


Link to the material:

<https://github.com/voltrondata-labs/2024-arrow-format-tutorial>





Joris Van den Bossche, Alenka Frim and Raúl Cumplido

A deep dive into the Arrow Columnar format with pyarrow and nanoarrow



VOLTRON DATA

24 April, 2024



Agenda

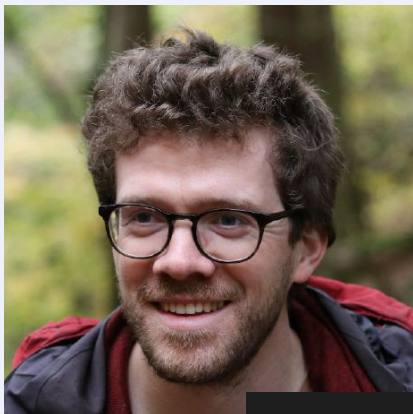
- 01 Who we are
- 02 Introduction
- 03 The Arrow Columnar format
- 04 Primitive layouts
- 05 Nested layouts
- 06 All types and extension types
- 07 The Arrow C Data Interface

[Link to more images](#)



01 Who we are

• Joris Van den Bossche



github.com/jorisvandenbossche

@jorisvandenbossche@fosstodon.org

1. Background

PhD bio-science engineer, air quality research

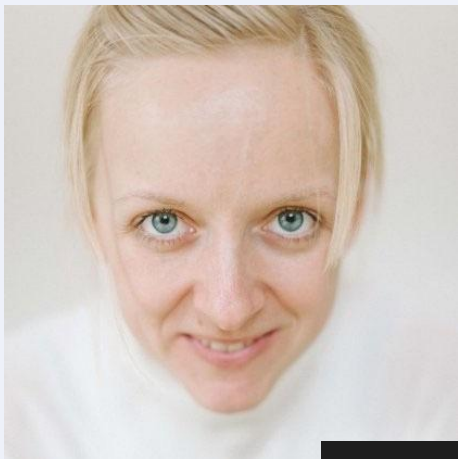
2. Open source

Open source enthusiast: core developer of pandas, GeoPandas, Shapely, Apache Arrow PMC, ...

3. Work

Software engineer at Voltron Data

• Alenka Frim



github.com/AlenkaF

1. Background

MSc in Mathematics, missing data

2. Open Source

Apache Arrow comitter, interested in interchanging data and supportive OS

3. Work

Software Engineer at Voltron Data

• Raúl Cumplido



github.com/raulcd

1. Background

Telecommunications engineer

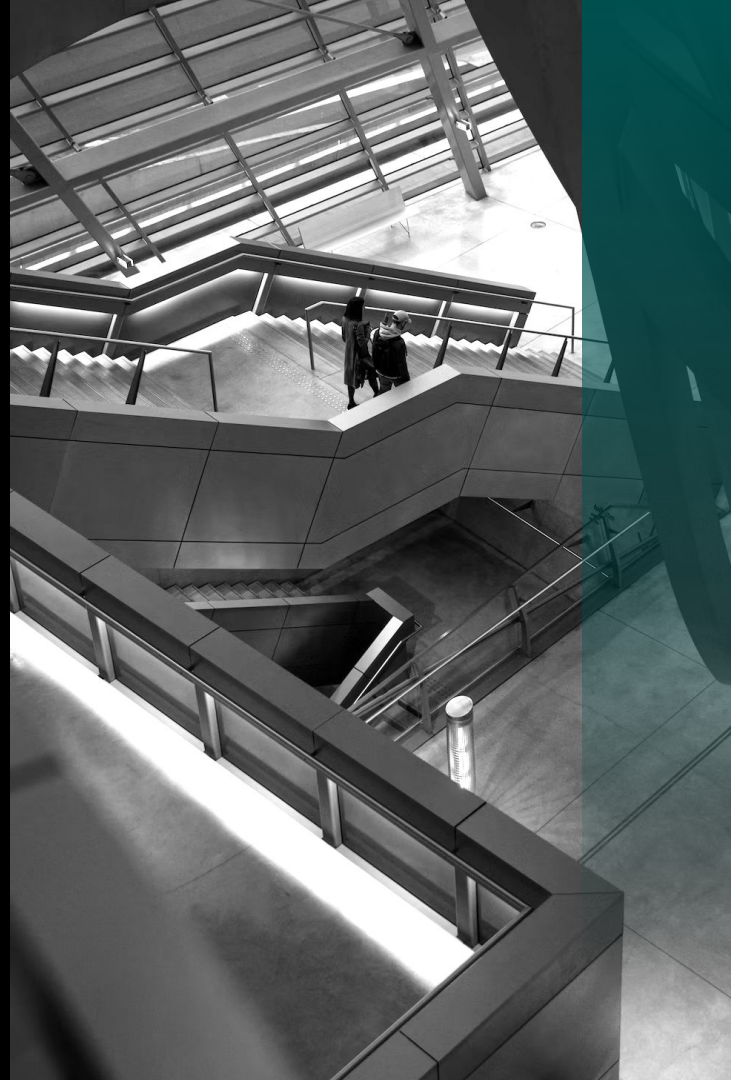
2. Open Source

Apache Arrow PMC - release manager

3. Work

Software Engineer at Voltron Data

Notebook time



Questions? 
Thank you! 

