Internship report

Student: Little Nathan - 14200

Academic year: 2017-20108

School: ECAM

Option: 3GEI

Period: 02/04/2018 au 18/05/2018

Company: EDITx

Address: Rue des Francs 79, 1040 Etterbeek

Mentor of the internship: Combéfis Sébastien

Supervisor of the internship: Flémal Clémence

Contents

1. Introduction

During our 3rd year at the ECAM we are invited to participate to an internship of 6 to 8 weeks in a company of our study field.

The objectives of the internship are:

- to discover the everyday life of an engineer by being paired with one.
- to follow each step of a concrete project and to be part of it.
- to discover the life in a company and teamworking.

My choice went towards the company EDITx which organise online IT challenges in different programming language and field (Project management, big data, etc.).

I was able to participate to a project called Pythia-EDITx and analyse the data of the participant to find relevant information.

This document contains a description of the company I was admit in, the tasks that were given to me and their explanation and my opinion over what I experienced.

2. Objectives of the internship

At the beginning of the internship, my master and I defined the global and specific objectives that I had to achieve during my internship:

Note d'activité :

Objectifs généraux	Objectifs spécifiques
Découverte de l'entreprise	- Structure et fonctionnement - Relations clients - Mission de chaque équipe
Découverte du métier d'ingénieur	- Implication dans un projet réel - Application d'outils de travail (Trello, Git,...) - Méthodologie Agile - Mettre en place, « vendre » un projet
Participation à des projets	- Projet pythia-EDITx : questions interactives de type coding - Participation développement plug-in Drupal

Each objective was validated except the Drupal plug-in that was replaced by the analysis of EDITx's data base due to a delay with the Pythia-EDITx project.

3. Presentation of the company

3.1. Introduction

EDITx was created in October 2016 as a SPRL (Ltd) with a capital of 18,600€ divided in 100 shares (186€ each). The two shareholders at this time were Serge Goffin (50%) and Alexandre Dembour (50%).

EDITx is an online platform where user can take IT challenges. Their motto is: "Challenges, Quizzes & Education in IT for every-one". So, behind the challenges there is also a will to spread IT knowledge and help users to learn new technologies.

They target students, professionals and educators in 2 different ways:

- The student and professionals can test their skills and gain knowledge by taking quizzes to prepare themselves for the challenges. They can win prizes and find employment opportunities.
- The educators contribute to EDITx by creating the multiple-choice questions of the challenges or quizzes. By doing so, they help to spread the knowledge of their

technology and they can also see the statistics about their questions. This can help to see the difficulty of the questions among the students or professionals.

3.2. Organigram

EDITx being a start-up, the number of employees is limited:

- Commercial & Operational director: Alexandre Dembour assisted by Julien Carlier
  Their role is to find new clients and sell challenges. They oversee all the challenge's process (finding the jury, contacting the finalists, …). They also manage the marketing campaigns.
- Financial & HR director: Serge Goffin
  He is in charge of all the financial aspects: checking if the payments are done by the clients, salaries, … He is also in charge of the recruitment.
- IT director: Nicolas Forêt
  He manages the EDITx platform and database.
- Academic director: Sébastien Combéfis
  Responsible of academic relationships and coordination of the IT challenges themes and juries.

For minor tasks such as graphic designer, they hire freelancer for short periods.

3.3. Product

EDITx sells challenges to companies. When a company buys a challenge, it becomes the sponsor of the challenge.

Sponsoring a challenge is interesting for a company in many ways:

- They show that they are researching and linked with a specific technology
- They can increase their visibility in the ICT community and market
- They have access to the profile of every participants and thus the possibility to recruit the ones that they are interested in.

A challenge is usually 15 weeks long with its preparation, set up and final.

First, there is a briefing during which they must choose the technology in which they want the competition to be in depending on their own research and needs and which country and public they are targeting (students, professionals, …). Then EDITx will find professionals and teachers to be part of the jury members who will have to write the questions to the challenge and the quizzes.

The challenges usually take place in two steps:

- During 4 to 6 weeks, the challenge is open on EDITx's platform and marketing campaigns are made to attract participants. They can first take quizzes to test their knowledge and see the type of questions that will be ask to them.
  The challenge consists of answering as much question as possible during 15min. The questions are multiple-choice questions and they have 3 jokers to help them (50/50, phone a friend and ask the audience).
- After this period, the top 10 or 20 of each category are invited to a final at the company's office. There, they will take a challenge but on paper this time and without jokers. The top 3 of each category usually win a prize from the sponsor.

3.4. Finances

In the December 2017 minute, some resolutions were taken, here are the most important ones:

- Each share was divided into 100 shares, so their value went from 186€ to 1,86€ with a total of 10,000 shares.
- The capital was raised to 23,553.15€ by creating 2,663 new shares.
  In those 2,663 shares:
  - 664 issued at 1.86€
  - 1,333 issued at premium of 61.91€ for a total of 63.77€ each
  - 666 issued at premium of 35.68€ for a total of 37.54€ each

The raised capital of 110,000€ is kept in the Paid-In Capital In Excess Par Value.

- Share subscription:
  - Mr. Lepère subscribed for the 1,333 for a total of 85,000€
  - Mr. Coster subscribed for the 666 shares for a total of 25,000€

- Mr. Forêt subscribed for 332 shares of the 664 for a total of 492.02€
- Mr. Combéfis subscribed for the other 332 for a total of 492.02€

The company's turnover for 2017 was 124,379€ with a loss of 78,487€. The loss came from the fact that EDITx is a start-up investing a lot in its platform and its new project that I had the chance to work on: Pythia-EDITx.

3.5. QHSE

EDITx being a start-up, they didn't think much about QHSE yet but here are the definition and what I saw or could be improve in each category:

- Quality: this means work quality but also quality product delivered to the client.
- Health & Safety: it helps to minimise the number of accident at work by training and securing the workplace. For example, during my internship we had an evacuation exercise to know the procedure in case of fire. There was also a cleaning team who would come each day after work to make sur that the office was healthful.
- Environment: the goal is to minimise the impact of the company on the environment. The impact of EDITx is very low because of its size and its product. They don't use much paper nor hazardous chemical so not much could be done here.

3.6. Astonishment report

Regarding my previous working experiences, I was not expecting much from my co-workers. I thought they would not treat me as a member of their team but just as an intern passing by.

But it wasn't the case, I was pleasantly surprise to be warmly welcome. I didn't feel any hierarchy because they were all openly speaking with everyone else and joking about things that could have been "touchy" in another company or with another boss.

My first impression is that every thing was perfectly planned. I was seeing my mentor only twice a week some times but he was always available by mail just in case. My tasks were

clear and precise and I always had the time to accomplish them before the deadline without working 10h per day.

But sometimes, there were expected tasks that were unexpectedly given to me by my boss that had to be done in a very short period of time. For example, the first report I had to do, my boss came to me 2h before he had to leave for the final and told me I had to do the report this time, knowing I had never done one, and asking me for perfection because it was for the client.

I succeeded in time but, in my opinion, didn't achieve perfection because with the stress that I had, I did some small mistakes that cost me more time than it should and thus couldn't do all of what I planned.

The second time he asked me for a report, without me complaining about what I thought on the condition in which I had to do the first one, he asked me to do it much earlier that day and I did a much better report.

This is a good example of what I would improve at EDITx.

I understand that sometimes the boss has to delegate some work and I'm ok with that if I have the abilities to achieve those tasks. But to delegate in last minute is not a good idea for the quality of work done and the state of mind of the worker.

In my opinion, if there is a lot of work to do during a specific day, all the team should gather together over a coffee and discuss about who will have to do what to equally distribute the work between everyone.

At the same time this would resolve the second reproach I have about EDITx, sometimes, as a simple intern, we are not given the global goal of the tasks we are given or we are not aware of what our other colleagues are working on. This could help us to give a better meaning on what we are doing and maybe help our colleagues.

4. Pyhtia-EDITx project

4.1. Introduction

For the moment, EDITx only provides multiple choice challenges but they would like them to be more interactive and personal for the user.

With an open question challenge, the user would have to solve a problem with a code of his own and in the programming language of his choice which would attract all participants from all code language.

Pythia on the other hand is already a testing-based code grader which can test and give feedbacks about a given code regarding the correctness of the tests, the time of execution, the memory used, etc. and works with multiple languages.

So, the solution to the problematic, without reinventing the wheel, was to merge EDITx platform with Pythia program.

The objective of the project is to create an environment where challengers could write code on the EDITx page and it would be send to the Pythia server where it would be processed and the grade and feedback resend to EDITx.

4.2. Pythia

Pythia is composed of 2 main components, a queue and pools.

The queue manages the incoming jobs or tasks that are requested and dispatch them to the available pools and then wait and send back the feedback to the submitter.

A pool consists of a virtual machine. There are 2 main files in a job: one to define the environment to work in (Java, Python, …) and the other to specified the task to execute. The pool creates the wanted environment and execute the demanded task. It then sends back the feedback to the queue.

It is also possible to add a server as a frontend to the queue to receive jobs throughout the web.

4.3. Structure and schedule

The project is to last 9 months, from April to December 2018. Two employees are working on the project: Mr. Combéfis (co-creator of Pythia) on the Pythia framework to add new functionalities and programming language and Mr. Forêt on the EDITx platform to implement a coding interface and on way to send the code to Pythia.

My task was just in the middle: to implement an API REST to link the 2 servers. An API is a software that simplify the interaction between a user and a server/application by normalizing the inputs and outputs. The user can then interact with a program or application without knowing its structure or how it is working.

The project was divided in 6 subparts called milestones:

1) Installation
    a. Installing Pythia on an EDITx server
    b. Drupal module to edit code and process request
    c. Basic echo API
2) Executing code
    a. Make the coding interface more user friendly
    b. Coding interface supporting multiple language
    c. API checking IP addresses
3) Consolidation
    a. API supporting tokens
    b. Coding interface with even more language and autocompletion
    c. Managing down server and adding Pythia status
4) Feedback
    a. Defining the feedbacks of Pythia
    b. Adding unit tests
5) Creating new questions
    a. Script packaging in root /api/tasks/create
    b. Standard question and test editor on EDITx
6) Validate questions creator
    a. Authors can edit, test and simulate their questions
    b. Bug fixing

Because my internship was only 7 weeks long, I could only participate at the first milestone of the project.

4.4. Task

I had to code an API in Go language as describe above. During my first week I had to learn and understand what an API was and how to code in Go which was a new programming language for me.

We divided this task in smaller steps and the first one was to have a simple echo route in the API: a JSON, which is a type of data format, POST request was sent to the API with a key "text" and the API had to echo the value in a "reply" key.

The second step was to have a HTTPS connection between the client and the server of the API. There too I didn't know much about this kind of protocol but with some research it got clearer. The HTTPS protocol was implemented but not tested because of the impossibility to have non-self-signed certificate authority (CA).

With the fast progression of the API and the delay on buying the servers at EDITx, my mentor thought it would be a good idea if I continued the API part of milestone 2 and 3 of the project by adding a middleware to only accept specific IP addresses and access token to authenticate the client sending the request after they would log in. In our case, we didn't implement the log in part which would have mean another route etc. so we generated the token in the API and hard coded it to client request to be authorized.

At the end of the internship, I had to merge my API with the Pythia framework which was a tough task. Even though we already had that kind of exercises of understanding the code of another programmer, the size and complexity of the Pythia made it even more difficult.

Hopefully my mentor who co-created Pythia could help and explain me the general structure of the framework.

5. Data analysis

5.1. Introduction

The second task given to me was to analyse the database of EDITx.

After finishing my API and because of a bug when installing Pythia due to an update of its Linux kernel, I couldn't directly implement my program.

So, my mentor thought it would be a good idea and an interesting task for the company to have me analysing the data of the participants of the challenges.

5.2. Description

We did a brainstorming together over what interesting information could be pull out of their database and we find a few questions. I then went to see my colleagues to ask them what they would want to know more about their users.

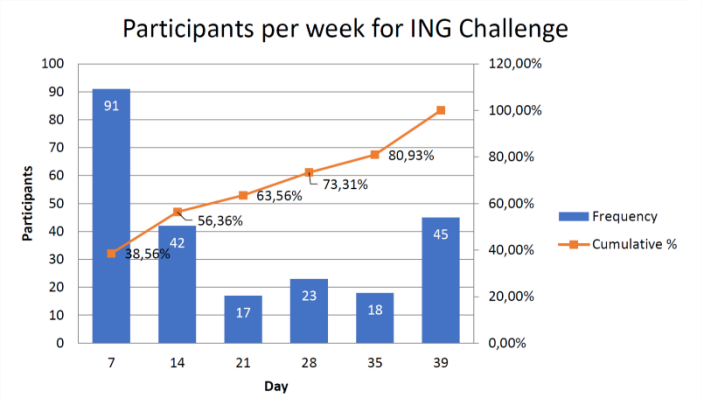
I had 6 questions to answer but because there was another intern with me at EDITx doing a thesis on user profiling, I decided to keep 3 of them that she was not or partially covering.

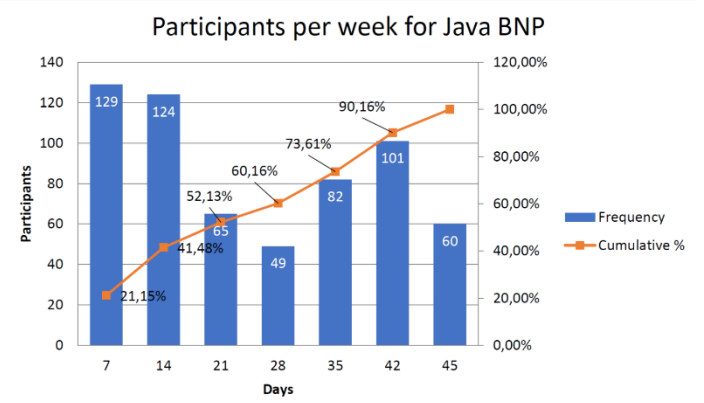
Even though my skills were poor, nay inexistent, I took the challenge and after a few MOOC lessons on EDx and advices from my co-intern, I started to analyse the data on Excel and Google Analytics.

5.3. Results

The report can find in the Annex but here is the summary of my research:

1) Distribution of the participations during challenges





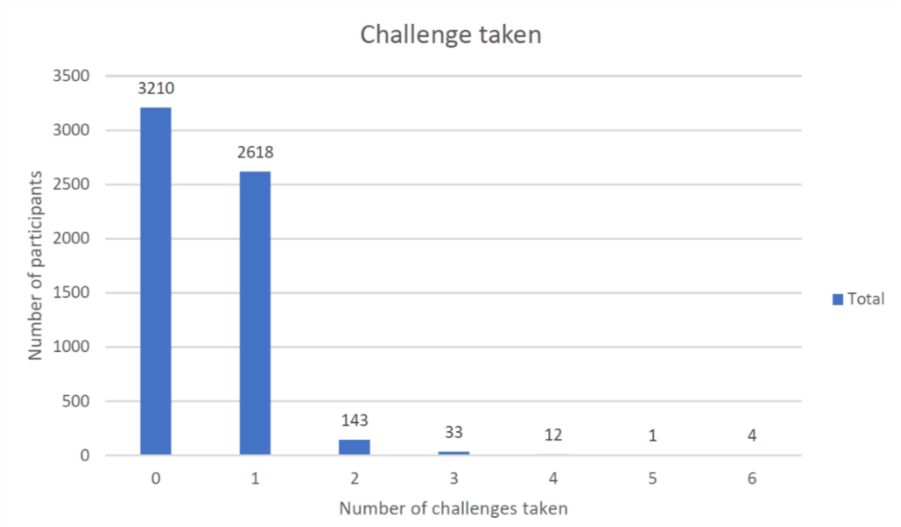
I plotted the number of participants on the Y-axis and the weeks on the X-axis and 5 out 6 diagrams have a U shape with smaller increase at the end.

We can see that between 20 to 46% of the total participants will take the challenge during the first week and that there is a final increase of challengers at the end because EDITx does a big final mailing to attract the last possible participants.

Another observation is that for the Java BNP challenge, the number of participants didn't fall under 49 per weeks which is a lot compare to the others. This comes from a marketing campaign from EDITx consisting of 2 banners on www.developpez.com website.

We can thus conclude on the fact that a campaign, other than mailing, should be undertaken during the mid-term of the challenge to counter the decrease of participants.

2) Are participants returning to EDITx to take other challenges?



As we can see, more than half of the users didn't take a single challenge and in the other half, 93% of them only took one.

There is a lack of users retaking challenges maybe due to the fact that they take the first one to test their knowledge in a certain field and then have no other reason to take another one.

With the prize pool only including the top 3 of each category, it's not the possibility of winning a prize that will attract them neither.

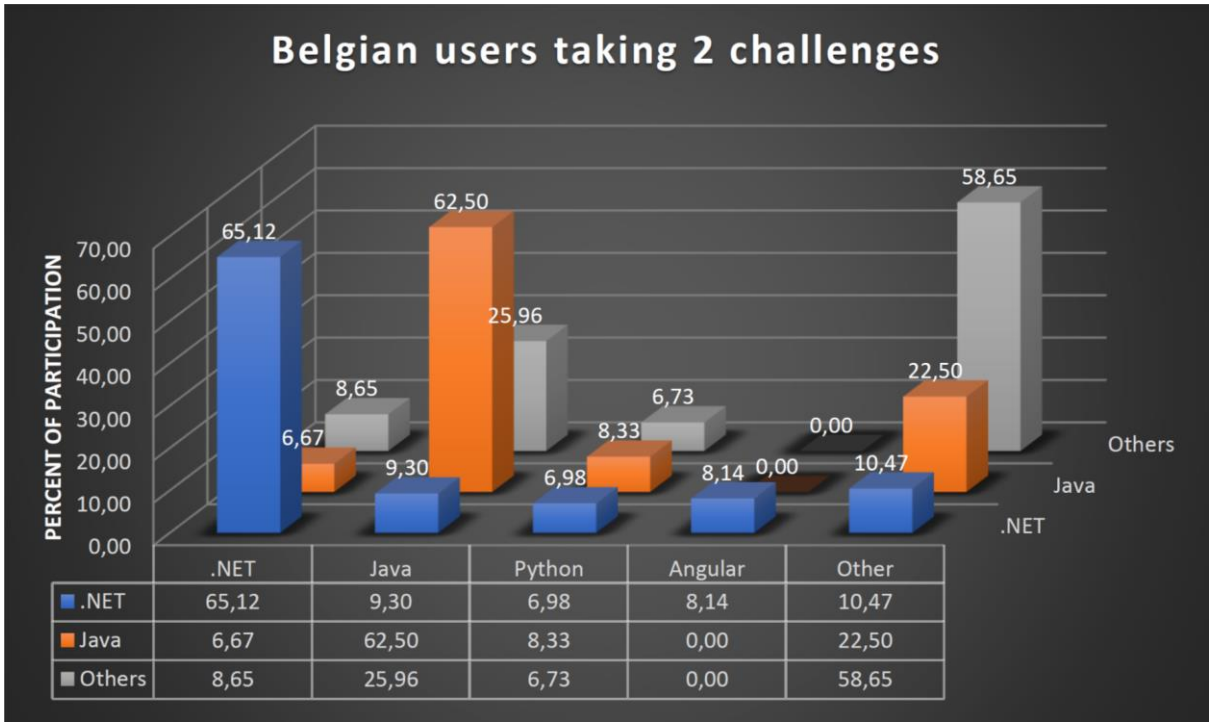
One of my hypothesis is to enlarge the prize pool to the top 50 or even more with smaller price as cinema tickets to give them a greater chance to win a prize and thus make them want to participate.

The second part of this problematic was to see if there were categories of users defined by the type of challenges they were taking.

I took the 111 Belgian users who took 2 challenges and found 3 categories:

- The .NET
- Java
- Challenges not related to a specific programming language (others)

Python and Angular were not considered as categories because EDITx only did one challenge in those fields so obviously the second challenge taken would have been different.



	.NET	Java	Python	Angular	Other
.NET	65,12	9,30	6,98	8,14	10,47
Java	6,67	62,50	8,33	0,00	22,50
Others	8,65	25,96	6,73	0,00	58,65

We can see that the participants tend to take a second challenge in the same field as their first one.

For example, among all the participants who took first a .NET challenge, 65.12% had retaken a .NET challenge.

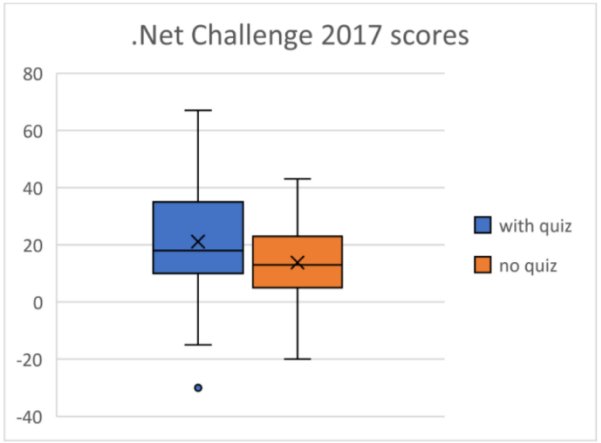
There is also the fact that, for a programming language, the users will prefer to take a challenge not related to another programming language than to change of language.

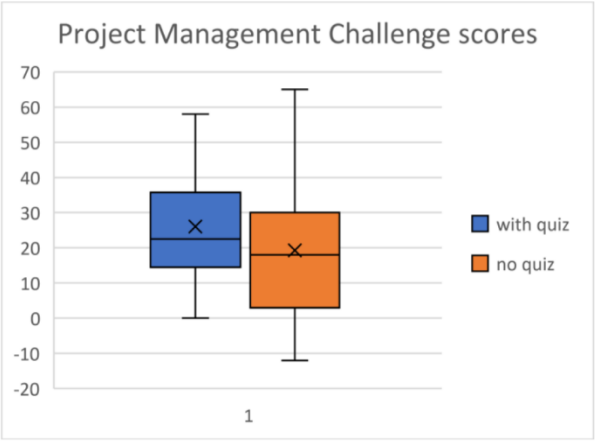
This shows the importance of the Pythia-EDITx project that will allow participants to code in the language they prefer.

3) Is there a correlation between doing quizzes and the score of a challenge?

We separated the participants who did at least a quiz from those who didn't for each challenge. Then we computed their mean and variance for a one-tail z test.

Here are 2 examples of results that were found:





z-Test: Two Sample for Means		
	with quiz	*no quiz*
Mean	21,15	12,50
Known Variance	323,55	164,09
Observations	71	139
Hypothesized Mean Difference	0	
z	3,61	
P(Z<=z) one-tail	0,00015	
z Critical one-tail	1,64	
P(Z<=z) two-tail	0,00030	
z Critical two-tail	1,96	

z-Test: Two Sample for Means		
	with quiz	*no quiz*
Mean	26,08	19,28
Known Variance	254,61	304,55
Observations	38	87
Hypothesized Mean Difference	0	
z	2,13	
P(Z<=z) one-tail	0,017	
z Critical one-tail	1,64	
P(Z<=z) two-tail	0,033	
z Critical two-tail	1,96	

At the end of the analysis, it was proven that there is a difference between the mean of those who did at least one quiz and the one who didn't.

This could be used as an advertisement just before a challenge to tell the challengers that the best participants did a quiz before taking a challenge.

This could then bring more people to do quizzes and spend more time on the platform.

6. Event participations

6.1. Sale support

I had the chance to be part of a sale support at Infrabel. The goal of this meeting was to explain in more details what EDITx can offer and to understand the expectations of the client.

I was with Mr. Dembour and Mr. Combéfis and we met with 3 directors from different departments.

First, one of the director explained what Infrabel was doing, that they were restructuring the company and aiming on reducing the number of employees on the railways and hire more IT profiles. Then Mr. Dembour explained what was EDITx and how it could be interesting for Infrabel to buy a challenge and the possibilities of recruiting it offers.

Afterward, the other directors asked their questions such as how they could target specific profiles, if they could place link to their job applications directly on EDITx website, etc…

At the end of the appointment, they look very interested about the challenge but still had to report to higher authorities before signing the contract.

6.2. CISCO Python final

At the end of the CISCO Python challenge, the 10 best participants of each category were invited to CISCO's office to participate at the final. I had the task to keep the answers of the final challenge and to give them to the jury members so they could correct it.

It was very interesting because it gathers students and professionals who are passionate by the same thing, in this case: Python. I got to learn more about the Python community in Belgium and the events that they were organizing.

The professionals also share their work experience and thoughts about the companies they worked for.

It is also the place where CISCO can approach the best professionals to schedule an appointment and maybe hire them. They also proposed internship or training camp to the students to keep in touch with potential future employees.

6.3. Preparing finals

I also had the task to prepare the report that summarize the talent pool and score of every participants that is given to the client during the final.

This means to gather the graphs and information from EDITx's platform and make a report that has a personal cover page with the colours and logo of the client's company.

7. Conclusion

It was a really interesting experience. I think I was particularly lucky about the company I was accepted in because it might be a start-up but you get to know and speak with everyone and there is a really nice atmosphere.

I was quite out of my comfort zone at the beginning because I knew nothing about Go or APIs and my tasks were the foundation of an import project for the company.

But with the Agile methodology, we went in step by step and it didn't seem so infeasible.

During the Pytiha-EDITx project I worked and learn a lot by myself which is not a downside on itself because I think it is important to fend for yourself.

On the other hand, on the Data analysis project I work a lot with another intern because she was doing her thesis on profiling and gamification. She told me how to use Excel and different analysis technics to finish my task.

I learned a lot of new tools too: Github to post my code so that my mentor could track my progess. I used cURL and Postman to test my API's requests.

For the data analysis, Excel and Google Analytics were used to gather and process the data etc…

I think that this internship really helped me to develop some skills and to learn more about my strengths and weakness but also to validate my choice to continue my studies in the IT department.

8. Bibliography

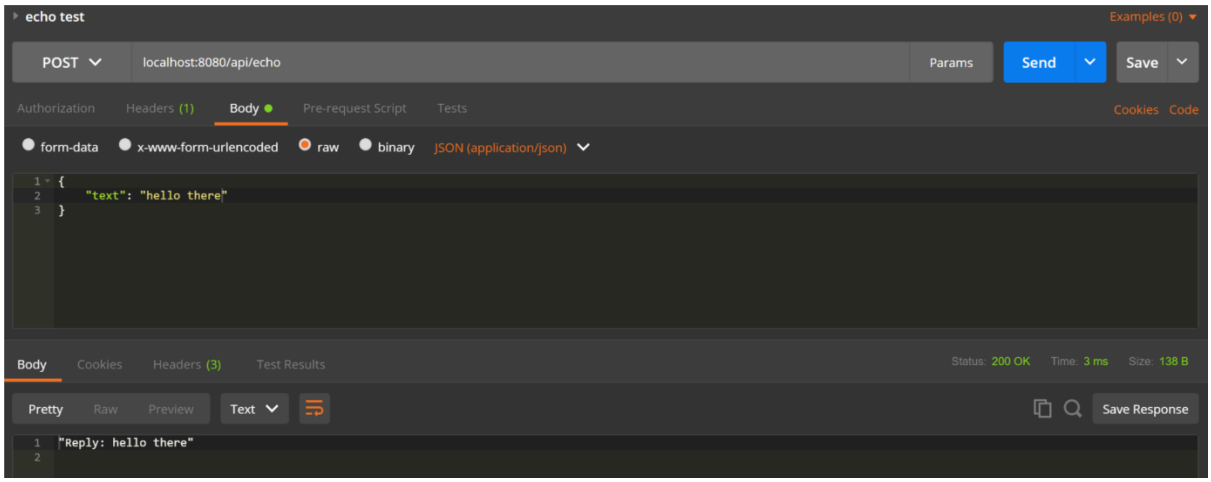
- "EDITx's Web Page." Accessed May 20, 2018. https://editx.eu/.
- "Qualité, Hygiène, Sécurité, Environnement." Accessed May 20, 2018. https://fr.wikipedia.org/wiki/Qualit%C3%A9,_hygi%C3%A8ne,_s%C3%A9curit%C3%A9,_environnement.
- "Why Shares Are Issued at a Premium." Accessed May 20, 2018. https://www.accountingtools.com/articles/why-are-shares-issued-at-a-premium.html.

9 Annex

- Sample of the API code in Go and working test on Postman:

```go
//Echo the given message in a JSON Message struct format
func Echo(rw http.ResponseWriter, r *http.Request) {
    var message map[string]string
    body, err := ioutil.ReadAll(r.Body)
    if err != nil {
        rw.WriteHeader(http.StatusInternalServerError)
    }
    if err := json.Unmarshal(body, &message); err != nil {
        Error422(rw, err)
        return
    }
    for key := range message {
        if key == "text" {
            if err := json.NewEncoder(rw).Encode("Reply: " + message["text"]); err != nil {
                panic(err)
            }
            return
        }
    }
    Error422(rw, err)

}

// Task function for the server.
func Task(rw http.ResponseWriter, req *http.Request) {
    log.Println("Client connected: ", req.URL)
    if req.Method != "POST" {
        rw.WriteHeader(http.StatusMethodNotAllowed)
        return
    }
    // Reading the task request
    body, err := ioutil.ReadAll(req.Body)
    if err != nil {
        rw.WriteHeader(http.StatusInternalServerError)
        return
    }
    var taskReq taskRequest
    if err := json.Unmarshal([]byte(body), &taskReq); err != nil {
        rw.WriteHeader(http.StatusBadRequest)
        return
    }
    // Connection to the pool and execution of the task
    conn := pythia.DialRetry(pythia.QueueAddr)
    defer conn.Close()
    content, err := ioutil.ReadFile("tasks/" + taskReq.Tid + ".task")
    if err != nil {
        Error422(rw, err)
        return
    }
    var task pythia.Task
    if err := json.Unmarshal([]byte(content), &task); err != nil {
        rw.WriteHeader(http.StatusInternalServerError)
        return
    }
    conn.Send(pythia.Message{
        Message: pythia.LaunchMsg,
        Id:      "test",
        Task:    &task,
        Input:   taskReq.Response,
    })
    if msg, ok := <-conn.Receive(); ok {
        switch msg.Status {
        case "success":
            fmt.Fprintf(rw, msg.Output)
        }
        return
    }
    rw.WriteHeader(http.StatusInternalServerError)
}
```

```go
//Routes is a list of Route
type Routes []Route

//NewRouter changed mux.Router func to work with the Rout struct
func NewRouter() *mux.Router {
    router := mux.NewRouter().StrictSlash(true)

    for _, route := range routes {
        router.Methods(route.Method).
            Path(route.Pattern).
            Name(route.Name).
            Handler(route.HandlerFunc)
    }

    return router
}

//MiddleWare check the IP of client with the list of IPs in conf.jdon
func MiddleWare(h http.Handler) http.Handler {

    return http.HandlerFunc(func(w http.ResponseWriter, r *http.Request) {

        addr := GetClientIPs(r)
        IPConf := GetConf().IP

        for _, ipConf := range IPConf {
            for _, ipClient := range addr {
                if ipConf == ipClient {
                    h.ServeHTTP(w, r)
                    return
                }
            }
        }
        http.Error(w, "Unauthorized IP address", 401)
        return
    })
}

//GetClientIPs returns the IPs address of client
func GetClientIPs(r *http.Request) []string {
    //If X-FORWARDED-FOR structure is respected (first IP is the client's private IP address)
    //and separate with ", "
    //Header.Get will have all other IPs but not the last one used (last proxy or client if empty)
    var IPs []string
    if allIP := r.Header.Get("W-FORWARDED-FOR"); len(allIP) > 0 {
        IPs = strings.Split(allIP, ", ")
    }
    ip, _, _ := net.SplitHostPort(r.RemoteAddr)
    IPs = append(IPs, ip)
    return IPs
}

var routes = Routes{
    Route{
        "Echo",
        "POST",
        "/api/echo",
        Echo,
    },
    Route{
        "Task",
        "POST",
        "/execute",
        Task,
    },
}
```



- Report of the data analysis: