Renewable Energy Share and Electrix Vehicle Adoption Across US States (2021-2023)

Example Solution 1

Part 0: libraries

```
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0
──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.1     ✔ stringr   1.5.2
✔ ggplot2   4.0.0     ✔ tibble    3.3.0
✔ lubridate 1.9.4     ✔ tidyr     1.3.1
✔ purrr     1.1.0
── 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(ggplot2)
library(dplyr)
library(RColorBrewer)
```

Part 1: Defining Research Question

This analysis explores the relationship between state-level renewable energy adoption and electric vehicle (EV) registrations in the United States from 2021 to 2023. The primary research question is:

How does the share of renewable energy relate to EV adoption and energy prices across U.S. states?

Chosen sub-questions: 1. How has the share of renewable energy changed from 2021 to 2023 across US states? 2. Do states with higher renewable energy shares also have more EVs registered in 2023? 3. Do states with higher renewable energy use have lower or higher average energy prices?

Part 2: Data Preparation and Cleaning

```r
renew_2021 <- read.csv("data/renew-use-2021.csv")
renew_2022 <- read.csv("data/renew-use-2022.csv")
renew_2023 <- read.csv("data/renew-use-2023.csv")
total_use_2021 <- read.csv("data/total-use-2021.csv" )
total_use_2022 <- read.csv("data/total-use-2022.csv" )
total_use_2023 <- read.csv("data/total-use-2023.csv" )
ev_regs <- read.csv("data/ev-registrations-by-state-2023.csv", skip=2)
avg_price <- read.csv("data/av-energy-price-2021-2023.csv", sep=",")

clean_colnames <- function(df){
    colnames(df) <- tolower(colnames(df))
    colnames(df)<- gsub(" ", "_", colnames(df))
    return(df)
}
renew_2021 <- clean_colnames(renew_2021)
renew_2022 <- clean_colnames(renew_2022)
renew_2023 <- clean_colnames(renew_2023)
avg_price <- clean_colnames(avg_price)
total_use_2021 <- clean_colnames(total_use_2021)
total_use_2022 <- clean_colnames(total_use_2022)
total_use_2023 <- clean_colnames(total_use_2023)
ev_regs <- clean_colnames(ev_regs)

state_map <-data.frame(
  state_name = c(state.name, "District of Columbia", "US"),
  state_abbr = c(state.abb, "DC", "US"))

renew_2021$state <- trimws(renew_2021$state)
renew_2022$state <- trimws(renew_2022$state)
renew_2023$state <- trimws(renew_2023$state)
renew_2021$state <- toupper(renew_2021$state)
renew_2022$state <- toupper(renew_2022$state)
renew_2023$state <- toupper(renew_2023$state)

renew_2021$state <- state_map$state_name[match(renew_2021$state,
state_map$state_abbr)]
renew_2022$state <- state_map$state_name[match(renew_2022$state,
state_map$state_abbr)]
renew_2023$state <- state_map$state_name[match(renew_2023$state,
state_map$state_abbr)]

clean_numeric <- function(x) {
  x <- gsub("[^0-9.]", "", x)
  as.numeric(x)
}
renew_2021$renewable_use_2021 <- clean_numeric(renew_2021$renewable_use_2021)
renew_2022$renewable_use_2022 <- clean_numeric(renew_2022$renewable_use_2022)
renew_2023$renewable_use_2023 <- clean_numeric(renew_2023$renewable_use_2023)
```

```
ev_regs$count.evs<- clean_numeric(ev_regs$count.evs)
avg_price$x2021 <- clean_numeric(avg_price$x2021)
avg_price$x2022 <- clean_numeric(avg_price$x2022)
avg_price$x2023 <- clean_numeric(avg_price$x2023)
```

```
Warning in clean_numeric(avg_price$x2023): NAs introduced by coercion
```

```
state_abbr <-
c("AK","AL","AR","AZ","CA","CO","CT","DC","DE","FL","GA","HI","IA","ID",

"IL","IN","KS","KY","LA","MA","MD","ME","MI","MN","MO","MS","MT","NC",

"ND","NE","NH","NJ","NM","NV","NY","OH","OK","OR","PA","RI","SC","SD",
                "TN","TX","UT","VA","VT","WA","WI","WV","WY","US")

state_full <-
c("Alaska","Alabama","Arkansas","Arizona","California","Colorado","Connecticut",
                "District of
Columbia","Delaware","Florida","Georgia","Hawaii","Iowa","Idaho",

"Illinois","Indiana","Kansas","Kentucky","Louisiana","Massachusetts","Maryland",

"Maine","Michigan","Minnesota","Missouri","Mississippi","Montana","North
Carolina",
                "North Dakota","Nebraska","New Hampshire","New Jersey","New
Mexico","Nevada",
                "New York","Ohio","Oklahoma","Oregon","Pennsylvania","Rhode
Island","South Carolina",
                "South
Dakota","Tennessee","Texas","Utah","Virginia","Vermont","Washington",
                "Wisconsin","West Virginia","Wyoming","US")


colnames(total_use_2021)[-1] <- state_full
colnames(total_use_2022)[-1] <- state_full
colnames(total_use_2023)[-1] <- state_full
avg_price$state <- state_full[match(avg_price$state, state_abbr)]
colnames(ev_regs)[1] <- "State"
colnames(avg_price)[1] <- "State"
head(avg_price)
```

```
       State x2021 x2022 x2023
1     Alaska 20.03 27.33    NA
2    Alabama 17.85 23.37 21.11
```

```
3   Arkansas 18.42 23.84 21.76
4    Arizona 25.07 31.72 30.28
5 California 28.44 37.35 35.72
6   Colorado 20.64 25.85 23.85
```

```
head(ev_regs)
```

```
       State count.evs
1    Alabama     13047
2     Alaska      2697
3    Arizona     89798
4   Arkansas      7108
5 California   1256646
6   Colorado     90083
```

```
head(renew_2021)
```

```
    state energy_source renewable_use_2021
1  Alaska       Biomass               3153
2  Alaska    Geothermal                186
3  Alaska    Hydropower               5763
4  Alaska  Solar Energy                 45
5  Alaska   Wind Energy                451
6 Alabama       Biomass             198543
```

```
head(total_use_2021)
```

```
           energy_source Alaska Alabama Arkansas Arizona California Colorado
1                   Coal  18694  309791   216123  160299      28244   252442
2           Natural Gas† 395590  739891   360545  484962    2172757   509970
3        Petroleum (BTU) 261094  583042   328271  606862    2959389   497788
4                nuclear      0  480115   141372  329868     171842        0
5 total_renewable_energy   9597  239817    89714   99266     810020   103955
  Connecticut District of Columbia Delaware Florida Georgia Hawaii   Iowa
1        2880                    0     4542  200193  203870  12566 264419
2      305184                28336    82708 1591864  773889    133 383424
3      284788                18439   113641 1748346  922503 223014 408385
4      179551                    0        0  307811  354085      0      0
5       49306                 2487     7150  297291  289113  20134 389787
   Idaho Illinois Indiana Kansas Kentucky Louisiana Massachusetts Maryland
1   3051   522809  753557 219031   548443     95856             0    69186
2 135176  1088485  869328 291797   365875   1862349        404301   299282
3 188263  1136797  712427 339006   584011   1840835        503312   433791
```

```
4      0  1011555       0  89426        0    179886             0   156369
5  74428   224106  157324 135551    71744    135905         75370    52732
   Maine Michigan Minnesota Missouri Mississippi Montana North Carolina
1   1588   436203    179055   616413       64446  122765         222501
2  57233   950364    523812   293633      576903   87105         637553
3 163991   814081    561731   607276      384328  176686         884299
4      0   358114    147286    44766      122771       0         449675
5  95141   194075    216113    88879       66134   56334         196973
  North Dakota Nebraska New Hampshire New Jersey New Mexico Nevada New York
1       361811   216298          3259      12586     133228  35910     5370
2       191168   191008         60116     697019     285809 305212  1359437
3       168682   237214        142030     749892     262885 286548  1237451
4            0    71758        102789     293494          0      0   325141
5        92653   158275         38479      70039      62210  63647   263977
     Ohio Oklahoma Oregon Pennsylvania Rhode Island South Carolina South
Dakota
1  575920   131695   1303       485193            0         162628
21589
2 1294814   745911 305665      1868137       105473         349990
96787
3 1028000   517408 317322      1047658        76464         508147
119505
4  182330        0      0       791587            0         560782
0
5  146858   177087 225544       179589        11798         143796
127382
  Tennessee   Texas   Utah Virginia Vermont Washington Wisconsin West Virginia
1    225784  968401 276159    68603       0      36943    286760        633582
2    413554 4773076 274420   699927   13801     384769    561076        277002
3    713210 6783182 304823   795296   72241     711662    533390        205005
4    368461  419363      0   297972       0      88764    103979             0
5    135841  654199  36050   174615   21430     394052    145936         26427
  Wyoming       US
1  376971 10548957
2  161580 31688203
3  146274 35250685
4       0  8130913
5   37734  7646167
```

Part 3: Joining / Pivoting Datasets for Analysis

```r
renew_2021_sum <- aggregate(as.numeric(renew_2021$renewable_use_2021),
                            by = list(State = renew_2021$state), sum, na.rm =
TRUE)
colnames(renew_2021_sum)[2] <- "Renewable_2021"

renew_2022_sum <- aggregate(as.numeric(renew_2022$renewable_use_2022),
                            by = list(State = renew_2022$state), sum, na.rm =
```

```r
TRUE)
colnames(renew_2022_sum)[2] <- "Renewable_2022"

renew_2023_sum <- aggregate(as.numeric(renew_2023$renewable_use_2023),
                            by = list(State = renew_2023$state), sum, na.rm =
TRUE)
colnames(renew_2023_sum)[2] <- "Renewable_2023"

total_use_2022[1] <- total_use_2021[1]
total_use_2023[1] <- total_use_2021[1]
numeric_2021 <- apply(total_use_2021[, -1], 2, as.numeric)
numeric_2022 <- apply(total_use_2022[, -1], 2, as.numeric)
numeric_2023 <- apply(total_use_2023[, -1], 2, as.numeric)
total_energy_2021 <- colSums(numeric_2021, na.rm = TRUE)
total_energy_2022 <- colSums(numeric_2022, na.rm = TRUE)
total_energy_2023 <- colSums(numeric_2023, na.rm = TRUE)

total_2021_long <- data.frame(State = names(total_use_2021)[-1],
                              Total_2021 = total_energy_2021)
total_2022_long <- data.frame(State = names(total_use_2022)[-1],
                              Total_2022 = total_energy_2022)
total_2023_long <- data.frame(State = names(total_use_2023)[-1],
                              Total_2023 = total_energy_2023)
head(total_2021_long)
```

```
                State Total_2021
Alaska         Alaska     684975
Alabama       Alabama    2352656
Arkansas     Arkansas    1136025
Arizona       Arizona    1681257
California California    6142252
Colorado     Colorado    1364155
```

```r
energy_2021 <- merge(renew_2021_sum, total_2021_long, by = "State", all.x =
TRUE)
energy_2021$Pct_Renew_2021 <- energy_2021$Renewable_2021 /
energy_2021$Total_2021

energy_2022 <- merge(renew_2022_sum, total_2022_long, by = "State", all.x =
TRUE)
energy_2022$Pct_Renew_2022 <- energy_2022$Renewable_2022 /
energy_2022$Total_2022

energy_2023 <- merge(renew_2023_sum, total_2023_long, by = "State", all.x =
TRUE)
energy_2023$Pct_Renew_2023 <- energy_2023$Renewable_2023 /
```

```r
energy_2023$Total_2023
head(energy_2021)
```

```
       State Renewable_2021 Total_2021 Pct_Renew_2021
1    Alabama         239816    2352656     0.10193415
2     Alaska           9598     684975     0.01401219
3    Arizona          99266    1681257     0.05904273
4   Arkansas          89714    1136025     0.07897185
5 California         810020    6142252     0.13187671
6   Colorado         103956    1364155     0.07620542
```

```r
energy_all <- Reduce(function(x, y) merge(x, y, by = "State", all = TRUE),
                     list(energy_2021, energy_2022, energy_2023))

energy_all <- merge(energy_all, ev_regs[, c("State", "count.evs")], by =
"State", all.x = TRUE)
energy_all <- merge(energy_all, avg_price[, c("State", "x2021", "x2022",
"x2023")],
                    by = "State", all.x = TRUE)

energy_all$Renew_Growth_21_23 <- (energy_all$Pct_Renew_2023 -
energy_all$Pct_Renew_2021) * 100
energy_all$EVs_per_BTU <- energy_all$count.evs / energy_all$Total_2023

head(energy_all)
```

```
       State Renewable_2021 Total_2021 Pct_Renew_2021 Renewable_2022
Total_2022
1    Alabama         239816    2352656     0.10193415         232035
2337513
2     Alaska           9598     684975     0.01401219          10410
730276
3    Arizona          99266    1681257     0.05904273         101214
1651857
4   Arkansas          89714    1136025     0.07897185          90824
1178115
5 California         810020    6142252     0.13187671         880995
6244174
6   Colorado         103956    1364155     0.07620542         114918
1411476
  Pct_Renew_2022 Renewable_2023 Total_2023 Pct_Renew_2023 count.evs x2021
x2022
1     0.09926576         222189    2265008     0.09809634     13047 17.85
23.37
2     0.01425488          10088     746979     0.01350507      2697 20.03
```

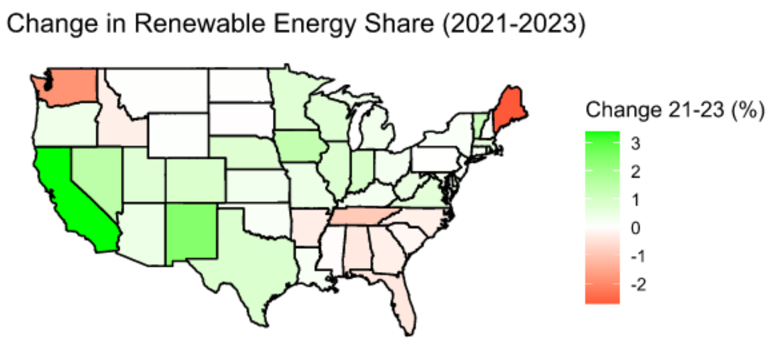
```
27.33
3     0.06127286         108445    1712667     0.06331937     89798 25.07
31.72
4     0.07709264          87277    1151062     0.07582302      7108 18.42
23.84
5     0.14109072        1065179    6429818     0.16566239   1256646 28.44
37.35
6     0.08141690         115062    1359507     0.08463509     90083 20.64
25.85
  x2023 Renew_Growth_21_23 EVs_per_BTU
1 21.11        -0.38378122 0.005760245
2    NA        -0.05071252 0.003610543
3 30.28         0.42766445 0.052431675
4 21.76        -0.31488310 0.006175167
5 35.72         3.37856750 0.195440369
6 23.85         0.84296754 0.066261520
```

Part 4: Mapping Visualization

```
us_map <- map_data("state")
energy_all$State_lower <- tolower(energy_all$State)

map_renew_growth <- us_map %>%
  left_join(energy_all, by = c("region" = "State_lower"))

ggplot(map_renew_growth, aes(x = long, y = lat, group = group, fill =
Renew_Growth_21_23)) +
  geom_polygon(color = "black") +
  coord_fixed(1.3) +
  scale_fill_gradient2(low = "red", mid = "white", high = "green",
                       midpoint = 0, na.value = "grey90",
                       name = "Change 21-23 (%)") +
  labs(title = "Change in Renewable Energy Share (2021-2023)") +
  theme_minimal() +
  theme(axis.text = element_blank(),
        axis.title = element_blank(),
        panel.grid = element_blank())
```



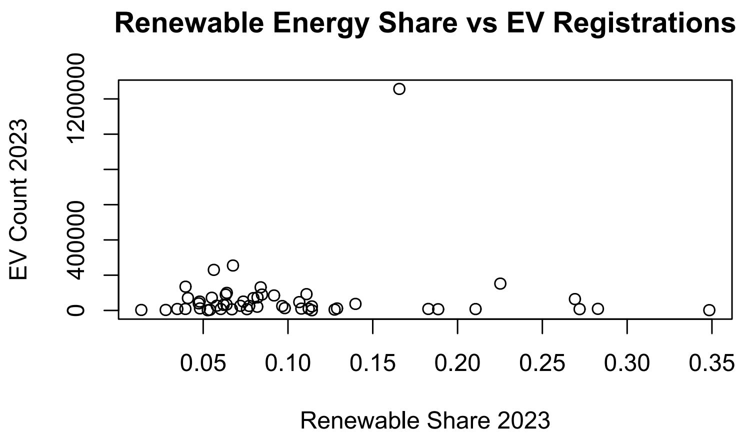

```
#Q1 Renewable Share Change 2021-2023

energy_all[, c("State", "Pct_Renew_2021", "Pct_Renew_2023",
"Renew_Growth_21_23")]
```

```
                  State Pct_Renew_2021 Pct_Renew_2023 Renew_Growth_21_23
1               Alabama     0.10193415     0.09809634       -0.383781215
2                Alaska     0.01401219     0.01350507       -0.050712516
3               Arizona     0.05904273     0.06331937        0.427664450
4              Arkansas     0.07897185     0.07582302       -0.314883105
5            California     0.13187671     0.16566239        3.378567501
6              Colorado     0.07620542     0.08463509        0.842967545
7           Connecticut     0.06000421     0.06203191        0.202769739
8              Delaware     0.03437303     0.03951112        0.513809132
9  District of Columbia     0.05048516     0.06035879        0.987362408
10              Florida     0.07171382     0.06755937       -0.415445407
11              Georgia     0.11366918     0.11092526       -0.274391494
12               Hawaii     0.07869547     0.07712717       -0.156829809
13                Idaho     0.18564395     0.18277623       -0.286771632
14             Illinois     0.05625526     0.06392519        0.766993423
15              Indiana     0.06311591     0.07191964        0.880372459
16                 Iowa     0.26955875     0.28276886        1.321011211
17               Kansas     0.12611520     0.12894756        0.283236831
18             Kentucky     0.04569405     0.04810930        0.241524928
19            Louisiana     0.03302809     0.03474218        0.171409421
20                Maine     0.29923291     0.27196959       -2.726331218
21             Maryland     0.05214068     0.05508126        0.294057456
22        Massachusetts     0.07667579     0.08203977        0.536397661
23             Michigan     0.07050000     0.07367796        0.317796485
24            Minnesota     0.13274779     0.13979975        0.705196496
25          Mississippi     0.05444918     0.05393394       -0.051524653
```

```
26          Missouri     0.05383390     0.05801603       0.418212083
27           Montana     0.12719863     0.12749761       0.029897959
28          Nebraska     0.18097817     0.18857022       0.759204849
29            Nevada     0.09206630     0.10666950       1.460319435
30     New Hampshire     0.11099220     0.11221148       0.121927455
31        New Jersey     0.03841901     0.03959792       0.117891117
32        New Mexico     0.08359942     0.10786734       2.426792137
33          New York     0.08271604     0.08388915       0.117310657
34    North Carolina     0.08238014     0.07963936      -0.274077871
35      North Dakota     0.11378166     0.11405974       0.027808255
36              Ohio     0.04549614     0.04791904       0.242290163
37          Oklahoma     0.11264289     0.11409188       0.144899207
38            Oregon     0.26539654     0.26920450       0.380796008
39      Pennsylvania     0.04107531     0.04099619      -0.007912541
40      Rhode Island     0.06089762     0.06691734       0.601971982
41    South Carolina     0.08334285     0.08187905      -0.146380104
42      South Dakota     0.34874050     0.34844050      -0.029999773
43         Tennessee     0.07315669     0.06376084      -0.939584823
44             Texas     0.04810916     0.05627777       0.816860435
45                US     0.08198330     0.08754108       0.555778751
46              Utah     0.04043964     0.04736720       0.692755346
47           Vermont     0.19941008     0.21062165       1.121157219
48          Virginia     0.08574636     0.09178959       0.604322821
49        Washington     0.24381539     0.22520904      -1.860635669
50     West Virginia     0.02314153     0.02790393       0.476239311
51         Wisconsin     0.08946805     0.09661296       0.714491536
52           Wyoming     0.05222273     0.05271054       0.048781088
```

```r
#Q2 Renewable share vs EVs
plot(energy_all$Pct_Renew_2023, energy_all$count.evs,
     xlab = "Renewable Share 2023", ylab = "EV Count 2023",
     main = "Renewable Energy Share vs EV Registrations")
```




```r
cor(energy_all$Pct_Renew_2023, energy_all$count.evs, use = "complete.obs")
```

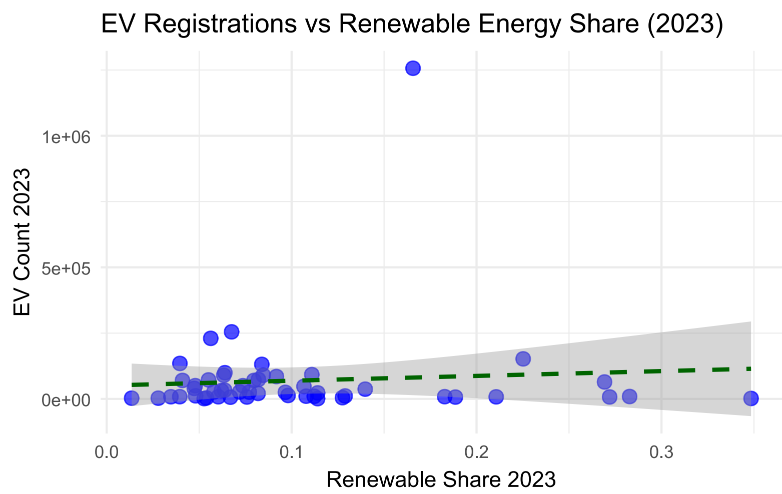
```
[1] 0.07415626
```

```r
ggplot(energy_all, aes(x = Pct_Renew_2023, y = count.evs)) +
  geom_point(color = "blue", size = 3, alpha = 0.7) +
  geom_smooth(method = "lm", color = "darkgreen", linetype = "dashed") +
  labs(title = "EV Registrations vs Renewable Energy Share (2023)",
       x = "Renewable Share 2023", y = "EV Count 2023") +
  theme_minimal()
```

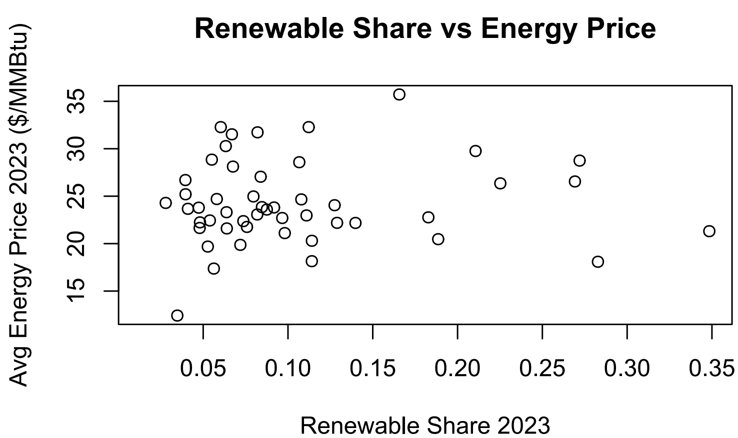
```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 1 row containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 1 row containing missing values or values outside the scale
range
(`geom_point()`).
```




```r
#Q3 Renewable share vs average price
plot(energy_all$Pct_Renew_2023, energy_all$x2023,
     xlab = "Renewable Share 2023", ylab = "Avg Energy Price 2023 ($/MMBtu)",
     main = "Renewable Share vs Energy Price")
```




```r
cor(energy_all$Pct_Renew_2023, energy_all$x2023, use = "complete.obs")
```

```
[1] 0.05423865
```

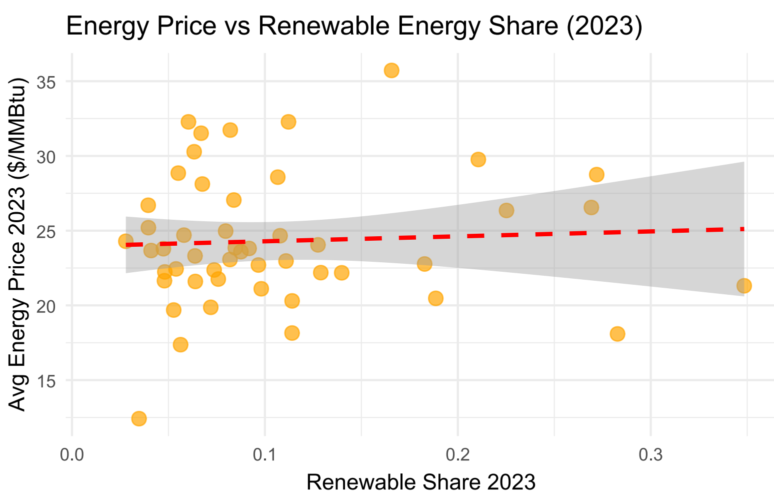
```r
ggplot(energy_all, aes(x = Pct_Renew_2023, y = x2023)) +
  geom_point(color = "orange", size = 3, alpha = 0.7) +
  geom_smooth(method = "lm", color = "red", linetype = "dashed") +
  labs(title = "Energy Price vs Renewable Energy Share (2023)",
```

```
    x = "Renewable Share 2023", y = "Avg Energy Price 2023 ($/MMBtu)") +
  theme_minimal()
```

```
`geom_smooth()` using formula = 'y ~ x'
```

```
Warning: Removed 3 rows containing non-finite outside the scale range
(`stat_smooth()`).
```

```
Warning: Removed 3 rows containing missing values or values outside the scale
range
(`geom_point()`).
```




```
cor(energy_all$Pct_Renew_2023, energy_all$x2023, use = "complete.obs")
```

```
[1] 0.05423865
```

```
energy_long <- energy_all %>%
  select(State_lower, Pct_Renew_2021, Pct_Renew_2022, Pct_Renew_2023) %>%
  tidyr::pivot_longer(cols = starts_with("Pct_Renew"),
                      names_to = "Year",
                      values_to = "RenewableShare")
```

Part 5: Observed Pattern

Renewable Energy Growth (2021–2023): California, New Mexico, and Vermont showed the largest increases in renewable share. Some states, like Alabama and Tennessee, saw small declines.

Renewable Share vs EVs: Weak positive correlation ($r \approx 0.074$), suggesting states with higher renewable share do not necessarily have substantially higher EV adoption. Visual inspection shows California as an outlier with high EV count and moderate renewable share.

Renewable Share vs Energy Price: Negligible correlation ($r \approx 0.054$), indicating renewable share does not strongly influence average energy price across states.

The map reinforces that states with high renewable energy share are geographically clustered. States with both high renewable share and significant EV adoption are limited (e.g., California, Vermont), highlighting regional differences in energy and transportation transitions.

In conclusion, while renewable energy adoption has grown modestly in several states, there is not a strong nationwide link between renewable share and EV registrations or energy prices. Geographic visualization helps identify states leading in renewable energy and supports more targeted policy interventions.