

understanding recall data

"recall\_number" vs "event\_id"

	recall_number	event_id
unique count	22,883	9099
variable definition	for specific product each # has different lots, different units, product has size 1, 2, 3)	for specific recall event (~specific brand name) 1 brand name can have different size, each size has different recall_number

A	B	C	D	E	F	G	H
event_id	recall_number	recall_initiation_date	center_classi	report_date	product_description	product_quantity	
85629	Z-1989-2020	20200430	20200514	20200520	Dover 100% Silicone Premium Urine Meter Tray 400mL, 16 Fr/Ch (5.3 mm), 5 mL Foley Catheter Pre-connected to 2000 mL Dr	8162 units	
85629	Z-1988-2020	20200430	20200514	20200520	Dover 100% Silicone Premium Urine Meter Foley Tray, 400mL, 14 Fr/Ch (4.7 mm), 5 mL Catheter Pre-connected to 2000 mL Dr	3470 units	
85629	Z-1985-2020	20200430	20200514	20200520	Dover 100% Silicone Silver-Hydrogel Coated Foley Catheter with Temperature Sensor 16FR 5CCCC2W, Product Code: 50516IC	480 units	
85629	Z-1993-2020	20200430	20200514	20200520	Dover 400 ml 100% Silicone Silver-Hydrogel Coated Premium Urine Meter Foley Tray,18 Fr/Ch (6.0 mm), 5 mL Catheter Pre-co	120 units	
85629	Z-1986-2020	20200430	20200514	20200520	Dover 100% Silicone Foley Catheter with Temperature 16FR 5CC2W, Product Code: 50516T - Product Usage: for Drainage of tl	3012 units	
85629	Z-1994-2020	20200430	20200514	20200520	Dover 400 ml 100% Silicone Silver-Hydrogel Coated Premium Urine Meter Foley Tray, 18 Fr/Ch (6.0 mm), 5 mL Catheter Pre-co	0	
85629	Z-1984-2020	20200430	20200514	20200520	Dover 100% Silicone Foley Catheter with Temperature Sensor 14FR 5CC2W, Product Code: 50514T - Product Usage: for Drainz	468 units	
85629	Z-1992-2020	20200430	20200514	20200520	Dover 100% Silicone Silver-Hydrogel Coated Premium Urine Meter Foley Tray 400mL, 16 Fr/Ch (5.3 mm), 5 mL Catheter Pre-co	46828 units	
85629	Z-1987-2020	20200430	20200514	20200520	Dover 100% Silicone Silver-Hydrogel Coated Foley Catheter with Temperature Sensor 18FR 5CCCC2W, Product Code: 50518IC1	0	
85629	Z-1990-2020	20200430	20200514	20200520	Dover 400mL, 100% Silicone Premium Urine Meter Foley Tray, 16 Fr/Ch (5.3 mm), 5 mL Catheter Pre-connected to 2000 mL D	23587 units	

for example: same event\_id: '85629', but different recall\_number  
In this example, specific brand\_name can be Dover Silicone device

dates

by time order	recall_init_date	center_classification_date	report_date	termination_date
---------------	------------------	----------------------------	-------------	------------------

```
In [42]: event_id_dup[event_id_dup['event_id'] == 85629].filter(col_interest)
Out[42]:
```

	recall_initiation_date	center_classification_date	report_date
0	20200430	20200514	20200520
1	20200430	20200514	20200520
2	20200430	20200514	20200520
3	20200430	20200514	20200520
4	20200430	20200514	20200520
5	20200430	20200514	20200520
6	20200430	20200514	20200520
7	20200430	20200514	20200520
8	20200430	20200514	20200520
9	20200430	20200514	20200520

### understanding 'status'

```
In [52]: data['status'].value_counts()
Out[52]:
```

Terminated	16507
Ongoing	5857
Completed	523

Name: status, dtype: int64

definition of categories in 'status' variable:

source: [https://www.fda.gov/safety/enforcement-reports/enforcement-report-information-and-definitions#weekly\\_report](https://www.fda.gov/safety/enforcement-reports/enforcement-report-information-and-definitions#weekly_report)

	terminated	ongoing	completed
count	16,507	5,857	523
definition	"A recall where FDA has determined that all reasonable efforts have been made to remove or correct the violative product in accordance with the recall strategy, and proper disposition has been made according to the degree of hazard"	"A recall which is currently in progress"	"A recall which has reached the point at which the firm has actually retrieved and impounded all outstanding product that could reasonably be expected to be recovered, or has completed all product corrections"

"gut check" to test if data following logic:

- if status == 'terminated' —> termination\_date != null ?

but here this is not the case because termination\_date missing 111 among 16,507 terminated

explain: only the 1st terminated status has termination\_date

A	B	C	D	E	F	G	H	I	J
event_id	recall_number	recall_initiat	center_classi	report_date	product_description	product_quantity	status	termination_date	
81932	Z-0831-2019	20190122	20190214	20190220	COBRA FUSION 150 Ablation System, Catalog # (REF) 700-001S	175 units	Terminated		
81932	Z-0830-2019	20190122	20190214	20190220	COBRA FUSION 150 Ablation System, Catalog # (REF) 700-001	6176 units	Terminated		
81932	Z-0829-2019	20190122	20190214	20190220	COBRA FUSION 50, Ablation System, Catalog # (REF) 700-002	1288 units	Terminated		
81932	Z-0832-2019	20190122	20190214	20190220	COBRA FUSION 150 Ablation System (International Only), Catalog # (REF) 700-001MI	765 units	Terminated		
75373	Z-0653-2017	20160922	20161122	20161130	COBRA Fusion 50 Ablation System, COBRA Fusion 150 Ablation System, and COBRA Fusion Magnetic	5,263	Terminated	20180208	

- if status == 'ongoing' —> termination\_date should be missing (check!)

\*\*\*\*\*

how to keep track of terminated status? do you need to update in real time?

Year of Recall Initiation Da..	Year of Center Classific..	Year of Report Date	Status	
2016	2019	2020	Ongoing	1
		2020	Ongoing	1
2017	2020	2020	Ongoing	2
2018	2020	2020	Completed	14
			Ongoing	14
			Terminated	2
2019	2019	2020	Completed	1
			Ongoing	44
			Terminated	1
	2020	2020	Completed	8
			Ongoing	583
2020	2020	2020	Terminated	29
			Ongoing	646
			Terminated	7

data in API update by 'report\_date'

```
In [66]: data.sort_values(by='report_date')['report_date']
```

```
Out[66]:
```

```
22181    20120620
22196    20120620
22193    20120620
22189    20120620
22188    20120620
```

```
...
```

```
71      20200527
72      20200527
73      20200527
48      20200527
763     20200527
```

```
Name: report_date, Length: 22887, dtype: int64
```

```
In [65]: data.sort_values(by='center_classification_date')['center_classification_date']
```

```
Out[65]:
```

```
22250    20120608
22294    20120608
22879    20120608
22881    20120608
22880    20120608
```

```
...
```

```
36      20200521
35      20200521
34      20200521
32      20200521
45      20200521
```

```
Name: center_classification_date, Length: 22887, dtype: int64
```

```
In [67]: data.sort_values(by='recall_initiation_date')['recall_initiation_date']
```

```
Out[67]:
```

```
22886    19301211
22885    20050701
22884    20050708
22883    20051227
22882    20060425
```

```
...
```

```
7      20200430
9      20200430
10     20200430
5      20200430
0      20200430
```

```
Name: recall_initiation_date, Length: 22887, dtype: int64
```