CIF21 DIBBs: EI: Vizier, Streamlined Data Curation

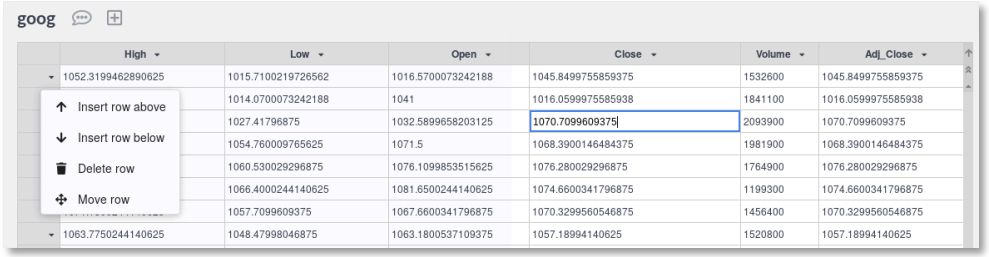
PI: Oliver Kennedy[1], **Co-Pis:** Juliana Freire[2], Boris Glavic[3]
Institutions: [1] University at Buffalo, [2] New York University, [3] Illinois Institute of Technology

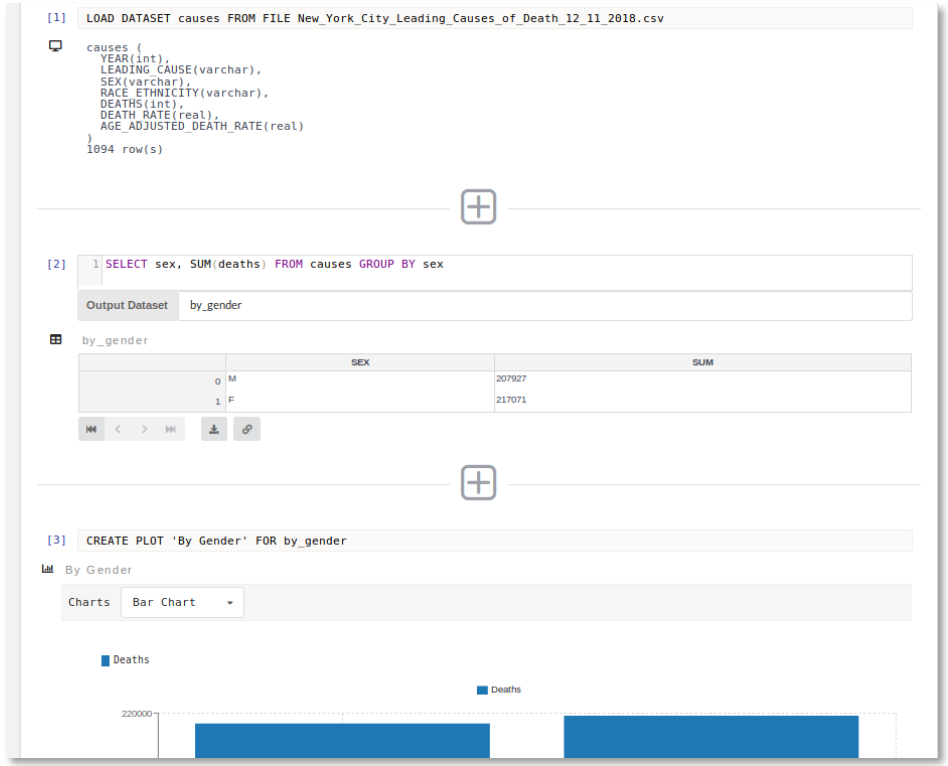
Award #: 1640864

- Scalable open-source platform for data debugging and exploration

  https://github.com/VizierDB

- Combines notebooks and spreadsheets for ease of use
- Captures provenance both for data and computations
  - Reproducibility, re-use, and transparency
  - Automated versioning
  - Tracking of potential data errors
- Vizier(Dirty Data) = Better Data + Provenance
  - Publish data with provenance and known caveats





https://vizierdb.info