



# Statistinio raštingumo ugdymas

METODINĖ PRIEMONĖ 9-12 KLASIŲ MOKYTOJAMS

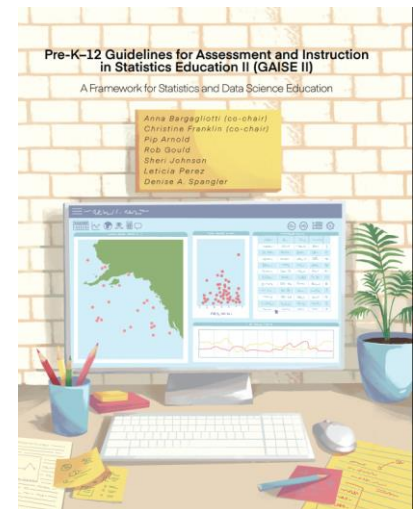
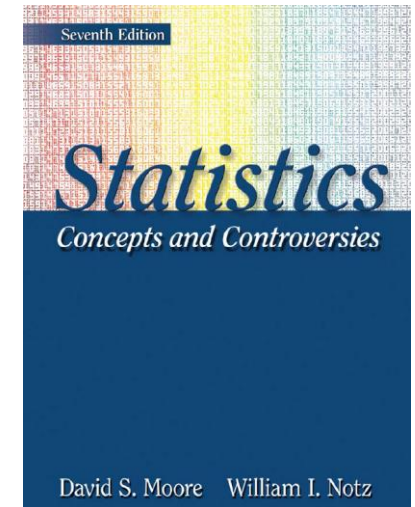
Sandra Aleksienė (*Garliavos Jonučių progimnazija*) ir Alfredas Račkauskas (*VU, MIF*)

2022

Statistika yra apie duomenis. Duomenys yra skaičiai, bet jie nėra „tik skaičiai“. Duomenys yra skaičiai su kontekstu. Pavyzdžiui, skaičius 2 neturi informacijos. Bet, jei išgirstame, kad kažkas loterijoje laimėjo 2 milijonus eurų, su pavydu sveikiname jį. O jei avarijoje žuvo 2 jaunuoliai – apima susirūpinimas. Kontekstas daro skaičius informatyvius. Bet išgauti informaciją iš jų – menas, reikalaujantis daug žinių ir specialiųjų gebėjimų.

Ši metodinė priemonė yra pagalba matematikos mokytojams, kurie rengiasi ugdyti moksleivių statistinį raštingumą pagal atnaujintą Matematikos mokymo programą. Ji rengta atsižvelgiant į programos turinio dalį „Duomenys ir interpretavimas“, skirtą 9, 10 ir 12 klasėms. Dauguma šioje metodinėje priemonėje pateiktų pavyzdžių paimta iš šių šaltinių:

1. David S. Moore and William I. Notz, *Statistics Concepts and Controversies*, 7 edition, H. Freeman and Company, New York, 2009.
2. *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II)*. A Framework for Statistics and Data Science Education. Writing Committee: Anna Bargagliotti (co-chair), Christine Franklin (co-chair), Pip Arnold, Sheri Johnson, Leticia Perez, Denise A. Spangler, Rob Gould. *National Council of Teachers of Mathematics*, 2020.



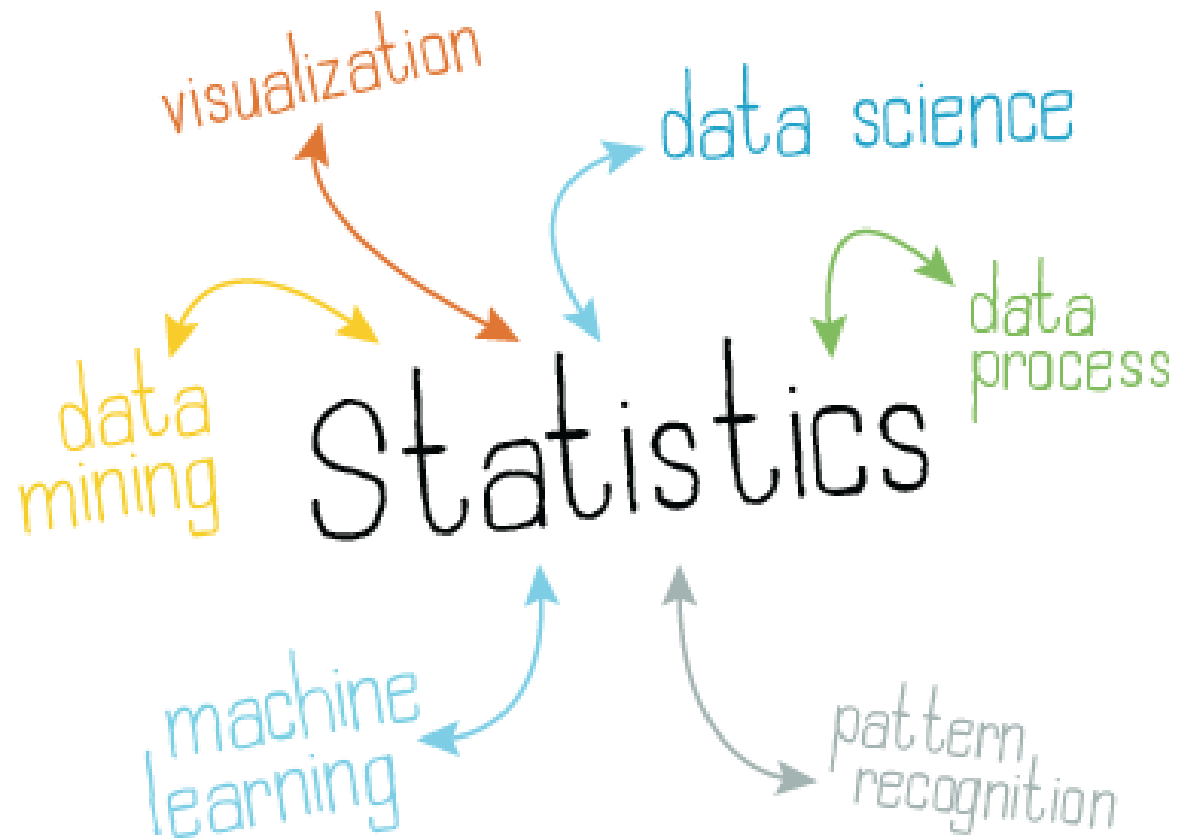
Terminas „Statistika” turi kelias prasmes.

**Statistika** - tai mokslas apie mokymąsi iš duomenų, informacijos rinkimą, sistemimą, analizavimą ir interpretavimą. Ji vysto duomenų rinkimo ir jų analizavimo metodus.

**Statistika** – tai matematinių metodų ir įrankių rinkinys, leidžiantis atsakyti į svarbius klausimus apie duomenis. Ji skirstomas į dvi kategorijas

- *Aprašomoji statistika* - siūlo metodus, kaip apibendrinti duomenis, paverčiant neapdorotus stebėjimus prasminga informacija, kurią lengva interpretuoti ir dalytis.
- *Išvadinė statistika* - siūlo metodus, kaip iš imties duomenų padaryti išvadas visai populiacijai.

**Statistika** - tai informacija apie veiklą ar procesą, išreikšta skaičiais, lentelėmis ar diagramomis.



Kraujo grupės



Nauja kometa

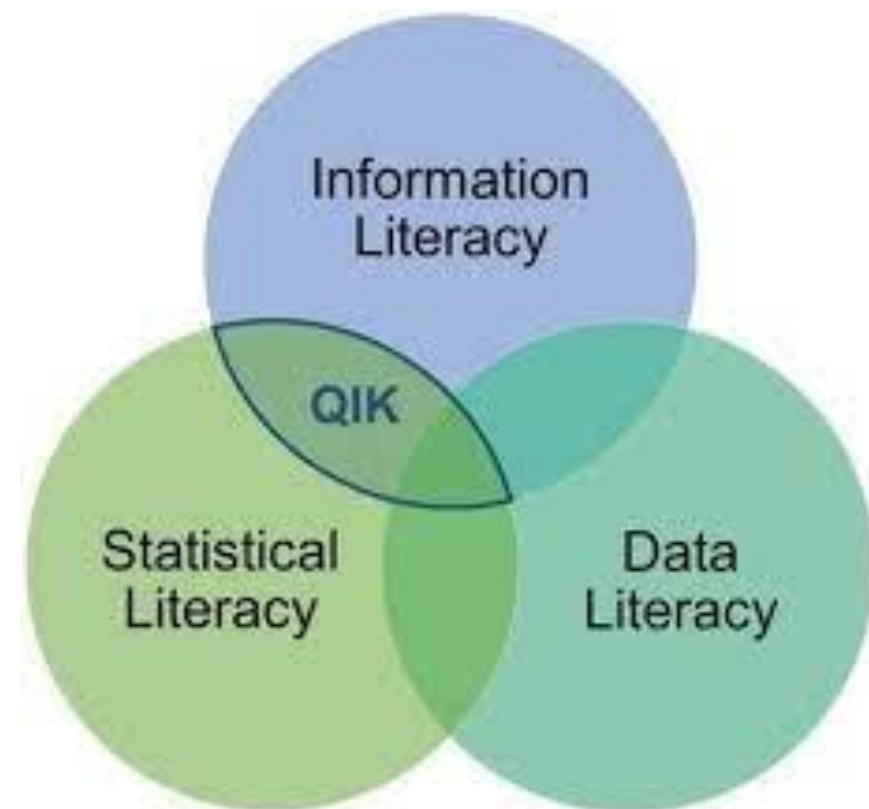
## MATEMATIKOS BENDROJI PROGRAMA

### II SKYRIUS TIKSLAS IR UŽDAVINIAI

**7. Matematikos dalyko tikslas** – sudaryti galimybę kiekvienam mokiniui, mokantis matematikos, *ugdytis* matematinę ir *statistinį raštingumą*, kuris šiame dokumente suprantamas kaip įgytas gebėjimas matematiškai samprotauti ir taikyti įgytas kompetencijas, sprendžiant įvairias realias, aktualias ir mokiniams suprantamas problemas.

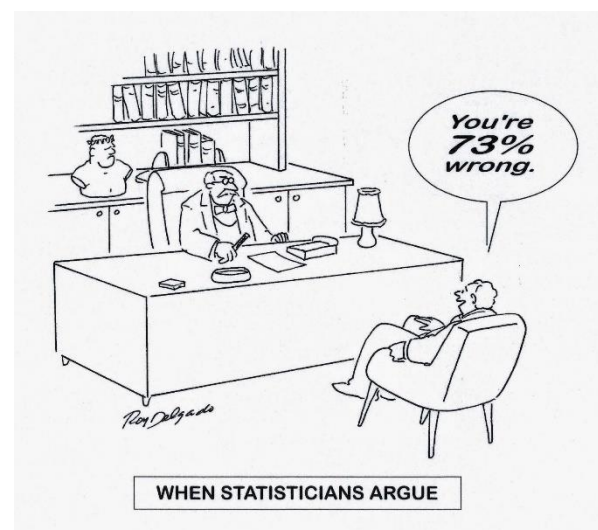


**Statistinis raštingumas** - pagrindinės statistikos kalbos bei pagrindinių statistikos idėjų supratimas (pvz., žinojimas, ką reiškia statistiniai terminai ir simboliai bei *mokėjimas skaityti statistinę informaciją*).



**Statistinis mąstymas** – tai mąstymo tipas, kurį statistikai naudoja aiškindamiesi ar spręsdami statistines problemas.

**Statistinis argumentavimas** - būdas, kaip žmonės argumentuoja statistinėmis idėjomis ir įprasmina statistinę informaciją. Tai apima argumentus, pagrįstus duomenų rinkiniais, grafiniais vaizdais ir statistinėmis suvestinėmis.



# Statistinio raštingumo ugdymas

Konceptuali statistinio raštingumo ugdymo struktūra turi dvi dalis:

- Statistinis problemų sprendimo procesas.
- Pasiekimų lygmenys.

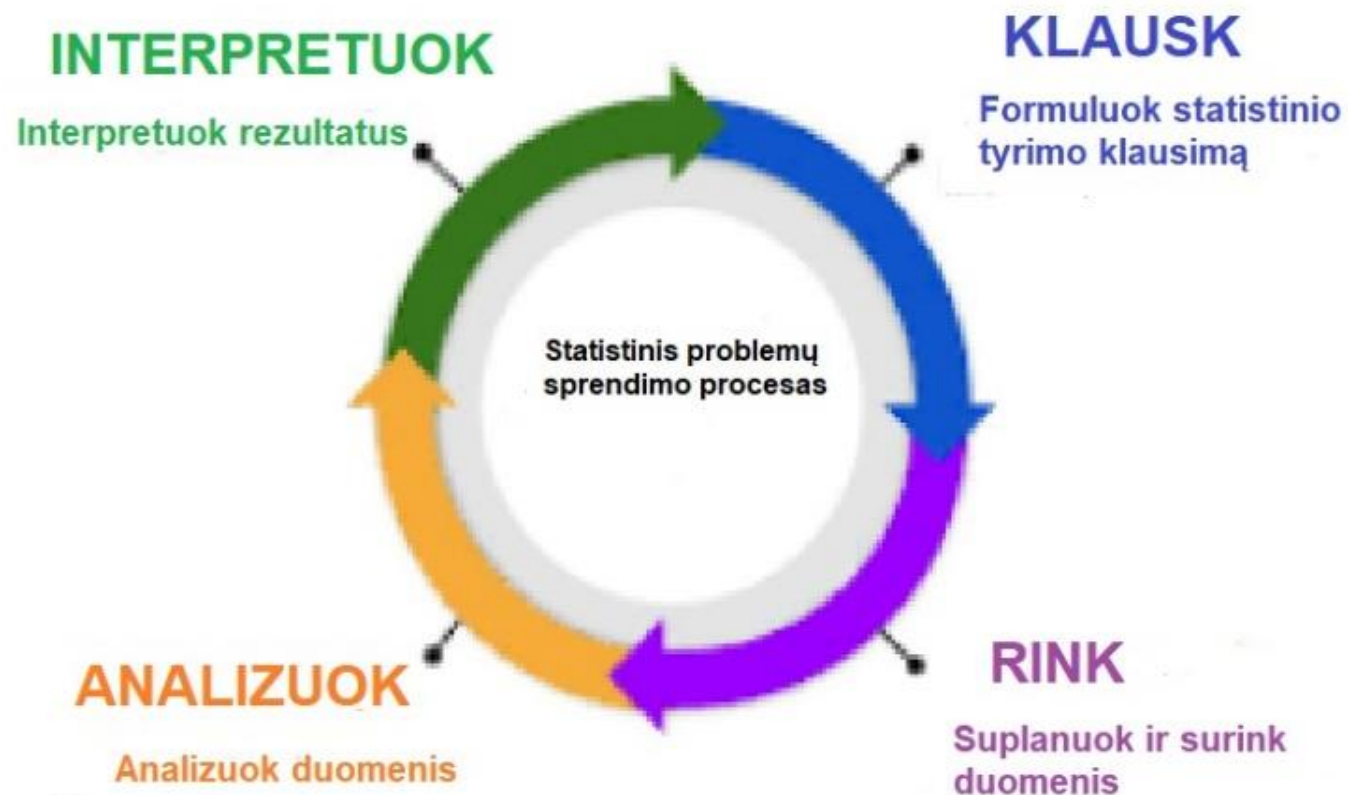


Problemomis grįstas mokymasis

Projektai grįstas mokymasis



# *KRAI: Klausk - Rink - Analizuok - Interpretuok*



# Statistinio raštingumo lygiai **A**, **B**, **C**

- (A) lygyje mokymasis yra labiau nukreiptas į mokytoją (labai sąlyginai, tikėtina, kad pasiekiamas 1-7(8) klasėse);
- (B) lygyje pereinama prie į moksleivį orientuoto darbo (labai sąlyginai, tikėtina, kad pasiekiamas 8-10 klasėse);
- (C) lygyje jis tampa labai orientuotas į moksleivį (taikomas labiau išplėstiniam kursams, arba 11-12 klasėms).

**! Šie trys lygiai pagrįsti statistikos mąstymo *raida*, o ne amžiumi !**

### Duomenys ir interpretavimas

Nagrinėjamos taškinės (sklaidos) diagramos, vaizduojančios statistinį ryšį tarp dviejų kintamųjų (stebimų požymių) reikšmių. Mokomasi iš sklaidos diagramos įvertinti šio ryšio buvimą ar nebuvimą, aptariama, kokiais atvejais kalbama apie kintamųjų koreliacinį ryšį. Detaliau aptariama tiesinė koreliacija. Mokomasi užrašyti sklaidos **diagramoje pavaizduotos tiesės lygtį**  $y = kx + b$ , koeficiento  $k$  reikšmę, interpretuoti šia lygtimi aprašomą duomenų ryšį. Aptariama, kodėl negalime daryti išvados apie tiesinės priklausomybės egzistavimą populiacijoje, jei **duomenys imtyje yra neatsitiktiniai** ar jų yra per mažai.

Šio etapo programai įgyvendinti reikės

šių sąvokų bei analizės įrankių:

- *statistinis kintamasis ir jų tipai;*
- *sklaidos diagrama;*
- *koreliacijos koeficientas;*
- *tiesinis sąryšis.*

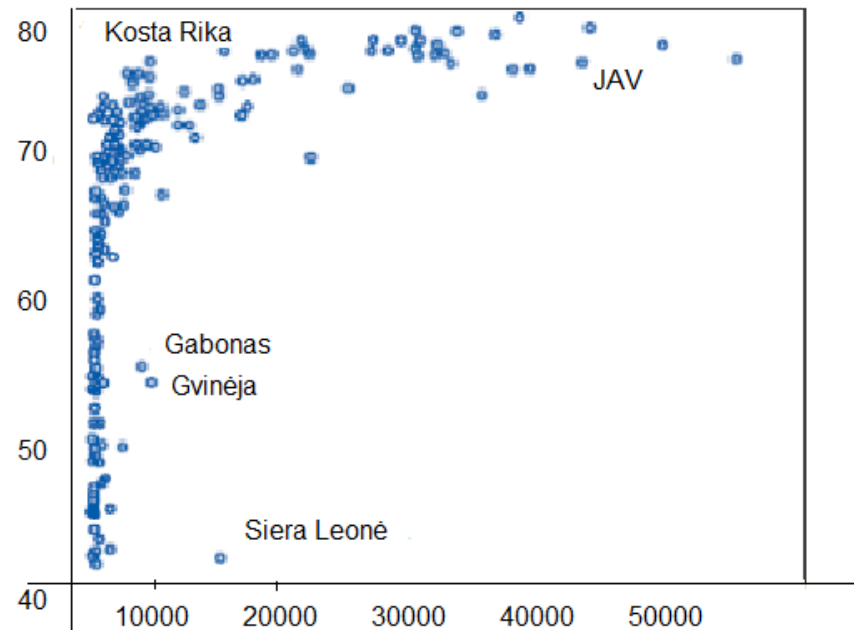
Siūlomas turinys.

1. Statistinių kintamųjų tipai.
2. Sklaidos diagrama ir jos interpretavimas.
3. Koreliacijos koeficiento samprata.
4. Tiesinis dviejų kintamųjų sąryšis.

## 1 Pavyzdys. Sveikata ir turtas

1 sklaidos diagramoje yra Pasaulio banko duomenų sklaidos diagrama. Individai yra visos pasaulio tautos, apie kurias yra prieinami duomenys. Aiškinantysis kintamasis yra šalies turtingumo matas: bendrasis vidaus produktas (BVP) tenkantis vienam gyventojui. BVP yra bendra šalyje pagamintų prekių ir suteiktų paslaugų vertė, konvertuota į šalies valiutą. Atsako kintamasis yra vidutinė gyvenimo trukmė metais.

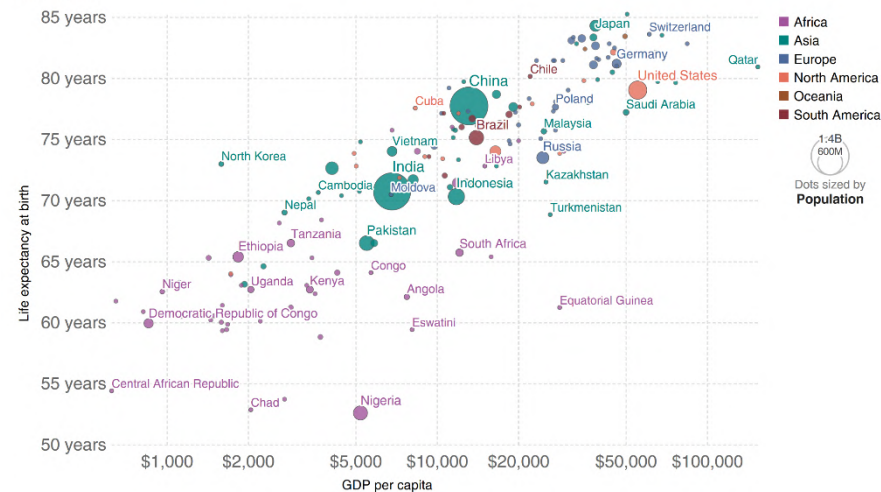
**Vilniaus  
universitetas**



Tikėtina, kad turtingesnėse šalyse žmonės gyvena ilgiau. Iš esmės, sklaidos diagrama tai ir rodo, tačiau sąsajos turi įdomią formą. Kai BVP didėja, vidutinė gyvenimo trukmė iš pradžių ilgėja labai greitai, bet vėliau greitis lėtėja arba augimas visai sustoja.

Life expectancy vs. GDP per capita, 2018

GDP per capita is measured in 2011 international dollars, which corrects for inflation and cross-country price differences.



**2 pratimas. Širdies ritmas.** Mokytojas Marius reguliariai nuplaukia 2000 metrų, bandydamas gerinti savo sportinę formą. Štai jo laikai (minutėmis) ir pulso dažnis (tvinksniai per minutę) po 23 plaukimo seansų:

<b>Laikas</b>	<b>34.12</b>	<b>35.72</b>	<b>34.72</b>	<b>34.05</b>	<b>34.13</b>	<b>35.72</b>	<b>36.17</b>	<b>35.57</b>
<b>Pulsas</b>	152	124	140	152	146	128	136	144
<b>Laikas</b>	35.37	35.57	35.43	36.05	34.85	34.70	34.75	33.93
<b>Pulsas</b>	148	144	136	124	148	144	140	156
<b>Laikas</b>	34.60	34.00	34.35	35.62	35.68	35.28	35.97	
<b>Pulsas</b>	136	148	148	132	124	132	139	

- Padarykite sklaidos diagramą. (Kas yra aiškinamasis kintamasis?)
- Ar sąsaja tarp šių kintamųjų yra teigiama ar neigiama? Paaiškinkite, kodėl tikite, kad sąsaja yra tokios krypties.
- Apibūdinkite sąsajos formą ir stiprumą.
- Skaičiuotuvu suraskite koreliacijos koeficientą  $r$ . Žvelgdami į sklaidos diagramą, paaiškinkite, kodėl ši  $r$  reikšmė yra pagrįsta.
- Tarkime, kad laikas buvo įrašytas sekundėmis. Pavyzdžiui, laikas 34,12 minutės būtų 2047 sekundės. Kaip pasikeistų  $r$  reikšmė?

# *Mokiniams: dainos trukmė ir MP3 failo dydis.*

Vilniaus  
universitetas

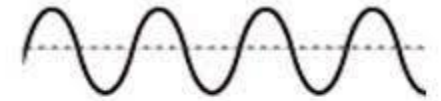
## *Statistinio klausimo formulavimas.*

Ar manote, kad yra ryšys tarp dainos trukmės minutėmis ir MP3 failo dydžio?

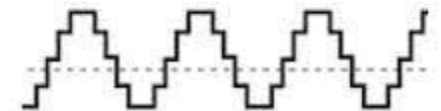
Galimi ir kitokie klausimai, pavyzdžiui:

- Kiek tikėtina dainų galite įsirašyti įrenginyje turinčiame 1 GB atminties?
- Ar kantri, ar roko dainos reikalauja daugiau atminties? Ir pan.

WAV/FLAC



MP3



Galima susieti su matematikos nauda skaitmeninant muziką ir ne tik!

## *Duomenų rinkimas, tvarkymas ir tikrinimas.*

Suplanuoti, kaip surinkti duomenis.

Kiekvienas mokinys gali duomenis susirinkti iš savo mobiliojo telefono ar kompiuterio.

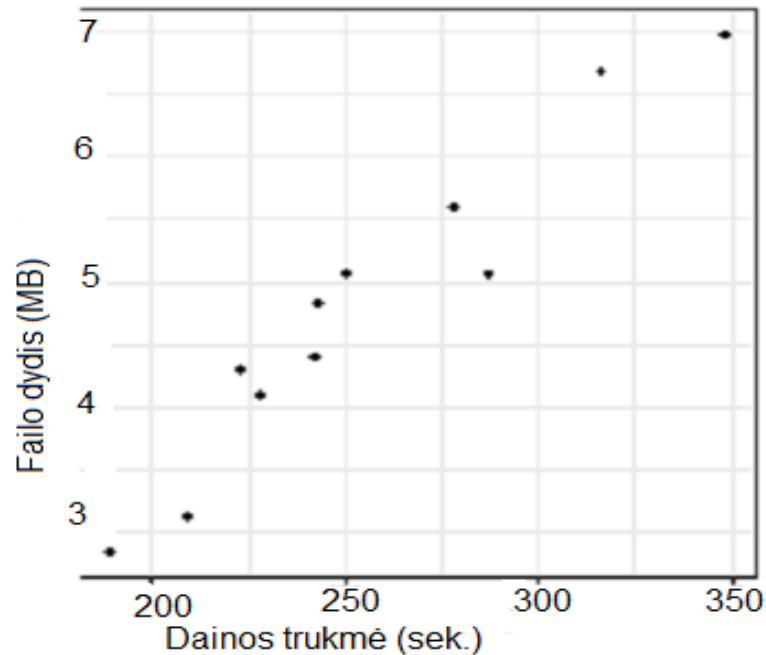
Jei mokytojas turi surinktų ar iš interneto ištrauktų duomenų, galima pradėti darbą su jais.



# Duomenų tyrinėjimas.

Vilniaus  
universitetas

Verta paaiškinti kas yra sklaidos diagrama, kam ir kaip ji naudojama.



Paklauskite, kokias mokiniai įžiūri tendencijas šioje sklaidos diagramoje?

**Galimas atsakymas:** Ilgesnė daina (sekundėmis) linkusi užimti daugiau atminties (MB).

Kas atsitinka su failo dydžiu, kai daina ilgėja?

**Galimas atsakymas:** Failo apimtis didėja.

Paaiškinkite, kad šioje sklaidos diagramoje dainos trukmė yra *nepriklausomas kintamasis* arba *aiškinantysis kintamasis*, o failo dydis – *priklausomas kintamasis* arba *atsako kintamasis*. Ko gero teks priminti ir sąvoką „statistinis kintamasis“ bei pakabėti apie jų tipus.

Kaip apibūdinsite sąsają tarp dainos ilgio ir failo apimties?

**Galimas atsakymas:** sąsaja yra teigiama.

Koks matematinis modelis tiktų ryšiui tarp dainos failo apimties ir dainos trukmės apibūdinimui?

**Galimas atsakymas:** tiesinis modelis turi prasmę, nes sklaidos diagramoje taškai yra glaudžiai išsidėstę apie tiesią liniją.



# *Rezultatų interpretavimas*

Grįžkite prie statistinio klausimo. Ar gavote atsakymą? Kokį atsakymą gavote? Paprašykite atliktą tyrimą pristatyti žodžiu. Patarkite, kaip tai galima padaryti.

Atlikę šią užduotį, moksleiviai sužino visas sąvokas ir instrumentus, kurie numatyti 9 klasės programoje. T.y.,

*statistinis kintamasis ir jų tipai;*

*sklaidos diagrama;*

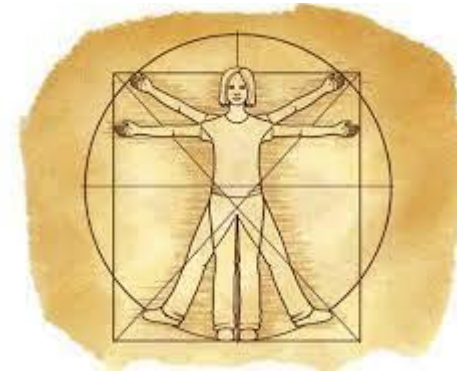
*statistinių kintamųjų sąryšiai;*

*tiesinė koreliacija.*

## 2 projektas. Ryšys tarp žmogaus ūgio ir rankų ilgio.

Galimi statistinio tyrimo klausimai:

- Ar žmogaus ūgis ir rankų ilgis susiję?
- Kiek stiprus yra ryšys tarp rankų ilgio ir ūgio?



**Vilniaus  
universitetas**

**Pastaba.** Rankų ilgį galima keisti kitu parametru, pavyzdžiui, batų dydžiu, galvos apimtimi ir pan.

## 3 projektas. Reikia ištirti ar matematika yra raktas į sėkmę?

Tai padėtų padaryti gauti atsakymai į šiuos klausimus.

- (a) Ar yra ryšys tarp matematikos ir kitų dalykų pasiekimų? Galima analizuoti vieną klasę, kelias klases arba visą mokyklą. Sąsajų ieškoti galima tarp matematikos pasiekimų ir kurio nors kito dalyko (užsienio kalbos, fizikos, chemijos ir kt.) arba kitų dalykų pasiekimų.
- (b) Kokie kintamieji galėtų paaiškinti ryšį tarp matematikos pasiekimų vidurinėje mokykloje ir sėkmės stojant į universitetus/kolegijas?

Problema gali būti duomenų surinkimas. Tam gali prireikti pagalbos iš mokyklos administracijos.



#### 4 užduotis. *Bebrai ir vabalai.*

Ekologai mūsų aplinkoje kartais randa gana keistų santykių. Vienas jų tyrimas rodo, kad bebrai naudingi vabalams. Tyrėjai parinko 23 apskritus sklypus, kurių kiekvienas buvo 4 metrų skersmens, tose vietose, kur bebrai kirto medvilnės medžius.

Kiekviename sklype jie skaičiavo bebrų nukirstų medžių kelmus ir vabalų lervų sankaupas.



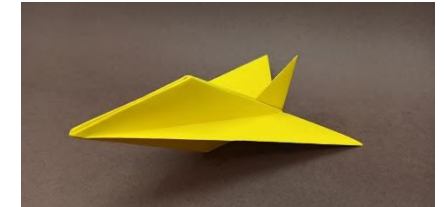
<b>Kelmai</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>3</b>	<b>3</b>	<b>4</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>5</b>	<b>1</b>	<b>3</b>
<b>Lervų sankaupos</b>	10	30	12	24	36	40	43	11	27	56	18	40
<b>Kelmai</b>	2	1	2	2	1	1	4	1	2	1	4	
<b>Lervų sankaupos</b>	25	8	21	14	16	6	54	9	13	14	50	

- a) Padarykite sklaidos diagramą, kuri parodytų, kiek bebrų padarytų kelmų skaičius turi įtakos vabalų lervų sankaupų skaičiui. Ką rodo diagrama? (Ekologai mano, kad nauji kelmų daigai yra švelnesni ir vabalams labiau patinka.)
- b) Regresijos linija yra  $lervų\ sankaupos = -1,286 + (11,894 \times kelmai)$ . Nubrėžkite šią liniją savo sklaidos diagramoje. (Norėdami nubrėžti liniją, naudokite lygtį  $y = -1.286 + 11.894x$ , ir raskite  $y$  imdami  $x = 1$  ir  $x = 5$ . Per surastus du  $(x, y)$  taškus nubrėžkite tiesią liniją.)
- c) Koreliacija tarp šių kintamųjų yra  $r = 0,916$ .
- d) Kiek procentų stebimo vabalų lervų sankaupų skaičiaus kitimo galima paaiškinti tiesinio modelio pagalba?
- e) Ar, remdamasis savo pastebėjimais (a), (b) ir (c) punktuose, manote, kad kelmų skaičiavimas yra patikimas būdas numatyti vabalų lervų sankaupas? Atsakymą pagrįskite.

# Kiti projektai

## Sudėtingesnis duomenų surinkimas

Ištirti sąsajas tarp skirtingų popierinių lėktuvėlių konstrukcijos ir įveikto atstumo bei skrydžio laiko.



Kiek trunka keleivių įlaipinimas?

Sąsajos tarp įlaipinimo į troleibusą ir autobusą.



# 10 ir II gimnazijos klasė

Vilniaus  
universitetas

Paaiškinama, kaip imties iš populiacijos sudarymas susijęs su **pagrįstų išvadų** darymu, ką vadiname duomenų rinkinių kintamumu, duomenų pasiskirstymu, kaip galima apibūdinti ir kiekybiškai interpretuoti duomenų rinkinius. Aptariamos sąvokos: dispersija, standartinis nuokrypis, skirstinys, **normalusis skirstinys**, simetriškasis skirstinys, asimetriškasis skirstinys. Nagrinėjant realaus gyvenimo konteksto pavyzdžius, **diskutuojama apie duomenų rinkimą ir analizavimą**. Svarstoma, kokias išvadas apie duomenis leidžia daryti jų pasiskirstymą aproksimuojančios kreivės forma ar apskaičiuotos duomenų centro (pavyzdžiui, vidurkio) ir sklaidos (pavyzdžiui, standartinio nuokrypio, kvartilių) charakteristikos. Analizuojamas statistinis patikimumas.

Šį turinį geriau išdėstyti prieš dviejų statistinių kintamųjų analizę, pvz., 9'oje klasėje. Taigi 9 klasės turinį reikėtų sukeisti vietomis su 10 klase.

- *Populiacija ir imtis*
- *Duomenų pasiskirstymas*
- *Duomenų pasiskirstymo parametrai*
- *Statistinis patikimumas*

Siūlome tokį turinį.

1. Imtys: geros ir blogos.
2. Duomenų pasiskirstymas: histograma ir medžio diagrama.
3. Duomenų pasiskirstymo parametrai: vidurkis, dispersija, standartinis nuokrypis.
4. Statistinis patikimumas.

## Problema. *Ar muzikos klausymas trukdo mokytis?*

**Vilniaus  
universitetas**

Įprasta, kad besimokydami mokiniai klausosi muzikos. Kai kurie mokiniai teigia, kad muzikos klausymas jiems padeda susikaupti ir nuteikia teigiamai. **Ar tikrai?** Reikia šią problemą ištirti. Tuo tikslu moksleiviai gali kurti eksperimentinį tyrimą, skirtą ištirti muzikos klausymosi poveikį žmogaus gebėjimui įsiminti žodžius.

*Statistinio klausimo formulavimas.*

**Ar klausydami muzikos mokiniai sugeba įsiminti daugiau žodžių nei nesiklausydami?**



shutterstock.com - 81372991

## *Duomenų surinkimas.*

Klasė turi sukurti eksperimentinių duomenų rinkimo strategiją, kad būtų galima atsakyti į tiriamąjį klausimą. Reikės, kiek įmanoma, pašalinti galimus pašalinius kintamumo šaltinius, kurie gali trukdyti interpretuoti rezultatus.

Vienas paprastas eksperimentas, kurį galima atlikti su klase: atsitiktinai padalyti klasę į dvi vienodo dydžio grupes.

Taip moksleiviams ugdomas supratimas, kad atsitiktinis priskyrimas yra svarbi eksperimentinio plano dalis, nes ji linkusi mokinių gebėjimus atskirti nuo kitų savybių, kurie gali paveikti atsakymus. Kaip tai padaryti?

Pavyzdžiui, mokytojas gali sumaišyti 28 kortas, 14 raudonų ir 14 juodų kortelių ir išdalinti po vieną kiekvienam mokiniui. Gavę raudoną kortelę dalyvaus muzikos klausytojų grupėje, o gavę juodą – tylos grupėje.



Klasė turėtų sukurti eksperimentinę procedūrą ir susitarti dėl jos detalių. Pavyzdžiui, kiekvienas dalyvis gali turėti dvi minutes 20 žodžių sąrašui išnagrinėti, po to daryti vienos minutės pauzę, o tada – per dvi minutes užrašyti kuo daugiau žodžių iš duoto sąrašo.

Viso eksperimento metu vienos grupės dalyviai klausysis tos pačios dainos, o kontrolinė grupė visą laiką tylės.

Teisingai įsimenamų žodžių skaičius yra dominantis statistinis kintamasis.

## *Duomenų analizavimas.*

Mokiniai turėtų apskaičiuoti suvestinę statistiką muzikos klausymo grupei ir tylos grupei.

Stačiakampės diagramos suteiks daug informacijos tiriamam klausimui atsakyti.

Jei duomenų yra daug, moksleiviai gali naudoti technologijas, svarbu žinoti, kaip jas naudoti.



*Rezultatų interpretavimas. Koks yra atsakymas į iškeltą klausimą?*

Mokiniam gali kilti klausimas, ar muzikos tipas buvo svarbus veiksnys rezultatui. Pavyzdžiui, ar žodžių buvimas dainoje prisideda prie skirtumų tarp dviejų grupių?

Galima kartoti eksperimentą su nauja moksleivių grupe, naudojant tik instrumentinę muziką.

Moksleiviai turėtų suprasti šio eksperimento išvadų apimtį. Atranka nebuvo atsitiktinė, jie buvo atsitiktinai priskirti tyrimo grupėms. Tai riboja išvadų apimtį klasei, tačiau ne bendrajai populiacijai. Eksperimentas gali būti plėtojamas kitoje klasėje.

## 2 projektas. Kiek trunka 1 minutė?

Šį projektą galima siūlyti jau nuo 5 klasės. Duomenims surinkti reikės chronometro. Imtis gali būti iš vienos klasės ar kelių klasių. Galima kelti, pavyzdžiui, tokius statistinius klausimus.

- (a) Koks yra klasės mokinių laiko suvokimas?
- (b) Ar vienos ir kitos klasės mokinių laiko suvokimas yra skirtingas?
- (c) Ar skiriasi mergaičių ir berniukų laiko suvokimas?



## 2 užduotis. Rašymo stiliaus statistika.

Skaitiniais duomenimis galima išskirti skirtingus rašymo tipus, o kartais net atskirus autorius. Štai moksleivių surinkti duomenys apie 1–15 raidžių žodžių procentą, žurnalo „Popular Science“ straipsniuose:

Žodžio ilgis	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Procentas	3.6	14.8	18.7	16.0	12.5	8.2	8.1	5.9	4.4	3.6	2.1	0.9	0.6	0.4	0.2

Statistinis kintamas: raidžių skaičius žodyje, žurnalo „Popular Science“ straipsniuose.

- Padarykite šio statistinio kintamojo pasiskirsymo histogramą. Apibūdinkite jo formą, centrą ir išplitimą.
- Raskite penkių skaičių suvestinę.
- Analogišką užduotį atlikite su Lietuviškais leidiniais (tai gali tikti ir projektiniam darbui).

# IV gimnazijos klasė. Bendrasis kursas

Vilniaus  
universitetas

Įvadas į taikomąją duomenų analizę. **Nagrinėdami straipsnius apie mokslo pasiekimus**, statistikos ir technologijų vaidmenį šiuolaikiniame pasaulyje, mokiniai sužino, kad funkcijos gali būti naudojamos ir duomenims apibūdinti, o jei duomenys susiję tiesiniu ryšiu, tai tas ryšys gali būti modeliuojamas tiese (regresijos tiese), o jo stiprumas ir kryptis išreikšti koreliacijos koeficientu. Visas naujas sąvokas mokiniai išsiaiškina, nagrinėdami konkrečius pavyzdžius, o reikiamai skaitinei informacijai gauti pasitelkia skaitmenines technologijas. Mokiniai išsiaiškina, kad statistinės analizės (regresinė analizė yra viena iš jos dalių) tikslas – ištyrus dalį respondentų (imtį), padaryti išvadą apie visą populiaciją. Aptariamos kintamojo, kintamojo matavimo skalių, duomenų sąvokos ir išsiaiškinama, kodėl tik **intervaliniams duomenims** taikomos vidurio, standartinio nuokrypio, koreliacijos (tiesinės koreliacijos koeficiento) skaičiavimo procedūros. Mokoma(si) praktiškai, naudojantis skaitmeninėmis technologijomis, apskaičiuoti duomenų rinkinio imties vidurkį, standartinį nuokrypį, interpretuoti, kaip jie charakterizuoja imtį. Nagrinėjami pavyzdžiai, kai sprendimui dėl kintamųjų ryšio ir jo stiprumo priimti naudojama koreliacija (pavyzdžiui, laiko ir pažymių, amžiaus ir atlyginimo, IQ ir darbo kompiuteriu). Atkreipiamas dėmesys, kad koreliacija nepaaiškina priežastingumo. Paaiškinama, kad **priežastingumui tarp kintamųjų nustatyti taikomas kitas matematinis modelis – tiesinė regresija**. Nagrinėjamos paprasčiausios tipinės situacijos, kai gali būti taikoma tiesinė regresija (pavyzdžiui, ar per egzaminą surinktų balų skaičius priklauso nuo socialinio statuso). Išsiaiškinama, kaip priimamas sprendimas, kuris kintamasis vadinamas priklausomu kintamuoju, o kuris – aiškinamuoju (regresoriumi). **Naudojantis skaičiuoklės programa**, demonstruojama, kaip atrodo grafinis duomenų rinkinio vaizdas („taškų debesėlis“). Nagrinėjama problema – ar įmanoma šiuos duomenis aprašyti modeliu (tiese). Išsiaiškinama, kad svarbiausia šio modelio (tiesės) charakteristika – determinacijos koeficientas (R kvadratas) ir mokomasi, jį žinant (suradus), priimti sprendimą dėl gauto modelio tinkamumo duomenims aprašyti.

## Kelios pastabos programos turiniui.

*„Paaiškinama, kad priežastingumui tarp kintamųjų nustatyti taikomas kitas matematinis modelis – tiesinė regresija.“* Šis teiginys yra klaidingas. Tiesinė regresija nepaaiškina priežastingumo. Yra daug taip vadinamų „klaidingos regresijos“ pavyzdžių. Klaidinga regresija yra statistinis modelis, rodantis klaidinančius statistinius tiesinio priežastinio ryšio įrodymus. Vienas iš daugybės pavyzdžių:

1. Metinis JAV eksporto indeksas (y) (1960-1990 metų duomenimis) ir Australijos vyrų gyvenimo trukmė (x): nustatytas tiesinis modelis

$$y = -2943. + 45.7974 x, \quad \text{koreliacijos koeficientas } r=0.9570 \text{ ir } R \text{ kvadratas} = 0.916.$$

1. JAV išlaidos gynybai (y), 1971-1990 metais ir PAR gyventojų skaičius: nustatytas tiesinis modelis

$$y = -368.99 + .0179 X, \quad R \text{ kvadratas} = .940, \quad r = .9694.$$

Intuityviai aišku, kad kartu didėjantys arba kartu mažėjantys kintamieji gali būti aprašyti tiesinės regresijos modeliu, tačiau nei vienas iš jų nebus priežastimi kito.

*„Išsiaiškinama, kaip priimamas sprendimas, kuris kintamasis vadinamas priklausomu kintamuoju, o kuris – aiškinamuoju (regresoriumi)“.*

Priežastingumui nustatyti dažniausiai naudojamas 2003 metų Nobelio premijos laureato Granger priežastingumo testas. Tačiau tai, mūsų nuomone, yra už mokyklinės programos ribų. Todėl lieka ekspertinis kintamųjų paskirstymas į priklausomus ir aiškinančiuosius.

# Problema.

**Vilniaus  
universitetas**

Švietimo administracija ir mokytojai yra susirūpinę dėl pastebimo mokinių miego trūkumo. Mokiniai klasėje atrodo labiau nei bet kada pavargę, sunkiai gali susikaupti ar ilgesnį laiką išlaikyti dėmesį. Dėl to sulaukiama tėvų skundų, kad mokiniams užduodama per daug namų darbų.

Tačiau manoma, kad yra keletas galimų veiksnių, kurie prisideda prie miego trūkumo ir tie veiksniai nėra akademiniai. Jie apima mokinių užklausinės veiklos, kurioje jie aktyviai dalyvauja, skaičių ir jiems skiriamą laiką bei laiką praleistą internete ar prie elektroninių prietaisų.

Nusprendžiama atlikti atrinktų mokyklų mokinių apklausą vienoje ar keliose savivaldybėse ir nustatyti galimus veiksnius, lemiančius vidurinės mokyklos mokinių miego trūkumą.





Tai labai svarbi problema, reikalaujanti rimto mokslininkų dėmesio. Prie jos sprendimo gali prisidėti ir patys moksleiviai, keldami paprastus, su šia problema susijusius klausimus.

Pavyzdžiui:

Kiek būrelių lanko 9-12 klasių mokiniai?

Kiek laiko užklasinėms veikloms (per savaitę; per mėnesį) skiria 9-12 klasių mokiniai?

Kiek valandų per dieną po pamokų skiriama internetui?

Kiek laiko per dieną praleidžiama žaidžiant kompiuterinius žaidimus?

Kiek laiko skiriama (matematikos) namų darbams?

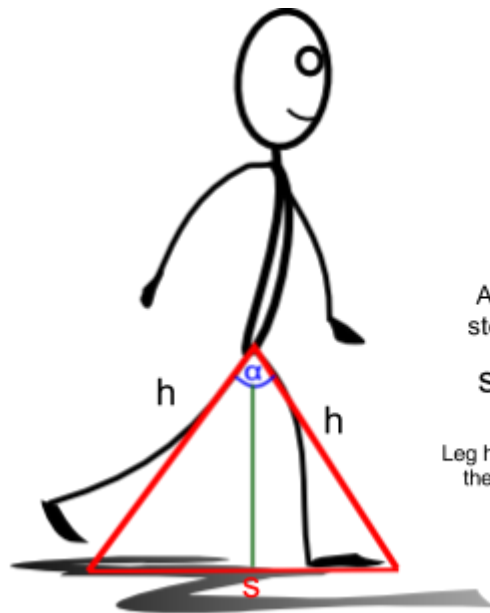
### 3 projektas. Kojų ilgis ir žingsnių ilgis.

Padalinkite klasę į grupes po 3 arba 4 mokinius.

Tegu grupės išmatuoja kiekvieno mokinio kojos ilgį nuo išorinio klubo kaulo iki grindų.

Tuo pačiu metu leiskite kiekvienam mokiniui eiti tam tikru maršrutu. Tas pats maršrutas kiekvienam mokiniui. Mokiniai tyliai suskaičiuoja žingsnių skaičių, eidami šiuo maršrutu.

Paprašykite mokinių įvesti duomenis į paprastą duomenų rinkimo formą.



Approximation for  
step length  $s < 1\text{m}$ :

$$s = 2 h \sin \frac{\alpha}{2}$$

Leg height  $h$  (from hip joint to  
the ground) and angle  $\alpha$   
of the swing



#### 4 uždutis.

Atlikdamas naminių bičių tyrimą, Seeley (2010) pastebėjo, kad bitės skautės šoka, kad praneštų bitėms, grįžusioms į lizdą, apie atstumą iki naujos lizdo vietos. Žemiau esančioje lentelėje parodytas septynių skirtingų bičių šokio trukmė (sekundėmis) ir atstumas iki naujos vietos (metrais).

Šokio trukmė (sec.)	0.40	0.45	0.95	1.30	2.00	3.10	4.10
Atstumas (metrai)	200	250	500	950	1950	3500	4300

Išanalizuokite duomenis . Panaudokite tiesinį modelį ir nuspėkite atstumą iki naujo lizdo, kai bitė šoka 1.5 sekundės.





**Vilnius  
University**

---

**[alfredas.rackauskas@mif.vu.lt](mailto:alfredas.rackauskas@mif.vu.lt)**

VU. MIF, Statistinės analizės katedra

Naugarduko 24

Vilnius