MAPPING PERCEIVED TRAVEL PROBLEMS ALONG OREGON'S ROADS

Geo-Coding Motor Carrier Survey Responses for the
Oregon Department of Transportation

Prepared by:

The UO InfoGraphics Lab
Department of Geography, University of Oregon

Eric Sproles, Graduate Research Fellow
Mike Engelmann, Project Manager

for

Oregon Department of Transportation
Planning Section
555 13th Street NE
Salem, Oregon

December 2004

Acknowledgements

The authors would like to thank the following people for their support and assistance in the preparation of this research report:

- Steve Kale, ODOT Freight Mobility Section
- Alan Kirk, ODOT Planning Section, Research Unit
- Jim Meacham, Director, UO InfoGraphics Lab

And for their GIS and cartographic work in preparing and analyzing the survey data:

- Lindsay Compton, Student Cartographer, UO InfoGraphics Lab
- Ann Huntington, Student Cartographer, UO InfoGraphics Lab
- Rita Pick, Student Cartographer, UO InfoGraphics Lab
- Darren Wyss, Graduate Research Fellow, UO InfoGraphics Lab

Disclaimer

Table of Contents

1.0 Introduction

In the fall of 2003, the Oregon Department of Transportation (ODOT) contracted with the InfoGraphics Lab at the University of Oregon for the creation and production of geo-spatial data and maps of responses to a motor carrier survey of perceived problems on Oregon's roads. This report addresses the methodologies the UO InfoGraphics Lab used to geo-code responses to the survey, challenges and problems encountered in the geo-coding process, selected results of geo-coding locations of problems, and recommendations for similar research in the future.

Portland State University (PSU) administered the initial survey in 2001 to gather information on problems that truck freight transportation firms encounter in moving freight on Oregon's roads.[1] A professional survey firm conducted telephone interviews of a randomly selected representative sample of motor carrier firms in each of the five ODOT administrative regions statewide.

The results of this study provide transportation planners and other readers with survey data that is linked to a geographic location. The study facilitates the spatial analysis of the perceptions of motor carriers on issues such as problems with roadway infrastructure, difficulties in permit processes or other administrative concerns, and the location of reported and previously unreported problems such as those related to congestion. Thus, the survey findings may help focus ODOT attention on potential opportunities for fixing the infrastructure, for improved communication with motor carriers, or for better customer service.

1.1 Mapping survey responses

The survey questionnaire posed several open-ended questions on what problems, if any, the respondents encountered on Oregon roads and where they occurred. Trained telephone interviewers recorded the responses, which the PSU principal investigator grouped into broad categories and recorded in a database.

To further investigate the locations of problems identified by survey respondents, the University of Oregon's InfoGraphics Lab created geo-spatial data and maps, graphs, and tables to portray

[1] Lawson, Catherine T., James G. Strathman, and Anne-Elizabeth Riis, *Survey Methods for Assessing Freight Industry Opinions,* Report No. FHWA-OR-RD-02-14, Oregon Department of Transportation, Salem, OR, 2002.

the data. The process of creating geo-spatial data consisted of staff from the UO InfoGraphics Lab assigning a geographic location to—or geo-coding— problems that motor carriers identified during the telephone interviews. After geo-coding the responses, staff from the InfoGraphics Lab displayed the data in a Geographic Information System (GIS), which allowed for analysis of the spatial location and distribution of roadway problems the motor carriers mentioned in their survey responses.

InfoGraphics Lab staff used the newly created geo-spatial data to produce a series of maps for ODOT. These maps provided a visual representation of the results of the motor carrier survey, and allowed for analysis of the spatial attributes of the motor carrier survey responses on the location of problems on Oregon's roads.

2.0 Data

The initial survey evolved from a series of pilot surveys that PSU performed in 2000-01. Based on results from the pilot surveys, ODOT and PSU agreed to conduct a statewide telephone survey of motor carriers. The statewide survey yielded responses from 1,872 firms out of a sample of 3,064 firms, for a response rate of 61 percent. Of the 1,872 firms responding, 617 indicated that they encountered no problems in the roadway system, and 1,255 registered one or more problems. Altogether, they reported 2,277 combined problems; this number exceeds the total number of responses since there could be more than one problem described per response.[2]

The survey data were compiled in an Excel spreadsheet consisting of 139 fields (see Appendix 1). For the mapping portion of this study, the InfoGraphics Lab staff used only 6 of the 139 fields, listed in Table 1. This included using only the responses to Question 1 in the geo-coding process.

The Excel spreadsheet was filtered to exclude data that could not be geo-coded, such as complaints about gas taxes or responses that did not include a geographic location.

[2] Lawson et al., 2002.

Table 2 represents a summary of the surveyed firms, respondents, and comments recorded for Question 1 (O_Q1). Of the 1,872 firms responding to the survey, 1,255 reported problems in responses for Question 1. Of the 1,255 firms reporting problems in responses for Question 1, 543 (43%) provided responses that contained one or more geo-codable problems, generating a total of 642 individual problem locations.[3]

Table 1
Survey Fields Used for Geo-coding Responses

Column	Name	Description
A	PSUID	UNIQUE INDENTIFIER
G	REGION	ODOT ADMINISTRATIVE REGION NUMBER
I	Respond1	SURVEY PARTICIPANT (problem reported or not)
J	Q1C_1	PROBLEM DESCRIPTION CODES FOR FIRST QUESTION
P	O_Q1	TOTAL FIRST QUESTION TEXT
R	O_Q2	SPECIFIC LOCATION TEXT

Table 2
Comments and Problems from Question 1 (O_Q1)

Number of firms surveyed	3064
Number of survey respondents	1872
Number of responses with problems identified in Question 1	1255
Number of Question 1 responses containing geo-codable problems	543
Number of individual problems from Question 1	642

Figure 1 and Table 3 provide context for the distribution of problems for Question 1 by ODOT region (column G). Table 4 shows the breakdown by problem type for responses to Question 1.

A problem was classified as "Infrastructure" if the respondent described a specific type of physical feature of the highway system (e.g., pavement conditions, curves, poor bridges, etc.). If the respondent used the term "congestion" or described traffic flow problems, the problem was classified as "Congestion." If the respondent specifically mentioned a regulatory restriction

[3]The PSU research team recorded problems for up to four questions for each respondent. To assess the geo-codability of responses, InfoGraphics staff focused on the problems identified for the first question only and not the problems identified for the second, third, and fourth questions. Of the 1,255 respondents reporting problems, 435 reported problems in their responses to the second, third, and fourth questions, as well as in their responses to Question 1.

Figure 1
ODOT Regions

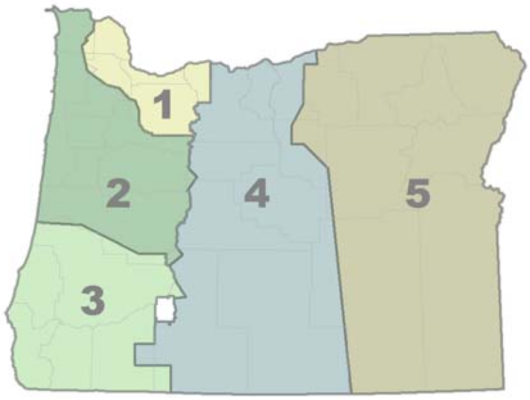


Table 3
Number of Respondents Reporting Problems in Responses to Question 1 (O_Q1), by ODOT Region

ODOT Region	# of Respondents	% of Total
1	286	23%
2	267	21%
3	232	18%
4	250	20%
5	220	18%
Total	1,255	100%

Table 4
Geo-coded Problems by Problem Type

Problem Type	Number of Individual Geo-coded Problems	% of Total
Infrastructure	347	54%
Congestion	120	19%
Restriction	62	10%
Other Drivers	15	2%
Construction	20	3%
Other	78	12%
Total	642	100%

(e.g., weight restrictions on bridges or truck length restrictions), the problem was classified as "Restriction." If the description referred to a problem related to the behavior of other drivers (e.g., general public, tourists, farm vehicles, rude drivers, etc.), the problem was classified as "Other Drivers". When a problem related to road construction was mentioned (e.g., traffic delays due to construction and lack of warnings at construction zones), the problem was classified as "Construction." Remaining issues were classified as "Other." Examples of the problems in this category included concerns regarding speed limits, permitting, bike traffic, taxes, operations, etc.[4] Appendix 2 shows a spatial and graphical representation of the 642 geo-codable problems by problem type and ODOT Region.

[4] Lawson et al., 2002.

3.0 Methodology

3.1 Data processing

From the initial master spreadsheet containing 139 fields, InfoGraphics Lab staff developed a smaller spreadsheet (*MC_survey_working.xls*) consisting of six fields as shown earlier in Table 1. The purpose of creating the smaller spreadsheet was to simplify the data-handling task and to reduce file size for processing. After geo-coding of the data was completed, the point locations were then linked back to the original (full) survey responses in the master spreadsheet according to the unique PSUID (column A).

Excel served as an efficient tool to examine the large number of survey responses. The main survey fields for determining the geographic location for responses to Question 1 were "O_Q1" and "O_Q2". All survey records with null values in these fields were filtered out of *MC_survey_ working.xls* in order to simplify the data processing and organization. This initial filter reduced the number of records down to 1,255, approximately 61% of the 1,872 total survey respondents. For the 1,255 respondents reporting problems, InfoGraphics staff developed separate spreadsheets (*MC_survey_region1.xls, MC_survey_region2.xls, etc.)* for each of ODOT's five administrative regions by filtering information according to region number as provided in column G of the initial spreadsheet.

Within each newly created spreadsheet, two additional fields were created, "link" and "comment_shape." In the link field a unique integer was assigned sequentially to each survey response, starting with 1 and continuing to all survey responses in each region. This link ID was the key attribute subsequently used to join the GIS shapefiles to the data in the master spreadsheet. The link field will be explored in greater detail later in Section 3.2.1 of this report.

For the "comment_shape" field for each response, InfoGraphics staff assigned a numeric code to describe the geometry of the problem(s) identified by the respondent, as shown in Table 5.

Table 5
Coding for the Comment_shape Field

Geometry	Comment_shape
Point	1
Line	2
Polygon	3
Multiple	4
No geo-codable problem location in column P or column R	5

The geometry of problems in each survey response was based on information in columns P and R of the initial spreadsheet. The initial survey responses served as the source data for this stage of coding. Based on the locational information in these columns, the problem's geometry was classified as a point, line, polygon, or a combination thereof. For instance, if a respondent described a problem on a particular section of road, the response was designated as a line feature, and assigned a "2" in the comment_shape field. If a respondent identified multiple problems of the same type, the response was assigned whatever comment_shape value (1, 2, or 3) applied. If a respondent identified multiple problems of differing types, the response was assigned a comment_shape value of "4." Appendix 3 provides specific examples of responses that were classified according to their comment_shape value, as well as responses that could not be geo-coded because 1) they did not define a discrete location or 2) defined a discrete location but were difficult to interpret meaningfully.

3.2 Geo-coding

The following section on geo-coding provides a description of the methodology used in this study. For a more detailed description of the fundamentals of geo-coding, please refer to the technical documentation provided by ESRI.[5]

[5] Shaner, Jeff, and Jennifer Wrightsell. *Editing in ArcMap*. ESRI Press: Redlands CA. 2000.

3.2.1 Creating each shapefile for geo-coding

Each survey response initially was geo-coded by assigning a point location in geographic space to each problem location, and then placing the points onto an existing map in a Geographic Information System (GIS). This project was set up and executed in ESRI ArcGIS (including ArcMap), though other similar GIS software packages could have been used.

The geo-coding was done on an ODOT region-by-region basis in order to create a manageable number of responses at any given time during the geo-coding process, as well as facilitate the distribution of work among multiple technicians.

A separate shapefile was created for each ODOT region (1-5) in ArcCatalog, and assigned a Lambert Conformal Conic projection and a North American 1983 datum. The Oregon Lambert Conformal Conic projection was used, as it is the standard projection for ODOT and the Oregon Geospatial Data Clearinghouse. For the parameters of this projection please refer to Appendix 4. The following naming convention was used: *region_#_master*.

An ArcMap map document (.mxd) was also created for each region, and set to an Oregon Lambert Conformal Conic projection and a North American 1983 datum. Other GIS layers were added to the base.mxd for additional reference information to aid in identifying and locating the features to be geo-coded. These layers were:

- ODOT's Highway Network (lines),
- Oregon Counties (polygons),
- Oregon Cities (points) and Urban Growth Boundaries (polygons)
- Populated Place Names (points) for smaller unincorporated communities.

Next, the shapefile of the region to be geo-coded was added to the map document (.mxd). An additional column, "Link," was added to each *region_#_master* shapefile using the "Add Field" option from the options menu on the corresponding attribute table. Table 6 shows the parameters used.

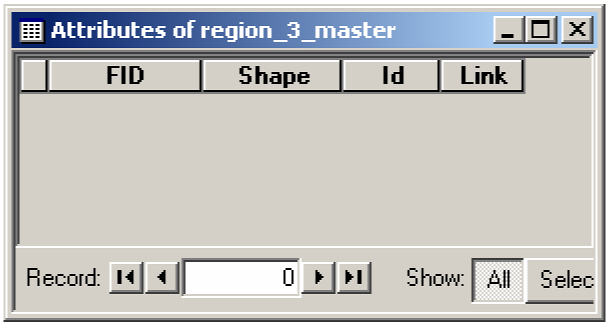
Table 6
Additional Column Added and Assigned Parameters

Column Name	Type	Precision
Link	Short Integer	4

These columns were added in order to allow the geo-coded locations to be connected back to the master spreadsheet data via the "Link" field. In selecting the field parameters, the "Type" choice governs the sort of data that will be contained in that field (e.g., short or long integer, floating point value, or text). For the integer or other numeric types, "Precision" specifies the number of digits or places allowed in each record.

At this point, the shapefile did not yet contain any data; it looked similar to the screen shown in Figure 2.

Figure 2
Sample Shapefile Screen Prior to Geo-Coding



Notice that the shapefile has four fields but no records. The FID, shape, and Id field are all auto-generated by the GIS software. The FID ("Feature ID") is a numeric value that is assigned sequentially to provide a unique identifier for each feature. The shape field is the geometry type (point, line, or polygon) for each feature, of which only one kind is allowed in a given shapefile; i.e., a single shapefile cannot contain a mix of geometry types. The ID field is also a sequentially assigned numeric value; however, it differs from the FID in that the user can subsequently edit the Id value but not the FID, which is managed internally by the software. The columns are important for the GIS to function properly, but are not part of the geo-coding or spatial analysis.

Of note is the field “Link” which was added to serve as the key identifier to join the records in the shapefile back to the original survey data. This procedure will be discussed in section 3.3: “Joining geo-coded points to data.”

3.2.2 Identifying geographic locations of survey responses

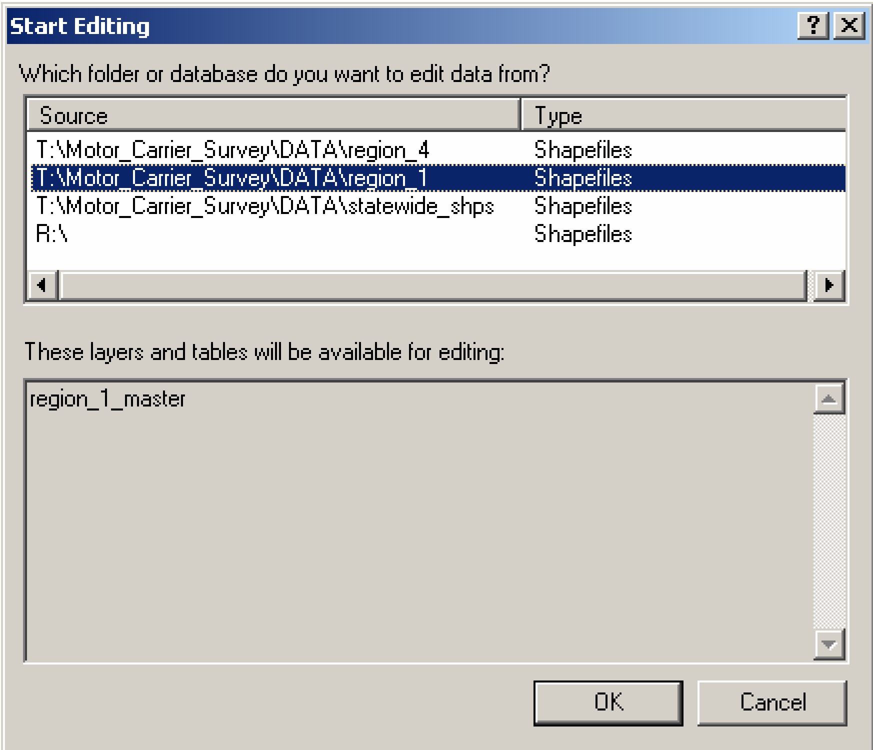
The initial phase of geo-coding assigned a point location for all comment_shape classifications regardless of their actual comment shape field designation (i.e., a point would be created at this stage even for a linearly defined problem). This process was implemented for two primary reasons. The first was efficiency on the part of the GIS technician, sparing him or her from having to manage several different feature classes concurrently. The second reason was to provide a common spatial framework by which all of the survey results could be compared as one set, particularly at a statewide scale such as in the case of small multiple maps identifying the distribution of different problem descriptions (infrastructure, congestion, construction, etc.) such as those shown in Appendix 2.

The first step in the initial phase of geo-coding involved identifying a geographic location for each survey comment. Placement of the feature geometry was accomplished via the *Editor* toolbar in ArcMap. Once the *Editor* toolbar was activated, editing was begun by selecting “Start Editing” from the *Editor* menu. When the dialog box similar to the one shown in Figure 3 appeared, the appropriate data source folder was chosen in order to edit the desired shapefile (in Figure 3, the Region 1 shapefile has been selected for editing).

Next, the snapping function was activated in *Editor* for the highway shapefile (“edge,” “vertex,” and “endpoint” all selected). This function enabled exact placement of the points along the highway linework.

The attribute table was then opened for the shapefile to be geo-coded. From the *Editor* toolbar, the *Sketch* tool was activated and set to *Create New Feature* from the drop-down menu. This series of steps was necessary in order to add a point into the *region_1_master.shp* using the base map of Oregon and the highways shapefile as spatial reference points.

Figure 3
Sample Data Editing Screen



In Excel, the corresponding survey file was opened, with the "comment_shape" and "Link" field completed (see Table 7). Survey response fields O_Q1 and O_Q2 (columns P and R) were used to determine the geographic location of the survey responses. The point was placed into the shapefile in ArcMap by using the *Sketch* tool on the highway line segment described in the Excel file. For polygon features the point was placed in the center of the polygon. For linear features, the point was placed in the approximate midpoint between the two nodes of the linear segment. For multiple survey responses where problems were identified in the same location, the points were slightly offset in order to distinguish them visually.

Table 7
Example of Comment_shape and Link Files in the Excel Spreadsheet

Problem Description (Column Q1C_1)	Link
I-5 CORRIDOR FROM PORTLAND TO EUGENE	1
THE SOUTHBOUND RAMP ONTO I-5 OUT AT COTTAGE GROVE	128

In the first example in Table 7, a point on I-5 approximately midway between Portland and Eugene would be added to the map, and one record would be added to the attribute table. Next, the Link field (in the attribute table of *region_#_master.shp*) was selected for the newly added survey response in ArcMap, and the corresponding link value (1) was added (from the Excel spreadsheet). Similarly, a point would be added at I-5 in Cottage Grove, and the link #128 would be added into the attribute table. Appendix 5 shows a screen capture illustrating the results of this process for selected problems in south central Oregon.

It was important that the link number in the attribute table match the record in the Excel file, as this was the key attribute used to join the data tables to the GIS shapefiles. The process was repeated in sequential order of the link numbers in order to reduce the likelihood of errors during the process.

The same procedure was used for all of the geo-codable survey responses and for each regional shapefile. If a single comment exhibited a one-to-many relationship as described in section 3.2, the same link value was entered into the attribute table for each geo-codable problem. The link value reflected the initial survey response record, not the location of the problem on the map. The geo-coding edits were saved frequently during the process.

Throughout the process, regular verification was performed to ensure the correct link value was entered into the attribute table for the geometry corresponding to the original survey entries. Once all of the geo-coding had been completed and all edits were saved, "Stop editing" was selected from the *Editor* toolbar to prevent further accidental edits. After editing was finished for each region, a supervising technician checked the dataset to look for discrepancies in the placement of features and to minimize errors in the interpretation of respondent descriptions of problem locations.

3.3 Joining geo-coded points to data

"Joining" geo-spatial data to the original survey response provides a method of extending the information about geographic features in a shapefile. In the motor carrier survey, the initial geo-coding of responses was based on 6 of the 139 fields in the master spreadsheet. Using the "join" function in ArcMap, the geo-coded features were connected back to their corresponding records in the master spreadsheet in order to create a dataset that was a combination of both the geo-coded responses and the initial survey data.

To prepare the data for joining, special steps were taken to avoid overwriting the geo-coding. One of the idiosyncrasies of the ESRI software is that it will only import external tables in dBase (.dbf) format, while each shapefile also consists of a .dbf component containing the embedded data for the corresponding attribute table. Consequently, if the Excel file for a region was exported to a .dbf with the same name as the shapefile for that region, the new table would overwrite the existing shapefile's component .dbf and the shapefile would become corrupted, losing all the data that had just been geo-coded. Thus, a back-up file was created for every *region_#_master.shp* file.

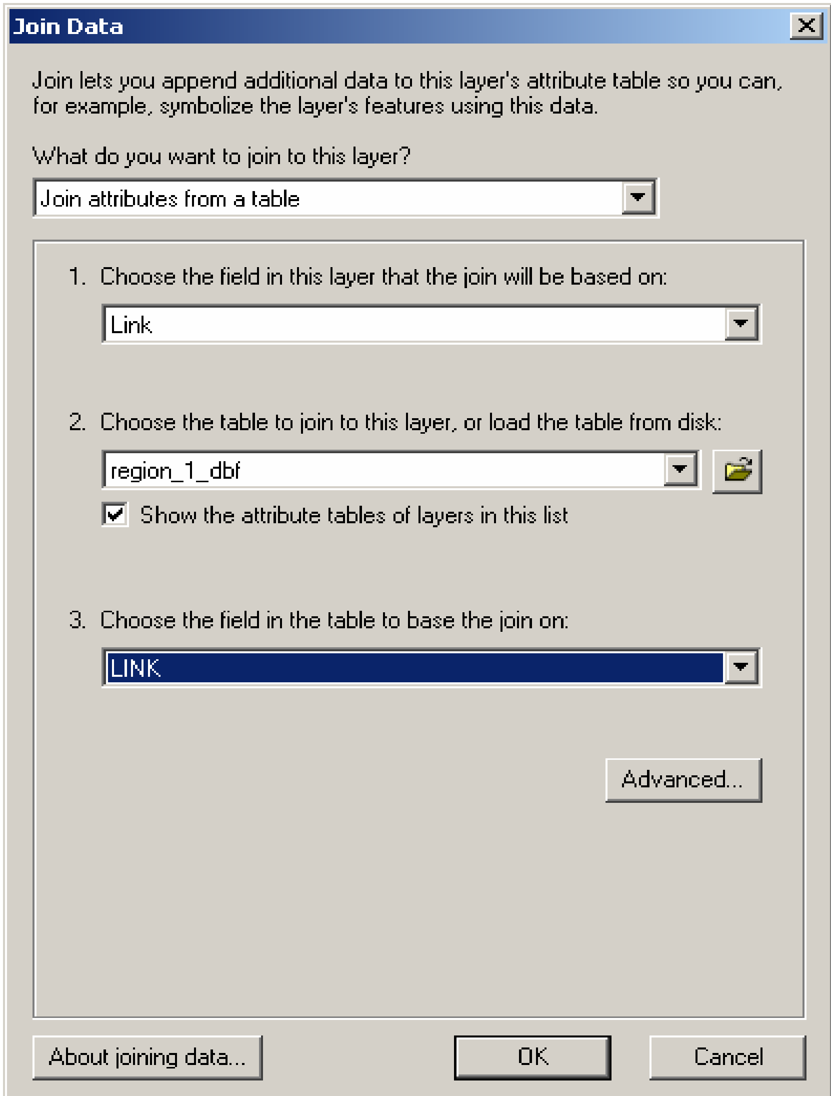
The Excel worksheet for each region was saved as a separate dBase file (.dbf), with a new, unique naming convention. This was important to avoid conflicting with the existing .dbf component of the corresponding shapefile. The newly created .dbf file was then added to the data frame for that region in ArcMap. The *region_#_master* shapefile was activated to join by selecting the shapefile layer for that region and choosing "Joins…" from the "Joins and Relates" menu. Figure 4 displays the appropriate settings used for joining the data.

At this point in the process, the *region_#_master.shp* file contained both the geo-coded location and the original survey responses appended to the geometry for each feature. This join was only a temporary function of the .mxd in ArcMap, however, so the joined *region_#_master.shp* file was then exported to create a new unique shapefile in which the joined data was permanently embedded.

3.4 Creation of linear data

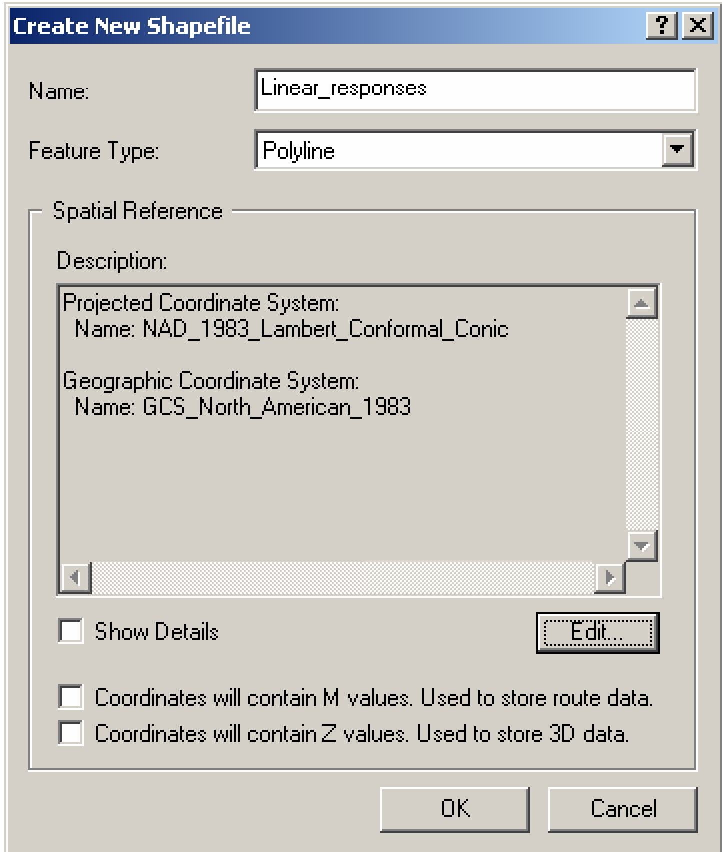
As mentioned in Section 3.2.2, the initial phase of geo-coding assigned point features for all geo-codable problems in order to show the generalized spatial distribution of problems identified by respondents. Subsequently, a linear road segment was assigned for all of the comments that were classified as linear features.

Figure 4
Sample Join Data Screen



The first step in the creation of the linework for road segments was to create a new shapefile (linear), set to Oregon Lambert Conformal Conic projection and NAD 1983 datum. Figure 5 shows the "Create New Shapefile" dialog box in ArcCatalog.

Figure 5
Sample Create New Shapefile Screen



The ArcMap document file that was used for the point geo-coding was reopened, and saved as a new file (*Linear_responses.mxd*), and the newly created shapefile was added. A "Link" field was added to the shapefile in ArcMap, as was done with the point shapefiles. This newly added shapefile was enabled for editing. The ODOT highways shapefile was designated as the only shapefile available to be selected.

Also opened was the master copy of all of the geo-coded points. These records were filtered to show the points, by querying for the linear features code (comment_shape = 2). From these records, the GIS technician noted the beginning and end point of the roadway segment based on the language recorded in each survey record.

In the ODOT highway shapefile, the road segments that corresponded to the described stretches of road were selected and then copied and pasted into *Linear_responses.shp*. Depending on the beginning and ending locations of each problem description, the ODOT highway segment(s) may have required clipping and/or merging so that every individual line feature matched the same extent as the corresponding problem. The link number associated with the original survey data was then added into the "Link" field of *Linear_responses.shp*. The result was a record in *Linear_responses.shp* that contains all linear attribute data associated with the ODOT shapefile. This record was joined to the motor carrier survey data via the Link field in the same manner as before.

For example, for the survey response "HWY 97 BETWEEN BEND AND REDMOND," the linework for Highway 97 between Bend and Redmond was copied from the highway shapefile and then pasted into *Linear_responses.shp*, creating an individual feature record. The endpoints for the segment were modified based upon the extent of each city's Urban Growth Boundary (UGB). In this instance, "between" is interpreted not to include the cities themselves, so the northern boundary of the Bend UGB would represent the southern endpoint of the problem segment, and the southern end of the Redmond UGB would be the northern endpoint. The Link field number from the Excel file for that region was then added to the record in *Linear_responses.shp*. After all of the data records were assigned a linear feature, the *Linear_responses.shp* would be joined to the original survey data using the same methodology described in Section 3.3, "Joining geo-coded points to data."

3.5 Hardware and software

All of the data processing and geo-coding was completed on a Windows 2000 system with at least 512 MB of RAM. The software used in the geo-coding portion of the study consisted of ESRI's ArcMap and ArcCatalog (version 8.3). Microsoft Excel was used for data sorting and filtering.

4.0 Challenges and problems encountered in geo-coding

While the methodology incorporated in this study was relatively straightforward, a few challenges were encountered, primarily due to the nature of the survey responses. Geo-coded data about problem locations were based on the PSU survey research team's interpretation of the respondents' answers. Staff from the UO InfoGraphics Lab interpreted and classified these responses, geo-coded the relevant locational information, and mapped the results. The discussion below outlines the major challenges and problems.

4.1 Inconsistent levels of specificity in the responses

The survey responses were inconsistent in their specificity about location of problem, with some responses very detailed while others were quite general. As discussed earlier in Section 3.1 of this report, InfoGraphics Lab staff classified responses according to whether the problems could be classified as occurring at a specific location (point), a specific stretch of roadway (line), or an area (polygon) within which identifying a specific location or stretch of highway was problematic. Several protocols for the comment_shape field were used to ensure that the problem location was consistently interpreted and geo-coded. These protocols were as follows:

- A point feature had to describe a distinct geographic point such as an intersection or milepost,
- A linear feature was required to have a discernable beginning and end point, and
- A polygon feature was required to mention a definable area such as an incorporated city or unincorporated community.

Collaboration among InfoGraphics Lab staff helped reduce the likelihood of inconsistently or incorrectly classifying responses.

4.2 Local roads

While most respondent-identified problems were on state highways, a few were on local roads. Because these comments represented only a small proportion (approximately 2-3%) of geo-codable problems, InfoGraphics staff chose not to analyze them further. These problems, however, could be geo-coded using other referencing systems such as street addresses or city/county data.

4.3 Geo-codability of the responses

Of the 1,255 respondents identifying at least one problem, 713 (57 percent), mentioned problems whose locations could not be geo-coded because 1) the respondent did not provide a description that specified a location for the problem according to the protocol outlined in Section 4.1, or 2) the respondent provided a geo-codable location that could not be interpreted meaningfully enough for the purposes of transportation planning and programming. See Appendix 3 for examples of problems that were not geo-codable.

In some cases, InfoGraphics Lab staff were able to figure out how to geo-code a problem where the locational information was not immediately clear from the survey response. For example, one respondent referred to a problem for "TERITORIAL RD BY BONITA AND EUGENE." Staff interpreted "Bonita" to mean "Veneta," a small community near Eugene. While staff was able to correctly identify the location of this problem, it may have "missed" the correct location of responses in other parts of the state. A critical factor in correctly geo-coding problem locations is a good understanding of local geography and local roads across the state. The better the understanding is of local geography and local roads, the better will be the locational interpretation of responses that are correctly reported, as well as those that are incorrectly reported.

Beyond identifying geographic features, the interpretation of what constituted a sufficiently "meaningful" location description was a difficult question. From a GIS standpoint the issue is that the level of spatial specificity required to place and/or edit the feature geometry for a problem description could be fundamentally different and more broad than that needed for transportation planning and programming. From a statewide mapping perspective InfoGraphics staff felt that some of the less specific geo-codable data still had significant potential to contribute to meaningful patterns. For applications that required a more specific scale, though, it might not be meaningful enough. In the absence of clear specification and definitions of "meaningful," this effort was "inclusive" rather than "exclusive" when deciding what to geo-code. Another iteration of analysis could filter out features depending on the needs of the application.

To be considered geo-codable, linear features needed specific beginning and ending points rather than a predetermined minimum or maximum length. Oregon has a wide range of road density -- very high in the Portland metro area, and very sparse in Eastern Oregon. In urban areas, with greater density of roads, intersections, and other significant transportation features, a shorter minimum length segment is needed in order to be meaningful than in Eastern Oregon where a long segment of a mountain or desert highway can be considered a single feature having a similar perceived problem. For example, rutting, snow removal, and tight curves may occur along extensive lengths of mountainous highways, so longer segment descriptions could meaningfully apply in areas of low road density.

Survey responses associated with polygon features tended to be more broad than specific, and often were difficult to geo-code meaningfully. For example, a problem identified only for the Portland area may not be meaningful for transportation planning and programming purposes because the specific location of the problem within the Portland area is not identified. As noted above, InfoGraphics staff chose to be “inclusive” rather than “exclusive.” With additional review, some of the responses classified as polygons might more appropriately be classified as not geo-codable for purposes of planning and programming.

5.0 Results/Recommendations

5.1 Results

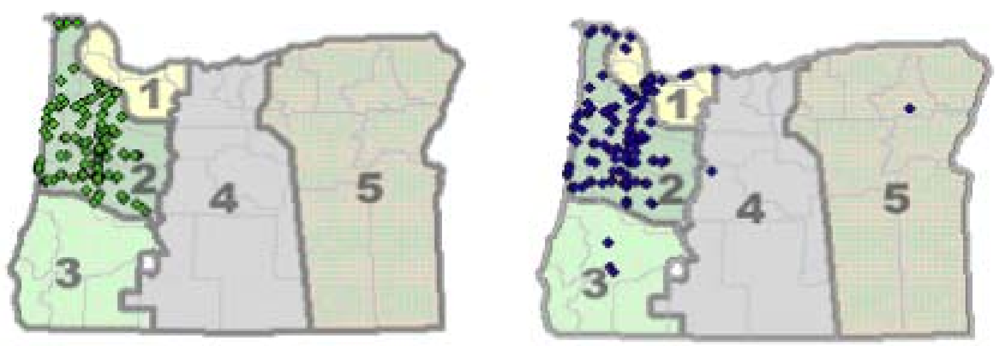
After the survey data were processed and geo-coded by the InfoGraphics Lab staff, they were compiled in tabular, graphical, and cartographic formats. The tables, graphs, and maps generally showed problem locations by ODOT region and by problem types, and help visualize trends within the data that may not be evident in text form, as graphics reveal data.[6]

[6] Tufte, Edward R. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press, 2001.

In interpreting the maps, it was important to establish whether the responses were being portrayed and analyzed by region of the problem identified, or by region of the respondent. Figure 6, for example, shows location of problems in ODOT Region 2 as identified by all respondents, and location of problems statewide as identified by respondents located in Region 2. The patterns for Region 2 are similar in both maps; the main difference is that the map on the left shows a few more problems in Region 2 than does the map on the right. The additional problems shown in Region 2 for the map on the left generally are in the same location as problems for the map on the right. Of the problems shown outside Region 2, as indicated by the map on the right, a majority are in Region 1, which is not unexpected given Region 1's proximity to Region 2 and the amount of truck travel between the two regions.

Figure 6
Example of Geo-codable Problem Locations

Location of Region 2 Problems for All Respondents | **Location of All Problems by Respondents in Region 2**



Based on the responses for Question 1, Table 8 shows, by ODOT region, the number and percentage of geo-codable responses, the number and percentage of geo-codable problems, and the number of responses that included more than one geo-codable problem. Region 4 accounted for the greatest number and percentage of responses with geo-codable problems, the greatest number and percentage of geo-codable problems, and the greatest number of responses with multiple geo-codable problems. Regions 1 and 5 generally had the lowest number and percentages of geo-codable responses and problems. Further analysis of the responses and other information is needed to identify possible underlying reasons for the differences among regions.

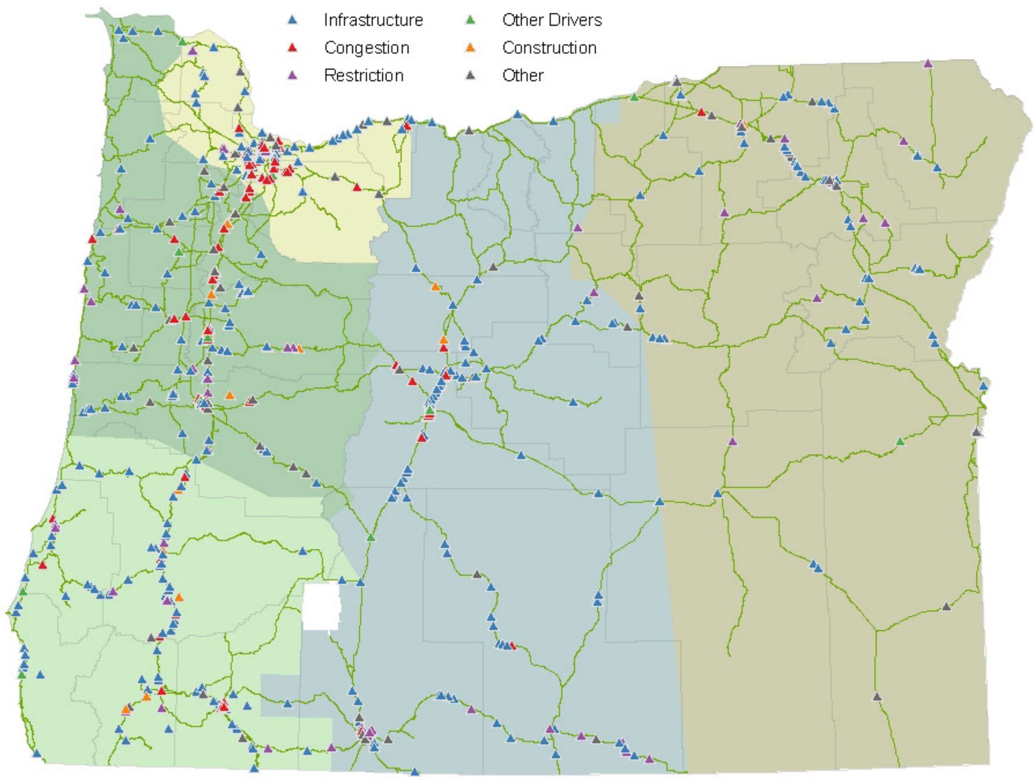
Table 8
Question 1 Geo-codable Responses and Problems by ODOT Region

	A	B	C	D	E	F
Region	**Number of Responses with Geo-codable Problems**	**% of Total**	**Number of Geo-codable Problems Reported by Respondents**	**% of Total**	**Number of Responses with Multiple Geo-codable Problems**	**Column E/ Column A**
1	90	17%	110	17%	9	10%
2	113	21%	139	22%	16	14%
3	109	20%	126	20%	15	14%
4	143	26%	167	26%	23	16%
5	88	16%	100	15%	10	11%
Total	543	100%	642	100%	73	n/a

Figure 7 shows the locations statewide of all 642 geo-coded problems by problem type. The map reveals trends that appear to be consistent with information available from other sources. For instance, congestion problems tend to be focused on the I-5 corridor between Portland and Salem, with the heaviest concentration in the Portland area. Infrastructure problems tend to be concentrated on the most heavily traveled routes--for example, I-5 and I-84, and US 97 in central Oregon, and selected locations along other routes. Patterns for other problem types are less clear, possibly due to the relatively small number of problems that were geo-codable for these types. Further analysis would be necessary to compare how the locations of perceived problems, as shown in Figure 7, compare with known locations of these problems as shown in other sources of information.

Table 9 provides a breakdown of the 642 geo-coded responses by region and problem type. Infrastructure, which ranged from 38 percent (Region 1) to 63 percent (Region 3) of geo-codable problems, was the dominant problem type in all regions except Region 1, where congestion was the most frequently stated geo-codable problem type. Congestion and restrictions generally were the second and third most important problem types in regions outside the Portland area.

Figure 7
Location of Geo-codable Problems by Problem Type



5.2 Recommendations

The mapping of problems identified in the motor carrier survey is intended to serve as a tool for transportation planners and others to visualize and interpret the spatial attributes of the data. The discussion of the geo-coding of problems identified by motor carriers may be useful for similar studies in the future. The difficulties encountered in the implementation of geo-coding were predominantly associated with 1) the lack of locational specificity found in the original survey data; and 2) challenges in distinguishing what locational information was meaningful to geo-code from what was not meaningful to geo-code. As noted before, this project only looked at Question 1 and its multiple problems. Additional geo-coding of responses for Questions 2 through 4 would be an important addition to the completeness of the project.

Table 9
Geo-coded Problems by Region and Problem Type

ODOT Region	Number of Geo-coded Problems	% of Region Total
1	**110**	
Infrastructure	42	38%
Congestion	47	43%
Restriction	2	2%
Other Drivers	4	4%
Construction	4	4%
Other	11	10%
2	**138**	
Infrastructure	65	47%
Congestion	32	23%
Restriction	15	11%
Other Drivers	4	3%
Construction	7	5%
Other	15	11%
3	**126**	
Infrastructure	79	63%
Congestion	15	12%
Restriction	18	14%
Other Drivers	2	2%
Construction	4	3%
Other	9	7%
4	**167**	
Infrastructure	102	61%
Congestion	24	14%
Restriction	13	8%
Other Drivers	3	2%
Construction	4	2%
Other	21	13%
5	**100**	
Infrastructure	59	59%
Congestion	2	2%
Restriction	14	14%
Other Drivers	2	2%
Construction	1	1%
Other	22	22%

For future studies, it would be beneficial for the interviewer to have the respondent pinpoint a discrete geographic location if possible. In addition, the establishment of clear goals for the use of the data and an adequate definition of "geospatially meaningful" that matched the intended use of the data would lead to a more successful geo-coding effort.

The initial survey did request a specific location (i.e., cities or highways). While many respondents provided problem locations, they often did not provide information that was specific enough to be geo-coded meaningfully. Future studies of a similar nature should explore possibilities for requesting a more specific description for problem locations. For example, when cities or towns are identified, trying to gain a specific road or location within the city would help the applicability of the response. Similarly for road segments, a beginning and ending point for problem locations would help clear any uncertainty and allow for better cartographic analysis. Attempting to obtain more detailed information could result in fewer responses and a lower response rate if the time and expense required per interview increased and if the respondents viewed a more lengthy survey as too much of an imposition.

Greater initial training for the interviewers could improve efficiency in gaining meaningfully geo-codable data. For example, this report identifies geo-codable and non-geo-codable responses, which could help interviewers for future surveys to understand and distinguish what constitutes a sufficiently specific description of a problem location.

The pre-processing of data prior to geo-coding also might help. For example, pre-processing might identify in advance which responses include sufficiently geo-codable data and which do not. For problems that are identified as potentially geo-codable, pre-processing might include developing ancillary data such as milepost numbers or other geo-referencing information that could enhance the geo-codability of responses.

The geo-coding process itself may be improved by employing a Linear Referencing System (LRS) to create a series of “event themes” in which each survey response is defined as an individual “event” along a single common road base (e.g., ODOT’s highway shapefile) as opposed to creating a new piece of separate geometry for each record. Since ODOT’s highway network is already routed and linearly referenced (by mileposts), the beginning and ending (milepost) values for each problem segment could be loaded into the LRS and then dynamically segmented onto the core linework. The “DynSeg” events would be based on a distinct mileage value (in the case of points), or a beginning and ending mileage range (in the case of a linear segment), calculated along a measured highway route.

Potential advantages would include greater flexibility and a more “lightweight” dataset because the individual feature geometry would not have to be replicated for each geo-coded record. One potential disadvantage is that it requires a common, routed LRS network on which to run. While that is not an issue for ODOT's highways, it could be when dealing with local or county roads for which a GIS network has not been established, or for those that use a different LRS. Another concern is that dynamic segmentation is still highly software-limited, whereas basic shapefiles have become a widely accepted standard even for non-ESRI users. For more information on Linear Referencing Systems please refer to *Linear Referencing and Dynamic Segmentation in ArcGIS 8.1.*[7]

[7] ESRI. *Linear Referencing and Dynamic Segmentation in ArcGIS* ., Redlands, CA: ESRI Press, 2001.

Appendices

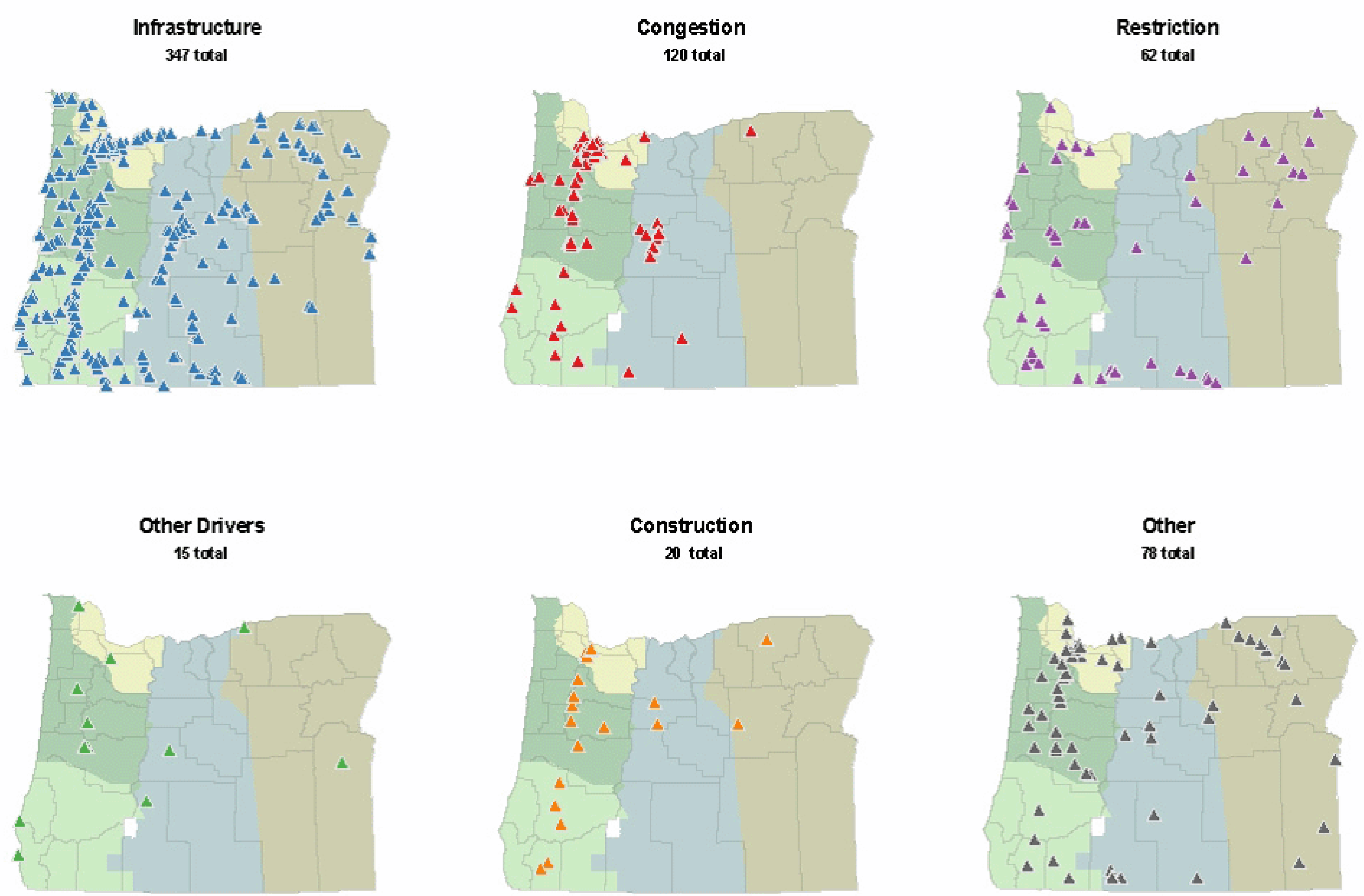
Appendix 1: Data Fields for Survey Responses

COLUMN	NAME	EXPLANATION
A	PSUID	UNIQUE INDENTIFIER
B	S_DAT	DAY OF INTERVIEW(LAST CONTACT)
C	S_JSE	DAY OF WEEK
D	S_HRD	BEGIN TIME
E	S_DUS	DURATION IN SECONDS
F	S_DUR	DURATION IN MINUTES
G	REGION	ODOT REGIONS
H	INTRO	FIRST CONTACT
I	respond1	SURVEY PARTICIPANT
J	Q1C_1	FIRST PROBLEM DESCRIPTION IN FIRST QUESTION
K	T_Q1C_1	TEXT (IF ADDITIONAL PROBLEMS DESCRIBED)
L	Q1C_2	SECOND PROBLEM DESCRIPTION IN FIRST QUESTION
M	T_Q1C_2	TEXT
N	Q1C_3	THIRD PROBLEM DESCRIPTION IN FIRST QUESTION
O	T_Q1C_3	TEXT
P	O_Q1	TOTAL FIRST QUESTION TEXT
Q	Q2	SPECIFIC LOCATION
R	O_Q2	SPECIFIC LOCATION TEXT
S	Q3	SPECIFIC TIME OF DAY
T	O_Q3	SPECIFIC TIME OF DAY TEXT
U	Q4	SPECIFIC TIME OF YEAR
V	O_Q4	SPECIFIC TIME OF YEAR TEXT
W	OTHER1	OTHER IMPACT
X	COST1A	COST ASSOCIATED WITH TAKING ANOTHER ROUTE
Y	COST1B	COST ASSOCIATED WITH FUEL
Z	EQUIP1	EQUIPMENT DAMAGE
AA	SAFETY1	SAFETY
AB	DRSTRS1	DRIVER STRESS
AC	COST1C	GENERAL COST
AD	DELAY1	TIME DELAY
AE	Q5	IMPACT DESCRIBED
AF	O_Q5	IMPACT DESCRIBED TEXT
AG	Q6	ALTERNATIVE
AH	O_Q6	ALTERNATIVE TEXT
AI	Q7A1	SECOND PROBLEM MENTIONED
AJ	Q7A1C_1	SECOND PROBLEM DESCRIPTION
AK	O_Q7A1	SECOND PROBLEM DESCRIPTION TEXT
AL	Q7A2	SPECIFIC LOCATION
AM	O_Q7A2	SPECIFIC LOCATION TEXT
AN	Q7A3	SPECIFIC TIME OF DAY
AO	O_Q7A3	SPECIFIC TIME OF DAY TEXT
AP	Q7A4	SPECIFIC TIME OF YEAR
AQ	O_Q7A4	SPECIFIC TIME OF YEAR TEXT
AR	OTHER2	OTHER IMPACT

AS	COST2A	COST ASSOCIATED WITH TAKING ANOTHER ROUTE
AT	COST2B	COST ASSOCIATED WITH FUEL
AU	EQUIP2	EQUIPMENT DAMAGE
AV	SAFETY2	SAFETY
AW	DRSTRS2	DRIVER STRESS
AX	COST2C	GENERAL COST
AY	DELAY2	TIME DELAY
AZ	A7Q5	IMPACT DESCRIBED
BA	O_A7Q5	IMPACT DESCRIBED TEXT
BB	A7Q6	ALTERNATIVE
BC	O_A7Q6	ALTERNATIVE TEXT
BD	Q7B1	THIRD PROBLEM MENTIONED
BE	Q7B1C_1	THIRD PROBLEM DESCRIPTION
BF	O_Q7B1	THIRD PROBLEM DESCRIPTION TEXT
BG	Q7B2	SPECIFIC LOCATION
BH	O_Q7B2	SPECIFIC LOCATION TEXT
BI	Q7B3	SPECIFIC TIME OF DAY
BJ	O_Q7B3	SPECIFIC TIME OF DAY TEXT
BK	Q7B4	SPECIFIC TIME OF YEAR
BL	O_Q7B4	SPECIFIC TIME OF YEAR TEXT
BM	OTHER3	OTHER IMPACT
BN	COST3A	COST ASSOCIATED WITH TAKING ANOTHER ROUTE
BO	COST3B	COST ASSOCIATED WITH FUEL
BP	EQUIP3	EQUIPMENT DAMAGE
BQ	SAFETY3	SAFETY
BR	DRSTRS3	DRIVER STRESS
BS	COST3C	GENERAL COST
BT	DELAY3	TIME DELAY
BU	B7Q5	IMPACT DESCRIBED
BV	O_B7Q5	IMPACT DESCRIBED TEXT
BW	B7Q6	ALTERNATIVE
BX	O_B7Q6	ALTERNATIVE TEXT
BY	Q7C1	FOURTH PROBLEM MENTIONED
BZ	Q7C1C_1	FOURTH PROBLEM DESCRIPTION
CA	O_Q7C1	FOURTH PROBLEM DESCRIPTION TEXT
CB	Q7C2	SPECIFIC LOCATION
CC	O_Q7C2	SPECIFIC LOCATION TEXT
CD	Q7C3	SPECIFIC TIME OF DAY
CE	O_Q7C3	SPECIFIC TIME OF DAY TEXT
CF	Q7C4	SPECIFIC TIME OF YEAR
CG	O_Q7C4	SPECIFIC TIME OF YEAR TEXT
CH	OTHER4	OTHER IMPACT
CI	COST4A	COST ASSOCIATED WITH TAKING ANOTHER ROUTE
CJ	COST4B	COST ASSOCIATED WITH FUEL
CK	EQUIP4	EQUIPMENT DAMAGE
CL	SAFETY4	SAFETY
CM	DRSTRS4	DRIVER STRESS
CN	COST4C	GENERAL COST

CO	DELAY4	TIME DELAY
CP	C7Q5	IMPACT DESCRIBED
CQ	O_C7Q5	IMPACT DESCRIBED TEXT
CR	C7Q6	ALTERNATIVE
CS	O_C7Q6	ALTERNATIVE TEXT
CT	Q8	TRUCK TYPE
CU	O_Q8	TRUCK TYPE TEXT
CV	Q9	TRIP LENGTH
CW	O_Q9	TRIP LENGTH TEXT
CX	Q10	NUMBER OF TRUCKS IN SHOP
CY	Q10A	IF Q10 = 9998 OR 9999 IS THAT
CZ	Q10C	COMBINED TRUCK GROUP
DA	Q11A	OWNER-OPERATOR
DB	O_Q11A	JOB TITLE TEXT
DC	Q11B	DRIVER OR DISPATCHER
DD	O_Q11B	JOB TITLE TEXT
DE	Q12	ANYTHING TO ADD?
DF	Q12C_1	FIFTH PROBLEM MENTIONED
DG	O_12	FIFTH PROBLEM MENTIONED TEXT
DH	Q12AA	ANOTHER TRANSPORTATION PROBLEM
DI	Q12A2	SPECIFIC LOCATION
DJ	O_Q12A2	SPECIFIC LOCATION TEXT
DK	Q12A3	SPECIFIC TIME OF DAY
DL	O_Q12A3	SPECIFIC TIME OF DAY TEXT
DM	Q12A4	SPECIFIC TIME OF YEAR
DN	O_Q12A4	SPECIFIC TIME OF YEAR TEXT
DO	OTHER5	OTHER IMPACT
DP	COST5A	COST ASSOCIATED WITH TAKING ANOTHER ROUTE
DQ	COST5B	COST ASSOCIATED WITH FUEL
DR	EQUIP5	EQUIPMENT DAMAGE
DS	SAFETY5	SAFETY
DT	DRSTRS5	DRIVER STRESS
DU	COST5C	GENERAL COST
DV	DELAY5	TIME DELAY
DW	Q12A5	IMPACT DESCRIBED
DX	O_Q12A5	IMPACT DESCRIBED TEXT
DY	Q12A6	ALTERNATIVE
DZ	O_Q12A6	ALTERNATIVE TEXT
EA	GENDR	RECORD GENDER
EB	TIME	TIME CALCULATION
EC	TIMEM	TIME IN MINUTES
ED	INTO1	COMPLETED SURVEY
EE	ATMPT	NUMBER OF ATTEMPTS
EF	ADD_TYPE	ODOT RECORDS
EG	ADD_TRKS	ODOT RECORD OF TRUCKS
EH	ADD_MLS	ODOT RECORD OF MILES
EI	ACT_ID	ODOT ACTS

Appendix 2: Statewide Points by Problem Type in Small Multiple Maps



Note: Since the locations of the problems may be similar or identical in some cases, the number of separate points visible on each map may not correspond to the total noted above the map.

Appendix 3: Examples of Responses with Geo-codable and Non-geo-codable Problem Locations

Responses with Geo-codable Problem Locations	Comment_Shape
I-5 IS BAD SHAPE AT THE TIGARD EXIT	Point
WHERE I-5 AND HWY 217 CROSS	Point
HGWAY 97, 12 MILES SOUTH OF BEND SUN RIVER	Point
EXITS 103 AND 136 ON I-5	Point
HWY 97 BETWEEN BEND AND REDMOND	Line
HWY 30 BETWEEN RAINIER AND CLATSKANIE	Line
FROM EUGENE TO SALEM ON I-5	Line
HWY 101 BETWEEN COOS BAY AND REEDSPORT … AND MYRTLE POINT TO ROSEBURG	Line
HOOD RIVER--WHOLE TOWN	Polygon

Responses with Non-geo-codable Problem Locations

THEY PULLED MY LICENSE, WOODBURN DRAG STRIP

OUT OF PHILOMATH PORTABLE SCALES

ROUGH ROADS ON HWY 99W

IN OREGON

ANYWHERE IN CENTRAL OREGON

NOT ENOUGH PASSING LANES; HWY 38

BAD ROADS; EVERYWHERE

TRAFFIC PROBLEMS; CENTRAL OREGON

Appendix 4: Parameters for the Oregon Lambert Conformal Conic Projection

Initial Projection Name: “Lambert_Conformal_Conic”

Linear Unit Name: “Foot” (0.3048 meters)

Projection Parameters:

Central_Meridian
-120.5

Latitude_of_Origin
41.75

Standard_Parallel_1
43.0

Standard_Parallel_2
45.5

False_Easting (feet)
1312335.958

False Northing
0.0

Scale_Factor
1.0

Geographic Coordinate System: “North American Datum 1983”

Appendix 5: ArcMap Screen of Geo-coding Process