Polars vs. pandas

Comparison with Pandas

Similarities

★ Data handling and analysis

★ DataFrame and Series objects

★ Many of the same operations

★ Familiar syntax

Polars Syntax

```
df.group_by('col1').agg(pl.col('col2').sum()).collect()
```

Group `df` by `col1`   Sum `col2` for each group   Evaluate results

Python Polars looks like a cross between **pandas** and **PySpark**.

Comparison with Pandas

Similarities

★ Data handling and analysis

★ DataFrame and Series objects

★ Many of the same operations

★ Familiar syntax

Differences

★ No index names

★ More parallelism

★ Optional lazy evaluation

★ Better syntax?

*"Came for the **speed**, stayed for the **syntax**."*

-Polars Enthusiasts

dKd

DR KIM DATA

dask

★ Out-of-core processing

★ Parallelism, lazy evaluation

★ Python vs. Rust

★ Cluster vs. single machine

★ Pandas vs. unique syntax

APACHE Spark

★ Performant for big data

★ Parallelism, lazy evaluation

★ Java/Scala vs. Rust

★ Cluster vs. single machine

dKd
DR KIM DATA