

Datos

Test

Lourdes Sofía Elizondo Guajardo Daniela Diaz Delgado Bernardo Ortega Chávez
Gabriel González Bataller

2021-06-23

Contents

1	Data Analysis	2
1.1	Data Reported in Sistema de Información Infonavit: 2013/01-2021/04	2
1.1.1	Number of Credits Granted	2
1.1.2	Credit Amount (MXN)	8
1.1.3	Subsidies	16
1.1.3.1	Number	16
1.1.3.2	Amount (\$MXN)	23
1.2	Data Reported in Sistema de Información Infonavit: Complete Portfolio Analysis	30
2	Sample data	31
2.1	Data Analysis	31
2.2	Adjust Sample to Match Population's Characteristics	33
2.3	Assign Variable Status	33
2.4	Subsidies	34
3	Credit Risk Model	34
3.1	Logistic Regression	34
3.2	Evaluation	34
3.3	Regional Analysis & Clusters	34
4	Building a Cross-Subsidy Scheme	34
4.1	Assigning an Interest Rate	34
5	Comparative Statics & Tests	34

1 Data Analysis

1.1 Data Reported in Sistema de Información Infonavit: 2013/01-2021/04

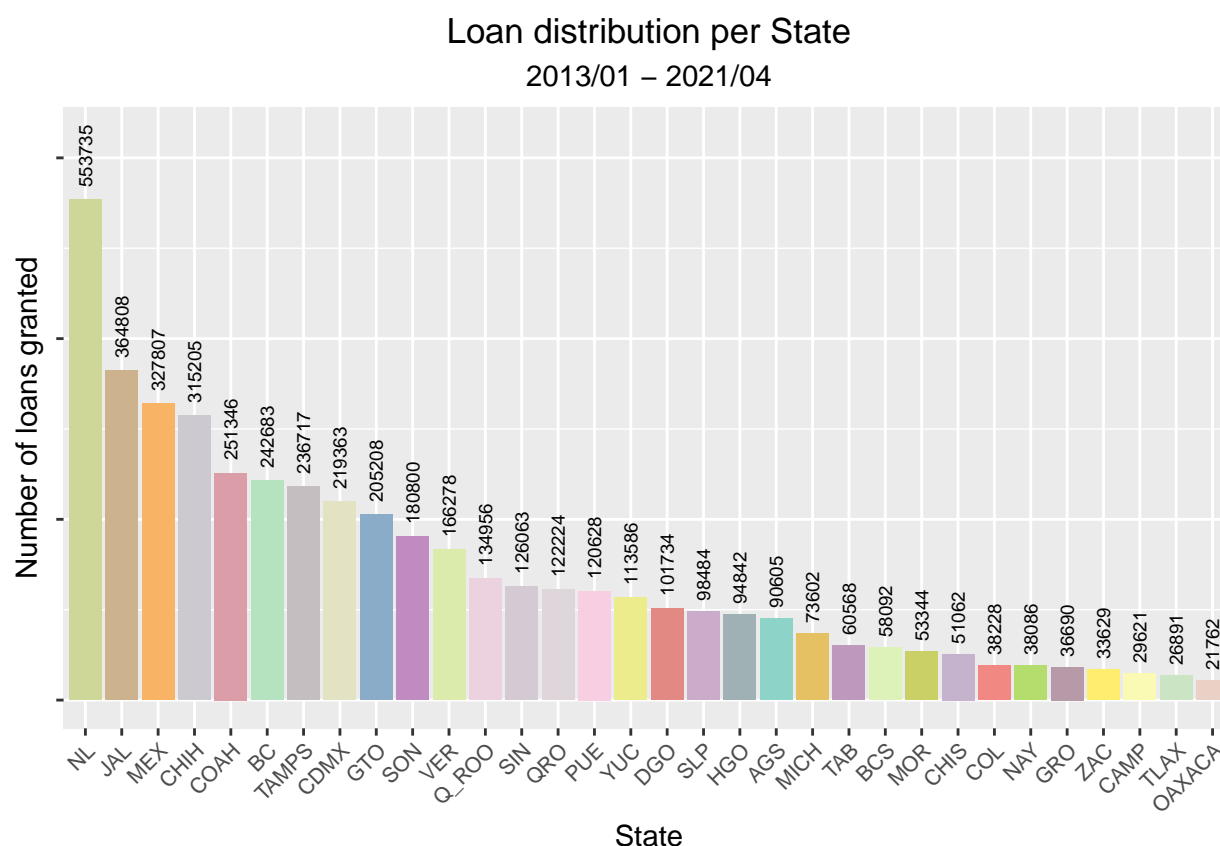
First we analyze information regarding all mortgages originated by Infonavit **from January 2013 to April 2021**. Source: https://portal.mx.infonavit.org.mx/wps/portal/infonavit.web/el-instituto/el-infonavit/SII/!ut/p/z1/pZBBDolwFETP4gn-FJHWJRosFQKiQUA3pitCoujCeH7RJcSKcXY_eTN_MmSoItPZR9vYe3vt7Lm_axOcAgnEa59lQh05iijwc873yAWncgCkyQIFdonEFuzAQOYfv_am-T0ByBVYJkUsEG4iFS3TuSbHB4U_9x8Bxh1fkkm8GC_oBF4TvQHXBt9a3C66V4VWNbMnVfmrRQ!!/dz/d5/L2dBISEvZ0FBIS9nQSEh/

Variables: NC: Número de créditos MC: Monto de créditos NS: Número de subsidios MS: Monto de subsidios

1.1.1 Number of Credits Granted

Website variable description: Número de créditos del Infonavit aprobados y signados. Incluye los productos hipotecarios para vivienda nueva o existente, créditos cofinanciados, créditos coparticipados, mejoramientos y autoproducción. Cifras preliminares.

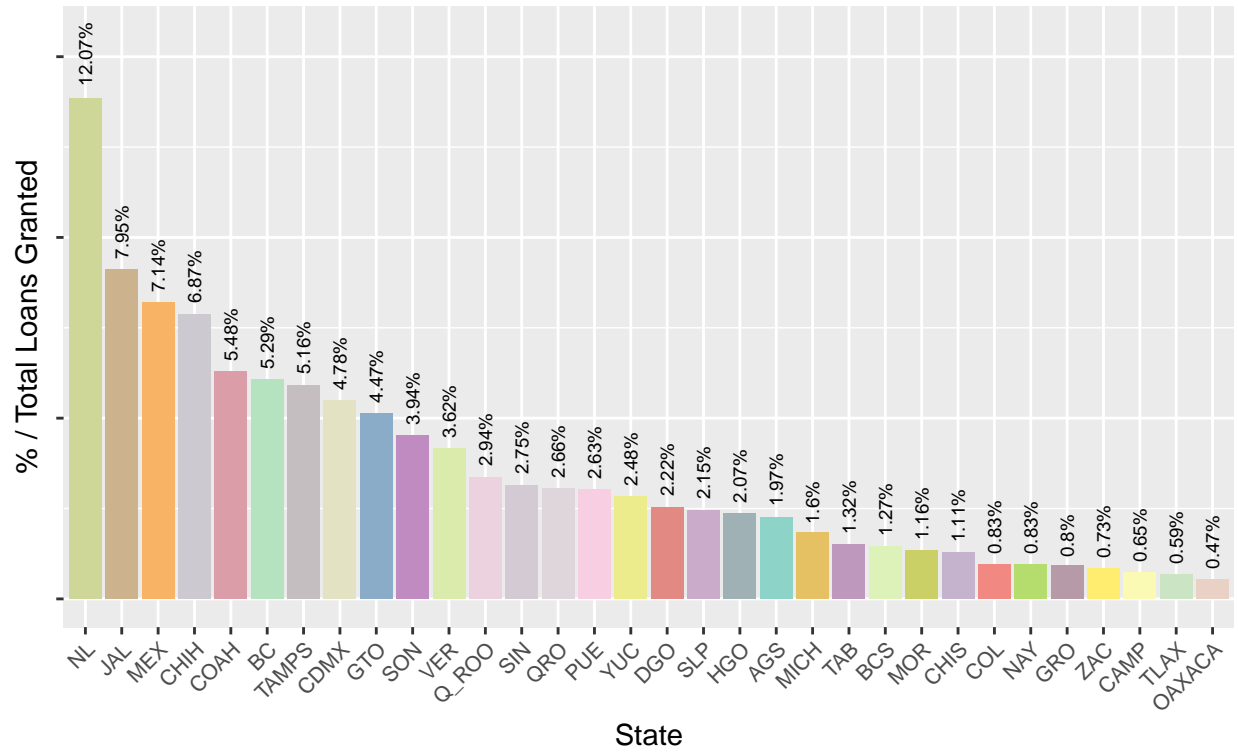
barplot_NC_Estado



barplot_NC_Estado_percent

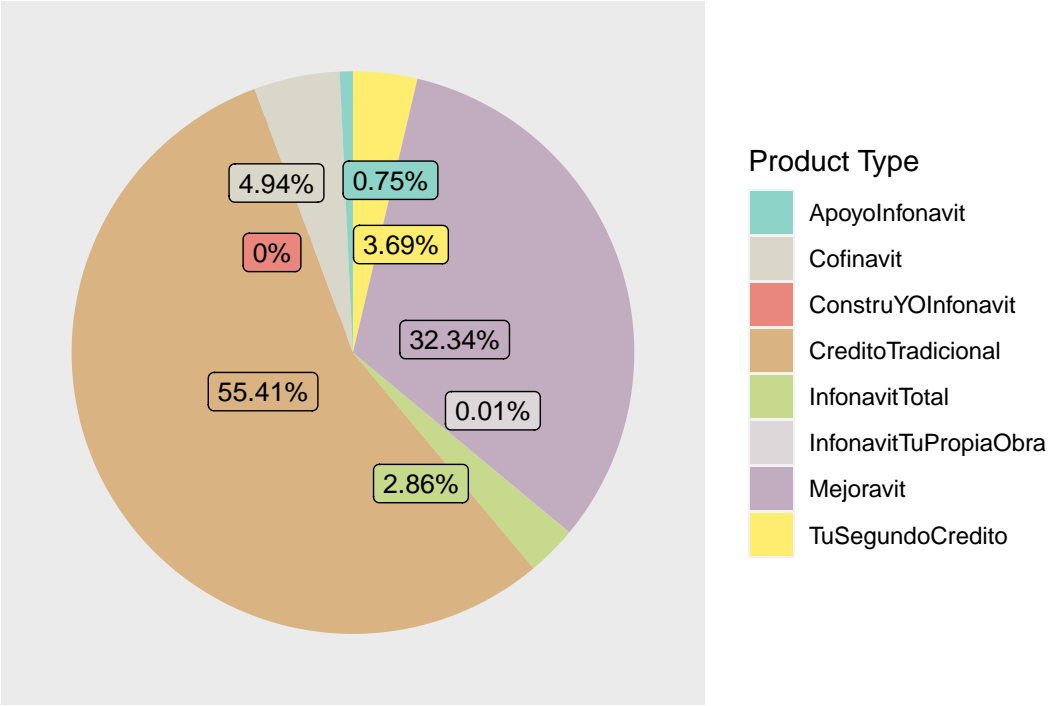
Loan distribution per State

2013/01 – 2021/04



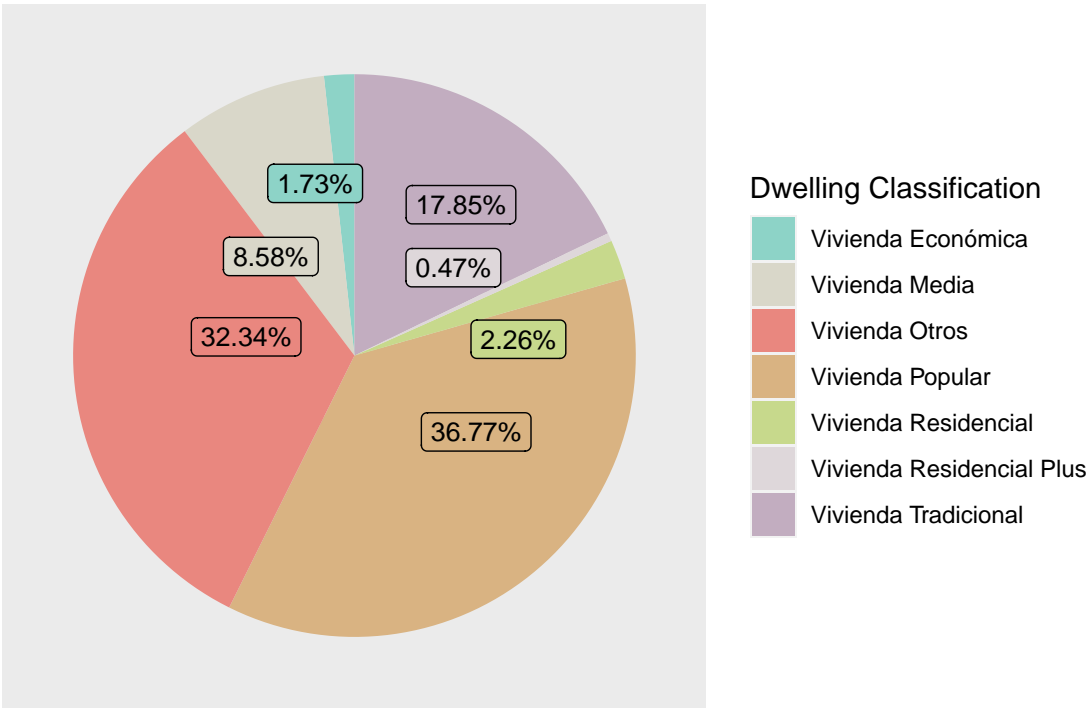
piechart_NC_TP

Loan Distribution by Product Type
2013/01 – 2021/04



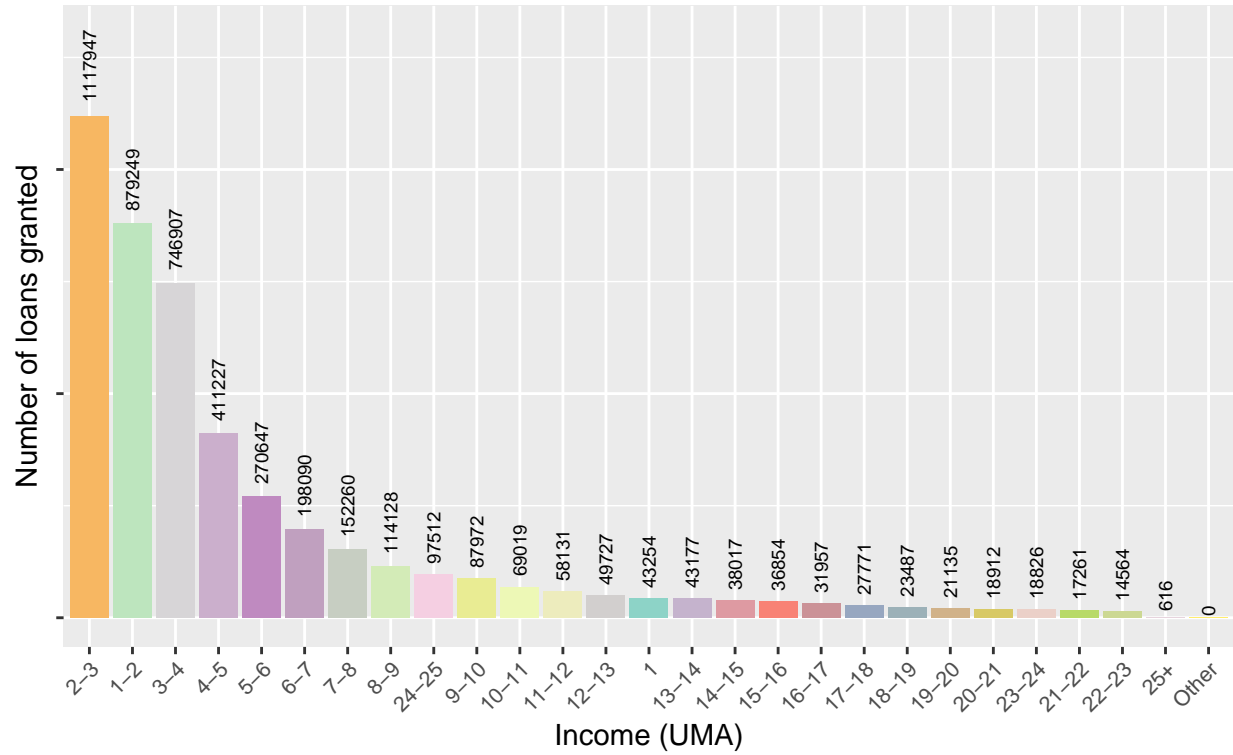
piechart_NC_CV

Loan Distribution by Dwelling Classification
2013/01 – 2021/04



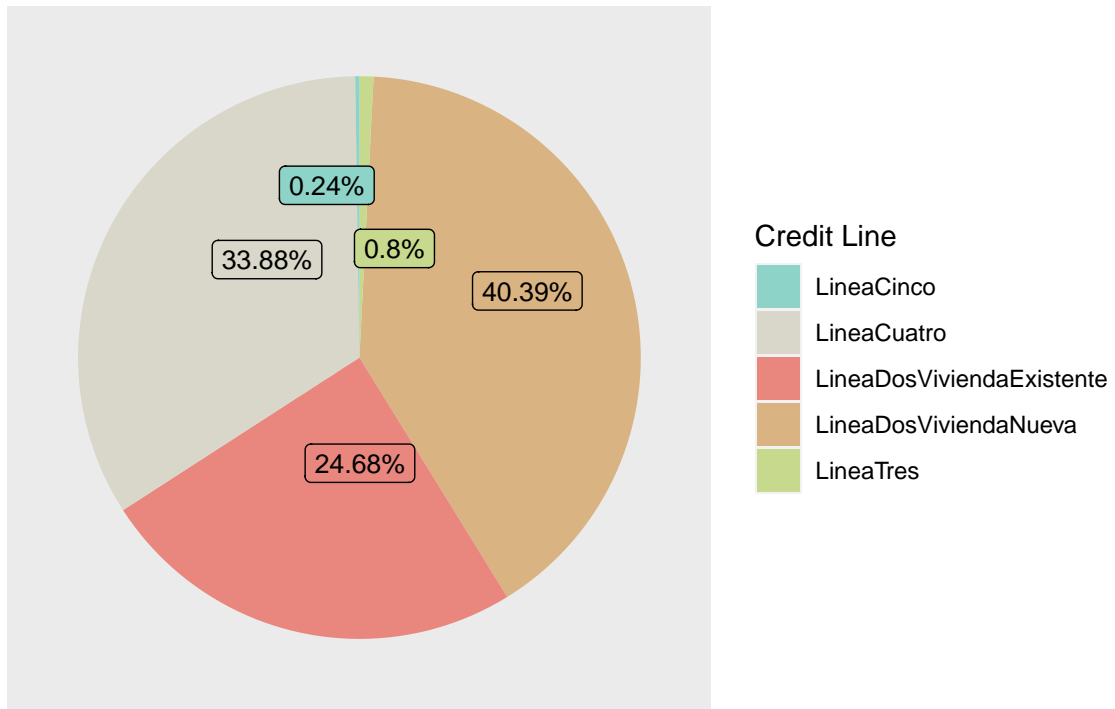
barplot_NC_I

Loans distribution by Income (UMA)
2013/01 – 2021/04

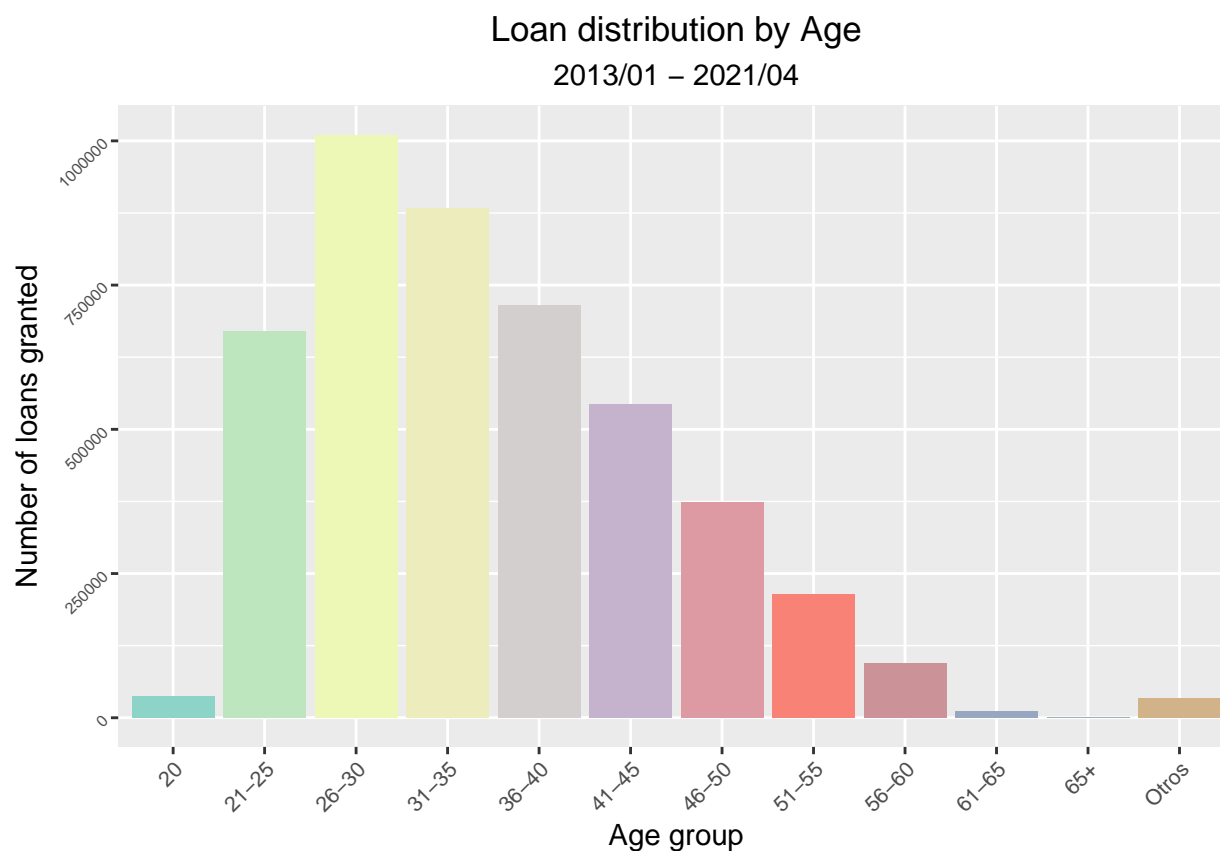


piechart_NC_L

Loan Distribution by Credit Line
2013/01 – 2021/04



barplot_NC_Edad

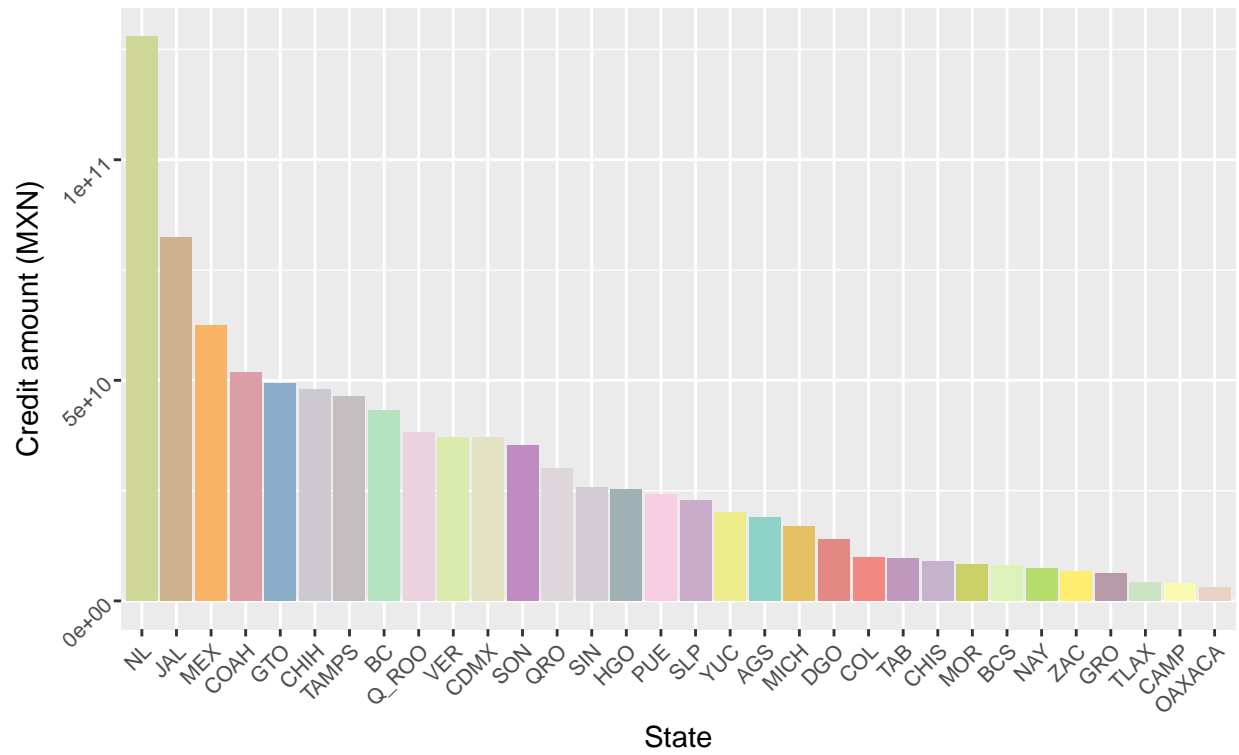


1.1.2 Credit Amount (MXN)

Website variable description: Recursos del fondo para financiar una acción de vivienda y que constituye el pasivo que el trabajador adquiere con el Instituto. En productos cofinanciados y coparticipados, se considera únicamente la parte financiada por el Infonavit. No se consideran productos donde el Instituto únicamente provee la garantía hipotecaria. A precios corrientes. Cifras preliminares.

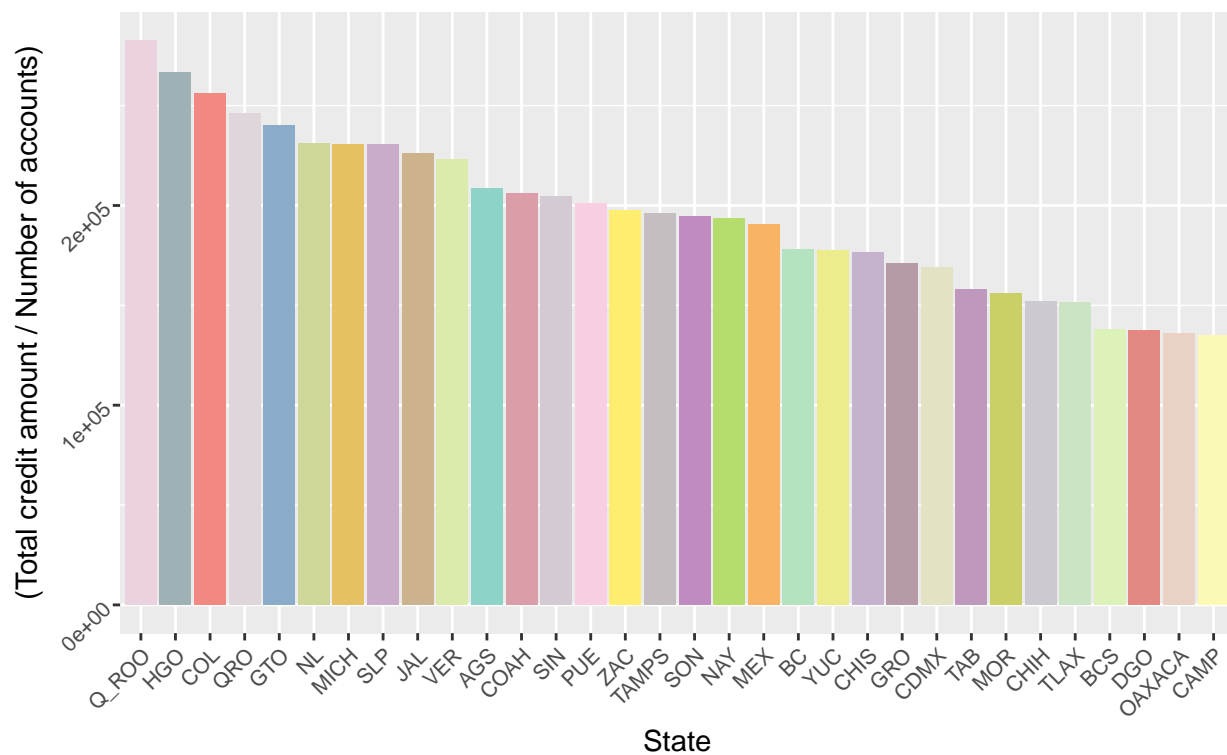
barplot_MC_Estado

Total Amount of Granted Credit by State
2013/01 – 2021/04



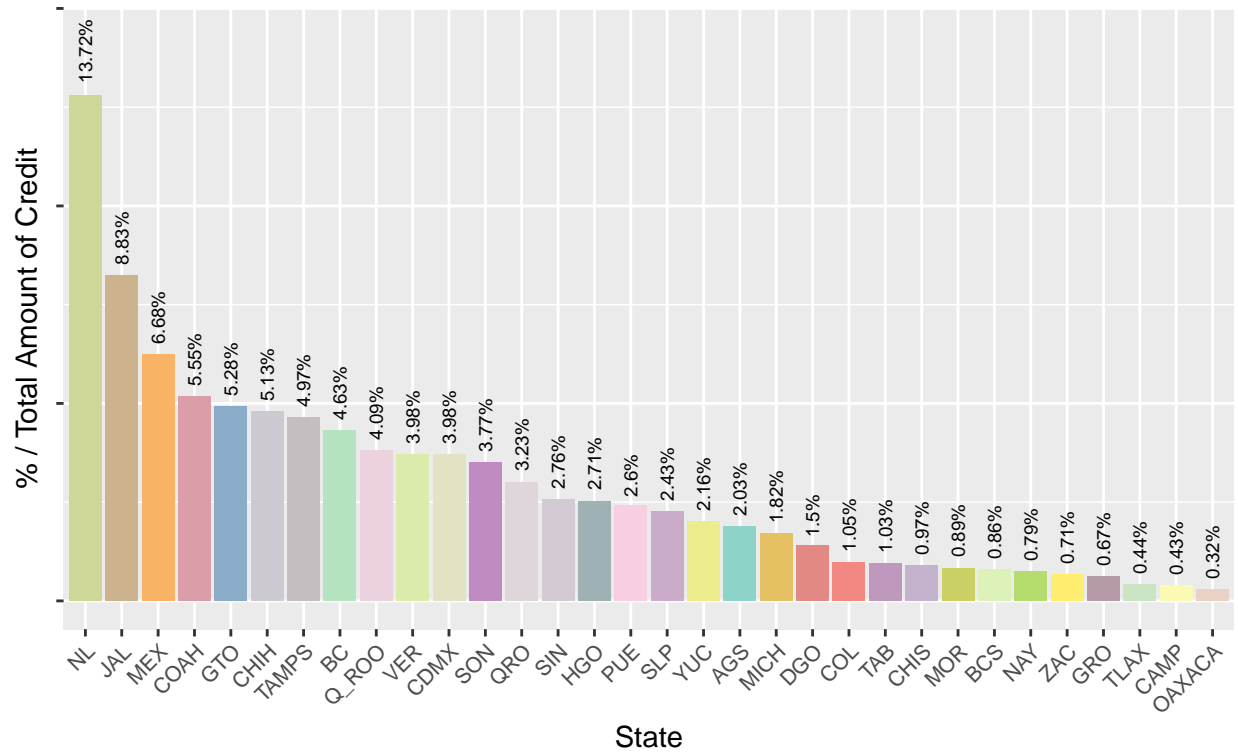
barplot_MC_Estado_av

Average credit amount per State
2013/01 – 2021/04



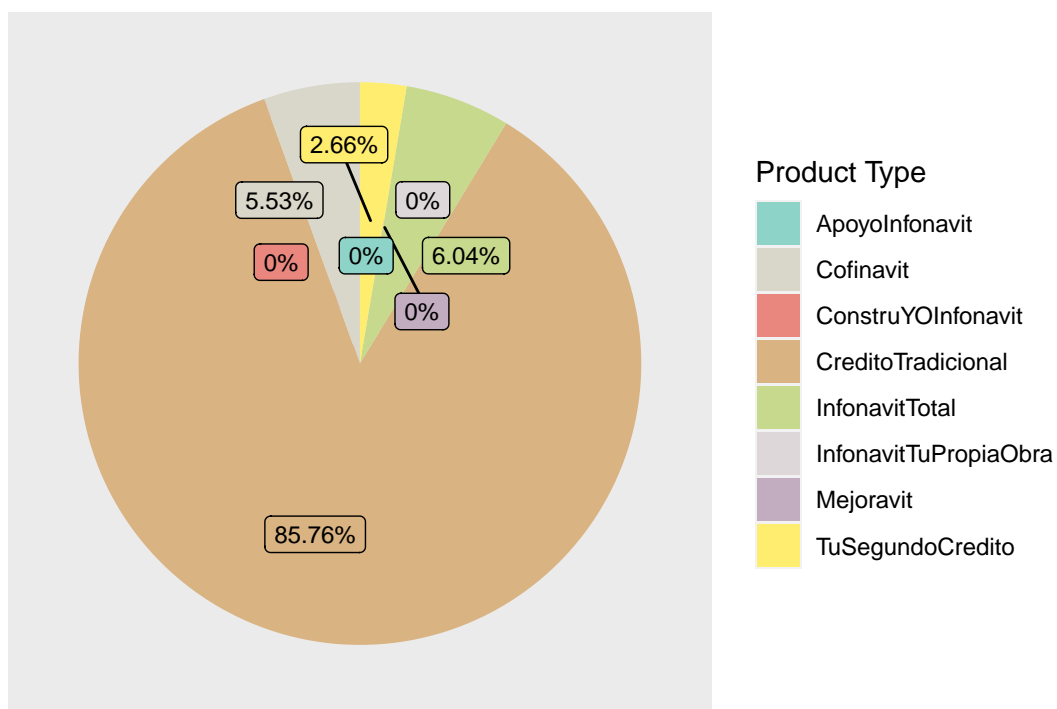
barplot_MC_Estado_percent

Total Amount of Granted Credit by State (%)
2013/01 – 2021/04



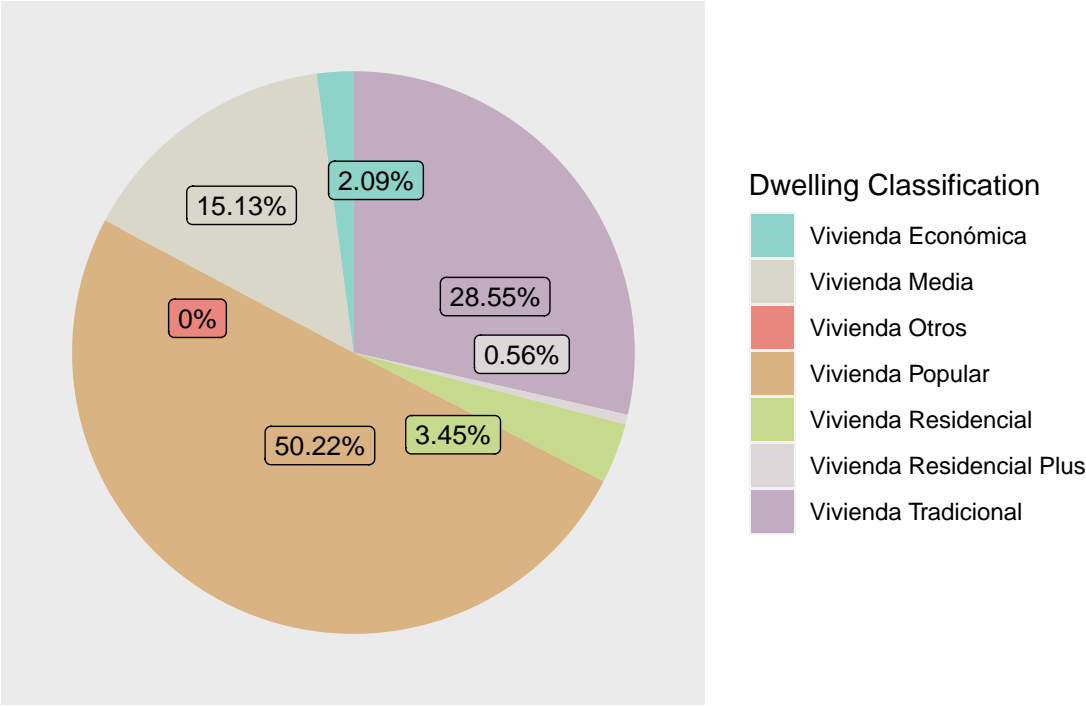
piechart_MC_TP

Total Amount of Granted Credit by Product Type 2013/01 – 2021/04



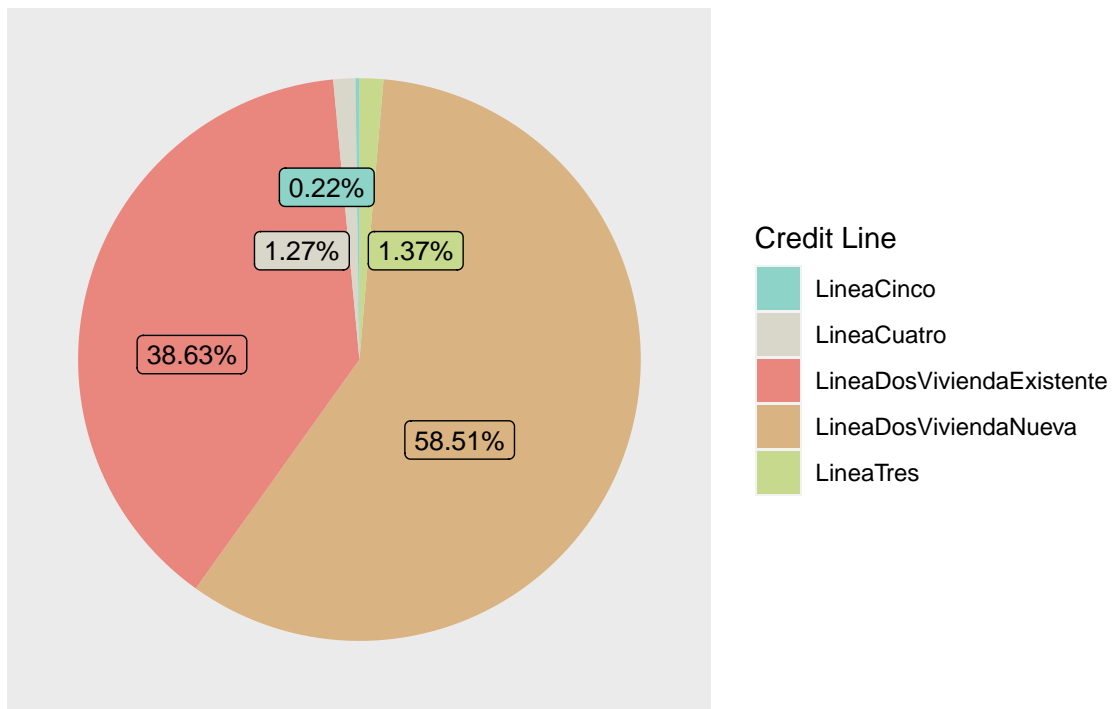
piechart_MC_CV

Total Amount of Granted Credit by Dwelling Classification
2013/01 – 2021/04



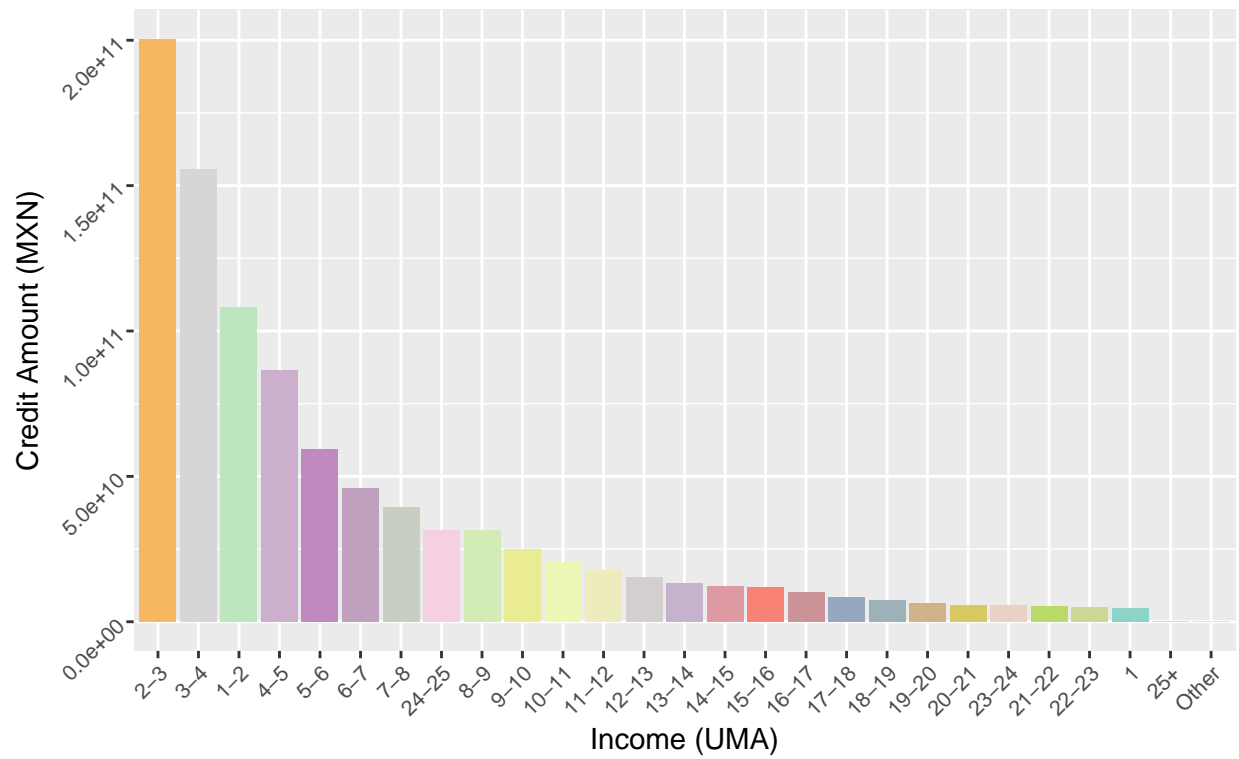
piechart_MC_L

Total Amount of Granted Credit by Credit Line
2013/01 – 2021/04

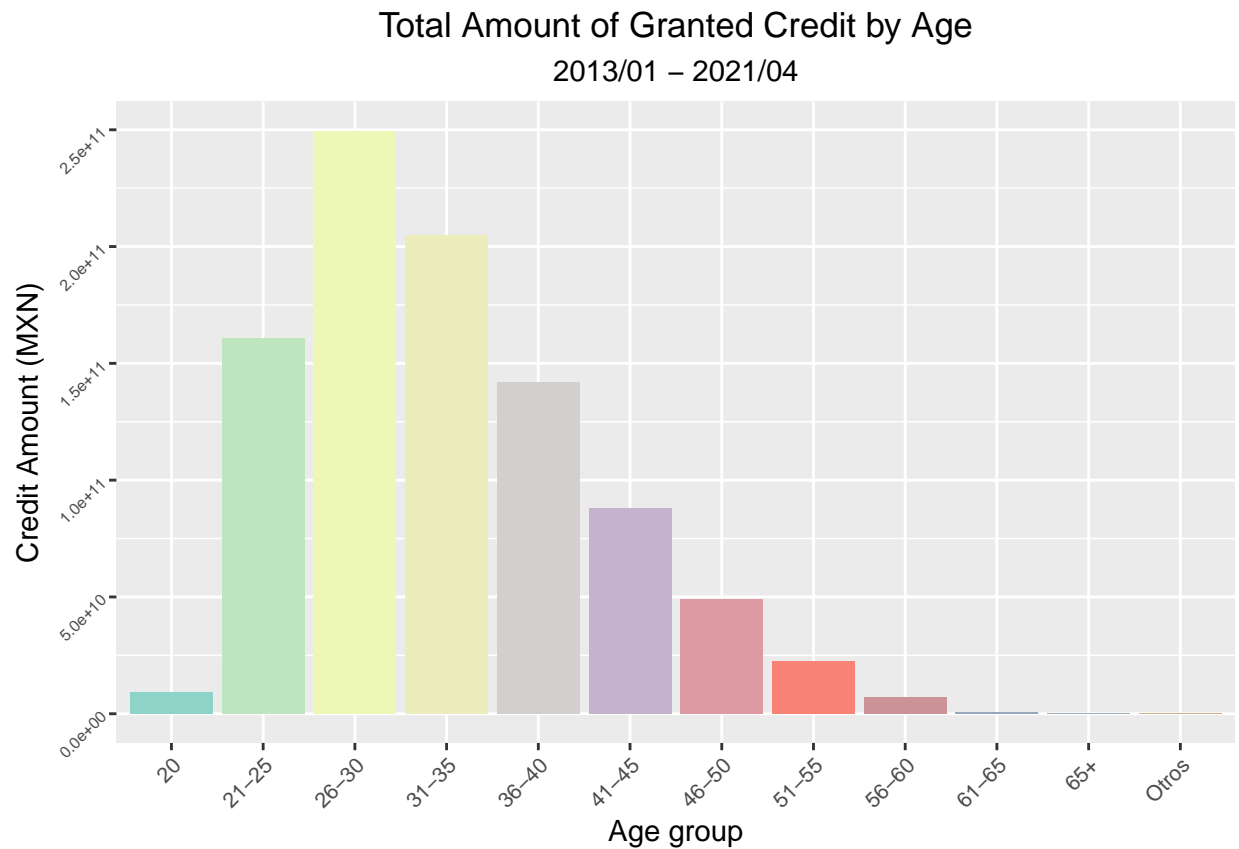


barplot_MC_I

Total Amount of Granted Credit by Income (UMA)
2013/01 – 2021/04



barplot_MC_Edad



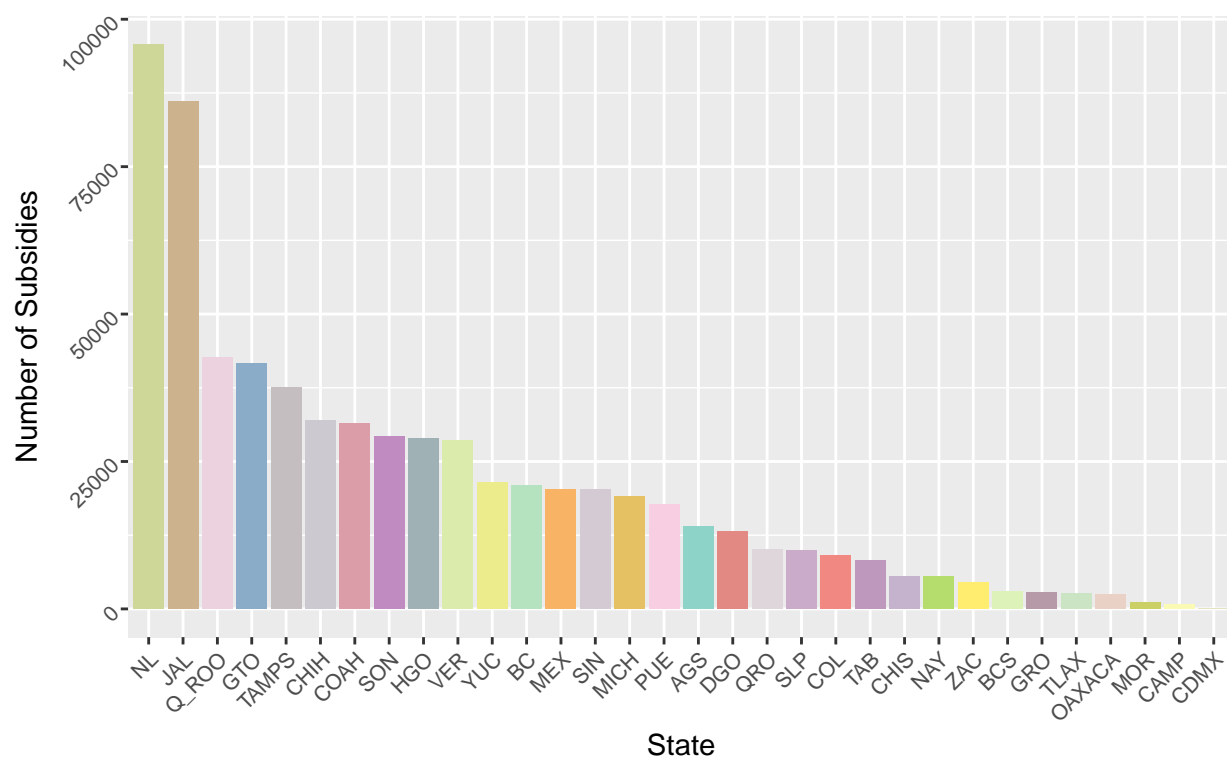
1.1.3 Subsidies

1.1.3.1 Number Website variable description: Número de créditos que recibieron el subsidio federal otorgado por la Comisión Nacional de Vivienda (Conavi). Se consideran únicamente los subsidios dispersados a través de un crédito de vivienda del Infonavit. Cifras preliminares.

barplot_NS_Estado

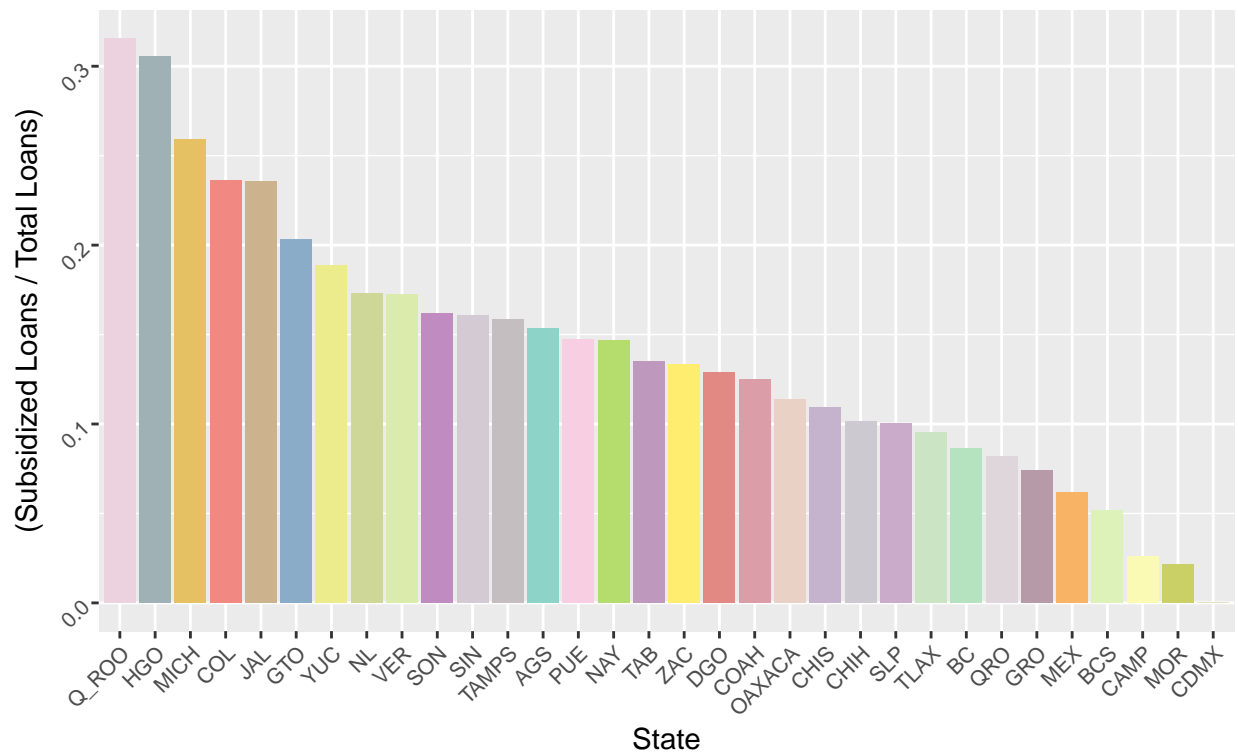
Number of Granted Subsidies by State

2013/01 – 2021/04



barplot_NS_Estado_prop

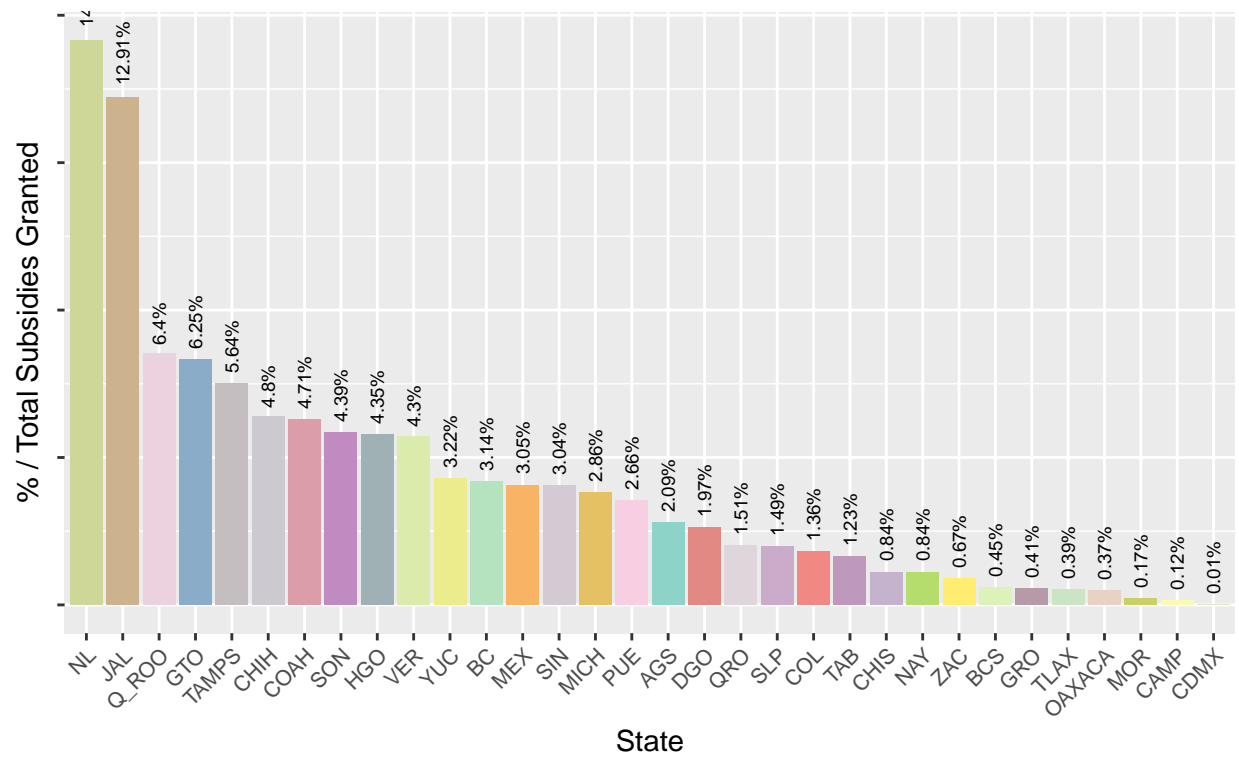
% of subsidized loans per State
2013/01 – 2021/04



barplot_NS_Estado_percent

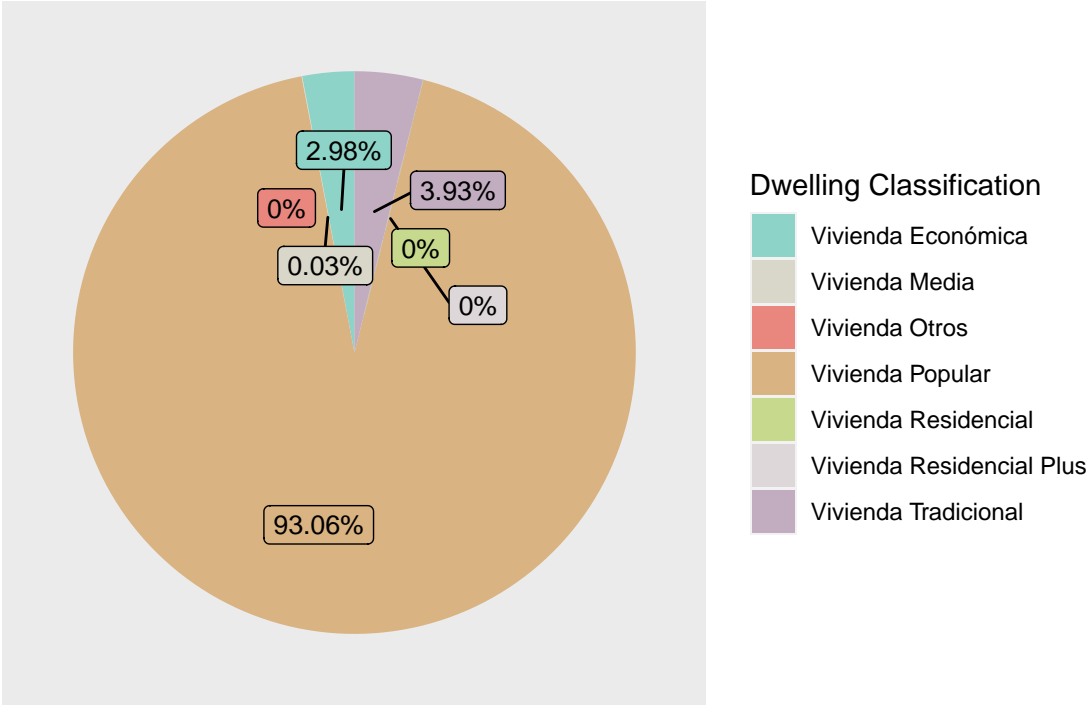
Number of Granted Subsidies by State (%)

2013/01 – 2021/04



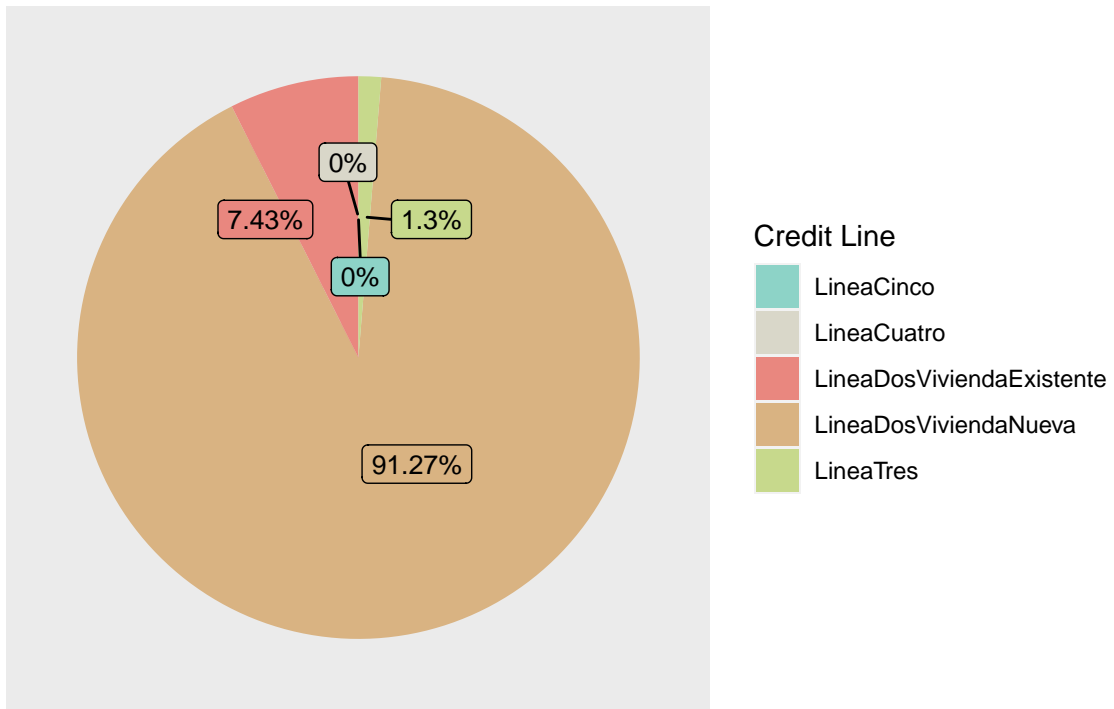
piechart_NS_CV

Distribution of Subsidies by Dwelling Classification
2013/01 – 2021/04



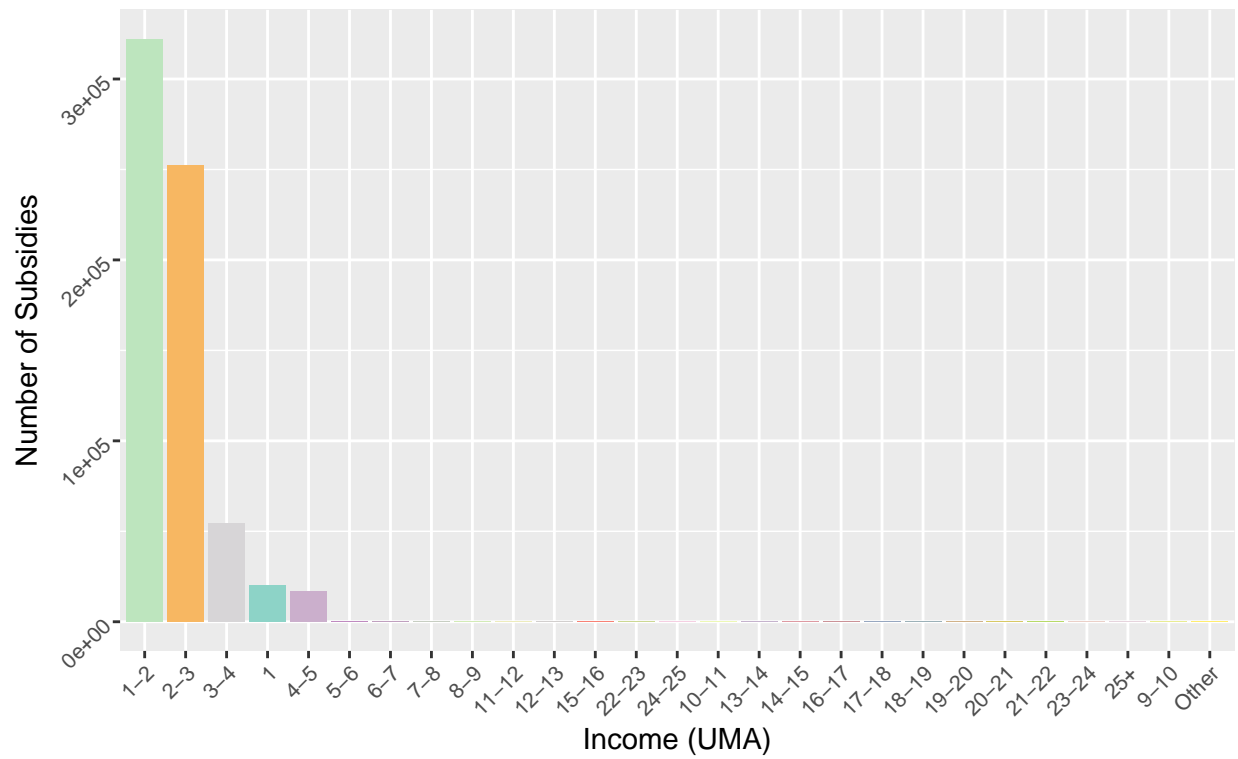
piechart_NS_L

Distribution of Subsidies by Credit Line
2013/01 – 2021/04

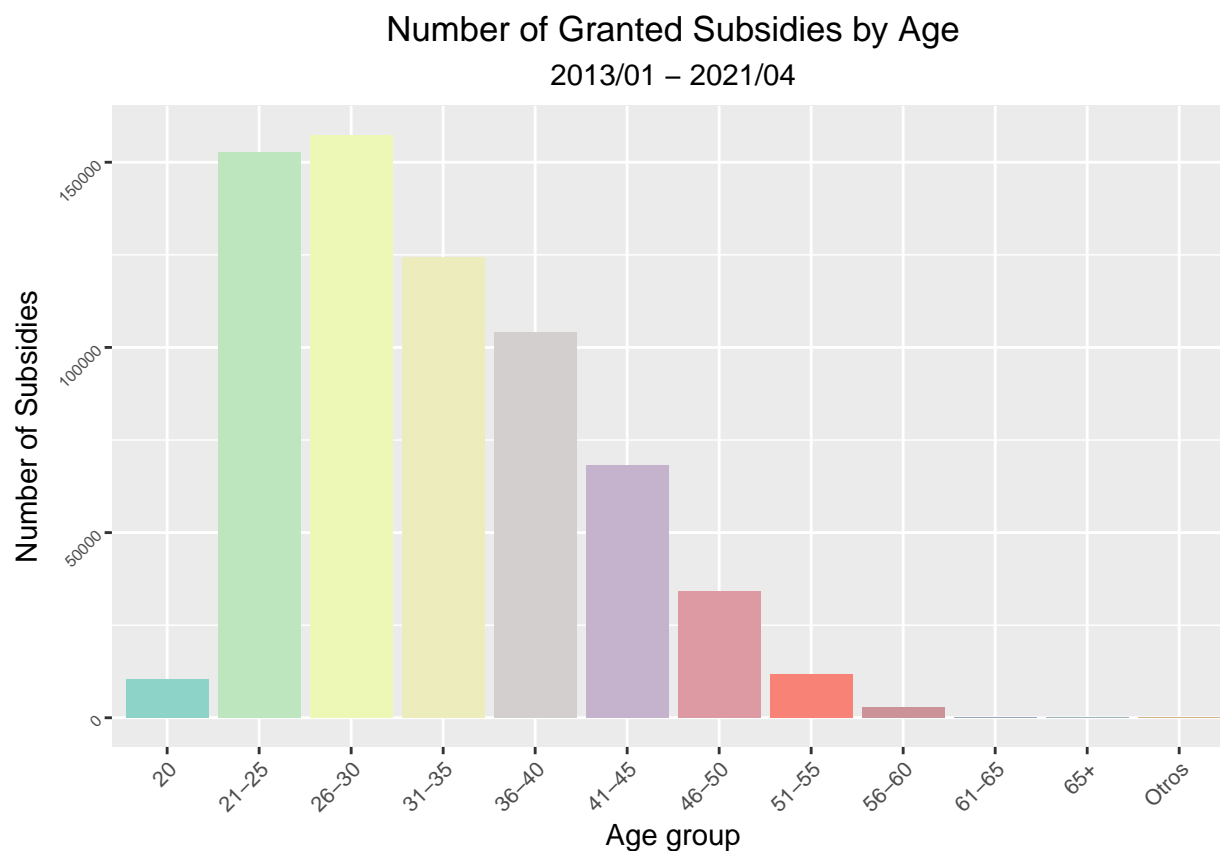


barplot_NS_I

Number of Granted Subsidies by Income (UMA)
2013/01 – 2021/04



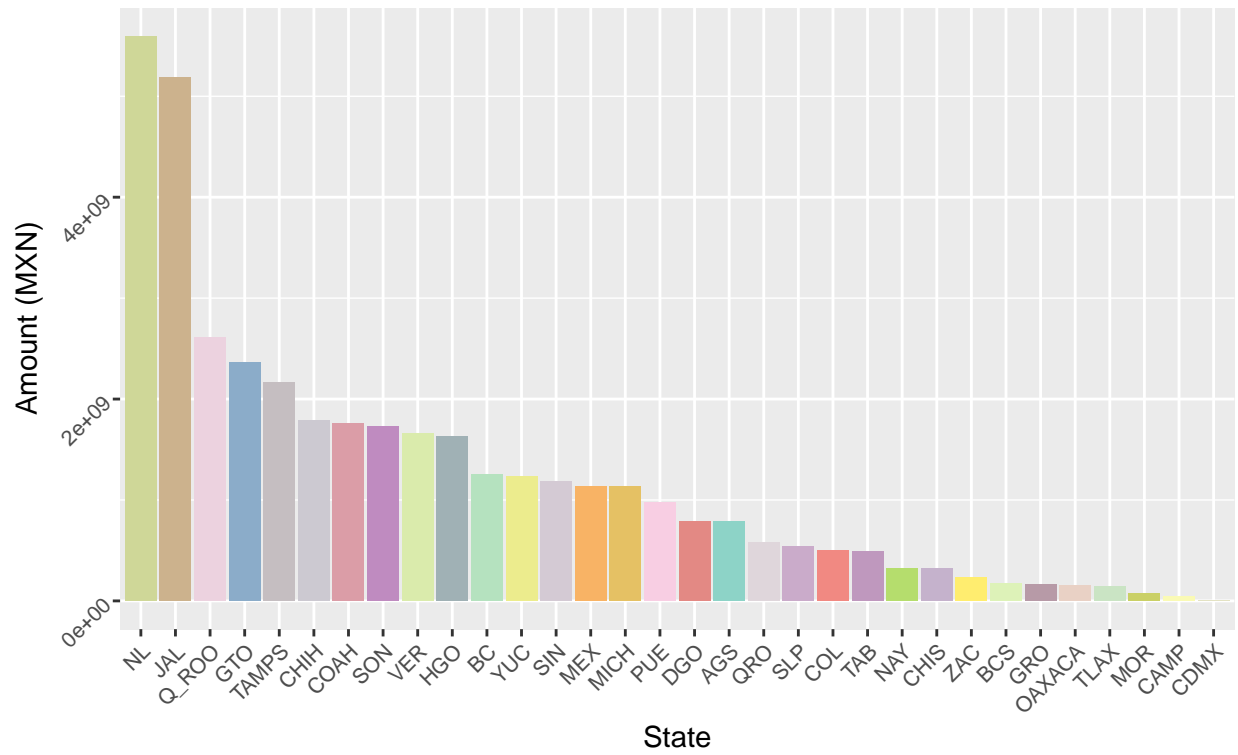
barplot_NS_Edad



1.1.3.2 Amount (\$MXN) Website variable description: Monto de subsidio del gobierno federal otorgado por la Comisión Nacional de Vivienda (Conavi), como complemento a un crédito Infonavit. Se consideran únicamente los subsidios dispersados a través de un crédito de vivienda del Infonavit. A precios corrientes. Cifras preliminares.

barplot_MS_Estado

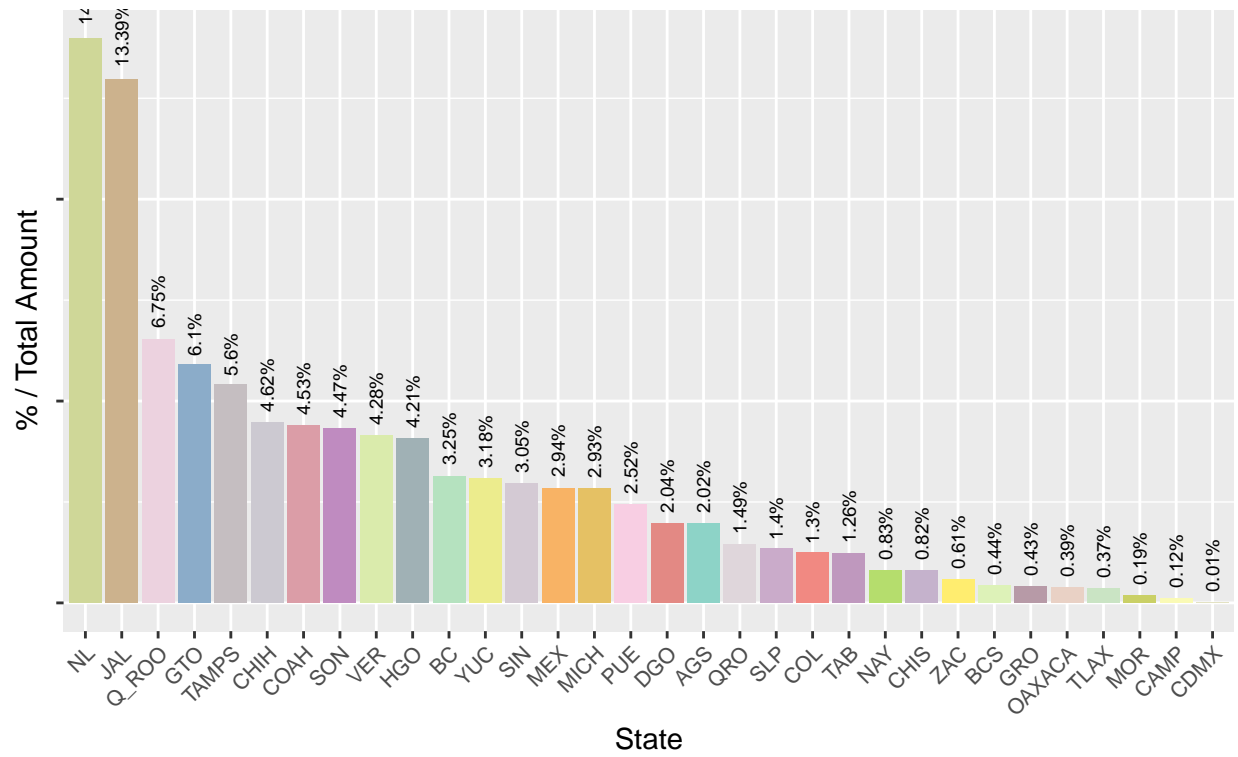
Amount of Granted Subsidies by State
2013/01 – 2021/04



barplot_MS_Estado_percent

Amount of Granted Subsidies by State (%)

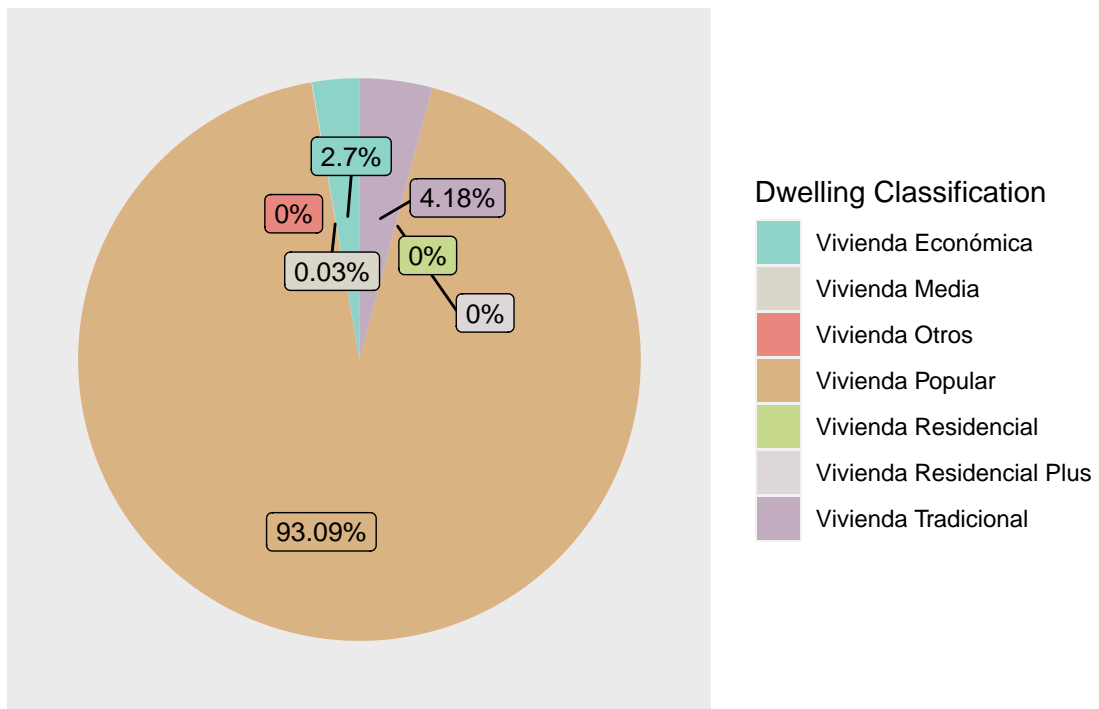
2013/01 – 2021/04



piechart_MS_CV

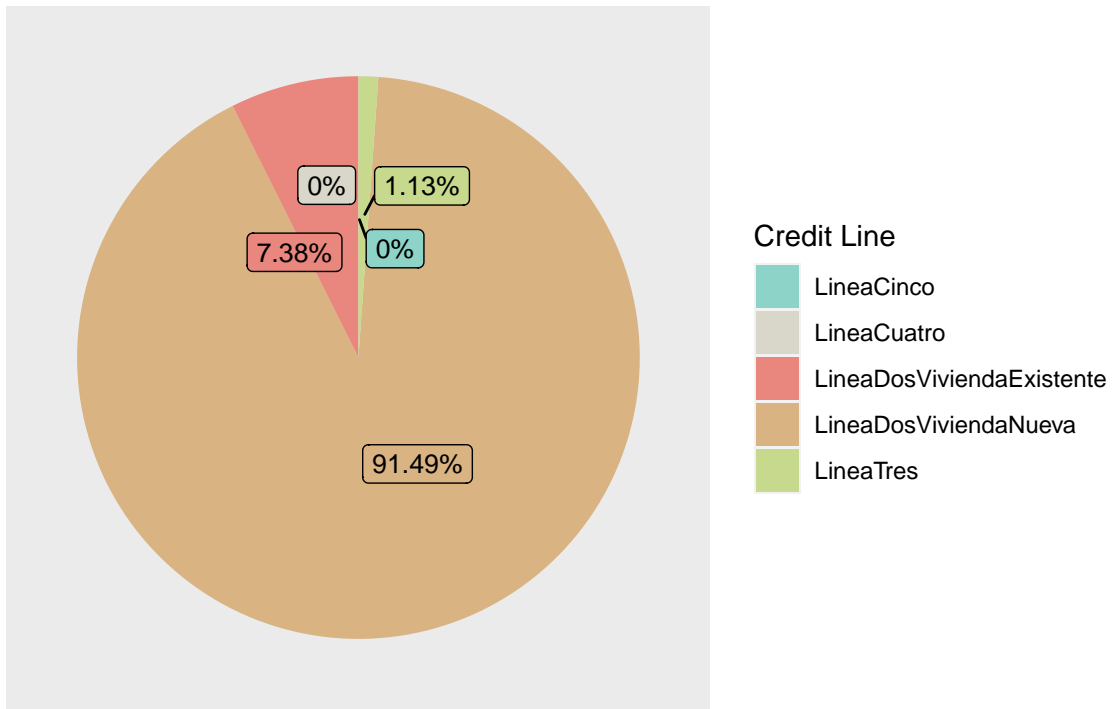
Amount of Subsidies by Dwelling Classification

2013/01 – 2021/04



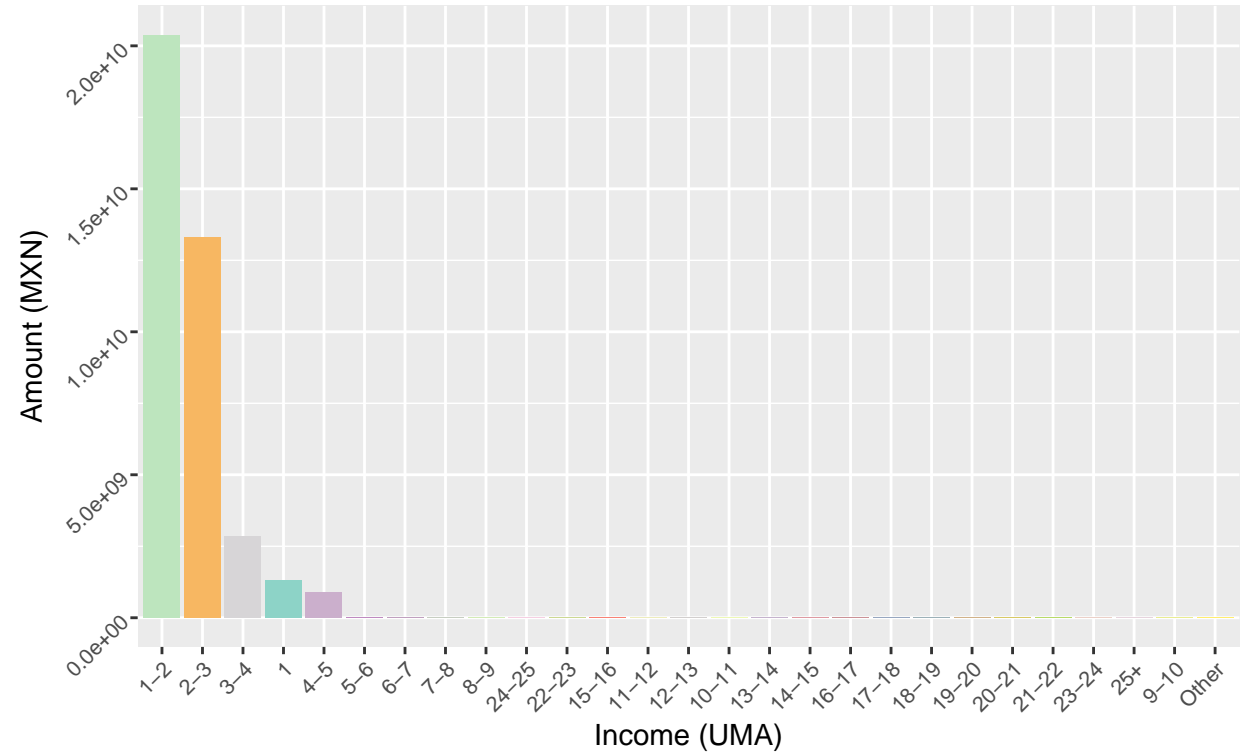
piechart_MS_L

Amount of Subsidies by Credit Line
2013/01 – 2021/04



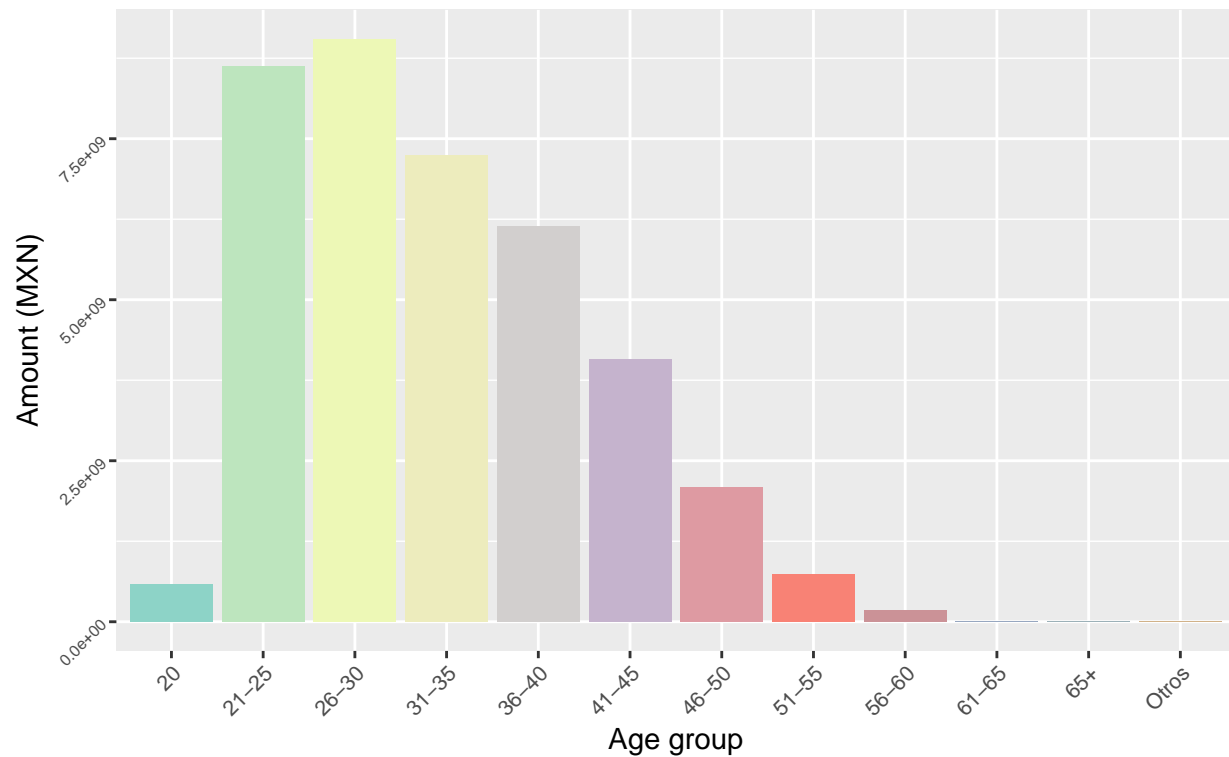
barplot_MS_I

Amount of Granted Subsidies by Income
2013/01 – 2021/04



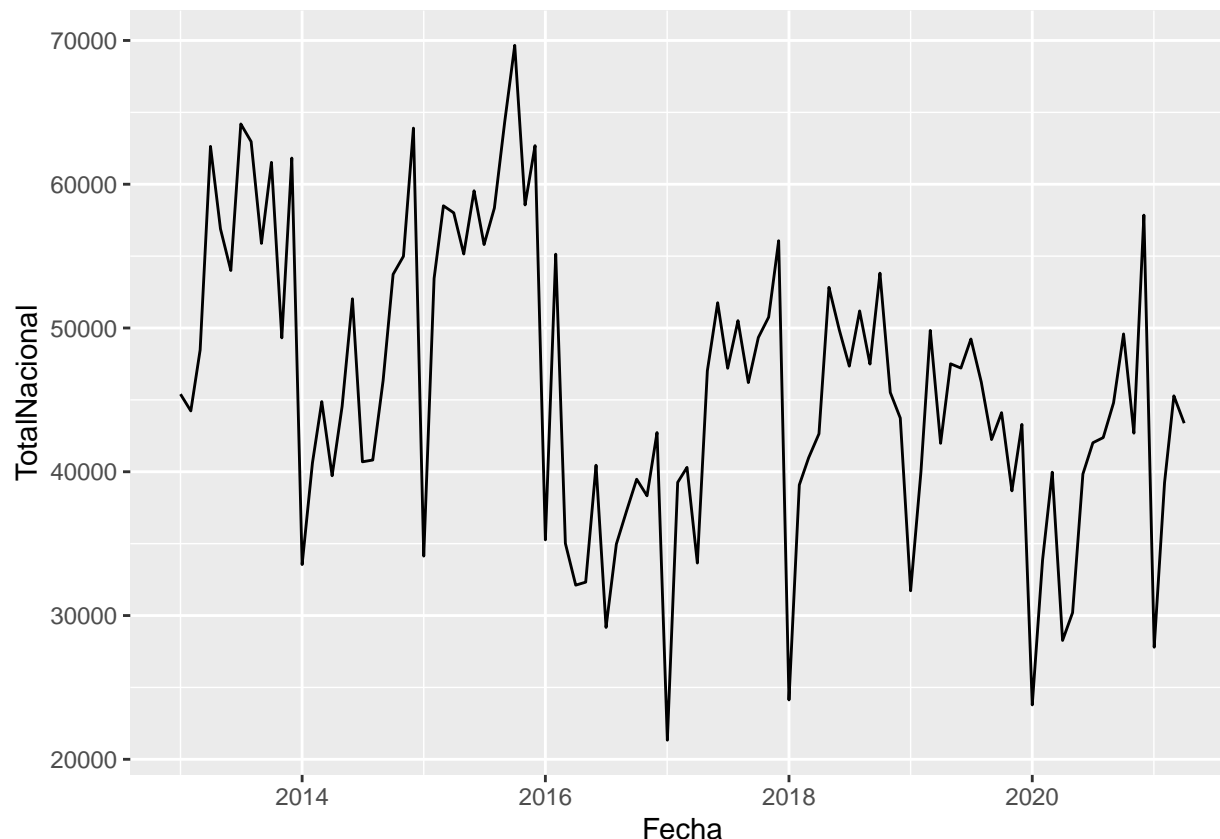
barplot_MS_Edad

Amount of Granted Subsidies by Age
2013/01 – 2021/04



Solo para ver **seasonality** en el número de créditos otorgados por mes:

```
timeseries_NCTotal<-ggplot(data=NCTotal)+
  geom_line(aes(x=Fecha,y=TotalNacional))
timeseries_NCTotal
```



1.2 Data Reported in Sistema de Información Infonavit: Complete Portfolio Analysis

Source: https://portal.mx.infonavit.org.mx/wps/portal/infonavit.web/transparencia/estadisticas/!ut/p/z1/jY_BCslwEEQ_aTdNSdZjUmtalKQFQ-tepCcJaPUgfr8iXjwYO7eBNzMMMIzA8_RIp-mervN0fvkDq6NyiE1VCh_MtsSebIdWd1IZDcMbKAjRWRTeUUNoNnVbr3aycFEAL8njDxlcls8AnK8fgL8nmCXC3u29rUKQtNYfIHfx38jtEuOIqX0C2O

Leaving behind the database of NC, MC, NS, MS -> we now consider the information reported in the section “Estadísticas” of Sistema de Información Infonavit, which gives operating indicators of Infonavit’s **complete** portfolio.

This database decomposes Infonavit’s portfolio by: past due portfolio (vencida), extension portfolio (prórroga) and current portfolio (vigente).

Note that at the beginning of this document we were analyzing credits originated from January 2013 to April 2021, but now this new database includes mortgages originated **before** 2013, considering the **complete portfolio** as of 30th of April, 2021.

Note: ICV stands for índice de cartera vencida

```
head(C)
```

```
##      Fecha Estado Vigente Prorroga Vencida Total      ICV
## 1 2021-04-01   AGS   92637    4684    9574 106895 0.08956453
## 2 2021-04-01    BC  210116   10479   34919 255514 0.13666179
## 3 2021-04-01   BCS   33791    1903    5228  40922 0.12775524
```

```
## 4 2021-04-01    CAMP    21013      1069      3301  25383 0.13004767
## 5 2021-04-01    COAH   235731     14253     40243 290227 0.13866043
## 6 2021-04-01     COL    47764      2238      6515  56517 0.11527505
```

```
head(S)
```

```
##      Fecha Estado  Vigente  Prorroga  Vencida    Total    ICV
## 1 2021-04-01    AGS 24253.25 1442.4995  3334.477 29030.225 0.1148622
## 2 2021-04-01     BC 59778.15 3404.6323 12985.804 76168.581 0.1704877
## 3 2021-04-01    BCS 10569.41  671.9220  2328.736 13570.068 0.1716083
## 4 2021-04-01    CAMP  5734.58  344.0202  1368.003  7446.603 0.1837084
## 5 2021-04-01    COAH 65262.28 4499.8120 14324.478 84086.573 0.1703539
## 6 2021-04-01     COL 13339.25  738.7096  2550.377 16628.333 0.1533754
```

2 Sample data

Now, we consider our main database. This is the data from where we intend to develop our credit risk model; it describes credits originated from January 2013 to April 2021 *individually*, characterizing each beneficiary by different variables like age, gender, type of dwelling, etc. Source: https://sniiv.conavi.gob.mx/reportes/datos_abiertos.aspx

In theory, this data base should be exactly the same as the data analyzed at the beginning, reported in “Sistema de Información Infonavit” for 2013/01-2021/04 in Number and Amount of Credits (NC & MC). However, we do not expect this database to be distributed exactly the same as the **complete portfolio** information. We intend to manipulate this sample to best resemble the complete portfolio, so that the credit risk model can more accurately represent reality.

2.1 Data Analysis

```
sample<-readRDS(file="DATA/MUESTRA.Rds")
head(sample)
```

```
##      AÑO MES_ACUMULADO CLAVE_ENTIDAD CLAVE_MUNICIPIO    MODALIDAD
## 1 2013              1              1              1 Viviendas nuevas
## 2 2013              1              1              1 Viviendas nuevas
## 3 2013              1              1              1 Viviendas nuevas
## 4 2013              1              1              1 Viviendas nuevas
## 5 2013              1              1              1 Viviendas nuevas
## 6 2013              1              1              5 Viviendas nuevas
##      DESTINO      TIPO GENERO      EDAD      INGRESO
## 1 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos 2.61 a 4.00
## 2 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos 4.01 a 6.00
## 3 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos 6.01 a 9.00
## 4 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos  M·s de 12
## 5 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos  M·s de 12
## 6 Vivienda nueva Cofinanciamientos y subsidios Hombre 29 o menos 6.01 a 9.00
##      VALOR_VIVIENDA      MONTO CLAVE
## 1      Tradicional  283223.1    1.1
## 2      Popular  1365780.9    1.1
## 3      Tradicional  803068.0    1.1
```

```
## 4 Tradicional 671815.4 1.1
## 5 Media 1650390.9 1.1
## 6 EconŰmica 214819.1 1.5
```

```
tail(sample)
```

```
##      AÑO MES_ACUMULADO CLAVE_ENTIDAD CLAVE_MUNICIPIO MODALIDAD
## 1397063 2021           1           28           27 Viviendas usadas
## 1397064 2021           1            3            1 Viviendas usadas
## 1397065 2021           2           14           97 Viviendas usadas
## 1397066 2021           2           28           35 Viviendas usadas
## 1397067 2021           1           30           68 Viviendas usadas
## 1397068 2021           1           14           44 Viviendas usadas
##      DESTINO TIPO GENERO EDAD INGRESO
## 1397063 Vivienda usada Credito individual Hombre 29 o menos 2.6 o menos
## 1397064 Vivienda usada Credito individual Hombre 29 o menos 2.6 o menos
## 1397065 Vivienda usada Credito individual  Mujer 29 o menos 2.6 o menos
## 1397066 Vivienda usada Credito individual  Mujer 29 o menos 2.6 o menos
## 1397067 Vivienda usada Credito individual  Mujer 29 o menos 2.6 o menos
## 1397068 Vivienda usada Credito individual  Mujer 29 o menos 2.6 o menos
##      VALOR_VIVIENDA MONTO CLAVE
## 1397063 EconŰmica 312690.0 28.27
## 1397064 EconŰmica 316308.3  3.1
## 1397065 EconŰmica 319551.4 14.97
## 1397066 EconŰmica 322922.0 28.35
## 1397067 EconŰmica 275690.8 30.68
## 1397068 EconŰmica 300287.6 14.44
```

```
str(sample)
```

```
## 'data.frame': 1397068 obs. of 13 variables:
## $ AÑO : Factor w/ 9 levels "2013","2014",...: 1 1 1 1 1 1 1 1 1 ...
## $ MES_ACUMULADO : Factor w/ 12 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 ...
## $ CLAVE_ENTIDAD : Factor w/ 32 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 ...
## $ CLAVE_MUNICIPIO: Factor w/ 316 levels "0","1","2","3",...: 2 2 2 2 2 6 7 12 12 12 ...
## $ MODALIDAD : Factor w/ 4 levels "Mejoramientos",...: 3 3 3 3 3 3 3 3 3 ...
## $ DESTINO : Factor w/ 6 levels "AutoproducciŰn",...: 5 5 5 5 5 5 5 5 5 ...
## $ TIPO : Factor w/ 2 levels "Cofinanciamientos y subsidios",...: 1 1 1 1 1 1 1 1 1 ...
## $ GENERO : Factor w/ 3 levels "Hombre","Mujer",...: 1 1 1 1 1 1 1 1 1 ...
## $ EDAD : Factor w/ 4 levels "29 o menos","30 a 59",...: 1 1 1 1 1 1 1 1 1 ...
## $ INGRESO : Factor w/ 7 levels "2.6 o menos",...: 2 3 4 6 6 4 5 3 3 5 ...
## $ VALOR_VIVIENDA : Factor w/ 7 levels "EconŰmica","Media",...: 7 4 7 7 2 1 7 4 7 7 ...
## $ MONTO : num 283223 1365781 803068 671815 1650391 ...
## $ CLAVE : Factor w/ 1911 levels "1.0","1.1","1.10",...: 2 2 2 2 2 8 9 4 4 4 ...
```

```
summary(sample)
```

```
##      AÑO MES_ACUMULADO CLAVE_ENTIDAD CLAVE_MUNICIPIO
## 2019 :188218 12 :133591 15 :149449 6 : 47494
## 2018 :186917 10 :125325 19 :115360 5 : 46046
## 2015 :186073 8 :119602 14 : 99518 2 : 44994
## 2017 :179491 2 :119336 30 : 81642 1 : 41861
```



```

## 2020 :179336 6 :118882 5 : 74938 4 : 37877
## 2016 :156016 9 :118240 11 : 64865 18 : 36384
## (Other):321017 (Other):662092 (Other):811296 (Other):1142412
## MODALIDAD DESTINO
## Mejoramientos :361602 Autoproducción : 5
## Otros programas : 9766 Con disponibilidad de terreno: 16626
## Viviendas nuevas:496191 Mejoramientos :361602
## Viviendas usadas:529509 Pago de pasivos : 9766
## Vivienda nueva :479560
## Vivienda usada :529509
##
## TIPO GENERO
## Cofinanciamientos y subsidios:500133 Hombre :858539
## Credito individual :896935 Mujer :531362
## No disponible: 7167
##
##
##
## EDAD INGRESO VALOR_VIVIENDA
## 29 o menos :417938 2.6 o menos :318690 Económica : 65335
## 30 a 59 :945999 2.61 a 4.00 :270913 Media :197201
## 60 o m's : 12652 4.01 a 6.00 :244657 No disponible :324806
## No disponible: 20479 6.01 a 9.00 :222882 Popular :376124
## 9.01 a 12.00 :134704 Residencial : 48282
## M's de 12 :205114 Residencial plus: 9013
## No disponible: 108 Tradicional :376307
## MONTO CLAVE
## Min. : -424633 14.120 : 16668
## 1st Qu.: 58243 2.4 : 16586
## Median : 352699 11.20 : 16072
## Mean : 763071 8.19 : 15985
## 3rd Qu.: 740000 19.6 : 15464
## Max. :180487760 22.14 : 15023
## (Other):1301270

```

```
# GRÁFICAS DE GABRIEL* Pegar su code*
```

2.2 Adjust Sample to Match Population's Characteristics

```
# Berny
```

2.3 Assign Variable Status

```
# Luly
```

2.4 Subsidies

3 Credit Risk Model

3.1 Logistic Regression

3.2 Evaluation

3.3 Regional Analysis & Clusters

4 Building a Cross-Subsidy Scheme

4.1 Assigning an Interest Rate

5 Comparative Statics & Tests