Heritage By The Numbers: Exploring The Harvard Art Museum Collections



This project introduces us to big data analysis by comparing organised visual parsing through the Harvard Art Museum website to automated machine parsing through CSV files on Jupyter Notebooks.

Process:

- Manually compared HAM collections via user-friendly website to an overwhelming CSV file.
- Leveraged Jupyter Notebooks on PositCloud to analyse extensive cultural data through the production of thematic word clouds.

Findings:

- Information availability, tagging, and categorisation is crucial to successful data analysis.
- Word clouds highlight thematic focuses for each culture giving insight to an institution's acquisition interests.





