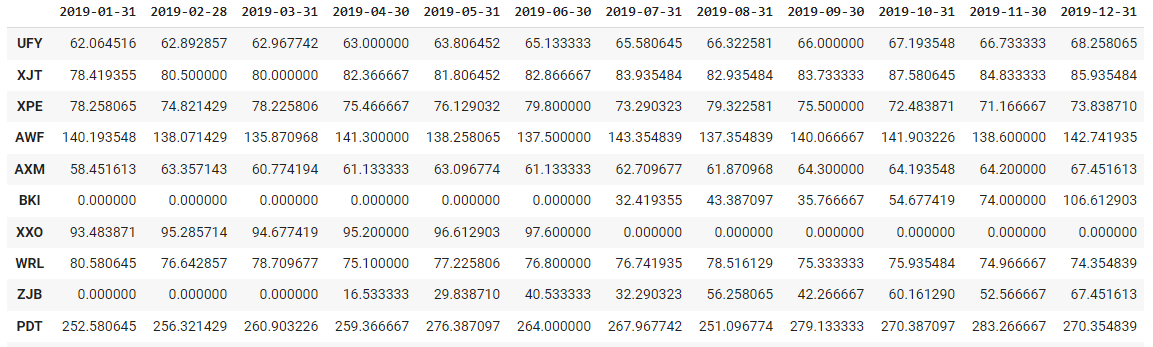
# Introduction

Data visualization is a way to represent the datas/informations in graphical view. There has several type of insight to show this graphical view. These are: Chart, Column Chart, Bar Graph, Stacked Bar Graph, Stacked Column Chart, Area Chart. Dual Axis Chart, Line Graph, Mekko Chart, Pie Chart, Waterfall Chart, Bubble Chart, Scatter Plot Chart, Bullet Graph, Funnel Chart, Heat Map, etc.

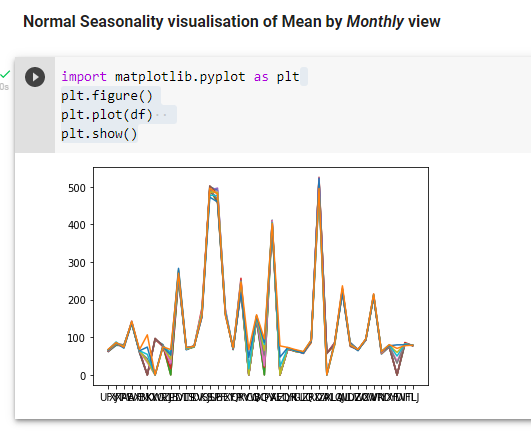
# Discussion

1st I read the daily visitors csv. And then stored in a new dataframe. Then found they have 40 unique Identifier and given a date column of 365(1 year). Then I compress the date to month. So total month is 12. And calculated MEAN. Then converted the rows(Date) to Columns and ID to rows.(Figure 1)



(Figure 1: Top 10 rows of Daily visitors monthly)

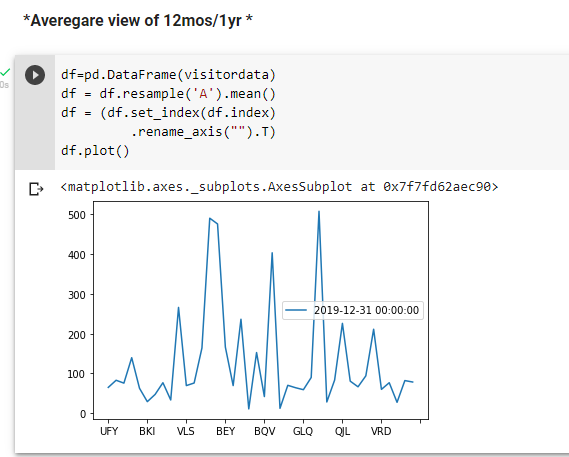
The reason of making monthly is to read easily and make visualizations very easily. And it will be very easy to find the progress of each month.



(Figure 2: Normal Seasonality visualization of Mean by Monthly view)

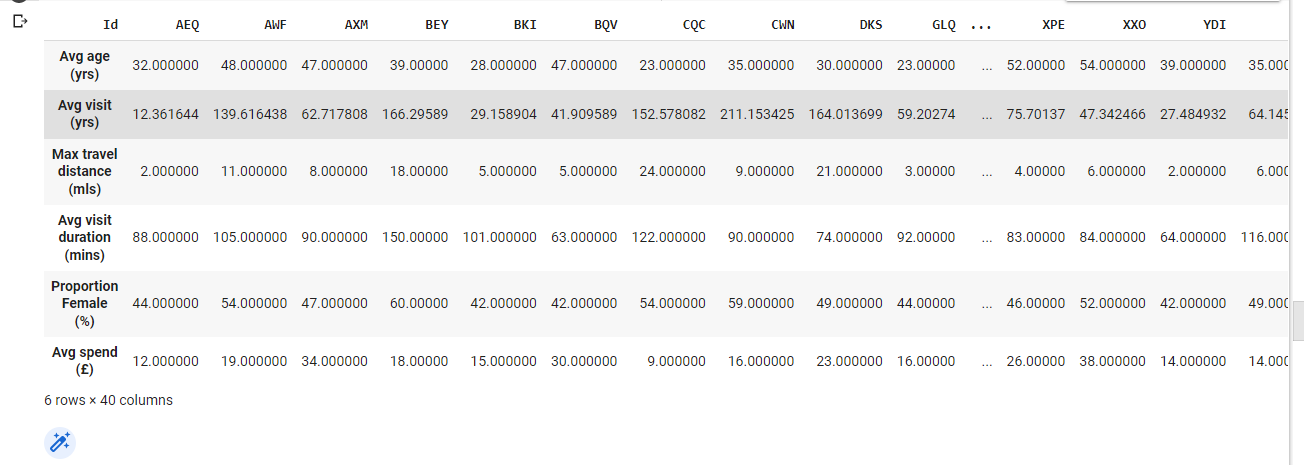
It will show all Venue visitors monthly average into this chart (Figure 2).

It is showing very ugly (Figure 2) but we can understand the progress of it. So, to fix it, I have used from 0 - 12 months (1 year) Average. So, now it will look beautiful and we can easily understand each venue Average (Figure 3).



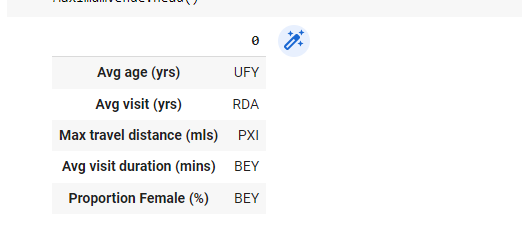
(Figure 3: Average view of 12mos/1yr)

### Summary Data

For summary I read all the others CSV. And then stored all into 2nd Dataframe. Then it became 6 rows x 40 columns. (Figure 4)

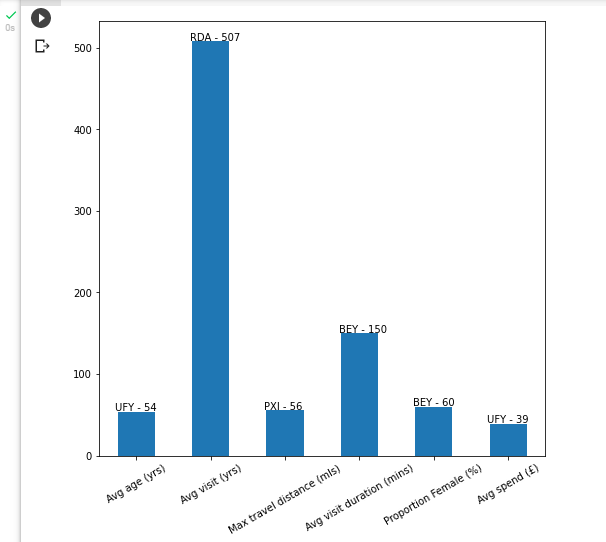
(Figure 4: Gathering all Datas into one DataFrame)

And the datas here are average. For the visit, I have calculated yearly Average (Figure 4). Then I calculated Maximum status for all the Venue ID for those 6 rows (Figure 5).



(Figure 5: Maximum status for all Venue ID for each summary.)

We need this because we are going to show the maximum status in chart. Otherwise it will be very messy with much data. So, the output looks like this:



(Figure 6: Maximum status for all the Venue ID for summary in graph)

# Critical review

I learnt Pandas, DataFrame, Patch(For setting color), Matplotlib, plot, read\_csv, resample, DataFrame mean, indexes, set\_indexes, rename\_axis, Append, pivot\_table, columns, reset\_index, iloc, loc, max, Average, to\_frame, tolist, head, get\_bbox, annotate, xticks etc.

I faced lots of error because of converting Rows to Columns and Converting columns to rows. But at the last I did it successfully.

# Data conclusions

The data are of 2019. And the company has 40 Venue.

* The average age (yrs) is 54 of UFY.
* The average visit (yrs) is 507 of RDA.
* The max travel distance is 56 miles of PXI.
* The average visit duration is 150 mins (2hrs and 30 mins) of BEY.
* The Proportion Female 60% of BEY.
* The Average spend is 39 euro of UFY.