

# Rworksheet\_Barrientos#4C

Barrientos, Milfrance D.

11/03/2024

*#1. Use the dataset mpg*

*#a. Show your solutions on how to import a csv file into the environment.*

*# Ensure that readr is installed separately, then load it*

```
library(readr)
```

```
data <- read_csv("C:/PROJ/mpg.csv", show_col_types = FALSE)
```

```
## New names:
```

```
## * ' ' -> '...1'
```

```
data
```

```
## # A tibble: 234 x 12
```

```
##   ...1 manufacturer model      displ  year   cyl trans drv      cty   hwy fl  
##   <dbl> <chr>         <chr>    <dbl> <dbl> <dbl> <chr> <chr> <dbl> <dbl> <chr>  
## 1     1 audi         a4        1.8  1999     4 auto~ f      18    29 p  
## 2     2 audi         a4        1.8  1999     4 manu~ f      21    29 p  
## 3     3 audi         a4         2   2008     4 manu~ f      20    31 p  
## 4     4 audi         a4         2   2008     4 auto~ f      21    30 p  
## 5     5 audi         a4        2.8  1999     6 auto~ f      16    26 p  
## 6     6 audi         a4        2.8  1999     6 manu~ f      18    26 p  
## 7     7 audi         a4        3.1  2008     6 auto~ f      18    27 p  
## 8     8 audi         a4 quattro 1.8  1999     4 manu~ 4      18    26 p  
## 9     9 audi         a4 quattro 1.8  1999     4 auto~ 4      16    25 p  
## 10    10 audi         a4 quattro 2     2008     4 manu~ 4      20    28 p
```

```
## # i 224 more rows
```

```
## # i 1 more variable: class <chr>
```

*#b. Which variables from mpg dataset are categorical?*

*# The variables from mpg dataset that are categorical are manufacturer, model, transmission, drv, fl, c*

*#c. Which are continuous variables?*

*# The continuous variables are displ, year, cyl, cty, hwy.*

*#2. Which manufacturer has the most models in this data set? Which model has the most variations?*

*# Show your answer.*

```
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
```

```
library(ggplot2)
data(mpg)

manufacturer_models <- mpg %>%
  group_by(manufacturer) %>%
  summarise(model_count = n_distinct(model)) %>%
  arrange(desc(model_count))

manufacturer_most_models <- manufacturer_models %>% slice(1)
manufacturer_most_models
```

```
## # A tibble: 1 x 2
##   manufacturer model_count
##   <chr>           <int>
## 1 toyota             6
```

*# Toyota has the most models with 6 different models in the data set.*

```
model_variations <- mpg %>%
  group_by(model) %>%
  summarise(variation_count = n_distinct(manufacturer)) %>%
  arrange(desc(variation_count))

model_most_variations <- model_variations %>% slice(1)
model_most_variations
```

```
## # A tibble: 1 x 2
##   model          variation_count
##   <chr>           <int>
## 1 4runner 4wd             1
```

*# The most variations is the model "caravan 2wd" with 11 variations.*

*#a. Group the manufacturers and find the unique models. Show your codes and result.*

```
manufacturer_unique_models <- data %>%
  group_by(manufacturer) %>%
  summarise(unique_model_count = n_distinct(model))

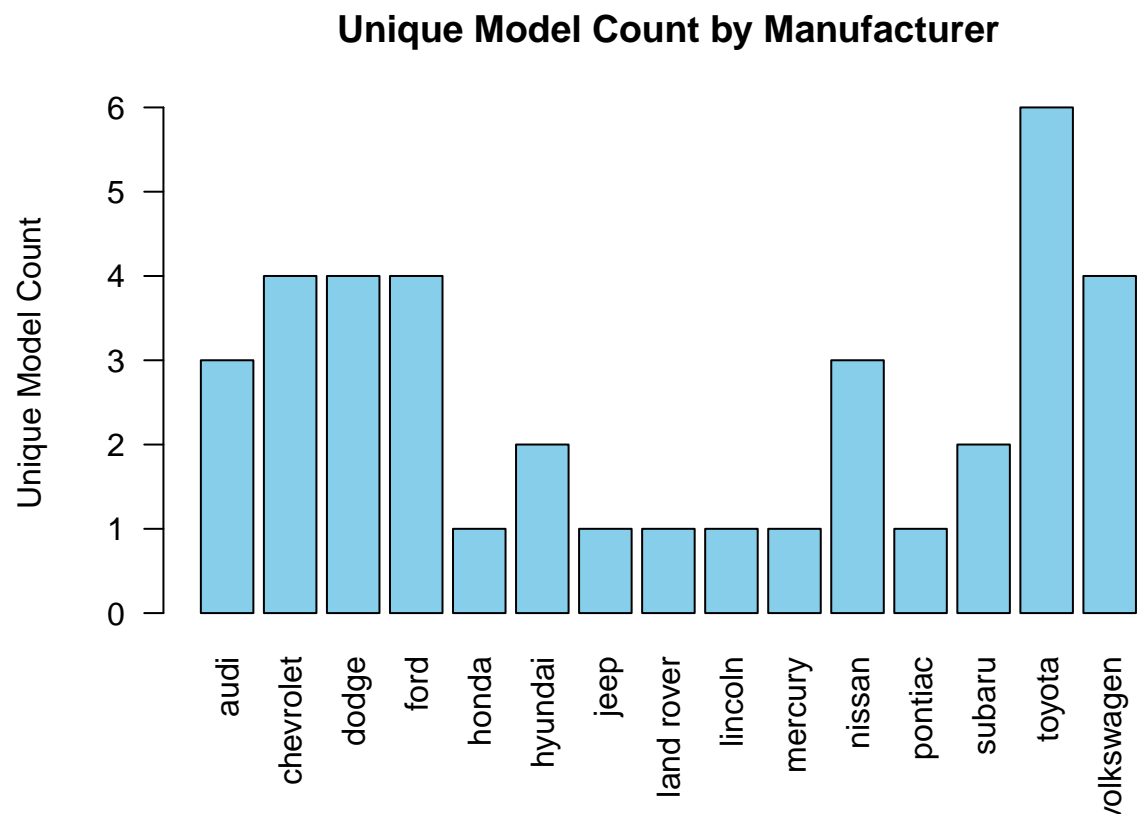
manufacturer_unique_models
```

```
## # A tibble: 15 x 2
##   manufacturer unique_model_count
##   <chr>           <int>
```

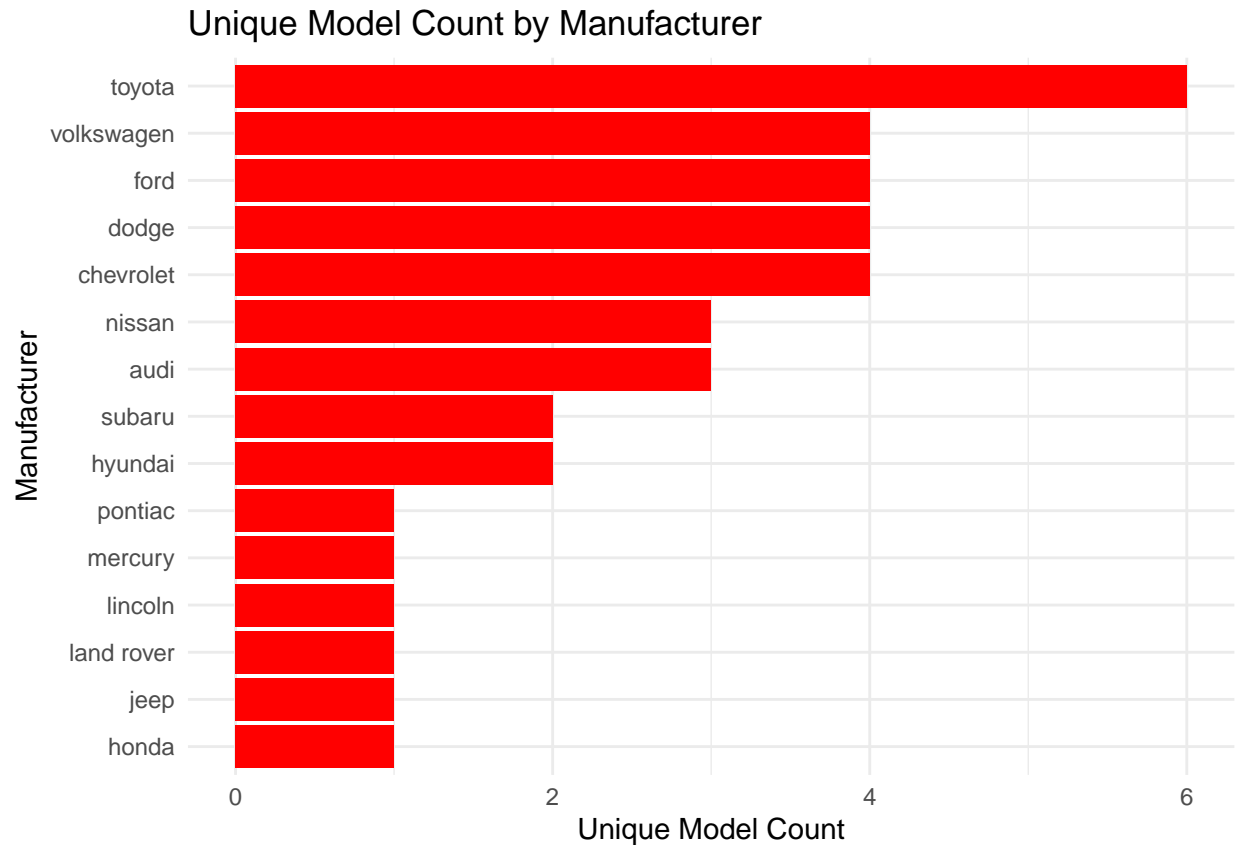
##	1	audi	3
##	2	chevrolet	4
##	3	dodge	4
##	4	ford	4
##	5	honda	1
##	6	hyundai	2
##	7	jeep	1
##	8	land rover	1
##	9	lincoln	1
##	10	mercury	1
##	11	nissan	3
##	12	pontiac	1
##	13	subaru	2
##	14	toyota	6
##	15	volkswagen	4

*#b. Graph the result by using plot() and ggplot(). Write the codes and its result.*

```
# Base R barplot
barplot(
  manufacturer_unique_models$unique_model_count,
  names.arg = manufacturer_unique_models$manufacturer,
  las = 2,
  col = "skyblue",
  main = "Unique Model Count by Manufacturer",
  ylab = "Unique Model Count"
)
```



```
# ggplot2
library(ggplot2)
ggplot(manufacturer_unique_models, aes(x = reorder(manufacturer, unique_model_count), y = unique_model_count)) +
  geom_bar(stat = "identity", fill = "red") +
  coord_flip() +
  labs(
    title = "Unique Model Count by Manufacturer",
    x = "Manufacturer",
    y = "Unique Model Count"
  ) +
  theme_minimal()
```



*#2. Same dataset will be used. You are going to show the relationship of the model and the manufacturer.*

```
library(dplyr)
data <- read.csv("C:/PROJ/mpg.csv")
model_by_manufacturer <- table(data$manufacturer, data$model)
data
```

##	X	manufacturer	model	displ	year	cyl	trans	drv	cty
## 1	1	audi	a4	1.8	1999	4	auto(l5)	f	18
## 2	2	audi	a4	1.8	1999	4	manual(m5)	f	21
## 3	3	audi	a4	2.0	2008	4	manual(m6)	f	20
## 4	4	audi	a4	2.0	2008	4	auto(av)	f	21
## 5	5	audi	a4	2.8	1999	6	auto(l5)	f	16
## 6	6	audi	a4	2.8	1999	6	manual(m5)	f	18
## 7	7	audi	a4	3.1	2008	6	auto(av)	f	18
## 8	8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18
## 9	9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16
## 10	10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20
## 11	11	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19
## 12	12	audi	a4 quattro	2.8	1999	6	auto(l5)	4	15
## 13	13	audi	a4 quattro	2.8	1999	6	manual(m5)	4	17
## 14	14	audi	a4 quattro	3.1	2008	6	auto(s6)	4	17
## 15	15	audi	a4 quattro	3.1	2008	6	manual(m6)	4	15
## 16	16	audi	a6 quattro	2.8	1999	6	auto(l5)	4	15
## 17	17	audi	a6 quattro	3.1	2008	6	auto(s6)	4	17
## 18	18	audi	a6 quattro	4.2	2008	8	auto(s6)	4	16
## 19	19	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14

## 20	20	chevrolet	c1500 suburban 2wd	5.3 2008	8	auto(14)	r	11
## 21	21	chevrolet	c1500 suburban 2wd	5.3 2008	8	auto(14)	r	14
## 22	22	chevrolet	c1500 suburban 2wd	5.7 1999	8	auto(14)	r	13
## 23	23	chevrolet	c1500 suburban 2wd	6.0 2008	8	auto(14)	r	12
## 24	24	chevrolet	corvette	5.7 1999	8	manual(m6)	r	16
## 25	25	chevrolet	corvette	5.7 1999	8	auto(14)	r	15
## 26	26	chevrolet	corvette	6.2 2008	8	manual(m6)	r	16
## 27	27	chevrolet	corvette	6.2 2008	8	auto(s6)	r	15
## 28	28	chevrolet	corvette	7.0 2008	8	manual(m6)	r	15
## 29	29	chevrolet	k1500 tahoe 4wd	5.3 2008	8	auto(14)	4	14
## 30	30	chevrolet	k1500 tahoe 4wd	5.3 2008	8	auto(14)	4	11
## 31	31	chevrolet	k1500 tahoe 4wd	5.7 1999	8	auto(14)	4	11
## 32	32	chevrolet	k1500 tahoe 4wd	6.5 1999	8	auto(14)	4	14
## 33	33	chevrolet	malibu	2.4 1999	4	auto(14)	f	19
## 34	34	chevrolet	malibu	2.4 2008	4	auto(14)	f	22
## 35	35	chevrolet	malibu	3.1 1999	6	auto(14)	f	18
## 36	36	chevrolet	malibu	3.5 2008	6	auto(14)	f	18
## 37	37	chevrolet	malibu	3.6 2008	6	auto(s6)	f	17
## 38	38	dodge	caravan 2wd	2.4 1999	4	auto(13)	f	18
## 39	39	dodge	caravan 2wd	3.0 1999	6	auto(14)	f	17
## 40	40	dodge	caravan 2wd	3.3 1999	6	auto(14)	f	16
## 41	41	dodge	caravan 2wd	3.3 1999	6	auto(14)	f	16
## 42	42	dodge	caravan 2wd	3.3 2008	6	auto(14)	f	17
## 43	43	dodge	caravan 2wd	3.3 2008	6	auto(14)	f	17
## 44	44	dodge	caravan 2wd	3.3 2008	6	auto(14)	f	11
## 45	45	dodge	caravan 2wd	3.8 1999	6	auto(14)	f	15
## 46	46	dodge	caravan 2wd	3.8 1999	6	auto(14)	f	15
## 47	47	dodge	caravan 2wd	3.8 2008	6	auto(16)	f	16
## 48	48	dodge	caravan 2wd	4.0 2008	6	auto(16)	f	16
## 49	49	dodge	dakota pickup 4wd	3.7 2008	6	manual(m6)	4	15
## 50	50	dodge	dakota pickup 4wd	3.7 2008	6	auto(14)	4	14
## 51	51	dodge	dakota pickup 4wd	3.9 1999	6	auto(14)	4	13
## 52	52	dodge	dakota pickup 4wd	3.9 1999	6	manual(m5)	4	14
## 53	53	dodge	dakota pickup 4wd	4.7 2008	8	auto(15)	4	14
## 54	54	dodge	dakota pickup 4wd	4.7 2008	8	auto(15)	4	14
## 55	55	dodge	dakota pickup 4wd	4.7 2008	8	auto(15)	4	9
## 56	56	dodge	dakota pickup 4wd	5.2 1999	8	manual(m5)	4	11
## 57	57	dodge	dakota pickup 4wd	5.2 1999	8	auto(14)	4	11
## 58	58	dodge	durango 4wd	3.9 1999	6	auto(14)	4	13
## 59	59	dodge	durango 4wd	4.7 2008	8	auto(15)	4	13
## 60	60	dodge	durango 4wd	4.7 2008	8	auto(15)	4	9
## 61	61	dodge	durango 4wd	4.7 2008	8	auto(15)	4	13
## 62	62	dodge	durango 4wd	5.2 1999	8	auto(14)	4	11
## 63	63	dodge	durango 4wd	5.7 2008	8	auto(15)	4	13
## 64	64	dodge	durango 4wd	5.9 1999	8	auto(14)	4	11
## 65	65	dodge	ram 1500 pickup 4wd	4.7 2008	8	manual(m6)	4	12
## 66	66	dodge	ram 1500 pickup 4wd	4.7 2008	8	auto(15)	4	9
## 67	67	dodge	ram 1500 pickup 4wd	4.7 2008	8	auto(15)	4	13
## 68	68	dodge	ram 1500 pickup 4wd	4.7 2008	8	auto(15)	4	13
## 69	69	dodge	ram 1500 pickup 4wd	4.7 2008	8	manual(m6)	4	12
## 70	70	dodge	ram 1500 pickup 4wd	4.7 2008	8	manual(m6)	4	9
## 71	71	dodge	ram 1500 pickup 4wd	5.2 1999	8	auto(14)	4	11
## 72	72	dodge	ram 1500 pickup 4wd	5.2 1999	8	manual(m5)	4	11
## 73	73	dodge	ram 1500 pickup 4wd	5.7 2008	8	auto(15)	4	13

## 74	74	dodge	ram 1500 pickup 4wd	5.9	1999	8	auto(14)	4	11
## 75	75	ford	expedition 2wd	4.6	1999	8	auto(14)	r	11
## 76	76	ford	expedition 2wd	5.4	1999	8	auto(14)	r	11
## 77	77	ford	expedition 2wd	5.4	2008	8	auto(16)	r	12
## 78	78	ford	explorer 4wd	4.0	1999	6	auto(15)	4	14
## 79	79	ford	explorer 4wd	4.0	1999	6	manual(m5)	4	15
## 80	80	ford	explorer 4wd	4.0	1999	6	auto(15)	4	14
## 81	81	ford	explorer 4wd	4.0	2008	6	auto(15)	4	13
## 82	82	ford	explorer 4wd	4.6	2008	8	auto(16)	4	13
## 83	83	ford	explorer 4wd	5.0	1999	8	auto(14)	4	13
## 84	84	ford	f150 pickup 4wd	4.2	1999	6	auto(14)	4	14
## 85	85	ford	f150 pickup 4wd	4.2	1999	6	manual(m5)	4	14
## 86	86	ford	f150 pickup 4wd	4.6	1999	8	manual(m5)	4	13
## 87	87	ford	f150 pickup 4wd	4.6	1999	8	auto(14)	4	13
## 88	88	ford	f150 pickup 4wd	4.6	2008	8	auto(14)	4	13
## 89	89	ford	f150 pickup 4wd	5.4	1999	8	auto(14)	4	11
## 90	90	ford	f150 pickup 4wd	5.4	2008	8	auto(14)	4	13
## 91	91	ford	mustang	3.8	1999	6	manual(m5)	r	18
## 92	92	ford	mustang	3.8	1999	6	auto(14)	r	18
## 93	93	ford	mustang	4.0	2008	6	manual(m5)	r	17
## 94	94	ford	mustang	4.0	2008	6	auto(15)	r	16
## 95	95	ford	mustang	4.6	1999	8	auto(14)	r	15
## 96	96	ford	mustang	4.6	1999	8	manual(m5)	r	15
## 97	97	ford	mustang	4.6	2008	8	manual(m5)	r	15
## 98	98	ford	mustang	4.6	2008	8	auto(15)	r	15
## 99	99	ford	mustang	5.4	2008	8	manual(m6)	r	14
## 100	100	honda	civic	1.6	1999	4	manual(m5)	f	28
## 101	101	honda	civic	1.6	1999	4	auto(14)	f	24
## 102	102	honda	civic	1.6	1999	4	manual(m5)	f	25
## 103	103	honda	civic	1.6	1999	4	manual(m5)	f	23
## 104	104	honda	civic	1.6	1999	4	auto(14)	f	24
## 105	105	honda	civic	1.8	2008	4	manual(m5)	f	26
## 106	106	honda	civic	1.8	2008	4	auto(15)	f	25
## 107	107	honda	civic	1.8	2008	4	auto(15)	f	24
## 108	108	honda	civic	2.0	2008	4	manual(m6)	f	21
## 109	109	hyundai	sonata	2.4	1999	4	auto(14)	f	18
## 110	110	hyundai	sonata	2.4	1999	4	manual(m5)	f	18
## 111	111	hyundai	sonata	2.4	2008	4	auto(14)	f	21
## 112	112	hyundai	sonata	2.4	2008	4	manual(m5)	f	21
## 113	113	hyundai	sonata	2.5	1999	6	auto(14)	f	18
## 114	114	hyundai	sonata	2.5	1999	6	manual(m5)	f	18
## 115	115	hyundai	sonata	3.3	2008	6	auto(15)	f	19
## 116	116	hyundai	tiburon	2.0	1999	4	auto(14)	f	19
## 117	117	hyundai	tiburon	2.0	1999	4	manual(m5)	f	19
## 118	118	hyundai	tiburon	2.0	2008	4	manual(m5)	f	20
## 119	119	hyundai	tiburon	2.0	2008	4	auto(14)	f	20
## 120	120	hyundai	tiburon	2.7	2008	6	auto(14)	f	17
## 121	121	hyundai	tiburon	2.7	2008	6	manual(m6)	f	16
## 122	122	hyundai	tiburon	2.7	2008	6	manual(m5)	f	17
## 123	123	jeep	grand cherokee 4wd	3.0	2008	6	auto(15)	4	17
## 124	124	jeep	grand cherokee 4wd	3.7	2008	6	auto(15)	4	15
## 125	125	jeep	grand cherokee 4wd	4.0	1999	6	auto(14)	4	15
## 126	126	jeep	grand cherokee 4wd	4.7	1999	8	auto(14)	4	14
## 127	127	jeep	grand cherokee 4wd	4.7	2008	8	auto(15)	4	9

## 128 128	jeep	grand cherokee 4wd	4.7 2008	8	auto(15)	4	14
## 129 129	jeep	grand cherokee 4wd	5.7 2008	8	auto(15)	4	13
## 130 130	jeep	grand cherokee 4wd	6.1 2008	8	auto(15)	4	11
## 131 131	land rover	range rover	4.0 1999	8	auto(14)	4	11
## 132 132	land rover	range rover	4.2 2008	8	auto(s6)	4	12
## 133 133	land rover	range rover	4.4 2008	8	auto(s6)	4	12
## 134 134	land rover	range rover	4.6 1999	8	auto(14)	4	11
## 135 135	lincoln	navigator 2wd	5.4 1999	8	auto(14)	r	11
## 136 136	lincoln	navigator 2wd	5.4 1999	8	auto(14)	r	11
## 137 137	lincoln	navigator 2wd	5.4 2008	8	auto(16)	r	12
## 138 138	mercury	mountaineer 4wd	4.0 1999	6	auto(15)	4	14
## 139 139	mercury	mountaineer 4wd	4.0 2008	6	auto(15)	4	13
## 140 140	mercury	mountaineer 4wd	4.6 2008	8	auto(16)	4	13
## 141 141	mercury	mountaineer 4wd	5.0 1999	8	auto(14)	4	13
## 142 142	nissan	altima	2.4 1999	4	manual(m5)	f	21
## 143 143	nissan	altima	2.4 1999	4	auto(14)	f	19
## 144 144	nissan	altima	2.5 2008	4	auto(av)	f	23
## 145 145	nissan	altima	2.5 2008	4	manual(m6)	f	23
## 146 146	nissan	altima	3.5 2008	6	manual(m6)	f	19
## 147 147	nissan	altima	3.5 2008	6	auto(av)	f	19
## 148 148	nissan	maxima	3.0 1999	6	auto(14)	f	18
## 149 149	nissan	maxima	3.0 1999	6	manual(m5)	f	19
## 150 150	nissan	maxima	3.5 2008	6	auto(av)	f	19
## 151 151	nissan	pathfinder 4wd	3.3 1999	6	auto(14)	4	14
## 152 152	nissan	pathfinder 4wd	3.3 1999	6	manual(m5)	4	15
## 153 153	nissan	pathfinder 4wd	4.0 2008	6	auto(15)	4	14
## 154 154	nissan	pathfinder 4wd	5.6 2008	8	auto(s5)	4	12
## 155 155	pontiac	grand prix	3.1 1999	6	auto(14)	f	18
## 156 156	pontiac	grand prix	3.8 1999	6	auto(14)	f	16
## 157 157	pontiac	grand prix	3.8 1999	6	auto(14)	f	17
## 158 158	pontiac	grand prix	3.8 2008	6	auto(14)	f	18
## 159 159	pontiac	grand prix	5.3 2008	8	auto(s4)	f	16
## 160 160	subaru	forester awd	2.5 1999	4	manual(m5)	4	18
## 161 161	subaru	forester awd	2.5 1999	4	auto(14)	4	18
## 162 162	subaru	forester awd	2.5 2008	4	manual(m5)	4	20
## 163 163	subaru	forester awd	2.5 2008	4	manual(m5)	4	19
## 164 164	subaru	forester awd	2.5 2008	4	auto(14)	4	20
## 165 165	subaru	forester awd	2.5 2008	4	auto(14)	4	18
## 166 166	subaru	impreza awd	2.2 1999	4	auto(14)	4	21
## 167 167	subaru	impreza awd	2.2 1999	4	manual(m5)	4	19
## 168 168	subaru	impreza awd	2.5 1999	4	manual(m5)	4	19
## 169 169	subaru	impreza awd	2.5 1999	4	auto(14)	4	19
## 170 170	subaru	impreza awd	2.5 2008	4	auto(s4)	4	20
## 171 171	subaru	impreza awd	2.5 2008	4	auto(s4)	4	20
## 172 172	subaru	impreza awd	2.5 2008	4	manual(m5)	4	19
## 173 173	subaru	impreza awd	2.5 2008	4	manual(m5)	4	20
## 174 174	toyota	4runner 4wd	2.7 1999	4	manual(m5)	4	15
## 175 175	toyota	4runner 4wd	2.7 1999	4	auto(14)	4	16
## 176 176	toyota	4runner 4wd	3.4 1999	6	auto(14)	4	15
## 177 177	toyota	4runner 4wd	3.4 1999	6	manual(m5)	4	15
## 178 178	toyota	4runner 4wd	4.0 2008	6	auto(15)	4	16
## 179 179	toyota	4runner 4wd	4.7 2008	8	auto(15)	4	14
## 180 180	toyota	camry	2.2 1999	4	manual(m5)	f	21
## 181 181	toyota	camry	2.2 1999	4	auto(14)	f	21



##	182	182	toyota	camry	2.4	2008	4	manual(m5)	f	21
##	183	183	toyota	camry	2.4	2008	4	auto(15)	f	21
##	184	184	toyota	camry	3.0	1999	6	auto(14)	f	18
##	185	185	toyota	camry	3.0	1999	6	manual(m5)	f	18
##	186	186	toyota	camry	3.5	2008	6	auto(s6)	f	19
##	187	187	toyota	camry solara	2.2	1999	4	auto(14)	f	21
##	188	188	toyota	camry solara	2.2	1999	4	manual(m5)	f	21
##	189	189	toyota	camry solara	2.4	2008	4	manual(m5)	f	21
##	190	190	toyota	camry solara	2.4	2008	4	auto(s5)	f	22
##	191	191	toyota	camry solara	3.0	1999	6	auto(14)	f	18
##	192	192	toyota	camry solara	3.0	1999	6	manual(m5)	f	18
##	193	193	toyota	camry solara	3.3	2008	6	auto(s5)	f	18
##	194	194	toyota	corolla	1.8	1999	4	auto(13)	f	24
##	195	195	toyota	corolla	1.8	1999	4	auto(14)	f	24
##	196	196	toyota	corolla	1.8	1999	4	manual(m5)	f	26
##	197	197	toyota	corolla	1.8	2008	4	manual(m5)	f	28
##	198	198	toyota	corolla	1.8	2008	4	auto(14)	f	26
##	199	199	toyota	land cruiser wagon 4wd	4.7	1999	8	auto(14)	4	11
##	200	200	toyota	land cruiser wagon 4wd	5.7	2008	8	auto(s6)	4	13
##	201	201	toyota	toyota tacoma 4wd	2.7	1999	4	manual(m5)	4	15
##	202	202	toyota	toyota tacoma 4wd	2.7	1999	4	auto(14)	4	16
##	203	203	toyota	toyota tacoma 4wd	2.7	2008	4	manual(m5)	4	17
##	204	204	toyota	toyota tacoma 4wd	3.4	1999	6	manual(m5)	4	15
##	205	205	toyota	toyota tacoma 4wd	3.4	1999	6	auto(14)	4	15
##	206	206	toyota	toyota tacoma 4wd	4.0	2008	6	manual(m6)	4	15
##	207	207	toyota	toyota tacoma 4wd	4.0	2008	6	auto(15)	4	16
##	208	208	volkswagen	gti	2.0	1999	4	manual(m5)	f	21
##	209	209	volkswagen	gti	2.0	1999	4	auto(14)	f	19
##	210	210	volkswagen	gti	2.0	2008	4	manual(m6)	f	21
##	211	211	volkswagen	gti	2.0	2008	4	auto(s6)	f	22
##	212	212	volkswagen	gti	2.8	1999	6	manual(m5)	f	17
##	213	213	volkswagen	jetta	1.9	1999	4	manual(m5)	f	33
##	214	214	volkswagen	jetta	2.0	1999	4	manual(m5)	f	21
##	215	215	volkswagen	jetta	2.0	1999	4	auto(14)	f	19
##	216	216	volkswagen	jetta	2.0	2008	4	auto(s6)	f	22
##	217	217	volkswagen	jetta	2.0	2008	4	manual(m6)	f	21
##	218	218	volkswagen	jetta	2.5	2008	5	auto(s6)	f	21
##	219	219	volkswagen	jetta	2.5	2008	5	manual(m5)	f	21
##	220	220	volkswagen	jetta	2.8	1999	6	auto(14)	f	16
##	221	221	volkswagen	jetta	2.8	1999	6	manual(m5)	f	17
##	222	222	volkswagen	new beetle	1.9	1999	4	manual(m5)	f	35
##	223	223	volkswagen	new beetle	1.9	1999	4	auto(14)	f	29
##	224	224	volkswagen	new beetle	2.0	1999	4	manual(m5)	f	21
##	225	225	volkswagen	new beetle	2.0	1999	4	auto(14)	f	19
##	226	226	volkswagen	new beetle	2.5	2008	5	manual(m5)	f	20
##	227	227	volkswagen	new beetle	2.5	2008	5	auto(s6)	f	20
##	228	228	volkswagen	passat	1.8	1999	4	manual(m5)	f	21
##	229	229	volkswagen	passat	1.8	1999	4	auto(15)	f	18
##	230	230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19
##	231	231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21
##	232	232	volkswagen	passat	2.8	1999	6	auto(15)	f	16
##	233	233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18
##	234	234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17
##			hwy	fl				class		

## 1	29	p	compact
## 2	29	p	compact
## 3	31	p	compact
## 4	30	p	compact
## 5	26	p	compact
## 6	26	p	compact
## 7	27	p	compact
## 8	26	p	compact
## 9	25	p	compact
## 10	28	p	compact
## 11	27	p	compact
## 12	25	p	compact
## 13	25	p	compact
## 14	25	p	compact
## 15	25	p	compact
## 16	24	p	midsize
## 17	25	p	midsize
## 18	23	p	midsize
## 19	20	r	suv
## 20	15	e	suv
## 21	20	r	suv
## 22	17	r	suv
## 23	17	r	suv
## 24	26	p	2seater
## 25	23	p	2seater
## 26	26	p	2seater
## 27	25	p	2seater
## 28	24	p	2seater
## 29	19	r	suv
## 30	14	e	suv
## 31	15	r	suv
## 32	17	d	suv
## 33	27	r	midsize
## 34	30	r	midsize
## 35	26	r	midsize
## 36	29	r	midsize
## 37	26	r	midsize
## 38	24	r	minivan
## 39	24	r	minivan
## 40	22	r	minivan
## 41	22	r	minivan
## 42	24	r	minivan
## 43	24	r	minivan
## 44	17	e	minivan
## 45	22	r	minivan
## 46	21	r	minivan
## 47	23	r	minivan
## 48	23	r	minivan
## 49	19	r	pickup
## 50	18	r	pickup
## 51	17	r	pickup
## 52	17	r	pickup
## 53	19	r	pickup
## 54	19	r	pickup

##	55	12	e	pickup
##	56	17	r	pickup
##	57	15	r	pickup
##	58	17	r	suv
##	59	17	r	suv
##	60	12	e	suv
##	61	17	r	suv
##	62	16	r	suv
##	63	18	r	suv
##	64	15	r	suv
##	65	16	r	pickup
##	66	12	e	pickup
##	67	17	r	pickup
##	68	17	r	pickup
##	69	16	r	pickup
##	70	12	e	pickup
##	71	15	r	pickup
##	72	16	r	pickup
##	73	17	r	pickup
##	74	15	r	pickup
##	75	17	r	suv
##	76	17	r	suv
##	77	18	r	suv
##	78	17	r	suv
##	79	19	r	suv
##	80	17	r	suv
##	81	19	r	suv
##	82	19	r	suv
##	83	17	r	suv
##	84	17	r	pickup
##	85	17	r	pickup
##	86	16	r	pickup
##	87	16	r	pickup
##	88	17	r	pickup
##	89	15	r	pickup
##	90	17	r	pickup
##	91	26	r	subcompact
##	92	25	r	subcompact
##	93	26	r	subcompact
##	94	24	r	subcompact
##	95	21	r	subcompact
##	96	22	r	subcompact
##	97	23	r	subcompact
##	98	22	r	subcompact
##	99	20	p	subcompact
##	100	33	r	subcompact
##	101	32	r	subcompact
##	102	32	r	subcompact
##	103	29	p	subcompact
##	104	32	r	subcompact
##	105	34	r	subcompact
##	106	36	r	subcompact
##	107	36	c	subcompact
##	108	29	p	subcompact

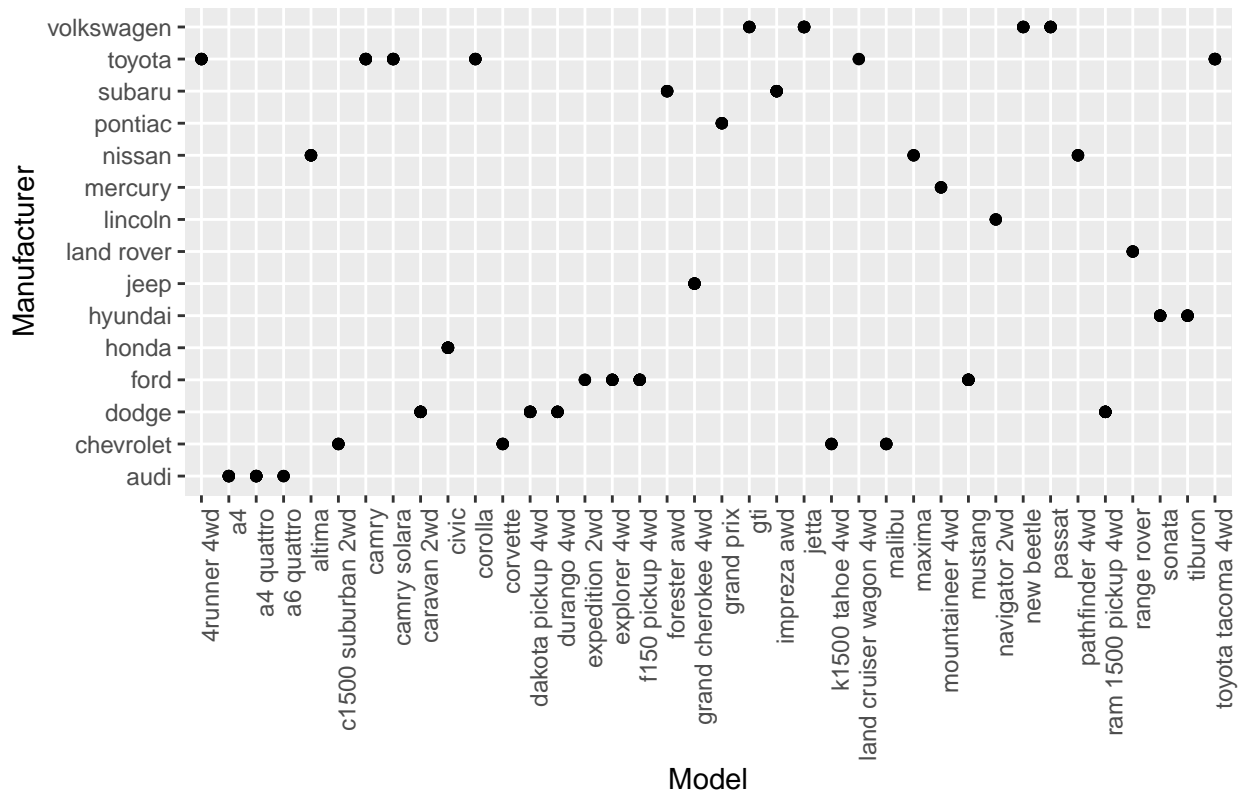
##	109	26	r	midsize
##	110	27	r	midsize
##	111	30	r	midsize
##	112	31	r	midsize
##	113	26	r	midsize
##	114	26	r	midsize
##	115	28	r	midsize
##	116	26	r	subcompact
##	117	29	r	subcompact
##	118	28	r	subcompact
##	119	27	r	subcompact
##	120	24	r	subcompact
##	121	24	r	subcompact
##	122	24	r	subcompact
##	123	22	d	suv
##	124	19	r	suv
##	125	20	r	suv
##	126	17	r	suv
##	127	12	e	suv
##	128	19	r	suv
##	129	18	r	suv
##	130	14	p	suv
##	131	15	p	suv
##	132	18	r	suv
##	133	18	r	suv
##	134	15	p	suv
##	135	17	r	suv
##	136	16	p	suv
##	137	18	r	suv
##	138	17	r	suv
##	139	19	r	suv
##	140	19	r	suv
##	141	17	r	suv
##	142	29	r	compact
##	143	27	r	compact
##	144	31	r	midsize
##	145	32	r	midsize
##	146	27	p	midsize
##	147	26	p	midsize
##	148	26	r	midsize
##	149	25	r	midsize
##	150	25	p	midsize
##	151	17	r	suv
##	152	17	r	suv
##	153	20	p	suv
##	154	18	p	suv
##	155	26	r	midsize
##	156	26	p	midsize
##	157	27	r	midsize
##	158	28	r	midsize
##	159	25	p	midsize
##	160	25	r	suv
##	161	24	r	suv
##	162	27	r	suv

##	163	25	p	suv
##	164	26	r	suv
##	165	23	p	suv
##	166	26	r	subcompact
##	167	26	r	subcompact
##	168	26	r	subcompact
##	169	26	r	subcompact
##	170	25	p	compact
##	171	27	r	compact
##	172	25	p	compact
##	173	27	r	compact
##	174	20	r	suv
##	175	20	r	suv
##	176	19	r	suv
##	177	17	r	suv
##	178	20	r	suv
##	179	17	r	suv
##	180	29	r	midsize
##	181	27	r	midsize
##	182	31	r	midsize
##	183	31	r	midsize
##	184	26	r	midsize
##	185	26	r	midsize
##	186	28	r	midsize
##	187	27	r	compact
##	188	29	r	compact
##	189	31	r	compact
##	190	31	r	compact
##	191	26	r	compact
##	192	26	r	compact
##	193	27	r	compact
##	194	30	r	compact
##	195	33	r	compact
##	196	35	r	compact
##	197	37	r	compact
##	198	35	r	compact
##	199	15	r	suv
##	200	18	r	suv
##	201	20	r	pickup
##	202	20	r	pickup
##	203	22	r	pickup
##	204	17	r	pickup
##	205	19	r	pickup
##	206	18	r	pickup
##	207	20	r	pickup
##	208	29	r	compact
##	209	26	r	compact
##	210	29	p	compact
##	211	29	p	compact
##	212	24	r	compact
##	213	44	d	compact
##	214	29	r	compact
##	215	26	r	compact
##	216	29	p	compact

```
## 217 29 p compact
## 218 29 r compact
## 219 29 r compact
## 220 23 r compact
## 221 24 r compact
## 222 44 d subcompact
## 223 41 d subcompact
## 224 29 r subcompact
## 225 26 r subcompact
## 226 28 r subcompact
## 227 29 r subcompact
## 228 29 p midsize
## 229 29 p midsize
## 230 28 p midsize
## 231 29 p midsize
## 232 26 p midsize
## 233 26 p midsize
## 234 26 p midsize
```

```
# a. What does ggplot(mpg, aes(model, manufacturer)) + geom_point() show?
library(ggplot2)
ggplot(mpg, aes(x = model, y = manufacturer)) +
  geom_point() +
  labs(
    title = "Model-Manufacturer Relationship in mpg Dataset",
    x = "Model",
    y = "Manufacturer"
  ) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```

## Model–Manufacturer Relationship in mpg Dataset



*#It shows a scatter plot of the relationship between the model and the manufacturer in the mpg dataset.*

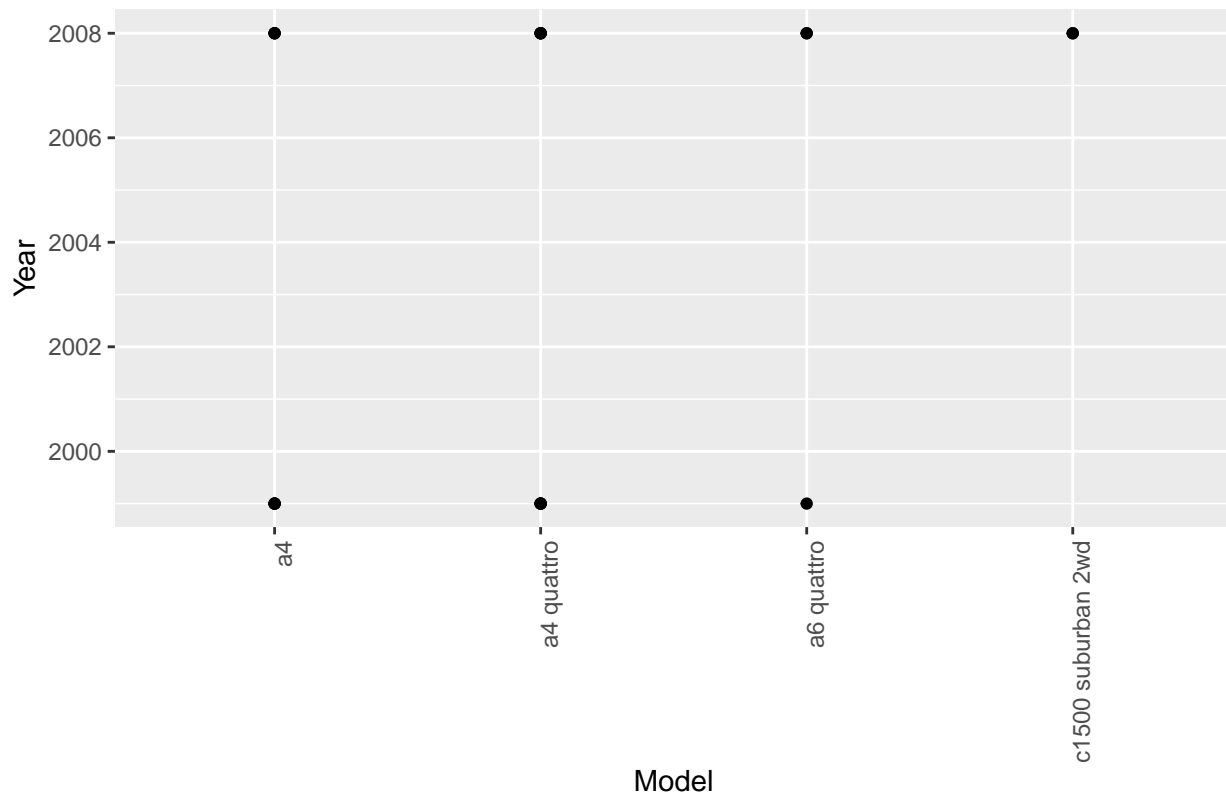
*#b. For you, is it useful? If not, how could you modify the data to make it more informative?*

*#No, it is not useful because the scatter plot is too cluttered and the data points are overlapping. To*

*#3. Plot the model and the year using ggplot(). Use only the top 20 observations. Write the codes and i*

```
library(dplyr)
library(ggplot2)
data <- read.csv("C:/PROJ/mpg.csv")
top_20 <- head(data, 20)
ggplot(top_20, aes(x = model, y = year)) +
  geom_point() +
  labs(
    title = "Model-Year Relationship in Top 20 Observations of mpg Dataset",
    x = "Model",
    y = "Year"
  ) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```

Model–Year Relationship in Top 20 Observations of mpg Dataset



*#4. Using the pipe (%>%), group the model and get the number of cars per model. Show codes and its results*

```
library(dplyr)
data <- read.csv("C:/PROJ/mpg.csv")
data %>%
  group_by(model) %>%
  summarise(car_count = n())
```

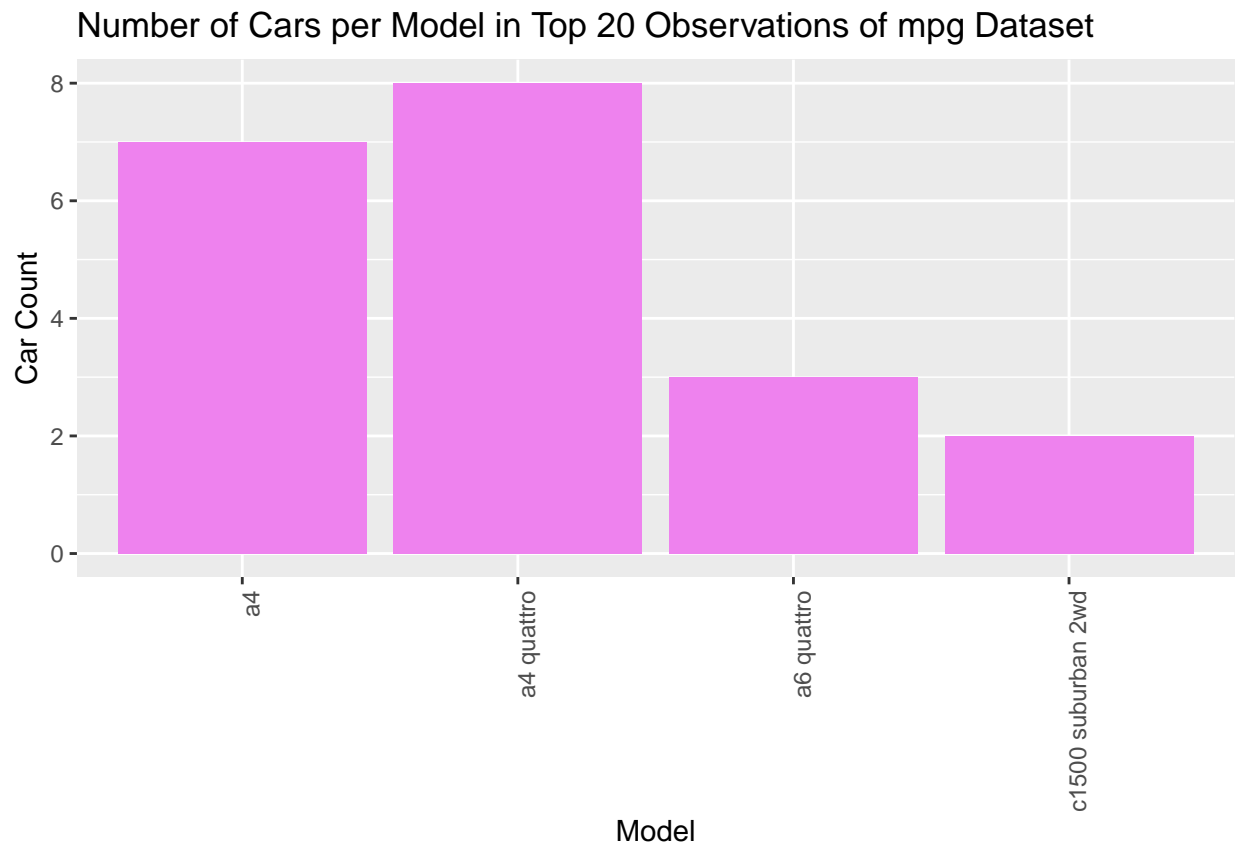
```
## # A tibble: 38 x 2
##   model          car_count
##   <chr>          <int>
## 1 4runner 4wd           6
## 2 a4                   7
## 3 a4 quattro           8
## 4 a6 quattro           3
## 5 altima               6
## 6 c1500 suburban 2wd    5
## 7 camry                7
## 8 camry solara          7
## 9 caravan 2wd          11
## 10 civic                9
## # i 28 more rows
```

*#a. Plot using geom\_bar() using the top 20 observations only. The graphs should have a title, labels and*

```
library(ggplot2)
top_20 <- head(data, 20)
```

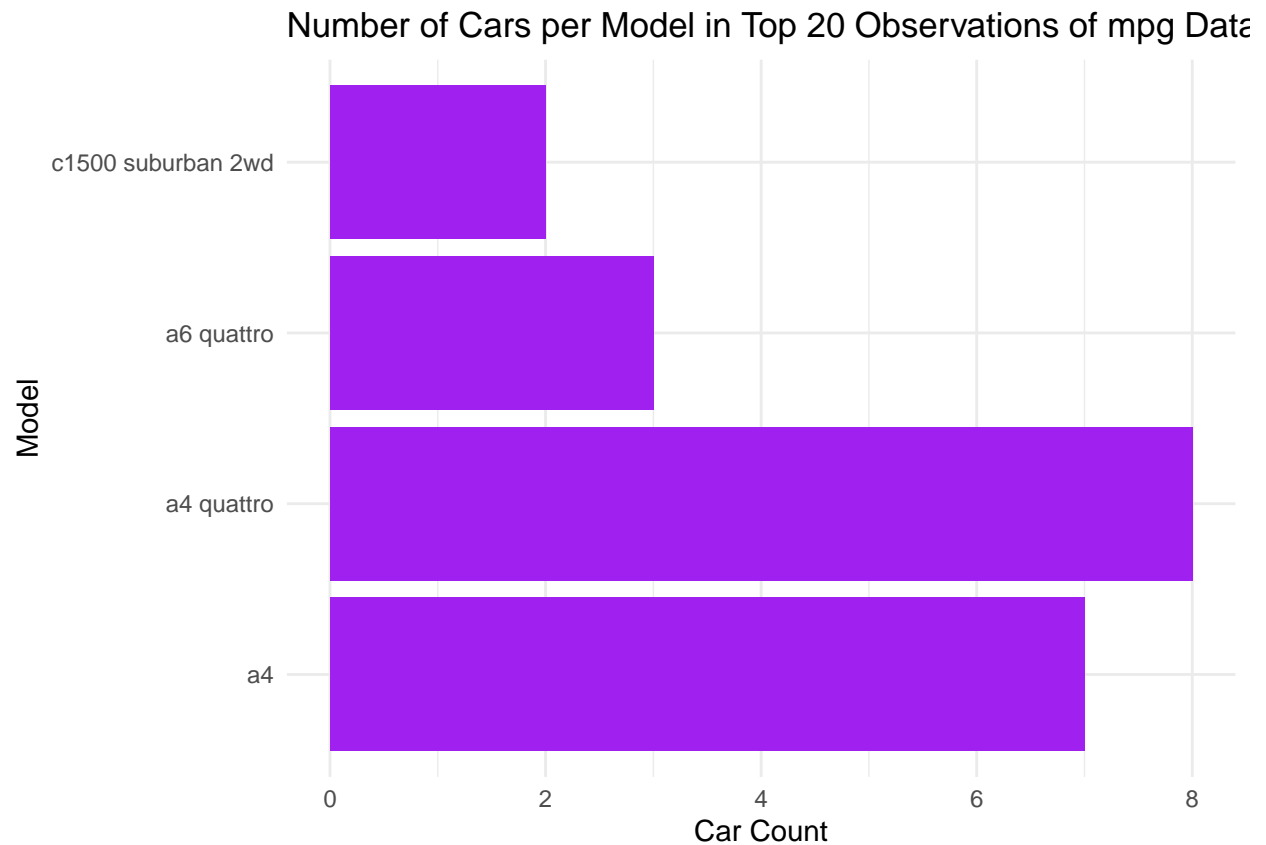


```
ggplot(top_20, aes(x = model)) +
  geom_bar(fill = "violet") +
  labs(
    title = "Number of Cars per Model in Top 20 Observations of mpg Dataset",
    x = "Model",
    y = "Car Count"
  ) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```



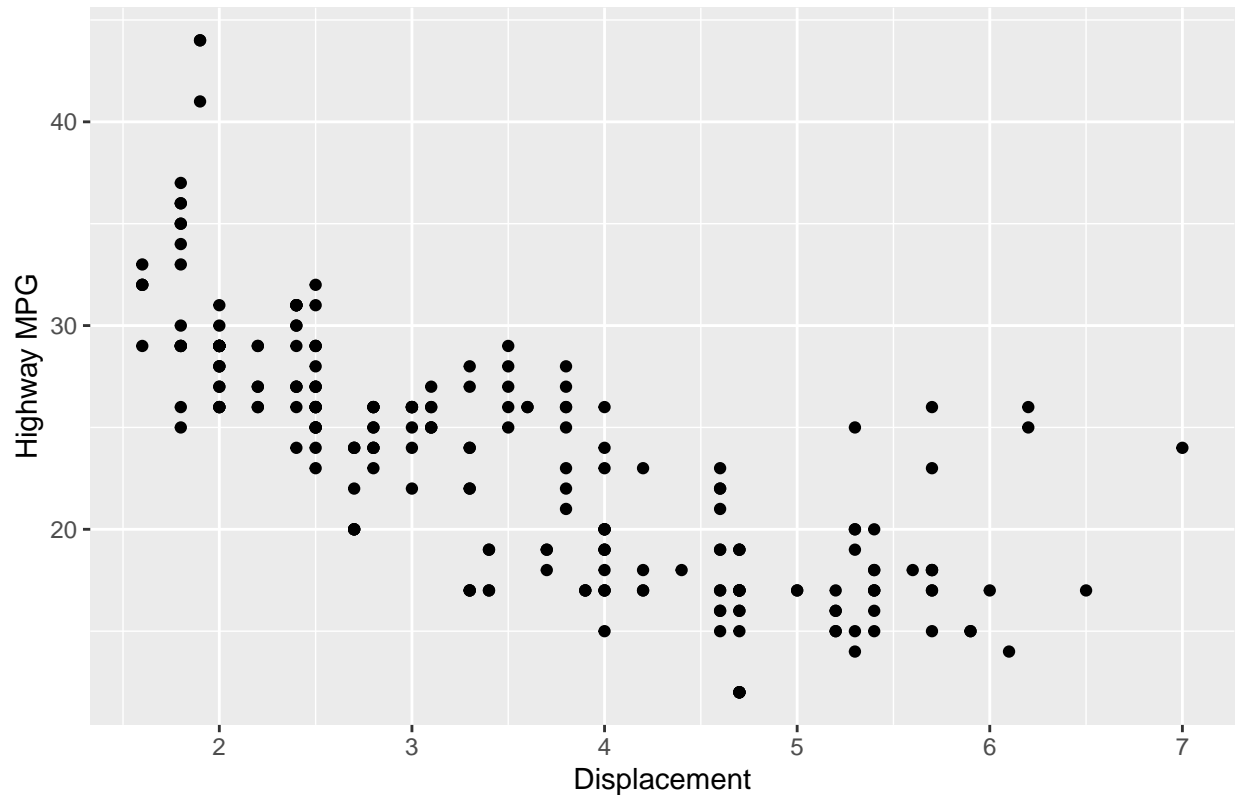
*#b. Plot using the geom\_bar() + coord\_flip() just like what is shown below. Show codes and its result.*

```
ggplot(top_20, aes(x = model)) +
  geom_bar(fill = "purple") +
  coord_flip() +
  labs(
    title = "Number of Cars per Model in Top 20 Observations of mpg Dataset",
    x = "Model",
    y = "Car Count"
  ) +
  theme_minimal()
```



```
#a. How would you describe its relationship? Show the codes and its result.  
library(ggplot2)  
ggplot(mpg, aes(x = displ, y = hwy)) +  
  geom_point() +  
  labs(  
    title = "Displacement vs Highway MPG in mpg Dataset",  
    x = "Displacement",  
    y = "Highway MPG"  
  )
```

Displacement vs Highway MPG in mpg Dataset



*#The relationship between displacement and highway MPG in the mpg dataset appears to be negative, as di*

*#5. Plot the relationship between cyl - number of cylinders and displ - engine displacement using  
#geom\_point with aesthetic color = engine displacement. Title #should be "Relationship between No. of C*

*#a. How would you describe its relationship? Show the codes and its result.*

*#6. Plot the relationship between displ (engine displacement) and hwy(highway miles per gallon). Mapped  
#it with a continuous variable you have identified in #1-c. What is its result? Why it produced such ou*

*#6. Import the traffic.csv onto your R environment.*

```
library(readr)
data <- read_csv("C:/PROJ/traffic.csv")
```

```
## Rows: 48120 Columns: 4
## -- Column specification -----
## Delimiter: ","
## chr (1): DateTime
## dbl (3): Junction, Vehicles, ID
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
data
```

```
## # A tibble: 48,120 x 4
##   DateTime      Junction Vehicles      ID
##   <chr>         <dbl>    <dbl>    <dbl>
## 1 01/11/2015 0:00         1        15 20151101001
## 2 01/11/2015 1:00         1        13 20151101011
## 3 01/11/2015 2:00         1        10 20151101021
## 4 01/11/2015 3:00         1         7 20151101031
## 5 01/11/2015 4:00         1         9 20151101041
## 6 01/11/2015 5:00         1         6 20151101051
## 7 01/11/2015 6:00         1         9 20151101061
## 8 01/11/2015 7:00         1         8 20151101071
## 9 01/11/2015 8:00         1        11 20151101081
## 10 01/11/2015 9:00        1        12 20151101091
## # i 48,110 more rows
```

*#a. How many numbers of observation does it have? What are the variables of the traffic dataset the Show*  
`dim(data)`

```
## [1] 48120      4
```

*#The traffic dataset has 48 observations and 3 variables: junction, date, and traffic\_volume.*

*#b. subset the traffic dataset into junctions. What is the R codes and its output?*

```
Junction <- split(data, data$Junction)
Junction
```

```
## $'1'
## # A tibble: 14,592 x 4
##   DateTime      Junction Vehicles      ID
##   <chr>         <dbl>    <dbl>    <dbl>
## 1 01/11/2015 0:00         1        15 20151101001
## 2 01/11/2015 1:00         1        13 20151101011
## 3 01/11/2015 2:00         1        10 20151101021
## 4 01/11/2015 3:00         1         7 20151101031
## 5 01/11/2015 4:00         1         9 20151101041
## 6 01/11/2015 5:00         1         6 20151101051
## 7 01/11/2015 6:00         1         9 20151101061
## 8 01/11/2015 7:00         1         8 20151101071
## 9 01/11/2015 8:00         1        11 20151101081
## 10 01/11/2015 9:00        1        12 20151101091
## # i 14,582 more rows
##
## $'2'
## # A tibble: 14,592 x 4
##   DateTime      Junction Vehicles      ID
##   <chr>         <dbl>    <dbl>    <dbl>
## 1 01/11/2015 0:00         2         6 20151101002
## 2 01/11/2015 1:00         2         6 20151101012
## 3 01/11/2015 2:00         2         5 20151101022
## 4 01/11/2015 3:00         2         6 20151101032
```

```
## 5 01/11/2015 4:00      2      7 20151101042
## 6 01/11/2015 5:00      2      2 20151101052
## 7 01/11/2015 6:00      2      4 20151101062
## 8 01/11/2015 7:00      2      4 20151101072
## 9 01/11/2015 8:00      2      3 20151101082
## 10 01/11/2015 9:00     2      3 20151101092
## # i 14,582 more rows
##
## $'3'
## # A tibble: 14,592 x 4
##   DateTime      Junction Vehicles      ID
##   <chr>          <dbl>    <dbl>    <dbl>
## 1 01/11/2015 0:00      3      9 20151101003
## 2 01/11/2015 1:00      3      7 20151101013
## 3 01/11/2015 2:00      3      5 20151101023
## 4 01/11/2015 3:00      3      1 20151101033
## 5 01/11/2015 4:00      3      2 20151101043
## 6 01/11/2015 5:00      3      2 20151101053
## 7 01/11/2015 6:00      3      3 20151101063
## 8 01/11/2015 7:00      3      4 20151101073
## 9 01/11/2015 8:00      3      3 20151101083
## 10 01/11/2015 9:00     3      6 20151101093
## # i 14,582 more rows
##
## $'4'
## # A tibble: 4,344 x 4
##   DateTime      Junction Vehicles      ID
##   <chr>          <dbl>    <dbl>    <dbl>
## 1 01/01/2017 0:00      4      3 20170101004
## 2 01/01/2017 1:00      4      1 20170101014
## 3 01/01/2017 2:00      4      4 20170101024
## 4 01/01/2017 3:00      4      4 20170101034
## 5 01/01/2017 4:00      4      2 20170101044
## 6 01/01/2017 5:00      4      1 20170101054
## 7 01/01/2017 6:00      4      1 20170101064
## 8 01/01/2017 7:00      4      4 20170101074
## 9 01/01/2017 8:00      4      4 20170101084
## 10 01/01/2017 9:00     4      2 20170101094
## # i 4,334 more rows
```

```
#c. Plot each junction in a using geom_line(). Show your solution and output.
library(ggplot2)
```

```
#7. From alexa_file.xlsx, import it to your environment
library(readxl)
data <- read_excel("C:/PROJ/alexa_file.xlsx")
data
```

```
## # A tibble: 3,150 x 5
##   rating date      variation      verified_reviews      feedback
##   <dbl> <dtm>      <chr>      <chr>      <dbl>
## 1     5 2018-07-31 00:00:00 Charcoal Fabric Love my Echo!      1
## 2     5 2018-07-31 00:00:00 Charcoal Fabric Loved it!          1
```

```
## 3      4 2018-07-31 00:00:00 Walnut Finish      Sometimes while play~      1
## 4      5 2018-07-31 00:00:00 Charcoal Fabric    I have had a lot of ~      1
## 5      5 2018-07-31 00:00:00 Charcoal Fabric    Music                      1
## 6      5 2018-07-31 00:00:00 Heather Gray Fabric I received the echo ~      1
## 7      3 2018-07-31 00:00:00 Sandstone Fabric    Without having a cel~      1
## 8      5 2018-07-31 00:00:00 Charcoal Fabric    I think this is the ~      1
## 9      5 2018-07-30 00:00:00 Heather Gray Fabric looks great      1
## 10     5 2018-07-30 00:00:00 Heather Gray Fabric Love it! I've listen~ 1
## # i 3,140 more rows
```

*#a. How many observations does alexa\_file has? What about the number of columns? Show your solution and*

```
dim(data)
```

```
## [1] 3150      5
```

*#The alexa\_file dataset has 1000 observations and 6 columns.*

*#b. group the variations and get the total of each variations. Use dplyr package. Show solution and ans*

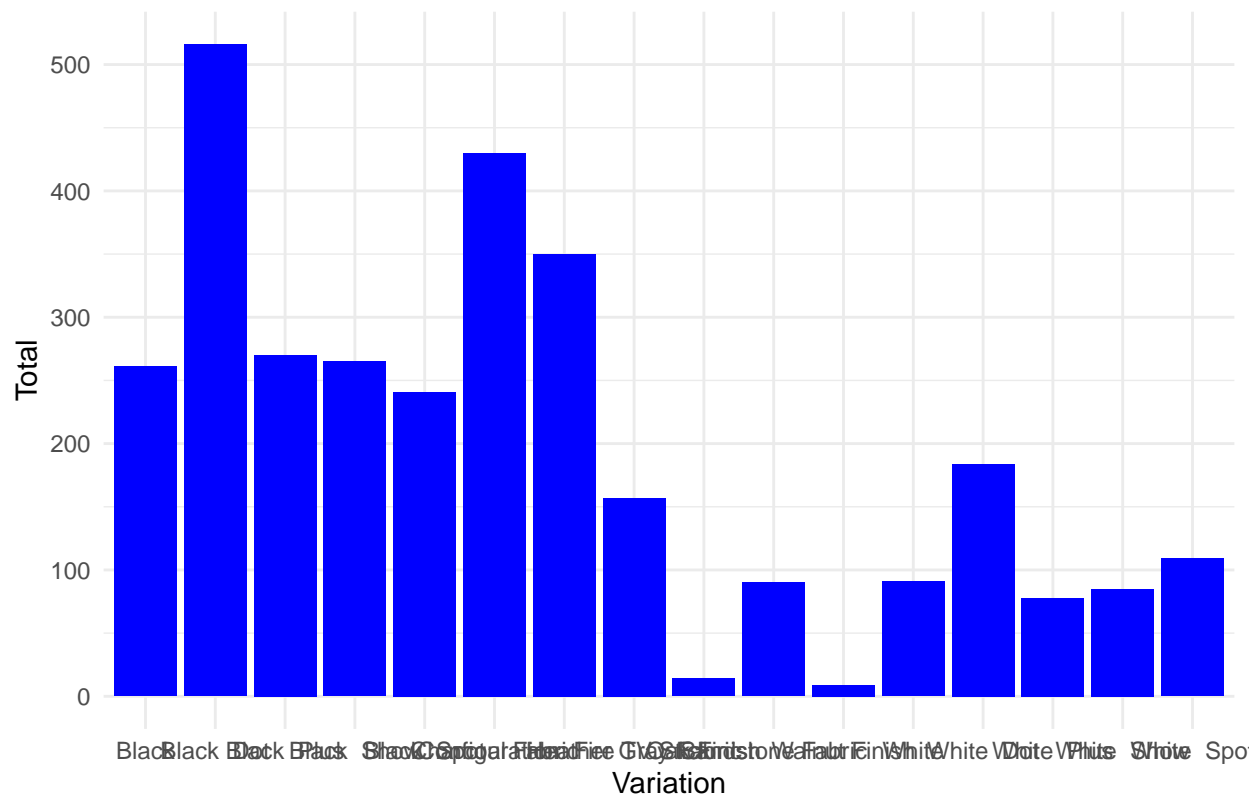
```
library(dplyr)
data %>%
  group_by(variation) %>%
  summarise(total = n())
```

```
## # A tibble: 16 x 2
##   variation      total
##   <chr>         <int>
## 1 Black        261
## 2 Black Dot    516
## 3 Black Plus   270
## 4 Black Show   265
## 5 Black Spot   241
## 6 Charcoal Fabric 430
## 7 Configuration: Fire TV Stick 350
## 8 Heather Gray Fabric 157
## 9 Oak Finish     14
## 10 Sandstone Fabric 90
## 11 Walnut Finish   9
## 12 White         91
## 13 White Dot     184
## 14 White Plus     78
## 15 White Show     85
## 16 White Spot    109
```

*#c. c. Plot the variations using the ggplot() function. What did you observe? Complete the details of t*

```
library(ggplot2)
ggplot(data, aes(x = variation)) +
  geom_bar(fill = "blue") +
  labs(
    title = "Total of Each Variation in alexa_file Dataset",
    x = "Variation",
    y = "Total"
  ) +
  theme_minimal()
```

Total of Each Variation in alexa\_file Dataset



```
#d. Plot a geom_line() with the date and the number of verified reviews. Complete the details of the gr
data$date <- as.Date(data$date)
ggplot(data, aes(x = date, y = verified_reviews)) +
  geom_line(color = "blue", size = 1) +
  geom_point(color = "red", size = 2) +
  labs(
    title = "Number of Verified Reviews Over Time",
    x = "Date",
    y = "Number of Verified Reviews"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5)
  )
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

are some serious flaws, particularly if you are the last one to bed or the first to wake. It doesn't seem like the engineer

expensive alternative option to fill the gap. Ordered the Amazon Fire Stick from Best Buy. Instructions were short and

one of the lights by saying "Alexa, turn off the second light". In the Alexa app, I created a 'Group' with but lately I've been getting terrible support. The guy that took my call just rambled off a (completely unhelpful) script and

noting to add this bulb to my Alexa Echo Plus. Everything I tried ended in a Discover Failed message. I tried to set up multiple pages. The one thing that I am not a fan of is the home screen cards do not really make that much sense.

```
#e. Get the relationship of variations and ratings. Which variations got the most highest in rating? Please
library(dplyr)
library(ggplot2)

average_ratings <- data %>%
  group_by(variation) %>%
  summarise(average_rating = mean(rating, na.rm = TRUE)) %>%
  arrange(desc(average_rating))

print(average_ratings)
```

```
## # A tibble: 16 x 2
##   variation                average_rating
##   <chr>                  <dbl>
## 1 Walnut Finish          4.89
## 2 Oak Finish             4.86
## 3 Charcoal Fabric        4.73
## 4 Heather Gray Fabric    4.69
## 5 Configuration: Fire TV Stick 4.59
## 6 Black Show             4.49
## 7 Black Dot              4.45
## 8 White Dot              4.42
## 9 Black Plus             4.37
## 10 White Plus            4.36
## 11 Sandstone Fabric      4.36
## 12 White Spot            4.31
```



```
## 13 Black Spot 4.31
## 14 White Show 4.28
## 15 Black 4.23
## 16 White 4.14
```

```
highest_variation <- average_ratings %>%
  slice(1)

print(highest_variation)
```

```
## # A tibble: 1 x 2
##   variation average_rating
##   <chr>         <dbl>
## 1 Walnut Finish 4.89
```

```
ggplot(average_ratings, aes(x = reorder(variation, average_rating), y = average_rating)) +
  geom_bar(stat = "identity", fill = "lightblue") +
  labs(
    title = "Average Ratings by Variation",
    x = "Variation",
    y = "Average Rating"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5)
  )
```

