Health Insurance Cost EDA and Analyzation

Introduction

Insurance premium and assured ammount largely depends on lifestyle and existing health condion of an individual. The data contains following information about people and their related insurance charges.

Columns

age: age of primary beneficiary

sex: insurance contractor gender, female, male

bmi: Body mass index, providing an understanding of body, weights that are relatively high or low relative to height, objective index of body weight

children: Number of children covered by health insurance / Number of dependents

smoker: Smoking

region: the beneficiary's residential area in the US, northeast, southeast, southwest, northwest.

charges: Individual medical costs billed by health insurance

Changing working directery

In [52]:
```
import os
os.chdir(r'C:\Desktop\Data Science\Python\Project')
os.getcwd()
```

Out[52]:
```
'C:\\Desktop\\Data Science\\Python\\Project'
```

In [53]:
```
# Removing/filtering unwanted warnings
import warnings
warnings.filterwarnings('ignore')
```

In [54]:

```
# Importing the numpy and pandas package
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O

!pip install https://github.com/pandas-profiling/pandas-profiling/archive/mas
import pandas_profiling

# Data Visualisation
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
Collecting https://github.com/pandas-profiling/pandas-profiling/archive/mas
ter.zip (https://github.com/pandas-profiling/pandas-profiling/archive/maste
r.zip)
  Using cached https://github.com/pandas-profiling/pandas-profiling/archiv
e/master.zip (https://github.com/pandas-profiling/pandas-profiling/archive/
master.zip)
Requirement already satisfied (use --upgrade to upgrade): pandas-profiling=
=2.12.0 from https://github.com/pandas-profiling/pandas-profiling/archive/m
aster.zip (https://github.com/pandas-profiling/pandas-profiling/archive/mas
ter.zip) in c:\programdata\anaconda3\lib\site-packages
Requirement already satisfied: joblib in c:\programdata\anaconda3\lib\site-
packages (from pandas-profiling==2.12.0) (0.17.0)
Requirement already satisfied: scipy>=1.4.1 in c:\programdata\anaconda3\lib
\site-packages (from pandas-profiling==2.12.0) (1.5.2)
Requirement already satisfied: pandas!=1.0.0,!=1.0.1,!=1.0.2,!=1.1.0,>=0.2
5.3 in c:\programdata\anaconda3\lib\site-packages (from pandas-profiling==
2.12.0) (1.1.3)
Requirement already satisfied: matplotlib>=3.2.0 in c:\programdata\anaconda
3\lib\site-packages (from pandas-profiling==2.12.0) (3.3.2)
Requirement already satisfied: confuse>=1.0.0 in c:\programdata\anaconda3\l
ib\site-packages (from pandas-profiling==2.12.0) (1.4.0)
Requirement already satisfied: jinja2>=2.11.1 in c:\programdata\anaconda3\l
ib\site-packages (from pandas-profiling==2.12.0) (2.11.2)
Requirement already satisfied: visions[type_image_path]==0.6.0 in c:\progra
mdata\anaconda3\lib\site-packages (from pandas-profiling==2.12.0) (0.6.0)
Requirement already satisfied: numpy>=1.16.0 in c:\programdata\anaconda3\li
b\site-packages (from pandas-profiling==2.12.0) (1.19.2)
Requirement already satisfied: attrs>=19.3.0 in c:\programdata\anaconda3\li
b\site-packages (from pandas-profiling==2.12.0) (20.3.0)
Requirement already satisfied: htmlmin>=0.1.12 in c:\programdata\anaconda3
\lib\site-packages (from pandas-profiling==2.12.0) (0.1.12)
Requirement already satisfied: missingno>=0.4.2 in c:\programdata\anaconda3
\lib\site-packages (from pandas-profiling==2.12.0) (0.4.2)
Requirement already satisfied: phik>=0.10.0 in c:\programdata\anaconda3\lib
\site-packages (from pandas-profiling==2.12.0) (0.11.2)
Requirement already satisfied: tangled-up-in-unicode>=0.0.6 in c:\programda
ta\anaconda3\lib\site-packages (from pandas-profiling==2.12.0) (0.0.7)
Requirement already satisfied: requests>=2.24.0 in c:\programdata\anaconda3
\lib\site-packages (from pandas-profiling==2.12.0) (2.24.0)
Requirement already satisfied: tqdm>=4.48.2 in c:\programdata\anaconda3\lib
\site-packages (from pandas-profiling==2.12.0) (4.50.2)
Requirement already satisfied: seaborn>=0.10.1 in c:\programdata\anaconda3
\lib\site-packages (from pandas-profiling==2.12.0) (0.11.0)
```

```
Requirement already satisfied: pytz>=2017.2 in c:\programdata\anaconda3\lib
\site-packages (from pandas!=1.0.0,!=1.0.1,!=1.0.2,!=1.1.0,>=0.25.3->pandas
-profiling==2.12.0) (2020.1)
Requirement already satisfied: python-dateutil>=2.7.3 in c:\programdata\ana
conda3\lib\site-packages (from pandas!=1.0.0,!=1.0.1,!=1.0.2,!=1.1.0,>=0.2
5.3->pandas-profiling==2.12.0) (2.8.1)
Requirement already satisfied: cycler>=0.10 in c:\programdata\anaconda3\lib
\site-packages (from matplotlib>=3.2.0->pandas-profiling==2.12.0) (0.10.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.3 in
 c:\programdata\anaconda3\lib\site-packages (from matplotlib>=3.2.0->pandas
-profiling==2.12.0) (2.4.7)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\programdata\anaconda
3\lib\site-packages (from matplotlib>=3.2.0->pandas-profiling==2.12.0) (1.
3.0)
Requirement already satisfied: certifi>=2020.06.20 in c:\programdata\anacon
da3\lib\site-packages (from matplotlib>=3.2.0->pandas-profiling==2.12.0) (2
020.6.20)
Requirement already satisfied: pillow>=6.2.0 in c:\programdata\anaconda3\li
b\site-packages (from matplotlib>=3.2.0->pandas-profiling==2.12.0) (8.0.1)
Requirement already satisfied: pyyaml in c:\programdata\anaconda3\lib\site-
packages (from confuse>=1.0.0->pandas-profiling==2.12.0) (5.3.1)
Requirement already satisfied: MarkupSafe>=0.23 in c:\programdata\anaconda3
\lib\site-packages (from jinja2>=2.11.1->pandas-profiling==2.12.0) (1.1.1)
Requirement already satisfied: networkx>=2.4 in c:\programdata\anaconda3\li
b\site-packages (from visions[type_image_path]==0.6.0->pandas-profiling==2.
12.0) (2.5)
Requirement already satisfied: imagehash; extra == "type_image_path" in
 c:\programdata\anaconda3\lib\site-packages (from visions[type_image_path]=
=0.6.0->pandas-profiling==2.12.0) (4.2.0)
Requirement already satisfied: idna<3,>=2.5 in c:\programdata\anaconda3\lib
\site-packages (from requests>=2.24.0->pandas-profiling==2.12.0) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
 c:\programdata\anaconda3\lib\site-packages (from requests>=2.24.0->pandas-
profiling==2.12.0) (1.25.11)
Requirement already satisfied: chardet<4,>=3.0.2 in c:\programdata\anaconda
3\lib\site-packages (from requests>=2.24.0->pandas-profiling==2.12.0) (3.0.
4)
Requirement already satisfied: six>=1.5 in c:\programdata\anaconda3\lib\sit
e-packages (from python-dateutil>=2.7.3->pandas!=1.0.0,!=1.0.1,!=1.0.2,!=1.
1.0,>=0.25.3->pandas-profiling==2.12.0) (1.15.0)
Requirement already satisfied: decorator>=4.3.0 in c:\programdata\anaconda3
\lib\site-packages (from networkx>=2.4->visions[type_image_path]==0.6.0->pa
ndas-profiling==2.12.0) (4.4.2)
Requirement already satisfied: PyWavelets in c:\programdata\anaconda3\lib\s
ite-packages (from imagehash; extra == "type_image_path"->visions[type_imag
e_path]==0.6.0->pandas-profiling==2.12.0) (1.1.1)
Building wheels for collected packages: pandas-profiling
  Building wheel for pandas-profiling (setup.py): started
  Building wheel for pandas-profiling (setup.py): finished with status 'don
e'
  Created wheel for pandas-profiling: filename=pandas_profiling-2.12.0-py2.
py3-none-any.whl size=243837 sha256=ff94167ff05269d096b1236da79957c65f13f3d
2a6ffb6356b4abe90b1a064bd
  Stored in directory: C:\Users\nilak\AppData\Local\Temp\pip-ephem-wheel-ca
che-htur3nv3\wheels\64\b6\85\dfc808b23666a5910371784e349d28818006ff63ed9cfe
ca59
Successfully built pandas-profiling
```

Reading csv file

In [55]:
```
df=pd.read_csv("C:\Desktop\Data Science\Python\project\insurance.csv")
```

Checking number of observation and columns ,index of columns, name of columns, number of non_null values and datatype, (memory usage,...)

In [56]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1338 entries, 0 to 1337
Data columns (total 7 columns):
 #   Column    Non-Null Count  Dtype  
---  ------    --------------  -----  
 0   age       1338 non-null   int64  
 1   sex       1338 non-null   object 
 2   bmi       1338 non-null   float64
 3   children  1338 non-null   int64  
 4   smoker    1338 non-null   object 
 5   region    1338 non-null   object 
 6   charges   1338 non-null   float64
dtypes: float64(2), int64(2), object(3)
memory usage: 73.3+ KB
```

checking the head of the data frame

In [57]:
```
df.head(10)
```

Out[57]:

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.90	0	yes	southwest	16884.92
1	18	male	33.77	1	no	southeast	1725.55
2	28	male	33.00	3	no	southeast	4449.46
3	33	male	22.70	0	no	northwest	21984.47
4	32	male	28.88	0	no	northwest	3866.86
5	31	female	25.74	0	no	southeast	3756.62
6	46	female	33.44	1	no	southeast	8240.59
7	37	female	27.74	3	no	northwest	7281.51
8	37	male	29.83	2	no	northeast	6406.41
9	60	female	25.84	0	no	northwest	28923.14

checking the yail of the data frame

In [58]:
```
df.tail(10)
```

Out[58]:

	age	sex	bmi	children	smoker	region	charges
1328	23	female	24.23	2	no	northeast	22395.74
1329	52	male	38.60	2	no	southwest	10325.21
1330	57	female	25.74	2	no	southeast	12629.17
1331	23	female	33.40	0	no	southwest	10795.94
1332	52	female	44.70	3	no	southwest	11411.68
1333	50	male	30.97	3	no	northwest	10600.55
1334	18	female	31.92	0	no	northeast	2205.98
1335	18	female	36.85	0	no	southeast	1629.83
1336	21	female	25.80	0	no	southwest	2007.94
1337	61	female	29.07	0	yes	northwest	29141.36

Getting idea about total number of observtion and column

In [59]:
```
df.shape
```

Out[59]:
```
(1338, 7)
```

In [60]:
```
#Getting number of rows(obs)
df.shape[0]
```

Out[60]:
```
1338
```

In [61]:
```
#Getting number of collumns
df.shape[1]
```

Out[61]:
```
7
```

Get to know the data type of each column

In [62]:
```
df.dtypes
```

Out[62]:
```
age           int64
sex          object
bmi         float64
children      int64
smoker       object
region       object
charges     float64
dtype: object
```

Finding number of missing values

In [63]:
```
df.isnull().sum()
```

Out[63]:
```
age         0
sex         0
bmi         0
children    0
smoker      0
region      0
charges     0
dtype: int64
```

We don't have any missing values in our data.

Summarizing and generating report for our dataset

In [64]:
```
from pandas_profiling import ProfileReport
```

In [65]:
```
pandas_profiling.ProfileReport(df)
```

A Jupyter widget could not be displayed because the widget state could not be found. This could happen if the kernel storing the widget is no longer available, or if the widget state was not saved in the notebook. You may be able to create the widget by running the appropriate cells.

A Jupyter widget could not be displayed because the widget state could not be found. This could happen if the kernel storing the widget is no longer available, or if the widget state was not saved in the notebook. You may be able to create the widget by running the appropriate cells.

A Jupyter widget could not be displayed because the widget state could not be found. This could happen if the kernel storing the widget is no longer available, or if the widget state was not saved in the notebook. You may be able to create the widget by running the appropriate cells.

Getting columns' names to makesure is it probaly assinned,if not we can modify accordingly

In [67]:
```
df.columns
```

Out[67]:
```
Index(['age', 'sex', 'bmi', 'children', 'smoker', 'region', 'charges'], dty
pe='object')
```

In [68]:
```
#I just want rename the following column
df=df.rename(columns={'children':'dependents'})
df.columns
```

Out[68]:
```
Index(['age', 'sex', 'bmi', 'dependents', 'smoker', 'region', 'charges'], d
type='object')
```

Finding Categorical variables

In [69]:
```
df_cat=df.select_dtypes(include='object')
df_cat.columns
```

Out[69]:
```
Index(['sex', 'smoker', 'region'], dtype='object')
```

In [70]:
```
#Filter categorical variables
categorical_columns= {x for x in df.dtypes.index if df.dtypes[x]=='object'}
```

In [71]:
```
#Print frequencey of categories
for col in categorical_columns:
    print ('\nFrequency of Categories for varible %s'%col)
    print(df[col].value_counts())
```

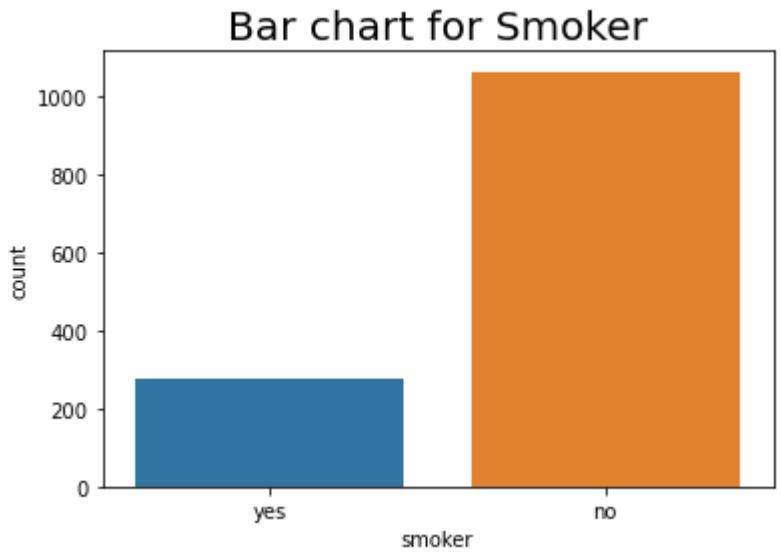
```
Frequency of Categories for varible region
southeast    364
northwest    325
southwest    325
northeast    324
Name: region, dtype: int64

Frequency of Categories for varible sex
male      676
female    662
Name: sex, dtype: int64

Frequency of Categories for varible smoker
no     1064
yes     274
Name: smoker, dtype: int64
```

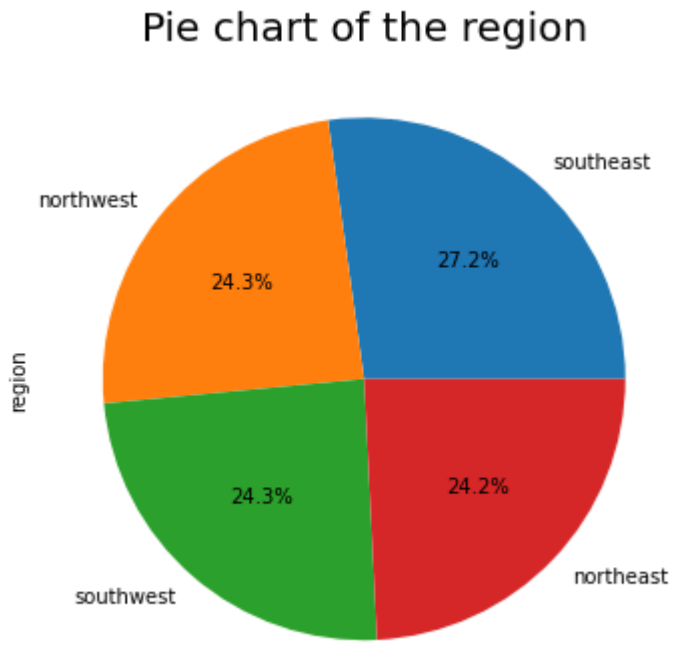
In [72]:
```
#Countplot
sns.countplot(df['smoker'])
plt.title("Bar chart for Smoker", y=1, fontdict={"fontsize": 20})
```

Out[72]: Text(0.5, 1, 'Bar chart for Smoker')



There are more non smoker than the smokers

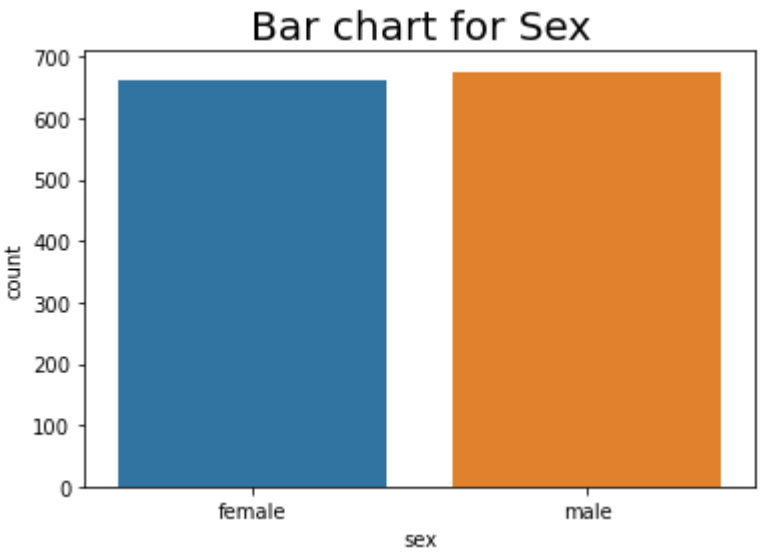
In [73]:
```
plt.figure(figsize=(6,6))
df.region.value_counts().plot.pie(autopct="%0.1f%%")
plt.title('Pie chart of the region',{'fontsize':20})
plt.show()
```



Number of people from southeast is little higher than the other region

In [74]:
```
#Countplot
sns.countplot(df['sex'])
plt.title("Bar chart for Sex", y=1, fontdict={"fontsize": 20})
```

Out[74]: Text(0.5, 1, 'Bar chart for Sex')



In [75]:
```
df.sex.value_counts(normalize=True)*100
```

Out[75]:
```
male      50.52
female    49.48
Name: sex, dtype: float64
```

The percentage of the male is little higher than the female

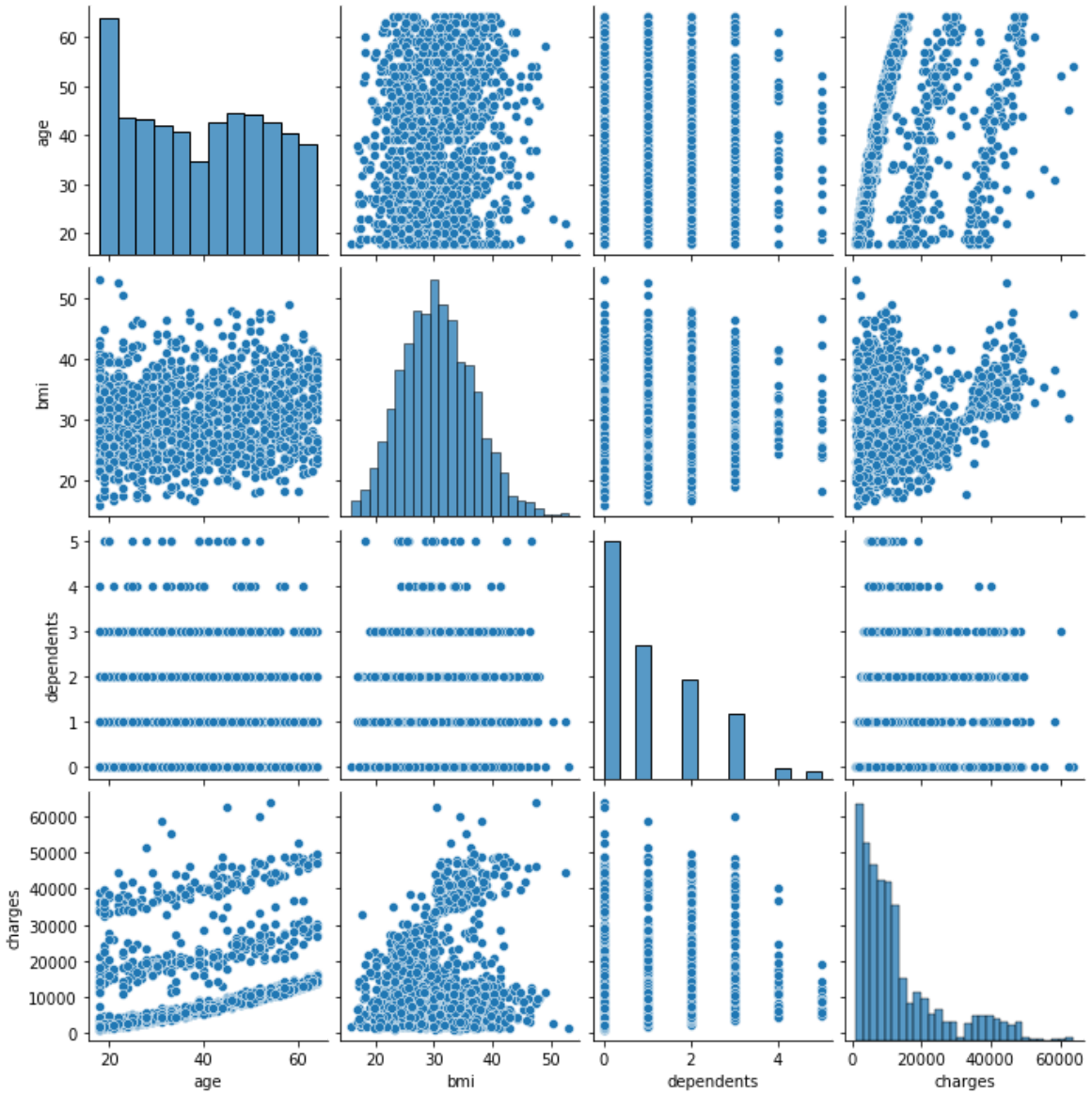
Finding numerical variables

In [76]:
```
df_num=df.select_dtypes(exclude='object')
df_num.columns
```

Out[76]: Index(['age', 'bmi', 'dependents', 'charges'], dtype='object')

In [77]:
```
sns.pairplot(df)
```

Out[77]: <seaborn.axisgrid.PairGrid at 0x1d7db90f910>

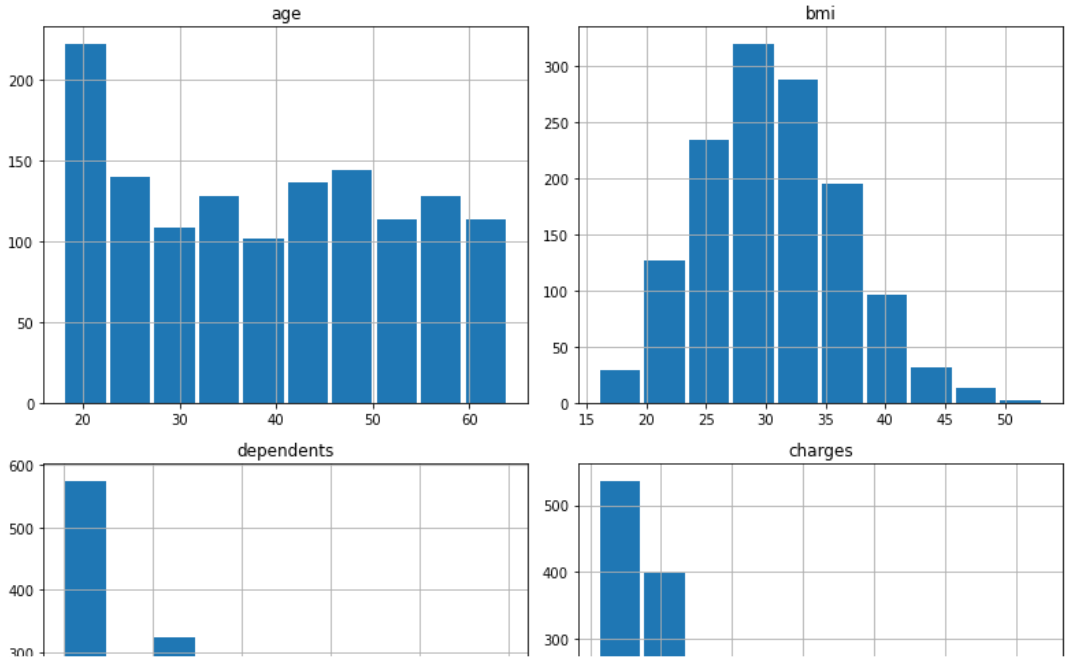


As there are only 4 attributes numeric in nature (int or float), the rest of them are not included in the pair plot. The diagonal plots represent distribution of the variable relative to itself. The plot between ‘age’ and ‘bmi’ does not indicate any relationship between them. The plot is more like a cloud and lacks direction. Between ‘age’ and ‘charges’, there is no clear relationship, though there seem to be 3 lines of positive relationship between them. It means, there are 3 sets of charges which increase gradually with age. No clear relation between ‘age’ and ‘dependents’ either. The range of

‘bmi’ decreases as dependents increases, however there are some extreme values in ‘bmi’ for children value 5. There is a little positive relation between ‘bmi’ and ‘charges’, although the plot is a cloud on initial values of ‘charges’. The range of ‘charges’ decreases as the value of ‘children’ increases.

Let's gets some idea about numerical variables

In [80]:
```
df.hist(rwidth=0.9,figsize=(11,9))
plt.tight_layout()
```



Getting the summary of Data for numerical Column with two decimal points

In [81]:
```
pd.options.display.float_format = "{:.2f}".format
df.describe()
```

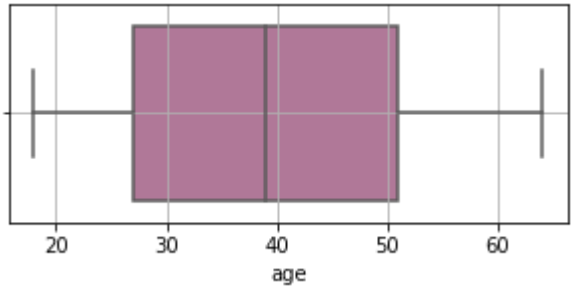
Out[81]:

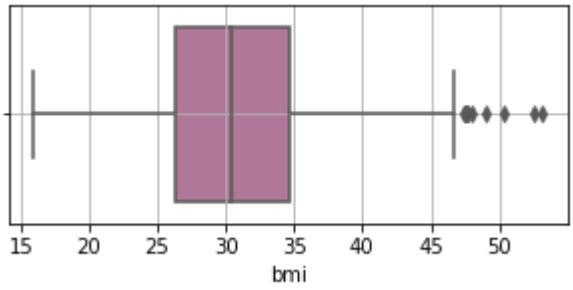
	age	bmi	dependents	charges
count	1338.00	1338.00	1338.00	1338.00
mean	39.21	30.66	1.09	13270.42
std	14.05	6.10	1.21	12110.01
min	18.00	15.96	0.00	1121.87
25%	27.00	26.30	0.00	4740.29
50%	39.00	30.40	1.00	9382.03
75%	51.00	34.69	2.00	16639.91
max	64.00	53.13	5.00	63770.43

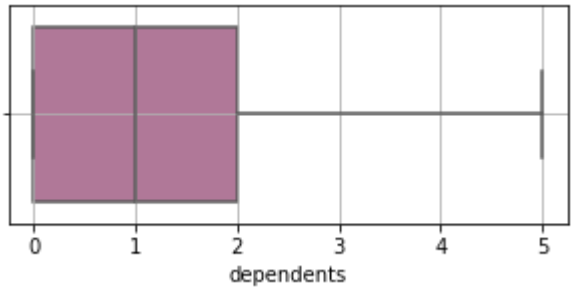
Accourding to this above discribtion, Age is between 18 to 64 and the mean is 39.21. BMI is between 15.96 to 53.13 and the mean is 30.66. Dependents are between 0 to 5 and the mean is 1.09=1. Charges are between $1121.87 to 63770.43$ and the mean is $13270.42.

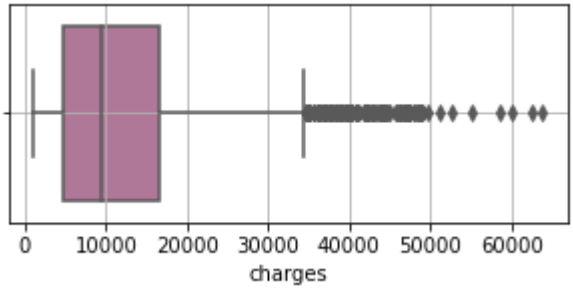
Finding outliers

In [82]:
```
for column in df.select_dtypes(include= ["float64", "int64"]).columns:
  plt.figure(figsize= (5,2))
  sns.boxplot(x= df[column], color = "#b96f99")
  plt.grid(True)
  plt.show()
```









BMI, and Charges in this dataset have potencial outliers. I will extract outliers to see the behavior

of them. We know, Potential outliers are observations that lie outside the lower and upper limits.

Lower limit = Q1 - 1.5 * IQR Upper limit = Q3 + 1.5 * IQR

Our target colomn is Charges. Let's take a closer look to the target column.

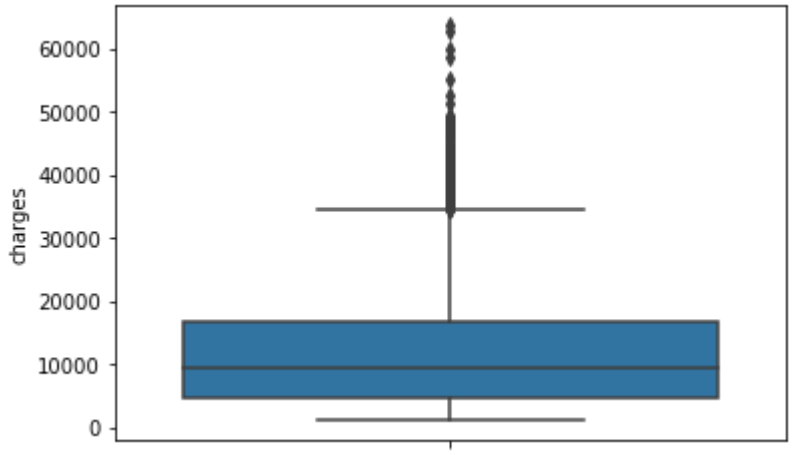
Our Taget column is Charges

In [83]:
```
df['charges'].describe()
```

Out[83]:
```
count    1338.00
mean    13270.42
std     12110.01
min      1121.87
25%      4740.29
50%      9382.03
75%     16639.91
max     63770.43
Name: charges, dtype: float64
```

In [84]:
```
sns.boxplot( y=df['charges'] )
```

Out[84]:
```
<AxesSubplot:ylabel='charges'>
```

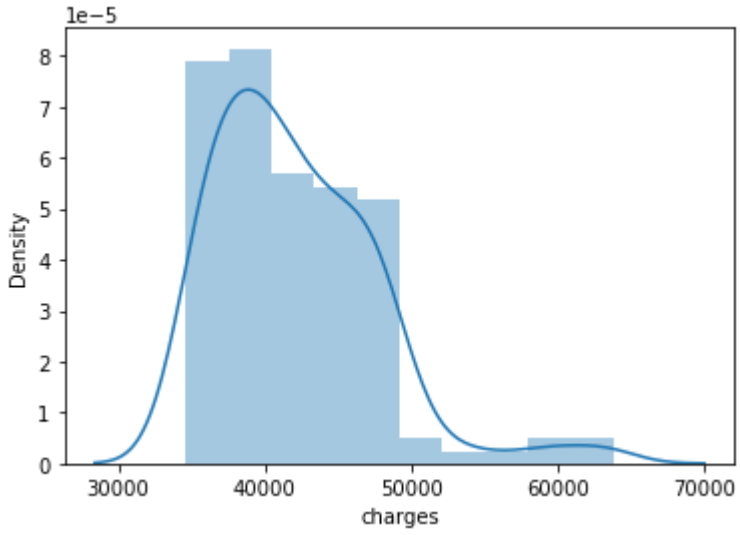


In [85]:
```
#Lower inner fence: Q1 - 1.5*IQR
#upper inner fence: Q3 + 1.5*IQR
def outlier_analysis(col):
    Q1=df[col].quantile(0.25)
    Q3=df[col].quantile(0.75)
    IQR=Q3-Q1
    UIF=Q3+1.5*(IQR)#UIF=upper inner fence
    LIF=Q1-1.5*(IQR) #LIF=lower inner fence
    df_out =df[(df[col]<LIF) | (df[col]>UIF)] # I created a new data set that
    sns.distplot(df_out[col])#Plotting univariate distributions.By default, t

    return df_out[col] .describe()
```

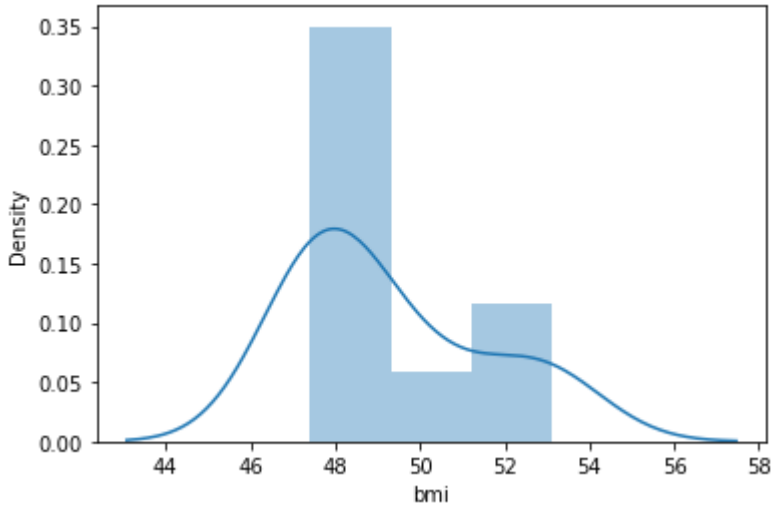
In [86]:
```
outlier_analysis('charges')
```

Out[86]:
```
count     139.00
mean    42103.95
std      5582.17
min     34617.84
25%     37786.15
50%     40974.16
75%     45786.71
max     63770.43
Name: charges, dtype: float64
```



In [87]:
```
outlier_analysis('bmi')
```

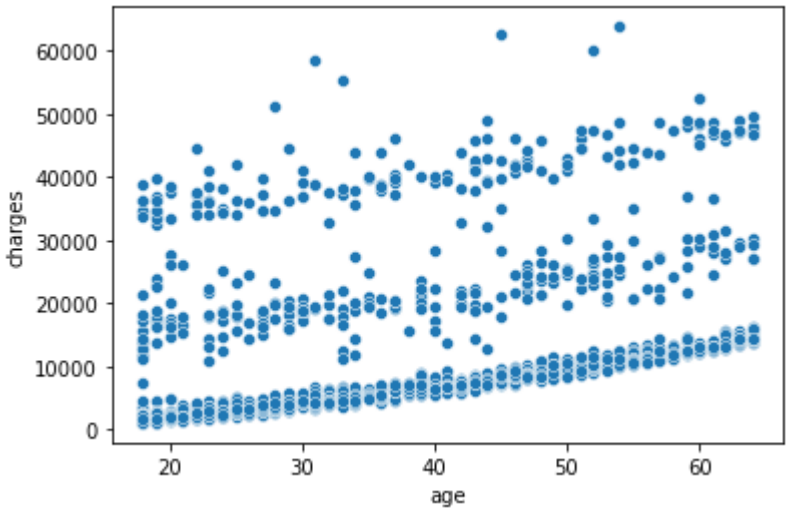
Out[87]:
```
count    9.00
mean    49.28
std      2.24
min     47.41
25%     47.60
50%     48.07
75%     50.38
max     53.13
Name: bmi, dtype: float64
```



Checking some variables how it effect the target(charges)amount

In [88]:
```
sns.scatterplot(x=df.age,y=df.charges)
```

Out[88]: <AxesSubplot:xlabel='age', ylabel='charges'>

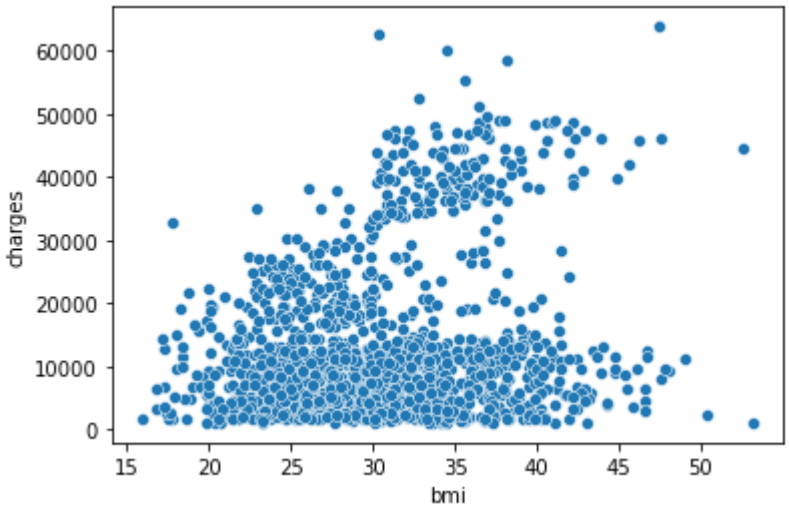


we can see there are 3 lines of positive relationship between age and charges. It means, there are 3 sets of charges which increase gradually with age.

Type *Markdown* and LaTeX: α^2

In [89]:
```
sns.scatterplot(x=df.bmi,y=df.charges)
```

Out[89]: <AxesSubplot:xlabel='bmi', ylabel='charges'>



There is a small positive relation between ‘bmi’ and ‘charges’.

In [90]:
```
df['bmi'].corr(df['charges'])
```

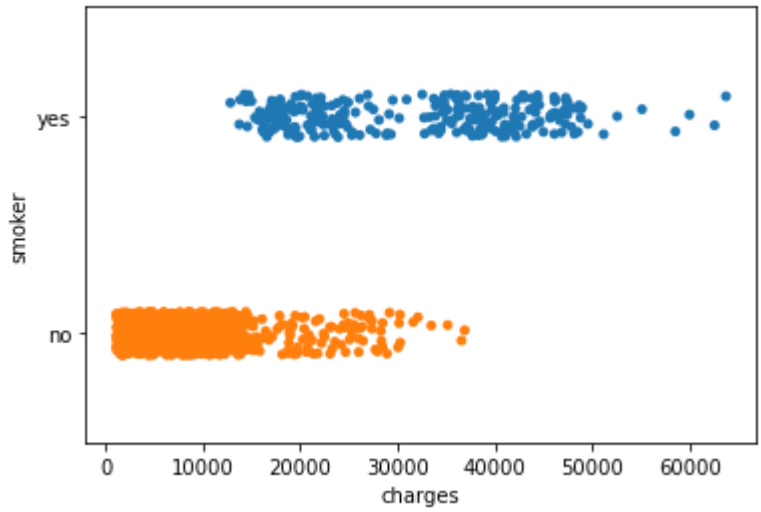
Out[90]: 0.19834096883362884

The correlation coefficient is 0.335 and the two-tailed p-value is .017. Since this p-value is less than .05, we would conclude that there is a statistically significant correlation between the two variables.

Type *Markdown* and LaTeX: α^2

In [91]:
```
sns.stripplot(x=df.charges,y=df.smoker)
```

Out[91]: <AxesSubplot:xlabel='charges', ylabel='smoker'>



People who smoke have high charges and the people who do not smoke have considerably very low charges. So, charges do differ for people who smoke from the people who do not smoke but not significantly as there is some intersection of values for both types of people.

```
Bivariate Analysis for sex Vs. smoker
```

Constructing the Contingency Table The next step is to format the data into a frequency count table. This is called a Contingency Table, we can accomplish this by using the pd.crosstab() function in pandas.

In [92]:
```
contingency_table1 = pd.crosstab(df['sex'],df['smoker'],margins = True)
contingency_table1
```

Out[92]:

smoker	no	yes	All
sex			
female	547	115	662
male	517	159	676
All	1064	274	1338

With Chi-square test understand the relationship between the categorical variable of the

dataset.By using this test we can estimate the level os correlation. this helps us analyze the dependence of the one category of the variable in the other independent category of the variable

In [95]:
```
from scipy.stats import chi2_contingency
def chi_square(c1,c2):
    chi_2, p_val, dof, exp_val = chi2_contingency(pd.crosstab(df[c1],df[c2],m

    print(exp_val)
    print('\nChi-square is : %f'%chi_2, '\n\np_value is : %f'%p_val, '\n\ndeg

    if p_val < 0.05:# consider significan level is 5%
        print("\nThere is some correlation between the two variables at signi
    else:
        print("\nThere is no correlation between the two variables")
```

In [96]:
```
chi_square("sex",'smoker')
```

```
[[526.43348281 135.56651719]
 [537.56651719 138.43348281]]

Chi-square is : 7.392911

p_value is : 0.006548

degree of freedom is : 1

There is some correlation between the two variables at significance level
0.05
```

In [97]:
```
chi_square("region",'sex')
```

```
[[160.30493274 163.69506726]
 [160.79970105 164.20029895]
 [180.09566517 183.90433483]
 [160.79970105 164.20029895]]

Chi-square is : 0.435137

p_value is : 0.932892

degree of freedom is : 3

There is no correlation between the two variables
```

In [98]:
```
pip install nbconvert
```

```
Requirement already satisfied: nbconvert in c:\programdata\anaconda3\lib\si
te-packages (6.0.7)
Requirement already satisfied: jinja2>=2.4 in c:\programdata\anaconda3\lib
\site-packages (from nbconvert) (2.11.2)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\programdata\anaco
nda3\lib\site-packages (from nbconvert) (1.4.3)
Requirement already satisfied: nbclient<0.6.0,>=0.5.0 in c:\programdata\ana
conda3\lib\site-packages (from nbconvert) (0.5.1)
Requirement already satisfied: entrypoints>=0.2.2 in c:\programdata\anacond
a3\lib\site-packages (from nbconvert) (0.3)
Requirement already satisfied: testpath in c:\programdata\anaconda3\lib\sit
e-packages (from nbconvert) (0.4.4)
Requirement already satisfied: pygments>=2.4.1 in c:\programdata\anaconda3
\lib\site-packages (from nbconvert) (2.7.2)
Requirement already satisfied: jupyter-core in c:\programdata\anaconda3\lib
\site-packages (from nbconvert) (4.6.3)
Requirement already satisfied: mistune<2,>=0.8.1 in c:\programdata\anaconda
3\lib\site-packages (from nbconvert) (0.8.4)
Requirement already satisfied: bleach in c:\programdata\anaconda3\lib\site-
packages (from nbconvert) (3.2.1)
Requirement already satisfied: traitlets>=4.2 in c:\programdata\anaconda3\l
ib\site-packages (from nbconvert) (5.0.5)
Requirement already satisfied: nbformat>=4.4 in c:\programdata\anaconda3\li
b\site-packages (from nbconvert) (5.0.8)
Requirement already satisfied: jupyterlab-pygments in c:\programdata\anacon
da3\lib\site-packages (from nbconvert) (0.1.2)
Requirement already satisfied: defusedxml in c:\programdata\anaconda3\lib\s
ite-packages (from nbconvert) (0.6.0)
Requirement already satisfied: MarkupSafe>=0.23 in c:\programdata\anaconda3
\lib\site-packages (from jinja2>=2.4->nbconvert) (1.1.1)
Requirement already satisfied: nest-asyncio in c:\programdata\anaconda3\lib
\site-packages (from nbclient<0.6.0,>=0.5.0->nbconvert) (1.4.2)
Requirement already satisfied: async-generator in c:\programdata\anaconda3
\lib\site-packages (from nbclient<0.6.0,>=0.5.0->nbconvert) (1.10)
Requirement already satisfied: jupyter-client>=6.1.5 in c:\programdata\anac
onda3\lib\site-packages (from nbclient<0.6.0,>=0.5.0->nbconvert) (6.1.7)
Requirement already satisfied: pywin32>=1.0; sys_platform == "win32" in
c:\programdata\anaconda3\lib\site-packages (from jupyter-core->nbconvert)
(227)
Requirement already satisfied: webencodings in c:\programdata\anaconda3\lib
\site-packages (from bleach->nbconvert) (0.5.1)
Requirement already satisfied: packaging in c:\programdata\anaconda3\lib\si
te-packages (from bleach->nbconvert) (20.4)
Requirement already satisfied: six>=1.9.0 in c:\programdata\anaconda3\lib\s
ite-packages (from bleach->nbconvert) (1.15.0)
Requirement already satisfied: ipython-genutils in c:\programdata\anaconda3
\lib\site-packages (from traitlets>=4.2->nbconvert) (0.2.0)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in c:\programdata\an
aconda3\lib\site-packages (from nbformat>=4.4->nbconvert) (3.2.0)
Requirement already satisfied: pyzmq>=13 in c:\programdata\anaconda3\lib\si
te-packages (from jupyter-client>=6.1.5->nbclient<0.6.0,>=0.5.0->nbconvert)
(19.0.2)
Requirement already satisfied: python-dateutil>=2.1 in c:\programdata\anaco
```

```
nda3\lib\site-packages (from jupyter-client>=6.1.5->nbclient<0.6.0,>=0.5.0-
>nbconvert) (2.8.1)
Requirement already satisfied: tornado>=4.1 in c:\programdata\anaconda3\lib
\site-packages (from jupyter-client>=6.1.5->nbclient<0.6.0,>=0.5.0->nbconve
rt) (6.0.4)
Requirement already satisfied: pyparsing>=2.0.2 in c:\programdata\anaconda3
\lib\site-packages (from packaging->bleach->nbconvert) (2.4.7)
Requirement already satisfied: attrs>=17.4.0 in c:\programdata\anaconda3\li
b\site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.4->nbconvert) (2
0.3.0)
Requirement already satisfied: setuptools in c:\programdata\anaconda3\lib\s
ite-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.4->nbconvert) (50.
3.1.post20201107)
Requirement already satisfied: pyrsistent>=0.14.0 in c:\programdata\anacond
a3\lib\site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.4->nbconver
t) (0.17.3)
Note: you may need to restart the kernel to use updated packages.
```