VISUALIZACION DE DATOS: PEC2

Autor: Francisco Ramirez Fominaya

Noviembre 2019

Contents

Planteamiento y carga de datos.

```
# Cargamos el juego de datos
file_csv <- "pax_all_agreements_data.xlsx"
df1 <- read_xlsx("pax_all_agreements_data.xlsx")
#
str(df1)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    1789 obs. of   266 variables:
##  $ Con         : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...
##  $ Contp       : chr  "Government" "Government" "Government" "Government" ...
##  $ PP          : num  2 2 2 2 2 2 2 2 2 2 ...
##  $ PPName      : chr  "Afghanistan: 2000s Post-intervention process" "Afghanistan: 2000s Post-interve
##  $ Reg         : chr  "Europe and Eurasia" "Europe and Eurasia" "Europe and Eurasia" "Europe and Eura
##  $ AgtId       : num  1739 1923 864 848 849 ...
##  $ Agt         : chr  "Agreement between the Islamic Republic of Afghanistan and Hizb-e-Islami" "Agr
##  $ Dat         : chr  "2016-09-22" "2014-09-21" "2012-07-08" "2011-12-05" ...
##  $ Status      : chr  "Multiparty signed/agreed" "Multiparty signed/agreed" "Multiparty signed/agree
##  $ Lgt         : num  4 4 14 7 10 9 4 7 4 3 ...
##  $ N_characters: num  10203 11110 39746 24106 21299 ...
##  $ Agtp        : chr  "Intra" "Intra" "InterIntra" "InterIntra" ...
##  $ Stage       : chr  "SubPar" "Imp" "Imp" "Pre" ...
##  $ StageSub    : chr  "MultIss" "ExtSub" "ExtSub" "Prin" ...
##  $ Part        : chr  "President of the IRA Mohammad Ashraf Ghani, chairman of HIA delegation Gulbudc
##  $ ThrdPart    : chr  NA "The foregoing signatures were witnessed by:\nH.E. Jan Kubis, Special Repres
##  $ OthAgr      : chr  NA "Page 3, D. Creation of the position of leader of the runner-up team\nIn li
##  $ Loc1ISO     : chr  "AFG" "AFG" "AFG" "AFG" ...
##  $ Loc2ISO     : chr  NA NA NA NA ...
##  $ Loc1GWNO    : num  700 700 700 700 700 700 700 700 700 700 ...
##  $ Loc2GWNO    : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ UcdpCon     : chr  "333" "333" "333" "333" ...
##  $ UcdpAgr     : chr  "1488" "NA" NA NA ...
##  $ PamAgr      : chr  NA "NA" NA NA ...
##  $ CowWar      : chr  "851" "225" "225" "225" ...
##  $ GCh         : num  0 2 1 1 0 2 1 0 0 0 ...
##  $ GChRhet     : num  0 1 1 1 0 0 1 0 0 0 ...
##  $ GChAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ GChSubs     : num  0 1 0 0 0 1 0 0 0 0 ...
##  $ GChOth      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ GDis        : num  0 2 0 0 0 0 0 0 0 0 ...
##  $ GDisRhet    : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ GDisAntid   : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
## $ GDisSubs     : num  0 1 0 0 0 0 0 0 0 0 ...
## $ GDisOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GAge         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GAgeRhet     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GAgeAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GAgeSubs     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GAgeOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GMig         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GMigRhet     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GMigAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GMigSubs     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GMigOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GRa          : num  1 0 0 0 0 0 0 0 0 0 ...
## $ GRaRhet      : num  1 0 0 0 0 0 0 0 0 0 ...
## $ GRaAntid     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GRaSubs      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GRaOth       : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GRe          : num  0 0 0 0 0 0 2 0 0 0 ...
## $ GReRhet      : num  0 0 0 0 0 0 1 0 0 0 ...
## $ GReAntid     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GReSubs      : num  0 0 0 0 0 0 1 0 0 0 ...
## $ GReOth       : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GInd         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GIndRhet     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GIndAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GIndSubs     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GIndOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GOth         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GOthRhet     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GOthAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GOthSubs     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GOthOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GRef         : num  2 0 1 1 1 1 1 1 0 0 ...
## $ GRefRhet     : num  0 0 1 1 1 1 0 1 0 0 ...
## $ GRefSubs     : num  1 0 0 0 0 0 0 0 0 0 ...
## $ GRefOth      : num  0 0 0 0 0 0 1 0 0 0 ...
## $ GSoc         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GSocRhet     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GSocAntid    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GSocSubs     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GSocOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeWom        : num  1 1 1 1 1 1 1 1 1 0 ...
## $ GeMe         : num  0 0 1 1 0 1 0 1 1 0 ...
## $ GeMeNu       : num  0 0 1 1 0 1 0 1 1 0 ...
## $ GeMeOth      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeSo         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeLgbti      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeLgbtiPos   : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeLgbtiNeg   : num  0 0 0 0 0 0 0 0 0 0 ...
## $ GeFa         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ StDef        : num  1 0 1 1 1 1 0 1 1 1 ...
## $ StGen        : num  0 0 1 1 1 1 0 1 1 1 ...
## $ StCon        : num  0 0 0 0 0 0 0 0 0 0 ...
## $ StSd         : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
##  $ StRef      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ StSym      : num  1 0 0 0 0 0 0 0 0 0 ...
##  $ StInd      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ StUni      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ StBor      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ StXbor     : num  0 0 1 0 0 0 0 0 0 0 ...
##  $ Pol        : num  1 1 0 0 0 0 0 0 1 0 ...
##  $ PolGen     : num  1 0 0 0 0 0 0 0 1 0 ...
##  $ PolNewInd  : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ PolNewTemp : num  0 1 0 0 0 0 0 0 0 0 ...
##  $ ConRen     : num  0 0 2 2 0 0 0 2 0 0 ...
##  $ Cons       : num  0 1 0 0 0 0 0 0 0 0 ...
##   [list output truncated]
```

```
# Comprabaciones para limpieza de datos
# Análisis de valores nulos
colSums(is.na(df1))
```

```
##        Con        Contp          PP       PPName         Reg
##          0            0           7            7           0
##      AgtId          Agt         Dat       Status         Lgt
##          0            0           0            0           0
## N_characters         Agtp       Stage     StageSub        Part
##          0            0           0           29          16
##    ThrdPart       OthAgr      Loc1ISO      Loc2ISO     Loc1GWNO
##        871          852          44         1557          20
##   Loc2GWNO       UcdpCon      UcdpAgr       PamAgr      CowWar
##       1557           39         1304         1553         514
##        GCh       GChRhet     GChAntid      GChSubs       GChOth
##          0            0           0            0           0
##       GDis       GDisRhet     GDisAntid     GDisSubs      GDisOth
##          0            0           0            0           0
##       GAge       GAgeRhet     GAgeAntid     GAgeSubs      GAgeOth
##          0            0           0            0           0
##       GMig       GMigRhet     GMigAntid     GMigSubs      GMigOth
##          0            0           0            0           0
##        GRa       GRaRhet      GRaAntid      GRaSubs       GRaOth
##          0            0           0            0           0
##        GRe       GReRhet      GReAntid      GReSubs       GReOth
##          0            0           0            0           0
##       GInd       GIndRhet     GIndAntid     GIndSubs      GIndOth
##          0            0           0            0           0
##       GOth       GOthRhet     GOthAntid     GOthSubs      GOthOth
##          0            0           0            0           0
##       GRef       GRefRhet     GRefSubs      GRefOth       GSoc
##          0            0           0            0           0
##    GSocRhet     GSocAntid     GSocSubs      GSocOth       GeWom
##          0            0           0            0           0
##       GeMe       GeMeNu       GeMeOth       GeSo        GeLgbti
##          0            0           0            0           0
##  GeLgbtiPos    GeLgbtiNeg      GeFa        StDef        StGen
##          0            0           0            0           0
##      StCon         StSd        StRef        StSym        StInd
##          0            0           0            0           0
##      StUni        StBor       StXbor        Pol         PolGen
```

```
##               0               0               0               0               0
##       PolNewInd      PolNewTemp          ConRen            Cons             Ele
##               0               0               0               0               0
##        ElecComm          PolPar     PolParTrans        PolParOth           Civso
##               0               0               0               0               0
##            Tral           Pubad           Polps           PpsSt          PpsSub
##               0               0               0               0               0
##           PpsEx          PpsOro         PpsOthPr          PpsVet          PpsAut
##               0               0               0               0               0
##          PpsInt          PpsOth           Terps          TpsSub          TpsLoc
##               0               0               0               0               0
##          TpsAut          TpsOth             Eps          EpsRes          EpsFis
##               0               0               0               0               0
##          EpsOth             Mps           MpsMe           MpsJt          MpsPro
##               0               0               0               0               0
##          MpsOth           HrGen           EqGen           HrDem            Prot
##               0               0               0               0               0
##         ProtCiv         ProtGrp         ProtLgl         ProtOth           HrFra
##               0               0               0               0               0
##           HrfSp          HrfBor         HrfTinc          HrfOth            HrCp
##               0               0               0               0               0
##         CprLife         CprTort           CprEq         CprSlav          CprLib
##               0               0               0               0               0
##          CprDet         CprFmov         CprFspe         CprFass         CprTria
##               0               0               0               0               0
##         CprPriv         CprVote         CprReli          CprOth           HrSec
##               0               0               0               0               0
##         SerProp         SerWork         SerHeal          SerEdu         SerStdl
##               0               0               0               0               0
##         SerShel           SerSs         SerCult          SerOth            HrNi
##               0               0               0               0               0
##          HrNiMe          HrNiNe         HrNiOth            HrIi         HrIiMon
##               0               0               0               0               0
##         HrIiBod         HrIiOth           HrMob           HrDet             Med
##               0               0               0               0               0
##          MedGov         MedSubs          MedLog          MedOth           HrCit
##               0               0               0               0               0
##          CitGen       CitRights          CitDef          CitOth           JusCr
##               0               0               0               0               0
##         JusCrSp        JusCrSys        JusCrPow           JusEm           JusJu
##               0               0               0               0               0
##          JusPri          JusTra             Dev          DevSoc          DevHum
##               0               0               0               0               0
##        DevInfra             NEC          NatRes           IntFu             Bus
##               0               0               0               0               0
##             Tax           TaxPo          TaxRef          TaxOth             Ban
##               0               0               0               0               0
##          CenBan         BanPers          BanInt           BanXb           LaRef
##               0               0               0               0               0
##        LaRefMan        LaRefRet        LaRefOth           LaNom            LaCH
##               0               0               0               0               0
##          LaCHTa          LaCHIt         LaCHPro         LaCHOth            LaEn
##               0               0               0               0               0
##             Wat          SsrGua              Ce          CeProv           CeGen
```

```
##          0              0              0              0              0
##     SsrPol         SsrArm         SsrDdr       DdrDemil        DdrProg
##          0              0              0              0              0
##     SsrInt         SsrPsf          SsrFf            Cor       SsrCrOcr
##          0              0              0              0              0
##   SsrDrugs           Terr          TjGen           TjAm        TjAmPro
##          0              0              0              0              0
##      TjSan        TjPower        TjAmBan          TjCou         TjJaNc
##          0              0              0              0              0
##      TjJaIc         TjMech        TjPrire          TjVet          TjVic
##          0              0              0              0              0
##       TjMis          TjRep         TjRSym         TjRMa           TjNR
##          0              0              0              0              0
##       ImUN          ImOth          ImRef           ImPK            ImE
##          0              0              0              0              0
##      ImSrc
##          0
```

```
# Análisis de valores vacíos
colSums(df1=="" | df1==" ")
```

```
##        Con          Contp             PP         PPName            Reg
##          0              0             NA             NA              0
##      AgtId            Agt            Dat         Status            Lgt
##          0              0              0              0              0
## N_characters          Agtp          Stage       StageSub           Part
##          0              0              0             NA             NA
##    ThrdPart         OthAgr        Loc1ISO        Loc2ISO       Loc1GWNO
##         NA             NA             NA             NA             NA
##   Loc2GWNO        UcdpCon        UcdpAgr         PamAgr        CowWar
##         NA             NA             NA             NA             NA
##        GCh         GChRhet       GChAntid        GChSubs         GChOth
##          0              0              0              0              0
##       GDis        GDisRhet       GDisAntid       GDisSubs        GDisOth
##          0              0              0              0              0
##       GAge        GAgeRhet       GAgeAntid       GAgeSubs        GAgeOth
##          0              0              0              0              0
##       GMig        GMigRhet       GMigAntid       GMigSubs        GMigOth
##          0              0              0              0              0
##        GRa         GRaRhet        GRaAntid        GRaSubs         GRaOth
##          0              0              0              0              0
##        GRe         GReRhet        GReAntid        GReSubs         GReOth
##          0              0              0              0              0
##       GInd        GIndRhet       GIndAntid       GIndSubs        GIndOth
##          0              0              0              0              0
##       GOth        GOthRhet       GOthAntid       GOthSubs        GOthOth
##          0              0              0              0              0
##       GRef        GRefRhet        GRefSubs        GRefOth           GSoc
##          0              0              0              0              0
##    GSocRhet       GSocAntid       GSocSubs        GSocOth         GeWom
##          0              0              0              0              0
##        GeMe          GeMeNu         GeMeOth           GeSo        GeLgbti
##          0              0              0              0              0
##  GeLgbtiPos      GeLgbtiNeg           GeFa          StDef          StGen
##          0              0              0              0              0
```

```
##        StCon        StSd        StRef       StSym        StInd
##            0           0           0           0           0
##        StUni        StBor       StXbor         Pol       PolGen
##            0           0           0           0           0
##    PolNewInd  PolNewTemp       ConRen        Cons         Ele
##            0           0           0           0           0
##     ElecComm       PolPar   PolParTrans   PolParOth       Civso
##            0           0           0           0           0
##         Tral       Pubad        Polps       PpsSt       PpsSub
##            0           0           0           0           0
##        PpsEx       PpsOro      PpsOthPr      PpsVet       PpsAut
##            0           0           0           0           0
##       PpsInt       PpsOth        Terps      TpsSub      TpsLoc
##            0           0           0           0           0
##       TpsAut       TpsOth          Eps      EpsRes      EpsFis
##            0           0           0           0           0
##       EpsOth          Mps        MpsMe       MpsJt      MpsPro
##            0           0           0           0           0
##       MpsOth        HrGen        EqGen       HrDem        Prot
##            0           0           0           0           0
##      ProtCiv      ProtGrp      ProtLgl     ProtOth       HrFra
##            0           0           0           0           0
##        HrfSp       HrfBor      HrfTinc      HrfOth        HrCp
##            0           0           0           0           0
##      CprLife      CprTort        CprEq      CprSlav      CprLib
##            0           0           0           0           0
##       CprDet      CprFmov      CprFspe      CprFass     CprTria
##            0           0           0           0           0
##      CprPriv      CprVote      CprReli       CprOth       HrSec
##            0           0           0           0           0
##      SerProp      SerWork      SerHeal      SerEdu      SerStdl
##            0           0           0           0           0
##      SerShel        SerSs      SerCult       SerOth        HrNi
##            0           0           0           0           0
##       HrNiMe       HrNiNe      HrNiOth        HrIi      HrIiMon
##            0           0           0           0           0
##      HrIiBod      HrIiOth        HrMob       HrDet         Med
##            0           0           0           0           0
##       MedGov      MedSubs       MedLog       MedOth       HrCit
##            0           0           0           0           0
##       CitGen    CitRights       CitDef       CitOth       JusCr
##            0           0           0           0           0
##      JusCrSp     JusCrSys     JusCrPow       JusEm       JusJu
##            0           0           0           0           0
##       JusPri       JusTra          Dev      DevSoc      DevHum
##            0           0           0           0           0
##     DevInfra          NEC       NatRes        IntFu         Bus
##            0           0           0           0           0
##          Tax        TaxPo       TaxRef       TaxOth         Ban
##            0           0           0           0           0
##       CenBan      BanPers       BanInt       BanXb       LaRef
##            0           0           0           0           0
##     LaRefMan     LaRefRet     LaRefOth       LaNom        LaCH
##            0           0           0           0           0
```

```
##      LaCHTa       LaCHIt      LaCHPro      LaCHOth          LaEn
##           0            0            0            0            0
##         Wat       SsrGua           Ce       CeProv        CeGen
##           0            0            0            0            0
##      SsrPol       SsrArm       SsrDdr     DdrDemil      DdrProg
##           0            0            0            0            0
##      SsrInt       SsrPsf        SsrFf          Cor     SsrCrOcr
##           0            0            0            0            0
##    SsrDrugs         Terr        TjGen         TjAm      TjAmPro
##           0            0            0            0            0
##       TjSan      TjPower      TjAmBan        TjCou       TjJaNc
##           0            0            0            0            0
##      TjJaIc       TjMech       TjPrire        TjVet        TjVic
##           0            0            0            0            0
##       TjMis        TjRep       TjRSym        TjRMa         TjNR
##           0            0            0            0            0
##        ImUN        ImOth        ImRef         ImPK          ImE
##           0            0            0            0            0
##       ImSrc
##           0
```

```
# Análisis de valores duplicados
nrow(df1[duplicated(df1), ])
```

```
## [1] 0
```

```
# ¿Con qué variables tendría sentido un proceso de discretización?
apply(df1,2, function(x) length(unique(x)))
```

```
##          Con        Contp           PP       PPName          Reg
##          169            5          152          152            6
##        AgtId          Agt          Dat       Status          Lgt
##         1789         1751         1470            4           70
## N_characters         Agtp        Stage     StageSub         Part
##         1656            4            7           22         1680
##      ThrdPart       OthAgr       Loc1ISO      Loc2ISO     Loc1GWNO
##          848          925           82           30           83
##      Loc2GWNO      UcdpCon       UcdpAgr       PamAgr      CowWar
##           30          111          294           35           65
##          GCh       GChRhet      GChAntid      GChSubs       GChOth
##            4            2            2            2            2
##         GDis      GDisRhet      GDisAntid     GDisSubs      GDisOth
##            4            2            2            2            2
##         GAge      GAgeRhet      GAgeAntid     GAgeSubs      GAgeOth
##            4            2            2            2            1
##         GMig      GMigRhet      GMigAntid     GMigSubs      GMigOth
##            4            2            2            2            1
##          GRa       GRaRhet      GRaAntid      GRaSubs       GRaOth
##            4            2            2            2            2
##          GRe       GReRhet      GReAntid      GReSubs       GReOth
##            4            2            2            2            2
##         GInd      GIndRhet      GIndAntid     GIndSubs      GIndOth
##            4            2            2            2            2
##         GOth      GOthRhet      GOthAntid     GOthSubs      GOthOth
##            4            2            2            2            1
##         GRef      GRefRhet      GRefSubs      GRefOth         GSoc
```

```
##          4           2           2           2           4
##    GSocRhet   GSocAntid    GSocSubs     GSocOth       GeWom
##          2           2           2           2           2
##        GeMe       GeMeNu      GeMeOth        GeSo     GeLgbti
##          2           2           2           2           1
##  GeLgbtiPos  GeLgbtiNeg        GeFa       StDef       StGen
##          1           1           2           2           2
##       StCon        StSd       StRef       StSym       StInd
##          2           2           2           2           2
##       StUni       StBor      StXbor         Pol      PolGen
##          2           2           2           4           2
##   PolNewInd  PolNewTemp      ConRen        Cons         Ele
##          2           2           4           4           4
##    ElecComm      PolPar  PolParTrans   PolParOth       Civso
##          4           4           2           2           2
##        Tral       Pubad       Polps       PpsSt      PpsSub
##          2           4           4           2           2
##       PpsEx      PpsOro     PpsOthPr      PpsVet      PpsAut
##          2           2           2           2           2
##      PpsInt      PpsOth       Terps      TpsSub      TpsLoc
##          2           2           4           2           2
##      TpsAut      TpsOth         Eps      EpsRes      EpsFis
##          2           2           4           2           2
##      EpsOth         Mps       MpsMe       MpsJt      MpsPro
##          2           4           2           2           2
##      MpsOth       HrGen       EqGen       HrDem        Prot
##          2           2           4           4           4
##     ProtCiv     ProtGrp     ProtLgl     ProtOth       HrFra
##          2           2           2           2           4
##       HrfSp      HrfBor     HrfTinc      HrfOth        HrCp
##          2           2           2           2           2
##     CprLife     CprTort       CprEq     CprSlav      CprLib
##          2           2           2           2           2
##      CprDet     CprFmov     CprFspe     CprFass     CprTria
##          2           2           2           2           2
##     CprPriv     CprVote     CprReli      CprOth       HrSec
##          2           2           2           2           2
##     SerProp     SerWork     SerHeal      SerEdu     SerStdl
##          2           2           2           2           2
##     SerShel       SerSs     SerCult      SerOth        HrNi
##          2           2           2           2           4
##      HrNiMe      HrNiNe     HrNiOth        HrIi     HrIiMon
##          2           2           1           4           2
##     HrIiBod     HrIiOth       HrMob       HrDet         Med
##          2           2           2           2           4
##      MedGov     MedSubs      MedLog      MedOth       HrCit
##          2           2           2           2           4
##      CitGen   CitRights      CitDef      CitOth       JusCr
##          2           2           2           2           2
##     JusCrSp     JusCrSys    JusCrPow       JusEm       JusJu
##          2           2           2           2           2
##      JusPri      JusTra         Dev      DevSoc      DevHum
##          2           2           4           2           2
##    DevInfra         NEC      NatRes       IntFu         Bus
```

```
##          2           2           2           2           4
##        Tax       TaxPo       TaxRef      TaxOth         Ban
##          4           2           2           2           2
##      CenBan     BanPers       BanInt       BanXb       LaRef
##          2           2           2           2           2
##    LaRefMan    LaRefRet     LaRefOth       LaNom        LaCH
##          2           2           2           2           2
##      LaCHTa      LaCHIt      LaCHPro     LaCHOth        LaEn
##          2           2           2           2           2
##        Wat      SsrGua          Ce      CeProv       CeGen
##          2           2           4           2           2
##      SsrPol      SsrArm      SsrDdr    DdrDemil     DdrProg
##          4           4           4           2           2
##      SsrInt      SsrPsf       SsrFf         Cor    SsrCrOcr
##          4           2           2           4           4
##    SsrDrugs        Terr       TjGen        TjAm     TjAmPro
##          4           2           4           4           2
##      TjSan     TjPower     TjAmBan       TjCou      TjJaNc
##          2           2           2           4           2
##      TjJaIc      TjMech      TjPrire       TjVet       TjVic
##          2           4           4           4           4
##       TjMis       TjRep      TjRSym       TjRMa        TjNR
##          4           4           2           2           4
##       ImUN       ImOth       ImRef        ImPK         ImE
##          2           2           2           2           2
##      ImSrc
##          2
```

```r
# Creamos una nueva tabla "df2" con todos los cambios para respetar la original
# Borramos campos sin interés para el análisis
borrar <- c("Part","ThrdPart","OthAgr","Loc1ISO","Loc2ISO","Loc1GWNO","Loc2GWNO","UcdpCon","UcdpAgr","Pa
df2 <- df1[ , !(names(df1) %in% borrar)]
#str(df2)
# Convertimos atributos numéricos en factor
df2$PP <- as.factor(df2$PP)
df2$AgtId <- as.factor(df2$AgtId)
df2$GCh <- as.factor(df2$GCh)
df2$GDis <- as.factor(df2$GDis)
df2$GAge <- as.factor(df2$GAge)
df2$GMig <- as.factor(df2$GMig)
df2$GRa <- as.factor(df2$GRa)
df2$GRe <- as.factor(df2$GRe)
df2$GInd <- as.factor(df2$GInd)
df2$GOth <- as.factor(df2$GOth)
df2$GRef <- as.factor(df2$GRef)
df2$GSoc <- as.factor(df2$GSoc)
df2$GeWom <- as.factor(df2$GeWom)
df2$GeMe <- as.factor(df2$GeMe)
df2$GeSo <- as.factor(df2$GeSo)
df2$GeFa <- as.factor(df2$GeFa)
df2$StDef <- as.factor(df2$StDef)
df2$ConRen <- as.factor(df2$ConRen)
df2$Cons <- as.factor(df2$Cons)
df2$Ele <- as.factor(df2$Ele)
```

```
df2$PolPar <- as.factor(df2$PolPar)
df2$Civso <- as.factor(df2$Civso)
df2$Tral <- as.factor(df2$Tral)
df2$Pubad <- as.factor(df2$Pubad)
df2$Polps <- as.factor(df2$Polps)
df2$PpsSt <- as.factor(df2$PpsSt)
df2$Eps <- as.factor(df2$Eps)
df2$Mps <- as.factor(df2$Mps)
df2$HrGen <- as.factor(df2$HrGen)
df2$EqGen <- as.factor(df2$EqGen)
df2$HrDem <- as.factor(df2$HrDem)
df2$Prot <- as.factor(df2$Prot)
df2$HrFra <- as.factor(df2$HrFra)
df2$HrCp <- as.factor(df2$HrCp)
df2$HrSec <- as.factor(df2$HrSec)
df2$HrNi <- as.factor(df2$HrNi)
df2$HrMob <- as.factor(df2$HrMob)
df2$HrDet <- as.factor(df2$HrDet)
df2$Med <- as.factor(df2$Med)
df2$HrCit <- as.factor(df2$HrCit)
#str(df2)
```

```
# Tratamiento de valores nulos.
# Para el caso de columnas que sean de tipo numéricas, o tipo string, pero no a los factores
df2$StageSub[is.na(df2$StageSub)] <- "Sin Especificar"
df2$PPName[is.na(df2$PPName)] <- "Plan de Paz sin nombre";
# Para el caso de columnas que sean de tipo Factor
# Obtener niveles y agregar adicionalmente (+1) el valor que se quiera incluir, en este caso "9999"
levels <- levels(df2$PP)
levels[length(levels) + 1] <- "9999"
# Refactorizar para incluir "9999" como un nivel de factor
# y reemplazar los nulos por el nuevo factor "9999"
df2$PP <- factor(df2$PP, levels = levels)
df2$PP[is.na(df2$PP)] <- "9999"
# Lo convierto a caracter (convertir en factor anteriormente ha sido ppor motivos academicos)
df2$PP <- as.character(df2$PP)
# Comprobamos que hemos quitado los valores nulos valores nulos
colSums(is.na(df2))
```

```
##          Con        Contp           PP       PPName          Reg
##            0            0            0            0            0
##        AgtId          Agt          Dat       Status          Lgt
##            0            0            0            0            0
## N_characters         Agtp        Stage     StageSub          GCh
##            0            0            0            0            0
##      GChRhet     GChAntid       GChSubs       GChOth         GDis
##            0            0            0            0            0
##     GDisRhet    GDisAntid      GDisSubs      GDisOth         GAge
##            0            0            0            0            0
##     GAgeRhet    GAgeAntid      GAgeSubs      GAgeOth         GMig
##            0            0            0            0            0
##     GMigRhet    GMigAntid      GMigSubs      GMigOth          GRa
##            0            0            0            0            0
##      GRaRhet     GRaAntid       GRaSubs       GRaOth          GRe
```

```
##             0           0           0           0           0
##         GReRhet     GReAntid     GReSubs      GReOth        GInd
##             0           0           0           0           0
##         GIndRhet    GIndAntid    GIndSubs     GIndOth        GOth
##             0           0           0           0           0
##         GOthRhet    GOthAntid    GOthSubs     GOthOth        GRef
##             0           0           0           0           0
##         GRefRhet     GRefSubs     GRefOth        GSoc     GSocRhet
##             0           0           0           0           0
##         GSocAntid    GSocSubs     GSocOth       GeWom        GeMe
##             0           0           0           0           0
##         GeMeNu       GeMeOth       GeSo       GeLgbti   GeLgbtiPos
##             0           0           0           0           0
##       GeLgbtiNeg       GeFa        StDef       StGen       StCon
##             0           0           0           0           0
##           StSd        StRef       StSym       StInd       StUni
##             0           0           0           0           0
##           StBor       StXbor        Pol       PolGen   PolNewInd
##             0           0           0           0           0
##       PolNewTemp      ConRen       Cons         Ele    ElecComm
##             0           0           0           0           0
##         PolPar   PolParTrans   PolParOth      Civso        Tral
##             0           0           0           0           0
##         Pubad        Polps       PpsSt       PpsSub       PpsEx
##             0           0           0           0           0
##         PpsOro      PpsOthPr      PpsVet       PpsAut      PpsInt
##             0           0           0           0           0
##         PpsOth        Terps      TpsSub       TpsLoc      TpsAut
##             0           0           0           0           0
##         TpsOth         Eps        EpsRes       EpsFis      EpsOth
##             0           0           0           0           0
##           Mps         MpsMe       MpsJt       MpsPro      MpsOth
##             0           0           0           0           0
##         HrGen        EqGen        HrDem        Prot      ProtCiv
##             0           0           0           0           0
##         ProtGrp      ProtLgl      ProtOth      HrFra       HrfSp
##             0           0           0           0           0
##         HrfBor       HrfTinc      HrfOth       HrCp      CprLife
##             0           0           0           0           0
##         CprTort       CprEq      CprSlav      CprLib      CprDet
##             0           0           0           0           0
##         CprFmov      CprFspe     CprFass      CprTria     CprPriv
##             0           0           0           0           0
##         CprVote      CprReli      CprOth      HrSec      SerProp
##             0           0           0           0           0
##         SerWork      SerHeal     SerEdu       SerStdl     SerShel
##             0           0           0           0           0
##         SerSs        SerCult     SerOth       HrNi        HrNiMe
##             0           0           0           0           0
##         HrNiNe       HrNiOth     HrIiMon      HrIiBod     HrIiOth
##             0           0           0           0           0
##         HrMob        HrDet         Med        MedGov     MedSubs
##             0           0           0           0           0
##         MedLog       MedOth       HrCit       CitGen     CitRights
```

```
##              0               0               0               0               0
##          CitDef          CitOth           JusCr         JusCrSp        JusCrSys
##              0               0               0               0               0
##         JusCrPow           JusEm           JusJu          JusPri          JusTra
##              0               0               0               0               0
##             Dev          DevSoc         DevHum        DevInfra             NEC
##              0               0               0               0               0
##          NatRes           IntFu             Bus             Tax           TaxPo
##              0               0               0               0               0
##          TaxRef          TaxOth             Ban          CenBan         BanPers
##              0               0               0               0               0
##          BanInt           BanXb           LaRef        LaRefMan        LaRefRet
##              0               0               0               0               0
##         LaRefOth           LaNom            LaCH          LaCHTa          LaCHIt
##              0               0               0               0               0
##         LaCHPro         LaCHOth            LaEn             Wat          SsrGua
##              0               0               0               0               0
##              Ce          CeProv           CeGen          SsrPol          SsrArm
##              0               0               0               0               0
##          SsrDdr        DdrDemil         DdrProg          SsrInt          SsrPsf
##              0               0               0               0               0
##           SsrFf             Cor        SsrCrOcr        SsrDrugs            Terr
##              0               0               0               0               0
##           TjGen            TjAm         TjAmPro           TjSan         TjPower
##              0               0               0               0               0
##          TjAmBan           TjCou          TjJaNc          TjJaIc          TjMech
##              0               0               0               0               0
##          TjPrire           TjVet           TjVic           TjMis           TjRep
##              0               0               0               0               0
##          TjRSym           TjRMa            TjNR            ImUN           ImOth
##              0               0               0               0               0
##           ImRef            ImPK             ImE           ImSrc
##              0               0               0               0
```

```r
# Renombrar columnas
df2 <- rename(df2, "Pais" = "Con",
            "Tipo_Conflicto" = "Contp",
            "PP_id" = "PP", "PP_Nombre" = "PPName", "Region" = "Reg", "Acuerdo_id" = "AgtId", "Acuerdo
            "Estado" = "Status", "Paginas" = "Lgt", "N_caracteres" = "N_characters", "Tpo_Acuerdo" = "
            "G_Jovenes" = "GCh", "G_Discapacidad" = "GDis", "G_Ancianos" = "GAge", "G_Migracion" = "GM
            "G_Indigenas" = "GInd", "G_Otros _Grupos" = "GOth", "G_Refugiados" = "GRef", "G_Clase_Soc
            "G_Masculino" = "GeMe", "G_Sesualidad" = "GeSo", "G_Familia" = "GeFa", "Def_Estado" = "Stk
            "Renov_Constitucional" = "Cons", "Elecciones" = "Ele", "Refor_PPoliticos" = "PolPar", "Soc
            "Compar_PoPolitic" = "Polps", "Reparto_Riqueza" = "PpsSt", "Compar_PoEconomico" = "Eps",
            "Compro_Democracia" = "HrDem", "Med_Proteccion" = "Prot", "Marco_Dhumanos" = "HrFra", "Ma
            "Control_Movilidad" = "HrMob", "Protocolos_Detencion" = "HrDet", "MediosComunica" = "Med"
#str(df2)
```

```r
# Transformar valores binarios
df2$G_Femenino <- ifelse(df2$G_Femenino=='1','Con Mencion','Sin Mencion')
df2$G_Masculino <- ifelse(df2$G_Masculino=='1','Con Mencion','Sin Mencion')
df2$G_Sesualidad <- ifelse(df2$G_Sesualidad=='1','Con Mencion','Sin Mencion')
df2$G_Familia <- ifelse(df2$G_Familia=='1','Con Mencion','Sin Mencion')
df2$Def_Estado <- ifelse(df2$Def_Estado=='1','Con Mencion','Sin Mencion')
df2$Inst_Politicas <- ifelse(df2$Inst_Politicas=='1','Con Mencion','Sin Mencion')
```

```r
df2$Soc_Civil <- ifelse(df2$Soc_Civil=='1','Con Mencion','Sin Mencion')
df2$Lideres <- ifelse(df2$Lideres=='1','Con Mencion','Sin Mencion')
df2$Compro_Dhumanos <- ifelse(df2$Compro_Dhumanos=='1','Con Mencion','Sin Mencion')
df2$Marco_DCiviles <- ifelse(df2$Marco_DCiviles=='1','Con Mencion','Sin Mencion')
df2$Marco_Dsocioecon <- ifelse(df2$Marco_Dsocioecon=='1','Con Mencion','Sin Mencion')
df2$Control_Movilidad <- ifelse(df2$Control_Movilidad=='1','Con Mencion','Sin Mencion')
df2$Protocolos_Detencion <- ifelse(df2$Protocolos_Detencion=='1','Con Mencion','Sin Mencion')
#str(df2)
```

```r
# Transformamos resto de valores
seleccion <- c("G_Jovenes", "G_Discapacidad","G_Ancianos","G_Migracion","G_Racistas","G_Religion","G_Ind
#df2[ , (names(df2) %in% seleccion)] <- ifelse(df2[ , (names(df2) %in% seleccion)]=='1','Mencion Retori
#df2[ , (names(df2) %in% seleccion)] <- ifelse(df2[ , (names(df2) %in% seleccion)]=='2','Mencion Media'
#df2[ , (names(df2) %in% seleccion)] <- ifelse(df2[ , (names(df2) %in% seleccion)]=='3','Mencion Sustan
```

```r
# Grabar los datos en excel
write.xlsx(df2, "dat_pax_agreements.xlsx", sheetName = "PeaceAgreements")
```