

# Matemātika un datu zinātne

LU Studentu zinātniskais seminārs

2024. gada 16.oktobris

Jānis Lazovskis

Printful + RTU Rīgas Biznesa Skola

# Ievads: Profesionālā pieredze

2009 - 2019:



2019 - 2024:



2020 - ... :



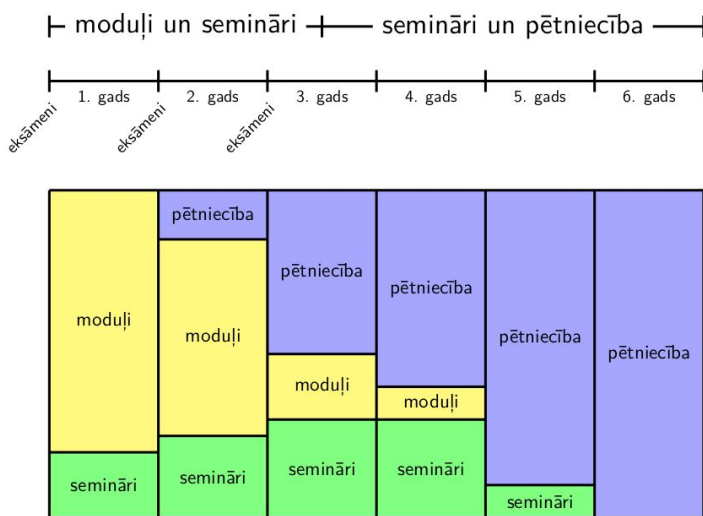
- Bakalaura grāds kā pilna laika darbs
- Maģistra grādu matemātika parasti izlaiž
- Doktora grāds apvieno studijas un pētniecību

Course	Description
AMATH 391	From Fourier to Wavelets
LAT 102	Introductory Latin 2
PMATH 499	Readings in Pure Mathematics
Course Topic:	Algebraic Topology

Course	Description
ACINTY 640	Academic Integrity Module
PMATH 665	Differential Geometry
PMATH 745	Groups and Representations
PMATH 955	Topics in Geometry
Course Topic:	Advanced Algebraic Geometry

**Winter**  
 Program: Pure Mathematics, Master of Mathematics  
 Attendance: Full-Time Term: 2.00 Status: Enrolment

Course	Description
PMATH 764	Algebraic Curves
PMATH 800	Topics in Real and Complex Analysis
Course Topic:	Riemann Surfaces
PMATH 955	Topics in Geometry
Course Topic:	Atiyah-Singer Index Theorem



Spring 2015 - Chicago  
 Graduate College  
 Mathematics  
 MATH 516 Second Abstract Algebra I  
 MATH 533 Real Analysis I  
 MATH 589 Teaching Mathematics  
 Ehrs: 10.00 GPA-Hrs: 10.00 QPts:

Spring 2015 - Chicago  
 Graduate College  
 Mathematics  
 MATH 517 Second Abstract Algebra II  
 MATH 535 Complex Analysis I  
 MATH 569 Adv Top In Geom&Diff Topology  
 MATH 596 Independent Study  
 Ehrs: 13.00 GPA-Hrs: 13.00 QPts:

Fall 2015 - Chicago  
 Graduate College  
 Mathematics  
 MATH 549 Differentiable Manifolds I  
 MATH 552 Algebraic Geometry I  
 MATH 595 Research Seminar  
 MATH 596 Independent Study  
 MATH 547 Algebraic Topology I  
 MATH 568 Topics Algebraic Topology  
 MATH 593 Graduate Student Seminar  
 Ehrs: 9.00 GPA-Hrs: 8.00 QPts:

Fall 2016 - Chicago  
 Graduate College  
 Mathematics  
 MATH 555 Complex Manifolds II  
 MATH 569 Adv Top In Geom&Diff Topology  
 MATH 596 Independent Study  
 Ehrs: 9.00 GPA-Hrs: 9.00 QPts:

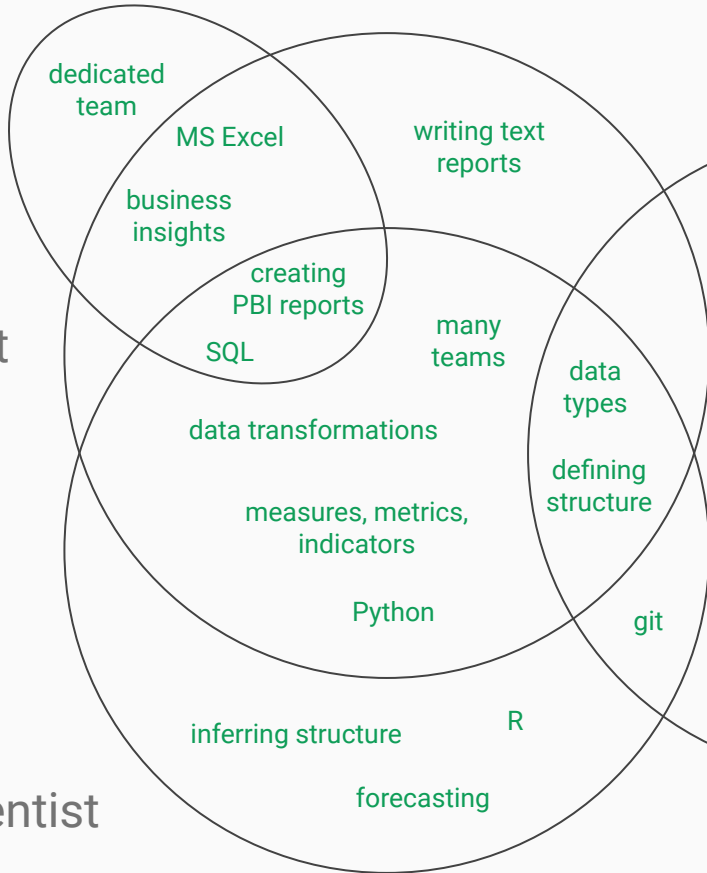
Spring 2017 - Chicago  
 Graduate College  
 Mathematics  
 MATH 553 Algebraic Geometry II  
 MATH 599 Thesis Research  
 Ehrs: 9.00 GPA-Hrs: 4.00 QPts:

# Datu zinātne

- Amata konteksts
- Vajadzības un atalgojums
- Situācija Latvijā un ārzemēs
- Kā alternatīva akadēmiskai karjerai

# Ar datiem saistīti amati

Business Analyst



Data Analyst

Data Scientist

writing text reports

dedicated team

MS Excel

business insights

creating PBI reports

SQL

data transformations

measures, metrics, indicators

Python

inferring structure

forecasting

many teams

data types

defining structure

git

R

ETL process

continuous update and delivery











data integrity







database maintenance

process monitoring

Data Engineer

# Ar datiem saistīti amati

	<b>Senior Data Scientist - MLOps Engineer</b> CV-Online Recruitment — Rīga Publicēts pirms apmēram 16 stundām   Beidzas: 11.11.2024 € 3500 – 4500
	<b>Senior Technical Business Analyst (Integrations)</b> Sapiens Software Solutions (Latvia) SIA — Rīga Publicēts pirms apmēram 13 stundām   Beidzas: 08.11.2024 € 3600 – 4700
	<b>AML Transaction Monitoring Business Analyst/ AML darījumu uzraudzības Biznesa Analītiķis   SEB, Rīga</b> SEB — Rīga Publicēts pirms apmēram 8 stundām   Beidzas: 14.11.2024 € 2100 – 3700
	<b>Data Analyst</b> SIA "Dynamatech" — Rīga Publicēts pirms 6 dienām   Beidzas: 04.11.2024 € 2200
	<b>Automation Business Analyst</b> Roche Latvija, SIA — Rīga Publicēts pirms 2 dienām   Beidzas: 04.11.2024 € 1900
	<b>Data Scientist</b> Arvato Systems Latvia SIA — Rīga Publicēts pirms 15 dienām   Beidzas: 01.11.2024 € 3800 – 5600
	<b>Junior BI Analyst</b> GoCardless — Rīga Publicēts pirms 16 dienām   Beidzas: 30.10.2024 € 1,2 – 1,33
	<b>Data Analyst</b> AS Mapon — Rīga Publicēts pirms apmēram 13 stundām   Beidzas: 01.11.2024 € 2200
	<b>Senior Business Analyst, TietoEVRY Banking</b> TietoEVRY — Rīga Publicēts pirms 1 dienas   Beidzas: 20.10.2024 € 3500 – 5000
	<b>Data Scientist</b> Eleving Consumer Finance AS — Rīga Publicēts pirms 29 dienām   Beidzas: 17.10.2024 € 3500 – 4300

	<b>Bioinformatics Senior R Engineer (Clinical Data Scientist)</b> EPAM Systems Latvia (Remote) 6 school alumni work here Viewed · Promoted · <b>Be an early applicant</b>
	<b>Sr. Data Scientist, NLP/AdTech (Remote)</b> PulsePoint European Economic Area (Remote) 1 school alum works here Promoted
	<b>Data Scientist</b> RED Global European Union (Remote) Promoted · Easy Apply
	<b>Data Scientist</b> Mapon Rīga, Rīga, Latvia (Hybrid) 4 company alumni work here Viewed
	<b>Data Scientist – Recommender Systems</b> ARRISE powering Pragmatic Play European Union (Remote) 1 company alum works here Promoted
	<b>Data Scientist</b> The Crypto Recruiters European Economic Area (Remote)

cv.lv / linkedin.com

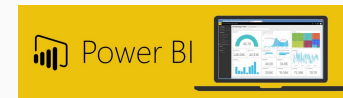
Kādiem uzņēmumiem vajadzīgi datu zinātnieki?

- Bankām
- Farmācijas uzņēmumiem
- Vairumtirdzniecībai
- Tehnoloģiju uzņēmumiem

Uzņēmumiem, kas to var atļauties (naudas un laika ziņā)

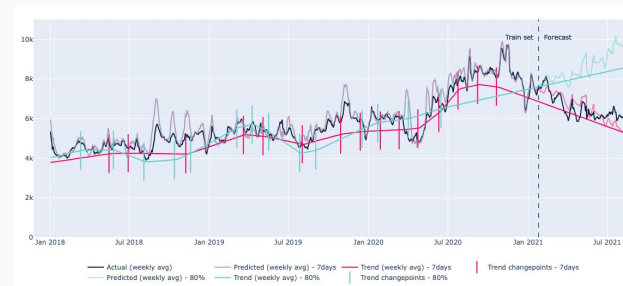
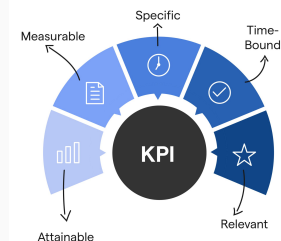
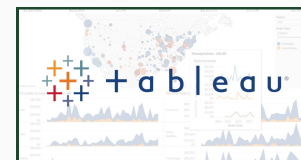
## Ikdienas darbi un atbildības

- Creation and maintenance of PowerBI reports for other teams
- Deep dives, analytical reports on open questions
- Monthly, quarterly tasks and progress tracking



## Rīki un programmēšana

- PowerBI (datu savienošana un vizualizācija)
- SQL (dažādos rīkos)
- Python, R, shell scripts
- Office tools (Google Docs, Sheets, Slides)



# Akadēmiskais darbs vs Darbs uzņēmumā: *Academia vs Industry*

## Academia positives

- High autonomy
- Innovative tasks, exploration
- Direct impact on local people
- Eventually high job security (EU, US)

## Industry positives

- Doing “real” work
- Good (LV) / high (EU, US) salary
- High level of activity
- Reasonable job security

## Academia negatives

- Low job security (LV)
- Aging, conservative system (LV)
- Closed loop science
- Medium (EU, US) / low (LV) salary
- Bureaucracy

## Industry negatives

- Medium / low autonomy
- Routine, repetitive tasks
- Capitalist machine
- Bureaucracy
- Indirect impact
- Academic achievements not valued



# Matemātikas pētniecība

- Algebriskā topoloģija
- Persistent homology
- Applications of computational topology

## Algebraic topology (*algebriskā topoloģija*)

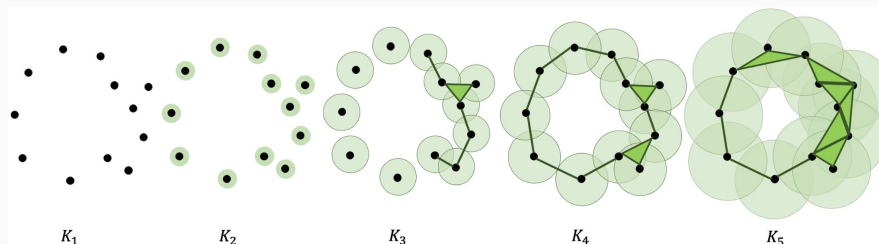
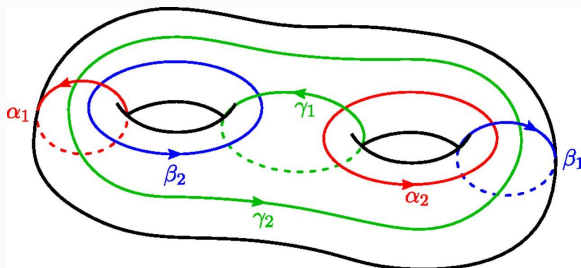
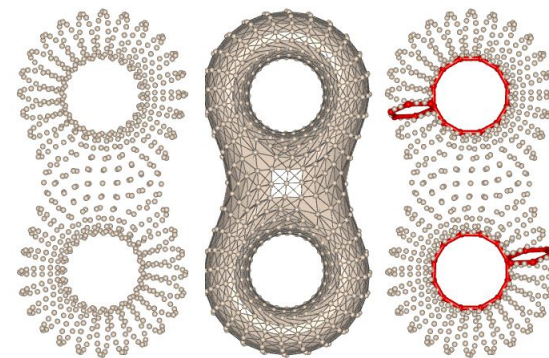
- Associating algebraic objects to topological spaces

## Topological data analysis (*topoloģiskā datu analīze*)

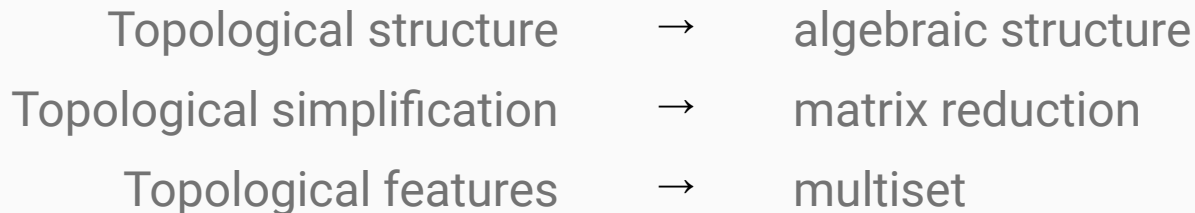
- Associating topological characteristics to data

## Computational geometry / computational topology

- Explicit descriptions of topological spaces



# Persistent homology & dynamic data



(a) A filtered simplicial complex:



(b) We put a total order on the simplices that is compatible with the filtration:



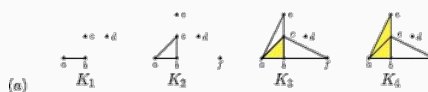
where  $\sigma_i$  denotes the  $i$ th simplex in this order.

(c) (Left) The boundary matrix  $B$  for the filtered simplicial complex in (a) with respect to order on simplices in (b), and (right) its reduction  $\bar{B}$  given by applying Algorithm 1 (one first adds column 5 to column 6, and then column 4 to column 6):

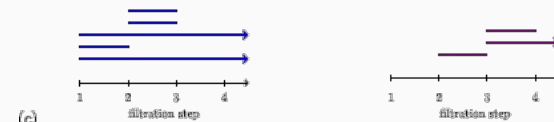
$$B = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \bar{B} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(d) We read off the following intervals from the matrix  $\bar{B}$  in (c):

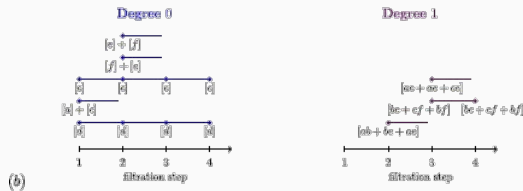
- $\sigma_1$  is positive, unpaired; this gives the interval  $[1, \infty)$  in  $H_0$ .
- $\sigma_2$  is positive, paired with  $\sigma_4$ ; this gives no interval, because  $\sigma_2$  and  $\sigma_4$  enter at the same time in the filtration.
- $\sigma_3$  is positive, paired with  $\sigma_7$ ; this gives the interval  $[3, 3]$  in  $H_0$ .
- $\sigma_6$  is positive, paired with  $\sigma_7$ ; this gives the interval  $[3, 4)$  in  $H_1$ .



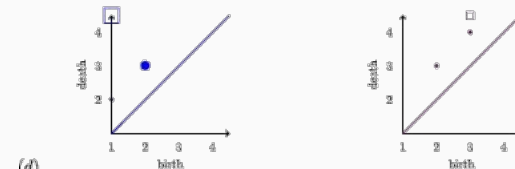
(a)



(c)



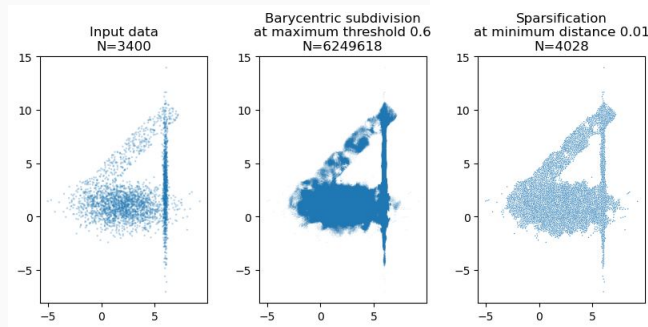
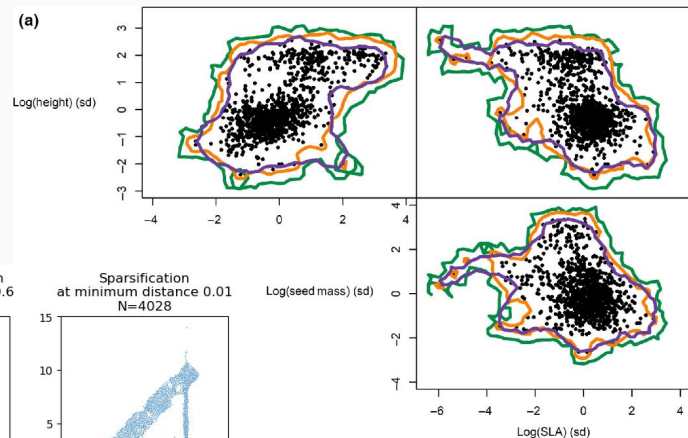
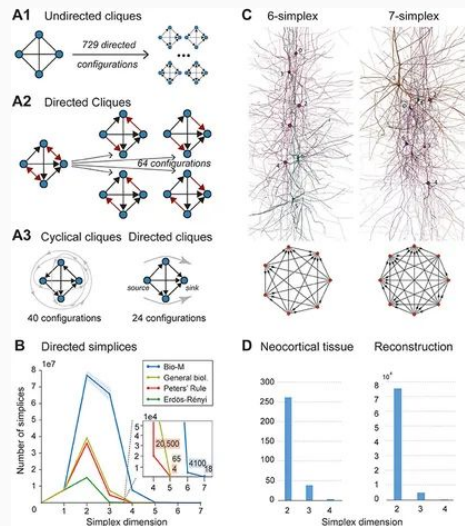
(b)



(d)

## Neuroscience

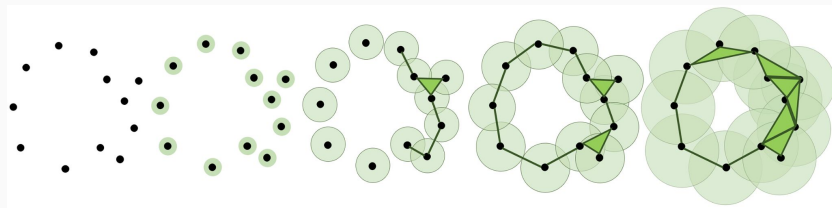
- The brain may be viewed as a directed graph
- The clique complex is a topological space
- Neuron activity gives time-dependent subgraphs



## Ecology

- Species observations are in a high dimensional space
- Voids in this space create opportunity for takeover

1. Datu zinātne
2. Matemātika un topoloģija



### Avoti:

- CV.lv, LinkedIn.com, Printful.com
- Allen Hatcher, [Algebraic Topology](#)
- Tamal Dey, Yusu Wang, [Computational Topology for Data Analysis](#)
- Otter et al, [A roadmap for the computation of persistent homology](#)
- Reimann et al, [Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function](#)
- Conceição et al, [An application of neighbourhoods in digraphs to the classification of binary dynamics](#)
- Blonder et al, [New approaches for delineating n-dimensional hypervolumes](#)