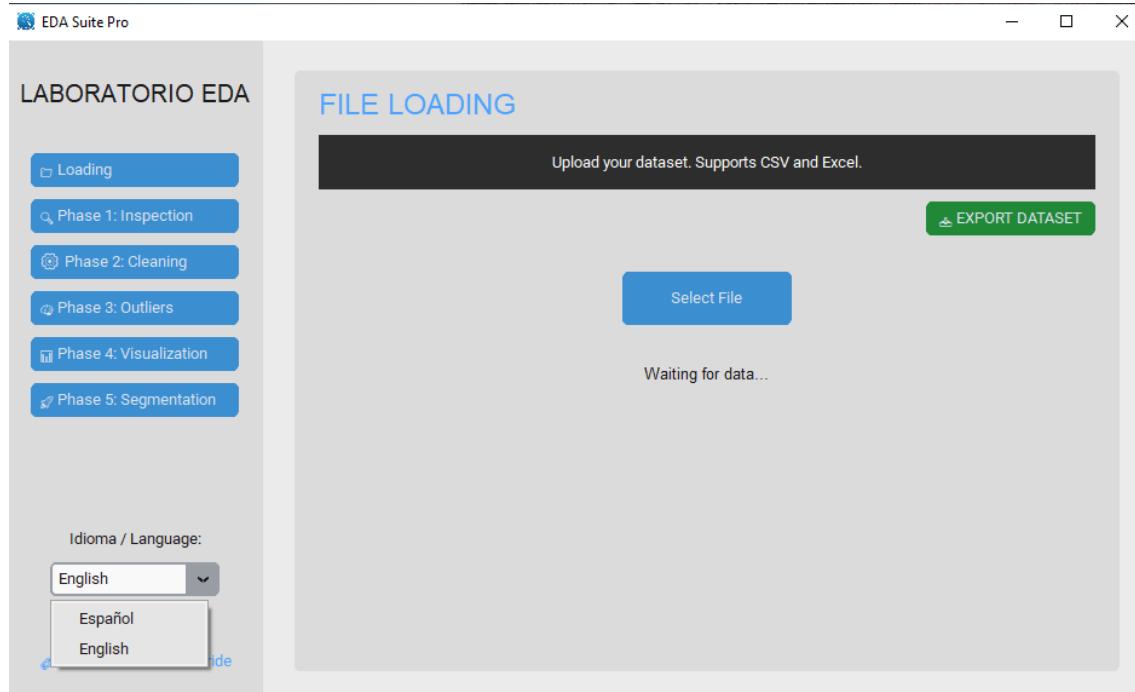


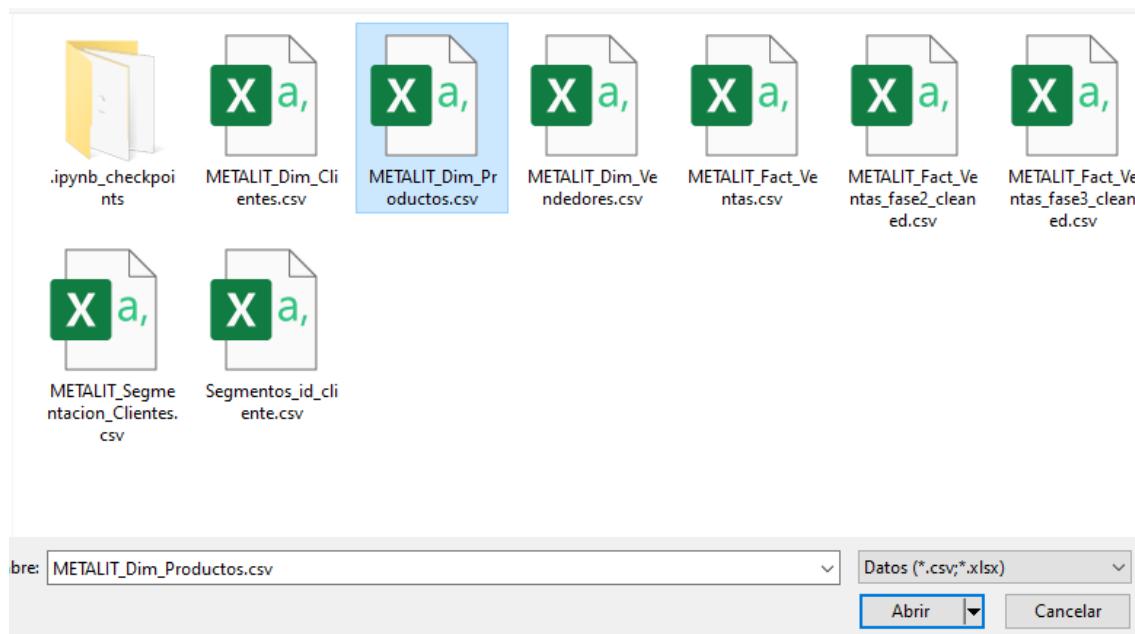
# EDA Suite Pro: Operation Manual and Visual User Guide

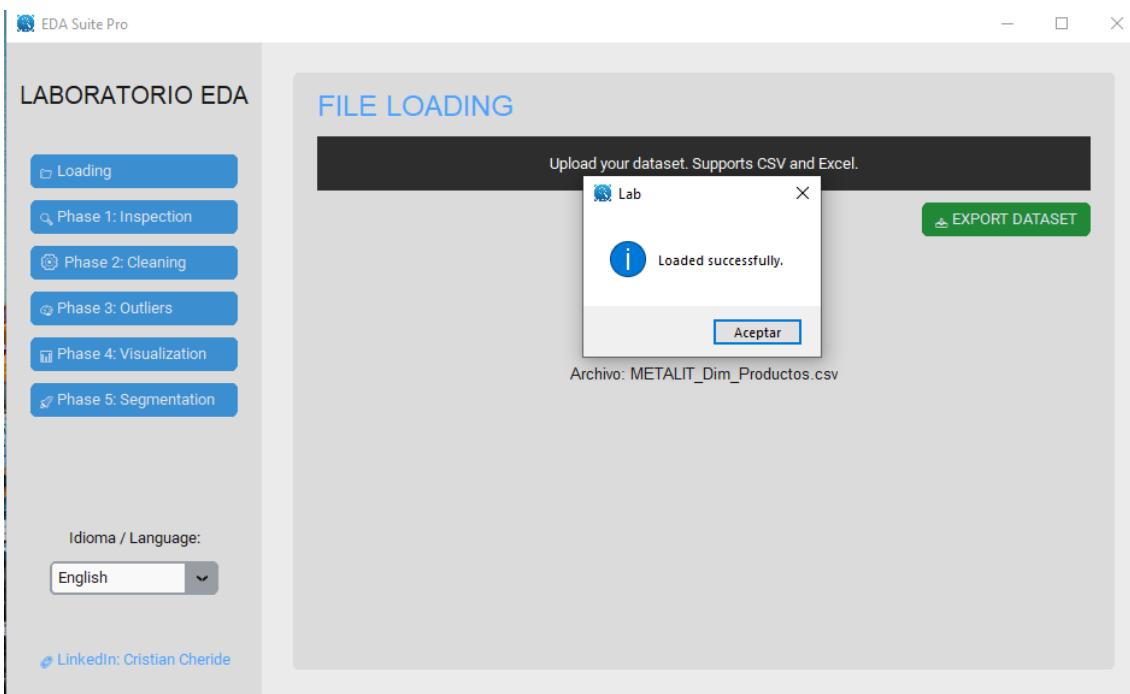
**Sidebar and Multilingual** "Navigation side panel with active language selector, allowing you to instantly switch the entire interface between Spanish and English."

**Phase 0 - Loading Module** "Main interface of Phase 0 ready to receive a dataset in CSV or Excel format."



**File Selection** "System file explorer allowing the user to choose the dataset for analysis."





**Phase 1 - Initial Diagnosis (360 Report)** "Phase 1 view showing the total count of rows, columns and an overall summary of the dataset's health."

The screenshot shows the EDA Suite Pro application window. The sidebar on the left is identical to the previous screenshot. The main area is titled 'PHASE 1: DIAGNOSTICS' and contains a sub-instruction 'Initial data quality analysis.'. Below this, there are three buttons: 'Reporte 360' (highlighted in blue), 'View Nulls', and 'View Duplicates'. A large text box displays diagnostic information:

```
Total Filas: 3086
Total Columnas: 14

Nulos:
id_transaccion      0
fecha               0
id_cliente          62
id_vendedor         93
id_producto         0
cantidad            0
precio_lista_unitario  0
descuento_pct       0
precio_final_unitario  0
total_venta         0
costo_total         0
margen_total        31
estado_pedido       0
canal_venta         0
dtype: int64

Duplicados: 46
```

**Phase 1 - Null and Duplicate Map** "Detailed audit that identifies exactly which columns have missing data and how many rows are repeated."

The screenshot shows the EDA Suite Pro application window. On the left sidebar, there is a vertical list of phases: "Loading", "Phase 1: Inspection", "Phase 2: Cleaning", "Phase 3: Outliers", "Phase 4: Visualization", and "Phase 5: Segmentation". Below this list is a "Idioma / Language:" label followed by a dropdown menu set to "English". At the bottom left, there is a LinkedIn link: "LinkedIn: Cristian Cheride". The main content area is titled "PHASE 1: DIAGNOSTICS" in blue text. Below the title, a dark box contains the text "Initial data quality analysis.". To the right of this box is a green button labeled "EXPORT DATASET" with a cloud icon. In the center of the main area are three blue buttons: "Reporte 360", "View Nulls", and "View Duplicates". To the left of these buttons, there is a section titled "MAPA DE NULOS:" with the following data:

id_cliente	62
id_vendedor	93
margin_total	31

 EDA Suite Pro

## LABORATORIO EDA

[Loading](#)[Phase 1: Inspection](#)[Phase 2: Cleaning](#)[Phase 3: Outliers](#)[Phase 4: Visualization](#)[Phase 5: Segmentation](#)

### PHASE 1: DIAGNOSTICS

Initial data quality analysis.

[EXPORT DATASET](#)

[Reporte 360](#)[View Nulls](#)[View Duplicates](#)

	id_transaccion	fecha	id_cliente	id_vendedor	id_producto	cantidad	precio_unitario	descuento_pct	precio_final	canal_venta	total_venta	costo_total	margin_to_tal
45	52649	2025-10-16		C0012	V009	1004	4						
	12000	0.0614			11263.20		45052.80		33600.0				11452.
80	Pendiente Stock	Directo		C0066	V009	1002	1						
77	51403	2024-12-13			8821700.00		8821700.00		7980000.0				841700.
	9500000	0.0714											
00	Entregado	Directo		C0205	V008	1004	19						
164	51070	2024-10-01			11151.60		211880.40		159600.0				52280.
	12000	0.0707											
40	Entregado	Televenta		C0064	V001	1008	87						
188	51439	2024-12-20			15728.40		1368370.80		1017900.0				350470.
	18000	0.1262											
80	Entregado	Televenta		C0009	V007	1008	128						
193	52324	2025-07-25			16673.40		2134195.20		1497600.0				636595.
	18000	0.0737											
20	Pendiente Stock	Televenta		C0150	V010	1005	63						
231	51139	2024-10-18											
	8500	0.1525			7203.75		453836.25		401625.0				52211.
25	Pendiente Stock	Televenta											
222	52523	2025-09-19		C0075	V005	1000	1						

Idioma / Language:

[LinkedIn: Cristian Cheride](#)

**Phase 2 - Text Normalization** "Using the 'UPPER & Strip' tool to clean up white spaces and standardize texts to uppercase in bulk."

The screenshot shows the EDA Suite Pro interface. On the left, a sidebar titled "LABORATORIO EDA" lists five phases: Loading, Phase 1: Inspection, Phase 2: Cleaning, Phase 3: Outliers, Phase 4: Visualization, and Phase 5: Segmentation. The "Phase 2: Cleaning" button is highlighted. The main window title is "PHASE 2: CLEANING" with the subtitle "Text normalization and format correction.". A toolbar at the top right includes an "EXPORT DATASET" button. Below the toolbar, there are three tabs: "Normalize", "Duplicates", and "Formulas". The "Normalize" tab is selected, showing a button labeled "UPPER & Strip". A dropdown menu below the button lists "0 - id\_transaccion". At the bottom of the main window, there are three buttons: "Date", "Integer", and "Decimal". On the far left of the main window, there is a "Idioma / Language:" dropdown set to "English" and a LinkedIn link.

This screenshot is identical to the one above, but the "Duplicates" tab in the toolbar is now selected instead of "Normalize". The "UPPER & Strip" button is still visible, and the dropdown menu shows "0 - id\_transaccion". The LinkedIn link and language selection remain the same.

**Phase 2 - Statistical Imputation of Nulls** "Selection menu to fill in missing data using methods such as Mean, Median, or Mode."

**Phase 2 - Custom Formula Editor** "Application of direct mathematical logic to the columns to resolve nulls through specific calculations."

The screenshot shows the EDA Suite Pro software interface. On the left, there's a sidebar titled 'LABORATORIO EDA' with the following buttons:

- Loading
- Phase 1: Inspection
- Phase 2: Cleaning (selected)
- Phase 3: Outliers
- Phase 4: Visualization
- Phase 5: Segmentation

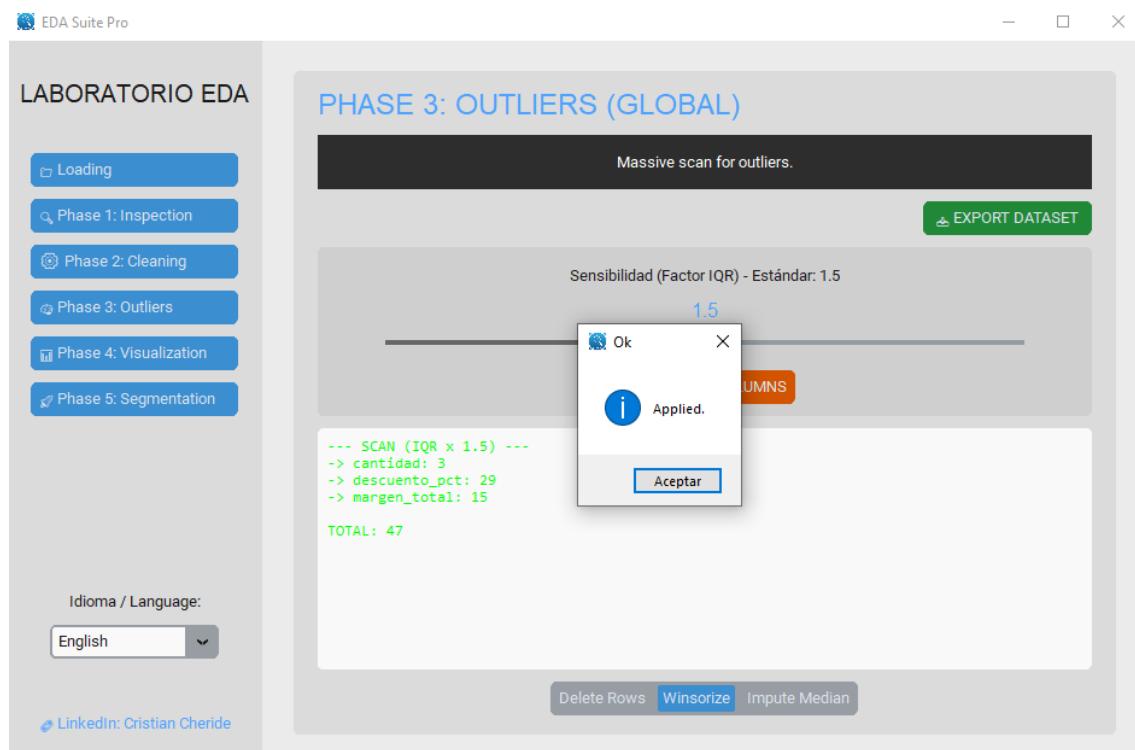
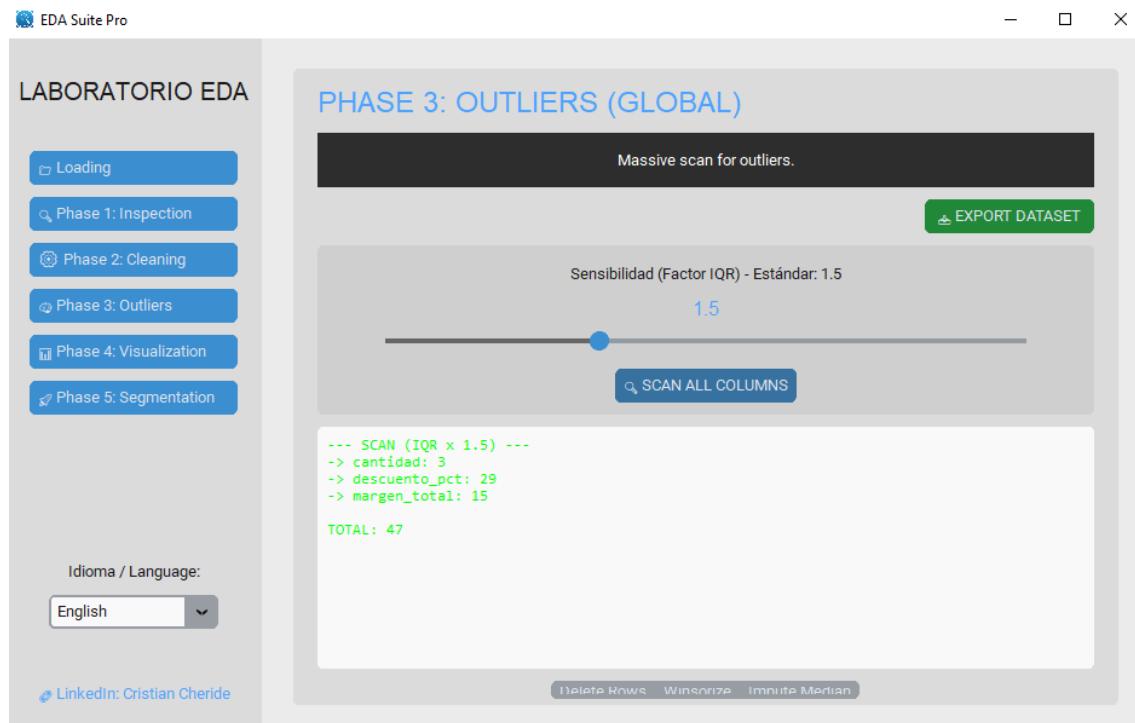
Below these buttons is a language selection dropdown labeled 'Idioma / Language:' with 'English' selected. At the bottom of the sidebar is a link to 'LinkedIn: Cristian Cheride'.

The main area is titled 'PHASE 2: CLEANING' and contains the following elements:

- A header bar with the text 'Text normalization and format correction.'
- A 'EXPORT DATASET' button with a file icon.
- A toolbar with three tabs: 'Normalize', 'Duplicates', and 'Formulas' (selected).
- A dropdown menu showing '11 - margen\_total'.
- Four buttons for statistical methods: 'Mean', 'Mediana', 'Moda', and 'Zero'.
- A text input field containing '9 - 10'.
- A blue button labeled 'Aplicar Fórmula Lab'.

**Phase 3 - Global Outlier Scan:** "Execution of the automatic diagnostic that analyzes all numerical columns simultaneously for outliers."

**Phase 3 - Sensitivity Adjustment (IQR):** "Manipulation of the sensitivity slider to define how strict the outlier detection filter (K-factor) will be."



**Phase 4 - BI Graphics Configuration** "Parameter selector for visualization: choice of axes (X, Y, Z), chart type and color variable."

EDA Suite Pro

### LABORATORIO EDA

- Loading
- Phase 1: Inspection
- Phase 2: Cleaning
- Phase 3: Outliers
- Phase 4: Visualization
- Phase 5: Segmentation

Idioma / Language:

English

LinkedIn: Cristian Cheride

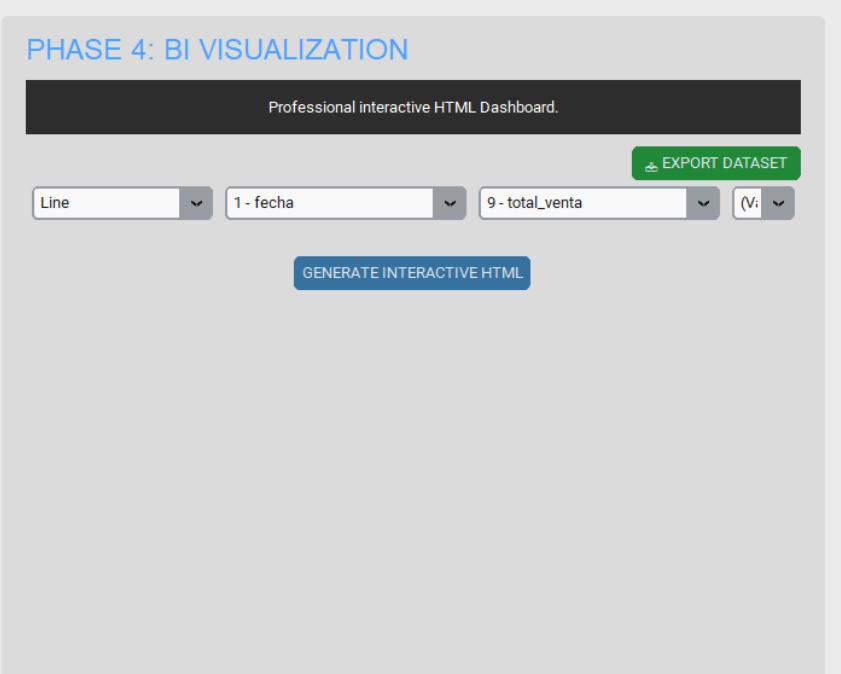
### PHASE 4: BI VISUALIZATION

Professional interactive HTML Dashboard.

EXPORT DATASET

Line 1 - fecha 9 - total\_venta (Vi)

GENERATE INTERACTIVE HTML



EDA Suite Pro

### LABORATORIO EDA

- Loading
- Phase 1: Inspection
- Phase 2: Cleaning
- Phase 3: Outliers
- Phase 4: Visualization
- Phase 5: Segmentation

Idioma / Language:

English

LinkedIn: Cristian Cheride

### PHASE 4: BI VISUALIZATION

Professional interactive HTML Dashboard.

EXPORT DATASET

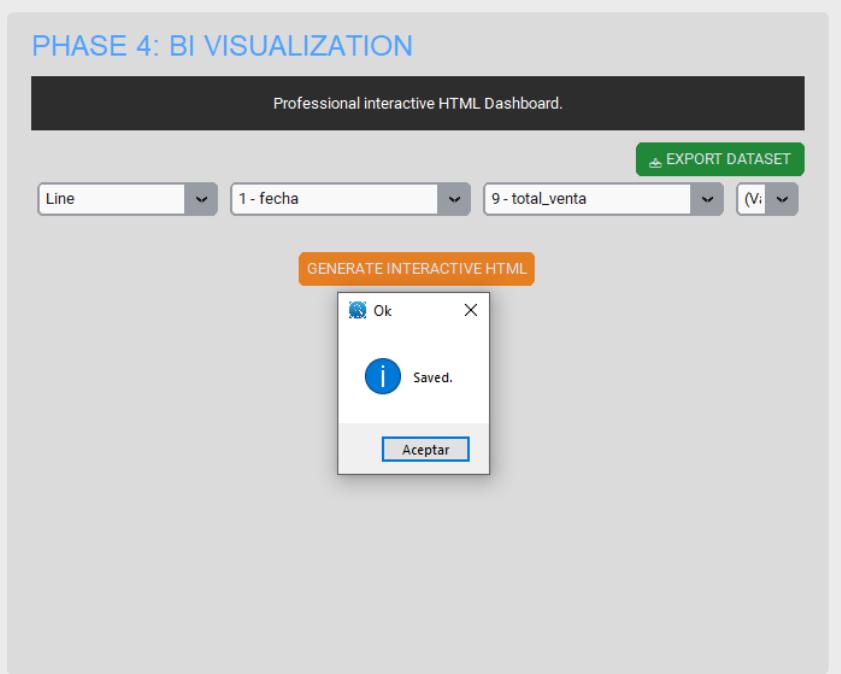
Line 1 - fecha 9 - total\_venta (Vi)

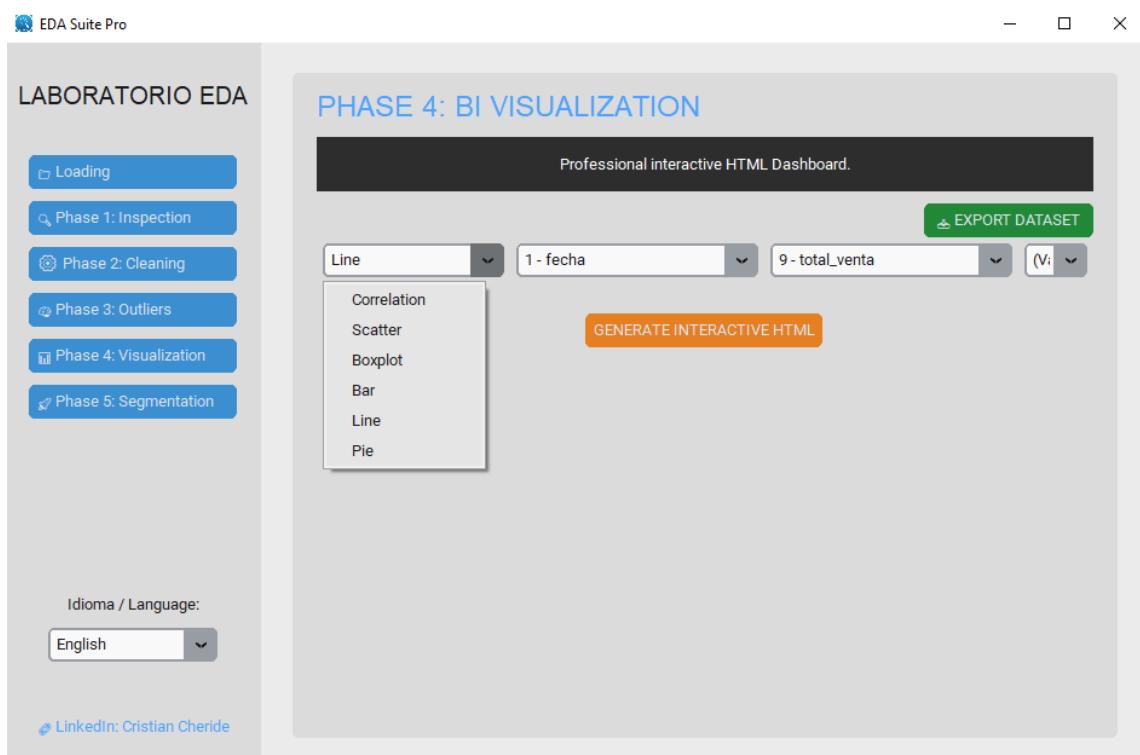
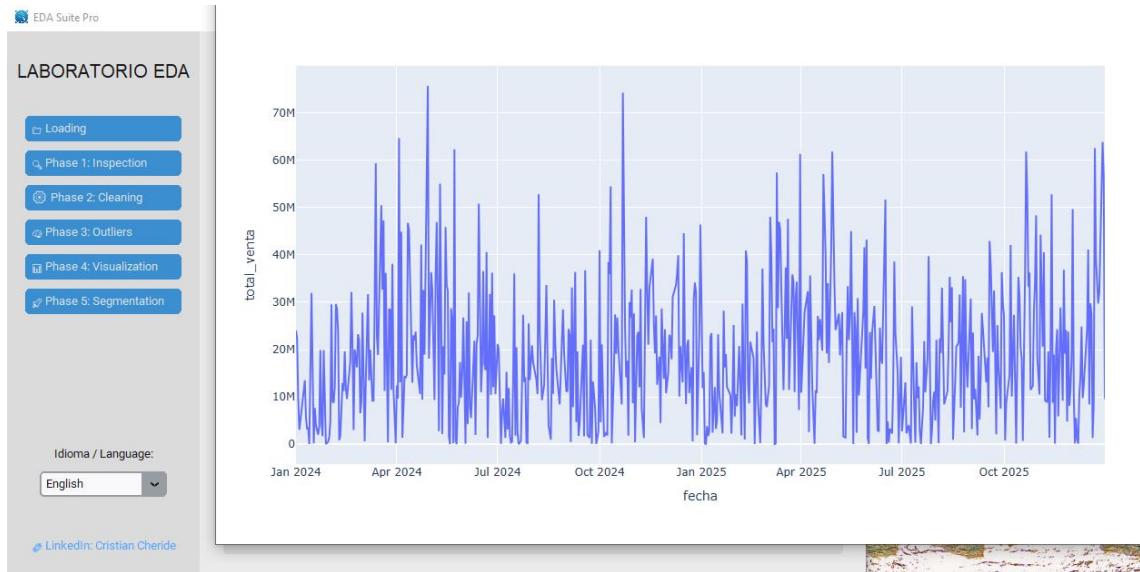
GENERATE INTERACTIVE HTML

Ok X

! Saved.

Aceptar

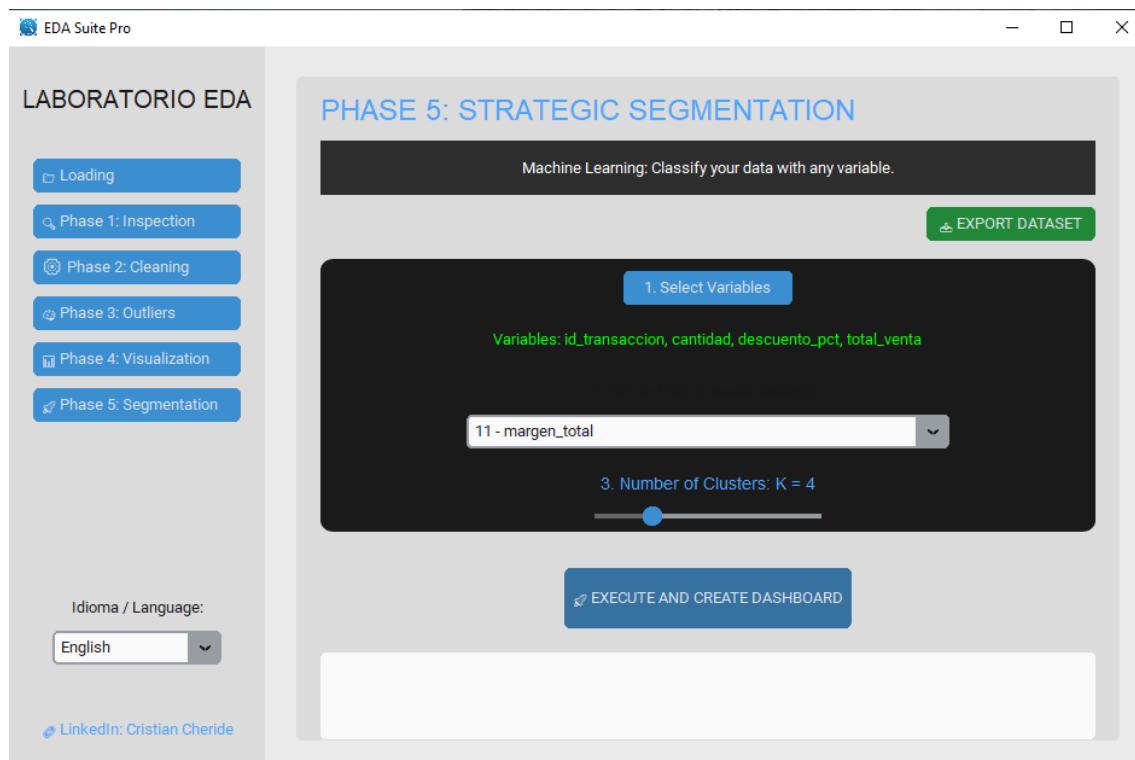
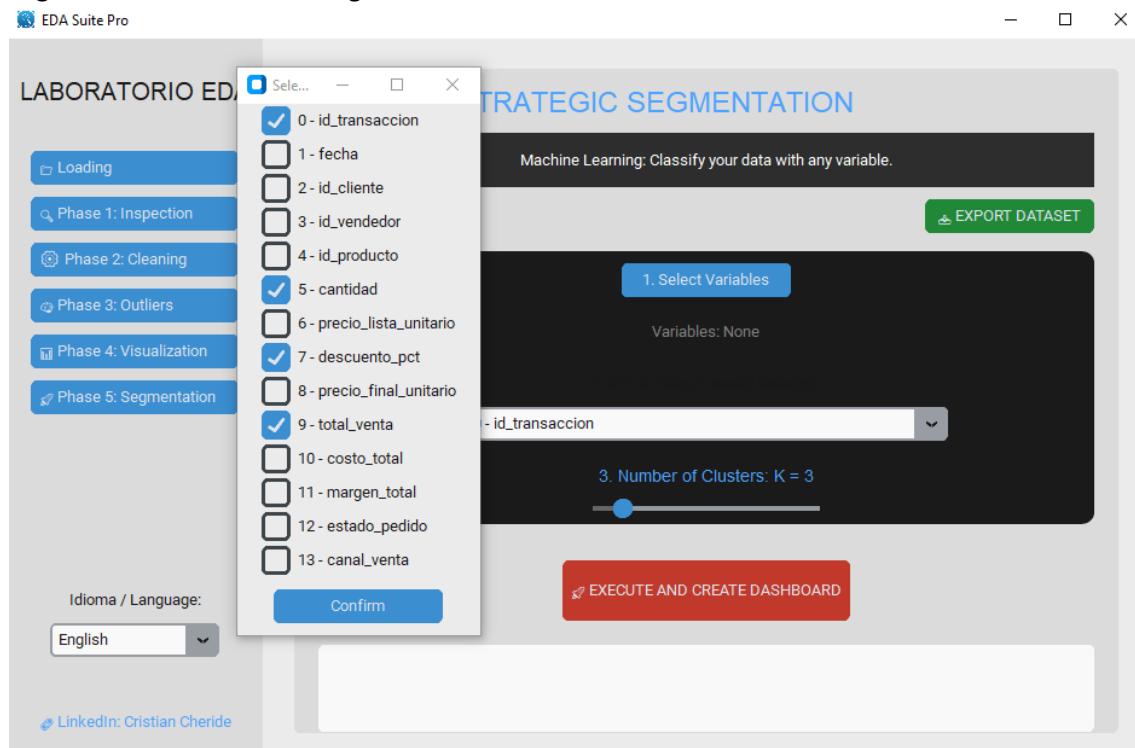




**Phase 5 - Variable Selection for Segmentation** "Pop-up window allowing the user to choose which columns (including IDs or text) will be used for the Machine Learning model."

**Phase 5 - Cluster Configuration (K=X)** "Adjustment of the desired number of clusters using the dynamic slider with real-time confirmation label."

**K-Means Algorithm Execution** "Results console displaying the averages for each generated segment and the cluster assignment to the dataset."



**Interactive 3D Dashboard** "Three-dimensional visualization of segmentation, where the size and color of the bubbles represent the selected success KPI."

