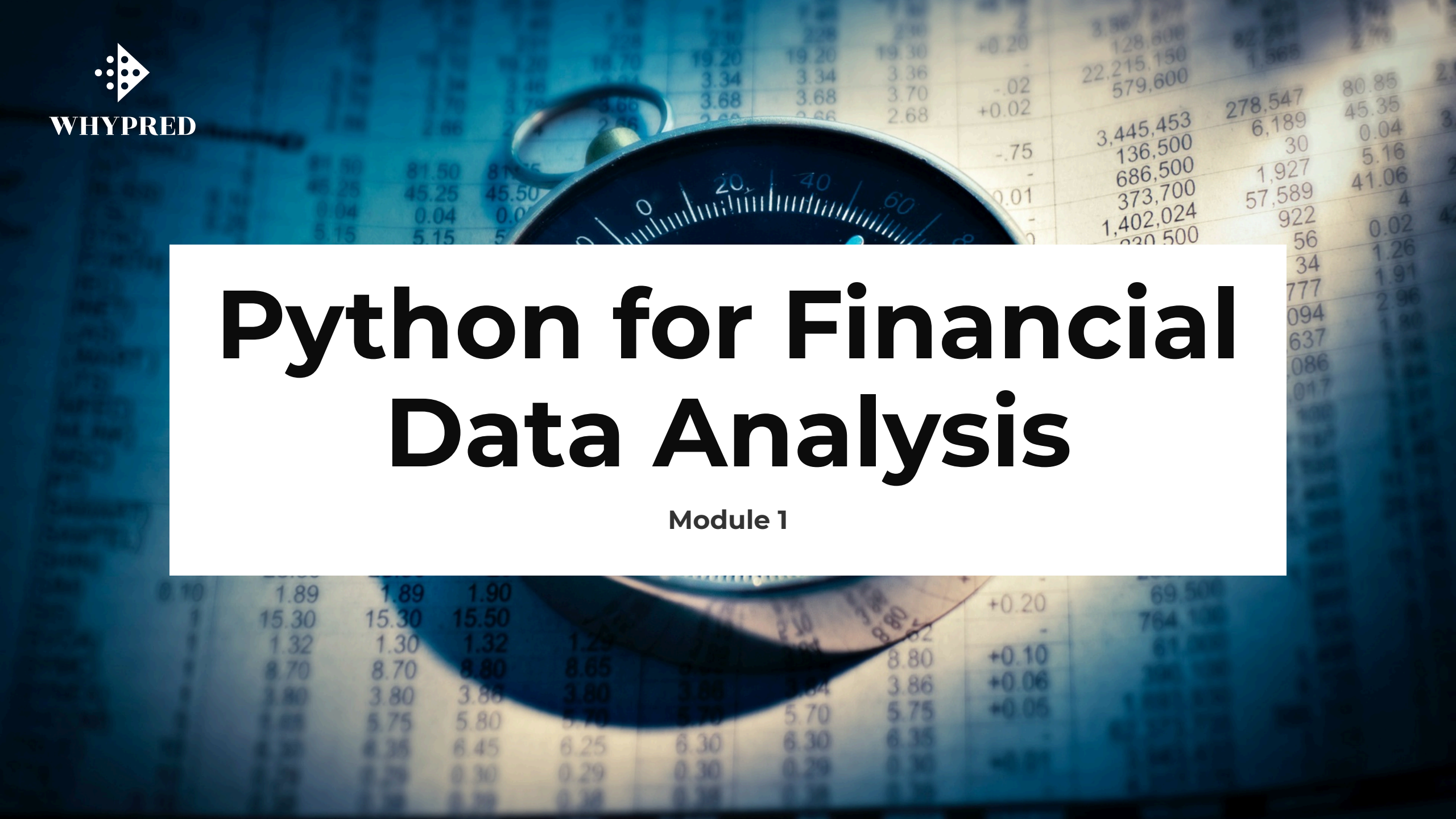


WHYPRED

Python for Financial Data Analysis

Module 1

Session Map

WHYPRED

1 **Welcome**

About, Outline, Status Quo

2 **Why Python?**

What makes it the tool of choice of data analysis

3 **Environment Setup**

Preparations and introduction to Google Colab

4 **Python Fundamentals**

Variables, Data Structures and Control Flows

BIO

Michael Wang

WHYPRED

- 14 years' experience across tech, investing, and teaching
- Worked as an analyst in the financial industry for almost 10 years, then made the career switch to being a data scientist
- Has taught python, data science, investments, AI at **University of Sydney**, **CFA Society Australia** and **Chartered Institute of Professional Certifications**. Currently teach at **Kaplan Business School**
- Founded **WhyPred** in 2021; consulting in investment analytics, AI, data science and learning and development

WHYPRED

“I'm a {x} , my role is to {y}. I have to work with a lot of data to, which can be repetitive and time consuming„

PROBLEM STATEMENT

Why Python for Financial Data Analysis?

WHYPRED

- There are many courses that teach python but the majority are taught by developers and/or data scientists with backgrounds ***rooted in tech/IT***.
- While this does give a pure, holistic approach to learning python, it may not address the need of many professionals who's occupation is not that of a developer or data scientist but still ***work a lot with data***.
- Their needs are rooted in being able to ***apply python*** to their specific use cases
- This course aims to equip you with the ***right parts of python*** to help you to ***work practically with data in your domain*** but at the same time ingrain some of the ***relevant best practices from tech***.

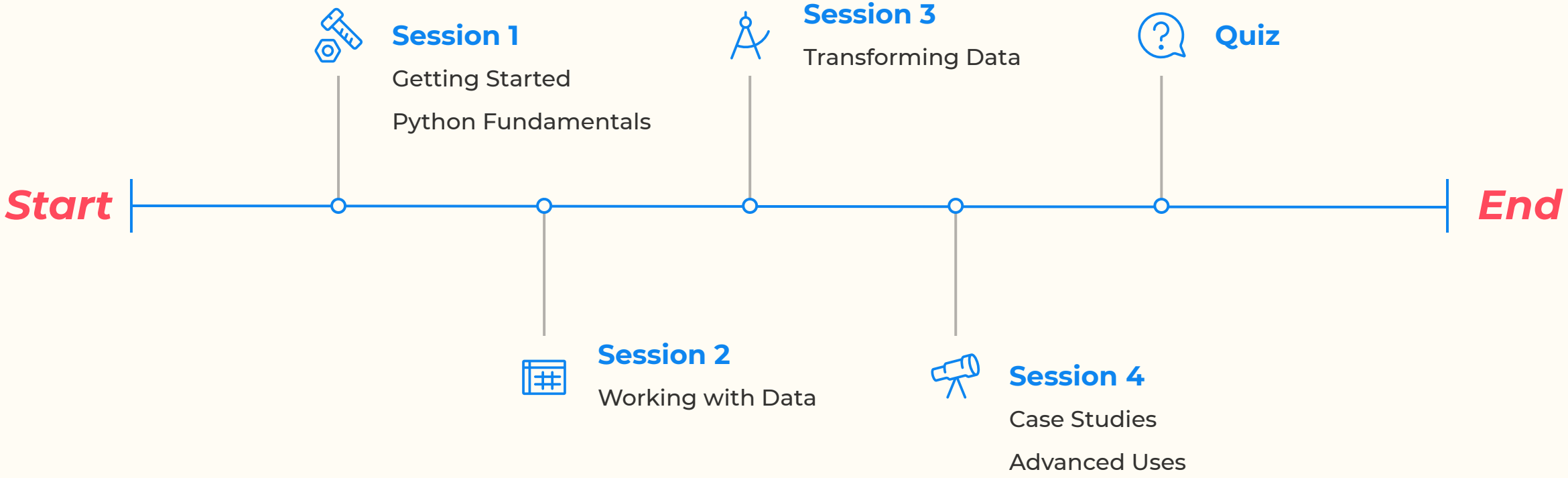
Why Python for Financial Data Analysis?

Domain Knowledge

Technical Expertise

Course Outline

WHYPRED



Status Quo

WHYPRED

- Excel has been the cornerstone of analysis and data storage for good reason:
  - Ease of use
  - Versatile
  - Transferrable across different organizations
  - Visual interface
- But the needs of organizations and analysts have evolved and Excel may not be the answer for every problem any more
  - The amount of data we're working with nowadays have increased
  - The types of calculations needed have become more complex
  - Lack of automation and reproducibility
  - Lack of controls and governance
  - Doesn't handle unstructured data well
  - Limited analytical capabilities

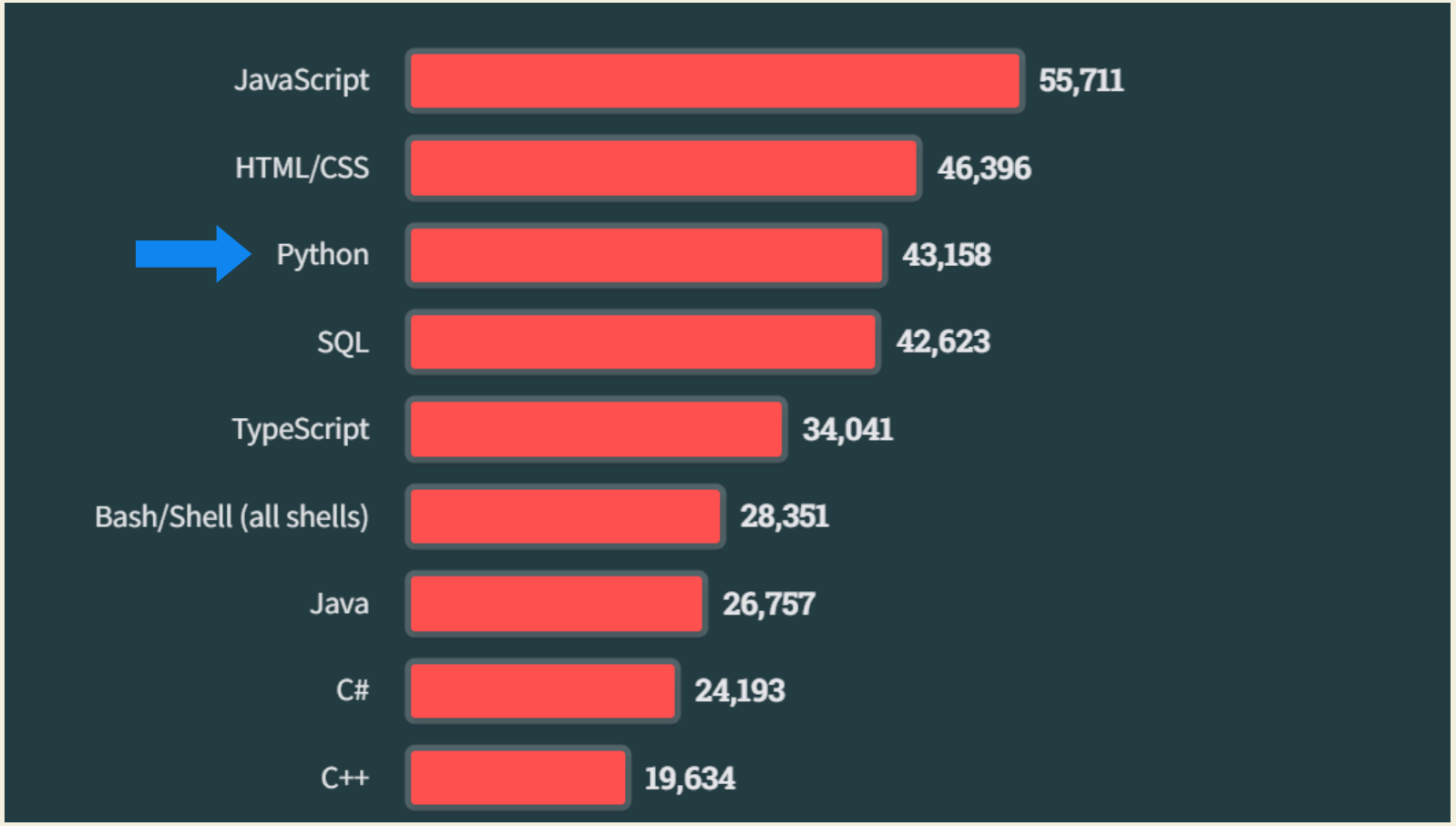
Is there a better way?

WHYPRED

Python is a high-level, interpreted programming language known for its simplicity, readability, and versatility. It has become the de facto language of AI, Data Science and Data Analysis

Stack Overflow Developer Survey 2023

WHYPRED



https://survey.stackoverflow.co/2023/

Features

WHYPRED

- **Automate Reporting**

  Using libraries python can streamline and automate repetitive processes like reporting and save time and reduce errors

- **Complex Calculations**

  Python enables efficient application of complex calculations on large datasets, making it ideal for financial modeling, scientific simulations, and statistical analysis.

- **Data Cleaning / Preprocessing**

  Python simplifies data cleaning tasks such as handling missing values, removing duplicates, reshaping data, and merging datasets. Its string manipulation capabilities also allow for processing of text data

- **Data Science / ML**

  Python offers rich and well-featured libraries for data science, machine learning and AI making it often the language of choice for these tasks.

Setting Up

Notebooks vs IDE

WHYPRED

There a few ways to work with Python, the easiest is with notebooks such as ***Jupyter Notebooks*** and ***Google Colab***.

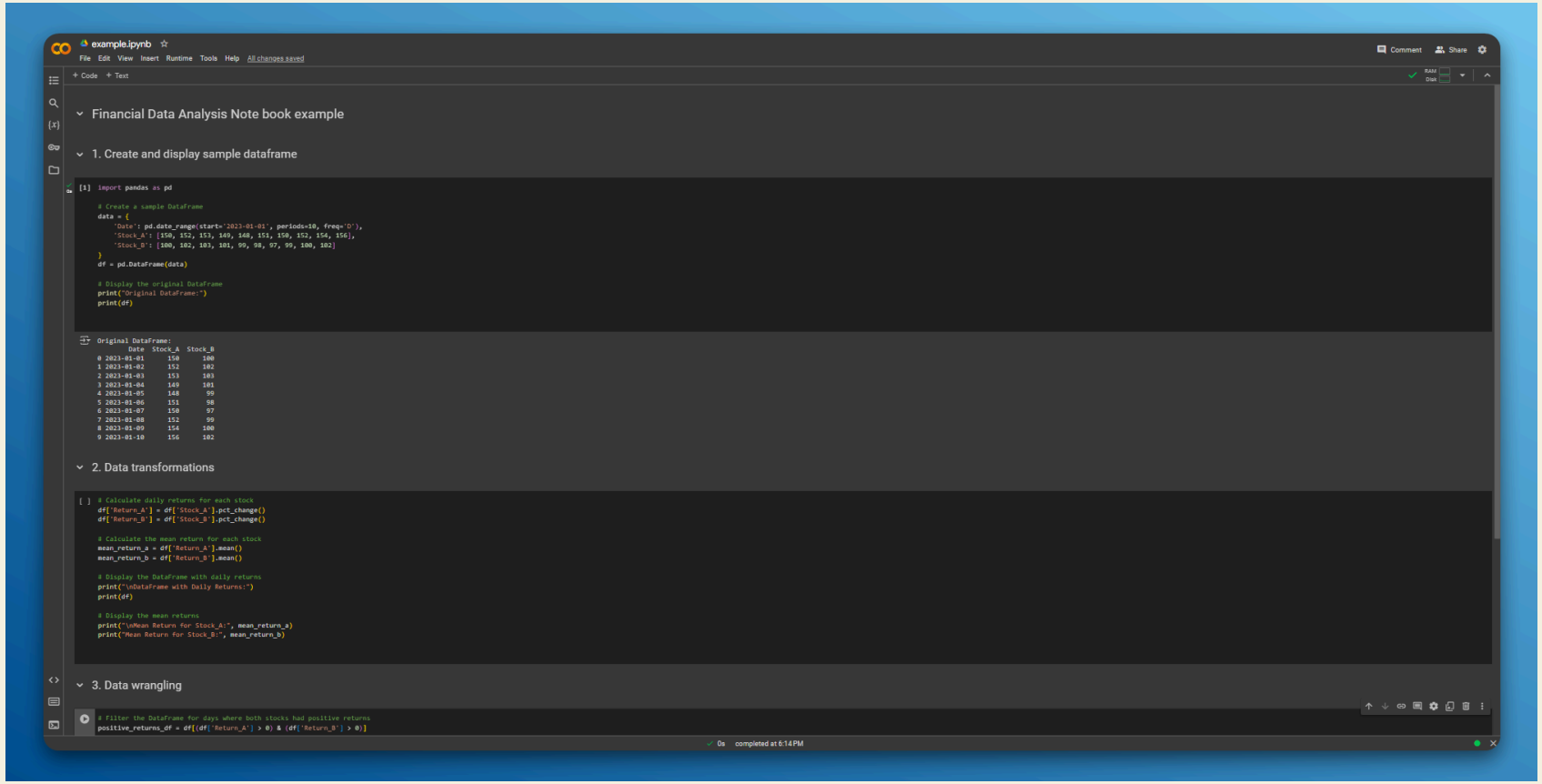
Notebooks are beginner friendly and integrate well with analytic workflows

It is recommended that as you get more proficient with python that the transition to IDEs are made.

The choice of IDE vs Notebook depends on which one you have access to in your environment and preference

WHYPRED

Notebook (Google Colab)



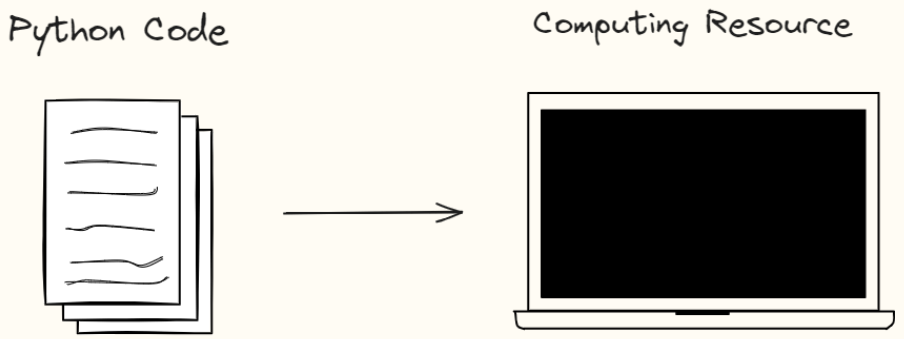
Google Colab Notebooks

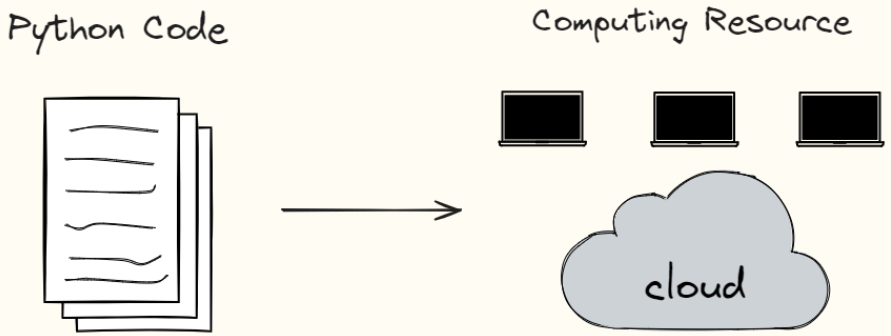
WHYPRED

- Google Colaboratory (Colab), is a cloud-based platform that allows users to write and execute Python code directly in their web browser.
- It is widely used for data analysis, machine learning, and scientific computing.
- Let's you work with python code like a regular notebooks, with code chunks, text, visualizations and charts

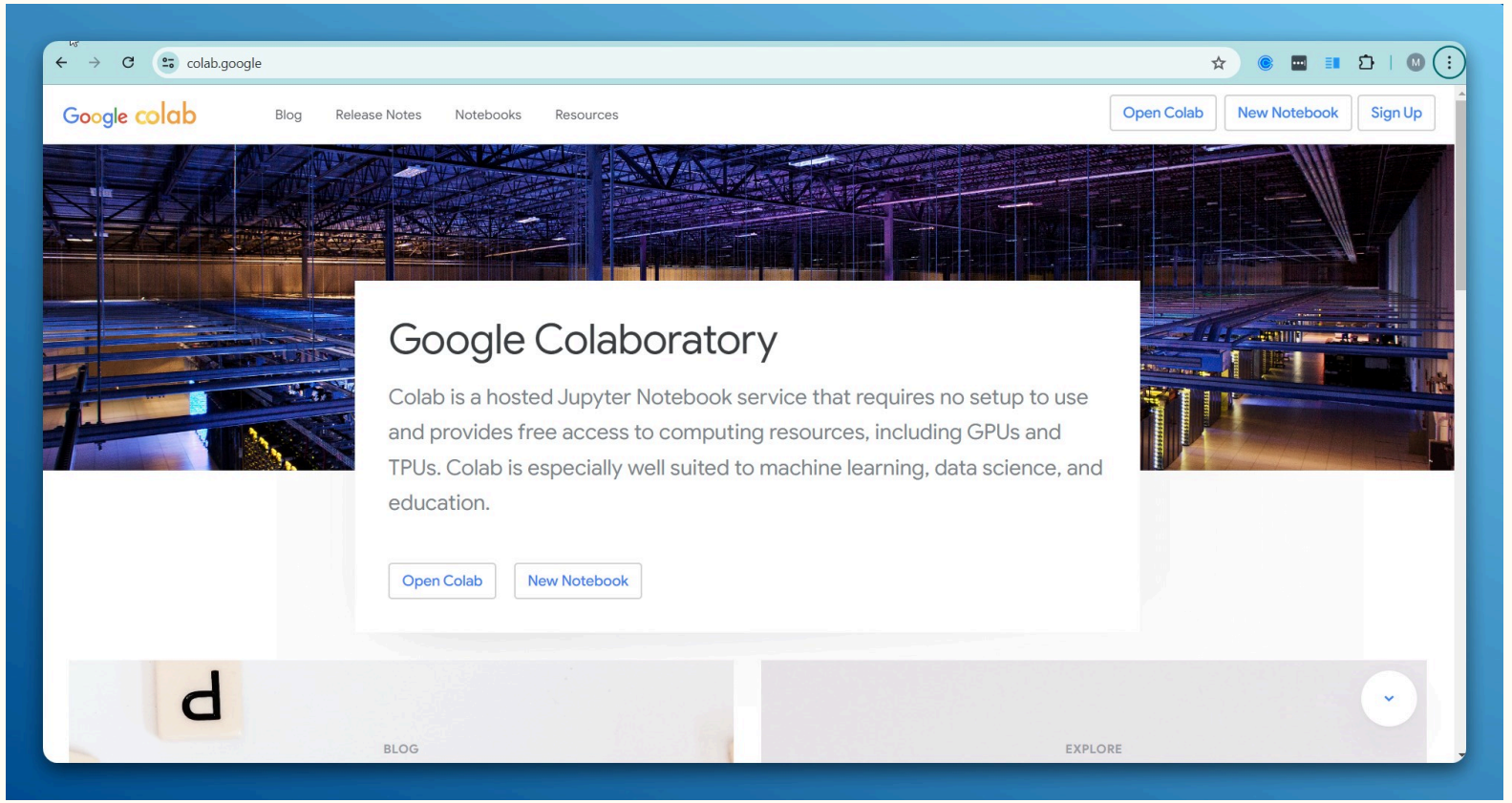
Google Colab Notebooks

WHYPRED





Getting Started



colab.google

Getting Started

WHYPRED

1.

You will need a Google Account, if you have one please make sure you're signed in if not create a Google account ***here***

2.

Navigate to colab.google

3.

Click "New Notebook" in top left corner. This will open a new notebook

4.

Let's create out first program and have a look around

Things to Know

WHYPRED

- ## Commenting

  We can leave comments in our code by using the the hash prefix "#", these act as references for you and other analysts when you have to collaborate on analysis e.g

  # this function helps us to calculate the rolling

  # average sales

- ## Indenting

  Python is sensitive to whitespaces. This is because indenting/spaces are used to determine the hierarchy of the code so extra spaces may cause the code not to run properly e.g

  print("Hello World")

  print("Hello Again")

- ## Text Values

  Text values can be created in a number of different ways, using single quotes, double quotes, or triple single quotes or triple double quotes eg

  'this is text' is the same as ***"this is text"***

  or we can have

  '''this is

  text'''

- ## Case Sensitive

  print("hello world") is not the same as ***Print("hello world")***

- ## Base 0 index

  Python starts counting at 0 not 1!

Getting Started

WHYPRED

1.

You will need a Google Account, if you have one please make sure you're signed in if not create a Google account ***here***

2.

Navigate to colab.google

3.

Click "New Notebook" in top left corner. This will open a new notebook

4.

Let's create out first program and have a look around

Getting Started

WHYPRED

1.

You will need a Google Account, if you have one please make sure you're signed in if not create a Google account ***here***

2.

Navigate to colab.google

3.

Click "New Notebook" in top left corner. This will open a new notebook

4.

Let's create out first program and have a look around

WHYPRED

Fundamentals

WHYPRED

```
############################################################
# Storing Variables

# Float
stock_price = 120.25

# Integer
stock_shares_held = 100

# String
stock_ticker = "NVDA"

# Boolean
stock_listed = True


print("Price:", stock_price)
print("Shares Held:", stock_shares_held)
print("Ticker:", stock_ticker )
print("Listed Stock:", stock_listed)


## Output

## Price: 120.25
## Shares Held: 100
## Ticker: NVDA
## Listed Stock: True

############################################################
```

Variables

- Variables are the fundamental units of storage in Python.
- They allow you to store data that can be referenced and manipulated throughout your code.
- Common types of variables encountered in data analysis include floats, integers, strings and booleans

Data Structures

WHYPRED

Data structures in Python are ways to organize and store data efficiently. They allow you to manage, access, and manipulate data in various formats

1. ***Lists:*** Ordered, mutable sequences that can contain elements of different types.
2. ***Tuples:*** Ordered, immutable sequences.
3. ***Dictionaries:*** Unordered collections of key-value pairs.
4. ***Sets:*** Unordered collections of unique elements.
5. ***Arrays:*** Efficient for storing large amounts of homogeneous data.

Operators

WHYPRED

Operators are symbols or keywords that perform operations on one or more inputs. They allow you to manipulate data and perform computations and apply logic. Operators can be categorized into several types

1. ***Arithmetic:*** Perform mathematical calculations (e.g., +, -, *, /, **)
2. ***Comparison:*** Compare values and return True or False results (e.g., ==, !=, <, >)
3. ***Logical:*** Combine True or False expressions, such as and, or, not
4. ***Assignment:*** Assign values to variables (e.g., =, +=, -=).

Control Flow

WHYPRED

The order in which individual statements are carried out in the code. It determines how a program proceeds and makes decisions based on certain conditions. Control flow structures allow you to do things like execute code conditionally, repeat parts of the code or skip to other parts of the program. Common structures in control flows include:

1. ***Loops:*** ***R***epeat operations (e.g. for loop, while loop)
2. ***Conditional Statements:*** Execute code based on whether conditions are met (e.g. if, elif, else)
3. ***Exception Handling:*** Manage errors and exceptional situations

Notebook Exercise

WHYPRED

1-1_python_fundamentals.ipynb

- All notebooks and slides are available here
- Remember Google Colab is a shared cloud service, everyone is looking at the same notebook!
- To prevent accidental changes the notebooks are read only, at the beginning of each exercise make sure you create a copy so that you can edit your own copy!

Notebook Exercise

WHYPRED

Functions are reusable chunks code of code that are defined using the ***def*** keyword

e.g. imagine we want to get the average monthly sales of business over 3 months, in python the code might look something like this:

```
total_sales = 100 + 200 + 300

average_sales = total_sales / 3
```

But we have to repeat the code for each quarter, solution:

```
def avg_sales(sales_m1, sales_m2, sales_m3):
    result = (sales_m1 + sales_m2 + sales_m3) / 3
    return(result)
```

That's it for module 1!

Disclaimer

WHYPRED

The information contained in this slide pack has been prepared by WhyPred Pty Ltd ("WhyPred") for informational purposes only. The data, analysis, and opinions expressed herein are based on sources believed to be reliable and provided in good faith, but no representation or warranty, express or implied, is made as to its accuracy, completeness, or correctness.

In addition, sample data has been used in this presentation and should not be relied upon for accuracy. This presentation does not constitute investment advice, nor is it an offer or solicitation to buy, hold, or sell any securities or financial instruments. Any projections, forecasts, or estimates herein are indicative only and subject to change without notice. Past performance is not indicative of future results.

Investors should seek their own independent financial, legal, and tax advice before making any investment decision. WhyPred and its affiliates, directors, employees, or agents accept no liability whatsoever for any loss or damage arising from any use of this document or its contents. By accessing and using this slide pack, you agree to the terms of this disclaimer.