

1. Use the dataset mpg

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

```
mpgdata <- read.csv("mpg.csv")
mpgdata
```

##	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(l4)	r	11
## 21	21	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14
## 22	22	chevrolet	c1500 suburban 2wd	5.7	1999	8	auto(l4)	r	13
## 23	23	chevrolet	c1500 suburban 2wd	6.0	2008	8	auto(l4)	r	12
## 24	24	chevrolet	corvette	5.7	1999	8	manual(m6)	r	16
## 25	25	chevrolet	corvette	5.7	1999	8	auto(l4)	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(l4)	4	14
## 30	30	chevrolet	k1500 tahoe 4wd	5.3	2008	8	auto(l4)	4	11
## 31	31	chevrolet	k1500 tahoe 4wd	5.7	1999	8	auto(l4)	4	11
## 32	32	chevrolet	k1500 tahoe 4wd	6.5	1999	8	auto(l4)	4	14
## 33	33	chevrolet	malibu	2.4	1999	4	auto(l4)	f	19
## 34	34	chevrolet	malibu	2.4	2008	4	auto(l4)	f	22
## 35	35	chevrolet	malibu	3.1	1999	6	auto(l4)	f	18
## 36	36	chevrolet	malibu	3.5	2008	6	auto(l4)	f	18
## 37	37	chevrolet	malibu	3.6	2008	6	auto(s6)	f	17
## 38	38	dodge	caravan 2wd	2.4	1999	4	auto(l3)	f	18
## 39	39	dodge	caravan 2wd	3.0	1999	6	auto(l4)	f	17
## 40	40	dodge	caravan 2wd	3.3	1999	6	auto(l4)	f	16
## 41	41	dodge	caravan 2wd	3.3	1999	6	auto(l4)	f	16
## 42	42	dodge	caravan 2wd	3.3	2008	6	auto(l4)	f	17
## 43	43	dodge	caravan 2wd	3.3	2008	6	auto(l4)	f	17
## 44	44	dodge	caravan 2wd	3.3	2008	6	auto(l4)	f	11
## 45	45	dodge	caravan 2wd	3.8	1999	6	auto(l4)	f	15
## 46	46	dodge	caravan 2wd	3.8	1999	6	auto(l4)	f	15
## 47	47	dodge	caravan 2wd	3.8	2008	6	auto(l6)	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(l4)	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(l4)	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(l4)	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(l4)	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(l5)	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(l5)	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						
##	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

```

b. Which variables from mpg dataset are categorical?

manufacturer - The car manufacturer (e.g., Audi, Chevrolet). - model - trans - drv - fl - class -
manufacturer

c. Which are continuous variables?

- displ
- year
- cyl
- 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

manufacturer_count <- mpgdata %>%
  group_by(manufacturer) %>%
  summarise(model_count = n_distinct(model)) %>%
  arrange(desc(model_count))

most_models_manufacturer <- manufacturer_count[1, ]

model_variation_count <- mpgdata %>%
  group_by(model) %>%
  summarise(variation_count = n()) %>%
  arrange(desc(variation_count))

most_variations_model <- model_variation_count[1, ]

most_models_manufacturer

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

most_variations_model

## # A tibble: 1 x 2
##   model          variation_count
##   <chr>             <int>
## 1 caravan 2wd          11
```

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

```
unique_models <- mpgdata %>%
  group_by(manufacturer) %>%
  summarise(unique_models = list(unique(model))) %>%
  arrange(manufacturer)

print(unique_models)

## # A tibble: 15 x 2
##   manufacturer unique_models
##   <chr>         <list>
## 1 audi         <chr [3]>
## 2 chevrolet   <chr [4]>
## 3 dodge       <chr [4]>
```

```
## 4 ford      <chr [4]>
## 5 honda     <chr [1]>
## 6 hyundai   <chr [2]>
## 7 jeep      <chr [1]>
## 8 land rover <chr [1]>
## 9 lincoln   <chr [1]>
## 10 mercury  <chr [1]>
## 11 nissan    <chr [3]>
## 12 pontiac  <chr [1]>
## 13 subaru   <chr [2]>
## 14 toyota   <chr [6]>
## 15 volkswagen <chr [4]>
```

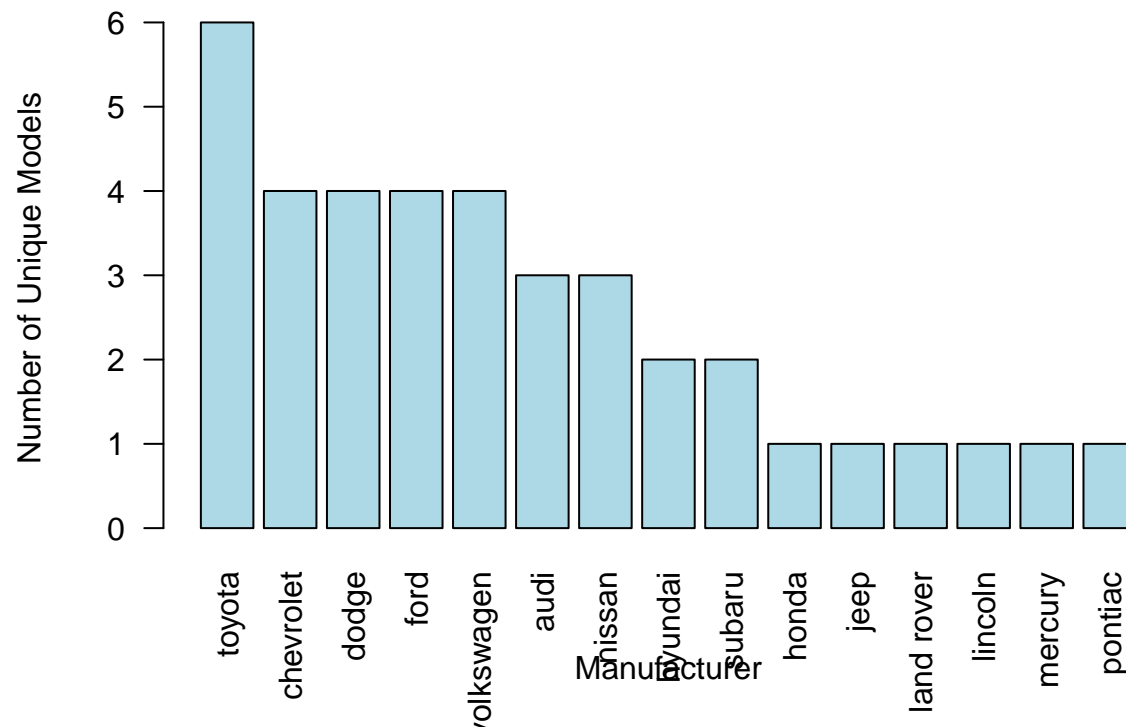
b. Graph the result by using `plot()` and `ggplot()`. Write the codes and its result.

```
model_counts <- mpgdata %>%
  group_by(manufacturer) %>%
  summarise(unique_model_count = n_distinct(model)) %>%
  arrange(desc(unique_model_count))

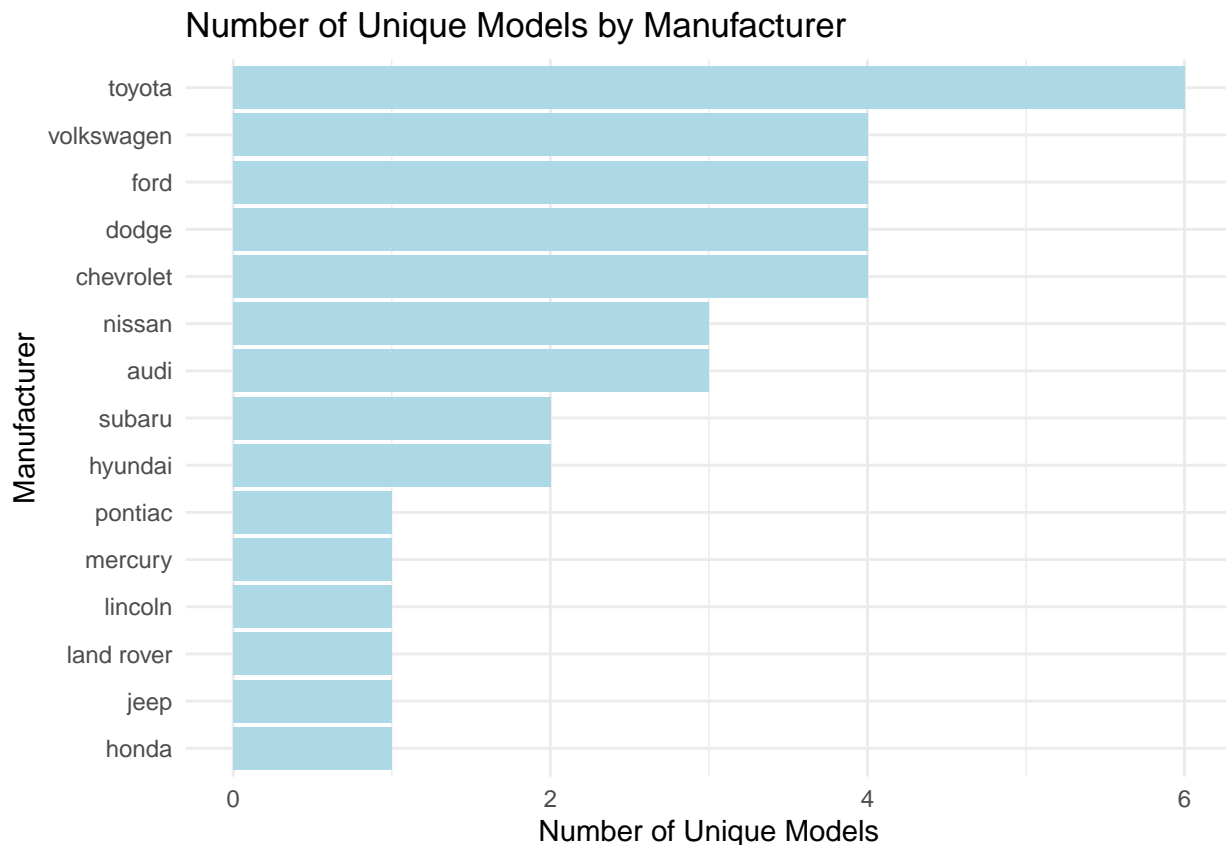
barplot(model_counts$unique_model_count,
        names.arg = model_counts$manufacturer,
        las = 2,
        col = "lightblue",
        main = "Number of Unique Models by Manufacturer",
        xlab = "Manufacturer",
        ylab = "Number of Unique Models")

library(ggplot2)
```

Number of Unique Models by Manufacturer



```
ggplot(model_counts, aes(x = reorder(manufacturer, unique_model_count), y = unique_model_count)) +  
  geom_bar(stat = "identity", fill = "lightblue") +  
  coord_flip() +  
  labs(title = "Number of Unique Models by Manufacturer",  
        x = "Manufacturer",  
        y = "Number of Unique Models") +  
  theme_minimal()
```



2.

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

- a. What does `ggplot(mpg, aes(model, manufacturer)) + geom_point()` show? This code generates a scatter plot using `ggplot2`, where:

x-axis: represents the model. y-axis: represents the manufacturer. However, since both model and manufacturer are categorical variables, the plot will display points for each combination of model and manufacturer. This can lead to a large number of overlapping points, making it difficult to discern any meaningful relationships or patterns.

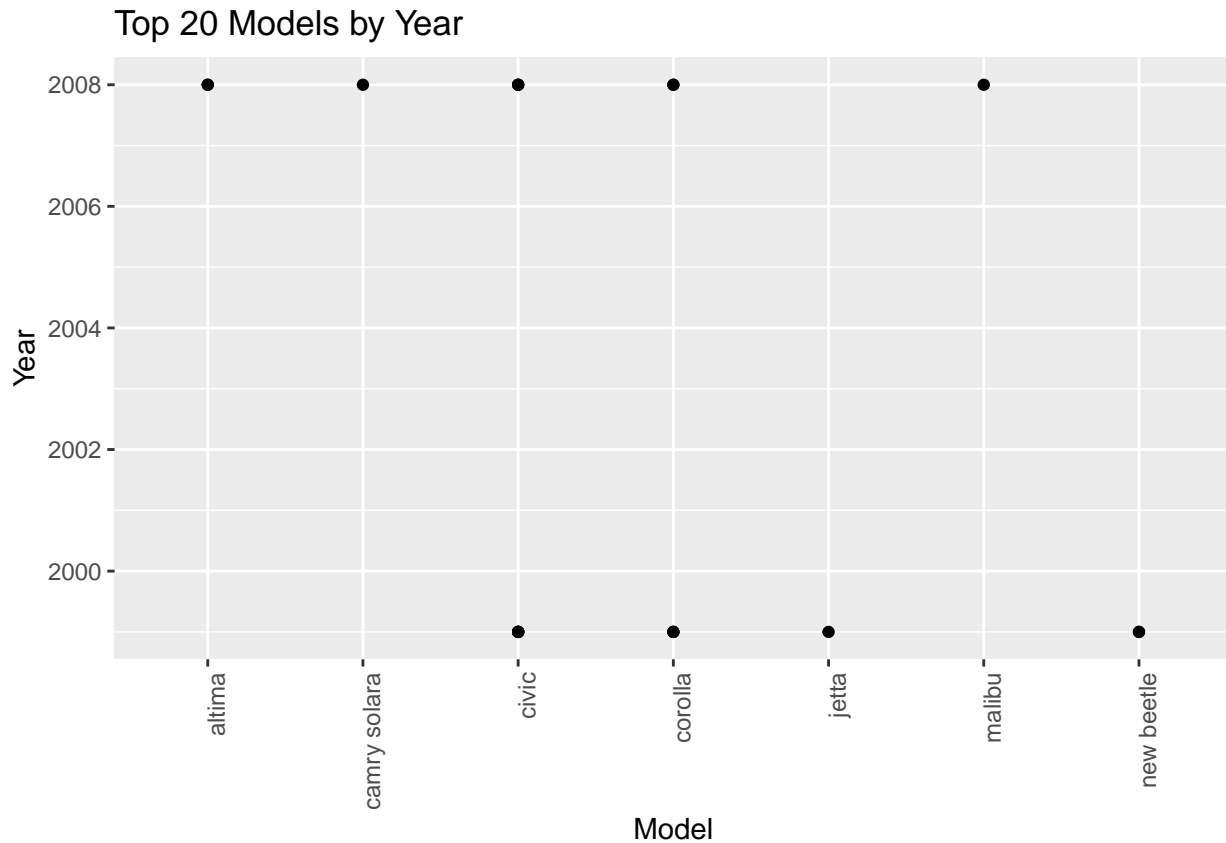
The result will likely look cluttered and may not effectively convey the relationship between the two variables.

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

In its current form, the plot is not very useful for understanding the relationship between model and manufacturer due to the overplotting of points and the lack of numerical or continuous variables that might provide more insights.

3. Plot the model and the year using `ggplot()`. Use only the top 20 observations. Write the codes and its results.

```
library(ggplot2)
library(dplyr)
top_20_mpg <- mpgdata %>%
  arrange(desc(cty)) %>%
  head(20)
ggplot(top_20_mpg, aes(x = model, y = year)) +
  geom_point() +
  labs(title = "Top 20 Models by Year", x = "Model", y = "Year") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```



4.

Using the pipe (`%>%`), group the model and get the number of cars per model. Show codes and its result

```
model_counts <- mpg %>%
  group_by(model) %>%
  summarise(number_of_cars = n()) %>%
  arrange(desc(number_of_cars))

print(model_counts)
```

```
## # A tibble: 38 x 2
##   model                number_of_cars
##   <chr>                  <int>
## 1 caravan 2wd             11
## 2 ram 1500 pickup 4wd     10
## 3 civic                   9
## 4 dakota pickup 4wd       9
## 5 jetta                   9
## 6 mustang                 9
## 7 a4 quattro              8
## 8 grand cherokee 4wd      8
## 9 impreza awd             8
## 10 a4                     7
## # i 28 more rows
```

- a. Plot using `geom_bar()` using the top 20 observations only. The graphs should have a title, labels and colors. Show code and results.

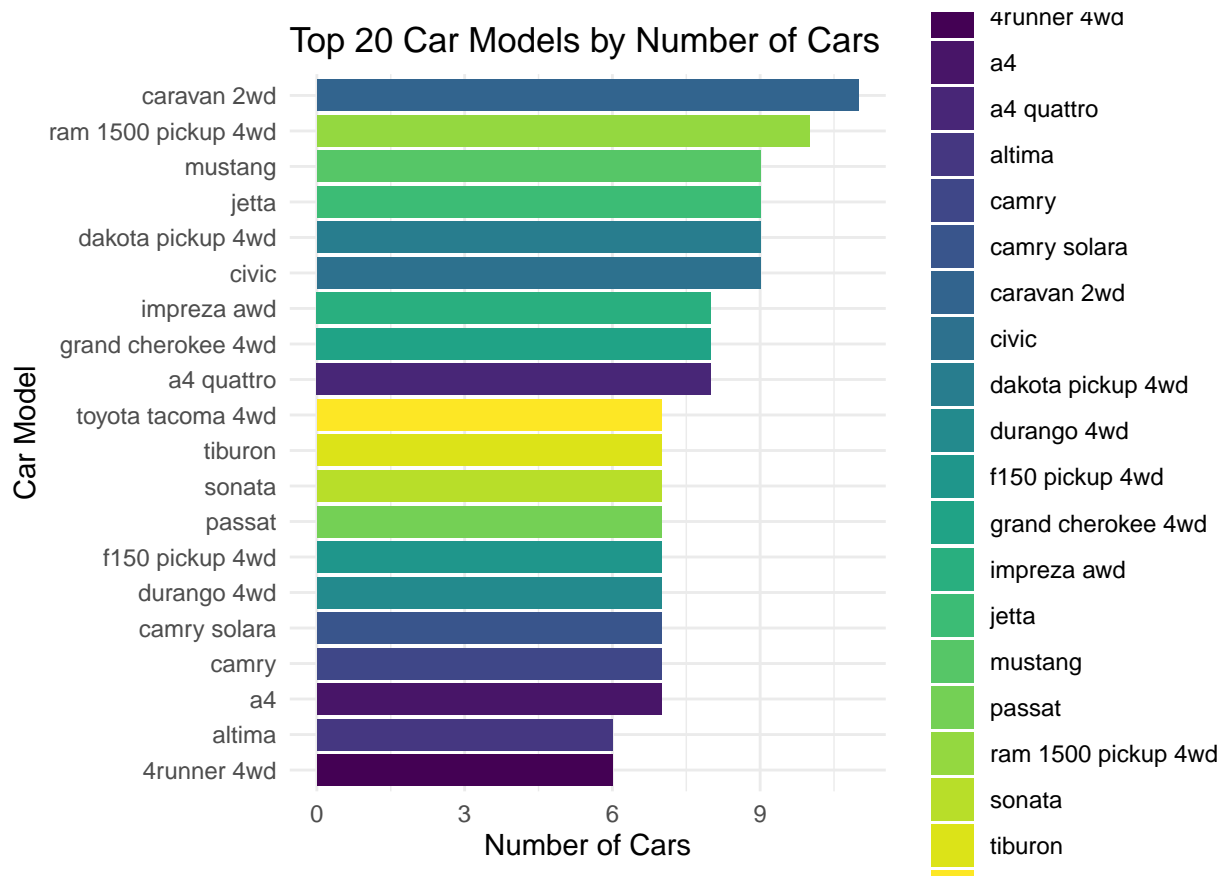
```
top_models <- mpgdata %>%
  group_by(model) %>%
```

```

summarise(number_of_cars = n()) %>%
arrange(desc(number_of_cars)) %>%
slice_head(n = 20)

ggplot(top_models, aes(x = reorder(model, number_of_cars), y = number_of_cars, fill = model)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(title = "Top 20 Car Models by Number of Cars",
       x = "Car Model",
       y = "Number of Cars",
       fill = "Model") +
  theme_minimal() +
  scale_fill_viridis_d()

```



b.

Plot using the `geom_bar()` + `coord_flip()` just like what is shown below. Show codes and its result.

```

library(viridis)

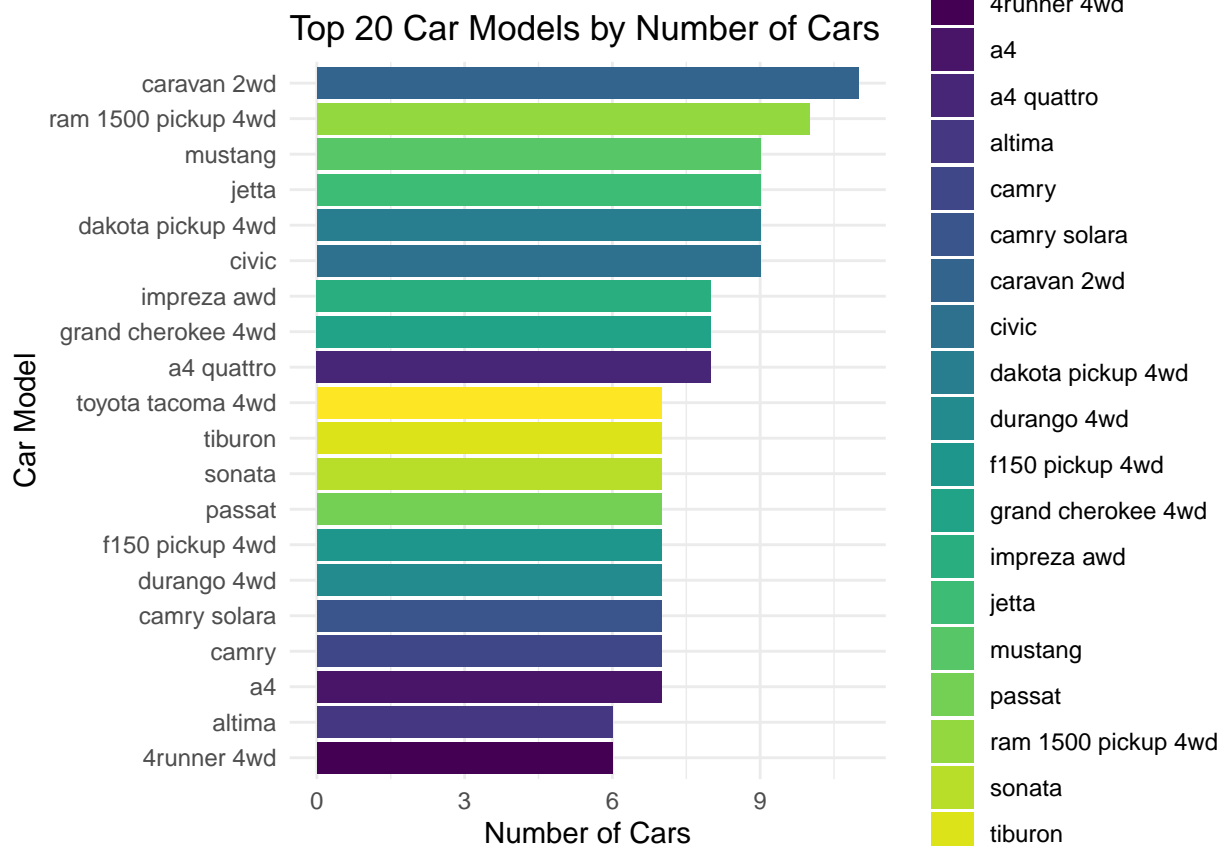
## Loading required package: viridisLite

library(viridisLite)
top_models <- mpgdata %>%
  group_by(model) %>%
  summarise(number_of_cars = n()) %>%
  arrange(desc(number_of_cars)) %>%
  slice_head(n = 20)

ggplot(top_models, aes(x = reorder(model, number_of_cars), y = number_of_cars, fill = model)) +

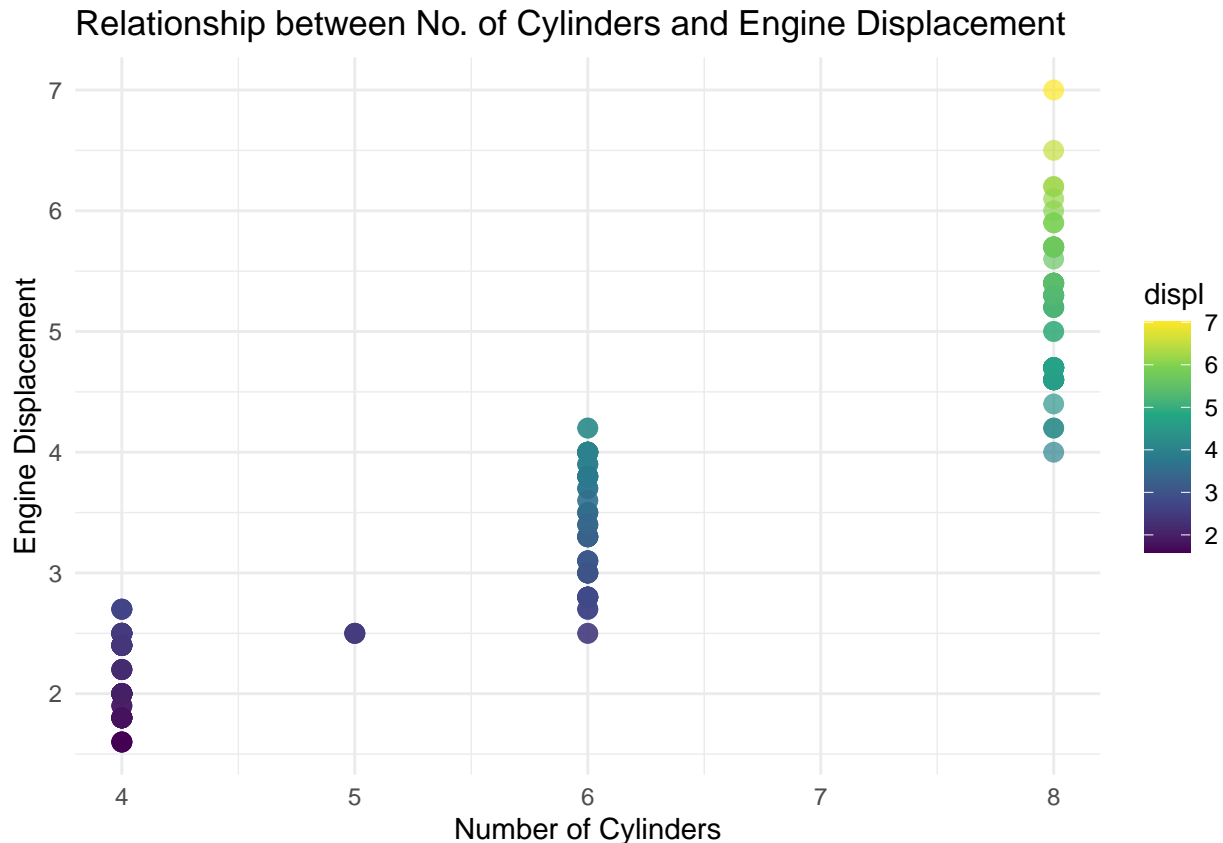
```

```
geom_bar(stat = "identity") +
coord_flip() +
labs(title = "Top 20 Car Models by Number of Cars",
     x = "Car Model",
     y = "Number of Cars") +
theme_minimal() +
scale_fill_viridis_d()
```



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 Cylinders and Engine Displacement".

```
ggplot(mpgdata, aes(x = cyl, y = displ, color = displ)) +
geom_point(size = 3, alpha = 0.7) +
labs(title = "Relationship between No. of Cylinders and Engine Displacement",
     x = "Number of Cylinders",
     y = "Engine Displacement") +
theme_minimal() +
scale_color_viridis_c()
```

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

General Trend: The scatter plot typically shows a positive correlation between the number of cylinders and engine displacement. As the number of cylinders increases, the engine displacement tends to increase as well.

Clustered Data: You'll likely see clusters of points, where vehicles with lower cylinder counts (e.g., 4 cylinders) have relatively lower displacement values, while vehicles with higher cylinder counts (e.g., 6 or 8 cylinders) are associated with larger displacement values.

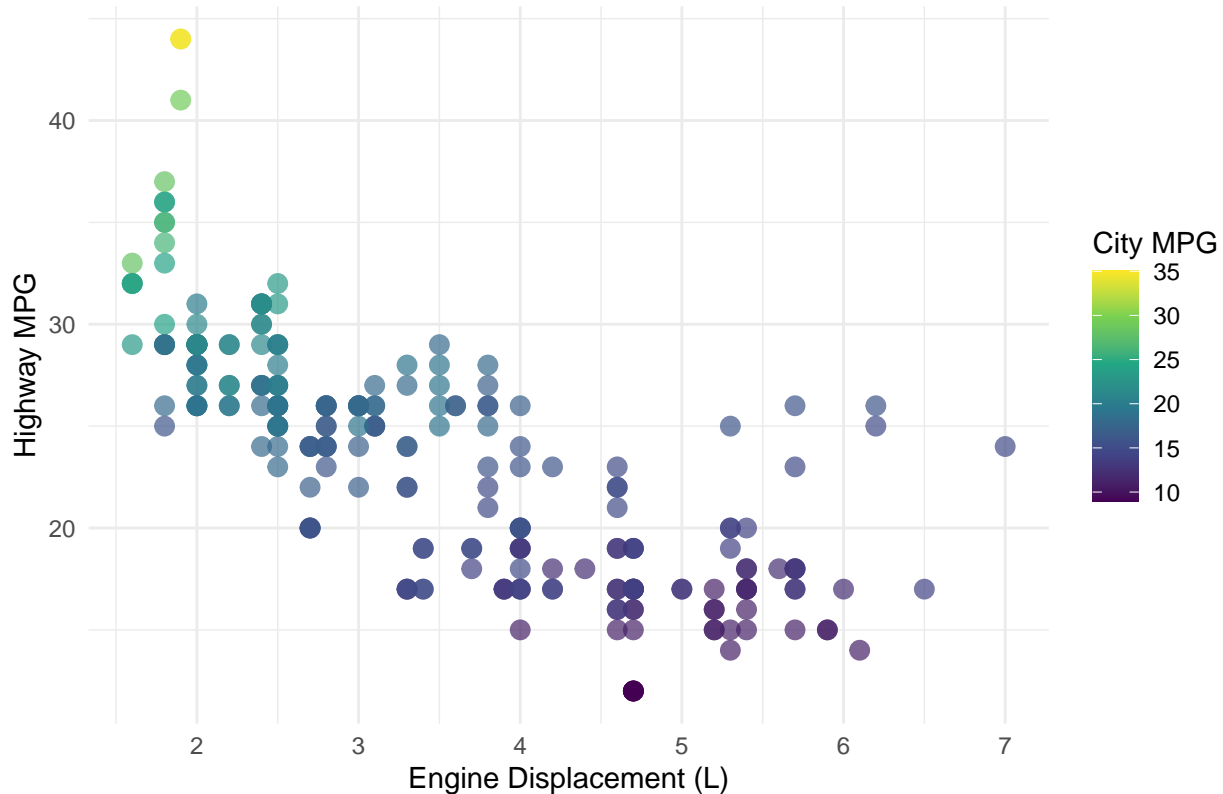
Variability: While the trend is generally positive, there may be some variability, particularly at higher cylinder counts where engine displacement can vary more widely.

Outliers: Pay attention to any outliers that might skew the perception of the relationship. For instance, a vehicle with a high displacement but a low cylinder count could indicate a performance-oriented engine.

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 output?

```
ggplot(mpg, aes(x = displ, y = hwy, color = cty)) +
  geom_point(size = 3, alpha = 0.7) +
  labs(title = "Relationship between Engine Displacement and Highway MPG",
       x = "Engine Displacement (L)",
       y = "Highway MPG",
       color = "City MPG") +
  theme_minimal() +
  scale_color_viridis_c()
```

Relationship between Engine Displacement and Highway MPG



6.

Import the traffic.csv onto your R environment.

```
library(readr)
traffic_data <- read_csv("traffic.csv")

## Rows: 48120 Columns: 4
## -- Column specification -----
## Delimiter: ","
## dbl (3): Junction, Vehicles, ID
## dtm (1): DateTime
##
## 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.
head(traffic_data)

## # A tibble: 6 x 4
##   DateTime          Junction Vehicles      ID
##   <dtm>            <dbl>    <dbl>    <dbl>
## 1 2015-11-01 00:00:00      1      15 20151101001
## 2 2015-11-01 01:00:00      1      13 20151101011
## 3 2015-11-01 02:00:00      1      10 20151101021
## 4 2015-11-01 03:00:00      1       7 20151101031
## 5 2015-11-01 04:00:00      1       9 20151101041
## 6 2015-11-01 05:00:00      1       6 20151101051
```

- How many numbers of observation does it have? What are the variables of the traffic dataset the Show your answer.

```
num_observations <- nrow(traffic_data)
variables <- colnames(traffic_data)
num_observations
```

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

```
variables
```

```
## [1] "DateTime" "Junction" "Vehicles" "ID"
```

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

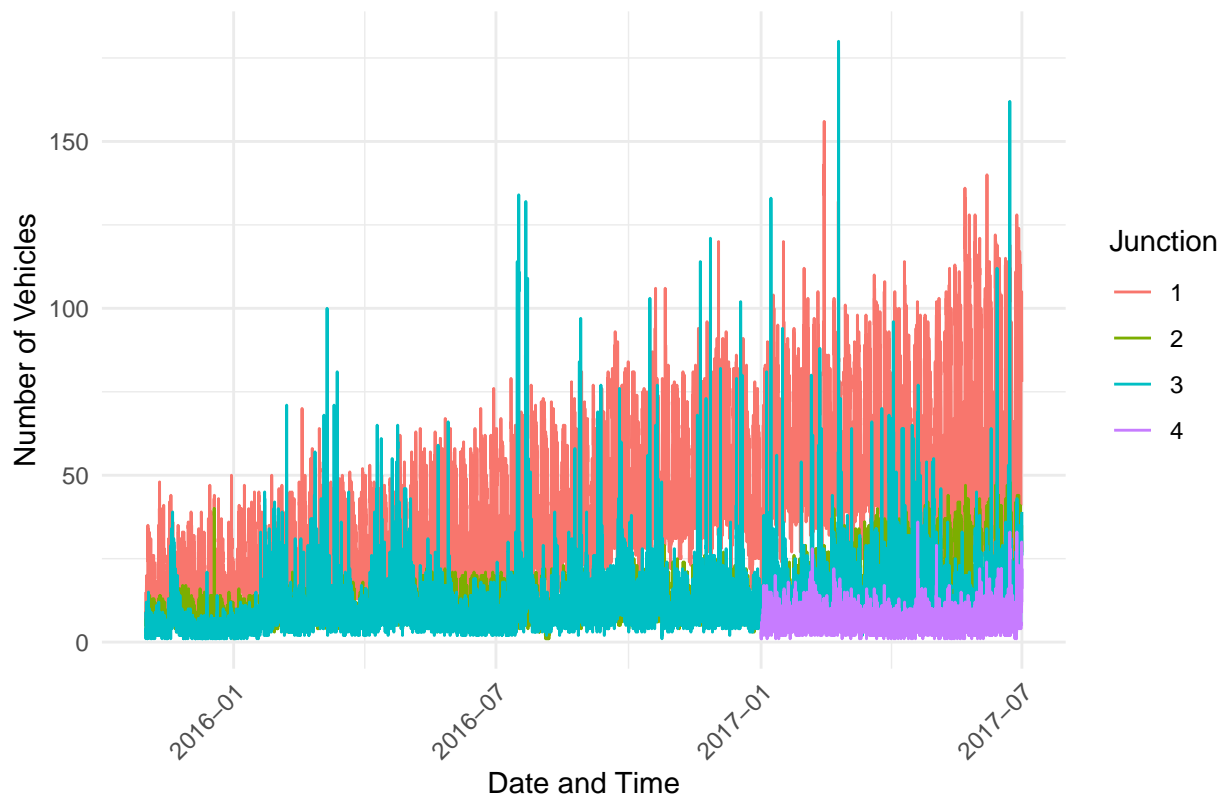
```
junctions_data <- subset(traffic_data, Junction == TRUE)
head(junctions_data)
```

```
## # A tibble: 6 x 4
##   DateTime          Junction Vehicles      ID
##   <dtm>          <dbl>    <dbl>    <dbl>
## 1 2015-11-01 00:00:00         1      15 20151101001
## 2 2015-11-01 01:00:00         1      13 20151101011
## 3 2015-11-01 02:00:00         1      10 20151101021
## 4 2015-11-01 03:00:00         1       7 20151101031
## 5 2015-11-01 04:00:00         1       9 20151101041
## 6 2015-11-01 05:00:00         1       6 20151101051
```

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

```
library(ggplot2)
ggplot(traffic_data, aes(x = DateTime, y = Vehicles, color = factor(Junction))) +
  geom_line() +
  labs(title = "Vehicle Counts at Junctions Over Time",
       x = "Date and Time",
       y = "Number of Vehicles",
       color = "Junction") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Vehicle Counts at Junctions Over Time



7.

From alexa_file.xlsx, import it to your environment

- a. How many observations does alexa_file has? What about the number of columns? Show your solution and answer.

```
library(readxl)
alexa_data <- read_excel("alexa_file.xlsx")
dimensions <- dim(alexa_data)
num_rows <- dimensions[1]
num_columns <- dimensions[2]

num_rows
```

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

```
num_columns
```

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

- b. group the variations and get the total of each variations. Use dplyr package. Show solution and answer.

```
library(dplyr)
variation_totals <- alexa_data %>%
  group_by(variation) %>%
  summarise(total = n())
variation_totals
```

```
## # A tibble: 16 x 2
```

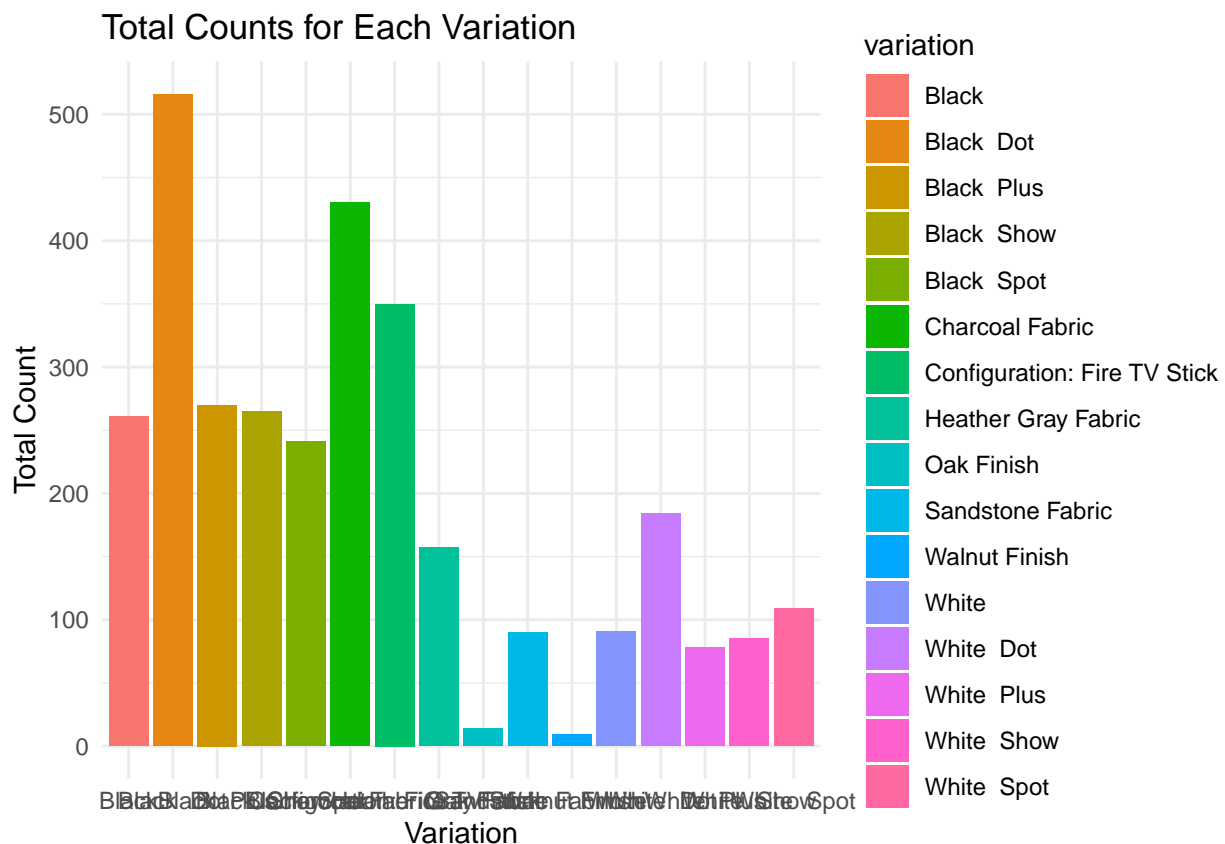
```
##   variation      total
##   <chr>      <int>
## 1 Black      261
```

```
## 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. Plot the variations using the `ggplot()` function. What did you observe? Complete the details of the graph. Show solution and answer.

```
library(ggplot2)

ggplot(variation_totals, aes(x = variation, y = total, fill = variation)) +
  geom_bar(stat = "identity") +
  labs(title = "Total Counts for Each Variation", x = "Variation", y = "Total Count") +
  theme_minimal()
```

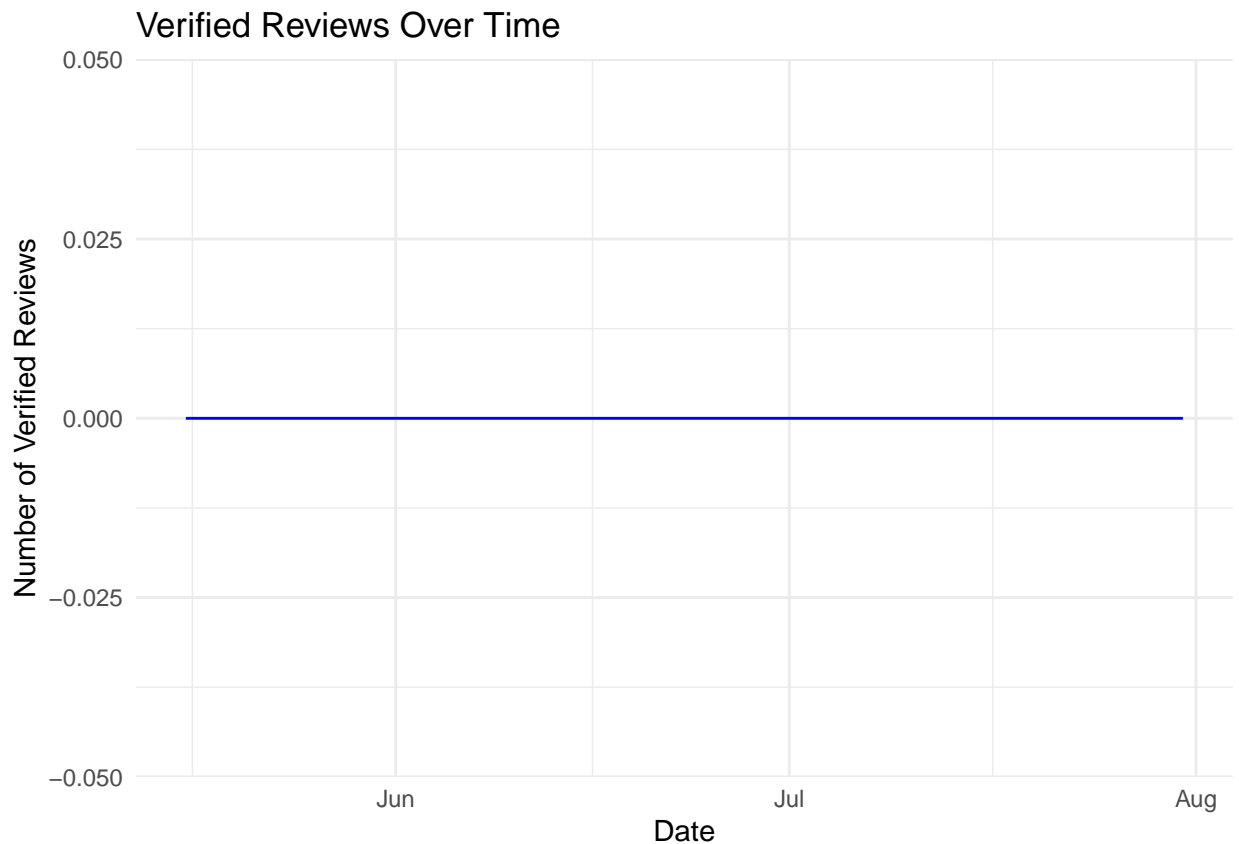


- d. Plot a `geom_line()` with the date and the number of verified reviews. Complete the details of the graphs. Show your answer and solution.

```
alex_data$verified_reviews <- as.numeric(alex_data$verified_reviews)
```

```
## Warning: NAs introduced by coercion
```

```
reviews_by_date <- alex_data %>%
  group_by(date) %>%
  summarise(total_verified_reviews = sum(verified_reviews, na.rm = TRUE))
library(ggplot2)
ggplot(reviews_by_date, aes(x = date, y = total_verified_reviews)) +
  geom_line(color = "blue") +
  labs(title = "Verified Reviews Over Time", x = "Date", y = "Number of Verified Reviews") +
  theme_minimal()
```



e. Get the relationship of variations and ratings. Which variations got the most highest in rating? Plot a graph to show its relationship. Show your solution and answer.

```
variation_ratings <- alex_data %>%
  group_by(variation) %>%
  summarise(average_rating = mean(rating, na.rm = TRUE))
ggplot(variation_ratings, aes(x = variation, y = average_rating, fill = variation)) +
  geom_bar(stat = "identity") +
  labs(title = "Average Rating by Variation", x = "Variation", y = "Average Rating") +
  theme_minimal()
```

