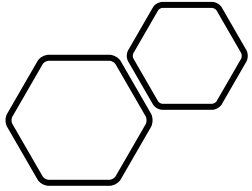


No galda datora līdz  
lieldatoram. Kas mūs kavē  
un pirmie soļi HPC  
izmantošanā

Jānis Bicāns (Janis.Bicans@rtu.lv)

21.11.2023





# Saturs

Eiropas konteksts un aktualitāte

---

Kas ir Superskaitļosana?

---

Pirmie soļi HPC izmantošanā



# Eiropas konteksts



- **Eiropas digitālā stratēģija :**
- Tehnoloģijas, kuras strādā cilvēkiem – iespējami plašam lokam
- Godīga un konkurējoša digitālā ekonomika – infrastruktūra un pakalpojumi ir plaši un caurskatāmi pieejami
- Atvērta, demokrātiska un ilgtspējīga digitālā sabiedrība
- Eiropa kā globāls digitāls spēlētājs – izglītība un inovācija ir centrālie dzinējspēki

<https://ec.europa.eu/digital-single-market/en/content/european-digital-strategy>

# Programma «Digitāla Eiropa» (DEP)

## - pieci konkrēto mērķu virzieni

### Augstas veiktspējas skaitļošana

- Eksalīmeņa iekārtu iepirkums
- Esošo superdatoru uzlabošana
- Kvantu skaitļošana
- Piekļuves un izmantošanas superskaitļošanai paplašināšana

### Kiberdrošība

- Kompetenču centru tīkla izvēršana
- Kiberdrošības vairogs (*Cybersecurity shield*), kvantu komunikācijas infrastruktūra
- Sertifikācijas shēmas
- Kiberdrošības rīki

### Eiropas digitālo inovāciju centri

- Vismaz viens katrā ES Dalībvalstī
- Vismaz viens katrā ES Dalībvalstī MI jomā

### Mākslīgais intelekts

Data4EU: kopīgas datu telpas, mākonis, platformas un infrastruktūra

- Liela mēroga Testēšanas un Eksperimentēšanas centri piecās jomās
- Eiropas MI platformas paplašināšana (*scale up*) piekļuvei testētām MI tehnoloģijām

### Augstas digitālās prasmes

- Maģistrantūra (*Master courses*)
- Īstermiņa apmācības
- Prakses vietas
- Prasmju un darbu platforma

### Izvēršana ar uzsvaru uz:

- Galamērķis Zeme (*Destination Earth*)
- Zaļās un viedās kopienas un mobilitāte
- Investēšanas turpināšana (CEF – ISA2)
- Vienreizes (*Once-only*) princips
- Blokķēde
- Uzticības veicināšana digitālai transformācijai



# Kas ir superskaitļošana?

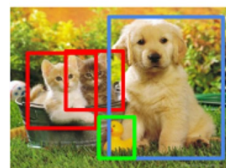
Augstas veikspējas datu  
analīze (High Performance  
Data Analysis)



Vēsturiskie dati

Mākslīgais Intelekts (Artificial  
Intelligence)

## Object Detection



CAT, DOG, DUCK



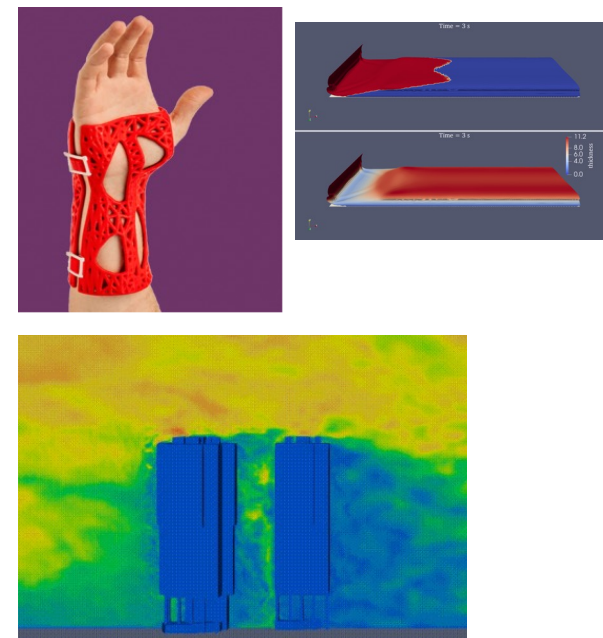
AI-Lost  
Detection of abandoned  
and removed objects



AI-Parking  
Monitoring parking  
areas, perimeter and  
non-perimeter

Reāla laika dati

Mudelēšana (Modelling)



Modelēšana un Simulācijas



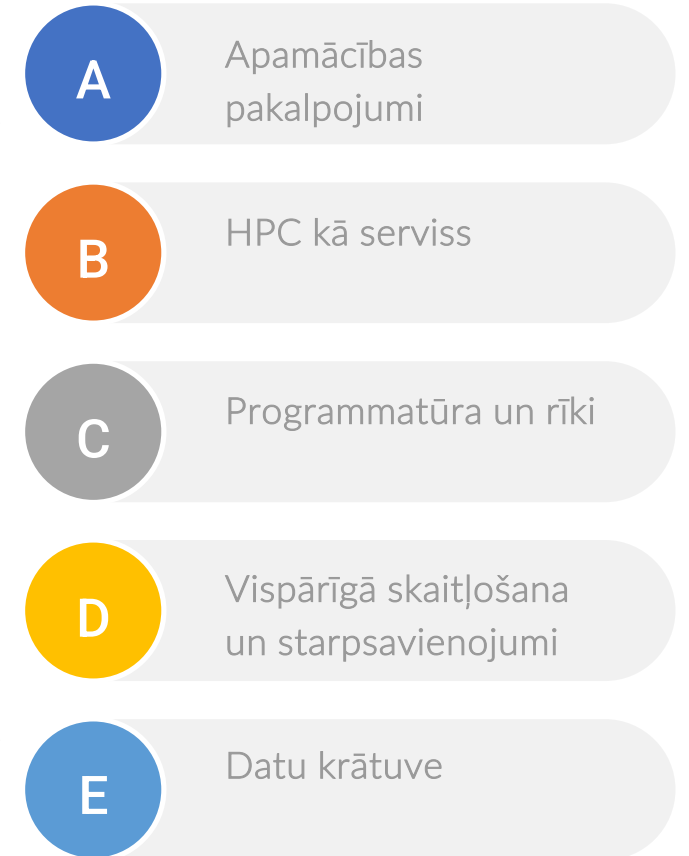
# HPC pakalpojumu grozs

Šobrīd pārāk ierobežots pakalpojums – pieejams lietotājiem ar specifiskām iemaņām

## Mērķis:

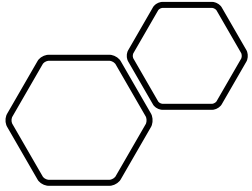
- Mākslīgā intelekta pētījumiem sagatavota programmatūras infrastruktūra;
- Salīdzinoši vienkārša saskarne plašam pielietojumam un pieejamībai;
- Masīvi paralēla skaitļošana, kā pakalpojums pētniekiem un uzņēmējiem;
- Vispārīgā skaitļošanas infrastruktūra;
- Datu īslaicīga un ilgtermiņa uzglabāšana;
- Starpsavienojumi ar pētniecības organizācijām;

**Nevienlīdzīgas iespējas** dažādām organizācijām piekļuvei datiem un skaitļošanas jaudām, kā arī ierobežotas iespējas izmantot tās zināšanu un iemaņu trūkuma dēļ.



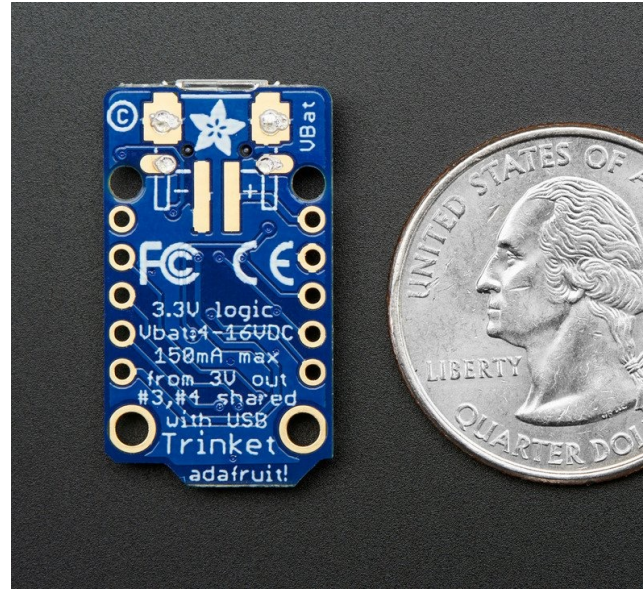
# HPC lietotāju gatavība

- **A-augstākā gatavība** (*high HPC-use readiness*), pati organizācija spēj atrisināt lieldatora izmantošanu. NCC sniedz konsultācijas, kuras ir saistītas ar specifiskiem lietošanas gadījumiem. Piemēram, kā labāk izpildīt uzdevumu resursu plānošanu, ja klasteris ir liels, u.tml. nišas jautājumi, kas ir nozīmīgi ekspertiem.
- **B –daļēja gatavība**, NCC nodrošina HPC mentora un stratēģiskā konsultanta atbalstu, lai definētu uzdevumu, tā pārveidošanu paralēlā skaitļošanas formātā un organizācijas personāla apmācībā un praktisko iemaņu koučingā, lai pārnestu skaitļošanu no galda datoriem uz HPC. Fokuss uz uzņēmumiem, kuriem jau ir konkurētspējīgs produkts un ir nepieciešams atbalsts, kuru nespēj atļauties uzņēmuma iekšienē.
- **C –neliela gatavība** vai gatavības trūkums, NCC nodrošina HPC mentora un stratēģiskā konsultanta atbalstu, lai definētu uzdevumu, tā pārveidošanu paralēlā skaitļošanas formātā un organizācijas personāla apmācībā un praktisko iemaņu koučingā. NCC nodrošina arī organizācijai trūkstošo **pētniecisko kompetenču** piesaisti, bez kurām zinātniski tehniskās problēmas risinājums nav iespējams (piemēram, TV torņa monitoringa sistēmas izveide, piesaistot mehāniskos, aerodinamiskās modelēšanas, vēja ātruma atmosfēras slāņos un laika prognožu, kā arī materiāli izpētes, noguruma analīzes u.c. speciālistus)
- **0-līmenis –bāzes gatavība**. Atspoguļo NCC sniegto bāzes informāciju (mājas lapa, vadlīnijas, ceļveži, demo piemēri, pamatapmācības programmas, HPC fizisko resursu pietiekama pieejamība). No NCC puses jāsniedz pirmais atbalsts, lai saprastu, kas ir vispār iespējams un rosinātu domāt par superskaitļošanu.



# Mākslīgais intelekts un Mašīnmācīšanās

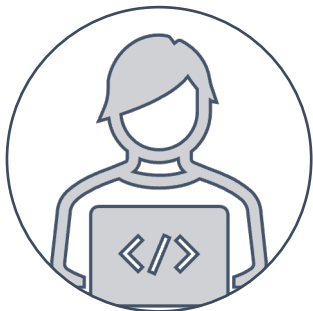
- MI un ML modeļu apmācībai izmanto jaudīgu infrastruktūru
- MI un ML modeļi var darboties arī mazās un mazjaudīgās ierīcēs
- Modeļu individualizācijai nepieciešami problēmvides specifiski un pareizi apstrādāti dati



	EDGE				FOG		PC/Cloud
Size	Nano	Micro	Milli	Server	Fog	Campus	Facility
Example	Adafruit Trinket	Particle.io Boron	Array of Things	Linux Box	Co-located Blades	1000-node cluster	Datacenter & Exascale
Memory	0.5 KB	256 KB	8 GB	32 GB	256 GB	32 TB	16 PB
Network	BLE	WiFi/LTE	WiFi/LTE	1 GigE	10 GigE	40 GigE	N*100 GigE
Cost	\$5	\$30	\$600	\$3K	\$50K	\$2M	\$1000M



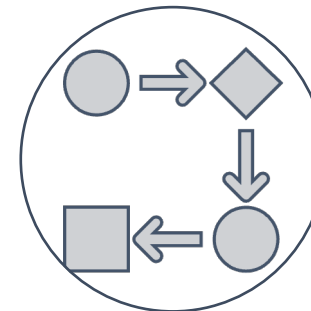
# Iespējamie veidi kā sākt darboties



Programmēt  
Python/R valodās



Vizuālā  
programmēšana



Izmantot gatavus  
risinājumus



Sarežģīti

Vienkārši

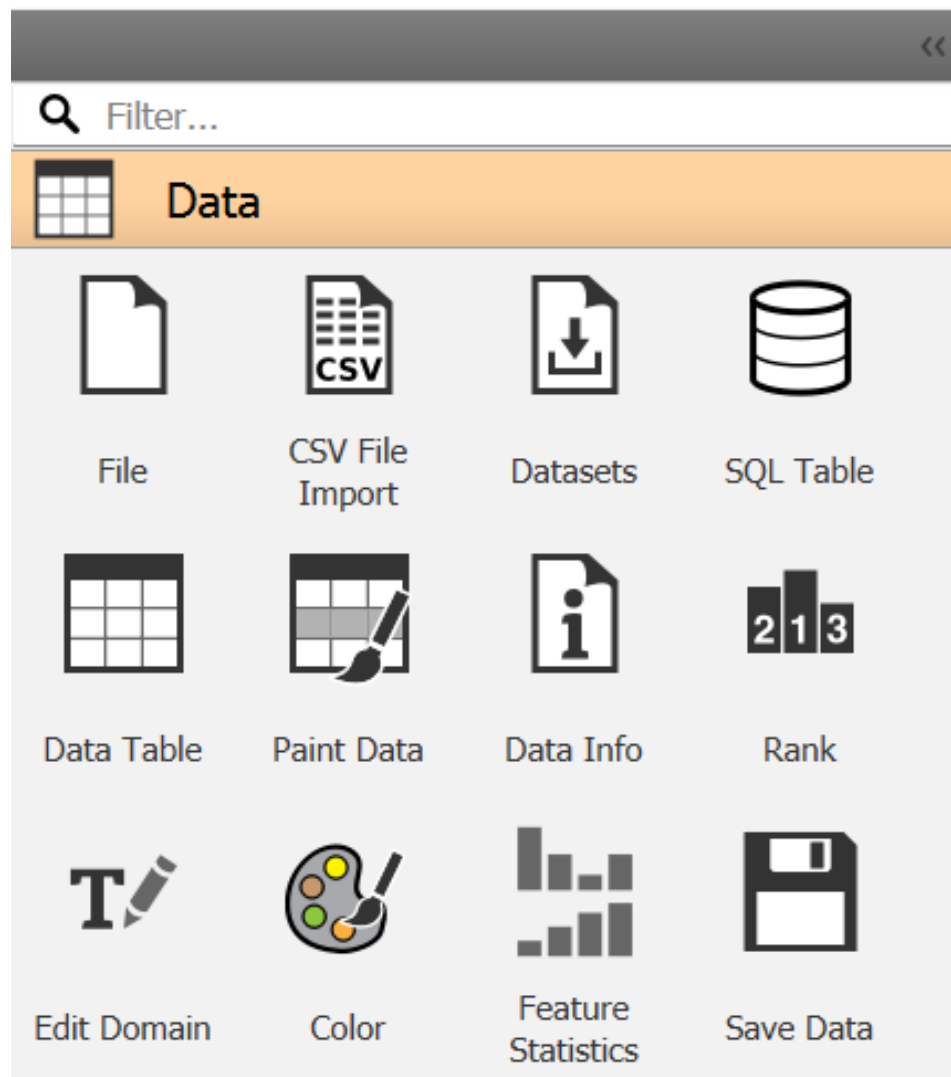
---

## Kā sākt savu ceļu mašīnmācīšanās pasaulē

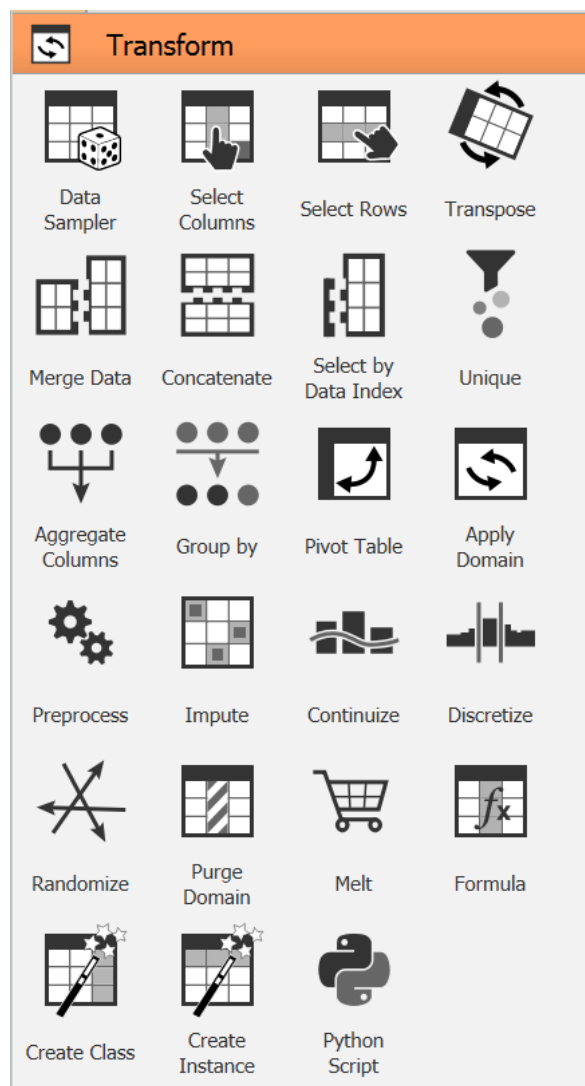
1. Noformulējam hipotēzi un apzinām kādas datu kopas var sniegt nepieciešamās atbildes
2. Izvēlamies piemērotu rīku kas spēj datus apstrādāt un vizualizēt
3. Veicam datu kopas vispārējo analīzi- identificējam atribūtus ar kuriem turpināt strādāt
4. Eksperimentēšana (dažādu algoritmu un metožu izmantošana)
5. Veicam rezultātu analīzi un modeļa pārbaudi



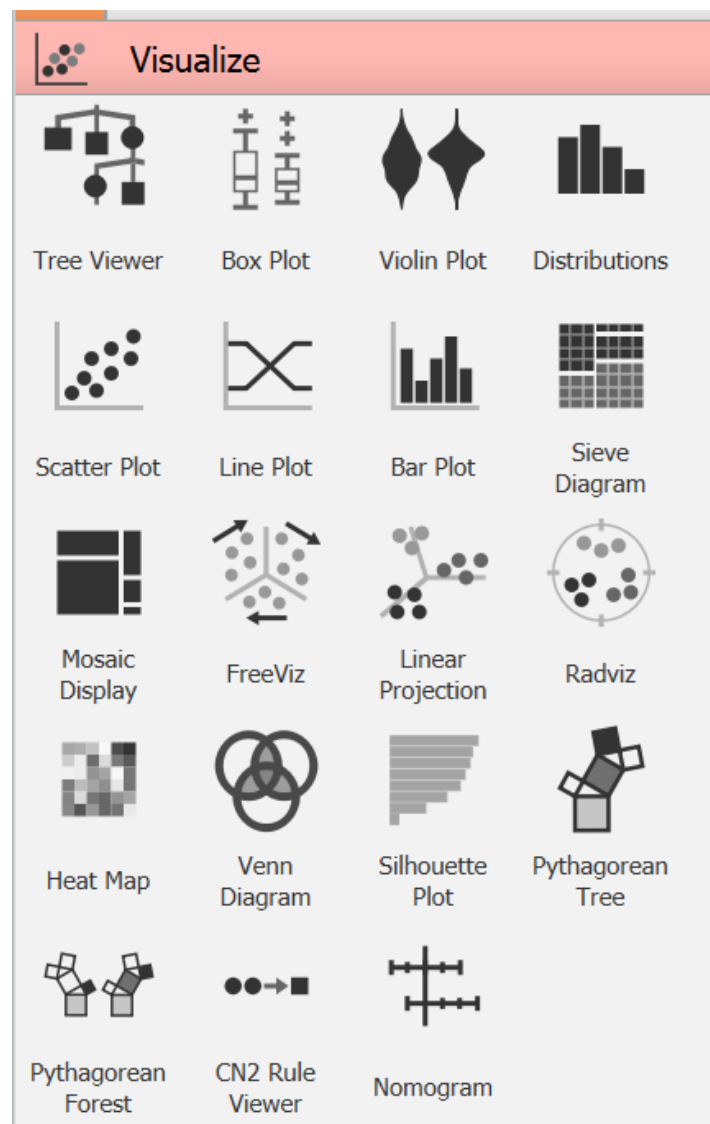
# Kādi var būt datu avoti?



# Ko ir iespējams veikt ar datiem?





















# Kā ir iespējams vizualizēt datus











# Datu modelēšana




















### Model

 Constant	 CN2 Rule Induction	 Calibrated Learner	 kNN
 Tree	 Random Forest	 Gradient Boosting	 SVM
 Linear Regression	 Logistic Regression	 Naive Bayes	 AdaBoost
 Curve Fit	 Neural Network	 Stochastic Gradient De...	 Stacking
 Save Model	 Load Model		

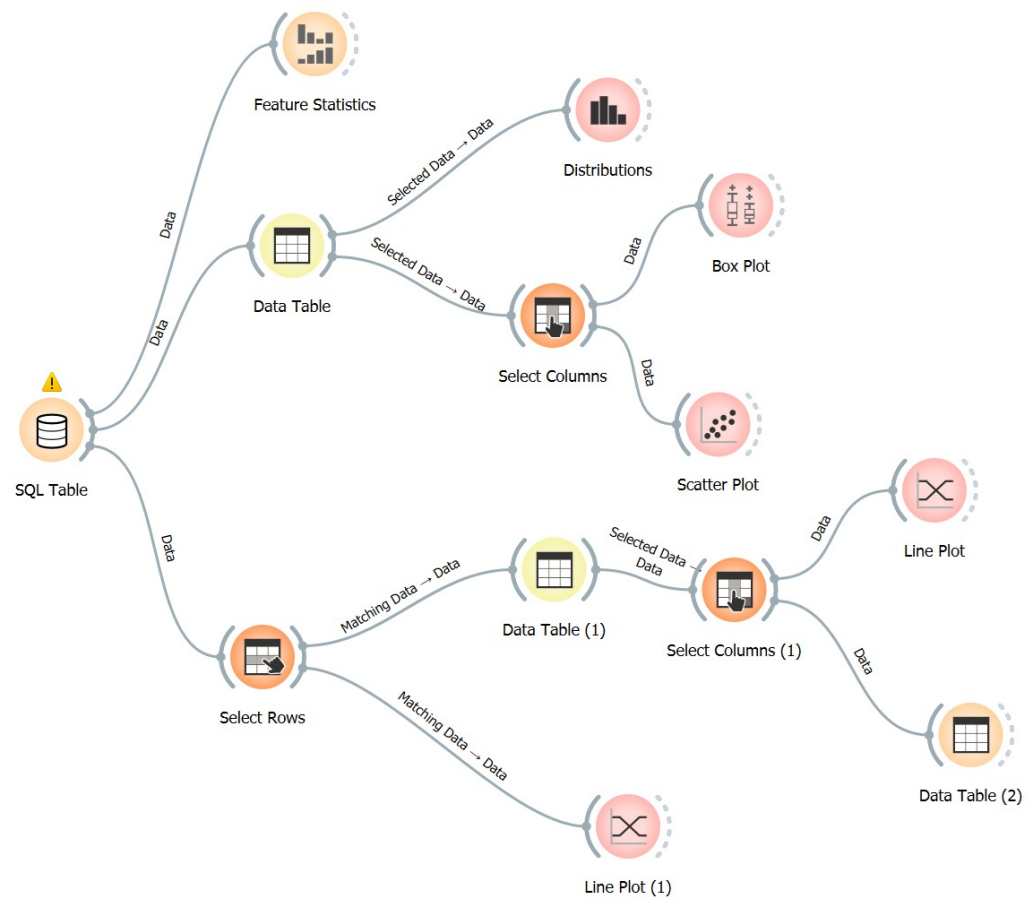
### Evaluate

 Test and Score	 Predictions	 Confusion Matrix	 ROC Analysis
 Performance Curve	 Calibration Plot		

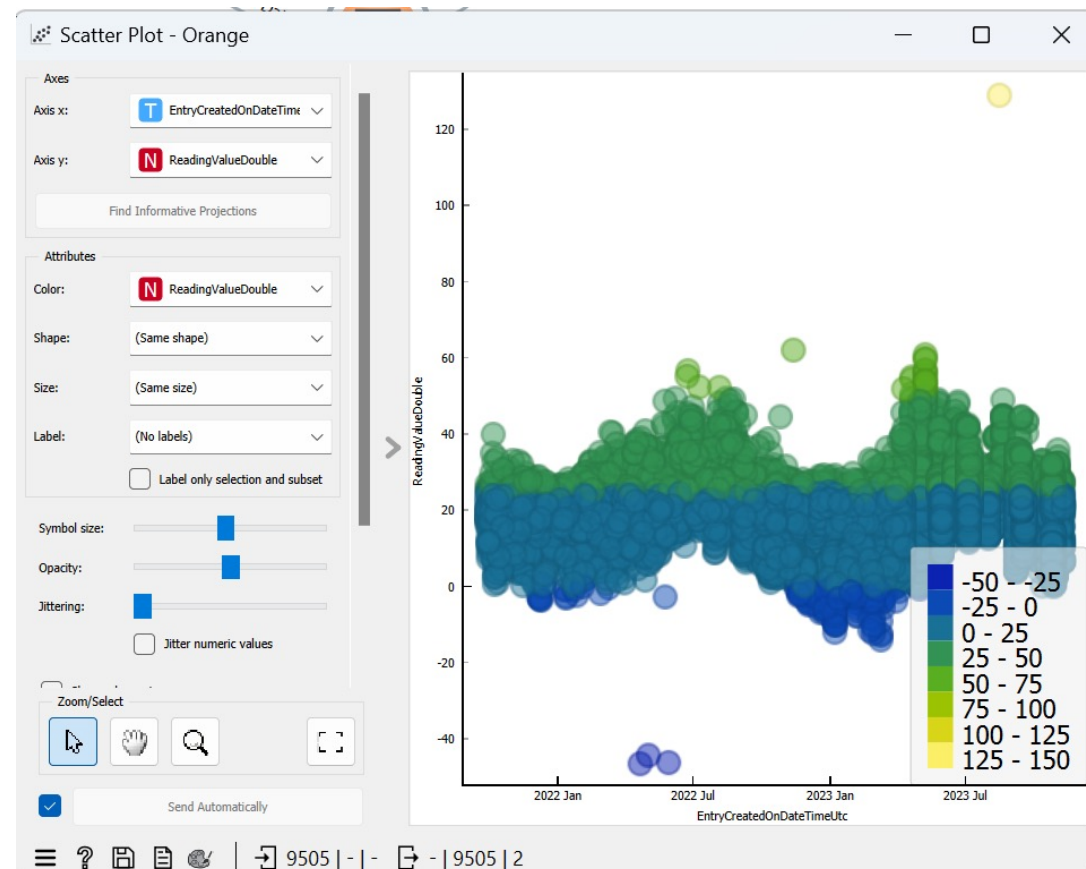
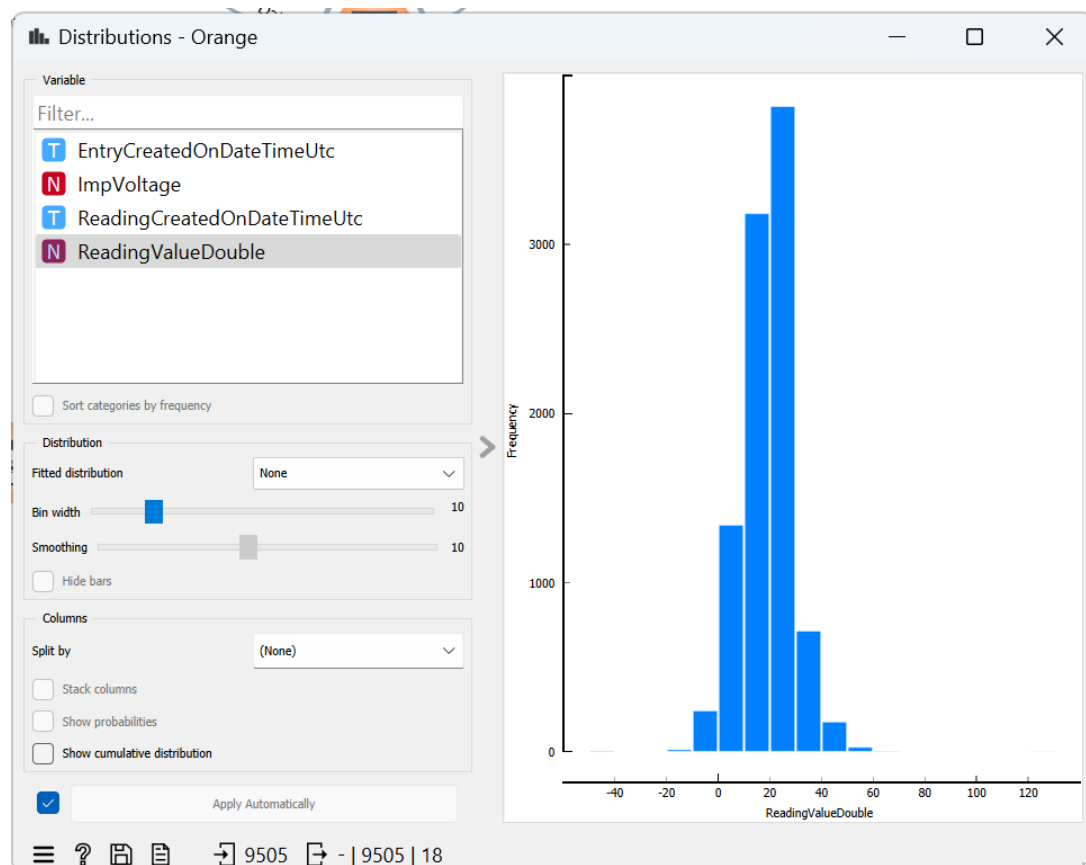
### Unsupervised

 Distance File	 Distance Matrix	 t-SNE	 Correlations
 Distance Map	 Hierarchical Clustering	 k-Means	 Louvain Clustering
 DBSCAN	 Manifold Learning	 Outliers	 PCA
 Neighbors	 Correspond... Analysis	 Distances	 Distance Transforma...
 MDS	 Save Distance Matrix	 Self-Organiz... Map	

# Darba plūsma

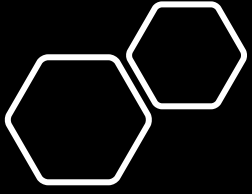


# Datu vizualizācijas piemērs



# Noslēgums

- Latvijā šobrīd ir nepietiekams, patstāvīgs un plašs superskaitļošanas pieprasījums no industrijas un valsts pārvaldes iestādēm
- Mākslīgā intelekta risinājumiem ir liels potenciāls procesu automatizācijas ieviešanā un pilnveidē
- Diemžēl organizācijas vēl pilnībā neapzinās savā rīcībā esošo datu un informācijas patieso vērtību
- Uzsākt apmācību un individuālus eksperimentus mašīnmācīšanās un mākslīgā intelekta tematikā var jau šobrīd un, izmantojot bezmaksas infrastruktūru
- Iesaku sākt Python programmēšanas valodas apgūšanu un piedalīties dažādās mūžizglītības digitālo prasmju apgūšanas programmās



# Paldies!



**EURO**



**Supers**

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 951732. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, United Kingdom, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Switzerland, Turkey, Republic of North Macedonia, Iceland, Montenegro.