

```
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import OneHotEncoder
from sklearn.preprocessing import LabelEncoder
```

```
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
from sklearn.svm import SVR
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import r2_score
```

```
data = pd.read_csv('insurance.csv')
print(data.head(15))
```

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520
5	31	female	25.740	0	no	southeast	3756.62160
6	46	female	33.440	1	no	southeast	8240.58960
7	37	female	27.740	3	no	northwest	7281.50560
8	37	male	29.830	2	no	northeast	6406.41070
9	60	female	25.840	0	no	northwest	28923.13692
10	25	male	26.220	0	no	northeast	2721.32080
11	62	female	26.290	0	yes	southeast	27808.72510
12	23	male	34.400	0	no	southwest	1826.84300
13	56	female	39.820	0	no	southeast	11090.71780
14	27	male	42.130	0	yes	southeast	39611.75770

```
In [23]: import seaborn as sns

figure, ax = plt.subplots(4,2,figsize=(12,24))

sns.distplot(data['charges'], ax = ax[0,0])
sns.distplot(data['age'], ax = ax[0,1])
sns.distplot(data['bmi'], ax = ax[1,0])
sns.distplot(data['children'], ax = ax[1,1])

sns.countplot(data['sex'], ax = ax[2,0])
sns.countplot(data['smoker'], ax = ax[2,1])
sns.countplot(data['region'], ax = ax[3,0])

sns.pairplot(data)

<seaborn.axisgrid.PairGrid at 0x7fe589866c18>
```

