

# Fernando Becerra

DATA ANALYSIS + VISUALIZATION

+56 9 3689 9526 | [✉ becerrafernando@gmail.com](mailto:becerrafernando@gmail.com) | [🏠 www.fernandobecerra.com](http://www.fernandobecerra.com) | [🐙 fbecerra](https://github.com/fbecerra) | [🌐 becerrafernando](https://www.linkedin.com/in/becerrafernando)

## Work Experience

---

### The New York Times

DATA ENGINEER, CONTRACTOR

*Aug 2023 to date*

- Run SQL queries in a 6TB database.
- Analyzed and visualized query results with Python and d3.js

### The Baltimore Banner

DATA VISUALIZATION DEVELOPER, CONTRACTOR

*Sep 2023 to date*

- Created an interactive dashboard using d3.js to show gun deaths in the Baltimore area.

### Epic Institute

DATA SCIENTIST

*Nov 2021 to date*

- 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.

### Urban Institute

DATA VISUALIZATION DEVELOPER, CONTRACTOR

*Oct 2021 to Sep 2023*

- 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.
- Developed a d3 interactive tool to explore how additional metrics in a high-quality accountability system for colleges affect how these institutions perform.
- Plotted quantities such as payments, repayment timelines, and debt balances for different scenarios in an income-driven repayment plan designs for student loan debt.

### USAFacts

DATA VISUALIZATION ENGINEER, CONTRACTOR

*Apr 2022 - Jun 2023*

- Created interactive weighted Voronoi treemap to show how inflation affects American of different ages.
- Compiled and processed air pollution, wages, housing, and school enrollment data for 2022 midterm elections tool.
- Helped implementing style and rolling average for work absences due to childcare reasons.
- Compiled and processed home prices and wages data at the County level for an article comparing how both of them are rising at different rates.
- Processed data from ACS regarding housing costs for renters and owners and created visualizations to explore the results.

### Planet Labs

DATA VISUALIZATION DEVELOPER, CONTRACTOR

*Sep 2021 - Jun 2022*

- 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.

## ProPublica

DATA VISUALIZATION DEVELOPER, CONTRACTOR

Mar 2022 - Jun 2022

- Designed and developed a story layout to compare two soldiers' parallel roads to more than 100 days of pre-trial confinement.

## Swayable

DATA VISUALIZATION DEVELOPER, CONTRACTOR

Feb 2022 - May 2022

- Processed and manipulated a dataset of more than a million survey responses to create quantities that allowed them to be positioned in the x-y-z plane using Jupyter notebooks.
- Used three.js to create a prototype of an interactive 3d cube that shows this data and reveals the relationship between all data points.

## Pontifical Catholic University of Chile

DATA SCIENTIST, CONTRACTOR

Sep 2020 - Oct 2021

- 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/disease).
- Calculated p-values and odds ratio and determine the risk of developing Non-Alcoholic Fatty Liver Disease and Non-Alcoholic Fatty Pancreas Disease 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 disease.

## Research Rabbit

DATA VISUALIZATION DEVELOPER, CONTRACTOR

Feb 2021 - Aug 2021

- 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.

## Golden Set Analytics

DATA SCIENTIST, CONTRACTOR

Jun 2020 - Jun 2021

- 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.

## Emteq Labs

DATA VISUALIZATION DEVELOPER, CONTRACTOR

Mar 2021 - May 2021

- 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 Atomic

DATA VISUALIZATION DEVELOPER, CONTRACTOR

Sep 2020 - May 2021

- 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.

## LA County's Department of Public Health

DATA VISUALIZATION CONSULTANT

May 2020 - Oct 2020

- 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 LA county.
- Added interactive tooltip that shows detailed information on demand.

## Fathom Information Design

DATA VISUALIZATION DEVELOPER

Boston, MA, USA

Jun 2018 - Jun 2019

- 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.
- Used Machine Learning libraries to extract and consolidate abstract, sections, and figures from academic papers.
- Scrapped all 403 transcripts from the show *The Joy of Painting* using the YouTube API and packages such as beautifulsoup.
- Developed front-end of the *The Joy of Parsing* app, back-end of *Myriscope* app, and back-end of *Laniakea* app

## Harvard University, Department of Astronomy

GRADUATE RESEARCH ASSISTANT

Cambridge, MA, USA

Aug 2012 - May 2018

- Lead group of collaborators to study the formation of stars and black holes in the early Universe.
- 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.
- Published findings in astronomy journals like *Monthly Notices of the Royal Astronomical Society* and *The Astrophysical Journal*.

## Teaching positions

---

### Universidad del Desarrollo

INFORMATION VISUALIZATION LECTURER, MASTER OF DATA SCIENCE

Jul 2022 to Oct 2023

### Columbia University, Graduate School of Journalism

TEACHING ASSISTANT, LEDE PROGRAM

Jun 2023 - Aug 2023

### Pontificia Universidad Católica de Chile

INTERACTIVE DATA VISUALIZATION WITH D3.JS GUEST LECTURER, DATA VISUALIZATION POSTGRADUATE DIPLOMA

Jun 2022

### Harvard University

TEACHING ASSISTANT, SPU19: THE ENERGETIC UNIVERSE AND AY17: GALACTIC AND EXTRAGALACTIC ASTRONOMY

Jan 2013 - Dec 2015

## Education

---

### Harvard University

PH.D. IN ASTRONOMY & ASTROPHYSICS

Cambridge, MA

May 2018

### Harvard University

A.M. IN ASTRONOMY & ASTROPHYSICS

Cambridge, MA

May 2014

### Universidad de Chile

M.Sc. IN ASTRONOMY

Santiago, Chile

Aug 2012

### Universidad de Chile

B.Sc. IN ASTRONOMY

Santiago, Chile

Dec 2009