Data Cleaning

NurAmniBahirah

5/31/2021

# 1. Data Cleaning

First, we read the dataset file. The data has been downloaded to the working directory. We also need to select only the rows and columns that are relevant to reduce the size of the dataset.

#set working directory  
setwd("~/R/GroupProject")  
  
#read the csv file  
data <- read.csv("unemployment\_rate.csv")

### Data Before Cleaning

The original dataset:

#show the dataset  
data

## Year State.Country Unemployment.Rate..Percentage.  
## 1 2019 Johor 2.7  
## 2 2019 Kedah 2.8  
## 3 2019 Kelantan 4.0  
## 4 2019 Malaysia 3.3  
## 5 2019 Melaka 1.1  
## 6 2019 Negeri Sembilan 3.1  
## 7 2019 Pahang 2.9  
## 8 2019 Perak 3.4  
## 9 2019 Perlis 3.3  
## 10 2019 Pulau Pinang 2.0  
## 11 2019 Sabah 5.8  
## 12 2019 Sarawak 3.1  
## 13 2019 Selangor 2.9  
## 14 2019 Terengganu 3.5  
## 15 2019 W.P Labuan 6.9  
## 16 2019 W.P. Kuala Lumpur 2.6  
## 17 2019 W.P Putrajaya 1.2  
## 18 2018 Johor 3.0  
## 19 2018 Kedah 2.9  
## 20 2018 Kelantan 4.0  
## 21 2018 Malaysia 3.3  
## 22 2018 Melaka 1.1  
## 23 2018 Negeri Sembilan 3.0  
## 24 2018 Pahang 2.6  
## 25 2018 Perak 3.3  
## 26 2018 Perlis 3.4  
## 27 2018 Pulau Pinang 2.2  
## 28 2018 Sabah 5.8  
## 29 2018 Sarawak 3.2  
## 30 2018 Selangor 2.8  
## 31 2018 Terengganu 4.8  
## 32 2018 W.P Labuan 7.1  
## 33 2018 W.P. Kuala Lumpur 2.4  
## 34 2018 W.P Putrajaya 1.3  
## 35 2017 Johor 3.4  
## 36 2017 Kedah 2.8  
## 37 2017 Kelantan 3.6  
## 38 2017 Malaysia 3.4  
## 39 2017 Melaka 1.0  
## 40 2017 Negeri Sembilan 2.8  
## 41 2017 Pahang 2.9  
## 42 2017 Perak 3.7  
## 43 2017 Perlis 3.5  
## 44 2017 Pulau Pinang 2.1  
## 45 2017 Sabah 5.6  
## 46 2017 Sarawak 3.0  
## 47 2017 Selangor 2.8  
## 48 2017 Terengganu 4.5  
## 49 2017 W.P Labuan 8.5  
## 50 2017 W.P. Kuala Lumpur 3.1  
## 51 2017 W.P Putrajaya 1.5  
## 52 2016 Johor 3.6  
## 53 2016 Kedah 2.9  
## 54 2016 Kelantan 3.8  
## 55 2016 Malaysia 3.4  
## 56 2016 Melaka 0.9  
## 57 2016 Negeri Sembilan 2.7  
## 58 2016 Pahang 2.6  
## 59 2016 Perak 3.4  
## 60 2016 Perlis 2.9  
## 61 2016 Pulau Pinang 2.1  
## 62 2016 Sabah 5.4  
## 63 2016 Sarawak 3.3  
## 64 2016 Selangor 3.2  
## 65 2016 Terengganu 4.2  
## 66 2016 W.P Labuan 7.8  
## 67 2016 W.P. Kuala Lumpur 3.3  
## 68 2016 W.P Putrajaya 1.7  
## 69 2015 Johor 3.1  
## 70 2015 Kedah 2.6  
## 71 2015 Kelantan 3.6  
## 72 2015 Malaysia 3.1  
## 73 2015 Melaka 1.0  
## 74 2015 Negeri Sembilan 3.1  
## 75 2015 Pahang 2.9  
## 76 2015 Perak 3.2  
## 77 2015 Perlis 2.9  
## 78 2015 Pulau Pinang 1.6  
## 79 2015 Sabah 5.0  
## 80 2015 Sarawak 3.5  
## 81 2015 Selangor 2.4  
## 82 2015 Terengganu 4.0  
## 83 2015 W.P Labuan 5.0  
## 84 2015 W.P. Kuala Lumpur 3.3  
## 85 2015 W.P Putrajaya 1.0  
## 86 2014 Johor 2.6  
## 87 2014 Kedah 2.7  
## 88 2014 Kelantan 3.3  
## 89 2014 Malaysia 2.9  
## 90 2014 Melaka 0.9  
## 91 2014 Negeri Sembilan 3.0  
## 92 2014 Pahang 2.7  
## 93 2014 Perak 3.4  
## 94 2014 Perlis 3.3  
## 95 2014 Pulau Pinang 1.6  
## 96 2014 Sabah 4.9  
## 97 2014 Sarawak 3.1  
## 98 2014 Selangor 2.0  
## 99 2014 Terengganu 4.1  
## 100 2014 W.P Labuan 5.1  
## 101 2014 W.P. Kuala Lumpur 2.9  
## 102 2014 W.P Putrajaya 1.2  
## 103 2013 Johor 2.8  
## 104 2013 Kedah 2.4  
## 105 2013 Kelantan 2.8  
## 106 2013 Malaysia 3.1  
## 107 2013 Melaka 0.7  
## 108 2013 Negeri Sembilan 3.1  
## 109 2013 Pahang 2.8  
## 110 2013 Perak 3.8  
## 111 2013 Perlis 3.7  
## 112 2013 Pulau Pinang 1.7  
## 113 2013 Sabah 5.3  
## 114 2013 Sarawak 3.7  
## 115 2013 Selangor 2.4  
## 116 2013 Terengganu 3.4  
## 117 2013 W.P Labuan 4.1  
## 118 2013 W.P. Kuala Lumpur 3.3  
## 119 2013 W.P.Putrajaya 1.2  
## 120 2012 Johor 3.1  
## 121 2012 Kedah 2.9  
## 122 2012 Kelantan 2.3  
## 123 2012 Malaysia 3.0  
## 124 2012 Melaka 0.5  
## 125 2012 Negeri Sembilan 3.0  
## 126 2012 Pahang 2.7  
## 127 2012 Perak 3.0  
## 128 2012 Perlis 3.9  
## 129 2012 Pulau Pinang 2.0  
## 130 2012 Sabah 5.6  
## 131 2012 Sarawak 3.5  
## 132 2012 Selangor 2.3  
## 133 2012 Terengganu 3.0  
## 134 2012 W.P Labuan 4.6  
## 135 2012 W.P. Kuala Lumpur 2.7  
## 136 2012 W.P.Putrajaya 1.5  
## 137 2011 Johor 2.5  
## 138 2011 Kedah 3.5  
## 139 2011 Kelantan 2.0  
## 140 2011 Malaysia 3.1  
## 141 2011 Melaka 0.7  
## 142 2011 Negeri Sembilan 3.0  
## 143 2011 Pahang 2.6  
## 144 2011 Perak 2.6  
## 145 2011 Perlis 4.0  
## 146 2011 Pulau Pinang 1.9  
## 147 2011 Sabah 5.6  
## 148 2011 Sarawak 4.1  
## 149 2011 Selangor 2.4  
## 150 2011 Terengganu 3.2  
## 151 2011 W.P Labuan 4.2  
## 152 2011 W.P. Kuala Lumpur 3.0  
## 153 2011 W.P.Putrajaya 1.1  
## 154 2010 Johor 2.4  
## 155 2010 Kedah 2.8  
## 156 2010 Kelantan 2.4  
## 157 2010 Malaysia 3.3  
## 158 2010 Melaka 1.0  
## 159 2010 Negeri Sembilan 3.1  
## 160 2010 Pahang 2.9  
## 161 2010 Perak 3.0  
## 162 2010 Perlis 3.3  
## 163 2010 Pulau Pinang 2.2  
## 164 2010 Sabah 5.5  
## 165 2010 Sarawak 4.4  
## 166 2010 Selangor 3.2  
## 167 2010 Terengganu 3.7  
## 168 2010 W.P Labuan 6.0  
## 169 2010 W.P. Kuala Lumpur 3.1  
## 170 2010 W.P.Putrajaya NA  
## 171 2009 Johor 3.2  
## 172 2009 Kedah 4.2  
## 173 2009 Kelantan 2.4  
## 174 2009 Malaysia 3.7  
## 175 2009 Melaka 2.4  
## 176 2009 Negeri Sembilan 3.9  
## 177 2009 Pahang 3.0  
## 178 2009 Perak 3.5  
## 179 2009 Perlis 3.3  
## 180 2009 Pulau Pinang 2.5  
## 181 2009 Sabah 5.5  
## 182 2009 Sarawak 4.6  
## 183 2009 Selangor 3.6  
## 184 2009 Terengganu 3.8  
## 185 2009 W.P Labuan 5.6  
## 186 2009 W.P. Kuala Lumpur 2.9  
## 187 2009 W.P.Putrajaya NA  
## 188 2008 Johor 2.2  
## 189 2008 Kedah 3.8  
## 190 2008 Kelantan 2.9  
## 191 2008 Malaysia 3.3  
## 192 2008 Melaka 1.1  
## 193 2008 Negeri Sembilan 3.1  
## 194 2008 Pahang 2.6  
## 195 2008 Perak 3.8  
## 196 2008 Perlis 2.9  
## 197 2008 Pulau Pinang 2.0  
## 198 2008 Sabah 4.9  
## 199 2008 Sarawak 5.3  
## 200 2008 Selangor 3.1  
## 201 2008 Terengganu 3.4  
## 202 2008 W.P Labuan 5.6  
## 203 2008 W.P. Kuala Lumpur 2.7  
## 204 2008 W.P.Putrajaya NA  
## 205 2007 Johor 2.0  
## 206 2007 Kedah 3.3  
## 207 2007 Kelantan 2.7  
## 208 2007 Malaysia 3.2  
## 209 2007 Melaka 2.1  
## 210 2007 Negeri Sembilan 3.6  
## 211 2007 Pahang 3.2  
## 212 2007 Perak 2.9  
## 213 2007 Perlis 3.0  
## 214 2007 Pulau Pinang 2.1  
## 215 2007 Sabah 5.5  
## 216 2007 Sarawak 4.3  
## 217 2007 Selangor 3.0  
## 218 2007 Terengganu 2.6  
## 219 2007 W.P Labuan 4.5  
## 220 2007 W.P. Kuala Lumpur 2.9  
## 221 2007 W.P.Putrajaya NA  
## 222 2006 Johor 2.1  
## 223 2006 Kedah 3.8  
## 224 2006 Kelantan 2.4  
## 225 2006 Malaysia 3.3  
## 226 2006 Melaka 1.3  
## 227 2006 Negeri Sembilan 4.0  
## 228 2006 Pahang 2.4  
## 229 2006 Perak 3.5  
## 230 2006 Perlis 3.2  
## 231 2006 Pulau Pinang 1.7  
## 232 2006 Sabah 5.7  
## 233 2006 Sarawak 4.1  
## 234 2006 Selangor 3.2  
## 235 2006 Terengganu 3.6  
## 236 2006 W.P Labuan 9.4  
## 237 2006 W.P. Kuala Lumpur 2.8  
## 238 2006 W.P.Putrajaya NA  
## 239 2005 Johor 3.3  
## 240 2005 Kedah 3.7  
## 241 2005 Kelantan 3.4  
## 242 2005 Malaysia 3.5  
## 243 2005 Melaka 1.2  
## 244 2005 Negeri Sembilan 3.6  
## 245 2005 Pahang 3.3  
## 246 2005 Perak 3.3  
## 247 2005 Perlis 2.9  
## 248 2005 Pulau Pinang 2.1  
## 249 2005 Sabah 6.4  
## 250 2005 Sarawak 4.1  
## 251 2005 Selangor 3.1  
## 252 2005 Terengganu 3.1  
## 253 2005 W.P Labuan 5.1  
## 254 2005 W.P. Kuala Lumpur 2.3  
## 255 2005 W.P.Putrajaya NA  
## 256 2004 Johor 3.6  
## 257 2004 Kedah 3.8  
## 258 2004 Kelantan 4.1  
## 259 2004 Malaysia 3.5  
## 260 2004 Melaka 1.6  
## 261 2004 Negeri Sembilan 3.5  
## 262 2004 Pahang 3.3  
## 263 2004 Perak 3.3  
## 264 2004 Perlis 3.3  
## 265 2004 Pulau Pinang 1.8  
## 266 2004 Sabah 6.5  
## 267 2004 Sarawak 3.8  
## 268 2004 Selangor 2.7  
## 269 2004 Terengganu 3.2  
## 270 2004 W.P Labuan 5.0  
## 271 2004 W.P. Kuala Lumpur 2.9  
## 272 2004 W.P.Putrajaya NA  
## 273 2003 Johor 3.5  
## 274 2003 Kedah 3.3  
## 275 2003 Kelantan 3.2  
## 276 2003 Malaysia 3.6  
## 277 2003 Melaka 2.8  
## 278 2003 Negeri Sembilan 3.5  
## 279 2003 Pahang 3.6  
## 280 2003 Perak 4.8  
## 281 2003 Perlis 4.4  
## 282 2003 Pulau Pinang 1.7  
## 283 2003 Sabah 5.4  
## 284 2003 Sarawak 3.9  
## 285 2003 Selangor 3.2  
## 286 2003 Terengganu 2.9  
## 287 2003 W.P Labuan 4.5  
## 288 2003 W.P. Kuala Lumpur 3.4  
## 289 2003 W.P.Putrajaya NA  
## 290 2002 Johor 2.9  
## 291 2002 Kedah 3.2  
## 292 2002 Kelantan 3.9  
## 293 2002 Malaysia 3.5  
## 294 2002 Melaka 2.9  
## 295 2002 Negeri Sembilan 3.7  
## 296 2002 Pahang 2.9  
## 297 2002 Perak 3.9  
## 298 2002 Perlis 3.1  
## 299 2002 Pulau Pinang 2.2  
## 300 2002 Sabah 5.9  
## 301 2002 Sarawak 4.0  
## 302 2002 Selangor 3.0  
## 303 2002 Terengganu 3.2  
## 304 2002 W.P Labuan 3.2  
## 305 2002 W.P. Kuala Lumpur 2.8  
## 306 2002 W.P.Putrajaya NA  
## 307 2001 Johor 3.4  
## 308 2001 Kedah 3.6  
## 309 2001 Kelantan 3.1  
## 310 2001 Malaysia 3.5  
## 311 2001 Melaka 1.8  
## 312 2001 Negeri Sembilan 3.2  
## 313 2001 Pahang 3.1  
## 314 2001 Perak 4.0  
## 315 2001 Perlis 3.5  
## 316 2001 Pulau Pinang 2.5  
## 317 2001 Sabah 6.6  
## 318 2001 Sarawak 4.4  
## 319 2001 Selangor 2.4  
## 320 2001 Terengganu 2.7  
## 321 2001 W.P Labuan 5.4  
## 322 2001 W.P. Kuala Lumpur 2.7  
## 323 2001 W.P.Putrajaya NA  
## 324 2000 Johor 2.2  
## 325 2000 Kedah 2.9  
## 326 2000 Kelantan 2.6  
## 327 2000 Malaysia 3.0  
## 328 2000 Melaka 1.3  
## 329 2000 Negeri Sembilan 3.4  
## 330 2000 Pahang 2.7  
## 331 2000 Perak 3.5  
## 332 2000 Perlis 2.9  
## 333 2000 Pulau Pinang 1.9  
## 334 2000 Sabah 5.8  
## 335 2000 Sarawak 4.3  
## 336 2000 Selangor 2.0  
## 337 2000 Terengganu 3.0  
## 338 2000 W.P Labuan 4.7  
## 339 2000 W.P. Kuala Lumpur 2.6  
## 340 2000 W.P.Putrajaya NA  
## 341 1999 Johor 3.0  
## 342 1999 Kedah 3.2  
## 343 1999 Kelantan 3.2  
## 344 1999 Malaysia 3.4  
## 345 1999 Melaka 1.9  
## 346 1999 Negeri Sembilan 3.8  
## 347 1999 Pahang 3.0  
## 348 1999 Perak 4.3  
## 349 1999 Perlis 3.4  
## 350 1999 Pulau Pinang 2.4  
## 351 1999 Sabah 5.6  
## 352 1999 Sarawak 4.6  
## 353 1999 Selangor 1.9  
## 354 1999 Terengganu 3.9  
## 355 1999 W.P Labuan 4.0  
## 356 1999 W.P. Kuala Lumpur 2.9  
## 357 1999 W.P.Putrajaya NA  
## 358 1998 Johor 1.7  
## 359 1998 Kedah 3.4  
## 360 1998 Kelantan 4.1  
## 361 1998 Malaysia 3.2  
## 362 1998 Melaka 2.5  
## 363 1998 Negeri Sembilan 3.6  
## 364 1998 Pahang 2.6  
## 365 1998 Perak 3.9  
## 366 1998 Perlis 4.3  
## 367 1998 Pulau Pinang 2.3  
## 368 1998 Sabah 4.9  
## 369 1998 Sarawak 5.0  
## 370 1998 Selangor 1.8  
## 371 1998 Terengganu 4.0  
## 372 1998 W.P Labuan 2.5  
## 373 1998 W.P. Kuala Lumpur 2.5  
## 374 1998 W.P.Putrajaya NA  
## 375 1997 Johor 1.7  
## 376 1997 Kedah 2.5  
## 377 1997 Kelantan 3.6  
## 378 1997 Malaysia 2.4  
## 379 1997 Melaka 1.7  
## 380 1997 Negeri Sembilan 2.6  
## 381 1997 Pahang 3.0  
## 382 1997 Perak 2.2  
## 383 1997 Perlis 2.3  
## 384 1997 Pulau Pinang 1.1  
## 385 1997 Sabah 4.6  
## 386 1997 Sarawak 3.9  
## 387 1997 Selangor 1.0  
## 388 1997 Terengganu 3.6  
## 389 1997 W.P Labuan 3.3  
## 390 1997 W.P. Kuala Lumpur 1.4  
## 391 1997 W.P.Putrajaya NA  
## 392 1996 Johor 2.1  
## 393 1996 Kedah 2.5  
## 394 1996 Kelantan 3.3  
## 395 1996 Malaysia 2.5  
## 396 1996 Melaka 1.5  
## 397 1996 Negeri Sembilan 2.5  
## 398 1996 Pahang 2.6  
## 399 1996 Perak 3.2  
## 400 1996 Perlis 2.0  
## 401 1996 Pulau Pinang 0.7  
## 402 1996 Sabah 3.9  
## 403 1996 Sarawak 4.4  
## 404 1996 Selangor 0.9  
## 405 1996 Terengganu 3.3  
## 406 1996 W.P Labuan 5.1  
## 407 1996 W.P. Kuala Lumpur 1.8  
## 408 1996 W.P.Putrajaya NA  
## 409 1995 Johor 2.6  
## 410 1995 Kedah 2.0  
## 411 1995 Kelantan 3.5  
## 412 1995 Malaysia 3.1  
## 413 1995 Melaka 2.3  
## 414 1995 Negeri Sembilan 3.1  
## 415 1995 Pahang 3.6  
## 416 1995 Perak 3.9  
## 417 1995 Perlis 1.4  
## 418 1995 Pulau Pinang 1.4  
## 419 1995 Sabah 5.5  
## 420 1995 Sarawak 4.8  
## 421 1995 Selangor 2.0  
## 422 1995 Terengganu 5.9  
## 423 1995 W.P Labuan 4.7  
## 424 1995 W.P. Kuala Lumpur 1.7  
## 425 1995 W.P.Putrajaya NA  
## 426 1993 Johor 2.9  
## 427 1993 Kedah 3.6  
## 428 1993 Kelantan 4.7  
## 429 1993 Malaysia 4.1  
## 430 1993 Melaka 3.2  
## 431 1993 Negeri Sembilan 2.8  
## 432 1993 Pahang 3.7  
## 433 1993 Perak 4.3  
## 434 1993 Perlis 1.4  
## 435 1993 Pulau Pinang 2.2  
## 436 1993 Sabah 7.1  
## 437 1993 Sarawak 7.0  
## 438 1993 Selangor 2.7  
## 439 1993 Terengganu 7.0  
## 440 1993 W.P Labuan 5.9  
## 441 1993 W.P. Kuala Lumpur 3.0  
## 442 1993 W.P.Putrajaya NA  
## 443 1992 Johor 2.2  
## 444 1992 Kedah 3.7  
## 445 1992 Kelantan 5.3  
## 446 1992 Malaysia 3.7  
## 447 1992 Melaka 3.3  
## 448 1992 Negeri Sembilan 3.3  
## 449 1992 Pahang 2.8  
## 450 1992 Perak 3.7  
## 451 1992 Perlis 2.7  
## 452 1992 Pulau Pinang 1.7  
## 453 1992 Sabah 6.2  
## 454 1992 Sarawak 7.8  
## 455 1992 Selangor 2.1  
## 456 1992 Terengganu 5.5  
## 457 1992 W.P Labuan 3.3  
## 458 1992 W.P. Kuala Lumpur 2.4  
## 459 1992 W.P.Putrajaya NA  
## 460 1990 Johor 3.7  
## 461 1990 Kedah 4.4  
## 462 1990 Kelantan 3.4  
## 463 1990 Malaysia 4.5  
## 464 1990 Melaka 4.2  
## 465 1990 Negeri Sembilan 3.8  
## 466 1990 Pahang 3.1  
## 467 1990 Perak 4.1  
## 468 1990 Perlis 4.3  
## 469 1990 Pulau Pinang 2.8  
## 470 1990 Sabah 9.0  
## 471 1990 Sarawak 9.9  
## 472 1990 Selangor 2.1  
## 473 1990 Terengganu 6.9  
## 474 1990 W.P Labuan 5.7  
## 475 1990 W.P. Kuala Lumpur 2.6  
## 476 1990 W.P.Putrajaya NA  
## 477 1989 Johor 4.6  
## 478 1989 Kedah 5.7  
## 479 1989 Kelantan 6.3  
## 480 1989 Malaysia 5.7  
## 481 1989 Melaka 6.0  
## 482 1989 Negeri Sembilan 4.9  
## 483 1989 Pahang 3.7  
## 484 1989 Perak 5.5  
## 485 1989 Perlis 6.4  
## 486 1989 Pulau Pinang 3.9  
## 487 1989 Sabah 7.4  
## 488 1989 Sarawak 9.6  
## 489 1989 Selangor 3.9  
## 490 1989 Terengganu 12.3  
## 491 1989 W.P Labuan 8.6  
## 492 1989 W.P. Kuala Lumpur 3.7  
## 493 1989 W.P.Putrajaya NA  
## 494 1988 Johor 5.4  
## 495 1988 Kedah 7.4  
## 496 1988 Kelantan 7.7  
## 497 1988 Malaysia 7.2  
## 498 1988 Melaka 8.1  
## 499 1988 Negeri Sembilan 6.0  
## 500 1988 Pahang 6.2  
## 501 1988 Perak 8.1  
## 502 1988 Perlis 7.6  
## 503 1988 Pulau Pinang 6.7  
## 504 1988 Sabah 8.2  
## 505 1988 Sarawak 10.3  
## 506 1988 Selangor 5.9  
## 507 1988 Terengganu 9.4  
## 508 1988 W.P Labuan 10.9  
## 509 1988 W.P. Kuala Lumpur 5.9  
## 510 1988 W.P.Putrajaya NA  
## 511 1987 Johor 6.3  
## 512 1987 Kedah 8.0  
## 513 1987 Kelantan 6.5  
## 514 1987 Malaysia 7.3  
## 515 1987 Melaka 8.0  
## 516 1987 Negeri Sembilan 7.6  
## 517 1987 Pahang 5.1  
## 518 1987 Perak 8.0  
## 519 1987 Perlis 6.6  
## 520 1987 Pulau Pinang 6.5  
## 521 1987 Sabah 9.0  
## 522 1987 Sarawak 9.8  
## 523 1987 Selangor 5.6  
## 524 1987 Terengganu 11.7  
## 525 1987 W.P Labuan 9.7  
## 526 1987 W.P. Kuala Lumpur 5.7  
## 527 1987 W.P.Putrajaya NA  
## 528 1986 Johor 6.8  
## 529 1986 Kedah 8.1  
## 530 1986 Kelantan 6.8  
## 531 1986 Malaysia 7.4  
## 532 1986 Melaka 9.5  
## 533 1986 Negeri Sembilan 7.0  
## 534 1986 Pahang 4.9  
## 535 1986 Perak 8.7  
## 536 1986 Perlis 8.0  
## 537 1986 Pulau Pinang 8.4  
## 538 1986 Sabah 8.2  
## 539 1986 Sarawak 9.0  
## 540 1986 Selangor 5.7  
## 541 1986 Terengganu 9.5  
## 542 1986 W.P Labuan 10.9  
## 543 1986 W.P. Kuala Lumpur 5.5  
## 544 1986 W.P.Putrajaya NA  
## 545 1985 Johor 5.3  
## 546 1985 Kedah 7.1  
## 547 1985 Kelantan 6.0  
## 548 1985 Malaysia 5.6  
## 549 1985 Melaka 5.5  
## 550 1985 Negeri Sembilan 4.7  
## 551 1985 Pahang 4.3  
## 552 1985 Perak 6.1  
## 553 1985 Perlis 7.6  
## 554 1985 Pulau Pinang 5.0  
## 555 1985 Sabah 6.4  
## 556 1985 Sarawak 7.1  
## 557 1985 Selangor 4.7  
## 558 1985 Terengganu 7.2  
## 559 1985 W.P Labuan 7.7  
## 560 1985 W.P. Kuala Lumpur 3.5  
## 561 1985 W.P.Putrajaya NA  
## 562 1984 Johor 4.5  
## 563 1984 Kedah 5.6  
## 564 1984 Kelantan 5.5  
## 565 1984 Malaysia 5.0  
## 566 1984 Melaka 5.4  
## 567 1984 Negeri Sembilan 4.9  
## 568 1984 Pahang 4.6  
## 569 1984 Perak 5.4  
## 570 1984 Perlis 4.6  
## 571 1984 Pulau Pinang 3.6  
## 572 1984 Sabah 6.0  
## 573 1984 Sarawak 6.7  
## 574 1984 Selangor 4.0  
## 575 1984 Terengganu 6.8  
## 576 1984 W.P Labuan 7.6  
## 577 1984 W.P. Kuala Lumpur 3.8  
## 578 1984 W.P.Putrajaya NA  
## 579 1983 Johor 4.2  
## 580 1983 Kedah 3.8  
## 581 1983 Kelantan 6.2  
## 582 1983 Malaysia 3.8  
## 583 1983 Melaka 2.8  
## 584 1983 Negeri Sembilan 3.8  
## 585 1983 Pahang 2.7  
## 586 1983 Perak 5.1  
## 587 1983 Perlis 3.2  
## 588 1983 Pulau Pinang 3.8  
## 589 1983 Sabah 4.1  
## 590 1983 Sarawak 5.2  
## 591 1983 Selangor 2.3  
## 592 1983 Terengganu 2.5  
## 593 1983 W.P Labuan NA  
## 594 1983 W.P. Kuala Lumpur 2.0  
## 595 1983 W.P.Putrajaya NA  
## 596 1982 Johor 3.5  
## 597 1982 Kedah 3.7  
## 598 1982 Kelantan 4.8  
## 599 1982 Malaysia 3.4  
## 600 1982 Melaka 4.0  
## 601 1982 Negeri Sembilan 2.9  
## 602 1982 Pahang 3.4  
## 603 1982 Perak 5.4  
## 604 1982 Perlis 4.1  
## 605 1982 Pulau Pinang 3.5  
## 606 1982 Sabah 3.0  
## 607 1982 Sarawak 2.8  
## 608 1982 Selangor 2.2  
## 609 1982 Terengganu 2.7  
## 610 1982 W.P Labuan NA  
## 611 1982 W.P. Kuala Lumpur 1.9  
## 612 1982 W.P.Putrajaya NA

We can see that the dataset contains the unemployment rate from the Year 1982 to 2019. In this project, we chose to start from the Year 2014 to 2019. Therefore, we need to clean the data. First, we rename the column to “Year”, “State” and the “Unemployment\_Rate”. Then by using the ‘dplyr’ packages, we remove the unwanted rows from Year 1982 to 2013. Lastly, using transpose to clean the data.

#rename the Columns  
names(data)[1] <- "Year"  
names(data)[2] <- "State"  
names(data)[3] <- "Unemployment\_Rate"  
head(data)

## Year State Unemployment\_Rate  
## 1 2019 Johor 2.7  
## 2 2019 Kedah 2.8  
## 3 2019 Kelantan 4.0  
## 4 2019 Malaysia 3.3  
## 5 2019 Melaka 1.1  
## 6 2019 Negeri Sembilan 3.1

#remove the Years from Year 1982 to 2013  
library(dplyr)

## Warning: package 'dplyr' was built under R version 4.0.5

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

data2 <- anti\_join(data,data[52:612,])

## Joining, by = c("Year", "State", "Unemployment\_Rate")

data2

## Year State Unemployment\_Rate  
## 1 2019 Johor 2.7  
## 2 2019 Kedah 2.8  
## 3 2019 Kelantan 4.0  
## 4 2019 Malaysia 3.3  
## 5 2019 Melaka 1.1  
## 6 2019 Negeri Sembilan 3.1  
## 7 2019 Pahang 2.9  
## 8 2019 Perak 3.4  
## 9 2019 Perlis 3.3  
## 10 2019 Pulau Pinang 2.0  
## 11 2019 Sabah 5.8  
## 12 2019 Sarawak 3.1  
## 13 2019 Selangor 2.9  
## 14 2019 Terengganu 3.5  
## 15 2019 W.P Labuan 6.9  
## 16 2019 W.P. Kuala Lumpur 2.6  
## 17 2019 W.P Putrajaya 1.2  
## 18 2018 Johor 3.0  
## 19 2018 Kedah 2.9  
## 20 2018 Kelantan 4.0  
## 21 2018 Malaysia 3.3  
## 22 2018 Melaka 1.1  
## 23 2018 Negeri Sembilan 3.0  
## 24 2018 Pahang 2.6  
## 25 2018 Perak 3.3  
## 26 2018 Perlis 3.4  
## 27 2018 Pulau Pinang 2.2  
## 28 2018 Sabah 5.8  
## 29 2018 Sarawak 3.2  
## 30 2018 Selangor 2.8  
## 31 2018 Terengganu 4.8  
## 32 2018 W.P Labuan 7.1  
## 33 2018 W.P. Kuala Lumpur 2.4  
## 34 2018 W.P Putrajaya 1.3  
## 35 2017 Johor 3.4  
## 36 2017 Kedah 2.8  
## 37 2017 Kelantan 3.6  
## 38 2017 Malaysia 3.4  
## 39 2017 Melaka 1.0  
## 40 2017 Negeri Sembilan 2.8  
## 41 2017 Pahang 2.9  
## 42 2017 Perak 3.7  
## 43 2017 Perlis 3.5  
## 44 2017 Pulau Pinang 2.1  
## 45 2017 Sabah 5.6  
## 46 2017 Sarawak 3.0  
## 47 2017 Selangor 2.8  
## 48 2017 Terengganu 4.5  
## 49 2017 W.P Labuan 8.5  
## 50 2017 W.P. Kuala Lumpur 3.1  
## 51 2017 W.P Putrajaya 1.5

### Data After Cleaning

#Using transpose data  
data3 <- tidyr::spread(data2, Year, Unemployment\_Rate)  
data3

## State 2017 2018 2019  
## 1 Johor 3.4 3.0 2.7  
## 2 Kedah 2.8 2.9 2.8  
## 3 Kelantan 3.6 4.0 4.0  
## 4 Malaysia 3.4 3.3 3.3  
## 5 Melaka 1.0 1.1 1.1  
## 6 Negeri Sembilan 2.8 3.0 3.1  
## 7 Pahang 2.9 2.6 2.9  
## 8 Perak 3.7 3.3 3.4  
## 9 Perlis 3.5 3.4 3.3  
## 10 Pulau Pinang 2.1 2.2 2.0  
## 11 Sabah 5.6 5.8 5.8  
## 12 Sarawak 3.0 3.2 3.1  
## 13 Selangor 2.8 2.8 2.9  
## 14 Terengganu 4.5 4.8 3.5  
## 15 W.P Labuan 8.5 7.1 6.9  
## 16 W.P Putrajaya 1.5 1.3 1.2  
## 17 W.P. Kuala Lumpur 3.1 2.4 2.6