Code for exp2:-

```
# ========================================
# STEP 1: Import Libraries
# ========================================

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression, Ridge, Lasso
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from sklearn.preprocessing import LabelEncoder, StandardScaler


# ========================================
# STEP 2: Load Dataset
# ========================================

df = pd.read_csv("/content/healthcare-dataset-stroke-data.csv")
print(df.head())


# ========================================
# STEP 3: Data Cleaning
# ========================================

# Drop unnecessary column
df.drop("id", axis=1, inplace=True)

# Fill missing BMI
df["bmi"].fillna(df["bmi"].mean(), inplace=True)


# ========================================
# STEP 4: Encode Categorical Data
# ========================================

le = LabelEncoder()

cat_cols = ["gender","ever_married","work_type","Residence_type","smoking_status"]
```

```
for col in cat_cols:
    df[col] = le.fit_transform(df[col])


# ========================================
# STEP 5: Feature & Target Selection
# ========================================

X = df.drop("bmi", axis=1)   # features
y = df["bmi"]                # target


# ========================================
# STEP 6: Train-Test Split
# ========================================

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)


# ========================================
# STEP 7: Feature Scaling (IMPORTANT)
# ========================================

scaler = StandardScaler()

X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)


# ========================================
# STEP 8: MULTIPLE LINEAR REGRESSION
# ========================================

lr = LinearRegression()
lr.fit(X_train, y_train)

y_pred_lr = lr.predict(X_test)

print("\n===== MULTIPLE LINEAR REGRESSION =====")
print("MAE:", mean_absolute_error(y_test, y_pred_lr))
print("MSE:", mean_squared_error(y_test, y_pred_lr))
print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred_lr)))
```

```
print("R2 Score:", r2_score(y_test, y_pred_lr))


# ========================================
# STEP 9: RIDGE REGRESSION
# ========================================

ridge = Ridge(alpha=1.0)   # alpha = regularization strength
ridge.fit(X_train, y_train)

y_pred_ridge = ridge.predict(X_test)

print("\n===== RIDGE REGRESSION =====")
print("MAE:", mean_absolute_error(y_test, y_pred_ridge))
print("MSE:", mean_squared_error(y_test, y_pred_ridge))
print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred_ridge)))
print("R2 Score:", r2_score(y_test, y_pred_ridge))


# ========================================
# STEP 10: LASSO REGRESSION
# ========================================

lasso = Lasso(alpha=0.1)
lasso.fit(X_train, y_train)

y_pred_lasso = lasso.predict(X_test)

print("\n===== LASSO REGRESSION =====")
print("MAE:", mean_absolute_error(y_test, y_pred_lasso))
print("MSE:", mean_squared_error(y_test, y_pred_lasso))
print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred_lasso)))
print("R2 Score:", r2_score(y_test, y_pred_lasso))


# ========================================
# STEP 11: MODEL COMPARISON
# ========================================

models = ["Linear", "Ridge", "Lasso"]
r2_scores = [
    r2_score(y_test, y_pred_lr),
    r2_score(y_test, y_pred_ridge),
    r2_score(y_test, y_pred_lasso)
```

```
]

plt.bar(models, r2_scores)
plt.title("Model Comparison (R2 Score)")
plt.xlabel("Models")
plt.ylabel("R2 Score")
plt.show()


# ========================================
# STEP 12: COEFFICIENT COMPARISON
# ========================================

print("\nFeature Coefficients:")

coeff_df = pd.DataFrame({
    "Feature": X.columns,
    "Linear": lr.coef_,
    "Ridge": ridge.coef_,
    "Lasso": lasso.coef_
})

print(coeff_df)
```