In [ ]:
```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import mean_squared_error, r2_score
import seaborn as sns
import matplotlib.pyplot as plt
```

In [ ]:
```
df=pd.read_csv("insurance.csv",sep=",")
```

In [ ]:
```
df
```

Out[ ]:

| | 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 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1333** | 50 | male | 30.970 | 3 | no | northwest | 10600.54830 |
| **1334** | 18 | female | 31.920 | 0 | no | northeast | 2205.98080 |
| **1335** | 18 | female | 36.850 | 0 | no | southeast | 1629.83350 |
| **1336** | 21 | female | 25.800 | 0 | no | southwest | 2007.94500 |
| **1337** | 61 | female | 29.070 | 0 | yes | northwest | 29141.36030 |

1338 rows × 7 columns

In [ ]:
```
print(df.isna().sum())
```

```
age         0
sex         0
bmi         0
children    0
smoker      0
region      0
charges     0
dtype: int64
```

In [ ]:
```
df=df.drop(["children","region"],axis=1)
```

In [ ]:
```
df
```

Out[ ]:

| | age | sex | bmi | smoker | charges |
|---|---|---|---|---|---|
| **0** | 19 | female | 27.900 | yes | 16884.92400 |
| **1** | 18 | male | 33.770 | no | 1725.55230 |
| **2** | 28 | male | 33.000 | no | 4449.46200 |
| **3** | 33 | male | 22.705 | no | 21984.47061 |
| **4** | 32 | male | 28.880 | no | 3866.85520 |
| **...** | ... | ... | ... | ... | ... |
| **1333** | 50 | male | 30.970 | no | 10600.54830 |
| **1334** | 18 | female | 31.920 | no | 2205.98080 |
| **1335** | 18 | female | 36.850 | no | 1629.83350 |
| **1336** | 21 | female | 25.800 | no | 2007.94500 |
| **1337** | 61 | female | 29.070 | yes | 29141.36030 |

1338 rows × 5 columns

In [ ]:
```
label_encoder=LabelEncoder()
df["sex"]= label_encoder.fit_transform(df["sex"])
```

In [ ]:
```
df['smoker'].unique()
```

Out[ ]:
```
array(['yes', 'no'], dtype=object)
```

In [ ]:
```
print(df['smoker'].value_counts(normalize=True))
```

```
smoker
no     0.795217
yes    0.204783
Name: proportion, dtype: float64
```

In [ ]:
```
df["smoker"]= label_encoder.fit_transform(df["smoker"])
```

In [ ]:
```
df
```

Out[ ]:

| | age | sex | bmi | smoker | charges |
|---|---|---|---|---|---|
| **0** | 19 | 0 | 27.900 | 1 | 16884.92400 |
| **1** | 18 | 1 | 33.770 | 0 | 1725.55230 |
| **2** | 28 | 1 | 33.000 | 0 | 4449.46200 |
| **3** | 33 | 1 | 22.705 | 0 | 21984.47061 |
| **4** | 32 | 1 | 28.880 | 0 | 3866.85520 |
| **...** | ... | ... | ... | ... | ... |
| **1333** | 50 | 1 | 30.970 | 0 | 10600.54830 |
| **1334** | 18 | 0 | 31.920 | 0 | 2205.98080 |
| **1335** | 18 | 0 | 36.850 | 0 | 1629.83350 |
| **1336** | 21 | 0 | 25.800 | 0 | 2007.94500 |
| **1337** | 61 | 0 | 29.070 | 1 | 29141.36030 |

1338 rows × 5 columns

In [ ]:
```
X = df.drop(["charges"],axis=1)
y = df["charges"].values


# Splitting data into training (80%) and testing (20%) sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=11 )

# Initializing the Linear regression model
model = LinearRegression()


# Training the model
model.fit(X_train, y_train)

# Predicting  test set
y_pred = model.predict(X_test)

# Evaluating accuracy
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f"Mean Squared Error: {mse:.2f}")
print(f"R² Score: {r2:.2f}")


# Making predictions for new cases
new_cases = np.array([
    [23,0,78.07,1],
    [23,0,28.06,0]
])
predictions = model.predict(new_cases)
print(f"Predicted : {predictions}")
```

```
Mean Squared Error: 27848938.96
R² Score: 0.80
Predicted : [43485.66944552  3457.91922976]
```

```
C:\Users\Asus\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\site-packages\sklearn\utils\validation.py:2739: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```

In [ ]:
```
results = []

for state in range(1, 201):
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=state)

    model = LinearRegression()
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    mse = mean_squared_error(y_test, y_pred)
    r2 = r2_score(y_test, y_pred)


    results.append((state, mse, r2))


sorted_results = sorted(results, key=lambda x: (x[1], -x[2]))

for state, mse, r2 in sorted_results:
    print(f"Random State: {state} → MSE: {mse:.2f}  R²: {r2:.2f}")
```

```
Random State: 120 → MSE: 27294427.61  R²: 0.78
Random State: 11 → MSE: 27848938.96  R²: 0.80
Random State: 36 → MSE: 28776886.44  R²: 0.75
Random State: 13 → MSE: 28840274.83  R²: 0.80
Random State: 115 → MSE: 29292112.24  R²: 0.78
Random State: 15 → MSE: 29357720.42  R²: 0.77
Random State: 32 → MSE: 29536475.99  R²: 0.78
Random State: 177 → MSE: 29819104.05  R²: 0.75
Random State: 112 → MSE: 29936878.79  R²: 0.80
Random State: 20 → MSE: 30628988.47  R²: 0.80
Random State: 75 → MSE: 31080214.21  R²: 0.77
Random State: 132 → MSE: 31092483.45  R²: 0.76
Random State: 17 → MSE: 31142354.76  R²: 0.72
Random State: 123 → MSE: 31168050.39  R²: 0.80
Random State: 122 → MSE: 31405117.17  R²: 0.80
Random State: 47 → MSE: 31515275.04  R²: 0.79
Random State: 130 → MSE: 31641651.21  R²: 0.72
Random State: 114 → MSE: 31796476.37  R²: 0.74
Random State: 86 → MSE: 32123031.47  R²: 0.78
Random State: 95 → MSE: 32207746.82  R²: 0.77
Random State: 9 → MSE: 32209749.91  R²: 0.77
Random State: 6 → MSE: 32421661.01  R²: 0.78
Random State: 185 → MSE: 32436726.05  R²: 0.77
Random State: 129 → MSE: 32458252.23  R²: 0.78
Random State: 84 → MSE: 32535902.10  R²: 0.76
Random State: 143 → MSE: 32576069.37  R²: 0.76
Random State: 125 → MSE: 32578100.45  R²: 0.77
Random State: 156 → MSE: 32589557.76  R²: 0.76
Random State: 173 → MSE: 32596159.36  R²: 0.72
Random State: 64 → MSE: 32599731.09  R²: 0.76
Random State: 100 → MSE: 32688554.50  R²: 0.79
Random State: 44 → MSE: 32746164.24  R²: 0.75
Random State: 90 → MSE: 32989945.82  R²: 0.75
Random State: 171 → MSE: 33183173.53  R²: 0.73
Random State: 66 → MSE: 33335542.42  R²: 0.78
Random State: 175 → MSE: 33425552.93  R²: 0.74
Random State: 194 → MSE: 33445523.40  R²: 0.77
Random State: 118 → MSE: 33462646.66  R²: 0.79
Random State: 5 → MSE: 33544222.56  R²: 0.78
Random State: 94 → MSE: 33585302.72  R²: 0.76
Random State: 25 → MSE: 33638125.12  R²: 0.76
Random State: 46 → MSE: 33718061.57  R²: 0.77
```

```
Random State: 182 → MSE: 33775040.48  R²: 0.78
Random State: 108 → MSE: 33837766.19  R²: 0.75
Random State: 37 → MSE: 33840127.01  R²: 0.71
Random State: 76 → MSE: 33853214.26  R²: 0.76
Random State: 178 → MSE: 33870809.77  R²: 0.76
Random State: 137 → MSE: 33943362.36  R²: 0.76
Random State: 157 → MSE: 34052410.76  R²: 0.76
Random State: 23 → MSE: 34120056.16  R²: 0.76
Random State: 59 → MSE: 34121833.39  R²: 0.80
Random State: 101 → MSE: 34199937.06  R²: 0.76
Random State: 159 → MSE: 34261691.10  R²: 0.77
Random State: 187 → MSE: 34289994.97  R²: 0.73
Random State: 105 → MSE: 34322952.74  R²: 0.78
Random State: 53 → MSE: 34369136.13  R²: 0.77
Random State: 141 → MSE: 34383115.95  R²: 0.70
Random State: 57 → MSE: 34501198.44  R²: 0.77
Random State: 42 → MSE: 34515553.67  R²: 0.78
Random State: 49 → MSE: 34540649.03  R²: 0.77
Random State: 72 → MSE: 34574081.62  R²: 0.74
Random State: 160 → MSE: 34723126.58  R²: 0.74
Random State: 110 → MSE: 34728617.64  R²: 0.75
Random State: 113 → MSE: 34735659.44  R²: 0.76
Random State: 93 → MSE: 34801234.15  R²: 0.76
Random State: 189 → MSE: 34890034.49  R²: 0.74
Random State: 24 → MSE: 34973867.31  R²: 0.77
Random State: 111 → MSE: 34981637.27  R²: 0.79
Random State: 50 → MSE: 35115198.05  R²: 0.78
Random State: 79 → MSE: 35196996.56  R²: 0.74
Random State: 68 → MSE: 35235189.77  R²: 0.77
Random State: 131 → MSE: 35453554.90  R²: 0.75
Random State: 144 → MSE: 35478219.86  R²: 0.74
Random State: 127 → MSE: 35535015.87  R²: 0.73
Random State: 80 → MSE: 35562363.79  R²: 0.75
Random State: 167 → MSE: 35569938.11  R²: 0.72
Random State: 163 → MSE: 35618481.87  R²: 0.72
Random State: 121 → MSE: 35658248.94  R²: 0.76
Random State: 41 → MSE: 35719727.77  R²: 0.68
Random State: 169 → MSE: 35728155.74  R²: 0.78
Random State: 179 → MSE: 35817773.22  R²: 0.78
Random State: 107 → MSE: 35861745.46  R²: 0.71
Random State: 154 → MSE: 35913070.47  R²: 0.74
Random State: 180 → MSE: 35922246.67  R²: 0.74
```

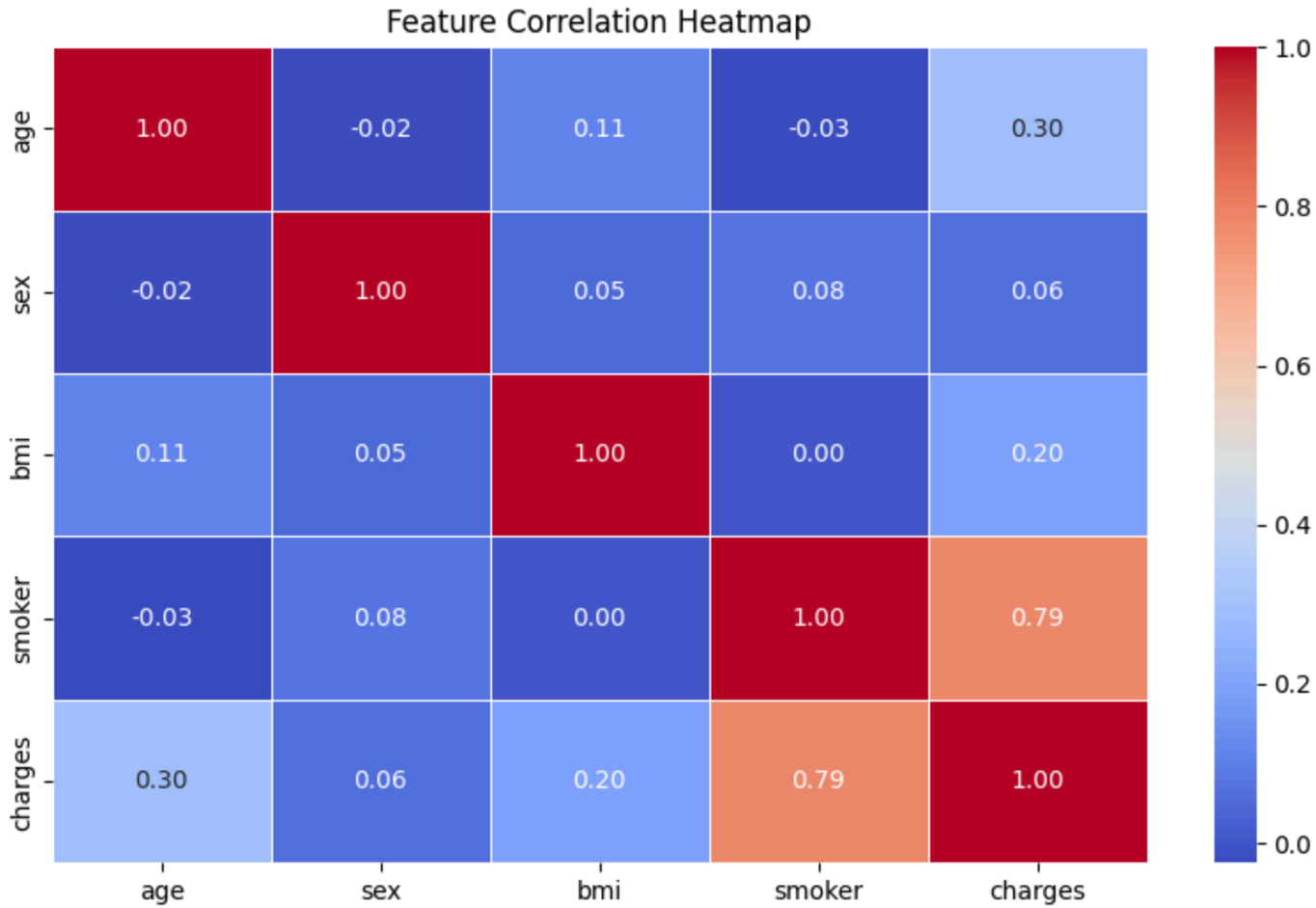
```
Random State: 63 → MSE: 35985128.06  R²: 0.73
Random State: 117 → MSE: 36210463.98  R²: 0.75
Random State: 78 → MSE: 36236492.73  R²: 0.73
Random State: 1 → MSE: 36348044.06  R²: 0.76
Random State: 22 → MSE: 36380593.81  R²: 0.75
Random State: 198 → MSE: 36429547.48  R²: 0.77
Random State: 62 → MSE: 36546696.13  R²: 0.68
Random State: 99 → MSE: 36581124.52  R²: 0.74
Random State: 77 → MSE: 36586920.22  R²: 0.74
Random State: 87 → MSE: 36696371.90  R²: 0.71
Random State: 200 → MSE: 36720378.85  R²: 0.77
Random State: 138 → MSE: 36881018.27  R²: 0.79
Random State: 38 → MSE: 36958294.64  R²: 0.71
Random State: 192 → MSE: 36971476.57  R²: 0.72
Random State: 128 → MSE: 37007519.39  R²: 0.78
Random State: 33 → MSE: 37053948.46  R²: 0.70
Random State: 165 → MSE: 37113094.19  R²: 0.75
Random State: 56 → MSE: 37167420.70  R²: 0.70
Random State: 21 → MSE: 37262392.11  R²: 0.73
Random State: 18 → MSE: 37265812.70  R²: 0.76
Random State: 88 → MSE: 37285586.59  R²: 0.77
Random State: 103 → MSE: 37381398.23  R²: 0.78
Random State: 183 → MSE: 37423330.15  R²: 0.74
Random State: 98 → MSE: 37427036.66  R²: 0.73
Random State: 136 → MSE: 37444263.41  R²: 0.73
Random State: 148 → MSE: 37564155.53  R²: 0.73
Random State: 30 → MSE: 37655437.06  R²: 0.76
Random State: 161 → MSE: 37740855.81  R²: 0.74
Random State: 14 → MSE: 37796917.50  R²: 0.73
Random State: 166 → MSE: 37855874.73  R²: 0.74
Random State: 124 → MSE: 37907710.13  R²: 0.75
Random State: 147 → MSE: 37931058.36  R²: 0.76
Random State: 168 → MSE: 37998493.68  R²: 0.77
Random State: 8 → MSE: 38032093.72  R²: 0.72
Random State: 28 → MSE: 38116188.49  R²: 0.70
Random State: 126 → MSE: 38119611.52  R²: 0.74
Random State: 2 → MSE: 38169988.40  R²: 0.75
Random State: 7 → MSE: 38196771.50  R²: 0.76
Random State: 82 → MSE: 38239630.82  R²: 0.75
Random State: 164 → MSE: 38296569.46  R²: 0.77
Random State: 191 → MSE: 38329531.43  R²: 0.74
Random State: 89 → MSE: 38384905.50  R²: 0.78
```

```
Random State: 134 → MSE: 38516593.08  R²: 0.76
Random State: 71 → MSE: 38765906.04  R²: 0.75
Random State: 58 → MSE: 38774365.24  R²: 0.75
Random State: 170 → MSE: 38813062.17  R²: 0.75
Random State: 45 → MSE: 38865732.19  R²: 0.70
Random State: 3 → MSE: 38940418.53  R²: 0.74
Random State: 162 → MSE: 38959583.46  R²: 0.76
Random State: 139 → MSE: 39071738.56  R²: 0.75
Random State: 196 → MSE: 39073866.42  R²: 0.74
Random State: 172 → MSE: 39114170.67  R²: 0.73
Random State: 96 → MSE: 39165452.47  R²: 0.66
Random State: 146 → MSE: 39213340.47  R²: 0.74
Random State: 83 → MSE: 39347449.18  R²: 0.76
Random State: 140 → MSE: 39363901.97  R²: 0.75
Random State: 4 → MSE: 39379442.87  R²: 0.69
Random State: 133 → MSE: 39405078.83  R²: 0.68
Random State: 102 → MSE: 39410357.00  R²: 0.72
Random State: 199 → MSE: 39426529.27  R²: 0.72
Random State: 52 → MSE: 39507400.82  R²: 0.69
Random State: 145 → MSE: 39508621.54  R²: 0.76
Random State: 116 → MSE: 39527732.99  R²: 0.77
Random State: 149 → MSE: 39628953.15  R²: 0.78
Random State: 40 → MSE: 39711574.80  R²: 0.73
Random State: 31 → MSE: 39755406.46  R²: 0.70
Random State: 67 → MSE: 39795845.14  R²: 0.74
Random State: 16 → MSE: 39843237.49  R²: 0.71
Random State: 184 → MSE: 39898206.83  R²: 0.77
Random State: 109 → MSE: 40058737.52  R²: 0.69
Random State: 190 → MSE: 40241951.50  R²: 0.70
Random State: 193 → MSE: 40343902.71  R²: 0.71
Random State: 81 → MSE: 40495810.77  R²: 0.72
Random State: 74 → MSE: 40560078.31  R²: 0.73
Random State: 27 → MSE: 40631631.28  R²: 0.72
Random State: 91 → MSE: 40634994.40  R²: 0.74
Random State: 153 → MSE: 40642411.40  R²: 0.70
Random State: 106 → MSE: 40795532.05  R²: 0.75
Random State: 39 → MSE: 40849494.22  R²: 0.74
Random State: 48 → MSE: 40973346.46  R²: 0.75
Random State: 26 → MSE: 41019664.33  R²: 0.73
Random State: 174 → MSE: 41020154.59  R²: 0.70
Random State: 85 → MSE: 41082447.97  R²: 0.72
Random State: 158 → MSE: 41110483.95  R²: 0.67
```

```
Random State: 104 → MSE: 41268026.19  R²: 0.71
Random State: 55 → MSE: 41347454.21  R²: 0.69
Random State: 65 → MSE: 41386213.54  R²: 0.73
Random State: 155 → MSE: 41472820.49  R²: 0.73
Random State: 19 → MSE: 41601834.52  R²: 0.75
Random State: 181 → MSE: 41921558.08  R²: 0.73
Random State: 34 → MSE: 42090165.41  R²: 0.74
Random State: 61 → MSE: 42124886.58  R²: 0.72
Random State: 186 → MSE: 42233419.97  R²: 0.73
Random State: 92 → MSE: 42252269.94  R²: 0.76
Random State: 73 → MSE: 42742886.49  R²: 0.71
Random State: 12 → MSE: 42910453.65  R²: 0.70
Random State: 10 → MSE: 42940195.80  R²: 0.69
Random State: 97 → MSE: 42944892.84  R²: 0.70
Random State: 60 → MSE: 43012277.83  R²: 0.66
Random State: 135 → MSE: 43041189.89  R²: 0.69
Random State: 142 → MSE: 43237481.22  R²: 0.74
Random State: 151 → MSE: 43317361.69  R²: 0.75
Random State: 70 → MSE: 43320323.29  R²: 0.67
Random State: 43 → MSE: 43358068.00  R²: 0.69
Random State: 188 → MSE: 43425508.71  R²: 0.73
Random State: 119 → MSE: 43546035.99  R²: 0.74
Random State: 35 → MSE: 43558906.13  R²: 0.72
Random State: 51 → MSE: 43840027.34  R²: 0.75
Random State: 195 → MSE: 44035167.90  R²: 0.70
Random State: 69 → MSE: 44115669.26  R²: 0.74
Random State: 197 → MSE: 44301270.81  R²: 0.71
Random State: 54 → MSE: 44851005.76  R²: 0.65
Random State: 29 → MSE: 45634067.83  R²: 0.71
Random State: 150 → MSE: 46364688.88  R²: 0.72
Random State: 152 → MSE: 46569818.75  R²: 0.75
Random State: 176 → MSE: 48785976.99  R²: 0.70
```

In [ ]:

```python
corr_matrix = df.corr()
plt.figure(figsize=(10, 6))
sns.heatmap(corr_matrix, annot=True, fmt=".2f", cmap="coolwarm", linewidths=0.5)
plt.title("Feature Correlation Heatmap")
plt.show()
```


Feature Correlation Heatmap
age
sex
bmi
smoker
charges
1.00
-0.02
0.11
-0.03
0.30
-0.02
1.00
0.05
0.08
0.06
0.11
0.05
1.00
0.00
0.20
-0.03
0.08
0.00
1.00
0.79
0.30
0.06
0.20
0.79
1.00
age
sex
bmi
smoker
charges
1.0
0.8
0.6
0.4
0.2
0.0