DORIS × The Data Mine

Jacob Alge, Dan Asrul, Oliver Bonilla, Yi-Chen Chiou, Priya Dave, Abolee Diwate, Oluwaseyi Jaiyeoba, Allison Kramp, Akshaya Kumar, Jerry Liu, Joaquin McCreary
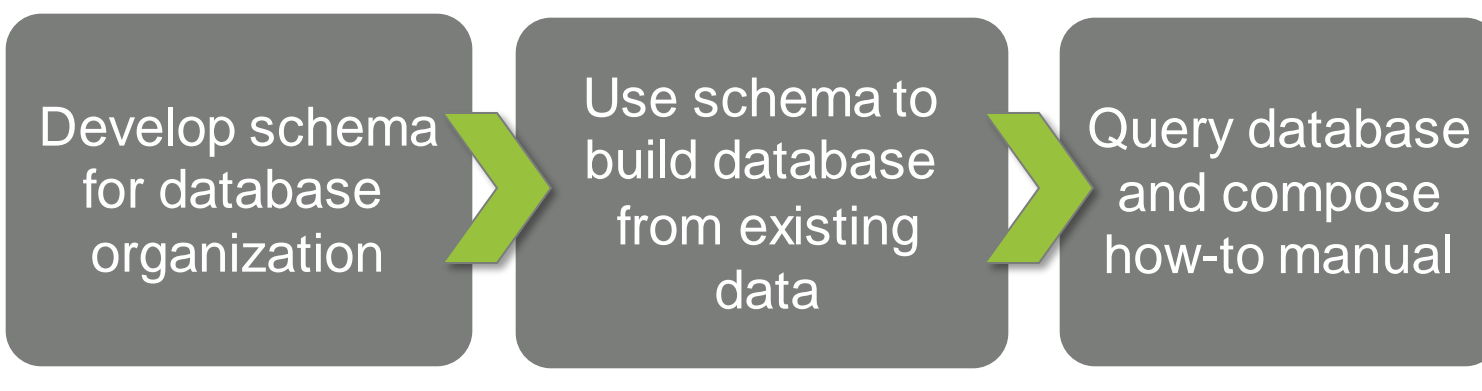
Introduction

DORIS has worked with client organizations since 2012 researching their people and workspaces. This research has resulted in over 4000+ hours of usability studies and 5000+ interviews.[1]

Our team streamlined DORIS' workflow by:

- building a customized master database
- analyzing the database for historic trends
- creating a dashboard illustrating key findings from client data
- enhancing visualizations from real use cases

Database Objectives

Develop schema for database organization → Use schema to build database from existing data → Query database and compose how-to manual

Data Visualization Objectives

Critique DORIS deliverables → Iterate on visualizations from real use cases → Produce analytics dashboard from database

Research Methodologies

Database

- Opted to use DBDiagram to generate expression tree
- Utilized Python and SQL to build database

Data Visualization

- Worked with GGPlot, Tableau, Infogram, and Beautiful.ai to design relevant plots
- Used Shiny in R to create interactive dashboard with data from master database [2]

Database Overview

Expression tree/entity-relation diagram which was used for the creation of a standardized schema that was applied to the database code.
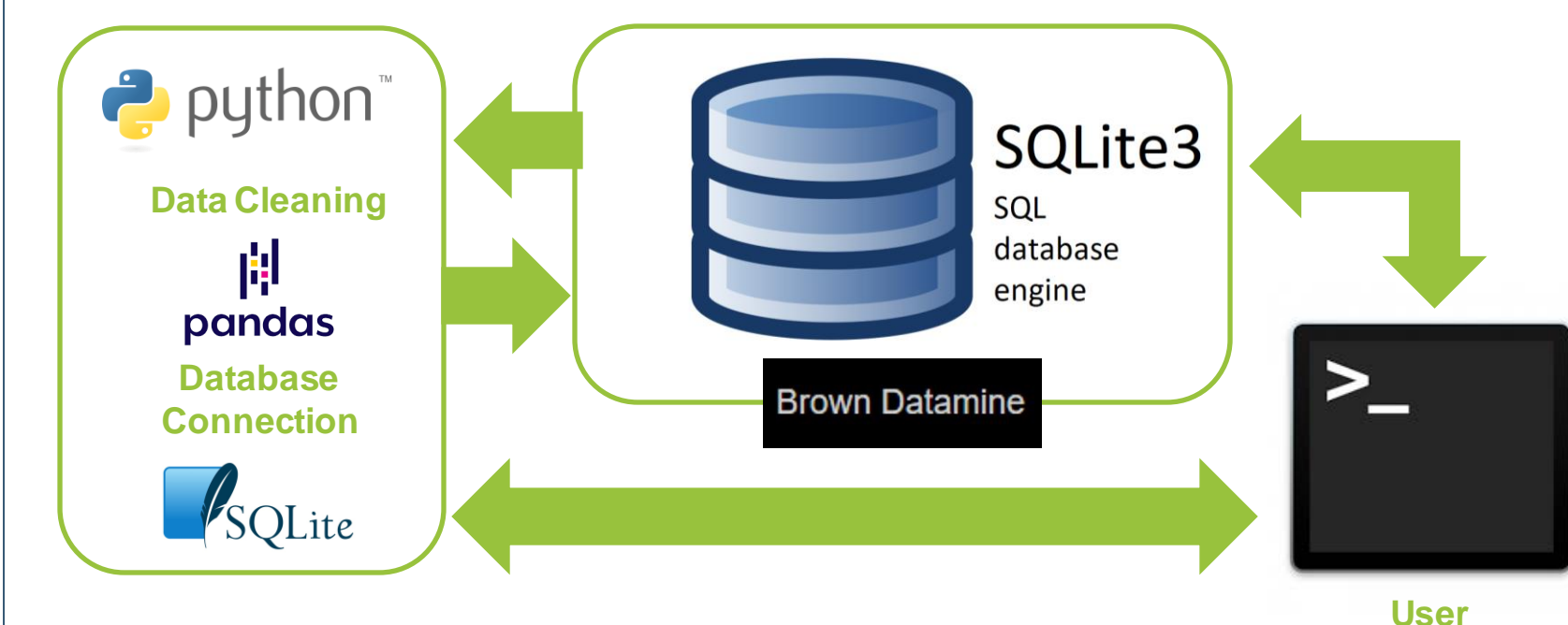
To operate the database, users will enter SQL queries through command line. SQL will then interact with the master database and return results.
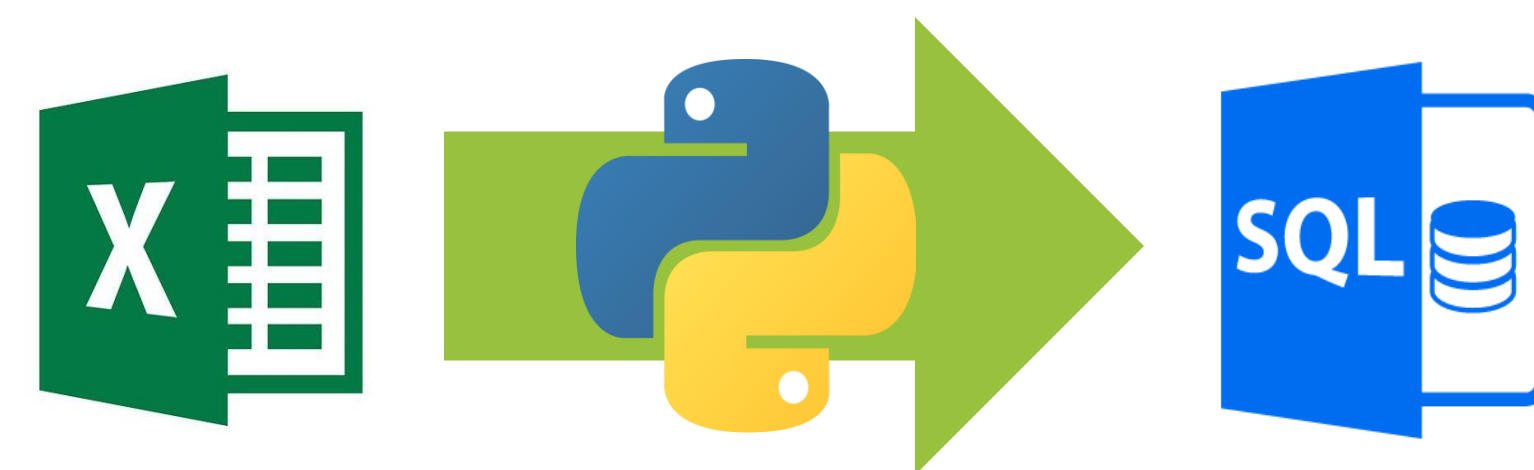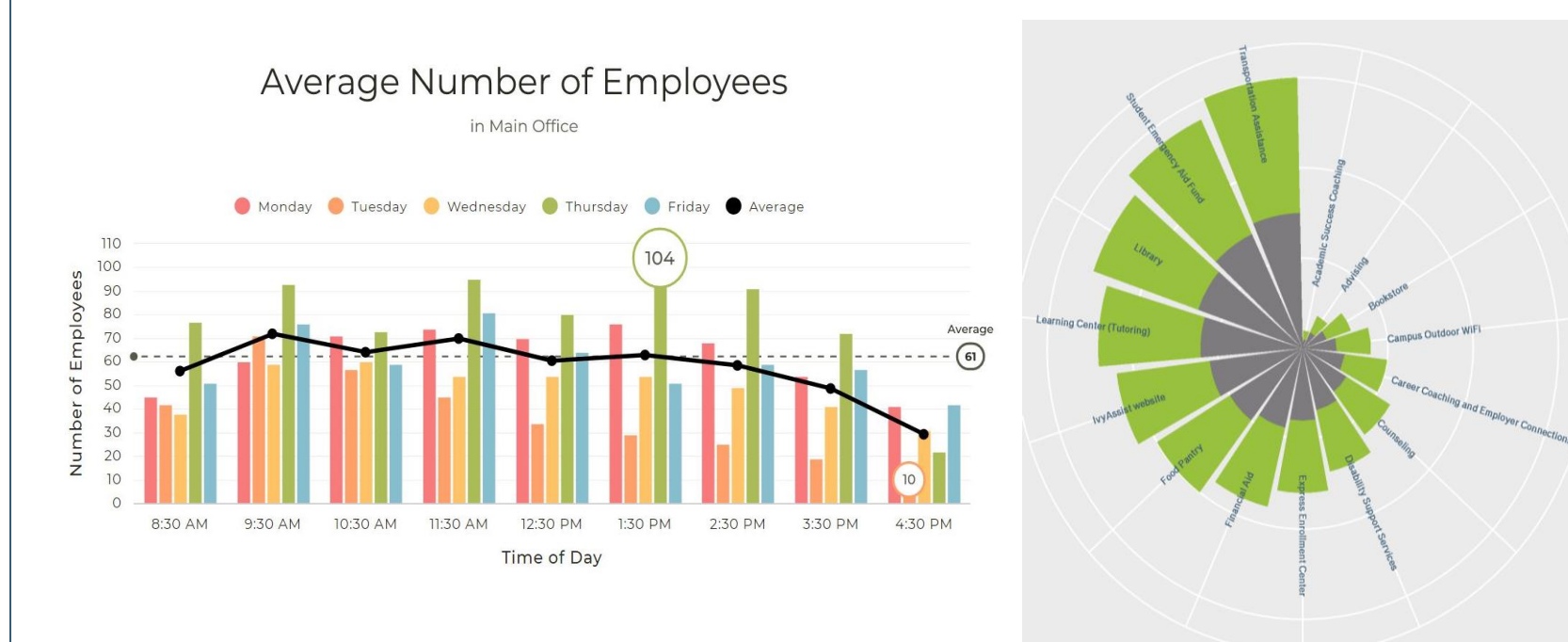

Used Python to read input data from DORIS data spreadsheets and generate master database according to schema in SQL.

Total Number of Datapoints: 427,325
Number of Companies Sampled: 18+

Data Visualization Overview

Real Use Cases

- Created new visualizations for real DORIS clients in order to analyze overall trends.
- Iterated on past DORIS deliverables to build a standard data visualization platform/template.

Top Left: Workweek employee attendance pattern for tax auditing firm
Top Right: College data highlighting virtual/in-person distribution.
Bottom: Identifying trends in workspace activity for tax auditing firm.

Dashboard

- Queried data from the SQL database and created dashboard with graphics in key areas of interest.
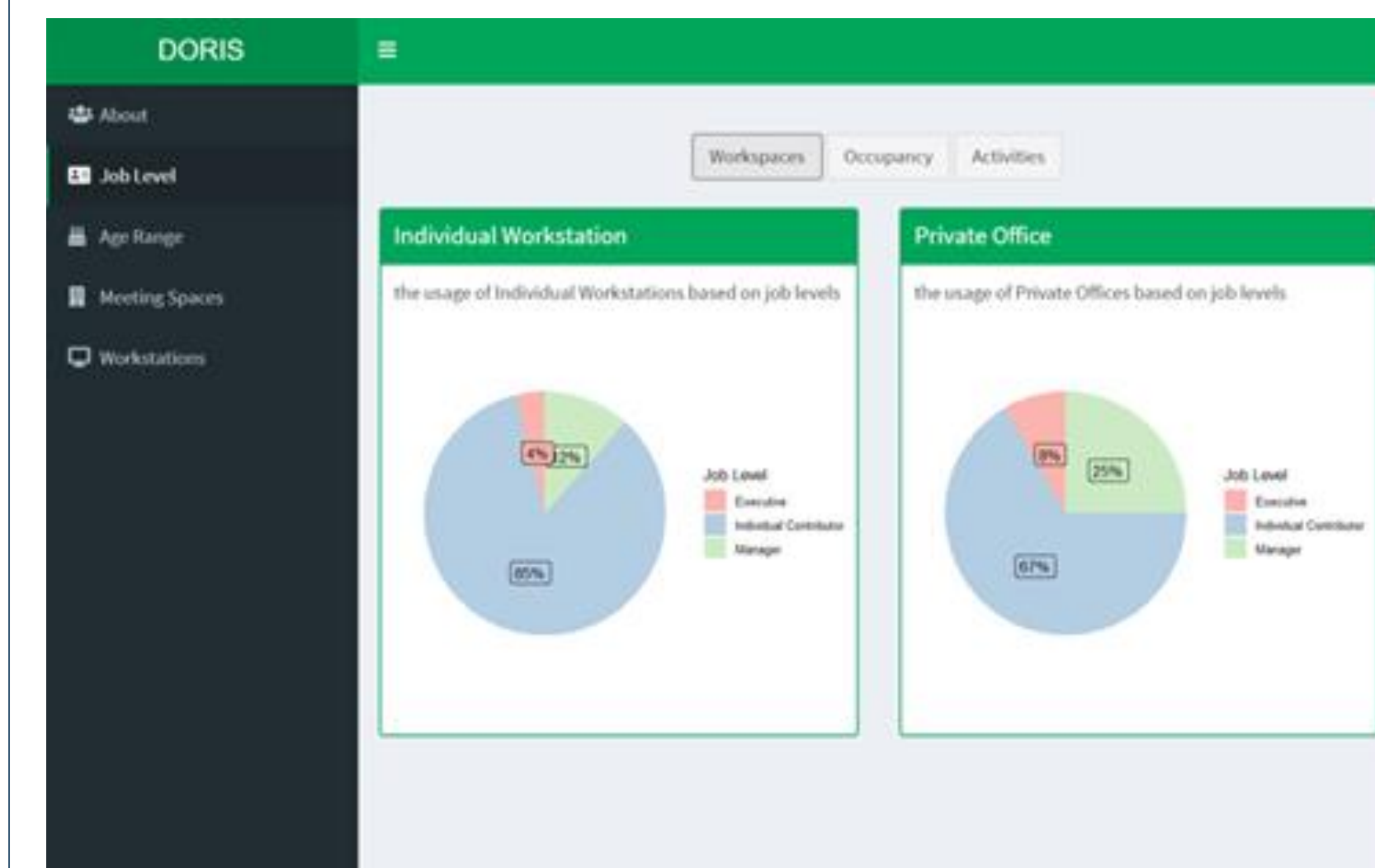
Conclusions

Database

- Once a generalized code was designed, it made the process of inserting every company into the database go smoothly.
- Not every company's data could be formatted to fit the standardized code, therefore flexibility needs to be made to be able to accommodate unforeseen changes in the data.

Data Visualization

- Visualizing data is extremely useful and emphasizes interesting trends that were otherwise hidden.
- The graphics chosen for data visualization have a big impact on the way data is interpreted, collected, and improved.

Next Steps

- Design a new, custom database schema for DORIS to collect data and conduct research on university organizations
- Reform DORIS method of adding new company data to existing master database

References

[1] "About DORIS." *DORIS*, 12 Jan. 2022, https://dorisresearch.com/about-doris-research/.

[2] "App Architecture." *Shiny*, RStudio, PBC, Jan. 2020, https://shiny.rstudio.com/.

Acknowledgements