Zadanie 1

Damian Kędzierski 260493

# Wprowadzenie

## Cel

* „sformułowanie problemów biznesowych i analitycznych dla dwóch przykładowych sytuacji biznesowych, łącznie ze wstępnym określeniem wskaźników realizacji celów biznesowych (KPI)”
* „wstępna **eksploracyjna analiza** i **przygotowanie danych**(EDA) do budowy wizualizacji w Excelu i Tableau oraz budowy modeli predykcyjnych w Tableau”

„Wykonanie zadania jest wstępem do opracowania założeń i projektów pulpitów menedżerskich - zestawów wizualizacji, przeznaczonych dla decydentów (zadanie 2), a następnie budowy pulpitów w Excelu z Power Pivot lub w Power BI (zadanie 3) oraz w Tableau (zadania 5-6)”

## Zbiory danych

1. Zbiór danych o klientach („klienci”) – Arkusz *„260493\_klienci.xlsx”*
2. Zbiór danych o lokatach („lokaty”) – Arkusz *„260493\_lokaty.xlsx”*

# Sformułowanie problemów biznesowych i problemów analitycznych

## Zbiór – klienci:

Ustalenie dla jakiego przedstawiciela, przeprowadzane rozmowy były najskuteczniejsze w określonym miesiącu.

WSKAŹNIKI:

* skuteczność jest wprost proporcjonalna do kwoty zakupu i odwrotnie proporcjonalna do czasu trwania rozmowy

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Średnia skuteczność rozmów przedstawicieli w określonym miesiącu** | | | | |
| **Etykiety wierszy** | **sierpień** | **wrzesień** | **październik** | **Średnia** |
| P01 | 3,99 | 6,09 | 4,78 | **4,95** |
| P02 | 2,68 | 3,39 | 2,80 | **2,97** |
| P03 | 3,51 | 2,71 | 2,58 | **2,91** |
| P04 | 2,74 | 4,79 | 2,20 | **3,10** |
| P05 | 5,06 | 3,49 | 3,92 | **4,15** |
| **Średnia** | **3,56** | **3,99** | **3,22** | **3,59** |

## Zbiór – lokaty:

Ustalenie do których grup klientów należy kierować akcje marketingowe

* Do jakiej grupy wiekowej jest kierowanych najwięcej akcji marketingowych?
* Która grupa wiekowa najchętniej zgadza się na założenie lokaty?

WSKAŹNIKI:

* Liczba lokat założonych w określonej grupie wiekowej
* Procent zawartych umów w określonej grupie wiekowej

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Liczba klientów reprezentujących określoną grupę wiekową, którzy zdecydowali się zapisać na lokatę terminową** | | | | |
| **Etykiety wierszy** | **no** | **yes** | **Suma końcowa** | **Procent yes** |
| **2008** | **26351** | **1339** | **27690** | **4,84%** |
| do 20 | 2 |  | 2 | 0,00% |
| 20-30 | 2835 | 163 | 2998 | 5,44% |
| **30-40** | **10852** | **591** | **11443** | 5,16% |
| 40-50 | 7513 | 334 | 7847 | 4,26% |
| 50-60 | 4986 | 241 | 5227 | 4,61% |
| 60-70 | 163 | 10 | 173 | 5,78% |
| **2009** | **9212** | **2228** | **11440** | **19,48%** |
| do 20 | 29 | 23 | 52 | 44,23% |
| 20-30 | 1628 | 494 | 2122 | 23,28% |
| **30-40** | **4016** | **779** | **4795** | 16,25% |
| 40-50 | 2071 | 347 | 2418 | 14,35% |
| 50-60 | 1057 | 321 | 1378 | 23,29% |
| 60-70 | 240 | 142 | 382 | 37,17% |
| 70-80 | 120 | 85 | 205 | 41,46% |
| 80-90 | 47 | 35 | 82 | 42,68% |
| ponad 90 | 4 | 2 | 6 | 33,33% |
| **2010** | **985** | **1073** | **2058** | **52,14%** |
| do 20 | 10 | 11 | 21 | 52,38% |
| 20-30 | 243 | 231 | 474 | 48,73% |
| **30-40** | **355** | **345** | **700** | 49,29% |
| 40-50 | 108 | 153 | 261 | 58,62% |
| 50-60 | 122 | 135 | 257 | 52,53% |
| 60-70 | 70 | 99 | 169 | 58,58% |
| 70-80 | 55 | 59 | 114 | 51,75% |
| 80-90 | 21 | 37 | 58 | 63,79% |
| ponad 90 | 1 | 3 | 4 | 75,00% |
| **Suma końcowa** | **36548** | **4640** | **41188** | **11,27%** |

# Wstępna analiza i przygotowanie danych – dane klienci

Zbiór posiada 1000 rekordów.

Fragment tabeli danych klientów

Obraz zawierający stół

Opis wygenerowany automatycznie

Tabela danych przedstawicieli

Obraz zawierający stół

Opis wygenerowany automatycznie

Struktura, rozmiar, typy zmiennych zbioru klientów

Obraz zawierający tekst

Opis wygenerowany automatycznie

Dane statystyczne (istotnych) zmiennych ilościowych

Obraz zawierający tekst

Opis wygenerowany automatycznie

Ilość wystąpień danego przedstawiciela w zbiorze (zmienna kategoryczna)

Obraz zawierający tekst

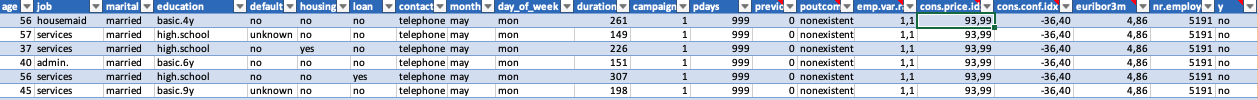
Opis wygenerowany automatycznie

Po wstępnej analizie można zaobserwować, że głównie, że najwięcej rozmów odbył przedstawiciel Krzysztof Bąk (P02), natomiast najlepszą skuteczność posiada przedstawiciel Jan Kowalski (P01). Ponadto najdłuższy czas rozmowy to nie całe 3 minuty, a najkrótszy to 1 sekunda, przy czym średnia długość to ok. 1.5 minuty. Najwyższa kwota zakupu wynosiła 225, najmniejsza 20, a średnia 124. Najwyższa uzyskana skuteczność to 158, najmniejsza 0.13, a średnia 3.59. Zbiór nie posiada braków danych.

# Wstępna analiza i przygotowanie danych – dane lokaty

Zbiór posiada 41 188 rekordów.

Fragment tabeli danych lokaty



Obraz zawierający stół

Opis wygenerowany automatycznie

Struktura, rozmiar, typy zmiennych zbioru lokaty

# Obraz zawierający tekst Opis wygenerowany automatycznie

Dane statystyczne (istotnych) zmiennych ilościowych

Obraz zawierający stół

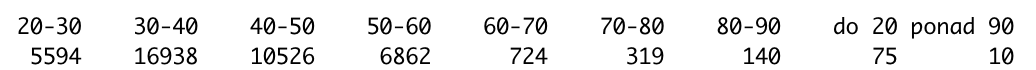
Opis wygenerowany automatycznie

Liczba klientów, którzy zdecydowali się zapisać na lokatę terminową (zmienna kategorialna)

Obraz zawierający tekst

Opis wygenerowany automatycznie

Podział grup wiekowych na określone przedziały (zmienna kategorialna)



Po wstępnej analizie można zauważyć, że z 41 188 klientów, tylko 4640 z nich zdecydowało się zapisać na lokatę terminową. Najmłodszy potencjalny klient miał 17 lat, a najstarszy 98. Natomiast największą grupę klientów docelowych stanowiła grupa od 30 do 40 lat.