Problem Solving Test

Lauw Dastin Falentino Laurianto

2023-08-25

```
#Setup Library
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.3
```

```
## Warning: package 'tibble' was built under R version 4.1.3
```

```
## Warning: package 'tidyr' was built under R version 4.1.3
```

```
## Warning: package 'readr' was built under R version 4.1.3
```

```
## Warning: package 'purrr' was built under R version 4.1.3
```

```
## Warning: package 'dplyr' was built under R version 4.1.3
```

```
## Warning: package 'stringr' was built under R version 4.1.3
```

```
## Warning: package 'forcats' was built under R version 4.1.3
```

```
## Warning: package 'lubridate' was built under R version 4.1.3
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to force all conflicts to become errors
```

```
#Read CSV
dataset_1 <- read.csv("Stockbit_Bibit_PST_dataset1.csv", header = TRUE)
dataset_2 <- read.csv("Stockbit_Bibit_PST_dataset2.csv", header = TRUE)

dataset_2  <- dataset_2[order(dataset_2$date, decreasing = TRUE), ]
unique_dataset_2 <- select(dataset_2, -c("date"))
unique_dataset_2 <- distinct(unique_dataset_2, user_id, .keep_all = TRUE)
```

```
i = 1
for (x in unique_dataset_2$user_id) {
    if (is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
            is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

      unique_dataset_2$flag[i] <- "Saham Only"

    } else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
            is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

      unique_dataset_2$flag[i] <- "Pasar Uang Only"

    } else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
            is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

      unique_dataset_2$flag[i] <- "Pendapatan Tetap Only"
```

```
} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "Campuran Only"

}  else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

  unique_dataset_2$flag[i] <- "Saham & PU"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

  unique_dataset_2$flag[i] <- "Saham & PT"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "Saham & Campuran"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

  unique_dataset_2$flag[i] <- "PU & PT"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "PU & Campuran"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "PT & Campuran"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == TRUE) {

  unique_dataset_2$flag[i] <- "Exclude Campuran"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == TRUE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "Exclude PT"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "Exclude PU"

} else if(is.na(unique_dataset_2$Saham_invested_amount[i]) == TRUE &&
    is.na(unique_dataset_2$Pasar_Uang_invested_amount[i]) == FALSE &&
    is.na(unique_dataset_2$Pendapatan_Tetap_invested_amount[i]) == FALSE &&
        is.na(unique_dataset_2$Campuran_invested_amount[i]) == FALSE) {

  unique_dataset_2$flag[i] <- "Exclude Saham"

} else {
```

```
      unique_dataset_2$flag[i] <- "Four of a Kind"
    }

  i <- i + 1
}

count(unique_dataset_2, flag)
```

```
##                     flag    n
## 1          Campuran Only   86
## 2       Exclude Campuran 7603
## 3             Exclude PT    3
## 4             Exclude PU   12
## 5          Exclude Saham   12
## 6         Four of a Kind  166
## 7          PT & Campuran   19
## 8          PU & Campuran   14
## 9                PU & PT  410
## 10       Pasar Uang Only 3525
## 11 Pendapatan Tetap Only  902
## 12      Saham & Campuran    8
## 13            Saham & PT  330
## 14            Saham & PU  422
## 15            Saham Only 1200
```

```
full_dataset <- left_join(unique_dataset_2, dataset_1, by = "user_id")
```

```
count(subset(full_dataset, flag == "Pasar Uang Only"), user_income_range)
```

```
##          user_income_range    n
## 1                < 10 Juta 1784
## 2            > Rp 1 Miliar    4
## 3 > Rp 100 Juta - 500 Juta  111
## 4  > Rp 50 Juta - 100 Juta  413
## 5 > Rp 500 Juta - 1 Miliar    7
## 6     Rp 10 Juta - 50 Juta 1206
```

```
count(subset(full_dataset, flag == "Pasar Uang Only"), user_age)
```

```
##    user_age   n
## 1        17  74
## 2        18 266
## 3        19 309
## 4        20 333
## 5        21 330
## 6        22 254
## 7        23 260
## 8        24 216
## 9        25 185
## 10       26 159
## 11       27 144
## 12       28 105
## 13       29 116
## 14       30 103
## 15       31  81
## 16       32  59
## 17       33  45
## 18       34  50
## 19       35  45
## 20       36  35
## 21       37  42
## 22       38  29
## 23       39  26
## 24       40  23
## 25       41  26
## 26       42  15
## 27       43  20
## 28       44  10
## 29       45  17
## 30       46  19
## 31       47   8
## 32       48  13
## 33       49  10
## 34       50   7
## 35       51  11
## 36       52  10
## 37       53  10
## 38       54  12
## 39       55   6
## 40       56   8
## 41       57   6
## 42       58   3
## 43       59   5
## 44       60   4
## 45       62   4
## 46       63   3
## 47       64   1
## 48       65   2
## 49       66   1
## 50       67   1
## 51       69   1
## 52       71   2
## 53       73   1
```

```
count(subset(full_dataset, flag == "Pendapatan Tetap Only"), user_income_range)
```

```
##          user_income_range   n
## 1                < 10 Juta 411
## 2            > Rp 1 Miliar   1
## 3 > Rp 100 Juta - 500 Juta  54
## 4  > Rp 50 Juta - 100 Juta 143
## 5 > Rp 500 Juta - 1 Miliar   4
## 6     Rp 10 Juta - 50 Juta 289
```

```
count(subset(full_dataset, flag == "Pendapatan Tetap Only"), user_age)
```

```
##    user_age  n
## 1        17 17
## 2        18 61
## 3        19 65
## 4        20 77
## 5        21 72
## 6        22 78
## 7        23 62
## 8        24 41
## 9        25 55
## 10       26 45
## 11       27 46
## 12       28 24
## 13       29 31
## 14       30 29
## 15       31 18
## 16       32 20
## 17       33 18
## 18       34 16
## 19       35 17
## 20       36  8
## 21       37  7
## 22       38 10
## 23       39  5
## 24       40  4
## 25       41  8
## 26       42  5
## 27       43  4
## 28       44  2
## 29       45  6
## 30       46  9
## 31       47  3
## 32       48  6
## 33       49  2
## 34       50  1
## 35       51  6
## 36       52  5
## 37       53  2
## 38       54  1
## 39       55  2
## 40       56  3
## 41       57  3
## 42       60  1
## 43       61  2
## 44       64  1
## 45       65  1
## 46       67  1
## 47       74  1
## 48       82  1
```

```
count(subset(full_dataset, flag == "Saham Only"), user_income_range)
```

```
##          user_income_range   n
## 1                < 10 Juta 555
## 2            > Rp 1 Miliar   3
## 3 > Rp 100 Juta - 500 Juta  82
## 4  > Rp 50 Juta - 100 Juta 164
## 5 > Rp 500 Juta - 1 Miliar  11
## 6     Rp 10 Juta - 50 Juta 385
```

```
count(subset(full_dataset, flag == "Saham Only"), user_age)
```

```
##    user_age   n
## 1        17  33
## 2        18  94
## 3        19  99
## 4        20  97
## 5        21 101
## 6        22  74
## 7        23  70
## 8        24  72
## 9        25  61
## 10       26  59
## 11       27  49
## 12       28  59
## 13       29  55
## 14       30  37
## 15       31  24
## 16       32  21
## 17       33  11
## 18       34  18
## 19       35  20
## 20       36  10
## 21       37  17
## 22       38   7
## 23       39  13
## 24       40  11
## 25       41  13
## 26       42  13
## 27       43   7
## 28       44  11
## 29       45   3
## 30       46   4
## 31       47   4
## 32       48   5
## 33       49   4
## 34       50   2
## 35       51   2
## 36       52   8
## 37       54   3
## 38       55   2
## 39       56   2
## 40       57   1
## 41       58   1
## 42       60   2
## 43       64   1
```

```
count(subset(full_dataset, flag == "Campuran Only"), user_income_range)
```

```
##          user_income_range  n
## 1                < 10 Juta 20
## 2            > Rp 1 Miliar  1
## 3 > Rp 100 Juta - 500 Juta 21
## 4  > Rp 50 Juta - 100 Juta 21
## 5 > Rp 500 Juta - 1 Miliar  4
## 6     Rp 10 Juta - 50 Juta 19
```

```
count(subset(full_dataset, flag == "Campuran Only"), user_age)
```

```
##    user_age n
## 1        18 2
## 2        19 3
## 3        20 1
## 4        21 6
## 5        23 1
## 6        24 5
## 7        25 4
## 8        26 5
## 9        27 1
## 10       28 7
## 11       29 4
## 12       30 2
## 13       31 5
## 14       32 1
## 15       33 4
## 16       34 2
## 17       35 3
## 18       36 3
## 19       37 2
## 20       38 3
## 21       39 1
## 22       40 2
## 23       41 2
## 24       42 1
## 25       43 2
## 26       44 1
## 27       45 1
## 28       46 1
## 29       47 2
## 30       48 1
## 31       49 2
## 32       50 1
## 33       52 1
## 34       53 1
## 35       56 1
## 36       57 1
## 37       76 1
```