Fintech 678 Final Project

Yan Wang

2023-12-12

##I. Abstract This project report presents an in-depth analysis of financial user data, employing advanced statistical modeling techniques within the RMarkdown environment. The report concludes with strategic recommendations for financial institutions seeking to optimize user engagement and deposit growth, grounded in the quantitative evidence gathered from the data analysis.

##II. Introduction Aimed at uncovering key insights in the fintech sector, the study focuses on the relationship between user deposits and several predictors including age, housing status, credit score, and zodiac sign. Utilizing linear regression models, multilevel models with random slopes and intercepts, and a comprehensive exploratory data analysis (EDA), the report offers a nuanced understanding of factors influencing user deposit behaviors. The models' adequacies were rigorously tested through ANOVA and convergence diagnostics, ensuring robustness in the findings. Key results indicate significant relationships between the predictors and deposit amounts, with detailed analysis provided on the variability due to individual differences and other socioeconomic factors.

##III. Method

```
##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages

##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages

##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages

##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

###Data source: https://www.kaggle.com/datasets/niketdheeryan/fintech-users-data

###overview of data:

```
head(fintech_data)
```

```
##    user churn age housing credit_score deposits withdrawal purchases_partners
## 1 55409     0  37      na           NA        0          0                  0
## 2 23547     0  28       R          486        0          0                  1
## 3 58313     0  35       R          561       47          2                 86
## 4  8095     0  26       R          567       26          3                 38
## 5 61353     1  27      na           NA        0          0                  2
## 6  3120     1  32       R          567        5          3                111
##   purchases cc_taken cc_recommended cc_disliked cc_liked cc_application_begin
## 1         0        0              0           0        0                    0
```

```
## 2           0        0          96         0         0                5
## 3          47        0         285         0         0                9
## 4          25        0          74         0         0               26
## 5           0        0           0         0         0                0
## 6           5        0         227         0         0               17
##   app_downloaded web_user app_web_user ios_user android_user registered_phones
## 1              1        1            1        1            0                 0
## 2              1        1            1        1            0                 0
## 3              1        0            0        0            1                 2
## 4              1        0            0        0            1                 0
## 5              1        1            1        0            1                 0
## 6              1        1            1        0            1                 0
##   payment_type waiting_4_loan cancelled_loan received_loan rejected_loan
## 1    Bi-Weekly              0              0             0             0
## 2       Weekly              0              0             0             0
## 3 Semi-Monthly              0              0             0             0
## 4    Bi-Weekly              0              0             0             0
## 5    Bi-Weekly              0              0             0             0
## 6    Bi-Weekly              0              0             0             0
##   zodiac_sign left_for_two_month_plus left_for_one_month rewards_earned
## 1         Leo                       1                  0             NA
## 2         Leo                       0                  0             44
## 3   Capricorn                       1                  0             65
## 4   Capricorn                       0                  0             33
## 5       Aries                       1                  0              1
## 6      Taurus                       0                  0             55
##   reward_rate is_referred
## 1        0.00           0
## 2        1.47           1
## 3        2.17           0
## 4        1.10           1
## 5        0.03           0
## 6        1.83           0
```

#The structure of the data

```
summary(fintech_data)
```

```
##       user           churn             age           housing
##  Min.   :    1   Min.   :0.0000   Min.   :17.00   Length:27000
##  1st Qu.:17810   1st Qu.:0.0000   1st Qu.:25.00   Class :character
##  Median :35749   Median :0.0000   Median :30.00   Mode  :character
##  Mean   :35423   Mean   :0.4139   Mean   :32.22
##  3rd Qu.:53244   3rd Qu.:1.0000   3rd Qu.:37.00
##  Max.   :69658   Max.   :1.0000   Max.   :91.00
##                                   NA's   :4
##   credit_score    highnetworth      withdrawal     purchases_partners
##  Min.   :  2.0   Min.   : 0.000   Min.   : 0.000   Min.   :   0.00
##  1st Qu.:507.0   1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.:   0.00
##  Median :542.0   Median : 0.000   Median : 0.000   Median :   9.00
##  Mean   :542.9   Mean   : 3.342   Mean   : 0.307   Mean   :  28.06
##  3rd Qu.:578.0   3rd Qu.: 1.000   3rd Qu.: 0.000   3rd Qu.:  43.00
##  Max.   :838.0   Max.   :65.000   Max.   :29.000   Max.   :1067.00
##  NA's   :8031
##    purchases         cc_taken       cc_recommended    cc_disliked
```

```
##  Min.    : 0.000   Min.    : 0.00000   Min.    :  0.00   Min.    : 0.00000
##  1st Qu.: 0.000   1st Qu.: 0.00000   1st Qu.: 10.00   1st Qu.: 0.00000
##  Median : 0.000   Median : 0.00000   Median : 65.00   Median : 0.00000
##  Mean   : 3.273   Mean   : 0.07378   Mean   : 92.63   Mean   : 0.05063
##  3rd Qu.: 1.000   3rd Qu.: 0.00000   3rd Qu.:164.00   3rd Qu.: 0.00000
##  Max.   :63.000   Max.   :29.00000   Max.   :522.00   Max.   :65.00000
##
##     cc_liked       cc_application_begin app_downloaded     web_user
##  Min.   : 0.00000   Min.   :  0.000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.: 0.00000   1st Qu.:  0.000   1st Qu.:1.0000   1st Qu.:0.0000
##  Median : 0.00000   Median :  4.000   Median :1.0000   Median :1.0000
##  Mean   : 0.01311   Mean   :  8.198   Mean   :0.9525   Mean   :0.6061
##  3rd Qu.: 0.00000   3rd Qu.: 11.000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :27.00000   Max.   :263.000   Max.   :1.0000   Max.   :1.0000
##
##   app_web_user       ios_user        android_user    registered_phones
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :1.0000   Median :0.0000   Median :1.0000   Median :0.0000
##  Mean   :0.5617   Mean   :0.3939   Mean   :0.5873   Mean   :0.4209
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :5.0000
##
##  payment_type      waiting_4_loan     cancelled_loan     received_loan
##  Length:27000     Min.   :0.000000   Min.   :0.00000   Min.   :0.00000
##  Class :character 1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.00000
##  Mode  :character Median :0.000000   Median :0.00000   Median :0.00000
##                   Mean   :0.001296   Mean   :0.01881   Mean   :0.01819
##                   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.00000
##                   Max.   :1.000000   Max.   :1.00000   Max.   :1.00000
##
##  rejected_loan      zodiac_sign       left_for_two_month_plus
##  Min.   :0.000000   Length:27000     Min.   :0.0000
##  1st Qu.:0.000000   Class :character 1st Qu.:0.0000
##  Median :0.000000   Mode  :character Median :0.0000
##  Mean   :0.004889                    Mean   :0.1734
##  3rd Qu.:0.000000                    3rd Qu.:0.0000
##  Max.   :1.000000                    Max.   :1.0000
##
##  left_for_one_month rewards_earned   reward_rate       is_referred
##  Min.   :0.00000   Min.   :  1.00   Min.   :0.0000   Min.   :0.000
##  1st Qu.:0.00000   1st Qu.:  9.00   1st Qu.:0.2000   1st Qu.:0.000
##  Median :0.00000   Median : 25.00   Median :0.7800   Median :0.000
##  Mean   :0.01807   Mean   : 29.11   Mean   :0.9077   Mean   :0.318
##  3rd Qu.:0.00000   3rd Qu.: 48.00   3rd Qu.:1.5300   3rd Qu.:1.000
##  Max.   :1.00000   Max.   :114.00   Max.   :4.0000   Max.   :1.000
##                    NA's   :3227
```
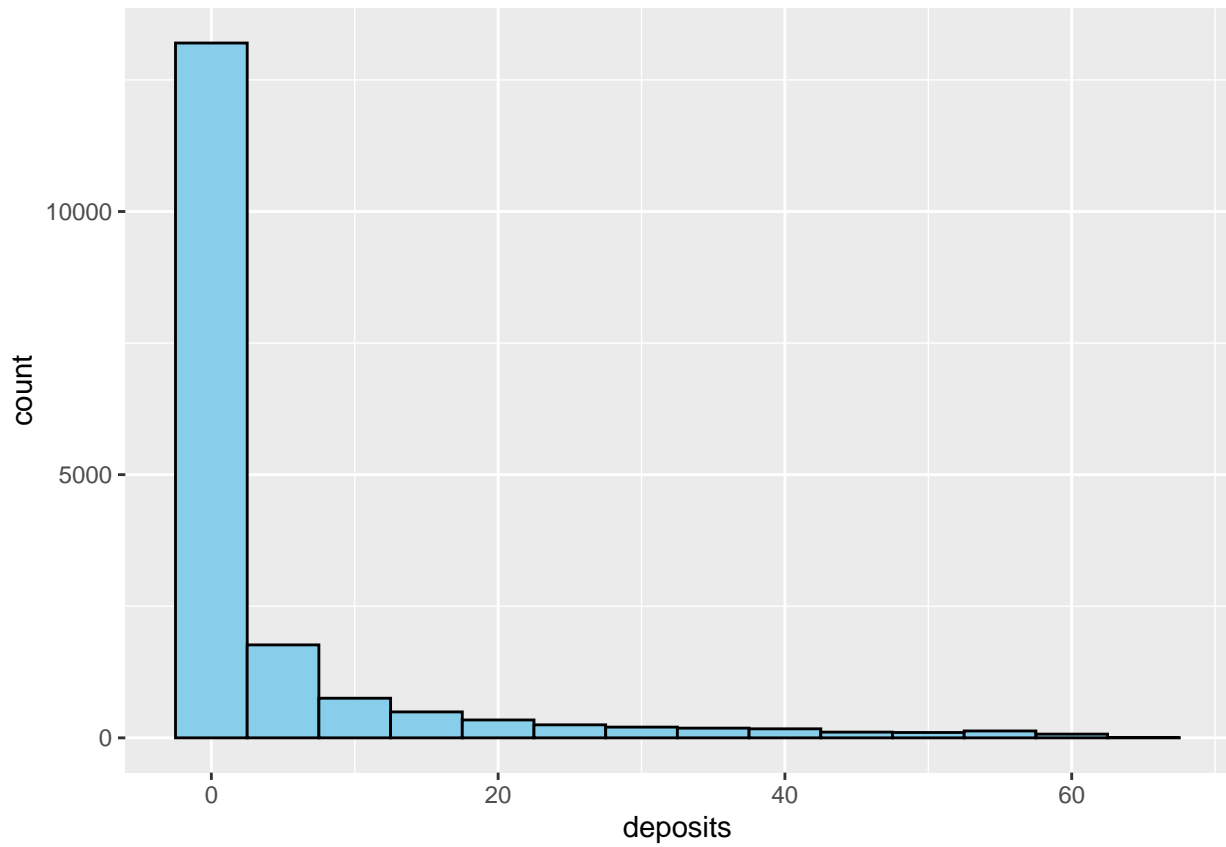
EDA

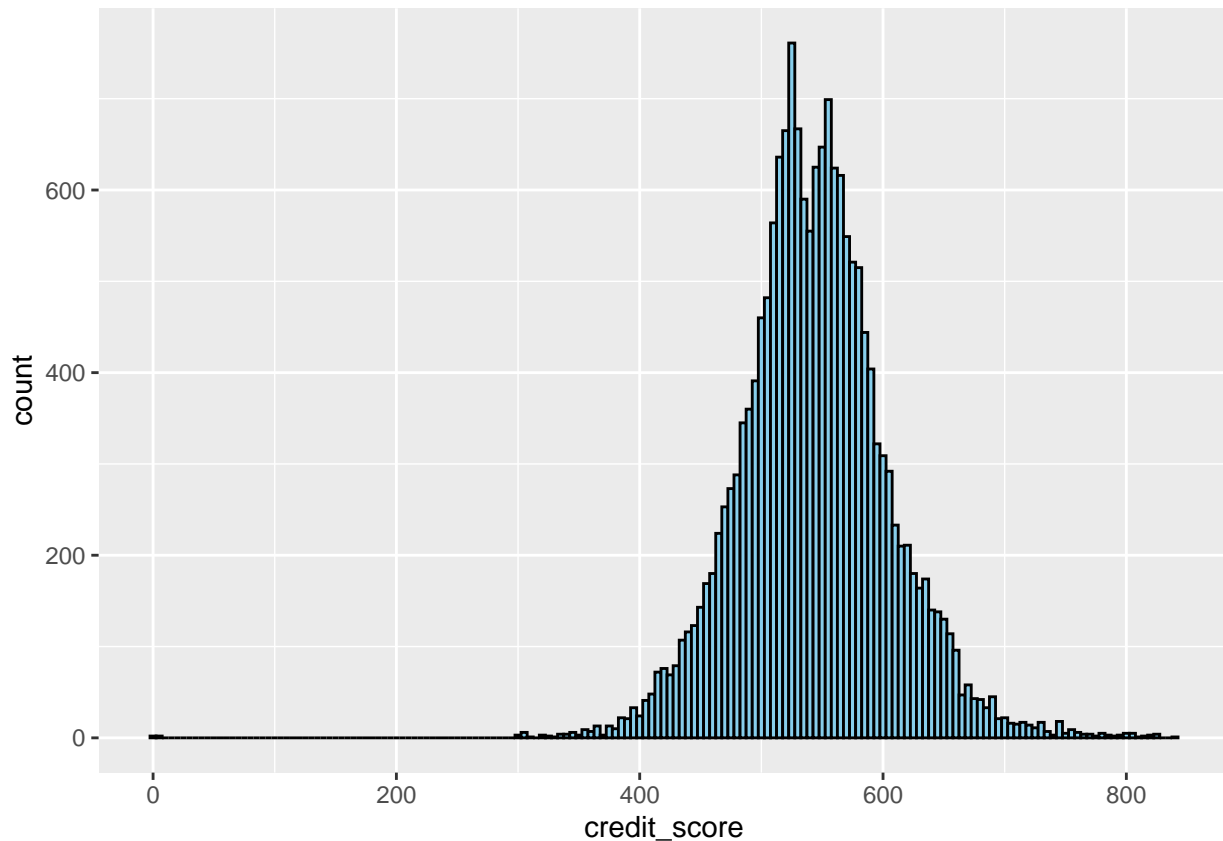Distribution of highnetworth/deposits

```
ggplot(highnetworthdata, aes(x = deposits)) +
  geom_histogram(binwidth = 5, color = "black", fill = "sky blue")
```

as we can see from the distribution plot of highnetworth/deposits, most deposits are distributed from 0 to 20, indicates the net worth of users of this financial technology company is generally not high. .
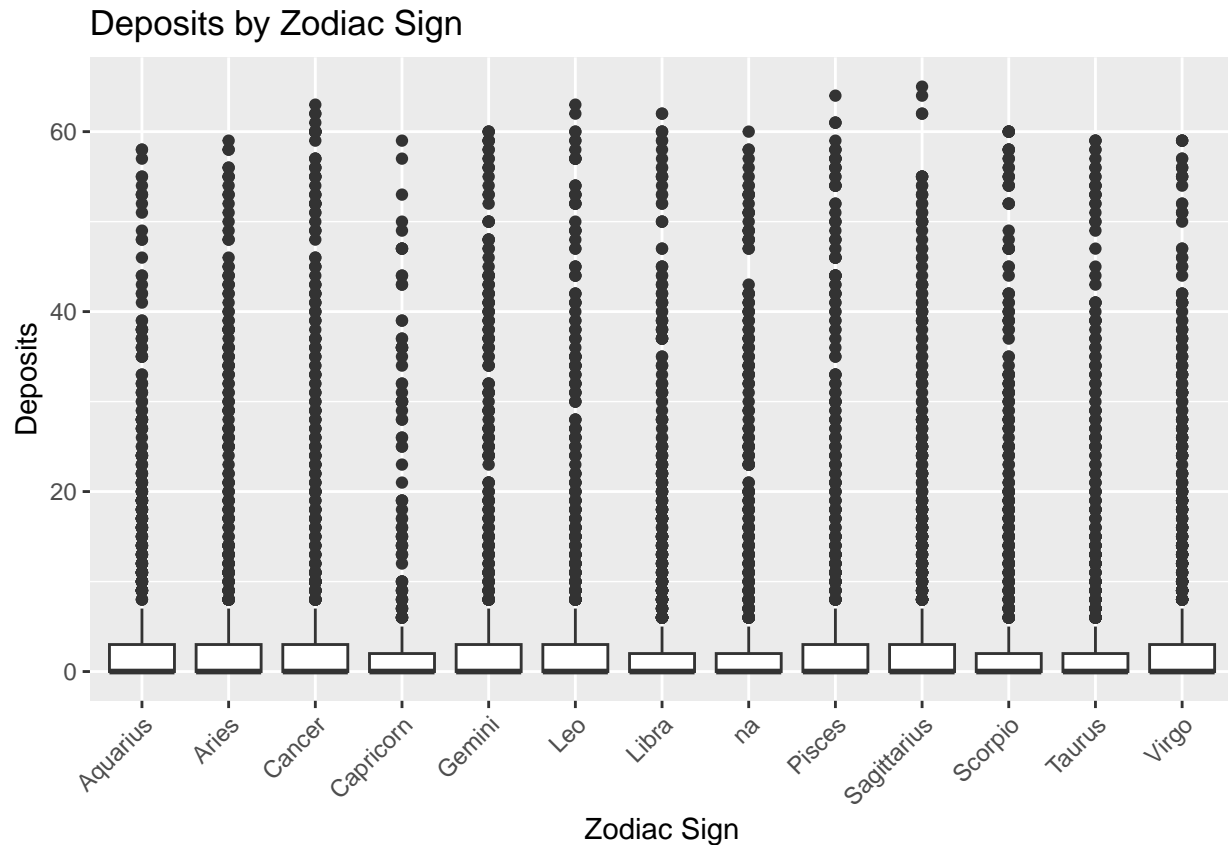
Distribution of creditscore:

in th histgram plot, the credit score distribution of high-net-worth users shows an obvious normal distribution trend, with the largest number of people in the 500-600 credit score range.

Distribution of age:

{r-} ggplot(highnetworthdata, aes(x = age)) +   geom_histogram(binwidth = 5, color = "black", fill = "sky blue") *in th histgram plot, the largest number of high-net-worth users are in the 25-30 age group, followed by the 30-25 age group, showing an overall younger trend.*

Deposits and zodiac_sign:

```r
ggplot(highnetworthdata, aes(x = zodiac_sign, y = deposits)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Deposits by Zodiac Sign", x = "Zodiac Sign", y = "Deposits")
```
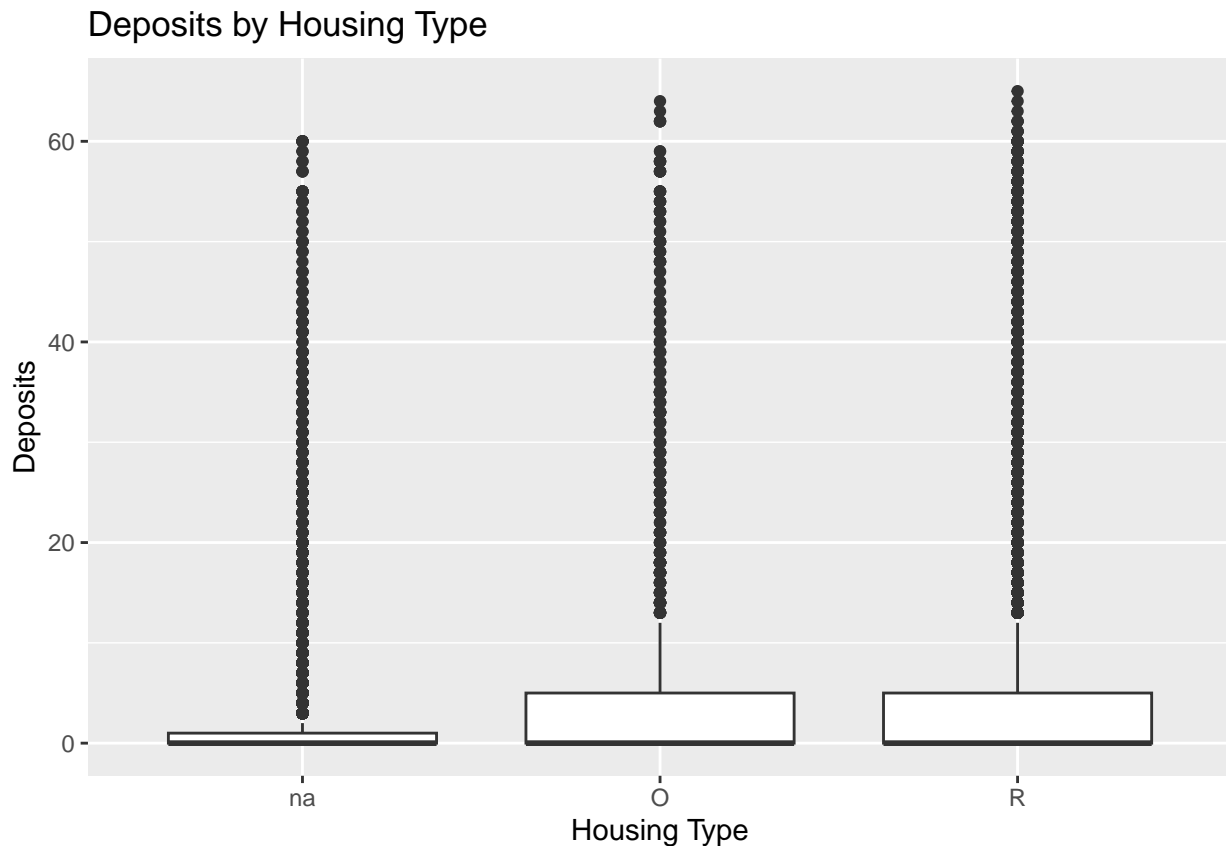
Deposits by Zodiac Sign

From the plot below, we can see the constellation distribution of the entire user does not reflect differences.

```r
highnetworthdata %>%
  group_by(zodiac_sign) %>%
  summarise(
    count = n(),
    mean_deposit = mean(deposits, na.rm = TRUE),
    median_deposit = median(deposits, na.rm = TRUE),
    sd_deposit = sd(deposits, na.rm = TRUE)
  )
```

```
## # A tibble: 13 x 5
##    zodiac_sign count mean_deposit median_deposit sd_deposit
##    <chr>       <int>        <dbl>          <dbl>      <dbl>
##  1 Aquarius     1409         3.84              0       8.99
##  2 Aries        1339         4.24              0       9.87
##  3 Cancer       1582         4.93              0      11.5
##  4 Capricorn     459         4.93              0      11.2
##  5 Gemini       1406         4.90              0      11.3
##  6 Leo          1568         4.70              0      10.6
##  7 Libra        1422         4.53              0      10.6
##  8 Pisces       1424         4.83              0      11.4
##  9 Sagittarius  1341         4.74              0      10.7
## 10 Scorpio      1385         4.59              0      10.7
## 11 Taurus       1458         4.04              0       9.89
## 12 Virgo        1575         4.25              0       9.83
## 13 na           1411         4.48              0      10.4
```

Deposits and housing:

```
ggplot(highnetworthdata, aes(x = housing, y = deposits)) +
  geom_boxplot() +
  labs(title = "Deposits by Housing Type", x = "Housing Type", y = "Deposits")
```


From the plot below, we can see the distribution of high-net-worth users of different housing types is significantly different

Before going further, we want to verify the distribution of the response variable– deposits, to test the assumption that it is normal distributed.
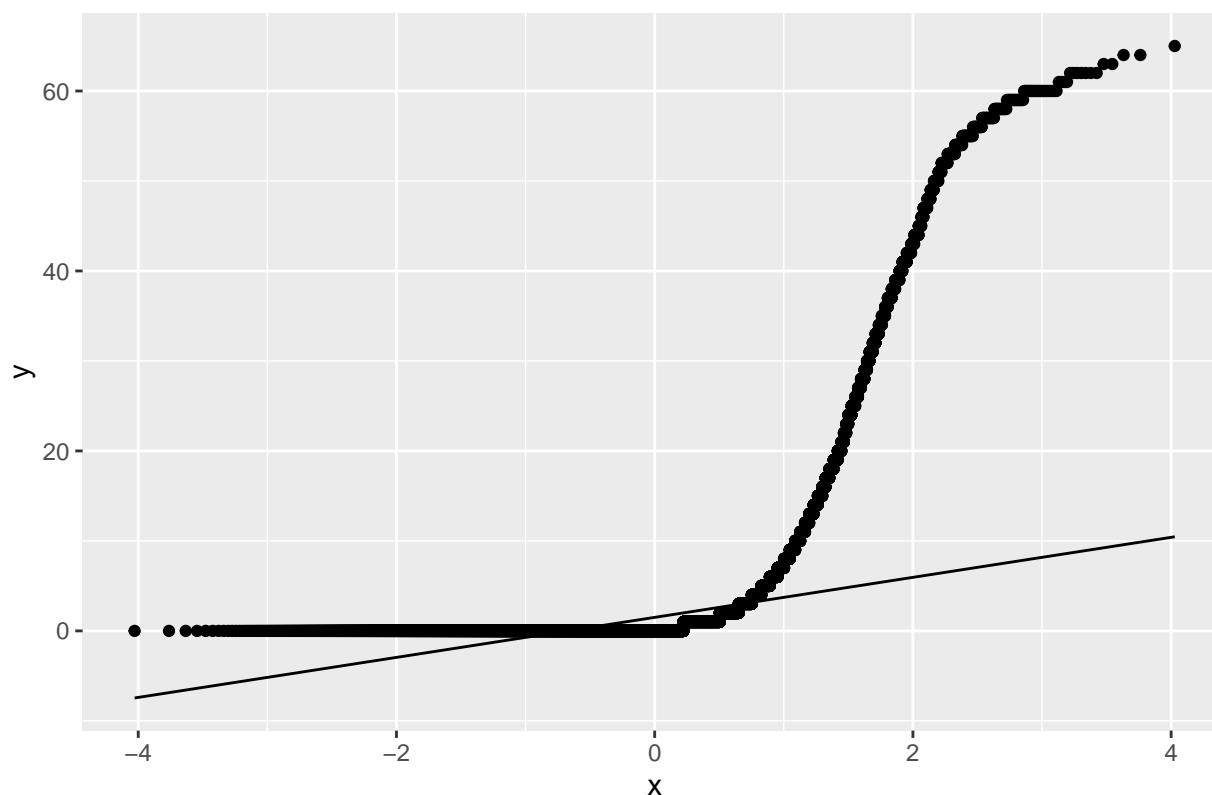
```
install.packages("nortest")
```

```
##
## The downloaded binary packages are in
##     /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```
library(nortest)

#A Q-Q plot can also help in visually inspecting the normality.

library(ggplot2)
ggplot(highnetworthdata, aes(sample = deposits)) +
  geom_qq() +
  geom_qq_line() +
  labs(title = "Q-Q Plot for Deposits")
```

Q–Q Plot for Deposits

From the Q-Q plot it looks like it follows normal distribution.

###Modelling:

####Model1: Simple linear regression:

```r
library(tidyverse)
library(ggplot2)
library(broom)
lm_model <- lm(deposits ~ age + housing + credit_score + zodiac_sign, data = highnetworthdata)
summary(lm_model)
```

```
##
## Call:
## lm(formula = deposits ~ age + housing + credit_score + zodiac_sign,
##     data = highnetworthdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.239  -5.249  -2.786  -0.501  58.319
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -5.588677   0.783792  -7.130 1.04e-12 ***
## age                0.128793   0.008284  15.547  < 2e-16 ***
## housing0           3.423417   0.275713  12.417  < 2e-16 ***
## housingR           3.731541   0.164537  22.679  < 2e-16 ***
## credit_score       0.005599   0.001272   4.403 1.08e-05 ***
## zodiac_signAries   0.342631   0.393536   0.871  0.38396
```
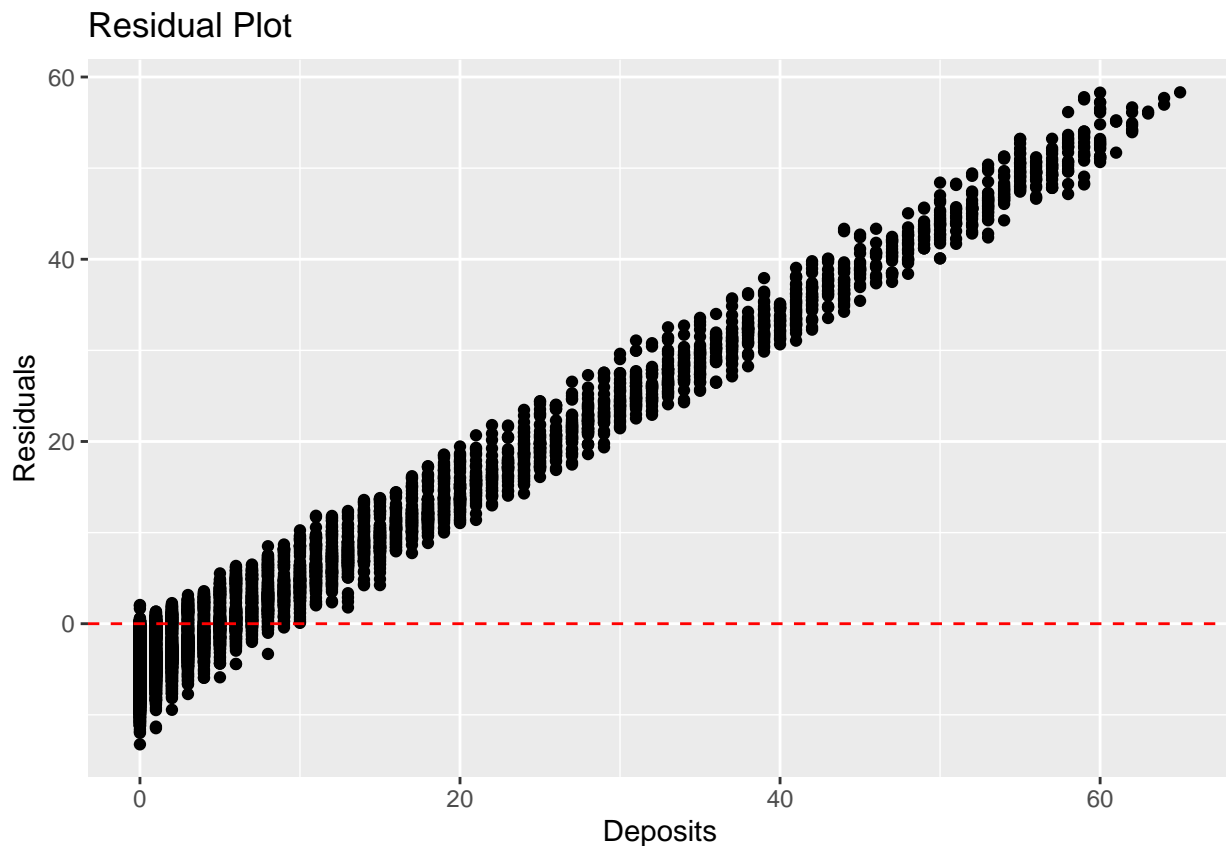
```
## zodiac_signCancer        1.047666   0.377753   2.773  0.00555 **
## zodiac_signCapricorn     0.855902   0.554260   1.544  0.12255
## zodiac_signGemini        1.090002   0.388816   2.803  0.00506 **
## zodiac_signLeo           0.845631   0.378513   2.234  0.02549 *
## zodiac_signLibra         0.751710   0.387600   1.939  0.05247 .
## zodiac_signna            0.685924   0.388427   1.766  0.07743 .
## zodiac_signPisces        0.921201   0.387506   2.377  0.01745 *
## zodiac_signSagittarius   1.063067   0.393454   2.702  0.00690 **
## zodiac_signScorpio       0.682742   0.390166   1.750  0.08016 .
## zodiac_signTaurus        0.244836   0.385235   0.636  0.52508
## zodiac_signVirgo         0.512169   0.378125   1.354  0.17560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.31 on 17762 degrees of freedom
## Multiple R-squared:  0.04531,    Adjusted R-squared:  0.04445
## F-statistic: 52.69 on 16 and 17762 DF,  p-value: < 2.2e-16
```

Create a residual plot to visualize the residuals from the linear model.
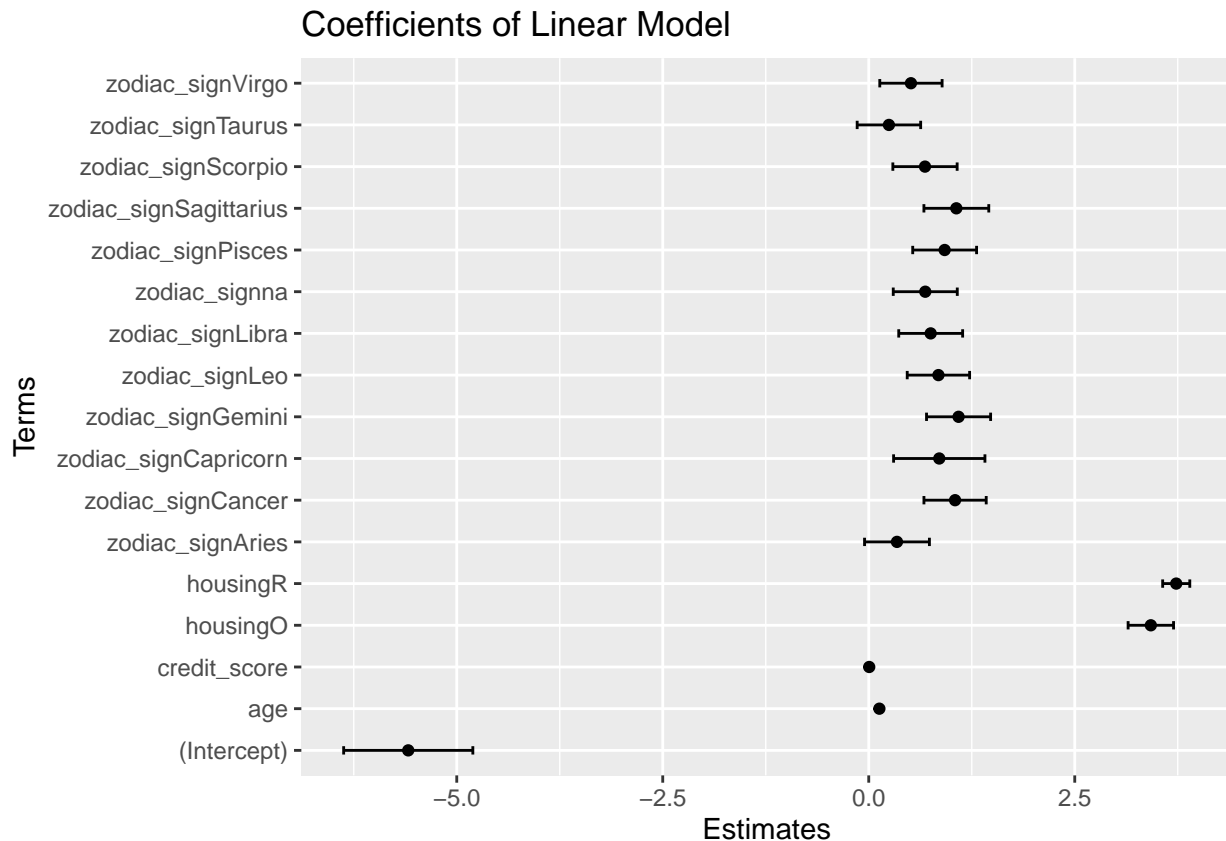
```
highnetworthdata <- highnetworthdata %>%
  add_column(residuals = resid(lm_model))

ggplot(highnetworthdata, aes(x = deposits, y = residuals)) +
  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  labs(title = "Residual Plot", x = "Deposits", y = "Residuals")
```

visualize the coefficients.

```
tidied_model <- tidy(lm_model)
ggplot(tidied_model, aes(x = term, y = estimate)) +
  geom_point() +
  geom_errorbar(aes(ymin = estimate - std.error, ymax = estimate + std.error), width = 0.2) +
  coord_flip() +
  labs(title = "Coefficients of Linear Model", x = "Terms", y = "Estimates")
```

Coefficients of Linear Model

From this simple linear regression, I found all the predictors are significant. Also the R-square in the model is 0.045, So the model is well fitted.

####Model2: Multilevel linear model with random intercept:

```
library(lme4)
library(ggplot2)
library(broom)

# Fit a multilevel model with random intercepts
mlm_model <- lmer(deposits ~ age + housing + credit_score + zodiac_sign + (1 | age), data = highnetworth
summary(mlm_model)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: deposits ~ age + housing + credit_score + zodiac_sign + (1 |     age)
##    Data: highnetworthdata
##
## REML criterion at convergence: 133420.4
##
## Scaled residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.1532 -0.4980 -0.2794 -0.0475  5.6649
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  age      (Intercept)   0.511   0.7148
##  Residual             106.003  10.2958
## Number of obs: 17779, groups:  age, 65
##
## Fixed effects:
##                        Estimate Std. Error t value
## (Intercept)           -4.740827   0.863138  -5.493
## age                    0.105005   0.011928   8.803
## housingO               3.416366   0.275531  12.399
## housingR               3.723617   0.164680  22.611
## credit_score           0.005578   0.001272   4.387
## zodiac_signAries       0.324019   0.393362   0.824
## zodiac_signCancer      1.032253   0.377427   2.735
## zodiac_signCapricorn   0.822644   0.553806   1.485
## zodiac_signGemini      1.088986   0.388566   2.803
## zodiac_signLeo         0.824279   0.378422   2.178
## zodiac_signLibra       0.735065   0.387454   1.897
## zodiac_signna          0.674818   0.388326   1.738
## zodiac_signPisces      0.910626   0.387283   2.351
## zodiac_signSagittarius 1.038934   0.393184   2.642
## zodiac_signScorpio     0.679210   0.390070   1.741
## zodiac_signTaurus      0.250643   0.385072   0.651
## zodiac_signVirgo       0.509509   0.377979   1.348
##
## Correlation matrix not shown by default, as p = 17 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```r
print(summary(lm_model), correlation = TRUE)
```

```
##
## Call:
## lm(formula = deposits ~ age + housing + credit_score + zodiac_sign,
##     data = highnetworthdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.239  -5.249  -2.786  -0.501  58.319
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -5.588677   0.783792  -7.130 1.04e-12 ***
## age                    0.128793   0.008284  15.547  < 2e-16 ***
## housingO               3.423417   0.275713  12.417  < 2e-16 ***
## housingR               3.731541   0.164537  22.679  < 2e-16 ***
## credit_score           0.005599   0.001272   4.403 1.08e-05 ***
## zodiac_signAries       0.342631   0.393536   0.871  0.38396
## zodiac_signCancer      1.047666   0.377753   2.773  0.00555 **
## zodiac_signCapricorn   0.855902   0.554260   1.544  0.12255
```

```
## zodiac_signGemini      1.090002    0.388816    2.803   0.00506 **
## zodiac_signLeo         0.845631    0.378513    2.234   0.02549 *
## zodiac_signLibra       0.751710    0.387600    1.939   0.05247 .
## zodiac_signna          0.685924    0.388427    1.766   0.07743 .
## zodiac_signPisces      0.921201    0.387506    2.377   0.01745 *
## zodiac_signSagittarius 1.063067    0.393454    2.702   0.00690 **
## zodiac_signScorpio     0.682742    0.390166    1.750   0.08016 .
## zodiac_signTaurus      0.244836    0.385235    0.636   0.52508
## zodiac_signVirgo       0.512169    0.378125    1.354   0.17560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.31 on 17762 degrees of freedom
## Multiple R-squared:  0.04531,    Adjusted R-squared:  0.04445
## F-statistic: 52.69 on 16 and 17762 DF,  p-value: < 2.2e-16
```

```r
install.packages("broom.mixed")
```

```
##
## The downloaded binary packages are in
##   /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
library(broom.mixed)

fixed_effects <- tidy(mlm_model, effects = "fixed")
ggplot(fixed_effects, aes(x = term, y = estimate)) +
  geom_point() +
  geom_errorbar(aes(ymin = estimate - std.error, ymax = estimate + std.error), width = 0.2) +
  coord_flip() +
  labs(title = "Fixed Effect Coefficients of Multilevel Model", x = "Terms", y = "Estimates")
```
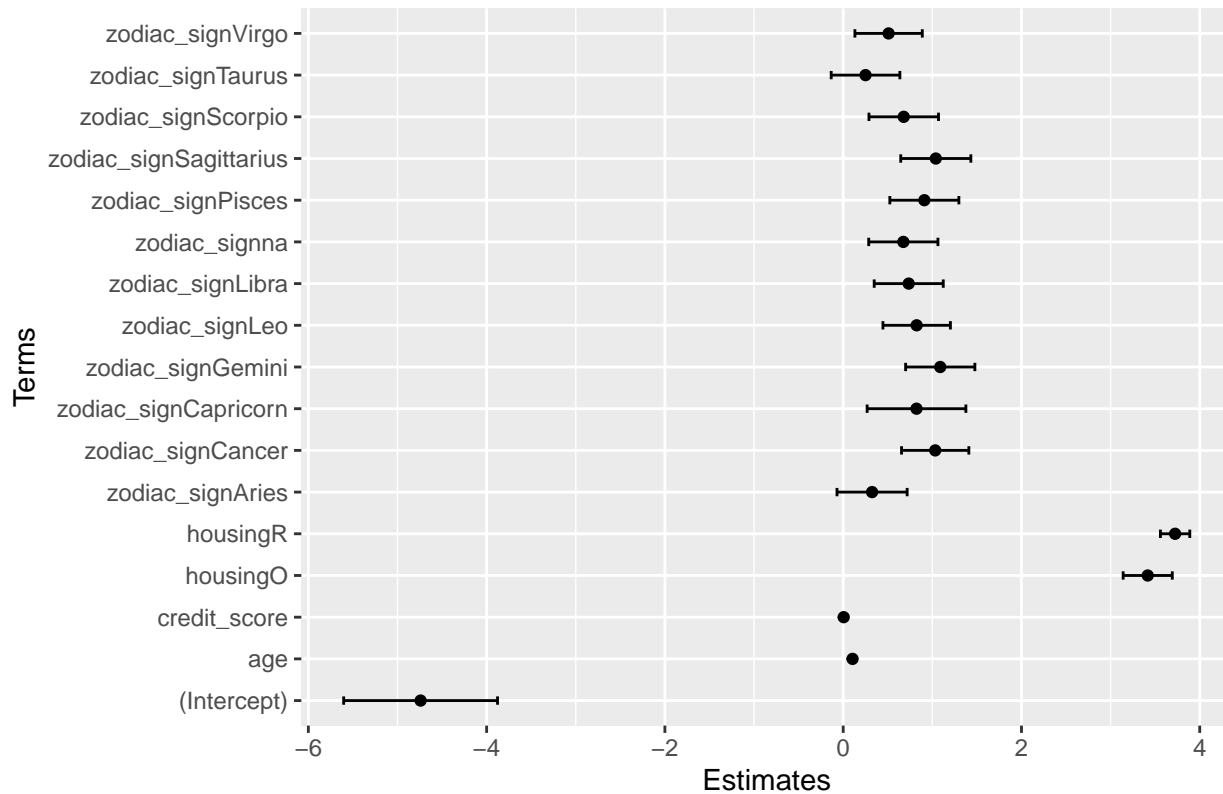
Fixed Effect Coefficients of Multilevel Model

Model3 : Multilevel linear model with random slope:

```r
install.packages("lme4")
```

```
##
## The downloaded binary packages are in
## /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
install.packages("broom.mixed")
```

```
##
## The downloaded binary packages are in
## /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
library(lme4)
library(broom.mixed)
library(ggplot2)

install.packages("broom.mixed")
```

```
##
## The downloaded binary packages are in
## /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
library(broom.mixed)
install.packages("lme4")
```

```
##
## The downloaded binary packages are in
## /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
library(lme4)

mlm_model2 <- lmer(deposits ~ age + housing + credit_score + zodiac_sign + (credit_score | age), data =
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 12.7435 (tol = 0.002, component 1)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly uniden
##   - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##   - Rescale variables?
```

```r
summary(mlm_model)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: deposits ~ age + housing + credit_score + zodiac_sign + (1 |       age)
##    Data: highnetworthdata
##
## REML criterion at convergence: 133420.4
##
## Scaled residuals:
##     Min     1Q  Median     3Q     Max
## -1.1532 -0.4980 -0.2794 -0.0475  5.6649
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  age      (Intercept)   0.511   0.7148
##  Residual             106.003  10.2958
## Number of obs: 17779, groups:  age, 65
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)           -4.740827   0.863138  -5.493
## age                    0.105005   0.011928   8.803
## housing0               3.416366   0.275531  12.399
## housingR               3.723617   0.164680  22.611
## credit_score           0.005578   0.001272   4.387
## zodiac_signAries       0.324019   0.393362   0.824
## zodiac_signCancer      1.032253   0.377427   2.735
## zodiac_signCapricorn   0.822644   0.553806   1.485
## zodiac_signGemini      1.088986   0.388566   2.803
## zodiac_signLeo         0.824279   0.378422   2.178
## zodiac_signLibra       0.735065   0.387454   1.897
## zodiac_signna          0.674818   0.388326   1.738
## zodiac_signPisces      0.910626   0.387283   2.351
## zodiac_signSagittarius 1.038934   0.393184   2.642
## zodiac_signScorpio     0.679210   0.390070   1.741
## zodiac_signTaurus      0.250643   0.385072   0.651
## zodiac_signVirgo       0.509509   0.377979   1.348
##
## Correlation matrix not shown by default, as p = 17 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```r
print(summary(mlm_model2), correlation = TRUE)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## deposits ~ age + housing + credit_score + zodiac_sign + (credit_score |
##     age)
##    Data: highnetworthdata
##
## REML criterion at convergence: 133414
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.1772 -0.4988 -0.2788 -0.0435  5.6666
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.  Corr
##  age      (Intercept) 1.632e+01 4.040353
##           credit_score 5.664e-05 0.007526 -0.99
##  Residual             1.059e+02 10.289010
## Number of obs: 17779, groups:  age, 65
##
## Fixed effects:
##                        Estimate Std. Error t value
## (Intercept)           -4.638721   1.087343  -4.266
## age                    0.108907   0.011367   9.581
## housing0               3.413470   0.275570  12.387
## housingR               3.724999   0.164638  22.625
## credit_score           0.005160   0.001832   2.816
## zodiac_signAries       0.323833   0.393213   0.824
## zodiac_signCancer      1.032497   0.377340   2.736
## zodiac_signCapricorn   0.810668   0.553674   1.464
## zodiac_signGemini      1.067781   0.388489   2.749
## zodiac_signLeo         0.814394   0.378252   2.153
## zodiac_signLibra       0.712410   0.387324   1.839
## zodiac_signna          0.662707   0.388157   1.707
## zodiac_signPisces      0.897928   0.387212   2.319
## zodiac_signSagittarius 1.012750   0.393088   2.576
## zodiac_signScorpio     0.669521   0.389933   1.717
## zodiac_signTaurus      0.247941   0.384982   0.644
## zodiac_signVirgo       0.485074   0.377878   1.284
##
## Correlation of fixed effects could have been required in summary()
##
## Correlation of Fixed Effects:
##             (Intr) age    hosng0 hosngR crdt_s zdc_sA zdc_sgnCn zdc_sgnCp
## age         -0.316
## housing0     0.014 -0.072
## housingR    -0.093  0.021  0.335
## credit_scor -0.884 -0.060 -0.048 -0.001
## zodc_sgnArs -0.175  0.000  0.002  0.000 -0.002
## zdc_sgnCncr -0.177 -0.002  0.011  0.003 -0.007  0.507
## zdc_sgnCprc -0.114 -0.007  0.001 -0.008 -0.009  0.346  0.360
## zodc_sgnGmn -0.168 -0.010  0.014  0.015 -0.009  0.493  0.514     0.350
## zodiac_sgnL -0.186  0.002  0.007 -0.004  0.002  0.507  0.527     0.360
## zodc_sgnLbr -0.185  0.004  0.005  0.002  0.004  0.495  0.515     0.351
```

```
## zodiac_sgnn -0.175 -0.005  0.020  0.008 -0.003  0.494  0.514     0.351
## zdc_sgnPscs -0.174 -0.005  0.009  0.002 -0.003  0.495  0.515     0.352
## zdc_sgnSgtt -0.174  0.013  0.006  0.015 -0.009  0.487  0.507     0.346
## zdc_sgnScrp -0.174 -0.004  0.002 -0.002 -0.003  0.492  0.512     0.349
## zodc_sgnTrs -0.176  0.002  0.007  0.009 -0.007  0.498  0.519     0.353
## zodc_sgnVrg -0.185  0.005  0.010  0.008 -0.001  0.507  0.528     0.360
##            zdc_sG zodc_sgnL zdc_sgnLb zdc_sg zdc_sP zdc_sgnSg zdc_sgnSc zdc_sT
## age
## housing0
## housingR
## credit_scor
## zodc_sgnArs
## zdc_sgnCncr
## zdc_sgnCprc
## zodc_sgnGmn
## zodiac_sgnL  0.513
## zodc_sgnLbr  0.501  0.514
## zodiac_sgnn  0.500  0.514      0.501
## zdc_sgnPscs  0.501  0.515      0.502      0.502
## zdc_sgnSgtt  0.494  0.506      0.495      0.494  0.495
## zdc_sgnScrp  0.498  0.511      0.499      0.498  0.499  0.491
## zodc_sgnTrs  0.504  0.517      0.506      0.505  0.506  0.498      0.502
## zodc_sgnVrg  0.513  0.527      0.515      0.514  0.515  0.507      0.512      0.518
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 12.7435 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```r
install.packages("broom")
```

```
##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
install.packages("broom.mixed")
```

```
##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```
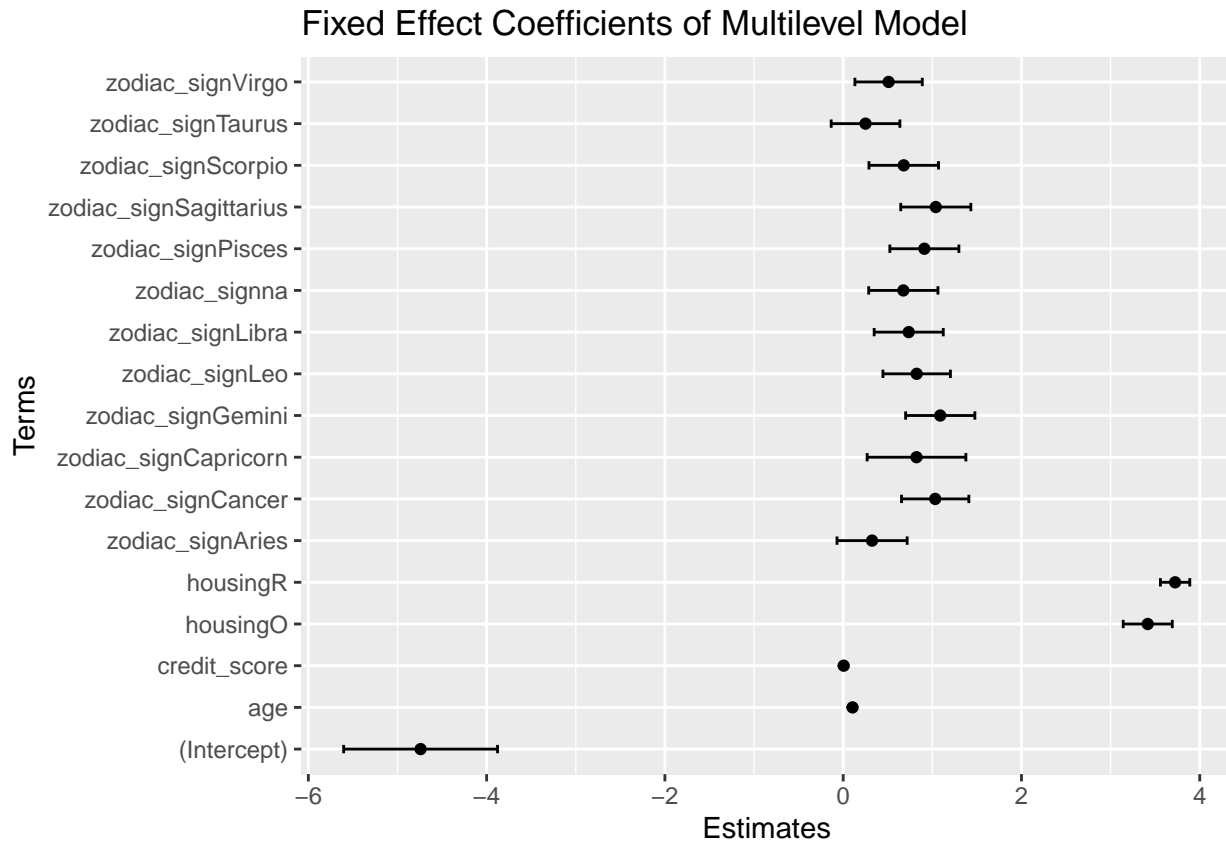
```r
install.packages("ggplot2")
```

```
##
## The downloaded binary packages are in
##  /var/folders/30/0l1gbkp92xz5l1zm5z2k5pvh0000gn/T//Rtmp2FbCDT/downloaded_packages
```

```r
library(broom)
library(broom.mixed)
library(ggplot2)

fixed_effects <- tidy(mlm_model, effects = "fixed")
ggplot(fixed_effects, aes(x = term, y = estimate)) +
  geom_point() +
  geom_errorbar(aes(ymin = estimate - std.error, ymax = estimate + std.error), width = 0.2) +
  coord_flip() +
```

```
labs(title = "Fixed Effect Coefficients of Multilevel Model", x = "Terms", y = "Estimates")
```


Fixed Effect Coefficients of Multilevel Model

This is a multilevel linear model with random slope. As we can see the model is not very fitted.

##IV. Result:

###Model Choice and Interpretation use Anova:

```
library(lme4)
library(Matrix)
library(broom.mixed)
tidied_model <- tidy(mlm_model)
fixed_effects <- fixef(mlm_model)
random_effects <- ranef(mlm_model2)
fixed_effects <- fixef(mlm_model2)
```

##V. Discussion:

###Implication *The result indicates the factor of credit score will determine the depositsd . Also, the factors age and housing are obvious to be significant.*

###Limitation *Due to the not good time management of the project, this report does not provide a good analysis of the model's fit.*

##VI.Reference and Appendix

https://www.kaggle.com/