Fernando Becerra

Astrophysicist · Data Scientist

Cerro del Paso Norte 831, San Bernardo, Santiago, Chile 8080982

+56 9 3689 9526 | becerrafernando@gmail.com | www.fernandobecerra.com | fbecerra | becerrafernando

Education

Harvard University — *Cambridge, MA*

Ph.D. in Astronomy & Astrophysics — *May 2018*

Harvard University — *Cambridge, MA*

A.M. in Astronomy & Astrophysics — *May 2014*

Universidad de Chile — *Santiago, Chile*

M.Sc. in Astronomy, with Highest Honors — *Aug 2012*

Universidad de Chile — *Santiago, Chile*

B.Sc. in Astronomy, with Highest Honors — *Dec 2009*

Work Experience

Freelance Developer — *Santiago, Chile*

Data Science and Data Visualization — *May 2020 to date*

- *Urban Institute*:
  - Created interactive visualization to show access to experienced teachers, Advanced Placement classes, and school counselors for students from different racial and ethnic backgrounds.
  - Used d3.js to visualize donations to charitable organizations and beyond, going from tax-exempt non-profits to include crowd-funding, impact investing, and political contributions.
  - Created interactive map to display gap between AP class enrollment and AP test taking for each racial or ethnic group and each district in Florida.
- *Epic Institute*:
  - Used Pandas library to analyze and process energy outputs dataset from different sources such as IEA.
  - Used NumPy and SciPy to optimize calculations for emissions model and Natural Climate Solutions adoption rates for the Positive Disruption 2022 (PD22) report.
  - Used d3.js to create an interactive data explorer to display compiled dataset about energy demand and supply, emissions, GHG concentration and temperature, and model outputs from PD22.
- *Planet Labs*:
  - Used Planet's Explorer to find and download satellite imagery to fulfill client's and internal requests.
  - Color-correct satellite imagery using Adobe Photoshop, Python, and GDAL.
  - Create publication-ready images for clients and internal use.
- *Golden Set Analytics*:
  - Used Pandas library to analyze and process a tennis matches database with more than 900,000 rows and 1,000 columns.
  - Created and documented a Python module based on NumPy, SciPy, Matplotlib and seaborn that calculates players ratings, computes accuracy of models, evaluates performance of processes, and creates plots to represent the results.
  - Used Machine Learning algorithms to run a hyperparameter optimization of models and evaluate their outcomes.
  - Developed reports and presentations to communicate my findings to the rest of the team.

- *Research Rabbit*:
  - Used d3.js implementation of a force-directed graph to make an interactive visualization of collaboration networks in Academia.
  - Represented authors and papers using nodes and labels that allows interactions such as clicking and hovering to get more detailed information about one item.
  - Showed collaboration between authors or citation metrics between papers using links between nodes.
  - Developed two views of the paper visualization: network and timeline, in which the latter orders the papers by date of publication.
- *Pontifical Catholic University of Chile*:
  - Processed and analyzed a cohort database that follows a group of Chilean people from their birth date until their 18th birthday.
  - Used Python libraries such as Pandas, NumPy and SciPy to calculate averages and standard deviation of variables throughout time for several subgroups (male/female, control/desease).
  - Calculated p-values and odds ratio and determine the risk of developing Non-Alcoholic Fatty Liver Desease and Non-Alcoholic Fatty Pancreas Desease based on fat and fat-free mass for each subject using SciPy and statsmodels modules.
  - Created Hattori plots using matplotlib to show the trajectory of fat and fat-free mass as a function of time for control group and group presenting the desease.
- *LA County Public Health Department:*
  - Replaced static graphics with interactive web visualizations that update itself once the dataset is updated.
  - Used d3.js to create interactive plots that show COVID-19 statistics such as testing numbers and mortality rates for Los Angeles county.
  - Added interactive tooltip that shows detailed information on demand.
- *Emteq Labs*:
  - Used d3.js to create an interactive plot that shows timeseries of measurements of user responses to immersive experiences in real time.
  - Used javascript to get data from API and update the plot parameters in real time.
  - Used HTML Canvas to optimize the performance of the plot by decreasing CPU requirements on the user end.
- *Copenhagen Atomics*:
  - Used d3.js to create an interactive line plot to show temperature from different sensors from a nuclear reactor in real time.
  - Updated time range shown in x-axis of the plot and time range selection tool based on data fed by the API.
  - Added option to save and load current view including zoom level, time range, and variables shown.
- *Needle Genomics*:
  - Created interactive visualization to explore single cell RNA-seq data by plotting their t-SNE coordinates.
  - Used javascript to get data from the API and d3.js to create the visualization.
  - Used jQuery to create menus to select properties to be shown in the visualization such as type of genes, coloring options, and coordinates to plot.

Fathom Information Design *Boston, MA, USA*

Data Visualization Developer *Jun 2018 - Jun 2019*

- Coded back end and designed front end prototype for *Laniakea* app (http://laniakea.fathom.info)
- Used Python packages such as spaCy and nltk to perform Natural Language Processing techniques on large document sets.
- Implemented topic modeling to group and classify more than 100,000 documents using LDA, NMF, and t-SNE.
- Optimized routines for fast processing with NumPy, SciPy, and multiprocessing, achieving a 100x speed increase.
- Coded back end and designed front end prototype for *Myriscope* app (http://myriscope.com).
- Used Machine Learning libraries to extract and consolidate abstract, sections, and figures from academic papers.
- Created prototype for front end employing Javascript, jQuery, CSS and HTML.
- Coded back end and front end for *The Joy of Parsing* (https://fathom.info/bobross/).
- Scrapped all 403 transcripts from the show *The Joy of Painting* using the YouTube API and packages such as beautifulsoup.
- Analyzed, grouped, and classified the transcripts using NLP techniques and Python packages like spaCy and nltk.
- Created interactive tool to explore paintings of the show using d3.js.

Harvard University, Department of Astronomy *Cambridge, MA, USA*

Graduate Research Assistant *Aug 2012 - May 2018*

- Explored the formation of stars and black holes in the early Universe.
- Lead, guided, directed, and managed group of collaborators to design and execute a research plan.
- Implemented new modules for primordial chemistry and sink particles in C for the *arepo* code to model behavior of black holes.
- Developed tools to generate plots, images, and videos of simulation outputs: the Python analysis tool *pacha* using packages like NumPy, SciPy, and matplotlib; and the parallel C analysis tool *sator* using MPI.
- Reported findings in astronomy journals like *Monthly Notices of the Royal Astronomical Society* and *The Astrophysical Journal*.
- Presented results in astronomy conferences across many continents.
- Mentored and supervised undergrad and graduate students.

Universidad de Chile, Department of Astronomy *Santiago, Chile*

Graduate Research Assistant *Mar 2010 - Aug 2012*

- Conducted independent research on the relation between star formation and properties of the host galaxy.
- Modified old modules and added new ones in C and fortran to the code *Enzo*.
- Developed the Python analysis package *pigs* based on the *yt* code to analyze simulation outputs.
- Coded analysis routines in IDL to examine simulation outputs from the code *Gadget*.
- Presented results in paper published in *The Astrophysical Journal*.

Skills

Programming	Python, C, fortran, IDL, MATLAB, Javascript, Java, LaTeX
Web	HTML5, CSS, jQuery, D3.js, Three.js, Processing, React
Software	Adobe Photoshop, Adobe Illustrator, Microsoft Office Suite
Languages	English, Spanish
Other	Landscape and Nature Photography