```
import pandas as pd
import numpy as np
import os
import warnings
from sklearn.ensemble import RandomForestRegressor
from xgboost import XGBRegressor
from sklearn.model_selection import cross_val_score
from sklearn.preprocessing import LabelEncoder
import plotly.express as px
from tqdm import tqdm
import xgboost as xgb
from tqdm.auto import tqdm
from google.colab import drive
import os
import time
import plotly.graph_objects as go
import matplotlib.pyplot as plt
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
pd.set_option('display.float_format', '{:.2f}'.format)
```

```
!ls "/content/drive/My Drive/EFasmaa"
```

```
df3_soum_data.csv
```

```
df3=pd.read_csv('/content/drive/My Drive/FEasmaa/df4_saved.csv')
```

```
df3.shape
```

```
(900000, 105)
```

1-Blood Pressure Features

Pulse Pressure

```
df3["Pulse_Pressure"] = df3["Systolic Pressure"] - df3["Diastolic Pressure"]
```

$$PP = \text{Systolic Pressure} - \text{Diastolic Pressure}$$

Mean Arterial Pressure (MAP)

```
df3["MAP"] = (df3["Systolic Pressure"] + 2 * df3["Diastolic Pressure"]) / 3
```

$$MAP = \frac{2 \times \text{Diastolic Pressure} + \text{Systolic Pressure}}{3}$$

BP Ratio

```
df3["BP_Ratio"] = df3["Systolic Pressure"] / df3["Diastolic Pressure"]
```

$$BP\ Ratio = \frac{\text{Diastolic Pressure}}{\text{Systolic Pressure}}$$

2-Kidney Function Features

```
df3["Urea_Creatinine_Ratio"] = df3["Urea in Serum"] / df3["Creatinine in Serum"
```

$$\frac{\text{Urea in Serum}}{\text{Creatinine in Serum}}$$

```
df3["BUN_eGFR_Ratio"] = df3["Blood Urea Nitrogen (BUN)"] / df3["Estimated Glome
```

$$\frac{\text{BUN}}{\text{eGFR}}$$

3. Lipid Profile Features

```
df3["Non_HDL_Cholesterol"] = df3["Cholesterol"] - df3["HDL Cholesterol"]
```

$$\text{Total Cholesterol} - \text{HDL Cholesterol}$$

```
df3["TG_HDL_Ratio"] = df3["Triglycerides (TG) in Serum"] / df3["HDL Cholesterol
```

$$\frac{\text{Triglycerides}}{\text{HDL Cholesterol}}$$

4. Glucose Control Features

```
df3["EAG"] = (28.7 * df3["Hb A1c %"]) - 46.7
```

$$EAG = (28.7 \times \text{Hb A1c}) - 46.7$$

```
df3["Glucose_HbA1c_Ratio"] = df3["Glucose in Plasma (Fasting)"] / df3["Hb A1c %
```

$$\frac{\text{Fasting Glucose}}{\text{Hb A1c \%}}$$

```
df3.isnull().sum().sum()
```

```
np.int64(37576155)
```

```
df3.shape
```

```
(900000, 114)
```