

Univerza v Ljubljani - Ekonomska fakulteta
Predmet: Poslovno obveščanje in analitika
Nosilca predmeta: **red. prof. dr. Jurij Jaklič**

Študijski program : **(IZREDNI)**
Študijsko leto: 2022/2023



Seminarska naloga

pri predmetu
POSLOVNO OBVEŠČANJE IN ANALITIKA

»Identifikacija navad strank v namen povečanja ponovnih nakupov strank«

Študent: Rok Zupan

Vpisna številka: 19610023

IZREDNI

Ljubljana, 22.03.2023

Kazalo

Uvod	3
Poslovno upravičenje.....	3
Zahteve	3
Analiza podatkov in metapodatkov	3
Ugotovitve in predlogi.....	8
Viri in literatura	9

Kazalo slik

Slika 1: Primerjava strank po segmentu stranke.....	5
Slika 2: Primerjava strank po načinu pošiljanja izdelka.....	6
Slika 3: Primerjava vsote prodaje skozi mesece.....	6
Slika 4: Primerjava povprečne cene izdelka skozi mesece.....	7
Slika 5: Primerjava sub-kategorij	7
Slika 6: Popularni izdelki za vračajoče stranke	8

Uvod

V projektnem poročilu "Identifikacija navad strank v namen povečanja ponovnih nakupov strank" bom uporabil podatkovni set o prodaji, ki vsebuje informacije o nakupih strank v veliki globalni trgovini. Cilj je razumeti navade strank pri nakupovanju in prepoznati vzorce, ki bi mi lahko pomagali pri povečanju ponovnih nakupov strank. Naša analiza bo vključevala uporabo različnih statističnih tehnik in raziskovalne analize, da bi bolje razumeli podatke in iz njih pridobili ugotovitve. Poročilo bo vsebovalo opis problema, poslovno upravičenje, analizo podatkov in predstavitev ugotovitev.

Poslovno upravičenje

Povečanje ponovnih nakupov strank ima lahko velik vpliv na poslovanje podjetja, saj lahko zmanjša stroške trženja in promocije, hkrati pa poveča prihodke. Zato je smiselno analizirati podatke o nakupih strank in prepoznati vzorce, ki nam lahko pomagajo pri načrtovanju marketinških strategij in programov zvestobe strankam. Analiza podatkov o nakupih strank nam lahko pomaga pri načrtovanju in izvajanju bolj usmerjenih kampanj, s katerimi bi povečali ponovne nakupe strank.

Zahteve

Za doseg ciljev projekta bomo potrebovali podatkovno bazo o prodaji, ki vsebuje informacije o nakupih strank v trgovinah. Te informacije bomo uporabili za identifikacijo navad strank pri nakupovanju, analizo vsote nakupov na mesečni ravni, pregled najbolj prodajanih izdelkov in filtriranje naročil glede na različne attribute kot so kategorije, podkategorije, zvezne države in vrste odpošiljanja. Uporabil bom vizualizacijsko orodje Power BI.

Analiza podatkov in metapodatkov

Podatkovni set vsebuje informacije o nakupih strank v trgovinah v Združenih Državah Amerike. Ima 9800 vrstic, ki ponazarjajo število prodanih artiklov v razmaku 4 let, med 3. januarjem 2015 in 30. decembrom 2018.

Atributi podatkovnega seta so:

- Row ID – številka vrstice, ki jo lahko uporabimo kot primarni ključ, saj unikatno določuje vsako vrstico podatkovnega seta.
- Order ID – številka naročila. V setu je 4922 naročil, kar pomeni, da ima lahko eno naročilo več številke vrstice oziroma več artiklov naročenih.
- Order Date – datum naročila. V setu je 1230 unikatnih vrednosti oziroma dni v katerih je bilo opravljeno vsaj eno naročilo.

- Ship Date – datum pošiljanja naročila. V setu je 1326 unikatnih vrednosti oziroma dni, v katerih je bilo opravljeno vsaj eno pošiljanje naročila. Na podlagi razlike med številom dni v katerih je bilo opravljeno naročilo ali poslano lahko ugotovimo, da je bilo skladišče oziroma prostor za odpremljanje odprt več časa kot dejanske trgovine. Vseh dni v razmaku med prvim in zadnjim datumom naročila je 1457. Vseh dni v razmaku med prvim in zadnjim datumom pošiljanja pa je 1459.
- Ship Mode – način pošiljanja. Naročilo je lahko poslano v 4 opcijah: First Class, Second Class, Standard Class, Same Day.
- Customer ID – identifikacijska številka stranke, unikatnih vrednosti 793. Vsaka stranka ima svojo identifikacijsko številko.
- Customer Name – ime in priimek stranke, unikatnih vrednosti 793.
- Segment – stranka je lahko v 3 segmentih: Consumer (potrošnik), Home Office (domača pisarna), Corporate (korporacija).
- Country – država je samo ena Združene Države Amerike.
- City – mesto, število mest je 529.
- State – zvezna država, imamo 49 zveznih držav.
- Postal Code – število poštних kod, 627 različnih vrednosti.
- Region – štiri različne regije (Central, South, West, East)
- Product ID – število produktnih kod. Vsak izdelek ima svojo kodo. 1861 unikatnih vrednosti.
- Category – število kategorij 3: Technology, Office Supplies and Furniture.
- Sub-Category - 17 podkategorij izdelkov.
- Product Name – ime produkta. Število unikatnih ime produktov je 1849.
- Sales – vrednost vrstice oziroma vrednost artikla.

Metadata podatkov: podatki se po napovedi avtorja Rohit Sahoo (Owner) nikoli ne bodo posodobili.

Razvoj rešitve

Zajem podatkov (Extract)

Podatke sem zajel na spletni platformi Kaggle. Datoteka je bila CSV (Comma Separated Values). To je pogosta oblika pri prenašanju podatkovnih setov. Ob nalaganju v Power BI sem naletel na težavo, saj je bil atribut Sales, ki je vseboval vrednost prodanega izdelka ločen z decimalno piko. Ob nalaganju v Power BI je odstranil piko in naložil celotno vrednost. To pomeni, da vrednost \$35.23 naložil kot 3523. To zadevo bi lahko odpravil v Power BI, če bi bilo število decimalnih mest za piko konstantno.

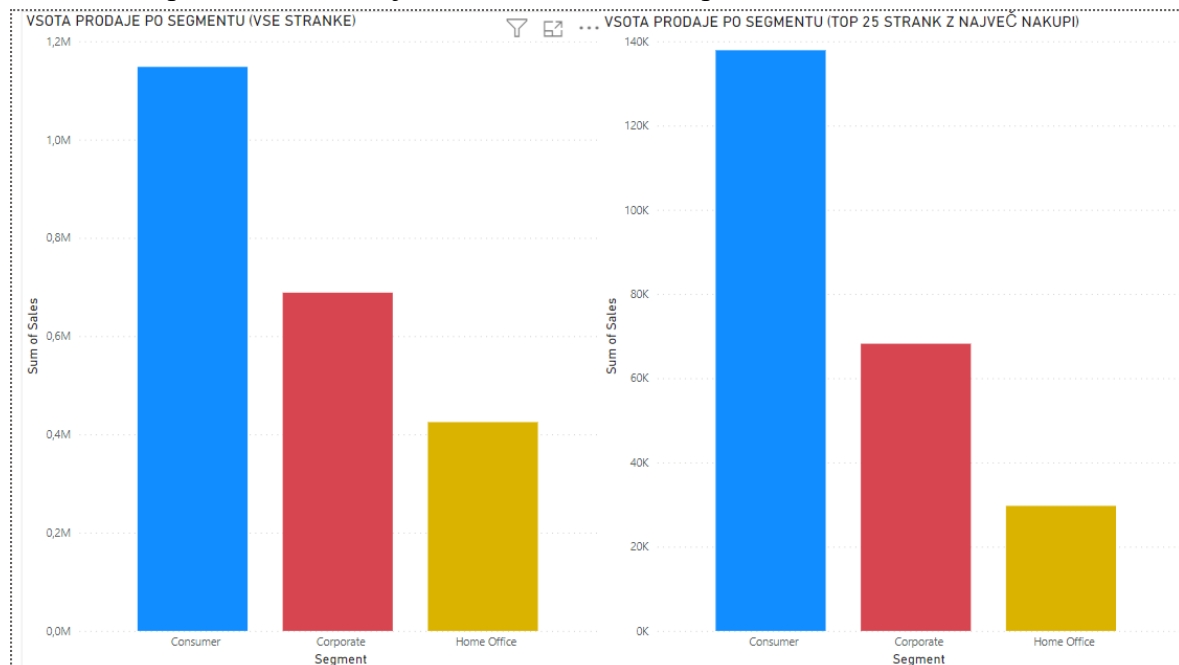
Priprava podatkov (Transform & Load)

Številko decimalnih mest za piko ni bilo konstantno: nekatera beleženja prodanih artiklov so imeli vrednost npr \$2341.4512 kar je pomenilo, da ob nalaganju CSV datoteke v Power BI se je vrednost zabeležila kot 23414512. Težavo sem rešil tako, da sem pretvoril datoteko iz

CSV datoteke v Excel datoteko in jo nato naložil v Power BI. Tako sem rešil težavo, nato sem »postrigel« vsa decimalna mesta na enako dolžino (2). Po transformaciji podatkov sem se spoznal s podatkovnimi tipi atributov in ugotavljal unikatne vrednosti datoteke.

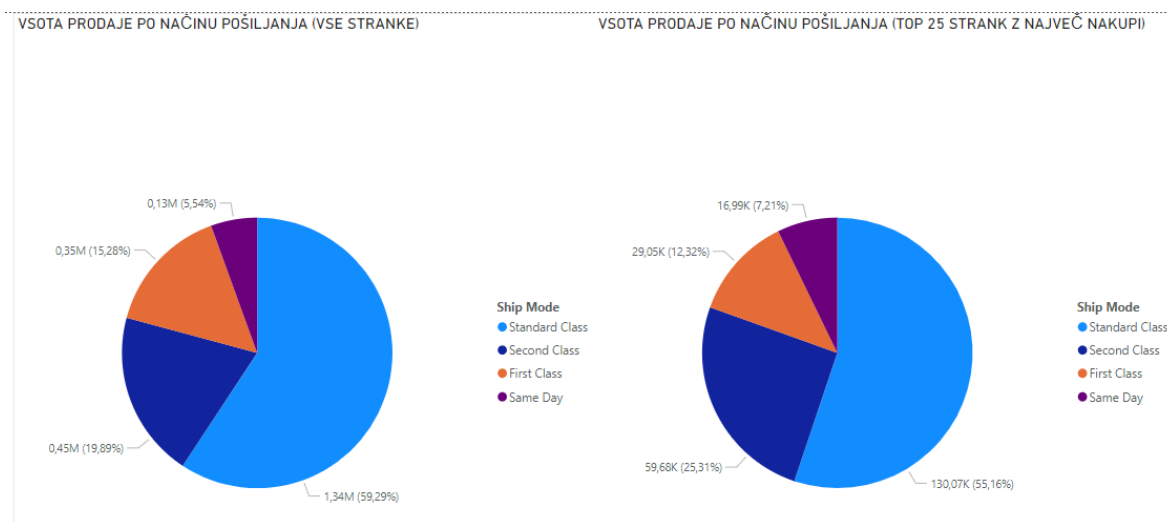
Rešitev

Namen podatkovne analize je bilo identificirati navade strank v namen povečanja ponovnih nakupov strank. Analiziral sem razlike v prodaji za vse stranke in za 25 strank z največ ponovnimi nakupi. Začel sem z primerjavo po kategorijah, kjer nisem ugotovil večjih razlik. Pri vseh primerjav sem uporabil tako celotno vsoto prodaje po kategorijah, kot povprečni vrednost prodaje na posamezen artikel. Prve omembe vredno razliko med vsemi strankami in najbolj pogostejše vračajočimi strankami sem opazil pri razliki v deležu vsote prodaje razdeljene po segmentu oziroma tipu stranke. Stranka se je lahko identificirala kot Consumer, Home Office ali Corporate. Najbolj vračajoče stranke so ime višji delež Consumer tipa stranke in manjši delež Home Office tipa stranke.



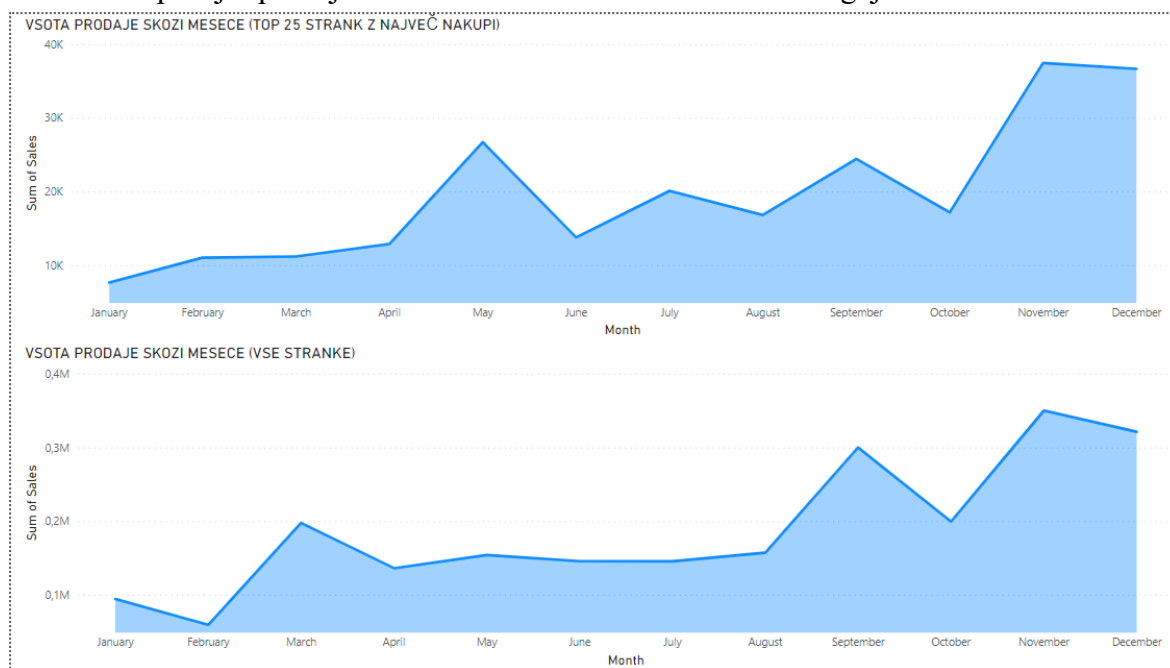
Slika 1: Primerjava strank po segmentu stranke

Naslednjo večjo razliko med najbolj vračajočimi strankami in vsemi strankami sem opazil v tipu pošiljanja izdelka. Vračajoče stranke so bolj posegale po posebnih vrstah pošiljanja kot vse stranke. Razlika ni velika, lahko pa poda izhodišče za prihodnje izboljšave.



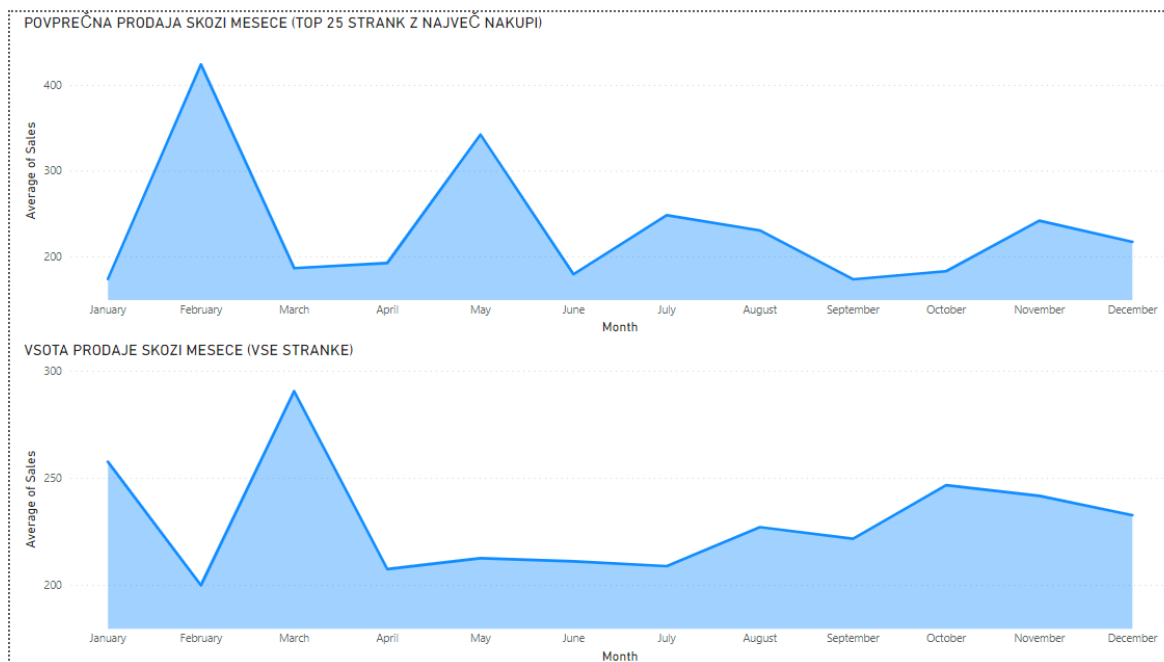
Slika 2: Primerjava strank po načinu pošiljanja izdelka

Nato sem analiziral razliko po vsoti prodaje na časovni skali po mesecih za vse stranke in vračajoče stranke. Vse stranke imajo močan padec prodaje v februarje, nato skok v marcu in nato se prodaja konstantno dviga skozi mesece do konca leta. 25 strank z najbolj pogostimi nakupi pa ne beležijo padca prodaje v februarju, ampak prav tako nimajo povišane prodaje v marcu, vendar imajo povišano prodajo v maju in juliju v primerjavi z drugimi meseci. Po tem se tudi pri njih prodaja s časom do konca leta konstantno dviguje.



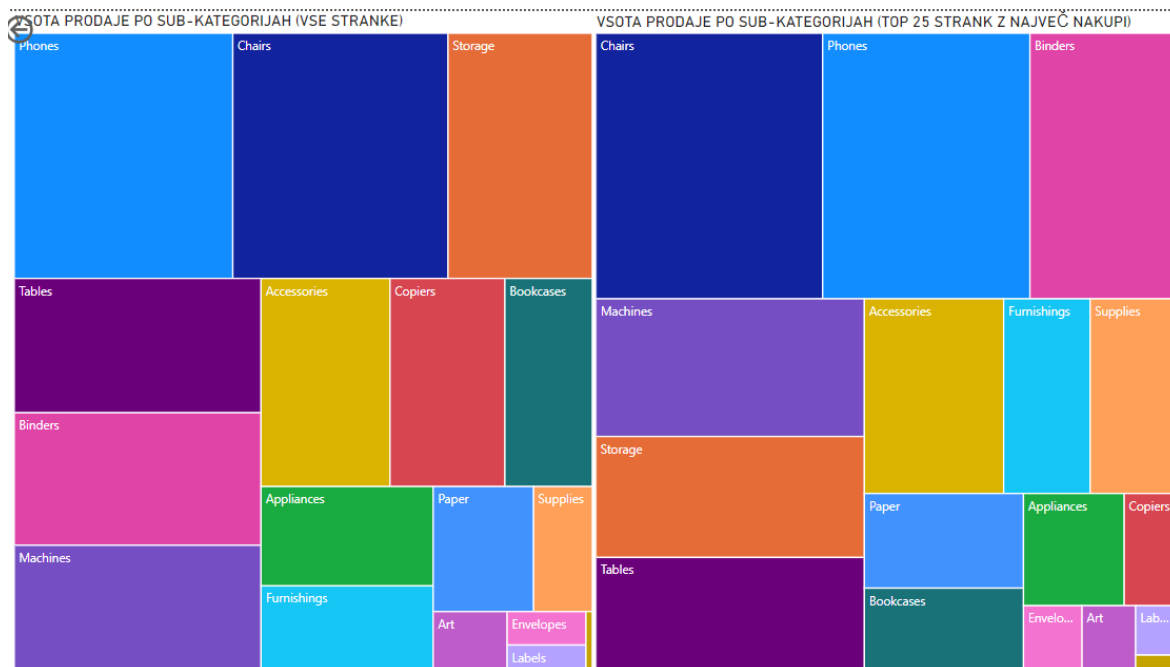
Slika 3: Primerjava vsote prodaje skozi mesece

Analiziral sem še eno povprečno prodajo skozi mesece za vse stranke in 25 najbolj pogosto vračajočih se strank. Zanimivo je, da je februar predstavljal najvišji povprečni nakup za pogosto vračajoče se stranke, medtem ko pri vseh strankah je februar predstavljal najnižji povprečni nakup. Povprečni nakup se je v maju dvignil za vračajoče stranke, za vse stranke pa je bil zelo uspešen marec.



Slika 4: Primerjava povprečne cene izdelka skozi mesece

Po tem sem primerjal še razlike med omenjenima tipoma strank v sub-kategorijah. Ugotovil sem, da so med vračajočimi strankami zelo popularni stoli in mape. Na podlagi preteklih rezultatov sem izbral produkte, ki so v sub-kategoriji mape in stoli, so bili kupljeni s strani 25 strank z največ nakupi in jih razvrstil po številu najbolj pogosto kupljenih.



Slika 5: Primerjava sub-kategorij

Product Name	Count of Order ID	Sum of Sales
DMI Arturo Collection Mission-style Design Wood Chair	4	1238,04
GBC Standard Therm-A-Bind Covers	4	134,57
Premium Transparent Presentation Covers by GBC	4	226,58
Avery Non-Stick Binders	3	46,70
Avery Reinforcements for Hole-Punch Pages	3	9,31
Fellowes PB200 Plastic Comb Binding Machine	3	1053,94
GBC DocuBind TL300 Electric Binding System	3	8252,31
Global Executive Mid-Back Manager's Chair	3	2269,64
Heavy-Duty E-Z-D Binders	3	43,64
Recycled Easel Ring Binders	3	67,14
Storex Dura Pro Binders	3	38,02
Storex DuraTech Recycled Plastic Frosted Binders	3	11,87
Zipper Ring Binder Pockets	3	16,22
Avery Heavy-Duty EZD Binder With Locking Rings	2	22,32
Total	183	74143,14

Slika 6: Popularni izdelki za vračajoče stranke

Ugotovitve in predlogi

Na podlagi analize bi vodstvu podjetja svetoval:

- Strankam, ki so tipa Home Office je potrebno ponuditi dodatno podporo in vodenje skozi nakup.
- Opravi se lahko dodatna analiza, ki se bo fokusirala na tip stranke Consumer pri strankah z vračajočimi nakupi, da vidimo nove vpoglede, ki pomagajo stranke, da se odloči za ponovni nakup.
- Pri načinu pošiljanja smo lahko zadovoljni z našimi Premium načini pošiljanja. Svetujem, da se fokusiramo na boljše izvajanje standardne storitve pošiljanja. Prav tako se lahko sestavi marketinška kampanja, ki promovira Premium načine pošiljanja, saj dobro delujejo v našem podjetju.
- Glede na celotno prodajo skozi mesece v podjetju, predlagam, da se padec prodaje v februarju ne bo zgodil, da se strankam, ki so opravile prvi nakup v mesecu decembru v preteklem letu ponudi popust za nakup v februarju.
- Glede na skok v prodaji pri vseh strankam v marcu v primerjavi z najbolj pogosto vračajočimi strankami predlagam, da se v mesecu marcu fokusiramo na podporo strankam in komunikaciji strankam, saj bo to tlakovalo pot za uspešno prodajno leto. Strank se na podlagi podatkov po slabši prodaji v mesecu februarju ponovno odločijo za nakup v marcu, zato so verjetno pri nakupu še bolj pozorne in natančne na napake.
- Povprečni nakup je pri vseh strankah v marcu visok, zato je potrebno dostavljati pakete natančno in zanesljivo.
- Pri sub-kategorijah je potrebno spodbuditi nakupe stolov in map, ki so izdelki, ki pripomorejo k vračanju in ponovnim nakupom strank.
- Popularne produkte za vračajoče stranke v sub-kategoriji stoli in mape lahko razdelimo na dva dela: cenejše izdelke lahko ponujamo kot dodatne nakupe ob

zaključku nakupa, medtem ko dražje izdelkom lahko glede na zanimanje in uspešnost ponudimo strankam, ki so kupile v naši trgovini le enkrat in niso imeli najboljših izkušenj.

Viri in literatura

<https://www.kaggle.com/datasets/rohitsahoo/sales-forecasting>