

```
In [1]: %matplotlib notebook
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

In [2]: df = pd.read_csv("E:\Protfolio\Depression Data Analysis Using 2012-2018 California Dataset/adult-depression-lghc-indicator-24.csv")

In [3]: df.head(20)
```

Out[3]:

	Year	Strata	Strata Name	Frequency	Weighted Frequency	Percent	Lower 95% CL	Upper 95% CL
0	2012	Total	Total	1920	NaN	11.74	11.11	12.37
1	2012	Sex	Male	561	1116664.0	8.12	7.32	8.92
2	2012	Sex	Female	1359	2163108.0	15.25	14.30	16.20
3	2012	Race-Ethnicity	White	1314	1806371.0	14.57	13.67	15.46
4	2012	Race-Ethnicity	Black	97	222022.0	13.54	10.44	16.65
5	2012	Race-Ethnicity	Hispanic	412	923174.0	9.98	8.91	11.05
6	2012	Race-Ethnicity	Asian/Pacific Islander	61	220418.0	5.48	3.92	7.03
7	2012	Race-Ethnicity	Other	36	107786.0	17.34	11.09	23.60
8	2012	Education	No High School Diploma	282	579047.0	14.43	12.60	16.26
9	2012	Education	High School Graduate or GED Certificate	337	665672.0	11.44	9.98	12.89
10	2012	Education	Some College or Tech School	563	947473.0	13.25	11.95	14.55
11	2012	Education	College Graduate or Post Grad	717	1040822.0	10.00	9.11	10.90
12	2012	Income	< \$20,000	642	1118292.0	16.97	15.42	18.51
13	2012	Income	\$20,000 - \$34,999	295	490532.0	12.39	10.67	14.10
14	2012	Income	\$35,000 - \$49,999	187	296225.0	9.88	8.16	11.60
15	2012	Income	\$50,000 - \$74,999	250	440798.0	12.65	10.82	14.49
16	2012	Income	\$75,000 - \$99,999	160	290492.0	10.17	8.26	12.07
17	2012	Income	\$100,000+	270	454444.0	8.31	7.17	9.46
18	2012	Age	18 to 34	219	705000.0	7.63	6.55	8.72
19	2012	Age	35 to 44	240	576639.0	11.13	9.57	12.69

```
In [4]: df=df.drop(['Strata Name', 'Weighted Frequency','Lower 95% CL','Upper 95% CL'],axis=1)
df.head(10)
```

Out[4]:

	Year	Strata	Frequency	Percent
0	2012	Total	1920	11.74
1	2012	Sex	561	8.12
2	2012	Sex	1359	15.25
3	2012	Race-Ethnicity	1314	14.57
4	2012	Race-Ethnicity	97	13.54
5	2012	Race-Ethnicity	412	9.98
6	2012	Race-Ethnicity	61	5.48
7	2012	Race-Ethnicity	36	17.34
8	2012	Education	282	14.43
9	2012	Education	337	11.44

```
In [5]: year_total=df.loc[df['Strata'] == 'Total']
```

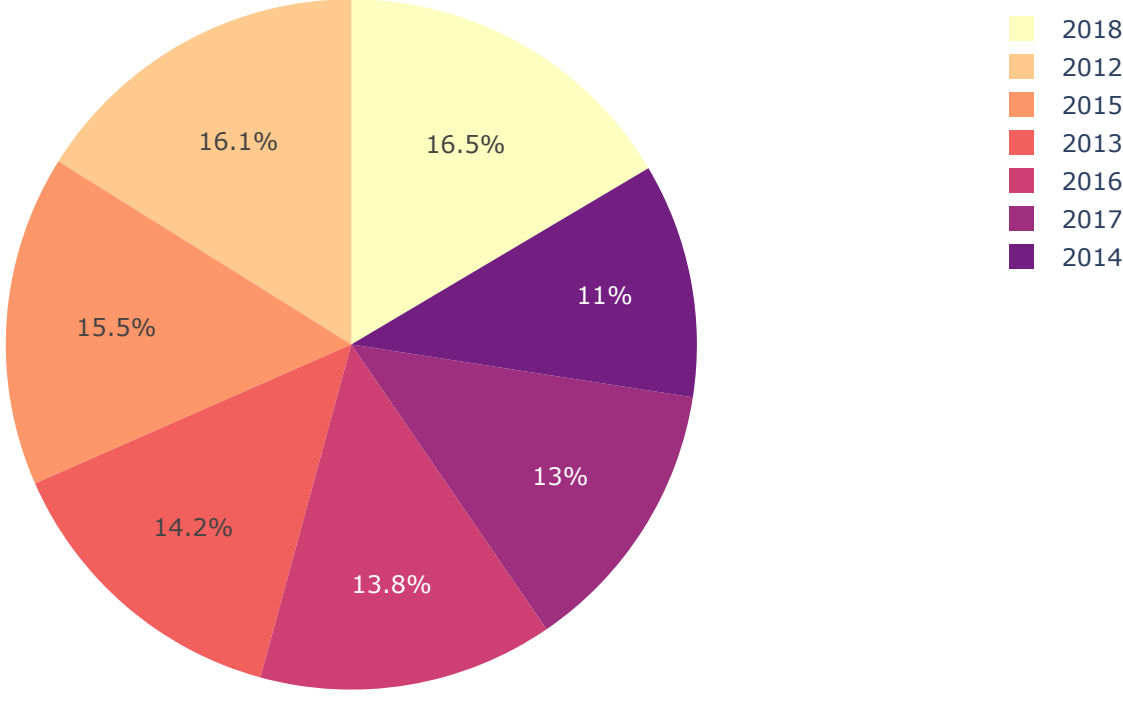
```
In [6]: year_total
```

Out[6]:

	Year	Strata	Frequency	Percent
0	2012	Total	1920	11.74
23	2013	Total	1689	13.08
46	2014	Total	1309	13.30
69	2015	Total	1848	12.92
92	2016	Total	1645	13.77
115	2017	Total	1550	19.04
138	2018	Total	1964	17.78

```
In [7]: fig = px.pie(year_total, values='Frequency',
names='Year', color_discrete_sequence=px.colors.sequential.Magma_r,
title='Percentage of Frequency per Year')
fig.show()
```

Percentage of Frequency per Year

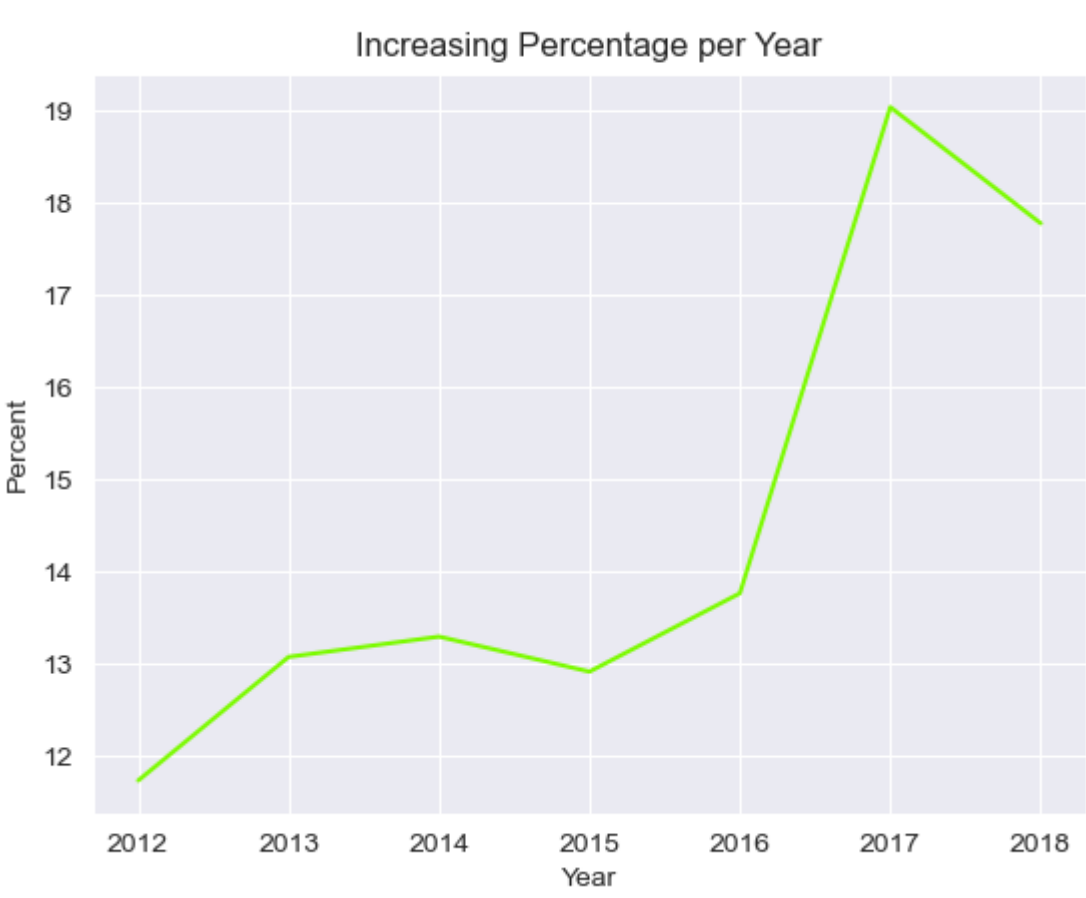


```
In [8]: data= year_total[['Year','Percent']]
data
```

Out[8]:

	Year	Percent
0	2012	11.74
23	2013	13.08
46	2014	13.30
69	2015	12.92
92	2016	13.77
115	2017	19.04
138	2018	17.78

```
In [9]: sns.set_style("darkgrid")
sns.lineplot(data=data, x="Year", y="Percent",markers=True,color='lawngreen').set(title='Increasing Per
centage per Year')
```



```
Out[9]: [Text(0.5, 1.0, 'Increasing Percentage per Year')]
```

```
In [10]: df_reason= df[['Strata','Frequency']]
```

```
In [11]: df_reason.head(20)
```

Out[11]:

	Strata	Frequency
0	Total	1920
1	Sex	561
2	Sex	1359
3	Race-Ethnicity	1314
4	Race-Ethnicity	97
5	Race-Ethnicity	412
6	Race-Ethnicity	61
7	Race-Ethnicity	36
8	Education	282
9	Education	337
10	Education	563
11	Education	717
12	Income	642
13	Income	295
14	Income	187
15	Income	250
16	Income	160
17	Income	270
18	Age	219
19	Age	240

```
In [12]: a=df_reason.groupby(['Strata']).sum().reset_index()
a[a.Strata != 'Total']
a
```

Out[12]:

	Strata	Frequency
0	Age	11925
1	Education	11283
2	Income	10270
3	Race-Ethnicity	11866
4	Sex	11925

```
In [13]: fig = px.pie(a,values='Frequency',names=px.Strata',
color_discrete_sequence=px.colors.sequential.RdBu,
title='Reason of Depression')
fig.show()
```

Reason of Depression

