**Project Plan**

Developing a Social Media Analytics Tool for finding

**Popular Restaurants in Amsterdam using the iens website**

Group Members:

Andrei Frunze - 551107

Tshen Xue - 521214

Nelson N Nkomo - 541687

Junior Eduado - 519516

Shamil Karimli - 523001

## Incentive

Social media is a powerful tool for delivering a message or an opinion, without any restrictions. The internet facilitates the freedom of speech, thus sometimes it could serve as a reliable source of information. There are lots of restaurants in Amsterdam, and all of them are advertised as being of high standards, just to attract more clients. Sometimes, people want to visit a nice restaurant, which will satisfy their desires, but it’s really hard to rely on any of the ratings out there.

In order to create a trustworthy ranking of these restaurants, it is necessary to consider the only thing that matters, the feedback given by real customers. Iens website provides an accessible way of leaving feedback and grades to the visited restaurants. This website, provides simple collection of restaurant reviews, each linked to a unique user, which could be used to compile a trustworthy ranking of the restaurants from Amsterdam. The creation of such ranking would prove useful, as everyone wants to enjoy a high quality evening at an appropriate restaurant.

## Objectives

This part will explain the goal of assignment referred to the incentive. Based on the needs described in the incentive, the goal of this project is to provide a ranking of restaurants from Amsterdam that are most popular by analyzing the feedback that the restaurants receive from their customers, and creating a weighted graph, with the grade received by the restaurant as weight, thus assisting the team in creating the ranking.

## Central research question and sub-questions

***Main question:***

*How can the data collected from iens website, facilitate the creation of a most popular restaurants ranking?*

***Sub-questions:***

1. *How can the feedback that the restaurants receive from their customers be efficiently extracted from iens website?*
2. *How can the collected data be presented by means of a weighted graph?*
3. *How can the collected data be analyzed in order to create the rankings?*
4. *How can an application compile and represent the data collected?*

## Description of assignment

The iens website is used to show restaurant opinion. It offers the possibility to grade and to leave reviews to the public users about their restaurant experiences. This project hopes to solve the question of which restaurant is the most popular in Amsterdam by using the data of rating and reviews from the iens website from its users. In order to answer the main question properly, sub questions will be used as part of this project. These sub questions help understand how the data may be collected from iens website and how a graph model can be used to run analysis of it. Furthermore a java application will be build and used as part of this project for the purpose of collecting data and helping make the graph model.

## Scope of assignment including preconditions

This section describes the scope of the project.

We will complete following tasks:

1. Analysis of Iens website to find out what type of information is available to gather based on what the customer have to say about the restaurant they have visited.
2. Analysis of current technology of extracting data from the web.
3. Analysis of the available data by the data analysis tools (e.g. EXCEL) and presenting a structured graph that shows which restaurants in the Amsterdam region is most popular in terms of how customers rate them on the Iens website.

The following task will not be part of the project:

Because there is a short research time, the project will analyze a specific region (Amsterdam) and part of the Iens website to gather the information needed.

If we have some time left, the following task will be complete:

1. Looking for more ways of ranking restaurants on the Iens website base on customer feedback, broadening the regions so that Amsterdam is not the only region the application can gather information from.
2. Create a graphical user interface that will be easy for not only private customer to use, but also for the public to use for future research.

## Deliverables

This project will be done with the hope of finding out how can iens website be used to create a model of Amsterdam restaurants based on popularity. To be completed successfully, this project aims to have three deliverables as follows:

1. A project report that shows in details how this project was carried out and what its outcome was. Part of which should be a graph model.
2. A java application that can be used to assist in data collection and graph representation of the outcome.
3. A presentation to show the outcome of the project.

## Risks & Solutions

The risk rating is calculated by multiplying the chance (rated 1 to 5) and impact (rated 1 to 6) of the chosen risk.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Description | Chance | Impact | Risk | Counter Measure |
| Sickness | 1 | 5 | 5 | The work gets redistributed between the other members |
| Skill problem | 4 | 5 | 20 | Members are responsible of refining their skills, in time apart from the project. The chance is dealt with by reading the recommended books and by choosing tools previously known to the team members. The impact is to be handled by splitting the task between the members, in order to ease task’s execution. |
| Hard disk | 1 | 6 | 6 | Keep multiple copies of project files and use online storage |
| Tools Malfunction | 2 | 4 | 8 | Make sure the project files are compatible with multiple platforms and tools |
| Inaccessible Data | 1 | 6 | 6 | Change the source of data in order to accommodate the project needs |
| Late deliverables | 2 | 5 | 10 | Create well organized documentation that will help preventing miscommunications and delays |

## Stakeholders & Communication Schedule

Name: Vera Hollink

Role in Project: Project Supervisor

Email: vera.hollink@inholland.nl

Phone Number: + (31)650277497

Department/Position: Lecturer and internship coordinator - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

Name: Margje Penning

Role in Project: Project Supervisor

Email: margje.penning@inholland.nl

Phone Number: + (31)620413644

Department/Position: Lecturer - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

Name: Andrei Frunze

Role in Project: Team Leader

Email: 551107@student.inholland.nl

Phone Number: + (31)616161325

Department/Position: Student - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

Name: Shamil Karimli

Role in Project: Developer

Email: 523001@student.inholland.nl

Phone Number: + (31)631992294

Department/Position: Student - Mathematical Engineering

Meeting Schedule: Meeting by appointment

Name: Tshen Xue

Role in Project: Archivist

Email: 521214@student.inholland.nl

Phone Number: + (31)652180913

Department/Position: Student - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

Name: Nelson N Nkomo

Role in Project: Developer

Email: 541687@student.inholland.nl

Phone Number: + (31)640347542

Department/Position: Student - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

Name: Junior Eduado

Role in Project: Developer

Email: 519516@student.inholland.nl

Phone Number: + (31)640026291

Department/Position: Student - Mathematical Engineering

Meeting Schedule: Scheduled meeting every Monday and by appointment

## Methods

In this section, all the necessary steps, methods and activities are described to complete the assignment which are related to research questions.  
 **How can the feedback that the restaurants receive from their customers be efficiently extracted from iens website?**

A preliminary study will be conducted to find how we can extract data from the website. First, we are going to find if we can get web content of iens website. If it is possible, then we are going to find and test what API can analyze the web content and help us to get the data we want. For example, an API called web crawler may be an option. This web crawler is an internet bot which systematically “crawls” the World Wide Web, typically for the purpose of Web indexing[[1]](#footnote-1). Therefore, we can use this kind of API to read Iens website to get all the data we need, like comments, ratings and users information, such as usernames, which looks for these certain classes and tags in the HTML file to find and collect the necessarythe data.

**How can the collected data be presented by means of a weighted graph?**A graph consists of an edges and vertices. There are many different graphs which have different advantages of using. The different graphs should be compared and use the one that is best suited for displaying users giving scores to restaurants. After deciding on a graph, decide which data should be where. For example, are the users a vertex or an edge?

**How can the collected data be analyzed in order to create the rankings?**There is a lot of information we can get from a graph. A user gives a score to a restaurant, if you add up all the scores a restaurant gets from the users and divide by every user that entered a score for that restaurant, we will get the average score of a restaurant. By analyzing which different restaurants the users give a score to, we can find out what kind of restaurant the user prefers.

**How can an application compile and represent the data collected?**The application will have to have a GUI where users can put restaurant names into the interface. Then, the application will have to connect to Iens by using the web crawler and return data which contain the user's feedbacks, usernames and rating. After that, the application will process the data and represent the result.

## Planning

|  |  |
| --- | --- |
| ***Week*** | ***Tasks*** |
| **Week 1**  **11 April** | Project kick-off  Idea brainstorming  Establishing ground rules  First Project Plan Draft |
| **Week 2**  **18 April** | Refine Project Plan  Dividing work  Setting up GIT repository  Literature Study on the subject |
| **Week 3**  **25 April** | First implementation of the code  Preliminary testing  Gather information  Research Report Draft |
| **Week 4**  **9 May** | Further code improvements  Analyzing gathered data  Research Report Writing |
| **Week 5**  **16 May** | Application GUI  Research Report Writing |
| **Week 6**  **23 May** | Create PowerPoint Presentation  Final touches of code and documentation |
| **Week 7**  **30 May** | Presentations  Final report touches  Submitting the deliverables |

**Detailed weekly planning**

*Tuesday - April 12th*

1. Submit a new idea for the project plan
2. Rewrite incentive
3. Set up GIT repository

*Sunday - April 17th*

1. Deadline new project plan
2. Research the main topic
3. Find sufficient data

*Monday - April 18th*

1. Feedback project plan
2. Discuss the models that will be used to evaluate data
3. Find sufficient literature for study

*Sunday - April 24th*

1. Improved Project Plan
2. Study iens website’s internal structure
3. Find efficient algorithms and data structures to be used in the tool

1. From Wikipedia, https://en.wikipedia.org/wiki/Web\_crawler [↑](#footnote-ref-1)