

巨資二 A 鄭雅綿 08170120

P1

1. 程式碼：ANS11 = FILTER users BY (state=='NY');

```
(Dan_Ehle,Stella Jones,NY)
(SuiGenerisHaHa,Juan Stewart,NY)
(laurennnwright,Harper Young,NY)
(danny4466,Christian Walker,NY)
(SESSIINAT,Alexis Thompson,NY)
(99_russell,Connor Johnson,NY)
(Sweetie_li_ciou,Ariana White,NY)
(the_niz_wiz,Brianna Thompson,NY)
(ANChums,Mackenzie Brooks,NY)
(kingling43,Zoe Wood,NY)
(Briley_M_,Caroline James,NY)
(CookieLokss,Andrew Thompson,NY)
(Keelinelise98,Leah Jackson,NY)
(candicerae08,London Watson,NY)
(sippey,Daniel Murphy,NY)
(meeshe,Layla Williams,NY)
```

- 程式碼：ANS1 = FOREACH ANS11 GENERATE login;

```
(xoalliyah)
(Dan_Ehle)
(SuiGenerisHaHa)
(laurennnwright)
(danny4466)
(SESSIINAT)
(99_russell)
(Sweetie_li_ciou)
(the_niz_wiz)
(ANChums)
(kingling43)
(Briley_M_)
(CookieLokss)
(Keelinelise98)
(candicerae08)
(sippey)
(meeshe)
```

2. 程式碼：ANS2 = FILTER tweets BY (content matches '*love.*');

```
(396124473390563328,"I love you Alicia! RT @dynamothory @GingerFC Quiet Sounder fan ,Imorequiethright)
(396124474736508928,"I love soccer, & soccer loves me ok.")
(396124475223060480,"Love my bf for being with me thru good and bad. He always makes me feel loved and k
now that he is always there.",tushi_20)
(396124475806474240,"Grouplove's new album is divine. ???",laylamariex3)
(396124481703661568,"Please love me. @Harry_Styles",jennaadams9)
(396124485390049280,"Yeah my boyfriend is amazing even when he should be super mad at me he still smiles
kisses me and tells me he loves me #sorrybabe #heskeeper",Justina_Marie84)
(396124497021247488,"Also keep snap chatting them and texting them. ??? that's lovely ? I love you so m
uch right now.",shmemmabee)
(396124502029250560,"I love @stefferranis ???",Scott_larocca)
(396124522165710848,"Y'all I'm such a animal lover. ???",MissSarah13)
(396124523621130240,"I love booty ???",RayPerez44)
(396124527329288192,"I love your attempts to try and work things out. #givetoeget",JustissMariee)
(396124529703276544,"#dolphinswin #godolphins #CINvsMIA #ilovefootball #overtime @ Sun Life Stadium ht
p://t.co/ajpBkebx2",zeekay88)
(396124529858072576,"With all these painted skulls on everyone's faces it looks like everyone is ready t
or d'a de Los muertos. Glad to know everyone loves it?",hannahhaun)
(396124533268422656,"I love you sooo much. @Harry_Styles",jennaadams9)
(396124535369396224,"Vampire diaries with @EliseRistau #justlikeoldtimes #loveher #suitemate",jhadd54)
(396124541208260608,"@reIatabie: can you imagine if someone sent you a list of all the reasons why they
love you.?",kputka)
(396124545809416192,"Wow I love everclear",jordannjayy)
(396124547734183936,"There is no fear now Let go and just be free I will love you unconditionally",Cassa
ndraCan_tu)
(396124551530037248,"Friend of a friend the one and only.. Will Ferrell! How can you not love him? http
://t.co/9phkxNHn116" Zachaouki8)
```

3. 程式碼：ANS3 = GROUP users BY state;

```
(VA,{(aKuulLuxuryTax,Natalie Peterson,VA),(ponchov69,William Taylor,VA),(iTweet4kicks_,Nolan Barnes,VA),
(Chyna_Munoz,Piper Stewart,VA),(katapaige5,John Sanders,VA),(jazalaniz15,John Evans,VA),(FUCKINNCompton,
Nathan Anderson,VA),(FlyGuy_Two,Gavin Watson,VA),(Dlbezner,Alyssa Washington,VA),(IBe_MissingYou,Cameron
Gonzalez,VA),(justinrenfroee,Evelyn Turner,VA),(FlippinUtheBIRD,Colton Harris,VA),(kid_kyra,Jacob Nelson
,VA),(buggabo0612,Brody Richardson,VA),(raley_wright, Lucy Hernandez,VA),(BryRB1277,Eva Edwards,VA),(carn
ifexkitty,Mason Thompson,VA),(sailorkathryn,Victoria Phillips,VA),(shanelevan,Kennedy Hernandez,VA)})
(VI,{(fondofanthonyq,Ashley Mitchell,VI),(TaeJnes,Parker Wilson,VI),(reesewong,Stella Gonzalez,VI),(Gal
lito_Zayy,Autumn Barnes,VI),(VanCanucks,Morgan Lee,VI),(tmt2011,Claire Harris,VI),(Ctweetinn,Faith Thomp
son,VI),(DaveLewissports,Dylan Lee,VI),(BixTweet062,Eva Davis,VI),(Mattyg61,Evelyn Scott,VI),(Limp_N_Zom
bie,Christopher Adams,VI),(AlexaDiana21,Sebastian Campbell,VI),(smorganwhite,Alex Williams,VI),(lukepaxt
91,Jackson Henderson,VI),(alissaliane,Aaliyah Taylor,VI)})
(VT,{(espi_olvera,Gabriella Hill,VT),(kelseymariee004,Thomas White,VT),(Aasholaay,William Morris,VT),(j_
ciay,Grace Wilson,VT),(savannnuhh18,Alexander Perez,VT),(tylermunford,Alyssa Bell,VT),(BsSoxPatsCelts,An
na Reed,VT),(Dre_Sosa,Brandon Howard,VT),(samdacat_,Grace Young,VT),(PunchZoneAris,Trinity Price,VT),(M
arkJrSays,Josiah Brown,VT),(Kimberley_Klein,Sofia Moore,VT),(Decemeber2013,Lillian Lee,VT),(ram_ash14,Am
elia Sanders,VT),(Smileynicki_,Jayden Phillips,VT),(erin_colombe,Ian Allen,VT),(Usnlen,Isaac Lee,VT),(dy
lanfurtado1,Wyatt Ross,VT),(lmgoshert93,Madeline Campbell,VT)})
(WA,{(Mbunch_Badger11,Melanie Young,WA),(adrian__209,Isabella Lee,WA),(Elena_Panaritis,Logan Smith,WA),
(Courty_Cat,Lillian Bell,WA),(diana22quizoz,Gavin Cooper,WA),(chelsiereynolds,Zoey Miller,WA),(Najely_ba
beyy,Mia Parker,WA),(infinitepharaoh,Nolan King,WA),(iLoveJMA,Jace Sanchez,WA),(SamanthaDanie15,Jack How
ard,WA),(TheReal_HTown,Natalie Alexander,WA),(Orie0,Ellie Harris,WA),(TylerComeback,Austin Gonzalez,WA)}
)
(WI,{(CUNTMAFIA,Savannah Wright,WI),(ChrisBenavides5,Matthew Campbell,WI),(JacobAGold,Andrew Gray,WI),(v
mrdll,Nolan Rodriguez,WI),(jessyj106,Connor Harris,WI),(BigBankBoosie,Parker Peterson,WI),(Too_Easyyyy,Pe
yton Sanders,WI),(WORLDWARJENN,Christian Lewis,WI),(Pim_BYExit_SwR,Sarah Baker,WI),(JodiRisden,Charles H
ayes,WI),(megtuusa,Brooklyn Peterson,WI),(ShygirlNathalie,Kaylee Rodriguez,WI),(mmcoates,Naomi Harris,WI
),(jlampee,Mason Walker,WI),(MsJaimy_G,Autumn Powell,WI),(Illmatic2,Sophia Peterson,WI),(ajreedxo,Anna J
ackson,WI),(MrLucrid,Serenity Wilson,WI),(tez_the_prez,Zoe Cooper,WI)})
(WV,{(walker_delaney,Zoey Hall,WV),(cmillzy12,Aubree Wood,WV),(Crain_Bird,Genesis Jones,WV),(EmmyJohnny1
3,Anna Ward,WV),(BixTweet292,Autumn Gray,WV),(Eric_Patry_,Bentley Gray,WV),(RayPerez44,Olivia Carter,WV)
,(paulinalynn40,Brayden Thomas,WV),(teti_ruiz,Benjamin Garcia,WV),(zeekay88,Sebastian Sanchez,WV),(miran
daaloves,Piper Lewis,WV),(GavinCampellone,Isaiah Perry,WV),(TeeeeeeJaaaaay,Gavin Jones,WV),(juice4u,Kay
lee Griffin,WV),(Johnzyyy,Joseph Diaz,WV),(mydaughterworld,Sofia Robinson,WV),(luketc1,Jaxon Morgan,WV),
(DonaldPrice11,Justin Scott,WV),(gabbyxdearest,Aaron Smith,WV),(brittanylyn13,Lydia Rivera,WV)})
(WY,{(Yazmean_,Autumn Young,WY),(Kennif_Kennif,Luke James,WY),(asialeemarie_,Kennedy Ross,WY),(Appaloosa
Guy,Ava Jackson,WY),(Big_DADDY_42,Landon Roberts,WY),(baltimore311,Angel Scott,WY),(thepink9975,Ryder He
rnandez,WY),(PangTJ,Autumn Peterson,WY),(AmberMcIarenxo,London Moore,WY),(HusDiggity,Logan Flores,WY),(_
EJEEZY,Joseph Brown,WY),(BixTweet011,Lucas Baker,WY),(doyalllllllllll,Jose Morgan,WY),(heardrwt,Isaiah Ro
gers,WY),(EcoMartinMtz,Parker Robinson,WY)})
```

4. 程式碼：users_join_tweets = JOIN users BY name LEFT, tweets BY user;

```
(ayslinzuniga,Savannah Peterson,OR,,,)
(zeekay88,Sebastian Sanchez,WV,,,)
(ThatGirl_Normaa,Victoria Campbell,CO,,,)
(sailorkathryn,Victoria Phillips,VA,,,)
(rayterrabossy,William Henderson,KY,,,)
(whhillsweather,Xavier Washington,DC,,,)
(SonOfAnAngel1,Abigail Richardson,ND,,,)
(C_Olivier_12,Alexander Martinez,TX,,,)
(BabyGotBackes,Alexandra Campbell,UT,,,)
(RigoFlores_XO,Allison Richardson,ME,,,)
(JustinTimeis410,Brayden Richardson,NH,,,)
(alexiselaine15,Christopher Butler,DE,,,)
(Jesse_Naylor,Christopher Foster,ME,,,)
(_samayoa_,Christopher Parker,AS,,,)
(DaniWhitee,Gabriel Washington,NH,,,)
(d0nstatus,Mackenzie Mitchell,MH,,,)
(U_HeartForevA_I,Mackenzie Peterson,MD,,,)
(AlexaDiana21,Sebastian Campbell,VI,,,)
(pdxsherrifflog,Sebastian Phillips,OK,,,)
(XavierMitchell_,Serenity Patterson,GU,,,)
(spencerkelley16,Serenity Rodriguez,PW,,,)
(KevinMoss924,Alexander Henderson,UT,,,)
(RIGHT_VISION,Caroline Richardson,ID,,,)
(Kareniracheta,Christopher Coleman,SD,,,)
(PowellOHWeather,Christopher Stewart,FL,,,)
(Kalvin_Ghostz,Katherine Rodriguez,NJ,,,)
(JordynNykoleee,Nathaniel Patterson,IA,,,)
(Lognoll30,Serenity Washington,AS,,,)
(itsmoorebaby,Christopher Richardson,NH,,,)

```

程式碼：name_group = GROUP users_join_tweets BY name;

```
(Nicholas Williams,{(zorilla87,Nicholas Williams,AZ,,)})
(Samantha Robinson,{(JustCallMeMarbs,Samantha Robinson,AL,,)})
(Savannah Peterson,{(ayslinzuniga,Savannah Peterson,OR,,)})
(Sebastian Sanchez,{(zeekay88,Sebastian Sanchez,WV,,)})
(Victoria Campbell,{(ThatGirl_Normaa,Victoria Campbell,CO,,)})
(Victoria Phillips,{(sailorkathryn,Victoria Phillips,VA,,)})
(William Henderson,{(rayterrabossy,William Henderson,KY,,)})
(Xavier Washington,{(whhillsweather,Xavier Washington,DC,,)})
(Abigail Richardson,{(SonOfAnAngel1,Abigail Richardson,ND,,)})
(Alexander Martinez,{(C_Olivier_12,Alexander Martinez,TX,,)})
(Alexandra Campbell,{(BabyGotBackes,Alexandra Campbell,UT,,)})
(Allison Richardson,{(RigoFlores_X0,Allison Richardson,ME,,)})
(Brayden Richardson,{(JustinTimeis410,Brayden Richardson,NH,,)})
(Christopher Butler,{(alexiselaine15,Christopher Butler,DE,,)})
(Christopher Foster,{(Jesse_Naylor,Christopher Foster,ME,,)})
(Christopher Parker,{(_samayoa_,Christopher Parker,AS,,)})
(Gabriel Washington,{(DaniWhitee,Gabriel Washington,NH,,)})
(Mackenzie Mitchell,{(d0nstatus,Mackenzie Mitchell,MH,,)})
(Mackenzie Peterson,{(U_HeartForevA_I,Mackenzie Peterson,MD,,)})
(Sebastian Campbell,{(AlexaDiana21,Sebastian Campbell,VI,,)})
(Sebastian Phillips,{(pdxsherrifflog,Sebastian Phillips,OK,,)})
(Serenity Patterson,{(XavierMitchell_,Serenity Patterson,GU,,)})
(Serenity Rodriguez,{(spencerkelley16,Serenity Rodriguez,PW,,)})
(Alexander Henderson,{(KevinMoss924,Alexander Henderson,UT,,)})
(Caroline Richardson,{(RIGHT_VISION,Caroline Richardson,ID,,)})
(Christopher Coleman,{(Kareniracheta,Christopher Coleman,SD,,)})
(Christopher Stewart,{(PowellOHWeather,Christopher Stewart,FL,,)})
(Katherine Rodriguez,{(Kalvin_GhosTz,Katherine Rodriguez,NJ,,)})
(Nathaniel Patterson,{(JordynNykoleee,Nathaniel Patterson,IA,,)})
(Serenity Washington,{(Lognoll30,Serenity Washington,AS,,)})
(Christopher Richardson,{(itsmoorebaby,Christopher Richardson,NH,,)})
```

程式碼：number_of_tweets = FOREACH name_group GENERATE group, COUNT(users_join_tweets.name)AS number;

```
(Gabriel Washington,1)
(Mackenzie Mitchell,1)
(Mackenzie Peterson,1)
(Sebastian Campbell,1)
(Sebastian Phillips,1)
(Serenity Patterson,1)
(Serenity Rodriguez,1)
(Alexander Henderson,1)
(Caroline Richardson,1)
(Christopher Coleman,1)
(Christopher Stewart,1)
(Katherine Rodriguez,1)
(Nathaniel Patterson,1)
(Serenity Washington,1)
(Christopher Richardson,1)
```

程式碼：tweets_include_favorite_ordered_by_id = ORDER number_of_tweets BY number;


```
(Ava Cox,1)
(Michael Bryant,2)
(Harper Washington,2)
(Trinity Price,2)
(Claire Gonzales,2)
(Evelyn Diaz,2)
(Mia Perez,2)
(Hailey Ross,2)
(Easton Cooper,2)
(Charlotte Cook,2)
(Claire Adams,2)
(Bailey Baker,2)
(Arianna Allen,2)
(Jocelyn Cooper,2)
(London Watson,2)
(Caroline James,2)
```

5. 程式碼：INNER_JOIN = JOIN users BY login, tweets BY user;

```
((rachelfischbach,Riley Reed,FM,396124509838655488,"@ShannonEDawes get used to it ;)",rachelfischbach)
((rachelfischbach,Riley Reed,FM,396124453823721472,"I just wanna let the world know I'm probably Instagram famous and vine famous from all the things I was in for #boorito tonight",rachelfischbach)
((sofia_olloqui09,Violet Morris,MA,396124505086509056,"#halloween2013 rap??? http://t.co/hrw1LFvKE6",sofia_olloqui09)
((spencerkelley16,Serenity Rodriguez,PW,396124494072647680,"Even white people can construct something beautiful! #THWTR http://t.co/b4pfb9oEgx",spencerkelley16)
((uu_Should_Adore,Jaxon Evans,SC,396124470785880064,"If he don't hurry TF up and get here ik something!",uu_Should_Adore)
((x_moneyMONROE_x,Lucas Barnes,UT,396124515081523201,"@BlaccPringles Lol my sister kept her for me but let me know you wanna get her fat ass",x_moneyMONROE_x)
((xosamanthalloyd,Claire Gonzales,GU,396124538070519808,"I just killed Andrew by starving him to death on sims. #payback",xosamanthalloyd)
((yvonne_suarez12,John Brown,AR,396124475181133824,"My feet hurt from all that walking & I didn't get candy at all ?",yvonne_suarez12))
```

程式碼：B = GROUP INNER_JOIN BY (login, name);

```
((rachelfischbach,Riley Reed),{(rachelfischbach,Riley Reed,FM,396124509838655488,"@ShannonEDawes get used to it ;)",rachelfischbach),(rachelfischbach,Riley Reed,FM,396124453823721472,"I just wanna let the world know I'm probably Instagram famous and vine famous from all the things I was in for #boorito tonight",rachelfischbach)})
((sofia_olloqui09,Violet Morris),{(sofia_olloqui09,Violet Morris,MA,396124505086509056,"#halloween2013 rap??? http://t.co/hrw1LFvKE6",sofia_olloqui09)})
((spencerkelley16,Serenity Rodriguez),{(spencerkelley16,Serenity Rodriguez,PW,396124494072647680,"Even white people can construct something beautiful! #THWTR http://t.co/b4pfb9oEgx",spencerkelley16)})
((uu_Should_Adore,Jaxon Evans),{(uu_Should_Adore,Jaxon Evans,SC,396124470785880064,"If he don't hurry TF up and get here ik something!",uu_Should_Adore)})
((x_moneyMONROE_x,Lucas Barnes),{(x_moneyMONROE_x,Lucas Barnes,UT,396124515081523201,"@BlaccPringles Lol my sister kept her for me but let me know you wanna get her fat ass",x_moneyMONROE_x)})
((xosamanthalloyd,Claire Gonzales),{(xosamanthalloyd,Claire Gonzales,GU,396124538070519808,"I just killed Andrew by starving him to death on sims. #payback",xosamanthalloyd)})
((yvonne_suarez12,John Brown),{(yvonne_suarez12,John Brown,AR,396124475181133824,"My feet hurt from all that walking & I didn't get candy at all ?",yvonne_suarez12)})
```

程式碼：ANS51= FOREACH B GENERATE group, COUNT(INNER_JOIN.id) AS number;

```
((grizzstraight44,Natalie Rodriguez),1)
((hausofhannnerrz,Caleb Coleman),1)
((jessicamurrayyy,Nathaniel Griffin),1)
((jorge_aenriquez,Easton Wright),1)
((kelseymariee004,Thomas White),1)
((laurenemilydean,Alexander Jenkins),2)
((lolololollollll,Kayden Lee),2)
((mydaughterworld,Sofia Robinson),1)
((prfessorbigsock,Luis Robinson),1)
((rachelfischbach,Riley Reed),2)
((sofia_olloqui09,Violet Morris),1)
((spencerkelley16,Serenity Rodriguez),1)
((uu_Should_Adore,Jaxon Evans),1)
((x_moneyMONROE_x,Lucas Barnes),1)
((xosamanthalloyd,Claire Gonzales),1)
((yvonne_suarez12,John Brown),1)
```

```
grunt> describe ANS51
ANS51: {group: (users::login: chararray,users::name: chararray),number: long}
```

程式碼：SubData = FILTER ANS51 BY(number is not null AND number>0L);

```
((grizzstraight44,Natalie Rodriguez),1)
((hausofhannnerrz,Caleb Coleman),1)
((jessicamurrayyy,Nathaniel Griffin),1)
((jorge_aenriquez,Easton Wright),1)
((kelseymariee004,Thomas White),1)
((laurenemilydean,Alexander Jenkins),2)
((lolololollollll,Kayden Lee),2)
((mydaughterworld,Sofia Robinson),1)
((prfessorbigsock,Luis Robinson),1)
((rachelfischbach,Riley Reed),2)
((sofia_olloqui09,Violet Morris),1)
((spencerkelley16,Serenity Rodriguez),1)
((uu_Should_Adore,Jaxon Evans),1)
((x_moneyMONROE_x,Lucas Barnes),1)
((xosamanthalloyd,Claire Gonzales),1)
((yvonne_suarez12,John Brown),1)
```

程式碼：STORE SubData INTO 'a.result' USING PigStorage(',');

```
2021-05-31 05:59:04,811 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,813 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,814 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,827 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,835 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,836 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Me
trics with processName=JobTracker, sessionId= - already initialized
2021-05-31 05:59:04,841 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapRed
uceLauncher - Success!
```

```
grunt> root@dc5878ce49b9:/# ls
a.result  dev    lib    mnt          pig_1622440496317.log  run    sys    users.csv
bin       etc    lib64  opt          proc                  sbin   tmp    usr
boot     home  media  pig_1622437508026.log  root                  srv    tweets.csv  var
```

程式碼：cat ./a.result/*

```
(grizzstraight44,Natalie Rodriguez),1
(hausofhannnerrz,Caleb Coleman),1
(jessicamurrayyy,Nathaniel Griffin),1
(jorge_aenriquez,Easton Wright),1
(kelseymariee004,Thomas White),1
(laurenemilydean,Alexander Jenkins),2
(lolololollollll,Kayden Lee),2
(mydaughterworld,Sofia Robinson),1
(prfessorbigsock,Luis Robinson),1
(rachelfischbach,Riley Reed),2
(sofia_olloqui09,Violet Morris),1
(spencerkelley16,Serenity Rodriguez),1
(uu_Should_Adore,Jaxon Evans),1
(x_moneyMONROE_x,Lucas Barnes),1
(xosamanthalloyd,Claire Gonzales),1
(yvonne_suarez12,John Brown),1
```