

°Número de casos confirmados por cada mes

°Promedio de casos confirmados por mes

°Dia del mes en el que se presentó el mayor número de casos confirmado

°Porcentaje de la población afectada por mes

°Función personalizada

```
In [1]: import pandas as pd
        from datetime import datetime
```

```
In [2]: df = pd.read_csv("Casos_Diarios_Estado_Nacional_Confirmados_20230308.csv")
        df.head()
```

```
Out[2]:
```

	cve_ent	poblacion	nombre	26-02-2020	27-02-2020	28-02-2020	29-02-2020	01-03-2020	02-03-2020	03-03-2020	...	26-02-2023
0	1	1434635	AGUASCALIENTES	0	0	0	0	0	0	0	...	4
1	2	3634868	BAJA CALIFORNIA	0	0	0	0	0	0	0	...	9
2	3	804708	BAJA CALIFORNIA SUR	0	0	0	0	0	0	0	...	10
3	4	1000617	CAMPECHE	0	0	0	0	0	0	0	...	1
4	7	5730367	CHIAPAS	0	0	0	1	0	0	0	...	0

5 rows × 1109 columns

```
In [3]: df.columns
```

```
Out[3]: Index(['cve_ent', 'poblacion', 'nombre', '26-02-2020', '27-02-2020', '28-02-2020', '29-02-2020', '01-03-2020', '02-03-2020', '03-03-2020', '...', '26-02-2023', '27-02-2023', '28-02-2023', '01-03-2023', '02-03-2023', '03-03-2023', '04-03-2023', '05-03-2023', '06-03-2023', '07-03-2023'],
              dtype='object', length=1109)
```

```
In [4]: df["26-02-2020"] + df["27-02-2020"]
```

```
Out[4]: 0      0
        1      0
        2      0
        3      0
        4      0
        5      0
        6      2
        7      0
        8      0
        9      0
       10      0
       11      0
       12      1
       13      0
       14      1
       15      0
       16      0
       17      0
       18      0
       19      0
       20      0
       21      0
       22      0
       23      0
       24      0
       25      0
       26      0
       27      0
       28      0
       29      0
       30      0
       31      0
       32      4
dtype: int64
```

```
In [5]: resultado = None
        for col_name in list(df.columns):
            if '2020' in col_name:
                if resultado is None:
                    resultado = df[col_name]
                else:
                    resultado = resultado + df[col_name]

        df["2020"] = resultado
```

# Número de casos confirmados por cada mes

```
In [6]: import datetime
from dateutil.relativedelta import relativedelta

first_month = datetime.datetime(2020,2,1)
last_month = datetime.datetime(2023,3,1)
months = {}

while first_month<= last_month:
    m = first_month.strftime("%m-%Y")
    months[m] = 0
    first_month = first_month + relativedelta(months = 1)

print(months)

{'02-2020': 0, '03-2020': 0, '04-2020': 0, '05-2020': 0, '06-2020': 0, '07-2020': 0, '08-2020': 0, '09-2020': 0, '10-2020': 0, '11-2020': 0, '12-2020': 0, '01-2021': 0, '02-2021': 0, '03-2021': 0, '04-2021': 0, '05-2021': 0, '06-2021': 0, '07-2021': 0, '08-2021': 0, '09-2021': 0, '10-2021': 0, '11-2021': 0, '12-2021': 0, '01-2022': 0, '02-2022': 0, '03-2022': 0, '04-2022': 0, '05-2022': 0, '06-2022': 0, '07-2022': 0, '08-2022': 0, '09-2022': 0, '10-2022': 0, '11-2022': 0, '12-2022': 0, '01-2023': 0, '02-2023': 0, '03-2023': 0}
```

```
In [7]: columns = list(df.columns)
columns.remove("cve_ent")
columns.remove("poblacion")
columns.remove("nombre")
columns.remove("2020")
# print(columns)
```

```
In [8]: for column in columns:
    month = column[3:]
    months[month] = months[month] + df[column]
```

```
In [9]: casos_por_mes_df = pd.DataFrame(months)
casos_por_mes_df ["nombre"] =df["nombre"]
casos_por_mes_df[["nombre","02-2020","03-2020","12-2020"]]
```

```
Out[9]:
```

	nombre	02-2020	03-2020	12-2020
0	AGUASCALIENTES	0	54	2815
1	BAJA CALIFORNIA	0	187	8061

2	BAJA CALIFORNIA SUR	0	42	3382
3	CAMPECHE	0	8	478
4	CHIAPAS	1	16	612
5	CHIHUAHUA	0	44	2680
6	DISTRITO FEDERAL	2	879	129018
7	COAHUILA	1	86	7792
8	COLIMA	0	5	716
9	DURANGO	0	15	3649
10	GUANAJUATO	0	120	19866
11	GUERRERO	0	34	2598
12	HIDALGO	1	40	5392
13	JALISCO	0	126	10658
14	MEXICO	3	518	38377
15	MICHOACAN	0	30	5157
16	MORELOS	0	22	2651
17	NAYARIT	0	12	678
18	NUEVO LEON	0	103	15556
19	OAXACA	0	29	4876
20	PUEBLA	0	168	8914
21	QUERETARO	0	38	11025
22	QUINTANA ROO	0	113	1590
23	SAN LUIS POTOSI	0	44	5707
24	SINALOA	0	88	2289
25	SONORA	0	37	10622
26	TABASCO	0	75	7072
27	TAMAULIPAS	0	24	5163
28	TLAXCALA	0	12	2213
29	VERACRUZ	0	53	3662
30	YUCATAN	0	75	2645
31	ZACATECAS	0	25	4082
32	Nacional	8	3122	329996

In [10]: `print(months)`

```
{'02-2020': 0      0
1      0
2      0
3      0
4      1
5      0
6      2
7      1
8      0
9      0
10     0
11     0
12     1
13     0
14     3
15     0
16     0
17     0
18     0
19     0
```

In [11]: `pd.DataFrame(months)`

Out[11]:

	02- 2020	03- 2020	04- 2020	05- 2020	06- 2020	07- 2020	08- 2020	09- 2020	10- 2020	11- 2020	...	06- 2022
0	0	54	261	670	1510	1812	1680	1649	2834	4714	...	3074
1	0	187	2411	3906	4486	4385	3202	2813	3118	4865	...	6483
2	0	42	312	339	1055	3059	3080	2523	2349	2027	...	8054
3	0	8	126	597	1878	2419	958	409	342	321	...	2334
4	1	16	219	1987	2799	1375	643	394	244	223	...	753
5	0	44	1101	2287	1782	2891	3024	3543	14193	8321	...	4906
6	2	879	8719	23580	24298	27227	26753	29774	37880	54145	...	101921
7	1	86	438	1119	4817	8619	8200	4638	8580	8715	...	6013
8	0	5	28	146	472	1435	1873	1389	1344	617	...	2644
9	0	15	70	434	1723	2243	2436	2640	6150	7030	...	2760
10	0	120	293	1692	7486	13581	10864	8276	8699	15391	...	8305
11	0	34	350	2368	4006	5111	3634	4368	2622	1626	...	1678
12	1	40	340	1706	2165	3137	3538	2388	2782	3575	...	4644
13	0	126	338	1880	5888	7051	6566	7024	7425	7996	...	8794

14	3	518	6002	17546	21647	18627	15638	14620	12143	16688	...	33216
15	0	30	385	1901	4352	4176	5717	5437	3920	3415	...	1922
16	0	22	536	1560	1395	1168	1147	784	978	1196	...	1851
17	0	12	116	534	1215	1804	1481	1006	664	506	...	2995
18	0	103	336	1373	6160	12703	10501	11171	13256	14402	...	20380
19	0	29	183	1922	4795	5125	3105	3481	3395	2827	...	6304
20	0	168	775	3066	8880	10053	6435	3926	3839	4672	...	5674
21	0	38	139	892	1193	1936	2705	2613	5048	8746	...	5503
22	0	113	862	1112	2079	4207	2466	1522	1229	935	...	11390
23	0	44	122	880	2563	8617	7490	4806	5866	6036	...	8403
24	0	88	1227	3139	5506	4377	3289	2751	2769	2227	...	18636
25	0	37	411	3661	8343	11777	7284	4364	4130	6071	...	6548
26	0	75	1179	3535	7436	10991	5707	3457	2198	3081	...	4759
27	0	24	557	1741	6388	10730	6159	4370	3333	3701	...	7394
28	0	12	253	969	1703	2401	1621	965	710	1016	...	1117
29	0	53	874	4296	7298	11376	6485	4762	3318	2741	...	10667
30	0	75	595	1385	2852	5955	4593	3452	2860	2420	...	10985
31	0	25	92	247	716	1906	2545	2153	3929	5792	...	1523
32	8	3122	29650	92470	158886	212274	170819	147468	172147	206038	...	321630

33 rows × 38 columns

## Promedio de Casos

```
In [12]: num_regions = len(df)
for month in months:
    months[month] /= num_regions
```

```
In [13]: promedios_x_mes_df = pd.DataFrame(months)
promedios_x_mes_df["nombre"] = df["nombre"]
```

In [14]: `print(promedios_x_mes_df)`

	02-2020	03-2020	04-2020	05-2020	06-2020	07-2020
0	0.000000	1.636364	7.909091	20.303030	45.757576	54.909091
1	0.000000	5.666667	73.060606	118.363636	135.939394	132.878788
2	0.000000	1.272727	9.454545	10.272727	31.969697	92.696970
3	0.000000	0.242424	3.818182	18.090909	56.909091	73.303030
4	0.030303	0.484848	6.636364	60.212121	84.818182	41.666667
5	0.000000	1.333333	33.363636	69.303030	54.000000	87.606061
6	0.060606	26.636364	264.212121	714.545455	736.303030	825.060606
7	0.030303	2.606061	13.272727	33.909091	145.969697	261.181818
8	0.000000	0.151515	0.848485	4.424242	14.303030	43.404040

## Día del mes en el que se presentó el mayor num de casos confirmados

In [20]: `totales_x_dia = df[columns].sum(axis=0)`

`indice_max = totales_x_dia.idxmax()`

`dia_con_mas_casos = indice_max[3:]`

`print("El día con más casos confirmados fue:", dia_con_mas_casos)`

El día con más casos confirmados fue: 01-2022

## Porcentaje de población afectada por mes

```
In [16]: df["poblacion_total"] = df["poblacion"].sum()

totales_x_mes = {}

for column in columns:
    mes = column [4:]
    if mes not in totales_x_mes:
        totales_x_mes[mes] = 0
    totales_x_mes[mes] += df[column].sum()
```

```
In [17]: porcentaje_x_mes = {}
for mes, total in totales_x_mes.items():
    porcentaje_x_mes[mes] = total / df["poblacion_total"].sum() * 100
```



```
In [18]: print ("El porcentaje de la población afectada por mes, fue: ")
for mes, porcentaje in porcentaje_x_mes.items():
    print(f"{mes}: {porcentaje:.2f}%")
```

El porcentaje de la población afectada por mes, fue:

2-2020: 0.01%  
3-2020: 0.00%  
4-2020: 0.00%  
5-2020: 0.00%  
6-2020: 0.00%  
7-2020: 0.01%  
8-2020: 0.00%  
9-2020: 0.00%  
0-2020: 0.00%  
1-2020: 0.00%  
1-2021: 0.01%  
2-2021: 0.01%  
3-2021: 0.00%  
4-2021: 0.00%  
5-2021: 0.00%  
6-2021: 0.00%  
7-2021: 0.01%  
8-2021: 0.01%  
9-2021: 0.01%  
0-2021: 0.00%  
1-2022: 0.03%  
2-2022: 0.01%  
3-2022: 0.00%  
4-2022: 0.00%  
5-2022: 0.00%  
6-2022: 0.01%  
7-2022: 0.02%  
8-2022: 0.00%  
9-2022: 0.00%  
0-2022: 0.00%  
1-2023: 0.00%  
2-2023: 0.00%  
3-2023: 0.00%

**Funcion que presente alguna otra informacion relevante de acuerdo con tus opiniones**

In [ ]:

