****

**MASTER THESIS**

|  |
| --- |
| **TITLE:  Project for the development of a Higher Education Management Dashboard in R**  **MASTER DEGREE: Master's degree in Applied Telecommunications and Engineering Management (MASTEAM)**  **AUTHOR: Dimitrios Chatzos**  **ADVISOR: Prof. Vicenç Fernández**  **DATE:** |

|  |
| --- |
| **Title:**  **Author:**  **Advisor:**  **Date:** |

**Abstract**

|  |
| --- |
| Dashboards are information management tools that are used for monitoring and analysing of different organization’s behaviour. The purpose of management Dashboards is to observe, identify and solve problems and to justify patterns with monitoring the key performance indicators.  The Dashboards are not just reporting tools they are interactive communication tools with meaning and purpose. They can analyse big amount of data and return as a result complete storyline of information letting the audience to understand how an organization works.  Thus, this Master Thesis aims to identify and analytically examine the key features, purposes, uses, and benefits of performance dashboards and to develop a Higher Education Management Dashboard in R for the needs of MASTEAM Master of the Polytechnic University ofCatalonia (UPC).  The Thesis is highlighting the methodology and the designing rules for creating dashboards and taking the full potential and benefits of them. Based on them, in the Thesis is implemented a real education management Dashboard in R programing language.  The fundamental rules are used for the implementation of the MASTEAM Master Dashboard. The Dashboard will include the most important features and information about the MASTEAM, which are pointed from the coordinator of the MASTEAM through an interview. The main goal is to give the coordinator the ability to observe and analyse the behaviour of the significant areas of the Master. For instance the enrolment status, student’s information, grades, courses, graduated rates, etc.  Finally, the results of the programing part will demonstrate the last version of the MASTEAM Dashboard web page that can be used from the coordinator. He can monitor the behaviour of the Master and take necessary decisions. According to the visualised data, the maintenance of the good condition of the existing Master is also possible. |

**CONTENTS**

[INTRODUCTION 4](#_Toc527329405)

[CHAPTER 1. Dashboard Definition 3](#_Toc527329406)

[1.1 The guidelines of the dashboard design 4](#_Toc527329407)

[1.2 The fundamentals of dashboard design 5](#_Toc527329408)

[1.2.1 Content 5](#_Toc527329409)

[1.2.1.1 Keep it relevant 5](#_Toc527329410)

[1.2.1.2 Keep it simple 5](#_Toc527329411)

[1.2.1.3 Iconography 6](#_Toc527329412)

[1.2.2 Layout 6](#_Toc527329413)

[1.2.2.1 Group related data together 6](#_Toc527329414)

[1.2.2.2 Keep uniform size 7](#_Toc527329415)

[1.2.2.3 Put limits to your data 7](#_Toc527329416)

[1.2.3 Colours 8](#_Toc527329417)

[1.2.3.1 Leverage contrast 8](#_Toc527329418)

[1.2.3.2 Use colours only for specific reasons 9](#_Toc527329419)

[1.2.3.3 Pay attention to the meaning of colours 10](#_Toc527329420)

[1.2.3.4 Test for colour blindness 10](#_Toc527329421)

[1.2.4 Fonts 11](#_Toc527329422)

[1.2.4.1 Keep standard your font sizes 11](#_Toc527329423)

[1.2.4.2 Combining fonts and colours together 12](#_Toc527329424)

[1.3 Visualization of the Dashboard 12](#_Toc527329425)

[1.4 Visualisation Techniques 14](#_Toc527329426)

[1.4.1 Static visualization 15](#_Toc527329427)

[1.4.2 Interactive visualization 15](#_Toc527329428)

[1.4.3 Drilldowns vs. Drill-Through 15](#_Toc527329429)

[1.4.3.1 Drilldowns 15](#_Toc527329430)

[1.4.3.2 Drill-through 15](#_Toc527329431)

[1.4.3.3 Zooming 15](#_Toc527329432)

[1.5 Validation of the Dashboard 15](#_Toc527329433)

[1.5.1 Build a Wireframe 16](#_Toc527329434)

[1.5.2 Build prototypes 17](#_Toc527329435)

[1.5.3 Regularly tests and upgrades 18](#_Toc527329436)

[1.6 Future work of Dashboard Design 18](#_Toc527329437)

[CHAPTER 2. Analysis of the case MASTEAM 20](#_Toc527329438)

[2.1 Creation of the Interview questions 20](#_Toc527329439)

[2.2 Interview part 22](#_Toc527329440)

[CHAPTER 3. Design a Dashboard 25](#_Toc527329441)

[3.1 Creation of the MASTEAM datasets on CSV format 25](#_Toc527329442)

[3.1.1 Enrolment CSV dataset 26](#_Toc527329443)

[3.1.2 Students information CSV datasets 26](#_Toc527329444)

[3.1.3 Students grades CSV datasets 27](#_Toc527329445)

[3.1.4 Courses and Grades CSV datasets 27](#_Toc527329446)

[3.1.5 Number of students per course CSV dataset 28](#_Toc527329447)

[3.2 Build the dataset Wireframe 29](#_Toc527329448)

[3.2.1 Main page Wireframe 29](#_Toc527329449)

[3.2.2 Students page Wireframe 29](#_Toc527329450)

[3.3.3 Courses page Wireframe 30](#_Toc527329451)

[CHAPTER 4. Implementation of the Dashboard 31](#_Toc527329452)

[4.1 Implementation of the main Dashboard page 31](#_Toc527329453)

[4.1.1 Enrolment View 34](#_Toc527329454)

[4.1.2 Enrolment per Year 34](#_Toc527329455)

[4.2 Implementation of the Students View page 36](#_Toc527329456)

[4.2.1 Students Grades subpage 36](#_Toc527329457)

[4.2.2 Geographical Map subpage 38](#_Toc527329458)

[4.2.3 Students Info subpage 38](#_Toc527329459)

[4.2.4 Students Grades subpage 39](#_Toc527329460)

[4.3 Implementation of the Courses View page 40](#_Toc527329461)

[4.3.1 Courses per Year View subpage 40](#_Toc527329462)

[4.3.1.1 Courses Students per Year 40](#_Toc527329463)

[4.3.1.2 Courses Info 41](#_Toc527329464)

[4.3.2 Courses Individual Info subpage 42](#_Toc527329465)

[4.3.2.1 Number of Students per Course 42](#_Toc527329466)

[4.3.2.2 Fail/Pass and Grades per Course 43](#_Toc527329467)

[4.3.3 Courses Students and Grades Info 44](#_Toc527329468)

[CHAPTER 5. Explanation of R code along with the Maintenance of the MASTEAM Dashboard 45](#_Toc527329469)

[5.1 -R- code explanation 45](#_Toc527329470)

[5.1.1 -R- Libraries explanations 46](#_Toc527329471)

[5.1.2 Manipulation of Datasets 47](#_Toc527329472)

[5.2 Maintenance of the Dashboard 49](#_Toc527329473)

[CONCLUSION 50](#_Toc527329474)

[REFERENCES 50](#_Toc527329475)

**LIST OF FIGURES**

[**Fig. 1.1** Iconography -icons- 6](#_Toc527326999)

[**Fig. 1.2** Layout -Wrong & Correct- 7](#_Toc527327000)

[**Fig. 1.3** Data limits 8](#_Toc527327001)

[**Fig. 1.4** Colour combination 9](#_Toc527327002)

[**Fig. 1.5** Colour combination 10](#_Toc527327003)

[**Fig. 1.6** Colour-blind Dashboard 11](#_Toc527327004)

[**Fig. 1.7** Font size 11](#_Toc527327005)

[**Fig. 1.8** Combination of fonts and backgrounds colours 12](#_Toc527327006)

[**Fig. 1.9** Heat maps 14](#_Toc527327007)

[**Fig. 1.10** Example of Dashboard Wireframe 17](#_Toc527327008)

[**Fig. 1.11** Dashboard prototype 17](#_Toc527327009)

[**Fig. 3.1** Enrolment dataset 26](#_Toc527327010)

[**Fig. 3.2** Students personal info year 2015 26](#_Toc527327011)

[**Fig. 3.3** Students Grades 27](#_Toc527327012)

[**Fig. 3.4** Courses Grades and enrolled students 28](#_Toc527327013)

[**Fig. 3.5** Number of student per course 28](#_Toc527327014)

[**Fig. 4.1** Main MASTEAM pages 31](#_Toc527327015)

[**Fig. 4.2** MASTEAM View Enrolment View 32](#_Toc527327016)

[**Fig. 4.3** MASTEAM View Enrolment per Year 33](#_Toc527327017)

[**Fig. 4.4** Enrolment Stack bar plot 35](#_Toc527327018)

[**Fig. 4.5** Gender bar plot 35](#_Toc527327019)

[**Fig. 4.6** Nationality bar plot 36](#_Toc527327020)

[**Fig. 4.7** Students Grades Box Plot 37](#_Toc527327021)

[**Fig. 4.8** Geographical Map 38](#_Toc527327022)

[**Fig. 4.9** Students Info 39](#_Toc527327023)

[**Fig. 4.10** Search box 39](#_Toc527327024)

[**Fig. 4.11** Students Grades 40](#_Toc527327025)

[**Fig. 4.12** Courses & Students bar plot 41](#_Toc527327026)

[**Fig. 4.13** Stack-Group bar plot 41](#_Toc527327027)

[**Fig. 4.14** Enrolled Students Bar plot 42](#_Toc527327028)

[**Fig. 4.15** Stack Fail/Pass bar plot 43](#_Toc527327029)

[**Fig. 4.16** Courses box plot 44](#_Toc527327030)

[**Fig. 4.17** Courses Data table 44](#_Toc527327031)

[**Fig. 5.1** Edit datasets -library (editData)- 48](#_Toc527327032)

[**Fig. 5.2** Edit datasets -command (editable = True)- 49](#_Toc527327033)

# INTRODUCTION

Nowadays, the big increase of digital data due to Information and Communication Technologies is a universal phenomenon, daily every company or organization is struggling to analyse all those big amount of data and to discover patterns that can help them to improve the productivity and the performance.

Thus, is why the analytics and management Dashboards were created and are needed to enhance decision making from humans, combining the management and the software programing languages. Hence, the development of Dashboards in increased rapidly almost in every organization to provide a quick overview of them.

The definition of Dashboards is to analyse and provide detailed information about the status of an association and to mark the key performance indicators (KPI). Those KPI’s can give the ability to the person who runs an organization to handle huge amount of data that is easier to manipulate them through Dashboards.

Organization they are not only the companies are also the universities that are handling in daily base big amount of data and they are something more than a school. Universities have different departments and people that are working for them from professors, managers, students, until economists. All those areas need somehow to be monitored and controlled from the university.

Thus, the purpose of this Master thesis is to create a higher educational management Dashboard for the needs of the Polytechnic University ofCatalonia (UPC) and more specific for the Master's degree in Applied Telecommunications and Engineering Management (MASTEAM) department.

The Polytechnic University of Catalonia (UPC) is among the greatest universities in the world and one of the best in Spain. UPC is between the 125 best in the world in Computer Science. UPC is also in Telecommunications, 47th globally and 1st in Spain according to the *“*[*National Taiwan University Ranking by subject National Taiwan University Ranking by subject”* and to the](https://www.upc.edu/ranquings/en/upc-at-international-rankings/rankings/national-taiwan-university-ranking-performance-ranking-of-scientific-papers) *“*[*Shanghai Ranking's Global Ranking of Academic Subjects”*](https://www.upc.edu/ranquings/en/upc-at-international-rankings/rankings/shanghairankings-global-ranking-of-academic-subjects).

UPC has several schools in Barcelona and in towns that are near like Castelldefels, Manresa, Sant Cugat de valles, Terrasa and Vilaniva i la Geltrú. All these schools have high education system and high reputation and history in the world.

The MASTEAM Master department belong to Castelldefels School of Telecommunications and Aerospace Engineering (EETAC). The EETAC is a technical school of higher education system of UPC and Barcelona TECH.

The EETAC has a various fields of education programs in the fields of Telecommunications and Aerospace Engineering. The school is famous for his strong commitment with educational innovation and quality for activities that are related with the industrial environment and with main goal to contribute in the Society making it better in several technological fields.

The high education management Dashboard for the MASTEAM case it will be implemented in R programming language and in the end it will provide a functional Dashboard to the coordinator Dr. David Rincon Rivera to monitor and analyse the behaviour of the Master.

To create a successful Dashboard for examining the performance of the MASTEAM is not enough to be just a good programmer, every Dashboard that is implemented need to have a reason and purpose, to be useful and to have the correct interactivity with the end user. Thus, this Thesis is trying to combine that with the management philosophy and to give a different approach to any digital application, because the ideology of the programmer and the ideology of the end user most of the times is different.

The Thesis is divided into 4 different categories the Dashboard’s designing rules and Wireframes, analysis of the MASTEAM case, design and implement the MASTEAM Dashboard and finally the R code with the maintenance of the Dashboard.

The work present all the fundamentals and designing rules to create useful Dashboards and the benefits of them. The main four pylons that need to be carefully consider to create a successful Dashboard are the Content, Layout, Colours and Fonts. Furthermore, the following step is to build the Dashboard’s Wireframes. Thus, are preprogramed Dashboard versions created on a paper and according to them it will be programed the digital view of the Dashboard’s web page.

The MASTEAM case contain an analysis of the Master that is done through an interview with the coordinator Dr. David Rincon Rivera. The interview will explain into details what are the important areas of the Master and what kind of data are needed for the Dashboard. Moreover, those data need to take the correct format and to be adapted inside the R program.

The preprogramed Wireframe versions of the MASTEAM need to be created also with the important datasets inside. According to those Wireframes the programing part it will start. A detailed explanation of every part of the MASTEAM Dashboard it will be provided by figures. The clarification of what is the reason and why everything were programed in the way that they are is also included.

Finally, the most important parts and parameters from the R code will be explained to provide a general idea how the program is working. The explanation of the R code is also important for the maintenance of the Dashboard, for future needs and can provide to the coordinator the ability to include new data inside for the new academicals years or to change the existed ones.

Concluding the last version of the MASTEAM Dashboard is available to this thesis ready for use and is consisted of three main categories the MASTEAM View page, the Students View age and the Courses View page.

# Dashboard Definition

The majority of the people had already use a reporting software and the question is why and for what is the need of a Dashboard at all. Many user can say that dashboards are just another way of representing the information that already is in reports. The truth is that the Dashboards **are not reports** they contain abilities that never can be detected in a normal reporting software.

The majority of reports until now are static without any interaction with the end user. Reports may provide useful information, but on their own, they show only one piece of the puzzle. What is more, many users (especially managers and executives) lack the time to decipher what a report or chart means. Reports do not allow these users to add more pieces of information, dig in to real-time data, or change how information displayed are. Furthermore, reports only include pre-determined datasets.

Dashboards are analytics tools that give users a unified view of the most important data. They consolidate real-time information in a simple, easy-to-understand, and dynamic format. Dashboards are especially useful when you are considering or comparing multiple datasets at time.

Dashboards can be detached according to role that they have and are either **strategic**, **analytical**, **operational**, or **informational**.

Strategic dashboardssupport managers at any level in an organization, and provide the quick overview that decision makers need to monitor the health and opportunities of the business. Dashboards of this type focus on high-level measures of performance, and forecasts. Strategic dashboards benefit from static snapshots of data (daily, weekly, monthly, and quarterly) that are not constantly changing from one moment to the next.

Dashboards for analytical purposesoften include more context, comparisons, and history, along with subtler performance evaluators. Analytical dashboards typically support interactions with the data, such as drilling down into the underlying details.

Dashboards for monitoring operations are often been designed differently from those that support strategic decision-making or data analysis and often require monitoring of activities and events that are constantly changing and might require attention and response at a moment's notice.

However, a dashboard is only as effective as its design and that design should be dictated by the needs of the users. That is why defining the audience is an essential first step in Dashboard design.

Recognize the different use cases in each department as well as the distinct needs of different users across the different organization. Nearly every analytic user will be into one of three categories or personas that are mentioned below:

* **Information consumers:** The majority of analytics users fall in this category. These users prefer to work with a predefined dashboard experience where they can regularly view, interact with, and personalize a preconfigured asset.
* **Content creators:** A creator’s self-service experience is more managed than predefined. These users want the ability to choose the data they need and supplement those existing dashboards and reports with their own metrics.
* **Data analysts:** An analyst’s experience is entirely self-directed. These power users prefer to bring their own data, build their own dashboards and reports, and mine insights they can share with others.

Once is decided what kind of dashboard it will be created according to his role (**strategic**, **analytical**, **operational**, or **informational**) and the users are placed into the three buckets above, the next step is to arrange interviews with each group of users about their requirements.

## 1.1 The guidelines of the dashboard design

After analysing the meaning of the Dashboards, define the audience and the purpose of the Dashboard is the time to create an attractive Dashboard for the user needs.

To create a successful Dashboard, need to be implement some general rules and ideas for the designing part like a guidelines through the designing part. These general ideas are the following:

* **Design for a target:** Simple to say it difficult to do it. The designer need to follow the goal of the Dashboard, and who is the target user otherwise will fill the dashboard with too much useless information.
* **Keep everything at a glance:** The main page of the Dashboard need to give the information directly with any additional clicks or scroll downs to the user.
* **Keep it simple:** Keep the Dashboard design simple and understandable to the end users.
* **Highlight the most relevant information:** The dashboard is like a page of a magazine, each location has its meaning and a different level of importance. Information and charts in random places are simple wrong.
* **Be clear:** The use of acronyms are bad. The use of legends is better.
* **Start from zero:** Chart axes must be used consciously. Every chart need numbers and limits.
* **Shorten the numbers:** Dashboards users want to see the overall picture.
* **Show the context:** Numbers carry their meaning only within their context.
* **Choose the right colours:** About 10% of the population suffers from colour blindness. Correct use of colours is important.
* **Design dashboards not reports:** Not all the details from a table are suitable for a dashboard.
* **Show variations:** Make the Dashboard able to make the math do not let users do the math.
* **Pick the right chart:** Each piece of information must be displayed using the right chart in the dashboard, the correct visualization technic need to be chosen correctly.

Each piece of the information and data for the Dashboard need to apply these general ideas.

## 1.2 The fundamentals of dashboard design

After defining the general rules and guidelines that need to follow, they are four important fundamentals of Dashboard design **content**, **layout**, **colour**, and **fonts** that the designer need to know. By manipulating each of these fundamentals, the Dashboard can be successful and useful to the user.

### 1.2.1 Content

The content is the general view of the Dashboard how it looks and is giving emphasis to the important information. The correct content need to follow some general rules like to keep the content relevant, to keep it simple and to use iconography.

#### 1.2.1.1 Keep it relevant

Already mentioned before that the content should be relevant to the defined audience. Part of accomplishing that is removing everything else. Simplify content and reduce visual elements to only the most critical pieces.

For example in a car the Speed meter that is the most important thing inside the car is huger and focused than all the other instruments. Finally, every Dashboard needs to focus in the important things and to give priority to the audience needs without any useless details than can confuse the audience.

#### 1.2.1.2 Keep it simple

Data visualisation is one of the most important designing part on the Dashboards. The use of visualisation techniques is important but do not overload the Dashboard. Visualization techniques can be charts, plots and graphs.

Visualization is at the same time good and bad for the reason that if the Dashboard is overloaded as was mentioned before can confuse the audience. The designing part must focus only to the main data that need to be visualized.

Visual effects like background gradients, shadows, and 3D elements are not good to be part of the Dashboard. The correct colours and elements need to be created to be able to give emphasis to the parts that the user need to focus.

#### 1.2.1.3 Iconography

The use of iconography is an important design part for the reason that the Dashboard content is limited to text and charts so need to contain the correct icons that will help the audience to navigate their self.

Iconography is important for both stylish and development reasons and icons like close, delete, eject, search and hide will make the Dashboard more useful and efficient for user experience.

For the iconography part, are two options, the first option to use an already existed icon font pack or the second option is to create custom images. The first option is more easy and fast for the Dashboard design the second option is time consuming and need drawing techniques.

The below figure is giving an example of some iconography icons.



**Fig. 1.1** Iconography -icons-

### 1.2.2 Layout

The layout of the Dashboard is sawing how the data are placed together or not, if they are in the correct order and size, and if they fill correct the entire Dashboard page without useless empty spaces. The layout need to follow some general rules that will be explained in details below.

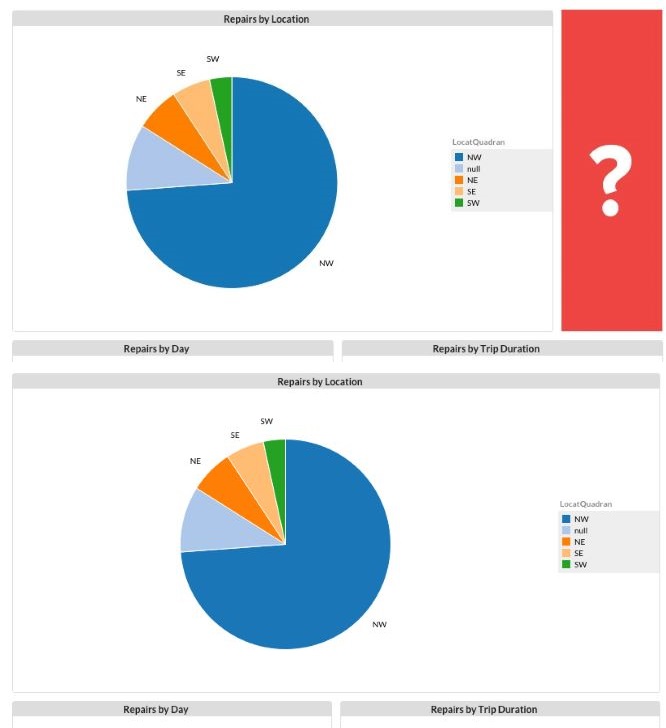
#### 1.2.2.1 Group related data together

Once all the data that we what to use for our Dashboard are gathered need to be grouped together and to be related to each other. The data need to be visualized with the right size and in the right position to be able to highlight the most important parts of the Dashboard.

#### 1.2.2.2 Keep uniform size

The size of the icons or the graphs is important to be the same and similar to reduce the distractions and to be able to identify all the features in the screen. This method helps to make the design easier and geometrically correct. Different visualizations techniques can be created and placed together on the same screen, this can give different points of view to the dataset that need to be visualised.

Furthermore, the figure below is giving an example of a Dashboard page with the incorrect and correct layout.

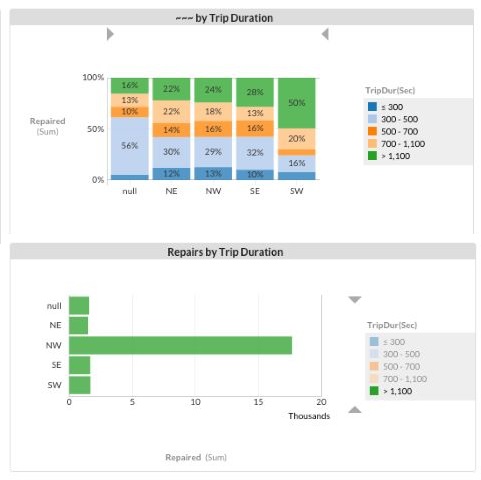


**Fig. 1.2** Layout -Wrong & Correct-

#### 1.2.2.3 Put limits to your data

One other parameter that need to be considered is to put reasonable limits to the data that will include on the Dashboard because the big amount of data may affect opposite and to distract the user.

The data need to visualised correct and to be understandable from the users as the below figure is sawing.



**Fig. 1.3** Data limits

### 1.2.3 Colours

One of the most important things to create a good Dashboard is the colours. Many users cannot understand the importance of choosing the correct colours and is the first mistake that are doing when they are designing a Dashboard. Below are some colour parameters that the user need to take in advance

#### 1.2.3.1 Leverage contrast

The correct and good choice of colours can be a good start for the Dashboard to make the data that need to be visualized clear to the audience.

Different colours combination need to be visualised to make the data easier for distinguishing the difference. The combination of the colours is important and need to be chosen carefully.

#### 1.2.3.2 Use colours only for specific reasons

The colours in the Dashboard need to be used only to serve a specific reason, for example, when is needed to highlight something and to put the attention of the audience to that specific panel of information.

In the other hand, colours can be used also when a group of data need to visualised together on the same chart and need be to separate from each other. From the different colours is easy to understand the different data information that are placed together.

The figure below is sawing a chart with different information in the same table with different colours.



**Fig. 1.4** Colour combination

A good method for choosing colours is to have not more than six different colours in the visualization part, because the use of many different colours will make difficult for the users to understand the difference between the meanings of the datasets.

The use of natural colours is good to distinguish the information and the use of bright or dark colours for highlighting the important things.

When the colours are used, the programmer need to be aware that the background colour is the correct also to make the other colours visible to the user to understand them.

Important here is to remember always that as long as the question ***“What purpose does this colour serve, and will it serve it effectively?”*** can be answered then the use of colour is been chosen correctly.

#### 1.2.3.3 Pay attention to the meaning of colours

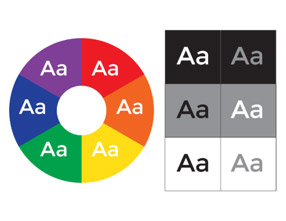
Important information about the colours is that they are not only colours but they have also a background meaning.

For instance, when the colours red and green will be used most of the people will understand that red is colour to highlight something bad or to put a warning and the green colour to highlight something good. Another example is with red and blue colours that anyone can associate them with hot (red) and cold (blue).

The use of the colours needs to be with a meaning for example the use of two different colours means that are two different information or for highlighting something. Never use different colour to represent the same thing.

Simple colours like black and white is good to be used for the Dashboard because sometimes the simplest is also and the best way to represent something.

The Figure below is representing some colour combinations with white and black letters.



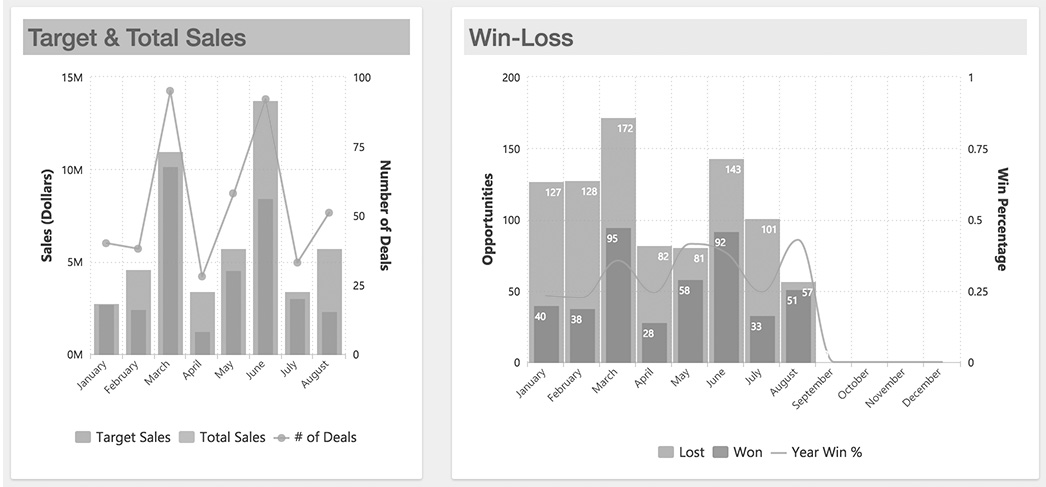
**Fig. 1.5** Colour combination

#### 1.2.3.4 Test for colour blindness

Another important thing that need to be taken in account to choose the colours for the Dashboard is for the users that have colour blindness, they cannot recognise well the different colours, like red and green that are almost the same for them.

Statistics shows that one in twenty users is colour-blind. The colour blindness problem can be solved thanks to some simulator programs like Colour Brewer, Vischeck e.t.c. Those programs can make the Dashboard visible to all the users and all of them to understand the differences between the colours.

Furthermore, the below Figure is an example of a Dashboard for colour-blind users.



**Fig. 1.6** Colour-blind Dashboard

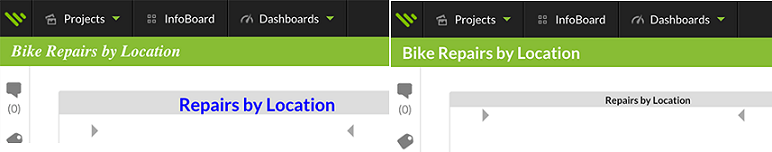
### 1.2.4 Fonts

After the colour part, another important thing for the design is the fonts of the letters that will be used. Most of the times the general rule for the Font size is, 12 for Times New Roman and 10 for Arial. Keep in mind than all the fonts must be maximum three times the size of the original Font when is needed to write something with bigger font.

#### 1.2.4.1 Keep standard your font sizes

The goal is to design a Dashboard with homogenous environment that will be easy to give emphasis to the words and let the users to understand it.

An example of the correct use of the fonts is given on the below figure.



**Fig. 1.7** Font size

There are two main categories of fonts the Serif and the Sans-Serif families. Both of them have a different purpose in the designing part.

For instance, Serif font is suitable for the main the need to be included because is easier for the user to read it. On the other hand, Sans-serif fonts are more suitable for short texts, like headlines or words that need to give emphasis because the Sans-serif font attract the attention of the user easier than the Serif Font.

The designer need to determine also how will be the space between the lines and to determine how will be the space between the words or the individual letters. Those simple details can make the text better or not.

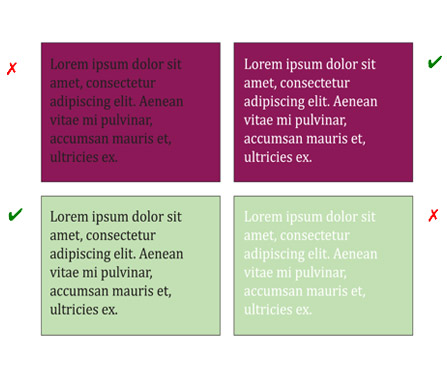
Most of the times the users tend to see first what is bigger in the screen without paying attention on the font, that is the reason that the headlines must be always bigger that the other text.

#### 1.2.4.2 Combining fonts and colours together

The final part is to combine the fonts and the colours together in the way that both of them can be combined well together. The colours can give a different point of view on the text, sometimes good and sometimes bad.

For instance, do not design dark text together with dark background or white text with a bright background. Another tip is not to use bright colours for the font with bright background, like yellow font and green background. This will confuse the users when they try to read something and probably will make the eye vision difficult.

The figure below is sawing some colour combinations between fonts and background



**Fig. 1.8** Combination of fonts and backgrounds colours

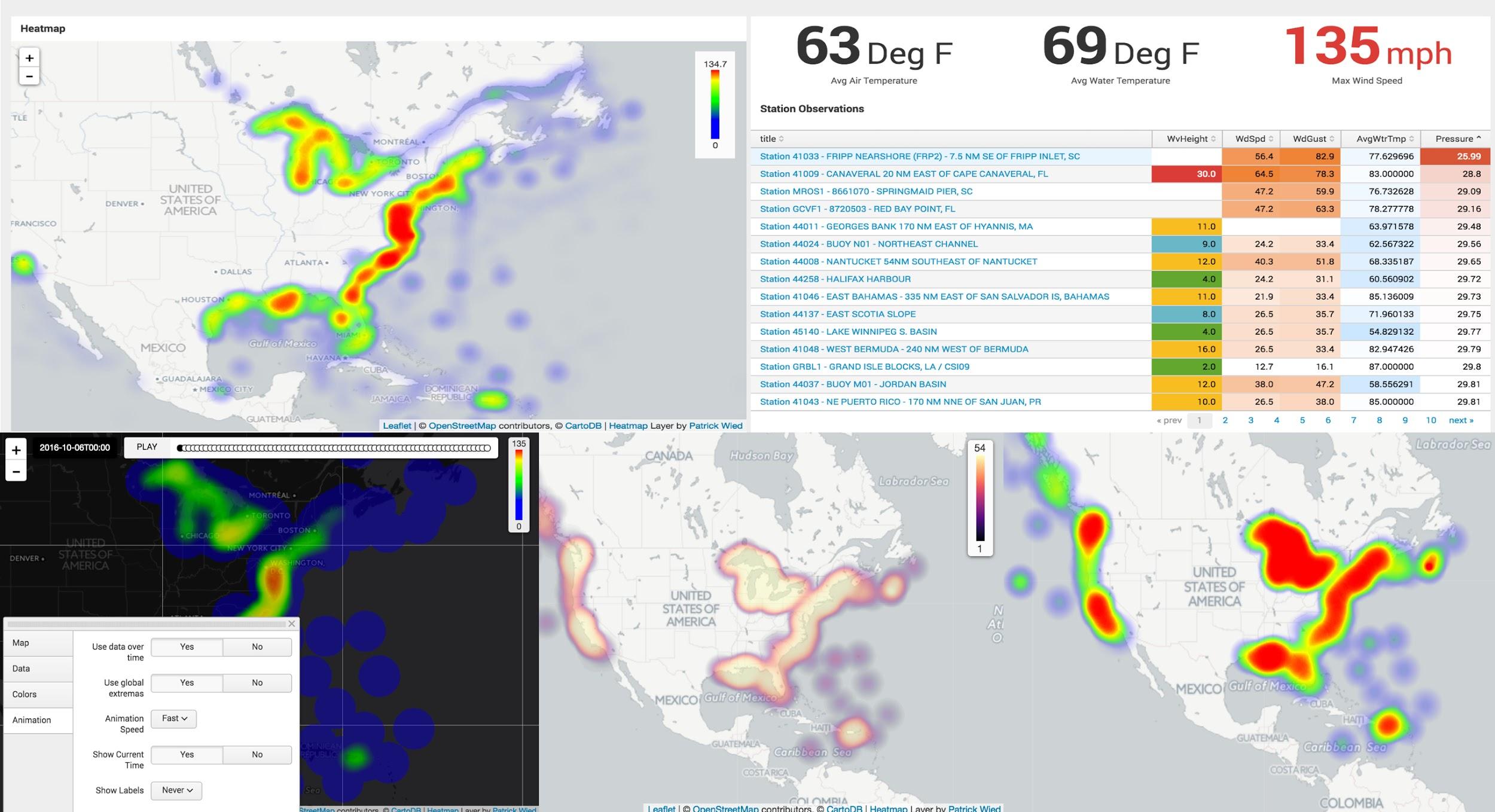
## 1.3 Visualization of the Dashboard

The part of visualization is the part that will be decide how will visualize the datasets, for instance with toolbars charts, pies, etc. On this part, it will be a combination between the information that had been gathered from the interview part with the visualization techniques that will be used to visualise that information.

The way that the designer it will represent the data is the key to visualize big amount of datasets together and let the users to analyse them easy and without any problems.

Furthermore, they are some main visualization types and charts that can be used for the designing part. Those visualization types are the bellow ones:

* **Tabular format:** This format is one of the best when we want to visualize exact amount of numbers because we can represent them in rows and columns and they can analyse any kind of information and the summary of them. Unfortunately, this format is not so good if we want to analyse and compare different data together because the Tabular format it will be very complicated and difficult for the user to analyse it.
* **Line charts:** This format is good when we want to visualize continuous data over time and to see the evolution of them. Line charts are very useful when we want to analyse our data, to observe them with different parameters, and to see if our goal is achieved or not.
* **Bar charts:** This format is the best when we want to analyse different data categories together in the same chart. Bar charts can be visualized horizontally or vertically and they fit very well when we have limited space on our Dashboard because of the reason that we can combine many data in the same chart. Most of the time one of the axis is the different data categories that we have and the other axis is the numerical data that they have for comparison.
* **Pie charts:** This format is the best when we want to compare different data together in the 100% and make the different parts of the charts important or not so important according to the percentage that they have from the total 100%. We need to keep in mind that until five different categories in the chart is acceptable after five categories the Pie will be more complicated and difficult from the user to understand it.
* **Area Charts**: They are like the line charts but with more information available inside.
* **Bubble Charts:** This format is to visualize the datasets in the three dimensions.
* **Funnel charts**: Are a type of chart, often used to represent stages in a sales process] and show the amount of potential revenue for each stage. This type of chart can also be useful in identifying potential problem areas in an organization’s sales processes. A funnel chart is similar to a stacked percent bar chart.
* **Heat Maps:** This format is suitable for showing a geographical area and the information about this area.



**Fig. 1.9** Heat maps

* **Radar chart:** This chart is a method of displaying multivariate data in the form of a two-dimensional chart of three or more quantitative variables represented on axes starting from the same point. The relative position and angle of the axes is typically uninformative. Alternative names include polar chart, web chart, spider chart, and star chart.
* **Pyramid Charts:** Are ideal for showing comparisons of data, using the thickness of layers to denote relative values.
* **Scatter Charts:**This chart is a type of a chart or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.
* **Sparkline Charts:** The Sparkline is a very small line chart, typically drawn without axes or coordinates. It presents the general shape of the variation (typically over time) in some measurement, such as temperature or stock market price, in a simple and highly condensed way.
* **Whisker Charts (Box Plots)**: This chart is a method for graphically depicting groups of numerical data through their quartiles. Box plots may also have lines extending vertically from the boxes (*whiskers*) indicating variability outside the upper and lower quartiles, hence the terms **box-and-whisker plot** and **box-and-whisker diagram**.

## 1.4 Visualisation Techniques

Let continue one-step further for the visualization part and decide how the Dashboard need to look. There are three visualisation categories the Static, the Animated and the combination of both of them together.

### 1.4.1 Static visualization

The Static visualization highlight only in the important information and keep the Dashboard static that means that the user is not able to do any additional actions because everything is on the main page. Static visualization can be a text or an image.

This visualization technique do not need any live browser or any add-ons because the data are designed not to change.

### 1.4.2 Interactive visualization

Interactive visualization give the option to the users to interact with the Dashboard. The Dashboard now looks animated with extra options like drill down and drill-through to explore all the information that are available.

For this visualization technique, the data can be also live-stream and to change through time. This visualization live-stream technique need some advance options like Flash player or HTML5.

### 1.4.3 Drilldowns vs. Drill-Through

As has been mentioned before the interactive Dashboard have some other options to achieve that like drill down, drill-through and zoom out/zoom in.

#### 1.4.3.1 Drilldowns

Drilldown technique means to be able to move from general data to more detail data, focusing on the things that the user want. The user is able here to search through deeper layers and to click to information in order to reveal more details.

#### 1.4.3.2 Drill-through

Drill-thought allow the user to jump from one report to another with the specific ability to focus on specific data. The users can right-click on the data that they want and to Drill Through the focused pages to get more detail that are available.

#### 1.4.3.3 Zooming

Zooming is the ability to dig deeper on the visualization part letting the user to click and see more details. For example, a map that the user can zoom into geographical places for more details, or a timeline chart with the ability to zoom and focus in a specific year and time for more details.

## 1.5 Validation of the Dashboard

The last and the most important step to finish the Dashboard is the validation part. The purpose of this step is to discover what the user’s opinion is for the Dashboard, because the user may have different opinion than the designer for the general view of the Dashboard.

Important for this part is the designer to star designing some beta versions of the Dashboard with the main ideas and datasets. This part is sawing in action how the combination between the main ideas and the datasets is together.

The last information that the designer need for the final design is to know how the Dashboard will interact with the user and if is good or not. The information it will come through the beta Dashboards that the designer will create and give to random users to evaluate it.

The validation part is especially important but need the correct use. Changing the Dashboard many times through the beta versions may affect the main idea of the Dashboard and destroy all the work that is done until now.

Through the evaluation of the Dashboard betas, the users can give some feedbacks for the design. The user experience is the most important parameter to understand how useful the Dashboard is or not.

Letting the users to be interacted with the Dashboard beta in daily base can make the designing part easier because through that the designer can ask the users opinion for how to make the Dashboard better and more useful.

The users can answer to some basic questions like the following one. “How they prefer the data to be visualized”, or “what kind of data and information are better”, or “what is important for them to understand the datasets”?

After the questions, the designer can also ask the users to draw their basic ideas if they want what, or to give some small examples about how they think that the Dashboard can be better.

During this procedure, the design of the Dashboard can be improved and the designer can solve in advance problems that has not thought before the final Dashboard design.

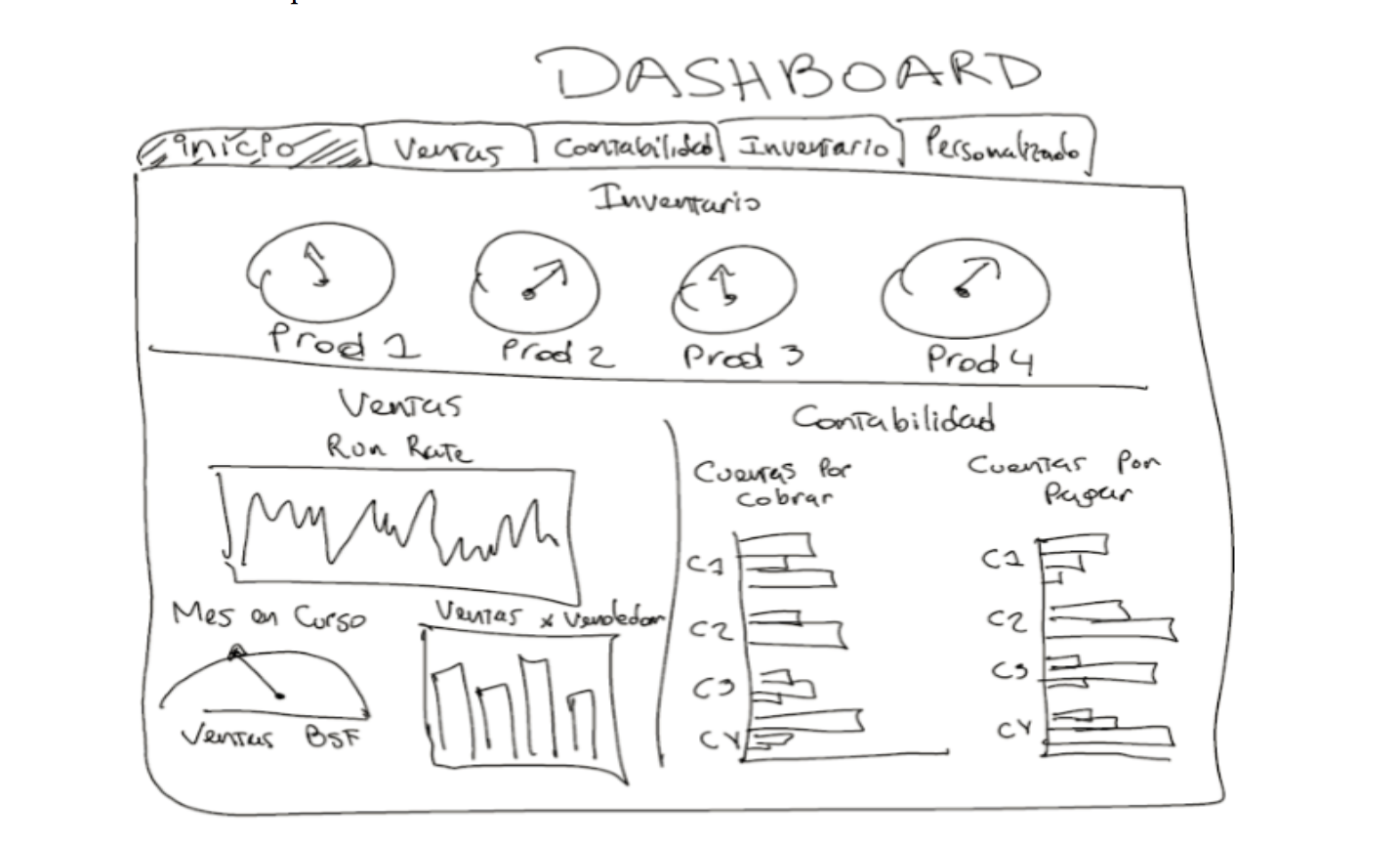
### 1.5.1 Build a Wireframe

Important step for the validation of the Dashboard is to create those beta versions of the Dashboard. The first step of those beta versions is called wireframe. Wireframe is nothing more than a basic design of the dashboard with all the datasets together to saw how they look when will be visualised

The wireframe more or less is the draft version of the Dashboard and looks more like a painting than an application, most of the times the main ideas can be designed first on a paper and then to proceed to the programing part.

The wireframe is sawing to the designer more or less how he will distribute the space that he has and how he will put the charts the buttons the images and the data all together. Knowing from the beginning where to put on the Dashboard all these charts or buttons or images etc. can save precious time from the programming part.

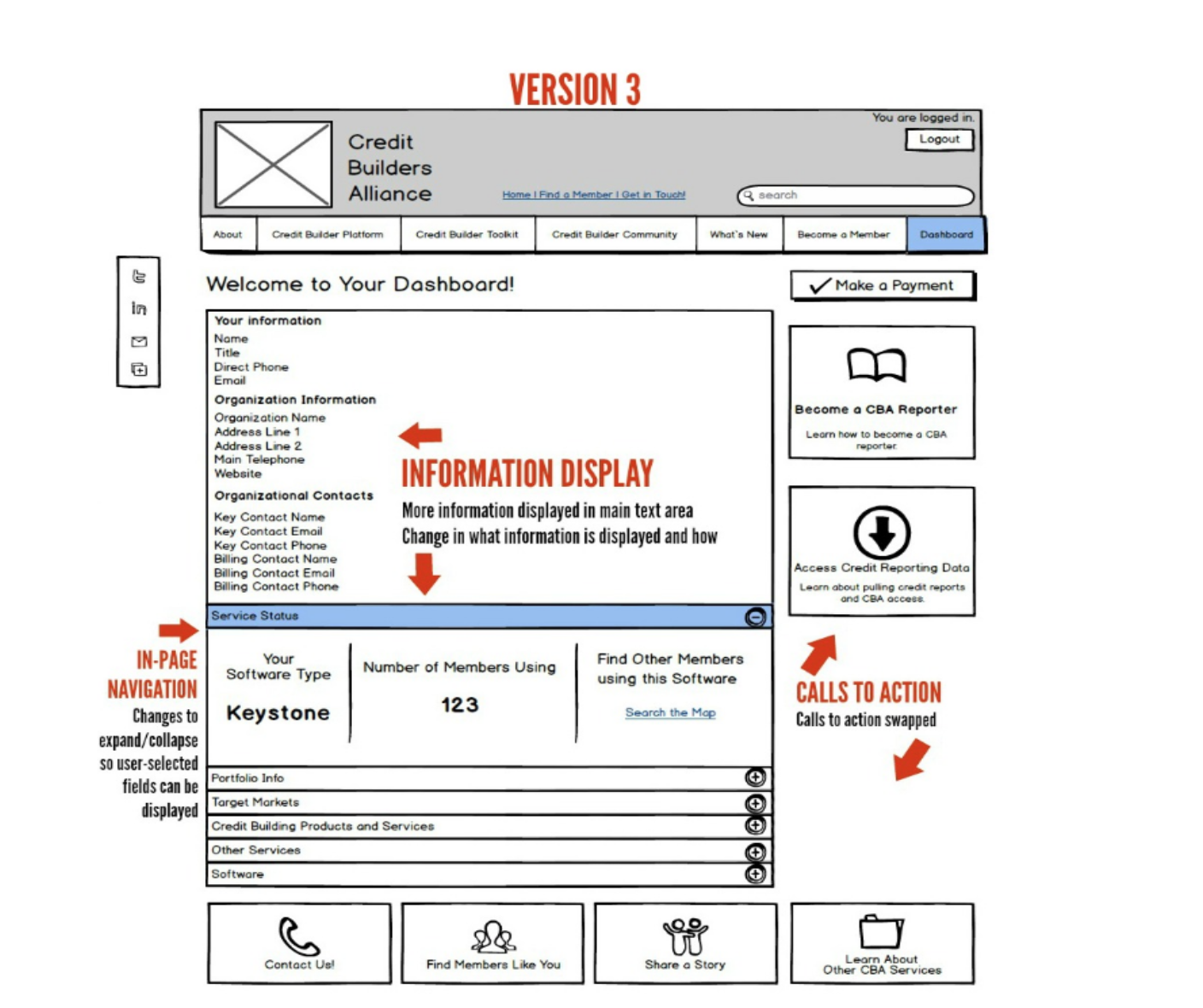
Furthermore the below figure is an example of a possible Dashboard wireframe.



**Fig. 1.10** Example of Dashboard Wireframe

### 1.5.2 Build prototypes

After the wireframe done, time now is to build the first prototype of the final view of the Dashboard, but not the final version of it. Below is an example of a prototype.



**Fig. 1.11** Dashboard prototype

As the ***Fig. 1.11*** is showing This will not going to be the final version of the Dashboard because can exist more than one versions, the design can be developed more and more over time and over the prototypes with improvements until to obtain the desirable design.

The feedbacks from the users is the important parameter here to know until when the Dashboard need to be improved. When the users star to give more positive feedbacks than negative ones then the improvement part is almost done is almost impossible the Dashboard to be for all the users perfect so is normal to have and some negative feedbacks.

### 1.5.3 Regularly tests and upgrades

Finally and important is when the design is finished and the final Dashboard we need still to not stop searching for what we can do better. We need to keep testing our Dashboard and still to take in account all the feedbacks from the entire user. We need to remember that everything needs to be evolved over time and to have new updated versions of our Dashboard in the future.

## 1.6 Future work of Dashboard Design

Everything that is evolved with the informatics and technology need to keep evolving through time, the same is with the Dashboards. Is important for every Dashboard to be the same useful like the first day until the end.

Below are some important things for the future of Dashboard design that designers need to do for future Dashboard designing part:

* **Simpler Interfaces:** Simpler interfaces means that the designer need to create a successful interface with a good user experience easy to understand and easy to navigate. No more big letters without a purpose, no more many colours, no more three-dimension treatments that can distract the user. Remember that simpler does not mean that just put text and that is all, simpler mean to provide an easy and useful user experience.
* **Needs of the On-the-Go User (Mobility):** As technological evolution keep growing so much the users are more and more attached to their phones and tables so that means that the Dashboard need to provide them good mobility experience.
* **Colour Balance:** As was mentioned before no more many colours without meaning. There is often problem in some Dashboards to be either too much colourful or to less. The design needs to be “Colour Balanced” with the meaning of monochromatic background colours in contrast with bright colours when something need to be highlighted on the datasets.
* **Iconography**: Less is more, is the main rule of the future Dashboards and Iconography is a part of that. Designers need to create more often Dashboards with small icons or photos with short descriptions to let the user to understand directly what they want to see without reading the content. Some “smart” photos can give an efficient and quicker navigation interface.
* **Dynamic Dashboards:** Dynamic Dashboards is the future of Dashboards for the reason that any dataset need to be able to be upgraded and interact with the Dashboard at any time. The designer need to be able to customize the Dashboard at any time and to provide new things to the users as the technology and the Internet evolves. Without a Dynamic Dashboard, it will impossible for the designer to upgrade the Dashboard for future needs.
* **Interactive and live data:** Interactive Dashboards are needed in our days for the reason that the users are more technological efficient than in the past and it will be more interesting for them to interact with the data and to have more options to analyse them. All the techniques that was analysed before like zooming, drilldowns and drill through plus elements like videos need to be includes in the design. The most important thing now is the Data to be live, users are more interested to see what is happening at the moment that they are searching for something than to have only past datasets.
* **Dashboards with High Location Technology:** In the past, the traditional computer Dashboards were using static maps for analysing different areas when the user wanted. Now with the mobility evolution the Dashboards need to be more efficient and with the use of mobile Gps to be able to analyse datasets that they are in the same geographical area automatically.
* **Future prediction on Dashboards:** Until now, the user is used to see on internet weather forecasts, election forecasts or in mobile phones possible text prediction according to the text history. The next big thing on the Dashboards is the data analytic predictions. Analysing information from different datasets can be useful to predict patterns and facts that are inside the data and to help the users or the companies to organise their plans better.

# Analysis of the case MASTEAM

One of the most important things for the creation of a successful Dashboard is to know in advance for what reason and why we are designing something. This will give us an advantage to create a Dashboard with the most relevant information that the user want to observe and analyse.

To know in advance that important information is needed to know how the user is thinking and what kind of information want to see on the Dashboard. To achieve that and to gather those information some interviews between the user and the programmer need to be arranged as was been explained already on the **CHAPTER 1**.

For the purpose of the MASTEAM Dashboard, an interview with the coordinator of the MASTEAM Master Mr. Rincon Rivera David was been arranged. This interview can help us to know in advance and gather that information that are important for him to create a successful Dashboard.

## 2.1 Creation of the Interview questions

After defining the user or users that we need for the interview part (Mr. David Rivera Rincon in our case), the next step is to prepare a list with possible interview questions. Those questions will refer to the MASTEAM topics and will have an explanation why are important for the designing part according to us.

Making correct or wrong questions is the same valuable, because from that it will be more understandable what is important for the Dashboard through the correct questions and what is not through the wrong questions.

The main idea for how the Dashboard need to be is most of the times different for the prospect of the designer and from the prospect of the user. Most of the times the designer probably does not know how important or not is what is designing or what is the purpose of putting a specific information on the Dashboard.

For the purpose of the MASTEAM interview, was been created a list with possible questions with an explanation to each of them why they are important to be included on the Dashboard. Following are the interview questions for the MASTEAM case.

* **How much are the Tuition Fees:** Is important to know how much the Master cost, or what are the other economical information that the Master include.
* **Enrolment status:** Maybe the most important thing that the user want to see on an Educational Dashboard because can see any time the current status of all the Master, students and enrolment applications.
* **How many are the courses on the MASTEAM and which are they:** Like in any education institute one of the most important thing that the user what to know is the courses that are offered.
* **The total grades of the courses:** Is important for the evaluation of any subject to know how the grades of any student are.
* **The percentage of the graduated student:** Important parameter for the evaluation of the master in total every year and how successful the Master is.
* **Information about the students like age, gender, nationality:** Any information about the students of a master are important, before the enrolment and during the Master for the evaluation system.
* **How many are the new students in each semester:** Important information because the master is divided in 2 periods and they are new students every autumn and spring semester.
* **How many internships the University is providing and what kind:** Useful information to know any time what are the internships that the MASTEA provide and to inform anyone that is interested.
* **How many students are going after the MASTEAM for a PHD:** Maybe important information to know what is the percentage of the students that are going for PHD?
* **How many are the professors that the master have and how many of them with a PHD degree:** Possible information from someone that is interested for the status of the professors of the master.
* **The professors that are involved in the research projects and the students that are participating:** Possible useful information about the professors and the students that are involved on university projects to know for evaluation purposes.
* **The range of the professor’s salary:** Possible useful information about the economic status of the master.
* **What are the external activities that the MASTEAM offer:** Useful information to inform the students about any activities that the master offer like seminars, workshops, etc.?
* **Motilities and Transportations about the students:** Like any Master degree they are students that are coming from other Universities for a double degree or for Erasmus or the opposite students of the MASTEAM are going to other Universities for the same purposes.
* **What is the Evaluation system of the MASTEAM:** Any information that can be included in the Dashboard about evaluation purposes is important.

The creation of the interview questions is a part that most of the times is individual and different for every designer that want to design a Dashboard. A successful formula for specific questions or instructions that the designer can follow does not exit, every Dashboard is different and with different purpose.

Every possible question that can be created for the interview part is totally individual and based on what the designer is thinking that is important for the Dashboard or not. All the real important parts of the Dashboard they will come through and after the interview part.

Furthermore we need to keep in mind that every possible question that the designer think that is important for him or not, maybe it will not going to be the same important for the user that is involved in the interview. Is possible to have similar thoughts or very different ones.

As was been mentioned before does not matter if the questions are correct or not, the purpose of the interview is to learn what are the most relevant information for the Dashboard and through them to evolve what is already planned.

During the interview part, not only the designer can be effected positive but also the user. The designer knows better what can be visualised or not, and can explain to the user through the interview, that every possible dataset can be visualised in ways that the user may does not thought before.

## 2.2 Interview part

After the end of the interview with Mr. David Rivera Rincon, many useful things were mentioned about the Master and for the important information that need to be included on the Dashboard, or which information does not need to be included.

The interview was very useful for understanding how an educational Dashboard needs to be, what the purpose of every information that is inside is and how need to be visualised.

The first mistake according to the questions that was made for the MASTEAM, was that all the questions were made from the perception of what a student want to see on the Dashboard and not the coordinator.

Let us now to take one by one all the questions that were asked to Mr. David Rivera Rincon and to point what were the most important information according to him that need to be include on the Dashboard.

According to the interview only the bellow information are very useful and important to be included on the Dashboard:

* **Enrolment:** The most important thing that need to be on the Dashboard because include all the information about the students, the old one and the new one that enrolled to study on the MASTEAM. Also, include the information about the percentage off the accepted applications and the rejected one.
* **The Courses:** The number of the courses, the information about them and the edition of the course according to the year and the semester.
* **Number of students:** The number of students that the MASTEAM has in total and in each course individual. Also important here is the number of new students that are enrolled in each semester.
* **Grades:** The grades of the students in average and in each course individual. Also the total average of the grade on each course individual.
* **Percentage of Graduated students:** Here is important to know the percentage of the graduated student, when they start and when they finish. Here need to be included also the percentage of the students that they failed to Graduate or they quit.
* **Information about the students:** Information that are important about the students are age, gender and the geographical region if they are European citizens or not. Important for this part is also the information about the universities that the students were studying before, for instance if they had studied on UPC or is a different Spanish university or studied on abroad universities.
* **Information about Internships:** The information that are important for this part is the students that were evolved in the internships and the evaluation that they have, also a list of the top companies that a student can do an internship.
* **Information about Motilities:** Important for this part, is the information about the students that are coming to study on MASTEAM which are from ERASMUS or from double degree programs and the opposite students from MASTEAM to go abroad. Important to know here are the number of the students, the grades on each lesson and the average grade in total.
* **The evaluation System:** Final but not least the Dashboard need to provide an evaluation system that every university has and need for each Master individual, this can evaluate all the activities that the MASTEAM provide and give a total score that is important for the Government.

After the end of the interview part, is clearer now what kind of information need to be included on the MASTEAM Dashboard and how they should be visualised. One important issue that exist for all the personal information is the human privacy and if someone is allowed to have those information or not.

Additionally, the possibility some of the data that are needed for the Dashboard to be accessible to gather them and to include them it will be the same with the data that will be closed and private due to data privacy.

Furthermore, another issue that may exist is that, if someone want to have access to datasets that contain information about the UPC University needs to be also an authorized user from the UPC directory.

This may be a problem for the programmer but not for the coordinator of the Master because he is an authorized user and he has the wright to have access to that kind of information.

Finally, after all those useful information that are learned during the interview part and also taking into account all the issues about the data privacy, we need to decide what datasets we will use for the MASTEAM Dashboard. Those datasets they will contain the most important information about the MASTEAM according to the interview part and from thus important data only the data that will be able to gather and to have the authority to use them.

# Design a Dashboard

The most important thing for the creation of a Dashboard is to know for what kind of data will be the designing part and what the purpose those data. Now that the interview part is finished, all those important parameters about the data are known and the creation of the MASTEAM Dashboard can start. Moreover, a list of the final datasets that will be included to the Dashboard need to be created, together with the way that those datasets will be visualized through the MASTEAM Dashboard.

Due to the data privacy issue is preferred the creation and the use of non-real datasets that they will contain fake but relevant information about the MASTEAM case. The creation of those possible non-real dataset examples thy will be explained below.

The datasets they will be on CSV format and they will include information about (i) the enrolment, (ii) all the information about the students including new and old students and (iii) all the information about the courses and the grades. Moreover, the created datasets were separated for the academicals years 2015, 2016 and 2017.

Additionally, need to mention here that the Dashboard and the program can manipulate any kind of information and datasets, not only those datasets that they will created for on the MASTEAM case. Thus, fake datasets are only an example which through that a real data can be implemented.

All the above information will be included on the Dashboard as individual categories and as combination of all the datasets together. The goal is to provide the general view of the MASTEAM only in one screen and latter to analyse the data into detail in different pages.

The main idea is to create a first Dashboard page that will contain all the important data about the MASTEAM combined in one page, this approach can show to the coordinator from the first moment the general view of the MASTEAM Master.

After the creation of the first Dashboard page, the next step is to continue with the creation of the other pages separating the Master into different categories and to examine them with further details. For instance one page with all the data about the students, one page about the grades or one page about the MASTEAM courses.

## 3.1 Creation of the MASTEAM datasets on CSV format

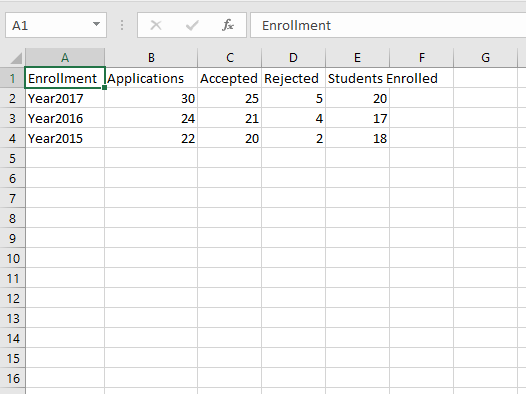
The creation of the Dashboard need now those non-real datasets that were mentioned before. Those datasets they will be on Excel CSV format and they will include data that are important according to the interview part but not the real one. The datasets are the following ones, Enrolment CSV dataset, Students Information CSV dataset, Students Grades CSV, Courses and Grades CSV datasets and Number of students per course CSV dataset.

### 3.1.1 Enrolment CSV dataset

The first created dataset was the enrolment CSV file that contains information about the students who enrolled the MASTEAM in years 2015, 2016 and 2017.

The enrolment CSV file contain the total number of applications that the interested students applied for the MASTEAM, the number of the applications that had been accepted, the number of the rejected one and from the accepted applications the number of the student that finally enrolled to the Master.

Furthermore, the below figure ***Fig. 3.1*** is an example of a possible CSV file with the enrolment information.

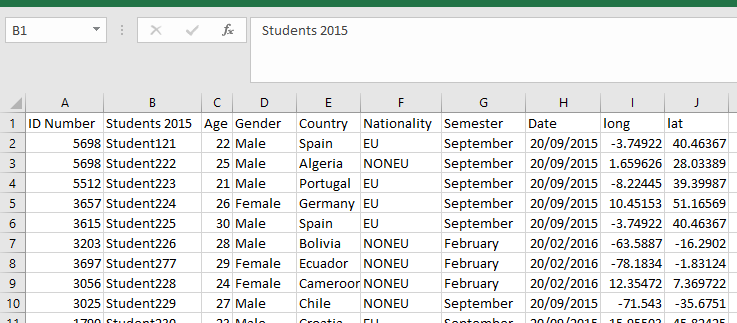


**Fig. 3.1** Enrolment dataset

### 3.1.2 Students information CSV datasets

The second created CSV datasets were about the personal information of each student of the MASTEAM for the years 2015, 2016 and 2017. Those CSV datasets contain the information about the current id number that the University is providing to them, the names of the students, the age, the gender, their countries with the geographical coordinates, the semester and the date that were enrolled for the Master and their nationalities (European or Non-European).

Furthermore, the below figure ***Fig. 3.2*** is showing an example of a possible CSV file that contain the student information for the academically year 2015.

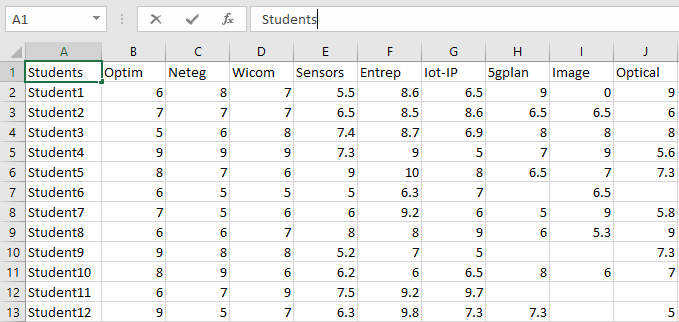


**Fig. 3.2** Students personal info year 2015

### 3.1.3 Students grades CSV datasets

The third created CSV datasets were about the individual grades of each student for all the courses that enrolled for the years 2015, 2016 and 2017. The CSV files contain the names of the students and the grades that obtained in all the courses that had enrolled.

Furthermore, the below figure ***Fig. 3.3*** is an example of a possible CSV file that contain the student grades from the year 2017.



**Fig. 3.3** Students Grades

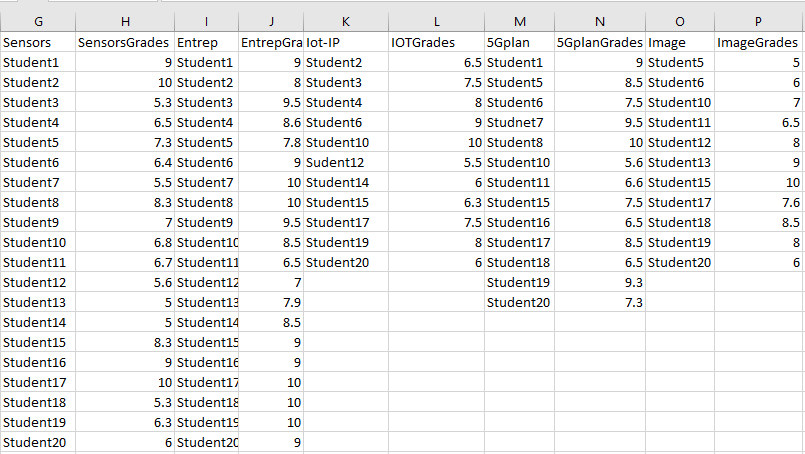
From the figure ***Fig. 3.3*** *Students Grades*, it can observed be that some squares are empty without any grade, which is because the students were not enrolled on these courses.

### 3.1.4 Courses and Grades CSV datasets

The fourth created CSV datasets were about the courses, the enrolled students per course and the grades that obtained for the years 2015, 2016 and 2017.

These datasets are similar to the ***3.1.3 Students grades CSV datasets***, but with the difference that, are showing information from course to course and not from student to student.

Furthermore, the below figure ***Fig. 3.4*** is an example of a possible CSV file that contain the courses with the enrolled students per course and the grades that obtained.

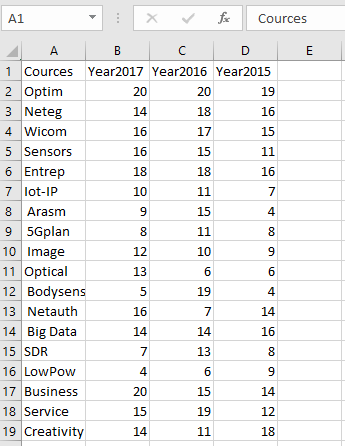


**Fig. 3.4** Courses Grades and enrolled students

### 3.1.5 Number of students per course CSV dataset

The fifth created dataset was about the number of students that each course had for the years 2015, 2016 and 2017.

Furthermore, the below figure ***Fig. 3.5*** is giving a possible CSV file that contain the number of students per course for the years 2015, 2016 and 2017.



**Fig. 3.5** Number of student per course

## 3.2 Build the dataset Wireframe

Now that the created datasets are ready for use, the next step is to create the Wireframes. The Wireframe is the procedure before the programming part, which the main ideas about how the visualization of the datasets need to be on the Dashboard, is going to be created on a piece of paper.

The paper or papers with the draws of the Dashboard with the datasets visualised on it, is what Wireframe means.

The Wireframes for the MASTEAM are the Main page Wireframe, the Students page Wireframe and the Courses page wireframe. Moreover, the three different Wireframes they will be explained in details below.

### 3.2.1 Main page Wireframe

The creation of the main page of the Dashboard is maybe the most important one, because through the main page the Dashboard need to provide to the coordinator the general view of the MASTEAM only from one page.

The coordinator need to understand from the beginning what is the status of the MASTEAM without spending additional time searching for important information about the general view of the Master. The main page it will contain information about the enrolment, information about the students and information about the graduation rate.

The enrolment part it will show the total number of the applications through the three academicals years 2015, 2016 and 2017, the total number of the accepted one and the rejected one again through the three years, the number of the enrolled one and the percentage of the graduated students for the three years.

The Main page Wireframe it will contain also the number of the male, female and European, Non-European students. Moreover, an addition information about the percentage of the male, female and European, Non-European it will be included.

Furthermore, the second part of the Main page in will contain the above information on detail with the enrolment status and the number of male, female and European, Non-European for each year separately.

### 3.2.2 Students page Wireframe

The Students page Wireframe it will contain the CSV files with all the personal information about the students and all the grades that they have for each course.

The first part of the Students page it will contain information about the grades that the students obtained and the average grade that they have for the years 2015, 2016 and 2017.

The second part of the Students page it will have a geographical map with the countries that the students are from for each year separately.

Moreover, the third part of the Students page it will contain all the student personal information like name, id number, age, gender, nationality, enrolment date and the semester that they will start.

Finally, the last part of the Students page it will include in details the grades of each student on each course that enrolled with three different data tables one for each year 2015, 2016 and 2017.

### 3.3.3 Courses page Wireframe

The last and final Wireframe is about the Courses page, this Wireframe it will consist all the information about the courses that the coordinator need to know like enrolled student per course, the grades of them and the average grade of each course for the years 2015, 2016 and 2017.

This part is similar to the Students wireframe because contain also the grades of the students, but it will be with different visualization technique as was mentioned on the subchapter ***3.1.4 Courses and Grades CSV dataset***.

Furthermore, the Courses page Wireframe it will have the number of enrolled student per course for each of the three academicals years. Moreover, for the enrolled students they will be additional information about the number of students that pass or fail the courses and average grade per course according to the grades that the students obtained in each course separately.

The Courses page Wireframe it will also have a comparison between the same courses for the three different years, for example it will show the status of one course for the three years together so the coordinator can observe the evolution of every course separately for the three years.

Conclusively, like every page’s Wireframe at the end it will be included in detail three different data tables with all the information for the courses of the three academicals years.

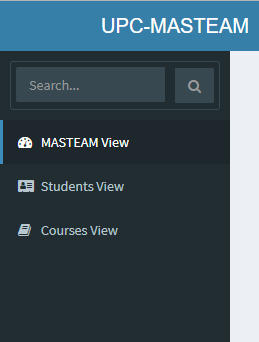
# Implementation of the Dashboard

The implementation of the Dashboard is ready now to begin. After the question creation for the interview, the interview part, the creation of the possible datasets and the creation of the wireframes if finally the time for the programing part and the creation of the MASTEAM Dashboard.

Now is the time to proceed with the implementation of the wireframes that were been explained on the subchapter ***3.2 Build the dataset Wireframe*,** to real Dashboard pages through the programing part with R-studio and the Shiny Dashboard.

The R language is a powerful programing platform for big data analyses and for complex mathematical equations. R-Shiny is an extension of R-studio that helps to make interactive web applications for visualizing datasets. R-Shiny can analyse the datasets that were been created and at the same time to visualise them with different visualization techniques.

The MASTEAM Dashboard contain three main pages that are following the same methodology according to the Wireframes. The first page that is the MASTEAM View page, the second page is the Students View page and the third page that is the Courses View page. Additionally, the below figure ***Fig. 4.1*** is showing the three different main MASTEAM Dashboard pages.

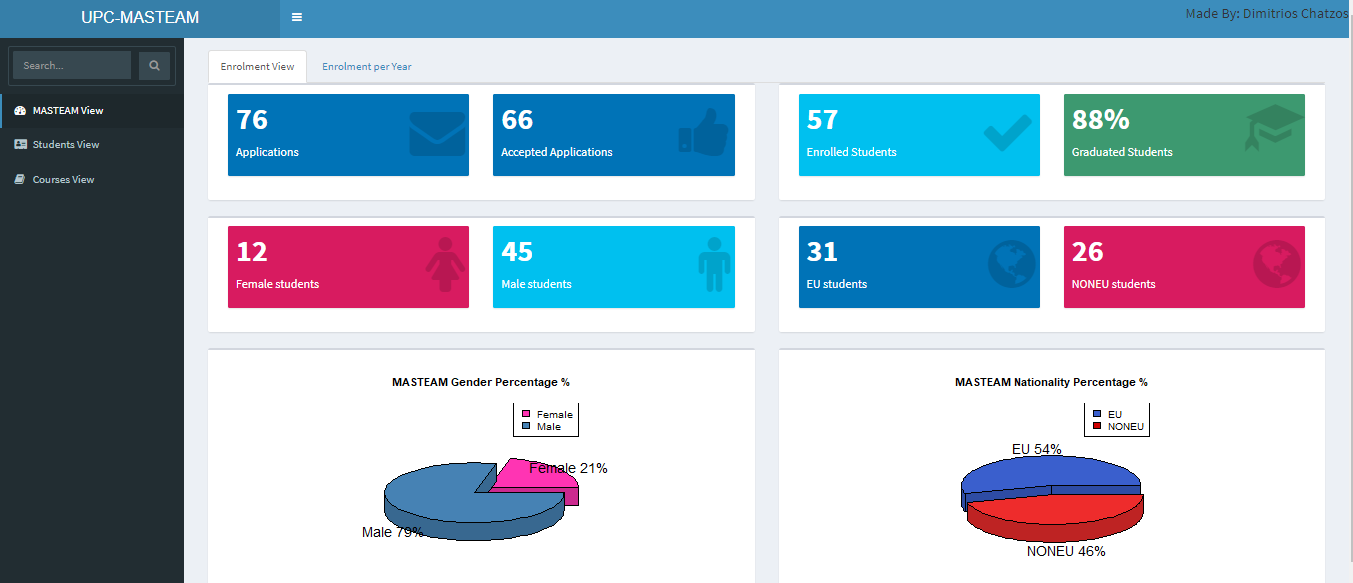


**Fig. 4.1** Main MASTEAM pages

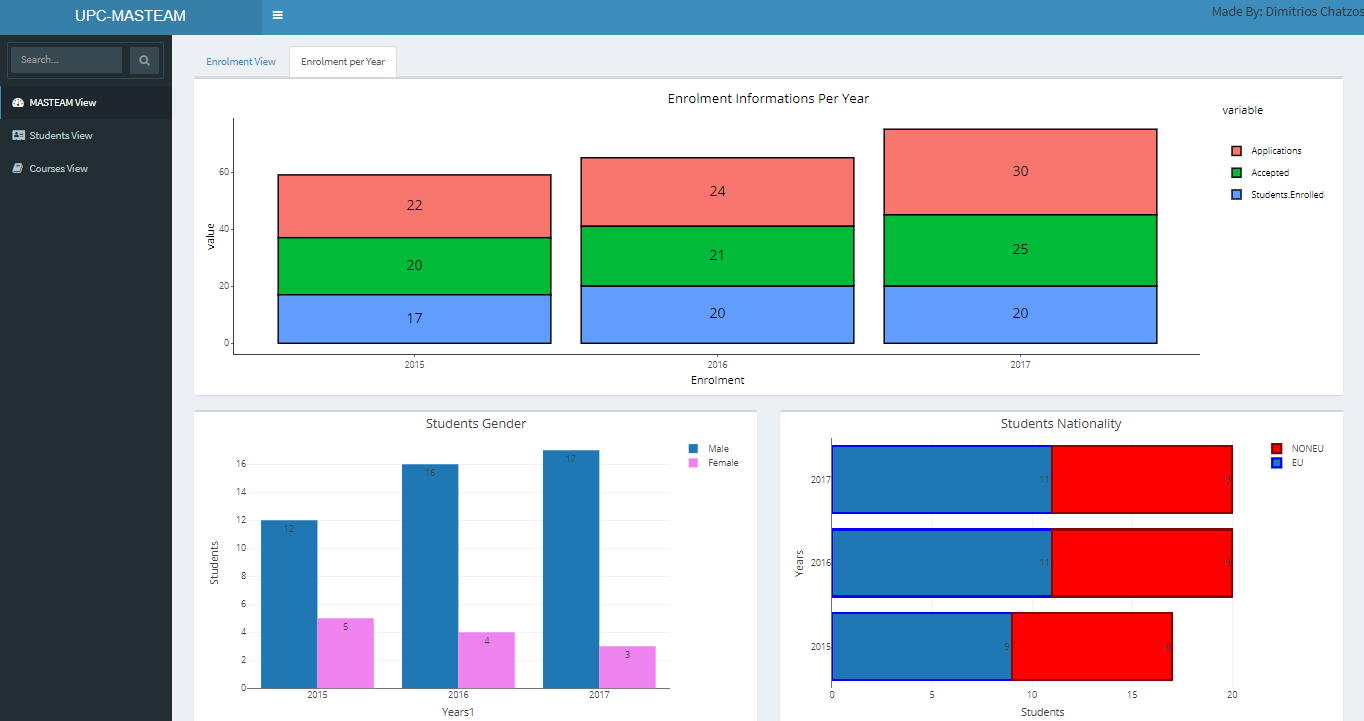
## 4.1 Implementation of the main Dashboard page

The implementation of the MASTEAM View page is accomplished, the programming is following the instructions that were explained on the subchapter ***3.2.1 Main page Wireframe***.

Moreover, the visualization of the MASTEAM View page can be observed on the below figures ***Fig. 4.2*** and ***Fig. 4.3***. Furthermore, every part of the page it will explained part by part on details.



**Fig. 4.2** MASTEAM View Enrolment View



**Fig. 4.3** MASTEAM View Enrolment per Year

### 4.1.1 Enrolment View

The Enrolment view page that was shown in ***Fig. 4.2*** contains the total number of submission letter applications for the MASTEAM, the total number of the accepted ones, the total number of the enrolled ones and the graduation rate.

Moreover, from the enrolled students the number of male, female and European, Non-European students are presented also in the boxes. The percentage of the gender and the nationality is also visualised through two pie charts.

To achieve thus results the three different datasets from the years 2015, 2016 and 2017 were merged together like one dataset. This technique is giving the ability to R program to search through only one dataset the information that were needed to be an output for the boxes.

The R program was programmed to count from the merged new dataset the total number of the female and male students and the total number of the European and Non-European students. The green box with the ratio of graduated students was programed to count all the students that had passed the Master Thesis searching through the students Master Thesis grades which of them had grade grater or equal to five.

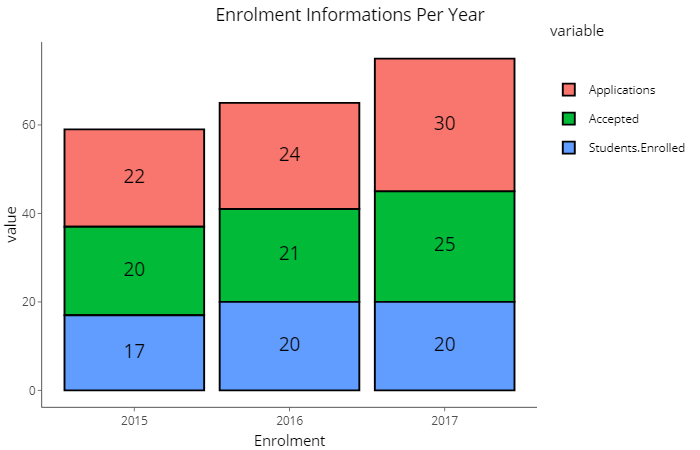
This page can demonstrate to the coordinator the information about the students of the MASTEAM and the enrolment status through time. According to the numbers, the coordinator can understand if any problem exist for the students of the MASTEAM. For example, the number of the male students on the MASTEAM is triple of the number of the female students. This information shows to the coordinator that is a problem with the few female students and additional decisions need to be taken for this to change.

### 4.1.2 Enrolment per Year

The second page of the MASTEAM View as was observed in the ***Fig. 4.3*** contain with more details what the first page is visualizing, but not for the three years together but for each year separately.

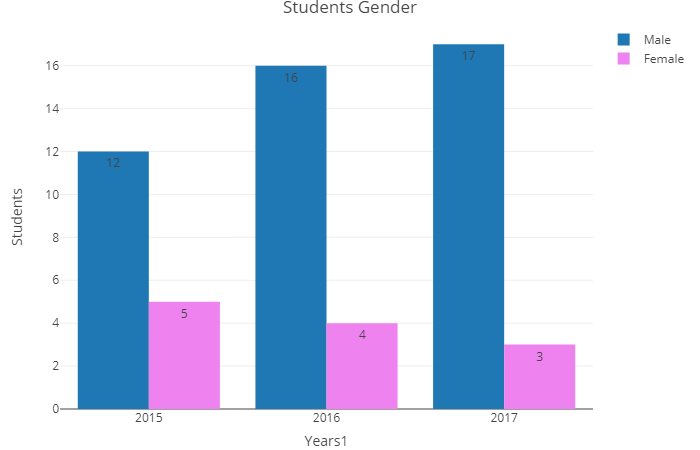
The upper part of the Enrolment per Year page contain a stacked bar plot with the number of applications, the accepted ones and the enrolled ones for every year. The down part of the page have two different bar plots one horizontal with the number of male and female students for each year and one vertical bar plot with the number of the European and Non-European students for each year again.

For further details the Enrolment stack bar plot, the gender bar plot and the nationality bar plot can be observed on the below figures ***Fig. 4.4***, ***Fig. 4.5*** and ***Fig. 4.6***.



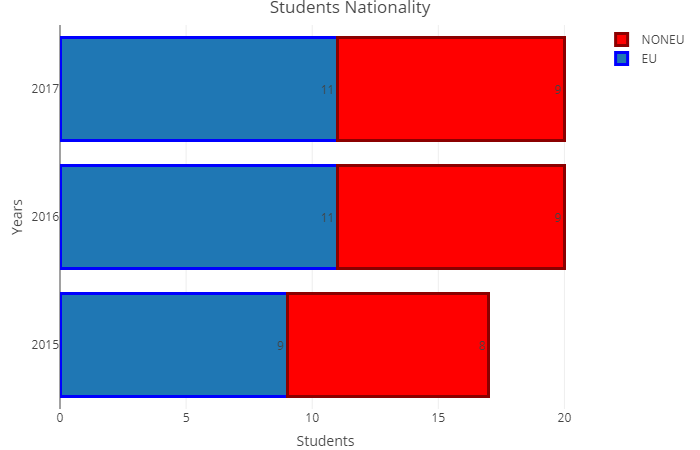
**Fig. 4.4** Enrolment Stack bar plot

From the stacked bar plot on the ***Fig. 4.4*** can be noticed on top the total number of Enrolment applications, below the number of the application that had been accepted from the Enrolment applications and on the bottom is the number of the final enrolled one of the accepted application.



**Fig. 4.5** Gender bar plot

From the bar plots on the ***Fig. 4.5***, can be observed with blue the number of male student and with pink the number of female. The ***Fig. 4.6*** below is displaying with blue the number of European student and with red the Non-European ones.



**Fig. 4.6** Nationality bar plot

## 4.2 Implementation of the Students View page

The second important page of the Dashboard, the Students View page is ready now to be programmed and to visualise the datasets, according to the wireframes on the subchapter ***3.1.2 Students information CSV dataset*** and ***3.1.3 Students grades CSV dataset.***

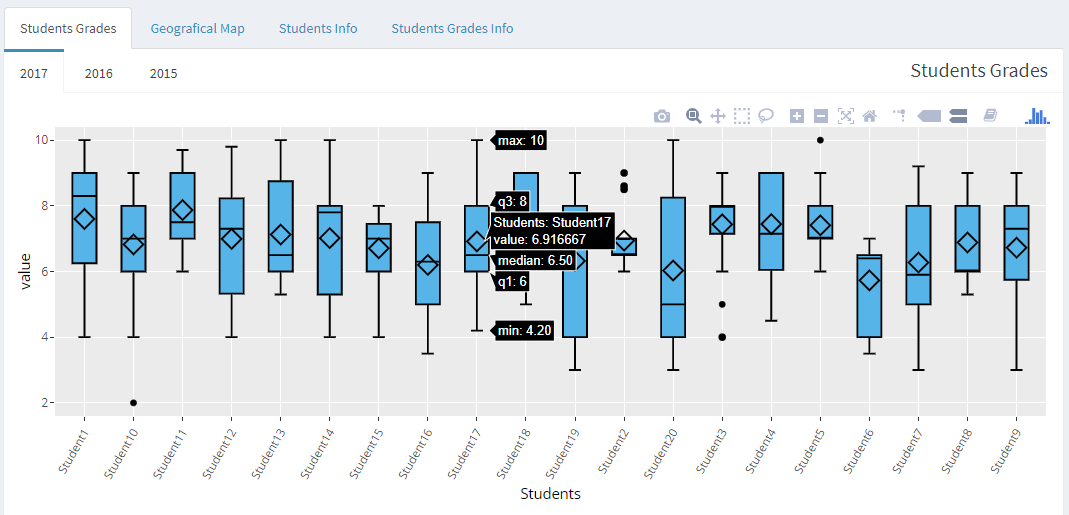
The Students View page is separated into four different subpages, the **Students Grades** that contain the student’s grade, the **Geographical Map** that contain the student’s countries, the **Students Info** that contain the student’s personal data, and the **Students Grades Info** that contain in details the students grades per course.

Additionally the above four subchapter will be explained on details following with all the visualization techniques and figure from the Students View page.

### 4.2.1 Students Grades subpage

The Students Grades subpage was programmed to visualise the student’s grades through box plots that are visualising the student’s top grades, middle grades, lower grades and the average grades. Thus, box plots can give to the coordinator the comparison between the student’s grades, the information about the top students and the students with the lower grades.

The visualisation with the student’s box plots can be examine on details on the below figure ***Fig. 4.7***.



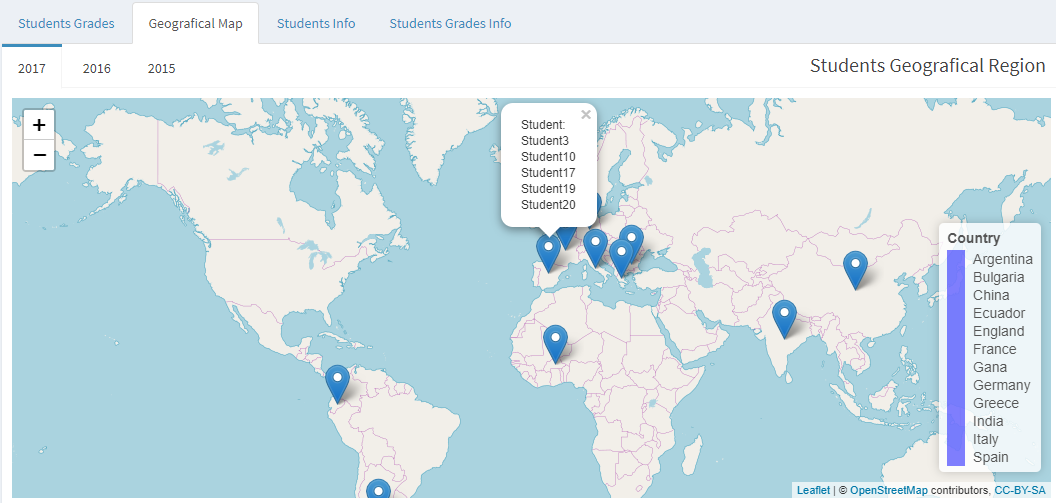
**Fig. 4.7** Students Grades Box Plot

From the ***Fig. 4.7*** can be noticed on the upper left part that are three different box plots pages for the years 2017, 2016 and 2015. Furthermore, below on the X-axis are the student’s names and on the Y-axis the grades from zero to ten, as was explained before every box plot contain the student’s grades that are hidden for the Students Grades box plot and can be shown automatically when the mouse courser is on the boxplot as is showed on the above figure. This technique was done to avoid any misunderstanding of the student’s grades if the page was full of grades in every box plot, and for aesthetic reasons.

### 4.2.2 Geographical Map subpage

The Geographical Map subpage contain the geographical map, that are pointed all the countries that the students are citizens. This visualization technique is taking the geographical coordinates from each student’s country and is putting a point on those countries. Every country has specific coordinates than can be taken through the geographical latitude and longitude. Those parameters were added into the ***3.1.2 Students information CSV datasets*** manually for every country.

The figure with the visualization of the Geographical Map subpage is given on the below



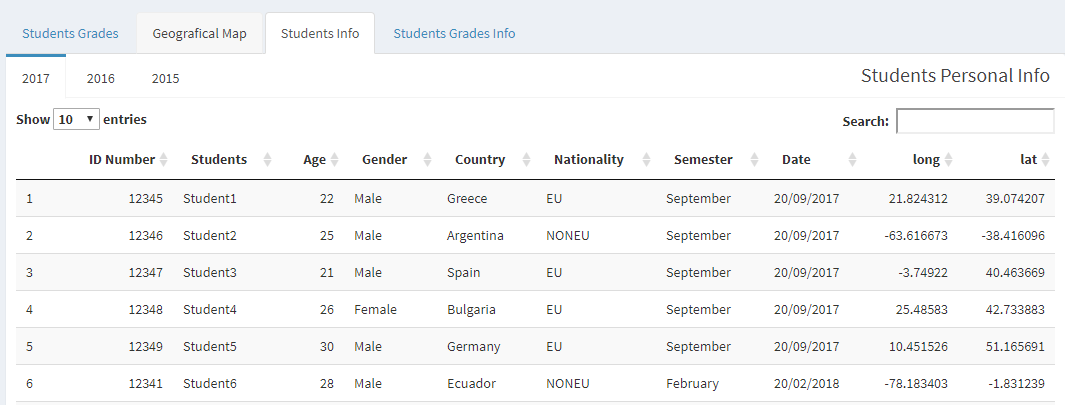
**Fig. 4.8** Geographical Map

From the ***Fig. 4.8*** can be noticed that again they are three different geographical maps for the three different academicals years. Every map has the ability for zoom in and zoom out for further details or less about the geographical areas that the coordinator wants to observe. The map is labelled with the names of the countries that the students are and every time a country point is clicked, is shown automatically the student or the students that are from this country as can be observed also on the above figure.

### 4.2.3 Students Info subpage

The Students info subpage have all the personal data from each student on the MASTEAM Master separately for the three years. The subpage include the: ID numbers, student’s names, ages, and genders, countries with the geographical coordinates, nationalities, semester, and the enrolment dates.

The view of the Students Info subpage can be observed on the below ***Fig. 4.9***.

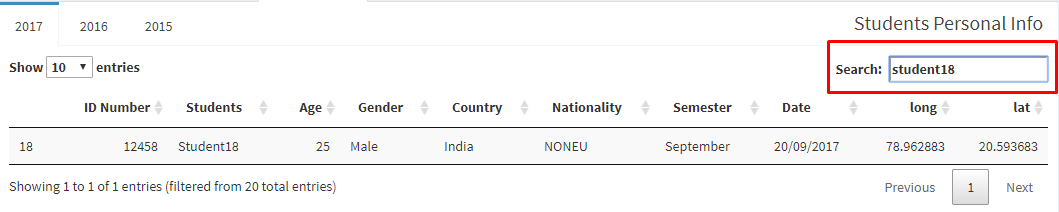


**Fig. 4.9** Students Info

The Students Info subpage was programmed to have the ability to sort all the data according to what the coordinator want, for instance sort the student names alphabetically or sort them from the age or from the gender, etc. The sorting part is coming from the small arrows that are next to each category and have the ability to sort the information downwards or upwards.

Finally, on the Students Info subpage is also important ability that the coordinator has to search for a specific name through the “Search box” that is on the top left part of the window.

The figure below can give an example of the “Search box” use.



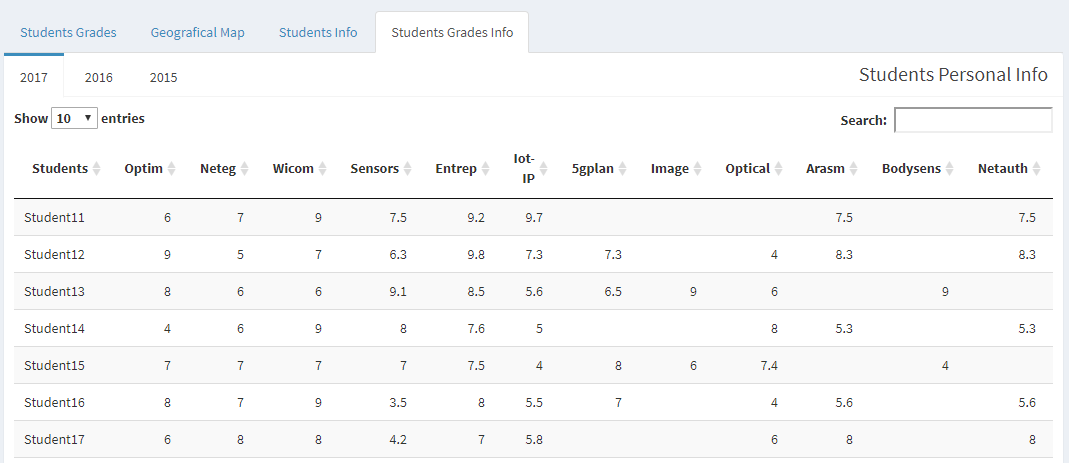
**Fig. 4.10** Search box

### 4.2.4 Students Grades subpage

The subpage Grades from the Students View page contains the information about the grades that the students obtained from each course that enrolled.

The dataset that was used for this visualization part is the dataset that was mentioned on the subchapter **3.1.3 Students grades CSV dataset**.

The figure ***Fig. 4.11*** is showing the Grades page, this page was programmed with the same way and has the same abilities like the ***4.2.3 Students Info subpage.*** Expect from the five first courses that are mandatory and all the students had enrolled, they are empty spaces on some courses and that is because the students had not enrolled for that courses.



**Fig. 4.11** Students Grades

## 4.3 Implementation of the Courses View page

The third and final important page of the MASTEAM Dashboard is the Courses View. This page was programed and visualised according to the wireframes on the subchapter ***3.1.4 Courses and Grades CSV dataset*** and ***3.*1.5 Number of students per course CSV dataset**.

The Courses View page is the page that contain the information about the MASTEAM courses, like grades, the average of every course, the number of students of each course and the number of the fail, pass students.

The Courses View page is separated in three different subpages, the **Courses per Year View**, the **Courses Individual Info** and the **Courses Students and Grades Info**. Furthermore, the three subpages they will be explain into details with figures on the following subchapters.

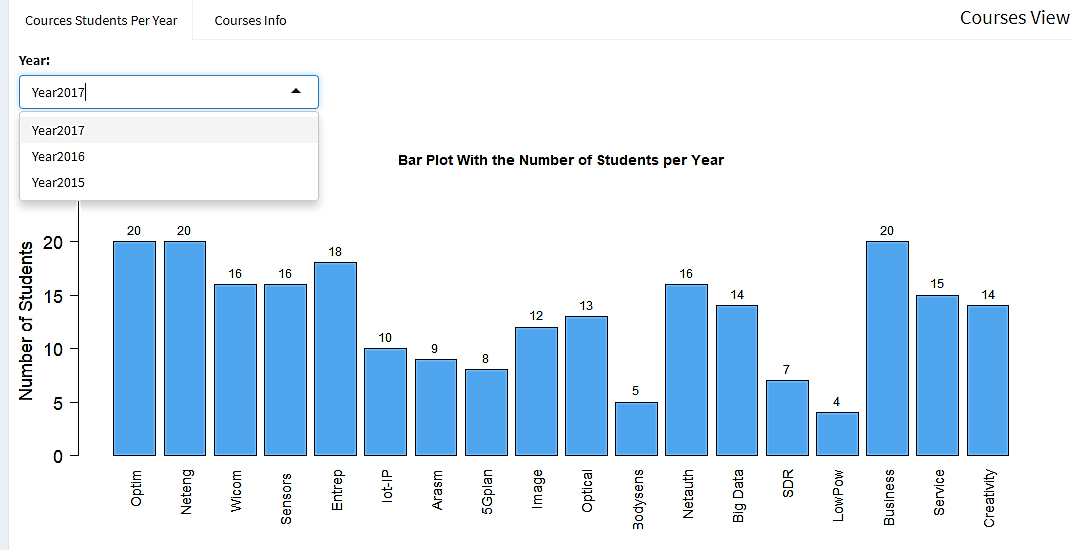
### 4.3.1 Courses per Year View subpage

The Courses per Year View subpage contain the number of the students that had enrolled in every course for the three years, moreover contain also the number of students who pass or fail for every course, and the average grade that every course has. This page is also separated into two different subpages, the **Courses Students per Year** and the **Courses info.**

#### 4.3.1.1 Courses Students per Year

The Courses Students per Year subpage was programed to contain a bar plot with all the courses of MASTEAM and the number of students that they had for every year separately.

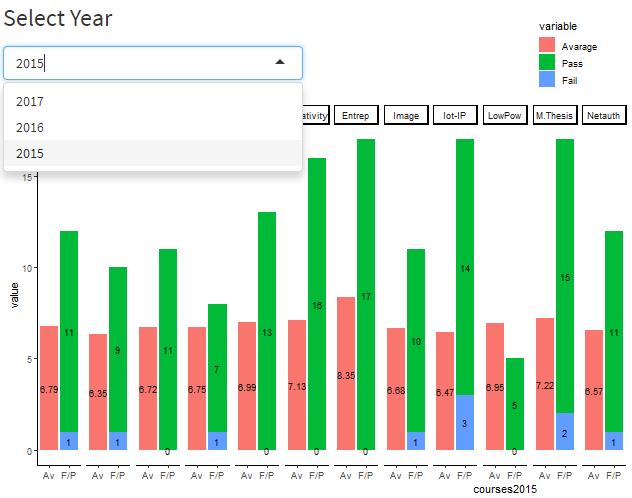
The bar plot on the ***Fig. 4.12*** can give a first and fast view of the number of student per course to the coordinator, letting him to understand which courses are popular and which are not. The coordinator can also choose and change the years through a slider.



**Fig. 4.12** Courses & Students bar plot

#### 4.3.1.2 Courses Info

The Courses Info subpage holds the number of students that pass or fail the courses and the average grade per course. The programing part for this subpage was different from the other pages because for this subpage a stack bar plot and a group bar plot were visualised together as can be noticed on the figure ***Fig. 4.13***.



**Fig. 4.13** Stack-Group bar plot

Moreover, from the ***Fig. 4.13*** can be observed the group and stack bar plot together for every course individual. The combination of thus two bar plots together does not have any official programing algorithm so additional algorithms were combined together for this result. That is the reason that this page is different from the others.

This page is important because the coordinator can understand directly the situation of the Master. For instance if the average grade of one course is low compared to the other courses then maybe this course need to be observed further on the future, or if the number of fail students is height in one course compared to the others then something is going wrong.

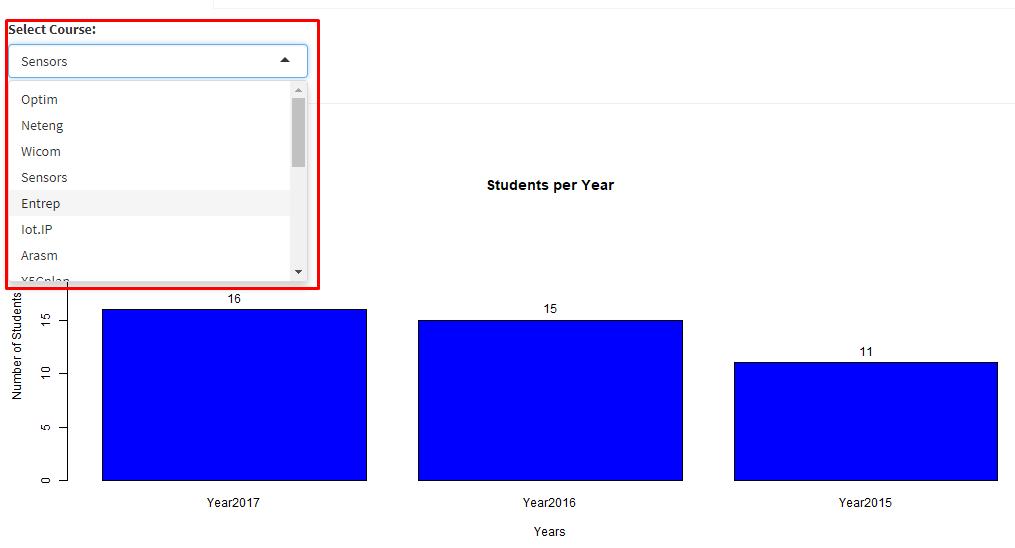
Like previous visualisation techniques, also here the coordinator has the ability to change and choose the year through a slider and to make a quick comparison between the previous year’s observing if any big change exist.

### 4.3.2 Courses Individual Info subpage

The Courses Individual Info subpage is similar to the previous subpage but is making a comparison for each individual course together for the three different years. The subpage contain for each course the number of the students that had enrolled per course with the three years together, moreover have the fail and pass students and the average of every course. This page is divided also into two different subpages, the **Number of Students per Course** and the **Fail/Pass and Grades per Course.**

#### 4.3.2.1 Number of Students per Course

This subpage contain a bar plot with the enrolled student per course for the years 2015, 2016 and 2017, with a slider that can change and choose the course that the coordinator want to observe. Moreover, the bar blot can be more understandable on the below figure ***Fig. 4.14***.



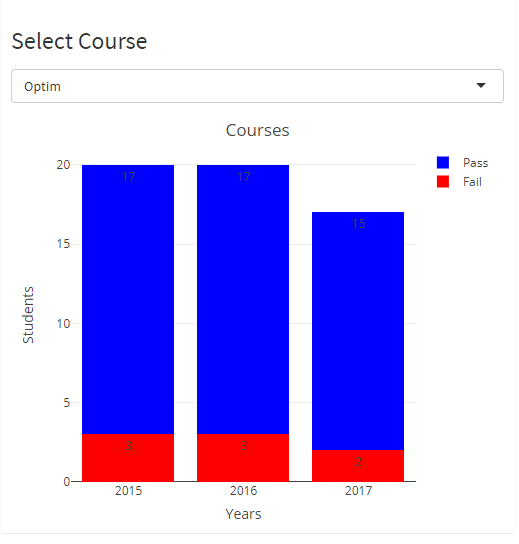
**Fig. 4.14** Enrolled Students Bar plot

#### 4.3.2.2 Fail/Pass and Grades per Course

The Fail/Pass and Grades per Course page include two different plots, one stack bar plot with the pass and fail student per course, and one box plot with the grades per course.

The stack bar plot was programed to count all the student’s grades per course that were equal or greater than five and the grades less than five. The stack bar plot have also a slider that can change the courses.

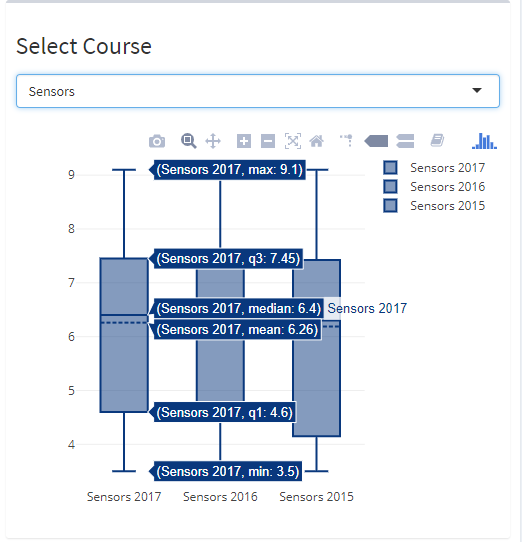
The results of the stack bar plot can be examine on the below figure.



**Fig. 4.15** Stack Fail/Pass bar plot

The other plot of the page is the box plot and contain the higher, lower and average grade per course compared with the three years together like stack bar plot above. The box plot was programed with the same way like the ***4.2.1 Students Grades subpage*** and when the curser is above the course automatically the grades appears.

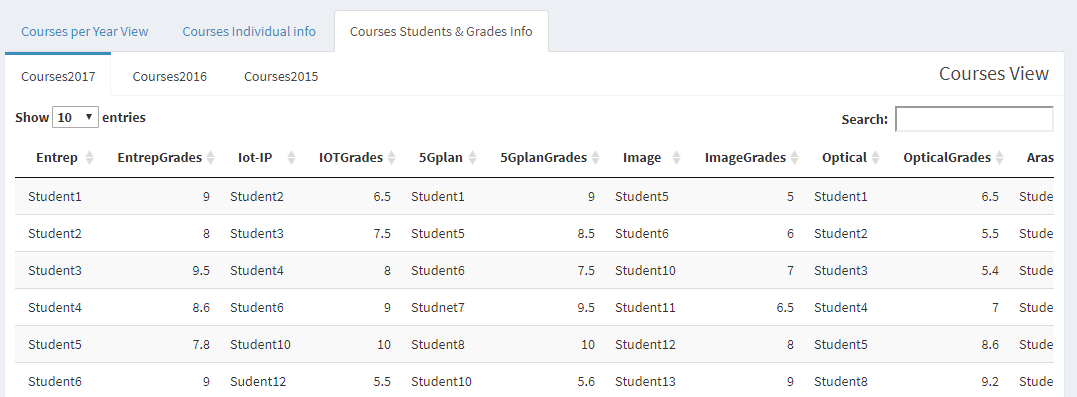
The box plot includes also the same slider to change the courses like before and can be observed better on the below figure.



**Fig. 4.16** Courses box plot

### 4.3.3 Courses Students and Grades Info

The last subpage of the **Courses view** page is the Courses Students and Grades Info, and is consisted of detailed data tables of the enrolled student per course with their grades together for the three academicals years. The abilities of thus data tables are the same like on ***4.2.4 Students Grades subpage*** furthermore the figure below is showing how the subpage looks like.



**Fig. 4.17** Courses Data table

# Explanation of R code along with the Maintenance of the MASTEAM Dashboard

The MASTEAM Dashboard is ready now for use with all the available datasets and features. The coordinator is able now to know what is the MASTEAM status and is able to take any additional decisions if is needed for making the Master better or to maintain it.

Until now, was explained all the procedure of making a Dashboard from the begging until the end, from the interview part ***Analysis of the case MASTEAM,*** the designing part ***Design a Dashboard*** until the implementation part ***Implementation of the Dashboard***. The only thing that need to be explained more is the important programing parts that the MASTEAM Dashboard have and how to maintain it for the future.

On this chapter further explanation about the important parts of the R code and how they are working it will be done. The important R code parts can be useful for the coordinator to maintain the Dashboard for the future and to understand how the MASTEAM Dashboard is working if is needed to change or add something.

## 5.1 -R- code explanation

This subchapter is explaining how the -R- code was programed on -R- studio (R version 3.5.1) along with the -R- Shiny Dashboard interface. Furthermore it will analyse the basic R code parts from the MASTEAM Dashboard and why were programmed like this.

Every programing platform to work and to make complicated calculation or to have visualization techniques need some libraries, those libraries contain inside different equations and pre-programing parts that can be used for different tasks inside the -R- code.

The creation of a Dashboard with the -R- Shiny interface is using the basic programing parts and parameters from –R- studio but is little bit different for the reason that here the program is creating a webpage and needs a server input and output to be able to visualise all the datasets into the MASTEAM page.

Additionally the R Shiny is divided in two parts the UI part and the server output part. The Dashboard is crated in two different programing parts the UI.R and server.R. The UI are acronyms from the “User Interface” and this part contain all the code about the Dashboards interface page, more specific how the MEASTEAM page looks like.

The server.R output is the part that contain all the equations and the visualization code techniques that are giving the correct visualization output to the web page combined together with the UI.R part. Moreover, the -R- platform has the ability to visualize inside the platform all those bar plots, box plots, etc. but for the web page, UI.R and server.R are needed.

### 5.1.1 -R- Libraries explanations

The -R- platform contain many libraries for different computation parts. All the libraries that were used for developing the MASTEAM Dashboard they will be explained in details on this subchapter.

To use a library in -R- is easy, writing in the begging of the code the command *“library()”* and inside the brackets the name of the library .

The libraries that the MASTEAM Dashboard use are the following one, with a brief explanation:

* **library(shiny):** The basic library to start a shiny interactive web application.
* **library(shinydashboard):** The library to create an interactive Dashboard with the Shiny package**.**
* **library(readxl):** The readxl package is reading the Excel datasets and is importing them into R.
* **library(dplyr):**  The dplyr package is transforming and summarizing data sets to be easier for use.
* **library(data.table):** This package can process big amount of data and is able to manipulate them for instance to add, to delete, to order, to separate the data inside the data tables.
* **library(ngram):** This package has the ability to process words and letters and to find and count from a text specific words or letters.
* **library(RColorBrewer):** Is a colour package that can create beautiful colour combinations for the different visualization techniques.
* **library(tuple):** This package can discover any duplicated or replicated data.
* **library(shinyWidgets):** This package contain more widgets for the Shiny Dashboard and can add more visualization techniques.
* **library(ggplot2):** ggplot2 is one of the most know and useful package for creating graphics of any kind of datasets and is based on the Grammar of Graphics.
* **library(memisc):** A package for managing data sets and transform them to data tables and data frames inside the R platform for manipulating easier the data for statistics purposes.
* **library(DT):** DT is the package that can take any data table on Excel format and to visualize it directly to PHP web format. DT has the ability to manipulate and change the data in multiply ways.
* **library(editData):** This package has the ability to change the datasets directly through R without any coding methods. Easy and fast edit, delete or update to every data table inside R platform.
* **library(eeptools):** A package that can analyse and visualize data sets.
* **library(ECharts2Shiny):** This package can make interactive charts (bar chart, pie chart, etc.) for all the Shiny web applications.
* **library(plotly):** Plotly is a powerful online platform that can make interactive graphs and plots, together with other platforms like R, Pythons, Java script.
* **library(lattice):** Is a package that is improving all the plots and graphs from the R platform and has the ability to analyse complex datasets.
* **library(plotrix):** Another package for graphs and plots with emphasis to labels and axis formats including also variety of colour combinations.
* **library(ggmap):** A package that can include datasets inside maps and analyse them.
* **library(maps):** Maps is the package than can create map plots.
* **library(ggrepel):** This package is a part of ggplot2 package and provide all the text and label variables to visualise correct plots without any missing variables.
* **library(leaflet):** [Leaflet](http://leafletjs.com/) package is the elite library to create interactive and three dimension maps directly from “Google maps”. This library is used from many famous web sites and online applications.
* **library(reshape2)**: One of the most important libraries to reshape datasets and to give them automatically the format that is needed to analyse and plot them.

### 5.1.2 Manipulation of Datasets

The datasets that the MASTEAM Dashboard is using are on Excel CSV format and can be maintained or edited directly from the Excel formats or through the R platform. The CSV datasets can be imported to -R- directly by the *“Import Dataset”* button that -R- have or through the code with the command *“read.csv (“datasetname.csv”)*, this command is easy to be found in any R documentation webpage.

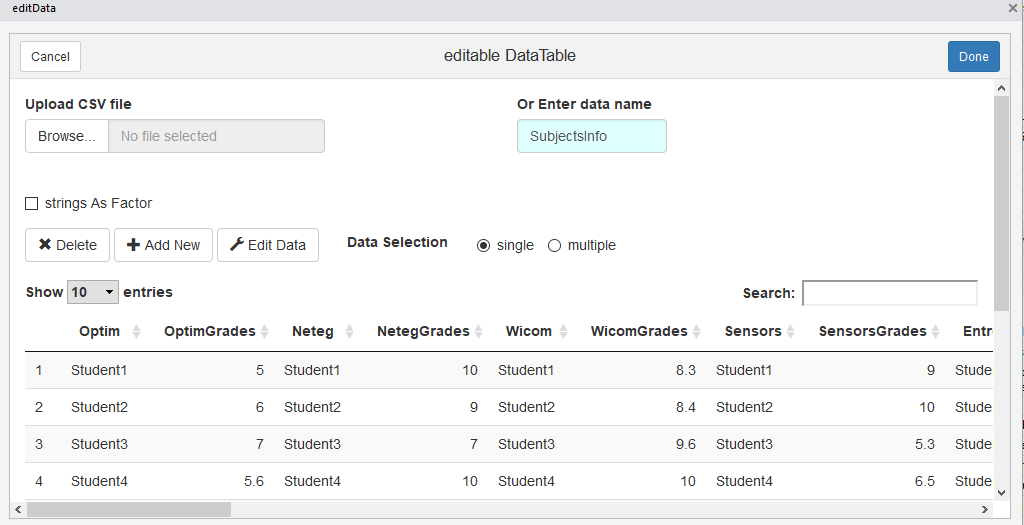
Important here is that the datasets are going to be imported inside the program need to be at the same file location with the application, because this gives the ability to the program to use the datasets any time inside the code only by the name of the dataset without any further instructions or commands.

The MASTEAM Dashboard is analysing the datasets mostly through the “library(DT)” so in many cases is preferred to import the datasets directly into -R- through the *“Import Dataset”* button and manipulate them.

To manipulate and edit the datasets as was mentioned before can be done directly through the Excel CSV files. The possibility to do that maybe some times it will not available so a different approach need to be done.

For the MASTEAM program, the datasets can be manipulated directly through the R studio without the need of the Excel program and the reason for that is the “library(editData)” . This library is giving the ability to manipulate the datasets directly through R.

Furthermore the figure below is showing an example how to edit data with the use of “library(editData)”.



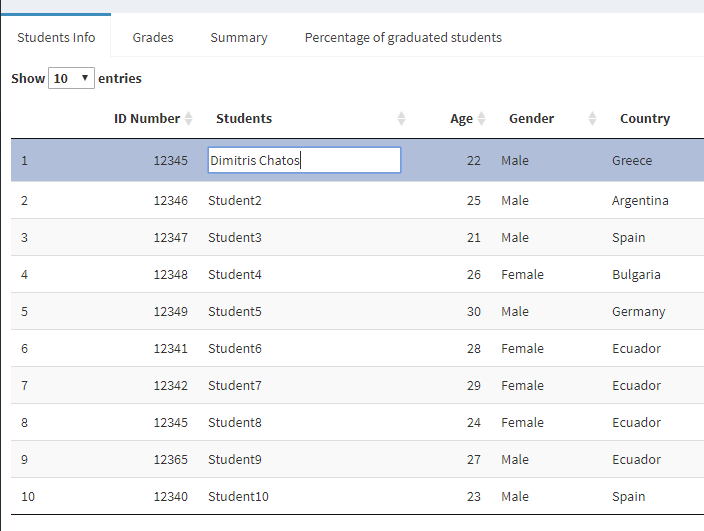
**Fig. 5.1** Edit datasets -library (editData)-

The “editData” library is not the only way to edit the datasets that were imported to the application. The MASTEAM Dashboard has the ability to let the coordinator to edit the data tables also directly from the web Dashboard page through the command “*editable = True”.*

This command is giving the ability to the coordinator to change the data that are inside the data tables only by double clicking them. This technique is available when the command is on the server output part of the code.

Moreover, an example of this technique can be observed on the below figure with changing a data with my own name.

.



**Fig. 5.2** Edit datasets -command (editable = True)-

## 5.2 Maintenance of the Dashboard

The maintenance of the Dashboard is the last step that the coordinator or the programmer need to do for keeping the MASTEAM Dashboard updated for future uses.

To maintain the datasets the user just need to update the datasets CSV files every six-month or every year when the new academically year start and to edit new information about the new students or when they want to manipulate the already existed datasets when that is needed. The procedure is the same like was been explained on ***5.1.2 Datasets edit and maintenance***subchapter.

The maintenance of the R application and the MASTEAM Dashboard is more or less the same, the user need to upgrade when is needed all the libraries to the new libraries versions that can give new abilities to the Dashboard and different or better visualization techniques. Important for the updating part is that the libraries and the version of the -R- platform need to be the same otherwise maybe some features will not be functional.

Last but not least, all the new datasets that the coordinator want to import to the program for the new academicals years can manipulate them and visualise them directly if they are following the same format and structure as the datasets that were explained on the subchapter ***3.1 Creation of the datasets on CSV format***.

Finally, if the coordinator wants to manipulate datasets that are not in the same format as the datasets on this Master Thesis, the MASTEAM Dashboard program is giving the possibility to do that. This is because the programing part thought -R- was made with all the basic and important Dashboard parameters and techniques, so only some additional parts of the program need to be changed according to the news datasets formats.

# CONCLUSION

# REFERENCES

https://cran.r-project.org/web/packages/shiny/index.html

http://genomicsclass.github.io/book/pages/dplyr\_tutorial.html

<https://www.rdocumentation.org/packages/data.table/versions/1.11.4>

<http://ugrad.stat.ubc.ca/R/library/RColorBrewer/html/ColorBrewer.html>

https://cran.r-project.org/web/packages/ngram/index.html

<https://cran.r-project.org/web/packages/tuple/index.html>

<https://ggplot2.tidyverse.org/>

https://cran.r-project.org/web/packages/tuple/tuple.pdf

https://cran.r-project.org/web/packages/shinyWidgets/index.html

[**https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf**](https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf)

[**https://cran.r-project.org/web/packages/memisc/memisc.pdf**](https://cran.r-project.org/web/packages/memisc/memisc.pdf)

<https://rstudio.github.io/DT/>

<https://cran.r-project.org/web/packages/editData/README.html>

<https://cran.r-project.org/web/packages/eeptools/eeptools.pdf>

<https://cran.r-project.org/web/packages/ECharts2Shiny/ECharts2Shiny.pdf>

<https://plot.ly/r/>

<https://cran.r-project.org/web/packages/lattice/lattice.pdf>

<https://cran.r-project.org/web/packages/plotrix/plotrix.pdf>

<https://cran.r-project.org/web/packages/ggmap/ggmap.pdf>

<https://cran.r-project.org/web/packages/maps/maps.pdf>

<https://cran.r-project.org/web/packages/ggrepel/ggrepel.pdf>

[**https://rstudio.github.io/leaflet/**](https://rstudio.github.io/leaflet/)

https://cran.r-project.org/web/packages/reshape2/index.html

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