



CSCI 5408 – Data Management, Warehousing, Analytics

Assignment 1

Work done by,

Name: Guturu Rama Mohan Vishnu

Banner ID: B00871849

Email: rm286720@dal.ca

DECLARATION

I, Guturu Rama Mohan Vishnu, declare that in assignment 1 of CSCI 5408 course, data scrapping is not done programmatically or using any online or offline tools. However, the webpages or the domain mentioned in this document are visited manually, and some useful information is gathered for education purpose only. Information, such as email, personal contact numbers, or names of people are not extracted. The course instructor or the Faculty of Computer Science cannot be held responsible for any misuse of the extracted data.

Problem #3: Ocean Tracking Data

Solution:

Due to changing climatic circumstances, marine conditions have deteriorated dramatically over time. As a result, one of the most crucial topics in the near future will be marine life management. This tracking technique was created as a result of the foresight to grasp undersea layout. The goal of the research is to surgically implant tags on marine animals so that acoustic receivers can track their movements. It also collects data on the species. The tags can also aid in the knowledge of ocean features such as underwater temperature, water salinity, and water currents.

The datasets and attributes I discovered are:

(a) otnunit_aat_animals_8dc3_4d15_c278

This dataset contains information regarding animal projects, data centers, animals, their scientific and vernacular names, length, weight, age and sex of these animals.

(b) otnunit_aat_datacenter_attributes_8a94_cefd_f8a3

This dataset contains all the detailed information about data centers like their name, abstract, keywords, reference, license, minimum and maximum values of geospatial longitude and latitude.

(c) otnunit_aat_detections_9062_5923_1394

Contains information about detection project reference, detection id, latitude and longitude, detection reference id, type and also about the uncertainties in latitude and longitude.

(d) otnunit_aat_manmade_platform_0735_7c9f_329c

It contains information regarding the platform such as reference, id, guide, type, depth, name, latitude and longitude.

(e) otnunit_aat_project_attributes_f29c_fb21_23a3

This file contains data about the project attributes such as reference, name of the project, abstract, citation of the project, geospatial minimum and maximum values of latitude and longitudes, etc.

(f) otnunit_aat_receivers_c595_05f4_68b2

It has the data of receivers like it's manufacture, model, it's coding scheme, serial number, latitude, longitude, time, recovery date and time of data in UTC, bottom depth and etc.

(g) otnunit_aat_recover_offload_details_4b23_f002_f89a

Includes information about project recovery such as recovered location, recovery comments, recovered by, recovery outcome, data offload time, etc.

(h) otnunit_aat_tag_releases_b793_03e7_a230

Includes information about tag releases such as tag device id, release reference id, latitude, longitude, time, manufacturer, tag model, tag serial number, frequency etc.

Cleaning and Transformation:

otnunit_aat_animals_8dc3_4d15_c278

The Primary Key for this table is “animal_reference_id”.

- (1) As we can see in this dataset, the column named “taxonrank” is blank. So, we can afford to delete the whole column.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	animal_project_ref	datacenter_ref	animal_reference	animal_guid	vernacularname	scientificname	taxonran	aphaid	tsn	animal_o	stock	length	length_ty	weight	life_stag
1	reference	reference	_id				k			rigin			pe		age
2															
3	FRO	OTN-Global	FRO-T14	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.561	FORK	2.35	NaN
4	FRO	OTN-Global	FRO-T2	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.576	FORK	2.53	NaN
5	FRO	OTN-Global	FRO-T25	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.57	FORK	2.41	NaN
6	FRO	OTN-Global	FRO-T26	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.6	FORK	2.8	NaN
7	FRO	OTN-Global	FRO-T27	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.55	FORK	2.17	NaN
8	FRO	OTN-Global	FRO-T28	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.49	FORK	1.55	NaN
9	FRO	OTN-Global	FRO-T29	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.45	FORK	1.24	NaN
10	FRO	OTN-Global	FRO-T5	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.521	FORK	1.89	NaN
11	FRO	OTN-Global	FRO-T1	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.556	FORK	2.29	NaN
12	FRO	OTN-Global	FRO-T41	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.54	FORK	2.15	NaN
13	FRO	OTN-Global	FRO-T8	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.588	FORK	2.69	NaN
14	FRO	OTN-Global	FRO-T7	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.64	FORK	3.44	NaN
15	FRO	OTN-Global	FRO-T4	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.54	FORK	2.1	NaN
16	FRO	OTN-Global	FRO-T6	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.612	FORK	3.02	NaN
17	FRO	OTN-Global	FRO-T9	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.6	FORK	2.85	NaN
18	FRO	OTN-Global	FRO-T10	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.544	FORK	2.15	NaN
19	FRO	OTN-Global	FRO-T13	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.586	FORK	2.66	NaN
20	FRO	OTN-Global	FRO-T11	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.596	FORK	2.8	NaN
21	FRO	OTN-Global	FRO-T12	OTN-GlobalFROFRC Arctic char	Salvelinus alpinus			127188	162001	W	UNK	0.54	FORK	2.1	NaN

- (2) If there are empty cells present in our file, then we replace them with NaN to maintain the uniformity between all the records.

- (3) Age is a column where there are values of NaN, but I think it is more reasonable to have the default value as 0 for age column.

otnunit_aat_animals_8dc3_4d15_c278 [Compatibility Mode] - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

Calibri 11 A A Wrap Text General Conditional Formatting Format as Table Cell Styles Insert Delete Format AutoSum Fill Sort & Find & Filter Select

O1 age

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	datacenter_ref	animal_reference_id	animal_guid	vernacularname	scientificname	aphiaid	tsn	animal_o_rigin	stock	length	length_type	weight	life_stage	age	sex	
1																
2																
3	OTN-Global	FRO-T14	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.561	FORK	2.35		0	U	
4	OTN-Global	FRO-T2	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.576	FORK	2.53		0	U	
5	OTN-Global	FRO-T25	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.57	FORK	2.41		0	U	
6	OTN-Global	FRO-T26	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.8		0	U	
7	OTN-Global	FRO-T27	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.55	FORK	2.17		0	U	
8	OTN-Global	FRO-T28	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.49	FORK	1.55		0	U	
9	OTN-Global	FRO-T29	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.45	FORK	1.24		0	U	
10	OTN-Global	FRO-T5	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.521	FORK	1.89		0	U	
11	OTN-Global	FRO-T1	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.556	FORK	2.29		0	U	
12	OTN-Global	FRO-T41	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.15		0	U	
13	OTN-Global	FRO-T8	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.588	FORK	2.69		0	U	
14	OTN-Global	FRO-T7	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.64	FORK	3.44		0	U	
15	OTN-Global	FRO-T4	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1		0	U	
16	OTN-Global	FRO-T6	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.612	FORK	3.02		0	U	
17	OTN-Global	FRO-T9	TN-GlobalFROFRO-1	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.85		0	U	
18	OTN-Global	FRO-T10	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.544	FORK	2.15		0	U	
19	OTN-Global	FRO-T13	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.586	FORK	2.66		0	U	
20	OTN-Global	FRO-T11	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.596	FORK	2.8		0	U	
21	OTN-Global	FRO-T12	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1		0	U	
22	OTN-Global	FRO-T33	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.53	FORK	1.97		0	U	

otnunit_aat_animals_8dc3_4d15_c

Ready Average: 0.210028879 Count: 3810 Sum: 800 100%

Type here to search

-5°C 8:08 PM 1/24/2022

(4) We can do the same replacement for length and weight columns also as it is more reasonable to have 0 for these values instead of NaN.

otnunit_aat_animals_8dc3_4d15_c278 [Compatibility Mode] - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

Calibri 11 A A Wrap Text General Conditional Formatting Format as Table Cell Styles Insert Delete Format AutoSum Fill Sort & Find & Filter Select

M1 weight

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
357	OTN-Global	WRS-10562	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.211	FORK	0		0		
358	OTN-Global	WRS-10563	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.202	FORK	0		0		
359	OTN-Global	WRS-10564	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.202	FORK	0		0		
360	OTN-Global	WRS-10565	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.204	FORK	0		0		
361	OTN-Global	WRS-10566	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.201	FORK	0		0		
362	OTN-Global	WRS-10567	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.202	FORK	0		0		
363	OTN-Global	WRS-10568	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.197	FORK	0		0		
364	OTN-Global	WRS-10569	I-GlobalWRSWRS-1C	Atlantic salmon	Salmo salar	127186	161996	W	WRSH	0.184	FORK	0		0		
365	OTN-Global	GEFT-23	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
366	OTN-Global	GEFT-21	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
367	OTN-Global	GEFT-20	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
368	OTN-Global	GEFT-19	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
369	OTN-Global	GEFT-18	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
370	OTN-Global	GEFT-17	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
371	OTN-Global	GEFT-25	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
372	OTN-Global	GEFT-24	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
373	OTN-Global	GEFT-22	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
374	OTN-Global	GEFT-26	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
375	OTN-Global	GEFT-29	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
376	OTN-Global	GEFT-28	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
377	OTN-Global	GEFT-30	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
378	OTN-Global	GEFT-16	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	
379	OTN-Global	GEFT-15	TN-GlobalGEFTGEFT-	leervis	Lichia amia	126810	168769	W		0		0		0	U	

otnunit_aat_animals_8dc3_4d15_c

Ready Average: 0.23679828 Count: 7620 Sum: 1803.9293 100%

Type here to search

-5°C 8:10 PM 1/24/2022

(5) Sex column has the values U, M and F. So for cells which are empty in that column, we will replace them with U.

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	datacenter_refere	animal_referenc	animal_guid	vernacularname	scientificname	aphiaid	tsn	animal_o	stock	length	length_ty	weight	life_stag	age	sex	
1	OTN-Global	FRO-T14	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.561	FORK	2.35		0	U	
2	OTN-Global	FRO-T2	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.576	FORK	2.53		0	U	
3	OTN-Global	FRO-T25	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.57	FORK	2.41		0	U	
4	OTN-Global	FRO-T26	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.8		0	U	
5	OTN-Global	FRO-T27	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.55	FORK	2.17		0	U	
6	OTN-Global	FRO-T28	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.49	FORK	1.55		0	U	
7	OTN-Global	FRO-T29	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.45	FORK	1.24		0	U	
8	OTN-Global	FRO-T5	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.521	FORK	1.89		0	U	
9	OTN-Global	FRO-T1	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.556	FORK	2.29		0	U	
10	OTN-Global	FRO-T41	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.15		0	U	
11	OTN-Global	FRO-T8	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.588	FORK	2.69		0	U	
12	OTN-Global	FRO-T7	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.64	FORK	3.44		0	U	
13	OTN-Global	FRO-T4	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1		0	U	
14	OTN-Global	FRO-T6	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.612	FORK	3.02		0	U	
15	OTN-Global	FRO-T9	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.85		0	U	
16	OTN-Global	FRO-T10	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.544	FORK	2.15		0	U	
17	OTN-Global	FRO-T13	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.586	FORK	2.66		0	U	
18	OTN-Global	FRO-T11	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.596	FORK	2.8		0	U	
19	OTN-Global	FRO-T12	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1		0	U	
20	OTN-Global	FRO-T33	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.53	FORK	1.97		0	U	

(6) For the columns “stock” and “life_stage”, there are values already with UNKNOWN and all empty values in these two columns can be changed to UNKNOWN.

otnunit_aat_animals_8dc3_4d15_c278 [Compatibility Mode] - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

life_stage

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	animal_referenc e_id	animal_guid	vernacularname	scientificname	aphiaid	tsn	animal_o rigin	stock	length	length_ty pe	weight	life_stage	age	sex		
1																
2																
3	FRO-T14	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.561	FORK	2.35	UNKNOWN	0	U		
4	FRO-T2	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.576	FORK	2.53	UNKNOWN	0	U		
5	FRO-T25	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.57	FORK	2.41	UNKNOWN	0	U		
6	FRO-T26	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.8	UNKNOWN	0	U		
7	FRO-T27	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.55	FORK	2.17	UNKNOWN	0	U		
8	FRO-T28	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.49	FORK	1.55	UNKNOWN	0	U		
9	FRO-T29	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.45	FORK	1.24	UNKNOWN	0	U		
10	FRO-T5	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.521	FORK	1.89	UNKNOWN	0	U		
11	FRO-T1	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.556	FORK	2.29	UNKNOWN	0	U		
12	FRO-T41	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.15	UNKNOWN	0	U		
13	FRO-T8	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.588	FORK	2.69	UNKNOWN	0	U		
14	FRO-T7	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.64	FORK	3.44	UNKNOWN	0	U		
15	FRO-T4	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1	UNKNOWN	0	U		
16	FRO-T6	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.612	FORK	3.02	UNKNOWN	0	U		
17	FRO-T9	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.6	FORK	2.85	UNKNOWN	0	U		
18	FRO-T10	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.544	FORK	2.15	UNKNOWN	0	U		
19	FRO-T13	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.586	FORK	2.66	UNKNOWN	0	U		
20	FRO-T11	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.596	FORK	2.8	UNKNOWN	0	U		
21	FRO-T12	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.54	FORK	2.1	UNKNOWN	0	U		
22	FRO-T33	TN-GlobalFROFRO-T	Arctic char	Salvelinus alpinus	127188	162001	W	UNK	0.53	FORK	1.97	UNKNOWN	0	U		

otnunit_aat_animals_8dc3_4d15_c

Ready Count: 32766 100%

Type here to search

-5°C 9:11 PM 1/24/2022

otnunit_aat_datacenter_attributes_8a94_cefd_f8a3

The Primary Key for this table is “datacenter_reference”.

(1) We have many blank fields in this data set. We replace them with NaN.

FileHomeInsertPage LayoutFormulasDataReviewViewHelpTell me what you want to do

(2) We have columns in our data set which doesn't have any columns at all like datacenter_distribution_statement, datacenter_date_modified, time_coverage_start, time_coverage_end. We can delete these columns from our table.

otnunit_aat_datacenter_attributes_8a94_cefd_f8a3 [Compatibility Mode] - Excel

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	datacenter_abs tract	datacenter_cita tion	datacenter_pi anization	datacenter_pi contact	datacenter_info url	datacenter_keywo rds	datacenter_keyw ords_voc abulary	datacenter er_doi	datacenter_licen se	datacenter_geospatial_l on_min	datacenter_geospatial_l on_max	datacenter_geospatial_lat min	datacenter_geospatial_lat max			
1																
2																
3	Halifax Nova Scork Data Centre -	lor of Data Managi	OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	NaN	NaN	NaN	NaN			
4	starting in 2008, cking Network Dor of Data Managi		OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	formed citations	-108.511	148.83	-50.9809	74.75011			
5	alifax Nova Scotivork Data Centreor of Data Managi		OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	16	36	-34.6304	-24.5			
6	Halifax Nova Scork Data Centre -	lor of Data Managi	OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	-176.9	-55.6	4.9932	60.7663			
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																

otnunit_aat_datacenter_attritut

(3) We have integer columns where NaN might not sound reasonable. We replace them with 0.

otnunit_aat_datacenter_attributes_8a94_cefd_f8a3 [Compatibility Mode] - Excel

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	datacenter_abs tract	datacenter_cita tion	datacenter_pi anization	datacenter_pi contact	datacenter_info url	datacenter_keywo rds	datacenter_keyw ords_voc abulary	datacenter er_doi	datacenter_licen se	datacenter_geospatial_l on_min	datacenter_geospatial_l on_max	datacenter_geospatial_lat min	datacenter_geospatial_lat max			
1																
2																
3	Halifax Nova Scork Data Centre -	lor of Data Managi	OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	0	0	0	0			
4	starting in 2008, cking Network Dor of Data Managi		OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	formed citations	-108.511	148.83	-50.9809	74.75011			
5	alifax Nova Scotivork Data Centreor of Data Managi		OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	16	36	-34.6304	-24.5			
6	Halifax Nova Scork Data Centre -	lor of Data Managi	OTN	otndc@dal.ca	members.ocean	NCE > BIOLOGICAL C	GCMD	10.1428	ers and to OTN by	-176.9	-55.6	4.9932	60.7663			
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																

otnunit_aat_datacenter_attritut

Average: 0 Count: 4 Sum: 0

otnunit_aat_detections_9062_5923_1394

The Primary Key for this table is “detecton_id”.

(1) There are many blank cells in this data set which can be replaced with NaN.

otnunit_aat_detections_9062_5923_1394 [Compatibility Mode] - Excel

Sign in

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Cut

Copy

Format Painter

Paste

Clipboard

Calibri

11

B

I

U

Wrap Text

General

\$ % , ' & # 0 00

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

AutoSum

Fill

Clear

Sort & Filter

Find & Select

Editing

O5

X

✓

f_x

NaN

	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
	detectio n_refer nce_id	detectio n_refer nce_type	transmitt er_code space	transmitt er_id	detectio n_transm item_nam e	detectio n_serial number	sensor_d ata	sensor_d ata_units	receiver log_id	deploym ent_id	detectio n_quality	depth	position_ data_sou rce	uncertain ty_in_lati tude	uncertain ty_in_lon gitude	depth_da ta_sourc e	uncertain ty_in_de pth	other_po sition_da ta	dataset_ quality		
1																					
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	m	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
3	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
5	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
6	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
7	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
8	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
9	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
10	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
11	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
12	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
13	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1085	NaN	NaN	NaN	SW1-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
14	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
15	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
16	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
17	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
18	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
19	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
20	WRS-10563	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	river Meta	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

otnunit_aat_detections_9062_592

+

1

Ready

Type here to search

W

X

10:04 PM

1/24/2022

-6°C

100%

(2) There are many columns in this data set which doesn't have any value at all. So as in the process of cleaning, we can delete those columns for good. Those columns are receiver_log_id, uncertainty_in_latitude, uncertainty_in_longitude, depth_data_source, uncertainty_in_depth, other_position_data, dataset_quality.

otnunit_aat_detections_9062_5923_1394 [Compatibility Mode] - Excel

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	detection_guid	time	latitude	longitude	tracker_reference	detection_reference_id	detection_reference_type	transmitter_codespace	transmitter_id	detection_transmitter_name	detection_serial_number	sensor_data	sensor_data_units	deployment_id	detection_quality	depth	position_data_source			
1																				
2	NaN	UTC	degrees_north	degrees_east	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	m	NaN			
3	ISA69-130:05-20T10:1	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
4	ISA69-130:05-20T07:3	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
5	ISA69-130:05-20T07:3	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
6	ISA69-130:05-20T14:3	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
7	ISA69-130:05-20T04:4	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
8	ISA69-130:05-20T04:4	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
9	ISA69-130:05-20T08:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
10	ISA69-130:05-20T08:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
11	ISA69-130:05-20T08:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
12	ISA69-130:05-20T17:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
13	ISA69-130:05-19T03:0	44.92408	-62.5425	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1085	NaN	NaN	NaN	SW1-VR2-	NaN	NaN	Receiver Metadata			
14	ISA69-130:05-19T04:0	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
15	ISA69-130:05-19T05:1	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
16	ISA69-130:05-19T06:1	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
17	ISA69-130:05-20T08:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
18	ISA69-130:05-20T10:3	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
19	ISA69-130:05-20T14:3	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			
20	ISA69-130:05-19T05:5	44.91786	-62.5363	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	NaN	NaN	NaN	SW2-VR2-	NaN	NaN	Receiver Metadata			

(3) As we know that sensor_data is an integer, it is more reasonable to have default value as 0 instead of NaN.

(4) In the detection_quality column, we replace NaN data with “no receiver found”.

The screenshot shows an Excel spreadsheet titled 'otnunit_aat_detections_9062_5923_1394 [Compatibility Mode] - Excel'. The table has columns for various tracking and detection data. The 'detection_quality' column (column R) contains the text 'No receiver found' for all rows. The 'position_data_source' column (column S) contains 'Receiver Metadata' for all rows.

	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	latitude	longitude	tracker_reference	detection_reference_id	detection_reference_type	transmitter_codespace	transmitter_id	detection_transmittername	detection_serial_number	sensor_data	sensor_data_units	deployment_id	detection_quality	position_data_source			
1	degrees_north	degrees_east															
2	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
3	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
4	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
5	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
6	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
7	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
8	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
9	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
10	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
11	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
12	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
13	44.92408	-62.5425	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1085	0	NaN	SW1-VR2	No receiver found	Receiver Metadata			
14	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
15	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
16	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
17	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
18	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
19	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			
20	44.91786	-62.53628	WRS	WRS-1056	ANIMAL	A69-1303	10563	9-1303-105	1136	0	NaN	SW2-VR2	No receiver found	Receiver Metadata			

otnunit_aat_manmade_platform_0735_7c9f_329c

The Primary Key for this table is “platform_guid”.

(1) First step is to replace blank cells with NaN to maintain same uniformity.

otnunit_aat_manmade_platform_0735_7c9f_329c [Compatibility Mode] - Excel

	A	B	C	D	E	F	G	H	I
50	SAF	OTN-Global	SAF-PA001	OTN-GlobalSAFSAF-PA001	Underwater mooring	28.4	SAF-PA001	-33.62376	26.89191
51	SAF	OTN-Global	SAF-PA002	OTN-GlobalSAFSAF-PA002	Underwater mooring	34.2	SAF-PA002	-33.62813	26.89851
52	SAF	OTN-Global	GSB-GSB002	OTN-GlobalSAFSAF-GSB002	Underwater mooring	29	GSB-GSB002	-34.70415	19.40503
53	SAF	OTN-Global	SAF-MOZ004	OTN-GlobalSAFSAF-MOZ004	Underwater mooring	44.9	SAF-MOZ004	-26.85447	32.91735
54	SAF	OTN-Global	SAF-PA003	OTN-GlobalSAFSAF-PA003	Underwater mooring	39.2	SAF-PA003	-33.63283	26.90616
55	SAF	OTN-Global	SAF-PA004	OTN-GlobalSAFSAF-PA004	Underwater mooring	50.2	SAF-PA004	-33.63732	26.91446
56	SAF	OTN-Global	SAF-PA005	OTN-GlobalSAFSAF-PA005	Underwater mooring	61.1	SAF-PA005	-33.64155	26.92158
57	SAF	OTN-Global	RECAPTURE-SAIAB-52-14	OTN-GlobalSAFRECAPTURE-SAIAB-52-14	Underwater mooring	NaN	RECAPTURE-SAIAB-52-14	-33.86534	25.63307
58	SAF	OTN-Global	RECAPTURE-SAIAB-r0-14	OTN-GlobalSAFRECAPTURE-SAIAB-r0-14	Underwater mooring	NaN	RECAPTURE-SAIAB-r0-14	-33.86534	25.63307
59	SAF	OTN-Global	SAF-MOZ005	OTN-GlobalSAFSAF-MOZ005	Underwater mooring	49.5	SAF-MOZ005	-26.85447	32.92342
60	SAF	OTN-Global	SAF-Sundays	OTN-GlobalSAFSAF-Sundays	Underwater mooring	1.5	SAF-Sundays	-33.69511	25.83502
61	SAF	OTN-Global	SAF-PSJ001	OTN-GlobalSAFSAF-PSJ001	Underwater mooring	20	SAF-PSJ001	-31.61339	29.58931
62	SAF	OTN-Global	SAF-PSJ002	OTN-GlobalSAFSAF-PSJ002	Underwater mooring	34	SAF-PSJ002	-31.61746	29.59521
63	SAF	OTN-Global	SAF-PSJ003	OTN-GlobalSAFSAF-PSJ003	Underwater mooring	42	SAF-PSJ003	-31.62116	29.60115
64	SAF	OTN-Global	SAF-PSJ004	OTN-GlobalSAFSAF-PSJ004	Underwater mooring	54	SAF-PSJ004	-31.62494	29.60773
65	SAF	OTN-Global	SAF-PSJ005	OTN-GlobalSAFSAF-PSJ005	Underwater mooring	70	SAF-PSJ005	-31.62851	29.61385
66	SAF	OTN-Global	RECAPTURE-SAIAB-23-14	OTN-GlobalSAFRECAPTURE-SAIAB-23-14	Underwater mooring	NaN	RECAPTURE-SAIAB-23-14	-33.72009	25.85276
67	SAF	OTN-Global	GSB-GSB003	OTN-GlobalSAFSAF-GSB003	Underwater mooring	29	GSB-GSB003	-34.71118	19.40718
68	SAF	OTN-Global	EST-Coega	OTN-GlobalSAFEST-Coega	Underwater mooring	3	EST-Coega	-33.79908	25.69785
69	SAF	OTN-Global	EST-Gamtoos	OTN-GlobalSAFEST-Gamtoos	Underwater mooring	1.6	EST-Gamtoos	-33.96096	25.01524
70	SAF	OTN-Global	EST-Gouritz	OTN-GlobalSAFEST-Gouritz	Underwater mooring	1	EST-Gouritz	-34.33985	21.881
71	SAF	OTN-Global	EST-Keiskamma	OTN-GlobalSAFEST-Keiskamma	Underwater mooring	1.3	EST-Keiskamma	-33.27771	27.48596
72	SAF	OTN-Global	EST-Knysna	OTN-GlobalSAFEST-Knysna	Underwater mooring	6	EST-Knysna	-34.06535	23.04641

Ready Average: 135.2607963 Count: 8940 Sum: 903136.3366

(2) In the integer columns we have (platform_depth, latitude, longitude), it makes sense to have default values as 0 instead of NaN.

otnunit_aat_manmade_platform_0735_7c9f_329c [Compatibility Mode] - Excel

	B	C	D	E	F	G	H	I	J
575	nep	VEDDER001	nepFRASERVEDDER001	Underwater mooring	3.4	VEDDER001	49.14286	-122.16423	
576	nep	VEDDER002	nepFRASERVEDDER002	Underwater mooring	4.6	VEDDER002	49.13172	-122.09554	
577	nep	DERBY-RKm51.3_01	nepFRASERDERBY-RKm51.3_01	Underwater mooring	1.5	DERBY-RKm51.3_01	49.20158	-122.596	
578	nep	NORTHARM-RKm17.8_01	nepFRASERNORTHARM-RKm17.8_01	Underwater mooring	3.7	NORTHARM-RKm17.8_01	49.18663	-122.99635	
579	nep	FRASER164	nepFRASERFRASER164	Underwater mooring	0	FRASER164	49.29075	-121.881	
580	nep	FRASER165	nepFRASERFRASER165	Underwater mooring	0	FRASER165	49.3155	-121.80394	
581	nep	FRASER166	nepFRASERFRASER166	Underwater mooring	0	FRASER166	51.18544	-122.12569	
582	nep	FRASER167	nepFRASERFRASER167	Underwater mooring	0	FRASER167	51.62772	-124.14219	
583	nep	FRASER168	nepFRASERFRASER168	Underwater mooring	0	FRASER168	52.61369	-121.55792	
584	nep	FRASER169	nepFRASERFRASER169	Underwater mooring	0	FRASER169	50.82889	-119.6955	
585	nep	FRASER170	nepFRASERFRASER170	Underwater mooring	0	FRASER170	50.88597	-119.55531	
586	nep	FRASER171	nepFRASERFRASER171	Underwater mooring	0	FRASER171	50.89108	-119.53592	
587	nep	FRASER172	nepFRASERFRASER172	Underwater mooring	0	FRASER172	50.89736	-119.54708	
588	nep	FRASER173	nepFRASERFRASER173	Underwater mooring	0	FRASER173	50.83703	-119.00133	
589	nep	FRASER174	nepFRASERFRASER174	Underwater mooring	0	FRASER174	50.79381	-118.99961	
590	nep	FRASER175	nepFRASERFRASER175	Underwater mooring	0	FRASER175	50.77028	-120.81306	
591	nep	FRASER176	nepFRASERFRASER176	Underwater mooring	0	FRASER176	50.35892	-121.39719	
592	nep	FRASER177	nepFRASERFRASER177	Underwater mooring	0	FRASER177	50.68403	-120.36144	
593	nep	FRASER178	nepFRASERFRASER178	Underwater mooring	0	FRASER178	50.66864	-119.59897	
594	nep	PORT MANN-RKm31.7_02	nepFRASERPORT MANN-RKm31.7_02	Underwater mooring	4	PORT MANN-RKm31.7_02	49.22137	-122.83256	
595	nep	SOUTHARM-RKm20.5_05	nepFRASERSOUTHARM-RKm20.5_05	Underwater mooring	7.3	SOUTHARM-RKm20.5_05	49.17241	-122.96179	
596	nep	SOUTHARM-RKm9.3_01	nepFRASERSOUTHARM-RKm9.3_01	Underwater mooring	11.9	SOUTHARM-RKm9.3_01	49.1217	-123.07061	
597	neo	SOUTHARM-RKm9.3_02	nepFRASERSOUTHARM-RKm9.3_02	Underwater mooring	14.6	SOUTHARM-RKm9.3_02	49.12257	-123.07166	

Ready

(3) There are few duplicate values in our table. We need to get rid off them.

otnunit_aat_manmade_platform_0735_7c9f_329c [Compatibility Mode] - Excel

	A	B	C	D	E	F	G	H	I
631	FWS	nep	Alex Fraser Bridge	nepFWSAlex Fraser Bridge	Underwater mooring	0	Alex Fraser Bridge	49.1632	-122.9417
632	FWS	nep	Coquihalla/Hope 1	nepFWSCoquihalla/Hope 1	Underwater mooring	0	Coquihalla/Hope 1	49.3798	-121.4797
633	FWS	nep	Coquihalla/Hope 2	nepFWSCoquihalla/Hope 2	Underwater mooring	0	Coquihalla/Hope 2	49.3996	-121.433
634	FWS	nep	Debris trap downstream	nepFWSDebris trap downstream	Underwater mooring	0	Debris trap downstream	49.3247	-121.6378
635	FWS	nep	Harrison	nepFWSHarrison	Underwater mooring	0	Harrison	49.2343	-121.9581
636	FWS	nep	Harrison 2	nepFWSHarrison 2	Underwater mooring	0	Harrison 2	49.2619	-121.9173
637	FWS	nep	Harrison Upstream	nepFWSHarrison Upstream	Underwater mooring	0	Harrison Upstream	49.2264	-121.9009
638	FWS	nep	Herrling 1	nepFWSHerrling 1	Underwater mooring	0	Herrling 1	49.2155	-121.7055
639	FWS	nep	Herrling 2	nepFWSHerrling 2	Underwater mooring	0	Herrling 2	49.2523	-121.6824
640	FWS	nep	Hunter Creek	nepFWSHunter Creek	Underwater mooring	0	Hunter Creek	49.3612	-121.5684
641	FWS	nep	Hunter mouth	nepFWSHunter mouth	Underwater mooring	0	Hunter mouth	49.3599	-121.5721
642	FWS	nep	Hwy 11	nepFWSHwy 11	Underwater mooring	0	Hwy 11	49.1248	-122.3098
643	FWS	nep	Hwy 11(2)	nepFWSHwy 11(2)	Underwater mooring	0	Hwy 11(2)	49.1248	-122.3108
644	FWS	nep	Jespersons 1	nepFWSJespersons 1	Underwater mooring	0	Jespersons 1	49.2144	-121.8788
645	FWS	nep	Jespersons 2	nepFWSJespersons 2	Underwater mooring	0	Jespersons 2	49.2097	-121.8439
646	FWS	nep	lower canyon	nepFWSlower canyon	Underwater mooring	0	lower canyon	49.4488	-121.4148
647	FWS	nep	Lower canyon	nepFWSLower canyon	Underwater mooring	0	Lower canyon	49.4488	-121.4148
648	FWS	nep	Pitt 1	nepFWSPitt 1	Underwater mooring	0	Pitt 1	49.2471	-122.7289
649	FWS	nep	Ruby 1	nepFWSRuby 1	Underwater mooring	0	Ruby 1	49.3553	-121.6031
650	FWS	nep	Ruby 2	nepFWSRuby 2	Underwater mooring	0	Ruby 2	49.3592	-121.5953
651	FWS	nep	Ruby 3	nepFWSRuby 3	Underwater mooring	0	Ruby 3	49.3563	-121.5847
652	FWS	nep	Seabird 1	nepFWSSeabird 1	Underwater mooring	0	Seabird 1	49.2933	-121.6783
653	FWS	neo	Seabird 2	neoFWSSeabird 2	Underwater mooring	0	Seabird 2	49.3004	-121.6795

otnunit_aat_manmade_platform_07

Ready Average: -23.98866667 Count: 18 Sum: -143.932 100%

11:32 PM 1/24/2022

otnunit_aat_project_attributes_f29c_fb21_23a3

The Primary Key for this table is “project_reference”.

(1) As we know the first step, we replace all the blank cells with NaN.

otnunit_aat_project_attributes_f29c_fb21_23a3 [Compatibility Mode] - Excel

FileHomeInsertPage LayoutFormulasDataReviewViewHelpTell me what you want to do

CutCopyFormat Painter

PasteFormat Painter

Calibri11

B

I

U

Font

Wrap Text

Merge & Center

Alignment

General

\$ % , ' 1000 00

Number

Conditional Formatting

Format as Table

Cell Styles

InsertDeleteFormat

Cells

AutoSum

FillClear

Sort & Find & FilterSelect

Editing

AA1time_coverage_end

	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
	project_k eywords	project_k eywords vocabulary	project_r reference s	project_d doi	project_l linestrings	project_d distribution statement	project_d date_modified	project_d atum	project_g eospacial lon_min	project_g eospacial lon_max	project_g eospacial lat_min	project_g eospacial lat_max	project_l linestrings	geospatial vertical al_min	geospatial vertical al_max	geospatial vertical al_posi tive	time_cov erage_st art	time_cov erage_en d			
1																					
2																					
3	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	18.45381	32.94051	-26.7436	-34.4094	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
4	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	30.82757	30.82757	-30.244	-30.244	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
5	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	-67	-65.5	47.4016	45	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
6	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	34.5	36	-23.5	-24.5	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
7	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	19.37508	19.37508	-34.6304	-34.6304	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
8	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	21.429	21.429	-34.381	-34.381	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
9	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	25.801	32.9	-26.84	-33.8343	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
10	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	16	34	-26	-35	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
11	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	30.84991	31.16991	-29.7018	-30.0218	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
12	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	19.19611	19.27611	-34.3775	-34.4575	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
13	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	35.36427	35.36427	-23.8562	-23.8562	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
14	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	-64.83	-63.55	48.36	47.54	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
15	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	26.28025	29.57	-31.62	-33.8343	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
16	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	18.51807	18.51807	-34.3496	-34.3496	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
17	; EARTH SC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	-70.6877	-69.9286	42.08232	41.51499	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
18	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	18.63217	25.83	-33.7874	-34.785	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
19	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	20.97	20.97	-34.45	-34.45	NaN	NaN	NaN	NaN	NaN	NaN	NaN		
20	ACOUSTIC GCMD		NaN	NaN	By access	NaN	NaN	WGS84	28.895	28.91	-32.25	-32.26	NaN	NaN	NaN	NaN	NaN	NaN	NaN		

otnunit_aat_project_attributes

Average: 45.13588909Count: 3010Sum: 14894.8434

Ready

11:39 PM 1/24/2022

(2) There are few columns in this data set which doesn't have any value at all.

So as in the process of cleaning, we can delete those columns for good.

Those columns are project_references, project_doi, project_distribution_statement, project_date_modified, project_linestrings, geospatial_vertical_positive, time_coverage_start, time_coverage_end.

(3) In the columns “geospatial_vertical_min” and “geospatial_vertical_max”, since these are integer type columns, we replace NaN values in this column with 0.

(4) The whole “geospatial_vertical_min” column is 0, so we can delete that column from our table.

The screenshot shows an Excel spreadsheet titled "otnunit_aat_project_attributes_f29c_fb21_23a3 [Compatibility Mode] - Excel". The spreadsheet contains a table with the following columns: project_p_i, project_p_i_organization, project_p_i_contact, project_i_nfourl, project_k_eywords, project_k_eywords_vocabulary, project_license, project_datum, project_geospatial_lon_min, project_geospatial_lon_max, project_geospatial_lat_min, project_geospatial_lat_max, project_geospatial_vertical_max, and project_geospatial_vertical_min. The 'geospatial_vertical_min' column is highlighted with a green box. The data rows show various project entries with their respective coordinates and other attributes.

	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	project_p_i	project_p_i_organization	project_p_i_contact	project_i_nfourl	project_k_eywords	project_k_eywords_vocabulary	project_license	project_datum	project_geospatial_lon_min	project_geospatial_lon_max	project_geospatial_lat_min	project_geospatial_lat_max	project_geospatial_vertical_max	project_geospatial_vertical_min	
1									degrees_east	degrees_east	degrees_north	degrees_north			
2															
3	Paul Cowle	SAIAB	sh@gmail.com	www.saiab.org	LOGICAL	GCMD	formed citations	WGS84	18.453813	32.94051	-26.7436	-34.40943	0		
4	Arrel Ande	DEA	environm	environn	LOGICAL	GCMD	formed citations	WGS84	30.82757	30.82757	-30.24398	-30.24398	0		
5	Allen Curry	UNB-F	curry@unbtrack.org	c/sSIFICAT	GCMD	formed citations	WGS84		-67	-65.5	47.4015978	45	0		
6	alum Muri	MMF	urrie@gmick.org	datSSIFICAT	GCMD	formed citations	WGS84		34.5	36	-23.5	-24.5	0		
7	ison Town	DICT	wner@gmick.org	daDLOGICAL	GCMD	formed citations	WGS84		19.37508	19.37508	-34.63038	-34.63038	0		
8	ve Lambe	DAFF	enl@daffrack.org	dDLOGICAL	GCMD	formed citations	WGS84		21.429	21.429	-34.381	-34.381	0		
9	Att Dicke	NMMU	ark@bayack.org	dDLOGICAL	GCMD	formed citations	WGS84		25.801	32.9	-26.84	-33.83425	0		
10	inus Beuke	TOA	ces@aqua	ww.aquari	SSIFICAT	GCMD	formed citations	WGS84		16	34	-26	-35	0	
11	ruce Mani	ORI	ce@ori.org	rv.seaworl	DLOGICAL	GCMD	formed citations	WGS84		30.849909	31.169909	-29.701825	-30.021825	0	
12	aghan McC	SASC	arkconser	sharkcons	SSIFICAT	GCMD	formed citations	WGS84		19.1961111	19.2761111	-34.3775	-34.4575	0	
13	NaN	MMF	NaN	ack.org/da	SSIFICAT	GCMD	formed citations	WGS84		35.364267	35.364267	-23.856172	-23.856172	0	
14	Doug Swait	DFO-GFC	ain@dfo-ntrack.org	DLOGICAL	GCMD	formed citations	WGS84		-64.83	-63.55	48.36	47.54	0		
15	Jeremy Clij	KZNSB	f@shark.co	www.shar	DLOGICAL	GCMD	formed citations	WGS84		26.280245	29.57	-31.62	-33.834252	0	
16	Alison Koclark	Spottekock	gmklarkspotte	DLOGICAL	GCMD	formed citations	WGS84		18.518072	18.518072	-34.349615	-34.349615	0		
17	Whoriska-NMFS-Nbriskey	key@d/https://	SSIFICAT	GCMD	formed citations	WGS84			-70.6876673	-69.9285511	42.0823174	41.5149887	0		
18	alcolm Smi	Bayworld	3bayworld	ack.org/dDLOGICAL	GCMD	formed citations	WGS84		18.632165	25.83	-33.78743	-34.78497	0		
19	aghan McC	SASC	arkconserrack.org	dDLOGICAL	GCMD	formed citations	WGS84		20.97	20.97	-34.45	-34.45	0		
20	lan Venter	ECPT	ven@emack.org	dDLOGICAL	GCMD	formed citations	WGS84		28.895	28.91	-32.25	-32.26	0		

otnunit_aat_receivers_c595_05f4_68b2

The Primary Key for this table is “deployment_guid”.

(1) As we know the first step, we replace all the blank cells with NaN.

otnunit_aat_receivers_c595_05f4_68b2 [Compatibility Mode] - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

deployment_guid

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
	deployment_ project_refer ence	datacent er_referenc e	deployment_ id	deployment_ guid	receiver_ manufactur er	receiver_ model	frequenc ies_monit ored	receiver_ coding_s cheme	receiver_ serial_nu mber	latitude	longitud e	time	recovery_ datetim e_utc	array_na me	receiver_ referenc e_type	receiver_ referenc e_id	bottom_ depth	depth	deploym ent_com ments	deployed_ by	expect _receiv _lif
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	grees_noregrees_ea	UTC	UTC	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	ASF	DTN-Globa	line D-VR2F-West Lin	VEMCO	VR2	NaN	NaN	NaN	2272	50.28333	-64.4356	06-13T18:008-06T20:0	ASF	madePlat	ASF-West l	20.9	7	NaN	NaN	NaN	NaN
3	ASF	DTN-Globa	-KedgewiRS POOL-K	VEMCO	VR2	NaN	NaN	NaN	2603	47.81337	-67.7385	05-23T03:010-05T03:0	ASF	madePlat	VERS POOL	NaN	NaN	NaN	NaN	NaN	NaN
4	ASF	DTN-Globa	-KedgewiRS POOL-K	VEMCO	VR2	NaN	NaN	NaN	3342	47.81336	-67.7385	05-27T03:010-01T03:0	ASF	madePlat	VERS POOL	NaN	NaN	NaN	NaN	NaN	NaN
5	ASF	DTN-Globa	f the Oxb/river of f	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	ASF	DTN-Globa	Falls-VR2F-Above Fi	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13	ASF	DTN-Globa	Island 1-Vnticosti Isl	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	ASF	DTN-Globa	Island 2-Vnticosti Isl	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
15	ASF	DTN-Globa	Island 3-Vnticosti Isl	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
16	ASF	DTN-Globa	Island 4-Vnticosti Isl	VEMCO	VR2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	SAF	DTN-Globa	aoos-VR2Wt-Gamtoo	NaN	VR2W	NaN	NaN	NaN	119101	-33.961	25.01524	07-04T11:502-25T14:2	SAF	madePlat	_EST-Gamf	1.6	1.6	NaN	NaN	NaN	NaN
18	SAF	DTN-Globa	-U22-001--Gouritz-U	NaN	U22-001	NaN	NaN	NaN	10098862	-34.3399	21.881	10-23T11:002-20T14:3	SAF	madePlat	_EST-Gou	1	1	NaN	NaN	NaN	NaN
19	SAF	DTN-Globa	ritz-VR2W-ST-Gouritz	NaN	VR2W	NaN	NaN	NaN	102853	-34.3399	21.881	10-23T11:002-20T14:3	SAF	madePlat	_EST-Gou	1	1	NaN	NaN	NaN	NaN
20	SAF	DTN-Globa	ana-U22-00eiskamma	NaN	U22-001	NaN	NaN	NaN	10098859	-33.2777	27.48596	02-17T12:503-12T13:0	SAF	madePlat	EST-Keiska	1.5	1.6	NaN	NaN	NaN	NaN

otnunit_aat_receivers_c595_05f4

Ready

Type here to search

-6°C 11:57 PM 1/24/2022

(2) There are few columns in this data set which doesn't have any value at all.

So as in the process of cleaning, we can delete those columns for good.

Those columns are frequencies_monitored, receiver_coding_scheme, deployed_by, expected_receiver_life.

otnunit_aat_receivers_c595_05f4_6bb2 [Compatibility Mode] - Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
		deployment_project_reference	datacenter_reference	deployment_id	deployment_guid	receiver_manufacturer	receiver_model	receiver_serial_number	latitude	longitude	time	recovery_datetime_utc	array_name	receiver_reference_type	receiver_reference_id	bottom_depth	depth	deployment_comments			
1		NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN			
2	ASF	DTN-Global-ine D-VR2F-West Lin			VEMCO	VR2	2272	50.28333	-64.4356	06-13T18:08-06T20:00	UTC	NaN	ASF	madePlatfASF-West L		20.9	7	NaN			
3	ASF	DTN-Global-KedgewiRS POOL-K			VEMCO	VR2	2603	47.81337	-67.7385	05-23T03:00-05T03:00	UTC	NaN	ASF	madePlatfERS POOL		NaN	NaN	NaN			
4	ASF	DTN-Global-KedgewiRS POOL-K			VEMCO	VR2	3342	47.81336	-67.7385	05-27T03:00-01T03:00	UTC	NaN	ASF	madePlatfERS POOL		NaN	NaN	NaN			
5	ASF	DTN-Global-Priver of the Oxbow			VEMCO	VR2	1135	46.95367	-65.8715	05-17T03:06-26T03:00	UTC	NaN	ASF	madePlatfUpriver c		1.538	NaN	NaN			
6	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	3335	50.53819	-64.0892	06-17T12:408-06T17:00	UTC	NaN	ASF	madePlatfASF-Above		1.2	1.2	NaN			
7	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	3335	50.53819	-64.0892	06-26T18:408-07T17:00	UTC	NaN	ASF	madePlatfASF-Above		1.2	1.2	NaN			
8	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	3342	50.53819	-64.0892	06-19T12:308-13T18:00	UTC	NaN	ASF	madePlatfASF-Above		1.2	NaN	NaN			
9	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	3349	50.53819	-64.0892	08-07T14:106-19T16:00	UTC	NaN	ASF	madePlatfASF-Above		0.6	0.6	NaN			
10	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	3357	50.53819	-64.0892	06-18T03:07-30T17:00	UTC	NaN	ASF	madePlatfASF-Above		1.5	1.5	NaN			
11	ASF	DTN-Global-Falls-VR2F-Above Falls			VEMCO	VR2	5032	50.53819	-64.0892	08-13T15:108-14T03:00	UTC	NaN	ASF	madePlatfASF-Above		NaN	NaN	NaN			
12	ASF	DTN-Global Island 1-Vnticosti Is			VEMCO	VR2	2587	49.89227	-64.5363	05-31T21:08-04T19:00	UTC	NaN	ASF	madePlatf-Anticosti		30	7	NaN			
13	ASF	DTN-Global Island 1-Vnticosti Is			VEMCO	VR2	2581	49.89673	-64.55	05-31T21:08-03T10:00	UTC	NaN	ASF	madePlatf-Anticosti		60	7	NaN			
14	ASF	DTN-Global Island 3-Vnticosti Is			VEMCO	VR2	3335	49.90093	-64.5633	05-31T21:08-03T18:00	UTC	NaN	ASF	madePlatf-Anticosti		70	7	NaN			
15	ASF	DTN-Global Island 4-Vnticosti Is			VEMCO	VR2	3346	49.90455	-64.5755	05-31T21:08-04T20:00	UTC	NaN	ASF	madePlatf-Anticosti		80	7	NaN			
16	SAF	DTN-Globaloos-VR2W-T-Gamtoos			NaN	VR2W	119101	-33.961	25.01524	07-04T11:502-25T14:00	UTC	NaN	SAF	madePlatf_EST-Gam		1.6	1.6	NaN			
17	SAF	DTN-Global-U22-001-Gouritz-L			NaN	U22-001	10098862	-34.3399	21.881	10-23T11:002-20T14:00	UTC	NaN	SAF	madePlatf_EST-Gou		1	1	NaN			
18	SAF	DTN-Global-U22-001-Gouritz-L			NaN	U22-001	102853	-34.3399	21.881	10-23T11:002-20T14:00	UTC	NaN	SAF	madePlatf_EST-Gou		1	1	NaN			
19	SAF	DTN-Global-U22-001-Gouritz-L			NaN	U22-001	10098859	-33.2777	27.48596	02-17T12:503-12T13:00	UTC	NaN	SAF	madePlatf_EST-Keiska		1.5	1.6	NaN			

(3) For integer type columns, it is better to replace NaN values with 0.

otnunit_aat_receivers_c595_05f4_6bb2 [Compatibility Mode] - Excel

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	datacenter_reference	deployment_id	deployment_guid	receiver_manufacturer	receiver_model	receiver_serial_number	latitude	longitude	time	recovery_datetime_utc	array_name	receiver_reference_type	receiver_reference_id	bottom_depth	depth	deployment_comments	
1	NaN	NaN	NaN	NaN	NaN	0	NaN	NaN	UTC	NaN	NaN	NaN	NaN	m	m	NaN	
2	OTN-Global	iF-West Line D-VR2-2272-205FASF-West Line D-V		VEMCO	VR2	2272	50.28333	-64.4356	06-13T18:08-06T20:00	UTC	ASF	madePlatfASF-West L		20.9	7	NaN	
3	OTN-Global	iRS POOL-KedgewiVR2-2 WYERS POOL-Kedgewi		VEMCO	VR2	2603	47.81337	-67.7385	05-23T03:00-05T03:00	UTC	ASF	madePlatfERS POOL		0	0	NaN	
4	OTN-Global	iRS POOL-KedgewiVR2-3 WYERS POOL-Kedgewi		VEMCO	VR2	3342	47.81336	-67.7385	05-27T03:00-01T03:00	UTC	ASF	madePlatfERS POOL		0	0	NaN	
5	OTN-Global	iUpriver of the Oxbow-VR2-1 km Upriver of the O		VEMCO	VR2	1135	46.95367	-65.8715	05-17T03:06-26T03:00	UTC	ASF	madePlatfUpriver c		1.538	0	NaN	
6	OTN-Global	iF-Above Falls-VR2-3335-205FASF-Above Falls-V		VEMCO	VR2	3335	50.53819	-64.0892	06-17T12:408-06T17:00	UTC	ASF	madePlatfASF-Above		1.2	1.2	NaN	
7	OTN-Global	iF-Above Falls-VR2-3335-205FASF-Above Falls-V		VEMCO	VR2	3335	50.53819	-64.0892	06-26T18:408-07T17:00	UTC	ASF	madePlatfASF-Above		1.2	1.2	NaN	
8	OTN-Global	iF-Above Falls-VR2-3342-205FASF-Above Falls-V		VEMCO	VR2	3342	50.53819	-64.0892	06-19T12:308-13T18:00	UTC	ASF	madePlatfASF-Above		1.2	0	NaN	
9	OTN-Global	iF-Above Falls-VR2-3349-205FASF-Above Falls-V		VEMCO	VR2	3349	50.53819	-64.0892	08-07T14:106-19T16:00	UTC	ASF	madePlatfASF-Above		0.6	0.6	NaN	
10	OTN-Global	iF-Above Falls-VR2-3357-205FASF-Above Falls-V		VEMCO	VR2	3357	50.53819	-64.0892	06-18T03:07-30T17:00	UTC	ASF	madePlatfASF-Above		1.5	1.5	NaN	
11	OTN-Global	iF-Above Falls-VR2-5032-205FASF-Above Falls-V		VEMCO	VR2	5032	50.53819	-64.0892	08-13T15:108-14T03:00	UTC	ASF	madePlatfASF-Above		0	0	NaN	
12	OTN-Global	iAnticosti Island 1-VR2-2587-2587-2587-2587-2587		VEMCO	VR2	2587	49.89227	-64.5363	05-31T21:08-04T19:00	UTC	ASF	madePlatf-Anticosti		30	7	NaN	
13	OTN-Global	iAnticosti Island 2-VR2-2581-2581-2581-2581-2581		VEMCO	VR2	2581	49.89673	-64.55	05-31T21:08-03T10:00	UTC	ASF	madePlatf-Anticosti		60	7	NaN	
14	OTN-Global	iAnticosti Island 3-VR2-3335-3335-3335-3335-3335		VEMCO	VR2	3335	49.90093	-64.5633	05-31T21:08-03T18:00	UTC	ASF	madePlatf-Anticosti		70	7	NaN	
15	OTN-Global	iAnticosti Island 4-VR2-3346-3346-3346-3346-3346		VEMCO	VR2	3346	49.90455	-64.5755	05-31T21:08-04T20:00	UTC	ASF	madePlatf-Anticosti		80	7	NaN	
16	OTN-Global	iST-Gamtoos-VR2W-119101-119101-119101-119101-119101		NaN	VR2W	119101	-33.961	25.01524	07-04T11:502-25T14:00	UTC	SAF	madePlatf_EST-Gam		1.6	1.6	NaN	
17	OTN-Global	iT-Gouritz-U22-001-10098862-10098862-10098862-10098862		NaN	U22-001	10098862	-34.3399	21.881	10-23T11:002-20T14:00	UTC	SAF	madePlatf_EST-Gou		1	1	NaN	
18	OTN-Global	iEST-Gouritz-VR2W-102853-102853-102853-102853-102853		NaN	VR2W	102853	-34.3399	21.881	10-23T11:002-20T14:00	UTC	SAF	madePlatf_EST-Gou		1	1	NaN	
19	OTN-Global	iKeiskamma-U22-001-10098859-10098859-10098859-10098859		NaN	U22-001	10098859	-33.2777	27.48596	02-17T12:503-12T13:00	UTC	SAF	madePlatf_EST-Keiska		1.5	1.6	NaN	

otnunit_aat_recover_offload_details_4b23_f002_f89a

The Primary Key for this table is “recovery_guid”.

(1) Basic step, to replace all blank values with NaN.

The screenshot shows a Microsoft Excel spreadsheet with the following columns: recovery_project_reference, datacenter_reference, recovery_id, deployment_id, recovery_guid, recovery_latitude, recovery_longitude, recovery_datetime_utc, recovery_outcome, data_offloaded, offload_datetime_utc, log_filenames, recovery_comments, and clock_synchronized. The recovery_guid column is highlighted in green. A Find and Replace dialog box is open, showing 'Find what:' as a blank cell and 'Replace with:' as 'NaN'. The 'Replace All' button is highlighted. A message box says 'All done. We made 77775 replacements.' The status bar at the bottom shows 'Ready' and 'Type here to search'.

(2) There are columns only with NaN values, which doesn't mean anything since there is no real data in it. So we delete these columns which are “clock_synchronized” and “recovered_by”.

otnunit_aat_recover_offload_details_4b23_f002_f89a [Compatibility Mode] - Excel													
File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do													
Calibri 11 A+ Wrap Text General Conditional Formatting Styles Cells Editing													
P6													
	recovery_id	deployment_id	recovery_guid	recovery_latitude	recovery_longitude	recovery_datetime_utc	recovery_outcome	data_offloaded	offload_datetime_utc	log_filenames	recovery_comments		
1													
2	NaN	NaN	NaN	NaN	NaN	UTC	NaN	NaN	UTC	NaN	NaN		
3	lar05-02-VR2W-12278	lar05-02-VR2W-12278	VR2W-122785-2016AE	67.2495	-60.38247	2017-09-22T23:18:00Z	RECOVERD	N	NaN	NaN	NaN		
4	EMP STAR ODDI-C813	EMP STAR ODDI-C813	136-2016-09-18T12AE	67.474117	-58.220567	2017-09-21T20:11:00Z	RECOVERD	N	NaN	NaN	NaN		
5	aff01-01-VR2W-12793	aff01-01-VR2W-12793	VR2W-127937-2016AE	71.97833	-71.44208	2017-10-01T15:46:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Up	
6	aff01-02-VR2W-12793	aff01-02-VR2W-12793	VR2W-127938-2016AE	72.00928	-71.34455	2017-10-01T14:09:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Up	
7	aff01-03-VR2W-12793	aff01-03-VR2W-12793	VR2W-127916-2016AE	72.04222	-71.24047	2017-10-01T14:27:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Down	
8	aff03-01-VR2W-12793	aff03-01-VR2W-12793	VR2W-127934-2016AE	70.68323	-66.8139	2017-09-29T23:15:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Up	
9	aff03-02-VR2W-12793	aff03-02-VR2W-12793	VR2W-127935-2016AE	70.68698	-66.76618	2017-09-29T23:42:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Up	
10	aff03-03-VR2W-12793	aff03-03-VR2W-12793	VR2W-127936-2016AE	70.6934	-66.68692	2017-09-30T00:15:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Down	
11	aff04-01-VR2W-11364	aff04-01-VR2W-11364	VR2W-113649-2016AE	68.9778	-64.08512	2017-10-07T20:20:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Down	
12	aff04-02-VR2W-11363	aff04-02-VR2W-11363	VR2W-113630-2016AE	68.97877	-64.18338	2017-10-07T21:11:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Up	
13	aff06-02-VR2W-12278	aff06-02-VR2W-12278	VR2W-122783-2016AE	67.68628	-62.37112	2017-09-25T18:43:00Z	RECOVERD	N	NaN	NaN	NaN		
14	aff06-03-VR2W-12279	aff06-03-VR2W-12279	VR2W-122792-2016AE	67.7546	-62.30852	2017-09-25T18:47:00Z	RECOVERD	N	NaN	NaN	NaN		
15	aff07-01-VR2W-12792	aff07-01-VR2W-12792	VR2W-127928-2016AE	66.83978	-59.43223	2017-09-24T21:31:00Z	RECOVERD	N	NaN	NaN	NaN		
16	aff07-02-VR2W-12793	aff07-02-VR2W-12793	VR2W-127931-2016AE	66.71638	-59.69172	2017-09-24T19:18:00Z	RECOVERD	N	NaN	NaN	NaN		
17	aff07-03-VR2W-12791	aff07-03-VR2W-12791	VR2W-127917-2016AE	66.59733	-59.96162	2017-09-24T17:10:00Z	RECOVERD	N	NaN	NaN	NaN		
18	Nar01-01-RBR-100816	Nar01-01-RBR-100816	RBR-100816-2016AE	68.1608	-59.7771	2017-09-20T21:01:00Z	RECOVERD	N	NaN	NaN	NaN		
19	lar01-01-VR2W-12792	lar01-01-VR2W-12792	VR2W-127924-2016AE	68.1608	-59.7771	2017-09-20T21:01:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Down	
20	lar01-02-VR2W-12990	lar01-02-VR2W-12990	VR2W-129908-2016AE	68.16305	-59.74157	2017-09-20T21:34:00Z	RECOVERD	N	NaN	NaN	NaN	it/Recovery Notes: Down	

(3) Integer type columns (recovery_latitude, recovery_longitude) should have the default values as 0 instead of NaN.

otnunit_aat_recover_offload_details_4b23_f002_f89a [Compatibility Mode] - Excel													
File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do													
Calibri 11 A+ Wrap Text General Conditional Formatting Styles Cells Editing													
G124													
14930	OTN-Global	32-126119-2015-2017-CJAY-M1902-126119-2015-2017-02-0		45.899475	-59.1420883		NaN	DOWNLOAD_ONL	Y	NaN	5119_2017	NaN	
14931	OTN-Global	31-106603-2015-2017-CJAY-M1901-106603-2015-2017-02-0		45.741415	-59.7417983		NaN	DOWNLOAD_ONL	Y	NaN	5603_2017	NaN	
14932	OTN-Global	30-127597-2015-2017-CJAY-M1900-127597-2015-2017-02-0		46.1298867	-59.3304817		NaN	DOWNLOAD_ONL	Y	NaN	7597_2017	NaN	
14933	OTN-Global	3-127314-2016-09-24T:946-VR2W-127314-2016-09-24T13:2018-03-2		46.16848333	-59.14149167		NaN	DOWNLOAD_ONL	Y	NaN	7314_2018	NaN	
14934	OTN-Global	3-127323-2016-09-23T:947-VR2W-127323-2016-09-23T18:2018-03-2		46.35565333	-58.72675667		NaN	DOWNLOAD_ONL	Y	NaN	7323_2018	NaN	
14935	OTN-Global	3-127320-2016-09-16T:949-VR2W-127320-2016-09-16T21:2018-03-2		43.609485	-62.87579		NaN	DOWNLOAD_ONL	Y	NaN	7320_2018	NaN	
14936	OTN-Global	3-127305-2016-11-11T:950-VR2W-127305-2016-11-11T15:2018-03-2		44.4626	-57.18323333		NaN	DOWNLOAD_ONL	Y	NaN	7305_2018	NaN	
14937	OTN-Global	3-127310-2016-09-24T:998-VR2W-127310-2016-09-24T16:2018-03-2		46.25	-59.140345		NaN	DOWNLOAD_ONL	Y	NaN	7310_2018	NaN	
14938	OTN-Global	3-127311-2016-09-23T:900-VR2W-127311-2016-09-23T12:2018-03-2		45.74142333	-59.741995		NaN	DOWNLOAD_ONL	Y	NaN	7311_2018	NaN	
14939	OTN-Global	3-127322-2016-09-23T:901-VR2W-127322-2016-09-23T15:2018-03-2		45.89957	-59.14176167		NaN	DOWNLOAD_ONL	Y	NaN	7322_2018	NaN	
14940	OTN-Global	32W-113051-2013-09-NWA-ROV-VR2W-113113051-2013-09-27BO		0	0		NaN	DOWNLOAD_ONL	Y	09-27T17:3	NaN	NaN	
14941	OTN-Global	32W-112343-2013-10-CS-Harpoon O-85-VR2-2013-10-18BOONWA		0	0		NaN	DOWNLOAD_ONL	Y	10-18T20:5	NaN	NaN	
14942	OTN-Global	32W-113609-2013-10-Deep Panuke-VR2W-09-2013-10-18BOONWA		0	0		NaN	DOWNLOAD_ONL	Y	10-18T20:5	NaN	NaN	
14943	OTN-Global	32W-119083-2013-10-CS-Harpoon O-85-VR2-2013-10-18BOONWA		0	0		NaN	DOWNLOAD_ONL	Y	10-18T20:5	NaN	NaN	
14944	OTN-Global	32W-120202-2013-10-Bay du Nord N-77-VI013-10-18BOONWA-F		0	0		NaN	DOWNLOAD_ONL	Y	10-18T20:5	NaN	NaN	
14945	OTN-Global	2W-119083-2013-10-1S-Federation K-87-VR013-10-18-2BOONWA		0	0		NaN	DOWNLOAD_ONL	Y	10-18T20:5	NaN	ttery died 2013-08-02	
14946	OTN-Global	32W-106614-2014-06-NWA-ROV-VR2W-106106614-2014-06-30BO		0	0		NaN	DOWNLOAD_ONL	Y	07-17T22:5	NaN	NaN	
14947	OTN-Global	32W-119091-2014-11-0ONWA-VR2W-119091W-119091-2014-11-06		43.71597	-59.842		NaN	DOWNLOAD_ONL	Y	11-06T20:4	NaN	NaN	
14948	OTN-Global	32W-110135-2015-03-M-4-WA-VR2W-11013V-110135-2015-03-02C		51.3821	-56.67063		NaN	DOWNLOAD_ONL	Y	03-02T19:40135_2015	NaN	NaN	
14949	OTN-Global	32W-110164-2015-03-0CM-3-VR2W-110164-2W-110164-2015-03-0		51.38008	-56.79143		NaN	DOWNLOAD_ONL	Y	03-02T21:30164_2015	NaN	NaN	
14950	OTN-Global	32W-110587-2015-03-CM-WA-VR2W-110587W-110587-2015-03-02		51.40695	-56.91525		NaN	DOWNLOAD_ONL	Y	03-02T19:40587_2015	NaN	NaN	
14951	OTN-Global	32W-110590-2015-03-0CM-1-VR2W-110590-2W-110590-2015-03-0		51.4333	-56.91958		NaN	DOWNLOAD_ONL	Y	03-02T19:30590_2015	NaN	NaN	
14952	OTN-Global	32W-112477-2015-03-CM-WA-VR2W-112477W-112477-2015-03-02		51.40688	-56.91538		NaN	DOWNLOAD_ONL	Y	03-02T19:2477_2015	NaN	NaN	

otnunit_aat_tag_releases_b793_03e7_a230

The Primary key for this table is “tag_device_id”.

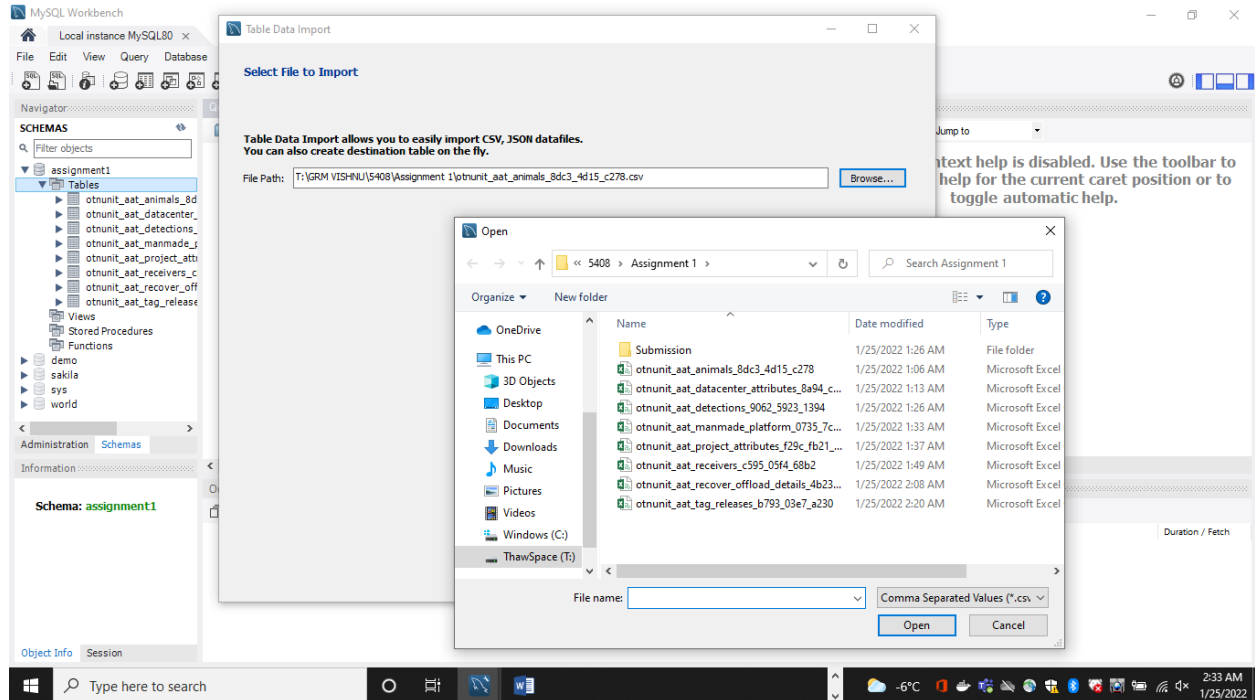
(1) Primary step is to replace empty cells with NaN.

The screenshot shows a Microsoft Excel spreadsheet with the following columns: release_project_reference, datacenter_reference, tag_device_id, release_guid, release_reference_id, release_type, latitude, longitude, time, expected_enddate, manufacturer, tag_model, tag_serial_number, tag_frequency, tag_coding_system, transmitter_id, transmitter_name, transmitter_type, and tag_programming_id. The data is organized into rows, with some cells containing NaN values. A 'Find and Replace' dialog box is open, showing 'Find what:' and 'Replace with:' fields, both set to 'NaN'. A message box is displayed over the spreadsheet, stating 'All done. We made 11526 replacements.' The status bar at the bottom indicates 'Ready' and 'Type here to search'.

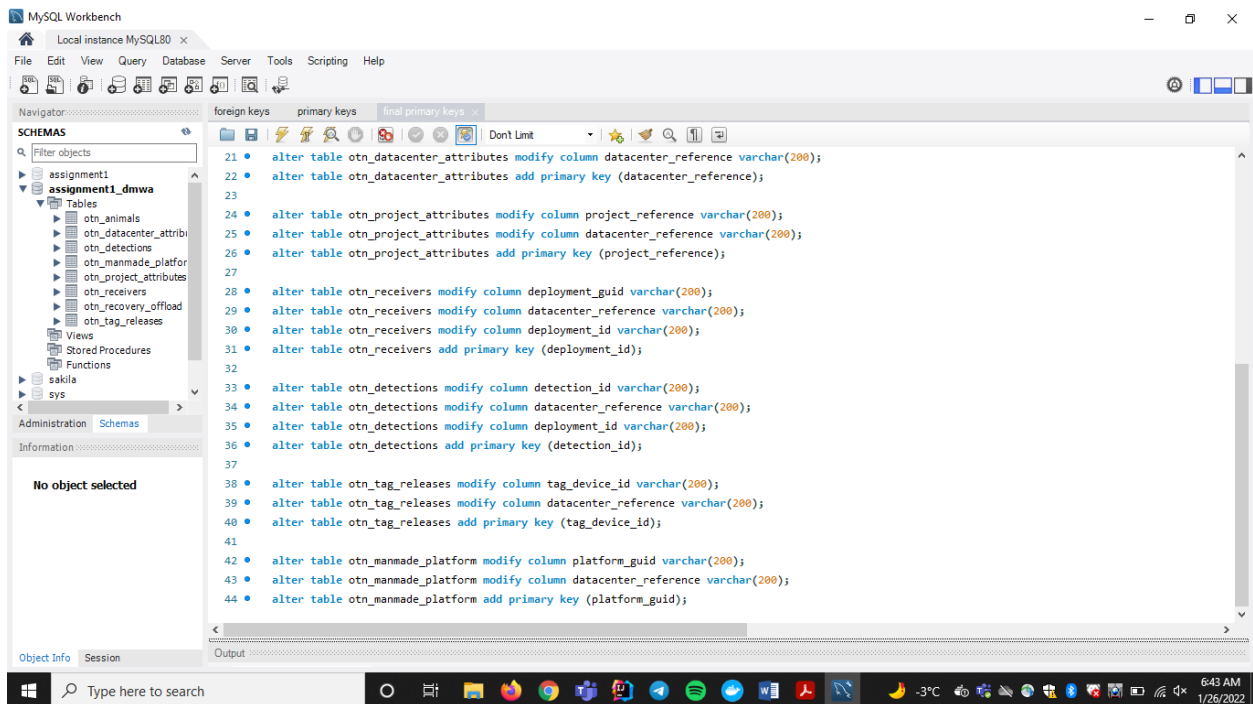
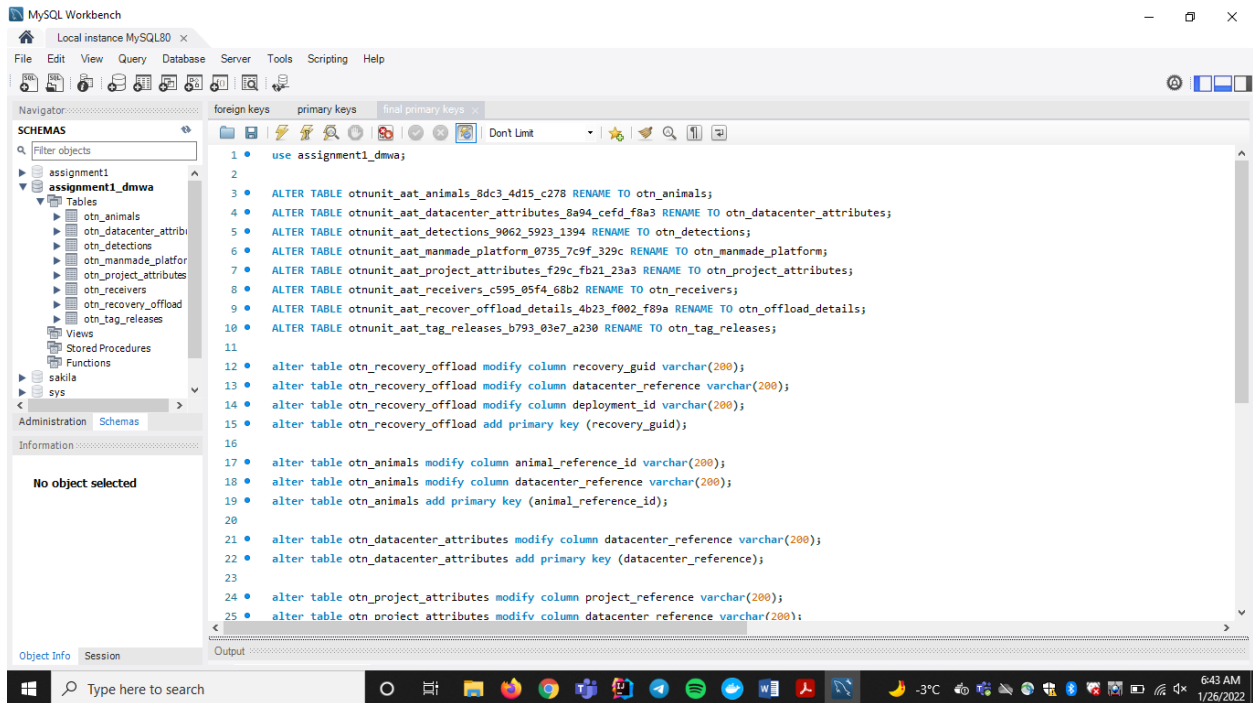
(2) Full columns with no values in it and which just have NaN values are to be deleted. The columns are tag_frequency, transmitter_type, tag_programming_id.

Importing csv files into workbench:

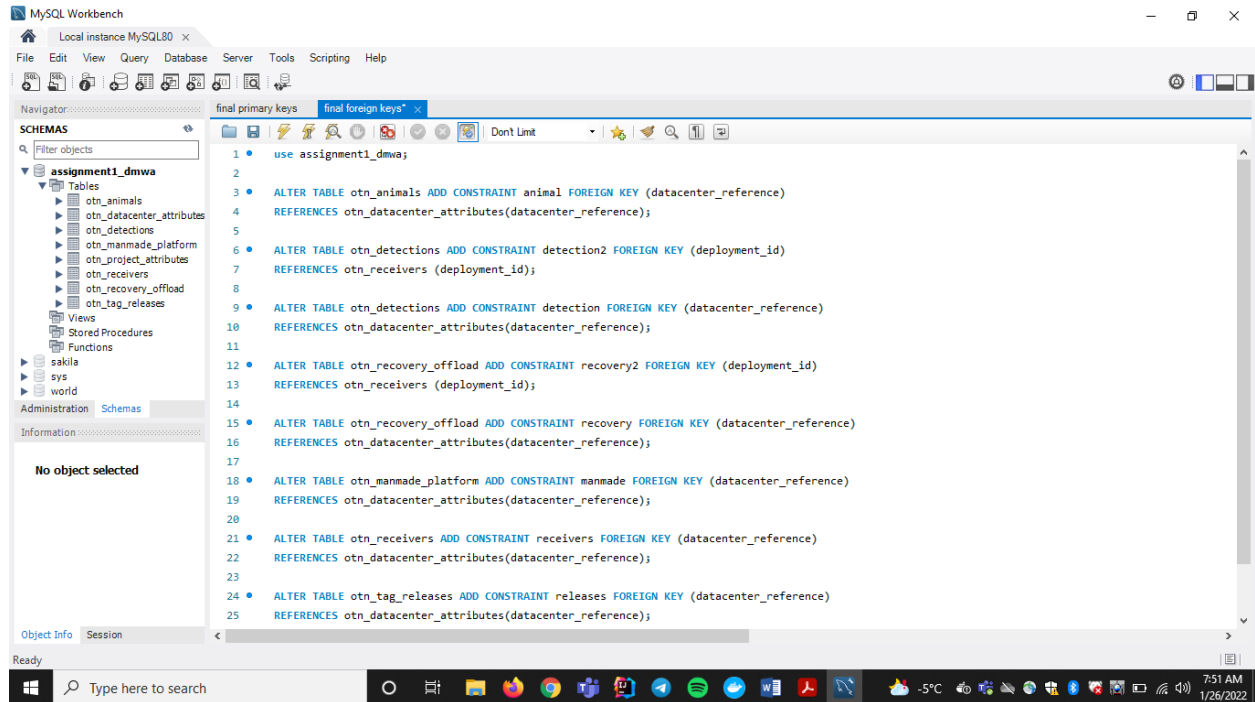
We have to create a schema in our database and then add all the files in tables.



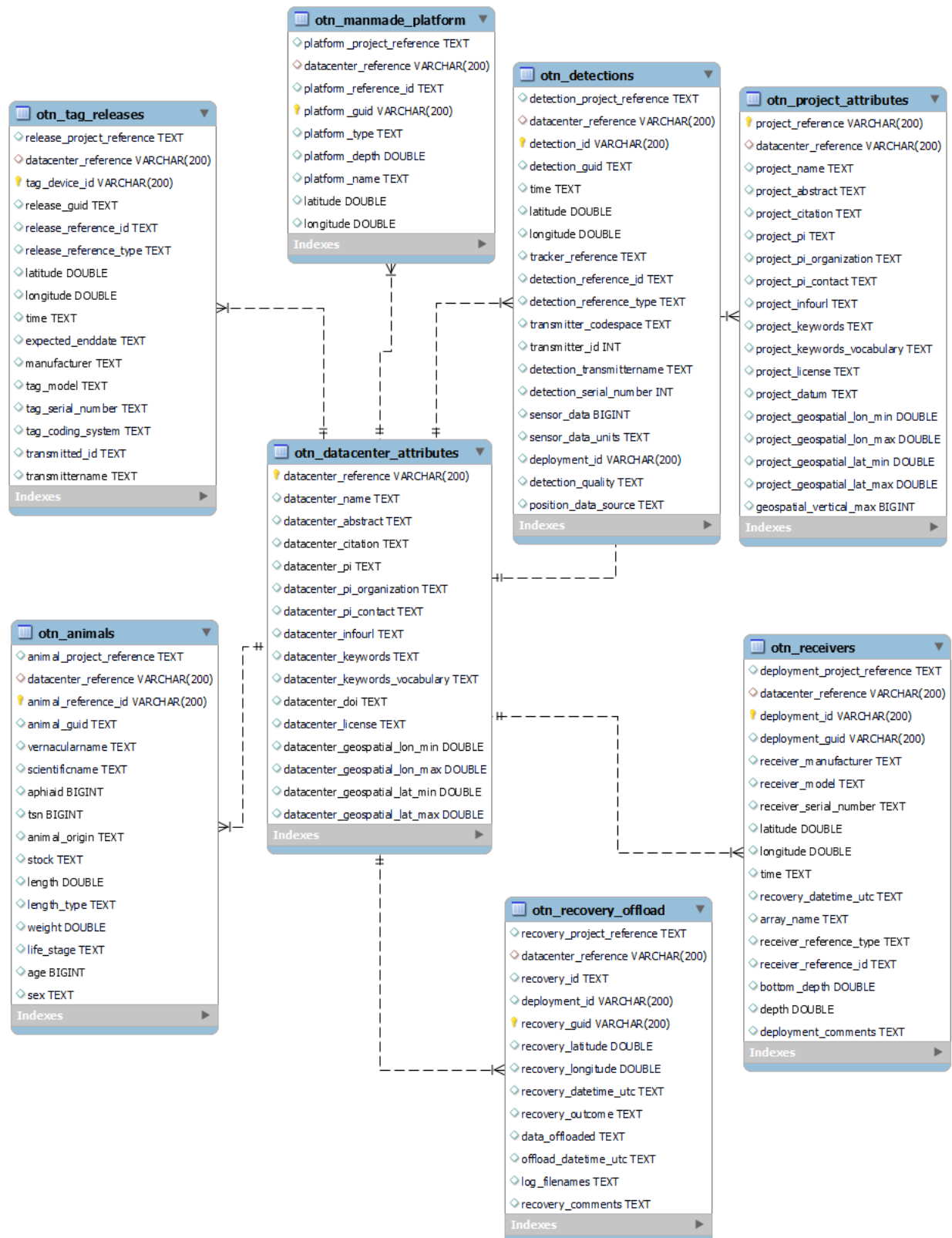
Now after importing table in to our workbench, the next step is to define primary keys for every table.



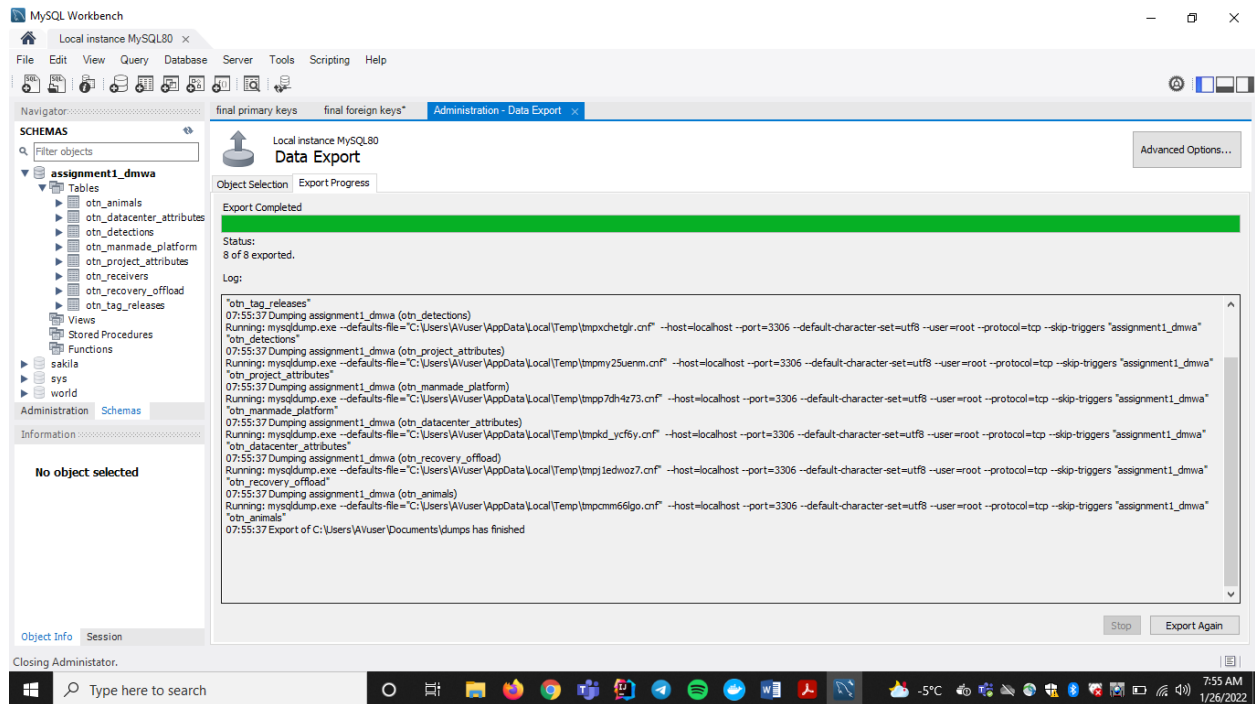
The next step is to define all the foreign keys in the tables.



The final ERD after reverse engineering is



SQL Dump of Table Structure and Values



The sql dump files have been attached to this document in the same zip file.

References:

- https://www.w3schools.com/sql/sql_foreignkey.asp