Using Python/R to build your first Machine Learning Model in Power BI

Ashraf Ghonaim

Ashraf Ghonaim

User Group Leader

LinkedIn /in/aghonaim

Strategic Management Consultant – City of Toronto

- Leader of Toronto Power BI, PowerApps and Flow User Group
- Organizer – Toronto Power BI World Tour Toronto and Montreal
- Data and Analytics Enthusiast
- Area of expertise Power BI, PowerApps and Microsoft Flow
- Speaker - Business Intelligence events and user groups
- Volunteer in several Open Data and Data for Public Good initiatives.
- Frequently provides mentoring/training to Analytics professionals.
- Microsoft MVP - Data Platform (Power BI)

Objective

This session is a gentle introduction about how to setup Python/R and how to use them in Power BI to build your first Machine Learning Model

Session Agenda

- Introduction to Power BI
- How to install Python/R
- How to use Python/R in Power BI
  - Get Data
  - Data Manipulation
  - Visualization
- Limitations

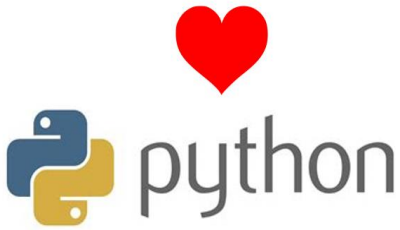
What is it ?

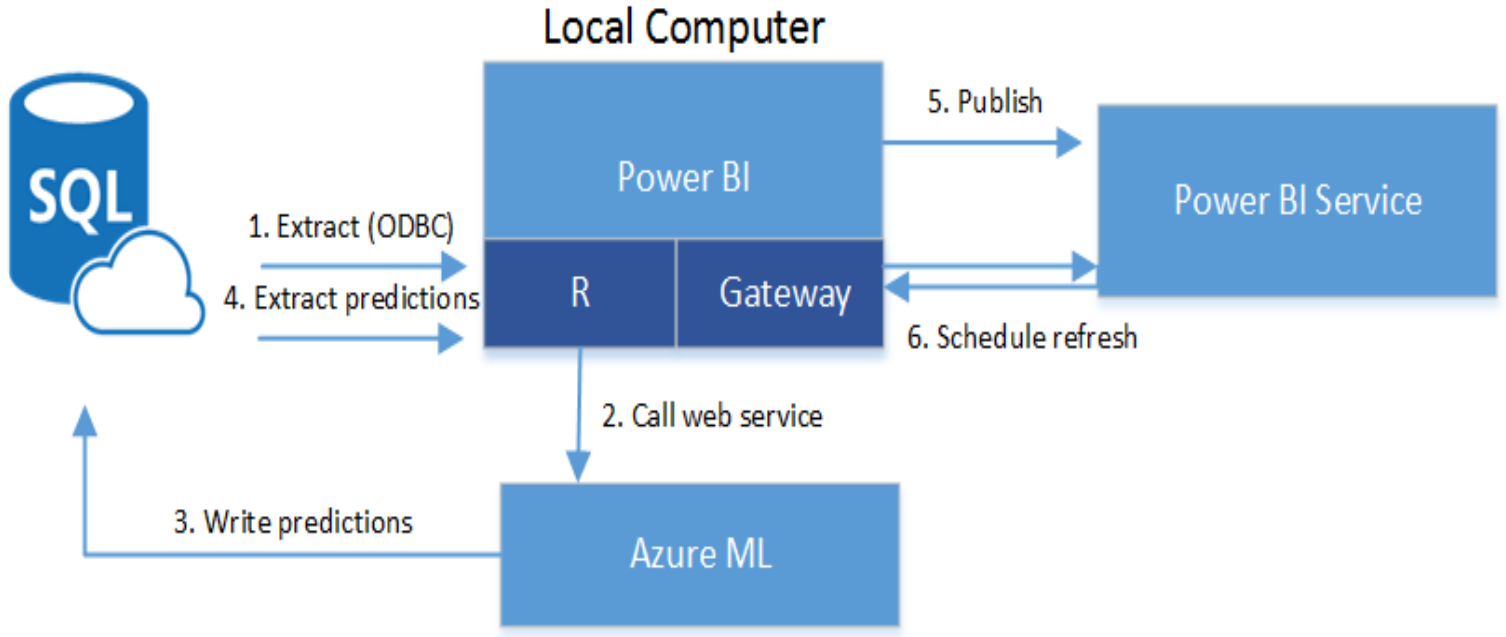
What is Power BI ?

- Power BI is Microsoft self service visualization tool
- Started as Microsoft Excel–based add-ins: Power Query, Power Pivot and Power View
- Later renamed to Power BI it was then unveiled by Microsoft in Sep 2013 as Power BI for Office 365.
- It was released as standalone visualization tool as a service (SaaS) to the general public on July 24, 2015

Machine Learning

Power BI





What is Next?

Power BI
PAGE TITLE
Label
Total Invoice
THIS MONTH
$2.22M
Discount Savings
THIS MONTH
$4.07M
Total Invoice
BY MONTH, CATEGORY
Total Invoice
BY SUBCATEGORY
Hardware
Indirect services
Other
Raw materials
Outsourced
Feedback Rating
BY SCORE
Account Count
BY INDUSTRY
Varience to Budget
BY MONTH
9:41 AM
100%
Executive Metrics
Expected Revenue
THIS YEAR
$85.22M
Total Invoice
BY MONTH, CATEGORY
Category
Direct
Indirect
Logistics
Other
Customer Feedback Trend
Sales Pipeline
Customer Size vs. Revenue

BLOG › ANNOUNCEMENTS

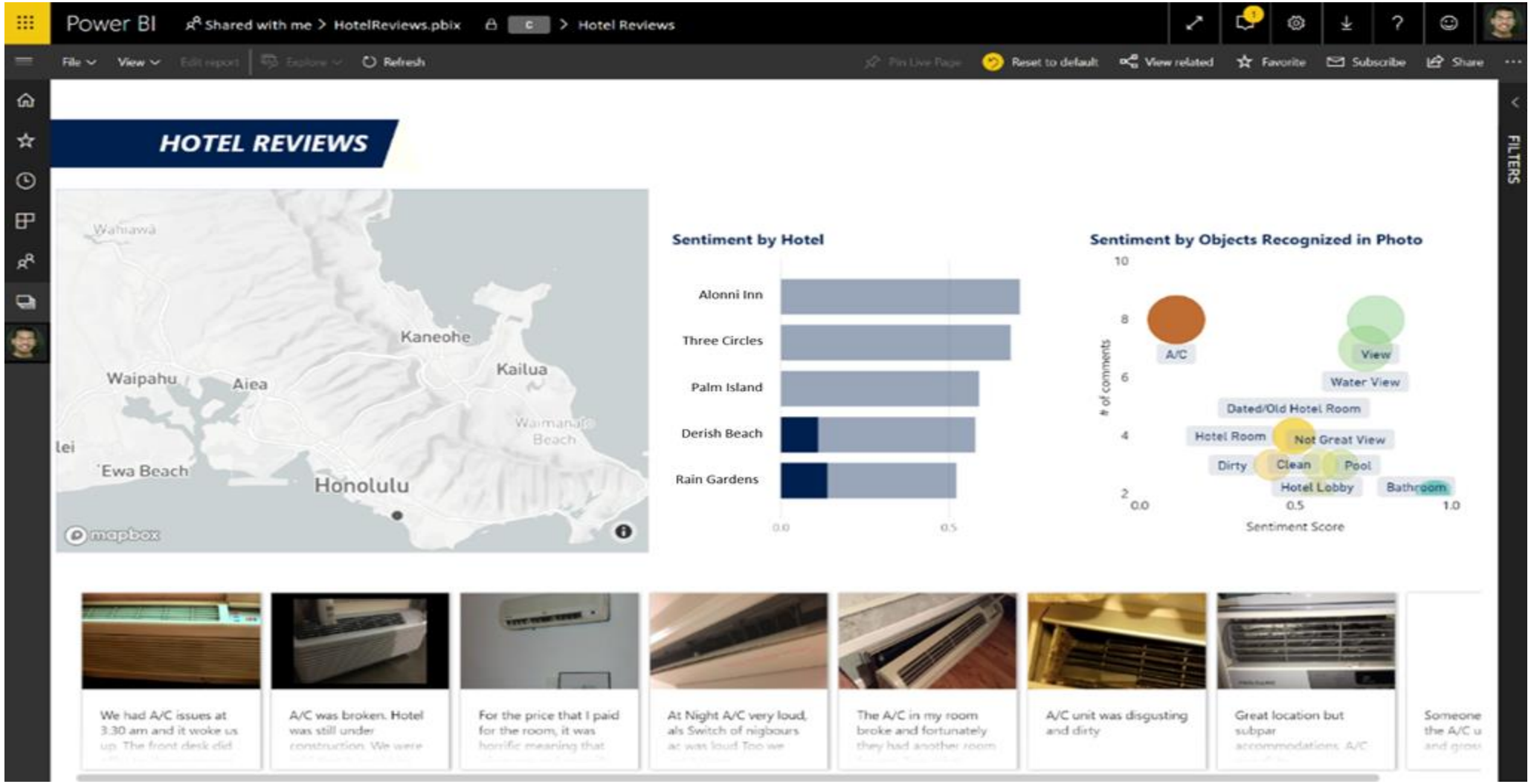
Announcing new AI Capabilities for Power BI to make AI Accessible for Everyone

Arun Ulag
General Manager, Engineering

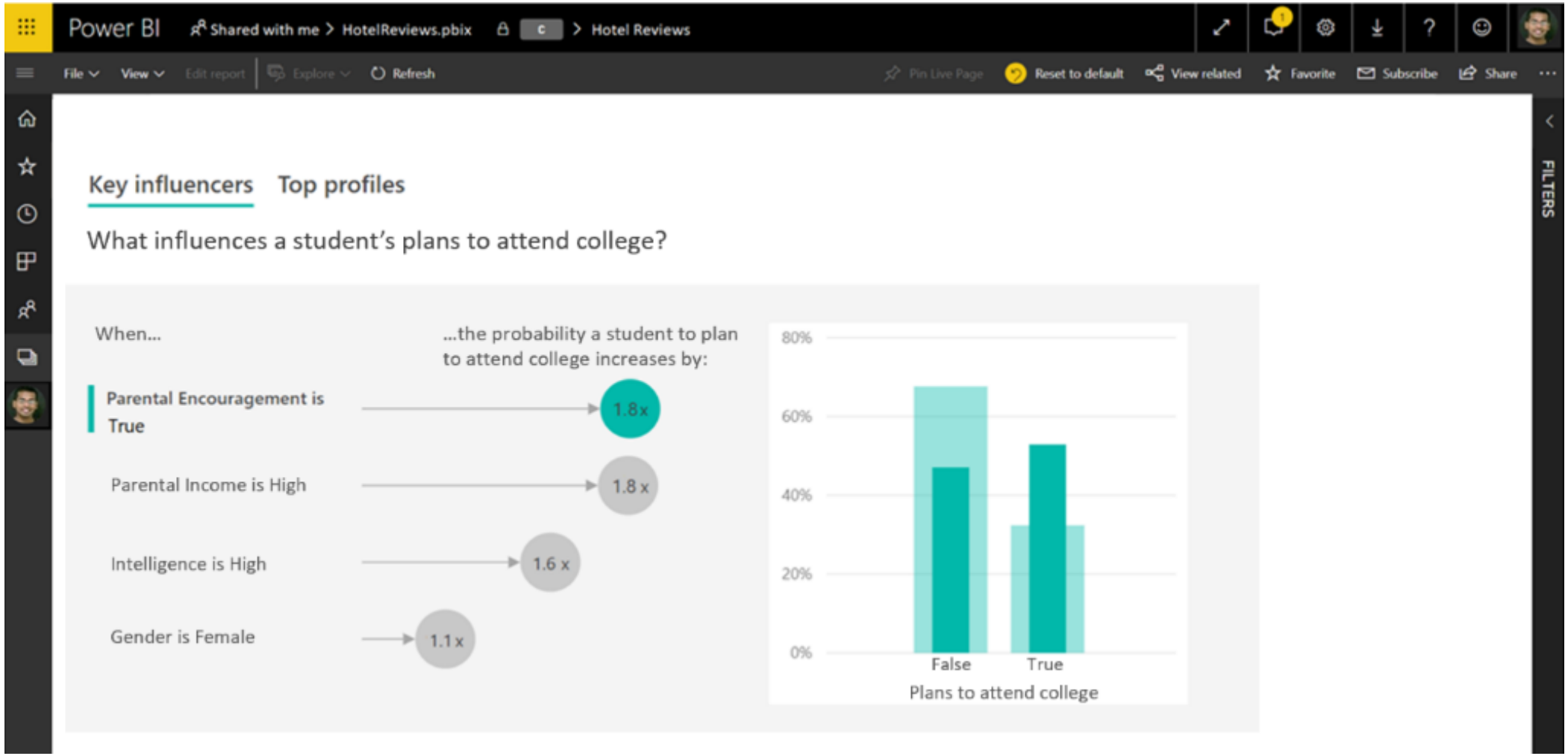
November 13, 2018

Share Tweet Like

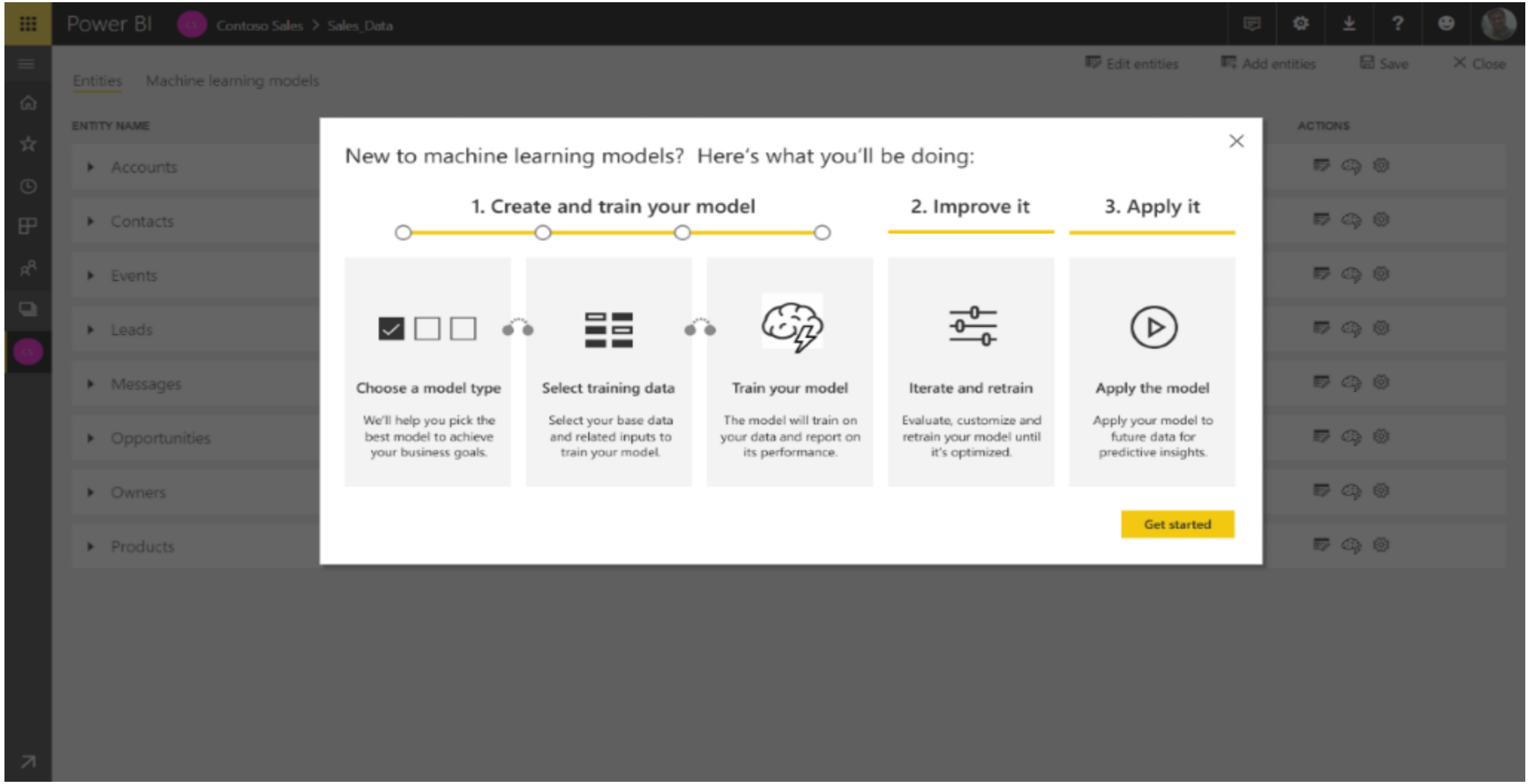
Azure Cognitive Services capabilities in Power BI



New Key Driver Analysis feature in Power BI



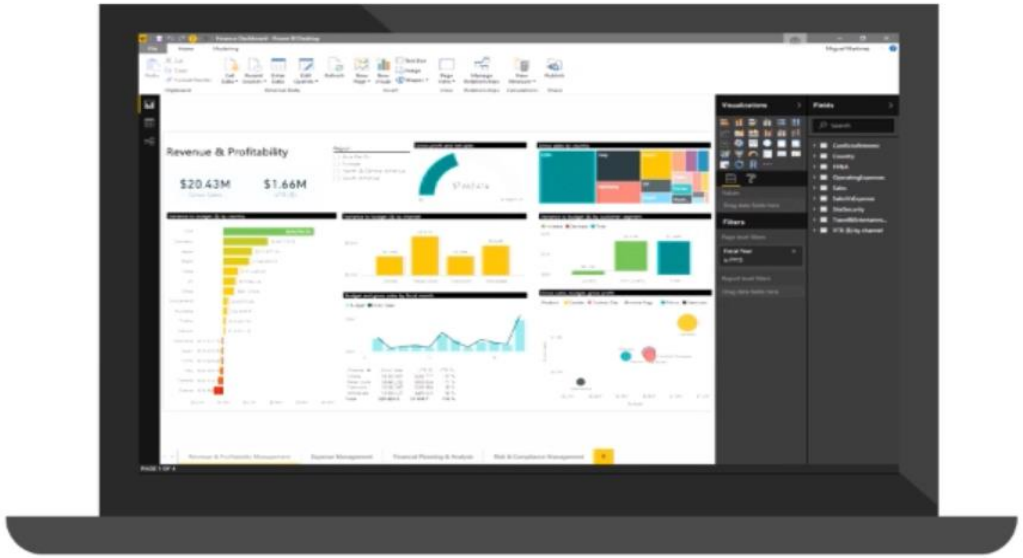
Build Your Own Machine Learning Models in Power BI



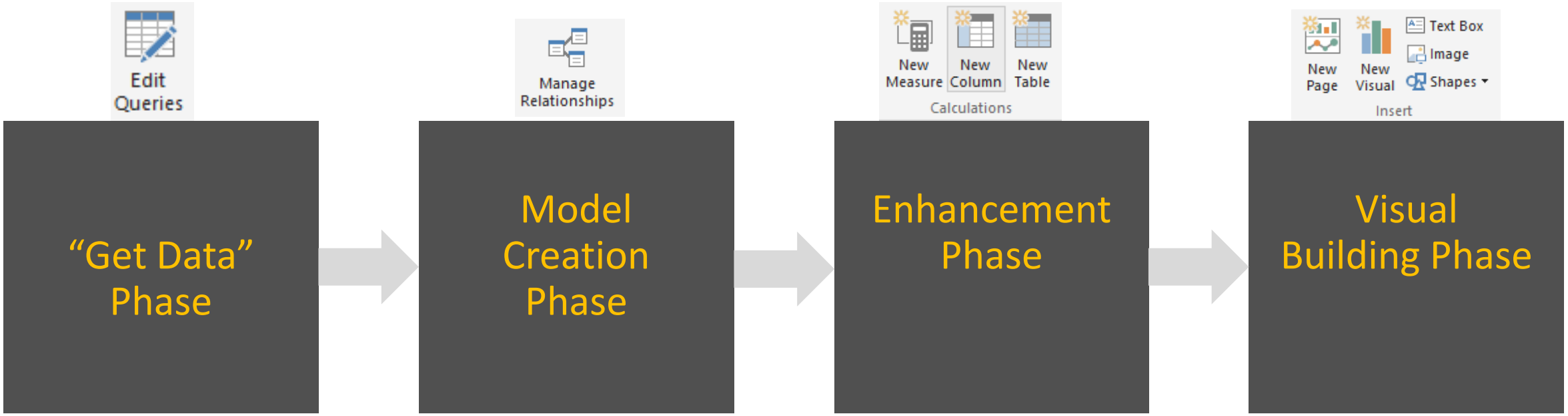
PBI Development Steps

What is Power BI Desktop?

- A free downloadable 32/64 bit desktop application optimized for use with the Power BI service & Power BI Report Server.
- Dedicated BI workflow – data connectivity, preparation, modeling and report authoring.
- Updated every month with new features (lots of them ☺)



Phases of Building a Power BI Report



Create Query in "M" and Data Transformation

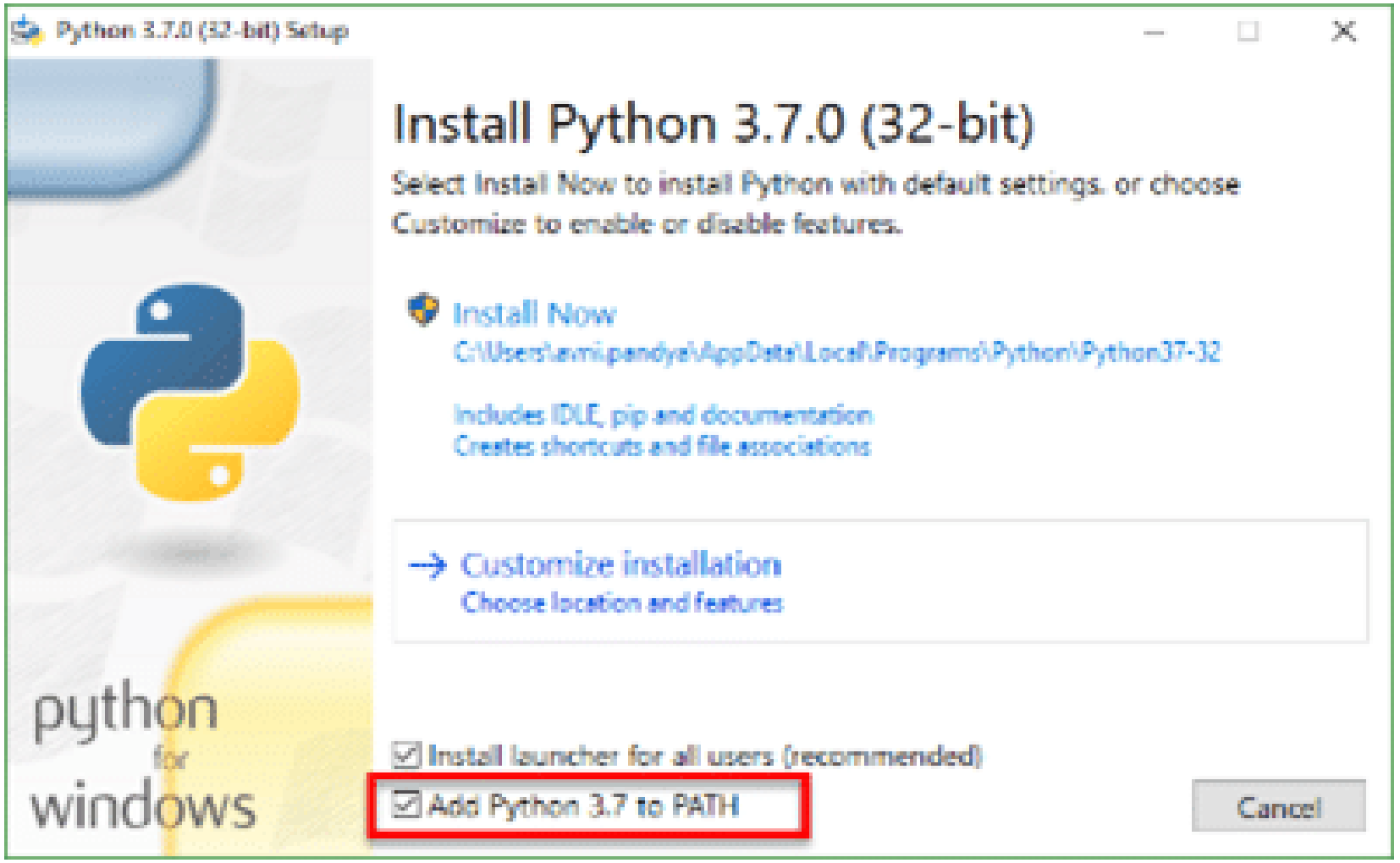
Build Fact and Dimensions and manage relationships

Add calculated columns, measures using DAX

Add visuals, Manage interactions and build Dashboards

Python & R Setup

Step 1: Install Python


Python 3.7.0 (32-bit) Setup
Install Python 3.7.0 (32-bit)
Select Install Now to install Python with default settings, or choose
Customize to enable or disable features.
Install Now
C:\Users\avnipandye\AppData\Local\Programs\Python\Python37-32
Includes IDLE, pip and documentation
Creates shortcuts and file associations
Customize installation
Choose location and features
python
for
windows
Install launcher for all users (recommended)
Add Python 3.7 to PATH
Cancel

Step 2: Install Libraries

- Data Prep/Analysis: NumPy, Pandas
- Visualization: matplotlib, Seaborn
- Machine Learning: scikit-learn
- Text Analytics: TextBlob, NLTK

- Run cmd as Administrator

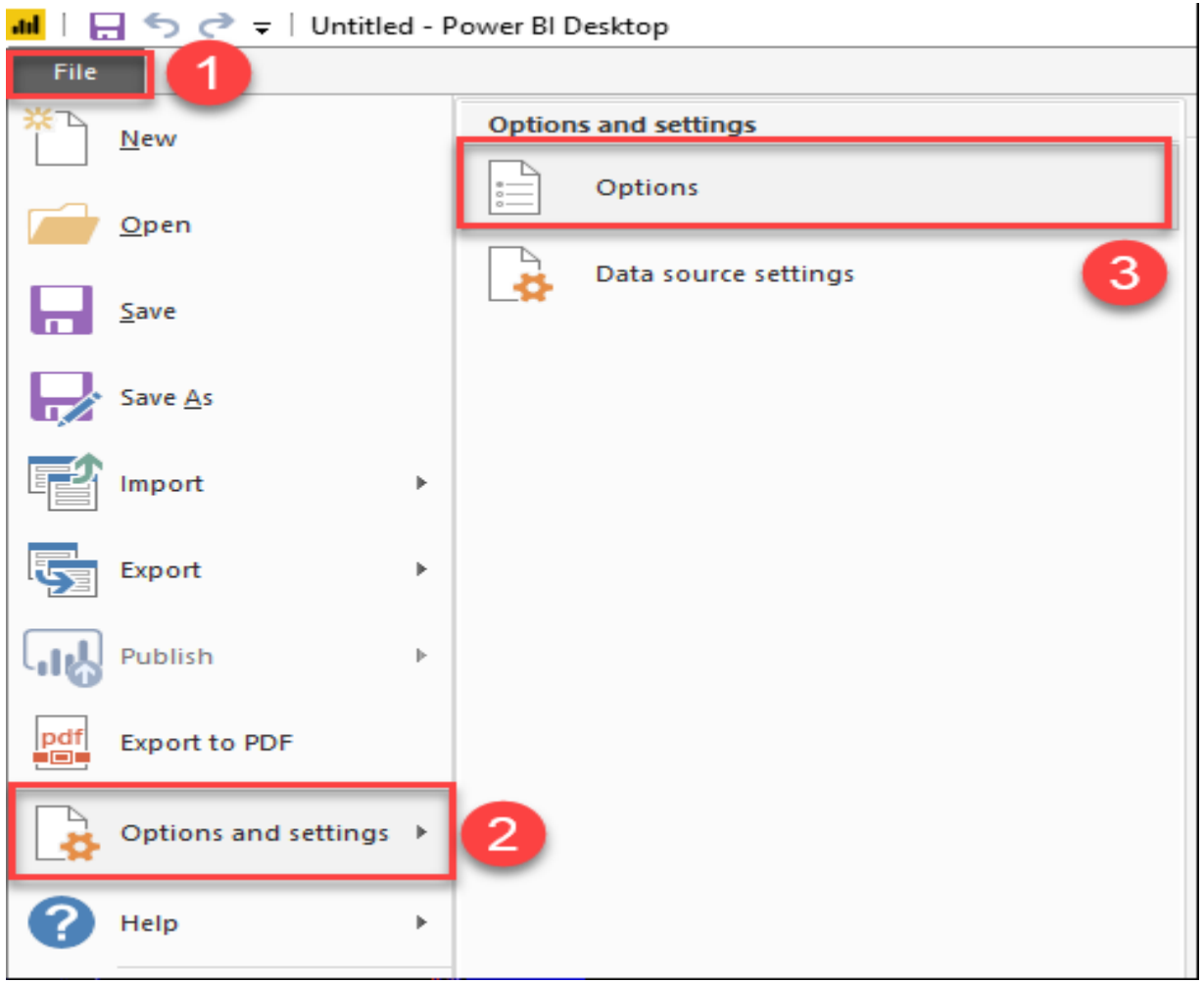
- Upgrade pip installation package:

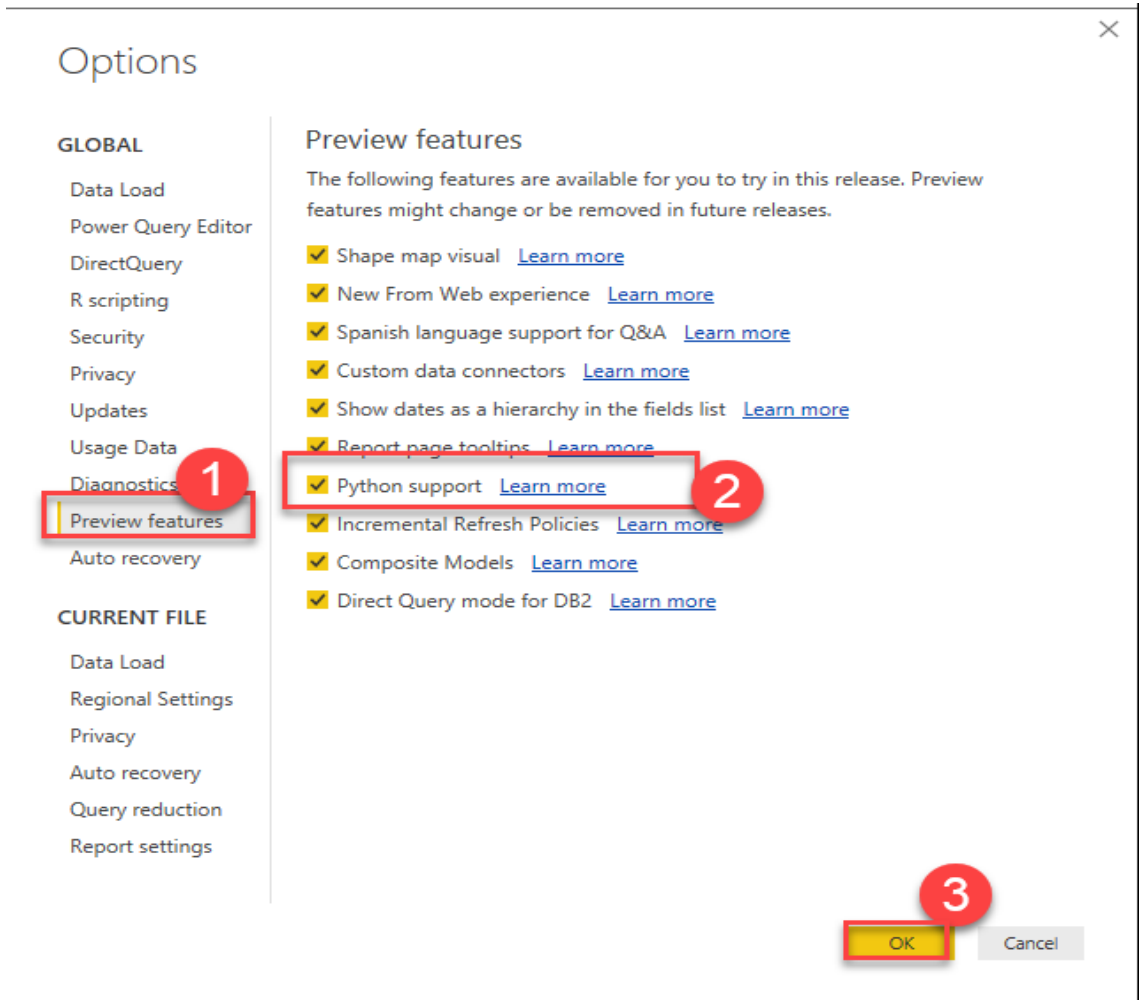
  > pip install --upgrade pip --proxy org\username:password@proxy.company.com:8080

- Install libraries:

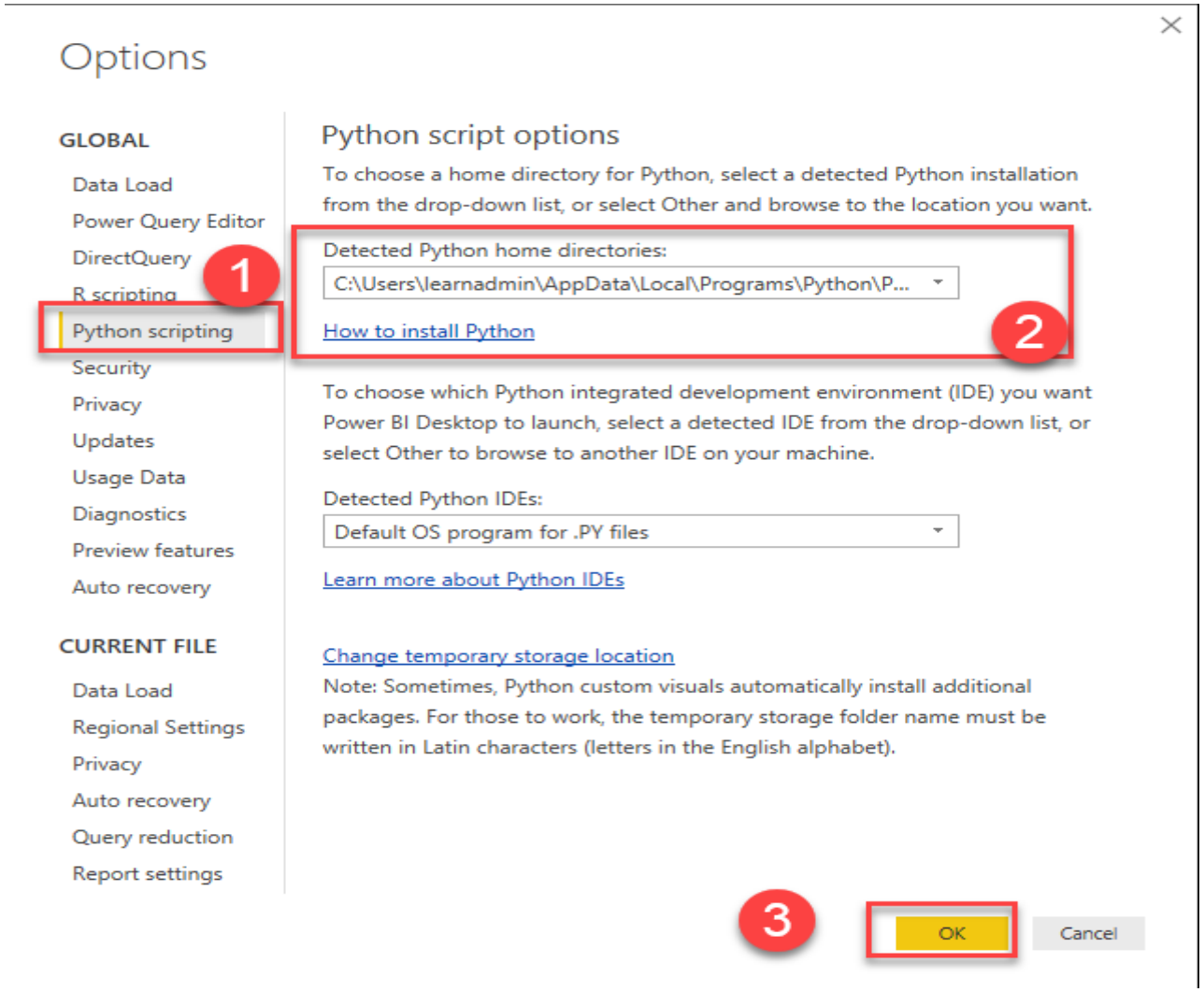
  > pip install pandas --proxy org\username:password@proxy.company.com:8080

Step 3: Enable Python in PBI



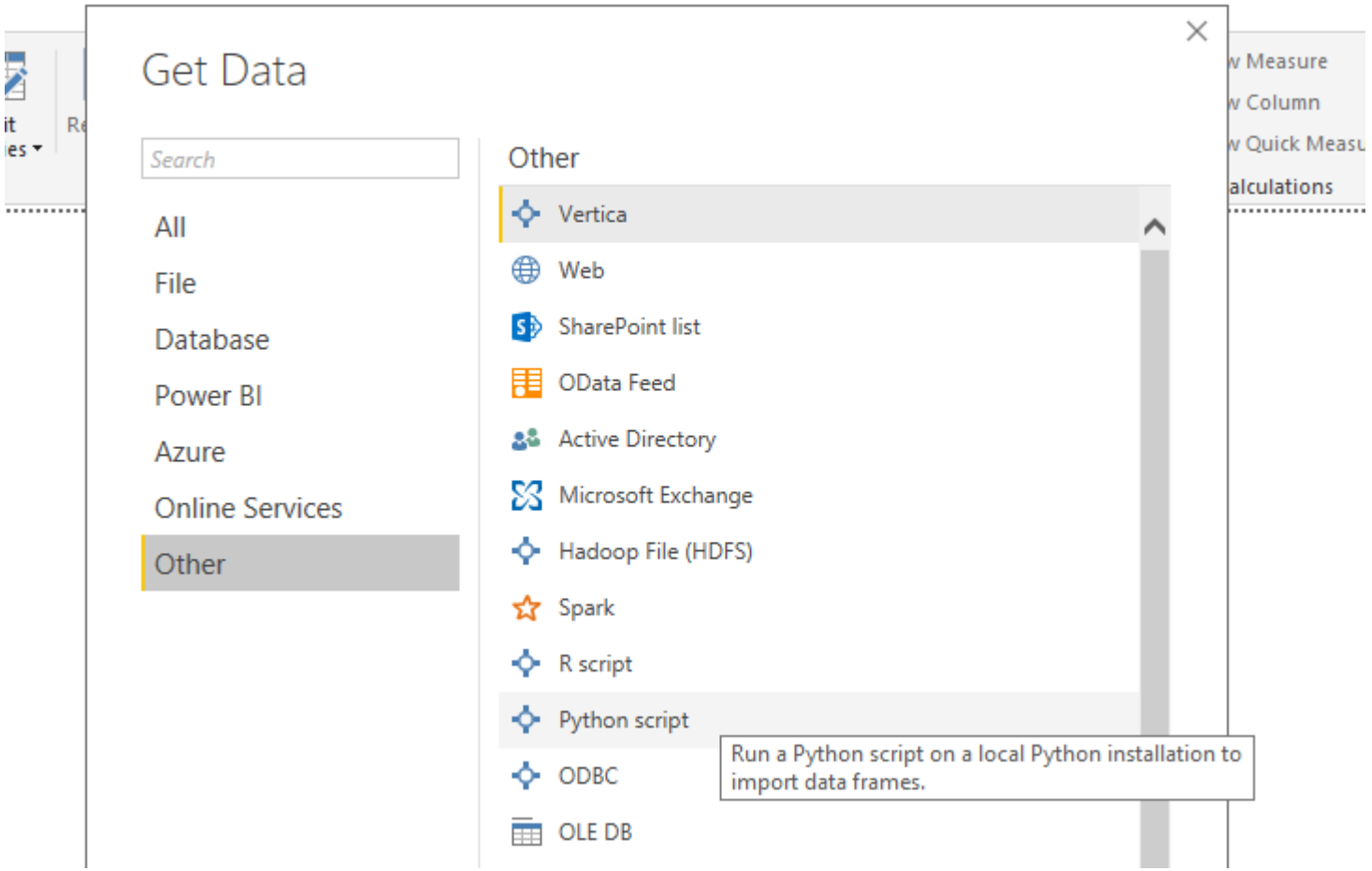


Step 3: Restart and check directory link



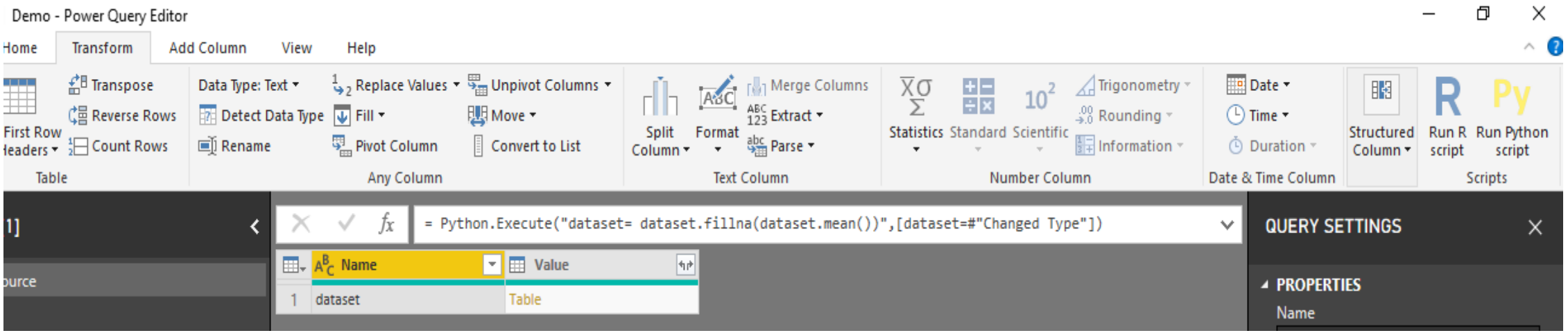
Where to use Python/R

1- Get Data



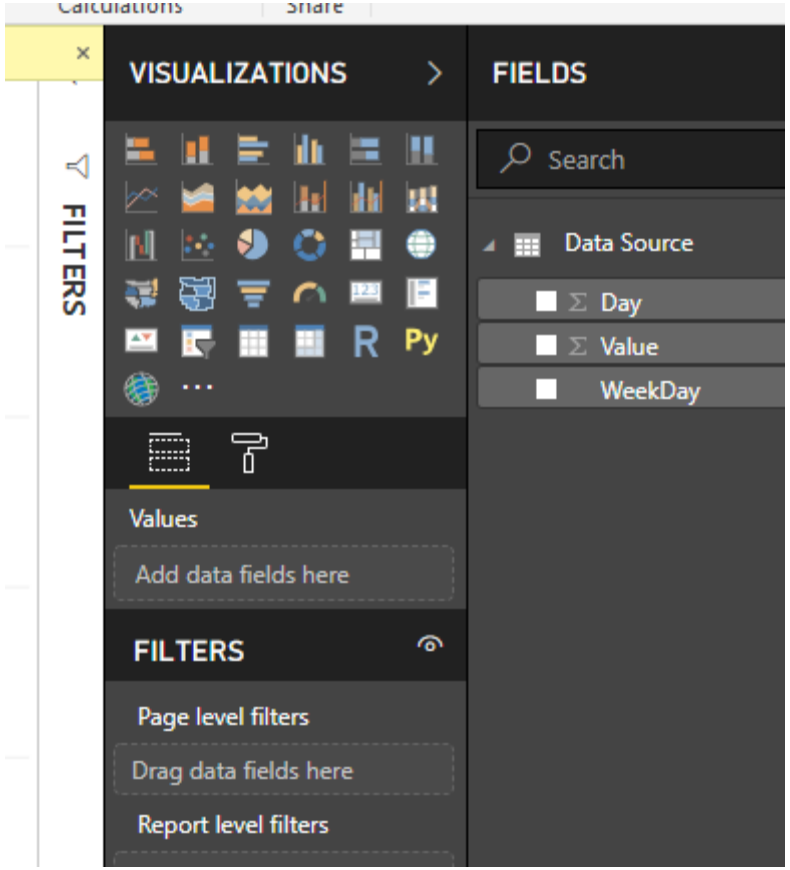
Example:
datasource = pandas.read_csv=('c:/datasource.csv')

2- Power Query



Example: dataset= dataset.fillna(dataset.mean())

3- Visualization



Example: sns.swarmplot(x="WeekDay", y="Value", data=dataset")

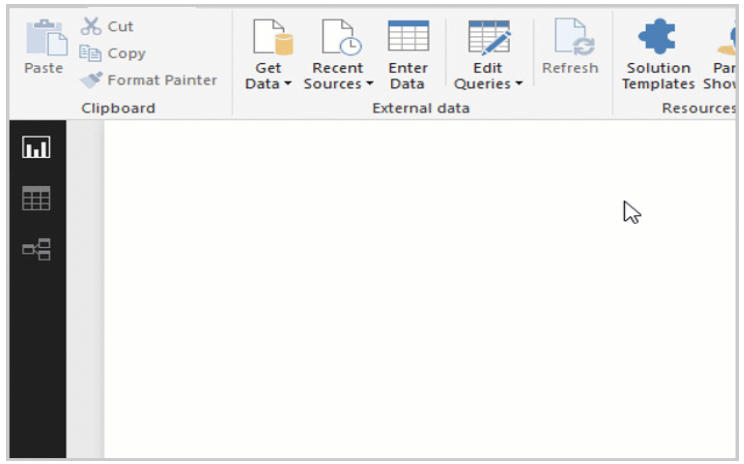
Limitations:

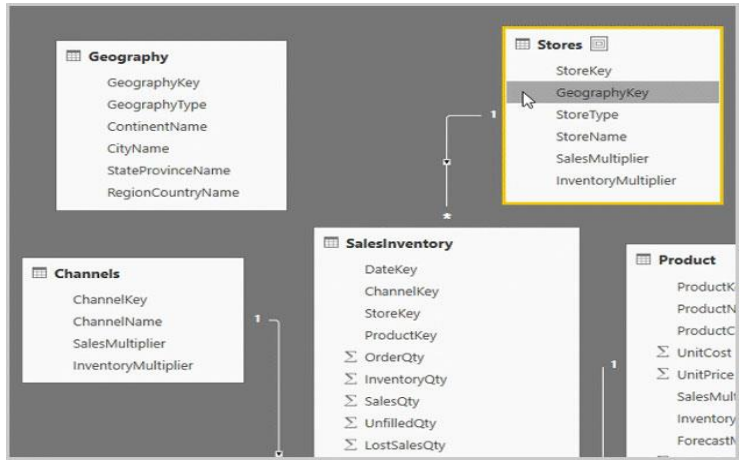
- Data size limitations – data used by the Python visual for plotting is limited to 150,000 rows. If more than 150,000 rows are selected, only the top 150,000 rows are used and a message is displayed on the image.
- Calculation time limitation – if a Python visual calculation exceeds five minutes the execution times out, resulting in an error.
- Relationships – as with other Power BI Desktop visuals, if data fields from different tables with no defined relationship between them are selected, an error occurs.
- Python visuals respond to highlighting other visuals, but you cannot click on elements in the Python visual in order to cross filter other elements.
- Can not be published to the web service yet.

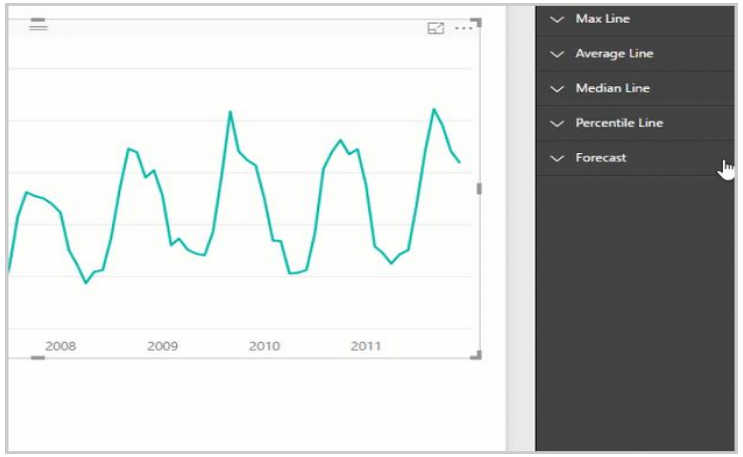
Facebook Sentiment Analysis Demo

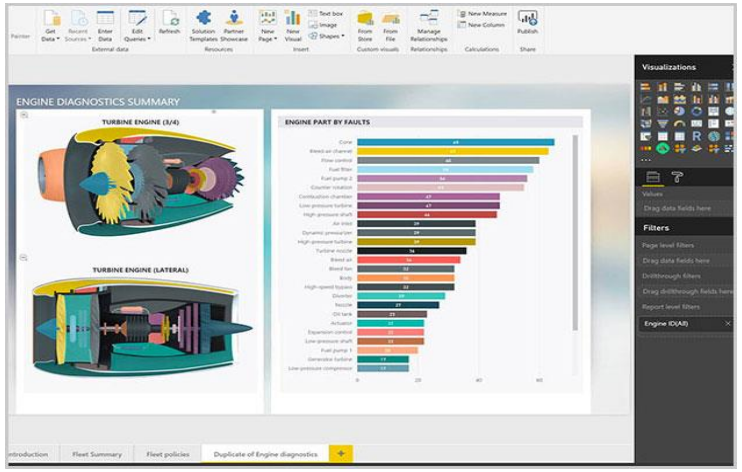
Q&A

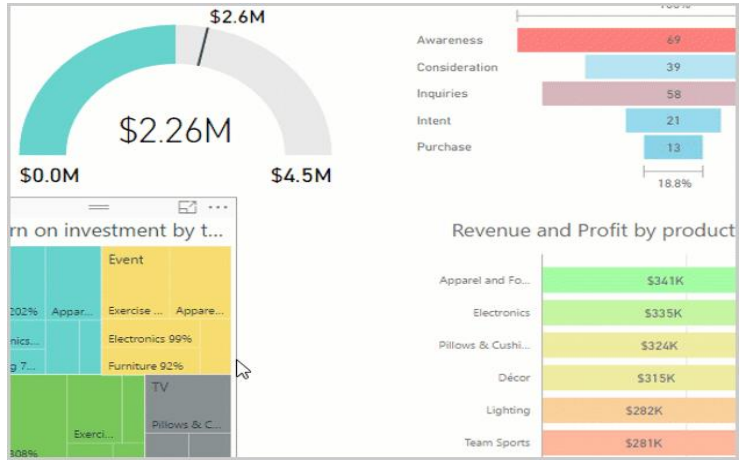
Power BI
Experience your data
Any data, any way, anywhere

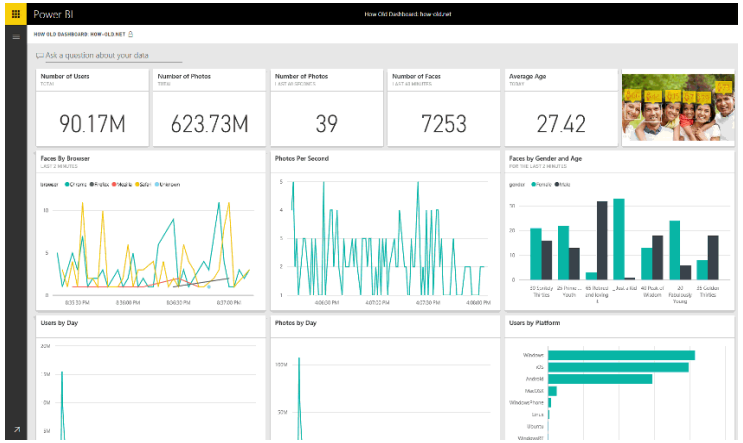












Thanks to the Organizers and Sponsors!