

## Cvičení 8: Množinové operátory UNION, EXCEPT a INTERSECT

- 1) Pomocí jednoho z množinových operátorů nalezněte seznam zákazníků (*Customer\_ID*), kteří ještě neuskutečnili žádnou objednávku.
- 2) Pomocí jednoho z množinových operátorů nalezněte seznam zákazníků (*Customer\_ID*), kteří uskutečnili alespoň jednu objednávku, ale údaj o nich chybí v detailní tabulce zákazníků.
- 3) Pomocí jednoho z množinových operátorů nalezněte seznam produktů (*Product\_ID*), které ještě nebyly nikým objednány. Poté najděte seznam těch, které najdeme v číselníku produktů a zároveň již někdy byly objednány.
- 4) Napište dotaz, který vrátí TOP 5 objednávek (*Order.Order\_ID*) s nejvyšší hodnotou, která se vypočítá vynásobením množství z tabulky objednávek (*Order.Quantity*) a jednotkové ceny produktu (*Product.Unit\_Price*). V tomto případě nás budou zajímat jen záznamy s jednotkovou cenou v CZK.
- 5) Dotaz z předchozí úlohy budeme chtít omezit další podmínkou, a to tak, že nás zajímají pouze objednávky se stavem 'Zaplacená'. Za tímto účelem bude potřeba připojit tabulku **OrderStatus**, v jejímž sloupci *OrderStatus\_Name* jsou právě zmíněné popisy stavů objednávek.
- 6) Nyní dotaz upravte tak, že bude vracet pouze průměrnou výši objednávky pro záznamy splňující podmínky definované v předchozí úloze. Výsledkem dotazu bude tedy jen jedna hodnota.
- 7) Do dotazu z předchozí úlohy přidejte nový sloupec (v pořadí by měl být jako první) s názvem *PopisVýpočtu*, který se bude plnit hodnotou 'Průměr ze zaplacených objednávek'.
- 8) Vytvořte obdobu dotazu z předchozí úlohy, který však bude počítat jen se záznamy se stavem objednávky 'Nezaplacená', a proto také hodnota sloupce *PopisVýpočtu* bude 'Průměr z nezaplacených objednávek'.

- 9) Dotazy z úloh 7 a 8 spojte do jedné tabulky jedním z množinových operátorů.
- 10) Vytvořte obdobu dotazu z úlohy 8, který však bude počítat jen se záznamy se stavem objednávky 'Zrušená', a proto také hodnota sloupce *PopisVýpočtu* bude 'Průměr ze zrušených objednávek'.
- 11) Dotazy z úloh 7, 8 a 10 spojte do jedné tabulky jedním z množinových operátorů.

## Cvičení 8: Množinové operátory UNION, EXCEPT a INTERSECT

- 1) Pomocí jednoho z množinových operátorů nalezněte seznam zákazníků (*Customer\_ID*), kteří ještě neuskutečnili žádnou objednávku.

Pokud hledáme záznamy patřící do jedné množiny, ale už ne do jiné, použijeme operátor **EXCEPT**. V tomto případě nejdříve vybereme množinu všech zákazníků z tabulky **Customer**, ze které odfiltrujeme ty, kteří se alespoň jednou vyskytli v tabulce **Order**, a tudíž provedli alespoň jednu objednávku. Tím získáme pouze zákazníky, kteří řádnou objednávku zatím neuskutečnili.

```
SELECT DISTINCT Customer_ID  
FROM [csv].[Customer]
```

```
EXCEPT
```

```
SELECT DISTINCT Customer_ID  
FROM [csv].[Order]
```

	Customer_ID
1	C230
2	C231
3	C233

- 2) Pomocí jednoho z množinových operátorů nalezněte seznam zákazníků (*Customer\_ID*), kteří uskutečnili alespoň jednu objednávku, ale údaj o nich chybí v detailní tabulce zákazníků.

V tomto případě bude postup přesně opačný než v předchozí úloze. Hledáme zákazníky z tabulky **Order**, kteří se nevyskytují v tabulce **Customer**.

Takové záznamy jsou bezesporu chybné a měly by být opraveny – buď bude zákazník přidán do tabulky **Customer**, nebo bude jeho kód v tabulce **Order** opraven, aby odpovídal nějaké již existující hodnotě.

```
SELECT DISTINCT Customer_ID  
FROM [csv].[Order]
```

```
EXCEPT
```

```
SELECT DISTINCT Customer_ID  
FROM [csv].[Customer]
```

	Customer_ID
1	C196
2	C197
3	C198
4	C199
5	C200
6	C201
7	C202
8	C203
9	C204
10	C205

- 3) Pomocí jednoho z množinových operátorů naleznete seznam produktů (*Product\_ID*), které ještě nebyly nikým objednány. Poté najděte seznam těch, které najdeme v číselníku produktů a zároveň již někdy byly objednány.

První dotaz je obdobou úlohy 1, jen je jeho výsledek prázdnou množinou, což znamená, že neexistují žádné produkty z tabulky **Product**, které by se již nevyskytovaly také v tabulce **Order**.

```
SELECT DISTINCT Product_ID FROM [csv].[Product]
EXCEPT
SELECT DISTINCT Product_ID FROM [csv].[Order]
```

V druhém dotazu použijeme operátor **INTERSECT**, jelikož hledáme **průnik** obou množin, tedy produktů z číselníku a z objednávek. Výsledný seznam obsahuje počet produktů shodný s tím v číselníku, a tedy všechny z nich již někdy byly objednané.

```
SELECT DISTINCT Product_ID FROM [csv].[Product]
INTERSECT
SELECT DISTINCT Product_ID FROM [csv].[Order]
```

- 4) Napište dotaz, který vrátí TOP 5 objednávek (*Order.Order\_ID*) s nejvyšší hodnotou, která se vypočítá vynásobením množství z tabulky objednávek (*Order.Quantity*) a jednotkové ceny produktu (*Product.Unit\_Price*). V tomto případě nás budou zajímat jen záznamy s **jednotkovou cenou v CZK**.

Za účelem výpočtu **výše objednávky** a také definování podmínky na **měnu** je nutné k tabulce **Order** připojit tabulku **Product**.

V tomto případě klauzule **INNER JOIN** i **LEFT JOIN** vrátí shodný výsledek, jelikož je v dotazu použita **WHERE** podmínka na konkrétní hodnotu určitého sloupce z připojované tabulky (**Product.Currency = 'CZK'**).

Ta nemůže být splněna žádným záznamem, pro který by neexistovala vazba mezi levou a pravou tabulkou, a proto obě **JOIN** klauzule povedou ke stejnému výsledku.

```
SELECT TOP(5) ord.Order_ID
              ,ord.Quantity * prod.Unit_Price AS VýšeObjednávky
FROM [csv].[Order] ord
```

```
LEFT JOIN [csv].[Product] prod
ON ord.Product_ID = prod.Product_ID

WHERE prod.Currency = 'CZK'
ORDER BY VýšeObjednávky DESC
```

	Order_ID	VýšeObjednávky
1	O1168	828000
2	O1142	476000
3	O1158	441000
4	O1141	433500
5	O1195	405000

- 5) Dotaz z předchozí úlohy budeme chtít omezit další podmínkou, a to tak, že nás zajímají pouze objednávky se stavem 'Zaplacená'. Za tímto účelem bude potřeba připojit tabulku **OrderStatus**, v jejímž sloupci **OrderStatus\_Name** jsou právě zmíněné **popisy stavů objednávek**.

Jediný způsob, jak lze k tabulce **Order** připojit tabulku **OrderStatus** je použít vztah mezi sloupci **Order.Order\_Status** a **OrderStatus.OrderStatus\_ID**. Poté můžeme sloupec **OrderStatus\_Name** libovolně používat k filtrování či jiným operacím v rámci daného dotazu.

```
SELECT TOP(5) ord.Order_ID
               ,ord.Quantity * prod.Unit_Price AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
ON ord.Product_ID = prod.Product_ID

WHERE prod.Currency = 'CZK'
AND status.OrderStatus_Name = 'Zaplacená'

ORDER BY VýšeObjednávky DESC
```

	Order_ID	VýšeObjednávky
1	O1168	828000
2	O1142	476000
3	O1158	441000
4	O1141	433500
5	O1195	405000

Výsledek dotazu z předchozí úlohy je shodný s touto úlohou, a proto víme, že všech 5 objednávek s nejvyšší hodnotou je již **zaplaceno**.

- 6) Nyní dotaz upravte tak, že bude vracet pouze průměrnou výši objednávky pro záznamy splňující podmínky definované v předchozí úloze. Výsledkem dotazu bude tedy jen jedna hodnota.

Podmínky i napojení tabulek zůstávají zachovány, jen z klauzule **SELECT** odstraníme sloupec **Order\_Id** a necháme zde pouze výpočet **výše objednávky**, který však zahrneme do **agregační funkce AVG**. Tím získáme průměr ze všech záznamů splňující dané **WHERE** podmínky.

```
SELECT AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
AND status.OrderStatus_Name = 'Zaplacená'
```

- 7) Do dotazu z předchozí úlohy přidejte nový sloupec (v pořadí by měl být jako první) s názvem *PopisVýpočtu*, který se bude plnit hodnotou 'Průměr ze zaplacených objednávek'.

Nyní do předchozího dotazu pouze **přidáme sloupec** s názvem ***PopisVýpočtu***, který však není založený na hodnotě nějakého existujícího sloupce nebo výpočtu, ale jedná se o konkrétní textový řetězec, který tedy definujeme v jednoduchých uvozovkách **před** klauzulí **AS** a **názvem tohoto nového sloupce**.

```
SELECT 'Průměr ze zaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
      AND status.OrderStatus_Name = 'Zaplacená'
```

	PopisVýpočtu	VýšeObjednávky
1	Průměr ze zaplacených objednávek	4957

- 8) Vytvořte obdobu dotazu z předchozí úlohy, který však bude počítat jen se záznamy se stavem objednávky 'Nezaplacená', a proto také hodnota sloupce *PopisVýpočtu* bude 'Průměr z nezaplacených objednávek'.

Tento dotaz je téměř stejný jako v předchozí úloze, jen je podmínka pro **stav objednávky** změněna na hodnotu '**Nezaplacená**' spolu s odpovídající úpravou textového řetězce pro sloupec ***PopisVýpočtu***.

```
SELECT 'Průměr z nezaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
      AND status.OrderStatus_Name = 'Nezaplacená'
```

	PopisVýpočtu	VýšeObjednávky
1	Průměr z nezaplacených objednávek	3042

9) Dotazy z úloh 7 a 8 spojte do jedné tabulky jedním z množinových operátorů.

Za účelem spojení výsledků více dotazů použijeme operátor **UNION**. Díky tomu, že jsou splněny všechny potřebné podmínky pro správné fungování množinových operátorů, vrátí dotaz tabulku složenou z obou záznamů.

```
SELECT 'Průměr ze zaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
      AND status.OrderStatus_Name = 'Zaplacená'

UNION

SELECT 'Průměr z nezaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
      AND status.OrderStatus_Name = 'Nezaplacená'
```

	PopisVýpočtu	VýšeObjednávky
1	Průměr z nezaplacených objednávek	3042
2	Průměr ze zaplacených objednávek	4957

10) Vytvořte obdobu dotazu z úlohy 8, který však bude počítat jen se záznamy se stavem objednávky 'Zrušená', a proto také hodnota sloupce *PopisVýpočtu* bude 'Průměr ze zrušených objednávek'.

```
SELECT 'Průměr ze zrušených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK'
      AND status.OrderStatus_Name = 'Zrušená'
```

	PopisVýpočtu	VýšeObjednávky
1	Průměr ze zrušených objednávek	3170

11) Dotazy z úloh 7, 8 a 10 spojte do jedné tabulky jedním z množinových operátorů.

Nakonec můžeme spojit všechny tři dotazy do jednoho souhrnného, který vrátí **průměrnou hodnotu** pro každý **stav objednávky** ve formě přehledné tabulky.

Záznamy výsledné tabulky můžeme **seřadit** dle libovolného sloupce z dotazů spojovaných operátorem **UNION**, ale daná **ORDER BY** podmínka může být napsána až za **posledním z těchto dotazů**.

```
SELECT 'Průměr ze zaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK' AND status.OrderStatus_Name = 'Zaplacená'

UNION

SELECT 'Průměr z nezaplacených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK' AND status.OrderStatus_Name = 'Nezaplacená'

UNION

SELECT 'Průměr ze zrušených objednávek' AS PopisVýpočtu
      ,AVG(ord.Quantity * prod.Unit_Price) AS VýšeObjednávky
FROM [csv].[Order] ord
LEFT JOIN [dbo].[OrderStatus] status
      ON ord.Order_Status = status.OrderStatus_ID
LEFT JOIN [csv].[Product] prod
      ON ord.Product_ID = prod.Product_ID
WHERE prod.Currency = 'CZK' AND status.OrderStatus_Name = 'Zrušená'

ORDER BY VýšeObjednávky DESC
```

	PopisVýpočtu	VýšeObjednávky
1	Průměr ze zaplacených objednávek	4957
2	Průměr ze zrušených objednávek	3170
3	Průměr z nezaplacených objednávek	3042