README

STATISTIC TOOL

# 📊 Statistical Tool – Streamlit App

An interactive, modular statistical analysis tool built in \*\*Python\*\* and \*\*Streamlit\*\*, designed to simplify the execution and interpretation of classical and modern statistical tests.

---

## 1. 📘 Introduction

This project presents a fully interactive web-based tool for statistical analysis. It allows students, analysts, and professionals to perform hypothesis testing, ANOVA, regression, correlation, outlier detection, and probability simulations \*\*without needing to write code\*\*.

Each test is implemented as a modular component, and the tool includes example datasets and visual outputs to guide users in learning and decision-making.

---

## 2. 📁 Dataset Description

The app accepts custom `.csv` files and provides a set of \*\*preloaded templates\*\* for quick testing:

- `descriptive\_template.csv`

- `ztest\_template.csv`

- `ttest\_template.csv`

- `proportion\_test\_template.csv`

- `chi\_square\_template.csv`

- `correlation\_template.csv`

- `sampling\_template.csv`

- `linear\_regression\_template.csv`

- `logistic\_regression\_template.csv`

- `encoding\_template.csv`

- ...

These templates ensure that each module receives well-formatted inputs for consistent analysis.

---

## 3. 🛠️ Tools and Technologies

| Component | Description |

|------------------|-----------------------------------------------|

| Language | Python 3.x |

| Web Framework | [Streamlit](https://streamlit.io) |

| Data Handling | `pandas`, `numpy` |

| Stats Engine | `scipy`, `statsmodels` |

| Visualization | `matplotlib`, `seaborn` |

| Export & Support | PDF/Excel export, glossary, statistical tables|

---

## 4. ⚙️ Procedure Overview

### ✔️ How it works:

1. \*\*Upload your dataset\*\* (or use a template).

2. \*\*Choose a statistical module\*\* from the sidebar.

3. \*\*Select variables and parameters\*\* interactively.

4. \*\*Run the analysis\*\* and visualize the results.

5. \*\*Export results\*\* to PDF or Excel (if available).

### 🔎 Each module includes:

- Data validation

- Test execution (Z, T, F, Chi², ANOVA, regression)

- P-value and statistical decision logic

- Confidence intervals, post-hoc tests

- Dynamic plots

- Result interpretation

---

## 5. 🏗️ High-Level Architecture

+-----------------------+

| 🌐 Web Interface |

| (Streamlit Frontend)|

+----------+------------+

|

User selects analysis,

uploads file, sets params

↓

┌─────────────────────────────────────────────────────┐

│ 🎛️ App Core (app.py) │

│ - Renders sidebar and navigation menu │

│ - Routes user actions to appropriate modules │

│ - Manages session state, input, and rendering │

└─────────────────────────────────────────────────────┘

↓

+------------+-------------+

| |

+-----------v------------+ +----------v----------+

| 📁 Modules (Test) | | 📦 Utility Modules |

| /modules/\*.py | | - PDF/Excel Export |

| | | - Reference Tables |

| - Hypothesis Tests | | - Interpretation |

| - ANOVA | | - Dataset Templates |

| - Correlation | +----------------------+

| - Regression |

| - Probability Tools |

| - Encoding / Outliers|

+-----------+-----------+

↓

+------------+------------+

| 📊 Stats Engine |

| (pandas, scipy, statsmodels) |

| - Data processing |

| - Z, T, F, χ² tests |

| - Regression models |

| - Confidence intervals |

+------------+------------+

↓

+------------+------------+

| 📈 Visualization Layer |

| (matplotlib, seaborn) |

| - Boxplots, Histograms |

| - Heatmaps, ANOVA plots|

| - Confidence curves |

+-------------------------+

---

## 6. 📂 Folder Structure

📁 root/

│

├── app.py

├── modules/

│ ├── descriptive.py

│ ├── proportions.py

│ ├── linear\_regression.py

│ ├── ...

│

├── templates/

│ ├── ztest\_template.csv

│ ├── proportion\_test\_template.csv

│ ├── ...

│

├── resources/

│ ├── statistical\_tables.pdf

│ ├── glossary.md

│

└── utils/

├── export\_pdf.py

├── load\_template.py

---

## 7. ✅ Results

Each module generates:

- Test statistics, confidence intervals

- p-values and decision conclusions

- Interpretable charts and tables

- Clear visual markers for hypothesis rejection

- Matching results with manual examples for validation

---

## 8. 📌 Conclusion

This statistical tool integrates essential analytical methods in an intuitive, modular format. Designed for teaching, learning, and practical data exploration, it offers:

- Guided input templates

- Real-time interactive results

- Expandable architecture (via `/modules`)

- Support for classic and inferential statistics

### 🚀 Future Enhancements

- Machine learning module integration

- Unified PDF report generator

- Full bilingual support (English/Spanish)

---

## 🧪 Try it locally

```bash

pip install -r requirements.txt

streamlit run app.py

## Contact

Developed by \*\*Mónica Prieto\*\*