



RIGA**CODING**SCHOOL

DATU ANALĪZES (ar PYTHON pamatiem)  
apmācības



# PAR MUMS

## SKOLA

Viena no pirmajām  
programmēšanas skolām Lietuvā  
un Latvijā

## FILIĀLES

Rīga, Vīļja, Klaipēda un  
Kauņa

## ABSOLVENTI

Vairāk nekā 700  
absolvētu



## KARJERAS CENTRS

Mēs mainam cilvēku pieredzi  
un profesionālo karjeru

## SADBĪBA

Vairāk nekā 70 sadarbības  
partneru

## PIEREDZE

Kopā 25 profesionāli treneru

# KURSI



Intensīvs un koncentrēts  
programmēšanas kurss



Apmācības sastāv no 20%  
teorijas un 80% prakses



Grupas tiek organizētas,  
atbilstoši zināšanu  
līmenim



Iespēja atkārtot kursu  
bezmaksas 1 gada laikā



Iespēja apmeklēt vieslekcijas  
ar mūsu sadarbības  
uzņēmumiem



Visas lekcijas tiek  
ierakstītas un pieejamas  
privātajā YouTube kanālā

# RĪGAS PROGRAMMĒŠANAS SKOLA - Karjeras centrs

- Tikšanās ar IT tirgus ekspertiem;
- CV un karjeras semināri;
- Palīdzam atrast prakses vietas IT uzņēmumos;
- Studenti tiek vērtēti apmācību laikā;
- Iespēja izmantot skolas telpas ārpus apmācībām;
- Izsniedzam **Sertifikātu** par 120 akadēmisko stundu apmeklējumu, kā arī pasniedzēju **rekomendācijas**.

# VALDIS



- Izglītība: Maģistra grāds datorzinātnēs
- Pieredze programmēšanā: 20+ gadi
- Specialitāte: grafu teorija sociālo tīklu analizēšanā
- Hobiji: prāta spēles, riteņbraukšana, šahs



# Data Lake



# Brief History of Data Analysis



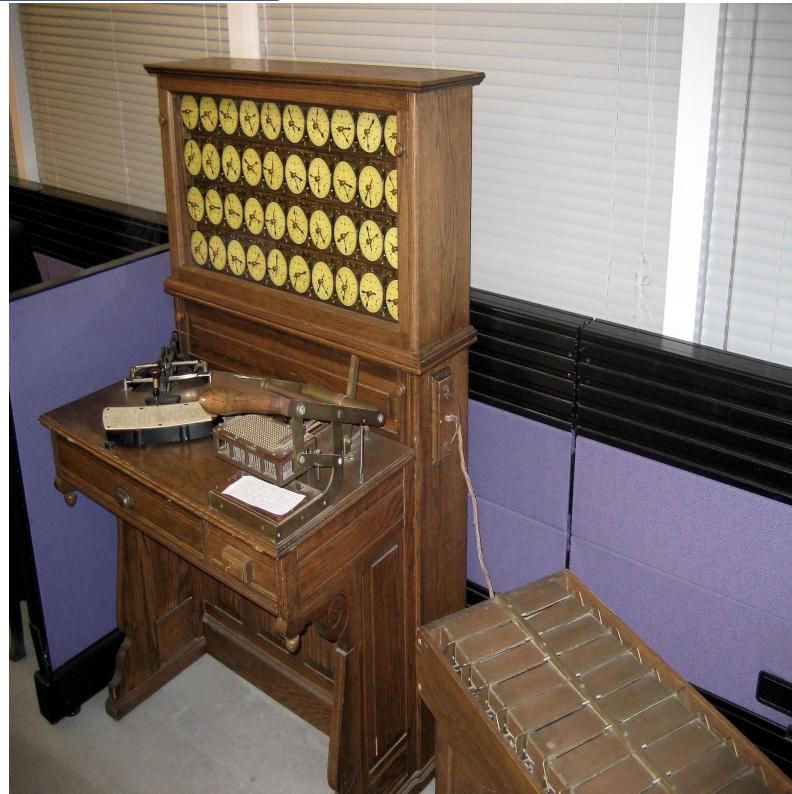
- ~ 18,000BC – Uganda, Ishango Bone
- ~ 2400BC – Babylon abacus, libraries
- 300BC – 48AD – Library of Alexandria
- ~ 100-200AD Antikythera Mechanism



# Brief History of Data Analysis II



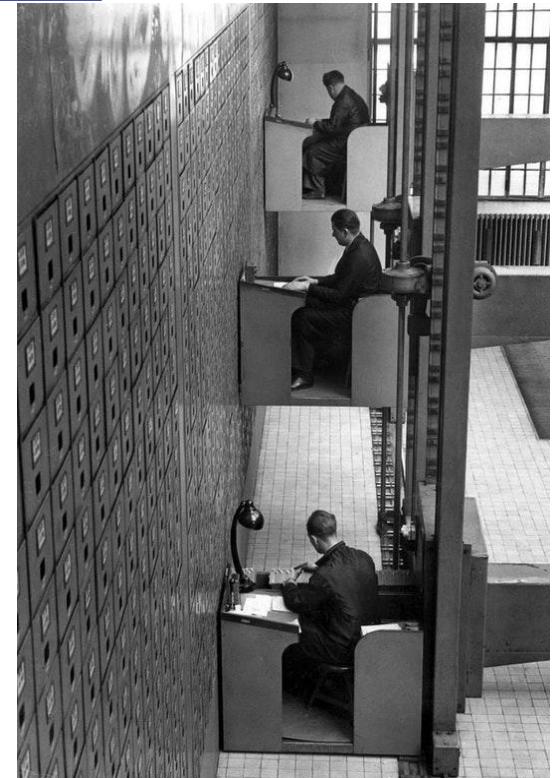
- 1663 – London, J.Graunt mortality analysis
- 1865 – banker H. Furnese business intelligence
- 1880-90 US Census Hollerith Machine -> IBM
- 1928 – F. Pfleumer magnetic tape invention



# Brief History of Data Analysis III



- 1950s - Flat Files
- 1958 – IBM's Luhn defines Business Intelligence
- 1960s - CODASYL
- 1970s – Codd's relational DBs -> SQL
- 1980s – Data Warehouses / Marts
- 2000s – Big Data / noSQL DBs



BIG DATA LANDSCAPE 2017



# Big Data

- Volume
- Velocity
- Variety
- Veracity



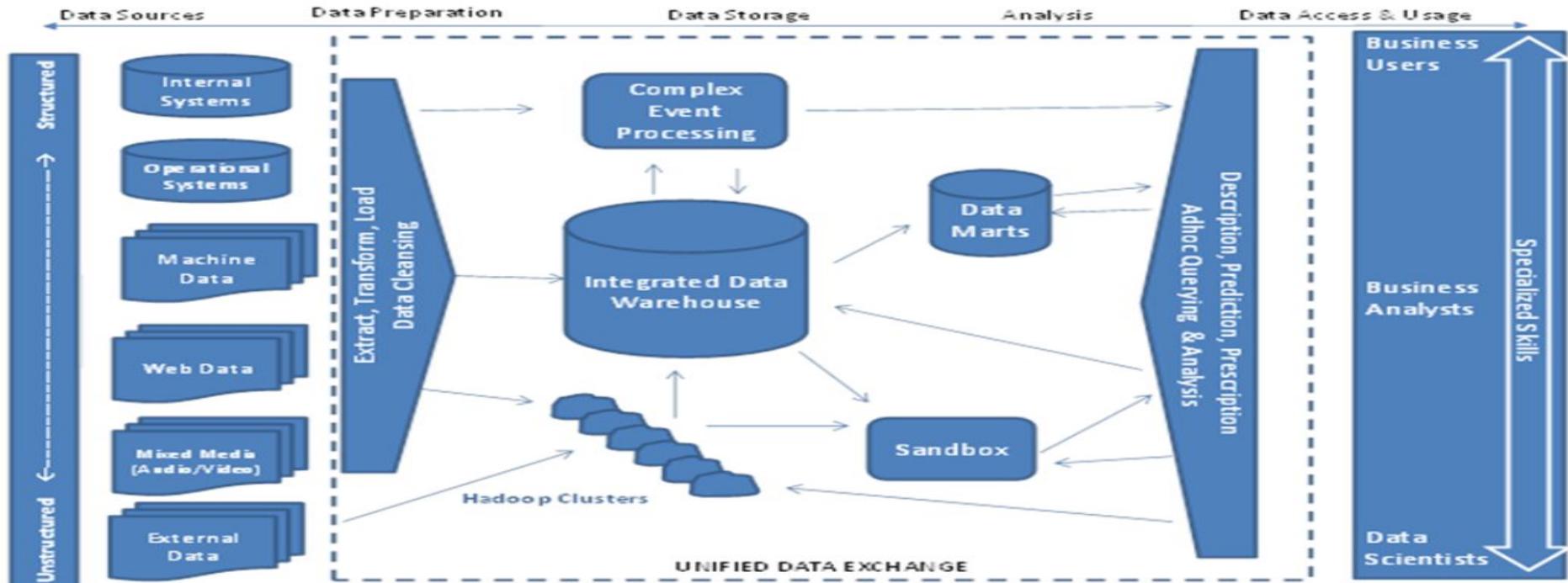
# Buzzword bingo

- Big Data
- Data Mining (datizrace)
- Machine Learning – subset of AI
- Data Science – statistics
- Big Data or Pokemon
- <https://pixelastic.github.io/pokemonorbigdata/>



Random Forests	Neural Network	Reinforcement Learning	Supervised Learning	Cognitive Computing
Caffe	Support Vector Machine	Artificial Intelligence	Python	Cloud
Unstructured Data	Bot	DATA SCIENCE BUZZWORD BINGO (free square)	K-means	GPU
Spark	Data Wrangling	Deep Learning	Ensemble	Machine Learning
Keras	Tensorflow	Big Data	Algorithm	Feature Engineering

# Full Analysis Framework



**BIG DATA MANAGEMENT and GOVERNANCE:** Strategic, Tactical and Operation Levels  
(Metadata, Data quality, Access, Use, Ethics, Privacy, and Security Management processes)

# Data Analysis Sandbox



# Data Mining



- Anomalies
- Classification
- Clusters
- Dimension Reduction
- Regression
- Relationship finding
- Summarization / Visualization



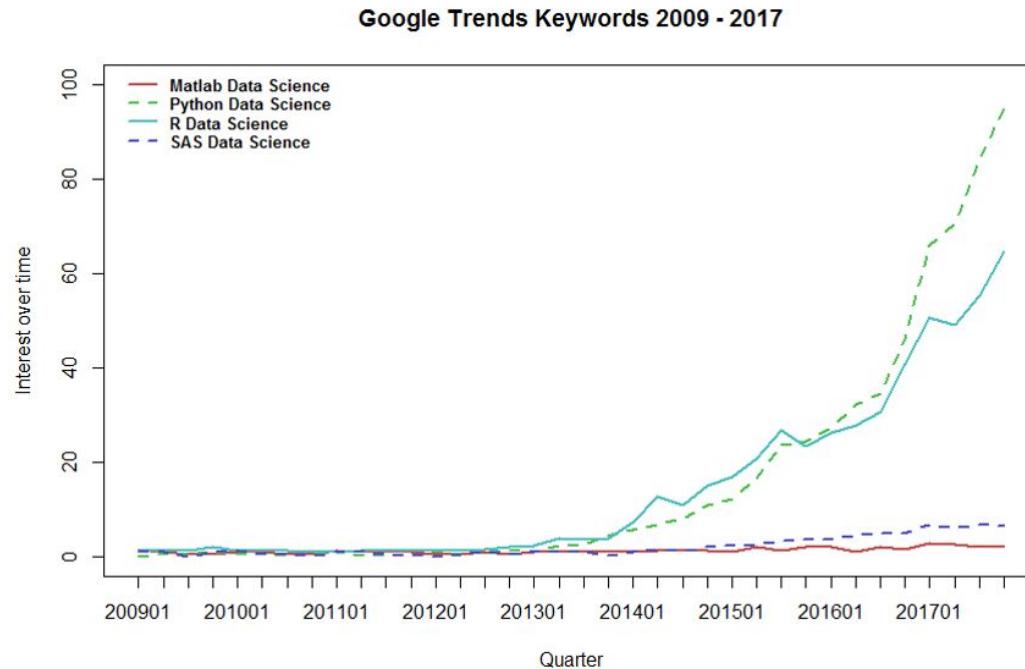
# Building a Pipeline



- Data Security
- Cleanup
- Organising Database
- Analysis
- Visualization – Dashboard
- Emphasis on Analysis less on Infrastructure



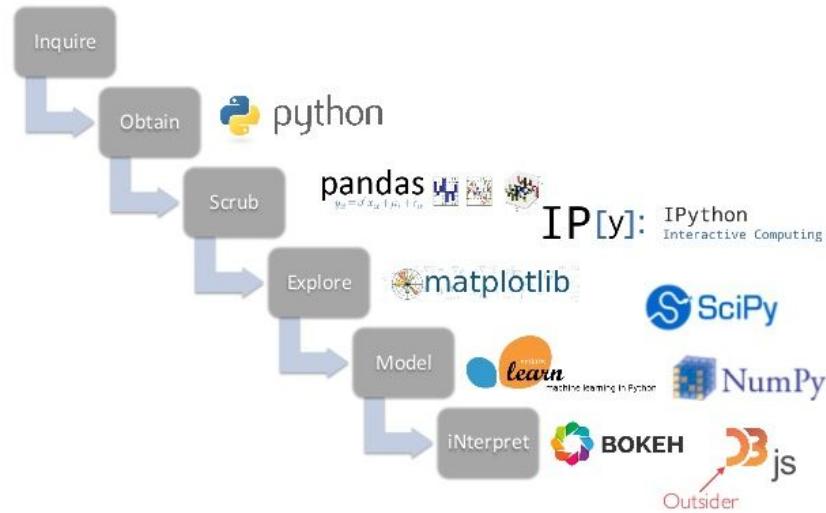
# Why Python?



# Python Ecosystem



PYTHON IS IOSEMN



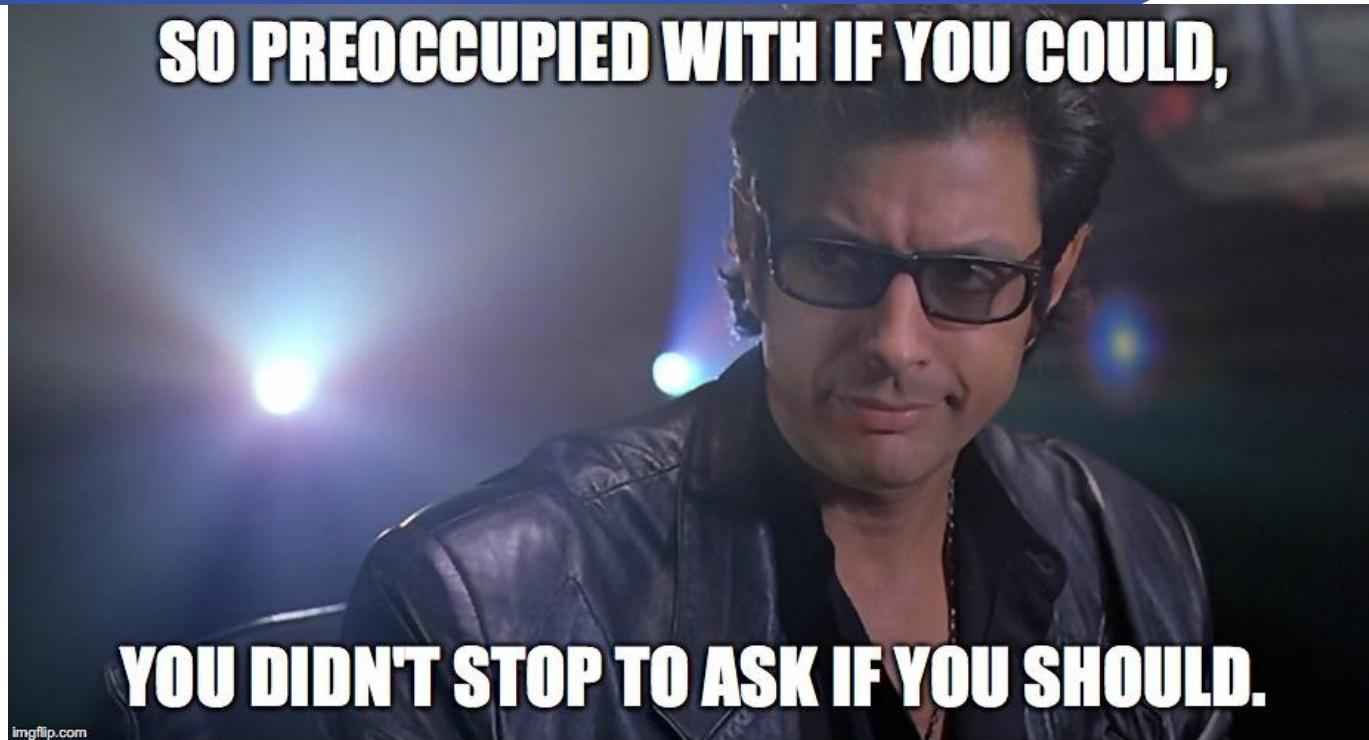
# Avoiding Pitfalls



- Overfitting
- Data Dredging / p-hacking
- <https://xkcd.com/882/>



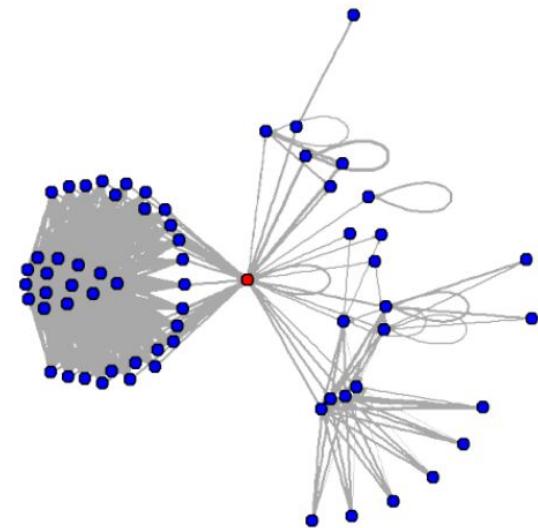
# GDPR, ethics



# Planned Projects



- Recommendation System / Churn Prediction
- Web Comments Sentiment Analysis
- Network Analysis (possibly some blockchain)
  
- Visualizations with PowerBI( or Tableau)
- Dashboards with Dash/plotly



# Requirements



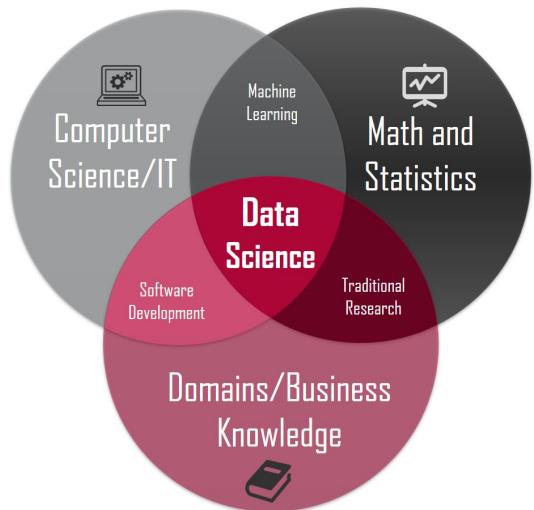
- Analytical / Logical Mind
- Helpful but NOT required knowledge:
- Python (will use Anaconda / Scipy)
- Comfortable in command line
- SQL
- Statistics
- Helpful: a computer with a minimum of 8GB RAM
- <https://www.anaconda.com/download/> (3.6+)



# Goals



- Access structured/unstructured data
- Clean data
- Apply correct methods for analysis
- Visualize Results



# Kādēļ Tu esi Šeit?



**Everybody in this country should  
learn to program a computer...  
because it teaches you how to think**

Steve Jobs, co-founder and CEO of Apple Inc. (1955 - 2011)





# Jautājumi?

