

LATVIJAS UNIVERSITĀTE
IZGLĪTĪBAS ZINĀTŅU UN PSIHOLOĢIJAS FAKULTĀTE
SKOLOTĀJU IZGLĪTĪBAS NODAĻA

MATEMĀTIKAS OPTIMĀLĀ LĪMEŅA EKSĀMENU ANALĪZE
PĒC URBANIZĀCIJAS PAKĀPES

Kursadarbs kursā
Matemātiskā statistika un tās metodika

Autore: **Liena Kāpiņa**
Studenta apliecības nr.: lk19101
Darba vadītāja: Leonora Pahirko

RĪGA 2025

Saturs

Ievads	2
1. Datu pirmapstrāde	3
2. Aprakstošā statistika	3
3. Secinošā statistika	5
4. Secinājumi	6
4.1. Diskusija	7
A Pielikums. Tabulas	9
B Pielikums. Programmas R kods	10

Ievads

Pētījumam izvēlēts valsts eksāmens matemātikā optimālajā līmenī. Šo eksāmenu kārtot skolēni (vai pieaugušie), kas apguvuši kursu matemātika I. Skolā to iespējams darīt, beidzot 11. klasi vai beidzot 12. klasi.

Dati pētījumam iegūti no Valsts izglītības satura centra (VISC) mājaslapas¹. Dati izvēlēti, jo tie ir publiski pieejami un gana uzticami, jo eksāmeni tiek laboti centralizēti.

Iepriekš jau veikti starptautiski pētījumi par skolēnu prasmēm dažādos vecumos un priekšmetos un pētnieki vērsuši uzmanību uz zemākiem sasniegumiem lauku skolās. Tāpat jau vairākus gadus publiskajā telpā notiek diskusijas par lauku skolu slēgšanu. Lauku skolās ir maz skolēnu, bet, piemēram, ēkas uzturēšanas izmaksas var būt līdzīgas kā lielākā skolā. Tāpat ir problēmas ar pedagogu piesaistīšanu, pedagogiem laukos ir zemākas slodzes nekā pilsētās. Izglītības un zinātnes ministrijas viens no uzdevumiem jau vairākus gadus ir bijis skolu tīkla optimizācija. Pirms skolas slēgšanas IZM jāizvērtē, vai tas ir optimāli un eksāmenu rezultāti varētu sniegt ieskatu izglītības kvalitātē skolā. Tā kā vidusskola nav obligāts izglītības posms un skolēni šajā vecumā ir patstāvīgāki, tad dati var sniegt ieskatu vidējās izglītības iestāžu akadēmiskajā sniegunā matemātikā.

Pētījuma problēma: Kā atšķiras rezultāti optimālā līmeņa matemātikas eksāmenā atkarībā no urbanizācijas pakāpes?

Pētījumā pārbaudītas trīs **hipotēzes**.

1. Eksāmenā iegūto punktu vidējās vērtība statistiski būtiski neatšķiras no 35.12^2 pie nozīmības līmeņa $\alpha = 0.05$.
2. Vidējais punktu skaits eksāmenā Rīgā un valsts nozīmes pilsētā nav statistiski nozīmīgi atšķirīgs pie nozīmības līmeņa $\alpha = 0.05$.
3. Eksāmena vidējais punktu skaits nav statistiski nozīmīgi atšķirīgs starp dažādām urbanizācijas pakāpēm pie nozīmības līmeņa $\alpha = 0.05$.

Pētījuma mērķauditorija: Politikas veidotāji, skolu vadītāji, skolotāji.

¹<https://www.visc.gov.lv/lv/valsts-parbaudes-darbi-20232024-mg-statistika>

²2023. gada matemātikas optimālā līmeņa vidējā punktu skaits, kā norādīts https://www.visc.gov.lv/sites/visc/files/media_file/cematol_3_dalu_salidzinajums.png

1. Datu pirmapstrāde

Datu kopa satur 13 633 skolēnu matemātikas optimālā līmeņa eksāmenu rezultātus atsevišķās eksāmena daļās - 1. zināšanas, izpratne un prasmes un 2. kompleksu problēmu risināšana, kā arī rezultāti par abām daļām kopā. Datiem ir 12 kolonnas.

Pirmā kolonna satur numuru pēc kārtas un tie ir ordināli dati. Kolonna “Norise” satur eksāmena nosaukumu un visiem ir vienāda, jo visi dati ir par matemātikas optimālo līmeni. Šie dati ir nomināli. Kolonna “pašvaldība” satur pašvaldības nosaukumu, kurā atrodas skola un tie ir nomināli. “Iestāde” satur iestādes nosaukumu, kurā eksāmens kārtots. Visbiežāk tā ir skola, kurā mācās, taču eksāmenus var kārtot arī pieaugušie, kas mācības skolā jau beiguši. Šie ir nomināli dati. Kolonna “Punkti Zināšanas. izpratne un prasmes” satur iegūtos punktus eksāmena 1. daļā un to vērtības ir skaitļi no 0 līdz 75. Šie ir skaitliski, diskrēti dati, jo valsts eksāmenus var iegūt tikai veselu skaitu punktu. “Procenti Zināšanas. izpratne un prasmes” satur procentus, ko persona ieguvusi eksāmena 1. daļā, tie ir skaitliski, diskrēti dati no 0 līdz 100. “Punkti Kompleksu problēmu risināšana” satur skaitliskus, diskrētus datus no 0 līdz 25. Atšķirībā no 1. daļas, 2. daļā ir iespējams iegūt arī puspunktu, taču sīkāka dalījuma nav. “Procenti Kompleksu problēmu risināšana” satur 2. daļā iegūtos procentus; arī šie ir skaitliski, diskrēti dati. “Punkti kopā” satur 1. un 2. daļas punktu summu, tātad tie ir skaitliski un diskrēti, turklāt ir robežās no 0 līdz 100. “Procenti kopā” ir diskrēti dati robežās 0 līdz 100. Tā kā eksāmena 2. daļā bija iespējams iegūt pusi no punkta, tad kopējais punktu skaits var būt daļskaitlis, taču kopējie procenti ir tikai šī skaitļa veselā daļa. “Tips” satur skolas veidu pēc klasifikācijas 3 - vidusskola, 4 - tālmācības vidusskola, 6 - valsts ģimnāzija, 7 - profesionālās izglītības iestāde, 9 - augstskola (personas, kas jau ieguvušas vidējo izglītību). Šie dati ir nomināli, jo iestāžu veidi neveido hierarhiju. Savukārt, “Urbanizācija” satur urbanizācijas pakāpi iestādes adresē pēc klasifikatora, 1 - Rīga, 2 - valstspilsētas, 3 - novadu pilsētas, 4 - lauki. Šie dati ir ordināli, jo norāda, cik urbanizētā vietā atrodas skolas, sākot no galvaspilsētas un beidzot ar lauku skolām.

No datiem izņemtas kolonnas ar kārtas numuru un eksāmena norisi, jo tās nebūs nepieciešamas.

Dati sadalīti grupās pēc urbanizācijas pakāpes - Rīga, valsts nozīmes pilsēta, novada nozīmes pilsēta un lauki. Lai to izdarītu, kolonna, kas satur urbanizācijas tipu programmā R pārdēvēta no tipa *double* uz tipu *character*. Pēc tam dati sadalīti tālākās apakšgrupās pēc iestādes tipa - Rīgas vidusskola, Rīgas tālmācības vidusskola, valsts pilsētas ģimnāzija utt. Arī kolonna, kas satur iestādes tipu pārdēvēta no *double* uz *character*, lai to paveiktu.

2. Aprakstošā statistika

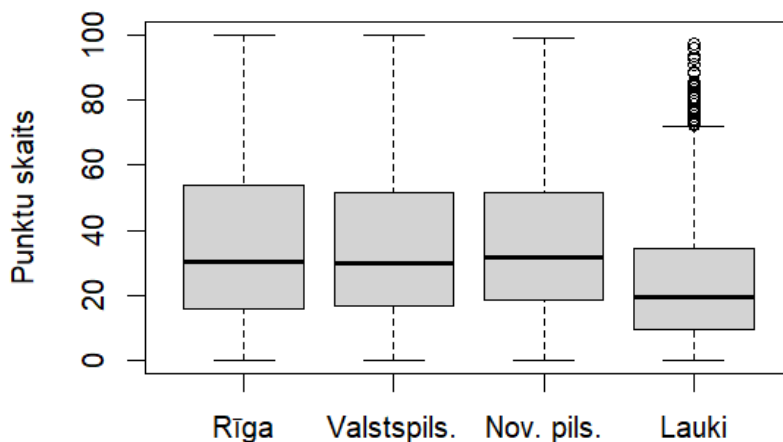
Dati tiks salīdzināti pēc urbanizācijas pakāpes. Pēc vidējām vērtībām (skat. 1. tabulu) Rīga, valsts nozīmes un novada nozīmes pilsētas ir līdzīgas, taču laukos vidējie iegūtie punkti ir zemāki.

Urbanizācija	Rīga	Valsts pilsēta	Novada pilsēta	Lauki
Punktu summa, vidējais	36.2	35.9	36.2	24.7
Punkti 1. daļā, vidējais	30.1	30.0	30.4	21.3
Punkti 2. daļā, vidējais	6.1	6.0	5.7	3.5

1. tabula Vidējais punktu skaits pēc urbanizācijas pakāpes

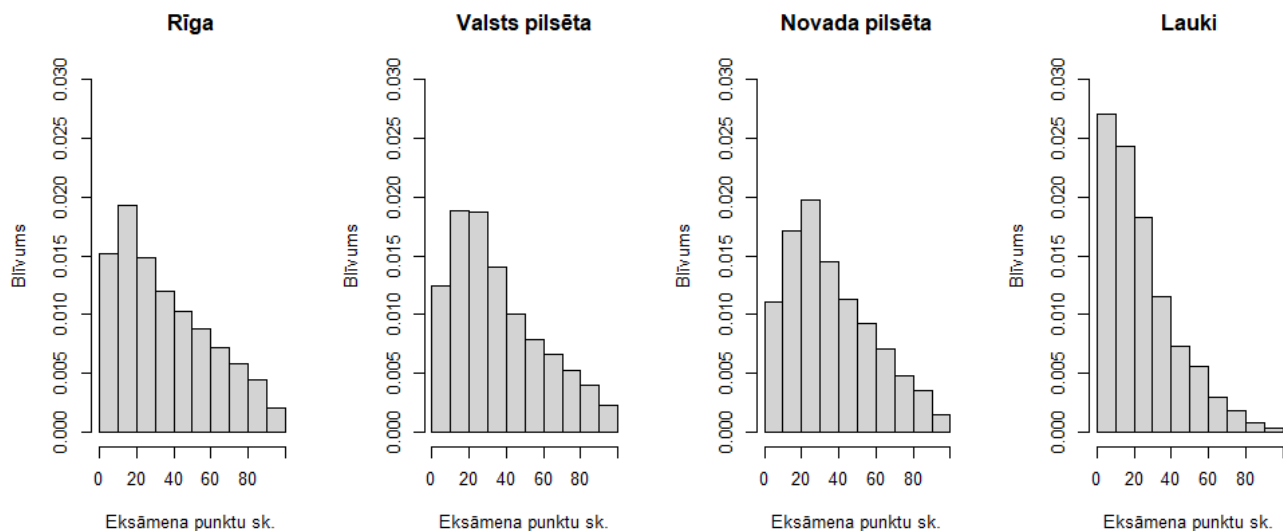
Kastu diagrammā (skat. 1. zīm.) redzams, ka Rīgā un pilsētās arī kvartiles ir līdzīgas. Savukārt, laukos kvartiles ir zemākas, taču ir arī izlecošās vērtības, kas nav pilsētām. Izlecošās

vērtības, kas saistītas ar netipiski augstiem rezultātiem, kopā ir 202 jeb 10.8% no visiem eksāmena kārtotājiem laukos. Par spīti izlecošajām vērtībām 0.95-tā kvantile saglabā kopējo tendenci, jo Rīgā tā ir 83.5 punkti, valsts nozīmes pilsētās - 82.5, novada nozīmes pilsētās - 80.2, bet laukos krietni zemāka - 62.5. Tas nozīmē, ka laukos labākie 5% no skolēniem ieguvuši vismaz 62.5 punktus, taču visās pilsētās šis rādītājs ir lielāks par 80. Kvartīļu precīzās vērtības iespējams aplūkot A pielikumā, 3. tabulā.



1. zīm. Kastu diagramma ar kopējo punktu skaitu eksāmenā pēc urbanizācijas pakāpes

2024. gadā bija nepieciešams iegūt 15 punktus no 100 visā eksāmenā, lai tas būtu nokārtots. Kopā eksāmenu nenokārtoja 21.4% no visiem licejiem, Rīgā sliekšni nesasniedza 21.3% skolēnu, valsts nozīmes pilsētās - 18.1%, novada nozīmes pilsētās - 16.1%, bet laukos - 36.7%.

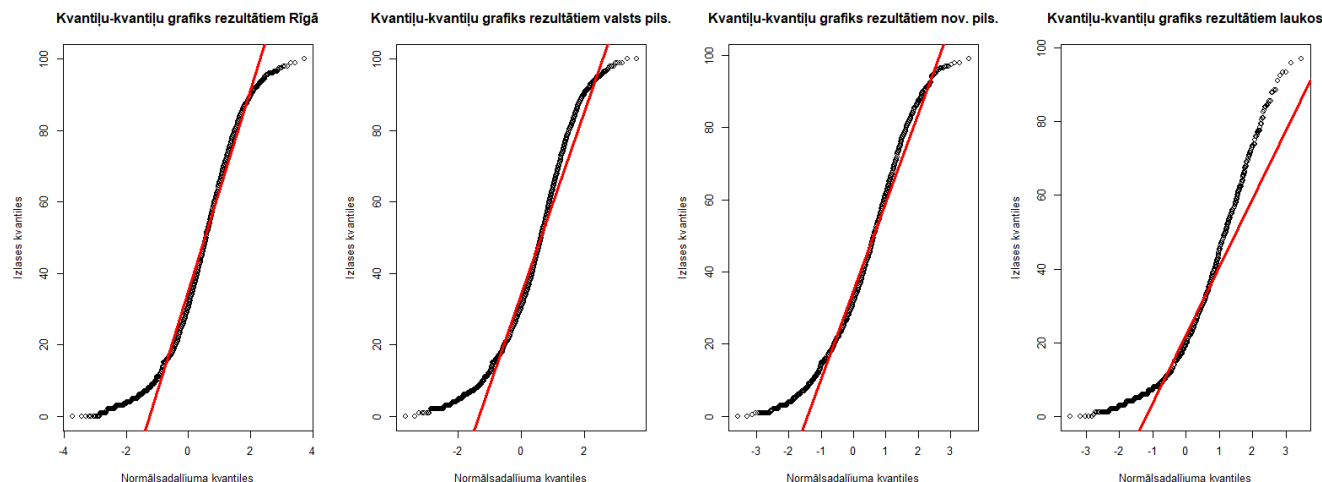


2. zīm. Kopējo iegūto punktu histogrammas pēc urbanizācijas pakāpes

Histogrammās, kas redzamas 2. zīm., var novērot, ka Rīgā visbiežāk iegūti punkti intervālā no 10 līdz 20 punktiem, tāpat tas ir valsts nozīmes pilsētās, novada nozīmes pilsētās - 20 līdz 30 punkti, bet laukos - 0 līdz 10 punkti.

Eksāmenā iegūtajiem punktiem varētu atbilst normālais sadalījums, taču visas histogrammas ir asimetriskas uz labo pusi. Asimetrija ir diezgan liela, jo vidējās vērtības stipri atšķiras no 50 un

arī no atbilstošajām mediānām jeb 0.5-ām kvantilēm. Normālsadalījuma neatbilstību apstiprina arī kvantiļu-kvantiļu grafiki 3. zīm., jo datu punkti sistemātiski neatrodas tuvu taisnei.



3. zīm. Kvantiļu-kvantiļu grafiki salīdzinot ar normālsadalījumu

Visi līdz šim aprēķinātie rādītāji norāda uz to, ka laukos ir zemāki eksāmena rezultāti, taču tas varētu būt saistīts ar to, kādās iestādēs skolēni mācās un pēc tam kārtot eksāmenu. Tā kā ģimnāzijās ir augstāki vidējie rezultāti (skat. A pielikumu, 4.), tad skolēnu skaits, kas kārtoja eksāmenu valsts ģimnāzijā, varētu ietekmēt visus statistiskos rādītājus reģionā.

	Rīga	Valsts pilsēta	Novada pilsēta	Lauki
Vidusskola	44.4%	37.8%	54.4%	35.0%
Tālmācības vidussk.	14.6%	0.9%	5.7%	37.2%
Prof. izgl. iestāde	32.3%	39.4%	15.8%	27.0%
Valsts ģimnāzija	7.8%	21.9%	24.1%	0%
Augstskola	0.9%	0%	0%	0%

2. tabula Eksāmena kārtotāju īpatsvars reģionos atkarībā no iestādes tipa

2. tabulā novērojams, ka Rīgā ir mazāks īpatsvars ar eksāmenu licējiem, kas to dara ģimnāzijā, salīdzinot ar valsts un novada pilsētām, bet laukos ģimnāziju nav vispār. Laukos ir liels tālmācības iestāžu īpatsvars. Tālmācības iestādēm savukārt ir zems vidējais punktu skaits (skat. A pielikumu, 4. tabulu). Laukos ir lielāks relatīvais biežums iestādēm ar zemākiem rezultātiem un nav ģimnāziju, kurām ir augstāks punktu skaits. A pielikumā iespējams aplūkot 5. tabulu, kurā ir iestāžu vidējais punktu skaits atkarībā no urbanizācijas pakāpes.

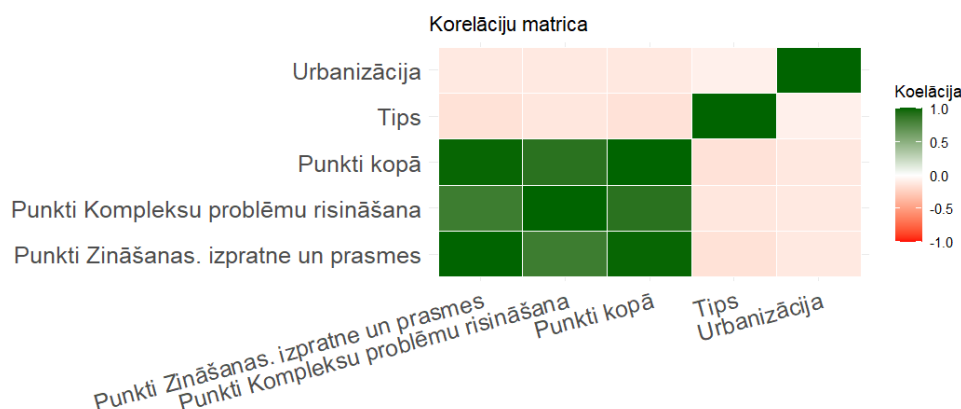
3. Secinošā statistika

Pirmā izvirzītā nulles hipotēze bija, ka kopējais iegūto punktu vidējais būtiski neatšķirsies no 35.12 pie nozīmības līmeņa $\alpha = 0.05$. Populācijai nav zināms datu sadalījums, tāpēc nav iespējams lietot t-testu un jālieto Vilksona tests. Lietojot Vilksona testu, iegūts, ka p -vērtība šai hipotēzei ir mazāka par $2.2 \cdot 10^{-16}$. Tātad hipotēze ir noraidīta pie nozīmības līmeņa $\alpha = 0.5$ un 2024. gada eksāmenu punktu vidējā vērtība ir statistiski nozīmīgi atšķirīga no 35.12.

Vidējais punktu skaits visā eksāmenā Rīgā un valsts nozīmes pilsētā ir līdzīgi, taču pie liela izlases apjoma arī ļoti neliela atšķirība var būt statistiski nozīmīga. Lai pārbaudītu šo hipotēzi, veikts Vilksoksona divu izlašu tests. Pēc testa statistikas $p = 0.6022$, tātad nulles hipotēzi (vidējās vērtības vienādas) noraidīt nevar. Rīgā un valsts nozīmes pilsētās iegūtais vidējais punktu skaits nav statistiski nozīmīgi atšķirīgi.

Lai pārbaudītu hipotēzi, ka vidējais punktu skaits eksāmenā nav atšķirīgs dažādās urbanizācijas pakāpēs, veikts Kruskala-Vollisa tests. Pēc šī testa $p < 2.2 \cdot 10^{-16}$, tāpēc hipotēze tiek noraidīta. Tas nozīmē, ka pastāv statistiski nozīmīgas atšķirības punktu vidējās vērtībās Rīgā, valsts pilsētā, novada pilsētā un laukos.

Aplūkotas arī korelācijas starp punktiem 1. daļā, 2. daļā, punktiem kopā, iestādes tipu un urbanizāciju. Pēc 4. zīm. novērojams, ka pastāv cieša korelācija starp 1. daļas punktiem un kopā iegūtajiem punktiem eksāmenā. Korelācija ir cieša arī starp 2. daļā iegūtajiem punktiem un kopējiem punktiem. Tāpat tā ir cieša starp 1. daļā un 2. daļā iegūtajiem punktiem.



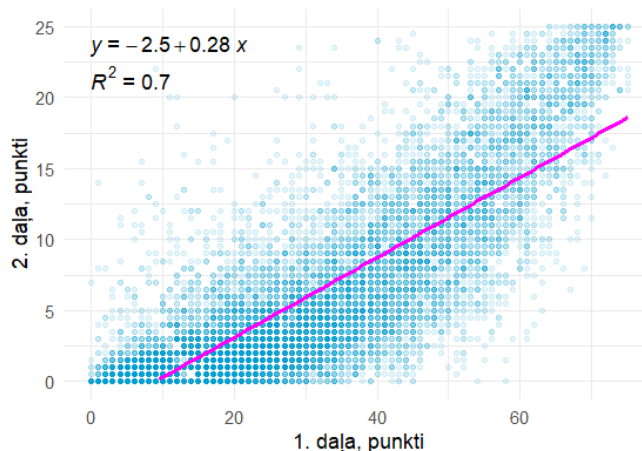
4. zīm. Enter Caption

5. zīm. izveidota izkļiedes diagramma iegūtajiem punktiem 1. un 2. daļā, kurā tumšāki punkti atspoguļo lielāku biežumu. Lielam skaitam eksāmena licēju ir mazāk par 30 punktiem 1. daļā un mazāk par 5 punktiem 2. daļā, kas apgrūtina sakarības novērtēšanu. Pārējie dati izskatās cieši korelēti, kā tas bija novērojams arī korelāciju matricā.

4. Secinājumi

Datu analīzes rezultātā formulēti sekojoši secinājumi.

1. Vidējie iegūtie punkti un attiecīgi vidējais vērtējums nav statistiski nozīmīgi atšķirīgi no iepriekšējā gada vidējā vērtējuma eksāmenā.
2. Rīgā un valsts nozīmes pilsētās nav statistiski būtiskas atšķirības starp vidējiem iegūtajiem punktiem. Tā kā novadu pilsētās vidējā vērtība ir vēl tuvāka vidējai vērtībai Rīgā, arī tā būtiski neatšķiras no Rīgas vidējās vērtības.
3. Starp vidējiem iegūtajiem punktiem dažādās urbanizācijas pakāpēs pastāv statistiski nozīmīga atšķirība. Visticamāk lauki statistiski nozīmīgi atšķiras no kāda no pilsētas veidiem (iespējams, vairākiem vienlaikus), jo laukos ir ievērojami zemāki vidējie iegūtie punkti un punktu mediāna.



5. zīm.: Izkliedes diagramma 1. un 2. daļā iegūtajiem punktiem ar pievienotu lineārās regresijas taisni

4. Kopumā laukos ir zemāki rezultāti nekā pilsētās, taču to varētu ietekmēt arī iestāžu ģeogrāfiskais sadalījums. Laukos nav ģimnāziju, kurām ir augstāki rezultāti, bet tālmācības vidusskolas, kurām ir zemāki rezultāti, sastāda vairāk kā trešdaļu no eksāmena licējiem. Nav iespējams noteikt, vai laukos ir zemāki rezultāti ģeogrāfiskā novietojuma un resursu, piemēram, kvalitatīvu skolotāju, pieejamības dēļ, vai lielā tālmācības skolēnu skaita dēļ.
5. Laukos ir ļoti dažāds sniegums, jo 10.8% no eksāmena rezultātiem laukos ir izlecošas vērtības.
6. Matemātikas optimālā līmeņa eksāmena punktu skaits nav normāli sadalīts.
7. Pastāv cieša korelācija starp 1. un 2. daļas punktiem un kopvērtējumu. Atsevišķo daļu kopvērtējumi cieši korelē ar punktu summu, jo punktu summa ir atkarīga no tiem, īpaši no 1. daļas, jo tā sastāda 75% no visiem eksāmena punktiem. Savukārt 1. un 2. daļas punktu korelācija liecina, ka pastāv kāda sakarība starp rezultātiem.

4.1. Diskusija

Tika aprēķinātas vidējās punktu vērtības pēc iestādes tipa un urbanizācijas pakāpes (skat. 5. tabulu A pielikumā), taču nebija iespējams noskaidrot, vai pastāv kāda sakarība starp iestādes veidu un urbanizācijas pakāpi. T.i., nav zināms, vai Rīgā vidusskolās ir augstāki rezultāti nekā vidusskolās valsts pilsētā.

Nav atrasts teorētiskais sadalījums, kas atbilstu datiem. Pēc informācijas iegūšanas veida datiem vajadzētu atbilst normālsadalījumam; testu rezultātus bieži apraksta normālsadalījums, taču šajā gadījumā normālsadalījums būtu ļoti deformēts.

Tā kā datiem nav normālā sadalījuma, tad arī iegūtās korelācijas un regresijas taisne neaprapsta datus tik labi, cik citi modeļi to varētu darīt. Pēc izkliedes diagrammas gan var redzēt, ka sakarība pastāv, bet iespējams tā nav lineāra.

Būtisks trūkums ir zema eksāmena licēju skaits no Rīgas valsts ģimnāzijām - vien 386 skolēni. Tas visticamāk pazemina vidējo punktu skaitu Rīgā, jo ģimnāzijās mācās vairāki tūkstoši skolēnu. Iespējams, ka Rīgas ģimnāzijās skolēni izvēlas mācīties matemātika II un nekārtot optimālā līmeņa eksāmenu 11. klasē, tā vietā kārtojot augstākā līmeņa eksāmenu. Tā kā 1. un 2. daļa abiem

eksāmeniem sakrīt un tos nevar kārtot vienā gadā, tad pētījumu varētu uzlabot pievienojot datus no augstākā līmeņa eksāmena 1. un 2. daļas.

A Pielikums. Tabulas

Kvantile	0.25	0.5	0.75	0.95
Rīga	16	30.5	54	83.5
Valsts pilsēta	17	30	51.5	82.5
Novada pilsēta	18.5	32.5	51.5	80.2
Lauki	9.5	19.5	34.5	62.5

3. tabula Eksāmenā iegūto punktu kvantiles pēc urbanizācijas pakāpes

Tips	Vidusskola	Tālmācības vidusskola	Valsts ģimnāzija	Prof. izgl. iest.	Augstskola
Vidējie punkti	39.9	17.4	56.1	23.6	18.4

4. tabula Vidējie eksāmenā iegūtie punkti pēc izglītības iestādes tipa

	Rīga	Valsts pilsēta	Novada pilsēta	Lauki
Vidusskola	44.7	37.0	37.2	36.6
Tālmācības vidussk.	16.7	13.0	20.1	17.6
Prof. izgl. iestāde	26.2	23.9	18.8	18.6
Valsts ģimnāzija	67.4	56.8	49.0	-
Augstskola	18.4	-	-	-

5. tabula Vidējais eksāmena punktu skaits pēc iestādes tipa un urbanizācijas pakāpes

B Pielikums. Programmas R kods

```
1 library(readxl)
2 data2024 <- read_excel("visi_dati_2024_0810.xlsx", sheet = "MATOL")
3 data2024$N.p.k. <- NULL
4 data2024$Norise <- NULL
5 as.character(data2024$Tips)
6 as.character(data2024$Urbanizcija)
7 split_data <- split(data2024, data2024$Urbanizcija)
8 data_riga <- split_data$"1"
9 data_pilseta <- split_data$"2"
10 data_nov <- split_data$"3"
11 data_lauki <- split_data$"4"
12
13 split_riga <- split(data_riga, data_riga$Tips)
14 rigas_vsk <- split_riga$"3"
15 rigas_talmaciba <- split_riga$"4"
16 rigas_gymn <- split_riga$"6"
17 rigas_prof <- split_riga$"7"
18 rigas_LU <- split_riga$"9"
19 mean(rigas_gymn$"Punkti kop")
20 mean(rigas_LU$"Punkti kop")
21 mean(rigas_prof$"Punkti kop")
22 mean(rigas_talmaciba$"Punkti kop")
23 mean(rigas_vsk$"Punkti kop")
24
25
26 split_pilseta <- split(data_pilseta, data_pilseta$Tips)
27 pilseta_vsk <- split_pilseta$"3"
28 pilseta_talmaciba <- split_pilseta$"4"
29 pilseta_gymn <- split_pilseta$"6"
30 pilseta_prof <- split_pilseta$"7"
31 pilseta_LU <- split_pilseta$"9"
32 mean(pilseta_gymn$"Punkti kop")
33 mean(pilseta_prof$"Punkti kop")
34 mean(pilseta_talmaciba$"Punkti kop")
35 mean(pilseta_vsk$"Punkti kop")
36
37 split_nov <- split(data_nov, data_nov$Tips)
38 nov_vsk <- split_nov$"3"
39 nov_talmaciba <- split_nov$"4"
40 nov_gymn <- split_nov$"6"
41 nov_prof <- split_nov$"7"
42 nov_LU <- split_nov$"9"
43 mean(nov_gymn$"Punkti kop")
44 mean(nov_prof$"Punkti kop")
45 mean(nov_talmaciba$"Punkti kop")
46 mean(nov_vsk$"Punkti kop")
47
48 split_lauki <- split(data_lauki, data_lauki$Tips)
49 lauki_vsk <- split_lauki$"3"
50 lauki_talmaciba <- split_lauki$"4"
51 lauki_gymn <- split_lauki$"6"
52 lauki_prof <- split_lauki$"7"
53 lauki_LU <- split_lauki$"9"
```

```

54 mean(lauki_prof$'Punkti kop')
55 mean(lauki_talmaciba$'Punkti kop')
56 mean(lauki_vsk$'Punkti kop')
57
58
59 split_skola <- split(data2024, data2024$Tips)
60 vsk <- split_skola$'3'
61 talmaciba <- split_skola$'4'
62 prof <- split_skola$'7'
63 gymn <- split_skola$'6'
64 LU <- split_skola$'9'
65 mean(vsk$'Punkti kop')
66 mean(talmaciba$'Punkti kop')
67 mean(prof$'Punkti kop')
68 mean(gymn$'Punkti kop')
69 mean(LU$'Punkti kop')
70
71 length(data_riga$'Punkti kop')
72 length(data_pilseta$'Punkti kop')
73 length(data_nov$'Punkti kop')
74 length(data_lauki$'Punkti kop')
75
76 par(mfrow = c(1, 4))
77 hist(data_riga$'Punkti kop', probability = TRUE, xlim = c(0, 100), ylim = c(0,
  0.03), main = "Rga", ylab = "Blvums", xlab = "Eksmena punktu sk.")
78 hist(data_pilseta$'Punkti kop', probability = TRUE, xlim = c(0, 100), ylim = c
  (0, 0.03), main = "Valsts pilsta", ylab = "Blvums", xlab = "Eksmena punktu
  sk.")
79 hist(data_nov$'Punkti kop', probability = TRUE, xlim = c(0, 100), ylim = c(0,
  0.03), main = "Novada pilsta", ylab = "Blvums", xlab = "Eksmena punktu sk.")
80 hist(data_lauki$'Punkti kop', probability = TRUE, xlim = c(0, 100), ylim = c(0,
  0.03), main = "Lauki", ylab = "Blvums", xlab = "Eksmena punktu sk.")
81
82 mean(data_riga$'Punkti kop')
83 mean(data_pilseta$'Punkti kop')
84 mean(data_nov$'Punkti kop')
85 mean(data_lauki$'Punkti kop')
86
87 mean(data_riga$'Punkti Zinanas. izpratne un prasmes')
88 mean(data_pilseta$'Punkti Zinanas. izpratne un prasmes')
89 mean(data_nov$'Punkti Zinanas. izpratne un prasmes')
90 mean(data_lauki$'Punkti Zinanas. izpratne un prasmes')
91
92 mean(data_riga$'Punkti Kompleksu problmu risinana')
93 mean(data_pilseta$'Punkti Kompleksu problmu risinana')
94 mean(data_nov$'Punkti Kompleksu problmu risinana')
95 mean(data_lauki$'Punkti Kompleksu problmu risinana')
96
97 par(mfrow = c(1, 1))
98 boxplot(data_riga$'Punkti kop', data_pilseta$'Punkti kop', data_nov$'Punkti kop
  ', data_lauki$'Punkti kop', ylab = "Punktu skaits", names = c("Rga", "
  Valstspils.", "Nov. pils.", "Lauki"))
99
100 outliers <- sum(data_lauki$'Punkti kop' > (1.5 * quantile(data_lauki$'Punkti
  kop', 0.75)))
101

```

```

102 quantile(data_riga$'Punkti kop', c(0, 0.25, 0.5, 0.75, 0.95, 1))
103 quantile(data_pilseta$'Punkti kop', c(0, 0.25, 0.5, 0.75, 0.95, 1))
104 quantile(data_nov$'Punkti kop', c(0, 0.25, 0.5, 0.75, 0.95, 1))
105 quantile(data_lauki$'Punkti kop', c(0, 0.25, 0.5, 0.75, 0.95, 1))
106
107 # eksmens skaits nokrtots no 15 punktiem.
108 nesekmigi <- sum(data2024$'Punkti kop' < 15) / length(data2024$'Punkti kop')
109 nesekmigi_riga <- (sum(data_riga$'Punkti kop' < 15) / length(data_riga$'Punkti
    kop'))
110 nesekmigi_riga
111 nesekmigi_pilseta <- (sum(data_pilseta$'Punkti kop' < 15) / length(data_pilseta
    $'Punkti kop'))
112 nesekmigi_pilseta
113 nesekmigi_nov <- (sum(data_nov$'Punkti kop' < 15) / length(data_nov$'Punkti kop
    '))
114 nesekmigi_nov
115 nesekmigi_lauki <- (sum(data_lauki$'Punkti kop' < 15) / length(data_lauki$'
    Punkti kop'))
116 nesekmigi_lauki
117
118 par(mfrow = c(1, 4))
119 qqnorm(data_riga$'Punkti kop', main = "Kvantu-kvantu grafiks rezulttiem Rg",
    ylab = "Izlases kvantiles", xlab = "Normlsadaljuma kvantiles")
120 qqline(data_riga$'Punkti kop', col = "red", lwd = 3)
121 qqnorm(data_pilseta$'Punkti kop', main = "Kvantu-kvantu grafiks rezulttiem
    valsts pils.", ylab = "Izlases kvantiles", xlab = "Normlsadaljuma kvantiles"
    )
122 qqline(data_pilseta$'Punkti kop', col = "red", lwd = 3)
123 qqnorm(data_nov$'Punkti kop', main = "Kvantu-kvantu grafiks rezulttiem nov.
    pils.", ylab = "Izlases kvantiles", xlab = "Normlsadaljuma kvantiles")
124 qqline(data_nov$'Punkti kop', col = "red", lwd = 3)
125 qqnorm(data_lauki$'Punkti kop', main = "Kvantu-kvantu grafiks rezulttiem
    laukos", ylab = "Izlases kvantiles", xlab = "Normlsadaljuma kvantiles")
126 qqline(data_lauki$'Punkti kop', col = "red", lwd = 3)
127
128
129 wilcox.test(data2024$'Punkti kop', mu = 35.12)
130 wilcox.test(data_riga$'Punkti kop', data_pilseta$'Punkti kop')
131 kruskal.test(data = data2024, 'Punkti kop' ~ 'Urbanizcija')
132
133 library(ggplot2)
134 library(reshape2)
135 library(Hmisc)
136
137 correlationdata <- dplyr::select(data2024, 'Punkti Zinanas. izpratne un prasmes
    ', 'Punkti Kompleksu problmu risinana', 'Punkti kop', 'Tips', 'Urbanizcija')
138 cor_matrix <- cor(correlationdata)
139 cor_data <- melt(cor_matrix)
140
141 ggplot(cor_data, aes(x = Var1, y = Var2, fill = value)) +
142   geom_tile(color = "white") +
143   scale_fill_gradient2(low = "red", mid = "white", high = "darkgreen",
144     midpoint = 0, limit = c(-1, 1), space = "Lab",
145     name = "Koelcija") +
146   theme_minimal() +
147   theme(axis.text.x = element_text(angle = 15, hjust = 1), axis.text = element_

```

```

148   text(size = 15)) +
149   labs(title = "Korelciju matrica", x = "", y = "")
150 library(ggpubr)
151 ggplot(data2024, aes(x = 'Punkti Zinanas. izpratne un prasmes', y = 'Punkti
    Kompleksu problmu risinana')) +
152   geom_point(color = "deepskyblue3", size = 1, alpha = 0.1) +           #
    Scatterplot points
153   geom_smooth(method = "lm", color = "magenta",                        # Regression line
154               se = TRUE, fill = "lightpink") +                        # Confidence interval
155   stat_regline_equation(aes(label = after_stat(eq.label)), color = "black") +
    # Regression equation
156   stat_regline_equation(aes(label = after_stat(rr.label)), color = "black",
157                         vjust = 2.5) +                                # R-squared value
158   labs(x = "1. daa, punkti",
159        y = "2. daa, punkti") +
160   theme_minimal() +
161   xlim(0, 75) +
162   ylim(0, 25)

```