

מערכות מסדי נתונים תרגיל בית 2 – תשפ"א – סמסטר אביב

חלק 1 – אלגברה רלציונית ו-SQL:

כתבו את השאילתות הבאות באלגברה רלציונית:

1. כתבו שאילתה המחזירה את פרטי העובדים (מספר ושם מלא) שהיו חלק מקבוצה שפיתחה לפחות 3 מוצרים שהמחיר שלהם מתחת ל-100 בשנתיים האחרונות.

$T1 \leftarrow \pi_{\text{product_ID}} (\sigma_{\text{Datediff}(\text{develop_finish_date}, \text{curdate}()) \leq 365 * 2 \wedge \text{product_consumer_price} < 100} (\text{Develops} \bowtie \text{Product}))$

$T2 \leftarrow \pi_{\text{team_id}} (\sigma_{\text{count}(\text{product_ID}) \geq 3} (\text{team_id } g \text{ count}(\text{product_ID}) (\text{Develops} \bowtie T1)))$

$\text{answer} \leftarrow \pi_{\text{employee_ID}, \text{employee_FirstName}, \text{employee_LastName}} (\text{Employee} \bowtie \text{RND_employee} \bowtie T2)$

2. כתבו שאילתה המחזירה את פרטי עובדי ה- Devops (מספר ושם מלא) שלא ניהלו אף מוצר שנוצר במפעלי AMDEL בשנה האחרונה.

$T1 \leftarrow \pi_{\text{devops_employee_ID}} (\sigma_{\text{Datediff}(\text{Production_date}, \text{curdate}()) \leq 365} (\text{Produced_In} \bowtie \text{Product}))$

$\pi_{\text{employee_ID}, \text{employee_FirstName}, \text{employee_LastName}} (\text{Employee} \bowtie \text{devops_employee})$

—

$\pi_{\text{employee_ID}, \text{employee_FirstName}, \text{employee_LastName}} (\text{employee} \bowtie T1)$

3. כתבו שאילתה שמחזירה את ת.ז. ושמות העובדים שעובדים כרגע בחברה בתור עובדי ייצור או עובדים שעבדו בעבר בתור עובדי rnd והיו מנהלים של צוותים.

$T1 \leftarrow \pi_{employee_ID, employee_FirstName, employee_LastName} (\sigma_{job_finish_date \text{ is null})$

$(Employee \bowtie job \bowtie Production_employee)$

$T2 \leftarrow \pi_{employee_ID, employee_FirstName, employee_LastName} (\sigma_{job_finish_date \text{ is not null})$

$(Employee \bowtie job)$

$T3 \leftarrow \pi_{employee_ID, employee_FirstName, employee_LastName} (\sigma_{manager_ID = rnd_employee_ID \wedge rnd_employee_ID = employee_ID})$

$(Rnd_team \bowtie Rnd_employee \bowtie Employee)$

$answer \leftarrow (T2 \cap T3) \cup T1$

כתוב את השאילתות הבאות בשפת SQL תוך שימוש ב-MYSQL. ניתן להשתמש בפונקציות שלא נלמדו בשיעור. במקרה כזה יש להוסיף לינק להגדרת הפונקציה על פי MYSQL. אין להשתמש ב view או temp- לצורך הפתרון אלא אם מצויין אחרת:

1. כתבו שאילתה המחזירה מספר העובדים, שמות ומספרי טלפון של עובדים השייכים לקבוצות שפיתחו לפחות עשרה מוצרי "CPU" בפחות מ-5 שנים, כך שבכל אחד מהם ה-CPU cores הוא לפחות 2.

```
#1
SELECT e.employee_ID, e.employee_FirstName, e.employee_LastName
from employee as e inner join rnd_employee as rnde on e.employee_ID=rnde.rnd_employee_ID
where rnde.team_id in ( SELECT d.team_id
                        from develops as d
                        inner join product as p on p.product_id = d.product_id inner join
                        cpu on cpu.cpu_product_id = p.product_id
                        where cpu.cpu_cores >=2 and
                        timestampdiff(day, d.develop_start_date, d.develop_finish_date )<5*365
                        group by d.team_id
                        having count(cpu.cpu_product_id)>=10);
```

2. כתבו שאילתה המחזירה את מספרי הזיהוי של צוותי ה-RND שעונים על כל התנאים הבאים:

- מספר העובדים בצוות ה-RND לא עולה על 7.
- לא היה אף מוצר שפיתחו ועלה יותר מ-250.

- פיתחו לפחות מוצר אחד מסוג Network controller שיש לו bandwidth של פחות מ- 25 והוא לא מיוצר במפעלי AMDEL.

#2

```
SELECT team_id
from rnd_employee
where (team_id not in ( SELECT distinct d.team_id
                        from develops as d
                        inner join product as p on p.product_id = d.product_id
                        where p.product_consumer_price > 250))
and
(team_id in ( SELECT distinct d.team_id
              from develops as d inner join network_controller as nc on nc.network_controller_product_id = d.product_id inner join
              supports as sup on sup.network_controller_product_id = nc.network_controller_product_id
              where sup.bandwidth < 25 and (nc.network_controller_product_id in (Select product_id from outsourced_to))))
group by team_id
having count(rnd_employee_ID)<=7;
```

3. כתבו שאילתה המחזירה פרטי הייצרנים החיצוניים. על השאילתה להחזיר דו"ח מהצורה הבאה. (אם ישנו יצרן שלא מייצר אף מוצר, אין להראות אותו בפתרון):

מספר יצרן	שם יצרן	סוג המוצר	כמות המוצרים	מוצר הכי זול שנוצר	המוצר הכי יקר שנוצר
1523001	intel	GPU	60	50	250
1523001	intel	CPU	70	50	100
1523001	intel	Network controller	150	120	300
2530020	Apple	GPU	290	300	690
2530020	Apple	CPU	320	200	800
1330002	NIVIDIA	GPU	120	110	290
...

#3

```
SELECT p.producer_id, p.producer_name, 'CPU' as product_type, count(prod.product_id), min(prod.product_consumer_price), max(prod.product_consumer_price)
from producer as p inner join outsourced_to as ot on p.producer_id = ot.producer_id inner join product as prod on prod.product_id = ot.product_id
inner join cpu on cpu.cpu_product_id = prod.product_id
group by p.producer_id, p.producer_name
union
SELECT p.producer_id, p.producer_name, 'GPU' as product_type, count(prod.product_id), min(prod.product_consumer_price), max(prod.product_consumer_price)
from producer as p inner join outsourced_to as ot on p.producer_id = ot.producer_id inner join product as prod on prod.product_id = ot.product_id
inner join gpu on gpu.gpu_product_id = prod.product_id
group by p.producer_id, p.producer_name
union
SELECT p.producer_id, p.producer_name, 'Network Controller' as product_type, count(prod.product_id), min(prod.product_consumer_price), max(prod.product_consumer_price)
from producer as p inner join outsourced_to as ot on p.producer_id = ot.producer_id inner join product as prod on prod.product_id = ot.product_id
inner join network_controller as nc on nc.network_controller_product_id = prod.product_id
group by p.producer_id, p.producer_name;
```

4. כתבו שאילתה המחזירה את פרטי המוצרים (המספר והשם של המוצר, ה-Gross profit ואיפה נוצר המוצר). יש להחזיר מוצרים שעונים על ההנחה הבאה:
במידה והחברה מוכרת מהם 100 יחידות אז ה-Gross profit שלהם מ-100 היחידות הוא יותר מ-20,000. יש לסדר את המוצרים כך שקודם מוצרים הרווחיים יותר.

תוצאת שאילתה לדוגמא:

Product id	Gross profit	Produced at
5	150,000	Amdel Fab
2	78,000	Ousourced
32	25,000	Ousourced
...

#4

```
select p.product_id, ((100*p.product_consumer_price)-(100*(pin.production_cost + amdf.operation_cost))) as gross_profit, 'Amdel Fab' as Produced_at
from product as p inner join produced_in as pin on p.product_id = pin.product_id inner join amdel_fab as amdf on amdf.amdel_fab_id = pin.amdel_fab_id
where ((100*p.product_consumer_price)-(100*(pin.production_cost+amdf.operation_cost)))>20000
union
#returns the gross profit when a product is produced out
select p.product_id, ((100*p.product_consumer_price)-(100*(outs.cost_per_unit))) as gross_profit, 'Outsourced' as Produced_at
from product as p inner join outsourced_to as outs on p.product_id = outs.product_id
where ((100*p.product_consumer_price)-(100*(outs.cost_per_unit)))>20000
order by gross_profit desc;
```

5. כתבו שאילתה המחזירה את המספר המזהה של המפעלים שלא יצרו אף מוצר מסוג CPU אי פעם, אך יצרו לפחות שני מוצרים מסוג GPU בחודש האחרון. יש למיין את התוצאות בסדר יורד לפי מספר המפעל.

#5

```
select AF.amdel_fab_id
from CPU inner join produced_in as PI on CPU.cpu_product_id=PI.product_id
right outer join amdel_fab as AF on PI.amdel_fab_id = AF.amdel_fab_id
where cpu_product_id is null and AF.amdel_fab_id in (
select PI.amdel_fab_id
from Network_controller as NC inner join produced_in as PI on NC.network_controller_product_id=PI.product_id
where PI.amdel_fab_id = AF.amdel_fab_id and timestampdiff(day,production_datetime,curdate())<=30
group by amdel_fab_id
having count(NC.network_controller_product_id)>=2)
order by AF.amdel_fab_id desc;
```

חלק 2 – NoSQL:

Property Graphs

להלן תיאור דרישות למסד נתונים:
חברת הביטוח ריקודביט מעוניינת להקים מערכת לאיתור ניסיונות הונאה. על המערכת לאפשר תיעוד לקוחות, פוליסות הביטוח שהם רכשו, והתביעות שהן הגישו בהקשר לפוליסה. ללקוחות יש לשמור את תעודת הזהות, השם, רחוב, מספר בית, עיר, מיקוד, תאריך לידה וקשרים משפחתיים ידועים לאנשים אחרים (בעל/אישה/ילד). לתביעות יש לשמור את הסכום, תיאור המקרה והאדם או חברה אליהם התשלום נועד. לעיתים אלה שונים מאשר בעל הפוליסה. למשל, אם בעל הפוליסה פגע בהולך רגל עם רכבו, הפיצוי מיועד להולך הרגל. יש לקטלג פוליסות לפי סוג הפוליסה ויש לשמור את סכום הפרמיה החודשית, מועד ההתחלה והסיום של הפוליסה.

- הציעו מודל property graph מתאים
- הגישו גרף לדוגמא המדגים את כל רכיבי המודל שלכם.
- כתבו שאילתת Cypher היוצרת את הגרף לדוגמא שהצעתם.
- כתבו שאילתת Cypher המחזירה טבלה ובה שלוש שורות ושתי עמודות מעל הגרף שהצעתם.

Document Databases

מיצאו באינטרנט אוסף קבצי JSON המכיל לפחות 3 קבצים באותו נושא.
א. רישמו בפורום שנפתח לנושא קישור לאוסף שמצאתם ותיאור קצר. אין לעשות שימוש באוסף שכבר נבחר על ידי קבוצה אחרת.

- ב. האם הסכמה של האוסף אחידה או שיש שוני בין מסמך למסמך? פרטו את ההבדלים אם ישנם.
- ג. הציעו שאילתת find() מעל האוסף שלכם שמחזירה רק את חלקו אבל יותר משני מסמכים.
- ד. הציעו מודל רלציוני מנורמל לשמירת המידע באוסף.