Findica Tech provides Software-as-a-service (SaaS) Customer Experience Management and software solutions to Restaurants, Hotels, Retail and Financial Services in the United States. The product “Gorilla.net” utilize sentiment and text analytics to automatically collect, categorize and report on structured and unstructured data. The product uses Big Data Eco System to process the large volumes of the data, Natural Language Processing and Machine Learning techniques to transform customer feedback into actionable insights and it acts a solution for the businesses to reduce churn rate that results in customer’s retention and increased revenue.

RESPONSIBILITIES:

1. Get URLs from yelp website through YahooBoss API
2. Parse Yelp reviews through Jsoup HTML parser
3. Apache POI to read and write excel files
4. XML parsers used to get long,lat of the location (SAX,Xpath)
5. Used Google API to get nearby locations

STARBUCKS vs CoffeeBean&TeaLeaf

1. Number of customers talking about the brand in the USA/PER STATE/PER CITY

Vs Total Population in the USA/PER STATE/PER CITY and the Customers Sentiment.

1. Number of Customers visiting the Brand on the DAY/HOLIDAY per STATE/CITY
2. Most Consumed Item among the MENU (FOOD , DRINK) / Most Negative Item among the MENU / correlate with most talked attribute (place, service)

DRINK - COFFEE, TEA, BEVERAGES

FOOD –

1. PRICE Perception on the BRAND/MENU
2. FOOD menu consumption among the customers
3. INTENT TO RETURN / if the customer concerns were considered % of INTENT TO RETURN

Read/write excel files using Apache POI

<https://gist.github.com/madan712/3912272>

COFFEE BEAN AND TEA LEAF LOCATION LIST USA

<http://locations.coffeebean.com/sitemap.html>

AZ – 4 (11) , CA- (250), HI- 3 (20),NV-7 (19)

STARBUCKS USA LOCATIONS COUNT IN EACH STATE

<http://www.businessinsider.in/Heres-The-Starbucks-Map-Of-America/articleshow/34111965.cms>

Get latitude and longitude for a location

<http://www.santhoshreddymandadi.com/java/java-program-to-get-latitude-longitude-points.html>

code from the below url worked fine

<https://javaho.wordpress.com/2011/07/03/server-side-sample-code-to-get-latitude-and-longitude-of-a-given-address-using-google-map-api/>

JSON PARSING TO JAVA

<http://examples.javacodegeeks.com/core-java/json/java-json-parser-example/>

READ EXCEL FILES

<http://stackoverflow.com/questions/1516144/how-to-read-and-write-excel-file-in-java>

R yelp analysis

<https://github.com/Yelp/dataset-examples>

MAXIR - <https://github.com/minimaxir/yelp-review-analysis/blob/master/yelp_review_analysis.R>

<http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>

<http://trendct.org/2015/04/03/using-python-and-r-to-pull-and-analyze-meetup-com-api-data/>

<https://rstudio-pubs-static.s3.amazonaws.com/124656_e2e77362772d40a4a7be244d2f114853.html#/4>

<http://cs229.stanford.edu/proj2014/Chen%20Li,%20Jin%20Zhang,%20Prediction%20of%20Yelp%20Review%20Star%20Rating%20using%20Sentiment%20Analysis.pdf>

<https://rstudio-pubs-static.s3.amazonaws.com/124656_e2e77362772d40a4a7be244d2f114853.html#/>

KEVIN - <https://github.com/kevin11h/YelpDatasetChallengeDataScienceAndMachineLearningUCSD>

R with MongoDb

<https://cran.r-project.org/web/packages/rmongodb/vignettes/rmongodb_cheat_sheet.pdf>

<http://www.r-bloggers.com/get-an-r-data-frame-from-a-mongodb-query/>

R – Java Integration

<http://stackoverflow.com/questions/7451716/java-r-integration>

R-Solr

<https://cran.r-project.org/web/packages/solr/solr.pdf>

R-Data visulaization

<http://www.revolutionanalytics.com/r-language-features-graphics-and-visualization>

<http://www.r-bloggers.com/playing-with-r-shiny-dashboard-and-google-analytics-data/>

R useful packages

<https://support.rstudio.com/hc/en-us/articles/201057987-Quick-list-of-useful-R-packages>

R 3d plots (IMP with code)

<http://www.rforscience.com/rpackages/visualisation/plot3d/>

<https://cran.r-project.org/web/packages/scatterplot3d/vignettes/s3d.pdf>

<http://pj.freefaculty.org/guides/Rcourse/plot-3d/plots-3d.pdf>

R US map

<http://stackoverflow.com/questions/27328372/r-create-a-spatial-bubble-plot-that-overlays-a-basemap-of-the-us-and-other-spati>

Sample MT cars data

<http://www.r-bloggers.com/getting-fancy-with-3-d-scatterplots/>

<http://pj.freefaculty.org/guides/Rcourse/plot-3d/lattice.example.R>

<http://rstudio-pubs-static.s3.amazonaws.com/15358_25737b089f0042e086e68b1609e4f20b.html>

plotly plots

<https://plot.ly/~MattSundquist/folder/home>

D3 js

<http://www.delimited.io/blog/2013/12/19/force-bubble-charts-in-d3>

<http://vallandingham.me/bubble_charts_in_d3.html> - blog

<http://blog.savageevan.com/> - blog

[**https://github.com/romsson/dragit/blob/master/example/test\_events.html**](https://github.com/romsson/dragit/blob/master/example/test_events.html)

<http://www.visualcinnamon.com/2015/10/different-look-d3-radar-chart.html> - blog

<http://www.visualcinnamon.com/2015/11/learnings-from-a-d3-js-addict-on-starting-with-canvas.html> - blog

<https://gist.github.com/nbremer?page=5> – blog

<http://christopheviau.com/d3list/> - blog

<http://stackoverflow.com/questions/12944024/d3-js-how-can-i-prevent-circles-text-overlapping>

<http://bl.ocks.org/ChrisJamesC/4474971>

<https://www.npmjs.com/package/react-bubble-chart> (react bubble chart)

<http://bl.ocks.org/mbostock/1249681> (multi-foci force layout)

<https://jsfiddle.net/ym04vunh/5/> (collision bubbles)

<https://jsfiddle.net/cyril123/LxfLv3wt/4/> (collision bubbles with text )

, <http://stackoverflow.com/questions/33928409/how-do-i-add-labels-to-d3-js-force-bubble-chart>

THREE JS

<http://mrdoob.github.io/three.js/examples/#canvas_interactive_particles>

VALID JS figure html pages for Reports

<http://www.localhost.com/score/starbucksD3Copy1.html>

[http://www.localhost.com/score/PlotlyStarbucks.html](http://www.localhost.com/score/starbucksPlotly.html)

<http://www.localhost.com/score/starbucksD3Copy2.html>

<http://www.localhost.com/score/starbucksD3Bubble1.html>

<http://www.localhost.com/score/PlotlyChipotle1.html>

<http://www.localhost.com/score/PlotlyChipotle.html>

stoccharts (pyramid charts)

<http://www.stocinn.com/stoccharts/js/stocChart.js>

<http://jsfiddle.net/akashtyagi40/fptv739g/3/>

CYGWIN LOCAL TO REMOTE  
scp -r <user\_name>@<host\_name or host\_ip>:<Path to your directory> <target\_path\_dir>

ERRORS

devtools::install\_github("ropensci/plotly")

Downloading GitHub repo ropensci/plotly@master

Error in curl::curl\_fetch\_memory(url, handle = handle) :

Peer certificate cannot be authenticated with given CA certificates

Solution : library(httr)

> set\_config( config( ssl\_verifypeer = 0L ) )

> devtools::install\_github("ropensci/plotly")

R + SOLR:

K-means CLUstering

<http://stackoverflow.com/questions/20446053/k-means-lloyd-forgy-macqueen-hartigan-wong>

MERGE DATAFRAMES

<http://stackoverflow.com/questions/1299871/how-to-join-merge-data-frames-inner-outer-left-right>

How can I do database style, i.e., [sql style, joins](http://en.wikipedia.org/wiki/Join_%28SQL%29)? That is, how do I get:

* An [inner join](http://en.wikipedia.org/wiki/Join_%28SQL%29#Inner_join) of df1 and df2:  
  Return only the rows in which the left table have matching keys in the right table.
* An [outer join](http://en.wikipedia.org/wiki/Join_%28SQL%29#Outer_join) of df1 and df2:  
  Returns all rows from both tables, join records from the left which have matching keys in the right table.
* A [left outer join (or simply left join)](http://en.wikipedia.org/wiki/Join_%28SQL%29#Left_outer_join) of df1 and df2  
  Return all rows from the left table, and any rows with matching keys from the right table.
* A [right outer join](http://en.wikipedia.org/wiki/Join_%28SQL%29#Right_outer_join) of df1 and df2  
  Return all rows from the right table, and any rows with matching keys from the left table.
* By using the merge function and its optional parameters:
* **Inner join:** merge(df1, df2) will work for these examples because R automatically joins the frames by common variable names, but you would most likely want to specify merge(df1, df2, by = "CustomerId") to make sure that you were matching on only the fields you desired. You can also use the by.x and by.y parameters if the matching variables have different names in the different data frames.
* **Outer join:** merge(x = df1, y = df2, by = "CustomerId", all = TRUE)
* **Left outer:** merge(x = df1, y = df2, by = "CustomerId", all.x = TRUE)
* **Right outer:** merge(x = df1, y = df2, by = "CustomerId", all.y = TRUE)
* **Cross join:** merge(x = df1, y = df2, by = NULL)

Machine Learning in Natural Language Processing using R

<https://ufal.mff.cuni.cz/mlnlpr13>

QLIK MAPS

<https://community.qlik.com/docs/DOC-6932?_ga=1.96003743.2069078398.1447840175>

HELPFUL DATA SCIENC E BLOG (USES PYTHON)

<http://www.racketracer.com/2015/03/02/predicting-and-plotting-crime-in-seattle/>

PYTHON

PLOTLY/IPYTHON

<https://plot.ly/python/ipython-notebook-tutorial/>

<https://pyscience.wordpress.com/2014/09/02/interactive-plotting-in-ipython-notebook-part-22-plotly-2/>

<https://github.com/ContinuumIO/anaconda-issues/issues/494>

<https://ipython.org/ipython-doc/1/interactive/public_server.html#running-a-public-notebook-server>

DIFFERENT PLOTS

<http://stackoverflow.com/questions/19949435/3d-plot-of-bivariate-distribution-using-r-or-matlab>

ERROR in ECLIPSE

D:\softwares\workspace\.metadata\.plugins\org.eclipse.wst.server.core\tmp0\wtpwebapps

Remove the webapp and re run on the server

STARBUCKS vs COFEE BEAN AND TEA LEAF

SAMPLE REPORT 1: (OVERALL BRAND MATCHUP)

1. START WITH THE BRAND SENTIMENT
2. MOST POPULAR COFFEE WITH THE SENTIMENT AND TOP MENU ITEMS OTHER THAN THE COFFEE
3. CATEGORY MATCHUP (FOOD,SERVICE,AMBIENCE,PRICE/VALUE FOR MONEY,DISCOUNTS)
4. LOYALTY (% INTENT TO RETURN)
5. IF SERVICE WAS GOOD %INTENT TO RETURN MIGHT BE IMPROVED (HOW CAN BE DONE?? ‘R’)

CHANGES TO REPORT1:

1. START WITH THE BRAND SENTIMENT
2. TOP 5 POPULAR COFFEE/TEA ITEMS WITH THE SENTIMENT
3. WHICH COFFEE IS GAINING POSITIVE SENTIMENT OVER THE PAST TWO YEARS/LOSING ITS SENTIMENT OVER THE PAST TWO YEARS.
4. THE CAUSE OF GAINING/LOSING ITS SENTIMENT (TASTE, FLAVOUR, QUALITY)

SAMPLE REPORT2: (LOCATION BASED MATCHUP)

1. START WITH THE BRAND PERFORMANCE SCORE
2. CATEGORIES INFLUENCING THE PERFORMANCE

(SERVICE, FOOD, VALUE FOR MONEY, AMBIENCE, DISCOUNTS)

TOP TRENDING FACTORS INFLUENCING THE EACH CATEGORY (1 +VE FEEDBACK FACTOR AND ONE -VE)

1. TINY FACTORS INFLUENCING THE PERFORMANCE

(ONLINE SHOPPING, DRIVE THRU, CUSTOMER SERVICE)

1. FACTORS SHOUBLE BE CONSIDERED TO IMPROVE PERFROMANCE.

THOUGHTS

1. Why not brand level comparison (considering 3 or more Mexican grills)