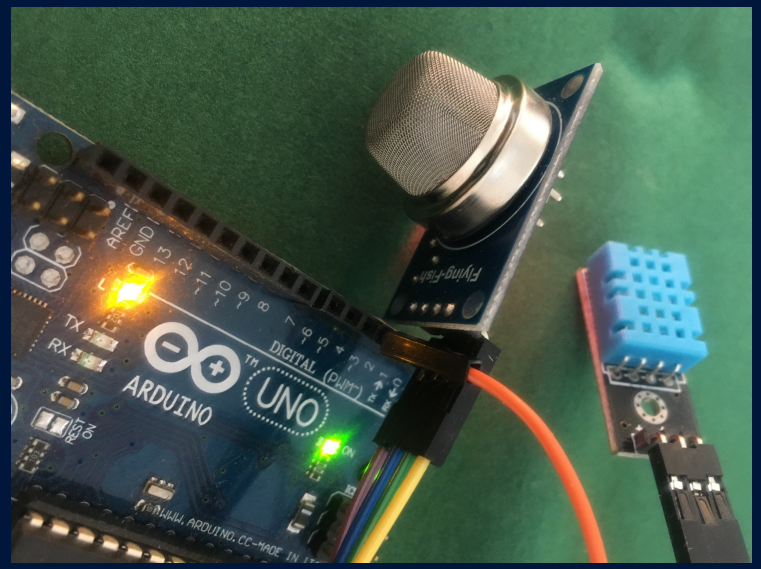
cphbusiness

COPENHAGEN BUSINESS ACADEMY

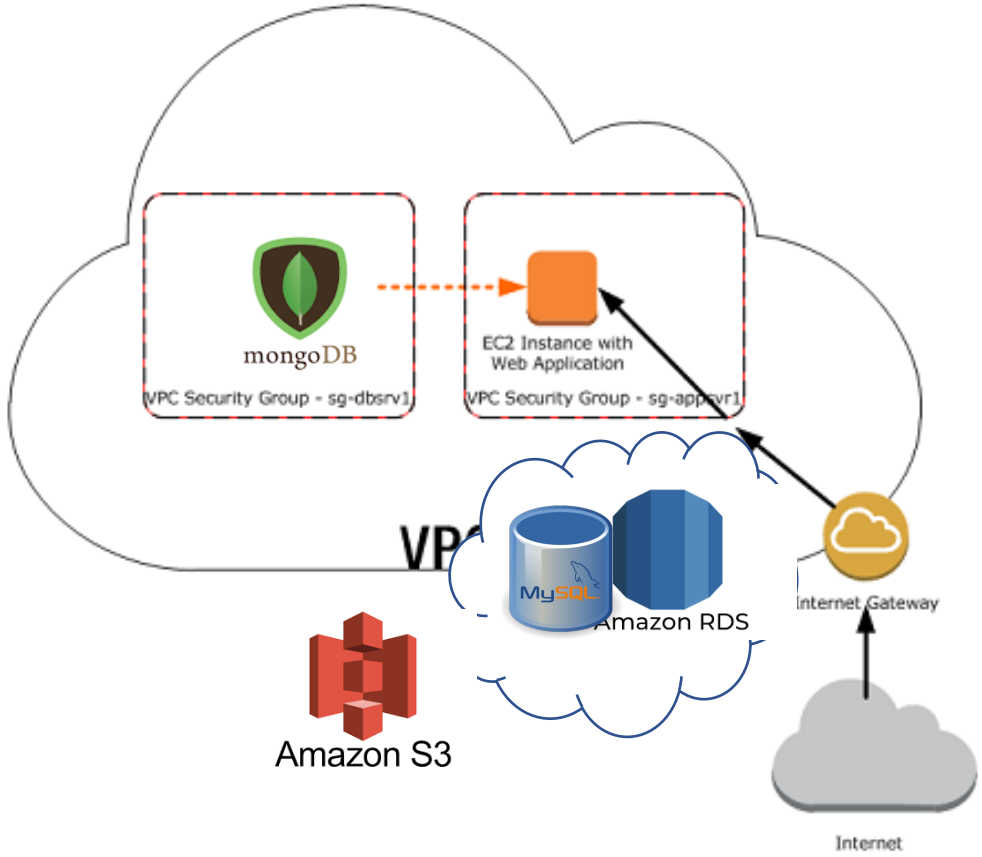
DATA ENGINEERING



Agenda – I gang med AWS

- AWS status og intro
  - services og instances
- Windows server
  - intro til OS samt server roller & services
  - powershell
    - øvelser
- Linux server
  - intro til OS samt server roller & services
  - terminalen
    - øvelser
- SQL og R på Windows / Linux
  - scateboard med csv

Bilbasen – slutmålet



- **MongoDB**
  - Create AWS EC2 instance
  - Get rdp-access
  - Client access
    - mongosh
  - Dev access
    - mongolite
- **SQL**
  - Create AWS RDS
  - Client access
    - MySQL-Workbench
  - Dev access
    - RMariaDB
- **Files**
  - Create AWS S3 bucket
  - Dev access
    - aws.s3

Linux vs Windows

~ødevppratik

Bash Command	Powershell Equivalent (alias)	Description
echo	Write-host (echo)	Outputs the strings that is passed as arguments
pwd	Get-Location (pwd)	Prints the current working directory
ls	Get-ChildItem (ls)	Lists the names of files in a particular directory
cd	Set-Location (chdir)	Change the current working directory to the argument passed
cp	Copy-Item	Copies the file or directory
rm	Remove-Item (rm)	Remove objects such as files, directories from file systems
touch	New-Item	Create, change and modify timestamps of a file
cat	Get-Content	Reads files sequentially and write them to standard output.
man	Get-Help	Displays the user manual of any command
cat	Get-Content	Reads files sequentially and write them to standard output.
man	Get-Help	Displays the user manual of any command

mkdir	Laver en folder

Linux

Bash Command	Description
echo	Outputs the strings that is passed as arguments
pwd	Prints the current working directory
ls	Lists the names of files in a particular directory
cd	Change the current working directory to the argument passed
cp	Copies the file or directory
rm	Remove objects such as files, directories from file systems
touch	Create, change and modify timestamps of a file
cat	Reads files sequentially and write them to standard output.
man	Displays the user manual of any command

~@devppratik

Lav **filer med indhold** med cat-kommandoen og redirection (slut med <ENTER> og derpå <ctrl+d>).

```
Administrator@DESKTOP-7QA27BH MINGW64 ~/test/filer/store
$ cat > test2.sh
#!/bin/bash
echo "hello world"
```

Øvelser

- Opret foldere DAL og tmp
- Cd til tmp og hent flg vha wget (eller curl)
  - https://raw.githubusercontent.com/cphstud/20m9596V2-uge5/master/cars.csv
    https://raw.githubusercontent.com/cphstud/20m9596V2-uge5/master/test2.sh
- Hvor stor er filen test2.sh?
- Hvordan vil du eksekvere scriptet?
- Modificer scriptet så det udskriver dit navn
- Lav en folder i tmp der hedder test og få vha chatgpt et script der kan lave 20 tomme filer i test-folderen
- Lokaliser mongodb's logfil og find den sidst tilføjede linje
- Hent cmd.sh fra https://github.com/cphstud/20m9596V2-uge5, læg det i tmp-folderen og undersøg hvordan det skal eksekveres

Windows

Powershell Equivalent (alias)	Description
Write-host (echo)	Outputs the strings that is passed as arguments
Get-Location (pwd)	Prints the current working directory
Get-ChildItem (ls)	Lists the names of files in a particular directory
Set-Location (chdir)	Change the current working directory to the argument passed
Copy-Item	Copies the file or directory
Remove-Item (rm)	Remove objects such as files, directories from file systems
New-Item	Create, change and modify timestamps of a file
Get-Content	Reads files sequentially and write them to standard output.
Get-Help	Displays the user manual of any command

~ødevppratik

Øvelser

- Opret foldere DAL og temp?
- Cd til temp og hent flg vha wget
  - https://raw.githubusercontent.com/cphstud/20m9596V2-uge5/master/cars.csv
    https://raw.githubusercontent.com/cphstud/20m9596V2-uge5/master/test.ps1
- Hvor stor er filen test.ps1?
- Hvordan vil du exekvere scriptet?
- Modificer scriptet så det udskriver dit navn
- Lav en foler i temp der hedder test og få vha chatgpt et script der kan lave 20 tomme filer i test-folderen
- Installér notepad++ og sørg for at notepad++ er føjet til stien i dit environment
- Lokaliser mongodb's logfil og find den sidst tilføjede linje

SQL og R

- Hello World med R
- Install packages
- command-lines arguments
- Read csv via commandline
- Init car-script
  - MySQL from commandline
  - create database
- RunScrape-Script

```
| => cat get.sql
CREATE OR REPLACE view getlatest as
SELECT c.mpg,c.milage,c.maketype,c.region,c.year,c.type,c.make,c.car_id,ph.price, ph.scrapedate, ph.sold
FROM cars c
INNER JOIN (
    SELECT car_id, MAX(scrapedate) AS max_date
    FROM pricehistory
    GROUP BY car_id
) latest_ph ON c.car_id = latest_ph.car_id
INNER JOIN pricehistory ph ON latest_ph.car_id = ph.car_id AND latest_ph.max_date = ph.scrapedate
WHERE ph.sold=0
```