

Úspešnosť telemarketingu v závislosti od vzdelania a mesiaca oslovenia

V tejto práci využijeme logistickú regresiu. Ide o štatistickú metódu, ktorá odhaduje pravdepodobnosť javu, kde závislá premenná je dichotomická. Na spracovanie údajov sme využili štatistický program Jasp 0.16.3.

Na vykonanie logistickej regresie sme pracovali s voľne-dostupným Bank Marketing Data datasetom, v ktorom boli údaje získané marketingovou divíziou portugalskej banky, ktorá svojou kampaňou chcela získať nových odberateľov k svojmu novému bankovému programu. Dataset disponuje širokou škálou údajov, ktoré boli získané počas procesu v roku 2014. Okrem základných údajov ako vek, zamestnanie či vzdelanie, obsahuje súbor aj dáta či daná osoba pôžičku alebo v ktorom mesiaci a dni bol človek kontaktovaný a koľkokrát.

Na základe vyššie uvedených údajov, ktoré patrili 41 188 participantom, sme sa rozhodli zamerať na 2 oblasti, a to vzdelanie a mesiac v roku. Zaujímalo nás, či tieto faktory nejako ovplyvnia rozhodnutie osloveného. Naša vzorka neobsahovala údaje o pohlaví, no vekový priemer bol 40 rokov (Min: 17r, Max: 98r). Vzdelanie bolo kategorizované do 8 oblastí, napríklad základná školská dochádzka 4,6 či 9 rokov či univerzitný titul. V rámci mesiacov neboli zastúpené prvé dva mesiace v roku.

Závislá premenná: Označená ako áno/ nie, ukazovala či sa oslovený človek nakoniec stal odberateľom (áno= stal sa, nie= nestal)

Kovariát: vek osoby

Faktory: Vzdelanie, mesiac v roku

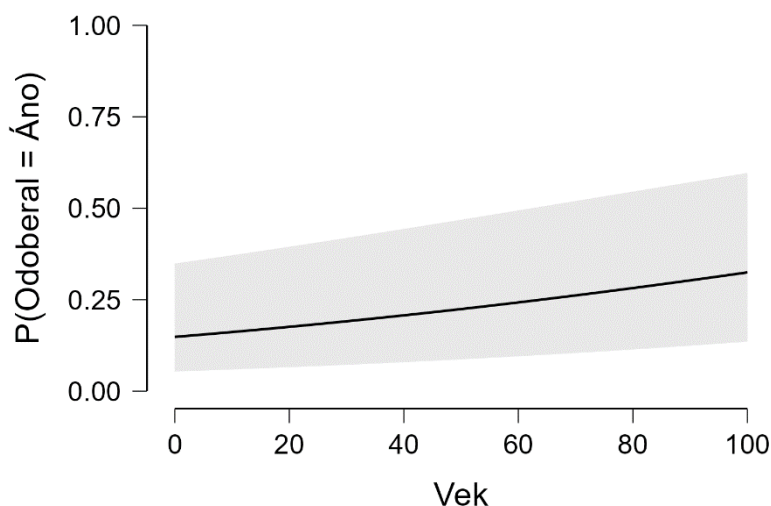
Tabuľka 1
Vzdelanie opýtaných

	Odhad	Štandard. Chyba	Pravdepodobnosť	p
Východzia hodnota*	-1.749	0.572	0.174	.002
Vek	0.010	0.002	1.010	< .001
Neuvedené	-0.473	0.572	0.623	.408
SŠ	-0.749	0.569	0.473	.188
Výučný list	-0.718	0.569	0.488	.207
VŠ	-0.488	0.568	0.614	.390
ZŠ 4	-0.911	0.570	0.402	.110
ZŠ 6	-1.081	0.573	0.339	.059
ZŠ 9	-1.119	0.570	0.327	.050

*Východzia hodnota uvádza kategóriu bez vzdelania ako normu.

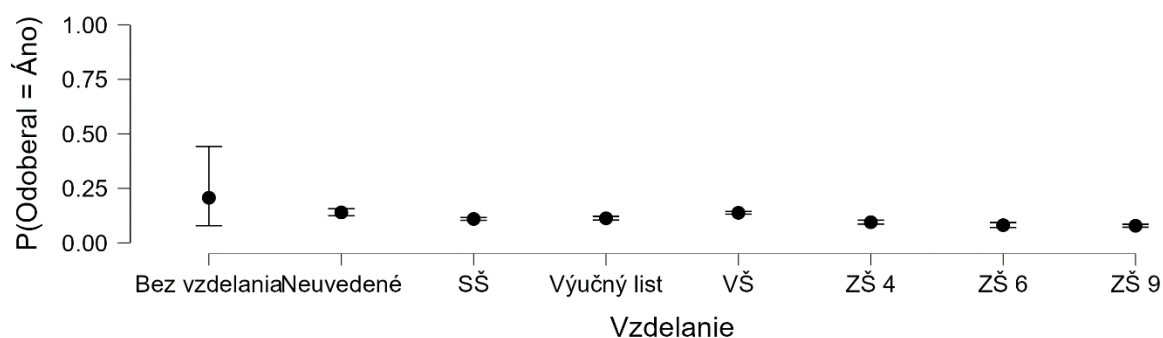
Kategória veku nám ukazuje, že čím je človek starší, tým má väčšiu pravdepodobnosť stať sa odberateľom. P je menšie ako 0,001 čo naznačuje, že ide o dôležitú časť modelu, nakoľko zmeny vo veku súvisia so zmenami v odpovedi. Z prieskumu sa ukazuje, banka mala najväčšiu šancu uspieť pri marketingovej kampani u ľudí bez uvedeného vzdelania alebo s vysokoškolským titulom. Naopak najnižšia šanca sa ukázala u klientov so základnou školou. Ani pri jednom údaji nie je p signifikantné (ZŠ 9 je diskutabilné), a tak nemôžeme s určitosťou tvrdiť, že vzdelanie je významný prediktor. Program Jasp nám vypočítal, že tento model má na 88.7% presnú predikciu správania.

Obrázok 1
Vek opýtaných



Obrázok 2

Vzdelanie opýtaných



V rámci logistickej regresie nás tiež zaujímalo, či mesiac, v ktorom je človek oslovený má vplyv na jeho rozhodnutie. Výsledky možno vidieť v tabuľke 2.

Tabuľka 2

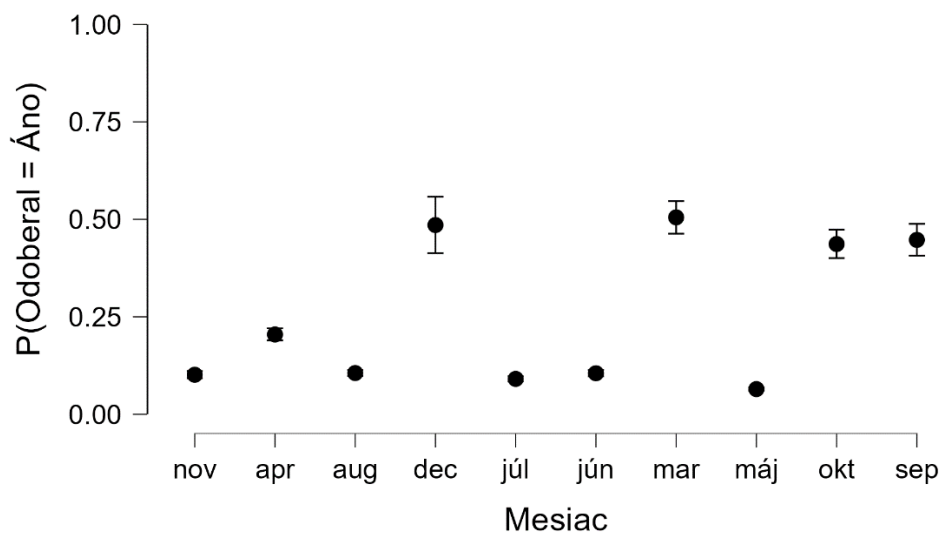
Mesiac v roku opýtania

	Odhad	Štandardná chyba	Pravdepodobnosť	p
Východzia hodnota*	-2.263	0.079	0.104	< .001
Vek	0.002	0.001	1.002	.168
Marec	2.204	0.100	9.060	< .001
Apríl	0.825	0.071	2.283	< .001
Máj	-0.492	0.062	0.612	< .001
Jún	0.042	0.068	1.043	.538
Júl	-0.123	0.066	0.884	.063
August	0.048	0.066	1.049	.469
September	1.972	0.099	7.184	< .001
Október	1.929	0.091	6.880	< .001
December	2.125	0.157	8.372	< .001

*Východzia hodnota uvádza mesiac november ako normu

Obrázok 3

Mesiac v roku opýtania



Logistická regresia nám ukázala, že mesiace v ktorých pravdepodobnosť, že sa klient stane odberateľom je najvyššia, sú marec, september, október a december. Ich pravdepodobnosť je vysoká, podložená rovnako významným p . Štatistický softvér nám vypočítal 88,7% šancu presnej predikcie. Tieto dáta boli zbierané na limitovanej vzorke portugalských občanov, a preto výsledky nemožno generalizovať.

Zdroje

Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22-31.

Dáta získané z: <https://archive.ics.uci.edu/ml/datasets/bank+marketing#>