Portfolio Landing - Data Analyst Case

One-line pitch

Data Analyst case built on historical real-world data (2021), transforming an academic project into a professional delivery with pipeline, inferential analysis, and dashboard.

Context and disclaimer

- Source: survey collected in 2021 during the COVID-19 peak.
- Sample: 178 responses (small, non-probabilistic).
- Current use: portfolio case to demonstrate professional methodology and execution.
- Inferential scope: associations within sample, not causal claims or strong population extrapolation.

What this project solves

- Structures and cleans real data with capture noise.
- Standardizes catalogs, anonymizes PII, and builds raw -> interim -> processed layers.
- Produces descriptive + inferential evidence (Chi-square, Spearman, exploratory logistic model).
- Delivers results through an executive notebook and interactive dashboard.

Key results (executed)

- Responses analyzed: 178
- Perceived negative impact: 51.98%
- Respondents reporting pandemic-related problems: 67.23%
- Final wellbeing at risk (No/Tal vez): 41.01%
- Significant associations:
- q4_problems vs q15_wellbeing_final (p=0.000033)
- q4_problems vs anxiety_high (p=0.034516)

Architecture (project phases)

- Phase 1: data/raw/ (immutable source).
- Phase 2: src/data/clean_survey.py -> data/interim/.
- Phase 3: analytical dataset in data/processed/.
- Phase 4: notebooks (notebooks/00_resumen_ejecutivo.ipynb, notebooks/01_eda.ipynb).
- Phase 5: Streamlit dashboard (app/streamlit_app.py).

Quick links

- Executive summary: RESUMEN_EJECUTIVO.md

- Findings report: reports/findings.md

- Executed executive notebook: notebooks/00_resumen_ejecutivo.executed.ipynb

- Executed EDA notebook: notebooks/01_eda.executed.ipynb

- Dashboard app: app/streamlit_app.py

Run locally

```
python3 -m venv .venv
source .venv/bin/activate
pip install -r requirements.txt
python3 src/data/clean_survey.py
jupyter notebook notebooks/00_resumen_ejecutivo.ipynb
streamlit run app/streamlit_app.py
```

Why this matters in a portfolio

This project is more than charts: it demonstrates data judgment, cleaning traceability, defensible analysis, and executive-style communication.