EV Power - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
library(tidyverse)
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0
──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.5.2     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.0
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts()
──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```
library(stringr)
```

Part 1: Defining Research Question

Chosen Question: What is the geographic distribution of the share of renewable energy in the United States? Is there a correlation between the share of renewable energy used and the price of renewable energy across states?

Part 2: Data Preparation and Cleaning

```
# getting working directory
getwd()
```

```
[1] "/Users/jacobbjurstrom/ev-power-JacobBjurstrom"
```

```
# listing files in current working directory
list.files()
```

```
[1] "data"                "ev-dashboard.qmd"    "project.qmd"
[4] "README.md"           "report_jb.pdf"       "report_jb.qmd"
[7] "report_jb.rmarkdown" "report.qmd"          "worksheet.qmd"
```

```
# setting new working directory
setwd("/Users/jacobbjurstrom/ev-power-JacobBjurstrom/data")

# loading data for renewable energy use between 2021 - 2023
renew_use_21 <- read.csv("renew-use-2021.csv")
renew_use_22 <- read.csv("renew-use-2022.csv")
renew_use_23 <- read.csv("renew-use-2023.csv")

# loading data for total energy use between 2021 - 2023
tot_use_21 <- read.csv("total-use-2021.csv")
tot_use_22 <- read.csv("total-use-2022.csv")
tot_use_23 <- read.csv("total-use-2023.csv")

# loading data for average energy prices and ev-registrations
av_en_price_21_23 <- read.csv("av-energy-price-2021-2023.csv", header = FALSE)
ev_regis <- read.csv("ev-registrations-by-state-2023.csv")

## Better understanding structure of each dataset
str(renew_use_21)
```

```
'data.frame':   260 obs. of  3 variables:
 $ State             : chr  "AK" "AK" "AK" "AK" ...
 $ Energy_Source     : chr  "Biomass" "Geothermal" "Hydropower" "Solar
Energy" ...
 $ Renewable_Use_2021: chr  "≈3153" "186 MMBtu" "5763 about" "~45" ...
```

```
str(renew_use_22)
```

```
'data.frame':   260 obs. of  3 variables:
 $ State             : chr  "AK" "AK" "AK" "AK" ...
 $ Energy_Source     : chr  "Biomass" "Geothermal" "Hydropower" "Solar
Energy" ...
 $ Renewable_Use_2022: chr  "≈3846" "$186" "$5846" "~57" ...
```

```
str(renew_use_23)
```

```
'data.frame':   260 obs. of  3 variables:
 $ State             : chr  "AK" "AK" "AK" "AK" ...
 $ Energy_Source     : chr  "Biomass" "Geothermal" "Hydropower" "Solar
Energy" ...
 $ Renewable_Use_2023: chr  "3404 kWh" "186.0" "6051" "67" ...
```

```
head(renew_use_21)
```

```
  State Energy_Source Renewable_Use_2021
1    AK       Biomass              ≈3153
2    AK    Geothermal           186 MMBtu
3    AK    Hydropower          5763 about
4    AK  Solar Energy                ~45
5    AK   Wind Energy             451 USD
6    AL       Biomass         198543 est.
```

```
head(renew_use_22)
```

```
  State Energy_Source Renewable_Use_2022
1    AK       Biomass              ≈3846
2    AK    Geothermal               $186
3    AK    Hydropower              $5846
4    AK  Solar Energy                ~57
5    AK   Wind Energy               $475
6    AL       Biomass          193932 USD
```

```
head(renew_use_23)
```

```
  State Energy_Source Renewable_Use_2023
1    AK       Biomass           3404 kWh
2    AK    Geothermal              186.0
3    AK    Hydropower               6051
4    AK  Solar Energy                 67
5    AK   Wind Energy                380
6    AL       Biomass          189040 kWh
```

```
str(tot_use_21)
```

```
'data.frame':   5 obs. of  53 variables:
 $ Energy_Source: chr  "Coal" "Natural Gas†" "Petroleum (BTU)" "nuclear" ...
 $ AK           : int  18694 395590 261094 0 9597
 $ AL           : int  309791 739891 583042 480115 239817
 $ AR           : int  216123 360545 328271 141372 89714
 $ AZ           : int  160299 484962 606862 329868 99266
 $ CA           : int  28244 2172757 2959389 171842 810020
 $ CO           : int  252442 509970 497788 0 103955
 $ CT           : int  2880 305184 284788 179551 49306
 $ DC           : int  0 28336 18439 0 2487
 $ DE           : int  4542 82708 113641 0 7150
 $ FL           : int  200193 1591864 1748346 307811 297291
 $ GA           : int  203870 773889 922503 354085 289113
 $ HI           : int  12566 133 223014 0 20134
 $ IA           : int  264419 383424 408385 0 389787
 $ ID           : int  3051 135176 188263 0 74428
 $ IL           : int  522809 1088485 1136797 1011555 224106
 $ IN           : int  753557 869328 712427 0 157324
 $ KS           : int  219031 291797 339006 89426 135551
 $ KY           : int  548443 365875 584011 0 71744
 $ LA           : int  95856 1862349 1840835 179886 135905
 $ MA           : int  0 404301 503312 0 75370
 $ MD           : int  69186 299282 433791 156369 52732
 $ ME           : int  1588 57233 163991 0 95141
 $ MI           : int  436203 950364 814081 358114 194075
 $ MN           : int  179055 523812 561731 147286 216113
 $ MO           : int  616413 293633 607276 44766 88879
 $ MS           : int  64446 576903 384328 122771 66134
 $ MT           : int  122765 87105 176686 0 56334
 $ NC           : int  222501 637553 884299 449675 196973
 $ ND           : int  361811 191168 168682 0 92653
 $ NE           : int  216298 191008 237214 71758 158275
 $ NH           : int  3259 60116 142030 102789 38479
 $ NJ           : int  12586 697019 749892 293494 70039
 $ NM           : int  133228 285809 262885 0 62210
 $ NV           : int  35910 305212 286548 0 63647
 $ NY           : int  5370 1359437 1237451 325141 263977
 $ OH           : int  575920 1294814 1028000 182330 146858
 $ OK           : int  131695 745911 517408 0 177087
 $ OR           : int  1303 305665 317322 0 225544
 $ PA           : int  485193 1868137 1047658 791587 179589
 $ RI           : int  0 105473 76464 0 11798
 $ SC           : int  162628 349990 508147 560782 143796
 $ SD           : int  21589 96787 119505 0 127382
 $ TN           : int  225784 413554 713210 368461 135841
 $ TX           : int  968401 4773076 6783182 419363 654199
 $ UT           : int  276159 274420 304823 0 36050
 $ VA           : int  68603 699927 795296 297972 174615
```

```
 $ VT           : int  0 13801 72241 0 21430
 $ WA           : int  36943 384769 711662 88764 394052
 $ WI           : int  286760 561076 533390 103979 145936
 $ WV           : int  633582 277002 205005 0 26427
 $ WY           : int  376971 161580 146274 0 37734
 $ US           : int  10548957 31688203 35250685 8130913 7646167
```

```
str(tot_use_22)
```

```
'data.frame':   5 obs. of  53 variables:
 $ Energy_Source: chr  "coal Consumption" "Natural-Gas" "petroleum (btu)"
"Nuclear Energy†" ...
 $ AK           : int  18615 437916 263335 0 10410
 $ AL           : int  297654 787300 578431 442093 232035
 $ AR           : int  211724 398099 327813 149654 90825
 $ AZ           : int  154007 468038 594859 333738 101215
 $ CA           : int  30049 2131372 3017944 183814 880995
 $ CO           : int  233256 524890 538413 0 114917
 $ CT           : int  0 307212 302881 172018 49084
 $ DC           : int  0 30174 18000 0 2622
 $ DE           : int  1846 89674 112026 0 7402
 $ FL           : int  171953 1659544 1815529 321468 304605
 $ GA           : int  180888 809618 940579 356001 293237
 $ HI           : int  7680 159 241994 0 20471
 $ IA           : int  227866 434374 423592 0 421784
 $ ID           : int  1881 141924 190635 0 78406
 $ IL           : int  496983 1134781 1138141 1032989 248541
 $ IN           : int  719238 913401 699235 0 170986
 $ KS           : int  226712 318779 346852 93844 151788
 $ KY           : int  523276 402534 580349 0 77517
 $ LA           : int  96914 2087166 1663129 168889 138209
 $ MA           : int  0 432442 529154 0 80700
 $ MD           : int  61932 310133 411842 154742 51255
 $ ME           : int  1269 62559 166724 0 93867
 $ MI           : int  423504 1087716 820709 271788 206811
 $ MN           : int  184517 535010 568916 153546 229769
 $ MO           : int  566940 322547 606374 92724 95312
 $ MS           : int  66214 617855 383366 89856 66614
 $ MT           : int  131345 93971 177009 0 60644
 $ NC           : int  163029 747187 906477 445547 198165
 $ ND           : int  369340 198986 170390 0 96024
 $ NE           : int  223571 199260 237556 58702 168382
 $ NH           : int  3864 60176 149025 114108 39863
 $ NJ           : int  6199 755048 769751 295875 73187
 $ NM           : int  138077 301279 255571 0 77286
 $ NV           : int  35835 302315 303234 0 72734
```

```
 $ NY           : int  6143 1403401 1321362 280133 269884
 $ OH           : int  539587 1422175 1031807 175806 155282
 $ OK           : int  106855 772405 521629 0 189654
 $ OR           : int  1066 297591 315400 0 237768
 $ PA           : int  435540 1936985 1108074 795783 182051
 $ RI           : int  0 93829 78260 0 13264
 $ SC           : int  150973 361249 495616 568055 145328
 $ SD           : int  24769 98288 118593 0 129978
 $ TN           : int  204725 440017 707095 372319 116472
 $ TX           : int  932569 5007366 6582173 434709 751680
 $ UT           : int  237870 287076 322387 0 37369
 $ VA           : int  67739 665869 798162 294606 185638
 $ VT           : int  0 14046 71534 0 22009
 $ WA           : int  42238 381886 725931 102929 418470
 $ WI           : int  232501 622144 535483 105285 150890
 $ WV           : int  536642 281657 201769 0 28391
 $ WY           : int  390303 172450 145723 0 42079
 $ US           : int  9885694 33361871 35330835 8061020 8107353
```

```
str(tot_use_23)
```

```
'data.frame':   5 obs. of  53 variables:
 $ Energy_Source: chr  "coal_usage" "NaturalGas" "petroleum (BTU)" "nuclear-
energy †" ...
 $ AK           : int  18414 448087 270391 0 10087
 $ AL           : int  224926 775747 565754 476392 222189
 $ AR           : int  180262 399566 327465 156492 87277
 $ AZ           : int  137885 537151 599712 329474 108445
 $ CA           : int  28746 2154533 2996168 185192 1065179
 $ CO           : int  204826 525446 514174 0 115061
 $ CT           : int  0 304924 292864 142873 48981
 $ DC           : int  0 26236 17292 0 2795
 $ DE           : int  338 84387 110721 0 8041
 $ FL           : int  129387 1673836 1835394 312935 286306
 $ GA           : int  177521 787361 980546 390663 291462
 $ HI           : int  0 152 251676 0 21046
 $ IA           : int  201276 446677 404172 0 414801
 $ ID           : int  1144 154150 189553 0 77128
 $ IL           : int  342683 1101064 1134461 1019691 245703
 $ IN           : int  613533 921814 695709 0 172891
 $ KS           : int  184614 309427 345807 107675 140268
 $ KY           : int  481815 369986 584722 0 72603
 $ LA           : int  58224 2055504 1620038 127634 138982
 $ MA           : int  0 386946 525647 0 81559
 $ MD           : int  30349 304669 429784 156610 53711
 $ ME           : int  1295 61045 177091 0 89444
```

```
 $ MI           : int  287490 1104234 810789 292615 198459
 $ MN           : int  148968 536789 567072 124626 223864
 $ MO           : int  442901 316512 612625 95947 90412
 $ MS           : int  49606 630107 378072 122807 67305
 $ MT           : int  130059 96777 173283 0 58470
 $ NC           : int  153784 662302 900241 442493 186804
 $ ND           : int  325716 220768 169307 0 92154
 $ NE           : int  195602 206276 233599 72391 164502
 $ NH           : int  1838 59589 147387 99658 38988
 $ NJ           : int  0 721282 787262 296162 74408
 $ NM           : int  75182 337083 251686 0 80278
 $ NV           : int  29284 301655 296155 0 74878
 $ NY           : int  4823 1346622 1341811 287690 272967
 $ OH           : int  413577 1448857 1009729 169392 153083
 $ OK           : int  63787 860217 515440 0 185378
 $ OR           : int  652 327164 313013 0 236062
 $ PA           : int  307604 1937041 1132958 787083 178035
 $ RI           : int  0 112499 76844 0 13579
 $ SC           : int  162323 346881 507146 581365 142486
 $ SD           : int  22246 99752 114623 0 126540
 $ TN           : int  202367 396870 702827 396522 115678
 $ TX           : int  805600 5284670 6752349 425186 791211
 $ UT           : int  174315 298976 324640 0 39674
 $ VA           : int  46785 655997 807547 310037 183979
 $ VT           : int  0 13001 70235 0 22209
 $ WA           : int  49523 403038 718277 88163 365956
 $ WI           : int  219995 565025 525386 101204 150965
 $ WV           : int  472309 309019 206969 0 28370
 $ WY           : int  366098 181395 143944 0 38474
 $ US           : int  8169673 33609104 35460356 8098974 8187317
```

```
head(tot_use_21)  # We will pivot longer these df
```

```
           Energy_Source     AK     AL     AR     AZ      CA     CO     CT
1                   Coal  18694 309791 216123 160299   28244 252442   2880
2            Natural Gas† 395590 739891 360545 484962 2172757 509970 305184
3         Petroleum (BTU) 261094 583042 328271 606862 2959389 497788 284788
4                nuclear      0 480115 141372 329868  171842      0 179551
5 total_renewable_energy   9597 239817  89714  99266  810020 103955  49306
     DC     DE      FL     GA     HI     IA     ID      IL     IN     KS
KY
1     0   4542  200193 203870  12566 264419   3051  522809 753557 219031
548443
2 28336  82708 1591864 773889    133 383424 135176 1088485 869328 291797
365875
3 18439 113641 1748346 922503 223014 408385 188263 1136797 712427 339006
```

```
584011
4     0      0  307811 354085      0      0      0 1011555      0  89426
0
5  2487   7150  297291 289113  20134 389787  74428  224106 157324 135551
71744
       LA     MA     MD     ME     MI     MN     MO     MS     MT     NC
ND
1   95856      0  69186   1588 436203 179055 616413  64446 122765 222501
361811
2 1862349 404301 299282  57233 950364 523812 293633 576903  87105 637553
191168
3 1840835 503312 433791 163991 814081 561731 607276 384328 176686 884299
168682
4  179886      0 156369      0 358114 147286  44766 122771      0 449675
0
5  135905  75370  52732  95141 194075 216113  88879  66134  56334 196973
92653
      NE     NH     NJ     NM     NV      NY      OH     OK     OR      PA
1 216298   3259  12586 133228  35910    5370  575920 131695   1303  485193
2 191008  60116 697019 285809 305212 1359437 1294814 745911 305665 1868137
3 237214 142030 749892 262885 286548 1237451 1028000 517408 317322 1047658
4  71758 102789 293494      0      0  325141  182330      0      0  791587
5 158275  38479  70039  62210  63647  263977  146858 177087 225544  179589
      RI     SC     SD     TN      TX     UT     VA    VT     WA     WI     WV
1      0 162628  21589 225784  968401 276159  68603     0  36943 286760 633582
2 105473 349990  96787 413554 4773076 274420 699927 13801 384769 561076 277002
3  76464 508147 119505 713210 6783182 304823 795296 72241 711662 533390 205005
4      0 560782      0 368461  419363      0 297972     0  88764 103979      0
5  11798 143796 127382 135841  654199  36050 174615 21430 394052 145936  26427
      WY       US
1 376971 10548957
2 161580 31688203
3 146274 35250685
4      0  8130913
5  37734  7646167
```

```
head(tot_use_22)  # We will pivot longer these df
```

```
     Energy_Source     AK     AL     AR     AZ      CA     CO     CT    DC
1 coal Consumption  18615 297654 211724 154007   30049 233256      0     0
2      Natural-Gas 437916 787300 398099 468038 2131372 524890 307212 30174
3  petroleum (btu) 263335 578431 327813 594859 3017944 538413 302881 18000
4  Nuclear Energy†      0 442093 149654 333738  183814      0 172018     0
5 total_renewables  10410 232035  90825 101215  880995 114917  49084  2622
      DE      FL     GA     HI     IA     ID      IL     IN     KS     KY
1   1846  171953 180888   7680 227866   1881  496983 719238 226712 523276
```

```
2  89674 1659544 809618    159 434374 141924 1134781 913401 318779 402534
3 112026 1815529 940579 241994 423592 190635 1138141 699235 346852 580349
4      0  321468 356001      0      0      0 1032989      0  93844      0
5   7402  304605 293237  20471 421784  78406  248541 170986 151788  77517
       LA     MA     MD     ME      MI     MN     MO     MS     MT     NC
1   96914      0  61932   1269  423504 184517 566940  66214 131345 163029
2 2087166 432442 310133  62559 1087716 535010 322547 617855  93971 747187
3 1663129 529154 411842 166724  820709 568916 606374 383366 177009 906477
4  168889      0 154742      0  271788 153546  92724  89856      0 445547
5  138209  80700  51255  93867  206811 229769  95312  66614  60644 198165
      ND     NE     NH     NJ     NM     NV      NY      OH     OK     OR
1 369340 223571   3864   6199 138077  35835    6143  539587 106855   1066
2 198986 199260  60176 755048 301279 302315 1403401 1422175 772405 297591
3 170390 237556 149025 769751 255571 303234 1321362 1031807 521629 315400
4      0  58702 114108 295875      0      0  280133  175806      0      0
5  96024 168382  39863  73187  77286  72734  269884  155282 189654 237768
       PA    RI     SC     SD     TN      TX     UT     VA    VT     WA     WI
1  435540     0 150973  24769 204725  932569 237870  67739     0  42238 232501
2 1936985 93829 361249  98288 440017 5007366 287076 665869 14046 381886 622144
3 1108074 78260 495616 118593 707095 6582173 322387 798162 71534 725931 535483
4  795783     0 568055      0 372319  434709      0 294606     0 102929 105285
5  182051 13264 145328 129978 116472  751680  37369 185638 22009 418470 150890
      WV     WY       US
1 536642 390303  9885694
2 281657 172450 33361871
3 201769 145723 35330835
4      0      0  8061020
5  28391  42079  8107353
```

```
head(tot_use_23)  # We will pivot longer these df
```

```
           Energy_Source     AK     AL     AR     AZ      CA     CO     CT
1             coal_usage  18414 224926 180262 137885   28746 204826      0
2             NaturalGas 448087 775747 399566 537151 2154533 525446 304924
3        petroleum (BTU) 270391 565754 327465 599712 2996168 514174 292864
4       nuclear-energy †      0 476392 156492 329474  185192      0 142873
5 total renewable-energy  10087 222189  87277 108445 1065179 115061  48981
     DC     DE      FL     GA     HI     IA     ID      IL     IN     KS
KY
1     0    338  129387 177521      0 201276   1144  342683 613533 184614
481815
2 26236  84387 1673836 787361    152 446677 154150 1101064 921814 309427
369986
3 17292 110721 1835394 980546 251676 404172 189553 1134461 695709 345807
584722
4     0      0  312935 390663      0      0      0 1019691      0 107675
```

```
0
5  2795   8041  286306 291462  21046 414801  77128  245703 172891 140268
72603
       LA     MA     MD     ME      MI     MN     MO     MS     MT     NC
1   58224      0  30349   1295  287490 148968 442901  49606 130059 153784
2 2055504 386946 304669  61045 1104234 536789 316512 630107  96777 662302
3 1620038 525647 429784 177091  810789 567072 612625 378072 173283 900241
4  127634      0 156610      0  292615 124626  95947 122807      0 442493
5  138982  81559  53711  89444  198459 223864  90412  67305  58470 186804
      ND     NE     NH     NJ     NM     NV      NY      OH     OK     OR
1 325716 195602   1838      0  75182  29284    4823  413577  63787    652
2 220768 206276  59589 721282 337083 301655 1346622 1448857 860217 327164
3 169307 233599 147387 787262 251686 296155 1341811 1009729 515440 313013
4      0  72391  99658 296162      0      0  287690  169392      0      0
5  92154 164502  38988  74408  80278  74878  272967  153083 185378 236062
       PA     RI     SC     SD     TN      TX     UT     VA    VT     WA
WI
1  307604      0 162323  22246 202367  805600 174315  46785     0  49523
219995
2 1937041 112499 346881  99752 396870 5284670 298976 655997 13001 403038
565025
3 1132958  76844 507146 114623 702827 6752349 324640 807547 70235 718277
525386
4  787083      0 581365      0 396522  425186      0 310037     0  88163
101204
5  178035  13579 142486 126540 115678  791211  39674 183979 22209 365956
150965
      WV     WY       US
1 472309 366098  8169673
2 309019 181395 33609104
3 206969 143944 35460356
4      0      0  8098974
5  28370  38474  8187317
```

```
str(ev_regis)
```

```
'data.frame':   54 obs. of  2 variables:
 $ electric.vehicle.registrations_by_state..2023.: chr  "" "STATE" "Alabama"
"Alaska" ...
 $ X                                             : chr  "" "Count-EVs "
"#13047" "~2697" ...
```

```
head(ev_regis)
```

```
  electric.vehicle.registrations_by_state..2023.          X
1
2                                            STATE Count-EVs
3                                          Alabama     #13047
4                                           Alaska      ~2697
5                                          Arizona      89798
6                                         Arkansas   7108 EVs
```

```
str(av_en_price_21_23)
```

```
'data.frame':   55 obs. of  1 variable:
 $ V1: chr  "Total energy average price, dollars per million Btu,,," ",,,"
"State,2021,2022,2023" "AK,$20.03 per MMBtu,$27.33,$23.84 est." ...
```

```
head(av_en_price_21_23)
```

```
                                                      V1
1 Total energy average price, dollars per million Btu,,,
2                                                   ,,,
3                                  State,2021,2022,2023
4                 AK,$20.03 per MMBtu,$27.33,$23.84 est.
5                    AL,about 17.85 USD,23.37 USD,≈21.11
6                      AR,$18.42,$23.84 per MMBtu,$21.76
```

```
## Starting to clean each dataset

## Renewable energy usage datasets
# We have to standardize the values in the column "Renewable_Use_2021"
renew_use_21u <- renew_use_21 %>% mutate(renewable_use_21 =
str_extract(Renewable_Use_2021, "[0-9]+"))

# repeating the same process for 2022 and 2023
renew_use_22u <- renew_use_22 %>% mutate(renewable_use_22 =
str_extract(Renewable_Use_2022, "[0-9]+"))

renew_use_23u <- renew_use_23 %>% mutate(renewable_use_23 =
str_extract(Renewable_Use_2023, "[0-9]+"), State = str_to_upper(State))

# Cleaning EV registrations by state dataframe

# Renaming columns and deleting first 2 rows
colnames(ev_regis)
```

```
[1] "electric.vehicle.registrations_by_state..2023."
[2] "X"
```

```
head(ev_regis)
```

```
  electric.vehicle.registrations_by_state..2023.          X
1                                                            
2                                           STATE Count-EVs
3                                         Alabama     #13047
4                                          Alaska      ~2697
5                                         Arizona      89798
6                                        Arkansas   7108 EVs
```

```
ev_regis_u <- ev_regis %>% rename(ev_registration_state =
electric.vehicle.registrations_by_state..2023., ev_count = X) %>% slice(-c(1,
2))

# Only extracting ev registrations values
ev_regis_by_state <- ev_regis_u %>% mutate(ev_count_u = str_extract(ev_count,
"[0-9]+"))


## Cleaning average energy price dataframe
colnames(av_en_price_21_23)
```

```
[1] "V1"
```

```
# removing first two rows
av_price <- av_en_price_21_23 %>% slice(-c(1:2))

# separating each column
colnames(av_price)
```

```
[1] "V1"
```

```
av_price_u <- av_price %>% separate(V1, into = c("State", "2021", "2022",
"2023"), sep = ",") %>% slice(-c(1))

# Only extracting relevant variables for the column data
av_priceu <- av_price_u %>% mutate(two_one = str_extract(`2021`, "[0-9]+\\.
[0-9]+"), two_two = str_extract(`2022`, "[0-9]+\\.[0-9]+"), two_three =
str_extract(`2023`, "[0-9]+\\.[0-9]+")) %>% select(-c(2, 3, 4))
```

Part 3: Joining / Pivoting Datasets for Analysis

```
## Joining the renewable energy usage datasets
head(renew_use_21u)
```

```
  State Energy_Source Renewable_Use_2021 renewable_use_21
1    AK       Biomass              ≈3153             3153
2    AK    Geothermal          186 MMBtu              186
3    AK    Hydropower         5763 about             5763
4    AK  Solar Energy                ~45               45
5    AK   Wind Energy            451 USD              451
6    AL       Biomass        198543 est.           198543
```

```
# using a left join to join 2021 and 2022 data
df_21_22 <- left_join(x=renew_use_21u, y=renew_use_22u, by = join_by(State ==
State, Energy_Source == Energy_Source)) %>% select(-c("Renewable_Use_2021",
"Renewable_Use_2022"))

# using another left join to join df_21_22 and 2023 data
df_21_23 <- left_join(x=df_21_22, y=renew_use_23u, by = join_by(State ==
State, Energy_Source == Energy_Source)) %>% select(-c("Renewable_Use_2023"))

## Calculating the total number of renewable energy used by state for each
respective year
tot_ren_state <- df_21_23 %>%
  mutate(across(starts_with("renewable_use"), as.numeric)) %>%
  group_by(State) %>%
  summarize(
    total_21 = sum(renewable_use_21, na.rm = TRUE),
    total_22 = sum(renewable_use_22, na.rm = TRUE),
    total_23 = sum(renewable_use_23, na.rm = TRUE)
  )

## Pivoting the total energy usage dataframes
tot_use_21l <- tot_use_21 %>% pivot_longer(cols = -Energy_Source, names_to =
"State", values_to = "Tot_usage_21") %>% group_by(State) %>% arrange(State)

## Repeating the same process for 2022 and 2023
tot_use_22l <- tot_use_22 %>% pivot_longer(cols = -Energy_Source, names_to =
"State", values_to = "Tot_usage_22") %>% group_by(State) %>% arrange(State)

tot_use_23l <- tot_use_23 %>% pivot_longer(cols = -Energy_Source, names_to =
"State", values_to = "Tot_usage_23") %>% group_by(State) %>% arrange(State)
```

Part 4: Mapping Visualization

```
# loading packages
library(maps)
```

```
Warning: package 'maps' was built under R version 4.3.3
```

```

Attaching package: 'maps'
```

```
The following object is masked from 'package:purrr':

    map
```

```
# creating maps
us_map <- map_data("state")

ggplot(us_map, aes(long, lat)) + geom_polygon(color = "white") +
coord_fixed(1.3) + theme_minimal() + labs(title = "Average Income by State")
```