
EXAMEN: Scriptingtalen

Prof. Dr. Peter Dawyndt
1^e Bachelor Informatica
groep 1-3

vrijdag 22-06-2007, 8:30h
academiejaar 2006-2007
eerste zittijd

Opgave 1

De Women's Tennis Association (WTA) is de overkoepelende organisatie voor professionele tennis-speelsters. De WTA organiseert jaarlijks een reeks van wedstrijden, waarbij het klassement dat door de WTA wordt opgesteld algemeen geldt als de officiële rangschikking van de beste tennisspeelsters ter wereld.

Het gegeven ZIP-bestand `WTA.zip` bevat een reeks bestanden `WTAYyyy`, waarbij elk bestand de rangschikking bevat van het WTA-klassement op het einde van het kalenderjaar `yyyy`. Deze statistieken werden opgemaakt voor elke jaargang tussen 1989 en 2006. Elk bestand `WTAYyyy` bevat het jaartal `yyyy` op een afzonderlijk regel, gevolgd door regels met de WTA top-20 van die jaargang. Elke regel bestaat uit drie velden, gescheiden door tabs, die respectievelijk i) de positie in de rangschikking, ii) de naam van de speelster, en iii) de nationaliteit van de speelster aangeven. Regels die informatie bevatten kunnen van elkaar worden gescheiden door één of meerdere lege regels.

Gevraagd wordt om een `awk`-script `WTAstats` te schrijven dat een overzicht geeft van de verschillende speelsters die in bepaalde jaargangen in de WTA top-20 stonden. Het script moet kunnen aangeroepen worden op de Unix-commandolijn als `WTAstats <files>`, waarbij een willekeurig aantal `WTAYyyy` bestanden als argument aan het script kan worden doorgegeven. Lege regels in de bestanden worden door het script genegeerd. Als we het script bijvoorbeeld met de volgende argumenten aanroepen

```
> WTAstats WTA2006 WTA2005 WTA2004
```

dan wordt een overzicht van alle speelsters die in minstens één van de opgegeven jaargangen in de WTA top-20 stonden naar standaard uitvoer geschreven, waarbij voor elke speelster haar top-20 posities worden weergegeven in de corresponderende kolommen (lege waarde voor jaargangen waarin de speelster niet in de top-20 stond). De eerste paar regels die als resultaat van het bovenstaande voorbeeld worden gegenereerd zijn

	2006	2005	2004
Justine Henin	1	6	8
Maria Sharapova	2	4	4
Amelie Mauresmo	3	3	2
Svetlana Kuznetsova	4	18	5
Kim Clijsters	5	2	

De eerste kolom bevat de namen van de speelsters, waarvoor 25 karakters worden gereserveerd, en is links uitgelijnd. Voor de kolommen van de jaargangen worden 9 karakters gereserveerd en worden de waarden rechts uitgelijnd. De eerste regel van de uitvoer bevat de verschillende jaargangen, opgemaakt zoals in het voorbeeld wordt aangegeven.

Het script `WTAstats` moet voorts een functie `trim` bevatten waaraan een tekenreeks `str` kan worden doorgegeven. De functie verwijdert alle spaties vooraan en achteraan de tekenreeks `str`. Indien de tekenreeks de naam van een speelster voorstelt waarvan de familienaam een koppelteken bevat, dan moet het gedeelte na het koppelteken verwijderd worden (inclusief koppelteken zelf). Zo wordt de tekenreeks `Justine Henin-Hardenne` bijvoorbeeld gereduceerd tot `Justine Henin`, maar blijft de tekenreeks `Anna-Lena Groenefeld` haar oorspronkelijke waarde behouden. Je mag ervan uitgaan dat

er in de voornaam van een speelster geen spaties voorkomen. Zorg ervoor dat alle lokale variabelen van de functie ook als dusdanig worden behandeld. De functie `trim` moet in het script worden aangeroepen voor een uniforme verwerking van de namen van speelsters.

Opgave 2

Voor deze opgave moet je een reeks Unix-commando's ontwikkelen voor het beheren en gebruiken van een eenvoudige citaten-databank. Deze databank wordt bewaard als een tekstbestand `citaten.txt` in de home directory van de gebruiker die de scripts aanroept. Elk citaat bestaat uit meerdere regels, waarvan de eerste regel telkens de auteur van het citaat bevat. Citaten worden van elkaar gescheiden door een regel die enkel twee procenttekens bevat (`%%`). Een voorbeeld van een bestand `citaten.txt` zou dus kunnen zijn

```
D. Adams
I may not have gone where I intended to go,
but I think I have ended up where I needed to be.
%%
A. Einstein
Not everything that counts can be counted,
and not everything that can be counted counts.
%%
S. Hawing (A Brief History of Time)
Someone told me that each equation I
included in the book would halve sales.
%%
T. Pratchett (Hogfather)
Real stupidity beats
artificial intelligence
every time.
%%
Matthew XXII.14
For many are called, but few are chosen.
%%
```

Gevraagd wordt

1. schrijf een **bash** shell script `insertCitaat` dat de auteur en de tekst van het citaat opvraagt aan de gebruiker. De tekst van het citaat kan uit meerdere regels bestaan, en moet door de gebruiker worden afgesloten met een regel die enkel twee procenttekens (`%%`) bevat. Zorg voor een verzorgde interactie met de gebruiker, waarbij gepaste vragen worden gesteld door het shell script. De volgende sessie zal bijvoorbeeld het eerste citaat uit het voorbeeld toegevoegen aan het bestand `citaten.txt`:

```
> insertCitaat
Geef de auteur van het citaat:
D. Adams
Geef de tekst van het citaat:
I may not have gone where I intended to go,
but I think I have ended up where I needed to be.
%%
>
```

Het shell script moet controleren of er reeds een bestand `citaten.txt` bestaat in de home directory van de gebruiker. Indien dit niet het geval is, dan moet dit bestand door het script worden aangemaakt.

2. schrijf een **bash** shell script `selectCitaat <nr>` dat voor een gegeven argument `nr` het citaat op de corresponderende positie in `citaten.txt` wegschrijft naar standaard uitvoer. Voor het uitknippen van het gepaste citaat mag je vrij kiezen of je daarvoor `sed` of `awk` gebruikt. Zorg ervoor dat het shell script controleert of het opgegeven argument binnen geldige onder- en bovengrenzen ligt, en dat er ook een gepaste opmerking wordt gegeven indien er geen argumenten worden meegegeven aan het script. De volgende sessie illustreert de functionaliteit die het shell script `selectCitaat` moet hebben voor het voorbeeldbestand `citaten.txt` dat hoger werd gegeven:

```
> selectCitaat
Syntax: selectCitaat <nr>
> selectCitaat 10
Syntax: selectCitaat <nr>
Waarschuwing: de waarde van nr moet in het interval [1,5] gelegen zijn.
> selectCitaat 4
T. Pratchett (Hogfather)
Real stupidity beats
artificial intelligence
every time.
>
```

3. schrijf een `sed` script `formatCitaat` waarmee de uitvoer van `selectCitaat` wordt opgemaakt zodat het citaat tussen dubbele quotes komt te staan, en de auteur achter de tekst van het citaat geplaatst wordt. Kies daarbij voor een oplossing die optimaal gebruik maakt van de functionaliteit die in `sed` zit ingebakken. De volgende sessie illustreert het gebruik van dit nieuwe script:

```
> selectCitaat 4 | sed -f formatCitaat
"Real stupidity beats
artificial intelligence
every time."
T. Pratchett (Hogfather)
>
```

Opgave 3

De DVD-verhuurketen die gebruikt maakt van de Sakila databank is ondertussen online gegaan. Films kunnen nu ook via de website van de keten besteld worden, en worden dan naar het thuisadres opgestuurd. Hiervoor werden twee extra tabellen toegevoegd aan de oorspronkelijke databank.

De tabel `online_customer` houdt een unieke lijst met IP-adressen van de klanten bij in de kolom `IP_ADDRESS`, die de primaire sleutel van de tabel vormt. Voor elk IP-adres bevatten de records voorts ook nog de hostname (`HOST_NAME`), het land (`COUNTRY`), de landcode (`COUNTRY_CODE`), de regio (`REGION`), de stad (`CITY`), de postcode (`POSTAL_ADDRESS`), de telefoon-prefix (`CALLING_CODE`), en de lengte- en breedtegraden van de geografische locatie waar de klant zich bevindt (`LONGITUDE` en `LATITUDE`). De primaire sleutel `IP_ADDRESS` van de tabel `online_customer` werd ook als vreemde sleutel toegevoegd aan de tabel `customer`, waarmee wordt uitgedrukt dat online klanten ook gewone klanten zijn (gewone klanten zijn niet noodzakelijk ook online klanten).

Voor elke online bestelling wordt een record toegevoegd aan de tabel `online_order`. Hierbij worden volgende gegevens geregistreerd: het IP-adres van de machine waarop de bestelling gebeurde (`IP_ADDRESS`), de unieke sleutel van de bestelde film (`FILM_ID`) en de dag waarop de bestelling gebeurde (`ORDER_TIME`).

Gevraagd wordt om SQL zoekopdrachten te formuleren die een antwoord bieden op de volgende vragen. Zorg er telkens voor dat de kolommen van de resulterende tabel een zinvolle naam krijgen. Geef op je antwoordblad ook aan hoeveel records de resultatentabel bevat.

1. Uit welke landen komen de meeste klanten die online kopen? Bepaal per land hoeveel online klanten er zijn. Sorteer de landen in dalende volgorde wat betreft de populariteit van de online verkoop.
2. Geef de voornaam, de familienaam en het land van herkomst van alle klanten die een online bestelling geplaatst hebben in de huidige maand (juni 2007). Sorteer het resultaat eerst per land, en daarna per familienaam.
3. Geef een overzicht van de landen van herkomst van alle klanten die de film `HOCUS FRIDA` of de film `DOGMA FAMILY` online hebben besteld. Geef daarbij voor elk land aan hoeveel verschillende klanten er minstens één van beide films online hebben besteld, en sorteer de landen volgens dalend aantal online bestellingen.
4. Geef de titels voor elk paar films die samen door minstens 8 klanten online werden besteld. De films hoeven hierbij door eenzelfde klant niet noodzakelijk op dezelfde dag besteld te zijn. Vermeld voor elk paar films ook het aantal keer dat ze aan eenzelfde klant werden verhuurd, en sorteer het resultaat volgens dalende frequentie.

Hint: De Sakila databank werd ingeladen in een MySQL RDBMS en is nog steeds bevroegbaar via de phpMyAdmin webapplicatie (<http://microlab.ugent.be/phpMyAdmin/index.php>). De guest-account die werd aangemaakt voor deze databank heeft als loginnaam *guest* en als paswoord *guest*.

Opgave 4

Het kroonjuweel van de exentrieke museumdirecteur Gaston Lagaffe is een zaal met vijftien Picasso's uit zijn blauwe periode. Lagaffe staat erom bekend dat hij zijn collectie permanent in beweging houdt, waartoe hij speciaal voor zijn tentoonstelling een roterend algoritme heeft bedacht, dat op de volgende manier werkt.

De haakjes waaraan de schilderijen worden opgehangen zijn genummerd van 1 tot 15. Onder elk haakje met nummer i heeft de museumdirecteur een briefje op de muur geplakt waarop staat: "Hang dit schilderij op plaats x_i ". De getallen x_i ($i = 1, \dots, 15$) vormen hierbij een permutatie van de natuurlijke getallen tussen 1 en 15. De suppoosten hebben als taak om iedere ochtend de Picasso's van hun haakje te halen en de opdracht op het briefje uit te voeren, zodat de bezoekers dagelijks tegen een opgefriste tentoonstelling aankijken.

Stel bijvoorbeeld dat onder de haakjes voor de schilderijen briefjes hangen met de nummers

14, 6, 15, 3, 4, 10, 1, 12, 5, 2, 7, 13, 8, 11, 9

en dat we voor het gemak de schilderijen aangeven met de letters A tot en met O, in de volgorde waarin ze hangen bij de opening van het museum:

A, B, C, D, E, F, G, H, I, J, K, L, M, N, O

De volgende dag worden de schilderijen dus verhangen volgens het roterende algoritme, en komen ze op de volgende plaatsen terecht:

G, J, D, E, I, B, K, M, O, F, N, H, L, A, C

Gevraagd wordt om een VBA procedure `BeheerMuseum` te schrijven die voor een gegeven permutatie van getallen $1, 2, \dots, k$ bepaalt hoeveel dagen het duurt vooraleer de schilderijen terug op hun oorspronkelijke plaats terechtkomen, en die deze waarde vervolgens met een zinvolle tekst weergeeft in een boodschapvenster.

De procedure moet de getallen van de permutatie inlezen uit de cellen op de eerste rij van het huidige Excel werkblad, startend vanaf cel `A1`. Omdat de collectie van schilderijen nog kan groeien, wordt de waarde k bepaald door de eerste lege cel op de eerste rij, die het einde van de permutatiereeks aangeeft. Om het roterende algoritme voor de museumdirecteur wat aanschouwelijker te maken, moet de procedure een nieuw werkblad aanmaken. Dit nieuwe werkblad wordt vervolgens op elke rij gevuld met de positie van de schilderijen op de corresponderende dag, totdat de schilderijen terug hun oorspronkelijke volgorde hebben gekregen.

EXAMEN: Scriptingtalen

Prof. Dr. Peter Dawyndt
1^e Bachelor Informatica
groep 4-5

vrijdag 22-06-2007, 14:00h
academiejaar 2006-2007
eerste zittijd

Opgave 1

Voor deze opgave moet je een reeks Unix-commando's ontwikkelen voor het beheren en gebruiken van een eenvoudig nieuwsforum. Berichten die op het forum worden geplaatst, worden bewaard in een tekstbestand `forum.txt` in de home directory van de gebruiker die de scripts aanroept. Elk bericht begint met de nickname van de persoon die het bericht gestuurd heeft, een regel met het onderwerp van het bericht, meerdere regels met de inhoud van het bericht, en wordt afgesloten met een regel waarop het tijdstip van inzenden staat. Berichten worden van elkaar gescheiden door een regel die enkel de tekst `-o-` bevat. Een voorbeeld van het bestand `forum.txt` zou kunnen zijn

```
atanenba (Andy Tanenbaum)
Linux is obsolete
Linus, I still maintain the point that designing a monolithic kernel in
1991 is a fundamental error. Be thankful you are not my student. You
would not get a high grade for such a design :-)
30 Jan 92 13:44:34 GMT
-o-
lthorval (Linus Thorvalds)
Re: LINUX is obsolete
That's ok. Einstein got lousy grades in math and physics.
30 Jan 92 18:57:28 GMT
-o-
atanenba (Andy Tanenbaum)
Re: LINUX is obsolete
And Dan Quayle got low grades in political science. I think that there
are more Dan Quayles than Einsteins out there... ;-)
31 Jan 92 09:49:37 GMT
-o-
kthompso (Ken Thompson)
Re: LINUX is obsolete
I would generally agree that microkernels are probably the wave of
the future. However, it is in my opinion easier to implement a
monolithic kernel. It is also easier for it to turn into a mess in
a hurry as it is modified.
3 Feb 92 23:07:54 GMT
-o-
```

Gevraagd wordt

1. schrijf een **bash** shell script `postMessage` dat de nickname, het onderwerp en de tekst van een bericht opvraagt aan de gebruiker. De tekst van het bericht kan uit meerdere regels bestaan, en moet door de gebruiker worden afgesloten met een regel die enkel de tekst `-o-` bevat. Zorg voor een verzorgde interactie met de gebruiker, waarbij gepaste vragen worden gesteld door het shell script. De volgende sessie zorgt er bijvoorbeeld voor dat het eerste bericht wordt toegevoegd aan het bestand `forum.txt`:

```

> postMessage
nickname: atanenba
onderwerp: LINUX is obsolete
bericht:
Linus, I still maintain the point that designing a monolithic kernel in
1991 is a fundamental error. Be thankful you are not my student. You
would not get a high grade for such a design ;-)
-o-
>

```

Het shell script moet controleren of er reeds een bestand `forum.txt` bestaat in de home directory van de gebruiker. Indien dit niet het geval is, dan moet dit bestand door het script worden aangemaakt. Ook het toevoegen van de datum vóór het afsluiten van het bericht moet door het shell script gebeuren.

2. schrijf een **bash** shell script `getMessage <nr>` dat voor een gegeven argument `nr` het bericht op de corresponderende positie in `forum.txt` wegschrijft naar standaard uitvoer. Voor het uitknippen van het gepaste bericht mag je vrij kiezen of je daarvoor **sed** of **awk** gebruikt. Zorg ervoor dat het shell script controleert of het opgegeven argument binnen geldige onder- en bovengrenzen ligt, en dat er ook een gepaste opmerking wordt gegeven indien er geen argumenten worden meegegeven aan het script. De volgende sessie illustreert de functionaliteit die het shell script `getMessage` moet hebben voor het voorbeeldbestand `forum.txt` dat hoger werd gegeven:

```

> getMessage
Syntax: getMessage <nr>
> getMessage 10
Syntax: getMessage <nr>
Waarschuwing: de waarde van nr moet in het interval [1,4] gelegen zijn.
> getMessage 3
atanenba (Andy Tanenbaum)
Re: LINUX is obsolete
And Dan Quayle got low grades in political science. I think that there
are more Dan Quayles than Einsteins out there... ;-)
31 Jan 92 09:49:37 GMT
>

```

3. schrijf een **sed** script `formatMessage` waarmee de uitvoer van `getMessage` wordt opgemaakt zodat de datum voor de tekst komt te staan, namen voor de headers worden geplaatst en een lege regel verschijnt voor de tekst van het bericht. Kies daarbij voor een oplossing die optimaal gebruik maakt van de functionaliteit die in **sed** zit ingebakken. De volgende sessie illustreert het gebruik van dit nieuwe script:

```

> getMessage 3 | sed -f formatMessage
FROM: atanenba (Andy Tanenbaum)
SUBJECT: Re: LINUX is obsolete
DATE: 31 Jan 92 09:49:37 GMT

And Dan Quayle got low grades in political science. I think that there
are more Dan Quayles than Einsteins out there... ;-)
>

```

Opgave 2

Tijdens de federale verkiezingen van 10 juni 2007 hebben de overheidsdiensten "Binnenlandse Zaken" en "Informatie- en Communicatietechnologie" de krachten gebundeld om aan elke burger de mogelijkheid te bieden om rechtstreeks, onder de vorm van cijfers of grafieken, kennis te nemen van de uitslagen van de verkiezingen voor Kamer en Senaat. De cijferuitslagen (per partij) en de naamstemmen werden gedurende de ganse verkiezingsnacht voortdurend gepubliceerd en om de 10 minuten bijgewerkt. Het bestand `verkiezingen2007.zip` bevat de volledige en geldig verklaarde uitslagen voor de Belgische Senaat, en dit voor elke provincie afzonderlijk. De provinciale uitslagen zijn opgeslagen in afzonderlijke bestanden die naar de desbetreffende provincies (of kieskringen) genoemd werden. Hieronder staat een deel van het bestand `west-vlaanderen` als voorbeeld.

```
-- WVL: West Vlaanderen --

CD&V NVA          302.930 37.57% +6.05%
CD&V                -      -      -
N-VA              -      -      -
VLAAMS BELANG    118.492 14.69% +0.68%
VLAAMS BLOK      -      -      -
open vld         137.877 17.10% -6.12%
VLD              -      -      -
...
```

Dergelijke bestanden starten met een regel die een beschrijving geeft van de desbetreffende provincie. Zoals je kan zien in het voorbeeld bevat deze regel zowel een drie-letter afkorting van de provincie als de volledige naam van de provincie. De overige regels bevatten de uitslagen per politieke partij voor die provincie. De velden in deze regels worden gescheiden door tabs, en bevatten volgende informatie: i) naam van de partij, ii) absoluut aantal voorkeurstemmen, iii) relatief aantal voorkeurstemmen en iv) winst of verlies ten opzichte van de verkiezingen in 2003. Regels die informatie bevatten kunnen van elkaar worden gescheiden door één of meerdere lege regels.

Gevraagd wordt om een `awk`-script `Vote2007` te schrijven dat het mogelijk maakt om een samenvatting te maken van een bepaalde statistiek voor verschillende provincies. Het script moet kunnen aangeroepen worden op de Unix-commandolijn als `Vote2007 <files>`, waarbij een willekeurig aantal bestanden (opgemaakt zoals hierboven aangegeven) als argument aan het script kan worden doorgegeven. Lege regels in de bestanden worden door het script genegeerd. Regels met een koppelteken (-) in het tweede veld corresponderen met oude namen van politieke partijen en moeten genegeerd worden bij het maken van het overzicht. De volgende sessie illustreert de functionaliteit van het script

```
> Vote2007 antwerpen brussel-halle-vilvoorde henegouwen
Welke statistiek wil je zien ?
  1) absoluut aantal voorkeurstemmen
  2) relatief aantal voorkeurstemmen
  3) winst tov verkiezingen 2003

> 1

          APN      BHV      HGW
CD&V NVA    326.740  102.037
VLAAMS BELANG 264.432   75.476
open vld    187.922   86.668
...
PTB+                3.540   7.282
FORCE NATIONALE    1.951   5.968
CDF                3.864   3.027
```


Het script vraagt eerst aan de gebruiker welke statistiek hij of zij wil bekijken. Nadat een correcte keuze werd gemaakt, worden alle partijen die in minstens één van de argumentbestanden voorkwamen opgelijst in een eerste kolom, en volgen daarna kolommen met de verschillende statistieken per provincie. De eerste kolom wordt links uitgelijnd en is 17 lettertekens breed. De overige kolommen worden rechts uitgelijnd en zijn telkens 9 lettertekens breed.

Het script `Vote2007` moet voorts een functie `trim` bevatten waaraan een tekenreeks `str` kan worden doorgegeven. De functie verwijdert alle spaties vooraan en achteraan de tekenreeks `str`. Indien de tekenreeks hierna enkel nog bestaat uit een koppelteken (dit komt voor als een bepaalde waarde niet van toepassing is), dan moet worden die vervangen door een lege tekenreeks. De functie `trim` moet in het script worden aangeroepen voor een uniforme verwerking van de veldinhouden.

Opgave 3

De DVD-verhuurketen die gebruikt maakt van de Sakila databank is ondertussen online gegaan. Films kunnen nu ook via de website van de keten besteld worden, en worden dan naar het thuisadres opgestuurd. Hiervoor werden twee extra tabellen toegevoegd aan de oorspronkelijke databank.

De tabel `online_customer` houdt een unieke lijst met IP-adressen van de klanten bij in de kolom `IP_ADDRESS`, die de primaire sleutel van de tabel vormt. Voor elk IP-adres bevatten de records voorts ook nog de hostname (`HOST_NAME`), het land (`COUNTRY`), de landcode (`COUNTRY_CODE`), de regio (`REGION`), de stad (`CITY`), de postcode (`POSTAL_ADDRESS`), de telefoon-prefix (`CALLING_CODE`), en de lengte- en breedtegraden van de geografische locatie waar de klant zich bevindt (`LONGITUDE` en `LATITUDE`). De primaire sleutel `IP_ADDRESS` van de tabel `online_customer` werd ook als vreemde sleutel toegevoegd aan de tabel `customer`, waarmee wordt uitgedrukt dat online klanten ook gewone klanten zijn (gewone klanten zijn niet noodzakelijk ook online klanten).

Voor elke online bestelling wordt een record toegevoegd aan de tabel `online_order`. Hierbij worden volgende gegevens geregistreerd: het IP-adres van de machine waarop de bestelling gebeurde (`IP_ADDRESS`), de unieke sleutel van de bestelde film (`FILM_ID`) en de dag waarop de bestelling gebeurde (`ORDER_TIME`).

Gevraagd wordt om SQL zoekopdrachten te formuleren die een antwoord bieden op de volgende vragen. Zorg er telkens voor dat de kolommen van de resulterende tabel een zinvolle naam krijgen. Geef op je antwoordblad ook aan hoeveel records de resultatentabel bevat.

1. Bereken de gemiddelde lengte- en breedtegraad per land op basis van de gegevens in het online klantenbestand. Sorteert hierbij de landen alfabetisch.
2. Geef de voornaam, de familienaam en het land van herkomst van alle klanten die ooit een online bestelling geplaatst hebben van een film waarvoor het woord `DARK` in de titel voorkomt. Sorteert het resultaat eerst per land, en daarna per familienaam.
3. Geef een overzicht van de landen van herkomst van alle klanten die zowel in de huidige maand (juni 2007) als in de vorige maand (mei 2007) een film online hebben besteld. Geef daarbij voor elk land aan hoeveel klanten in beide maanden een film hebben besteld, en sorteert de landen volgens dalende frequentie.
4. Geef de naam en het land van herkomst van elk paar klanten dat minstens 40 keer eenzelfde film online besteld heeft. De klanten hoeven daarbij eenzelfde film niet noodzakelijk op dezelfde dag besteld te hebben. Vermeld voor elk paar klanten ook het aantal films dat ze gemeenschappelijk besteld hebben, en sorteert het resultaat volgens dalende frequentie.

Hint: De Sakila databank werd ingeladen in een MySQL RDBMS en is nog steeds bevroegbaar via de phpMyAdmin webapplicatie (<http://microlab.ugent.be/phpMyAdmin/index.php>). De guest-account die werd aangemaakt voor deze databank heeft als loginnaam `guest` en als paswoord `guest`.

Opgave 4

Scripting wordt heel vaak gebruikt om gegevens van het ene formaat naar een ander formaat te vertalen. Gevraagd wordt om een **VBA** procedure `Excel2LaTeX` te schrijven die de waarden extraheert uit de rechthoekige selectie in het actieve werkblad van Excel, en omzet naar een tabel in $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ formaat. De procedure moet zelf de positie en de grootte van het geselecteerde gebied bepalen. Hierbij wordt via boodschapvensters aan de gebruiker gevraagd of de bovenste rij een hoofding voorstelt en of de linkerkolom een hoofding voorstelt. Hoofdingen moeten worden opgemaakt met vette tekst. Het resultaat kan je wegschrijven naar het venster Direct, waarna de gebruiker het zelf kan kopiëren naar het klembord. Dit is echter niet de meest ideale oplossing. Als uitbreiding kan je het resultaat ook rechtstreeks proberen weg te schrijven naar het klembord (beschrijf ook wat er nodig is om tot deze oplossing te komen).

Beschouwen we bij wijze van illustratie de volgende twee voorbeelden:

	A	B	C	D	E	F	G	H	I	J	K
1		1	2	3	4	5	6	7	8	9	10
2	1	1	2	3	4	5	6	7	8	9	10
3	2	2	4	6	8	10	12	14	16	18	20
4	3	3	6	9	12	15	18	21	24	27	30
5	4	4	8	12	16	20	24	28	32	36	40
6	5	5	10	15	20	25	30	35	40	45	50
7	6	6	12	18	24	30	36	42	48	54	60
8	7	7	14	21	28	35	42	49	56	63	70
9	8	8	16	24	32	40	48	56	64	72	80
10	9	9	18	27	36	45	54	63	72	81	90
11	10	10	20	30	40	50	60	70	80	90	100

	A	B	C	D	E	F	G	H	I	J	K
1		1	2	3	4	5	6	7	8	9	10
2	1	1	2	3	4	5	6	7	8	9	10
3	2	2	4	6	8	10	12	14	16	18	20
4	3	3	6	9	12	15	18	21	24	27	30
5	4	4	8	12	16	20	24	28	32	36	40
6	5	5	10	15	20	25	30	35	40	45	50
7	6	6	12	18	24	30	36	42	48	54	60
8	7	7	14	21	28	35	42	49	56	63	70
9	8	8	16	24	32	40	48	56	64	72	80
10	9	9	18	27	36	45	54	63	72	81	90
11	10	10	20	30	40	50	60	70	80	90	100

Indien voor het linkse voorbeeld wordt aangegeven dat zowel de bovenste rij als de linkerkolom een hoofding vormen, en indien voor het rechtse voorbeeld wordt aangegeven dat enkel de bovenste rij een hoofding vormt, dan genereert de procedure `Excel2LaTeX` de gevraagde $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ code die respectievelijk de volgende tabellen voorstelt:

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

6	7	8	9
6	7	8	9
12	14	16	18
18	21	24	27
24	28	32	36
30	35	40	45
36	42	48	54
42	49	56	63
48	56	64	72
54	63	72	81
60	70	80	90