CHRONIC KIDNEY DISEASE

SUBMITTED BY:
VARUN MAHESHWARI

VARIABLES SELECTION

- Remove ID
- Remove Weight, Height, Obese Columns (included in BMI)
- Remove HDL, LDL, and Total Chol (included in dyslipidemia column)
- Remove SBP and DBP (included in Hypertension column)
- Not including Education. Income, Insured, Unmarried

  - Left with 20 variables
- Screening: Removed Null values in the dependent variable: CKD

Treatment of Data

Name of Variable	Variable Type	Treatment
Age	Numerical Converting into Categorical *Base: 1: 18-44*	Converting into Age category with following levels: 18-44 as 1 44-65 as 2 65+ as 3
BMI	Numerical Converted to Categorical Base: Normal: 25-30	Converted into BMI category with levels: <18.5, 18.5-25, 25-30, 30+
Hypertension, Fam Hypertension	Clubbed in 1 column: Hyper Categorical Base: None	Levels: None Only in Individual Only in Family In both
Diabetes, Fam Diabetes	Clubbed in 1 column: db Categorical Base: None	Levels: None Only in Individual Only in Family In both
CVD, Fam CVD	Clubbed in 1 column: cv Categorical Base: None	Levels: None Only in Individual Only in Family In both

Hypothesis Testing

```
#Significant
chisq.test(ck1$Agecat, ck1$CKD)    #p-value < 0.00000000000000022
chisq.test(ck1$Racegrp, ck1$CKD)     #p-value = 0.00000000001476
chisq.test(ck1$CareSource, ck1$CKD)     #p-value = 0.0000000006604
chisq.test(ck1$PVD, ck1$CKD)          #p-value < 0.00000000000000022
chisq.test(ck1$Activity, ck1$CKD)      #p-value = 0.000000000001826
chisq.test(ck1$PoorVision, ck1$CKD)     #p-value = 0.0000000000000235
chisq.test(ck1$Smoker, ck1$CKD)        #p-value = 0.00002722
chisq.test(ck1$hyper, ck1$CKD)        #p-value < 0.00000000000000022
chisq.test(ck1$Stroke, ck1$CKD)      #p-value < 0.00000000000000022
chisq.test(ck1$CHF, ck1$CKD)       #p-value < 0.00000000000000022
chisq.test(ck1$db, ck1$CKD)      #p-value < 0.00000000000000022
chisq.test(ck1$cvd, ck1$CKD)       #p-value < 0.00000000000000022
chisq.test(ck1$Anemia, ck1$CKD)      #p-value = 0.001072

#Insignificant
chisq.test(ck1$Female, ck1$CKD)   #p-value = 0.5798
chisq.test(ck1$Dyslipidemia, ck1$CKD)    #p-value = 1
chisq.test(ck1$BMIcat, ck1$CKD)     #p-value = 0.04349
```

Removed Insignificant Variables

Model Building

```
Coefficients:
                   Estimate Std. Error z value             Pr(>|z|)
(Intercept)        -5.00219    0.40832 -12.251 < 0.0000000000000002 ***
Racegrpblack       -0.36881    0.17874  -2.063              0.039074 *
Racegrphispa       -0.86753    0.17719  -4.896           0.000000978 ***
Racegrpother       -0.81718    0.54262  -1.506              0.132070
CareSourceclinic    0.28462    0.33975   0.838              0.402169
CareSourceDrHMO     0.23773    0.32091   0.741              0.458811
CareSourceother     0.27142    0.41926   0.647              0.517385
PVD1                0.67999    0.19266   3.529              0.000416 ***
Activity2          -0.36053    0.13773  -2.618              0.008851 **
Activity3          -0.79185    0.23470  -3.374              0.000741 ***
Activity4          -0.78941    0.48152  -1.639              0.101125
PoorVision1         0.38102    0.19168   1.988              0.046841 *
Smoker1            -0.01275    0.12804  -0.100              0.920656
Stroke1             0.19824    0.29998   0.661              0.508707
CHF1                0.48511    0.24732   1.961              0.049826 *
Anemia1             1.63347    0.36568   4.467           0.000007936 ***
Agecat2             1.60659    0.30508   5.266           0.000000139 ***
Agecat3             3.07951    0.30194  10.199 < 0.0000000000000002 ***
hyperIn both        0.45735    0.28637   1.597              0.110247
hyperOnly in fam   -0.55123    0.44214  -1.247              0.212497
hyperOnly individual  0.80368    0.16234   4.951           0.000000740 ***
dbIn both           0.36359    0.19479   1.867              0.061964 .
dbOnly in fam      -0.02377    0.16041  -0.148              0.882208
dbOnly individual   0.62338    0.21830   2.856              0.004296 **
cvdIn both          0.71480    0.36485   1.959              0.050096 .
cvdOnly in fam      0.27258    0.23444   1.163              0.244955
cvdOnly individual  0.60585    0.24814   2.442              0.014625 *
---
```

Significant variables:

Positively related with respect to base:
PVD, Anemia, Age category, hyperInBoth, hyperOnlyIndividual, dbInBoth, dbIndividual,
cvdInBoth, cvdIndividual

Negatively related with respect to base:
Race group, Activity, Anemia

Odds Ratio

```
> logistic.display(log)

Logistic regression predicting CKD : 1 vs 0

                        crude OR(95%CI)      adj. OR(95%CI)      P(Wald's test) P(LR-test)
Racegrp: ref.=white                                                             < 0.001
   black                0.69 (0.51,0.92)     0.69 (0.49,0.98)    0.039
   hispa                0.36 (0.26,0.49)     0.42 (0.3,0.59)     < 0.001
   other                0.26 (0.09,0.7)      0.44 (0.15,1.28)    0.132

CareSource: ref.=noplace                                                        0.861
   clinic               4.82 (2.6,8.93)      1.33 (0.68,2.59)    0.402
   DrHMO                5.91 (3.3,10.61)     1.27 (0.68,2.38)    0.459
   other                4.41 (2.08,9.37)     1.31 (0.58,2.98)    0.517

PVD: 1 vs 0             7.17 (5.13,10.02)    1.97 (1.35,2.88)    < 0.001        < 0.001

Activity: ref.=1                                                                0.002
   2                    0.61 (0.48,0.77)     0.7 (0.53,0.91)     0.009
   3                    0.3 (0.2,0.46)       0.45 (0.29,0.72)    < 0.001
   4                    0.13 (0.05,0.31)     0.45 (0.18,1.17)    0.101

PoorVision: 1 vs 0      3.34 (2.42,4.61)     1.46 (1.01,2.13)    0.047          0.051

Smoker: 1 vs 0          1.61 (1.29,2.01)     0.99 (0.77,1.27)    0.921          0.921

Stroke: 1 vs 0          6.17 (4.2,9.07)      1.22 (0.68,2.2)     0.509          0.508

CHF: 1 vs 0             5.92 (3.96,8.86)     1.62 (1,2.64)       0.05           0.054

Anemia: 1 vs 0          2.53 (1.47,4.36)     5.12 (2.5,10.49)    < 0.001        < 0.001

Agecat: ref.=1                                                                  < 0.001
   2                    8.12 (4.57,14.42)    4.99 (2.74,9.07)    < 0.001
   3                    54.43 (31.63,93.69)  21.75 (12.03,39.3)  < 0.001
```

```
hyper: ref.=None                                                                < 0.001
   In both              4.14 (2.86,6)        1.58 (0.9,2.77)     0.11
   Only in fam          0.42 (0.2,0.89)      0.58 (0.24,1.37)    0.212
   Only individual      6.97 (5.25,9.26)     2.23 (1.62,3.07)    < 0.001

db: ref.=None                                                                   0.014
   In both              3.03 (2.19,4.19)     1.44 (0.98,2.11)    0.062
   Only in fam          0.98 (0.74,1.3)      0.98 (0.71,1.34)    0.882
   Only individual      4.42 (3.05,6.4)      1.87 (1.22,2.86)    0.004

cvd: ref.=None                                                                  0.066
   In both              3.82 (2.36,6.17)     2.04 (1,4.18)       0.05
   Only in fam          0.67 (0.51,0.88)     1.31 (0.83,2.08)    0.245
   Only individual      6.68 (4.74,9.42)     1.83 (1.13,2.98)    0.015

Log-likelihood = -911.6742
No. of observations = 5166
AIC value = 1877.3484
```

Inferences

- People with PVD are 1.97 times more likely to have CKD than people not having PVD
- People lifting loads are 0.7 times less likely and people climbing stairs and doing heavy workout are 0.45 times less likely to have CKD than people who mostly sit.
- Blacks are 0.69 times less likely to have CKD and Hispanic are 0.42 times less likely to have CKD than Whites.
- Anemic patients (in past 3 months) are 5.12 times more likely to have CKD than who aren't.
- People with age 44-65 are 4.99 times more likely and people with age 65+ are 21.75 times more likely to have CKD.
- People having hypertension (with no family history) are 2..23 times more likely to have CKD than people who do not have hypertension nor have a family history.
- People having diabetes (with no family history) are 1.87 times more likely to have CKD than people who do not have diabetes nor have a family history.
- People having CVD (with family history) are 2.04 times more likely to have CKD than people who do not have CVD nor have a family history. People having CVD (with no family history) are 1.83 times more likely to have CKD than people who do not have CVD nor have a family history.