

What is the prevalence of different chronic conditions of interest by year?

Chronic conditions:

- Hypertension
- Diabetes (excluding gestational diabetes)
- Cancers
- Musculoskeletal
- Cardiovascular
- Hyperlipidemia
- Asthma
- COPD
- Mental Health
- Sleep Disorders (sleep apnea, insomnia, narcolepsy) - Chronic Kidney Disease.

Calculate percentage of unique HCAs with a condition of interest. Use all HCAs including those who have not generated a claim as the denominator

Notes about the included chronic conditions

The chronic conditions we have listed above do not correspond to what we have developed now. **Do you want to update this list on this question?** I will continue to answer this question with the conditions we have discussed and incorporated in the data.

Chronic conditions listed above but not included as categories in our data:

- Hypertension
- Hyperlipidemia

In [2]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [190]:

```
med = pd.read_csv('../data/reshaped_med.csv')
enroll = pd.read_csv('../data/enroll.csv')
```

Total number of HCAs each year

In [186]:

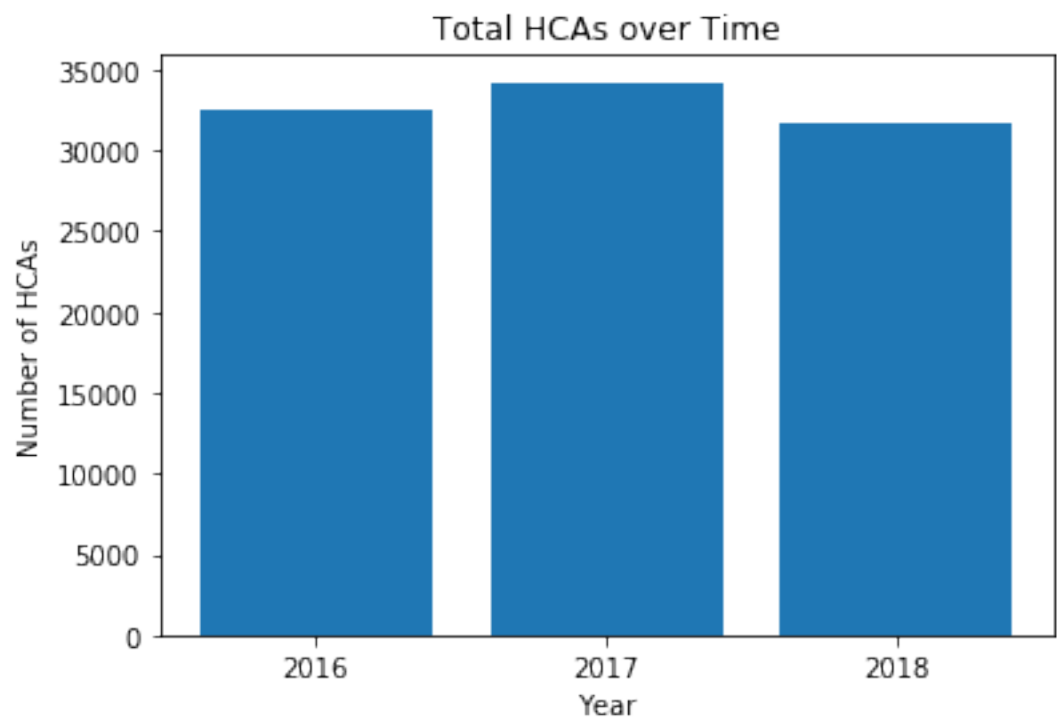
```
# Store total HCAs for each year
years = enroll.groupby(['Incurred Year']).size().reset_index(name='total')
years.loc[:, 'Incurred Year'] = ['2016', '2017', '2018']
years
```

Out[186]:

	Incurred Year	total
0	2016	32433
1	2017	34210
2	2018	31739

In [189]:

```
plt.bar(years['Incurred Year'], years['total'])
plt.title('Total HCAs over Time')
plt.xlabel("Year")
plt.ylabel("Number of HCAs")
plt.show()
```



Total number of claims each year

In [184]:

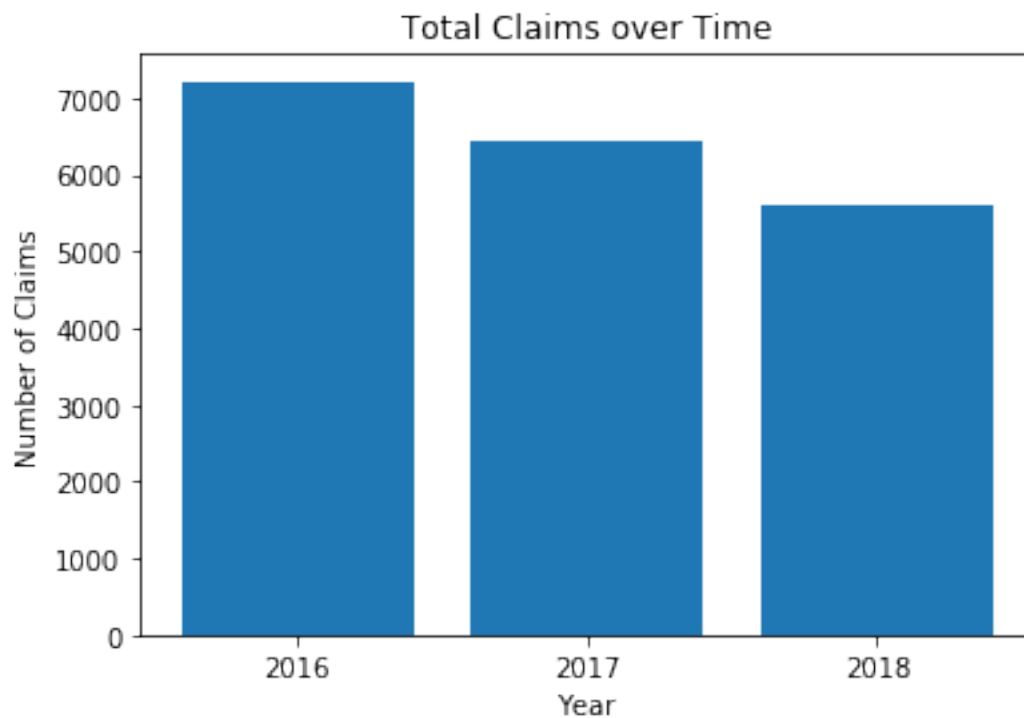
```
num_claims = med.groupby(['year']).size().reset_index(name='total')
num_claims.loc[:, 'year'] = ['2016', '2017', '2018']
num_claims
```

Out[184]:

	year	total
0	2016	7213
1	2017	6449
2	2018	5601

In [185]:

```
plt.bar(num_claims['year'], num_claims['total'])
plt.title('Total Claims over Time')
plt.xlabel("Year")
plt.ylabel("Number of Claims")
plt.show()
```



Calculating percentage

In [91]:

```
asthma = med.groupby(['year'])['binary_asthma'].sum().reset_index(name = 'asthma')
bp = med.groupby(['year'])['binary_bp'].sum().reset_index(name = 'bp')
cancer = med.groupby(['year'])['binary_cancer'].sum().reset_index(name = 'cancer')
cardiovascular = med.groupby(['year'])['binary_cardiovascular'].sum().reset_index(name = 'cardiovascular')
copd = med.groupby(['year'])['binary_copd'].sum().reset_index(name = 'copd')
diabetes = med.groupby(['year'])['binary_diabetes'].sum().reset_index(name = 'diabetes')
kidney = med.groupby(['year'])['binary_kidney'].sum().reset_index(name = 'kidney')
obesity = med.groupby(['year'])['binary_obesity'].sum().reset_index(name = 'obesity')
musculoskeletal = med.groupby(['year'])['binary_musculoskeletal'].sum().reset_index(name = 'musculoskeletal')
cholesterol = med.groupby(['year'])['binary_cholesterol'].sum().reset_index(name = 'cholesterol')
mental_disorder = med.groupby(['year'])['binary_mental_disorder'].sum().reset_index(name = 'mental_disorder')
```

Total claims for each chronic condition each year

In [144]:

```
totals_by_year = pd.concat([years, asthma, bp, cancer, cardiovascular, copd, diabetes, kidney, obesity, musculoskeletal, cholesterol, mental_disorder], axis=1)
totals_by_year
```

Out[144]:

	Incurred Year	total	asthma	bp	cancer	cardiovascular	copd	diabetes	kidney	ok
0	2016	32433	356	1252	350	1099	247	928	138	48
1	2017	34210	365	1304	329	1139	250	961	149	51
2	2018	31739	319	1154	251	1027	173	844	131	54

Percentages for each chronic condition each year

In [145]:

```
percentages = totals_by_year.iloc[:,2:].div(totals_by_year['total'], axis=0) * 100
percentages.loc[:, 'year'] = ['2016', '2017', '2018']
percentages
```

Out[145]:

	asthma	bp	cancer	cardiovascular	copd	diabetes	kidney	obese
0	1.097647	3.860266	1.079148	3.388524	0.761570	2.861283	0.425493	1.483061
1	1.066939	3.811751	0.961707	3.329436	0.730780	2.809120	0.435545	1.508541
2	1.005073	3.635905	0.790825	3.235767	0.545071	2.659189	0.412741	1.726541

In [148]:

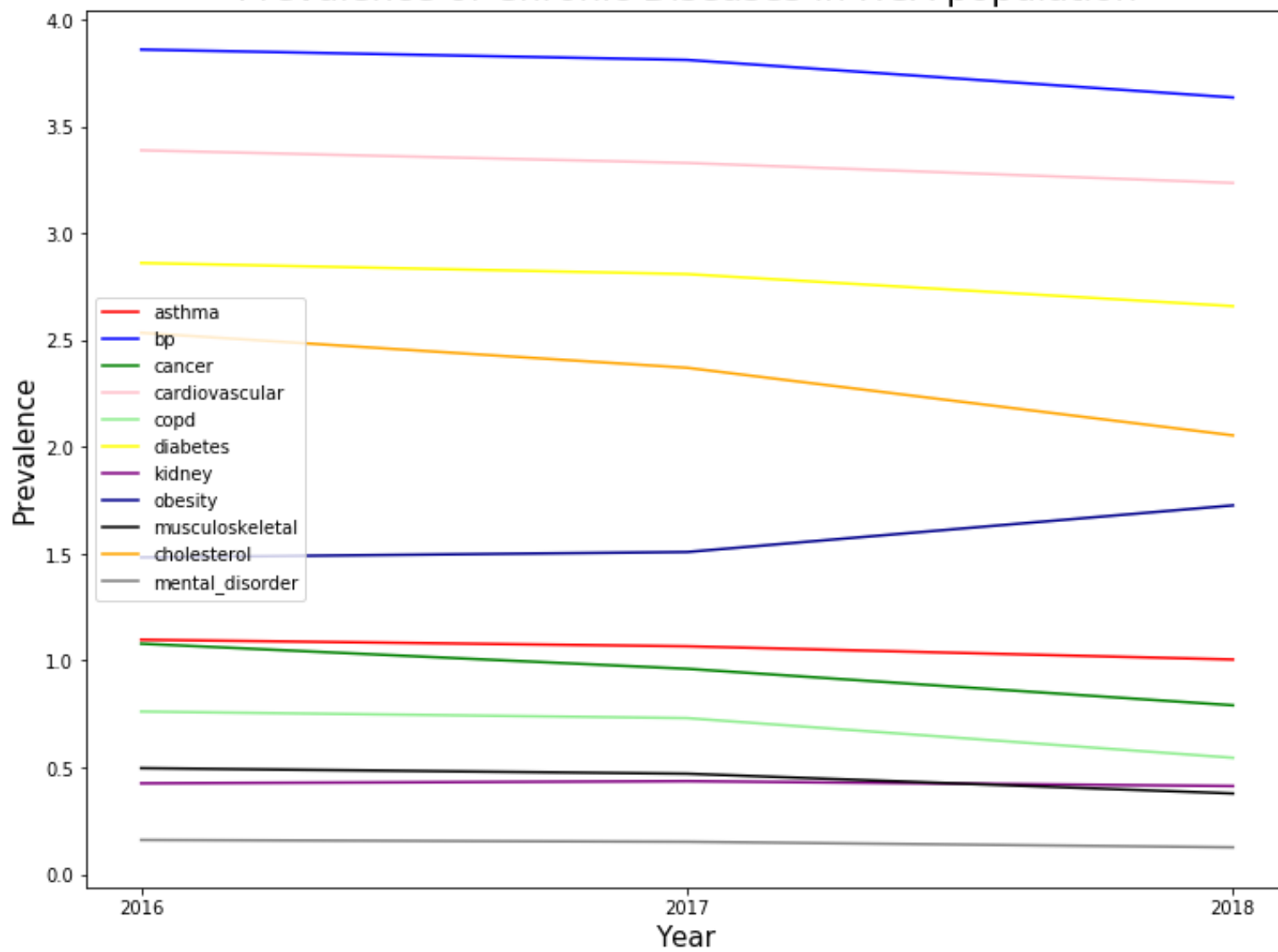
```
fig = plt.figure(figsize = (12, 9))

plt.plot(percentages['year'], percentages['asthma'], label = 'asthma', c = 'red'
)
plt.plot(percentages['year'], percentages['bp'], label = 'bp', c = 'blue')
plt.plot(percentages['year'], percentages['cancer'], label = 'cancer', c = 'green')
plt.plot(percentages['year'], percentages['cardiovascular'], label = 'cardiovascular', c = 'pink')
plt.plot(percentages['year'], percentages['copd'], label = 'copd', c = 'lightgreen')
plt.plot(percentages['year'], percentages['diabetes'], label = 'diabetes', c = 'yellow')
plt.plot(percentages['year'], percentages['kidney'], label = 'kidney', c = 'purple')
plt.plot(percentages['year'], percentages['obesity'], label = 'obesity', c = 'darkblue')
plt.plot(percentages['year'], percentages['musculoskeletal'], label = 'musculoskeletal', c = 'black')
plt.plot(percentages['year'], percentages['cholesterol'], label = 'cholesterol', c = 'orange')
plt.plot(percentages['year'], percentages['mental_disorder'], label = 'mental_disorder', c = 'gray')

# labels
plt.title('Prevalence of Chronic Diseases in HCA population', size = 20)
plt.xlabel("Year", size = 15)
plt.ylabel("Prevalence", size = 15)
plt.legend(['asthma', 'bp', 'cancer', 'cardiovascular', 'copd', 'diabetes', 'kidney', 'obesity', 'musculoskeletal',
           'cholesterol', 'mental_disorder'])

plt.show()
```

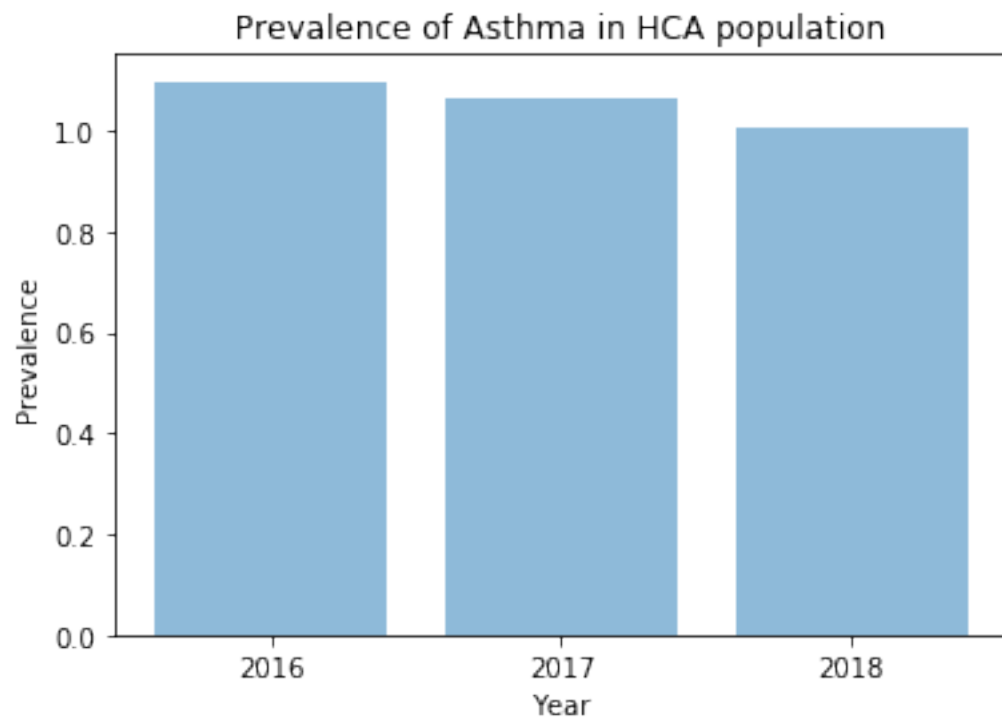
Prevalence of Chronic Diseases in HCA population



asthma

In [138]:

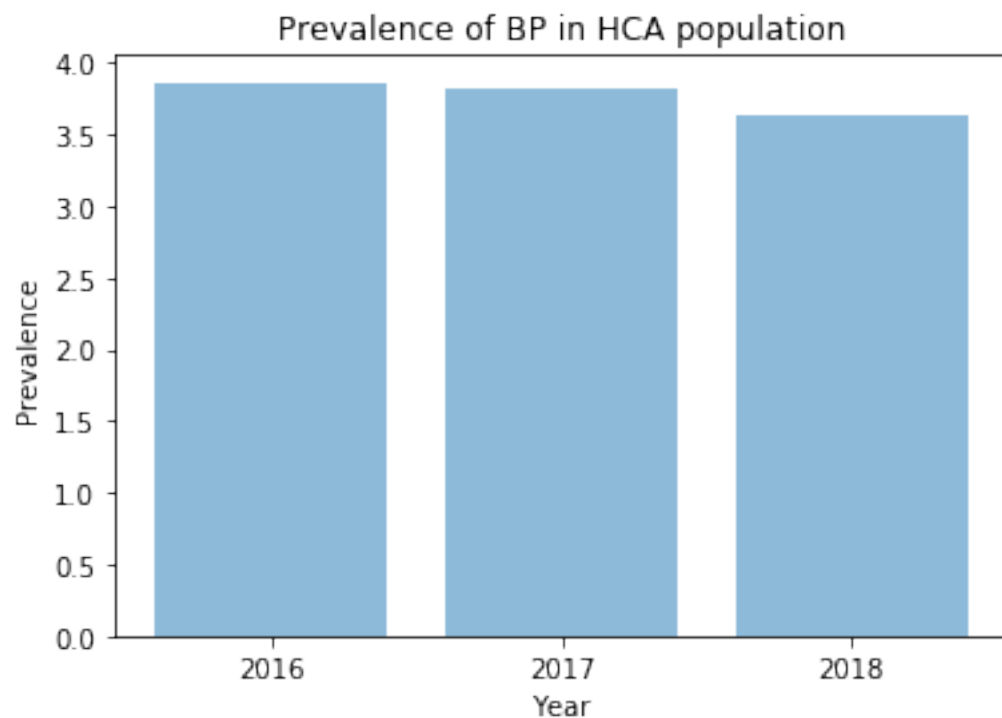
```
plt.bar(percentages['year'], percentages['asthma'], align='center', alpha=0.5)
plt.title('Prevalence of Asthma in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



bp

In [139]:

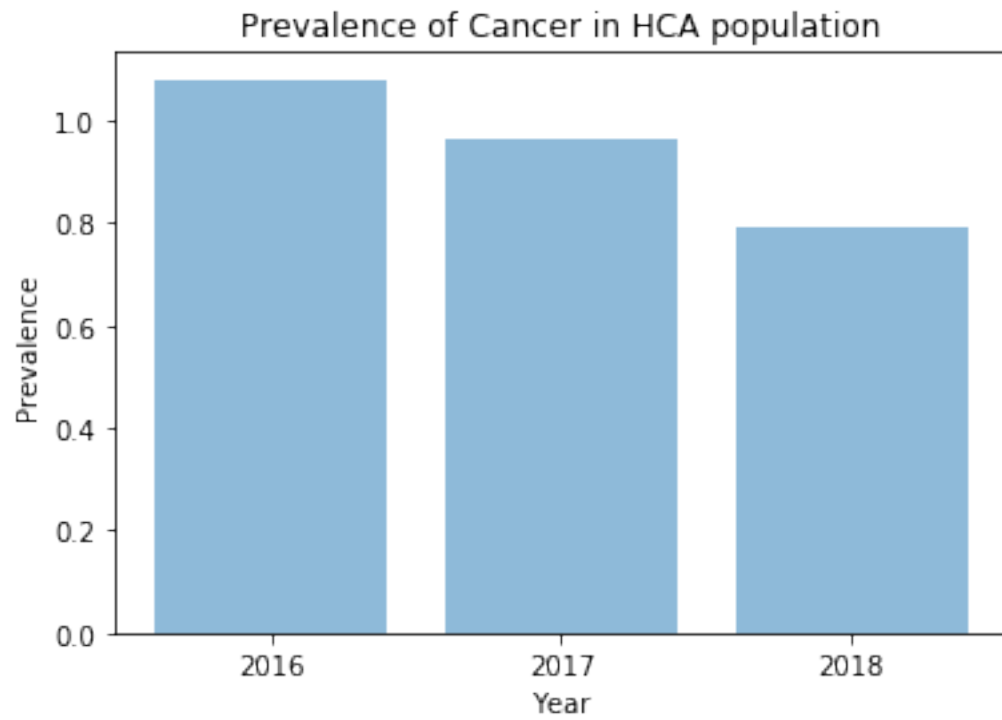
```
plt.bar(percentages['year'], percentages['bp'], align='center', alpha=0.5)
plt.title('Prevalence of BP in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



cancer

In [141]:

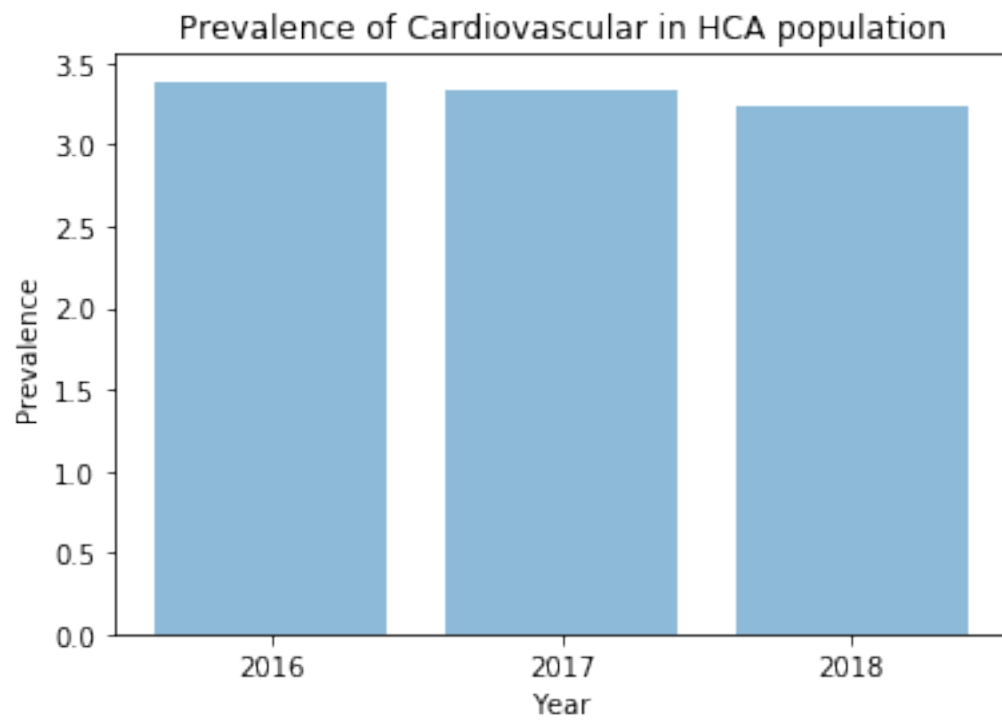
```
plt.bar(percentages['year'], percentages['cancer'], align='center', alpha=0.5)
plt.title('Prevalence of Cancer in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



cardiovascular

In [142]:

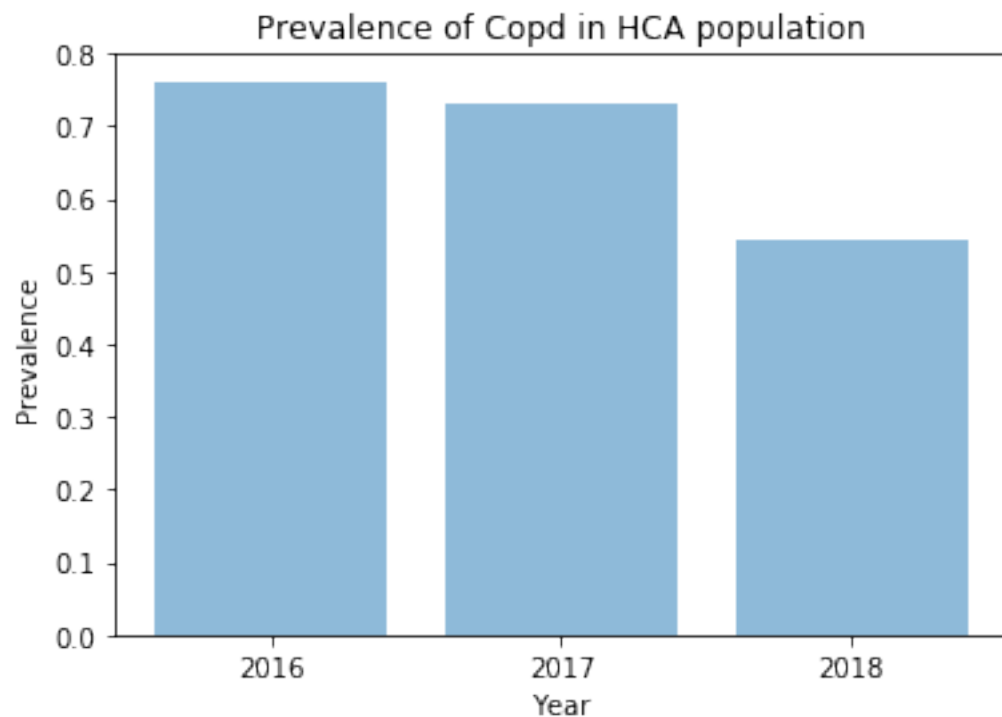
```
plt.bar(percentages['year'], percentages['cardiovascular'], align='center', alpha=0.5)
plt.title('Prevalence of Cardiovascular in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



copd

In [149]:

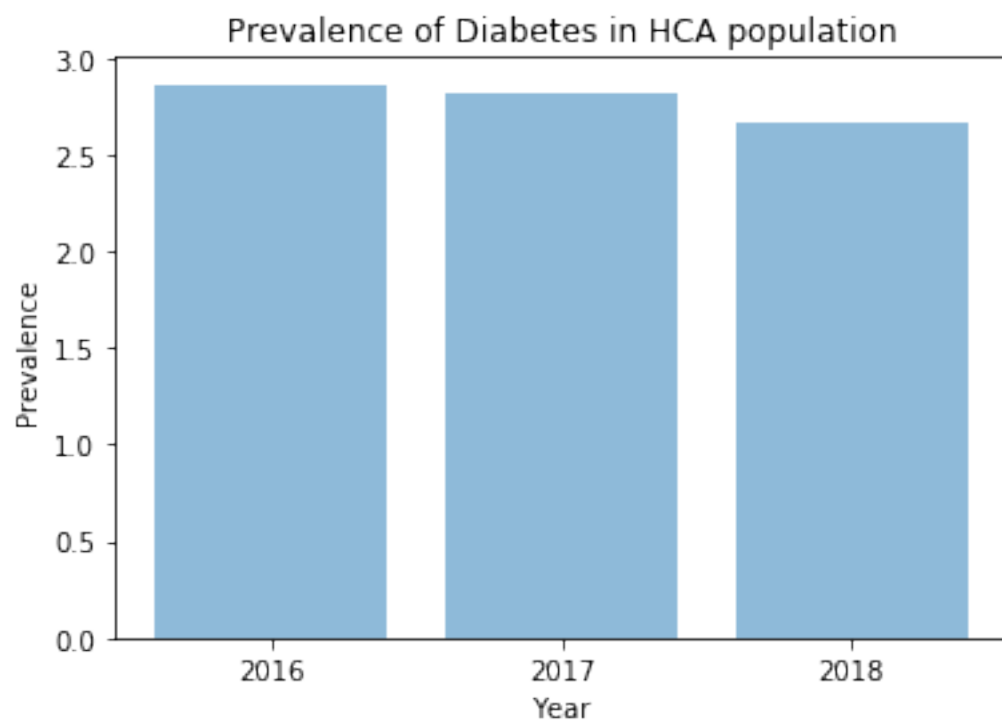
```
plt.bar(percentages['year'], percentages['copd'], align='center', alpha=0.5)
plt.title('Prevalence of Copd in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



diabetes

In [150]:

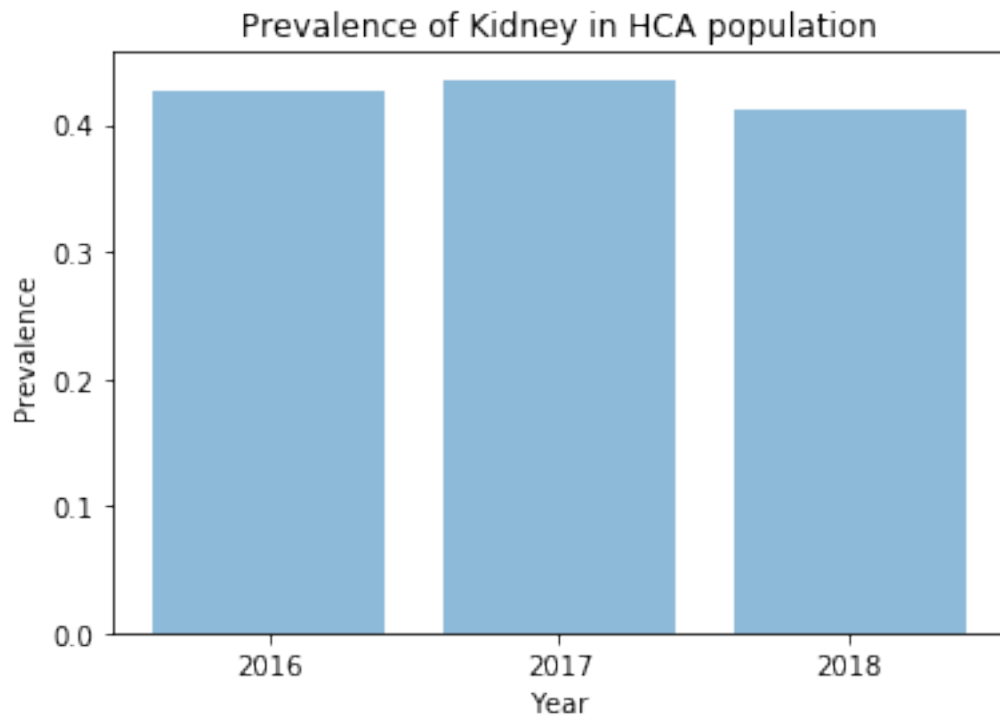
```
plt.bar(percentages['year'], percentages['diabetes'], align='center', alpha=0.5)
plt.title('Prevalence of Diabetes in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



kidney

In [151]:

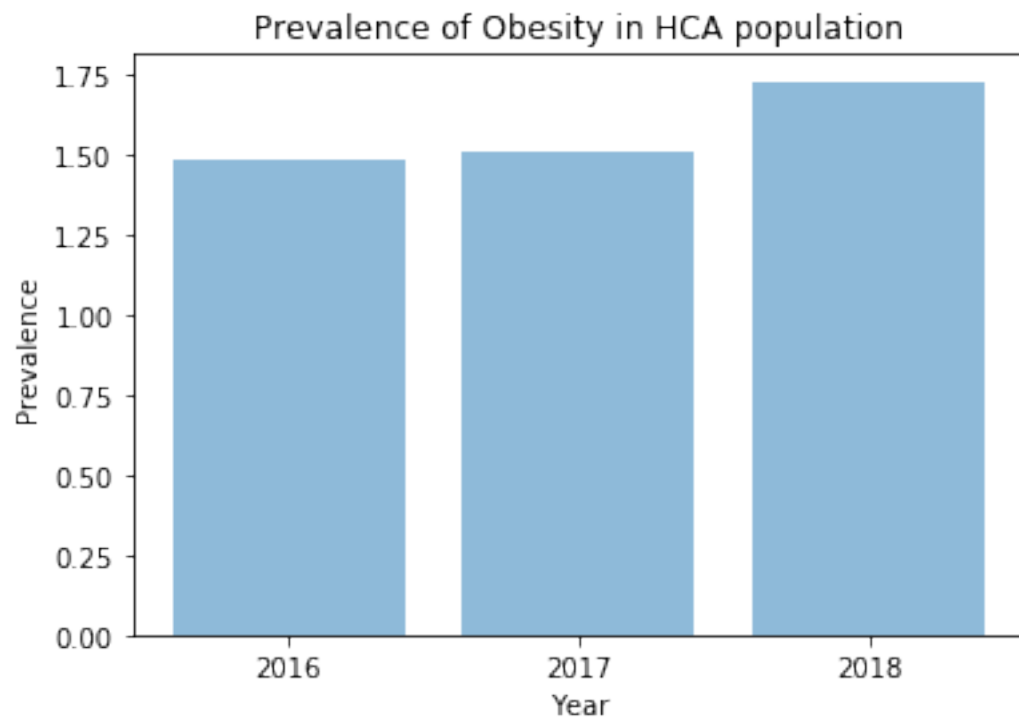
```
plt.bar(percentages['year'], percentages['kidney'], align='center', alpha=0.5)
plt.title('Prevalence of Kidney in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



obesity

In [152]:

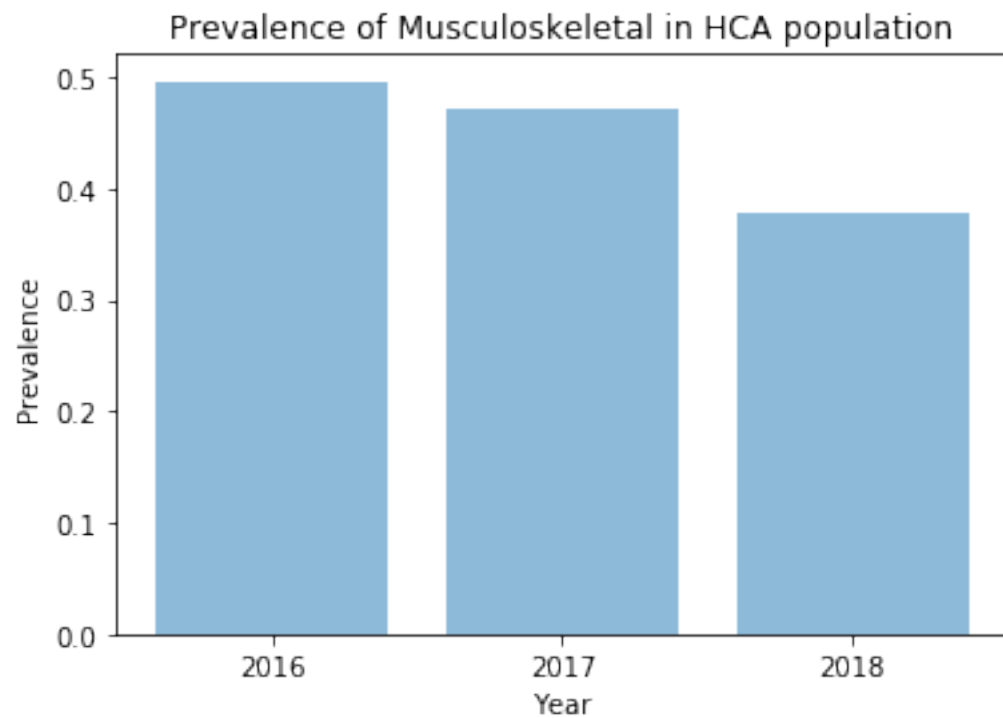
```
plt.bar(percentages['year'], percentages['obesity'], align='center', alpha=0.5)
plt.title('Prevalence of Obesity in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



musculoskeletal

In [153]:

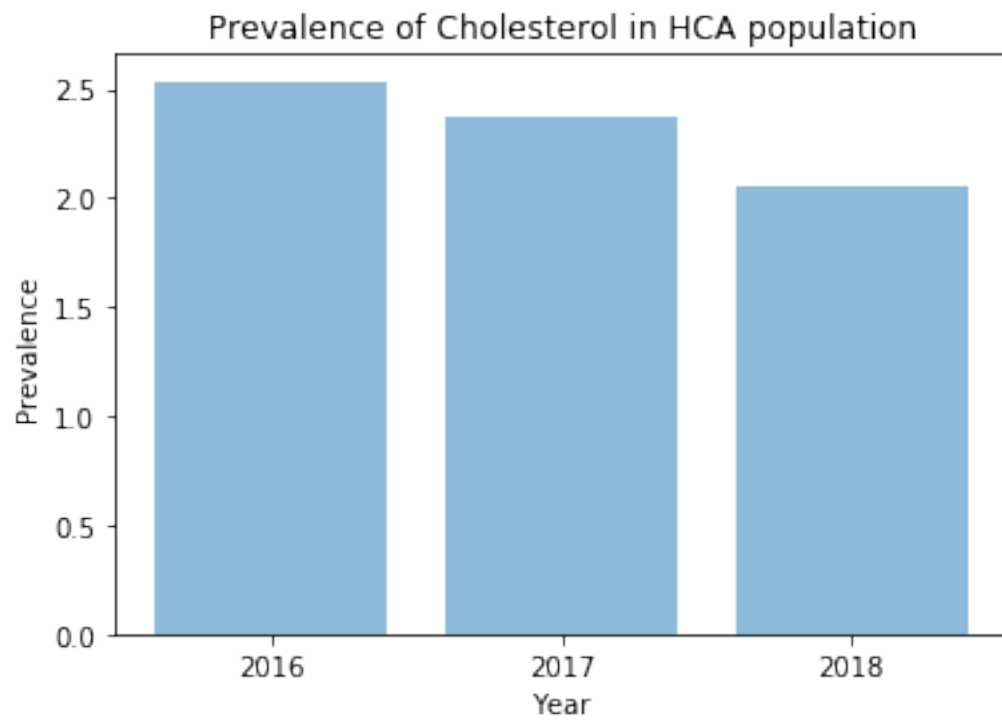
```
plt.bar(percentages['year'], percentages['musculoskeletal'], align='center', alpha=0.5)
plt.title('Prevalence of Musculoskeletal in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



cholesterol

In [156]:

```
plt.bar(percentages['year'], percentages['cholesterol'], align='center', alpha=0.5)
plt.title('Prevalence of Cholesterol in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```



mental_disorder

In [171]:

```
plt.bar(percentages['year'], percentages['mental_disorder'], align='center', alpha=0.5)
plt.title('Prevalence of Mental Disorder in HCA population')
plt.xlabel("Year")
plt.ylabel("Prevalence")
plt.show()
```

