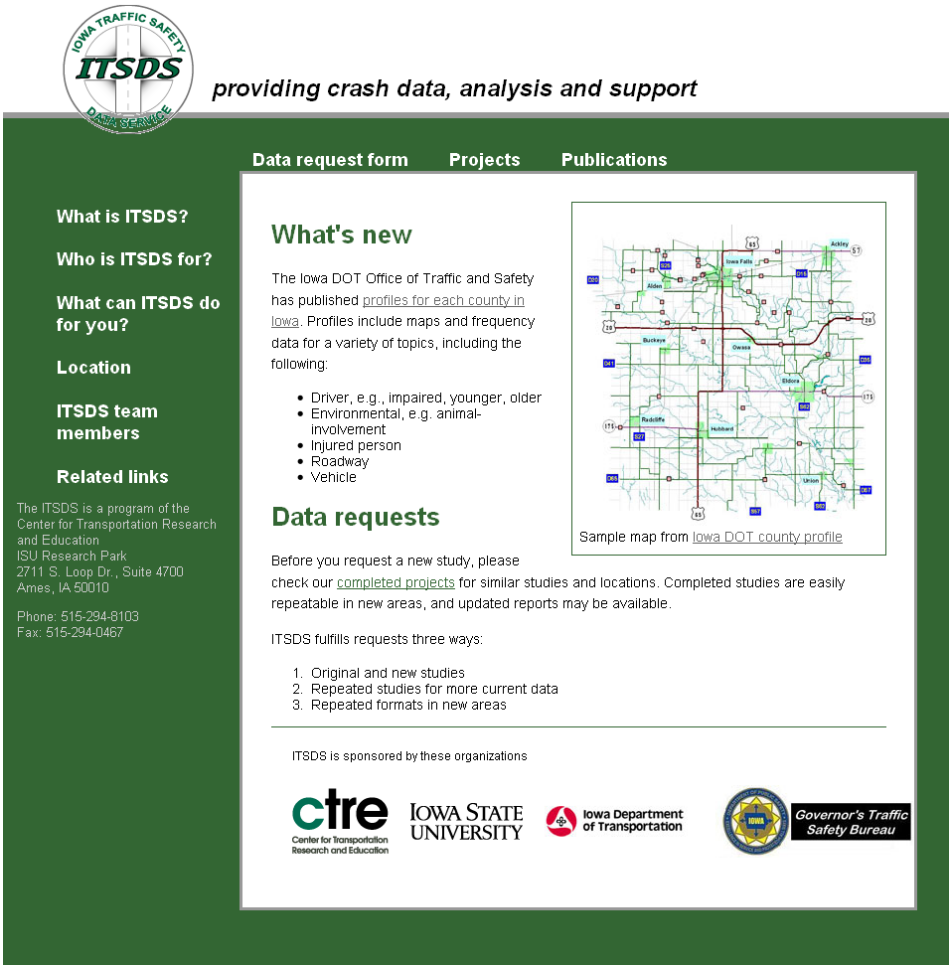


Figure 2.3 Screen capture of Iowa Traffic Safety Data Service

2.2.2 Kentucky

Kentucky Collision Analysis for the Public[4] website is developed and maintained by the Kentucky State Police. The website has the ability to query the database, generate maps and export data extracts. The portal has two main sections – dashboard analysis and collision data analysis. The dashboard analysis provides summary charts and information for a number of preset queries (crashes by light condition, weather condition, roadway conditions, etc). A screen capture of this dashboard analysis is illustrated in Figure 2.4. The site also includes the ability to build specific queries and generate raw data extracts of those as well as maps of those crashes. The query builder is flexible and allows the user to filter based on any searchable field in the Kentucky collision database.

[4] http://crashinformationky.org/KCAP/Public/Home.aspx

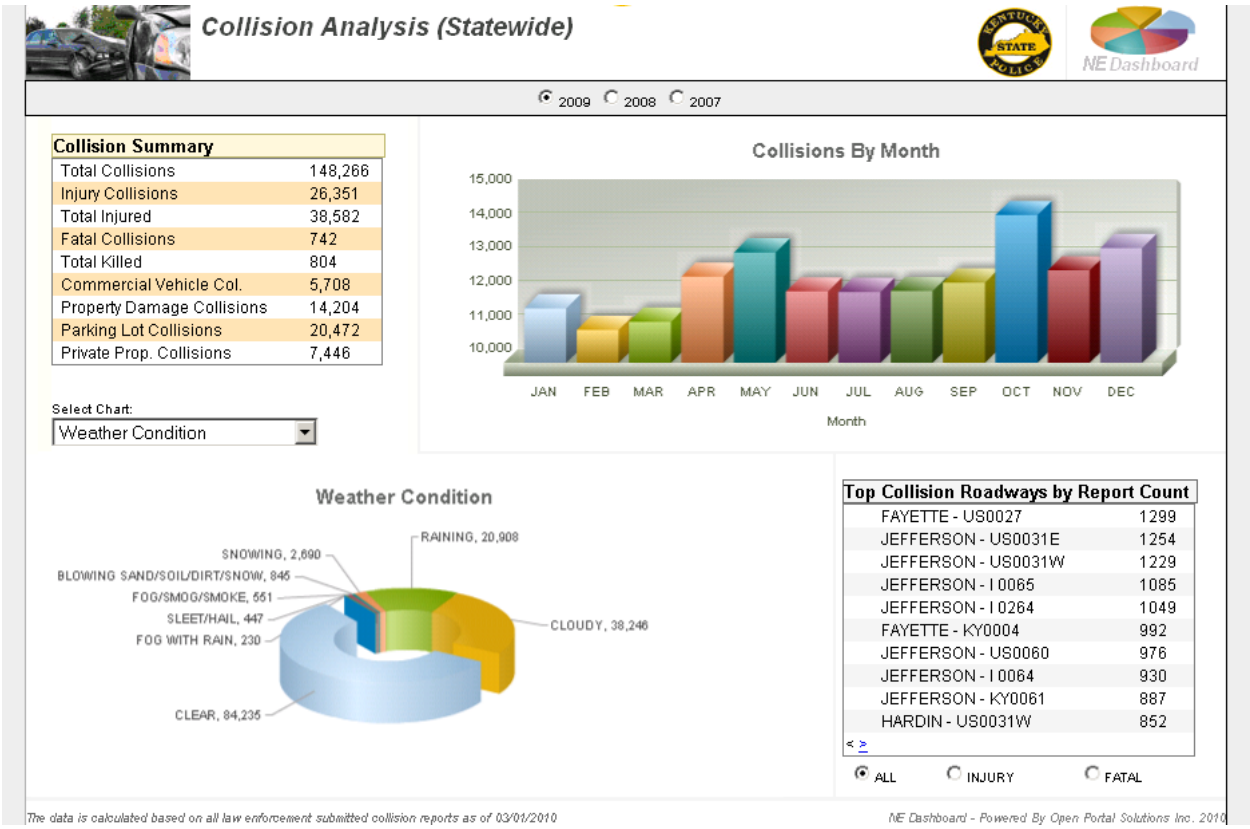


Figure 2.4 Screen capture of Kentucky Collision Analysis for the Public

2.2.3 Michigan

The Michigan Traffic Crash Facts (MTCF) site contains a statewide crash manual for historical information, and different crash causes and variables for specific years[5]. County and city data on traffic crashes, fatalities and injuries are listed in detail. Fact sheets are categorized under different categories, such as age and alcohol. The website contains a link to the "Data Query Tool," which allows user access to an interactive query builder and map. Like the Kentucky online archive, the query builder tool is very flexible and allows users to filter crash data across multiple years, geographic area and any filterable record in the database. The user also has the ability to save queries, as long as cookies are not removed in their browser's memory. A screen capture of the tool is shown in Figure 2.5.

[5] http://www.michigantrafficcrashfacts.org/

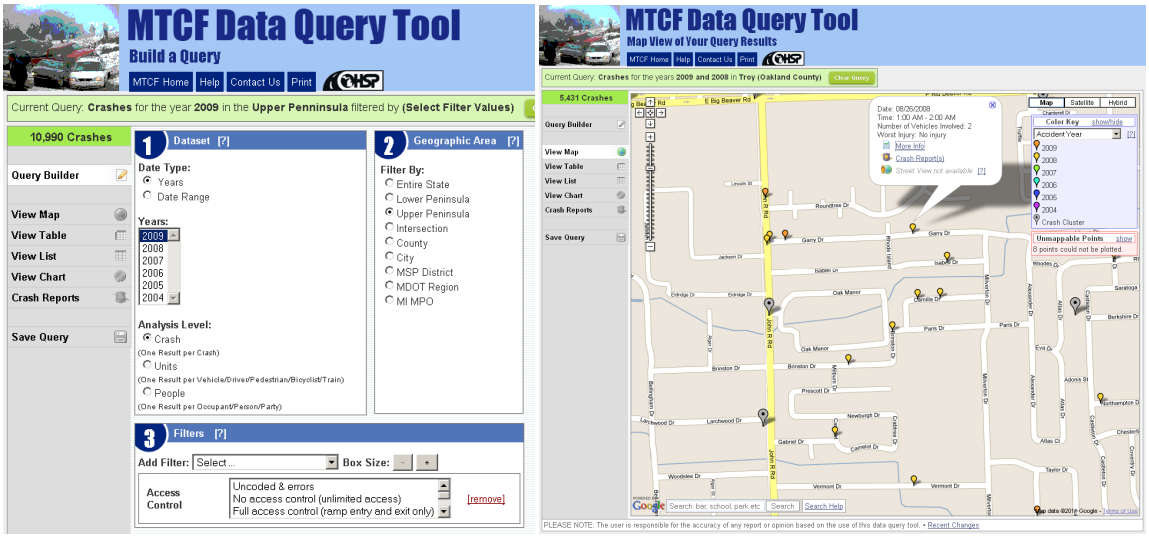


Figure 2.5 Screen capture of Michigan TCF Data Query Tool

2.2.4 New Mexico

A series of comprehensive traffic crash data files of the state of New Mexico is maintained by the Division of Government Research (DGR)[6]. Annual crash reports, community reports and DWI (Driving While Impaired) reports are available in PDF format from 1995-2006. Data files include accident-level analysis files, occupant-level analysis files, detail-level analysis files and DWI-citation tracking files. Reports consist of tables, graphs and charts classified by districts in PDF format. Maps of traffic crash data (state, regional, county and urban) are in both PDF and Geographic Information System (GIS) formats. Traffic safety-related websites are listed.

Crash data files are maintained in Statistical Analysis System (SAS) format. These files can be converted into other formats such as dBase (DBF) and Arc/INFO-ArcView. Traffic crash maps can be viewed by selecting from a list of state, regional and county maps or from a table of urban maps. Raw data is not accessible without a request. Data are all displayed in the form of summaries (tables, graphs and maps). Cooperative data-use agreements can be arranged so that local government agencies can obtain this data. User can obtain crash data reports and access crash manuals for 1995-2006. These files could be in PDF format or as a GIS map for the location chosen by a user. A sample web-based display of crashes is shown in Figure 2.6.

[6] http://www.unm.edu/~dgrint/tcd.html

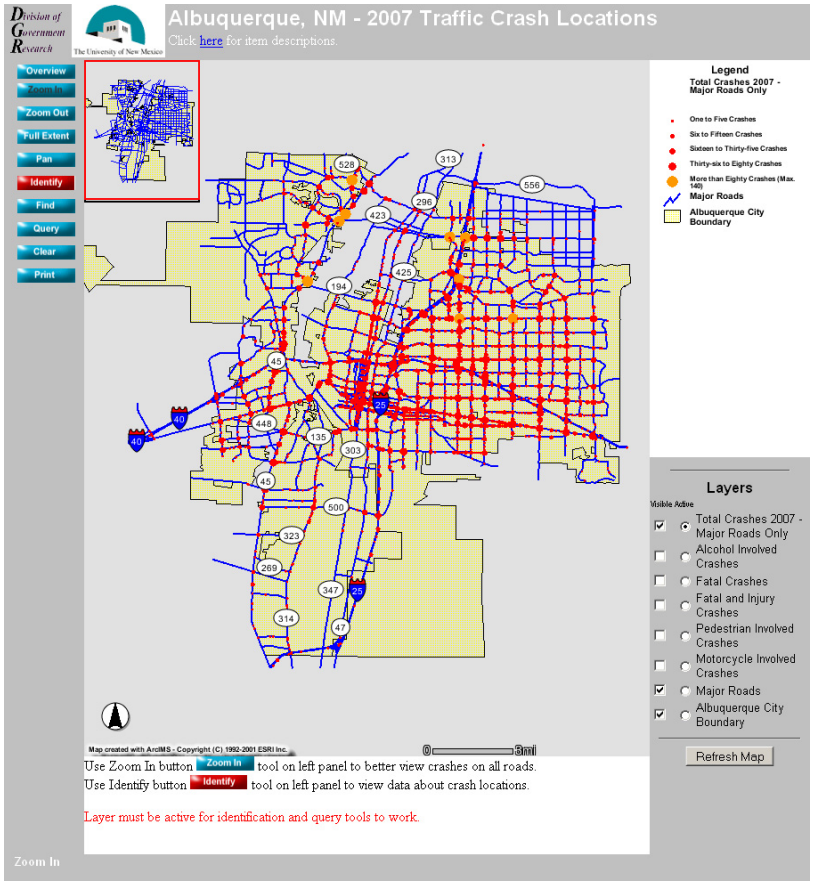


Figure 2.6 Screen capture of New Mexico's web-based GIS map of crash data

2.2.5 North Carolina

A data analysis tool in the site enables the creation of tables reflecting crashes in North Carolina from 2001-2009[7]. A screen capture of the query builder tool is shown in Figure 2.7. Tables reflect data for the whole state or a specific city, county, highway or division. Also, there is help for a way to use the form and variables explanation. Data is archived through a static copy of 2001-2009 NCDOT live-crash databases and is available in the form of a summary table, depending on the query variables. There is no access to raw data over the Internet. In this form of table output, users find reportable/non-reportable crashes for 2001-2006, worst injury, causes of crashes (age, alcohol, speeding, etc.), and crash data regarding certain variables (weather, work zone, etc.). Users can also develop general descriptions of crashes, including problem identification.

[7] http://buffy.hsrc.unc.edu/crash/

Figure 2.7 Screen capture of North Carolina query tool

2.2.6 Oregon

The Oregon DOT provides external users access to a variety of standard crash reports and summaries, both on state highways and local roads. Users are provided the data in either a neatly formatted PDF report or in Excel format that matches internal DOT formatted reports. The Excel data tables can be fed into other tools created by ODOT for analysis. Rather recently, the ability to query local roads has been added to the tool. A screen capture of the user interface for state highway crash reports is shown in Figure 2.8. The reports are provided through an application

gateway software tool (https://keiko36.odot.state.or.us/). The crash data system reports do not work with the Mozilla browsers.

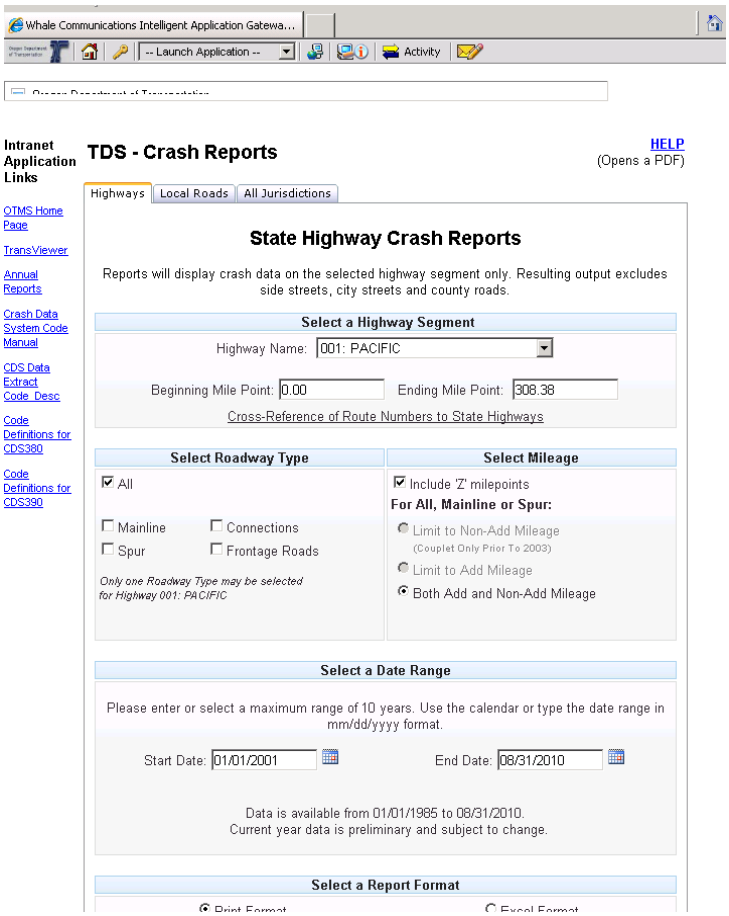


Figure 2.8 Screen capture of Oregon TransViewer crash site

2.3 SUMMARY

From our review, it appears that interactive crash portals are restricted to a handful of states, agencies or research university partners. There are likely more data portals in development than are captured in this review. Liability issues most likely contribute to a reluctance to share collision data. The crash data archives and portals offer various ideas and tools to convey safety data. Some of the sites allow users to customize summary queries and tabulations (an advancement on the static crash summaries typically produced by crash reporting agencies) while others include map-based formats for interactive selection of crash data. A handful of sites provided access to the raw data files for more advanced users to analyze in detail. None of the archives appeared to merge multiple traffic or highway safety information in one location. While not reviewed here, similar data archives and map tools (http://www.crimereports.com/) are available for other public safety data such as crime mapping. In the next chapter, the data assembled for display in OrTSDA are described.

3.0 ASSEMBLY OF DATA

This chapter describes the data that has been collected and assembled for the archive. For the most part, we have compiled data from 1995-2007. At a minimum, data can be spatially aggregated at the county level; however, many elements can be aggregated to the city level. Only the crash data has a latitude-longitude location which allows grouping at any spatial level.

3.1 OREGON STATEWIDE CRASH DATA SYSTEM

Crash records were obtained from ODOT in a Microsoft Access file format for 1995-2007. There were nearly 622,774 crash records on state highways during that time period. Two units within ODOT have responsibility and oversight for crash reporting, the Driver and Motor Vehicles (DMV) Services Division and the Crash Analysis and Reporting (CAR) Unit. Currently, private citizens are require to file an Oregon Traffic Accident and Insurance Report within 72 hours if they are involved in a crash that results in injury, death, more than $1,500 damage to their vehicle, or more than $1,500 damage and towing of another vehicle. These reporting thresholds changed in 1998 from $500 to $1,000 and in 2004 from $1,000 to $1,500. If a police officer responds to the scene, he or she completes the Oregon Police Traffic Crash Report, which is more detailed than the citizen report. A citizen must file a report even if a police officer is present and completes a form. In general, however, nearly 70 percent of the crash data that is entered in the statewide file for crash records comes from citizen reports. The state Crash Data System (CDS) is a relational database that contains three primary tables to describe a crash:

- *Crash table* - contains summary information about the crash, including road condition, location, time of day, day of week and other variables. There is one record in the crash table for each crash event (622,774 records).
- *Vehicle table* - contains data about each vehicle in the crash, including information about movements of vehicles, possible vehicle-related errors (mechanical failures), actions of the vehicle during the crash, and any objects hit by the vehicle causing injury or property damage. If more than one vehicle is involved there is one record for each vehicle (1,153,638 records).
- *Participant table* - includes data about each participant who was present at the crash (unless data are not presented or recorded about them on the report). The participant table includes the sex of each driver, their residence status, driver's license status, injury level, data about the use of safety equipment, and other participant-related data. All tables are related by a unique number assigned to each crash (1,332,907 records).

3.1.1 Geo-location Information

Prior to 2007, crashes were located in the state crash file with three different methods based on jurisdictional roadway:

- State highways – by a route (internal highway number) and milepost

- County roads – by a route milepost, if the county had mile-posted roads
- City streets – by distance from an intersection (defined by a unique street number)

These location methods (with the exception of state highways) do not lend themselves to easy-display spatial maps.

However, in 2007 the state started using a new tool in the coding process that allowed it to include latitude and longitude information in the crash data. The new process has not been perfect and considerable delays were introduced in the release of the geo-coded data. For example, as of this report 2009 data is still currently undergoing quality control processing nearly one year after release of the non-geocoded quality checked data. The process is improving, however, and the state expects much faster production of the 2010 data.

Two additional sources of geo-coding information exist for crashes prior to 2007. All crashes that occur on state highways can be spatially located using the state's linear-referenced highway network file. Information can be produced by mapping these files, then exporting the latitude-longitude from the GIS software location. In addition, the City of Portland has developed an internal process to map all crashes within city limits. These data were shared and included in the archive. Table 3.1 shows a summary of the geo-located crashes included in the archive. Note that there are a portion of crashes (because of incomplete crash reports) that are not locatable. The table shows that prior to 2007, the OrTSDA archive provides lat-long locations for about 40 percent of crashes statewide. In 2007, nearly 90 percent of crashes can be mapped and improvement is expected for data after 2008.

Table 3.1 Crash Geo-Location Summary

Year	Total Reported Crashes	Unknown Location Crashes	Crashes w/ lat-long	Percent Mapped
1995	49055	1482	18950	39.8%
1996	53251	1906	20843	40.6%
1997	49595	1911	19377	40.6%
1998	51784	2315	20260	41.0%
1999	48570	2484	18923	41.1%
2000	46841	2721	19614	44.5%
2001	48138	2402	20571	45.0%
2002	48282	1523	20218	43.2%
2003	51707	1074	19641	38.8%
2004	41440	728	16346	40.2%
2005	44878	1104	17565	40.1%
2006	45071	1117	17222	39.2%
2007	44162	1487	37770	88.5%

3.1.2 Disclaimer

The data provided in the OrTSDA archive is generated from the annual snapshot of data provided by ODOT. Internally, however, continuous improvement is being made to the data. As such, the OrTSDA data snapshot will not match the state's record exactly. To make it clear to the user that there is another source of more recent data, the following disclaimer is posted on all pages:

> *Disclaimer: All crash data provided from the Oregon Department of Transportation's Crash Analysis Reporting Unit's annual snapshot. Data provided here may not reflect corrections made after the annual release.*

3.2 POPULATION

City and county population records were obtained from the Population Research Center (PRC), College of Urban and Public Affairs, Portland State University. The PRC serves as the Oregon State Data Center for the census and is the official provider of population estimates in Oregon. These data are used to generate exposure measures and available from 1995 to 2007.

3.3 TRAFFIC VOLUMES

Estimated vehicle miles traveled (VMT) data for state highways was obtained from the Oregon DOT Traffic Monitoring Group for 1995 to 2007. Summarized measures of travel exposure are only available for state highways and aggregations are provided for by county but not by city. Future efforts will seek to develop estimates at the city level using spatial analysis of recent efforts to develop a statewide inventory of ADTs counts on all functionally classified roads in the state.

3.4 TRAFFIC CRIME

The source of traffic crime measures is information provided by annual reports published by the Law Enforcement Data Systems (LEDS) Section using data collected through the Oregon Uniform Crime Reporting (OUCR) Program by state law enforcement agencies. From the annual reports, we have extracted statewide regulatory offenses - traffic crimes as well as fish and game and marine violations. These offenses include misdemeanor and felony hit-and-run, reckless driving, eluding, misdemeanor and felony driving while suspended, and failure to display a driver's license. These data are provided for each county but not each city. Location information for these offenses is not available. New electronic citation data is capturing this information, but it is not yet available in a statewide or standard format.

3.5 DUII MEASURES

DUII statistics were obtained from the DUII Data Book for Oregon counties that has been compiled by ODOT staff in the Transportation Safety Division (ODOT-TSD). The data book contains a wealth of measures related to DUII activities in the state. A selected number of these measures, not captured in the state reported crash database, are extracted from the DUII fact book. These are:

- Enforcement Activity: Total DUII offenses, DUII arrests (drugs other than alcohol), 3. Breath test refused, DWS/DWR (misdemeanor) offenses, DWS/DWR (felony) offenses),
- Implied Consent: Suspensions issued.
- Diversion Program: Diversion agreements, Diversion enrollments, Education, Diversion enrollments, Rehabilitation.
- DUII Convictions: All convictions, First DUII, Second DUII, Third or subsequent DUII, Convicted DUII enrollments, Education, Convicted DUII enrollments, Rehabilitation.
- Additional Court Activity: DWS/DWR (violation) convictions, DWS/DWR (misdemeanor) convictions, DWS/DWR (felony) convictions, DUII (felony) convictions).

4.0 ONLINE DATA ARCHIVE

This chapter describes the various features that are currently available in the Oregon Transportation Safety Data Archive (OrTSDA), which can be accessed at http://web.cecs.pdx.edu/~ortsda. Because the archive is in a constant state of development, readers who visit the archive should expect that features and text described in this report may have changed.

4.1 SOFTWARE

OrTSDA uses a MySQL database backend maintained by the PSU MCECS Computer Action Team (CAT) to store and retrieve data files. The database currently consists of 73 tables (many are look-up tables for the crash database). The website interface uses PHP and JAVA scripting language. Presently, we have chosen to use Google Maps for the spatial representation of data. Our initial web graphs were produced using *jpgraph*. This tool lacked the ability for interactive display and high-quality data visualization tools. Newer developments are being produced with the *Protovis*[8] package (developed by the Stanford Visualization Group) which uses JavaScript and SVG for web-native visualizations. SVG is for modern browsers and there are known browser compatibility issues with versions of Internet Explorer below 9.0.

4.2 USER LOGIN

We are requiring all users to create a user name and password to login to the OrTSDA database. There is no approval process and the access is instant, but at the development stage it is important to log the users who have requested an account. Once logged in, the user is taken to a main page (dashboard) and a navigation menu that contains the following main choices:

- Dashboard
- Crash Patterns
- Google Map Tool
- Download Raw Data

Each of these submenus is described in the following sections.

4.3 DASHBOARD

The main page of the archive is a "dashboard" that displays time-series trends of key data elements that we have identified. An interactive dashboard has been created based on these recommendations and data availability. The default measures are produced statewide, but the

[8] http://vis.stanford.edu/protovis/

users can select any city or county geographic aggregation (though some measures are not available at the city level).

A good set of performance measures helps governmental agencies identify problematic areas and predict future trends. One of the main characteristics of an indicator is that it can be influenced by policy measures (Hermans et al., 2009). To decide which measures to produce on the dashboard, we briefly reviewed traffic safety performance measures for motor vehicle safety, pedestrian safety and bicyclist safety. The National Highway Traffic Safety Administration (NHTSA) and the Governors Highway Safety Association (GHSA) have agreed on a minimum set of performance measures to be used by state and federal agencies in the development and implementation of behavioral highway safety plans and programs (Hedlund, 2008). The initial minimum set contains 14 measures (10 core outcome measures, one core behavior measure and three activity measures). Primarily based on these recommendations and others, the following 15 measures are shown on the dashboard (which can be presented at the state, county or city level (if data exists)).

4.3.1 Crash Outcomes

Crash Severity

Count of all motor vehicle crashes with 1) fatal and 2) injury type A reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Fatalities per 10,000 Population

A common measure, the fatality rate per 10,000 population is the count of annual fatalities reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, divided by the population, aggregated by year.

Fatalities by Age Group

Count of fatalities by age groups (1-15, 16-21, 22-34, 35-54, 55-64, 65-74, 75-99) reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Motorcyclist Fatalities

Count of all fatal motorcyclist-involved crashes reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Pedestrian Fatalities

Count of all fatal pedestrian-involved crashes reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Cyclist Fatalities

Count of all fatal cyclist-involved crashes reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

4.3.2 Behavioral

Traffic Crime

Count of total traffic crimes reported to the Oregon State Police as part of the OUCR. The traffic crime measures the total behavioral traffic crimes each year. Traffic crime data does not include any spatial location data other than county or city of infraction.

Occupants with Fatal or Injury A Severity Coded for No or Improper Seat-Belt Use:

Count of participants (occupants) with fatal or injury A severity level who were reported to have either no seat-belt use or improper seat-belt usage reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Total DUII Offenses

Count of DUII offenses compiled by the Oregon Transportation Safety Division for the DUII fact book. The source of these offenses is the Law Enforcement Data System and includes juvenile offenses. If an arrest includes higher charges than a DUI, it will not be recorded as a DUII. Offenses, therefore, captures all arrested for DUII.

DUII Convictions

Count of DUII conviction by three groups (first DUII, second DUII and third or subsequent DUII) compiled by the Oregon Transportation Safety Division for the DUII fact book. These convictions can be reported at the county level (adjudicated by the circuit courts).

4.3.3 Emergency Services

Out-of-Hospital Time

The standard deviation of the out-of-hospital time for reported fatal crashes calculated from FARS data. Time attributes of the emergency call time, ambulance arrival time and hospital arrival time were extracted from the FARS database, converted into minute units, and filtered out unknown and illegitimate times. Future efforts will seek to gather data from Oregon's EMS system, which does not routinely publish data on response times.

4.3.4 Emphasis Crash Outcomes

Heavy Vehicle Crashes

Count of all crashes involving a truck reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Run-off-the-Road Crashes

Count of all crashes coded as fixed-object or non-collision reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

Head-On Crashes

Count of all collision types coded as head-on reported to the Oregon DOT, Crash Analysis and Reporting, Statewide Crash Data System and archived in OrTSDA, aggregated by year.

A screenshot of the dashboard is shown in Figure 4.2. Rather than display count data, each of the performance metrics are normalized by displaying an index calculated as:

$$Index = value\ in\ year\ /\ value\ in\ reference\ year \qquad (1)$$

A horizontal line is drawn at 1.0. Indices above 1.0 indicate an increase from the reference year, and the values below 1.0 indicate a decrease in from the reference year. The graphs include a dynamic feature where the user can change the reference year by mousing over the graph. The index is recalculated relative to that reference year. While the trendline does not change, the user can easily see the change relative to the base year. In Figure 4.2, the left plot shows the user moused over 1997, the right plot shows the user moused over 2001. The index plot adjusts to 1.0 for the selected year; thus, the user can quickly see the crash trend relative to the selected year. For example, in the left figure, the count of injury A crashes in 2007 was nearly the same as 1997 (it is not visible in the static snapshot, but if the user mouses over the *x*-axis they will know that the chart is from 1995-2007). However, in the right panel it can be seen that injury A crashes are about 10 percent higher in 2007 than 2001.

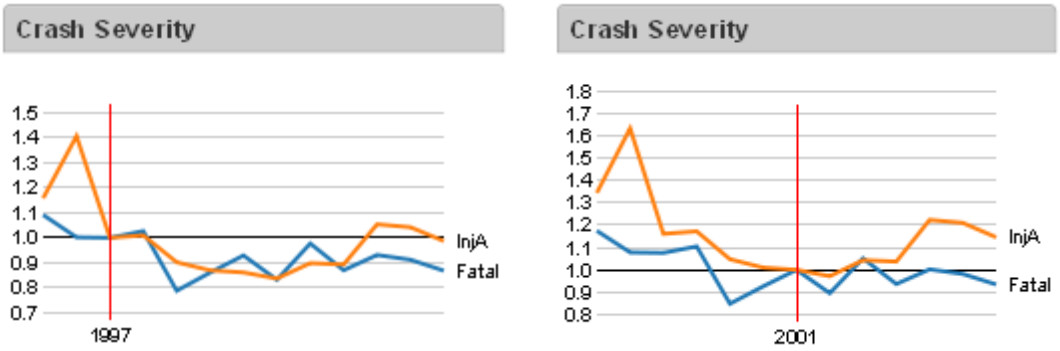


Figure 4.1 Sample index chart for crash severity showing user interaction

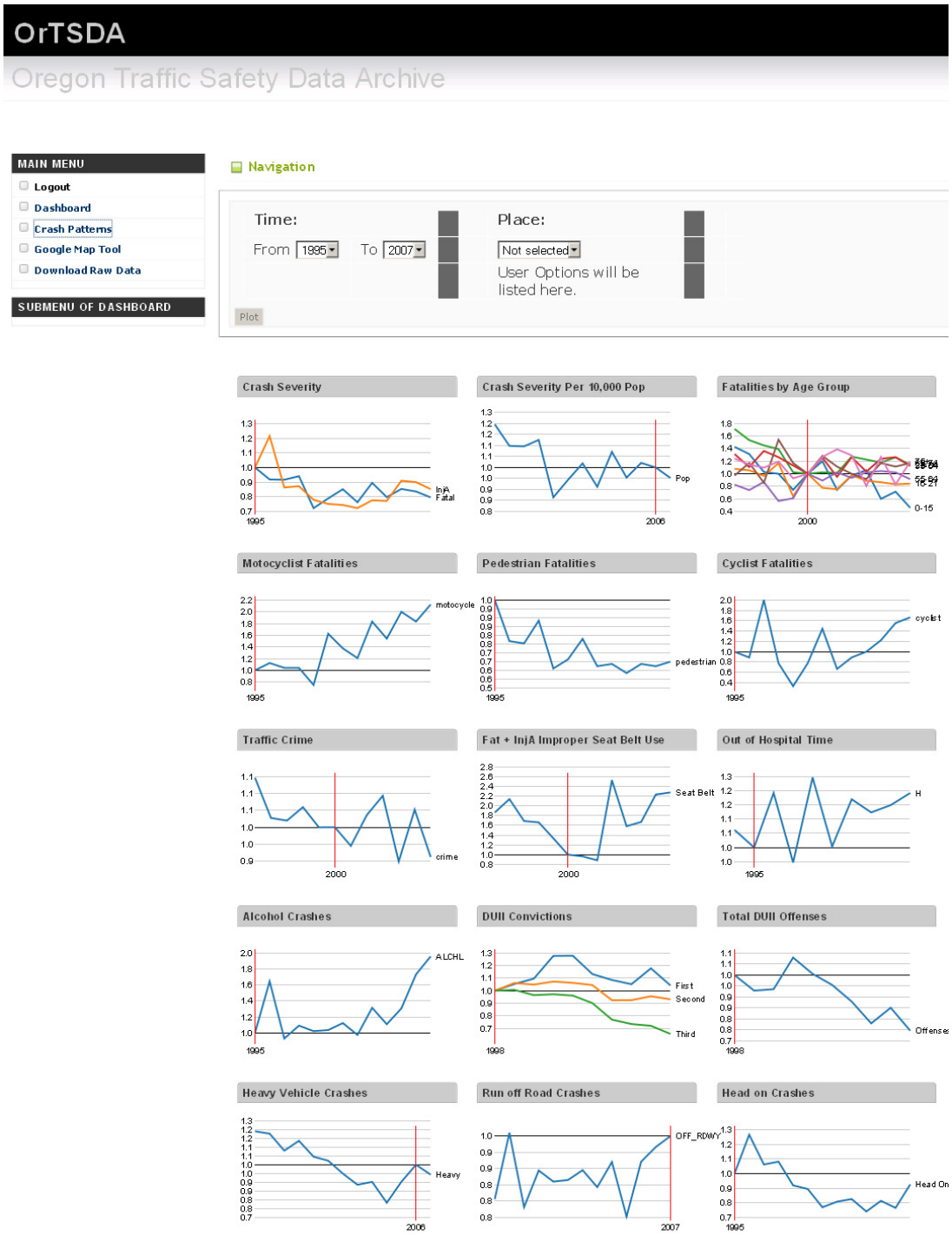


Figure 4.2 Screen capture of OrTSDA dashboard

4.4 CRASH PATTERNS

The crash patterns section of the archive is a set of pre-built data queries that users can filter by year, month and geographic area. Once accessed from the main menu, a submenu appears that lists the various patterns that can be summarized. At the time of this report, we have not yet migrated the graphics to *Protovis,* so the screen captures are in the older, less dynamic *jpgraph* tool. In Figure 4.3, a screen capture is shown of a time-series plot of all crashes by severity for the city of Beaverton that is accessed in the "Crash by Severity" link in the submenu. In Figure 4.4, a screen capture is shown of the "Crash Table" link. Users can summarize crashes by type and severity for any year, month, and county or city geographic features. In future revisions, we will include a query-type builder, such as demonstrated on the FARS, Kentucky, North Carolina and other crash portals.

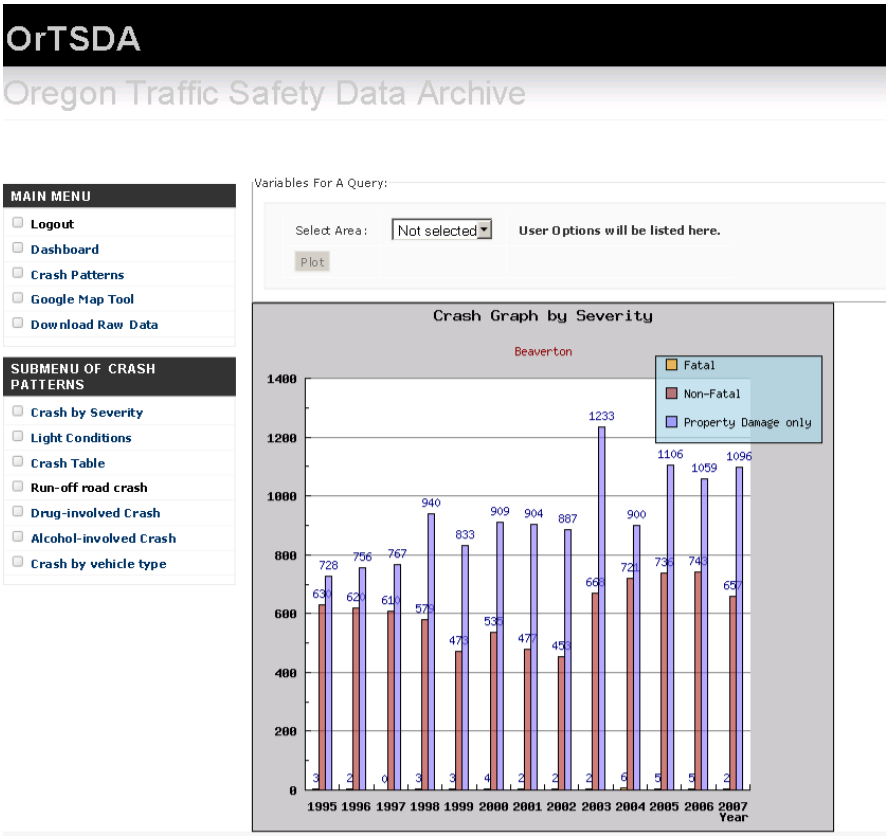
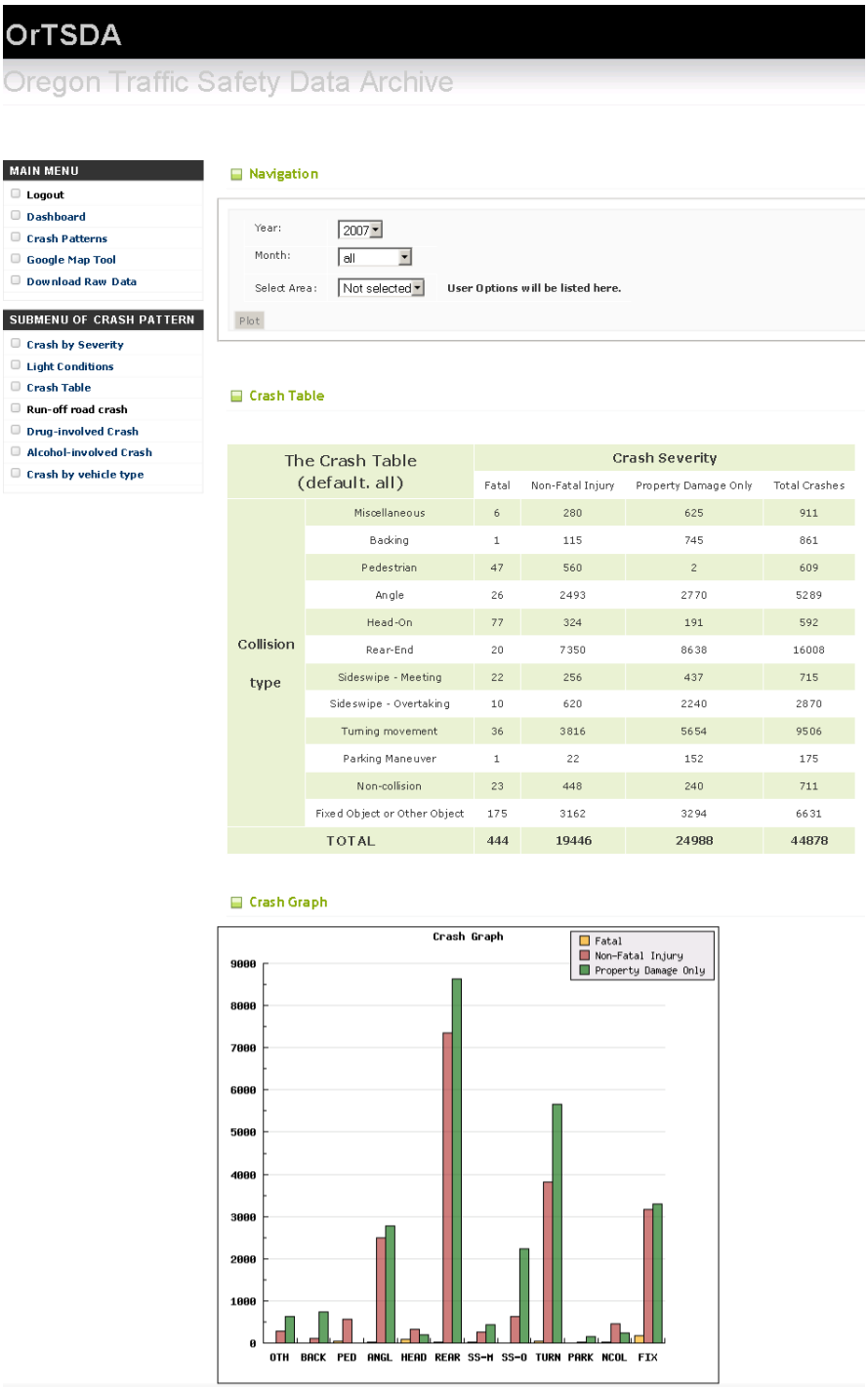


Figure 4.3 Screen capture of crash severity time series.



The Crash Table (default. all)		Crash Severity			
		Fatal	Non-Fatal Injury	Property Damage Only	Total Crashes
Collision type	Miscellaneous	6	280	625	911
	Backing	1	115	745	861
	Pedestrian	47	560	2	609
	Angle	26	2493	2770	5289
	Head-On	77	324	191	592
	Rear-End	20	7350	8638	16008
	Sideswipe - Meeting	22	256	437	715
	Sideswipe - Overtaking	10	620	2240	2870
	Turning movement	36	3816	5654	9506
	Parking Maneuver	1	22	152	175
	Non-collision	23	448	240	711
	Fixed Object or Other Object	175	3162	3294	6631
TOTAL		444	19446	24988	44878

Figure 4.4 Screen capture of result of crash patterns type and severity

4.5 GOOGLE MAPS TOOL

One of the primary challenges with accessing and analyzing crash data is the ability to select the records of interest. Most users will be concerned with crashes that have happened at a particular location or intersection. Clearly, a map-based interface makes selection of specific crash records an easy task for most users.

We have developed a Google map-based tool in Oregon Traffic Safety Data Archive (OrTSDA) to select crashes for further analysis. A Google maps interface has many appealing features – primarily the interface's familiarity to a wide variety of users, integration with satellite photos and street-level imagery. One of the main limitations, however, is a restriction on the number of pins that can be placed on a map view at one time. To solve this issue, we have used cluster tools that group crashes together at large zoom scales to limit the number of pins placed (the clusters are shown as blue pins in Figure 4.5). As the user zooms in, the cluster separates out the crash pins so that individual crashes can be seen. These crashes are shown as dots that are color coded based on severity (●Fatal Crash, ●Non-Fatal Injury, ●Property Damage Only) and can also be seen in Figure 4.5. Appendix A documents our exploration of various cluster tool performance.

Another approach that we employed to work within the pin limitation is to first allow the user to filter the crash type that they are interested in exploring. Thus, we have created two versions of maps that are provided for different user demands. Map 1 (Single Year) provides an option to display only one crash-year data. Map v.2 (Multi-year) allows the user to filter the selection by crash type, crash severity and multiyear time slice. Since most users are interested only in a small geographic area, the user can also filter by city or county to speed plotting of the necessary crash data.

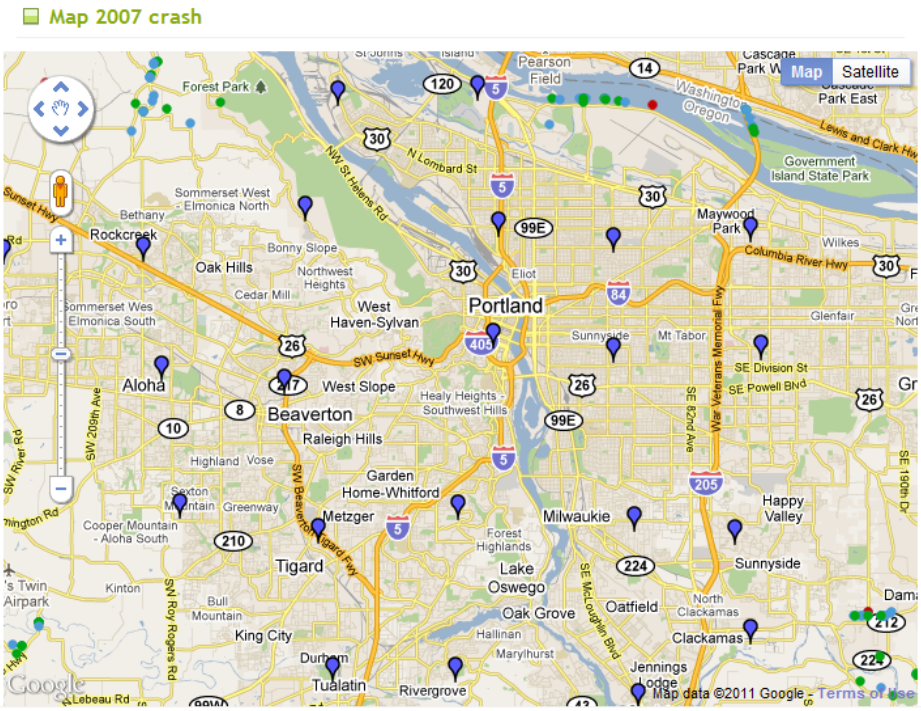


Figure 4.5 Sample crash map from Google map

The key addition to our Google Maps tool is a function that allows users to extract crash data from any irregular geographic region within Oregon by simply drawing a box on the web browser. This function works in both the single-year and multiyear maps. After selecting and plotting the crash data, when the user clicks anywhere in the map window a red pin is created (Figure 4.6). The selection is completed by drawing a boundary around the desired crashes. The polygon vertices cannot be moved once the user has clicked, but the "Delete Polygon" button will delete the drawn polygon. When the user has completed the polygon, the red pins disappear and the region turns transparent red. To extract the data within the polygon, the user only needs to select the "Extract Data in Polygon" button. The user is then taken to a page that contains summary graphs and the ability to download the CSV files of the selected crashes.

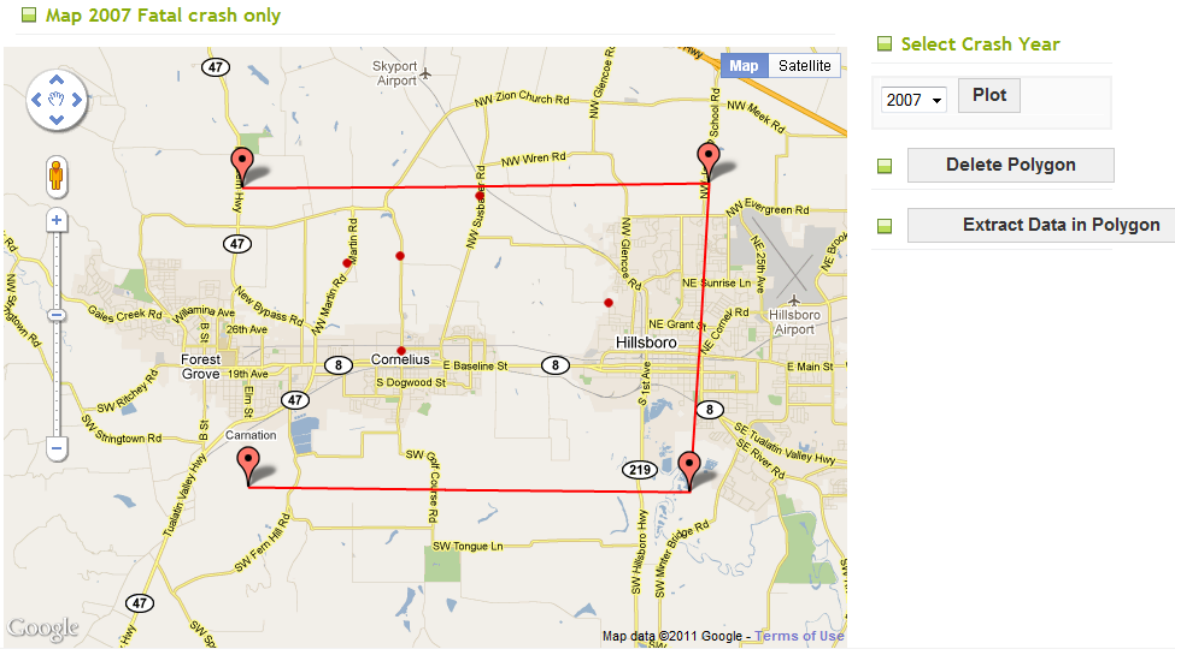


Figure 4.6 Step 1: Drawing the selection polygon

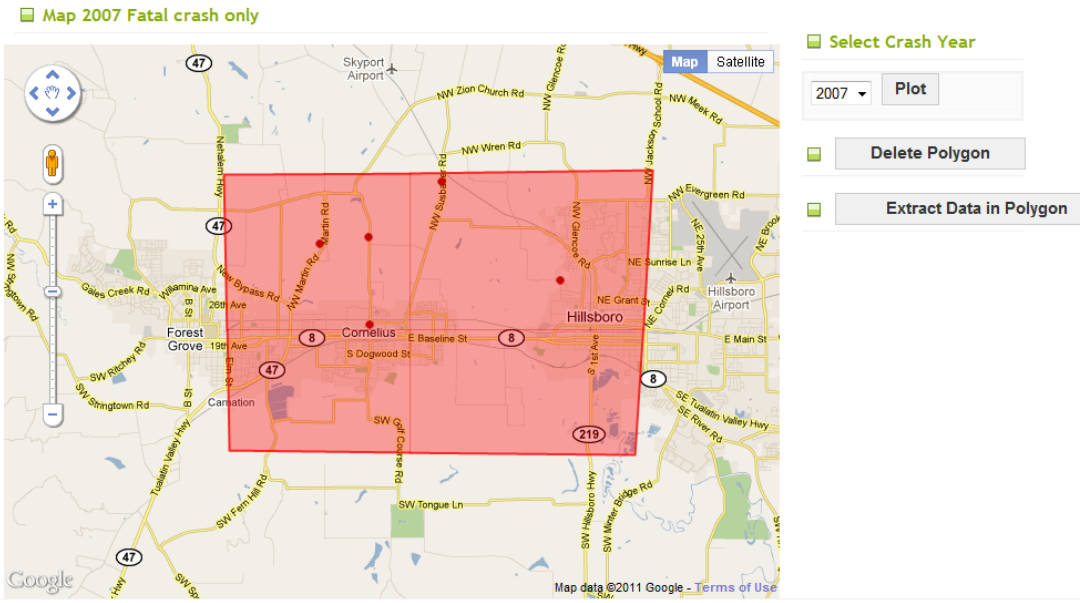


Figure 4.7 Step 2: Selection of crashes within enclosed polygon region

4.6 RAW DATA

Raw data is accessible from both the Google Maps tool (as described previously) and the "download raw data" menu. Raw data for download are archived in GIS files as shape format and tables as CSV format.

4.6.1 Crash Data

The user is able to download the raw data in a CSV file format based on a set of filters. The filter requires the user to specify a spatial area (statewide, county or city), crash type (all, fatal, injury, pedestrian, bicycle or pedestrian) and a date range. After selecting the filters, the user clicks the submit button and CSV tables for the crash level, participant level and vehicle level are provided in the "output" section of the web page. As a reminder, the three filters applied are shown in the submitted parameters part of the web page. A screen capture of the interface is shown in Figure 4.8.

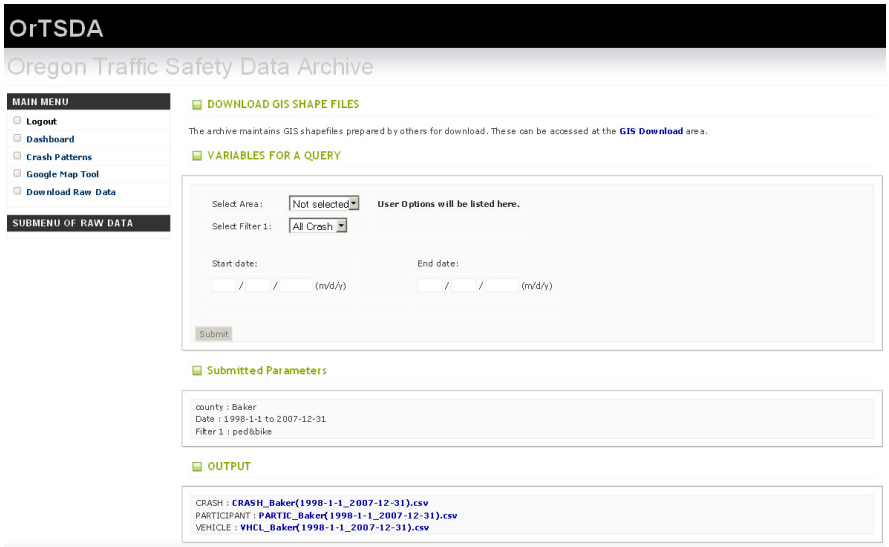


Figure 4.8 Screen capture of download raw data user interface

4.6.2 GIS Archive

In the archive's development, a number of GIS files have been obtained and produced by others. We have elected to post these files with a simple description, type of projection and the date posted in the download window. Presently, the GIS files consist of Portland city crash data, statewide data and tri-county geo-coded crash data from 2007-2008, which are available as compressed files in ZIP format and ready for use.

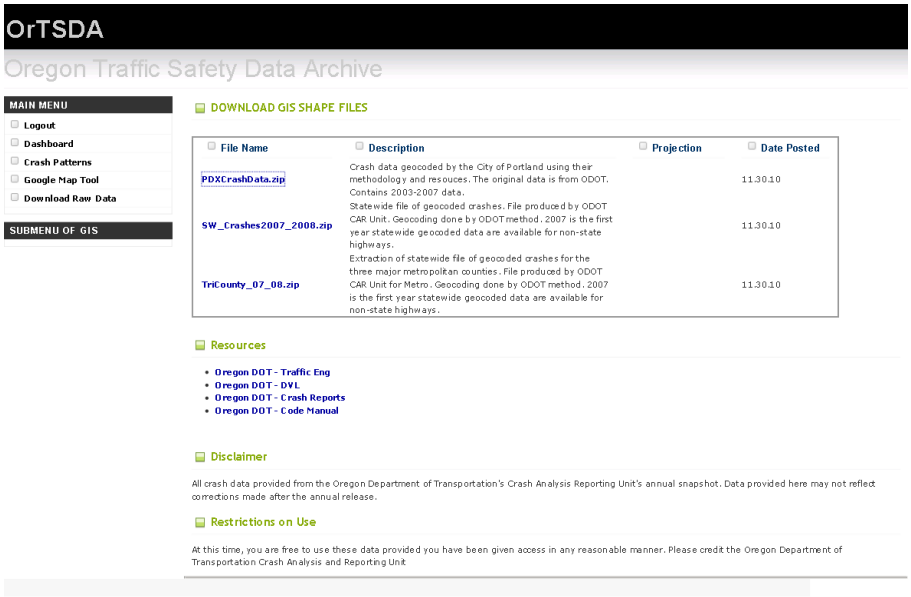


Figure 4.9 Screen capture of GIS archive

5.0 CONCLUSIONS

This report has described the preliminary work to develop the Oregon Traffic Safety Data Archive. From our review of other safety portals, it appears that interactive crash portals are restricted to a handful of state agencies or research university partners. Concerns of liability most likely contribute to a reluctance to share collision data more broadly. Of those states providing direct access to the crash data, there is a common theme of maps, data filters and crash summaries.

The cooperative nature of Oregon's public agencies that maintain the various data sources assembled for this archive made obtaining the data relatively easy. Using these data, we have built an online archive to house and disseminate safety-related information. To date, the primary effort (after obtaining and uploading the data) has focused on building the Google Maps interface. During our development, the Google Maps API changed slightly, which required some rewriting of our code. The map interface allows users to select crash data by a simple spatial selection tool. This relieves the user of needing to know route, milepost or intersection numeric identifiers to retrieve their data, and should make the crash data available to a wider audience. The crash-patterns section of the archive allows users to analyze motor vehicle crashes by a number of prebuilt queries with some user-adjustable filters. More demanding and experienced users can have access to the full database through the raw data archive.

Future improvements will include new enhanced data visualizations and improvements to the map interface. Explorations of additional data sources (such as location-based citations) will continue.

6.0 REFERENCES

Bertini, R.L., Hansen, S., Byrd, A. and Yin, T. "PORTAL: Experience Implementing the ITS Archived Data User Service in Portland, Oregon." *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1917,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 90-99.

Delucia, B. and B. Scopatz. (2005) Crash Records Systems. NCHRP Synthesis 350. Transportation Research Board of the National Academies, Washington, D.C., 2003

Griffith, M.S., C. Hayden, and H. Kalla, "Data is Key to Understanding and Improving Safety," *Public Roads*, Vol. 6, No. 4, 2003, pp. 42–47 [Online]. Available:http://www.tfhrc.gov/pubrds/03jan/09.htm [2003, Jan./Feb.].

Hauer, E. (1997). *Observational Before-After Studies in Road Safety*, Hauer, E. Observational Before-After Studies in Road Safety. Elvsevier Science Ltd.

Hauer, E. (1999) "Overview " *The Traffic Safety Toolbox: A Primer of Traffic Safety*. ITE. pp xi-vvi.

Hauer, E. (2005) *Fishing for Safety Information in Murky Waters. Journal of Transportation Engineering,* Volume 131, Issue 5, pp. 340-344.

Hedlund, James H.. (2008). *Traffic safety performance measures for states and federal agencies*. United States. National Highway Traffic Safety Administration, 2008.

Hermans E., Brijs T., Wets G., and Vanhoof K. (2009), Benchmarking road safety: lessons to learn from a data envelopment analysis. In: *Accident Analysis and Prevention*, Vol. 41(1), pp. 174-182.

Identification of Liability-Related Impediments to Sharing 409 Safety Data Among Transportation Agencies And A Synthesis Of Best Practices (2006). NCHRP Research Results Digest 306**.**

APPENDIX A - Handling Large Amounts of Markers

1. The Goal

Display all crashes per year. We can perform this with displaying markers of crashes on Google Map.

2. Problem

When you have a reasonable amount of markers, it is fairly trivial. But once you have more than few hundreds of markers, it will degrade on performance. Currently we have about 20,000 markers to display for a year; and it is going to be double up. So eventually we need to pin just above 40,000 markers which means that we have performance problem when we plot all markers.

3. Plan

There are a few approaches to increase performance. We are going to test all approaches which are available. The approaches are the following.

1) Marker Manager

   Description: Rather than adding each marker individually to the map, the marker manager keeps track of all your markers.

2) Marker Light

   Description: It render less complex markers, thus it will increase performance. The trade-off is that it does not allow us to interact with it because the markers are just an image on the map.

3) Cluster Marker

   Description: It bonds certain amount of markers into a group. That is, it detects markers that interest each other and cluster these into a single cluster marker.

4) Cluster2

   Description: It is similar to Cluster Marker. It adds two special properties to the function of cluster marker as the following.

   - Only the markers currently visible actually get created.
   - If too many markers would be visible, then they are grouped together into cluster markers.

4. Test & Result

- Test Environments:

  Windows XP
  Intell(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz
  2.95 GB of RAM

- Result

	Test 1 (168 markers)		**Test 2 (21220 markers)**	
	Loading Time	Performance For zoom in/out	Loading Time	Performance For zoom in/out
Marker Manager	1162 ms	Works well	No answer	Doesn't work
Light Markers	130 ms	Works well	22909 ms	Doesn't work
Cluster Marker	136 ms	Works well	4126 ms	Works, but too slow
Cluster2	**84 ms**	**Works well**	**3360 ms**	**Works well**

(Notice that the loading time only measures up to the time that markers are passed to the map.)

5. Conclusion

In the tests Cluster2 was the fastest technique of them because it has a good performance even if it has 20,000 markers to pin.

OTREC
OREGON TRANSPORTATION RESEARCH AND EDUCATION CONSORTIUM

P.O. Box 751
Portland, OR 97207

OTREC is dedicated to stimulating and conducting collaborative multi-disciplinary research on multi-modal surface transportation issues, educating a diverse array of current practitioners and future leaders in the transportation field, and encouraging implementation of relevant research results.