Why we take each decision in the dashboard?

1) Why we use Power BI

We chose Power BI because it offers the ideal balance between **speed of development**, **visual clarity**, and **analytical capability** for a clinical dashboard. Unlike purely programming tools, Power BI allows you to build clean, consistent visualisations in a very short time, while maintaining precise control over filters, interactions and design. Its interface makes it easy to integrate multiple graphs on a single page without losing aesthetic consistency, and its data engine allows you to handle complex clinical tables without the need for additional code. Furthermore, Power BI is a widely used tool in healthcare and management environments, which reinforces reproducibility and familiarity for non-technical audiences. Overall, it provides an optimal environment for transforming statistical analyses into a clear, professional, decision-oriented visual narrative.

2) Why we choose each graphic

Scatter plot: Max HR vs ST Depresion

Why is suitable?

- **Represents 2 continuous variables and allows you to observe patterns, clusters, and correlations**
- **Is the best channel to show the functional relationship between phisiologic variables**

Box plots: Cholesterol vs Heart Disease / Blood Pressure vs Heart Disease

Why are suitable?

- **Allow you to compare distibutions between 2 groups (Presence vs Absence)**
- **Show median, IQR & outliers**
- **It's the standard graph in clinical analysis for continuous variables**

Bar plot: ST Slope vs Heart Disease

Why is suitable?

- **ST Slope is cathegorical, the barplot is the most clear channel to compare frequencies**
- **Allows to see what patterns of ST Slope are associated more with the Heart Disease**

3) Why we choose the title of the Dashboard and titles of the graphs, axes & legends

For the general title and subtitle:

- It is clinical, direct, and professional.
- It summarises the purpose of the dashboard without unnecessary technicalities.
- It avoids overly long or specific titles that limit interpretation.
- It allows the dashboard to be reused in other contexts or cohorts.
- The subtitle adds context: it is not just a descriptive dashboard, but one oriented towards clinical patterns.

All graph titles follow a pattern:

- For relationship between 2 continuous variables (stress content): **Variable X vs Variable Y during exercise stress**
- For distributions of continuous biomarkers at rest: **Variable Levels at rest or Variable Distribution at rest**
- For prevalence or counts across categorical ECG features: **Counts of category during exercise stress**

For the titles of the axes:

- Must show clinical clarity
- Must show the variable with its unit (if applicable)
- Must have consistency in all graphs
- Must avoid long phrases

For the title of the legend:

- It must be brief and clear
- It must be precise
- It must be consistent with all the dashboard

The title of the legend will be **Heart Disease Status.**

The four repeated legends are replaced by a single global legend, making the dashboard clearer, more consistent and visually economical.

The legend is no longer a redundant element within each graph and becomes a cross-cutting marker, explaining the colour code for the entire panel in a centralised manner. This reduces visual noise, frees up space in each visualisation and reinforces the hierarchy: first the context (Heart Disease Status) is understood, and then the graphs are interpreted.

Each structure is chosen to match the visual grammar of the plot

4) Why we display the units of the variables in the axes

- In clinical dashboards, numerical values without units can lead to misinterpretation
- Showing units reinforces that the variables are **not directly comparable** and belong to different measurement systems.
- Showing units aligns perfectly with the principles: clarity, autonomy and low cognitive friction
- Showing units ensures reproducibility and transparency
- Showing units ensures that the graph is still interpretable outside the dashboard context

5) Why we choose the specific structure in the dashboard (the distribution of the layout)

- **First row: Stress Test Makers**
  Max Heart Rate vs ST Depression during exercise stress (Scatter plot)
  Counts of ST Segment Slope during exercise stress (Bar plot)

  Purpose:
  - **Shows dynamic cardiac response**
  - **Highlights ischemic patterns and ECG morfology**
  - **Connects stress physiology with disease presence**

- **Second row: Resting Risk Factors**
  Cholesterol levels at rest (Box plot)
  Blood Pressure distribution at rest (Box plot)

  Purpose:
  - **Shows baseline cardiovascular risk**
  - **Highlights distribution shifts**
  - **Complements stress-test findings with resting biomarkers**

STORYTELLING OF THESE PLOTS:

The story of the dashboard begins with the variables measured during exercise. These first two graphs show us how the heart responds when we take it out of resting conditions, always comparing patients with Absence and Presence of disease.

In the scatter plot of Max Heart Rate vs ST Depression, we see how the heart behaves under stress. Here, a clear pattern emerges: patients with Presence tend to be in areas of lower frequency response and greater ST depression, while those with Absence show a broader and more physiological response. Without the need for clinical interpretation, the scatter plot begins to separate and gives us the first diagnostic clue: when the heart is stressed, the disease limits its ability to respond.

The second graph, the ST Segment Slope bar plot, reinforces this idea from a categorical perspective. The morphology of the ST segment is a classic marker of ischaemia during exertion, and here we see how the distribution changes between groups: in Presence, flat or descending patterns increase, while in Absence, ascending patterns predominate. A simple variable becomes a powerful indicator that complements what we see in the scatter plot.

With these two graphs, we construct the first part of the narrative: how the heart behaves when we challenge it, and how disease modifies that response.

The second row brings us back to rest, to the risk factors that condition this behaviour. The Cholesterol and Blood Pressure box plots compare how these biomarkers are distributed between Absence and Presence. The differences are not huge, but they are consistent: in Presence, we see slight increases in blood pressure and cholesterol, which help to contextualise what we saw during the exercise. These are factors that do not manifest themselves directly in the stress test, but which shape the vulnerability of the cardiovascular system in the long term.

6) Why we choose the specific colours for representing the variables in the graphs

We chose **light green** for **Absence** and **light pink** for **Presence**.

The justification for this choice is:

- **Emotional neutrality:**
  - Light green conveys calm, stability, absence of alertness.
  - Light pink is soft, not aggressive, but introduces contrast.

- **Compatibility with clinical environments:**
  - These colours are frequently used in clinical dashboards, where the use of saturated colours is avoided.
  - They do not interfere with label reading or data interpretation.

- **Visual accessibility:**
  - Light colours are easily visible on screens and do not cause eye strain.
  - They are compatible with colour-blind readers when combined with clear patterns or labels.

- **Narrative coherence:**
  - Light green conveys ‘normality’ or ‘no findings’
  - Light pink conveys ‘presence’ without drama

7) Why we choose the specific typography, styles and captions

IMPORTANT NOTE: A single font family (Tahoma) is used throughout the dashboard to ensure visual consistency, reduce cognitive load, and maintain a clean and professional appearance.

- Title of the Dashboard:
  - Font: **Tahoma**
  - Size: **14 pt**
  - Weigth: **Bold**
  - Color: **Soft black**

  This is the highest level of visual hierarchy
  14 pt bold provides presence without overwhelming the layout
  Works perfectly in compact clinical dashboards

- Chart titles:
  - Font: **Tahoma**
  - Size: **11 pt**
  - Weigth: **Bold**
  - Color: **Soft black**

  This is the second level of hierarchy
  Must stand out clearly above the chart content
  11 pt bold is a standard in professional dashboards

- Axis titles (X & Y):
  - Font: **Tahoma**
  - Size: **8 pt**
  - Weigth: **Regular**
  - Color: **Dark gray**

  Axis titles must be **visible but not dominant**
  Bold would compete with chart titles and break hierarchy
  Dark gray reduces visual weight while maintaining readability

- Axis tick labels:
  - Font: **Tahoma**
  - Size: **8 pt**
  - Weigth: **Regular**
  - Color: **Black, 80% lighter**

  These should be the quietest text elements
  They support interpretation but should never attract attention
  Slightly lighter gray keeps them subtle

- Legends:
  - Font: **Tahoma**
  - Size: **9 pt**
  - Weigth: **Regular**

    - Color: **Dark gray**

    Same visual level as axis titles
    Ensures consistency across the dashboard
    Avoids competing with chart titles

- Plot descriptions:
    - Font: **Tahoma**
    - Size: **8 pt**
    - Weigth: **Regular**
    - Color: **Dark gray**

    Prevent text from competing with graphics

8) Why we choose a white theme

The advantages of using a white theme are:

- Maximum visual clarity
- Optimal contrast with text and graphics
- Facilitates printing and exporting
- Prevents interference with data colours

9) Why we choose this distribution of the titles

The titles of the graphs are aligned to the left at the top.

- Reinforces left-to-right reading (natural in Spanish and English)
- Visually align with the Y-axis and the rest of the dashboard.
- Enables quick scanning on dashboards with multiple charts

10) Why we change the range of the Y-axis in the boxplots & scatterplot

- Manual axis ranges ensure that the visualization remains clinically meaningful and avoids distortions caused by automatic scaling
- Manual ranges maintain visual coherence across the dashboard, preventing misleading comparisons caused by inconsistent auto-scaling
- Manual axis limits prevent extreme values from distorting the main distribution, ensuring that the boxplot remains readable and statistically informative

- Adding controlled visual margins enhances readability and prevents the plot from feeling cramped or cut off
- Manual axis ranges follow a documented, reproducible criterion that enhances methodological transparency

We manually adjust the Y-axis range to ensure clinical interpretability, maintain visual consistency across the dashboard, prevent distortion caused by outliers, and provide adequate visual margins. This approach avoids the inconsistencies of automatic scaling and ensures that the boxplots remain readable, comparable, and methodologically transparent.

BLOOD PRESSURE:

The observed values range from **94 to 200 mmHg**. A manual axis from **90 to 210 mmHg** captures the full distribution while adding a small visual margin above and below the extremes.

Values below 90 mmHg or above 210 mmHg are extremely rare in resting blood pressure and would not appear in this dataset. Setting the axis within this physiological window ensures that the plot remains realistic and interpretable.

CHOLESTEROL:

Cholesterol values range from **4.9 to 6.3 mg/dL**. A manual axis from **4.8 to 6.4 mg/dL** captures:

- the full distribution
- all outliers
- a clean margin above the maximum

Values below 4.8 mg/dL or above 6.4 mg/dL are extremely uncommon in typical clinical cohorts. Using this range ensures that the graph remains anchored in realistic biomedical limits.

MAX HEART RATE:

We set the Y-axis for Maximum Heart Rate to **60–210 bpm** to capture the full distribution of values while maintaining a clinically realistic range. This prevents distortion caused by extreme values, preserves the readability of the scatter plot, and ensures visual consistency across the dashboard. The manual range also enhances the interpretation of the relationship between ST Depression and Maximum Heart Rate, which forms the starting point of the clinical narrative.

11) Why we choose to display outliers in the boxplots

Outliers in biomedical variables are not “noise”: they often represent **clinically meaningful extreme cases** (e.g., very high cholesterol, hypertensive crises, unusually low blood pressure). Hiding them could mask important patterns or risk profiles.

Displaying outliers allows the dashboard to:

- detect extreme physiological values
- identify potential high-risk patients
- preserve the integrity of the clinical distribution

Outliers help the viewer understand:

- the shape of the distribution
- the presence of skewness
- the existence of extreme values

They add information without saturating the graph, as long as they are displayed discretely.

FORMAT:

- We use small circular markers because they provide a clean, neutral representation of extreme values without distracting from the main distribution.
- Outliers inherit the color of their category to maintain visual coherence and avoid introducing unnecessary visual elements.
- We use small markers so that outliers remain informative but do not visually compete with the main structure of the boxplot.

12) Why we avoid slicers & KPIs

We intentionally avoided slicers because the dashboard is explanatory rather than exploratory. The legend already provides a clear categorical distinction, and additional filters would add unnecessary visual noise without improving interpretability.

We chose not to include KPIs because the purpose of the dashboard is to reveal relationships and distributional patterns, not to monitor performance or highlight single aggregated values. KPIs belong to exploratory or operational dashboards, not to explanatory clinical visualizations.

13) Giving a gray background to the rest test plots

We applied a light grey background (#EEEEEE) to the resting-test box plots to visually distinguish them from the exercise-stress visualizations. This subtle background layer creates a clear thematic separation between the two physiological contexts (stress and rest) without introducing visual noise or competing with the colour encoding of Heart Disease Status. The grey tone is soft enough to maintain a clean clinical aesthetic while helping the viewer immediately recognise that the second row corresponds to baseline risk factors rather than stress-induced responses. This improves visual grouping and strengthens the narrative structure of the dashboard.

14) Tooltips in each graph

SCATTER PLOT

We enriched the scatter plot tooltip with Age and Sex, in addition to the core variables (Max Heart Rate, ST Depression, and Heart Disease Status). These two contextual variables improve clinical interpretation without adding visual noise, allowing the user to understand how stress-test responses vary across demographic subgroups.

BAR PLOT

We enriched the ST Segment Slope bar chart tooltip with the patient count and the percentage within each slope category, in addition to the core variables (ST Segment Slope and Heart Disease Status). These contextual indicators enhance the clinical interpretation of the slope distribution without introducing visual clutter, allowing the user to understand how the prevalence of heart disease varies within each morphological slope type.

BOX PLOTS

We enriched the boxplot tooltips with key statistical indicators: median, quartiles, non-outlier range, and sample size; in addition to the category label. These values support the interpretation of distributional differences between groups without adding visual clutter to the figure. The tooltip provides precise statistical context and preserving an align design of the boxplots.

15) Why we apply a log-transformation for Cholesterol

A log-transformation will be applied to “Cholesterol” because of its strong right-skewed distributions with a small number of extreme high values. The log transform reduces skewness, compresses large values, and stabilizes the variance, resulting in more normally distributed features.

16) Why we decide to add a short text/description of each graph

- They guide visual reading without overwhelming the reader
- They reinforce the narrative without invading the space
- Separate exploration from inference
- Increase academic defensibility
- They facilitate translation into reports or presentations

17) Why we decide to add statistical tests and p-values

Adding statistical values directly to each graph allows visual evidence to be linked to analytical evidence, reinforcing interpretation without forcing the reader to consult another section of the report. This improves methodological traceability, facilitates academic defence, and ensures that each visualisation is supported by an appropriate test. Furthermore, presenting the p-value and the type of test in the graph itself provides transparency, avoids ambiguity, and helps distinguish between merely descriptive patterns and relationships that have formal statistical support.