

La Serena school for data science: Applied tools for data-driven sciences

V. Mesa^{1,2,3}, G. Damke⁴, A. Bayo⁵, M. J. Graham⁶, D. Norman⁷, M. Cerda⁸, F. Förster⁹ and C. Ibarlucea⁴

¹Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile.

²Association of Universities for Research in Astronomy (AURA), Avda. Juan Cisternas 1500, La Serena, Chile.

³Grupo de Astrofísica Extragaláctica-IANIGLA, CONICET, Universidad Nacional de Cuyo (UNCuyo), Gobierno de Mendoza, Argentina.

⁴Cerro Tololo Interamerican Observatory, NSF's NOIRLab, Av. J. Cisternas 1500 N, 1720236, La Serena, Chile.

⁵European Southern Observatory, Karl-Schwarzschild Str. 2, 85748 Garching bei Muenchen, Germany.

⁶California Institute of Technology, 1200 East California Boulevard Pasadena, California 91125, USA.

⁷NSF's NOIRLab, 950 North Cherry Avenue Tucson Arizona 85719 USA.

⁸Instituto de Ciencias Biomédicas y Centro de Informática Médica y Telemedicina, Facultad de Medicina, Universidad de Chile.

⁹Data and Artificial Intelligence Initiative (D&IA), Center for Mathematical Modeling (CMM), U. Chile.

Keywords: *astronomical data bases: surveys, methods: data analysis, methods: statistical, software: data analysis, sociology of astronomy*

Corresponding author: valeria.mesa@userena.cl

The volume and complexity of astronomical data continues to grow as the current generation of surveys come online (PanSTARRS, Gaia, Dark Energy Survey, VISTA). Astronomers will need to work with giga, tera and even petabytes of data in real time (LSST). This poses the challenges of developing and using new tools for data discovery, access, and analysis. At the same time, there are new opportunities for interdisciplinary research in applied mathematics, statistics, machine learning, crowd-sourcing, etc. Astronomy provides a sand-box where scientists can come together from diverse fields to address common challenges within the “Big Data” paradigm.

In response, since 2013 the AURA-O/NOIRLab together with other Chilean organizations such as the ULS and Universidad de Chile, have implemented the “La Serena School for Data Science: Tools for Data-Driven Sciences”. In this contribution we present our initiative, which is developed annually at the AURA campus in La Serena, Chile. Between July and August, the school receives 32 students and 10 expert professors for 10 days, following an intensive program of introductory classes with an important component of hands-on learning. This is materialized mainly by the development of group projects, where students solve complex problems in groups, applying what they have learned and guided by the teachers. Participants come from various countries such as the US (NSF-funded), Chile, and other Latin American countries such as Ecuador, Brazil, Mexico or Colombia. Students from different disciplines are encouraged to participate, including (but not limited to) astronomy,

physics, computer science, statistics, mathematics, biology medical informatics.

Since LSSDS began 10 years ago, we have trained more than 230 students in data science. As seen in the statistics, efforts have been made to promote diversity in all possible areas. Also, the alumni of the school has been particularly active not only “returning” as Teaching Assistants, but also replicating parts of their schools at their home institutes.

The interest of the scientific community is increasing, with a record of applicants in the last year, mainly from Latin American countries.

We are very pleased to announce that the NSF has just approved the budget to fund LSSDS through 2024. ¹ We invite all Latin American institutions to support this initiative!

■ Acknowledgements

Thanks to our funding from NSF (Award Number (FAIN): 1637359) and ANID (Gemini/ANID program, Chile N° ASTRO21-0067), along with other international partners, applications from students in US, Chilean, Ecuadorian (members of CEDIA) and Latin American (members of RedCLARA) institutions are eligible for full scholarships that cover all school expenses. We also thank NSF's NOIRLab, AURA Observatory in Chile, Universidad de Chile, Universidad de La Serena, NIC Chile, REUNA, Millennium Institute of Astrophysics (MAS). V.M. also acknowledges support from project DIDULS Regular N° PR2353857.

¹Additional information and applications: http://www.aura-o.aura-astronomy.org/winter_school/