

TP Infrastructure

Clarence Charles & Ewen Quimerc'h

1. SQL

Exercice 0: Décrivez les tables et les attributs.

```
exec sp_help
exec sp_columns tNames
exec sp_columns tPrincipals
exec sp_columns tTitles
```

Exercice 1: Visualisez l'année de naissance de l'artiste `Jude Law`.

```
select birthYear from [dbo].[tNames] where dbo.tnames.primaryName = 'Jude Law';
```

1972

Exercice 2: Comptez le nombre d'artistes présents dans la base de donnée.

```
select count(nconst) from [dbo].[tNames] ;
```

438361

Exercice 3: Trouvez les noms des artistes nés en `1960`, affichez ensuite leur nombre.

```
select primaryName from [dbo].[tNames] where dbo.tnames.birthYear = 1960;
select count(nconst) from [dbo].[tNames] where dbo.tnames.birthYear = 1960;
```

890

primaryName
Antonio Banderas
Kenneth Branagh
Colin Firth
Julianne Moore
Kristin Scott Thomas
Jean-Claude Van Damme

Exercice 4: Trouvez l'année de naissance la plus représentée parmi les acteurs (sauf 0!), et combien d'acteurs sont nés cette année là.

```
select count(nconst), birthYear from [dbo].[tNames] where birthYear <> 0 group by birthYear order by count(nconst) DESC;
```

Nombre	Année
1775	1980
1727	1979
1716	1978

Exercice 5: Trouvez les artistes ayant joué dans plus d'un film

```
select dbo.tNames.primaryName, number
from dbo.tNames
inner join
(
    select nconst, count(nconst) as number
    from dbo.tPrincipals
    where category = 'acted in'
    GROUP BY nconst
    HAVING count(nconst) > 1
) AS temp
on dbo.tNames.nconst = temp.nconst
```

primaryName	number
Gong Li	5
John Cleese	8
Brad Pitt	12
Woody Allen	2
Gillian Anderson	9
Pamela Anderson	3

Exercice 6: Trouvez les artistes ayant eu plusieurs responsabilités au cours de leur carrière (acteur, directeur, producteur...).

```

select primaryName
from (
    select nconst
    from tPrincipals
    group by nconst
    having count(distinct(category))>1
) as solution
join dbo.tNames
on dbo.tNames.nconst = solution.nconst;

```

primaryName
Brad Pitt
Woody Allen
Luc Besson
Kenneth Branagh
Pierce Brosnan
George Clooney

Exercice 7: Trouver le nom du ou des film(s) ayant le plus d'acteurs (i.e. uniquement *acted in*).

```

select tTitles.primaryTitle, actorsNumber
from tTitles
join (
    select tconst, count(nconst) as actorsNumber
    from tPrincipals
    where category = 'acted in'
    group by tconst
) as res
on res.tconst = tTitles.tconst
order by actorsNumber desc

```

primaryTitle	actorsNumber
P.O.W. Open Air am Meer	10
MacGuffin	10
Aakhri Pal	10
Hollywood Casting Confessions	10
Corazones en Llamas 3	10
Corazones en Llamas 2	10

Exercice 8: Montrez les artistes ayant eu plusieurs responsabilités dans un même film (ex: à la fois acteur et directeur, ou toute autre combinaison) et les titres de ces films.

```

select primaryName, primaryTitle
from (
    select nconst, tconst
    from tPrincipals
    group by nconst, tconst
    having count(distinct(category))>1
) as solution
join dbo.tTitles
on tTitles.tconst = solution.tconst
join dbo.tNames
on dbo.tNames.nconst = solution.nconst;

```

primaryName	primaryTitle
Samuthirakani	Vinodhaya Sitham
Brian George	Refinery Surveyor Black
Lom Harsh	Yeh Hai India
Trista Suke	Foxy
Chiwetel Ejiofor	The Boy Who Harnessed the Wind
Andy Barker	Rest Area

2. Cypher

Exercice 1: Ajoutez une personne ayant votre prénom et votre nom dans le graphe. Vérifiez que le nœud a bien été créé.

```

CREATE(Clarance: Name {nconst: 'cccccccc' , primaryName:'Clarence Charles' ,
birthYear: 1999})

```

```

MATCH (n:Name) WHERE n.primaryName = 'Clarence Charles' RETURN n

```

Exercice 2: Ajoutez un film nommé `L'histoire de mon 20 au cours Infrastructure de données`

```

CREATE(myFilm: Title {tconst: '11111111' , primaryTitle:"L'histoire de mon 20 en
Infra" , startYear: 2022})

```

Exercice 3: Ajoutez la relation `ACTED_IN` qui modélise votre participation à ce film en tant qu'acteur/actrice

```
MATCH
  (actor:Name{primaryName: 'Clarence Charles'}),
  (film:Title{primaryTitle: "L'histoire de mon 20 en Infra"})
CREATE (actor)-[r:acted_in]->(film)
RETURN type(r)
```

```
MATCH (n:Name{primaryName: 'Clarence Charles'})-[r]->(t:Title) RETURN n, r, t
```

Exercice 4: Ajoutez deux de vos professeurs/enseignants comme réalisateurs/réalisatrices de ce film.

```
CREATE(Name {nconst: 'cabaretl' , primaryName:'Laurent Cabaret' , birthYear:
1980})
```

```
CREATE(Name {nconst: 'quercini' , primaryName:'Gianluca Quercini' , birthYear:
1980})
```

```
MATCH
  (actor:Name),
  (film:Title{primaryTitle: "L'histoire de mon 20 en Infra"})
WHERE actor.nconst = 'cabaretl' OR actor.nconst = 'quercini'
CREATE (actor)-[r:directed]->(film)
RETURN type(r)
```

```
MATCH (n:Name)-[r]->(t:Title{tconst: 'l1111111'}) RETURN n, r, t
```

Exercice 5: Affichez le noeud représentant l'acteur nommé `Jude Law`, et visualisez son année de naissance.

```
MATCH(n:Name{primaryName:'Jude Law'}) RETURN n.birthYear
```

1972

Exercice 6: Visualisez l'ensemble des films.

```
MATCH(t:Title) return t limit 5000
```

Sans `limit`, la requête plante (timeout).

Exercice 7: Trouvez les noms des artistes nés en `1960`, affichez ensuite leur nombre.

```
MATCH (n:Name{birthYear: 1960}) RETURN count(n)
```

Exercice 8: Trouver l'ensemble des acteurs (sans entrées doublons) qui ont joué dans plus d'un film.

```
MATCH (u:Title)<-[:acted_in]-(n:Name)-[:acted_in]->(t:Title)
RETURN distinct n.primaryName limit 500
```

Exercice 9: Trouvez les artistes ayant eu plusieurs responsabilités au cours de leur carrière (acteur, directeur, producteur...).

```
MATCH ()<-[r]-(n:Name)-[s]->()
WHERE type(r) <> type(s)
RETURN n
```

Exercice 10: Montrez les artistes ayant eu plusieurs responsabilités dans un même film (ex: à la fois acteur et directeur, ou toute autre combinaison) et les titres de ces films.

```
MATCH (t:Title)<-[r]-(n:Name)-[s]->(t:Title)
WHERE type(r) <> type(s)
WITH DISTINCT n, t
RETURN n.primaryName, t.primaryTitle
```

Exercice 11: Trouver le nom du ou des film(s) ayant le plus d'acteurs.

```
MATCH (actor:Name)-[:acted_in]->(movie:Title)
RETURN
    movie.primaryTitle,
    COLLECT(DISTINCT actor.nconst),
    COUNT(DISTINCT actor) as COUNTING
ORDER BY COUNTING DESC LIMIT 5;
```

Alone 65 actors

3. Gremlin

Exercice 1: Ajoutez une personne ayant votre prénom et votre nom dans le graphe. Vérifiez que le noeud a bien été créé.

```
g.v().addV('Name').property(T.id, "cc").property("primaryName", "Clarence
Charles").property("birthYear", 1999).property("pk", "pk3")
```

Vérification:

```
g.v().has('primaryName', 'Clarence Charles')
```

Exercice 2: Ajoutez un film nommé `L'histoire de mon 20` au cours `Infrastructure de donnees`

```
g.v().addV('Title').property(T.id, "hh20").property("primaryTitle", "L'histoire de mon 20 en Infra").property("startYear", 2022).property("pk", "pk11")
```

Exercice 3: Ajoutez la relation `ACTED_IN` qui modélise votre participation à ce film en tant qu'acteur/actrice

```
g.v()
  .hasLabel('Name').has('primaryName', 'Clarence Charles').as('actor')
.v().hasLabel('Title').has('primaryTitle', "L'histoire de mon 20 en Infra").as('film')
.addE('acted_in').from('actor').to('film')
```

```
g.v().has('primaryName', 'Clarence Charles').outE('acted_in')
```

Exercice 4: Ajoutez deux de vos professeurs/enseignants comme réalisateurs/réalisatrices de ce film.

```
g.addV('Name').property(T.id, "lc").property('primaryName', 'Laurent Cabaret').property('birthYear', 1980).property("pk", "pk4")
```

```
g.addV('Name').property(T.id, "gq").property('primaryName', 'Gianluca Quercini').property('birthYear', 1980).property("pk", "pk5")
```

```
g.v().has('primaryTitle', "L'histoire de mon 20 en Infra").as('film')
.v().where(__.has(id, 'lc').or().has(id, 'gq')).as('actor')
.addE('directed').from('actor').to('film')
```

```
g.v().has('id', 'lc').outE('directed')
```

Exercice 5: Affichez le noeud représentant l'acteur nommé `Jude Law`, et visualisez son année de naissance.

```
g.v().hasLabel('Name').has('primaryName', eq('Jude Law'))
  .project('birthYear').by('birthYear')
```

1972

Exercice 6: Visualisez l'ensemble des films.

```
g.v().hasLabel('Title')
```

Exercice 7: Trouvez les noms des artistes nés en 1960, affichez ensuite leur nombre.

```
g.v().hasLabel('Name').has('birthYear', 1960).count()
```

890

Exercice 8: Trouver l'ensemble des acteurs (sans entrées doublons) qui ont joué dans plus d'un film.

```
g.v().hasLabel('Title').inE('acted_in').as('film_in_acted_in')
.outV().as('n')
.hasLabel('Name').outE('acted_in').as('film_out_acted_in')
.inV().hasLabel('Title')
.where(__.select('film_in_acted_in').where(neq('film_out_acted_in'))).select('n')
.project('primaryName').by('primaryName').dedup()
```

Exercice 9: Trouvez les artistes ayant eu plusieurs responsabilités au cours de leur carrière (acteur, directeur, producteur...).

```
g.v().inE().as('r')
.outV().hasLabel('Name').as('persons')
.outE().as('s')
.inV().where(__.select('r').where(neq('s'))).select('persons')
```

Exercice 10: Montrez les artistes ayant eu plusieurs responsabilités dans un même film (ex: à la fois acteur et directeur, ou toute autre combinaison) et les titres de ces films.

```
g.v().hasLabel('Title').as('t')
.v().hasLabel('Title').as('t2')
.v().inE().as('r')
.v().outE().as('s')
.v().hasLabel('Name').as('n')
.where(__.select('t2').where(eq('t')).where(__.select('r').where(neq('s')))).select('n')
.project('n').by('primaryName').dedup()
```

Exercice 11: Trouver le nom du ou des film(s) ayant le plus d'acteurs.

```
g.v().hasLabel('Name').as('actor')
.outE('acted_in').inV().hasLabel('Title').as('movie')
.select('movie', 'actor').group().by('primaryTitle')
.by(__.fold().project('primaryTitle', 'COLLECT(DISTINCT id)', 'COUNTING'))
.dedup().fold()
.by(__.unfold().select('actor').dedup().count()).unfold().select(values).order(
).by(__.select('COUNTING'), desc)
```


