Contents

[Making IsThisTaxiLegal application 1](#_Toc503847185)

[Making Proof of Concept with data 1](#_Toc503847186)

[Step 1 : acquiring data from 1](#_Toc503847187)

[***Step 2*** : cleaning data - read with Word and transform into CSV. 2](#_Toc503847188)

[***Step 3*** : Making objects to support 2](#_Toc503847189)

[***Step 4:***  Consolidating data from CSV to objects 2](#_Toc503847190)

[***Step 5***: Application Web + WebAPI to display data 2](#_Toc503847191)

[***Step 6***: Deploy sources to GitHub 2](#_Toc503847192)

[***Step 7***: Create an application visible on internet 2](#_Toc503847193)

[***Step 8***: Document the API 3](#_Toc503847194)

[***Step 9*** : Remake the documentation 3](#_Toc503847195)

# Making IsThisTaxiLegal application

# Making Proof of Concept with data

Finding PDF about taxi licence for Bucuresti : PDF at <http://www.pmb.ro/adrese_utile/transport_urban/autorizatii_taxi/autorizatii_TAXI.php>

Thinking about application + web site about this -with a simple search and possible OCR ( take image of the car plate and recognizing number) . Also for international ( GPS for phones to know the city). Also IOT for monitoring the illegal taxi movement( could be also illegal to do this)

The Proof of Concept: A WebAPI + WebSite( that can be accessed also by mobile) to enter a plate number and find if it is legal. Maybe later an Android App.

Step 1 : acquiring data from <http://www.pmb.ro/adrese_utile/transport_urban/autorizatii_taxi/autorizatii_TAXI.php>

. Download situatia\_autorizatiilor\_taxi\_20171208.PDF

## ***Step 2*** : cleaning data - read with Word and transform into CSV.

Trying <table>. ConvertToText - not good, it preserves the return character and data can not be read safely after that.

The solution: read row by row and replace CR/LF with empty for each cell.

After this, problem with Bell character – replace this also the 7 character

After this, problem with repeating headers for each table in the csv

This is the VBA code

( maybe doing same in R : <https://datascienceplus.com/extracting-tables-from-pdfs-in-r-using-the-tabulizer-package/> )

## ***Step 3*** : Making objects to support

First creating objects to support this . Car, City, LicenseState, Licensee, TaxiAutorization.

Creating test - the City should be unique - even if Bucarest is multiple times, the City object should be the same.

## ***Step 4:*** Consolidating data from CSV to objects

Creating objects to mimic CSV data . Return to step 2 and put separator | instead of ,

Problem with parsing data

- a date could not exist, so it will be nullable

– could be either dd.mm.yyyy , either d.mm.yyyy either 27.02.202

Separate the lines with errors from the lines without errors . And return the result as a tuple

Making test in order to see the errors

## ***Step 5***: Application Web + WebAPI to display data

Making WEBAPI for seeing all and some taxis

See that valid taxis are not parsed correctly ( Validat vs Valid) . Modify test

Discover there are some licenses with state…

## ***Step 6***: Deploy sources to GitHub

Easy to do – create an application on GithUb, connecting with Visual Studio, submit and sync

## ***Step 7***: Create an application visible on internet

You can create an account at appharbor.com , integrate with GitHub and deploy there : <http://isthistaxilegal.apphb.com/API/Home/Index>

## ***Step 8***: Document the API

Swagger / Swashbuckle is the easy way to do this.

## ***Step 9*** : Remake the documentation

Mention all the documentation in all places ( API, GitHub, others)

Mention the contact name if something is wrong

Add API for enums

# Thinking about application

## What remains to be done / how can be improved

It can be improved so many ways:

* Making OCR ( to recognize the plates number)
* Making application Android / iPhone
* Making Windows (Store ) application ( not really necessary, but anyway)
* Finding if another cities have taxi licenses available
* Publicity
* Improving with ratings for taxi drivers
* Find a lawyer that knows what to do next
* Improving the documentation
* Passing from csv to sql lite
* Downloading automatically the pdf with licences from [www.pmb.ro/adrese\_utile/transport\_urban/autorizatii\_taxi/autorizatii\_TAXI.php](http://www.pmb.ro/adrese_utile/transport_urban/autorizatii_taxi/autorizatii_TAXI.php)
* Making all this points on <https://github.com/ignatandrei/IsThisTaxiLegal> as issues
* Running the test automatically