

SOURCE CONTROL AND MOBILE
BUSINESSES:
BUSINESS LISTING GENERATION
Technical Memorandum

Prepared for: King County and Washington State
Department of Ecology

Project No. 200212 • December 2022 FINAL





SOURCE CONTROL AND MOBILE BUSINESSES: BUSINESS LISTING GENERATION

Technical Memorandum

Prepared for: King County and Washington State
Department of Ecology

Project No. 200212 • December 2022 FINAL

Aspect Consulting, LLC

A blue ink signature of James J. Packman, written in a cursive style.

James Packman, PMP
Associate Hydrologist
jpackman@aspectconsulting.com

A blue ink signature of Parker Wittman, written in a cursive style.

Parker Wittman
Senior Associate Data Scientist
pwittman@aspectconsulting.com

V:\200212 KC Mobile Businesses SAM Study\Deliverables\Business List Process Memo\Business Listing memo_FINAL.docx



Contents

Acronyms	ii
Executive Summary.....	ES-1
1 INTRODUCTION.....	1
1.1 Mobile Businesses and Source Control SAM Study	1
1.2 Regulatory Requirements and Background	2
1.3 Task 3: Business Listing Generation	3
2 DATA SOURCE AND SEARCHING.....	4
2.1 Data Source: Business License Data	4
2.1.1 Data Availability	4
2.1.2 Data Usability	5
2.2 Data Analysis: Mobile Business Fingerprinting	6
3 RESULTS AND MOBILE BUSINESS IDENTIFICATION	8
3.1 Results	8
3.2 Mobile Business Identification Tool.....	9
4 DISCUSSION.....	10
4.1 Key Outcomes.....	10
4.2 Recommendations.....	10
References	12
Limitations.....	13

List of Tables

1	BLS Partner Cities in King County
2	Permit Appendix 8 NAICS with Risk Ranking
3	Keywords and Weighting Criteria
4	Search Results by NAICS
5	Search Results by City

List of Appendices

A	Appendix 8 of the Municipal Stormwater Permits: Businesses and Activities that are Potential Sources of Pollutants
B	Instructions for Downloading Data from the DOR BLS Data Portal

Acronyms

Aspect	Aspect Consulting, LLC
BIG	Business Inspection Group
BLS	Business Licensing Service
BMP	Best Management Practice
DOR	Washington State Department of Revenue
Ecology	Washington State Department of Ecology
ETP-ID	Endorsement Type Identifier
IDDE	Illicit Discharge Detection and Elimination
MS4	Municipal Separate Storm Sewer System
NAICS	North American Industry Classification System
NPDES	National Pollutant Discharge Elimination System
QA	Quality Assurance
SAM	Stormwater Action Monitoring
SCIP	Source Control Inspection Program
SIC	Standard Industrial Code
Source ID	Source Identification
Study	This SAM study
TAC	Technical Advisory Committee
WSC	Washington Stormwater Center

Executive Summary

This technical memorandum provides a summary of a search procedure developed to identify mobile businesses in business license lists. In this context, mobile businesses are defined by travel to customer sites to provide services, and a subset of mobile businesses is of interest for managing stormwater. The search procedure was developed as part of a [Stormwater Action Monitoring SAM study](#) (Study) funded by municipal stormwater permittees in western Washington and administered by the Washington State Department of Ecology (Ecology).

The search procedure helps to identify mobile businesses that are of interest to stormwater source control. Source control refers to efforts to prevent, treat, or reduce stormwater pollutants at their source. In the context of stormwater management, any commercial enterprise, including businesses that are mobile in nature, are of interest for source control if there is potential to cause pollution by the business activities (i.e., illicit discharges as defined in the Ecology municipal stormwater permits).

The search procedure is based on trying to identify mobile businesses using combinations of keywords that may be present in the business's own narrative description of its services as provided in the business license application. A list of 66 keywords or phrases was developed for the search procedure, 50 of which indicate a possible mobile business (e.g., mobile repair, carpet cleaning, and maintenance), and 16 of which indicate a mobile business not of interest to stormwater source control (e.g., wireless, mobile home, and ride-based businesses like Uber). Essentially, this "fingerprint" search leverages a business's own description of its service to identify the likelihood that the business is mobile in nature.

The source data that was searched was provided by the Washington State Department of Revenue (DOR) in July 2022. The data represent business licensees for the 25 cities in King County participating in DOR's Business License Service (BLS) at that time. The data were obtained via a peer-to-peer request between the Washington State Department of Ecology (Ecology) and DOR. The business license data available from BLS does not represent all cities, however, since some cities manage their own business license records or use a third-party service, such as FileLocal. Other data sources of business lists were explored, but ultimately just business license data was used for this effort.

Our procedure demonstrates that the information collected for business licensure is sufficient to identify mobile businesses of interest to stormwater. In total, nearly 30,000 possible mobile businesses were identified among the 25 BLS cities in King County. The search procedure is being prepared as a spreadsheet tool for use with the Study's Pilot Program. The spreadsheet tool will also provide a resource to aid permittees with developing and updating their site inventories (per permit requirements).

The type of keyword "fingerprint" search used here could also be used with other business lists and to help with several aspects of permittee site inventory development and updating. Potential uses include:

ASPECT CONSULTING

- Searching business listings from other sources (e.g., FileLocal) that contain business services descriptions
- Identifying specific types of businesses that use common words to describe their business activities.
- Prioritizing sites based on businesses activities; for example, distinguishing “refuelers” who deliver fuel to sites versus fixed refueling stations.
- Catching businesses that may have slipped through other listing procedures because they selected the wrong NAICS code (or didn’t select one at all).
- Using a customized list of keywords to identify other general business types

In addition, the guidance available by Ecology and others for how to use business license data in developing site inventories would benefit from being revisited and updated.

Recommendations in this regard include:

- Revisit the list of NAICS codes in the Permit Appendix 8 that are of interest to stormwater source control; provide guidance on how to assemble, sort, and prioritize the qualifying sites.
- Provide alternative guidance for site inventory development for entities that don’t collect or wish to use business license records.
- Continue developing the peer-to-peer relationship between Ecology Stormwater programs and DOR; this could be helpful in future revisions to the municipal stormwater permit, the business licensing process, and with obtaining business license data for local government efforts to create and update site inventories for stormwater management.

1 INTRODUCTION

This technical memorandum provides a summary of task work for a Stormwater Action Monitoring (SAM) study on stormwater source control and mobile businesses. The SAM studies are administered by the Washington State Department of Ecology (Ecology) and support municipal separate storm sewer system (MS4) permittees of the National Pollutant Discharge Elimination System (NPDES) in Washington State. Funding for the Study was provided by more than 90 western Washington stormwater permittees themselves via the SAM program. Ecology's stated goal for the SAM program is to improve stormwater management to reduce pollution, improve water quality, and reduce flooding.

1.1 Mobile Businesses and Source Control SAM Study

As part of the Source Identification (Source ID) subset of SAM studies, this SAM study (Study) is focused on how permittees are implementing or plan to implement stormwater source control activities with mobile businesses. As a subset of commercial enterprises, some mobile businesses have unique challenges related to stormwater and pollution prevention due to their mobile nature and working at dispersed sites in multiple jurisdictions.

The overall goal of the Study is to provide insight and tools to help permittees identify, inspect, and coordinate in regard to mobile businesses and municipal source control activities. The Study also directly addresses two priority topics identified by permittees and other stakeholders during the proposal and planning stages of the current round of SAM studies:

- What additional regional or statewide regulatory systems or approaches would likely support local government oversight of mobile businesses that discharge waste to the MS4?
 - How can the business licensure process and requirements support proper waste handling?
 - What are barriers to proper handling of waste?
- Evaluate the IDDE data reported by permittees and gather additional information needed to identify mobile and other multi-jurisdictional business violations, to support coordinated and effective multi-jurisdiction enforcement.

A working definition of mobile businesses in the context of stormwater source control was developed in the Study survey (Aspect, 2022a). The definition was refined in discussion with the Study's technical advisory committee (TAC). The current working definition includes the following:

Mobile businesses are defined by travel to customer sites to provide services, which typically include sites outside of the businesses' home jurisdiction. Like

any commercial enterprise, mobile businesses are of interest to stormwater management if there is potential to cause illicit discharges. This includes via business activities that generate, store, transfer, treat, and/or discharge pollution-generating material or waste that could enter the municipal storm drainage system or natural water bodies. Examples include mobile equipment repair services, mobile vehicle repair and detailing, landscaping services, onsite carpet cleaning and pet groomers, untreated pressure washing waste, and exhaust hood cleaning, among others.

The outcomes from this business listing generation effort (Task 3 of the Study) will aid MS4 permittees with their site/business inventories, which permittees are required create and update.

1.2 Regulatory Requirements and Background

The NPDES municipal stormwater permits require a Source Control for Existing Development program. The program is outlined in section S5.C.8 in both of the western Washington municipal stormwater permits (Phase I and Phase II permits, Ecology 2019a and Ecology 2019b, respectively). Specifically, the permits require that:

Permittees shall establish an inventory that identifies publicly and privately owned institutional, commercial, and industrial sites which have the potential to generate pollutants to the MS4. The inventory shall include:

(a) Businesses and/or sites identified based on the presence of activities that are pollutant generating (refer to Appendix 8).

(b) Other pollutant generating sources, based on complaint response, such as: home-based businesses and multi-family sites.

The method and procedures for creating site, or business inventories are not indicated in the stormwater permits, but there is guidance available. Ecology provides Appendix 8 of the permits with lists of business types “with potential outdoor pollutant-generating sources...and should be included in the Permittee’s source control inventory” (Ecology 2019c). The lists of business types in the permit appendix are oriented around the classification codes of the North American Industry Classification System (NAICS) and the older Standard Industrial Classification (SIC) system. These codes, especially the NAICS codes, are also used for taxation determination based on where a business is located and what it does as listed in its business license. The permit’s Appendix 8 with the business inventory guidance is provided in full in Appendix A of this document.

Other relevant guidance on developing site inventories is available from the Washington Stormwater Center (WSC), especially the Source Control Inspection Program (SCIP) Guidance Manual (WSC, 2022). The SCIP Manual has a chapter on site inventory development, updates, and prioritization that provides detailed guidance, including noting multiple potential data sources that can be used to identify qualifying businesses and multiple prioritization methods to rank for pollution potential. Regarding mobile businesses, the SCIP Manual includes a recommendation for designating known active mobile businesses in the site inventory database, and supplemental material is provided with both local and non-local (California) examples of outreach items for mobile

automobile washers, food vendors, pressure washing, and mobile cleaners. However, the SCIP guidance doesn't address how to identify mobile businesses or specific best management practices (BMPs) for the range of mobile business activities that could impact stormwater.

The Business Inspection Group (BIG) report (WSC, 2020) is also relevant as background. It reports permittee survey responses to questions about their source control programs. Question 2.1 of the survey asks how permittees handle complaints for mobile businesses coming from outside their jurisdictions. While many respondents indicated they would try to contact the mobile business that caused the complaint, most indicated the only recourse once the mobile business left the site was to contact the property owner or the hiring party. The BIG report also provides an appendix for Source Control Inventory Development. The appendix from the BIG report contains a list of nine prompting questions to help permittees begin their inventory development and search for qualifying sites among multiple data sources. None of the prompting questions, however, are related to mobile businesses or contain details about how to build the database of sites.

1.3 Task 3: Business Listing Generation

Task 3 of this Study is to analyze business listing data to determine a process for identifying mobile businesses of interest for stormwater source control. The process is intended to provide a method for analyzing business listing data from business license records or other publicly available business listings. The goal of the business listing identification process is to generate a list of businesses that potentially have mobile operations that would be of interest to municipal source control programs—specifically mobile businesses with the potential for generating stormwater pollution.

The business listing process was developed using a spreadsheet database and Python coding to parse business license datasets and identify likely mobile businesses. The search procedure went through an interactive trial-and-error process to try to identify a reasonable “fingerprint” of mobile business using available data and select keywords—including business descriptions, names, and regulatory codes. While in development, the procedure is being applied in this Study to test and refine the process. The final listing process will also be used to aid the mobile business source control pilot program planned in Task 6 of the Study, where it will help identify mobile businesses in participating jurisdictions. Once completed and tested, the procedure will be prepared as a spreadsheet tool for permittees to help parse their business license datasets and update source control site inventories.

2 DATA SOURCE AND SEARCHING

This section describes the data source that was used, the steps taken to identify and obtain the data, and the search procedure developed for mobile business “fingerprinting.”

2.1 Data Source: Business License Data

Among the potential business listing data sources, we focused on business license data for this work. Initially, we considered proprietary business listings (e.g., Esri/Data Axle) and the potential scraping of data from business review websites (e.g., Yelp, Google). However, we learned that most proprietary business listing datasets are largely sourced from business license data and supplemented with crowd-sourced information, such as reviews. Since permittees already have access to business license data, it seemed unlikely that they would switch to proprietary sources for which they would have to pay to obtain data (as is the case with Esri/Data Axle). Additionally, reliance on “scraped” website data would pose a considerable challenge in terms of consistency and reliability as an ongoing concern.

The maintainer and provider of business license data for most jurisdictions in Washington is the state Department of Revenue (DOR). To obtain the DOR business license data and better understand its potential uses, we established a working relationship with a representative from DOR, with whom the Study team had several meetings. Through this process, a peer-to-peer relationship was established between DOR and the Ecology Stormwater/SAM program.

Typically, business license data are obtained directly by permittees for their own jurisdiction via the DOR data portal. But for this Study, the peer-to-peer relationship between Ecology and DOR helped with our unique data request for a larger geographic area than a single jurisdiction. This was needed to capture a broad range of business types and was something only DOR could provide. We anticipate that in the longer term, the relationship between Ecology Stormwater and DOR will grow based upon the recommendations from this Study, which will be helpful for local jurisdictions who work with business license data in their source control programs.

2.1.1 Data Availability

The business license data available from DOR for municipal jurisdictions can be obtained in more than one way. The first is via a public information request for publicly available data, such as business name and address. Alternatively, DOR offers a Business Licensing Service (BLS) for cities, which includes a special website (portal) through which customer cities may search and obtain lists of business licensees in their jurisdiction. The data that cities can obtain via the BLS portal includes both public and non-public data, which includes certain business attributes that are not on the public record, including a description of the business in the applicant’s own words. In order to view, obtain, or use non-public data from business license data, the Study team members signed a *Tax and License Secrecy Clause* confidentiality agreement.

The BLS program links taxpayer business registration data with geographic location, business type, and many other attributes. However, not all cities in Washington State use

DOR’s BLS service to obtain tax and license information for businesses in their jurisdiction. Some cities, such as Seattle, are customers of FileLocal instead—a private registrar of business license information—which they use to obtain, organize, and report on business licensing. FileLocal data are not used in this Study, thus, the data analyzed as part of this study do not represent cities that use the FileLocal service instead of DOR’s BLS service. We consider the DOR BLS data to be sufficiently representative of business license information for the purposes of developing a mobile business search process.

2.1.2 Data Usability

The BLS data from DOR was reviewed for relevance to the Study and potential usefulness in identifying mobile businesses. While the distinction between mobile and fixed location businesses is not directly captured in business license data, we identified three data fields within the DOR business license database that, individually or in combination, would be useful in identifying mobile businesses.

1. **NAICS Codes**, a public data field of the federal standard business establishment classification system used in the collection and evaluation of statistical data related to the U.S. economy. At varying degrees of accuracy and resolution, these codes can be used to identify business types and associated likely activities.
2. **An Endorsement Type Identifier (ETP-ID) specific to “non-resident” businesses**, a public data field that indicates that a business is applying for a license in a city in which their business does not reside—which is an indicator that a business is located outside the licensed jurisdiction to do work. While a non-resident endorsement alone does not indicate a mobile business, it would be a likely attribute of the type of mobile businesses of interest to this study assuming they applied for business licenses in the cities in which they operate.
3. **The Product Description field**, a non-public data field that provides the applicant-entered description of their business. This was the data field that made the keyword search possible. Using non-public data has constraints, however, because the information is confidential and not allowed to be publicly disseminated. Specifically, DOR requires that BLS partner agencies sign a confidentiality agreement and limit the distribution of non-public information (*Secrecy Clause* mentioned above). The agreement allows for non-public data to be used within a BLS partner agency for internal purposes. Additionally, because business license data are user-entered and unconstrained, information with typos, jargon, or other inaccuracies exist. Nonetheless, for many businesses, it is a useful field for understanding the type of business and potential activities.

In addition to meetings and email communications with DOR, we also had a screenshare session with one of the Study TAC members from the City of Redmond (a BLS partner) to better understand the user interface and potential use of the BLS online portal. Screenshots from that session are provided in Appendix B of this memorandum (only public information shown). The screenshots include annotation and instructions for BLS partners to follow to obtain business licensing data in the necessary format to use with the spreadsheet tool being developed.

Ultimately, the team requested a dataset from DOR that included business license information for all BLS partner cities in King County (25). The upcoming pilot project of the Study is focused in King County, and, as such, King County was selected as the geographic scope of the BLS data to obtain. Additionally, given its relative size, King County represented a reasonably large (but manageable) area, while still providing a representative dataset with which to work. Table 1 provides a summary of the number of business license records received from DOR (83,319 total) and a breakdown of how many records indicated with a “non-resident” endorsement (42,815) that were searched.

2.2 Data Analysis: Mobile Business Fingerprinting

Once the relevant data fields from the business license dataset were identified, the team developed a process to search the data fields by keywords and sort the data to identify a “fingerprint” of mobile businesses. This involved the following seven steps:

1. **Extracted the records with non-resident status in the ETP-ID field**
 - a. Converted the list of endorsement codes into parsable tables using Pandas (a Python data analysis library).
 - b. Filtered out the non-resident business types and gathered the applicable ETP-ID for each target city.
2. **Sorted and highlighted the filtered dataset by a risk ranking applied to the Permit Appendix 8 table of NAICS codes** (see Table 2). The ranking of relatively higher, medium, or lower risk was based on the experience and best judgment of the team.
 - a. Higher Priority NAICS Codes: 16 NAICS codes representing 7 business categories with perceived high risk for stormwater pollution generating activities, including transportation-related businesses like gas stations or auto repair and those based on outdoor business operations.
 - b. Medium Priority NAICS Codes: 40 NAICS codes representing 24 business categories with moderate risk for stormwater pollution generating activity, including manufacturing, construction, and product manufacturing or equipment rental.
 - c. Lower Priority NAICS Codes: 3 NAICS codes representing 3 business categories with perceived low risk of stormwater pollution generating activities, including printing, food service places, and museums.
3. **Sorted the dataset so that businesses with potential hazardous material use/production are spotlighted** (HazardousFlammableMtrl field = TRUE)

Note: the use of hazardous materials by a business triggers additional regulations beyond stormwater. The search output can be sorted on this field to see which businesses may use hazardous materials and may fall under additional regulations for pollution prevention. As noted, when this field was “true” the record was highlighted, but the search was not limited to just those businesses with hazardous materials.
4. **Created a keyword list** to identify keywords in the Product Description field to better isolate potential mobile businesses.

5. **Assigned a normalized weight between -1.0 and 1.0 for each keyword** (negatives for exclusions, positive for inclusions, assigned in increments of 0.25)
6. **Applied keyword search scores to each business** to reflect the number of keyword occurrences
 - a. Search rating = the sum of non-weighted keyword occurrences
 - b. Weighted search rating = the sum of the weighted keyword occurrences

The list of keywords developed for the search are provided in Table 3. These keywords were identified through multiple sources, including using commonly known mobile business types and the results of the Study survey (Aspect, 2022a) and IDDE data analysis (Aspect, 2022b), in which types and characteristics of mobile businesses of interest to source control were identified. In addition, a natural language machine learning algorithm (KeyBERT) was used with the data, which provided a rating value associated with the frequency of occurrence of each keyword. The final list of keywords was developed through an iterative process of reviewing search results and identifying additional keywords or phrases (2-word combinations) and then revising and re-running the search. The number of occurrences of each keyword and the natural language rating is also provided in Table 3.

Table 3 also includes the weighting factors used to increase or decrease the final keyword search score. A relatively larger weight (closer to 1.0) reflects an increased likelihood of a business being mobile (e.g., mobile cleaners), and a relatively smaller weight (closer to 0) indicates the business is less likely to be mobile (e.g., mobile home). The list of keywords includes two types: those that potentially indicate that a business is mobile (positive weighting) and those that indicate the business is not likely mobile and should be excluded from the search results (negative weighting).

A search score for each record was determined by summing the occurrence of the keywords, and a weighted search score was likewise determined by summing the weighted occurrence of the keywords. The search scores are provided in Tables 4 and 5, which provide the minimum, average, and maximum search scores by NAICS code and by City, respectively. In addition, Table 5 provides the weighted search scores by city along with a count of the businesses that fell in each of the lower, medium, and higher risk NAICS categories.

In addition, as a quality assurance (QA) step, the keyword search output was reviewed to determine how successful the search was in identifying likely mobile businesses. The QA review was qualitative and involved reading the Product Description field to subjectively assess the likelihood of a business being mobile (or not). The review targeted up to 10 percent of the records (randomly selected) per jurisdiction with smaller cities having relatively more records reviewed than larger cities. The QA review included records with keyword search scores that were positive (likely mobile business), negative (likely not mobile business), and zero (no keywords found). Web searches were also done on selected businesses due to inconclusive information (e.g., no or minimal product description). The QA review was done iteratively in conjunction with refining the list of keywords and adjustments to the weighting criteria.

3 RESULTS AND MOBILE BUSINESS IDENTIFICATION

This section describes the results from the keyword search and how the procedure can be used with business listings.

3.1 Results

The results of the keyword search process are shown in Table 4 by NAICS codes and in Table 5 by city. Table 4 also summarizes the results by the NAICS risk levels (same as Table 2) and provides the total number of non-resident businesses along with summary statistics of the minimum, mean, median, and maximum search score by NAICS category (regardless of score value). Table 5 also includes the count of businesses, but by city, and summary statistics for just positive search scores. Both the arithmetic average (mean) and geometric average (median) were calculated, but the mean is more informative since the median values are nearly the same across the board (either 0 or 0.25 for most NAICS codes and cities) due to the large number of search scores with those values.

The results indicate many potential mobile businesses in all cities across many NAICS categories. Most NAICS categories had positive mean search scores with the highest falling under automotive dealers and repair and maintenance services along with some manufacturing business types. Notably, some of the higher risk NAICS codes had negative search scores (transportation-related). This is due to some keywords in those business descriptions with negative weighting criteria to indicate likely not a mobile business of interest to stormwater source control (e.g., trucking).

The search results summarized by city (Table 5) provides the count of businesses for each BLS city and summed by the NAICS risk level (higher, medium, and lower). Table 5 also includes summary statistics, but of just the positive search scores. The mean search scores were all between 0.39 and 0.45, indicating most business descriptions included two or more keywords. As noted above, the median score is the same for each city due to the preponderance of that value in the keyword search output. Likewise, the minimum value is also 0.25 and the same for each city because only positive search results were used to summarize by city.

The QA review of the keyword search results confirmed that nearly all records with positive scores were mobile businesses. A small number (less than 1 percent) of businesses with positive scores were deemed not likely mobile, which was due to low positive scores (e.g., 0.25) indicating just one keyword match. The larger positive scores (over 1.0) were all associated with mobile businesses of interest to source control as higher scores indicate more than one keyword match.

The QA review of records with negative keyword search scores were confirmed to be unfaillingly not mobile businesses. As businesses that had one or more keyword match, one of those keywords was on the exclusion list (see Table 3) and assigned a negative weight indicating not likely a mobile business. The review of the negative keyword results (the fewest in number) confirmed this outcome.

The review of zero values from the keyword search, however, showed a mix of most businesses not being of the mobile type of interest here, but also approximately 1 in 10 negative search scores were inconclusive. This was due to lack of keywords (and obvious business name) or typos in the business description (e.g., “service” instead of service, which was a not an uncommon occurrence).

3.2 Mobile Business Identification Tool

A spreadsheet tool is being developed to identify mobile businesses based on the search procedure developed here. The tool will be useable directly with business license records. The tool is under development and will be used during this Study’s Pilot Program (Task 6).

To aid in making the tool easy to use and work correctly, we have prepared instructions for local jurisdictions who participate in the DOR BLS program for downloading business license data in the necessary format. The instructions are in Appendix B of this memorandum and have annotated screenshots from the BLS portal. In the Study’s Pilot Program, the spreadsheet tool will be used to create lists of potential mobile businesses in each BLS partner city in King County. These lists will be shared with the BLS partner cities for their use in identifying and including mobile businesses in their source control programs.

In addition to its use in the Study’s Pilot Program, the tool’s primary audience is municipal stormwater permittees to aid them in sorting business license data when developing and updating their site inventories for source control programs. The tool is still in development, though it will likely have tabs including:

- **Instructions:** will provide basic instructions on how to use the spreadsheet tool to run a search and provide an output listing of businesses that meet the search criteria.
- **Dataset:** will have fixed headers and is ready to be populated with a user-provided dataset in a specific format and layout, such as from BLS portal.
- **Output summary tables:** will have a search score, weighted search score, and total number of businesses by NAICS code.
- **Output listing of potential mobile businesses:** will provide positive search scores and added data fields to help sort the data, including NAICS risk ranking and the weighted search score, which can be sorted to identify the most likely mobile businesses (highest scores).

The spreadsheet tool will also be useable with other business or site listings that contain the NAICS code and a text field describing the business type or activities. However, using other data sources will require formatting the data to work with the search procedure. The formatting can be done by copying the data into the Dataset tab of the spreadsheet tool and adjusting or relocating data fields to match the fixed headers associated with the search. Changing or removing data fields will cause the search to fail.

4 DISCUSSION

This section provides a discussion of the business listing search findings, including key outcomes and recommendations.

4.1 Key Outcomes

The results of the search process show a significant number of likely mobile businesses throughout the BLS partner jurisdictions in King County. The search process successfully used keywords to identify a “fingerprint” of how mobile businesses are described in business license applications.

The list of keywords used to identify potential mobile businesses was developed through an iterative procedure that identified a reasonable range of mobile business types. The keyword list included words that refer to a possible mobile business and words that indicate not a mobile business (such as “mobile home”). In addition, the keywords found in the business descriptions had a weighting applied to them to raise or lower the search score to indicate relatively higher or lower likelihood of a business being mobile and of interest to stormwater source control. Negative values for some keywords helped sift out and exclude business that may have keyword matches but are not of interest to stormwater (e.g., “cellular”).

The search process focused intentionally on “non-resident” businesses to capture the businesses that travel outside their home jurisdictions to provide services and who have some risk for stormwater pollution generation. While not all mobile businesses necessarily travel outside their home jurisdictions, the focus on businesses that are outside of their home jurisdiction provided the best subset of data in business license records on which to perform the search and test the procedure.

The outcomes from this effort support the use of business licensing records to identify mobile business of interest to stormwater source control. While not the only source of business listing data worth considering, it provides a useful basis for identifying licensed businesses that may have source control needs, including those of a mobile nature.

4.2 Recommendations

We provide recommendations below based on the outcomes of this effort. Recommendations include ways to enhance the keyword search procedure and for Ecology’s and permittees’ consideration in developing and updating source control site inventories.

- The keyword search and weighting approach could be used with other business lists and to help with several aspects of site inventory development and updating. Other business listings (such as from FileLocal) with business descriptions may be searched with this method. Potential uses include:
 - Identifying specific types of businesses using common descriptive words.
 - Prioritizing sites based on businesses activities; for example, distinguishing “refuelers” who deliver fuel to sites versus fixed refueling stations.

- Catching businesses that may have slipped through other sorting procedures because they selected the wrong NAICS code (or didn't select one at all).
- Identifying and weeding out certain business types that may seem similar but are not of interest to stormwater source control; for example, "equipment repair" would be of interest for heavy equipment repair (like some mobile repair businesses offer), but not of interest for onsite computer equipment repair.
- The keyword identification process could further include:
 - Collecting the entire product descriptions of the initially targeted business and use natural language processing / keyword parsing to parse additional potential keywords of use.
 - Using natural language processing libraries to capture synonyms and related words to the keywords provided.
 - Using other natural language processing techniques and libraries to capture wildcards, word combos, typos, etc.
- The keyword search results of zero (representing the largest portion of records) could be further explored to identify if a significant number of mobile businesses exist in that subset of the data that may have been missed by the keyword search.

Recommendations regarding Permit Appendix 8 and site inventory guidance include:

- Revisit the NAICS codes that should be included with permittee site inventories. Some of them capture unnecessary business types that make it more difficult for permittees to sift through to find businesses that do have stormwater pollution generation potential. Also, where possible, use the full 6-digit NAICS code in the appendix.
- Provide an alternative site inventory development guidance for entities that don't collect or wish to use business license records.
- Maintain and grow the peer-to-peer relationship between Ecology Stormwater and DOR. This will be helpful as Ecology continues to update and refine the source control program permit requirements.
 - If/when the Appendix 8 guidance is updated, it will help to understand what additional information in business license data (even confidential non-public data) may be useful for local governments when building site inventories.
 - It may be possible to streamline the source control site inventory development process for local governments by adding or revising features in the DOR BLS portal. Suggestions include adding an option to indicate if one is a mobile business when applying for a license, and adding features to support sorting data by stormwater management considerations.

References

- Aspect 2022a, Source Control and Mobile Businesses: Survey of Municipal Programs Technical Memorandum, Aspect Consulting, LLC, Seattle, Washington, prepared for King County and Washington State Department of Ecology, May 2022. [LINK](#)
- Aspect 2022b Source Control and Mobile Businesses: IDDE Data Analysis Technical Memorandum, Aspect Consulting, LLC, Seattle, Washington, prepared for King County and Washington State Department of Ecology, July 2022. [LINK](#)
- Ecology 2019a. Western Washington Phase I Municipal Stormwater Permit, effective 2019-2024, Washington State Department of Ecology. [LINK](#)
- Ecology 2019b. Western Washington Phase II Municipal Stormwater Permit, effective 2019-2024, Washington State Department of Ecology. [LINK](#)
- Washington Stormwater Center, 2020. Business Inspection Program Report. [LINK](#)
- Washington Stormwater Center, 2022. Source Control Inspection Program Guidance Manual. [LINK](#)

Limitations

Work for this project was performed for King County and the Washington State Department of Ecology (Client), and this report was prepared in accordance with generally accepted professional practices for the nature and conditions of work completed in the same or similar localities, at the time the work was performed. This report does not represent a legal opinion. No other warranty, expressed or implied, is made.

All reports prepared by Aspect Consulting for the Client apply only to the services described in the Agreement(s) with the Client. Any use or reuse by any party other than the Client is at the sole risk of that party, and without liability to Aspect Consulting. Aspect Consulting's original files/reports shall govern in the event of any dispute regarding the content of electronic documents furnished to others.

TABLES

Table 1. BLS Partner Cities in King County

Project No. 200212

BLS Partner City in King County	Total Business License Records	Records with ETP-ID of "Non-Resident"
Algona	347	262
Black Diamond	636	385
Bothell	3,663	2,057
Burien	5,579	3,056
Carnation	789	638
Clyde Hill	966	854
Covington	2,475	1,661
Duvall	1,187	769
Enumclaw	1,462	720
Federal Way	8,969	3,880
Issaquah	6,868	4,119
Kenmore	1,094	0*
Kirkland	12,583	5,472
Maple Valley	2,453	1,294
Mercer Island	5,588	3,927
Milton	816	435
Newcastle	2,047	1,469
Normandy Park	687	504
North Bend	1,300	694
Pacific	817	583
Redmond	6,494	3,714
Sammamish	4,515	1,772
Seatac	4,496	3,074
Tukwila	4,853	252
Woodinville	2,635	1,224
TOTAL	83,319	42,815

*Kenmore does not have a Non-Resident ETP-ID code

Table 1

Table 2. Permit Appendix 8 NAICS with Risk Ranking

Project No. 200212

NAICS Group Description	NAICS Number
<i>Support Activities for Animal Production</i>	1152
<i>Construction of Buildings</i>	236
<i>Heavy and Civil Engineering Construction</i>	237
<i>Specialty Trade Contractors</i>	238
<i>Beverage, Food, and Tobacco Manufacturing</i>	311, 312
<i>Wood Product Manufacturing</i>	321
<i>Paper Manufacturing</i>	3221, 3222
<i>Printing and Related Support Activities</i>	323
<i>Chemical Manufacturing</i>	325
<i>Petroleum and Coal Products Manufacturing</i>	3241
<i>Plastics and Rubber Product Manufacturing</i>	326
<i>Leather and Allied Product Manufacturing</i>	316
<i>Nonmetallic Mineral Product Manufacturing</i>	327
<i>Primary Metal Manufacturing</i>	331
<i>Fabricated Metal Product Manufacturing</i>	332
<i>Machinery, Computer, and Electronic Product manufacturing</i>	333, 334
<i>Electrical Equipment, Appliance, and Component Manufacturing</i>	335
<i>Transportation Equipment Manufacturing</i>	336
Rail Transportation	482
Transit and Ground Passenger Transportation	485
Truck Transportation and Warehousing	484, 493
Support Activities for Transportation	4881, 4882, 4884, 4889
Utilities	2211
<i>Wholesale Trade – Durable Goods</i>	423140, 423930, 423110, 4233, 4237, 4238
<i>Wholesale Trade – Nondurable Goods</i>	424930, 4244, 4246, 4247, 4248
<i>Building Materials, Hardware, Garden Supplies Dealers</i>	444
<i>Food and Beverage Stores</i>	445
Automotive Dealers and Gasoline Service Stations	441, 447
<i>Food Services and Drinking Places</i>	722
<i>Rental and Leasing Services</i>	5321, 5324
Repair and Maintenance	811192, 8111, 8112, 8113, 8114
<i>Ambulatory Health Care Services and Hospitals</i>	621910
<i>Educational Services</i>	6111, 6112, 6113, 6115
<i>Museums, Historical Sites, and Similar Institutions</i>	712

Legend:

Higher risk NAICS
<i>Medium risk NAICS</i>
<i>Lower risk NAICS</i>

Notes:

NAICS numbers are listed as in Permit Appendix 8.

When codes of less than 6-digits are shown, the number represents the first 5, 4, or 3 digits of the code, corresponding to successively less specific business types

Risk levels were assigned for this Study (see memorandum section 2.2)

Table 2

Table 3. Keywords and Weighting Criteria

Project No. 200212

Keywords to <u>Include</u>	Weighting Factor	Occurrences	Natural Language Rating
animal groom	0.25	13	NA
appliance repair	0.75	22	NA
arborist	0.5	43	NA
asphalt	0.25	364	0.1569
auto	0.25	483	0.1331
carpet cleaners	0.75	0	NA
cleaners	0.25	3	0.1835
concrete	0.25	846	NA
contractor	0.25	6226	0.2002
detail	0.25	118	NA
equipment	0.25	1645	0.24
fog	0.25	0	NA
food truck	1	124	0.3408
food vendor	0.25	110	NA
grease	0.25	8	NA
handyman	0.25	428	0.1777
hazardous	0.5	63	NA
herbicide	0.5	2	NA
hvac	0.25	1821	0.3142
installation	0.25	3760	NA
landscape	0.5	883	NA
maintenance	0.5	1698	0.2745
masonry	0.25	130	NA
mechanic	0.25	360	0.2488
mobile	0.5	315	NA
mobile auto	1	48	NA
mobile cleaning	1	2	NA
mobile service	1	2	NA
oil	0.25	299	NA
on site	0.5	68	NA
paint	0.25	1317	NA
pesticide	0.5	51	NA
pet groom	0.25	9	NA
petroleum	0.25	59	NA
pipe clean	0.25	0	NA
pool	0.25	152	NA
pressure washing	1	338	NA
repair	0.25	3465	0.1989
service	0.25	10964	0.291
sewer	0.25	265	0.1738
spa	0.25	186	0.1786
storm	0.25	102	NA
street	0.25	140	0.1434
street sweep	1	47	NA
transportation	0.25	246	0.3132
tree service	0.5	169	NA
utility	0.25	200	0.2139
vactor	1	16	NA
vent hood	1	0	NA
window washing	1	37	NA

Keywords to <u>Exclude</u>	Weighting Factor	Occurrences	Natural Language Rating
ambulance	-1	10	0.1836
auto + salvage	-1	0	NA
auto + wrecking	-1	4	NA
cabulance	-1	17	0.1303
cellular	-1	128	NA
drive	-1	141	0.2089
fats	-1	0	NA
fireplace	-1	179	NA
mobile + modular	-1	7	NA
mobile home	-1	9	NA
plumb	-1	2238	NA
ride	-1	19	0.1413
roadside	-1	26	0.2224
special needs	-1	25	NA
towing	-1	38	0.2964
wireless	-1	195	NA

Notes:

Keywords with 'NA' natural language ratings were not within the first 2000 keyword results from the NLP algorithm or were unable to be adequately parsed by the NLP algorithm.

Table 4. Search Results by NAICS

FINAL

Project No. 200212

NAICS Group Description	NAICS Number	Number of "Non-Resident" Businesses	Minimum Weighted Search Score	Mean Weighted Search Score	Median Weighted Search Score	Maximum Weighted Search Score
Support Activities for Animal Production	1152	8	-1	-0.38	0	0
Construction of Buildings	236	6611	-0.5	0.14	0	2.5
Heavy and Civil Engineering Construction	237	1097	-1	0.18	0	1.5
Specialty Trade Contractors	238	16209	-1	0.18	0	1.75
Beverage, Food, and Tobacco Manufacturing	311, 312	140	0	0.03	0	0.5
Wood Product Manufacturing	321	27	0	0.11	0	0.75
Paper Manufacturing	3221, 3222	2	0	0	0	0
Printing and Related Support Activities	323	30	0	0	0	0
Chemical Manufacturing	325	22	0	0	0	0
Petroleum and Coal Products Manufacturing	3241	0	NA	NA	NA	NA
Plastics and Rubber Product Manufacturing	326	5	0	0.25	0	1
Leather and Allied Product Manufacturing	316	0	NA	NA	NA	NA
Nonmetallic Mineral Product Manufacturing	327	53	0	0.08	0	1
Primary Metal Manufacturing	331	9	0	0.06	0	0.25
Fabricated Metal Product Manufacturing	332	85	0	0.12	0	0.5
Machinery, Computer, and Electronic Product manufacturing	333, 334	75	-0.75	0.12	0	0.75
Electrical Equipment, Appliance, and Component Manufacturing	335	6	0	0.04	0	0.25
Transportation Equipment Manufacturing	336	3	0.25	0.25	0.25	0.25
Rail Transportation	482	1	0	0	0	0
Transit and Ground Passenger Transportation	485	132	-4	-1.05	-1	0
Truck Transportation and Warehousing	484, 493	299	-2	-0.12	0	1.5
Support Activities for Transportation	4881, 4882, 4884, 4889	77	-2	-0.52	-0.75	0.5
Utilities	2211	54	0	0.16	0.25	0.25
Wholesale Trade – Durable Goods	423140, 423930, 423110, 4233, 4237, 4238	511	0	0.19	0	1.5
Wholesale Trade – Nondurable Goods	424930, 4244, 4246, 4247, 4248	455	0	0.05	0	0.75
Building Materials, Hardware, Garden Supplies Dealers	444	418	0	0.15	0	1.5
Food and Beverage Stores	445	193	-1	0.00	0	1.5
Automotive Dealers and Gasoline Service Stations	441, 447	92	-0.5	0.33	0.25	1.5
Food Services and Drinking Places	722	527	-1	0.08	0	1.5
Rental and Leasing Services	5321, 5324	381	-1	0.17	0	0.75
Repair and Maintenance	811192, 8111, 8112, 8113, 8114	683	-1	0.46	0.25	4.5
Ambulatory Health Care Services and Hospitals	621910	9	-0.75	-0.17	0	0
Educational Services	6111, 6112, 6113, 6115	17	0	0	0	0
Museums, Historical Sites, and Similar Institutions	712	2	0	0	0	0

Legend:

Higher risk NAICS
Medium risk NAICS
Lower risk NAICS

Notes:

- NAICS numbers are listed as in Permit Appendix 8.
- When codes of less than 6-digits are shown, the number represents the first 5, 4, or 3 digits of the code, corresponding to successively less specific business types.
- Risk levels were assigned for this Study (see memorandum section 2.2)

Table 4

Table 5. Search Results by City

Project No. 200212

City	High Risk NAICS	Medium Risk NAICS	Low Risk NAICS	Minimum Weighted Search Score	Mean Weighted Search Score	Median Weighted Search Score	Maximum Weighted Search Score
Algona	8	153	2	0.25	0.44	0.25	1.75
Black Diamond	5	237	2	0.25	0.45	0.25	1.5
Bothell	47	1419	23	0.25	0.43	0.25	2.75
Burien	94	1875	36	0.25	0.41	0.25	2.75
Carnation	18	349	10	0.25	0.39	0.25	1.5
Clyde Hill	19	571	6	0.25	0.45	0.25	2.75
Covington	62	986	22	0.25	0.45	0.25	3
Duvall	25	434	14	0.25	0.43	0.25	2
Enumclaw	15	471	13	0.25	0.42	0.25	1.5
Federal Way	112	2386	56	0.25	0.42	0.25	4.5
Issaquah	104	2269	57	0.25	0.41	0.25	2.75
Kirkland	111	3614	49	0.25	0.40	0.25	3.5
Maple Valley	43	659	33	0.25	0.43	0.25	2.75
Mercer Island	83	2531	56	0.25	0.42	0.25	3.5
Milton	10	306	0	0.25	0.43	0.25	2.25
Newcastle	37	875	18	0.25	0.43	0.25	2
Normandy Park	7	393	1	0.25	0.41	0.25	1.75
North Bend	14	421	7	0.25	0.40	0.25	2
Pacific	21	315	2	0.25	0.45	0.25	1.75
Redmond	73	2309	44	0.25	0.40	0.25	2.25
Sammamish	35	1105	13	0.25	0.44	0.25	2.75
Seatac	125	1701	40	0.25	0.40	0.25	3.75
Tukwila	87	1584	44	0.25	0.40	0.25	2.75
Woodinville	22	761	20	0.25	0.44	0.25	2.75

Notes:

Only positive search scores are included in the minimum, average, and maximum values (since negative scores indicate likely not a mobile business).

APPENDIX A

Appendix 8 of the Municipal Stormwater Permits: Businesses and Activities that are Potential Sources of Pollutants

APPENDIX 8 – Businesses and Activities that are Potential Sources of Pollutants

Use this appendix to help identify businesses and/or activities with potential outdoor pollutant-generating sources that discharge to the MS4 and should be included in the Permittee’s source control inventory, developed pursuant to S5.C.8.b.ii. The Standard Industrial Code (SIC), Major Group, and NAICS numbers are provided for reference. Permittees may include additional outdoor pollutant-generating sources that are located within their jurisdictions.

Group Description	SIC Major Group	SIC Industry Group No.	NAICS Major Group
Support Activities for Animal Production		074, 075	1152xx,
Construction of Buildings	15		236
Heavy and Civil Engineering Construction	16		237
Specialty Trade Contractors	17		238
Beverage, Food, and Tobacco Manufacturing	20		311, 312
Wood Product Manufacturing	24		321
Paper Manufacturing	26		3221xx, 3222xx
Printing and Related Support Activities	27		323
Chemical Manufacturing	28		325
Petroleum and Coal Products Manufacturing	29		3241xx
Plastics and Rubber Product Manufacturing	30		326
Leather and Allied Product Manufacturing	31		316
Nonmetallic Mineral Product Manufacturing	32		327
Primary Metal Manufacturing	33		331
Fabricated Metal Product Manufacturing	34		332
Machinery, Computer, and Electronic Product manufacturing	35		333, 334
Electrical Equipment, Appliance, and Component Manufacturing	36		335
Transportation Equipment Manufacturing	37		336
Rail Transportation	40		482

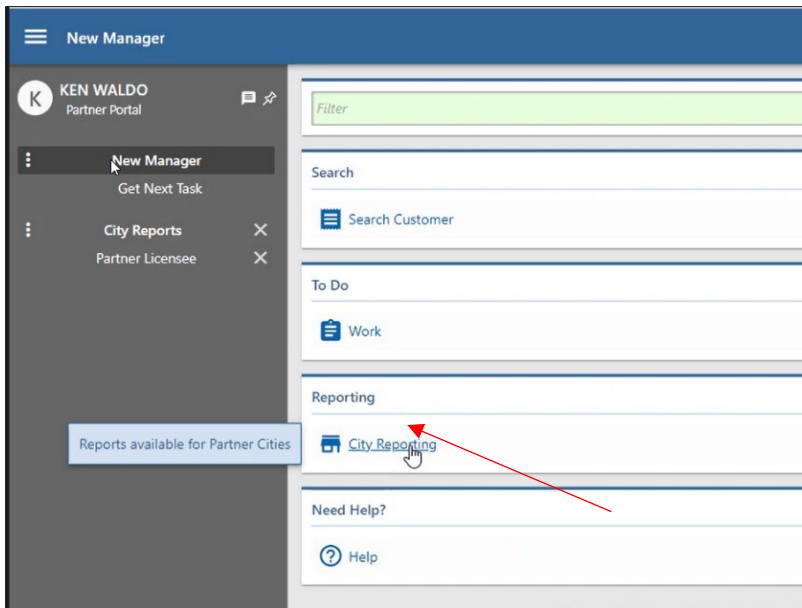
Group Description	SIC Major Group	SIC Industry Group No.	NAICS Major Group
Transit and Ground Passenger Transportation	41		485
Truck Transportation and Warehousing	42		484, 493
Support Activities for Transportation		473, 474, 478	4881xx, 4882xx, 4884xx, 4889xx,
Utilities	49		2211xx
Wholesale Trade – Durable Goods		501, 503, 505, 506, 507, 509	423140, 423930, 423110, 4233xx, 4237xx, 4238xx,
Wholesale Trade – Nondurable Goods		514, 515, 516, 517, 518, 519	424930, 4244xx, 4246xx, 4247xx, 4248xx,
Building Materials, Hardware, Garden Supplies Dealers		521, 523, 526	444
Food and Beverage Stores	54		445
Automotive Dealers and Gasoline Service Stations	55		441, 447
Food Services and Drinking Places	58		722
Rental and Leasing Services		735	5321xx, 5324xx
Repair and Maintenance	75		811192, 8111xx, 8112xx, 8113xx, 8114xx,
Ambulatory Health Care Services and Hospitals		806, 807	621910,
Educational Services	82		6111xx, 6112xx, 6113xx, 6115xx
Museums, Historical Sites, and Similar Institutions		842	712

APPENDIX B

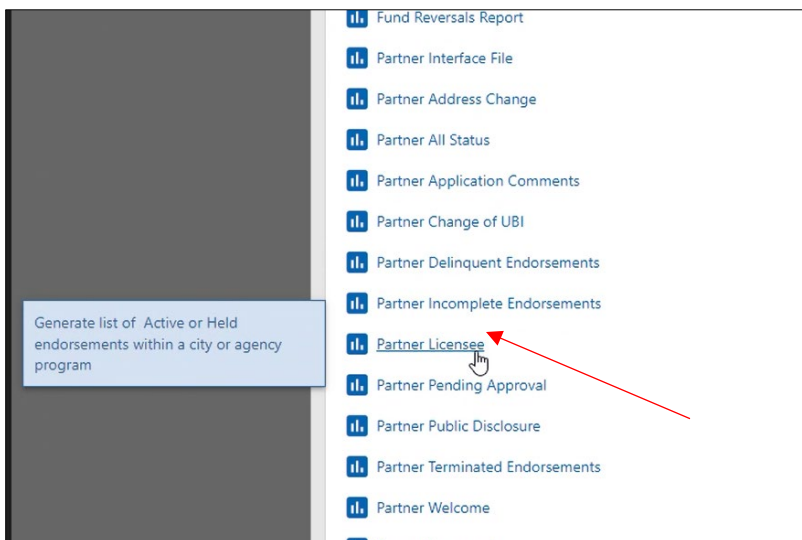
Instructions for Downloading Data from the DOR BLS Data Portal

Directions for Downloading Business Licensing Data from the Department of Revenue Business License Service For Use with the Mobile Business Identification Spreadsheet Tool

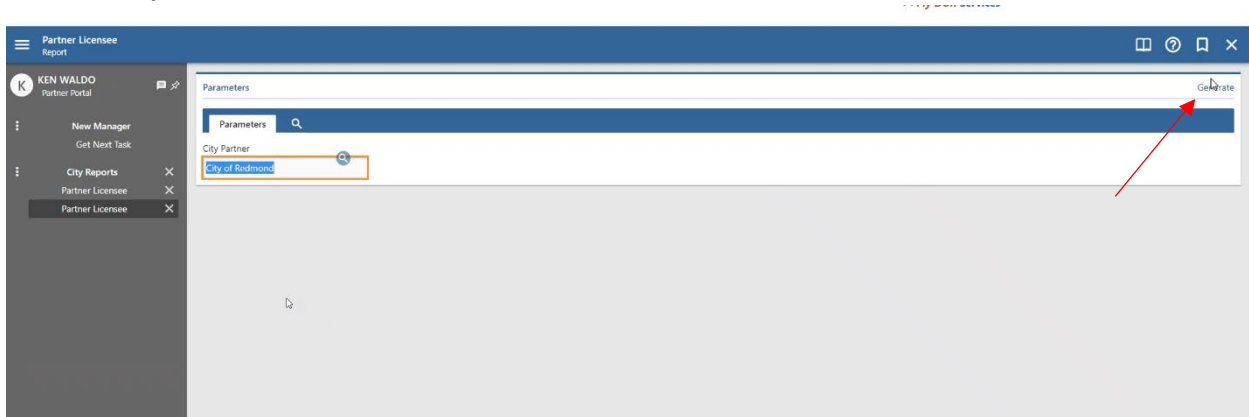
1. Log onto the BLS Portal and navigate to the Reporting tab and select City Reporting



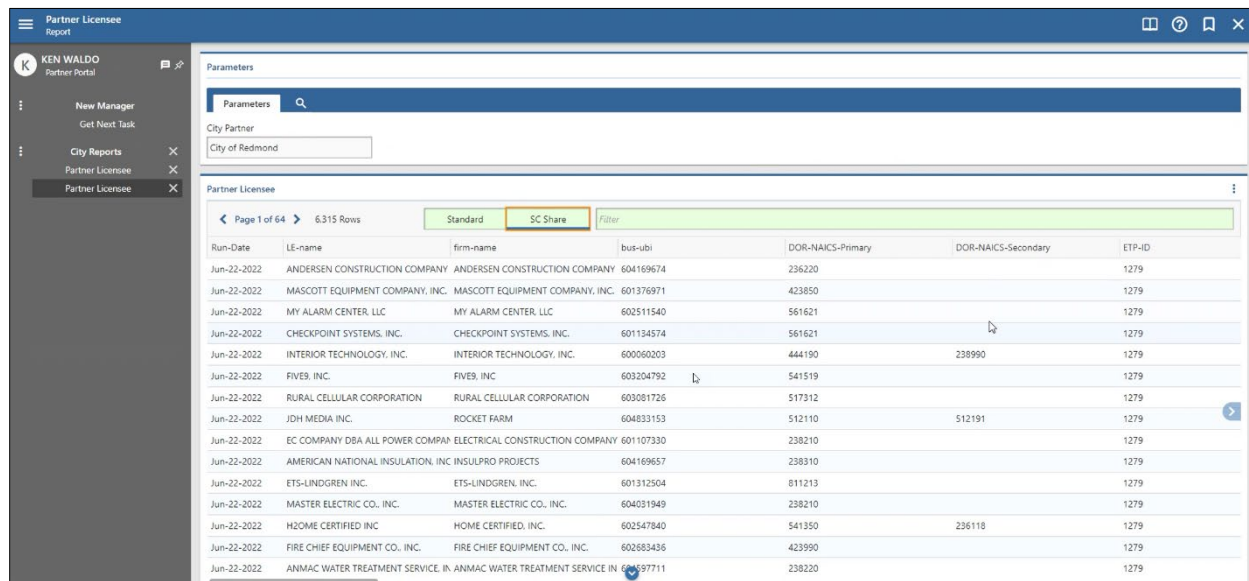
2. Select the Partner Licensee option



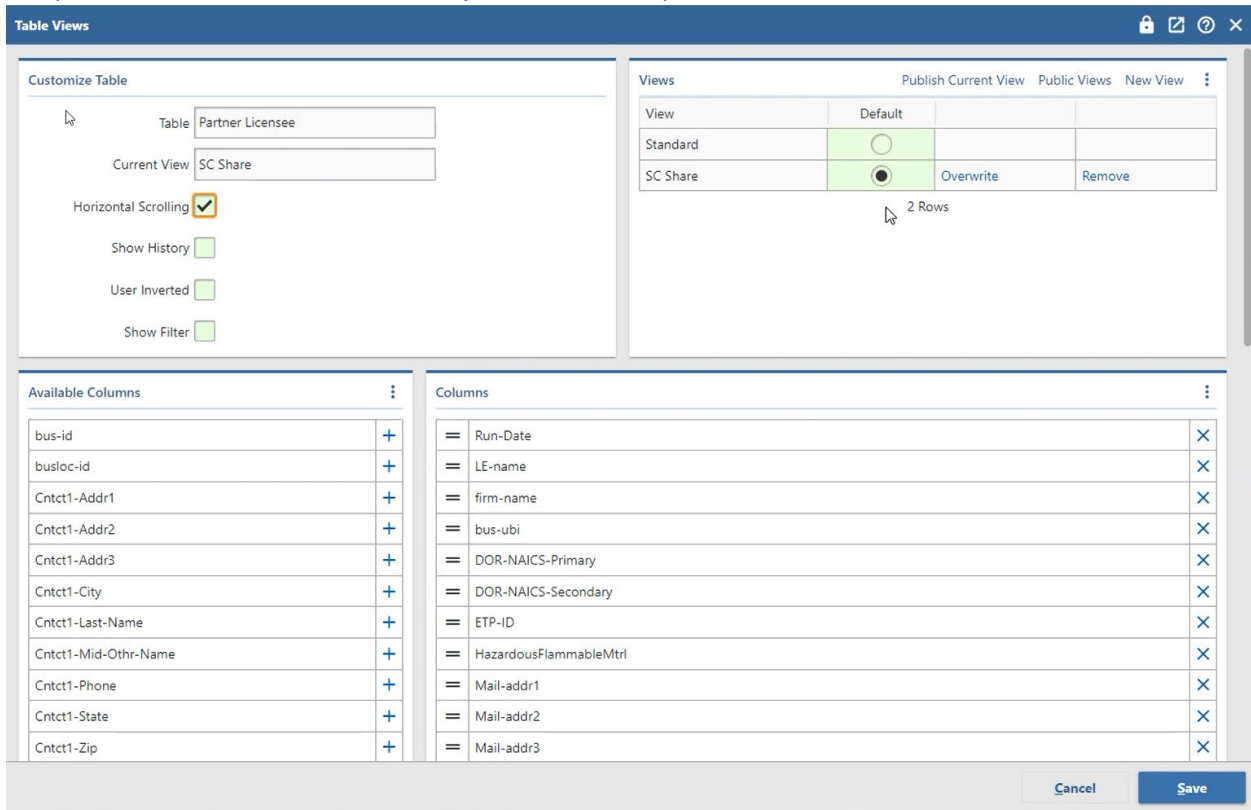
3. Enter the jurisdiction name and click Generate



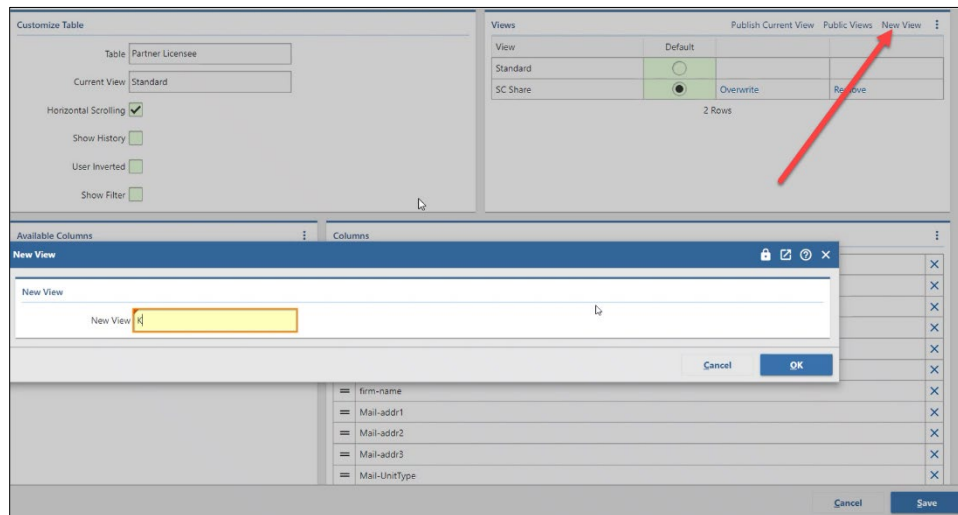
4. All current business licenses in the jurisdiction will appear in a preview window.



5. Open the Table Views window to adjust the data output



6. Open a New View to customize the table and name it "BLS license download_[6-digit date]"



7. Add (+ symbol) or remove (x symbol) columns to include only the following data fields. If necessary, rearrange the data fields to be in this order (using the parallel lines symbol on the left of the column name):

- *bus-ubi*, the business license number
- *LE-name*, legal business name
- *firm-name*, the “doing business as” name
- *Mail-addr1*, the mailing address of the business
- *Mail-addr2*, the mailing address of the business
- *Mail-addr3*, the mailing address of the business
- *Mail-UnitType*, the mailing address of the business
- *Mail-Unit*, the mailing address of the business
- *Mail-City*, the mailing address of the business
- *Mail-State*, the mailing address of the business
- *Mail-Zip*, the mailing address of the business
- *Mail-Zip+4*, the mailing address of the business
- *Phys-addr1*, the physical address of the business
- *Phys-addr2*, the physical address of the business
- *Phys-addr3*, the physical address of the business
- *Phys-UnitType*, the physical address of the business
- *Phys-Unit*, the physical address of the business
- *Phys-City*, the physical address of the business
- *Phys-State*, the physical address of the business
- *Phys-Zip*, the physical address of the business
- *Phys-Zip+4*, the physical address of the business
- *ETP-ID*, the endorsement identifiers for the business (includes “non-resident”)
- *HazardousFlammableMtrl*, indicates if the business uses hazardous or flammable materials
- *ProdDescription*, provides the business description as written by the applicant
- *DOR-NAICS-Primary*, indicates the primary NAICS code for the business
- *DOR-NAICS-Secondary*, indicates the secondary NAICS code for the business

8. Save the custom table that you created with the columns listed above

The screenshot shows the 'Table Views' application interface. The top bar contains the title 'Table Views' and standard window controls. The main area is divided into several sections:

- Customize Table:** Includes a 'Table' dropdown set to 'Partner Licensee', a 'Current View' dropdown set to 'SC Share', and several checkboxes: 'Horizontal Scrolling' (checked), 'Show History', 'User Inverted', and 'Show Filter'.
- Views:** A table with columns 'View', 'Default', and two unlabeled columns. It shows 'Standard' and 'SC Share' views. The 'SC Share' row is selected, and the 'Default' column has a radio button. Below the table, it says '2 Rows'.
- Available Columns:** A list of columns with a '+' icon next to each, including bus-id, busloc-id, Cntct1-Addr1, Cntct1-Addr2, Cntct1-Addr3, Cntct1-City, Cntct1-Last-Name, Cntct1-Mid-Othr-Name, Cntct1-Phone, Cntct1-State, and Cntct1-Zip.
- Columns:** A list of columns with a '-' icon next to each, including Run-Date, LE-name, firm-name, bus-ubi, DOR-NAICS-Primary, DOR-NAICS-Secondary, ETP-ID, HazardousFlammableMtrl, Mail-addr1, Mail-addr2, and Mail-addr3.

At the bottom right, there are 'Cancel' and 'Save' buttons. A red arrow points to the 'Save' button.