DEMO cum IMPORTANT JOB AND REAL-TIME CURRENT [PRE-REQUISITE FOR CLASSES]

VINAY TECH
www.vinaytechhouse.com

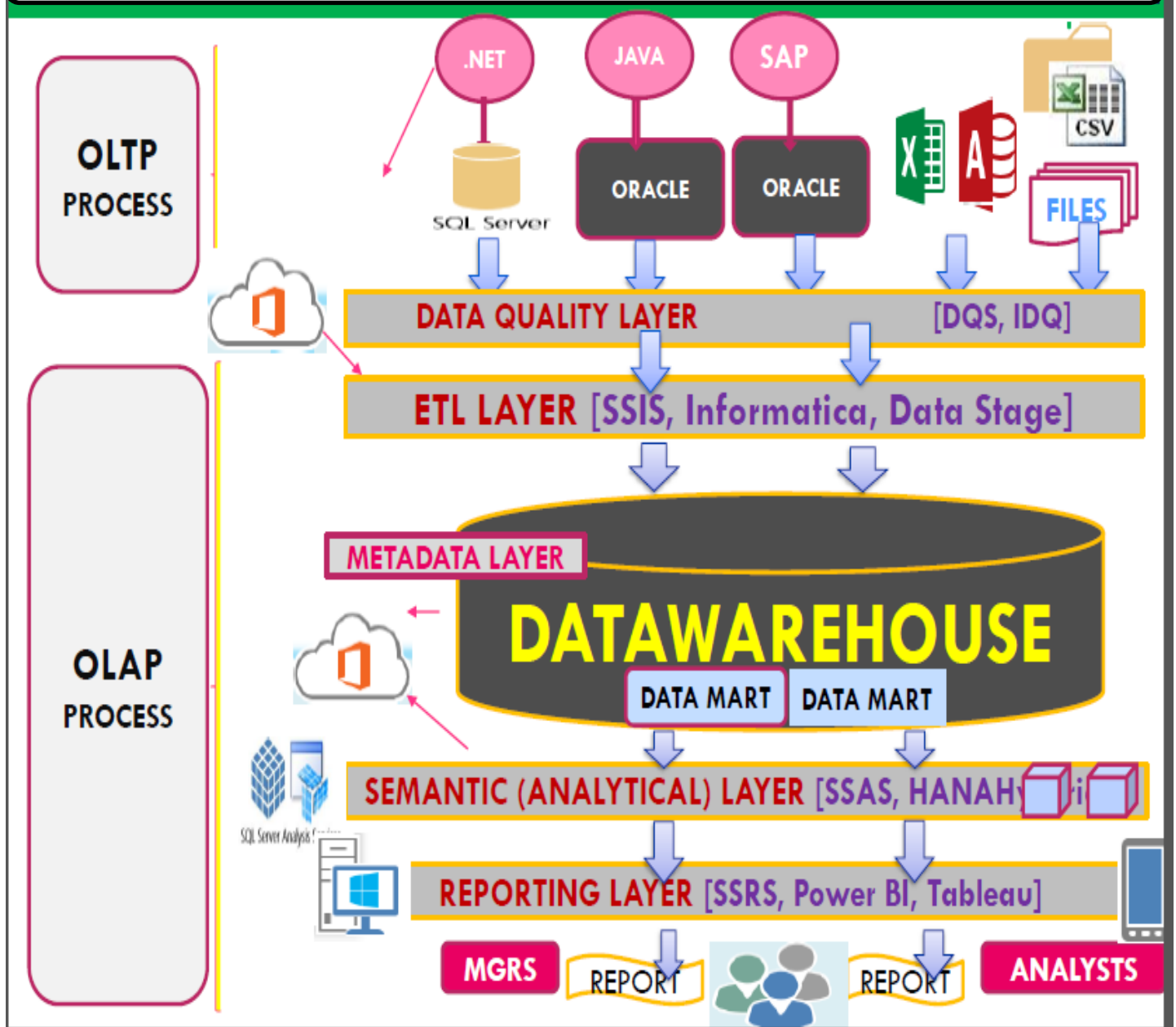## MSBI-DEMO
9573168449 / 040 66638869

NON IT GRADUATE	IT GRADUATE
What is Data?	What is BI? And What is AI?
Types of data available in IT?	What is Data Warehouse? Data Lake?
What is storage area?	Legacy DWH process (End-End)?
Types of storage areas available?	Modern DWH process?
What is IT application?	Data Science vs. DWH-BI?
Major IT application processes?	MSBI hw and software requirements?
	MSBI tools & utilities in real time?
	MSBI Projects real time flow.
	MSBI Opportunities

IT APPLICATIONS DATA STORAGE AND DECISION MAKING

BUSINESS DATA STORAGE	BUSINESS DATA ANALYSIS
These applications called as "OLTP [Online Transactional Processing]"	These applications called as "OLAP [Online Analytical Processing]"
Front-End and Back-End Approach	Layered Approach
Transactional or Operational System	DWH and BI System
Example:	
Irctc application [Web, Desktop, APP]	Irctc Confirmation Probability
Tatkal ticket booking @10 AM stores PNR, Amount, Start Station, Destination, Amount, Trainer Number Etc...	Tatkal analysis Number of trains having full booking Number of trains not having booking Station with huge demand

MODERN DATAWAREHOUSE AND BUSINESS INTELLIGENCE SYSTEM

Data warehouse: Data "Huge Storage"

Usually a database acts like warehouse if it stores more data.

Databases used for warehouses: Teradata, Oracle, Sqlserver, DB2, Sybase etc...

ETL: Extract Transform Load:

Retrieves the data, apply Transformations, and then Loads the data.

Tools:

Informatica, SSIS, Datastage, Abinitio, ODI, Talend, SAS-ETL, Pentaho ETL, etc...

Semantic Layer:

Converts normal data into decision making data.

   There are two types of decision conversion

   a) Multidimensional data presentation

   b) Tabular data presentation

Tools:

SSAS, Hyperion Essbase, SAP BO Cubes etc...

Reporting Layer:

Represents the data with visuals [Chart, Matrix, Table, KPIs, Gauges etc...]

Tools:

SSRS, Power BI, Cognos, SAP BO, Microstrategy, OBIEE, SAS, Tableau, Qlik View etc...

Data Quality:

Ensures data quality of before processing through ETL by "profiling", "scrubbing", and "cleansing"

Tools:

DQS, IDQ.

Metadata Management:

Manages master data or metadata of the objects in the data warehouse.

Tools:

MDS, MDM.

MSBI [MICROSOFT BUSINESS INTELLIGENCE]

PROPER PRONUNCIATION & TERMINOLOGY

MSBI is a SUITE / STACK / PACK / TOOLS which perform the below operations using 5 different kinds of ingredients.

Operations:

a) Data Quality Assurance

b) Data Integration and gathering [ETL & ELT] for the available data

c) Master Data or Metadata management

d) Semantic Model creation for better decision support & SSBI

e) On Premises (majorly), Cloud storage based Analytical and Visualization Reporting tool

MSBI [MICROSOFT BUSINESS INTELLIGENCE]

PROPER PRONUNCIATION & TERMINOLOGY

MSBI is a SUITE / STACK / PACK / TOOLS which perform the below operations using 5 different kinds of ingredients

Components

SERVER INTEGRATION SERVICES [SSIS]

SQL SERVER ANALYSIS SERVICES [SSAS]

SQL SERVER REPORTING SERVICES [SSRS]

DATA QUALITY SERVICES [DQS]

MASTER DATA SERVICES [MDS]

DAY-DAY BUSINESS PROCESS [OLTP]	DAY-DAY ANALTYLICAL PROCESS [OLAP]
Front-End App----->Back-end DB	Layered Process
(Ex: .Net) (Ex: Sql Server)	OLTP→ODS→Stage→EDW→DW
• Web or Desktop pages, apps etc...	• Analytical and reporting tools
• Daily data storage	• Daily data analysis
• Less volume of data	• More volume of data
• Current data [last one month, quarter or year]	• Years and years data
Database in less size	**Database in more size**
Database in Normalized format	Database in DE normalized format

IMPORTANT TERMINOLOGIES

[DWH] DATA WAREHOUSE [Data "Huge Storage"]

A <u>database</u> which has <u>more storage</u> and designed for <u>decision making</u>.

[BI] BUSINESS INTELLIGENCE [Collect, Analyze and Present]

The term Business Intelligence (BI) refers to technologies, applications and practices for the collection, integration, analysis, and presentation of business information for better business decision making.

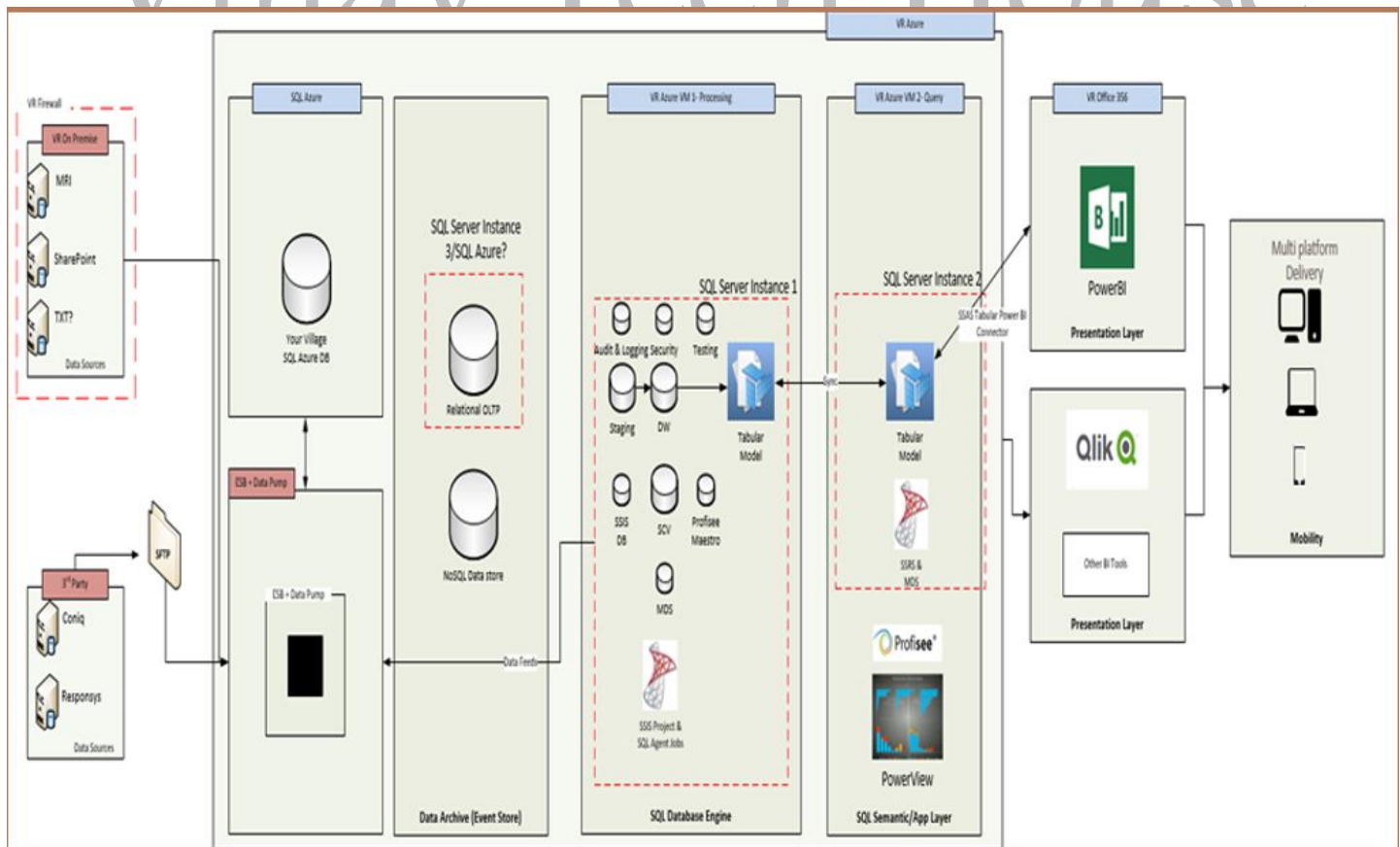### DATA SCIENCE [Math + Stats + Processes + Machine Methods]

Data science is a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data.

[AI] ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is an area of computer science that emphasizes the creation of intelligent machines that work and react like humans. Some of the activities computers with artificial intelligence are designed for include: a) Speech recognition b) Learning c) Planning d) Problem solving

MSBI Design Reason and end-end steps

1. **Get the data from OLTP**
2. **Process in multiple stages**
3. **Load into Data warehouse**
4. **Create cubes [Tabular or Multi]**
5. **Generate visualization reports**
6. **Publish Reports to Web Portal**
7. **Pull or push [Subscription] reports**
8. **Mobile presentation of reports**
9. **Decision Making [Business Intelligence]**
10. **Customer reviews**

POWER BI [MICROSOFT POWER BI] —SSBI

DEFINITION

POWER BI is a

"   **SELF SERVICE,**
Cloud based and On-Premises,
Any Type of Data load,
Data Model Supportive,
Analytical, and
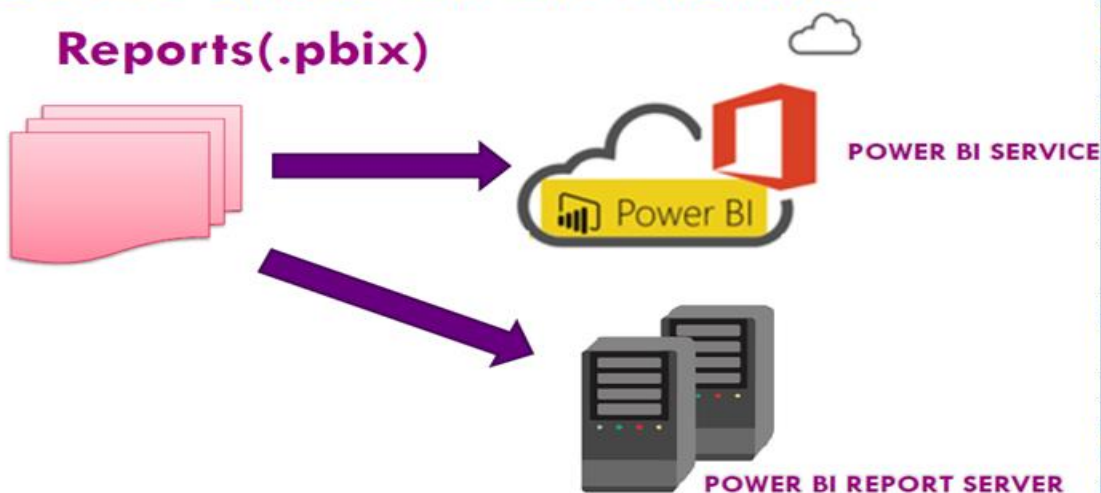**Visualization Reporting tool**   "

Designed for DWH & BI solution, and Data "Insights and Analytics".

SELF SERVICE

Personal Analysis

Team Analysis

Corporate Analysis

Designer Analysis ***

Customer or user analysis ***

CLOUED BASED OR ON-PREMISES

Reports(.pbix)

POWER BI SERVICE

Power BI

POWER BI REPORT SERVER

Any Type Of Data Load

a) Structured b)Semi-Structured c)Unstructured

Structured Data

0.103	0.176	0.387	0.300	0.379
0.333	0.384	0.564	0.587	0.857
0.421	0.309	0.654	0.729	0.228
0.266	0.750	1.056	0.936	0.911
0.225	0.326	0.643	0.337	0.721
0.187	0.586	0.529	0.340	0.829
0.153	0.485	0.560	0.428	0.628

Unstructured Data

Any Type Of Data Load

Unstructured
PDFs, JPEGs, MP3, Movies, ...

Semi-structured
CSV, JSON, XML, MongoDB, ...

Structured
Oracle, MSSQL, MySQL, DB2, ...

Analytical [Browse, view, and take decisions]

< Back to Report   NET PAYMENT FLOW   BY MONTH AND YEAR

● Increase ● Decrease ● Other ● Total

Month	2,017.00	2,018.00	2,019.00
	5,000.00	5,000.00	
January	-1,509.00	-503.00	-503.00
February	-5,010.00	-1,670.00	-1,670.00
March	14,406.00	4,802.00	4,802.00
April	-3,594.00	-1,198.00	-1,198.00
May	-10,578.00	-3,526.00	-3,526.00
June	5,478.00	1,826.00	1,826.00

MSBI VS POWER BI, MSBI VERSION CHANGES AND POWER BI INTRODUCTION

MSBI	POWER BI
END_END full depth BI Tool [ETL, SEMANTIC, REPORTING, QUALITY, and MASTER DATA MGMT]	Analytical reporting tool with partial support of ETL, Model and Master data.
2005 introduced	2013 announced, official release 2015
Equivalent to Power BI: SSRS	SSRS enhanced version is Power BI SSRS 2012 + Power View + Power Query + DAX = Power BI
Less standard visuals	More inbuilt and support custom visuals
On-Premises storage	Initially cloud, now adays on premises also

POWER BI HISTORY AND EVALUATION

PowerPivot – Excel add-in using xVelocity(Vertipaq) columnar engine

SQL Server 2012 SSAS Tabular Model

Microsoft BI Semantic Model (BISM)

PowerView

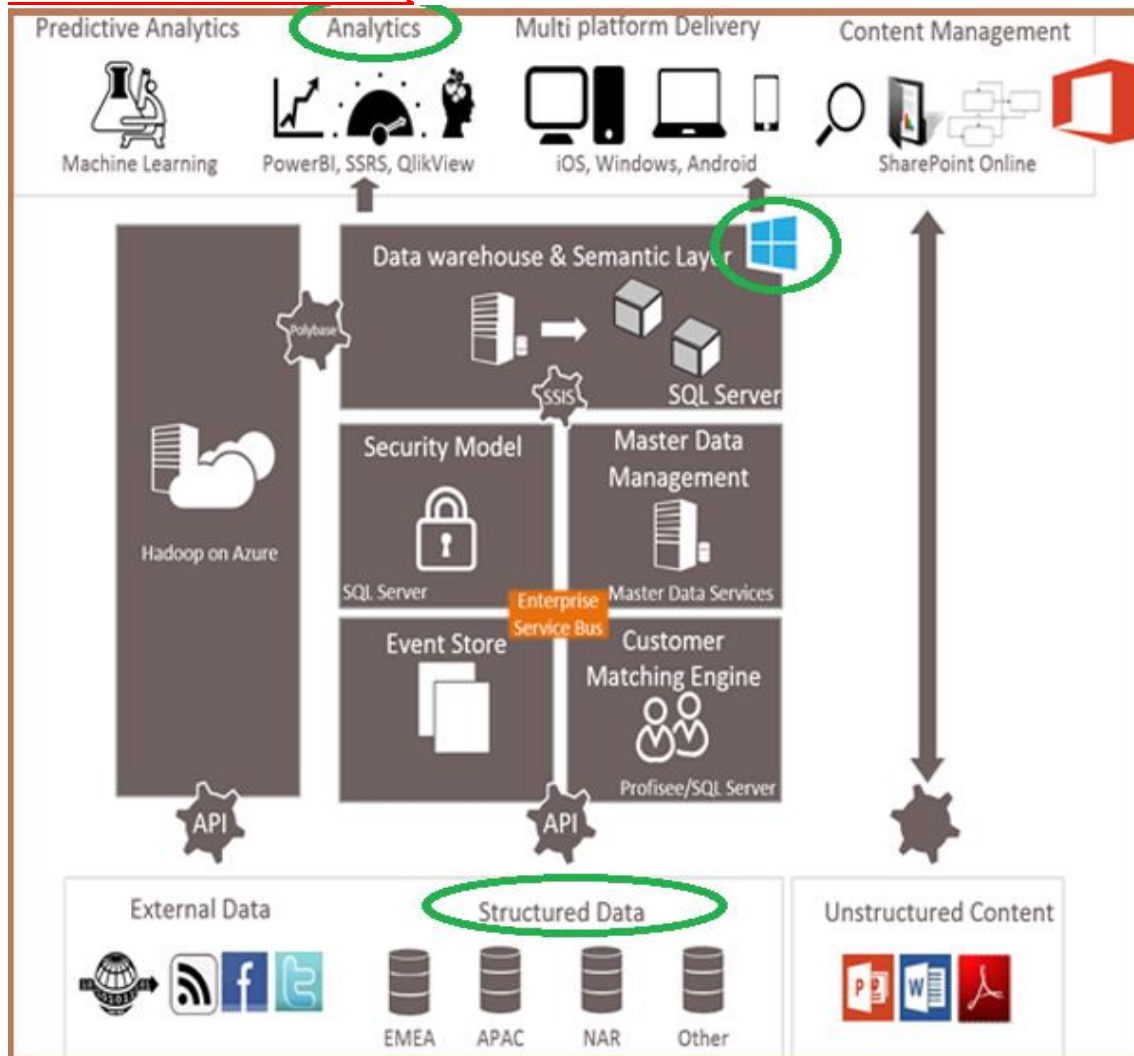Re-branding to Power BI (notice the space)

SharePoint and Office 365 integration (Q&A, Power Query, and Power Map)

Power BI v2 Preview (December 2014)

Power BI v2 Service & Desktop GA (July 2015)

REAL-TIME COMPANY FLOW [

MSBI [MICROSOFT BUSINESS INTELLIGENCE] –SSBI

SQL HISTORY → SQL SERVER HISTORY → SQL SERVER BI HISTORY [MSBI]

PARTIAL-BI [before Sql Server 2000]

DATA TRANSFORMATION SERVICES [DTS]
MICRO SOFT ANALYSIS SERVICES [MSAS]

FULL BI [After 2005]
MSBI 2005
MSBI 2008
MSBI 2008 R2
MSBI 2012
MSBI 2014
MSBI 2016
MSBI 2017
MSBI 2019

MSBI REQUIREMENTS AND INSTALLATION PROCESS

MSBI-2012 Windows 7 Service pack1 or above
SINGLE "SETUP.EXE"--install SSDT [IS, AS, RS], DQS, Database Engine

MSBI-2016 Windows 10 or above
Two separate components need to install
a) SSDT. Exe b) SSMS

MSBI-2017 Windows 10 or above
Three separate components need to install
a) SSDT. Exe b) SSMS c) Sql Server Reporting Services

Note: Installation Steps
www.vinaytechhouse.com → Faqs→

MSBI DAY-DAY COMPONENTS IN REAL-TIME

SSDT-SQL SERVER DATA TOOLS

To work with SSIS, SSAS [Tabular and Multidimensional] and SSRS
applications.

SSMS-SQL SERVER MANAGEMENT STUDIO
To work with SSIS, SSAS [Tabular and Multidimensional], SSRS, DQS and
MDS databases.

COMMAND LINE UTILITIES & COMMANDS
SSIS→ DTEXEC, DTEXECUI, DTUTIL
SSAS→ASCMD
SSRS→RS, RS CONFIG, RSKEYMGMT
SQL SERVER PROFILER
To improve the performance of Database Engine and Analysis
Services databases.
SQL SERVER AGENT
To schedule IS, AS, RS, SQL and other Micro Soft Applications

MSBI Opportunities

EXPERIENCE BELOW 7YEARS	SENIOR ROLES (>5 years)
MSBI Suite Developer	MSBI Solution Specialist
MSBI Combination Deve (Any 2)	MSBI Designer & Modeler
MSBI Single Component Developer	MSBI BI Architect
.Net + (SSIS or SSRS)	
MSBI Testers (ETL Testers)	**MSBI + POWER BI Opportunities**
SQLServer Developer + MSBI	MSBI + Power BI Knowledge
SQLServer DBA + MSBI	MSBI + Power BI 2 years
MSBI Freshers	
DWH (ETL or Reporting) + MSBI	

Vinay Tech House

MSBI BI DEMO- FAQS

Q: What are the types of data available in current IT industry?

A:

Three types

a) Unstructured

b) Semi structured

c) Structured

Q: What are the storage areas available in IT industry?

 A:

Three types

a) File Stream Storage

b) Databases

c) Cloud storage

Q: What is Power BI?

A:

POWER BI is a SELF SERVICE Cloud based Data Load and Model supportive Visualization reporting tool.

Q: What is MSBI?

A:

MSBI is a **SUITE / STACK / PACK / TOOLS** which has the below ingredients

SERVER INTEGRATION SERVICES [SSIS]

SQL SERVER ANALYSIS SERVICES [SSAS]

SQL SERVER REPORTING SERVICES [SSRS]

DATA QUALITY SERVICES [DQS]

MASTER DATA SERVICES [MDS]

Q: What is Self Service BI?

A:

If an application support Personal, Team and Corporate analysis, then that application is SSBI application.

With respect to Power BI, SSBI indicate accessing, creation, managing{content packs and apps), reports and dash boards.

Note: Users who have Pro account, they consider as SSBI professionals.

Q: What is OLTP?

A: Online Transaction Processing, this process helps to store daily business data.

Q: What is OLAP?

A: Online Analytical Processing, this process gathers, convers and presents data.

Q: What is BI?

A:

It is the process of gathering, converting and presenting business data for decision making purpose.

Q: What is Data warehouse? [Business data]

A:

Data Huge Storage Area.

Q: What is a Modern Data Warehouse? [Social media, IOT, Business data and Cloud data {Lake}]

Modern Data warehouses comprised of multiple programs impervious to User.

Polyglot persistence encourages the most suitable data storage technology based on data. This "best-fit engineering" aligns multi-structure data into data lakes and considers NoSQL solutions for JSON formats. Pursuing a polyglot persistence dat strategy benefits from virtualization and takes advantage of the different infrastructure. Modern DW requires Petabytes of storage and more optimized techniques to run complex analytic queries.

Q: How many types of Data warehouses available?

A:

a) Structured storage Data warehouse [Databases used for this purpose]

b) Unstructured storage Data warehouse [File Streams and Cloud Storages used]

     Ex: Data Lake

Q: What is Internet of things [IoT]?

A:

The Internet of things (IoT) refers to the concept of extending Internet connectivity beyond conventional computing platforms such as personal computers and mobile devices, and into any range of traditionally "dumb" or non-internet-enabled physical devices and everyday objects. Embedded with electronics, Internet connectivity, and other forms of hardware (such as sensors), these devices can communicate and interact with others over the Internet, and they can be remotely monitored and controlled

Q: What is Data Insight?

A: The capacity to gain an accurate and deep understanding of someone or something

Q: What do you mean by Analytics?

A: The systematic computational analysis of data or statistics.

Information resulting from the systematic analysis of data or statistics.

"these analytics can help you decide if it's time to deliver content in different ways"

Analytics is the scientific process of discovering and communicating the meaningful patterns which can be found in data.

Q: What is Polybase?

A:

PolyBase is a technology that accesses and combines both non-relational and relational data, all from within SQL Server. It allows you to run queries on external data in Hadoop or Azure blob storage. The queries are optimized to push computation to Hadoop.

Q: What is Hadoop?

Hadoop is an open source distributed processing framework that manages data processing and storage for big data applications running in clustered systems. It provides massive storage for any kind of data, enormous processing power and the ability to handle virtually limitless concurrent tasks or jobs.

Q: What is Azure?

A:

Microsoft Azure is a cloud computing service created by Microsoft for building, testing, deploying, and managing applications and services through Microsoft-managed data centers.

Q: What can Microsoft Azure Do?

Microsoft maintains a growing directory of Azure services, with more being added all the time. All the elements necessary to build a virtual network and deliver services or applications to a global audience are available, including:

Virtual machines. Create Microsoft or Linux virtual machines (VMs) in just minutes from a wide selection of marketplace templates or from your own custom machine images. These cloud-based VMs will host your apps and services as if they resided in your own data center.

SQL databases. Azure offers managed SQL relational databases, from one to an unlimited number, as a service.This saves you overhead and expenses on hardware, software, and the need for in-house expertise.

Azure Active Directory Domain services. Built on the same proven technology as Windows Active Directory, this service for Azure lets you remotely manage group policy, authentication, and everything else. This makes moving and existing security structure partially or totally to the cloud as easy as a few clicks.

Application services. With Azure it's easier than ever to create and globally deploy applications that are compatible on all popular web and portable platforms. Reliable, scalable cloud access lets you respond quickly to your business's ebb and flow, saving time and money. With the introduction of Azure WebApps to the Azure Marketplace, it's easier than ever to manage production, testing and deployment of web applications that scale as quickly as your business. Prebuilt APIs for popular cloud services like Office 365, Salesforce and more greatly accelerate development.

Visual Studio team services. An add-on service available under Azure, Visual Studio team services offer a complete application lifecycle management (ALM) solution in the Microsoft cloud. Developers can share and track code changes, perform load testing, and deliver applications to production while collaborating in Azure from all over the world. Visual Studio team services simplify development and delivery for large companies or new ones building a service portfolio.

Storage. Count on Microsoft's global infrastructure to provide safe, highly accessible data storage. With massive scalability and an intelligent pricing structure that lets you store infrequently accessed data at a huge savings, building a safe and cost-effective storage plan is simple in Microsoft Azure.

Q: What s Data Lake?

Lake is a storage area for structured, semi structured and unstructured data.

There are multiple types of lakes.

A Hadoop data lake is a data management platform comprising one or more Hadoop clusters. It is used principally to process and store non relational data, such as log files, internet clickstream records, sensor data, JSON objects, images and social media posts.

Azure Data Lake Service is Microsoft's internal big data system where

Users can store structured, semi-structured or unstructured data produced from applications including social networks, relational data, sensors, videos, web apps, mobile or desktop devices.

A single Azure Data Lake Store account can store trillions of files where a single file can be greater than a petabyte in size.

What is Azure Data Lake Analytics?

Azure Data Lake Analytics is a parallel on-demand job service.

What is U-Sql?

A query language that combines SQL with C#.

Users can develop and run parallel data transformation and processing programs in U-SQL.

Vinay Tech House