AI4D-Lab Tanzania - The University of Dodoma, CIVE

Artificial Intelligence Short Course

Launching a Career as a Data Scientist

Zephania Reuben

June 28, 2022

Outline

1. Build your Career as a Data Scientist
   - How do I land a career as a data scientist?
   - Learning approaches
   - Building portifolio of real world projects
2. Roles and Requirements
   - How much math should I learn for DS / ML?
   - What makes a good data scientist?
   - Am I too old / too young to become a data scientist?
3. Future Proofing Your Career
   - Libraries / pre-existing solutions, or code from scratch?
   - How can I future-proof my skills and career?
   - Making money from home / Are there remote opportunities?
4. General Advice for
   - People/students with business/computer studies background
   - Someone with no relevant work experience

Build your Career as a Data Scientist

What is Data Science?

Data Science is a multi-desciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data. Wikipedia

How do I land a Career as a Data Scientist?

All Data Scientists

- Data Analysis / Exploratory Analysis
- Data Preprocessing
- Applied Machine Learning

How do I Land a Career as a Data Scientist?

Business Data Scientists

- Insight knowledge
  - Marketing
  - Strategy
  - Operations
  - Inversiting
- Domain knowledge
- Communication and presentations

How do I land a Career as a Data Scientist?

Product Data Scientists

- Product knowledge
  - E-commerce
  - Entertainment
  - Banking
  - SaaS
- Software development basics
- Data pipelines

Building Portifolio of Real World Projects

Steps

1. Learn the core skills for data science and applied machine learning - Roadmap
2. Pick out a dataset to start with - Datasets
3. Start your project in Jupyter Notebook.
4. Explore the data and make sure you understand the features.
5. Define an interesting objective to pursue
6. Clean the data, engineer features.
7. Complete your analysis / train your models.
8. Write about your project directly inside your Jupyter notebook.
9. Upload your project to Github. Repeat 1 - 9

Roles and Requirements

How Much Mathematics Should I Learn for DS / ML?

- Probably very little Mathematics foundation.
- Depends on your goal and learning approach.

Roles and Requirements

What Makes a Good Data Scientist?

- Always drive toward business value.
- Develop elite communication skills.
- Know when when NOT to use ML.
- Understand tradeoffs / deadlines.
- Focus on feature engineering.
- Consider themselves lifelong students

Roles and Requirements

Am I too old / too young to become a data scientist?

Anyone can become a data scientist, regardless of age, educational background, or prior work experience.

- Develop real skills.
- Build a portfolio of projects that help to prove the real skills.

Future Proofing Your Career

Career Paths in Data Science

- Level Up
- Choose your own adventure

Future Proofing Your Career

Libraries / pre-existing solutions, or code from scratch?

For learning purposes, you can choose to code a few of your favorite algorithms from scratch.

Future Proofing Your Career

Coding From Scratch - Advantages

- Learning how the algorithms works.
- Customizable implementation.
- Potentially faster implementations.

Coding From Scratch - Disadvantages

- Higher math programming requirements.
- Takes a long time.
- Difficult to beat pre-existing libraries.

Future Proofing Your Career

Libraries and Pre-existing Solutions - Advantages

- Easier to learn.
- Much more commercial demand.
- Pre-optimized implementations allow you to focus on the application and building better models.

Libraries and Pre-existing Solutions - Disadvantages

- Cannot customize implementations.
- Cannot see each step of the algorithms.
- Limited in functionality by what's already there in the pre-existing library.

Future Proofing Your Career

How can I Future-proof my Skills and Career?

- Opportunity assessment
- Creativity
- Domain expertise
- Nuanced decision making
- Empathy

Future Proofing Your Career

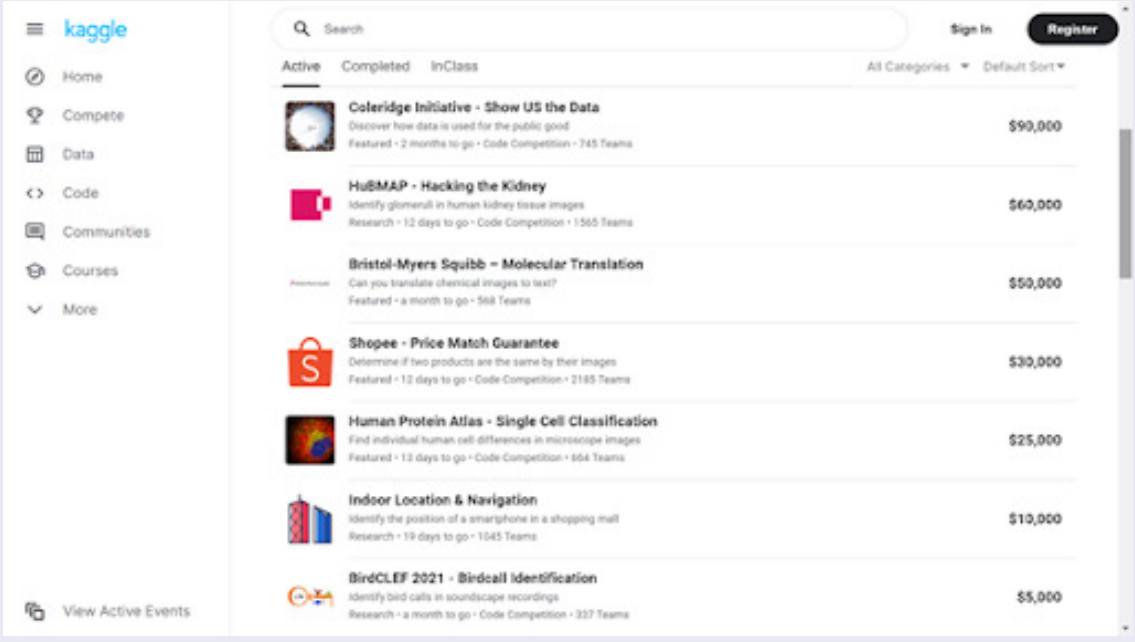
Data Science Competitions

- Data Science Competition is the process of participating in data science challenge to solve complex business problems and share top solutions to the organizations that own the data.
- Data science competitions help organizations find solutions to their complex business problems while enabling data scientists to learn from the experience and win awards/prizes.

Future Proofing Your Career

Data Science Competition Platforms

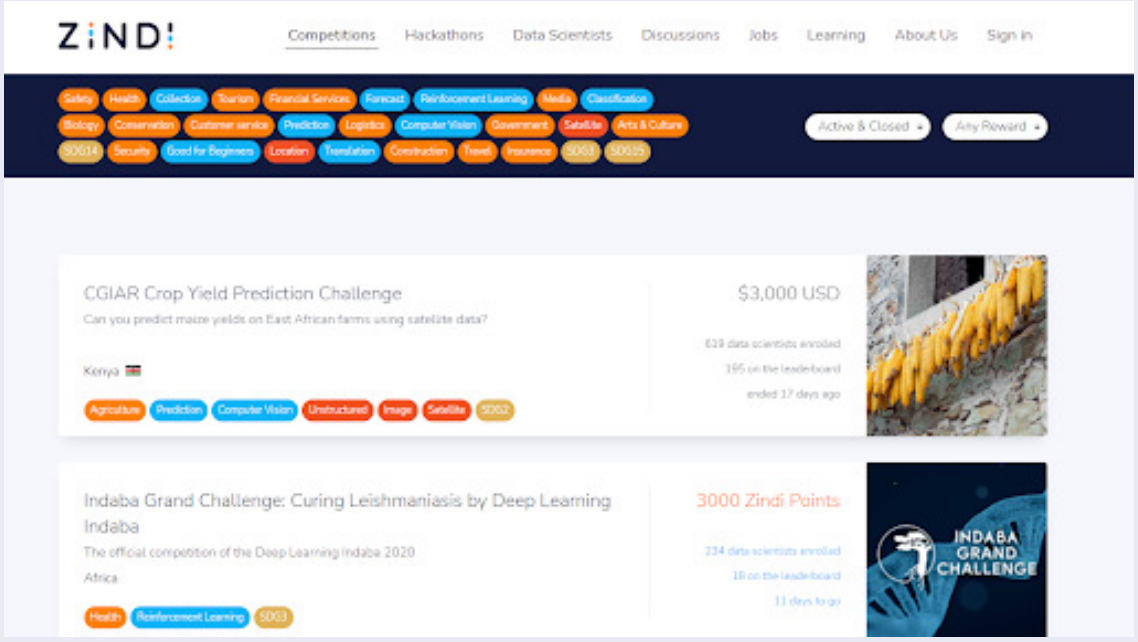
- Kaggle (Level: Beginner, Intermediate Advanced)



Future Proofing Your Career

Data Science Competition Platforms

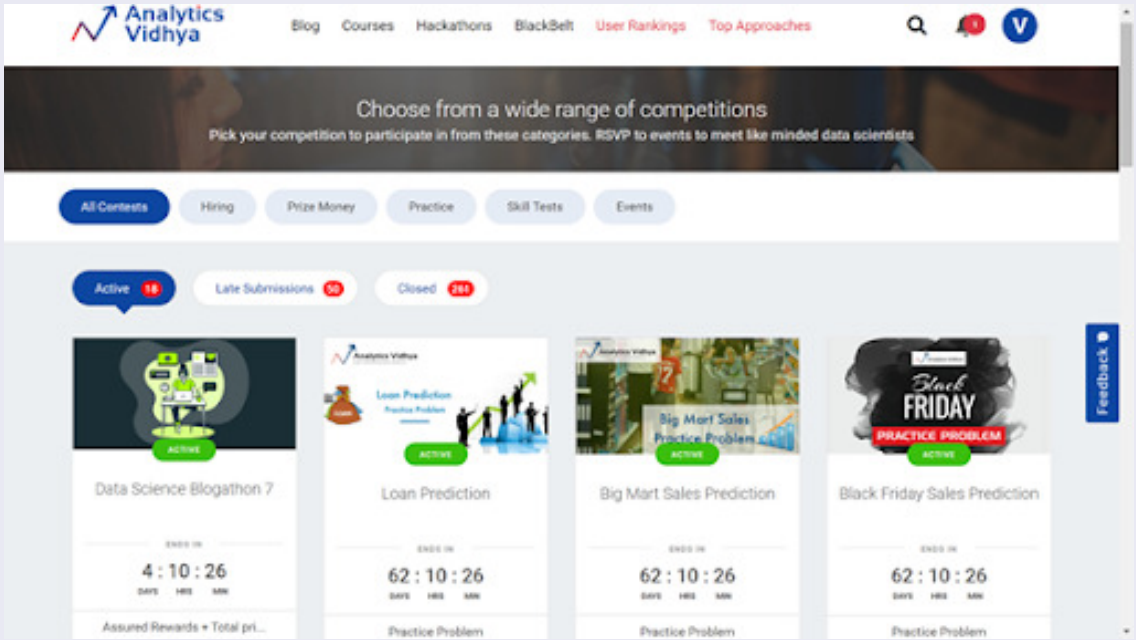
- Zindi Africa (Level : Beginner, Intermediate Advanced))



Future Proofing Your Career

Data Science Competition Platforms

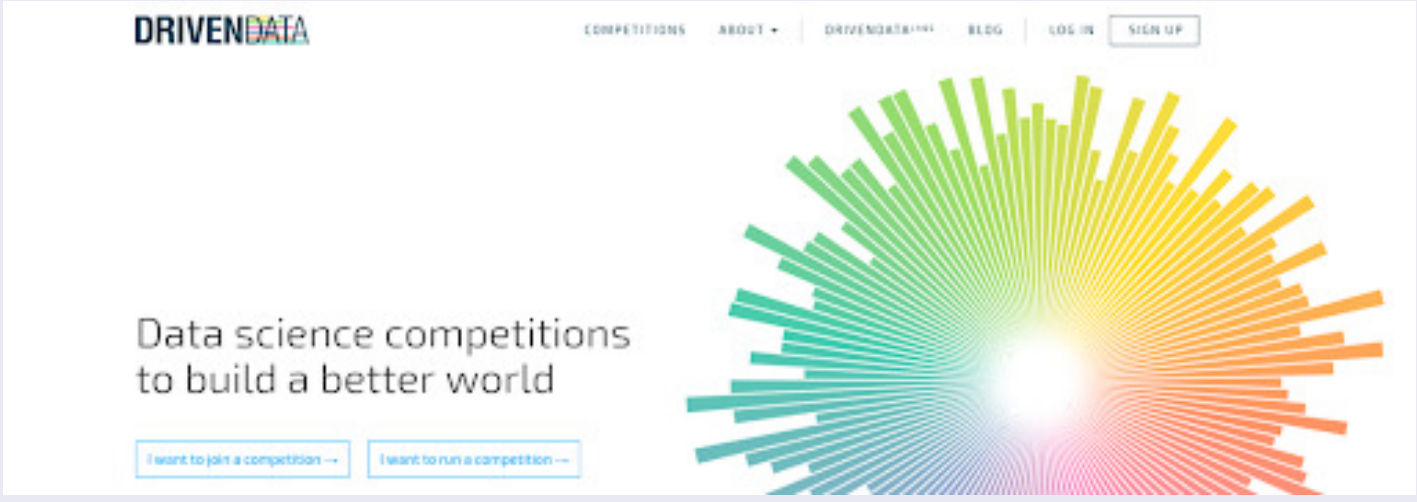
- Data Hack by Analytics Vidhya (Level: Beginner Intermediate))



Future Proofing Your Career

Data Science Competition Platforms

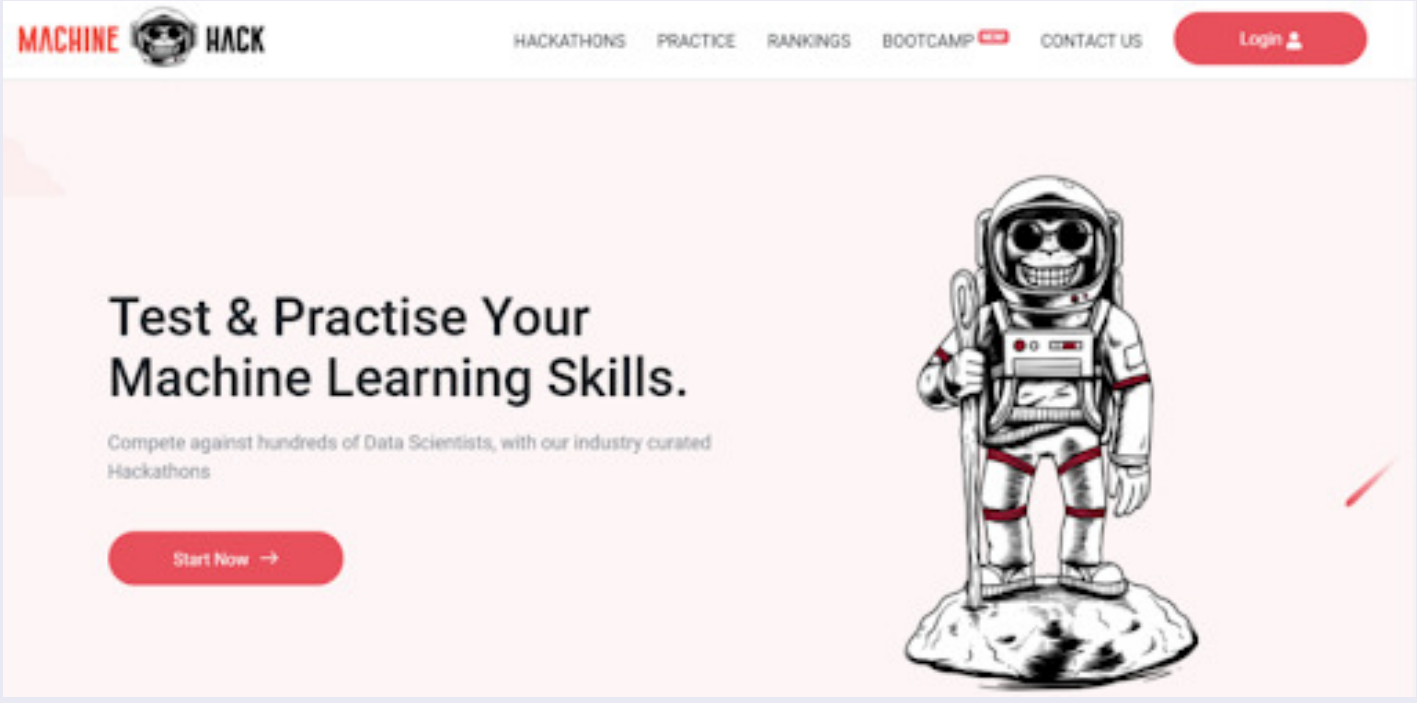
- DrivenData(Level: Beginner, Intermediate Advanced))



Future Proofing Your Career

Data Science Competition Platforms

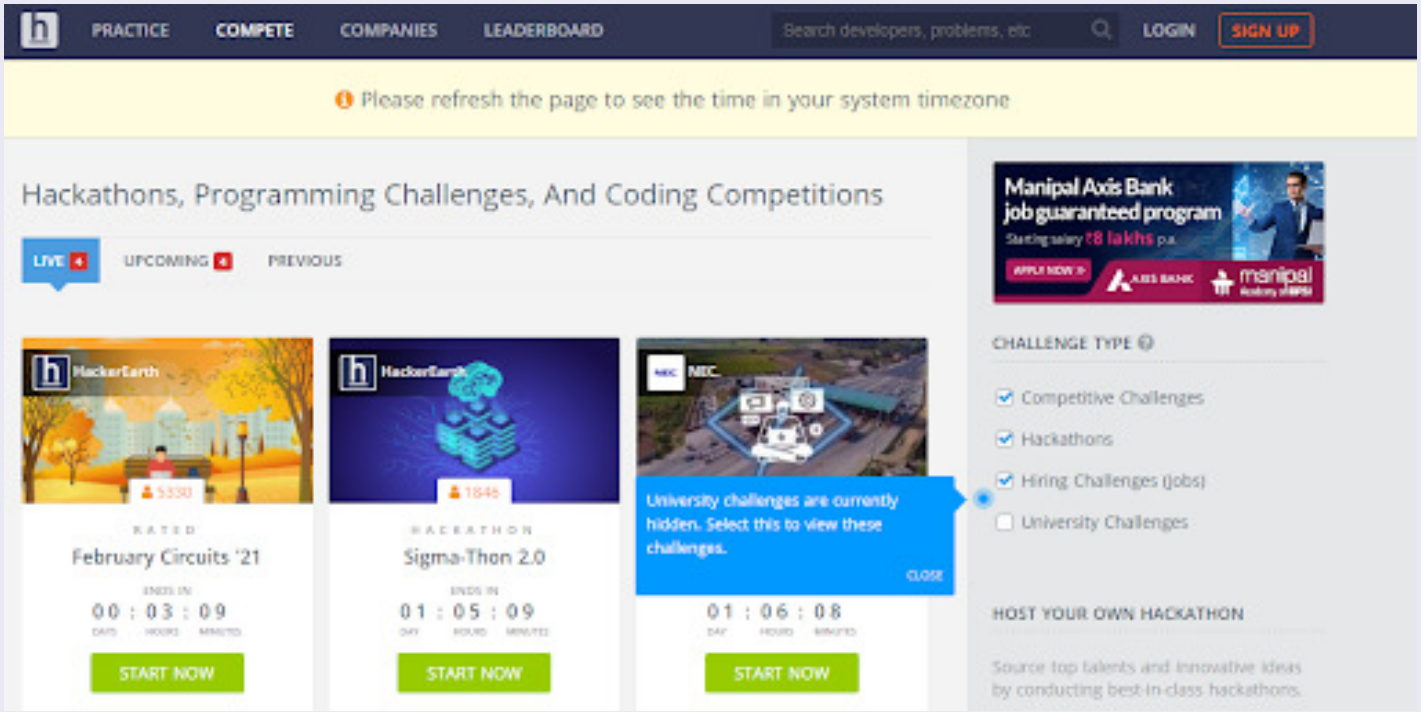
- Machine Hack (Level: Beginner and Intermediate))



Future Proofing Your Career

Data Science Competition Platforms

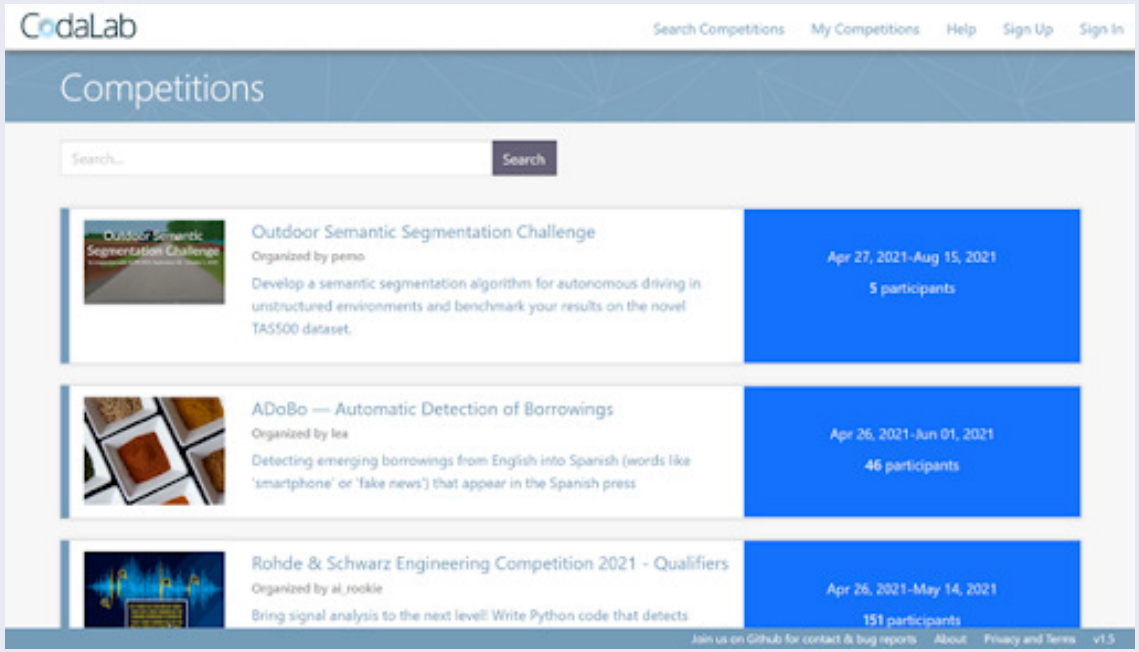
- HackerEarth(Level: Beginner and Intermediate)



Future Proofing Your Career

Data Science Competition Platforms

- CodaLab(Level: Advanced))



Future Proofing Your Career

Benefits of Participating in Data Science Competitions

- Test your skills against Top Talent.
- Learn by doing and gain exposure.
- Earn income by doing what you enjoy and do best.
- Build your Profile and attract potential employers.
- Apply for Data Science job opportunities.
- Learn from others through collaboration and discussion.
- Networking with like-minded people.

Future Proofing Your Career

Types of Data Science Competitions

- Prize Competitions.
- Points or Medals Competitions.
- Knowledge Competitions. Recommended for beginners

Future Proofing Your Career

Making Money From Home / Are There Remote Opportunities?

- Freelancing: UpWork, Toptal
- Competitions and Hackathons e.g Zindi.
- Working with enterprise customers/ small businesses to solve really hard customer problems in the various domains. Doing consultancy work in data science.
- Work as part of a core ML/ Research team in an academic team at a major university.
- Writing technical articles, data science educational practice, and data blogging to teach, earn, and learn as well.
- Public speaking at paid large tech/ Machine Leanring conferences, workshops, bootcamps, and meetups.

General Advice for

People/Students with IT/Business Studies Background

1. Data science is not only machine learning, analytical skills are crucial.
2. Skip most of the Math.
3. Be prepared for the mindset difference between software development and data science.
4. Domain knowledge can help you stand out big time
5. Practice your communication skills.

General Advice for

Someone with no Relevant Work Experience

- Develop the real skills capable of driving business value.
- Follow a top-down approach.


THE UNIVERSITY OF DODOMA
UDOM
Embracing Knowledge


The Nelson Mandela
African Institution of Science and Technology
Arusha
Mt Kilimanjaro
2009
Academia for Society and Industry
N M
A I S T