

VAEX:

**1 BILLION ROWS, 1 LAPTOP,
SERIOUS DATA SCIENCE**

JOVAN VELJANOSKI

Sr. Data Scientist @ XebiaLabs

UNCOMFORTABLY LARGE DATA

Working with %i samples without going to the cloud:

- ➔ $< 1_000_000$ samples
- ➔ $\sim 10_000_000$ samples
- ➔ $\sim 100_000_000$ samples
- ➔ $\sim 1_000_000_000$ samples
- ➔ larger datasets

VAEX.IO: WHO ARE WE?



Jovan Veljanoski

Former astrophysicist

Sr. Data Scientist @ XebiaLabs

Co-founder of vaex.io

✉ jovan.veljanoski@gmail.com

🌐 <https://www.linkedin.com/in/jovanvel/>



Maarten Breddels

Former astrophysicist

Freelancer / consultant / data scientist

Core Jupyter-Widgets developer

Founder of vaex.io

Principal author of [vaex](https://vaex.io)

✉ maartenbreddels@gmail.com

🌐 www.maartenbreddels.com

🐦 [@maartenbreddels](https://twitter.com/maartenbreddels)

🐙 github.com/maartenbreddels



Yonatan Alexander

Head of Data Science at BuiltOn

✉ jonathan@xdss.io

🌐 <https://www.linkedin.com/in/xdssio/>



Mario Buikhuizen

Freelancer / consultant

Front-end / dashboards / widgets specialist

✉ mbuikhuizen@gmail.com

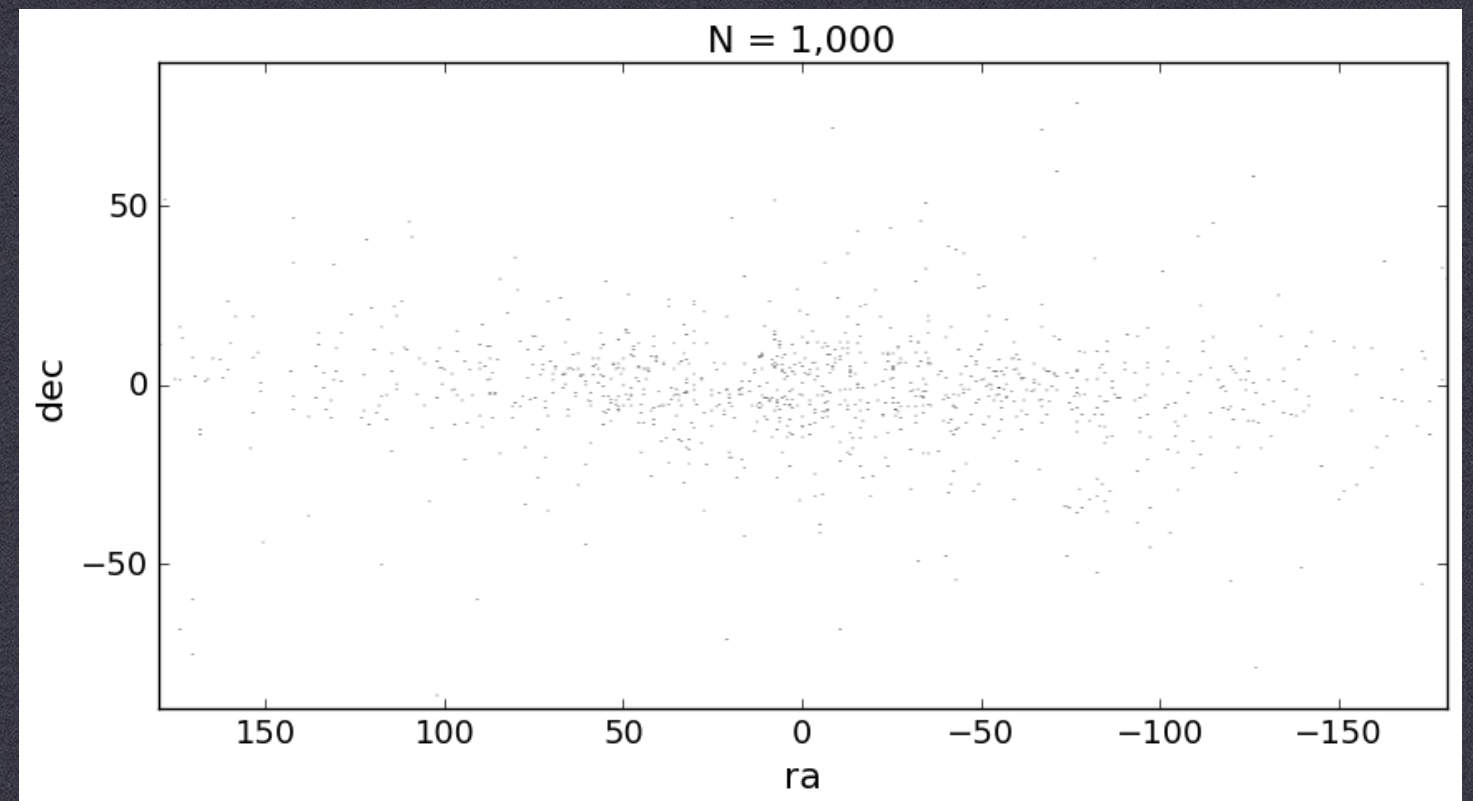
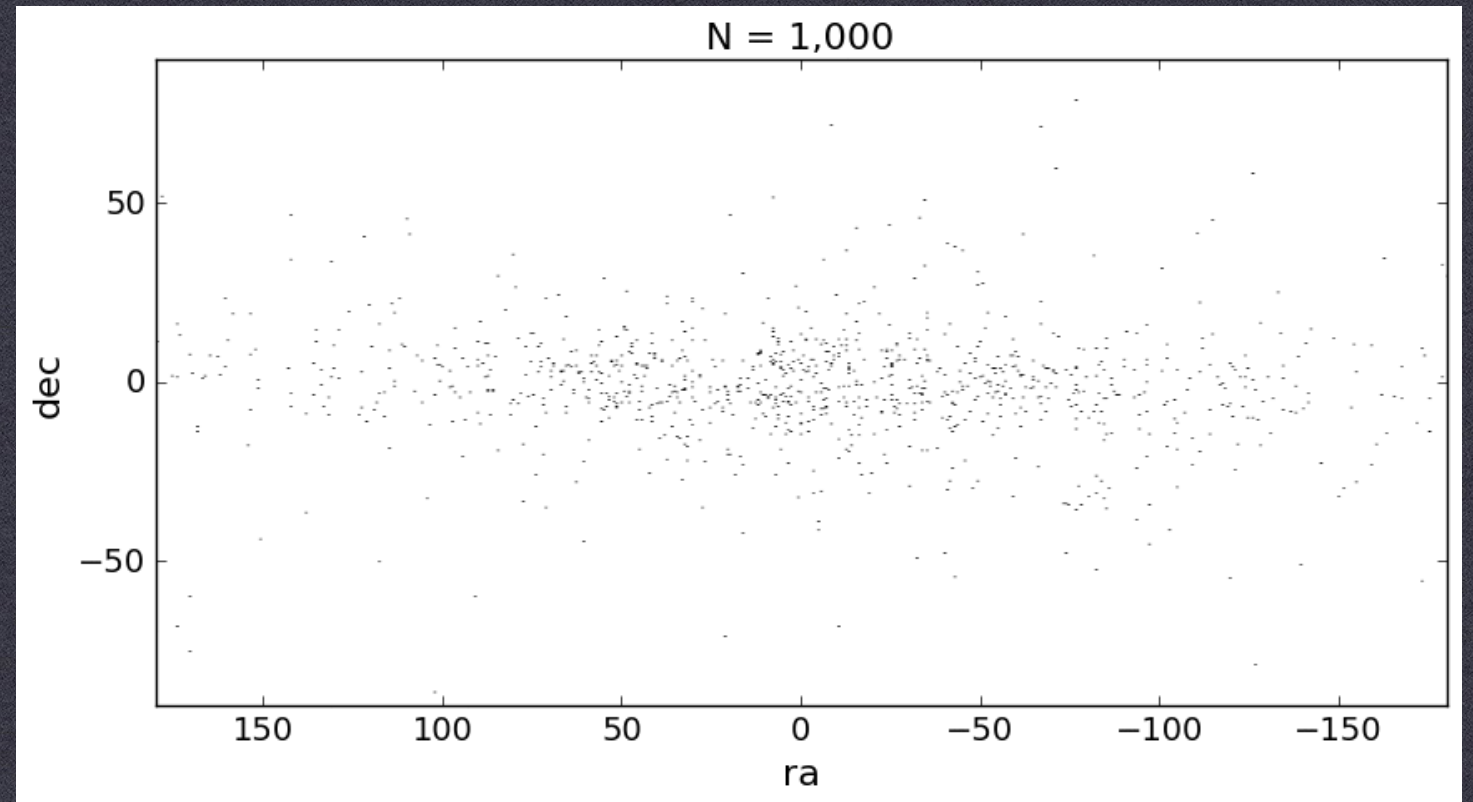
THE NEED FOR VAEX



The Gaia satellite:

More than 1 billion observations of stars in our Galaxy!

How do we work (explore, filter, visualize, analyze) with such data?



LIVE DEMO

The Jupyter notebooks presented at the live demo can be found at:

<https://github.com/vaexio/vaex-talks>