VAERS

https://github.com/jgarza9788/vaers

Disclaimer

THIS IS NOT MEDICAL ADVICE
if you want medical advice please go see your doctor.

please also read the disclaimer on the VAERS data on their website.
VAERS website

get started

1. get the data

download the data from VAERS website
link to VAERS website

make sure to download the 2020-2022 data 2022VAERSVAX.csv
2022VAERSSYMPTOMS.csv
2022VAERSDATA.csv
2021VAERSVAX.csv
2021VAERSSYMPTOMS.csv
2021VAERSDATA.csv
2020VAERSVAX.csv
2020VAERSSYMPTOMS.csv
2020VAERSDATA.csv

2. move and unzip

move the AllVAERSDataCSVS.zip to VAERS folder and unzip it

3. check the files

you should have at least 6 files

```
.../VAERS/AllVAERSDataCSVS/
    2022VAERSVAX.csv
    2022VAERSSYMPTOMS.csv
    2022VAERSDATA.csv
    ...
    2020VAERSVAX.csv
```

2020VAERSSYMPTOMS.csv
2020VAERSDATA.csv

THE CODE!

setup/imports

directory variables

DIR (String) : the current directory this file is in
DATAPATH (String) : the location where the data files are stored

```
DATAPATH:  .\AllVAERSDataCSVS
```

Saving and Loading JSON files

processing files

the functions flow will process the csv files into one all_data.json file

get data

the below will get data from the files or from all_data.json.

> please note i am only using 25% of the data below, but you can see the .pdf for how this will look with all the data, or edit and run the notebook on your own.

df (DataFrame) : contains all the datafrom VAERS files listed below

- .../VAERS/AllVAERSDataCSVS/
  - 2022VAERSVAX.csv
  - 2022VAERSSYMPTOMS.csv
  - 2022VAERSDATA.csv
  - ...
  - 2020VAERSVAX.csv
  - 2020VAERSSYMPTOMS.csv
  - 2020VAERSDATA.csv

```
loading all_data.json (15sec-30sec)

loaded 946,527 records/rows

columns:
 ['VAERS_ID', 'RECVDATE', 'STATE', 'AGE_YRS', 'CAGE_YR', 'CAGE_MO', 'SEX', 'RPT_DATE', 'SYMPTOM_TE
XT', 'DIED', 'DATEDIED', 'L_THREAT', 'ER_VISIT', 'HOSPITAL', 'HOSPDAYS', 'X_STAY', 'DISABLE', 'REC
```

```
OVD', 'VAX_DATE', 'ONSET_DATE', 'NUMDAYS', 'LAB_DATA', 'V_ADMINBY', 'V_FUNDBY', 'OTHER_MEDS', 'CUR
_ILL', 'HISTORY', 'PRIOR_VAX', 'SPLTTYPE', 'FORM_VERS', 'TODAYS_DATE', 'BIRTH_DEFECT', 'OFC_VISI
T', 'ER_ED_VISIT', 'ALLERGIES', 'VAX_TYPE', 'VAX_MANU', 'VAX_LOT', 'VAX_DOSE_SERIES', 'VAX_ROUTE',
'VAX_SITE', 'VAX_NAME', 'SYMPTOMS', 'COVID_VAX']


df.head(10):
```

	VAERS_ID	RECVDATE	STATE	AGE_YRS	CAGE_YR	CAGE_MO	SEX	RPT_DATE	SYMPTOM_TEXT	DIED	DATEDIED	L_THRE
0	902418	12/15/2020	NJ	56.00	56.00	NaN	F	NaN	Patient experienced mild numbness traveling from injection site up and down arm that subsided over 20 minutes.	N	NaN	
1	902440	12/15/2020	AZ	35.00	35.00	NaN	F	NaN	C/O Headache	N	NaN	

gets a list of symptoms

df_symptoms (DataFrame) : a list of all the symtoms and the counts of each all_symptoms (list): a list of all the symtoms and the counts of each

note:
symptoms might be medical jargon or plain english
i.e. "RASH","ERYTHEMA", and "ITCHY RED SKIN"
would be reported as different items (for now)

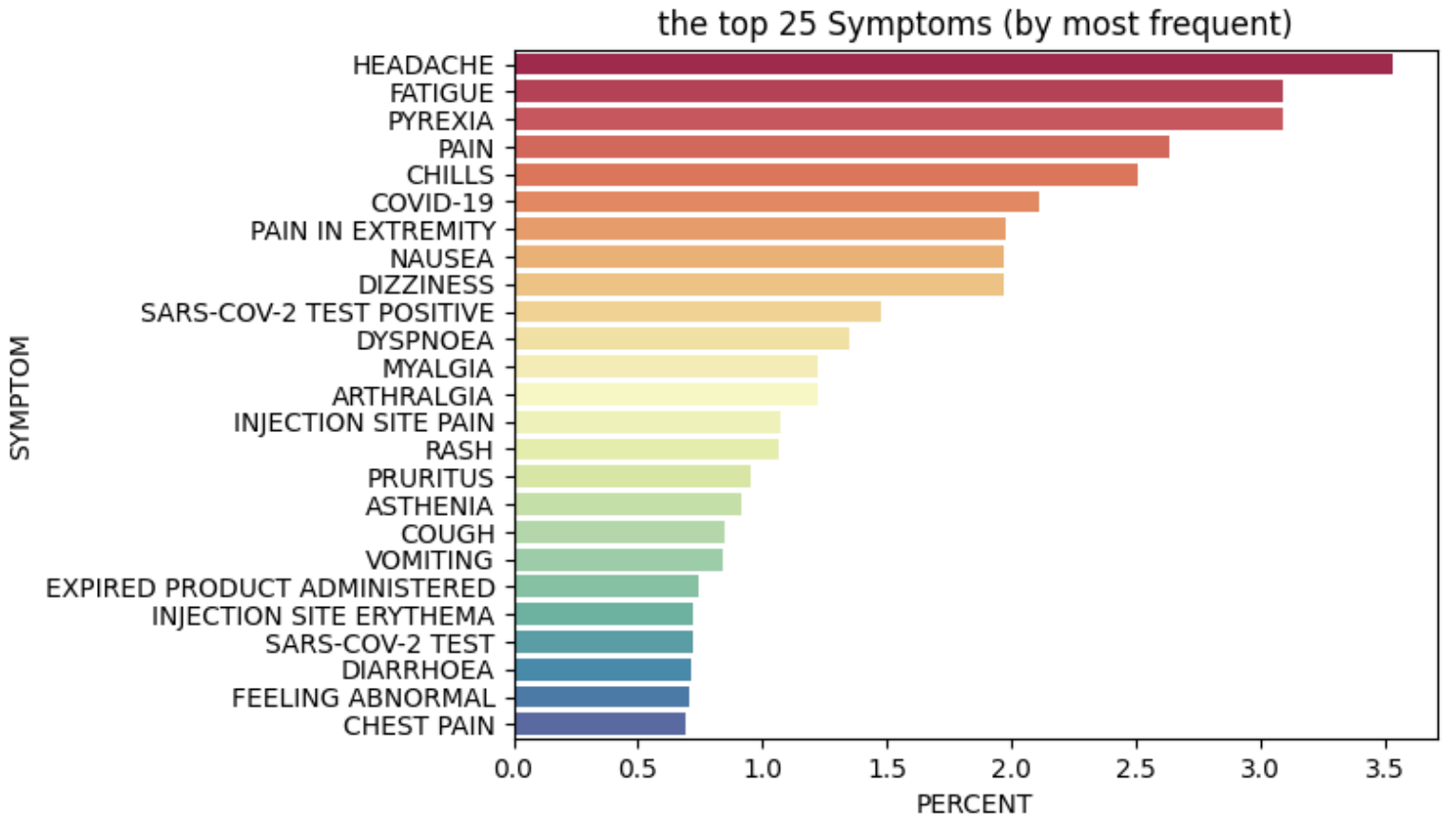
note:
the counts/percentages below are of the symptoms.
and one adverse reaction can have multiple symptoms.

```
saved:  .\AllVAERSDataCSVS\symptoms.csv
below are the top 25 symptoms
```

	SYMPTOM	COUNT	PERCENT
0	HEADACHE	144140	3.53
1	FATIGUE	126025	3.09
2	PYREXIA	126019	3.09
3	PAIN	107523	2.63
4	CHILLS	102390	2.51
5	COVID-19	86311	2.11
6	PAIN IN EXTREMITY	80762	1.98
7	NAUSEA	80440	1.97
8	DIZZINESS	80340	1.97

9	SARS-COV-2 TEST POSITIVE	60134	1.47
10	DYSPNOEA	54926	1.34
11	MYALGIA	50024	1.22
12	ARTHRALGIA	49897	1.22
13	INJECTION SITE PAIN	43769	1.07
14	RASH	43330	1.06
15	PRURITUS	39047	0.96
16	ASTHENIA	37502	0.92
17	COUGH	34580	0.85
18	VOMITING	34188	0.84
19	EXPIRED PRODUCT ADMINISTERED	30502	0.75
20	INJECTION SITE ERYTHEMA	29368	0.72
21	SARS-COV-2 TEST	29300	0.72
22	DIARRHOEA	29093	0.71
23	FEELING ABNORMAL	28701	0.70
24	CHEST PAIN	28067	0.69

```
[Text(0.5, 1.0, 'the top 25 Symptoms (by most frequent)')]
```



break down functions...

these functions will help me breakdown the data

break_down_columns

break_down_buckets

Additional Numbers

~~these numbers are **not** part of the vaers data,
however they are important to analyzing the data~~

~~vaxx (int) : the number of vaccinated (1 or more shots) US citizens according to a quick google search (on 8/3/2021)~~

~~**google no longer shows partially vaxxed** so we are using the **Total doses given** number
"Total doses given" shows the number of vaccine doses given to people. Since some vaccines require more than 1 dose, the number of fully vaccinated people is likely lower. "People fully vaccinated" shows how many people have received the full amount of doses for the COVID-19 vaccine. ~~

~~This number was updated on 06/28/2022~~

~~full_vaxx (int) : the number of vaccinated (2 or more shots) US citizens according to a quick google search (on 06/28/2022)~~

the numbers were taken from https://usafacts.org/visualizations/covid-vaccine-tracker-states on 11/18/2022

Analyzing the df (all_data.json)

this is just a generic break down of the data

df_death (DataFrame) : adverse reactions that resulted in a death

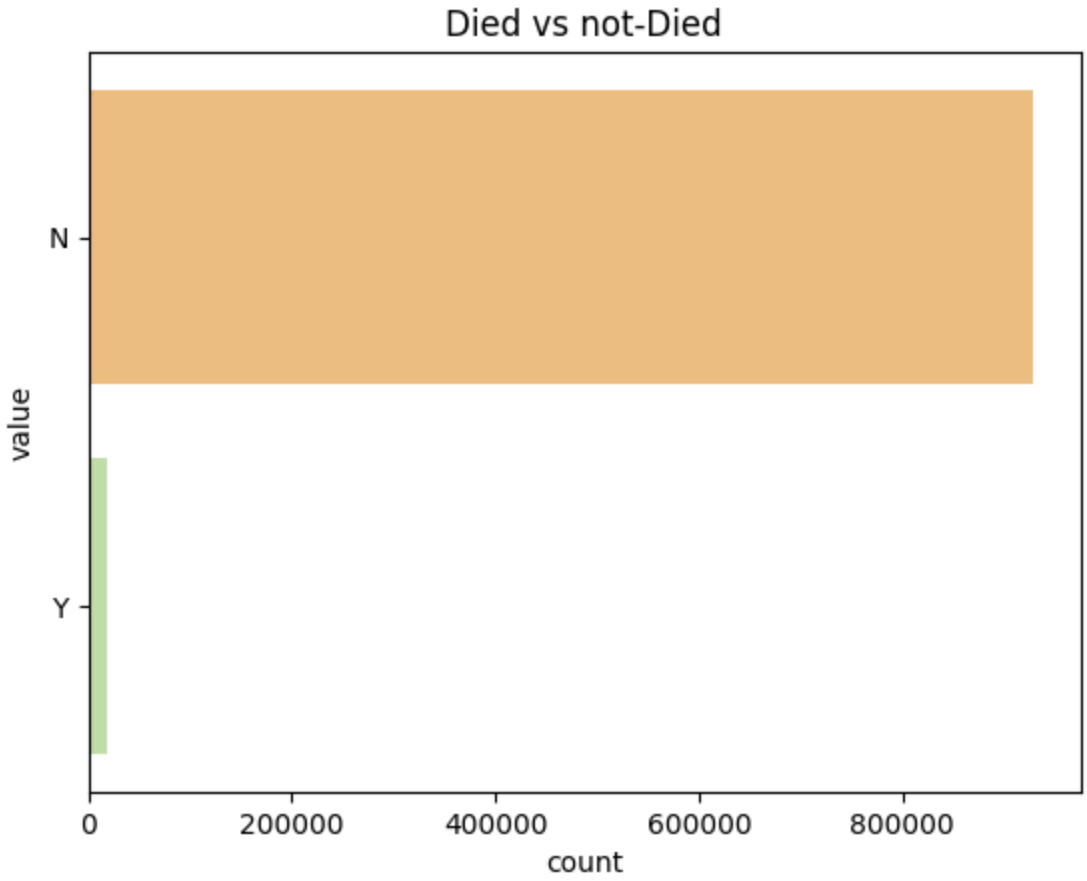
DIED column

```
break down of DIED
```

	column	value	count	percent
0	DIED	N	928180	98.06
1	DIED	Y	18347	1.94

```
-------------------------

[Text(0.5, 1.0, 'Died vs not-Died')]
```



ER_VISIT column

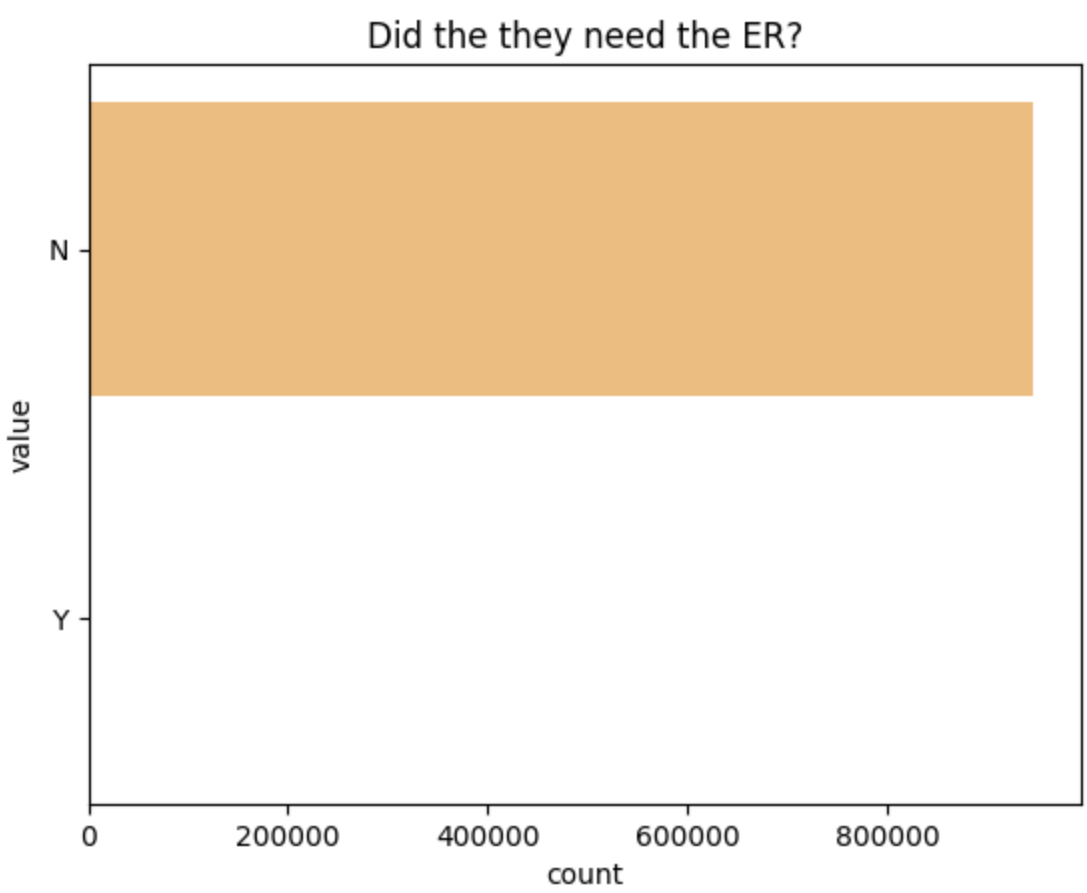
```
did the adverse reaction result in an ER Visit

break down of ER_VISIT
```

	column	value	count	percent
0	ER_VISIT	N	946399	99.99
1	ER_VISIT	Y	128	0.01

```
-------------------------

[Text(0.5, 1.0, 'Did the they need the ER?')]
```



L_THREAT column

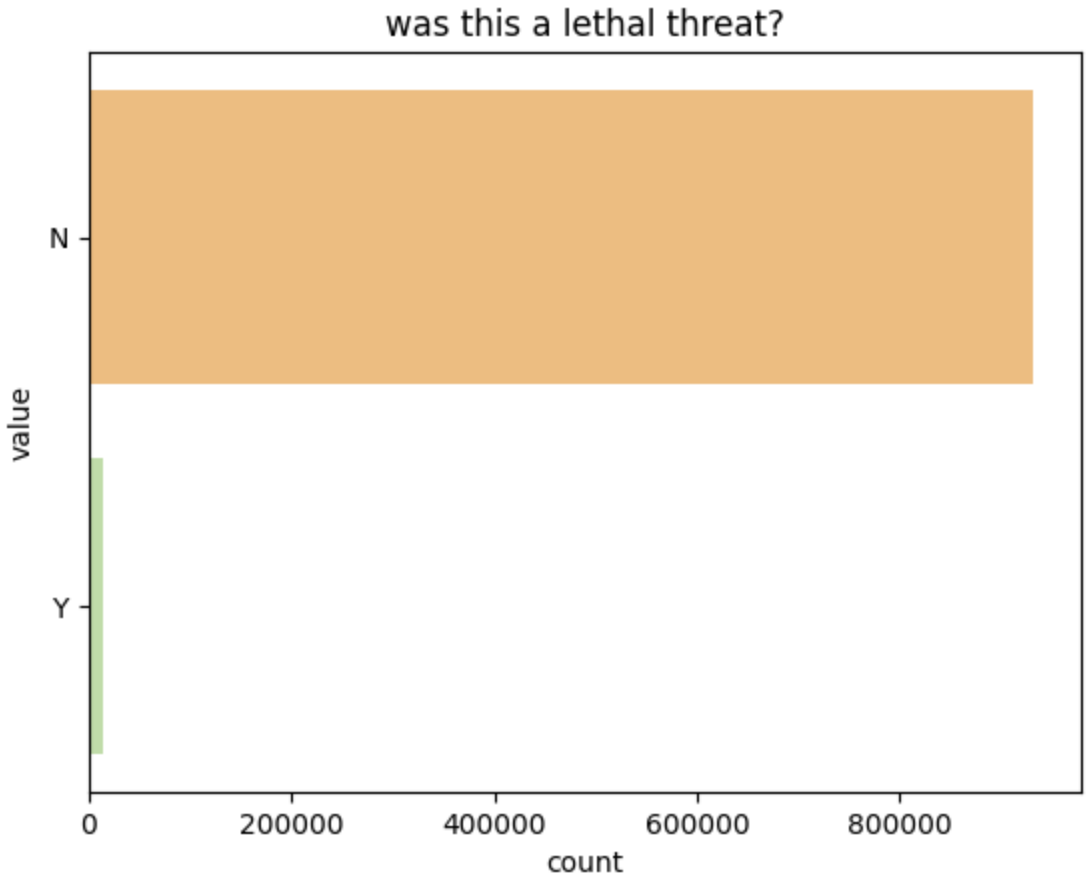
```
Life Threatening/Leathal Threat

break down of L_THREAT
```

	column	value	count	percent
0	L_THREAT	N	931317	98.39
1	L_THREAT	Y	15210	1.61

```
-------------------------

[Text(0.5, 1.0, 'was this a lethal threat?')]
```



RECOVD column

```
did the patient recover?

break down of RECOVD
```

	column	value	count	percent
0	RECOVD	Y	71201	30.09
1	RECOVD	U	56518	23.88
2	RECOVD	N	79462	33.58
3	RECOVD	nan	29450	12.45

```
-------------------------
```

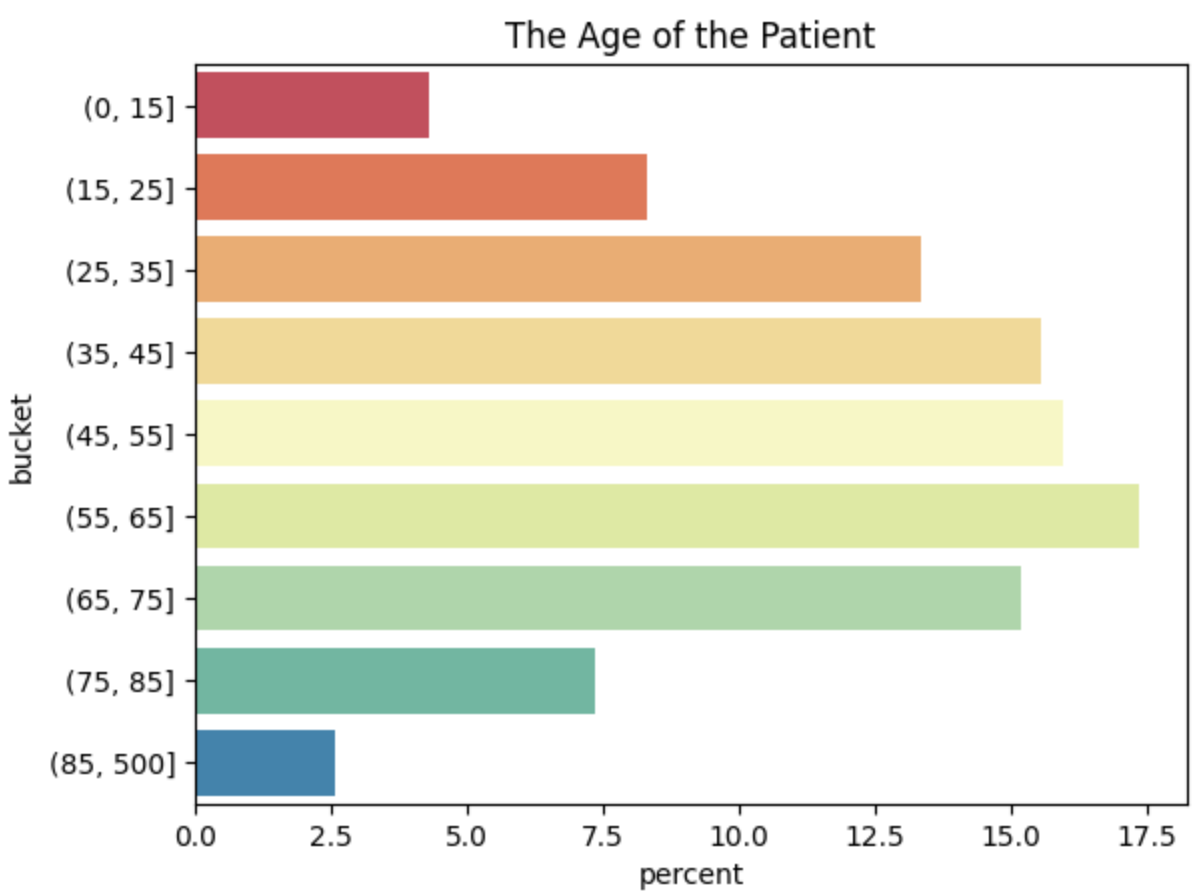
the Age of the patient

```
the Age of the patient

column:  AGE_YRS
buckets:  [0, 15, 25, 35, 45, 55, 65, 75, 85, 500]
```

bucket	AGE_YRS	percent
(0, 15]	36851	4.32
(15, 25]	70859	8.30
(25, 35]	114041	13.36
(35, 45]	132833	15.56
(45, 55]	136185	15.95
(55, 65]	148283	17.37
(65, 75]	129763	15.20
(75, 85]	62625	7.34
(85, 500]	22119	2.59

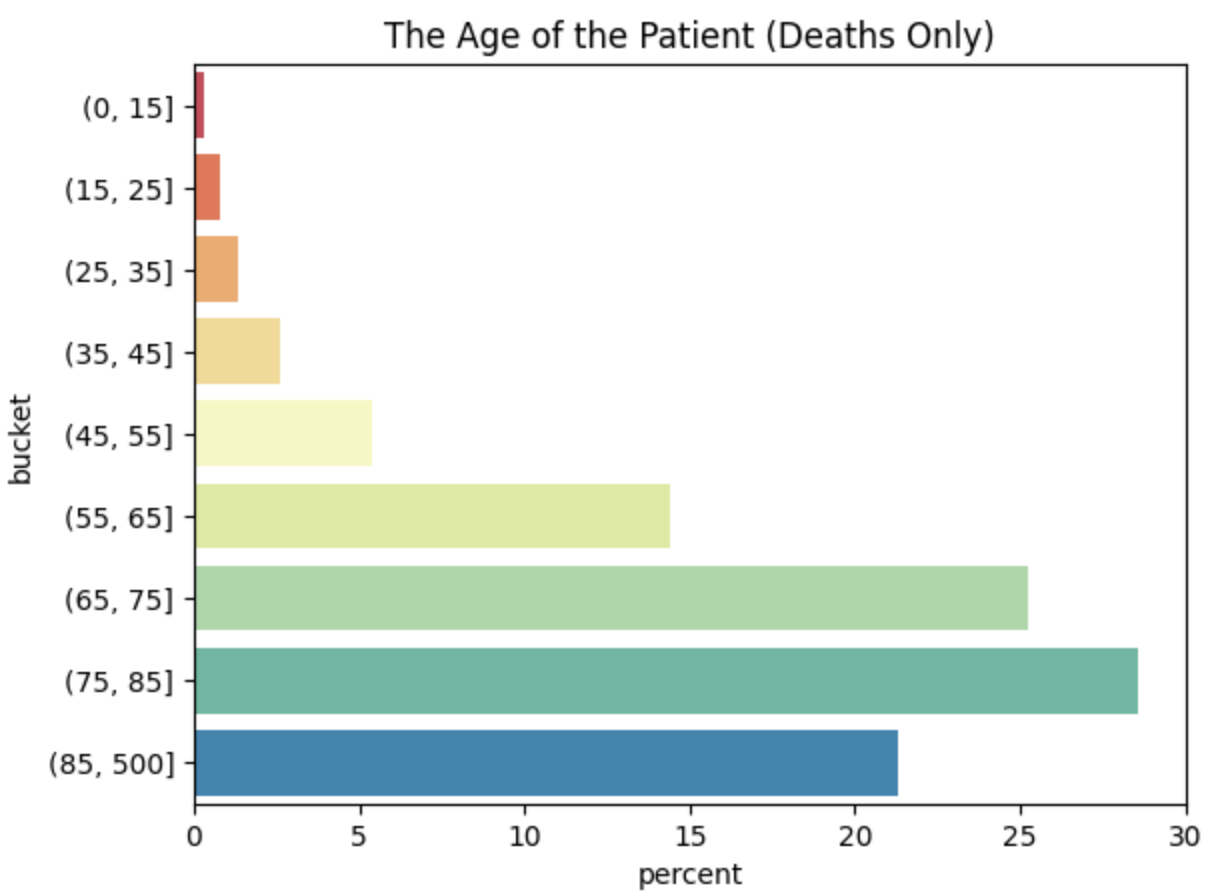
```
--------------------------
```




```
 ***deaths only***
column:  AGE_YRS
buckets:  [0, 15, 25, 35, 45, 55, 65, 75, 85, 500]
```

bucket	AGE_YRS	percent
(0, 15]	49	0.29
(15, 25]	139	0.82
(25, 35]	230	1.36
(35, 45]	440	2.61
(45, 55]	913	5.42
(55, 65]	2428	14.41
(65, 75]	4252	25.23
(75, 85]	4817	28.58
(85, 500]	3586	21.28

```
-------------------------
```



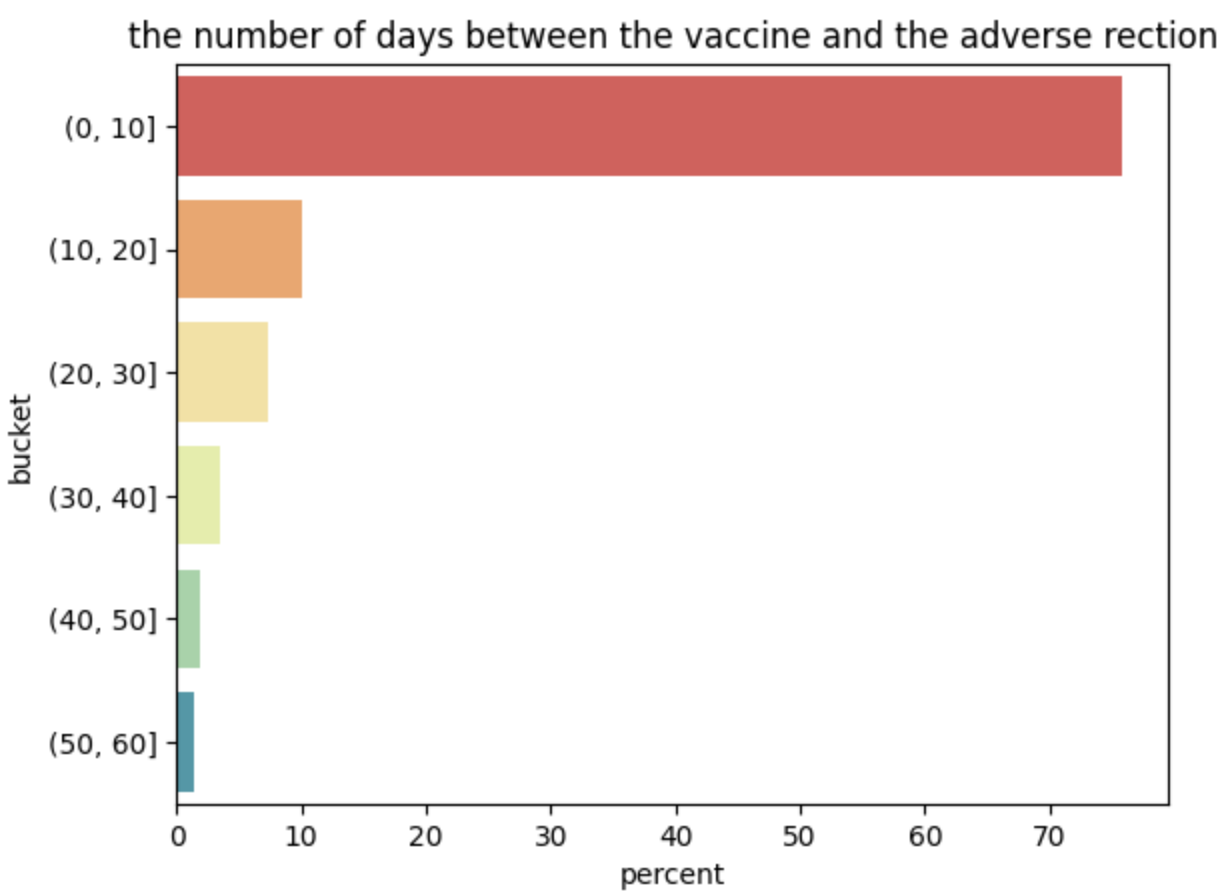
the number of days between the vaccine and the adverse rection

```
the number of days between the vaccine and the adverse rection

column:  NUMDAYS
buckets:  [0, 10, 20, 30, 40, 50, 60]
```

bucket	NUMDAYS	percent
(0, 10]	288294	75.73
(10, 20]	38190	10.03
(20, 30]	28138	7.39
(30, 40]	13633	3.58
(40, 50]	6921	1.82
(50, 60]	5492	1.44

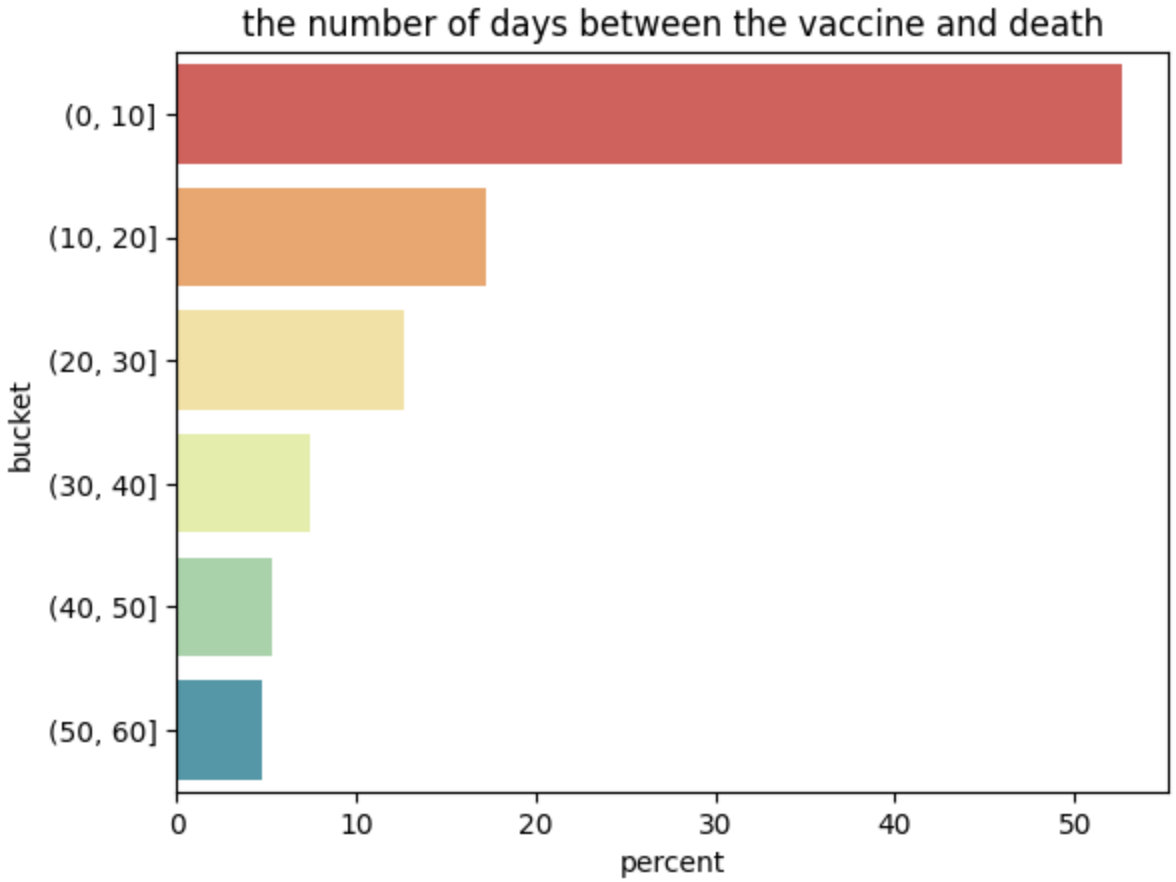
```
------------------------
```




```
 ***deaths only***
column:  NUMDAYS
buckets:  [0, 10, 20, 30, 40, 50, 60]
```

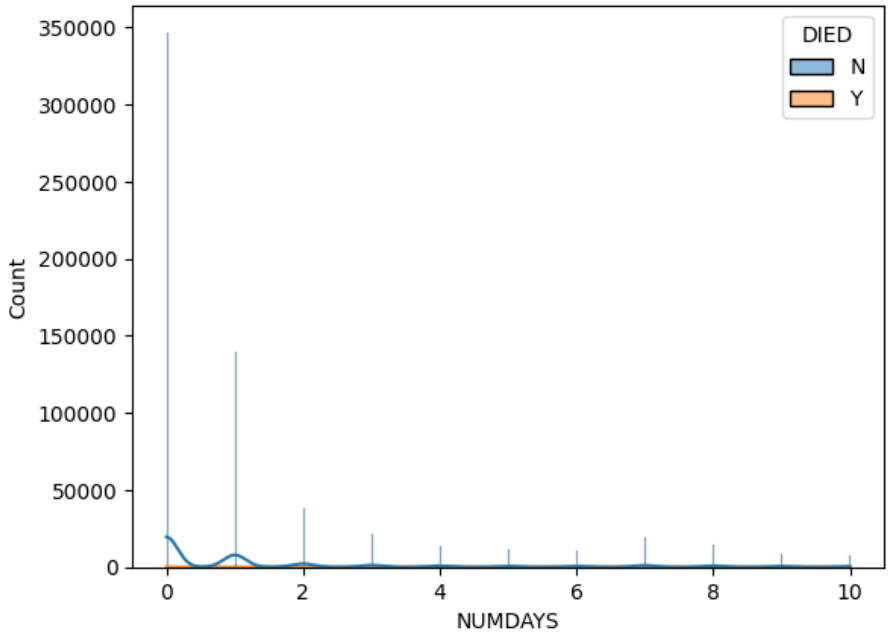
bucket	NUMDAYS	percent
(0, 10]	3752	52.59
(10, 20]	1233	17.28
(20, 30]	900	12.62
(30, 40]	532	7.46
(40, 50]	379	5.31
(50, 60]	338	4.74

```
------------------------
```




```
[Text(0.5, 1.0, 'showing number of days between vaccine and adverse reaction ( only 10 days or les
s, broke out death into other color)')]
```

showing number of days between vaccine and adverse reaction (only 10 days or less, broke out death into other color)



break down of the VAX_NAME column

```
break down of the VAX_NAME column

break down of VAX_NAME
```

	column	value	count	percent
0	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH))	437667	46.24

1	VAX_NAME	COVID19 (COVID19 (UNKNOWN))	4579	0.48
2	VAX_NAME	COVID19 (COVID19 (MODERNA))	423732	44.77
3	VAX_NAME	COVID19 (COVID19 (JANSSEN))	71507	7.55
4	VAX_NAME	COVID19 (COVID19 (NOVAVAX))	154	0.02
5	VAX_NAME	COVID19 (COVID19 (MODERNA BIVALENT))	4025	0.43
6	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH BIVALENT))	4863	0.51

```
***deaths only***

break down of VAX_NAME
```

	column	value	count	percent
0	VAX_NAME	COVID19 (COVID19 (MODERNA))	7783	42.42
1	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH))	8608	46.92
2	VAX_NAME	COVID19 (COVID19 (UNKNOWN))	64	0.35
3	VAX_NAME	COVID19 (COVID19 (JANSSEN))	1825	9.95
4	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH BIVALENT))	46	0.25
5	VAX_NAME	COVID19 (COVID19 (MODERNA BIVALENT))	21	0.11

```
calculating the death ratio ( death_count / count )
```

	column	value	count	death_count	death_ratio	death_percent
0	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH))	437667	8608	0.02	1.97
1	VAX_NAME	COVID19 (COVID19 (UNKNOWN))	4579	64	0.01	1.40
2	VAX_NAME	COVID19 (COVID19 (MODERNA))	423732	7783	0.02	1.84
3	VAX_NAME	COVID19 (COVID19 (JANSSEN))	71507	1825	0.03	2.55
4	VAX_NAME	COVID19 (COVID19 (MODERNA BIVALENT))	4025	21	0.01	0.52
5	VAX_NAME	COVID19 (COVID19 (PFIZER-BIONTECH BIVALENT))	4863	46	0.01	0.95

```
-------------------------
```

Approximating actual adverse reaction numbers

VAERS only contains reported data and
'*...fewer than 1% of vaccine adverse events are reported.*'
Source: https://digital.ahrq.gov/sites/default/files/docs/publication/r18hs017045-lazarus-final-report-2011.pdf (page 6)

we will me multiplying the counts by 80 and 120,
in order to get an approzimate min and max of what the numbers might actually be.

```
( approx adverse reactions / vaxxed [1 or more shots] ) * 100
 min: ( 75,722,160 / 595,828,629 ) * 100
 max: ( 113,583,240 / 595,828,629 ) * 100
 12.71 % - 19.06 %

( approx adverse deaths / vaxxed [1 or more shots] ) * 100
 min: ( 1,467,760 / 595,828,629 ) * 100
```

```
 max: ( 2,201,640 / 595,828,629 ) * 100
 0.25 % - 0.37 %

( approx no recovery / vaxxed [1 or more shots] ) * 100
 min: ( 25,463,760 / 595,828,629 ) * 100
 max: ( 38,195,640 / 595,828,629 ) * 100
 4.27 % - 6.41 %

( approx no recovery + (50% of unknowns) / vaxxed [1 or more shots] ) * 100
 min: ( 34,511,440.0 / 595,828,629 ) * 100
 max: ( 51,767,160.0 / 595,828,629 ) * 100
 5.79 % - 8.69 %
```

Women's Reproductive Symptoms

Why? I have women in my life that were curious about this.

WRS_list (list) : a list of symptoms that effect or could cause effects to a women's reproductive system
df_WRS (DataFrame) : a dataframe that contains VAERS events that have at least 1 of the WRS symptoms

```
this could take between 20sec-60sec
df_WRS.head(5)
```

	VAERS_ID	RECVDATE	STATE	AGE_YRS	CAGE_YR	CAGE_MO	SEX	RPT_DATE	SYMPTOM_TEXT	DIED	DATEDIED	L_THRE
0	902796	12/16/2020	TX	34.00	34.00	NaN	F	NaN	I inserted my NuvaRing birth control on 12/14/2020. I have the Covid19 vaccine on 12/15/2020 at about 8:30am. The injection site was mildly sore and that continued into the next day. I woke up on 12/16/2020 feeling a little off and it progressed throughout the day. I felt a headache that I knew was turning into a migraine, threw up a few times (this is normal for me when I get migraines), and started having hot flashes. My temperature never went above 98. After sleeping a few hours and taking some Excedrin	N	NaN	

									migraine I was feeling much better, though still a little sickly, by 4pm. I really think this is due to my migraines (which I get roughly once a month) even though it was a day later than normal.		
1	903202	12/17/2020	NE	41.00	41.00	NaN	F	NaN	Severe dizzy spell about 5-10 following injection, helped to the floor, this lasted approximately 30-40 minutes after start of symptom. Hot flashes and visual disturbance lasting into following day.	N	NaN
2	903247	12/17/2020	AR	42.00	42.00	NaN	F	NaN	I had arm pain. Yesterday I had a hot flash with throwing up and diarrhea.	N	NaN
3	903329	12/18/2020	PA	46.00	46.00	NaN	F	NaN	hot flushing feeling, light headed, legs heavy gave patient a chair to sit and candy symptoms resolved by 11:36	N	NaN
4	903345	12/18/2020	VA	37.00	37.00	NaN	F	NaN	Right arm soreness, headache, hot flashes.	N	NaN

What percent of women experienced WRS during their adverse reactions?

(women experiencing reproductive symptoms ➗ number of women in VAERS data) * 100

WRS_ratio (float) : count WRS / count VAERS

```
3.41 %
VAERS records of women experiencing reproductive symptoms have been saved.
```

```
saved:  .\AllVAERSDataCSVS\WRS.csv
```

Approximate the number of WRS in reality

since '~63% of the people who are vaccinated are women' source: https://www.statista.com/statistics/1212103/share-of-persons-initiating-covid-vaccinations-by-gender-us-first-month/

we will be multiplying the total number of people vaccinated (1 or more shots) by 0.63 to get the count of women vaccinated.

women_vaxx (float) : an approximate number of women who have had 1 or more vaccine shot.

> note: 0.63 is an estimate, and the actual could be somewhere between 0.50 and 0.70

```
375372036.27
```

approximate real WRS

if we assume that the VAERS data is a random sample (or close to it)
then the ratio of WRS systems should be the same...

and thus we can get an approximate number of women that would be experiencing reproductive symptoms by multiplying the number of vaxxed women by the ratio

WRS (float) : the approximate number of actual women experiencing reproductive symptoms min_WRS (float) : WRS *0.80 max_WRS (float) : WRS* 1.20

```
total vaxxed (1 or more)           ---  595,828,629.00
women vaxxed ~0.63%                ---  375,372,036.27
repro sympt / women count          ---  0.0341
(repro sympt / women count) * 100  ---  3.41 %
women w/ repro symptoms            ---  12,814,887.41
min women w/ repro symptoms        ---  10,251,909.93
max women w/ repro symptoms        ---  15,377,864.89
```

Where do the WRS (Women reproductive symptoms) rank against with the other symptoms?

	index	symptoms	count	percent
0	128	HEAVY MENSTRUAL BLEEDING	5,670.00	47.05
1	170	MENSTRUATION IRREGULAR	4,329.00	35.92
2	172	HOT FLUSH	4,243.00	35.21
3	194	MENSTRUAL DISORDER	3,258.00	27.03
4	246	DYSMENORRHOEA	2,384.00	19.78
5	251	BREAST PAIN	2,311.00	19.18
6	277	VAGINAL HAEMORRHAGE	1,979.00	16.42
7	305	INTERMENSTRUAL BLEEDING	1,678.00	13.92
8	474	BREAST SWELLING	905.00	7.51

9	488	POLYMENORRHOEA	854.00	7.09
10	531	OLIGOMENORRHOEA	763.00	6.33
11	591	BREAST TENDERNESS	641.00	5.32
12	815	MATERNAL EXPOSURE BEFORE PREGNANCY	392.00	3.25
13	1245	MENOPAUSE	192.00	1.59
14	1265	OVARIAN CYST	188.00	1.56
15	1468	UTERINE LEIOMYOMA	145.00	1.20
16	1631	ABNORMAL UTERINE BLEEDING	120.00	1.00
17	2998	BACTERIAL VAGINOSIS	34.00	0.28
18	3238	ANOVULATORY CYCLE	29.00	0.24
19	3466	BIOPSY UTERUS	24.00	0.20
20	3614	OVULATION DELAYED	22.00	0.18
21	5157	GYNAECOLOGICAL EXAMINATION ABNORMAL	9.00	0.07