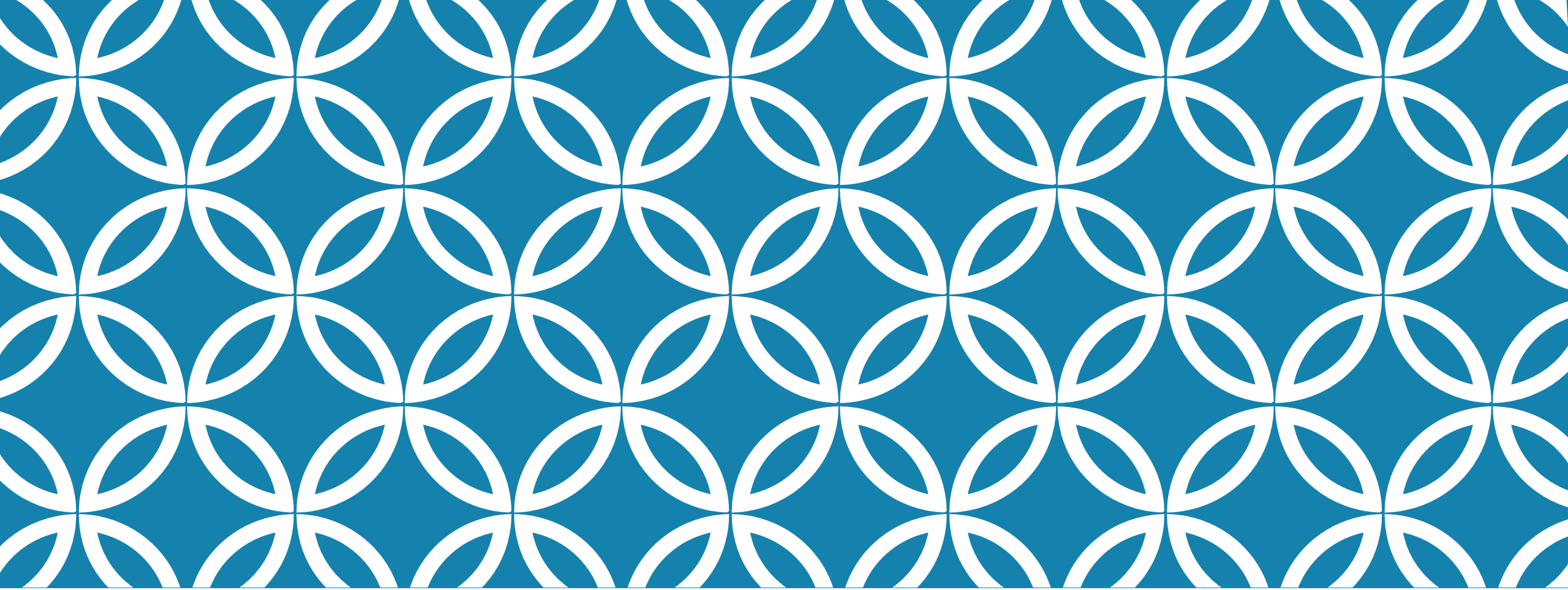


HOW TO GET A "REAL JOB" IN DATA SCIENCE

Dr. Anastasia Karavdina

MY CAREER PLAN BACK IN UNIVERSITY

WHAT IS DATA SCIENCE

Data science is an interdisciplinary academic field that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data.

(Wikipedia)

My definition:

Data Science is a field, where Data is used to automize and/or quantify decisions

DATA SCIENCE IS EVERYWHERE

- **E-commerce:** Personalized product recommendations, optimizing sales strategies, and inventory management
- **Healthcare:** Predictive analytics for disease outbreaks, early stage cancer diagnosis, patient care improvement, and drug discovery
- **Finance:** Fraud detection, credit scoring, and algorithmic trading
- **Transportation:** Optimizing routes for delivery, predicting vehicle maintenance, and autonomous vehicle algorithms
- **Energy:** Predicting equipment failures, optimizing grid performance, and renewable energy forecasting
- **Retail:** Supply chain optimization, customer purchase behavior, and store layout optimization

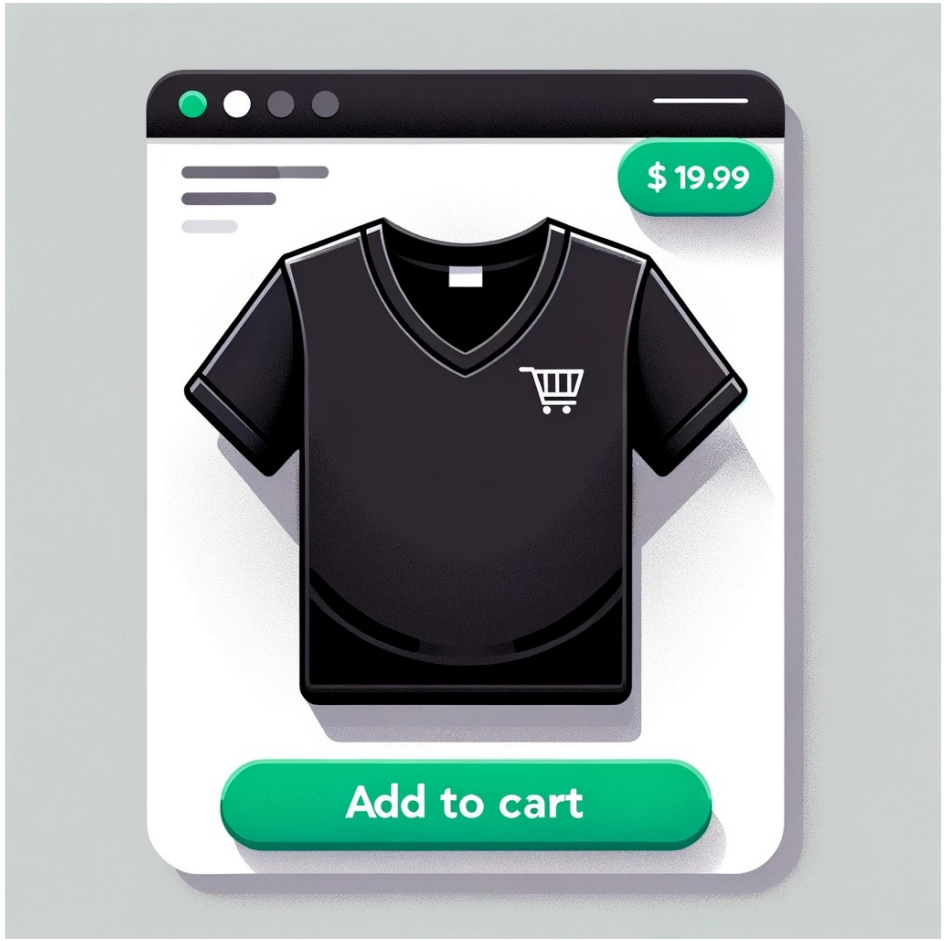
REMEMBER THIS T-SHIRT YOU BOUGHT WITH HUGE DISCOUNT?

Pricing is complex topic. Goals:

- remain atractive for customers
- sell all stocks by the end of the season
- maximize the profit

Two step solution:

1. Demand model (price, time of year, t-shirt)
2. Optimal price for the business objective based on demand model



T-SHIRT PRICING SOLUTION STEPS

HISTORICAL DATA

Data collection
Data pipeline

Data Analysis

T-SHIRT PRICING SOLUTION STEPS

HISTORICAL DATA ⟶ **DEMAND MODELING**

Data collection
Data pipeline

Data Analysis

ML modeling

T-SHIRT PRICING SOLUTION STEPS

HISTORICAL DATA → **DEMAND MODELING** → **OPTIMAL PRICE**

Data collection
Data pipeline

Data Analysis
ML modeling

Optimisation with constraints

T-SHIRT PRICING SOLUTION STEPS

HISTORICAL DATA → **DEMAND MODELING** → **OPTIMAL PRICE** → **FUNCTIONALITY IN PRODUCTION**

Data collection
Data pipeline

Data Analysis
ML modeling

Optimisation with constraints

New system design or incorporation into existing system

T-SHIRT PRICING SOLUTION STEPS

HISTORICAL DATA → **DEMAND MODELING** → **OPTIMAL PRICE** → **FUNCTIONALITY IN PRODUCTION** → **PERFORMANCE MEASURMENT**

Data collection
Data pipeline

Data Analysis
ML modeling

Optimisation with constraints

New system design or incorporation into existing system

Data Analysis

ROLES INVOLVED INTO T-SHIRT PRICING SOLUTION

HISTORICAL DATA → **DEMAND MODELING** → **OPTIMAL PRICE** → **FUNCTIONALITY IN PRODUCTION** → **PERFORMANCE MEASURMENT**

Data collection
Data pipeline

Data Analysis
ML modeling

Optimisation with constrains

New system design or incorporation into existing system

Data Analysis

Data Engineer

Data Analyst

Data Scientist

Machine Learning Engineer

Product Manager (Data-Driven Products)

Data Analyst

Business Intelligence Analyst

ROLES INVOLVED INTO T-SHIRT PRICING SOLUTION

HISTORICAL DATA → **DEMAND MODELING** → **OPTIMAL PRICE** → **FUNCTIONALITY IN PRODUCTION** → **PERFORMANCE MEASURMENT**

Data collection
Data pipeline

Data Analysis
ML modeling

Optimisation with constrains

New system design or incorporation into existing system

Data Analysis

Data Engineer

Data Analyst

Data Scientist

Machine Learning Engineer

Product Manager (Data-Driven Products)

Data Analyst

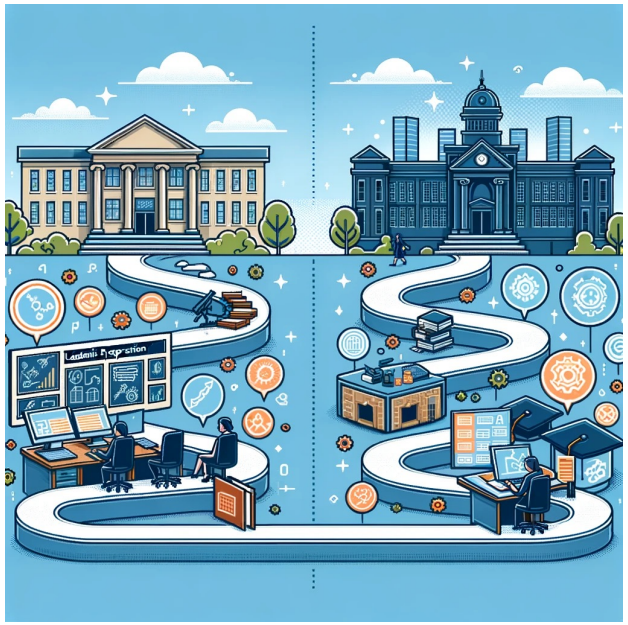
Business Intelligence Analyst

Former physicists also excel as Data Science Consultants

HOW HAS ACEDEMIA ALREADY PREPARED ME FOR THIS ROLE?

Making it through a Ph.D. (and beyond) requires significant mental and emotional growth.
If you're surviving the challenges of independent research, then you surely possess:

- The ability to distill interesting things from large amounts of information
- Structural thinking
- Extreme attention to detail
- Strong planning and organization attitude
- Written and verbal communication
- Serious mental fortitude to push through (and learn from) failure



WHERE ACADEMICS STRUGGLE?

For successful transition to industry you will have to unlearn some of the mentality inherited in fundamental research

- The speed-accuracy tradeoff
- Implementing the results of an analysis
- Focusing on team above self
- Expanding skillset in Computer Science

REACH OUT FOR HELP

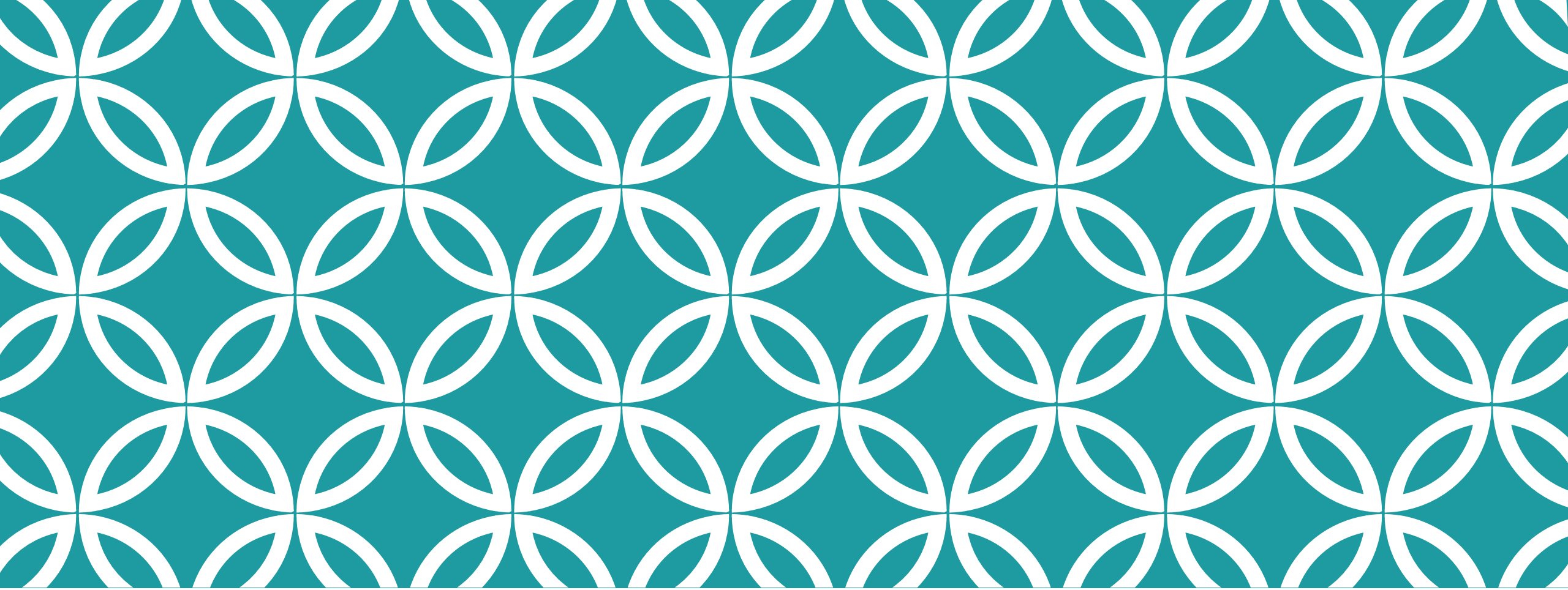
- Find someone, who already working in Data Science with similar backgroud as yours and ask them how they did it
- Ask if a senior member of the group knows someone who might be hiring
- Inform yourself about companies around you: what they do, which roles they have
- Go to workshops, meet-ups, etc
- Hire a mentor

REACH OUT FOR HELP

- Find someone, who already working in Data Science with similar backgroud as yours and ask them how they did it
- Ask if a senior member of the group knows someone who might be hiring
- Inform yourself about companies around you: what they do, which roles they have
- Go to workshops, meet-ups, etc
- Hire a mentor

Thank you for your attention!

Let's connect on LinkedIn: Dr. Anastasia Karavdina



BACK UP

SUITABLE ROLLES FOR ACADEMICS

- Data Scientist
  - Applies ML, statistics, etc to make predictions for the future based on the past
- **Data Analyst**
  - Transforming data into actionable insights
- Business Intelligence Analyst
  - More focus on dashboards and visualizations
- Machine Learning Engineer
  - Implements ML systems, SE with ML focus
- Data Engineer
  - Builds and maintains the infrastructure for data collection&processing
- (Applied) Research Scientist
  - Works on cutting-edge research, very strong knowledge of the field expected
- Product Manager (Data-Driven Products)
  - Defines product strategies
- **Data Science Consultant**
  - Assists organizations in applying DS, often implement solutions tailored to the client's need