

# Rworksheet\_4c

Arjay De Guzman

2023-11-22

#1. #a.

```
data <- read.csv("mpg.csv")
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(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(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(l4)	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(l4)	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(l5)	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(l4)	f 18
## 156 156	pontiac	grand prix	3.8 1999	6 auto(l4)	f 16
## 157 157	pontiac	grand prix	3.8 1999	6 auto(l4)	f 17
## 158 158	pontiac	grand prix	3.8 2008	6 auto(l4)	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(l4)	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(l4)	4 20
## 165 165	subaru	forester awd	2.5 2008	4 auto(l4)	4 18
## 166 166	subaru	impreza awd	2.2 1999	4 auto(l4)	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(l4)	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(l4)	4 16
## 176 176	toyota	4runner 4wd	3.4 1999	6 auto(l4)	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(l5)	4 16
## 179 179	toyota	4runner 4wd	4.7 2008	8 auto(l5)	4 14
## 180 180	toyota	camry	2.2 1999	4 manual(m5)	f 21
## 181 181	toyota	camry	2.2 1999	4 auto(l4)	f 21
## 182 182	toyota	camry	2.4 2008	4 manual(m5)	f 21
## 183 183	toyota	camry	2.4 2008	4 auto(l5)	f 21
## 184 184	toyota	camry	3.0 1999	6 auto(l4)	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(l4)	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(l4)	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(l3)	f 24
## 195 195	toyota	corolla	1.8 1999	4 auto(l4)	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(l4)	f 26
## 199 199	toyota	land cruiser wagon 4wd	4.7 1999	8 auto(l4)	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(l4)	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(l4)	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(l5)	4	16
##	208	208	volkswagen	gti	2.0	1999	4	manual(m5)	f	21
##	209	209	volkswagen	gti	2.0	1999	4	auto(l4)	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

1b. Which variables from mpg dataset are categorical?

```
str(data)

## 'data.frame': 234 obs. of 12 variables:
## $ X : int 1 2 3 4 5 6 7 8 9 10 ...
## $ manufacturer: chr "audi" "audi" "audi" "audi" ...
## $ model : chr "a4" "a4" "a4" "a4" ...
## $ displ : num 1.8 1.8 2 2 2.8 2.8 3.1 1.8 1.8 2 ...
## $ year : int 1999 1999 2008 2008 1999 1999 2008 1999 1999 2008 ...
## $ cyl : int 4 4 4 4 6 6 6 4 4 4 ...
## $ trans : chr "auto(l5)" "manual(m5)" "manual(m6)" "auto(av)" ...
## $ drv : chr "f" "f" "f" "f" ...
## $ cty : int 18 21 20 21 16 18 18 18 16 20 ...
## $ hwy : int 29 29 31 30 26 26 27 26 25 28 ...
## $ fl : chr "p" "p" "p" "p" ...
## $ class : chr "compact" "compact" "compact" "compact" ...
```

ans. manufacturer, model, trans, drv, fl, class variables are categorical

1c. Which are continuous variables? ans. X, displ, year, cyl, cty, hwy are continuous variables

2. Which manufacturer has the most models in this data set? Which model has the most variations? Show your answer.

```
manufacturer_counts <- table(data$manufacturer)

most_models_manufacturer <- as.character(names(manufacturer_counts)[which.max(manufacturer_counts)])

model_counts <- table(data$model)

most_variations_model <- as.character(names(model_counts)[which.max(model_counts)])

print(paste("The manufacturer with the most models is:", most_models_manufacturer))

## [1] "The manufacturer with the most models is: dodge"

print(paste("The model with the most variations is:", most_variations_model))

## [1] "The model with the most variations is: caravan 2wd"
```

2a. Group the manufacturers and find the unique models. Show your codes and result.

```
dfmpg <- data.frame(Manufacturer = data$manufacturer, Model = data$model)
unique_mpg <- unique(dfmpg)
unique_mpg

##      Manufacturer      Model
## 1          audi          a4
## 8          audi    a4 quattro
## 16         audi    a6 quattro
## 19    chevrolet c1500 suburban 2wd
## 24    chevrolet      corvette
## 29    chevrolet k1500 tahoe 4wd
```

```

## 33      chevrolet          malibu
## 38        dodge      caravan 2wd
## 49        dodge    dakota pickup 4wd
## 58        dodge      durango 4wd
## 65        dodge    ram 1500 pickup 4wd
## 75         ford      expedition 2wd
## 78         ford      explorer 4wd
## 84         ford      f150 pickup 4wd
## 91         ford      mustang
## 100        honda          civic
## 109        hyundai      sonata
## 116        hyundai      tiburon
## 123         jeep    grand cherokee 4wd
## 131    land rover      range rover
## 135        lincoln    navigator 2wd
## 138        mercury    mountaineer 4wd
## 142         nissan          altima
## 148         nissan          maxima
## 151         nissan    pathfinder 4wd
## 155        pontiac      grand prix
## 160         subaru    forester awd
## 166         subaru    impreza awd
## 174         toyota      4runner 4wd
## 180         toyota      camry
## 187         toyota    camry solara
## 194         toyota      corolla
## 199         toyota    land cruiser wagon 4wd
## 201         toyota      toyota tacoma 4wd
## 208    volkswagen          gti
## 213    volkswagen          jetta
## 222    volkswagen      new beetle
## 228    volkswagen      passat

```

2b. Graph the result by using `plot()` and `ggplot()`. Write the codes and its result

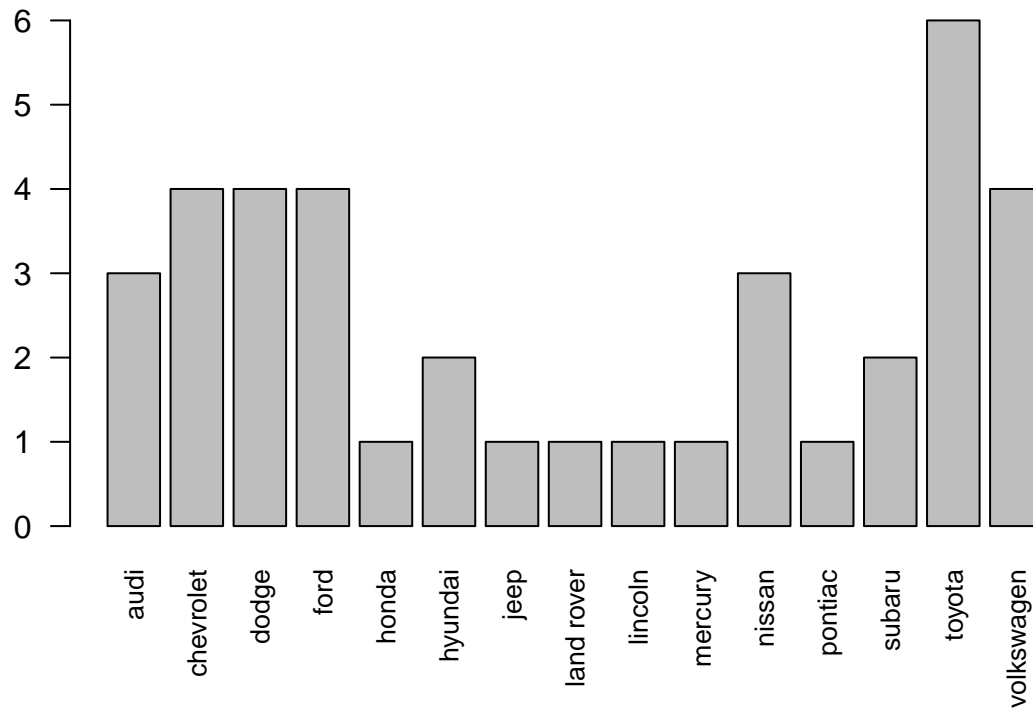
```
library(ggplot2)
```

```

factoredManufacturer <- as.factor(unique_mpg$Manufacturer)
manufacturerPlot <- plot(as.factor(factoredManufacturer),
  main = "Unique Model per Manufacturer",
  cex.names = 0.8, las = 2)

```

## Unique Model per Manufacturer



```
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
```

```
unique_count <- unique_mpg %>%
  count(unique_mpg$Manufacturer)
unique_count
```

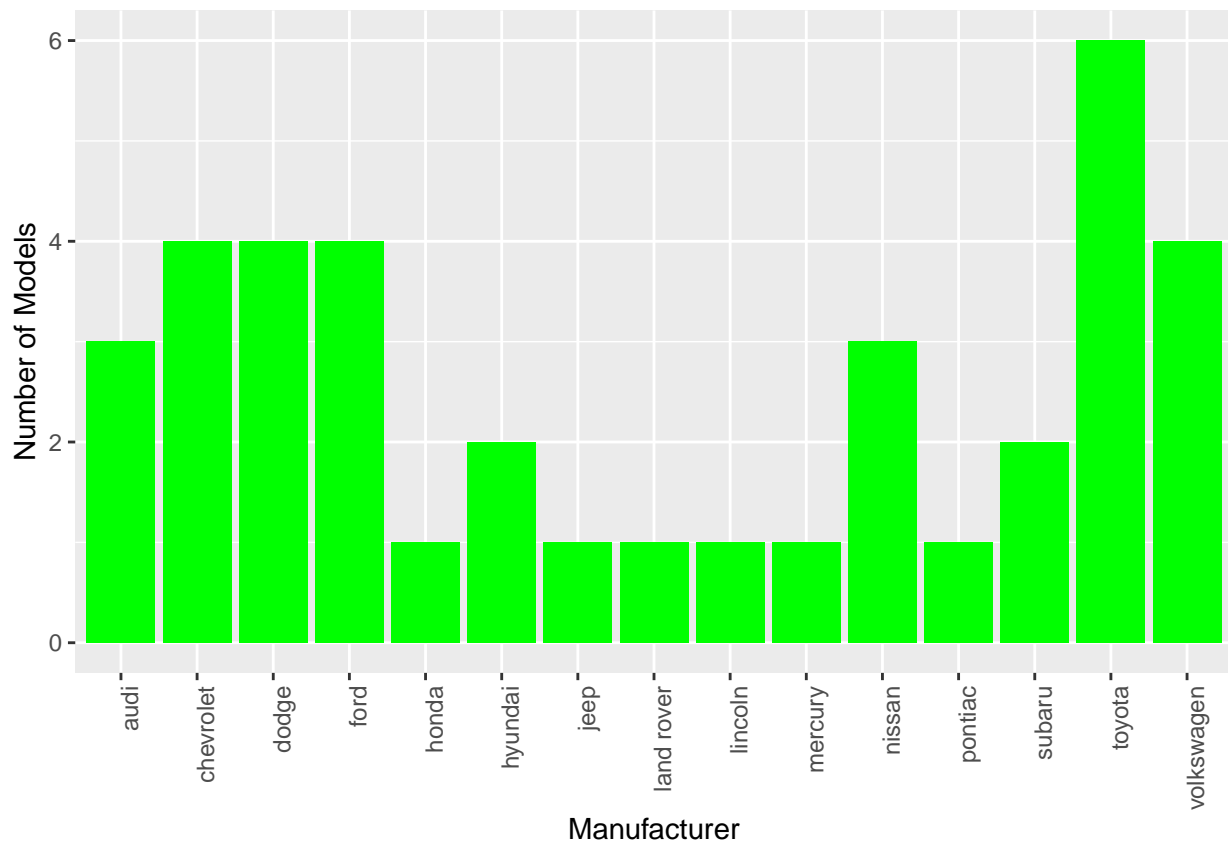
```
##   unique_mpg$Manufacturer n
## 1             audi 3
## 2          chevrolet 4
## 3             dodge 4
## 4             ford 4
## 5             honda 1
## 6          hyundai 2
## 7             jeep 1
## 8        land rover 1
## 9            lincoln 1
## 10          mercury 1
## 11            nissan 3
## 12          pontiac 1
## 13          subaru 2
```

```
## 14          toyota 6
## 15          volkswagen 4
```

```
names(unique_count) <- c("Manufacturer", "Number of Models")
unique_count
```

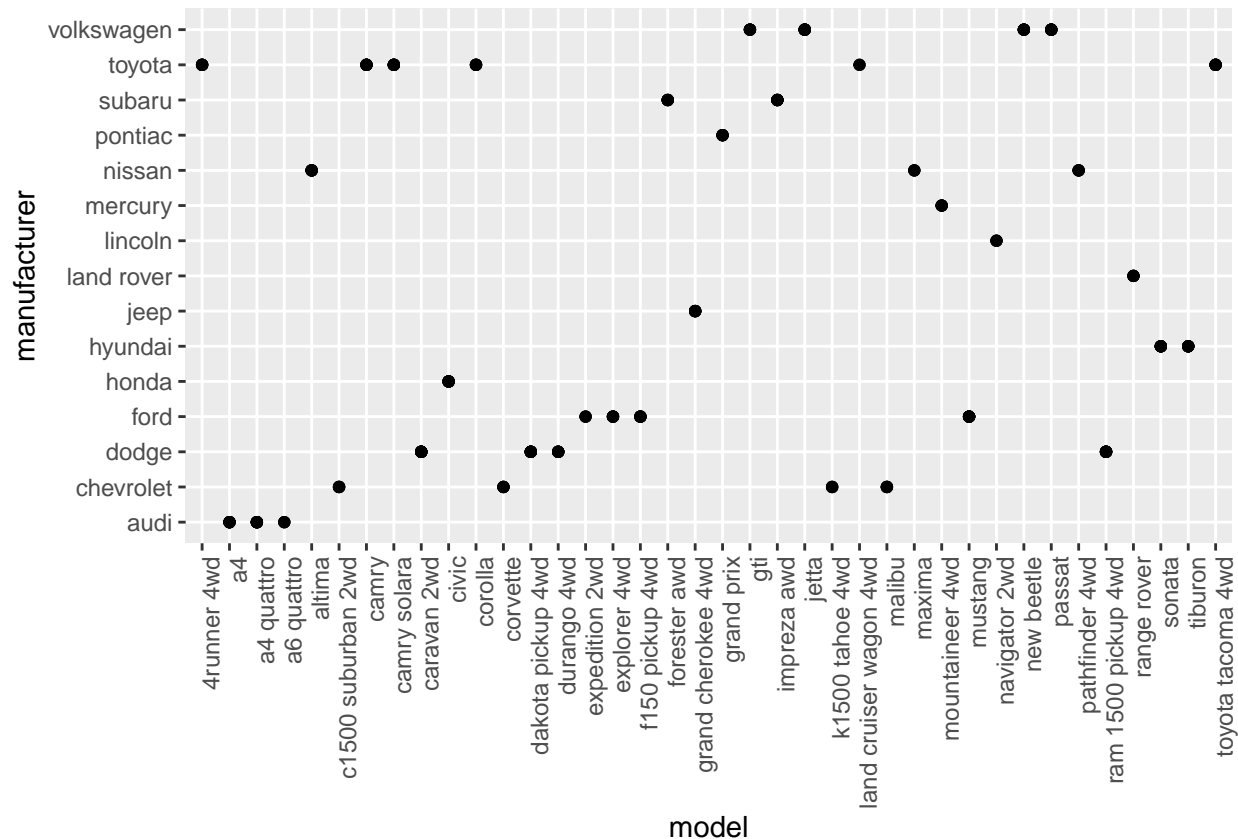
```
##      Manufacturer Number of Models
## 1         audi          3
## 2      chevrolet          4
## 3         dodge          4
## 4          ford          4
## 5         honda          1
## 6        hyundai          2
## 7          jeep          1
## 8    land rover          1
## 9         lincoln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

```
ggplot(unique_count, aes(x = Manufacturer, y = `Number of Models`)) +
  geom_bar(stat = "identity", fill = "green") + theme(plot.title = element_text(hjust = 0.5),
    axis.text.x = element_text(angle = 90, hjust = 1))
```



2. Same dataset will be used. You are going to show the relationship of the model and the manufacturer.  
 2a. What does `ggplot(mpg, aes(model, manufacturer)) + geom_point()` show?

```
library(ggplot2)
ggplot(data, aes(model, manufacturer)) + geom_point() + theme(plot.title = element_text(hjust = 0.5),
  axis.text.x = element_text(angle = 90, hjust = 1))
```



ans. Plots the mpgdataset with model on the x-axis and manufacturer on the y-axis. Each point represents a car model

2b. For you, is it useful? If not, how could you modify the data to make it more informative?

It's useful to know how many models each manufacturer produces. To make it more interesting, I can color code each point based on the class variable and add a label using the cty variable.

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

```
library(ggplot2)
```

```
Aobj <- head(data, 20)
```

```
Aobj
```

##	X	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl
## 1	1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p
## 2	2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p
## 3	3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p
## 4	4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p
## 5	5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p
## 6	6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p
## 7	7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p
## 8	8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p
## 9	9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p

```

## 10 10      audi      a4 quattro  2.0 2008  4 manual(m6)  4 20 28 p
## 11 11      audi      a4 quattro  2.0 2008  4  auto(s6)  4 19 27 p
## 12 12      audi      a4 quattro  2.8 1999  6  auto(l5)  4 15 25 p
## 13 13      audi      a4 quattro  2.8 1999  6 manual(m5)  4 17 25 p
## 14 14      audi      a4 quattro  3.1 2008  6  auto(s6)  4 17 25 p
## 15 15      audi      a4 quattro  3.1 2008  6 manual(m6)  4 15 25 p
## 16 16      audi      a6 quattro  2.8 1999  6  auto(l5)  4 15 24 p
## 17 17      audi      a6 quattro  3.1 2008  6  auto(s6)  4 17 25 p
## 18 18      audi      a6 quattro  4.2 2008  8  auto(s6)  4 16 23 p
## 19 19  chevrolet c1500 suburban 2wd  5.3 2008  8  auto(l4)  r 14 20 r
## 20 20  chevrolet c1500 suburban 2wd  5.3 2008  8  auto(l4)  r 11 15 e
##      class
## 1 compact
## 2 compact
## 3 compact
## 4 compact
## 5 compact
## 6 compact
## 7 compact
## 8 compact
## 9 compact
## 10 compact
## 11 compact
## 12 compact
## 13 compact
## 14 compact
## 15 compact
## 16 midsize
## 17 midsize
## 18 midsize
## 19  suv
## 20  suv

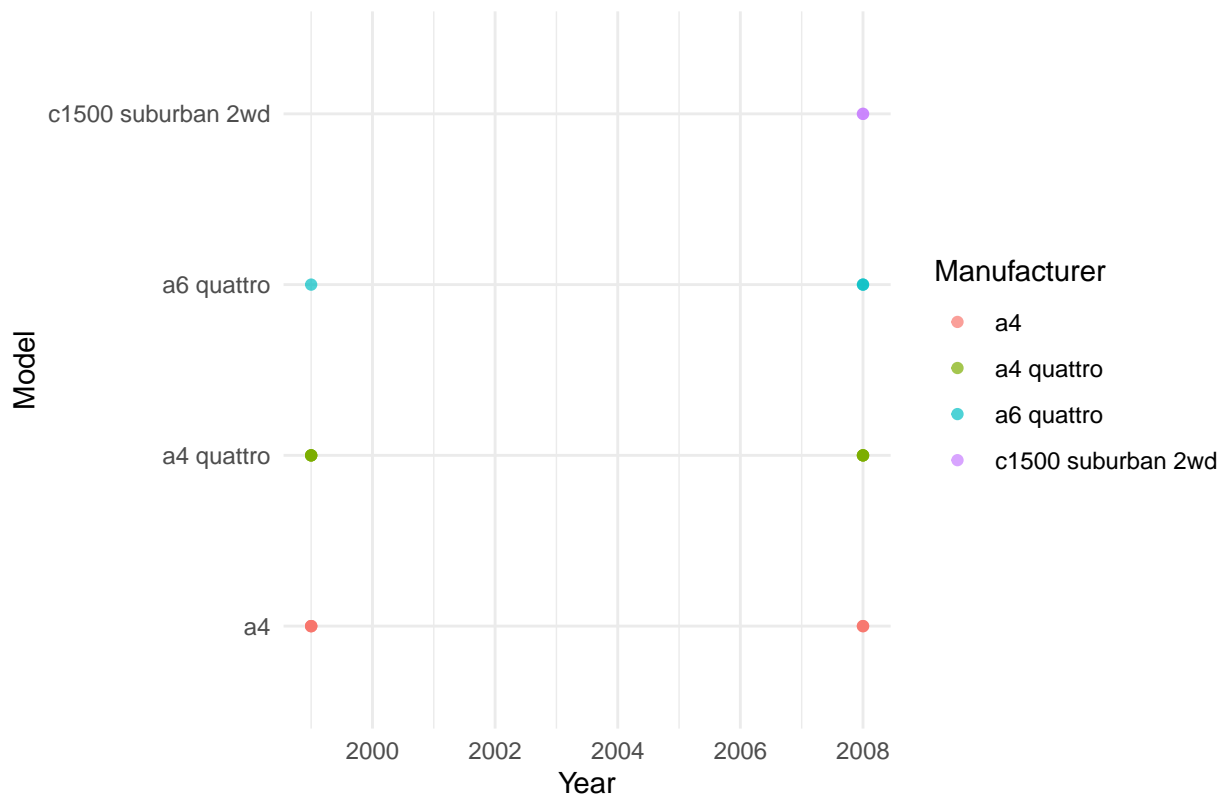
```

```

ggplot(data = Aobj, aes(x = year, y = model, color = as.factor(model))) +
  geom_point(alpha = 0.7) +
  labs(title = "Relationship between Model, Year, and Manufacturer",
       x = "Year",
       y = "Model",
       color = "Manufacturer") +
  theme_minimal()

```

## Relationship between Model, Year, and Manufacturer



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

```
library(dplyr)
library(ggplot2)

head(data, 20)
```

##	X	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl
## 1	1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p
## 2	2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p
## 3	3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p
## 4	4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p
## 5	5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p
## 6	6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p
## 7	7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p
## 8	8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p
## 9	9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p
## 10	10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p
## 11	11	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p
## 12	12	audi	a4 quattro	2.8	1999	6	auto(l5)	4	15	25	p
## 13	13	audi	a4 quattro	2.8	1999	6	manual(m5)	4	17	25	p
## 14	14	audi	a4 quattro	3.1	2008	6	auto(s6)	4	17	25	p
## 15	15	audi	a4 quattro	3.1	2008	6	manual(m6)	4	15	25	p
## 16	16	audi	a6 quattro	2.8	1999	6	auto(l5)	4	15	24	p
## 17	17	audi	a6 quattro	3.1	2008	6	auto(s6)	4	17	25	p
## 18	18	audi	a6 quattro	4.2	2008	8	auto(s6)	4	16	23	p
## 19	19	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14	20	r



```
## 20 20    chevrolet c1500 suburban 2wd    5.3 2008    8    auto(14)    r 11 15 e
##      class
## 1 compact
## 2 compact
## 3 compact
## 4 compact
## 5 compact
## 6 compact
## 7 compact
## 8 compact
## 9 compact
## 10 compact
## 11 compact
## 12 compact
## 13 compact
## 14 compact
## 15 compact
## 16 midsize
## 17 midsize
## 18 midsize
## 19    suv
## 20    suv
```

```
data %>%
  group_by(model) %>%
  summarize(count = n()) -> grouped_data
```

```
grouped_data
```

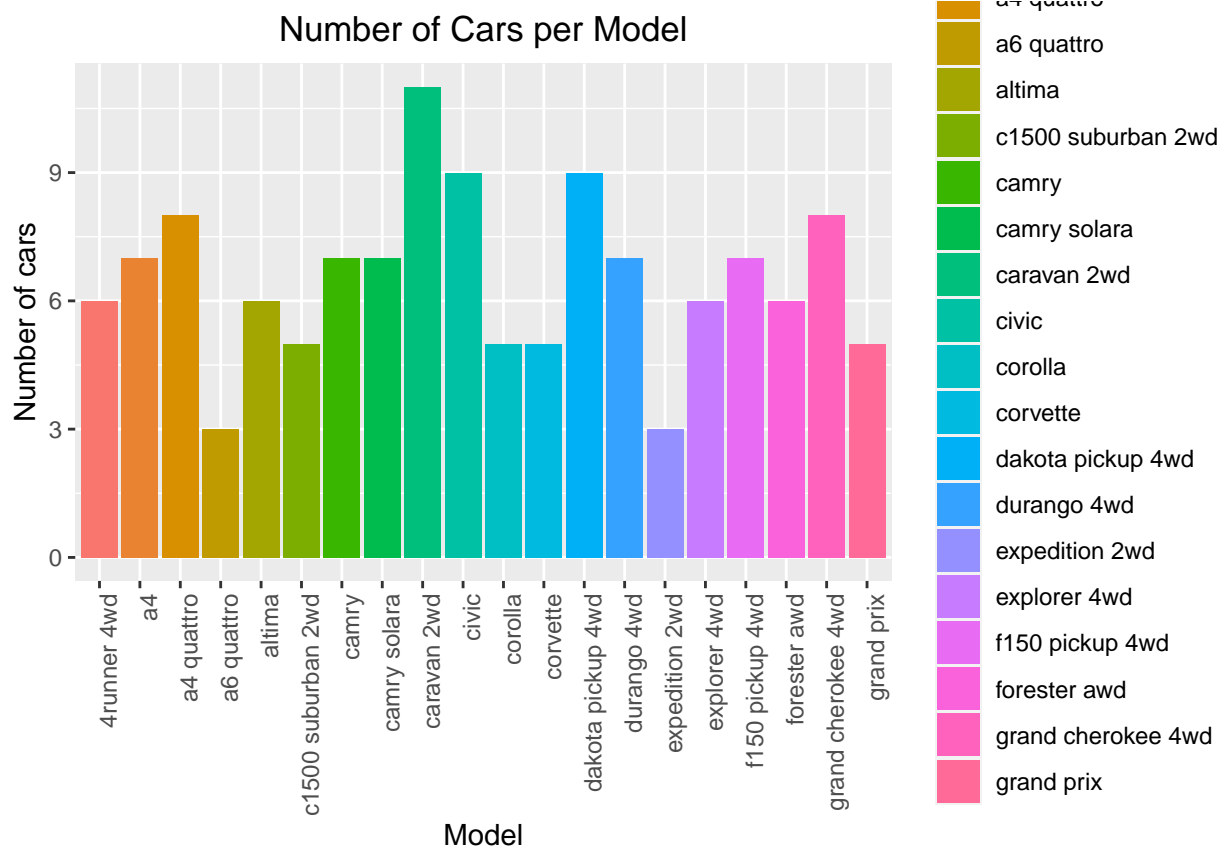
```
## # A tibble: 38 x 2
##   model      count
##   <chr>      <int>
## 1 4runner 4wd         6
## 2 a4              7
## 3 a4 quattro        8
## 4 a6 quattro        3
## 5 altima           6
## 6 c1500 suburban 2wd  5
## 7 camry            7
## 8 camry solara       7
## 9 caravan 2wd       11
## 10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

4a. Plot using `geom_bar()` using the top 20 observations only. The graphs should have a title, labels and colors. Show code and results.

```
mpg_observation_top20 <- head(grouped_data, 20)

mpg_observation_top20 <- ggplot(mpg_observation_top20, aes(x = model, y = count, fill = model)) +
  geom_bar(stat = "identity") +
  labs(title = "Number of Cars per Model",
       x = "Model" ,
       y = "Number of cars")+ theme(plot.title = element_text(hjust = 0.5),
                                   axis.text.x = element_text(angle = 90, hjust = 1))
```

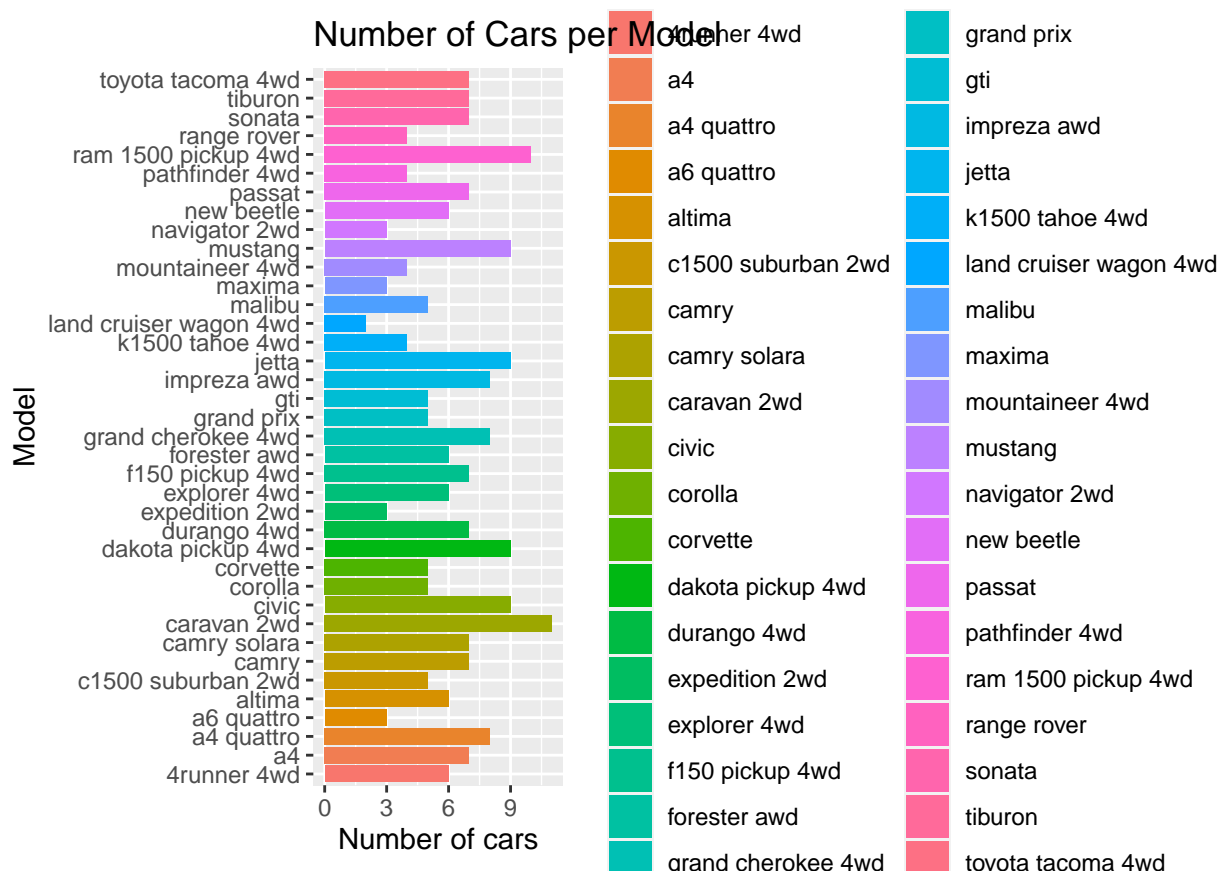
mpg\_observation\_top20



4b. Plot using the `geom_bar()` + `coord_flip()` just like what is shown below. Show codes and its result.

```
library(ggplot2)
mpg_top20 <- ggplot(grouped_data, aes(x = model, y = count, fill = model)) +
  geom_bar(stat = "identity") +
  labs(title = "Number of Cars per Model",
       x = "Model" ,
       y = "Number of cars") +
  coord_flip()
```

mpg\_top20

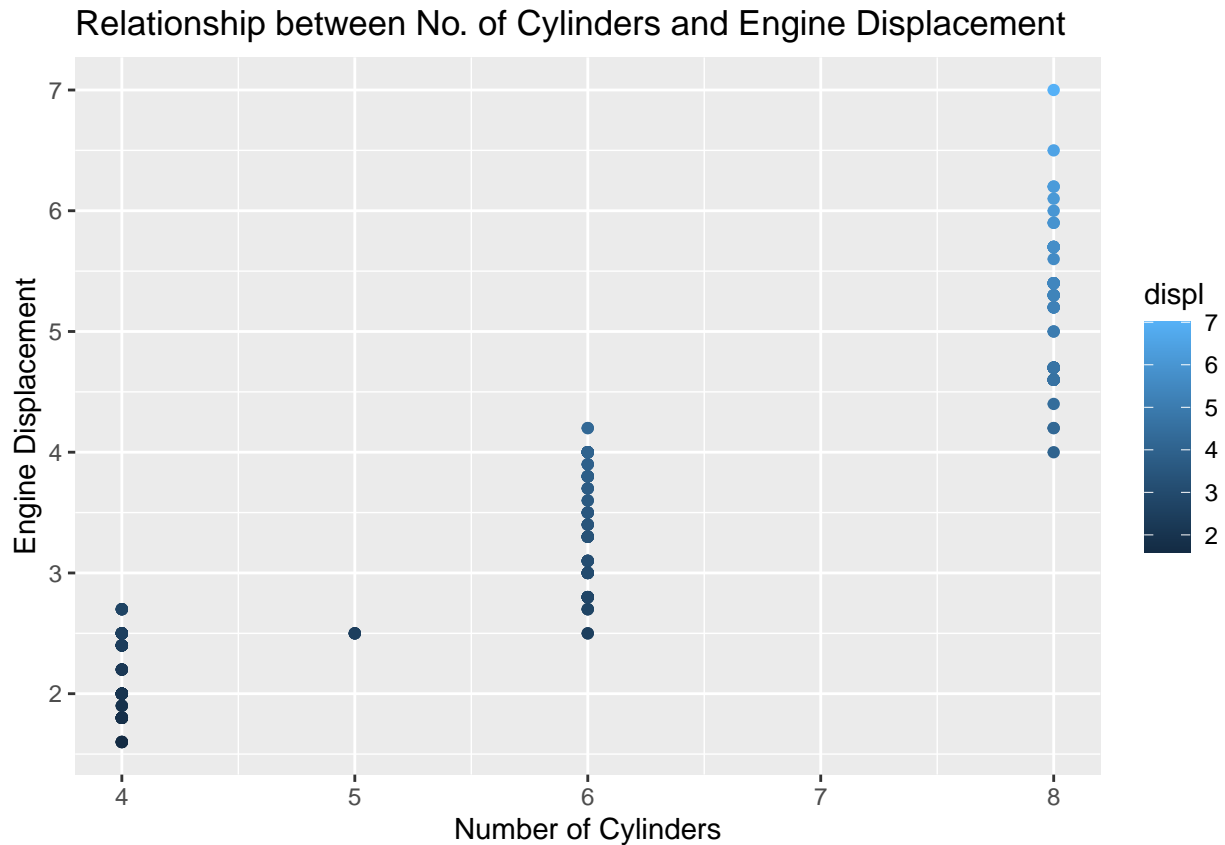


5. Plot the relationship between cyl - number of cylinders and displ - engine displacement using `geom_point` with `aesthetic color = engine displacement`. Title should be "Relationship between No. of Cylinders and Engine Displacement".

```
relationship_cyl_displ <- ggplot(data, aes(x = cyl,
                                           y = displ,
                                           color = displ)) +

  geom_point() +
  labs(title = "Relationship between No. of Cylinders and Engine Displacement",
       x = "Number of Cylinders",
       y = "Engine Displacement")

relationship_cyl_displ
```



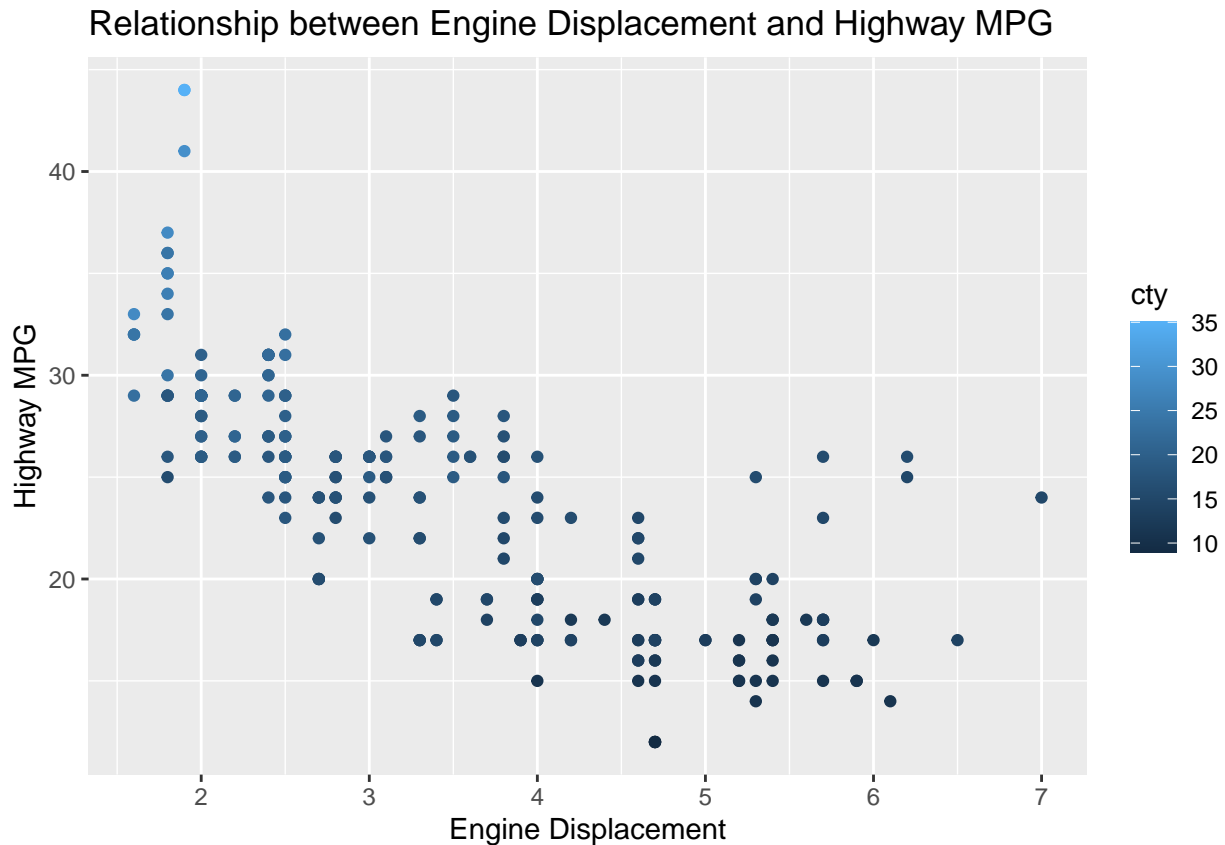
5a. How would you describe its relationship? Show the codes and its result.

The scatter figure illustrates the relationship between engine displacement and cylinder count, showing that larger engines with more cylinders often have higher displacement. As cylinder count increases, so does the displacement.

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?

```
displ_hwy_relationship <- ggplot(data, aes(x = displ, y = hwy, color = cty)) +
  geom_point() +
  labs(title = "Relationship between Engine Displacement and Highway MPG",
        x = "Engine Displacement",
        y = "Highway MPG")

displ_hwy_relationship
```



Each point in the scatterplot represents a city's fuel efficiency, and it displays the relationship between engine displacement and highway miles per gallon. Insight into vehicle fuel efficiency for various engine sizes can be gained by visualizing the relationship between various city fuel efficiency levels and displ and hwy.

6.Import the traffic.csv onto your R environment. a.How many numbers of observation does it have? What are the variables of the traffic dataset the Show your answer.

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

num_obs <- nrow(traffic)
num_obs
```

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

```
num_vars <- ncol(traffic)
num_vars
```

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

```
vars <- colnames(traffic)
vars
```

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

6b. subset the traffic dataset into junctions. What is the R codes and its output?

```
junctions_subset1 <- subset(traffic, Junction == 1)
```

```
junctions_subset2 <- subset(traffic, Junction == 2)
```

```
junctions_subset3 <- subset(traffic, Junction == 3)
```

```
junctions_subset4 <- subset(traffic, Junction == 4)
```

6c. Plot each junction in a using `geom_line()`. Show your solution and output.

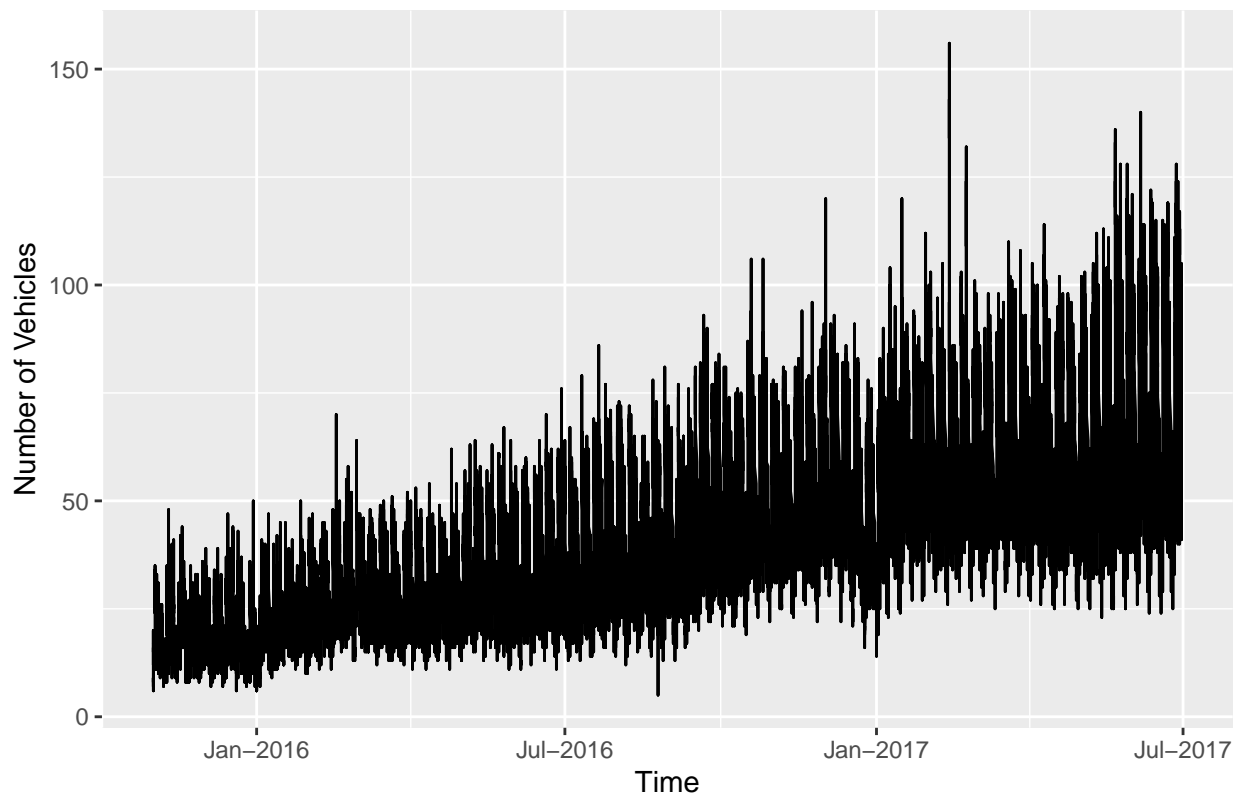
```
junction_plot_1 <- ggplot(junctions_subset1,  
  aes(x = as.Date(junctions_subset1$DateTime),  
      y = Vehicles)) +  
  
  geom_line() +  
  scale_x_date(date_labels = "%b-%Y") + theme(legend.position = "none") +  
  labs(title = "Junction 1", x = "Time", y = "Number of Vehicles")
```

```
junction_plot_1
```

```
## Warning: Use of `junctions_subset1$DateTime` is discouraged.
```

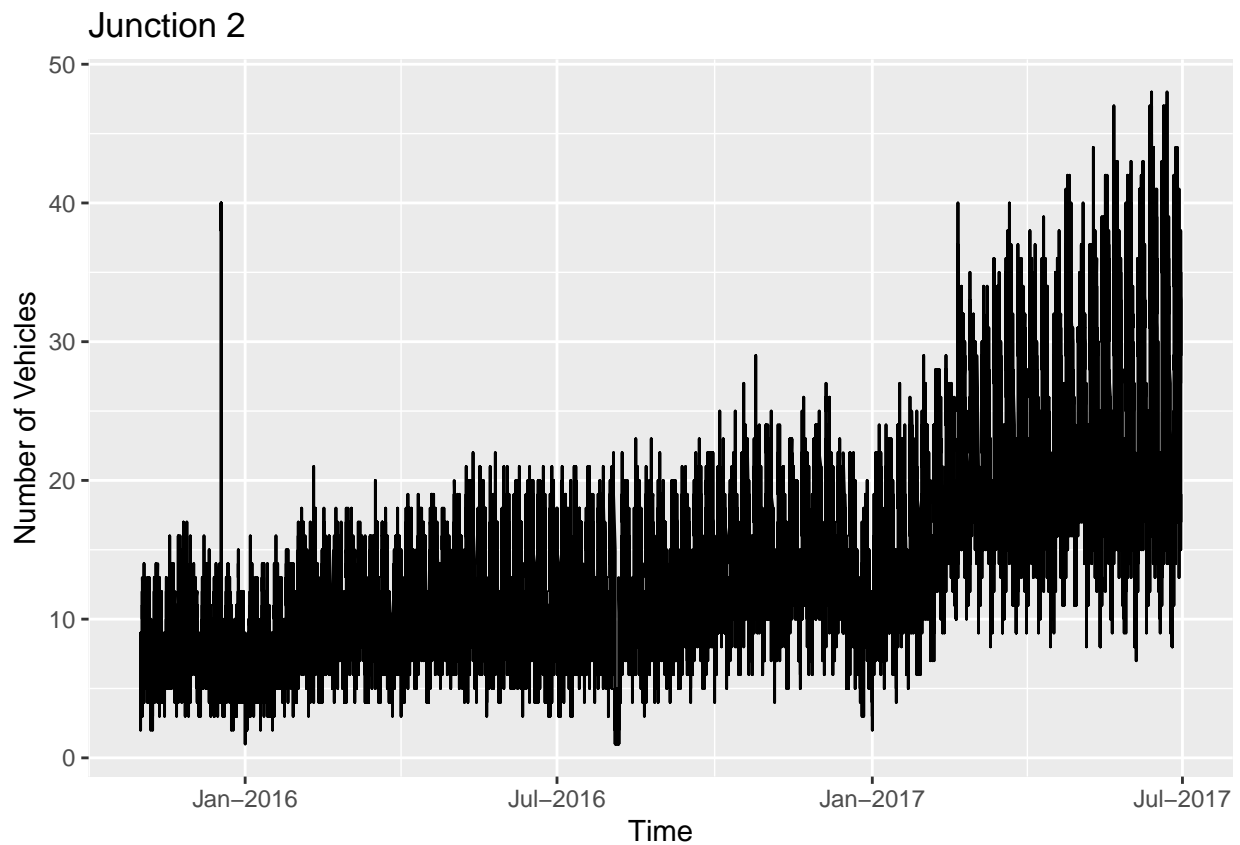
```
## i Use `DateTime` instead.
```

### Junction 1



```
junction_plot_2 <- ggplot(junctions_subset2,  
  aes(x = as.Date(junctions_subset2$DateTime),  
      y = Vehicles)) +  
  
  geom_line() +  
  scale_x_date(date_labels = "%b-%Y") + theme(legend.position = "none") +  
  labs(title = "Junction 2", x = "Time", y = "Number of Vehicles")
```

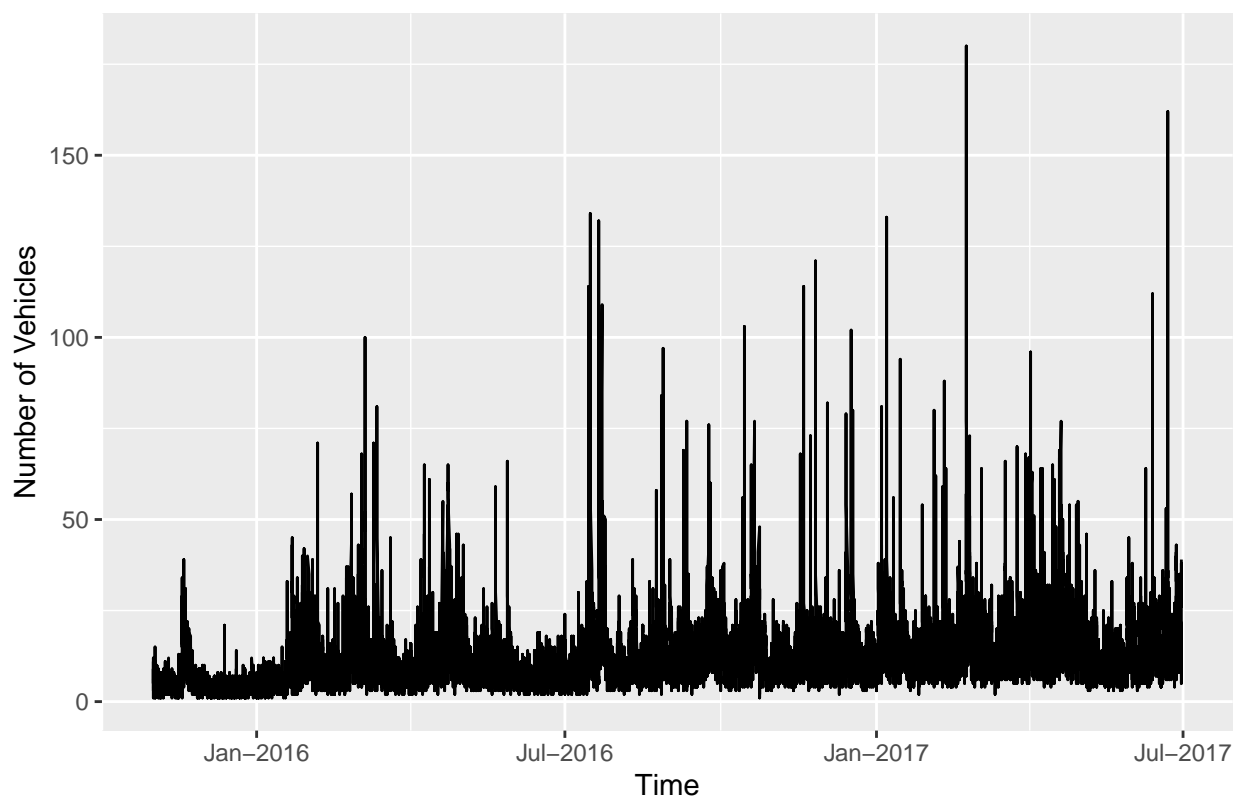
```
junction_plot_2
```



```
junction_plot_3 <- ggplot(junctions_subset3,
  aes(x = as.Date(junctions_subset3$DateTime),
    y = Vehicles)) +
  geom_line() +
  scale_x_date(date_labels = "%b-%Y") + theme(legend.position = "none") +
  labs(title = "Junction 3", x = "Time", y = "Number of Vehicles")

junction_plot_3
```

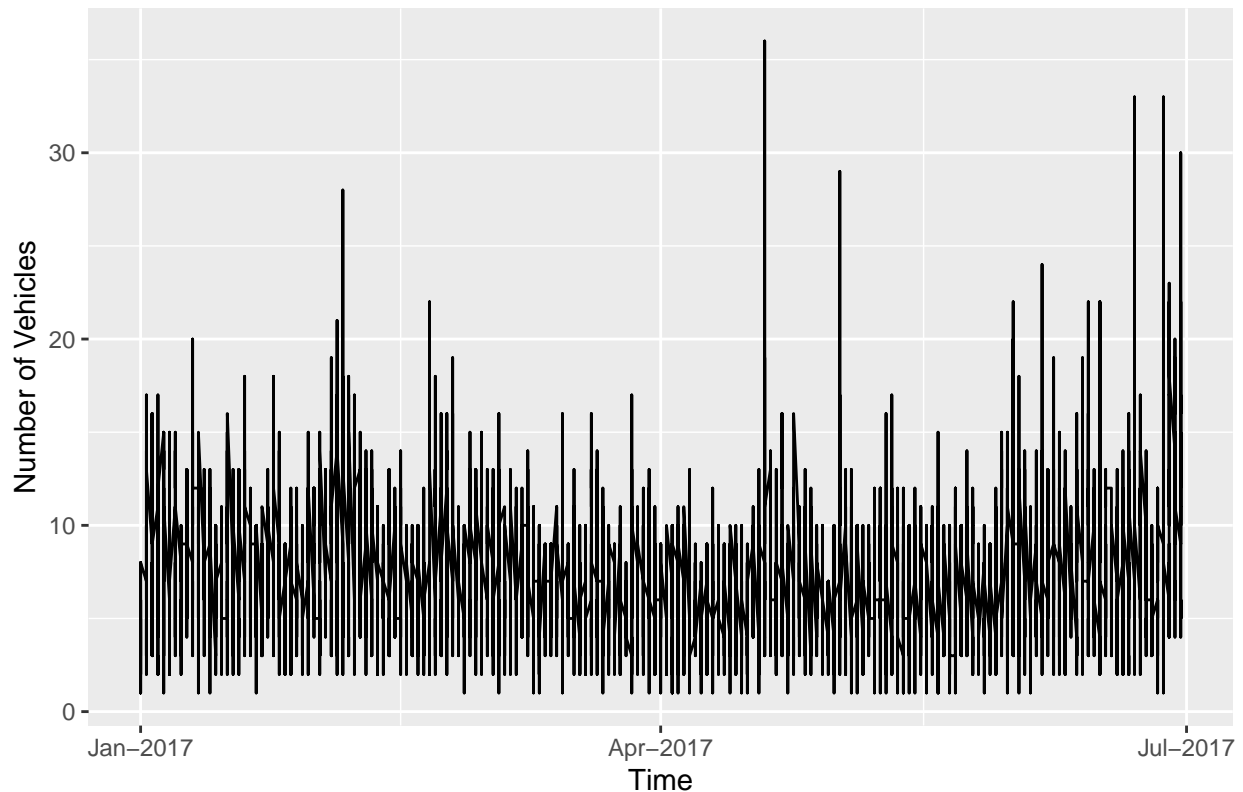
### Junction 3



```
junction_plot_4 <- ggplot(junctions_subset4,  
  aes(x = as.Date(junctions_subset4$DateTime),  
      y = Vehicles)) +  
  geom_line() +  
  scale_x_date(date_labels = "%b-%Y") + theme(legend.position = "none") +  
  labs(title = "Junction 4", x = "Time", y = "Number of Vehicles")  
  
junction_plot_4
```



## Junction 4



7. From alexa\_file.xlsx, import it to your environment

7a. How many observations does alexa\_file have? What about the number of columns? Show your solution and answer.

```
library(readxl)

open_alexa_data <- read_excel("/cloud/project/Worksheet 4/Rworksheet_4c/alexa_file.xlsx")

num_obs <- nrow(open_alexa_data)
num_obs
```

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

```
num_cols <- ncol(open_alexa_data)
num_cols
```

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

7b. group the variations and get the total of each variations. Use dplyr package. Show solution and answer

```
var_counts <- open_alexa_data %>%
  count(variation)

var_counts
```

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

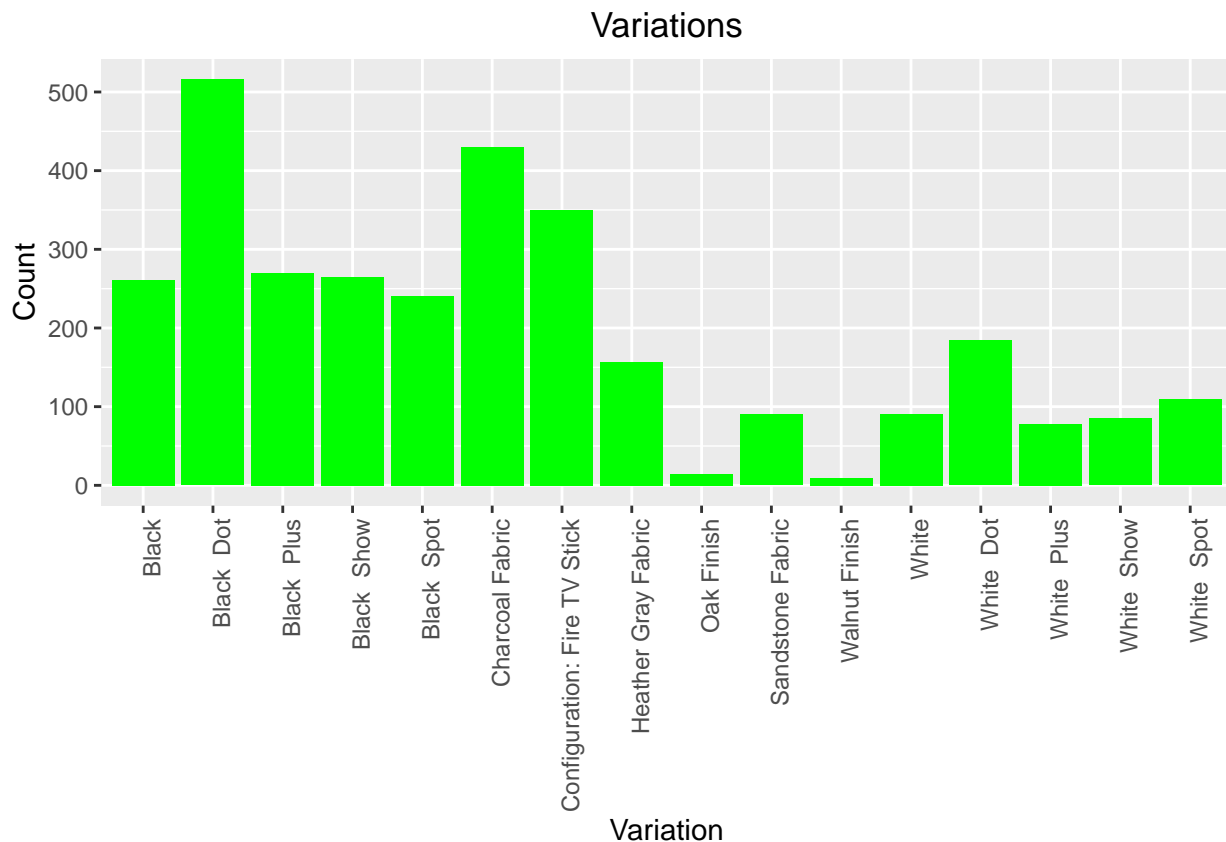
```
##   variation      n
##   <chr>      <int>
## 1 Black      261
## 2 Black Dot  516
```

```
## 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
```

7c. Plot the variations using the `ggplot()` function. What did you observe? Complete the details of the graph. Show solution and answer.

```
plot_alexafile <- ggplot(open_alex_data, aes(x = variation)) +
  geom_bar(fill = "green") +
  labs(title = "Variations",
       x = "Variation",
       y = "Count") + theme(plot.title = element_text(hjust = 0.5),
                           axis.text.x = element_text(angle = 90, hjust = 1))
```

plot\_alexafile



Each bar on the graph represents a distinct variation, and the height of a bar indicates the number of that variation. The graph shows the distribution of variations and their counts.

7d. Plot a `geom_line()` with the date and the number of verified reviews. Complete the details of the graphs. Show your answer and solution

```
library(dplyr)

open_alex_data$date <- as.Date(open_alex_data$date)

open_alex_data$month <- format(open_alex_data$date, "%m")

alex_month <- open_alex_data %>%
  count(month)
alex_month

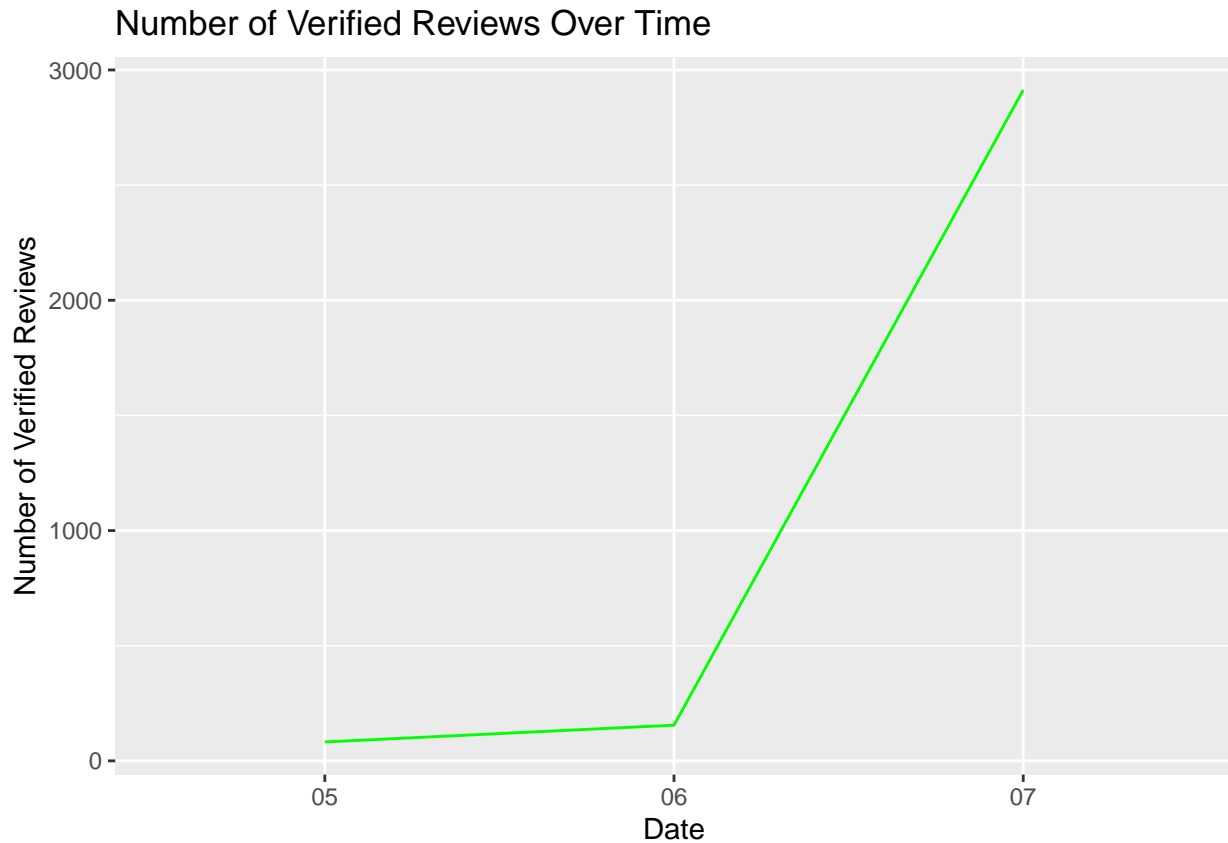
## # A tibble: 3 x 2
##   month     n
##   <chr> <int>
## 1 05      82
## 2 06     155
## 3 07     2913

alex_monthly_reviews <- table(alex_month)
alex_monthly_reviews

##           n
## month 82 155 2913
##    05 1 0 0
##    06 0 1 0
##    07 0 0 1

alex_lineplot <- ggplot(alex_month, aes(x = month, y = n, group = 1)) +
  geom_line(color = "green") +
  labs(title = "Number of Verified Reviews Over Time",
       x = "Date",
       y = "Number of Verified Reviews")

alex_lineplot
```



7e. 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.

```
alex_a_variation_ratings <- open_alex_a_data %>%
  group_by(variation) %>%
  summarise(avg_rating = mean(rating))
```

```
alex_a_variation_ratings
```

```
## # A tibble: 16 x 2
##   variation          avg_rating
##   <chr>             <dbl>
## 1 Black             4.23
## 2 Black Dot         4.45
## 3 Black Plus        4.37
## 4 Black Show        4.49
## 5 Black Spot        4.31
## 6 Charcoal Fabric   4.73
## 7 Configuration: Fire TV Stick 4.59
## 8 Heather Gray Fabric 4.69
## 9 Oak Finish         4.86
## 10 Sandstone Fabric  4.36
## 11 Walnut Finish     4.89
## 12 White             4.14
## 13 White Dot         4.42
## 14 White Plus        4.36
## 15 White Show        4.28
## 16 White Spot        4.31
```

```
highest_ratings <- alexa_variation_ratings %>%
  filter(avg_rating == max(avg_rating))
```

```
highest_ratings
```

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

```
# The walnut finish variation has the highest rating
```

```
ggplot(alexa_variation_ratings, aes(x = variation, y = avg_rating)) +
  geom_bar(stat = "identity", fill = "green") +
  labs(title = "Average Ratings by Variation",
       x = "Variation",
       y = "Average Rating") + theme(plot.title = element_text(hjust = 0.5),
                                     axis.text.x = element_text(angle = 90, hjust = 1))
```

