



Migrating agricultural data management to a cloud-hosted relational database

MSU Cloud Computing Fellowship Symposium

May 5th, 2023

Sarah Manski

Department of Statistics and Probability



About the fellow

My background

- PhD Candidate in the Department of Statistics and Probability, expected graduation August, 2023
- Dissertation research areas include Bayesian statistics applied to agriculture and statistics education
- Bachelor's degree in mathematics and computer science and Master's degree in mathematics
- 2022 - 2023 Cloud Computing Fellow



Project motivation

- Large-scale Bayesian modeling is highly computationally intensive
- My agricultural research currently encompasses over a million fields over 5 states and 17 years and is continually expanding
- Modeling and prediction of each county takes approximately 90 minutes HPCC compute time
- Input and output data is stored in thousands of CSVs in Google Drive and is infeasible to access on personal computers

Considerations made using cloud

Automating Bayesian modeling
and prediction in cloud



Building a cloud-hosted
database for our extensive
input and output data

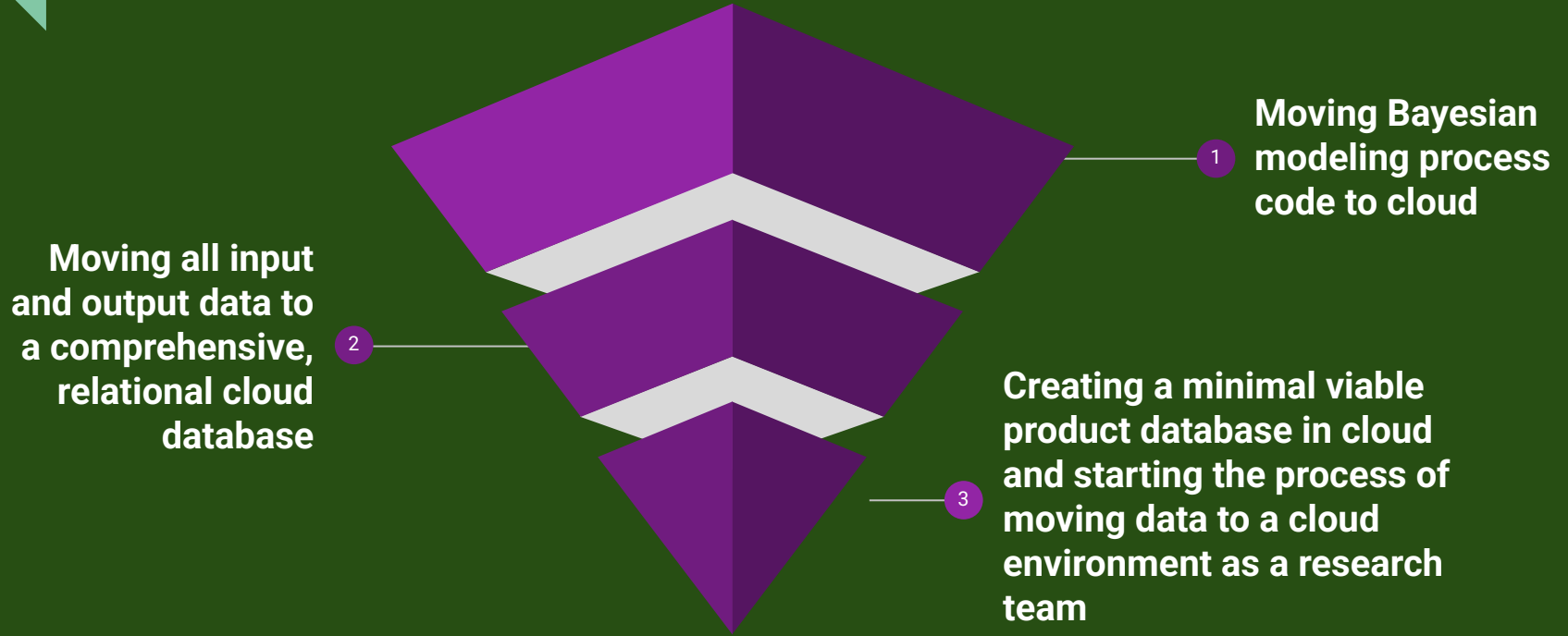
Starting local and
migrating to cloud

Data security

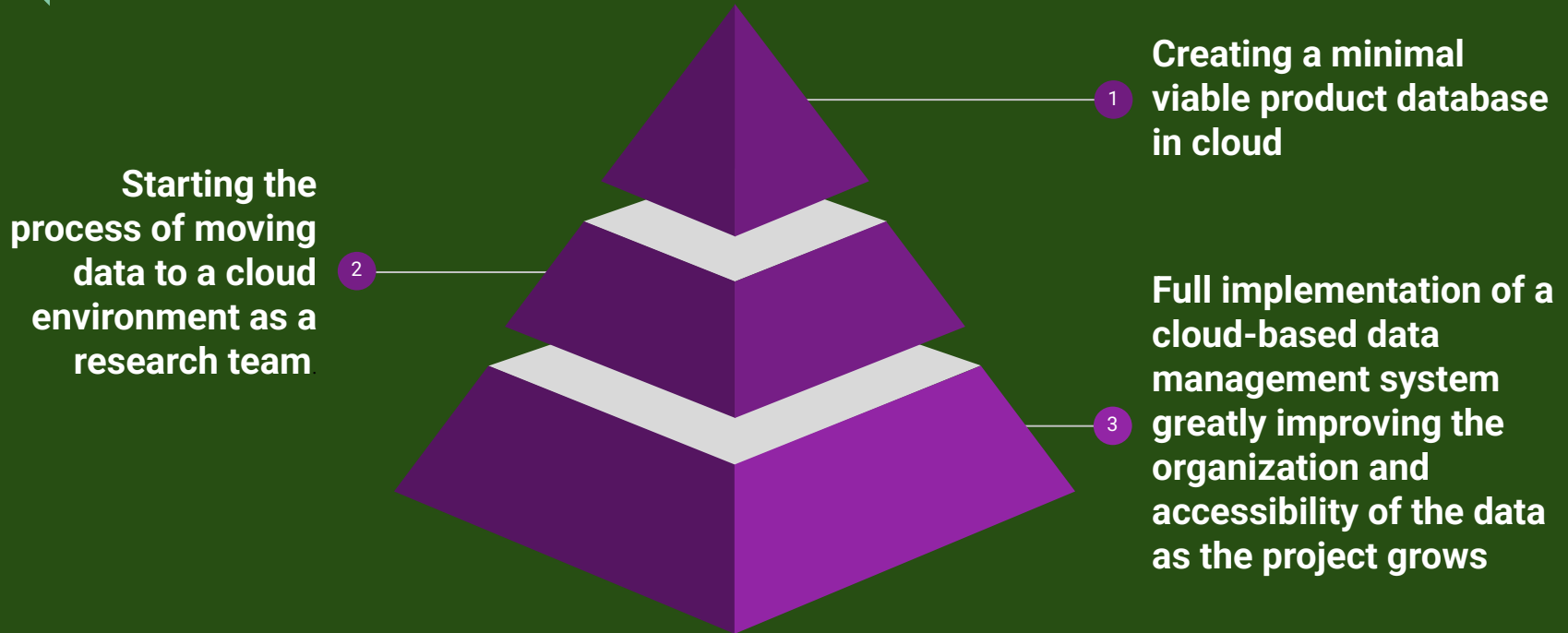
Cloud platform

Type of database
(PostgreSQL flexible)

Project Scope



Project (Future) Impact





Personal Impacts

- Learning about the realities and practical application of cloud computing
- Learning about the Azure portal, PostgreSQL, RPostgres, psql, and pgAdmin
- Networking with the CCF mentors and fellows and a graduate student with database experience
- Expanding my data management toolkit for future work



Thanks!

- To the CCF fellows and mentors, especially Pat and Mahmoud
- To Joe and my research team
- To you for attending the CCF Symposium

Keep in touch!

Before August, 2023:

Sarah Manski

Email: manskisa@msu.edu

Graduate Assistant

Department of Statistics and
Probability, Wells Hall

After August, 2023:

Dr. Sarah Manski

Email: manskisa@msu.edu

Research Associate

Center for Statistical Training and
Consulting (CSTAT), Giltner Hall