Result of summary(Log_Reg)…

```
Call:
glm(formula = health_ins ~ age + maritl + wage + logwage + race +
    education + jobclass, family = binomial(link = logit), data = Health.train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4620  -0.7820  -0.5714   0.9000   2.3889

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)              20.428433   2.188121   9.336  < 2e-16 ***
age                      -0.014547   0.004503  -3.231 0.001235 **
maritl2. Married          0.313469   0.127756   2.454 0.014141 *
maritl3. Widowed         -0.355122   0.640348  -0.555 0.579185
maritl4. Divorced        -0.061557   0.218209  -0.282 0.777866
maritl5. Separated        0.426080   0.342054   1.246 0.212892
wage                      0.021227   0.005101   4.161 3.17e-05 ***
logwage                  -4.910664   0.589347  -8.332  < 2e-16 ***
race2. Black             -0.002317   0.155556  -0.015 0.988118
race3. Asian              0.365297   0.189906   1.924 0.054409 .
race4. Other              0.099236   0.382079   0.260 0.795074
education2. HS Grad      -0.326479   0.159350  -2.049 0.040480 *
education3. Some College -0.440662   0.175771  -2.507 0.012175 *
education4. College Grad -0.415392   0.182523  -2.276 0.022856 *
education5. Advanced Degree -0.319218   0.217627  -1.467 0.142428
jobclass2. Information   -0.346349   0.097556  -3.550 0.000385 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3322.0  on 2699  degrees of freedom
Residual deviance: 2842.2  on 2684  degrees of freedom
AIC: 2874.2

Number of Fisher Scoring iterations: 4
```

Result of summary(Gam_Reg)…

```
Call: gam(formula = health_ins ~ s(age) + s(logwage) + s(wage) + maritl +
    race + education + jobclass, family = binomial(link = logit), data = Health.train)
```

Deviance Residuals:
   Min   1Q Median   3Q   Max
-2.1095 -0.7672 -0.5641 0.8624 2.3175

(Dispersion Parameter for binomial family taken to be 1)

   Null Deviance: 3322.047 on 2699 degrees of freedom
Residual Deviance: 2824.68 on 2675 degrees of freedom
AIC: 2874.681

Number of Local Scoring Iterations: 5

Anova for Parametric Effects
           Df Sum Sq Mean Sq F value   Pr(>F)
s(age)   1   29.88 29.885 29.9111 4.942e-08 ***
s(logwage)   1 244.12 244.118 244.3343 < 2.2e-16 ***
s(wage)   1   1.22   1.220   1.2207 0.2693322
maritl   4   17.39   4.348   4.3517 0.0016475 **
race   3   4.07   1.357   1.3581 0.2538100
education   4   8.56   2.140   2.1424 0.0731192 .
jobclass   1   12.20 12.195 12.2059 0.0004841 ***
Residuals 2675 2672.63   0.999

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Anova for Nonparametric Effects
           Npar Df Npar Chisq P(Chi)
(Intercept)
s(age)   3   9.3275 0.02524 *
s(logwage)   3   3.6985 0.29592
s(wage)   3   2.9160 0.40476
maritl
race
education
jobclass

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Result of summary(RandomForest_Reg)…

             Length Class Mode
call   3   -none- call
type   1   -none- character

```
predicted      2700   factor numeric
err.rate       1500   -none- numeric
confusion         6   -none- numeric
votes          5400   matrix numeric
oob.times      2700   -none- numeric
classes           2   -none- character
importance        7   -none- numeric
importanceSD      0   -none- NULL
localImportance   0   -none- NULL
proximity         0   -none- NULL
ntree             1   -none- numeric
mtry              1   -none- numeric
forest           14   -none- list
y              2700   factor numeric
test              0   -none- NULL
inbag             0   -none- NULL
terms             3   terms  call
```

Results of summary(Tree2_Reg)…

```
Classification tree:
tree(formula = health_ins ~ age + maritl + wage + logwage + race +
    education + jobclass, data = Health.train)
Variables actually used in tree construction:
[1] "wage"
Number of terminal nodes:  3
Residual mean deviance:  1.075 = 2900 / 2697
Misclassification error rate: 0.2559 = 691 / 2700
```

Results of summary(Rpart_Reg)…

```
Call:
rpart(formula = health_ins ~ age + maritl + wage + logwage +
    race + education + jobclass, data = Health.train)
  n= 2700

         CP nsplit rel error    xerror       xstd
1 0.12864078      0 1.0000000 1.0000000 0.02903828
2 0.03276699      1 0.8713592 0.8774272 0.02792308
3 0.01334951      2 0.8385922 0.8446602 0.02758318
4 0.01000000      3 0.8252427 0.8507282 0.02764753
```

```
Variable importance
 logwage      wage       age education
      46        46         6         2

Node number 1: 2700 observations,    complexity param=0.1286408
  predicted class=1. Yes  expected loss=0.3051852  P(node) =1
    class counts:  1876   824
   probabilities: 0.695 0.305
  left son=2 (1988 obs) right son=3 (712 obs)
  Primary splits:
      wage      < 87.91729 to the right, improve=140.20990, (0 missing)
      logwage   < 4.476396 to the right, improve=140.20990, (0 missing)
      age       < 24.5     to the right, improve= 53.09782, (0 missing)
      education splits as  RRLLL,       improve= 36.99797, (0 missing)
      jobclass  splits as  RL,          improve= 24.98751, (0 missing)
  Surrogate splits:
      logwage   < 4.476396 to the right, agree=1.000, adj=1.000, (0 split)
      age       < 24.5     to the right, agree=0.773, adj=0.140, (0 split)
      education splits as  RLLLL,       agree=0.750, adj=0.053, (0 split)
      race      splits as  LLLR,        agree=0.737, adj=0.004, (0 split)

Node number 2: 1988 observations
  predicted class=1. Yes  expected loss=0.2087525  P(node) =0.7362963
    class counts:  1573   415
   probabilities: 0.791 0.209

Node number 3: 712 observations,    complexity param=0.03276699
  predicted class=2. No   expected loss=0.4255618  P(node) =0.2637037
    class counts:   303   409
   probabilities: 0.426 0.574
  left son=6 (477 obs) right son=7 (235 obs)
  Primary splits:
      wage      < 68.25382 to the right, improve=30.509870, (0 missing)
      logwage   < 4.223231 to the right, improve=30.509870, (0 missing)
      age       < 24.5     to the right, improve= 8.339122, (0 missing)
      education splits as  RLLLL,       improve= 3.663245, (0 missing)
      maritl    splits as  RLLLR,       improve= 2.362453, (0 missing)
  Surrogate splits:
      logwage < 4.223231 to the right, agree=1.000, adj=1.000, (0 split)
      age     < 19.5     to the right, agree=0.677, adj=0.021, (0 split)

Node number 6: 477 observations,    complexity param=0.01334951
  predicted class=1. Yes  expected loss=0.4716981  P(node) =0.1766667
    class counts:   252   225
```

probabilities: 0.528 0.472
left son=12 (206 obs) right son=13 (271 obs)
Primary splits:
age < 41.5 to the right, improve=2.963951, (0 missing)
maritl splits as RRLLR, improve=2.440026, (0 missing)
education splits as RLRLR, improve=1.870319, (0 missing)
wage < 74.17434 to the right, improve=1.162474, (0 missing)
logwage < 4.306413 to the right, improve=1.162474, (0 missing)
Surrogate splits:
maritl splits as RLRLR, agree=0.675, adj=0.248, (0 split)
education splits as RRRRL, agree=0.585, adj=0.039, (0 split)
wage < 84.20789 to the right, agree=0.572, adj=0.010, (0 split)
logwage < 4.433288 to the right, agree=0.572, adj=0.010, (0 split)
race splits as RLRR, agree=0.572, adj=0.010, (0 split)

Node number 7: 235 observations
predicted class=2. No expected loss=0.2170213 P(node) =0.08703704
class counts: 51 184
probabilities: 0.217 0.783

Node number 12: 206 observations
predicted class=1. Yes expected loss=0.407767 P(node) =0.0762963
class counts: 122 84
probabilities: 0.592 0.408

Node number 13: 271 observations
predicted class=2. No expected loss=0.4797048 P(node) =0.1003704
class counts: 130 141
probabilities: 0.480 0.520