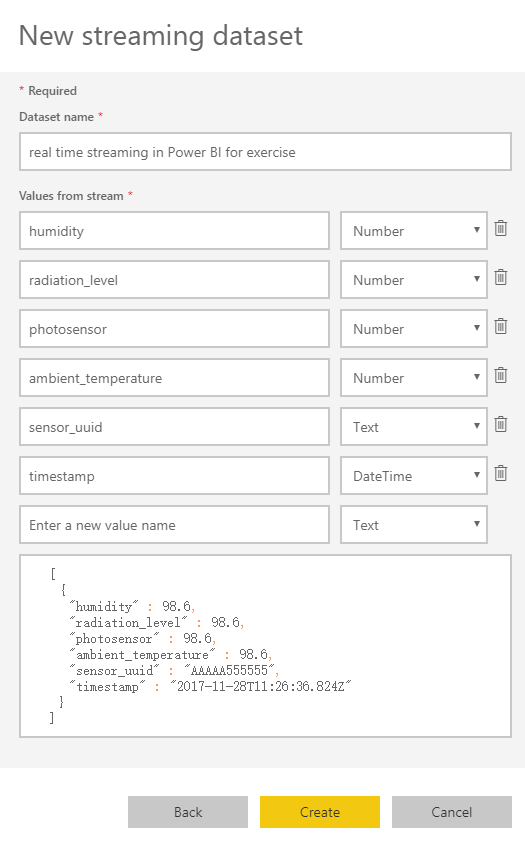
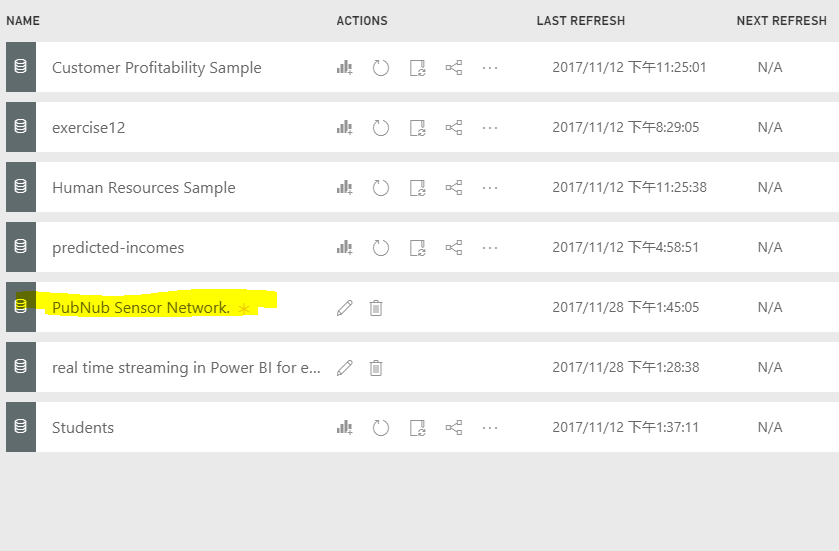
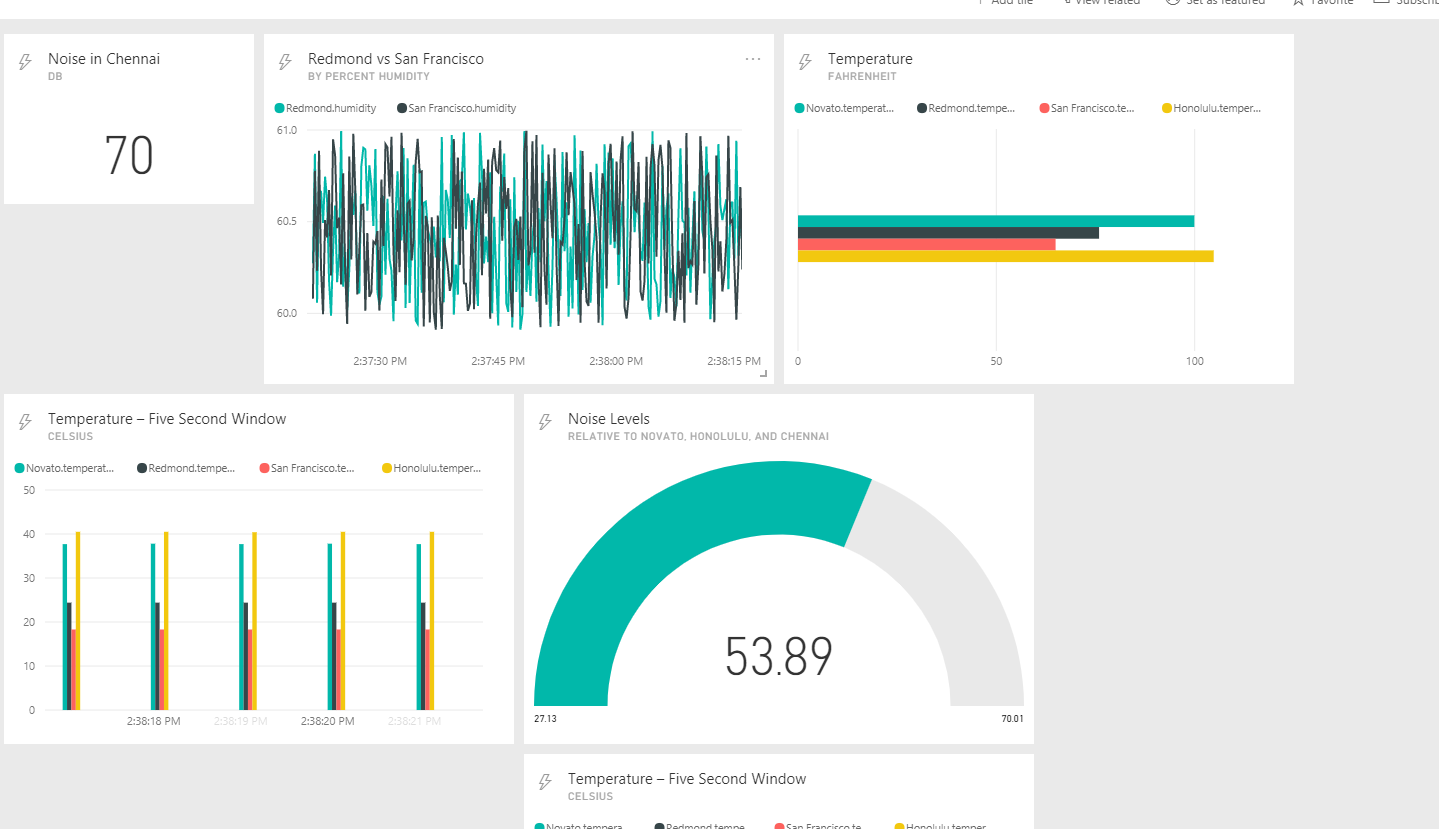
41. Follow the instructions and feed PubNub data streams into Power BI Service.



After the operations, we can get the stream data in datasets



It is able to add a line chart such as：



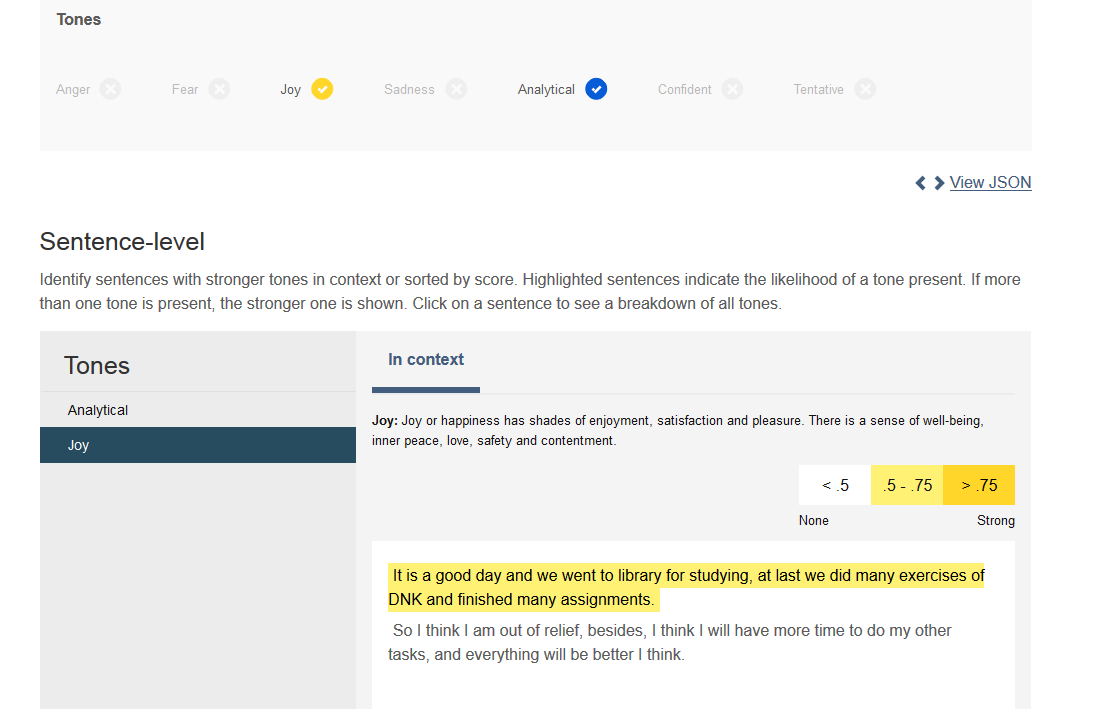
42. Write your own text and analyze it within IBM Tone Analyzer and IBM Personality Insights

The text that I input is:

‘It is a good day and we went to library for studying, at last we did many exercises of DNK and finished many assignments.

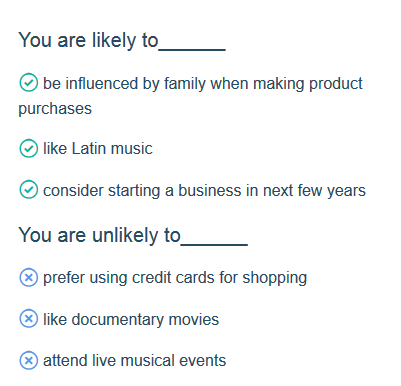
So I think I am out of relief, besides, I think I will have more time to do my other tasks, and everything will be better I think.’

* The IBM Tone Analyzer result is:

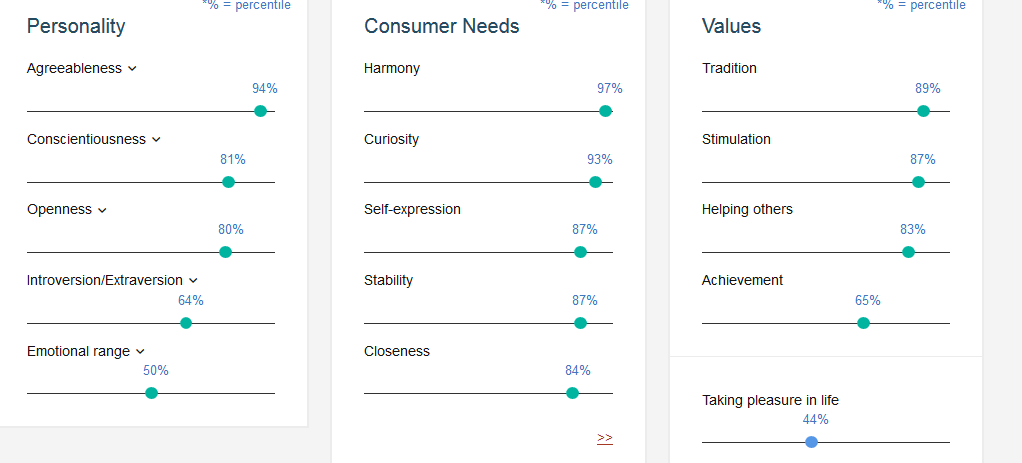


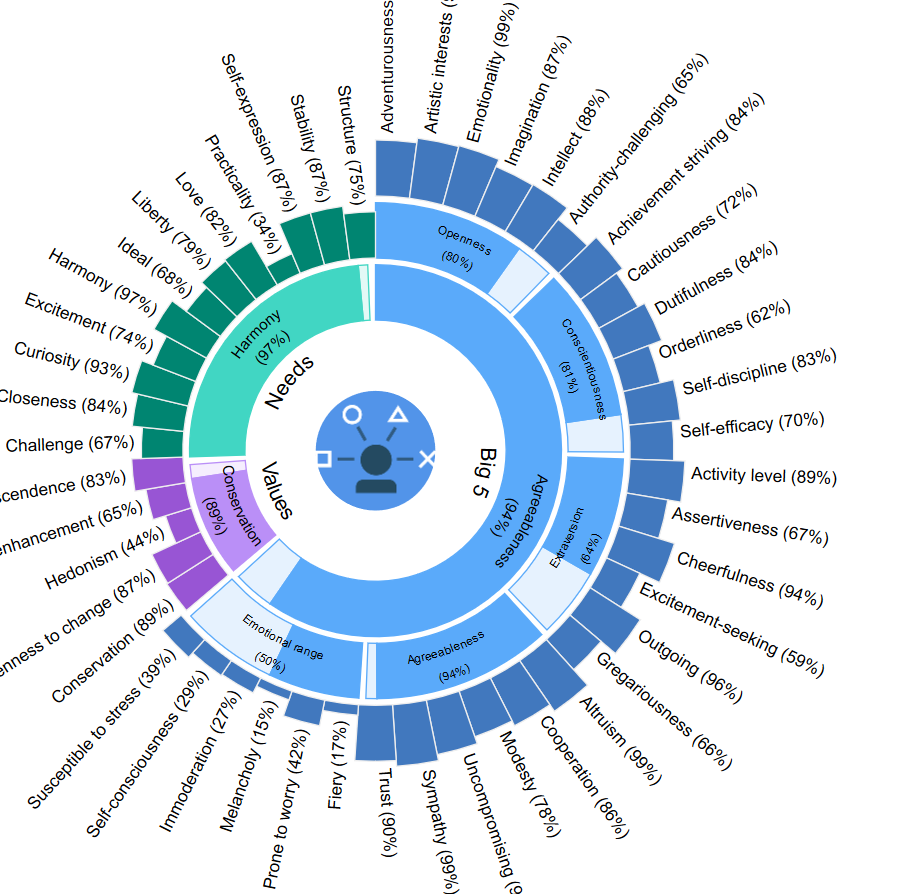
So it can clearly get the emotion expressed in the context and get the result for: Joy. It is able to understand emotions and communication style in text. The emotions are joy and analytical.

* IBM Personality Insights (I choosed demo as analysis text)

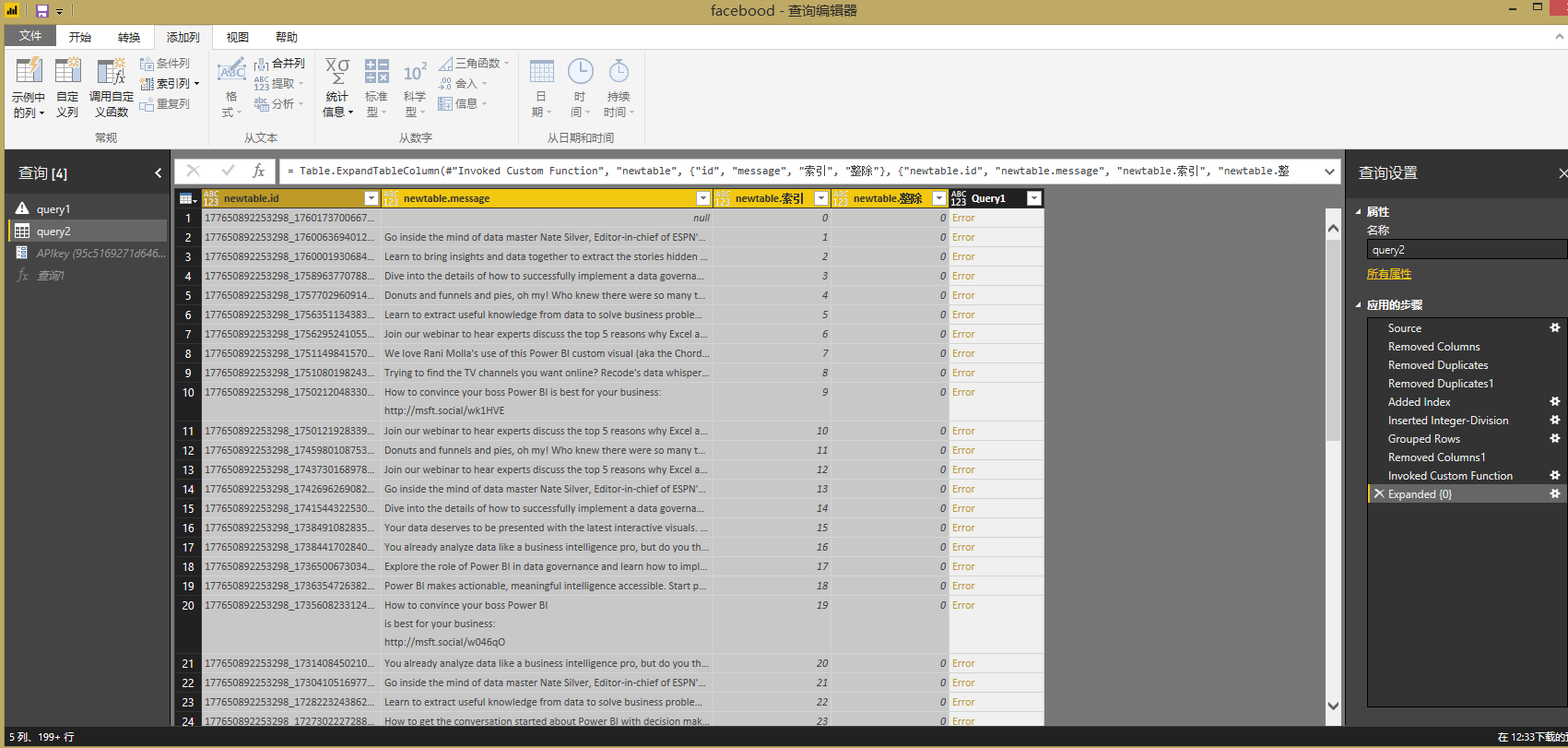


We can see that it can it can gain insight into how and why people think, act, and feel the way they do. This service applies linguistic analytics and personality theory to infer attributes from a person's unstructured text. So it helps user to see more clear about what attributes and characteristics they have. According to those analysis,



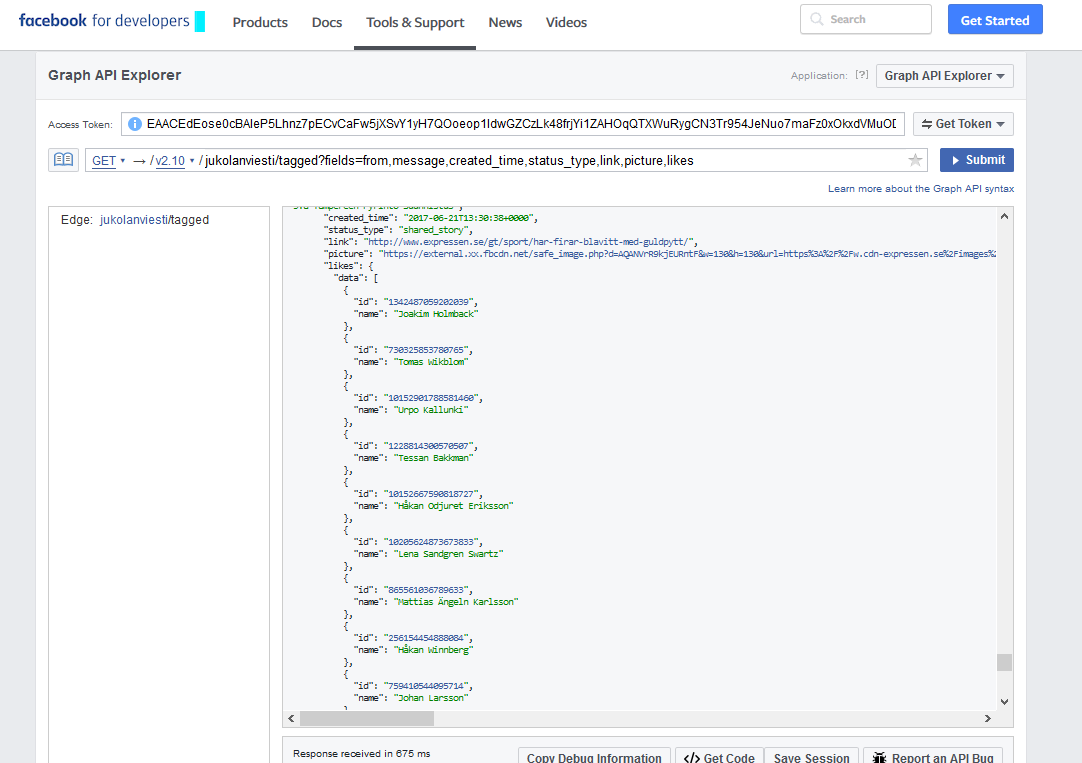


The graph give us a clear view that in this person’s personality, the Altruism, Emotionality take highest weight which are 99%. And this person is bad at emotional range, the weight of this person in emotional range is lowest among all of the personalities.

43. 

As I follow the instruction to get the analysis, I got error from cognitive service, but if it is not error, I can get the analysis from Text Analysis API, and get the information from it.

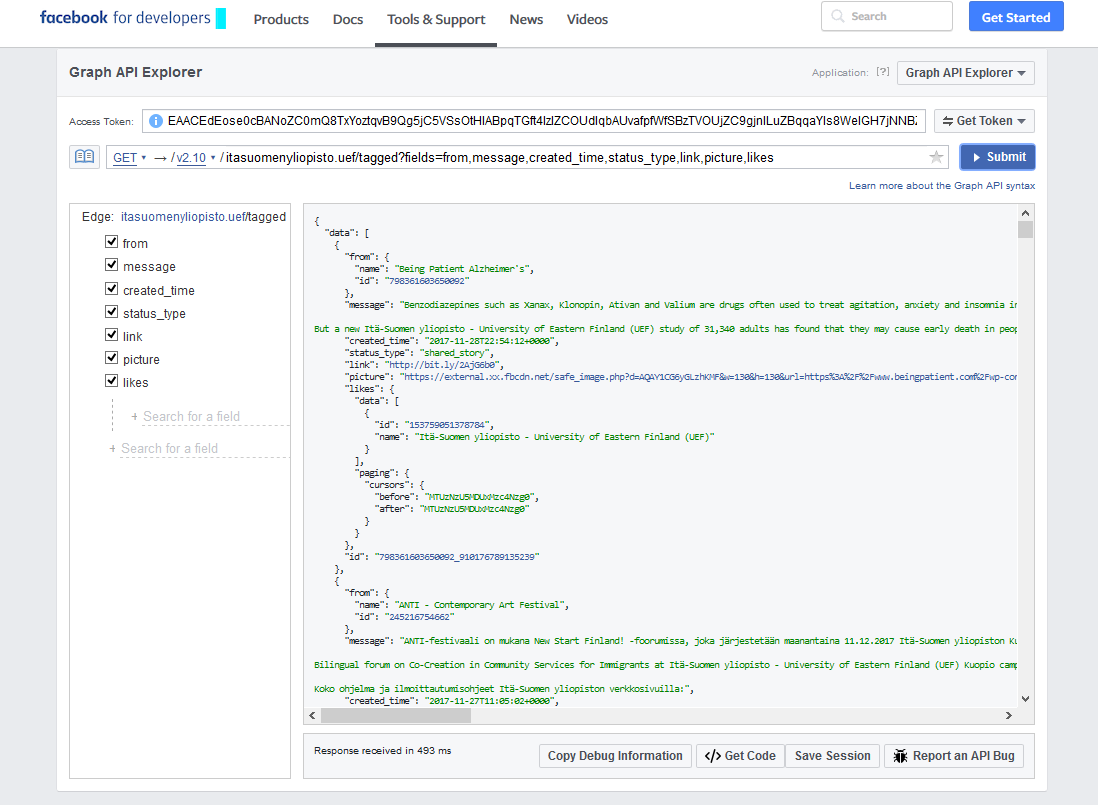
44. When I apply the method, in Graph API Explorer: we can see



When I apply the command:

‘itasuomenyliopisto.uef/tagged?fields=from,message,created\_time,status\_type,link,picture,likes’

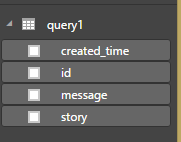
I got this result:



We can see that the data is analyzed with JSON format, and according to the message and created time , status\_type, and picture, The Graph API is the primary way to get data out of, and put data into, Facebook's platform. It's a low-level HTTP-based API that you can use to programmatically query data, post new stories, manage ads, upload photos, and perform a variety of other tasks that an app might implement.

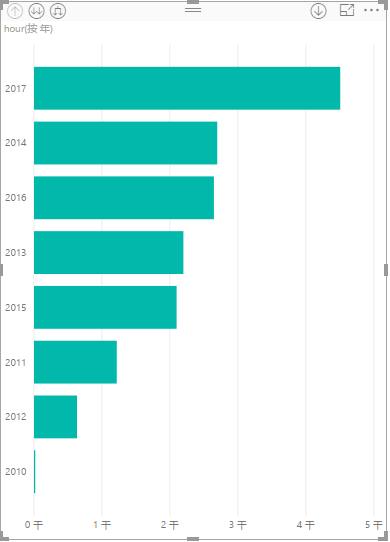
Facebook analysis:

In this part, when I get data, the version has changed from v2.2 to v2.8, so some columns has changed, different from the tutorials, there are totally 4 columns in the facebook posts data:



* Treemap visualization and bar chart





Cause the data has changed since version changed, we can’t get the share column, so I used the ‘hour’ column to show the analysis as a bar chart.

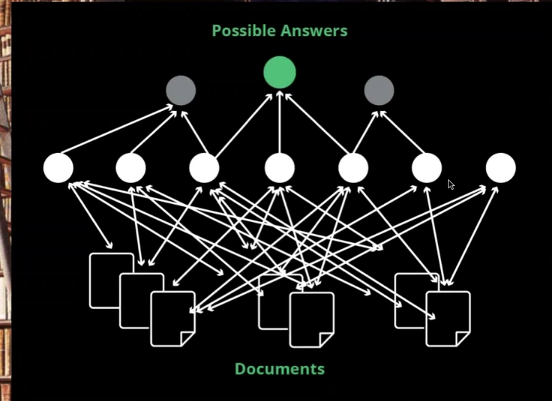
According to the visualization, we can see that sharing a link takes the most percentage, then is the updating of photos, next is Microsoft Ingite’s video… Also, with the increasing of year, more and more hours are spent on the using of facebook, especially 2017.

45. the core meaning of AI:

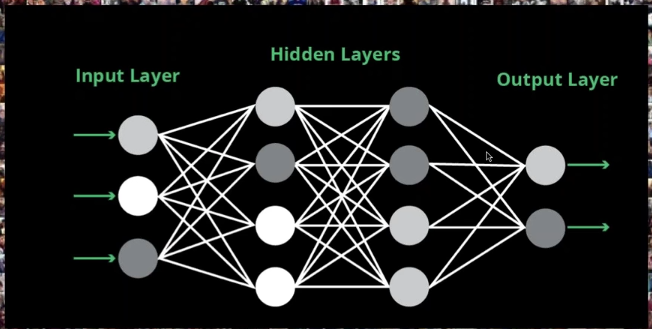
There are 3 core functions that consist of what AI can do today:

* Assess:

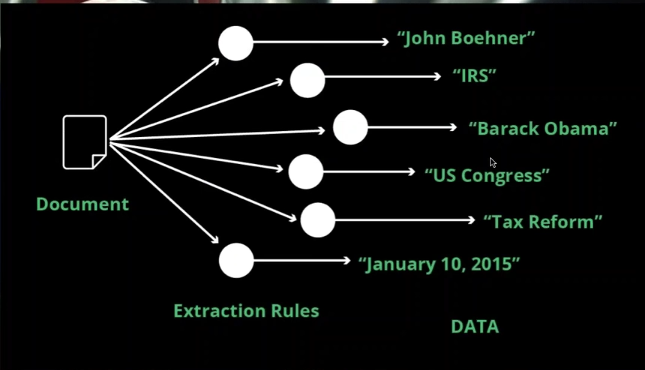
1. Machine learning: analysis to get some relation between data elements in AI, extend our knowledge by looking at these data and find the pattern. Take the unstructured data and form it into structured data and find the trend.
2. Reasoning from evidence: Given a question you try to answer, looking for what kinds of evidence can be put together and aggregate to get the answer,



1. Deep learning: give the capability to recognize what kind of data; based on a lot of features, take these features and reduces them down, move through these layers and build a network, until I got what I want.

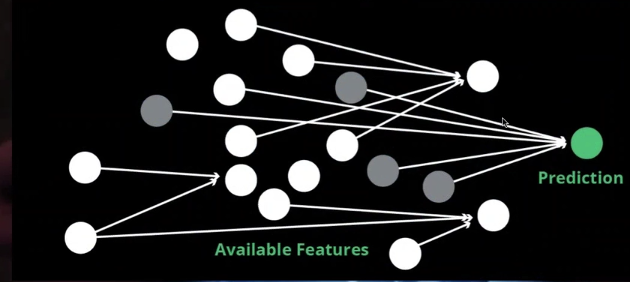


1. Language understanding: there are some clus, by extracting these clus from language, there are some keywords that can be used for analysis and understanding.
2. Natural language processing:

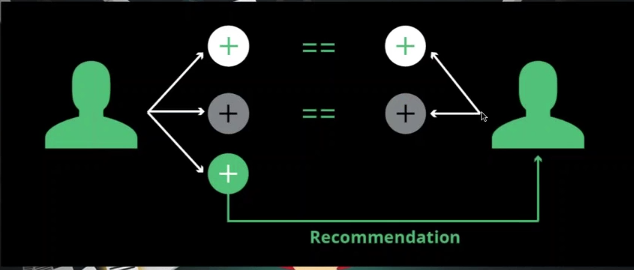


* Predict: extend view to future

1. predictive analysis: the features that related to what we know about tomorrow, can be link to the things that we want to know.

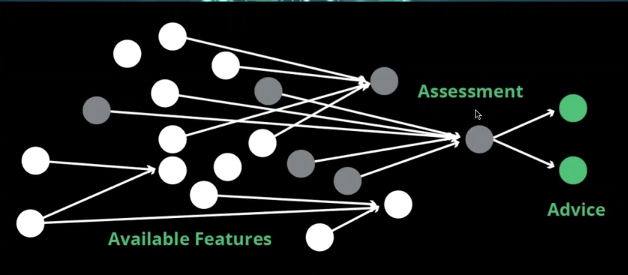


Recommendation: Having an idea what the item are, and who you are; the transactions that I have can be linked to some other people.

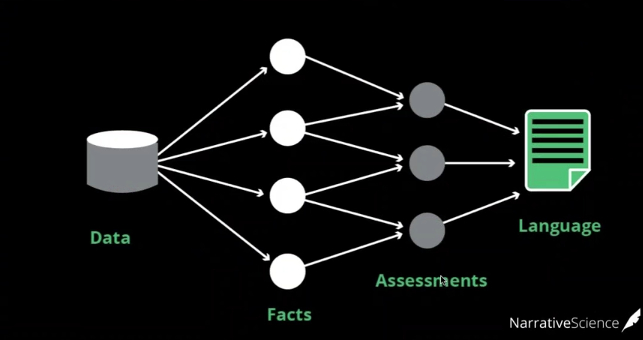


* Advise:

advisory system: different kind of data, use data about your goals, driven not only by data, but also such as business rules.



language generations: take the process of data and turn it into language.

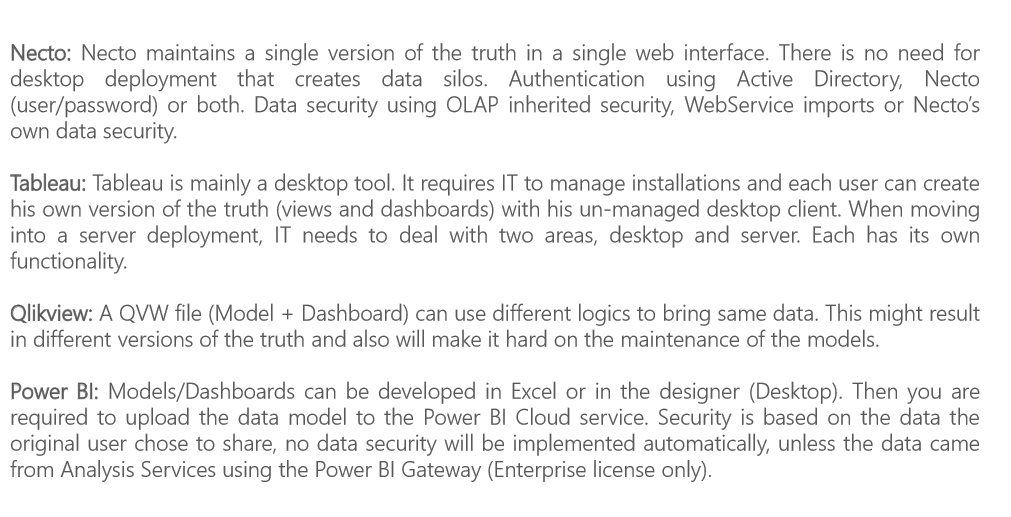


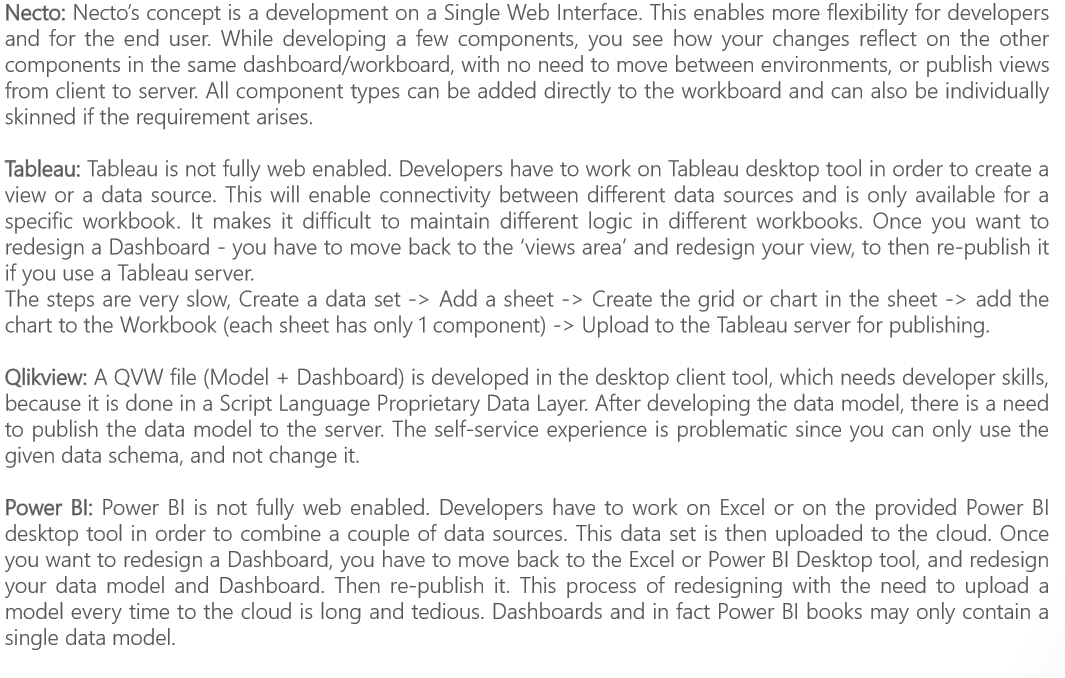
We need them to communication, co-worker, how they figure out what we know; since we never want data, we wanted insight & Action; we are about to enter the world that machine can better understand data than us, and the core meaning of AI is to let the machines be able to think like a human. Learning and improving like a human.

46..

* BI Platform Administration, Security and Architecture.
* Cloud BI
* Data Source Connectivity and Ingestion
* Metadata Management
* Self-Contained Extraction, Transformation and Loading (ETL) and Data Storage.
* Self-Service Data Preparation
* Embedded Advanced Analytics
* Analytic Dashboards
* Interactive Visual Exploration
* Smart Data Discovery
* Mobile Exploration and Authoring
* Embedding Analytic Content
* Publish, Share and Collaborate on Analytic Content
* Platform Capabilities and Workflow
* Ease of Use and Visual Appeal

Panorama Necto is a secure, centralized and state of the art BI tool.





As can be seen,

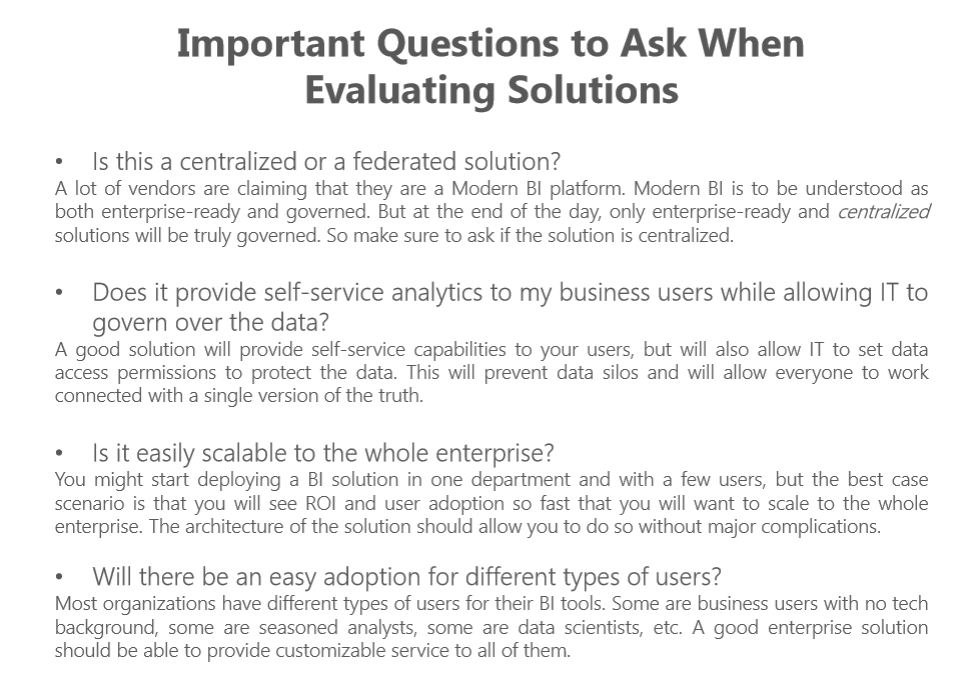
Necto is a single web interface, its concept is a development on a single web interface. Business users can visualize their insights and process with a full infographic suite. It supports a fully collaborative decision making capability in context. It provides business users with the ability to define alerts and notifications within the system based on time, thresholds, exceptions, etc.

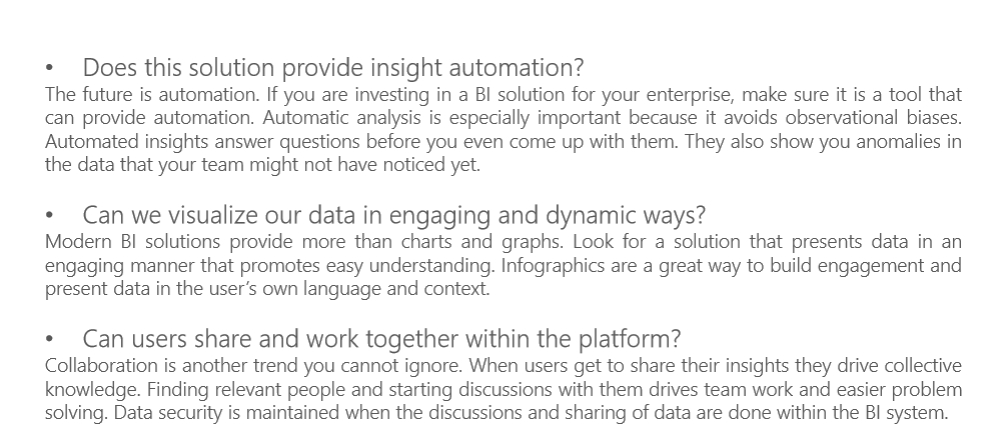
Tableau is a desktop tool, it is not fully web enabled. It supports infographics using some basic shapes. It has no collaboration functionality aside from the share and annotations that Necto supports as well. Only systems managers can define exceptions and emails, hindering the ability for the end users to work and report on their needs.

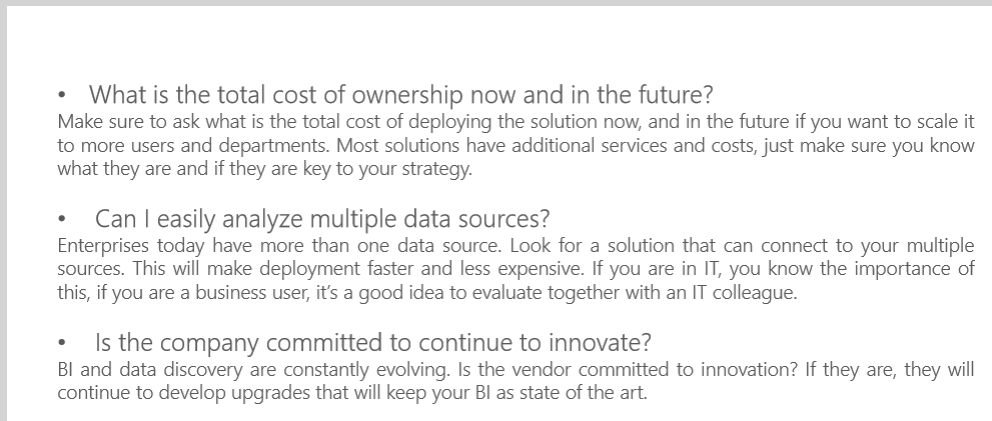
Qlikview is a file, it is developed in a desktop client tool, the inner infographics suite is limited. It supports: build and share analysis objects; share analysis with bookmarks; notes and comments; share sessions, those methods of collaborations. Limited in it, available in full mode by using printing tool as an additional plug-in

PowerBI includes models and dashboard, it is not fully web enabled, developers have to work on Excel or on the provided PowerBI desktop tool in order to combine a couple of data sources. It supports some basic and minimal infographics. Users can create new comments on a workbook or insert text boxes near elements as annotations on a component level only. No ability to get reports by mail from the web.

Take into consideration when evaluating solutions:







47. If you will evaluate data science platforms what you should take into consideration?

* Does it aggregate the steps and big data tools necessary to get from raw data to production ready application?
* Is it able to make the models and predicted values accessible to other business applications.
* Is it able to publish the models and predicted values to a variety of other destinations.
* Does it provides an interactive visual interface where they can point, click , and build or use languages.
* Does it provides tools to draft data preparation and modulization in seconds.
* Does it include its vast customer base and continued innovation of its data.

The firms covered:

Leaders

Challengers

Visionaries

Niche players



It is able for us to evaluate the platform with the figure above on commercial vendors.

Also it is able for us to compare by time such as 2017 and 2016, we can compare with the completeness of vision, how is the ability to execute.

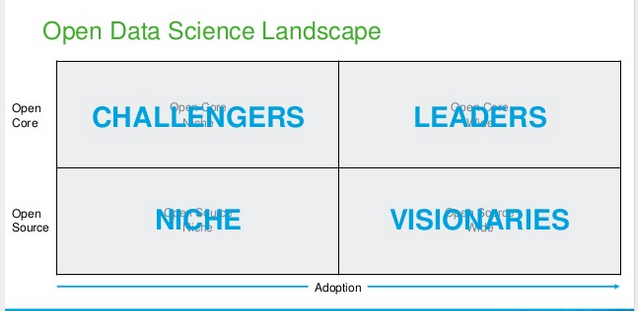


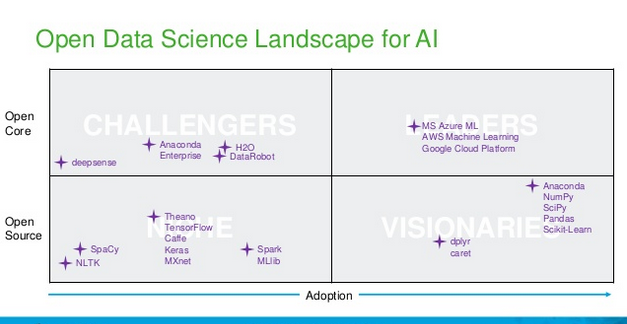
48.

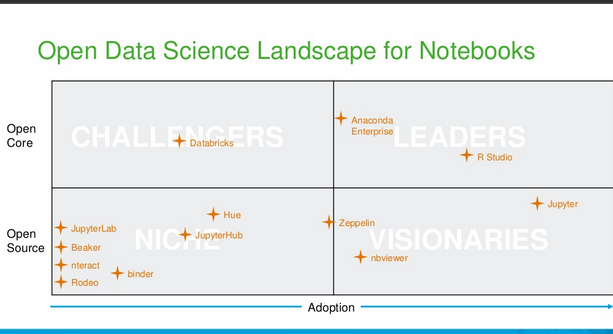
* + What is a compendium?

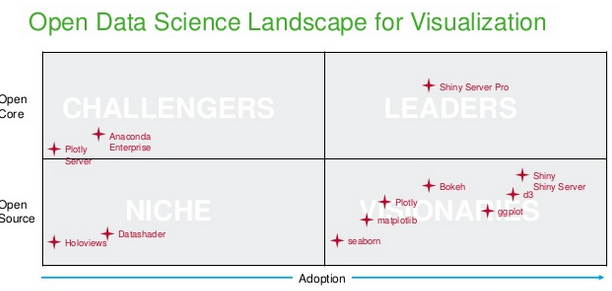
A solution for providing a standard and easily recognisable way for organising the digital materials of a research project to enable other researchers to inspect , reproduce, and extend the research.

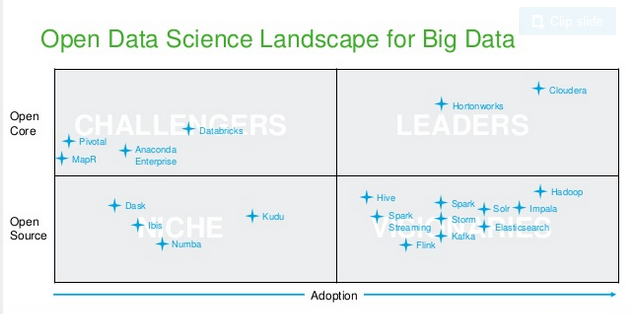
* + What is open core and open source?

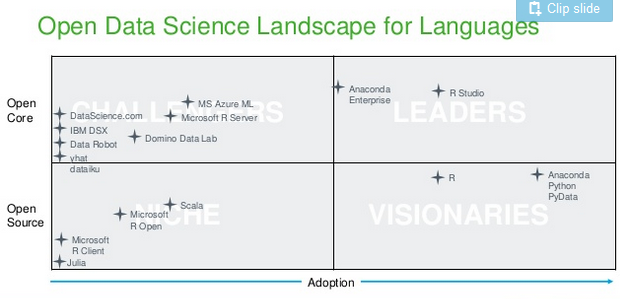


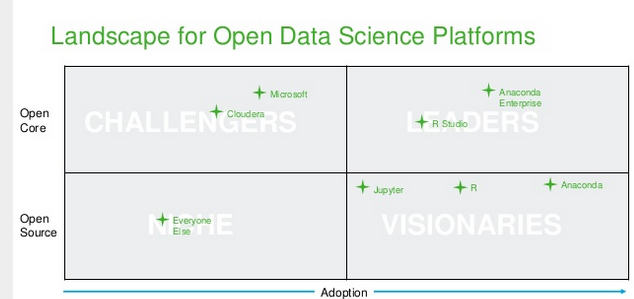












* + Open core[1] is a business model for the monetization of commercially produced open-source software. Coined by Andrew Lampitt in 2008, the open-core model primarily involves offering a "core" or feature-limited version of a software product as free and open-source software, while offering "commercial" versions or add-ons as proprietary software. open core providers realize that enterprises want to move toward a true open source model, they face the reality of their new business model - offering a renewable subscription which relies on superior support, ongoing certifications and constant innovation.
  + The open-source[2] model is a decentralized software development model that encourages open collaboration. A main principle of open-source software development is peer production, with products such as source code, blueprints, and documentation freely available to the public. The open-source movement in software began as a response to the limitations of proprietary code. The model is used for projects such as in open source appropriate technology, and open source drug discovery.

As can be seen from the figures above,

Open core focus on challengers and leaders, in AI, there are deep sense, anaconda etc. Open Source focus on NICHE and VISIONARIES; in AI, there are spaCy, NLTK, etc.

* + Why containers are leveraged in data science?

Containerization technology has revolutionized how developers build, ship, and deploy applications such as Doker. Containerization is one solution to solve the challenge that many open data and their dependencies can make it hard to control the analytics and development environment. It is able to be one way to support multiple projects that require different versions of libraries. It provedes the strategy to easily reproduce their entire environment.

49. Try to find out do they leverage generative models and discriminative models or something else.

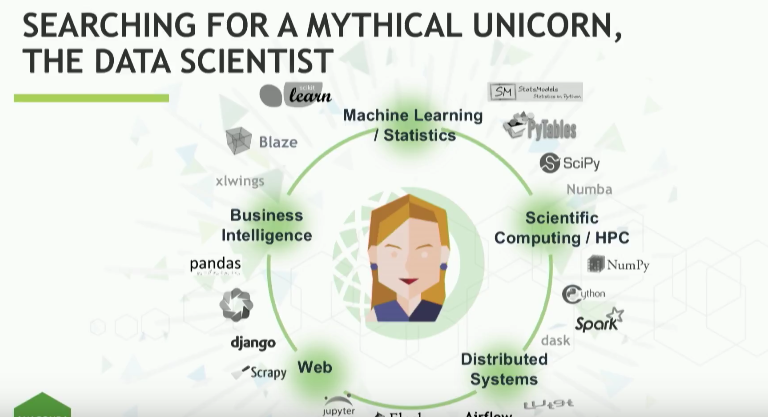
Yes, Deep Convolutional Generative Adversarial Networks (DCGAN), Generative Adversarial Networks (GAN) have used generative models

According to the link , I have found different models such as:

* Pre-trained mode (this model differs slightly in that they no longer split the convolution layers into two groups (model parallelism), as a result this model has very slightly more paramaters but achieves better accuracy.)
* Acoustic model;
* Deep Structured Semantic Model
* ResNet\_18 model
* BN-Inception model
* Inception V3 model

...

Review Supercharging Excel with Anaconda Fusion | Christine Doig & Fabio Pliger |AnacondaCON 2017:



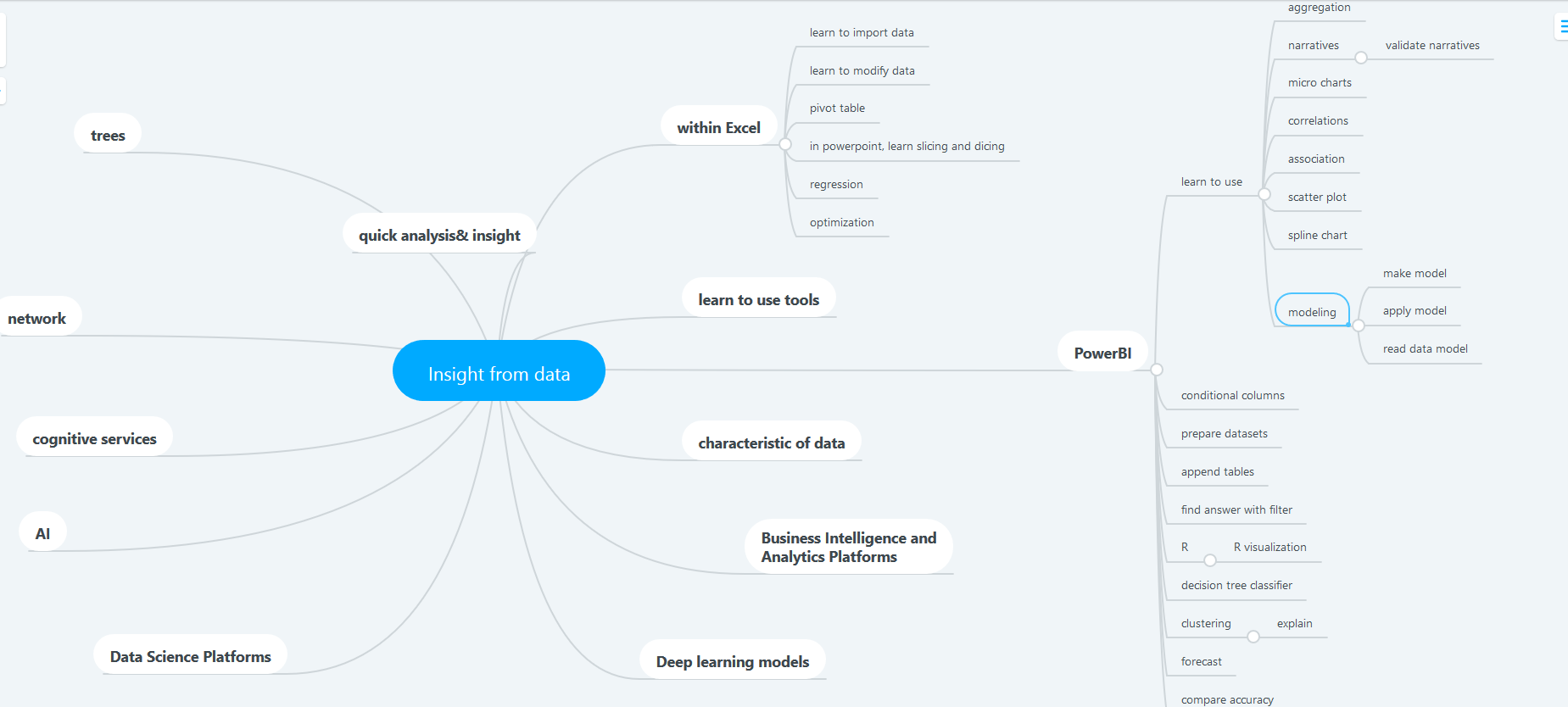
Open data science in organizations, it is a team work, also collaboration is needed, business analysits and data scientists; business analysts are being left out of the open data science revolution, anaconda fusion brings analysts to the open data science table, it is a revolutionizing how business analysts do data science. In Anaconda fusion, analysts and data scientists can keep using their preferred tools.

Review of the demos:

50.

Make your own construction that help you to get insights from data

MindMap:



Explaination:

Summary: This mind map demonstrates in my understanding, the consists of insight of data.

Description: This mind map includes the network, cognitive services, AI, use of powerBI and Excel, in order to get the insight of data, we are able to use these theories and tools above to achieve the goal.

Data: In the lecture, there are many datasets can be choosed, such as UCI data sets

Data visualization: Not only powerBI, but also excel provides tools to make the data visualized, in order to help us get clearer understanding of the data we are operating, and help us to find the answer we want such as forecasting.

Finally, when we are able to handle these tools and understand to meaning of AI, deep learning, machine learning, etc. We can get clearer insight of data.

Reference:

[1] <https://en.wikipedia.org/wiki/Open_core>

[2] <https://en.wikipedia.org/wiki/Open-source_model>