

# Diabetes Research Code

## Load Required library

```
library(readr)
```

Warning: package 'readr' was built under R version 4.3.3

```
library(car)
```

Warning: package 'car' was built under R version 4.3.3

Loading required package: carData

Warning: package 'carData' was built under R version 4.3.2

```
library(ggplot2)
```

Warning: package 'ggplot2' was built under R version 4.3.3

```
library(dplyr)
```

Warning: package 'dplyr' was built under R version 4.3.3

Attaching package: 'dplyr'

The following object is masked from 'package:car':

recode

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(caret)
```

Warning: package 'caret' was built under R version 4.3.3

Loading required package: lattice

Warning: package 'lattice' was built under R version 4.3.3

```
library(corrplot)
```

Warning: package 'corrplot' was built under R version 4.3.3

corrplot 0.95 loaded

```
library(pROC)
```

Warning: package 'pROC' was built under R version 4.3.3

Type 'citation("pROC")' for a citation.

Attaching package: 'pROC'

The following objects are masked from 'package:stats':

cov, smooth, var

```
set.seed(140)
```

## Load Dataset

```
dataset <- read_csv("diabetes_dataset.csv")
```

Rows: 9538 Columns: 17

-- Column specification -----

Delimiter: ","

dbl (17): Age, Pregnancies, BMI, Glucose, BloodPressure, HbA1c, LDL, HDL, Tr...

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

## Preview Dataset

```
head(dataset)
```

# A tibble: 6 x 17

	Age	Pregnancies	BMI	Glucose	BloodPressure	HbA1c	LDL	HDL	Triglycerides
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	69	5	28.4	130.	77	5.4	130.	44	50
2	32	1	26.5	116.	72	4.5	87.4	54.2	130.
3	89	13	25.3	101	82	4.9	112.	56.8	178.
4	78	13	29.9	146	104	5.7	50.7	39.1	117
5	38	8	24.6	103.	74	4.7	102.	29.1	146.
6	41	10	17.5	67	71	4.2	105.	58.8	141.

# i 8 more variables: WaistCircumference <dbl>, HipCircumference <dbl>,

# WHR <dbl>, FamilyHistory <dbl>, DietType <dbl>, Hypertension <dbl>,

# MedicationUse <dbl>, Outcome <dbl>

## Run initial model

```
model <- glm(Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL + HDL + Triglycerides)
```

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
summary(model)
```

Call:

```
glm(formula = Outcome ~ Age + Pregnancies + BMI + BloodPressure +  
     HbA1c + LDL + HDL + Triglycerides + WaistCircumference +  
     HipCircumference + WHR + factor(FamilyHistory) + factor(DietType) +  
     factor(Hypertension) + factor(MedicationUse), family = "binomial",  
     data = dataset)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-3.180e+01	4.697e+00	-6.772	1.27e-11	***
Age	4.810e-02	8.786e-03	5.475	4.38e-08	***
Pregnancies	-7.460e-03	1.480e-02	-0.504	0.614	
BMI	1.725e-01	2.615e-02	6.597	4.20e-11	***
BloodPressure	-1.992e-03	7.387e-03	-0.270	0.787	
HbA1c	4.956e+00	2.530e-01	19.592	< 2e-16	***
LDL	2.057e-03	2.410e-03	0.853	0.393	
HDL	-4.771e-04	4.762e-03	-0.100	0.920	
Triglycerides	1.823e-03	1.438e-03	1.268	0.205	
WaistCircumference	4.643e-02	4.347e-02	1.068	0.285	
HipCircumference	-4.784e-02	4.120e-02	-1.161	0.246	
WHR	-4.682e+00	4.653e+00	-1.006	0.314	
factor(FamilyHistory)1	3.066e+01	4.348e+02	0.071	0.944	
factor(DietType)1	-5.074e-02	1.614e-01	-0.314	0.753	
factor(DietType)2	-4.974e-02	2.570e-01	-0.194	0.847	
factor(Hypertension)1	1.364e-01	1.304e+00	0.105	0.917	
factor(MedicationUse)1	-7.161e-02	3.000e-01	-0.239	0.811	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 12279.5 on 9537 degrees of freedom  
Residual deviance: 1336.2 on 9521 degrees of freedom  
AIC: 1370.2

Number of Fisher Scoring iterations: 20

```
vif(model)
```

	GVIF	Df	GVIF^(1/(2*Df))
Age	3.670190	1	1.915774
Pregnancies	1.006458	1	1.003224
BMI	3.173633	1	1.781469
BloodPressure	1.595918	1	1.263296
HbA1c	1.025278	1	1.012560
LDL	1.007540	1	1.003763
HDL	1.009917	1	1.004946
Triglycerides	1.019409	1	1.009658
WaistCircumference	74.626529	1	8.638665
HipCircumference	47.807387	1	6.914289
WHR	67.792078	1	8.233594
factor(FamilyHistory)	1.000001	1	1.000001
factor(DietType)	1.025875	2	1.006407
factor(Hypertension)	1.054709	1	1.026990
factor(MedicationUse)	3.357799	1	1.832430

## AIC - Backward

```
backward_model <- step(model, direction = "backward")
```

Start: AIC=1370.16

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(DietType) + factor(Hypertension) +  
factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(DietType)	2	1336.3	1366.3
- HDL	1	1336.2	1368.2
- factor(Hypertension)	1	1336.2	1368.2
- factor(MedicationUse)	1	1336.2	1368.2
- BloodPressure	1	1336.2	1368.2
- Pregnancies	1	1336.4	1368.4
- LDL	1	1336.9	1368.9
- WHR	1	1337.2	1369.2
- WaistCircumference	1	1337.3	1369.3
- HipCircumference	1	1337.5	1369.5
- Triglycerides	1	1337.8	1369.8
<none>		1336.2	1370.2
- Age	1	1368.0	1400.0
- BMI	1	1381.8	1413.8
- HbA1c	1	2004.0	2036.0
- factor(FamilyHistory)	1	12024.3	12056.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1366.28

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(Hypertension) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(Hypertension)	1	1336.3	1364.3
- HDL	1	1336.3	1364.3
- factor(MedicationUse)	1	1336.3	1364.3
- BloodPressure	1	1336.3	1364.3
- Pregnancies	1	1336.5	1364.5
- LDL	1	1337.0	1365.0
- WHR	1	1337.3	1365.3
- WaistCircumference	1	1337.4	1365.4
- HipCircumference	1	1337.6	1365.6
- Triglycerides	1	1337.9	1365.9
<none>		1336.3	1366.3
- Age	1	1368.2	1396.2
- BMI	1	1382.1	1410.1
- HbA1c	1	2004.7	2032.7
- factor(FamilyHistory)	1	12026.0	12054.0

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1364.28

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- HDL	1	1336.3	1362.3
- factor(MedicationUse)	1	1336.3	1362.3
- BloodPressure	1	1336.4	1362.4

- Pregnancies	1	1336.5	1362.5
- LDL	1	1337.0	1363.0
- WHR	1	1337.3	1363.3
- WaistCircumference	1	1337.4	1363.4
- HipCircumference	1	1337.6	1363.6
- Triglycerides	1	1337.9	1363.9
<none>		1336.3	1364.3
- Age	1	1368.2	1394.2
- BMI	1	1382.3	1408.3
- HbA1c	1	2004.7	2030.7
- factor(FamilyHistory)	1	12028.9	12054.9

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1362.29

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
 Triglycerides + WaistCircumference + HipCircumference + WHR +  
 factor(FamilyHistory) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
 Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(MedicationUse)	1	1336.3	1360.3
- BloodPressure	1	1336.4	1360.4
- Pregnancies	1	1336.6	1360.6
- LDL	1	1337.0	1361.0
- WHR	1	1337.3	1361.3
- WaistCircumference	1	1337.5	1361.5
- HipCircumference	1	1337.7	1361.7
- Triglycerides	1	1337.9	1361.9
<none>		1336.3	1362.3



```

- Age                1    1368.2  1392.2
- BMI                1    1382.3  1406.3
- HbA1c              1    2005.0  2029.0
- factor(FamilyHistory) 1  12029.1 12053.1

```

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1360.35

```

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +
          Triglycerides + WaistCircumference + HipCircumference + WHR +
          factor(FamilyHistory)

```

```

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```

	Df	Deviance	AIC
- BloodPressure	1	1336.4	1358.4
- Pregnancies	1	1336.6	1358.6
- LDL	1	1337.1	1359.1
- WHR	1	1337.4	1359.4
- WaistCircumference	1	1337.5	1359.5
- HipCircumference	1	1337.7	1359.7
- Triglycerides	1	1337.9	1359.9
<none>		1336.3	1360.3
- BMI	1	1382.4	1404.4
- Age	1	1424.3	1446.3
- HbA1c	1	2005.3	2027.3
- factor(FamilyHistory)	1	12030.3	12052.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1358.41

Outcome ~ Age + Pregnancies + BMI + HbA1c + LDL + Triglycerides +  
WaistCircumference + HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- Pregnancies	1	1336.7	1356.7
- LDL	1	1337.2	1357.2
- WHR	1	1337.4	1357.4
- WaistCircumference	1	1337.6	1357.6
- HipCircumference	1	1337.8	1357.8
- Triglycerides	1	1338.0	1358.0
<none>		1336.4	1358.4
- BMI	1	1390.7	1410.7
- Age	1	1443.5	1463.5
- HbA1c	1	2005.4	2025.4
- factor(FamilyHistory)	1	12030.5	12050.5

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1356.67

Outcome ~ Age + BMI + HbA1c + LDL + Triglycerides + WaistCircumference +  
HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- LDL	1	1337.4	1355.4
- WHR	1	1337.7	1355.7
- WaistCircumference	1	1337.8	1355.8
- HipCircumference	1	1338.0	1356.0
- Triglycerides	1	1338.3	1356.3
<none>		1336.7	1356.7
- BMI	1	1391.0	1409.0
- Age	1	1444.1	1462.1
- HbA1c	1	2005.4	2023.4
- factor(FamilyHistory)	1	12033.9	12051.9

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1355.42

Outcome ~ Age + BMI + HbA1c + Triglycerides + WaistCircumference +  
HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- WHR	1	1338.4	1354.4
- WaistCircumference	1	1338.5	1354.5
- HipCircumference	1	1338.7	1354.7
- Triglycerides	1	1339.0	1355.0
<none>		1337.4	1355.4
- BMI	1	1391.9	1407.9
- Age	1	1444.9	1460.9
- HbA1c	1	2005.7	2021.7
- factor(FamilyHistory)	1	12036.8	12052.8

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1354.42

Outcome ~ Age + BMI + HbA1c + Triglycerides + WaistCircumference +  
HipCircumference + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- WaistCircumference	1	1338.6	1352.6
- HipCircumference	1	1339.3	1353.3
- Triglycerides	1	1340.0	1354.0
<none>		1338.4	1354.4
- BMI	1	1393.2	1407.2
- Age	1	1446.1	1460.1
- HbA1c	1	2006.1	2020.1
- factor(FamilyHistory)	1	12037.5	12051.5

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1352.6

Outcome ~ Age + BMI + HbA1c + Triglycerides + HipCircumference +  
factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- HipCircumference	1	1339.4	1351.4
- Triglycerides	1	1340.2	1352.2
<none>		1338.6	1352.6
- BMI	1	1432.0	1444.0
- Age	1	1446.5	1458.5
- HbA1c	1	2006.2	2018.2
- factor(FamilyHistory)	1	12038.2	12050.2

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1351.44

Outcome ~ Age + BMI + HbA1c + Triglycerides + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- Triglycerides	1	1341.0	1351.0
<none>		1339.4	1351.4
- Age	1	1447.2	1457.2
- BMI	1	1469.5	1479.5
- HbA1c	1	2007.3	2017.3
- factor(FamilyHistory)	1	12040.3	12050.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1351.03

Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
<none>		1341.0	1351.0
- Age	1	1449.0	1457.0
- BMI	1	1470.9	1478.9
- HbA1c	1	2007.7	2015.7
- factor(FamilyHistory)	1	12040.7	12048.7

```
summary(backward_model)
```

Call:

glm(formula = Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory),

```
family = "binomial", data = dataset)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-35.951179	1.471747	-24.428	<2e-16 ***
Age	0.045863	0.004829	9.497	<2e-16 ***
BMI	0.166579	0.015576	10.695	<2e-16 ***
HbA1c	4.906516	0.249002	19.705	<2e-16 ***
factor(FamilyHistory)1	30.631935	435.532368	0.070	0.944

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 12280 on 9537 degrees of freedom  
Residual deviance: 1341 on 9533 degrees of freedom  
AIC: 1351

Number of Fisher Scoring iterations: 20

```
vif(backward_model)
```

	Age	BMI	HbA1c
	1.115542	1.117578	1.002088
factor(FamilyHistory)	1.000001		

## AIC - Bidirectional

```
both_model <- step(model, direction = "both")
```

Start: AIC=1370.16

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(DietType) + factor(Hypertension) +  
factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(DietType)	2	1336.3	1366.3
- HDL	1	1336.2	1368.2
- factor(Hypertension)	1	1336.2	1368.2
- factor(MedicationUse)	1	1336.2	1368.2
- BloodPressure	1	1336.2	1368.2
- Pregnancies	1	1336.4	1368.4
- LDL	1	1336.9	1368.9
- WHR	1	1337.2	1369.2
- WaistCircumference	1	1337.3	1369.3
- HipCircumference	1	1337.5	1369.5
- Triglycerides	1	1337.8	1369.8
<none>		1336.2	1370.2
- Age	1	1368.0	1400.0
- BMI	1	1381.8	1413.8
- HbA1c	1	2004.0	2036.0
- factor(FamilyHistory)	1	12024.3	12056.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1366.28

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(Hypertension) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(Hypertension)	1	1336.3	1364.3
- HDL	1	1336.3	1364.3
- factor(MedicationUse)	1	1336.3	1364.3
- BloodPressure	1	1336.3	1364.3
- Pregnancies	1	1336.5	1364.5
- LDL	1	1337.0	1365.0
- WHR	1	1337.3	1365.3
- WaistCircumference	1	1337.4	1365.4
- HipCircumference	1	1337.6	1365.6
- Triglycerides	1	1337.9	1365.9
<none>		1336.3	1366.3
+ factor(DietType)	2	1336.2	1370.2
- Age	1	1368.2	1396.2
- BMI	1	1382.1	1410.1
- HbA1c	1	2004.7	2032.7
- factor(FamilyHistory)	1	12026.0	12054.0

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1364.28

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred



Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- HDL	1	1336.3	1362.3
- factor(MedicationUse)	1	1336.3	1362.3
- BloodPressure	1	1336.4	1362.4
- Pregnancies	1	1336.5	1362.5
- LDL	1	1337.0	1363.0
- WHR	1	1337.3	1363.3
- WaistCircumference	1	1337.4	1363.4
- HipCircumference	1	1337.6	1363.6
- Triglycerides	1	1337.9	1363.9
<none>		1336.3	1364.3
+ factor(Hypertension)	1	1336.3	1366.3
+ factor(DietType)	2	1336.2	1368.2
- Age	1	1368.2	1394.2
- BMI	1	1382.3	1408.3
- HbA1c	1	2004.7	2030.7
- factor(FamilyHistory)	1	12028.9	12054.9

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1362.29

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
Triglycerides + WaistCircumference + HipCircumference + WHR +  
factor(FamilyHistory) + factor(MedicationUse)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- factor(MedicationUse)	1	1336.3	1360.3
- BloodPressure	1	1336.4	1360.4
- Pregnancies	1	1336.6	1360.6
- LDL	1	1337.0	1361.0
- WHR	1	1337.3	1361.3
- WaistCircumference	1	1337.5	1361.5
- HipCircumference	1	1337.7	1361.7
- Triglycerides	1	1337.9	1361.9
<none>		1336.3	1362.3
+ HDL	1	1336.3	1364.3
+ factor(Hypertension)	1	1336.3	1364.3
+ factor(DietType)	2	1336.2	1366.2
- Age	1	1368.2	1392.2
- BMI	1	1382.3	1406.3
- HbA1c	1	2005.0	2029.0
- factor(FamilyHistory)	1	12029.1	12053.1

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1360.35

Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
Triglycerides + WaistCircumference + HipCircumference + WHR +  
factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- BloodPressure	1	1336.4	1358.4
- Pregnancies	1	1336.6	1358.6
- LDL	1	1337.1	1359.1
- WHR	1	1337.4	1359.4
- WaistCircumference	1	1337.5	1359.5
- HipCircumference	1	1337.7	1359.7
- Triglycerides	1	1337.9	1359.9
<none>		1336.3	1360.3
+ factor(MedicationUse)	1	1336.3	1362.3
+ HDL	1	1336.3	1362.3
+ factor(Hypertension)	1	1336.3	1362.3
+ factor(DietType)	2	1336.2	1364.2
- BMI	1	1382.4	1404.4
- Age	1	1424.3	1446.3
- HbA1c	1	2005.3	2027.3
- factor(FamilyHistory)	1	12030.3	12052.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1358.41

Outcome ~ Age + Pregnancies + BMI + HbA1c + LDL + Triglycerides +  
WaistCircumference + HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- Pregnancies	1	1336.7	1356.7
- LDL	1	1337.2	1357.2
- WHR	1	1337.4	1357.4
- WaistCircumference	1	1337.6	1357.6
- HipCircumference	1	1337.8	1357.8
- Triglycerides	1	1338.0	1358.0
<none>		1336.4	1358.4
+ BloodPressure	1	1336.3	1360.3
+ factor(MedicationUse)	1	1336.4	1360.4
+ HDL	1	1336.4	1360.4
+ factor(Hypertension)	1	1336.4	1360.4
+ factor(DietType)	2	1336.3	1362.3
- BMI	1	1390.7	1410.7
- Age	1	1443.5	1463.5
- HbA1c	1	2005.4	2025.4
- factor(FamilyHistory)	1	12030.5	12050.5

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1356.67

Outcome ~ Age + BMI + HbA1c + LDL + Triglycerides + WaistCircumference +  
HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- LDL	1	1337.4	1355.4
- WHR	1	1337.7	1355.7
- WaistCircumference	1	1337.8	1355.8
- HipCircumference	1	1338.0	1356.0
- Triglycerides	1	1338.3	1356.3
<none>		1336.7	1356.7
+ Pregnancies	1	1336.4	1358.4
+ BloodPressure	1	1336.6	1358.6
+ factor(MedicationUse)	1	1336.6	1358.6
+ HDL	1	1336.7	1358.7
+ factor(Hypertension)	1	1336.7	1358.7
+ factor(DietType)	2	1336.5	1360.5
- BMI	1	1391.0	1409.0
- Age	1	1444.1	1462.1
- HbA1c	1	2005.4	2023.4
- factor(FamilyHistory)	1	12033.9	12051.9

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1355.42

Outcome ~ Age + BMI + HbA1c + Triglycerides + WaistCircumference +  
HipCircumference + WHR + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- WHR	1	1338.4	1354.4
- WaistCircumference	1	1338.5	1354.5
- HipCircumference	1	1338.7	1354.7
- Triglycerides	1	1339.0	1355.0
<none>		1337.4	1355.4
+ LDL	1	1336.7	1356.7
+ Pregnancies	1	1337.2	1357.2
+ BloodPressure	1	1337.4	1357.4
+ factor(MedicationUse)	1	1337.4	1357.4
+ HDL	1	1337.4	1357.4
+ factor(Hypertension)	1	1337.4	1357.4
+ factor(DietType)	2	1337.3	1359.3
- BMI	1	1391.9	1407.9
- Age	1	1444.9	1460.9
- HbA1c	1	2005.7	2021.7
- factor(FamilyHistory)	1	12036.8	12052.8

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1354.42

Outcome ~ Age + BMI + HbA1c + Triglycerides + WaistCircumference +  
HipCircumference + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- WaistCircumference	1	1338.6	1352.6
- HipCircumference	1	1339.3	1353.3
- Triglycerides	1	1340.0	1354.0
<none>		1338.4	1354.4
+ WHR	1	1337.4	1355.4
+ LDL	1	1337.7	1355.7
+ Pregnancies	1	1338.2	1356.2
+ BloodPressure	1	1338.4	1356.4
+ factor(MedicationUse)	1	1338.4	1356.4
+ HDL	1	1338.4	1356.4
+ factor(Hypertension)	1	1338.4	1356.4
+ factor(DietType)	2	1338.3	1358.3
- BMI	1	1393.2	1407.2
- Age	1	1446.1	1460.1
- HbA1c	1	2006.1	2020.1
- factor(FamilyHistory)	1	12037.5	12051.5

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1352.6

Outcome ~ Age + BMI + HbA1c + Triglycerides + HipCircumference +  
factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- HipCircumference	1	1339.4	1351.4
- Triglycerides	1	1340.2	1352.2
<none>		1338.6	1352.6
+ LDL	1	1337.9	1353.9
+ Pregnancies	1	1338.3	1354.3
+ WaistCircumference	1	1338.4	1354.4
+ WHR	1	1338.5	1354.5
+ BloodPressure	1	1338.5	1354.5
+ factor(MedicationUse)	1	1338.6	1354.6
+ HDL	1	1338.6	1354.6
+ factor(Hypertension)	1	1338.6	1354.6
+ factor(DietType)	2	1338.5	1356.5
- BMI	1	1432.0	1444.0
- Age	1	1446.5	1458.5
- HbA1c	1	2006.2	2018.2
- factor(FamilyHistory)	1	12038.2	12050.2

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1351.44

Outcome ~ Age + BMI + HbA1c + Triglycerides + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred



Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
- Triglycerides	1	1341.0	1351.0
<none>		1339.4	1351.4
+ HipCircumference	1	1338.6	1352.6
+ LDL	1	1338.7	1352.7
+ WHR	1	1338.8	1352.8
+ Pregnancies	1	1339.2	1353.2
+ WaistCircumference	1	1339.3	1353.3
+ BloodPressure	1	1339.4	1353.4
+ HDL	1	1339.4	1353.4
+ factor(MedicationUse)	1	1339.4	1353.4
+ factor(Hypertension)	1	1339.4	1353.4
+ factor(DietType)	2	1339.3	1355.3
- Age	1	1447.2	1457.2
- BMI	1	1469.5	1479.5
- HbA1c	1	2007.3	2017.3
- factor(FamilyHistory)	1	12040.3	12050.3

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Step: AIC=1351.03

Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory)

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

	Df	Deviance	AIC
<none>		1341.0	1351.0
+ Triglycerides	1	1339.4	1351.4
+ HipCircumference	1	1340.2	1352.2
+ LDL	1	1340.4	1352.4
+ WHR	1	1340.4	1352.4
+ Pregnancies	1	1340.8	1352.8
+ WaistCircumference	1	1340.9	1352.9
+ BloodPressure	1	1341.0	1353.0
+ factor(MedicationUse)	1	1341.0	1353.0
+ HDL	1	1341.0	1353.0
+ factor(Hypertension)	1	1341.0	1353.0
+ factor(DietType)	2	1341.0	1355.0
- Age	1	1449.0	1457.0
- BMI	1	1470.9	1478.9
- HbA1c	1	2007.7	2015.7
- factor(FamilyHistory)	1	12040.7	12048.7

```
summary(both_model)
```

Call:

```
glm(formula = Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory),
     family = "binomial", data = dataset)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-35.951179	1.471747	-24.428	<2e-16 ***
Age	0.045863	0.004829	9.497	<2e-16 ***
BMI	0.166579	0.015576	10.695	<2e-16 ***
HbA1c	4.906516	0.249002	19.705	<2e-16 ***
factor(FamilyHistory)1	30.631935	435.532368	0.070	0.944

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 12280 on 9537 degrees of freedom  
Residual deviance: 1341 on 9533 degrees of freedom

AIC: 1351

Number of Fisher Scoring iterations: 20

### Anova test

```
anova(backward_model, model)
```

#### Analysis of Deviance Table

Model 1: Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory)

Model 2: Outcome ~ Age + Pregnancies + BMI + BloodPressure + HbA1c + LDL +  
HDL + Triglycerides + WaistCircumference + HipCircumference +  
WHR + factor(FamilyHistory) + factor(DietType) + factor(Hypertension) +  
factor(MedicationUse)

	Resid. Df	Resid. Dev	Df	Deviance
1	9533	1341.0		
2	9521	1336.2	12	4.871

### Correlation for final dataset

```
final_model_dataset <- dataset[,c("Age", "HbA1c", "FamilyHistory", "BMI")]  
cor(final_model_dataset)
```

	Age	HbA1c	FamilyHistory	BMI
Age	1.00000000	0.396354303	-0.012887442	0.021793701
HbA1c	0.39635430	1.000000000	-0.007980574	0.459851892
FamilyHistory	-0.01288744	-0.007980574	1.000000000	0.008037471
BMI	0.02179370	0.459851892	0.008037471	1.000000000

### Finalize the model

```
final_model <- glm(Outcome ~ Age + HbA1c + factor(FamilyHistory) + BMI, data = dataset, fami.
```

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
summary(final_model)
```

Call:

```
glm(formula = Outcome ~ Age + HbA1c + factor(FamilyHistory) +  
    BMI, family = "binomial", data = dataset)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-35.951179	1.471747	-24.428	<2e-16 ***
Age	0.045863	0.004829	9.497	<2e-16 ***
HbA1c	4.906516	0.249002	19.705	<2e-16 ***
factor(FamilyHistory)1	30.631935	435.532368	0.070	0.944
BMI	0.166579	0.015576	10.695	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 12280 on 9537 degrees of freedom  
Residual deviance: 1341 on 9533 degrees of freedom  
AIC: 1351

Number of Fisher Scoring iterations: 20

```
vif(final_model)
```

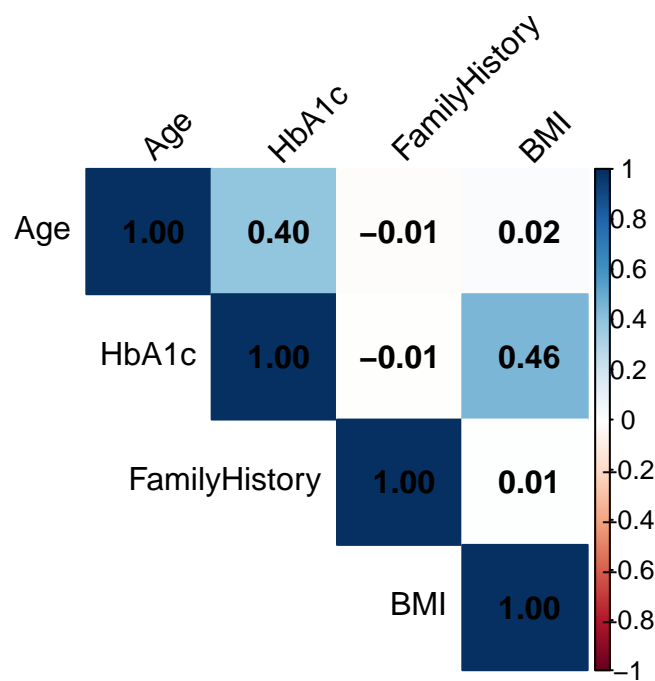
	Age	HbA1c	factor(FamilyHistory)
	1.115542	1.002088	1.000001
BMI	1.117578		

### Create correlation heatmap

```
cor_matrix <- cor(final_model_dataset, use = "pairwise.complete.obs")  
diag(cor_matrix) <- 1  
  
#png(  
# "correlation_heatmap_large.png",
```

```
# width = 2000,
# height = 2000,
# res = 120)

heatmap <- corrplot(
  cor_matrix,
  method = "color",
  type = "upper",
  addCoef.col = "black",
  tl.col = "black",
  tl.srt = 45,
  diag = TRUE)
```



```
#dev.off()
heatmap
```

```
$corr
      Age      HbA1c FamilyHistory      BMI
Age      1.00000000 0.396354303 -0.012887442 0.021793701
HbA1c    0.39635430 1.000000000 -0.007980574 0.459851892
FamilyHistory -0.01288744 -0.007980574 1.000000000 0.008037471
BMI      0.02179370 0.459851892 0.008037471 1.000000000
```

```
$corrPos
```

	xName	yName	x	y	corr
1	Age	Age	1	4	1.000000000
2	HbA1c	Age	2	4	0.396354303
3	HbA1c	HbA1c	2	3	1.000000000
4	FamilyHistory	Age	3	4	-0.012887442
5	FamilyHistory	HbA1c	3	3	-0.007980574
6	FamilyHistory	FamilyHistory	3	2	1.000000000
7	BMI	Age	4	4	0.021793701
8	BMI	HbA1c	4	3	0.459851892
9	BMI	FamilyHistory	4	2	0.008037471
10	BMI	BMI	4	1	1.000000000

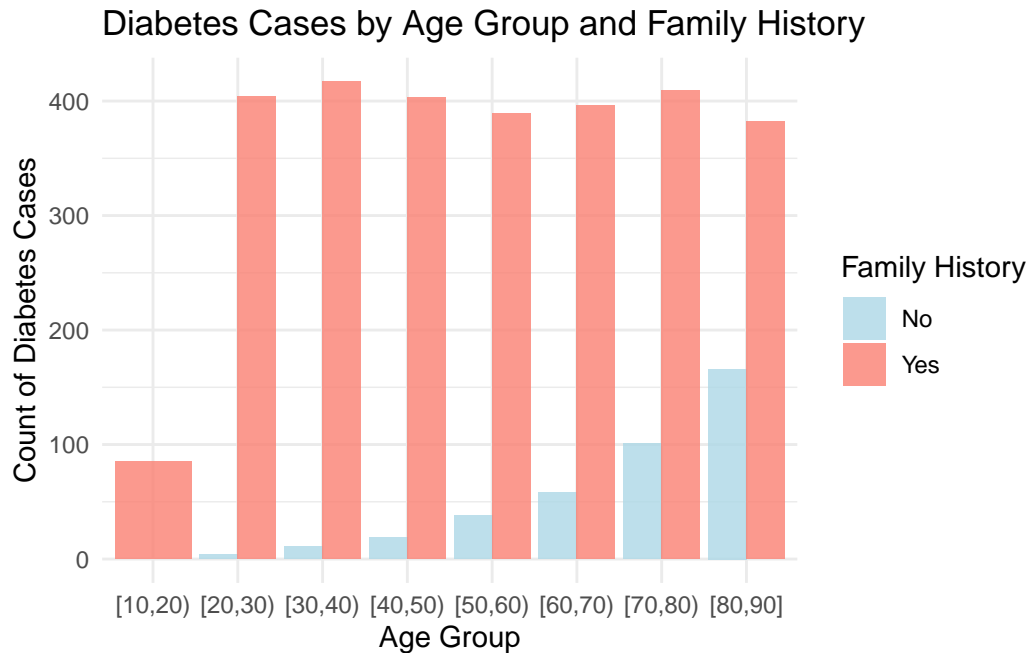
```
$arg
$arg$type
[1] "upper"
```

## Visualization

```
temp_data <- dataset %>%
  mutate(AgeGroup = cut(Age, breaks = seq(10, 90, by = 10), include.lowest = TRUE, right = F),
         FamilyHistory = factor(FamilyHistory, levels = c(0, 1), labels = c("No", "Yes")))

diabetes_data <- temp_data %>% filter(Outcome == 1)

ggplot(diabetes_data, aes(x = AgeGroup, fill = FamilyHistory)) +
  geom_bar(position = "dodge", alpha = 0.8) +
  labs(title = "Diabetes Cases by Age Group and Family History",
       x = "Age Group",
       y = "Count of Diabetes Cases",
       fill = "Family History") +
  scale_fill_manual(values = c("No" = "lightblue", "Yes" = "salmon")) +
  theme_minimal()
```

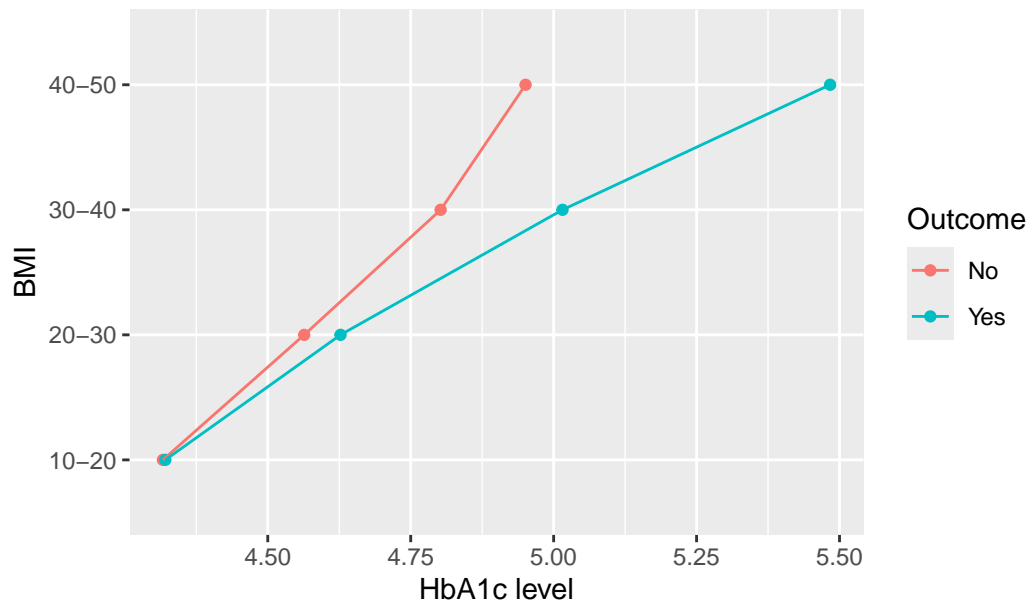


```
temp_dataset <- dataset %>%
  mutate(BMI_category = cut(BMI,
                             breaks = c(10, 20, 30, 40, 50),
                             labels = c("10-20", "20-30", "30-40", "40-50"),
                             include.lowest = TRUE))

temp_dataset$Outcome <- factor(temp_dataset$Outcome, levels = c(0, 1), labels = c("No", "Yes"))

ggplot(data=temp_dataset,
       mapping = aes(x = HbA1c,
                     y = factor(BMI_category),
                     group = Outcome,
                     color = Outcome)) +
  stat_summary(fun = "mean", geom = "point") +
  stat_summary(fun = "mean", geom = "line") +
  labs(title = "Interaction between HbA1c and BMI category on Diabetes Outcome",
       x = "HbA1c level",
       y = "BMI",
       color = "Outcome")
```

Interaction between HbA1c and BMI category on Diabetes O



## Model Evaluation

```
train_index <- createDataPartition(dataset$Outcome, p=0.7, list= FALSE)
train_data <- dataset[train_index,]
test_data <- dataset[-train_index,]

train_model <- glm(Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory), data = train_data, f
```

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
predictions <- predict(train_model, newdata = test_data, type = "response")
predicted_classification <- ifelse(predictions > 0.5, 1, 0)

conf_matrix <- confusionMatrix(as.factor(predicted_classification), as.factor(test_data$Outc
conf_matrix
```

## Confusion Matrix and Statistics

Reference



Prediction	0	1
0	1905	59
1	20	877

Accuracy : 0.9724  
 95% CI : (0.9657, 0.9781)  
 No Information Rate : 0.6728  
 P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.9366

McNemar's Test P-Value : 1.909e-05

Sensitivity : 0.9896  
 Specificity : 0.9370  
 Pos Pred Value : 0.9700  
 Neg Pred Value : 0.9777  
 Prevalence : 0.6728  
 Detection Rate : 0.6659  
 Detection Prevalence : 0.6865  
 Balanced Accuracy : 0.9633

'Positive' Class : 0

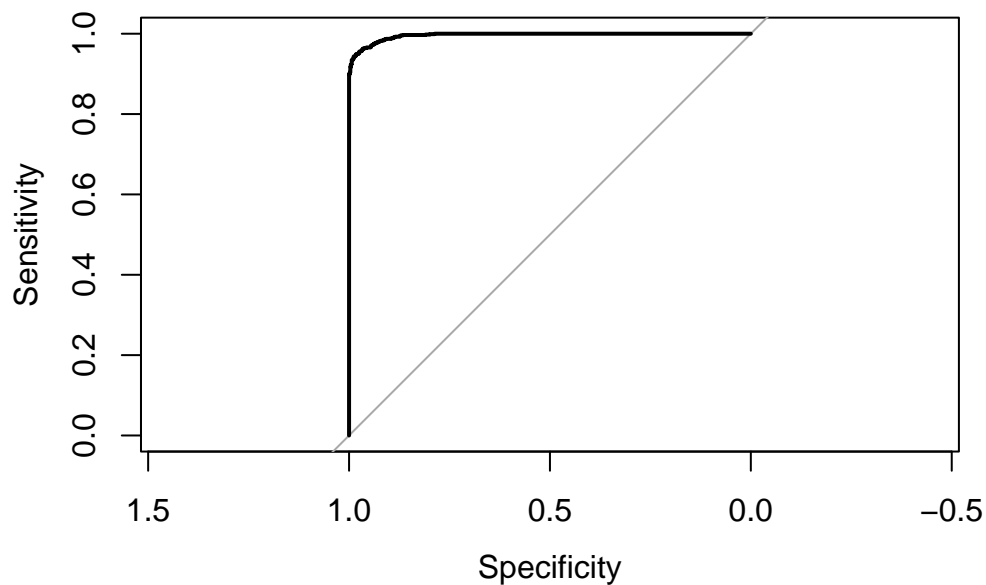
## Plotting ROC curve

```
roc_curve <- roc(test_data$Outcome, predictions)
```

Setting levels: control = 0, case = 1

Setting direction: controls < cases

```
plot(roc_curve)
```



```
auc_value <- auc(roc_curve)
auc_value
```

Area under the curve: 0.9959

### 5-fold cross validation

```
train_control <- trainControl(method = "cv", number = 5)
cv_model <- train(Outcome ~ Age + BMI + HbA1c + factor(FamilyHistory),
  data = dataset,
  method = "glm",
  family = "binomial",
  trControl = train_control)
```

Warning in train.default(x, y, weights = w, ...): You are trying to do regression and your outcome only has two possible values Are you trying to do classification? If so, use a 2 level factor as your outcome column.

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
cv_model
```

Generalized Linear Model

```
9538 samples
  4 predictor
```

```
No pre-processing
Resampling: Cross-Validated (5 fold)
Summary of sample sizes: 7630, 7630, 7631, 7630, 7631
Resampling results:
```

RMSE	Rsquared	MAE
0.1456907	0.9059009	0.04161462

```
cv_accuracy <- cv_model$results
cv_accuracy
```

	parameter	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD
1	none	0.1456907	0.9059009	0.04161462	0.005608905	0.006149663	0.001311478