

# Term Deposit Prediction & Marketing Analysis

Targeting Customers that are likely to opt into a Term Deposit service

*Business Science*

*2/13/2019*

## Executive Summary

- Duration since last contact and success of prior enrollments are correlated to success of Term Deposit opt-in
- A strategy targeting customers that haven't been contacted in 319 days and have those with prior enrollments yields 32% positive outcomes (enrollments) vs 4.3% otherwise.

## Analysis

```
# Libraries
library(tidyverse)
library(readxl)
library(recipes)
library(tidyquant)
library(ggrepel)

# 1.0 READ EXCEL SHEETS ----
path  <- "data/bank_term_deposit_marketing_analysis.xlsx"
sheets <- excel_sheets(path)

# # 2.0 INVESTIGATE DATA FOR EACH SHEET ----
# sheets %>%
#   map(~ read_excel(path = path, sheet = .)) %>%
#   set_names(sheets)

# 3.0 PERFORM VLOOKUP EQUIVALENT ----
data_joined_tbl <- sheets[4:7] %>%
  map(~ read_excel(path = path, sheet = .)) %>%
  reduce(left_join)

# 4.0 ANALYZE ----

# 4.1 Prepare Data ----
# recipes reference: https://tidymodels.github.io/recipes/reference/index.html

recipe_obj <- recipe(~ ., data = data_joined_tbl) %>%
  step_rm(ID) %>%
  step_discretize(all_numeric(), options = list(min_unique = 1)) %>%
```

```

    step_dummy(all_nominal(), one_hot = TRUE, naming = partial(dummy_names, sep = "__")) %>%
    prep()

data_transformed_tbl <- data_joined_tbl %>%
  bake(recipe_obj, new_data = .)

# 4.2 Correlation Analysis ----

# Prepare Correlations
correlation_tbl <- data_transformed_tbl %>%
  cor(y = data_transformed_tbl$TERM_DEPOSIT__yes) %>%
  as_tibble(rownames = "feature") %>%
  rename(TERM_DEPOSIT__yes = V1) %>%
  separate(feature, into = c("feature", "bin"), sep = "__") %>%
  filter(!is.na(TERM_DEPOSIT__yes)) %>%
  filter(!str_detect(feature, "TERM_DEP")) %>%
  arrange(abs(TERM_DEPOSIT__yes) %>% desc()) %>%
  mutate(feature = as_factor(feature) %>% fct_rev())

```

Correlation of various features to Term Deposit enrollment. Duration since last contact and success of prior enrollments are correlated to success of Term Deposit opt-in.

```

# Visualize Correlations
correlation_tbl %>%

  ggplot(aes(TERM_DEPOSIT__yes, y = feature, text = bin)) +

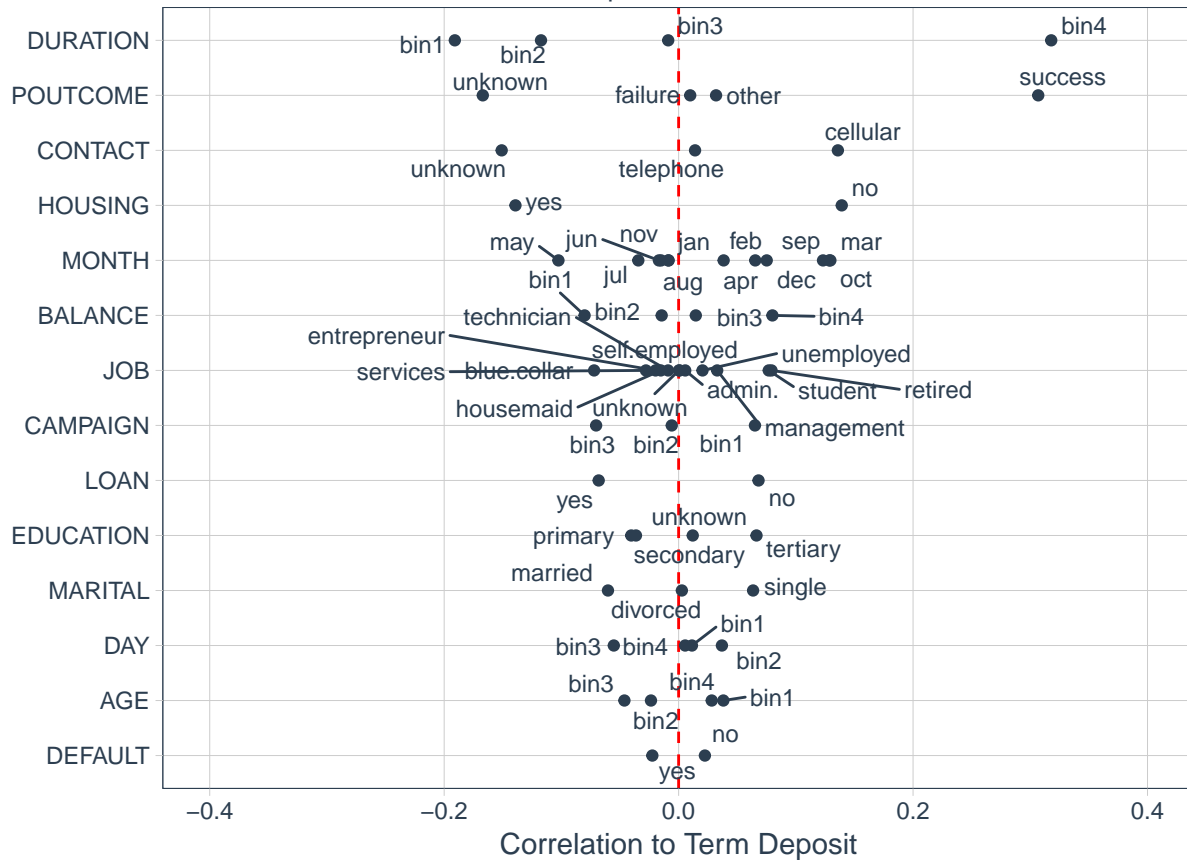
  # Geometries
  geom_vline(xintercept = 0, linetype = 2, color = "red") +
  geom_point(color = "#2c3e50") +
  geom_text_repel(aes(label = bin), size = 3, color = "#2c3e50") +

  # Formatting
  expand_limits(x = c(-0.4, 0.4)) +
  theme_tq() +
  labs(title = "Bank Marketing Analysis",
       subtitle = "Correlations to Enrollment in Term Deposit",
       y = "", x = "Correlation to Term Deposit")

```

## Bank Marketing Analysis

Correlations to Enrollment in Term Deposit



## Strategy

A strategy targeting customers that haven't been contacted in 319 days and have those with prior enrollments yields 32% positive outcomes (enrollments) vs 4.3% otherwise.

```
strategy_tbl <- data_joined_tbl %>%
  select(DURATION, POUTCOME, TERM_DEPOSIT) %>%
  mutate(POTENTIAL = case_when(
    DURATION > 319 ~ "High Potential",
    POUTCOME == "success" ~ "High Potential",
    TRUE ~ "Normal"
  )) %>%
  group_by(POTENTIAL) %>%
  count(TERM_DEPOSIT) %>%
  mutate(prop = n / sum(n)) %>%
  ungroup() %>%
  mutate(label_text = str_glue("n: {n}
                                prop: {scales::percent(prop)}"))

strategy_tbl %>%
  ggplot(aes(POTENTIAL, prop, fill = TERM_DEPOSIT)) +
```

```

geom_col() +
geom_label(aes(label = label_text), fill = "white", color = "#2c3e50") +
scale_fill_tq() +
scale_y_continuous(labels = scales::percent_format()) +
theme_tq() +
labs(title = "Bank Marketing Strategy",
      subtitle = str_glue("Targeting customers that haven't been contacted in 319 days
                           or those with prior enrollments yields 32% vs 4.3%")
)

```

