

Poznej svého zákazníka v Power BI

Ing. Lucie Šperková Ph.D.

BI specialista

lucie.sperkova@prg.aero

 www.linkedin.com/in/luciesperkova/

O mně



- BI specialista v Rozvoji technologií na Letišti Praha
- Lektor Power BI v počítačové škole Gopas
- BI konzultant (retail/e-commerce, digitální agentury)
- Akademický pracovník VŠE (zákaznická analytika)

Obsah přednášky

- **Jaké pohledy lze na základě transakčních dat dělat**
- **Jak připravit data (potřebné hodnoty)**
- **Jak segmentovat dle RFM**

Zákazník?

- Kdo je aktivní zákazník?
- Kdo je věrný zákazník?
- Kdy je považován za ztraceného?
- Jak zákazníka identifikovat v B2B a B2C?
- Kde se nachází v rámci zákaznického životního cyklu?

AKVIZICE -> AKTIVACE -> RETENCE

Je pro váš byznys důležitější získávání nových zákazníků nebo udržení těch stávajících?

Nutné minimum – kdo, kdy, za kolik

- Transakční data
 - datum
 - označení zákazníka
 - označení objednávky
 - částka
- Možné rozšíření: co a jak
 - zakoupené produkty
 - způsoby platby
 - pobočky
 - ...

Vytvoření datumové tabulky

- Pokud nevyužíváte Time Intelligence, potřeba vytvořit 2 datumové dimenzionální tabulky Date a Min_Date pro datum prvního nákupu kvůli kohortám, kde datумы stojí proti sobě

```
date =  
ADDCOLUMNS (  
    CALENDAR (  
        MIN ( rfa[date] ),  
        MAX ( rfa[date] )  
    ),  
    "Year", YEAR ( [Date] ),  
    "Month", FORMAT ([Date], "mmmm"),  
    "Month Number", MONTH ( [Date] ),  
    "Quarter", QUARTER ( [Date] )  
)
```

```
min_date =  
ADDCOLUMNS (  
    CALENDAR (  
        MIN ( customer[min_date] ),  
        MAX ( customer[min_date] )  
    ),  
    "Year", YEAR ( [Date] ),  
    "Month", FORMAT ([Date], "mmmm"),  
    "Month Number", MONTH ( [Date] ),  
    "Quarter", QUARTER ( [Date] )  
)
```

Zákaznický cyklus

<https://www.daxpatterns.com/new-and-returning-customers/>

- **Nový zákazník**

- nakoupil v daném období poprvé

- **Vracející se zákazník**

- již v minulosti nakoupili

- **Ztracený zákazník**

- nenakoupil už určité období (např. poslední rok)

- **Reaktivovaný zákazník**

- byl považován za ztraceného, ale znovu nakoupil v dalším období



Počet nových zákazníků

```
New customers =  
COUNTROWS (  
    FILTER (  
        ADDCOLUMNS (  
            VALUES ( rfa[customer_id] ),  
            "PreviousOrders", CALCULATE (  
                COUNTROWS ( rfa ),  
                FILTER ( ALL ( 'date' ), 'date'[Date] < MIN ( 'date'[Date] ) ) )  
        ),  
        [PreviousOrders] = 0  
    )  
)
```


Počet vracejících se zákazníků

Returning customers =

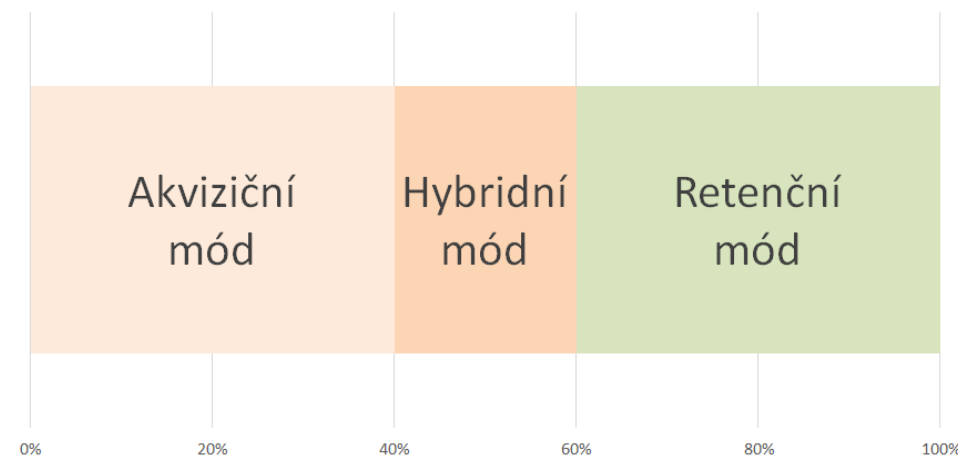
```
COUNTROWS (
    CALCULATETABLE (
        VALUES ( rfa[customer_id] ),
        VALUES ( rfa[customer_id] ),
        FILTER (
            ALL ( 'date' ),
            'date'[Date] < MIN ( 'date'[Date] )
        )
    )
)
```

■ MRR: Monthly Repurchase Rate

- Kolik % zákazníků z tohoto měsíce nakoupilo i měsíc minulý

■ ARR: Anual Repurchase Rate

- Kolik % letošních zákazníků nakoupilo i loni



Počet zákazníků, kteří nakoupili v dalším období

Rok	2017	2018	2019	2020
⊕ 2017	8506	1134	767	387
⊕ 2018		11804	1153	434
⊕ 2019			10711	621
⊕ 2020				7494

% zákazníků, kteří nakoupili v dalším období

Rok	2017	2018	2019	2020
⊕ 2017	100,00%	13,33%	9,02%	4,55%
⊕ 2018		100,00%	9,77%	3,68%
⊕ 2019			100,00%	5,80%
⊕ 2020				100,00%

% zákazníků, kteří nakoupili v předchozím období

Rok	2017	2018	2019	2020
⊕ 2017	100,00%	8,76%	6,07%	4,33%
⊕ 2018		91,24%	9,13%	4,86%
⊕ 2019			84,80%	6,95%
⊕ 2020				83,86%

Vracející se zákazníci z minulého roku

Returning customers LY =

```
COUNTROWS (
    CALCULATETABLE (
        VALUES ( rfa[customer_id] ),
        VALUES ( rfa[customer_id] ),
        FILTER (
            ALL ( 'date' ),
            'date'[Date] < MIN ( 'date'[Date] )
            && 'Date'[year]
                = SELECTEDVALUE ( 'Date'[year] ) - 1
        )
    )
)
```

AAR =

```
DIVIDE ( [Returning customers LY], [Customers] )
```

Po jak dlouhé době zákazník nakupuje?

date	customer_id	order_id	price
05.06.2019	50882	SO64610	1519
05.11.2018	51051	SO36080	1519
02.04.2019	6371	SO57786	1519
11.01.2019	54390	SO49078	1519
25.01.2019	37694	SO50800	1519
14.09.2018	33504	SO34212	1519
22.03.2018	45347	SO22973	1519

previous_date	transaction_order	difference
27.09.2018	2	251
18.10.2018	3	18
30.12.2018	4	93
	1	
04.01.2019	9	21
	1	
08.03.2018	2	14



dopočítávané sloupce

Pořadí objednávky

```
transaction_order =  
(  
    RANKX (  
        FILTER ( rfa, rfa[customer_id] = EARLIER ( rfa[customer_id] ) ),  
        rfa[date],  
        ,  
        ASC,  
        SKIP  
    )  
)
```

Doba od předchozí objednávky

```
difference =  
VAR previous_date =  
    VAR currentDate = rfa[date]  
    VAR currentCustomer = rfa[customer_id]  
    RETURN  
        CALCULATE (  
            MAX ( rfa[date] ),  
            FILTER (  
                ALL ( rfa ),  
                rfa[date] < currentDate  
                    && rfa[customer_id] = currentCustomer  
            )  
        )  
    )  
RETURN  
    DATEDIFF ( previous_date, rfa[date], DAY )
```

Podíl objednávek dle časové prodlevy

Za jak dlouho udělal zákazník opakovanou objednávku (% uživatelů, objednávek, profitu)

Doba_od_předchozí_objednávky	Počet uživatelů	% uživatelů	Počet objednávek	% objednávek	Profit	% profitu
b) stejný den	3054	33,41%	3810	29,26%	8 459 048 Kč	30,84%
c) den - 1 týden	881	9,64%	949	7,29%	1 855 255 Kč	6,76%
d) 1 týden - 1 měsíc	820	8,97%	913	7,01%	1 902 813 Kč	6,94%
e) 1 měsíc - 3 měsíce	819	8,96%	891	6,84%	1 886 889 Kč	6,88%
f) 3 měsíce - půl roku	888	9,71%	947	7,27%	1 968 915 Kč	7,18%
g) půl rok - 1 rok	2430	26,58%	2636	20,24%	4 506 235 Kč	16,43%
h) 1 rok - 2 roky	1523	16,66%	1607	12,34%	3 710 326 Kč	13,53%
i) více jak 2 roky	1265	13,84%	1269	9,75%	3 135 226 Kč	11,43%
Celkem	9141	100,00%	13022	100,00%	27 424 705 Kč	100,00%

Základ RF analýzy časových prodlev mezi objednávkami

Počty objednávek: Doba od předchozí objednávky / O kolikátou objednávku jde

Doba od předchozí objednávky	b) druhá	c) třetí	d) 4. - 6.	e) 7. - 10.	f) 11 a více	Celkem
b) stejný den	2432	785	477	72	44	3810
c) den - 1 týden	604	189	106	23	27	949
d) 1 týden - 1 měsíc	550	165	128	27	43	913
e) 1 měsíc - 3 měsíce	535	186	116	21	33	891
f) 3 měsíce - půl roku	576	196	137	23	15	947
g) půl rok - 1 rok	1743	531	316	32	14	2636
h) 1 rok - 2 roky	1098	304	177	22	6	1607
i) více jak 2 roky	968	209	87	4	1	1269
Celkem	8506	2565	1544	224	183	13022

Rychlé kategorie pomocí IF

```
poradi_objednavky =
IF (
  rfa[transaction_order] = 1,
  "a) první",
  IF (
    rfa[transaction_order] = 2,
    "b) druhá",
    IF (
      rfa[transaction_order] = 3,
      "c) třetí",
      IF (
        rfa[transaction_order] > 3
          && rfa[transaction_order] <= 6,
        "d) 4. - 6.",
        IF (
          rfa[transaction_order] > 6
            && rfa[transaction_order] <= 10,
          "e) 7. - 10.",
          IF ( rfa[transaction_order] > 10, "f) 11 a více" )
        ))))
))))
```

Zakategorizování pořadí objednávky jako počítaný sloupec

```
poradi_objednavky =  
CALCULATE (  
    VALUES ( c_poradi[poradi] ),  
    FILTER (  
        c_poradi,  
        AND (  
            c_poradi[from] <= rfa[transaction_order],  
            c_poradi[to] >= rfa[transaction_order]  
        )  
    )  
)  
)
```

Pořadí objednávky přes ID do číselníku

```
transaction_order_id =  
CALCULATE (  
    DISTINCT ( c_poradi[id] ),  
    FILTER (  
        c_poradi,  
        AND (  
            c_poradi[from] <= rfa[transaction_order],  
            c_poradi[to] >= rfa[transaction_order]  
        )  
    )  
)  
)
```

Kohorty

- Počty a podíly objednávek
 - Počty a podíly zákazníků
 - Celkové příjmy a podíly příjmů
 - Průměrné příjmy
-
- dle časové prodlevy mezi objednávkami
 - dle časové prodlevy od první objednávky
 - dle časové prodlevy od poslední objednávky
- VS**
- o kolikátou objednávku jde

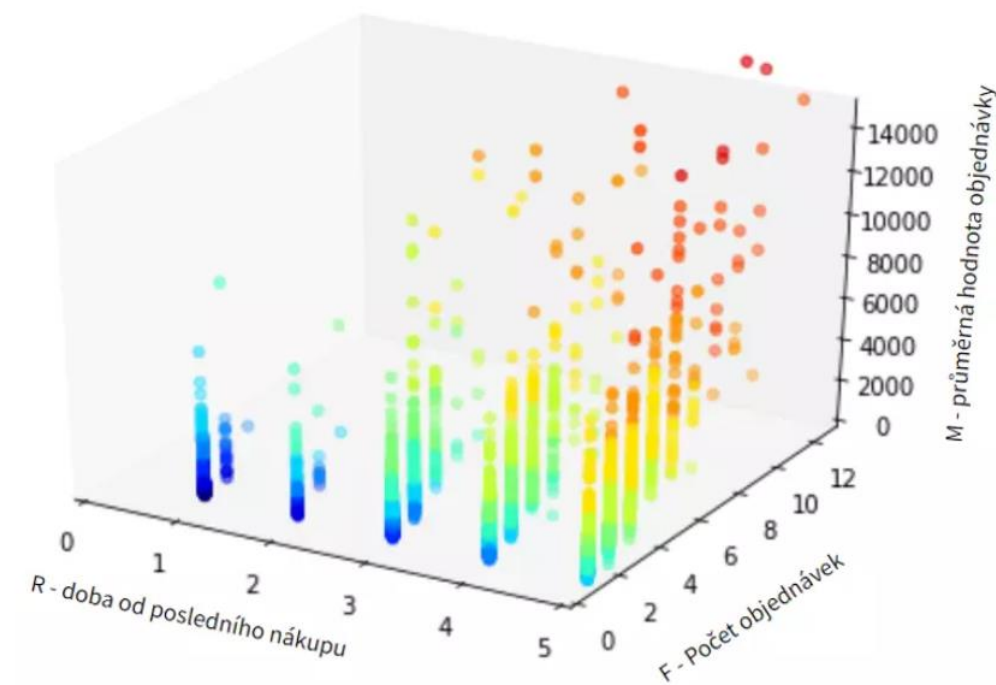
RFM

3 faktory rozlišující hodnotu zákazníka:

- **Recency**
 - ♦ **Kdy naposledy** zákazník nakoupil? Jsme schopni identifikovat aktivní a ztracené zákazníky?
 - ♦ *Rozdíl mezi datem posledního nákupu a datem analýzy*
- **Frequency**
 - ♦ **Jak často** zákazník nakupuje? Jsme schopni identifikovat ty, co nakupují často od jednorázových zákazníků?
- **Monetary**
 - ♦ **Kolik** zákazník **utratil**?
 - ♦ *Suma nebo průměr částek z objednávek*

Analýza RFM

- Cílem je rozdělit zákazníky s podobným nákupním chováním do několika segmentů a dle toho na ně cílit
- Zvýšit loajalitu, konverzní poměr, zisky
- RFM ukazatele rozdělíme na intervaly, do kterých zařadíme zákazníky



Potřebné hodnoty

- Datum
- Id objednávky
- Id zákazníka
- Částka

Dopočítávané údaje

- Pořadí objednávky
- Doba od předchozí objednávky
- Datum první objednávky
- Doba od poslední objednávky (R)
- Celkový počet objednávek na zákazníka (F)
- Suma hodnot objednávek (M)

Kde napočítat

- v SQL
- v Power Query
- v DAX

Jak napočítat

- staticky
- dynamicky

RFM hodnoty na úrovni zákazníka

```
recency =  
VAR max_date =  
    LASTDATE ( rfa[date] )  
RETURN  
    DATEDIFF ( max_date, TODAY () - 1, DAY )
```

```
frequency =  
COUNTROWS ( RELATEDTABLE ( rfa ) )
```

```
monetary =  
SUMX ( RELATEDTABLE ( rfa ), rfa[price] )
```


Tradiční skóringový pohled na RFM

- Každý pohled si rozdělíte na 5 rovnoměrných částí
- Zákazníkům dáváte skóre 1-5 pro R, 1-5 pro F, 1-5 pro M
- $524 = R_5F_2M_4$ = nakoupil nedávno, už párkrát nakupoval za velké peníze
- Do kampaně si vybíráte zákazníky jen dle určitých úrovní RFM skóre
- Někdo skóre váží a sčítá $524 = 11$

Recency	HODNOTA
1	5
30	4
180	3
360	2
450	1

Frequency	HODNOTA
0	1
1.21	2
1.77	3
2	4
2.35	5

Monetary	HODNOTA
0	1
2400	2
3226	3
7400	4
20000	5

Monetary skóre jako počítaný sloupec v tabulce zákazník

```
monetary_score =  
CALCULATE (  
    VALUES ( c_monetary[score] ),  
    FILTER (  
        c_monetary,  
        AND (  
            c_monetary[from] <= customer[monetary],  
            c_monetary[to] >= customer[monetary]  
        )  
    )  
)
```

RFM segmenty

SEGMENT

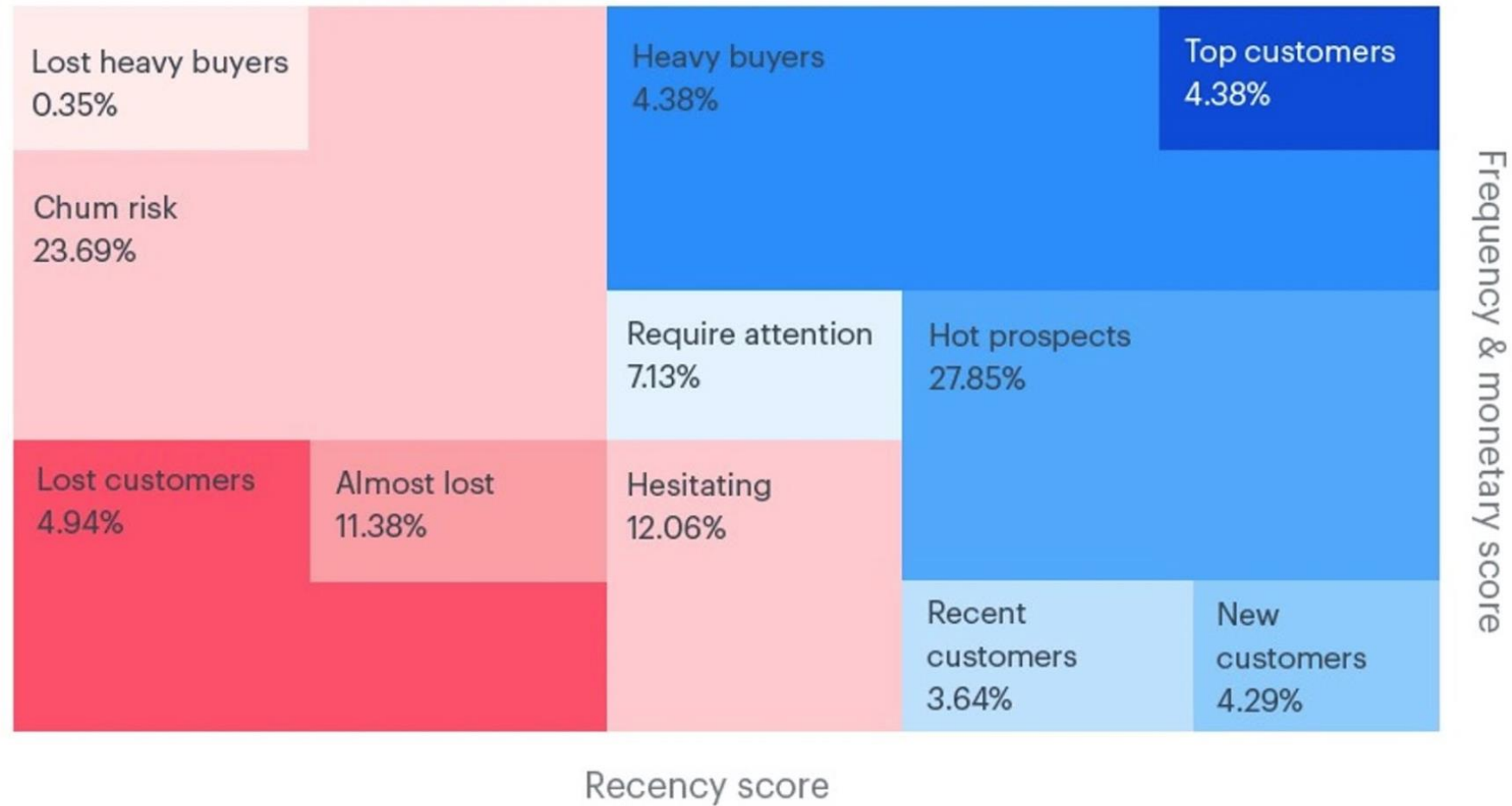
VIP
Loajální zákazníci
Potenciální loajální
Noví zákazníci
Slibní
Zákazníci, kteří potřebují pozornost
Spící
V riziku
Neztraťte je!
Hibernace
Ztracení
K analýze
K navolání

AKTIVITA

Nakoupili v poslední době, nakupují často a utrácejí nejvíce.
Utrácejí často slušné peníze. Reagují na kampaně.
Nedávni zákazníci, co utratili slušné peníze a nakoupili více než jednou.
Nakoupili nedávno, ale nekupují často.
Nejnovější zákazníci, kteří moc neutratili.
Nadprůměrná aktuálnost, frekvence i peněžní hodnota, ale nakoupili již před delší dobou.
Pod průměrnou aktuálností, frekvencí i peněžní hodnotou. Ztratíme je, pokud nebudou znovu aktivováni.
Utratili hodně peněz a nakupovali často. Potřebujeme je vrátit zpět!
Nakupovali často a měli největší útraty, ale dlouho se nevrátili.
Poslední nákup byl dávno v minulosti, nakupovali za málo a zřídka.
Nejmenší aktuálnost, frekvence a peněžní hodnota.
prověření objednávek

rfm_score	segment
555	VIP
554	VIP
553	Loajální zákazníci
552	Loajální zákazníci
551	K analýze
545	VIP
544	VIP
543	Loajální zákazníci
542	Slibní
541	Slibní

RFM ANALYSIS



Děkuji za pozornost

Ing. Lucie Šperková Ph.D.

BI specialista

lucie.sperkova@prg.aero

 www.linkedin.com/in/luciesperkova/