

# DEMOCRATIZING DATA SCIENCE WITH



Trifacta is an intelligent cloud data service used to explore, clean, and prepare structured and unstructured data for analysis, machine learning, and reporting. Trifacta makes messy data understandable with ensured data quality and automates data pipelines in just a few clicks. The AI-driven data-wrangling built-ins allows both technical and non-technical users to get creative and manipulate the data they need. The power of insights drawn from data now falls into the hands of all your employees, with lower cost and accelerated time.



## Why Trifacta?

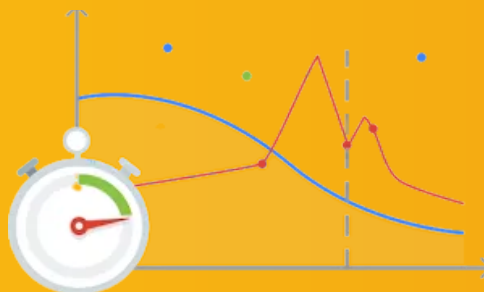
### Sever-less and Interoperable

- Seamless user experience without the need for up-front hardware or software installations
- Fully managed platform that scales easily as your data size grows
- The only cloud native data preparation product across all leading clouds



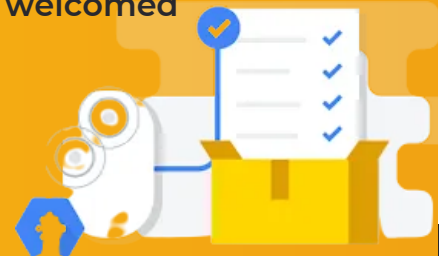
### Fast and Intelligent

- Instant data visualizations about your data distributions
- Automatically detects schemas, data types, anomalies such as missing values and outliers so you can skip time assessing your data and go right into your analysis



### Automated Pipelines

- Artificial Intelligence automatically predicts your next ideal data transformation, enabling processing structured or unstructured data with just a few clicks, not blocks of code
- Data engineering pipelines can be built at scale in minutes not month. Non-IT users also welcomed



**Contact US**  
Team 6



Katherine Li  
lix5213@umn.edu



Xiaoshuo Liu  
liu00860@umn.edu

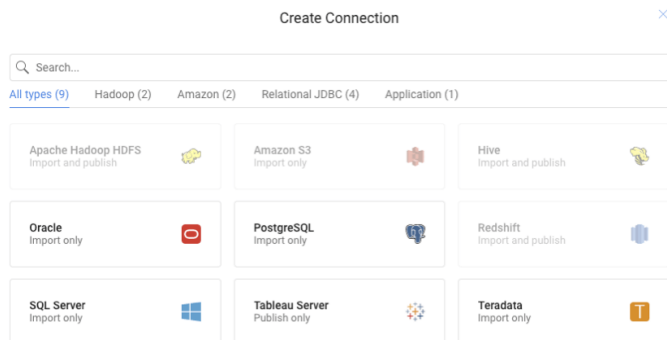


Kyle Kuznia  
kuzni076@umn.edu

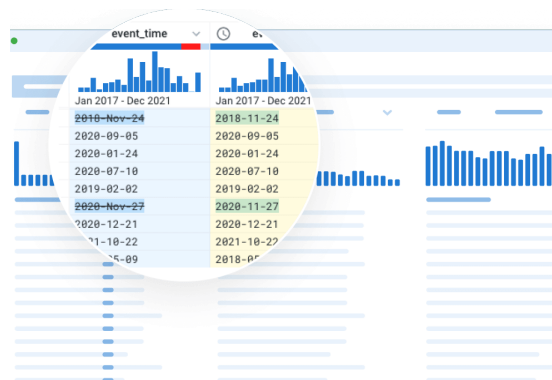


Vikram Payyavula  
payya005@umn.edu

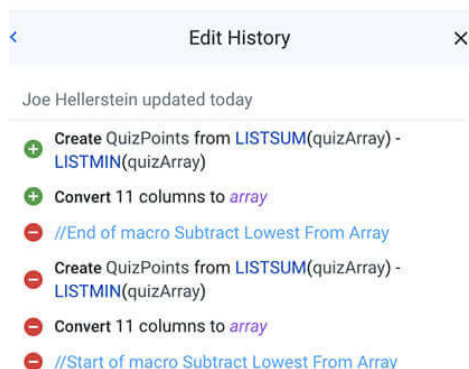
# How Does It Work?



## 1. Connect



## 2. Transform



## 3. Collaborate



## 4. Automate

## Use Cases

### Consumer Goods

**Challenges:** PepsiCo needs to analyze performance sales from 10+ retailers. Each analyst has to build and update 4+ reports weekly, which was slow and laborious.

**Solution:** Trifacta helped analysts to wrangle data at the organizational level. Reporting time was reduced by 70%; build time was reduced by 90%, this allowed Pepsi to respond to sales trends quickly to reduce cost.



### Research

**Challenges:** Epidemiological data sources are diverse and messy, so manual data preparation was time-consuming. Analysts spend weeks to merge data and can't spot data issue easily.

**Solution:** Trifacta allowed CDC to accomplish the same work that took 3 months in 3 days, with fewer resources. Data issues can be quickly detected and subject experts have increased access to view and manipulate raw data.



### Software

**Challenges:** Data prep for machine learning (fraud detection) takes about 4 weeks. Data inaccuracies risk millions of dollars, and the BI team couldn't prepare data without requesting to IT.

**Solution:** Trifacta was implemented from the AWS marketplace. Model can be developed in several days now. Millions in fraud loss saved with the new and fast model. Product and BI teams can work directly with raw data.



## Rated #1 by Users



Rated 4.6 out of 5 in Gartner Peer Insights for Data Preparation Tools



Dresner

Highest ranked Data Preparation Vendor by Dresner Advisory Services for 6 years



Highest ranked Data Prep Vendor by users in G2