

Tartu Ülikool
Loodus- ja täppisteaduste valdkond
Matemaatika ja statistika instituut

Tarbijate ostukäitumine ühe kaubandusettevõtte põhjal

Projekt aines Rakendustarkvara: R

Autor: Anastassia Gorškova

Tartu 2020

Sissejuhatus

Projekti eesmärgiks on uurida tarbijate ostukäitumist ettevõtte kolme kaupluse põhjal, mille andmed olid kogutud 2019. aastal. Andmestik sisaldab andmeid 1000 ostu kohta, nagu näiteks ostu teinud kliendi sugu ja tüüp, ostu asukoht, kauba hind jne. Kokku on andmestikus 17 tunnust ja 1000 vaatlust, mida moodustab kaupluses sooritatud ostude arv 3 kuu jooksul. Oma projekti raames püstitasin järgmised uurimisküsimused:

1. Mis on keskmine kauba hind igas tootekategoorias? Milline tootekategooria on kõrgeima keskmise hinnaga?
2. Kas on seos tarbija soo ja tema ostukäitumise vahel?
3. Millisel ettevõtte kauplusel on kõige kehvem teeninduskvaliteedi reiting?
4. Kas on seos makseviisi ja ostu teostamise kuu vahel?
5. Kui palju oste päevas oli sooritatud kolme kuu jooksul?

Andmestik

Andmed 2019. aasta kaubandusettevõtte klientide üksikute ostude kohta on saadud addressilt <https://www.kaggle.com/aungpyaeap/supermarket-sales>.

Kokku on andmestikus 17 tunnust:

1. Invoice ID – arvuti poolt genereeritud arve number
2. Branch – harukauplus (võimalikud väärtused on A, B või C)
3. City – linn, kus asub kauplus (kaupluse A juhul on see Yangon, B – Mandalay, C – Naypyitaw)
4. Customer type – külastaja tüüp (st kliendikaardiga või ilma kliendikaardita külastaja)
5. Gender – külastaja sugu (mees või naine)
6. Product line – tootekategooria
7. Unit price – kauba hind (dollarites)
8. Quantity – ostetud toodete arv
9. Tax 5% - toodete maks (5%)
10. Total – kogusumma (maks on kaasa arvatud)
11. Date – ostu teostamise kuupäev
12. Time – ostu teostamise aeg (10:00 kuni 21:00)
13. Payment – makseviis (võimalikud väärtused on Cash, Credit Card või Ewallet)
14. COGS – kulud müüdud kaubale
15. Gross margin percentage – kasumi marginaali protsent
16. Gross income - brutotulu
17. Raiting – külastaja rahulolu saadud teeninduse kvaliteediga (skaalas 1 – 10)

Kaupade maksumuse ülevaade

Alguses vaatleme, mis on keskmine ostetud kaupade hind, lähtudes sellest, millisse kategooriasse tooted kuuluvad.

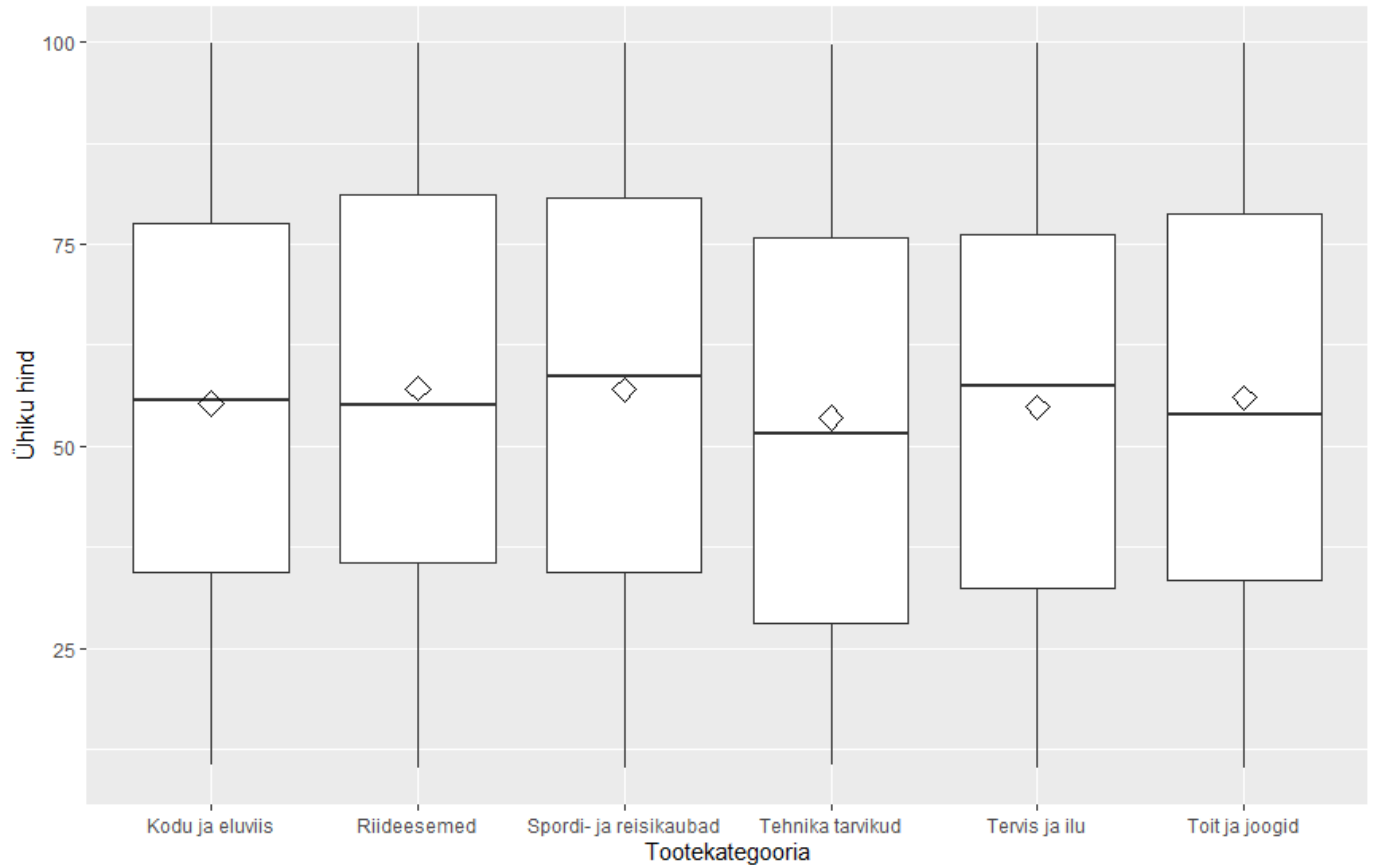
Tootekategooria	Keskmine kauba hind (\$)	Minimaalne hind (\$)	Maksimaalne hind (\$)
Tehnika tarvikud	53.55	10.56	99.73
Riideesemed	57.15	10.18	99.89
Toit ja joogid	56.01	10.13	99.79
Tervis ja ilus	54.85	10.08	99.96
Kodu ja eluviis	55.32	10.53	99.92
Spordi- ja reisikaubad	56.99	10.17	99.96

Tabel 1, kaupade hindade keskmine, miinimum ja maksimum

Tabelist 1 võib näha, et ostetud tooted on enam-vähem samasuguse keskmise hinnaga. Kõige kõrgem keskmine hind on kaupadel riideesemete kategooriast (keskmiselt 57.15 dollarit kauba ühiku eest), ning kõige madalam keskmine hind on kaupadel tehnika tarvikute kategooriast (keskmiselt 53.55 dollarit kauba ühiku eest). Paistab imelik see fakt, et tooted toidu ja jookide kategooriast on keskmiselt sama kallid, kui teiste kategooriate tooted.

Võiks ka uurida, mis on minimaalne ja maksimaalne kauba hind igas tootekategoorias. Enne seda oleks tähtis märkida, et siin on tegemist juhusliku valimiga, kuhu kuuluvad küllastajate ostud, mis olid teostatud kolme kuu jooksul. See tähendab, et ettevõtte kauplustes võivad olla ka tooted madalama või kõrgema hinnaga.

Võime nüüd ka saadud tulemusi esitada karpdiagrammi abil.



Joonis 1, kaupluste hindade karpdiagrammid

Karpdiagrammid (Joonis 1) näitavad ettevõtte kaupluste hindade jaotust tootekategooria põhjal, nimelt näitab valge ristkülik alumist (25%) ja ülemist (75%) kvartiile, väike romb näitab hindade keskmist ja joon ristkülikus näitab mediaani. Tabelist (Tabel 1) ja karpdiagrammist (Joonis 2) näeme, et igas tootekategoorias ei ületa hinnad 100 dollarit (ümardatud) ning samas nad ei ole madalamad, kui 10 dollarit (ümardatud), mis on päris imelik näha toidu ja jookide kategooria puhul, kuid nii kõrge maksimaalne toote hind võiks olla põhjustatud näiteks kalli alkoholi müügist.

Ostukäitumine külastaja soo järgi

Eisialgu vaatleme mees- ja naissoost külastajate osakaalu.

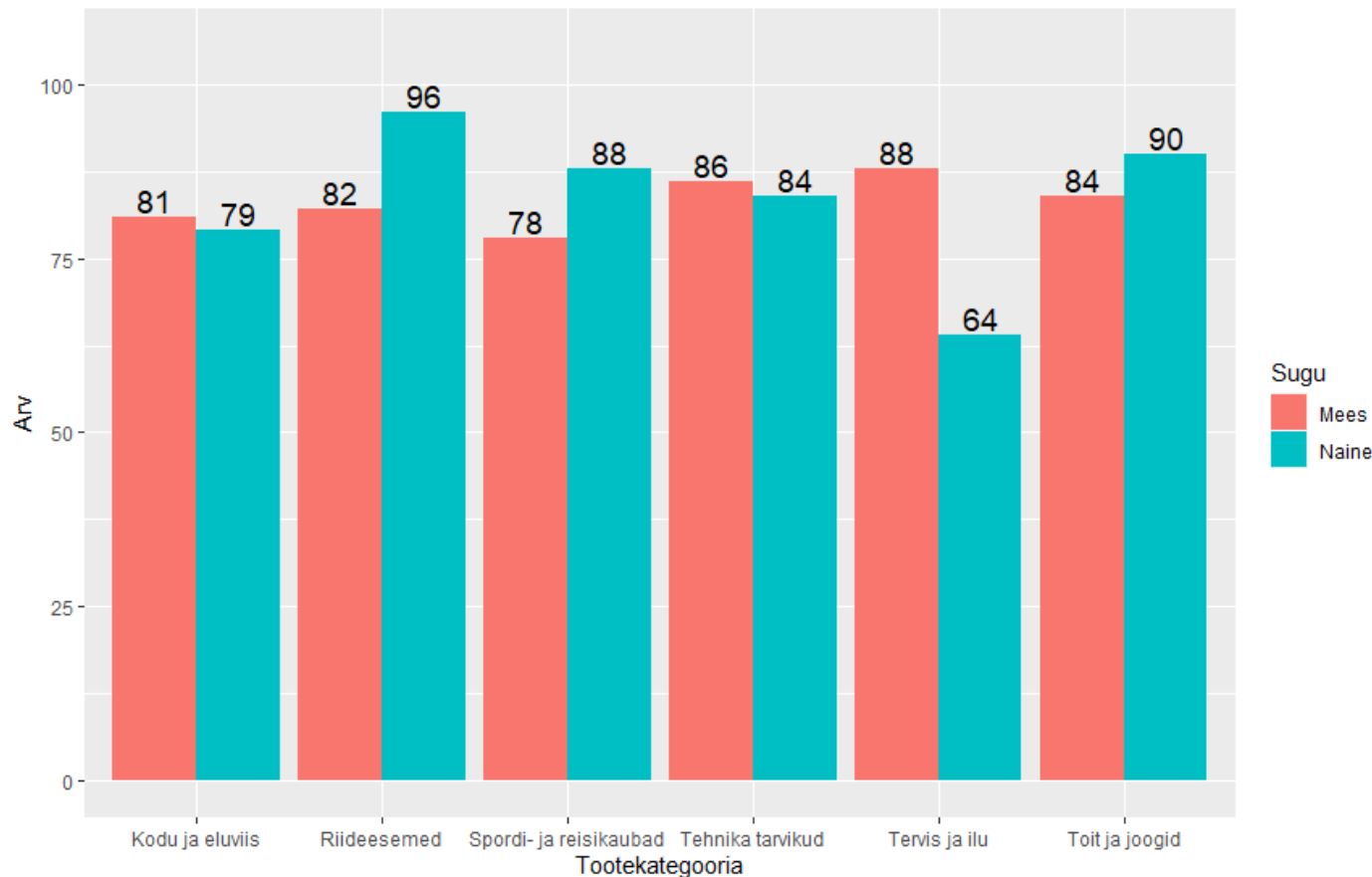
Külastaja sugu	Jaotus soo järgi	Jaotus soo järgi protsentides (%)	Jaotus liikme staatuse järgi	Jaotus liikme staatuse järgi protsentides (%)
Mees	499	49.9	240	47.9
Naine	501	50.1	261	52,1

Tabel 2, meeste ja naiste osakaalud kolme kaupluse põhjal

Tabelist 2 on hästi näha, et meeste ja naiste jaotus on peaaegu võrdne. Ettevõtte kauplusi külastas 2019. aastal 499 meest ja 501 naist kolme kuu jooksul.

Oleks huvitav ka uurida, keda on liikme staatusega rohkem, kas mehi või naisi.

Nagu Tabelist 2 on näha, kokku on ettevõtte kaupluse liikme staatusega 501 inimest, mis on peaaegu pool külastajate koguhulgast (1000). Tabeli 2 põhjal näeme ka, et naised vormistavad endale liikme staatust sagedamini, kui mehed. See võib olla seotud sellega, et naised tavaliselt külastavad kauplusi rohkem, kui mehed, ning suhtuvad ostudesse suurema vastutustundega, eriti kui tegemist on toiduainete ostuga.



Joonis 2, meeste ja naiste ostude arv tootekategooria põhjal

Jooniselt 2 on näha, et 2019. aasta andmete põhjal ostsid naised riideesemeid (96 kauba ühikut), toitu (90 kauba ühikut) ja sporditarbeid (88 kauba ühikut) rohkem, kui mehed. Päril imelik tundub see, et tervise ja ilu kategooriasse kuuluvaid tooteid ostsid mehed küllalt rohkem, võrreldes naissoost küllastajatega (vastavalt 88 toodet ja 64 toodet). Selgus, et tervise ja ilu kategooria oli kõige ebapopulaarsem tootekategooria naiste seas, mis osutus omakorda kõige ostetavamaks kategooriaks meessoost küllastajate seas. Populaarseim tootekategooria naiste seas on seotud riideesemetega (96 müüdud kauba ühikut), mis on üldiselt oodatav. Kõige vähem ostsid mehed spordi ja reisi kategooriasse kuuluvaid tooteid (78 müüdud kauba ühikut).

Teeninduskvaliteedi reitingu ülevaade

Esiialgu vaatleme, milline keskmine teeninduskvaliteedi reiting on igal ettevõtte kauplusel. Huvi võiks pakkuda ka see, millist kõrgeimat ja madalaimat hinangut on iga ettevõtte kauplus saanud.

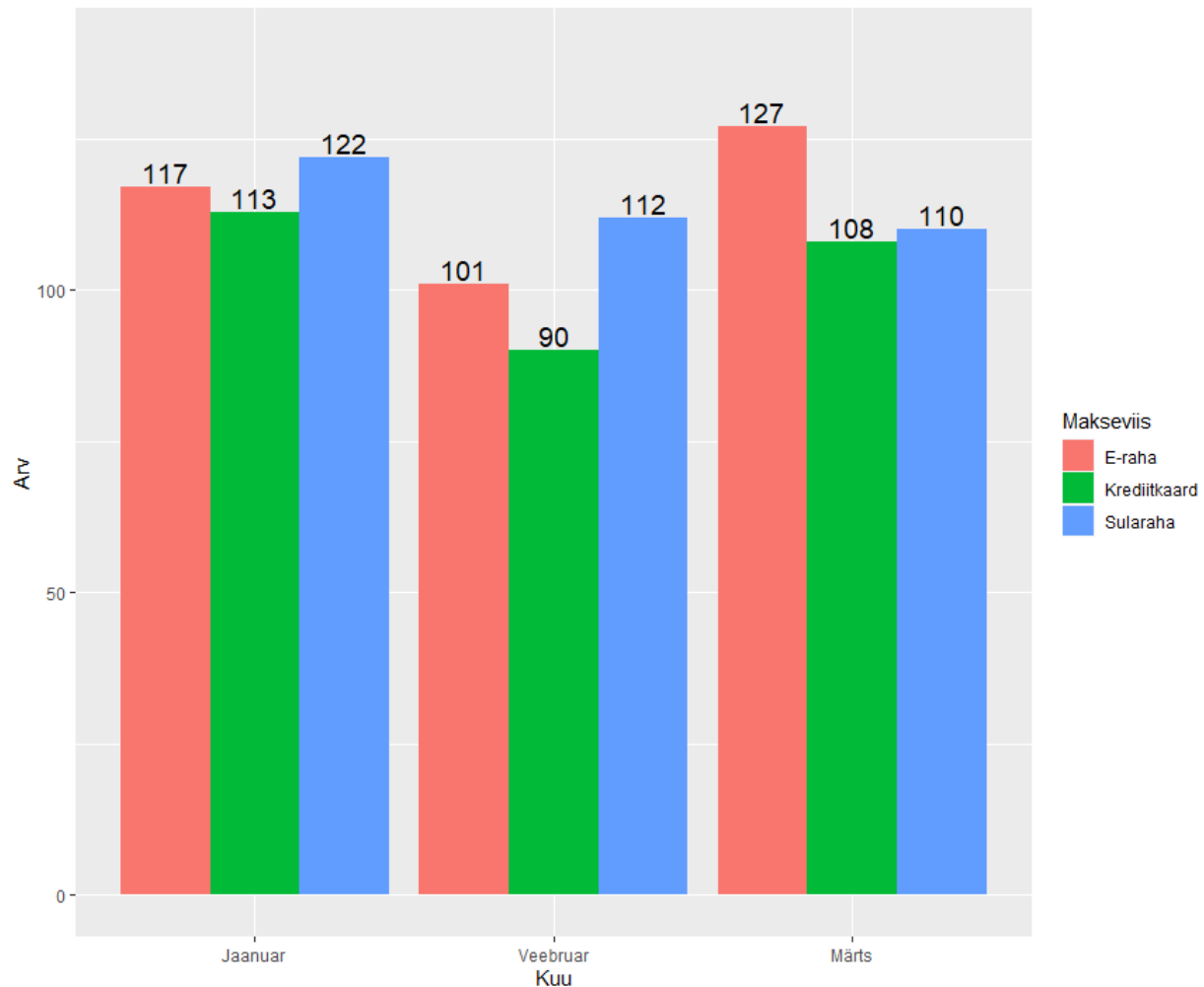
Kauplus	Keskmine teeninduskvaliteedi reiting	Minimaalne hinnang	Maksimaalne hinnang
A	7.03	4.0	10.0
B	6.82	4.0	10.0
C	7.07	4.0	10.0

Tabel 3, keskmine teeninduskvaliteedi reiting ettevõtte kaupluse põhjal

Tabelist 3 võib kohe märgata, et külastajate teeninduskvaliteedi keskmine hinnang on 2019. aastal enam-vähem samasugune, erandiks võib lugeda ainult kauplust B, mille keskmine reiting on natuke madalam (6.82), kui ülejäänud kauplustel.

Tabelist 3 näeme ka, et kõigis vaadeldud kauplustes A, B ja C muutuvad reitingu hinnangud 4 ja 10 vahel. Saadud tulemustega võib ka põhjendada, miks kaupluste teeninduskvaliteedi reitingud on peaaegu võrdsed.

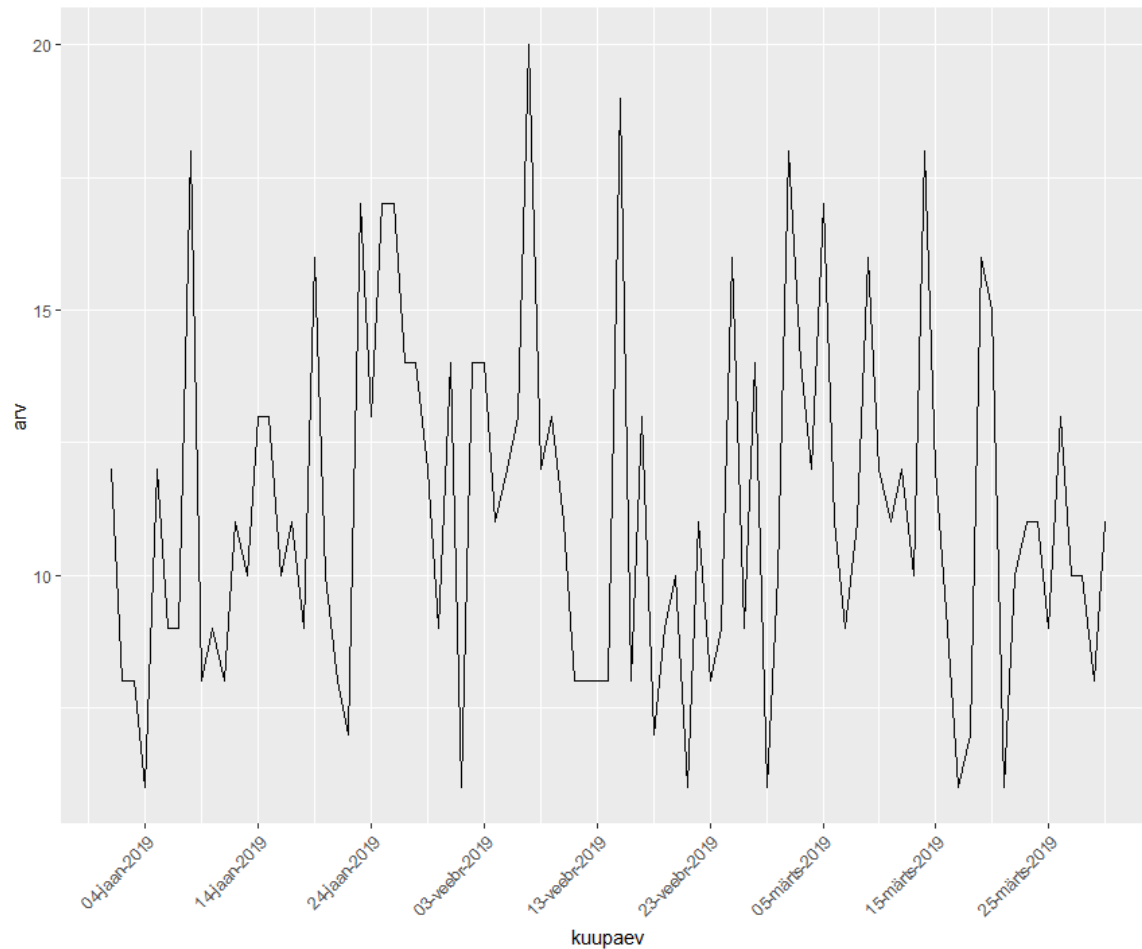
Makseviisid ostu teostamise kuu järgi



Joonis 3, ostude arv kuu ja makseviisi järgi

Joonise 3 põhjal näeme, et kolme kuu jooksul on krediitkaardiga makse teostamine jäänud kõige ebapopulaarsemaks makseviisiks. Kokku oli kolme kuuga sooritatud 311 krediitkaardiga makset, võrreldes 344 sularahamakse ja 345 e-rahamaksega. Jaanuari ja veebruari jooksul kasutati kõige rohkem sularahamakset, kuid märtsis on e-rahakoti kasutamine märgatavalt suurenenud (127 makset), võrreldes eelmiste kuudega (117 ja 101 makset vastavalt jaanuaris ja veebruaris).

Ostude arv kuupäeva järgi



Joonis 4, Ostude arv kuupäeva järgi alates 1. jaanuarist 2019 kuni 30. märtsini 2019

Jooniselt 4 näeme, et 3 kuu jooksul on ettevõttele kaupluste ostude arv pidevalt kõikunud. Kõige rohkem oste oli 7. veebruaril (20 ostu päeva eest) ning kõige vähem 4. jaanuaril, 1. veebruaril, 21. veebruaril, 28. veebruaril, 17. märtsil ja 21. märtsil (6 ostu päeva eest). Saadud tulemuste põhjal võiks oletada, et ettevõtte ei ole väga suur ja populaarne, kauplustesse käib üsna piiratud arv inimesi.

Kokkuvõtte

1. Ettevõtte kaupluste toodete hindu uurides, avastasin, et eri tootekategooriate ostetud kaubad on enam-vähem võrdse keskmise hinnaga. Kõrgeima keskmise hinnaga tootekategooriaks osutusid riideesemed.
2. Edasisel uurimisel selgus, et ettevõtte kauplustes käis 49.9% meessoost küllastajatest ning vastavalt 50.1% naissoost küllastajatest, mis on peaaegu võrdseks jaotuseks. Umbes pool kaupluste küllastajatest omab liikme staatust, mille hulgas on omakorda rohkem naisi, kui mehi.
3. Uurides kaupluste teeninduskvaliteedi reitingut, selgus, et kolme kaupluste reitingud ei ole väga erinevad omavahel. Kauplusel B on kvaliteedi hinnang natuke kehvem, võrreldes kauplustega A ja C.
4. Vaadeldes seost makseviisi ja ostu teostamise kuu vahel, avastasin, et kõige ebapopulaarsemaks makseviisiks on kolme kuu jooksul (jaanuarist märtsini) jäänud krediitkaardi kasutamine. Küllastajad eelistavad kasutada rohkem sularaha ja e-rahakotti, eriti on see näha veebruari kuus.
5. Kolme kuu jooksul ostude arv on olnud üsna ebastabiilne. Ostude arv osutus maksimaalseks 7. veebruaril, millal küllastajad on teinud 20 ostu. Kõige vähem osteti 4. jaanuaril, 1. veebruaril, 21. veebruaril, 28. veebruaril, 17. märtsil ja 21. märtsil, millal oli tehtud ainult 6 ostu ühe päeva eest.