HW1 Solutions

Problem 1:

1a

```
#1a. Import the data and give appropriate column names.

#read abalone data file
abalone_df <- read.table("C:/Users/hyperion/University of Michigan Dropbox/Deepan Islam/Fall

#Assign column names according to abalone.names
colnames(abalone_df) <- c("Sex", "Length", "Diameter", "Height", "Whole_Weight", "Shucked_Wei

head(abalone_df) #lets take a prelim look.
```

```
  Sex Length Diameter Height Whole_Weight Shucked_Weight Viscera_Weight
1   M  0.455    0.365  0.095       0.5140         0.2245         0.1010
2   M  0.350    0.265  0.090       0.2255         0.0995         0.0485
3   F  0.530    0.420  0.135       0.6770         0.2565         0.1415
4   M  0.440    0.365  0.125       0.5160         0.2155         0.1140
5   I  0.330    0.255  0.080       0.2050         0.0895         0.0395
6   I  0.425    0.300  0.095       0.3515         0.1410         0.0775
  Shell_Weight Rings
1        0.150    15
2        0.070     7
3        0.210     9
4        0.155    10
5        0.055     7
6        0.120     8
```

1b

```r
sex_counts <-table(abalone_df$Sex)
sex_counts
```

```
   F    I    M 
1307 1342 1528 
```

```r
males_count <-sum(abalone_df$Sex=="M")
females_count <-sum(abalone_df$Sex=="F")
infants_count <-sum(abalone_df$Sex=="I")

cat("\nThe number of males is:",males_count,
    "\nThe number of females is", females_count,
    "\nThe number of infants is", infants_count)
```

```

The number of males is: 1528 
The number of females is 1307 
The number of infants is 1342
```

1c (1)

Which weight has the highest correlation with rings?

```r
#get weights from dataframe
weights <- names(abalone_df)[5:8]

#do correlations with ring
weight_correlations <-sapply(abalone_df[weights], function(x) cor(x, abalone_df$Rings))

#find which weight has the highest correlation
highest_corr_weight <-names(which.max(weight_correlations))

cat("The weight with the highest correlation to rings is:", highest_corr_weight, "and the co
```

```
The weight with the highest correlation to rings is: Shell_Weight and the correlation value
```

1c (2)

For that weight, which sex has the highest correlation?

```
#We know the weight with the highest correlation to rings is the Shell-Weight.

correlations_by_sex <- sapply(c("F", "I", "M"), function(s) {
  subset_data <- abalone_df[abalone_df$Sex == s, ]
  cor(subset_data$Shell_Weight, subset_data$Rings)
})

print(correlations_by_sex)
```

```
        F         I         M
0.4059070 0.7254357 0.5109967
```

```
highest_corr_sex <- names(which.max(correlations_by_sex))
max_corr <- max(correlations_by_sex)

cat("\nFor the shell weights,", highest_corr_sex, "has the highest correlation to rings with
```

```

For the shell weights, I has the highest correlation to rings with a value of  0.7254357
```

1c (3)

What are the weights of the abalone with the most rings?

```
# Find max number of rings
max_rings <- max(abalone_df$Rings)
rows_max  <- which(abalone_df$Rings == max_rings)

cat("Max number of rings:", max_rings, "\n")
```

```
Max number of rings: 29
```

```
# Get all abalones with max rings
max_rings_data <- abalone_df[abalone_df$Rings == max_rings,]

# Display their weights
cat("Abalone(s) with most rings (", max_rings, " rings):\n", sep = "")
```

```
Abalone(s) with most rings (29 rings):
```

```r
print(abalone_df[rows_max, c("Sex", weights), drop = FALSE])
```

```
    Sex Whole_Weight Shucked_Weight Viscera_Weight Shell_Weight
481   F       1.8075         0.7055         0.3215        0.475
```

1c (4)

What percentage of abalones have a viscera weight larger than their shell weight?

```r
viscera_weight_comp <- mean(abalone_df$Viscera_Weight>abalone_df$Shell_Weight)*100

cat("Percentage with viscera weight > shell weight:", round(viscera_weight_comp, 3), "%")
```

```
Percentage with viscera weight > shell weight: 6.512 %
```

1d

Create a table of correlations between weights and rings, within each sex. The columns should be the four weights, and the rows should be the sexes. (This table does not need to be "fancy" but should clearly identify what each value represents.)

```r
# Create correlation table

sexes <-names(table(abalone_df$Sex))
#weights were defined before

corr_table <-matrix(nrow = length(sexes), ncol=length(weights))
rownames(corr_table) <- sexes
colnames(corr_table) <- weights

# populate the table
for(sx in sexes) {
  subset_data <- abalone_df[abalone_df$Sex == sx, ]
  for(weight in weights) {
    corr_table[sx, weight] <- cor(subset_data[[weight]], subset_data$Rings)
  }
}
```

```
# Display result
print(corr_table)
```

```
  Whole_Weight Shucked_Weight Viscera_Weight Shell_Weight
F    0.2667585     0.09484802      0.2116154    0.4059070
I    0.6963268     0.62024577      0.6732727    0.7254357
M    0.3721966     0.22239382      0.3209535    0.5109967
```

1e

Carry out a series of t-tests to examine whether the number of rings differs across the three sexes. Present the R output and interpret the results. (You may use an existing R function to carry out the t-test, or for minor extra credit, manually write your own calculation of the t-test p-values.)

So we need to do the following t-tests: 1. Female vs Infant 2. Female vs Male 3. Infant vs Male

```
#let's define the rings here so we don't need to retype over and over
female_rings<-abalone_df[abalone_df$Sex=="F","Rings"]
infant_rings<-abalone_df[abalone_df$Sex=="I","Rings"]
male_rings<-abalone_df[abalone_df$Sex=="M","Rings"]

#Test 1: Female vs Infant
t_test_FI <- t.test(female_rings, infant_rings)
print(t_test_FI)
```

```

	Welch Two Sample t-test

data:  female_rings and infant_rings
t = 29.477, df = 2508.9, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3.023380 3.454304
sample estimates:
mean of x mean of y 
11.129304  7.890462 
```

```r
#Test 2: Female vs Male
t_test_FM <- t.test(female_rings, male_rings)
print(t_test_FM)
```

```

    Welch Two Sample t-test

data:  female_rings and male_rings
t = 3.6657, df = 2742.4, p-value = 0.0002514
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1971045 0.6505082
sample estimates:
mean of x mean of y 
  11.1293   10.7055 
```

```r
#Test 3: Infant vs Male
t_test_IM <- t.test(infant_rings,male_rings)
print(t_test_IM)
```

```

    Welch Two Sample t-test

data:  infant_rings and male_rings
t = -27.221, df = 2859, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.017808 -2.612263
sample estimates:
mean of x mean of y 
 7.890462 10.705497 
```

```r
#analysis of p-value

#Case 1:
if (t_test_FI$p.value<0.05) {
  cat("There is a statistically significant difference between Female and Infant ring counts
} else {
  cat("There is no statistically significant difference between Female and Infant ring count
}
```

```
There is a statistically significant difference between Female and Infant ring counts.
```

```
#Case 2:
if (t_test_FM$p.value<0.05) {
  cat("There is a statistically significant difference between Female and Male ring counts\n
} else {
  cat("There is no statistically significant difference between Female and Male ring counts.
}
```

```
There is a statistically significant difference between Female and Male ring counts
```

```
#Case 3:
if (t_test_IM$p.value<0.05) {
  cat("There is a statistically significant difference between Infant and Male ring counts.\
} else {
  cat("There is no statistically significant difference between Infant and Male ring counts.
}
```

```
There is a statistically significant difference between Infant and Male ring counts.
```

Problem 2

2a

```
#import dataset
food_data <- read.csv("C:/Users/hyperion/University of Michigan Dropbox/Deepan Islam/Fall 20

(food_data)
```

```
     ID What.is.your.age.
1     1                68
2     2                88
3     3                82
4     4                73
5     5                89
6     6                18
7     7                38
8     8                28
```

9	9	16
10	10	84
11	11	46
12	12	29
13	13	90
14	14	22
15	15	70
16	16	3
17	17	31
18	18	23
19	19	86
20	20	81
21	21	51
22	22	70
23	23	39
24	24	79
25	25	33
26	26	19
27	27	47
28	28	42
29	29	48
30	30	85
31	31	59
32	32	48
33	33	37
34	34	58
35	35	79
36	36	75
37	37	150
38	38	58
39	39	23
40	40	61
41	41	68
42	42	58
43	43	62
44	44	77
45	45	46
46	46	51
47	47	30
48	51	21
49	52	31
50	53	27
51	54	67

```
52   55                 24
53   56                 15
54   57                 38
55   58                 67
56   59                 42
57   60                 65
58   61                 22
59   62                 21
60   64                 74
61   65                 34
62   66                 44
63   67                 81
64   68                 73
65   69                 65
66   70                 74
67   71                 NA
68   72                 66
69   73                 42
70   74                 18
71   75                150
72   77                 70
73   78                 83
74   79                 71
75   80                 61
76   81                 69
77   82                 45
78   83                 18
79   84                 NA
80   85                 80
81   86                 34
82   87                 66
83   88                 84
84   89                 27
85   90                 81
86   92                  3
87   93                 19
88   94                  7
89   95                 30
90   96                  6
91   97                 59
92   98                 17
93   99                 85
94  100                150
```

```
95  101                 20
96  102                  8
97  103                 43
98  104                 46
99  105                 81
100 106                 79
101 108                 17
102 109                 34
103 110                 83
104 111                 63
105 112                 88
106 113                 77
107 114                 35
108 115                 47
109 116                 23
110 117                 57
111 118                 49
112 119                 76
113 120                  8
114 121                 30
115 122                 51
116 123                 80
117 124                 28
118 125                 19
119 126                 45
120 127                  6
121 130                 30
122 131                 27
123 132                 77
124 133                 78
125 134                 85
126 135                 19
127 136                  2
128 137                 31
129 138                 28
130 139                 74
131 140                 19
132 141                 60
133 142                 82
134 143                 73
135 144                 NA
136 145                 67
137 146                 58
```

```
138 147                67
139 149                44
140 150                42
141 151                51
142 152                61
143 153                61
144 154                48
145 155                15
146 156                46
147 157                58
148 161                67
149 163                64
150 164                65
151 166                90
152 167                30
153 168                 8
154 169                83
155 170                40
156 171                84
157 172                31
158 173                33
159 174                NA
160 175                39
161 176                16
162 177                90
163 178                68
164 179                49
165 180                52
166 181                53
167 182                18
168 183                47
169 184                47
170 185                90
171 187                52
172 188                66
173 189                NA
174 190                16
175 191                68
176 192                54
177 193                NA
178 194                68
179 195                89
180 197                86
```

```
181 198                30
182 199                90
183 200                 4
184 201                77
185 202                52
186 203                17
187 204                NA
188 205                54
189 207                64
190 208                54
191 209                68
192 210                70
193 211                 3
194 212                50
195 213                16
196 215                38
197 217                56
198 218                 3
199 219                42
200 220                51
201 221                85
202 222                21
203 224                26
204 225                28
205 226                38
206 227                 5
207 228                51
208 230                83
209 231                74
210 232                76
211 233                77
212 234                25
213 235                45
214 236                37
215 237                83
216 238                71
217 239                72
218 240                 4
219 241                16
220 242                NA
221 243                78
222 244                77
223 245                16
```

```
224 246                18
225 247                48
226 248                54
227 249                23
228 250                82
229 252                45
230 253                26
231 254                58
232 255                81
233 256                48
234 257                32
235 258                44
236 259                16
237 260                88
238 261                81
239 262                44
240 263                44
241 264                87
242 265                40
243 266                57
244 267                76
245 268                89
246 269                50
247 270                36
248 271                84
249 272                35
250 273                54
251 274                85
252 275                83
253 276                79
254 277                78
255 278                55
256 279                34
257 280                47
258 281                29
259 283                 6
260 284                NA
261 286                33
262 287                59
    How.many.individuals.live.in.your.household.for.which.you.are.responsible.for.food.expend
1
2
3
```

```
262
    What.state.do.you.live.in.
1                           LA
2                           WA
3                           MS
4                           AK
5                           IN
6                           WI
7                           DC
8                           ID
9                           SC
10                          HI
11                          ND
12                          UT
13                          DC
14                          NV
15                          WA
16                          ME
17                          WI
18                          VA
19                          RI
20                          MT
21                          AK
22                          IL
23                          PA
24                          OR
25                          WA
26                          CT
27                          GA
28                          NV
29                          MT
30                          NC
31                          TX
32                          NE
33                          NY
34                          CO
35                          DC
36                          WY
37                          IL
38                          IA
39                          NV
40                          DC
41                          FL
```

42	SC
43	AZ
44	NM
45	WA
46	AL
47	MI
48	NV
49	IL
50	MN
51	NV
52	WI
53	WV
54	
55	OH
56	LA
57	NJ
58	OK
59	ME
60	RI
61	AL
62	WV
63	FL
64	CA
65	
66	AL
67	XX
68	NE
69	KS
70	UT
71	PA
72	WY
73	VT
74	DE
75	XX
76	IL
77	PA
78	XX
79	ND
80	MN
81	KY
82	ND
83	CO
84	IL

85	PR
86	AL
87	MO
88	PR
89	TX
90	NE
91	MA
92	HI
93	NH
94	IA
95	CA
96	TN
97	KS
98	MA
99	IL
100	MD
101	IL
102	MS
103	AZ
104	VA
105	OH
106	TN
107	PA
108	XX
109	VT
110	WA
111	AL
112	AK
113	CO
114	OR
115	CA
116	CT
117	PR
118	DE
119	GA
120	GA
121	OR
122	LA
123	CT
124	LA
125	SC
126	IA
127	NC

128	CT
129	MA
130	VT
131	MD
132	MD
133	VT
134	AL
135	AZ
136	WV
137	TX
138	
139	MA
140	ID
141	MS
142	PA
143	RI
144	OH
145	OK
146	WI
147	AL
148	VT
149	MI
150	ID
151	
152	MD
153	NH
154	WI
155	IA
156	WI
157	CO
158	FL
159	LA
160	AZ
161	TX
162	IN
163	PR
164	
165	WY
166	AZ
167	DE
168	AL
169	VT
170	AK

171	TN
172	NJ
173	OK
174	ID
175	ME
176	NY
177	DC
178	ID
179	ID
180	MT
181	VA
182	VT
183	PA
184	DC
185	KS
186	WV
187	WY
188	SC
189	FL
190	WI
191	VT
192	MN
193	MT
194	WI
195	NM
196	AZ
197	OH
198	NC
199	PA
200	NJ
201	FL
202	MS
203	WI
204	AZ
205	WI
206	SD
207	CO
208	WV
209	MS
210	SC
211	WI
212	RI
213	WA

214	IA
215	IA
216	VT
217	AZ
218	NY
219	TX
220	ID
221	SD
222	WA
223	DC
224	XX
225	OR
226	LA
227	AR
228	LA
229	WV
230	WV
231	DC
232	PA
233	ND
234	KS
235	RI
236	GA
237	MS
238	WY
239	IN
240	CT
241	OK
242	PR
243	UT
244	XX
245	NY
246	DC
247	MS
248	KY
249	KS
250	MN
251	TN
252	AR
253	MN
254	VA
255	MA
256	LA

```
257                         ND
258                         KY
259                         NC
260                         AK
261                         AZ
262                         CA
    What.currency.are.you.reporting.your.food.expenditures.in.
1                                                          USD
2                                                          USD
3                                                          USD
4                                                          USD
5                                                          USD
6                                                          EUR
7                                                          USD
8                                                          USD
9                                                          USD
10                                                         USD
11                                                         USD
12                                                         USD
13                                                         USD
14                                                         USD
15                                                         USD
16                                                         USD
17                                                         USD
18                                                         USD
19                                                         USD
20                                                         USD
21                                                         USD
22                                                         USD
23                                                         EUR
24                                                         USD
25                                                         USD
26                                                         USD
27                                                         USD
28                                                         EUR
29                                                         USD
30                                                         USD
31                                                         USD
32                                                         USD
33                                                         USD
34                                                         USD
35                                                         USD
36                                                         USD
```

37	USD
38	CAD
39	USD
40	CAD
41	USD
42	CAD
43	USD
44	USD
45	EUR
46	CAD
47	USD
48	USD
49	USD
50	USD
51	USD
52	USD
53	USD
54	USD
55	USD
56	USD
57	CAD
58	USD
59	USD
60	USD
61	USD
62	USD
63	CAD
64	USD
65	USD
66	USD
67	USD
68	CAD
69	USD
70	USD
71	USD
72	USD
73	USD
74	CAD
75	USD
76	USD
77	USD
78	USD
79	USD

80	USD
81	EUR
82	USD
83	USD
84	USD
85	USD
86	EUR
87	USD
88	USD
89	USD
90	USD
91	CAD
92	USD
93	USD
94	USD
95	USD
96	USD
97	CAD
98	USD
99	USD
100	USD
101	CAD
102	USD
103	USD
104	USD
105	USD
106	USD
107	USD
108	USD
109	USD
110	USD
111	USD
112	USD
113	EUR
114	USD
115	USD
116	USD
117	USD
118	USD
119	USD
120	USD
121	USD
122	USD

123	USD
124	USD
125	USD
126	USD
127	USD
128	USD
129	USD
130	USD
131	USD
132	USD
133	USD
134	USD
135	USD
136	USD
137	USD
138	USD
139	USD
140	USD
141	CAD
142	USD
143	USD
144	USD
145	USD
146	CAD
147	EUR
148	USD
149	USD
150	USD
151	USD
152	USD
153	USD
154	USD
155	CAD
156	USD
157	USD
158	USD
159	USD
160	USD
161	USD
162	USD
163	USD
164	USD
165	USD

166	USD
167	EUR
168	USD
169	USD
170	USD
171	USD
172	USD
173	USD
174	USD
175	CAD
176	USD
177	USD
178	USD
179	USD
180	USD
181	USD
182	USD
183	EUR
184	USD
185	USD
186	USD
187	USD
188	USD
189	USD
190	USD
191	USD
192	USD
193	USD
194	USD
195	USD
196	USD
197	USD
198	USD
199	USD
200	USD
201	USD
202	USD
203	USD
204	CAD
205	USD
206	USD
207	USD
208	USD

209	USD
210	USD
211	USD
212	USD
213	USD
214	USD
215	USD
216	USD
217	CAD
218	USD
219	USD
220	USD
221	USD
222	USD
223	USD
224	USD
225	USD
226	USD
227	USD
228	USD
229	USD
230	CAD
231	USD
232	USD
233	USD
234	USD
235	USD
236	USD
237	USD
238	USD
239	CAD
240	USD
241	USD
242	USD
243	USD
244	USD
245	USD
246	USD
247	USD
248	USD
249	USD
250	USD
251	USD

```
252                                                        USD
253                                                        USD
254                                                        CAD
255                                                        USD
256                                                        USD
257                                                        USD
258                                                        USD
259                                                        CAD
260                                                        USD
261                                                        USD
262                                                        EUR
    What.was.your.total.food.expenditure.in.the.last.week.
1                                                   436.35
2                                                         
3                                                    279.1
4                                                   -20.98
5                                                   494.87
6                                                   276.32
7                                                   318.79
8                                                   304.52
9                                                   325.71
10                                                  332.08
11                                                 -201.52
12                                                  622.58
13                                                  292.08
14                                                  505.11
15                                                  311.84
16                                                  555.39
17                                                  529.38
18                                                   404.6
19                                                  561.31
20                                                   604.2
21                                                  794.98
22                                                  284.72
23                                                  184.15
24                                                  789.86
25                                                  141.85
26                                                  865.36
27                                                  457.64
28                                                  762.41
29                                                  514.22
30                                                  431.93
31                                                   44.13
```

32	~350
33	-25
34	398.68
35	477.72
36	346.09
37	-25
38	852.93
39	366
40	156.05
41	113.97
42	366.45
43	358.13
44	654.01
45	364.97
46	
47	0.46
48	0
49	574.63
50	481.04
51	-25
52	257.59
53	537.65
54	434.91
55	427.24
56	-25
57	200.74
58	253.33
59	151.63
60	331.69
61	377.16
62	808.2
63	299.44
64	903.34
65	300.85
66	554.75
67	0
68	438.77
69	10.69
70	-25
71	182.65
72	659.87
73	779.49
74	596.91

75	350.89
76	-113.7
77	0
78	
79	631.88
80	338.99
81	346.87
82	510.02
83	291.98
84	476.51
85	669.93
86	360.95
87	232.01
88	796.02
89	68.05
90	
91	468.61
92	312.88
93	628.59
94	474.25
95	554.42
96	-28.59
97	375.79
98	358.42
99	493.97
100	674.45
101	412.68
102	357.18
103	289.94
104	
105	38.9
106	80.85
107	348.55
108	602.64
109	816.52
110	43.14
111	-51.81
112	365.79
113	222.67
114	446.7
115	-25
116	479.69
117	-25

118	0
119	831.67
120	360.1
121	0
122	763.08
123	-25
124	152.31
125	301.83
126	191.01
127	445.18
128	428.16
129	594.57
130	-88.1
131	701.31
132	394.33
133	980.51
134	1049.19
135	9999999
136	469.82
137	273.56
138	93.39
139	0
140	519.49
141	279.85
142	866.78
143	244.06
144	271.92
145	288.34
146	172.83
147	
148	458.84
149	517.2
150	125.16
151	277.78
152	159.28
153	469.54
154	177.01
155	638.23
156	831.7
157	0
158	222.07
159	339.24
160	447.78

161	454.66
162	
163	419.23
164	102.96
165	-28.99
166	520.64
167	191.02
168	476.38
169	575.32
170	149.26
171	730.91
172	545.7
173	629.39
174	430.66
175	472.25
176	523.62
177	281.21
178	299.39
179	628.22
180	926.01
181	408.82
182	109.44
183	663.37
184	394.14
185	
186	-31.23
187	418.79
188	337.2
189	0
190	24.86
191	469.73
192	
193	914.78
194	503.33
195	330.42
196	504.49
197	706.36
198	0
199	338.39
200	13.88
201	526.39
202	458.28
203	296.91

204	157.65
205	439.92
206	554.18
207	407.43
208	413.68
209	206.1
210	702.22
211	614.19
212	503.78
213	99.47
214	194.13
215	440.67
216	553.9
217	249.49
218	
219	570.85
220	334.96
221	372.02
222	449.52
223	695.13
224	829.06
225	677.2
226	645.8
227	531.27
228	318.53
229	326.43
230	283.38
231	171.17
232	191.14
233	-101.25
234	285.71
235	619.39
236	687.45
237	393.96
238	284.15
239	346.25
240	804.71
241	0
242	405.04
243	797.61
244	180.68
245	298.1
246	643.22

```
247                                                      0
248                                                   84.8
249                                                 350.91
250                                                 524.49
251                                                  394.3
252                                                      0
253                                                  455.1
254                                                 864.72
255                                                 566.29
256                                                  70.29
257                                                 316.39
258                                                 635.88
259                                                 434.91
260                                                 423.81
261                                                      0
262                                                 469.17
    What.was.your.total.food.expenditures.at.grocery.stores.in.the.last.week.
1                                                                      168.59
2                                                                      452.10
3                                                                      301.66
4                                                                      139.66
5                                                                          NA
6                                                                      394.44
7                                                                      153.49
8                                                                      286.70
9                                                                      484.22
10                                                                     236.68
11                                                                      40.54
12                                                                     144.16
13                                                                     168.88
14                                                                     381.19
15                                                                     212.63
16                                                                     280.91
17                                                                     139.08
18                                                                     243.86
19                                                                     444.56
20                                                                     254.94
21                                                                     409.40
22                                                                     197.66
23                                                                     280.08
24                                                                     269.49
25                                                                     265.90
26                                                                     438.04
```

27	220.24
28	266.37
29	105.19
30	244.75
31	347.12
32	227.74
33	66.24
34	186.02
35	237.40
36	429.12
37	64.36
38	216.58
39	248.39
40	388.40
41	139.72
42	79.78
43	161.24
44	278.24
45	458.23
46	350.09
47	198.30
48	175.13
49	135.49
50	381.95
51	373.16
52	62.73
53	226.99
54	253.20
55	34.22
56	155.34
57	213.79
58	360.63
59	184.29
60	172.17
61	123.85
62	453.84
63	152.47
64	535.71
65	215.11
66	101.25
67	6.92
68	424.70
69	429.88

70	312.83
71	403.44
72	77.55
73	226.23
74	74.52
75	155.17
76	189.40
77	192.51
78	167.49
79	234.08
80	392.50
81	196.75
82	655.63
83	234.25
84	287.35
85	378.69
86	214.35
87	582.38
88	216.26
89	230.87
90	452.31
91	141.87
92	462.08
93	190.77
94	21.99
95	380.56
96	327.06
97	112.27
98	195.66
99	313.15
100	249.78
101	327.76
102	181.17
103	26.66
104	313.34
105	231.06
106	304.51
107	222.32
108	250.90
109	233.00
110	289.77
111	273.09
112	217.12

113	189.21
114	67.61
115	64.35
116	155.32
117	103.58
118	255.80
119	223.32
120	NA
121	174.57
122	131.63
123	24.90
124	NA
125	236.43
126	140.44
127	NA
128	308.43
129	322.71
130	311.37
131	292.02
132	75.19
133	399.70
134	335.26
135	341.89
136	288.69
137	121.73
138	305.19
139	99.79
140	138.13
141	133.26
142	136.25
143	106.52
144	143.34
145	292.15
146	191.59
147	165.44
148	323.71
149	181.40
150	9.45
151	261.15
152	51.79
153	53.47
154	458.42
155	523.44

156	154.98
157	25.42
158	27.85
159	392.50
160	155.02
161	100.61
162	356.81
163	51.60
164	167.14
165	240.29
166	271.85
167	282.43
168	378.96
169	167.92
170	-73.48
171	231.92
172	40.75
173	356.17
174	222.54
175	347.08
176	154.13
177	102.73
178	223.92
179	211.46
180	149.17
181	229.45
182	115.02
183	318.27
184	198.25
185	167.71
186	269.22
187	308.26
188	207.84
189	304.87
190	230.58
191	362.11
192	164.08
193	283.95
194	448.04
195	130.88
196	389.02
197	117.70
198	421.68

199	358.79
200	390.23
201	300.39
202	143.68
203	415.10
204	384.96
205	129.28
206	651.01
207	155.91
208	239.02
209	416.60
210	354.07
211	182.34
212	248.54
213	168.89
214	122.03
215	310.17
216	305.57
217	239.71
218	164.02
219	104.97
220	NA
221	14.68
222	25.66
223	285.61
224	353.68
225	193.65
226	502.64
227	191.46
228	301.81
229	607.20
230	187.75
231	500.33
232	62.87
233	NA
234	292.33
235	110.08
236	191.32
237	238.79
238	320.24
239	345.89
240	90.24
241	119.05

```
242                                                                    286.88
243                                                                     12.54
244                                                                    358.72
245                                                                    439.47
246                                                                    296.92
247                                                                    154.38
248                                                                    361.79
249                                                                     85.39
250                                                                    150.62
251                                                                    209.59
252                                                                     61.57
253                                                                    190.14
254                                                                    407.74
255                                                                    360.76
256                                                                    270.44
257                                                                    517.47
258                                                                    658.51
259                                                                    162.45
260                                                                     15.22
261                                                                    387.72
262                                                                     37.74
    What.was.your.food.expenditure.while.dining.out.in.the.last.week.
1                                                              140.71
2                                                              192.94
3                                                              239.84
4                                                               69.19
5                                                              191.72
6                                                              283.20
7                                                              104.05
8                                                                  NA
9                                                              289.89
10                                                             105.59
11                                                              10.57
12                                                              58.50
13                                                              64.77
14                                                             121.20
15                                                              93.26
16                                                              63.78
17                                                                 NA
18                                                             159.71
19                                                             186.67
20                                                             152.27
21                                                             123.04
```

22	122.72
23	149.95
24	199.48
25	58.71
26	148.28
27	63.68
28	142.55
29	48.59
30	130.94
31	58.75
32	78.38
33	30.79
34	123.60
35	NA
36	240.54
37	13.93
38	95.28
39	61.93
40	162.57
41	67.64
42	39.53
43	118.86
44	185.93
45	186.53
46	81.04
47	56.36
48	101.38
49	99.38
50	207.97
51	184.65
52	34.88
53	131.19
54	45.38
55	21.13
56	98.89
57	175.49
58	103.60
59	163.52
60	76.16
61	66.73
62	134.64
63	70.78
64	177.89

65	162.36
66	32.45
67	4.02
68	104.74
69	114.11
70	174.82
71	302.01
72	25.79
73	145.32
74	51.07
75	32.77
76	87.00
77	60.96
78	67.61
79	54.77
80	190.88
81	126.31
82	116.01
83	48.84
84	142.44
85	263.26
86	37.66
87	224.92
88	89.80
89	136.63
90	NA
91	89.95
92	217.19
93	71.82
94	4.71
95	147.83
96	98.62
97	56.91
98	130.40
99	149.92
100	200.70
101	158.38
102	130.43
103	18.79
104	73.57
105	72.45
106	194.44
107	73.81

108	125.99
109	98.00
110	115.19
111	74.79
112	58.61
113	64.20
114	13.85
115	40.28
116	100.28
117	18.38
118	49.53
119	85.16
120	8.09
121	128.09
122	63.33
123	12.24
124	65.87
125	143.16
126	61.75
127	47.02
128	226.07
129	214.61
130	215.30
131	102.61
132	21.87
133	136.78
134	95.41
135	150.29
136	152.09
137	44.72
138	61.04
139	51.68
140	63.15
141	60.22
142	75.12
143	38.18
144	91.71
145	96.25
146	100.21
147	43.97
148	168.76
149	86.95
150	4.33

151	131.70
152	39.51
153	11.10
154	137.36
155	115.95
156	138.23
157	10.64
158	13.32
159	157.53
160	123.40
161	67.30
162	213.07
163	42.67
164	100.07
165	63.20
166	68.08
167	198.83
168	146.60
169	47.37
170	NA
171	42.75
172	19.38
173	127.45
174	136.85
175	157.12
176	50.25
177	12.20
178	168.69
179	92.17
180	66.20
181	133.52
182	32.37
183	129.63
184	109.50
185	51.24
186	151.78
187	99.96
188	107.22
189	87.67
190	118.06
191	202.42
192	80.31
193	90.40

194	191.17
195	99.86
196	102.25
197	57.16
198	254.45
199	209.13
200	187.16
201	178.98
202	66.84
203	152.52
204	178.98
205	61.36
206	341.93
207	40.42
208	88.30
209	171.82
210	110.85
211	159.21
212	40.31
213	66.95
214	83.11
215	108.71
216	70.33
217	189.22
218	141.08
219	38.44
220	52.73
221	5.78
222	6.59
223	136.73
224	256.65
225	100.05
226	231.80
227	91.23
228	106.93
229	221.89
230	80.25
231	335.12
232	22.55
233	63.64
234	59.80
235	24.62
236	69.49

```
237                                                            127.89
238                                                            192.50
239                                                             83.07
240                                                             19.23
241                                                             40.16
242                                                            111.79
243                                                              5.53
244                                                            172.49
245                                                             88.47
246                                                            234.98
247                                                             90.03
248                                                            210.05
249                                                             74.27
250                                                             32.44
251                                                            115.50
252                                                             22.05
253                                                            106.51
254                                                            283.14
255                                                            186.40
256                                                                NA
257                                                            111.68
258                                                            238.98
259                                                             55.88
260                                                              9.26
261                                                            106.12
262                                                             13.59
    What.was.your.food.expenditure..miscellaneous..in.the.last.week.
1                                                             109.77
2                                                                 NA
3                                                             103.94
4                                                              44.84
5                                                             172.31
6                                                             114.06
7                                                              39.21
8                                                              24.61
9                                                             145.01
10                                                             38.86
11                                                             40.24
12                                                             14.73
13                                                             29.12
14                                                             76.45
15                                                            111.79
16                                                                NA
```

```
17                                                            24.01
18                                                            94.78
19                                                               NA
20                                                           145.16
21                                                           132.14
22                                                            66.61
23                                                            76.74
24                                                           154.01
25                                                           203.34
26                                                               NA
27                                                           181.99
28                                                           184.82
29                                                            47.80
30                                                           131.07
31                                                           146.84
32                                                            55.25
33                                                               NA
34                                                           153.23
35                                                            32.02
36                                                           208.01
37                                                            59.46
38                                                           125.48
39                                                           123.18
40                                                            81.65
41                                                            17.44
42                                                            52.27
43                                                            33.27
44                                                            77.01
45                                                           135.61
46                                                           327.59
47                                                           184.63
48                                                           155.85
49                                                            20.42
50                                                            57.90
51                                                           -22.47
52                                                            52.37
53                                                           -19.81
54                                                            70.11
55                                                             5.64
56                                                            80.35
57                                                            94.52
58                                                            -0.02
59                                                            66.11
```

60	60.66
61	53.43
62	NA
63	103.36
64	NA
65	127.87
66	30.16
67	5.22
68	89.23
69	NA
70	-1.68
71	64.67
72	10.18
73	44.29
74	19.85
75	128.25
76	196.43
77	NA
78	47.66
79	NA
80	257.16
81	158.98
82	128.92
83	160.69
84	256.41
85	49.67
86	82.36
87	NA
88	63.43
89	72.37
90	113.17
91	99.39
92	7.54
93	80.38
94	15.93
95	62.63
96	89.20
97	23.37
98	43.42
99	238.45
100	149.13
101	NA
102	93.66

103	10.31
104	124.93
105	NA
106	78.32
107	99.41
108	120.22
109	196.39
110	81.82
111	83.01
112	NA
113	100.95
114	40.32
115	23.40
116	106.35
117	29.25
118	64.36
119	77.38
120	15.91
121	19.06
122	-6.85
123	9.10
124	45.97
125	186.30
126	120.73
127	141.64
128	96.81
129	209.08
130	206.58
131	232.53
132	11.86
133	316.47
134	311.18
135	89.82
136	48.47
137	105.44
138	38.13
139	13.55
140	34.12
141	11.06
142	20.63
143	69.52
144	8.83
145	41.77

146	123.31
147	69.92
148	96.45
149	NA
150	4.09
151	16.93
152	18.66
153	21.34
154	146.98
155	131.59
156	78.54
157	NA
158	13.44
159	298.49
160	108.53
161	34.32
162	77.22
163	30.30
164	-1.66
165	49.79
166	130.76
167	107.13
168	16.33
169	NA
170	NA
171	76.35
172	39.40
173	181.37
174	NA
175	NA
176	24.81
177	NA
178	135.80
179	65.54
180	5.44
181	125.36
182	96.87
183	51.01
184	31.98
185	37.59
186	74.02
187	145.00
188	-15.63

189	86.73
190	73.83
191	-22.92
192	58.32
193	269.16
194	3.50
195	36.73
196	228.84
197	41.69
198	39.58
199	225.76
200	-7.62
201	75.05
202	44.52
203	NA
204	58.50
205	NA
206	134.28
207	178.87
208	118.91
209	100.48
210	206.87
211	62.81
212	97.45
213	149.65
214	58.90
215	40.97
216	73.73
217	154.34
218	89.56
219	55.39
220	126.68
221	NA
222	NA
223	109.09
224	179.00
225	115.74
226	93.19
227	6.84
228	235.96
229	NA
230	67.31
231	190.55

```
232                                                            49.63
233                                                            57.85
234                                                           208.02
235                                                           109.15
236                                                           163.87
237                                                            99.20
238                                                            24.91
239                                                            79.68
240                                                               NA
241                                                            24.78
242                                                           198.15
243                                                             4.63
244                                                           -15.50
245                                                               NA
246                                                           109.39
247                                                            23.13
248                                                            -5.97
249                                                            33.44
250                                                               NA
251                                                           123.41
252                                                            54.88
253                                                           123.85
254                                                           122.20
255                                                            96.85
256                                                            62.93
257                                                               NA
258                                                               NA
259                                                            97.09
260                                                             3.81
261                                                               NA
262                                                            10.23
    How.many.times.did.you.dine.out.last.week.
1                                            4
2                                            1
3                                            9
4                                            2
5                                            3
6                                            6
7                                            1
8                                           10
9                                            1
10                                           9
11                                           1
```

12	5
13	8
14	8
15	30
16	0
17	2
18	10
19	6
20	10
21	1
22	2
23	5
24	9
25	9
26	10
27	1
28	9
29	15
30	20
31	8
32	7
33	7
34	10
35	20
36	5
37	6
38	8
39	30
40	9
41	7
42	5
43	3
44	30
45	1
46	2
47	2
48	4
49	9
50	1
51	2
52	7
53	20
54	0

55	10
56	15
57	10
58	3
59	8
60	15
61	0
62	10
63	2
64	15
65	4
66	4
67	3
68	7
69	5
70	8
71	9
72	9
73	10
74	7
75	9
76	10
77	2
78	7
79	6
80	3
81	4
82	6
83	3
84	4
85	7
86	3
87	5
88	2
89	6
90	3
91	6
92	3
93	15
94	9
95	10
96	7
97	5

98	3
99	1
100	5
101	10
102	9
103	3
104	3
105	5
106	5
107	1
108	10
109	4
110	1
111	4
112	15
113	5
114	20
115	15
116	7
117	8
118	4
119	5
120	5
121	4
122	8
123	7
124	10
125	1
126	4
127	5
128	7
129	5
130	8
131	8
132	8
133	20
134	0
135	9
136	4
137	9
138	15
139	6
140	4

141	6
142	15
143	2
144	3
145	2
146	6
147	4
148	5
149	7
150	7
151	7
152	7
153	5
154	5
155	0
156	2
157	10
158	9
159	20
160	10
161	15
162	0
163	20
164	8
165	10
166	15
167	3
168	7
169	2
170	4
171	10
172	2
173	15
174	8
175	6
176	3
177	6
178	15
179	2
180	4
181	4
182	15
183	3

184	1
185	4
186	20
187	8
188	0
189	15
190	20
191	7
192	15
193	2
194	1
195	4
196	2
197	7
198	8
199	15
200	6
201	9
202	15
203	2
204	8
205	10
206	10
207	6
208	7
209	0
210	0
211	8
212	4
213	1
214	4
215	10
216	1
217	15
218	15
219	6
220	2
221	5
222	4
223	8
224	8
225	2
226	3

```
227                                           4
228                                           8
229                                           0
230                                           7
231                                          20
232                                           2
233                                           6
234                                           8
235                                           1
236                                           5
237                                          20
238                                           8
239                                           7
240                                           4
241                                           9
242                                          20
243                                          15
244                                           9
245                                           0
246                                           1
247                                          20
248                                           9
249                                           6
250                                          20
251                                           8
252                                           4
253                                           4
254                                           9
255                                          15
256                                          20
257                                           2
258                                           6
259                                           9
260                                           3
261                                           2
262                                           0
    Are.you.including.alcohol.in.your.food.expenditures.
1                                                    Yes
2                                                Unknown
3                                                    Yes
4                                                Unknown
5                                                    Yes
6                                                Unknown
```

7	Yes
8	No
9	
10	Unknown
11	No
12	No
13	Yes
14	No
15	Yes
16	N
17	Unknown
18	No
19	Unknown
20	No
21	No
22	No
23	Unknown
24	Yes
25	No
26	Yes
27	Yes
28	No
29	Yes
30	Yes
31	Yes
32	No
33	No
34	Yes
35	Yes
36	Unknown
37	
38	N
39	Yes
40	Yes
41	
42	No
43	Y
44	
45	
46	Yes
47	No
48	No
49	No

50	Yes
51	N
52	No
53	Yes
54	No
55	Yes
56	Yes
57	No
58	No
59	Yes
60	Yes
61	No
62	Yes
63	Y
64	Yes
65	Yes
66	Yes
67	No
68	Unknown
69	No
70	Yes
71	Yes
72	Yes
73	N
74	Yes
75	No
76	Yes
77	Yes
78	Yes
79	No
80	No
81	
82	No
83	No
84	Yes
85	Yes
86	Yes
87	Unknown
88	No
89	Unknown
90	N
91	Yes
92	No

93	Unknown
94	No
95	Y
96	Unknown
97	
98	No
99	Unknown
100	
101	Unknown
102	
103	No
104	Unknown
105	No
106	No
107	No
108	No
109	No
110	No
111	No
112	Yes
113	Y
114	Yes
115	No
116	
117	No
118	Unknown
119	Yes
120	Unknown
121	No
122	No
123	No
124	Yes
125	Yes
126	No
127	Unknown
128	No
129	No
130	Yes
131	No
132	N
133	Yes
134	Yes
135	No

136	Yes
137	Yes
138	Y
139	No
140	Yes
141	Yes
142	No
143	Yes
144	N
145	Unknown
146	Unknown
147	No
148	Y
149	Yes
150	No
151	N
152	No
153	Unknown
154	No
155	N
156	No
157	
158	Yes
159	Yes
160	Yes
161	No
162	Yes
163	Yes
164	No
165	
166	
167	No
168	No
169	Yes
170	Yes
171	No
172	Yes
173	No
174	No
175	Yes
176	N
177	No
178	

179	No
180	No
181	No
182	N
183	Unknown
184	Yes
185	No
186	No
187	Yes
188	No
189	No
190	Yes
191	Yes
192	
193	Yes
194	N
195	No
196	Yes
197	Yes
198	No
199	No
200	Yes
201	No
202	Yes
203	
204	No
205	Unknown
206	Yes
207	No
208	No
209	Yes
210	N
211	No
212	No
213	No
214	Yes
215	No
216	Y
217	Yes
218	No
219	No
220	No
221	Yes

```
222                                                   Yes
223                                                   Yes
224                                                   Yes
225                                                   Yes
226                                                    No
227
228                                                   Yes
229                                                   Yes
230                                                   Yes
231                                                    No
232                                                   Yes
233                                                   Yes
234                                                     N
235                                                   Yes
236                                               Unknown
237                                                    No
238                                                    No
239                                                   Yes
240                                                   Yes
241                                                   Yes
242                                                    No
243                                                    No
244                                                    No
245                                                    No
246                                                   Yes
247                                                    No
248                                                   Yes
249                                                   Yes
250                                                   Yes
251                                                   Yes
252                                               Unknown
253                                                    No
254                                                   Yes
255                                                   Yes
256                                                   Yes
257                                                    No
258                                                    No
259                                               Unknown
260                                                   Yes
261                                               Unknown
262                                                   Yes
    What.food.assistance.programs..if.any..did.you.use.for.your.food.expenditures.last.week.
1                                                                                       None
```

2	SNAP
3	None
4	None
5	None
6	Food Pantry
7	SNAP
8	None
9	WIC
10	None
11	None
12	None
13	SNAP
14	None
15	SNAP
16	None
17	None
18	None
19	Food Pantry
20	None
21	Food Pantry
22	None
23	None
24	School Meals
25	None
26	Food Pantry
27	None
28	None
29	None
30	Food Pantry
31	SNAP
32	None
33	None
34	SNAP
35	None
36	WIC
37	None
38	None
39	None
40	?
41	None
42	SNAP
43	None
44	None

45	None
46	None
47	None
48	None
49	None
50	None
51	None
52	SNAP
53	None
54	None
55	Food Pantry
56	None
57	None
58	School Meals
59	None
60	None
61	School Meals
62	None
63	None
64	None
65	None
66	None
67	None
68	Food Pantry
69	None
70	?
71	None
72	?
73	WIC
74	None
75	None
76	None
77	None
78	School Meals
79	None
80	Food Pantry
81	None
82	WIC
83	None
84	None
85	None
86	Food Pantry
87	None

```
88                School Meals
89                        None
90                        None
91                        None
92                School Meals
93                        None
94                        None
95                        None
96                        None
97                           ?
98                        None
99                        SNAP
100                       None
101                       None
102                       SNAP
103                       None
104                Food Pantry
105                       None
106                       None
107                       None
108               School Meals
109                       None
110                       None
111                       None
112                       None
113                       None
114                       None
115                       None
116                       None
117                       None
118                       None
119                       None
120                          ?
121                       None
122                Food Pantry
123                       SNAP
124                       None
125                       None
126                       None
127                          ?
128                       None
129                       None
130                       None
```

131	None
132	Food Pantry
133	None
134	None
135	None
136	None
137	School Meals
138	None
139	None
140	None
141	None
142	None
143	None
144	Food Pantry
145	None
146	SNAP
147	None
148	None
149	None
150	SNAP
151	None
152	SNAP
153	None
154	None
155	None
156	School Meals
157	None
158	SNAP
159	None
160	WIC
161	None
162	None
163	Food Pantry
164	None
165	?
166	None
167	None
168	None
169	None
170	Food Pantry
171	WIC
172	?
173	None

174	None
175	None
176	Food Pantry
177	None
178	None
179	None
180	School Meals
181	WIC
182	None
183	None
184	None
185	Food Pantry
186	None
187	None
188	None
189	None
190	None
191	?
192	None
193	None
194	None
195	None
196	None
197	None
198	SNAP
199	None
200	School Meals
201	Food Pantry
202	None
203	SNAP
204	None
205	None
206	None
207	None
208	WIC
209	None
210	SNAP
211	WIC
212	None
213	None
214	None
215	None
216	SNAP

217	Food Pantry
218	None
219	Food Pantry
220	None
221	SNAP
222	None
223	None
224	None
225	SNAP
226	?
227	None
228	None
229	None
230	Food Pantry
231	None
232	None
233	None
234	None
235	None
236	None
237	None
238	None
239	None
240	None
241	None
242	SNAP
243	None
244	WIC
245	None
246	None
247	School Meals
248	None
249	None
250	SNAP
251	Food Pantry
252	None
253	SNAP
254	None
255	None
256	SNAP
257	None
258	SNAP
259	None

```
260                                                                                     None
261                                                                                     None
262                                                                                     None
```

2b

Clean up the variable names. Simplify them.

```
#cleanup the variable names

colnames(food_data)<- c("ID", "Age", "Household_Size", "State", "Currency",
                 "Total_Food_Exp", "Grocery_Exp", "Dining_Out_Exp","Misc_Food_Exp","Dining_Ou

head(food_data)
```

```
  ID Age Household_Size State Currency Total_Food_Exp Grocery_Exp
1  1  68              7    LA      USD         436.35      168.59
2  2  88              5    WA      USD                     452.10
3  3  82              3    MS      USD          279.1      301.66
4  4  73              8    AK      USD         -20.98      139.66
5  5  89              0    IN      USD         494.87          NA
6  6  18              6    WI      EUR         276.32      394.44
  Dining_Out_Exp Misc_Food_Exp Dining_Out_Count Alcohol_Included
1         140.71        109.77                4              Yes
2         192.94            NA                1          Unknown
3         239.84        103.94                9              Yes
4          69.19         44.84                2          Unknown
5         191.72        172.31                3              Yes
6         283.20        114.06                6          Unknown
  Food_Assistance
1            None
2            SNAP
3            None
4            None
5            None
6     Food Pantry
```

2c

Restrict the data to those paying in US dollars (USD). Show that it worked by confirming the number of observations before and after restricting the data.

```
#Count before restriction
n_before <- nrow(food_data)
cat("Initial number of observations is:",n_before,"\n")
```

```
Initial number of observations is: 262
```

```
#restriction
food_data<-food_data[food_data$Currency=="USD",]
n_after <- nrow(food_data)
cat("After restricting to USD, number of observations is:",n_after,"\n")
```

```
After restricting to USD, number of observations is: 230
```

2d We will only consider people older than 18 but less than 100. BUT WE HAVE NA values. How do we deal with them? I googled this and found this reference: https://uc-r.github.io/missing_values

We can use `is.na()`.

```
food_data <- food_data[food_data$Age >= 18 & food_data$Age <= 100 & !is.na(food_data$Age), ]
cat("After age cleaning, number of observations left is:", nrow(food_data), "\n")
```

```
After age cleaning, number of observations left is: 196
```

2e. The variable related to state.

Got to deal with missing states, invalid states and empty strings. Talking to gpt => R has a inbuilt state.abb. we can check `food_data$State` against it to make sure all state abbreviations are valid.

`%in%` inspiration: https://www.r-bloggers.com/2022/07/how-to-use-in-operator-in-r/?utm_source=chatgpt.com

```
valid_states <- state.abb
food_data <- food_data[ !is.na(food_data$State) &
                        food_data$State != "" &
                        food_data$State %in% valid_states, ]
cat("After state cleaning, number of observations left is:", nrow(food_data), "\n")
```

```
After state cleaning, number of observations left is: 176
```

2f. The four variables related to food expenditures.

Rules: Only consider positive values, discard missing values, discard any values above 3000/week for all 4 variables. Doesn't make sense to spend $3000 a week on any of them.

NOTE THAT *TOTAL_FOOD_EXP* is stored as a character! we got to make it numerical. Basically the R equivalent of casting.

```
food_data$Total_Food_Exp   <- as.numeric(food_data$Total_Food_Exp)
```

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

```
food_data$Grocery_Exp      <- as.numeric(food_data$Grocery_Exp)
food_data$Dining_Out_Exp   <- as.numeric(food_data$Dining_Out_Exp)
food_data$Misc_Food_Exp    <- as.numeric(food_data$Misc_Food_Exp)
```

```
#Cleaning time

# Total_Food_Exp
food_data <- food_data[ !is.na(food_data$Total_Food_Exp) &
                         food_data$Total_Food_Exp >= 0 &
                         food_data$Total_Food_Exp <= 3000, ]
cat("After Total_Food_Exp clean, n is:", nrow(food_data), "\n")
```

```
After Total_Food_Exp clean, n is: 157
```

```
# Grocery_Exp
food_data <- food_data[ !is.na(food_data$Grocery_Exp) &
                         food_data$Grocery_Exp >= 0 &
                         food_data$Grocery_Exp <= 3000, ]
cat("After Grocery_Exp clean, n is:", nrow(food_data), "\n")
```

```
After Grocery_Exp clean, n is: 154
```

```
# Dining_Out_Exp
food_data <- food_data[ !is.na(food_data$Dining_Out_Exp) &
                         food_data$Dining_Out_Exp >= 0 &
                         food_data$Dining_Out_Exp <= 3000, ]
cat("After Dining_Out_Exp clean, n is:", nrow(food_data), "\n")
```

```
After Dining_Out_Exp clean, n is: 151
```

```r
food_data <- food_data[ !is.na(food_data$Misc_Food_Exp) &
                         food_data$Misc_Food_Exp >= 0 &
                         food_data$Misc_Food_Exp <= 3000, ]
cat("After Misc_Food_Exp clean. n is:", nrow(food_data), "\n")
```

```
After Misc_Food_Exp clean. n is: 122
```

2g

Dining out cleaning

Value has to be greater than 0 and less than or equal to 21 i.e. 3 meals a day *7 = 21

```r
food_data <- food_data[ !is.na(food_data$Dining_Out_Count) &
                         food_data$Dining_Out_Count >= 0 &
                         food_data$Dining_Out_Count <= 21, ]

cat("After the dining out cleaning, n is:", nrow(food_data), "\n")
```

```
After the dining out cleaning, n is: 119
```

2h

The final number of observations is 119.

Problem 3: Collatz Conjecture

3a

Define the function `nextCollatz`

```r
#' Function to compute the next number in its Collatz sequence
#' Collatz sequence: n/2 if n is even or 3*n+1 if n is odd
#'
#' @param n A single positive integer
#' @return A single positive integer, the next Collatz value.
#'
```

```
nextCollatz <- function(n) {
  if (!is.numeric(n) || length(n) != 1 || n <= 0 || n != floor(n)) {
    stop("n is not a single positive integer")
  } #If n is even divide by 2 or do 3n+1
  if (n %% 2 == 0) {
    return(n / 2)
  } else {
    return(3 * n + 1)
  }
}

#Example
nextCollatz(5)
```

```
[1] 16
```

```
nextCollatz(16)
```

```
[1] 8
```

3b

Create a function `collatzSequence` that returns the Collatz sequence for a given input.

```
#' Function to return the Collatz sequence for a given input.
#' Applies nextCollatz() repeatedly to produce the full Collatz sequence from starting value
#' @param n A single positive integer
#' @return A list containing the vector of the entries in the Collatz sequence beginning at
#'
#' Example: collatzSequence(5)$sequence should return [5, 16, 8, 4, 2, 1]
#'          collatzSequence(5)$length should return 6

collatzSequence <- function(n) {
  # input validation (same as above)
  if (!is.numeric(n) || length(n) != 1 || n <= 0 || n != floor(n)) {
    stop("n is not a single positive integer")
  }
  seq_vals <- n
  # build the sequence from n to 1
  while (n != 1) {
    n <- nextCollatz(n)
    seq_vals <- c(seq_vals, n)
  }
  return(list(sequence = seq_vals, length = length(seq_vals)))
}

collatzSequence(5)
```

```
$sequence
[1]  5 16  8  4  2  1

$length
[1] 6
```

```
collatzSequence(19)
```

```
$sequence
 [1] 19 58 29 88 44 22 11 34 17 52 26 13 40 20 10  5 16  8  4  2  1

$length
[1] 21
```

3c

find the shortest and longest Collatz sequence starting with values between 100 and 500, inclusive. In the case of ties, report the lowest starting value.

```
start_vals <- 100:500
lengths <- sapply(start_vals, function (n) collatzSequence(n)$length)

#Find shortest sequence length and the smallest n that produces this
min_length <- min(lengths)
min_n <- (start_vals)[which(lengths==min_length)[1]]

#Find longest sequence length and smallest n that produces this
max_length <- max(lengths)
max_n <- (start_vals)[which(lengths==max_length)[1]]

cat("The shortest sequence length is:",min_length, "and starts at ",min_n,"\n")
```

```
The shortest sequence length is: 8 and starts at  128
```

```
cat("The longest sequence length is:", max_length, "and starts at", max_n, "\n")
```

```
The longest sequence length is: 144 and starts at 327
```