Presentation and Visualization

Delivery 2: Storytelling

10th January 2024 – Dafni Tziakouri, Victor Fayos, Alejandro Vara, Pere Blanch

Initially, we decided to persist with the "Adults" dataset chosen during Delivery 1 for the development of our presentation.

Dataset link: https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data

As agreed before in the first delivery, we assume that the main audience is a government organization or department responsible for labor and income-related policies and analysis (Government Analyst). Their expertise lies in data analysis, statistics, and policy research, with a specific focus on matters related to labor and income, but not necessarily being an expert in chart design. In contrast, our external company "Income Pulse Unit" specializes in visually representing demographic disparities within the population.

In their work, the Government Analyst primarily engages with visualizations on desktop computers or internal government systems. These visualizations serve as a crucial resource for their regular consultations, supporting decision-making processes and policy analysis within their domain of responsibility. In addition, we decided our second audience to be policy makers, researchers, and analysts involved in labor and income-related matters.

Therefore, after all the rounds of validation we have chosen the following theme for our presentation: Our objective is to unveil perspectives on income disparities, employment patterns, and the influence of education, marital status, and various factors on income levels within the population. Our emphasis will be on examining gender inequality. Some questions that could define our message is the following:

- "How does gender affect income levels?"
- "What is the income distribution among various education levels in our population?"
- "How does the marital status of each individual affect their income?"
- "Are there regional variations in income within our country?"

Furthermore, some user goals can be the following:

- **Gain insight into gender-based income disparities:**
    - Understand how gender influences income levels.
    - Identify areas of income inequality between male and female individuals.
- **Explore educational impact on income:**
    - Examine income distribution across various education levels.
    - Determine the correlation between education and income disparities.
- **Similar with occupation and marital status:**
    - Identify trends and disparities in income based on occupation.
    - Analyze the impact of marital status on individual income.
- **Examine regional variations in Income:**
    - Explore geographic differences in income distribution.

Inclusion of Pilot Users

For each round of validation and co-creation, we engaged both a government analyst and a policy maker as pilot users, ensuring a balanced representation of our target audience. Striking a balance was essential, given the analyst's emphasis on data-centric perspectives and the policy maker's focus on actionable government interventions. This approach allowed us to cater to the nuanced needs of both user groups effectively.

1st Round

During the initial round, the objective is to define the output and questions that will shape the direction of the presentation. The following highlights key questions posed, and insights gained during this phase.

Starting with the government analyst, he was asked about the relevance of the selected dataset to their usual analytical work and inquired if there are specific data points or variables they commonly analyze in their reports. As an insightful response, the analyst conveyed that the "Adults" dataset is well-stratified and adequately represented, negating the necessity for additional data sources. Additionally, the analyst emphasized key variables frequently considered, including gender, education, marital status, occupation, and country of origin. Subsequently, the decision was made to concentrate on these significant variables.

The analyst was queried about the types of visualizations or charts deemed most effective in conveying analytical insights, aligning with their accustomed modes of data interpretation. Additionally, he was prompted to identify any specific aspects of data visualizations that pose challenges in interpretation. The feedback received indicated a preference for simplicity and conciseness in plots, highlighting the need for straightforward visual representations, considering the diverse array of projects they engage in and the potential cognitive load on their minds. This ensured effective communication of findings to our audience.

Afterward, the policy maker was questioned about her objectives, and she expressed her aim to promote economic equality of opportunity for individuals with diverse demographics, all the while ensuring continued voter support for the upcoming candidature in two years. As a result, our focus pivoted towards highlighting income inequalities among various demographic groups. In particular, we made the strategic decision to place special emphasis on gender equality, as it addresses a significant public concern in contemporary discourse. Furthermore, we decided to present the information in a coherent manner, starting from basic disparities between genders and progressively delving into more complex aspects.

When questioned about existing policies, the policy maker noted ongoing efforts to tackle income inequality based on country of origin. She suggested performing a routine check for updates to contrast their current perspective without placing undue emphasis on the process.

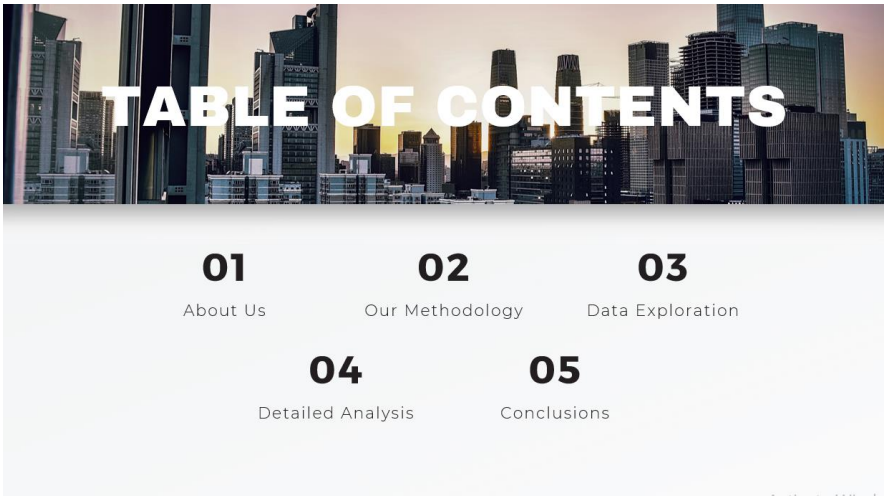
As we transition to the next round, our focus shifts towards transforming these valuable insights into a tangible prototype that mirrors the preferences and goals articulated by both the government analyst and policy maker. This prototype will serve as the bedrock for subsequent rounds of validation and co-creation.

2nd Round

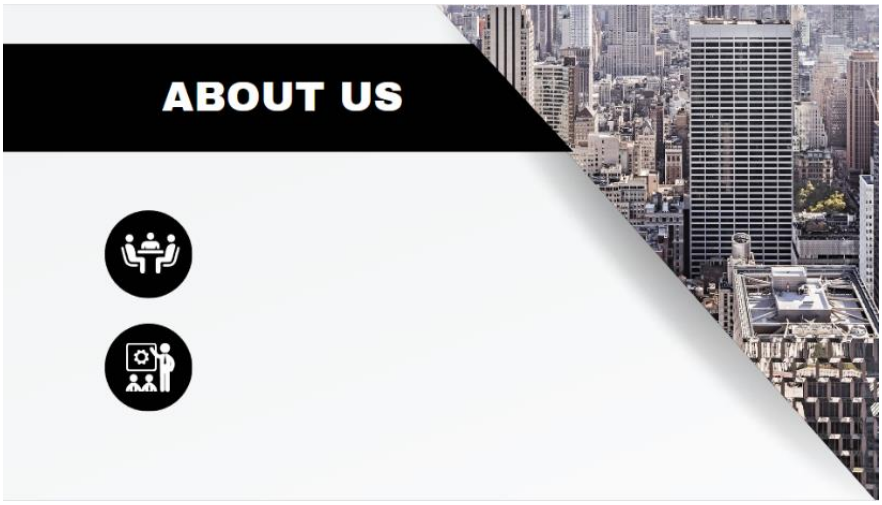
In crafting our presentation, we decided on adopting a design characterized by simplicity and minimalism. Our objective is to deliver the optimal content to our audience while maintaining a professional and disciplined demeanor. The following images are the initial ideas for our presentation: the first illustrates a draft of the main slide, the second the table of contents, and the third presents an idea of the subsequent slides.

Main slide:

Table of contents:



Additional slides:



Our intention is to incorporate images with a consistent theme across the majority of slides to enhance cohesion and create visually appealing presentations. Alternatively, if this approach is not feasible, we will keep a white background to maintain simplicity and cleanliness. Overall, each slide will feature a title along with the main key points of the corresponding discussion. This format aims to facilitate easy comprehension, sustain audience interest, and ensure that important details and information are effectively communicated.

Upon sharing the design with both pilot users, they expressed their approval. However, they recommended steering clear of overloaded slides, particularly when visualizing analytical insights. In response, we made the decision to remove background images in such slides, ensuring a clear and uncluttered presentation of our analytical findings.

Following that, we commenced the development of the prototype. One can review it by examining the attached *Storytelling_Prototype.pptx* file, which was presented to both pilot users. The slides are intentionally not filled with all the information; only selected graphs have been included. This approach serves as a guide for the pilot users, providing a preview of how the slides, the coherence and organization of the presentation will take shape.

The pilot users expressed positive feedback regarding the inclusion of information about our external company and our methodology. This transparency aims to further establish credibility and confidence in the forthcoming presentation. Additionally, discussing the dataset and clarifying the way the data is presented aims to prevent any potential confusion.

While the plots were thoroughly explained in the brief presentation of the prototype, they advised us to revise the titles for a better understanding. Furthermore, they suggested altering the color scheme between males and females to avoid potential bias and recommended incorporating consistent legends in each plot. Concerning the map illustrating the percentage of high-income individuals from different countries, they expressed concerns about its clarity. Consequently, we decided to replace it with a simpler plot, opting to represent the continent of origin instead of individual countries.

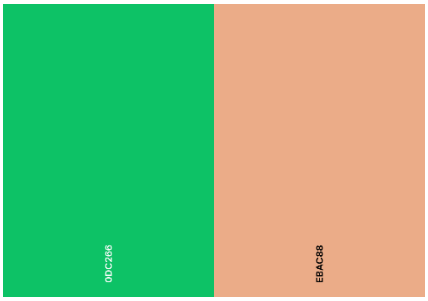
Both pilot users appreciated the structure of the detailed analysis, commencing with an overview of general disparities between genders and progressing to an exploration of the specific attributes contributing to the notable discrepancies. The adoption of causality storytelling not only fosters a more profound comprehension of the main ideas but also aligns with our commitment to clarity and coherence in delivering insights. Their positive feedback on this approach reinforces our strategy of guiding the audience through a logical sequence for a better understanding and memorability, ensuring that the key ideas presented leave a lasting impression on the audience.

Based on their reactions, we also recognized that there is some potential for improved engagement. Acknowledging the importance of allowing time for them to contemplate their initial hypotheses about the topic, we made the decision to incorporate a quiz before delving into the detailed analysis. This interactive element not only enhances engagement but also serves as a strategic tool to encourage active participation and reflection among the listeners. The quiz will be presented to them in the third round.

3rd Round

In the third round, we have implemented all the recommendations put forth by our pilot users. This includes adjusting the color schemes for males and females, enhancing the clarity of titles, incorporating consistent legends for each visualization, and replacing the map with a simpler and more intuitive representation, more precisely with pie charts.

The color scheme underwent a transformation, incorporating two distinct colors and various tonalities depending on the visualization. It is worth noting that both colors are complementary, positioned opposite each other on the color wheel, thereby enhancing their distinguishability between them and against the white background. In line with our commitment to formality and minimalism, black and white persisted throughout the entirety of the presentation, contributing to a cohesive and refined visual aesthetic. This thoughtful adjustment not only catered to visual clarity but also maintained a consistent and professional tone across the entire presentation.



We introduced the concept of the quiz to the pilot users, featuring the following set of questions:

Q1. What percentage of each gender earns over $50K?

Q2. For people with a master's degree, what percentage of each gender earns over $50K?

Q3. Considering married individuals, what percentage of each gender earns over $50K?

Q4. Guess the continent of origin with the lowest share of individuals earning over $50K.

Following each question, participants were presented with the actual answer, providing a valuable contrast between their assumptions and the factual data. Additionally, we shared the cumulative results of the entire audience, offering insights into the collective perspectives of their peers. This interactive approach not only encouraged individual reflection but also fostered a sense of communal understanding and awareness among the participants.

The pilot users noted a significant enhancement in the overall presentation with the incorporation of the quiz. Their acknowledgment underscored the positive impact and effectiveness of this interactive element.

Regarding storytelling, we adjusted it accordingly, following the outlined sequence:

1. Introduction: Discuss our identity and methodology.

2. Dataset Explanation: Provide insights into the dataset, avoiding misinterpretations.

3. Quiz Application: Engage the audience with a quiz, fostering both individual and collective reflection.

4. Gender Inequalities Discussion: Explore gender inequalities and highlight the demographic attributes most influential in these disparities.

5. Conclusions: Summarize key findings.

At this point we deliver the storytelling task as a zip composed by this report, the python notebook from which the plots were extracted, the first prototype *Storytelling_Prototype.pptx* and the final presentation *Storytelling.pptx*.

Explanation of the duties of each member of the group

Every team member actively participated in all facets of the project, contributing to the creation of the dashboard, the development of the storytelling presentation, and the creation of the reports. Throughout the project, constant communication fostered collaboration among the team.

Here are the designated positions for each team member as required in the delivery:

- Victor Fayos: Data Analyst. Responsibilities included analyzing and interpreting the data.

- Pere Blanch: Presentation Designer. Charged with designing visually appealing and effective presentations, ensuring consistency in design elements such as color schemes and fonts.

- Dafni Tziakouri: Content Strategist. Responsible for crafting engaging and informative narratives for presentations, ensuring that the storyline aligned with the goals and objectives of the presentation.

- Alejandro Vara: QA and Feedback Coordinator. Held the responsibility of conducting quality assurance checks on the presentation, collecting and consolidating feedback from pilot users and team members, and coordinating revisions and improvements based on the feedback received.