Real-time data analytics/visualisation of big-data in Industry 4.0

Zhicheng Lin, Payam Soulatiantork, Mohammad Zandi, Joan Cordiner and Peyman Z. Moghadam

Department of Chemical and Biological Engineering, University of Sheffield, S1 3JD, UK
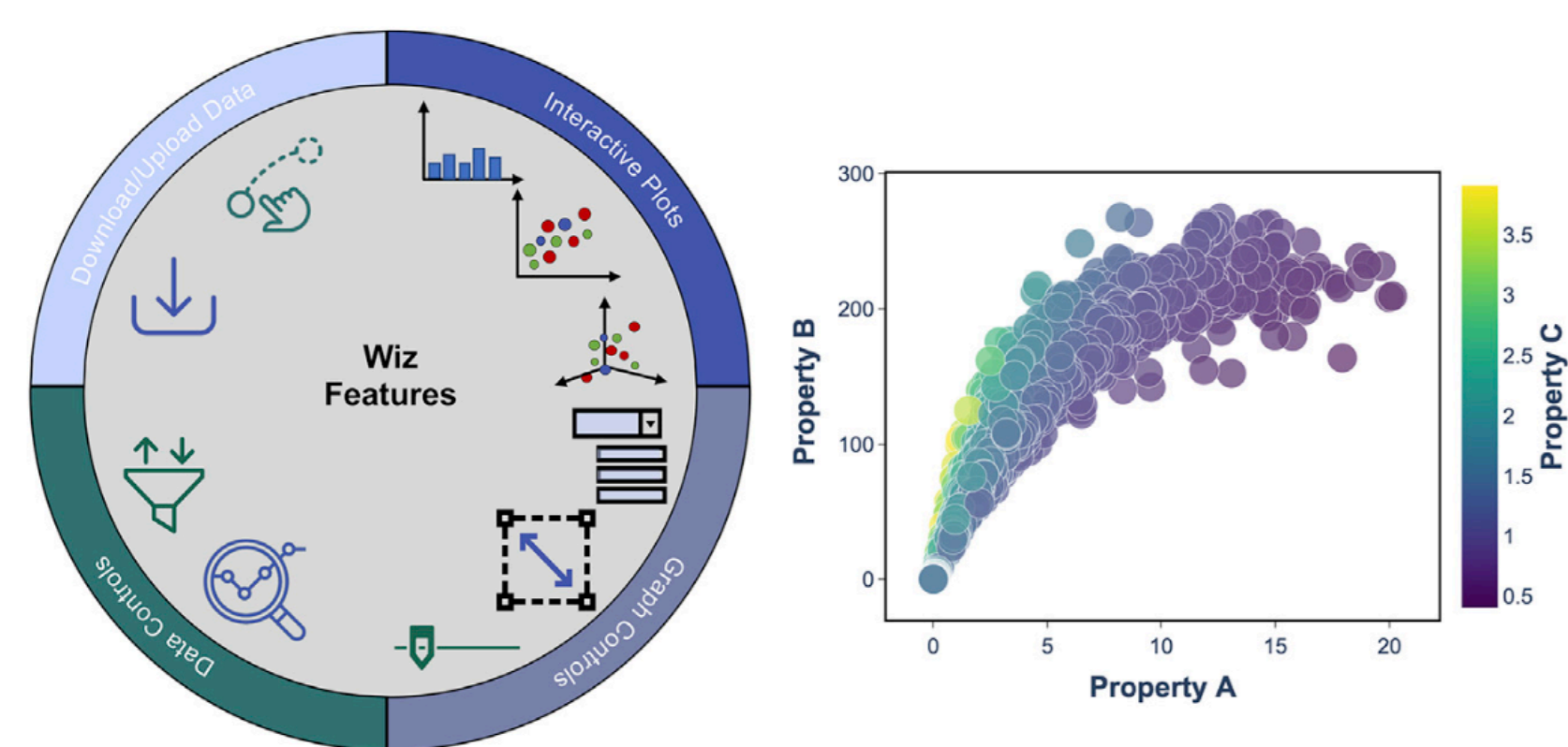
BACKGROUND

- Scientific data become bigger and have high dimension with complex relationship between variables.
- New simulation techniques led to the emerging of large-scale computational screening includes properties of millions materials.
- Standard data visualization fail to capture the full story of a complex dataset.
- Large datasets requires some degree of interactivity or user-controlled experience.
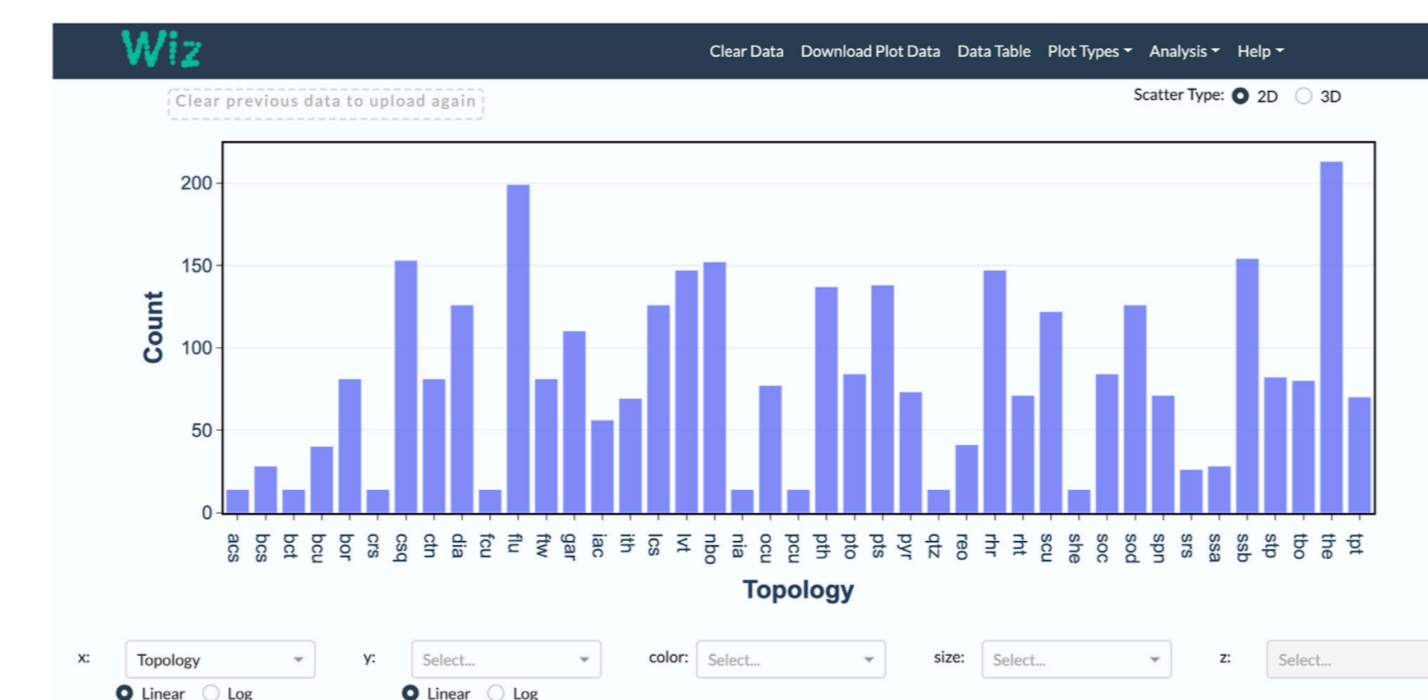
MOTIVATION

- Give readers increased ability to understand author's conclusion and living figures to enhance the underlying dataset.
- Create a web-based data visualization tool named Wiz with Dash to realize the live and high-dimension data analysis.

OVERVIEW OF WIZ

- Wiz provides a range of graph types with hover- and click-driven insights.
- Allow users to quickly switch between datasets and focus on what they need.
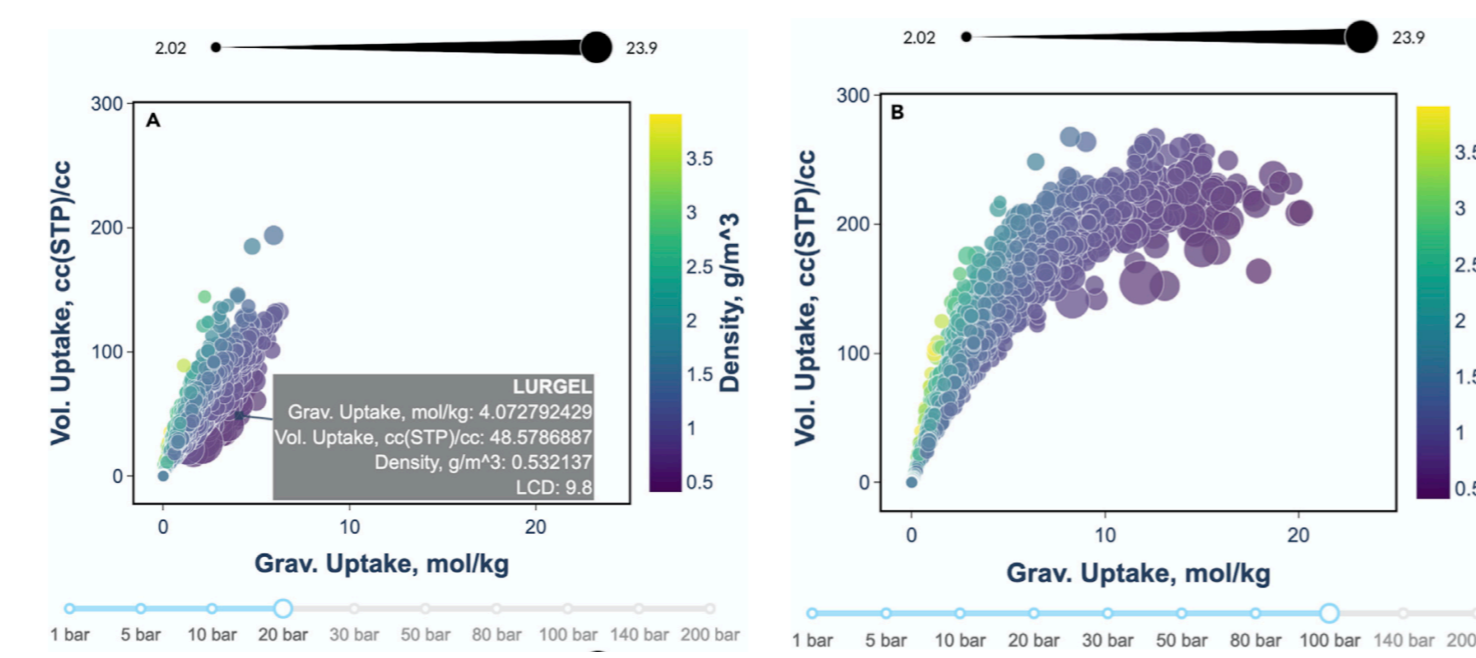- User-friendly.

FEATURES

Histogram of categorical data of over 3,000 metal-organic frameworks displayed in multiple graph types in Wiz.

- Explore the relationships across large and complex data easily.
- Web-based and accessible online.
- Not require users have programming skills to do visualization.
- Live visualization.
- High-dimension analysis.
- Built-in analysis and data filtering

- High-Throughput Screening for Materials Design and Discovery

Oxygen storage capacity for 3,000 metal-organic frameworks at different pressures plotted in Wiz. [1]
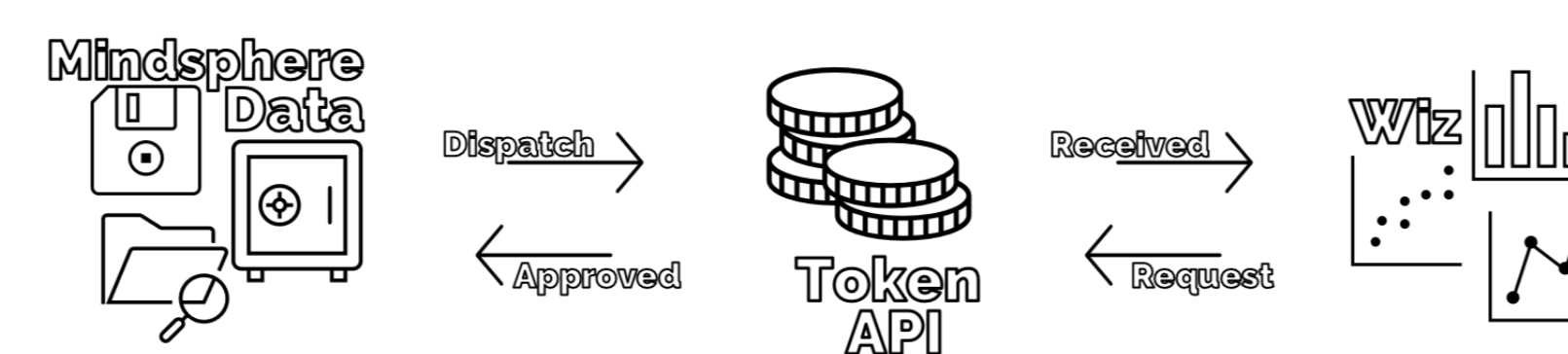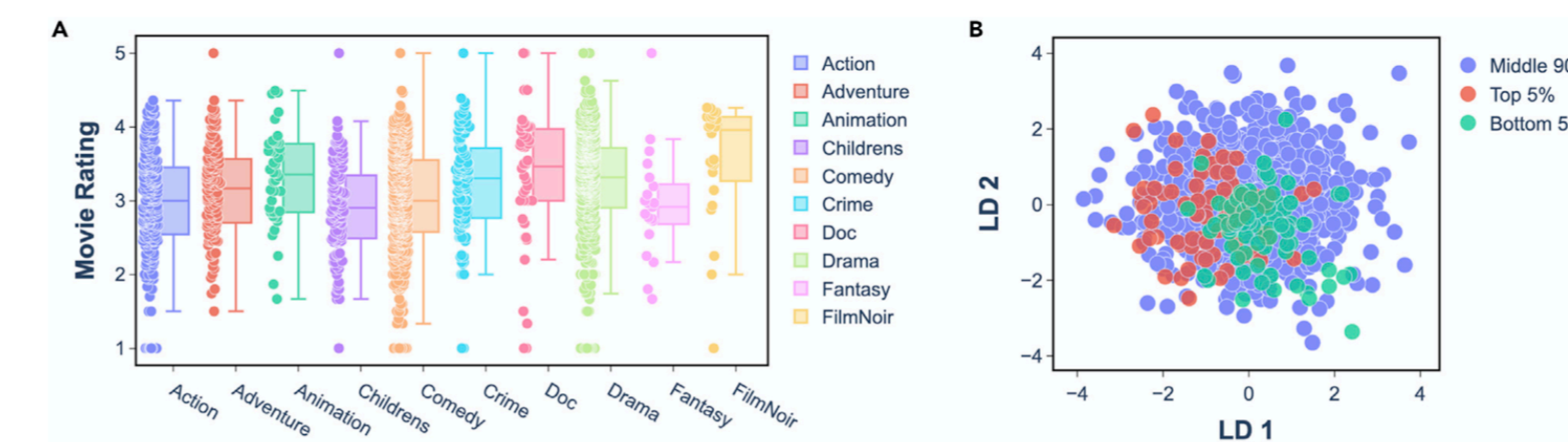

Diagram demonstrating how Wiz interacts with Mindsphere and requests and receives data.

- Access data through Siemens' Mindsphere using token management APIs for security purposes.

VISUALIZATION

- The following example illustrates how Wiz can be used at different stages of the machine learning pipeline.[1]
- Data from Movie Lens Dataset collected by Grouplens at the University of Minnesota.
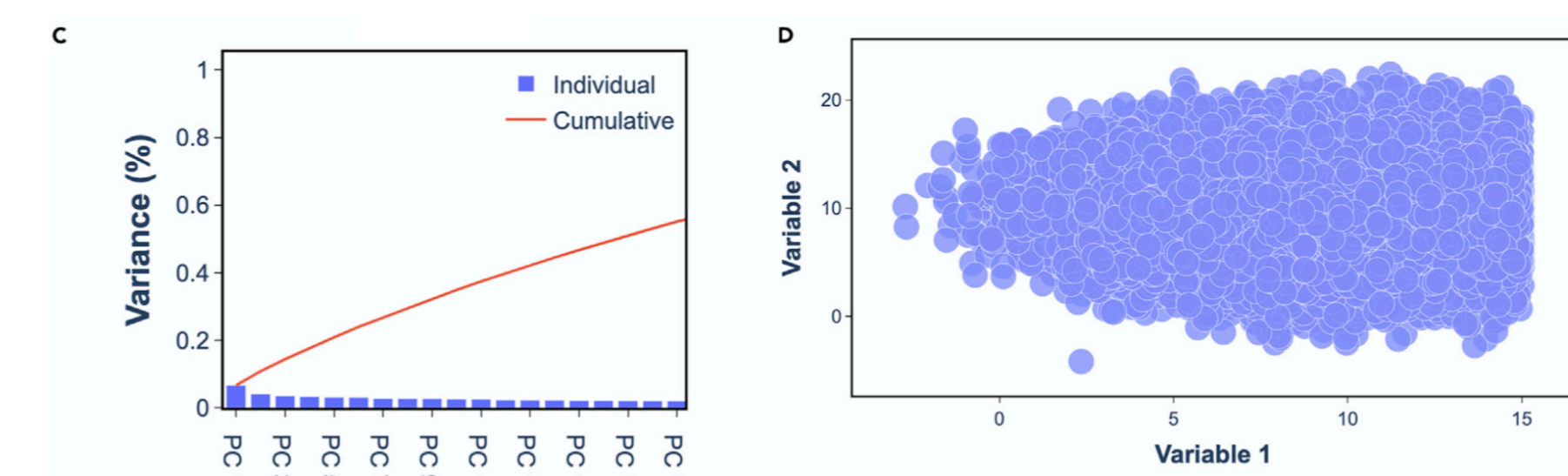

Average movie rating versus movie genre for dataset.

LDA projection using the top and bottom 5% rated movies as classes.

- Understanding the distribution and outlying data can aid in data processing before training.
- Wiz makes analysis of the structure fast and easy.
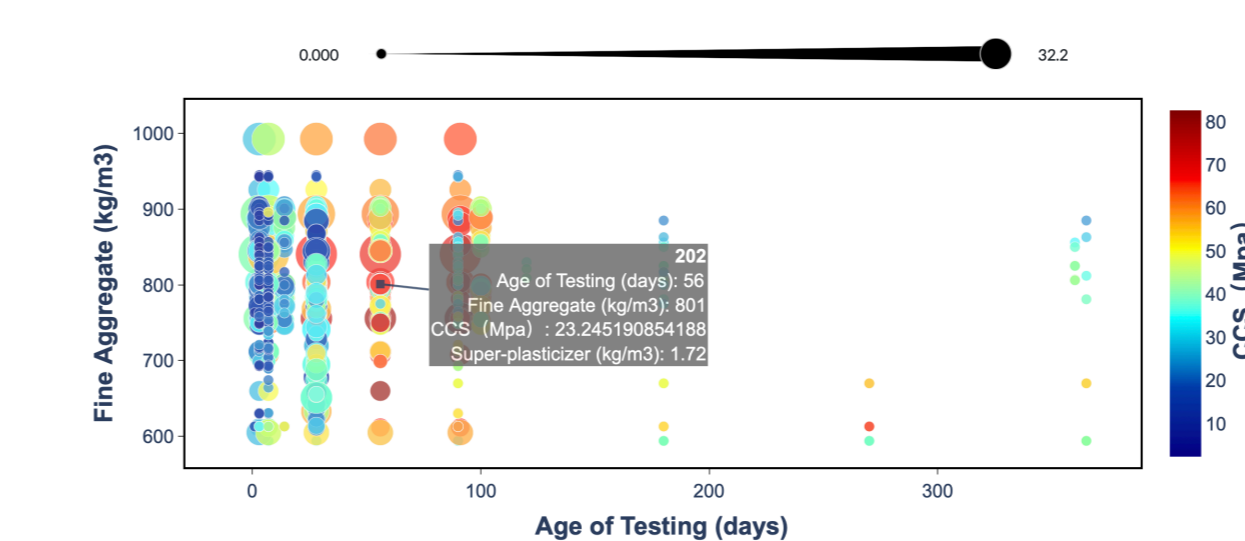

- PCA results (first 20 of 50 PCs shown).

Filtering process: the data are filtered such that data points above a threshold value of 15 (x-axis) is not plotted.

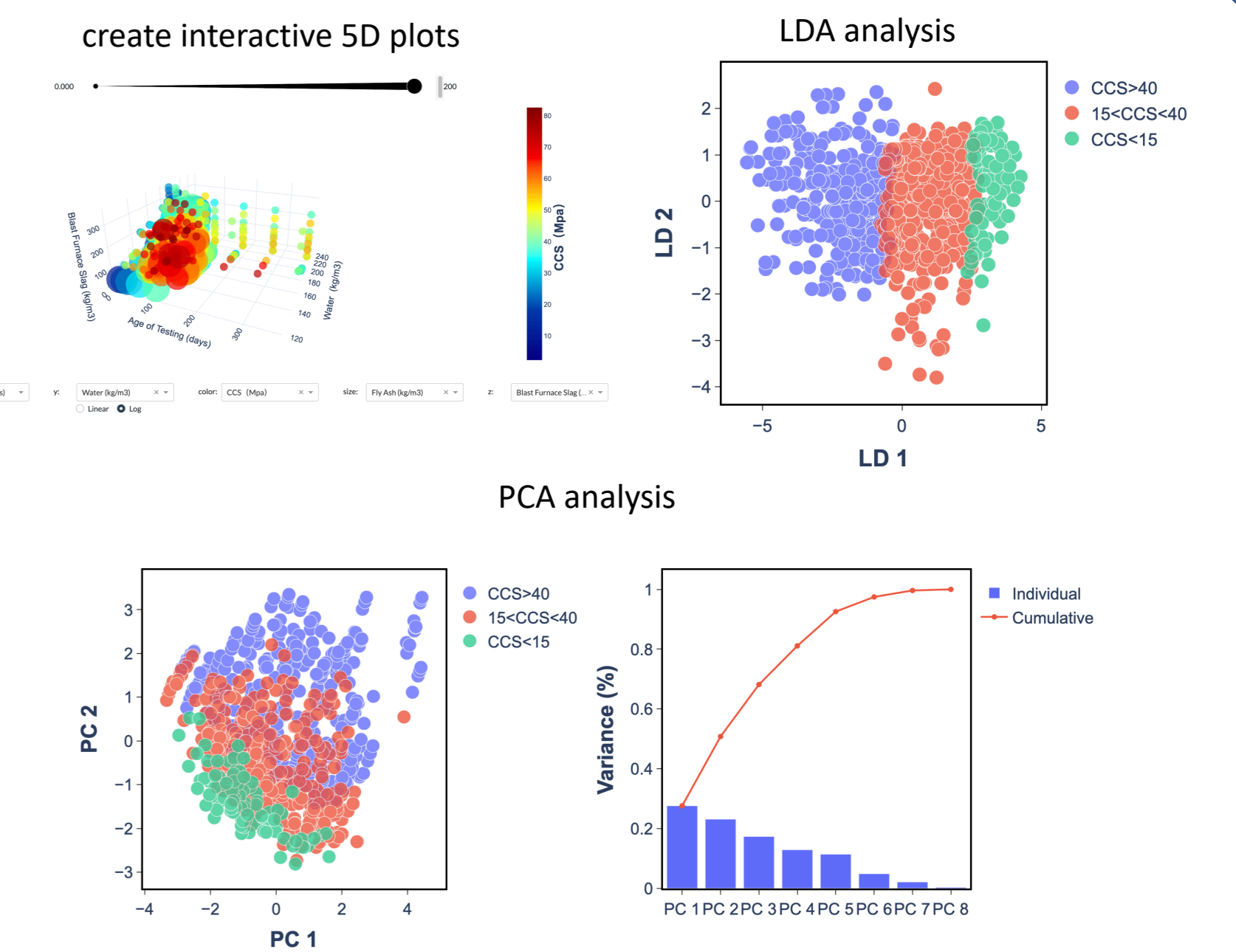- Filter data easily without editing the underlying dataset.

- Users not with relevant knowledge can also quickly find the relationship between relevant variables.

Example : Research of CCS.[2]
The concrete compressive strength (CCS) is a function of age and ingredients.

- Easy access to achieve each point of information.
- Visualize relationships.
- Transform combination between different variables.
- 4D visualization.

CONCLUSION

- Wiz is highly modular on the backend.
- Each page within the app contains different essential components that makeup the layout (i.e., links, dropdowns, upload buttons) that are implemented using Dash.
- Wiz removes the need to program routines for data upload, data filtering/processing, and the plotting commands.
- Combined with robust hosting through the University of Sheffield, Wiz is a one-of-a-kind multi-user platform.
- Wiz plays an important role in big data analysis and machine learning in the industry 4.0 era.

REFERENCE

1. Balzer et al., Wiz: A Web-Based Tool for Interactive Visualization of Big Data, Patterns (2020), https://doi.org/ 10.1016/j.patter.2020.100107
2. I-Cheng Yeh, (1998) "Modeling of strength of high-performance concrete using artificial neural networks", Cement and Concrete Research, 28, 1797-1808.

ACKNOWLEDGE