Excel and PowerBI

Data Analysis and Visualization

Analysis & Visualization in Excel

> Data Formatting – Tables

> Data Cleanse – Duplication, Missing Data

> Derived Data – Formula's & In-Built Functions

> Conditional Formatting

> Spark Lines

> Charts

> Slicing and Dicing – Pivot Tables & Charts

> **A multi-disciplinary field:**
> - Statistics & Probability
> - Machine Learning and Data Mining
> - Pattern Recognition
> - Natural Language Processing and Text Mining
> - Statistical Computing Languages

orion™
Business Innovation

STATISTICS AND PROBABILITY

> Probability deals with predicting likelihood of future events

> Statistics help us understand the past events better.

> Statistics is treated as one of the pillars of Data Science.

> Probability is an important part too.

orion™
Business Innovation

Basic Statistics

- Numerical - Continuous (defined range) and Discreet
- Strings – Categorical
- Data Categories
  - Population (referred with Capital case)
  - Sample (referred with small case)
- Central Tendency
  - Mean – Average value, formula : Total / (Count)
  - Median – is the Middle Value, Median Position : (Count+1)/2 , Median value is data in that Median position
  - Mode – Most Frequent value
- Normal Distribution (Bell Curve) - Mean, Median & Mode are close to each other & 90% of your data should fall between -2 & +2 Standard Deviation
  - Other distributions are Binomial, Degenerate, Poisson, Chi-Squared, Gamma
- Other Calculations
  - Range – Max value – Min Value
  - Variance – For Population : [Sum of (each value – Mean)^2] / (Count) & For Sample : [Sum of (each value – Mean)^2] / (Count -1)
  - Standard Deviation – Square Root of Variance
  - Standard Error – Its used in Sample, SD/(Square root of Count of Sample Data)
  - P-value – by what margin the comparative data can be off depends on the business area
- Comparative Statistics
  - Z-Test or T-Test : To look at past history and predicting for future
  - df : degrees of Freedom

Various Statistical Values

> **Mean**
>> Population Mean
$$\mu = \frac{\sum\limits_{i=1}^{N} X_i}{N}$$
>> Sample Mean $\longrightarrow$ $\bar{x} = \frac{\sum\limits_{i=1}^{n} x_i}{n}$

> **Median**
>> Position – (Total Count + 1)/2
>> Median Value is the value at Median Position

> **Mode**
>> Most frequent visible value

Range and Variance

> **Range**
>> Max value – Min value

> **Variance**
>> Formula $\sigma^2 = \dfrac{\sum\limits_{i=1}^{N}(x_i - \mu)^2}{N}$

> **Standard Deviation**
>> Formula $\sigma = \sqrt{\sigma^2}$

> **Standard Error**
>> Formula $SE = \dfrac{s}{\sqrt{n}}$

orion™
Business Innovation

Power BI

> Power BI is a suite of business analytics tools to analyze data and share insights. Power BI dashboards provide a 360-degree view for business users with their most important metrics in one place, updated in real time, and available on all of their devices.

> Flavors
>> PowerBI Free
>> PowerBI Pro

> Various components
>> PowerBI Desktop – Feature Rich data mashup and report authoring tool
>> PowerBI Mobile Apps – Access your reports anywhere with real-time updates
>> PowerBI Service – allows to publish reports securely to the organization with automatic refresh

> Demo 1 - https://fast.wistia.net/embed/iframe/y7obyb10u1 (from 34 sec)

orion™
Business Innovation

Flavors of Power BI

Free
- Data refresh frequency: Daily
- Data capacity Limit: 1GB/user
- Streaming rate: 10K rows/hour
- Data sources are limited to content packs for services and importing files

Pro
- Data refresh frequency: Hourly
- Data Capacity Limit: 10GB/user
- Streaming rate: 10M rows/hour
- Data Sources include free ones plus direct query dataset and **on-premises data**
- Collaboration with content packs

Power BI – Component Integration
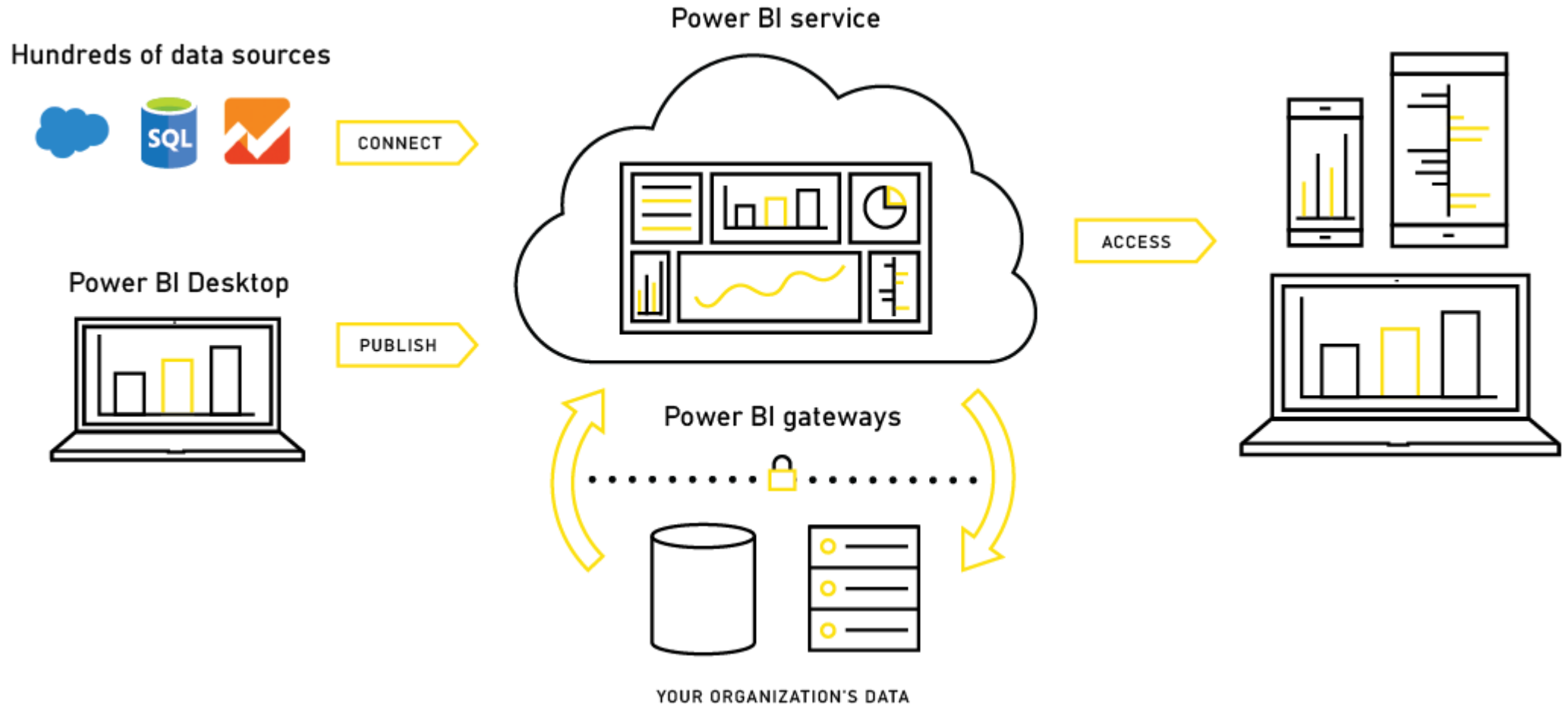
Why use PowerBI over Excel

› Excel over the years have improved a lot with Power Pivot, Power Map, Power View and Power Query

But PowerBI is more interactive for data designers with additional features

› Impressive list of Data Source Integrations (https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-data-sources/)

› You can connect to On-Premise and Cloud data

› Capability to easily edit the data in-turn interactively changing visualizations

› Live Dashboards – with streaming real-time data

› Mobile Friendly reports

orion™
Business Innovation

Datasets

- Import
  - Files
    - Local
    - OneDrive
    - SharePoint – Team Sites

▶ Connect

  ▶ Services

    ▶ WebTrends

    ▶ SalesForce

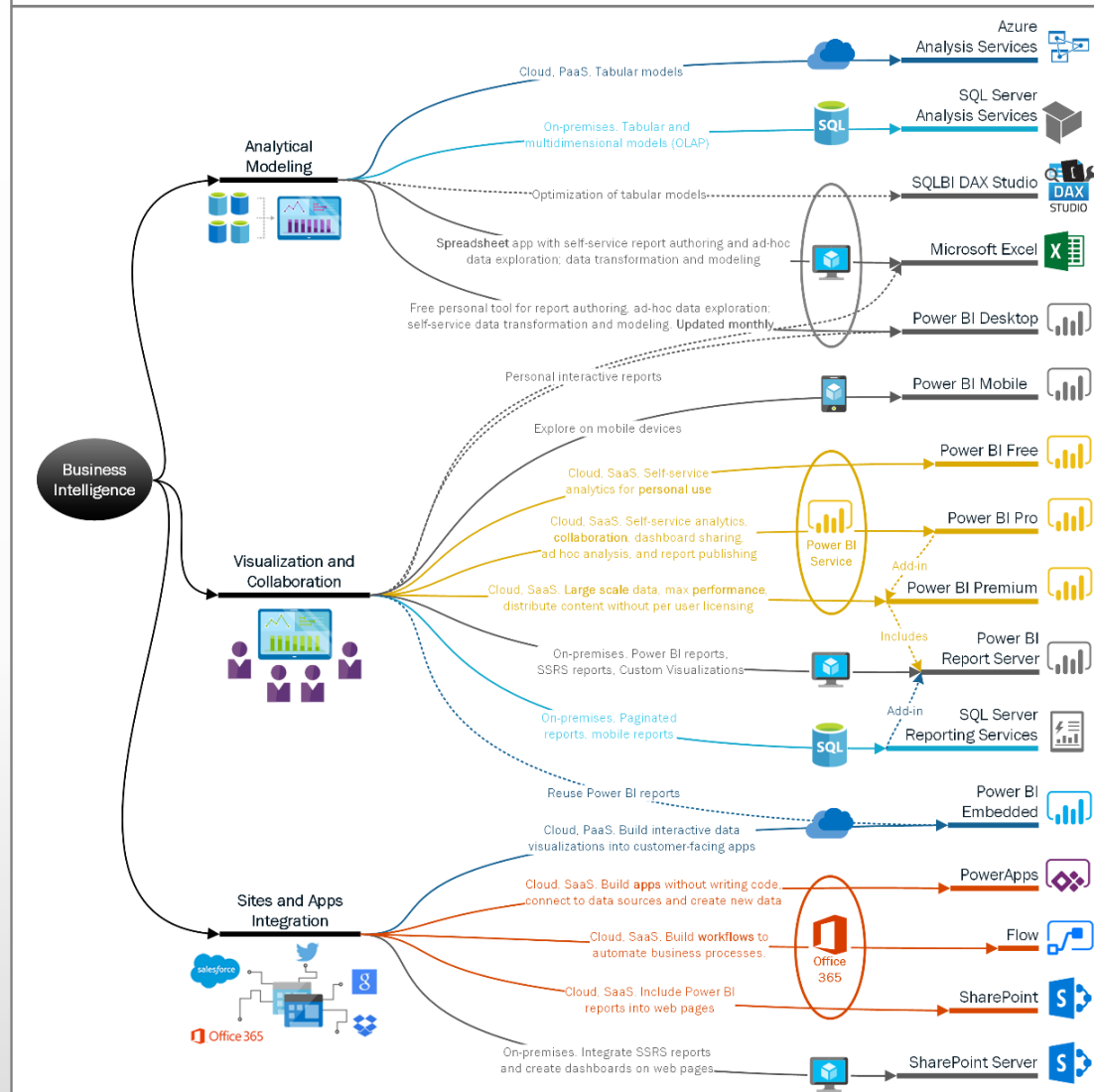   ▶ Google Analytics

  ▶ Databases

    ▶ Azure SQL Database

    ▶ Azure SQL Data Warehouse
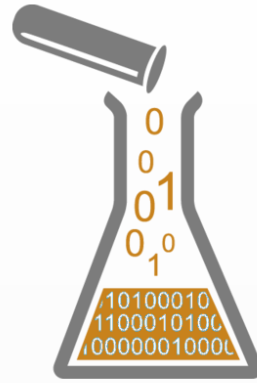
    ▶ SQL Server Analysis Services (Tabular)

  ▶ Content Pack

orion™
Business Innovation

Business Intelligence Solutions Decision Tree

Version 1.07. May 18, 2017. Author: Ivan Kosyakov, Ph.D., Technical Architect in Microsoft Technology Center New York

Business Intelligence

Analytical Modeling

- Cloud, PaaS. Tabular models → Azure Analysis Services
- On-premises. Tabular and multidimensional models (OLAP) → SQL Server Analysis Services
- Optimization of tabular models → SQLBI DAX Studio
- Spreadsheet app with self-service report authoring and ad-hoc data exploration; data transformation and modeling → Microsoft Excel
- Free personal tool for report authoring, ad-hoc data exploration; self-service data transformation and modeling. **Updated monthly** → Power BI Desktop

Visualization and Collaboration

- Personal interactive reports → Power BI Mobile
- Explore on mobile devices
- Cloud, SaaS. Self-service analytics for **personal** use → Power BI Free
- Cloud, SaaS. Self-service analytics, **collaboration**, dashboard sharing, ad hoc analysis, and report publishing → Power BI Pro
- Cloud, SaaS. **Large scale** data, max performance, distribute content without per user licensing → Power BI Premium
- Power BI Service
- Add-in
- Includes
- On-premises. Power BI reports, SSRS reports, Custom Visualizations → Power BI Report Server
- On-premises. Paginated reports, mobile reports → SQL Server Reporting Services
- Add-in
- Reuse Power BI reports → Power BI Embedded
- Cloud, PaaS. Build interactive data visualizations into customer-facing apps

Sites and Apps Integration

- Cloud, SaaS. Build **apps** without writing code, connect to data sources and create new data → PowerApps
- Cloud, SaaS. Build **workflows** to automate business processes. → Flow
- Cloud, SaaS. Include Power BI reports into web pages → SharePoint
- Office 365
- On-premises. Integrate SSRS reports and create dashboards on web pages → SharePoint Server

Legend

- **Microsoft Azure**
- **SQL Server (on-prem)**
- **Office 365 components**
- **Local or virtual machine**
- **Power BI Service**
- **Mobile devices**

Download Business Intelligence Solutions Decision Tree at http://biz-excellence.com/2017/05/16/bi-decision-tree

orion™ Business Innovation

THANK YOU

Kunal Kumar Shah

https://www.linkedin.com/in/kunalkumarshah/

orion™
Business Innovation