Federico Castro, Ph.D.
Data Analyst

Mexico City, Mexico

fcastro.biol@gmail.com | GitHub Page

ES | EN | DE | FR | JA

Summary

PhD-trained scientist specialized in data analysis and statistical modeling. Proficient in R programming and data visualization, with a strong record of interdisciplinary collaboration on complex data-driven projects. Author of peer-reviewed publications with experience across Europe and Asia. Comfortable working across a wide range of cultural and professional settings. Eager to apply quantitative expertise to advance evidence-based decision-making in industry.

Education

Ph.D. in Biological Sciences
2014–2022 / Germany

Freie Universität Berlin

M.Sc. in Evolution, Ecology, and Systematics
2011–2013 / Germany

Ludwig-Maximilians-Universität München

B.Sc. in Biology
2005–2011 / Mexico

Universidad Nacional Autónoma de México

Recent Work Experience

Postdoctoral Researcher
2022–2024 / Mexico

Biology Institute, UNAM

- Conducted epidemiological and data-driven research using large-scale biological datasets.
- Applied advanced statistical modeling and predictive analytics in R to identify trends and correlations.
- Published findings in high-impact international journals.
- Collaborated with interdisciplinary teams across Europe and Latin America.

Lecturer
2022–2024 / Mexico

Postgraduate Program in Biomedical Studies, UNAM

- Taught graduate-level courses in statistical analysis and data visualization using R.
- Supervised projects and mentored students on how to perform data wrangling and hypothesis testing.
- Oversaw data analysis projects, defining goals, deliverables, and review checkpoints.
- Developed course materials that integrated theory with applied case studies and real-world datasets.

Skills

Core Analytical Skills	EDA	Modeling	Hypothesis Testing	Trend Analysis	Data Visualization
Programming & Tools	R	Bash	Excel	SQL	Git
Data Management	Data Cleaning	ETL	Database Querying	Reproducible Research	Dashboard Design
Project Management	Scrum	Kanban
Languages	*Native:* Spanish	*Academic:* English	*Conversational:* German, Japanese

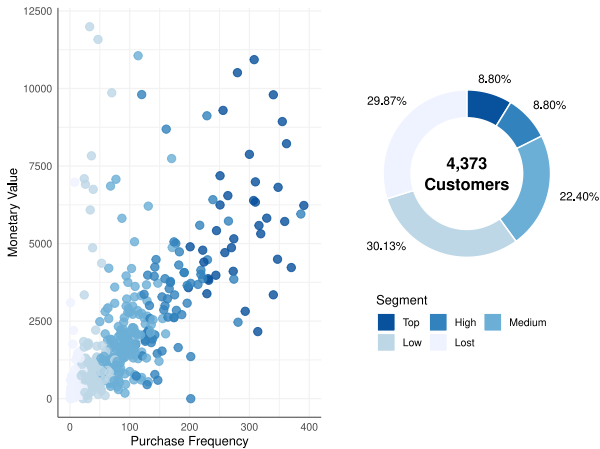## Selected Publications

- Tracking *Batrachochytrium dendrobatidis* infection across the globe. Link
- *Batrachochytrium salamandrivorans'* amphibian host species and invasion range. Link
- Crypticity in Biological Invasions. Link

Federico Castro, PhD
Data Scientist

Mexico City, Mexico

fcastro.biol@gmail.com

```
SELECT
  DATE(InvoiceDate) AS sale_date,
  ROUND(SUM(Quantity * UnitPrice), 2) AS daily_sales
FROM sales_data
WHERE InvoiceNo NOT LIKE 'C%'
  AND Quantity > 0
  AND UnitPrice > 0
  AND CustomerID IS NOT NULL
  AND InvoiceDate IS NOT NULL
GROUP BY sale_date
ORDER BY sale_date;
```
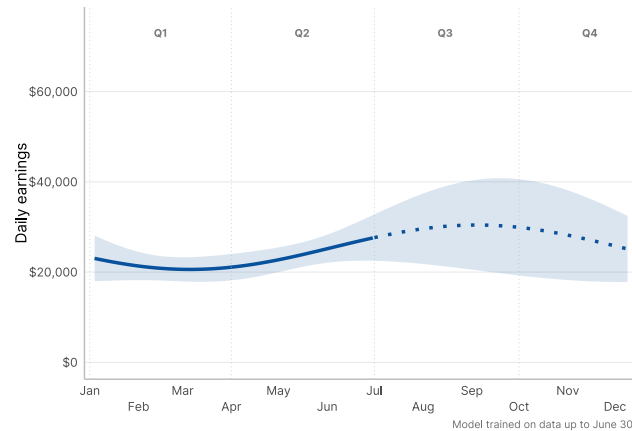
SQL Data Wrangling

This SQL query aggregates valid daily sales while filtering out returns, nulls, and zero values to ensure clean, reliable data for analysis.

Exploratory Data Analysis (EDA)

The ridge plot reveals monthly sales patterns, showing cyclical behavior and highlighting months with unusual spikes or dips.
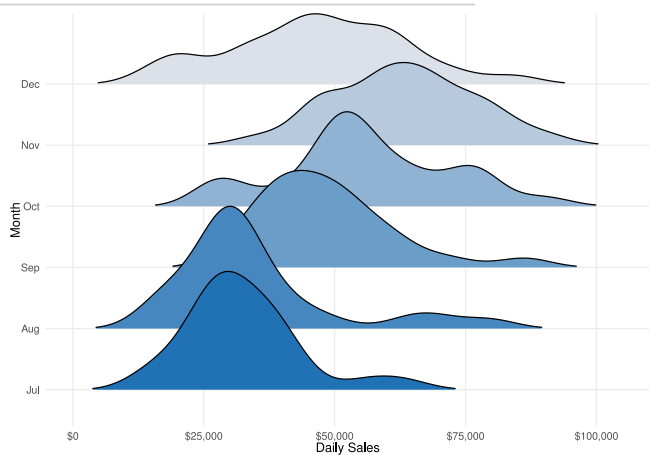

Customer Segmentation (RFM)

Customers are segmented by Recency, Frequency, and Monetary value. The scatterplot shows behavioral clusters, and the donut chart visualizes their proportions.
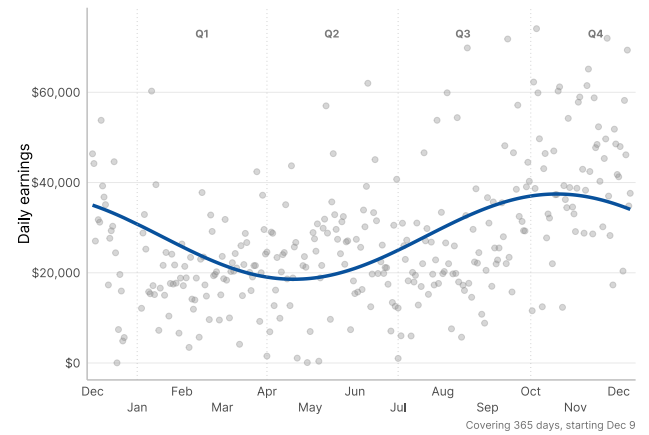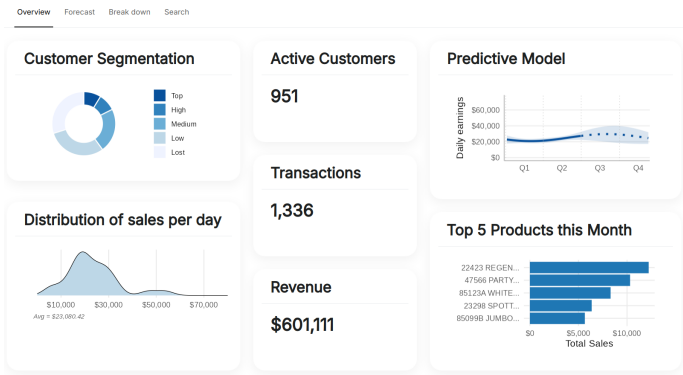

Explanatory Modeling

A regression model captures seasonal fluctuations, showing an upward trend during peak months and confirming recurring demand cycles.

Predictive Modeling & Forecast

A predictive model trained on Q1Q2 data forecasts Q3Q4 sales, providing forward-looking insights for inventory and marketing decisions.

Interactive Dashboard

The interactive dashboard combines multiple analyses, enabling stakeholders to explore trends and uncover actionable insights.