IO-AVSTATS-DB - User Manual

Release 23.10.01

IO Aeronautical Autonomy Labs, LLC

© IO-Aero 2023

1. Introduction

The National Transportation Safety Board (NTSB) investigates all aviation accidents in the U.S. and makes the investigation results available on their website (data.ntsb.gov, n.d.) in Microsoft Access database files for public use. The NTSB provides aviation event data from 1982 through 2007 in the file **Pre2008.zip** (version 9/30/2020 12:51:56 PM). Data since 2008 are available first in the overall **avall.zip** file, which is updated monthly, and second with a quasi-weekly amendment file each month on the 1st, 8th, 15th, and 22nd, e.g., **up22JUN.zip**.

IO-AVSTATS-DB is IO-Aero's database version of NTSB's aviation event data. **IO-AVSTATS-DB** is based on a PostgreSQL database (The PostgreSQL Global Development Group, 2019) and contains, in addition to the data from NTSB IO-Aero specific extensions and data from the following sources:

- Aviation Occurrence Categories (AVIATION OCCURRENCE CATEGORIES DEFINITIONS AND USAGE NOTES, 2013)
- GeoDatos (Geodatos.net, 2019)
- opendatasoft (Opendatasoft, n.d.)
- simplemaps (simplemaps.com, n.d.)
- United States Zip Codes.org (UnitedStatesZipCodes, n.d.)

The **IO-AVSTATS-DB** is updated on the 8th, 15th, and 22nd of each month with the update file provided by NTSB. Every first of the month a new database version is created from the two files **Pre2008.zip** and **avall.zip**. This ensures that the deletions made by NTSB, which are not included in the update file, are reflected in the **IO-AVSTATS-DB** database.

2. Getting Started

IO-Aero provides the database files in PostgreSQL format. The development of **IO-AVSTATS-DB** was done using PostgreSQL version 15 in the form of the official Docker images available on DockerHub (hub.docker.com. (n.d.)). This is also the recommended installation form for using **IO-AVSTATS-DB**.

3. Database Schema

3.1. NTSB

The documentation provided by NTSB on the data provided consists of the following two documents (data.ntsb.gov. (n.d.)):

- codman.pdf —-aviation coding manual
- eadmspub.pdf – database schema diagram

For the data provided by NTSB in the form of MS Access databases, the data dictionary was extracted using RazorSQL (razorsql.com, n.d.) and transferred to the PostgreSQL database **IO-AVSTATS-DB**. This includes the following tables:

- `aircraft`
- `dt_aircraft`
- `dt_events`
- `dt_flight_crew`
- `engines`
- `events`
- `Events_Sequence`
- `Findings`
- `Flight_Crew`
- `flight_time`
- `injury`
- `narratives`
- `NTSB_Admin`
- `Occurrences`
- `seq_of_events`

During the transfer to PostgreSQL, the spelling of all table and column names was unified to lowercase letters. Furthermore, the following adjustments were made to the data types:

MS Access	PostgreSQL
`bit`	`boolean`
`byte`	`char(1)`
`datetime`	`timestamp`
`double`	`float`
`integer`	`int`
`longchar`	`text`

3.2. IO-Aero specific supplements

3.2.1. Column io_last_seen_ntsb

This database column of the timestamp data type has been added to all NTSB tables to document when a specific row was created or last modified.

3.2.2. Indexes

table / view	column(s)
aircraft	acft_category
events_sequence	occurrence_code
findings	finding_code
io_app_ae1982	cictt_codes
io_app_ae1982	ev_highest_injury
io_app_ae1982	ev_type
io_app_ae1982	far_parts
io_app_ae1982	inj_f_grnd
io_app_ae1982	inj_tot_f
io_app_ae1982	latlong_acq
io_app_ae1982	no_aircraft
io_app_ae1982	preventable_events
io_app_ae1982	tll_parameters

3.2.3. Materialized View io_app_ae1982

This materialized view provides all necessary data for IO-Aero's data analysis applications.

3.2.3.1. Original Data - Unconditional

- **events.ev_highest_injury**
- **events.ev_id**
- **events.ev_month**
- **events.ev_type**
- **events.ev_year**
- **events.inj_f_grnd**
- **events.inj_tot_f**
- **events.io_dec_lat_lng_actions** as **dec_lat_lng_actions**
- **events.io_invalid_latitude** as **is_invalid_latitude**
- **events.io_invalid_longitude** as **is_invalid_longitude**
- **events.io_invalid_us_city** as **is_invalid_us_city**
- **events.io_invalid_us_city_zipcode** as **is_invalid_us_city_zipcode**
- **events.io_invalid_us_state** as **is_invalid_us_state**
  events.io_invalid_us_zipcode as **is_invalid_us_zipcode**
- **events.io_nearest_airport_distance** as **nearest_airport_distance**
- **events.io_nearest_airport_global_id** as **nearest_airport_global_id**
- **events.lat_lng_acq**
- **events.ntsb_no**
- **io_airports.ident** as **nearest_airport_ident**
- **io_airports.servcity** as **nearest_airport_servcity**

3.2.3.2. Original Data - with Conditions

- **events.ev_dow** - capital letters
- **events.io_city** - if empty **events.ev_city**
- **events.io_country** - if empty **events.ev_country**
- **events.io_dec_latitude** - if empty **events.ev_dec_latitude**
- **events.io_dec_latitude_deviating** - if empty **0**
- **events.io_dec_longitude** - if empty **events.ev_dec_longitude**
- **events.io_dec_longitude_deviating** - if empty **0**
- **events.io_site_zipcode** - if empty **events.ev_site_zipcode**
- **events.io_state** - if empty **events.ev_state**

3.2.3.3. Logical Variables

3.2.3.3.1. Level Base

- **acft_categories** - an array of **aircraft.acft_category**
- **all_category_codes** - an array of **findings.category_no**
- **all_category_cause_codes** - an array of **findings.category_no** where **cause_factor** is '**C**'
- **all_category_factor_codes** - an array of **findings.category_no** where **cause_factor** is '**F**'
- **all_category_none_codes** - an array of **findings.category_no** where **cause_factor** is neither '**C**' nor '**F**'

- **all_eventsoe_codes** - an array of **events_sequence.eventsoe_no**
- **all_eventsoe_false_codes** - an array of **events_sequence.eventsoe_no** where **defining_ev** is not true
- **all_eventsoe_true_codes** - an array of **events_sequence.eventsoe_no** where **defining_ev** is true
- **all_finding_codes** - an array of **findings.finding_no**
- **all_finding_cause_codes** - an array of **findings.finding_no** where **cause_factor** is '**C**'
- **all_finding_factor_codes** - an array of **findings.finding_no** where **cause_factor** is '**F**'
- **all_finding_none_codes** - an array of **findings.finding_no** where **cause_factor** is neither '**C**' nor '**F**'
- **all_modifier_codes** - an array of **modifiers.modifier_no**
- **all_modifier_cause_codes** - an array of **modifiers.modifier_no** where **cause_factor** is '**C**'
- **all_modifier_factor_codes** - an array of **modifiers.modifier_no** where **cause_factor** is '**F**'
- **all_modifier_none_codes** - an array of **modifiers.modifier_no** where **cause_factor** is neither '**C**' nor '**F**'
- **all_occurrence_codes** - an array of **events_sequence.occurrence_no**
- **all_occurrence_false_codes** - an array of **events_sequence.occurrence_no** where **defining_ev** is not true
- **all_occurrence_true_codes** - an array of **events_sequence.occurrence_no** where **defining_ev** is true
- **all_phase_codes** - an array of **events_sequence.phase_no**
- **all_phase_false_codes** - an array of **events_sequence.phase_no** where **defining_ev** is not true
- **all_phase_true_codes** - an array of **events_sequence.phase_no** where **defining_ev** is true
- **all_section_codes** - an array of **sections.section_no**
- **all_section_cause_codes** - an array of **sections.section_no** where **cause_factor** is '**C**'
- **all_section_factor_codes** - an array of **sections.section_no** where **cause_factor** is '**F**'
- **all_section_none_codes** - an array of **sections.section_no** where **cause_factor** is neither '**C**' nor '**F**'
- **all_subcategory_codes** - an array of **subcategorys.subcategory_no**
- **all_subcategory_cause_codes** - an array of **subcategorys.subcategory_no** where **cause_factor** is '**C**'
- **all_subcategory_factor_codes** - an array of **subcategorys.subcategory_no** where **cause_factor** is '**F**'
- **all_subcategory_none_codes** - an array of **subcategorys.subcategory_no** where **cause_factor** is neither '**C**' nor '**F**'
- **all_subsection_codes** - an array of **subsections.subsection_no**
- **all_subsection_cause_codes** - an array of **subsections.subsection_no** where **cause_factor** is '**C**'
- **all_subsection_factor_codes** - an array of **subsections.subsection_no** where **cause_factor** is '**F**'
- **all_subsection_none_codes** - an array of **subsections.subsection_no** where **cause_factor** is neither '**C**' nor '**F**'
- **cictt_codes** - an array of **events_sequence.cictt_code**

- **description_main_phase_defining** - an array of **dewcription_main_phase** where **defining_ev** is true
- **dest_countries** - an array of **aircraft.dest_country**
- **dprt_countries** - an array of **aircraft.dprt_country**
- **ev_highest_injury**: Categorization of missing data as '**n/a**'.
- **far_parts** - an array of **aircraft.far_part**
- **finding_codes** - an array of encoded **findings.finding_code**:
  - '**PARAMS_ALT**' if position 1 to 8 of **finding_code** is '**01062012**'
  - '**PARAMS_DEC_RATE**' if position 1 to 8 of **finding_code** is '**01062037**'
  - '**PARAMS_DEC_APP**' if position 1 to 8 of **finding_code** is '**01062040**'
  - '**PARAMS_AoA**' if position 1 to 8 of **finding_code** is '**01062042**'
  - '**ENV_TER** if position 1 to 6 of **finding_code** is '**030210**'
  - '**ENV_OAS**' if position 1 to 6 of **finding_code** is '**030220**'
- **is_attitide_controllable** –if:
  - position 1 to 6 of **finding_code** is one of
    - '**010355**', '**010357**', '**030320**', '**030330**'
  - or position 1 to 8 of **finding_code** is one of
    - '**01061040**', '**03031025**', '**03034020**', '**03034030**', '**03034050**', '**03034060**'
  - or **eventsoe_no** is one of
    - **'210'**, **'245'**, **'333'**, **'361'**,
  - or **eventsoe_no** is one of '**337**', '**338**' and **defining_ev**
- **is_midair_collision** –if:
  - **eventsoe_no** is '**250**' and **defining_ev**
- **is_narrative_stall** –if:
  - **narratives.narr_accp** is like '**%STALL%**
- **is_pilot_issue** –if:
  - **modifier_no** is one of '**44**', '**45**', '**46**'
- **latlong_acq** is '**none**' if
  - **latlong_acq** is null
  - and **io_latlong_acq** is null
  - and one of **dec_latitude**, **dec_longitude**, **io_dec_latitude**, **io_dec_longitude** is not null
- **no_aircraft** –the number of aircraft involved in the event
- **occurrence_codes** - an array of encoded **events_sequence.occurrence_code**:
  - '**INIT_CLIMB**' if **phase:no** is '**350**'
  - '**MAN_LALT**' if **phase:no** is '**452**'
  - '**FINAL_APP**' if **phase:no** is '**502**'
  - '**CFIT**' if **eventsooe_no** is '**120**'
  - '**LALT**' if **eventsooe_no** is '**220**'
  - '**LOC-I**' if **eventsooe_no** is '**240**'
  - '**STALL**' if **eventsooe_no** is '**241**'
  - '**MIDAIR**' if **eventsooe_no** is '**250**'
  - '**UIMC**' if **eventsooe_no** is '**401**'
  - '**CAA**' if **eventsooe_no** is '**420**'
  - '**BIRD**' if **eventsooe_no** is '**901**'
- **oper_countries** - an array of **aircraft.oper_country**
- **owner_countries** - an array of **aircraft.owner_country**
- **phase_codes_defining** - an array of **events_sequence.phase_no** where **defining_ev** is true
- **regis_countries** - an array of
  - '**USA**' if **regis_no** has one of the formats '**N99999**', '**N9999A**', '**N999AA**',
  - '**NON-US**' else

- **regis_nos** - an array of **aircraft.regis_no**

3.2.3.3.2. Level 1

- **is_dest_country_usa** –if '**USA**' is in **dest_countries**
- **is_dprt_country_usa** –if '**USA**' is in **dprt_countries**
- **is_emergency_landing** –if:
  - **modifier_no** is one of
    - '**01**', '**02**', '**03**', '**06**', '**22**', '**23**', '**24**', '**25**', '**26**'
  - or position 1 to 6 of **finding_code** is one of
    - '**010224**', '**010228**', '**010230**', '**010461**', '**010462**', '**010465**', '**010467**', '**010571**', '**010572**', '**010573**', '**010574**', '**010578**', '**010579**', '**010581**', '**010585**'
  - or position 1 to 8 of **finding_code** is one of
    - '**01050000**', '**01061010**', '**01061020**', '**01062025**', '**01071000**', '**01071010**'
  - or **phase_no** is '**600**'
  - or **eventsoe_no** is one of
    - '**130**', '**140**', '**190**', '**191**', '**192**', '**193**', '**194**', '**340**', '**341**', '**342**', '**343**', '**440**', '**441**', '**500**', '**901**',
  - or **eventsoe_no** is one of '**337**', '**338**' and **defining_ev**
- **is_far_part_091x** –if **far_parts** is any of
  - '**091**'
  - '**091F**'
  - '**091K**'
- **is_far_part_121** –if '**121**' is in **far_parts**
- **is_far_part_135** –if '**135**' is in **far_parts**
- **is_oper_country_usa** –if '**USA**' is in **oper_countries**
- **is_owner_country_usa** –if '**USA**' is in **owner_countries**
- **is_regis_country_usa** –if '**USA**' is in **regis_countries**
- **is_spin_stall** –if
  - **(** '**PARAMS_AoA**' is in **findings_codes**
  - or '**STALL**' is in **occurrence_codes**
  - or '**LOC-I**' is in **occurrence_codes** and **is_narrative_stall)**
  - and not **(occurrence_codes** is any of '**CAA**', '**CFIT**' **)**

3.2.3.3.3. Level 2

- **is_altitude_controllable** –if one of the following logical variables is true:
  - **is_attitude_controllable**
  - **is_emergency_landing**
  - **is_spin_stall**
- **is_altitude_low** –if:
  - **(occurrence_codes** is one of
    - '**CAA**',
    - '**CFIT**',
    - '**FINAL_APP**',
    - '**INIT_CLIMB**',
    - '**LALT**',
    - '**MAN_LALT**'
  - or **findings_codes** is '**ENV_OAS**' and **occurrence_codes** is not '**BIRD**'
  - or **findings_codes** is one of
    - '**ENV_TER**'
    - '**PARAMS_ALT**'

        - '**PARAMS_DEC_APP**'
        - '**PARAMS_DEC_RATE**')
    - and not (**occurrence_codes** is '**MIDAIR**' and **is_spin_stall** is false)

3.2.3.3.4. Level 3

- **is_rss_forced_landing** –if one of the following logical variables is true:
    - **is_attitude_controllable**
    - **is_emergency_landing**
- **is_rss_spin_stall_prevention_and_recovery** –if one of the following logical variables is true:
    - **is_attitude_controllable**
    - **is_spin_stall**
- **is_rss_terrain_collision_avoidance** –if one of the following logical variables is true:
    - **is_altitude_controllable**
    - **is_altitude_low**
    - **is_attitude_controllable**

3.2.3.3.5. Level 4

- **cictt_codes**: Categorization of missing data as '**no data**'.
- **far_parts**: Categorization of missing data as '**no data**'.
- **is_lp_n_a** – a logical parameter is not available if none of the following logical variables is true:
    - **is_altitude_low**
    - **is_attitude_controllable**
    - **is_emergency_landing**
    - **is_midair_collision**
    - **is_pilot_issue**
    - **is_rss_forced_landing**
    - **is_rss_spin_stall_prevention_and_recovery**
    - **is_rss_terrain_collision_avoidance**
    - **is_spin_stall**
- **is_rss_n_a** – a required safety system is not available if none of the following logical variables is true:
    - **is_midair_collision**
    - **is_rss_forced_landing**
    - **is_rss_spin_stall_prevention_and_recovery**
    - **is_rss_terrain_collision_avoidance**
- **phase_codes_defining**: Categorization of missing data as '**no data**'.
- **preventable_events**:
    - **Airborne collision** if **is_midair_collision**
    - **Forced landing** if **is_rss_forced_landing**
    - **Spin / stall** if **is_rss_spin_stall_prevention_and_recovery**
    - **Terrain collision** if **is_rss_terrain_collision_avoidance**
    - **Not preventable** if none of the above
- **tll_parameters** – top logical parameters:
    - **Aerodynamic spin / stall** if **is_spin_stall**
    - **Aircraft can climb** if **is_altitude_controllable**
    - **Aircraft has degraded control failure** if **is_emergency_landing**
    - **Altitude too low** if **is_altitude_low**
    - **Attitude is controllable** if **is_attitude_controllable**
    - **Pilot is able to perform maneuver** if **is_pilot_issue**
    - **n/a** if none of the above

Sample data:

Name	Value
ev_id	20221112106276
ntsb_no	CEN23MA034
ev_type	ACC
ev_year	2022
ev_month	11
ev_dow	SA
country	USA
state	TX
city	Dallas
zip	75237
acft_categories	{AIR,AIR}
all_category_codes	{}
all_category_cause_codes	{}
all_category_factor_codes	{}
all_category_none_codes	{}
all_eventsoe_codes	{ee_250_a}
all_eventsoe_false_codes	{}
all_eventsoe_true_codes	{ee_250_t}
all_finding_codes	{}
all_finding_cause_codes	{}
all_finding_factor_codes	{}
all_finding_none_codes	{}
all_modifier_codes	{}
all_modifier_cause_codes	{}
all_modifier_factor_codes	{}
all_modifier_none_codes	{}
all_occurrence_codes	{eo_452250_a}
all_occurrence_false_codes	{}
all_occurrence_true_codes	{eo_452250_t}
all_phase_codes	{ep_452_a}
all_phase_false_codes	{}
all_phase_true_codes	{ep_452_t}
all_section_codes	{}
all_section_cause_codes	{}
all_section_factor_codes	{}
all_section_none_codes	{}
all_subcategory_codes	{}
all_subcategory_cause_codes	{}
all_subcategory_factor_codes	{}
all_subcategory_none_codes	{}
all_subsection_codes	{}
all_subsection_cause_codes	{}
all_subsection_factor_codes	{}
all_subsection_none_codes	{}
cictt_codes	MAC
dec_lat_lng_actions	
dec_latitude	32.673779
dec_latitude_deviating	0.0
dec_longitude	-96.862801
dec_longitude_deviating	0.0
description_main_phase_defining	Manuevering
dest_countries	{n/a,n/a}
dprt_countries	{n/a,n/a}
ev_highest_injury	FATL

far_parts	091
finding_codes	{}
inj_f_grnd	0
inj_tot_f	6
is_altitude_controllable	true
is_altitude_low	true
is_attitude_controllable	true
is_dest_country_usa	false
is_dprt_country_usa	false
is_emergency_landing	false
is_far_part_091x	true
is_far_part_121	false
is_far_part_135	false
is_invalid_latitude	false
is_invalid_longitude	false
is_invalid_us_city	false
is_invalid_us_city_zipcode	false
is_invalid_us_state	false
is_invalid_us_zipcode	false
is_lp_n_a	false
is_midair_collision	true
is_narrative_stall	false
is_oper_country_usa	true
is_owner_country_usa	true
is_pilot_issue	false
is_regis_country_usa	true
is_rss_forced_landing	false
is_rss_n_a	false
is_rss_spin_stall_prevention_and_recovery	false
is_rss_terrain_collision_avoidance	true
is_spin_stall	false
latlong_acq	MEAS
nearest_airport_distance	0.6260786976106293
nearest_airport_global_id	18C8BDB4-18B7-44…
nearest_airport_ident	RBD
nearest_airport_servcity	DALLAS
no_aircraft	2
occurrence_codes	{MAN_LALT,MAN_LALT}
oper_countries	{USA,USA}
owner_countries	{USA,USA}
phase_codes_defining	452
preventable_events	Airborne collision,
regis_countries	{NON-US,USA}
regis_nos	{N6763,N7227C}
tll_parameters	Aircraft can climb,

3.2.4. Table events

3.2.4.1. Column io_latlong_acq

Extending the categories used in the **latlong_acq** column:

- **LOLA**: Correction based on swapped latitude and longitude
- **LALO**: Correction based on latitude and longitude
- **ERRA**: Invalid latitude string
- **ERRO**: Invalid longitude string

3.2.4.2. Corrections with csv files

The values of the columns **city**, **country**, **latitude**, **longitude**, **site_zipcode** and **state** can be corrected via csv files and the changed value is used in all subsequent processing. The changed values are stored in the following IO-Aero specific columns:

- **io_city**
- **io_country**
- **io_latitude**
- **io_longitude**
- **io_site_zipcode**
- **io_state**

3.2.4.3. Determination of the missing values of dec_latitude / dec_longitude

- **io_dec_lat_lng_actions** - shows the actions performed to find the missing values in the **dec_latitude** and **dec_longitude** columns.
- **io_dec_latitude** - the found value for the **dec_latitude** column.
- **io_dec_latitude_deviating** - Difference between latitude and decimal latitude if a predefined threshold value is exceeded.
- **io_dec_longitude** - the found value for the **dec_longitude** column.
- **io_dec_longitude_deviating** - Difference between longitude and decimal longitude if a predefined threshold value is exceeded.

3.2.4.4. Error Flags

- **io_invalid_latitude** - invalid value in column **latitude**.
- **io_invalid_longitude** - invalid value in column **longitude**.
- **io_invalid_us_city** - unknown value in **ev_city** column.
- **io_invalid_us_city_zipcode** - unknown combination of values in columns **ev_site_zipcode** and **ev_city**.
- **io_invalid_us_state** - unknown value in **ev_state** column.
- **io_invalid_us_zipcode** - unknown value in **ev_site_zipcode** column.

3.2.4.5. Nearest airport

These columns determine the nearest airport for events on U.S. soil.

- **io_nearest_airport_distance** - distance in miles to the nearest airport.
- **io_nearest_airport_global_id** – global identification of the nearest airport.

3.2.5. Table io_airports

This database table contains data on airports provided by the FAA.

	Name	Type	Primary	Not Null
1	global_id	varchar(254)	☑	☑
2	airanal	varchar(254)	☐	☐
3	country	varchar(254)	☐	☐
4	dec_latitude	double precision	☐	☐
5	dec_longitude	double precision	☐	☐
6	dodhiflib	smallint	☐	☐
7	elevation	double precision	☐	☐
8	far91	smallint	☐	☐
9	far93	smallint	☐	☐
10	iapexists	smallint	☐	☐
11	ident	varchar(254)	☐	☐
12	latitude	varchar(254)	☐	☐
13	longitude	varchar(254)	☐	☐
14	max_runway_comp_code	varchar(254)	☐	☐
15	max_runway_length	double precision	☐	☐
16	mil_code	varchar(50)	☐	☐
17	name	varchar(254)	☐	☐
18	operstatus	varchar(254)	☐	☐
19	privateuse	smallint	☐	☐
20	servcity	varchar(254)	☐	☐
21	state	varchar(254)	☐	☐
22	type_code	varchar(254)	☐	☐
23	first_processed	timestamp without time zone	☐	☑
24	last_processed	timestamp without time zone	☐	☐

3.2.5.1. Sample data:

```
Name                |Value                                |
--------------------+-------------------------------------+
global_id           |68E20291-FCD0-46C3-91B3-4A0B3CCA9D82|
airanal             |NOT ANALYZED                         |
country             |USA                                  |
dec_latitude        |36.1984000191282                     |
dec_longitude       |-95.8881130841036                    |
dodhiflib           |1                                    |
elevation           |677.5                                |
far91               |0                                    |
far93               |0                                    |
iapexists           |0                                    |
ident               |TUL                                  |
latitude            |36-11-54.2170N                       |
longitude           |095-53-17.1800W                      |
max_runway_comp_code|CONC                                 |
max_runway_length   |10000.0                              |
```

```
mil_code            |CIVIL                                |
name                |Tulsa Intl                           |
operstatus          |OPERATIONAL                          |
privateuse          |0                                    |
servcity            |TULSA                                |
state               |OK                                   |
type_code           |AD                                   |
first_processed     |2023-04-19 15:51:05.225              |
last_processed      |2023-04-19 15:51:09.217              |
```

3.2.5.2. Data source - FAA Airports:

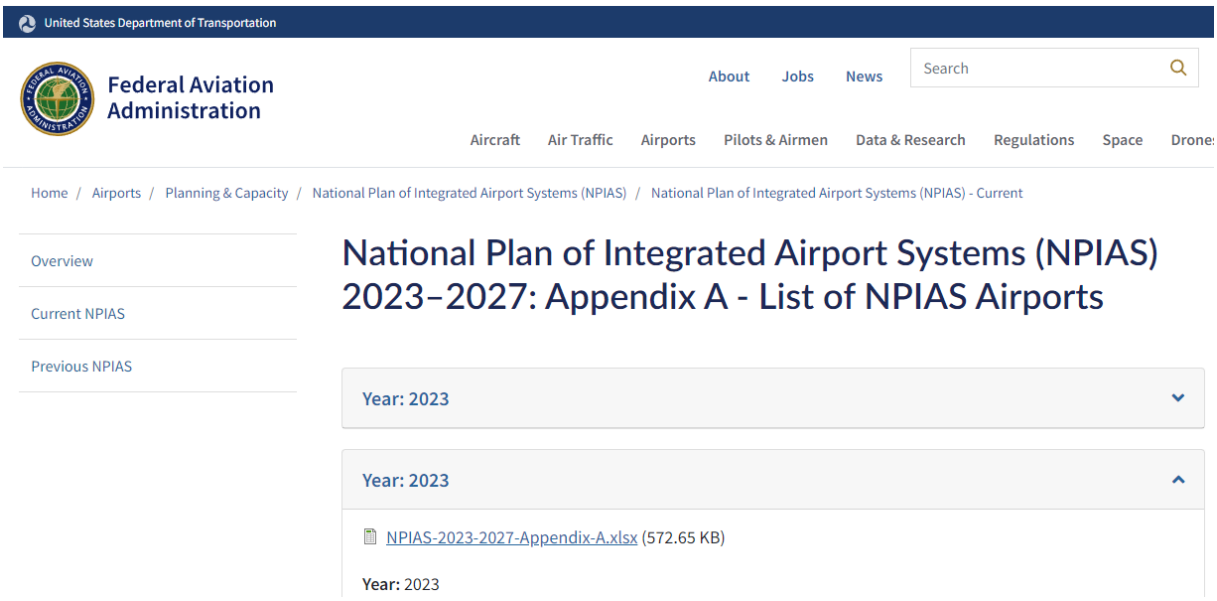
The csv file from this FAA Website has been downloaded and converted into a MS Excel file

https://adds-faa.opendata.arcgis.com/datasets/faa::airports-1/explore?location=0.158824%2C-1.633886%2C2.00

and the MS Excel file can be loaded into the database table.

X	Y	OBJECTID	GLOBAL_ID	IDENT	NAME
-176.6424981	51.88358162	1	656D38F0-F1FE-49A8-AB4F-677281616EF8	ADK	Adak
-154.1825759	56.9386921	2	F39AFCD2-D07F-4F41-96CF-08B79A271EAB	AKK	Akhiok
-161.4933513	60.91380952	3	C0EE48D3-E3AD-404E-945D-F404E345020D	Z13	Akiachak
-161.4350994	60.90786453	4	26D96486-FA29-4866-93EB-2EEEB7FA7144	KKI	Akiachak
-161.2306055	60.90289706	5	C7DF2896-135E-487B-9172-2FF613E7EAF1	AKI	Akiak
-165.7784016	54.13389989	6	81424AAF-E13A-4719-AD36-D26A582192AB	KQA	Akutan
-165.6041266	54.14461102	7	2441EA57-02E3-47E2-AC9C-1DC54444A884	7AK	Akutan
-164.7222175	62.68305467	8	978B7D5B-2C47-4B0F-BFBA-63975F26AD11	AUK	Alakanuk

3.2.5.3. Data source - FAA List of NPIAS Airports:



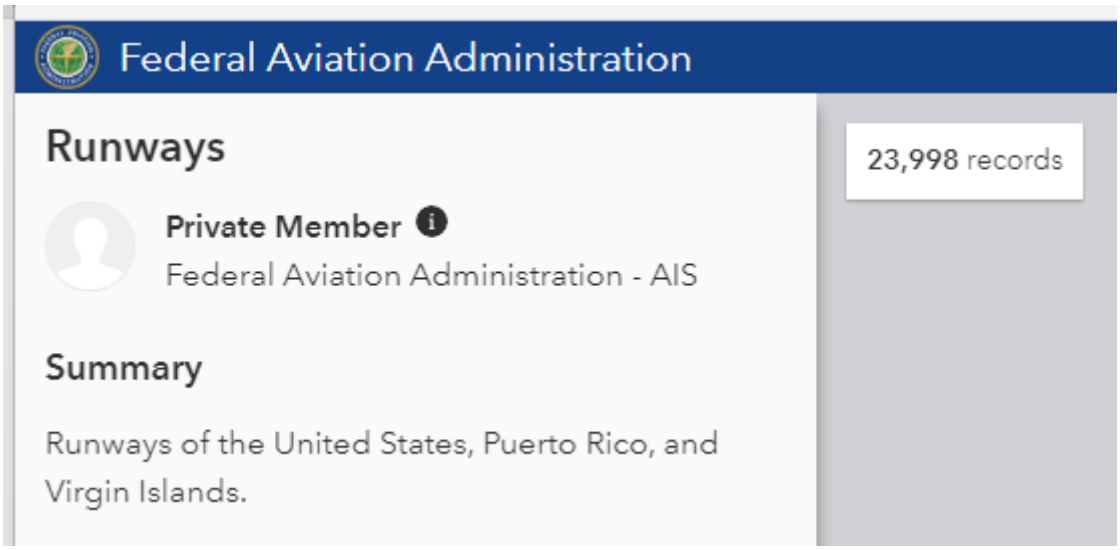
The MS Excel file from this FAA Website has been downloaded

https://www.faa.gov/sites/faa.gov/files/2022-09/NPIAS-2023-2027-Appendix-A.xlsx

and only the airports in US states listed here are considered for the database table **`io_airports`**.

	State	City	Airport	LocID
15	NE	Scottsbluff	Western Nebraska Regional/William B Heilig Field	BFF
16	NE	Scribner	Scribner State	SCB
17	NE	Seward	Seward Municipal	SWT
18	NE	Sidney	Sidney Municipal/Lloyd W Carr Field	SNY
19	NE	Superior	Superior Municipal	12K
50	NE	Tecumseh	Tecumseh Municipal	0G3
51	NE	Tekamah	Tekamah Municipal	TQE
52	NE	Thedford	Thomas County	TIF
53	NE	Valentine	Miller Field	VTN

3.2.5.4. Data source - FAA Runways:



The csv file from this FAA Website has been downloaded and converted into a MS Excel file

https://adds-faa.opendata.arcgis.com/datasets/faa::runways/explore?location=0.059024%2C-1.628764%2C2.00

and selected data from the MS Excel file will be loaded into the database table.

OBJECTID	GLOBAL_ID	AIRPORT_ID	DESIGNATOR	LENGTH	WIDTH	DIM_UOM
1	C747AD56-6CCF-4130-9255-643A4306CA51	656D38F0-F1FE-49A8-AB4F-677281616EF8	05/23	7790	200	FT
2	EA83C6B6-306D-4AD0-9BBC-3534B05CA636	F39AFCD2-D07F-4F41-96CF-08B79A271EAB	04/22	3120	50	FT
3	6D8AC274-5B38-4994-8559-05B048032A21	C0EE48D3-E3AD-404E-945D-F404E345020D	01/19	3300	60	FT
4	84E294FE-4120-472E-9186-7FE999B4BA7E	26D96486-FA29-4866-93EB-2EEEB7FA7144	E/W	5000	300	FT
5	135362D2-1BD6-45A1-86B6-0B7BA2314384	26D96486-FA29-4866-93EB-2EEEB7FA7144	NW/SE	5000	300	FT
6	0687A386-40EF-4AD0-A4A2-FEA0EC331264	C7DF2896-135E-487B-9172-2FF613E7EAF1	03/21	3200	76	FT
7	C66D674A-7288-401E-A195-DECEE910F5C4	81424AAF-E13A-4719-AD36-D26A582192AB	E/W	10000	300	FT
8	76B94CF6-3823-4E9F-B465-EAD81574CA53	2441EA57-02E3-47E2-AC9C-1DC54444A884	09/27	4500	75	FT
9	5514C208-64C5-4740-8506-A43151DF0742	978B7D5B-2C47-4B0F-BFBA-63975F26AD11	16/34	4000	75	FT
10	1E3CAEB1-5D3E-419A-9638-731AD9F7FB09	C3052BE5-3177-4343-9D5A-42B0046D57EE	11/29	1250	300	FT
11	0C3A88B2-5EEA-4CC2-AEF4-39BA897B85A3	C3052BE5-3177-4343-9D5A-42B0046D57EE	18/36	850	300	FT
12	7F724C9E-92A3-403F-9682-E775CB4F945A	EA0EB415-23A2-412F-9363-CA84BDB2F0E0	E/W	10000	300	FT

3.2.6. Table io_aviation_occurrence_categories

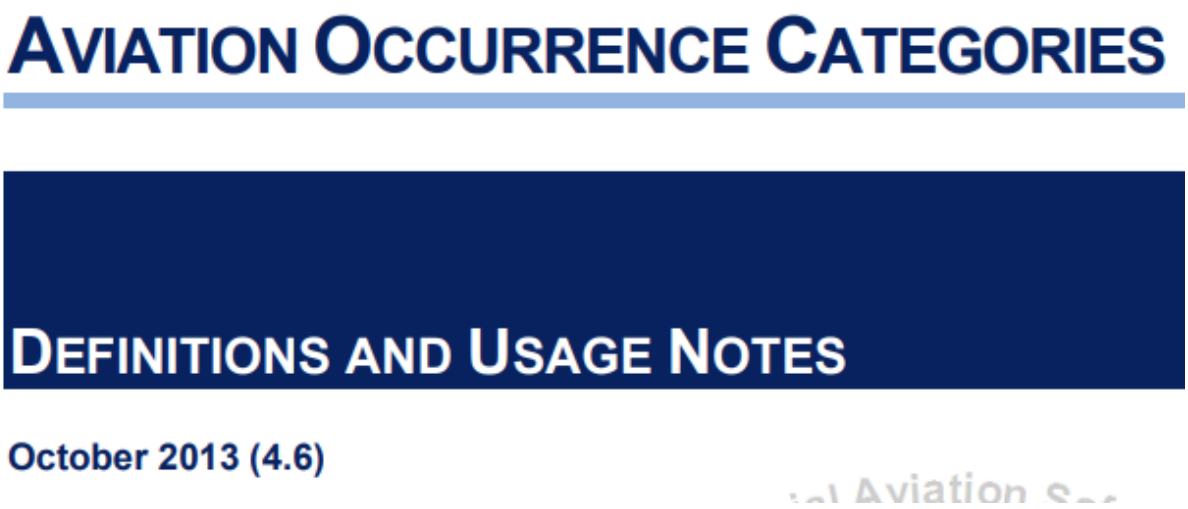
This table contains CICTT codes from the document Aviation Occurrence Categories (AVIATION OCCURRENCE CATEGORIES DEFINITIONS AND USAGE NOTES):

	Name	Type	Primary	Not Nul
1	cictt_code	varchar(10)	☑	☑
2	identifier	varchar(100)	☐	☑
3	definition	text	☐	☑
4	first_processed	timestamp without time zone	☐	☑
5	last_processed	timestamp without time zone	☐	☐
6	last_seen	timestamp without time zone	☐	☐

3.2.6.1. Sample data:

```
Name           |Value
---------------+------------------------------------------------------------
cictt_code     |LOC-I
identifier     |LOSS OF CONTROL–INFLIGHT
definition     |Loss of aircraft control while, or deviation from inten ...
first_processed|2023-01-18 11:07:38.756
last_processed |
last_seen      |2023-01-18 10:07:37.837
```

3.2.6.2. Data source:



The data from this document (AVIATION OCCURRENCE CATEGORIES DEFINITIONS AND USAGE NOTES, 2013) has been extracted into the MS Excel file

`Aviation_Occurrence_Categories/aviation_occurrence_categories.xlsx`

and can be loaded from there into the database table.

CICTT Code	Identifier	Definition
ADRM	AERODROME	Occurrences involving Aerodrome design, service, or functionality issues
AMAN	ABRUPT MANEUVER	The intentional abrupt maneuvering of the aircraft by the flight crew.
ARC	ABNORMAL RUNWAY CONTACT	Any landing or takeoff involving abnormal runway or landing surface contact.
ATM	ATM/CNS	Occurrences involving Air Traffic Management (ATM) or Communication, Navigation, Surveillance (CNS) service issues.
BIRD	BIRD	Occurrences involving collisions/near collisions with bird(s).
CABIN	CABIN SAFETY EVENTS	Miscellaneous occurrences in the passenger cabin of transport category aircraft.
CFIT	CONTROLLED FLIGHT INTO OR TOWARD TERRAIN	In-flight collision or near collision with terrain, water, or obstacle without indication of loss of control.
CTOL	COLLISION WITH OBSTACLE(S) DURING TAKEOFF AND LANDING	Collision with obstacle(s) during takeoff or landing while airborne.

3.2.7. Table io_countries

This table mainly contains data on latitudes and longitudes of countries. Currently, only the values of the USA are stored.

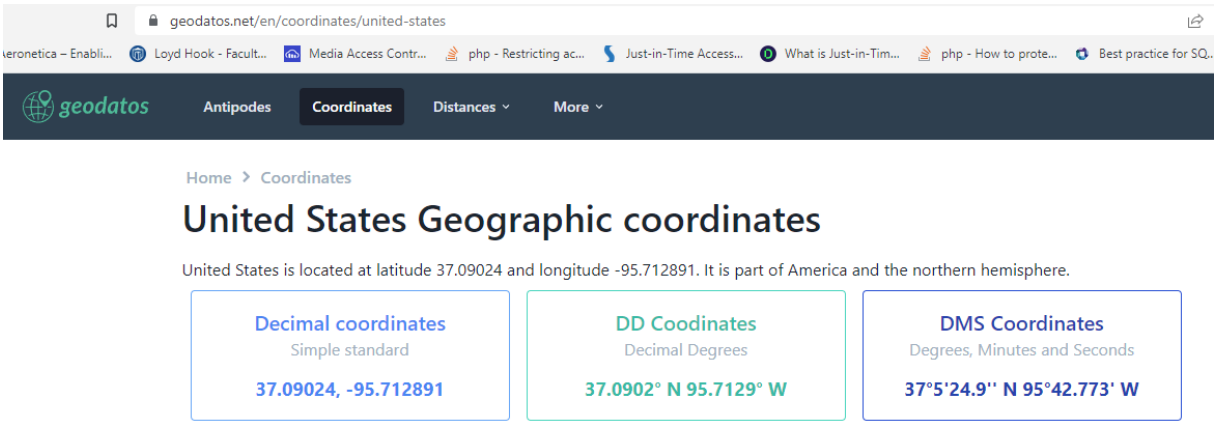
The latitude and longitude stored here will be applied to US events that do not themselves contain decimal latitude and longitude and which cannot be determined at the zip code, city or state level.

	Name	Type	Primary	Not Null
1	country	varchar(4)	☑	☑
2	country_name	varchar(100)	☐	☐
3	dec_latitude	double precision	☐	☑
4	dec_longitude	double precision	☐	☑
5	first_processed	timestamp withou	☐	☑
6	last_processed	timestamp withou	☐	☐

3.2.7.1. Sample data:

```
country|country_name |dec_latitude|dec_longitude|
-------+-------------+------------+-------------+
USA    |United States|    37.09024|   -95.712891|
```

3.2.7.2. Data source:



The data must be provided in a JSON file with the following format:

```
[
  {
    "type": "country",
    "country": "USA",
    "country_name": "United States",
    "dec_latitude": 37.09024,
    "dec_longitude": -95.712891
  },
  {
    "type": "state",
    "country": "USA",
    "state": "AK",
    "state_name": "Alaska",
    "dec_latitude": 63.7431630974,
    "dec_longitude": -151.594035116
  },
  {
    "type": "state",
    "country": "USA",
    "state": "AL",
    "state_name": "Alabama",
    "dec_latitude": 32.7570463396,
    "dec_longitude": -86.844525962
  },
...
]
```

3.2.8. Table io_lat_lng

This is the main table for determining missing decimal latitude or longitude for US events. The table is fed by the following sources:

- simplemaps United States Cities Database (simplemaps.com. (n.d.))
- simplemaps US Zip Codes Database (simplemaps.com. (n.d.))
- Zip Codes.org ZIP Code Database (UnitedStatesZipCodes. (n.d.))

In addition, the missing values for cities are averaged from the values of the matching zip codes.

	Name	Type	Primary	Not Null
1	city	varchar(50)	☐	☐
2	country	varchar(4)	☐	☐
3	dec_latitude	double precision	☐	☑
4	dec_longitude	double precision	☐	☑
5	first_processed	timestamp withou	☐	☑
6	id	serial	☑	☑
7	last_processed	timestamp withou	☐	☐
8	source	varchar(50)	☐	☑
9	state	varchar(2)	☐	☐
10	type	varchar(7)	☐	☑
11	zipcode	varchar(10)	☐	☐

3.2.8.1. Sample data:

```
Name           |Value                              |
---------------+-----------------------------------+
id             |13620                              |
type           |ZIPCODE                            |
country        |USA                                |
state          |SC                                 |
city           |ABBEVILLE                          |
zipcode        |29620                              |
dec_latitude   |34.1813                            |
dec_longitude  |-82.42804                          |
source         |simplemaps US Zip Codes Database|
first_processed|2022-11-29 13:11:16.347            |
last_processed |2022-11-29 13:28:49.520            |
```

3.2.8.2. Data source simplemaps

Data from the simplemaps flat files **`uscities.xlsx`** and **`uszips.xlsx`** is loaded into the PostgreSQL database using the following processing logic:

1. the existing data with column **`source`** equal to '**`simplemaps United States Cities Database'`** or '**`simplemaps US Zip Codes Database`**' are deleted,
2. the database United States Cities is processed, whereby for each zip code in the column **`zips`** an entry in the PostgreSQL table **`io_lat_lng`** is created (not updated) - as

latitude and longitude the corresponding values of the affected city are used, with column **type** equal to '**ZIPCODE**' and column **source** equal to '**simplemaps United States Cities Database**',

3. the database US Zip Codes is processed, whereby for each zip code an entry in the PostgreSQL table **io_lat_lng** is created or updated with column **type** equal to '**ZIPCODE**' and column **source** equal to '**simplemaps US Zip Codes Database**',
4. the database United States Cities is processed, whereby for each city an entry in the PostgreSQL table **io_lat_lng** is created or updated with column **type** equal to 'CITY' and column **source** equal to '**simplemaps United States Cities Database**',
5. the existing rows of the database table **io_lat_lng** with column **source** equal to '**average**' are deleted,
6. from the lines of the database table **io_lat_lng** of type '**ZIPCODE**' the average of latitude and longitude per city is calculated and this value is stored as latitude and longitude of this city in the database table **io_lat_lng** with **source** equal to '**average**'.

3.2.8.2.1. Resulting database table **io_lat_lng**:

id	type	country	state	city	zipcode	dec_latitude	dec_longitude	source
9692	ZIPCODE	USA	LA	ABBEVILLE	70510	29.89493	-92.19569	simplemaps US Zip Codes Database
31690	ZIPCODE	USA	MS	ABBEVILLE	38601	34.48146	-89.45166	simplemaps US Zip Codes Database
78821	CITY	USA	LA	ABBEVILLE		29.9751	-92.1265	simplemaps United States Cities Database
17136	ZIPCODE	USA	GA	ABBEVILLE	31001	31.97595	-83.33979	simplemaps US Zip Codes Database
17578	ZIPCODE	USA	AL	ABBEVILLE	36310	31.60296	-85.21229	simplemaps US Zip Codes Database
13620	ZIPCODE	USA	SC	ABBEVILLE	29620	34.1813	-82.42804	simplemaps US Zip Codes Database
84792	CITY	USA	GA	ABBEVILLE		31.9925	-83.3068	simplemaps United States Cities Database
81898	CITY	USA	SC	ABBEVILLE		34.1787	-82.3774	simplemaps United States Cities Database
85176	CITY	USA	AL	ABBEVILLE		31.5664	-85.2528	simplemaps United States Cities Database
97190	CITY	USA	MS	ABBEVILLE		34.5034	-89.5025	simplemaps United States Cities Database
224446	CITY	USA	ME	ABBOT		45.2186	-69.57816	average
42778	ZIPCODE	USA	ME	ABBOT	04406	45.2186	-69.57816	simplemaps US Zip Codes Database
19368	ZIPCODE	USA	WI	ABBOTSFORD	54405	44.96289	-90.2797	simplemaps US Zip Codes Database
19369	ZIPCODE	USA	WI	ABBOTSFORD	54421	44.9435	-90.3174	simplemaps United States Cities Database
86704	CITY	USA	WI	ABBOTSFORD		44.9435	-90.3174	simplemaps United States Cities Database
31885	ZIPCODE	USA	TX	ABBOTT	76621	31.87789	-97.08891	simplemaps US Zip Codes Database
97352	CITY	USA	TX	ABBOTT		31.8868	-97.0799	simplemaps United States Cities Database
224225	CITY	USA	IL	ABBOTT PARK		42.32	-87.86	average

3.2.8.2.2. Data source United States Cities Database

city_ascii	state_id	state_name	county_fips	county_name	lat	lng
New York	NY	New York	36081	Queens	40.6943	-73.9249
Los Angeles	CA	California	06037	Los Angeles	34.1141	-118.4068
Chicago	IL	Illinois	17031	Cook	41.8375	-87.6866
Miami	FL	Florida	12086	Miami-Dade	25.784	-80.2101
Dallas	TX	Texas	48113	Dallas	32.7935	-96.7667
Houston	TX	Texas	48201	Harris	29.786	-95.3885
Philadelphia	PA	Pennsylvania	42101	Philadelphia	40.0077	-75.1339
Atlanta	GA	Georgia	13121	Fulton	33.7628	-84.422
Washington	DC	District of Columbia	11001	District of Columbia	38.9047	-77.0163
Boston	MA	Massachusetts	25025	Suffolk	42.3188	-71.0852
Phoenix	AZ	Arizona	04013	Maricopa	33.5722	-112.0892
Detroit	MI	Michigan	26163	Wayne	42.3834	-83.1024
San Francisco	CA	California	06075	San Francisco	37.7558	-122.4449
Seattle	WA	Washington	53033	King	47.6211	-122.3244
San Diego	CA	California	06073	San Diego	32.8313	-117.1222
Minneapolis	MN	Minnesota	27053	Hennepin	44.9635	-93.2678

3.2.8.2.3. Data source US Zip Codes Database

zip	lat	lng
00601	18.18027	-66.75266
00602	18.36075	-67.17541
00603	18.45744	-67.12225
00606	18.16585	-66.93716
00610	18.2911	-67.12243
00611	18.27698	-66.80688
00612	18.41283	-66.7051
00616	18.41878	-66.6679
00617	18.44598	-66.56006
00622	17.98892	-67.1566
00623	18.08429	-67.15336
00624	18.05905	-66.71932

3.2.8.3. Data source Zip Codes.org

Data from the Zip Codes.org flat file **`zip_code_database.xls`** is loaded into the PostgreSQL database using the following processing logic:

1. the existing data with column source equal to 'Zip Codes.org ZIP Code Database' are deleted,
2. the Excel rows of **`type`** equals '**`STANDARD`**' and **`country`** equals '**`US`**' are inserted or updated in the database table **`io_lat_lng`** with column **`source`** equal to '**`Zip Codes.org ZIP Code Database`**',

3. the existing rows of the database table **`io_lat_lng`** with column **`source`** equal to '**`average`**' are deleted,
4. from the lines of the database table **`io_lat_lng`** of **`type`** '**`ZIPCODE`**' the average of latitude and longitude per city is calculated and this value is stored as latitude and longitude of this city in the database table **`io_lat_lng`** with **`source`** equal to '**`average`**'.

3.2.8.3.1. Resulting database table **`io_lat_lng`**:

id	type	country	state	city	zipcode	dec_latitude	dec_longitude	source
211079	CITY	USA	CA	29 PALMS		34.24	-115.86	average
183583	ZIPCODE	USA	CA	29 PALMS	92277	34.24	-115.86	Zip Codes.org ZIP Code Database
162642	ZIPCODE	USA	KY	AARON	42602	36.69	-85.13	Zip Codes.org ZIP Code Database
202361	CITY	USA	KY	AARON		36.69	-85.13	average
25811	ZIPCODE	USA	PA	AARONSBURG	16820	40.91995	-77.38702	simplemaps US Zip Codes Database
92194	CITY	USA	PA	AARONSBURG		40.9042	-77.4513	simplemaps United States Cities Database
207851	CITY	USA	GA	ABAC		31.46	-83.51	average
157371	ZIPCODE	USA	GA	ABAC	31794	31.46	-83.51	Zip Codes.org ZIP Code Database
100551	CITY	USA	AL	ABANDA		33.0926	-85.5253	simplemaps United States Cities Database
35631	ZIPCODE	USA	AL	ABANDA	36276	33.0926	-85.5253	simplemaps United States Cities Database
9692	ZIPCODE	USA	LA	ABBEVILLE	70510	29.89493	-92.19569	simplemaps US Zip Codes Database
31690	ZIPCODE	USA	MS	ABBEVILLE	38601	34.48146	-89.45166	simplemaps US Zip Codes Database

3.2.8.3.2. Data source ZIP Code Database

zip	type	decommissioned	primary_city	state	county	country	latitude	longitude
00501	UNIQUE	0	Holtsville	NY	Suffolk County	US	40.81	-73.04
00544	UNIQUE	0	Holtsville	NY	Suffolk County	US	40.81	-73.04
00601	STANDARD	0	Adjuntas	PR	Adjuntas Municipio	US	18.16	-66.72
00602	STANDARD	0	Aguada	PR	Aguada Municipio	US	18.38	-67.18
00603	STANDARD	0	Aguadilla	PR	Aguadilla Municipio	US	18.43	-67.15
00604	PO BOX	0	Aguadilla	PR		US	18.43	-67.15
00605	PO BOX	0	Aguadilla	PR		US	18.43	-67.15
00606	STANDARD	0	Maricao	PR	Maricao Municipio	US	18.18	-66.98
00610	STANDARD	0	Anasco	PR	Anasco Municipio	US	18.28	-67.14
00611	PO BOX	0	Angeles	PR		US	18.28	-66.79
00612	STANDARD	0	Arecibo	PR	Arecibo Municipio	US	18.45	-66.73
00613	PO BOX	0	Arecibo	PR		US	18.45	-66.73
00614	PO BOX	0	Arecibo	PR		US	18.45	-66.73
00616	STANDARD	0	Bajadero	PR	Arecibo Municipio	US	18.42	-66.67
00617	STANDARD	0	Barceloneta	PR	Barceloneta Municipio	US	18.45	-66.56
00622	STANDARD	0	Boqueron	PR	Cabo Rojo Municipio	US	17.99	-67.15
00623	STANDARD	0	Cabo Rojo	PR	Cabo Rojo Municipio	US	18.08	-67.14
00624	STANDARD	0	Penuelas	PR	Penuelas Municipio	US	18.06	-66.72

3.2.9. Table io_md_codes_category

This table contains the category codes extracted from the **`findings`** table.

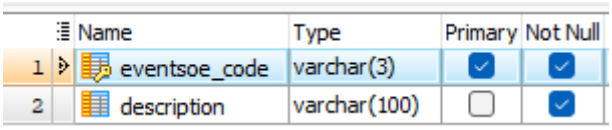
	Name	Type	Primary	Not Null
1	category_code	varchar(2)	☑	☑
2	description	varchar(255)	☐	☑

3.2.9.1. Sample data:

```
Name         |Value   |
-------------+--------+
category_code|01      |
description  |Aircraft|
```

3.2.10. Table io_md_codes_eventsoe

This table contains the eventsoe codes extracted from the **events_sequence** table.



	Name	Type	Primary	Not Null
1	eventsoe_code	varchar(3)	☑	☑
2	description	varchar(100)	☐	☑

3.2.10.1. Sample data:

```
Name          |Value                          |
--------------+-------------------------------+
eventsoe_code|490                             |
description   |Collision during takeoff/land|
```

3.2.11. Table io_md_codes_modifier

This table contains the modifier codes extracted from the **`findings`** table.

	Name	Type	Primary	Not Null
1	modifier_code	varchar(3)	☑	☑
2	description	varchar(100)	☐	☑

3.2.11.1. Sample data:

```
Name         |Value  |
-------------+-------+
modifier_code|01     |
description  |Failure|
```

3.2.12. Table io_md_codes_phase

This table contains the phase codes extracted from the **events_sequence** table. From a suitable MS Excel file the column description_main_phase is added.

	Name	Type	Primary	Not Null
1	phase_code	varchar(3)	☑	☑
2	description	varchar(100)	☐	☑
3	description_main_phase	varchar(100)	☐	☐

3.2.12.1. Sample data:

```
Name                  |Value           |
----------------------+----------------+
phase_code            |453             |
description           |Maneuvering-hover|
description_main_phase|Manuevering     |
```

3.2.12.2. Data source:

phase_code	description	main_phase
100	Prior to flight	Pre-taxi
150	Standing	Pre-taxi
151	Standing-engine(s) not oper	Pre-taxi
152	Standing-engine(s) start-up	Pre-taxi
153	Standing-engine(s) operating	Pre-taxi
154	Standing-engine(s) shutdown	Pre-taxi
200	Pushback/towing	Pre-taxi
201	Pushback/tow-engine not oper	Pre-taxi
202	Pushback/tow-engine start-up	Pre-taxi
203	Pushback/tow-engine oper	Pre-taxi
250	Taxi	Taxi
251	Taxi-to runway	Taxi
252	Taxi-into takeoff position	Taxi

3.2.13. Table io_md_codes_section

This table contains the section codes extracted from the **`findings`** table.

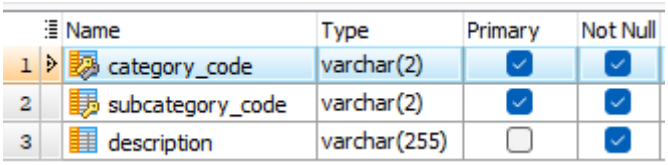
	Name	Type	Primary	Not Null
1	category_code	varchar(2)	☑	☑
2	subcategory_code	varchar(2)	☑	☑
3	section_code	varchar(2)	☑	☑
4	description	varchar(255)	☐	☑

3.2.13.1. Sample data:

```
Name            |Value    |
----------------+---------+
category_code   |01       |
subcategory_code|01       |
section_code    |00       |
description     |(general)|
```

3.2.14. Table io_md_codes_subcategory

This table contains the subcategory codes extracted from the **`findings`** table.



	Name	Type	Primary	Not Null
1	category_code	varchar(2)	☑	☑
2	subcategory_code	varchar(2)	☑	☑
3	description	varchar(255)	☐	☑

3.2.14.1. Sample data:

```
Name            |Value        |
----------------+-------------+
category_code   |02           |
subcategory_code|02           |
description     |Psychological|
```

3.2.15. Table io_md_codes_subsection

This table contains the subsection codes extracted from the **`findings`** table.

	Name	Type	Primary	Not Null
1	category_code	varchar(2)	☑	☑
2	subcategory_code	varchar(2)	☑	☑
3	section_code	varchar(2)	☑	☑
4	subsection_code	varchar(2)	☑	☑
5	description	varchar(255)	☐	☑

3.2.15.1. Sample data:

```
Name            |Value                       |
----------------+----------------------------+
category_code   |01                          |
subcategory_code|02                          |
section_code    |37                          |
subsection_code |10                          |
description     |Vacuum distribution system|
```

3.2.16. Table io_processed_files

This table documents which external files were processed to create and maintain the database content.

	Name	Type	Primary	Not Null
1	file_name	varchar(100)	☑	☑
2	first_processed	timestamp without time zone	☐	☑
3	last_processed	timestamp without time zone	☐	☐
4	counter	integer	☐	☑

3.2.16.1. Sample data:

```
Name           |Value                  |
---------------+-----------------------+
file_name      |up22NOV                |
first_processed|2022-11-27 13:40:18.094|
last_processed |2022-12-12 12:08:15.915|
counter        |2                      |
```

3.2.17. Table io_sequence_of_events

This table enables the assignment of CICTT codes via the event sequence number.

	Name	Type	Primary	Not Null
1	soe_no	integer	☑	☑
2	cictt_code	varchar(10)	☐	☐
3	meaning	varchar(100)	☐	☑
4	first_processed	timestamp without time zone	☐	☑
5	last_processed	timestamp without time zone	☐	☐
6	last_seen	timestamp without time zone	☐	☐

3.2.17.1. Sample data:

```
soe_no|cictt_code|meaning
------+----------+----------------------------------------
     0|UNK       |Unknown or undetermined
    10|RAMP      |Aircraft loading event
    20|RAMP      |Aircraft servicing event
    30|          |Preflight or dispatch event
    40|          |Aircraft maintenance event
    50|          |Aircraft inspection event
    60|          |Attempted remediation/recovery
    70|ADRM      |Airport occurrence
    80|RAMP      |Ground handling event
    81|RAMP      |AC/prop/rotor contact w person
    82|RAMP      |Prop/jet/rotor blast/suction
```

3.2.17.2. Data source CICTT_SOE_MAP.csv (file given by NTSB):

eventsoe_no	meaning	CICTT_Code	CICTT_Description
0	Unknown or undetermined	UNK	Unknown
10	Aircraft loading event	RAMP	Ground Handling
20	Aircraft servicing event	RAMP	Ground Handling
30	Preflight or dispatch event		
40	Aircraft maintenance event		
50	Aircraft inspection event		
60	Attempted remediation/recovery		
70	Airport occurrence	ADRM	Aerodrome
80	Ground handling event	RAMP	Ground Handling
81	AC/prop/rotor contact w person	RAMP	Ground Handling
82	Prop/jet/rotor blast/suction	RAMP	Ground Handling
90	Abnormal runway contact	ARC	Abnormal Runway Contact
91	Tailstrike	ARC	Abnormal Runway Contact
92	Hard landing	ARC	Abnormal Runway Contact
93	Dragged wing/rotor/float/other	ARC	Abnormal Runway Contact
94	Landing gear collapse	ARC	Abnormal Runway Contact

3.2.18. Table io_states

This table mainly contains data on latitudes and longitudes of US states.

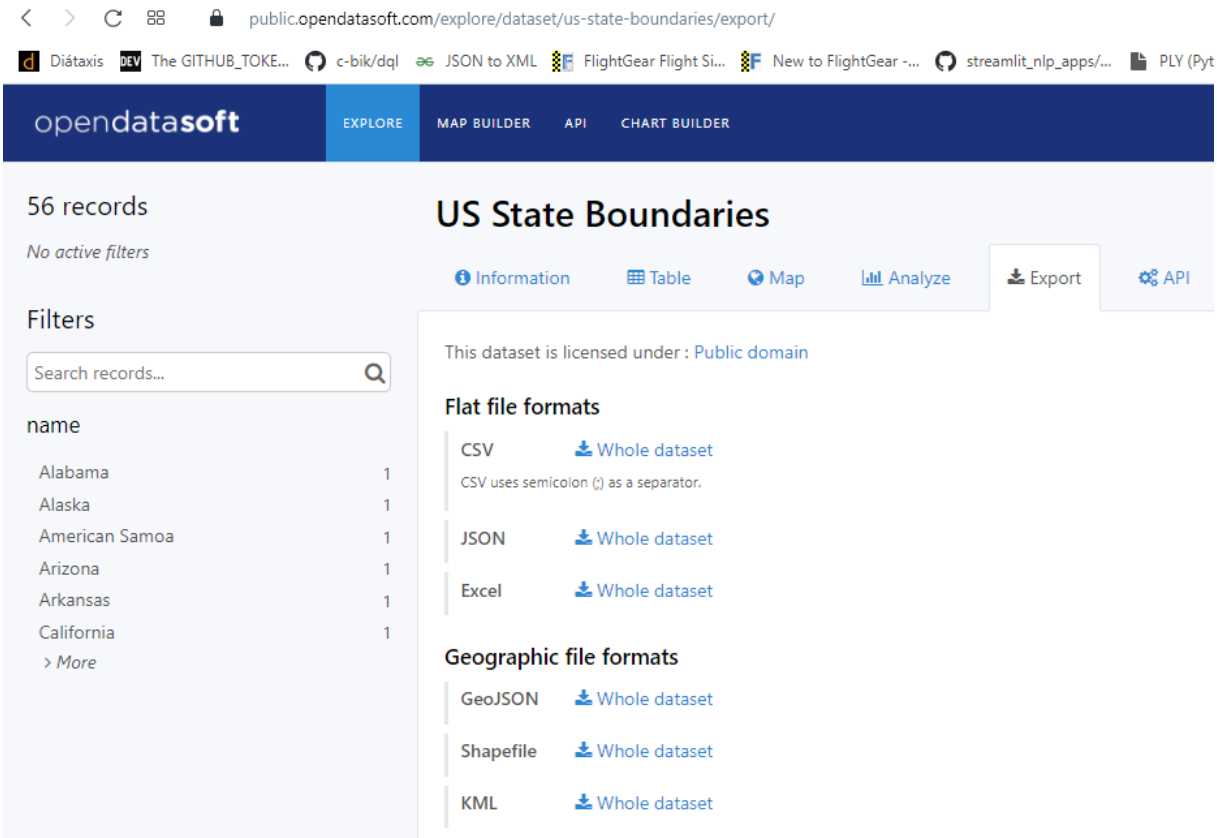
The latitude and longitude stored here will be applied to US events that do not themselves contain decimal latitude and longitude and which cannot be determined at the zip code or city level.

	Name	Type	Primary	Not Null
1	country	varchar(4)	☑	☑
2	dec_latitude	double precision	☐	☑
3	dec_longitude	double precision	☐	☑
4	first_processed	timestamp withou	☐	☑
5	last_processed	timestamp withou	☐	☐
6	state	varchar(2)	☑	☑
7	state_name	varchar(100)	☐	☐

3.2.18.1. Sample data:

```
country|state|state_name           |dec_latitude |dec_longitude |
-------+-----+---------------------+-------------+--------------+
USA    |AK   |Alaska               |63.7431630974|-151.594035116|
USA    |AL   |Alabama              |32.7570463396| -86.844525962|
```

3.2.18.2. Data source:



Same JSON file as with table `io_countries`.

3.2.19. View io_lat_lng_issues

This view shows the related data from the tables **`events`**, **`io_countries`**, **`io_lat_lng`** and **`io_states`** for those events in the USA where either the decimal latitude or longitude is missing from the NTSB data.

1	ev_id	varchar(14)
2	ev_country	varchar(4)
3	ev_state	varchar(2)
4	ev_city	varchar(50)
5	ev_site_zipcode	varchar(10)
6	ev_dec_latitude	double precision
7	ev_dec_longitude	double precision
8	ev_latitude	varchar(7)
9	ev_longitude	varchar(8)
10	io_country	varchar(4)
11	io_state	varchar(2)
12	io_city	varchar(50)
13	io_site_zipcode	varchar(10)
14	io_dec_latitude	double precision
15	io_dec_longitude	double precision
16	io_latitude	varchar(7)
17	io_longitude	varchar(8)
18	io_dec_lat_lng_actions	text
19	io_dec_latitude_deviating	double precision
20	io_dec_longitude_deviating	double precision
21	io_invalid_latitude	boolean
22	io_invalid_longitude	boolean
23	io_invalid_us_city	boolean
24	io_invalid_us_city_zipcode	boolean
25	io_invalid_us_state	boolean
26	io_invalid_us_zipcode	boolean
27	country	varchar(4)
28	state	varchar(2)
29	city	varchar(50)
30	site_zipcode	varchar(10)
31	latitude	varchar(7)
32	longitude	varchar(8)
33	state_name	varchar(100)
34	zipcode_dec_latitude	double precision
35	zipcode_dec_longitude	double precision
36	city_dec_latitude	double precision
37	city_dec_longitude	double precision
38	state_dec_latitude	double precision
39	state_dec_longitude	double precision
40	country_dec_latitude	double precision
41	country_dec_longitude	double precision

3.2.19.1. Sample data:

Name	Value
ev_id	20030618X00903
ev_country	USA
ev_state	AK
ev_city	Anchorage
ev_site_zipcode	99515
ev_dec_latitude	
ev_dec_longitude	
ev_latitude	610624N
ev_longitude	1495152W
io_country	
io_state	
io_city	
io_site_zipcode	
io_dec_latitude	61.10666666666667
io_dec_longitude	-149.86444444444444
io_latitude	
io_longitude	
io_dec_lat_lng_actions	INFO.00.037 Correction based on latitude and l…
io_dec_latitude_deviating	
io_dec_longitude_deviating	
io_invalid_latitude	
io_invalid_longitude	
io_invalid_us_city	
io_invalid_us_city_zipcode	
io_invalid_us_state	
io_invalid_us_zipcode	
country	USA
state	AK
city	Anchorage
site_zipcode	99515
latitude	610624N
longitude	1495152W
state_name	Alaska
zipcode_dec_latitude	61.11733
zipcode_dec_longitude	-149.88894
city_dec_latitude	61.1508
city_dec_longitude	-149.1091
state_dec_latitude	63.7431630974
state_dec_longitude	-151.594035116
country_dec_latitude	37.09024
country_dec_longitude	-95.712891

4. Data Cleansing

4.1. Decimal latitude and longitude

The values in the following columns will be filled when executing task **c_l_l** (Correct decimal US latitudes and longitudes):

- **io_dec_latitude**
- **io_dec_longitude**
- **io_dec_lat_lng_actions**

The columns **dec_latitude** and **dec_longitude** were added by NTSB in 2008. For data from previous years, the two columns remained empty (NULL). **IO-Aero** pursues the following two goals with the **c_l_l** task:

- to check the validity of the existing values in the columns **latitude** and **longitude** and
- to find the missing values in the columns **dec_latitude** and **dec_longitude** for rows with **ev_country** equal to **'USA'**.

The underlying algorithm logs its processing steps in the **io_dec_lat_lng_actions** column - see the following examples:

```
io_dec_lat_lng_actions                 |
---------------------------------------------------------------------------+
ERROR.00.916 Unknown US state and city & INFO.00.035 Correction based on US state|
ERROR.00.922 Invalid US state id & INFO.00.033 Correction based on US zip code   |
---------------------------------------------------------------------------+
INFO.00.033 Correction based on US zip code                                 |
INFO.00.034 Correction based on US state and city                           |
INFO.00.037 Correction based on latitude and longitude                      |
```

4.2. Trimmed Database Columns

The following database columns are trimmed because they contain different whitespace elements:

Table	Column
aircraft	acft_category
aircraft	dest_country
aircraft	dprt_country
aircraft	far_part
aircraft	oper_country
aircraft	owner_country
aircraft	regis_no
events	ev_city
events	ev_site_zipcode
events	latitude
events	longitude
events_sequence	occurrence_code

5. References

adds-faa.opendata.arcgis.com. (n.d.). Airports. [online] Available at: https://adds-faa.opendata.arcgis.com/datasets/faa::airports-1/explore?location=0.158824%2C-1.633886%2C2.00 [Accessed 14 Apr. 2023].

adds-faa.opendata.arcgis.com. (n.d.). Runways. [online] Available at: https://adds-faa.opendata.arcgis.com/datasets/faa::runways/explore?location=0.059024%2C-1.628764%2C2.00 [Accessed 19 Apr. 2023].

AVIATION OCCURRENCE CATEGORIES DEFINITIONS AND USAGE NOTES. (2013). Available at: https://www.ntsb.gov/safety/data/Documents/datafiles/OccurrenceCategoryDefinitions.pdf [Accessed 28 Mar. 2023].

data.ntsb.gov. (n.d.). MDB Download Directory - NTSB.ADMS.DataTransfer.Web. [online] Available at: https://data.ntsb.gov/avdata.

Docker Documentation. (2022). Get Docker. [online] Available at: https://docs.docker.com/get-docker/.

Faa.gov. (2015). National Plan of Integrated Airport Systems (NPIAS) | Federal Aviation Administration. [online] Available at: https://www.faa.gov/airports/planning_capacity/npias.

Geodatos.net. (2019). GeoDatos - Information and maps of the world. [online] Available at: https://www.geodatos.net/en.

hub.docker.com. (n.d.). Docker Hub. [online] Available at: https://hub.docker.com/_/postgres.

members.io-aero.com. (n.d.). Streamlit. [online] Available at: http://members.io-aero.com/ [Accessed 28 Mar. 2023].

Opendatasoft. (n.d.). Create the best data experiences. [online] Available at: https://www.opendatasoft.com/?hsLang=en [Accessed 28 Mar. 2023].

The PostgreSQL Global Development Group (2019). PostgreSQL: The world's most advanced open source database. [online] Postgresql.org. Available at: https://www.postgresql.org/.

razorsql.com. (n.d.). RazorSQL - SQL Query Tool and SQL Editor for Mac, Windows, and Linux. [online] Available at: https://razorsql.com/.

simplemaps.com. (n.d.). Interactive HTML5 and JavaScript Maps for Websites | Simplemaps.com. [online] Available at: https://simplemaps.com [Accessed 28 Mar. 2023].

UnitedStatesZipCodes. (n.d.). U.S. ZIP Codes: Free ZIP code map and zip code lookup. [online] Available at: https://www.unitedstateszipcodes.org.

www.ntsb.gov. (n.d.). U.S. Civil Aviation Fatalities and Flight Activity Decreased in 2020. [online] Available at: https://www.ntsb.gov/news/press-releases/Pages/NR20211117.aspx.