Abstract – 200 words

Aim and Objectives

Provide a solution to the business analysts of Yusen logistics that help them to efficiently and accurately generate straightforward reports in shorter amount of time and show the business performance by visualise main KPI indicators for their customers and suppliers.

Motivation for the project

Why developing web based application

Flexibility

Keywords: web application development

The method for achieving this aim is creating a web based application running on local host which can be used to pull out data and generate reports in PDF format. This application enables users to obtain KPI indicators automatically and generate graphs for further analysis. Through the development of this application individuals would now be able to easily and quickly generate reports and deliver it to their customers and suppliers further their understanding of business performance.

Introduction

It is very crucial that know your business better. At the same time let your customers see your performance and improvement so that can boost the confidence of your clients and win a stable and long term business relationship. That is called business intelligence.

The report shows four value driver areas such as financial value, productivity, risk and trust. clarify general business objectives and corresponding performance metrics and indicators. define achievable targets and use the tools to inform the decision-making processes.

KPI indicators are essential and important for users and their clients to know the business better. While time spend on generating such report should not take long. Efficient way of Benefit from technology automation let the application helps to do the repetitive work for you.

KPI dash board can offer users an intuitive way to review the business performance and even could help to do an accurate and practical projection for the company.

In order to offer some straightforward feedbacks and to improve the communication with Yusen’s clients, Yusen logistics is looking for an efficient and automated solution to generate monthly, quarterly or annual report based on user’s choice in shorter time. The report would show required KPI indicators and performances of related hauliers. In order to present the feedbacks to their customers, Yusen is looking forward to an auto generated PDF format or similar report in the end.

At the moment, Yusen’s analysts using excel with VBA function that creating macros to produce locate the data and generating graphs then save it to PDF format. Due to all the business documents come from different countries and they have not yet unified the reporting language, manual organizing the data is time wasting. However, for this problem cannot be solved at this stage. Apart from different languages reporting, it is also time consuming for analyst to locate history data then generate longer or random time span reports.

Requirement:

An excel based or ERP system supportive application that can be implement on computers in Milton Keynes office.

Create a report template with same format of current reporting format that made with excel.

Reduce time of creating the report. Analysts now from Yusen now need 2 to 4 days to generate the report due to large amount of manual typing. Though it is evitable for the moment, they are looking forward to shorten the time of generating the relative graphs.

Generate report by customize time span easily and flexibly for instance user can generate monthly, quarter, annual report or other time periods they want.

Clients would like to have menus to choose related business performers, markets, MMW and month.

Clients want to have a click view section that they can hover to see over on the district map that shows some data of different area of one particular market.

No extra costs for creating the application.

Ideally the application could be benefit both goals of short term and long term.

My solution

Initial idea

Creating a web application automatically pull out the certain data that match user’s requirements. Using python as the programming language with assists from other open source python library like pandas for data manipulation and matplotlib for data visualization.

This KPI dashboard generating web application will be written with python programming language and flask for building the web framework allows the application works remotely or on the local host. Since Yusen’s updating the system, this application will run locally without internet. The application will get HTTP requests which is users’ input in this case, and dispatch code that generates HTML which is rending the customize html template through the flask framework, and creates an HTTP response with that content in this web application is to write the response to a csv file then save it locally.

Rationale for developing web based application

For Yusen logistics, it is essential to save time and costs on generating report. With the data sets get bigger, it is difficult to find the history data or even some current monthly data. Analysts would have to look through the whole file or using excel filtering and pivot table to get the data they need. And along with the growing files, macros in excel will experience serval crushes and will need more downtime to recover the system.

While using application, users will find it more flexible and much easier to generate the report since all the data selection choice is predefined also users can choose a random time span, different hauliers and the other KPI indicators of hauliers.

According to Yusen, their working system is updating and tending to create a click view reporting system in the future.

With application

Design

Read the original report

Get to know the format for the report.

It requires to show key highlights of current month, including MMW performances of different hauliers, performances of IMC in different markets along with procurement and quality and those makes the first part of the report.

In second part of the report, line chart will show planning performance of all the linehauls through 12 months until current month. And a table will be shown to present the 12 KPI indicators of hauliers separately from current month. Also the report will present the reasons of late arrivals, dispatches and late 1st hub arrivals, time of late arrivals/dispatches by either bar graph or line chart. Then the trending of arrivals, dispatch and 1st hub arrivals of different hauliers will be presented by different line charts. At the end of this part, 3 lines charts will be presented to show the difference between planned and actual vehicles they used, the average LFF per haulier and average pallets they used during the transportation.

In the third part of this business report, it focuses on IMC’s performances of different markets. It will compare IMCs’ number of outbound deliveries, number of handling units, total weight, number of returns, number of pallets and number of parcels that IMCs’ carried. Also it presents the percentages of the all the IMCs on time data, carrier- related on time and fully delivery data and OT and IF target data. Finally, it will show IMC s’ spending and their special costs pattern.

The report will end with projection of Yusen for the next business time periods.

Get to know the details of charts

Chart type

The report includes line chart for showing the trending, bar graph and pie chart.

Chart details

The charts show different KPI indicators in several parts. Information that are going to be presented including time span and number of hauliers. Those two factors vary from part to part through the whole report. In some of the charts, only data of current month is required however it requires a full year data in any others. Applying different numbers of hauliers happens as well since part of the report only need the data of all the hauliers while the others expecting 4 or 5 hauliers each time.

Project management

In order to achieve the developing goals on time, applying project management will be helpful to track the progress in each stage.

Development planning

Stage one: Jun – July

Understanding the clients, looking into multiple solutions and programming language and tools for developing the application.

Communicating with clients frequently insure that requirements are fully understood. And offering a projection about its functionalities that can be applied by the end of the project.

understand the data sets including the source customer requirement and do research for proper programming tool and learn to use necessary libraries to assist in developing the application.

After doing research about tools for data extraction, manipulation and analysis online, I found out that python library Pandas is one of the most useful tools for fast data analysis and manipulation*.* Pandaswith powerful built in functions that enabling users to carry out the entire data analysis workflow in Python. In this application development process, Pandas would be the main tool to make data processing easily and automatically.

Stage two: July -August

data manipulation, plot graph and final report generation

The strength of Pandas seems to be in the data manipulation side, but it comes with very handy and easy to use tools for data analysis, providing wrappers around standard statistical methods in [stats models](http://statsmodels.sourceforge.net/).

It is also handy to use read\_csv or read\_excel function from python Pandas library to pull out the required data from excel files. And then applying graphing methods of [matplotlib](http://matplotlib.org/) to the plot the charts clients required.

At this stage, based on the report requirements including writing time periods, KPI calculation formulas and total number of each performance that carried out by different hauliers or overall perforce of them all, certain built in functions are applied to filter, do simple calculation and separate or group the data sets.

Stage three:

Integrate the user interface

Creating relative html static pages to get and store user’s selection or input data

Using Gantt chart to manage the development process

During the developing stage, Gantt chart will be used to document and schedule the project for it is good at showing timing and progress clearly which helps project tracking.

In the KPI project Gantt chart, it shows key events, duration and milestones. Along with more developing issues shown up, recording the problems raised during the developing process and possible solutions to the issues.

Also keep more details logs about problems and solutions to each part of the codes on jupyter notebook for later review and update.

GitHub version control

Communication with customers – make sure fully understand the requirements

Onsite

Email

After onsite communication and read through the current report template, new questions related to the report raised. and later then solved by Questions, confirm details

Telephone

Suggestions to customers that could make the application work and the reason expected result of this application.

Communication and discussion with supervisor

Achievement each week

Problems encountered

When step into the application development, at the first stage of developing circle, using UML to structure and visualize the requirements of application.

UML

Method for development

Why Agile

Small project

Application

Operating environment

Windows 10 laptop

Spyder or windows command line

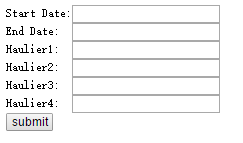
Programming with Python version 3.5 and python library including Pandas, Matplotlib for visualize, web application framework like flask and html template like Jinja2.

Code explanation

Taking codes for generating MMW OT Arrival Trend Germany, %, MMW OT Despatch Trend Germany, % and 1st Hub OT Arrival Trend Germany, % trending charts as an example. For this action, html pages, python and a main flask script will be applied.

Creating a form for users to input required information in a static html page. Bonding the http GET method with the submit button to store the inputs from users to a local csv file. In html script, using form tag, input tag, p tag and label tag to structure the input area. And apply CSS script to style the user input page and align the input to left.

Execute MMW.py in virtual environment first. Enter <http://127.0.0.1:5000/> in the web browser to get to the input page. Users can type in up to 4 hauliers in any preference order at the same time and customize time span in star date and end date form. After click the submit button, if the submission was successful, users will be lead to an ok page that confirms the success of collecting data from the input form. All the inputs from users will be stored in a csv file called MMW.csv in the same folder with MMW.py script.



Users’ inputs are processed by http method ‘GET’ and writer function from csv module in the python script. CSV writerow helps to write the inputs to MMW.csv in one row so that we could easily locate the inputs for later data manipulation.



Executing the MMWChart.py script to access the MMW.csv file to get the hauliers and the time period they want to show and then to generate trending charts for each hauliers.



In the first part of MMWChart.py, using Pandas library to read user inputs csv file MMW.csv and copy of super template kpi.csv file. By using df. iloc function locating the certain dates and names of hauliers in template file and read them into Pandas pkl format by for later plotting.

In order to plot multiple hauliers in one chart, concat and group by are used to group the data together. For keeping the charts clear, legend is added to present each hauliers by different colours and the legend is set to automatically adjust its position wherever it suits the best. In the end, plotting the data by using plot function and save them to statics folder for later use.

Problems with the doc and code:

Cannot plot the second input haulier.

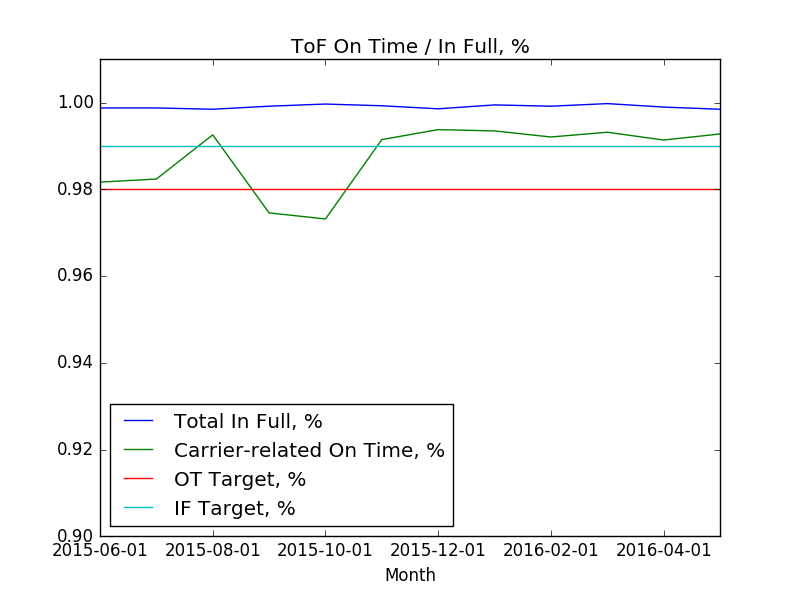
MMWChart.py

Legend needs to relocate. When save the charts figures the legends stay in the middle of the graphs.

On Time to 1st Hub, % and average LFF indicators cannot be plotted. Has not yet found the reason.

TOFTMONtime.py

In this python script, it produces on time performance of different IMCs. Once the user inputs dates and names of IMC, the application would locate the eligible contents from csv files. User could either get the inputs directly running the program at back end or get information from html page then read the csv file that stores user inputs and locate relative data from in market csv file then start to plot.



And relative tables would be created and added to html format report via jinjia2 template.

Problems with doc and code:

IF and OT target data need to be filled in excel table or the line will not show in the chart.

Axis cannot be auto adjust by different range of data each time periods. Or if using default y axis then the largest figures would be the maximum number of the axis and results in the charts become too full.

Once the time span goes over 12 months, the x axis would not show the months fully rather than show the month of every two months.

imc.py

Fuel surcharge should be filled in the files or it will report error.

Cannot plot special cost and plot the overall results.

Form.py

Back-end

In input html script, the code can only add 3 of the selected objects to the array sometimes including un wanted ones.

Front -end: html page

Design

Introduction

Simple instruction

Might not render in new pages

user Interface:

design

how it should work

problems

drop down menu to select

output the report to PDF

testing

method

problem

cannot render more than one html that generated by Pandas in the report page.

outlook and improvement for the application

improvement based on current application

create another button allow users to add more hauliers and IMC they need.

In some circumstance, it would be possible that Yusen uses more than four hauliers sometime. Adding more haulier

improving with how the data are stored database, cloud

Learning outcomes-reflection-has developed insight and understanding

Reflection on project

Structure the code as soon as possible

Reflection on learning outcomes

Learning new tool for web application developing, consolidate and apply the knowledge learned through the year to practice. Apply theories of soft engineering developing to this application development. Learn to use libraries to help develop the application

Instruction for users

In order to run the application, users can download Anaconda python environment from <https://www.continuum.io/downloadsor>. The anaconda has already included both pandas and flask library. Or using windows command line directly then install those libraries via pip in virtual environment. Under this circumstance, users have to install python3 or advance version get the pip.

how to install essential tools

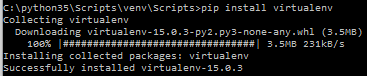
python:

download compatible version of python from <https://www.python.org/downloads/>

click the package to install the python in suggested directory

virtualenv:

for later installation and running of flask, virtual environment needs to be installed through pip. Find your pip python script in the python directory then run *pip install virtualenv* in command line.



install flask:

after installing the virtualenv module, users need to find the activate command script in virtualenv folder to activate the virtual environment. Once users see the (venv) at the front of the directory, flask can now be installed via *pip install Flask.*





install pandas:

Stay with the pip and run pip install pandas.



install matplotlib:

run pip install matplotlib.



The program would perform slower after times of running with anaconda or running scripts directly via windows command line, It happens to both anaconda has already had those libraries as package in the its system,

how to use the application

start command line on user’s computer.

Start the virtual environment

Running the python scripts in following order.

Precautions:

Before starting to extract the data, save the sheets of excel file to separate csv files and renames the files to line haul and in market. Or users can read the excel spreadsheets directly by using read\_excel function then save the them to csv files.

In saved csv files, changing the default format of month column from JUN-2015 to 2015-06-01. And change the columns with percentages (99%) to float number (0.99).

where they can keep original file and be ready for exacting by the application

save all the file in the same folder and create related subfolder for instance under the main KPI folder, creating a statics folder for storing generated charts and template folder for html written report template.

important message:

month data format

float number format

blank cell

Problems:

Flexibility

Performance

Command line needs to be restarted after 20 times running.

Cost

Code not flexible. Could choose different time span but cannot choose number of hauliers or IMC.

Run at back end cannot run from user interface for now.

Reference

<http://machinelearningmastery.com/quick-and-dirty-data-analysis-with-pandas/>

<http://pandas.pydata.org/index.html>

<http://pandas.pydata.org/pandas-docs/stable/merging.html>

<https://virtualenv.pypa.io/en/stable/installation/> install

<http://flask.pocoo.org/docs/0.11/installation/>

<https://github.com/ehmatthes/pcc/blob/master/chapter_15/README.md#checking-if-matplotlib-is-already-installed>