Project Proposal - Droppe Market Data Crawler

1. Introduction

Wouldn't it be cool if you could pick any product online and know whether it's cheaper to buy it today or tomorrow? Have you wondered how **web crawlers** work? Are you interested in **data visualisation**? Would you want to learn more about **machine learning** based predictions? Would you like to combine this into a **data collection & prediction** based visualisation platform? Then we have the perfect project for you and your team.

But first, what's [Droppe](#)? We're the one-stop shop for all your industrial wholesale. In other words, a business-to-business marketplace, connecting Nordic businesses to Europe's top equipment manufacturers. This means we tackle all kinds of cool problems with tech, like fintech, logistics and e-commerce all connected together. This includes giving buyers the opportunity to pay when they want, e.g. in 30 days with a business invoice, but instantly sending the money to our suppliers, therefore removing the risk from both sides—a total gamechanger. Our buyers can pick products from tens of different suppliers in different countries and checkout with the click of a button and we handle the rest. Think about ordering work gloves from Germany and overalls from Sweden but getting it all on one invoice. And as a matter of fact, in the last year we've grown to 400 B2B customers which include some of the biggest & coolest companies in Finland including—Valio, Oura, Myllyn Paras, Specsavers, Hesburger, Verkkokauppa.com, HKScan and hundreds more.

The team that chooses Droppe will have the opportunity to learn about cutting edge B2B marketplace development and building solutions used by some of the biggest companies in Finland & Sweden. **Droppe is hiring multiple ambitious developers to continue full time after the project.**

2. Project goals

On the abstract level the project is to build a proof-of-concept-level platform for collecting data on prices and availability for different product groups. The user of the platform could check the price (or other data) for a certain product in the same way as you can check on [flights.google.com](#) if the price for your flight is more affordable than usual. The project has proposed sub-goals as described below as guidance but the team has free hands for it's creative direction.

Suggested optional subgoals:

1. **Crawl & Store**—Platform that crawls & stores the overall top-down situation on the market
   - Crawls certain geographical regions' sellers e.g. all Finnish e-commerce stores and online wholesalers for data such as stock availability, price, demand and other interesting data in real time for specified product categories. Crawl, save in the database, store historical data of the current situation for later use.
   - Awesome way to learn about crawlers & scrapers on the web as well as storing structured data.

2. **Visualise**—Show and make sense of the collected data
   - Gives the ability to visualise and browse the collected data in a user friendly way. Dashboard for sorting & filtering the results. Analyse & visualise the obtained data in a nice UI.
   - Great way to take a deep dive into data visualization and play around with some of the newest libraries & frameworks as well as learn about designing dashboards.

3. **Predict**—Simple predictions that use existing ML algorithms and data from previous stages
   - This subgoal gives the team extremely free hands for exploring & implementation of open source algorithms. For example, trends and patterns in demand spikes could be found from the data and used for future prediction (What happens in a day, a week, a month? Will the demand spread to

other products? Which ones?). This is a far reaching and advanced subgoal for very ambitious teams.
- Awesome way to dip your toes into the field of predictive analytics & machine learning related topics.

3. Technologies

The team can propose the technologies they are most familiar with or which would suit the project goals best. The project is quite full-stack and provides a good opportunity to play with and learn from a variety of domains such as crawling, storage, analytics, frontend implementation, data visualisation, prediction.

The platform should be a browser based application (preferably NextJs / React, Node, etc.). Familiarity with cloud services (AWS, Azure) will be helpful for students. We expect the use of git as version control (GitHub) and proper public CI tools (e.g. Travis/CircleCI) for testing and deploying the software. Code reviews in the form of pull requests. Additionally, the code must be well documented, tested and additional documentation of the system constructed as a part of the project. Some Data Science / Machine Learning experience (or interest to learn PyTorch, Julia or similar tech) is a strong bonus if the team wants to pursue the prediction subgoal.

4. Requirements for the students

1. First and foremost we look for impact driven, motivated and ambitious developers who want to learn in a completely new field.
2. Open minded & flexible, it's very likely in these projects that the scope changes possibly multiple times—vision that extends over the very next project subgoal, an eye for the bigger picture.
3. Working as a team—this is how all successful software project teams have worked.
4. Willingness to question things—what to build and why to build it, ways of working, processes, tedious tasks.
5. Being creative & challenging the status quo

We also appreciate close co-operation with Droppe & stakeholders to lead us all in the correct direction. Project difficulty level is moderate (first stage) / demanding (later escalating to very demanding).

5. Legal Issues

Intellectual Property Rights (IPR): 2. The client gets all IPRs to the results.

Confidentiality: 1. Signing the NDA included in the Aalto's contract template is required. The client might share some confidential information with the students in order to enable the student team to work with enough knowledge to guarantee success in the project. No other legal issues.

6. Client

The team will be given a cloud server to develop and run the application.

Spaces for meetings or working might be available depending on how the corona situation develops in Finland & the current university & company guidelines. Currently all meetings at Droppe are held remotely.

However, we accept that Scrum meetings can take place outside the office hours during the evenings or weekends, if it is more convenient to the team. In addition, the use of hackathon / code sprint / specified office days is greatly encouraged (success oriented approach in previous SWP projects).

Client representative(s)

Henrik Helenius will be the Product Owner & technical stakeholder in the project and can support in web technologies if needed. He is very familiar with the Aalto SWP as he acted as the Scrum Master (and part time developer) in Software Project 2018 and was awarded with the Accenture Software Quality Award for Sievo's Feedbacker Forum.

Henrik Helenius, technical stakeholder
henrik@droppe.fi, +358 440 889 676

Johannes Salmisaari, financial stakeholder
johannes@droppe.fi, +1 424 450 8192

Both Johannes & Henrik are well experienced with the domain of the project. The outcome of the project has real value to our business and we are committed to supporting the project team as required. We believe that communication and staying in sync is key and as such both the student group and we get the most benefit by discussing the project regularly on a biweekly or weekly basis.

Preselected Student Team Members

None

7. Additional information

Technical implementation, documentation, and source code must be written in English.

If you have any questions regarding this project, please do not hesitate to contact the client representatives. As mentioned the Product Owner has participated in this course before, and should be able to answer any related questions you might have. We are happy to discuss the proposal further in a casual setting, for example: henrik@droppe.fi or @henrikhelenius on telegram.

We're actively hiring ambitious talented developers. A good skill set, plain enthusiasm & true will to **solve big problems** with tech is what we appreciate the most. This is a great opportunity for us to get to know each other and for you to learn more about Droppe and the mission-driven environment for curious people we aim to create. **We want you to succeed together with us and therefore a successful completion of the project will be celebrated**.

Find out more about us and our story online at https://droppe.fi/