skyscanner

F I B UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

SKYSCANNER

IN COLLABORATION WITH

UNIVERSITAT POLITÈCNICA DE CATALUNYA (UPC)

FINAL DEGREE PROJECT

Skyscanner Heatmap

Skyscanner's data domain representation

Author:
Fèlix Arribas

Director:
Francisco López
University supervisor:
Maria José Casañ

A Project for the Computer Engineering Degree *in the*

Software Engineering and Information Systems department

Facultat d'Informàtica de Barcelona (FIB)

working with

DeLorean squad *from* Marketplace Engine tribe

Wednesday 10th January, 2018

Universitat Politècnica de Catalunya (UPC)

Abstract

Facultat d'Informàtica de Barcelona (FIB)
Software Engineering and Information Systems department

Computer Engineering Degree

Skyscanner Heatmap

by Fèlix Arribas

In the last century, the world has became smaller. Communications are easier and faster than fifty years ago. Back then, you could talk through a fix phone, but you were not able to send any kind of media, like photos, videos, etc. Only the latest technology of that moment was able to do that. Since the smart phone revolution in 2007 almost everyone can text messages, sending images, share live videos or almost whatever you can imagine in less than a second.

But the internet, phones and communications are not the only thing that made the world smaller. Ways of traveling helped to this earth flattering too. In 1918 visiting another place was very difficult. If you wanted to go through the sea, you had to do it by boat. The fastest way to travel very far in a continent was by train, but not all places were connected with rails. Nowadays, all along with the internet revolution, anyone can travel to the other side of the world in less than a day by plane. Even for traveling inside the same country people use planes.

Is the air industry as efficient as it could? Are all airplane users satisfied with their purchases and possibilities? Skyscanner provides an easy to use tool to search cheap flights from any airport to another. Sadly, sometimes is difficult for users to find what they really want.

This project wants to help solving this problem, providing a heat map to explore differences and similarities between what users search, finds and finally buys and what airlines provides. Being able to compare between specific dates to guess user behavior.

Contents

Chapter 1

Introduction

This is a project developed in *Skyscanner* and evaluated by the *Universitat Politècnica de Catalunya (UPC)* as a Final Degree Project. The main goal is creating a tool for *Skyscanner* providers to ease the comparison of routes by different parameters (for instance, user demand and flights provided), so then the flights advertisement is improved according to user demand. Also being programmer oriented, so the company can develop complex software using the Application Program Interface it will provide.

1.1 Context

1.1.1 Economy

1.1.2 Evolution

1.1.3 Complexity

1.2 Motivation

1.3 Vision

Chapter 2

Market

Chapter 3

Project scope

3.1 Objectives

3.2 Scope

3.2.1 Data warehouse

3.2.2 Visual representation

3.2.3 Other applications

3.3 Risks

3.4 Methodology

3.4.1 Tribes and Squads

3.4.2 Agile

3.4.3 GitLab

3.4.4 Jira

3.4.5 Other tools

Chapter 4

Requirements analysis

4.1 Actors

Initially it seemed difficult to find stakeholders and actors in these project apart from the providers. It is not a tool for the user of Skyscanner so, as explained before, one risk of these project was not finding enough support.

After walking with the Squad Lead and then the Product Own of DeLorean squad a lot of stakeholders appeared: DeLorean Squad, Marketing Automation Squad, State Machine Squad, etc. Each of these stakeholders has different use cases and the project became very interesting for a considerable part of Skyscanner .

4.1.1 DeLorean Squad

The mission of DeLorean squad is to provide the best data and services around the routes, timetables and modes of transportation to go from one point on Earth to another.

Right now, DeLorean squad provides a very fast service that serves flights logistic information between a given origin and destination. Some information you can find in a route is the fight number, carriers, stops, date ranges, etc. These flights are just **single ticket flights**. The squad is currently working on more complex routes, constructed routes[1].

The constructed routes timetable is under construction, so is not available for these project. The Heatmap relies on the timetable of Single Flight Number service, also known as Timetable SFN Service.

Timetable SFN Service

The *timetable SFN* endpoint returns details for time tabled Single Flight Number itineraries series. Note that SFNs are not ticket-able, so they do not include itineraries which cannot be bought on their own, neither the price nor restrictions.

Timetable SFN Service provides all the **current** flights. This is a little bit of a problem when trying to get past routes: Timetable SFN Service does not provide past flights information, it is always **up-to-date**. In order to get this data it is needed to go one step back in the whole DeLorean data processing: Timetable Ingest Pipeline.

---

[1]Contructed Timetable contains constructed routes, routes composed by two or more single ticket flights.

Tech specifications of Timetable SFN Services are explined in section

Timetable Ingest

This phase, basically collects all the OAG[2] from a provider and maps it into routes
in JSON[2] format. For each different version of the OAG file, the pipeline creates a
new file with all the routes.

Then, in order to get past routes, the heatmap reference in old versions of the file
created by the Timetable Ingest Pipeline.

Tech specifications of Timetable Ingest Pipeline are explined in section

4.1.2   Product Owner

Jen Agerton is the Product Owner of DeLorean Squad,

4.1.3   DeLorean's Squad Lead

Francisco López, who is also the supervisor of this project, and I has the initial idea
for this project. He oriented it for a Machine Learning purpose.

4.1.4   Neil Ross

Neil Ross is currently the Squad Lead of 4.1.4, this contact was given to me from my
4.1.2 Jen Agerton.

After some talks through Slack[3], he forward me extracts of the email he sent when
he was in RaTS Squad, regarding to their use cases. Most of the use cases have been
used for this project.

Fuel RaTS Squad

Routes and Timetable Servies Squad provides the best data and services around the
routes, timetables and modes of transportation to go from one point on Earth to an-
other. Fuel RaTS has the same mission as DeLorean Squad, but develop different ser-
vices. Since Fuel RaTS provides basic routes data, pricing, live update information
and multi-destination combinationcs, DeLorean squad provides a very fast service
for only routes.

Marketing Automation squad

Marketing Automation squad enables scalable growth by automating workflows,
and the collection of insightful data. They have three main goals:

- Provide data to support decision making

- Automated, data driven campaign management

- Budget process automation

---

[2]OAG file (also know as WTF file), is a CSV[1] file which each row represents a timetable for a
Single Flight Number.

MAS provides fourteen

4.1.5 Providers

4.1.6 Competitors

4.2 Functional requirements

4.3 Non functional requirements

4.4 Use cases

Bibliography

[1] Y. Shafranovich. *Common Format and MIME Type for Comma-Separated Values (CSV) Files*. 2005. URL: https://tools.ietf.org/html/rfc4180 (visited on 01/30/2018).

[2] JSON.org. *The JavaScript Object Notation (JSON) Data Interchange Format*. 2017. URL: https://tools.ietf.org/html/rfc8259 (visited on 01/30/2018).

[3] Slack Technologies. *Slack: Where Work Happens*. 2013. URL: https://slack.com/ (visited on 01/30/2018).