'Real-life problem solving with data'
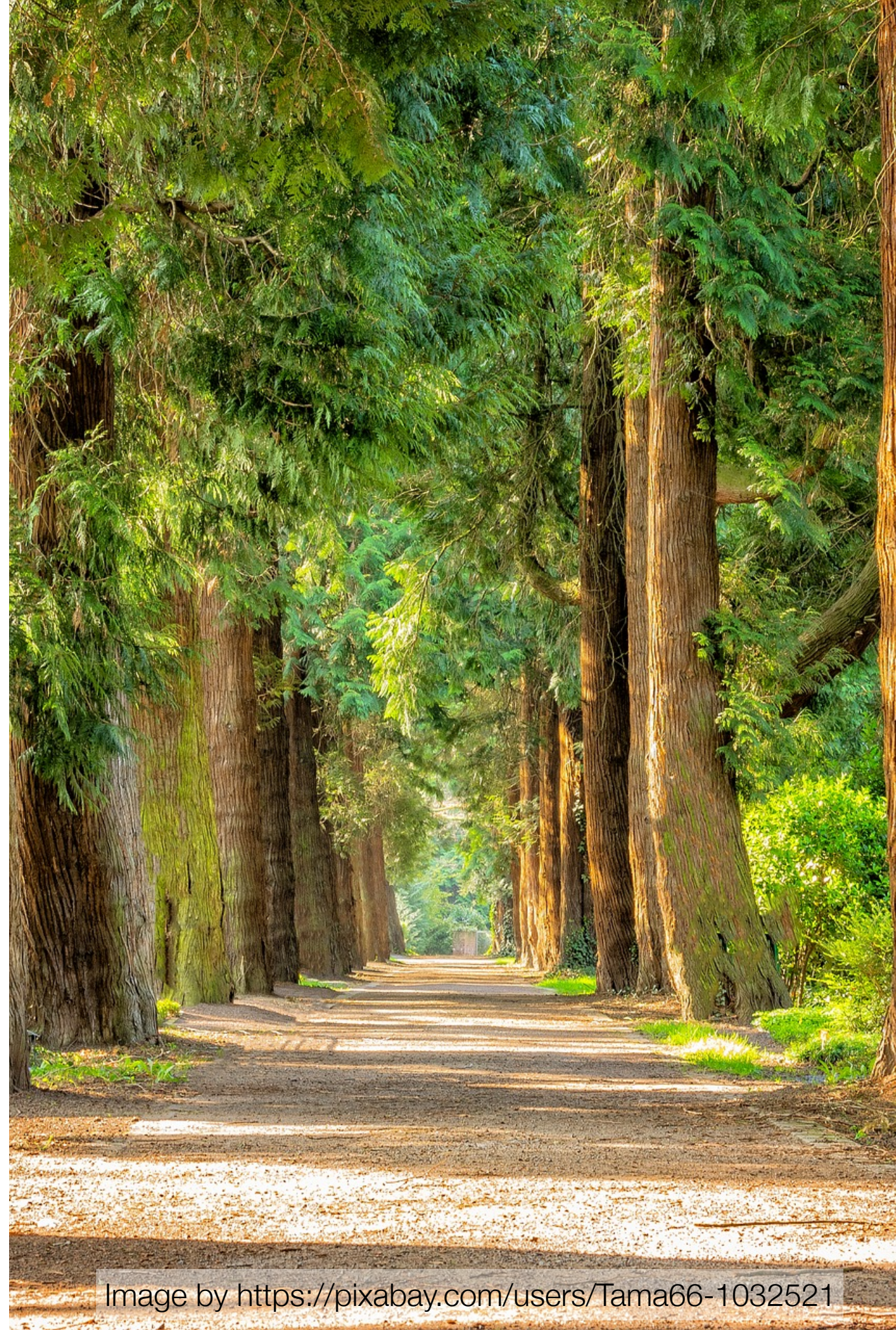
Dr Elena Hensinger
Dr Joanne Kitson

Toumetis Ltd.

WTH Workshop Wednesdays
11th Sept 2019

Image by https://pixabay.com/users/Tama66-1032521

Tools

Do you have a Google account?

YES —> sit back and relax

NO —> please register for one now:
https://accounts.google.com/signup/v2

WHY? We will be working with the free online coding environment Colaboratory, which will save your notebooks (i.e. code) in your Google Drive.

Objectives of today's workshop

Introduction to data analysis

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data
- read in data

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data

- read in data
- explore

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data

- read in data
- explore
- clean, preprocess

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data

- read in data
- explore
- clean, preprocess
- model in order to answer a problem statement

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data

- read in data
- explore
- clean, preprocess
- model in order to answer a problem statement
- assess answer, evaluate model

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data
- read in data
- explore
- clean, preprocess
- model in order to answer a problem statement
- assess answer, evaluate model

Learn some Python commands for data analysis

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data
- read in data
- explore
- clean, preprocess
- model in order to answer a problem statement
- assess answer, evaluate model
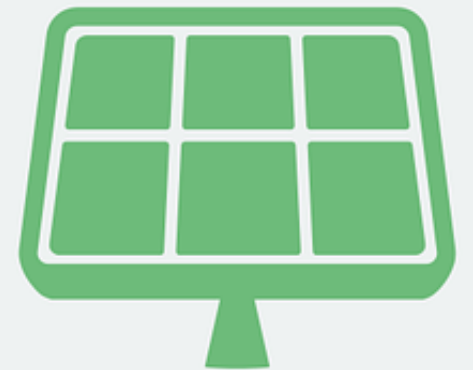
Learn some Python commands for data analysis

Eat some pizza ~ 7.45pm

All set to go?

Problem statement

Support Bristol's citizens to save money on energy usage

Image by https://pixabay.com/users/rawpixel-4283981

Problem statement

Support Bristol's citizens to save money on energy usage

Participants monitored energy usage at different times of the day —> usage per typical day

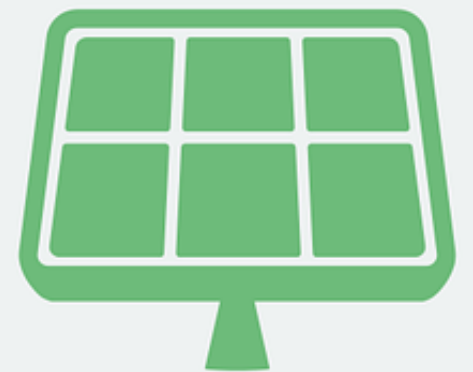# Problem statement

Support Bristol's citizens to save money on energy usage

Participants monitored energy usage at different times of the day —> usage per typical day

Task: analyse the data, detect some patterns

Image by https://pixabay.com/users/rawpixel-4283981

Problem statement

Support Bristol's citizens to save money on energy usage

Participants monitored energy usage at different times of the day —> usage per typical day

Task: analyse the data, detect some patterns

How can we support participants to reduce their usage and / or associated cost?
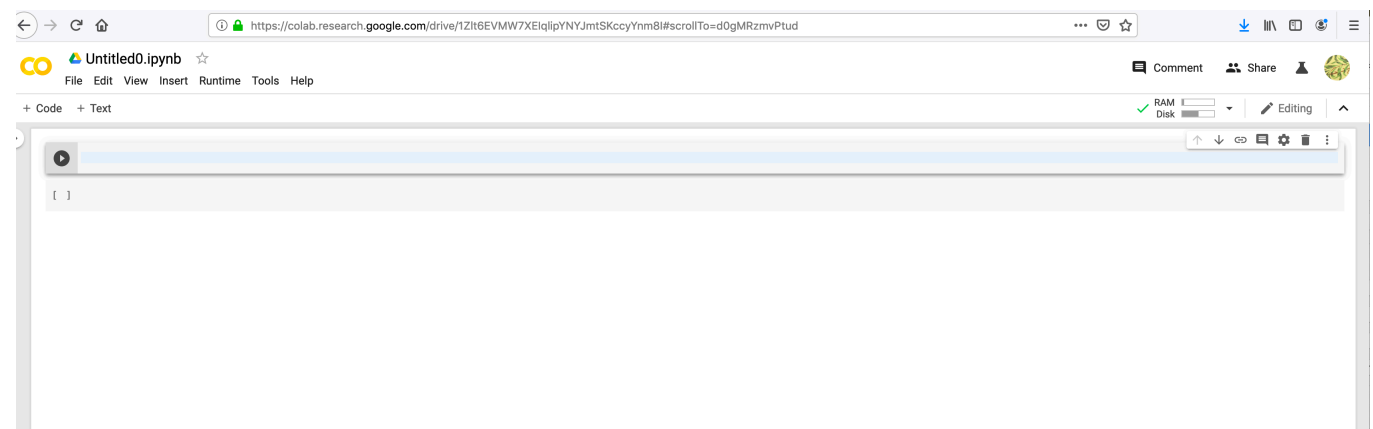
Image by https://pixabay.com/users/rawpixel-4283981

Tools

Google Colaboratory - interactive environment to run Python code

https://colab.research.google.com/notebooks/welcome.ipynb

- Follow URL to homepage
- Right-hand corner: Sign in
- You should get something like this

Tools, cont.

Colab is your own working space

- Change name for notebook into something like 'Workshop_WTH_IntroData'

- Save

- Cells = code is executed here

What to analyse?

Browser: https://github.com/Toumetis/
2019_09_11_WorkshopWednesdays

Download the data file in 'datasets/workshop_data1.csv'
('Clone or download' —> download zip file —> unzip)

Remember where you stored the file

All set to go?

Let's work together

Objectives of today's workshop

Introduction to data analysis

Experience typical stages of working with data
- read in data
- explore
- clean, preprocess
- model in order to answer a problem statement
- assess answer, evaluate model

Learn some Python commands for data analysis

Eat some pizza ~ 7.45pm

Hands-on demo

Download zip of repository at https://github.com/Toumetis/2019_09_11_WorkshopWednesdays —> unzip

Go to Colab and log in with your Google account https://colab.research.google.com/notebooks/welcome.ipynb

'File' —> 'Upload notebook' —> choose the file 'Data_analysis_1.ipynb' from the directory 'notebooks/workshop/'

Run cell by cell

Data Science

You did some Data Science!

Data Science

You did some Data Science!

Data science is the field of study that combines domain expertise, programming skills, and knowledge of math and statistics to **extract meaningful insights from data**. (Source: https://www.datarobot.com/wiki/data-science/)
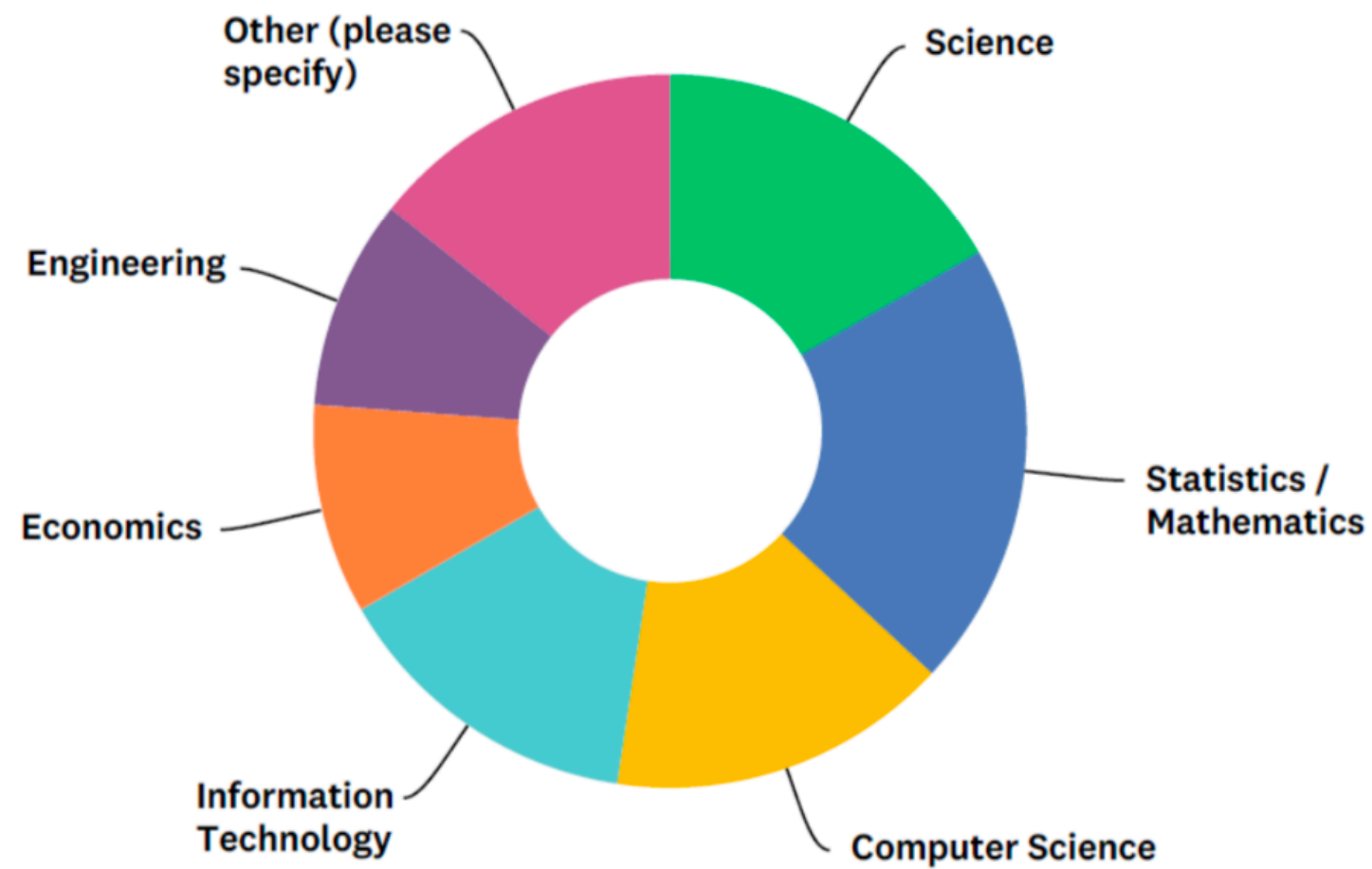
Data Science

You did some Data Science!

Data science is the field of study that combines domain expertise, programming skills, and knowledge of math and statistics to **extract meaningful insights from data**. (Source: https://www.datarobot.com/wiki/data-science/)
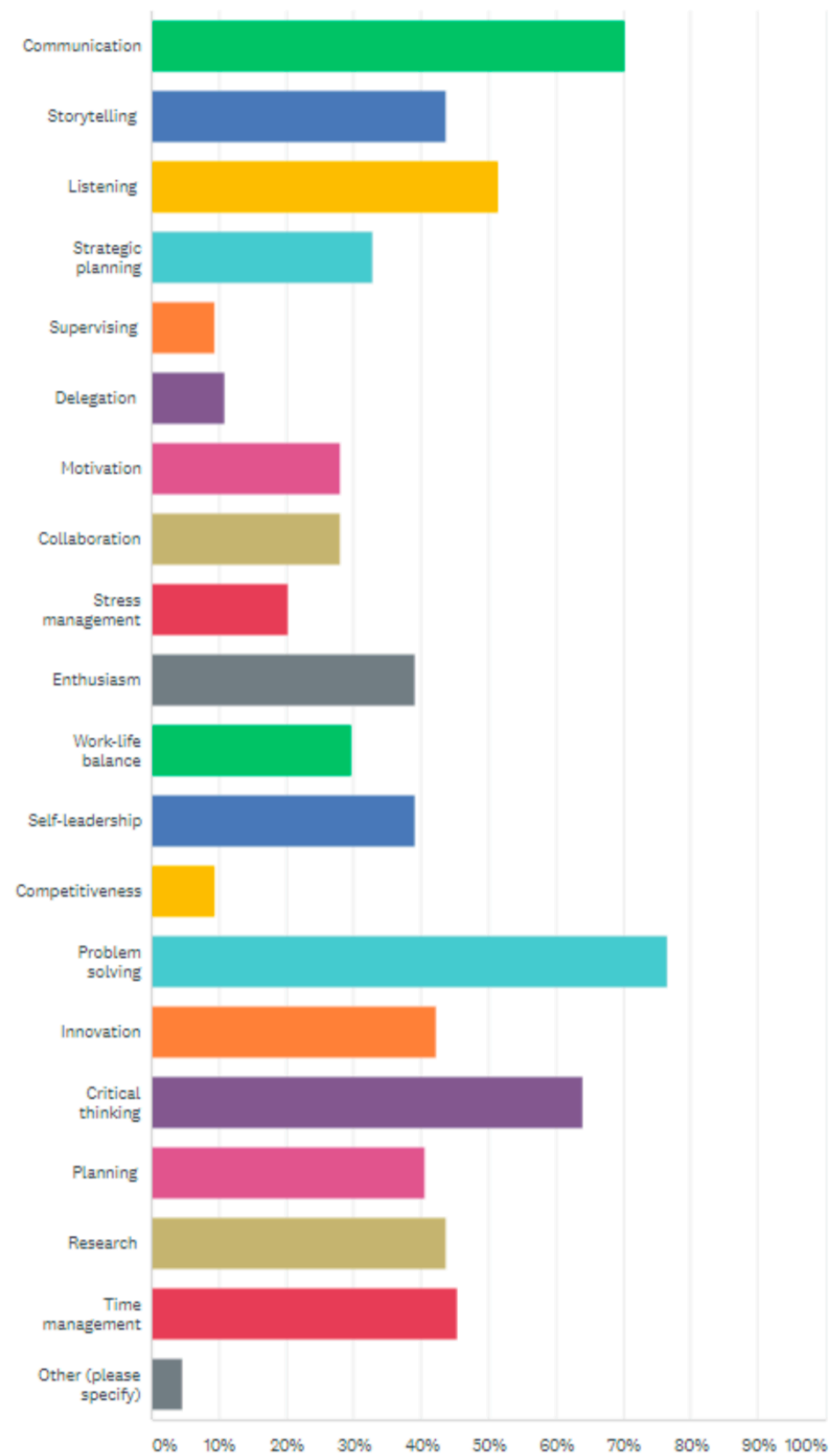
A bit more about the people who do Data Science…

Background of data scientists

Source: 'The Timeline of Data Science - Reviewing the results' (Wade Macdonald recruiting)
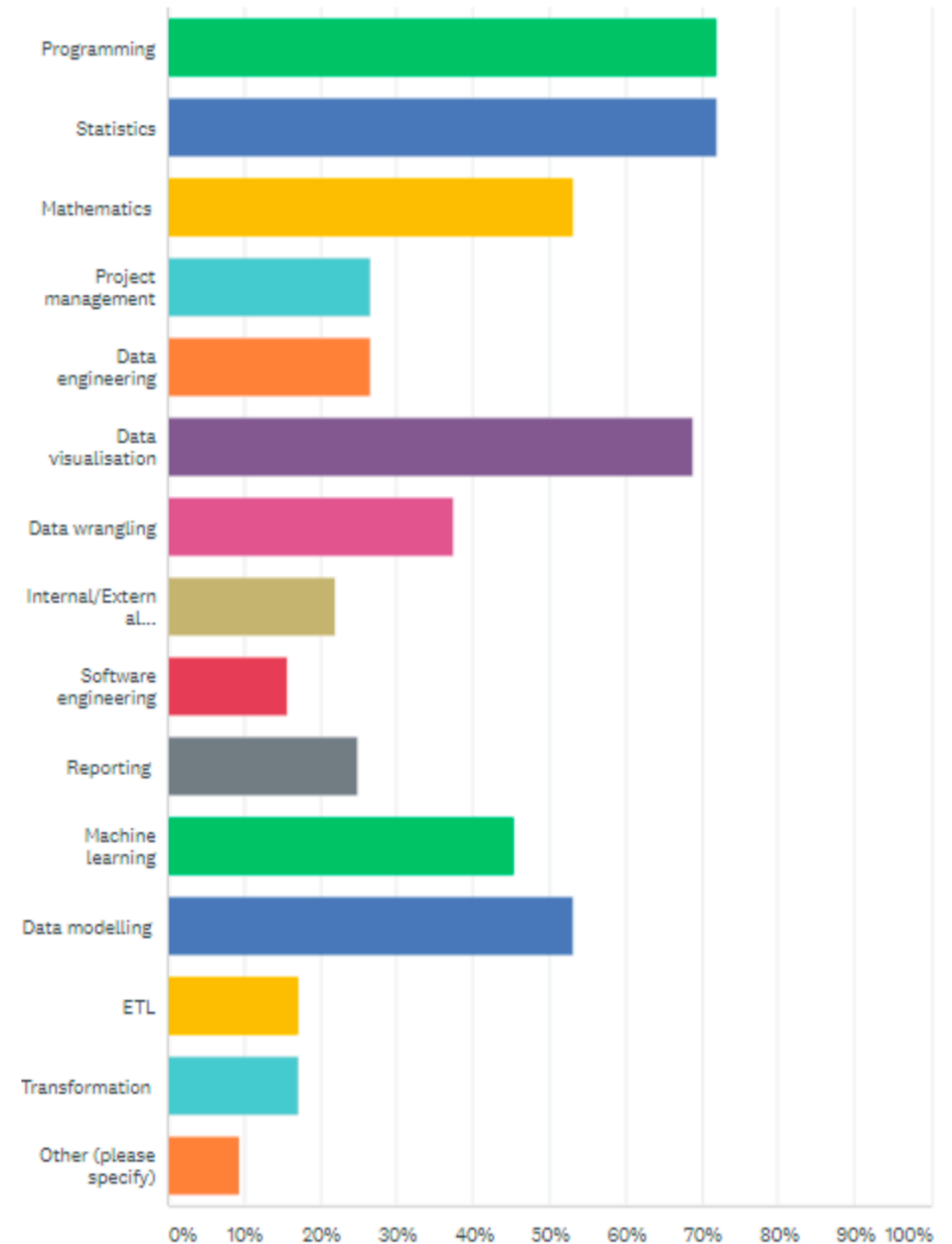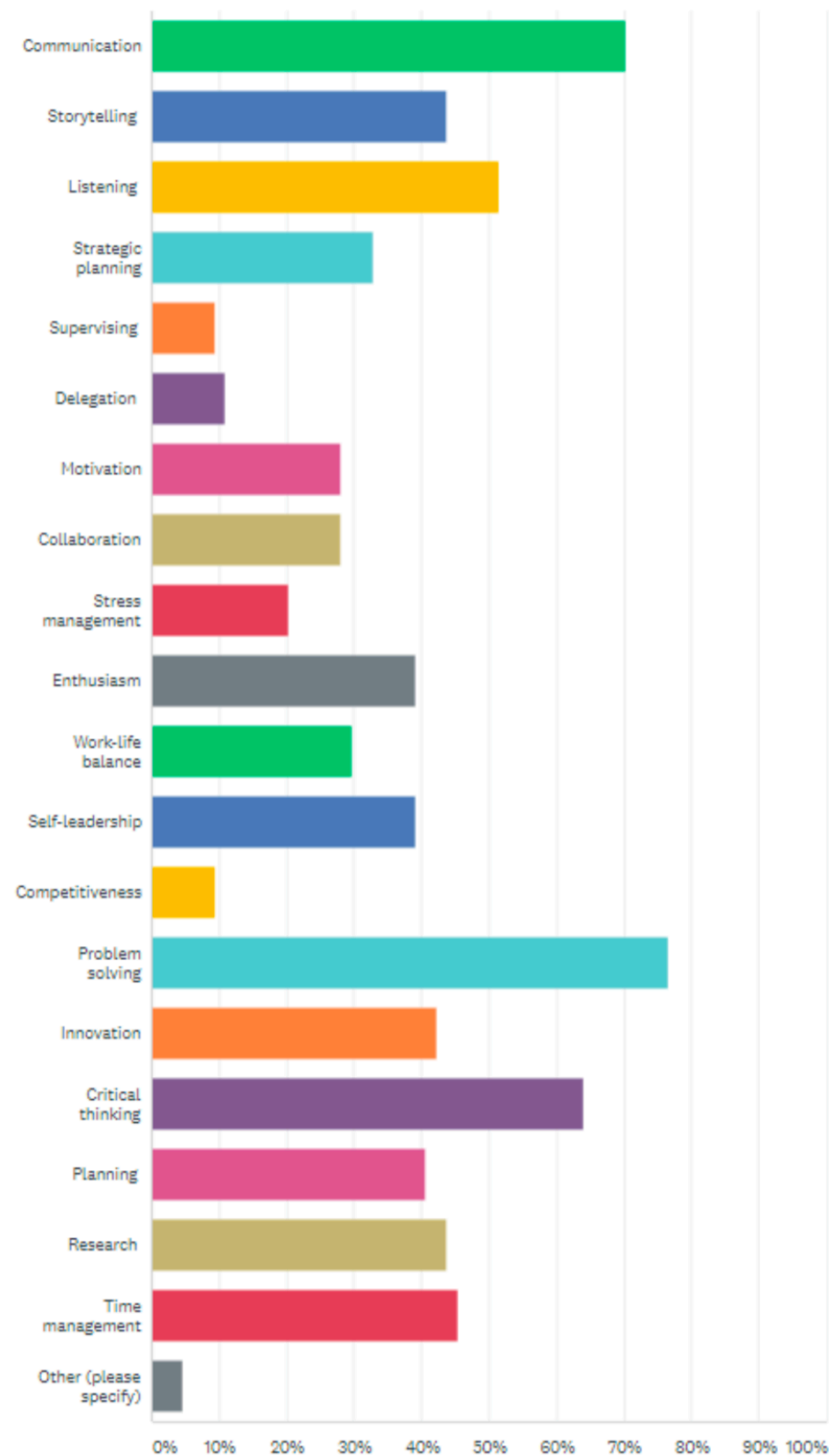
Skills required

Skills required

Source: 'The Timeline of Data Science - Reviewing the results' (Wade Macdonald recruiting)

Skills required

Source: 'The Timeline of Data Science - Reviewing the results' (Wade Macdonald recruiting)

Want to learn more?

- Women Who Code Data Science Slack channel: questions; info about events; webinars and recordings; resources; academic papers reading group; career topics; job opportunities

- Kaggle: datasets, tutorials, competitions

- Medium 'Towards Data Science': inspiration, visualisations

- To learn Python: https://pythonprogramming.net/ and the youtube channel by the same author https://www.youtube.com/user/sentdex/featured

- https://www.fast.ai/ Various free courses, amongst other on Machine Learning, Deep Learning and Natural Language Processing

- explore local Meetup groups for Machine Leaning and Data Science talks, workshops and learning groups

Stay in touch

Dr Elena Hensinger
elena@toumetis.com and on LinkedIn

Toumetis Ltd. - hiring
https://toumetis.bamboohr.com/jobs/