



PRELIMINARY PROPOSAL FOR LICENSING  
AND OPERATION  
SALES TAX DATA MANAGEMENT SYSTEM

---

**PUBLIC DOCUMENT**

*PREPARED FOR*  
STREAMLINED SALES TAX GOVERNING BOARD  
EXECUTIVE COMMITTEE

RESPECTFULLY SUBMITTED NOVEMBER 16, 2009 BY:

THE FEDERAL TAX AUTHORITY (FED-TAX.NET)

CORNELIUS WILLIS  
CHIEF SALES AND MARKETING OFFICER  
CORNELIUS@FED-TAX.NET

*This page intentionally left blank*

# TABLE OF CONTENTS

1	EXECUTIVE SUMMARY	1
2	CSDPUB FEATURES AND BENEFITS	3
2.1	Service Administration and Controls .....	4
2.2	Rate and Boundary Data Validation Service (RBV) .....	4
2.3	Taxability Matrix Editor and Database (TMED).....	5
2.4	Data Distribution .....	5
3	TECHNICAL DESCRIPTIONS	6
3.1	System Requirements.....	6
3.2	Data Formats .....	6
4	CHANGE MANAGEMENT AND DOCUMENTATION CONSIDERATIONS	7
5	SUPPORT AND MAINTENANCE	7
6	PROPOSED COMMERCIAL TERMS	8
7	TIMING CONSIDERATIONS	8
	APPENDIX A - JUNE 2009 CDIR	9

*This page intentionally left blank*

## 1 EXECUTIVE SUMMARY

---

The Streamlined Sales and Use Tax Agreement (SSUTA) requires Member States to periodically publish rate and boundary data and taxability matrices. This information, (herein referred to as Critical States Data or CSD), is used by CSPs, Volunteer and Nexus Sellers to accurately calculate local sales tax rates. This data is currently managed in a “loosely-coupled” manner, creating opportunities for error, variances and increased workload for sellers, CSP’s and tax administration staff.

The current framework for CSD publishing does not provide automated validation or provide taxability matrices in machine-readable form, and does not provide tools to allow states to manage the creation of CSD or facilities to allow SSTGB to manage publishing of CSD. Consequently calculating interstate sales tax requires multiple transcriptions from human-readable to machine-readable data representations, by sellers, states, SSTGB staff and CSPs. This automation gap creates additional cost and workload, and increases the potential for clerical errors and interpretation disputes. The multiple sources of taxability information, the need to transform this data into multiple formats and the dependency of multiple business entities and regulatory processes on this data all indicate a requirement for a referenceable **system of record** for sales tax data provided by states. Such a system of record would remove or reduce the potential for errors or disputes about semantic definitions, data formatting, deadlines and approval workflows, dates of effectiveness and interpretation generally.

As part of the TaxCloud sales tax automation service currently undergoing certification as a CAS, Fed-Tax.net has created a system that ingests, updates, validates the syntactic integrity of, and manages the publishing of CSD in machine-readable form. Fed-Tax incorporates reports from this system in the “Critical Data Irregularities Reports” provided to the Streamlined Sales Tax Governing Board (“SSTGB”) by Fed-Tax when errors or omissions are discovered in the rates and boundaries component of CSD. Previously provided examples of this report are included as Exhibits to this proposal. This same system also provides tools for editing, managing and publishing taxability matrices in machine-readable form as stipulated by section 328.A of the SSUTA.

Based upon ongoing discussions with the SSUTA Certification Committee, Fed-Tax has prepared this preliminary proposal to license and/or operate a stand-alone version of the sub-system describe above to serve as a system of record for sales tax data, herein referred to as CSDPub, for independent operation by the SSTGB. Such a system could be implemented as an extension to the SSTGB website, with similar appearance and branding. Fed-Tax anticipates that SSTGB will designate an appropriate procurement process should it elect to provide such a system.

The stand-alone CSDPub service described here will help SSUTA Member States to provide syntactically accurate and complete Critical States Data to the SSTGB, to ensure the accuracy of tax calculation and the overall integrity of the automated sales tax system envisioned by the SSUTA. By

providing editing and validation tools directly to the states for creating CSD, CSDPub will also reduce workloads for state tax administrators and SSTGB technical staff.

CSDPub could be used to both capture and validate CSD in advance of publication via the SSTGB web site, and could also be extended with basic authentication and tracking techniques to allow the SSTGB to monitor maintenance of CSD by states, and utilization of CSD by vendors, CSPs and others to assist in compliance tracking.

Fed-Tax.net proposes licensing CSDPub, and/or operating and supporting it on behalf of the SSTGB to help ensure conformance and completeness of Critical States Data, increase the accuracy of sales tax collection and allocation by all SSUTA participants, and increase overall compliance with the SSUTA.

## 2 CSDPUB FEATURES AND BENEFITS

CSDPub is implemented as four subsystems:

- 1) **Service Administration and Controls**, a subsystem which manages entities and individuals that have read and write access to the system, e.g. the system allows security groups per state or other entity, such that administration-privileged users can manage system access by others in their security group (e.g. a work group of tax administration officials working for a particular state)
- 2) **Rates and Boundaries Data Validation (RBV)** checks the integrity and accuracy of uploaded Rate and Boundary Files,
- 3) **Taxability Matrix Editor and Database (TMED)** provides forms-driven data entry for use by states to populate taxability matrix data, with output provided in machine readable or human readable (PDF) form.
- 4) **Data Publishing and Syndication System** automatically provides rate and boundary data and taxability matrices to CSPs, states and the general public in multiple machine-readable and human-readable formats, and includes tracking mechanism to monitor utilization of CSDPub supplied data.

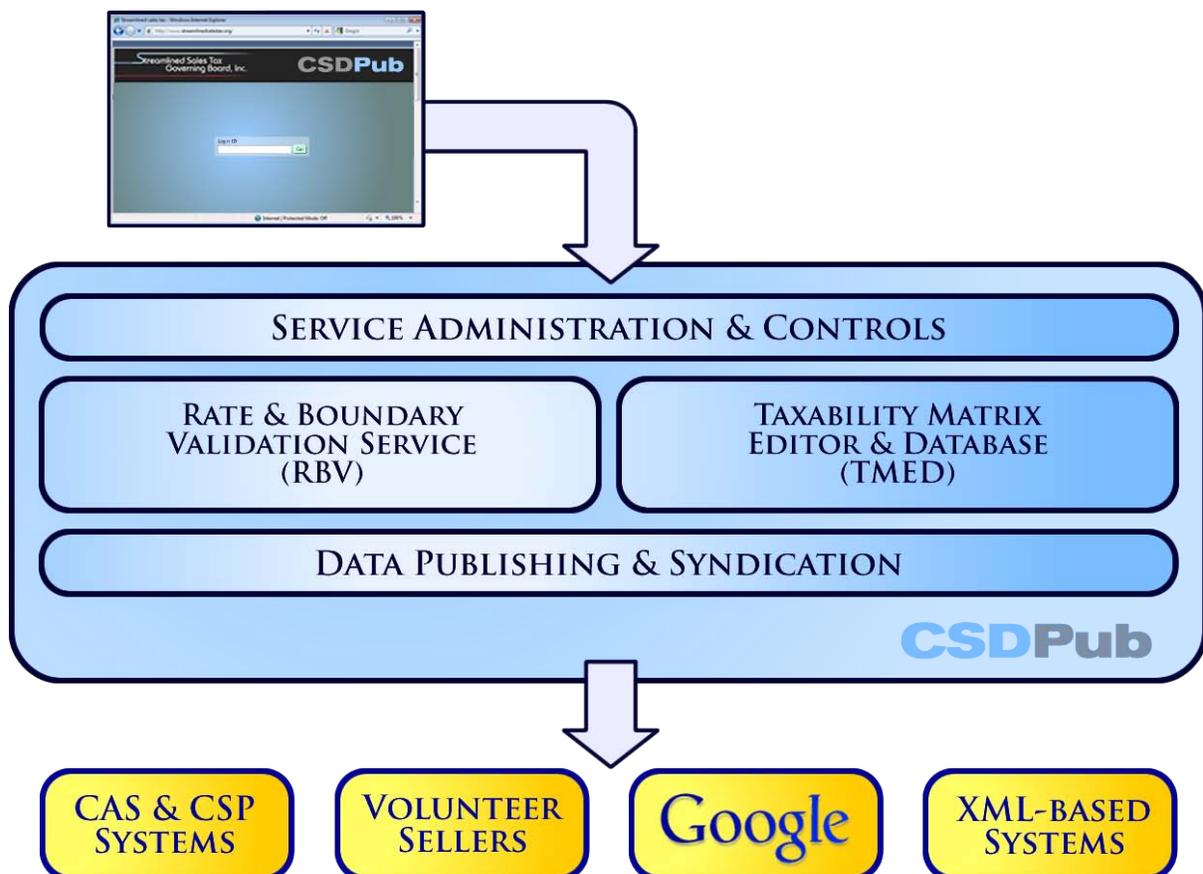


Figure 1 - CSDPub Logical Architecture

## 2.1 SERVICE ADMINISTRATION AND CONTROLS

The Service Administration and Controls subsystem manages authenticated access and security groups, which allow individual states to control access to the system. For example, an administrative user in the security group for a particular state can create and manage user accounts for different state tax administration staff, but these users can only update data for their state. States and administrative staff log-in to the RBV to upload and validate files. Super-user privileges are provided to SSTGB staff to allow management of all states' tax information.

## 2.2 RATE AND BOUNDARY DATA VALIDATION SERVICE (RBV)

The RBV subsystem automatically scans new submissions (uploads) for irregularities before accepting submissions, forcing submitting states to take responsibility for data correctness. Specific checks the RBV performs include:

Type of Check	Explanation	Alert Threshold
Syntactic validation of rate files	Ensure correct field layout, field size and expected data types of documented fields	Any error discovered raises an alert and must be corrected
Syntactic validation of boundary files	Ensure correct field layout, field size and expected data types of documented fields	Any error discovered causes an alert and must be corrected
Completeness control	Checks to determine if there are zip codes for which no sales tax rates exist	If more than 5% (adjustable) of a state's zip published codes are missing, when compared to the immediately preceding period, than an alert is raised, and must be addressed.
FIPS code matching	For Rates Data files, compare Jurisdictional FIPS with Jurisdiction Type (and correct if necessary) to ensure leading zeros have not been inadvertently "trimmed" by well-intentioned software (example: FIPS 047 is distinctly different from FIPS 47)	Any error discovered can be corrected automatically by the system, but the user/submitter will be alerted to the matter. The user can suspend processing and resubmit, or simply acknowledge the correction made by the system.
Finding potentially inaccurate rates	Rates outside of normal ranges usually indicate data entry or formatting errors	Issues warnings on rates that are outside of acceptable range (i.e. 5% to 17%)

Figure 2 - RBV Standard Controls

**Automatic reporting and notification:** The RBV produces reports for each submission, indicating irregularities as they are discovered. The RBV can also be configured to automatically notify SSTGB staff of updated, erroneous or out of date data.

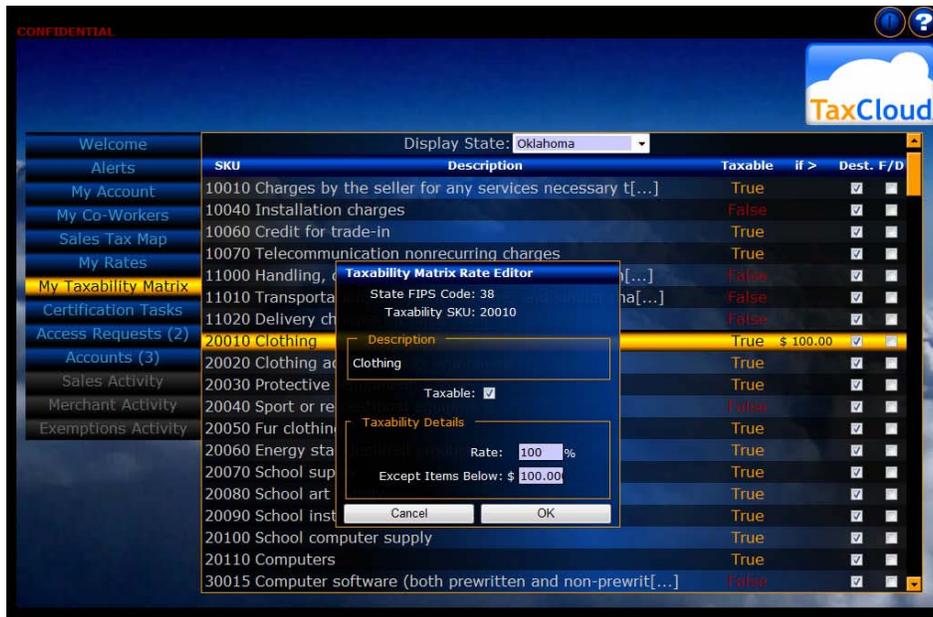
**Issue Levels:** The RBV supports configurable issue levels, including Informational alerts, Warnings and Errors. Information includes number of boundaries and jurisdictions. Warnings include missing zip codes and inaccurate rates, and Errors include missing columns.



Figure 3 - Printable PDF Irregularity Reports

## 2.3 TAXABILITY MATRIX EDITOR AND DATABASE (TMED)

The TMED subsystem depends on the Service Administration and Controls system to limit access to authorized state tax personnel. Once authenticated, state tax administration officials can edit and update taxability data for their state, or post it for review by colleagues or SSTGB. This system



produces output in both machine-readable XML through the data distribution system described below, enabling easy incorporation into third party systems. The system also produces human-readable PDF files that closely resemble the forms currently used to capture and communicate taxability matrices today.

Figure 4 - Screen capture of TaxCloud 0.9a Taxability Matrix Editor

Taxability matrices are complex data structures that include conditionality and other special relationships. Further, both the structure and content of taxability matrices change as SSUTA policy and state tax rates change. As part of an operational contract, Fed-Tax can continue to maintain and update and test this system to ensure up-to-date compliance with the SSUTA. TMED can be configured to notify SSTGB staff of accesses or updates to the system by state personnel.

## 2.4 DATA PUBLISHING AND SYNDICATION

CSDPub can optionally be extended to be used as a secure public access point for validated Critical States Data. If this option is preferred, the SSTGB could require simple registration (with a valid email address) to view/download files. This would give the SSTGB further visibility into who is accessing and/or relying upon this data. The site could also be used to reformat validated Critical States Data into one or more alternative XML formats, to encourage automated consumption by third-party systems (e.g. enabling population of a Google search result with taxability rates). The monitoring and reporting functionality allows SSTGB to monitor maintenance and utilization of CSD by states, vendors, CSPs and others to assist in compliance tracking.

## 3 TECHNICAL DESCRIPTIONS

---

### 3.1 SYSTEM REQUIREMENTS

---

CSDPub is implemented as a web application with the following characteristics:

- Authentication relies upon a unique email address for Login ID and a strong password.
- The RBV web application logic is implemented in ASP.NET 2.0x, running from Microsoft Internet Information Server v.6 on Windows Server 2003 Datacenter Edition.
- The underlying RBV database engine is Microsoft SQL 2005 Express Edition.
- The RBV is operated entirely from a cloud computing environment, utilizing one or more purpose-built EC2 Amazon Machine Images (AMIs).

### 3.2 DATA FORMATS

---

Data Publishing formats will likely evolve to be numerous, however to get started we plan on developing a Really Simple Syndication (RSS) 2.0 syntax which may look approximately as shown in Figure 5.

```
<?xml version="1.0"?>
<rss version="2.0" xmlns:t="http://taxcloud.net/CSDfeed/1.0">
  <channel>
    <title>Local Sales Tax Rate</title>
    <link>http://taxcloud.net</link>
    <description>Determine Local Sales Tax Rates for
Free!</description>
    <item>
      <title>98199-1451</title>
      <link>http://taxcloud.net/myrate/?z=98199-1451</link>
      <description>Local sales tax rate information for 98199-1451
- learn more at http://taxcloud.net</description>

      <t:image_link>https://taxcloud.net/taxcloud_logo.gif</t:image_link>
      <t:basic_aggr_rate>9.5</t:basic_aggr_rate>
      <t:effective_date>2010-01-01</t:effective_date>
      <t:expiration_date>9999-01-01</t:expiration_date>
      <t:is_destination_based>true</t:is_destination_based>
      <t:map_url>http://maps.google.com/maps?f=d&q=98199-
1451&source=s_d&saddr=&daddr=&geocode=&hl=en&mra=mr&sll=47.65892,-
122.401997&sspn=0.003563,0.006588&ie=UTF8&ll=47.660806,-
122.397523&spn=0.003562,0.006588&z=17</t:map_url>
    </item>
  </channel>
</rss>
```

Figure 5 - Example RSS 2.0 Sales Tax Rate Feed

Making this data available in this manner will allow third-party automated systems to begin "consuming" or otherwise ingesting this information for inclusion within their own systems. For example, registering this sales tax rates feed with Google Base (<http://base.google.com>) could result in auto-answer search results.

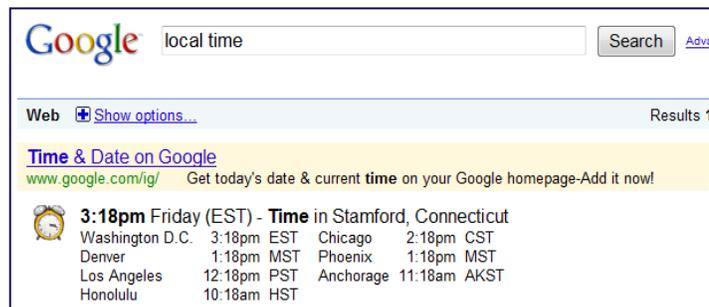


Figure 6 - Example Google Auto-Answer for "Local Time"

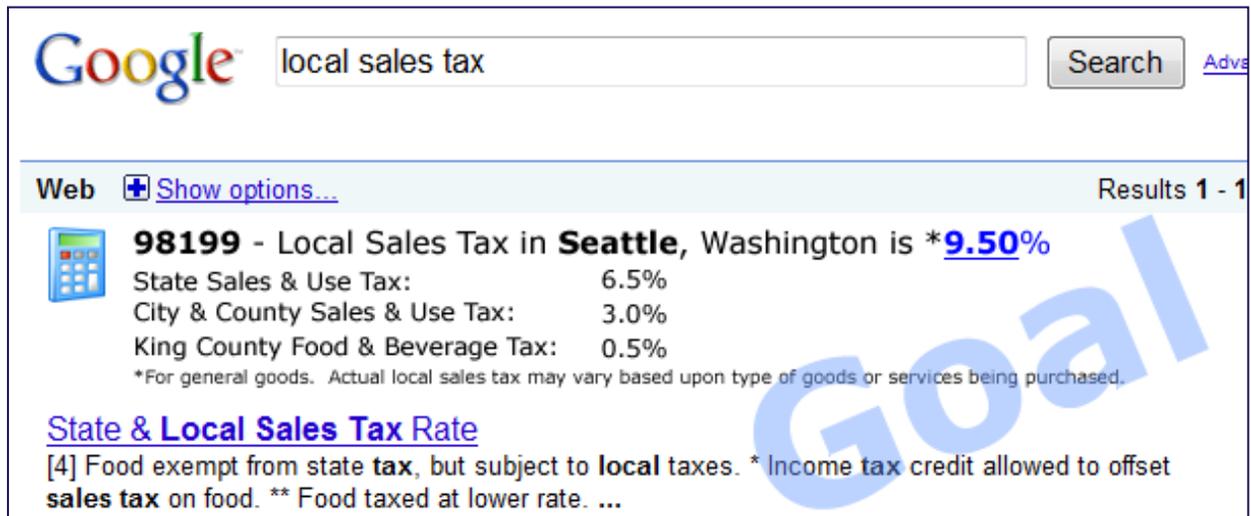


Figure 7 - A syndicated source *could* enable Auto-Answer for "local sales tax"

## 4 CHANGE MANAGEMENT AND DOCUMENTATION CONSIDERATIONS

To make use of the RBV subsystem of CSDPub, state tax administration personnel must be familiar with the concept of uploading a file to a website, using the same file selection metaphor commonly used to upload images or documents to web services. Consequently training and documentation for this system will consist of explanations of error messages and remediation processes.

TMED uses drop down menus and textboxes, and closely approximates the existing non-automated PDF forms currently used to capture and communicate this data. Taxability matrix data will be pre-populated in these forms for each state at system launch, so that initially the states simply log in to the system and review, edit and approve the data provided.

## 5 SUPPORT AND MAINTENANCE

Fed-Tax.net can provide multi-tier technical support, including operation of a ticketing and tracking system, as part of an operational contract. Support could include level 1, 2 and 3 technical support as described in the table below, and can be provided at various service levels (response times and availability) per SSTGB requirements.

Tier 1 – Front Line End User Support	Assisting with user account creation, user on-boarding and usage trouble shooting. Gather diagnostic information on usage issues and troubleshoot by proposing known solutions to common user issues.
Tier 2 – Technical Analysis	Escalation from Tier 1, including further determination of root cause. This may include software repair or reinstallation, use of diagnostic tools and access to user accounts.
Tier 3 – Quick Fix Engineering	Repair and resolution of discovered defects, regression testing fixes in the TaxCloud test environment and deploying to production systems

Figure 8 - Support Options

## 6 PROPOSED COMMERCIAL TERMS

---

Licensing, operations, support and maintenance provisions to be discussed.

## 7 TIMING CONSIDERATIONS

---

Fed-Tax.net would encourage the SSTGB to consider immediate adoption of the rates and boundaries validation system such that the next round of quarterly updates to Critical States Data (anticipated in December) can be validated in advance of widespread distribution via the SSTGB website.

The Taxability Matrix Editor and other elements of the system can be brought online with up-to-date taxability matrix information as early as February 15<sup>th</sup> 2010.



## JUNE 2009 CRITICAL STATES DATA IRREGULARITIES

INTERNAL USE ONLY

REVISION HISTORY:

DATE	AUTHORING PARTIES	DESCRIPTION
7/15/2009	CHRIS GREATHOUSE CGREATHOUSE@FED-TAX.NET	NOTED ALGORITHM FLAW IN BOUNDARY COMPLETENESS CONTROL
8/20/2009	R. DAVID L. CAMPBELL DCAMPBELL@FED-TAX.NET	REVISED VALIDATION SYSTEMS OUTPUT WITH PARTIAL MANUAL INSPECTION

**CONFIDENTIALITY NOTICE:**

THIS DOCUMENT IS CONSIDERED PRIVILEGED "CONFIDENTIAL INFORMATION" AND IS INTENDED ONLY FOR REVIEW BY AUTHORIZED INDIVIDUALS AND ORGANIZATIONS BOUND BY TERMS OF CONFIDENTIALITY BETWEEN RECIPIENT AND FED-TAX.NET. ANY DISSEMINATION, CITATION, PUBLICATION, DISTRIBUTION, OR COPYING OF THIS DOCUMENT, OR ANY PART THEREOF, (IN ELECTRONIC FORM OR OTHERWISE) IS STRICTLY PROHIBITED.

# TABLE OF CONTENTS

1	ABOUT THIS REPORT	1
1.1	Introduction.....	1
1.2	Database Setup.....	1
1.3	Import Method.....	1
1.4	Irregularities.....	1
1.5	Irregularities Reporting.....	1
1.6	Questions?.....	1
2	STATE PROVIDED RATE FILE(S)	2
2.1	North Carolina Rate Data Irregularities.....	2
2.2	North Carolina Rate Data Irregularities Resolution.....	2
3	STATE PROVIDED BOUNDARY FILE(S)	3
3.1	Indiana Boundary Data Irregularities.....	3
3.2	Indiana Boundary Data Irregularities Resolution.....	3
3.3	West Virginia Boundary Data Irregularities.....	4
3.4	West Virginia Boundary Data Irregularities Resolution.....	4
4	QUESTIONS ABOUT MISSING ZIP CODES	5
4.1	State of Washington Missing Zip Codes.....	6
4.1.1	<i>State of Washington DOR - SSUTA provided data.....</i>	6
4.1.2	<i>State of Washington DOR Tax Rate Lookup Tool.....</i>	6
4.1.3	<i>State of Washington DOR Published Data:.....</i>	7
4.1.4	<i>Inquiry and Response from State of Washington DOR.....</i>	7
4.2	Other States with Apparently "Missing" Zip Codes.....	8
4.2.1	<i>Nebraska.....</i>	8
4.2.2	<i>Oklahoma.....</i>	8
4.2.3	<i>Vermont.....</i>	8
4.2.4	<i>Utah.....</i>	8
4.2.5	<i>Minnesota.....</i>	8
4.2.6	<i>Tennessee.....</i>	8
4.2.7	<i>North Dakota.....</i>	8
4.2.8	<i>Ohio.....</i>	8
4.2.9	<i>Arkansas.....</i>	9
4.2.10	<i>Kansas.....</i>	9
4.2.11	<i>Wyoming.....</i>	9

# 1 ABOUT THIS REPORT

---

## 1.1 INTRODUCTION

Fed-Tax.net's TaxCloud service automatically retrieves and syntactically validates tax rates and boundaries data ("Critical States Data") made available by the states via the SSUTA Governing Board's web site (at URL [https://streamlinedsalestax.org/Extract\\_Files/](https://streamlinedsalestax.org/Extract_Files/)). Subsequent updates to published Critical States Data are also automatically retrieved, syntactically validated, and differentially compared.

## 1.2 DATABASE SETUP

The boundary and rate database was created based upon SSUTA provided guidance from "Streamlined Sales Tax Project : Rates and Boundary Databases Instructional Paper (August, 2005)" (at URL: <http://www.streamlinedsalestax.org/Technology/RatesandBoundariesClean082605.pdf>).

## 1.3 IMPORT METHOD

Fed-Tax.net currently uses SQL Server 2005 to store all databases. Our goal is to rely primarily upon built-in tools and commands when importing Critical States Data. Due to the large amount of data to be imported, the primary command used is the BULK INSERT command.

## 1.4 IRREGULARITIES

Any data format or syntax errors detected in Critical States Data cause a "Severity 2" alert, which is immediately noted in an advisory log, with an automated email message sent to the Technical Operations team and the CTO.

## 1.5 IRREGULARITIES REPORTING

In addition to the Severity 2 system alert, Critical States Data Irregularity incidents also trigger immediate investigation by the Release Manager and/or Engineering Team. Findings of such investigation(s) are then reported to the CTO for review. This document is just such a Critical States Data Irregularity Report for State Rate and Boundary Files released in **June 2009 - Last Retrieved/Downloaded August 6th, 2009.**

## 1.6 QUESTIONS?

For details regarding the findings outlined in this report, please contact the preparer indicated on the first page, or contact Fed-Tax.net's CISO/CTO:

Paul E. Onnen  
Email: [paul@fed-tax.net](mailto:paul@fed-tax.net)  
Telephone: (206) 390-1535

## 2 STATE PROVIDED RATE FILE(S)

---

On or about June 1, 2009, the following SSUTA Member States released updated Sales Tax Rate Files. **Grey** Member States did not release updated Rate files. **Bold Red** Member States released updated Rate files which contain Irregular or Invalid Data.

Arkansas	Indiana	Iowa	Kansas	Kentucky
Michigan	Minnesota	Nebraska	Nevada	New Jersey
<b>North Carolina</b>	North Dakota	Oklahoma	Ohio	Rhode Island
South Dakota	Tennessee	Utah	Vermont	Washington
West Virginia	Wisconsin	Wyoming		

### 2.1 NORTH CAROLINA RATE DATA IRREGULARITIES

The rate file for North Carolina from the June 1, 2009 release is labeled NCR102008.csv. The problem encountered with this file is that there were some rows that had too many commas. This extra comma was causing SQL to assume that an additional column of data was available.

**UPDATE: On August 10, 2009, North Carolina released a revised Sales Tax Rate File (via the SSUTA rates URL) labeled NCR092009.csv, which is no longer exhibiting any irregularities.**

Sample from NCR102008.csv:

```
line 192 : 37,00,185,0.02500,0.02500,,,20050801,20080930
line 193 : 37,00,185,0.02250,0.02250,,,20081001,99991231,
line 194 : 37,00,187,0.02500,0.02500,,,20050801,20080930,
```

In the sample lines above the trailing commas at the end of the row are causing the problems. When trying to import as is, SQL Server will generate the following error:

SQL Server 2005 error:

NCR102008.csv - Bulk load data conversion error (type mismatch or invalid character for the specified codepage) for row 193, column 9 (END).

### 2.2 NORTH CAROLINA RATE DATA IRREGULARITIES RESOLUTION

To remedy this column count irregularity the following changes were made by hand:

Lines 192 through 208 were altered by removing the trailing comma on each line before issuing the BULK INSERT command - which then succeeded without error(s).

**Resolution proposed:** 6/18/2009 by cgreathouse@fed-tax.net

**Proposed resolution accepted:** 6/19/2009 by paul@fed-tax.net

**Proposed resolution implemented:** 6/21/2009 by cgreathouse@fed-tax.net

**Resolution verified:** 6/21/2009 by cgreathouse@fed-tax.net

**Irregularity Corrected/Removed by North Carolina August 10, 2009**

### 3 STATE PROVIDED BOUNDARY FILE(S)

---

On or about June 1, 2009, the following SSUTA Member States released updated Boundary Files. **Grey** Member States did not release updated Boundary files. **Bold Red** Member States released updated Boundary files which contain Irregular or Invalid Data.

Arkansas	<b>Indiana</b>	Iowa	Kansas	Kentucky
Michigan	Minnesota	Nebraska	Nevada	New Jersey
North Carolina	North Dakota	Oklahoma	Ohio	Rhode Island
South Dakota	Tennessee	Utah	Vermont	Washington
<b>West Virginia</b>	Wisconsin	Wyoming		

#### 3.1 INDIANA BOUNDARY DATA IRREGULARITIES

The boundary file for Indiana from the June 1, 2009 release is labeled INB012005.csv. The problem encountered with this file is that there were not enough commas (columns) in the CSV file. These missing commas was causing problems during the import.

Sample from INB012005.csv:

```
4,20050101,29991231,,,,,,,,,,,,,46001,0,47997,9999,,18,18  
Z,20050101,29991231,,,,,,,,,,,,,46001,,47997,,,18,18
```

Referring to the sample lines above, during import the SQL Server gave the following error:

```
INB012005.csv - Bulk load data conversion error (truncation) for row 1,  
column 24 (FIPSSI).
```

This error is based upon the fact that the TaxCloud SQL Server is attempting to import twenty-four columns of data into the TaxCloud boundary database which is setup to expect 89 columns of data (based upon SSUTA documentation referenced previously in Section 1.2).

#### 3.2 INDIANA BOUNDARY DATA IRREGULARITIES RESOLUTION

To remedy this column count irregularity the following changes were made by hand:

To fix this problem, 65 more commas were added to the end of each row before issuing the BULK INSERT command - which then succeeded without error(s).

**Resolution proposed:** 6/18/2009 by cgreathouse@fed-tax.net

**Proposed resolution accepted:** 6/19/2009 by paul@fed-tax.net

**Proposed resolution implemented:** 6/21/2009 by cgreathouse@fed-tax.net

**Resolution verified:** 6/21/2009 by cgreathouse@fed-tax.net

### 3.3 WEST VIRGINIA BOUNDARY DATA IRREGULARITIES

The boundary file for West Virginia from the June 1, 2009 release is labeled WVB072005.csv. The problem encountered with this file is that there were not enough commas (columns) in the CSV file. These missing commas was causing problems during the import.

**UPDATE:** On July 16, 2009, West Virginia released a revised Boundaries Data File (via the SSUTA boundaries URL) labeled WVB072009, which is no longer exhibiting any irregularities.

Sample from WVB072005.csv:

```
Z,20050701,99991231,,,,,,,,,,,,,ONEGO,,,26886,,26886,,,54,54  
4,20050701,99991231,,,,,,,,,,,,,24700,1,26899,9999,,54,54
```

Referring to the sample lines above , during import the SQL Server gave the following error:

```
WVB072005.csv - Bulk load data conversion error (truncation) for row 1,  
column 24 (FIPSSI).
```

This error is based upon the fact that the TaxCloud SQL Server is attempting to import twenty-four columns of data into the TaxCloud boundary database which is setup to expect 89 columns of data (based upon SSUTA documentation referenced previously in Section 1.2).

### 3.4 WEST VIRGINIA BOUNDARY DATA IRREGULARITIES RESOLUTION

To remedy this column count irregularity the following changes were made by hand:

To fix this problem, 65 more commas were added to the end of each row before issuing the BULK INSERT command - which then succeeded without error(s).

**Resolution proposed:** 6/18/2009 by cgreathouse@fed-tax.net

**Proposed resolution accepted:** 6/19/2009 by paul@fed-tax.net

**Proposed resolution implemented:** 6/21/2009 by cgreathouse@fed-tax.net

**Resolution verified:** 6/21/2009 by cgreathouse@fed-tax.net

## 4 QUESTIONS ABOUT MISSING ZIP CODES

---

While working on import of the most recent set of Critical States Data, Fed-Tax.net encountered an issue of sincere concern regarding apparently "missing" zip codes.

Our boundary completeness control is based intrinsically upon the current release of Critical States Data (June 2009). This completeness test first builds a list of expired zip codes, and then builds a list of current zip codes. Simple comparison of these lists yields a list of potentially new or missing current zip codes. On a per-state basis, deviation or variation between expired and current zip codes in excess of 5% causes a Severity 2 system alert, requiring immediate investigation. We generally expect that it may be natural for some level of churn due to USPS management of zip codes, which likely involves periodic reassignment, expiration, and creation of zip codes - we have arbitrarily chosen 5% as our alert threshold.

The initial implementation of our completeness control showed an extraordinarily large volume (over 180,000) of possibly missing zip codes. Upon further review, we uncovered an oversight in the algorithm we developed for building these lists of expired and current zip codes - however we alerted SSUTA regarding our findings and provided a copy of this document on July 16, 2009. During our discussion with Scott Peterson and David Thompson on July 16, 2009 we were advised that for all instances of data irregularity and incompleteness, our regular course of action should include immediate notification to David Thompson, SSUTA Chief Technology Officer.

We are pleased to report that we have corrected our algorithm, however we are still detecting an extremely large volume of apparently missing zip codes (**112,916**). Of the 23 SSUTA states, 12 appear to have missing zip codes, though only three are exceeding our 5% alert threshold: Washington, Nebraska, and Oklahoma.

This section 4 describes the current state of our technical investigation(s) related to boundary completeness alerts generated the latest version of our boundary data completeness control.

## 4.1 STATE OF WASHINGTON MISSING ZIP CODES

The TaxCloud completeness control has indicated that exactly **38,587** zip codes (693 five-digit and 37,894 plus4) were no longer defined in the State of Washington's SSUTA published boundary data (of 190,513 total reported zip codes - approx. **20% missing**).

### 4.1.1 STATE OF WASHINGTON DOR - SSUTA PROVIDED DATA

Specifically, referring to zip code data from the Washington State boundary data file provided by SSUTA Governing Board at:

[https://streamlinedsalestax.org/Extract\\_Files/BOUNDARIES/WAB072009.zip](https://streamlinedsalestax.org/Extract_Files/BOUNDARIES/WAB072009.zip)

While inspecting this data for general completeness, we were unable to locate a current entry for an empirically sampled plus4 zip code **98282-7279**. Here are several relevant lines from that file:

WAB072009.csv - Lines 506 - 509:

```
Z,20090101,20090331,,,,,,,,,,,,,98281,,98281,,03737,53,53,073,,,,,ST,L3737,45,,,,,
Z,20080101,20090331,,,,,,,,,,,,,98282,,98282,,09134,53,53,,,,,ST,L4200,45,,,,,
Z,20090401,99991231,,,,,,,,,,,,,98282,,98282,,09134,53,53,,,,,ST,L4200,45,,,,,
Z,20090401,99991231,,,,,,,,,,,,,98283,,98283,,09135,53,53,073,,,,,ST,L4200,45,,,,,
```

What can be seen from lines 506-509 (the straight 5 digit Zip Code section of the Boundaries file), is that the five digit zip code 98282 has a new boundary definition effective 4/1/2009. However, looking further down the file in the plus4 section of the Boundaries file, we cannot find a valid entry for the plus4 zip code for 98282-7279:

WAB072009.csv - Lines 61,034 - 61,036:

```
4,20090401,99991231,,,,,,,,,,,,,98282,7270,98282,7270,01500,53,53,029,,,,,ST,L4200,45,,,,,
4,20080101,20090331,,,,,,,,,,,,,98282,7271,98282,7296,01500,53,53,029,,,,,ST,L4200,45,,,,,
4,20090701,99991231,,,,,,,,,,,,,98282,7296,98282,7296,01500,53,53,029,,,,,ST,L4200,45,,,,,
```

As can be seen, the only line applicable to 98282-7279 appears to have expired on 3/31/2009.

### 4.1.2 STATE OF WASHINGTON DOR TAX RATE LOOKUP TOOL

In order to cross-check our interpretation, we then checked the State of Washington Department of Revenue website to see if the plus4 zip code existed. Using the Washington State DOR tool provided at <http://dor.wa.gov/content/findtaxesandrates/salesandusetaxrates/lookupataxrate/>, we

were able to enter in the plus4 zip code 98282-7279 and get a rate of 8.4%. However, when entering simply the five digit zip code (a fall-back position suggested by missing plus4 data in the SSUTA provided Washington State Boundaries Data), we are returned a rate of 7.7%. This seems to indicate a distinct need for the complete plus4 data to determine the correct current tax rate.

#### 4.1.3 STATE OF WASHINGTON DOR PUBLISHED DATA:

To further validate (or alleviate) our concerns, we then checked the Washington State DOR databases downloadable from:

<http://dor.wa.gov/content/FindTaxesAndRates/SalesAndUseTaxRates/stdownloads.aspx>

Specifically, in the Washington State DOR provided file "State\_09Q2.txt" we were able to find the plus4 zip code 98282-7279.

State\_09Q2.txt - Line 583,160:

370,370,E,NE CAMANO DR,WA,98282,7279,Q22009,1500,N,Island PTBA,

#### 4.1.4 INQUIRY AND RESPONSE FROM STATE OF WASHINGTON DOR

In an effort to alleviate (or validate) our concerns, we initiated contact with the State of Washington Department of Revenue (via email on July 2nd) to disclose our findings (as described in this section) and solicit assistance to better understand what we were seeing. After several exchanges our concerns were validated by **the State of Washington, which has acknowledged and agreed with our findings.**

Specifically, on July 13th, the State of Washington (via David Wright - email:davidwr@dor.wa.gov) replied via email:

*"Yes we did find a error in the file. For some reason this record didn't get the plus 4 appended to it properly as it passed through the system. We are taking a look at this issue and will use this as a test record to see if we can find the core issue."*

We are currently awaiting further communication from the State of Washington Department of Revenue regarding next steps, or (hopefully) notification of publication of updated/corrected SSUTA file(s).

On August 20th, we sent another email to the State of Washington to inquire on the status of the corrected/updated data, and received the following response:

*"I did forward this issue to our web-services group since they build that file out. I will contact him again and see where he is sitting. Right now we are preparing the next quarter's files so he may be focusing on resolving the issue for that release."*

## 4.2 OTHER STATES WITH APPARENTLY "MISSING" ZIP CODES

### 4.2.1 NEBRASKA

The TaxCloud completeness control has indicated that exactly **26,962** zip codes (127 five -digit and 26,835 plus<sup>4</sup>) were no longer defined in the State of Nebraska's SSUTA published boundary data (of 98,643 total reported zip codes - approx. **27% missing**).

### 4.2.2 OKLAHOMA

The TaxCloud completeness control has indicated that exactly **10,484** zip codes (2 five -digit and 10,482 plus<sup>4</sup>) were no longer defined in the State of Oklahoma's SSUTA published boundary data (of 155,989 total reported zip codes - approx. **6.7% missing**).

### 4.2.3 VERMONT

The TaxCloud completeness control has indicated that exactly **52** zip codes (18 five -digit and 34 plus<sup>4</sup>) were no longer defined in the State of Vermont's SSUTA published boundary data (of 1,169 total reported zip codes - approx. 4.5% missing).

### 4.2.4 UTAH

The TaxCloud completeness control has indicated that exactly **36,247** zip codes (65 five-digit and 36,182 plus<sup>4</sup>) were no longer defined in the State of Utah's SSUTA published boundary data (of 1,187,324 total reported zip codes - approx. 3% missing).

### 4.2.5 MINNESOTA

The TaxCloud completeness control has indicated that exactly **443** zip codes (112 five-digit and 331 plus<sup>4</sup>) were no longer defined in the State of Minnesota's SSUTA published boundary data (of 21,487 total reported zip codes - approx. 2% missing).

### 4.2.6 TENNESSEE

The TaxCloud completeness control has indicated that exactly **46** zip codes (23 five -digit and 23 plus<sup>4</sup>) were no longer defined in the State of Tennessee's SSUTA published boundary data (of 80,154 total reported zip codes - less than 1% missing).

### 4.2.7 NORTH DAKOTA

The TaxCloud completeness control has indicated that exactly **3** zip codes (all five -digit) were no longer defined in the State of North Dakota's SSUTA published boundary data (of 15,505 total reported zip codes - less than 1% missing).

### 4.2.8 OHIO

The TaxCloud completeness control has indicated that exactly **62** zip codes (all five -digit) were no longer defined in the State of Ohio's SSUTA published boundary data (of 143,492 total reported zip codes - less than 1% missing).

#### *4.2.9 ARKANSAS*

The TaxCloud completeness control has indicated that exactly **26** zip codes (all five -digit) were no longer defined in the State of Arkansas' SSUTA published boundary data (of 533,627 total reported zip codes - less than 1% missing).

#### *4.2.10 KANSAS*

The TaxCloud completeness control has indicated that exactly **2** zip codes (both plus4) were no longer defined in the State of Kansas' SSUTA published boundary data (of 64,526 total reported zip codes - less than 1% missing).

#### *4.2.11 WYOMING*

The TaxCloud completeness control has indicated that exactly **2** zip codes (both plus4) were no longer defined in the State of Wyoming's SSUTA published boundary data (of 626 total reported zip codes - less than 1% missing).

