Course Introduction

What (and how) are we going to learn?

Yordan Darakchiev

Technical Trainer

iordan93@gmail.com

Table of Contents

Course Objectives

Working with data in a disciplined way

Course Objectives

- Learn how to apply the scientific method to
  - Ask the right questions
  - Obtain and clean up data
  - Explore and analyze data
  - Make the correct conclusions
- Write your own research
  - Learn how to create evidence-based, reproducible research
- Learn how to create a complete solution
  - Incorporate best practices in software design and science
- Communicate and compare results with the community

Prerequisites

Programming Basics

- Understand what variables and for-loops are
- Software development experience is a plus but not required

Math Concepts

- Know a little about algebra and statistics
- Have basic logic and intuition

Intermediate English

- Understand what is written on the slides

Scientific Mindset

- Be open to (and not afraid of) challenges

Course Format Details

Curriculum, schedule, trainer, lecture format, exam

Curriculum

- Course introduction
- Data acquisition
- Data tidying and cleaning
- Data visualization. Exploratory data analysis
- Case studies, part 1: Images and text
- Case studies, part 2: Spatial data and networks
- Modelling basics: regression models
- Best practices in software and science
- Final exam

Course Schedule

- Lessons
  - 8 lectures x 4 hours each on-site
- Homework
  - 4 hours+ – the more, the better
  - Quiz – 0.25-0.5 hours
    - Questions to check your understanding
  - Lab – 1-6 hours
    - Problems related to real-life data science
- Extracurricular activities: 0+ hours
- Practical exam
  - Preparation at home – 3+ hours
  - On-site defense: 10 minutes

Final Exam

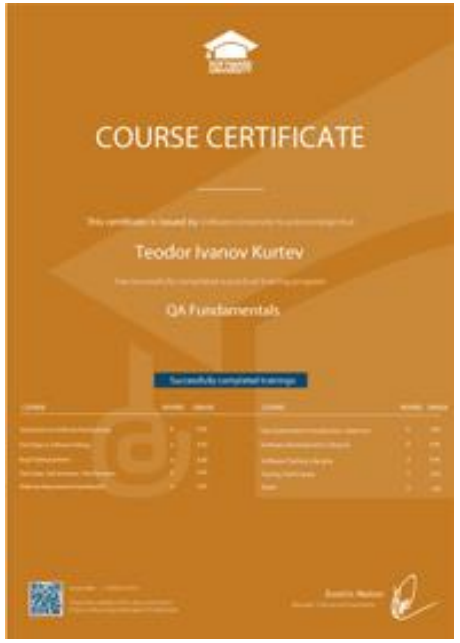
- Practical project
  - Work on your own
    - No teams allowed
  - Present your results (documentation, code, etc.) in a limited amount of time
- Work on a given assignment
  - Assignment release time: at second lecture
  - Perform research
    - Scientific papers, community forums, etc.
  - Analyze the data
  - Write code
  - Communicate the results

Grading Scheme

- Quizzes: up to 10%
  - Due 2 weeks after the lecture date
  - 3 tries per quiz
- Labs: up to 20%
  - Due 2 weeks after the lecture date
- Final exam: up to 70%
  - Develop at your own pace
  - Upload deadline: 9 February 2018, 23:59:59 GMT+2
  - On-site defense: 11 February 2018, 10:00:00 GMT+2
  - To qualify: at least 5/30 points from quizzes and labs
- Forum activity: bonus up to 10%

Grading and Course Certificate

- All students will be graded on a scale from 2.00 to 6.00
  - The same way the standard grading in Bulgaria works
- Everyone who scores ≥ 5.00 (total) on the course will get a certificate from Softuni
  - Starting point for a new career or continuing education in your current field
  - Career assistance
    - The SoftUni career center will help you find work
  - Official and recognizable
    - Employers value certificates
  - Proof of hard work :)
    - Shareable and verifiable



Who Am I?

- Programmer
  - .NET Web developer
  - Graduated Telerik Academy in Oct 2013 with distinction
- Trainer
  - Various programming courses
    - To beginners and experienced developers
  - Scientific (and popular) lectures
- Scientist / Enthusiast
  - BSc in Astrophysics (July 2016)
  - Currently pursuing a MSc in Astrophysics
- Overall nerd
  - Curious and skeptical

Learning Resources

Learn more
and share your knowledge

SoftUni Course Pages

- Official Web page of this course
  - https://softuni.bg/trainings/1816/data-science-november-2017
- Forum category
  - https://softuni.bg/forum/categories/96/data-science
  - Ask and answer questions
    - I will try to answer your questions as well
  - Post what you've learned
    - Links to resources, code snippets, ideas, tips and tricks
  - Share your problems (homework or not) and help solve them
  - Create and maintain a community
    - A **critical part** of doing science

Online Resources

- Books
  - "How not to be wrong" – Jordan Ellenberg
  - "Learning Data Mining with Python" – Robert Layton
  - … and anything else you can find
- Websites
  - Khan Academy
  - Communities: Kaggle, Quora, Stack Exchange
  - Online courses: Coursera, edX, MIT OCW, Stanford, etc.
- YouTube
  - FunFunFunction, Daniel Shiffman, Siraj Raval, AsapSCIENCE, Veritasium, Vsauce, TedEd, CrashCourse, Mind Your Decisions, Infinite Series, Numberphile, Computerphile, Vi Hart, 3Blue1Brown, blackpenredpen, Mathologer, and many more

Questions?