



## DatabaseUSA Relies on Alteryx to Ensure Clients Get the Most Accurate Lists and Data Possible

### CHALLENGE

DatabaseUSA wanted a flexible ETL solution that would increase the quality and accuracy of its data while reducing the time and money spent on data processing tasks and related training.

### SOLUTION

Using Alteryx in a unique self-hosted environment, DatabaseUSA created an elegant workflow that easily blends hundreds of data sources and ensures industry-leading accuracy with little to no coding required.

### RESULTS

- **Deeper Insights:** Blended hundreds of data sources, including call center data, online data, and utility data, into a clean, accurate database upon which customers can depend
- **Hours vs. Weeks:** Compressed complex processes from several weeks into hours, enabling the delivery of ad hoc requests as the DatabaseUSA team tests new business rules
- **Intuitive Workflow:** Replaced script and low-level code with a simple drag-and-drop visual workflow

### Introduction

DatabaseUSA.com is a leading provider of business and personal mailing lists, email lists, sales leads, and related database products serving most of the Fortune 1000. The company's lists and online products are used by customers large and small for everything from marketing, analytics, and customer profiling to background checks, business verifications, and employee queries—and have been highlighted on the television news program “60 Minutes.” One of the few database providers that still compiles its own complete databases from original sources, DatabaseUSA performs rigorous verifications online, via its call center, and by way of proprietary algorithms, resulting in greater than 95 percent data accuracy.

### Situation

In the competitive landscape of database marketing, offering the “biggest lists” alone is no longer enough. “Even though DatabaseUSA is only four-and-a-half years old, many of us are industry veterans with 10 to 20 years of experience,” relates Jerry Dailey, Vice President of Product and Content at DatabaseUSA. “That’s a lot of accumulated knowledge. We know firsthand that more data is not better—higher quality is better. To that end, we’ve sought to provide more features, greater depth, and better data accuracy than any of our competitors.”

“Prior to Alteryx, raising that bar of data quality was not easy,” recalls Erich Kaiser, Vice President of Data Processing and Special Projects at DatabaseUSA. “Our consumer database has details on 238 million individuals and 164,000 households, and our business database details approximately 14 million businesses. That’s a lot of data to process, check for accuracy, and update on a monthly basis.”

With so much raw data as part of the monthly build process, hitting deadlines for fresh builds was a Herculean task. To deal with this reality, DatabaseUSA sought individuals with high-level data expertise—but that brought high-level cost. Remembers Kaiser, “When we started out, we had to hire people familiar with a Linux environment and shell scripting to support our database builds. Unfortunately, those people don’t come cheaply; each hire cost us \$100,000 to \$150,000 and forced us to put all our eggs in one basket because of the intense specialization required.”

Because the company was working only with flat files from hundreds of different sources and multiple directories, including data from utility companies and call



*“We never really had a locked-down process before. But now, with Alteryx, we can string these components together, create a build process, and lock down that process. And if there is any anomaly during quality control, we know exactly what component to go back to and correct based on what we are seeing—and we don’t miss deadlines.”*

—Erich Kaiser, Vice President of  
Data Processing and Special Projects  
at DatabaseUSA

# alteryx

3345 Michelson Dr., Ste. 400, Irvine, CA 92612  
+1 714 516 2400  
www.alteryx.com

data from its call center, it needed to develop a significant amount of new, low-level code to get the data into its database successfully—and accurately. “Every time we introduced new data we had to custom-code solutions to make it fit,” continues Kaiser. “Not only was our initial platform expensive, but also the knowledge required to build and maintain it was prohibitive and difficult to scale.”

## Solution

Driven by the need for data accuracy and scalability and to simplify and codify its processes, DatabaseUSA actively searched for solutions before landing on *Alteryx Analytics*. “We did an extensive review and comparison of databases, platforms, and tools, including SQL Server, SSIS, Clover, and Melissa,” recalls Dailey. “But, ultimately, Alteryx did a better job of scrubbing and enhancing the data and providing the ETL we needed.”

While enthusiastic about Alteryx, Dailey and his team weren’t certain how much of the heavy lifting Alteryx would ultimately be able to provide them in the areas of ETL and overall data blending. “Initially, we weren’t sure how far we could take Alteryx and thought we might have to use other tools to help us with our needs. But, so far, with the exception of our platform and server environment, we’ve now used Alteryx to do just about everything. Alteryx helps us blend, clean, input, and update the data, test its accuracy, maintain the required database relationships, and deliver the data the way we need it, when we need it.”

Kaiser agrees, elaborating that Alteryx’s successful integration into DatabaseUSA’s methodology stems from how well it handles tasks like standardization and duplicates. “We use Alteryx where all of our data sources come together. For example, address data can be represented in a variety of different ways in terms layouts, formats, values, and fields,” explains Kaiser. “After we use Alteryx to standardize, clean, and de-dupe the data from the different sources, we bring it into our own database to identify which records are truly new.”

DatabaseUSA appreciated how amenable the Alteryx team was in supporting its unconventional deployment requirements—namely, to a hosted server

environment. “Although most customers access Alteryx remotely, we needed to run it on our hosted servers due to data sensitivity issues,” says Daily. “The Alteryx team understood our needs and helped us make this a reality. As of now, we have two rack servers, two application hosted servers, and a data analytics database server with Alteryx in the mix.”

## Results

Since deploying Alteryx, Kaiser has been impressed with how well the solution has risen to new challenges. “There are several tools in the ETL realm,” Kaiser notes, “but none near as strong for data quality, and that’s key for us. All of our ETL work is now done by Alteryx, which blends and cleanses all our external sources and call center data.”

Dailey agrees, adding, “Before Alteryx, I had to do a lot of low-level complex and custom coding. But now, with Alteryx, using fuzzy matching we can apply several different types of matching algorithms on our data by simply changing a configuration setting.”

Prior to Alteryx, DatabaseUSA dealt with each step in its build process as a discrete code issue. “Each step, including transforming, merging, identifying duplicates, removing duplicates, and assigning address standardization enhancements, was its own challenge and deadlines were easy to miss,” recalls Kaiser. “But now, with Alteryx, we can string these components together, create a build process, and lock down that process. And if there is any anomaly during quality control, we know exactly what component to go back to and correct based on what we are seeing—and we don’t miss deadlines.”

“Likewise, with our process locked down and simplified, database build processes have easily shortened from a month to a week,” asserts Dailey. Simplification of the workflow has also relieved the pressures and costs of hiring high-level expertise. “By using Alteryx, I can now recruit new, young computer engineering students, train them, and get them up to speed applying business rules in a very short time. We can accomplish more in two months now with Alteryx than we could in six months with our former environment, custom coding, and expensive hires.”