



Fact of Death Mortality Index



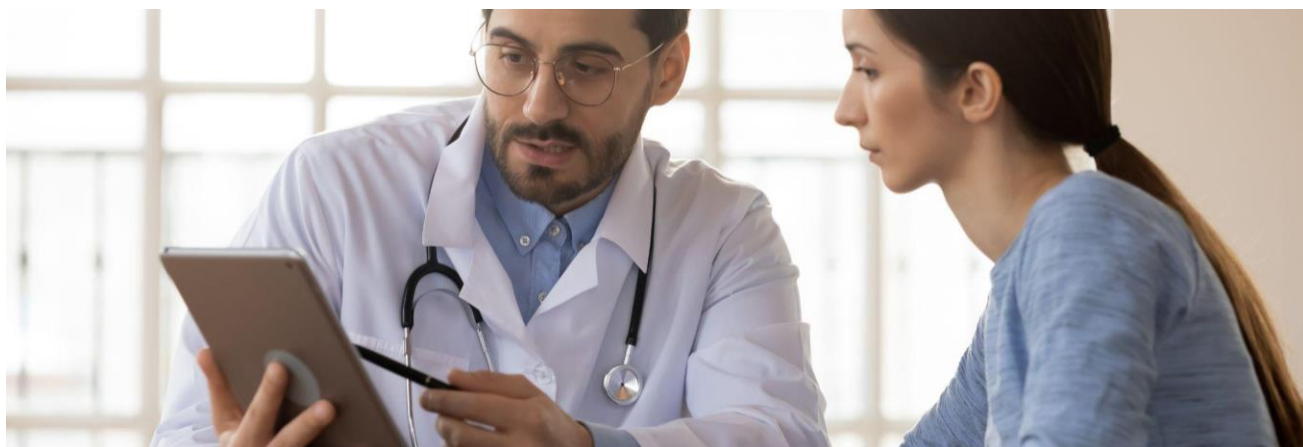


Table of contents

Overview	2
Product Attributes	2
Completeness	2
Timeliness	3
Quality Monitoring and Internal Validation	3
External Validation	4
Transparency	4
Delivery Parameters	5
Data Preparation and Curation	5
Data Dictionary	6

Overview

Datavant has entered into partnership with Veritas Data Research to offer Veritas' Fact of Death mortality index as the preferred (non-exclusive) source of death data in Datavant's Ecosystem. Datavant customers may now subscribe to Veritas' Fact of Death and receive weekly/monthly/quarterly refreshes.

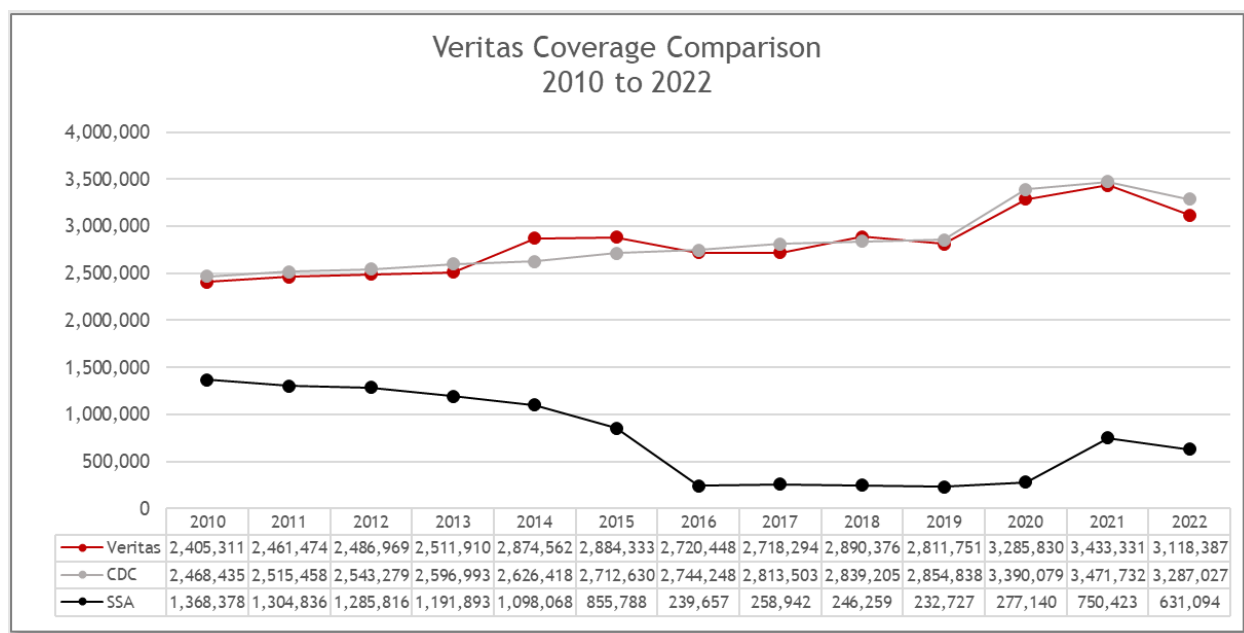
Veritas Data Research's Fact of Death contains over 35 million death records from 2010 to present and is the most complete, timely, accurate, transparent, and representative index of U.S. mortality data.

Note: The full historical data from the Social Security Administration's death master file is also available from 1935 onwards.

Product Attributes

Completeness

Veritas indexes death data from over 40,000 online and offline sources. When compared to official government sources, Veritas counts are a close match to the U.S. Centers for Disease Control and Prevention's (CDC) National Death Index (NDI), which is considered the "gold standard" for mortality data.

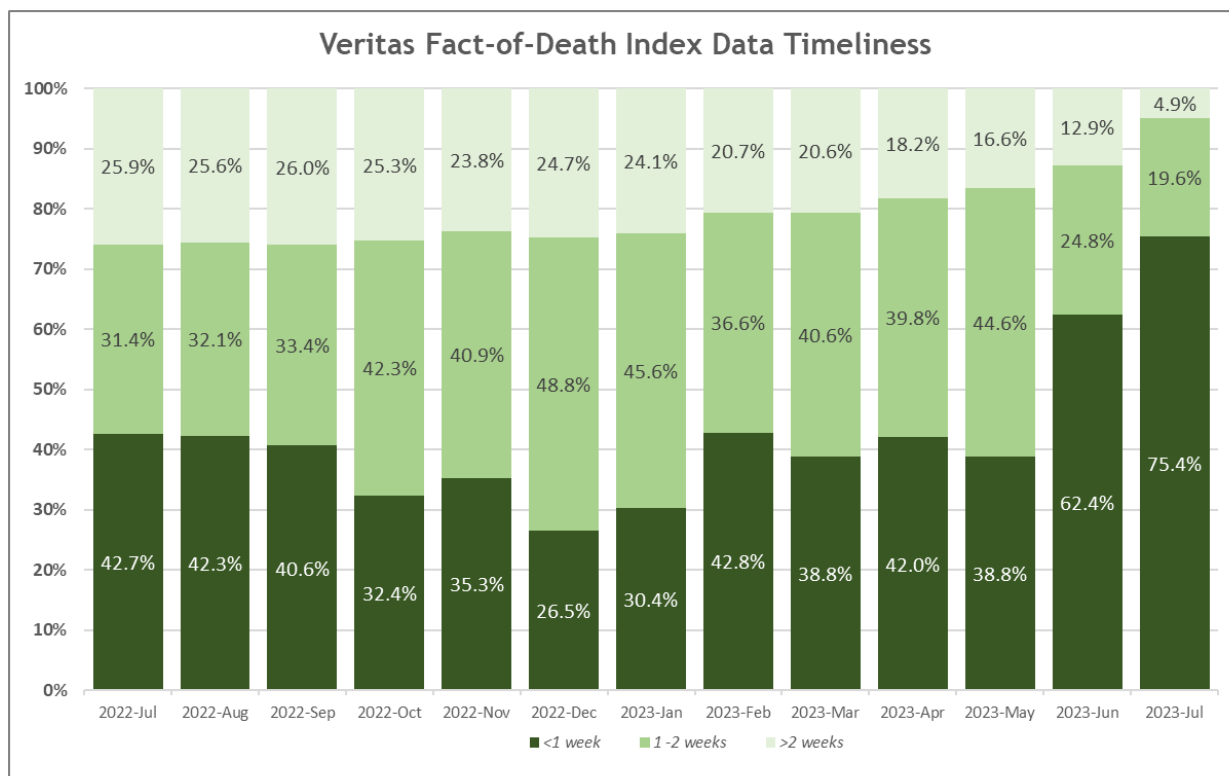


Last updated: August 5, 2023



Timeliness

Veritas indexes 90% of all deaths in Fact of Death within 30 days of an individual's death, making it one of the timeliest mortality data sources available.



Last updated: August 5, 2023

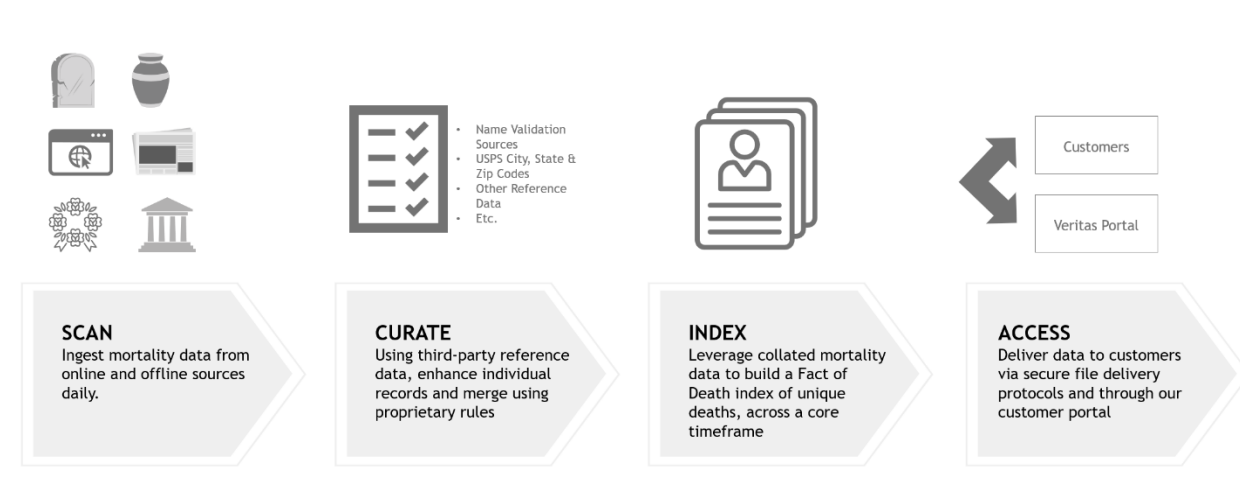
Quality Monitoring and Internal Validation

By continually indexing data from multiple sources, Veritas is able to cross-check the quality and accuracy of records. The Fact of Death leverages highly credible industry-standard reference data to cleanse and improve the dataset's quality. For example, all dates are reformatted into a standard YYYY-MM-DD format, all names are cross-checked against a list of valid names provided by the U.S. Census Bureau, and all geography is standardized and compared to the U.S. Postal Service (USPS)

reference dataset. Data not meeting Veritas’s rigorous data quality standards are re-checked for validation on a set schedule.

Fact of Death contains records from multiple sources, including the Social Security Administration’s Limited Access Death Master File (“SSA LADMF”), and public announcements of deaths, burials/internments and other death rituals (“Other sources”). Over 70% of Fact of Death index records have more than one source reporting the death, and utilizing death reports from multiple sources is particularly important as a single source can be unreliable or provide inconsistent information.

Fact of Death records are de-duplicated and aggregated into a single curated index that can be licensed for commercial and research use cases. **The indexing and curation process is outlined in the diagram below:**



External Validation

Fact of Death has been successfully reviewed by multiple customers and healthcare researchers for its veracity. In addition, several national studies of composite sources of mortality data have shown that sources that combine public death notices and data from government sources outperform stand-alone data sources, including mortality information collected from electronic health records (EHR) ([da Graca, Filardo, and Nicewander 2013](#); [Curtis et al. 2018](#); [Zhang 2021](#)).

Transparency

Fact of Death is designed to meet the emerging FDA regulatory guidelines for data provenance, quality, and transparency, and Veritas is able to provide traceability to original sources for every record for additional research and validation.

Contributing Sources

Veritas indexes records from government and public notifications/announcements of death. A **representative sample** of sources are provided below:

- **Government and Public Health Agencies** (requires use-case based permissions)
 - Social Security Limited Access Death Master File (SSA LADMF)
 - State death registry files covering over 20% of the U.S. population
- **Death Notices**
 - Obituaries
 - Memorial services
 - Death notices in public news reports
- **Internments**
 - Public cemetery rosters
 - Military cemetery rosters
 - Crematoria

Delivery Parameters

Data Preparation and Curation

The Fact of Death is available in a de-identified, tokenized format in the Datavant Ecosystem, and subscriptions are available for 2010-current and pre-2010 records.

Prior to de-identification and tokenization, Veritas ensures each record has these four fields:

1. First Name
2. Last Name
3. Date of Birth
4. Date of Death



Additionally, fields are standardized and validated against private and referential data sources to ensure quality and improve matching. Only records meeting a minimum threshold are included in the final Fact of Death offering.

Data Dictionary

Field	Description	Format	2010-Current Fill Rate	Pre-2010 Fill Rate
token_1	Datavant Token 1: <ul style="list-style-type: none"> Last Name First Name (First Character) Gender DOB 	Datavant Token	99.6%	100%
token_2	Datavant Token 2: <ul style="list-style-type: none"> Last Name Soundex First Name Soundex Gender DOB 	Datavant Token	98.8%	97.3%
token_4	Datavant Token 4: <ul style="list-style-type: none"> Last Name First Name Gender DOB 	Datavant Token	98.8%	97.3%
token_5	Datavant Token 5: <ul style="list-style-type: none"> SSN Gender DOB 	Datavant Token	26.3%	98.9%
token_16	Datavant Token 16: <ul style="list-style-type: none"> SSN First Name 	Datavant Token	25.9%	96.2%
token_101	Datavant Token 101: <ul style="list-style-type: none"> Last Name First Name (Confident Resolution) DOB 	Datavant Token	98.8%	97.3%
token_102	Datavant Token 102: <ul style="list-style-type: none"> Last Name First Name (Fuzzy Resolution) DOB 	Datavant Token	98.8%	97.3%
token_103	Datavant Token 103: <ul style="list-style-type: none"> Last Name First Name (Fuzzy Resolution, Metaphone) DOB 	Datavant Token	98.8%	97.3%
token_106	Datavant Token 106: <ul style="list-style-type: none"> Last Name 	Datavant Token	88.8%	95.7%

Field	Description	Format	2010-Current Fill Rate	Pre-2010 Fill Rate
	<ul style="list-style-type: none"> First Name (Fuzzy Resolution, Metaphone) DOB ZIP3 			
token_encryption_key	Proprietary Datavant Information	Datavant Info	NA	NA
traceabilityNbr	Datavant-specific “upsert” tracker	String (#####)	100%	100%
deathDate	Date of death	Date (YYYY-MM-DD)	100%	100%
gender	Highest probability gender identification	String (M/F)	100%	100%
genderProbabilityScore	Probability of gender identification provided	String (0 – 100)	100%	100%
state	2-digit US state code	String (xx)	89.0%	97.4%
zip3	3-digit USPS zip code	String (000 – 999)	88.3%	97.4%
dobImputeFlags	Date of birth imputation flags. 3-letter combination of “Y” and “N” <ul style="list-style-type: none"> 1st letter indicates if year was imputed. 2nd letter indicates if month was imputed 3rd letter indicates if day was imputed 	String (xxx)	100%	100%
dodImputeFlags	Date of death imputation flags. 3-letter combination of “Y” and “N” <ul style="list-style-type: none"> 1st letter indicates if year was imputed. 2nd letter indicates if month was imputed 3rd letter indicates if day was imputed 	String (xxx)	100%	100%

Last updated: August 5, 2023

About Veritas Data Research

Founded by experts in the data analytics industry, Veritas uses cutting-edge technology and efficient workflow design to collect, curate, and distribute foundational reference datasets.

Contact Veritas



<https://www.veritasdataresearch.com>



info@veritasdataresearch.com

About Datavant

Datavant's mission is to connect the world's health data to improve patient outcomes. Datavant works to reduce the friction of data sharing across the healthcare industry by building technology that protects the privacy of patients while supporting the linkage of patient health records across datasets.

Contact Datavant



<https://www.datavant.com>



info@datavant.com