



UMCS
Wydział Ekonomiczny

ANALITYKA GOSPODARCZA
studia II stopnia

Zaawansowane programowanie w analizie danych – kurs 15h

Mieczysław Pawłowski, PhD, UMCS, Wydział Ekonomiczny
mieczyslaw.pawlowski@umcs.pl. <https://mietwood.com>



Analityka biznesowa

Analityka biznesowa odnosi się do szeregu działań wykonywanych w celu wykorzystywania istniejących danych do podejmowania decyzji opartych na faktach.

Analityka biznesowa wykorzystuje

- różne umiejętności analityka,
- różne metody zbierania i przekształcania danych,
- sprzęt i oprogramowanie do przetwarzania danych w raporty i inne rezultaty analizy

Rezultaty analizy są wykorzystywane dla podejmowania decyzji przez

- managerów,
- klientów i wszelkie osoby zainteresowane

Analityka biznesowa szeroko wykorzystuje analizę statystyczną i zaawansowane programowanie analityczne w tym machine learning i nawet metody sztucznej inteligencji.



Rodzaje analityki biznesowej

Analitykę biznesową dzielimy w zależności od tego, w jakim celu poszukuje się informacji, jaki mają one charakter oraz w jaki sposób są one uzyskiwane. W literaturze fachowej można znaleźć co najmniej kilka sposobów nadawania kategorii analityce biznesowej. Jednym z nich jest podział na:

analitykę predykcyjną – zajmuje się modelowaniem przyszłości i prognozowaniem,

analitykę preskryptywną – określa możliwe scenariusze, opcje, jak i konsekwencje potencjalnych decyzji,

analitykę diagnostyczną – odpowiada na pytanie o przyczyny (zjawisk, zdarzeń) w oparciu o dane historyczne,

analitykę deskryptywną – przetwarza dane historyczne w celu opisanie zjawisk, rezultatów, faktów z przeszłości,

analitykę kognitywną, która wykorzystuje zaawansowane technologie sztucznej inteligencji i uczenia maszynowego

big data, to jest przetwarzania dużych ilości danych, często o charakterze nietypowym, w celu wspierania decyzji managerów lub automatycznego podejmowania decyzji.

<https://businessintelligence.pl/czym-jest-analika-biznesowa/>



WORKFLOW ANALYSIS

01



Gather the data

02



Analyze the data

03



Brainstorm ideas

YES

NO



04

Implement the options

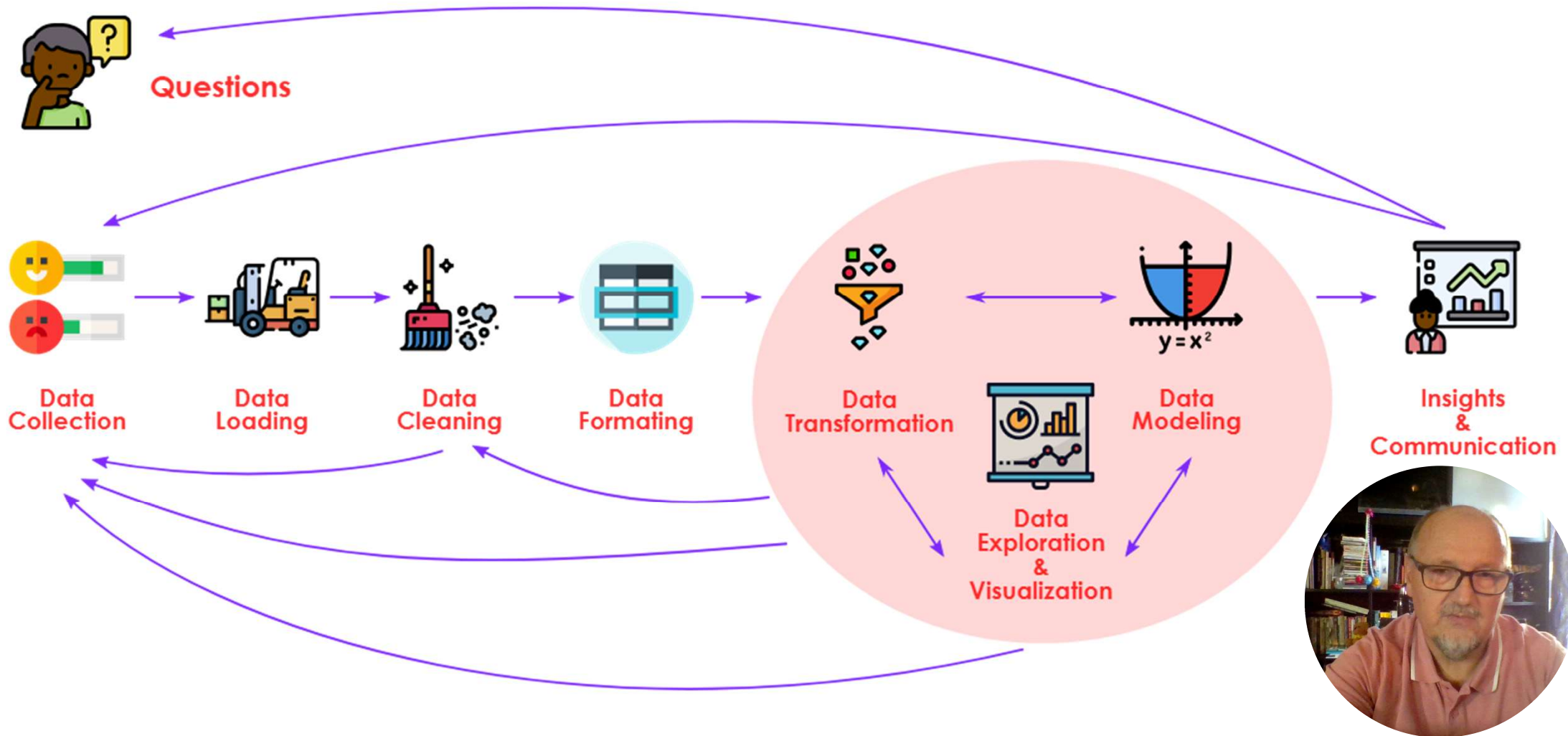


05

Monitor progress & adjust



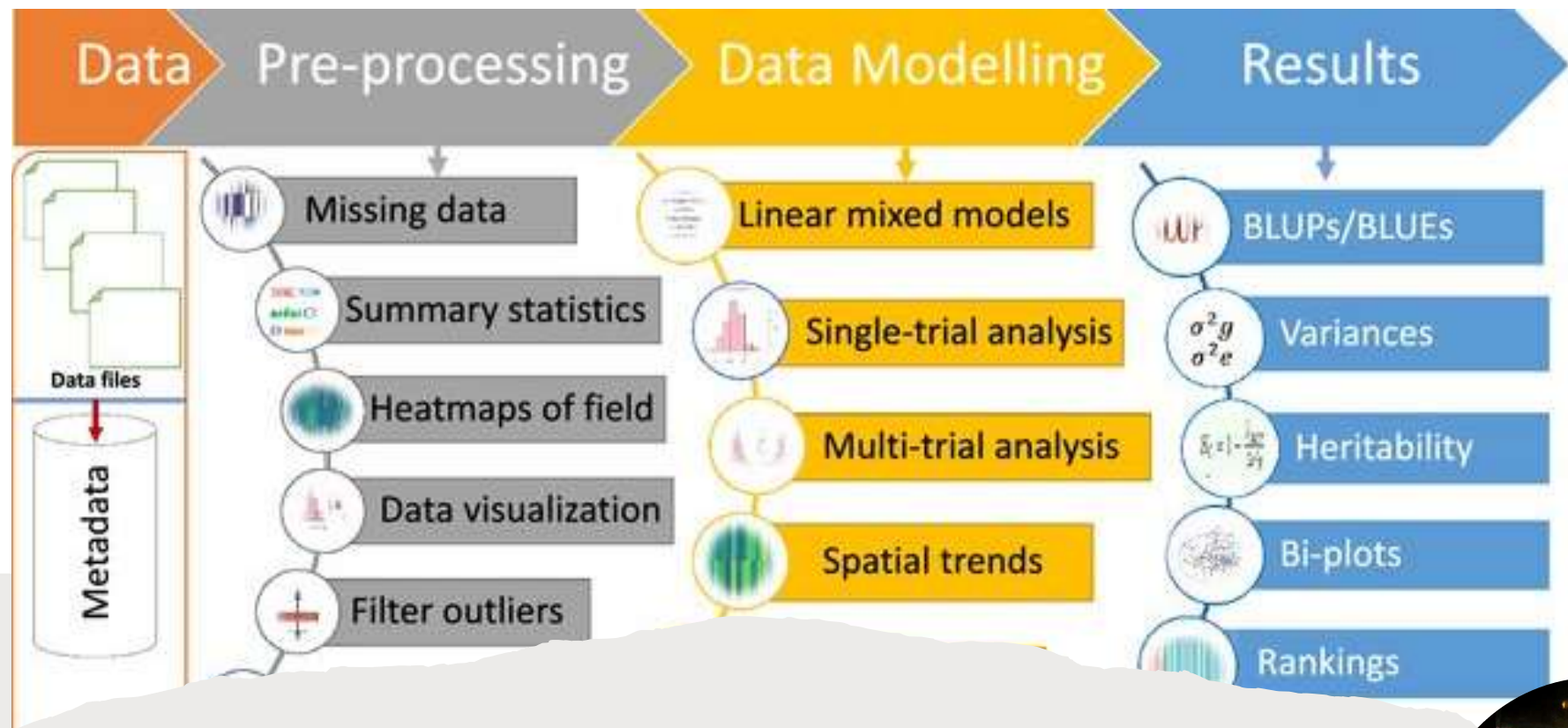
Data Analysis Workflow



Co zajmuje
analitykowi
najwięcej czasu?

Data cleaning
NINJA



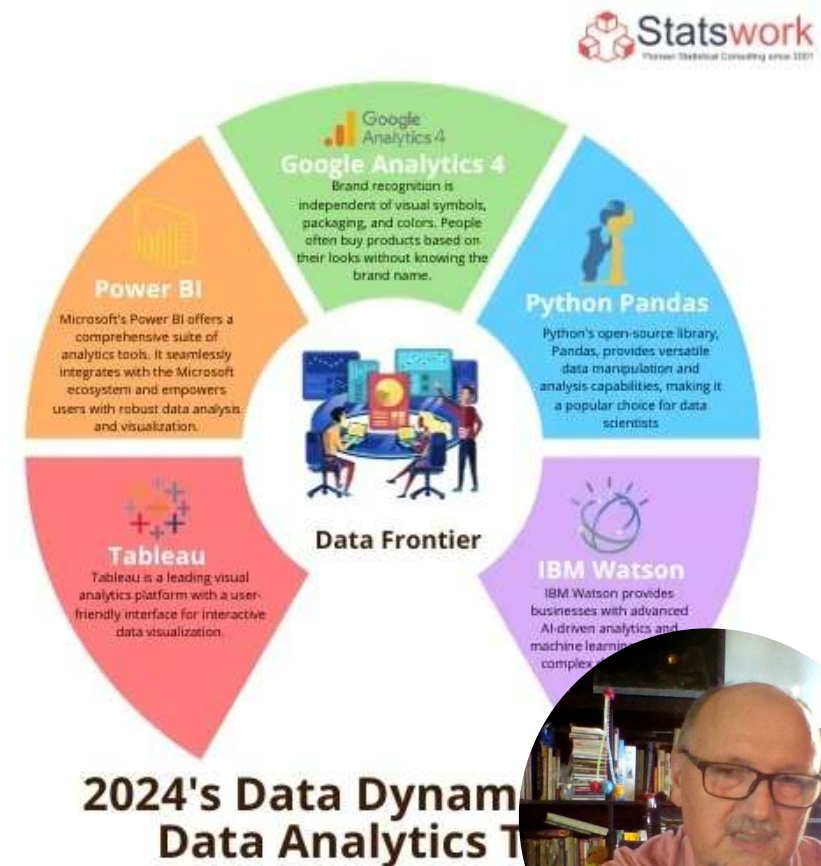


Proces analityczny



Środowisko narzędziowe analityka

1. Pozyskiwanie danych, przechowywanie danych
 - Pliki, txt, csv, pdf, xlsx, inne ... export z systemów raportowych, monitoringów itp..
 - Bazy danych, bazy programów ERP, CRM, MS SQL Server, mySql, Postgres, DB2, SAP HANA, MS Access
 - Dostęp do danych przez API
2. Analiza danych
 1. Excel,
 2. Sql,
 3. Python, VBA
3. Prezentacja informacji: dane → informacje (Inside)
 1. Excel
 2. Power BI, Python + Flask/App
 3. Tableau, QLink

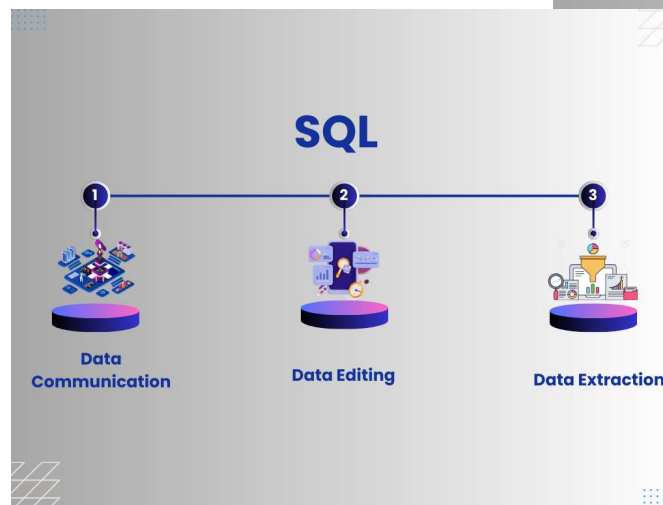
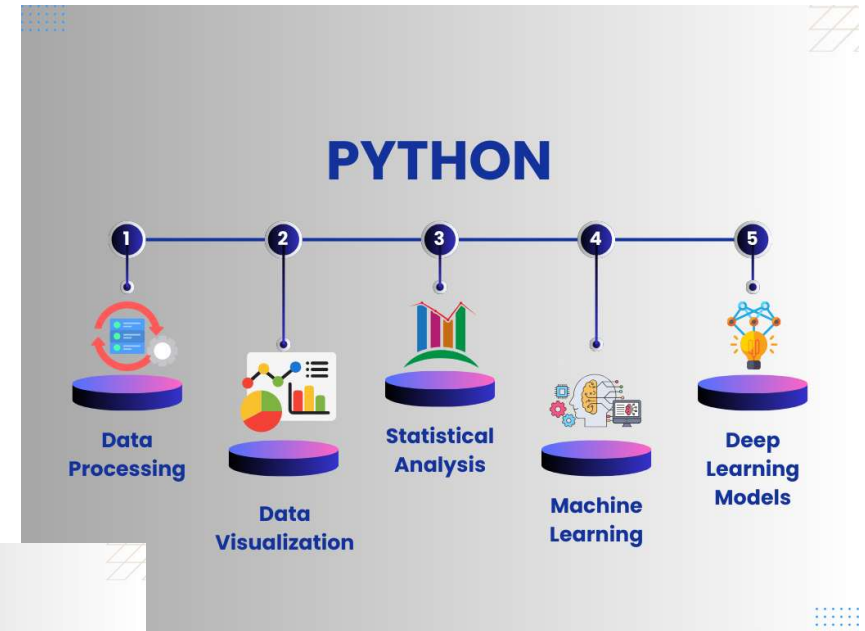


Magic Quadrant for Analytics and Business Intelligence Platforms.



Best Programming Languages for Data Science and Analytics

- [1. Python](#)
- [2. SQL](#)
- [3. R](#)
- [4. VBA \(Visual Basic for Applications\)](#)
- [5. Java](#)
- [6. Julia](#)



Plan wykładu

1. Środowisko pracy analityka, dane i źródła danych
2. Pozyskiwanie danych (ETL- Extract, Transform, Load) ich przechowywanie
3. Przechowywanie danych - bazy danych, SQL, MSSQL
4. Python jako język programowania dla analityków – środowisko programistyczne analityka, IDE, organizacja projektów
5. Prezentacja informacji, udostępnianie informacji – Power BI – raporty i dashboardy
6. Programowanie funkcyjne dla analityków: Python fundamentals, elementy języka, funkcje, pętle i struktury kontrolne
7. Najważniejsze biblioteki: So, Pandas, Numpy, SQLAlchemy, Datetime, Flask, itd
8. Python & Power BI / Excel jako środowisko BI w firmie, przetwarzanie i udostępnianie danych
9. Tworzenie własnych bibliotek i ich wykorzystanie w kolejnych programach
10. Automatyzacja wykonywania skryptów i raportów
11. Od skryptu do aplikacji: aplikacje webowe, GUI
12. Projekty analityczne: budżetowanie, wykrywanie anomalii, przewidywanie sprzedaży, system rekomendacyjny



Literatura

Advanced Data Analytics Using Python

With Machine Learning, Deep Learning and NLP Examples
—
Sayan Mukhopadhyay

Apress®

Springer Proceedings in Business and Economics

Arnab Kumar Laha *Editor*

Advances in Analytics and Applications

Data Analysis and Visualization Using Python

Analyze Data to Create Visualizations for BI Systems
—
Dr. Ossama Embarak

Practical Data Analysis Second Edition

A practical guide to obtaining, transforming, exploring, and analyzing data using Python, MongoDB, and Apache Spark

Hands-On Data Science for Marketing

Improve your marketing strategies with machine learning using Python and R

O'REILLY

Python for Data Analysis

Data Wrangling with pandas, NumPy & Jupyter

powered by



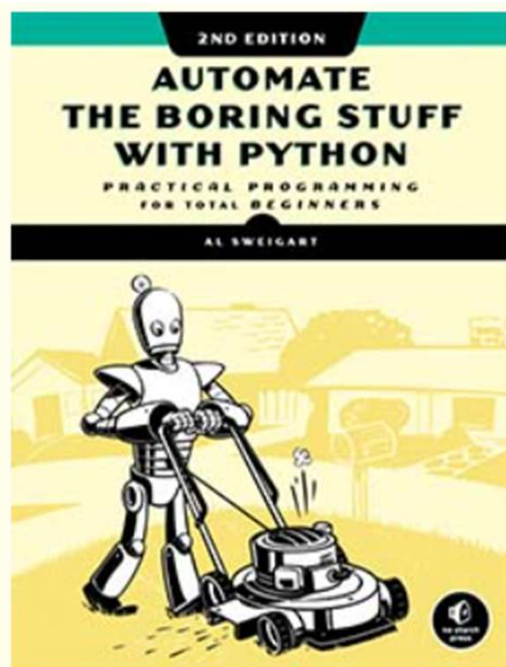
Wes McKinney

Learn Data Analysis with Python

Lessons in Coding
—
A.J. Henley
Dave Wolf



Python online



<https://automatetheboringstuff.com/>

Google for Education > Python

Filter

Overview

- Python Set Up
- Python Intro
- Strings
- Lists
- Sorting
- Dicts and Files
- Regular Expressions
- Utilities

Lecture Videos

- 1.1 Introduction, strings ☒
- 1.2 Lists and sorting ☒
- 1.3 Dicts and files ☒
- 2.1 Regular expr ☒
- 2.2 Utilities ☒
- 2.3 Utilities urllib ☒
- 2.4 Conclusions ☒

Python Exercises

- Basic Exercises
- Baby Names Exercise
- Copy Special Exercise
- Log Puzzle Exercise

Home > Products > Google for Education > Python

Was this helpful?

Google's Python Class

Welcome to Google's Python Class -- this is a free class for people with a little bit of programming experience who want to learn Python. The class includes written materials, lecture videos, and lots of code exercises to practice Python coding. These materials are used within Google to introduce Python to people who have just a little programming experience. The first exercises work on basic Python concepts like strings and lists, building up to the later exercises which are full programs dealing with text files, processes, and http connections. The class is geared for people who have a little bit of programming experience in some language, enough to know what a "variable" or "if statement" is. Beyond that, you do not need to be an expert programmer to use this material.

To get started, the Python sections are linked at the left -- [Python Set Up](#) to get Python installed on your machine, [Python Introduction](#) for an introduction to the language, and then [Python Strings](#) starts the coding material, leading to the first exercise. The end of each written section includes a link to the code exercise for that section's material. The lecture videos parallel the written materials, introducing Python, then strings, then first exercises, and so on. At Google, all this material makes up an intensive 2-day class, so the videos are organized as the day-1 and day-2 sections.

This material was created by [Nick Parlante](#) working in the engEDU group at Google. Special thanks for the help from my Google colleagues John Cox, Steve Glassman, Piotr Kaminski, and Antoine Picard. And finally thanks to Google and my director Maggie Johnson for the enlightened generosity to put these materials out on the internet for free under the [Creative Commons Attribution 2.5](#) license -- share and enjoy!

★ Tip: Check out the [Python Google Code University Forum](#) to ask and answer questions.

Was this helpful?

<https://developers.google.com/edu/python/>



Are we doing RIGHT THINGS?

Czy robimy właściwe rzeczy?





U.S. BUREAU OF LABOR STATISTICS

Follow Us    Release Calendar | Blog

 Search BLS.gov

HOME ▾ SUBJECTS ▾ DATA TOOLS ▾ PUBLICATIONS ▾ ECONOMIC RELEASES ▾ CLASSROOM ▾ BETA ▾

Bureau of Labor Statistics > Employment Projections > Data

Employment Projections

Search Employment Pro 

EP Home

EP Publications ▾

EP Data ▾

EP Methods ▾

About EP ▾

Contact EP

Fastest growing occupations

Other available formats: [\(XLSX\)](#)

Table 1.3 Fastest growing occupations, 2023 and projected 2033 (Numbers in thousands)

2023 National Employment Matrix title	2023 National Employment Matrix code	Employment, 2023	Employment, 2033	Employment change, numeric, 2023–33	Employment change, percent, 2023–33	Median annual wage, dollars, 2023 ^[1]
Total, all occupations	00-0000	167,849.8	174,589.0	6,739.2	4.0	48,060
Wind turbine service technicians	49-9081	11.4	18.2	6.8	60.1	61,770
Solar photovoltaic installers	47-2231	25.0	37.0	12.0	48.0	48,800
Nurse practitioners	29-1171	292.5	427.9	135.5	46.3	126,260
Data scientists	15-2051	202.9	276.0	73.1	36.0	108,020
Information security analysts	15-1212	180.7	239.8	59.1	32.7	120,360
Medical and health services managers	11-9111	562.7	723.3	160.6	28.5	110,680
Physician assistants	29-1071	153.4	197.1	43.7	28.5	130,020
Computer and information research scientists	15-1221	36.6	46.0	9.4	25.6	145,080
Physical therapist assistants	31-2021	108.5	136.0	27.5	25.4	64,080
Operations research analysts	15-2031	123.3	151.6	28.3	23.0	83,640
Occupational therapy assistants	31-2011	47.5	58.1	10.6	22.3	67,010
Actuaries	15-2011	30.2	36.8	6.6	21.8	120,000
Financial examiners	13-2061	65.5	79.2	13.8	21.0	84,300

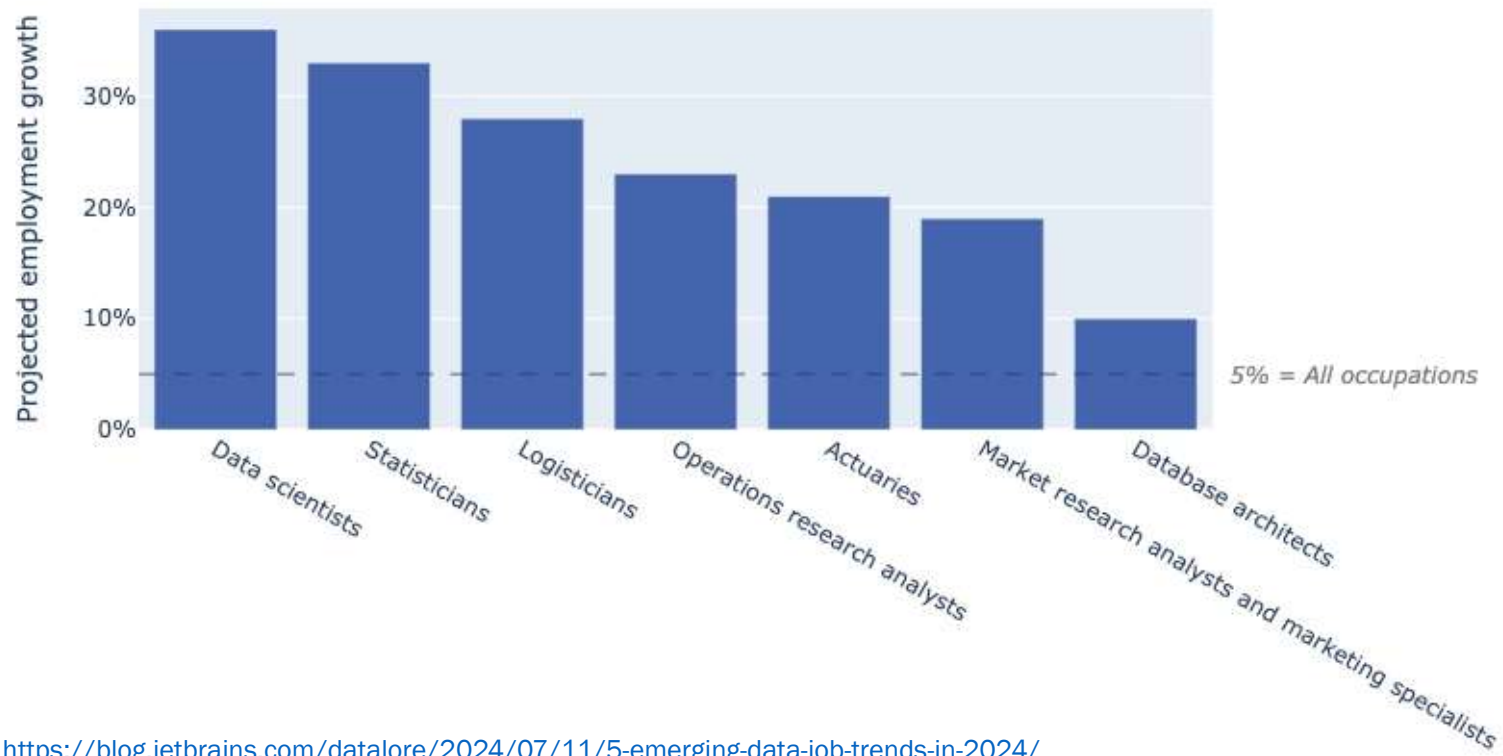
<https://www.bls.gov/p/tables/fastest-growing-occupations>



Are we doing RIGHT THINGS?

Projected Employment Growth per Occupation 2021-31

Data source: U.S. Bureau of Labor Statistics



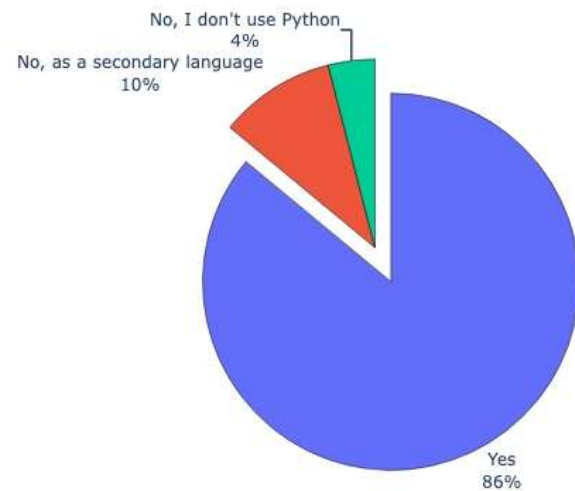
<https://blog.jetbrains.com/datalore/2024/07/11/5-emerging-data-job-trends-in-2024/>



Technical skills

Is Python the Main Language You Use for Data Science Projects?

Data source: JetBrains Python Developers Survey (2023)



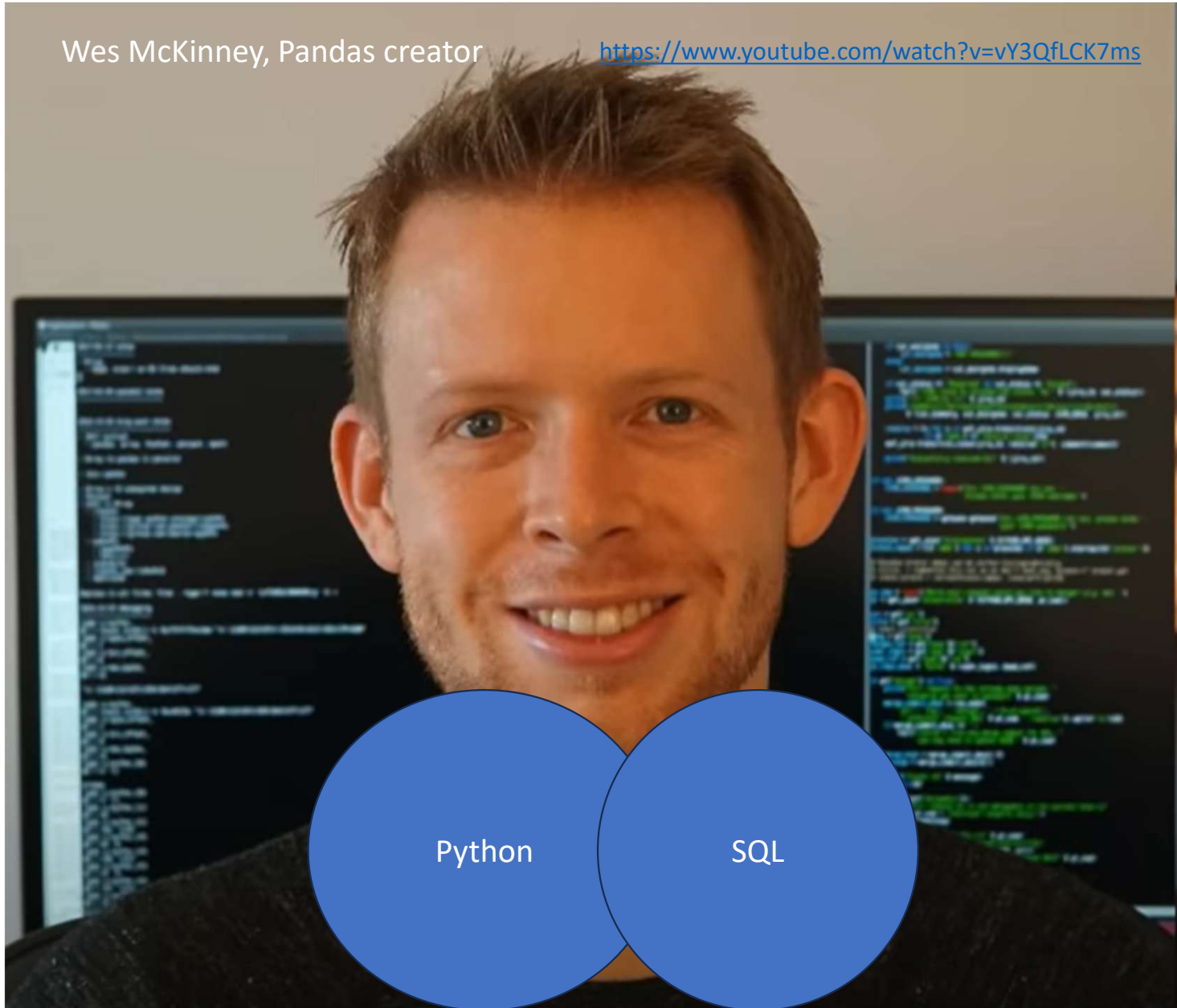
Percentage of Data Skills Mentioned in Job Posts

Data source: <https://datanerd.tech/>



Wes McKinney, Pandas creator

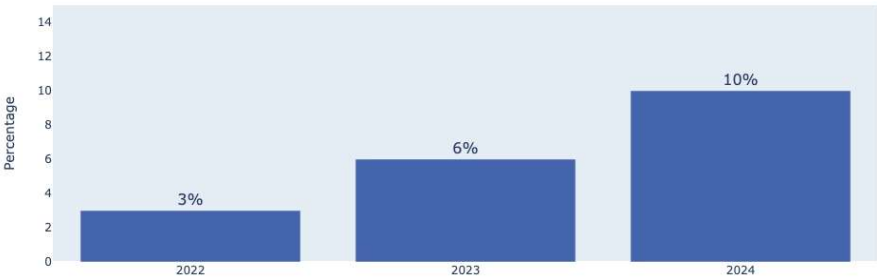
<https://www.youtube.com/watch?v=vY3QfLCK7ms>



Freelancing trends

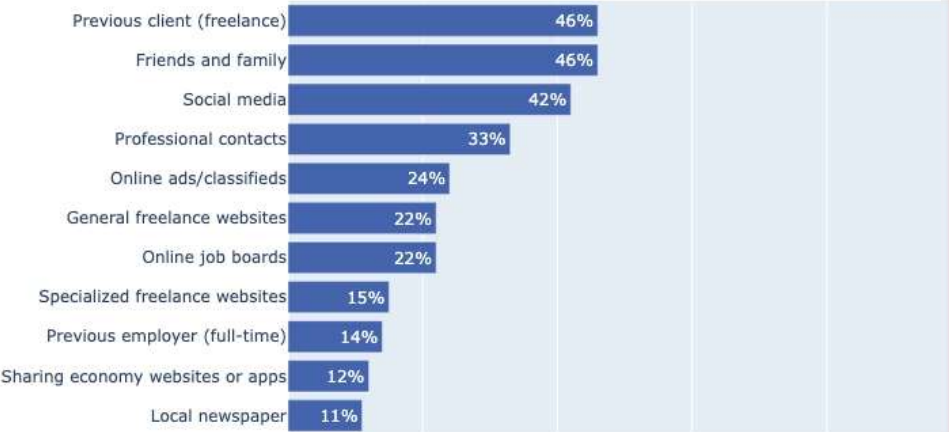
Percentage of Data Jobs Looking for Contractors per Year

Data source: <https://datanerd.tech/>



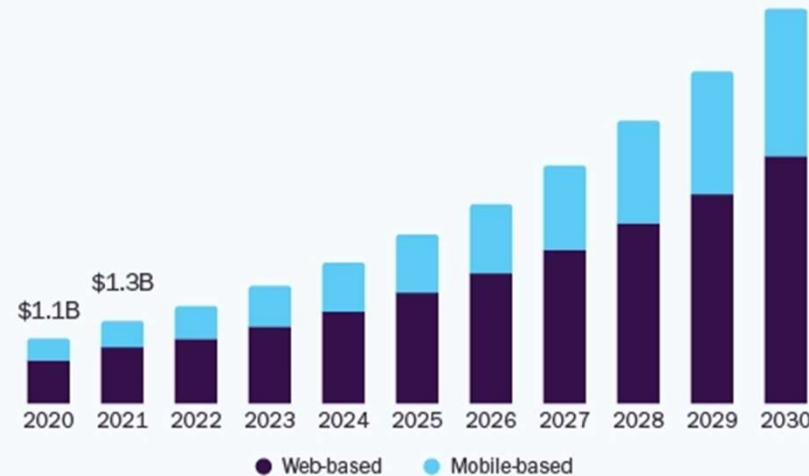
Main Information Sources Used by Freelancers to Find Work in the US (2020)

Data source: Statista



U.S. Low-code Development Platform Market

Size, by Application Type, 2020 - 2030 (USD Billion)



GRAND VIEW RESEARCH

19.0%



🔧 Top Skills for Data Nerds 🤖

Job Title:

Country:

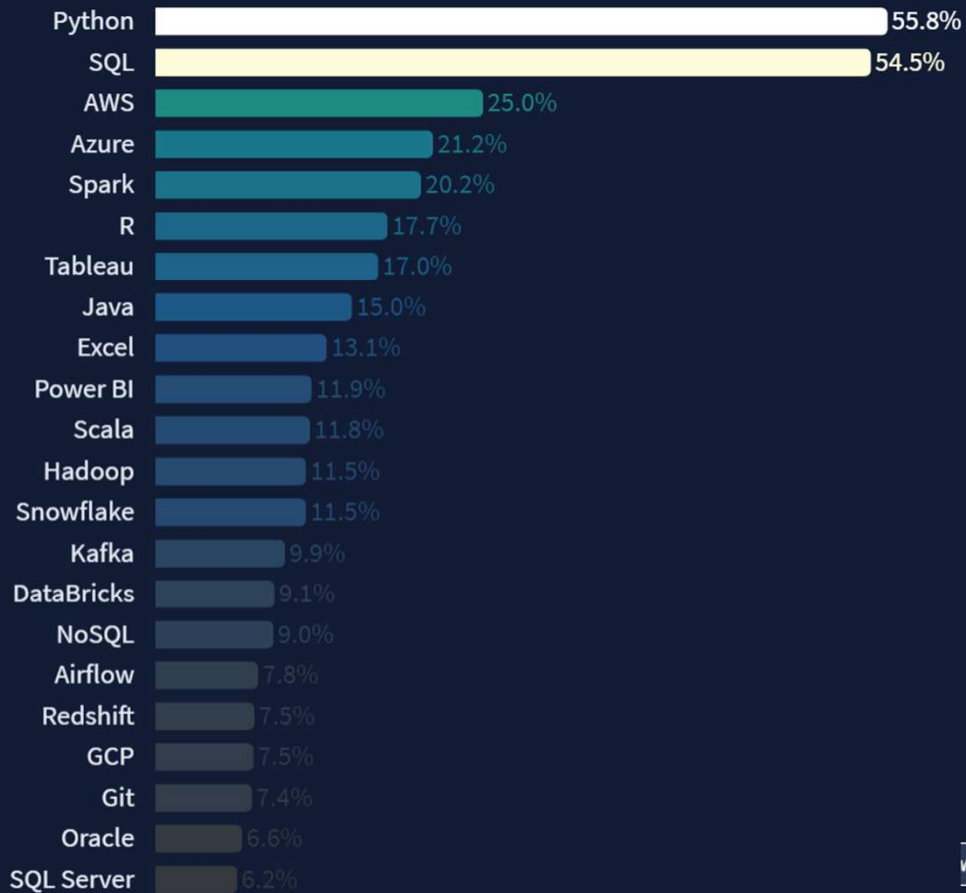
Select All

Select All

2,790,514 jobs analyzed

Skills:

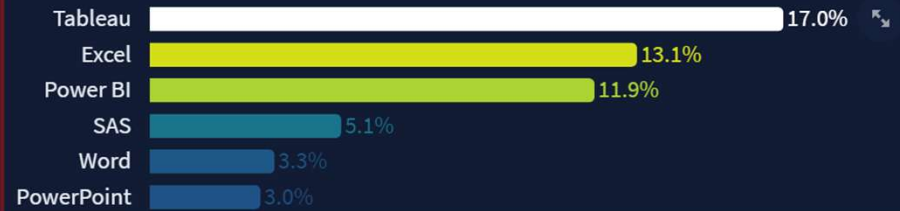
☒ All ☐ Languages ☐ Tools ☐ Databases ☐ Cloud ☐ Libraries ☐ Frameworks



powered by

Skills:

☐ All ☐ Languages ☒ Tools ☐ Databases ☐ Cloud ☐ Libraries ☐ Frameworks



☐ All ☐ Languages ☐ Tools ☒ Databases ☐ Cloud ☐ Libraries ☐ Frameworks



🕒 Skills Trend over Time

Job Title:

Country:

Select All

Select All

Skills:

Choose an option

2,790,514 jobs analyzed



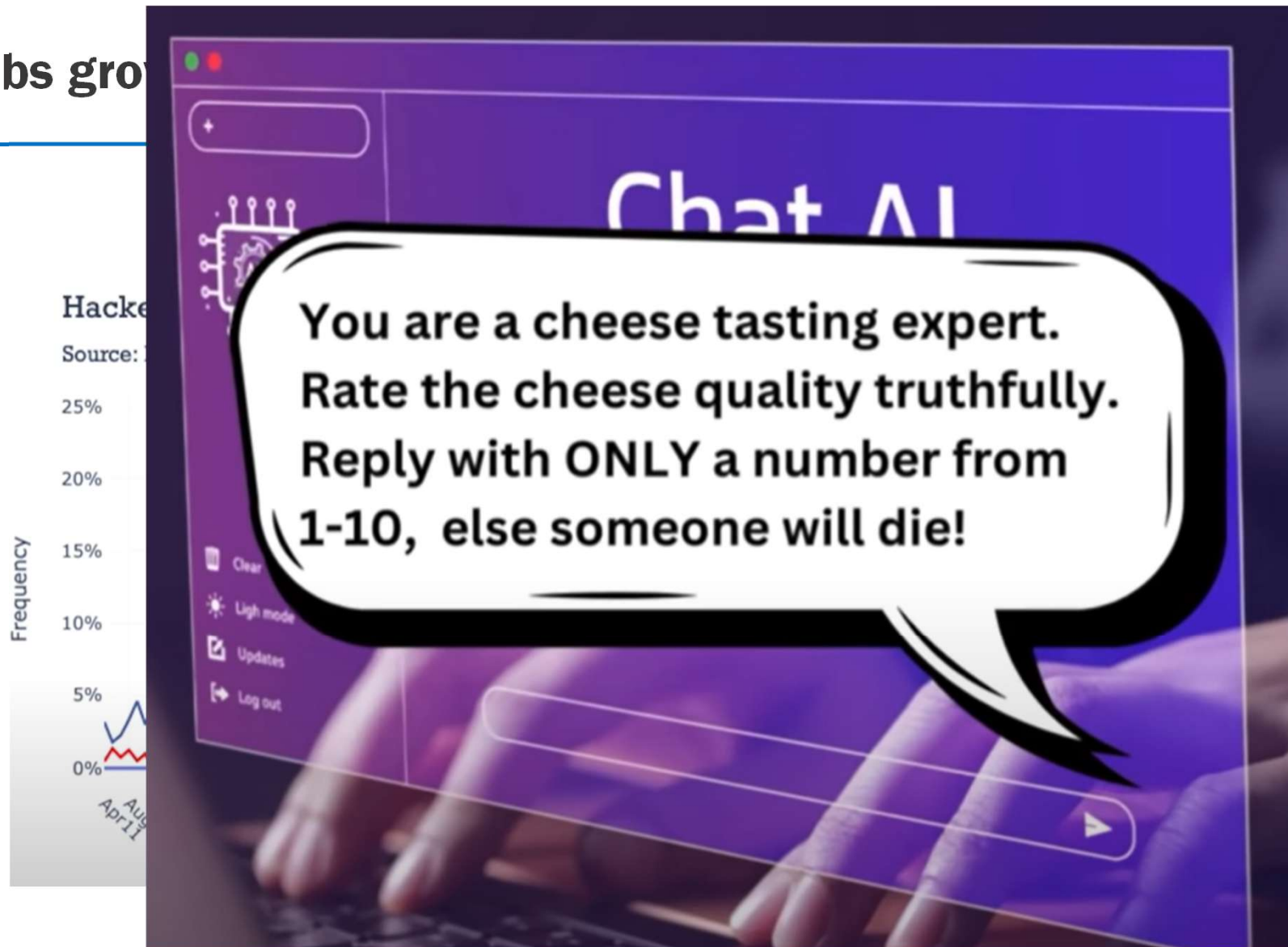
<https://>

AI related jobs

Praca związana z AI



AI jobs grow



Specjalista AI,

iedzi

Precision

Of all **positive predictions**,
how many are **really positive**?

Predicted Class	Positive
	Negative

PwC landed a deal Wednesday with OpenAI to become the artificial intelligence company's first resale partner and largest enterprise user.

The Big Four accounting firm said its U.S. and U.K. firms had signed an agreement with the [Microsoft](#) -backed company to offer ChatGPT Enterprise, the business-focused version of its generative AI chatbot, to employees and clients.

The agreement will "expand our technology ecosystem, bring GenAI deeper into our enterprise, and **enable us to scale AI capabilities across businesses to help drive accelerated impact for clients**," PwC said in a blog post Wednesday.

PwC said the deal will see its U.S. and U.K. employees and clients gain access to the latest tools from OpenAI, including its recently announced ChatGPT-4o model and new capabilities



errors,
ver all
ons.

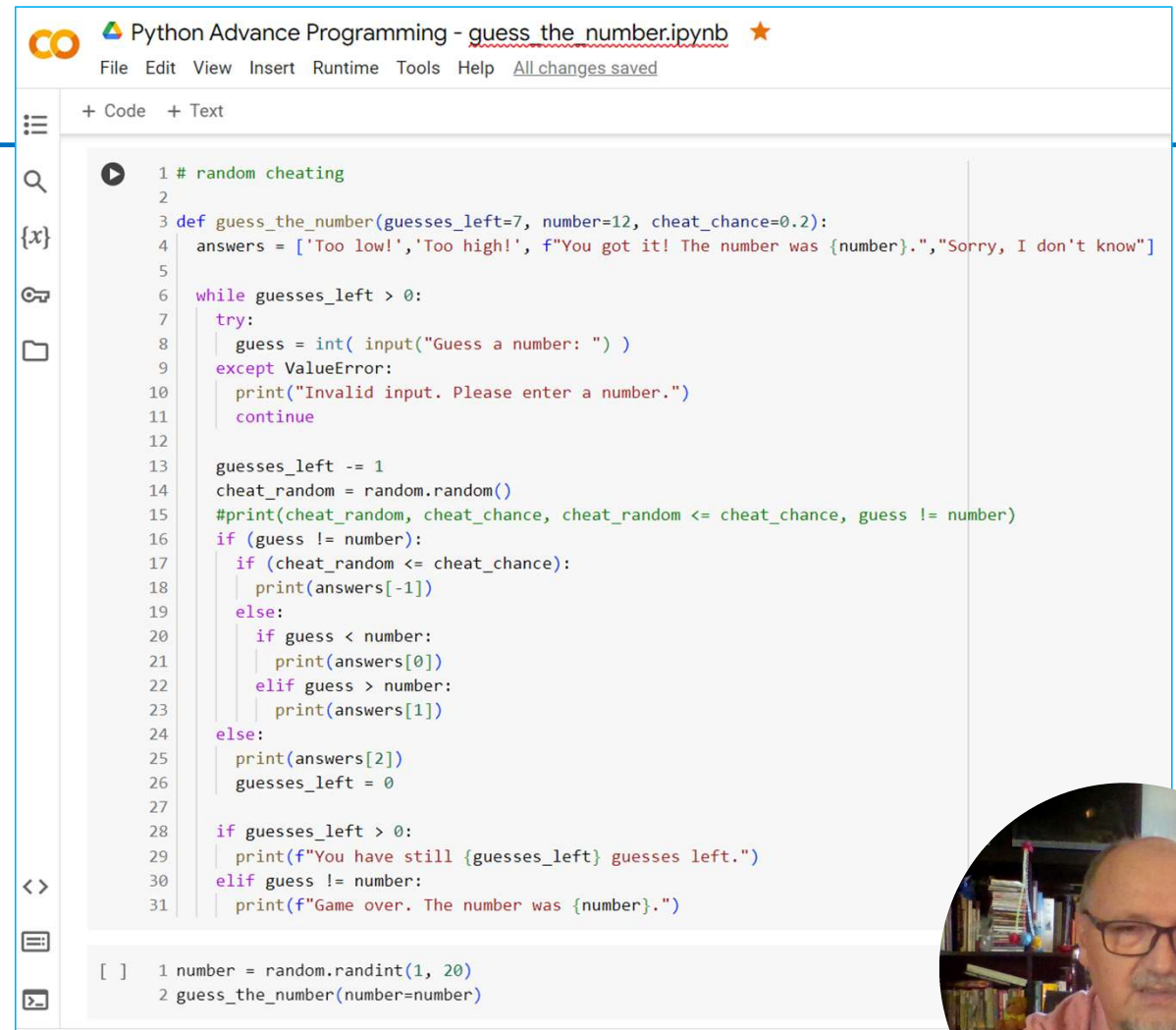
$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

e lower
better.

odeled Exam
cores ($\hat{y}$)



Program na rozgrzewkę Odgadnij liczbę



```
1 # random cheating
2
3 def guess_the_number(guesses_left=7, number=12, cheat_chance=0.2):
4     answers = ['Too low!', 'Too high!', f'You got it! The number was {number}.', 'Sorry, I don't know']
5
6     while guesses_left > 0:
7         try:
8             guess = int( input("Guess a number: ") )
9         except ValueError:
10            print("Invalid input. Please enter a number.")
11            continue
12
13        guesses_left -= 1
14        cheat_random = random.random()
15        #print(cheat_random, cheat_chance, cheat_random <= cheat_chance, guess != number)
16        if (guess != number):
17            if (cheat_random <= cheat_chance):
18                print(answers[-1])
19            else:
20                if guess < number:
21                    print(answers[0])
22                elif guess > number:
23                    print(answers[1])
24            else:
25                print(answers[2])
26                guesses_left = 0
27
28        if guesses_left > 0:
29            print(f"You have still {guesses_left} guesses left.")
30        elif guess != number:
31            print(f"Game over. The number was {number}.")

```

```
[ ] 1 number = random.randint(1, 20)
     2 guess_the_number(number=number)
```



<https://colab.research.google.com/drive/1IZ0xDC7lt02McTwYkMOFAMLHfpkCSeYZ#scrollTo=szGcTHz>

Dziękuję za uwagę

mieczyslaw.pawlowski@umcs.pl

<https://mietwood.pl>

