

DEVELOPMENT OF THE OPERATIONAL PLATFORM

D20 (PROTOTYPE)

IDEALVis Consortium

<http://idealvis.inspirecenter.org/>



European Union
European Regional
Development Fund



Republic of Cyprus



Structural Funds
of the European Union in Cyprus



**ΙΔΡΥΜΑ
ΕΡΕΥΝΑΣ ΚΑΙ
ΚΑΙΝΟΤΟΜΙΑΣ**

Executive Summary

This deliverable (prototype) provides a brief overview of the operational IDEALVis platform and its main components. Its objective is to guide the reader on the main development phases and provide a quick introduction to the platform.

The IDEALVis platform was developed to support the research and innovation activities of the project. As such, its main development phases coincide with the order of the three studies: (i) *user modelling*, which extracts the required user characteristics defined by the multi-dimensional human-centered user model; (ii) *visualisations engine*, which generates visualizations of different types and with different elements; and (iii) *adaptation engine*, which realizes the human-centred personalized visualizations. To support these phases, several additional components were developed including user interaction tracking, user management, security and data protection. Additionally, to support the final pilot study, important alterations were performed on the final platform.

The deliverable provides information about the methodology and each development phase, describing the main functionality developed at each step.

Table of Contents

EXECUTIVE SUMMARY	2
LIST OF FIGURES.....	4
1 Introduction	6
2 Development Methodology	6
3 Development Phases.....	6
3.1 User Modelling Phase.....	7
3.2 Visualizations Engine Phase.....	9
3.3 Adaptation Engine Phase	11
4 Platform Configuration for Final Pilot	16
5 Other Functionality	18

List of Figures

Figure 1 - Interface for creating or editing a research study.....	7
Figure 2 - Prompt that enables the user to join a user study.....	7
Figure 3 - Interface where the user can select to complete a research test from a specific user study.....	8
Figure 4 – Example psychometric test for Working Memory.....	8
Figure 5 – Example questionnaire for collecting data about data analysis tasks.....	8
Figure 6 - Statistics page given to a user once a test is completed.....	9
Figure 7 - Interface for viewing the results of user studies.....	9
Figure 8 - Interface for exporting the collected data from a research test found in a research study	9
Figure 9 – Examples of different visualization types produced by the Visualization Engine	10
Figure 10 – Examples of adapted visual elements (left: gridlines, right: dark theme enabled)	10
Figure 11 - Registration Page	11
Figure 12 - Login Page	11
Figure 13 - Activate Account Email	11
Figure 14 - Privacy Statement Page.....	11
Figure 15 - Terms and Conditions Page.....	11
Figure 16 - Security Tips Page	11
Figure 17 - Contact Support Form	11
Figure 18 - Privacy Settings Management.....	11
Figure 19 - Analysis Wizard – New/Load Analysis	12
Figure 20 - Analysis Wizard – Dataset Selection	12
Figure 21 - Analysis Wizard – Analysis Task Selection.....	12
Figure 22 - Analysis Wizard – Attribute selection and aggregation options	13
Figure 23 - Analysis Wizard – Filters settings	13

Figure 24 - Analysis Wizard –Output Method selection.....	14
Figure 25 - Analysis Wizard – Final Result	14
Figure 26 - Analysis Wizard – Save Analysis Workflow	14
Figure 27 - Analysis Wizard – Load Existing Analysis Workflow	14
Figure 28 - Active (primary) Dashboard	15
Figure 29 - Analysis Tracker Dashboard	15
Figure 30 - Analysis Tasks examples.....	16
Figure 31 –Task Questions Management (left: Create new Task, right: Management Panel).....	17
Figure 32 - Quick Tutorial – Pilot Study Tutorial Slides	17
Figure 33 - Administration – Manage Users.....	18
Figure 34 - Administration – Manage User Groups.....	18
Figure 35 - Administration – Manage Organizations.....	18
Figure 36 - Manage Data Protection Requests.....	18

1 Introduction

The IDEALVis platform was developed to support the research and innovation activities of the project. The main activities of the project included three user studies, which supported the construction of the multi-dimensional model, the assessment of the impact of the cognitive factors on the data visualization's types and elements and the evaluation of the adaptation mechanisms to improve the performance and accuracy of users for understanding data visualizations. To accomplish the above objectives and deliver the desired functionality to the end users, several components were developed with complex interactions and dependencies. The following sections, provide an overview of the main aspects concerning the development of the platform, including the software methodology and the main development phases. For the IDEALVis platform deployment setup, please refer to deliverable D23.

2 Development Methodology

The methodology used for the development of the IDEALVis platform was the agile methodology. The agile methodology was essential because of the dynamic nature of the project, which required development decisions to be taken after the outcomes of research tasks (e.g., analysis and assessment of user study results) and innovation tasks (e.g., selection of appropriate code structures to support specific software qualities, such as flexibility).

An adapted Scrum framework was selected to realize the IDEALVis platform, composed of three main phases to support the studies, where each phase focused on a specific software feature defined by multiple user/system requirements. Furthermore, each phase consisted of several "sprints", smaller iterative and incremental updates until the desired functionality was met. This approach has been proven effective in a plethora of projects and was able to cope with changes that arose during the development of the project.

3 Development Phases

The three main phases of development were:

1. **User Modelling Phase:** development of the user modelling platform to extract the required user characteristics defined by the multi-dimensional human-centered user model and allow users to assess their individual and group performance.
2. **Visualizations Engine Phase:** development of the first version of the IDEALVis platform that generates visualizations of different types and with different elements and tracks user interactions to support the second study.
3. **Adaptation Engine Phase:** development of the final version of the IDEALVis platform, which orchestrates all components to deliver human-centred personalized visualizations.

Finally, we also present specific configurations that were performed on the final version of the platform to support the final pilot study.

3.1 User Modelling Phase

This phase developed all the required mechanisms that would be used for building user models. User Modelling (i.e., the process of capturing the user's characteristics or attributes) in IDEALVis is done through the IDEALVis User Modelling System (IDEALVis-UMS). IDEALVis-UMS is an independent platform, which enables the creation of centralised user models. From the early stages of the project, it was decided that IDEALVis-UMS should be built and deployed as a separate service (decoupled from IDEALVis) able to communicate user model information to multiple systems via secured and well-defined API endpoints. The rationale behind this decision, was that after the project is finalized, IDEALVis-UMS can support future research activities by interacting with external platforms.

The user modelling system allows researchers to create “research studies”, which are collections of various “tests” (e.g., questionnaires, psychometric tests). As illustrated in Figure 1, the system allows researchers to create a study by providing a name, a description, a publication date (i.e., when the study will become available) and an expiration date (i.e., when the study will become unavailable). The access mode of the study can be set to two options: (i) open, which means that anyone with the link can access the study, or (ii) closed, which allows only selected user emails to receive the request. Regardless of the mode selected, all study participants need to register on the platform (or login if they have already participated in a previous study).

The screenshot shows a web form titled "Research Study". It has a "View All" button in the top right. The form contains several input fields: "Research Study Name" (with a placeholder "Please enter Research Study Name"), "Description" (with a placeholder "Please enter Description"), "Publication Date" (with a "Select Publication Date" dropdown), and "Expiration Date" (with a "Select Expiration Date" dropdown). Below these is a "Study Availability" section with two radio buttons: "Open Study" (selected) and "Closed Study". Under "Open Study", there is a text input for "Select Research Tests (Press and hold Ctrl (Windows) / Command (Mac) to select multiple tests)". A list of tests is shown below: "Business Role Questionnaire", "Data Analysis Tasks Questionnaire", "Data Visualization Experience Questionnaire", "Control of Attention", and "Speed of Processing". At the bottom, there are three buttons: "Publish", "Create Draft", and "Back to List".

Figure 1 - Interface for creating or editing a research study

When the study becomes available, the participants can access the study through a link that prompts them to login or register as illustrated in Figure 2.

The screenshot shows a "Join Research Study" prompt. It displays the "Research Study Name: IDEALVis Study" and the "IDEALVis Study" link. Below this, it shows the "Duration of study" as "01/01/2019 09:00 - 02/02/2020 17:00". At the bottom, there are two buttons: "Login" and "Register", with the text "to join the study" next to them.

Figure 2 - Prompt that enables the user to join a user study

The user is then able to access the study and check which tests are completed and which ones are pending. In the example presented in Figure 3, the participant has joined a study named IDEALVis Study, which is composed of three tests (questionnaires) related to understanding the participant's Business Role, performed Data Analysis Tasks and current Data Visualization

Experience. As it can be seen, the Business Role Questionnaire has been completed, while the Data analysis Tasks Questionnaire has not been started yet and the Data Visualization Experience Questionnaire is partially completed at 93%. The platform also assists the user to quickly assess the current level of completion by providing colour mappings to indicate the status of each test (i.e., Blue: Completed, Green: Not Started and Orange: Pending).

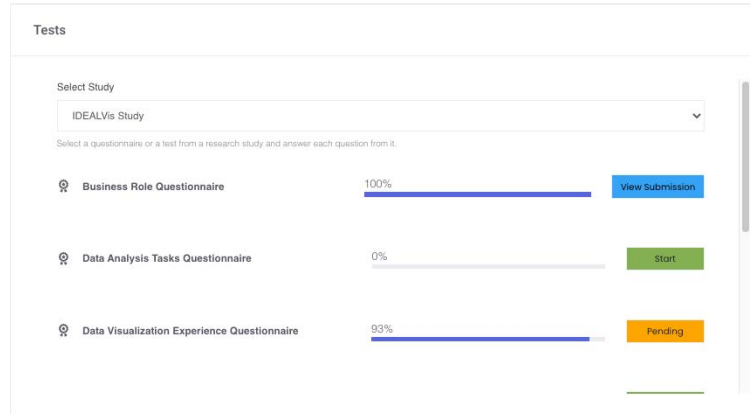


Figure 3 - Interface where the user can select to complete a research test from a specific user study

Some examples of specific tests' parts are presented in Figure 4 and Figure 5, displaying a part of a psychometric test for working memory and a questionnaire for collecting data about data analysis tasks.

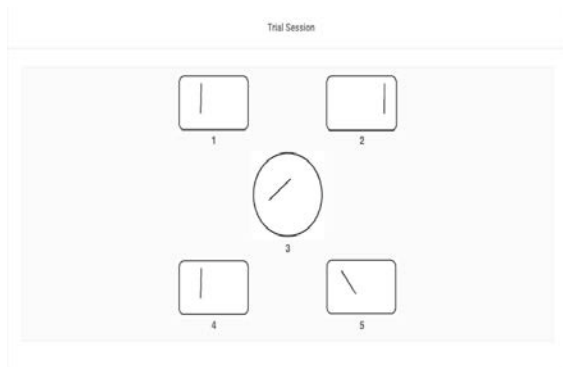


Figure 4 – Example psychometric test for Working Memory



Figure 5 – Example questionnaire for collecting data about data analysis tasks

After a test has finished, a participant can obtain some statistics about her/his performance and accuracy in some questions. This is used as means for self-reflection (i.e., to understand her/his performance/accuracy individual and within the study group) and gamification. Note that this of course applies only to tests that monitor time performance and accuracy.

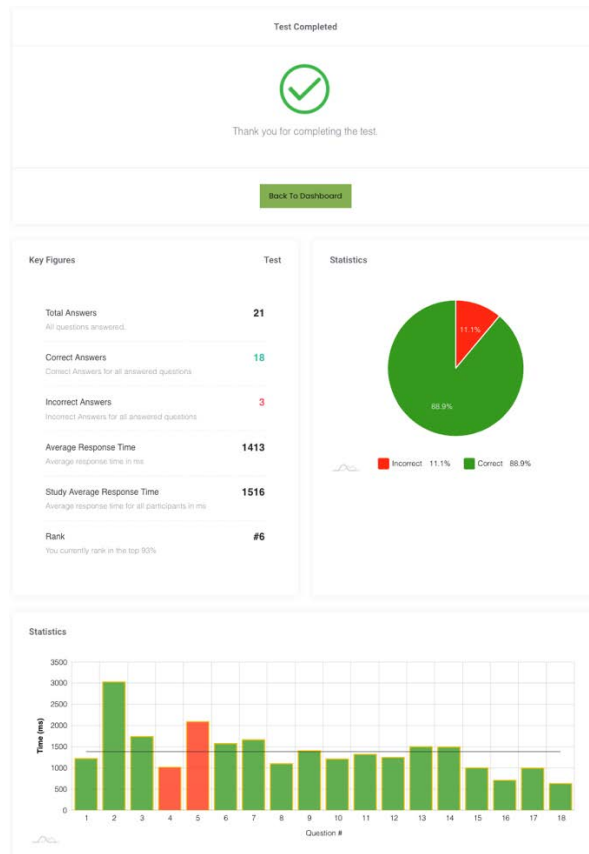


Figure 6 - Statistics page given to a user once a test is completed

Finally, a researcher can use a dedicated page (see Figure 7) to monitor the collection of responses for all studies, view overall statistics and export the results of a study (see Figure 8).

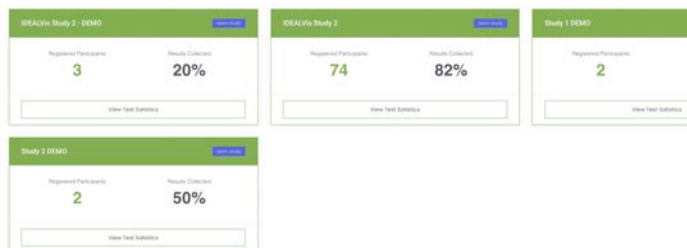


Figure 7 - Interface for viewing the results of user studies

Training Results for Business Risk Specialists

Participant ID	Gender	Nationality	Age	Age Range	Age Mile	Age Milestone	Organization Name	Business Role
1019	F	IT	31	30	30	30	ITMAD	Director
1019	F	IT	31	30	30	30	ITMAD	Manager
1017	F	IT	30	29	29	29	ITMAD	Analyst
1016	M	IT	30	29	29	29	ITMAD	Senior Analyst
1016	M	IT	30	29	29	29	ITMAD	Manager
1015	M	IT	30	29	29	29	ITMAD	Senior Analyst Digital Transformation
1015	M	IT	30	29	29	29	ITMAD	Analyst IT
1014	F	IT	30	29	29	29	ITMAD	Senior Manager Security Services
1014	F	IT	30	29	29	29	ITMAD	Senior Manager
1014	F	IT	30	29	29	29	ITMAD	Product Risk Analyst
1013	F	IT	30	29	29	29	ITMAD	Manager
1013	F	IT	30	29	29	29	ITMAD Digital Ltd	Manager
1013	F	IT	30	29	29	29	ITMAD	Senior Manager in the Financial Risk Management Department of ITMAD
1012	F	IT	30	29	29	29	ITMAD	Manager
1012	F	IT	30	29	29	29	ITMAD Limited Company	Business role in a group of related roles with no level of authority to perform a particular

Figure 8 - Interface for exporting the collected data from a research test found in a research study

Overall, IDEALVis-UMS is a flexible platform that can be easily extended with new tests and additional functionality for data pre-processing analysis and reporting for the purpose of retrieving the participants' user model information.

3.2 Visualizations Engine Phase

This phase includes the development of the IDEALVis platform to support the second study. It includes the development of one of the main components of the platform, the visualizations engine, which can produce visualizations of different types (e.g., bar, line, pie) and manipulate several of their elements (e.g., colour, bar width, gridlines). Several examples of different visualization types are illustrated in Figure 9.

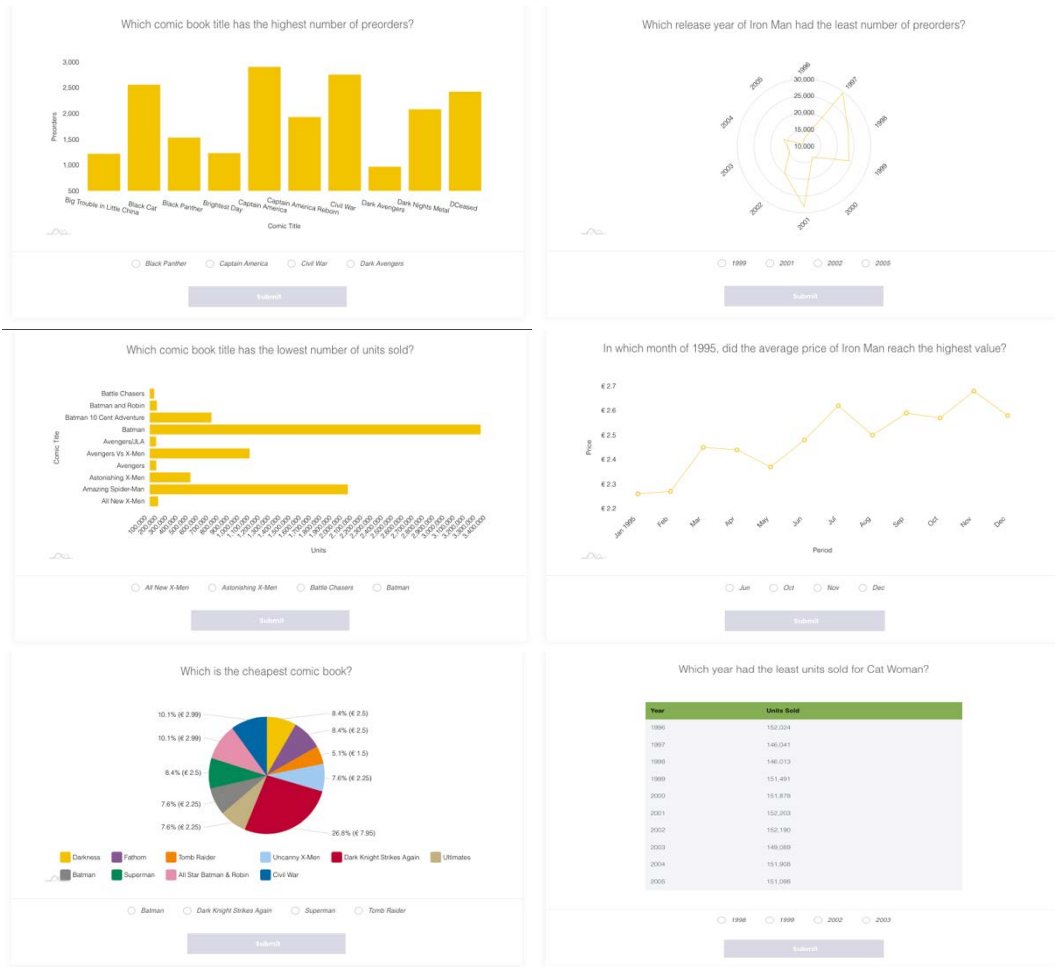


Figure 9 – Examples of different visualization types produced by the Visualization Engine

Visualization elements can be adjusted according to different visualization types. For example, proximity of bars can only be adapted for column and bar charts, while element size (i.e., line width or bar/column width) can only be adapted for line, radar, bar and column charts. Figure 10 shows examples of two adapted elements, gridlines and dark theme.

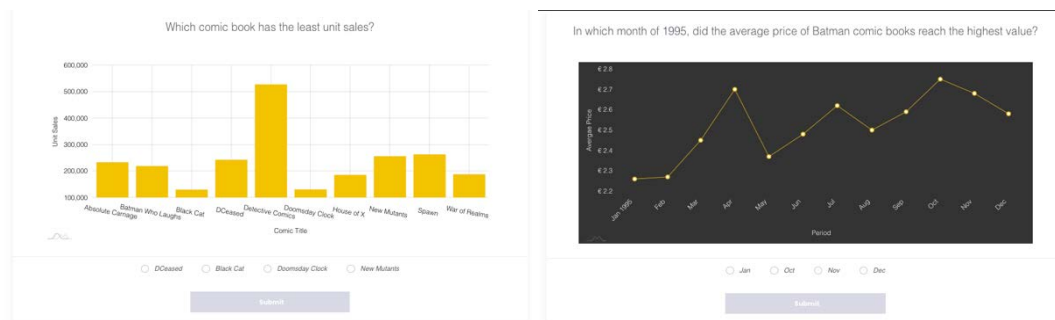


Figure 10 – Examples of adapted visual elements (left: gridlines, right: dark theme enabled)

Finally, this phase included the construction of many components to support different operations of users, such as registration, login and activation (Figure 11, Figure 12, Figure 13), information pages (Figure 14, Figure 15, Figure 16), support page (Figure 17) and privacy settings management (Figure 18).



Figure 11 - Registration Page



Figure 12 - Login Page



Figure 13 - Activate Account Email



Figure 14 - Privacy Statement Page



Figure 15 - Terms and Conditions Page



Figure 16 - Security Tips Page



Figure 17 - Contact Support Form

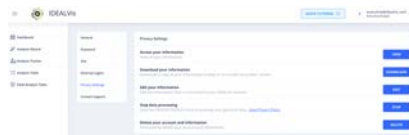


Figure 18 - Privacy Settings Management

3.3 Adaptation Engine Phase

This phase concerns the development of the final version of the IDEALVis platform to support the third pilot study. It includes the development of the adaptation engine that orchestrates the final versions of all components to deliver personalized visualizations to the users. Furthermore, it finalizes all system related aspects of the platform, including amongst others dashboard design, analytical workflows, security and data protection. For the internal mechanisms and algorithmics of the adaptation process, please refer to deliverable D15.

The development started with designing the analytical workflows, enabling the users to construct visual data analysis reports through an intuitive user interface without having to specify the underlying analysis queries using a specific query language. This was accomplished by first introducing the Analysis Wizard and the Query Engine, which are amongst the key technological innovations of the project. The Analysis Wizard features a step-by-step process for constructing data pipelines to visually explore data and discover patterns and hidden insights. The Query Engine is responsible for translating the user's Analysis Wizard interactions (e.g., selecting a dataset) to a realizable query language for execution. More information on the underlying mechanisms of the Query Engine can be found in deliverable D19.

The Analysis Wizard features a highly flexible and open design, which allows developers to easily incorporate more analysis functions (e.g., data mining and machine learning algorithms) to process data. The analytical workflow starts from the selection of a dataset as shown in Figure 20. The datasets available are preloaded in the knowledge base and are available for selection from the dropdown list.



Figure 19 - Analysis Wizard – New/Load Analysis

To emulate a realistic data processing platform, IDEALVis allows users to either create new analysis workflows or load existing ones (see Figure 19 bottom right) to save time. Next, the wizard provides some basic descriptive statistics for the dataset and also provides the ability to check descriptive statistics for each attribute.

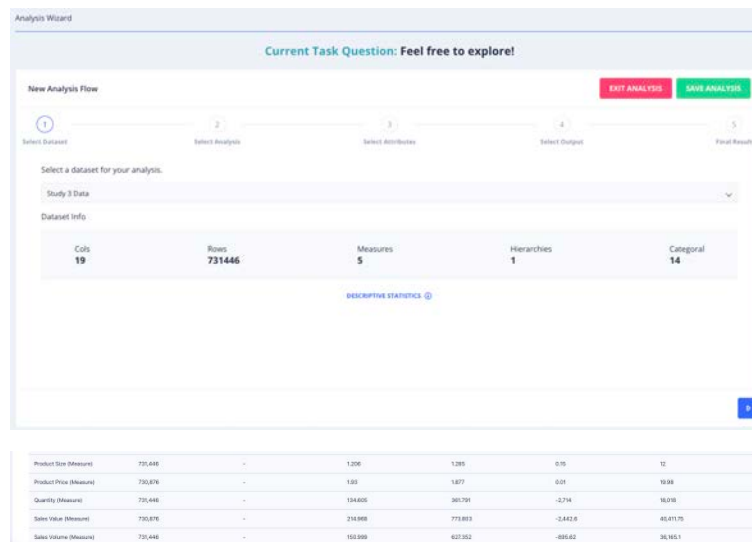


Figure 20 - Analysis Wizard – Dataset Selection

Next the user can select from a plethora of different analysis tasks as illustrated in Figure 21. The wizard supports simple (descriptive) analytical tasks, such as list report and top-k, and advanced ones, such as forecasting, distribution analysis and correlation analysis.

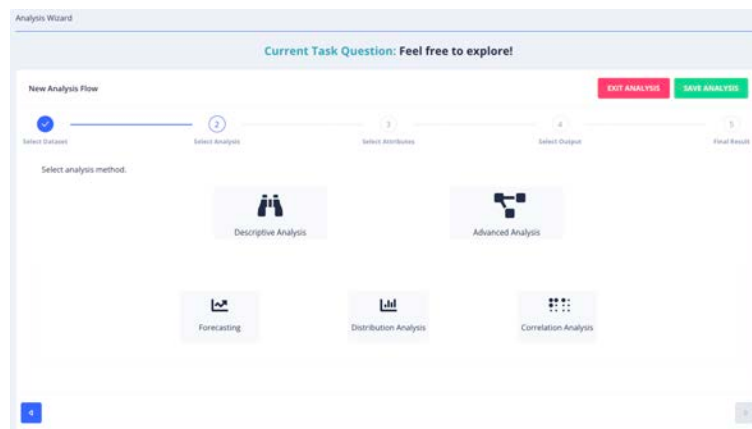


Figure 21 - Analysis Wizard – Analysis Task Selection

The wizard then allows the selection of attributes that will be used for the analysis task. The attributes should be appropriate for the selected task and be assigned appropriate roles. For

example, in the case of a list report that will be exported as a pie chart, the wizard requires one attribute that will serve as the category and another attribute that will serve as the measure. In the example illustrated in Figure 22, the user has selected a categorical value, the “Area Name” to serve as the category and the “Sales Value” attribute as the measure value. Additionally, the user has selected that the “Sales Value” be aggregated with the “Sum” aggregation function. As presented in Figure 22, the wizard supports multiple types of aggregation, such as Count, Min, Max, Unique (distinct values), etc.

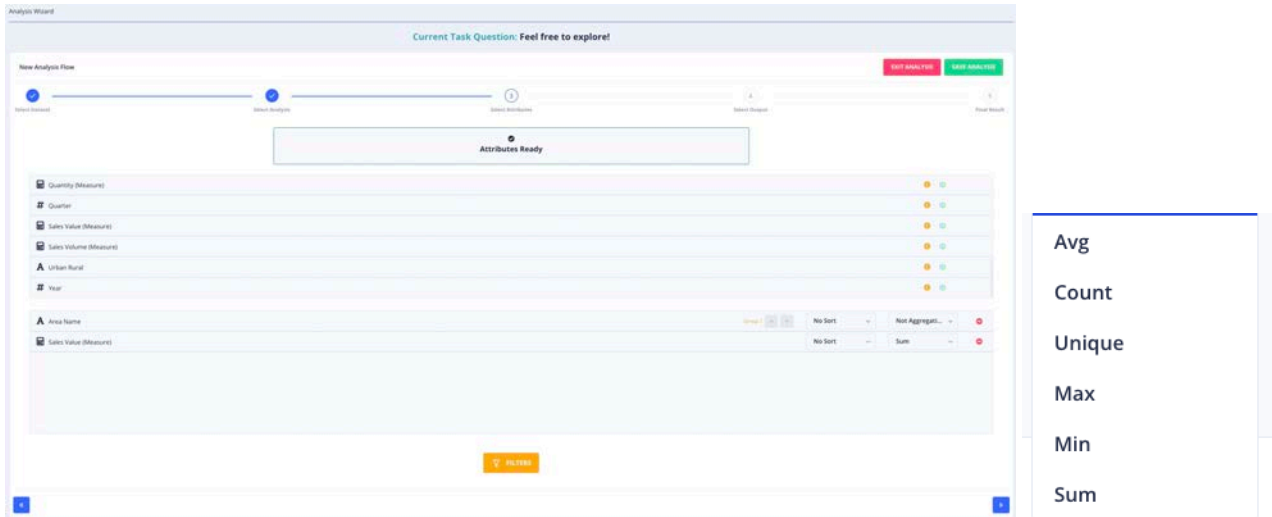


Figure 22 - Analysis Wizard – Attribute selection and aggregation options

Within the same step, the user can provide specific filters to select specific data to be pre-processed. Figure 23 shows an example where the user intends to create a filter to limit the results to a specific brand. Similar to modern platforms, the system facilitates the process by retrieving the unique values of the selected attribute (i.e., brand) for easy selection and also enables filtering at real-time using partial string matching. As illustrated at the bottom of Figure 23, the user can create multiple filters using the AND/OR logical operators.

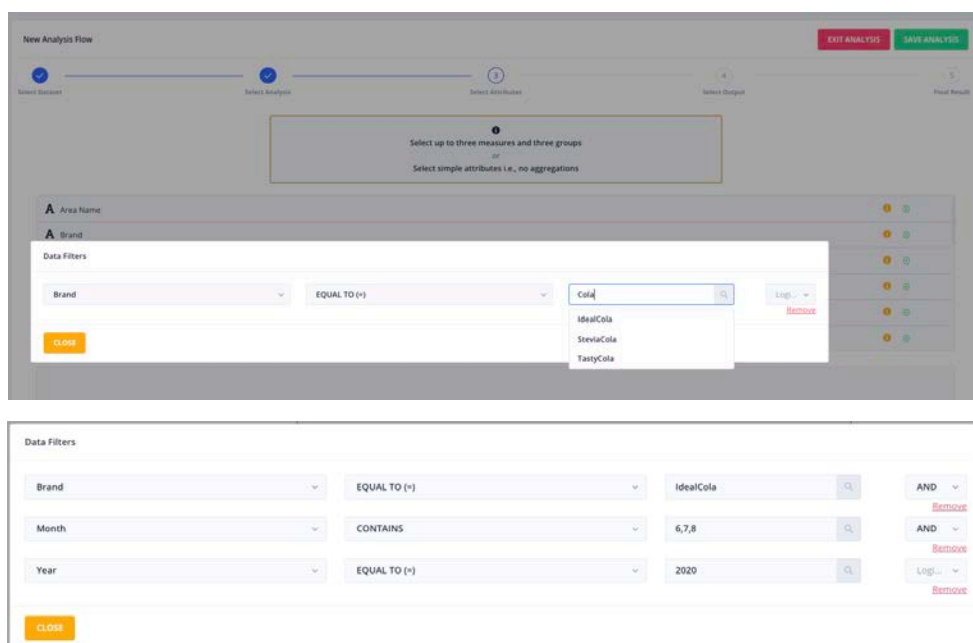


Figure 23 - Analysis Wizard – Filters settings

Next, the wizard allows users to select the output method as seen in Figure 24 and then generates the required output. Output methods include standard visualizations (e.g., table, pie chart, line chart, bar chart) and more advanced ones, such as tree map and radar chart. Specific analytical tasks limit the options of the output. For example, the distribution analysis utilizes box plots for output.

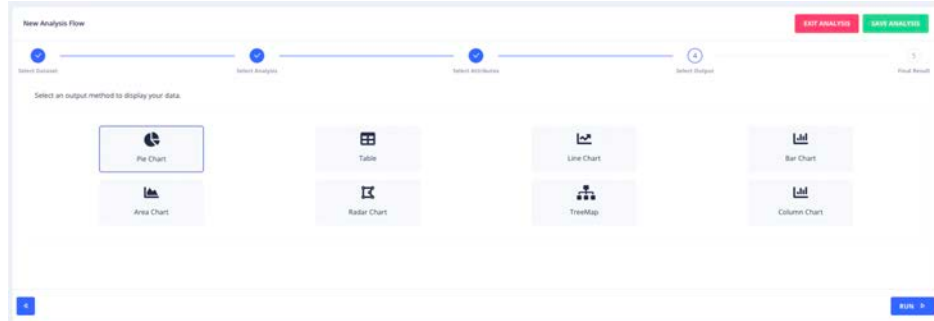


Figure 24 - Analysis Wizard –Output Method selection

Finally, the desired output is presented to the user. Figure 25 demonstrates an example of a generated pie chart. Once an analysis report is rendered the user can either explore it using the interactive visualization tools (e.g., zooming, removing series etc.), or even pin the analysis result to their personal dashboard (Figure 28).



Figure 25 - Analysis Wizard – Final Result

During this step, or any of the previous ones, the user can also save the analysis workflow so it can be loaded at any given time, as presented in Figure 19. Figure 26 and Figure 27 show an example of saving an analysis workflow and then loading it as a template for a new analysis back in the first step of the Analysis Wizard.

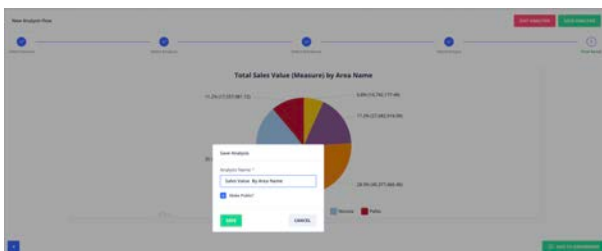


Figure 26 - Analysis Wizard – Save Analysis Workflow

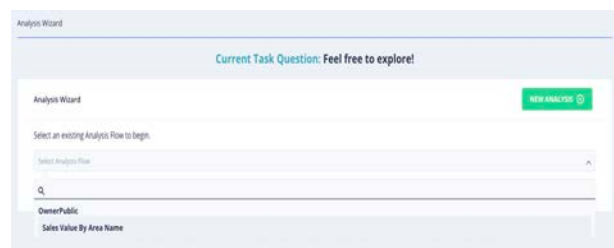


Figure 27 - Analysis Wizard – Load Existing Analysis Workflow

The IDEALVis platform supports an active (primary) dashboard and several custom (secondary) dashboards where users can store the results of their analytical workflows for further analysis.

Figure 28 presents an example of the dashboard, which integrates the results of two analysis. The user is able to change the location, using drag and drop functionality and also resize or delete the visualization using the hamburger menu on the top right.



Figure 28 - Active (primary) Dashboard

Finally, during this phase, the analysis tracking mechanisms were enhanced to monitor and record the user's data analysis interactions. These mechanisms enable the platform to track the different steps involved when a user is constructing a data analysis report via the Analysis Wizard and track the time a user is looking at a specific visual report. Using these data, the platform can provide an assessment on the efficiency of the user's data analysis process based on the performance level of an individual user but also based on the user's peers i.e., other system users. Such assessment of interaction is provided to users via the Analysis Tracker Dashboard (Figure 29), which helps users to view and understand their analytic interaction i.e., their analysis efficiency/performance, as well as assess their interaction patterns across specific analysis tasks.

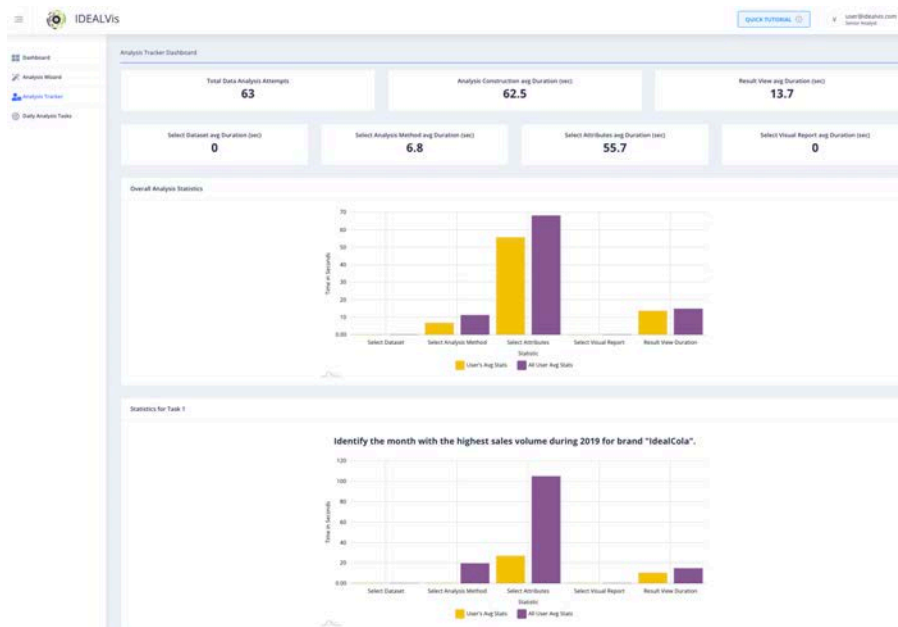


Figure 29 - Analysis Tracker Dashboard

4 Platform Configuration for Final Pilot

The pilot study was conducted on the final version of the platform with specific configurations to facilitate the study objectives. First, the dataset used in the pilot study was constructed in collaboration with the consortium's industry partners. The dataset is a realistic dataset composed of 731,446 observations, representing Soft Drinks sales. Several attributes are included in the dataset, representing important dimensions such as time, outlet and product. All participants assumed the role of a Brand Manager responsible to analyze the dataset and provide answers to specific analytical tasks (i.e., business questions). For more information about the pilot study scenario and settings, please refer to deliverable D24.

To support the pilot setting, the team created several business questions that the participants needed to answer by analysing specific dimensions of the dataset. The interface for selecting the task questions is presented in Figure 30, where users were able to select their active/current task in a dedicated page ("Daily Analysis Tasks").

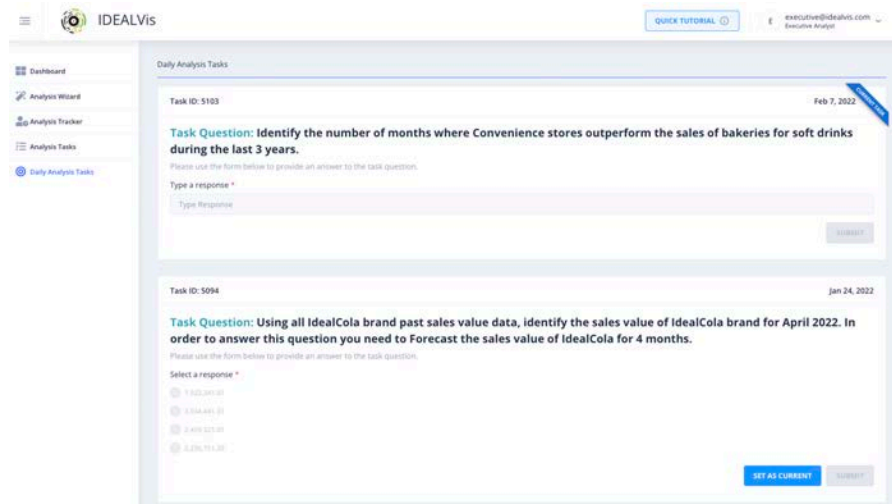


Figure 30 - Analysis Tasks examples

The user can select a task ("Task Question") and set it as the active/current one so that the Analysis Wizard defaults its parameters to the active task. When the users are finished with their analysis, they can visit the Daily Analysis Tasks and complete a text-based or multiple-choice-based answer for the task.

Formulating the task questions was performed in collaboration with the industry partners so that they represent real-world problems. To facilitate the task management process, a new component was developed allowing executing users to create new analysis tasks and manage them as illustrated in Figure 31.

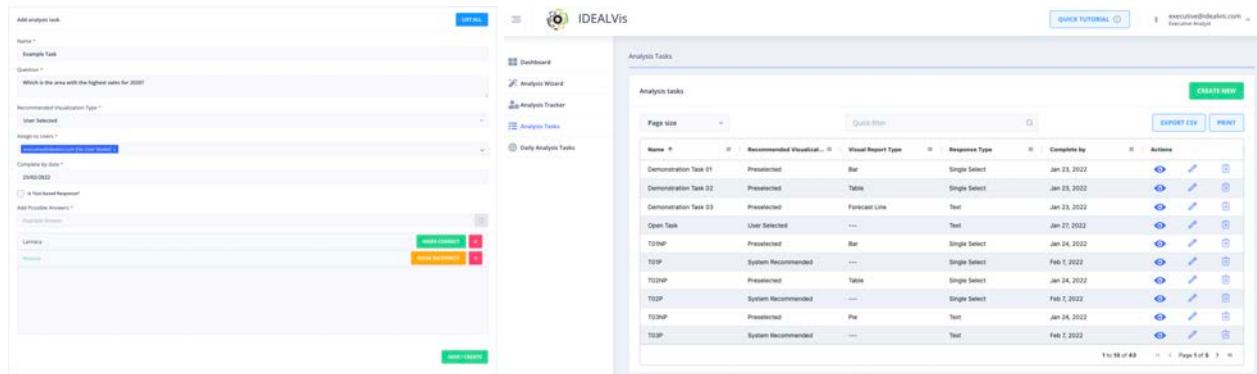


Figure 31 –Task Questions Management (left: Create new Task, right: Management Panel)

To allow training of the users on the platform and allow them to quickly learn the environment, a quick tutorial was introduced, demonstrating all features of the system. The tutorial was available to newcomer users when they logged in to the platform for the first time as illustrated in Figure 32.

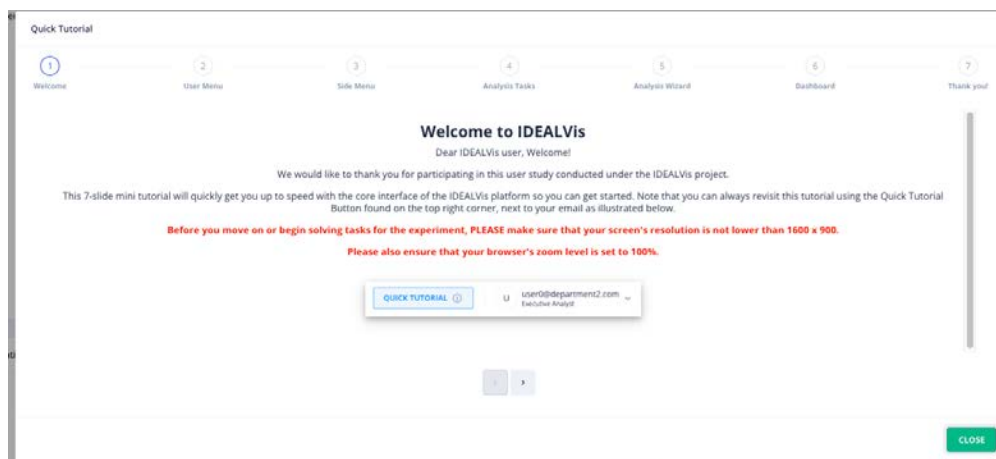


Figure 32 - Quick Tutorial – Pilot Study Tutorial Slides

Finally, the following important alterations to the IDEALVis platform were performed:

- The “Select Dataset” of the Analysis Wizard was disabled for the participants. Instead, the analysis dataset was pre-selected.
- The “Select Output Method” of the Analysis Wizard was hidden when a particular task had a predefined data visualization type, or if the adaptation mechanism was enabled for the task.
- Only one active dashboard is available for each “Analysis Task”. When a user switches to a different “Analysis Task” the dashboard of that task replaces the active dashboard.

5 Other Functionality

To support all phase operations, several components were developed to manage users and other related data in the platform. Below we provide some examples of those components and illustrate secondary system features that enable the smooth and secure operation of the platform.

Figure 33 - Administration – Manage Users

Figure 34 - Administration – Manage User Groups

Figure 35 - Administration – Manage Organizations

Figure 36 - Manage Data Protection Requests