(e) Coding question: Please find the GradientDescentAlgorithms.R file attached. Steps taken to solve this question:

1. Removed the 'name' column from Auto dataset and ran linear Regression using the inbuilt lm function with mpg as the target variable. Output is as follows:

```
Call:
lm(formula = mpg ~ ., data = num_data)

Residuals:
    Min      1Q  Median      3Q     Max
-9.5903 -2.1565 -0.1169  1.8690 13.0604

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -17.218435   4.644294  -3.707  0.00024 ***
cylinders     -0.493376   0.323282  -1.526  0.12780
displacement   0.019896   0.007515   2.647  0.00844 **
horsepower    -0.016951   0.013787  -1.230  0.21963
weight        -0.006474   0.000652  -9.929  < 2e-16 ***
acceleration   0.080576   0.098845   0.815  0.41548
year           0.750773   0.050973  14.729  < 2e-16 ***
origin         1.426141   0.278136   5.127 4.67e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.328 on 384 degrees of freedom
Multiple R-squared:  0.8215,    Adjusted R-squared:  0.8182
F-statistic: 252.4 on 7 and 384 DF,  p-value: < 2.2e-16
```

2. Identified two most important features: weight, acceleration (most significant p-values)
3. For running the gradient descent algorithm, the data should be scaled. Created a function to scale the data using mean and standard deviation. Applied the function to all the columns in the data frame.
4. Ran the linear regression again with the scaled data and stored the value of coefficients in a vector (coef). Output:

```
Call:
lm(formula = mpg ~ ., data = scaled_auto)

Residuals:
     Min       1Q   Median       3Q      Max
-1.22873 -0.27630 -0.01498  0.23946  1.67334

Coefficients:
                        Estimate            Std. Error t value             Pr(>|t|)
(Intercept)   0.000000000000001046  0.021534039819788978   0.000              1.00000
cylinders    -0.107827322503740650  0.070653302697663053  -1.526              0.12780
displacement  0.266746678346135080  0.100756850629747971   2.647              0.00844 **
horsepower   -0.083596230321067999  0.067991407222685846  -1.230              0.21963
weight       -0.704556535669697892  0.070960987295151878  -9.929 < 0.0000000000000002 ***
acceleration  0.028481431902585324  0.034939082986361801   0.815              0.41548
year          0.354342869554390760  0.024057831271216041  14.729 < 0.0000000000000002 ***
origin        0.147185265760343131  0.028705120433303387   5.127          0.000000467 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4264 on 384 degrees of freedom
Multiple R-squared:  0.8215,    Adjusted R-squared:  0.8182
F-statistic: 252.4 on 7 and 384 DF,  p-value: < 0.00000000000000022
```

5. Divided x and y matrix. y contains only the mpg values, x is all the columns without mpg. Added a column containing only '1' in x (this will be a constant value used in matrix multiplication of theta_0)
6. Initialized theta vector with 8 values as 0 and alpha as 0.05.
7. Batch gradient descent algorithm is iterated 1500 times (epoch), once for each feature (j) and updates the theta vector using batch gradient descent formula. Saved the final theta values in 'theta'

8. Stochastic gradient descent algorithm is iterated 1500 times (epoch), once for each observation (i) and for each feature (j) and updates the theta vector using stochastic gradient descent formula. Saved the final theta values in 'theta_s'
9. Comparison of coef, theta and theta_s:

```
                              coef                      theta       theta_s
(Intercept)   0.00000000000001046426  0.00000000000000005756906 -0.0005117713
cylinders    -0.10782732250374064964 -0.09332602238795681404 15 -0.0952076883
displacement  0.26674667834613507988  0.23783705088573969810 28  0.2363532243
horsepower   -0.08359623032106799911 -0.08126802494728950498 89 -0.0814973431
weight       -0.70455653566969789203 -0.69445246855867848179 14 -0.6984949801
acceleration  0.02848143190258532395  0.02641678910638807498 13  0.0283697452
year          0.35434286955439076027  0.35369812734886996175 02  0.3536331567
origin        0.14718526576034313135  0.14513629176506007145 70  0.1456598476
```

coef: coefficients for lm
theta: coefficients of batch gradient descent algorithm
theta_s: coefficients of stochastic gradient descent algorithm

Interpretation:
Except the intercept in stochastic gradient descent, all the coefficients are very similar to each other. Weight's coefficient is approximately -0.7 and acceleration's coefficient is approximately 0.02 in all the algorithms.