TERM DEPOSIT ANALYSIS

INTRODUCTION

What is a term deposit?

- A term deposit is a type of deposit account held at a financial institution where money is locked up for some set period of time.

- Usually, short-term deposits with maturities ranging from one month to a few years

INTRODUCTION

Why use a term deposit?

- Could help you achieve your savings goals

- In exchange for locking your money away with a bank for a set period of time, you could obtain a competitive fixed interest rate.

- Therefore, you make some money off what you put away with the bank

MY DATASET

- A COLLECTION OF DATA FOCUSED ON A BANK'S MARKETING CAMPAIGN TO ACQUIRE DEPOSITS FROM CUSTOMERS.

- CONTAINS VARIABLES LIKE AGE, JOB, MARITAL STATUS, EDUCATION, WHETHER THEY HAVE A LOAN OR NOT, OUTCOME OF THE MARKETING CAMPAIGN, ETC.

- VALUABLE FOR ANALYZING THE FACTORS INFLUENCING CUSTOMERS' DECISIONS TO SUBSCRIBE TO TERM DEPOSITS, AS WELL AS FOR PREDICTING CUSTOMER BEHAVIOR IN FUTURE SIMILAR MARKETING CAMPAIGNS

- UNDERSTANDING THIS DATASET, BANKS OR MARKETING ANALYSTS CAN OPTIMIZE THEIR MARKETING STRATEGIES TO ENHANCE THE SUCCESS OF DEPOSIT CAMPAIGNS IN THE FUTURE

CLEANING

- A lot of my data had answers like yes or no…

housing	loan	deposit
yes	no	yes
no	no	yes
yes	no	yes
yes	no	yes
no	no	yes
yes	yes	yes
yes	yes	yes
yes	no	yes
yes	no	yes
yes	no	yes
yes	no	yes

CLEANING

- To fix this…

```
#Add a column that is a binary form of deposit:
#0 = No Deposit, 1 = Deposit
deposit_data_file$deposit_binary=0
deposit_data_file$deposit_binary[which(deposit_data_file$deposit=="yes")]=1
head(deposit_data_file)
View(deposit_data_file)

#Add a column that is a binary form of housing:
#0 = No housing Loan, 1 = Housing Loan
deposit_data_file$housing_loan_binary=0
deposit_data_file$housing_loan_binary[which(deposit_data_file$housing=="yes")]=1

#Add a column that is a binary form of loan:
#0 = No personal Loan, 1 = personal Loan
deposit_data_file$personal_loan_binary=0
deposit_data_file$personal_loan_binary[which(deposit_data_file$loan=="yes")]=1
```

CLEANING

- Results: three new columns

deposit_binary	housing_loan_binary	personal_loan_binary
1	1	0
1	0	0
1	1	0
1	1	0
1	0	0
1	1	1

- Luckily, no missing data to clean

MODELING

- Started with doing a linear model with all the variables predicting whether a customer made a term deposit (deposit_binary)

- About ¾ of the variable were significant

- Made a second model taking out all the terms that weren't significant in the first one.

- When looking into the VIF, I found one variable that was above 2

MODELING

	GVIF
job	2.659136
marital	1.127448
education	2.264530
balance	1.024645
housing	1.141547
loan	1.030310
duration	1.277427
campaign	1.026697
previous	1.457097
poutcome	1.683507
deposit	1.480699

- For the third linear model, I removed job due to the VIF being above 2

- To make sure I had the best model, I decided to create a logistic regression with all the variables that were highly significant in the linear models

COMPARING THE TWO MODELS

- Ran ANOVA on the linear and logistic regression models
- P-Value is highly significant at 2.2e-16 which shows there is a difference in the two models
- Model 1, the linear model has an RSS of 0 indicating it is better than model 2

```
Analysis of Variance Table

Model 1: deposit_binary ~ (age + job + marital + education + default +
    balance + housing + loan + contact + day + month + duration +
    campaign + pdays + previous + poutcome + deposit + housing_loan_binary +
    personal_loan_binary) - month - age - pdays - default - contact -
    day - housing_loan_binary - personal_loan_binary - job
Model 2: deposit_binary ~ balance + campaign + housing_loan_binary + personal_loan_binary +
    marital + poutcome
  Res.Df   RSS Df Sum of Sq        F    Pr(>F)
1  56357     0
2  56363 47795 -6    -47795 5.4092e+32 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

COMPARING THE TWO MODELS

- Ran AIC on the linear and logistic regression models

- Although I found the linear model to be better, the logistic regression has a lower degrees of freedom at 10 compared to the linear model at 17

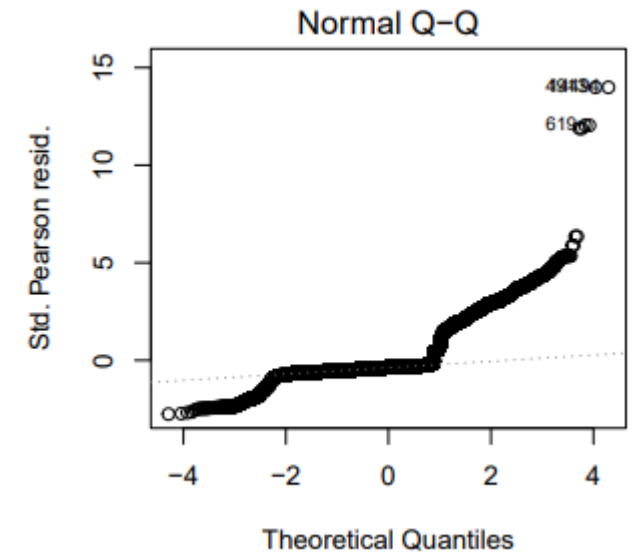- The linear model beats the logistic model in AIC with a very negative amount

```
                 df          AIC
linear_model3 17 -3582488.20
glm.fits      10     47815.13
```

COMPARING THE TWO MODELS

- The normal Q-Q for the linear model has a very straight line which is desirable

- Compared to the logistic model which has a zigzag or crooked line
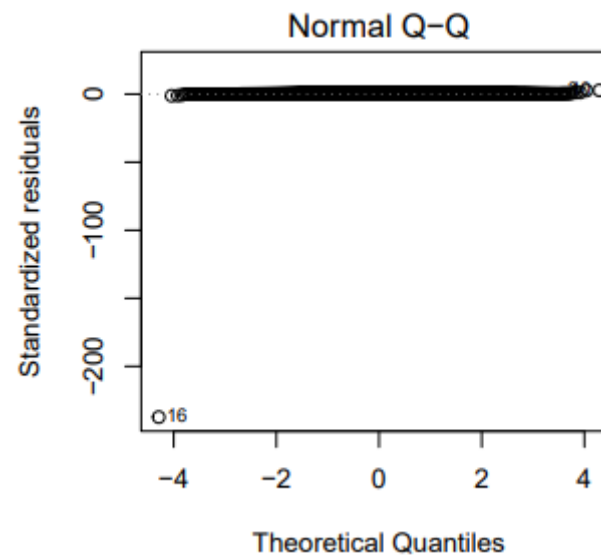
LINEAR MODEL

LOGISTIC MODEL

WHAT WAS LEARNED?

- Variables that were very significant in all models for whether a customer made a term deposit were: what the balance was in the customers account, if they had a personal or housing loan, how many times they were contacting during the time of the banks marketing campaign, and their marital status

WHAT WAS LEARNED?

Variables that surprisingly didn't influence whether a customer made a term deposit or not were: the customers age, job, outcome of the previous marketing campaign, and whether the customer has credit in default

SOURCES

- "Bank Target Marketing." Kaggle, 29 Feb. 2024, www.kaggle.com/datasets/seanangelonathanael/bank-target-marketing?resource=download.

- Limited, Bank of The Bahamas. "Term Deposits." Bank of The Bahamas Limited, www.bankbahamas.com/term-deposits. Accessed 18 Apr. 2024.

- "The Pros and Cons of Investing in Term Deposit: Qudos Bank." The Pros and Cons of Investing in Term Deposit | Qudos Bank, www.qudosbank.com.au/blog/blogs/investing-term-deposits/#:~:text=If%20you're%20looking%20for,a%20competitive%20fixed%20interest%20rate. Accessed 18 Apr. 2024.

- "Term Deposit vs Fixed Deposit - a Detailed Comparison." Term Deposit Vs Fixed Deposit: A Detailed Comparison | HDFC Bank, www.hdfcbank.com/personal/resources/learning-centre/save/term-deposit-vs-fixed-deposit#:~:text=In%20Term%20Deposits%2C%20the%20sum,up%20to%20a%20particular%20term. Accessed 18 Apr. 2024.