2.0 Overview of Survey Methods

In October 2006, Washington State Ferries (WSF) conducted an origin-destination onboard survey,[8] which collected ridership pattern information from Washington State Ferries passengers. The methods employed in this survey are documented in detail in the *2006 Travel Survey Technical Report of Methods*.[9]

This report summarizes the survey design, procedures, and data processing. There are some key differences in the conduct of the survey that differ from the previous 1999 and 1993 surveys, to enhance the type of data collected or to improve the accuracy of the data for use in future ferry planning activities. These differences are highlighted in this section and are references in the analysis report where the comparative results are affected by these differences. The two primary differences are the use of a web-based field management system to improve the management of the field data collection, and the use of more detailed expansion factors to expand the data to a full day of travel rather than the p.m. half-day expansion in 1999 (the 1993 survey data was not expanded). In addition, weekend data was collected on Saturdays in 2006 to capture the higher volumes (weekend data was collected on Sundays in the previous surveys). These differences are described in more detail in this section.

The primary objective of the survey was to capture accurate measures of ferry rider travel patterns. These surveys allow WSF to accomplish the following objectives:

- Match services with customer needs;
- Run an efficient ferry operation; and
- Update data inputs to ferry travel model for use in WSF Long Range Plan.

The survey was conducted during the p.m. peak-period for weekday travel and during the a.m. peak-period for weekend travel to capture behavior for the highest ridership periods. The data collected included, but was not limited to:

- Geographic markets (trip origin and destination);
- Demographic characteristics (age, sex, employment status, household size, income, etc.);

---

[8] NuStats conducted the survey under the sponsorship of Washington State Ferries and the Washington State Department of Transportation and as subconsultants to Cambridge Systematics. RST International also provided technical oversight. The project team was led by Celine Gihring at WSF.

[9] NuStats, *2006 Travel Survey Technical Report of Methods Final Report*, prepared for the Washington State Ferry System, February 2007.

- Ferry travel characteristics (trip purpose, fare type, frequency, boarding wait time, boarding method, number of passengers traveling together, etc.); and
- Access and egress characteristics (access and egress modes, parking location and cost, monthly transit costs, etc.).

Surveys were conducted on all 12 WSF routes during the period from October 14, 2006 to November 4, 2006. A total of 31,663 questionnaires were distributed to boarding passengers age 16 years or older. A total of 13,801 questionnaires were returned, for an overall participation rate[10] of 44 percent. Of the returned questionnaires, 11,844 were completed for analysis purposes (i.e., usable) after data cleaning and geocoding, for a completion rate[11] of 37 percent.

2.1 SURVEY DESIGN

The questionnaires were developed jointly by WSF and its stakeholders and the consulting team and consisted of 26 core questions grouped into five categories:

- Origin and Destination (four questions);
- Ferry Trip Characteristics (nine questions);
- Access, Egress, and Boarding Characteristics (four questions);
- Demographics (seven questions); and
- Follow-up Research and Comments (two questions).

An insert to the main questionnaire also was used on the routes passing through Seattle terminals (Seattle-Bainbridge, Seattle-Bremerton, Seattle-Vashon Passenger-Only, and Fauntleroy-Vashon-Southworth routes). The insert was developed at the request of the Colman Dock project team and included seven questions collecting more detailed information about access and egress modes and options, and personal safety and policy issues.

Surveys were sequentially numbered with bar codes to manage the distribution of questionnaires and monitor the number of returned questionnaires. The survey instrument (both the main survey and the insert) included a postage-paid return-mail option for passengers who were unable to complete the questionnaire while on-board the ferry or who chose to fill out the questionnaire at a later date. The questionnaire also featured a contest for participants to register to win a pair of round-trip airline tickets to any location in the continental United States. These tickets were awarded to a lucky couple of

[10] Participation rate is defined as returned questionnaires/distributed questionnaires. This definition is used to be consistent with prior surveys.

[11] Completion rate is defined as usable questionnaires/distributed questionnaires. Similar to the participation rate, completion rate is defined in this way for consistency with prior surveys.

respondents. The contest helped to increase completion rates, capture valid home address location information, and capture valid telephone or e-mail contact information for follow-up.

Three questionnaires were produced for the study (see Appendix A). There was the core questionnaire that was used on most routes and two alternate versions:

- Core routes (Edmonds-Kingston, Keystone-Port Townsend, Mukilteo-Clinton, Point-Defiance-Tahlequah, Seattle-Bainbridge, and Seattle-Bremerton);
- Alternate for Anacortes-San Juan Islands-Sidney routes; and
- Alternate for Fauntleroy-Vashon-Southworth and Seattle-Vashon Passenger-Only routes.

The alternate versions included additional questions that captured origin and destination terminals on sailings serving multiple terminals. The insert that was included with the Seattle-Bainbridge, Seattle-Bremerton, Seattle-Vashon Passenger-Only, and Fauntleroy-Vashon-Southworth routes also is provided in Appendix A.

2.2 SURVEY PROCEDURES AND METHODOLOGY

The survey was conducted as a self-administered survey using paper and pencil questionnaires. The sampling frame was comprised of vessel sailings on each WSF route, surveyed on a representative weekday and a representative Saturday[12] during the survey period. For weekdays, all p.m. peak sailings were sampled and a systematic sampling of p.m. off-peak sailings. For Saturdays, a.m. peak-period sailings were sampled, as well as some additional sampling of a.m. off-peak and midday sailings.

All passengers were offered questionnaires before boarding (whether they were driving or walking on) a sampled sailing. In addition, surveyors were stationed on vessels to distribute questionnaires to late-arriving passengers, to encourage and help passengers complete questionnaires, and to collect completed questionnaires. A crew of 125 surveyors were hired from temporary employment agencies to distribute and collect the surveys, and to count all boarding passengers (16 years and older) and all boarding vehicles.

In mid September 2006, a pilot test of the survey was conducted to test field procedures and the survey questionnaire (content, design, layout, etc.). The pilot test results were reviewed and resulted in some changes in procedures, as well as to modify and add some questions to address comments and suggestions

[12] Weekend data was collected on Sundays in the previous 1999 survey and on Saturdays for this 2006 survey because Saturdays offered higher ridership and more resident travel than Sundays.

from stakeholders.

Surveys were conducted for 12 ferry routes on weekdays and weekends for the universe of sailings identified in Table 2.1. There were over 300 sailings surveyed during the three week survey period. Of the 51, 170 total boardings, 31, 663 questionnaires were distributed and 13,801 surveys were returned.

Table 2.1 Boardings and Sailings Universe

Boarding Mode	Weekday P.M. Peak	Weekday Off-peak	Saturday	Total
Vehicle Driver	10,092	3,195	9,421	22,708
Vehicle Passenger	4,681	1,389	6,918	12,988
Bus Driver	48	16	7	71
Bus Passenger	600	234	66	900
Truck	300	122	99	521
Motorcycle	560	66	171	797
Bicycle	351	28	92	471
Walk-on	7,624	1,076	4,014	12,714
Total Boardings	24,256	6,126	20,788	51,170
Questionnaires Distributed	14,869	4,389	12,405	31,663
Questionnaires Returned	6,899	1,456	5,446	13,801
Total Sailings	123	57	121	301

A web-based field management system was used to manage the field effort. The system contained information on all weekday and weekend sailings and enabled the field managers to record serial numbers for questionnaires distributed on each sailing by mode. Hand-held electronic palm devices were used to record passenger and vehicle counts by mode. This management system provided a more accurate recording of counts and questionnaires recorded for each sailing than had been possible in the past.

2.3 SURVEY EXPANSION

The 2006 survey was expanded to match average annual counts from ticket counts for both weekday and weekend, and to include round trips by time period rather than one-way trips. This procedure created survey trips in all time periods (not just p.m.) and almost doubles the total trips in the survey. The actual number of "reversed" records was 9,394 (79 percent of all usable records). To fill out the survey day appropriately, the reversed records were paired on the basis of sailing time period as follows:

- Night paired with a.m. off-peak;
- A.M. peak with p.m. peak;

- A.M. off-peak with p.m. off-peak;
- P.M. off-peak with a.m. off-peak; and
- P.M. peak with a.m. peak.

The assumption was that the respondents would return to their starting location, so origins and destinations were reversed for the reversed records. The interisland San Juan Island trips were not included in the expansion because the data collected did not provide enough samples for reasonable expansion and the detailed counts needed for expansion were not available for this analysis.

Approximately 80 percent of all respondents returned on the same day, but those returning on a different day required additional processing. All of the reversed weekday trips remained weekday trips. However, 412 of the weekend trips indicated that the first leg of their journey had taken place on the previous day (Friday), so the reversed records were coded as weekday trips.

After all the reversed records had been generated, the expansion process began. Initially, we hoped to match commercial vehicle records to commercial vehicle counts, but ultimately there were too few survey records. On several routes, no commercial vehicle records were collected, making expansion along this dimension problematic and in certain cases impossible. Ultimately, we simply controlled for motorized and non-motorized counts during expansion. Generally, there were sufficient weekday trips to span all time periods. However, due to the survey collection process for the weekend, certain time periods (typically Night) do not have any valid records. This meant that in order to match daily counts, the counts from the unobserved time periods had to be covered by observed time periods. Thus, time period information for weekend travel is unreliable, though the daily totals are accurate. The final weights allow us to generate records for the region that match the counts by time period, direction, route, and boarding mode (motorized or non-motorized), although weekend time period information was not considered to be reliable so was not provided separately in this report.

It is difficult to present the full range of expansion factors based on this process, since there are different factors for each time period, direction, route, and boarding mode. Instead of reporting the expansion factors, we report the results of the expansion factor process in Table 2.2. This table presents the average expansion weights calculated as total number of expanded trips divided by total number of survey trips that are a result of the expansion factor process. Table 2.2 also presents the number of survey responses and the average annual counts used in the expansion factor process.

Table 2.2 Average Expansion Factor Weights

	Number of Survey Responses			Average Annual Counts			Average expansion weight		
	Weekday	Weekend	Total	Weekday	Weekend	Total	Weekday	Weekend	Total
Anacortes-San Juan Islands	1,129	693	1,822	4,243	4,955	9,198	3.76	7.15	5.05
Anacortes-Sidney	131	315	446	262	344	606	2.00	1.09	1.36
Edmonds-Kingston	1,311	1,372	2,683	11,068	13,932	25,000	8.44	10.15	9.32
Fauntleroy-Southworth	1,084	317	1,401	2,920	2,122	5,042	2.69	6.69	3.60
Fauntleroy-Vashon	1,364	615	1,979	5,718	5,382	11,100	4.19	8.75	5.61
Mukilteo-Clinton	2,771	776	3,547	10,891	11,911	22,802	3.93	15.35	6.43
Pt Defiance-Tahlequah	393	270	663	1,769	1,991	3,760	4.50	7.37	5.67
Port Townsend-Keystone	648	576	1,224	1,864	2,692	4,556	2.88	4.67	3.72
Seattle-Bainbridge	3,023	1,266	4,289	18,174	16,584	34,758	6.01	13.10	8.10
Seattle-Bremerton	1,909	794	2,703	6,902	5,982	12,884	3.62	7.53	4.77
Seattle-Vashon Passenger-Only	368	0	368	497	0	497	1.35	n/a	1.35
Southworth-Vashon	140	46	186	519	552	1,071	3.71	12.00	5.76
Total	14,271	7,040	21,311	64,827	66,447	131,274	4.54	9.44	6.16

This expansion process for the 2006 survey is quite a bit more detailed than the expansion process for 1999. The 1999 survey was expanded to match survey day counts (using the most reliable information for each route, which in some cases was boarding counts and other cases ticket counts) for weekday. Expansion factors were by route and time period. However, weekend surveys were not expanded due to issues with counts. In order to generate daily ridership for weekdays, the expanded survey results were doubled. The 1993 survey was not expanded. Thus, the comparison with the 1999 and 1993 surveys were done entirely on distributions of travelers, rather than on total ridership.

2.4 SURVEY PRECISION

Response rates for this survey were calculated as the ratio of the number of returned questionnaires to the total number of questionnaires distributed. This is defined as the participation rate and was 44 percent for the full survey.

Usable questionnaires are the total number of questionnaires that have complete information for analysis, defined as the completion rate. These were defined initially as all records with sufficient data on origin and destination, trip purpose, boarding method, vehicle type/occupancy, and access/egress mode. Confidence intervals were calculated to the 95 percent level of confidence using the following equation:

$$\textbf{Confidence Interval for Population Proportion} = [1.96^2 * (PQ/n)]^{1/2}$$

Where:

- P is the population percent;
- Q is 100-P; and
- n is the sample size.

Since PQ is largest when they are equal to 50 percent, the confidence interval has been set to a conservative choice, where P is set as close to 50 percent as is likely to occur. This confidence interval is most critical for the p.m. peak-period, so these are presented in Table 2.2. Weekday p.m. off-peak completion rates were 24 percent, with a confidence interval of +/-1.1 percent. Saturday completion rates were 22 percent, with a confidence interval of +/- 1.5 percent. These are documented by route in the Nustats *Technical Report of Methods.*[13]

[13] NuStats, *2006 Travel Survey Technical Report of Methods Final Report*, prepared for the Washington State Ferry System, February 2007.

Table 2.3 Weekday P.M. Peak Usable Samples and Precision Estimates

Route	P.M. Peak Boardings	P.M. Peak Usable Questionnaires	Response Rate	Confidence Interval for Population Proportion[a]
Anacortes-San Juan Islands-Sidney	1,006	409	41%	± 4.8
Edmonds-Kingston	3,515	523	15%	± 4.3
Fauntleroy-Southworth[b]	1,637	502	31%	± 4.4
Fauntleroy-Vashon[b]	2,105	567	27%	± 4.1
Southworth-Vashon [b]	370	62	17%	± 12.4
Keystone-Port Townsend	520	231	44%	± 6.4
Mukilteo-Clinton[b]	5,871	1,164	20%	± 2.9
Point Defiance-Tahlequah	610	146	24%	± 8.1
Seattle-Bainbridge	5,480	1,319	24%	± 2.7
Seattle-Bremerton	2,747	952	35%	± 3.2
Seattle-Vashon Passenger-Only	395	212	54%	± 6.7
Total	24,256	6,087	25%	± 1.3

[a] For the 95 percent confidence level based on number of p.m. peak usable questionnaires, assuming it is applied to binomial proportions.

[b] Based on two weekday survey days.

2.5 SURVEY DATA FILES AND GEOCODING

2.5.1 TAZ and District Schemes

Transportation Analysis Zone is the smallest geographic unit used to analyze the origin-destination data. WSF TAZ system, developed for travel demand forecasting, consists of 1,014 zones that covers 12 counties in western Washington State and British Columbia of Canada. The core of WSF TAZ system is the PSRC four-county system which has 938 zones in King, Snohomish, Pierce, and Kitsap. There are eight counties in the outlying areas that include San Juan, Island, Skagit, Whatcom, Thurston, Mason, Jefferson, and Clallam. The number of zones defined in each county reflects its land use density, socioeconomic importance, and relevance to ferry travel shed. The TAZ numbering system is presented in Table 2.3 and Figure 2.1.

For better analyzing and presenting survey results, TAZs were aggregated into origin districts. Although a general district scheme for the entire system was developed, emphasis was placed on developing route-specific district schemes for travel analysis presentation at the route level. The route-specific districts were sized with the intention of providing more detail closer to the route's ferry terminals. The district boundaries, names, and numbers by route can be found in the route specific analysis in later chapters. Figure 2.2 is an example of district system for systemwide analysis. The complete details of district scheme and maps are presented in Appendix B – TAZ and District System.

Table 2.4 WSF Transportation Analysis Zone System

Zone Number Range	Area Coverage	Number of Zones
1-530	King County	530
531-671	Snohomish County	141
672-877	Pierce County	206
878-938	Kitsap County	61
957-976	Thurston County	20
977-987	Skagit County	11
988-991	Island County	4
992-994	Mason County	3
995-998	Jefferson County	4
999-1002	Clallam County	4
1003-1006	San Juan County	4
1007-1009	Whatcom County	3
1010-1014	British Columbia	5

Figure 2.1 WSF 12-County TAZ System

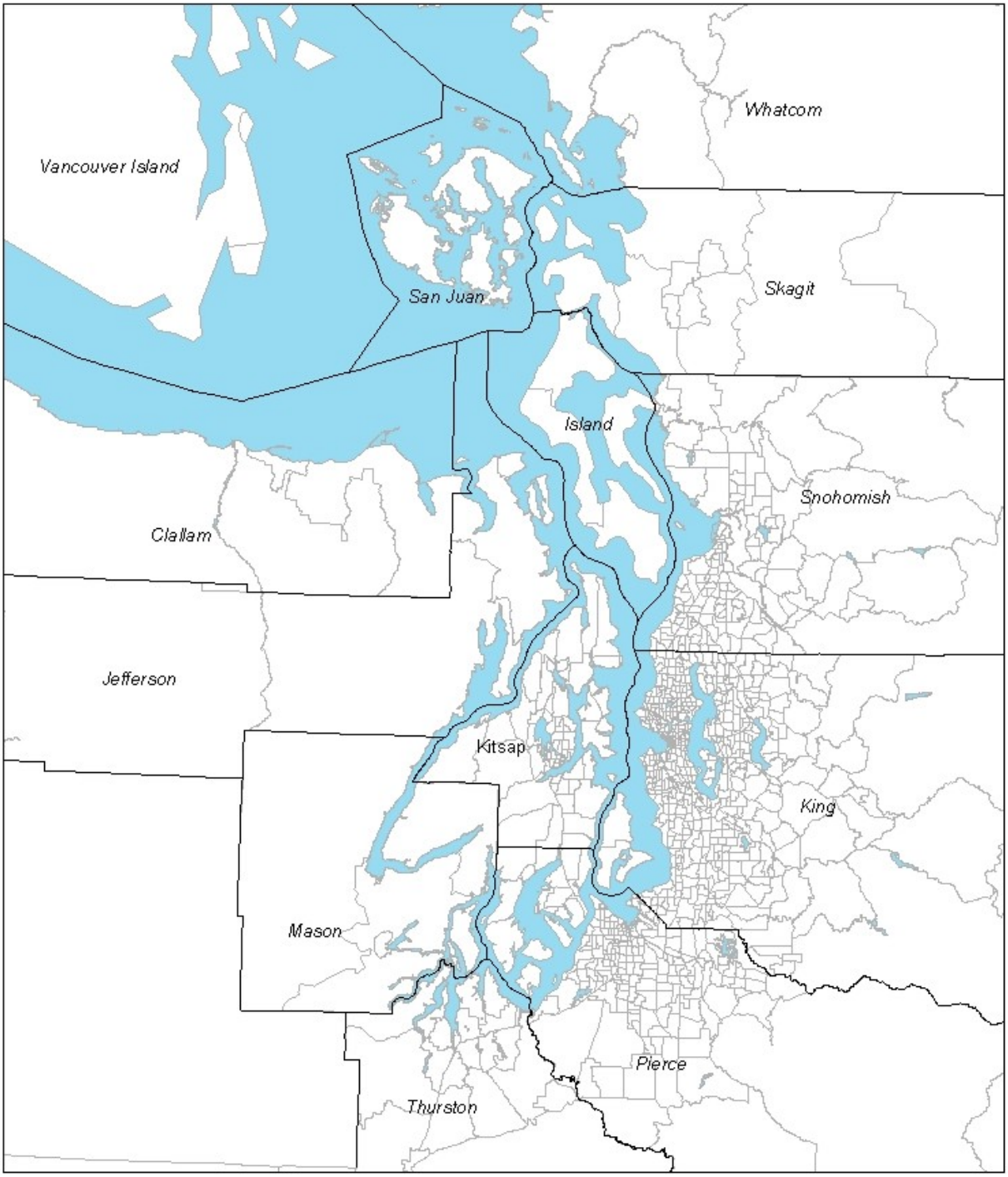
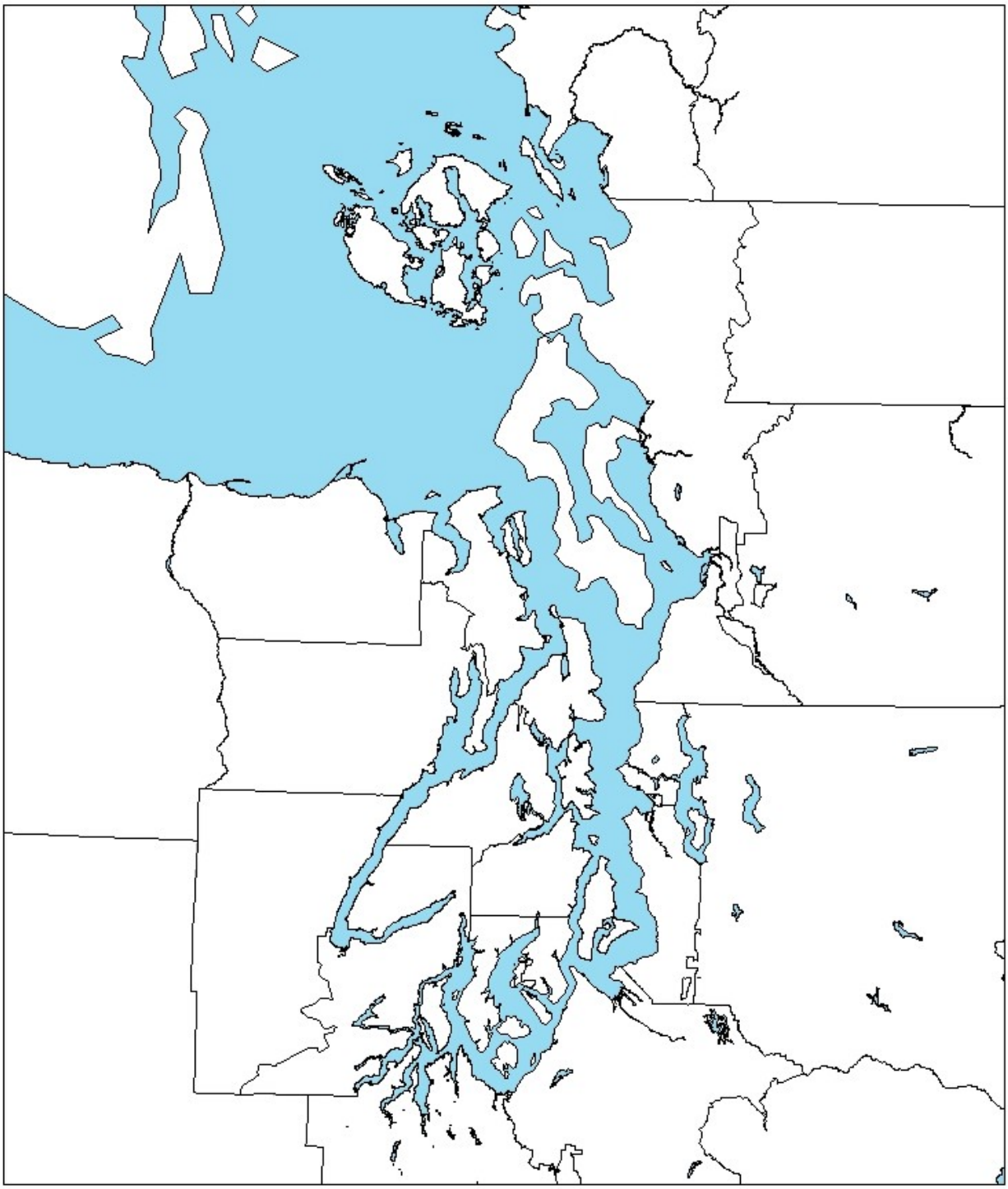


Figure 2.2 Districts for Systemwide Analysis



2.5.2 Geocoding Procedures and Results

Geocoding is a process to locate the X-Y coordinates for origin and destination address reported by the ferry riders. Once origin and destination points are geocoded, they can be easily analyzed on a geographic information system (GIS) to perform spatial analysis, such as: define TAZ; calculate travel distance and time; and mapping. Geocoding is also a process to check the data quality by viewing the data for its reasonableness of origin and destination locations, given other characteristics about the trips. The result of this validation process is summarized in Table 2.5. The process is further described below.

Table 2.5 Geocoding Result Summary

Validation Category	Number of Records	Percent of Total
Unknown origin and destination	392	2.84%
Unknown destination	876	6.35%
Destination outside study area	241	1.75%
Unknown origin	155	1.12%
Origin outside study area	300	2.17%
Origin and destination are identical	220	1.59%
Origin and destination are too similar	213	1.54%
O/D are inconsistent with sailing direction	165	1.20%
Unusable O/D	2,562	18.56%
Usable O/D	11,239	81.44%
Survey dataset	13,801	100.00%

Geocoding was an iterative process that involves both automated and manual procedures. It also used multiple GIS-based tools such as, ArcView/ArcMap, UFOSNET, and web-based address finder, including Google map, Microsoft map, and USPS zip code lookup. The geocoding steps can be briefly described as follows:

- Set up geocoding platform in ArcView using census TIGER street database. Then let ArcView automatically matches home, origin, destination, and work addresses. This step was repeated several times with reduced precision tolerance in each until no more addresses can be matched. In all, 40 percent of addresses were matched successfully in this automated procedure.
- After automated matching was done, the remaining addresses were manually matched in UFOSNET. It was done by interactively approximating the locations of addresses on screen, one at a time, using all available information reported by riders. For incomplete or wrongly coded addresses, corrections were attempted using local knowledge and other relevant data from the survey. The ultimate goal of this manual process was

to ensure addresses can be assigned correct TAZs, even though they might be off by several blocks.

- The last step in geocoding was to view geocoded address points on screen and overlaying TAZ and district boundaries. Each origin-destination pair was checked for reasonableness in association with other survey characteristics such as, ferry route, sailing direction, and boarding terminals. The result was used to determine the usefulness of an O-D record for the subsequent travel pattern and market analyses.

2.6 SURVEY SUMMARY TABLES

All of the survey summary tables provided in this report are based on expanded survey results, as described in Section 2.3. All of the summary tables are separated by the weekday and weekend surveys, except for time-of-day summaries, where weekend summaries were not included because of concerns with the validity of the weekend time period expansion factors.

The aggregate data summaries from the 1993 and 1999 surveys are also provided in each table so that changes occurring between 1993, 1999, and 2006 surveys can be identified. Although the expansion procedures for 1993 and 1999 are different than in 2006, we felt that these comparisons were still valuable. There are notations where we felt that the differences in expansion procedures might be contributing to differences in the survey results.

There are also notations at the bottom of each tabular summary identifying the total number of weekday and weekend expanded survey trips included in the analysis and the total number of surveys that had missing trips for that summary. Missing trips are those surveys where the respondent did not answer that particular question, expanded using the expansion procedures documented in Section 2.3.

2.7 SURVEY CONCLUSIONS

There were more than 50,000 boarding passengers age 16 or over during the survey period, and 13,801 surveys were returned, with 11,844 completed and available for analysis purposes. This resulted in an overall response rate of 23 percent, which is good for this type of survey.

There were some important lessons learned during the survey process that are summarized as follows:

- Surveyors are critical to the success of the project, screening, training and keeping surveyors throughout the survey period are important for consistency and efficiency.
- The prize drawing was important to generate interest in the survey and encourage riders to complete the home address and contact information. This was used to follow-up with 1,000 respondents who provided unclear or

incomplete information. These records were restored to usable status following this process.

- The use of improved field technology (hand-held devices, bar-coded questionnaires, scanning in the field and the web-based management system) allowed the team to complete the survey in less time and allowed the delivery of the processed data in a shorter time period (three months in 2006 compared to nine months in the 1999 survey).

A few of the questions yielded unreliable results, indicating that the respondents misunderstood the questions. For example, some respondents misunderstood the question on how many travelers in different age groups are traveling with you (not including yourself). This question was subsequently reprocessed to include the respondent.

Finally, while the response rate for the survey was 23 percent (higher than average for surveys of this type), it still only represents one in four sampled passengers. Given this result, it is important to consider the nonparticipants in the study and whether there are any systematic biases in the results.