

# Analyse og innsikt for alle

---

Azure Synapse Analytics og Power BI

09.11.2022

# Hva vi skal se på

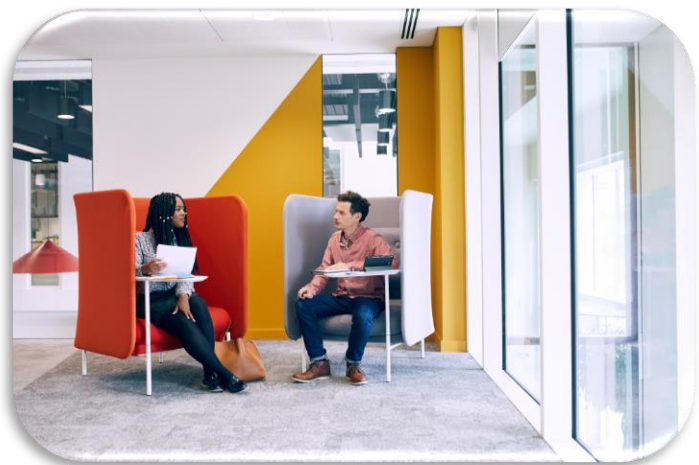
Azure Synapse Analytics  
Power BI tjenester

Komponenter og dere funksjoner og bruksområder  
Hvordan tjenestene kan benyttes i både små og store bedrifter

Kostnader



# Synapse og Power BI følger deg i vekst



Startup /  
Liten bedrift



Mellomstor  
bedrift



Stort konsern

# 01

## Startups / Liten bedrift

Og hvordan Synapse og Power BI passer inn fra starten

## Data kilder



eksterne sky tjenester



Excel



Database

## Azure Synapse Analytics

### Admin



Synapse Studio



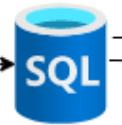
Synapse Monitor

### Data bevegelse

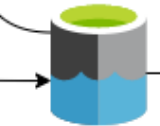


Synapse Pipelines

### Serverless SQL Pool



SQL



Data Lake Gen 2

## Data konsumering



Excel



Power BI Desktop

## Azure Support services



Azure AD



Key Vault

# Karakteristikk Liten Bedrift

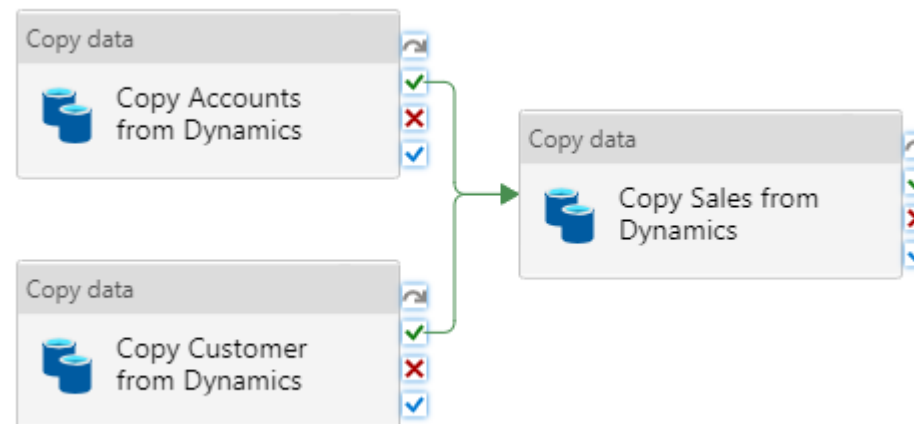
- Få kilder og begrenset data
- Samle rapportering (One truth)
- Data tilgjengelig for alle
- Små IT budsjetter
- Få IT ressurser

# Synapse Pipelines

- Orkestrering
- Kobling mot mye eksterne datakilder
- Kan starte flere ulike compute prosesser
- No-code verktøy

**\$1 pr x1000 aktivitet kjøres**

Jobb kjører hver 30min:  
(3 x 48 x 30 dager) / 1000 = **\$4.32**



**\$0.25 pr DIU**

Jobb kjører i 3 min (0.05 DIU) hver 30min:  
(0.05 x \$0.25 x 48 x 30 dager) = **\$18**

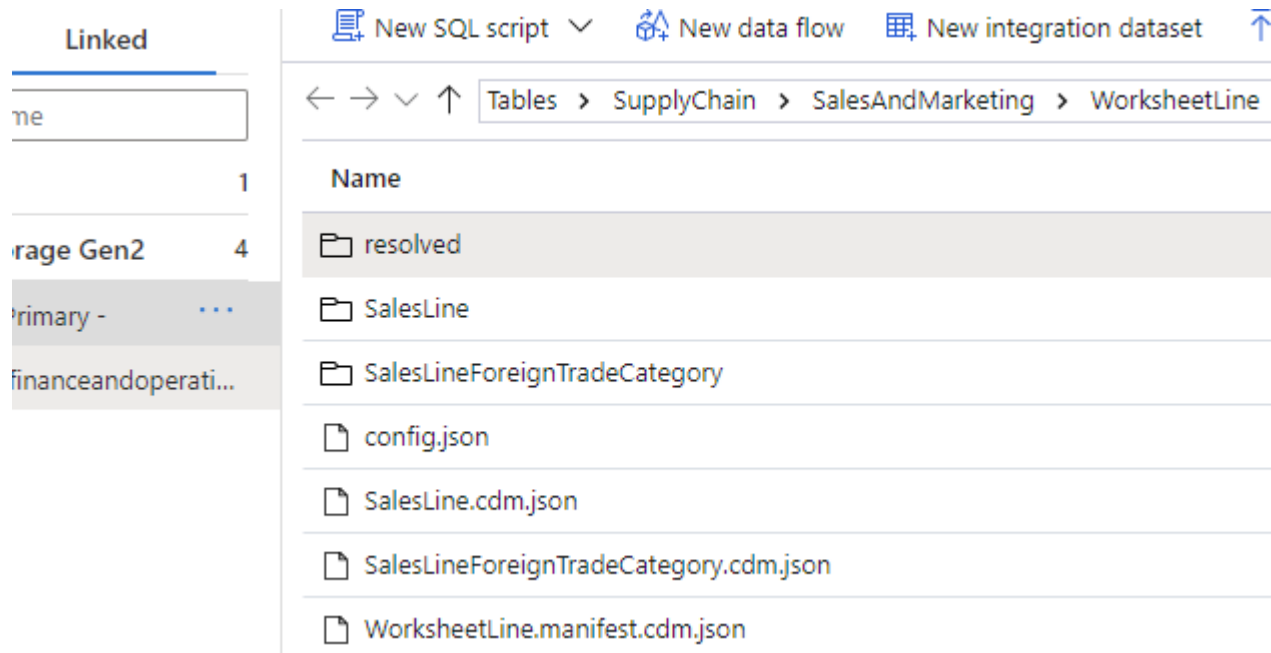
0,0167 DIU	00:01:00
0,05 DIU	00:03:00
0,05 DIU	00:03:00

00:00:41	57173528-45bc-...	✓ Succeeded
00:02:32	ExecutePipelines...	✓ Succeeded
00:02:44	ExecutePipelines...	✓ Succeeded

DIU = Data Integration Units = CPU, Minne og nettverk pr time

# Azure Data Lake

- Hierarkisk lagring av data
- Ustrukturert -> strukturert
- God ytelse på les og skriv
- Tilgangsstyring på objekt nivå
- Parquet fil format



**\$22,53**  
pr TB data

Område:	Type:	Nivå:	Storage-kontotype:
North Europe	Data Lake Storage Gen2	Standard	Generell bruk V2
Tilgangsnivå:	Redundans: ⓘ	Filstruktur:	
Varm	LRS	Hierarkisk navneområde	
Kapasitet			
1 TB			

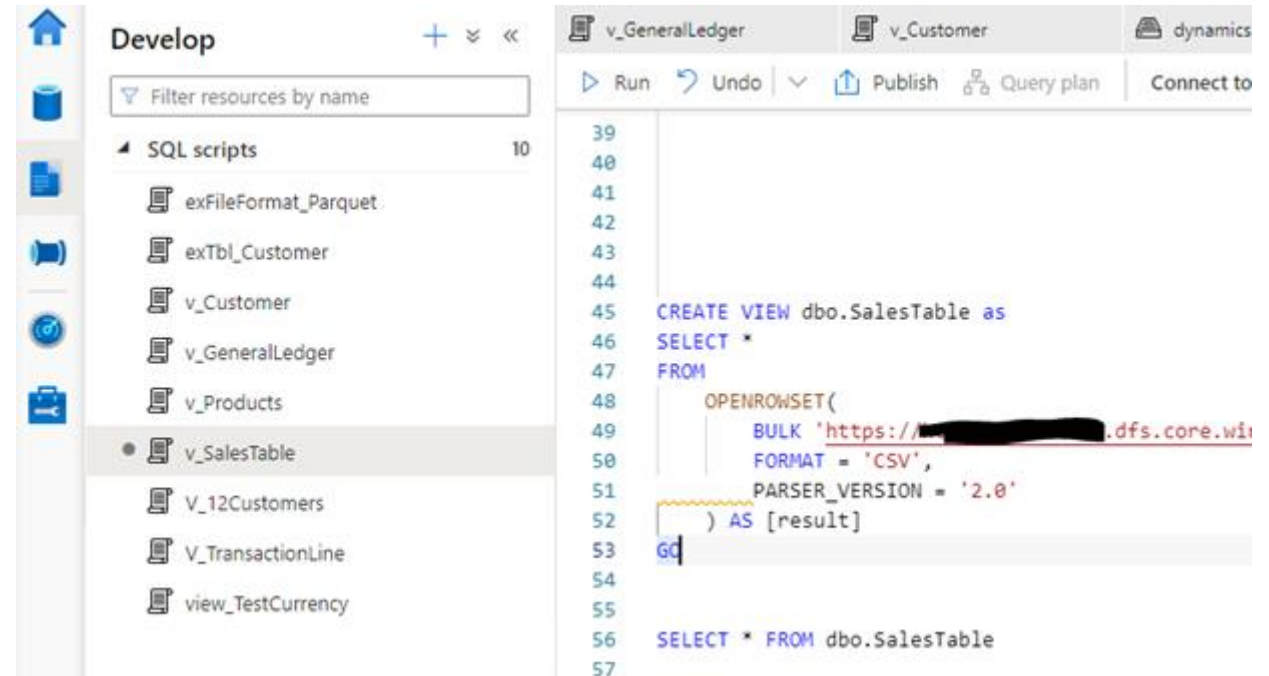
**= USD 22,53**  
Gjennomsnitt per måned  
(USD 0,00 forhåndsbetalt)

# Synapse Serverless SQL Pool

- On-demand SQL prosessering
- Ad-Hoc SQL analytics
- Ingen fysisk lagring av data

**\$5 /TB**  
**data prosessert**

Mulig med kost kontroll



```
39
40
41
42
43
44
45 CREATE VIEW dbo.SalesTable as
46 SELECT *
47 FROM
48     OPENROWSET(
49         BULK 'https://[redacted].dfs.core.wi
50         FORMAT = 'CSV',
51         PARSER_VERSION = '2.0'
52     ) AS [result]
53 GO
54
55
56 SELECT * FROM dbo.SalesTable
57
```

Prosessert data = data midlertidig lagret med en spørring kjører

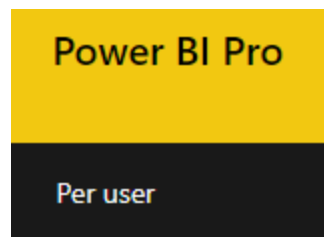
- Data lest fra lager (data og metadata)
- Data mellomlagret under prosessering
- Data som skrives til lager



# Power BI Desktop

- Windows app for utvikling av visuell data og innsikt

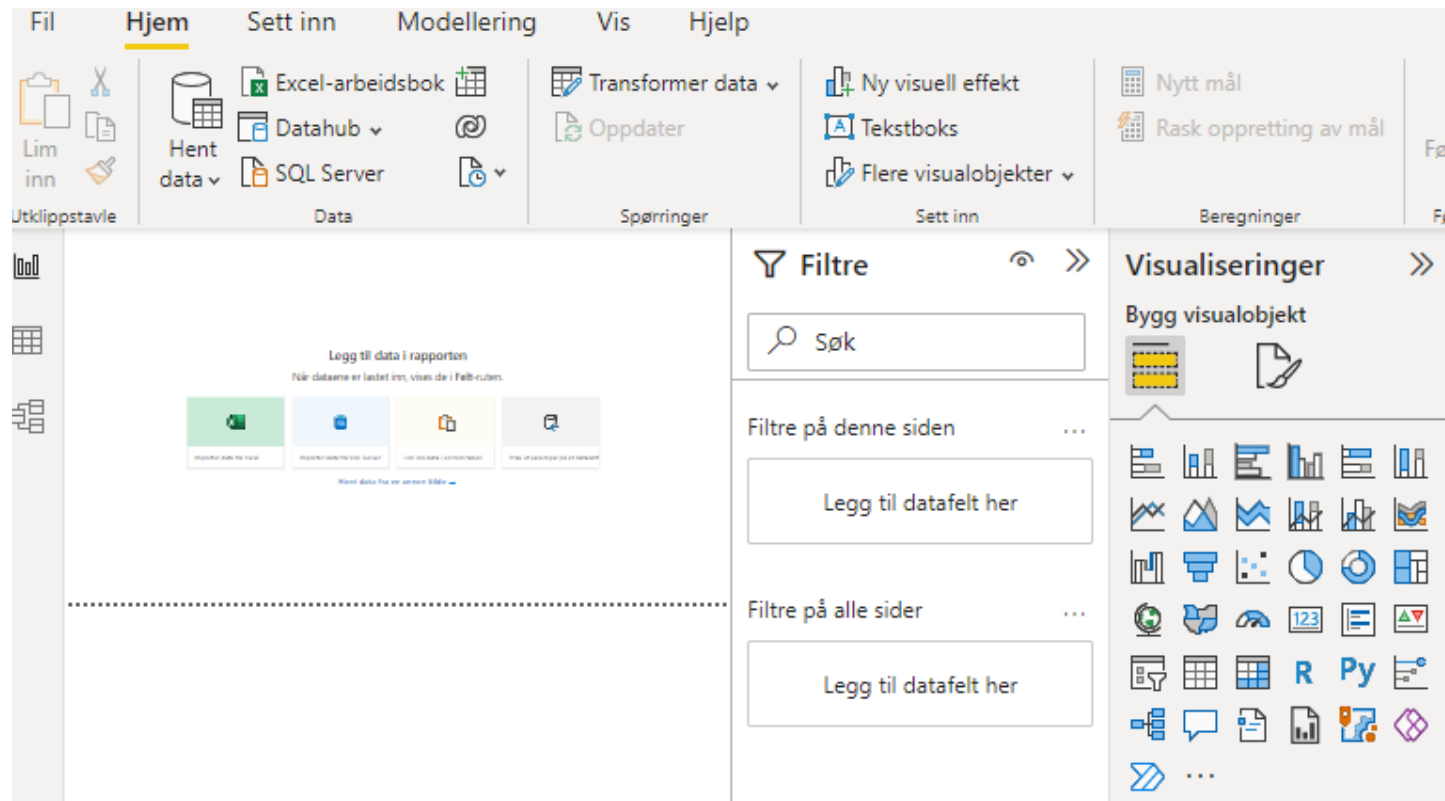
- Desktop er gratis
- Deling koster



Alle som skal se delte rapporter må ha PRO lisens

**\$9.99**

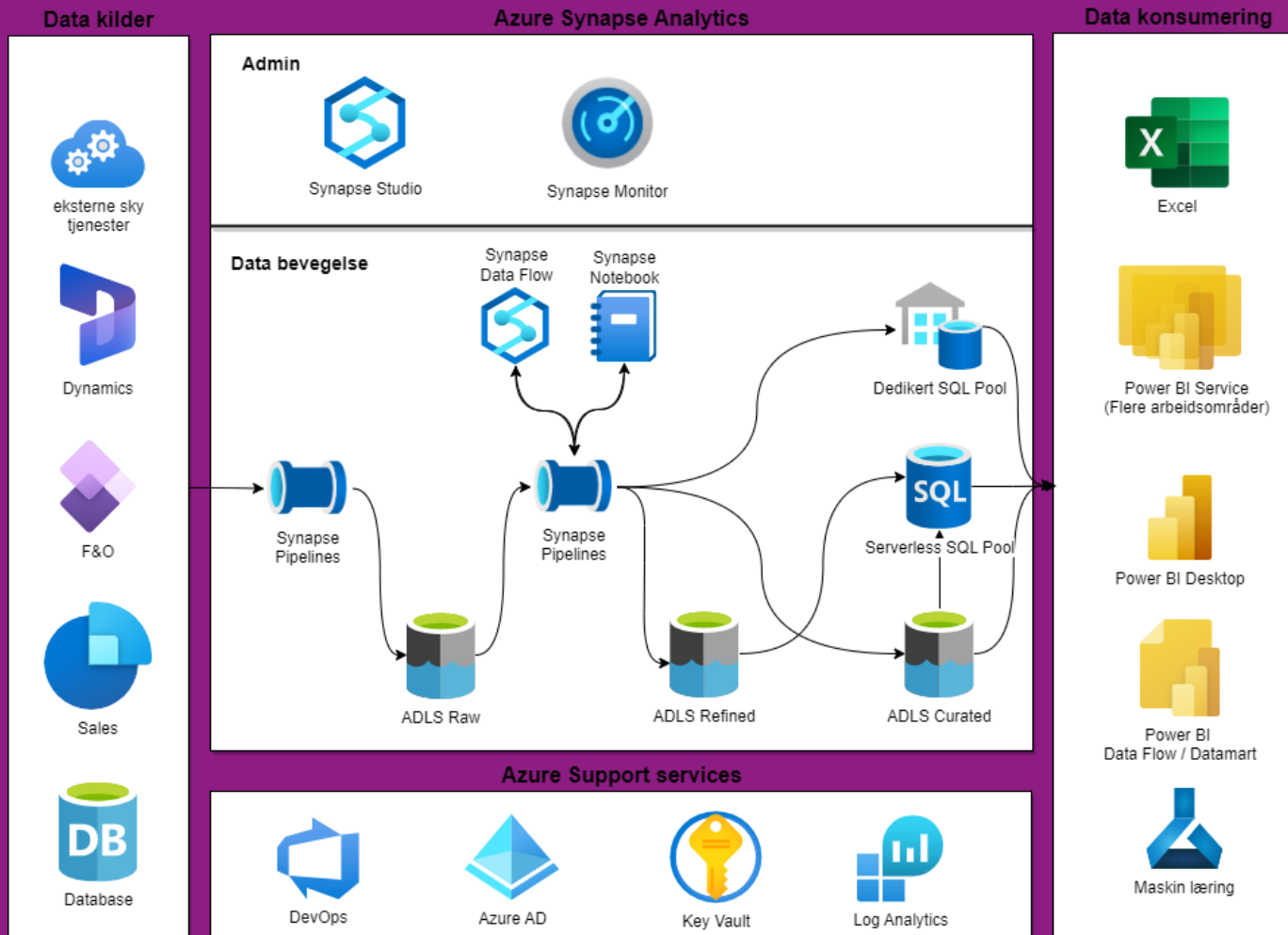
Per user/month



# 02

## Mellomstor bedrift

Og hvordan Synapse og Power BI kan vokse med deg



# Karakteristikk Mellomstor bedrift

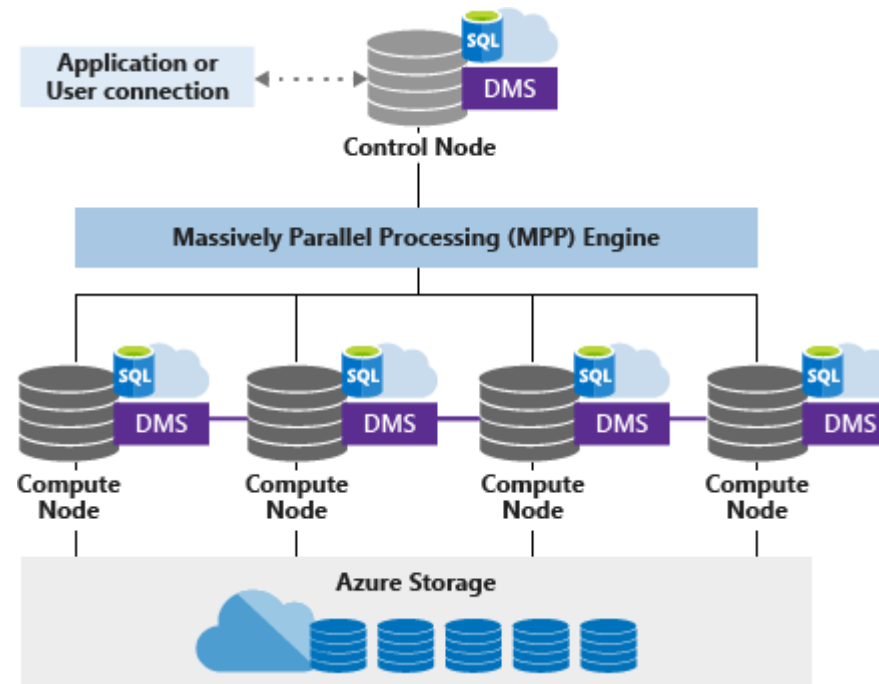
- Flere forretningsområder og data kilder
- Voksende datamengde
- Behov for roller og tilgangsstyring
- Ulike behov for data og innsikt
- Egen dedikert IT organisasjon

# Dedicated SQL Pool

- SQL basert datavarehus /database
- Scale-out arkitektur
- Kan skaleres raskt og enkelt
- Tilgangsstyring rad og kolonner
- Maskering- og kryptering av felter m.m. (GDPR)

NB! Faktura skalerer med ressurser

DWU = Data Warehouse Units = CPU, Minne og IO



DWU blocks:

100	730	Hours	×	\$1.510	Per hour
100					
200					
300					
400					
500					
1000					
1500					
2000					
2500					
3000					
5000					
6000					
7500					
10000	10		×	\$23.00	Per TB
15000					
30000					

undant Disaster Recovery

### Savings Options ⓘ

- Pay as you go = \$1,102.30
- 1 year reserved (~37% savings) Average per month (\$0.00 charged upfront)
- 3 year reserved (~65% savings)

# Synapse Notebook og Data Flow

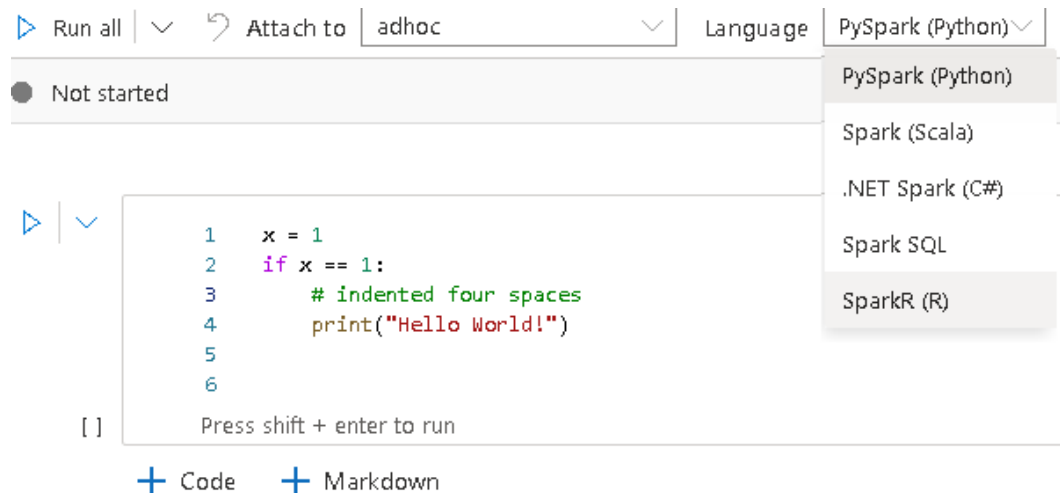
## Synapse Notebook

- Verktøy for avansert transformasjon med kode
- Flere kode språk

## Synapse Data Flows

- No-code verktøy for visuell transformasjon av data
- Tilknyttet Synapse pipelines

Begge bygger på spark teknologi, men kan ikke dele Spark ressurser



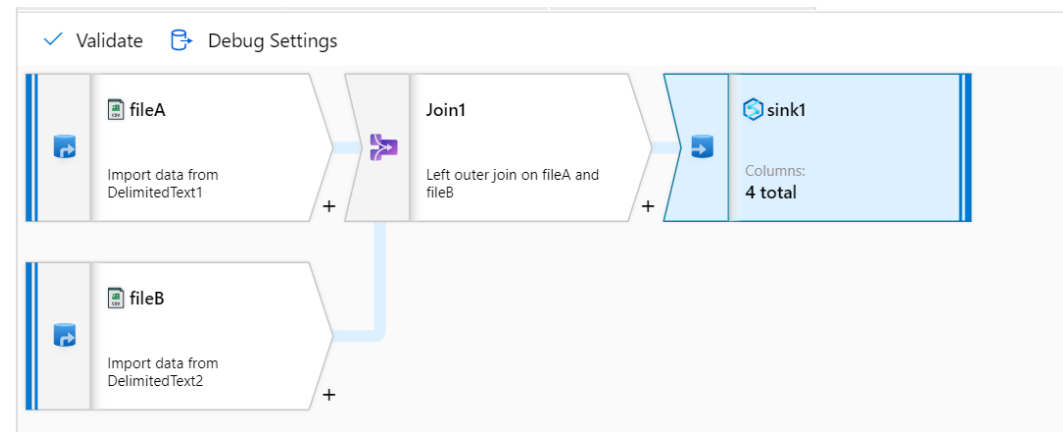
Run all | Attach to | `adhoc` | Language | PySpark (Python)

● Not started

```
1 x = 1
2 if x == 1:
3     # indented four spaces
4     print("Hello World!")
5
6
```

[ ] Press shift + enter to run

+ Code + Markdown



# Kostnader - Synapse Notebooks og Data Flow

## Notebooks

- Minimum 4 og opp til 64 kjerner
- Spesifisert minne 32 til 448GB (XLarge)
- Koster pr hele påbegynte time

## Scenario

- 1 instans og 8 kjerner
- 3 timer transformasjon hver morgen

## Data Flows

- Minimum 8 og opp til 68 kjerner
- Kost pr hele påbegynte time

**\$93 VS \$193**

## Apache Spark Pool

### Memory optimized

Node size:

Medium (8 vCores / 64 GB) ▾

1 × 90 Hours ▾ = \$92.74  
Instances

## Data Flows

▴ Data Flows

### Basic vCores

Instance:

8 vCores ▾

1 × 90 Hours ▾ = \$192.77  
Instances

# Power BI Service

- Løsningen for å dele Power BI rapporter med andre
- Mulig med tilgang til eget arbeidsområde uten lisens
- Tilgangsstyring på arbeidsområdet
- Konfigurering av bl.a.
  - Gateway mot on-prem
  - Power BI Apps
  - Row Level Security
- PRO eller Premium lisens

Mitt arbeidsområde

+ Ny ▾

☰ Vis ▾    🔍 Filtre    🔍 Søk

Alle   Innhold   Datasett + dataflyter

	Navn	Type	Eier	Oppdatert	Neste oppdatering
	PBI_Test	Rapport	Jon Terje TJØNN	8.11.22, 19:50:24	—
	PBI_Test	Datasett	Jon Terje TJØNN	8.11.22, 19:50:24	ikke aktuelt

Oppdater a...

☰ Vis ▾    🔍 Filtre    ⚙ Innstillinger    👤 Tilgang    ⋮    🔍 Søk

Nytt datamart    ✕

# Power BI Dataflow / Datamart

## Dataflow

- Selvbetjent opprettelse av gjenbrukbare transformasjoner
- Kilde for datasett eller andre Dataflows

Data sources

Common Data Service Power Platform	Excel File	Folder File	JSON File	PDF File	SharePoint folder File	Text/CSV File
XML File	Access Database	Amazon Redshift Database	Google BigQuery Database	IBM Db2 database Database	Impala Database	MySQL database Database
Oracle database Database	PostgreSQL database Database	Power BI dataflows Power Platform	Power Platform dataflows Power Platform	SAP BW Application Server Database	SAP BW Message Server Database	SAP HANA database Database
SQL Server database Database	Sybase database Database	Teradata database Database	Vertica Database	Azure Blobs Azure	Azure Data Explorer (Kusto) Azure	Azure Data Lake Storage Gen2 Azure
Azure HDInsight Spark Azure	Azure SQL database Azure	Azure Synapse Analytics (SQL DW) Azure	Azure Tables Azure	Microsoft Exchange Online Online services	Salesforce objects Online services	Salesforce reports Online services
SharePoint Online list Online services	Smartsheet Online services	Active Directory Other	OData Other	Odbc Other	SharePoint list Other	Spark Other
Web API Other	Web page Other	Blank table Other	Blank query Other			

## Datamart

- Selvbetjent løsning for modellering av data på web
- Opptil 100GB på SQL Database
- Mao. tilgang med SQL
- Kan hente data fra Dataflows
- Krever Premium lisens

Power BI Premium	
Per user	Per capacity
\$20	from \$4,995
Per user/month <sup>2</sup>	Per capacity/month

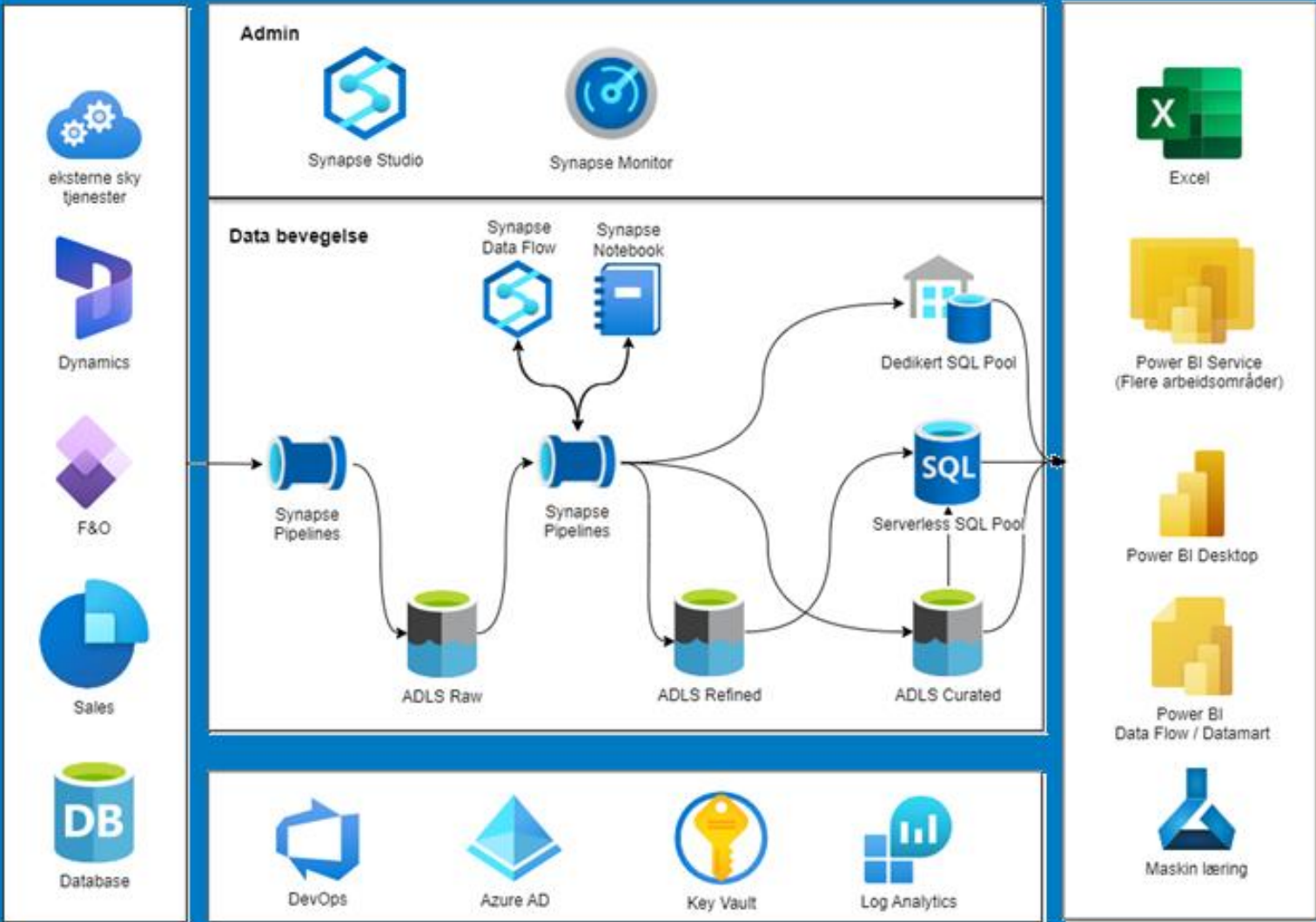
- Tenkt brukt til selvbetjening for å «slippe å vente på IT»
- **Krever teknisk kompetanse**



# 03

## Stor bedrift / Konsern

Og hvordan Synapse og Power BI kan utnyttes fullt ut



# Karakteristikk Konsern

- Organisasjoner med høy kompleksitet
- Svært ulike data behov i ulike forretningsområder
- Egen IT hos hvert forretningsområdet (FR)
- Samme data på ulikt vis
- Fokus på styring og kontroll av IT i FR og på tvers i konsern

**Spørsmål?**

# Støtte

- [Azure pris kalkulator](#)
- [PBI Dataflow Best Practices](#)



# Takk for meg!

---

Jon Terje Tjønn

[jon.tjonn@soprasteria.com](mailto:jon.tjonn@soprasteria.com)