

Mastering Data Science Engineering

Roman Golovnya
27 April 2019



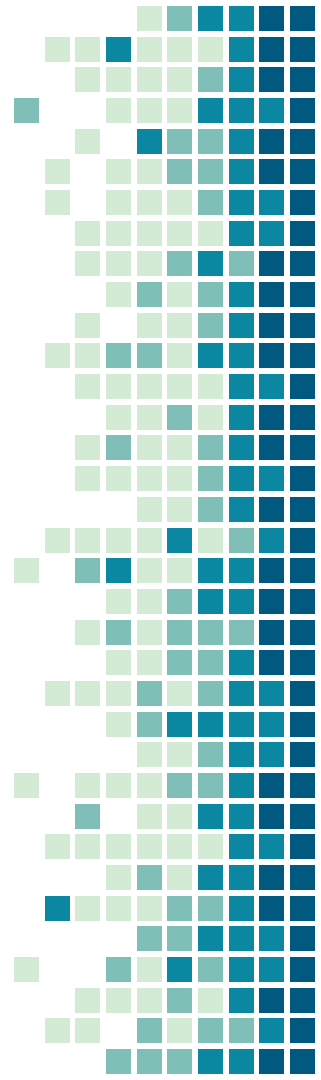
About me

Roman Golovnya

(<https://www.linkedin.com/in/romangolovnya>) graduated with degree in Finance and IT. Followed by post-graduate qualifications in Corporate Treasury, Data Analytics and Computer Science (ongoing). He is a life-learner and practitioner. Currently, he works in Hertz as Data Engineer. He develops scalable data solutions using python, Apache Spark and AWS services. In a free time, he organises the events for this meetup group, plays table tennis and participate in Kaggle competitions.



The goals of the Data Science Engineering Fundamentals program is to review Data Science fundamentals via presentations and workshops. We will have a good mix of theory and hands-on practical sessions. The participants will be encouraged to share their experience and work on the projects of their choice.



In the last 12 months – 18 events

- Data Science Engineering Fundamentals – Toolset, Git, Docker
- Data Wrangling – Programming with Python
- Getting and importing datasets from multiple sources
- Data Science and Engineering Club @DogPatch Labs-Data Visualisations
- Building Web Applications
- Applied Math, Probability and Statistics for Data Science
- Cloud Computing & Feature Engineering
- Data Science Engineering Dojo Financial Data – Segmentation project
- Data Engineering Workshop – Terraform
- Machine Learning Workshop
- Deep Learning Workshop
- Data Engineering Workshop – Apache Spark



Results:

- Financial data - Segmentation project
- As a team, won Silver medal at Kaggle competition
- Partnership with european tech conferences



Mastering Data Science Engineering

- Work in teams to build Data Science products - MVP
- Rapid prototyping - fast POC
- Apply best practices in Data Engineering and Data Science
- Using the latest technologies and techniques
- Be technology agnostic and agile
- Present your MVP to audience and potentially to investors
- Participate and win medals in kaggle competitions
- Have fun!



Future events topics:

- Building Visualisations and Dashboards Dojo
- Real-time data pipelines
- Contenerizations of the apps
- Advance Cloud computing services
- Machine Learning



What's next?

- Join or form a team
- Based on data related project topic
- Programing languages and technology stack
- Skills level
- Interests



Kaggle team

- Data Analyst(s) (EDA, Insights - know data, feature engineering, visualisations)
- Data Engineer (modularize code, create pre-processing and ML pipelines, cloud and tech support)
- Data Scientist - Machine Learning Engineer (modeling, leading the team)

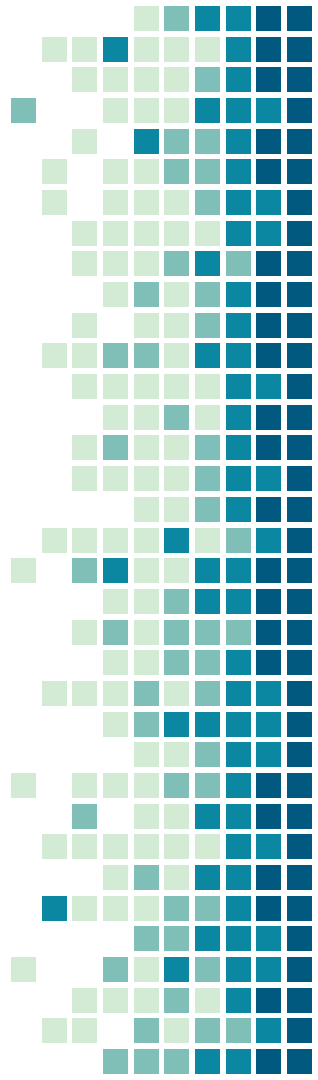


Data Science Engineering project – team

- Data analyst(s) and Visualisation specialist(s) (Dashboards)
- Junior Data Engineer(s) (pipelines, databases)
- Software Engineer (Front-End, APIs)
- Data Engineer (Cloud, Scalable Computing, Microservices)
- Data Scientist - Machine Learning Engineer
- Data Science Engineer - Data Architect

Resources:

- <https://www.kaggle.com/competitions>
- <https://eu.udacity.com/course/data-engineer-nanodegree--nd027>
- <https://www.coursera.org/specializations/gcp-data-machine-learning>
- <https://www.coursera.org/learn/machine-learning>
- <https://www.coursera.org/learn/competitive-data-science>
- <https://www.edx.org/course/subject/data-analysis-statistics>
- <https://www.datacamp.com/>
- <https://www.udemy.com/learn-devops-the-complete-kubernetes-course/>
- <https://aws.amazon.com/training/learning-paths/machine-learning/>



Next event – Data Visualisation Dojo

- You can apply any technology
- Building Visualisations and Dashboards
- Kaggle competitions data and kaggle datasets
- Open data sources
- Data from data pipelines
- Real-time data



Contact me:

roman.golovnya@gmail.com

<https://www.linkedin.com/in/romangolovnya>

[Via meetup.com/messages](https://www.meetup.com/messages)

<https://twitter.com/rgolovnya>

<https://github.com/dseclub>



Takeaways:

- We need you.
- Start small and iterate
- Do not be afraid to fail, learn from mistakes
- Join or form a team
- Choose together a project
- Have fun!

Thank you!

Any Questions?