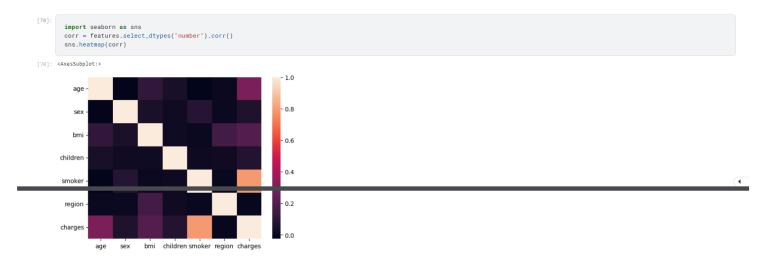


Label Encoding the data

Studying the correlation through heatmap



we see no correlation

```
features.info()
cclass '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 1 int64
1 sex 1338 non-null int64
2 bmi 1338 non-null int64
3 children 1338 non-null int64
4 smoker 1338 non-null int64
5 region 1338 non-null int64
5 charges 1338 non-null int64
6 charges 1338 non-null int64
types: float64(2), int64(5)
memory usage: 73.3 KS
```

Linear regression

```
from sklearn.linear_model import LinearRegression
from sklearn.pipeline import Pipeline, make_pipeline
from sklearn.linear_model import Ridge
model = LinearRegression()
```

features.head()

```
[74]: age sex bmi children smoker region
      0 19 0 27.900 0 1 3 16884.92400
      1 18 1 33.770 1 0 2 1725.55230

    2
    28
    1
    33.000
    3
    0
    2
    4449.46200

    3
    33
    1
    22.705
    0
    0
    1
    21984.47061

       4 32 1 28.880 0 0 1 3866.85520
```

```
target=features["charges"]
            target.head()
[75]: 0 16884.92400
1 1725.55230
2 4449.46200
3 21984.47061
4 3866.85520
Name: charges, dtype: float64
            lr=model.fit(features, target)
            predictions = model.predict(features)
predictions = pd.Series(predictions)
predictions.head()
         0 16884.92400
1 1725.55230
2 4449.46200
3 21984.47061
4 3866.85520
dtype: float64
[77]: 0
[78]:
             from sklearn.metrics import mean_absolute_error as mae
             \verb|print("Mean Absolute Error:", mae(target, predictions))|\\
         Mean Absolute Error: 5.475661113700914e-12
```