Agricultural Dataset of Bangladesh

Jawwad Ahmed Ornob

2024-09-10

```r
# Loadiing Necessary Librariies

library(readr)
library(dplyr)
```

```
##
## 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
```

```r
library(ggplot2)
library(ggcorrplot)
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```r
library(shiny)


# Import Dataset

Agri_Data <- read.csv("C:/Users/HP/R Programming/Agricultural Dataset.csv")


head(Agri_Data)
```

```
##      District year area avg_rainfall max_temperature min_temperature  aus
## 1       dhaka 2008   NA         2385            34.2            12.5  804
## 2     gazipur 2008   NA         2197            30.2            19.5 4328
## 3   narsingdi 2008   NA         2197            34.2            12.5  484
## 4 narayangonj 2008   NA         2197            34.2            12.2 1617
```

```
## 5     tangail 2008   NA         1856           34.1            10.2   447
## 6  mymenshing 2008   NA         2239           29.6            21.6 88198
##     aman   boro wheat potato   jute humidity storm   urea   tsp    mp   DAP
## 1   9691 233939  1129  33679  29825       71   yes  25967  8262  4808  1573
## 2  93956 208434   479   4300  15072       71   yes  29264  9225  9057  5721
## 3  93958 207669  2073  14323  42440       71    no  30273  5068  5113  3031
## 4  14485 129390  4585  73525   3620       71   yes  16992  7954  7026   443
## 5 170918 672851 17477  56693 125166       78   yes  91562 19022 20143  8809
## 6 491413 812189  6111  43755  50856       79    no 120279 34007 47121 28723
##   inundationland_Highland inundationland_mediumhighland inundationland_lowland
## 1                   30118                         32245                 20539
## 2                   88346                         30449                 15220
## 3                   23779                         27893                 22923
## 4                    9076                         13243                 14834
## 5                   91012                        133622                 66018
## 6                  163675                        148743                 68689
##   inundationland_mediumlowland inundationland_verylowland Miscellaneous.Land
## 1                        41335                          0              30137
## 2                        20654                          0              16542
## 3                        20368                          0              20574
## 4                        22629                          0              15871
## 5                        15887                          0              31462
## 6                         9442                          0              44852
##   Calcareous.Alluvium Noncalcareous.Alluvium Acid.Basin.Clay
## 1                4869                   6782            6253
## 2                   0                    648           28245
## 3                   0                   8247            4942
## 4                   0                   2396            3263
## 5                   0                  28897            7786
## 6                   0                  11988           32531
##   Calcareous.Brown.Floodplain.Soil Calcareous.Grey.Floodplain.Soil
## 1                             7691                           26289
## 2                                0                               0
## 3                                0                               0
## 4                                0                               0
## 5                                0                               0
## 6                                0                               0
##   Calcareous.Dark.Grey.Floodplain.Soil Noncalcareous.Grey.Floodplain.Soil
## 1                                28109                              87891
## 2                                    0                              14655
## 3                                    0                              41976
## 4                                    0                              18095
## 5                                    0                             128008
## 6                                    0                              88491
##   Noncalcareous.Dark.Grey.Floodplain.Soil  Peat Made.Land
## 1                                    1876 13900      7048
## 2                                    4899     0         0
## 3                                   27689     0         0
## 4                                   26925     0       492
## 5                                   47714     0         0
## 6                                  165795     0         0
##   Noncalcareous.Brown.Floodplain.Soil Shallow.Red.Brown.Terrace.Soil
## 1                                   0                              0
## 2                                  13                          29623
```

```
## 3                                   0                            881
## 4                                   0                           2800
## 5                                 566                          14764
## 6                               14725                          12158
##   Deep.Red.Brown.Terrace.Soil Brown.Mottled.Terrace.Soil
## 1                           0                          0
## 2                       46500                       2414
## 3                        9800                        601
## 4                        4840                         30
## 5                       47378                       2423
## 6                       28174                       2689
##   Shallow.Grey.Terrace.Soil Deep.Grey.Terrace.Soil Grey.Valley.Soil
## 1                         0                      0                0
## 2                      8375                   4660            14637
## 3                         0                    140              687
## 4                       213                      0              728
## 5                      1971                   6390            20642
## 6                      1846                  10150             7915
##   Brown.Hill.Soil Grey.Piedmont.Soil soil.moisture
## 1               0                  0        124237
## 2               0                  0        154669
## 3               0                  0         94963
## 4               0                  0         59782
## 5               0                  0        306539
## 6             442              13645        390549
```

```r
ncol(Agri_Data)
```

```
## [1] 44
```

```r
colnames(Agri_Data)
```

```
##  [1] "District"
##  [2] "year"
##  [3] "area"
##  [4] "avg_rainfall"
##  [5] "max_temperature"
##  [6] "min_temperature"
##  [7] "aus"
##  [8] "aman"
##  [9] "boro"
## [10] "wheat"
## [11] "potato"
## [12] "jute"
## [13] "humidity"
## [14] "storm"
## [15] "urea"
## [16] "tsp"
## [17] "mp"
## [18] "DAP"
## [19] "inundationland_Highland"
## [20] "inundationland_mediumhighland"
## [21] "inundationland_lowland"
```

```
## [22] "inundationland_mediumlowland"
## [23] "inundationland_verylowland"
## [24] "Miscellaneous.Land"
## [25] "Calcareous.Alluvium"
## [26] "Noncalcareous.Alluvium"
## [27] "Acid.Basin.Clay"
## [28] "Calcareous.Brown.Floodplain.Soil"
## [29] "Calcareous.Grey.Floodplain.Soil"
## [30] "Calcareous.Dark.Grey.Floodplain.Soil"
## [31] "Noncalcareous.Grey.Floodplain.Soil"
## [32] "Noncalcareous.Dark.Grey.Floodplain.Soil"
## [33] "Peat"
## [34] "Made.Land"
## [35] "Noncalcareous.Brown.Floodplain.Soil"
## [36] "Shallow.Red.Brown.Terrace.Soil"
## [37] "Deep.Red.Brown.Terrace.Soil"
## [38] "Brown.Mottled.Terrace.Soil"
## [39] "Shallow.Grey.Terrace.Soil"
## [40] "Deep.Grey.Terrace.Soil"
## [41] "Grey.Valley.Soil"
## [42] "Brown.Hill.Soil"
## [43] "Grey.Piedmont.Soil"
## [44] "soil.moisture"
```

```r
summary(Agri_Data)
```

```
##    District              year           area          avg_rainfall
##  Length:70          Min.   :2008   Min.   :   21.0   Min.   :1181
##  Class :character   1st Qu.:2010   1st Qu.:  290.5   1st Qu.:1635
##  Mode  :character   Median :2012   Median : 1085.0   Median :1848
##                     Mean   :2012   Mean   : 7475.2   Mean   :1838
##                     3rd Qu.:2015   3rd Qu.:18910.5   3rd Qu.:2115
##                     Max.   :2017   Max.   :39131.0   Max.   :2385
##                                    NA's   :35
##  max_temperature min_temperature      aus               aman
##  Min.   :22.20   Min.   :10.10   Min.   :    49   Min.   :  9691
##  1st Qu.:29.48   1st Qu.:11.73   1st Qu.:   634   1st Qu.: 28041
##  Median :30.70   Median :12.60   Median :  1518   Median :108443
##  Mean   :30.63   Mean   :13.72   Mean   : 18252   Mean   :189617
##  3rd Qu.:32.35   3rd Qu.:14.50   3rd Qu.: 33841   3rd Qu.:261409
##  Max.   :35.60   Max.   :21.60   Max.   :108383   Max.   :647789
##
##       boro              wheat              potato            jute
##  Min.   :  95582   Min.   :  125.0   Min.   :  2711   Min.   :  2313
##  1st Qu.: 205682   1st Qu.:  487.5   1st Qu.: 18306   1st Qu.: 14460
##  Median : 232829   Median : 2334.5   Median : 40718   Median : 54384
##  Mean   : 469225   Mean   : 4171.2   Mean   : 38164   Mean   : 69623
##  3rd Qu.: 700441   3rd Qu.: 4846.8   3rd Qu.: 50869   3rd Qu.: 98398
##  Max.   :1425741   Max.   :17477.0   Max.   :100547   Max.   :250120
##
##     humidity         storm                 urea              tsp
##  Min.   :55.00   Length:70           Min.   : 16920   Min.   : 5044
##  1st Qu.:65.00   Class :character   1st Qu.: 29300   1st Qu.: 8369
##  Median :68.20   Mode  :character   Median : 32011   Median : 9642
```

```
##  Mean   :69.47                     Mean   : 54897   Mean   :15643
##  3rd Qu.:78.00                     3rd Qu.: 89272   3rd Qu.:19156
##  Max.   :80.00                     Max.   :130393   Max.   :49248
##
##        mp              DAP        inundationland_Highland
##  Min.   : 4808   Min.   :  443   Min.   :  9076
##  1st Qu.: 5738   1st Qu.: 1930   1st Qu.: 15749
##  Median : 9155   Median : 5834   Median : 30118
##  Mean   :15570   Mean   : 8273   Mean   : 60251
##  3rd Qu.:18656   3rd Qu.: 9003   3rd Qu.: 91012
##  Max.   :53832   Max.   :31785   Max.   :163675
##
##  inundationland_mediumhighland inundationland_lowland
##  Min.   : 13243                Min.   :14834
##  1st Qu.: 27893                1st Qu.:15220
##  Median : 32245                Median :22923
##  Mean   : 63121                Mean   :36198
##  3rd Qu.:133622                3rd Qu.:66018
##  Max.   :148743                Max.   :68689
##
##  inundationland_mediumlowland inundationland_verylowland Miscellaneous.Land
##  Min.   : 9442                Min.   :    0             Min.   : 9927
##  1st Qu.:15887                1st Qu.:    0             1st Qu.:17551
##  Median :20654                Median :    0             Median :30076
##  Mean   :29758                Mean   : 4364             Mean   :27556
##  3rd Qu.:41335                3rd Qu.:    0             3rd Qu.:31929
##  Max.   :77992                Max.   :30550             Max.   :50861
##
##  Calcareous.Alluvium Noncalcareous.Alluvium Acid.Basin.Clay
##  Min.   :   0.0      Min.   :  591          Min.   : 2189
##  1st Qu.:   0.0      1st Qu.: 2216          1st Qu.: 6253
##  Median :   0.0      Median : 6804          Median : 8655
##  Mean   : 701.2      Mean   : 8499          Mean   :13990
##  3rd Qu.:   0.0      3rd Qu.:10668          3rd Qu.:27877
##  Max.   :4952.0      Max.   :31611          Max.   :34907
##
##  Calcareous.Brown.Floodplain.Soil Calcareous.Grey.Floodplain.Soil
##  Min.   :   0                     Min.   :    0
##  1st Qu.:   0                     1st Qu.:    0
##  Median :   0                     Median :    0
##  Mean   :1101                     Mean   : 3761
##  3rd Qu.:   0                     3rd Qu.:    0
##  Max.   :7720                     Max.   :26371
##
##  Calcareous.Dark.Grey.Floodplain.Soil Noncalcareous.Grey.Floodplain.Soil
##  Min.   :    0                        Min.   : 12314
##  1st Qu.:    0                        1st Qu.: 21337
##  Median :    0                        Median : 86102
##  Mean   : 4085                        Mean   : 70607
##  3rd Qu.:    0                        3rd Qu.:103945
##  Max.   :28987                        Max.   :158905
##
##  Noncalcareous.Dark.Grey.Floodplain.Soil      Peat          Made.Land
##  Min.   :  1876                          Min.   :    0   Min.   :   0.0
```

```
##  1st Qu.:  4930                          1st Qu.:    0   1st Qu.:   0.0
##  Median : 27363                          Median :    0   Median :   0.0
##  Mean   : 49371                          Mean   : 2020   Mean   :1078.3
##  3rd Qu.: 83093                          3rd Qu.:    0   3rd Qu.: 418.2
##  Max.   :189477                          Max.   :14277   Max.   :7157.0
##
##  Noncalcareous.Brown.Floodplain.Soil Shallow.Red.Brown.Terrace.Soil
##  Min.   :    0.0                     Min.   :    0.00
##  1st Qu.:    0.0                     1st Qu.:   39.25
##  Median :   17.5                     Median : 2569.00
##  Mean   : 2527.1                     Mean   : 8650.96
##  3rd Qu.:  522.0                     3rd Qu.:14063.00
##  Max.   :20826.0                     Max.   :32923.00
##
##  Deep.Red.Brown.Terrace.Soil Brown.Mottled.Terrace.Soil
##  Min.   :    0               Min.   :   0
##  1st Qu.: 1088               1st Qu.:   0
##  Median : 9023               Median : 716
##  Mean   :18737               Mean   :1170
##  3rd Qu.:40005               3rd Qu.:2402
##  Max.   :50661               Max.   :3080
##
##  Shallow.Grey.Terrace.Soil Deep.Grey.Terrace.Soil Grey.Valley.Soil
##  Min.   :   0.0            Min.   :    0          Min.   :    0
##  1st Qu.:   0.0            1st Qu.:    0          1st Qu.:    0
##  Median : 249.5            Median :  142          Median :  789
##  Mean   :1737.2            Mean   : 3347          Mean   : 6076
##  3rd Qu.:1909.8            3rd Qu.: 6299          3rd Qu.:11650
##  Max.   :9873.0            Max.   :16808          Max.   :20834
##
##  Brown.Hill.Soil  Grey.Piedmont.Soil soil.moisture
##  Min.   :  0.00   Min.   :    0      Min.   : 59606
##  1st Qu.:  0.00   1st Qu.:    0      1st Qu.: 94967
##  Median :  0.00   Median :    0      Median :154705
##  Mean   : 58.87   Mean   : 2128      Mean   :193794
##  3rd Qu.:  0.00   3rd Qu.:    0      3rd Qu.:306446
##  Max.   :493.00   Max.   :18396      Max.   :391055
##
```

```
# Adding Average Tmperatrue in Dataset

Agri_Data <- Agri_Data %>%
  mutate(Avg_temp = (min_temperature+max_temperature)/2)

head(Agri_Data)
```

```
##      District year area avg_rainfall max_temperature min_temperature   aus
## 1       dhaka 2008   NA         2385            34.2            12.5   804
## 2     gazipur 2008   NA         2197            30.2            19.5  4328
## 3   narsingdi 2008   NA         2197            34.2            12.5   484
## 4 narayangonj 2008   NA         2197            34.2            12.2  1617
## 5     tangail 2008   NA         1856            34.1            10.2   447
## 6  mymenshing 2008   NA         2239            29.6            21.6 88198
##     aman    boro wheat potato   jute humidity storm   urea    tsp     mp    DAP
```

```
## 1   9691 233939  1129  33679  29825       71   yes  25967  8262  4808  1573
## 2  93956 208434   479   4300  15072       71   yes  29264  9225  9057  5721
## 3  93958 207669  2073  14323  42440       71    no  30273  5068  5113  3031
## 4  14485 129390  4585  73525   3620       71   yes  16992  7954  7026   443
## 5 170918 672851 17477  56693 125166       78   yes  91562 19022 20143  8809
## 6 491413 812189  6111  43755  50856       79    no 120279 34007 47121 28723
##   inundationland_Highland inundationland_mediumhighland inundationland_lowland
## 1                   30118                         32245                 20539
## 2                   88346                         30449                 15220
## 3                   23779                         27893                 22923
## 4                    9076                         13243                 14834
## 5                   91012                        133622                 66018
## 6                  163675                        148743                 68689
##   inundationland_mediumlowland inundationland_verylowland Miscellaneous.Land
## 1                        41335                          0              30137
## 2                        20654                          0              16542
## 3                        20368                          0              20574
## 4                        22629                          0              15871
## 5                        15887                          0              31462
## 6                         9442                          0              44852
##   Calcareous.Alluvium Noncalcareous.Alluvium Acid.Basin.Clay
## 1                4869                   6782            6253
## 2                   0                    648           28245
## 3                   0                   8247            4942
## 4                   0                   2396            3263
## 5                   0                  28897            7786
## 6                   0                  11988           32531
##   Calcareous.Brown.Floodplain.Soil Calcareous.Grey.Floodplain.Soil
## 1                             7691                           26289
## 2                                0                               0
## 3                                0                               0
## 4                                0                               0
## 5                                0                               0
## 6                                0                               0
##   Calcareous.Dark.Grey.Floodplain.Soil Noncalcareous.Grey.Floodplain.Soil
## 1                                28109                              87891
## 2                                    0                              14655
## 3                                    0                              41976
## 4                                    0                              18095
## 5                                    0                             128008
## 6                                    0                              88491
##   Noncalcareous.Dark.Grey.Floodplain.Soil  Peat Made.Land
## 1                                    1876 13900      7048
## 2                                    4899     0         0
## 3                                   27689     0         0
## 4                                   26925     0       492
## 5                                   47714     0         0
## 6                                  165795     0         0
##   Noncalcareous.Brown.Floodplain.Soil Shallow.Red.Brown.Terrace.Soil
## 1                                   0                              0
## 2                                  13                          29623
## 3                                   0                            881
## 4                                   0                           2800
## 5                                 566                          14764
```

```
## 6                              14725                          12158
##   Deep.Red.Brown.Terrace.Soil Brown.Mottled.Terrace.Soil
## 1                           0                          0
## 2                       46500                       2414
## 3                        9800                        601
## 4                        4840                         30
## 5                       47378                       2423
## 6                       28174                       2689
##   Shallow.Grey.Terrace.Soil Deep.Grey.Terrace.Soil Grey.Valley.Soil
## 1                         0                      0                0
## 2                      8375                   4660            14637
## 3                         0                    140              687
## 4                       213                      0              728
## 5                      1971                   6390            20642
## 6                      1846                  10150             7915
##   Brown.Hill.Soil Grey.Piedmont.Soil soil.moisture Avg_temp
## 1               0                  0        124237    23.35
## 2               0                  0        154669    24.85
## 3               0                  0         94963    23.35
## 4               0                  0         59782    23.20
## 5               0                  0        306539    22.15
## 6             442              13645        390549    25.60
```

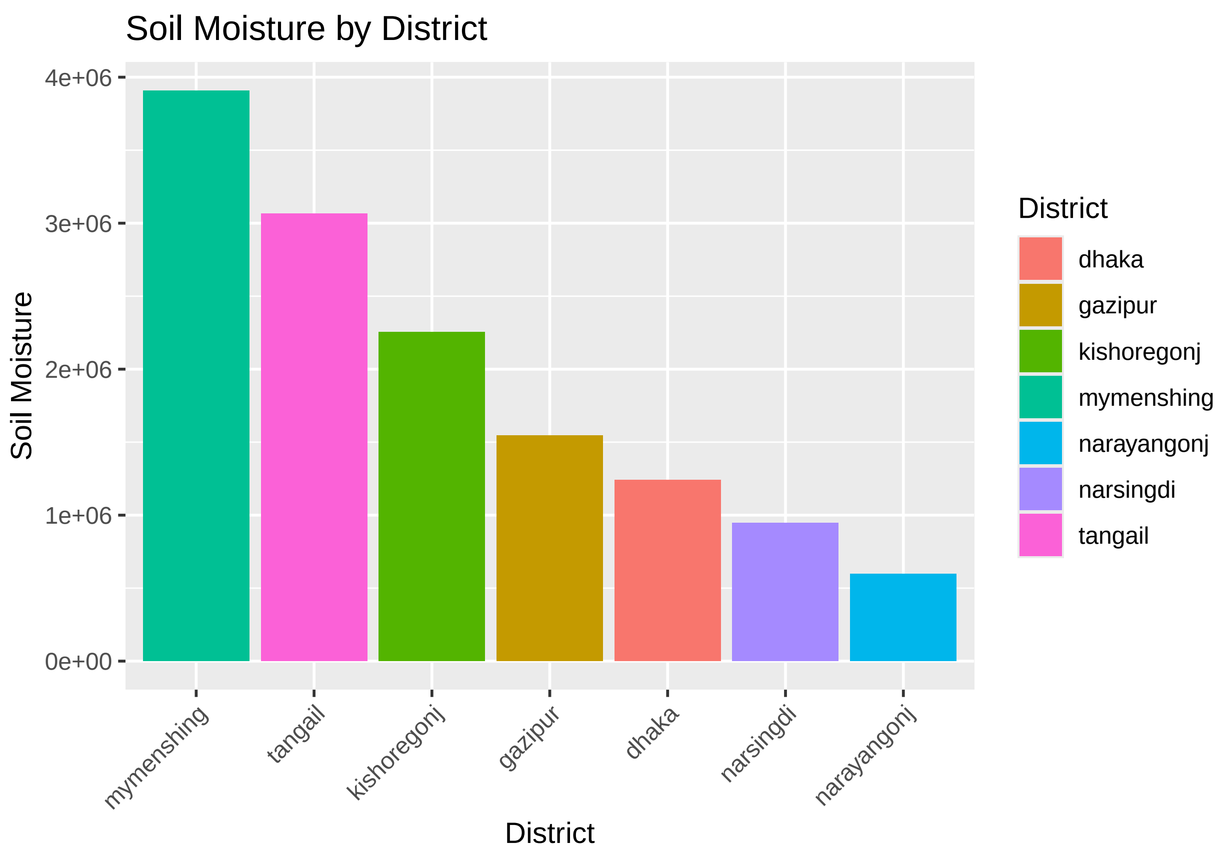
```r
View(Agri_Data)

# Soil Moisture by District

soil_moisture_by_district<- Agri_Data %>%
  arrange(desc(soil.moisture)) %>%
  select(District, soil.moisture,year)



# Highest Soil Moisture
head(soil_moisture_by_district,n=1)
```

```
##     District soil.moisture year
## 1 mymenshing        391055 2010
```

```r
# Lowest Soil Moisture
tail(soil_moisture_by_district,n=1)
```

```
##       District soil.moisture year
## 70 narayangonj         59606 2016
```

```r
# Data Visualization


ggplot(Agri_Data, aes(x = reorder(District, -soil.moisture), y = soil.moisture, fill = District)) +
  geom_bar(stat = "identity") +
  labs(title = "Soil Moisture by District", x = "District", y = "Soil Moisture") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



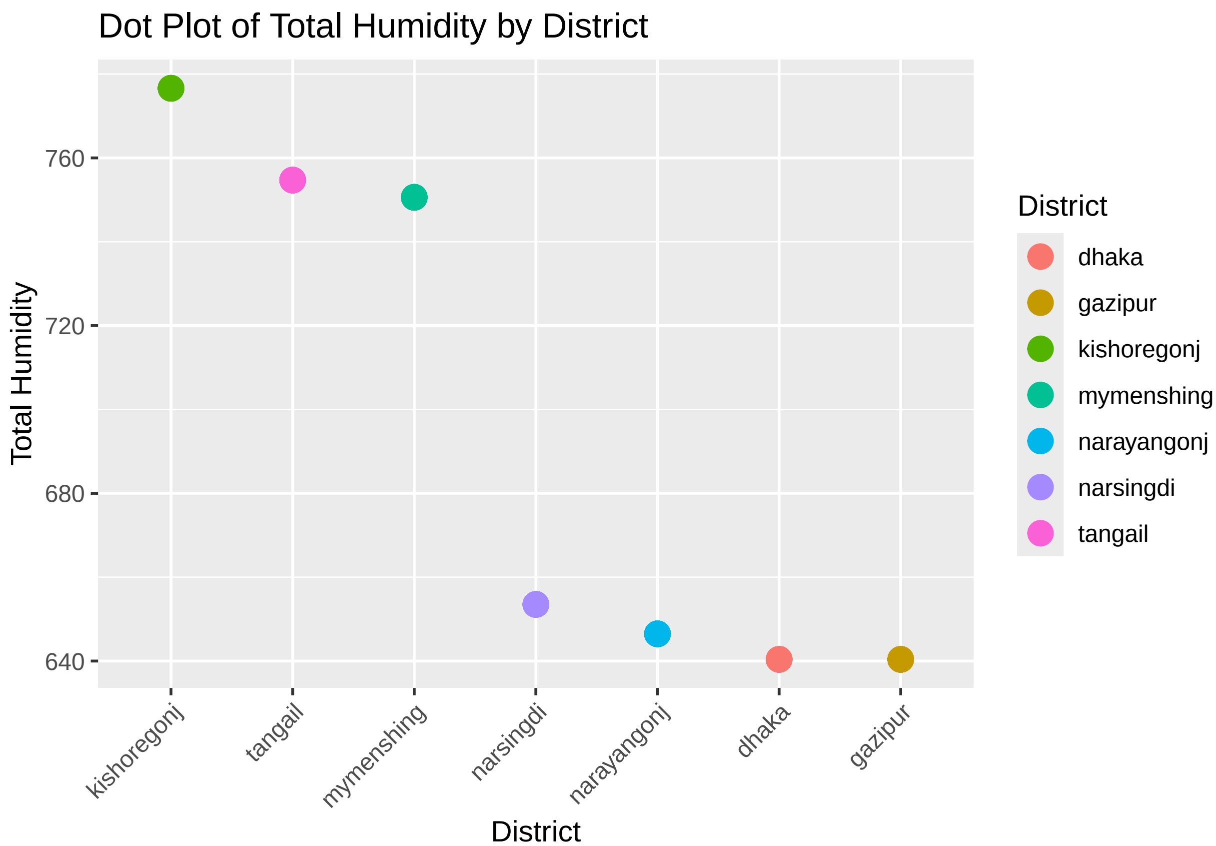

```r
# Group by district and calculate the total (sum) of Humidity for each district
humidity_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_humidity = sum(humidity, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(humidity_sum_by_district, by = "District")



# Total Rainfall

total_humidity_by_district <- Agri_Data %>%
  arrange(desc(total_humidity)) %>%
  select(District, total_humidity)


# Highest Total Humidity by District
head(total_humidity_by_district,n=1)
```

```
##      District total_humidity
## 1 kishoregonj          776.6
```

```r
# Lowest Total Rainfall By District
tail(total_humidity_by_district, n=1)
```

```
##    District total_humidity
## 70  gazipur          640.4
```

```r
# Data Visualization

ggplot(Agri_Data, aes(x = reorder(District, -total_humidity), y = total_humidity, color = District)) +
  geom_point(size = 4) +
  labs(title = "Dot Plot of Total Humidity by District", x = "District", y = "Total Humidity") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
# Group by district and calculate the total (sum) of rainfall for each district
rainfall_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_rainfall = sum(avg_rainfall, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(rainfall_sum_by_district, by = "District")



# Total Rainfall

total_rainfall_by_district <- Agri_Data %>%
  arrange(desc(total_rainfall)) %>%
  select(District, total_rainfall)
```

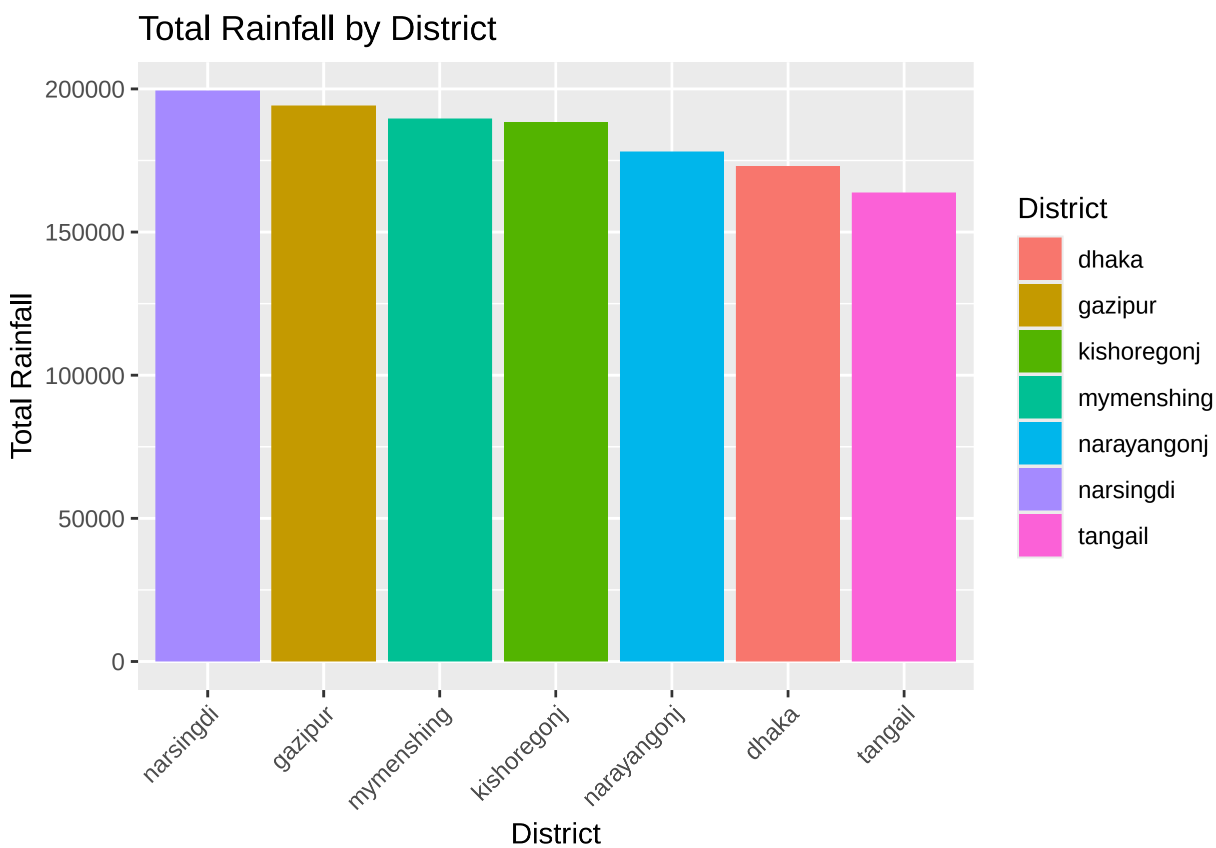
```r
# Highest Total rainfall by District
head(total_rainfall_by_district,n=1)
```

```
##    District total_rainfall
## 1 narsingdi          19935
```

```r
# Lowest Total Rainfall By District
tail(total_rainfall_by_district, n=1)
```

```
##    District total_rainfall
## 70  tangail          16380
```

```r
# Data Visualization

ggplot(Agri_Data, aes(x = reorder(District, -total_rainfall), y = total_rainfall, fill = District)) +
  geom_bar(stat = "identity") +
  labs(title = "Total Rainfall by District", x = "District", y = "Total Rainfall") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
# Group by district and calculate the total (sum) of Tempeture for each district

temp_sum_by_district <- Agri_Data %>%
```

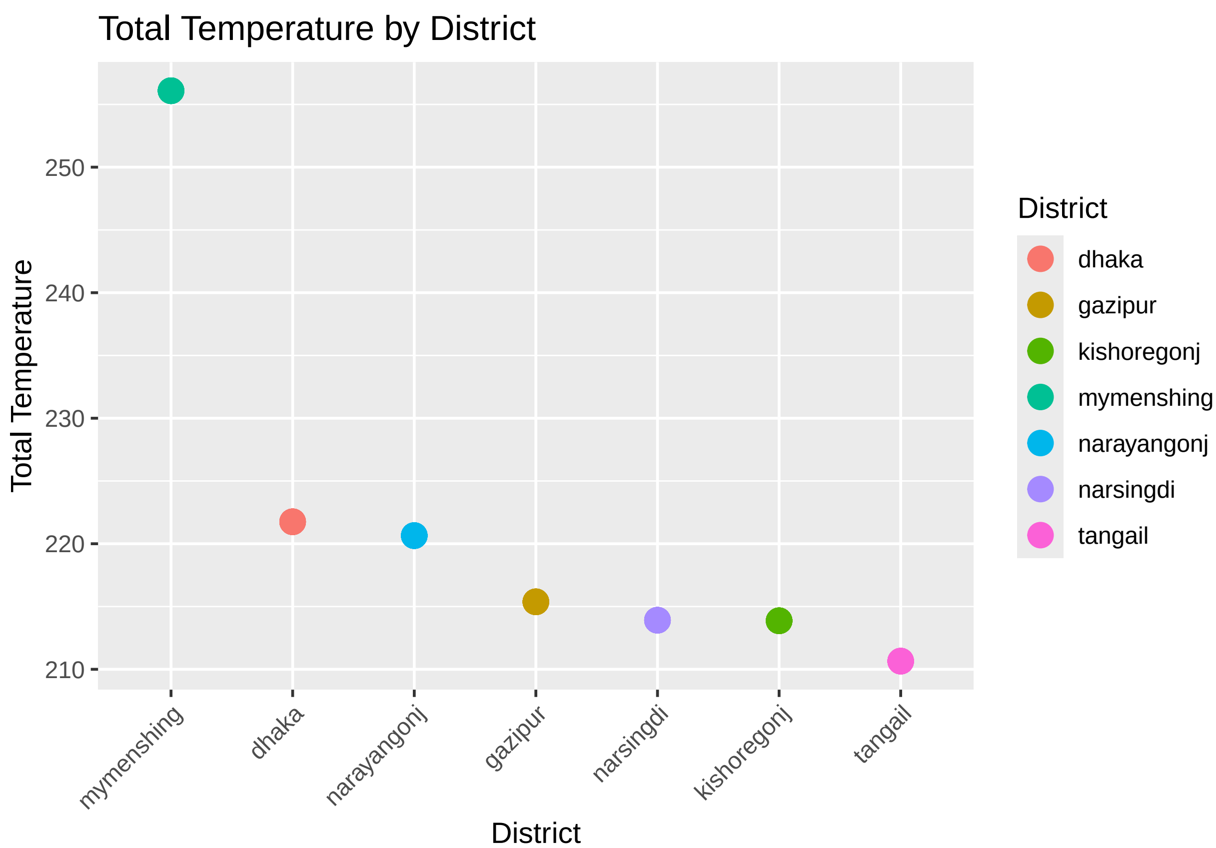
```r
  group_by(District) %>%
  summarise(total_temp = sum(Avg_temp, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(temp_sum_by_district, by = "District")


# Total Temperature

total_temp_by_district <- Agri_Data %>%
  arrange(desc(total_temp)) %>%
  select(District, total_temp)


# Highest Total Temperature by District
head(total_temp_by_district,n=1)
```

```
##     District total_temp
## 1 mymenshing     256.09
```

```r
# Lowest Total Temperature by District
tail(total_temp_by_district, n=1)
```

```
##    District total_temp
## 70  tangail     210.65
```

```r
# Data Visualization

ggplot(Agri_Data, aes(x = reorder(District, -total_temp), y = total_temp)) +
  geom_point(aes(color = District), size = 4) +
  labs(title = "Total Temperature by District", x = "District", y = "Total Temperature") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



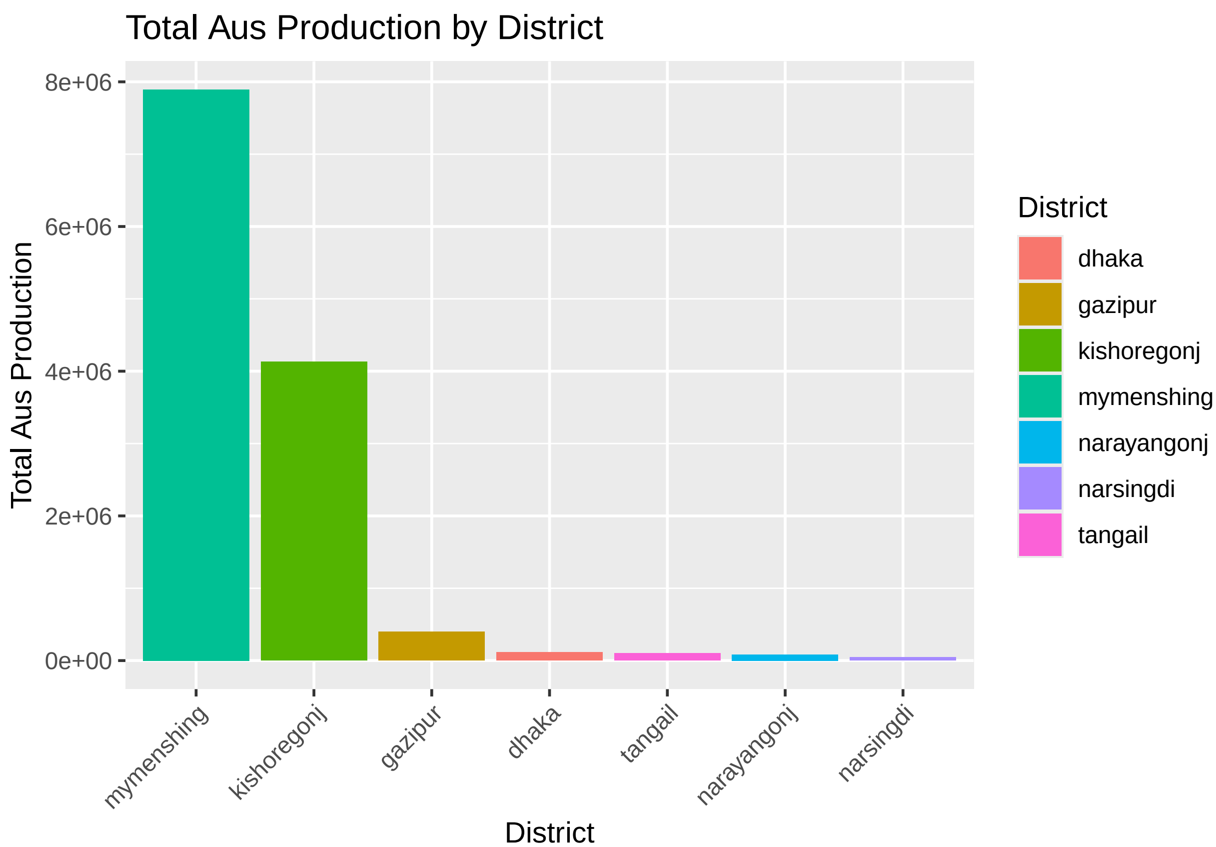

```r
# Group by district and calculate the total (sum) of Aus for each district

aus_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_aus = sum(aus, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(aus_sum_by_district, by = "District")


# Total Aus Production by District

total_aus_by_district <- Agri_Data %>%
  arrange(desc(total_aus)) %>%
  select(District, total_aus)


# Highest total Aus Production by District
head(total_aus_by_district,n=1)
```

```
##     District total_aus
## 1 mymenshing    788904
```

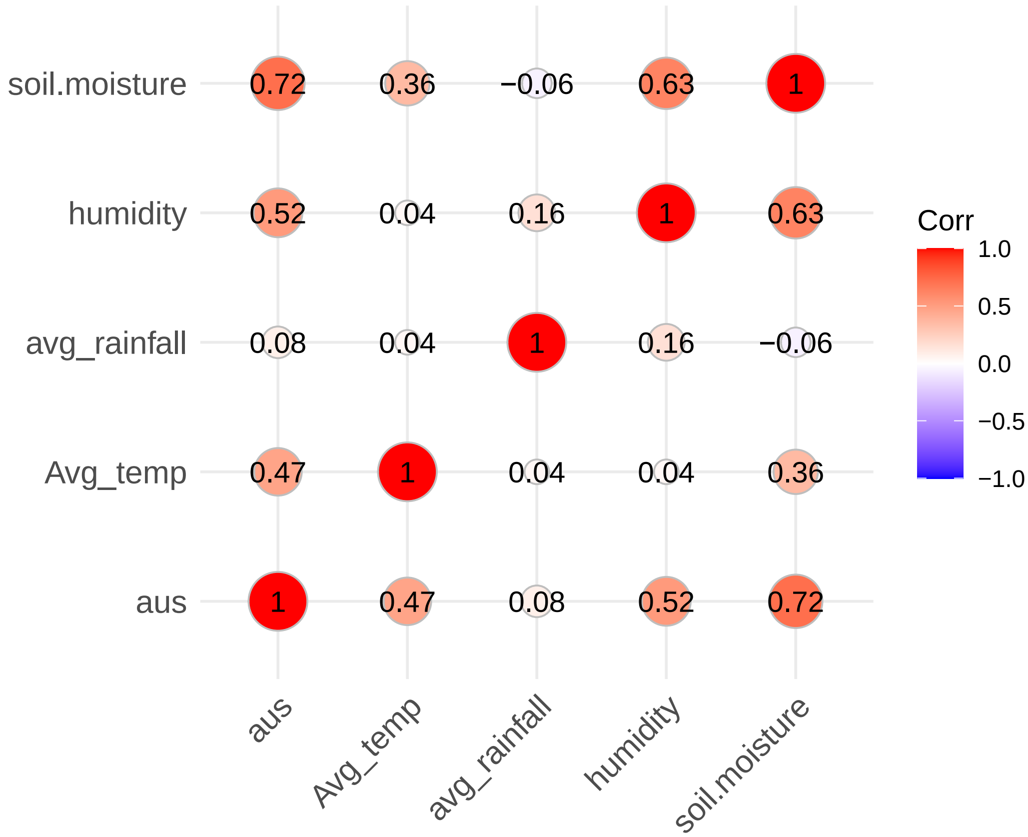
```r
# Lowest Total Aus Production by District
tail(total_aus_by_district, n=1)
```

```
##     District total_aus
## 70 narsingdi      4813
```

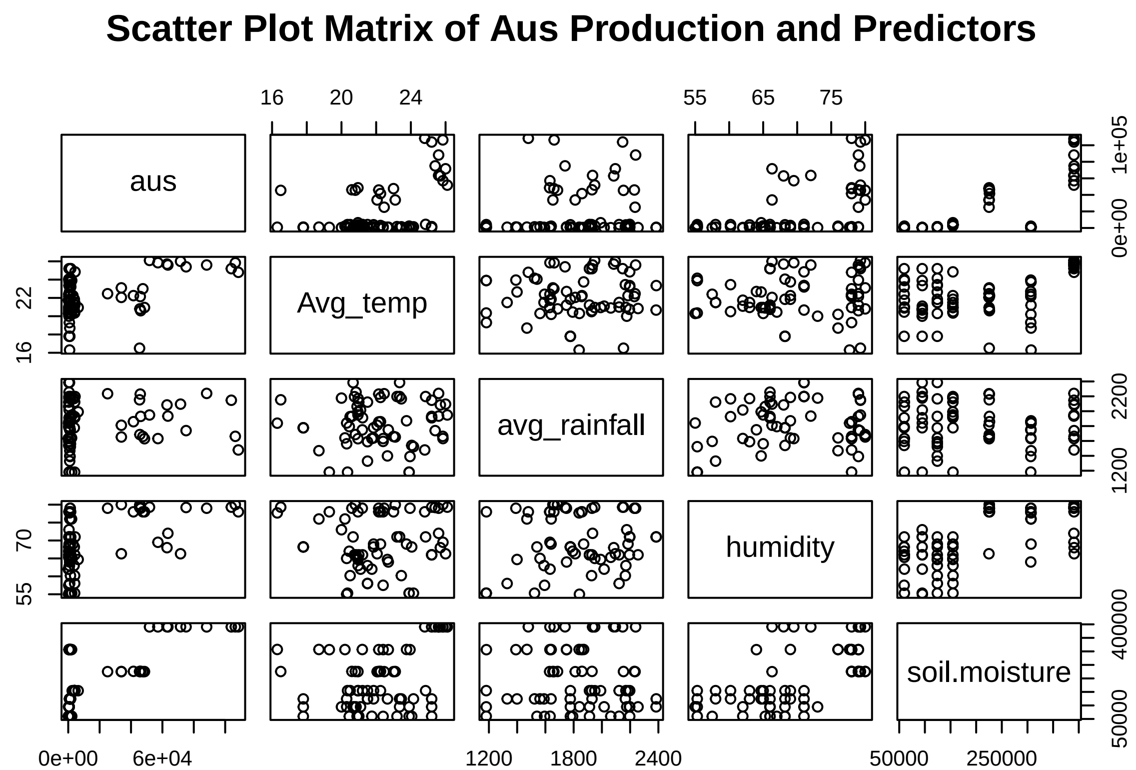
```r
# Data Visualization for Total Aus Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_aus), y = total_aus)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Aus Production by District", x = "District", y = "Total Aus Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
## Data Analysis of Aus Production based on Average Temperature, Average Rainfall, Humidity and Soil Mo

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(aus, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```

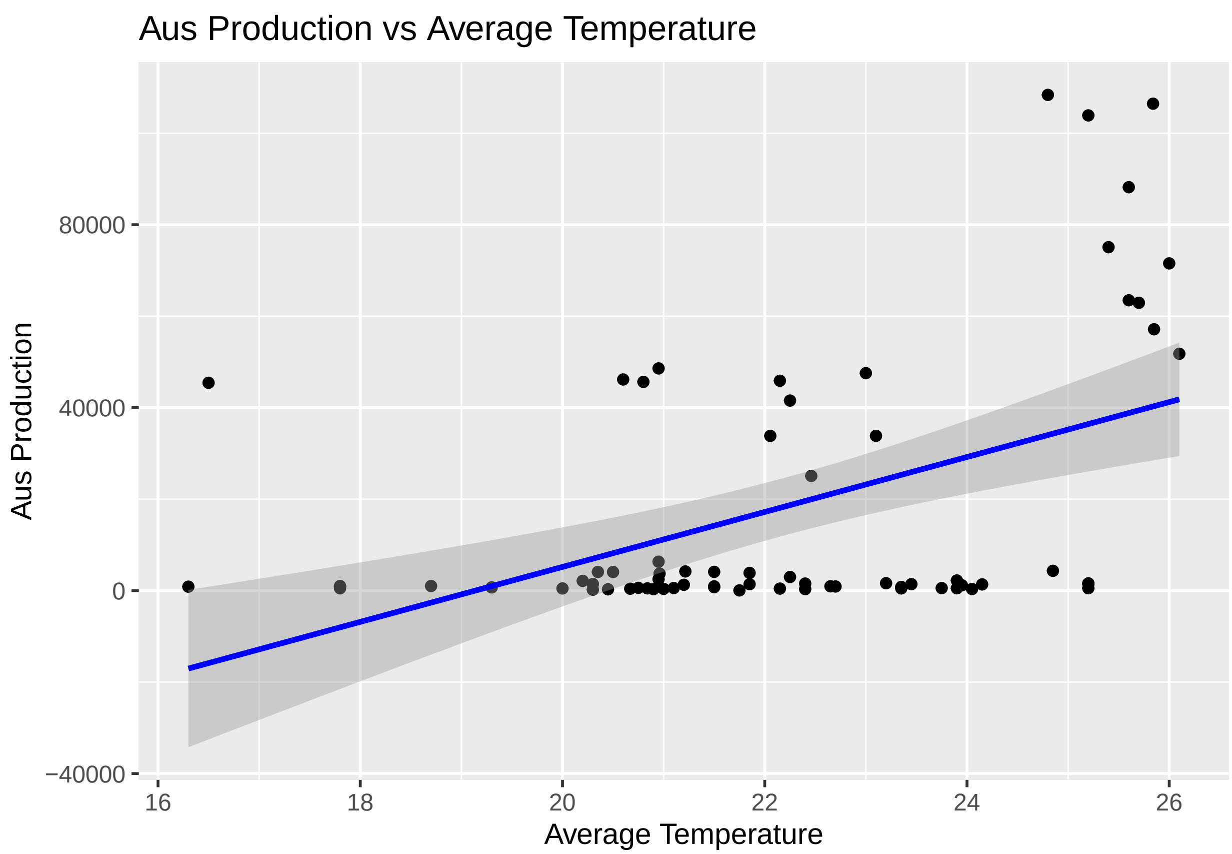



```r
# Scatter plot matrix to visualize relationships
pairs(~ aus + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Aus Production and Predictors")
```



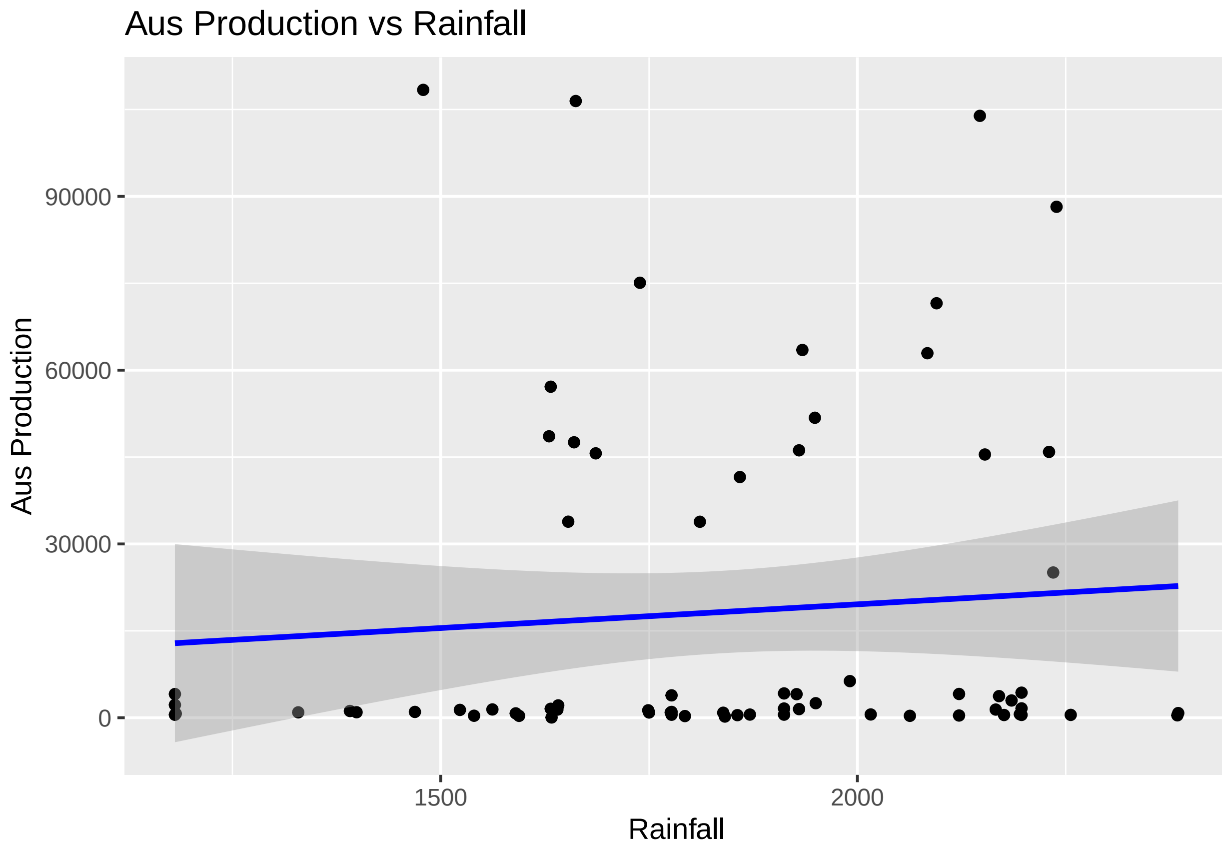

```
## Indiividual Scatter Plots
#  Aus Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = aus)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aus Production vs Average Temperature", x = "Average Temperature", y = "Aus Production")
```

```
## ‘geom_smooth()‘ using formula = ’y ~ x’
```



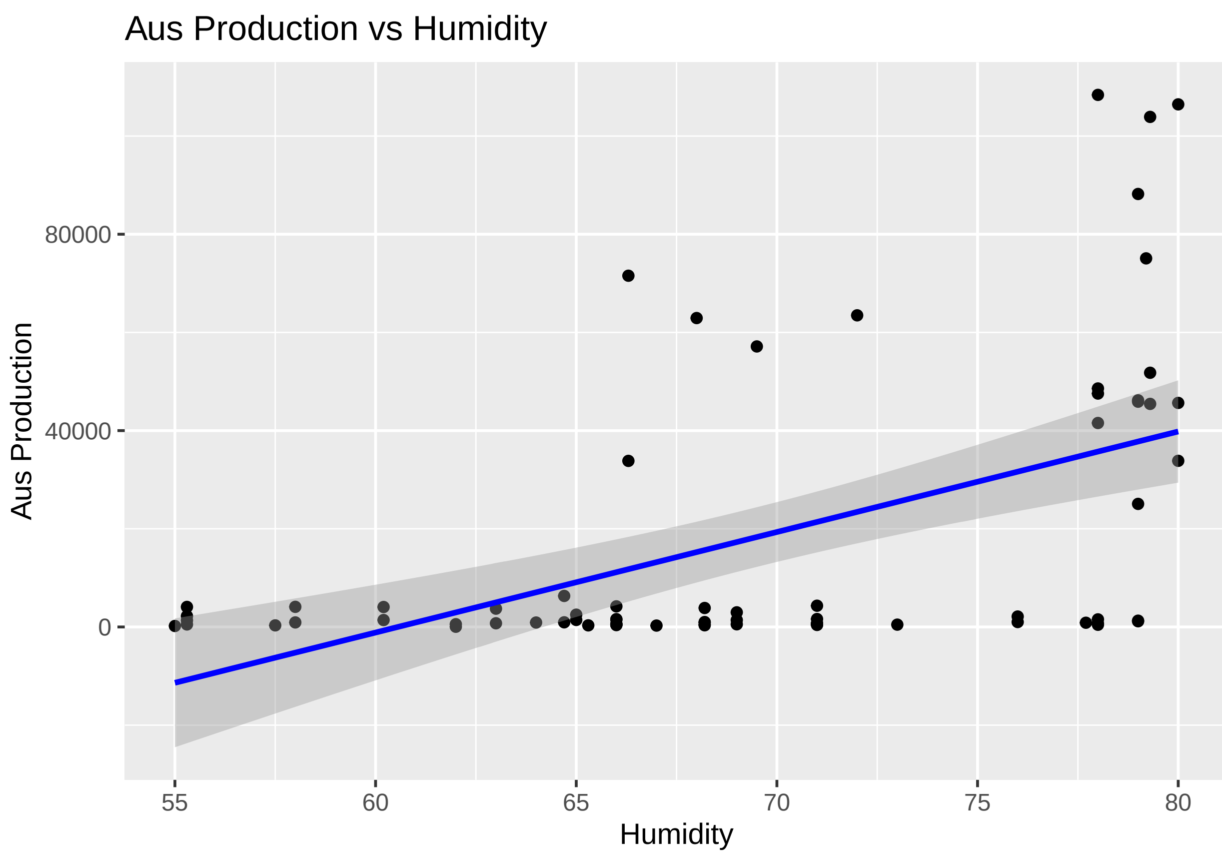

```r
# Aus Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = aus)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aus Production vs Rainfall", x = "Rainfall", y = "Aus Production")
```

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



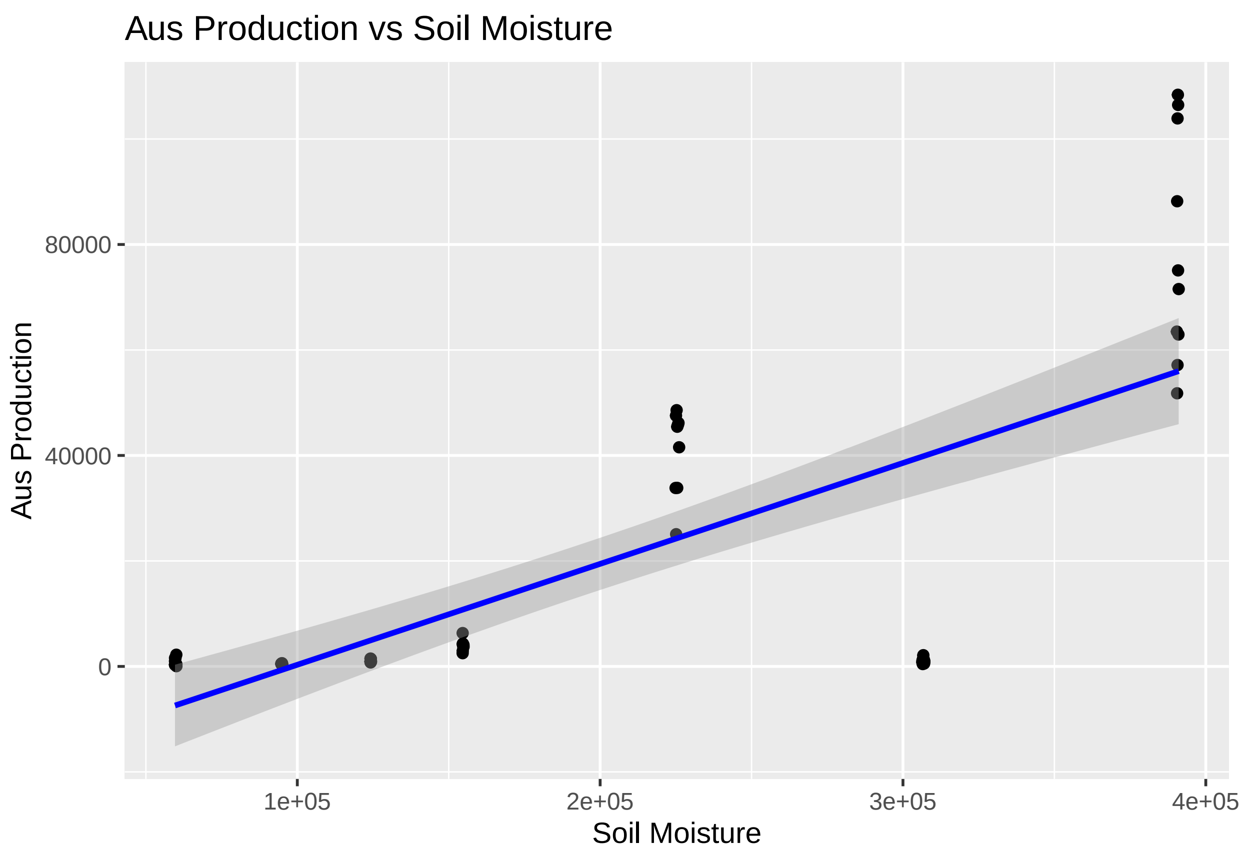

```r
# Aus Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = aus)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aus Production vs Humidity", x = "Humidity", y = "Aus Production")
```

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




```r
#  Aus Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = aus)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aus Production vs Soil Moisture", x = "Soil Moisture", y = "Aus Production")
```

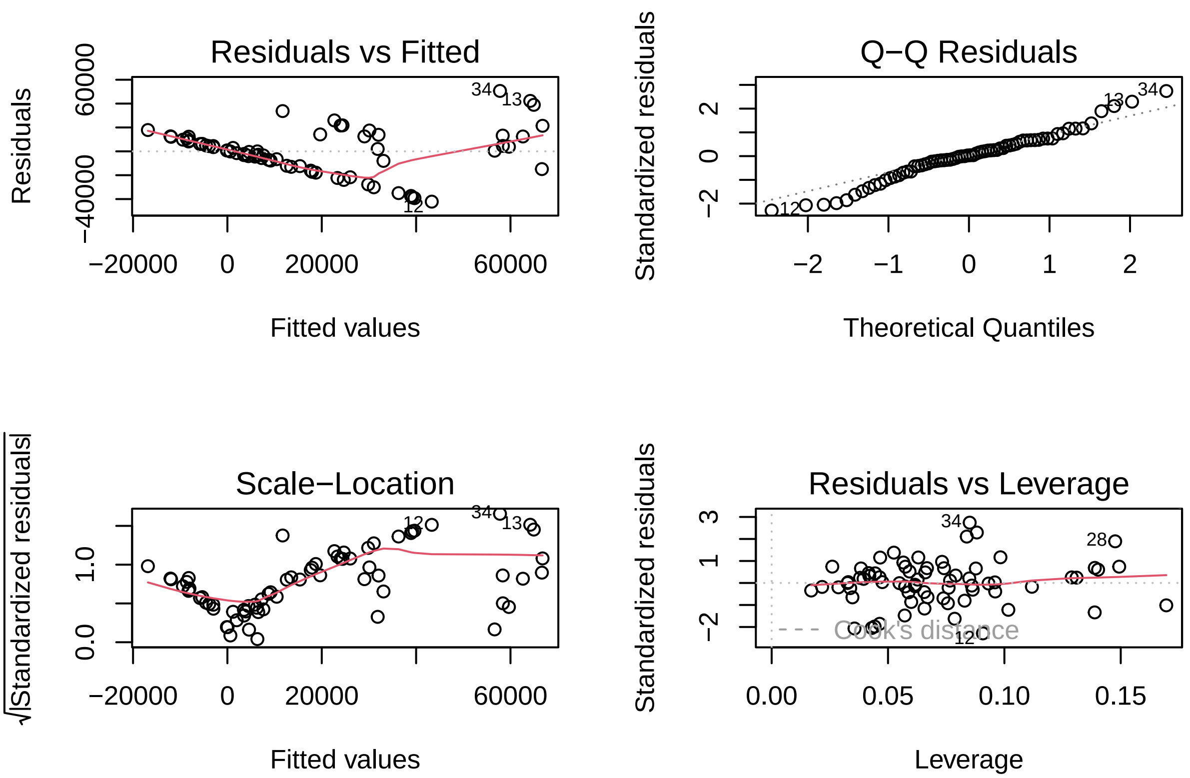
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(aus ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = aus ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -42153  -7946    606  11815  50606
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.484e+05  3.827e+04  -3.879 0.000247 ***
## Avg_temp       3.530e+03  1.115e+03   3.167 0.002347 **
## avg_rainfall   7.275e+00  8.041e+00   0.905 0.368943
## humidity       6.989e+02  4.216e+02   1.658 0.102193
## soil.moisture  1.366e-01  3.032e-02   4.505 2.83e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 19310 on 65 degrees of freedom
## Multiple R-squared:  0.6022, Adjusted R-squared:  0.5777
## F-statistic: 24.59 on 4 and 65 DF,  p-value: 2.014e-12
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
# Group by district and calculate the total (sum) of Aman for each district

aman_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_aman = sum(aman, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(aman_sum_by_district, by = "District")


# Total Aus Production by District

total_aman_by_district <- Agri_Data %>%
  arrange(desc(total_aman)) %>%
  select(District, total_aman)


# Highest total Aus Production by District
head(total_aman_by_district,n=1)
```

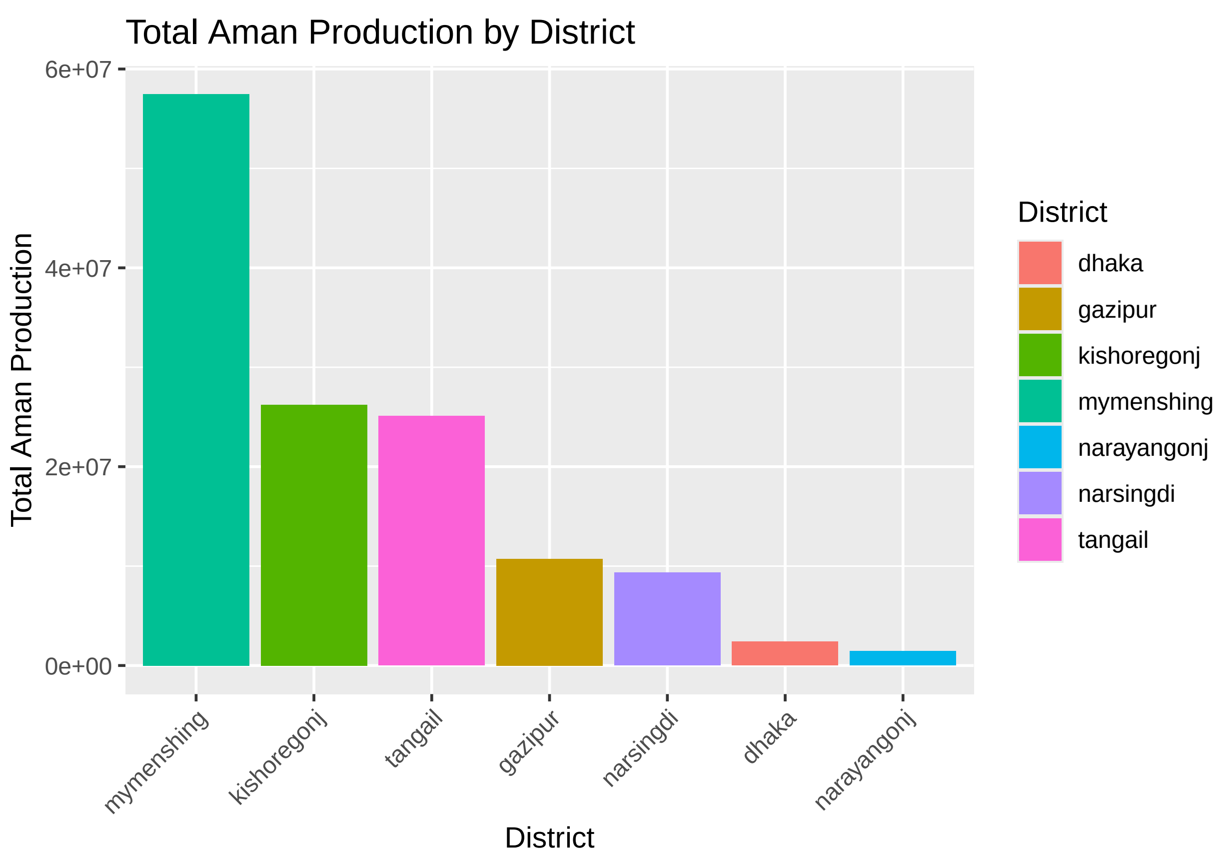
```
##      District total_aman
```

```
## 1 mymenshing    5744216
```

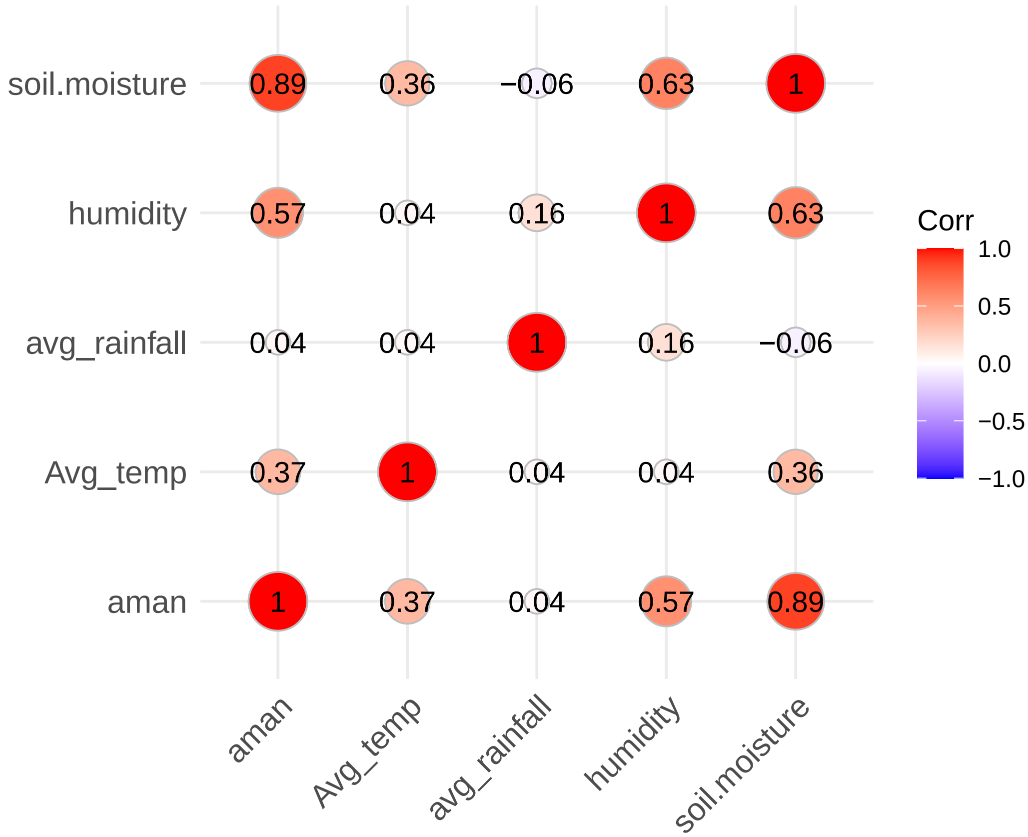
```r
# Lowest Total Aus Production by District
tail(total_aman_by_district, n=1)
```

```
##       District total_aman
## 70 narayangonj     148773
```

```r
# Data Visualization for Total Aus Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_aman), y = total_aman)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Aman Production by District", x = "District", y = "Total Aman Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

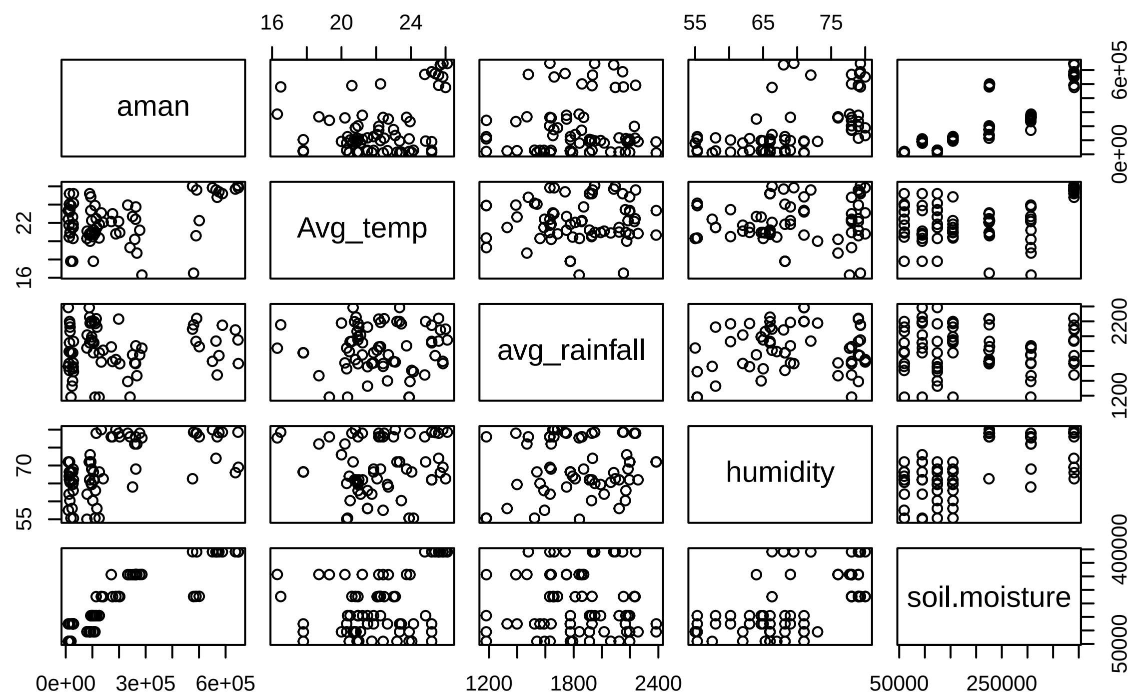



```r
## Data Analysis of Aus Production based on Average Temperature, Average Rainfall, Humidity and Soil Mo

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(aman, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```



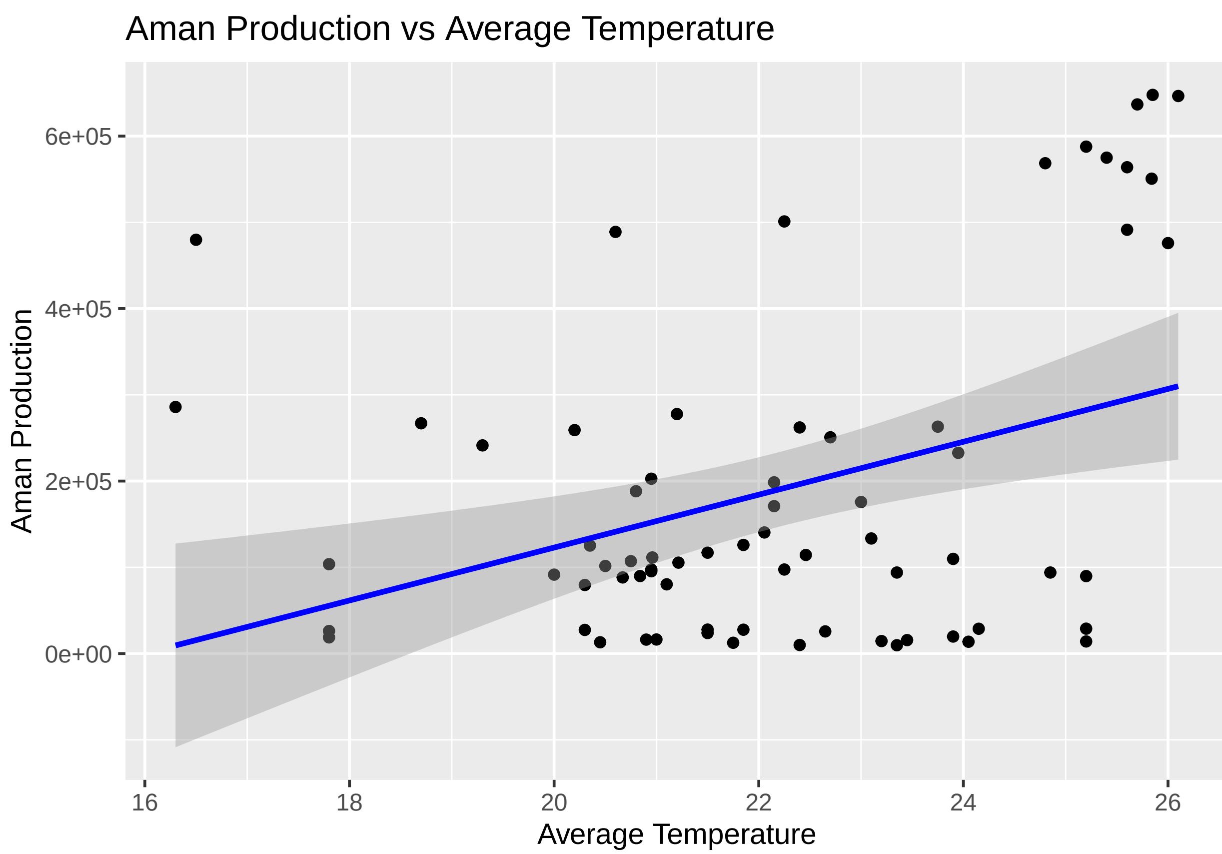

```r
# Scatter plot matrix to visualize relationships
pairs(~ aman + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Aman Production and Predictors")
```

Scatter Plot Matrix of Aman Production and Predictors



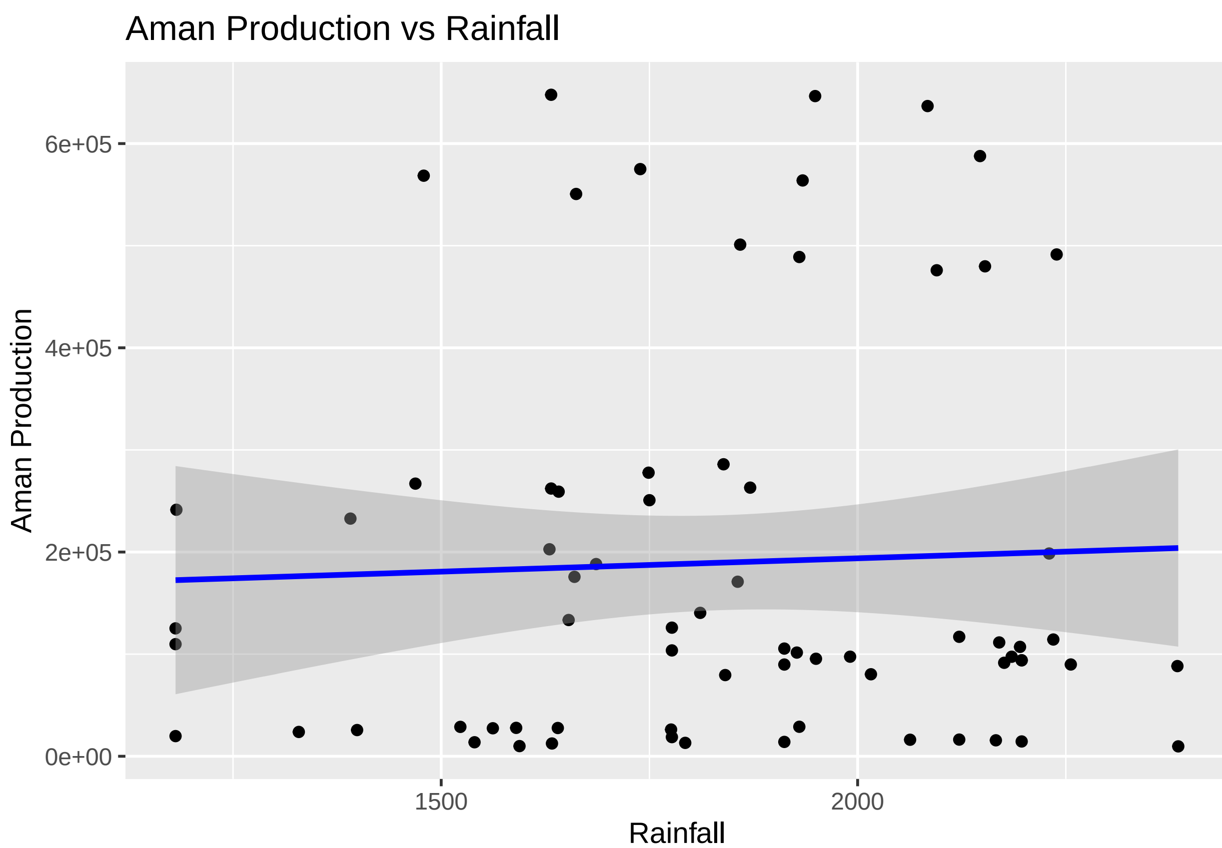

```
## Indiividual Scatter Plots
#  Aman Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = aman)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aman Production vs Average Temperature", x = "Average Temperature", y = "Aman Productio
```

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



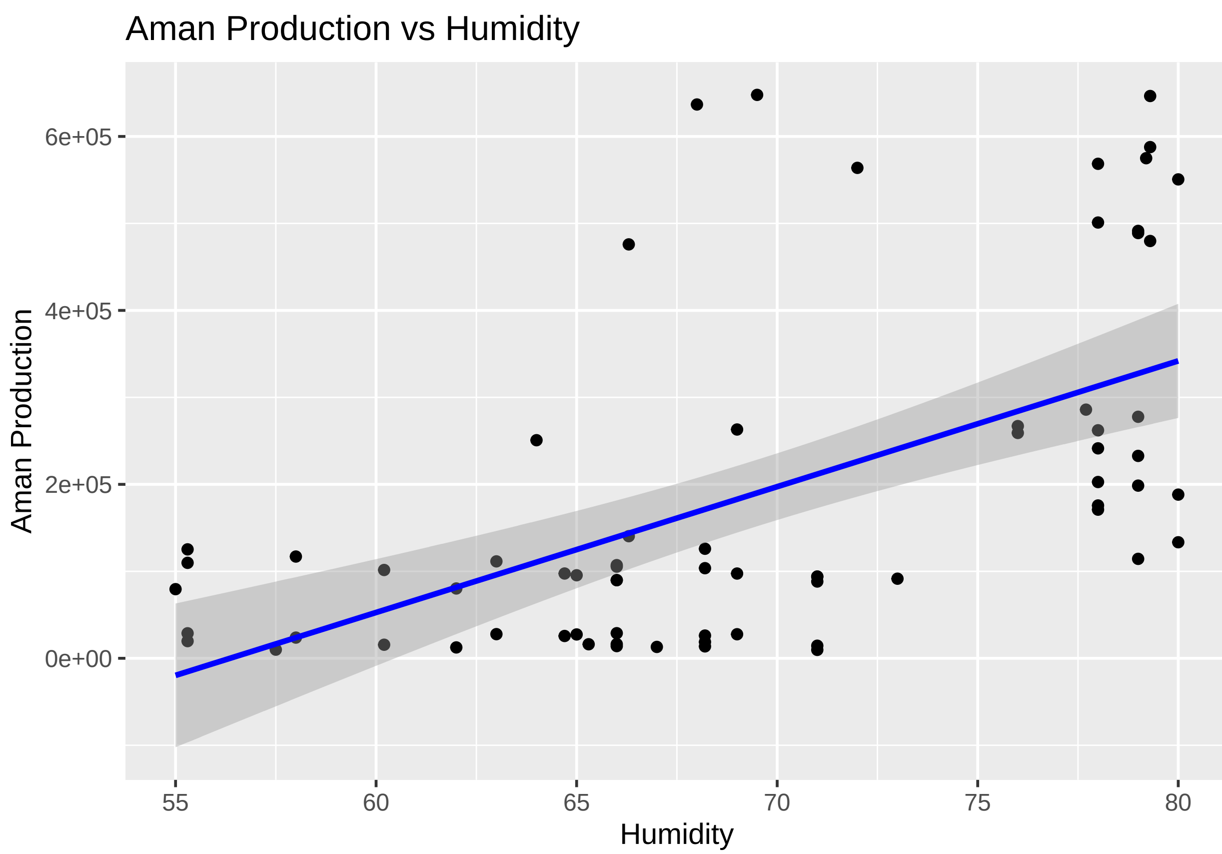

```r
# Aus Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = aman)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aman Production vs Rainfall", x = "Rainfall", y = "Aman Production")
```

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



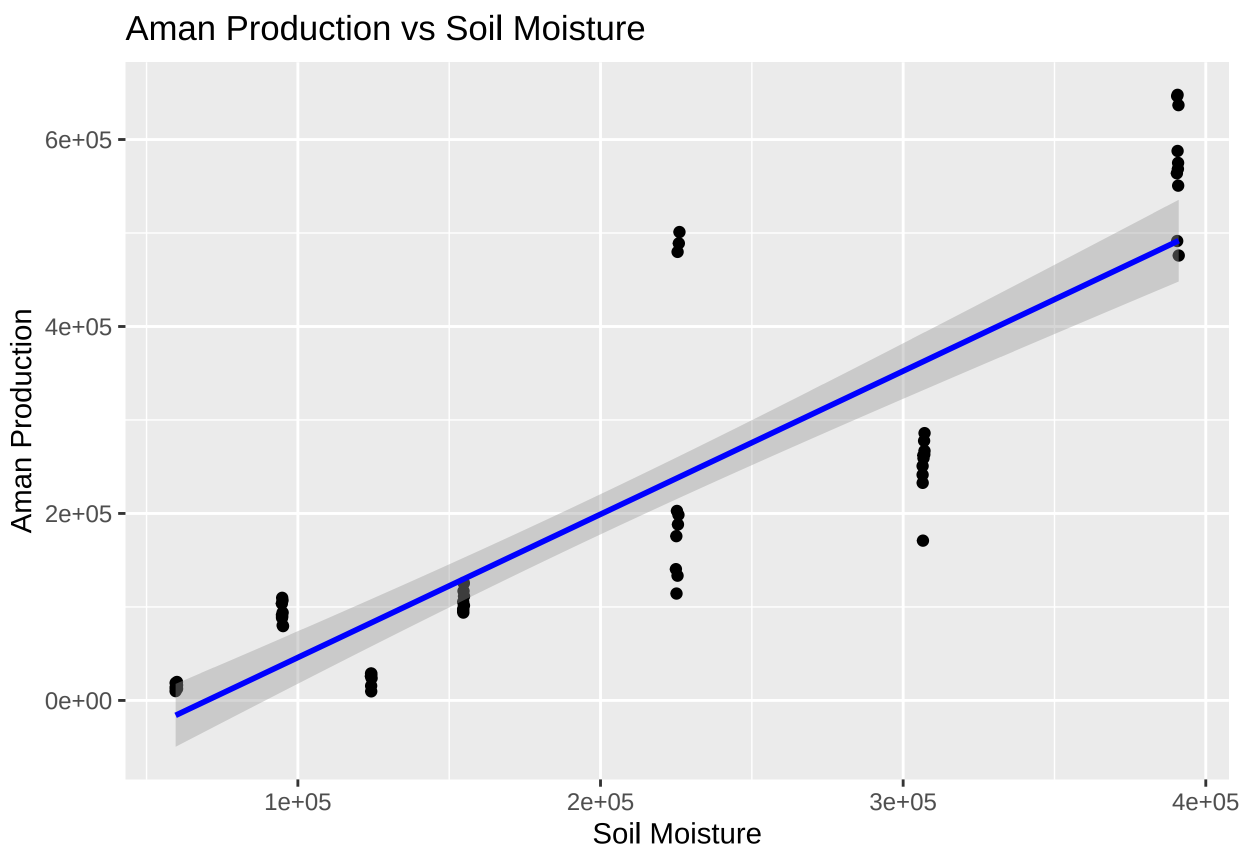

```r
# Aus Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = aman)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aman Production vs Humidity", x = "Humidity", y = "Aman Production")
```

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




```r
#  Aus Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = aman)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Aman Production vs Soil Moisture", x = "Soil Moisture", y = "Aman Production")
```

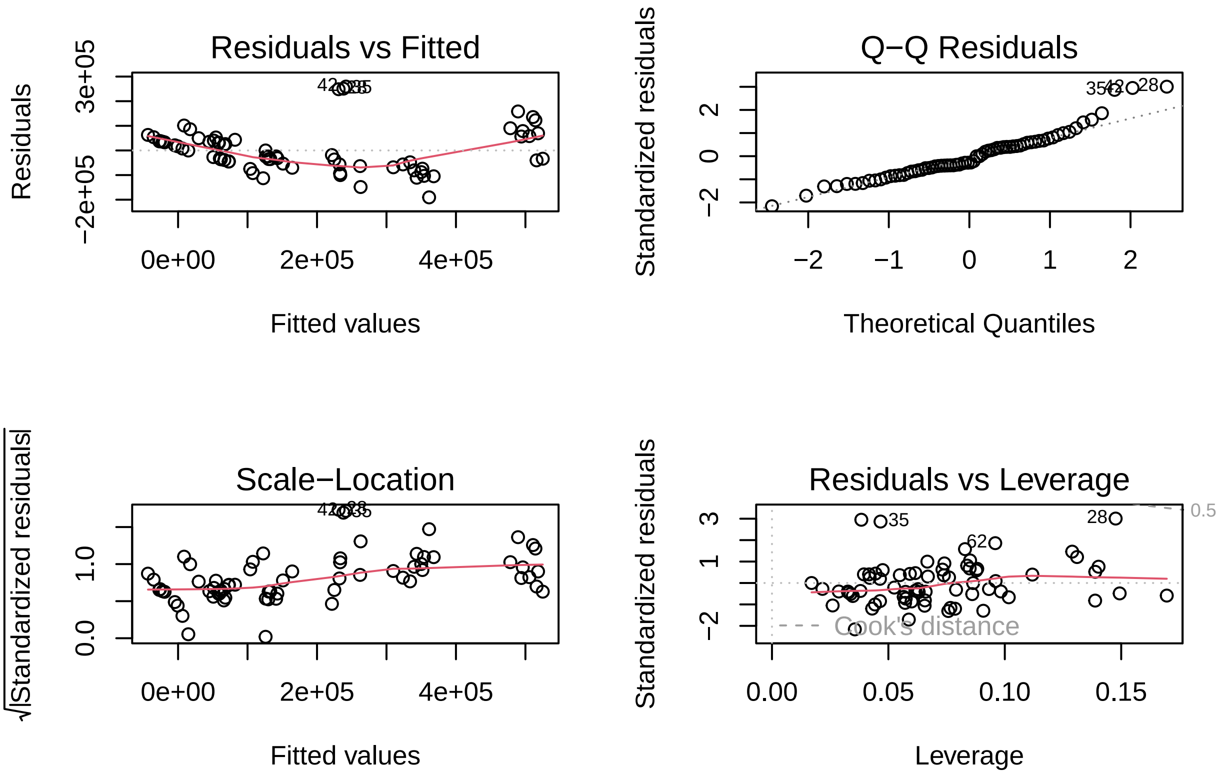
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(aman ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = aman ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -190463  -55857  -24297   42493  259341
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.256e+05  1.776e+05  -1.833   0.0714 .
## Avg_temp       4.642e+03  5.176e+03   0.897   0.3731
## avg_rainfall   5.668e+01  3.733e+01   1.518   0.1338
## humidity       2.637e+02  1.957e+03   0.135   0.8932
## soil.moisture  1.495e+00  1.408e-01  10.623 7.61e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 89630 on 65 degrees of freedom
## Multiple R-squared:  0.798,  Adjusted R-squared:  0.7856
## F-statistic:  64.2 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
## Group by district and calculate the total (sum) of Boro for each district

boro_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_boro = sum(boro, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(boro_sum_by_district, by = "District")


# Total Boro Production by District

total_boro_by_district <- Agri_Data %>%
  arrange(desc(total_boro)) %>%
  select(District, total_boro)


# Highest total Boro Production by District
head(total_boro_by_district,n=1)
```

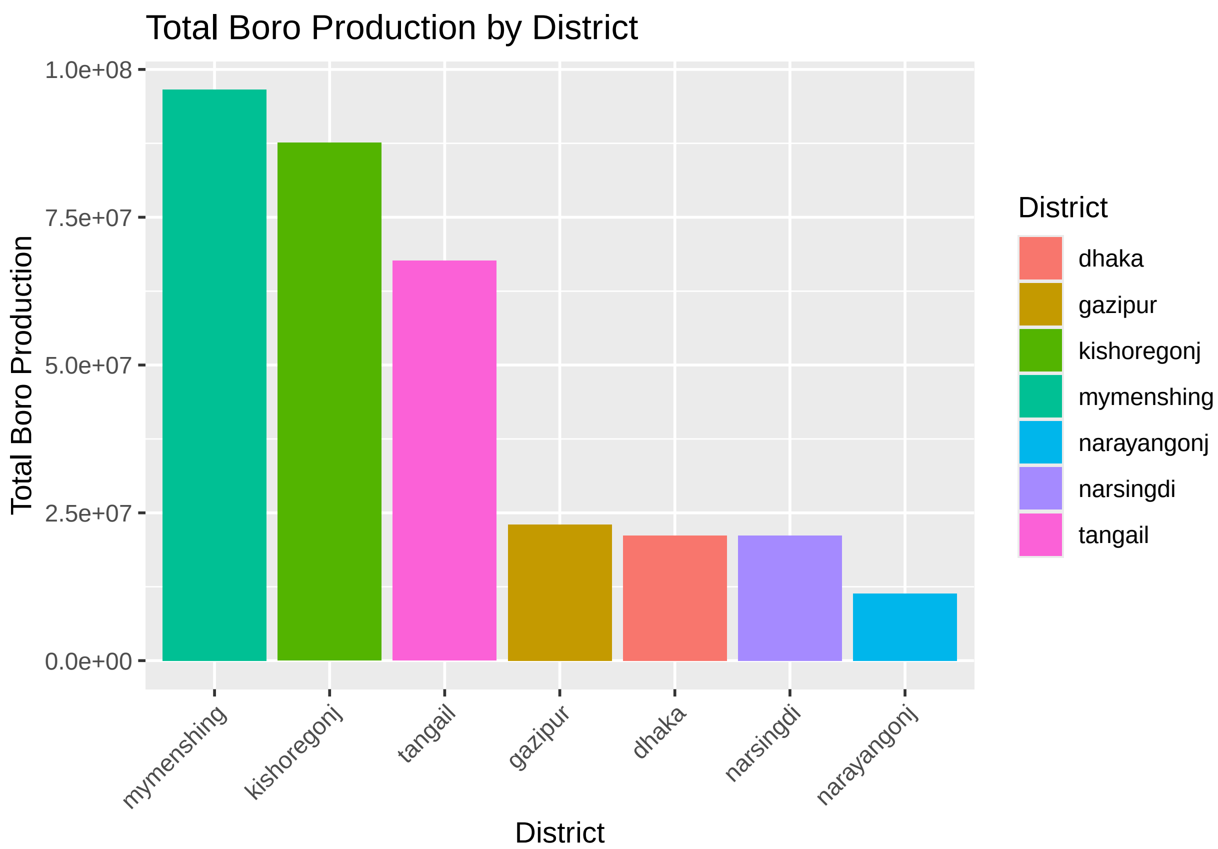
```
##      District total_boro
```

```
## 1 mymenshing    9653940
```

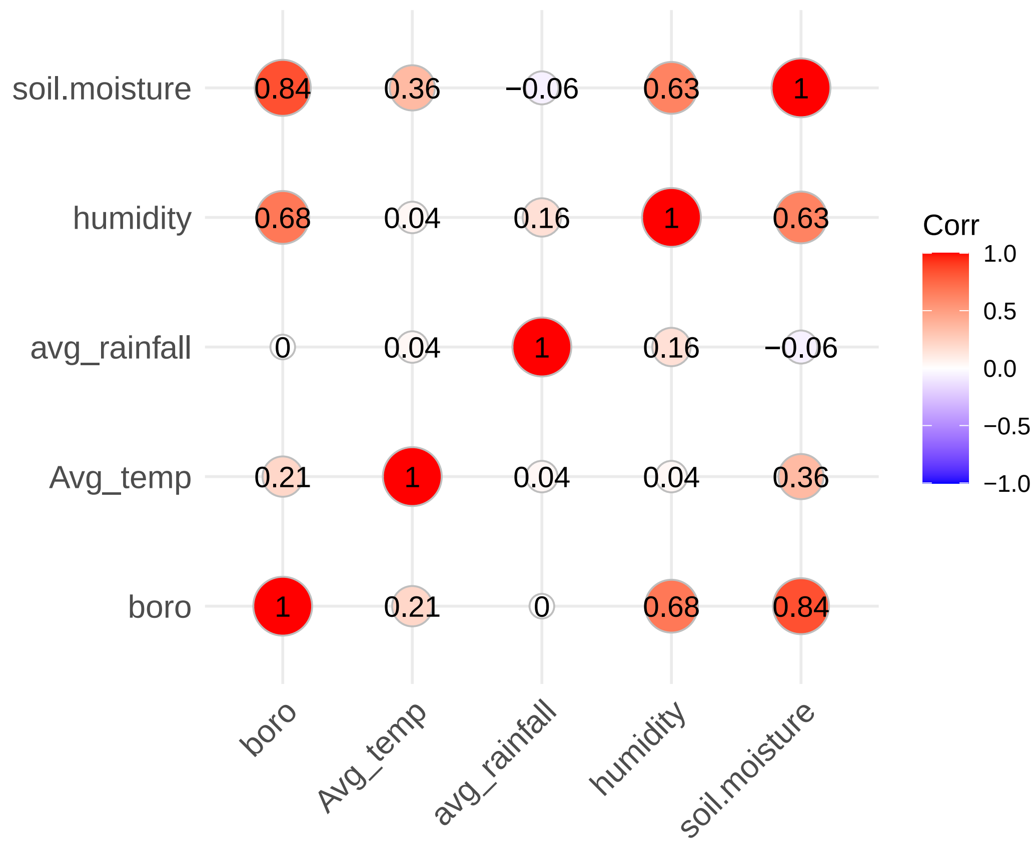
```r
# Lowest Total Boro Production by District
tail(total_boro_by_district, n=1)
```

```
##        District total_boro
## 70 narayangonj    1136944
```

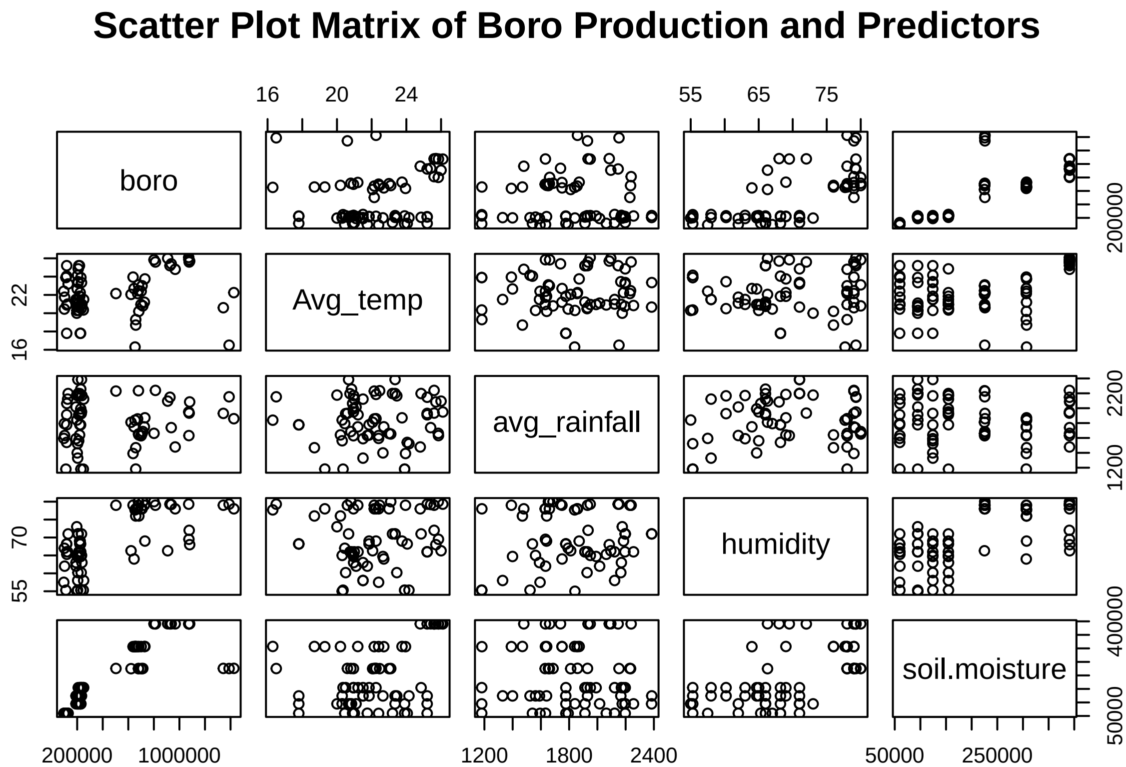
```r
# Data Visualization for Total Aus Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_boro), y = total_boro)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Boro Production by District", x = "District", y = "Total Boro Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
## Data Analysis of Boro Production based on Average Temperature, Average Rainfall, Humidity and Soil Mo

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(boro, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```

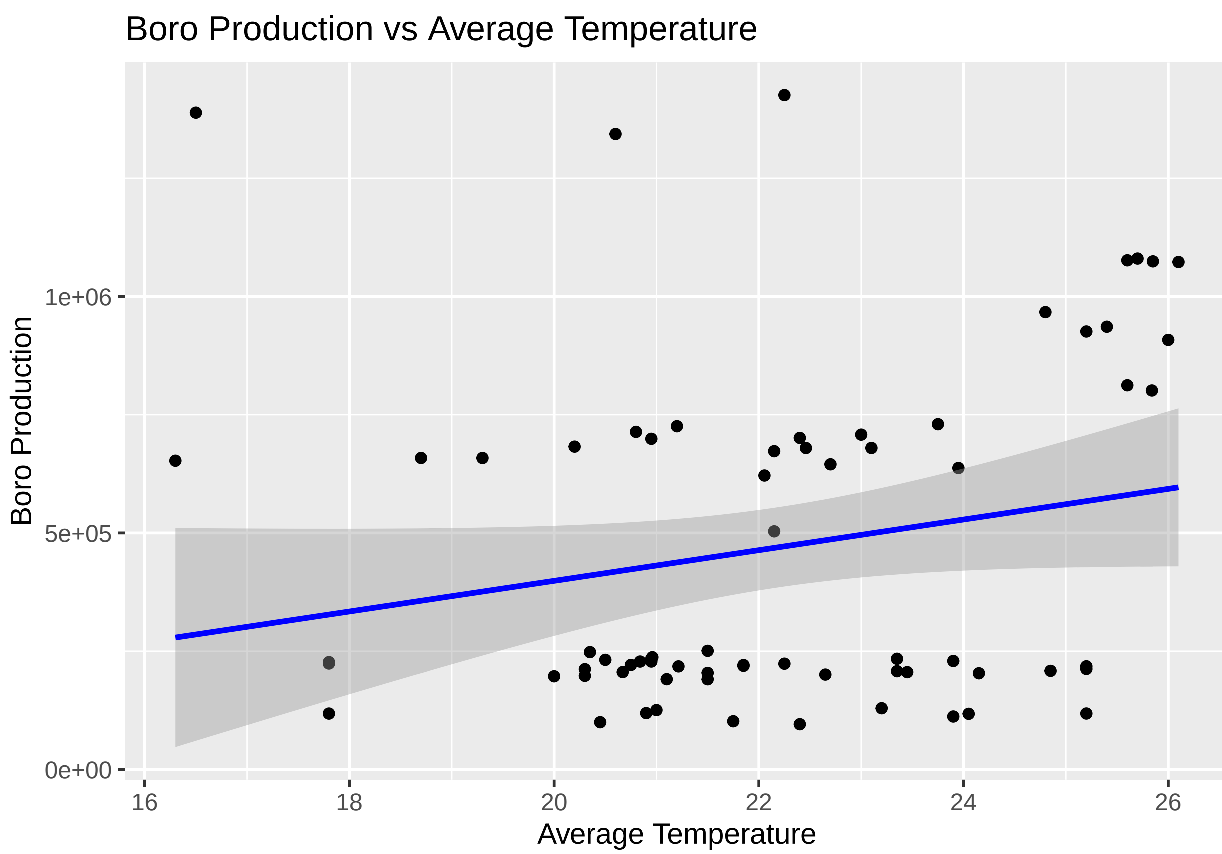



```r
# Scatter plot matrix to visualize relationships
pairs(~ boro + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Boro Production and Predictors")
```



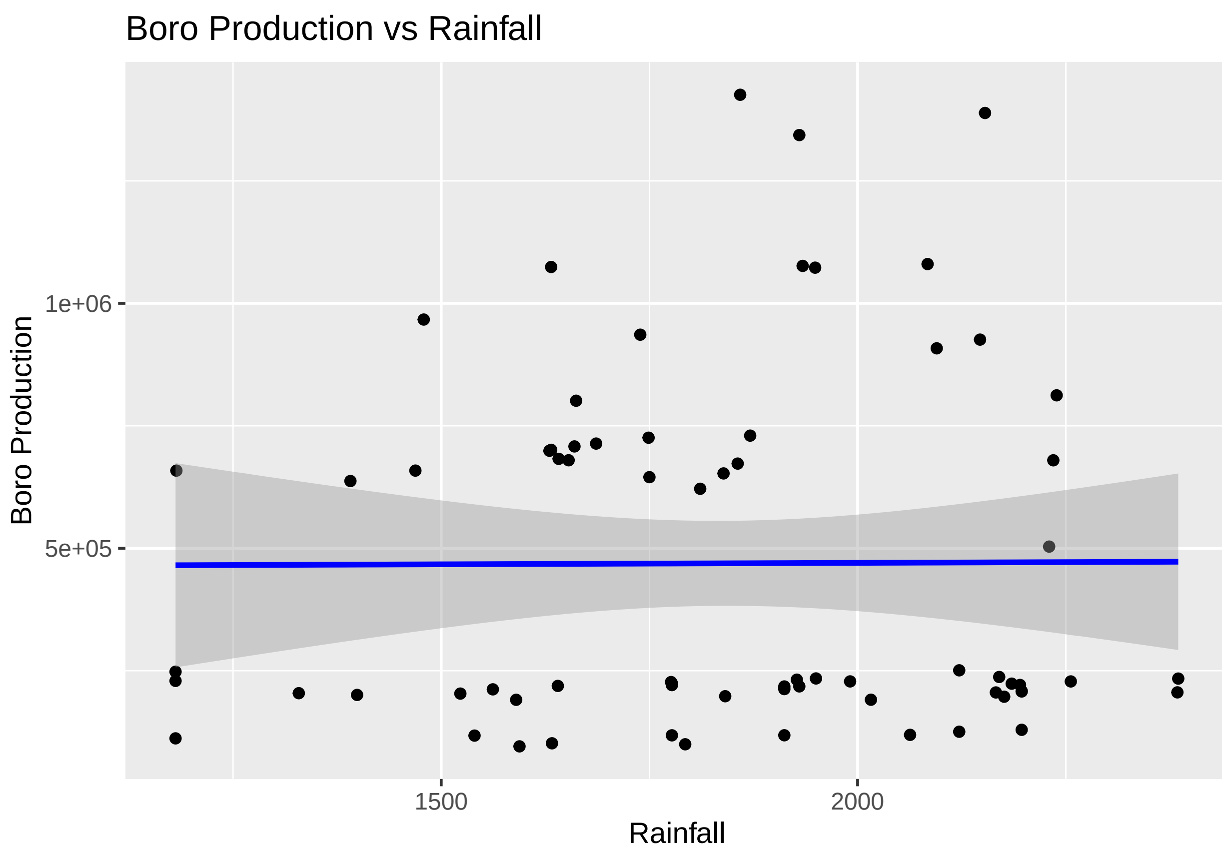

```r
## Indiividual Scatter Plots
#  Boro Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = boro)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Boro Production vs Average Temperature", x = "Average Temperature", y = "Boro Productio
```

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



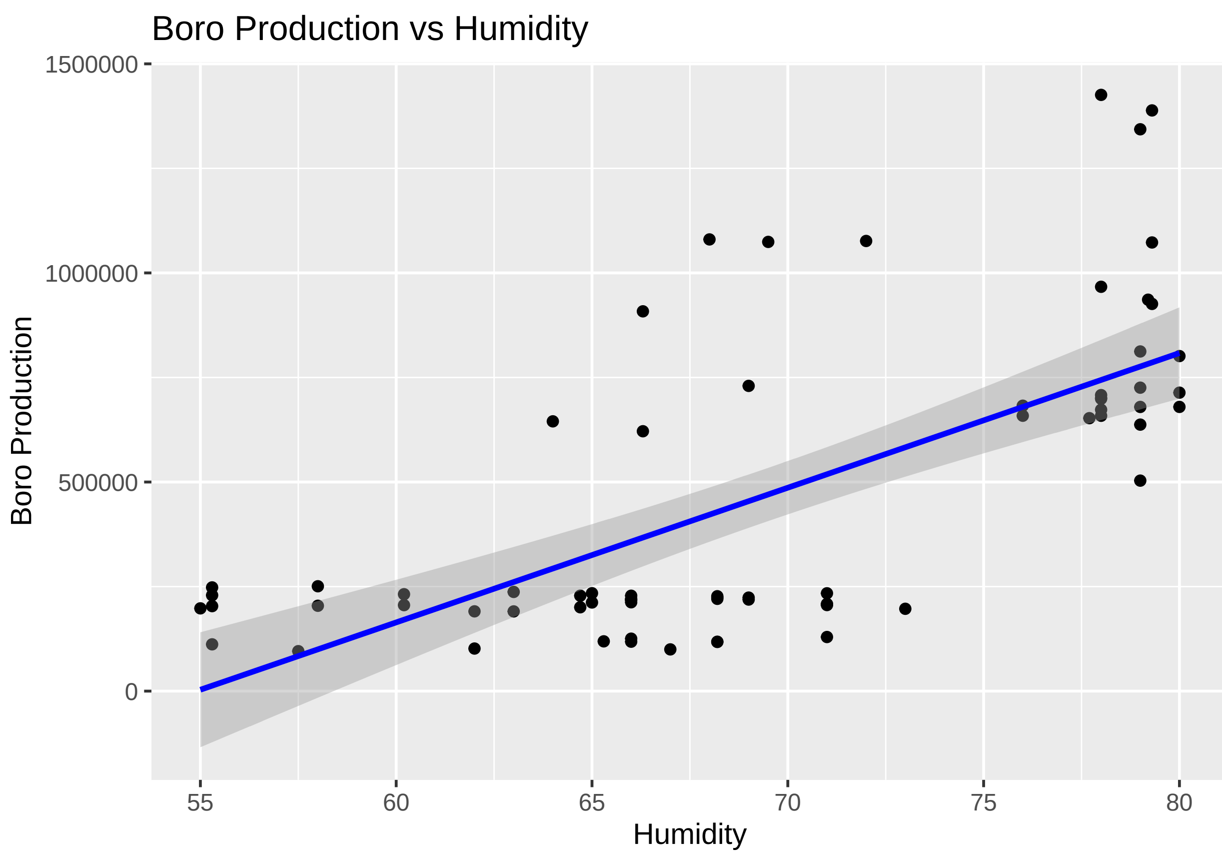

```
# Boro Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = boro)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Boro Production vs Rainfall", x = "Rainfall", y = "Boro Production")
```

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



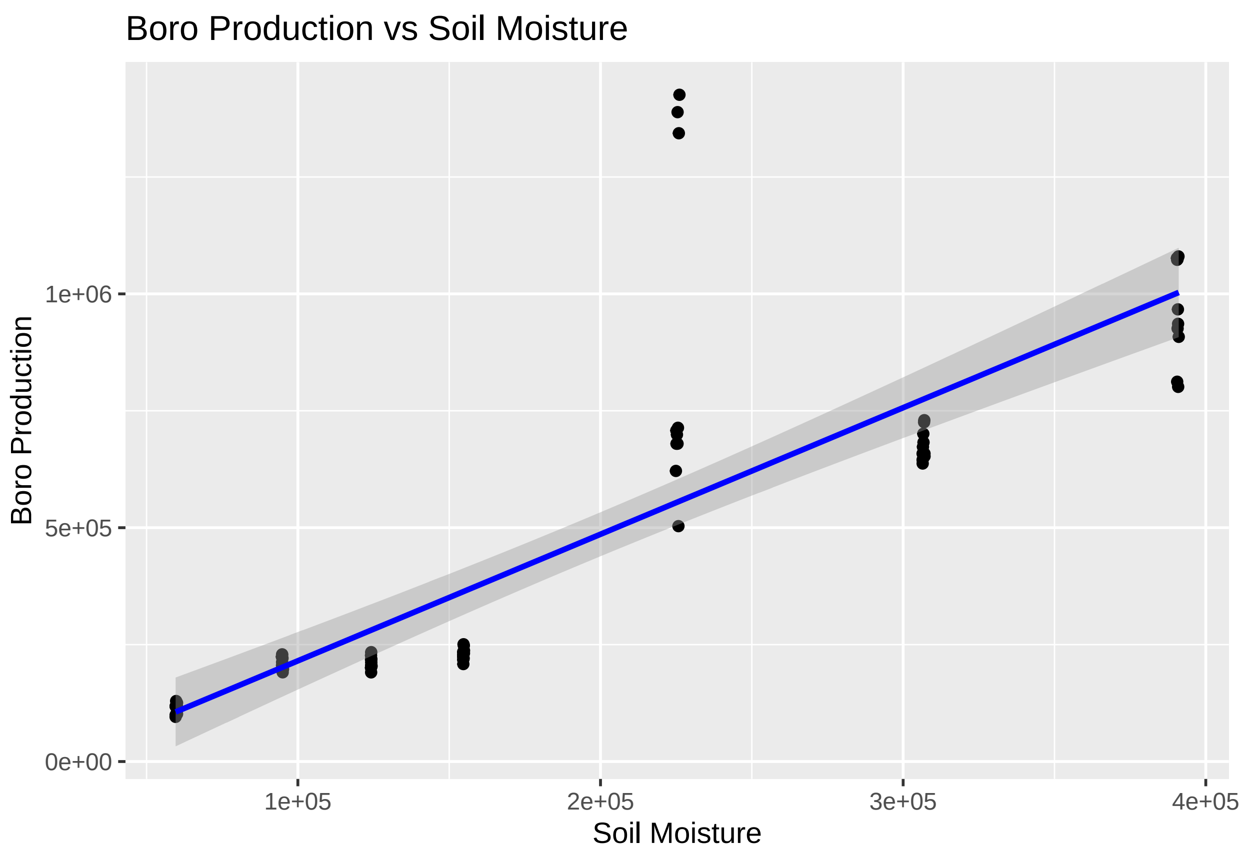

```r
# Boro Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = boro)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Boro Production vs Humidity", x = "Humidity", y = "Boro Production")
```

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




```r
#  Aus Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = boro)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Boro Production vs Soil Moisture", x = "Soil Moisture", y = "Boro Production")
```

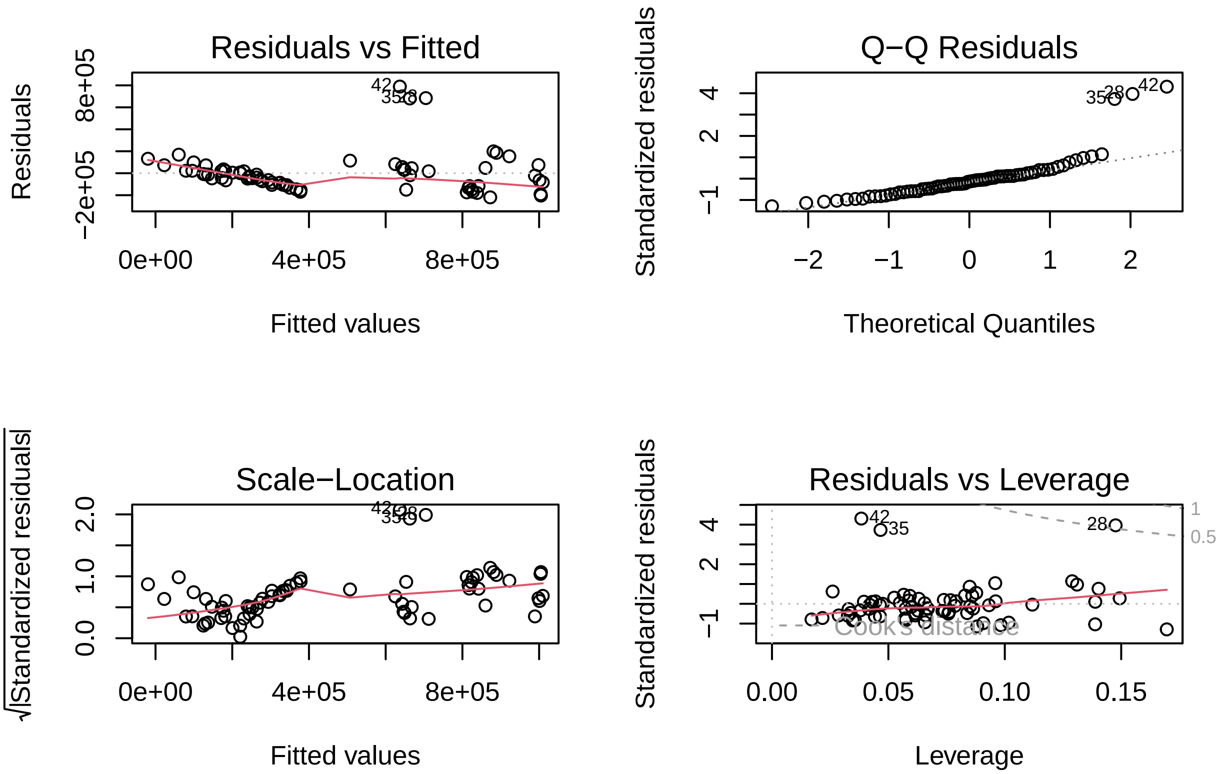
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(boro ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = boro ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -219837 -106185  -24759   34688  788853
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -5.714e+05  3.702e+05  -1.544   0.1275
## Avg_temp      -8.791e+03  1.078e+04  -0.815   0.4180
## avg_rainfall   1.493e+01  7.778e+01   0.192   0.8484
## humidity       1.095e+04  4.079e+03   2.685   0.0092 **
## soil.moisture  2.309e+00  2.933e-01   7.871 4.95e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 186800 on 65 degrees of freedom
## Multiple R-squared:  0.7477, Adjusted R-squared:  0.7321
## F-statistic: 48.15 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
## Group by district and calculate the total (sum) of Wheat for each district

wheat_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_wheat = sum(wheat, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(wheat_sum_by_district, by = "District")


# Total Wheat Production by District

total_wheat_by_district <- Agri_Data %>%
  arrange(desc(total_wheat)) %>%
  select(District, total_wheat)


# Highest total Boro Production by District
head(total_wheat_by_district,n=1)
```

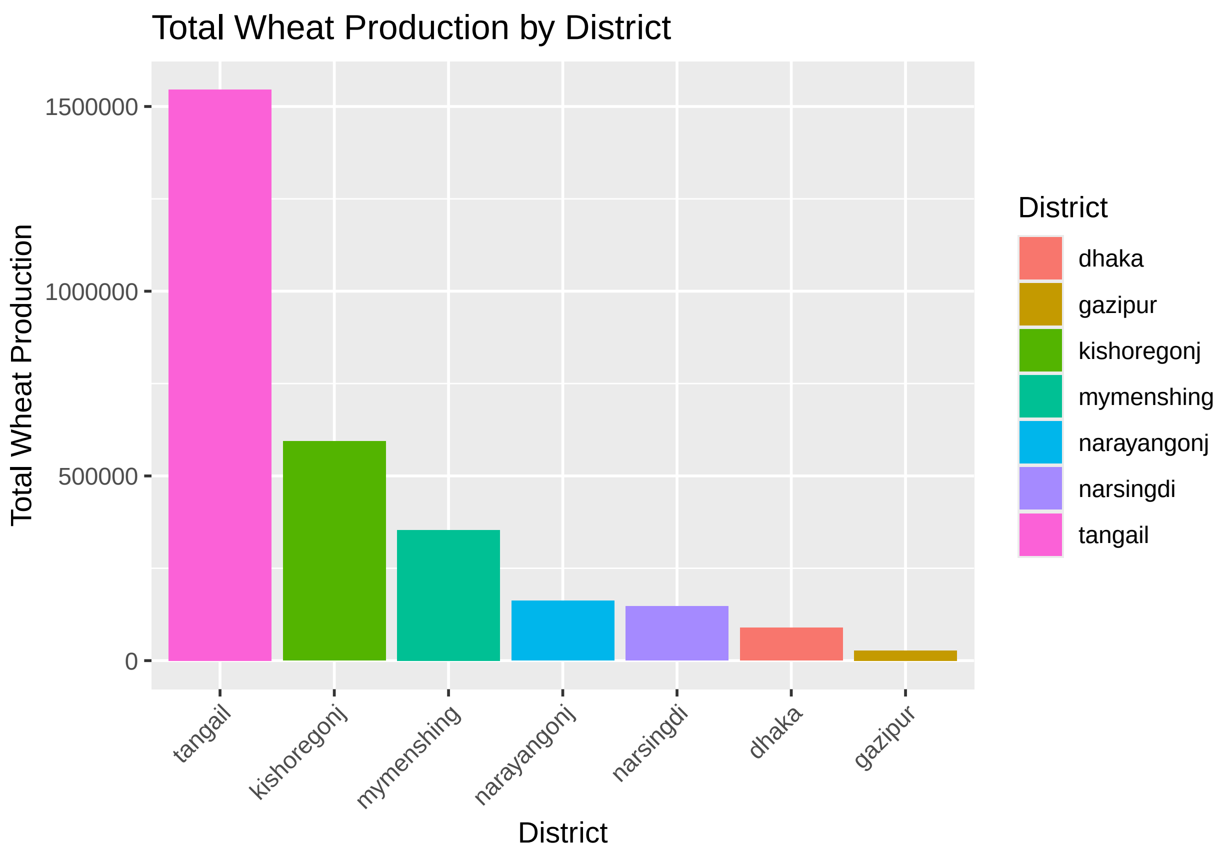
```
##   District total_wheat
```

```
## 1  tangail      154497
```

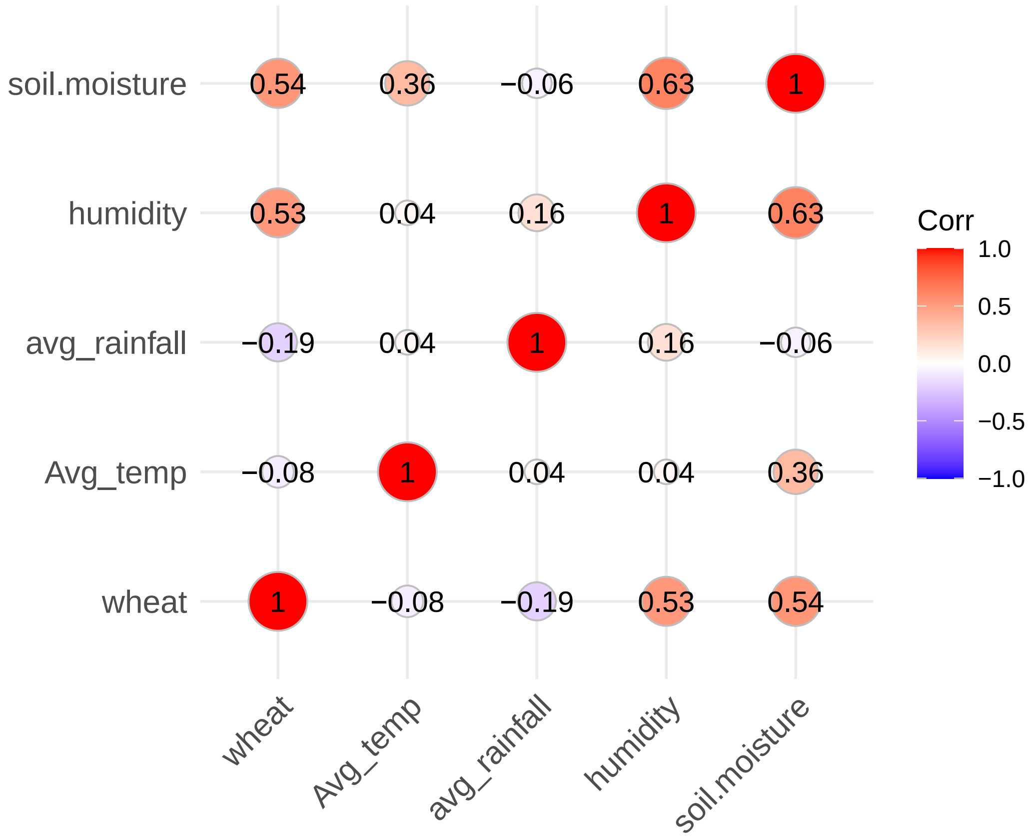
```r
# Lowest Total Boro Production by District
tail(total_wheat_by_district, n=1)
```

```
##    District total_wheat
## 70  gazipur        2737
```

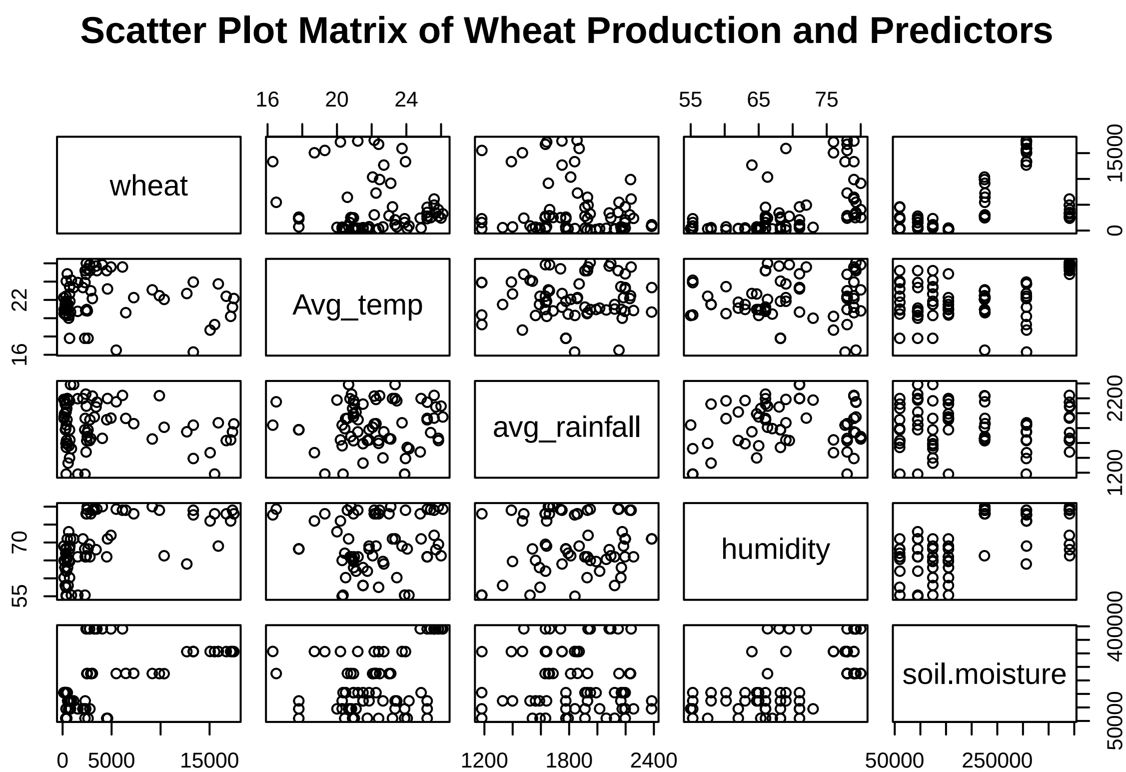
```r
# Data Visualization for Total Wheat Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_wheat), y = total_wheat)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Wheat Production by District", x = "District", y = "Total Wheat Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
## Data Analysis of Wheat Production based on Average Temperature, Average Rainfall, Humidity and Soil M

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(wheat, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```

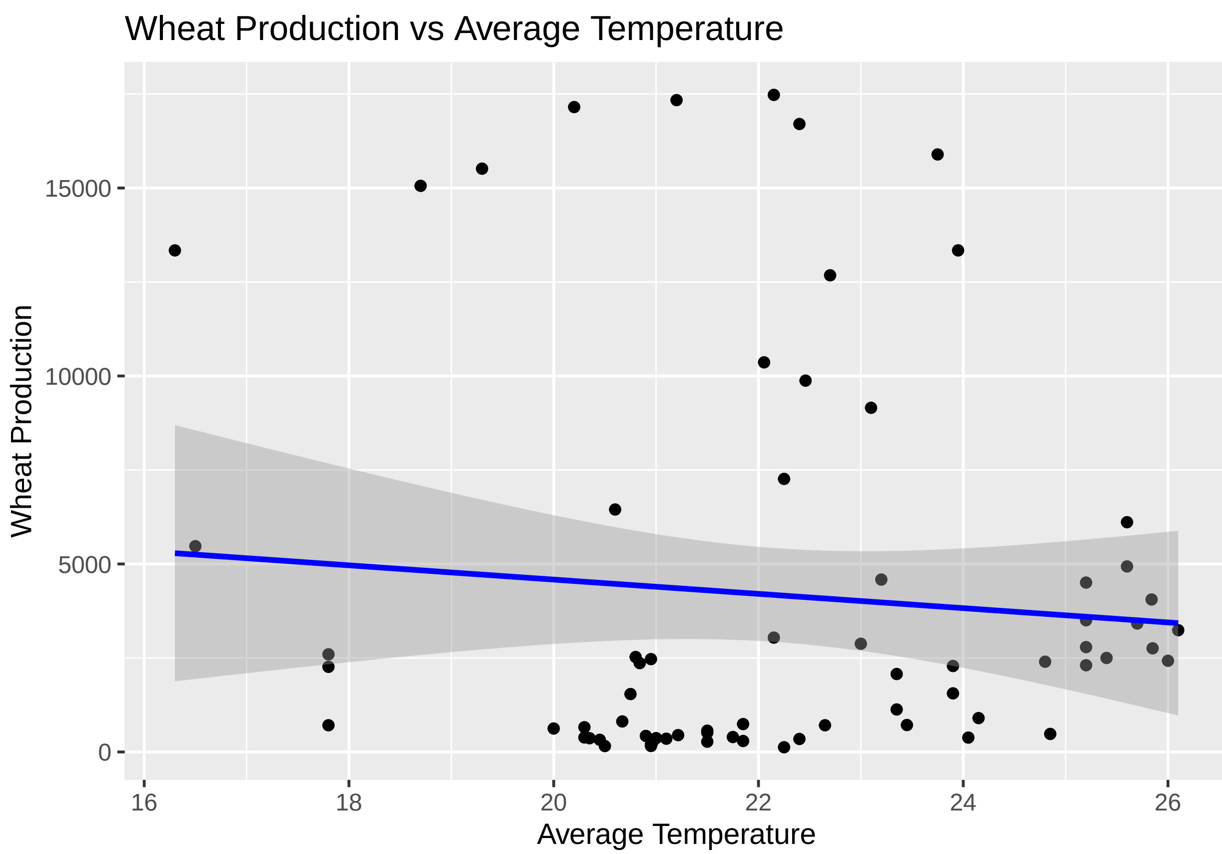



```r
# Scatter plot matrix to visualize relationships
pairs(~ wheat + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Wheat Production and Predictors")
```



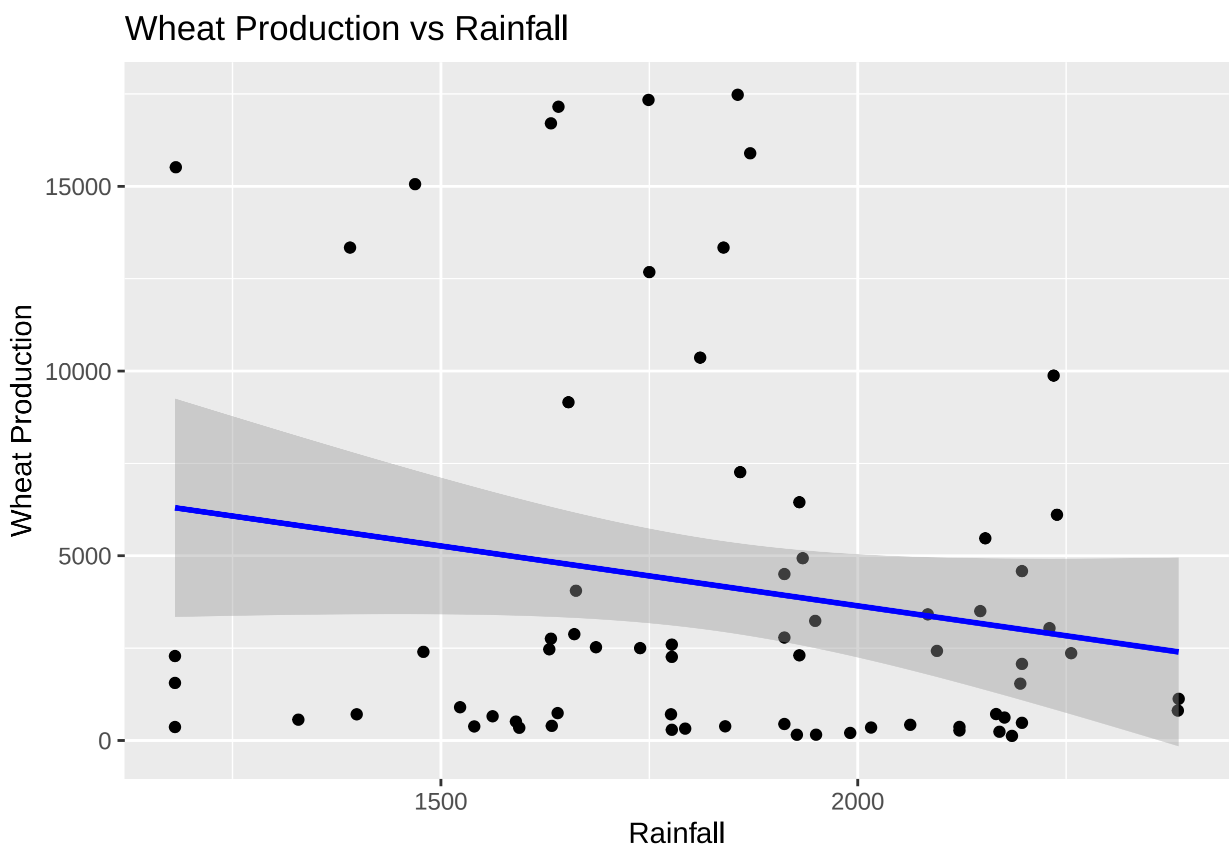

```
## Indiividual Scatter Plots
#  Wheat Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = wheat)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Wheat Production vs Average Temperature", x = "Average Temperature", y = "Wheat Producti
```

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



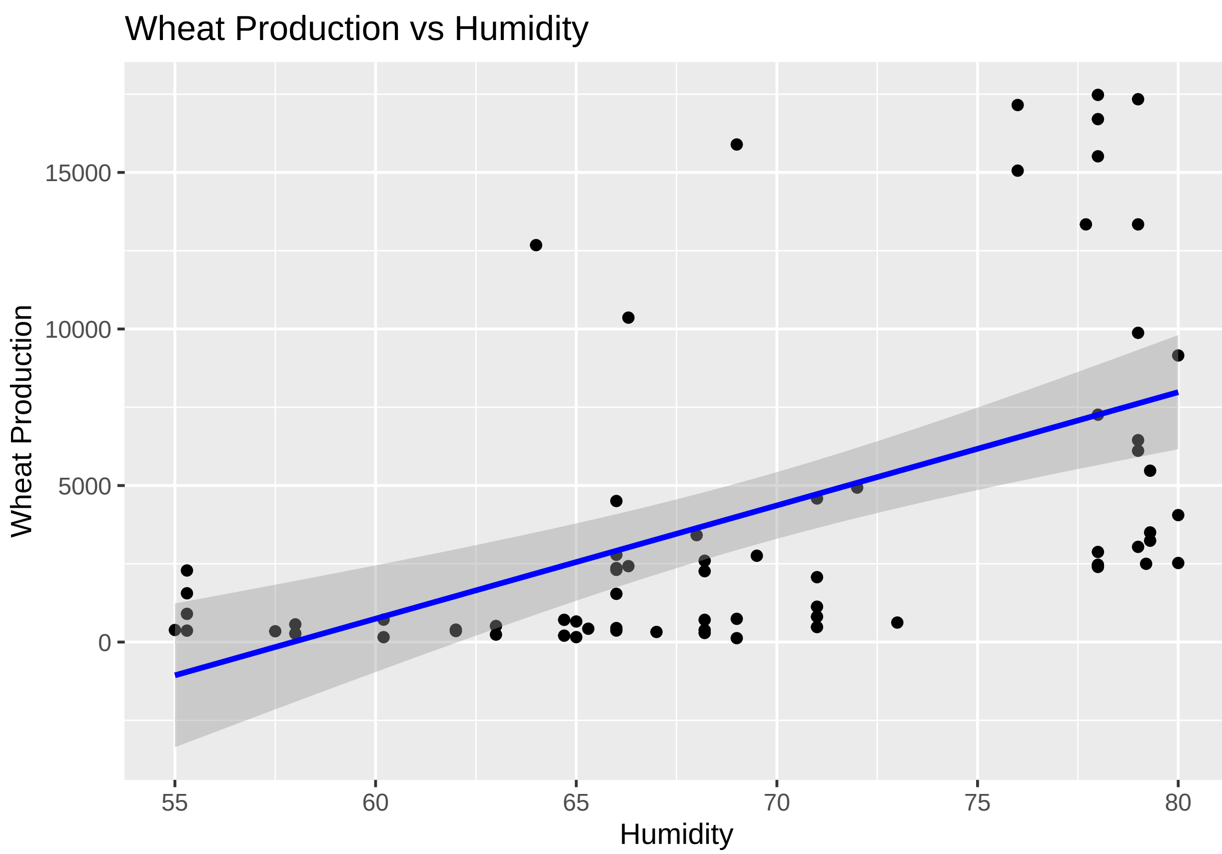

```
# Wheat Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = wheat)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Wheat Production vs Rainfall", x = "Rainfall", y = "Wheat Production")
```

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



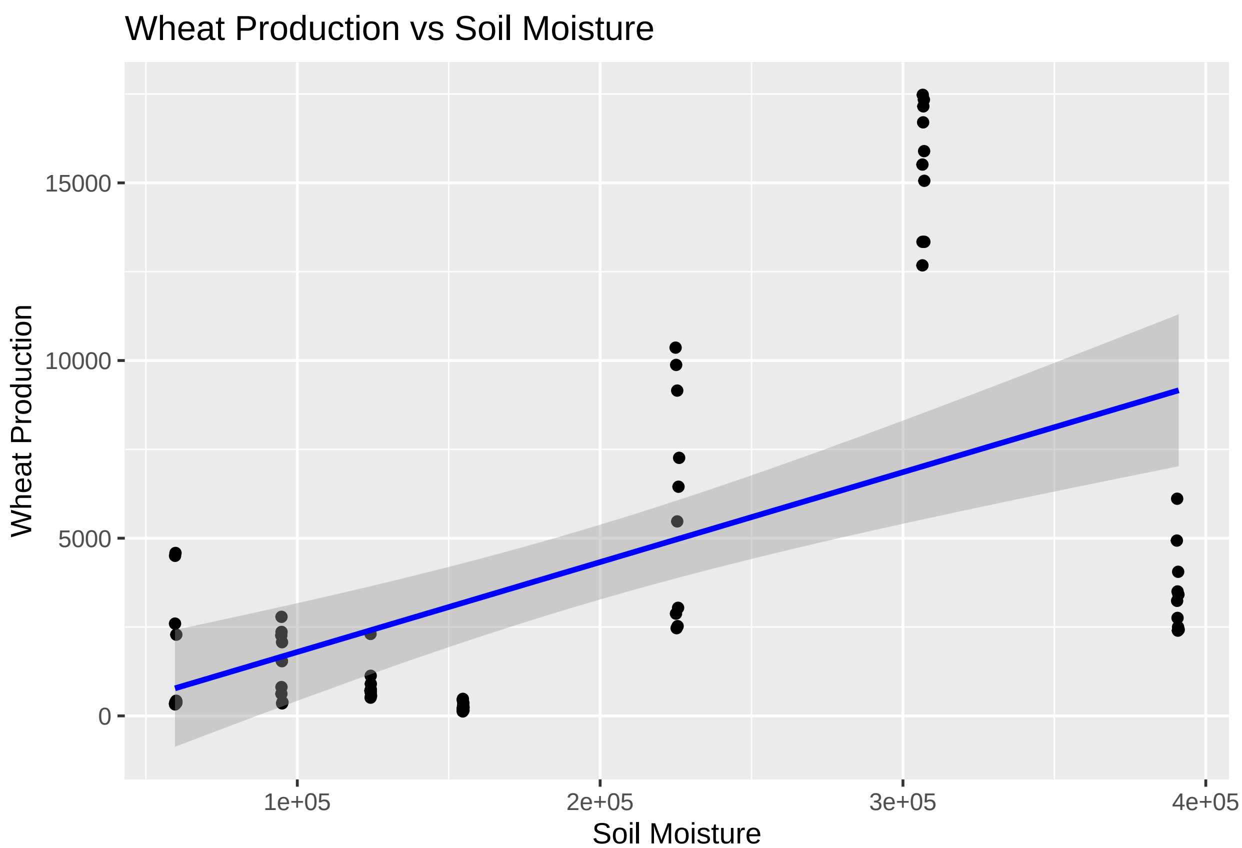

```r
# Wheat Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = wheat)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Wheat Production vs Humidity", x = "Humidity", y = "Wheat Production")
```

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




```
#  Wheat Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = wheat)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Wheat Production vs Soil Moisture", x = "Soil Moisture", y = "Wheat Production")
```

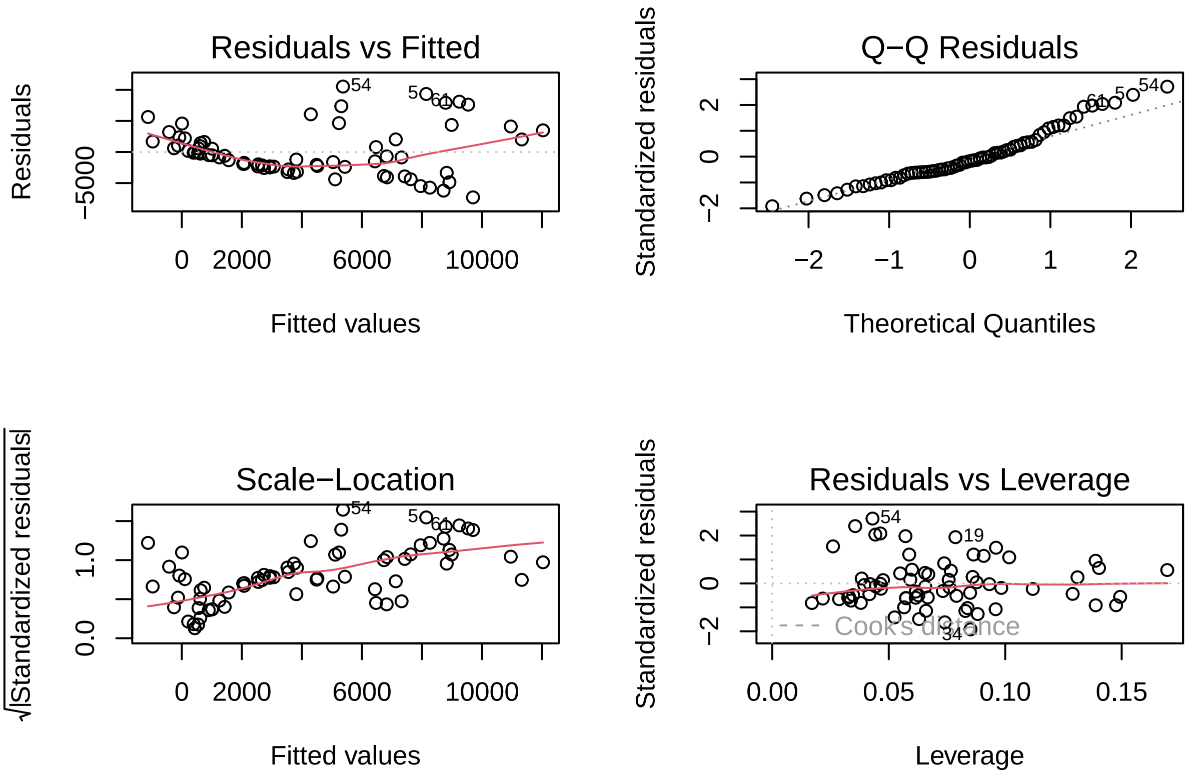
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(wheat ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = wheat ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7296.8 -2381.2  -663.2  1936.8 10526.6
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.345e+03  7.878e+03   0.552  0.58311
## Avg_temp      -5.424e+02  2.295e+02  -2.363  0.02111 *
## avg_rainfall  -3.478e+00  1.655e+00  -2.101  0.03950 *
## humidity       2.069e+02  8.680e+01   2.384  0.02006 *
## soil.moisture  1.998e-02  6.242e-03   3.200  0.00212 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3974 on 65 degrees of freedom
## Multiple R-squared:  0.452,  Adjusted R-squared:  0.4183
## F-statistic:  13.4 on 4 and 65 DF,  p-value: 5.085e-08
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
## Group by district and calculate the total (sum) of Potato for each district

potato_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_potato = sum(potato, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(potato_sum_by_district, by = "District")


# Total Wheat Production by District

total_potato_by_district <- Agri_Data %>%
  arrange(desc(total_potato)) %>%
  select(District, total_potato)


# Highest total Potato Production by District
head(total_potato_by_district,n=1)
```

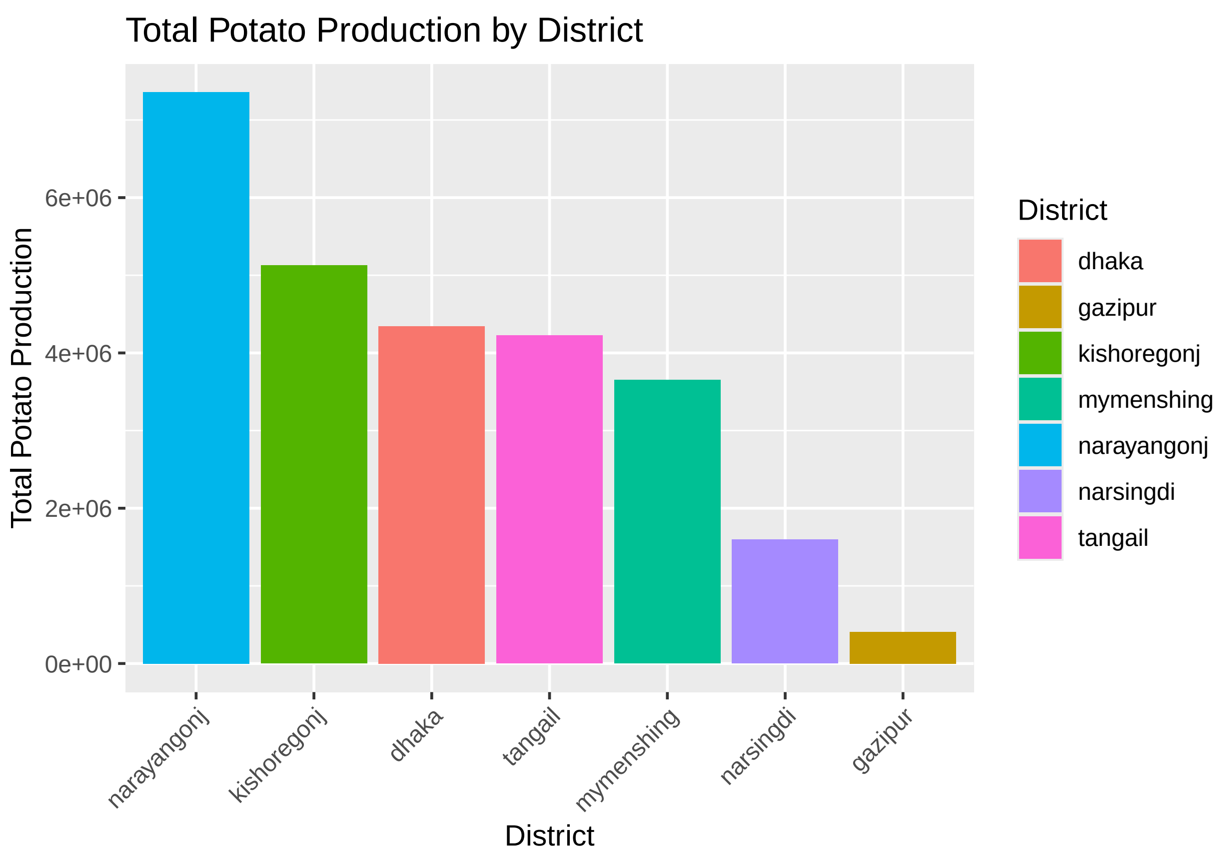
```
##       District total_potato
```

```
## 1 narayangonj       735442
```

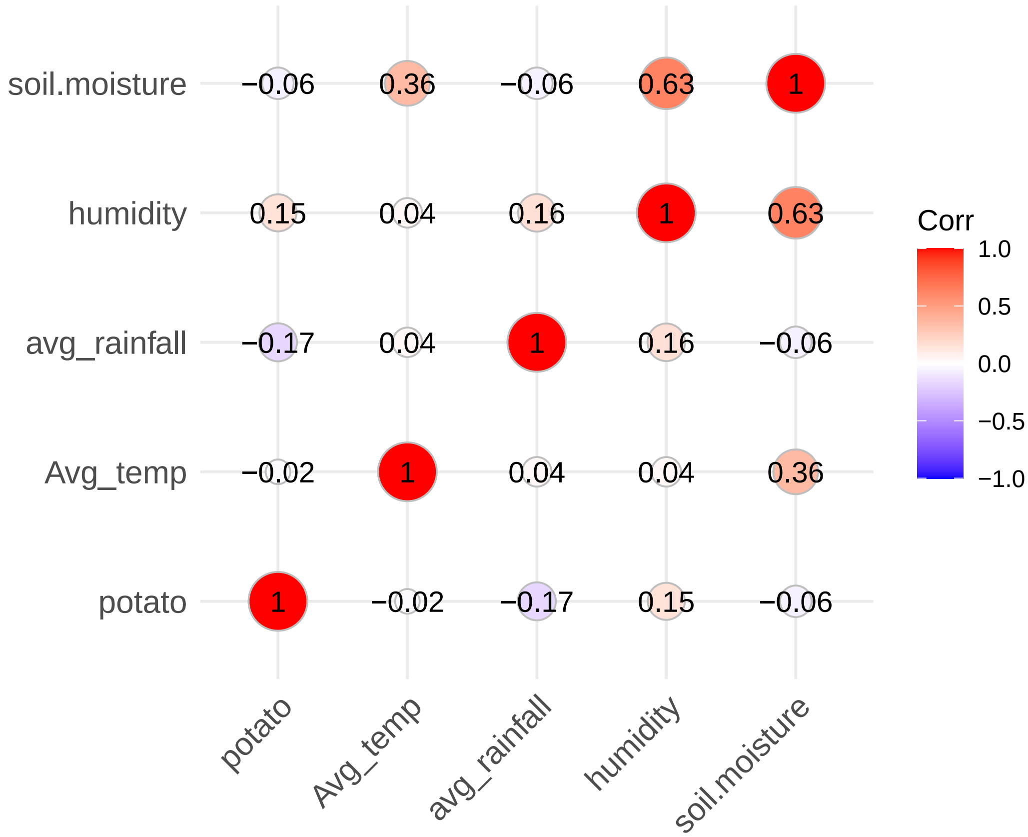
```r
# Lowest Total Potato Production by District
tail(total_potato_by_district, n=1)
```

```
##    District total_potato
## 70  gazipur        40640
```

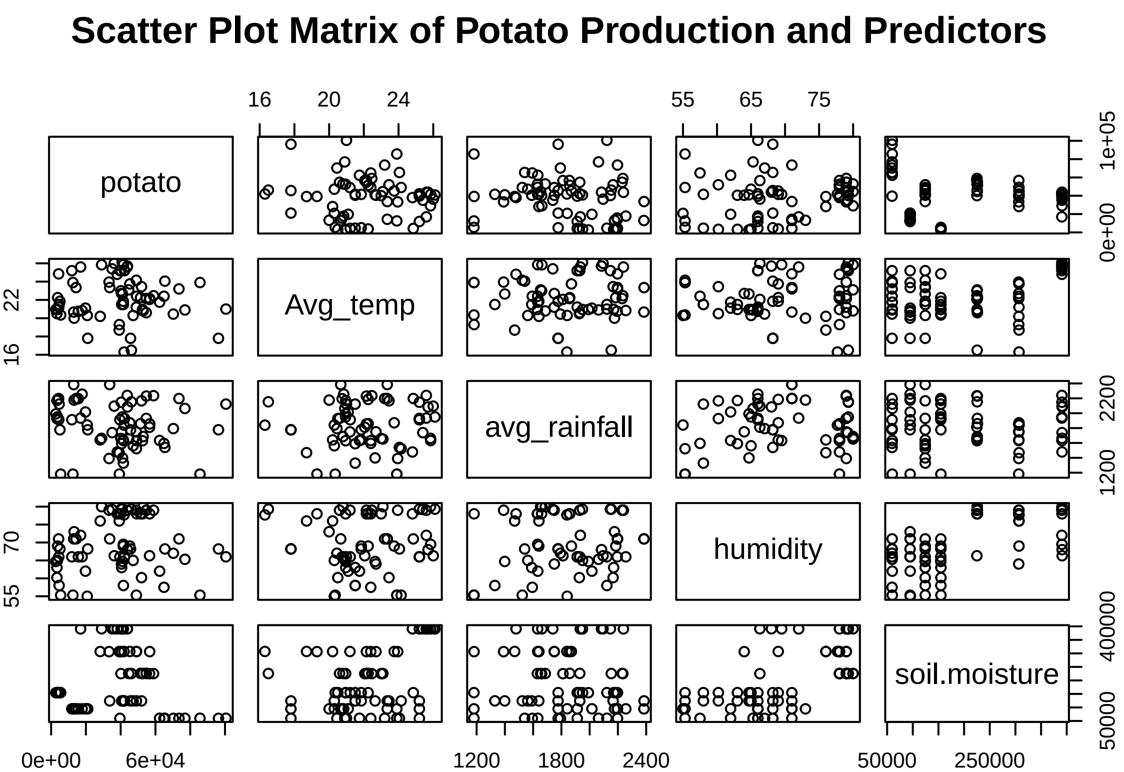
```r
# Data Visualization for Total Wheat Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_potato), y = total_potato)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Potato Production by District", x = "District", y = "Total Potato Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
## Data Analysis of Potato Production based on Average Temperature, Average Rainfall, Humidity and Soil

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(potato, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```

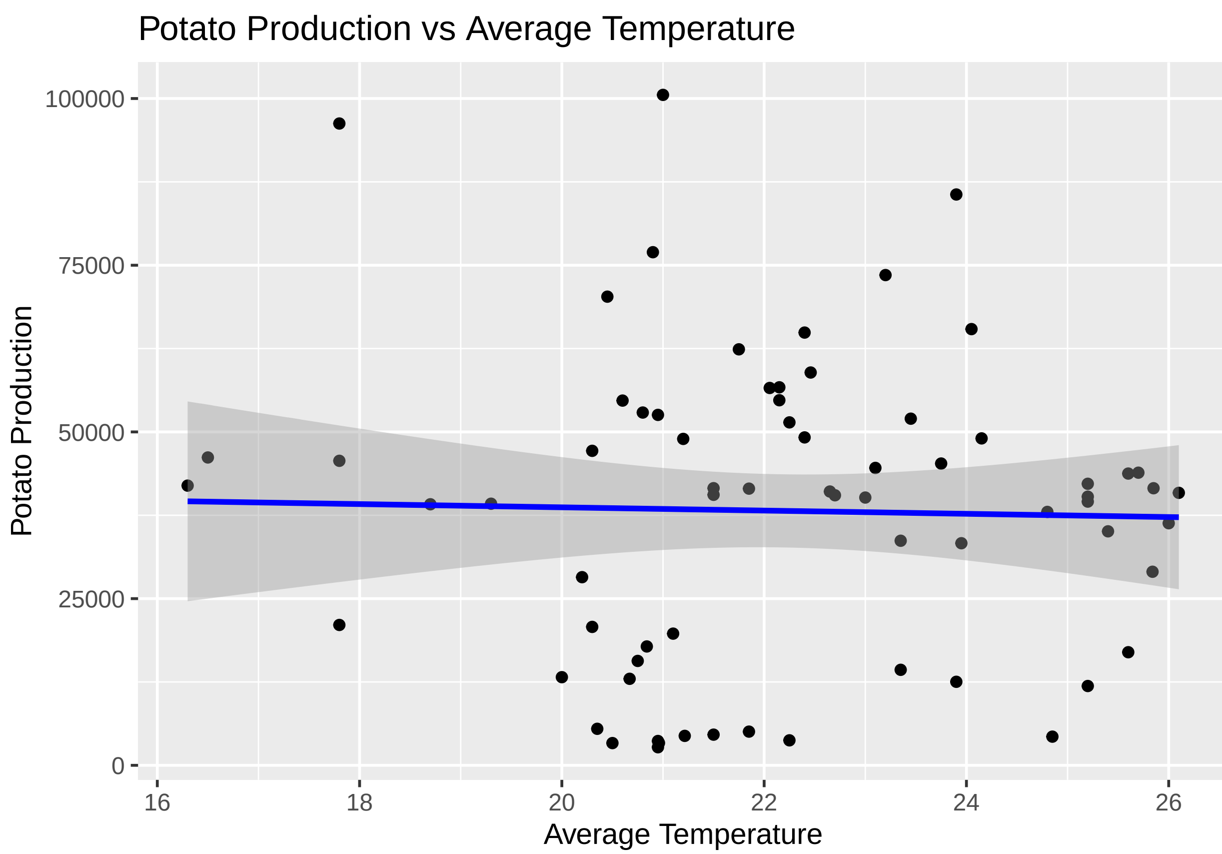



```r
# Scatter plot matrix to visualize relationships
pairs(~ potato + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Potato Production and Predictors")
```



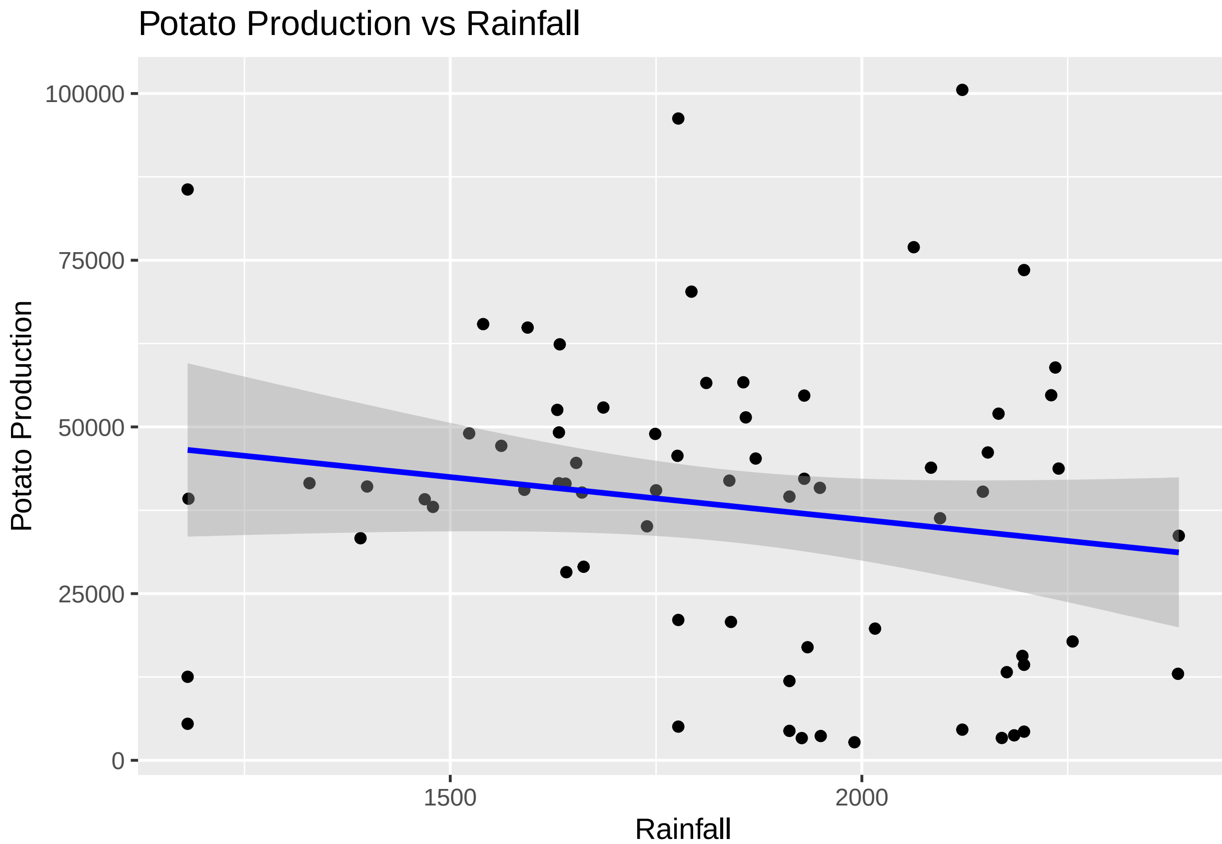

```r
## Indiividual Scatter Plots
#  Potato Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = potato)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Potato Production vs Average Temperature", x = "Average Temperature", y = "Potato Produ
```

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



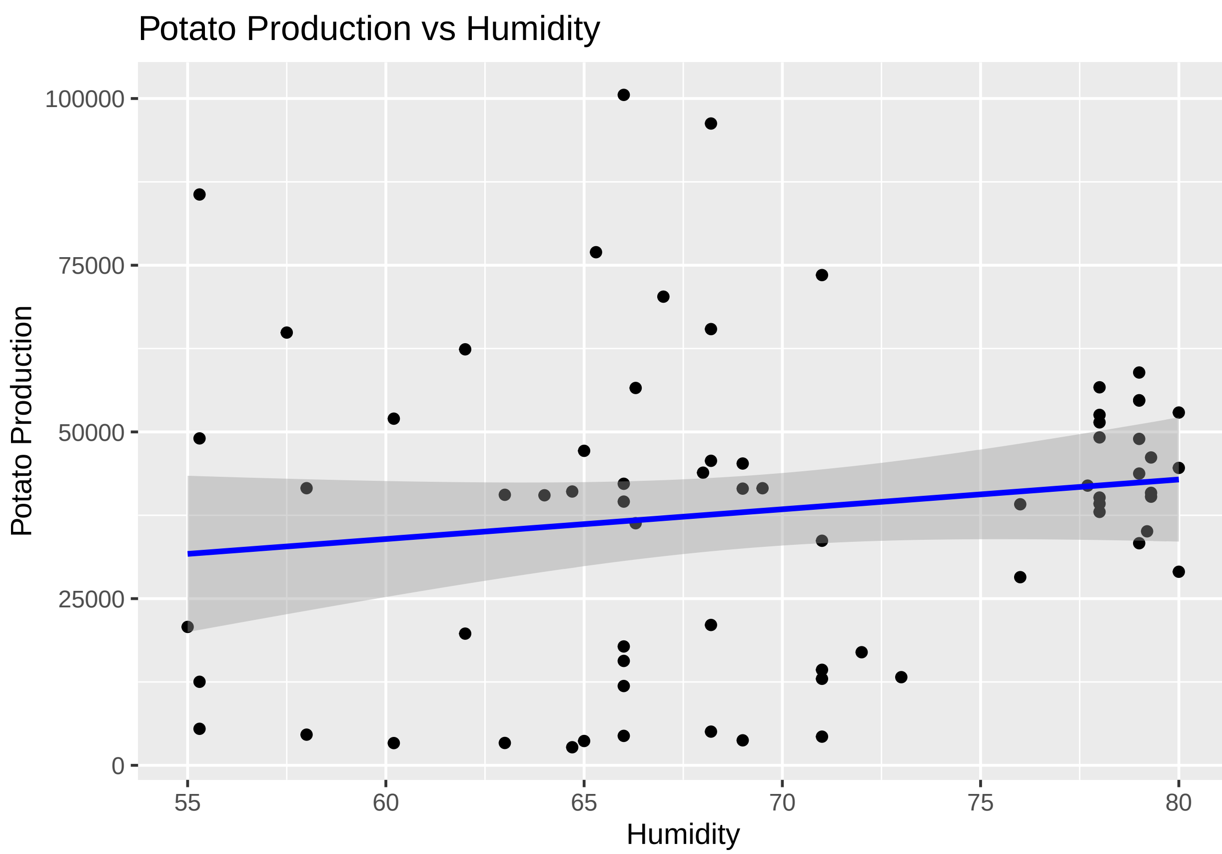

```
# Potato Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = potato)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Potato Production vs Rainfall", x = "Rainfall", y = "Potato Production")
```

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



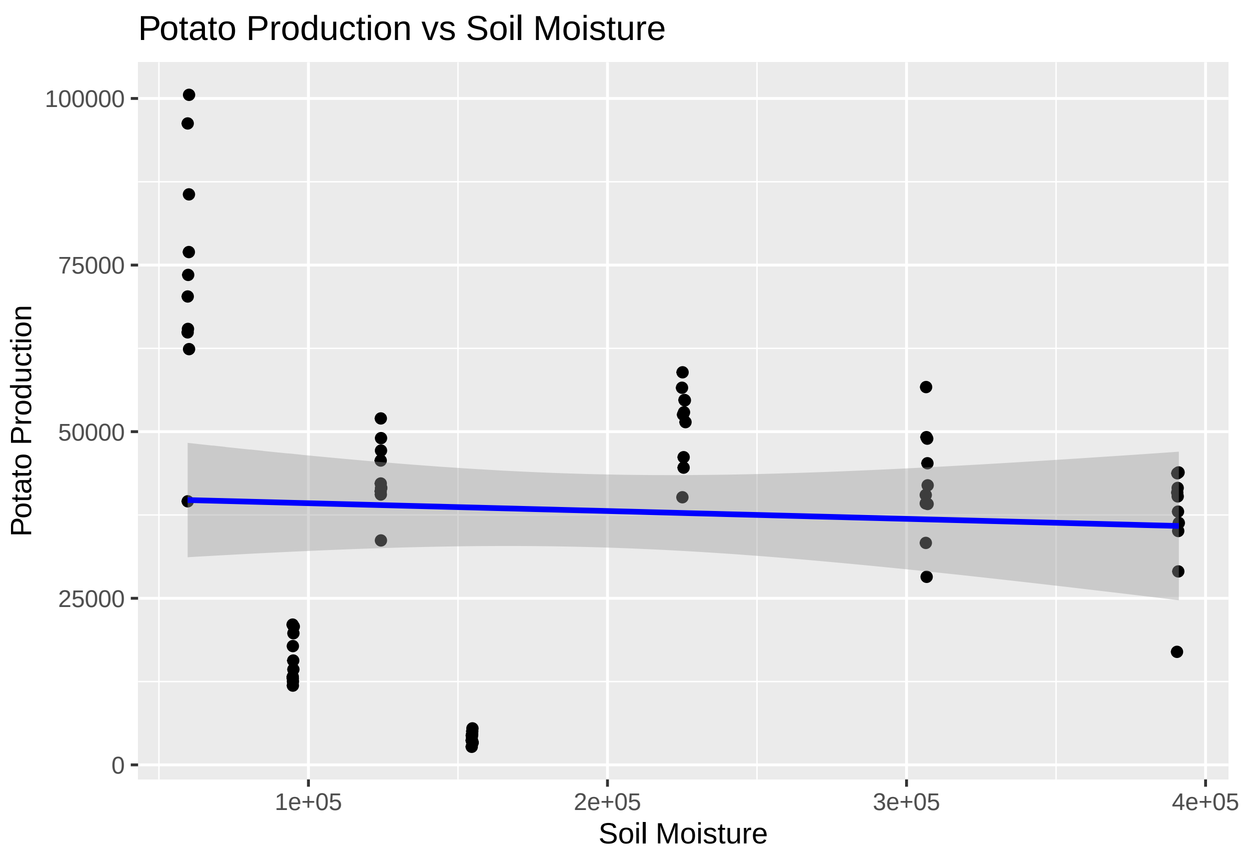

```r
# Potato Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = potato)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Potato Production vs Humidity", x = "Humidity", y = "Potato Production")
```

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




```
#  Potato Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = potato)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Potato Production vs Soil Moisture", x = "Soil Moisture", y = "Potato Production")
```

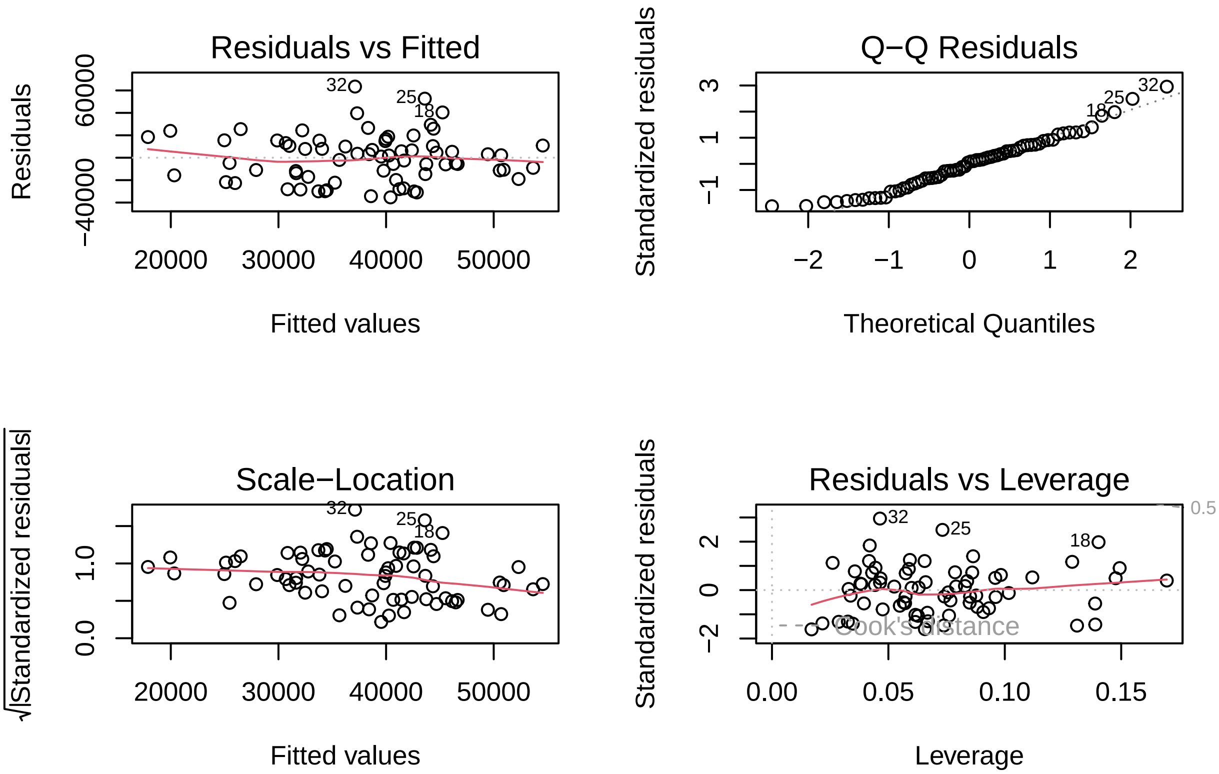
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(potato ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = potato ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35360 -15445   1492  14381  63433
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.039e+04  4.359e+04  -0.468   0.6414
## Avg_temp       9.999e+02  1.270e+03   0.787   0.4339
## avg_rainfall  -1.991e+01  9.159e+00  -2.174   0.0334 *
## humidity       1.263e+03  4.803e+02   2.629   0.0107 *
## soil.moisture -7.600e-02  3.454e-02  -2.200   0.0313 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 21990 on 65 degrees of freedom
## Multiple R-squared:  0.1262, Adjusted R-squared:  0.07238
## F-statistic: 2.346 on 4 and 65 DF,  p-value: 0.0637
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
## Group by district and calculate the total (sum) of Jute for each district

jute_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_jute = sum(jute, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(jute_sum_by_district, by = "District")


# Total Wheat Production by District

total_jute_by_district <- Agri_Data %>%
  arrange(desc(total_jute)) %>%
  select(District, total_jute)


# Highest total Jute Production by District
head(total_jute_by_district,n=1)
```

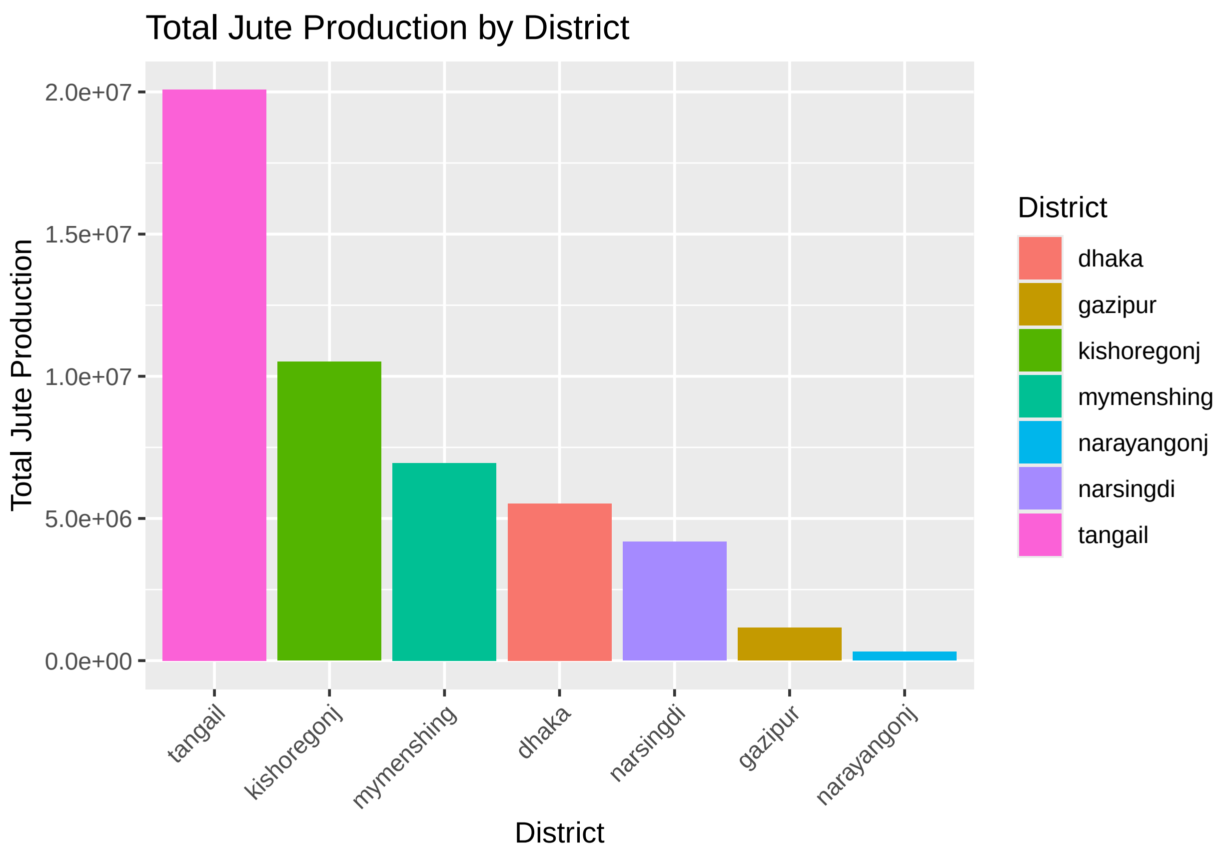
```
##    District total_jute
```

```
## 1  tangail    2007216
```

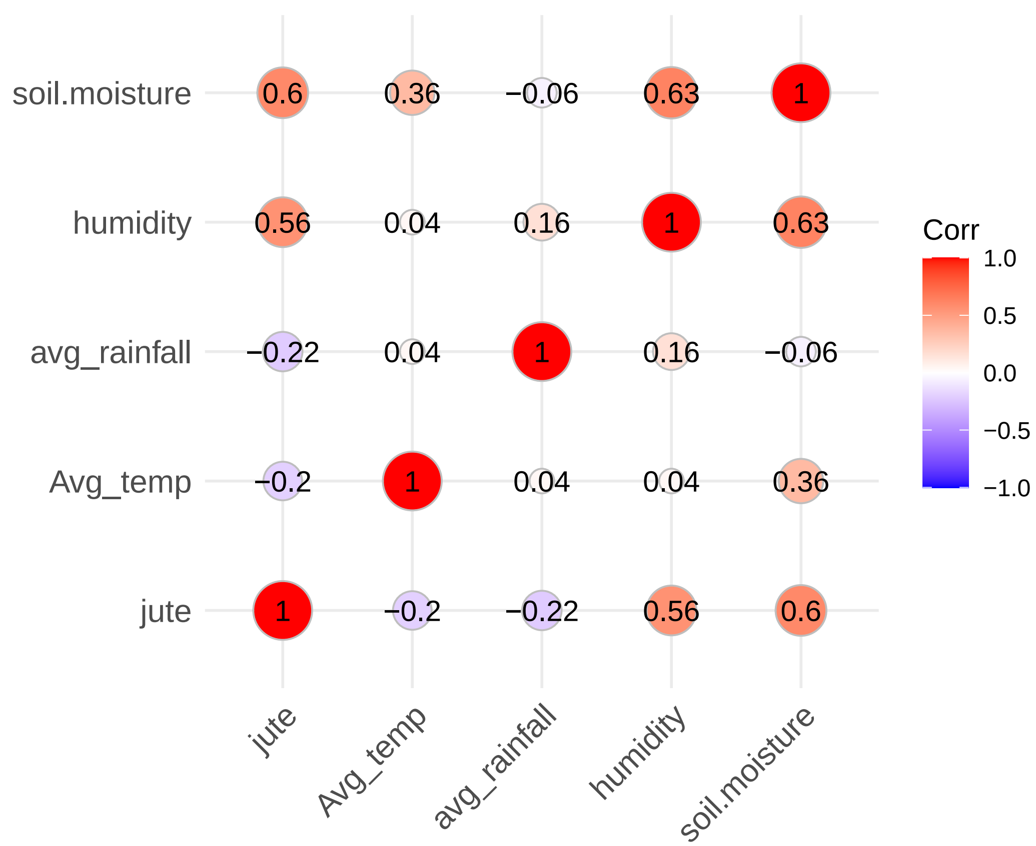
```
# Lowest Total Jute Production by District
tail(total_jute_by_district, n=1)
```

```
##       District total_jute
## 70 narayangonj      32793
```

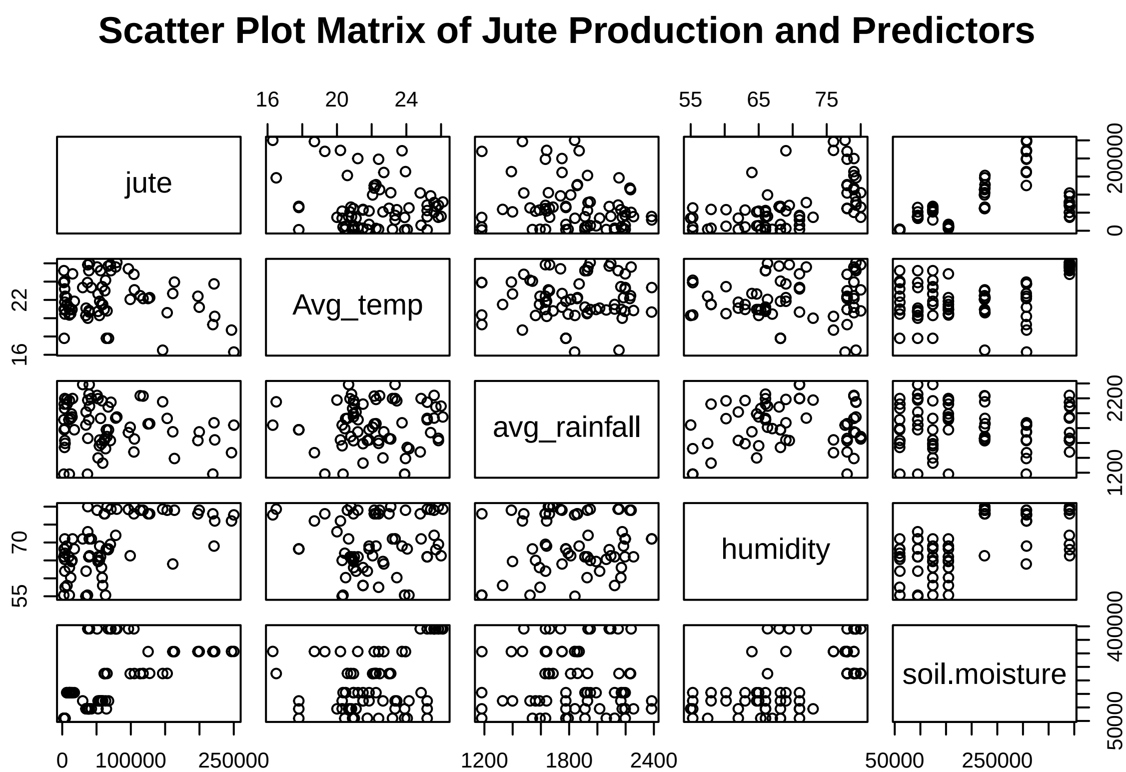
```
# Data Visualization for Total Wheat Production by District
ggplot(Agri_Data, aes(x = reorder(District, -total_jute), y = total_jute)) +
  geom_bar(stat = "identity", aes(fill = District)) +
  labs(title = "Total Jute Production by District", x = "District", y = "Total Jute Production") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```
## Data Analysis of Jute Production based on Average Temperature, Average Rainfall, Humidity and Soil M

# Correlation Matrix

analysis_data <- Agri_Data %>%
  select(jute, Avg_temp, avg_rainfall , humidity, soil.moisture)
correlation_matrix <- cor(analysis_data, use = "complete.obs")
ggcorrplot(correlation_matrix, method = "circle", lab = TRUE)
```

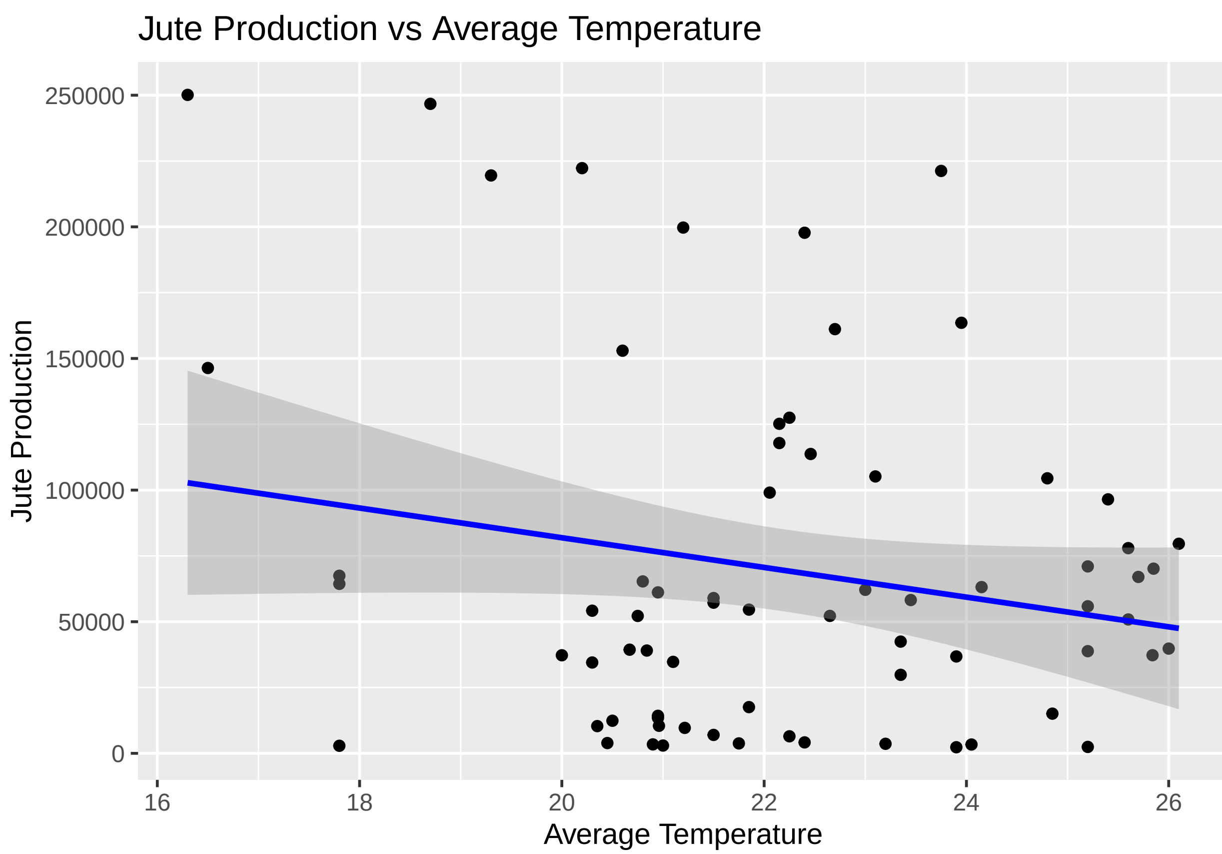



```r
# Scatter plot matrix to visualize relationships
pairs(~ jute + Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data,
      main = "Scatter Plot Matrix of Jute Production and Predictors")
```



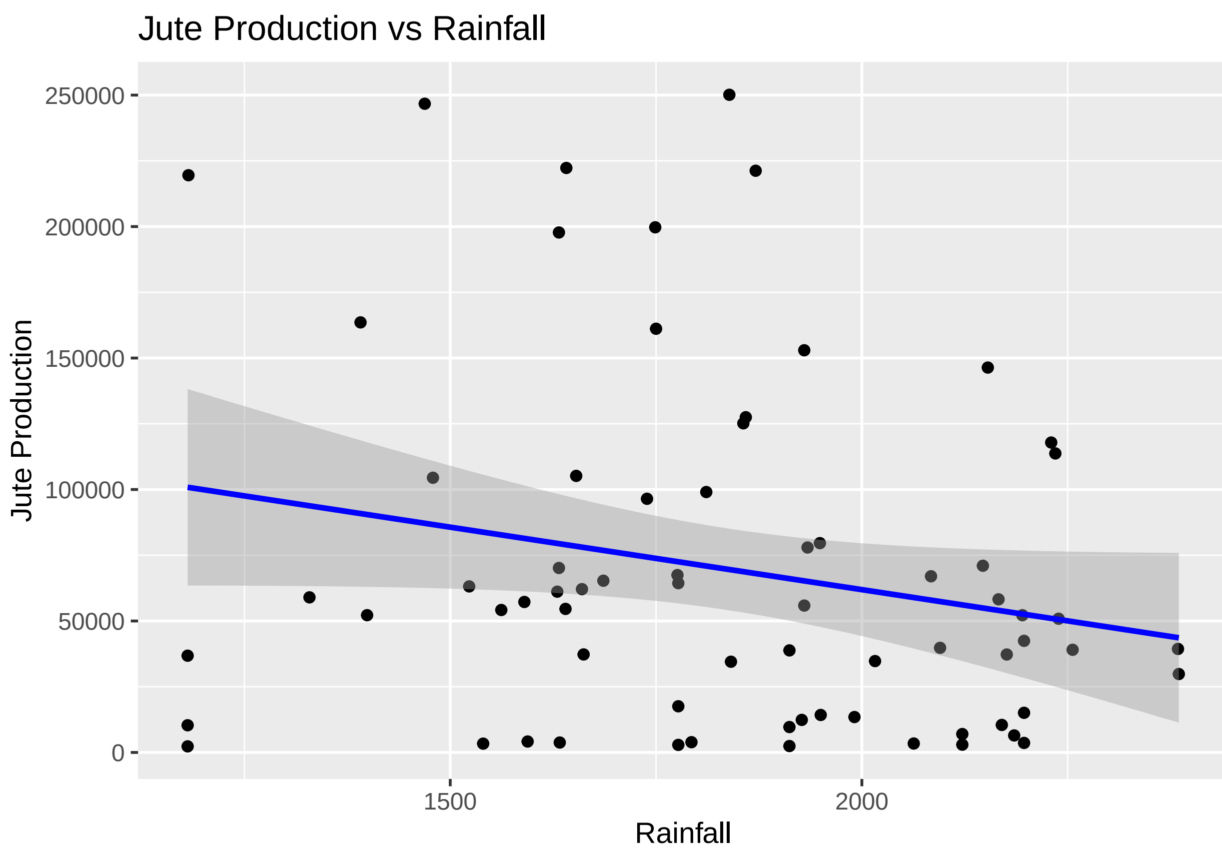

```r
## Indiividual Scatter Plots
#  Jute Production vs. Average Temperature:
ggplot(Agri_Data, aes(x = Avg_temp, y = jute)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Jute Production vs Average Temperature", x = "Average Temperature", y = "Jute Productio
```

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



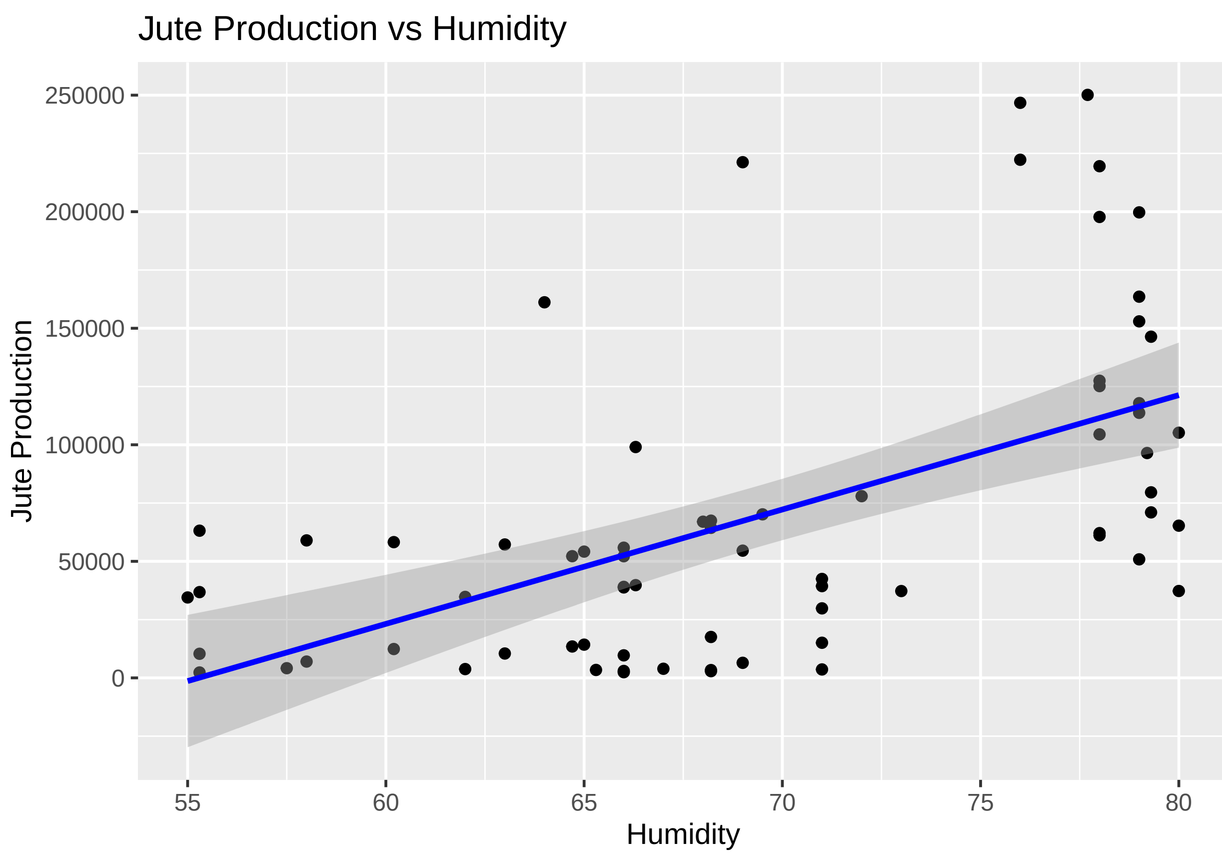

```r
# Jute Production vs. Rainfall:
ggplot(Agri_Data, aes(x = avg_rainfall, y = jute)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Jute Production vs Rainfall", x = "Rainfall", y = "Jute Production")
```

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



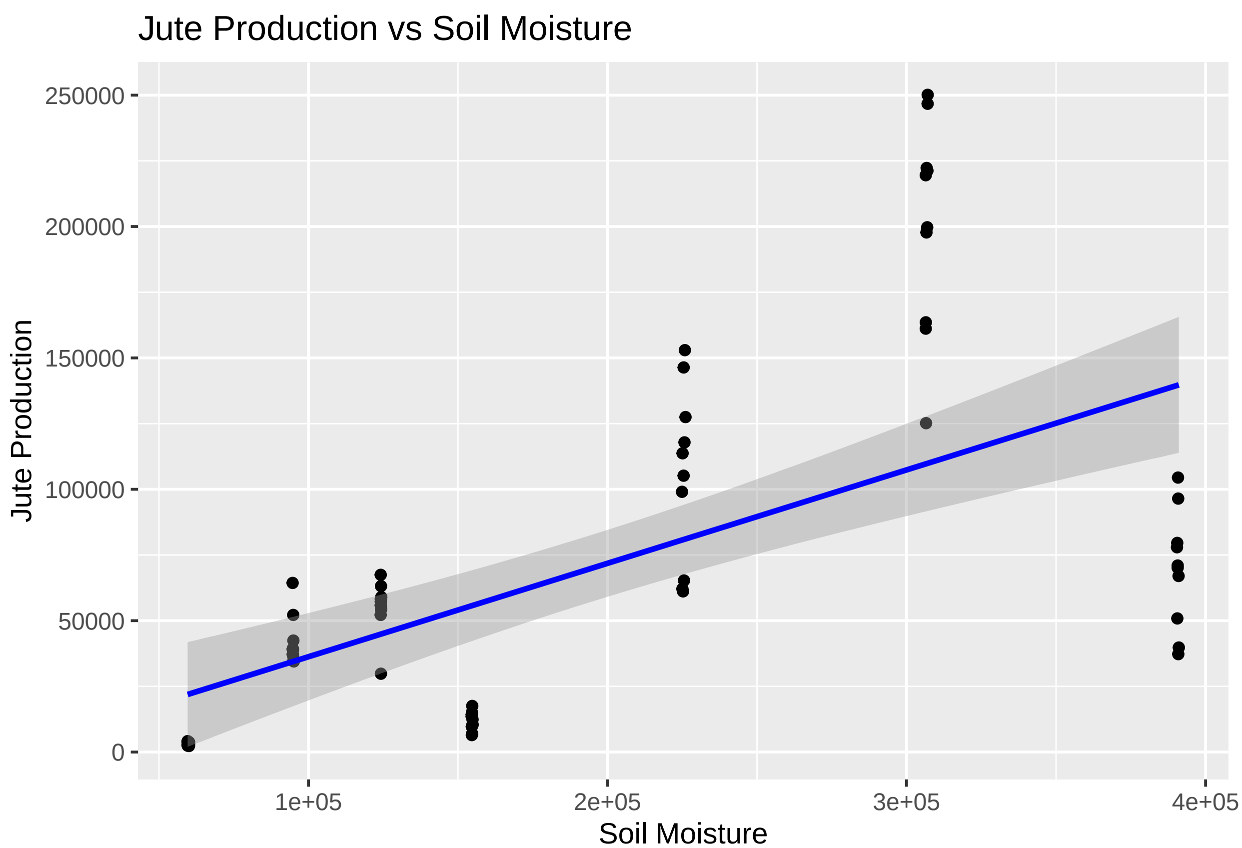

```r
# Jute Production vs. Humidity:
ggplot(Agri_Data, aes(x = humidity, y = jute)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Jute Production vs Humidity", x = "Humidity", y = "Jute Production")
```

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




```r
#  Jute Production vs. Soil Moisture:
ggplot(Agri_Data, aes(x = soil.moisture, y = jute)) +
  geom_point() +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Jute Production vs Soil Moisture", x = "Soil Moisture", y = "Jute Production")
```

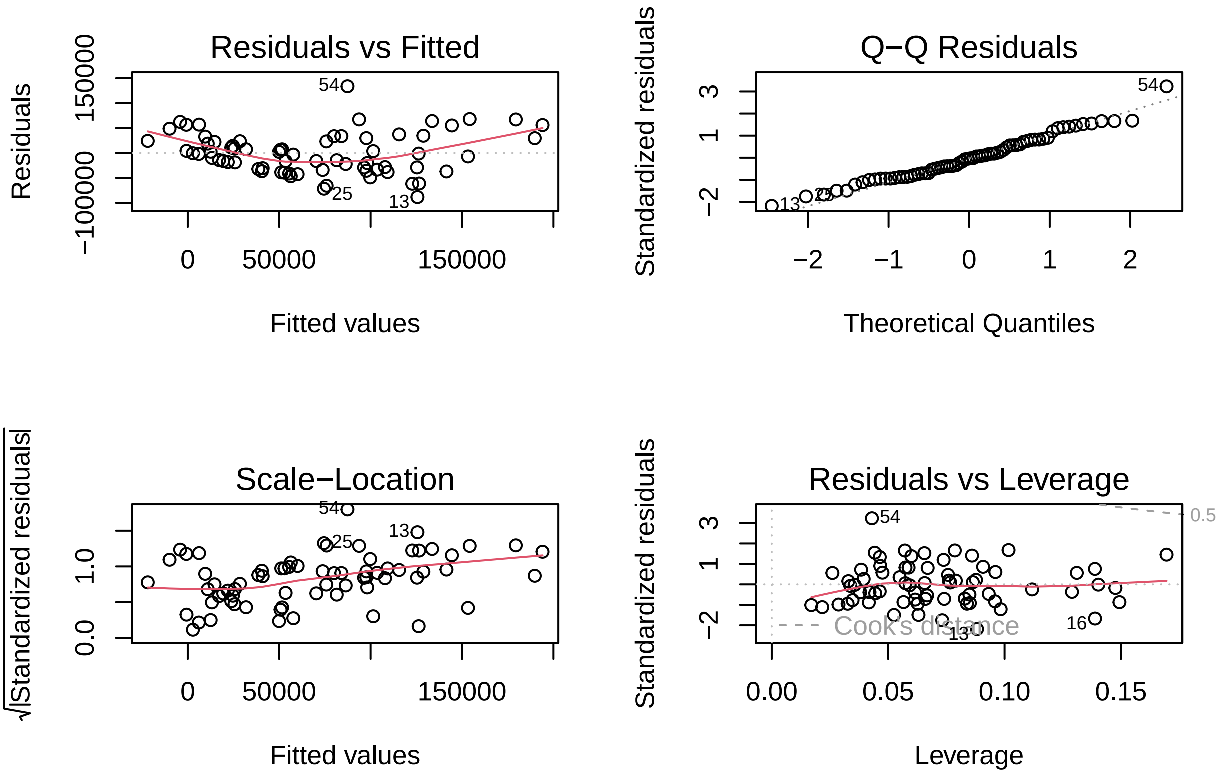
```
## `geom_smooth()` using formula = 'y ~ x'
```




```r
# Linear regression model
model <- lm(jute ~ Avg_temp + avg_rainfall + humidity + soil.moisture, data = Agri_Data)
summary(model)
```

```
##
## Call:
## lm(formula = jute ~ Avg_temp + avg_rainfall + humidity + soil.moisture,
##     data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -88356 -31751  -1507  28370 133872
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.894e+05  8.394e+04   2.256   0.0274 *
## Avg_temp      -1.153e+04  2.445e+03  -4.717 1.31e-05 ***
## avg_rainfall  -4.553e+01  1.764e+01  -2.581   0.0121 *
## humidity       2.215e+03  9.249e+02   2.395   0.0195 *
## soil.moisture  3.397e-01  6.651e-02   5.107 3.08e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 42350 on 65 degrees of freedom
## Multiple R-squared:  0.6148, Adjusted R-squared:  0.5911
## F-statistic: 25.94 on 4 and 65 DF,  p-value: 7.182e-13
```

```r
# Diagnostic plots for the linear regression model
par(mfrow = c(2, 2))
plot(model)
```




```r
## Identify the most and least stormy districts

storm_summary <- Agri_Data %>%
  group_by(District) %>%
  summarise(Storm_Count = sum(storm == "yes", na.rm = TRUE)) %>%
  arrange(desc(Storm_Count))

print(storm_summary)
```

```
## # A tibble: 7 x 2
##   District     Storm_Count
##   <chr>              <int>
## 1 dhaka                 10
## 2 gazipur               10
## 3 kishoregonj           10
## 4 mymenshing             6
## 5 tangail                5
## 6 narayangonj            3
## 7 narsingdi              1
```

```r
most_stormy_district <- storm_summary[1, ]
least_stormy_district <- storm_summary[nrow(storm_summary), ]
print(paste("The most stormy district is", most_stormy_district$District, "with", most_stormy_district$S
```

```
## [1] "The most stormy district is dhaka with 10 storms."
```

```r
print(paste("The least stormy district is", least_stormy_district$District, "with", least_stormy_distri
```

```
## [1] "The least stormy district is narsingdi with 1 storms."
```

```r
## Finding Out the Ideal Fertilizer for Each Crops


## Aus Production

aus_model <- lm(aus ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(aus_model)
```

```
##
## Call:
## lm(formula = aus ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32931  -7808  -3351   2249  37959
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2384.6162  4468.1217  -0.534   0.5954
## urea           -0.1679     0.1303  -1.289   0.2019
## tsp            -0.4764     0.7104  -0.671   0.5048
## mp              1.6262     0.7398   2.198   0.0315 *
## DAP             1.4493     0.9567   1.515   0.1347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15830 on 65 degrees of freedom
## Multiple R-squared:  0.7324, Adjusted R-squared:  0.716
## F-statistic: 44.48 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
## mp has a significant positive coefficicent indicating it's beneficial for Aus Production
## DAP also has a positive coefficient but not as much as mp
## urea & tsp are not actually statistically significant


## Aman Production

aman_model <- lm(aman ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(aman_model)
```

```
##
```

```
## Call:
## lm(formula = aman ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -105391  -35719   -7011   10190  292051
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.142e+04  1.951e+04  -1.610  0.11227
## urea         2.303e+00  5.689e-01   4.048  0.00014 ***
## tsp          3.302e+00  3.103e+00   1.064  0.29109
## mp          -3.281e+00  3.231e+00  -1.015  0.31367
## DAP          1.137e+01  4.178e+00   2.720  0.00836 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 69140 on 65 degrees of freedom
## Multiple R-squared:  0.8798, Adjusted R-squared:  0.8724
## F-statistic: 118.9 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
## Urea and DAP are both statistically significant and have positive impacts on Aman production.Urea and
## TSP and MP do not show statistically significant effects on Aman production in this model.


## Boro Production

boro_model <- lm(boro ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(boro_model)
```

```
##
## Call:
## lm(formula = boro ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -201676  -91210  -30880    6532  842329
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -44328.824  55334.749  -0.801    0.426
## urea             9.619      1.613   5.963 1.12e-07 ***
## tsp             10.241      8.798   1.164    0.249
## mp              -6.465      9.162  -0.706    0.483
## DAP             -8.954     11.849  -0.756    0.453
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 196100 on 65 degrees of freedom
## Multiple R-squared:  0.7219, Adjusted R-squared:  0.7048
## F-statistic: 42.18 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
## Urea is the only fertilizer that has a significant positive impact on crop production. Its p-value i
## TSP, MP, and DAP are not statistically significant, meaning their impacts on crop production cannot 


## Wheat Production

wheat_model <- lm(wheat ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(wheat_model)
```

```
##
## Call:
## lm(formula = wheat ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6063.0 -1478.9   225.3  1090.0  5286.6
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.947e+03  6.935e+02  -5.691 3.26e-07 ***
## urea         2.768e-01  2.022e-02  13.688  < 2e-16 ***
## tsp          3.930e-02  1.103e-01   0.356    0.723
## mp          -9.613e-02  1.148e-01  -0.837    0.406
## DAP         -7.487e-01  1.485e-01  -5.041 3.94e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2457 on 65 degrees of freedom
## Multiple R-squared:  0.7905, Adjusted R-squared:  0.7776
## F-statistic: 61.32 on 4 and 65 DF,  p-value: < 2.2e-16
```

```r
## Urea has a strong positive and highly significant effect on crop yield. Increasing urea usage is lik
## DAP has a strong negative and highly significant effect on crop yield. Increasing DAP usage is likel
## TSP and MP are not statistically significant in this model, meaning their impact on crop production 



## Potato Production

potato_model <- lm(potato ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(potato_model)
```

```
##
## Call:
## lm(formula = potato ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -27381 -11285   1571   6320  41372
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 15513.4140  4240.2970   3.659  0.00051 ***
```

```
## urea           -0.1709     0.1236  -1.382  0.17165
## tsp             3.0621     0.6742   4.542 2.47e-05 ***
## mp              3.6282     0.7021   5.168 2.45e-06 ***
## DAP            -8.7471     0.9080  -9.634 3.85e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15020 on 65 degrees of freedom
## Multiple R-squared:  0.5921, Adjusted R-squared:  0.567
## F-statistic: 23.59 on 4 and 65 DF,  p-value: 4.458e-12
```

```
## TSP and MP have strong positive effects on crop yield. Their p-values are highly significant, indica
## DAP has a strong negative effect on crop yield. It is also highly significant, suggesting that incre
## Urea has a negative coefficient, but its effect is not statistically significant in this model, mean


## Jute Production

jute_model <- lm(jute ~ urea + tsp + mp + DAP, data = Agri_Data)
summary(jute_model)
```

```
##
## Call:
## lm(formula = jute ~ urea + tsp + mp + DAP, data = Agri_Data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -68515  -8919  -1330   7588  59657
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.000e+04  6.770e+03  -5.908 1.39e-07 ***
## urea         3.718e+00  1.974e-01  18.838  < 2e-16 ***
## tsp          2.131e+00  1.076e+00   1.980 0.051939 .
## mp          -4.157e+00  1.121e+00  -3.708 0.000434 ***
## DAP         -7.629e+00  1.450e+00  -5.263 1.71e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23990 on 65 degrees of freedom
## Multiple R-squared:  0.8764, Adjusted R-squared:  0.8688
## F-statistic: 115.2 on 4 and 65 DF,  p-value: < 2.2e-16
```

```
## Urea has a strong positive effect on crop yield and is highly significant, indicating that increasin
## TSP has a marginally positive effect on yield, with borderline statistical significance. Its impact
## MP and DAP both have strong negative effects on crop yield and are statistically significant. Increa

### Land Types

## Inundationland_Highland
highland_data_by_district<- Agri_Data %>%
  arrange(desc(inundationland_Highland)) %>%
  select(District, inundationland_Highland)
```

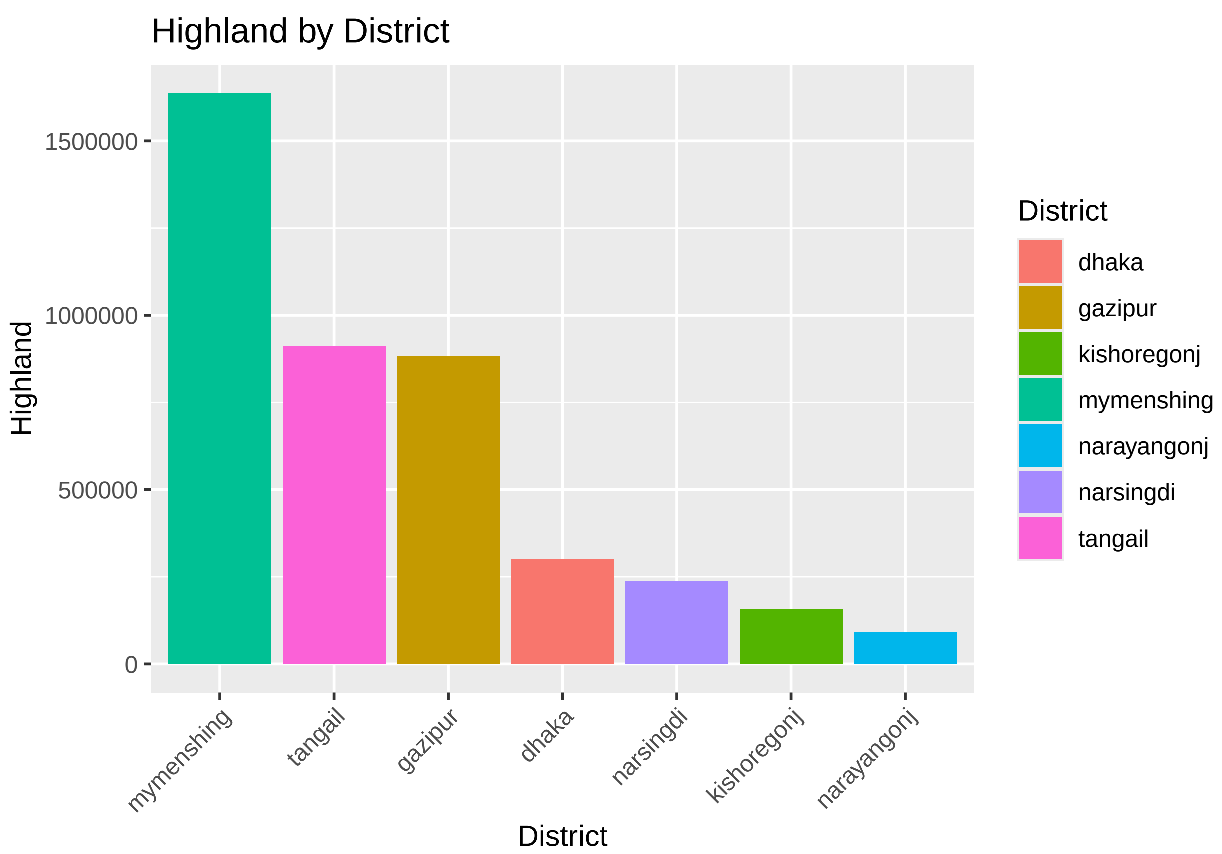
```r
# Most highland District
head(highland_data_by_district, n=1)
```

```
##     District inundationland_Highland
## 1 mymenshing                  163675
```

```r
# Least Highland District
tail(highland_data_by_district, n=1)
```

```
##       District inundationland_Highland
## 70 narayangonj                    9076
```

```r
# Highland Chart Visualization


ggplot(Agri_Data, aes(x = reorder(District, -inundationland_Highland), y = inundationland_Highland, fil
  geom_bar(stat = "identity") +
  labs(title = "Highland by District", x = "District", y = "Highland") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
## Inundationland_Mediumhighland

mediumhighland_data_by_district<- Agri_Data %>%
```

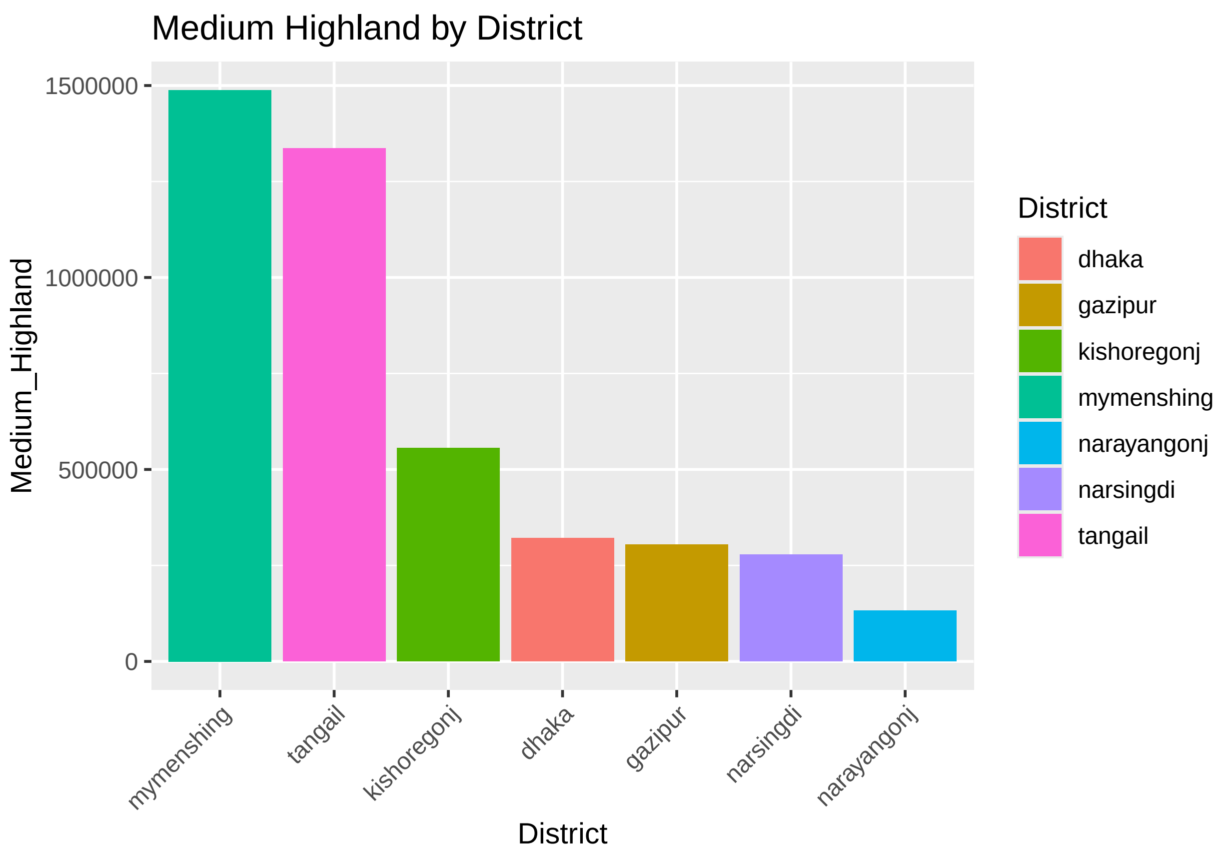
```r
  arrange(desc(inundationland_mediumhighland)) %>%
  select(District, inundationland_mediumhighland)

# Most highland District
head(mediumhighland_data_by_district, n=1)
```

```
##     District inundationland_mediumhighland
## 1 mymenshing                        148743
```

```r
# Least Highland District
tail(mediumhighland_data_by_district, n=1)
```

```
##       District inundationland_mediumhighland
## 70 narayangonj                         13243
```

```r
# MediumHighland Chart Visualization


ggplot(Agri_Data, aes(x = reorder(District, -inundationland_mediumhighland), y = inundationland_mediumhi
  geom_bar(stat = "identity") +
  labs(title = "Medium Highland by District", x = "District", y = "Medium_Highland") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



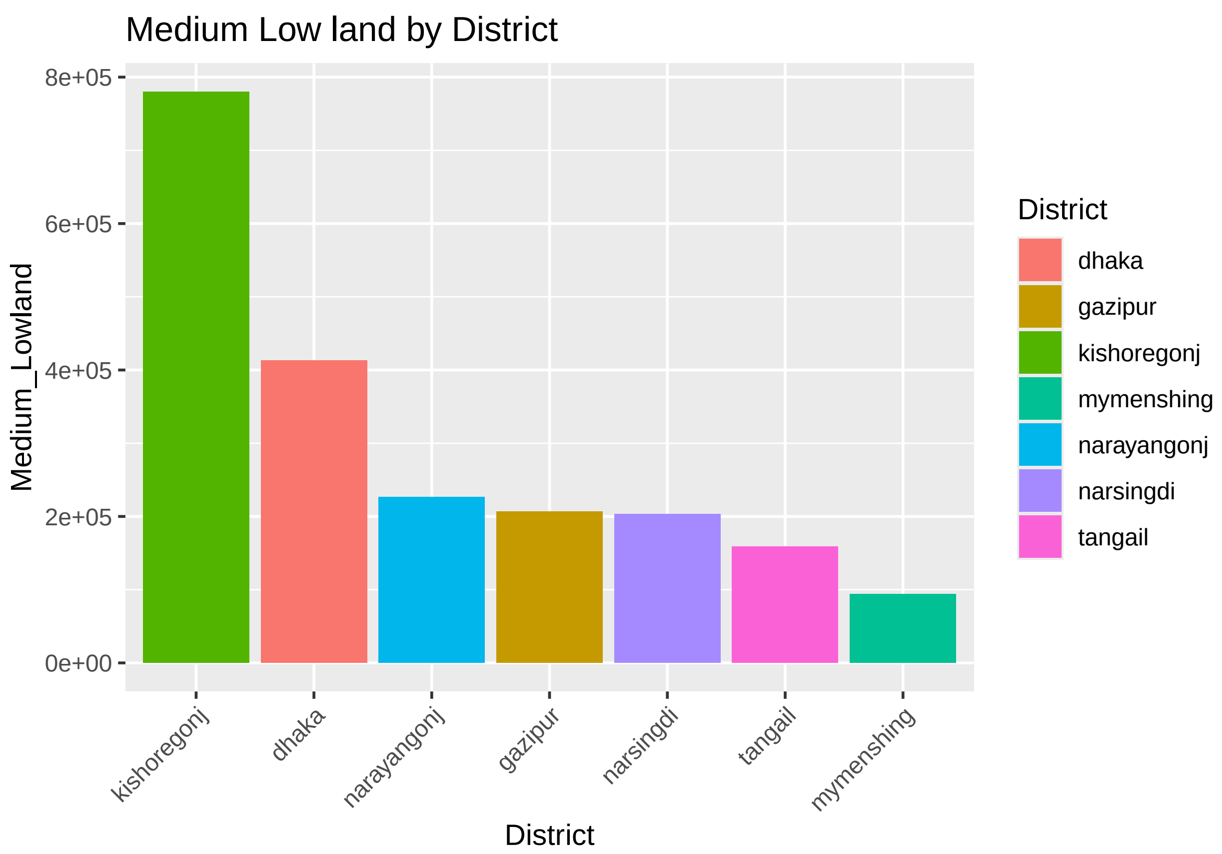
```r
## Inundationland_Medium Low land

mediumlowland_data_by_district<- Agri_Data %>%
  arrange(desc(inundationland_mediumlowland)) %>%
  select(District, inundationland_mediumlowland)

# Most Low land District
head(mediumlowland_data_by_district, n=1)
```

```
##      District inundationland_mediumlowland
## 1 kishoregonj                        77992
```

```r
# Least Low land District
tail(mediumlowland_data_by_district, n=1)
```

```
##      District inundationland_mediumlowland
## 70 mymenshing                         9442
```

```r
# Medium Low land Chart Visualization


ggplot(Agri_Data, aes(x = reorder(District, -inundationland_mediumlowland), y = inundationland_mediumlo
  geom_bar(stat = "identity") +
  labs(title = "Medium Low land by District", x = "District", y = "Medium_Lowland") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



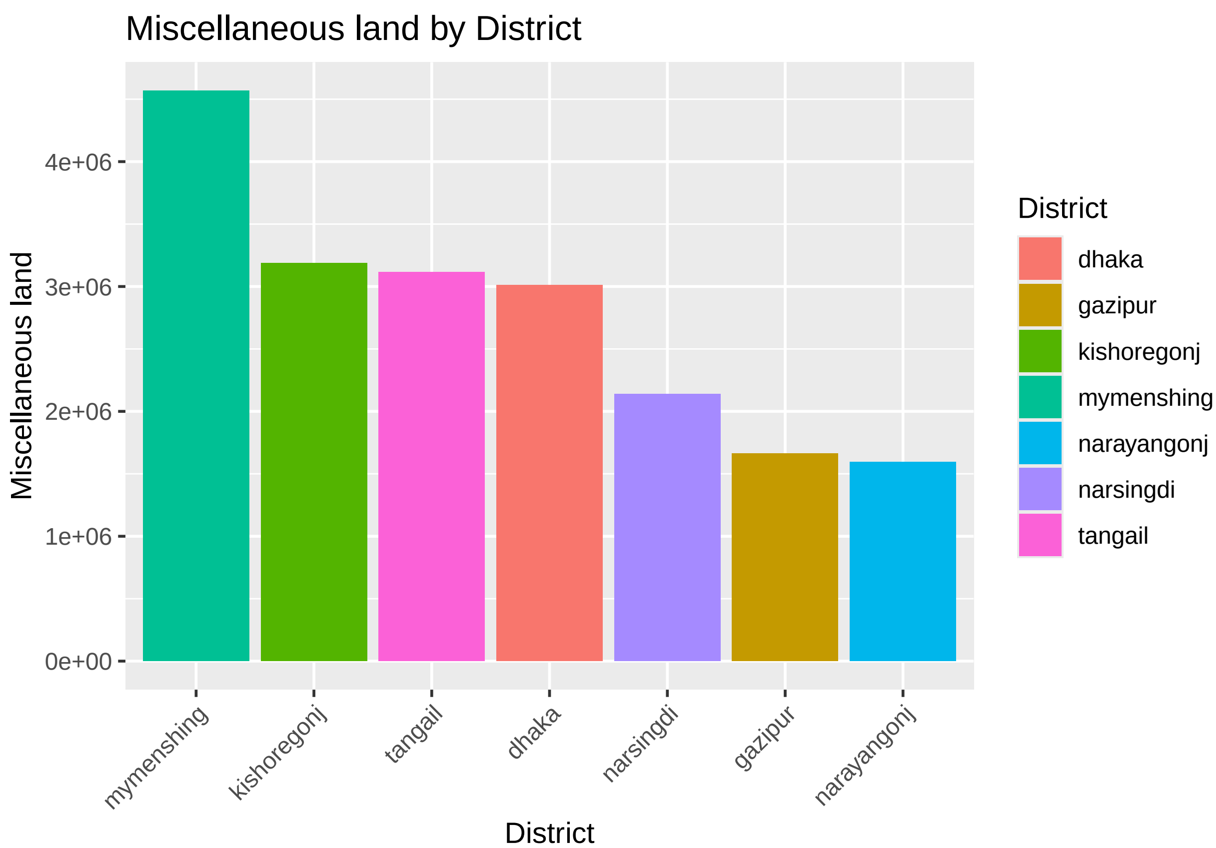
```r
## Miscellaneous land


# Group by district and calculate the total (sum) of Miscellaneous Land for each district
miscellaneousland_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_miscellaneousland = sum(Miscellaneous.Land, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(miscellaneousland_sum_by_district, by = "District")


miscellaneousland_data_by_district<- Agri_Data %>%
  arrange(desc(total_miscellaneousland)) %>%
  select(District, total_miscellaneousland)

# Most Miscellaneous land District
head(miscellaneousland_data_by_district, n=1)
```

```
##     District total_miscellaneousland
## 1 mymenshing                  456786
```

```r
# Least Miscellaneous land District
tail(miscellaneousland_data_by_district, n=1)
```

```
##       District total_miscellaneousland
## 70 narayangonj                  159565
```

```r
# Miscellaneous land Chart Visualization
ggplot(Agri_Data, aes(x = reorder(District, -total_miscellaneousland), y = total_miscellaneousland, fil
  geom_bar(stat = "identity") +
  labs(title = "Miscellaneous land by District", x = "District", y = "Miscellaneous land") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```
### Soil Type


# Group by district and calculate the total (sum) of Noncalcareous Alluvium soil for each district
noncalcareous_alluvium_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_noncalcareous_alluvium= sum(Noncalcareous.Alluvium, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(noncalcareous_alluvium_sum_by_district, by = "District")


noncalcareous_alluvium_sum_by_district<- Agri_Data %>%
  arrange(desc(total_noncalcareous_alluvium)) %>%
  select(District, total_noncalcareous_alluvium)

# Most Noncalcareous Alluvium soil District
head(noncalcareous_alluvium_sum_by_district, n=1)
```

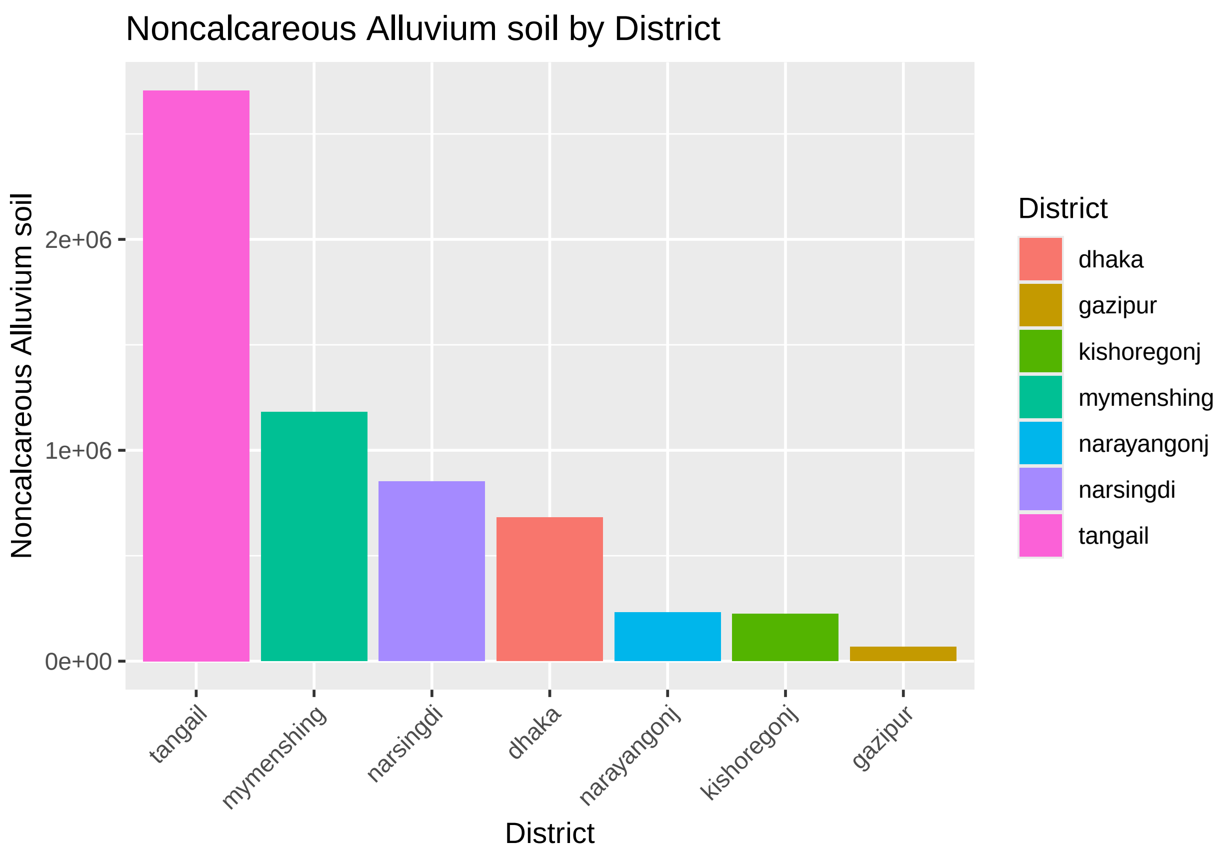
```
##   District total_noncalcareous_alluvium
## 1  tangail                       270457
```

```
# Least Noncalcareous Alluvium soil District
tail(noncalcareous_alluvium_sum_by_district, n=1)
```

```
##    District total_noncalcareous_alluvium
```

```
## 70   gazipur                          6947
```

```
# Noncalcareous Alluvium soil Chart Visualization
ggplot(Agri_Data, aes(x = reorder(District, -total_noncalcareous_alluvium), y = total_noncalcareous_allu
  geom_bar(stat = "identity") +
  labs(title = "Noncalcareous Alluvium soil by District", x = "District", y = "Noncalcareous Alluvium so
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```
# Group by district and calculate the total (sum) of Acid Basin Clay for each district
acid_basin_clay_sum_by_district <- Agri_Data %>%
  group_by(District) %>%
  summarise(total_acid_basin_clay= sum(Acid.Basin.Clay, na.rm = TRUE))

Agri_Data <- Agri_Data %>%
  left_join(acid_basin_clay_sum_by_district, by = "District")


acid_basin_clay_sum_by_district<- Agri_Data %>%
  arrange(desc(total_acid_basin_clay)) %>%
  select(District, total_acid_basin_clay)

# Most Acid Basin Clay District
head(acid_basin_clay_sum_by_district, n=1)
```

```
##     District total_acid_basin_clay
## 1 mymenshing                329078
```

```r
# Least Acid Basin Clay District
tail(acid_basin_clay_sum_by_district, n=1)
```

```
##       District total_acid_basin_clay
## 70 narayangonj                 31545
```

```r
# Noncalcareous Alluvium soil Chart Visualization
ggplot(Agri_Data, aes(x = reorder(District, -total_acid_basin_clay), y = total_acid_basin_clay, fill = D
  geom_bar(stat = "identity") +
  labs(title = "Acid Basin Clay by District", x = "District", y = "Acid Basin Clay") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```