Informācijas dizains un vizualizācijas

**Anotācija:** Datu mācība, spēja uzskatāmi apkopot ļoti lielu informācijas apjomu, ļauj skolēniem orientēties tajās jomās, ko nevar tieši sajust ar redzi, dzirdi vai ožu. Tā ir demogrāfija, klimata izmaiņas, tendences ekonomikā un darba tirgū, informācijas lietošana (un arī publiskošana) sociālajos tīklos - šīs prasmes jau šobrīd ir vajadzīgas atbildīgiem sabiedrības locekļiem un to loma varētu vēl pieaugt.

## Ievads: Kas ir datu mācība?

Kādēļ vajadzīgs atsevišķs kurss par informāciju un vizualizācijām? Par datiem interesējas vairākas savstarpēji saistītas zinātnes, kuras jau tagad māca skolā. Visu, kas skolā būtu jāzina par datiem, nevar saukt par vienotu zinātni (ar tai raksturīgām pētīšanas metodēm un pieejām), bet drīzāk par savstarpēji saistītu zināšanu kopumu, ko angliski sauc par “data science”, bet latviski - par datu mācību. (Līdzīgi kā teātris nav pilnīgi atsevišķa mākslas nozare, bet gan “sintētiska māksla”, kas sintezē dramaturģiju, aktiermākslu, scenogrāfiju, deju, mūziku, gaismošanu, u.c.). Par datiem ir dažādi skolas priekšmeti:

* **Matemātika** aplūko operācijas, ko var veikt ar datiem (aritmētiskās darbības, ģeometriskie pārveidojumi, funkcijas, u.c.)
* **Dabaszinātnes** interesējas par datu iegūšanu mērījumu ceļā, mērvienībām un mērījumu kļūdām.
* **Sociālās zinātnes** precīzas atziņas balsta uz sabiedriskās domas mērījumiem, aptaujām un marketinga pētījumiem.
* **Valodniecība** apkopo datus, kas nepieciešami automatizētajai tulkošanai, kvalitatīvu vārdnīcu sastādīšanai (korpuslingvistika) un citiem komunikācijas mērķiem (pilna teksta meklētāji, u.c.).
* **Statistika,** ko parasti māca matemātikas kursa ietvaros, bet kuras lietojumi un metodes iziet tālu ārpus matemātikas, aplūko secinājumus, kas iegūstami no liela novērotu datu daudzuma.
* **Informātikā** līdzās citām tēmām apgūst informācijas strukturēšanu ar biroja programmām (teiksim, Microsoft Office vai LibreOffice). Izklājlapas (spreadsheets) nereti kalpo datu sākotnējai ievadei, ļauj iegūt vienkāršas diagrammas. Lielākus datu apjomus skolēni noglabā datubāzēs (piemēram, Microsoft Access). Tīmekļa (Web) mājaslapas šādi strukturēto informāciju viegli padara pieejamu ikvienam, kam ir Interneta pieslēgums.
* **Vizuālajā mākslā** apgūst attēlu proporcijas, krāsas (to dalījumu siltajās un aukstajās, pretnostatījumu pēc to piesātinātības, spilgtuma un kontrasta), piemērotu fontu izvēli, lai teksti būtu viegli uztverami.

Datu apstrādes metodes ir labi zināmas, bet skolēni tās nereti apgūst fragmentāri un dažādos mācību priekšmetos: Par kļūdu apstrādi mēdz mācīt fizikas stundās; MS Excel grafikus apgūst informātikā, par demogrāfiskām eglītēm var stāstīt ekonomikā bet par vides datiem – ģeogrāfijā. Lai dotu iespēju ieinteresētajiem skolēniem iepazīt datu apstrādi kā vienotu zināšanu sistēmu, piedāvājam fakultatīvu par datu mācību un nosaucam to “Informācijas dizains un vizualizācijas”, lai priekšlaicīgi nenobaidītu skolēnus, kuri netic savām matemātikas vai informātikas spējām.

## Par ko būs šis fakultatīvs

Formāli runājot, fakultatīva tēmu var saukt par aprakstošo statistiku (*descriptive statistics*) – t.i. novērojumos iegūto datu apkopošanu un vispārināšanu (*aggregation and summarization of data observations*). Ir cita statistikas nozare – inferenciālā statistika, t.i. hipotēžu pārbaude – piemēram, vai medikaments ir efektīvs noteiktas slimības ārstēšanai. Tomēr šajā ievadkursā hipotēžu pārbaudei pievērsīsimies tikai pavisam nedaudz, jo skolēniem arī vecākajās klasēs (izņemot dažus gadījumus ZPD – zinātniski pētnieciskajos darbos) biežāk jāsaskaras ar datu attēlošanu, vizualizāciju un intuitīvu izprašanu, nevis jaunu zināšanu iegūšanu no datiem vai jaunu zinātnisku teoriju pārbaudi.

Minēsim dažas svarīgas tēmas, kuras skolā pagaidām māca nepietiekami, kas var noderēt dažādos mācību priekšmetos, un kur bērni samērā īsā laikā var sasniegt jūtamu progresu:

1. **Infografikas:** Sarežģītu tendenču vieglai un saprotamai attēlošanai vajag daudzveidīgas diagrammas, animācijas un to radošu sintēzi - tās bieži sauc par infografikām. Stabiņu un sektoru diagrammas ir tikai sākums. Datu pratība (literacy) nozīmē vismaz spēju nolasīt sarežģītu diagrammu. Ideālā gadījumā - arī spēt “piepildīt” ar saviem datiem kādu diagrammas vai animācijas šablonu - izstāstīt pašam savu stāstu ar datiem. Labs piemērs, kur šādi šabloni ir pieejami, ir Gapminder - <https://www.gapminder.org/tools/> . Tās ir gan “burbuļu diagrammas”, gan interaktīvas kartes, gan dažādu tendenču attēlojumi divās vai trīs dimensijās, ar vai bez kustīgiem elementiem, ar radošu fontu un krāsu lietojumu. Šī ir joma, kur var izpausties cilvēki ar vizuālās mākslas talantiem.
2. **Datu žurnālistika un hakatoni:** Mūsdienu valsts iestādes, mediji, sociālie tīkli, uzņēmumi un pat skolas ik dienas saražo milzumdaudz dažādu datu. Daļa no tiem ir publiski pieejami, citiem piekļuvi var palūgt, ja tas vajadzīgs pētnieciskiem mērķiem. Diemžēl, liela daļa no šiem datiem ir pirmajā acumirklī grūti uztverami. Lai šādi dati kļūtu noderīgi gala lietotājam, tie jāpārvērš datoram saprotamā formā un jāapstrādā, par tiem jāizveido "stāsts”. Šī ir joma, kurā bieži sadarbojas žurnālisti un programmētāji. Arī skolu vecākajās klasēs tie skolēni, kurus nebaida ne humanitārā, ne eksaktā joma, var veidot kopīgus projektus. Tad arī it kā “garlaicīgi” dati (piemēram, kādas pašvaldības veikti publiskie iepirkumi) var pēkšņi “ierunāties”. Var izrādīties, ka pašvaldība tērē naudu nesaprotamiem mērķiem, remontē tikai dažas skolas, dokumentos minētie darbu apjomi nesakrīt ar dabā esošajiem, utml. Šādi pētījumi var kalpot par pamatu skolēnu ZPD, programmēšanas projektiem un dažām citām vajadzībām.
3. **Datu vākšanas praktiskie darbi:** Vienalga, vai ir runa par ķīmijas eksperimentiem, sensoru tīkliem, meklējumiem Internetā vai marketinga aptaujām, spējīgākos bērnus ir jāiepazīstina ar datu ievākšanas metodēm, ar pašu mirkli, kad dati rodas. Vajadzīgi vingrinājumi, kuri (pat īpaši nekoncentrējoties uz konkrētās nozares saturu) parāda mērījumu metodes vai datu ievākšanas metodikas vājās vietas. Pat ja vingrinājums ir tik vienkāršs kā pulsa skaitīšana skolasbiedram, skolēnu veidotas mājaslapas apmeklējumu statistika ar Google Analytics vai temperatūras un CO2 sensors skolas telpās, tas parāda reālas datu plūsmas uzvedību.
4. **Melošana ar statistiku:** Diemžēl, statistikas metodes un arī pilnīgi nepamatotas manipulācijas ar datiem nereti izmanto negodprātīgi cilvēki, lai vazātu citus aiz deguna. Lai skolēnus sagatavotu šādām lietām, ir vēlams atsevišķi aplūkot visādu “datos balstītu demagoģiju”, dīvaini izveidotas diagrammas un citas vizualizācijas, kas notušē autoram nevēlamo vai izceļ vēlamo. Materiāli ieinteresētu cilvēku nespēja vienoties par sarežģītu datu interpretāciju ir pamatā strīdiem par cilvēka izraisītām klimata izmaiņām, riskiem (vai arī otrādi - risku neesamību), kas saistās ar masu migrāciju, drošību un noziedzību, iedzīvotāju izmiršanu u.c. demogrāfiskām izmaiņām, makroekonomiskiem rādītājiem, valsts budžeta skaitļiem, tēriņiem par izglītību un citām politiski jūtīgām tēmām.

Minētos 4 lielos mērķus paredzam sasniegt, attīstot dažas konkrētas prasmes sešu nodarbību ciklā:

1. **Infografikas: Datu tabulas:** Spēt ielasīt R programmā CSV failu kā “data frame”, saprast tajā esošos datu tipus (skaitļi, klasifikatori, teksti, datumi), spēt to filtrēt, apkopot, sapludināt ar citām tabulām (aggregate, aggregate, merge). **Vēlamais galaiznākums:** Skolēns spēj no sākotnējām datu tabuliņām (t.s. “raw data” jeb “jēlajiem datiem”) aši saražot jaunas tabuliņas - atbilstoši uzdevuma nosacījumiem.
2. **Infografikas: Datu vizualizācija:** Spēt izvēlēties piemērotu datu attēlošanas šablonu jeb diagrammas veidu un piepildīt to ar “data frame” datiem, kas iegūts iepriekšējā nodarbībā. Gūt priekšstatu par meklējumu telpas sašaurināšanu (drill down).   
   **Vēlamais galaiznākums:** Skolēns izveido vizualizācijas kā rastra grafiku PNG, vai kā (interaktīvu) vektorgrafiku SVG. Ja šablons to paredz, tad tam var būt arī iespēja navigēt starp vizualizācijām, uzstādīt parametru vērtības ar vadības elementiem bīdāmu kloķīšu vai teksta ievades lauku veidā.
3. **Hakatoni: Pārskata veidošana:** Apskatīt kādas valsts iestādes vai komerciālu datu kopu (vēlēšanu dati, KNAB dati, Centrālās Statistikas Pārvaldes tabuliņas - piemēram, par demogrāfiju, nekustamo īpašumu vai automašīnu cenas sludinājumu serverī SS.LV, iepazīšanās portāli, vidusskolu centralizēto eksāmenu rezultāti). Veikt šo datu analīzi, izveidot infografikas (un saprast, kurā virzienā žurnālistam vajadzētu rakt tālāk).  
   **Vēlamais galaiznākums:** Skolēns izveido pārskatu (report), kur infografikas kombinētas ar skaidrojošu tekstu, ko var izvietot Internetā.
4. **Hakatoni: Ģeogrāfiskie dati:** Izmantojot ģeogrāfiskās koordinātes, apzīmēt punktus Google kartē. Izmantot arī ESRI Shapefile vai līdzīgus vektoru datus, lai veidotu kartes, kurās vizualizēta kaut kāda informācija - piemēram, ar krāsas toni vai intensitāti norādīta kāda parametra vidējā vērtība attiecīgajā apgabalā, interaktīvās kartēs apzīmētas skolas vai pilsētas, u.c.  
   **Vēlamais galaiznākums:** Skolēns izveido intensitāšu karti (choropleth map) vai interaktīvu karti ar karodziņiem. Tā var būt Latvijas novadu, pagastu, citu teritoriālo vienību karte, Eiropas karte vai kas līdzīgs.
5. **Praktiskie darbi:** Fizikas vai ķīmijas praktiskais darbs, kurā notiek atkārtoti mērījumi, tiek izrēķinātas vidējās vērtības, standartnovirzes, noskaidroti ticamības intervāli, izmantojot Stjūdenta koeficientus. Var būt arī praktiskais darbs, kurā starp neatkarīgo un atkarīgo mainīgo pastāv lineāra sakarība, kur caur datu punktiem jāvelk taisne pēc mazāko kvadrātu likuma.   
   **Vēlamais galaiznākums:** Skolēns izveido pārskatu (report), kas ir kulturāli noformēts skolas laboratorijas darba apraksts.
6. **Melošana ar statistiku:** Aplūkojam dažādus piemērus medijos, kuros ir manipulatīvi pasniegti kaut kādi dati, mācām saskatīt tajos izmantotos “trikus”. Veidojam arī savas infografikas vai datu interpretācijas, lai no kādas reālas datu kopas iegūtu tendenciozus (nevis objektīvi pamatotus) secinājumus.

## Motivācijas sadaļa

Svarīgs jautājums - kādēļ bērniem visas augšminētās lietas būtu jāmācās.

Pēdējo gadu gaitā datu mācība (data science) izraisījusi lielu ažiotāžu. ASV un Ķīna sacenšas mākslīgā intelekta lietojumu attīstīšanā (piemēram, cilvēku seju atpazīšana ielu videokamerās), “deep learning” jeb dziļā mācīšanās, izmantojot neironu tīklus ir izveidojusi Alpha Zero - datorprogrammu, kas spēj uzvarēt jebkuru cilvēku un jebkuru agrāku datorprogrammu tādās intelektuālās spēlēs kā šahs un go, Facebook un citu sociālo tīklu dati palīdzējuši politiskajiem konsultantiem (piemēram, Cambridge Analytics) viņu politiskajā aktīvismā - veidojot priekšvēlēšanu stratēģijas ASV prezidenta vēlēšanās 2016.gadā vai Brexit referendumā Lielbritānijā (arī 2016.g.).

Datu analīze pēdējo gadu gaitā pārstājusi būt šauru speciālistu (matemātiķu vai statistiķu) atbildība, dati jāapkopo un jāapstrādā gandrīz ikvienā profesijā. Ar matemātiskām metodēm savā darbā saskaras ne vien inženieri un dabaszinātnieki, bet arī ģeogrāfi, ekonomisti, finansisti, pārdošanas/marketinga speciālisti, sociologi, demogrāfi, psihologi, ārsti, u.c. Datu apstrāde var atvieglot arī skolotāja ikdienu, palīdz sagatavot interesantas ilustrācijas savos mācību materiālos. Jau tagad vidusskolas matemātikas un informātikas kursi piedāvā daudzas no tām prasmēm, kas vajadzīgas datu analīzes uzdevumu risināšanai.

Svarīga mērķauditorija ir arī eksakto mācību priekšmetu (arī ģeogrāfijas, ekonomikas, informātikas u.c.) skolotāji. Viņiem varētu būt interese par datu prasmēm kā vienojošo elementu, mācot līdzīgas tēmas dažādos mācību priekšmetos. IZM par skolotāju tālākizglītību atbild Aivars Opincāns ([aivars.opincans@izm.gov.lv](mailto:aivars.opincans@izm.gov.lv) ). Ja šo kursu saskaņo atbilstoši Ministru kabineta noteikumiem nr. 431 (2001.g. 7.jūnijs) vai citā līdzīgā veidā, tad to var piedāvāt ne vien spējīgākajiem ģimnāziju audzēkņiem, bet, iespējams, arī kā skolu pedagogu profesionālās kvalifikācijas kursu.

**Nepieciešamās priekšzināšanas:** Vēlamas priekšzināšanas vidusskolas matemātikas kursa apjomā (it īpaši aritmētika, algebra, funkciju pētīšana, planimetrija, Dekarta koordinātes, vienādojumi un to sistēmas, nevienādības). Vajadzīgas pamatiemaņas darbā ar datoru – spēja orientēties Windows vai Linux operētājsistēmā, darbināt lietojumprogrammas (piemēram, izklājlapas) un pildīt instrukcijas. Vēlama arī spēja orientēties Internetā, kā arī interese par datu analīzes lietojumiem.

**Apmācības ilgums:** 24 akadēmiskās stundas plus aptuveni 8 akadēmiskās stundas, ko skolēni velta kursam patstāvīģi (piemēram, 6 reizes pa 4x40 minūtēm). Aptuveni atbilst 2 kredītpunktu kursam augstskolā. Nodarbības var noturēt, piemēram, trīsreiz nedēļā (brīvdienu laikā, pēc darba utml) - divas nedēļas.

**Aprīkojums:** Katram kursa klausītājam nepieciešams dators. Uz datora uzstādāmas divas bezmaksas atvērtā koda lietojumprogrammas – „LibreOffice Calc”/”MS Excel” (izklājlapas datu ievadei) un R/RStudio grafiskā vide (aprēķinu veikšanai, statistikai, grafiku un diagrammu zīmēšanai).

**Kursa sekmīgas pabeigšanas kritēriji:** Izpildīti nodarbībās aplūkotie vingrinājumi, īpaša eksāmena nav.

## Datu vizualizācijas vingrinājumi

Visticamāk, ka visas šīs vizualizācijas fakultatīvā iekļaut neizdosies, jo fakultatīva laiks ir ļoti ierobežots. Bet šeit minētās lietas var raksturot, ko varētu uzbūvēt ar kursā’iegūstamajām zināšanām.

1. Elementāru funkciju grafiki. (Parādām, kā uzzīmēt visas pamata funkcijas - pakāpes, eksponentfunkcijas, logaritmiskās, trigonometriskās, inversi trigonometriskās.) Var parādīt, kā datoram var uzticēt tipveida uzdevumu risināšanu. Skolēns ievada kvadrātvienādojumu un saņem pārskatu, kur tas risināts “pilnā pierakstā”.
2. Līzinga un kredīta kalkulatori un ikmēneša maksu vizualizācijas. Eksperimentēšana ar dažādiem aizdevumu procentiem un komisijas maksām – cik kredīta ņēmējs pārmaksā.  
   Dažādu gabaliem lineāru funkciju attēlošana un praktiskie lietojumi.
3. Demogrāfiskās eglītes. (Zīmējam simulācijas dažādiem demogrāfiskajiem jautājumiem.)  
   Varbūtiski eksperimenti; gaidīšanas laiki (sakrītošās dzimšanas dienas; nāves brīdis, utml.); pensijas aprēķināšana – un pensijas lieluma saistība ar paredzamo mūža ilgumu. 1., 2., un 3. līmeņa pensijas.
4. Kustīgo burbuļu diagrammas. (Zīmējam dažas diagrammas ar burbulīšiem – vēlams par Latvijas reģioniem vai pasaules valstīm – makroekonomiski vai demogrāfiski dati.) Parādām, ka ar animētas diagrammas palīdzību var veidot dinamisku priekšnesumu. Sal. GapMinder.org – Hans Rosling video-lekcijas, kas plaši izmanto šo vizualizācijas tehnku.
5. Fraktāļi un citas jocīgas matemātiskas bildītes. Kuros gadījumos iteratīvi procesi rada haosu (3 ķermeņu problēma fizikā; atraktori; daļveida dimensija; Hilberta līkne, kas aizpilda plakni vai telpu; nepārtrauktas un nevienā punktā neatvasināmas funkcijas jeb t.s. “eži”).
6. Valsts iestāžu datu vizualizācijas. Saeimas vēlēšanu vizualizācijas (neparasti liels plusu/svītrojumu skaits atsevišķos iecirkņos – aizdomas par balsu pirkšanu vai balsotāju atvešanu ar autobusiem; vēlēšanu kandidātu ranžēšana – ranga numura iespaids uz izredzēm tikt ievēlētam). KNAB datubāze. Kādos veidos un šķērsgriezumos var parādīt ziedojumu datus. Ieejošie dati – KNAB datubāze un CVK – dati par deputātu kandidātiem. Rīgas pašvaldības iepirkumu datubāze – anomālijas iepirkumos (neparasti augsta vai zema cena; konkursa uzvarētājs, kurš nupat kā reģistrēts uzņēmums, utml.)
7. Ģeogrāfiskie dati. Klimats no meteo.lv (Nokrišņi, vidējās temperatūras, utml.) Klimatgeitas stāsts. Kartes zīmēšana. Vēlēšanu rezultātu, centralizēto matemātikas eksāmenu rezultātu utml. attēlošana kartē.
8. Fizioloģiski dati. (Piemēram, asisnsspiediena vai pulsa mērījumi, ko var veikt klasē – pirms un pēc kāda fiziska vingrinājuma). Praktiskas iemaņas ātri nosakot pulsu – saskaitot to 15 sekundes (vai pilnu minūti). Vai skaitīšanas ilgums iespaido precizitāti?

## Teorijas tēmas matemātikā, statistikā un programmēšanā

1. **Tēma:** Darbības ar skaitļiem, dubultā precizitāte.   
   <http://rwiki.sciviews.org/doku.php?id=misc:r_accuracy> – stāsti par to, kā reālie skaitļi tiek attēloti ar „double” mainīgajiem. Skaitļu binārais pieraksts.   
   Augstas precizitātes aritmētika. Lieli skaitļi. Racionāli skaitļi. Kompleksie skaitļi.   
   **Vingrinājumi valodā R:** Dažu lielu pirmskaitļu aprēķināšana (2^n-1 utml.). Racionālu skaitļu aritmētika; vērtības, kas attēlo „bezgalību” vai „nedefinētu skaitli” (NaN). Bibliotēka “gmp” (Gnu multiple precision) – darbības ar ļoti lieliem veseliem skaitļiem.
2. **Tēma:** Darbības ar stringiem, regulāras izteiksmes.   
   Datu tabulas dažādos formātos, un to pārveidošana viegli apstrādājamā veidā (CSV, Tab-delimited). Stringu apstrāde – konkatenācija, apakšstringi, meklēšana stringā, sadalījums burtos, utt.   
   **Vingrinājumi valodā R:** Valodas R literāļu sintakse. Regulāro izteiksmju u.c. instrumentu atbalsts pašā valodā R. Reālās pasaules datu tabulu sagatavošana ielasīšanai.
3. **Tēma:** Vektori, matricas, pārveidojumi. Datu struktūras “Vector” un “Data frame” valodā R.  
   Pagriezienu matrica ar piemēriem. Vispārīgs lineārs pārveidojums, un ko tas izdara ar dažādām bildītēm (burtu „R” piemēram). Šķiebšana, spiešana, rotācija, paralēlā pārnese. Homogēnās koordinātes.  
   **Vingrinājumi valodā R:** Operācijas ar vektoriem un matricām; lineāru vienādojumu sistēmas. Papildu koordinātu sistēmu ieviešana; aplīšu utml. Figūru zīmēšana 2D un3D Dekarta koordinātēs.
4. **Tēma:** Elementārās funkcijas.   
   Kā saprātīgi izvēlēties koordinātu intervālus, kur zīmēt grafiku (Y pieskaņošana X). Grafiku pārveidojumi (paralēlā pārnese pa X un Y asi; saspiešana pa ikvienu no asīm, utml.). Krāsainas grafu bildītes – dažādu masu uzdevumu atrisinājumi (uzzīmēt grafus visādām parabolām un racionālām daļām).   
   **Vingrinājumi valodā R:** Izteiksmes valodā R. Dažādu funkciju pieraksts; grafiku zīmēšanas komandas; iespēja saglabāt rezultātu kā attēlu; R koda ielasīšana no failu sistēmas vai Weba. Sastādīt astronomiskas tabulas vai planētu stāvokļus/mēness fāzes, zīmēt bioritmus.
5. **Tēma:** Funkciju atvasinājumi.   
   Funkciju pētīšana un grafiku zīmēšana, t.sk. izmantojot atvasinājumu. Teilora rindas.  **Vingrinājumi valodā R:** Sinusa tuvināšana ar Teilora rindas parciālsummām. Kā jāņem vērā noapaļošanas kļūdas.
6. **Tēma:** Grafu teorijas elementi.   
   Dažādu veidu grafi (koki, pilni grafi, cikli, utml.) Iespējas šos grafus attēlot, zīmējot plaknē. Transporta shēmas; grafi ar svērtām šķautnēm, grafu algoritmi.   
   **Vingrinājumi valodā R:** Kā zīmēt divās dimensijās – virsotnes, šķautnes, burti, utml. Sarežģītākas animācijas. Īsākā ceļa meklēšana grafā.
7. **Tēma:** Datu tabulu apstrāde. ESRI Shapefile formāts un darbības ar krāsainiem punktiem, līnijām, daudzstūrīšiem. Kalendāru diagrammas, kustīgo burbuļu diagrammas, utml. Datu meklēšanas un apkopošanas veidi, izmantojot Internetu. Kvantitatīvi un kvalitatīvi dati. Ģenerālkopa un iztvērums, indivīdi un to atribūti. Datu filtrēšana un pieprasījumi to izvilkšanai no datubāzēm.  
   **Vingrinājumi valodā R:** Datu filtrēšana ar SQL līdzīgiem pieprasījumiem (ierakstu atlase, kolonnu atlase). Infografikas, ko var sazīmēt ar R. Datu vektoru filtrēšana, grupēšana. Kartes, piemēram, vēlēšanu rezultātu vai tautas skaitīšanas rezultātu attēlošana kartē.
8. **Tēma:** Līknes, iteratīvi procesi, fraktāļi. Iespēja uzzīmēt dažādas jocīgas līknes, lauztas līnijas, tuvināti attēlot fraktāļus, utml. Interpolācijas polinomi.   
   **Vingrinājumi valodā R:** Komplekso skaitļu apstrāde. Polārās koordinātas. Līkņu zīmēšana. Vienkāršas animācijas (piemēram, Koha sniegpārsliņa). Hilberta līkne 2 un 3 dimensijās.
9. **Tēma:** Varbūtiski modeļi.   
   Monētas atkārtota mešana – nejauša klaiņošana. Dzimšanasdienu uzdevums. Kombinatorisku lielumu tuvināta novērtēšana. Uzdevumi ar spēļu kārtīm. Binomiāls (Bernulli) sadalījums. Centrālā robežteorēma.

**Vingrinājumi valodā R:** Varbūtisku notikumu (monētas, metamā kauliņa utml.) simulēšana. Kā izpaužas “lielo skaitļu likums”, kuros gadījumos rodas normālais sadalījums.

1. **Tēma:** Demogrāfiski dati. Histogrammas un demogrāfiskas eglītes.   
   **Vingrinājumi valodā R:** Demogrāfiskas prognozes (ekstrapolācija pie zināmiem dzimstības un mirstības koeficientiem; kopējā depopulācija; nepilsoņu depopulācija; „dependency ratio” ekstrēmi; strādājošo/pensionāru skaita attiecība stacionārā gadījumā; zema dzimstība ar imigrāciju).
2. **Tēma:** Integrālrēķini. Daži nepārtraukti sadalījumi: Puasona sadalījums. Hi kvadrāta sadalījums. Normālais sadalījums. Centrālā robežteorēma.  
   Integrālsummas. Histogrammas. 2.kosmiskais ātrums un Zemes atmosfēra.   
   **Vingrinājumi valodā R:** Cikli, integrāļu tuvināta aprēķināšana. Kas ir labāk – saņemt 15000 latu tagad; vai pa 1000 latiem turpmākos 25 gadus? Uzdevumi par kredītiem un līzingu. Dažu funkciju tuvināta integrēšana. Z-indekss un novērtējumi, kāda daļa no novērojumiem būs noteiktā intervālā, ja zināma normāli sadalīta gadījumlieluma vidējā vērtība un dispersija.
3. **Tēma:** Fizikālu mērījumu apstrāde. Stjūdenta koeficients (kā darboties tad, ja dispersija nav zināma). Fizikāli eksperimenti, kuros rodas vajadzība pēc lineāras regresijas.   
   **Vingrinājumi valodā R:** Aplūkojam fizikas vai ķīmijas laboratorijas darbus (piemēram, kristālhidrāta karsēšana – lai uzzinātu, cik ūdens molekulas piesaistītas katrai kristālhidrāta sāls molekulai; titrēšana – lai uzzinātu kādas vielas koncentrāciju šķīdumā – piemēram cukuru saturošā dzērienā). Fizikas eksperimenti ar pieckārtīgi atkārtotiem mērījumiem. Fizikas eksperimenti, kuros jāvelk lineārās regresijas taisne. Kļūdu novērtēšana fizikas eksperimentā, Stjūdenta koeficients utml.
4. **Tēma:** Kvantitatīvu un kvalitatīvu datu analīze. Socioloģiskas aptaujas. Datumu, veselo un daļskaitļu attēlošana datos. Nedefinētās vērtības. Datu virknes jeb vektori (vectors); datu tabulas jeb ietvari (data frames). Vidējā vērtība, mediāna, kvantiles, kvadrātiskā novirze. „Stumbra un lapu” diagrammas (stem and leaf plot), „kastes un ūsu diagrammas” (box and whisker plot). Teksta vērtības un klasifikatori. Kvalitatīvi dati tabulās un diagrammās (joslu un sektoru diagrammas).   
   **Vingrinājumi valodā R:** Socioloģisku aptauju veidošana. Aptauju datu attēlošana diagrammās. Datu kārtošana augošā, dilstošā secībā. Histogrammas. Virknes maksimums, minimums, summas, parciālsummas.