

Korszerű felhőalapú analitikai megoldások kórházi környezetben

Abronits Péter

Technikai projektvezető

Grape Solutions Zrt.

Bemutatózás



BLACK BELT
Partner



Microsoft Partner

Gold Data Analytics
Gold Application Development
Gold Collaboration and Content
Silver Cloud Platform



Tartalom



mediViR

Környezet bemutatása

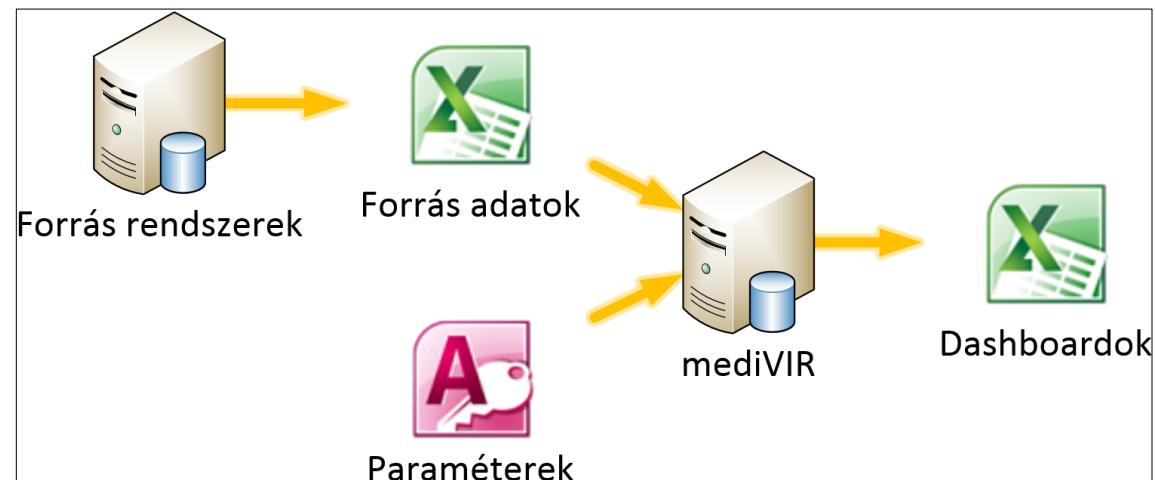
Esettanulmány

Q&A



Környezet bemutatása - a rendszer

- **Gazdasági modul**
- **Medikai modul**
 - KPI 3 Betegforgalmi teljesítmény adatok
 - KPI 4 Súlyszám, gyógyszerforgalom
 - KPI 5 Ambuláns kimutatások, elemzések
 - KPI 6 Igénybevett és végzett diagnosztika
 - KPI 7 Várólista
- **Humán modul**
- **Orvosi információs modul**
- **Területi ellátási kötelezettség lefedettsége (térképes, megye, kerület)**



Éves 2-4%
megtakarítás
(0,2-0,4 Mrd)

Környezet bemutatása - forrásrendszerek

Gazdasági modul

- EcoSTAT (xls)
- SAP (connector)

Medikai modul

- MedWorks (csv, xls, Oracle)
- HIS (csv)
- eMedSolution (csv, xls)
- PharmaGlobe (xls)
- GYURIKA-HospiPharma (csv)
- OEP visszaigazolások (DAT, DBF, PDF)

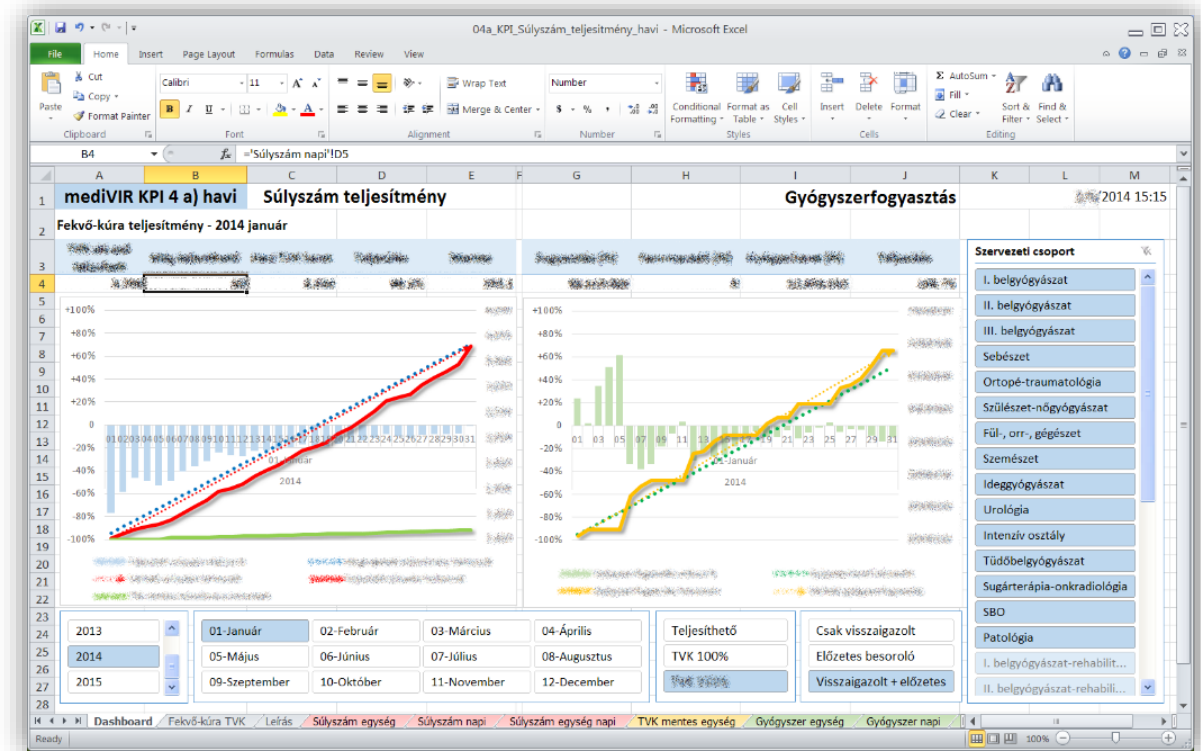
Humán modul

- KIRA (xml)
- JDolBer (xls)

Környezet bemutatása - keretgazdálkodás

- **Medikai modul**

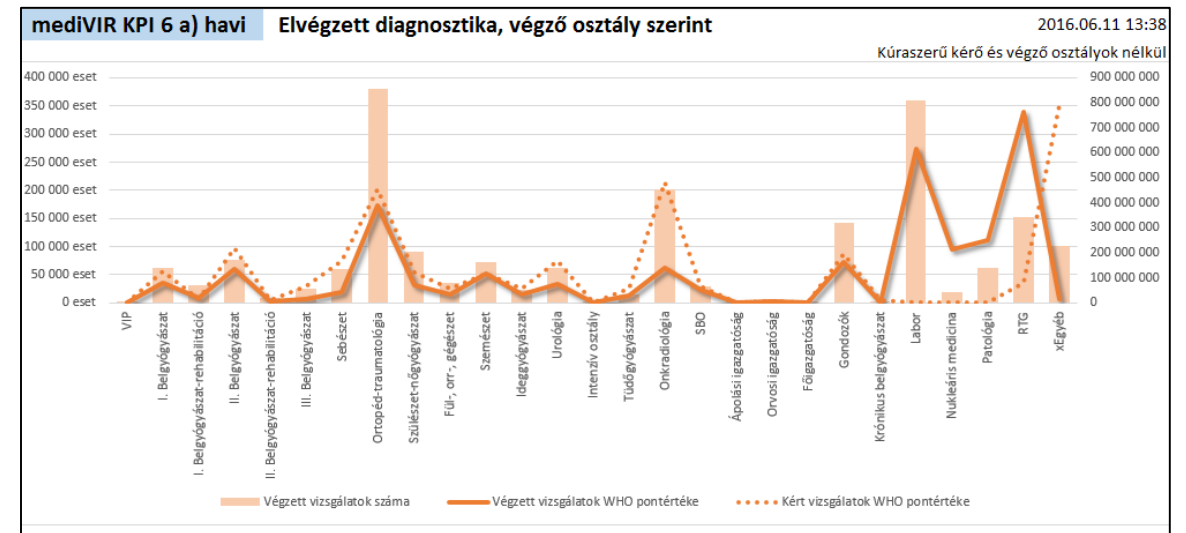
- KPI 3 Betegforgalmi teljesítmény adatok
- **KPI 4 Súlyszám, gyógyszerforgalom**
- KPI 5 Ambuláns kimutatások, elemzések
- KPI 6 Igénybevett és végzett diagnosztika
- KPI 7 Várólista



Környezet bemutatása - diagnosztika

- **Medikai modul**

- KPI 3 Betegforgalmi teljesítmény adatok
- KPI 4 Súlyszám, gyógyszerforgalom
- KPI 5 Ambuláns kimutatások, elemzések
- **KPI 6 Igénybevett és végzett diagnosztika**
- KPI 7 Várólista



Esettanulmány - helyzetkép

- ETL tulajdonságok
 - Napi egyszer
 - Cél: **a tegnapi adatok** munkakezdésig elérhetővé váljanak (pl. reggel 8 órára)
- Adatok mennyisége
 - Adatok 2011-ig visszamenőleg elemezhetőek
 - Kórház összevonások (pl.: Bács-Kiskun Megyei Kórház)
- Betöltéséhez szükséges idő
 - ETL + MOLAP Process = 100 + 12 = **112 perc**
 - Kocka process alatt az elemzés csak offline lehetséges

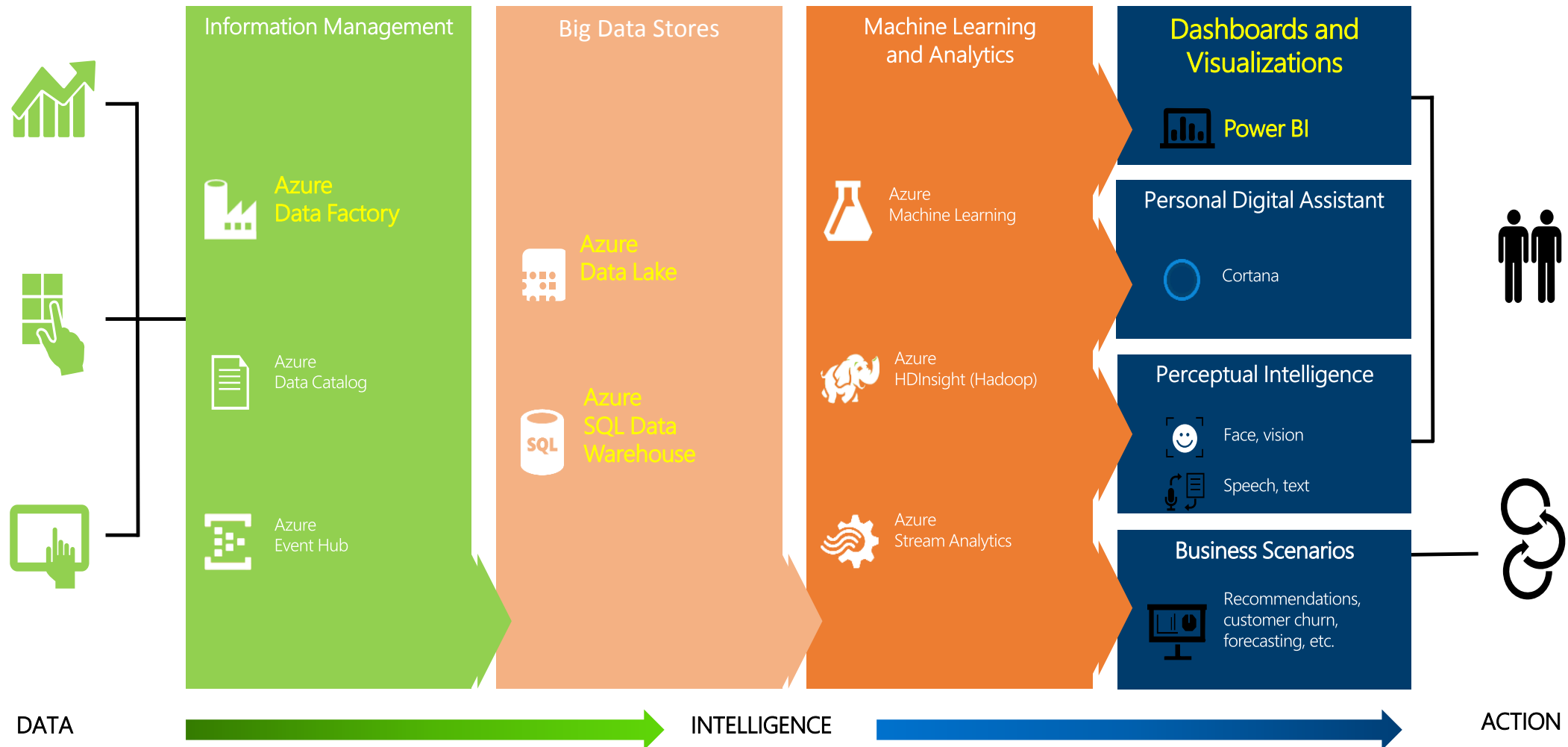
Esettanulmány - igények

- Betöltéséhez szükséges idő csökkentése
- Lekérdezések gyorsítása

Esettanulmány - megoldási lehetőségek

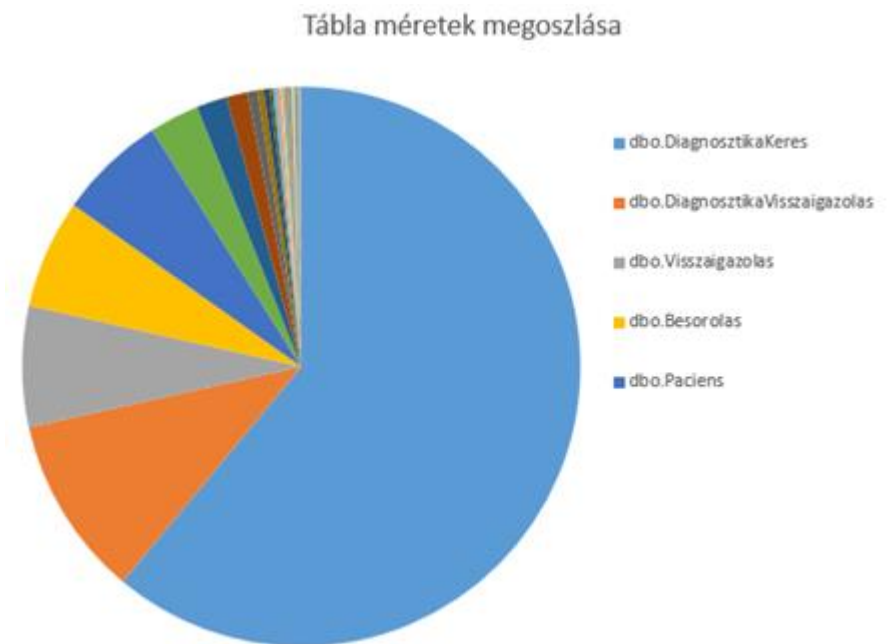
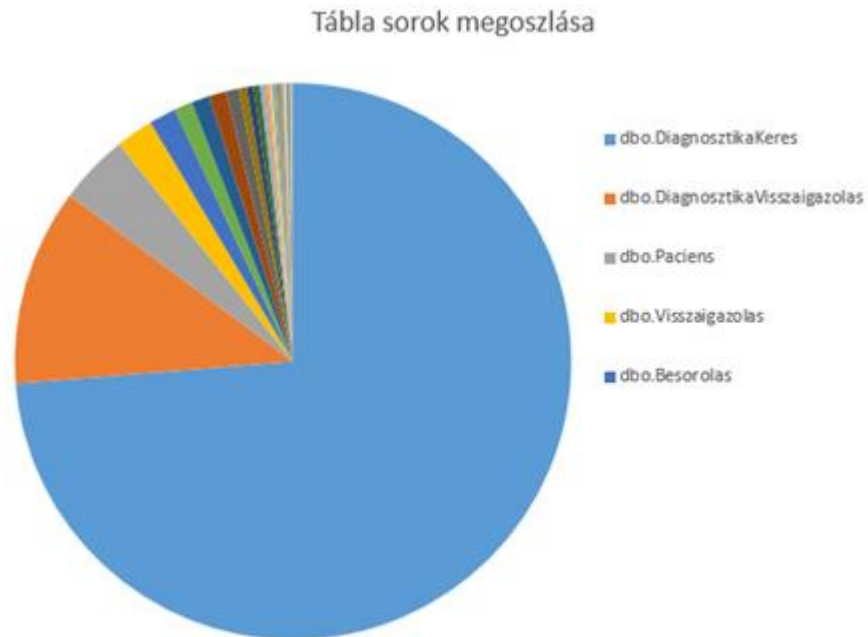
- Helyi kiszolgáló felfelé skálázása
- **Cortana Analytics Suite**

Cortana Analytics Suite



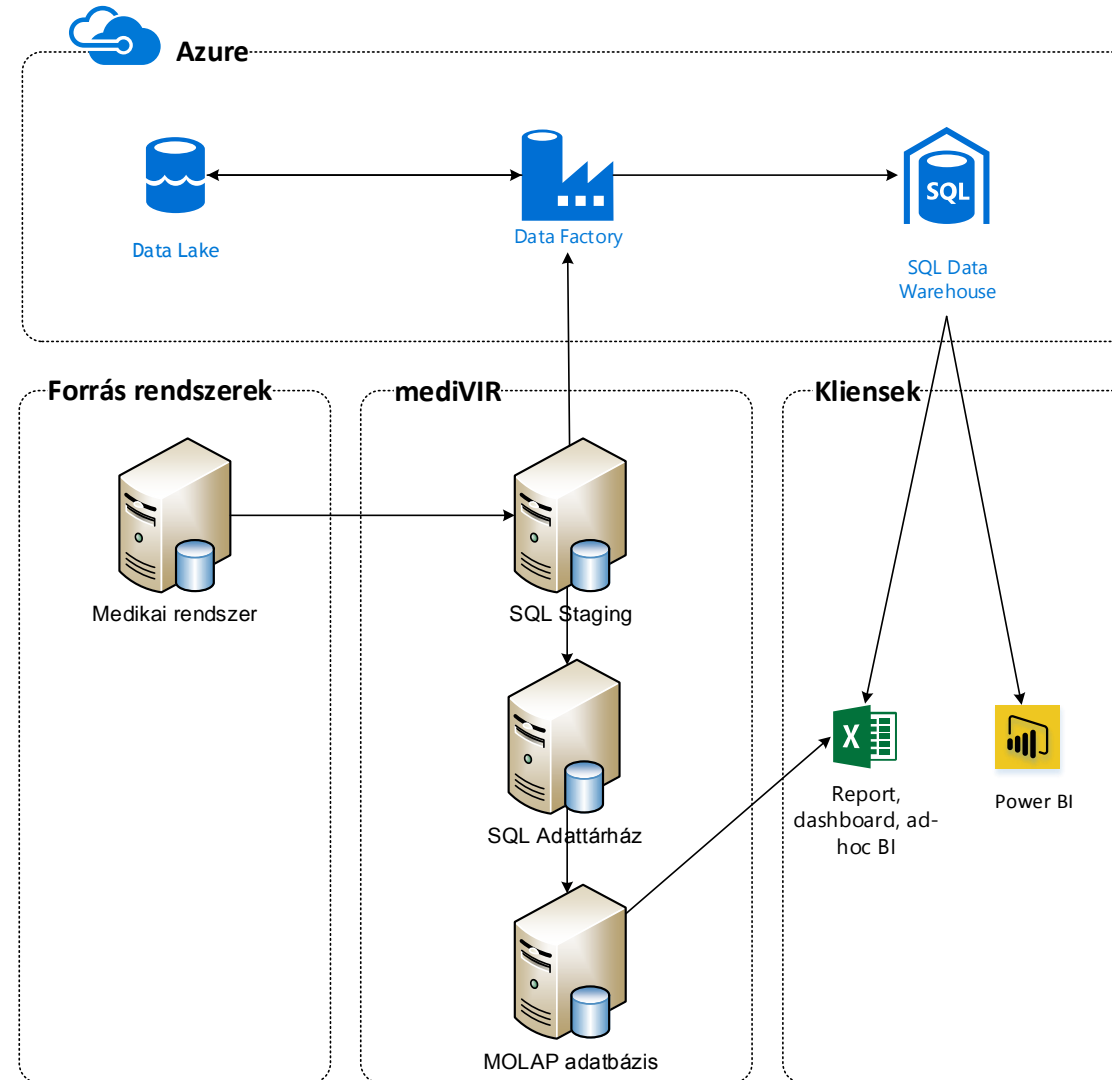
Esettanulmány - tervezés

- Adat körök meghatározása



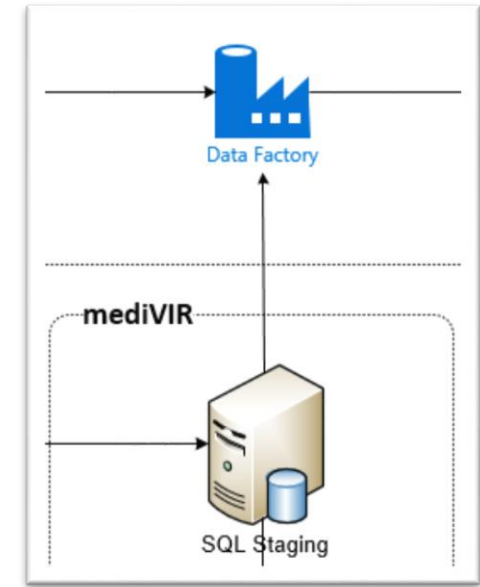
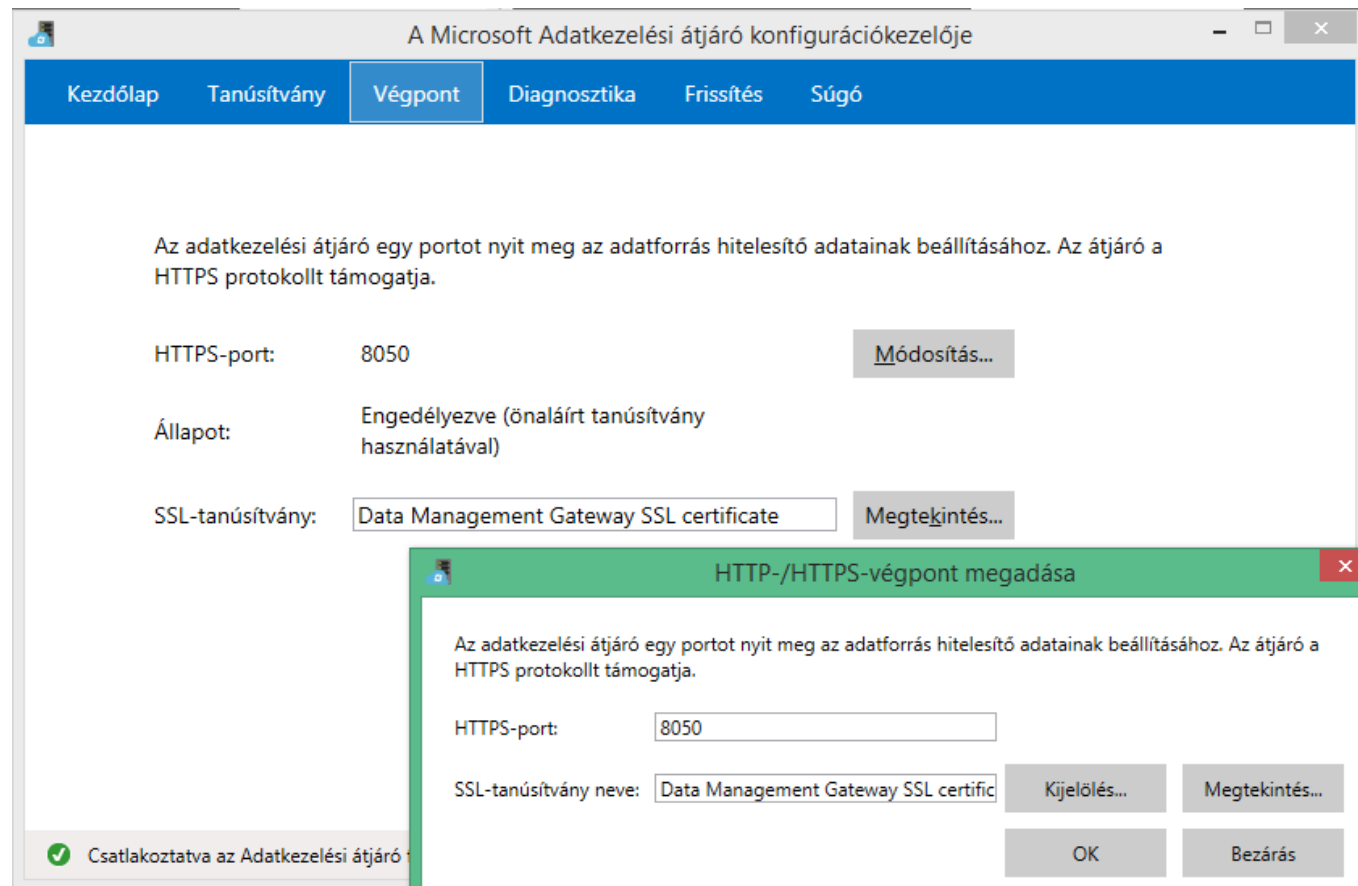
Diagnostika kérések és visszaigazolások

Esettanulmány - architektúra



Azure Data Factory (ADF)

- Data Management Gateway



Azure Data Factory (ADF)

The screenshot shows the Azure Data Factory (ADF) portal interface. At the top, the resource group is identified as 'medivir-dwh' (Data factory). Below this, there are 'Delete' and 'Move' buttons. The 'Essentials' section displays key information: Resource group (Demo), Location (northeurope), Subscription name (Grape Sponsorship IUR 6000\$), Provisioning state (Succeeded), and Subscription id (8c06f91-694b-4d82-9dfd-225dfef8c8a3). An 'All settings' button is visible. The 'Actions' section contains several tiles: 'Author and deploy', 'Copy data (PREVIEW)' (highlighted with an orange box), 'Monitor & Manage', 'Sample pipelines', 'Diagram', and 'Metrics and operations'.

medivir-dwh
Data factory

Delete Move

Essentials ^

Resource group
Demo

Location
northeurope

Subscription name
Grape Sponsorship IUR 6000\$

Provisioning state
Succeeded

Subscription id
8c06f91-694b-4d82-9dfd-225dfef8c8a3

All settings →

Actions

Author and deploy

Copy data (PREVIEW)

Monitor & Manage

Sample pipelines

Diagram

Metrics and operations

Azure Data Factory (ADF)

1 Properties

2 Source

SQL Server

3 Destination

Data Lake Store

5 Summary

COPY DATA

Enter name and description for the copy data task and specify how often you want to run the task.

Task name **(required)** i

CopyPipeline-ouc

Task description

Enter description here

Task cadence (or) Task schedule

- Run once now
 Run regularly on schedule

Recurring pattern

Daily every 1 day

Start date time (UTC)

06/12/2016 07:27 pm

End date time (UTC)

12/30/2099 11:00 pm

Previous

Next

Azure Data Factory (ADF)

1 Properties

Recurring copy

2 Source

○ Connection

● Dataset

3 Destination

Data Lake Store





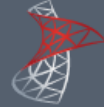









5 Summary

COPY DATA

Specify the source data store for the copy task. You can use an existing data store linked service (or) specify a new data store.

FROM EXISTING LINKED SERVICES

CONNECT TO A DATA STORE

						
Azure Blob Storage	Azure Table Storage	Azure SQL Database	Azure SQL Data Warehouse	On-premises SQL Server	Azure Data Lake	Oracle
						
MySQL	DB2	Sybase	PostgreSQL	Teradata	HDFS	File

Azure Data Factory (ADF)

1 Properties

Recurring copy

2 Source

SQL Server

3 Destination

○ Connection

● Dataset

4 Performance

Polybase, Staging

5 Summary

COPY DATA

Specify the Azure SQL Data Warehouse

Linked service name *(required)*



OutputLinkedService-ouc

Server / database selection method *(required)*



From Azure subscriptions

Server name *(required)*

medivir

Database name *(required)*

medivir_dwh

User name *(required)*


Password *(required)*

Azure Data Factory (ADF)



- 1 Properties
Recurring copy
- 2 Source
SQL Server
- 3 Destination
Azure SQL Data Warehouse
- 4 Performance
Polybase, Staging
- 5 Summary




COPY DATA

Deployment complete



 **On-premises SQL Server**
9 table(s) 
Region: On premise

Copy Run Time Region: On premise

 **Azure SQL Data Warehouse**
9 table(s) 
Region: northeurope

▼ Registering Connections 
▼ Creating Datasets 
▼ Creating Pipelines 

Azure Blob Storage Azure Table Storage Azure SQL Database **Azure SQL Data Warehouse** On-premises SQL Server

Azure Data Lake File

Azure Data Factory (ADF)

medivir-dwh / CopyPipeline-608

Start time (UTC): 03/07/2016 08:50 pm End time (UTC): 06/08/2016 07:50 pm Apply Next scheduled run at 6/8/2016 12:00 AM.

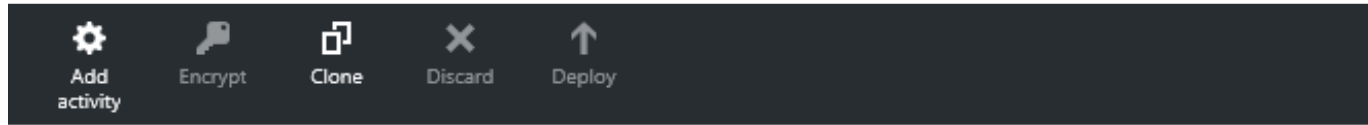
```
graph LR; I1[InputDatasets-608-SzervezetiEgyseg...] --> CA8[CopyActivity-8]; I2[InputDatasets-608-SzervezetiCsoport...] --> CA7[CopyActivity-7]; CA8 --> O1[OutputDatasets-608-SzervezetiEgyseg]; CA7 --> O2[OutputDatasets-608-SzervezetiCsoport];
```

ACTIVITY WINDOWS

1 filter applied. Last refreshed a few seconds ago.

Pipeline	Activity	Window Start	Window End	Status	Type	Attempt Start	Attempt End	Duration	Retry Attempts
CopyPipeline-...	CopyActivity-1	06/06/2016 12:...	06/07/2016 12:...	In progress	Copy	06/07/2016 9:...	--	00:01:00	1
CopyPipeline-...	CopyActivity-4	06/06/2016 12:...	06/07/2016 12:...	In progress	Copy	06/07/2016 9:...	--	00:01:00	1
CopyPipeline-...	CopyActivity-0	06/06/2016 12:...	06/07/2016 12:...	In progress	Copy	06/07/2016 9:...	--	00:01:01	1
CopyPipeline-...	CopyActivity-7	06/06/2016 12:...	06/07/2016 12:...	In progress	Copy	06/07/2016 9:...	--	00:01:00	1

Azure Data Factory (ADF)



```
},  
{  
  "type": "Copy",  
  "typeProperties": {  
    "source": {  
      "type": "SqlSource",  
      "sqlReaderQuery": "select * from [DiagnosztikaKeres_Cloud]"  
    },  
    "sink": {  
      "type": "SqlDWSink",  
      "allowPolyBase": true,  
      "polyBaseSettings": {},  
      "writeBatchSize": 0,  
      "writeBatchTimeout": "00:00:00"  
    }  
  },  
  "inputs": [  
    {  
      "name": "InputDatasets-2e7-DiagnosztikaKeres_Cloud"  
    }  
  ],  
  "outputs": [  
    {  
      "name": "OutputDatasets-2e7-DiagnosztikaKeres"  
    }  
  ],  
}
```

MSSQL – Integration Services (push)

The image displays three overlapping windows from Microsoft Visual Studio, illustrating the configuration and execution of an SSIS Data Flow Task.

- Left Window (Design View):** Shows the SSIS Data Flow Task editor for 'Package1.dtsx'. It features an 'ADO NET Source' connected to an 'ADO NET Destination'.
- Center Window (ADO.NET Destination Editor):** A dialog box titled 'ADO.NET Destination Editor' with the instruction: 'Configure the properties used to insert data into a destination using ADO.NET provider.' It includes:
 - Connection Manager:** A dropdown menu set to 'medivir.database.windows.net.medivir_dwh.SQLAdmin 1'.
 - Use a table or view:** A dropdown menu set to 'dbo.Calendar'.
 - Preview:** A button to preview the data.
 - Use Bulk Insert when possible:** A checked checkbox.
 - Buttons:** 'OK', 'Cancel', and 'Help' at the bottom.
- Right Window (Running View):** Shows the SSIS Data Flow Task editor for 'Package.dtsx' in a running state. The 'ADO NET Source' and 'ADO NET Destination' are both marked with green checkmarks. A data flow arrow between them is labeled '4 744 973 rows'.

Azure SQL Data Warehouse (ADW)

- Elszámolási egysége DWU
- Tervezési megkötések
 - Adattípusok
 - DISTRIBUTION
 - HASH
 - ROUND ROBIN
 - PARTITION

```
1 CREATE TABLE [dbo].[DiagnosztikaKeres]  
2 (  
3     [DiagnosztikaKeresID] [int] NOT NULL,  
4     [DiagnosztikaKeroSzervezetID] [int] NOT NULL,  
5     [KeroSzervezetiEgysegID] [int] NOT NULL,  
6     [VegzoSzervezetiEgysegID] [int] NOT NULL,  
7     [TeritesiKategoriaID] [int] NOT NULL,  
8     [KeroOrvosID] [int] NOT NULL,  
9     [PaciensID] [int] NOT NULL,  
10    [IranditoszamID] [int] NOT NULL,  
11    [KezdesiIdo] [datetime] NOT NULL,  
12    [WHOPontszam] [int] NOT NULL,  
13    [NaploSorszam] [nvarchar](20) NOT NULL,  
14    [TimestampDatum] [datetime] NOT NULL,  
15    [TimestampHonap] [datetime] NOT NULL,  
16    [Letrehozva] [datetime] NOT NULL,  
17    [BeavatkozasKodID] [int] NOT NULL,  
18    [VegzoOrvosID] [int] NOT NULL,  
19    [BeavDB] [int] NOT NULL,  
20    [CalendarID] int NOT NULL  
21 )  
22 WITH  
23 (  
24     CLUSTERED COLUMNSTORE INDEX,  
25     DISTRIBUTION = HASH([DiagnosztikaKeresID])  
26 );
```

Azure SQL Data Warehouse (ADW)

- Tervezési megkötések

- Ajánlott rekord és tranzakció szintű globális azonosítók használata
- CTE

Common table expressions followed by INSERT, UPDATE, DELETE or MERGE are not supported in this version.

- Default adatbázis collation **SQL_Latin1_General_CP1_CI_AS**

'Hungarian_CI_AS' is not a supported collation.

- DEFAULT VALUE

An expression cannot be used with a default constraint. Specify only constants for a default constraint.

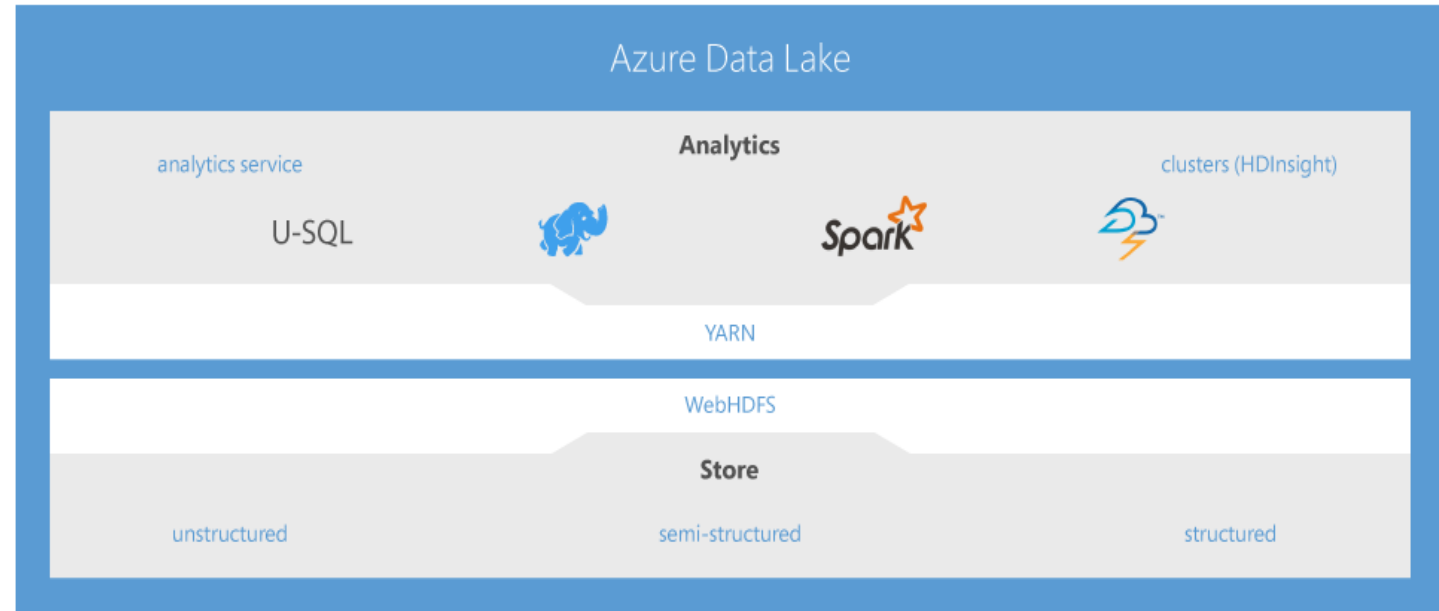
Azure SQL Data Warehouse (ADW)

- Nem támogatott funkciók/innovatív újítások:
 - Identities (workaround)
 - **Primary keys**
 - **Foreign keys**
 - Check constraints
 - Unique constraints
 - Unique indexes
 - Computed columns
 - Sparse columns
 - **User-defined types**
 - Indexed views
 - Sequences
 - Triggers
 - Synonyms

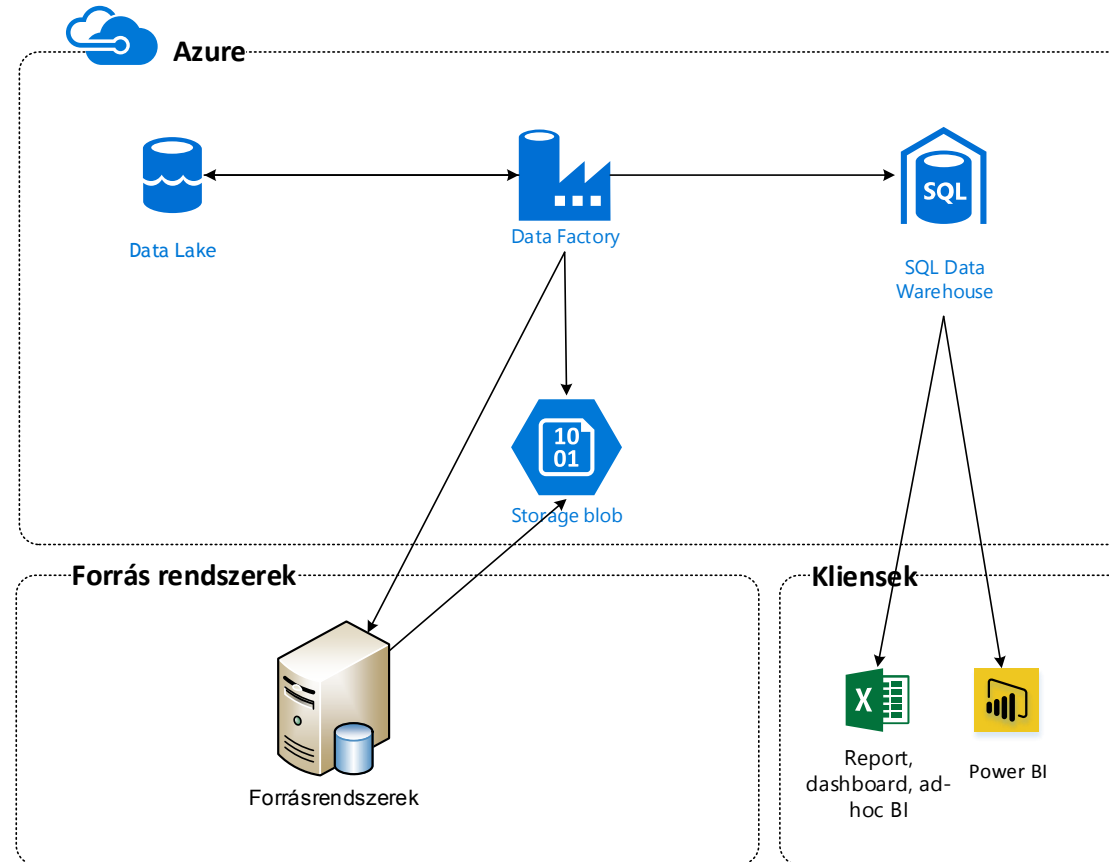


Azure Data Lake (ADL)

- Storage
 - WebHDFS
 - Tárolásért nem kell fizetni
- Analytics
 - U-SQL
 - Job futtatásért kell fizetni

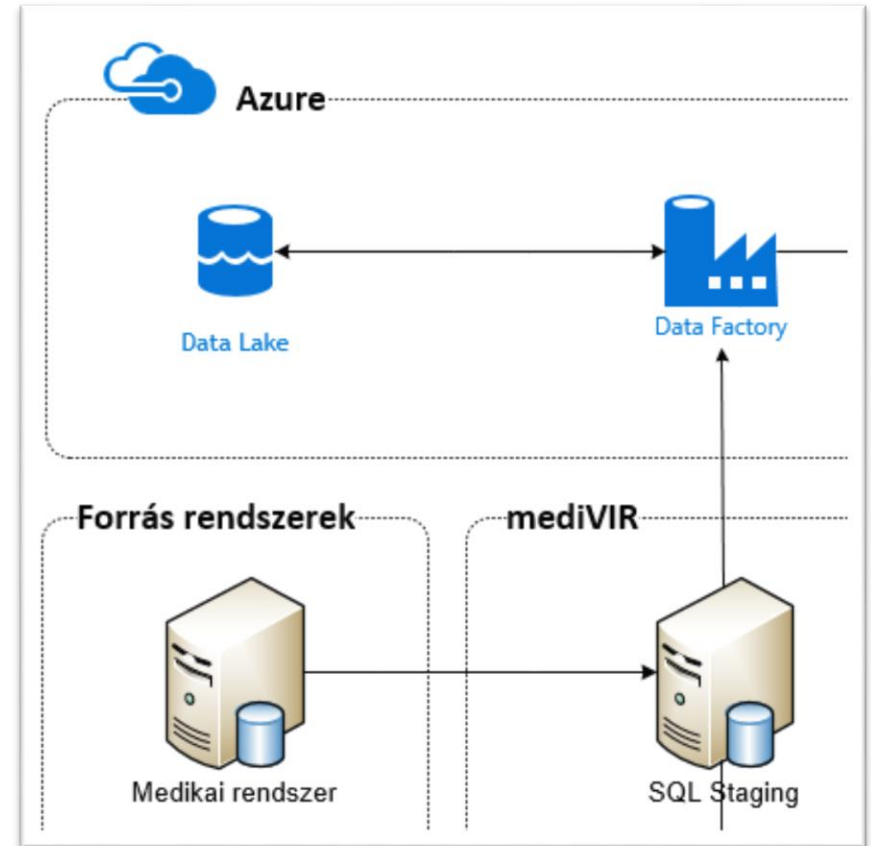


Azure Data Lake (ADL)



Azure Data Lake (ADL)

- Staging funkciók kiváltása
- Csoportosítások, aggregációk számítása



Azure Data Lake (ADL)

- Adatbázis létrehozása

```
Submit | medivir | master | dbo |
1 CREATE DATABASE mediVIR_Staging;
2
3 USE DATABASE mediVIR_Staging;
4
5 CREATE TABLE DiagnosztikaKeres (
6     DiagnosztikaKeresID int,
7     DiagnosztikaKeroSzervezetID int,
8     KeroSzervezetiEgysegID int,
9     VegzoSzervezetiEgysegID int,
10    TeritesiKategoriaID int,
11    KeroOrvosID int,
12    PaciensID int,
13    IranyitoszamID int,
14    KezdesiIdo DateTime,
15    WHOPontszam int,
16    NaploSorszam string,
17    TimestampDatum DateTime,
18    TimestampHonap DateTime,
19    Letrehozva DateTime,
20    BeavatkozasKodID int,
21    VegzoOrvosID int,
22    BeavDB int,
23    CalendarID int,
24
25    INDEX IX_DiagnosztikaKeresID
26        CLUSTERED (DiagnosztikaKeresID ASC)
27        PARTITIONED BY HASH (DiagnosztikaKeresID)
28 );
```

- Aggregáció képzése

```
Submit | medivir | master | dbo |
29
30 @aggregation =
31 SELECT
32     KeroSzervezetiEgysegID,
33     VegzoSzervezetiEgysegID,
34     TeritesiKategoriaID,
35     TimestampHonap,
36     SUM(WHOPontszam) AS TotalWHOPontszam,
37     SUM(BeavDB) AS TotalBeavDB
38 FROM
39     mediVIR_Staging.dbo.DiagnosztikaKeres
40 GROUP BY
41     KeroSzervezetiEgysegID,
42     VegzoSzervezetiEgysegID,
43     TeritesiKategoriaID,
44     TimestampHonap;
45
46 OUTPUT @output
47 TO "mediVIR/Aggregations/DiagnosztikaKeres.csv"
48 ORDER BY TimestampHonap DESC
49 USING Outputters.Csv();
```

Power BI

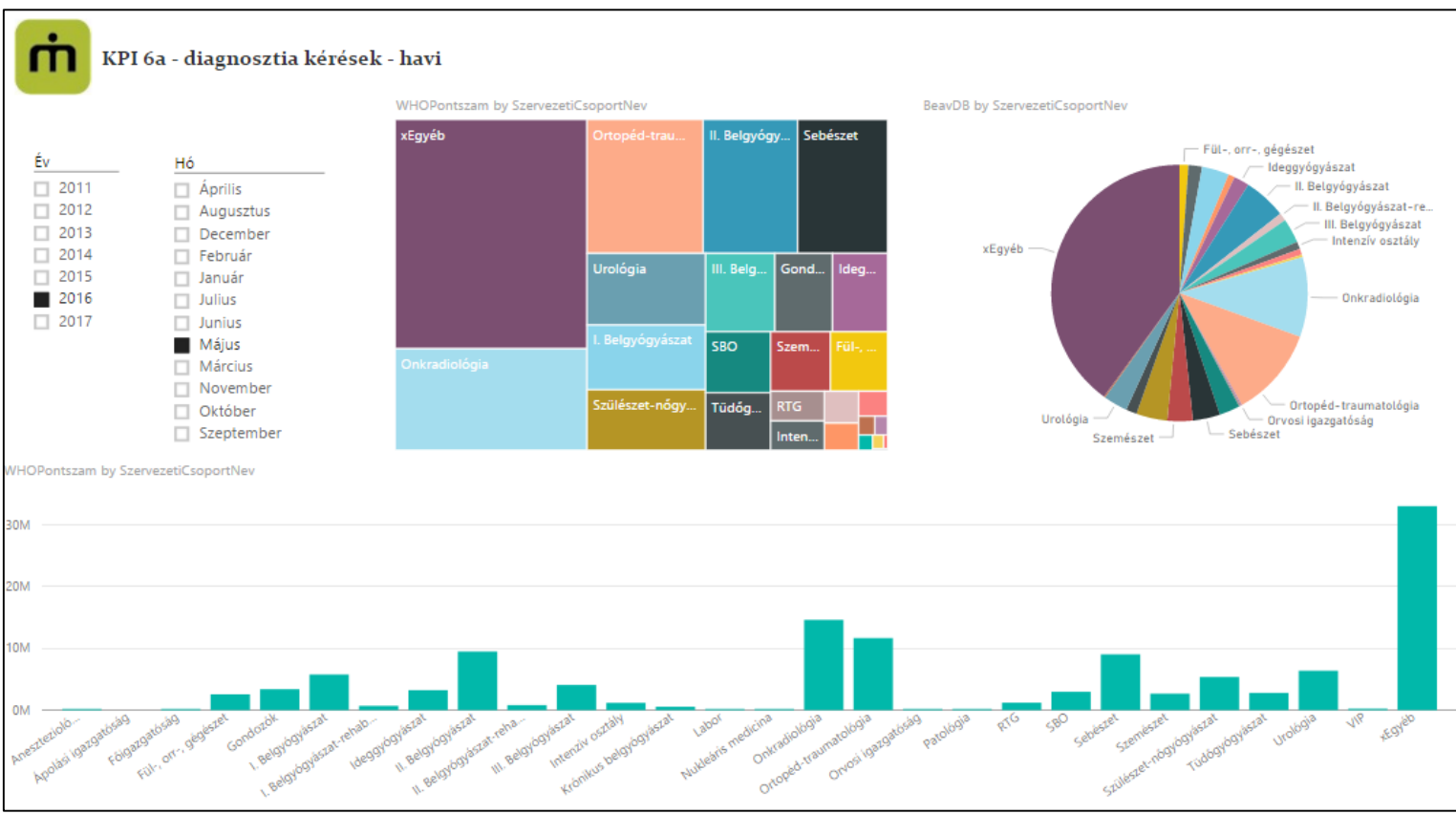
Felhő alapú megoldás

Webes megjelenítés

Desktop önkiszolgáló BI eszköz

Mobil kliensek (iOS, Android, Win10)

Természetes nyelvű Q&A



Power BI

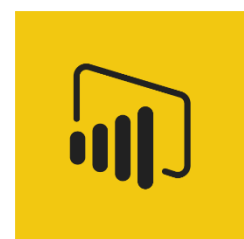
Felhő alapú megoldás

Webes megjelenítés

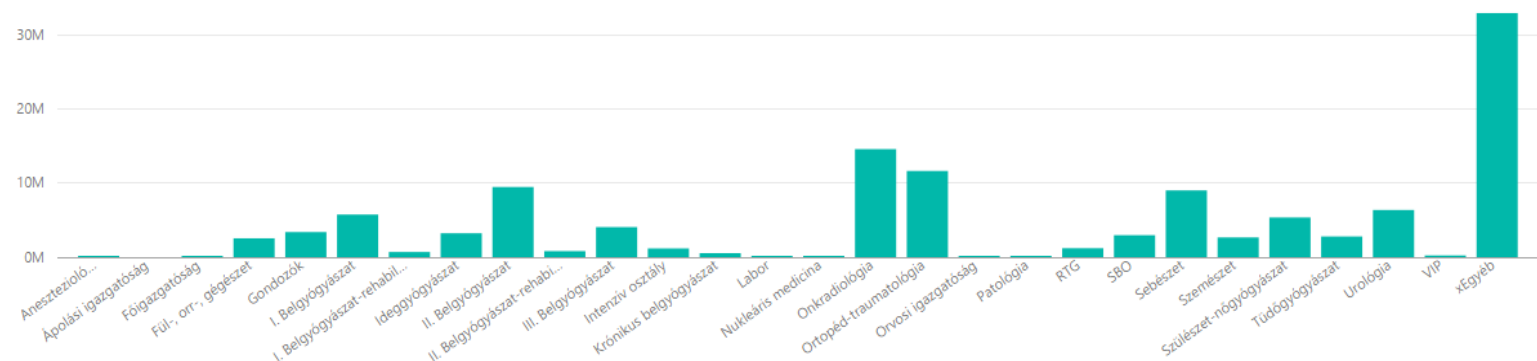
Desktop önkiszolgáló BI eszköz

Mobil kliensek (iOS, Android, Win10)

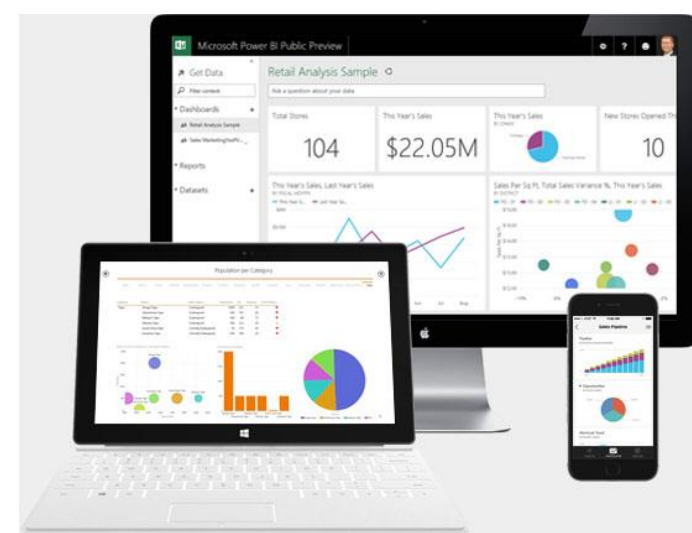
Természetes nyelvű Q&A



< Back to Report | WHOPONTSZAM BY SZERVEZETICSOPORTNEV



SzervezetiCsoportNev	WHOPontszám
Aneszteziológia	159088
Ápolási igazgatóság	0
Főigazgatóság	37790
Fül-, orr-, gégeész	2547923
Gondozók	3391782
I. Belgyógyászat	5761282
I. Belgyógyászat-rehabilitáció	693362
Ideggyógyászat	3231947
II. Belgyógyászat	9487470
II. Belgyógyászat-rehabilitáció	817213
III. Belgyógyászat	4067029
Intenzív osztály	1176657
Krónikus belgyógyászat	540939
Labor	115855
Nukleáris medicina	15760
Onkradiológia	14603182
Ortopéd-traumatológia	11660192



Esettanulmány - eredmények

- Migráció 3GB adattal

- 1 Mbit/sec

Forrás	Cél	Eszköz	Idő
Helyi SQL szerver	ADW - North EU	ADF	2:53:00
Helyi SQL szerver	ADW - North EU	SSIS	2:59:00
Helyi SQL szerver	ADL - East US	ADF	4:12:00
ADW - North EU	ADL - East US	ADF	0:11:00

- Napi betöltés

- Helyi és ADW műveletek párhuzamosíthatóak
- ADW tetszőleges időközönként futtatható

- Funkciók

- Közel valós idejű analízis
- Megszakítások nélkül

Művelet (perc)	Régi	Új	
	Helyi	Helyi	ADW (15)
ETL	100	9	3
MOLAP process	12	8	0
Összesen	112	17	3





Abronits Péter

Technikai projektvezető

abronits.peter@grape.hu



Referenciák

1. <https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-develop-table-design/>
2. <https://www.microsoft.com/en-us/download/details.aspx?id=39717>
3. <https://blogs.msdn.microsoft.com/sqlcat/2015/08/11/choosing-hash-distributed-table-vs-round-robin-distributed-table-in-azure-sql-dw-service/>
4. <https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-develop-table-partitions/>
5. <https://msdn.microsoft.com/en-us/library/azure/mt591959.aspx>